TABLE DES MATIÈRES

Page

INTRODUCTION 1

CHAPITRE 1 Revue de littérature 5
1.1 Modélisation hydrologique 5
1.1.1 Types de modèles hydrologiques 6
1.1.2 Calage des modèles hydrologiques 7
1.2 Méthodes d'optimisation 11
1.2.1 Origine de la recherche opérationnelle 11
1.2.2 Types de méthodes d'optimisation 12
1.3 Choix des méthodes pour le calage automatique des modèles hydrologiques et problématique de recherche 13

CHAPITRE 2 Algorithmes d'optimisation sélectionnés.....15
2.1 Justification des choix des algorithmes d'optimisation sélectionnés.....15
2.2 Shuffled Complex Evolution of the University of the Arizona (SCEUA) 16
2.3 Dynamically Dimensioned Search (DDS) 19
2.4 Mesh Adaptive Direct Search (MADS) 21
2.5 Algorithme Complètement Aléatoire (ACA).....26

CHAPITRE 3 Objectifs de la recherche.....29
3.1 Objectif principal 29
3.2 Objectifs secondaires 30

CHAPITRE 4 Méthodologie 31
4.1 Bassins versants étudiés.....31
4.1.1 Description générale des bassins versants 31
4.1.2 Données hydrométéorologiques historiques 35
4.1.3 Justification du choix des bassins versants 35
4.2 Modélisation hydrologique 36
4.2.1 HYDROTEL 37
4.2.2 HSAMI..... 39
4.2.3 Justification du choix des modèles utilisés 40
4.3 Calibration des modèles hydrologiques 41
4.3.1 Définition des problèmes d'optimisation..... 41
4.3.1.1 Évaluation des méthodes de calage pour les modèles hydrologique coûteux en temps de calcul 41
4.3.1.2 Études des possibilités de configuration de l'optimisation dans NOMAD 43
4.3.1.3 Évaluation de l'impact du type de modèle hydrologique sur le comportement des algorithmes d'optimisation 44
4.3.2 Configuration des algorithmes..... 46

4.3.2.1 Shuffled Complex Evolution of the University of the Arizona (SCEUA) ... 46
4.3.2.2 Dynamically Dimensioned Search (DDS) ... 46
4.3.2.3 Nonlinear Optimization by Mesh Adaptive Direct Search (NOMAD) ... 48
4.3.2.4 Algorithme Complètement Aléatoire (ACA) ... 50
4.4 Évaluation de l'efficacité des algorithmes d'optimisation sélectionnés ... 51
4.4.1 Profils de performance ... 51
4.4.2 Analyse de dispersion des valeurs finales de la fonction objectif ... 54
4.4.3 Étude de la progression de la fonction objectif selon le temps de calcul ... 55

CHAPITRE 5 Résultats ... 57
5.1 Détermination de l'efficacité des algorithmes d'optimisation sélectionnés pour le calage des modèles hydrologiques coûteux en temps de calcul ... 57
5.1.1 Profils de performance ... 57
5.1.2 Analyse de dispersion des valeurs finales de la fonction objectif pour les différents algorithmes d'optimisation ... 63
5.1.3 Étude de la progression de la fonction objectif selon le temps de calcul ... 64
5.2 Études des possibilités de configuration de l'optimisation dans NOMAD ... 68
5.2.1 Profils de performance ... 68
5.2.2 Analyse de dispersion des valeurs finales de la fonction objectif ... 70
5.2.3 Étude de la progression de la fonction objectif selon le temps de calcul ... 71
5.3 Impact du type de modèle hydrologique sur le comportement des approches ... 73
5.3.1 Profils de performance ... 73
5.3.2 Analyse de dispersion des valeurs finales de la fonction objectif ... 76
5.3.3 Étude de la progression de la fonction objectif selon le temps de calcul ... 77

CHAPITRE 6 Discussion ... 81
6.1 Stratégies employées par les méthodes d'optimisation ... 81
6.2 Discussion sur les configurations de NOMAD ... 84
6.3 Discussion sur le type de modèle hydrologique ... 86
6.3.1 Complexité des problèmes de calage ... 86
6.3.2 Difficultés rencontrées par les méthodes d'optimisation : dimensionnalité, interdépendance et équifinalité ... 92
6.4 Efficacité en fonction du budget de calage ... 97
6.5 Vers le développement de nouvelles approches de calage ... 100
6.6 Choix méthodologiques de l'étude et autres pistes de recherche possibles ... 103

CONCLUSION ... 105

ANNEXE I PARAMÈTRES DES MODÈLES HYDROLOGIQUES ... 109

ANNEXE II BORNES DE CALAGE DES PROBLÈMES D'OPTIMISATION ... 113

ANNEXE III JEUX DE PARAMÈTRES INITIAUX DES PROBLÈMES D'OPTIMISATION ... 117

ANNEXE IV PSEUDO CODE DES ALGORITHMES D'OPTIMISATION SÉLECTIONNÉS 123

LISTE DE RÉFÉRENCES BIBLIOGRAPHIQUES.....129

Tableau 4.1 Regroupement des paramètres selon leur impact dominant sur le régime hydrique : (1) fonte de neige, (2) volume annuel, (3) pointes de crues et (4) débits d'étiages.....49

LISTE DES FIGURES

Page

Figure 2.1 Directions visitées lors de l'étape de la recherche sur le motif de points (F_k) à partir du meilleur point trouvé (p_k) pour un problème à deux paramètres où les points t_1, t_2 et t_3 sont tirés aléatoirement sur le motif de points (F_k) 23

Figure 2.2 Directions orthogonales 2N visitées lors de la génération du motif de points (F_k) à partir du meilleur point trouvé (p_k) pour un problème à deux paramètres où les points t_1 et t_3 sont orthogonales entre eux, de même que les points t_2 et t_4 24

Figure 2.3 Mise à jour du dimensionnement du maillaige (Δ_k) à l'itération de MADS k+1 dans l'éventualité où la recherche ou la sonde soit déclaré réussi (étape 3a) et dans l'éventualité où la recherche et la sonde soient déclarés insatisfaisants (étape 3b) 25

Figure 4.1 Emplacement du sous-bassin versant Cowansville sur le bassin versant de la rivière Yamaska.....32

Figure 4.2 Emplacement du sous-bassin versant Ceizur sur le bassin versant de la rivière Gatineau.....34

Figure 4.3 Schéma des sous-modèles employés pour la simulation à l'aide du modèle hydrologique HYDROTEL dans le cadre de cette étude).....38

Figure 4.4 Cadre expérimental pour l'évaluation des méthodes d'optimisation pour le calage de modèles hydrologiques coûteux en temps de calcaul 43

Figure 4.5 Cadre expérimentale pour l'étude des fonctionnalités d'optimisation de NOMAD 44

Figure 4.6 Cadre expérimentale pour l'écaluation de l'impact du type de modèle hydrologique sur le comportement des algorithmes d'optimisation.....45

Figure 4.7 Comportement moyen de DDS issu de 32 calages de la fonction test GRIEWANK (Griewank, 1981) à 10 dimensions où le nombre maximal d'évaluations de la fonction objectif (m) est égal 200, 500, 1000 et 2000; les mêmes 32 jeux de paramètres initiaux ont été repris pour chaque essai 47

Figure 4.8 Profils de performance de quatre algorithmes d'optimisation issus de 32 minimisations de la fonction test SCHEFEL (Schwefel, 1981) à 10 dimensions où le nombre d'évaluations est égal à 2000; les 32 jeux de

paramètres de départ ont été initialisés aléatoirement pour chaque problème d'optimisation 53

Figure 5.1 Profils de performance des algorithmes SCEUA, DDS, MADS et ACA pour le bassin versant Cowansville avec le modèle hydrologique HYDROTEL à 10 paramètres.....59

Figure 5.2 Profils de performance des algorithmes SCEUA, DDS, MADS et ACA pour le bassin versant Cowansville avec le modèle hydrologique HYDROTEL à 19 paramètres.....60

Figure 5.3 Profils de performance des algorithmes SCEUA, DDS, MADS et ACA pour le bassin versant Ceizur avec le modèle hydrologique HYDROTEL à 10 paramètres.....61

Figure 5.4 Profils de performance des algorithmes SCEUA, DDS, MADS et ACA pour le bassin versant Ceizur avec le modèle hydrologique HYDROTEL à 19 paramètres.....62

Figure 5.5 Dispersion du critère objectif du « Nash-Sutcliffe Efficiency » (NSE) à 2000 évaluations considérant les 32 calages effectués avec chacun des algorithmes SCEUA, DDS et MADS, sur les bassins versants Cowansville et Ceizur, pour les modèles hydrologiques HYDROTEL à 10 et 19 paramètres 64

Figure 5.6 Progression de la fonction objectif 1- « Nash-Sutcliffe Efficiency » (1-NSE) en fonction du nombre d'évaluations du modèle hydrologique pour les algorithmes SCEUA, DDS, MADS et ACA sur le bassin versant Cowansville pour les modèles hydrologiques HYDROTEL à 10 et 19 paramètres 66

Figure 5.7 Progression de la fonction objectif 1- « Nash-Sutcliffe Efficiency » (1-NSE) en fonction du nombre d'évaluations du modèle hydrologique pour les algorithmes SCEUA, DDS, MADS et ACA sur le bassin versant Ceizur pour les modèles hydrologiques HYDROTEL à 10 et 19 paramètres 67

Figure 5.8 Profils de performance des algorithmes MADS, MADS N+1 et MADS VG pour le bassin versant Cowansville avec le modèle hydrologique HYDROTEL à 10 paramètres.....69

Figure 5.9 Dispersion du critère objectif du « Nash-Sutcliffe Efficiency » (NSE) à 2000 évaluations considérant les 32 calages effectués avec les algorithmes MADS, MADS N+1 et MADS VG sur le bassin versant Cowansville pour le modèle hydrologique HYDROTEL à 10 paramètres 70

Figure 5.10 Progression de la fonction objectif 1- « Nash-Sutcliffe Efficiency » (1-NSE) en fonction du nombre d'évaluations du modèle hydrologique pour les algorithmes MADS, MADS N+1 et MADS VG sur le bassin versant Cowansville pour le modèle hydrologique HYDROTEL à 10 paramètres 72

Figure 5.11 Profils de performance des algorithmes SCEUA, DDS, MADS et ACA pour le bassin versant Cowansville avec le modèle hydrologique HSAMI à 23 paramètres 74

Figure 5.12 Profils de performance des algorithmes SCEUA, DDS, MADS et ACA pour le bassin versant Ceizur avec le modèle hydrologique HSAMI à 23 paramètres 75

Figure 5.13 Dispersion du critère objectif du « Nash-Sutcliffe Efficiency » (NSE) à 2000 évaluations considérant les 32 calages effectués avec les algorithmes SCEUA, DDS et MADS sur les bassins versants Cowansville et Ceizur pour le modèle hydrologique HSAMI à 23 paramètres 76

Figure 5.14 Progression de la fonction objectif 1- « Nash-Sutcliffe Efficiency » (1-NSE) en fonction du nombre d'évaluations du modèle hydrologique pour les algorithmes SCEUA, DDS, MADS et ACA sur le bassin versant Cowansville pour les modèles hydrologiques HYDROTEL à 10 et 19 paramètres, et HSAMI à 23 paramètres 78

Figure 5.15 Progression de la fonction objectif 1- « Nash-Sutcliffe Efficiency » (1-NSE) en fonction du nombre d'évaluations du modèle hydrologique pour les algorithmes SCEUA, DDS, MADS et ACA sur le bassin versant Ceizur pour les modèles hydrologiques HYDROTEL à 10 et 19 paramètres, et HSAMI à 23 paramètres 79

Figure 6.1 Teste de convergence de l'algorithme DDS. Critère objectif du « Nash-Sutcliffe Efficiency » (NSE) à 2000 évaluations avec DDS, critère du NSE à 4000 évaluations avec DDS et MADS (jeux de paramètres finaux de DDS recalés à l'aide de MADS) et gain sur le NSE suite au test de convergence sur les bassins versants Cowansville et Ceizur pour le modèle hydrologique HYDROTEL à 10 paramètres 82

Figure 6.2 Progression de la fonction objectif 1- « Nash-Sutcliffe Efficiency » (1-NSE) sur 10 000 du modèle HSAMI à 23 paramètres, pour chacun des algorithmes SCEUA, DDS-2000, DDS-10000, MADS et ACA sur le bassin versant Cowansville 88

Figure 6.3 Progression de la fonction objectif 1- « Nash-Sutcliffe Efficiency » (1-NSE) sur 10 000 du modèle HSAMI à 23 paramètres, pour chacun

Rapport-Gratuit.com

des algorithmes SCEUA, DDS-2000, DDS-10000, MADS et ACA sur le bassin versant Ceizur 89

Figure 6.4 Dispersion du critère objectif du « Nash-Sutcliffe Efficiency » (NSE) à 2000 évaluations et à 10 000 évaluations considérant les 32 calages effectués avec les algorithmes SCEUA, DDS et MADS sur les bassins versants Cowansville et Ceizur pour le modèle hydrologique HSAMI à 23 paramètres 91

Figure 6.5 Étude de l'équifinalité des jeux de paramètres finaux issus des 96 calages à l'aide de DDS (courbes bleues), SCEUA (courbes pourpres) et MADS (courbes orangées) sur le bassin versant Cowansville avec les modèles hydrologiques HYDROTEL à 10 et 19 paramètres, et HSAMI à 23 paramètres 94

Figure 6.6 Étude de l'équifinalité des jeux de paramètres finaux issus des 96 calages à l'aide de DDS (courbes bleues), SCEUA (courbes pourpres) et MADS (courbes orangées) sur le bassin versant Ceizur avec les modèles hydrologiques HYDROTEL à 10 et 19 paramètres, et HSAMI à 23 paramètres 95

Figure 6.7 Lien entre l'efficacité des méthodes d'optimisation et la disponibilité du budget de calage.....99

LISTE DES ABRÉVIATIONS, SIGLES ET ACRONYMES

ACA	Algorithme Complétement Aléatoire
CEHQ	Centre d'Expertise Hydrique du Québec
CCE	Competitive Complex Evolution algorithm
DDS	Dynamically Dimensioned Search
GERAD	Groupe d'Études et de Recherche en Analyse des Décisions
INRS	Institut National de la Recherche Scientifique
MADS	Mesh Adaptive Direct Search
MADS VG	Mesh Adaptive Direct Search employant les groupes de variables
MADS N+1	Mesh Adaptive Direct Search employant les directions orthogonales N+1
NOMAD	Nonlinear Optimization by Mesh Adaptive Direct Search
NSE	Critère du Nash-Sutcliffe (Nash-Sutcliffe Efficiency)
RMSE	Root Mean Square Error
SCEUA	Shuffled Complex Evolution of the University of the Arizona
SIG	Systèmes d'Information Géographique
UHRH	Unité Hydrologique Relativement Homogène

Rapport-gratuit.com
LE NUMERO 1 MONDIAL DU MÉMOIRES

INTRODUCTION

Le territoire québécois s'appuie sur la disponibilité abondante et l'accessibilité à la ressource en eau pour produire de l'hydroélectricité de manière à satisfaire la demande énergétique de toute la province du Québec, tout en contribuant à satisfaire celle du Nord-Est du continent nord-américain. Cette source d'énergie, très souvent qualifiée de renouvelable et propre, est devenue l'un des champs d'expertise les plus reconnus et prospères du sol québécois (Hydro-Québec, 2014). La soixantaine de centrales hydroélectriques implantées sur la zone géographique du Bouclier canadien québécois alimente cette production d'énergie via l'emmagasinement de la ressource dans d'immenses réservoirs naturels ou artificiels dans le but d'exploiter chaque litre d'eau précipité sur les bassins versants qui desservent finalement les différentes centrales. La gestion efficace des niveaux des réservoirs des centrales est primordiale pour les producteurs d'hydroélectricité puisqu'elle minimise les coûts liés à la production favorisant ainsi une meilleure rentabilité dans la vente de son énergie. Cette gestion passe par une prévision hydrologique des apports naturels aux réservoirs (débits en rivière), issue des simulations produites par des outils indispensables, les modèles hydrologiques.

De manière à rendre cet outil utilisable et performant, tout modèle hydrologique doit être adapté au bassin versant modélisé en ajustant les valeurs des paramètres internes, lesquelles sont soit difficiles, soit impossibles à mesurer. L'ajustement de ces paramètres, appelé calage, est crucial pour l'utilisateur d'un modèle hydrologique puisqu'il détermine le niveau de confiance envers l'outil de simulation. L'ajustement manuel des paramètres peut être complexe et difficile à effectuer (essais et erreurs) et le recours à des algorithmes d'optimisation est très souvent la méthode envisagée. Ces algorithmes sont essentiellement des procédures de raisonnements mathématiques, lesquels permettent de trouver une combinaison de paramètres qui rend les débits simulés par le modèle hydrologique le plus similaires possible aux débits observés en rivière.

Les modèles hydrologiques ne servent pas qu'à la prédiction d'apports aux réservoirs des centrales hydroélectrique, mais également à la prévision hydrologique des inondations et des étiages, à l'étude de l'impact des changements climatiques et à la gestion des quelques 5 500 barrages du Québec qui assurent la sécurité civile (Centre d'expertise hydrique du Québec, 2003). Les simulations obtenues à l'aide de ces modèles mèneront notamment à la prise de décision quant à la gestion des ressources en eau et à la construction en milieu hydrique. Cette utilisation diversifiée des modèles hydrologiques a mené à la confection de plusieurs types de modèles qui permettent de mieux répondre aux besoins des utilisateurs. Parmi ces différents types, certains s'appuient sur la résolution d'équations à base plus physique et emploient des bases de données issues de la télédétection à fine résolution pour spatialiser la modélisation à l'échelle du bassin versant. Cependant, les temps de calcul des simulations à l'aide de ces modèles sont généralement importants et ces temps de calcul se répercutent inévitablement sur le calage des paramètres internes à l'aide d'algorithmes d'optimisation. L'efficacité des méthodes d'optimisation employées pour le calage peut alors jouer un rôle décisif quant à l'utilisation de ces modèles.

Les principaux objectifs de ce mémoire sont : (1) d'analyser l'efficacité d'algorithmes d'optimisation sélectionnés, soient « Mesh Adaptive Direct Search » (MADS) et « Dynamically Dimensionned Search » (DDS), lorsqu'employés pour le calage de modèles hydrologiques coûteux en temps de calcul; (2) d'évaluer diverses possibilités de configuration de l'algorithme MADS, dans le but d'identifier les configurations propices au calage de modèles hydrologiques; et, finalement, (3) d'étudier l'impact du type de modèle hydrologique sur le comportement des algorithmes sélectionnés. De manière à atteindre ces objectifs, il est nécessaire de choisir un cadre expérimental représentatif des conditions de calage normales auxquelles un hydrologue peut être confronté. Ce faisant, les combinaisons de quelques bassins versants (2) et modèles hydrologiques (3) ont fait l'objet de bancs d'essai pour l'exécution des calages à l'aide de plusieurs approches d'optimisation (6) jugées potentiellement efficaces. Les choix effectués quant à la conception des problèmes d'optimisation constituent des hypothèses plausibles formant le cadre expérimental et celles-ci ont été posées de manière à ne pas favoriser ou désavantager l'une ou l'autre des méthodes

d'optimisation étudiées. Les résultats obtenus sont considérés et analysés du point de vue de l'optimisation, essentiellement. Ces travaux analysent le comportement des méthodes d'optimisation et visent à permettre d'identifier les stratégies employées qui conviennent ou non au contexte de la problématique.

Le texte de ce mémoire est organisé comme suit. Un premier chapitre expose une revue de la littérature scientifique en lien avec la problématique examinée dans cette étude. Une section fort importante pour l'atteinte des objectifs présentés ci-haut décrit ensuite les procédures et stratégies d'optimisation employées par les diverses méthodes d'optimisation sélectionnées. Les objectifs et choix méthodologiques sont présentés respectivement aux chapitres 3 et 4. S'ensuit le chapitre 5 avec la présentation des résultats obtenus pour chacun des trois objectifs de recherche et, finalement, la discussion et les interprétations des résultats sont présentées au chapitre 6 de ce mémoire.

CHAPITRE 1

Revue de littérature

Ce chapitre passe en revue la littérature pertinente liée à la recherche présentée dans ce mémoire, en abordant trois thèmes principaux : la modélisation hydrologique, les méthodes d'optimisation et le choix des méthodes pour le calage des modèles hydrologiques.

1.1 Modélisation hydrologique

Le terme « modélisation hydrologique » prend naissance de la conception d'une suite d'équations mathématiques qui représente les divers processus hydrologiques à l'échelle d'un bassin versant, dans le but de reproduire la relation qui existe entre les précipitations et les débits en rivière à l'échelle de ce territoire (Singh et Woolhiser, 2002). Par ces processus hydrologiques, tels que l'évapotranspiration, l'infiltration, la fonte de la neige, le ruissellement de surface, etc., le modèle simule la séparation, le transport, la rétention et la transformation de l'eau de manière à reproduire la réponse hydrologique sur un bassin versant étudié. Également, la modélisation hydrologique requiert des séries de données météorologiques (précipitations, températures, etc.) et des caractéristiques intrinsèques au bassin versant qui agissent comme compléments essentiels à la constitution d'un modèle hydrologique (Singh et Woolhiser, 2002; Moradkhani et Sorooshian, 2009). Cet outil numérique que représente le modèle hydrologique a une utilité importante pour le domaine de l'ingénierie civile puisque, d'une part, il devient un outil de design d'infrastructures telles que les ponts, les digues, les routes et les systèmes de drainage. D'autre part, les modèles hydrologiques permettent l'analyse temporelle des comportements hydriques (prévision hydrologique, études d'impacts des changements climatiques) qui mènent à la prise de décision et à la planification de la gestion de la ressource (Duan et al., 1994; Singh et Woolhiser, 2002; Beven, 2006).

1.1.1 Types de modèles hydrologiques

Même si l'histoire des modèles hydrologiques ne débute que vers les années 1850 avec l'arrivée de la méthode rationnelle développée par Mulvany (1850), les recherches sur le domaine de l'hydrologie ont mené au développement de plusieurs types de modèles hydrologiques qui se différencient, notamment, par leur manière de représenter les processus hydrologiques et leur degré de spatialisation. En ce qui a trait à la représentation des processus hydrologiques, il existe essentiellement trois catégories de modèles : (1) empiriques, (2) conceptuels et (3) à base physique (Clarke, 1973; Singh et Woolhiser, 2002; Moradkhani et Sorroshian, 2009; Pechlivanidis et al., 2011). Un modèle est qualifié d'empirique lorsque celui-ci établit un lien direct entre les précipitations et les débits en rivière, en faisant abstraction de l'existence des processus hydrologiques, ce qui peut cependant apporter un grave problème d'adaptabilité à des bassins versants différents (Clarke, 1973). Le recours à des modèles hydrologiques un peu plus élaborés et adaptatifs est envisageable par la conceptualisation des processus hydrologiques; c'est-à-dire que chacun des principaux processus hydrologiques est symbolisé par des éléments conceptuels, tels que des réservoirs à niveaux variables, et des paramètres sans signification physique, non mesurables dans la réalité, mais tout de même liés à des propriétés intrinsèques du modèle hydrologique (Clarke, 1973; Singh et Woolhiser, 2002; Moradkhani et Sorooshian, 2009; Pechlivanidis et al., 2011). Les modèles conceptuels démontrent une certaine forme de compréhension du cycle hydrologique. Pourtant, ceux-ci ne font pas appel aux lois de la physique connues comme c'est le cas des modèles hydrologiques dits à base physique, dont les paramètres sont plus généralement mesurables indépendamment de la réponse hydrologique du bassin versant, et/ou ont un sens physique (Clarke, 1973; Singh et Woolhiser, 2002; Moradkhani et Sorooshian, 2009; Pechlivanidis et al., 2011).

Par ailleurs, la simulation des processus hydrologiques peut être effectuée suivant différents degrés de spatialisation (ou découpage spatial du bassin versant en sous unités de modélisation) au sein des outils de modélisation. La spatialisation peut être inexistante dans certains modèles, c'est-à-dire que le bassin versant est considéré comme une seule entité dite

globale (Singh et Woolhiser, 2002; Das et al., 2008; Moradkhani et Sorooshian, 2009; Pechlivanidis et al., 2011). Il est également possible de sous-diviser le bassin versant en sous-bassins liés entre eux, de façon à mieux simuler les procédés de séparation, de transport, de rétention et de transformation de l'eau au sein de la modélisation. Cette échelle, semi-distribuée, fait souvent appel à notre niveau de connaissances quant à la division du territoire en sous-bassins versant (Singh et Woolhiser, 2002; Das et al., 2008; Pechlivanidis et al., 2011). Enfin, les SIG supportent la possibilité de construire des modèles hydrologiques complètement distribués qui séparent le bassin versant en mailles ou sous-bassins versant élémentaires de dimension réduite, diminuant considérablement l'échelle de simulation des processus hydrologiques. La configuration de ce type de modèle permet une meilleure représentation de la variabilité spatiale des caractéristiques physiographiques intrinsèque au bassin versant (altitude, types de sol, occupation de territoire, etc.). Cependant, cette démarche est plus exigeante puisque les données physiographiques nécessitent une collecte et un prétraitement avant d'être introduites au modèle (Singh et Woolhiser, 2002; Das et al., 2008; Moradkhani et Sorooshian, 2009; Pechlivanidis et al., 2011). Des associations systématiques et conséquentes entre la manière de représenter les processus hydrologiques et le degré de spatialisation des modèles hydrologiques ont été identifiées par Singh et Woolhiser (2002). Notamment, les modèles hydrologiques conceptuels et globaux ainsi que les modèles à base physique et distribués sont des types de modèles largement répandus et employés en recherche et en milieu opérationnel. Ces associations ne sont toutefois pas les seules possibles puisque le choix du type de modèle dépend d'abord des moyens dont l'utilisateur dispose (connaissance du bassin versant et données disponibles), et de l'expertise de celui-ci.

1.1.2 Calage des modèles hydrologiques

De manière à rendre un modèle hydrologique utilisable et performant, il est nécessaire d'adapter celui-ci au bassin versant étudié en ajustant les valeurs des paramètres internes intégrés au sein des équations mathématiques qui régissent la représentation des processus hydrologiques. Qu'ils soient de nature empirique ou physique, ces paramètres sont souvent

Rapport-Gratuit.com

difficiles (voire impossibles) à mesurer, mais sont essentiels à la performance des simulations hydrologiques. Ils nécessitent alors un ajustement de façon à ce qu'un modèle hydrologique, appliqué à un bassin versant étudié, reproduise le plus fidèlement possible des conditions historiques observées (généralement des débits en rivière). Ce processus d'ajustement, communément appelé calage ou calibration, constitue un problème d'optimisation visant à identifier un jeu de paramètres qui minimise l'écart entre la réponse hydrologique simulée par le modèle et les observations sur le bassin versant. Cette comparaison entre les données simulées et observées se fait par l'intermédiaire d'un repère objectif qui sert d'indicateur de la performance du modèle. Ce repère cible est couramment appelé fonction objectif, mais les noms de critère objectif ou métrique de calage sont également employés (Duan et al., 1994; Moriasi et al., 2006; Pushpalatha et al., 2012). Il existe une multitude de fonctions objectifs permettant ainsi de poser un regard de performance sur nos modèles hydrologiques bien que des critères comme le « *Nash-Sutcliffe Efficiency* » (NSE) ou le « *Root Mean Square Error* » (RMSE) soient employés depuis longtemps. Ces fonctions objectifs sont devenues incontournables pour les hydrologues qui, au fil du temps, ont étudié, accepté et recommandé celles-ci dans la littérature. Cependant, les possibilités d'évaluation de la performance ne se limitent pas à ce qui est normalement utilisé par l'ensemble de la communauté. D'autres métriques de calculs statistiques sont envisageables et peuvent s'avérer plus adéquates selon les objectifs de modélisation poursuivis (Moriasi et al., 2006). La calibration des modèles hydrologiques est cruciale puisqu'elle détermine le niveau de confiance d'un utilisateur envers ces outils qui, comme mentionné précédemment, serviront à mener des études sur lesquelles reposent des décisions quant à la gestion des ressources en eau et la construction d'infrastructures en milieu hydrique.

Deux principales avenues peuvent être empruntées pour effectuer le calage des modèles hydrologiques : le calage manuel et le calage automatique. Dans le premier cas, l'ajustement des paramètres est effectué par essai-erreur et requiert l'exécution manuelle de plusieurs simulations consécutives où l'utilisateur devra inspecter et juger chacune des simulations du modèle. L'ampleur du travail requis dépend notamment du nombre de paramètres à ajuster et du niveau d'expertise de l'usager. La tâche est laborieuse, complexe à enseigner aux

prochains utilisateurs et difficilement reproductible (Boyle et al., 2000; Moradkhani et Sorooshian, 2009; Pechlivanidis et al., 2011). Le second cas implique l'utilisation d'algorithmes d'optimisation qui, de façon objective et reproductible, explorent l'espace des solutions (jeux de paramètres) afin de minimiser la valeur de la fonction objectif. Pareillement à une calibration manuelle, l'optimiseur automatique génère des jeux de paramètres qu'il incorpore ensuite à l'exécution du modèle hydrologique pour évaluer le critère objectif qui devient le seul critère d'appréciation quant à la performance du jeu de paramètres inséré. Ces étapes sont répétées à maintes reprises tant que le critère d'arrêt de l'optimiseur n'est pas atteint. Les simulations et calculs sont entièrement automatisés via des codes informatiques qui, selon leur suite d'instructions, arrivent à accomplir une calibration sans intervention de l'utilisateur. Contrairement au calage manuel, les algorithmes d'optimisation n'exploitent aucune connaissance hydrologique pour accomplir le processus de calage. Ils suivent leur propre ligne de conduite pour orienter les modifications des paramètres vers un modèle hydrologique performant. L'emploi d'algorithmes d'optimisation est relativement simple une fois configurés, mais nécessite une compréhension adéquate de la problématique de calage (Singh et Woolhiser, 2002; Moradkhani et Sorooshian, 2009; Pechlivanidis et al., 2011).

Comme le rapportent Moradkhani et Sorooshian (2009), le problème d'optimisation que pose la calibration des modèles hydrologiques est un problème dit inverse; c'est-à-dire qu'à partir d'un modèle composé de plusieurs concepts, paramètres plus ou moins physiques et alimenté de données météorologiques (et physiographiques si requis), l'on tente d'imiter au meilleur de nos capacités une situation réelle et observable. L'on cherche donc la combinaison des meilleurs éléments qui provoquent le résultat d'un problème (l'observation) contrairement à une résolution de problème classique, où l'on cherche la solution du problème à partir d'une combinaison d'informations spécifiques. Il est alors difficile de définir si les concepts, paramètres et données positionnés à l'intérieur de notre modèle hydrologique sont la bonne représentation de la réalité. La modélisation hydrologique est issue de la science fondamentale qui vise à décrire le sens physique réel d'une existence. Pourtant, il y a une perception contradictoire, car de multiples représentations de la réalité sont définies par

Rapport-gratuit.com
LE NUMERO 1 MONDIAL DU MÉMOIRES

plusieurs combinaisons différentes quasi équivalentes (Beven, 2006). Autrement dit, la réalité observable et unique (souvent une série temporelle de débits en rivière) peut être modélisée à l'aide de divers types de modèles hydrologiques différents, diverses séries de données météorologiques et divers jeux de paramètres totalement distincts qui fourniront des réponses hydrologiques très similaires; c'est le concept de l'équifinalité. De façon plus générale, l'on peut décrire l'équifinalité comme étant un même résultat qui peut être obtenu par des voies et des conditions initiales différentes.

Le concept de l'équifinalité a été largement examiné par Beven (2001; 2006). Selon cet auteur, l'évidence de l'équifinalité au sein des modèles hydrologiques apporte un dilemme où l'utilisateur peut soit : (1) tenter de l'éviter et poursuivre la recherche du jeu de paramètres qui offrira la valeur optimale de la fonction objectif, ou (2) l'accepter et décider de l'exploiter à son avantage. L'acceptation de l'équifinalité commence par l'acceptation de l'erreur au sein des modèles hydrologiques. Aucun modèle, conceptuel comme physique, global comme distribué, n'est une parfaite représentation de la réalité. Plusieurs sources d'incertitude interviennent dans l'exercice de modélisation : erreurs sur les données de forçage météorologique et sur les données observées de débit (erreurs de mesure, notamment), imperfection de la structure interne du modèle (biais au niveau de la représentation des processus hydrologiques) et choix des jeux de paramètres (calibration) (Moradkhani et Sorooshian, 2009; Pechlivanidis et al., 2011). Cette incertitude, globalement, est difficilement évaluable et continue de poser un défi d'envergure. Toutefois, il est possible de tirer parti de l'équifinalité, entre autres, par le développement de méthodes d'optimisation qui utilisent la disponibilité de plusieurs jeux de paramètres de qualité pour améliorer le calage final. Ou encore, l'équifinalité peut démontrer aux modélisateurs que la surparamétrisation de leur modèle et la conceptualisation de ceux-ci exigent des améliorations à la structure interne des modèles. L'apprivoisement de l'équifinalité à notre avantage est donc l'une des avenues qui mènera à améliorer la recherche sur la modélisation hydrologique et non pas, un obstacle que nous devons contourner. (Beven, 2006)

Malgré tous les défis que pose la calibration de modèles hydrologiques, les méthodes de résolution de problèmes (algorithmes d'optimisation) développées au cours des dernières décennies réussissent à produire des résultats qui satisfont les utilisateurs de modèles hydrologiques. Évidemment, les algorithmes d'optimisation que nous employons aujourd'hui pour résoudre nos problèmes d'optimisation sont issus d'un historique de recherche auquel il est intéressant de s'attarder brièvement.

1.2 Méthodes d'optimisation

1.2.1 Origine de la recherche opérationnelle

Lors de la Seconde Guerre mondiale, les britanniques et les américains ont réuni un large éventail de scientifiques pour répondre à un besoin urgent d'efficacité en approvisionnement des ressources pour les diverses opérations et activités militaires (Hillier et Lieberman, 2010). Les objectifs visés étaient la rapidité d'exécution et la diminution des coûts; deux missions simples en apparence, mais complexes à mener à bien étant donné la quantité d'approvisionnements nécessaires. Le succès de cette division spéciale a mené cette première escouade scientifique en recherche opérationnelle à étendre ses mandats à d'autres champs d'expertise. Au début des années 1950, l'introduction de la recherche opérationnelle se fait progressivement dans les affaires, l'industrie et même le gouvernement. En moins d'un siècle, la demande de méthodes rigoureuses d'optimisation est répandue à travers des disciplines telles que la construction, le transport, la manufacture, les télécommunications, la santé, les services publiques, etc. (Hillier et Lieberman, 2010). La recherche de l'optimalité dans les problèmes devient une caractéristique fortement importante que les méthodes d'optimisation doivent garantir. Cependant, la modernisation et l'évolution technologique constante amènent une complexité au sein des problèmes à résoudre qui ne cesse de croître. Bien que les efforts scientifiques soient soutenus, les méthodes d'optimisation ne peuvent devancer cette progression de complexité empêchant ainsi le développement de méthodes qui garantissent continuellement l'optimalité (Hillier et Lieberman, 2010).

1.2.2 Types de méthodes d'optimisation

Dans le cadre de cette étude et pour des fins de simplifications, la classification des méthodes d'optimisation peut essentiellement être divisée selon trois classes : les méthodes exactes, les méthodes inexactes et les méthodes semi-exactes. La première classe regroupe les approches d'optimisation qui vont garantir à tout coup une solution optimale sur l'ensemble du domaine de solutions (optimum global). Cette classe apporte toutefois d'importantes limitations au niveau de la complexité des problèmes pouvant être résolus puisqu'elles reposent sur des hypothèses très contraignantes comme par exemple : fonction objectif convexe, existence des dérivées, etc. Inversement, les méthodes inexactes ne peuvent garantir l'optimalité de la solution trouvée, mais font preuve d'astuces pour identifier rapidement une solution de bonne qualité et ce, même lorsque la taille du problème est importante. Finalement, les méthodes semi-exactes joignent les deux opposés en donnant une solution qui satisfait les conditions d'optimalité non pas globale, mais locale; en d'autres termes, on aura seulement la preuve qu'aucune solution voisine n'est meilleure que celle trouvée.

Parmi les méthodes d'optimisation inexactes, on retrouve les méthodes heuristiques qui consistent en une procédure de recherche spécifique au problème à résoudre et qui devrait découvrir une excellente solution en un nombre raisonnable d'exécutions du problème (budget de résolution) (Hillier et Lieberman, 2010; Rios et Sahinidis, 2013). Toutefois, ces méthodes sont difficilement transposables pour d'autres problèmes que ceux pour lesquels elles ont été pensées. Depuis maintenant quelques années, le développement et la recherche ont amené le domaine de la recherche opérationnelle à concevoir des méthodes dites métaheuristiques. Il existe une panoplie de métaheuristiques, basées sur des principes de recherche simples et intuitifs directement inspirés de comportements que l'on observe dans la nature. Plus ces méthodes disposent d'une quantité d'exécutions élevée, plus la solution finale devrait s'approcher d'un point optimal. Ces méthodes reposent sur une structure générale dont les lignes directrices et stratégies de recherche permettent d'adapter les métaheuristiques à certains types de problèmes d'optimisation (Hillier et Lieberman, 2010). La recherche, dans l'espace des solutions, est donc plus robuste, ce qui devrait empêcher les

approches de fournir des solutions de mauvaise qualité en exploitant des procédés d'amélioration locale et d'exploration plus efficaces. Les métaheuristiques sont donc devenues une ressource importante de techniques d'optimisation pour les praticiens en recherche opérationnelle (Hillier et Lieberman, 2010).

1.3 Choix des méthodes pour le calage automatique des modèles hydrologiques et problématique de recherche

Le calage des modèles hydrologiques est un problème d'optimisation spécifique ayant des caractéristiques propres pour lequel certaines méthodes d'optimisation sont appropriées. Premièrement, il s'agit d'un problème d'optimisation sans dérivée; c'est-à-dire qu'il est impossible pour l'algorithme d'optimisation d'exploiter la structure de la fonction objectif en dérivant l'équation de celle-ci pour trouver un jeu de paramètres optimal. Cette particularité du calage des outils hydrologiques rend donc certaines méthodes exactes complètement inefficaces puisque leur stratégie de recherche basée sur l'exploitation de la surface de réponse devient caduque. Le processus d'optimisation sans dérivée, également appelé optimisation de « boîte noire », utilise pour seule information la valeur du critère objectif obtenue pour un jeu de paramètres par simulation du modèle hydrologique (Audet et al., 2010; Shan et Wang, 2010; Audet et al., 2012; Rios et Sahinidis, 2013). Deuxièmement, ce problème d'optimisation que représente le calage des modèles hydrologique est caractérisé par son nombre important de paramètres à optimiser (pouvant aller de 4 à plus de 30 paramètres) et par la grande variabilité de ceux-ci (Duan et al., 1994; Madsen, 2003; Shan et Wang, 2010). La variabilité de chaque paramètre est limitée par des valeurs maximales et minimales qui constituent les seules contraintes du problème d'optimisation. Ceci donne lieu à des espaces paramétriques à explorer qui sont très vastes et qui sont impossibles à examiner dans leur totalité. Troisièmement, l'équifinalité que l'on retrouve comme caractéristique dans la surface de réponse des modèles hydrologiques apporte une complexité supplémentaire puisque plusieurs jeux de paramètres offrent des performances similaires (Freer et Beven, 2001; Beven, 2006), ce qui, du point de vue du domaine de l'optimisation, s'apparente au concept de dégénérescence de l'optimisation. Les caractéristiques mentionnées jusqu'alors résignent déjà l'utilisateur à adopter des méthodes d'optimisation qui font partie des classes

heuristiques et métaheuristiques, manifestement mieux adaptées à la problématique du calage des modèles hydrologiques.

L'utilité de plusieurs algorithmes d'optimisation récemment développés a déjà été démontrée pour le calage de modèles conceptuels et agrégés spatialement (Duan et al., 1994; Arsenault et al., 2013; Minville et al., 2014). Cependant, avec la croissance de la puissance des ressources informatiques, on observe le recours de plus en plus fréquent à des modèles hydrologiques à base physique et/ou distribués spatialement. Ceux-ci offrent certains avantages comme la possibilité d'améliorer les prévisions hydrologiques à une échelle beaucoup plus fine que les modèles globaux ou encore l'opportunité d'étudier certains processus hydrologiques non perceptibles à une échelle globale (par exemple, retrait des couverts glaciers) (Singh et Woolhiser, 2002; Carpenter et Georgakakos, 2006; Pechlivanidis et al., 2011). Lorsqu'il s'agit du calage de modèles plus complexes, à base physique et distribués, les coûts associés au temps de calcul augmentent de manière importante. Comme l'exécution d'une seule simulation à l'aide de ce type de modèle peut prendre plusieurs minutes, voire quelques heures, et que l'optimisation peut requérir des milliers de simulations, l'approche retenue se doit de tendre à minimiser le nombre d'exécutions du modèle hydrologique (Mungunthan et al., 2005; Mungunthan et Shoemaker, 2006; Zhang et al., 2009; Razavi et al., 2010). Un algorithme nécessitant le moins de simulations possibles est donc source d'une grande efficacité en matière d'économie de temps de calcul. De plus, cette efficacité est également définie par la capacité de la méthode de calage à trouver un jeu de paramètres de bonne qualité. Autrement dit, l'heuristique ou la métaheuristique optimale pour le calage des modèles hydrologiques distribués et à base physique sera celle répondant au meilleur compromis entre l'efficacité en temps de calcul (le moins de simulations possible) et la valeur de la fonction objectif (qualité du jeu de paramètres final). Ce compromis recherché sera donc l'objectif principal de cette étude qui évaluera les méthodes d'optimisation pertinentes en termes d'efficacité pour le calage des modèles hydrologiques coûteux en temps de calcul.

CHAPITRE 2

Algorithmes d'optimisation sélectionnés

Ce chapitre justifie la sélection des algorithmes utilisés pour cette étude et décrit leur fonctionnement afin d'en illustrer les stratégies de recherche. La configuration des algorithmes (paramètres internes) est plutôt discutée au chapitre 4 exposant la méthodologie de l'étude.

2.1 Justification des choix des algorithmes d'optimisation sélectionnés

Les dernières décennies ont fourni aux hydrologues une quantité importante d'heuristiques et de métaheuristiques pouvant répondre à la problématique du calage des modèles hydrologiques. Puisque la totalité de ces algorithmes ne peut être sélectionnée pour cette étude en raison des limites imposées par le temps de calcul importants des calages, certains ont été ciblés. Quatre approches de calage de natures différentes, (1) évolutive, (2) à recuit simulé, (3) à descente direct et (4) aléatoire, ont été sélectionnées de manière à couvrir une étendue plus vaste des stratégies d'optimisation applicables aux problèmes de calage des modèles hydrologiques.

La première métaheuristique sélectionnée est le très répandu « Shuffled Complex Evolution method developed at the University of Arizona » (SCEUA). Depuis son apparition dans la littérature (Duan et al., 1993), cet algorithme est devenu une méthode d'optimisation très largement utilisée pour le calage des modèles hydrologiques. Par sa forte popularité tant dans le domaine de la recherche que dans le milieu opérationnel, l'algorithme SCEUA apporte un standard au niveau de la performance des méthodes d'optimisation auquel il est convenable de comparer d'autres approches de calage. Bien que sa renommée soit faite pour le calage de modèles hydrologiques conceptuels et globaux, SCEUA offre une stratégie d'optimisation efficace qui pourrait également répondre à la problématique de calage de modèles hydrologiques plus complexes.

Le second algorithme choisi est une méthode d'optimisation inexacte (métaheuristique) prometteuse récemment développée : « Dynamically Dimensionned Search » (DDS) (Tolson et Shoemaker, 2007). Il a été observé que, de par sa structure interne, DDS performe efficacement en ciblant très rapidement des jeux de paramètres ayant une valeur de fonction objectif située très près des optima de l'espace paramétrique. Cette méthode a démontré des gains en temps de calcul pour le calage du modèle SWAT (distribué et à base physique; coûteux en temps de calcul) variant de 15 à 59 % du temps requis comparativement à des algorithmes d'optimisation typiquement utilisés (Tolson et Shoemaker, 2007).
Contrairement aux deux algorithmes précédents, le troisième est une méthode d'optimisation classée parmi les méthodes semi-exactes; c'est-à-dire que celle-ci garantira les conditions d'optimalité locale du jeu de paramètres issu du processus de calage si ce dernier est mené à terme (Abramson et Audet, 2006). Cet algorithme appelé « Mesh Adaptive Direct Search » (MADS) (Audet et Dennis, 2006), se retrouve au sein d'un logiciel d'optimisation qui offre diverses fonctionnalités permettant à l'utilisateur d'adapter la méthode d'optimisation au contexte du problème de calage. Ce logiciel, appelé NOMAD pour « Nonlinear Optimization by Mesh Adaptive Direct Search », aide à solutionner un large éventail de problèmes d'optimisation étant donné ces diverses fonctionnalités (Le Digabel, 2011; Audet, 2014).

Finalement, cette étude emploiera un algorithme complètement aléatoire (ACA) qui déterminera le seuil minimal de performance des algorithmes de manière à apporter un point de comparaison standard (exempt de toute stratégie de recherche intelligente) pour les trois approches d'optimisation sélectionnés au sein de ce mémoire. Les sections qui suivent présentent la description générale des différentes étapes de calcul associées au fonctionnement de SCEUA, DDS, MADS et ACA.

2.2 Shuffled Complex Evolution of the University of the Arizona (SCEUA)

Étape 0 - Initialisation

Pour lancer l'algorithme, l'utilisateur doit fournir des valeurs pour ses paramètres internes :

$maxn$: le nombre maximal d'évaluations du modèle hydrologique;

p : le nombre de groupes de points (groupes de jeux de paramètres), identifié comme étant le nombre de complexes;

α : le nombre de transpositions générées au cours d'une évolution des points;

β : le nombre d'évolutions au sein d'un complexe avant le remaniement des complexes;

de même que ces vecteurs associés aux paramètres du modèle hydrologique à caler :

X_{min} : le vecteur constitué des valeurs minimales que peuvent prendre les paramètres;

X_{max} : le vecteur constitué des valeurs maximales que peuvent prendre les paramètres;

X_0 : le vecteur constitué du jeu de paramètres initial (optionnel).

Étape 1 - Échantillonnage

Cet algorithme travaille avec une population de jeux de paramètres en générant un échantillon de points, qui représente la population de travail, selon le nombre de complexes défini (p) et le nombre de paramètres (n) que l'on retrouve au sein du problème d'optimisation. La taille de l'échantillon initial (z) devient :

$$z = p * (2n + 1) \tag{2.1}$$

La génération des jeux de paramètres qui constitueront l'échantillon se fait selon une distribution de probabilité uniforme; autrement dit, l'algorithme construit un échantillon où chaque jeu de paramètres a la même probabilité d'être choisi. Si l'utilisateur souhaite fournir un jeu de paramètre initial (X_0), celui-ci est incorporé comme l'un des points de l'échantillon. Finalement, chaque point est évalué au sein du modèle hydrologique pour en connaître son critère d'évaluation (valeur de la fonction objectif).

Étape 2 - Classement des points

Une fois la population initiale créée et l'évaluation initiale de chaque point effectuée, l'algorithme effectue un classement, dit primaire, des points selon la valeur de la fonction objectif associée de manière décroissante; c'est-à-dire que la méthode attribue le rang premier au meilleur point et le rang dernier au plus mauvais.

Rapport-Gratuit.com

Étape 3 - Partitionnement des points

Par la suite, SCEUA subdivise cette population classée pour former les complexes en distribuant équitablement les points parmi tous les complexes et en s'assurant que chacun comporte des points de bonne et de mauvaise qualité. Chaque complexe a un nombre de points (m) équivalent et de grandeur :

$$m = 2n + 1 \tag{2.2}$$

Le premier complexe se verra distribuer les points ayant les rangs :

$$p * (k - 1) + 1 \tag{2.3}$$

où p représente le nombre total de complexes et k est un paramètre incrémenté ($k = 1, \dots m$), le second complexe aura les points ayant les rangs :

$$p * (k - 1) + 2 \tag{2.4}$$

le troisième complexe aura les rangs :

$$p * (k - 1) + 3 \tag{2.5}$$

et ainsi de suite.

Étape 4 - Évolution des complexes

La constitution des complexes étant complétée, l'algorithme peut maintenant faire évoluer chacun des complexes selon le « Competitive Complex Evolution algorithm » (CCE). Celui-ci est basé sur le système de recherche de descente de Nelder et Mead (1965) qui est une méthode heuristique inspirée du principe de recuit simulé. Sommairement, l'algorithme CCE sélectionne aléatoirement un sous-groupe de points issus du même complexe et le plus mauvais de ces points est transposé par réflexion dans l'espace paramétriques en direction du centre de masse composé par les autres points du sous-groupe. Ainsi, CCE effectue un nombre α de transpositions au sein du même sous-groupe de façon à faire évoluer les points en direction des zones de bonne qualité de l'espace paramétrique. Le travail effectué par CCE est repris β fois au sein de chacun des complexes.

Étape 5 - Remaniement des complexes

Lorsque chaque complexe a subi le procédé d'évolution des points, tous les complexes sont rassemblés en une seule population et le processus de classement des points décrit à l'étape 2 est recommencé ainsi que le processus de répartition des points de l'étape 3.

Étape 6 - Vérification du critère d'arrêt

Finalement, une vérification est faite quant au nombre d'évaluations du modèle hydrologique effectué jusqu'à ce stade de l'optimisation en le comparant avec le nombre d'évaluations maximal ($maxn$). Si le nombre maximal d'évaluations ($maxn$) est atteint (critère d'arrêt), la recherche de l'algorithme SCEUA est terminée et la meilleure solution trouvée (jeu de paramètres) est retenue. Sinon, il y a retour à l'étape 4 pour recommencer le processus d'évolution des complexes.

2.3 Dynamically Dimensioned Search (DDS)

Étape 1 - Identification des intrants

Pour lancer l'algorithme, l'utilisateur doit fournir des valeurs pour ses paramètres internes :

r : le pas de perturbation des paramètres;

m : le nombre maximal d'évaluations du modèle hydrologique;

de même que ces vecteurs associés aux paramètres du modèle hydrologique à caler :

X_{min} : le vecteur constitué des valeurs minimales que peuvent prendre les paramètres;

X_{max} : le vecteur constitué des valeurs maximales que peuvent prendre les paramètres;

X_0 : le vecteur constitué du jeu de paramètres initial (optionnel).

Étape 2 - Évaluation initiale

Cet algorithme travaille avec un seul individu (point) qu'il perturbera par la suite. Il calcule d'abord la valeur de la fonction objectif ($F(X_0)$) pour un jeu de paramètres initial (X_0). Ce dernier peut être fourni par l'utilisateur ou généré aléatoirement par DDS. Puisqu'aucune perturbation n'a encore été effectuée, la meilleure valeur de la fonction objectif trouvée (F_{best}) jusqu'alors est égale à la valeur de la fonction objectif initiale ($F(X_0)$), le meilleur jeu

Rapport-gratuit.com
LE NUMERO 1 MONDIAL DU MÉMOIRES

de paramètres trouvé (X_{best}) est égal au jeu de paramètres initial (X_0) et le nombre d'évaluations du modèle est égal à un ($i = 1, \dots, m$).

Étape 3 - Sélection des paramètres perturbés

L'algorithme utilise une loi de probabilité définie comme suit pour sélectionner lesquels des paramètres, qui constituent le vecteur du meilleur point trouvé (X_{best}), seront perturbés :

$$P(i) = 1 - Ln(i) \,/\, Ln(m) \tag{2.6}$$

C'est donc dire que chaque paramètre (X_{best}^{j} $o\grave{u}\, j = 1, \dots, nombre\ de\ param\grave{e}tres$) du meilleur point trouvé a une probabilité de sélection égale à l'équation (2.6). Selon la loi définie précédemment, plus le nombre d'évaluations du modèle hydrologique (i) augmente, plus la probabilité de sélection de chaque paramètre ($P(i)$) devient faible. Si aucun paramètre n'est choisi en raison d'une probabilité de sélection ($P(i)$) trop faible, l'algorithme sélectionnera un seul paramètre de façon aléatoire parmi le meilleur jeu de paramètres trouvé (X_{best}). Tous les paramètres sélectionnés feront partie d'un ensemble (Y') composé des paramètres (X_{best}^{j}) qui seront appelés à subir une perturbation.

Étape 4 - Perturbation des paramètres

Chaque paramètre sélectionné pour subir une perturbation ($X_{best}^{j} \in Y'$) est modifié selon la formule suivante :

$$X_{new}^{j} = X_{best}^{j} + r * (X_{max} - X_{min}) * N(0,1) \tag{2.7}$$

où X_{new}^{j} représente le paramètres après la perturbation et $N(0,1)$ représente une distribution normale centrée réduite. Suivant la formule décrite ci-haut, il pourrait advenir que le nouveau paramètre modifié (X_{new}^{j}) dépasse les bornes maximales (X_{max}^{j}) ou minimales (X_{min}^{j}) fixées pour ce paramètre. Dans une telle situation, un processus de réflexion est opéré. D'une part, si la valeur du nouveau paramètre (X_{new}^{j}) est plus grande que la valeur maximale permise (X_{max}^{j}), la réflexion est effectuée comme suit :

$$X_{new}^{j} = X_{max}^{j} - \left(X_{new}^{j} - X_{max}^{j}\right) \tag{2.8}$$

Le processus de réflexion ne doit toutefois pas entrainer un dépassement de la valeur minimale (X_{min}^{j}) du paramètre. Si tel est le cas, la valeur du nouveau paramètre est repositionnée sur la valeur maximale (X_{max}^{j}). Si la valeur du nouveau paramètre (X_{new}^{j}) est plus petite que la valeur minimale permise (X_{min}^{j}), la réflexion est alors :

$$X_{new}^{j} = X_{min}^{j} + \left(X_{min}^{j} - X_{new}^{j}\right) \tag{2.9}$$

De même, si le processus de réflexion mène à un dépassement de la valeur maximale (X_{max}^{j}), la valeur du nouveau paramètre est repositionnée sur la valeur minimale (X_{min}^{j}).

Étape 5 - Évaluation après perturbation

Une fois les perturbations effectuées et le processus de réflexion complété, l'algorithme évalue la valeur de la fonction objectif ($F(X_{new})$) du nouveau jeu de paramètres (X_{new}). Si celle-ci ($F(X_{new})$) est mieux que la meilleure valeur de la fonction objectif trouvée (F_{best}) à ce stade de l'optimisation, le meilleur jeu de paramètres (X_{best}) devient égal au nouveau jeu de paramètres (X_{new}). Sinon, le nouveau jeu de paramètres (X_{new}) est rejeté et le meilleur jeu de paramètres trouvé, à ce stade, est conservé.

Étape 6 - Vérification du critère d'arrêt

Finalement, une vérification est faite quant au nombre d'évaluations du modèle hydrologique effectué (i) jusqu'alors en le comparant avec le nombre d'évaluations maximal permis (m). Si ce nombre (m) est atteint, la recherche de l'algorithme DDS est terminée. Sinon, il revient à l'étape 3 pour recommencer le processus de sélection et de perturbation.

2.4 Mesh Adaptive Direct Search (MADS)

Étape 0 - Initialisation

Comme pour les deux précédents algorithmes, l'utilisateur doit d'abord fournir des valeurs pour les paramètres internes de MADS :

m : le nombre maximal d'évaluations du modèle hydrologique;

Δ_{k_0} : le dimensionnement initial du maillage de l'espace paramétrique (valeur fournie ou valeur par défaut);

de même que ces vecteurs associés aux paramètres du modèle hydrologique à caler :

X_{min} : le vecteur constitué des valeurs minimales que peuvent prendre les paramètres;

X_{max} : le vecteur constitué des valeurs maximales que peuvent prendre les paramètres;

S_0 : le vecteur constitué des jeux de paramètres initiaux (un seul ou plusieurs).

L'algorithme évalue le ou les jeux de paramètres initiaux (S_0) et identifie le plus performant de ceux-ci comme étant le meilleur jeu de paramètre (p_k).

Étape 1 - La recherche

Tout d'abord, il est à noter que par les diverses fonctionnalités qu'offre le logiciel d'optimisation NOMAD, l'étape de la recherche peut être remplacée complètement par toute autre stratégie de recherche que l'utilisateur souhaite incorporer à l'algorithme d'optimisation MADS (Le Digabel, 2011). De plus, lors de la première itération de cette méthode d'optimisation, l'étape de la recherche n'est pas effectuée et l'algorithme passe directement à l'étape 2.

MADS est basée sur le principe de séparation de l'espace paramétrique en un réseau de mailles à dimensions constantes (le treillis). Le paramètre interne Δ_k gère le dimensionnement du maillage et chaque jeu de paramètres considéré par l'algorithme se trouve à l'intersection des mailles (Audet et Dennis, 2006; Le Digabel, 2011).

L'étape de la recherche génère aléatoirement un certain nombre de jeux de paramètres positionnés à une distance égale à la dimension du maillage (Δ_k) autour du meilleur point trouvé jusqu'ici (p_k). La figure 2.1 présente cette méthode de génération des points.

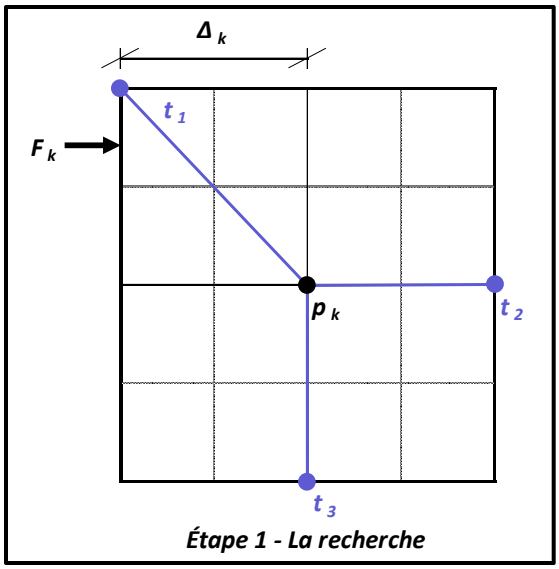

Figure 2.1 Directions visitées lors de l'étape de la recherche sur le motif de points (F_k) à partir du meilleur point trouvé (p_k) pour un problème à deux paramètres où les points t_1, t_2 et t_3 sont tirés aléatoirement sur le motif de points (F_k)

Les points sont alors évalués selon un ordonnancement établi à partir d'un modèle quadratique[1] qui se perfectionne à mesure que le nombre de points évalués au sein du problème réel d'optimisation augmente. Les points générés aléatoirement sur le motif et ayant un meilleur potentiel (bonne valeur de la fonction objectif selon le modèle quadratique) sont ainsi évalués au sein du problème réel dans un premier temps. Si l'algorithme obtient un jeu de paramètres de meilleure qualité, il est retenu comme étant le nouveau meilleur jeu de paramètres trouvé (p_k), l'étape de la recherche est déclarée comme étant réussie et l'algorithme passe à l'étape 3a. Sinon, cette étape est déclarée insatisfaisante et l'algorithme poursuit à l'étape 2.

---

[1] Modèle quadratique : fonction mathématique qui approxime localement la surface de l'espace paramétrique sous la forme $a_i * x_i^2 + b_i * x_i * x_j + c_i * x_i + d_i$, où $i \wedge j = 1, \ldots, nombre\ de\ paramètres$, $i < j$ et où a, b, c et d sont des constantes ($\neq 0$).

Étape 2 - La sonde

La satisfaction des conditions d'optimalité (obtention d'un optimum local) est issue de la stratégie utilisée à l'étape de sonde (Audet et Dennis, 2006). L'algorithme construit un motif de points (F_k) autour du meilleur jeu de paramètres trouvé (p_k) à une distance égale à la dimension du maillage (Δ_k) suivant différentes directions orthogonales de l'espace paramétrique. La figure 2.2 présente un exemple de directions orthogonales envisageables lors de la construction d'un motif de points (F_k).

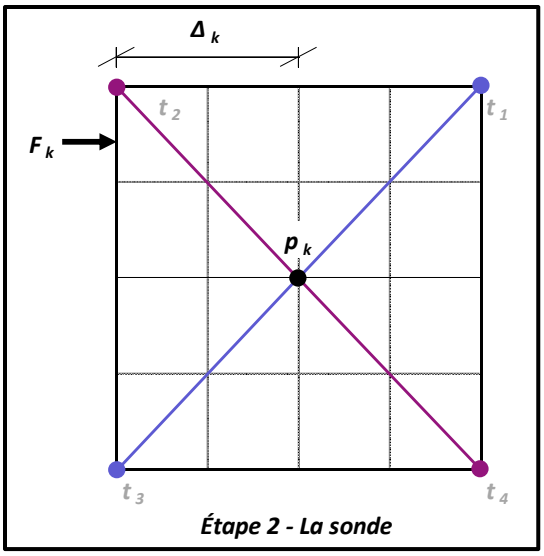

Figure 2.2 Directions orthogonales 2N visitées
lors de la génération du motif de points (F_k)
à partir du meilleur point trouvé (p_k)
pour un problème à deux paramètres (N = 2) où
les points t_1 et t_3 sont orthogonaux
avec les points t_2 et t_4

De cette façon, l'algorithme explore le voisinage situé à une distance égale à la dimension du maillage (Δ_k) depuis le meilleur jeu de paramètres trouvé (p_k). MADS visite un certain nombre de points dans les directions orthogonales faisant partie du motif (F_k) et ceux-ci sont alors évalués. De la même manière que pour l'étape de la recherche, l'ordonnancement des points se fait à partir du modèle quadratique. Si l'algorithme obtient un jeu de paramètre de meilleure qualité au sein du motif de points (F_k), il est retenu comme étant le nouveau

meilleur jeu de paramètres trouvé (p_k), l'étape de sonde est déclarée comme étant réussie et l'algorithme passe à l'étape 3a. Sinon, cette étape est déclarée insatisfaisante et il passe à l'étape 3b.

Étape 3a - Mise à jour des paramètres (réussie)

Si l'étape de la recherche ou celle de la sonde a été déclarée comme étant réussie, la dimension du maillage (Δ_k) est augmentée; c'est-à-dire que les prochains points générés seront à une distance plus grande du meilleur jeu de paramètres trouvé (p_k). Le facteur d'agrandissement du maillage (τ) est de 4 fois la dimension du maillage existant (Δ_k) (valeur par défaut de l'algorithme). Ce faisant, on accélère la recherche si on vient de trouver une nouvelle direction de descente. La figure 2.3 illustre le procédé de mise à jour du dimensionnement du maillage (Δ_k) qui est utilisé à l'itération subséquente de l'algorithme. L'algorithme passe ensuite à l'étape 3c.

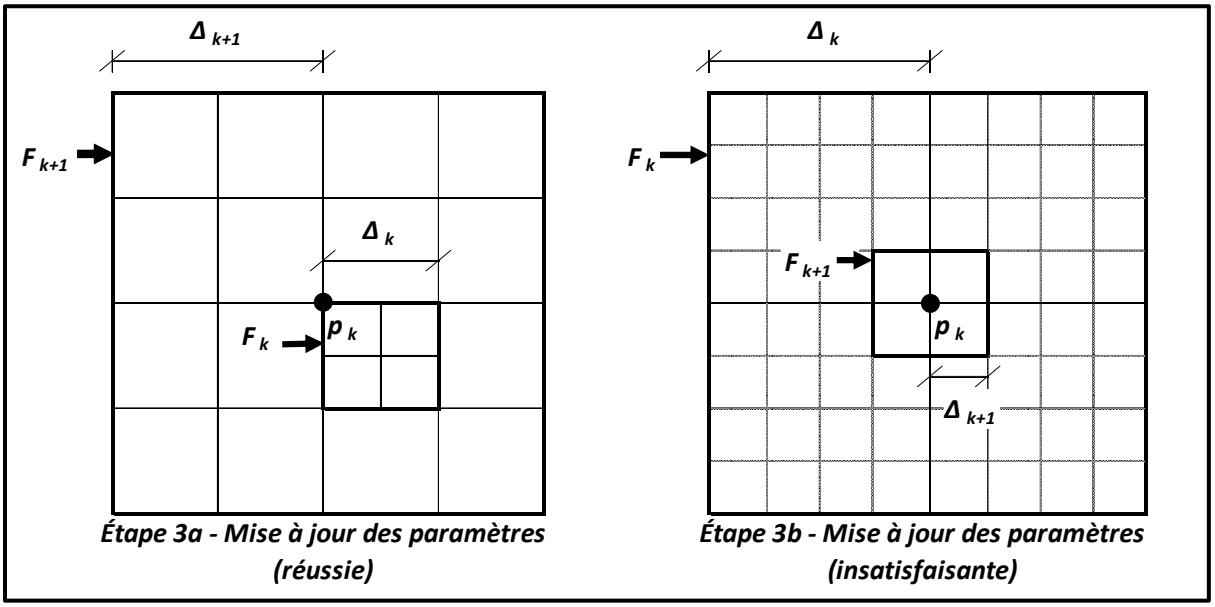

Figure 2.3 Mise à jour du dimensionnement du maillage (Δ_k) à l'itération de MADS k+1 dans l'éventualité où la recherche ou la sonde soit déclaré réussi (étape 3a) et dans l'éventualité où la recherche et la sonde soient déclarés insatisfaisants (étape 3b)

Étape 3b - Mise à jour des paramètres (insatisfaisante)

Si l'étape de la recherche et celle de la sonde ont été déclarées comme étant insatisfaisantes, la dimension du maillage (Δ_k) est réduite de façon à raffiner le procédé de recherche autour du meilleure jeu de paramètres (p_k). Le facteur de réduction du maillage ($1/\tau$) est de ¼ fois la dimension du maillage existant (Δ_k) (valeur par défaut de l'algorithme). C'est ce procédé de convergence vers une dimension toujours plus faible du maillage qui permet de qualifier la méthode de semi-exacte puisqu'elle garantit d'une certaine façon que le jeu de paramètres final est le meilleur point du voisinage. La figure 2.3 illustre le procédé de mise à jour du dimensionnement du maillage (Δ_{k+1}) qui est utilisé à l'itération subséquente de l'algorithme. L'algorithme passe ensuite à l'étape 3c.

Étape 3c - Vérification des critères d'arrêt

Finalement, une vérification est faite quant au nombre d'évaluations du modèle hydrologique effectué jusqu'alors en le comparant avec le nombre d'évaluations maximal du modèle hydrologique (m). Si le nombre maximal d'évaluations (m) est atteint (critère d'arrêt), la recherche de l'algorithme MADS est terminée. Sinon, il y a retour à l'étape 1 pour recommencer le processus de recherche. Signalons qu'il peut y avoir un arrêt si la maille minimale est atteinte (la valeur employée est celle par défaut), et ceci, même si la condition d'arrêt sur m n'est pas satisfaite. Seul l'arrêt conditionnel à l'atteinte de la maille minimale assure l'obtention d'un optimum local.

2.5 Algorithme Complètement Aléatoire (ACA)

Étape 0 - Initialisation

Pour lancer l'algorithme, l'utilisateur doit fournir une valeur pour le paramètre interne :

m : le nombre maximal d'évaluation du modèle hydrologique;

de même que ces vecteurs associés aux paramètres du modèle hydrologique à caler :

X_{min} : le vecteur constitué des valeurs minimales que peuvent prendre les paramètres;

X_{max} : le vecteur constitué des valeurs maximales que peuvent prendre les paramètres;

X_0 : le vecteur constitué du jeu de paramètres initial.

Étape 1 - La recherche aléatoire

L'algorithme évalue la fonction objectif (F_0) du jeu de paramètre initial (X_0) et l'identifie comme étant le meilleur point trouvé (X_{best}) associé au meilleur critère objectif trouvé (F_{best}). Pour chaque évaluation du modèle hydrologique ($i = 1, \dots, m$) qui suivra, l'algorithme construit un jeu de paramètre (X_i) qui est tiré aléatoirement suivant une distribution uniforme entre les bornes minimales (X_{min}) et maximales (X_{max}). La fonction objectif ($F(X_i)$) est alors évaluée en effectuant une simulation du modèle hydrologique. Si celle-ci ($F(X_i)$) est plus performante que la meilleure fonction objectif trouvée (F_{best}), le meilleur jeu de paramètres (X_{best}) devient égal au jeu de paramètres construit aléatoirement (X_i). Sinon, le jeu de paramètres aléatoire (X_i) est rejeté et le meilleur jeu de paramètres trouvé à ce stade est conservé. L'algorithme recommence la génération aléatoire d'un point (X_{i+1}) jusqu'à ce que le nombre d'évaluations du modèle hydrologique (i) soit égal au nombre maximal d'évaluations permis (m).

Rapport-Gratuit.com

CHAPITRE 3

Objectifs de la recherche

À la lumière de ce qui précède, ce chapitre expose l'objectif principal de ce mémoire ainsi que les objectifs secondaires qui en découlent. Ces objectifs ont été identifiés suite à la définition de la problématique du calage des modèles hydrologiques coûteux en temps de calcul.

3.1 Objectif principal

La recherche vise à évaluer le comportement des algorithmes d'optimisation sélectionnés lorsqu'ils sont employés pour le calage de modèles hydrologiques coûteux en temps de calcul, considérant deux versions du modèle HYDROTEL (distribué et à base physique; Fortin et al., 2001), l'une à 10 paramètres et l'autre à 19 paramètres. On cherche d'abord à connaître laquelle (ou lesquelles) des méthodes testées offrira(ront) la plus grande efficacité pour le calage des modèles hydrologiques distribués et à base physique. Comme il a été mentionné à la section 1.3, l'évaluation de l'efficacité se caractérise par le meilleur compromis entre deux critères, soient le potentiel de l'algorithme à diminuer le nombre d'évaluations du modèle hydrologique (indicateur du temps de calcul), que l'on souhaite le plus faible possible, et la qualité de la solution finale (valeur de la fonction objectif), que l'on souhaite la meilleure possible. Pour ce faire, l'examen et l'analyse du comportement des algorithmes tout au long de la procédure d'optimisation seront effectués. Ce faisant, il sera possible de différencier les stratégies de recherches qui sont appropriées à la problématique de cette étude de celles qui ralentissent le processus de calage. Les résultats obtenus serviront à formuler des recommandations quant à l'utilisation des algorithmes considérés et à identifier des avenues prometteuses qui serviront potentiellement au développement de nouvelles méthodes d'optimisation plus adaptées. Bien qu'en modélisation hydrologique, la performance des calages est très souvent validée à l'aide d'une seconde série de données observées, cet aspect ne sera pas couvert dans le cadre de cette étude.

Rapport-gratuit.com
LE NUMERO 1 MONDIAL DU MÉMOIRES

3.2 Objectifs secondaires

En ce qui a trait aux objectifs secondaires, le premier consiste à exploiter les diverses fonctionnalités ou options disponibles au sein du logiciel NOMAD pour identifier la configuration de MADS la plus efficace pour le calage de modèles hydrologiques. N'ayant pas été développé spécifiquement pour le calage de modèles hydrologiques, NOMAD sera plus largement étudié quant à sa mise en place. D'abord, le recours aux groupes de variables sera examiné. Cette fonctionnalité permet de créer des sous-groupes distincts de paramètres pour lesquels MADS génère indépendamment les directions orthogonales qui composent le motif de points. Cette subdivision de l'espace des solutions lors de la génération des directions orthogonales pourrait permettre de mieux orienter les jeux de paramètres dans le processus de calage et ainsi devancer l'obtention d'un jeu de paramètres de bonne qualité. Cette manière d'aborder la problématique pourrait mener à une réduction potentielle du nombre d'évaluations du modèle hydrologique d'où l'intérêt d'examiner cette avenue. Il sera présenté à la section 4.3.2.3 comment la séparation des groupes de variables a été effectuée. Ensuite, une option supplémentaire quant à la façon de générer le motif de points au sein de l'étape de sonde sera étudiée. Une version récemment développée de MADS considère un motif dans lequel la moitié des directions orthogonales de sonde (celles contraires à la direction de descente des dernières améliorations) sont agrégées en une seule direction, ce qui pourrait, encore une fois, permettre de réduire le temps de calcul.

Le deuxième objectif secondaire consiste à évaluer l'impact du type de modèle hydrologique sur le comportement des algorithmes d'optimisation sélectionnés. Pour ce faire, ceux-ci seront évalués pour le calage du modèle hydrologique conceptuel et global HSAMI (Bisson et Roberge, 1983; Fortin, 2000), qui dispose d'une structure interne complètement différente de celle d'HYDROTEL. Ces expérimentations seront effectuées dans le but d'observer les similitudes et les différences de comportements de chacune des approches d'optimisation sur différents types de modèle. Cet objectif secondaire informera sur le niveau de complexité du problème de calage des modèles hydrologiques distribués et à base physique par rapport à celui du calage des modèles hydrologiques conceptuels et globaux.

CHAPITRE 4

Méthodologie

Ce chapitre décrit la méthodologie employée pour cette étude. La présentation des bassins versant étudiés est d'abord abordée, puis la méthodologie relative à la modélisation hydrologique est exposée. Une description détaillée des problèmes de calage des modèles hydrologiques définis pour ces travaux et la configuration des approches de calage sélectionnées sont ensuite présentées. Finalement, les moyens entrepris pour la présentation des résultats et l'analyse de ceux-ci terminent la description globale de la méthodologie.

4.1 Bassins versants étudiés

Cette section présente une description générale des bassins versants étudiés dans le cadre de la présente recherche : les emplacements et les particularités géographiques tels que la topographie, les types de sols et l'occupation du territoire; les données météorologiques et hydrométriques utilisées; et la justification du choix des bassins versants.

4.1.1 Description générale des bassins versants

Le premier bassin choisi est le bassin versant Cowansville, sous-bassin de la rivière Yamaska située dans la région de la Montérégie au Québec. La figure 4.1 illustre son emplacement sur le territoire de la rivière Yamaska ainsi que les exutoires (1) de la rivière Yamaska, lequel se déverse au fleuve St-Laurent, et (2) du bassin versant Cowansville. Celui-ci étant un bassin de tête de la rivière Yamaska, il se situe entre les latitudes 45°7' et 45°20' et les longitudes -72°44' et -72°11', et sa superficie est de 215 km².

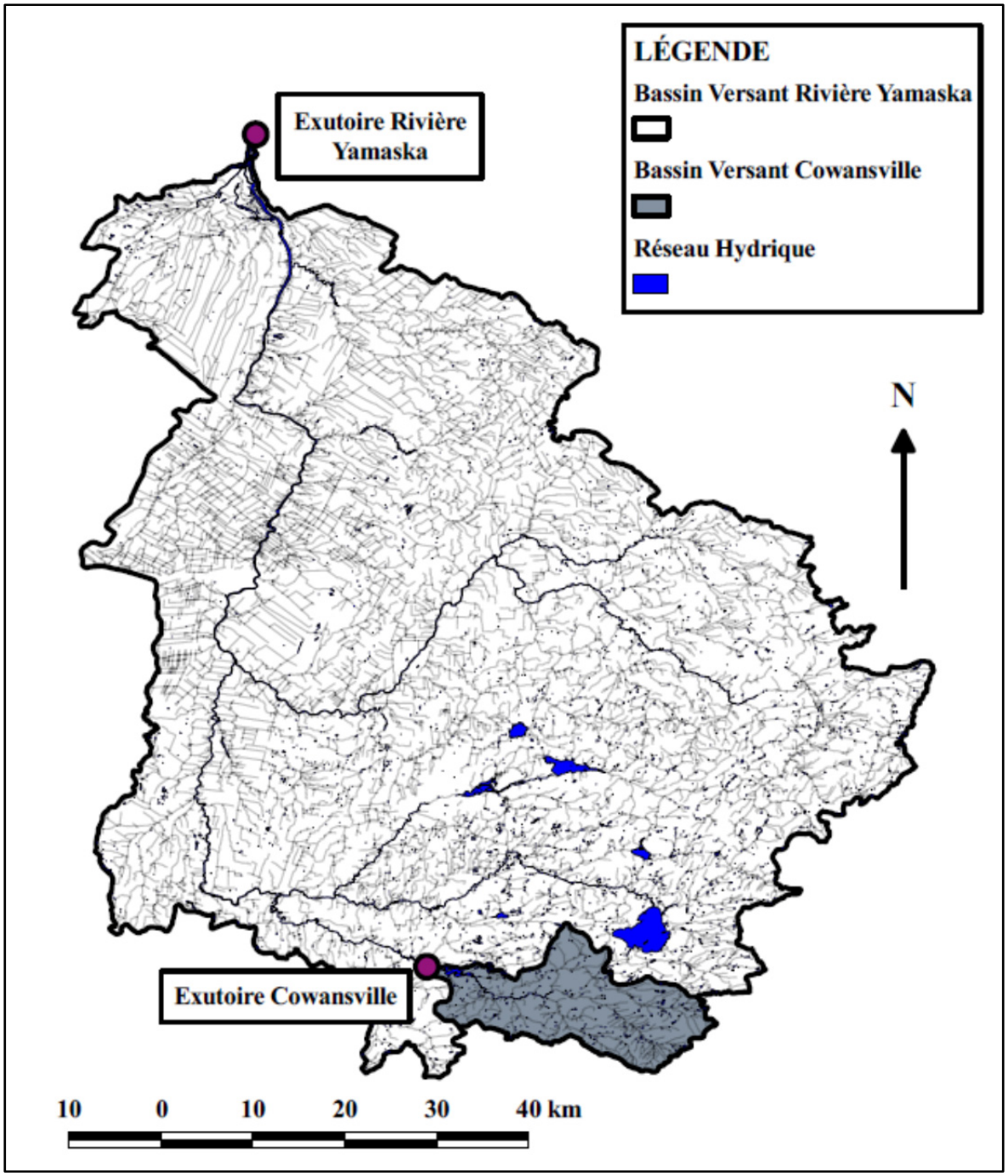

Figure 4.1 Emplacement du sous-bassin versant Cowansville
sur le bassin versant de la rivière Yamaska

Le bassin versant Cowansville recoupe deux grandes régions physiographiques : les Basses Terres du St-Laurent et les Appalaches. La section sud-est du bassin versant atteint la région

appalachienne et l'on y retrouve une altitude supérieure à 275m. La portion restante du bassin est associée à la physiographie des Basses Terres du St-Laurent où la topographie est plutôt plane et l'altitude varie entre 100 et 200m. Bien que le relief soit peu pentu, une légère fraction de celui-ci est occupée par des terres agricoles ou des terres d'élevages. Le bassin versant est donc composé en grand partie de zones forestières denses à composition majoritairement feuillues. Le sol est supposé uniforme sur l'ensemble du bassin versant et composé de sable et de loam (d'après la base de données utilisée pour la modélisation hydrologique).

Le second bassin versant sélectionné est Ceizur, sous-bassin de la rivière Gatineau située dans la région de l'Outaouais au Québec. La figure 4.2 montre l'emplacement du bassin versant Ceizur sur le territoire de la rivière Gatineau. L'exutoire du bassin versant Ceizur se trouve directement en amont du réservoir Baskatong. Ce dernier constitue une importante étendue ayant une superficie de plus de 400 km² et créée artificiellement par l'implantation de la centrale hydroélectrique Mercier. Ceizur reste cependant un bassin versant à apports totalement naturels et il est le bassin le plus au Nord, se trouvant en tête de la rivière Gatineau. Il se situe entre les latitudes 47°7' et 48°8' et entre les longitudes -75°50' et -74°20'; sa superficie est de 6928 km². Le Bouclier canadien est la région physiographique qui recouvre la totalité du bassin versant Ceizur. Bien que sa superficie soit très grande, la variation de l'altitude est relativement faible, soit d'environ 500m au Nord et de 315m au Sud. Ce bassin versant quasi-inhabité et sans terres agricoles comporte un pourcentage de conifères très élevé au nord du territoire et on dénote une augmentation progressive du ratio de feuillus lorsque l'on se dirige vers le Sud jusqu'à une dominance nette des feuillus à l'extrême sud du bassin. Toujours d'après la base de données employée pour la modélisation, le sol comporte deux zones homogènes : du sable limoneux au Nord et du loam sableux au Sud.

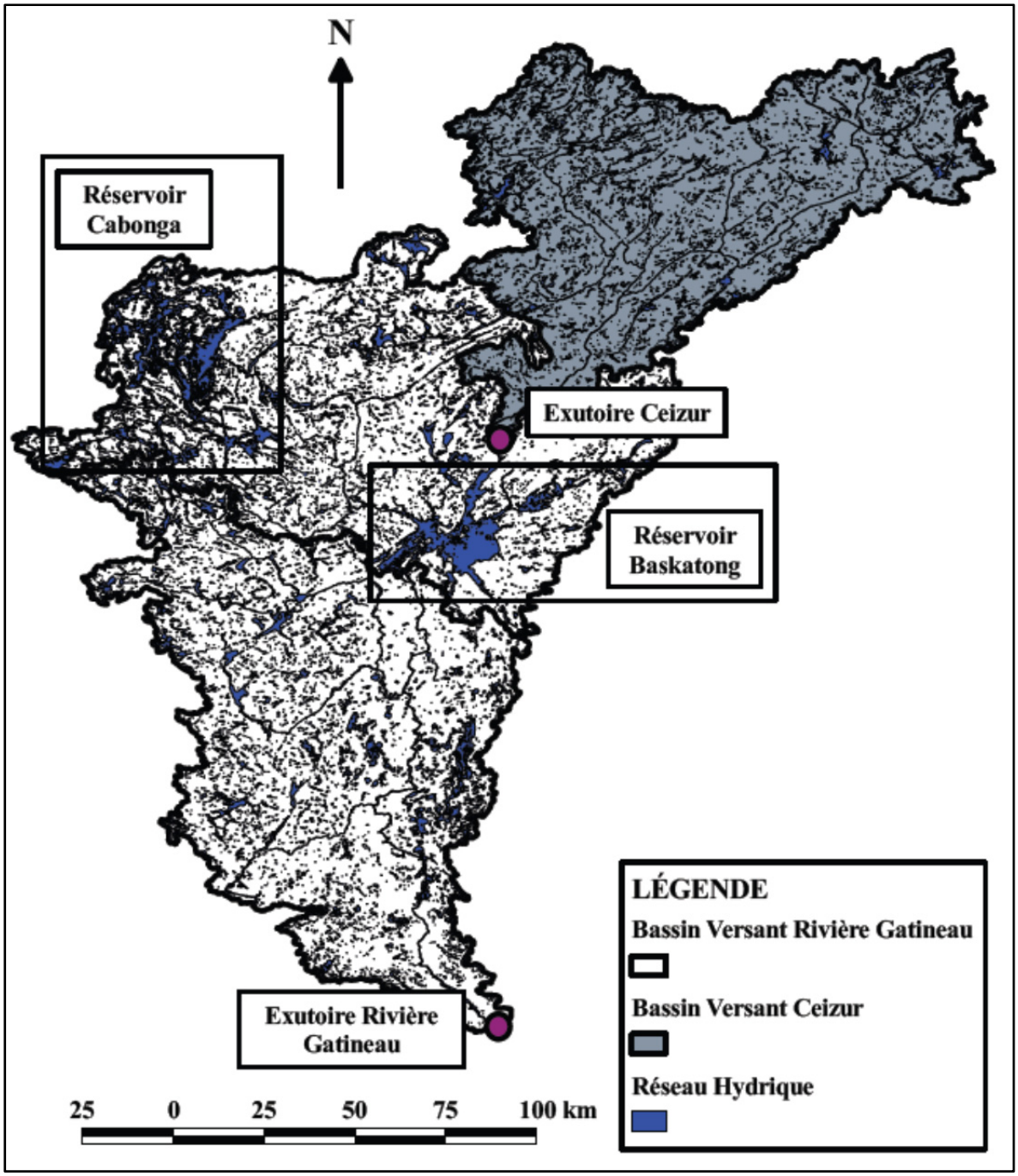

Figure 4.2 Emplacement du sous-bassin versant Ceizur
sur le bassin versant de la rivière Gatineau

4.1.2 Données hydrométéorologiques historiques

Les données météorologiques utilisées pour la modélisation des deux bassins versants incluent, sur une base journalière, les températures minimales, les températures maximales ainsi que les précipitations liquides totales. En ce qui concerne la modélisation hydrologique pour le bassin de Cowansville, les données météorologiques proviennent de la base de données du Centre d'expertise hydrique du Québec (CEHQ - voir la section Remerciements) où les températures et précipitations sont disponibles sur une grille à une résolution de 10 kilomètres. Il s'agit de données interpolées au moyen de la technique du krigeage, à partir des données disponibles aux stations météorologiques du Programme de surveillance du climat du ministère du Développement durable, de l'environnement, de la faune et des parcs du Québec. Compte-tenu de la taille réduite de Cowansville, uniquement 26 stations météorologiques virtuelles (points de grille) ont une incidence sur le bassin. Pour Ceizur, ces données proviennent plutôt de la base de données du National Land and Water Information Service (NLWIS; [www.agr.gc.ca/nlwis-snite]) qui fournit également les températures et les précipitations journalières sur une grille à une résolution de 10 kilomètres. Ces données émanent des archives climatiques de Ressources naturelles Canada et sont interpolées à partir du logiciel ANUSPLIN qui emploie la méthode d'interpolation des splines plaque mince (Hutchinson, 1989). La superficie importante de Ceizur entraîne la présence de 210 stations météorologiques virtuelles sur le territoire. Les données hydrométriques journalières observées à l'exutoire du bassin Cowansville proviennent du CEHQ, et celles du bassin Ceizur proviennent d'Hydro-Québec.

4.1.3 Justification du choix des bassins versants

Premièrement, les bassins versants Cowansville et Ceizur ont été sélectionnés puisqu'ils ont tous deux des caractéristiques physiques (superficie, végétation, occupation du territoire, etc.) très différentes (section 4.1.1). Deuxièmement, la mise en place de la modélisation hydrologique distribuée et à base physique nécessite un large éventail de données physiographiques à une échelle spatiale très fine, ainsi que le prétraitement de ces données. Dans le cas des deux bassins à l'étude, les étapes de prétraitement et la constitution des bases

de données nécessaires à la modélisation hydrologique distribuée avaient déjà été effectués. Troisièmement, ces deux bassins versants ont été choisis puisque de bonnes connaissances quant à la qualité des bassins versants modélisés sont acquises. En ce qui a trait au bassin versant Cowansville, il est difficile d'obtenir, suite au processus de calibration, quel qu'il soit, un jeu de paramètres d'aussi bonne qualité que pour Ceizur. Cela pourrait être lié entre-autres à la qualité des données physiographiques et météorologiques ou à l'utilisation d'un pas de temps journalier pour la modélisation, lequel peut s'avérer grossier pour un bassin de l'ordre de 200 km². De façon générale, la valeur de la fonction objectif finale obtenue à partir du critère de Nash-Sutcliffe sur ce bassin versant est tout près de 0,65. Selon les critères établis par Ritter et Munoz-Carpena (2013), la qualité de la modélisation hydrologique du bassin versant Cowansville se situe à la limite entre une modélisation faible et moyenne. Contrairement à ce dernier, le bassin versant Ceizur offre facilement des jeux de paramètres finaux associés à un critère du Nash-Sutcliffe de 0,85. Toujours selon Ritter et Munoz-Carpena (2013), la qualité de la modélisation hydrologique du bassin Ceizur est plutôt considérée comme étant bonne. Toutes ces distinctions entre les deux bassins versants sélectionnés permettront de mieux supporter les analyses et conclusions par rapport à la performance des méthodes d'optimisation lorsqu'employés pour le calage des modèles hydrologiques.

4.2 Modélisation hydrologique

Premièrement, l'utilisation d'un modèle hydrologique coûteux en temps de calcul doit être évidemment incluse dans le cadre de ces travaux pour répondre à l'objectif principal, soit d'évaluer le comportement de différentes méthodes d'optimisation lorsqu'employés à des fins de calage. Le modèle hydrologique distribué et à base physique HYDROTEL (Fortin et al., 2001) a été identifié comme modèle central de ces travaux. Celui-ci pourra d'autant plus servir à l'étude de la configuration du logiciel d'optimisation NOMAD qui offre une variété de fonctionnalités et d'options de calage. Deuxièmement, dans le but d'évaluer l'impact du type de modèle hydrologique sélectionné sur le comportement des méthodes d'optimisation, le modèle hydrologique conceptuel et global HSAMI (Bisson et Roberge, 1983; Fortin et al.,

2000) a été utilisé. Une brève description de chacun des modèles hydrologiques utilisés est exposée ci-après et des explications sont données quant à leur mise en place.

4.2.1 HYDROTEL

HYDROTEL (Fortin et al., 2001) a été développé par l'Institut national de la recherche scientifique (INRS). Ce modèle est spécialement conçu pour exploiter les méthodes de télédétection ainsi que les SIG dans le but de concevoir un modèle ayant une résolution spatiale très fine par rapport à la superficie des bassins versants, ce qui impose toutefois des coûts en temps de calcul.

À l'aide d'un logiciel complémentaire, appelé PHYSITEL, un bassin versant est construit à partir de propriétés physiographiques telles que la topographie, l'occupation du sol, le réseau hydrographique et le type de sol en exploitant, de ce fait, la télédétection et les SIG. À la sortie de PHYSITEL, le bassin est discrétisé en plusieurs unités spatiales de calcul nommées UHRH (Unités Hydrologiques Relativement Homogènes). La figure 4.3 représente les divers sous-modèles choisis dans le cadre de cette étude pour constituer le schéma de simulation du modèle hydrologique HYDROTEL. Les détails concernant ces sous-modèles sont présentés par Fortin et al. (2001) et d'autres sous-modèles employables y sont également définis. Suite à la simulation hydrologique, il est alors possible de comparer les débits simulés par le modèle avec les données hydrométriques observées en rivière de manière à évaluer la qualité du modèle hydrologique construit à partir d'HYDROTEL. Par ce type de modélisation, la division spatiale des bassins versants en diverses UHRH rend possible l'observation du comportement hydrologique non pas uniquement à l'exutoire du bassin, mais également à la sortie de chaque unité spatiale de calcul (UHRH).

Rapport-Gratuit.com

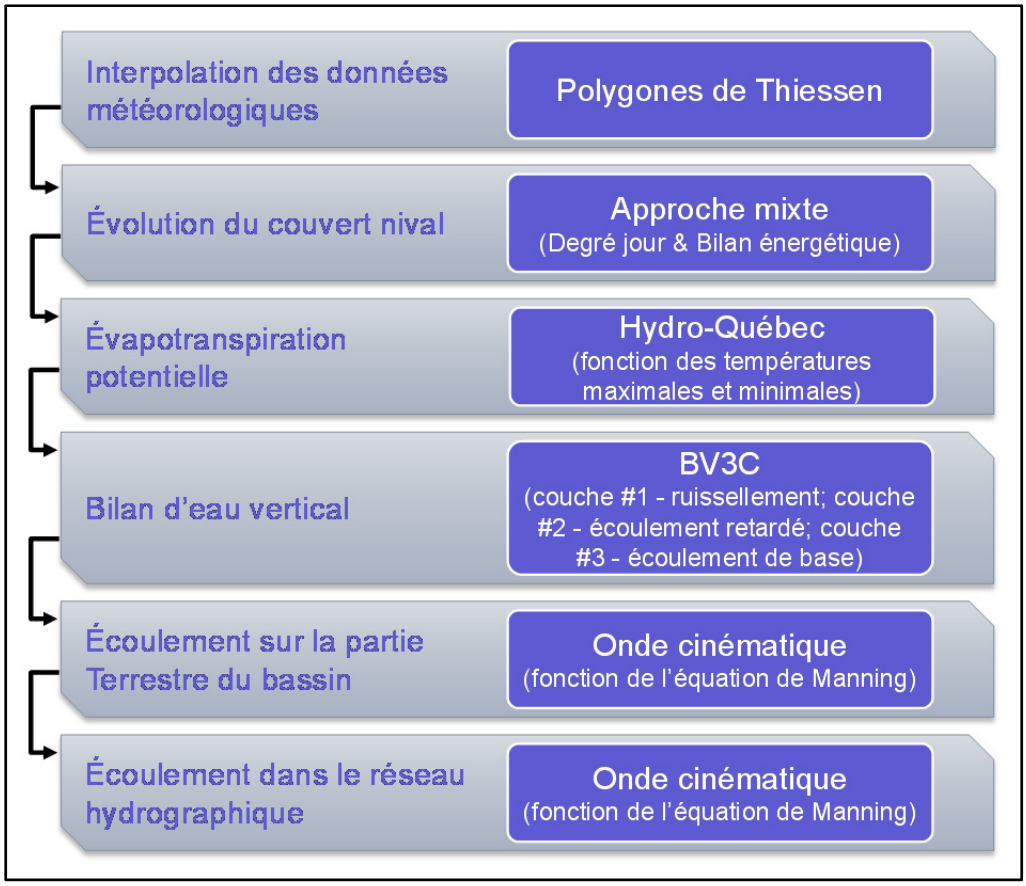

Figure 4.3 Schéma des sous-modèles employés pour la simulation à l'aide du modèle hydrologique HYDROTEL dans le cadre de cette étude

Les sous-modèles d'HYDROTEL utilisés dans cette étude comportent 27 paramètres internes qui sont détaillés dans le tableau-A I-1, à l'Annexe I. Bien qu'HYDROTEL soit identifié comme étant un modèle hydrologique à base plus physique, certains des paramètres représentent plutôt des coefficients additifs ou correctifs qui ne sont pas essentiels à un calage de bonne qualité. Certains des paramètres ont pu être fixés sur la base d'expérience antérieure acquise avec le modèle (Fortin et al., 2001; Turcotte et al., 2007; Poulin et al., 2011) et/ou parce qu'ils affichent une moins grande sensibilité lors du calage. Par conséquent, deux versions d'HYDROTEL ont été mise en place, soit : l'une à 19 paramètres où cinq paramètres additifs ou correctifs et trois paramètres additionnels sont fixés (voir tableau-A I-2 de l'Annexe I), et l'autre à 10 paramètres où cinq paramètres additifs ou correctifs et douze paramètres additionnels sont fixés (voir tableau-A I-2 de l'Annexe I). Il

est à noter que, lorsque modifiés, les paramètres du sous-modèle d'écoulement sur la partie terrestre du bassin inclus au sein de la version d'HYDROTEL à 19 paramètres entraînent une augmentation du temps de calcul variant de 15 à 25 % pour une seule simulation de ce dernier relativement à une simulation à l'aide du modèle à 10 paramètres. Cette augmentation de temps de calage est causée par la nécessité de régénérer, pour chaque UHRH, les ordonnées d'un hydrogramme qui décrit la structure temporelle d'écoulement (hydrogramme géomorphologique - écoulement sur la partie terrestre du bassin, figure 4.3).

Les deux versions du modèle hydrologique HYDROTEL ont été employés comme deux problèmes d'optimisation distincts servant, d'une part, à l'évaluation des comportements des algorithmes d'optimisation pour le calage de modèles hydrologiques coûteux en temps de calcul et, d'autre part, pour l'analyse de l'impact du type de modèle sur le comportement des méthodes de calage. Finalement, la version d'HYDROTEL à 10 paramètres a servi également à l'étude des différentes possibilités de configuration de l'optimisation au moyen de NOMAD.

4.2.2 HSAMI

Le modèle hydrologique HSAMI (Bisson et Roberge, 1983; Fortin et al., 2000), employé par la société d'état Hydro-Québec, est de type conceptuel et global. Contrairement aux modèles distribués qui nécessitent une quantité importante d'information sur le bassin versant, le modèle global HSAMI exige uniquement la superficie du bassin versant modélisé, la latitude moyenne, les températures moyennes journalières et les précipitations journalières sous forme solide et liquide moyennées sur l'ensemble du bassin versant. Le principe de fonctionnement de HSAMI est basé sur une suite conceptuelle de réservoirs et d'hydrogrammes unitaires qui simulent les divers processus hydrologiques tels que l'accumulation et la fonte de neige, l'évapotranspiration, l'infiltration ainsi que les écoulements verticaux et horizontaux. Le modèle HSAMI comporte 23 paramètres qui sont détaillés dans le tableau-A I-3, à l'Annexe I. La plupart de ceux-ci sont sans signification

Rapport-gratuit.com
LE NUMERO 1 MONDIAL DU MÉMOIRES

physique, mais sont malgré tout liés à des propriétés intrinsèques du modèle. Aucun paramètre n'a été fixé au sein du modèle lors des calages.

4.2.3 Justification du choix des modèles utilisés

Tout d'abord, la justification principale du choix d'HYDROTEL comme modèle hydrologique coûteux en temps de calcul vient du fait qu'une seule simulation du modèle pour une année de données sur un bassin versant de taille moyenne (5 000 km²) nécessite environ une minute sur un ordinateur à processeur Intel Core i7 de 3.40-GHz. Comme le processus de calage peut nécessiter plusieurs milliers de simulations, le temps de calcul atteint rapidement les cent heures de calage. De plus, sa nature plus physique, qui permet d'exploiter les connaissances des utilisateurs par rapport à leur bassin versant, a permis la mise en place de deux versions d'HYDROTEL (à 10 et 19 paramètres) qui serviront comme deux bancs d'essai distincts pour les expérimentations. Ces choix ont été faits considérant que ces deux versions permettent d'évaluer le comportement des algorithmes d'optimisation sélectionnés lorsqu'ils sont confrontés au calage d'un modèle « rigide » (HYDROTEL à dix paramètres) et au calage d'un modèle « souple » (HYDROTEL à dix-neuf paramètres). Le modèle est qualifié de « rigide » lorsqu'il est constitué d'un nombre restreint de paramètres. Ceci tend à contraindre l'espace de recherche des solutions et occasionne inévitablement une réduction de l'équifinalité par rapport à un modèle comportant un plus grand nombre de dimensions (paramètres). À l'inverse, la qualification « souple » provient du fait que le nombre de paramètres est beaucoup plus élevé entraînant ainsi une équifinalité plus importante.

Bien que la problématique principale soit associée aux modèles coûteux en temps de calcul, tel que mentionné précédemment, le modèle conceptuel et global HSAMI permet d'évaluer l'impact du type de modèle hydrologique sélectionné sur le comportement des méthodes d'optimisation. HSAMI est grandement employé au Québec. Par sa grande facilité d'adaptation à tous les types de bassins versants, HSAMI est sans équivoque l'un des modèles les plus employés parmi la recherche scientifique québécoise et le milieu

opérationnel de la gestion de la production d'hydroélectricité. De plus, le nombre de paramètres (23 paramètres) que l'on retrouve au sein de HSAMI offre une dimensionnalité du même ordre de grandeur que celle offerte par la version d'HYDROTEL à 19 paramètres, appuyant ainsi les comparaisons faites entre les deux modèles.

4.3 Calibration des modèles hydrologiques

Cette section de la méthodologie présente les hypothèses sous-jacentes aux problèmes d'optimisation dont il est question dans le cadre de cette recherche scientifique. Il est très important de noter que les hypothèses posées lors de la construction des problèmes de calage s'appuient sur une problématique de calage réelle à laquelle un hydrologue pourrait être confronté. Ce contexte dit réaliste a été choisi de manière à ne favoriser aucun algorithme d'optimisation d'autant plus qu'aucune étude sur la performance de ces algorithmes n'a encore été publiée relativement à la problématique de calage du modèle hydrologique HYDROTEL. Les méthodes d'optimisation ont donc été adaptées au contexte de la problématique et non l'inverse (contexte adapté aux algorithmes d'optimisation). Par la suite, les hypothèses de configuration pour chacun des algorithmes sélectionnés seront illustrées en respectant toujours cette impartialité entre les diverses méthodes de calage.

4.3.1 Définition des problèmes d'optimisation

4.3.1.1 Évaluation des méthodes de calage pour les modèles hydrologique coûteux en temps de calcul

En premier lieu, il est nécessaire de bien comprendre la notion de temps de calcul telle que considérée dans le cadre de cette recherche. Le temps n'est pas évalué en termes d'unités temporelles (secondes, minutes, etc.), mais plutôt en nombre d'évaluations du modèle hydrologique. Ceci permet de généraliser la comptabilisation du temps de calcul, peu importe les infrastructures de calcul utilisées ou la façon dont les approches de calage sont programmées (interprétées vs compilées, par exemple).

Par ailleurs, la méthodologie proposée jusqu'alors nécessite une puissance de calcul non négligeable et les choix méthodologiques qui suivent ont été faits en considérant que l'ensemble des expérimentations devait pouvoir être réalisé à l'intérieur d'une période d'au plus un an à l'aide des installations informatiques à notre disposition.[2] Sur la base d'études réalisées antérieurement à cette recherche (Tolson et Shoemaker, 2007; Poulin et al., 2011), il a été établi que chaque méthode d'optimisation effectuerait un processus de calage comportant deux mille évaluations du modèle hydrologique HYDROTEL. La période de calage considérée pour le bassin versant Cowansville s'étend du 1er octobre 2000 au 30 septembre 2005, soit cinq années de données à un pas de temps de vingt-quatre heures. Celle pour le bassin versant Ceizur s'étend du 1er octobre 1988 au 30 septembre 1992, soit quatre années de simulation à un pas de temps de vingt-quatre heures. Les bornes inférieures et supérieures pour les versions à 10 et à 19 paramètres d'HYDROTEL sont spécifiées au tableau-A II-1 et au tableau-A II-2, à l'Annexe II, et ont été adaptées en fonction de connaissances acquises antérieurement avec ce modèle hydrologique. Chaque algorithme d'optimisation reprend 32 fois le processus de calage pour toutes les combinaisons 'modèle hydrologique-bassin versant' de manière à assurer un échantillonnage représentatif des tendances des comportements des méthodes de calage. Les 32 jeux de paramètres initiaux pour les problèmes à 10 et à 19 paramètres ont été définis de façon aléatoire, mais sont identiques pour chaque méthode d'optimisation (voir le tableau-A III-1, le tableau-A III-2a et le tableau-A III-2b, à l'Annexe III). La figure 4.4 représente le cadre expérimental pour l'évaluation des différentes méthodes d'optimisation pour le calage du modèle hydrologique coûteux en temps de calcul HYDROTEL.

---

[2] Serveur Intel Xeon X735 de 2.93-GHz (4 processeurs, 16 nœuds de calcul)

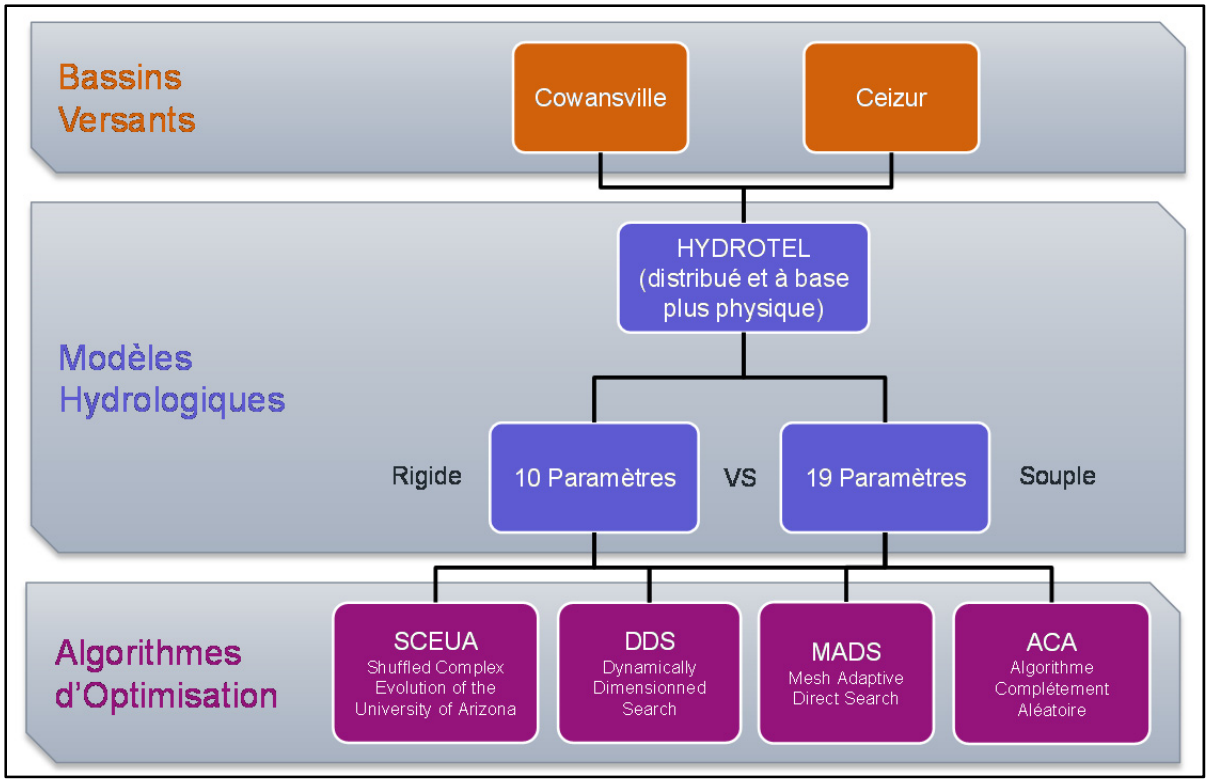

Figure 4.4 Cadre expérimental pour l'évaluation des méthodes d'optimisation pour le calage de modèles hydrologiques coûteux en temps de calcul

4.3.1.2 Étude des possibilités de configuration de l'optimisation dans NOMAD

Le logiciel d'optimisation NOMAD a été développé dans le but de répondre à différents problèmes d'optimisation de type « boîte noire ». Celui-ci comporte plusieurs fonctionnalités qui permettent son adaptation à la problématique de calage étudiée. Certaines fonctionnalités ont été ciblées comme étant potentiellement adaptées aux problèmes d'optimisation de cette étude et sont évaluées uniquement sur le bassin versant Cowansville de la rivière Yamaska modélisé par HYDROTEL à 10 paramètres afin de limiter les temps de calcul impliqués (voir section 4.2.1). La période de calage ainsi que les bornes inférieures et supérieures des paramètres d'HYDROTEL sont les mêmes que précédemment (section 4.3.1.1 et Annexe II). Pour chaque fonctionnalité étudiée, le calage à l'aide de NOMAD est repris 32 fois. Les 32 jeux de paramètres initiaux (même qu'à la section 4.3.1.1) sont définis de façon aléatoire, mais sont identiques pour chaque configuration du logiciel NOMAD qui est étudiée (voir le

tableau-A III-1, à l'Annexe III). La figure 4.5 représente le cadre expérimental pour l'étude des fonctionnalités d'optimisation de NOMAD. Les diverses fonctionnalités étudiées seront présentées plus tard (section 4.3.2.3).

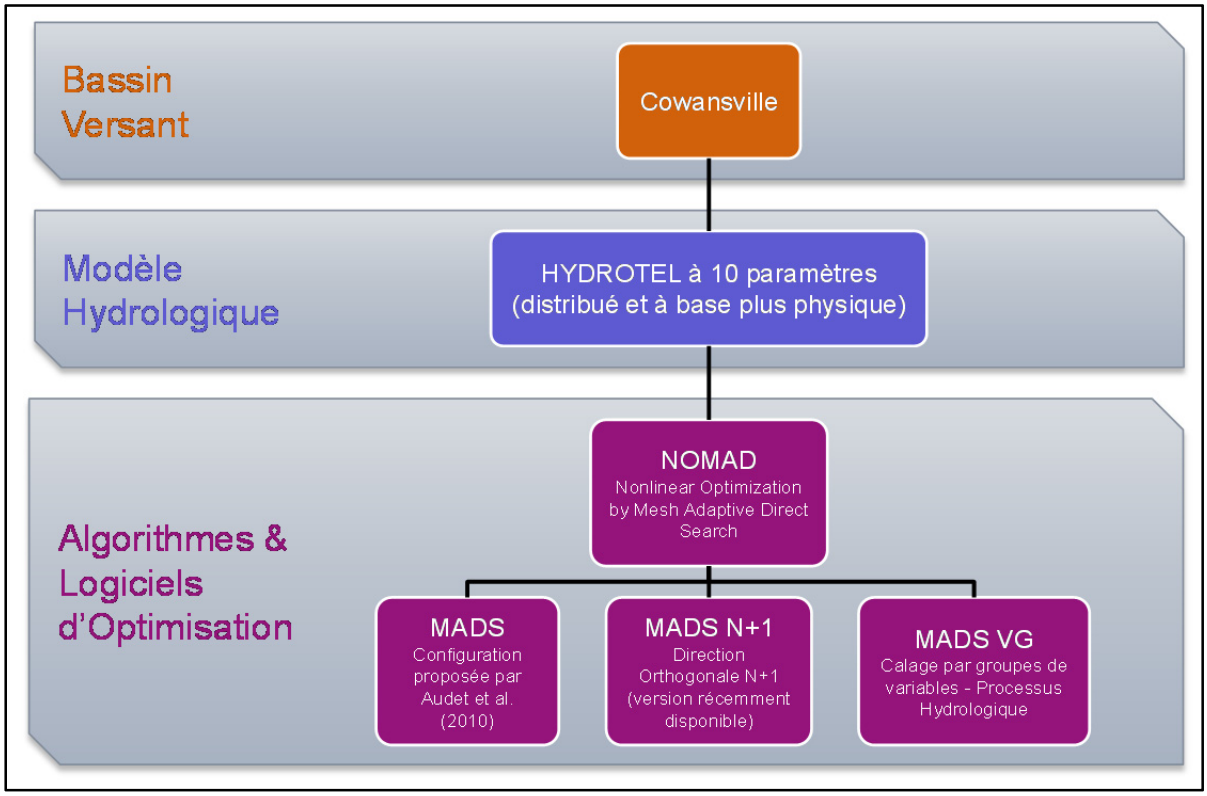

Figure 4.5 Cadre expérimental pour l'étude des fonctionnalités d'optimisation de NOMAD

4.3.1.3 Évaluation de l'impact du type de modèle hydrologique sur le comportement des algorithmes d'optimisation

Les travaux d'Arsenault et al. (2013) ont démontré que plusieurs algorithmes d'optimisation sont en mesure de bien caler le modèle hydrologique HSAMI lorsque l'on permet un nombre très élevé d'évaluations du modèle (plus de vingt mille). Toutefois, il est peu envisageable de permettre un nombre aussi élevé d'évaluations pour le calage d'HYDROTEL étant donné son large coût en temps de calcul et ainsi permettre une comparaison juste avec les travaux de ces auteurs. Pour évaluer l'impact du type de modèle sur le comportement des algorithmes d'optimisation sélectionnés, le problème de calage d'HSAMI a été limité au même nombre

d'évaluations permises que pour le calage d'HYDROTEL bien que ce modèle conceptuel et global ne consomme que très peu de temps de calcul (une simulation du modèle sur plusieurs années de données coûte moins d'une seconde). De cette façon, il est possible d'évaluer s'il existe des similitudes, ou différences, entre les problèmes d'optimisation d'HYDROTEL, distribué et à base physique, et d'HSAMI, conceptuel et global. Ainsi, les méthodes d'optimisation sont limitées à deux mille évaluations du modèle. Les bornes inférieures et supérieures pour les 23 paramètres d'HSAMI sont présentées au tableau-A II-3, à l'Annexe II. Les périodes de calage pour les bassins versants Cowansville et Ceizur sont les mêmes que celles employés pour le calage d'HYDROTEL (section 4.3.1.1). Chaque algorithme d'optimisation reprend 32 fois le processus de calage pour chacun des deux bassins versants modélisés. Les 32 jeux de paramètres initiaux pour le modèle HSAMI sont définis de façon aléatoire, mais sont identiques pour chaque méthode d'optimisation (voir le tableau-A III-3a, le tableau-A III-3b et le tableau-A III-3c, à l'Annexe III). La figure 4.6 représente le cadre d'expérimentation pour l'étude de l'impact du type de modèle sur les comportements des algorithmes d'optimisation.

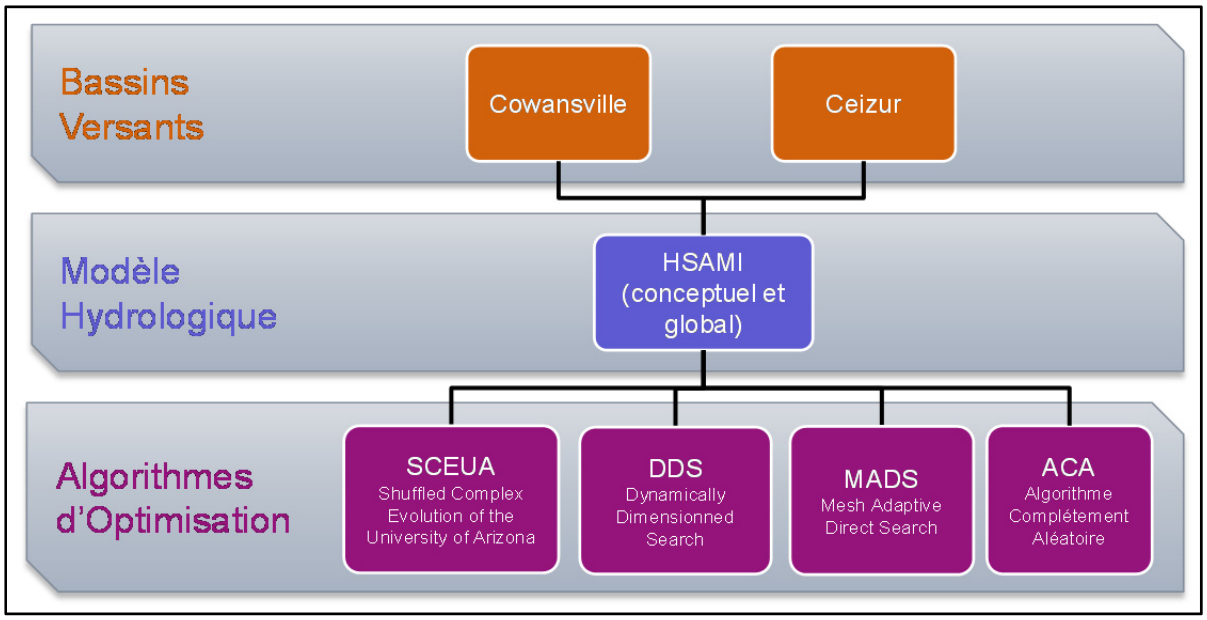

Figure 4.6 Cadre expérimental pour l'évaluation de l'impact du type de modèle hydrologique sur le comportement des algorithmes d'optimisation

4.3.2 Configuration des algorithmes

4.3.2.1 Shuffled Complex Evolution of the University of the Arizona (SCEUA)

Comme il a été mentionné au chapitre 2, la méthode d'optimisation SCEUA comporte des paramètres (3) dont il faut préalablement fixer les valeurs avant d'exécuter le processus de calage. Duan et al. (1994) ont procédé à une étude quant à la configuration de ces paramètres pour ainsi mieux orienter l'utilisateur. Ce sont ces valeurs de paramètres qui ont été utilisées dans la présente recherche. L'hypothèse de mettre en place l'algorithme selon les recommandations des auteurs vient illustrer l'esprit d'impartialité de cette étude envers les différentes approches de calage sélectionnées. Un pseudocode de l'algorithme SCEUA est présenté à l'Annexe IV (algorithme-A IV-1a-b), et les valeurs de paramètres utilisées y sont également indiquées.

4.3.2.2 Dynamically Dimensioned Search (DDS)

L'algorithme d'optimisation DDS est d'une grande simplicité en termes de mise en place. L'unique paramètre de DDS, qui correspond au pas de perturbation des paramètres du modèle hydrologique à calibrer, a été largement étudié par ses auteurs. Selon Tolson et Shoemaker (2007), la valeur du paramètre de DDS doit être suffisamment élevée pour obliger l'algorithme à s'enfuir des zones d'optima de piètre qualité, mais assez faible pour envisager l'exploration des zones étroites au voisinage des optima locaux. La valeur recommandée du pas de perturbation est de 0,2. Dans l'éventualité où le jeu de paramètre fourni initialement est déjà de bonne qualité, il est plutôt recommandé d'utiliser un pas de perturbation plus faible (r = 0,1) pour ainsi concentrer la recherche autour de ce point. Dans le cadre de la définition des problèmes d'optimisation de cette étude, le jeu de paramètre initial fourni lors du processus de calage n'est pas encore évalué, donc pas en état d'être jugé sur sa qualité. Par conséquent, la valeur du pas de perturbation du jeu de paramètre est de 0,2 tel que recommandé par Tolson et Shoemaker (2007). Un pseudocode de DDS est présenté à l'Annexe IV (algorithme-A IV-2).

Outre le pas de perturbation, DDS comporte une particularité dont il faut sans contredit tenir compte dans l'évaluation et l'analyse du comportement de ce dernier. À l'inverse des autres méthodes d'optimisation sélectionnées, la stratégie de recherche de DDS est adaptée au nombre maximal d'évaluations du modèle hydrologique fourni par l'utilisateur. Ce principe d'adaptation provient de la loi de probabilité présentée plus tôt à l'équation (2.6) du chapitre 2. Celle-ci calcule la probabilité de sélection des paramètres du modèle hydrologique qui seront perturbés. Ce calcul dépend justement du nombre maximal d'évaluations du modèle hydrologique et constitue la pierre angulaire de la stratégie de recherche adaptative de DDS. Ainsi, DDS assure, en fonction du budget de calage, l'obtention d'un jeu de paramètres final le plus satisfaisant possible. La figure 4.7 démontre ce comportement adaptatif où, par exemple, un budget de calage de 200 évaluations procure de façon moyenne un meilleur jeu de paramètres qu'un budget de calage de 500 interrompu après 200 évaluations.

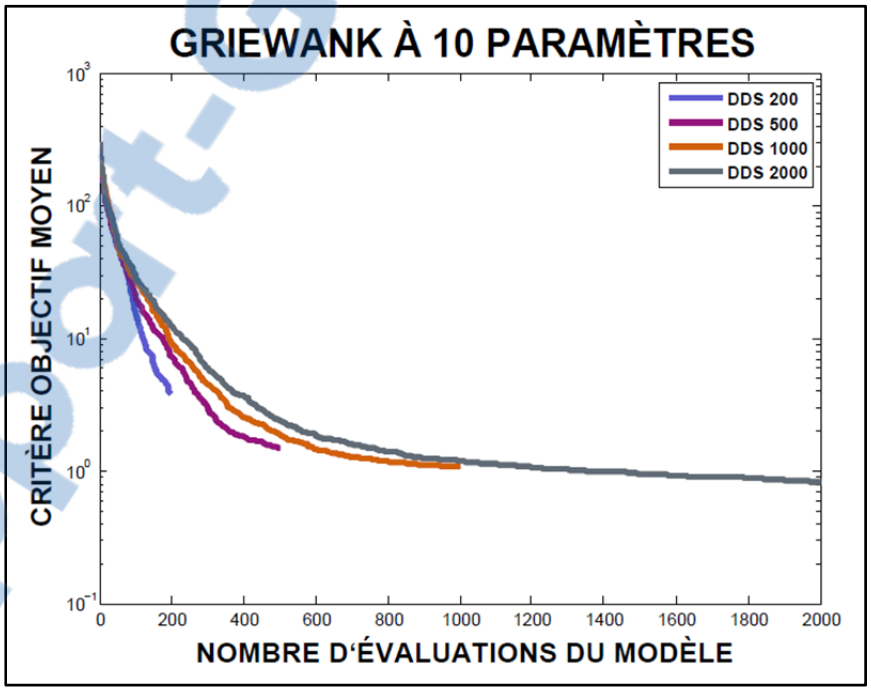

Figure 4.7 Comportement moyen de DDS issu de 32 calages de la fonction test GRIEWANK (Griewank, 1981) à 10 dimensions où le nombre maximal d'évaluations de la fonction objectif (m) est égal à 200, 500, 1000 et 2000; les mêmes 32 jeux de paramètres initiaux ont été repris pour chaque essai

Rapport-Gratuit.com

4.3.2.3 Nonlinear Optimization by Mesh Adaptive Direct Search (NOMAD)

Comme pour les méthodes d'optimisation présentées précédemment (sections 4.3.2.1 et 4.3.2.2), la configuration de NOMAD est celle qui est préconisée par les auteurs, Audet et al. (2010). Les paramètres de NOMAD sont donc ajustés aux valeurs par défaut. MADS, l'algorithme d'optimisation de NOMAD, est une méthode d'optimisation à descente directe; c'est-à-dire qu'elle trouve les directions dans l'espace paramétrique dans lesquelles elle pourra améliorer son critère objectif le plus rapidement possible. Cependant, comme sa stratégie de calage est basée sur la progression du meilleur jeu de paramètres trouvé, il devient alors victime de la qualité du jeu de paramètres fourni en début de calage. Cela occasionne donc un retard de progression inévitable par rapport aux autres algorithmes d'optimisation qui eux exploitent le principe de la recherche aléatoire en début de processus de calage pour mieux explorer l'espace de solution et ainsi cibler les secteurs qui sont de meilleure qualité. La flexibilité des fonctionnalités de NOMAD permet d'incorporer une manœuvre de recherche supplémentaire qui place MADS sur la même ligne de départ que les autres algorithmes. Cette méthode de recherche supplémentaire, appelée le « Latin Hypercube Search » (McKay et al., 1979), permet à NOMAD de générer des jeux de paramètres aléatoires au début de chaque itération de la procédure de recherche de MADS. Le paramètre de MADS qui correspond au nombre de points générés selon la méthode « Latin Hypercube Search » est réglé à une seule génération de manière éviter que trop d'importance soit accordée à cette stratégie de recherche aléatoire compte tenu des nombreuses itérations de l'algorithme MADS. De cette façon, NOMAD ne devrait pas strictement dépendre du jeu de paramètres initial puisqu'il a la chance, ou non, de générer un jeu de paramètres de meilleure qualité situé dans une zone de l'espace de solutions plus prisée. Un pseudocode de NOMAD (algorithme-A IV-3a-b) est présenté à l'Annexe IV. Cette configuration est celle qui est employée tant pour l'évaluation de l'efficacité des méthodes d'optimisation lors du calage des modèles hydrologiques coûteux en temps de calcul (objectif principal, section 3.1) que pour l'étude de l'impact du type de modèle hydrologique sur le comportement des algorithmes sélectionnés (section 3.2). En ce qui a

trait à l'étude des possibilités de configuration de NOMAD (section 3.2), deux autres ajustements sont examinés et présentés ci-dessous.

Calage par groupes de variables

La calibration par groupes de variables est explorée et celle-ci est basée sur la modélisation des processus hydrologiques au sein du modèle HYDROTEL. L'approche employée est établie de façon à rassembler les paramètres qui ont un impact dominant par processus hydrologique ou section de l'hydrogramme. En fonction des connaissances que nous détenons et des informations fournies par Fortin et al. (2001), les 10 paramètres d'HYDROTEL ont été répartis en quatre groupes formés en fonction de leur impact dominant sur : (1) la fonte de la neige, (2) les pointes de crues, (3) le volume d'eau annuel et (4) les débits d'étiage. Le tableau 4.1 présente le regroupement des paramètres.

Tableau 4.1 Regroupement des paramètres selon leur impact dominant sur le régime hydrique : (1) fonte de neige, (2) volume annuel, (3) pointes de crues et (4) débits d'étiage

HYDROTEL à 10 paramètres		
Numéro	**Description des paramètres**	**Groupe de variables**
3	Température de passage de la pluie en neige (°C)	1
7	Taux de fonte pour les conifères (mm/jour*°C)	1
8	Seuil de fonte en milieux de conifères (°C)	1
11	Taux de fonte pour le milieu non-forestier (mm/jour*°C)	1
12	Seuil de fonte en milieu non-forestier (°C)	1
13	Coefficient multiplicatif de l'évapotranspiration potentielle	2
14	Bilan vertical - limite inférieure de la couche #1 (m)	3
15	Bilan vertical - limite inférieure de la couche #2 (m)	3
16	Bilan vertical - limite inférieure de la couche #3 (m)	2
18	Bilan vertical - coefficient de récession (m/heure)	4

Le calage par groupes de variables pourrait offrir une amélioration de l'efficacité de NOMAD puisque cette façon d'aborder le problème d'optimisation permet de mieux orienter les jeux de paramètres dans le processus de calage et ainsi devancer l'obtention d'un jeu de

Rapport-gratuit.com
LE NUMERO 1 MONDIAL DU MÉMOIRES

paramètres de bonne qualité. Cette subdivision de l'espace des solutions permet de créer des sous-groupes distincts de paramètres où MADS génère, de façon indépendante pour chacun, les directions orthogonales qui composent le motif de points. MADS est alors en mesure de trouver plus facilement les directions qui accentuent la descente vers les optima locaux.

Mise à jour récente du logiciel NOMAD

Au cours de cette étude, une version récemment développée de l'algorithme MADS a été mise à la disposition des divers utilisateurs. Cette version a donc également été testée pour évaluer son potentiel d'efficacité. Dans cette nouvelle version, la construction du motif de points est modifiée (voir la description de l'algorithme MADS au Chapitre 2). La version antérieure générait le motif de points dans des directions orthogonales tout autour du meilleur point trouvé selon une quantité égale à deux fois le nombre de paramètres à optimiser. Un tel motif de points est nommé « directions orthogonales 2N » (Audet et al., 2010). La version récemment développée génère une quantité de points égale à une fois le nombre de paramètres à optimiser dans les directions orthogonales ayant une composante orientée dans le sens du vecteur de déplacement du meilleur point trouvé. Un point additionnel est également tiré de manière à agréger toutes les directions orthogonales non visitées en une seule direction résultante. Un tel motif de points est nommé « directions orthogonales N+1 » (Audet et al., 2014). Cette version de MADS donne lieu à une configuration différente de NOMAD, laquelle est intéressante pour cette étude. En effet, la diminution du nombre de points générés par motif pourrait faire en sorte que la convergence de l'algorithme vers un optimum local soit plus rapide (moins de points à évaluer à chaque itération de l'algorithme), donc plus efficace en termes de temps de calcul.

4.3.2.4 Algorithme Complètement Aléatoire (ACA)

L'ACA ne nécessite aucune configuration particulière puisqu'il ne fait que générer des jeux de paramètres aléatoirement à chaque évaluation du modèle hydrologique. Aucun paramètre

de l'algorithme ne doit être ajusté pour l'exécution du processus de calage. Un pseudo code de l'ACA est présenté à l'Annexe IV (algorithme-A IV-4).

4.4 Évaluation de l'efficacité des algorithmes d'optimisation sélectionnés

Pour répondre aux objectifs décrits précédemment au chapitre 3, les résultats des expérimentations (chapitre 5) seront présentés selon certaines méthodes d'analyse qui permettront de soutenir les conclusions apportées par cette recherche scientifique. Cette section présente les trois principales méthodes d'analyse employées : les profils de performance, l'analyse de la dispersion de la fonction objectif finale ainsi que l'étude de la progression de la fonction objectif en fonction du nombre d'évaluations du modèle hydrologique.

4.4.1 Profils de performance

Les profils de performance ont été initialement introduits par Dolan et Moré (2002) comme méthode d'analyse comparative pour la performance de différentes approches d'optimisation lorsqu'appliquées à une population de problèmes d'optimisation. Les profils de performance consistent à mettre en relation trois mesures de performance : la valeur de la fonction objectif, la probabilité de succès de l'algorithme d'optimisation et le temps de calcul. De manière à rendre l'analyse comparative visuellement agréable, l'une des variables est fixée à une valeur où l'on souhaite focaliser l'analyse. Il existe deux types de profils de performance soient ceux où l'on fixe la mesure temporelle à un moment précis de la procédure d'optimisation, soient ceux où l'on fixe la valeur de la fonction objectif à un seuil de référence. Dans les deux cas, aucune information n'est mise à l'écart, mais les résultats de l'analyse sont présentés de manière à accentuer l'examen sur la mesure qui est invariable. Dans le cadre de cette étude, la mesure temporelle est fixée puisque l'évaluation de la performance des algorithmes d'optimisation est principalement axée sur l'efficacité en temps de calcul. De manière à faciliter les explications sur le fonctionnement des profils de performance, un exemple est détaillé ci-après.

Prenons quatre algorithmes d'optimisation (AO_i où $i = 1, \dots, 4$) qui procèdent à la résolution de 32 problèmes d'optimisation ($p_j \in P'$ où $j = 1, \dots, 32$) représentés par la minimisation de la fonction test SCHWEFEL (Schwefel, 1981) à 10 paramètres (32 jeux de paramètres initiaux différents) avec un budget de deux mille évaluations pour chaque problème. Dans cet exemple, on cherche à analyser la performance des quatre approches à la fin de l'optimisation, soit à deux mille évaluations de la fonction SCHWEFEL. Les profils de performance obtenus sont illustrés à la figure 4.8.

En premier lieu, pour chacun des trente-deux problèmes d'optimisation ($p_j \in P'$), l'analyse comparative attribue le statut de vainqueur à l'algorithme d'optimisation qui a obtenu la meilleure valeur de la fonction objectif ($v^*_{p_j}$). La probabilité de succès initiale est alors calculée à partir de la fraction du nombre de victoires de l'algorithme d'optimisation sur le nombre de problèmes d'optimisation total. La probabilité de succès initiale est établie sur l'axe vertical. La figure 4.8 montre que l'algorithme AO_1 a été déclaré vainqueur 24 fois sur les 32 problèmes d'optimisation, pour une probabilité de succès initiale égale à 75 %. L'AO_3 a été victorieux huit fois sur 32 pour une probabilité de succès initiale de 25 % tandis que AO_2 et AO_4 n'ont obtenu aucune victoire les amenant ainsi à une probabilité de succès initiale nulle.

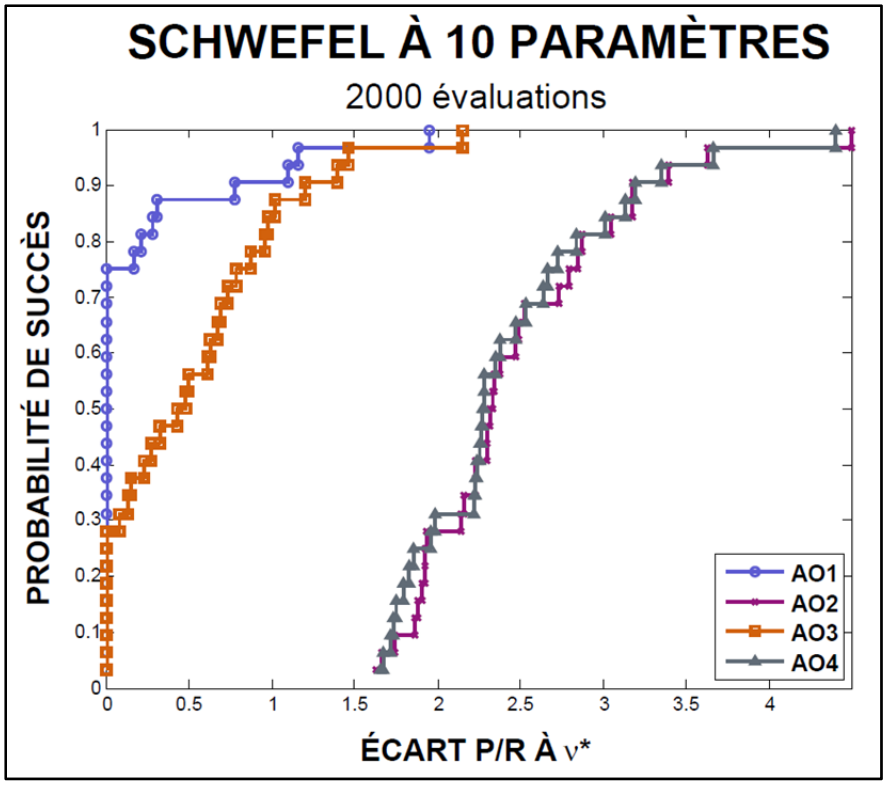

Figure 4.8 Profils de performance de quatre algorithmes d'optimisation issus de 32 minimisations de la fonction test SCHWEFEL (Schwefel, 1981) à 10 dimensions où le nombre d'évaluations est égal à 2000; les 32 jeux de paramètres de départ ont été initialisés aléatoirement pour chaque problème d'optimisation

En second lieu, l'analyse comparative indique l'évolution de la probabilité de succès des algorithmes d'optimisation si l'on consent à s'éloigner de la meilleure fonction objectif trouvée pour chaque problème ($v^*_{p_j}$). Un ratio ($r^i_{p_j}$) est calculé entre la valeur de la fonction objectif obtenue par l'algorithme d'optimisation ($v^i_{p_j}$) et la meilleure fonction objectif trouvée par toutes les approches de calage ($v^*_{p_j}$) pour chaque problème d'optimisation ($p_j \in P'$). Ce ratio ($r^i_{p_j}$) représente la grandeur de l'écart qui existe entre la fonction objectif obtenue par l'algorithme d'optimisation ($v^i_{p_j}$) et la meilleure fonction objectif trouvée ($v^*_{p_j}$). Les profils de performance démontrent alors la progression de la probabilité de succès à mesure que le ratio ($r^i_{p_j}$) augmente; autrement dit, plus on s'éloigne de la meilleur valeur de la fonction objectif trouvée ($v^*_{p_j}$), plus la probabilité de succès des algorithmes

d'optimisations augmentera jusqu'à l'obtention d'une probabilité de succès de 100 %, là où l'on obtient le plus grand ratio ($r_{p_j}^i$) pour chacun des algorithmes. Dans un souci de présenter les résultats le plus clairement possible, une transformation logarithmique en base deux a été effectué sur les ratios calculés ($r_{p_j}^i$) de manière à ajuster l'échelle horizontale. La figure 4.8 illustre que l'AO_1 obtient une probabilité de succès de 100 % avec son plus grand ratio ($r_{p_j}^1 = 1{,}924$) qui est plus faible que celui de l'AO_3 ($r_{p_j}^3 = 2{,}192$), et que l'AO_2 et l'AO_4 obtiennent leur probabilité de succès de 100 % avec des écarts ($r_{p_j}^2 = 4{,}497$ et $r_{p_j}^4 = 4{,}411$) beaucoup plus élevés.

Considérant les deux aspects de l'analyse comparative présentés précédemment, on cherche à identifier l'algorithme d'optimisation qui offrira, d'une part, le plus haut taux de succès initial, c'est-à-dire le plus grand nombre de déclarations victorieuses par rapport au nombre total de problèmes d'optimisation et, d'autre part, une probabilité de succès de 100 % acquise avec le moins d'écart possible par rapport à la meilleure valeur de la fonction objectif trouvée. Dans le cadre de cette étude, les profils de performance sont élaborés au nombre maximal d'évaluations permis pour le processus de calage (2000 évaluations), mais également en cours d'optimisation, lorsque le nombre d'évaluations est égal à 100, 250, 500, 1000 et 1500 évaluations, de manière à observer s'il y a des variations de performance de la part des algorithmes lorsque les budgets de calage sont réduits. Les profils de performance permettent d'identifier distinctement laquelle des méthodes d'optimisation choisies démontre la plus grande capacité à trouver un jeu de paramètres de bonne qualité et à réduire potentiellement le nombre requis d'évaluations du modèle hydrologique, ce qui cadre directement avec la définition de l'efficacité établie dans cette recherche.

4.4.2 Analyse de dispersion des valeurs finales de la fonction objectif

Cette méthode d'analyse consiste à étudier la distribution des données représentées par les 32 valeurs de la fonction objectif finale produite pour chaque combinaison 'algorithme-modèle hydrologique-bassin versant'. Cette analyse de variabilité de données est illustrée sous forme graphique en utilisant les boîtes à moustaches, développées par John Wilder Tukey dans les

années 1970 (McGill et al., 1978). Les limites de la boîte à moustaches sont le 25ième centile (premier quartile) et le 75ième centile (troisième quartile). Le 50ième centile représente la médiane et les valeurs extrêmes sont placées à l'extérieur des boîtes à moustaches si elles se trouvent en-deçà de 1,5 fois l'écart interquartile. D'un point de vue efficacité, l'algorithme d'optimisation le plus performant est celui qui montre la médiane la plus élevée (maximisation du critère de Nash-Sutcliffe) et qui obtient la plus petite dispersion, montrée d'une part par l'écart interquartile et d'autre part par l'étendue des moustaches. Cette méthode d'analyse met l'emphase sur l'évaluation de l'efficacité des approches de calage en fonction de leur capacité à trouver un jeu de paramètres de bonne qualité.

4.4.3 Étude de la progression de la fonction objectif selon le temps de calcul

Les résultats de calage obtenus par les différents algorithmes d'optimisation sont également présentés graphiquement sous la forme de courbes de progression de la meilleure valeur du critère objectif trouvé (1 moins le Nash-Sutcliffe) en fonction du nombre d'évaluations du modèle hydrologique. Pour chaque combinaison 'algorithme-modèle hydrologique-bassin', une courbe de progression moyenne de la fonction objectif pour les 32 calages est calculée, de même que deux courbes définissant l'enveloppe de progression, l'une étant la limite inférieure (les pires valeurs rencontrées) et l'autre étant la limite supérieure (les meilleures valeurs rencontrées). Cette représentation des résultats permet non seulement de comparer entre elles les courbes de progression moyennes de chacune des méthodes d'optimisation, mais également d'identifier les segments du processus de calage où les algorithmes d'optimisation adoptent des stratégies de recherche bénéfiques à la problématique d'efficacité en temps de calcul. Connaissant le fonctionnement interne de chacune des approches de calage, cette étude de la progression moyenne du critère objectif rend possible l'association de ces segments, où les méthodes fournissent (ou non) un rendement de progression souhaitable, aux stratégies de recherche employées par les algorithmes d'optimisation.

CHAPITRE 5

Résultats

Ce chapitre présente les résultats associés aux trois objectifs énoncés au chapitre 3 selon les méthodes d'analyse décrites à la section 4.4. Les observations pertinentes sur les résultats produits sont exposées.

5.1 Détermination de l'efficacité des algorithmes d'optimisation sélectionnés pour le calage des modèles hydrologiques coûteux en temps de calcul

5.1.1 Profils de performance

Les figures 5.1, 5.2, 5.3 et 5.4 présentent les profils de performances pour les bassins versants de Cowansville et de Ceizur modélisés à l'aide des versions à 10 et à 19 paramètres d'HYDROTEL. On constate d'abord que les résultats exposés sur ces quatre figures partagent des similitudes en ce qui a trait aux performances des algorithmes d'optimisation (efficacité). Les figures démontrent, de façon générale, que l'algorithme aléatoire (ACA) est d'une efficacité comparable à celle des méthodes de calage plus intelligentes telles que SCEUA et MADS en début de processus de calage seulement, soit avant 250 évaluations du modèle. Après 250 évaluations, l'algorithme aléatoire se distance peu à peu des autres méthodes et devient clairement une approche de calage moins adaptée aux problèmes d'optimisation définis dans ces travaux.

Également, les résultats démontrent sans contredit que DDS est la méthode de calage qui offre la plus grande efficacité puisqu'il obtient, d'une part, la plus haute probabilité de succès initiale pour la majorité des nombres d'évaluations considérés sur chaque figure (combinaisons 'modèle hydrologique-bassin versant'). Les cas de 1500 et 2000 évaluations pour la combinaison 'HYDROTEL 10 paramètres-Ceizur' (figure 5.3) font toutefois exception. Cette dernière représente un problème d'optimisation qualifié de « rigide » et de « facile » à caler où uniquement MADS est en mesure de surpasser DDS pour ces deux

Rapport-Gratuit.com

exceptions. Ce cas particulier tend à favoriser MADS puisque la rigidité diminue la victimisation liée au jeu de paramètre initial (ce point sera abordé plus en détails dans la discussion, à la section 6.1) et l'atteinte d'un jeu de paramètres de bonne qualité est beaucoup plus simple (facile). D'autre part, DDS obtient une probabilité de succès de 100 % avec le moins d'écart possible par rapport à la meilleure valeur de la fonction objectif trouvée (v^*) pour l'ensemble des résultats présentés aux figures 5.2, 5.3 et 5.4. On observe toutefois une situation différente à la figure 5.1 puisque, à 500 évaluations de modèle et plus, DDS acquiert la probabilité de succès de 100 % avec un plus grand écart par rapport à v^* que SCEUA et MADS. Cette situation est causée par un seul des 32 processus de calage considérés où DDS semble avoir éprouvé des difficultés. Par ailleurs, la figure 5.1 démontre que DDS obtient tout de même une forte probabilité de succès, soit 96,875 %, avec un écart plus petit par rapport à v^* pour tous les nombres d'évaluations. Les profils de performance dévoilent donc que DDS est l'algorithme d'optimisation le plus performant pour le calage des modèles hydrologiques coûteux en temps de calcul définis dans le cadre de cette étude.

De plus, on observe que, pour les deux bassins, l'algorithme SCEUA obtient plus rapidement la probabilité de succès de 100 % (plus faible ratio par rapport à v^*) que MADS pour tous les nombres d'évaluations lorsque le problème d'optimisation comporte un plus grand nombre de paramètres (19; figures 5.2 et 5.4). À l'inverse, MADS acquiert l'avantage sur SCEUA lors des processus de calage d'HYDROTEL à 10 paramètres. La dimensionnalité du problème d'optimisation semble donc avoir une influence sur le classement de SCEUA et MADS (voir la discussion à la section 6.3). DDS et ACA, pour leur part, gardent toujours leurs rangs respectifs de premier et dernier.

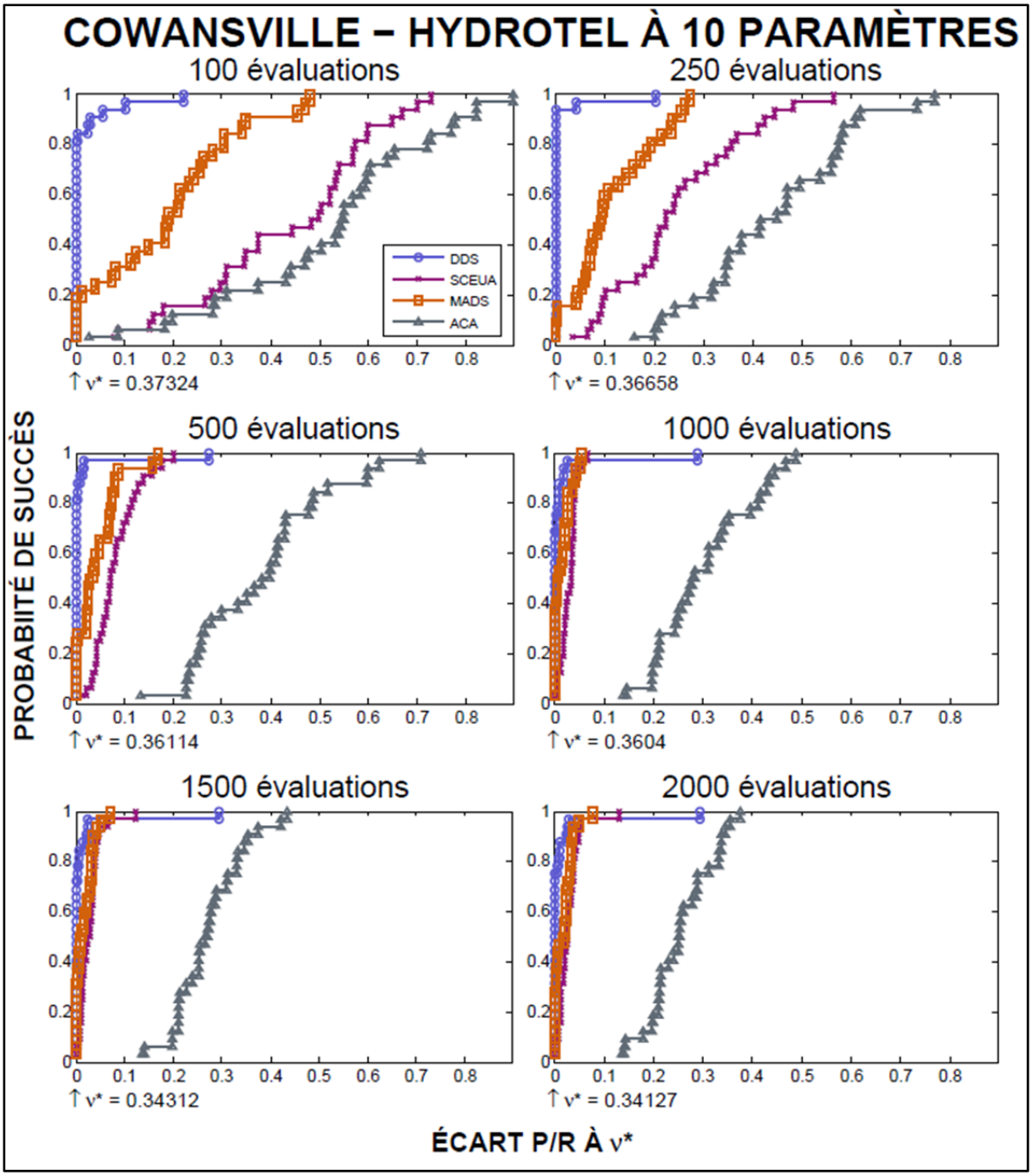

Figure 5.1 Profils de performance des algorithmes SCEUA, DDS, MADS et ACA pour le bassin versant Cowansville avec le modèle hydrologique HYDROTEL à 10 paramètres

Rapport-gratuit.com
LE NUMERO 1 MONDIAL DU MÉMOIRES

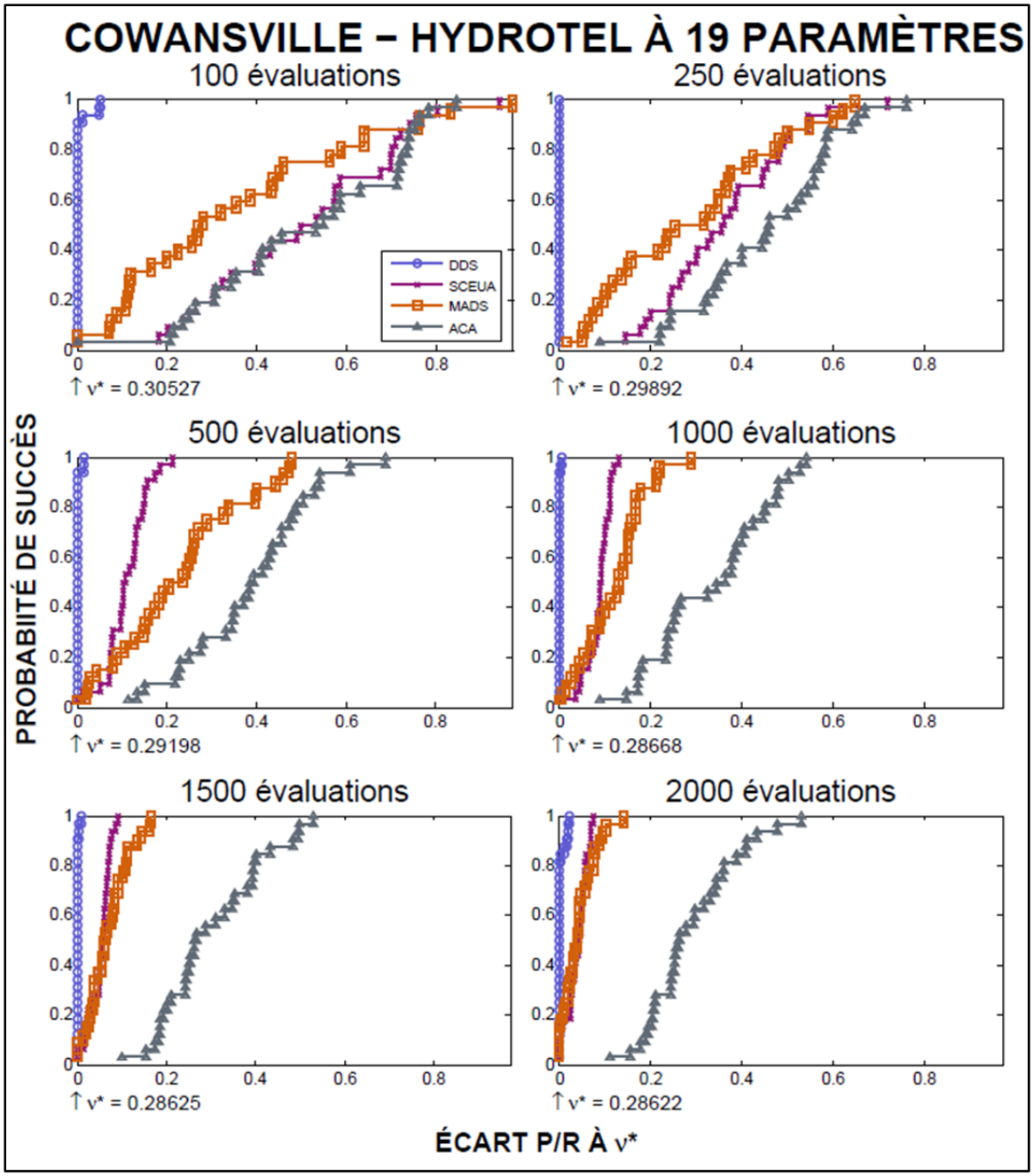

Figure 5.2 Profils de performance des algorithmes SCEUA, DDS, MADS et ACA pour le bassin versant Cowansville avec le modèle hydrologique HYDROTEL à 19 paramètres

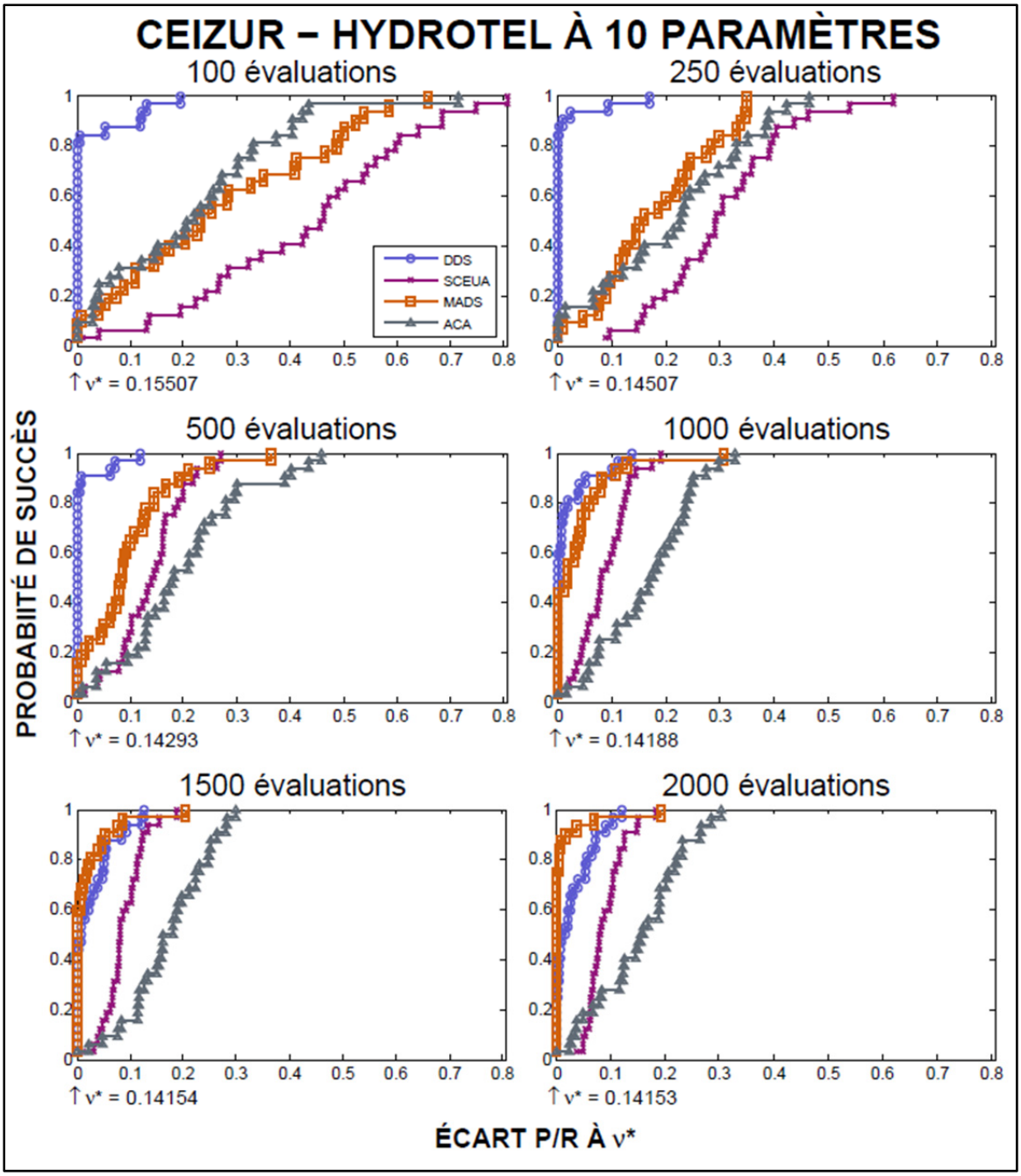

Figure 5.3 Profils de performance des algorithmes SCEUA, DDS, MADS et ACA pour le bassin versant Ceizur avec le modèle hydrologique HYDROTEL à 10 paramètres

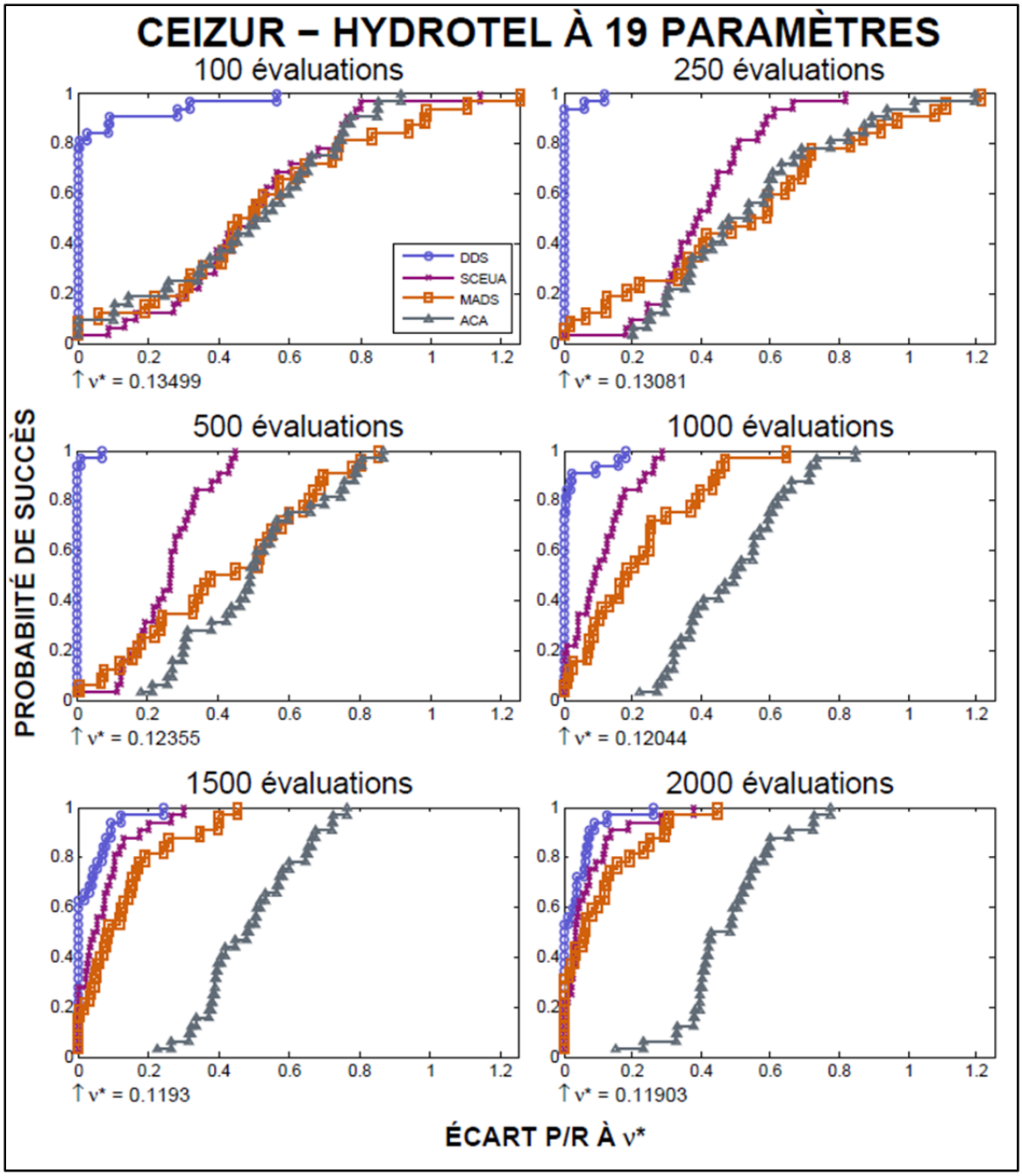

Figure 5.4 Profils de performance des algorithmes SCEUA, DDS, MADS et ACA pour le bassin versant Ceizur avec le modèle hydrologique HYDROTEL à 19 paramètres

5.1.2 Analyse de dispersion des valeurs finales de la fonction objectif pour les différents algorithmes d'optimisation

La figure 5.5 présente les boîtes à moustaches des valeurs de la fonction objectif (critère de Nash-Sutcliffe) obtenues pour les 32 calages au moyen de chacun des algorithmes, DDS, SCEUA et MADS, après 2000 évaluations du modèle hydrologique. Ces résultats démontrent en premier lieu que les trois approches de calage offrent, après 2000 évaluations, des valeurs finales de la fonction objectif qui sont très semblables. Bien que l'on puisse identifier que DDS semble avoir une performance légèrement supérieure à SCEUA et MADS dans la plupart des combinaisons 'modèle hydrologique-bassin versant' (à l'exception de 'HYDROTEL 10 paramètres-Ceizur' où MADS surpasse DDS), la différence observée entre les trois médianes est de l'ordre de ± 0,01 (ordre de grandeur très faible pour ce critère objectif). Ainsi, pour les cas examinés dans cette étude, les trois méthodes d'optimisation montrent des performances équivalentes quant à la valeur du critère objectif après 2000 évaluations des modèles.

En second lieu, on observe une dispersion très faible de la part de tous les algorithmes pour tous les bancs d'essai à l'exception de MADS pour le calage du modèle HYDROTEL à 19 paramètres (pour les deux bassins) où la dispersion est plus importante (voir discussion à la section 6.3). Outre cette exception, la dispersion généralement étroite de chacun des algorithmes appuie l'affirmation présentée ci-haut à l'effet que ces trois algorithmes, à 2000 évaluations du modèle hydrologique, présentent des efficacités similaires. Une attention particulière doit être portée aux échelles des axes des ordonnées à la figure 5.5 puisqu'elles sont différentes d'un bassin à l'autre (étant donné les plages de valeurs différentes prises par le critère de Nash-Sutcliffe). Cette différence doit être prise en compte si l'on compare visuellement les dispersions entre les deux bassins.

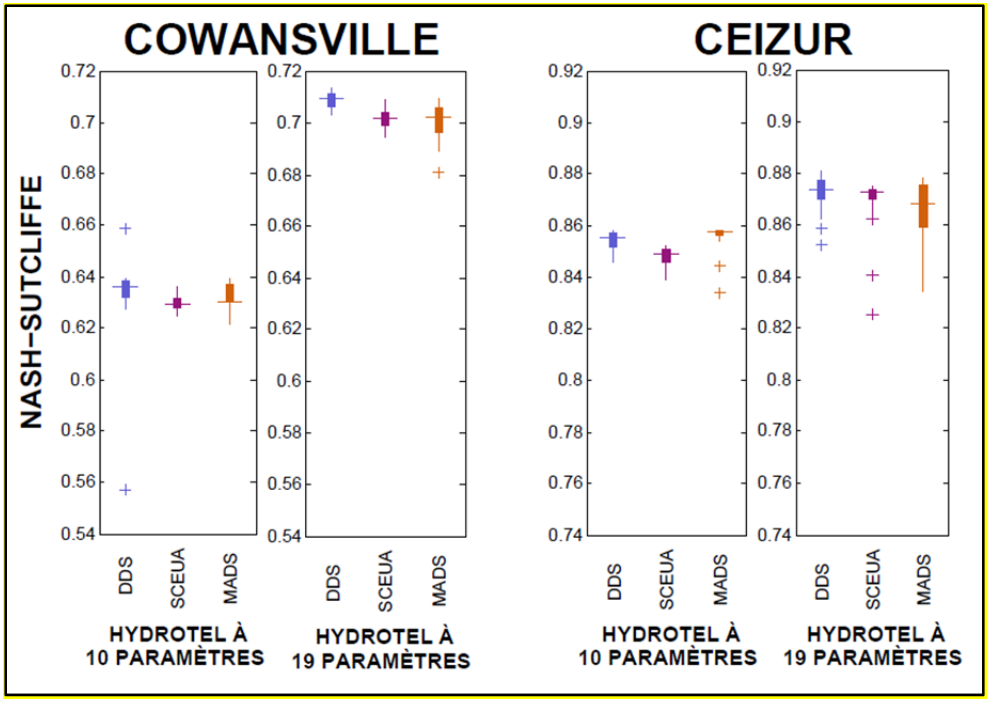

Figure 5.5 Dispersion du critère objectif du « Nash-Sutcliffe Efficiency» (NSE) à 2000 évaluations considérant les 32 calages effectués avec chacun des algorithmes SCEUA, DDS et MADS, sur les bassins versants Cowansville et Ceizur, pour les modèles hydrologiques HYDROTEL à 10 et à 19 paramètres

5.1.3 Étude de la progression de la fonction objectif selon le temps de calcul

Les figures 5.6 et 5.7 présentent les résultats des courbes de progression de la fonction objectif (un moins le critère de Nash-Sutcliffe) en fonction du nombre d'évaluations du modèle hydrologique pour toutes les combinaisons 'HYDROTEL-bassin versant'. Ces figures démontrent clairement que l'algorithme DDS explore l'espace des paramètres plus efficacement que SCEUA et MADS puisqu'il obtient beaucoup plus rapidement une valeur de fonction objectif de bonne qualité. La courbe moyenne de progression de DDS trace distinctement une courbe d'allure asymptotique où, en quelques évaluations du modèle (moins de 250 évaluations), l'algorithme a trouvé des zones d'optima de bonne qualité. Pour ce qui est de SCEUA et MADS, leurs descentes ne sont pas aussi prononcées. Mise à part la combinaison 'HYDROTEL à 19 paramètres-Cowansville', la courbe moyenne de SCEUA atteint tout de même une cessation de descente avant les 2000 évaluations du budget de calage. Pour MADS, une tel cessation de progression ne s'observe que pour 'HYDROTEL à

10 paramètres-Cowansville'. Cependant, une vérification a été faite quant aux processus de calage effectués à l'aide de MADS de manière à savoir si ceux-ci avaient été interrompus par le critère du nombre maximal d'évaluations du modèle hydrologiques (m) ou par l'atteinte du dimensionnement minimal du maillage ($\Delta_{k_{min}}$). Une investigation sur ces résultats a permis d'identifier que, dans environ 65 % des problèmes d'optimisation d'HYDROTEL à 10 paramètres et dans environ 45 % des problèmes d'HYDROTEL à 19 paramètres, les calibrations ont été interrompues avant l'atteinte du nombre maximal d'évaluations permis (m) puisque la dimension minimale du maillage ($\Delta_{k_{min}}$) était égalée. Cette apparence de comportement « en progression » de la fonction objectif à 2000 évaluations, sur les figures 5.6 et 5.7, peut être expliquée par le fait que les courbes présentent des comportements moyens de MADS. On en déduit également que le nombre d'évaluations permis est suffisant et adéquat pour que MADS puisse atteindre ou pratiquement atteindre un optimum local. Il est également pertinent d'observer le gain obtenu par les diverses méthodes d'optimisation dites rigoureuses par rapport à ACA (gain d'environ + 0,07 de la fonction objectif pour Cowansville et d'environ + 0,02 à + 0,05 pour Ceizur).

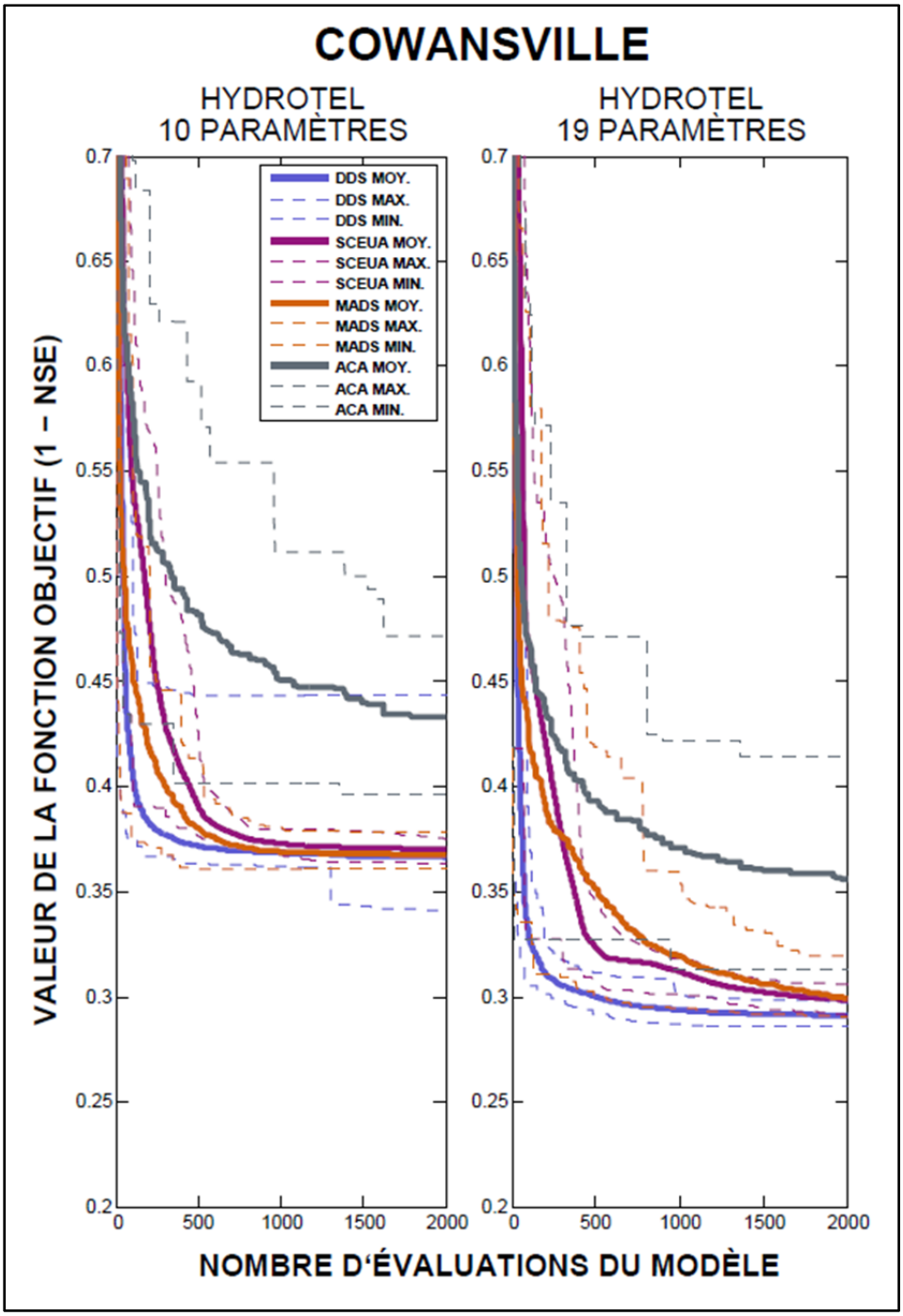

Figure 5.6 Progression de la fonction objectif 1- « Nash-Sutcliffe Efficiency » (1-NSE) en fonction du nombre d'évaluations du modèle hydrologique pour les algorithmes SCEUA, DDS, MADS et ACA sur le bassin versant Cowansville pour les modèles hydrologiques HYDROTEL à 10 et à 19 paramètres

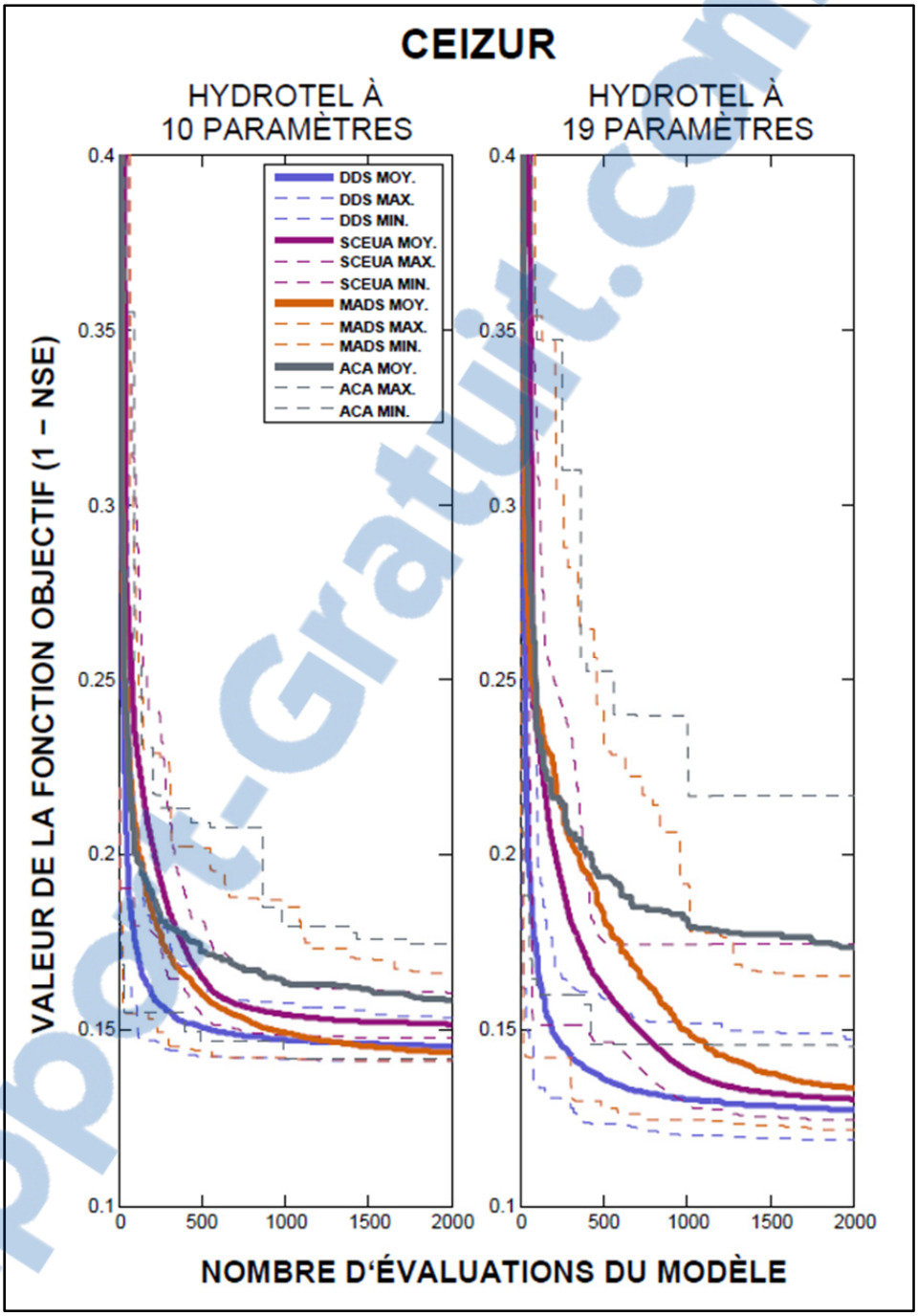

Figure 5.7 Progression de la fonction objectif 1- « Nash-Sutcliffe Efficiency » (1-NSE) en fonction du nombre d'évaluations du modèle hydrologique pour les algorithmes SCEUA, DDS, MADS et ACA sur le bassin versant Ceizur pour les modèles hydrologiques HYDROTEL à 10 et à 19 paramètres

Rapport-Gratuit.com

5.2 Études des possibilités de configuration de l'optimisation dans NOMAD

5.2.1 Profils de performance

La figure 5.8 illustre les profils de performance des différentes configurations du logiciel d'optimisation NOMAD étudiées pour le modèle hydrologique HYDROTEL à 10 paramètres sur le bassin versant Cowansville. Dans un premier temps, on dénote sur cette figure que les deux configurations proposées, soient MADS avec les groupes de variables (MADS VG) et MADS avec les directions orthogonales « N+1 » (MADS N+1), sont distinctement plus efficaces en début de calage (avant 1000 évaluations du modèle hydrologique) que la configuration par défaut (MADS). Toujours avant 1000 évaluations, MADS VG et MADS N+1 atteignent les probabilités de succès de 100 % à des valeurs similaires d'écart par rapport à v^*, mais MADS VG offre une plus haute probabilité de succès initiale. Dans un second temps, bien que les deux configurations explorées apportent une meilleure performance avec un petit budget de calage, il n'en demeure pas moins que MADS devance ces deux approches lorsque ce budget devient plus important (plus de 1000 évaluations du modèle).

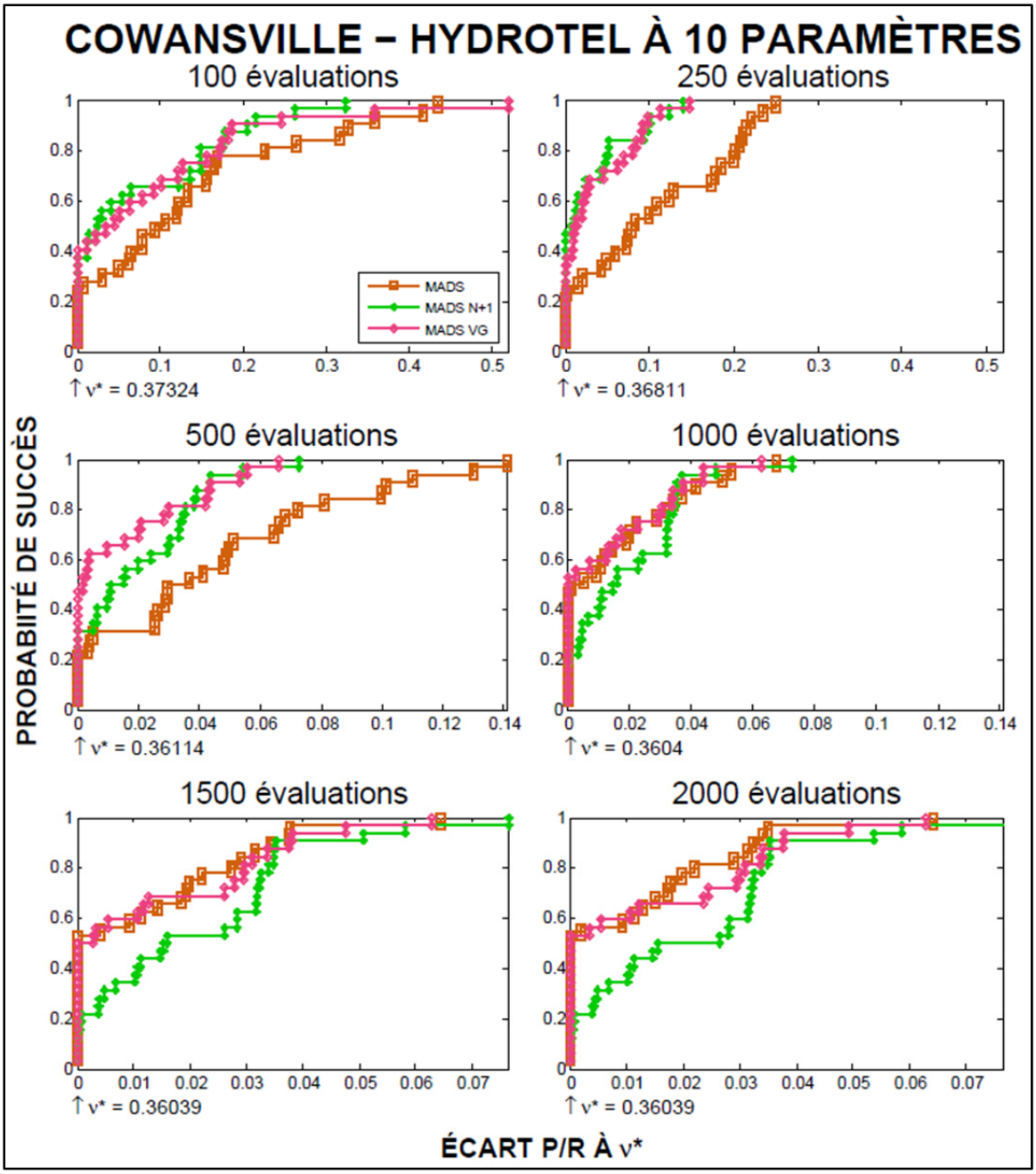

Figure 5.8 Profils de performance des algorithmes MADS, MADS N+1 et MADS VG pour le bassin versant Cowansville avec le modèle hydrologique HYDROTEL à 10 paramètres

Rapport-gratuit.com
LE NUMERO 1 MONDIAL DU MÉMOIRES

5.2.2 Analyse de dispersion des valeurs finales de la fonction objectif

La figure suivante (5.9) illustre les boîtes à moustaches des valeurs de la fonction objectif (critère de Nash-Sutcliffe) obtenues par MADS, MADS N+1 et MADS VG après 2000 évaluations du modèle hydrologique, pour les 32 calages effectués dans chacun des cas. Ces résultats démontrent clairement, encore une fois (même observation que section 5.1.2), que les diverses méthodes d'optimisations obtiennent toutes une efficacité équivalente en ce qui a trait à la valeur finale de la fonction objectif étant donné le minuscule écart entre les trois médianes et les similitudes entre les écarts interquartiles et les étendues des moustaches. Cependant, MADS N+1 semble offrir une très légère diminution d'efficacité par rapport à MADS et MADS VG (voir la discussion à la section 6.2).

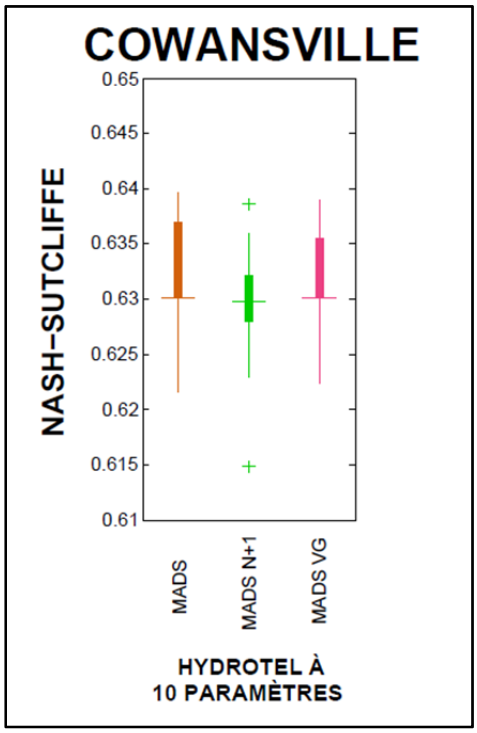

Figure 5.9 Dispersion du critère objectif du « Nash-Sutcliffe Efficiency» (NSE) à 2000 évaluations considérant les 32 calages effectués avec les algorithmes MADS, MADS N+1 et MADS VG sur le bassin versant Cowansville pour le modèle hydrologique HYDROTEL à 10 paramètres

5.2.3 Étude de la progression de la fonction objectif selon le temps de calcul

La figure 5.10 présente les résultats des courbes de progression de la fonction objectif (un moins le critère de Nash-Sutcliffe) en fonction du nombre d'évaluations du modèle hydrologique. De façon complémentaire à ce que l'on observe à la figure 5.8, la figure 5.10 démontre que les méthodes de calage MADS N+1 et MADS VG sont plus efficaces en termes d'exploration de l'espace des paramètres puisque leurs courbes moyennes de progression offrent une descente plus prononcée que la configuration de MADS par défaut. Ces deux méthodes de calage offrent donc un potentiel de réduction du temps de calcul plus importante que la version par défaut de MADS. Toutefois, après 750 évaluations, la courbe moyenne de progression de MADS se superpose à celles des deux autres approches.

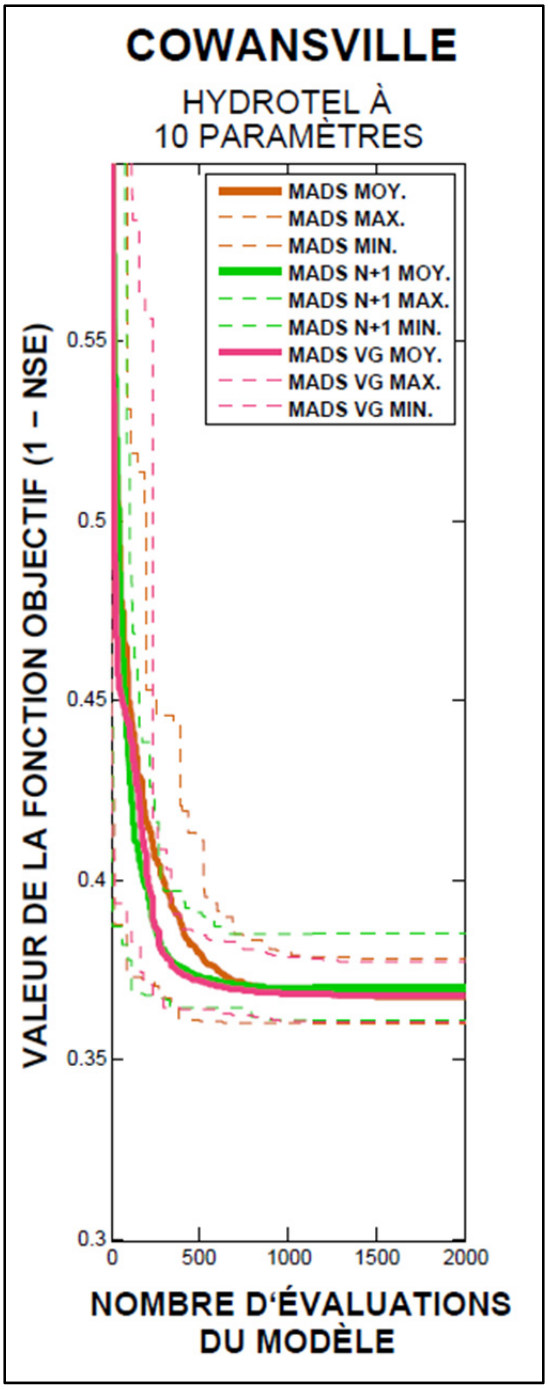

Figure 5.10 Progression de la fonction objectif 1- « Nash-Sutcliffe Efficiency » (1-NSE) en fonction du nombre d'évaluations du modèle hydrologique pour les algorithmes MADS, MADS N+1 et MADS VG sur le bassin versant Cowansville pour le modèle hydrologique HYDROTEL à 10 paramètres

5.3 Impact du type de modèle hydrologique sur le comportement des approches

5.3.1 Profils de performance

Les figures 5.11 et 5.12 présentent les profils de performance pour les bassins versants de Cowansville et de Ceizur modélisés à l'aide d'HSAMI à 23 paramètres. En comparant les profils de performance présentés ci-dessous avec ceux des figures 5.1 à 5.4, on constate une certaine similarité à l'égard du classement des différentes méthodes de calage selon leur efficacité. DDS domine constamment tous les autres algorithmes et l'algorithme aléatoire demeure d'une efficacité comparable à celles de SCEUA et MADS en début de calage (moins de 250 évaluations du modèle hydrologique), mais devient inefficace par la suite. Également, SCEUA et MADS se positionnent respectivement en deuxième et troisième places tout comme lors du calage d'HYDROTEL à 19 paramètres (figures 5.2 et 5.4). Tel qu'observé précédemment (section 5.1.1), la dimensionnalité plus grande du problème d'optimisation que représente le calage d'HSAMI à 23 paramètres semble poser un obstacle pour MADS qui serait ralenti dans son processus de calibration (voir la discussion à la section 6.3).

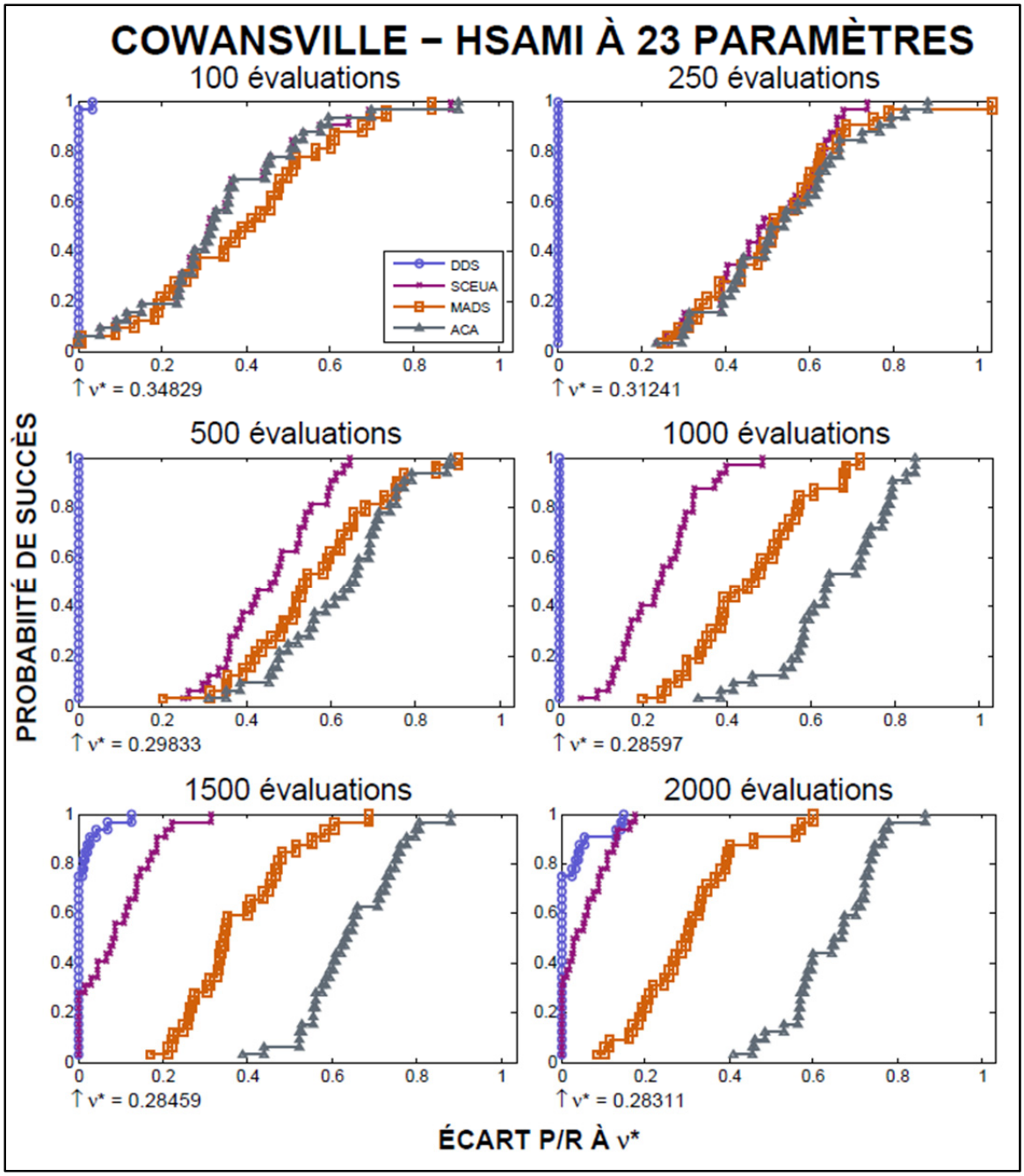

Figure 5.11 Profils de performance des algorithmes SCEUA, DDS, MADS et ACA pour le bassin versant Cowansville avec le modèle hydrologique HSAMI à 23 paramètres

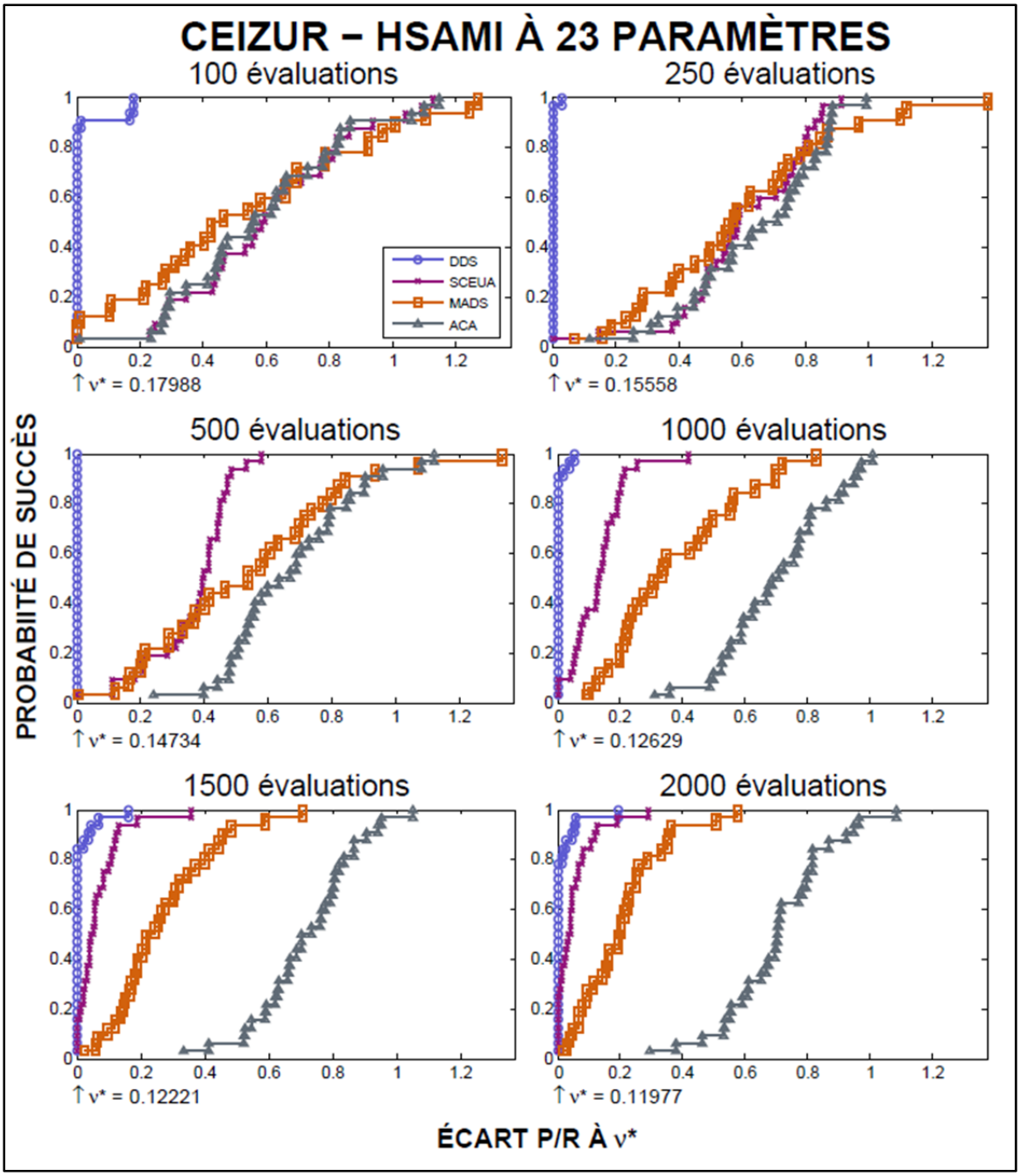

Figure 5.12 Profils de performance des algorithmes SCEUA, DDS, MADS et ACA pour le bassin versant Ceizur avec le modèle hydrologique HSAMI à 23 paramètres

5.3.2 Analyse de dispersion des valeurs finales de la fonction objectif

La figure 5.13 illustre la dispersion des valeurs de la fonction objectif (critère de Nash-Sutcliffe) obtenues par DDS, SCEUA et MADS après 2000 évaluations du modèle hydrologique HSAMI, pour les 32 processus de calage effectués sur chacun des bassins d'étude. Ces résultats démontrent une différence notable par rapport à ceux qui sont présentés à la figure 5.5 en ce qui concerne l'étendue des moustaches, les écarts interquartiles et la différence entre les médianes des trois algorithmes. L'analyse de la figure révèle une dispersion beaucoup plus grande que celles obtenues avec HYDROTEL à 10 et 19 paramètres (figure 5.5) et ce, pour les deux bassins versants (à noter que la dispersion de MADS est encore plus importante). De plus, l'écart entre les différentes médianes est plus important, passant de ± 0,01 dans l'analyse de dispersion d'HYDROTEL (10 et 19 paramètres) à ± 0,05 dans celle d'HSAMI. Le type de modèle et la dimensionnalité semblent donc avoir un impact sur l'efficacité des méthodes d'optimisation sélectionnées.

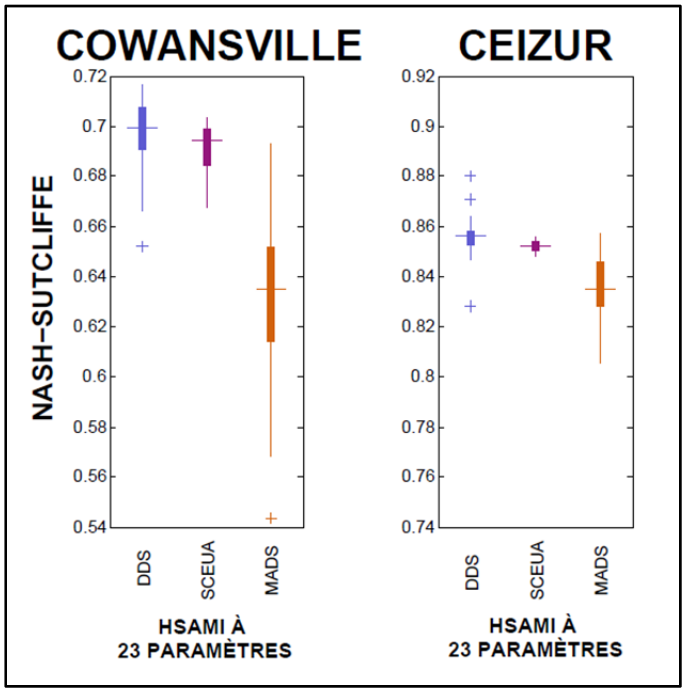

Figure 5.13 Dispersion du critère objectif de « Nash-Sutcliffe Efficiency» (NSE) à 2000 évaluations considérant les 32 calages effectués avec les algorithmes SCEUA, DDS et MADS sur les bassins versants Cowansville et Ceizur pour le modèle hydrologique HSAMI à 23 paramètres

5.3.3 Étude de la progression de la fonction objectif selon le temps de calcul

Les figures 5.14 et 5.15 reprennent les figures 5.6 et 5.7, auxquelles ont été ajoutés les résultats obtenus avec le modèle HSAMI. L'analyse de ces figures permet de soutenir les observations mentionnées à la section précédente (section 5.3.2) concernant l'impact du type de modèle sur le comportement ou l'efficacité des méthodes de calage. Tout d'abord, la performance de ACA indique qu'il est plus difficile d'obtenir un bon jeu de paramètres de façon aléatoire lors de l'emploi d'HSAMI comparativement à HYDROTEL. Le gain obtenu par les diverses méthodes d'optimisation par rapport à ACA est beaucoup plus important dans le cas du calage de HSAMI que dans celui d'HYDROTEL (gain d'environ + 0,15 de la fonction objectif pour Cowansville et d'environ + 0,09 pour Ceizur avec HSAMI comparativement à + 0,07 et + 0,02 à 0,05, respectivement, avec HYDROTEL). Ces observations appuient d'autant plus les affirmations précédentes (section 5.3.2) quant à la dissemblance remarquée au sein des problèmes d'optimisation à dimensionnalité et types de modèles différents (voir la discussion à la section 6.3).

Même si le classement des approches de calage reste inchangé avec l'utilisation de HSAMI, les courbes de progression de chacune des méthodes employées adoptent une descente encore moins prononcée que lors du calage d'HYDROTEL. En fait, on observe que, pour le calage de HSAMI, les algorithmes SCEUA, MADS et ACA semblent ne pas être arrivés au terme de leur processus de calage puisque les courbes de progression montrent une descente qui semble vouloir continuer même à 2000 évaluations du modèle. Il est à noter qu'aucune calibration à l'aide de MADS n'a été interrompue avant les 2000 évaluations par cause d'atteinte du critère d'arrêt du dimensionnement minimal du maillage ($\Delta_{k_{min}}$), contrairement à ce qui avait été observé avec HYDROTEL (section 5.1.3). Cela laisse croire que le processus de calage pour le modèle hydrologique HSAMI nécessite plus de 2000 évaluations pour que les méthodes d'optimisation puissent atteindre les zones d'optima, et ce faisant, fournir un jeu de paramètres de qualité. Quant à l'algorithme DDS, le comportement asymptotique démontré lors du calage d'HYDROTEL est sans contredit moins dominant étant donné le ralentissement observé dans la progression de la fonction objectif. Cela

Rapport-Gratuit.com

corrobore également les précédentes observations à l'effet qu'il existe des différences entre les deux problèmes d'optimisation représentés par le calage des modèles HYDROTEL et HSAMI (voir la discussion à la section 6.3).

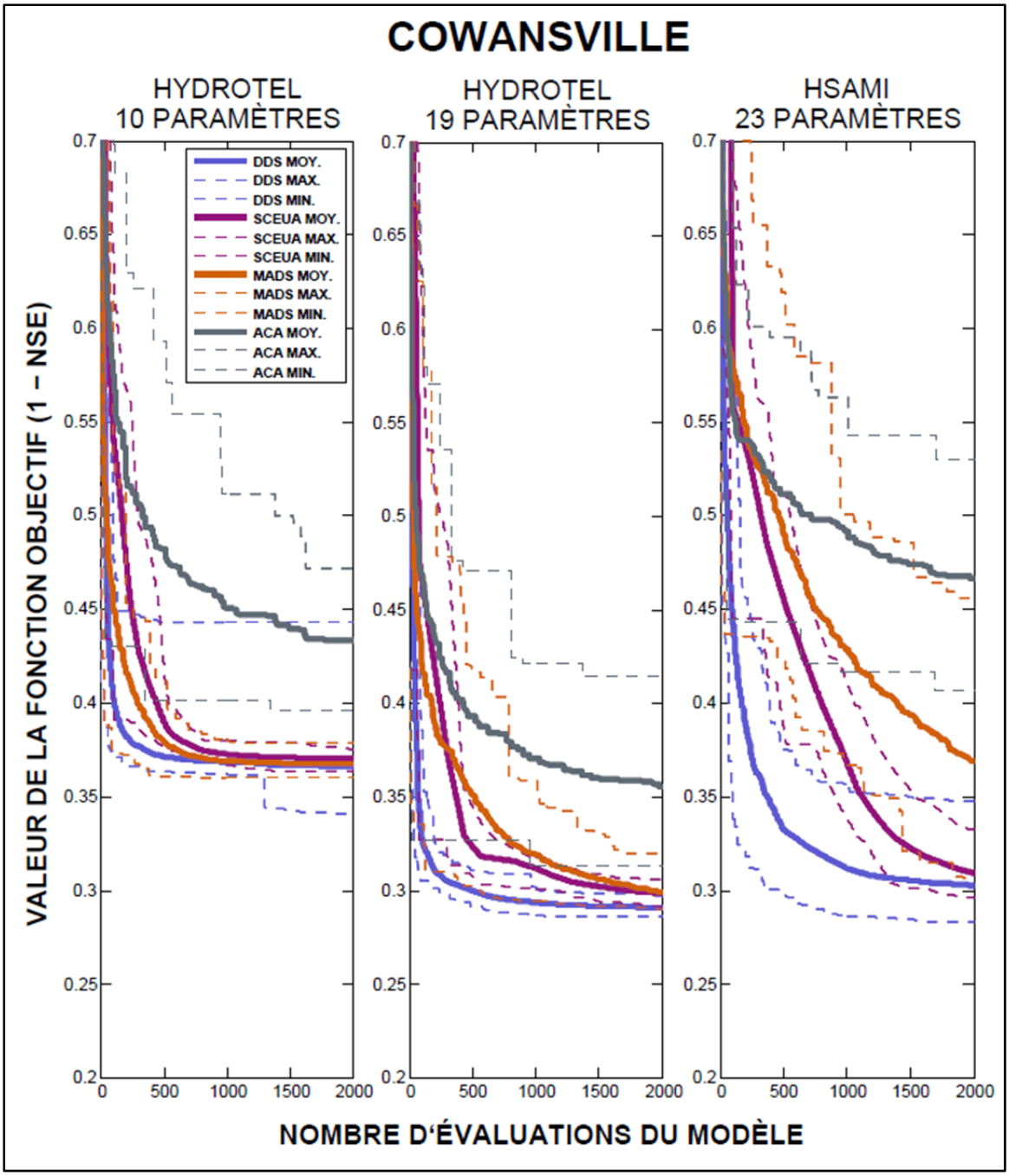

Figure 5.14 Progression de la fonction objectif 1- « Nash-Sutcliffe Efficiency » (1-NSE) en fonction du nombre d'évaluations du modèle hydrologique pour les algorithmes SCEUA, DDS, MADS et ACA sur le bassin versant Cowansville pour les modèles hydrologiques HYDROTEL à 10 et à 19 paramètres, et HSAMI à 23 paramètres

Rapport-gratuit.com
LE NUMERO 1 MONDIAL DU MÉMOIRES

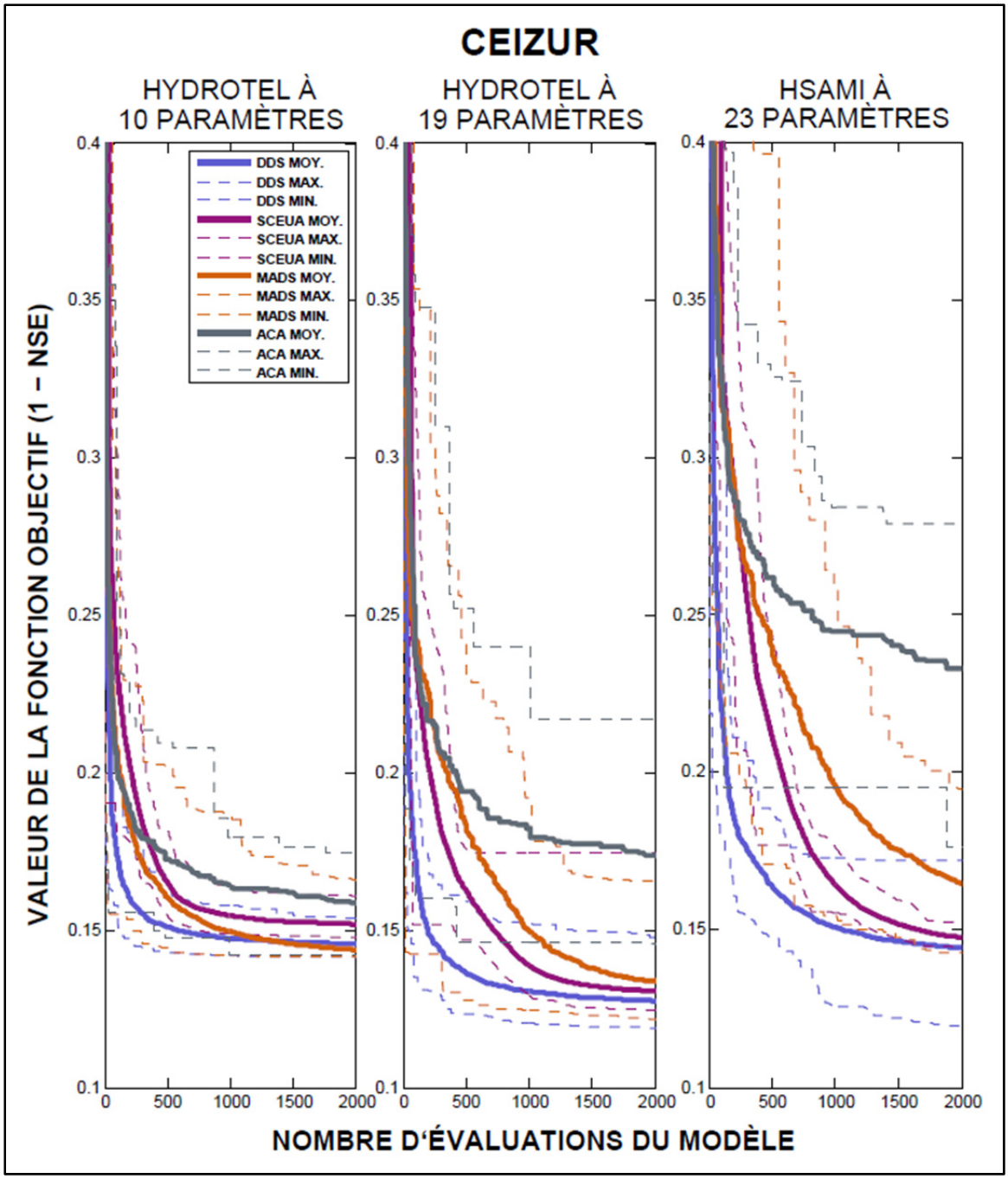

Figure 5.15 Progression de la fonction objectif 1- « Nash-Sutcliffe Efficiency » (1-NSE) en fonction du nombre d'évaluations du modèle hydrologique pour les algorithmes SCEUA, DDS, MADS et ACA sur le bassin versant Ceizur pour les modèles hydrologiques HYDROTEL à 10 et à 19 paramètres, et HSAMI à 23 paramètres

CHAPITRE 6

Discussion

Ce chapitre présente une discussion des résultats obtenus ainsi que les interprétations importantes que les expérimentations ont mises de l'avant. Les limitations et l'estimé des retombées de la recherche y sont également exposés.

6.1 Stratégies employées par les méthodes d'optimisation

Les résultats présentés à la figure 5.5 ont illustré des éléments intéressants quant aux valeurs finales de la fonction objectif obtenues par les trois approches de calage sélectionnées après avoir atteint 2000 évaluations du modèle hydrologique. Comme il a été expliqué précédemment (section 2.4), les stratégies de recherche employées par MADS apportent un processus de raffinement qui fournit un jeu de paramètres final satisfaisant les conditions d'optimalité relativement aux optima locaux (pour des raisons de concision, à partir de ce point, l'expression « preuve de convergence » sera souvent employée en remplacement de « satisfaction des conditions d'optimalité »). Au moment de lancer cette étude, il était ainsi anticipé que MADS, au terme des processus de calage, fournisse des jeux de paramètres dont la valeur de la fonction objectif surpasse celle des jeux de paramètres finaux trouvés par DDS et SCEUA puisqu'il est doté de stratégies de raffinement robustes (méthode semi-exacte). Les résultats de cette recherche démontrent toutefois que DDS offre des points finaux de qualité supérieure, bien que le fonctionnement de DDS décrit plus tôt (section 2.3) ne semble prévoir aucun procédé ou stratégie particulière démontrant une robustesse par rapport à la convergence vers des optima.

Une vérification de cette hypothèse a donc été faite en effectuant des processus de calage à l'aide de MADS à partir des jeux de paramètres finaux trouvés par DDS. De cette façon, il est possible d'identifier si les points finaux trouvés par DDS correspondent à des optima locaux (satisfaction des conditions d'optimalité) tout comme ceux trouvés par MADS (si aucune amélioration n'est remarquée). La figure 6.1 illustre les résultats de cette

expérimentation où les 32 jeux de paramètres finaux de DDS ont été calés avec MADS pour le modèle hydrologique HYDROTEL à 10 paramètres sur les bassins Cowansville et Ceizur. Il est à noter que, bien qu'il ait été attribué un budget de recalage de 2000 évaluations pour MADS, chacun des processus de calage a été prématurément interrompu pour cause de satisfaction du critère d'arrêt de la dimension minimale du maillage ($\Delta_{k_{min}}$). L'interruption du processus de calage par l'atteinte de ce critère permet d'affirmer que les points finaux recalés à l'aide de MADS satisfont les conditions d'optimalité.

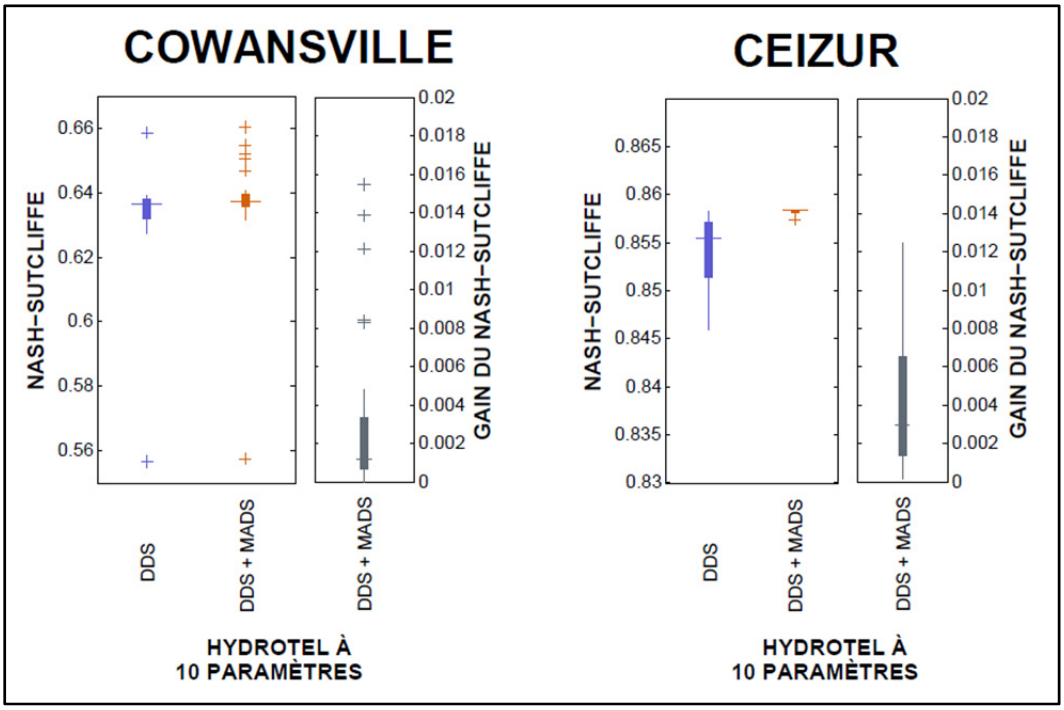

Figure 6.1 Test de convergence de l'algorithme DDS. Critère objectif du « Nash-Sutcliffe Efficiency» (NSE) à 2000 évaluations avec DDS, critère du NSE à 4000 évaluations avec DDS et MADS (jeux de paramètres finaux de DDS recalés à l'aide de MADS) et gain sur le NSE suite au test de convergence sur les bassins versants Cowansville et Ceizur pour le modèle hydrologique HYDROTEL à 10 paramètres

La figure 6.1 démontre que chacun des jeux de paramètres finaux trouvés par DDS a obtenu un gain supplémentaire de la valeur de la fonction objectif suite aux recalages effectués par MADS. Si petit soit le gain obtenu, de cette expérimentation, il est possible d'affirmer que l'hypothèse soulevée est juste; DDS n'est pas en mesure de fournir un jeu de paramètres final

satisfaisant les conditions d'optimalité avec un budget de 2000 évaluations contrairement à MADS où, comme il a été mentionné à la section 5.1.3, la plupart des calibrations effectuées ont été prématurément interrompues par le critère d'arrêt du dimensionnement minimal du maillage ($\Delta_{k_{min}}$) (65 % pour les problèmes d'optimisation d'HYDROTEL à 10 paramètres et 45 % pour HYDROTEL à 19 paramètres).

Comme il a été démontré au sein des résultats de cette étude, DDS est sans-contredit un algorithme qui excelle en recherche globale (exploration de l'espace des paramètres) puisqu'il peut identifier les zones d'optima de bonne qualité en un nombre limité d'évaluations du modèle hydrologique. Son comportement adaptatif vient également accentuer sa performance en recherche globale par rapport aux autres méthodes d'optimisation sous budget de calage restreint. Les stratégies globales sur lesquels DDS s'appuie font de lui un algorithme efficace qui cible de très bons jeux de paramètres, et ce, même s'il ne possède pas de stratégies de recherche équivalentes à celles associées à la « preuve de convergence ».

De plus, il a été expliqué précédemment que MADS est une méthode de descente directe basée sur la progression du meilleur jeu de paramètres trouvé en empruntant des directions de descente dans l'espace paramétrique. Les résultats ont illustré que MADS souffre effectivement d'un retard quant à la recherche globale justifiant ainsi l'emploi du « Latin Hypercube Search » qui a été préalablement configuré au sein de MADS pour diminuer la victimisation due, d'une part, à la qualité du jeu de paramètres fourni comme point de départ et d'autre part, au processus de recherche qui oblige l'algorithme à converger vers un optimum local, ralentissant ainsi la progression de la valeur de la fonction objectif. Cependant, bien que le « Latin Hypercube Search » contribue à l'exploration de l'espace de recherche, il s'avère que la configuration utilisée demeure insatisfaisante comparativement aux stratégies de recherche globale employées par DDS. Autrement dit, MADS ne parvient malheureusement pas à éclipser totalement la victimisation causant le ralentissement en recherche globale malgré le support que le « Latin Hypercube Search » peut apporter.

Il est clair que le principe de raffinement du réseau de maillages autour du meilleur point trouvé est une méthode de convergence robuste qui apporte à un optimiseur une qualité intéressante lorsqu'utilisé pour le calage d'un modèle hydrologique. Toutefois, la « preuve de convergence » au sein des stratégies de recherche de l'algorithme n'est pas annonciatrice d'une grande efficacité en temps de calcul (comme le démontrent les résultats de cette étude), ni de la bonne qualité de la solution trouvée (la convergence pouvant se faire sur un optimum médiocre) si elle n'est pas accompagnée d'une stratégie de recherche globale efficace. Certaines caractéristiques particulières de l'espace des paramètres des problèmes de calage des modèles hydrologiques semblent grandement avantager l'emploi de stratégies axées sur la recherche globale comme celles employées par DDS. Il serait certainement pertinent et intéressant d'approfondir la recherche scientifique quant à la caractérisation de l'espace de recherche de manière à mieux comprendre les particularités du problème d'optimisation que représente le calage des modèles hydrologiques. Plusieurs métriques de caractérisation de l'espace de recherche peuvent être utilisées et adaptées au besoin dans le but d'apporter des informations sur la topologie générale de la fonction objectif, le bruit, l'interdépendance entre les différents paramètres, etc. (Abell et al., 2012; Malan et Engelbrecht, 2013).

6.2 Discussion sur les configurations de NOMAD

Certaines configurations de NOMAD ont été explorées au sein de cette étude et ont permis d'identifier un potentiel de réduction en temps de calcul en utilisant des fonctionnalités offertes par le logiciel. Les deux configurations exploitées, soit le calage par groupes de variables (MADS VG) et la version récemment développée de l'algorithme (MADS N+1), sont toutes deux arrivées à fournir une meilleure exploration globale par rapport à la version par défaut de MADS, et ainsi permettre un potentiel de réduction du temps de calcul. Toutefois, ces deux configurations peuvent perdre de leur intérêt du point de vue d'un utilisateur pour qui l'efficacité du calage de son modèle hydrologique est orientée plutôt vers la rapidité d'exécution uniquement, ou encore vers la seule qualité du jeu de paramètres final contrairement à celle décrite dans cette étude (meilleur compromis entre la réduction du temps de calage et la performance accrue en termes de valeur de la fonction objectif finale).

Malgré le potentiel de réduction en temps de calcul montré, les résultats des calibration de la version de MADS à génération de motif « N+1 » ont démontré une légère diminution d'efficacité en terme de la valeur de la fonction objectif finale par rapport à la version de MADS par défaut. Cette constatation au sein des processus de calage effectués avec MADS N+1 pourrait dissuader l'utilisateur possédant un grand budget de calage d'utiliser cette configuration puisqu'il identifie un risque d'obtenir un jeu de paramètres de moins bonne qualité qu'avec la version par défaut de MADS. Cette situation observée dans le comportement de MADS N+1 est simple à expliquer : la configuration du motif selon les directions orthogonales « N+1 » diminue de presque la moitié les directions orthogonales visitées par rapport à la version par défaut de MADS (directions orthogonales « 2N »). Lorsque MADS N+1 entre en processus de raffinement du treillis autour du meilleur jeu de paramètres trouvé, il obtient alors une probabilité moindre d'identifier des orientations de descente qui sont à contre-courant des directions enregistrées précédemment étant donné qu'un seul point résultant des directions orthogonales non visités est généré. Il raffine alors prématurément la taille du motif de points alors qu'il omet de visiter quelques directions orthogonales qui pourraient diriger la descente de l'algorithme vers un autre secteur de l'espace des paramètres. De ce fait, la version MADS N+1 ne semble pas appropriée pour la problématique de calage des modèles hydrologiques advenant que la qualité de la valeur de la fonction objectif finale trouvée compte pour beaucoup auprès de l'évaluation de l'efficacité du calage.

Contrairement à MADS N+1, le calage par groupes de variables (MADS VG) ne semble pas être affecté en termes de qualité de la valeur de la fonction objectif finale lors du processus de raffinement tout en démontrant des gains potentiels de réduction du temps de calcul. Cependant, la constitution des groupes de variables a nécessité certaines connaissances pour que cette configuration de l'algorithme soit efficace : les liens existant entre les différents paramètres du modèle hydrologique et leur influence sur le régime hydrique. Cette compréhension plus approfondie du modèle hydrologique a permis la constitution de groupes de variables qui ont favorisé le processus de calage de MADS dans le contexte de la

problématique étudiée, mais l'utilisation des groupes de variables n'est pas un gage de succès à tout coup. Cette méthode de calage reste spécifique à chaque structure de problème d'optimisation (en l'occurrence, le modèle hydrologique utilisé). Des combinaisons irréfléchies des paramètres pourraient certainement diminuer le potentiel de cette configuration à réduire le temps de calcul et pourraient peut-être même devenir contraignantes pour le processus de calage en ralentissant la progression de la fonction objectif.

Les observations et conclusions mises de l'avant quant à l'étude des possibilités de configurations de NOMAD apportent une certaine limitation qu'il ne faut surtout pas négliger. Les temps de calculs importants des expérimentations ont limité la possibilité d'effectuer l'étude sur plusieurs combinaisons 'modèle hydrologique-bassin versant'. L'utilisation d'un seul banc d'essai (combinaison) ne permet pas de généraliser les conclusions, cependant les résultats et analyses exposés demeurent tout de même des avenues qu'il serait intéressant d'investiguer davantage de manière à tirer des conclusions mieux fondées et généralisées.

6.3 Discussion sur le type de modèle hydrologique

Les prochaines sections discutent de la complexité des problèmes de calages et des principales difficultés auxquelles les algorithmes d'optimisation doivent faire face : l'interdépendance entre les paramètres, la dimensionnalité du problème et de l'équifinalité.

6.3.1 Complexité des problèmes de calage

Les résultats présentés à la section 5.3 quant à l'impact du type de modèle hydrologique sur le comportement des algorithmes de calage ont révélé qu'il existait une différence certaine entre les problèmes de calage des modèles hydrologiques HYDROTEL et HSAMI. Cette différence est remarquée par l'identification de comportements distincts chez les mêmes algorithmes d'optimisation lorsqu'employés pour le calage des deux types de modèles (distribué et à base physique, vs global et conceptuel, respectivement). En premier lieu,

rappelons qu'il a été observé que les courbes de progression moyenne de la fonction objectif pour les diverses approches de calage pour le problème d'optimisation d'HSAMI semblent interrompues prématurément contrairement à ce que les figures montrent lors du calage d'HYDROTEL (voir les figures 5.14 et 5.15, à la section 5.3.3). De manière à vérifier si les processus de descente d'HSAMI sont bels et bien inachevés à 2000 évaluations, le calage de ce modèle a été poursuivi jusqu'à un budget de 10 000 évaluations. Les figures 6.2 et 6.3 présentent les portions des figures 5.14 et 5.15 se rapportant à HSAMI ainsi que les résultats obtenus lors de la poursuite des processus de calage pour les algorithmes SCEUA, MADS et ACA. En ce qui a trait à DDS, étant donné qu'il s'adapte au nombre d'évaluations du modèle fourni, les calages ont été relancés à partir des mêmes points initiaux, mais chaque fois avec un budget de calage de 10 000 évaluations.

Rapport-Gratuit.com

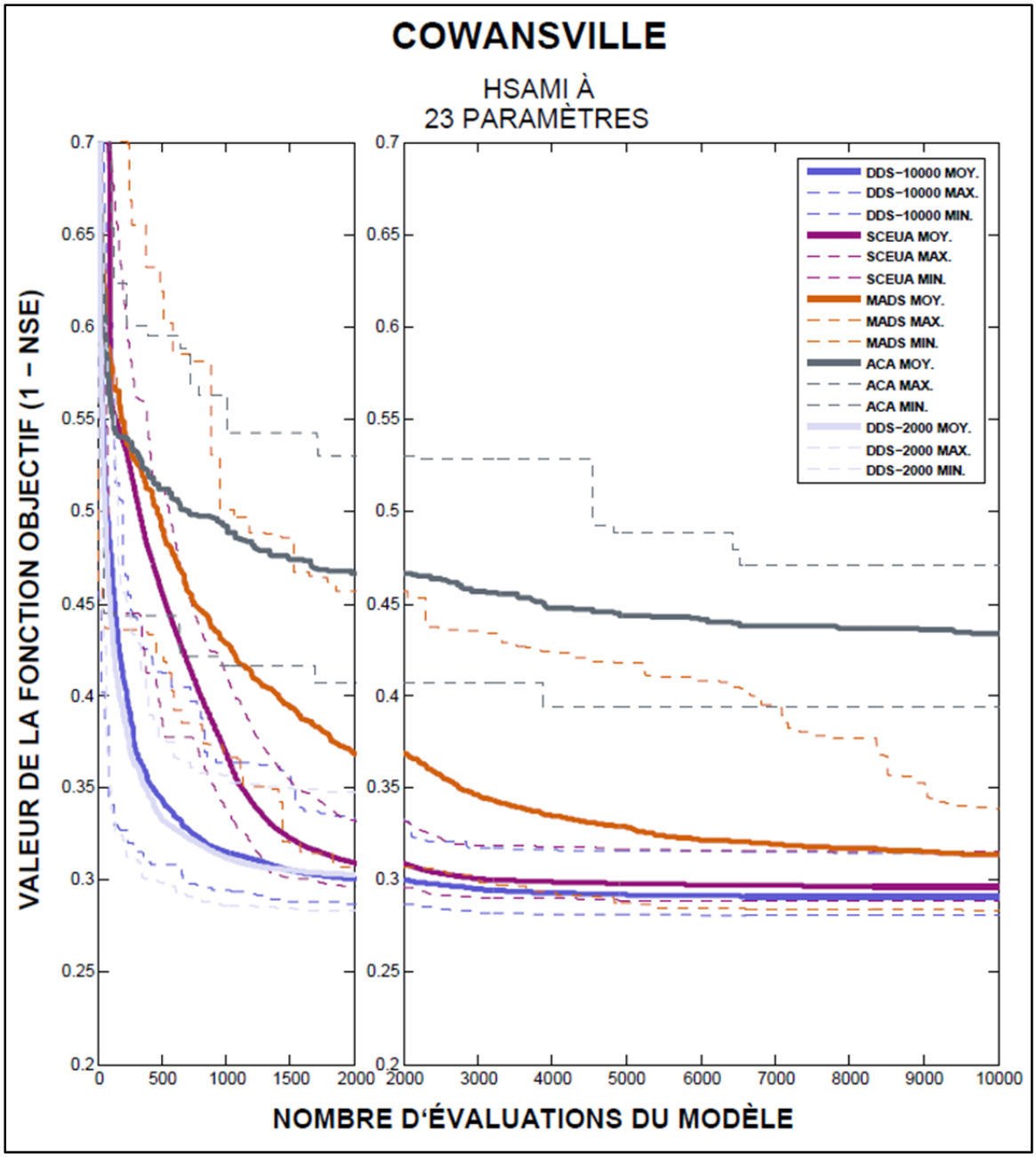

Figure 6.2 Progression de la fonction objectif 1- « Nash-Sutcliffe Efficiency » (1-NSE) sur 10 000 évaluations du modèle HSAMI à 23 paramètres, pour chacun des algorithmes SCEUA, DDS-2000, DDS-10000, MADS et ACA sur le bassin versant Cowansville

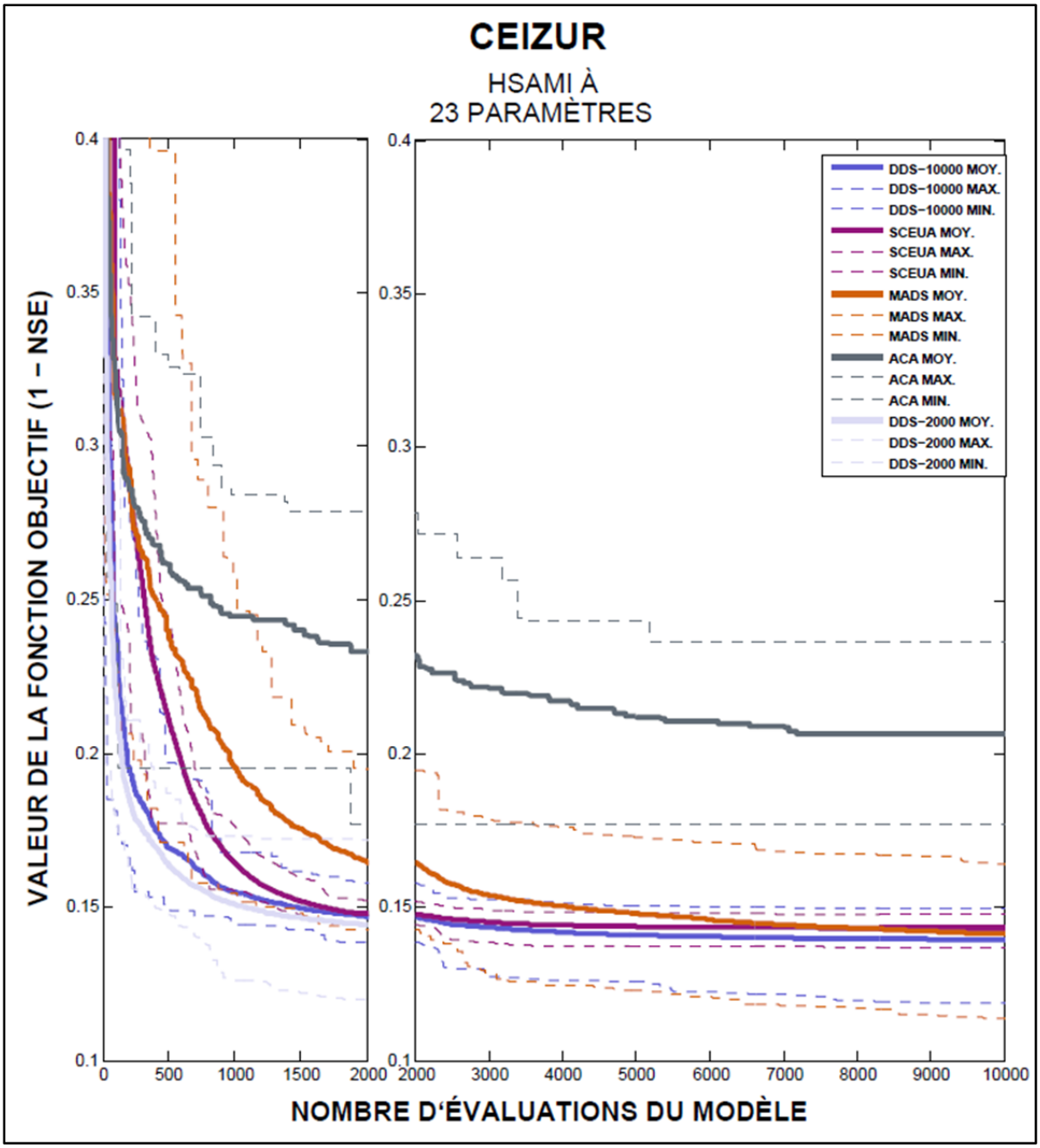

Figure 6.3 Progression de la fonction objectif 1- « Nash-Sutcliffe Efficiency » (1-NSE) sur 10 000 évaluations du modèle HSAMI à 23 paramètres, pour chacun des algorithmes SCEUA, DDS-2000, DDS-10000, MADS et ACA sur le bassin versant Ceizur

Les figures 6.2 et 6.3 démontrent clairement que, pour les algorithmes MADS et ACA, la valeur de la fonction objectif continue de décroître au-delà de 2000 évaluations du modèle hydrologique, voir même jusqu'à 10 000 évaluations. Pour ce qui est de SCEUA, la valeur de

Rapport-gratuit.com
LE NUMERO 1 MONDIAL DU MÉMOIRES

la fonction objectif continue de décroître après les 2000 évaluations, mais le gain occasionné par la poursuite du processus de calage n'est pas aussi perceptible que pour MADS et ACA. Il est également important de mentionner qu'uniquement 10 % des calibrations effectuées à l'aide de MADS ont été interrompues avant l'atteinte des 10 000 évaluations permises contrairement à ce qui avait été observé lors des calages du modèle HYDROTEL (section 5.1.3) où, rappelons-le, MADS avait atteint le critère d'arrêt du dimensionnement minimale du maillage ($\Delta_{k_{min}}$) avant 2000 évaluations dans plus de la moitié des calibrations. On peut alors comprendre que cette méthode semi-exacte éprouve quelques difficultés à satisfaire aux conditions d'optimalité du jeu de paramètres final lors du calage d'HSAMI. Ceci pointe vers une différence quant au niveau de complexité qui existe au sein du problème de calage d'HSAMI par rapport à HYDROTEL (topologie plus accidentée de l'espace paramétrique). Ceci laisse entrevoir que la complexité est plus importante au sein d'un modèle comme HSAMI où le nombre de paramètres est plus important et/ou le type de modèle (modélisation et spatialisation des processus hydrologiques) est à base conceptuelle et global.

En second lieu, il est possible d'observer que, malgré un budget de calage de 10 000 évaluations, le gain obtenu par les diverses approches de calage par rapport à ACA demeure important (gain d'environ + 0,14 de la fonction objectif pour Cowansville et d'environ + 0,07 pour Ceizur pour un budget de 10 000 évaluations comparativement aux gains de + 0,15 et + 0,09 obtenus à 2000 évaluations pour les mêmes bassins). Ceci corrobore les propos liés à l'impact du type de modèle sur le comportement des algorithmes d'optimisation, mais surtout sur la variation des niveaux de complexité au sein des problèmes de calage occasionnés par les diverses manières dont la paramétrisation des modèles hydrologiques sont effectuées et par la dimensionnalité du modèle. Comme le gain acquis par les approches de calage intelligentes est plus important avec HSAMI, le comportement de la recherche aléatoire tend à admettre que le niveau de complexité du problème d'optimisation d'HSAMI est plus élevé que celui d'HYDROTEL (difficulté plus importante à trouver les zones d'optima).

En dernier lieu, les dispersions des valeurs de la fonction objectif à 2000 évaluations pour les divers algorithmes (figure 5.13 de la section 5.3.2) ont été identifiées comme étant beaucoup

plus étendues que celles obtenus dans le cadre du calage d'HYDROTEL. À titre de comparaison, la figure 6.4 présente les dispersions des valeurs finales de la fonction objectif pour les différents algorithmes à 2000 et 10 000 évaluations pour le calage d'HSAMI sur les bassins versants Cowansville et Ceizur.

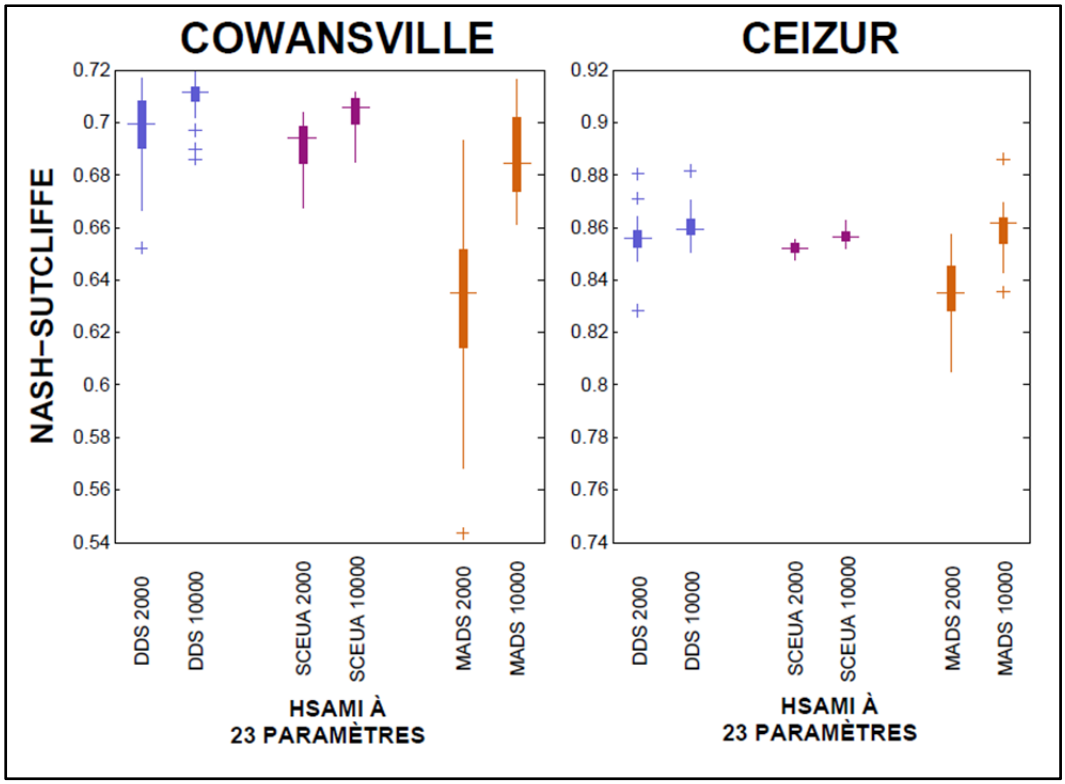

Figure 6.4 Dispersion du critère objectif de « Nash-Sutcliffe Efficiency» (NSE) à 2000 évaluations et à 10 000 évaluations considérant les 32 calages effectués avec les algorithmes SCEUA, DDS et MADS sur les bassins versants Cowansville et Ceizur pour le modèle hydrologique HSAMI à 23 paramètres

Cette dernière figure montre que les dispersions des valeurs finales de la fonction objectif des jeux de paramètres à 10 000 évaluations sont, de façon générale, plus faibles que celles obtenues à 2000 évaluations. De plus, il est à noter que les boîtes à moustaches de la fonction objectif à 10 000 évaluations ressemblent davantage à celles obtenues lors des calages d'HYDROTEL (voir la figure 5.5 à la section 5.1.2) en termes de dispersion des valeurs finales. La faible dispersion des boîtes à moustaches peut être vue comme une forme d'indication de la fin des processus de calages puisque les jeux de paramètres finaux sont

tous positionnés à l'intérieur des zones d'optima de qualité quasi-équivalente en termes de valeurs de la fonction objectif, mais dispersées en divers endroits de l'espace paramétrique. Comme l'apparition des faibles dispersions arrive plutôt vers les 10 000 évaluations du modèle, ces observations laissent croire, encore une fois, que la complexité du problème de calage d'HSAMI est supérieure à celle du problème de calage d'HYDROTEL.

6.3.2 Difficultés rencontrées par les méthodes d'optimisation : dimensionnalité, interdépendance et équifinalité

Les résultats et les analyses ont démontré que le type de modèle utilisé pour la modélisation des bassins versants a un impact non négligeable sur le comportement des méthodes d'optimisation. Chaque problème d'optimisation, ici défini par le problème de calage du modèle hydrologique, comporte des particularités et caractéristiques propres quant à la surface de l'espace paramétrique qui apportent un niveau spécifique de complexité. Dans le cadre de cette étude, la variation du niveau de complexité entre les différents problèmes d'optimisation (HYDROTEL à 10 et 19 paramètres et HSAMI à 23 paramètres) est certainement liée à la dimensionnalité du problème d'optimisation (nombre de paramètres). Avec un nombre important de paramètres de calage viennent la surparamétrisation des modèles hydrologiques, l'existence de liens d'interdépendance entre les paramètres et l'équifinalité, ce qui est d'autant plus vrai que les modèles sont généralement calés sur une seule variable hydrologique (les débits à l'exutoire d'un bassin). L'interdépendance est issue de la façon dont la paramétrisation est effectuée au sein de la structure interne du modèle hydrologique et la dimensionnalité provient quant à elle du nombre de paramètres à caler. Ces deux éléments contribuent à l'augmentation du niveau de complexité et par le fait même, à l'équifinalité.

Comme il a été mentionné dans la revue de littérature (section 1.1.2), l'équifinalité est généralement bien présente au sein des problèmes de calage des modèles hydrologiques et cette caractéristique peut renseigner sur la complexité du problème d'optimisation. Une analyse des jeux de paramètres finaux trouvés par les approches SCEUA, DDS et MADS a été réalisée et est présentée aux figures 6.5 et 6.6 pour illustrer la répartition de chacun des

points dans l'espace paramétrique. Les valeurs finales de chacun des paramètres ont été normalisées sur une échelle allant de zéro à un. D'un point de vue hydrologique, il a été jugé que tous les jeux de paramètres finaux, incluant les valeurs aberrantes, ont une valeur de fonction objectif équivalente étant donné la faible dispersion des valeurs finales (se référer aux figures 5.5 et 6.4). Ceci permet tout de même de donner un aperçu de l'équifinialité présente dans chacune des combinaisons 'modèle hydrologique-bassin versant'.

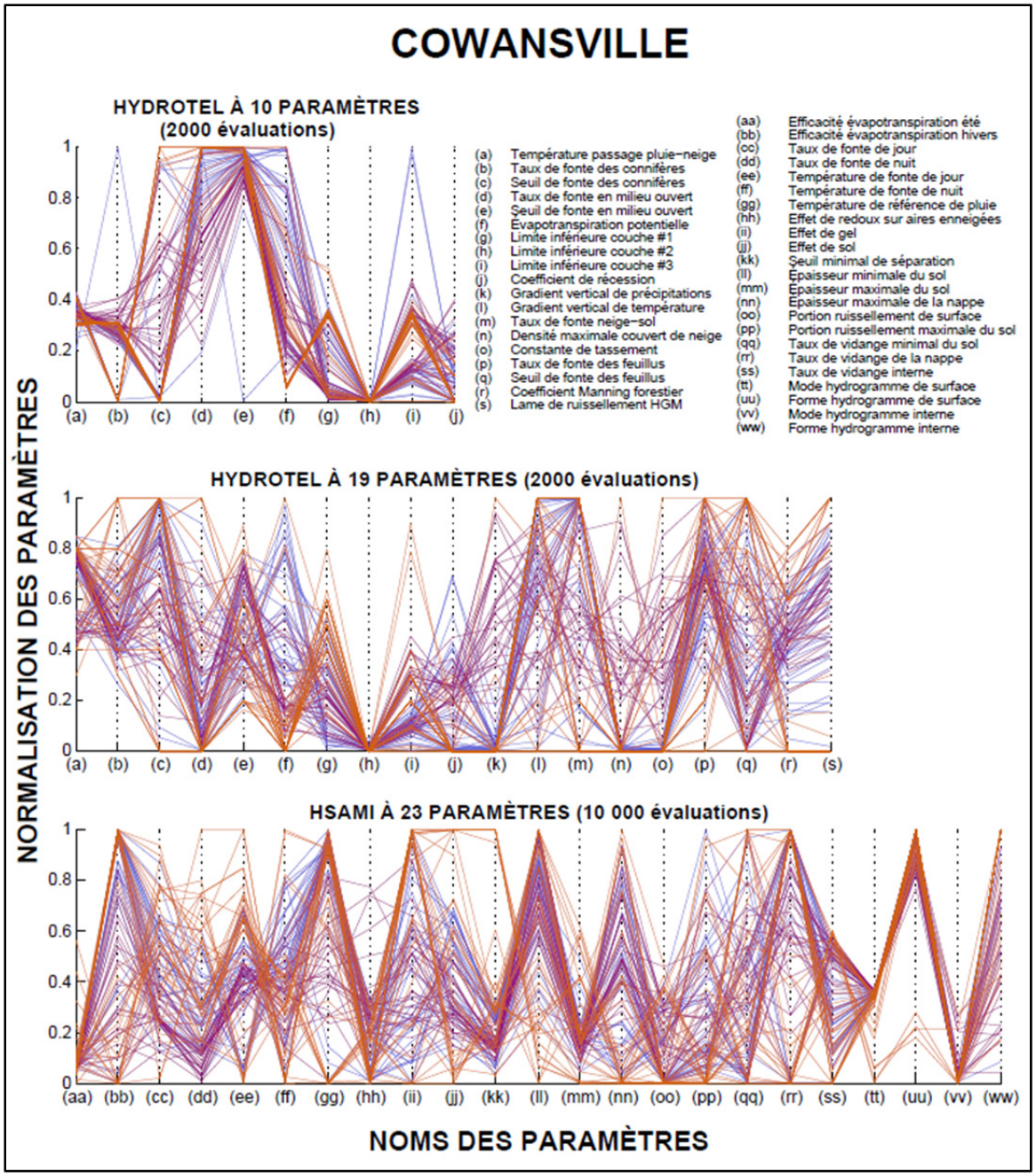

Figure 6.5 Étude de l'équifinalité des jeux de paramètres finaux issus des 96 calages à l'aide de DDS (courbes bleues), SCEUA (courbes pourpres) et MADS (courbes orangées) sur le bassin versant Cowansville avec les modèles hydrologiques HYDROTEL à 10 et 19 paramètres, et HSAMI à 23 paramètres

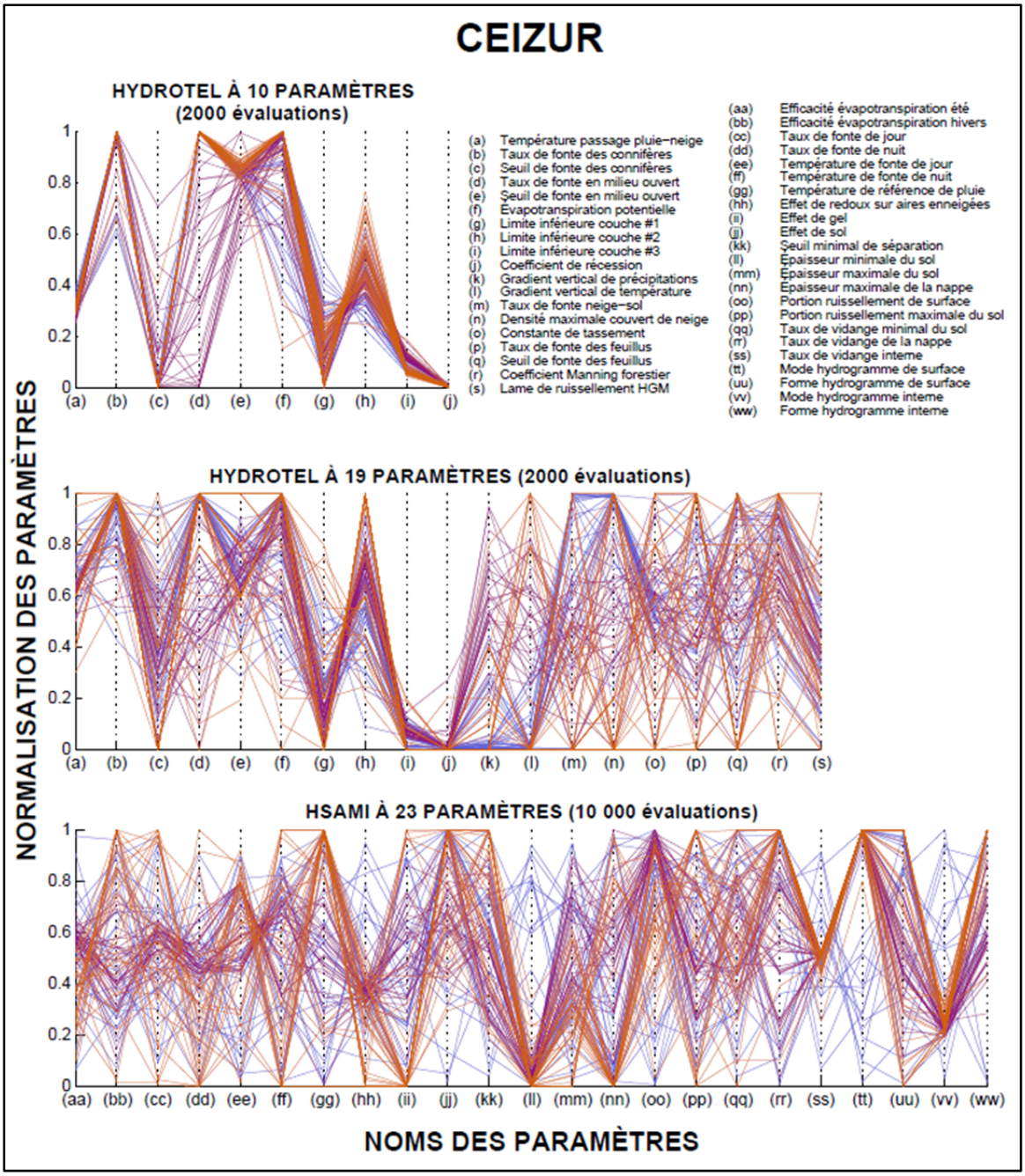

Figure 6.6 Étude de l'équifinalité des jeux de paramètres finaux issus des 96 calages à l'aide de DDS (courbes bleues), SCEUA (courbes pourpres) et MADS (courbes orangées) sur le bassin versant Ceizur avec les modèles hydrologiques HYDROTEL à 10 et 19 paramètres, et HSAMI à 23 paramètres

L'aspect général des figures 6.5 et 6.6 présentées ci-haut démontre clairement qu'il existe plusieurs jeux de paramètres finaux quasi-équivalents en termes de valeur de la fonction

objectif, se situant dans des zones de l'espace paramétrique éloignées les unes des autres, et ce, pour les trois modèles hydrologiques étudiés. L'effet d'éparpillement des différents jeux de paramètres finaux témoigne de la présence de l'équifinalité au sein des problèmes de calage de l'étude.

Il est possible de constater que le modèle hydrologique HYDROTEL à 10 paramètres illustre une équifinalité moindre par rapport à celle des deux autres modèles hydrologiques vu la variabilité généralement plus faible des valeurs de la majorité des paramètres. Cette situation est expliquée par la rigidité du modèle (fixation de plusieurs paramètres du modèle, donc plus faible dimensionnalité) qui contraint les algorithmes d'optimisation à se diriger vers des zones d'optima moins nombreuses. Le degré de rigidité dépend entre autres de la sensibilité des paramètres qui ont fait partie du processus d'ajustement, de la quantité de paramètres qui ont été fixés et/ou du lien d'interdépendance et de corrélation qui peut exister entre les divers paramètres fixés et libres. Dans le cas d'HYDROTEL à 10 paramètres, des contraintes sont imposées sur ces trois facteurs, ce qui tend à réduire l'équifinalité, par rapport à la version à 19 paramètres. Cela vient donc confirmer en quelque sorte la pertinence du qualificatif de « rigide » donné au modèle de HYDROTEL à 10 paramètres comme il a été discuté à la section 4.2.3.

Également, on observe que les neuf paramètres supplémentaires ajoutés au processus de calage d'HYDROTEL (version à 19 versus celle à 10 paramètres) ont augmenté et/ou modifié la variabilité des valeurs des 10 paramètres que l'on retrouve au sein des deux versions de ce modèle (paramètres « a » à « j »). En ce qui concerne le bassin versant Cowansville, on note des changements importants de comportement pour les paramètres « a » à « e » tandis que les paramètres « f » à « j » montrent surtout une variabilité plus étendue (à l'exception de « h »). Pour le bassin versant Ceizur, on observe principalement une plus grande variabilité pour les 10 paramètres (« a » à « j »), avec la version d'HYDROTEL à 19 paramètres. L'étude de l'équifinalité démontre que les neuf dimensions supplémentaires viennent indéniablement diminuer le phénomène de contraintes, auparavant observé dans la version d'HYDROTEL à 10 paramètres, étant donné l'augmentation et/ou le

changement de la variabilité des valeurs des paramètres. Le modèle HYDROTEL à 19 paramètres offre à l'algorithme d'optimisation beaucoup plus de souplesse dans la recherche des zones d'optima puisqu'il est moins contraint de toujours se diriger vers les mêmes secteurs de qualité de l'espace paramétrique. La souplesse au sein du problème de calage engendre alors un plus grand nombre de zones d'optima quasi-équivalents en termes de valeur de la fonction objectif.

Par ailleurs, HSAMI comporte quatre paramètres de plus que la version d'HYDROTEL à 19 paramètres, mais la variabilité de leurs paramètres semblent définir un niveau d'équifinalité comparable. De façon générale, les résultats de cette étude (figures 6.5 et 6.6) laissent percevoir une augmentation de l'équifinalité (en lien avec la variabilité des paramètres) à mesure que le nombre de dimensions augmente et ce, peu importe le type de modèle employé. Il faut cependant garder en tête que ces conclusions sont émises en considérant encore une fois les limitations quant au nombre de modèles utilisés et au nombre de bassins versant modélisés.

6.4 Efficacité en fonction du budget de calage

Les résultats et les analyses présentés précédemment démontrent abondamment que l'efficacité en temps de calcul est étroitement liée à un facteur déterminant et essentiel : le budget de calage imposé par l'utilisateur. Il a été observé que le choix de la méthode d'optimisation importe puisque l'écart de performance entre les diverses méthodes de calage sélectionnées varie en fonction du nombre d'évaluations permises. Il apparaît donc pertinent de comprendre le lien qui lie le budget d'évaluations aux fondements des approches de calage.

Tout d'abord, ces travaux ont établi que la recherche aléatoire était tout aussi performante que les autres approches de calage lorsque le budget de l'utilisateur était grandement limité (inférieur à 100 évaluations pour la problématique examinée dans cette étude). Comme les méthodes ne disposent pas de suffisamment d'évaluations pour mettre à profit leurs stratégies

Rapport-Gratuit.com

d'optimisation, aucune différence notable en termes d'efficacité ne se dégage du choix de l'approche de calage sélectionnée. La recherche aléatoire peut donc être employée sans « complexe d'infériorité » pour les budgets très restreints.

Également, cette étude a démontré qu'une méthode telle que DDS, dont la stratégie d'optimisation repose sur l'exploration globale de l'espace des paramètres, est avantagée lors de budget limité (entre 100 et 1500 évaluations), comparativement à une méthode à descente directe (MADS), une méthode évolutionnaire (SCEUA) ou une méthode aléatoire (ACA). Les résultats de cette étude suggèrent que les méthodes axées sur la recherche globale proposent un compromis très intéressant pour l'utilisateur, soit un accommodement équitable entre le temps de calcul requis et la qualité satisfaisante des jeux de paramètres trouvés. De plus, le comportement adaptatif au nombre spécifié d'évaluations du modèle est également une caractéristique de DDS qui valorise sans contredit la sélection de cette méthode d'optimisation.

Malgré l'intérêt des stratégies d'exploration globale de l'espace des paramètres en termes d'économie en temps de calcul, l'utilisation de méthodes de recherche locale peut très certainement apporter une robustesse supplémentaire au processus de calage lorsque le budget d'évaluations est élevé. L'expérimentation présentée à la figure 6.1, sur la « preuve de convergence », démontre que des approches robustes en recherche locale peuvent contribuer à faire progresser la fonction objectif une fois les stratégies de recherche globale épuisées. Bien que le gain obtenu à la figure 6.1 semble relativement faible, certains autres problèmes de calage où l'épuisement des stratégies de recherche globale survient plus tôt pourraient grandement profiter des méthodes de raffinement local robuste telles que les stratégies offertes par MADS. L'emploi des méthodes locales devient donc une solution qui accompagne la recherche globale et non un choix concurrentiel. Lorsque le budget de calage est suffisamment élevé, l'algorithme a alors avantage à transiter de stratégies d'exploration globale vers l'exploitation de processus de raffinement de la valeur de la fonction objectif de manière à fournir un calage de qualité supérieure.

La figure 6.7 illustre la synthèse de la précédente discussion dans le but de mieux visualiser le lien suggéré par les résultats de cette recherche entre le budget de calage et les méthodes d'optimisation à favoriser. Une échelle du nombre d'évaluations a également été mise en place à titre de référence approximative pour le calage du modèle hydrologique coûteux en temps de calcul HYDROTEL. Bien entendu, des études complémentaires, impliquant des bassins versants supplémentaires ainsi que d'autres algorithmes d'optimisation, permettraient de bonifier et de préciser la figure 6.7.

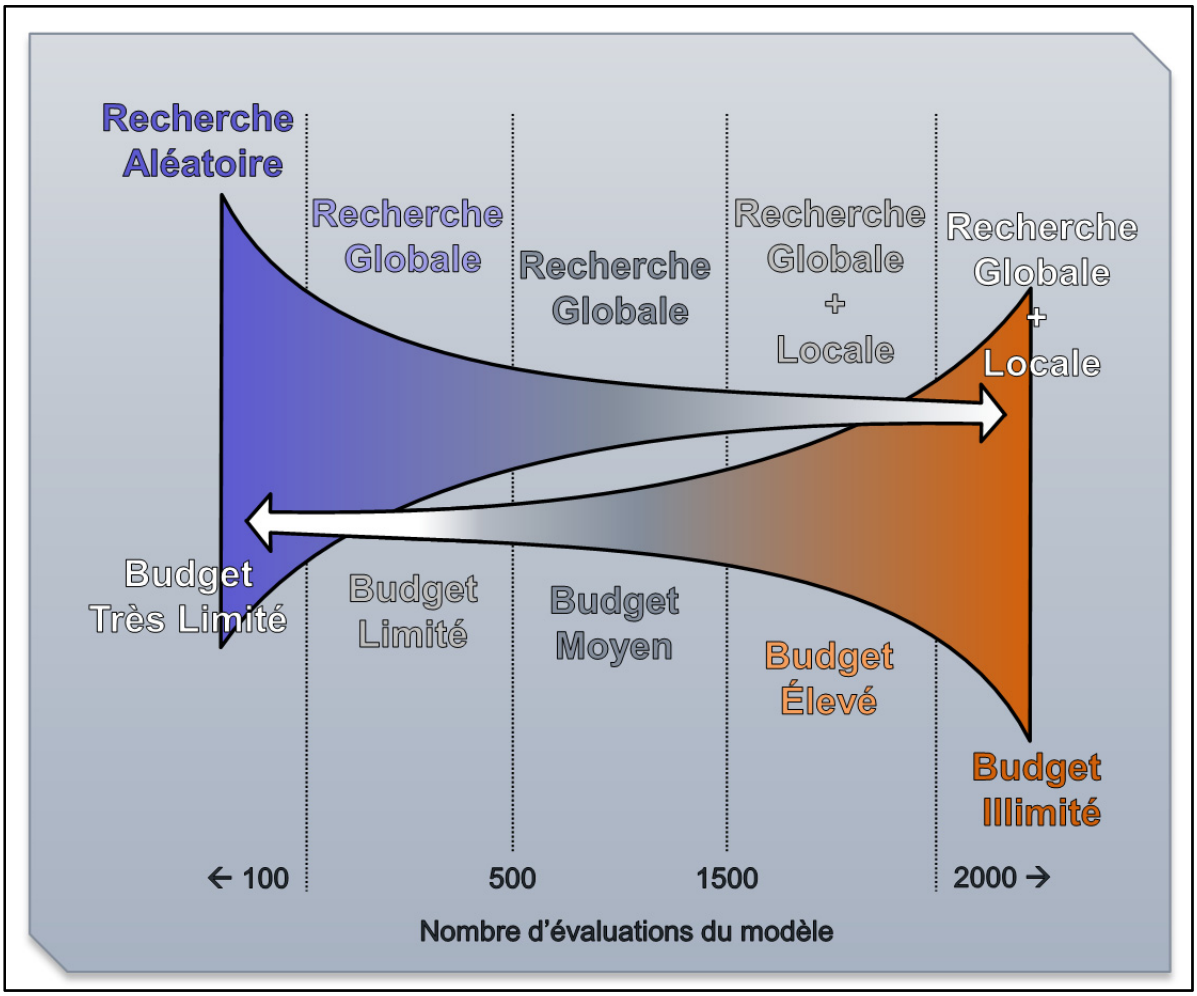

Figure 6.7 Lien entre l'efficacité des méthodes d'optimisation et la disponibilité du budget de calage

Rapport-gratuit.com
LE NUMERO 1 MONDIAL DU MÉMOIRES

6.5 Vers le développement de nouvelles approches de calage

Les analyses et résultats issus de cette étude apportent un meilleur éclairage sur l'efficacité des diverses méthodes d'optimisation sélectionnées pour le calage de modèles hydrologiques coûteux en temps de calcul. En guise de suite possible à ces travaux, des idées et recommandations sont ici présentées, lesquelles pourraient permettre le développement de nouvelles approches de calage adaptées à la problématique étudiée dans ce mémoire.

Tout d'abord, comprenant maintenant l'importance d'une recherche globale aléatoire en début de calage, il serait avantageux de modifier légèrement la proposition faite à la section 4.3.2.3 qui consiste à l'exploitation du « Latin Hypercube Search ». Au lieu de procéder à une recherche « Latin Hypercube Search » d'une seule évaluation au début de chaque itération, il pourrait être envisagé d'initier le calage avec un premier « Latin Hypercube Search » d'une centaine d'évaluations (ce nombre d'évaluations est à analyser) pour ensuite ne poursuivre qu'avec l'étape de sonde (étapes 2 et 3, section 2.4). Les quelques 1900 évaluations restantes serviraient alors à creuser la meilleure solution trouvée à partir du « Latin Hypercube Search » à l'aide des stratégies de raffinement local. De cette façon, il est possible d'éviter la consommation inutile d'un nombre important d'évaluations en raffinant localement des jeux de paramètres de mauvaise qualité. Ceci devrait alors contribuer à améliorer les performances de MADS.

Malgré le potentiel montré par DDS pour la réduction du nombre d'évaluations du modèle hydrologique, les analyses ont permis de démontrer que les solutions trouvées par cet algorithme pourraient poursuivre leur progression si la méthode de calage était pourvue d'une plus grande capacité de recherche locale. Une piste de développement intéressante pourrait alors être celle d'associer les stratégies d'optimisation relatives au processus de raffinement de MADS aux procédés de recuit simulé de DDS qui sont responsables de la capacité d'exploration globale. Ce fusionnement entre les deux méthodes permettrait à la fois d'obtenir une descente rapide de la valeur de la fonction objectif, garantie par les procédés internes de DDS, et une recherche locale où les conditions d'optimalité du jeu de paramètres

final sont remplies par les stratégies de raffinement de MADS. Comme il a été identifié précédemment, l'épuisement précoce des stratégies de recherche globale n'est pas inévitable et la transition vers des méthodes de recherche locale peut très certainement mener au développement d'un algorithme à plus grande efficacité et robustesse. Il est à noter que le logiciel NOMAD permet déjà la substitution complète de l'étape de la recherche au sein de MADS (section 2.4) selon le désir de l'utilisateur. Il serait alors possible d'intégrer les composantes spécifiques de DDS au sein de l'algorithme MADS de manière à bien imbriquer les deux méthodes d'optimisation.

Par ailleurs, comme il a été mentionné à la section 6.2, l'algorithme d'optimisation MADS N+1 peut souffrir d'un léger complexe en matière de raffinement local par rapport à la version par défaut de MADS (« 2N »), mais l'utilisation des directions orthogonales « N+1 » offre un potentiel intéressant de réduction du nombre d'évaluations du modèle hydrologique dont l'utilisateur peut certainement bénéficier. Il pourrait être pertinent d'examiner la combinaison des avantages des deux configurations. Une première avenue serait celle de séparer le processus de calage en deux phases : la descente rapide et le raffinement local. D'une part, MADS pourrait traiter une certaine fraction du nombre d'évaluations permises en descente directe à l'aide des directions orthogonales « N+1 » de manière à accélérer la descente de la valeur de la fonction objectif. D'autre part, MADS poursuivrait son processus de calage avec la fraction restante du nombre d'évaluations à l'aide des directions orthogonales « 2N » de façon à obtenir un meilleur raffinement local. Pour cette configuration, l'utilisateur devrait uniquement spécifier le budget de calage défini pour chacune des phases. Une seconde avenue serait celle d'implanter une forme de génération de directions orthogonales dynamique où l'algorithme fait évoluer le nombre de directions orthogonales visitées en fonction de la progression du processus d'optimisation. Débutant avec les directions orthogonales « N+1 », MADS pourrait ajouter une direction résultante à chaque fois que le dimensionnement du maillage (Δ_k) diminue et, à l'inverse, en enlever une si la dimension du maillage augmente, jusqu'à l'obtention d'au plus le nombre de directions orthogonales équivalent à « 2N ». De cette façon, la transition entre la descente directe et le raffinement local serait non pas en fonction de la décision provenant de l'utilisateur, comme

il est illustré dans la proposition précédente, mais plutôt en fonction de l'avancement du processus de calage.

Enfin, comme il a été noté précédemment, le logiciel d'optimisation NOMAD offre une souplesse de configuration qui passe par le remplacement de l'étape de la recherche au sein de MADS. NOMAD permet donc l'implantation d'une méthode d'évaluation des jeux de paramètres supplémentaire qui peut se substituer complètement à l'étape de la recherche. De façon générale, l'on évalue la performance d'un jeu de paramètres en effectuant une simulation du modèle hydrologique. Toutefois, il existe un moyen de remplacer cette manière d'évaluer la performance des jeux de paramètres en employant des fonctions ou modèles substituts qui ont la faculté de réduire le temps de calcul. Un substitut peut être, d'une part, une fonction mathématique qui reproduit approximativement la surface de réponse obtenue à partir d'un échantillon de simulations du modèle hydrologique (préalablement à l'optimisation et/ou en cours d'optimisation). D'autre part, les modèles substituts peuvent également être des simplifications du processus de modélisation; par exemple, le jeu de paramètres pourrait être évalué en premier lieu sur un modèle hydrologique à courte période de données observées (mais stratégiquement choisie) pour en connaître le potentiel de qualité du critère objectif. Advenant que ce jeu de paramètres est jugé potentiellement de bonne qualité, il peut alors être simulé au sein du modèle réel (plus longue période d'observation) de manière à évaluer le véritable critère objectif. De cette façon, le substitut filtre les jeux de paramètres qui sont de très mauvaise qualité pour attribuer le plus possible le temps de calcul aux jeux de paramètres de bonne qualité (Audet et al., 2012; Razavi et al., 2012). Dans les deux cas, les fonctions ou modèles substituts sont des méthodes qui pré-évaluent et identifient les jeux de paramètres prometteurs (potentiel de performance) pour ensuite les diriger vers le processus de calage régulier. Les fonctions et modèles substituts sont donc une avenue identifiée comme étant très prometteuse et cette voie pourrait grandement contribuer au développement de nouvelles approches de calage adaptées à la problématique de ces travaux.

6.6 Choix méthodologiques de l'étude et autres pistes de recherche possibles

Cette étude comporte certaines limitations quant aux choix méthodologiques effectués et qui ont un impact certain sur les résultats obtenus. Il est donc important de souligner ces choix qui, en contrepartie, représentent des axes de recherche pertinents pour de possibles travaux futurs. Tous les choix méthodologiques ont été justifiés au chapitre 4, mais la plupart de ceux-ci, rappelons-le, ont été faits en considérant le temps de calcul important qu'exigent les expérimentations de cette étude. Les présents travaux se sont penchés sur l'impact du type de modèle hydrologique sur le comportement des algorithmes d'optimisation, mais d'autres analyses fort intéressantes auraient pu être menées, par exemple, en considérant différentes fonctions objectifs d'évaluation de la performance du modèle, ou encore, en examinant l'impact du choix des périodes de calage (longueur ou caractéristiques des séries de données observées) et des bassins versants. Également, d'autres algorithmes d'optimisation pourraient être considérés, de même que différents jeux de données météorologiques (par exemple, des réanalyses). De façon complémentaire à la présente recherche, ces ajouts permettraient de mieux cerner les éléments qui influencent le comportement des algorithmes d'optimisation.

CONCLUSION

La présente recherche s'est concentrée, en premier lieu, sur l'évaluation du comportement et de l'efficacité de trois principaux algorithmes d'optimisation, SCEUA, DDS et MADS, lorsqu'employés pour le calage de deux versions d'HYDROTEL (10 et 19 paramètres), un modèle hydrologique coûteux en temps de calcul. Également, une approche de calage complétement aléatoire (ACA) a été utilisée dans le but de placer un seuil minimal de performance en termes d'efficacité. Deux bassins versants de caractéristiques physiographiques différentes ont été étudiés. Les résultats ont démontré que la méthode d'optimisation DDS offre un potentiel de réduction du temps de calcul (nombre d'évaluations du modèle hydrologique) plus intéressant que les deux autres approches. La capacité d'exploration globale de l'espace des paramètres de DDS en fait une méthode de calage robuste qui surpasse les stratégies employées par les deux autres concurrents dans le cadre de la problématique examinée. Cependant, il existe un lien étroit entre le budget de calage alloué par l'utilisateur (nombre d'évaluations du modèle hydrologique) et le choix des méthodes d'optimisation puisqu'il a été montré que, à partir d'un certain budget de calage (dans le cadre de cette recherche, 1500 évaluations du modèle), les diverses méthodes d'optimisation apportent des jeux de paramètres de qualité comparable. À travers la discussion des résultats obtenus, les stratégies d'optimisation des trois algorithmes ont été rigoureusement analysées, et un guide (ou synthèse) a été proposé quant aux types de stratégies à favoriser selon l'ampleur du budget de calage disponible (figure 6.7). Les stratégies liées aux recherches aléatoires, globales et locales ont ainsi été positionnées sur une échelle de budget de calage.

De façon complémentaire, deux configurations supplémentaires du logiciel NOMAD (à partir duquel MADS a été exécuté) ont également été étudiées quant à leur capacité à fournir rapidement un jeu de paramètres de bonne qualité. D'une part, la configuration à groupes de variables (MADS VG) a su exploiter la connaissance du modèle hydrologique à son avantage en regroupant les paramètres selon leur impact dominant, soit par processus hydrologique ou section de l'hydrogramme de ruissellement. D'autre part, la configuration à directions orthogonales « N+1 » (MADS N+1) a, quant à elle, permis de réduire le nombre

d'évaluations du modèle hydrologique lors de la génération d'un motif de points. Les deux configurations ont montré un potentiel de réduction du nombre d'évaluations de la fonction objectif, en permettant à MADS de progresser plus rapidement dans les directions de descente de la fonction objectif. Toutefois, d'une part, MADS N+1 démontre une légère diminution de la qualité de la solution finale trouvée et, d'autre part, MADS VG nécessite une bonne connaissance du modèle hydrologique pour que les groupes de variables soient constitués adéquatement.

Ensuite, les comportements des algorithmes d'optimisation sélectionnés ont également été étudiés pour le calage d'HSAMI (conceptuel et global), dans le but de mieux comprendre l'impact du type de modèle sur la complexité du problème d'optimisation. L'étude des comportements des algorithmes a montré que le problème de calage du modèle HSAMI exige un nombre d'évaluations du modèle hydrologique plus important que pour le calage d'HYDROTEL. Ces résultats suggèrent donc que les deux modèles hydrologiques (HYDROTEL et HSAMI) donnent lieu à des problèmes d'optimisation à caractéristiques dissemblables en termes de surface de l'espace paramétrique, ce qui entraîne inévitablement des niveaux de complexité différents (dimensionnalité et type de modèle). Cette portion de la présente recherche a par ailleurs permis d'examiner l'impact de la dimensionnalité (nombre de paramètres de calage) sur le niveau de complexité du problème d'optimisation et, par le fait même, sur l'équifinalité. Entre, d'une part, la version d'HYDROTEL à 10 paramètres et, d'autre part, la version à 19 paramètres de même que HSAMI à 23 paramètres, il est observé que l'équifinalité augmente.

Ces travaux ont également étudié les stratégies de recherche employées par les divers algorithmes. Ceux-ci montrent que MADS tend à être victime, d'une part, de la qualité du jeu de paramètres initial (point de départ de l'optimisation) puisque cette méthode de raffinement local n'est pas accompagnée d'une méthode de recherche globale robuste qui identifie préalablement une zone d'optima de bonne qualité, dans le contexte du calage des modèles hydrologiques. Comme il a été mentionné au sein de la discussion, le « Latin Hypercube Search » pourrait être une méthode de recherche globale robuste qui complémente le

processus de raffinement de MADS d'autant plus que celle-ci se retrouve déjà au sein des fonctionnalités disponibles de NOMAD. D'autre part, MADS est également victime de son obligation à converger vers un optimum local, ce qui entraîne inévitablement un ralentissement dans la progression de la valeur de la fonction objectif contrairement aux heuristiques et métaheuristiques qui n'ont pas cette contrainte au sein de leur processus de calage. De plus, ces méthodes sont orientées vers des approches exploratrices qui les conduisent à performer de manière plus efficace en termes de progression de la valeur de la fonction objectif en fonction du temps de calcul. Notons entre autre le comportement adaptatif et la capacité à trouver les zones d'optima de bonne qualité de DDS.

Finalement, l'aspect exploratoire de ces travaux permet de fournir plusieurs pistes pertinentes qui peuvent servir au développement de nouvelles approches de calage plus adaptées au contexte de la problématique de calage des modèles hydrologiques coûteux en temps de calcul. Les propositions faites à la section 6.5 présentent des avenues intéressantes qui seront assurément analysées en profondeur dans le cadre de prochaines études scientifiques, à savoir : la combinaison des stratégies pertinentes de DDS et de MADS en un seul algorithme, la séparation du processus de calage de MADS en deux phases, l'une étant à directions orthogonales « N+1 » et l'autre à « 2N », la constitution d'un nouveau motif de points selon des directions orthogonales dynamiques (transition progressive de « N+1 » vers « 2N »), l'utilisation d'un « Latin Hypercube Search » en début de processus de calage et le recours aux fonctions substituts de manière à mieux identifier les points à évaluer au sein du modèle hydrologique.

Rapport-Gratuit.com

ANNEXE I

PARAMÈTRES DES MODÈLES HYDROLOGIQUES

Cette annexe présente les paramètres des modèles hydrologiques employés pour cette étude, soient HYDROTEL (à dix et dix-neuf paramètres) et HSAMI.

Tableau-A I-1 Identification des vingt-sept paramètres d'HYDROTEL

HYDROTEL	
Numéro	**Description des paramètres**
1	Gradient vertical des précipitations (mm/jour)
2	Gradient vertical de la température (°C/100m)
3	Température de passage de la pluie en neige (°C)
4	Taux de fonte entre la neige et le sol (mm/jour)
5	Densité maximale du couvert nival (kg/m³)
6	Constante de tassement
7	Taux de fonte pour les connifères (mm/jour*°C)
8	Seuil de fonte de la température des connifères (°C)
9	Taux de fonte pour les feuillus (mm/jour*°C)
10	Seuil de fonte de la température des feuillus (°C)
11	Taux de fonte pour le milieu non-forestier (mm/jour*°C)
12	Seuil de fonte de la température du milieu non-forestier (°C)
13	Coefficient multiplicatif de l'évapotranspiration potentielle
14	Bilan vertical - limite inférieure de la zone #1 (m)
15	Bilan vertical - limite inférieure de la zone #2 (m)
16	Bilan vertical - limite inférieure de la zone #3 (m)
17	Bilan vertical - coefficient d'extinction
18	Bilan vertical - coefficient de récession (m/heure)
19	Coefficient additif d'optimisation de la classe de sol
20	Coefficient multiplicatif d'optimisation de l'assèchement
21	variation maximale de l'humidité relative par pas de temps
22	Coefficient de Manning du milieu forestier
23	Coefficient de Manning de l'eau
24	Coefficient de Manning des autres milieux
25	Lame de ruissellement de l'hydrogramme géomorphologique
26	Coefficient d'optimisation de la rugosité
27	Coefficient d'optimisation des largeurs de rivières

Rapport-gratuit.com
LE NUMERO 1 MONDIAL DU MÉMOIRES

Tableau-A I-2 Paramètres de la version du modèle hydrologique HYDROTEL à dix variables et de la version du modèle hydrologique HYDROTEL à dix-neuf variables

HYDROTEL			
Numéro	Description des paramètres	Version à 10 paramètres	Version à 19 paramètres
1	Gradient vertical des précipitations (mm/jour)		●
2	Gradient vertical de la température (°C/100m)		●
3	Température de passage de la pluie en neige (°C)	●	●
4	Taux de fonte entre la neige et le sol (mm/jour)		●
5	Densité maximale du couvert nival (kg/m³)		●
6	Constante de tassement		●
7	Taux de fonte pour les connifères (mm/jour*°C)	●	●
8	Seuil de fonte de la température des connifères (°C)	●	●
9	Taux de fonte pour les feuillus (mm/jour*°C)		●
10	Seuil de fonte de la température des feuillus (°C)		●
11	Taux de fonte pour le milieu non-forestier (mm/jour*°C)	●	●
12	Seuil de fonte de la température du milieu non-forestier (°C)	●	●
13	Coefficient multiplicatif de l'évapotranspiration potentielle	●	●
14	Bilan vertical - limite inférieure de la zone #1 (m)	●	●
15	Bilan vertical - limite inférieure de la zone #2 (m)	●	●
16	Bilan vertical - limite inférieure de la zone #3 (m)	●	●
17	Bilan vertical - coefficient d'extinction		
18	Bilan vertical - coefficient de récession (m/heure)	●	●
19	Coefficient additif d'optimisation de la classe de sol		
20	Coefficient multiplicatif d'optimisation de l'assèchement		
21	variation maximale de l'humidité relative par pas de temps		
22	Coefficient de Manning du milieu forestier		●
23	Coefficient de Manning de l'eau		
24	Coefficient de Manning des autres milieux		
25	Lame de ruissellement de l'hydrogramme géomorphologique		●
26	Coefficient d'optimisation de la rugosité		
27	Coefficient d'optimisation des largeurs de rivières		

Tableau-A I-3 Identification des vingt-trois paramètres de HSAMI

HSAMI	
Numéro	**Description des paramètres**
1	Efficacité de l'évapotranspiration l'été
2	Efficacité de l'évapotranspiration l'hivers
3	Taux de fonte de la neige le jour (cm/Δ°C/jour)
4	Taux de fonte de la neige la nuit (cm/Δ°C/jour)
5	Température de fonte de la neige le jour (°C)
6	Température de fonte de la neige la nuit (°C)
7	Température de référence pour calculer le flux de chaleur pluie-couvert nival (°C)
8	Effet de redoux sur les aires enneigées
9	Effet de gel sur le ruissellement
10	Effet de sol sur le ruissellement (cm)
11	Seuil minimal du sol saturé sur le ruissellement (cm)
12	Épaisseur minimale du sol - quantié d'eau dans la zone vadose qui ne peut s'écouler par gravité (cm)
13	Épaisseur maximale du sol - quantité d'eau que peut contenir la zone vadose (cm)
14	Épaisseur maximale de la nappe (cm)
15	Portion de ruissellement de surface
16	Portion de ruissellement dans le sol maximale
17	Taux de vidange de la zone vadose vers la nappe (1/jour)
18	Taux de vidange de la nappe (1/jour)
19	Taux de vidange de la réserve intermédiaire (1/jour)
20	Mode de l'hydrogramme de surface (jour)
21	Forme de l'hydrogramme de surface
22	Mode de l'hydrogramme intermédiaire (jour)
23	Forme de l'hydrogramme intermédiaire

ANNEXE II

BORNES DE CALAGE DES PROBLÈMES D'OPTIMISATION

Cette annexe présente les bornes inférieures et supérieures des modèles hydrologiques employés pour cette étude, soient HYDROTEL (à dix et dix-neuf paramètres) et HSAMI.

Tableau-A II-1 Identification des bornes inférieures et supérieures pour le modèle hydrologique HYDROTEL à dix paramètres

HYDROTEL à 10 paramètres			
Numéro	Description des paramètres	Bornes Inférieures	Bornes Supérieures
3	Température de passage de la pluie en neige (°C)	-6	4
7	Taux de fonte pour les connifères (mm/jour*°C)	0,01	6,39
8	Seuil de fonte de la température des connifères (°C)	1,61	4
11	Taux de fonte pour le milieu non-forestier (mm/jour*°C)	6,41	14
12	Seuil de fonte de la température du milieu non-forestier (°C)	-6	1,59
13	Coefficient multiplicatif de l'évapotranspiration potentielle	0,7	1,2
14	Bilan vertical - limite inférieure de la zone #1 (m)	0,05	1
15	Bilan vertical - limite inférieure de la zone #2 (m)	0,05	1
16	Bilan vertical - limite inférieure de la zone #3 (m)	0,05	4
18	Bilan vertical - coefficient de récession (m/heure)	0,00000001	0,0001

Tableau-A II-2 Identification des bornes inférieures et supérieures pour le modèle hydrologique HYDROTEL à dix-neuf paramètres

HYDROTEL à 19 paramètres			
Numéro	Description des paramètres	Bornes Inférieures	Bornes Supérieures
1	Gradient vertical des précipitations (mm/jour)	0	2
2	Gradient vertical de la température (°C/100m)	-0,8	-0,3
3	Température de passage de la pluie en neige (°C)	-6	4
4	Taux de fonte entre la neige et le sol (mm/jour)	0,4	0,6
5	Densité maximale du couvert nival (kg/m³)	300	550
6	Constante de tassement	0,005	0,015
7	Taux de fonte pour les connifères (mm/jour*°C)	0,01	6,39
8	Seuil de fonte de la température des connifères (°C)	0,01	1,99
9	Taux de fonte pour les feuillus (mm/jour*°C)	0,01	3,99
10	Seuil de fonte de la température des feuillus (°C)	0,01	1,99
11	Taux de fonte pour le milieu non-forestier (mm/jour*°C)	0,01	3,99
12	Seuil de fonte de la température du milieu non-forestier (°C)	-6	1,59
13	Coefficient multiplicatif de l'évapotranspiration potentielle	0,7	1,2
14	Bilan vertical - limite inférieure de la zone #1 (m)	0,05	1
15	Bilan vertical - limite inférieure de la zone #2 (m)	0,05	1
16	Bilan vertical - limite inférieure de la zone #3 (m)	0,05	4
18	Bilan vertical - coefficient de récession (m/heure)	0,00000001	0,0001
22	Coefficient de Manning du milieu forestier	0,041	0,3
25	Lame de ruissellement de l'hydrogramme géomorphologique	0,0002	0,002

Tableau-A II-3 Identification des bornes inférieures et supérieures pour le modèle hydrologique HSAMI à vingt-trois paramètres

HSAMI à 23 paramètres			
Numéro	**Description des paramètres**	**Bornes Inférieures**	**Bornes Supérieures**
1	Efficacité de l'évapotranspiration l'été	0,6	3
2	Efficacité de l'évapotranspiration l'hivers	0	0,3
3	Taux de fonte de la neige le jour (cm/Δ°C/jour)	0,05	0,4
4	Taux de fonte de la neige la nuit (cm/Δ°C/jour)	0,05	0,5
5	Température de fonte de la neige le jour (°C)	-6	7
6	Température de fonte de la neige la nuit (°C)	-6	6
7	Température de référence pour calculer le flux de chaleur pluie-couvert nival (°C)	-6	4
8	Effet de redoux sur les aires enneigées	0,8	5
9	Effet de gel sur le ruissellement	0,8	15
10	Effet de sol sur le ruissellement (cm)	10	45
11	Seuil minimal du sol saturé sur le ruissellement (cm)	1	8
12	Épaisseur minimale du sol - quantié d'eau dans la zone vadose qui ne peut s'écouler par gravité (cm)	0	7
13	Épaisseur maximale du sol - quantité d'eau que peut contenir la zone vadose (cm)	3	25
14	Épaisseur maximale de la nappe (cm)	4	30
15	Portion de ruissellement de surface	0,15	0,7
16	Portion de ruissellement dans le sol maximale	0,3	1
17	Taux de vidange de la zone vadose vers la nappe (1/jour)	0,009	0,07
18	Taux de vidange de la nappe (1/jour)	0,006	0,018
19	Taux de vidange de la réserve intermédiaire (1/jour)	0,6	1,2
20	Mode de l'hydrogramme de surface (jour)	0,3	5
21	Forme de l'hydrogramme de surface	0,4	1,5
22	Mode de l'hydrogramme intermédiaire (jour)	1,5	13
23	Forme de l'hydrogramme intermédiaire	0,15	1,5

ANNEXE III

JEUX DE PARAMÈTRES INITIAUX DES PROBLÈMES D'OPTIMISATION

Cette annexe présente les jeux de paramètres initiaux générés aléatoirement employés pour le calage d'HYDROTEL à dix et dix-neuf paramètres et de HSAMI à vingt-trois paramètres.

Tableau-A III-1 Identification des trente-deux jeux de paramètres initiaux pour le problème d'optimisation d'HYDROTEL à dix paramètres

HYDROTEL à 10 paramètres										
N° / JP	(3)	(7)	(8)	(11)	(12)	(13)	(14)	(15)	(16)	(18)
1	-4,2661	2,5042	3,5970	12,5075	-5,5410	0,8996	0,5505	0,4460	2,6446	6,2801E-05
2	-3,0802	2,7639	1,6470	13,8790	-4,7312	0,7531	0,4038	0,2382	1,9843	3,3956E-05
3	3,5163	5,8817	1,7359	12,0103	-3,9574	0,9114	0,5705	0,9456	1,7001	9,8305E-05
4	-2,9855	4,4830	3,2025	10,5020	-0,7014	1,0333	0,2192	0,1716	3,9964	1,712E-05
5	-5,6740	3,5905	3,7177	11,4890	-4,5546	0,8845	0,4877	0,9826	0,6678	8,5554E-05
6	0,4476	2,4106	2,0663	9,6604	-2,3415	0,7603	0,6100	0,2649	1,5692	5,8303E-05
7	-3,4819	1,8630	3,0848	8,4235	0,2570	1,1913	0,7437	0,3767	2,3571	1,0786E-05
8	3,0631	5,6222	3,5644	8,3889	-1,4888	0,7113	0,4540	0,3471	0,6879	1,7885E-05
9	-1,7711	0,6112	3,0405	9,9843	-0,7177	1,0499	0,6566	0,0819	0,3218	3,1967E-05
10	-0,6914	4,1854	2,5842	12,6337	-0,5477	1,1843	0,5548	0,3589	0,4672	6,11E-05
11	1,7880	2,7116	1,8271	8,4325	-4,8337	0,8405	0,4681	0,5508	1,8568	8,7538E-05
12	-0,8195	6,0303	3,1341	13,6789	-4,1730	1,0381	0,3246	0,6882	2,7958	6,8086E-06
13	-3,4521	1,4394	3,2061	12,8189	-3,3855	1,0903	0,6916	0,0564	2,4286	3,8683E-05
14	3,1599	0,0173	2,7153	9,6308	-2,5016	1,0851	0,3563	0,7955	1,9119	3,5859E-06
15	-4,2413	4,6148	2,7416	7,5692	-3,4109	1,0037	0,2322	0,7515	1,0093	9,1743E-05
16	-3,3094	4,8939	2,0609	8,5921	-5,3084	0,9881	0,6992	0,5693	1,7316	6,4448E-05
17	0,4762	4,3421	3,1295	13,5839	-4,4142	1,0546	0,2744	0,1634	2,4489	4,5019E-05
18	-1,4127	4,2332	3,4510	9,0682	-0,9753	0,9081	0,8498	0,8413	1,0629	6,135E-05
19	-0,1775	3,4599	3,6892	8,4197	-3,5858	0,7596	0,9428	0,6633	1,9439	6,3935E-05
20	-0,5528	4,1398	2,9099	11,8827	-2,0343	1,1969	0,2577	0,1505	0,4833	6,3685E-06
21	-1,9542	2,8706	2,4843	12,2050	-1,2343	1,0860	0,9362	0,9741	0,8085	1,3896E-05
22	0,9627	0,6086	2,8657	10,4353	0,5361	0,9424	0,4238	0,6879	2,9780	5,201E-05
23	-2,5229	0,9670	3,0108	8,3997	-5,6626	1,0775	0,2806	0,4703	2,7668	3,5929E-05
24	1,3634	2,5282	3,2434	11,7537	-2,6429	0,7098	0,3643	0,4531	1,1176	1,9713E-05
25	2,2172	2,7529	3,7318	9,3791	-0,1624	0,8984	0,8181	0,7673	1,5407	2,161E-05
26	1,9041	6,0666	2,3929	11,5049	-2,6707	1,1168	0,7804	0,2089	3,4548	9,8987E-05
27	-0,8558	5,6517	3,0154	7,5846	-4,4830	0,9035	0,7613	0,8343	3,1704	3,1859E-05
28	-0,6594	0,5839	1,8770	7,4445	-0,8490	0,9476	0,2302	0,5203	0,6331	5,5069E-06
29	2,5071	3,5864	3,8318	11,6977	-1,5766	1,1077	0,8851	0,9895	0,0521	8,6545E-05
30	0,1257	6,3259	2,8712	10,0496	0,0822	0,8139	0,5232	0,9058	2,3199	8,4519E-05
31	1,3864	3,7486	2,1997	11,4681	-5,3664	1,0130	0,6779	0,7433	3,5685	9,823E-05
32	1,6903	3,7196	3,8287	10,8129	-5,8711	0,7604	0,8696	0,5101	3,3872	2,0948E-05

Rapport-gratuit.com

Tableau-A III-2a Identification des jeux de paramètres initiaux pour le problème d'optimisation d'HYDROTEL à dix-neuf paramètres

HYDROTEL à 19 paramètres										
N° / JP	(3)	(7)	(8)	(11)	(12)	(13)	(14)	(15)	(16)	(18)
1	3,2274	5,1164	0,5762	2,1738	1,4745	1,0578	0,8470	0,4616	1,9090	5,6076E-05
2	1,8030	4,2751	0,2743	0,0958	-1,7508	0,8504	0,9424	0,9819	1,1822	8,0084E-05
3	0,4630	3,3353	0,7472	3,7398	0,2962	1,1245	0,4039	0,6135	3,4966	9,3351E-05
4	1,5675	2,6708	1,9341	3,9421	0,5589	0,8944	0,4820	0,2844	3,1485	8,8285E-05
5	2,8249	1,8280	1,3430	2,6538	-5,0678	0,9037	0,3115	0,7308	1,1694	8,9621E-05
6	1,1380	5,6525	1,4373	0,0841	-0,8784	0,9193	0,4659	0,1612	3,2680	3,2492E-05
7	3,5092	4,6186	0,8022	3,3208	-4,9804	0,7302	0,1300	0,2057	1,3307	3,018E-05
8	-0,4646	3,2986	0,6648	1,7214	-2,2672	0,7355	0,8934	0,1114	1,7729	8,2665E-05
9	-0,9616	3,9197	1,6325	2,1269	-4,4663	0,9269	0,4565	0,9678	2,4992	6,9542E-05
10	1,3139	2,3070	0,9093	1,5478	-0,1135	1,0671	0,4588	0,7091	3,7836	7,8425E-05
11	2,6371	0,5081	1,3347	2,0008	-4,3454	0,9858	0,1661	0,6876	2,4184	5,6071E-06
12	3,0805	0,6991	1,0337	0,5798	-1,7544	0,7023	0,7783	0,8563	3,6714	9,8697E-05
13	-4,2895	5,9980	1,1792	1,7637	1,1492	1,0280	0,4793	0,8477	2,1539	5,5393E-05
14	-1,9861	5,3269	0,8092	1,5629	-3,2642	0,7701	0,2971	0,1325	1,7461	2,5736E-05
15	-5,9661	1,4179	0,0126	0,7629	-4,9185	0,8340	0,2161	0,1817	2,4156	9,0107E-05
16	-5,7389	6,1008	0,8626	3,8370	-0,2133	0,7037	0,6960	0,7207	2,5983	5,5235E-05
17	3,1020	5,8100	1,1814	1,3336	0,4748	0,9212	0,9091	0,0815	2,1531	7,1653E-05
18	-3,5850	1,5613	0,3152	3,8165	1,1017	1,1094	0,7418	0,2170	1,4735	1,8887E-05
19	1,6241	3,6852	1,4904	2,5792	-5,0648	0,9522	0,3799	0,1375	0,6340	1,9825E-05
20	-2,0732	0,1686	1,3394	3,3419	1,3737	0,7285	0,4778	0,6033	2,7622	7,1946E-05
21	3,6216	1,1953	0,3922	1,3697	1,0807	0,8953	0,3096	0,1943	1,6186	3,7479E-05
22	-0,3064	0,4017	0,9927	2,5664	-4,3206	1,1185	0,9725	0,8541	2,0487	2,7895E-05
23	3,1107	4,4737	1,4459	0,9249	-1,6278	1,1053	0,4337	0,9890	0,4055	3,2101E-05
24	-1,4779	4,8092	0,2275	0,4468	-3,9516	0,9623	0,9740	0,7249	1,2818	2,9153E-05
25	-0,1536	1,8290	1,6489	0,7701	-2,6412	0,8967	0,8352	0,6930	0,8700	3,1817E-05
26	-5,4614	5,1463	0,9037	1,5329	-0,0066	0,8821	0,5557	0,7261	3,4923	3,2876E-05
27	-4,1343	4,9937	0,3977	3,9596	0,0892	0,9121	0,7424	0,5234	3,2455	3,5657E-05
28	-0,4118	1,1829	0,9959	2,0710	1,5463	1,1274	0,9643	0,6950	1,6438	9,3499E-05
29	3,6387	0,7477	0,1119	1,2213	-1,5963	0,9655	0,9061	0,5635	1,7563	5,4271E-05
30	-3,2122	1,2747	0,3962	1,3108	0,6818	0,9356	0,4338	0,2203	3,8773	4,0752E-05
31	-2,6567	3,8166	1,7959	2,8042	-3,1351	1,0675	0,9564	0,5657	2,1834	3,1118E-05
32	-3,9153	2,9127	0,2620	0,0444	-0,4815	0,8771	0,7914	0,4648	1,7744	4,9308E-06

Tableau-A III-2b Identification des jeux de paramètres initiaux pour le problème d'optimisation d'HYDROTEL à dix-neuf paramètres

HYDROTEL à 19 paramètres									
N° / JP	(1)	(2)	(4)	(5)	(6)	(9)	(10)	(22)	(25)
1	0,5382	-0,4255	0,5008	461,70	0,0081	0,5621	0,9516	0,1349	0,001619
2	1,7922	-0,5012	0,5768	535,93	0,0105	2,9090	1,1520	0,0477	0,001004
3	1,3369	-0,6966	0,5308	318,01	0,0091	2,6644	1,8588	0,2510	0,001072
4	1,8274	-0,5209	0,5198	337,22	0,0140	1,8026	0,4172	0,2740	0,001573
5	1,6532	-0,6050	0,4996	473,70	0,0133	2,4363	1,1480	0,1254	0,001022
6	0,4925	-0,6286	0,4751	436,64	0,0106	1,5854	0,7983	0,1745	0,001384
7	0,0234	-0,5300	0,4191	336,63	0,0113	3,4301	1,9390	0,1888	0,001994
8	0,7891	-0,4933	0,5637	521,56	0,0143	0,7693	0,5220	0,2735	0,001268
9	1,4403	-0,6266	0,5034	439,17	0,0066	2,2470	1,3857	0,1515	0,001705
10	1,4111	-0,7453	0,4780	447,73	0,0096	0,2104	0,4628	0,2571	0,000228
11	0,1127	-0,7237	0,4039	408,79	0,0133	2,4672	1,0399	0,2647	0,000376
12	1,0103	-0,6643	0,4202	426,96	0,0109	3,0463	0,1743	0,2124	0,001131
13	1,3601	-0,6164	0,4479	444,73	0,0137	1,6290	0,2330	0,1560	0,000740
14	0,5951	-0,5876	0,4238	423,77	0,0121	0,9794	1,5644	0,0602	0,000909
15	1,8788	-0,6894	0,4965	394,00	0,0102	1,0642	0,1453	0,1540	0,000513
16	0,4362	-0,4138	0,4456	392,72	0,0139	3,4184	0,8068	0,1234	0,001296
17	0,3586	-0,6317	0,4375	380,48	0,0090	2,1933	0,1065	0,1842	0,000695
18	0,0024	-0,6418	0,5399	456,31	0,0104	1,7574	0,5791	0,1709	0,001571
19	1,3445	-0,5842	0,5389	364,20	0,0051	2,1285	0,5632	0,2861	0,001832
20	1,3001	-0,4365	0,4748	445,40	0,0062	0,2395	1,9499	0,1148	0,001271
21	0,2622	-0,5825	0,4183	453,66	0,0051	2,2916	1,5737	0,1020	0,001006
22	1,4932	-0,6815	0,5915	455,07	0,0110	0,6970	0,1889	0,1071	0,001745
23	1,0228	-0,7697	0,5451	439,14	0,0103	3,3133	1,7103	0,2454	0,000772
24	1,7007	-0,3442	0,5279	363,84	0,0059	3,3463	1,1677	0,2866	0,000310
25	0,2676	-0,4643	0,5142	342,44	0,0065	1,9048	1,8080	0,1840	0,000259
26	1,3002	-0,3126	0,4152	446,75	0,0091	1,2404	0,5324	0,2375	0,001991
27	0,1465	-0,5045	0,5820	348,44	0,0093	2,9917	0,0876	0,2861	0,001575
28	0,9590	-0,6841	0,4793	476,27	0,0106	3,0214	1,9811	0,2903	0,001163
29	1,4248	-0,7917	0,5602	335,63	0,0098	1,0322	0,7408	0,2124	0,000505
30	1,6890	-0,4923	0,4753	519,30	0,0128	1,8605	1,6217	0,2737	0,000973
31	0,1425	-0,7090	0,4186	415,87	0,0051	3,6518	1,2826	0,0414	0,000255
32	0,0993	-0,7544	0,5188	360,27	0,0134	3,4217	1,9180	0,1676	0,000597

Rapport-gratuit.com
LE NUMERO 1 MONDIAL DU MÉMOIRES

Tableau-A III-3a Identification des jeux de paramètres initiaux pour le problème d'optimisation d'HSAMI à vingt-trois paramètres

HSAMI à 23 paramètres										
JP \ N°	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
1	2,7748	0,0982	0,2056	0,2658	3,5580	-5,8192	-5,2230	3,0325	5,6084	44,6378
2	1,5523	0,0867	0,1544	0,3037	-4,5831	-2,9075	-2,3844	1,7912	4,3088	14,4612
3	0,7474	0,2450	0,2552	0,1895	-4,4529	2,5422	-2,6930	1,9175	4,1129	34,4472
4	2,4477	0,1356	0,1017	0,1727	-4,0062	3,6781	0,0510	1,8665	1,8002	11,6846
5	2,2693	0,2658	0,0795	0,4286	-5,6684	2,6922	-2,9334	3,8304	9,6478	26,6210
6	1,0867	0,2795	0,1890	0,1188	3,1225	1,1999	-3,6757	1,9095	13,4687	32,5185
7	2,3489	0,1431	0,0831	0,2263	-5,6823	-1,2363	0,1452	2,1712	6,4243	32,0402
8	1,8918	0,2688	0,3908	0,1539	-3,5933	-2,2052	-5,8618	3,3102	5,7644	30,4921
9	1,9754	0,2307	0,2183	0,4553	2,2317	-0,2039	-5,6027	1,6763	7,7765	20,8967
10	2,9331	0,2880	0,3345	0,2358	5,7235	-0,6017	2,2343	3,4486	13,3579	22,7726
11	1,3336	0,0660	0,2568	0,3362	4,1239	-5,3350	3,8689	1,4284	13,3484	21,3321
12	1,8939	0,0873	0,3802	0,2070	5,5418	-3,1476	4,3692	2,6894	4,9242	30,9564
13	1,3378	0,0194	0,3701	0,0608	4,4697	3,8398	0,8283	3,6320	9,1576	10,4509
14	0,7462	0,0893	0,1824	0,1526	-5,4013	-5,2551	3,6426	3,1557	8,4593	22,6788
15	2,0625	0,2746	0,2768	0,3692	-2,7942	2,1036	0,2068	2,9850	7,1279	33,5787
16	1,6613	0,0792	0,3062	0,1410	-2,9093	-2,1998	-4,9676	2,8920	8,6791	33,8000
17	2,6483	0,1534	0,2050	0,1260	1,8312	-5,0956	2,0323	1,1562	7,7134	23,6401
18	1,0795	0,0737	0,1721	0,2803	5,4832	-4,8296	4,9460	3,7939	4,7495	38,0967
19	1,3695	0,1384	0,2108	0,3571	4,3844	-3,3122	-3,3705	0,8985	7,5095	20,6437
20	2,9370	0,1321	0,3370	0,2866	2,3686	0,7601	-0,8542	1,4219	11,8169	32,4373
21	1,0623	0,2280	0,3871	0,1397	0,8877	-5,7638	-1,2201	2,7538	3,8736	40,5199
22	1,9995	0,0297	0,1428	0,2732	0,5964	-1,3318	3,2731	1,9666	8,0344	33,7444
23	1,1293	0,0987	0,1631	0,3796	-2,1456	5,9215	-1,7456	3,2372	13,0522	30,6720
24	1,4939	0,1435	0,2934	0,1142	1,5922	5,4497	3,3135	3,7788	5,5805	27,0345
25	1,4593	0,1577	0,0838	0,3768	3,3674	1,3068	-1,6673	1,3053	7,9944	21,9521
26	1,1815	0,1440	0,3468	0,4281	-0,9082	-0,7059	-4,1134	2,1036	9,6895	43,8879
27	1,6870	0,2443	0,1774	0,2755	0,6594	4,5607	-0,8522	3,0569	11,3747	12,4872
28	2,7569	0,1350	0,3363	0,1467	-1,3111	-4,8310	1,1442	3,4639	11,3170	15,4279
29	1,8401	0,2008	0,3209	0,0635	3,7418	4,8166	3,7821	1,9947	13,4726	39,5598
30	1,8350	0,1348	0,0686	0,2403	-4,2229	-5,2787	-1,7468	0,9928	6,9215	32,7110
31	2,4719	0,0108	0,2571	0,0829	0,6410	1,6995	-0,0379	2,0191	2,7077	24,9181
32	1,1243	0,0787	0,3739	0,4332	3,1302	3,5202	-3,1696	3,5185	8,6464	27,3641

Tableau-A III-3b Identification des jeux de paramètres initiaux pour le problème d'optimisation d'HSAMI à vingt-trois paramètres

HSAMI à 23 paramètres										
N° / JP	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)	(19)	(20)
1	1,4774	5,6338	17,0108	28,7085	0,2842	0,4268	0,0265	0,0158	0,7516	3,0315
2	1,2455	4,3375	19,2736	6,9778	0,5157	0,3480	0,0606	0,0079	1,0561	1,4882
3	5,9527	4,1406	23,7153	8,0962	0,5585	0,3724	0,0584	0,0132	0,9843	0,7796
4	4,1281	0,5442	19,9820	10,3978	0,2497	0,6769	0,0630	0,0150	1,1462	1,1926
5	5,4833	6,2695	10,4274	4,4454	0,1525	0,6069	0,0520	0,0087	1,1419	0,3684
6	7,8904	4,1584	19,3618	28,1292	0,6988	0,6498	0,0351	0,0155	1,1111	4,9731
7	4,6401	5,4419	17,7249	4,2729	0,3560	0,9953	0,0644	0,0132	0,8090	2,7127
8	6,2126	1,1112	8,6283	26,2491	0,2431	0,8885	0,0545	0,0112	0,8647	4,5219
9	3,9506	6,3302	18,0025	19,8750	0,5941	0,9310	0,0296	0,0177	0,6676	2,7273
10	3,1015	5,7429	6,3310	29,6799	0,2186	0,8019	0,0144	0,0103	0,8329	2,0979
11	2,5991	3,1523	5,7428	20,9373	0,1557	0,6881	0,0628	0,0157	0,6385	1,6163
12	5,7186	5,7378	3,5886	6,1864	0,5273	0,6534	0,0686	0,0177	1,0219	0,7508
13	6,2156	5,5795	7,5030	18,4969	0,6169	0,8448	0,0112	0,0073	0,9936	2,1421
14	5,3284	2,1798	4,6210	20,4606	0,6297	0,5729	0,0191	0,0174	0,8901	1,6259
15	4,0822	2,1318	18,8618	13,3126	0,5986	0,9237	0,0541	0,0063	0,8556	1,9903
16	2,6707	6,5572	18,4066	14,9562	0,1548	0,8628	0,0298	0,0077	0,9713	1,3164
17	6,7316	3,9310	7,4990	14,5144	0,1851	0,8481	0,0280	0,0144	1,1490	3,6075
18	6,3072	2,2505	11,4894	11,3351	0,2958	0,8275	0,0355	0,0122	0,7707	2,1507
19	7,5877	5,6170	11,0884	7,0474	0,2843	0,8370	0,0544	0,0122	0,8547	3,6961
20	6,6669	2,5351	9,7696	21,9801	0,4205	0,8111	0,0178	0,0094	0,6373	3,1802
21	7,2075	2,9141	20,5604	11,4237	0,6718	0,3971	0,0289	0,0092	0,9735	3,9716
22	3,1242	6,0697	15,2782	23,9900	0,4943	0,7380	0,0094	0,0062	0,9674	0,5750
23	3,1125	2,5474	10,5762	4,7956	0,5938	0,4693	0,0303	0,0063	0,9244	0,6936
24	2,0047	1,4753	13,8723	24,8896	0,4119	0,5705	0,0263	0,0176	0,6354	3,4723
25	7,3508	3,6286	21,2369	9,1256	0,6849	0,7551	0,0506	0,0166	0,6952	2,6583
26	5,4592	6,0806	5,9715	20,6742	0,2688	0,7125	0,0497	0,0167	0,6074	2,2301
27	2,3065	5,0785	21,6061	18,0529	0,1814	0,7867	0,0342	0,0074	1,1972	2,3062
28	4,4446	6,0081	8,8122	28,0197	0,1907	0,9102	0,0278	0,0133	0,9614	3,2202
29	5,3191	5,8211	21,7696	8,2479	0,2307	0,5362	0,0313	0,0158	0,9119	2,1881
30	3,5271	4,2072	5,1255	10,3479	0,3736	0,4847	0,0167	0,0073	1,0614	2,1176
31	1,2318	4,7246	23,2252	14,1095	0,4409	0,6082	0,0247	0,0060	1,0978	2,3612
32	4,9039	6,8980	3,3008	26,5200	0,5125	0,8710	0,0559	0,0123	0,8326	3,8911

Tableau-A III-3c Identification des jeux de paramètres initiaux pour le problème d'optimisation d'HSAMI à vingt-trois paramètres

HSAMI à 23 paramètres			
N° / JP	(21)	(22)	(23)
1	4,0163	8,4311	1,4920
2	1,9571	11,5400	1,1740
3	4,3066	5,7242	0,9023
4	0,8890	8,7429	1,4427
5	1,2777	9,6709	1,1813
6	1,7073	2,8390	1,1951
7	1,9880	12,1421	1,3382
8	3,2056	7,6964	0,2152
9	4,2749	12,9578	0,1639
10	4,9530	1,7857	1,1455
11	4,2762	5,5958	0,1753
12	0,4343	3,5161	0,9451
13	2,7055	7,6456	0,9598
14	4,2316	6,1698	0,9289
15	4,3940	7,7316	1,0784
16	2,4229	12,4314	0,9300
17	3,2711	5,2427	1,4113
18	0,9706	2,8338	1,0657
19	1,3972	3,8951	1,0754
20	1,3725	3,1183	0,4550
21	1,0034	12,8277	0,9931
22	1,5127	2,7766	0,1876
23	4,7945	8,3327	0,8770
24	4,9575	3,4632	1,4972
25	4,3423	3,2547	0,4131
26	1,1274	12,2891	1,2905
27	3,7813	9,6232	0,4364
28	4,5325	11,9302	0,8457
29	0,5272	7,0506	0,6922
30	4,0200	3,3628	1,3505
31	1,9375	3,6746	0,5937
32	1,2454	10,1299	0,4121

ANNEXE IV

PSEUDO CODE DES ALGORITHMES D'OPTIMISATION SÉLECTIONNÉS

Cette annexe présente le pseudo code des approches de calage sélectionnées pour la résolution des problèmes d'optimisation de cette étude.

Algorithme-A IV-1a « Shuffled Complex Evolution
of the University of the Arizona » (SCEUA)

Étape 0 - INITIALISATION

- Nomdre d'évaluations du modèle : $maxn = 2000$
- Nombre de complexes : $p = 2$
- Nombre de transpositions générés au cours d'une évolution : $\alpha = 1$
- Nombre d'évolutions au sein d'un complexe : $\beta = 2 * length(X_0) + 1$
- Valeurs minimales des paramètres : X_{min} = **[voir Annexe II]**
- Valeurs maximales des paramètres : X_{max} = **[voir Annexe II]**
- Valeur du jeu de paramètres initial : X_0 = **[voir Annexe III]**

Étape 1 - ÉCHANTILLONNAGE

- Définition de la taille de la population de travail : $z = p * (2 * length(X_0) + 1)$
- Définition de la population : la population S' est constitué de la génération aléatoire de $z - 1$ points et du point X_0
- Évaluation de la population : $F(S_i)$ pour $i = 1, \ldots, z$

Étape 2 - CLASSEMENT DES POINTS

- Classement de la population : la population S' est triée selon la valeur de $F(S_i)$, du plus grand au plus petit (supposant que l'on souhaite maximiser le critère objectif)

Étape 3 - PARTIONNEMENT DES POINTS

- Nombre de point au sein d'un complexe : $m = 2 * length(X_0) + 1$
- Rangs alloués au premier complexe : les S_i alloués au complexe p_1 ont les rangs $p*(k - 1) + 1$, où $k = 1, \ldots, m$
- Rangs alloués au deuxième complexe : les S_i alloués au complexe p_2 ont les rangs $p*(k - 1) + 2$, où $k = 1, \ldots, m$

Algorithme-A IV-1b « Shuffled Complex Evolution
of the University of the Arizona » (SCEUA)

Étape 4 - ÉVOLUTIONS DES COMPLEXES

• Pour chaque complexe :	$FOR\ u = 1, \ldots, \beta$
• Construction d'un sous-complexe :	selon une probabilité trapézoïdale, sélection aléatoire de $length(X_0) + 1$ points à partir de p_i, $= G'$
• Évolution d'un sous-complexes :	$FOR\ t = 1, \ldots, \alpha$
• Sélection du plus mauvais points:	sélection du plus mauvais G_{worst},
• Centre de masse :	calcul du centre de masse de G' sans
• Réflexion :	G_{worst}, réflexion de G_{worst} par rapport au centre de masse.
• Vérification des limites :	$IF\ X_{min} \leq G_{reflex} \leq X_{max}$, ok
• Génération d'un point aléatoire si hors limite:	$ELSE$ un point aléatoire est généré qui remplace G_{reflex} dans G' et fin de l'évolution t.
• Évaluation du point réfléchit :	$IF\ F(G_{reflex}) \geq F(G_{worst})$, G_{reflex} remplace G_{worst} dans G'
• Réflexion à mi-distance si échoué:	$ELSE$ on refait une réflexion de G_{worst} mais à mi-chemin entre le centre de masse et G_{worst}
• Évaluation du point semi-réfléchit :	$IF\ F(G_{reflex½}) \geq F(G_{worst})$, $G_{reflex½}$ remplace G_{worst} dans G'
• Génération d'un point aléatoire si échoué:	$ELSE$ un point aléatoire est généré qui remplace $G_{reflex½}$ dans G' et fin de l'évolution t.
	$NEXT\ u$

Étape 5 - REMANIEMENT DES COMPLEXES

• Rassemblement des complexes :	$S' = p_1 + p_2$
• Classement de la population :	la population S' est triée à nouveau selon la valeur de $F(S')$

Étape 6 - VÉRIFICATION DU CRITÈRE D'ARRÊT

• Vérification du critère d'arrêt :	IF nombre d'évaluations $= maxn$, arrêt du processus de calage $ELSE$ retour à l'*Étape 4*

Algorithme-A IV-2 « Dynamically Dimensioned Search » (DDS)

Étape 1 - IDENTIFICATION DES INTRANTS

- Nomdre d'évaluations du modèle : $m = 2000$
- Pas de perturbation : $r = 0,2$
- Valeurs minimales des paramètres : X_{min} = **[voir Annexe II]**
- Valeurs maximales des paramètres : X_{max} = **[voir Annexe II]**
- Valeur du jeu de paramètre initial : X_0 = **[voir Annexe III]**

Étape 2 - ÉVALUATION INITIALE

- Nombre d'évaluations du modèle effectuées : $i = 1$
- Évaluation du jeu de paramètre initial : $F_{best} = F(X_0)$ et $X_{best} = X_0$

Étape 3 - SÉLECTION DES PARAMÈTRES PERTURBÉS

- Calcul de la probabilité de sélection : $P(i) = 1 - (Ln(i) / Ln(m))$
- Sélection des paramètres à perturbés : $FOR\ j = 1, \ldots, length(X_{best})$,
  ajout de X^j_{best} à Y' selon
  la probabilité $P(i)$
  $IF\ Y'$ est vide, on tire au hasard un
  seul paramètre de X_{best}

Étape 4 - PERTURBATION DES PARAMÈTRES

- Perturbation des paramètres sélectionnés : FOR chaque paramètre de Y'
  $X^j_{new} = X^j_{best} + \sigma^j * N(0,1)$
  où $\sigma^j = r*(X_{max} - X_{min})$
- Vérification des limites des bornes :
  (Processus de réflexion)
  $IF\ X^j_{new} < X^j_{min}$,
  $X^j_{new} = X^j_{min} + (X^j_{min} - X^j_{new})$
  $IF\ X^j_{new} > X^j_{max},\ X^j_{new} = X^j_{min}$
  $IF\ X^j_{new} > X^j_{max}$,
  $X^j_{new} = X^j_{max} - (X^j_{new} - X^j_{max})$
  $IF\ X^j_{new} < X^j_{min},\ X^j_{new} = X^j_{max}$

Étape 5 - ÉVALUATION APRÈS PERTURBATION

- Évaluation du nouveau jeu de paramètres : $F(X_{new})$
- Mise à jour du meilleur jeu de paramètres : $IF\ F(X_{new}) \leq F_{best}$,
  $F_{best} = F(X_{new})$ et $X_{best} = X_{new}$

Étape 6 - VÉRIFICATION DU CRITÈRE D'ARRÊT

- Augmentation du nombre d'évaluations effectuées : $i = i + 1$
- Vérification du critère d'arrêt : $IF\ i = m$, arrêt du processus de calage
  $ELSE$ retour à l'***Étape 3***

Algorithme-A IV-3a Logiciel : « Nonlinear Optimization by Mesh Adaptive Direct Search » (NOMAD)

Étape 0 - INITIALISATION

- Nomdre d'évaluations du modèle : $m = 2000$
- Dimensionnement du maillage de l'espace paramétrique : $\Delta_k = 1$ (valeur par défaut)
- Paramètre de mise à jour du dimensionnement du maillage : $\tau = 4$ (valeur par défaut)
- Valeurs minimales des paramètres : X_{min} = **[voir Annexe II]**
- Valeurs maximales des paramètres : X_{max} = **[voir Annexe II]**
- Ensemble de jeux de paramètres initiaux : S'_0 = **[voir Annexe III]**
  (un seul point a été fourni)
- Paramètres du "Latin Hypercube Search" : $p_{_0} = 0$ et $p_{_1} = 1$
- Évaluation des jeux de paramètres initiaux : $F(S'_0)$
- Identification du meilleur point : $F(p_k)$ = meilleur de $F(S'_0)$
  p_k = meilleur de S_0

Étape 0.5 - LATIN HYPERCUBE SEARCH

- Génération d'un point selon la méthode d'échantillonnage LHS : séparation de l'espace paramétrique en $p_{_1}$ sous-espace et génération d'un point X_{LHS} aléatoirement dans chacune des sous-espace
- Mise à jour du meilleur jeu de paramètres : $IF\ F(X_{LHS}) \leq F(p_k)$,
  $F(p_k) = F(X_{LHS})$ et $p_k = X_{LHS}$

Étape 1 - LA RECHERCHE (passer cette étape lors de la première itération)

- Génération de jeux de paramètres aléatoires autour du meilleur point trouvé: $R' = p_k + \Delta_k * Z'$ où Z' représente les directions orthogonales aléatoires non-visitées lors de la sonde
- Évaluation des jeux de paramètres : $F(R')$
- Mise à jour du meilleur jeu de paramètres : $IF\ F(R_j) \leq F(p_k)$,
  $F(p_k) = F(R_j)$ et $p_k = R_j$, aller à l'***Étape 3*** et déclaré l'***Étape 1*** réussie

Étape 2 - LA SONDE

- Génération de jeux de paramètres sur le motif autour du meilleur point trouvé: $L' = p_k + \Delta_k * Z''$ où Z'' représente les $2n$ directions orthogonales sélectionnées où $n = length(X_0)$
- Évaluation des jeux de paramètres : $F(L')$
- Mise à jour du meilleur jeu de paramètres : $IF\ F(L_a) \leq F(p_k)$,
  $F(p_k) = F(L_a)$ et $p_k = L_a$, aller à l'***Étape 3*** et déclaré l'***Étape 2*** réussie

Algorithme-A IV-3b Logiciel : « Nonlinear Optimization by Mesh Adaptive Direct Search » (NOMAD)

Étape 3 - MISE À JOUR DES PARAMÈTRES

• Modification de la dimension du maillage de l'espace paramétrique :	***IF*** l'***Étape 1*** ou ***2*** a été déclaré réussie, $\Delta_{k+1} = \Delta_k * \tau$ ***ELSE*** $\Delta_{k+1} = \Delta_k / \tau$
• Vérification du critère d'arrêt :	***IF*** nombre d'évaluations $= m$, ***and IF*** $\Delta_k \leq \Delta_{k\,min}$ (valeur par défaut), arrêt du processus de calage ***ELSE*** retour à l'***Étape 0.5***

Algorithme-A IV-4 « Algorithme Complètement Aléatoire » (ACA)

Étape 0 - INITIALISATION

• Nomdre d'évaluations du modèle :	$m = 2000$
• Valeurs minimales des paramètres :	X_{min} = **[voir Annexe II]**
• Valeurs maximales des paramètres :	X_{max} = **[voir Annexe II]**
• Valeur du jeu de paramètre initial :	X_0 = **[voir Annexe III]**

Étape 1 - LA RECHERCHE ALÉATOIRE

• Évaluation du jeu de paramètre initial :	$F_{best} = F(X_0)$ et $X_{best} = X_0$
• Nombre d'évaluations du modèle effectuées :	$i = 2$
• Génération de points aléatoires :	***FOR*** $i = 2, \ldots, m$ ***FOR*** $j = 1, \ldots, length(X_0)$ r^j = valeur aléatoire entre 0 et 1 $X^j_i = X^j_{min} + r^j * (X^j_{max} - X^j_{min})$
• Évaluation du jeu de paramètres aléatoire :	$F(X_i)$
• Mise à jour du meilleur jeu de paramètres :	***IF*** $F(X_i) \leq F_{best}$, $F_{best} = F(X_i)$ et $X_{best} = X_i$
• Augmentation du nombre d'évaluations effectuées :	***NEXT i***

Rapport-Gratuit.com

LISTE DE RÉFÉRENCES BIBLIOGRAPHIQUES

Abell, T., Malitsky, Y., Tierney, K. (2012). « Fitness Landscape Based Features for Exploring Black-Box Optimization Problem Structure ». Coll. « IT University Technical Report Series », TR-2012-163. Copenhagen (Denmark) : IT University of Copenhagen, 13p.

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.

Arsenault, R., Poulin, A., Côté, P., Brissette, F. (2013). « A comparison of stochastic optimization algorithms in hydrological model calibration ». *Journal of Hydrologic Engineering*, DOI: 10.1061/(ASCE)HE.1943-5584.0000938. (Nov. 6, 2013)

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

Audet, C., Dang, C.-K, Orban, D. (2010). « Algorithmic Parameter Optimization of the DFO Method with the OPAL Framework ». In *Software Automatic Tuning: From Concepts to State-of-the-Art Results (chapter 15)*, sous la dir. de Naono, K., Teranishi, K., Cavazos, J., Suda, R., p. 255-274. Montréal (Canada) : Springer. DOI: 10.1007/978-1-4419-6935-4_15.

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.

Beven, K.J. (2006). « A manifesto for the equifinality thesis ». *Journal of Hydrology*, Volume 320, p. 18-36. DOI: 10.1016/j.jhydrol.2005.07.007.

Bisson, J.L., Roberge, F. (1983). « Prévisions des apports naturels : Expérience d'Hydro-Québec ». Proc., *Workshop on Flow Prediction, Institute of Electrical and Electronics Engineers (IEEE)*, Toronto.

Rapport-gratuit.com
LE NUMERO 1 MONDIAL DU MÉMOIRES

Boyle, D.P., Gupta, H.V., Sorooshian, S. (2000). « Toward improved calibration of hydrologic models: Combining the strengths of manual and automatic methods ». *Water Resources Research*, Volume 36, n°2, p. 3663-3674. DOI: 10.1029/2000WR900207.

Carpenter, T.M., Georgakakos, K.P (2006). « Intercomparison of lumped versus distribute hydrologic model ensemble simulations on operational forecast scales ». *Journal of Hydrology*, Volume 329, n°1, p. 174-185. DOI: 10.1016/j.jhydrol.2006.02.013.

Centre d'expertise hydrique du Québec (2003). « Le Centre ». In *Le site du Centre d'expertise hydrique du Québec*. En ligne. < https://www.cehq.gouv.qc.ca/mission/index.htm >. Consulté le 12 juin 2014.

Clarke, R.T. (1973). « A review of some mathematical models used in hydrology, with observations on their calibration and use ». *Journal of Hydrology*, Volume 19, n°1, p. 1-20. DOI: 10.1016/0022-1694(73)90089-9.

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.

Dolan, E.D., Moré, J.J. (2002). « Benchmarking optimization software with performance profiles ». *Mathematical Programming*, Volume 91, n°2, p. 201-213. DOI: 10.1007/s101070100263.

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.

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.

Fortin, V. (2000). Le modèle météo apport HSAMI : historique, théorie et application. Institut de Recherche d'Hydro-Québec (IREQ), Varennes. 68 p.

Freer, J.E., Beven, K.J. (2001). « Equifinality, data assimilation, and uncertainty estimation in mechanistic modelling of complex environmental systems using the GLUE methodology ». *Journal of Hydrology*, Volume 249, n°1-4, p. 11-29. DOI: 10.1016/S0022-1694(01)00421-8.

Griewank, A.O. (1981). « Generalized Descent for Global Optimization ». *Journal of Optimization Theory and Applications*, Volume 34, p. 11-39.

Hillier, F.S., Lieberman, G.J. (2010). *Introduction to operations research*, 9th ed. New-York (NY) : McGraw-Hill, 1047 p.

Hutchinson, M.F. (1989). « A new objective method for spatial interpolation of meteorological variables from irregular networks applied to the estimation of monthly mean solar radiation, temperature, precipitation and windrun ». *CSIRO Division of Water Ressources Technical Memorandum*, volume 89, n°5, p. 95-104.

Hydro-Québec. 2014. « À propos d'Hydro-Québec : Hydro-Québec en bref ». In *Le site d'Hydro-Québec*. En ligne. <http://www.hydroquebec.com/a-propos-hydro-quebec/qui-sommes-nous/hydro-quebec-bref.html >. Consulté le 12 juin 2014.

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.

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.

Malan, K.M., Engelbrecht, A.P. (2013). « A survey of techniques for characterising fitness landscapes and some possible ways forward ». *Information Sciences*, Volume 241, p. 148-163. DOI: 10.1016/j.ins.2013.04.015.

McGill, R., Tuckey, J.W., Larsen, W.A. (1978). « Variations of Box Plots ». *The American Statistician*, Volume 32, n°1, p. 12-16.

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.

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.

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.

Mugunthan, P., Shoemaker, C.A. (2006). « Assessing the impacts of parameter uncertainty for computationally expensive groundwater models ». *Water Resources Research*, Volume 42, n°10, W10428. DOI: 10.1029/2005WR004640.

Mulvany, T.J. (1850). « On the use of self registering rain and flood gauges ». *Institution of Civil Engineering (Dublin)*, Volume 4, p. 1-8.

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

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.

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.

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 preemption ». *Water Resources Research*, Volume 46, W11523. DOI: 10.1029/2009WR008957.

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

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.

Ritter, A., Munoz-Carpena, R. (2013). « Performance evaluation of hydrological models: Statistical significance for reducing subjectivity in goodness-of-fit assessments ». *Journal of Hydrology*, Volume 480, p. 33-45. DOI: 10.1016/j.jhydrol.2012.12.004.

Schwefel, H.P. (1981). *Numerical optimization of computer models*. Chichester, WS, UK: Wiley.

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.

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).

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.

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 ». *Nord. Hydrol.*, Volume 38, n°1, p. 211-234.

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.