

**Low-cost and representative surrogate hydrological models. Part II –Use within calibration frameworks**

P.-L. Huot, A. Poulin,  
C. Audet, S. Alarie

G-2019-08

January 2019

---

La collection *Les Cahiers du GERAD* est constituée des travaux de recherche menés par nos membres. La plupart de ces documents de travail a été soumis à des revues avec comité de révision. Lorsqu'un document est accepté et publié, le pdf original est retiré si c'est nécessaire et un lien vers l'article publié est ajouté.

**Citation suggérée :** P.-L. Huot, A. Poulin, C. Audet, S. Alarie (Janvier 2019). Low-cost and representative surrogate hydrological models. Part II –Use within calibration frameworks, Rapport technique, Les Cahiers du GERAD G-2019-08, GERAD, HEC Montréal, Canada.

**Avant de citer ce rapport technique**, veuillez visiter notre site Web (<https://www.gerad.ca/fr/papers/G-2018-08>) afin de mettre à jour vos données de référence, s'il a été publié dans une revue scientifique.

The series *Les Cahiers du GERAD* consists of working papers carried out by our members. Most of these pre-prints have been submitted to peer-reviewed journals. When accepted and published, if necessary, the original pdf is removed and a link to the published article is added.

**Suggested citation:** P.-L. Huot, A. Poulin, C. Audet, S. Alarie (January 2019). Low-cost and representative surrogate hydrological models. Part II –Use within calibration frameworks, Technical report, Les Cahiers du GERAD G-2019-08, GERAD, HEC Montréal, Canada.

**Before citing this technical report**, please visit our website (<https://www.gerad.ca/en/papers/G-2019-08>) to update your reference data, if it has been published in a scientific journal.

---

La publication de ces rapports de recherche est rendue possible grâce au soutien de HEC Montréal, Polytechnique Montréal, Université McGill, Université du Québec à Montréal, ainsi que du Fonds de recherche du Québec – Nature et technologies.

Dépôt légal – Bibliothèque et Archives nationales du Québec, 2019  
– Bibliothèque et Archives Canada, 2019

The publication of these research reports is made possible thanks to the support of HEC Montréal, Polytechnique Montréal, McGill University, Université du Québec à Montréal, as well as the Fonds de recherche du Québec – Nature et technologies.

Legal deposit – Bibliothèque et Archives nationales du Québec, 2019  
– Library and Archives Canada, 2019



# Low-cost and representative surrogate hydrological models. Part II

## –Use within calibration frameworks

Pierre-Luc Huot <sup>a</sup>

Annie Poulin <sup>a</sup>

Charles Audet <sup>b,c</sup>

Stéphane Alarie <sup>b,d</sup>

<sup>a</sup> Construction Engineering, École de Technologie Supérieure, Montréal (Québec) Canada, H3C 1K3

<sup>b</sup> GERAD, Montréal (Québec), Canada, H3T 2A7

<sup>c</sup> Department of Mathematics and Industrial Engineering, Polytechnique Montréal, Montréal (Québec) Canada, H3C 3A7

<sup>d</sup> Institut de recherche d'Hydro-Québec, Varennes (Québec) Canada, J3X 1S1

pierre-luc.huot.1@ens.etsmtl.ca

annie.poulin@etsmtl.ca

charles.audet@gerad.ca

alarie.stephane@ireq.ca

January 2019

Les Cahiers du GERAD

G–2019–08

Copyright © 2019 GERAD, Huot, Poulin, Audet, Alarie

Les textes publiés dans la série des rapports de recherche *Les Cahiers du GERAD* n'engagent que la responsabilité de leurs auteurs. Les auteurs conservent leur droit d'auteur et leurs droits moraux sur leurs publications et les utilisateurs s'engagent à reconnaître et respecter les exigences légales associées à ces droits. Ainsi, les utilisateurs:

- Peuvent télécharger et imprimer une copie de toute publication du portail public aux fins d'étude ou de recherche privée;
- Ne peuvent pas distribuer le matériel ou l'utiliser pour une activité à but lucratif ou pour un gain commercial;
- Peuvent distribuer gratuitement l'URL identifiant la publication.

Si vous pensez que ce document enfreint le droit d'auteur, contactez-nous en fournissant des détails. Nous supprimerons immédiatement l'accès au travail et enquêterons sur votre demande.

The authors are exclusively responsible for the content of their research papers published in the series *Les Cahiers du GERAD*. Copyright and moral rights for the publications are retained by the authors and the users must commit themselves to recognize and abide the legal requirements associated with these rights. Thus, users:

- May download and print one copy of any publication from the public portal for the purpose of private study or research;
- May not further distribute the material or use it for any profit-making activity or commercial gain;
- May freely distribute the URL identifying the publication.

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

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

**Keywords:** Distributed hydrological model, computationally-intensive simulation model, efficient calibration, surrogate model, reduced-fidelity model, representativeness

---

**Acknowledgments:** The authors would like to sincerely thank provider of the sources codes of the DDS algorithm, Bryan A. Tolson, and the precious help from Christophe Tribes with the NOMAD software. The meteorological gridded datasets and daily observed streamflows for studied watersheds have been provided by the *Direction de l'Expertise Hydrique* (DEH) and *Hydro-Québec*. Funding for this study has been provided by a PhD scholarship from National Sciences and Engineering Research Council (NSERC) of Canada.

# 1 Introduction and objectives

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

Part I of this research studies the construction of various low-cost and representative surrogate models, from the family of reduced-fidelity physical-based models. The hydrological model used for the exploration of different approaches of reduced-fidelity models is the HYDROTEL model (Fortin et al., 2001a-b), a distributed and computationally-intensive model. Two versions of the HYDROTEL model with 10 and 19 parameters are used, called hereafter HYDROTEL 10 and HYDROTEL 19, respectively. Three watersheds from the province of Québec are modeled, namely Cowansville, Ceizur and Toulmoustou watersheds (see Section 3 in Part I). The six combinations of “Watershed-HYDROTEL problem” define the calibration problems used as computational experiments in this study. Part I demonstrates the potential of using reduced-fidelity models by combining three approaches (further described in Section 3.1. Results show significant reduction of computational time and high-levels of representativeness between final combined reduced-fidelity surrogates and original models. These two characteristics suggest the potential to efficiently use them within hydrological model calibration processes.

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

## 1.1 Literature review

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

The literature on response surface surrogates is rich and encompasses a wide range of application fields. Popular response surface surrogates include Gaussian kriging techniques (Krige, 1951; Lophaven et al., 2002), polynomial expressions (Hussain et al., 2002; Lophaven et al., 2002; Fen et al., 2009), Radial Basis Functions (RBF, Hussain et al., 2002; Mugunthan et al., 2005; Regis and Shoemaker, 2007), Artificial Neural Networks (ANNs, Papadrakakis et al., 1998) and Support Vector Machines (SVMs, Zhang et al., 2009). Some papers present more exhaustive overviews of different types of surrogate models (e.g. Queipo et al., 2005; Forrester and Keane, 2009; Razavi et al., 2012a).

There are several uses of adaptive functions that deal with computationally intensive problems including hydrological or environmental modelling. Booker et al. (1999) propose a surrogate management framework using kriging or polynomial models for optimization processes solving design problems for helicopter rotor blades. Recently, Audet et al. (2018) proposed an algorithmic framework to combine various types of surrogates. Closer

to the scope of this study, Regis and Shoemaker (2004, 2007, 2009, 2013) make extensive use of response surface surrogates such as RBFs in hydrological modelling and underground water restoration processes. They improve the final values of the objective functions while reducing the computational time. Khu et al. (2004) demonstrate that the use of ANNs inside the optimization process applied to hydrological models reduces the computational time by up to 40% in comparison with a standard metaheuristics optimization algorithm. Razavi et al. (2012b) published a comparative study of three different adaptive functions; RBFs, ANNs and Kriging, which indicates that the use of adaptive functions may decrease the CPU time compared to existing optimization algorithms. Wang et al. (2014) identified Gaussian kriging as the best adaptive function to significantly reduce the computational time of a conceptual rainfall-runoff model. Finally, some existing algorithms integrate adaptive functions as sequencing strategy in order to provide better optimization efficiency. The MADS algorithm (Audet and Dennis, 2006) is adapted to exploit a quadratic polynomial model to locally represent the parametric space by estimating the objective function values of future candidates and sequencing those that will be evaluated in the optimization process according to the improvement potential (Conn and Le Digabel, 2013). All studies cited above prove that CPU time gains may be obtained by using response surface surrogates within an efficient calibration framework or in sensitivity analysis applications. Many other studies also support this fact (Marsden et al., 2004; Neelin et al., 2010; Mousavi and Shourian, 2010; Castelletti et al., 2010; Yazdi and Salehi Neyshabouri, 2014; Brunetti et al., 2017; Dan Lu et al., 2017; Willers Moore et al., 2016).

Reduced-fidelity models are a priori design models which cannot be updated nor improved during the optimization process. They represent a simplification of the original simulation model while maintaining representativeness and reducing the computational time as much as possible (Regis and Shoemaker, 2007; Castelletti et al., 2012; Razavi et al., 2012a, Razavi and Tolson, 2013; Wang et al., 2014). This type of surrogates is commonly used to pre-evaluate the potential of some optimized parameter sets before running a simulation with computationally-intensive model. Reduced-fidelity models with high level of representativeness may be used during the entire calibration process, or partially. Razavi and Tolson (2013) illustrate the case where early in the optimization process, improvements of the objective function value can be achieved by using the surrogate model only, which generally leads to very similar results as when the original model is used. But at some point during the optimization process, no further improvement is obtained with the surrogate model due to a non-perfect representation of the original model. Therefore, an important challenge is to correctly manage the use of the reduced-fidelity surrogate versus the original model during optimization.

Some papers explore reduced-fidelity models in the specific context of hydrological modelling. Razavi and Tolson (2013) demonstrate that a reduced-fidelity hydrological model, characterized by a shorter but carefully selected simulation time-period, used within an efficient calibration framework, results in interesting gains on objective function values relatively to a standard optimization algorithm, both limited to the same computational time budget. Haghnegahdar et al. (2015) study various levels of spatial discretization (modelling scales) on different computational time budgets to evaluate the impact on final values of the objective function, but do not focus on reduced-fidelity models within calibration processes. Other studies propose strategies for the implementation of reduced-fidelity surrogates within efficient calibration frameworks outside of the hydrological modelling domain (Polak and Wetter, 2006; Castelletti et al., 2012; Koziel and Leifsson, 2013; Leifsson et al., 2014; Huang et al., 2015).

## 1.2 Paper organization

This paper is organized as follows. Section 2 exposes the two benchmark algorithms and the calibration framework used to exploit surrogate models. Section 3 presents five calibration frameworks exploiting surrogate models and they are applied to six different “Watershed-HYDROTEL problem” combinations. Performance of calibration frameworks is evaluated in terms of computational time and final objective function values. Discussion and future works conclude this paper.

## 2 Benchmark algorithms and calibration framework

Previous works of Huot et al. (2017) study the evaluation of existing calibration algorithms and develop an efficient calibration approach for computationally-intensive hydrological problems. Two different calibration algorithms serve as benchmark algorithms in this present study: the DDS algorithm (Tolson and Shoemaker, 2007) and the

hybrid DDS-MADS calibration approach (Huot et al., 2017). The hybrid DDS-MADS approach exploits surrogate models developed in this two-part research study. Brief descriptions of these two algorithms are presented below.

## 2.1 The dynamically dimensioned search (DDS) algorithm

DDS is a global search heuristic optimization algorithm that evolves the current best single solution. All DDS internal parameters have default values, except for the allowable number of calls to the simulation model representing the total (maximum) simulation budget. Given a user-specified simulation budget, this algorithm dynamically adjusts the search from global to local by gradually reducing the number of perturbed model parameters (search dimensions of the problem). DDS starts by exploring the entire space of variables, then gradually focuses on a series of subspaces, until it reaches a one-dimensional subspace. DDS always exploits the totality of the user-specified simulation budget (the stopping criteria), i.e. it does not base the optimization search strategies on quality of current best solution. For the calibration of computationally-intensive hydrological models, the user needs to carefully select the simulation budget in order to control the overall computational time. Studies on hydrological modelling (Razavi et al. 2010, Arsenault et al. 2014, Huang et al., 2014; Huot et al., 2017) report that the global search strategies used by DDS converge more rapidly to good-quality solutions in comparison with some other existing algorithms.

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

## 2.2 The hybrid DDS-MADS optimization approach

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

The MADS algorithm is a direct search method that discretizes the solution space into a grid mesh. It performs an adaptive search on this mesh; i.e. the mesh becomes coarser when the current best solution is improved and is refined when the local search around the current best solution fails to identify a better solution. The MADS algorithm is designed to automatically terminate the calibration as soon as some local optimality conditions are satisfied, guaranteeing that the final solution satisfies local optimality conditions. Moreover, the MADS algorithm is implemented in the NOMAD software (Le Digabel, 2011) with many functionalities including a response surface function integrated to the direct search method. This response surface function is a quadratic model and is used to order the  $2N$  (where  $N$  is the number of calibrated parameters) trial points on the mesh according to orthogonal directions from the current best solution in a new iteration. The quadratic model is an adaptive function constructed with past-evaluated trial points that influence the local domain around the current best solution point. Best trial point performances in the quadratic model are simulated first to avoid several computationally-intensive simulations. Conn and Le Digabel (2013) demonstrate that the use of quadratic models in the trial points ordering improves the performance of the MADS algorithm significantly by reducing the total number of objective function evaluations. Although this paper does not specifically examine response surface surrogate, the MADS algorithm within the DDS-MADS approach use this functionality in its optimization process. For more details about the optimization strategies and/or functionalities of the MADS algorithm, please see the relevant papers (Audet and Dennis, 2006; Abramson et al., 2009; Conn and Le Digabel, 2013; Audet et al., 2016).

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

the MADS step is set according to a spatial variability analysis from good-quality points registered in the cache file of the DDS step; in other words, the position in the solution space of all good-quality points serve to adjust independently each search dimension of the mesh. The total user-specified simulation budget is first attributed to the DDS step. When the stagnation parameter interrupts the DDS step, the remaining simulation budget is transferred to the MADS step. However, the MADS step does not necessarily consume all the remaining simulation budget depending on the necessary number of simulations to identify a nearby local optimum. Finally, numerical noise in the objective function surface may cause a continuous adaptation of the mesh, delaying the local convergence process to a local optimum from the MADS step. Rounding the objective function value to an appropriated number of digits (depending on the objectives pursued by the calibration process) attenuates this numerical noise and stabilizes the performance of the algorithm.

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

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

### 3 Use of surrogate models

A review of surrogate highlights obtained with the reduced-fidelity models developed in Part I of this study is first presented. The objective of this review is to guide the implementation of potential calibration frameworks. Next, five calibration frameworks exploiting surrogate models are defined and tested. All combinations of “Watershed-HYDROTEL problem” are used for numerical testing. Comparisons between calibration frameworks and/or benchmark algorithms according to the computational time reductions and quality of final objective function values conclude this section.

#### 3.1 Review of reduced-fidelity models highlights

Part I of this study focused on the development of reduced-fidelity models which present interesting characteristics for usage them within calibration frameworks: low-cost and high representativeness. The CPU time ratio (ratio between the CPU times of the original and surrogate models) was used to evaluate the potential of computational time reductions of reduced-fidelity models. A terminology regarding the levels of CPU time ratio has been developed and the representativeness was evaluated according to the Spearman’s rank correlation coefficient ( $R_s$ ) and R-Square coefficient ( $R^2$ ) calculated from the objective function  $1 - NSE$  obtained with original and surrogate models. The reader is invited to consult Section 2.2 of Part I for more details. Three types of reduced-fidelity models and their combination have been evaluated: (1) reduction of the number of “pseudo-meteorological” stations on the territory, (2) the reduction of the calibration time-period and (3) the reduction of the level of spatial discretization by decreasing the number of simulation units, referring to RHHUs. Two main results from Part I are highlighted herein.

First, all three types of reduced-fidelity models provide generally high  $R_s$  and  $R^2$  coefficients. However, their individual use is not recommended either because of the poor CPU time ratios that were obtained (refer to Section 2.2 from Part I), or because of the very low representativeness coefficients that were obtained in some instances (although interesting time ratios were observed in some of these poorly representative cases).

Second, the combination of the three types of reduced-fidelity models were evaluated in Part I. The final reduced-fidelity models retained were the combination of all three types of reduced-fidelity models for the HYDROTEL 10 problem. For the HYDROTEL 19 problem, it was the merge of the reduction of the number of meteorological stations and the reduction of the calibration time-period (no modification to spatial discretization). These final



reduced-fidelity models are potentially usable within calibration frameworks since they provide the best compromise between high representativeness coefficients and a decrease in CPU time. For the HYDROTEL 10 problem, the final reduced-fidelity models provided representativeness coefficients over or close to the 0.9 value for the Ceizur and Touloustouc watersheds as recommended by Toal (2015). The reduced-fidelity models for the the Cowansville watershed appeared to be less representative with  $R_s$  of 0.694 and  $R^2$  of 0.475. However, all three modelled watersheds with HYDROTEL 10 provided ideal CPU time ratios, being 16 to 44 times less computationally-intensive than the original models. For the HYDROTEL 19 problem, similar results are obtained for the levels of representativeness (over or close to 0.9 for Ceizur and Touloustouc watersheds with smaller values for the Cowansville watershed). Nonetheless, the final reduced-fidelity models developed for HYDROTEL 19 are all approximately 2 times less computationally-intensive; i.e. resulting as poor CPU time ratios.

## 3.2 Calibration framework experiments

Five different calibration frameworks exploiting final reduced-fidelity models and response surface functions within the hybrid DDS-MADS optimization approach are tested. As mentioned in Section 2.2, the DDS-MADS approach is a two-step method; DDS and MADS steps focusing on global and local searches respectively. This approach provides many possibilities to manage surrogate models for calibration problems due to this clear division of global/local searches and intrinsic characteristics of each algorithm. Tables 1 and 2 summarized the three (**A**, **B** and **C**) and five (**A** to **E**) calibration frameworks tested on the original HYDROTEL 10 and 19 problems respectively. A description of each calibration framework is presented next.

### Calibration framework A

The calibration framework **A** allies the global search strategy of the DDS step performing on the final reduced-fidelity models and the local search strategy of the MADS step performing on the original models. The objective is to save computational time during the global search, using low-cost and highly representative reduced-fidelity models, and then to identify optimal solutions on the original models. Given that DDS step performs on reduced-fidelity models and the MADS step on original models; the initial mesh size adjustment cannot be used in calibration framework **A**. As mentioned in Section 2.2, the MADS algorithm use the quadratic model to order the  $2N$  trial points on the mesh. Considering the local search role of the MADS algorithm, the use of quadratic models would result in reducing the total number of computationally-intensive simulations. In the worst situation, all trials points are simulated as would have been the case if the quadratic models had not been exploited. The calibration framework **A** is performed on the original HYDROTEL 10 and 19 problems and is summarized in Tables 1 and 2 under the column (A).

### Calibration framework B

Calibration framework **B** aims to fully exploit the low-cost and the high representativeness of the final reduced-fidelity models (especially for HYDROTEL 10). Thus, the calibration framework **B** is the hybrid DDS-MADS optimization approach performing strictly on final reduced-fidelity models without considering the original models in the calibration process. The initial mesh size adjustment and the quadratic models within the MADS step are used. At the end of the optimization process, only one last simulation is evaluated on the original models to obtain the final objective function value. Moreover, results of the calibration framework **B** may provide more information about the objective function gap between local optima in the final reduced-fidelity models and their counterpart in the original models. The calibration framework **B** is performed on both HYDROTEL versions and is summarized in Tables 1 and 2.

### Calibration framework C

The first step of the calibration framework **C** is the calibration framework **B**; i.e. the hybrid DDS-MADS approach is performed strictly on the final reduced-fidelity models. The solution obtained from this first step is then used as a starting point for the second step, which launches a local search with the MADS algorithm (using quadratic models for trial points ordering) on the original models. The objective is to exploit the calibration framework **B** as the global search strategy and then use the local search features of the MADS algorithm to fill the representativeness

gap between local optimal solutions in the final reduced-fidelity models and their counterpart in the original models. If this representativeness gap is low, the required CPU time to compensate for it with the launch of the MADS algorithm is expected to be low. Conversely, if the representativeness gap is high, the local search is expected to be time-consuming and/or to lead to low-quality solutions. The calibration framework **C** is performed on the HYDROTEL 10 and 19 versions and is summarized in Tables 1 and 2.

**Table 1: Three calibration frameworks exploiting reduced-fidelity models and response surface functions tested on the original HYDROTEL 10 problems.**

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

### Calibration framework D

The calibration framework **D** is performed only on the HYDROTEL 19 version and is summarized in Table 2. This framework exploits a fourth type of reduced-fidelity model; i.e. a reduction of the number of calibration parameters. As mentioned in Section 3.1, the ideal CPU times ratios obtained by final reduced-fidelity models on HYDROTEL 10 problems are much more interesting than the poor CPU time ratios obtained on HYDROTEL 19 problems. Recall that this important gap in terms of CPU time ratios is mainly caused by the production of computationally-intensive geomorphological hydrographs. These hydrographs are produced on every new combination of 2 calibration parameters included in the overland routing of the HYDROTEL 19 version (see Section 2.1 in Part I). The calibration framework **D** is designed to take advantage of this gap in terms of CPU time ratios using the final reduced-fidelity models obtained on HYDROTEL 10 rather than those on HYDROTEL 19. The first step of the calibration framework **D** uses the calibration framework **B** as global search strategy, but performs on the HYDROTEL 10 final reduced-fidelity models (instead of HYDROTEL 19 final reduced-fidelity models). The initial mesh size adjustment and the quadratic models are both used in this DDS-MADS first step. The second step of the framework **D** uses the solution from the first step as a starting point, and then launches a local search with the MADS algorithm (using quadratic models for trial points ordering) on HYDROTEL 19.

### Calibration framework E

The calibration framework **E** is based on same principle as previous framework **D** and is performed only using the HYDROTEL 19 version. As mentioned in the calibration framework **D**, two calibration parameters are mainly responsible of the poor CPU time ratios. To isolate these two time-consuming calibration parameters, the final reduced-fidelity models on the HYDROTEL 19 have been modified to create a 17 calibration parameters version, named hereafter HYDROTEL 17. No experiments have been completed to evaluate the level of representativeness and the CPU time ratios of the reduced-fidelity HYDROTEL 17 models, but this version is expected to yield good performance on CPU time ratio and representativeness. The calibration framework **E** is identical to the calibration framework **D** except that the final reduced-fidelity models on HYDROTEL 10 is replaced by the HYDROTEL 17 version. As for the framework **D**, the main objective is to exploit the final reduced-fidelity models on HYDROTEL 17 which are less computationally-intensive (similar to the HYDROTEL 10 version), but more representative than the HYDROTEL 10 version. Table 2 summarizes the calibration framework **E**.

**Table 2: Five calibration frameworks exploiting reduced-fidelity models and response surface functions tested on the original HYDROTEL 19 problems.**

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

### 3.3 Results from the calibration frameworks on HYDROTEL 10

Figure 1 presents the final  $1 - NSE$  results from 32 calibration trials using calibration frameworks **A** to **C** for the three modeled watersheds with HYDROTEL 10. Calibration frameworks results are also compared with the two benchmark optimization algorithms DDS-MADS-Bench and DDS-Bench as presented in Section 2. The DDS-MADS-Bench algorithm is used for the computational times comparison while the DDS-Bench is rather for the final objective function values comparison. Table 3 presents minima, maxima and averages of computational times (hours) related to Figure 1 for each of calibration frameworks **A** to **C**. The percentage in computational time reductions (%) according to the computational time of DDS-MADS-Bench are also presented in the neighbouring columns. Results for the final objective function values (Figure 1) and related computational times (Table 3) are discussed in parallel for each calibration framework.

#### Calibration framework A

Based on medians and interquartile ranges, Figure 1 shows that calibration framework **A** produces slightly better objective function values than the DDS-MADS-Bench ( $-0.005$ ), and slightly worst values than those of DDS-Bench ( $+0.005$ ) for all three modelled watersheds. Table 3 suggests that on average, framework **A** decreases the computational times compared to those of DDS-MADS-Bench; i.e. ranging from 2% for the Cowansville watershed to 14% for the Ceizur watershed reductions in the computational times. These computational time reductions by the calibration framework **A** are probably associated to the global search strategy of DDS step performing on final reduced-fidelity models. Nonetheless, calibration framework **A** provides similar performances to both benchmark algorithms regarding the final  $1 - NSE$  values, with small computational time reductions.

#### Calibration framework B

Calibration framework **B** offers the worst performance of the three frameworks **A** to **C** regarding the final objective functions values, but the difference between the medians of framework **B** and DDS-Bench is less than  $+0.02$  for the Ceizur and Toulmoustouc watersheds and  $+0.04$  for the Cowansville watershed (Figure 1). Considering that this framework only performed on reduced-fidelity models, it may be confirmed that local optimal solutions of reduced-fidelity models are still good-quality solutions (even if not optimal) in the original models. Although the calibration framework **B** does not present the best final objective function values, Table 3 shows a significant decrease in computational times for all three watersheds. Approximately 30 minutes are needed to complete all calibration

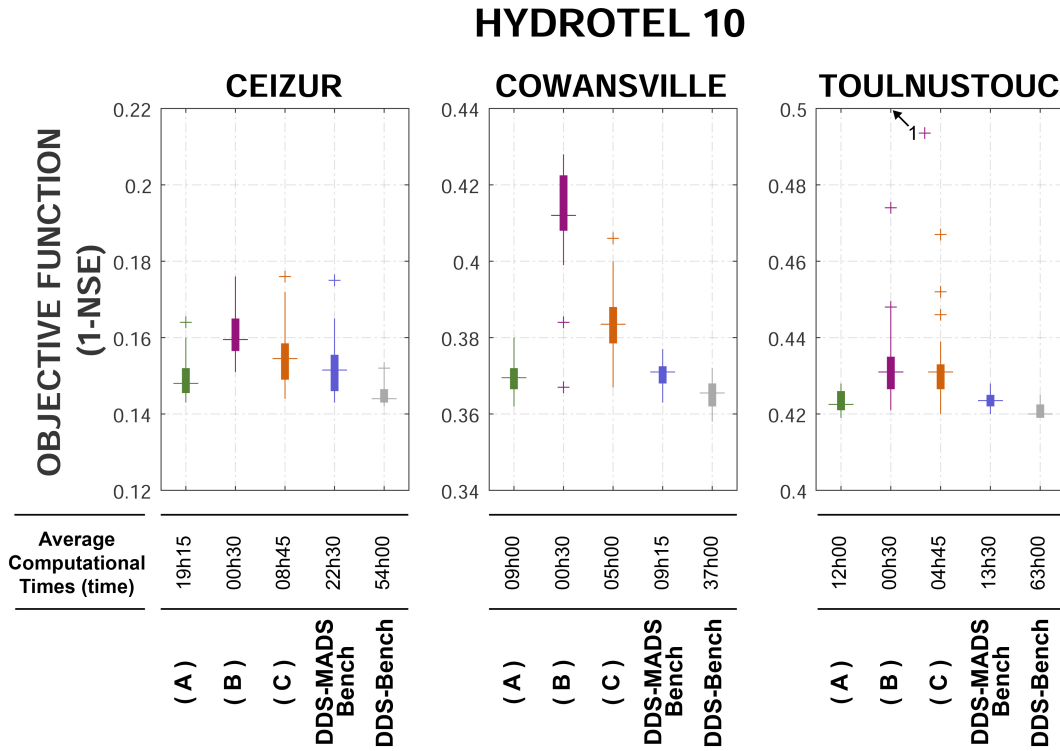


Figure 1: Final  $1 - NSE$  results from calibration frameworks A, B and C (boxplots), and for DDS-MADS-Bench and DDS-Bench. Average computational times are listed for all optimization methods. Calibrations were performed on all three modelled watersheds: Ceizur, Cowansville and Touloustouc with HYDROTEL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

trials with the framework **B** on any modelled watershed, representing reductions of computational time of 94% to 98% in comparison with DDS-MADS-Bench. As the calibration framework **B** only performs on reduced-fidelity models (except for the very last evaluation), these high computational time reductions are explained by the ideal CPU time ratios for all three modelled watersheds (see Section 4.4 in Part I). It is important to keep in mind that framework **B** does not provide a local optimal solution for the original model.

### Calibration framework C

Figure 1 shows that the medians of framework **C** are at a distance of +0.01 of the DDS-Bench  $1 - NSE$  values for the Ceizur and Touloustouc watersheds and at a distance of +0.02 for the Cowansville watershed. These  $1 - NSE$  gaps are greater with calibration framework **A**, and interquartile ranges of framework **C** are also wider. These differences in the results between frameworks A and C still remain very small from a hydrological perspective. As described, calibration framework **C** uses the framework **B** in a first step and then, the MADS algorithm performs a local descent to optimal solutions on the original models. Gaps between final results from frameworks **B** and **C** illustrate the additional labor of the MADS algorithm to reduce the  $1 - NSE$  gaps between framework **B** and benchmark algorithms at the expense of higher computational times. Table 3 shows that framework **C** is 45% (Cowansville watershed) to 64% (Touloustouc watershed) less time-consuming in comparison with the DDS-MADS-Bench. Computational times associated to the labor of the MADS algorithm are illustrated in Table 3 by the difference between the values of framework **B** and framework **C**. Overall, calibration framework **C** offers an intermediate compromise between frameworks **A** and **B**; i.e. higher computational time reductions than framework **A** but slightly lower performance on final  $1 - NSE$  values, and lower computational time reductions than framework **B** but better performance in terms of objective function values.

**Table 3: Minima, maxima and averages of computational times (hours) and computational time reductions (%) for each calibration framework A, B and C in comparison with DDS-MADS-Bench and DDS-Bench. Ceizur, Cowansville and Toulnostouc watersheds are modelled on HYDROTEL 10.**

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

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

### 3.4 Results from the calibration frameworks on HYDROTEL 19

Calibration results of the frameworks **A** to **E** for the three modelled watersheds with HYDROTEL 19 are presented in Figure 2. Comparison with the DDS-MADS-Bench for the computational times and with the DDS-Bench for the final objective function values are discussed in this section. Boxplots in Figure 2 present the results from 32 calibration trials for each combination of “Watershed-Framework” for the HYDROTEL 19 version. Computational times and percentage of computational time reductions related to Figure 2 for each of the frameworks **A** to **E** are presented in Table 4.

#### Calibration framework A

Figure 2 shows that calibration framework **A** performs differently on each modelled watershed. For the Ceizur watershed, the median and interquartile range are similar to those from DDS-MADS-Bench and slightly bigger than those from DDS-Bench (+0.01). A gap of +0.02 of  $1 - NSE$  values between medians of framework **A** and DDS-Bench is shown for the Cowansville watershed, along with similar interquartile range magnitudes. For the Toulnostouc watershed, the interquartile range of framework **A** is much wider than that of DDS-Bench and the median is at a distance of +0.04 of  $1 - NSE$  value from the DDS-Bench median. Table 4 shows that framework **A** is slightly less time-consuming than the DDS-MADS-Bench with averages of 3% (Cowansville watershed) to 9% (Ceizur watershed) of computational time reductions. These reductions are lower than those obtained for the same framework on HYDROTEL 10 (Table 1). This is probably due to poor CPU time ratios obtained by the final reduced-fidelity models on HYDROTEL 19 in comparison with the ideal CPU time ratios on HYDROTEL 10 (as shown in Section 4.4 of Part I).

#### Calibration framework B

Calibration framework **B** presents in Figure 2 a decrease in performance in terms of final objective function values in comparison with all other frameworks. For all three modelled watersheds, interquartile ranges are the widest of all calibration frameworks and medians are at least a distance of +0.05 of  $1 - NSE$  value (Cowansville watershed) from the DDS-Bench median value. These final  $1 - NSE$  values are less interesting than those obtained on HYDROTEL 10 for calibration framework **B**. This is probably due to a lower level of representativeness obtained by reduced-fidelity models on HYDROTEL 19. Yet computational times related to framework **B** are very favorable and provide

on average 54% to 58% reductions for all modelled watersheds (Table 4), although they are smaller than those obtained on HYDROTEL 10 (Table 3). This difference may be explained by the poor CPU times ratios (around 2) for all reduced-fidelity models on HYDROTEL 19. Despite the wider gaps of  $1 - NSE$  in comparison with other calibration frameworks and benchmark algorithms, framework **B** is still the less time-consuming framework for HYDROTEL 19.

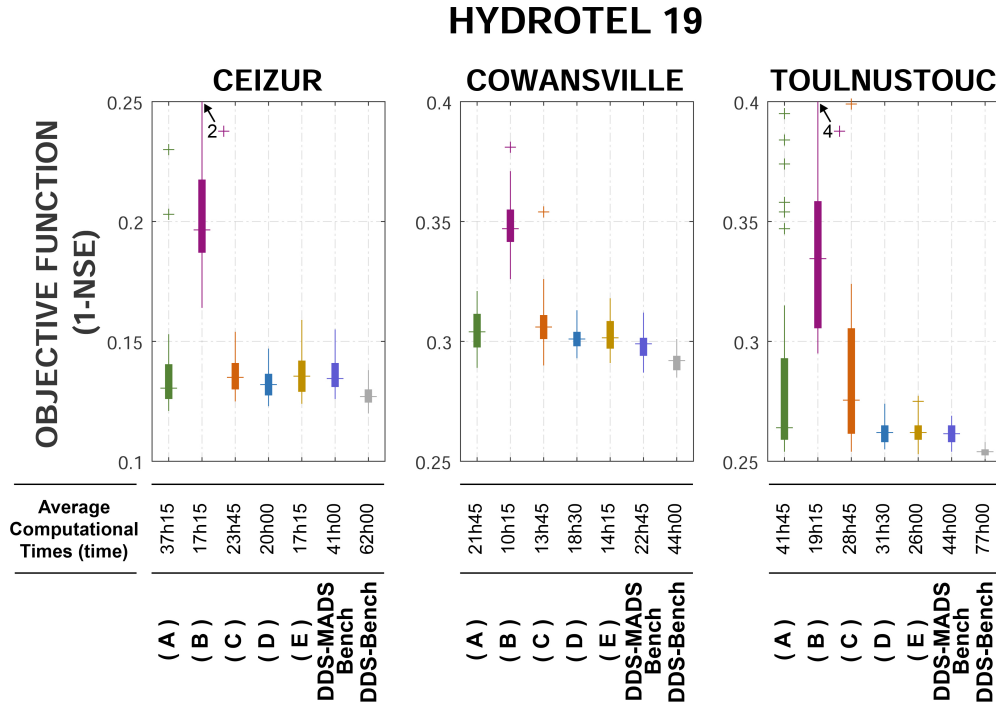


Figure 2: Final  $1 - NSE$  results from calibration frameworks **A** to **E** (boxplots), and for the DDS-MADS-Bench and DDS-Bench. Average computational times are listed for all optimization methods. Calibrations were performed on all three modelled watersheds: Ceizur, Cowansville and Toulnostouc with HYDROTEL 19.

### Calibration framework C

Calibration framework **C** presents an uneven final performance from one watershed to the other regarding the final objective function values. For the Ceizur watershed, Figure 2 shows similar median and interquartile range between framework **C** and DDS-MADS-Bench while a short gap (+0.01 of  $1 - NSE$  value) is observed between medians for the Cowansville watershed. The boxplot for the Toulnostouc watershed shows a much wider interquartile range. Moreover, a gap of +0.08 of  $1 - NSE$  value is obtained between the medians of calibration framework **C** and DDS-Bench for this watershed. Launching the MADS algorithm on the original models (as additional step with respect to framework **B**) labor well for the Ceizur and Cowansville watersheds and results in shrinking the interquartile ranges and increasing medians of the objective function values (first step of framework **C**), but not for the Toulnostouc watershed. It seems that this last watershed poses more difficulties to the local strategies of the MADS algorithm, which ends-up being trapped in poorer-quality zones of the parametric space. Table 4 reports that average computational time reductions are however fairly homogenous, ranging from 35% for Toulnostouc watershed to 42% for the Ceizur watershed. Despite these interesting computational time reductions, calibration framework **C** provides an ambiguous performance in terms of the objective function values.

### Calibration framework D

Figure 2 shows that calibration framework **D** is slightly less performant than DDS-Bench with objective function values of +0.01 of  $1 - NSE$ , but leads to similar results as those obtained with DDS-MADS-Bench for all three modelled watersheds. Given that reduced-fidelity models on HYDROTEL 10 are used instead of HYDROTEL 19,



**Table 4: Minima, maxima and averages of computational times (hours) and computational time reductions (%) for the calibration frameworks A to E in comparison with DDS-MADS-Bench and DDS-Bench. Ceizur, Cowansville and Toulnostouc watersheds are modelled on HYDROTEL 19.**

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

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

interquartile ranges and medians are very close to DDS-MADS-Bench results performing strictly on original HYDROTEL models. This performance indicates that localization of good-quality zones in the parametric space are probably aligned in both HYDROTEL problems for a same watershed. Table 4 shows a reduction in computation times of 18% for the Cowansville watershed and 51% for the Ceizur watershed in comparison with the average computational time of DDS-MADS-Bench. As results suggest, the ideal CPU time ratios obtained with the reduced-fidelity models on HYDROTEL 10 allow a larger decrease in computational time without sacrificing the final  $1 - NSE$  values. Thus, the MADS algorithm (second step of framework D) is able to fill the representativeness gaps between the reduced-fidelity models on HYDROTEL 10 and the original models on HYDROTEL 19.

### Calibration framework E

Calibration framework E performs similarly to calibration framework D, as seen in Figure 2. Boxplots of final  $1 - NSE$  values present similar interquartile ranges and median values are positioned slightly higher, in comparison with DDS-MADS-Bench. DDS-Bench results remain marginally better. As in the case of framework D, the MADS algorithm is able to fill the representativeness gaps between the reduced-fidelity HYDROTEL 17 models and the original HYDROTEL 19 models. Calibration framework E stands out from framework D in the computational time reductions. Table 4 presents average computational time decreases ranging from 38% for the Cowansville watershed to 58% for the Ceizur watershed. Moreover, it is observed that the average computational time from framework E for the Ceizur watershed (17.24 hours) is equivalent to that from the framework B (17.27 hours) which strictly uses the reduced-fidelity HYDROTEL 19 models (except for the last simulation which is conducted on original HYDROTEL 19 models).

## 4 Discussion and conclusion

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

On all HYDROTEL problems, the tested frameworks offer a range of alternatives for the calibration of computationally-intensive hydrological models. Depending on the user's calibration objectives and optimization

constraints, the three calibration frameworks (**A** to **C**) tested on HYDROTEL 10 problems offer different tradeoffs between CPU time and quality of final solutions. If the calibration objectives are focused on the final objective function value, framework **A** is more appropriate with in addition slight computational time reductions. Conversely, calibration framework **B** is the dominant method according to the computational time reduction, but a slight decrease in final solutions quality must be accepted. Calibration framework **C** is in between frameworks **A** and **B** with a more balanced compromise.

In a perspective to decrease as much as possible the computational times, calibration framework **B** remains the less time-consuming framework on average for all three watersheds when tested on HYDROTEL 19 problems. However, computational time reductions are not as high as those obtained on HYDROTEL 10 problems, and final objective function values are much more degraded. Calibration frameworks **D** and **E** represent the best compromise between final objective function value and computational time reduction. Framework **D** presents slightly better performance in terms of the final objective function values than framework **E**, but this latter framework offers a higher percentage of computational time reduction. Either may be chosen depending on user' calibration objectives and optimization constraints. The main strategy behind these frameworks is to use the HYDROTEL 10 and HYDROTEL 17 versions as reduced-fidelity models which both exclude the 2 particularly time-consuming calibration parameters responsible of the production of geomorphological hydrographs. With these 2 isolated calibration parameters, reduced-fidelity models used in these two frameworks combine the advantage of high CPU time ratios and high representativeness. These performances of frameworks **D** and **E** in comparison with frameworks **A** to **C** raise the importance of using reduced-fidelity models with CPU time ratios qualified as ideal (over a factor of 15). Moreover, for all calibration frameworks tested on HYDROTEL 10 and 19 problems, all ideal CPU time ratios lead to significant decreases in the computational times. Conversely, poor CPU time ratios obtained with the reduced-fidelity HYDROTEL 19 models are not interesting enough to justify their implementation even though they can provide high representativeness. This paper proposes, as in Part I, a terminology to qualify the CPU time ratios between reduced-fidelity and original models but further work could study more precisely how different levels of CPU time ratio can impact the computational time reductions.

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

The HYDROTEL model shows particularities in its internal structure which lead to the construction of the final reduced-fidelity models used in this paper. Nevertheless, other computationally-intensive hydrological models present different challenges to achieve the construction of low-cost and representative reduced-fidelity models. They could be study in the same perspective to exploit them within efficient calibration frameworks to decrease the computational times. Moreover, a wider range of modelled watersheds or objective functions bring also the opportunity to extend the conclusions of this study.

## References

- Abramson, M.A., Audet, C., Dennis Jr., J.E., Le Digabel, S. (2009). OrthoMads: A deterministic MADS instance with orthogonal directions. *SIAM Journal on Optimization*, 20(2), 948–966.
- Arsenault, R., Poulin, A., Côté, P., Brissette, F. (2014). Comparison of Stochastic Optimization Algorithms in Hydrological Model Calibration. *Journal of Hydrologic Engineering*, 19(7), 1374–1384. DOI: 10.1061/(ASCE)HE.1943-5584.0000938.
- Audet, C., Dennis Jr., J.E. (2006). Mesh adaptive direct search algorithms for constrained optimisation. *SIAM Journal on Optimization*, 17(1), 188–217. DOI: 10.1137/040603371.
- Audet, C., Le Digabel, S., Tribes, C. (2016). Dynamic scaling in the Mesh Adaptive Direct Search algorithm for blackbox optimization. *Optimization and Engineering*, 17(2), 333–358.
- Audet, C., Kokkolaras, M., Le Digabel, S., Talgorn, B. (2018). Order-based error for managing ensembles of surrogates in mesh adaptive direct search. *Journal of Global Optimization*, Volume 70(3), 645–675. DOI: 10.1007/s10898-017-0574-1.



- Booker, A.J., Dennis, Jr., J.E., Frank, P.D., Serafini, D.B., Torczon, V., Trosset, M.W. (1999). A rigorous framework for optimization of expensive functions by surrogates. *Structural and Multidisciplinary Optimization*, 17(1), 1–13.
- Brunetti, G., Simunek, J., Turco, M., Piro, P. (2017). On the use of surrogate-based modeling for the numerical analysis of Low Impact Development techniques. *Journal of Hydrology*, 548, 263–277. DOI: 10.1016/j.jhydrol.2017.03.013.
- Castelleti, A., Pianosi, F., Soncini-Sessa, R., Antenucci, J.P. (2010). A multiobjective response surface approach for improved water quality planning in lakes and reservoirs. *Water Resources Research*, 46, W06502. DOI: 10.1029/2009WR008389.
- Castelleti, A., Galelli, S., Ratto, M., Soncini-Sessa, R., Young, P.C. (2012). A general framework for Dynamic Emulation Modelling in environmental problems. *Environmental Modelling & Software*, 34, 5–18. DOI: 10.1016/j.envsoft.2012.01.002.
- Conn, A.R., Le Digabel, S. (2013). Use of quadratic models with mesh-adaptive direct search for constrained black box optimization. *Optimization Methods and Software*, 28(1), 139–158. DOI: 10.1080/10556788.2011.623162.
- Dan Lu, M.X., Gui, D., Qi, Z., Zhang, G. (2017). Calibration of an agricultural-hydrological model (RZWQM2) using surrogate global optimization. *Journal of Hydrology*, 544, 456–466. DOI: 10.1016/j.jhydrol.2016.11.051.
- Fen, C.S., Chan, C.C., Cheng, H.C. (2009). Assessing a response surface-based optimization approach for soil vapor extraction system design. *Journal of Water Resources Planning and Management*, 135(3), 198–207. DOI: 10.1061/(ASCE)0733-9496(2009)135:3(198).
- Forrester, A.I.J., Keane, A.J. (2009). Recent advances in surrogate-based optimization. *Progress in Aerospace Sciences*, 45(1), 50–79. DOI: 10.1016/j.paerosci.2008.11.001.
- Fortin, J.P., Turcotte, R., Massicotte, S., Moussa, R., Fitzback, 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, 6(2), 91–99. DOI: 10.1061/(ASCE)1084-0699(2001)6:2(91).
- Fortin, J.P., Turcotte, R., Massicotte, S., Moussa, R., Fitzback, 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, 6(2), 100–108. DOI: 10.1061/(ASCE)1084-0699(2001)6:2(100).
- Haghnegahdar, A., Tolson, B.A., Craig, J.R., Paya, K.T. (2015). Assessing the performance of a semi-distributed hydrological model under various watershed discretization schemes. *Hydrological Processes*, 29, 4018–4031. DOI: 10.1002/hyp.10550.
- Huang, X., Liao, W., Lei, X., Jia, Y., Wang, Y., Wang, X., Jiang, Y., Wang, H. (2014). Parameter optimization of distributed hydrological model with a modified dynamically dimensioned search algorithm. *Environmental Modelling & Software*, 52, 98–110.
- Huang, E., Xu, J., Zhang, S., Chen, C.-H. (2015). Multi-Fidelity Model Integration for Engineering Design. *Procedia Computer Science*, 44, 336–344. DOI: 10.1016/j.procs.2015.03.002.
- Hussain, M.F., Barton, R.R., Joshi, S.B. (2002). Metamodeling: Radial basis functions, versus polynomials. *European Journal of Operational Research*, 138(1), 142–154. DOI: 10.1016/S0377-2217(01)00076-5.
- Huot, P.-L., Poulin, A., Audet, C., Alarie, S. (2017). A Hybrid Optimization Approach for Efficient Calibration of Computationally Intensive Hydrological Models. *Les Cahiers du Gerad*, G-2017-56.
- Krige, D.G. (1951). A statistical approach to some mine valuation and allied problems on the Witwatersrand. Master's thesis of the University of Witwatersrand, 272 p.
- Khu, S.T., Savic, D., Liu, Y., Madsen, H. (2004). A fast Evolutionary-based Meta-Modelling Approach for the Calibration of a Rainfall-Runoff Model. Paper presented at The First Biennial Meeting of the International Environmental Modelling Software Society, Osnabruck, Germany.
- Koziel, S., Leifsson, L. (2013). Multi-level CFD-based Airfoil Shape Optimization with Automated Low-fidelity Model Selection. *Procedia Computer Science*, 18, 889–898. DOI: 10.106/j.procs.2013.05.254.
- Le Digabel, S. (2011). Algorithm 909: NOMAD: Nonlinear Optimization with the MADS algorithm. *ACM Transactions on Mathematical Software*, 37(4), 44:1–44:15. DOI: 10.1145/1916461.1916468.
- Leifsson, L., Koziel, S., Bekasiewicz, A. (2014). Fast Low-Fidelity Wing Aerodynamics Model for Surrogate-Based Shape Optimization. *Procedia Computer Science*, 29, 811–820. DOI: 10.1016/j.procs.2014.05.073.
- Lophaven, S.N., Nielsen, H.B., Sondergaard, J. (2002). DACE – A Matlab Kriging Toolbox, Version 2.0. Report IMM-REP-2002-02, Informatics and mathematical Modelling, DTU. 36 p.

- Marsden, A.L., Wang, M., Dennis Jr., J.E., Moin, P. (2004). Optimal Aeroacoustic Shape Design Using the Surrogate Management Framework. *Optimization and Engineering*, 5(2), 235–262. DOI: 10.1023/B:OPTE.0000033376.89159.65.
- Mousavi, S.J., Shourian, M. (2010). Adaptive sequentially space-filling metamodeling applied in optimal water quantity allocation at basin scale. *Water Resources Research*, 46, W03520. DOI: 10.1029/2008WR007076.
- Mugunthan, P., Shoemaker, C.A., Regis, R.G. (2005). Comparison of function approximation, heuristic, and derivative-based methods for automatic calibration of computationally expensive groundwater bioremediation models. *Water Resources Research*, 41(11), W11427. DOI: 10.1029/2005WR004134.
- Neelin, J.D., Bracco, A., Luo, H., McWilliams, J.C., Meyerson, J.E. (2010). Considerations for parameter optimization and sensitivity in climate models. *Proceedings of the National Academy of Sciences of the United States of America*, 107(50), 21349–21354. DOI: 10.1073/pnas.1015473107.
- Papadrakakis, M., Lagaros, N.D., Tsompanakis, Y. (1998). Structural optimization using evolution strategies and neural networks. *Computer Methods in Applied Mechanics and Engineering*, 156(1–4), 309–333. DOI: 10.1016/S0045-7825(97)00215-6.
- Polak, E., Wetter, M. (2006). Precision Control for Generalized Pattern Search Algorithms with Adaptive Precision Function Evaluations. *SIAM Journal on Optimization*, 16(3), 650–669. DOI: 10.1137/040605527.
- Queipo, N.V., Haftka, R.T., Shyy, W., Goel, T., Vaidyanathan, R., Tucker, P.K. (2005). Surrogate-based analysis and optimization. *Progress in Aerospace Sciences*, 41(1), 1–28. DOI: 10.1016/j.paerosci.2005.02.001.
- Razavi, S., Tolson, B.A., Matott, L.S., Thomson, N.R., Maclean, A., Seglenieks, F.R. (2010). Reducing the computationally cost of automatic calibration through model pre-emption. *Water Resources Research*, 46, W11523. DOI: 10.1029/2009WR008957.
- Razavi, S., Tolson, B.A., Burn, D.H. (2012a). Review of surrogate modeling in water resources. *Water Resources Research*, 48, W07401. DOI: 10.1029/2011WR011527.
- Razavi, S., Tolson, B. A., Burn, D.H. (2012b). Numerical assessment of metamodelling strategies in computationally intensive optimization. *Environmental Modelling & Software*, 34, 67–86. DOI: 10.1016/j.envsoft.2011.09.010.
- Razavi, S., Tolson, B.A. (2013). An efficient framework for hydrologic model calibration on long data periods. *Water Resources Research*, 49, 8418–8431. DOI: 10.1002/2012WR013442.
- Regis, R.G., Shoemaker, C.A. (2004). Local function approximation in evolutionary algorithms for the optimization of costly functions. *IEEE Transaction on Evolutionary Computation*, 8(5), 490–505. DOI: 10.1109/TEVC.2004.835247.
- Regis, R.G., Shoemaker, C. (2007). A Stochastic Radial Basis Function Method for the Global Optimization of Expensive Functions. *INFORMS Journal on Computing*, 19(4), 497–509. DOI: 10.1287/ijoc.1060.0182.
- Regis, R.G., Shoemaker, C.A. (2009). Parallel stochastic global optimization using radial basis functions. *INFORMS - Journal on Computing*, 21(3), 411–426. DOI: 10.1287/ijoc.1090.0325.
- Regis, R.G., Shoemaker, C.A. (2013). A quasi-multistart framework for global optimization of expensive functions using response surface models. *Journal of Global Optimization*, 56(4), 1719–1753. DOI: 10.1007/s10898-012-9940-1.
- Talgorn, B., Audet, C., Le Digabel, S., Kokkolaras, M. (2017). Locally weighted regression models for surrogate-assisted design optimization. *Optimization and Engineering*, 19(1), 213–238. DOI: 10.1007/s11081-017-9370-5.
- Toal, D.J.J., (2015). Some considerations regarding the use of multi-fidelity Kriging in the construction of surrogate models. *Structural and Multidisciplinary Optimization*, 51(6), 1223–1245. DOI: 10.1007/s00158-014-1209-5.
- Tolson, B.A., Shoemaker, C.A. (2007). Dynamically dimensioned search algorithm for computationally efficient watershed model calibration. *Water Resources Research*, 43(1), W01413. DOI: 10.1029/2005WR004723.
- Wang, C., Duan, Q., Gong, W., Ye, A., Di, Z., Miao, C. (2014). An evaluation of adaptive surrogate modeling based on optimization with two benchmark problems. *Environmental Modelling and Software*, 60, 197–179. DOI: 10.1016/j.envsoft.2014.05.026.
- Willers Moore, W.O., Mala-Jetmarova, H., Gebreslassie, M., Tabor, G.R., Belmont, M.R., Savic, D.A. (2016). Comparison of multiple surrogates for 3D CFD model in tidal farm optimisation, *Procedia Engineering*, 154, 1132–1139. DOI: 10.1016/j.proeng.2016.07.523.
- Yazdi, J., Salehi Neyshabouri, S.A.A. (2014). Adaptive surrogate modeling for optimization of flood control detention dams. *Environmental Modelling & Software*, 61, 106–120. DOI: 10.1016/j.envsoft.2014.07.007.
- Zhang, X., Srinivasan, R., Van Liew, M. (2009). Approximating SWAT Model Using Artificial Neural Network and Support Vector Machine. *JAWRA Journal of the American Water Resources Association*, Volume 45(2), 460–474. DOI: 10.1111/j.1752-1688.2009.00302.x.