Les Cahiers du GERAD ISSN: 0711–2440

Low-cost and representative surrogate hydrological models. Part I –Construction of surrogates

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

G–2019–07

January 2019

Citation suggérée : P.-L. Huot, A. Poulin, C. Audet, S. Alarie (Janvier 2019). Low-cost and representative surrogate hydrological models. Part I –Construction of surrogates, Rapport technique, Les Cahiers du GERAD G–2019–07, GERAD, HEC Montréal, Canada.

Suggested citation: P.-L. Huot, A. Poulin, C. Audet, S. Alarie (January 2019). Low-cost and representative surrogate hydrological models. Part I –Construction of surrogates, Technical report, Les Cahiers du GERAD G–2019–07, GERAD, HEC Montréal, Canada.

Avant de citer ce rapport technique, veuillez visiter notre site Web (https://www.gerad.ca/fr/papers/G-2018-07) afin de mettre à jour vos données de référence, s'il a été publié dans une revue scientifique.

Before citing this technical report, please visit our website (https://www.gerad.ca/en/papers/G-2019-07) 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.

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.

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

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

GERAD HEC Montréal
3000, chemin de la Côte-Sainte-Catherine
Montréal (Québec) Canada H3T 2A7

Tél. : 514 340-6053
Téléc. : 514 340-5665
info@gerad.ca
www.gerad.ca

Low-cost and representative surrogate hydrological models. Part I –Construction of surrogates

Pierre-Luc Huot [a]
Annie Poulin [a]
Charles Audet [b,c]
Stéphane Alarie [b,d]

[a] *Construction Engineering, École de Technologie Supérieure, Montréal (Québec) Canada, H3C 1K3*

[b] *GERAD, Montréal (Québec), Canada, H3T 2A7*

[c] *Department of Mathematics and Industrial Engineering, Polytechnique Montréal, Montréal (Québec) Canada, H3C 3A7*

[d] *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–2019–07

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

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

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

1 Introduction

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

Diversification in hydrological model applications has led to the design of several types of models according to their level of spatial discretization (Moradkhani and Sorooshian, 2008). Ranging from a single and global territory entity (global models) to a fully-distributed area into fine mesh elements (distributed models), each spatial discretization level brings its pros and cons depending of its use. Global models require a small amount of input data for modeling and are fast to compute, but the spatial variability in the catchment is ignored. Distributed models often rely on the use of fine-resolution databases from remote sensing or geographic information systems (GIS) to spatialize catchment scale modelling and offer a good variability in the representation of some hydrological processes (Singh and Whoolhiser, 2002; Moradkhani and Sorooshian, 2008; Das et al., 2008; Pechlivanidis et al., 2011). Nonetheless, this advantage brings an important cost in computational time relatively to the chosen spatial discretization level. One simulation requiring several minutes to compute leads to computationally-intensive issue in a full calibration process (Mugunthan et al., 2005; Zhang et al., 2009; Razavi et al., 2010, Huot et al., 2017). This can provide a major impact in operational or research contexts, which forces users to choose computationally efficient calibration methods.

1.1 Litterature review

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

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

Razavi et al. (2012) define surrogate models as a simplification or approximation of a simulation model and having the characteristic of being less intensive in terms of computational time than the original simulation model. Audet and Hare (2017) define a surrogate as a function that shares similarities with the objective function of a simulation model, which is less intensive in computational time and which is useful to exploit for reducing the number of simulation calls within an optimization process.

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

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

1.2 Research objectives and contributions

The present work is part of a larger research project composed of two papers aiming to combine the use of surrogate models inside more efficient optimization strategies to deal with the computationally-intensive calibration process of a hydrological model. The first part studies the development of reduced-fidelity models from the hydrological model HYDROTEL (Fortin et al., 2001a). This paper explores, evaluates and combines three different methods to construct surrogate models from the original computationally-intensive hydrological model HYDROTEL. The first type of surrogate model reduces the number of meteorological gridpoints from a gridded dataset on the territory of the modeled watershed, the second type reduces the calibration time-period and the third one reduces the watersheds spatial discretization by decreasing the number of simulation units within the modelling. Representativeness and computational time for each type of surrogate model and for the combination of all of them are evaluated and analysed. The second part of this project consist of a paper that presents the implementation of reduced-fidelity surrogate models within the hybrid optimization approach DDS-MADS (Huot et al., 2017), which has proven to be efficient for the calibration of HYDROTEL.

1.3 Paper organization

The present paper is organized as follows. Section 2 presents the hydrological model HYDROTEL and its particular model structure, followed by the method used to evaluate the representativeness and the computational time consumed of the surrogate models. Section 3 exposes the hydrological modelling methodology: modeled watersheds, computational time benchmark, necessary input data and objective function. Section 4 details the computational experiments evaluating the representativeness and computational time for each type of reduced-fidelity surrogate models and all combined. Connections with the follow-up paper are discussed in the closing section.

2 The hydrotel problems

This project considers the distributed, physically-based and computationally intensive hydrological model, named HYDROTEL, as the main model to be calibrated (Fortin et al., 2001a, b). This model is based on a modular approach with various submodel options, and is compatible with remote sensing and/or GIS data. The spatial variability of land cover, land use, natural hydrographic system, soil type and topography is distributed on the territory into a gridded mesh at a low distance resolution. Cells of this gridded mesh are rallied in several simulation units named RHHUs (Relatively Homogenous Hydrological Units), similarly to a division of the territory into many sub-basins. All hydrological processes are simulated independently on each RHHU excepting the channel water routing on existing hydrological system. In a first paper, Fortin et al. (2001a) describes mathematical equations governing each hydrological submodels available, as much as possible physically-based, but also conceptually or empirically in nature. A second paper gives a step-by-step watershed modelling example (Fortin et al., 2001b).

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

2.1 Model structure

The HYDROTEL modular structure provides flexibility on simulating various available submodels and allows the addition of new ones by the user. Hydrological submodels used in this study relate to the three computing steps presented in Table 1. The HYDROTEL structure sequences six submodels simulating the hydrological processes: (1) interpolation of meteorological data; (2) snow cover estimation; (3) potential evapotranspiration; (4) vertical water balance; (5) overland water routing and (6) channel water routing on the existing hydrographic system. The first four submodels are simulated independently on each RHHU and represent the first computing step of the hydrological model routine. The second step consists of the overland water routing simulated on each RHHUs towards it river reach associated. The third one calculates the channel water routing on the entire existing hydrographic system. The three sequential computing steps provide the opportunity to re-simulate the following computing step without having to simulate again the previous one when no calibration parameter was modified.

Table 1: Submodels used in each computing steps of the HYDROTEL routine and impact on computational time consumed when using the proposed surrogate models.

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

A particularity in the overland water routing (step 2) requires special attention. First, the production of a specific geomorphological hydrograph based on the routing of a reference water depth over all cells in the flow direction structure of each RHHU is necessary (shaded section in Table 1, step 2). Then, the water depths available, at each time step, are distributed according to these hydrograph patterns to provide the lateral inflow to the hydrographic reach to which the RHHU is associated. The production of the geomorphological hydrograph is computationally-intensive especially because it is independently calculated on each RHHU. The creators of HYDROTEL established a file storage procedure to decrease this time-consuming step. On every simulation, the HYDROTEL routine chooses the appropriate RHHU hydrograph file according to the 4 related calibration parameters of the overland water routing (see the Appendix I for their definition). If a new combination of these related calibration parameters is to be simulated, a new RHHU geomorphological hydrograph is calculated for each RHHU and they are all stored in a single new file. In the HYDROTEL 10 problem, none of these 4 related calibration parameters are included, but 2 are included in the HYDROTEL 19 problem. So, a geomorphological hydrograph for each RHHU will be produced only on the first simulation of HYDROTEL 10. New hydrographs are produced on each new combination of these 2 related calibration parameters for the HYDROTEL 19 problem. This situation can generate a 5% to 25% CPU time increase for a simulation with HYDROTEL 19 versus HYDROTEL 10.

2.2 Representativeness of surrogates versus computational time

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

Representativeness of surrogates

The representativeness of forthcoming surrogate models will be assessed by simulating 5,000 uniformly distributed parameter sets in the parametric space generated with the latin hypercube sampling (McKay et al., 1979), comparing both objective function values of the original and surrogate models. As for Razavi and Tolson (2013), the R-Square coefficient (R^2) is commonly used for qualifying the level of representativeness since it provides information on the goodness of fit between the original and the surrogate models. As guideline, Toal (2015) suggests to use surrogates that have a reasonably high R^2 (greater than 0.9). It is important to note that is a guideline for response surface function surrogates, thus not directly specified for reduced-fidelity models. However, as can be seen in Figure 1, a high R^2 value does not guarantee a high representativeness near the optimal zones. Overconfidence in the R^2 obtained by sampling 5,000 points from the parametric space could lead to a surrogate that quickly becomes useless in a calibration process, particularly if the optimal zones are not well represented. Figure 1 presents two hypothetical cases in which the surrogate function A gives a better R^2 coefficient than the surrogate B. Despite this, important gaps are observed between the different optima on surrogate A. It would be preferable to use the surrogate B in a calibration process since the optima of the surrogate and original functions are aligned. The surrogate model must share similarities with the original model meaning that the behavior of the surrogate model must represent as much as possible the behavior of the original model, including the alignment and the ordering of the optimal zones. Hence, sharing similarities does not mean that the objective function values obtained on the surrogate versus the original must be as similar as possible.

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

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

Computational time

Each surrogate model designed is simulated 1,000 times. The average computational time reduces variations in computer performance and computational time related to the production of geomorphological hydrographs. Obviously, the lower the computational time is, the more interesting the surrogate model is with respect to a full calibration process. The ratio between the CPU time of the original and surrogate models, named

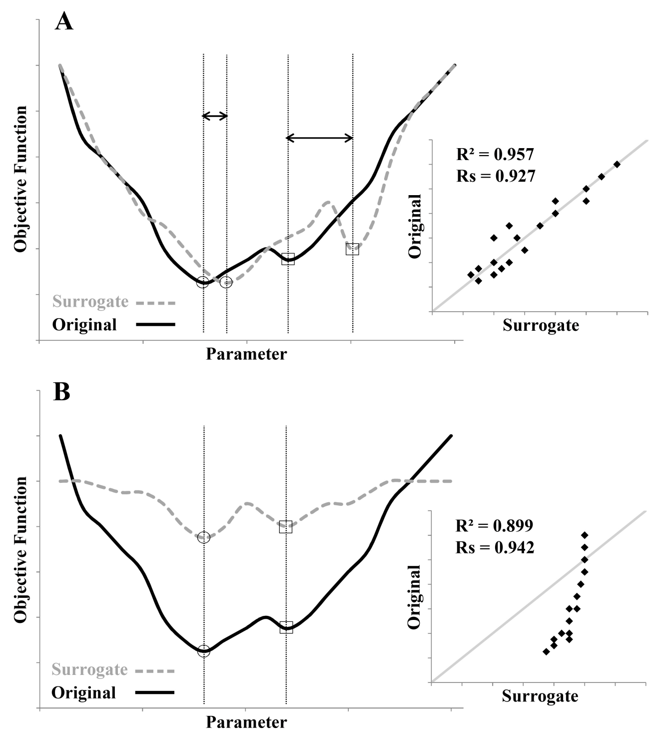

Figure 1: Representativeness of two hypothetical cases illustrating original versus surrogate functions (A & B) with R-Square coefficient (R^2) and Spearman's rank correlation coefficient (Rs).

hereafter the CPU time ratio, has a major impact on the potential for the use of each surrogate model. Higher the CPU time ratio is, more surrogate simulations can be rooted without weighing down the calibration process. Razavi and Tolson (2013) propose a hydrological model calibration framework exploiting short but representative data timeseries to calibrate a problem for which longer data timeseries are available. The CPU time ratio between short data timeseries model (surrogate model) and long data timeseries model (original model) may be evaluated between 6 to 8 times less intensive in computational time. This level of CPU time ratio conduct to more efficient calibration tools.

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

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

2.3 Potential surrogate models

We propose three potential simplifications to design surrogate models: (1) reducing the input meteorological dataset (i.e. reducing the spatial resolution of the gridded dataset) to be interpolated by HYDROTEL on the watersheds; (2) using a shorter calibration time-period; (3) decreasing the spatial discretization by reducing the number of RHHUs (sub-basins). Individual and combinations of these simplifications are also assessed to obtain the best CPU time ratio available while maintaining a high level of representativeness.

Number of meteorological stations

The HYDROTEL problems are designed with a meteorological gridded dataset at a low distance resolution, which results in a large number of “pseudo-meteorological” stations (hereafter simply called stations or meteorological stations) on the territory. In operational or research contexts, this of course can contribute to an adequate simulation of the streamflows, but for calibration purposes, the surrogate model does not need the same accuracy. Table 1 shows that reducing the number of meteorological stations on the territory provides a decrease in computational time only in the interpolation of the meteorological data submodel. Therefore, one could expect a poor or medium CPU time ratio. A checkerboard pattern is used to withdraw meteorological stations in order to maintain a good distribution of the dataset on the territory and to simplify the design of this surrogate model type.

Calibration time-period

Obtaining gain in computational time with a shorter calibration time-period is not surprising. However, a high level of representativeness may be more difficult to reach considering the interannual variability of the hydrologic cycle. Initial conditions on the watershed of a short calibration time-period have also an important impact on representativeness, so these initial conditions need to be adequately set on every surrogate models. On the HYDROTEL problems, running a short calibration time-period should decrease the computational time for all submodels resulting in a medium or superior CPU time ratio (see Table 1). For this type of surrogate model, the representativeness level is more challenging than obtaining interesting CPU time ratios.

Number of RHHUs

In the HYDROTEL problems, the division of the territory into several RHHUs is based on the overland flow directions structure obtained with the modified digital elevation model and the digitized river network. The RHHUs division is automatically produced in the preparation of a watershed database using a threshold, specified by the user, indicating the maximum number of cells upstreaming each RHHUs. This preparation parameter gives the opportunity to modify the number of RHHUs within the watershed territory by adjusting its value. As the hydrological processes are simulated on each RHHU, a smaller number of RHHUs could speed-up the simulation for all submodels. However, Table 1 suggests the opposite for the production of the geomorphological hydrographs. A smaller number of RHHUs results in RHHUs with a larger distance between the upstream and the downstream cells. Routing the reference water depth over all cells on a longer flow distance on each RHHU is more time-consuming. Utilization of this type of surrogates leads to the following question: could the gain in computational time obtained in the routine of all submodels with a smaller number of RHHUs compensate the intensive computational time assigned to the production of longer geomorphological hydrographs? Obviously, this situation is relevant to HYDROTEL 19 since the hydrographs are frequently re-calculated. With a high level of uncertainty in this surrogate type, poor CPU time ratio for HYDROTEL 19 may be anticipate, but a medium or superior ratio for HYDROTEL 10.

3 Original calibration problems

This section describes the six original calibration problems studied in this paper. The two formulations of the HYDROTEL problem with 10 and 19 calibration parameters in Section 3 are tested on three watersheds with natural streamflows, all located in the province of Québec in Canada: the Ceizur watershed with a 6,928 km^2 area which is an upstream sub-basin of the Gatineau River (west of the province), the Cowansville watershed with a 215 km^2 area which is an upstream sub-basin of the Yamaska River (south of the province) and the Toulnustouc watershed with a 8,109 km^2 area which is an upstream sub-basin of the Manicouagan River (north of the province). All three watersheds differ in their soil types, land cover, land uses and topography. The six combinations of “Watershed-HYDROTEL problem” establish the original calibration problems and are originally developed by Huot et al. (2017).

3.1 Computational time

As for the potential surrogate models in Section 2.3, the computational time for one single simulation of the original hydrological model is obtained by averaging the computational time of 1,000 simulations with random calibration parameters on a 3.40 GHz Intel Core i7 processor with 12 Go of RAM. Table 2 summarizes all averaged computational times for all combinations of "Watershed-HYDROTEL problem". Recall that HYDROTEL 19 includes the production of the geomorphological hydrographs, which is intensive in computational time when related calibration parameters are modified.

3.2 Meteorological data and calibration time-periods

Meteorological data, including daily minimum and maximum temperatures and daily precipitations, were obtained from gridded datasets (see the acknowledgements section for data sources) and Table 2 shows the number of meteorological stations in each watershed. Calibration time-period spans from October 1, 1988 to September 30, 1992 for the Ceizur watershed (4 years), from October 1, 2000 to September 30, 2005 for the Cowansville watershed (5 years) and from October 1, 1984 to September 30, 1988 for the Toulnustouc watershed (4 years). Daily streamflow at the outlets are observed on the same time-periods. The initial number of RHHUs for each watershed was previously designed by industrial partners (see acknowledgments section for data source) in order to optimize the quality of modelling and to maintain moderate computational time. The number of RHHUs obtained are also shown in Table 2.

Table 2: Summary features of the six original simulation model problems relative to three studied watersheds modeled on HYDROTEL 10 and HYDROTEL 19.

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

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

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

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

3.3 Objective function

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

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

on daily time-steps $i = 1, 2, 3, \ldots, N$ of the calibration time-period, and where the residual variance is calculated between the observed streamflow value O_i and the simulated (predicted) streamflow value S_i and the observed data variance is calculated between the observed streamflow value O_i and the mean of observed streamflows $\overline{O}$. The value $1 - NSE$ determines the ratio of the residual variance (noise) to the observed data variance (information). The value $1 - NSE$ ranges from 0 to ∞, with $1 - NSE = 0$ indicating a perfect fit between observed and simulated values. When $1 - NSE > 1$, the mean of observed streamflows displayed on all the time-period is considered as a better quality hydrograph than the simulated streamflows.

4 Computational experiments

Each computational experiment presented below is based on simulation results of 5,000 parameter sets uniformly distributed in the parametric space generating with the Latin Hypercube Sampling (McKay et al., 1979). The objective function values $(1 - NSE)$ obtained with the original (horizontal axis) and surrogate (vertical axis) models are graphically compared in scatter plots supporting by the Spearman coefficient (Rs) and the R-Square coefficient (R^2) to evaluate the representativeness. As mentioned previously, it is more important to focus on the representativeness of the optima zones in a perspective of using surrogate models in calibration processes. Therefore, the scatter plots inside figures are presented on adjusted axis zooming on optima zones. A good representativeness is considered when the 0.9 value is reached by Rs and R^2 as recommended by Toal (2015). The slanted line represents a perfect fit between original and surrogate models. The gray area is delimited by the best known value of the objective function for each original "Watershed-HYDROTEL problem". The average computational time by one single surrogate simulation and the CPU time ratios are presented in following tables. Results are presented on the Toulnustouc watershed because it represents the worst watershed candidate in terms of the representativeness for all types of surrogates.

4.1 Reducing the number of meteorological stations

Figure 2 illustrates the representativeness between surrogates having different levels of meteorological station removal and the original models for the Toulnustouc watershed on HYDROTEL 10 and 19. The withdrawing process of meteorological stations generates many possibilities of dataset distribution on the territory and variability into the number of meteorological stations selected. 4 different instances have been made on each level of meteorological station removal. Consequently, this figure presents the worst and the best instances for each level of meteorological station removal in terms of the values of the Rs and R^2 coefficients.

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

Second, Figure 2 shows that when more meteorological stations are removed, more the representativeness may decrease significantly (worst instances) or slightly (best instances), sign of some variability in the results. Example, a wide difference in representativeness is obtained between worst and best instances at 87.5% level of stations removal for HYDROTEL 10. Indeed, as more stations are removed, the selection of meteorological stations must be made with vigilance. Importance is granted to a systematic station removing process to maintain a good distribution of the data on the territory and to integrate the geographic and meteorological knowledge by the user on the territory.

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

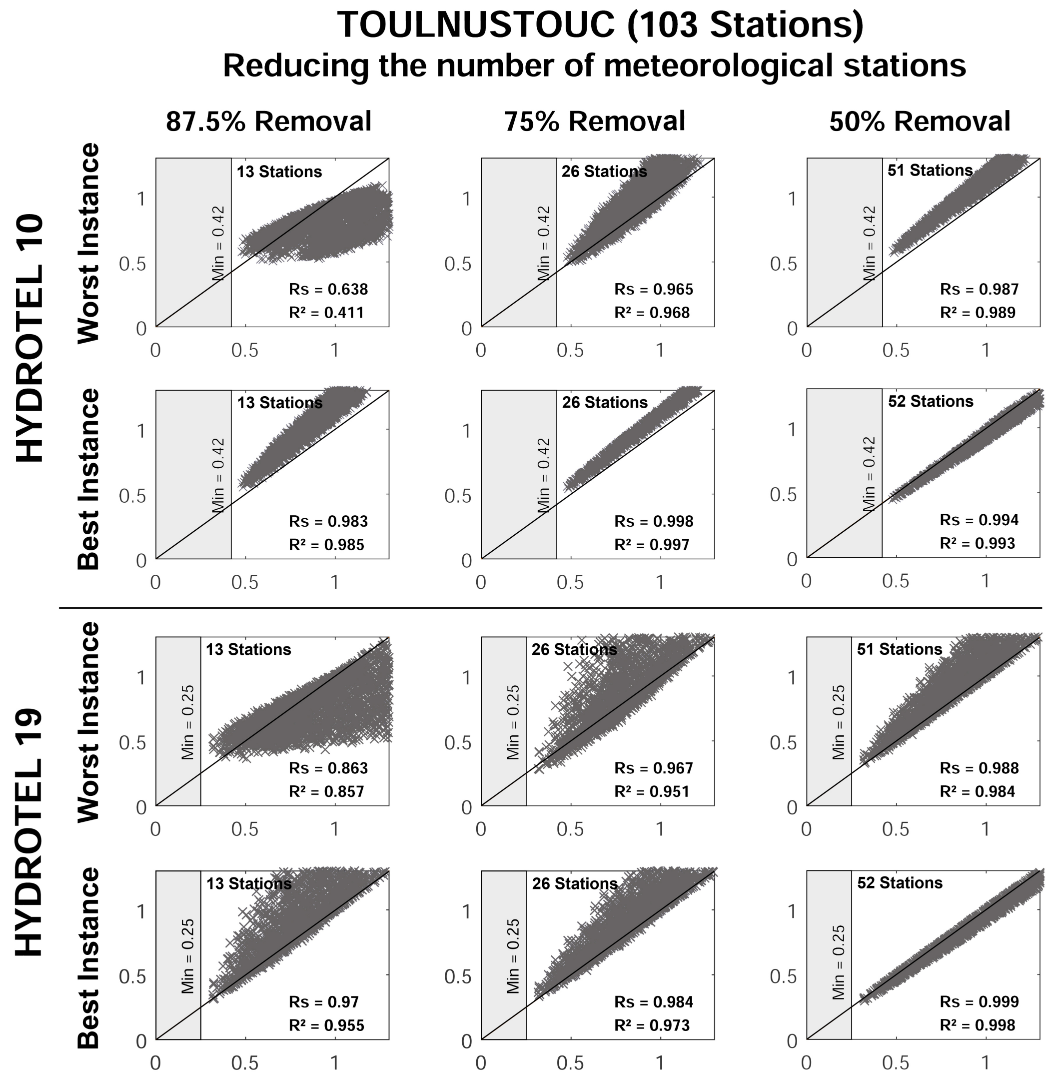

Figure 2: $1 - NSE$ **simulation results obtained for HYDROTEL 10 and 19 problems on the Toulnustouc watershed comparing the surrogates having a reduced number of meteorological stations with the original models.** Rs **and** R^2 **coefficients are calculated for the representativeness evaluation.**

Table 3 suggests that this type of surrogate decreases the computational time but does not provide better than poor CPU time ratios for any instance. Obviously, better CPU time ratios are associated to the surrogates having 87.5% of meteorological stations removed. Surrogates on the Ceizur watershed offer the best CPU time ratios, then the Toulnustouc on second rank and Cowansville on last one. This ranking is related to the initial number of meteorological stations on the original models; i.e. the interpolation of meteorological data from a huge number of stations consumes a big slice of pie in the total computational time of the HYDROTEL simulating routine, as for the Ceizur watershed. Even if the percentage of meteorological stations removed is the same, the CPU time consumed decrease further on a huge number of stations. Conversely, the interpolation of meteorological data from a small number of stations as on the Cowansville watershed represents a very short computational time, about a few seconds, in the HYDROTEL simulating routine.

With results presented above, the best compromise for all "Watershed-HYDROTEL problem" combinations between representativeness and a decrease of CPU time is the surrogates having 87.5% of meteorological stations removed. High representativeness may be obtained if station removing process is completed thoughtfully, but CPU time ratios are clearly not as good as we would like to use this type of surrogates inside efficient calibration processes.

Table 3: Average CPU time for surrogate simulation having a reduced number of meteorological stations and CPU time ratios for all "Watershed-HYDROTEL problem" combinations.

		HYDROTEL - 10 Parameters			HYDROTEL - 19 Parameters		
		CEIZUR	COWANSVILLE	TOULNUSTOUC	CEIZUR	COWANSVILLE	TOULNUSTOUC
Benchmark Models		120.84 sec	26.81 sec	49.43 sec	116.18 sec	40.96 sec	89.05 sec
Surrogates Number of Meteo Stations	50% removal	79.75 sec	26.73 sec	38.27 sec	92.35 sec	40.45 sec	77.22 sec
		poor	unusable	unusable	unusable	unusable	unusable
	75% removal	69.90 sec	26.20 sec	36.15 sec	85.51 sec	39.53 sec	74.78 sec
		poor	unusable	unusable	unusable	unusable	unusable
	87.5% removal	66.13 sec	26.44 sec	25.34 sec	76.13 sec	40.66 sec	64.28 sec
		poor	unusable	poor	poor	unusable	unusable

CPU Time Ratio: -∞, 1.5] = unusable,]1.5, 4] = poor,]4, 8] = medium,]8, 15] = superior,]15, +∞ = ideal

4.2 Reducing the calibration time-period

Figure 3 illustrates the representativeness between surrogates with four different calibration time-periods: 6, 9, 12 and 18 months, and the original models for the Toulnustouc watershed on HYDROTEL 10 and 19. The original models for Toulnustouc includes a calibration time-period of four years from 1 October 1984 to 30 September 1988. This same period has been divided into many calibration time-periods to perform the surrogate models. Start and end dates are carefully chosen to avoid starting or interrupting a hydrological event, especially on calibration time-period under 1 year. All calibration time-periods of surrogate models begins on a dry period and finish after a flow recession period. No initialization period has been rooted first before the calibration time-period considered because the initial known conditions are carefully set inside the HYDROTEL routine. Two instances with 18 months are selected: October 1, 1984 to March 31, 1986 and October 1, 1986 to March 31, 1988. Four instances with 12 months are considered beginning on October 1 of each year and ending on September 30 of the following year. Four instances with 9 months beginning on October 1 of each year and ending on June 30 of the following year and four instances with 6 months beginning on January 1 of each year and ending on June 30 of the same year are also considered. As the three watersheds chosen are highly influenced by the accumulation and melting of snow, calibration time-periods of 6 and 9 months focus on the period where spring flood occurs of peak flows. Other instances, not included in this paper, have been tested on short time-periods of 6 or 9 months focusing on other hydrological events (dry period) and worst representativeness have been obtained on every trial. Considering that several sub-time-periods can be partitioned inside the original 4-year time-period, Figure 3 presents only two different instances (the worst and the best values of Rs and R^2) on each calibration time-period previously described.

Figure 3 shows that the representativeness of the surrogate models may be importantly impacted by a calibration time-period under 1 year. Surrogate models with a calibration time-period of 6 months never have a Rs value exceeding 0.512, nor a R^2 value exceeding 0.392. The scatter plots also present a wide extend of discrepancy. With such results, it is clear that a short time-period of 6 months covering only part of the snow accumulation period (November and December snow is missing) and the melting period does not adequately represent 4 completed annual hydrological cycles of the original calibration time-period.

For the calibration time-period of 9 months, Figure 3 shows that acceptable representativeness may be obtained on both HYDROTEL problems achieving Rs and R^2 coefficients both close to the 0.9 limit value. However, much lower coefficients are also obtained. This lack of consistency makes difficult it to consider this calibration time-period as a potential useful surrogate model inside an efficient calibration process. Conclusions obtained on the Ceizur and the Cowansville watersheds for this two time-periods are similar.

Figure 3 shows that the best representativeness is reached with the 12 months time-period where the Rs and R^2 coefficients are very close or superior to the 0.9 limit value recommended by Toal (2015). This high level of representativeness obtained is related to the coverage of all hydrological events within the calibration time-period. Thus, it is easier to represent the average behaviour of the original calibration time-period consisting of 4 annual hydrological cycles in this case. It is important not to conclude that a time-period of 12 months is sufficient to well represent an original calibration period, but rather that the ratio between the original and surrogate calibration time-periods is short enough to maintain a high representativeness. Representativeness on the time-period of 12 months for the Cowansville watershed are slightly worst because the original calibration time-period is composed of 5 annual hydrological cycles unlike 4 years for the other two watersheds.

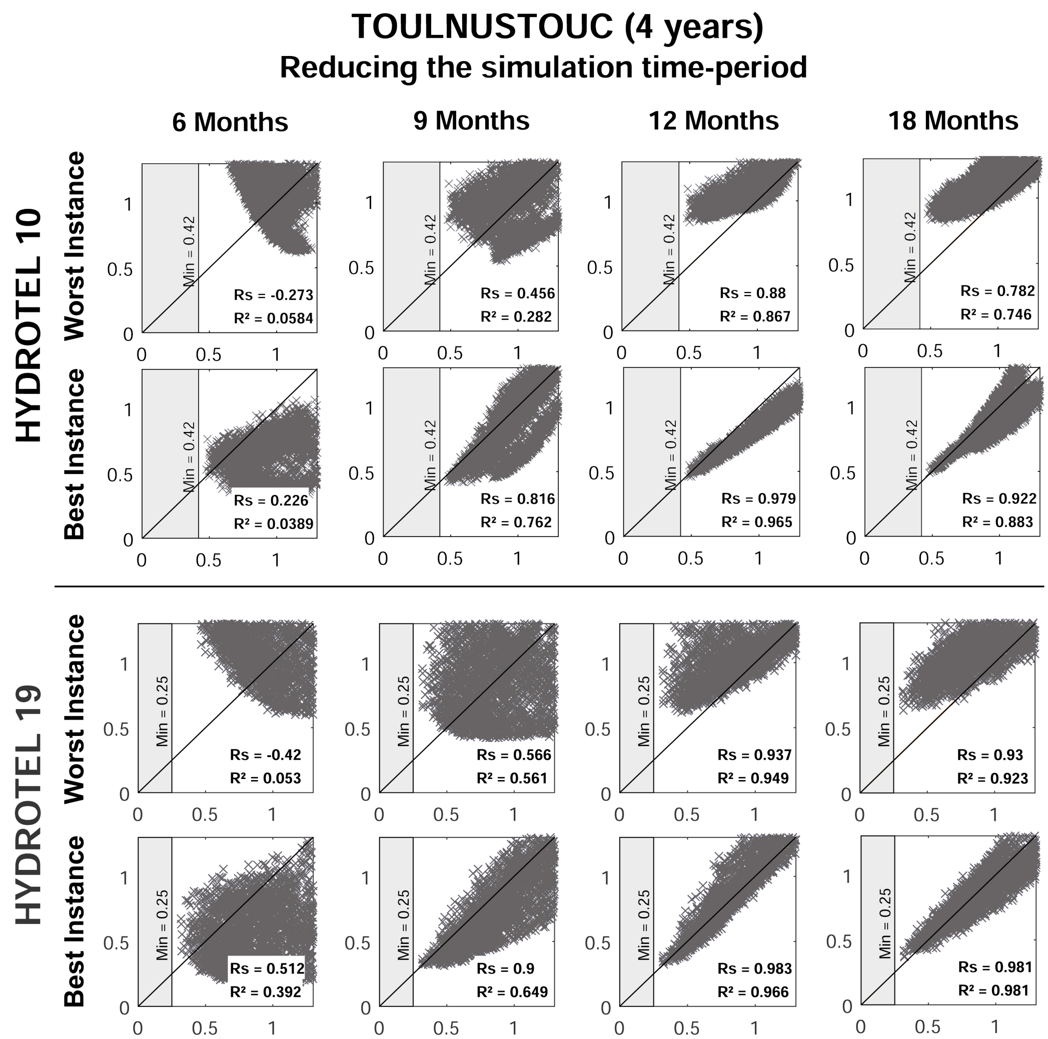

Figure 3: $1 - NSE$ **simulation results obtained for HYDROTEL 10 and 19 problems on the Toulnustouc watershed comparing the surrogates having a reduced calibration time-period with the original models.** Rs **and** R^2 **coefficients are calculated for the representativeness evaluation.**

Finally, Figure 3 shows that the calibration time-period of 18 months provides lower representativeness coefficients than the time-period of 12 months. As with the 6 or 9 months time-periods, an incomplete annual hydrological cycle in the calibration time-period of 18 months slightly reduces the level of representativeness.

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

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

Table 4: Average CPU time for one single simulation and CPU time ratio of surrogates having a reduced calibration time-period for all six "Watershed-HYDROTEL problem" combinations.

		HYDROTEL - 10 Parameters			HYDROTEL - 19 Parameters		
		CEIZUR	COWANSVILLE	TOULNUSTOUC	CEIZUR	COWANSVILLE	TOULNUSTOUC
Benchmark Models		120.84 sec	26.81 sec	49.43 sec	116.18 sec	40.96 sec	89.05 sec
Surrogates Simulation Time-Period	18 months	46.13 sec	8.12 sec	16.38 sec	66.93 sec	21.72 sec	55.15 sec
		poor	poor	poor	poor	poor	poor
	12 months	34.81 sec	5.96 sec	12.71 sec	59.46 sec	19.57 sec	51.48 sec
		poor	medium	poor	poor	poor	poor
	9 months	27.64 sec	5.00 sec	8.14 sec	54.06 sec	18.64 sec	47.08 sec
		medium	medium	medium	poor	poor	poor
	6 months	15.96 sec	3.54 sec	5.38 sec	47.83 sec	17.28 sec	43.84 sec
		medium	medium	superior	poor	poor	poor

CPU Time Ratio: -∞, 1.5] = unusable,]1.5, 4] = poor,]4, 8] = medium,]8, 15] = superior,]15, +∞ = ideal

4.3 Reducing the number of RHHUs

The automatic process for the RHHUs division of the territory is based on the overland flow directions structure. Therefore, the preparation of this type of surrogates requires a lot of trials to balance the size of each RHHU versus the number of RHHUs. In this particular case, it is more difficult to anticipate the impact of RHHUs reduction on representativeness of the surrogates and CPU time ratios. Figure 4 shows results for the Toulnustouc watershed on HYDROTEL 10 and 19 at different percentages of reduction of the number of RHHUs. Unlike the previous figures, Fig.4 does not show the worst and best instances, because a single instance is considered.

The figure reveals that the representativeness decreases gradually as the number of RHHUs decreases. However, the representativeness on all instances is not so impacted by the RHHUs reduction process. Figure 4 presents an overall Rs greater than 0.961 and R^2 greater than 0.945 even with a high percentage of reduction of the number of RHHUs (87.5% of RHHUs removed on HYDROTEL 19). Higher representativeness is obtained on all instances of the Ceizur and Cowansville watersheds for both HYDROTEL problems.

A second observation is that the extent of discrepancy between the original and surrogate models on HYDROTEL 19 is more important than on HYDROTEL 10. Considering this, six instances have been made for HYDROTEL 19 with a gradual reduction on the number of RHHUs for better representation and only three instances with more aggressive percentages of reduction for HYDROTEL 10. According to our analysis, this difference is assigned to the production of the geomorphologic hydrographs (step 2 in Table 1) inside the HYDROTEL routine when the calibration parameters relating with them are modified. On HYDROTEL 10, the calibration parameters related to the production of geomorphological hydrographs are not modified on the 5,000 comparison simulations, but the hydrographs are still produced once because the RHHUs subdivision of the territory have changed following the reduction process. So, on this HYDROTEL problem, the representativeness is evaluated between two different RHHUs subdivisions with their own hydrographs patterns. On HYDROTEL 19, the calibration parameters are modified many times on the 5,000 comparison simulations which leads to the production of new hydrographs for the original and surrogate models every time the calibration parameters related with them are modified. The representativeness is evaluated between the original RHHUs subdivision having a multitude of hydrographs patterns depending on the values of the calibration parameters and the new RHHUs subdivision (surrogate model) having also a multitude of different calculated hydrographs patterns. This may explain why the impact on representativeness is higher on the HYDROTEL 19 problem.

Table 5 presents two opposite scenarios for HYDROTEL 10 versus 19. Results for HYDROTEL 10 suggest that reducing the number of RHHUs decreases the computational time and improves the CPU time ratios. Medium CPU time ratio is obtained for the Cowansville watershed and poor ratios are obtained for Ceizur and Toulnustouc with 87,5% RHHUs removal. A high representativeness and better CPU time ratios may indicate that this surrogate may provide acceptable assets to use it inside efficient calibration processes for HYDROTEL 10.

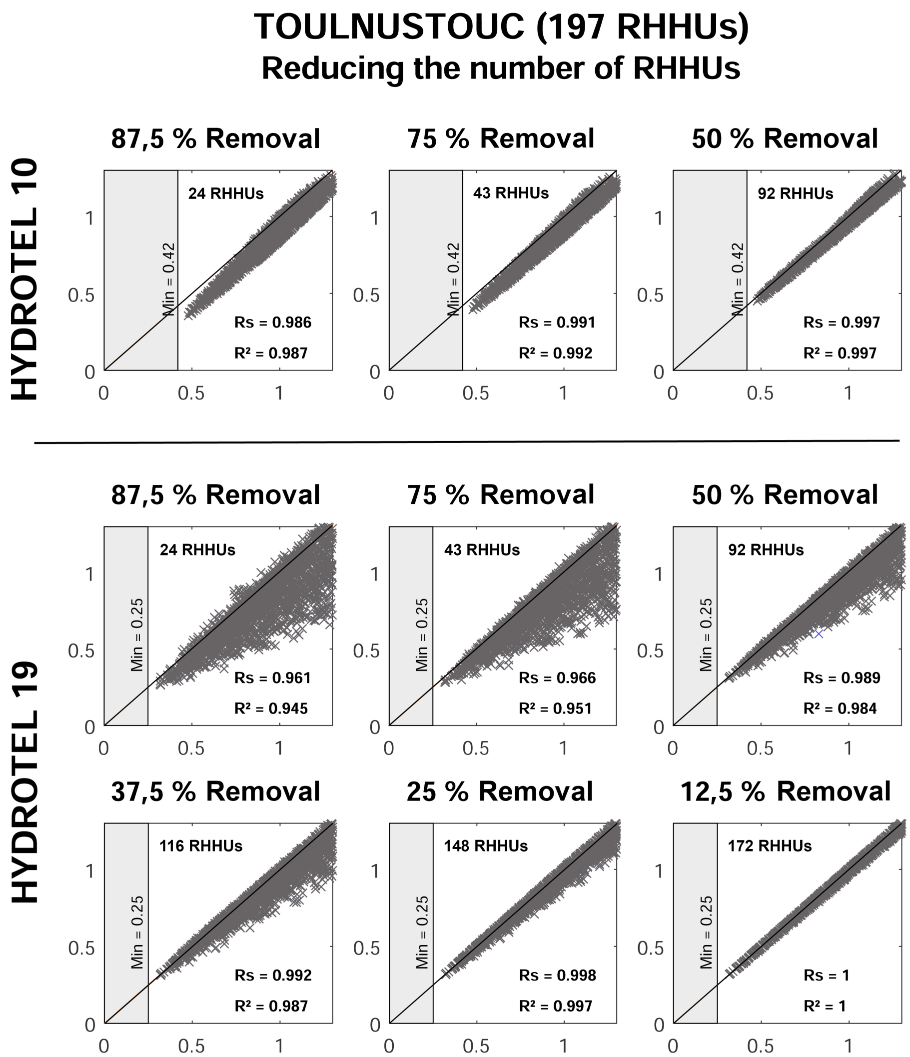

Figure 4: $1 - NSE$ simulation results obtained for HYDROTEL 10 and 19 problems on the Toulnustouc watershed comparing the surrogates having a reduced number of RHHUs with the original models. Rs and R^2 coefficients are calculated for the representativeness evaluation.

Results for HYDROTEL 19 show that reducing the number of RHHUs is problematic. As mentioned in Section 3.3, the main question is about the possibility of reaching a low-cost surrogate simulation with the particular routine of HYDROTEL 19, i.e. that introduces less computing of hydrological processes due to a smaller number of RHHUs, but that also involves the production of longer geomorphological hydrographs. Table 5 gives the answer to this question with unusable CPU time ratios for each instance. Slight gains are achieved through wide reductions in RHHUs, but the production of longer hydrographs almost balances the computational time consumed. It possible to observe a slightly peak gain in computational time at 75% RHHUs removal for Ceizur and Cowansville watersheds and at 37,5% for Toulnustouc. However, some instances on Toulnustouc watershed are more time-consuming than the original models. High representativeness is obtained on HYDROTEL 19, but gains in computational time is too little to use this type of surrogate inside an efficient calibration process.

Table 5: Average CPU time for surrogate simulation having a reduced number of RHHUs and CPU time ratios for all "Watershed-HYDROTEL problem" combinations.

		HYDROTEL - 10 Parameters			HYDROTEL - 19 Parameters		
		CEIZUR	COWANSVILLE	TOULNUSTOUC	CEIZUR	COWANSVILLE	TOULNUSTOUC
Benchmark Models		120.84 sec	26.81 sec	49.43 sec	116.18 sec	40.96 sec	89.05 sec
Surrogates Number of RHHUs	12.5% less	- - -	- - -	- - -	105.99 sec unusable	37.53 sec unusable	80.31 sec unusable
	25% less	- - -	- - -	- - -	103.69 sec unusable	34.90 sec unusable	78.68 sec unusable
	37.5% less	- - -	- - -	- - -	103.41 sec unusable	31.26 sec unusable	75.44 sec unusable
	50% less	91.58 sec unusable	15.61 sec poor	31.72 sec poor	108.61 sec unusable	31.51 sec unusable	85.54 sec unusable
	75% less	67.76 sec poor	9.91 sec poor	22.77 sec poor	100.27 sec unusable	29.70 sec unusable	112.42 sec unusable
	87.5% less	58.92 sec poor	6.37 sec medium	17.72 sec poor	104.34 sec unusable	34.11 sec unusable	124.74 sec unusable

CPU Time Ratio: -∞, 1.5] = unusable,]1.5, 4] = poor,]4, 8] = medium,]8, 15] = superior,]15, +∞ = ideal

4.4 Final combined surrogate models

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

Individually, each type of surrogate model has shown that reduced-fidelity models can be produced with high representativeness, but the CPU time ratios obtained are mostly less than that would be expected if to be used within an efficient calibration process. In order to amplify the CPU time ratio, combinations of the three different types of surrogates is tested on each original "Watershed-HYDROTEL problem" to produce final combined surrogate models. For all watersheds on HYDROTEL 10, final combined surrogate models are set at 87.5% of meteorological stations removal, a 12 months' calibration time-period and at 87.5% of RHHUs removal. For HYDROTEL 19, the variability on the best level of RHHUs removal depends on modeled watershed. This situation generates two different experiments. The first experiments tested the final combined surrogate models at 87.5% of meteorological stations removal, a 12 months' calibration time-period and at the best level of RHHUs removal depending on modeled watershed (75% for Ceizur and 37.5% for Cowansville and Toulnustouc). The second tested the final combined surrogate models at 87.5% of meteorological stations removal, a 12 months' calibration time-period and at the original number of RHHUs; i.e. that the final combined surrogate models are set without reducing the original number of RHHUs.

Table 6 presents that the CPU time ratios obtained for HYDROTEL 10 are all qualified as ideal when the best level of reduced-fidelity models is chosen. Meaning that the final combined surrogate models are at least 15 times less time-consuming than the original hydrological models. More precisely, the CPU time ratios are 44 for the Ceizur watershed, 16 for the Cowansville watershed and 32 for the Toulnustouc watershed. The CPU time ratios obtained for HYDROTEL 19 are, on the opposite, qualified as poor for both experiments, where the CPU time ratios are around 2. Table 6 also illustrates that it is more interesting, in terms of the CPU time ratios, to set the final combined surrogate models without reducing the original number of RHHUs. As reported previously, the production of the geomorphological hydrographs inside the HYDROTEL 19 routine brings uncertainty and variability in the computational time. Moreover, results in Table 6 relates that there is no benefit in reducing the original number of RHHUs even when the best level of RHHUs removal is selected per watershed. Consequently, the experiments with final combined surrogate models set without reducing the original number of RHHUs are selected for HYDROTEL 19.

Table 6: Average CPU time for final combined surrogate simulation and CPU time ratio for all six "Watershed-HYDROTEL problem" combinations.

		HYDROTEL - 10 Parameters			HYDROTEL - 19 Parameters		
		CEIZUR	COWANSVILLE	TOULNUSTOUC	CEIZUR	COWANSVILLE	TOULNUSTOUC
Benchmark Models		120.84 sec	26.81 sec	49.43 sec	116.18 sec	40.96 sec	89.05 sec
Final Combined Surrogate Models	Best Level of Reduced-Fidelity Models	2.69 sec	1.61 sec	1.54 sec	46.64 sec	22.54 sec	51.22 sec
		ideal	ideal	ideal	poor	poor	poor
	Without RHHUs Reduction	- - -	- - -	- - -	47.46 sec	19.79 sec	45.82 sec
					poor	poor	poor

CPU Time Ratio: -∞, 1.5] = unusable,]1.5, 4] = poor,]4, 8] = medium,]8, 15] = superior,]15, +∞ = ideal

Figure 5 presents the representativeness of the final combined surrogate models for each combination "Watershed-HYDROTEL problem" and the summary of its characteristics regarding each type of surrogate models.

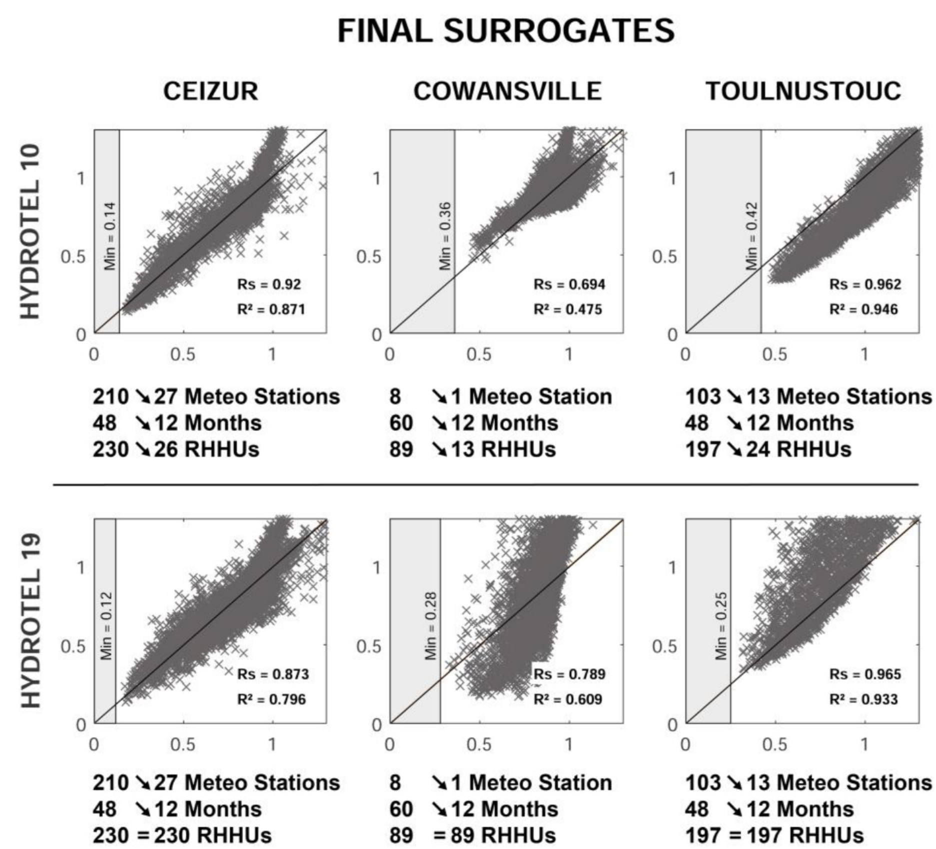

Figure 5: $1 - NSE$ simulation results obtained for HYDROTEL 10 and 19 problems on the Ceizur, Cowansville and Toulnustouc watersheds comparing the final selected surrogates with the original models. Rs and R^2 coefficients are calculated for the representativeness evaluation. The reduction levels are also indicated below each case presented in the figure.

The combination of the three types of surrogates (final combined surrogates) provides higher CPU time ratios in comparison with only one type of surrogate model as presented in Tables 3 to 6. However, representativeness coefficients of the final combined surrogates illustrated in Figure 5 are at a lower-quality and 3 out of 6 surrogate models on this same figure yield representativeness coefficient less than 0.9.

5 Discussion

This paper explores and develops three different types of reduced-fidelity surrogate models and their combination for the original hydrological model HYDROTEL. The surrogates reduce the number of meteorological stations inputting the modelling, the calibration time-periods and the watersheds spatial discretization. The representativeness and the CPU time ratios between original and surrogate models has been compared and evaluated to use it within calibration processes.

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

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

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

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

Thirdly, the reduced-fidelity models developed in this paper raise questions about the original formulation of the HYDROTEL problems. Final reduced-fidelity models suggest high representativeness of the original models, which means that this reduced-modelling level could also well perform in hydrological modelling. Future work will explore the opportunities to use rightfully the reduced-fidelity models as original hydrological models in various hydrological modelling context.

A Survey of all calibration parameters inside the HYDROTEL problems

Submodels Used	N°	HYDROTEL Model Parameters	Units [a]	10 Parameters [b]	19 Parameters [b]	Lower Bound	Upper Bound
Thiessen Polygons	1	Lapse Rate of Precipitation	mm/day		X	0	2
	2	Lapse Rate of Temperature	°C/100m		X	-0,8	-0,3
	3	Base Refreezing Temperature	°C	X	X	-6	-4
Mixed (degree-day) energy-budget method	4	Melt Rate Coefficient at the Soil-Snow Interface	mm/day		X	0,4	0,6
	5	Maximum Density of Snow Cover	kg/m³		X	300	550
	6	Ratio of Snow Compaction			X	0,005	0,015
	7	Degree-Day-Factor - Conifers Trees	mm/day/°C	X	X	0,01	6,39
	8	Temperature Threshold for Melt - Conifers Trees	°C	X	X	0,01	1,99
	9	Degree-Day-Factor - Deciduous Trees	mm/day/°C		X	0,01	3,99
	10	Temperature Threshold for Melt - Deciduous Trees	°C		X	0,01	1,99
	11	Degree-Day-Factor - Non-Forest Environments	mm/day/°C	X	X	0,01	3,99
	12	Temperature Threshold for Melt - Non-Forest Environments	°C	X	X	-6	1,59
Hydro-Quebec method	13	Potential Evapotranspiration Factor		X	X	0,7	1,2
BV3C	14	Lower limit of soil layer #1	m	X	X	0,05	1
	15	Lower limit of soil layer #2 [c]	m	X	X	0,05	1
	16	Lower limit of soil layer #3 [d]	m	X	X	0,05	4
	17	Extinction Coefficent of Solar Radiation by Vegetation					
	18	Recession Coefficient	m/hour	X	X	0,00000001	0,0001
	19	Additive Coefficient in Soil Type Index					
	20	Drying Effect Factor from Evaporation at Ground Surface					
	21	Maximum Variation in Relative Humidity per time step					
Kinematic Wave Equation	22	Manning's Roughness Coefficient - Forest Cover			X	0,04	0,3
	23	Manning's Roughness Coefficient - Water					
	24	Manning's Roughness Coefficient - Other Environments					
	25	Depth of Runoff of the Geomorphological Hydrograph	m		X	0,0002	0,002
Kinematic Wave Equation	26	Roughness Optimization Factor					
	27	River Widths Optimization Factor					

[a] Blank spaces mean no unit
[b] Blank spaces mean fixed parameter
[c] Lower Limit of Soil Layer #2 = 0,05 ≤ x15 ≤ 1 + x14
[d] Lower Limit of Soil Layer #3 = 0,05 ≤ x16 ≤ 4 + x15

References

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., Orban, D. (2006). Finding Optimal Algorithmic Parameters Using Derivative-Free Optimization. SIAM Journal on Optimization, 17(3), 642–664. DOI:10.1137/040620886.

Audet, C. (2014). A survey on direct search methods for blackbox optimization and their applications. In Pardalos, P.M., Rassias, T.M. (eds): Mathematics without boundaries: Surveys in interdisciplinary research (chapter 2), 648p. Springer, New-York.

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

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

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

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.

Das, T., Bardossy, A., Zehe, E., He, Y. (2008). Comparison of conceptual model performance using different representations of spatial variability. Journal of Hydrology, 356, 106–118. DOI:10.1016/j.jhyfrol.2008.04.008.

Fieller, E.C., Hartley, H.O., Pearson, E.S. (1957). Test for rank correlation coefficients. I. Biometrika, 44(3–4), 470–481. DOI:10.1093/biomet/44.3-4.470.

Fortin, J.P., Turcotte, R., Massicotte, S., Moussa, R., Fiztback, J. (2001a). A Distributed Watershed Model Compatible with Remote Sensing and GIS Data, Part 1: Description of the model. Journal of Hydrologic Engineering, American Society of Civil Engineering, 6(2), 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, 6(2), 100–108.

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.

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

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

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, 21(2), 239–245.

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, Directed by 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., Harmel, R.D., Veith, T.L. (2007). Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. Transactions of the ASABE, 50(3), 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, 41(11), W11427. DOI:10.1029/2005WR004134.

Nash, J.E., Sutcliffe, J.V. (1970). River flow forecasting through conceptual models part I - A discussion of principles. Journal of Hydrology, 10(3), 282–290.

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, 13(3), 193–214.

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

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 basin. Journal of Hydrology, 409(3–4), 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, 420–421, 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 pre-emption. 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.

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. (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.

Ricard, S., Bourdillon, R., Roussel, D., Turcotte, R. (2012). Global Calibration of Distributed Hydrological Models for Large-Scale Applications. Journal of Hydrologic Engineering, 18(6), 719–721. DOI:10.1061/(ASCE)HE.1943-5584.0000665.

Servat, E., Dezetter, A. (1991). Selection of calibration objective functions in the context of rainfall-runoff modelling in a sudanese savannah area. Hydrological Sciences Journal, 36(4), 307–300. DOI:10.1080/02626669109492517.

Shoemaker, C.A., Regis, R.G., Fleming, R.C. (2007). Watershed calibration using multistart local optimization and evolutionary optimization with radial basis functions approximation. Hydrological Sciences Journal, 52(3), 450–465.

Singh, V., Woolhiser, D. (2002). Mathematical Modeling of Watershed Hydrology. Journal of Hydrology Engineering, 7(4), 270–292.

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.

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, 45(2), 460–474. DOI:10.1111/j.1752-1688.2009.00302.x.