A complementary modelling approach to manage uncertainty of computationally expensive models

Z. Vojinovic
Department of Hydroinformatics and Knowledge Management, UNESCO-IHE, Institute for Water Education, Westvest 7, 2611 AX Delft, The Netherlands (E-mail: z.vojinovic@unesco-ihe.org)

Abstract The fact that the models applied in the ‘water domain’ are far from reality can be attributed to many reasons. In this context, a systematic analysis of uncertainties reflected by the model error can provide insight into the level of confidence in the model results and how to approach estimation of optimal model parameters. This paper discusses the four commonly used approaches for estimation of model parameters and suggests that an alternative complementary modelling approach should be considered in cases where the traditional model calibration gives limited results and particularly in cases where the computationally expensive models are concerned. It treats uncertainty as modelling the total discrepancy between the model and physical process. The proposed approach combines the results from a physically-based model and Support Vector Machine model into the final solution.

Keywords Complementary modelling; computationally expensive models; uncertainty; urban drainage models

Introduction
Since the models are approximation of reality, the uncertainty of model results is inevitable, and therefore, making decisions on a single model output without the consideration of uncertainties and risk involved can be very dangerous and misleading. Typically, for computationally expensive models, as it is in the case of urban drainage models, the degree of uncertainty will largely depend on model physics used to describe the processes, initial conditions, boundary conditions and accuracy of the measurements, which are normally used as an input (e.g., rainfall) and/or as a basis to estimate model parameters (e.g., flow and depth measurements). There are several ways of improving outputs from such models and most of them relate to the use of different calibration methods.

Traditionally, by model calibration we understand the process of adjusting model parameters that would enable to closely match the observed (or measured) behaviour of the real system. Lately, however, an extended notion of calibration is used: calibration is understood as combining all relevant information into a parameter probability distribution, or some other function reflecting their uncertainty. Such understanding of calibration is based on the assumption that it is difficult, if not impossible, to determine an ‘optimal’ parameters set, and therefore, it is reasonable to consider their distribution. There has been an argument over the years whether an optimal set of model calibration parameters exist. Since it is possible to derive literally thousands of different parameter sets, where all of them can satisfy the given calibration criteria, Beven and Binley (1992) have been arguing that an alternative approach to the traditional calibration approach must be adopted. Furthermore, following the increase in computer power, the research into automatic calibration procedures has led to the use of different automatic parameter optimisation approaches. These approaches are based on a general task of optimising (minimising or maximising, as appropriate) the value of objective function, which is a
numerical measure of the difference between observed and simulated data. Automatic parameter optimisation has its advantages, when compared to the manual calibration, in that it is faster since it is computer based, is less subjective, and the confidence of the model simulation can be explicitly stated. However, due to the difficulties arising from defining the best value or criterion to be optimised, the search for the global optimum when many parameters are involved and mutually dependent, and the impossibility to differentiate between different error sources the single-criteria methods used for automatic calibration have found limited application. Subsequently, automatic calibration procedures started to employ multiple performance criteria in the search for optimal parameter set. Certainly, assuming that the measurements are reliable, it could be further argued that inability of achieving better calibration results resides on the simplicity of model physics and as such any further adjustment of model calibration parameters might be pointless.

From the practical point of view, there are several approaches that can be applied to almost all computationally inexpensive models. However, there are several issues associated with most of them when it comes to computationally expensive models, which is certainly the case of urban drainage models. The difficulty comes either due to the computational time required for a single simulation or the complexity of the processes being modelled. Typically, these models solve full dynamic de Saint Venant equations that can be written as shown in equations (1) and (2).

\[
\begin{align*}
\frac{\partial A}{\partial t} + \frac{\partial Q}{\partial x} &= q \\
\frac{\partial Q}{\partial t} + \frac{\partial}{\partial x} \left( \frac{Q^2}{A} \right) + gA \left[ \frac{\partial y}{\partial x} - s_0 + \frac{Q|Q|}{K^2} \right] &= 0
\end{align*}
\]

where \( Q \), discharge; \( y \), water depth; \( A \), cross-sectional area; \( K \), conveyance.

With such models, as oppose to computationally inexpensive models, thousands of simulations are impractical to carry out. Also, the processes modelled in urban drainage systems are complex, subject to local heterogeneity across the catchment area and the dynamics between the rainfall and system’s response (usually reflected by the system’s overflows) are highly nonlinear (often involving unknown stochastic processes). The present paper intents to demonstrate the use of support vector machine, as one of the more promising data-driven techniques, in a complementary modelling fashion with a computationally expensive model (i.e., 1D hydrodynamic model) of an urban drainage system.

**Errors and uncertainty**

Modelling a physical system such as an urban drainage catchment involves uncertainties that arise from various sources. An understanding of implications of these uncertainties is particularly important when the results of models are used in the decision-making process. Depending on the purpose of the modelling task, a systematic analysis of uncertainties provides insight into the level of confidence in the model results, and furthermore, it can assist in the identification of the key sources of uncertainty that may or may not be important for the given task. In order to carry out systematic analysis of uncertainties it is important to bear in mind that uncertainty can arise from two sources, data as well as the model.

**The data**

The data gathering (or the measurement) process has a number of steps where at each step there is a potential to incur an error. Such errors together with the precision of the
measuring equipment are the main sources for measurement uncertainty. In order to minimise measurement uncertainty it is important to use the best available equipment for the selected gauging location and to introduce regular site inspections. On the other hand, errors that are associated with the temporal and/or spatial resolution of sampling are the main sources of sampling uncertainty. Hence, in order to minimise the sampling uncertainty the data must be collected from sufficiently large number of locations and at carefully selected time intervals.

The model
The models applied in modelling of urban drainage systems carry uncertainty in terms of their structure (structural uncertainty), spatial or temporal resolution (spatial or temporal uncertainty), simplification of the system’s geometry and catchment characteristics, parameter values (parametric uncertainty) and the level of extrapolation. Successful model calibration requires that measured conditions at a site be properly characterized in the first place. Lack of proper site characterization may result in a model that is calibrated to a set of conditions, which are not representative of actual conditions. Furthermore, in urban drainage modelling it can happen that due to the complexity of sub-processes (or in other words, simplicity of model physics) the hydrodynamic model cannot be calibrated with satisfactory results.

Methods for estimation of model parameter values
From the literature review and industry practice to date the following four approaches to the estimation of model parameter values can be found: trial-and-error manual calibration, automatic calibration, the so-called ‘GLUE’ approach and an approach based on best literature/industry code of practice. In a case study by Vojinovic and Solomatine (2006), it was demonstrated how the results of models using these four approaches may have significant differences in terms overflow frequency and spilling volumes, despite their ‘satisfactory’ calibration results. Hence, the fifth approach which treats the model error as uncertainty (i.e., a complementary modelling approach) is proposed as a means to overcome deficiencies.

Trial-and-error manual calibration
In urban drainage modelling, the model parameters are assigned to each subcatchment (or loading point) and by using this information the model calculates values of dry and wet weather flows. The calculated values are then compared with measured values and the model parameters are then adjusted according to the modeller’s judgement in order to provide a better match. In the trial-and-error approach, such procedure is repeated until the sufficient match (giving an acceptable error) is obtained. This approach has a number of features. It is a solution-oriented (it makes no attempt to discover why a solution works, merely that it is a solution), problem-specific (it makes no attempt to generalise a solution to other problems), non-optimal (it is an attempt to find a solution, not all solutions, and certainly not the best solution) and needs little knowledge (it can proceed even when there is a little or no knowledge of the subject). Despite its strengths, the practice has proved that it is often time consuming and does not guarantee to locate an optimal parameter set, and therefore, such work can certainly benefit by considering the use of automatic calibration routines.

Automatic calibration by global optimisation
With automatic calibration, the difference between the model results and measurements is normally minimised by solving an optimisation algorithm, with independent variables
being the unknown model parameters. This is nowadays becoming more attractive approach since more and more practitioners do not want to rely solely on the trial-and-error approach, and show interest in using more systematic procedures for calibration. The possibilities of using optimisation in model calibration have been demonstrated by several authors (e.g., Solomatine, 1998; Solomatine et al., 1999). This approach has also been implemented in some of the standard commercial software packages. For example, the shuffled complex evolutionary algorithm has been built within the MOUSE package.

Calibration and uncertainty analysis by using the GLUE methodology

It can be argued that identifying a single parameter set is not a desirable option, or it is difficult or impossible to build or to calibrate a model that would be good for all possible situations. In this case, a number of models can be built and their outputs can then be averaged by a weighting scheme where the weights are determined on the basis of the model skill (accuracy). One of the approaches following this general idea is the so-called Generalised Likelihood Uncertainty Estimation (GLUE) approach introduced by Beven and Binley (1992). The GLUE method provides tools for sensitivity analysis and uncertainty estimation using the results of Monte Carlo simulations. In the GLUE terminology the rejection of an idea that an optimum parameter set exist is in the favour of non-uniqueness, ambiguity or non-identifiability of parameter sets (called the equifinality). This implies that there could be several satisfying calibration parameter sets estimated on one data set, but they could all produce significantly different results on a another data set. This is certainly a valid comment and it is the basis used for GLUE methodology. However, it should be noted that the GLUE method contains a number of subjective elements such as the choice of prior parameter ranges, the choice of likelihood measure employed and the choice of threshold of acceptability. However, it does force those choices to be made explicitly. The major limitation of the GLUE methodology is the dependence on Monte Carlo simulation. The Monte Carlo sampling requires a large number of computer runs, particularly for models with a large number of parameters, and the method is more suited for implementation on parallel computer systems that greatly facilitate Monte Carlo methods. For complex models (e.g., urban drainage models), which require a great deal of computer time for a single run, it is not feasible to fully explore high order parameter response surfaces. In such case, there is a greater dependency on parallel computing or otherwise such method will never be able to find its applicability with the computationally expensive models.

Literature/industry based parameter values

‘Good practice’ is really state of the art at a given time (historically speaking). In engineering, it is nothing other than a certain way to apply norms, carry out calculations, etc., that is recognised as an ‘acceptable’ and ‘correct’ by the profession. It is presumed that, if an engineer follows ‘good practice’ rules, his project will not fail (Cunge, 2003). In true ‘good practice’, a lack of adequate data necessitates the use of the most advanced and reliable modelling tools over the greatest range of uncertainty in the data in order to compute (or assess) the level of uncertainty in the results of the coarse model. And thus the cost is increased by inadequate data. With this approach, the argument runs on the basis that in order to calibrate parameters that will subsequently be modified during the exploitation runs used for simulating the impact of future projects is most often useless and costly exercise. It further argues that the meaningful calibration is not possible, at least for most cases, because of the lack of appropriate data, or the cost of their acquisition. This view is supported by examples on how ‘obvious’ applications of a paradigm including calibration may well lead to serious errors because of the belief that calibration
is meaningful or because of a wrong choice of calibrated parameters. What is therefore suggested is a new paradigm in deterministic modelling based on validation and not on calibration. Certainly, the question that could be posed to this approach is whether we should ignore the data (i.e., measurements) that are available or not, and what should be the strategy if the local knowledge/experience does not exist? In such situations where the measurement data exist and if it is of a reliable quality, such data should not be disregarded, but indeed the way it should be utilised in the process of estimating model parameters could be open for further discussion.

**Uncertainty as an error modelling problem**

As discussed in the opening section of the paper, employing only a deterministic model to describe the process accurately will not always generate satisfactory results. The residual errors between model outputs and the measurements are the best indicators of the discrepancy between the model and real world. The greater this discrepancy is, the greater the uncertainty associated with the particular model. On this basis, a complementary modelling approach which treats uncertainty as modelling the total discrepancy between the model and the physical process can be used as a means of handling model uncertainty. In this approach, deterministic processes are described by a physically based model and stochastic processes (i.e., error of a physically based model) are represented by a data-driven error prediction model. The final solutions from both models are superimposed into a combined solution, Figure 1 (see also Vojinovic et al., 2003).

This approach may be found particularly relevant for real-time forecast applications (Babovic et al., 2001). The following case study demonstrates the performance of such complementary modelling approach when applied in the area of urban drainage modelling.

**Case study**

A separate sewer catchment, located in Seaforth area (Sydney, Australia), with relatively high inflow/infiltration rates (as it is often the case with separate sewer systems) has been chosen for the case study. The Seaforth catchment is located in the Northern Beaches in
Sydney and covers an area of approximately 356 hectares, with a residential population of approximately 10,000. The 16 temporary gauges were installed within the Seaforth sewer to record the flow data, Figure 2.

The subcatchment, monitored by the gauge 8BBC08, is used in the following analysis as a typical ‘leaf subcatchment’ and covers an area of approximately 20 hectares. Also, as a typical ‘subtract catchment’, the subcatchment monitored by the gauge 8BBC09 (which covers an area of about 150 hectares) was used in the same analysis, Figure 2.

For the physically based process modelling, the deterministic modelling system MOUSE (developed by DHI Water & Environment) was applied. As an error-correction tool, the Support Vector Machine (SVM) model, developed and described in earlier work by Vojinovic and Kečman (2004), was applied. The SVM model used here utilises the quadratic programming method for finding the optimal number of kernel functions for a given accuracy of a learning machine. The regression function is designed to minimise the following:

\[
\min_{\mathbf{w}, \xi} \frac{1}{2} \|\mathbf{w}\|^2 + C \sum_{i=1}^{l} \xi_i + \sum_{i=1}^{l} \xi_i^* 
\]

subject to the following constraints:

\[
y_i - \mathbf{w}^T \mathbf{x}_i - b \leq \xi_i \quad i = 1, l, \\
\mathbf{w}^T \mathbf{x}_i + b - y_i \leq \xi_i^* \quad i = 1, l, \\
\xi_i \geq 0 \quad i = 1, l, \\
\xi_i^* \geq 0 \quad i = 1, l,
\]

where \(\xi_i\) and \(\xi_i^*\) are slack variables for measurements above/below an \(\varepsilon\)-tube, \(\mathbf{w}\) represents the unknown model parameters and \(\mathbf{x}\) denotes model input variables. Lagrange multipliers that are introduced during the minimisation are \(\alpha_i\) and \(\alpha_i^*\) corresponding to \(\xi_i\) and \(\xi_i^*\). The following form of a SVM model was examined:

\[
e = f(u_1, u_2, u_3)
\]

**Figure 2** The flow gauge schematics
where \( e \) denotes the physically based (i.e., deterministic) model error, \( u_1 \) denotes rainfall, \( u_2 \) denotes flow measurements and \( u_3 \) denotes the physically-based model output at a selected site. Therefore, the SVM applied considers approximating function of the form:

\[
    f(\mathbf{x}, \mathbf{w}) = \sum_{i=1}^{P} w_i \varphi_i(\mathbf{x})
\]

where \( \varphi \) is the Gaussian function. Also, the SVM uses Vapnik’s error (loss) function, the so-called, \( e \)-insensitivity loss function.

\[
    |y - f(x, w)|_e = \begin{cases} 
    0 & \text{if } |y - f(x, w)| \leq e \\
    |y - f(x, w)| - e & \text{otherwise}
    \end{cases}
\]

Thus, the loss is equal to 0 if the difference between the predicted \( f(x, w) \) and the measured value is less than \( e \). In high-dimensional vector regression problems, the training vectors are mapped into a higher dimensional space by the nonlinear kernel function, then SVM finds a linear separating hyperplane with the maximal margin in the higher dimensional space. In this work, the radial basis kernel function (RBF) was used, which can be expressed by the following notation:

\[
    K(x_i, x_j) = \exp(-\varphi\|x_i - x_j\|^2), \quad \varphi > 0
\]

Several MOUSE model calibrations were carried out on a separate data set from the validation data set and on both data sets, the statistical analysis between the measured and modelled time series were carried out. At the same time, a separate data set was used for the training of the SVM model and a separate data set was used for the test phase. On both data sets, training and test, the SVM model received rainfall, flow measurements and MOUSE model discharge outputs as inputs, whereas the associated MOUSE model error values (being the difference between the measurements and model results at each time step) are used as output variable. All data sets were scaled between \(-1\) and \(1\). During the training phase, the \( e \)-values were varied between 0.5 and 1 for all simulations. From this, the resulting number of support vectors was found to be between 1\% and 5\% of the total number of training data.

The main intention of this experimental work was to improve the dry weather and wet weather flow predictions of MOUSE model at the two gauging locations (8BBC08 and 8BBC09).

### Table 1: Physically-based (MOUSE) model results

<table>
<thead>
<tr>
<th></th>
<th>Calibration data set</th>
<th>Validation data set</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>8BBC08</td>
<td>8BBC09</td>
</tr>
<tr>
<td>( E(\exp d\text{r}) )</td>
<td>19%</td>
<td>25%</td>
</tr>
<tr>
<td>CoE</td>
<td>0.79</td>
<td>0.72</td>
</tr>
</tbody>
</table>

### Table 2: Some of the statistical characteristics (min, max and mean) of the data sets used and complementary modelling results for the two gauging sites 8BBC08 and 8BBC09

<table>
<thead>
<tr>
<th></th>
<th>8BBC08</th>
<th></th>
<th>8BBC09</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Training</td>
<td>Test</td>
<td>Training</td>
<td>Test</td>
</tr>
<tr>
<td>Minimum</td>
<td>-0.64107</td>
<td>-0.85696</td>
<td>-0.033483</td>
<td>-0.047708</td>
</tr>
<tr>
<td>Maximum</td>
<td>0.021697</td>
<td>0.041924</td>
<td>0.097167</td>
<td>0.12701</td>
</tr>
<tr>
<td>Mean</td>
<td>-0.0024723</td>
<td>-0.00068248</td>
<td>0.0000033392</td>
<td>-0.0013111</td>
</tr>
<tr>
<td>( E(\exp d\text{r}) )</td>
<td>6.1%</td>
<td>7.3%</td>
<td>6.4%</td>
<td>7.7%</td>
</tr>
<tr>
<td>CoE</td>
<td>0.99</td>
<td>0.98</td>
<td>0.99</td>
<td>0.98</td>
</tr>
</tbody>
</table>
Figure 3 MOUSE model errors on the validation set at the gauging site 8BBC08

Figure 4 MOUSE model errors on the validation set at the gauging site 8BBC09

Figure 5 Comparison of model results for gauging sites 8BBC08 (left) and 8BBC09 (right)
All model results are calculated in terms of their error function, \( E(\exp \, d\alpha) \) (i.e., signal to noise ratio in percentage) which proposed and described in the earlier work by Vojinovic et al. (2004), and the standard coefficient of efficiency (CoE), refer Table 1. The accuracy of SVM model on the training data set was found to be below 5% when expressed in terms of \( E(\exp \, d\alpha) \) and between 5% and 10% on the test data set (refer Table 2).

The resulting MOUSE model errors from the validation phase (Figures 3 and 4) are the target for SVM data-driven model predictions.

From the overall comparison of MOUSE model results and complementary model results (refer Tables 1, 2; and scatterplots in Figure 5), it can be observed that the physically based model (i.e., MOUSE model) results are not as good as the results obtained from the complementary modelling approach (i.e., the results corrected by the SVM model) as they displayed biased values towards overprediction for the site 8BBC08 and underprediction for the site 8BBC09. By contrast, the complementary modelling approach generated much more accurate and unbiased predictions than those of the physically based model alone.

**Conclusion**

There is an inherent limitation in both, models and gauging capabilities. This fact together with the need to make more cost effective decisions that would lead to system improvements makes it difficult to apply a physically based model without some form of calibration. When dealing with computationally expensive models, there are several difficulties that limit model calibration capabilities. Therefore, the research to date shows that a physically based model can be improved and the associated uncertainty can be systematically reduced by the combined use of a data-driven model of errors. The results from a case study presented here indicate that a complementary modelling approach based on the physically based and data-driven models could be effectively applied to minimise the uncertainty of computationally expensive urban drainage models.

**References**


