9th International Conference on Urban Drainage Modelling
Belgrade 2012

Urban drainage models - making uncertainty analysis simple

Luca Vezzaro[1], Peter Steen Mikkelsen[1], Ana Deletic[2], David McCarthy[2]

[1] Department of Environmental Engineering (DTU Environment), Technical University of Denmark, Building 113, DK-2800 Kgs. Lyngby, Denmark, luve@env.dtu.dk; psmi@env.dtu.dk
[2] Centre for Water Sensitive Cities, Department of Civil Engineering, Monash University, Clayton, Victoria, Australia, david.mccarthy@monash.edu; ana.deletic@monash.edu

ABSTRACT

There is increasing awareness about uncertainties in modelling of urban drainage systems and, as such, many new methods for uncertainty analyses have been developed. Despite this, all available methods have limitations which restrict their widespread application among practitioners. Here, a modified Monte-Carlo based method is presented which reduces the subjectivity inherent in typical uncertainty approaches (e.g. cut-off thresholds), while using tangible concepts and providing practical outcomes for practitioners. The method compares the model's uncertainty bands to the uncertainty inherent in each measured/observed datapoint; an issue which is commonly overlook in the uncertainty analysis of urban drainage models. This comparison allows the user to intuitively estimate the optimum number of simulations required to conduct uncertainty analyses. The output of the method includes parameter probability distributions (often used for sensitivity analyses) and prediction intervals. To demonstrate the new method, it is applied to a conceptual rainfall-runoff model using a dataset collected from Melbourne, Australia.

KEYWORDS

Drainage, modelling, urban, uncertainty, sensitivity, parameter estimation

1 INTRODUCTION

The increased application of modelling tools in the urban drainage field has raised the awareness of the various sources of uncertainty influencing the modelling results and their reliability. A general framework for assessing uncertainty in urban drainage has been recently presented by Deletic et al. (2012), and the existing literature reports a great variety of studies investigating uncertainties of urban drainage models: including, among others, runoff predictions (Dotto et al., 2009), stormwater quality estimation (McCarthy et al., 2010; Lindblom et al., 2011) and integrated water quality management (Freni et al., 2009; Vezzaro et al. 2012). The choice of available methods for uncertainty estimation is wide, while the possibility of comparing the advantages and disadvantages of each approach is limited due to their intrinsic differences. However, Dotto et al. (2012) attempted to compare four different uncertainty analysis methods applied in combination with a simple runoff and water quality model. The study showed how the different methods, when using comparable assumptions, provide similar results. The final user (researcher, practitioner) can thus select the most appropriate approach according to a series of criteria, such as the computational requirement, the software availability, and the difficulty of implementation.

Despite the great research efforts in this field, all the available methods have limitations which hamper their widespread application among practitioners (Dotto et al., 2012). These limitations include (i) the complexity of the methods, (ii) the difficulties in their implementation, (iii) the need for expert knowledge, (iv) the number of underlying assumptions which are difficult to meet given other criteria, and (v) difficulties in interpreting both the model parameter sensitivity and the estimated model prediction bounds in a practical manner.

The most widely used approaches (by researchers and practitioners alike) all typically follow Monte-Carlo techniques; this is especially the case since these methods are simple and easy to implement. A good example of a simple and flexible approach is the Generalized Likelihood Uncertainty Estimation (GLUE; Beven and Binley, 1992). However, the results of this method are strongly affected by subjective choices made by the modeller (specifically on the acceptance threshold; Freni et al., 2008), which are difficult to justify and to convey in a practical context.

In this paper, a modified Monte-Carlo based method is presented which (i) helps to reduce the subjectivity inherent in these types of approaches, while also (ii) following tangible concepts and providing practical outcomes for urban drainage modellers. The method also introduces a performance criterion which (iii) assesses the model's performance by taking into account the uncertainty inherent in each measured/observed datapoint; an issue which is commonly overlooked in uncertainty analysis of urban drainage models. These three improvements will eventually lead to a wider application of uncertainty analysis tools, providing results which both practitioners and researchers can easily understand.

2 METHODOLOGY

2.1 Modified Monte Carlo approach for estimation of uncertainty bounds

The proposed approach expands the traditional Monte Carlo methods for uncertainty estimation by utilising the concept of observation coverage initially conducted by Blasone et al. (2008) in a GLUE framework. The proposed approach is outlined below (summarised in Figure 1):

1. Define the prior distributions (and ranges) of the parameters included in the model and generate *N* parameter sets. For example, in this study, we assumed all parameters came from uniform distributions, and using these distributions, 10,000 random parameter sets were generated.

2. Run the model with the *1st* parameter set and evaluate its performance using a chosen objective/criteria function which compares the predicted and the observed data. In this study, we chose the Nash-Sutcliffe efficiency criterion (*E*; Nash and Sutcliffe, 1970), mainly to facilitate comparisons with other studies. However, this function should be selected based on the objectives of the modelling exercise: for example, if the focus is on prediction of high flows, an objective function favouring peaks should be selected.

3. Repeat Step 2 for the remaining *N*-1 generated parameter sets; in our case study, we had 10,000 parameters sets, each with their respective *E* value. Rank the *N* parameter sets according to the value of the selected objective/criteria function, from high to low, such that k =1 has the best performance and $k = N$ has the worst performance; in our case, we ranked our 10,000 parameter sets using their corresponding E values, from high (k=1, highest E) to low (k=10,000, lowest E).

4. For each of the *J* observed datapoints (obs_j), estimate an uncertainty interval such that the datapoint falls between a lower value ($obs_{j,5}$) and an upper value ($obs_{j,95}$). For example, our 4th observed datapoint (i.e. obs_4) had a flow rate of 0.010 m^3/s and based upon known errors in flow sensors and the Law of Propagating Uncertainties (see McCarthy et al., 2008 for more information) we estimated that the uncertainty interval for this observation was between 0.003 m^3/s ($obs_{4,5}$) and 0.018 m^3/s ($obs_{4,95}$). In general, these observation errors will depend on the measurement method and can include both systematic and random errors.

5. Select the *K* most highly ranked model simulations (as per the ranking in Step 3) such that *K* timeseries of modelled values are selected, with $Mod_{k,j}$ representing the j^{th} value in the modelled dataset from the k^{th} ranked simulation (i.e. $Mod_{2,4}$ represents the 4th modelled datapoint in the 2nd best model). Use these simulations to create modelled 5th and 95th prediction bounds for each timestep *j*:

$$U_{5,j} = Mod_{k,j}|_{k=R_{5,j}} \quad (1)$$

$$U_{95,j} = Mod_{k,j}|_{k=R_{95,j}} \quad (2)$$

where $U_{5,j}$ and $U_{95,j}$ is the 5th and 95th percentile prediction bounds (respectively) for the j^{th} modelled datapoint, $R_{5,j}$ and $R_{95,j}$ are the rank-order position of the 5th and 95th percentile flow at timestep *j* out of the *K* currently selected simulations (respectively). In the situation where only a few simulations are selected (i.e. when K<4), $U_{5,j}$ is assumed to equal the minimum of the *K* selected simulations at timestep *j*, while $U_{95,j}$ is assumed to equal the maximum of the *K* simulations at timestep *j*.

6. Evaluate the intersection λ [%] between the prediction bounds and observation error interval as per the following formula:

$$\lambda = \frac{1}{J}\sum_{j=1}^{J} \mathrm{I}_j \text{ with } \begin{cases} \mathrm{I}_j = 0 \ for \ \ U_{95,j} < obs_{j,5}.OR.obs_{j,95} < U_{5,j} \\ \mathrm{I}_j = 1 \ for \ U_{95,j} \geq obs_{j,5}.AND.obs_{j,95} \geq U_{5,j} \end{cases} \quad (3)$$

where *J* is the number of observations and I_j indicates whether the modelled prediction bounds for the j^{th} observation are overlapping the associated observation error interval. The initial

concept of including observation data within the modelled prediction bounds proposed by Blasone et al. (2008) is thus extended by considering the observation measurement error.

7. Repeat Steps 5 and 6 with different values of K; K= 2, 3, 4…N; that is, start by including just the top two ranked parameter sets (i.e. K=2) and estimate λ, next include the top three ranked parameter sets (K=3) and estimate λ, …, finally include all tested parameter sets (K=N) and estimate λ. Ideally, all the observations should be intersected by the modelled prediction bounds (i.e. λ=100%) when the sample size of included simulations is sufficiently representative of parameter uncertainty. However, it is unlikely that this will occur due to the presence of other sources of uncertainty, especially model structural errors. The user is able to plot λ against the number of included ranked simulations (K) to gain a better idea about how many ranked simulations should be included for parameter prediction bounds. Indeed, this plot is expected to either have an obvious peak or will be characterised by a clear diminishing point of return curve. In either case, the user can objectively define the acceptance threshold γ, which represents the best information which can be obtained with the available data and model. This procedure has the benefit of expressing the threshold γ in a more tangible manner compared to other methods, such as GLUE, where γ is subjectively defined by the modeller.

8. Estimate model prediction bounds by including the K model simulations required to produce an intersection (λ) which equals the defined acceptance threshold (γ). Furthermore, the associated parameters sets used to estimate these prediction bounds could be used to create parameter distributions, which helps understand the uncertainty/sensitivity of the model's parameters.

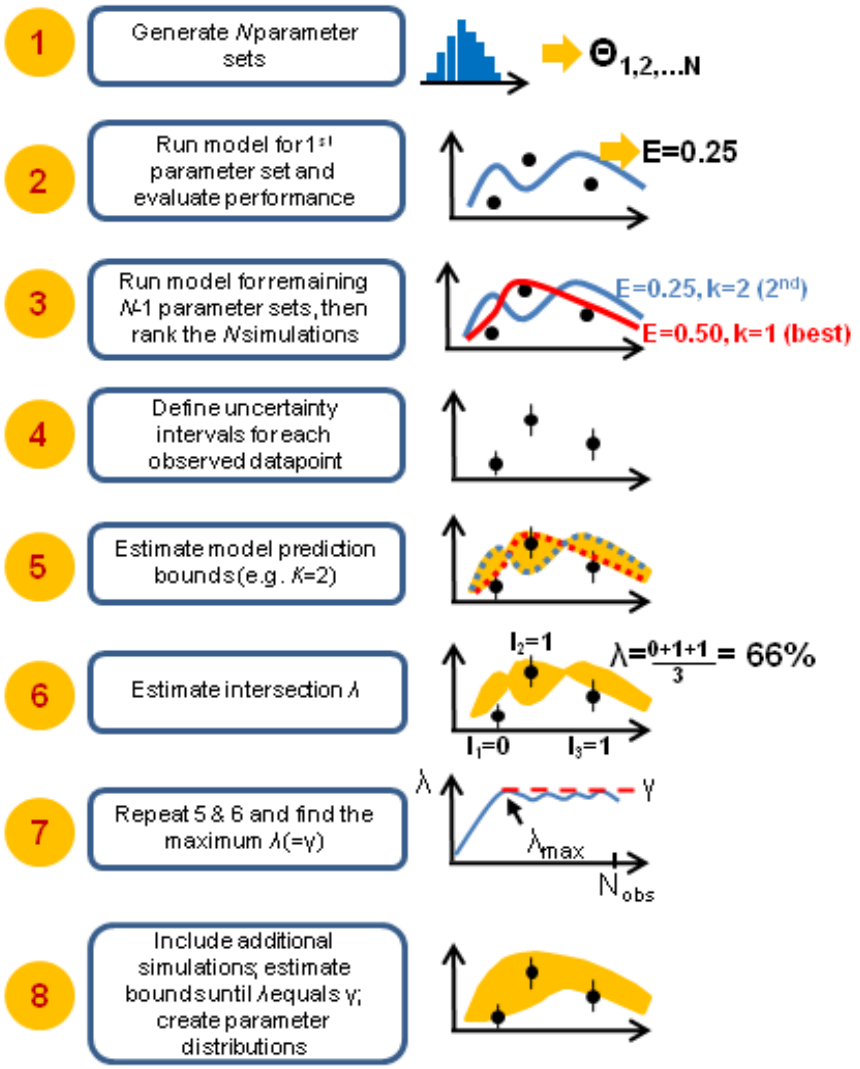

Figure 1. Steps of the proposed procedure for the estimation of model prediction bounds.

2.2 The case study

The proposed approach was applied using an existing rainfall-runoff model (MOPUS; McCarthy et al., 2011) to predict stormwater flows leaving an urbanised catchment in Clayton, Victoria, Australia (area 28ha, imperviousness 80%; for more information, see McCarthy et al., 2011). Doppler area-velocity flow meters were used to measure flows at the catchment's outlet, while a standard 0.2mm tipping bucket rainfall gauge was installed near the centroid of the catchment; both flow and rainfall were logged at one minute intervals.

MOPUS's rainfall-runoff model presents a modified version of Sim-Hyd (Chiew et al.,2002), which is widely used in Australian practice. MOPUS simulates processes on both pervious and impervious surfaces, contains two corresponding stores and a simple routine to route the runoff to the catchment outlet. There are six model parameters: (1) *IMP* – effective imperviousness of the catchment, (2) S_{impmax} – depression storage of the impervious surface, (3) $S_{pervmax}$ – capacity of the pervious soil store, (4) K_R – reservoir routing coefficient, (5) m – non-linear reservoir routing exponent and (6) *TOC* – time of concentration for the catchment. The model requires 1 minute rainfall intensity [mm/hr] as

input data and predicts flow rates at the catchment outlet [L/s]. For further information about the model, please refer to McCarthy et al. (2011).

Uniform distributions for the six parameters were used to generate N=10,000 parameter sets (upper and low parameter bounds were set based upon prior knowledge). To obtain observation error intervals, the uncertainty in the flow measurements was assumed to follow a linear relationship, where a ±75% error was assumed for low flows (between 0.003 m^3/s and 1 m^3/s) and a ±20% error was assumed for high flows (above 1 m^3/s - see McCarthy et al. 2008 for uncertainty estimation). This assumption roughly account for heteroscedasticity in flow measurements, where major relative errors are observed for low flows.

3 RESULTS AND DISCUSSION

3.1 Determining cut-off threshold

Figure 2 shows λ as function of the number of included ranked simulations (K) and indicates that the model prediction bounds fail to intersect all the observations, as the maximum value of λ is 97 %. This indicates that sources of uncertainty which were not included in the analysis are affecting the model (e.g. model structural error, input uncertainties, etc.). This situation can be regarded as common in urban drainage modelling, especially when dealing with stormwater quality models. This underlines the usefulness of the proposed approach in this context. Based on the data shown in Figure 2, the threshold γ was set to 97%; this intersection point occurred when the top 500 simulations were used to create the model prediction bounds. Looking at each simulation individually, the best ranked simulation obtained an E of 0.76, while the 500^{th} ranked simulation achieved an E of 0.58.

A second interesting result is the fact that the performance of the model (as assessed using intersection/λ) is not linearly proportional to the number of simulated parameter sets; indeed, a clear peak is observed in Figure 2. Initially, the inclusion of additional ranked simulations leads to an enlargement of the of the prediction bounds, causing a sharp increase in the number of observation error intervals intersected by the prediction bounds (i.e. increasing λ with increasing K). After this, an increasing number of poorer performing simulations are included in the analysis. As these simulations have been ranked by E which favours peak flows, the subsequent inclusion of more simulations causes a shift in the prediction bounds toward low flow values. This suggests that a great number of parameter sets evenly sampled across the parameter space, without considering the model performance, does not necessarily imply a good coverage of the observed values.

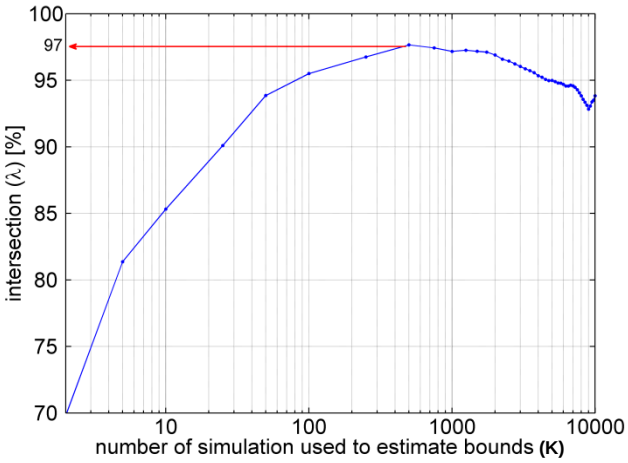

Figure 2. Behaviour of intersection as function of the number of model simulations used to create the model prediction intervals (N.B. the simulations have been ranked according to the Nash-Sutcliffe efficiency criterion and the highest ranked simulations are included first, with the poorest ones left until last).

It was hypothesised that using a different method of ranking would influence the pattern observed in Figure 2. Indeed, using an objective function which is less biased toward peak flows (for example, one which considers all observations equal) may produce a more log-linear relationship, meaning that every increase in the number of included simulations increases the intersection (λ) but with a diminishing point of return. Furthermore, other methods could also be considered for determining the optimal number of simulations to be used for the uncertainty analysis; for example, instead of maximising λ (as done in the above analysis), a cut-off intersection threshold of λ=90% could be adopted (in this case, the top 24 ranked simulations would be used; Figure 2). However, for this case study, we will continue to use the maximising approach (i.e. λ=97% and top 500 simulations).

3.2 Parameter distributions

A total of 500 simulations were necessary to reach the desired threshold of 97% intersection. The parameter sets for these top 500 simulations were used to create distributions of each of the six model parameters (Figure 3). In general, flat distributions suggest that the parameter could take any value within the given range and still adequately calibrate the model, indicating that the model's outputs are not significantly impacted by this parameter. Distributions with peaks suggest the opposite; that adequate calibration only occurs toward certain parameter values and indicates the model's outputs are sensitive toward this parameter. Using this, the outputs of the model appear to be sensitive to some model parameters (IMP, K, TOC) but not to others (S_{impmax}, $S_{pervmax}$). In fact, these results confirm previous research showing that, for similar models, parameters describing the percentage of directly connected impervious area (i.e. IMP) were important, while parameters describing the impervious and pervious area storage thresholds were not (Dotto et al., 2011).

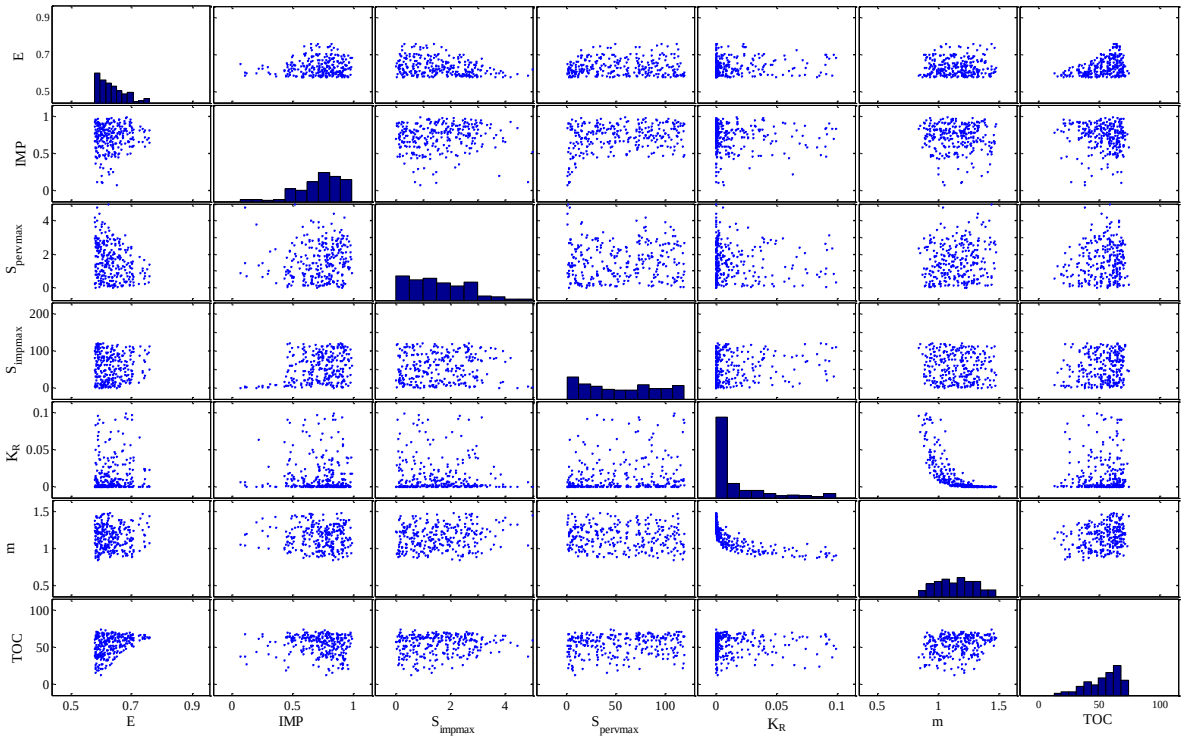

Figure 3. Matrix plot of efficiency scatter plots and posterior histograms (efficiency density in the upper left corner, parameters posterior likelihood densities in the other diagonal places) of model parameters.

3.3 Model prediction bounds

Based on the information obtained on the model performance, the model prediction bounds were estimated using the top 500 ranked simulation results. Examples of the estimated model prediction bounds for selected rain events are shown in Figure 4. One of the major benefits of the proposed approach is the possibility of expressing the estimated prediction bounds in a more intuitive manner. The bounds shown in Figure 4 are in fact the results of the simulations which intersect 97% of the observation error interval (i.e. they reflect the prediction bounds when 500 simulations are used, which was the number required to produce the maximum λ). This method of determining the number of simulations to include is easier to convey to users with limited knowledge in uncertainty estimation methods. It is especially more objective than developing prediction bounds based on a likelihood threshold (like in the GLUE methodology), which is an abstract value difficult to relate to physical variables.

Figure 5 illustrates how the model prediction bounds are generally smaller than the assumed observation interval errors for flows below 750 l/s; it is interesting to note that 99% of the measured low values are less than 750 l/s Uncertainty bounds dramatically increase for flows above 1000 l/s; this increase could be a reflection of the inherent limitations of the conceptual, spatially lumped, rainfall-runoff model which has been shown to have issues modelling peak flows (McCarthy et al., 2011).

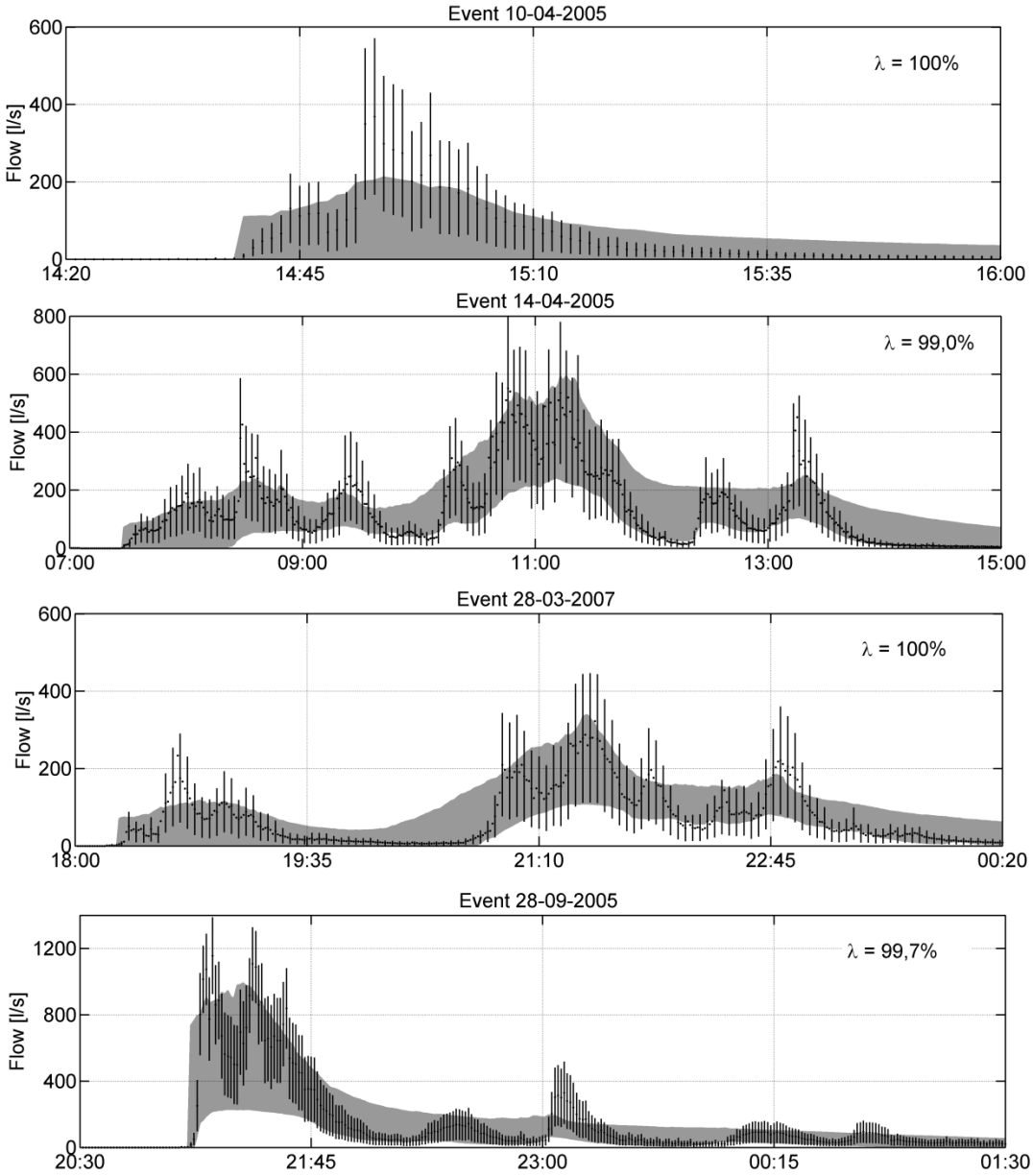

Figure 4. Examples of estimated model prediction bounds for four different rain events (grey area) and respective observed flow (observation error interval are illustrated by the black lines).

By including observation errors in the analysis, it is possible to assess that the performance of the model can be regarded as acceptable. This is another benefit of the proposed approach, as the inclusion of measurement uncertainty improves the understanding of modelling results and it provides a better background for an assessment of the overall performance of the model.

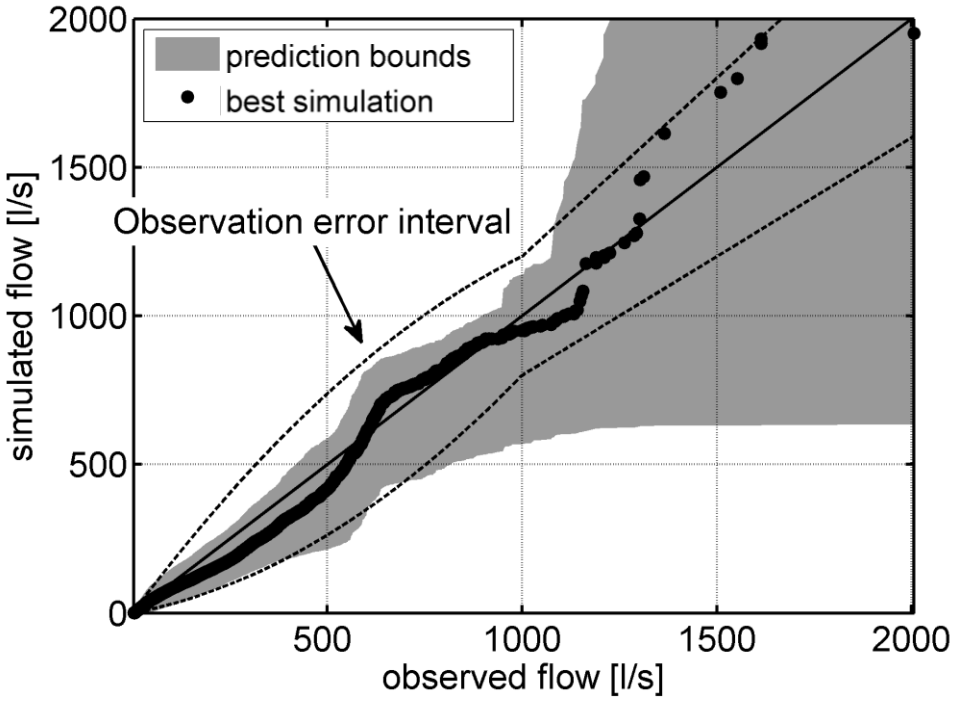

Figure 5. Q-Q Plot showing the best simulation result (black dots), observation error interval (dashed lines) and prediction bounds (grey shaded area).

4 CONCLUSION

This article presents a modified Monte Carlo approach for estimating model uncertainty bounds. This approach overcomes some of the major issues which limit the use of uncertainty estimation methods in practice by using practical and tangible concepts, as intersection between measurements and model prediction bounds. While not eliminating subjectivity, common to all the available methods for uncertainty analysis, this approach expresses in a simpler and more intuitive manner the various choices that a modeller makes during the process of uncertainty analysis. This explicit process increases the confidence of the user on the results of the uncertainty analysis and it enables their usage for practical applications.

The proposed approach also includes observation errors in the estimation of uncertainty bounds, thus incorporating this commonly overlooked issue in the quantification of model uncertainties. In fact, simulations are evaluated by considering also the quality of the available observation, leading to a "fairer" evaluation of model performances. The improvements provided by the proposed approach will eventually lead to a wider application of uncertainty analysis tools, providing results which both practitioners and researchers can easily create and understand.

5 REFERENCES

Beven, K. and Binley, A. (1992), Future of distributed models: Model calibration and uncertainty prediction, *Hydrological Processes*, 6(3), 279-298.

Blasone, R. S., Vrugt, J. A., Madsen, H., Rosbjerg, D., Robinson, B. A. and Zyvoloski, G. A. (2008), Generalized Likelihood Uncertainty Estimation (GLUE) Using Adaptative Markov Chain Monte Carlo Sampling, *Advances in Water Resources*, 31(4), 630-648.

Chiew, F. H. S., Peel, M. C. & Western, A. W. (2002) Application and testing of the simple rainfall-runoff model SimHyd. IN Singh, V. P. & Frevert, D. K. (Eds.) *Mathematical models of small watershed hydrology and applications*. Water Resources Publication, Colorado.

Deletic, A., Dotto, C. B. S., McCarthy, D. T., Kleidorfer, M., Freni, G., Mannina, G., Uhl, M., Henrichs, M., Fletcher, T. D., Rauch, W., Bertrand-Krajewski, J. L. and Tait, S. (2012), ssessing uncertainties in urban drainage models, *Physics and Chemistry of the Earth, Parts A/B/C*, 42-44, 3-10.

Dotto, C. B. S., Deletic, A. and Fletcher, T. D. (2009), Analysis of parameter uncertainty of a flow and quality stormwater model, *Water Science and Technology*, 60(3), 717-725.

Dotto C.B.S., Kleidorfer M, Deletic A, Rauch W., McCarthy D.T., Fletcher T.D. (2011), Performance and sensitivity analysis of stormwater models using a Bayesian approach and long-term high resolution data, *Environmental Modelling and Software* 26 (10), 1225-1239

Dotto, C. B. S., Mannina, G., Kleidorfer, M., Vezzaro, L., Henrichs, M., McCarthy, D. T., Freni, G., Rauch, W. and Deletic, A. (2012), Comparison of different uncertainty techniques in urban stormwater quantity and quality modelling, *Water Research*, 46(8), 2545 -2558

Freni, G., Mannina, G. and Viviani, G. (2008), Uncertainty in urban stormwater quality modelling: The effect of acceptability threshold in the GLUE methodology, *Water Research*, 49(8-9), 2061-2072.

Freni, G., Mannina, G. and Viviani, G. (2009), Uncertainty assessment of an integrated urban drainage model, *Journal of Hydrology*, 373(3-4), 392-404.

Lindblom, E., Ahlman, S. and Mikkelsen, P. S. (2011), Uncertainty-based calibration and prediction with a stormwater surface accumulation-washoff model based on coverage of sampled Zn, Cu, Pb and Cd field data, *Water Research*, 45(13), 3823-3835.

McCarthy, D.T., Deletic, A., Mitchell, V.G., Fletcher, T.D. & Diaper, C. (2008) Uncertainties in stormwater E. coli levels. Water Research, 42 (6-7), 1812-1824.

McCarthy, D. T., Deletic, A., Mitchell, V. G. and Diaper, C. (2010), Sensitivity analysis of an urban stormwater microorganism model, *Water Science and Technology*, 62(6), 1393-1400.

McCarthy DT, Deletic A, Mitchell VG, Diaper C (2011) Development and testing of a model for Micro-Organism Prediction in Urban Stormwater (MOPUS) *Journal of Hydrology* 409 (1-2), 236-247.

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

Vezzaro, L., Ledin, A. and Mikkelsen, P. S. (in press), Integrated modelling of Priority Pollutants in stormwater systems, *Physics and Chemistry of the Earth, Parts A/B/C*. 42-44, 42-51