QUICS: Quantifying Uncertainty in Integrated Catchment Studies

D.1.1 Report on uncertainty frameworks

D.4.2 Report on application of uncertainty frameworks, potential improvements

Lead: USFD and TUD
Revision: 29/05/2017

Report Details

Title: Report on uncertainty frameworks and Report on application of uncertainty frameworks, potential improvements.

Deliverable Number (If applicable): 1.1 and 4.2

Author(s): Ambuj Sriwastava (AS) and Antonio Moreno (AM)

Document History

Version	Date	Status	Submitted by	Checked by	Comment
v 0.0	18/04/2017	Structure	AM	AS S. Tait	
v 1.0-1.1	12/05/2017	Full Draft	AS, AM	S.Tait	
v 1.2	21/05/2017	Full draft	AM	A Schellart	
v 1.3	23/05/2017	1, 4.2, 5, Ex. Summary Final version	AS, AM	J. Langeveld, F. Clemens	

Acronyms and Abbreviations

CSO	Combined Sewer Overflow
GAMU	Global Assessment of Modelling Uncertainties
GLUE	Generalized Likelihood Uncertainty Estimation
GSA	Global Sensitivity Analysis
LHS	Latin hypercube sampling
QUICS	Quantifying Uncertainty in Integrated Catchment Studies
SCEM-UA	Shuffled Complex Evolution Metropolis algorithm
UPA	Uncertainty Propagation Analysis
WWTP	Wastewater Treatment Plant

Acknowledgements

This project has received funding from the European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement no 607000.

Executive Summary

This report combines deliverables D1.1 and D4.2 and focuses on uncertainty frameworks and their applications in urban drainage modelling.

We review the research found in literature dealing with uncertainty quantification or analysis of water quantity and quality simulations using urban drainage models followed by a critical analysis of uncertainty framework proposals by Refsgaard et al. (2006), Refsgaard et al. (2007) and Deletic et al. (2012) by discussing the merits and limitations of these frameworks.

This report further suggests potential future improvements which should be incorporated in upcoming uncertainty frameworks. These improvements have been suggested by addressing the limitations of the three frameworks through latest research found in the literature on uncertainty analysis. In addition, certain merits which are not common to all the three frameworks are also suggested as an area of potential improvement.

Contents

1. Uncertainty in urban drainage and environmental modelling

Decision making in urban drainage infrastructure management is mainly performed by urban drainage managers and engineering professionals. The definition of an efficient decision often strongly revolves around strict compliance with regulatory/policy guidelines while satisfying organisational budget constraints. Regulatory authorities set certain environmental thresholds for water utility companies such that the combined sewer overflows (CSO) and the discharges from wastewater treatment plants (WWTP) managed by the utility companies must conform to the regulations. These utility companies face the risk of paying a penalty and/or negative publicity if they breach the permitted environmental thresholds for the release of untreated and treated wastewater into receiving waters in the natural environment. Decisions on measures to control untreated sewage discharge to the receiving surface water bodies are often based on simulation models which are mathematical representations of the physical processes involved. However, these numerical models should also provide uncertainties accompanied with the model predictions because any unaccounted uncertainty in these model predictions could have a significant effect on the outcome of the decision making process of water utilities that manage the drainage infrastructure. An uncertainty analysis provides information on the limitation of models and simulation methodologies employed to predict the physical processes.

The modelling uncertainty can be classified into two broad categories, aleatory and epistemic (Kiureghian & Ditlevsen, 2009). Aleatory uncertainty refers to the inherent randomness in any physical process while the epistemic uncertainty arises from a lack of knowledge about the physical process in question. Kiureghian & Ditlevsen (2009) argue that the categorization of the uncertainties in any modelling study depends on the choices a modeller makes. Typically, a modeller should categorise uncertainties as aleatory uncertainties when they cannot be reduced by improved knowledge through additional data collection or model structure or calibration improvement. In contrast to aleatory uncertainty, epistemic uncertainty is associated with the assumptions and simplifications made while formulating the mathematical equations to represent the physical processes. As a result, epistemic uncertainty can be reduced by various measures such as enhanced calibration using better or more measurements and improvement of the underlying mathematical relationships.

In this report, we define uncertainty, wherever mentioned from here on, as epistemic uncertainty arising from model simulations. Among the various sources contributing to this type of uncertainty in model calculations, Refsgaard et al. (2007) classified the uncertainty in modelling into model input data, model parameters and model structure uncertainties whereas, Deletic et al. (2012) expressed the uncertainty in the calibration of model parameters as a separate source of modelling uncertainty. In this report, the classification of uncertainty proposed by Refsgaard et al. (2007) is used by treating any calibration uncertainty as a subset of model parameter uncertainty.

The major input to urban drainage models is rainfall whose uncertainty can have a significant effect on the overall model uncertainty (Hoppe, 2008). Rainfall data uncertainties are characterised by spatial and temporal variability and the error in data measurement. Thorndahl and Willems (2008) looked at the effect of the uncertainty in rainfall estimation on the failures of urban drainage system where they looked at localised flooding failures by modelling one manhole at a time. It was found that synthetic rainfall events generated using depth and duration of measured rainfall events were sufficient to simulate the occurrence of manhole surcharge and flooding events. Freni et al. (2010) studied the impact of rainfall time resolution on urban water quality assessments and concluded that rainfall temporal resolution had a greater effect on water quality sub-models than the structure of the water quantity sub-models due to the dependency of the wash-off sub-model on the rainfall intensity. However, when a lower temporal resolution of rainfall is applied, parameter calibration compensates this lack of information in rainfall data by adjusting the parameter values to reflect the real world behaviour. The physical significance of the parameters might have been lost as a result of this forced calibration adjustment of parameter values.

Most water quality and hydraulic models have model parameters which are required to be empirically estimated. Model parameter uncertainty is the uncertainty associated with the estimation of each parameter in the model. Input data such as roughness values, geometrical parameters which remain fixed during the single model run are grouped as model parameters in this report. Often these parameters are estimated using monitoring data; however, if there is no evidence or data available and the parameter value has to be assumed, this may result in an unknown and possibly higher degree of uncertainty. In some cases, the desired parameter is estimated by combining two or more parameters from different sub-models because there is no direct measure for its estimation. The uncertainty associated with the parameter of interest consists not only of the uncertainties in the values of sub-model parameters but also of how these uncertainties propagate between sub-models. For example, roughness in sewer pipes can be estimated using the Colebrook-White equation which uses geometrical and flow characteristics in the pipe (Swaffield & Bridge, 1983). The potential uncertainty in these characteristics affect the estimated values of pipe roughness.

The steps involved in uncertainty analysis are directly linked to the estimation of the values of parameters. Standard statistical methods result in a point estimate and a measure of precision around this point estimate, for example, a 95% confidence interval. However, within a multivariate framework, an additional measure 'covariance' is also generated which reflects the relationship among parameters. Representation of the parameter uncertainty depends on the method applied. In a deterministic uncertainty analysis, the uncertainty is expressed as the possible range based on the belief about the parameter, however, in probabilistic uncertainty analysis, a distribution is produced by specifying the distribution parameters which provide the likelihood of the parameter values. Uncertainty distribution around the 'true' parameter value (expected value) can be expressed through either Bayesian or frequentist approach. It is suggested that the assumptions to specify the probability distribution should follow standard statistical methods, for example, one may use a beta distribution for binomial data, or a gamma or lognormal for the right skewed parameter (O'Hagan et al., 2006; Briggs et al., 2012).

Sometimes there is very little information available about the parameter because either there is no data or there are not many studies related to its estimation. In such cases, a conservative approach can be followed by relying upon expert opinion and the uncertainty can be explained through an appropriate range of possible estimates elicited from each expert (Garthwaite, Kadane, & O'Hagan, 2005). If formal elicitation is not feasible, a wide uniform distribution can be assumed to account for the uncertainty around this parameter. There has been a lack of studies where prior parameter distribution was estimated from field measurements representing the measured behaviour of a parameter. Studies such as Freni et al. (2008), Korving et al. (2002), and Vezzaro et al. (2013) proceeded with the prior assumption that the input and model parameters followed uniform or normal distributions which may not reflect reality. These distributions do not account for uncertainty at the extremes or beyond the range specified. In addition, continuous distributions which give a reliable estimate of uncertainty around the expected parameter value should be preferred. Instead of using a triangular distribution (Iooss & Lemaître, 2014) while performing three-point estimates, it is recommended to use PERT distribution which is a special case of a Beta distribution (Benke, Lowell, & Hamilton, 2008). The distribution is specified by assigning maximum, minimum values and the mode which is the most likely value. The scale parameter λ for the height of the distribution is taken as 4 by default (Vose, 2000). The PERT distribution has a distinct advantage over a triangular distribution because it can be changed from a symmetrical distribution to a skewed distribution by changing the mode. It can be used instead of a normal distribution when the parameters take values within a specified range and the extreme values are not important (Benke, Lowell, & Hamilton, 2008).

The choice of methods to represent the uncertainty in input data and model parameters in model simulations depends on the computational requirements and complexity in implementing such methods (Dotto et al., 2012). Monte Carlo simulation is one such method which is non-intrusive, meaning it does not require modifications to the model structure. However, Monte Carlo is not easy to implement for computationally expensive models, hence this technique is usually applied to simplified models. The Monte Carlo method involves repeated simulations with samples of the selected input/ model parameters drawn from the parameter space. This results in a mapping of input/model parameters to the desired model output. To cut down the required number of simulations, Latin hypercube sampling (LHS) can be used instead of random sampling because the LHS method results in a better convergence than random sampling approach for models which require long simulation time and it has the ability to generate samples representing the entire parameter space (Helton & Davis, 2003). Korving et al. (2002) propagated the uncertainty in model parameters to simulate combined sewer overflow volume using Monte Carlo simulations where the sewer system was simply represented as a reservoir connected to an external weir and a pump. Alternately, model reduction techniques have been used for complex models, for example Schellart et al. (2010) used a response database for model reduction before applying Monte Carlo simulations for uncertainty propagation in an integrated catchment model which comprised a rainfall generator, a simplified hydrological model, a computationally expensive sewer hydrodynamic model and a simple river impact model to estimate water quality failures in a receiving watercourse over an extended time period. Model reduction is an approximation of a complex model and introduces additional uncertainty in the realisation of the physical system on top of the uncertainty in the complex model.

A further approach to quantify the uncertainty of the output for complex models is to select only a small subset of dominant model inputs and parameters which can explain the model output variance for uncertainty analysis or parameter estimation (Wainwright et al., 2014). Key processes to be included in the uncertainty analysis were identified by ranking all the parameters using sensitivity analyses. This reduced the computational cost by only including the most significant parameters in the uncertainty analysis. There are several methods proposed in the literature for performing sensitivity analysis which can be broadly classified as Global Sensitivity Analysis (GSA) or Local Sensitivity Analysis (Saltelli et al., 2000). Local sensitivity analysis is performed to study the effect of small input perturbations on the model output and has been performed around a point in the parameter space whereas a GSA has been performed over the whole parameter space of model inputs considered for study (Gamerith et al., 2013; Iooss & Lemaître, 2014; Borgonovo & Plischke, 2016). Global sensitivity analysis is performed using different approaches e.g. Standard regression coefficients (SRC) (Saltelli et al., 2008), Extended-FAST method (Saltelli et al., 1999), Morris Screening method (Morris, 1991), Sobol' indices (Sobol, 2001). Although Vanrolleghem et al. (2015) preferred Extended-FAST over SRC and Morris Screening method for water quality simulation in the catchment and sewer network, they concluded that for water quantity simulations all three methods Extended-FAST, SRC, and Morris Screening produced similar results. Kroll et al. (2016) further demonstrated that Morris Screening performed on par with Extended-FAST while ranking the influence of parameters on CSO volume. It can be concluded that Morris Screening is an appropriate method for performing GSA because it is computationally cheap and it performs at a level with other available more computationally expensive methods.

Model structure uncertainty corresponds to the inaccuracy in the simulation used to represent the process. This uncertainty can also arise from inappropriate methods to define the boundary conditions, errors in construction plan databases which are transferred to the geometrical model structure of pipelines (Clemens, 2001) or the choice of numerical solution techniques (Deletic et al., 2009). These different uncertainties get coupled with each other and propagate through to the model outputs (Swayne et al., 2010).

2. Uncertainty quantification frameworks

Simulations from urban drainage and water quality models are subjected to significant levels of uncertainty. This is due to the inherent characteristics of those systems; complex processes represented with limited knowledge, relationships calibrated with small data sets, and linked simulations carried out over a wide range of spatiotemporal scales. During the last decade, there has been a clear agreement within the environmental engineering and urban drainage scientific community regarding the need to quantify and communicate modelling uncertainties. This was further recognised by the development of several frameworks directed to provide the community with a common uncertainty language and repository of methods. Refsgaard et al. (2006) brought the focus of uncertainty analysis to the quantification of model structure errors. This is recognised as a significant uncertainty source in most water quality modelling applications. Refsgaard's work discusses the process to assess model structural uncertainty in cases in which data are not available. Refsgaard et al. (2007) postulated a more comprehensive framework for uncertainty tractability in the environmental modelling process. Several methods were qualitatively presented and classified depending on their relation with the model conceptualization stage and the type of source uncertainty. This work discussed uncertainty analysis not as an additional product to be added to the finished model, but rather as a process that should happen in parallel to the problem identification, model design, build and operation.

Deletic et al. (2012) presented a **G**lobal **A**ssessment of **M**odelling **U**ncertainties (GAMU) produced by the IWA/IAHR Joint Committee on Urban Drainage. This was an effort to provide urban drainage modellers with a unifying terminology and understanding on uncertainty analysis. Within GAMU, modelling uncertainties were classified in three main sources; I) Model input uncertainties, II) Calibration uncertainties and III) Model structure uncertainties. The links between uncertainty sources are depicted in Figure 1.

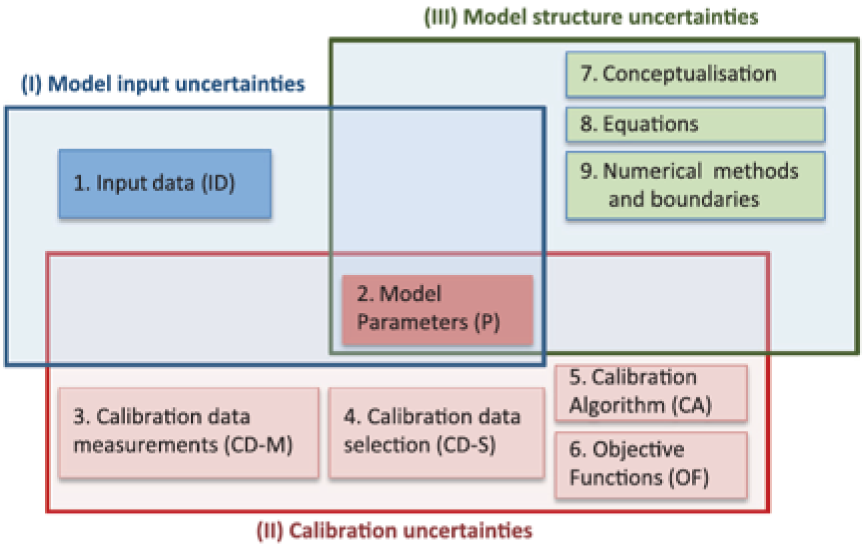

Figure 1. Key sources of uncertainties in urban drainage models. From Deletic et al. (2012).

The GAMU framework focuses on statistical uncertainties at the calibration and prediction phases of the model, thus diverting from the process established by Refsgaard et al. (2007).

The GAMU framework defined a procedure for uncertainty analysis composed of three steps:

1. **Selecting analysis tools and data sets to minimise uncertainties:** Each type of model/objective influences the requirement/suitability of calibration tools (CA), objective functions (OF) and the calibration-validation dataset. The prior selection and justification of them is directed to minimise uncertainties in the modelling process.
2. **Creating probability distributions of model parameters while simultaneously propagating all data uncertainties**: The GAMU scheme recommends proceeding with the calibration/inference of parameters while simultaneously propagating uncertainties due to input data. The calibrated model is then used to determine model predictive uncertainties. Total predictive uncertainties are assessed using the residuals of the validation process.
3. **Comparing different model structures for similar scenarios:** Uncertainties due to model structure could only be assessed by comparing the performance of different model conceptualisations under the same conditions. GAMU

appoints the modeller to propose several model structures which total uncertainty is compared by applying step 1 and step 2.

3. Advances in uncertainty analysis methodology

After the presentation of the GAMU and Refsgaard et al. (2007) frameworks, several studies appeared addressing some aspects of uncertainty quantification in urban drainage and water quality modelling. This is not an exhaustive list yet it aims to provide the reader with an overview of the research interests after the initial frameworks.

Addressing parameter inference/calibration tools

Following the release of the GAMU framework, Dotto et al. (2012) presented a comparative analysis of four techniques for parametric inference/calibration in urban drainage. This work discussed the suitability of GLUE, SCEM-UA under GLUE, AMALGAM and MICA to fit parameter vectors to observed data and to evaluate parametric correlations. Algorithms were classified according to their ability to identify parameters values, correlation, availability and the required user skills. This study refers to the first step of the GAMU framework which discusses how tools for uncertainty analysis should be carefully chosen in order to minimise biased outcomes.

It should be noted though, that the applicability of GLUE-based techniques and non-formal likelihood measures are still subjected to debate in the literature (Stedinger et al. 2008; Freni et al. 2008; Mantovan and Todini, 2006 and Beven et al. 2008). Formal Bayesian methods are discussed in the GAMU framework as one of the main ways to infer probability density functions of parameter spaces. However, the influence of likelihood selection and posterior validation was not discussed. Schoups and Vrugt, (2010) proposed a generalised likelihood relaxing the assumptions of residual gaussianity, autocorrelation and homoscedasticity. Del Giudice et al. (2013) proposed a statistical description of bias in a likelihood function for urban drainage in which autocorrelation of errors are taken into account. Freni and Mannina (2012) explored the influence of the use of Box-cox transformations in Bayesian inference for water quality modelling.

Uncertainty analysis and propagation

Willems (2008) and Willems (2012), presented a methodology and application to quantify the contribution of different sources of uncertainty in urban drainage models. This method is based on a variance decomposition approach, which separates the total variance presented by the residuals in different characteristic sources; Input, the parameter (expert elicited), and structural uncertainty (quantifiable uncertainty). Nevertheless, the variance decomposition approach described is subjected to several assumptions. First, it requires a homogeneous variance of the model-observations residuals. This is seldom found in real applications, thus Box-cox transformation was used in order to stabilise residual variance (reaching homoscedasticity). Secondly, variance decomposition relies strictly on the independence of error terms. This fact was pointed by Freni and Mannina (2010) by comparing the relationship between the sum of partial variances and total variance. This difference indicates non-independency of the error terms. They concluded that the applicability of this method is increasingly reduced when propagating uncertainties downstream of the sub-model chain, where correlation amongst parameter uncertainties appears to be higher. The variance decomposition method provides a valuable source of knowledge by pointing out the relative importance of each contributor when the assumptions are met.

Most of the above-mentioned studies have considered the probabilistic representation of uncertainty in different model components. Fu et al. (2011) argued that the type of uncertainty in urban drainage modelling is quite broad and cannot be expressed adequately by probabilistic measures alone. They proposed a mathematical framework which facilitated the inclusion of vagueness in expert knowledge about model parameters using fuzzy sets and imprecise rainfall data using probabilities. This framework suggests the use of imprecise probabilities for input rainfall data when more than one probability distributions fit the data. For model parameters, data scarcity to characterise their uncertainty is widely cited as one of the major problems. This framework includes a fuzzy set representation for such model parameter uncertainty. These two different types of uncertainty representations are combined in a joint random set using two methods, discretization and Monte Carlo method where the latter was found to be more computationally efficient.

Input and measured data

The GAMU frameworks proposed the definition of error models for input and measured data, which should be propagated together in the total uncertainty analysis. Dotto, et al. (2014) explored the impacts of data uncertainty on urban stormwater models, concluding that although random errors could easily be filtered by the parameterization, systematic errors had a significant impact. Biases in the measured data led to distinct parameter distributions. Del Giudice et al. (2016) used a stochastic model for rainfall inputs in a water quantity urban model. Outflow data in the sewer system was used to constrain the characteristics of the rainfall stochastic process. Uncertainties in model inputs are often represented by a correcting factor (eg. rainfall multipliers) which limits the variability of the measured time-series to places where measurements are not zero. This prevents rainfall measurement failures to be represented in the input uncertainty representation. Findings from Del Giudice et al. (2016) indicate that this extra flexibility can help to better refine the rainfall-runoff and urban drainage model parameters. Although it can be arguable that providing too much flexibility to an input process should not lead to issues of process identifiability. Sun and Bertrand-Krajewski (2013) analysed the effect on water quantity modelling uncertainties of small urban catchments of rainfall input, concluding that the contribution of rainfall input to uncertainties is relatively minor.

4. Discussion and future prospects

The following section is framed as a discussion on the possible shortcomings and gaps identified in the current frameworks for uncertainty assessment. Deletic et al. (2012) serves as a base to reach a common definition of uncertainty terminology. It correctly discusses the need for model identification, calibration and validation. However, the applicability of the current framework for real cases is probably limited. Additionally, some concepts might lead to confusion within the scientific community and so might need further revision.

(1) Uncertainty Quantification vs Uncertainty Analysis

There is still certain confusion on the use of terminology for uncertainty analysis, propagation and quantification. Those terms are still used interchangeably although they address fundamentally different (but not exclusive) problems. The GAMU framework dealt only with the propagation of statistical uncertainties. Thus directing the modeller to the following process:

1. Define uncertainty distributions for input data sources
2. Identification of important input/model parameters through a sensitivity analysis
3. Infer parameter probability density functions from the comparison of a given model structure vs measured data (inverse modelling).
4. Propagation of uncertainties in the selected input/model parameters, model structure and input data for a validation data set (forward modelling).
5. Comparison of total uncertainty propagation from two or more different model conceptualisations.

This, when carried out appropriately could provide an estimation of the model predictive uncertainties. The quantification of total predictive uncertainties is recognised as a necessary exercise for a transparent communication of model results. However, this quantification alone will hardly provide any additional knowledge to the modeller beyond an assessment of how reliable a model is in a given context (boundary conditions and input data). Perhaps even of more importance from the scientific and operational perspective is the case of model structure improvement. Every modeller seeks to increase its capability to reproduce reality; this is done by applying all available theoretical and experimental knowledge to the service of a good conceptualization of the phenomena, often this imperfect knowledge is translated to a structural misfit (often referred as epistemic uncertainty and model structural error). With the proposition of the GAMU only a comparison of the relative performance between two already conceptualised strategies can be performed, requiring a “blind” updating of the model structure. GAMU suggests the need to perform sensitivity analyses to identify important input and model parameters. However, this represents an inefficient strategy to lead to model improvement, as long as the detection of individual uncertainty contributors remains unaddressed.

As mentioned previously, Willems, P. (2008) and Willems, P. (2012) discussed a methodology to decompose the total uncertainty in different contributors, which should be further tested and developed. Other approaches as in Reichert and Mieleitner (2009), focused on using time-dependent parameters as a proxy to detect temporal windows of structural mismatch, pointing therefore at examining particular physical processes.

In the field of hydrology, Gupta et al. (2008) discussed the need for developing a new philosophy of model evaluation, as a “diagnostic approach”. They acknowledge the poor capability of current data assimilation methods to pin-down and detect incorrect individual structural definitions. Formal likelihood-based inference methods fail to produce this information since they evaluate the measurement n-dimensional informative space through a one-dimensional criterion, which by definition generate an ill-posed system. The alternative proposed is based on the extraction of the relevant patterns and signatures from data. If those pattern signature statistics are carefully chosen to be discriminatorily sensitive to certain process-based parameters, this brings the possibility to detect which parameter cluster is not able to be fitted to the data, pointing then to a possible main contributor to the epistemic uncertainty. An early adoption of this idea can be found in Vrugt and Sadegh (2013) and Sadegh and Vrugt (2014), in which an implementation of this philosophy is done for an Approximate Bayesian Computation (ABC) scheme. Although many issues are still to be addressed:

- how to deal with input uncertainties, the need to include an error-generating process (not included in Sadegh and Vrugt (2014))
- how to perform the correct detection of epistemic errors
- how to define a sufficient number of pseudo-orthogonal parameter signatures which provide a diagnostic information extraction

Further research on uncertainty analysis should focus on uncertainty source decomposition and its use to provide valuable information to the modeller, incrementing the overall knowledge on the processes studied, directing monitoring network design and understanding the characteristics of the particular system.

(2) Assessment of uncertainties

The second limitation of the uncertainty quantification framework proposed by Deletic et al. (2012) is that it recommends uncertainty analysis as a standalone and separate process than the usual modelling workflow. Model calibration using observed data is nothing but correcting the model in order to generate predictions which are as close as possible to the real world. However, it does not provide any information about the accuracy of the model predictions such as by which amount the model predictions would deviate from this ‘corrected’ model prediction and what is the likelihood of such deviations. An uncertainty analysis acknowledges the limitation of the model predictions by providing an error band with the corresponding likelihood of model predictions. This additional information gives more confidence to the modellers on their model performance which would further facilitate a better-informed decision making. Therefore, it is recommended that uncertainty analysis should be treated as a process which runs parallel and is integrated into the model definition, building, calibration, and validation stages.

For example, Deletic et al. (2012) recommend that parameter uncertainty should be quantified using observed data through Bayesian inference which provides posterior distribution for such parameters by tuning the prior parameter distribution using the observed data. This process indirectly addresses the discussion on the validation of quantified uncertainty in model predictions against the observations in the real world. Even if there is limited data available about parameters, expert elicitation or literature references can be used as a prior in the Bayesian inference so that the resulting posterior distribution encompasses the expert knowledge as well as the added information from the limited available data. Therefore, instead of a traditional model calibration process giving ‘corrected’ parameter values, a Bayesian inference may be applied to generate a posterior probability distribution of model parameters along with the information about the correlation between the model parameters followed by uncertainty propagation of these uncertainties. This will ensure the consideration of the real world observations into the uncertainty analysis process and will also ensure that local catchment and environmental conditions are well reflected into the uncertainty predictions.

In cases where there is vagueness in the available information about the uncertainty in model components, the framework proposed by Fu et al. (2011) should be integrated into the wider uncertainty analysis framework to characterize the uncertainty in different sources using fuzzy sets, interval based probabilities, imprecise probabilities etc. depending on the type of uncertainty.

(3) When model inference is not viable

The procedure proposed in the GAMU framework is often not applicable to full-scale urban water quality models in real applications. The methodology depicted relies on the possibility to infer converged probability distributions of the sensitive parameters.

In most cases, this requires a prohibitive number of model runs, which seldom will be affordable for a modeller in a real sale study. Uusitalo et al. (2015) presented a comprehensive review of strategies for uncertainty analysis in slow paced models. Some examples with direct application to hydrology and urban drainage could be listed as:

1. Conceptual model schemes: Langeveld et al. (2013)
2. Model data-driven emulation: By interpolating a model output metric like in Schellart et al. (2010), emulation of the posterior probability density function or likelihood (Dietzel and Reichert, 2012) which can speed up sampling from the likelihood.

   Often, the uncertainty analysis interest lays in a dynamic model response (output time-series). This presents the added difficulty of having a multi-variable process with heavy autocorrelation structures (time structure). Some authors have proposed strategies to deal with those cases; Carbajal et al. (2017) compared a physically based emulator (which merges a simplified physical model and error interpolation) with a fully data-driven emulator (based on Gaussian process interpolation of a decomposed time-series) for an urban drainage case. Conti and O'Hagan (2010) presented three strategies to deal with multi-output or dynamic simulators a multivariate Gaussian process, ensembles of single-output emulators or the use of time as an extra dimension. Emulation of time-dynamic processes can also be proposed by the use chaos polynomial expansions as in Xiu and Karniadakis (2003). This technique can also provide sensitivity analysis result, therefore, both processes SA and emulation can be done under the same model sampling scheme. However, the integration of dynamic input uncertainties in emulation based problems is still not readily solved the problem, limiting the process to parametric uncertainty propagation.
3. Parallel model sampling algorithms: which speeds up the convergence of posterior exploration (Laloy and Vrugt 2012; Dejanic et al. 2017)
4. Postprocessor based strategies; slow models which are used operationally (eg. flood forecasting, resources estimation etc.) often present large databases of measurement and model results. This information can be used to estimate the model behaviour against new scenarios. If model residual

patterns for similar conditions are present in the database, this can be directly used for estimation of the current uncertainties (eg. Wani et al. 2017). Those strategies are reserved for the case in which the interest is only on approximating total predictive uncertainties.

5. Conclusion

The objective of this document is to discuss the validity of currently available frameworks for uncertainty analysis in urban drainage and water quality modelling based on their impact in the scientific bibliography and to identify gaps for improvement. Some issues still remain poorly addressed and they require further attention, some of the most relevant ones can be listed as:

- Uncertainty analysis should be discussed as a process in parallel to the modelling exercise. Aiming to evaluate the degree of uncertainty of model outputs, the decomposition of source contributors, and to use this knowledge to direct needs in monitoring data acquisition and model structure improvement.
- There is a lack of a common methodology to decompose uncertainties. The studied frameworks only discussed quantification of total uncertainty, neglecting the separation of contributing sources.
- The methodologies proposed for inference of parameter probability distributions will often not be applicable to many integrated catchment studies (computationally demanding). Further research in model emulation and model reduction is required.
- Future frameworks should provide the capability to include different ways of representing uncertainty in modelling on the similar lines to the framework proposed by Fu et al. (2011). This ensures a more suitable representation of the available knowledge about the uncertainty.

Nevertheless, it is necessary to mention that the 3 frameworks discussed in this document represented an excellent starting point for discussion in a field in which uncertainty analysis is still not of wide use in practice or research.

6. References

Anne Dietzel, Peter Reichert. (2012) Calibration of computationally demanding and structurally uncertain models with an application to a lake water quality model, Environmental Modelling & Software, Volume 38 , Pages 129-146, ISSN 1364-8152, https://doi.org/10.1016/j.envsoft.2012.05.007.

Benedetti, L., Batstone, D. J., De Baets, B., Nopens, I., & Vanrolleghem, P. A. (2012). Uncertainty analysis of WWTP control strategies made feasible. *Water Quality Research Journal of Canada, 47*(1), 14. doi: 10.2166/wqrjc.2012.038

Borgonovo, E., & Plischke, E. (2016). Sensitivity analysis: A review of recent advances. *European Journal of Operational Research*, *248*(3), 869–887. http://doi.org/10.1016/j.ejor.2015.06.032

Briggs, A. H., Weinstein, M. C., Fenwick, E. a. L., Karnon, J., Sculpher, M. J., & Paltiel, a. D. (2012). Model Parameter Estimation and Uncertainty Analysis: A Report of the ISPOR-SMDM Modeling Good Research Practices Task Force Working Group-6. *Medical Decision Making*, *32*(5), 722–732. http://doi.org/10.1177/0272989X12458348

Carbajal, J. P., Leitão, J. P., Albert, C., & Rieckermann, J. (2017). Appraisal of data-driven and mechanistic emulators of nonlinear simulators: The case of hydrodynamic urban drainage models. Environmental Modelling & Software, 92, 17-27.

Clemens, F. H. L. R. (2001). *Hydrodynamic Models in Urban Drainage: Application and calibration.* (PhD Thesis), Technische Universiteit Delft, Delft.

Conti, S., O'Hagan, A. (2010) Bayesian emulation of complex multi-output and dynamic computer models, Journal of Statistical Planning and Inference, Volume 140, Issue 3, Pages 640-651, ISSN 0378-3758, http://dx.doi.org/10.1016/j.jspi.2009.08.006.

Dejanic, S., Scheidegger, A., Rieckermann, J., and Albert, C. Appraisal of jump distributions in ensemble based sampling algorithms 2017, Paper in preparation.

Del Giudice, D., Honti, M., Scheidegger, A., Albert, C., Reichert, P., & Rieckermann, J. (2013). Improving uncertainty estimation in urban hydrological modeling by statistically describing bias. Hydrology and Earth System Sciences, 17(10), 4209-4225.

Del Giudice, D., C. Albert, J. Rieckermann, and P. Reichert (2016), Describing the catchment-averaged precipitation as a stochastic process improves parameter and input estimation, Water Resour. Res., 52, 3162–3186, doi:10.1002/ 2015WR017871.

Deletic, a., Dotto, C. B. S., Mccarthy, D. T., Kleidorfer, M., Freni, G., & Mannina, G. (2009). Defining Uncertainties in Modelling of Urban Drainage Systems. *8th UDM*, 1–12.

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

Dotto, C.B., 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), pp.2545-2558.

Dotto, C. B. S., Kleidorfer, M., Deletic, A., Rauch, W., & McCarthy, D. T. (2014). Impacts of measured data uncertainty on urban stormwater models. *Journal of Hydrology, 508*, 28-42. doi: 10.1016/j.jhydrol.2013.10.025

Freni, G., Mannina, G., & Viviani, G. (2008). Uncertainty in urban stormwater quality modelling: the effect of acceptability threshold in the GLUE methodology. *Water Res, 42*(8-9), 2061-2072. doi: 10.1016/j.watres.2007.12.014

Freni, G., and Mannina, G. (2010). Uncertainty in water quality modelling: The applicability of Variance Decomposition Approach. *Journal of Hydrology, 394*(3-4), 324-333. doi: 10.1016/j.jhydrol.2010.09.006

Freni, G. and G. Mannina (2012). "Uncertainty estimation of a complex water quality model: The influence of Box–Cox transformation on Bayesian approaches and comparison with a non-Bayesian method." Physics and Chemistry of the Earth, Parts A/B/C 42-44: 31-41.

Fu, G., Butler, D., Khu, S.-T., & Sun, S. (2011). Imprecise probabilistic evaluation of sewer flooding in urban drainage systems using random set theory. *Water Resources Research*, *47*(2), n/a-n/a. http://doi.org/10.1029/2009WR008944

Gamerith, V., Neumann, M. B., & Muschalla, D. (2013). Applying global sensitivity analysis to the modelling of flow and water quality in sewers. *Water Research*, *47*(13), 4600–11. http://doi.org/10.1016/j.watres.2013.04.054

Garthwaite, P. H., Kadane, J. B., & O'Hagan, A. (2005). Statistical methods for eliciting probability distributions, (139911), 1–74. http://doi.org/10.1198/016214505000000105

Helton, J. C., & Davis, F. J. (2003). Latin hypercube sampling and the propagation of uncertainty in analyses of complex systems. *Reliability Engineering and System Safety*, *81*(1), 23–69. http://doi.org/10.1016/S0951-8320(03)00058-9

Hoppe, H. (2008). Impact of input data uncertainties on urban drainage models : climate change - a crucial issue ? In *11th International Conference on Urban Drainage* (pp. 1–10). Edinburgh, UK.

Iooss, B., & Lemaître, P. (2014). A review on global sensitivity analysis methods. *arXiv Preprint arXiv:1404.2405*, (around 30), 23. http://doi.org/10.1007/978-1-4899-7547-8_5

Keith J. Beven, Paul J. Smith, Jim E. Freer, So just why would a modeller choose to be incoherent?, Journal of Hydrology, Volume 354, Issue 1, 2008, Pages 15-32, ISSN 0022-1694, http://dx.doi.org/10.1016/j.jhydrol.2008.02.007.

Korving, H., Van Gelder, P., Van Noorwijk, J., & Clemens, F. (2002). Influence of model parameter uncertainties on decision- making for sewer system management. *Fifth International Conference on Hydroinformatics*, *42*(13). http://doi.org/10.1016/j.watres.2008.05.006

Kroll, S., Dirckx, G., Donckels, B. M. R., Van Dorpe, M., Weemaes, M., & Willems, P. (2016). Modelling real-time control of WWTP influent flow under data scarcity. *Water Science and Technology*, *73*(7), 1637–1643. http://doi.org/10.2166/wst.2015.641

Langeveld, J. , L. Benedetti , J. J. M. de Klein , I. Nopens , Y. Amerlinck , A. van Nieuwenhuijzen , T. Flameling , O. van Zanten & S. Weijers (2013) Impact-based integrated real-time control for improvement of the Dommel River water quality, Urban Water Journal, 10:5, 312-329, DOI: 10.1080/1573062X.2013.820332

Laloy, E., and J. A. Vrugt (2012), High-dimensional posterior exploration of hydrologic models using multiple-try DREAM(ZS) and high-performance computing, Water Resour. Res., 48, W01526, doi:10.1029/2011WR010608.

Morris, M. D. (1991). Factorial Sampling Plans for Preliminary Computational Experiments. *Technometrics*, *33*(2), 161–174.

Pietro Mantovan, Ezio Todini, Hydrological forecasting uncertainty assessment: Incoherence of the GLUE methodology, Journal of Hydrology, Volume 330, Issue 1, 2006, Pages 368-381, ISSN 0022-1694, http://dx.doi.org/10.1016/j.jhydrol.2006.04.046.

Reichert, P., & Mieleitner, J. (2009). Analyzing input and structural uncertainty of nonlinear dynamic models with stochastic, time-dependent parameters. Water Resources Research, 45(10), n/a-n/a. doi: 10.1029/2009wr007814

Refsgaard, J. C., van der Sluijs, J. P., Brown, J., & van der Keur, P. (2006). A framework for dealing with uncertainty due to model structure error. *Advances in Water Resources*, *29*(11), 1586–1597. http://doi.org/10.1016/j.advwatres.2005.11.013

Refsgaard, J. C., van der Sluijs, J. P., Højberg, A. L., & Vanrolleghem, P. A. (2007). Uncertainty in the environmental modelling process – A framework and guidance. *Environmental Modelling & Software*, *22*(11), 1543–1556. http://doi.org/10.1016/j.envsoft.2007.02.004

Renard, B., D. Kavetski, G. Kuczera, M. Thyer, and S. W. Franks (2010), Understanding predictive uncertainty in hydrologic modeling: The challenge of identifying input and structural errors, Water Resour. Res., 46, W05521, doi:10.1029/2009WR008328.

Saltelli, A., Chan, K., & Scott, E. M. (2000). *Sensitivity analysis*. Chichester, UK: John Wiley & Sons.

Saltelli, A., Ratto, M., Andres, T., Campolongo, F., Cariboni, J., Gatelli, D., … Tarantola, S. (2008). *Sensitivity Analysis: From Theory to Practice*. *Global Sensitivity Analysis. The Primer*. Wiley. http://doi.org/10.1002/9780470725184.ch6

Saltelli, A., Tarantola, S., & Chan, K. P.-S. (1999). A Quantitative Model-Independent Method for Global Sensitivity Analysis of Model Output. *Technometrics*, *41*(1), 39–56. http://doi.org/10.1080/00401706.1999.10485594

Schellart, A. N., Tait, S. J., & Ashley, R. M. (2010). Towards quantification of uncertainty in predicting water quality failures in integrated catchment model studies. *Water Res, 44*(13), 3893-3904. doi: 10.1016/j.watres.2010.05.001

Schoups, G., & Vrugt, J. A. (2010). A formal likelihood function for parameter and predictive inference of hydrologic models with correlated, heteroscedastic, and non-Gaussian errors. *Water Resources Research, 46*(10), n/a-n/a. doi: 10.1029/2009wr008933

Sobol, I. M. (2001). Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates. *Mathematics and Computers in Simulation*, *55*(1–3), 271–280. http://doi.org/10.1016/S0378-4754(00)00270-6

Sun, S., & Bertrand-Krajewski, J.-L. (2013). Separately accounting for uncertainties in rainfall and runoff: Calibration of event-based conceptual hydrological models in small urban catchments using Bayesian method. *Water Resources Research, 49*(9), 5381-5394. doi: 10.1002/wrcr.20444

Stedinger, J. R., R. M. Vogel, S. U. Lee, and R. Batchelder (2008), Appraisal of the generalized likelihood uncertainty estimation (GLUE) method, Water Resour. Res., 44, W00B06, doi:10.1029/2008WR006822.

Swaffield, J. A., & Bridge, S. (1983). Applicability of the Colebrook-White Formula to Represent Frictional Losses in Partially Filled Unsteady Pipeflow. *Journal of Research of the National Bureau of Standards*, *88*(6), 389. http://doi.org/10.6028/jres.088.019

Swayne, D., Yang, W., Voinov, A., Rizzoli, a, & Filatova, T. (2010). Uncertainty propagation throughout an integrated water-quality model. *International Environmental Modelling and Software Society*, (2006), 1–8. Retrieved from http://www.iemss.org/iemss2010/papers/S20/S.20.03.Uncertainty propagation throughout an integrated waterquality model.

Uusitalo, L., Lehikoinen, A., Helle, I., & Myrberg, K. (2015). An overview of methods to evaluate uncertainty of deterministic models in decision support. Environmental Modelling and Software, 63, 24–31. http://doi.org/10.1016/j.envsoft.2014.09.017

Vanrolleghem, P. A., Mannina, G., Cosenza, A., & Neumann, M. B. (2015). Global sensitivity analysis for urban water quality modelling: Terminology, convergence and comparison of different methods. *Journal of Hydrology*, *522*(September), 339–352. http://doi.org/10.1016/j.jhydrol.2014.12.056

Vezzaro, L., Mikkelsen, P. S., Deletic, A., & McCarthy, D. (2013). Urban drainage models--simplifying uncertainty analysis for practitioners. *Water Science and Technology : A Journal of the International Association on Water Pollution Research*, *68*(10), 2136–43. http://doi.org/10.2166/wst.2013.460

Wainwright, H. M., Finsterle, S., Jung, Y., Zhou, Q., & Birkholzer, J. T. (2014). Making sense of global sensitivity analyses. *Computers & Geosciences*, *65*, 84–94. http://doi.org/10.1016/j.cageo.2013.06.006

Wani, O., Beckers, J. V., Weerts, A. H., & Solomatine, D. P. Residual uncertainty estimation using instance-based learning with applications to hydrologic forecasting.

Willems, P. (2008). Quantification and relative comparison of different types of uncertainties in sewer water quality modeling. *Water Res, 42*(13), 3539-3551. doi: 10.1016/j.watres.2008.05.006

Willems, P. (2012). "Model uncertainty analysis by variance decomposition." Physics and Chemistry of the Earth, Parts A/B/C 42-44: 21-30.

Xiu, D., & Karniadakis, G. E. (2003). Modeling uncertainty in flow simulations via generalized polynomial chaos. Journal of computational physics, 187(1), 137-167.