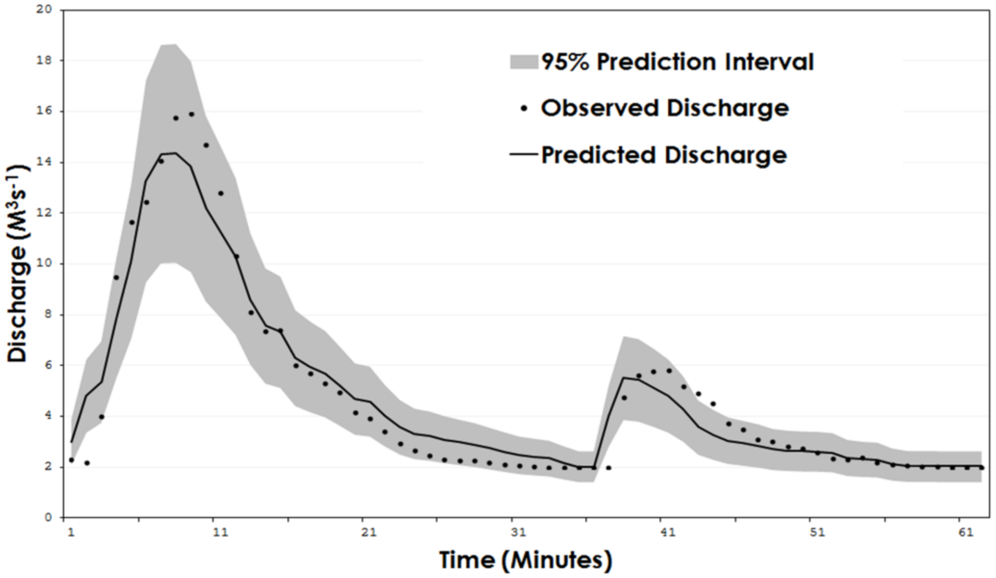
Prepared
enabling change

Uncertainty Quantification and Data Assimilation for Real-Time Modelling:

Technical Guidelines



The European Commission is funding the Collaborative project 'PREPARED Enabling Change'

SEVENTH FRAMEWORK
PROGRAMME

COLOPHON

Title
Uncertainty Quantification and Data Assimilation for Real-Time Modelling: Technical Guidelines

Report number
PREPARED 2014.050

Deliverable number
D3.6.4

Author(s)
Hutton C.J., Kapelan Z., Vamvakeridou-Lyroudia L., and Savić D.

Document history

Version	Team member	Status	Date update	Comments
1	CJH	Draft	30/01/2014	First Draft

This report is:
PU = Public

Contents

	Contents	**1**
1	**Introduction**	**2**
2	**Uncertainties in Urban Water Systems' Modelling**	**4**
2.1	Water Distribution System Model Uncertainties	4
2.2	Urban Rainfall Runoff Model Uncertainties	5
3	**Technical Guidelines Overview**	**9**
4	**Model Calibration**	**16**
4.1	Formal Bayesian Calibration	17
4.2	Informal Bayesian Calibration	19
4.3	Water Distribution System Model Calibration	19
4.4	Urban Rainfall-Runoff Model Calibration	20
5	**Data Assimilation**	**21**
5.1	Water Distribution Systems	23
5.2	Urban rainfall runoff models	24
6	**Model Forecasting**	**26**
6.1	Error-Correction Procedures	26
6.2	Approximating Ensemble Forecasts	27
6.3	Data-Driven Modelling	27
6.4	Forecast Evaluation	28
7	**References**	**29**

1 Introduction

Adaptive strategies are required to enable cities to better manage the potential risks imposed by climatic change, allowing them to become more resilient. Such resilience may be obtained by making appropriate capital investments into water supply, urban drainage and sanitation infrastructure. To avoid an unnecessary waste of resources, and postpone infrastructure investments until more knowledge is available, methods may be applied to optimise the performance of existing infrastructure.

Mathematical Models have been, and continue to be applied to improve the operation of Urban Water Systems. Models are essentially applied to understand how a given Urban Water System (UWS) may act under future conditions. Such models are applied either at longer time scales (> 10 years) for urban water system design, and increasingly at shorter timescales (sub daily and weekly forecasts) to inform and improve the real-time control of urban water systems. It is through such control that the performance of existing infrastructure may be optimised. Models can therefore form a central component in the development of adaptive strategies.

Despite the benefits that models may provide to the development of adaptive strategies, like the impacts of climate change themselves, Urban Water Systems' models contain considerable aleatory and epistemic uncertainties. Uncertainty in a model's predictive performance provides a constraint on how useful and reliable such a prediction may be when used to inform the decision making process. Neglecting to quantify and communicate model predictive uncertainty, or at least failing to consider its magnitude, at best represents bad practise. At worst, a failure to consider uncertainty can lead to poor control decisions, and undesired consequences. Given the importance of water, both as a vital resource and as a natural hazard, it is very important that methods are applied to reduce the various forms of model uncertainty, and their influence on predictions of water quantity and quality. Furthermore, it is arguably more important that model uncertainty is adequately quantified, such that model predictions can be used robustly to inform the decision making process.

Research has been conducted to address the challenge of quantifying and reducing model predictive uncertainty in Urban Water Systems' models, such that model predictions may be used more robustly to inform real-time control (adaptive) strategies. Such research has been conducted in two principal areas:

1. **Model Calibration** – where the predictive performance of a model is optimised by adjusting model parameters, and the associated uncertainties in the model parameters and predictions quantified.
2. **Real-Time Modelling** – where newly available data may be used (assimilated) to correct the states and parameters of a model to derive more accurate real-time predictions and forecasts of future system conditions.

Based on the research conducted in PREPARED Work Package 3.6, this report provides technical guidelines in applying uncertainty quantification, data assimilation and error-correction techniques in UWS modelling. The purpose of these guidelines is to provide recommendations of best practice for uncertainty quantification to help end users in their own applications.

A review of existing methods for uncertainty quantification, technical guidelines for best practise, and recommendations has been published based on the PREPARED work conducted in WP3.6 (Hutton et al., 2012). The technical guidelines presented in Hutton et al. (2012) form the basis for this report. These guidelines have been extended based on more recently conducted work, and in this report also extended to consider Urban Rainfall-Runoff (URR) Models alongside Water Distribution System Models. Furthermore, some additional material is also re-presented, where appropriate from Deliverable 3.6.1 and Deliverable 3.6.2. This report is therefore structured as follows:

First, an overview is provided in Section 2 on typical uncertainties that need to be considered when modelling Water Distribution Systems and Urban Rainfall Runoff Models. Second, an overview of the technical guidelines is presented in Section 3, which specifies the general guidelines for Model Calibration, Data Assimilation and Model Forecasting. Further details for each of these areas may then be found in sections 4, 5 and 6, respectively.

2 Uncertainties in Urban Water Systems' Modelling

Uncertainty may be described as a situation where we have incomplete or imprecise knowledge about a system or a specific state of that system. Such uncertainty may be divided into two categories (Hall 2003):

- **Aleatory uncertainty** refers to natural variability in a system that by its very nature is irreducible. For example, variability in water demand or rainfall.
- **Epistemic uncertainty** results from incomplete knowledge of a system. Numerical models contain epistemic uncertainty as they are, somewhat inherently, incomplete representations of a system – a known simplification. Such uncertainty may be reduced by deriving greater understanding of a system.

Both sources of uncertainty are present when modelling Urban Water Systems, and are often specific to the type of system in question. In the context of numerical modelling, these uncertainties are manifest in the ***Model Structures***, ***Model Parameters***, and also in ***Measurement Data***. To provide greater context for the technical guidelines below, this section briefly considers the major types of uncertainties that are to be dealt with in Water Distribution Systems and Urban Rainfall-Runoff Systems.

2.1 Water Distribution System Model Uncertainties

Model Structural Uncertainty refers to errors in the mathematical representation of reality, and is a form of epistemic uncertainty, where the model will never equal reality. Examples include:

- ***Skeletonisation*** - the removal of pipes not considered essential for system analysis – represents one of the key WDS model structural uncertainties. Skeletonised models may neglect dead ends and high elevation nodes in the network, and adversely affect pressure surges (Boulos et al. 2004), demand satisfaction predictions (Walski et al. 2003), contaminant consequence assessment (Bahadur et al. 2006) and chlorine decay simulation (Menaia et al. 2003). Whilst skeletonised models may be hydraulically equivalent to all pipes models for steady state conditions (Jung et al. 2007; Preis et al. 2011), they can perform poorly in transient conditions (Jung et al. 2007). All pipes models application (e.g. Jacobsen and Kamojjala 2009) may be computationally unfeasible in real-time, while increased data requirements for calibration may outweigh structural uncertainty.
- ***Water demand*** is typically aggregated at junction nodes in WDS models, yet consumers extract water from along pipes within the network. Although head loss corrections to overcome this simplification have been developed (Giustolisi 2010), accurate specification of distributed demand is difficult.
- ***Demand driven models*** may be considered valid for normal operating conditions in well designed and maintained WDS, whereas ***pressure driven*** solutions are more appropriate in cases of fire flow, pipe leakage and valve closure (Giustolisi et al. 2008a; Giustolisi et al. 2008b). The latter approach requires additional data to determine the relationship

between pressure head and flow (Ozger and Mays 2004), which are not usually accurate, and increased computational time, which is not always available for real time computations.

Parameter Uncertainty reflects uncertainty in equation variables used to represent system components (e.g. pipe roughness). Such uncertainty is aleatory, as parameters can vary over space and time, and epistemic as system discretisation is space and time can result in a failure to reconcile the scale observations with model parameters. Parameter values are often therefore 'effective' (Lane, 2005) in that they produce the correct prediction, but often have little physical meaning.

- ***Pipe roughness*** is problematic to identify accurately in WDS models as it cannot be directly measured, and because of pipe deterioration (Boulos et al. 2004; Kleiner and Rajani 2001), roughness changes with pipe age. Roughness values calibrated using junction pressure measurements (Kapelan et al. 2007; Savic et al. 2009) will reflect uncertainties in system specification, roughness pipe grouping, and data uncertainty.
- Due to wear (Hirschi et al. 1998) ***pumps*** typically do not operate at the efficient point supplied by the manufacturer (Walski et al. 2003), and alongside ***valve settings***, may need to be considered in the calibration problem.

Measurement/Data Uncertainty refers to uncertainty in quantities used to define initial conditions, model inputs, and model state observations utilised in evaluation of model predictions. Such uncertainties result from instrumentation errors, and mismatches between the scale of observations and predictions (e.g., demand lumping and disaggregation).

- ***Aleatory demand uncertainty*** is large in WDS models, as demand fluctuates over a variety of temporal and spacial scales depending on consumer type (Davidson and Bouchart 2006; Herrera et al. 2010)
- ***Epistemic demand uncertainty*** results from a low density of metered houses, and the difficulty of obtaining such information in real-time. Demand is more readily inferred by calibration to measured pipe flow, water quality and DMA measurements (Branisavljevic et al. 2009; Jonkergouw et al. 2008; Kang and Lansey 2009). However, such approaches may require downscaling to individual network nodes (Kang and Lansey 2009).

2.2 Urban Rainfall Runoff Model Uncertainties

A number of modelling approaches have moved towards integrated modelling of urban rainfall runoff models (Butler and Schutze 2005; Vanrolleghem et al. 2005; e.g. WEST and SIMBA; Rauch et al. 2002). Such models are required to help optimise the performance of existing urban waste water systems by explicitly accounting for interactions between different components of the system (Butler and Schutze 2005), and in doing so facilitate incremental adaptation (Butler and Parkinson 1997). However, the integrated urban waste water system is complex, involving a number of epistemic and aleatory uncertainties (Benedetti et al. 2008; Korving et al. 2003). In addition to that presented herein, further details on Urban Drainage Model uncertainties can be found in Deletic et al. (2012).

Model Structural Uncertainty in sewer systems reflects uncertainty in the manner in which the complex processes governing flow hydraulics, water quality, biochemistry and sediment transport are represented. In general model complexity may be increased, if possible, to reduce structural uncertainty, however this comes at the expense of needing to constrain more parameters, which due to data limitations are themselves uncertain. If model structural complexity is reduced to a simpler conceptual approach it is often difficult to infer the physical meaning of model parameters, which require sufficient data for calibration. The most important examples of model structural uncertainty include:

- ***Sewer System Discretisation.*** URR systems are typically branched in nature, similar to natural rainfall-runoff systems. Similar to skeletonisation in WDS models, URR systems may be represented in a distributed way, but often the hydraulic/hydrologic response of components of a network may be represented in a lumped manner.
- ***Model Equations.*** Individual pipes may be represented in a sewer system model, however various forms of simplification of the fully dynamic 1D St Venant equations to derive diffusive wave and kinematic wave equations have been applied to sewer systems, and conceptual store models when computational times and data are not availablerequired to support a more detailed model representation (Vaes and Berlamont 1999). Furthermore there are epistemic uncertainties surrounding the representation of other sewer system processes, such as sediment transport (Ashley et al. 1999; De Sutter et al. 2003).
- ***Waste Water Treatment*** processes are complex and difficult to measure. Given the complexity of biological processes a certain amount of greyness may need to be introduced into process representation. For example, Black-box models, calibrated based on input and output data may provide better system representations in cases when white-box models fail to correctly describe all system dynamics (Gernaey et al. 2004). Further details of structural uncertainties in the WWTP may be found else ware (e.g. Gernaey et al. 2004; Rauch et al. 1999)

Model Parameter Uncertainty in URR models reflects uncertainty in equation variables used to represent system components (e.g. pipe roughness). As above, such uncertainty is both aleatory and epistemic in nature; parameters may naturally vary over both time and space, and system discretisation can result in scale dependent model parameter values. Whilst the modelling basis of many of the model structural components is well understood (low epistemic uncertainty) due to sound conceptual and mathematical understanding of the system (e.g. St. Venant equations) there are a number of significant problems in deriving empirical information (Ashley et al. 1999). Such uncertainties include:

- ***Pipe roughness*** is a key parameter value in URR models, which is difficult to constrain and identify in a distributed manner owing to strong spatial variability and a difficulty of measuring sewer system processes. Such variability in part results within pipe sediment

(Pomeroy 1967) and biofilm formation (Guzman et al. 2007), which when deposited also modifies sewer pipe geometry and the pipe depth-discharge relationship (Ackers et al. 1996).

- ***Waste Water Treatment models*** are difficult to parameterise due to parameter demands that are often substantial and difficult to constrain (Sin et al. 2009), in part due to the complex processes that may need to be represented. For example, parameters governing the active sludge process are often determined from laboratory studies (Van Veldhuizen et al. 1999), which may not be representative of field conditions. Further parameter uncertainty may occur when models, which are often calibrated for dry flow conditions, are applied to wet flow conditions (Gernaey et al. 2004).

Measurement/Data Uncertainty results from errors in data used to constrain and calibrate models, and also in data used as inputs to drive the model. These uncertainties result principally from the difficulty of actually measuring and monitoring sewer system processes in a distributed manner, with subsequent difficulties that result in model calibration. Therefore, even if model description may be considered perfect (no structural uncertainty), models are often heavily reliant on quality data that is unavailable to constrain parameter values, and drive the model (e.g. rainfall). The most common forms of Data Uncertainty refer to:

- ***Aleatory Rainfall Uncertainty*** relates to natural spatial and temporal variability in rainfall. Temporally, rainfall varies over a range of different timescales due to different processes (Rodriguez-Puebla et al. 1998) (Kutiel and Sharon 1980; Kutiel and Sharon 1981) which results in difficulties in forecasting rainfall, and therefore in forecasting the urban rainfall-runoff systems response. For example the magnitude of the first flush phenomenon is strongly dependent on the length of the antecedent dry period (Krein et al. 2007). Spatially, rainfall varies over a range of different scales. The urban environment may induce spatial variability itself in rainfall due to convective heating (Jauregui and Romales 1996; Thielen and Gadian 1997).
- ***Epistemic Rainfall Uncertainty*** results from measurement errors and errors in the spatial and temporal resolution of the phenomena. Point rainfall measurements are typically obtained from rain gauges. Measurement accuracy may be compromised by wind speed (Sevruk 1996; Sevruk et al. 1994; Sevruk and Nespor 1998), rainfall intensity (Ciach 2003), evapotranspiration, calibration (Rauch et al. 1998; Stransky et al. 2007), and random errors due to data transmission, mechanical problems and clogging (Rauch et al. 1998). For integrated urban modelling rainfall time series with a temporal resolution of the order of minutes are potentially required (Rauch et al. 1998), as temporal resolution of rainfall has been shown to affect urban drainage model performance and uncertainty (Aronica et al. 2005). Point rainfall measurements require spatial interpolation for input to sewer models. The assumption of uniform rainfall, which is often made due to a low resolution of rain gauges, introduces significant error. As in hydrological applications (Yatheendradas et al. 2008), rainfall uncertainty may dominate over model and parameter uncertainty for the prediction of sewer flow emissions (Willems 1999). Rainfall radar measurements can provide a greater spatial coverage in comparison to point rainfall

measurements, yet the algorithm used to convert a radar signal to rainfall intensity often requires bias correction due to uncertain parameters (Vieux and Vieux 2005). Point gauge measurements are typically used for bias correction (Campolongo et al. 2007), which as discussed above are themselves uncertain. Such uncertainty needs to be propagated through UWS models (Collier 2009).

- ***Dry Weather Flow*** uncertainty is similar to water consumption (demand) uncertainty in the WDS, and is both aleatory, reflecting changing consumer inputs over different timescales, and epistemic because of the difficulty in quantifying the volume and quality of waste water from consumers and industry. Domestic wastewater may be made up of contributions from a variety of different household appliances (e.g. WC, Shower, Dishwasher, Sink, Washing Machine), each with their own patterns of use that vary between weekday and weekend (Butler 1993; Friedler et al. 1996), and diurnally (Figure 8; Figure 9; Almeida et al. 1999). Uncertainty in the temporal sequence of pollution also results as different types of pollutants are produced by different appliances in different quantities, which each may have multiple functions, and therefore loads (Almeida et al. 1999; Friedler and Butler 1996). Further aleatory uncertainty results from different usage amongst different users, which has been found to dominate uncertainty introduced by different types of WC when considering risks of overloading a treatment plant (Wong and Mui 2007). There is significant epistemic uncertainty in the nature of DWF from domestic properties, owing to the difficulty of measuring actual discharge per household. Actual volume and pollutant loads have been determined by consumer survey (Almeida et al. 1999; Wong and Mui 2007), coupled with appliance measurement for average usage and literature figures for different pollutants (Siegrist et al. 1976) . Therefore, there is uncertainty regarding the reliability of multiplying up short period measurements with consumer survey information, as both may not be representative of reality.
- ***Measurment uncertainty*** in sewer systems is often high due to the difficulty of actually measuring water discharge and quality in real-time (Bertrand-Krajewski et al. 2003; Lepot et al. 2013). For further information on measurement uncertainty please refer to Deliverable 3.1.6.

3 Technical Guidelines Overview

Section 2 introduced the key types of uncertainty that exist in the context of urban systems modelling. Methods are required to move beyond deterministic model predictions, and a simple understanding of the likely model uncertainty, to methods that formally represent our uncertain knowledge of system variables, states, and parameters mathematically. Probability theory has been the dominant paradigm for representing uncertainty in the broader Hydroinformatics literature (Hall 2003). Instead of making deterministic model predictions based on a single set of optimal parameters, a range of models (mathematical formulations) and/or parameter sets within each model are considered and assigned a probability. The model(s) are run forwards in time to make an ensemble prediction, which is used to construct a probabilistic prediction from the model for the variable (model state) of interest.

It is within a probabilistic framework, and more specifically a Bayesian probabilistic framework, that the majority of methods considered in these technical guidelines have been developed. Figure 1 shows a flow diagram that summarises the technical guidelines and considers some of the key uncertainties and decisions that need to be made when considering how to deal with uncertainty. The most appropriate techniques for dealing with model uncertainty that should be applied depend on the stage of model development and application. These technical guidelines are therefore divided into ***Model Calibration***, ***Data Assimilation*** and ***Model Forecasting***. These guidelines, alongside some more generic principles, are summarised in specific text-boxes here in section 3, and expanded upon in Sections 4, 5 and 6.

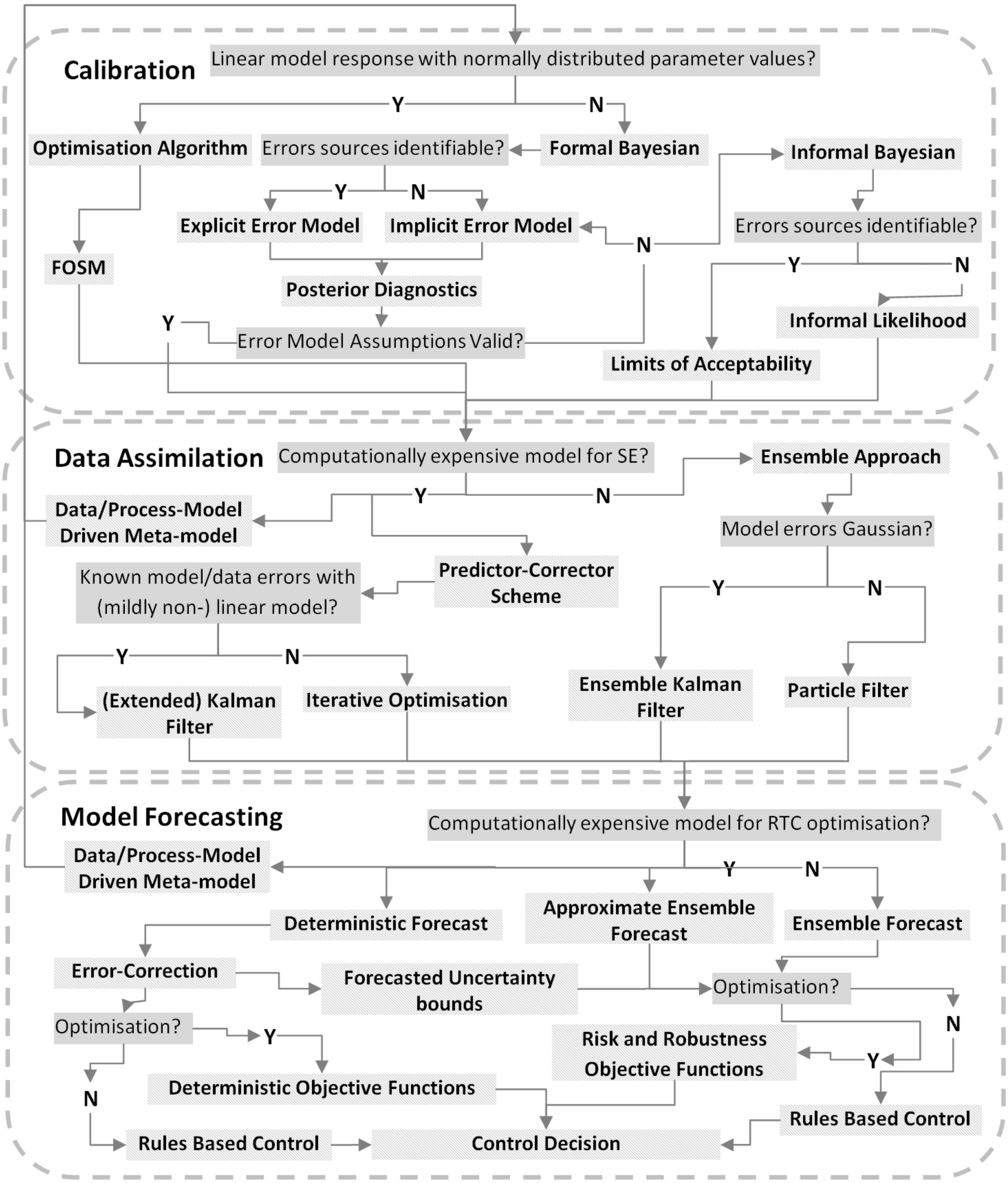


Figure 1. Flow diagram illustrating the key decisions and assumptions that need to be considered when dealing with the propagation of model uncertainty (Hutton et al., 2012).

General Guidelines

- Applying uncertainty quantification and reduction methods may necessarily be iterative, until the assumption made in applying a given method in a particular setting have been fully evaluated.
- The assumptions of a particular method for quantifying model uncertainty should be evaluated with posterior residual diagnostic checks (Section 4.2) and also evaluations of the model forecasting uncertainty bounds (Section 6.4).
- Uncertainty quantification and reduction methods should not be seen as a final stage, or add on procedure during model calibration and subsequent application. Rather, evaluation of the assumptions made during model application, through residual checks, and sensitivity analysis, should form part of an iterative diagnostic approach towards the development of better models and improved methods for quantifying and reducing model error.
- Access to the underlying code of a model may be required to modify system states in the model, which if not available will restrict the application of some Data Assimilation techniques.
- During Data Assimilation, modifications to model states may cause instabilities in the numerical solution of the constitutive equations.
- Computational limitations represent a key constraint on the application of different uncertainty quantification methods, particularly during real-time application. In such circumstances, alternative approaches such as data-driven models may be applied as an alternative to the original model, as a means to approximate uncertainty bounds derived offline, and as a model error-correction procedure to reduce forecasting uncertainty.

It is important to emphasise that the most appropriate uncertainty quantification and reduction method to apply in any given setting may not at first seem obvious until the assumptions made in applying the method have been evaluated. Whilst for some methods it might be known a priori that they are likely to be inappropriate, for example the strong assumptions made in a formal Bayesian likelihood functions, for other methods experimentation will be required. Therefore, applying methods to deal with uncertainty quantification and reduction, at each stage from model development to forecasting and informing the decision making process, may necessarily be iterative.

Iterative application of uncertainty quantification methods is exacerbated by the need to consider not only the validity of different assumptions made in a particular methodology, but also the practical considerations of coupling a specific system model, which may be proprietary in nature, to a methodology for uncertainty quantification and reduction. These practical issues are threefold: First, access to the model code may be restricted, depending on the type of model that is to be used. Such constraint may not be such a problem for model calibration, which is typically run on a batch (time-series) of data, and model parameters can be changed each time the model is run. However, in real-time model application, data assimilation methodologies use near real-time observations to update model states or parameters. Second, depending on the mathematical formulation of the model, and the imposed numerical solution to the constitutive equations, modifications to model states in real-time may induce numerical instabilities. In the above two circumstances a data-driven model may be calibrated to a residual error time-series between model predictions and observations, and added to the model forecast to try and reduce model forecasting error, without the need to modify the original model.

Computational resources may limit the effectiveness of some of the uncertainty quantification and reduction methodologies. Such limitations depend on the computational resources available in each specific problem. The problem may be an issue during offline modelling if the model is too complex and takes too long to run for each parameter sample. This may lead to inadequate sampling of the posterior distribution and insufficient characterisation of the model uncertainty. However, computational limitations are more likely to be a constraint during real-time modelling. During real-time modelling there is limited time, and therefore computational budget between receiving an observation, assimilating the observation to improve the states of the model, and running a model forecast. In the decision making context a number of model forecasts may be required to identify near optimal controls (e.g. pump settings). Whilst ideally an ensemble of model predictions may be run in real-time to represent different forms of model, parameter and input uncertainty for each forecast, as considered in Hutton et al. (2012), the limited computational budget may need to be divided carefully. Where computational resources are limited, less demanding models computationally, such as conceptual models and data driven models may be applied, alongside other methods to approximate model uncertainty (See Section 6). An important and case specific question to address concerns the application of real-time modelling under a limited computational budget: For a fixed computational budget does a deterministic model prediction and correction procedure provide better predictions to inform the control decision in comparison to a simpler system meta-model where the computational budget is used to better quantify and reduce the model uncertainty?

The best method for a particular application may not at first be apparent, and therefore that application of different methods to any particular problem is encouraged to adequately understand the validity of the assumptions. For example, probability theory, though widely applied to quantify different forms of uncertainty has a number of limitations, which should be understood (Hall et al. 2003). Exploration and exposition of methodological assumptions is of particular importance when models are to be then used to inform the decision making process, as inadequate characterisation of model uncertainties may lead to inappropriate control decisions.

Finally, it is important to consider that uncertainty quantification and reduction within modelling is not merely a final stage, or add on procedure to a deterministic model, or should at least not be depending on practical considerations. Analysing model performance, and in particular adequately characterising and reducing the magnitude of residual errors between model predictions and observations, should form an essential stage in model development, and model application towards identifying better models and better predictions.

Model Calibration

- Deterministic Model calibration is limited by the implicit assumption that there are no other forms of uncertainty, besides that of the model parameter values. Probabilistic Bayesian approaches attempt to quantify the effect of different sources of model uncertainty during model calibration.
- It is recommended that the uncertainties present in data used to drive model simulations (e.g. rainfall, dry weather flow and water demand) and in the measurements used to evaluate model performance (e.g. flow rates, tank water levels, pressure heads) are quantified where possible prior to model calibration.
- Model calibration and performance should, where possible be evaluated with additional validation data – e.g. the model should be applied to data not used during calibration.
- It is recommended that Formal Bayesian approaches are first considered for model calibration.
- During Formal Bayesian calibration, the assumptions made in applying a Formal Bayesian error model should be fully evaluated with posterior diagnostic checks of model residual errors (Section 4.2) and also of the derived prediction bounds (Section 6.4).
- If, as is often the case, the residual assumptions of a Formal Bayesian Calibration are deemed invalid, then an Informal Bayesian Calibration should be applied.
- During informal Bayesian calibration it is important to evaluate the sensitivity of the uncertainty bounds to the subjective assumptions concerning the behavioural threshold, and also to the chosen likelihood.
- During Water Distribution System model calibration the uncertainty in specifying water demand should be jointly considered alongside pipe roughness calibration.

Data Assimilation

- Regardless of the specific method applied to assimilate observations, a key factor governing the success of the assimilation procedure is the ability to define error terms of both the observations and also of the model.
- As with model calibration, it is recommended that uncertainties present in data used to drive model simulations (e.g. rainfall, dry weather flow and water demand) and in the measurements used in Data Assimilation (e.g. flow rates, tank water levels, pressure heads) are quantified where possible prior to model calibration.
- As computational resources represent a key constraint in real-time modelling, deterministic assimilation procedures may necessarily have to be applied. However, ideally an ensemble approach will be applied (e.g. Ensemble Kalman Filter, and Particle Filter or variants thereof) to quantify and better account for model uncertainty.
- When applying Data Assimilation procedures the physical realism of the state correction needs to be considered, which may lead to physically unrealistic states in the model.
- Data Assimilation may lead to numerical instability in the applied model – this should also be evaluated.
- The presence of control structures in a system model such as pumps, weirs and gates will strongly influence Data Assimilation performance and may act to decouple model states from observations. Their influence on Data Assimilation performance should be evaluated.
- The model process time-lag between a model state to be updated and the observations used in the update should be accounted for when applying Data Assimilation methods to maintain some physical realism in the update, and to avoid the potential for oscillatory behaviour in the model induced by independent state updates.
- Given flow attenuation in many models, multiple downstream observations through time may be informative of the value of an upstream model state. This should be considered as a means to reduce further state uncertainty.

Model Forecasting

- Data-Driven models can be applied as error correction procedures to correct model forecasts at the observation location, and typically represent a more computationally efficient approach for model error reduction.
- During application of error-correction procedures, it is important not to apply the error correction model outside the range of conditions for which the model was calibrated.
- If computational resources are limited, then approximate ensemble forecasts can be applied. Model sensitivity analysis is recommended as a means to identify the key sources of model error to represent. Alternatively data-driven models may be applied to approximate uncertainty bounds derived from offline calibration.
- Data-driven models in general can be applied to overcome the computational limitations of real-time modelling and forecasting, through error-correction, calibration to offline prediction bounds, and also as alternatives of the originally applied process-based model. Therefore, methods for model calibration and data assimilation presented should also be applied when developing an error-correction model.
- Finally, during probabilistic model forecasting the reliability and sharpness of the model forecast at different time-horizons should be evaluated.

4 Model Calibration

Calibration represents a fundamental stage in model application and is a necessary step, typically taken offline prior to applying a model to understand current and future system conditions. At its most basic level calibration attempts to identify an optimal set of model parameters (θ) that minimise the difference between a vector of observations and a vector of corresponding model predictions. E.g. that minimises the vector of model residuals. In the deterministic approach the residual vector is typically summarised by a global metric such as the Root Mean Square Error (RMSE), which is forced as close to zero as possible by adjusting model parameters using some optimisation procedure (e.g. Genetic Algorithm; Savic et al., 2009).

The limitation of deterministic approaches in that there is an implicit assumption that there are no other forms of uncertainty. More specifically, that the model structure ($\mathbf{M}$) equates to reality; that the initial states ($\mathbf{X}_0$) of the system represent the true system states; the input drivers ($\mathbf{D}$) such as rainfall, water demand and dry weather flow represent the true drivers of the system; and the measurements ($\mathbf{Y}$) that are compared to the model predictions are also error free. This is not the case, however; a model will never equate to reality and measurements used to constrain model inputs and evaluate model performance are rarely free from errors. The result is that as model parameter space is explored to minimise the RMSE for example, multiple parameter sets may be identified in parameter space that may produce equally likely model predictions, a form of model parameter equifinality (Beven 2006).

In light of parameter equifinality it is more useful and interesting to obtain the posterior parameter probability density function (PDF). In WDS models this has typically been achieved post calibration with the **F**irst **O**rder **S**econd **M**oment (FOSM) method (Bush and Uber 1998; Lansey et al. 2001), as in Figure 1. However, FOSM makes strong assumptions including model linearity, independence and normality of measurement errors and parameter values, and can be expensive computationally when calculating derivatives (Kapelan et al. 2007). Further, it may not be applicable if the posterior parameter distribution deviates from the multi-normal distribution (Vrugt et al. 2003).

In a range of modelling disciplines the preferred method for quantifying model parameter and predictive uncertainty is to obtaining the posterior PDF is via Bayes' equation, which can be specified considering the joint inference of both model structure and parameters (Draper 1995):

$$P(\mathbf{M}, \boldsymbol{\theta} | \mathbf{Y}, \mathbf{X_0}, \mathbf{D}) \propto P(\mathbf{Y} | \boldsymbol{\theta}, \mathbf{M}, \mathbf{X_0}, \mathbf{D}) P(\boldsymbol{\theta} | \mathbf{M}) P(\mathbf{M}) \quad (1)$$

The first right hand term represents the likelihood function, the second term the prior parameter distribution, and the third term the prior distribution of possible model structures. In natural catchment systems a range of model structures might be considered as different hypotheses of how a system operates (e.g. Krueger et al. 2010), which provides a means to identify where different models perform best, and therefore a means to try and understand epistemic uncertainties. In should be noted though that all model structures are wrong, and thus one cannot formlly specify a prior on the model structure as this cannot contain a ***true*** model. Modelling in a bayesian framework in this sense falls better within a hypothetico-deductivist framework (Gelman and Shalizi, 2012), whereby model structures are applied to identify improvements in predictive performance with increasing complexity. In most applications in urban systems, however, a single model is typically applied, which collapses this final term in Equation 1 to a single set of structural assumptions. This is because the modelling focus is less on improving understanding per se, but in developing good predictive models to be applied to inform system control. This is not to say, however that model structures cannot be improved in UWS models – posterior model predictive checks provide a means to identify where improvements may occur (see below).

Solving Bayes' equation analytically is typically intractable, and therefore some form of posterior sampling is conducted; methods include random Monte Carlo sampling and Latin hypercube sampling (Kang et al. 2009), and a more advanced family of **M**arkov **C**hain **M**onte **C**arlo (**MCMC**) methods, which use past information derived from the posterior distribution to inform the nature of further sampling (see Vrugt et al. (2009a) for recent developments and methodological issues).

Two general Bayesian approaches have been widely applied for model calibration, which may generally be referred as Formal and Informal Bayesian approaches. Such approaches generally differ in how the likelihood in Equation 1 is specified. The guidelines presented in Figure 1 recommend these approaches when the assumption of the FOSM method are invalid. Informal Bayesian approaches were, in some respects, initially developed in light of dissatisfaction with Formal Bayesian approaches, and the strong assumptions that may have to be made in applying the approach.

4.1 Formal Bayesian Calibration

The classical formal approach requires specification of a likelihood function which necessarily makes strong assumptions regarding the nature of model errors. The standard approach assumes that residual errors are mutually independent, Gaussian-distributed and homoscedastic, leading to a Gaussian error model (or log likelihood for convenience; Vrugt et al. 2009b). Unfortunately residual errors often do not conform to such a simple distribution (Thyer et al. 2009), which can lead to bias in the posterior parameter PDF, and predictive distribution (Beven et al. 2008). To approaches may be applied to overcome this issue:

1. Apply a ***data transformation***, such as the Box-Cox transformation to reduce/remove Heteroscedasticity and non-Gaussianity from the residual distribution (Box and Cox 1982; Freni and Mannina 2010).
2. ***Modify the likelihood function*** to account better for heavy tailed distributions, and skew and kurtosis in the residuals. Recently Schoups and Vrugt (2010) included

heteroscedastic, skew, kurtosis and bias parameters into a likelihood function that, alongside model parameters, had to be jointly inferred from the sampling procedure (Schoups and Vrugt 2010).

Within Formal Bayesian calibration the ***Explicit Formal Bayesian*** approach seeks to separate out and represent different forms of error within the modelling framework, as in the Bayesian Total Error Framework **BATEA.** Within this idea, uncertainty in driving conditions (such as rainfall), measurement error, and model structural error, may be represented separately, via additional "latent parameters" (Thyer et al. 2009; Vrugt et al. 2009b), that have to be inferred jointly with model parameters during the calibration procedure. Such an approach will be most successful where there is additional information on for example measurement uncertainty. The reason is that the approach is vulnerable to ill-posedness that results from the difficulty of specifying a priori that nature of input and structural errors (Renard et al. 2010) such that it may be difficult to explicitly separate out, and represent different sources of error.

In contrast to the Formal bayesian approach the ***Implicit Formal Bayesian*** approach dos not attemp to separate out sources of uncertainty, but rather attempts to use the likelihood function to quantify different forms of uncertainty, and their influence on the parameter and predictive distributions. The parameters of such a procedure may be more readily derived for representation of total model uncertainty (Schoups and Vrugt 2010).

An integral yet frequently ignored aspect of the formal inference procedure is the application of posterior diagnostic checks to evaluate the error model hypothesises encapsulated in the likelihood function (Thyer et al. 2009). Such checks are also important to identify how a model prediction is defficient, and therefore form an important part of a diagnostic approach for improving the model. These checks include:

- **Q-Q plots:** should be applied to evaluate whether the empirical residual distribution conforms to the (theoretical) error model distribution (Thyer et al. 2009).
- **Heteroscedasticity:** May be checked by plotting the model residuals as a function of the observations. Ideally the residuals should plot randomly about 0 (on the y-axis), and show no trend with the magnitude of observations.
- **Auto-correlation:** should be evaluate by calculating the autocorrelation function (ACF) and partial autocorrelation function (PACF) – the latter calculates the autocorrelation at a given lag once autocorrelation at smaller lags has been accounted for. Such metrics should be calculated to evaluate temporal auto-correlation, and also spatial auto-correlation if more than one observation location is used in calibration.

Such checks are vitally important to evaluate the validity of the assumptions made in the model likelihood function, and to help identify defficiencies within the model. Such checks should be made during calibration but also when applied to a set of validation data to understand the validity of the calibration and associated prediction bounds perform during prediction; e.g. against a set of data not used for the calibration. Posesive

4.2 Informal Bayesian Calibration

The residual assumptions made in the formal bayesian approach may, following posterior predictive checks, be difficult to justify, as residual distributions often do not conform to the specfid distributions. The Generalised Likelihood Uncertainty Estimation Procedure (GLUE) (Beven and Binley 1992), more recently referred to as an informal Bayesian approach (Smith et al. 2008), seeks to find "behavioural" parameter sets; that is, seeks to identify parameter sets consistent with the observations according to an informal likelihood function (see Smith et al. (2008) for a review). Following some posterior sampling, using an informal likelihood, parameter sets whose likelihood is greater than a user defined threshold are retained, whilst those that do not meet the treshold are rejected. The likelihoods of the retained parameter sets are normalised to unity to derive probabilistic information for both the model parameters and the model predictions. Such an approach does not make assumptions about the nature of residual errors in the likelihood function, and in doing so can help avoid the potential over-conditioning of the posterior parameter distribution (Beven et al. 2008). However, in doing so GLUE also maps all uncertainty onto the posterior parameter distribution, which may lead to poorly constrained parameters. A key aspect when applying an informal likelihood function is that subjective decisions need to be made regarding the behavioural threshold used. It is important to evaluate the sensitivity of the model calibration to the retained threshold.

Recent extensions of the GLUE framework include the limits of acceptability approach (Beven 2006), which like some formal Bayesian approaches, attempts to incorporate the effects of observation error in model evaluation. The approach calculates a normalised evaluation score at each time-step to evaluate where during a simulation a model is behavioural (Liu et al. 2009). The approach has also been extended within the GLUE framework to consider model structural uncertainties (Krueger et al. 2010). The MCMC SCEM-UA algorithm has also been adapted to adequately explore posterior parameter space within GLUE (McMillan and Clark 2009).

Much debate exists in the modelling literature regarding the most appropriate application of formal and informal Bayesian approaches (Beven et al. 2008; Mantovan and Todini 2006; Stedinger et al. 2008; Vrugt et al. 2009b), and the suitability of simpler approximations for quantifying parameter uncertainty (Gallagher and Doherty 2007; Kang et al. 2009). All methods are currently limited by our inability to separate model structural error from input error a priori. Regardless of the applied method, data used for calibration should be representative of future forecasting conditions, and assumptions made during calibration should be explicitly evaluated.

4.3 Water Distribution System Model Calibration

A comparison of formal and informal Bayesian calibration approaches when applied to a Water Distribution Network hydraulic model calibration was made in Hutton et al. (2013). Both methods were applied to calibrate ten pipe roughness groups using a distributed set of 27 observation locations, and a time-series of 24 hourly observations at each location. In the formal Bayesian calibration a Gaussian likelihood function was assumed, and in the informal procedure a likelihood function based on the Nash Sutcliffe efficiency statistic was applied. Both methods were found to identify similar posterior parameter distributions. The informal Bayesian approach, however, produced uncertainty bounds which were too narrow and produced insufficient variability in the model predictions to adequately bracket the observations. During calibration, joint inference of water demand and pipe roughness should be considered, as assuming a fixed demand can lead to bias and lead to insufficient representation of true model uncertainty. In contrast to the informal prediction bounds, the prediction bounds derived from the formal Bayesian approach provided more adequate coverage of the observations and therefore of the prediction uncertainty. The reason is that the error model standard deviation was also inferred during the model calibration, alongside the roughness parameters. That said, posterior diagnostic checks revealed that the Gaussian assumption for the residual errors, which is widely applied during calibration, was insufficient to adequately describe the heavier tailed residual distributions. An additional consideration is that spatial correlation between observation locations should also be considered during calibration.

4.4

4.5 Urban Rainfall-Runoff Model Calibration

First order approximations are likely to be inappropriate to adequately characterise parameter and predictive uncertainty in urban rainfall runoff models given their nonlinearity (Vrugt and Bouten, 2002). Bayesian approaches have been applied more widely to urban rainfall runoff model calibration, including the application of formal and informal Bayesian approaches (Freni et al., 2009; Sun and Bertrans-Krajewski 2013). For explicit representation of rainfall uncertainty in a formal Bayesian procedure, the rainfall multiplier approach is fairly well established (Mcmillan et al 2011) and has recently been applied to urban catchments (Sun and Bertrans-Krajewski 2013). For explicit representation of discharge rating curve uncertainty, a number of papers have applied and evaluated in both formal and informal Bayesian settings (McMillan et al 2010; Liu et al. 2009). Model structural uncertainties have been dealt with explicitly with stochastic models using a grey-box modelling approach. Further details of uncertainty assessment methods can be found in Deletic et al. 2012).

5 Data Assimilation

Numerical models, once calibrated off-line, may be applied in real-time to understand the current state of the system being modelled, and also to make forecasts of future system states, which may be useful to inform the development of real-time control strategies. Whilst off-line calibration of a model may help to improve model accuracy, and also help to quantify predictive error, unfortunately in real-time modelling such offline calibration cannot correct for bias in the model predictions, which typically leads to divergence between what the model predicited states, and the true states of the system.

Advances in the collection of data in (near) real-time from system sensors using telemetric methods (see: Hart and Murray 2010; Ruggaber et al. 2007; Storey et al. 2011) , coupled with automated methods for data validation (Branisavljevic et al. 2010; Schilperoort et al. 2008), help to provide observations of the system. Such measurements can be compared to the equivalent predictions from the system model, which when combined, can be used to derive a near optimal estimate of the system state.

Data Assimiltation refers to a range of methods that combine uncertain models with often uncertain data to derive a refined estimate of the current system state.

In general terms a model is propagated forwards in time from a set of initial conditions at a previous time-step ($\mathbf{X}_{t-1}$) to the current time (t), as driven by a set of driving conditions ($\mathbf{D}$), and dependent on a vector of model parmeters (θ):

$$\boldsymbol{X}_t = \mathbf{M}(\boldsymbol{\theta}, \mathbf{X}_{t-1}, \mathbf{D}) + \omega_t \tag{2}$$

where ω_t is the model error term. In order to assimilate obsevations ($\mathbf{Y}_t$) to update the model they need to be related to the model states and parameters, through an observation operator ($\mathbf{H}$):

$$\mathbf{Y}_t = \mathbf{H}(\boldsymbol{\theta}, \mathbf{X}_t) + \varepsilon_t \tag{3}$$

where ε_t denotes the observation error. Model states are then updated through Data Assimilation considering the relative magnitude of the model error and observation error. ***Regardless of the specific method applied to assimilate observations, a key factor governing the success of the assimilation procedure is the ability to define the distributions ω_t and ε_t.*** Well defined observations with small errors will lead to greater corrections of the model. However, if the observations are unreliable, then smaller corrections to the model will be made. Miss-specification of the observation error term may lead to a model that is over-corrected.

Computational resources are often limited in real-time modelling. During real-time modelling a number of tasks may need to be scheduled and performed at regular intervals: Data processing and validation; assimilation of data to update the model states; model propagation to the next observation timestep, and subsequently into a model forecast of future system states; potentially repeated model forecasts if the model is to be used as part of a control optimisation procedure. Thus, computational resources need to be divided into a number of tasks – the computational time available will depend upon the frequency of observations, and the frequency with which forecasts of future system states are required, and therefore the potential frequency at which warnings of future system states, or new control settings are required. ***Computational resources therefore represent a key constraint on the type of Data Assimilation procedure that may be applied in real-time modelling.***

Given computational limitations, deterministic error correction procedures have been widely applied. Such procedures may be referred to as Predictor-Corrector schemes, where a model prediction is made, and subsequently the observations are used to correct the model. The most popular deterministic correction procedure is the Kalman Filter (KF). The forecasted states of a model are corrected with an additive term that uses the observation operator to map the observations to the model states; the relative strength of the correction depends on the relative error of the model and the observations. The method works provided it is a Gaussian Linear System (Burgers et al. 1998). The Extended Kalman Filter (EKF) was developed to work better than the KF in cases of non-linearity, where the model is approximated with a tangent linear operator (Jacobian; Evensen 2003), however is unsuitable in the case of large system non-linearities (Hoteit et al. 2005).

Other deterministic procedures have applied itterative approaches in the correction step, including Preis et al (2011) who applied a genetic algorithm in real time, and Kang and Lansey (2009) who applied an itterative tracking state estimator. These methods may be better applied when the system is non-linear. From the computational perspective, however, an itterative solution may also be computationally expensive. As per the guidleines set out in Figure 2, if computational resources are constrained, and the issues of model linearity are a problem that invalidate the application of deterministic procedures, then some form of reduced or data driven model that provides an approximation to the often physically based model being applied, may be appropriate (See section 6).

If computational resources are less restrictive, ensemble assimilation procedures are popular. A key reason for their popularity is their improved performance in dealing with model non-linearity, and by using an ensemble of models to represent model uncertainy. The Ensemble Kalman Filter (EnKF) has recieved much attention and application in hydrology. The model error covariance is represented by the ensemble, which avoids the computational costs associated with propagating the error covariance matrix (Burgers et al. 1998). A number of variants and extensions of the approach have been made, including the Ensemble Square root Filter (Tippett et al. 2003). A key limitation of the EnKF is that optimal performance is restricted to to multi-gaussian distributed states and parameters (Liu et al. 2012).

An alternative ensemble procedure to the EnKF is Sequential Monte Carlo sampling, also known as the Particle Filter (Arulampalam et al. 2002). In the Particle Filter the posterior density of model states s represents by an ensemble of models, or particles, each with an associated weight. As the model is propagated forwards in time, observations are used to re-weight each particle and provide an updated estimate of the system states. The Particle filter is potentially more flexibly than the Ensemble Kalman filter in that a range of non-gaussian likelihoods may be used to re-weight the ensemble, and the method is applicably to non-gaussian state space models. The particle filter in its most basic form however, is rarely applied in practise – the ensemble may diverge from the true states of the model, and therefore requires some correction to prevent divergence. Typically this takes the form of duplicating better performing models at the expense of poorer performing models. However, this can lead to an additional problem of insufficient diversity in the ensemble. More recent procedures for updating inclue applying some form of MCMC sampling (Moradkhani et al. 2012), and retrospective approaches, which seek to average the weights applied to each ensemble member over a horizon of observations (Noh et al. 2011).

A number of papers have applied and compared ensemble approaches for Data Assimilation, with often different results as to which methods work best (Noh et al. 2013; Pasetto et al. 2012; Pham 2001). A key issue is that there are often different implementations of these methods, specific to certain settings determined by which components of the ensemble are perturbed, alongside the specific implementation of the Data Assimilation scheme, including approaches that merge different methods (van Delft et al. 2009). Therefore, it is difficult to generalise about the most preferred method without more detailed application in a specific setting. During one comparison of the EnKF and the Particle Filter by Pasetto et al (2012) however, a key issue arose during the updating of a Physically based catchment model. Problems with the state update from the EnKF resulted in problems in the non-linear solver of model equations, because the Gaussian assumptions in the update led to physically unrealistic states. Therefore, if the assumptions made in the correction step are physically inconsistent, then such corrections can lead to increased computational time, and importantly physically inconsistent models. Thus, it is important to consider the physical consistency during application of Data Assimilation method.

5.1 Water Distribution Systems

In Water Distribution system models, which are generally fairly linear, deterministic correction procedures are the most widely applied Data Assimilation method. The Kalman Filter has been applied for demand estimation, and performed well in branched networks, but struggled more when applied to looped networks (Kang and Lansey 2009) because of increased nonlinearity. The EKF has been applied for water demand estimation (Nasseri et al. 2010) and also by Shang et al. (2006) for real-time update of demand estimates. Predictor-corrector schemes have been applied by Preis et al (2011) who applied a GA to update Demand Multiplication Factors for a skeletonised network in real-time, with an objective function that accounted for measurement noise.

In general the time-lag between observation locations in a system and the system states to be updated does not need to be considered, particularly in application of Data Assimilation techniques to steady-state models. However, a key factor that will affect the mapping of observation information to state space is the presence of control structures in the distribution network. Valve closure, and also the operation of pumps will determine how 'connected' different locations are within the network, and therefore will affect how well it is possible to map observation information to update the estimates of different parts of the network. It is therefore important to consider, during the operation of control structures, to regularly update the covariance matrix relating locations in the network.

5.2 Urban rainfall runoff models

Data Assimilation schemes have been widely applied to natural catchment systems (Liu et al. 2012), yet have received relatively little attention in urban settings. A number of key issues need to be considering when applying Data Assimilation procedures to urban rainfall runoff models, somewhat independent of the specific assimilation scheme employed including the observation frequency, the lag time in the system, the speed of process operation and the presence of control structures in the system.

Lag-time between downstream observations and upstream model states. In urban rainfall runoff systems, there is a lag time between when the rainfall hits the catchment, travels down through different parts of the system, and arrives downstream at the point at which the system is observed – typically a pipe flow rate or storage tank water level. Likewise, the model of the system will also propagate flow downstream. Part of the residual error downstream between observed and predicted flow will therefore result from errors in the upstream representation of the catchment system, and it is typically these states that require correction. If the speed of the processes occurring are slow, then using the downstream observation to update the state of the system upstream at the same time-step may work and lead to improved predictions (Borup et al. 2011). However, if the processes occurring in the system are faster – e.g. the passing of a convective rain cell, and the subsequently derived hydrograph - is faster than the lag-time in the system, then the discharge rate downstream may have no relation to the volume of water in the upstream part of the catchment. Therefore, if the instantaneous relationship between the observation location downstream and the upstream model state to be updated is used, they may have no relationship, and result in physically unrealistic data assimilation, and potentially lead the model to diverge further from the predictions (McMillan et al. 2013). Furthermore, if frequency of assimilated observations in the model is smaller than the lag time in the system, then when updating upstream states at the current time-step, updates will occur without the system having sufficient time to allow these updates to propagate downstream to affect the residual error at the observation location. The result is that independent updates are made to the same model state, which can result in oscillatory system behaviour (McMillan et al. 2013).

To overcome these limitations the lag time in the system should be considered: once the observation time has been reached, the model states should be re-set to their previous values at several time-steps in the past, depending on the lag in the system. The model states can then be updated, and the model re-propagated forwards in time to the next observation, to make a model forecast. A variant on this approach include making multiple re-propagations of the model forwards in time from each time-lag in the past to account for different time-lags in the system (McMillan et al. 2013), and their affect on the residual error downstream at a point in time.

Another way of looking at this problem is to consider the flow attenuation in the system. In the same way as a single observation at a point in time downstream is influenced by multiple upstream model states, typically at different time-lags, so because of flow attenuation does a model state at a single point in time influence multiple observations at different times at the same location downstream. A method of accounting for flow attenuation is therefore to run a retrospective assimilation approach and during updating use all observations during a window to provide an updated estimate of the system state (Noh et al. 2011).

Finally, control structures in the system, such as gates, weirs and tanks can effectively decouple the signal received downstream from upstream, if for example water is diverted via a weir into a storage tank. The result is that errors in the upstream model are no longer propagated downstream, and therefore the downstream observation does not provide any information that can be used to improve the model prediction upstream. It is important, therefore, to consider the potential influence of control structures on the performance of an assimilation scheme. Ideally multiple observations will be available in a system to alleviate this problem. However, the issue should be considered during application, and ideally during sensor location.

6 Model Forecasting

Following Data Assimilation a model may be propagated forwards into the future, driven by a set of either measured or simulated driving conditions (e.g. rainfall or water demand), to make predictions of future system states. Such predictions are typically useful over some operating time horizon, such that they may be used to inform some real-time control strategy. Data Assimilation techniques, by incorporating the latest system observations, may be applied to update the states of the system, but only up to the point of the most recent observation in the system. Thus Data Assimilation can only correct the initial conditions of the model; over time these conditions are washed out, as the different sources of input and model structural error lead to divergence between what the model predicts and the true system (Madsen and Skotner 2005).

6.1 Error-Correction Procedures

Error-correction procedures have been developed, where a form of data driven model, such as an Artificial Neural Network (ANN) is calibrated to a recent time-series of model residual errors, and applied as an additional term to the model prediction. For example, Abebe and Price (2003) improved 1-6 hour rainfall-runoff model forecasts by adding an ANN model prediction, which was calibrated to past model residuals. Thus the method is effectively used to predict future differences between the model and the observations, and is thus likely to be most effective when these differences are persistent – e.g. the model residuals show strong autocorrelation. Other error correction methods include autoregressive time series models (Hostache et al. 2011; Lekkas et al. 2001) and genetic programming (Khu et al. 2001). Furthermore, Madsen and Skotner (2005) extended error correction procedures to also update model states; pre-determined gain functions were used to update state variables using the innovation determined off-line at measurement locations. A slightly different approach is to apply a multiplicative term to correct the bias in the model prediction (Smith et al. 2012).

Such error-correction procedures are likely to be more computationally efficient than ensemble data assimilation procedures, and have the additional advantage of not interfering with the model code, which has additional practical advantages. Furthermore, unlike Data Assimilation methods, error correction procedures correct the acual forecast of the system. It is important to consider, however, that during more extreme events, previous shortcomings in the system model being corrected may not previously have been experienced, which may render the error correction inappropriate, particularly if the residual errors are non-stationary. This point is general to all models during calibration – a model should not generally be applied outside of the conditions for which it was calibrated.

Furthermore, it is pertinent to consider that error-correction models are models themselves in their own right. Considering them as such means that as with the model that seeks to represent the processes themsevles, error-correction models should also be calibrated considering the uncertainty in the predictions and tehcniques for calibration, as considered in section 4. Furthermore, error correction models are often adaptive in the sense that they are calibrated to a set of most recent observations. It may be difficult to adapt in real-time through offline calibration approaches given computational cost, and thus Data Assimilation techniques may also be appropriate to apply to the error-correction model (Smith et al. 2012). Additionally, if the error-correction procedure is also applied with a conventional Data Assimilation procedure, it is important to take into account the effect of the Data Assimilation procedure on the residual errors when developing and calibrating the residual error model, as the time-series of past residual errors will result from application of the model plus Data Assimilation.

6.2 Approximating Ensemble Forecasts

In addition to error correction procedures that may be applied to deterministic models when running an ensemble of physical models is too computationally inefficient, approached have been applied than make approximate forecasts of the ensemble. These approaches may be broadly divided into approaches that attempt to fit a data driven model to the uncertainty bounds derived from an offline calibration – such as Shrestha et al. (2009) who trained an ANN to reproduce the upper and lower prediction intervals (90%) derived from GLUE calibration, which was then ran alongside the deterministic model to produce upper and lower forecast predictio bounds – and also methods that attempt to make an approximate ensemble forecast by chosing a sub-set of models to represent key sources of model error during error propagation and forecasting (Pappenberger et al. 2005; Zappa et al. 2011). The application of model sensitivity analysis is recommended as a means to derive a representative approximate ensemble of different sources of model error.

6.3 Data-Driven Modelling

Data-driven models are models that have no physical structure informed by the processes they attempt to represent – rather they seek to simulate the relationship between a set of input and output variables. As discussed above such models can be applied in a number of ways to supplement the application of the system model to derive improved understanding of the model uncertainties, and also to reduce those model uncertainties:

- As an alternative to the process model of the system.
- As an error correction procedure.
- As a model of prediction bounds derived offline.

As mentioned above, regardless of how the data driven model is applied, whilst it may be used to reduce or represent model uncertainty, the model fitting itself contains additional uncertainty depending on the time series of values it is trying to represent – be it the system observations, the residuals between the system model’s predictions and these observations, or prediction bounds of the system model about these observations. As a consequence the guidelines for model calibration presented in Section 4 should also be considered when also applying Data Driven models.

6.4 Forecast Evaluation

During forecasting, and in particular ensemble forecasting, it is important to evaluate the quality of the model predictions, particularly if these are to be used in the decision making process. Therefore, in addition to the Residual analysis metrics considered in Section 4.2, which are well suited to deterministic model predictions, and evaluations of Maximum Likelihood Estimates, additional forecasting measures are recommended, further details of which can be found in Gneiting et al (2007), with example application in Breinholt et al (2012):

- **Reliability Bias:** Reliability bias measures the discrepancy between the true proportion of observations that fall within the prescribed uncertainty bounds, and those defined by the bounds. For example, a reliable 90% prediction bound would bound 90% of the observations.
- **Sharpness:** Measures the average width of a particular uncertainty bound over a set of observations. A good forecast may therefore be considered to score well in terms of both sharpness and reliability, for which there is an inevitable trade-off.
- **Interval Skill Score Criterion:** Seeks to combine reliability and sharpness into a unified measure. A forecast prediction bound is penalised proportional to its percentile.

It is important to calculate these metrics for different forecasting horizons, and also for different periods during a simulation. For example, in a water demand forecasting model these metrics are likely to change by time of day between peak and minimum demand. Additionally, in forecasting catchment discharge, differences are likely to occur over different portions of the hydrograph that may be evaluated with these metrics of performance.

7 References

Abebe, A. J., and Price, R. K. (2003). "Managing uncertainty in hydrological models using complementary models." *Hydrological Sciences Journal-Journal Des Sciences Hydrologiques*, 48(5), 679-692.

Ackers, J. C., Butler, D., and May, R. W. P. (1996). "Design of sewers to control sediment problems." *Report Number CIRIA 141*, Construction Industry research and Information Association, London.

Almeida, M. C., Bulter, D., and Frielder, E. (1999). "At-source domestic wastewater quality." *Urban Water*, 1, 49-55.

Aronica, G., Freni, G., and Oliveri, E. (2005). "Uncertainty analysis of the influence of rainfall time resolution in the modelling of urban drainage systems." *Hydrological Processes*, 19(5), 1055-1071.

Arulampalam, M. S., Maskell, S., Gordon, N., and Clapp, T. (2002). "A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking." *Ieee Transactions on Signal Processing*, 50(2), 174-188.

Ashley, R. M., Hvitved-Jacobsen, T., and Bertrand-Krajewski, J. L. (1999). "Quo vadis sewer process modelling?" *Water Science and Technology*, 39(9), 9-22.

Bahadur, R., Johnson, J., Janke, R., and Samuels, W. B. (2006). "Impact of Model Skeletonization on Water Distribution Model Parameters as Related to Water Quality and Contaminant Consequence Assessment." *8th Annual Water Distribution Systems Analysis Symposium, CIncinati, Ohio, USA*.

Benedetti, L., Bixio, D., Claeys, F., and Vanrolleghem, P. A. (2008). "Tools to support a model-based methodology for emission/immission and benefit/cost/risk analysis of wastewater systems that considers uncertainty." *Environmental Modelling & Software*, 23(8), 1082-1091.

Bertrand-Krajewski J.-L., Bardin J.-P., Mourad M., Béranger Y. (2003). Accounting for sensor calibration, data validation, and measurement and sampling uncertainties in monitoring of urban drainage systems. *Water Science and Technology*, 47(2), 95-102. ISSN 0273-1223.

Beven, K. (2006). "A manifesto for the equifinality thesis." *Journal of Hydrology*, 320(1-2), 18-36.

Beven, K., and Binley, A. (1992). "The Future of Distributed Models - Model Calibration and Uncertainty Prediction." *Hydrological Processes*, 6(3), 279-298.

Beven, K. J., Smith, P. J., and Freer, J. E. (2008). "So just why would a modeller choose to be incoherent?" *Journal of Hydrology*, 354(1-4), 15-32.

Borup, M., Grum, M., and Mikkelsen, P. S. (2011). "Real time adjustment of slow changing flow components in distributed urban runoff models." *12th International Conference on Urban Drainage*, Porto Alegre, Brasil.

Boulos, P. F., Lansey, K. E., and Karney, B. W. (2004). *Comprehensive Water Distribution Systems Analysis Handbook For Engineers and Planners*, MWH Soft, Inc., Pasadena, California.

Box, G. E. P., and Cox, D. R. (1982). "An Analysis of Transformations Revisited, Rebutted." *Journal of the American Statistical Association*, 77(377), 209-210.

Branisavljevic, N., Kapelan, Z., and Prodanovic, D. (2010). "Online time data series pre-processing for the improved performance of anomaly detection methods." *Integrating water systems: Proceedings of the Tenth International Conference on Computing and Control for the Water Industry, CCWI 2009 'Integrating Water Systems'*, Sheffield, UK.

Branisavljevic, N., Prodanovic, D., and Ivetic, M. (2009). "Uncertainty reduction in water distribution network modelling using system inflow data." *Urban Water Journal*, 6(1), 69-79.

Breinholt, A., Møller, J. K., Madsen, H., & Mikkelsen, P. S. (2012). A formal statistical approach to representing uncertainty in rainfall-runoff modelling with focus on residual analysis and probabilistic output evaluation-distinguishing simulation and prediction. *Journal of Hydrology*.

Burgers, G., van Leeuwen, P. J., and Evensen, G. (1998). "Analysis scheme in the ensemble Kalman filter." *Monthly Weather Review*, 126(6), 1719-1724.

Bush, C. A., and Uber, J. G. (1998). "Sampling design methods for water distribution model calibration." *Journal of Water Resources Planning and Management-Asce*, 124(6), 334-344.

Butler, D. (1993). "The influence of dwelling occupancy and day of the week on domestic appliance wastewater discharges " *Building and Environment*, 28, 73-79.

Butler, D., and Parkinson, J. (1997). "Towards sustainable urban drainage." *Water Science and Technology*, 35(9), 53-63.

Butler, D., and Schutze, M. (2005). "Integrating simulation models with a view to optimal control of urban wastewater systems." *Environmental Modelling & Software*, 20(4), 415-426.

Campolongo, F., Cariboni, J., and Saltelli, A. (2007). "An effective screening design for sensitivity analysis of large models." *Environmental Modelling & Software*, 22(10), 1509-1518.

Ciach, G. J. (2003). "Local random errors in tipping-bucket rain gauge measurements." *Journal of Atmospheric and Oceanic Technology*, 20(5), 752-759.

Collier, C. G. (2009). "On the propagation of uncertainty in weather radar estimates of rainfall through hydrological models." *Meteorological Applications*, 16(1), 35-40.

Davidson, J. W., and Bouchart, F. J. C. (2006). "Adjusting nodal demands in SCADA constrained real-time water distribution network models." *Journal of Hydraulic Engineering-Asce*, 132(1), 102-110.

De Sutter, R., Rushforth, P., Tait, S., Huygens, M., Verhoeven, R., and Saul, A. (2003). "Validation of existing bed load transport formulas using in-sewer sediment." *Journal of Hydraulic Engineering-Asce*, 129(4), 325-333.

Deletic, A., Dotto, C. B. S., McCarthy, D. T., Kleidorfer, M., Freni, G., Mannina, G., ... & Tait, S. (2012). Assessing uncertainties in urban drainage models. *Physics and Chemistry of the Earth, Parts A/B/C*, *42*, 3-10.

Draper, D. (1995). "Assessment and Propagation of Model Uncertainty." *Journal of the Royal Statistical Society Series B-Methodological*, 57(1), 45-97.

Evensen, G. (2003). "The Ensemble Kalman Filter: theoretical formulation and practical implementation." *Ocean Dynamics*, 53, 343-367.

Freni, G., and Mannina, G. (2010). "Bayesian approach for uncertainty quantification in water quality modelling: The influence of prior distribution." *Journal of Hydrology*, 392(1-2), 31-39.

Friedler, E., Brown, D. M., and Butler, D. (1996). "A study of WC derived sewer solids." *Water Science and Technology*, 33(9), 17-24.

Friedler, E., and Butler, D. (1996). "Quantifying the inherent uncertainty in the quantity and quality of domestic wastewater." *Water Science and Technology*, 33(2), 65-78.

Gallagher, M., and Doherty, J. (2007). "Parameter estimation and uncertainty analysis for a watershed model." *Environmental Modelling & Software*, 22(7), 1000-1020.

Gelman, A., & Shalizi, C. R. (2012). Philosophy and the practice of Bayesian statistics. *British Journal of Mathematical and Statistical Psychology*.

Gernaey, K. V., van Loosdrecht, M. C. M., Henze, M., Lind, M., and Jorgensen, S. B. (2004). "Activated sludge wastewater treatment plant modelling and simulation: state of the art." *Environmental Modelling & Software*, 19(9), 763-783.

Giustolisi, O. (2010). "Considering Actual Pipe Connections in Water Distribution Network Analysis." *Journal of Hydraulic Engineering-Asce*, 136(11), 889-900.

Giustolisi, O., Kapelan, Z., and Savic, D. (2008a). "Extended Period Simulation Analysis Considering Valve Shutdowns." *Journal of Water Resources Planning and Management-Asce*, 134(6), 527-537.

Giustolisi, O., Savic, D., and Kapelan, Z. (2008b). "Pressure-driven demand and leakage simulation for water distribution networks." *Journal of Hydraulic Engineering-Asce*, 134(5), 626-635.

Gneiting, T., Balabdaoui, F., & Raftery, A. E. (2007). Probabilistic forecasts, calibration and sharpness. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *69*(2), 243-268.

Guzman, K., La Motta, E. J., McCorquodale, J. A., Rojas, S., and Ermogenous, M. (2007). "Effect of biofilm formation on roughness coefficient and solids deposition in small-diameter PVC sewer pipes." *Journal of Environmental Engineering-Asce*, 133(4), 364-371.

Hall, J. W. (2003). "Handling uncertainty in the hydroinformatic process." *Journal of Hydroinformatics*, 5, 215-232.

Hart, W. E., and Murray, R. (2010). "Review of Sensor Placement Strategies for Contamination Warning Systems in Drinking Water Distribution Systems." *Journal of Water Resources Planning and Management-Asce*, 136(6), 611-619.

Herrera, M., Torgo, L., Izquierdo, J., and Perez-Garcia, R. (2010). "Predictive models for forecasting hourly urban water demand." *Journal of Hydrology*, 387(1-2), 141-150.

Hostache, R., Matgen, P., Montanari, A., Montanari, M., Hoffmann, L., and Pfister, L. (2011). "Propagation of uncertainties in coupled hydro-meteorological forecasting systems: A stochastic approach for the assessment of the total predictive uncertainty." *Atmospheric Research*, 100(2-3), 263-274.

Hoteit, I., Korres, G., and Triantafyllou, G. (2005). "Comparison of extended and ensemble based Kalman filters with low and high resolution primitive equation ocean models." *Nonlinear Processes in Geophysics*, 12(5), 755-765.

Hutton, C.J., Kapelan, Z., Vamvakeridou-Lyroudia, L., Savic, D. (2012) Dealing with Uncertainty in Water Distribution Systems' Models: a Framework for Real-Time Modeling and Data Assimilation. *Journal of Water Resources Planning and Management*, http://ascelibrary.org/doi/abs/10.1061/%28ASCE%29WR.1943-5452.0000325

Hutton, C.J., Kapelan, Z., Vamvakeridou-Lyroudia, L., Savic, D. (2013) The Application of Formal and Informal Bayesian Methods for Water Distribution Hydraulic Model Calibration. *Journal of Water Resources Planning and Management*, http://ascelibrary.org/doi/abs/10.1061/%28ASCE%29WR.1943-5452.0000412

Jacobsen, L. B., and Kamojjala. (2009). "Tools and processes for calibrating large all-pipes models." *Urban Water Journal*, 6(1), 29-38.

Jauregui, E., and Romales, E. (1996). "Urban effects on convective precipitation in Mexico city." *Atmospheric Environment*, 30(20), 3383-3389.

Jonkergouw, P. M. R., Khu, S. T., Kapelan, Z. S., and Savic, D. A. (2008). "Water quality model calibration under unknown demands." *Journal of Water Resources Planning and Management-Asce*, 134(4), 326-336.

Jung, B. S., Boulos, P. F., and Wood, D. J. (2007). "Pitfalls of water distribution model skeletonization for surge analysis." *Journal American Water Works Association*, 99(12), 87-98.

Kang, D., and Lansey, K. (2009). "Real-Time Demand Estimation and Confidence Limit Analysis for Water Distribution Systems." *Journal of Hydraulic Engineering-Asce*, 135(10), 825-837.

Kang, D. S., Pasha, M. F. K., and Lansey, K. (2009). "Approximate methods for uncertainty analysis of water distribution systems." *Urban Water Journal*, 6(3), 233-249.

Kapelan, Z. S., Savic, D. A., and Walters, G. A. (2007). "Calibration of water distribution hydraulic models using a Bayesian-Type procedure." *Journal of Hydraulic Engineering-Asce*, 133(8), 927-936.

Khu, S. T., Liong, S. Y., Babovic, V., Madsen, H., and Muttil, N. (2001). "Genetic programming and its application in real-time runoff forecasting." *Journal of the American Water Resources Association*, 37(2), 439-451.

Kleiner, Y., and Rajani, B. (2001). "Comprehensive review of structural deterioration of water mains: statistical models." *Urban Water* 3, 131-150.

Korving, H., van Noortwijk, J. M., van Gelder, P. H. A. J. M., and Parkhi, R. S. (2003). "Coping with uncertainty in sewer system rehabilitation." *Safety and Reliability, Vols 1 and 2*, 959-967.

Krein, A., Salvia-Castellvi, M., Iffly, J. F., Pfister, L., and Hoffmann, L. (2007). "The importance of precedent hydro-climatological conditions for the mass transfer of pollutants in separated sewer systems and corresponding tributaries during storm events." *Water Air and Soil Pollution*, 182(1-4), 357-368.

Krueger, T., Freer, J., Quinton, J. N., Macleod, C. J. A., Bilotta, G. S., Brazier, R. E., Butler, P., and Haygarth, P. M. (2010). "Ensemble evaluation of hydrological model hypotheses." *Water Resources Research*, 46.

Kutiel, H., and Sharon, D. (1980). "Diurnal-Variation of Rainfall in Israel." *Archiv Fur Meteorologie Geophysik Und Bioklimatologie Serie a-Meteorologie Und Geophysik*, 29(4), 387-395.

Kutiel, H., and Sharon, D. (1981). "Diurnal-Variation in the Spatial Structure of Rainfall in the Northern Negev Desert Israel." *Archives for Meteorology Geophysics and Bioclimatology Series B-Theoretical and Applied Climatology*, 29(3), 239-243.

Lansey, K. E., El-Shorbagy, W., Ahmed, I., Araujo, J., and Haan, C. T. (2001). "Calibration assessment and data collection for water distribution networks." *Journal of Hydraulic Engineering-Asce*, 127(4), 270-279.

Lekkas, D. F., Imrie, C. E., and Lees, M. J. (2001). "Improved non-linear transfer function and nerual network methods of flow routing for real-time forecasting." *Journal of Hydroinformatics*, 3, 153-164.

Lepot M., Aubin J.-B., Bertrand-Krajewski J.-L. (2013). Accuracy of different sensors for the estimation of pollutant concentrations (total suspended solids, total and dissolved chemical oxygen demand) in wastewater and stormwater. *Water Science and Technology*, 68(2), 462-471. doi: 10.2166/wst.2013.276.

Liu, Y., Weerts, A. H., Clark, M., Franssen, H. J. H., Kumar, S., Moradkhani, H., Seo, D. J., Schwanenberg, D., Smith, P., van Dijk, A. I. J. M., van Velzen, N., He, M., Lee, H., Noh, S. J., Rakovec, O., and Restrepo, P. (2012). "Advancing data assimilation in operational hydrologic forecasting: progresses, challenges, and emerging opportunities." *Hydrology and Earth System Sciences*, 16(10), 3863-3887.

Liu, Y. L., Freer, J., Beven, K., and Matgen, P. (2009). "Towards a limits of acceptability approach to the calibration of hydrological models: Extending observation error." *Journal of Hydrology*, 367(1-2), 93-103.

Madsen, H., and Skotner, C. (2005). "Adaptive state updating in real-time river flow forecasting - a combined filtering and error forecasting procedure." *Journal of Hydrology*, 308(1-4), 302-312.

Mantovan, P., and Todini, E. (2006). "Hydrological forecasting uncertainty assessment: Incoherence of the GLUE methodology." *Journal of Hydrology*, 330(1-2), 368-381.

McMillan, H., and Clark, M. (2009). "Rainfall-runoff model calibration using informal likelihood measures within a Markov chain Monte Carlo sampling scheme." *Water Resources Research*, 45.

McMillan, H., Jackson, B., Clark, M., Kavetski, D., & Woods, R. (2011). Rainfall uncertainty in hydrological modelling: An evaluation of multiplicative error models. *Journal of Hydrology*, *400*(1), 83-94.

McMillan, H. K., Hreinsson, E. O., Clark, M. P., Singh, S. K., Zammit, C., and Uddstrom, M. J. (2013). "Operational hydrological data assimilation with the recursive ensemble Kalman filter." *Hydrology and Earth System Sciences*, 17(1), 21-38.

Menaia, J., Coelho, S. T., Lopes, A., Fonte, E., and Palma, J. (2003). "Dependency of bulk chlorine decay rates on flow.velocity in water distribution networks." *3rd World Water Congress: Water Services Management, Operations and Monitoring*, 3(1-2), 209-214.

Moradkhani, H., DeChant, C. M., and Sorooshian, S. (2012). "Evolution of ensemble data assimilation for uncertainty quantification using the particle filter-Markov chain Monte Carlo method." *Water Resources Research*, 48.

Nasseri, M., Moeini, A., and Tabesh, M. (2010). "Forecasting monthly urban water demand using Extended Kalman Filter and Genetic Programming." *Expert systems with Applications*, Article in Press, corrected proof.

Noh, S. J., Tachikawa, Y., Shiiba, M., and Kim, S. (2011). "Applying sequential Monte Carlo methods into a distributed hydrologic model: lagged particle filtering approach with regularization." *Hydrology and Earth System Sciences*, 8, 3383-3420.

Noh, S. J., Tachikawa, Y., Shiiba, M., and Kim, S. (2013). "Sequential data assimilation for streamflow forecasting using a distributed hydrologic model: particle filtering and ensemble Kalman filtering." *Floods: From Risk to Opportunity*, 357, 341-349.

Ozger, S., and Mays, L. W. (2004). "Optimal Location of Isolation Valves: A Reliability Approach." Water Supply Systems Security, L. W. Mays, ed., McGraw-Hill, New York.

Pappenberger, F., Beven, K. J., Hunter, N. M., Bates, P. D., Gouweleeuw, B. T., Thielen, J., and de Roo, A. P. J. (2005). "Cascading model uncertainty from medium range weather forecasts (10 days) through a rainfall-runoff model to flood inundation predictions within the European Flood Forecasting System (EFFS)." *Hydrology and Earth System Sciences*, 9(4), 381-393.

Pasetto, D., Camporese, M., and Putti, M. (2012). "Ensemble Kalman filter versus particle filter for a physically-based coupled surface-subsurface model." *Advances in Water Resources*, 47, 1-13.

Pham, D. T. (2001). "Stochastic methods for sequential data assimilation in strongly nonlinear systems." *Monthly Weather Review*, 129(5), 1194-1207.

Pomeroy, R. D. (1967). "Flow Velocities in Small Sewers." *Journal Water Pollution Control Federation*, 39(9), 1525-&.

Preis, A., Whittle, A. J., Ostfeld, A., and Perelman, L. (2011). "Efficient Hydraulic State Estimation Technique Using Reduced Models of Urban Water Networks." *Journal of Water Resources Planning and Management-Asce*, 137(4), 343-351.

Rauch, W., Aalderink, H., Krebs, P., Schilling, W., and Vanrolleghem, P. (1998). "Requirements for integrated wastewater models - Driven by receiving water objectives." *Water Science and Technology*, 38(11), 97-104.

Rauch, W., Bertrand-Krajewski, J. L., Krebs, P., Mark, O., Schilling, W., Schutze, M., and Vanrolleghem, P. A. (2002). "Deterministic modelling of integrated urban drainage systems." *Water Science and Technology*, 45(3), 81-94.

Rauch, W., Vanhooren, H., and Vanrolleghem, P. A. (1999). "A simplified mixed-culture biofilm model." *Water Research*, 33(9), 2148-2162.

Renard, B., Kavetski, D., Kuczera, G., Thyer, M., and Franks, S. W. (2010). "Understanding predictive uncertainty in hydrologic modeling: The challenge of identifying input and structural errors." *Water Resources Research*, 46, -.

Rodriguez-Puebla, C., Encinas, A. H., Nieto, S., and Garmendia, J. (1998). "Spatial and temporal patterns of annual precipitation variability over the Iberian Peninsula." *International Journal of Climatology*, 18(3), 299-316.

Ruggaber, T. P., Talley, J. W., and Montestruque, L. A. (2007). "Using Embedded Sensor Networks to Monitoring, Control, and Reduce CSO Events: A Pilot Study." *Environmental Engineering Science*, 24(2), 172-182.

Savic, D. A., Kapelan, Z. S., and Jonkergouw, P. M. R. (2009). "Quo vadis water distribution model calibration?" *Urban Water Journal*, 6(1), 3-22.

Schilperoort, R. P. S., Dirksen, J., and Clemens, F. H. L. R. (2008). "Practical aspects for long-term monitoring campaigns: pifals to avoid to maximise data yield." *11th International Conference on Urban Drainage*, Edinburgh, Scotland, UK.

Schoups, G., and 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, -.

Sevruk, B. (1996). "Adjustment of tipping-bucket precipitation gauge measurements." *Atmospheric Research*, 42(1-4), 237-246.

Sevruk, B., Hertig, J. A., and Tettamanti, R. (1994). "The Effect of Orifice Rim Thickness on the Wind-Speed above Precipitation Gauges." *Atmospheric Environment*, 28(11), 1939-1944.

Sevruk, B., and Nespor, V. (1998). "Empirical and theoretical assessment of the wind induced error of rain measurement." *Water Science and Technology*, 37(11), 171-178.

Shang, F., Uber, J. G., van Bloemen Waanders, B. G., Boccelli, D., and Janke, R. (2006). "Real Time Water Demand Estimation in Water Distribution System." *8th Annial water Distribution Systems Analysis Symposium*, Cincinnati, Ohio, USA.

Shrestha, D. L., Kayastha, N., and Solomatine, D. P. (2009). "A novel approach to parameter uncertainty analysis of hydrological models using neural networks." *Hydrology and Earth System Sciences*, 13(7), 1235-1248.

Siegrist, R., Witt, M., and Boyle, W. C. (1976). "Characteristics of Rural Household Wastewater." *Journal of the Environmental Engineering Division-Asce*, 102(3), 533-548.

Sin, G., Gernaey, K. V., Neumann, M. B., van Loosdrecht, M. C. M., and Gujer, W. (2009). "Uncertainty analysis in WWTP model applications: A critical discussion using an example from design." *Water Research*, 43(11), 2894-2906.

Smith, P., Beven, K. J., and Tawn, J. A. (2008). "Informal likelihood measures in model assessment: Theoretic development and investigation." *Advances in Water Resources*, 31(8), 1087-1100.

Smith, P. J., Beven, K. J., Weerts, A. H., and Leedal, D. (2012). "Adaptive correction of deterministic models to produce probabilistic forecasts." *Hydrology and Earth System Sciences*, 16(8), 2783-2799.

Stedinger, J. R., Vogel, R. M., Lee, S. U., and Batchelder, R. (2008). "Appraisal of the generalized likelihood uncertainty estimation (GLUE) method." *Water Resources Research*, 44, -.

Storey, M. V., van der Gaag, B., and Burns, B. P. (2011). "Advances in on-line drinking water quality monitoring and early warning systems." *Water Research*, 45(2), 741-747.

Stransky, D., Bares, V., and Fatka, P. (2007). "The effect of rainfall measurement uncertainties on rainfall-runoff processes modelling." *Water Science and Technology*, 55(4), 103-111.

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.

Thielen, J., and Gadian, A. (1997). "Influence of topography and urban heat island effects on the outbreak of convective storms under unstable meteorological conditions: a numerical study." *Meteorological Applications*, 4, 139-149.

Thyer, M., Renard, B., Kavetski, D., Kuczera, G., Franks, S. W., and Srikanthan, S. (2009). "Critical evaluation of parameter consistency and predictive uncertainty in hydrological modeling: A case study using Bayesian total error analysis." *Water Resources Research*, 45, -.

Tippett, M. K., Anderson, J. L., Bishop, C. H., Hamill, T. M., and Whitaker, J. S. (2003). "Ensemble square root filters." *Monthly Weather Review*, 131(7), 1485-1490.

Vaes, G., and Berlamont, J. (1999). "Emission predictions with a multi-linear reservoir model." *Water Science and Technology*, 39(2), 9-16.

van Delft, G., El Serafy, G. Y., and Heemink, A. W. (2009). "The ensemble particle filter (EnPF) in rainfall-runoff models." *Stochastic Environmental Research and Risk Assessment*, 23(8), 1203-1211.

Van Veldhuizen, H. M., Van Loosdrecht, M. C. M., and Heijnen, J. J. (1999). "Modelling biological phosphorus and nitrogen removal in a full scale activated sludge process." *Water Research*, 33(16), 3459-3468.

Vanrolleghem, P. A., Benedetti, L., and Meirlaen, J. (2005). "Modelling and real-time control of the integrated urban wastewater system." *Environmental Modelling & Software*, 20(4), 427-442.

Vieux, B. E., and Vieux, J. E. (2005). "Statistical evaluation of a radar rainfall system for sewer system management." *Atmospheric Research*, 77(1-4), 322-336.

Vrugt, J. A., & Bouten, W. (2002). Validity of first-order approximations to describe parameter uncertainty in soil hydrologic models. *Soil Science Society of America Journal*, *66*(6), 1740-1751.

Vrugt, J. A., Gupta, H. V., Bouten, W., and Sorooshian, S. (2003). "A Shuffled Complex Evolution Metropolis algorithm for optimization and uncertainty assessment of hydrologic model parameters." *Water Resources Research*, 39(8), -.

Vrugt, J. A., ter Braak, C. J. F., Diks, C. G. H., Robinson, B. A., Hyman, J. M., and Higdon, D. (2009a). "Accelerating Markov Chain Monte Carlo Simulation by Differential Evolution with Self-Adaptive Randomized Subspace Sampling." *International Journal of Nonlinear Sciences and Numerical Simulation*, 10(3), 273-290.

Vrugt, J. A., ter Braak, C. J. F., Gupta, H. V., and Robinson, B. A. (2009b). "Equifinality of formal (DREAM) and informal (GLUE) Bayesian approaches in hydrologic modeling?" *Stochastic Environmental Research and Risk Assessment*, 23(7), 1011-1026.

Willems, P. (1999). "Stochastic generation of spatial rainfall for urban drainage areas." *Water Science and Technology*, 39(9), 23-30.

Wong, L. T., and Mui, K. W. (2007). "Modeling water consumption and flow rates for flushing water systems in high-rise residential buildings in Hong Kong." *Building and Environment*, 42(5), 2024-2034.

Yatheendradas, S., Wagener, T., Gupta, H., Unkrich, C., Goodrich, D., Schaffner, M., and Stewart, A. (2008). "Understanding uncertainty in distributed flash flood forecasting for semiarid regions." *Water Resources Research*, 44(5), -.

Zappa, M., Jaun, S., Germann, U., Walser, A., and Fundel, F. (2011). "Superposition of three sources of uncertainties in operational flood forecasting chains." *Atmospheric Research*, 100(2-3), 246-262.