



Real-time Modelling and Data Assimilation Techniques for Improving the Accuracy of Model Predictions:

Scientific report



© 2010 PREED

The European Commission is funding the Collaborative project 'PREPARED Enabling Change' (PREPARED, project number 244232) within the context of the Seventh Framework Programme 'Environment'. All rights reserved. No part of this book may be reproduced, stored in a database or retrieval system, or published, in any form or in any way, electronically, mechanically, by print, photoprint, microfilm or any other means without prior written permission from the publisher

COLOPHON

Title

Real-time modelling and Data Assimilation techniques for improving the accuracy of model predictions: Scientific report

Report number PREPARED 2011.010

Deliverable number

D3.6.2

Author(s)

Hutton, C.J., Vamvakeridou-Lyroudia, L.S., Kapelan, Z., and Savic, D.A. (UNEXE)

Quality Assurance

By Jean-Luc Bertrand-Krajewski (INSA)

Document history

Version	Team member	Status	Date update	Comments
1	Hutton, C.J.	Draft	1/4/2011	
1.1	Kapelan, Z.	Draft	14/6/2011	Review, Minor Corrections
1.2	Vamvakeridou-Lyroudia, L.S.	Final	30/8/2011	Review, Minor corrections

This report is:

PU = Public

Summary

The Scientific report of Data Assimilation in Real-time modelling fulfils the requirements of Deliverable 3.6.2 within work package 3.6 of the PREPARED Enabling change project (EC Seventh Framework Programme Theme 6). This report has evaluated existing methods to assimilate data and correct predictive errors to improve the application of numerical models in real-time. Though the optimisation of Urban Water Systems through real-time modelling has received increasing interest in recent years, such modelling approaches often do not consider multiple sources of system uncertainty that affect our ability to identify optimal operational solutions.

Data Assimilation (DA) approaches have been applied and developed most widely in related scientific disciplines for updating model predictions in real-time as new measurements become available. Kalman Filtering, notably Ensemble Kalman Filtering, and Particle Filtering are promising approaches for propagating system uncertainty in real-time. Such methods provide the potential to account for uncertainty in model structure and uncertainty associated with input forecasts. Similar to many of the methods reviewed in Hutton et al. (2011) for quantifying model uncertainty, such methods are potentially demanding computationally for Data Assimilation.

Error-correction methodologies are relatively simple to implement and provide the ability to extend beyond DA approaches by reducing forecast error where observational data are unavailable. Such methods can implicitly account for a range of uncertainties provided these uncertainties are manifest in the deterministic model residual time-series derived off-line prior to application. Error-correction has been applied to update system states; however, the reviewed methods mostly provide deterministic corrections to output time-series, despite the methods themselves containing uncertainties.

Joint state and parameter estimation approaches have also been developed where DA filters have been applied within calibration frameworks. A hierarchical approach for dealing with model uncertainty combining model calibration, data assimilation and error-correction applied at different temporal scales, blending different representations of uncertainty, may provide an optimal framework to account for uncertainty in real-time modelling and control in Urban Water Systems. Although the methods presented here, as well as the techniques and methodologies that will be implemented in Task 3.6.2 can be considered as generic, the final selection of the methodologies to be applied depends also on the specific requirements of the PREPARED cities selected for demonstration.

Contents

	Summary	4
	Contents	5
1	Introduction	7
1.1	Introduction to PREPARED	7
1.2	Report Structure	9
2	Uncertainty In Urban Water Systems Models	11
2.1	Introduction	11
2.2	Types of Uncertainty	11
2.3	Sources of Uncertainty in Water Distribution Network modelling	12
2.4	Sources of Uncertainty in Urban Waste Water Systems Modelling	15
2.5	Summary	19
3	Real-time Modelling in Urban Water Systems	20
3.1	Introduction	20
3.2	Drivers of Real-Time Modelling	21
3.3	Real-Time Control	22
3.4	Real-Time Data Acquisition and Processing	25
3.5	Real-Time Model Forecasting	27
3.5.1	Rainfall Forecasting	28
3.5.2	Water Demand Forecasting	31
3.5.3	Real-Time Models	32
3.6	Summary	35
4	Data Assimilation and Real-Time Error Reduction Methods	36
4.1	Introduction	36
4.2	Kalman Filtering	38
4.3	Sequential Monte Carlo Sampling (Particle Filtering)	41
4.4	Variational Data Assimilation	43
4.5	Joint State and Parameter Estimation	44
4.6	Forecasting Error-Correction	45
4.7	Summary	47
5	Conclusion	49
6	References	51

7	Appendix A: Tabular Classification of Uncertainty Methodologies	71
8	Appendix B: Glossary of Terms	75

1 Introduction

This report fulfils the requirements of Deliverable 3.6.2 within work package 3.6 of the PREPARED Enabling change project (EC Seventh Framework Programme Theme 6). The report discusses the application of modelling tools in real-time, and evaluates data assimilation techniques for improving the accuracy of model predictions. Data assimilation techniques are much less developed in UWS modelling than methods for uncertainty quantification. The report will therefore review and evaluate a number of data assimilation techniques developed and applied more widely in related scientific disciplines, and assess their efficiency in improving model accuracy.

1.1 Introduction to PREPARED

Projected climatic change over the 21st century is predicted to manifest itself regionally through changes in water availability; Northern Europe and Southern Europe are projected to experience, respectively, an increase and decrease in mean precipitation, as well as an increase in the magnitude and frequency of extreme events (e.g. extreme precipitation events for Northern Europe and drought conditions in Central and Southern Europe; Christensen et al. 2007). Through impacts on the availability and quality of water in the water cycle (Figure 1), such changes will have direct consequences for what the World Health Organisation (WHO) considers the foundation of public health and development: the provision of drinking water and sanitation (WHO 2009). In Urban Environments drinking water is provided by the Water Distribution System (WDS) to consumers and industry, and sanitation chiefly provided for by the sewer network (Figure 1.). Adaptive strategies are required to reduce the vulnerability of UWS to climatic variability and change.

The aim of PREPARED is to show that the water supply and sanitation systems of cities and their catchments can adapt and be resilient to the challenges of climate change. In order to respond to the risks posed by climatic change, the impacts of which are currently surrounded by uncertainty, adaptive strategies are required that move beyond the current approach of building larger infrastructure that cannot be relied upon to deliver acceptable risk over the medium timescale. Strategies are required to better manage potential risk. Strategies that can be optimised as new information becomes available to avoid two potential scenarios: First, the potential for under-investment as climate change impacts are underestimated; Second, the potential for over-investment, and an unnecessary use of resources.

PREPARED, which has taken an industry/end-user driven approach will seek to build the resilience of UWS, initially in a number of demonstration cities, in two primary ways:

1. Through optimisation of existing water supply and sanitation systems, to postpone investments in new infrastructure until investment risks are lower as more knowledge is available.
2. Second, in the case where optimisation is not sufficient, PREPARED will provide guidance and produce frameworks to aid utilities in building more resilient water supply and sanitation systems.

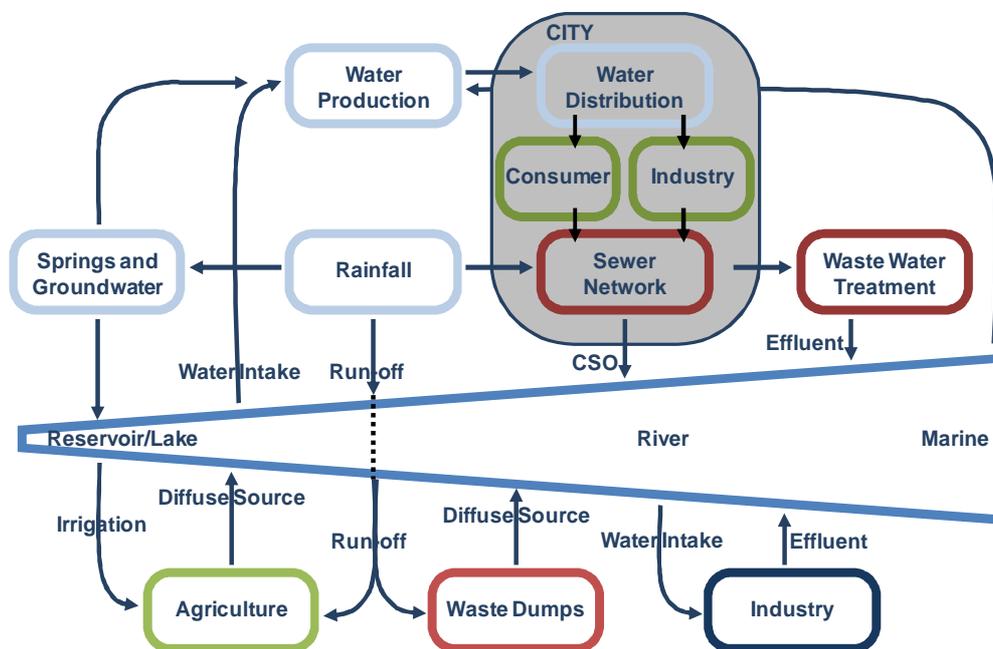


Figure 1. Position of the Urban Water System (Grey Shaded Region) within the water Cycle.

Developing approaches for optimal management of UWS requires a detailed understanding of how such systems operate. Conventional management approaches have typically focussed on solving isolated technical problems, in what has been termed a “command and control” approach (Pahl-Wostl et al. 2007). Dealing with problems in such a way neglects system complexity and the potential for complex system feedbacks that may result in unexpected consequences (Pahl-Wostl 2007). Such management therefore represents poor management of risk and resources. Although subjectively defined, risk may be generally considered as the consequence combined with the probability of

occurrence of a particular event. The identification and potential reduction of risk associated with Urban Water System management (PREPARED work package 2.3) requires a deeper, holistic understanding of inherent system complexity and uncertainty to better inform an understanding of the probability of event occurrence.

An essential and innovative aspect of PREPARED is the development of a toolbox for real time monitoring and modelling (Work area 3.6). The toolbox is required to increase the technological capacity of existing water supply and sanitation systems to deal with changes in the quality and quantity of system input resulting from climatic change, alongside potential changes in demand. Such demands call for an integrated real time control strategy, supported by monitoring and modelling approaches, to provide decision support in the face of inherent system uncertainty. Towards this end, Work package 3.6 will investigate methodologies for uncertainty quantification in UWS modelling, and identify possible steps to reduce uncertainty through real-time modelling, calibration and data assimilation.

1.2 Report Structure

Section 2 defines types of uncertainty that affect modelling from a systems perspective. In order to understand the potential application of different methods for quantifying and reducing uncertainty in UWS, Section 2 first reviews the types of uncertainty affecting our understanding, and therefore ability to model UWS.

Section 3 provides some context to understand the potential application of Data Assimilation methods in UWS, and reviews the aims and drivers real-time modelling, and issues of model forecasting.

Section 4 reviews Data Assimilation methods and Error-Correction methods developed and applied specifically to deal with quantifying and reducing uncertainty in (near) real-time. The methods considered in Section 4 are those considered most applicable for dealing with uncertainty in Real-time WDS modelling.

Section 5 provides a summary of the conclusions of the report.

Section 6 References.

Section 7 Appendix A. Provides a tabular classification of uncertainty methodologies that may be applied in UWS modelling, extending the Table presented in Hutton et al. (2011).

Section 8 Appendix B. Provides a Glossary of terms to facilitate understanding of the issues and methods presented in the report, extending the glossary presented in Hutton et al. (2011).

2 Uncertainty In Urban Water Systems Models

2.1 Introduction

Uncertainty may be defined as a state where we have incomplete knowledge of a system. Uncertainty is typically divided into **aleatory uncertainty**, which refers to irreducible natural variability in measurements (e.g. rainfall; Hall 2003; Helton and Burmaster 1996); and **epistemic uncertainty**, which results from incomplete system knowledge, often pertaining to numerical system models.

In order to address uncertainty when modelling UWS in real-time, three areas need to be considered: understanding, quantification, and reduction of uncertainty (Liu and Gupta 2007). The report that fulfils the requirements of REPAED deliverable 3.6.1, entitled: *Uncertainty Quantification and Reduction in Urban Water Systems Modelling: Evaluation Report* (Hutton et al. 2011), provides a thorough review of existing uncertainties affecting our ability to accurately simulate UWS. This section provides an abridged version of Section 2 of Hutton et al. 2011 to provide context, and aid in understanding the potential application of the real-time approaches reviewed in this report.

2.2 Types of Uncertainty

From a systems theory perspective, a model may be considered as composed of six different components (Figure 2):

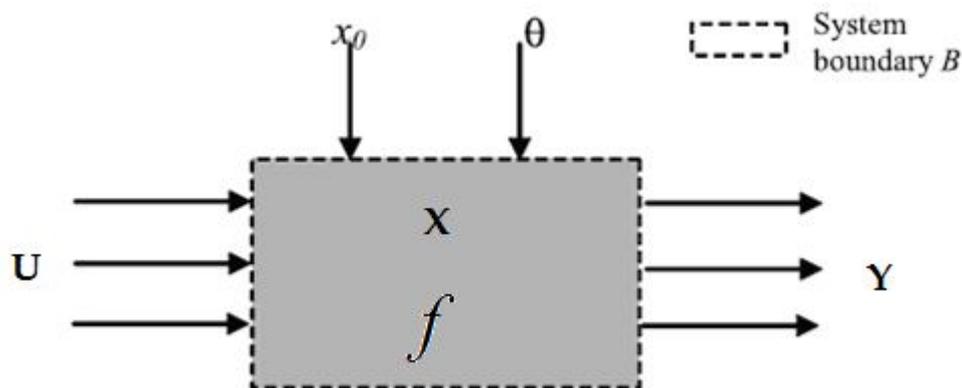


Figure 2. A schematic systems representation of model components (modified from Liu and Gupta 2007).

where B is the system boundary; $\mathbf{U}=(u_1,\dots,u_n)$ and $\mathbf{Y}=(y_1,\dots,y_n)$ represent model inputs and outputs, with length n , as fluxes of mass or energy into and out of

the system; x_0 represents the model initial conditions; $\theta = (\theta_1, \dots, \theta_m)$ are model parameters (e.g. pipe roughness) with length m , which are typically considered time invariant during simulations, but in real-time applications that seek to reduce uncertainty, they may be time varying (Moradkhani et al. 2005b); $\mathbf{X} = (x_1, \dots, x_n)$ represents model system states (e.g. pressure or head in a WDS model), which are stored in the system boundary, and alongside \mathbf{Y} , evolve over time when the model system equations (f) are conditioned on model parameters and inputs:

$$\mathbf{Y}, \mathbf{X} = f(\mathbf{U}, x_0, \theta, B) \quad (2.1)$$

The model equations (f) may be considered as a formalised mathematical representation of reality that seek to make the correct mapping from system inputs to states and system outputs. In general, there are three different types of model uncertainty, which incorporate the model system components described above: **Structural uncertainty** (B, f) refers to errors in the mathematical representation of reality, and is a form of epistemic uncertainty; **Parameter uncertainty** reflects uncertainty in the value of variables used in equations to represent model system components (e.g. pipe roughness; θ); **Measurement/data uncertainty** refers to uncertainty in the quantities used to define initial conditions (x_0), model inputs (\mathbf{U}) and observations used to evaluate model predictions (either system states (\mathbf{X}) or outputs (\mathbf{Y})). A first step towards reducing the effects of these sources of uncertainty on predictions of system states (\mathbf{X}) and outputs (\mathbf{Y}) is to first understand sources of uncertainty in UWS, and uncertainties in the models typically used to represent them.

2.3 Sources of Uncertainty in Water Distribution Network modelling

The primary objective of the Water Distribution Network is to provide clean, potable water at sufficient pressure and volume for end users (domestic and industrial). To meet this demand a WDS typically consists of a number of links (pipes, pumps and valves) that are joined at junction nodes, and control distribution of drinking water, via storage tanks, from a water production to the consumer. Under normal design (steady state conditions), the network must be capable of supplying anticipated demands with adequate pressures. Network models that seek to represent the WDS consist of a collection of pipes, pumps and valves, which are connected together at a series of nodes, where consumer demand is specified. The detail with which the original WDS is represented in both time and space depends on the purpose to which the model is to be used. For a given demand (pattern) the system equations

conserving mass at junction nodes and energy along pipes may be solved for steady state and extended period simulation (e.g. EPANET2; Rossman 2000). In pressure deficient situations resulting from leakage or fire flow, pressure driven modelling approaches may be applied where the assumption that demand is met at each node is inadequate (Giustolisi et al. 2008). In the case where the transition between hydraulic conditions is important the governing equations of mass and momentum need to be solved to simulate pressure wave propagation (Boulos et al. 2004; Jung et al. 2007). A number of uncertainties affect our ability to model UWS:

- **Skeletonisation** involves the removal of pipes that for computational constraints and reasons of parameterisation are not considered essential to model system performance. For steady-state simulations skeletonisation may not have large impacts on model performance; however for surge analysis such simplification may inadequately represent maximum surge head by neglecting dead ends and high elevation nodes (Jung et al. 2007). Skeletonisation by trimming also results in the need to re-allocate demand, which results in modifications to pipe velocities, the potential for inaccurate contaminant consequence assessment (Bahadur et al. 2006), and may incorrectly assume sufficient pressure at every trimmed location in the network (Walski et al. 2003).
- **Demand** uncertainty consists of aleatory uncertainty surrounding the natural variability of consumer demand, which varies over a range of different timescales (Buchberger and Wells 1996; Davidson and Bouchart 2006; Herrera et al. 2010) reflecting work, commercial and domestic usage throughout the day and week, and changes in response to seasonal and climatic changes over the year. There is significant epistemic uncertainty surrounding understanding of this inherent system variability, which is most often constrained through measurements of system state (e.g. Davidson and Bouchart 2006), or by predictive models (Cutore et al. 2008; Herrera et al. 2010), which are themselves uncertain. Demand is typically expressed at network nodes; however, consumers extract water along the pipes within the network. Although this error may be small relative to uncertainty surrounding the actual amount of demand, it may lead to errors in the prediction of system head loss (Giustolisi and Todini 2009).
- **Pipes** form an integral part of the WDS, and their **Roughnesses**, alongside demand, are one of the most significant sources of uncertainty in WDS modelling. Pipe roughness changes over time due to pipe deterioration and the deposition of material (Boulos et al.

2004), which reduced pipe diameter. Pipe deterioration depends on a range of factors, including material and water quality, and thus roughness is difficult to predict with increasing pipe age, not least due to the difficulty of measurement of what is an effective parameter with limited physical meaning. Pipe roughness values, like nodal demand patterns are often grouped to reduce the dimensions of the calibration problem (Mallick et al. 2002). However, as the number of parameters reduces, so does model accuracy in representing true variability that controls the functioning of the WDS.

- **Pumps, valves and tanks** are key system components allowing managers to control the movement of water in the distribution network. In practice pumps do not typically operate at the curve efficiency supplied by the manufacturer (Walski et al. 2003), and over time performance will deteriorate due to cavitation and wear (Hirschi et al. 1998). Valves control the flow of water through the WDS, and operate in different ways depending on their purpose. The effect of some valves may be adequately represented by a minor loss coefficient, and potentially incorporated into a pipe roughness coefficient, emphasising the effectiveness of pipe roughness values. Tanks store water in the distribution network, and are characterised by a rating curve between head and storage volume.
- **Water Quality** predictions are described by the advection-dispersion-reaction equation (Blokker et al. 2008). Given water quality is dominated by advective transport (Pasha and Lansey 2005), the dispersion terms are neglected in EPANET2 (Rossman 2000). Whilst this is a reasonable assumption for turbulent flows, dispersion is important in laminar flows (Blokker et al. 2008). Further, the perfect mixing assumption applied at network junctions in EPANET2 is inaccurate (Austin et al. 2008; Romero-Gomez et al. 2008), and may lead to erroneous predictions of pollutant concentration. Underlying hydraulic model uncertainty, resulting from demand uncertainty and skeletonisation, affects the velocity predictions required for predicting the fate of contaminants, chlorine decay rates (Menaia et al. 2003), and for quantifying the population actually affected by a given contamination event may be incorrect (Bahadur et al. 2006). Relatively little attention has been given to joint calibration of WDS and water quality models (Savic et al. 2009). Pipe wall chlorine decay depends on the pipe age and material, which as for roughness is difficult to quantify for all network pipes, and results in pipe grouping (Munavalli and Kumar 2005).

2.4 Sources of Uncertainty in Urban Waste Water Systems Modelling

UWWS consist of three principal components: Sewer System, Wastewater treatment plant, and receiving water body (Figure 1), and have been designed to complement the WDS by mitigating flooding, and providing good sanitation (Korving et al. 2003). Many major cities around the world have combined sanitary and storm-water flows. During rainfall events the WWTW has to deal with a larger volume of relatively dilute wastewater, increasing processing costs, and the potential for Combined Sewerage Overflow (CSO) discharges, with potentially detrimental impacts on water quality (Casadio et al. 2008).

Traditionally, each component of the UWWS was managed separately, often by a different company, with different management aims (Devesa et al. 2009). A number of approaches moving towards integrated modelling of UWWS have been developed both in the research literature (Butler and Schutze 2005; Vanrolleghem et al. 2005), and commercially (e.g. WEST and SIMBA; Rauch et al. 2002), to address the concerns for the vulnerability of water quality (Beck 2005), as exemplified by the introduction of the Water Framework Directive (WFD; Bloch 1999); and second to meet public expectation in attaining higher levels of service (Pahl-Wostl 2005). Such models are required to facilitate incremental adaptation (Butler and Parkinson 1997), and help optimise the performance of existing UWWS, by explicitly accounting for interactions between different components of the system (Butler and Schutze 2005). However, the integrated UWWS is complex, involving a number of epistemic and aleatory uncertainties (Benedetti et al. 2008; Korving et al. 2003):

- **Rainfall** represents the key input to UWWS during Wet Weather Flow (WWF), and during storm events may cause sewers to exceed their hydraulic capacity, resulting in surcharge and CSO discharges. 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). Aleatory rainfall uncertainty relates to natural temporal variability in rainfall over annual timescales reflecting seasonal variations and climatic circulation patterns (Rodriguez-Puebla et al. 1998); over daily timescales due to convective processes in the atmosphere (Kutiel and Sharon 1980; Kutiel and Sharon 1981); and over storm event timescales relating to the movement of clouds/rain cells (Morin et al. 2006). Rainfall also varies spatially over large scales relating to climatic patterns (e.g. over the Iberian peninsula; Rodriguez-Puebla et al. 1998) and continental topography (Jang 2010); over sub-catchment scales in

response to local topographic forcing (Chaubey et al. 1999) and wind shelter (Sevruk and Nevenic 1998); and over short distances (10²m) at event timescales in response to the spatial structure of convective rainfall cells (Faures et al. 1995). Spatial patterns in rainfall may also be controlled by the presence of the urban area itself (Jauregui and Romales 1996; Rosenfeld 2000; Thielen and Gadian 1997).

Epistemic uncertainties in rainfall measurements result from measurement errors and errors in the spatial and temporal resolution of the phenomena. Point rain gauge measurements are subject to both systematic errors (Ciach 2003; Rauch et al. 1998; Sevruk 1996; Sevruk et al. 1994; Sevruk and Nespor 1998; Stransky et al. 2007), and random errors (Rauch et al. 1998). Such measurements may not have a sufficient temporal resolution for accurate modelling (Aronica et al. 2005), and require interpolation (Goovaerts 2000; Willems and Berlamont 1998). Rainfall radar has been used increasingly alongside point rainfall measurements (Vieux and Vieux 2005); however, the algorithm used to convert a radar signal to rainfall intensity often requires bias correction due to uncertain parameters (Vieux and Vieux 2005), and runoff predictions may be sensitive to the resolution of radar measurements (Ogden and Julien 1994). Point gauge measurements are typically used for bias correction (Campolongo et al. 2007), which as discussed above are themselves uncertain.

Many urban sewer systems are situated in wider hydrological catchments; there is considerable uncertainty surrounding the simulation of the rainfall-runoff process (Wagener et al. 2003), and even more uncertainty concerning the transport of sediment and pollutants during runoff, both from agriculture (Beven et al. 2005) and urban environments (Deletic et al. 2000). This is a particular problem for understanding the potential impacts of CSOs during wet periods, as the state of the river will be independently altered by rainfall-runoff.

- **Dry Weather Flow** (DWF) consists of flow outputs from domestic and industrial users into the UWWs (Figure 1). Similar to water consumption (demand) in the WDS, uncertainty in DWF is both aleatory, reflecting changing consumer inputs over different timescales. Domestic wastewater may be made up of contributions from a variety of different household appliances (e.g. WC, Shower), each with their own patterns of use that vary between weekday and weekend (Butler 1993; Friedler et al. 1996), and diurnally (Almeida et al. 1999). Aleatory uncertainty also results from different usage amongst different users (Wong and Mui 2007).

There is significant epistemic uncertainty in the nature of DWF from domestic properties, owing to the difficulty of measuring discharge volume and content 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).

- **Sewer Systems**, though initially designed to quickly removing storm water (Delleur 2003), cannot simply be seen as inert conveyors of material. Sewer processes are complex, with the following key components (Ashley et al. 1999): hydraulics, sediment transport, advection-dispersion and biochemical water quality processes (Figure 10). Whilst the modelling of most of these components is well developed (e.g. St. Venant equations for sewer hydraulics) there are a number of significant problems in deriving empirical information, which constrains our ability to apply complex models (Ashley et al. 1999): First, there are logistical difficulties of actually measuring certain processes within the sewer system; Second, even when such processes can be measured, economic or logistical issues prevent extensive distributed measurements; Third, extreme spatial and temporal variability in sewer systems poses difficulties for constraining parameter and system state uncertainty in distributed sewer models (Jack et al. 1996).

Structural uncertainty, however, does exist in sewer system models, first, because of a lack of understanding of a number of processes. There is uncertainty regarding the nature of sediments in transport near the bed (De Sutter et al. 2003), and many existing models do not represent sufficient size fractions for sediment transport prediction, and cohesive sediment transport and deposition (Ashley et al. 1999). Second, model simplifications are necessary due to system complexity and computational resources (Fischer et al. 2009), leading to known structural uncertainty. For example, simplifications of the fully dynamic 1D St Venant equations have been applied to sewer systems, in addition to conceptual store models (Vaes and Berlamont 1999).

- **Waste Water Treatment Works (WWTW)** represent a number of different components for deriving clean water, including a clarifier, an active sludge model, hydraulic model, oxygen transfer model, and sedimentation tank model. The WWTW is subject aleatory input uncertainties associated with dry weather flow and rainfall input, as well as potential modification of flow volume and quality in the sewer

system due to sewer residence times and within sewer processes (Nielsen et al. 1992; Van Veldhuizen et al. 1999).

There are difficulties in applying complex WWTW models (Gernaey et al. 2004) due to parameter demands that are often substantial and difficult to constrain (Sin et al. 2009). 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). 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). Only in particular cases are hydraulic models applied explicitly to simulate flow through reactors (De Clercq et al. 1999), which are typically assumed instantaneous (Rauch et al. 2002). Further details of structural uncertainties in the WWTW may be found elsewhere (e.g. Gernaey et al. 2004; Rauch et al. 1999).

- **Rivers** are the primary receiving water bodies for many UWWs, and are vulnerable to oxygen depletion and eutrophication owing to sewer and WWTW effluent nutrient loads (Harremoes and Rauch 1999). Rivers have the same general input and structural uncertainties as described for sewer systems, in addition to uncertain water volume (and quality) derived from non-urban sources (e.g. agricultural; Bilotta and Brazier 2008; Bilotta et al. 2008). Issues surrounding epistemic uncertainties in river models are similar to those in sewer system models (see also: Reichert et al. 2001; Reichert and Vanrolleghem 2001); model complexity may be increased 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.

Despite river water quality being one of the key policy drivers to evaluate UWWs performance, data relating to the relationship between water quality and river properties, such as ecology, is often lacking, because knowledge of such processes is uncertain (Bilotta and Brazier 2008; Borchardt and Statzner 1990). For example, different organisms respond differently to certain flow dynamics/exposures, and may have different recovery times. The determination of ecologically meaningful hydrological parameters and thresholds is

difficult owing to nonlinear dynamics and multiple causes (Groffman et al. 2006), and limited to specific case studies (Borchardt and Statzner 1990). Furthermore, traditional measures of pollution impact (emission standards), such as the frequency or volume of CSO spill (Lau et al. 2002), may not be compatible with measures of stream water quality standard (Freni et al. 2010), nor reflect actual pollution (Lau et al. 2002).

2.5 Summary

Models of both WDS networks and UWS are subject to structural, parameter and data uncertainties. Natural variability, particularly in demand and dry weather flow, represent consumer driven uncertainties, whilst UWS also have to deal with natural rainfall variability. There are further difficulties in measuring such phenomena accurately in both time and space. Models applied to UWS have known structural uncertainties that often must be tolerated for computational reasons and data constraints. Further, the difficulty of obtaining distributed measurements results in uncertain and often poorly constrained model parameters. Formal methods are required to deal with this uncertainty in real-time approaches that may be employed to optimally manage UWS.

3 Real-time Modelling in Urban Water Systems

3.1 Introduction

Real-time refers to a state where data referring to a system is analysed and updated at the rate at which it is received (i.e. at the rate at which the system operates). Real-time modelling (also referred to as online modelling) refers to the process of employing numerical models to make predictions about current or near future system states (**X**) and outputs (**Y**) based on newly received (and forecasted) data. Real-time modelling is employed in a range of environmental fields, including meteorology (Golding 2000; Thorndahl et al. 2010), hydrology (Cloke and Pappenberger 2009; Collier 2007), and in UWS (Fu et al. 2008; Shang et al. 2006), typically for one of the following purposes:

- To provide warnings of future events (e.g. flash flooding) such that evacuation and mitigation can take place (Penning-RowSELL et al. 2000).
- To inform management of future system states and potential anomalies, such that control intervention of system states and outputs can take place (Ingeduld and Turton 2002; Shang et al. 2008).
- To explore a range of possible control strategies such that the optimum control solution that minimises some function (typically a system property such as CSO discharge, or operational cost) is implemented (Darsono and Labadie 2007; Rao and Salomons 2007).

Real-time modelling may also be referred to as online modelling (Machell et al. 2009) as data capture and processing is directly coupled with model application. This approach is in contrast to many of the calibration approaches considered in Hutton et al. (2011), in which model application may be considered offline, in that the models are applied using existing time-series, typically for model parameter calibration prior to application. In this sense offline modelling may be seen as a complementary step prior to the online model methodologies considered in this report. Section 3 of this report will provide some context for the application of real-time modelling in UWS. First the section will consider the drivers of real-time modelling (3.2) and the issues surrounding real-time control in UWS (3.3). Section 3 will also discuss data collection in real-time (3.4), and forecasting and real-time modelling issues (3.5) that provide the context for the potential application of the Data Assimilation and Error-correction techniques considered in Section 4.

3.2 Drivers of Real-Time Modelling

The application of real-time modelling in UWS for water management has received growing interest in the research literature in a range of European countries including the UK (Machell et al. 2009), Denmark (Harremoes and Rauch 1999) Germany (Schroeder and Pawlowsky-Reusing 2005), France (Stinson et al. 2000), Belgium (Vanrolleghem et al. 2005), and Spain (Ocampo Martinez 2007), and also in the U.S.A. (Darsono and Labadie 2007) and Canada (Pleau et al. 2005). The increase in application of real time modelling has been driven by a need for greater efficiency of UWS system operation to meet a number of system demands.

Greater management of UWS is required to meet the demands imposed by policy. In the context of UWWS the European Water Framework Directive (Bloch 1999) requires greater integration of catchment management, and improved qualitative and quantitative water quality by 2015. Meeting these requirements means a reduction in the quantity and improved quality of water derived from the sewer network. Similar Guidelines and requirements have also been implemented in the United States by the Environmental Protection Agency, including the nine minimum controls to have been implemented in 1997 (USEPA 1995a), as well as longer term CSO control plans (USEPA 1995b), which include modelling of the combined sewer system.

Both UWWS and WDS also have to meet demands of regulatory authorities, such as Ofwat in the UK, public expectation of a high quality service (Pahl-Wostl 2005), and when the management companies of such systems are privatised, shareholder satisfaction (Ogden and Watson 1999). Such demands are often competing, as in general improved water quality, both that derived for potable consumption and that leaving the sewer system, comes at a cost. Further, the driving factors considered here are closely related; for example, adherence to the WFD will come at a cost to the water industry (Ofwat 2005), a cost ultimately borne by consumers. Other drivers include greater concern of the potential for deliberate contamination events, which has increased research into real-time monitoring and modelling of water quality in WDS (Davidson et al. 2005; Panguluri et al. 2005). Effective and economic monitoring and optimisation of system performance is required, which is currently delivered by Real-Time Control (RTC).

3.3 Real-Time Control

Real-Time Control (RTC) refers to the process by which system states are monitored in real time, and regulatory devices are operated in response to measurements to control the system and its states (Figure 3; Schutze et al. 2004). The main objective of RTC is to maximise the use of regulatory devices (actuators), such as pumps, gates, valves, and treatment plants to meet customer and regulatory demands at minimal costs.

In the context of WDS, RTC control is required in WDS to reduce pumping costs (e.g. by filling tanks in low tariff periods) whilst maintain adequate system pressure to meet fluctuating consumer demands (Davidson and Bouchart 2006). Higher system pressures than necessary are normally maintained by controllers due to current control limitations, which leads to higher leakage losses from the system (Jamieson et al. 2007). Further, pump-scheduling for system control typically takes the form of lapsed-time control in response to average demand curves over a 24 hour period (Rao et al. 2007), which does not take full advantage of on-line monitoring data.

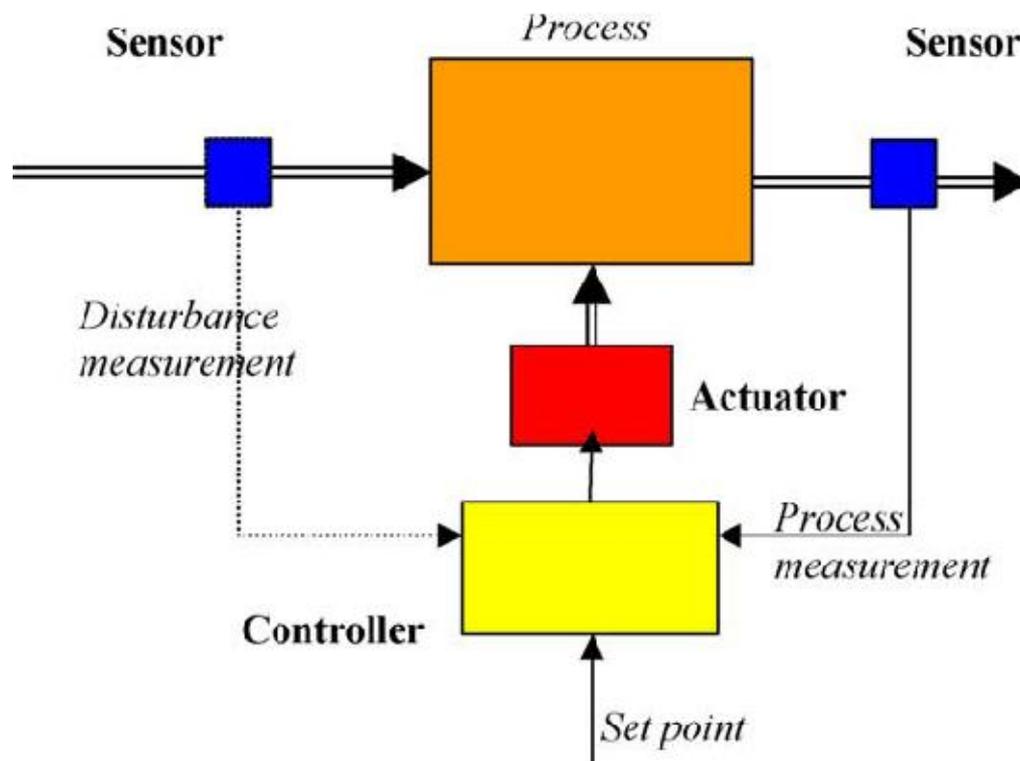


Figure 3. Schematic illustration of real-time system control. Feed-forward (disturbance measurement) and feedback (process measurement) control loop. Simple arrows indicate data flow, double arrows indicate hydrodynamic action. Bold letters indicate hardware and italic letters indicate variables (Schutze et al. 2004).

In UWWS the volume and quality of CSO discharge needs to be minimised by optimally using regulatory devices (e.g. gates, weirs, pumps and treatment works) to manage the flux of sewerage within the wastewater system, through for example, inline storage (Darsono and Labadie 2007). There are three basic approaches for RTC (Vanrolleghem et al. 2005): volume-based, pollution-based and emission based. Although volume-based approaches do not necessarily minimise pollution impact (Lau et al. 2002; Rauch and Harremoes 1999), a relative dearth of data means volume-based approaches are often the most practical approach.

Although state of the art control and monitoring equipment has been installed in some sewerage systems (e.g. Branisavljevic et al. 2010; Stoianov et al. 2006), data have often been considered underused (Weywand 2002), without demonstrating the full potential of on-line monitoring data such as that provided by Supervisory Control And Data Acquisition (SCADA) systems (Kang and Lansey 2009). Over recent years the cost and performance benefits of real time control have been demonstrated (Broks et al. 1995; Campisano et al. 2000; Entem et al. 1998), with implementation in some European Cities, including Vienna (Fuchs and Beeneken 2004), Barcelona (Ocampo Martinez 2007), and cities in Germany (Weywand 2002), and the development of guidelines and predictive tools to demonstrate the benefits of RTC (Messmer et al. 2008; Schutze et al. 2008). In recognition of the benefits, real-time control of WDS has also been realised in many cities worldwide (Panguluri et al. 2005; Patterson et al. 2005; Shin et al. 2009; Zhao et al. 2005).

It has been argued that in many current systems much of the RTC management is limited to local control (Ocampo Martinez 2007; Pleau et al. 2005; Schutze et al. 2004). There has been an increased recognition, particularly over the last decade, of the need for integrated system control in the research literature to meet the aforementioned demands for economic and efficient system performance of UWS (Butler and Schutze 2005). Integrated control in the UWWS requires integration of WWTW, sewer system and river data to optimise controls of, for example, inline sewer storage in response to wider system states and controls (Darsono and Labadie 2007). Finding optimal solutions for integrated systems control is a more complex, and a more computationally demanding problem than localised optimisation, because of the large number of variables that can potentially influence system state. However, at the same time, a larger number of control options, and the potential to observe system state upstream of a given location provides greater potential to optimise system performance. Such operation may therefore require a move beyond heuristic trial-and-error control strategies based on the experience of the operator.

In contrast to ‘natural’ systems such as hydrological catchments, UWS are essentially manmade with built in control structures for system optimisation. Therefore in addition to real-time modelling approaches that forecast future system states, models have become increasingly important to investigate the implications of different control strategies. **Model-based Predictive Control** (MPC; Fu et al. 2008; Rao and Salomons 2007; Rauch and Harremoes 1999) employs models to find the optimal control decision that best meets future demands on the system. Such modelling can take the form of offline modelling, where models are applied to identify optimum system control strategies based on integrated system models (Butler and Schutze 2005; Jamieson et al. 2007; Rauch et al. 2002). Such models may be applied to determine rules for rule-based system control (Fuchs and Beeneken 2005), or through exploration of a range of possible operational scenarios based on system optimisation with reference to cost and performance.

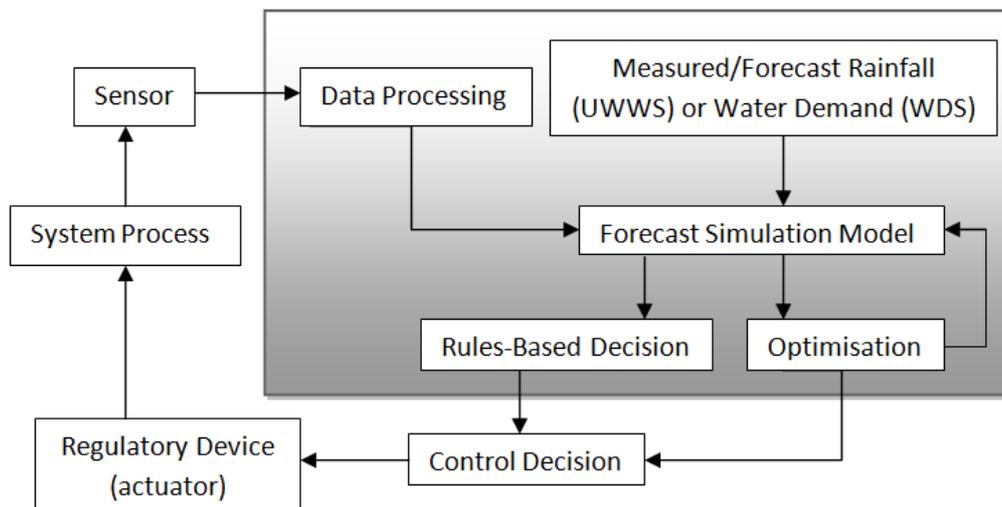


Figure 4. Flow diagram showing the location of simulation modelling in real-time system control. The large grey box contains the processes that must be conducted between receiving system data and modifying actuator settings. The forecast model may be applied iteratively in an optimisation procedure to find the best control scenario, or once to predict future system state and inform a rules-based decision strategy.

As an alternative to offline modelling, MPC models have also been applied online in real-time to forecast both future system states, based on current conditions and forecasted drivers (e.g. water demand or rainfall; Fuchs and Beeneken 2004; Rao and Salomons 2007), or to explore in real-time the implications of future operational scenarios, to seek the best management condition (Figure 4). The processes contained within the grey box in Figure 4

are the stages that must be conducted between receiving system information and producing system forecasts to inform control decisions. As identified in a number of studies, the time available for system modelling is constrained, which has led to the application of alternative models (Rao and Salomons 2007). Perhaps in part because of these temporal constraints, the application of Data Assimilation techniques for dealing with uncertainty in real-time modelling has been lacking. Of central concern for the application of modelling in real-time are the spatial extent, frequency and quality of data, and the need for efficient, automated methods for processing sensor data.

3.4 Real-Time Data Acquisition and Processing

Data acquisition in real time, from measuring the given phenomenon (e.g. discharge or a given water quality parameter), to providing data for modelling and control decisions involves a number of stages: 1) Sensing the given phenomena; 2) converting this measurement, typically by calibration, into the property of interest; 3) Transfer the data to a central receiver/storage device; 4) anomaly detection and data processing.

A wide variety of devices are available for measuring properties of the flow field. Many of these devices do not measure directly the property of the flow field of interest, and require calibration. Of wide application in Sewer systems and WDS is the measurement of the physical flow properties to quantify discharge volumes. Measuring free surface flow in sewer systems may be achieved by combining velocity measurements with flow depth measurements (Fulford and Gonzalez-Castro 2009). Errors exist in such measurements due to sediment deposition induced changes in cross-section area (Larrarte and Chanson 2008), and for example, uncertainties in the representativeness and settings of Doppler velocity measurements (Larrarte 2006; Larrarte et al. 2008).

A number of different sensors have also been applied to measure water quality parameters, including the application of UV/VIS/NIR spectrometers in sewers (Gruber et al. 2005; Stumwohrer et al. 2003), rivers (Barillon et al. 2010) and in WWTW (Rieger et al. 2008), and the application of turbidity probes (Joannis et al. 2008). Many sensors suffer reliability problems because of the corrosive and destructive nature of the monitoring environments, with regular cleaning often required to prevent clogging (Gruber et al. 2005).

Spectrometers, like other sensors, require calibration to relate the signal received to a water parameter of interest. Manufacturer and laboratory calibration prior to installation may be inadequate to represent the specific

conditions in the location of sensor deployment; calibration for local conditions such as temperature for turbidity probes may be required (Joannis et al. 2008). Further, in local application noise may dominate measurements (Maribas et al. 2008), and measuring devices may not adequately deal with shifts in the wastewater matrix resulting from different event timings and magnitudes (Stumwohrer et al. 2003). Also drift in sensor calibration may occur, which may result in the need for regular recalibration (Rieger et al. 2008). Methods are required to quantify such measurement uncertainties (Bertrand-Krajewski et al. 2003). Recent advances in online drinking water quality monitoring are reviewed in Storey et al. (2011).

Data transfer for real-time application has received more interest in recent years following technological advances in telemetry, and an increased interest in integrated system control. Such control requires communication of large amounts of data over potentially long distances, and synchronisation of such data for further application. A range of methods have been employed to communicate information from sensors, including telephone lines, radio, satellite, WAP wireless networks and the internet (Castro et al. 2008; Ruggaber et al. 2007; Stoianov et al. 2006). A key difficulty is the power required, often in remote and difficult to reach locations (e.g. in underground pipes), to communicate the volumes of collected data for real-time application. Recent advances in technology include local low energy data transmission from sensors to above ground data gatherers via Bluetooth, which then manage time synchronisation and control long range communication to servers where power resources are greater (Stoianov et al. 2006). Such technologies though not widely applied and often in stages of development, are often low cost. Such developments will facilitate wider application of real time data collection, and application of real-time modelling.

Given the potential problems in measuring data, validation and processing are essential steps to perform prior to utilising data derived from sensors for real-time modelling application, as there may be a large redundancy in data collected (Schilperoort et al. 2008). Procedures for removing zero values, filling in single gaps and double entries may be readily automated and executed on data time-series (Schilperoort et al. 2008), and are often required prior to application of more advanced techniques for data validation (Branisavljevic et al. 2010). Recent developments for automatic detection of anomalous data include application of Artificial Neural Networks (ANN), and context classification of data prior to detection (Branisavljevic et al. 2009). Romano et al. (2010) applied a wavelet approach to de-noise WDS time series data. Such methods are required as manual (visual) techniques, though potentially effective, are not suitable for real time application.

3.5 Real-Time Model Forecasting

The basis of applying models in real-time control is to make predictions of future events (forecasts) and to inform controllers of the best strategies for dealing with those future events. Although model forecasting is applied in a range of scientific fields, some generic issues govern the applicability of different methods for model forecasting (see Collier 2007; Pappenberger et al. 2005; Todini 2004). For a forecast to be useful it must be made before an appropriate time horizon (also termed operating horizon) for the system in question. This horizon is determined by the time required to initiate control actions to mitigate the impacts of a forecasted event. For example, if demand in a WDS is projected to increase, pump settings may be modified to fill storage tanks to meet this future demand, and if the warning is sufficiently early, this pumping may occur in low tariff periods (Salomons et al. 2007). Larger storage tanks provide some leeway in meeting future demands, which would reduce the necessary time horizon required to take action. Therefore the time horizon is specific to the control system in question.

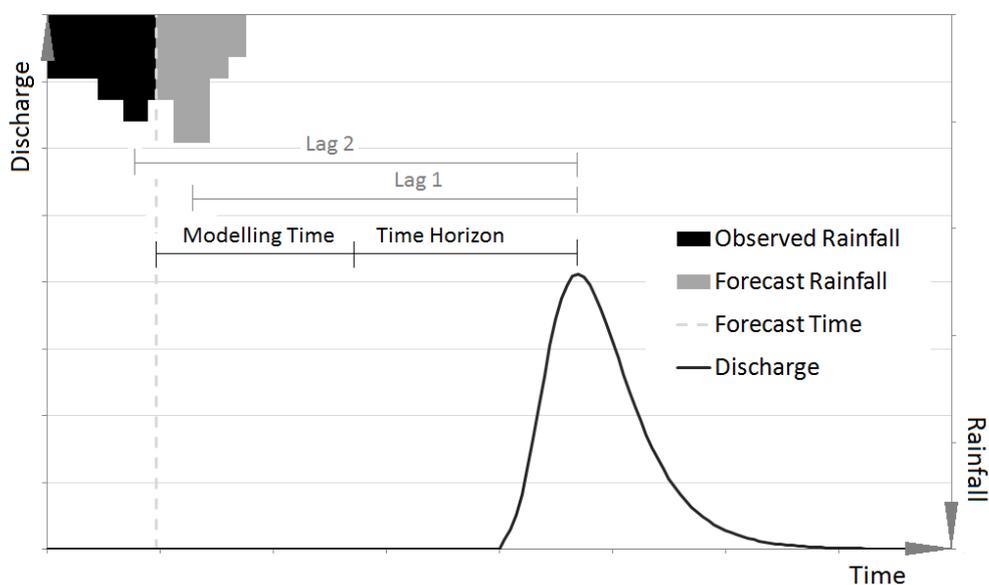


Figure 5. Schematic illustration of model forecasting. The forecast time is the time at which system modelling starts to predict future events or optimise the control strategy to be implemented before the time horizon. If the natural lag in the system is longer than the modelling time plus time horizon (lag 2), then system observations can be used to drive the modelling. If the lag time is shorter than the modelling time plus time horizon (lag 1), then input forecasts (e.g. rainfall) are required to initiate the modelling optimisation in enough time prior to the time horizon.

The system forecast is made based on the input drivers of the system. In the case of UWWS the key driver of wet weather flow is rainfall, and for WDS, water demand. The system lag time (also termed Concentration Time) refers to the characteristic time for a response to an input at a given location in the system. If the lag time is greater than the time horizon plus the time required to generate the system forecast (Figure 5: Lag 2), then predictions based on existing measurements of input conditions up to the time of prediction can be made. However, if the lag time is shorter than the time horizon and the modelling time required to optimise the system (Figure 5: Lag1), in order to make timely predictions (i.e. before the time horizon) forecasts of input drivers, such as rainfall, will be required (Todini 2004). In urban catchments due to the presence of impermeable surfaces, lag times are typically lower than non-urban catchments, which results in the need for rainfall nowcasts (short term forecasts up to 6 hours ahead) for real-time model application. As outlined above, the time required for demand forecasting in WDS depends on the specific system. As demand driven WDS models are typically in equilibrium, there is, in effect, no lag time in the system, which in addition to the difficulties of quantifying demand, means water demand forecasts are required for all WDS models.

3.5.1 Rainfall Forecasting

Errors and uncertainties associated with rainfall measurement have been outlined in Section 2.4; errors derived from uncertain rainfall predictions will propagate into hydrological models and models of UWWS (Achleitner et al. 2008; Collier 2009). As catchment size increases, discharge is less sensitive to spatial and temporal rainfall uncertainty; however, in urban catchments, with short lag time responses (Berne et al. 2004), hydrographs are more sensitive to the temporal and spatial distribution of rainfall (Segond et al. 2007), as interconnected impervious surfaces decrease system response time (Mejia and Moglen 2010). Such conditions require finely resolved predictions of rainfall intensity for accurate flow prediction (Berne et al. 2004).

A number of methods have been applied for forecasting rainfall. For short timescales (nowcasting) correlation based approaches have been developed, which evaluate consecutive rainfall radar images to derive vectors of the rainfall motion, which are used to propagate the radar image forward in time (Thorndahl et al. 2009; Verworn and Kramer 2005). Advanced correlation methods include filtering features of different scale for vector calculation, which slows the decrease in forecast skill over time (Van Horne et al. 2006).

Kalman filtering has also been applied to a rainfall field decomposed into rain cells (Barillec and Cornford 2009).

A limitation of extrapolation approaches is the inability to simulate the development and decay of rainfall. Further, the uncertainty associated with radar based forecasts is not typically considered, and will be a combination of the uncertainties associated with relating radar reflectivity to rainfall, and those associated with the extrapolation itself. What is often required for effective management is the probability of extreme events, as to be dealt with effectively, such events may require more work and earlier warning (Figure 6; Fabry and Seed 2009). Uncertainty in nowcast rainfall has been investigated with the Generalised Likelihood Uncertainty Estimation (GLUE) procedure (Thorndahl et al. 2010). In application of a tracking approach for rainfall forecast in Linz, Austria, a forecast horizon of greater than 90 minutes introduced intolerable uncertainties into sewer flow predictions (Achleitner et al. 2008). Fabry and Seed (2009) conclude that for radar forecasts there will always be significant errors in short term high resolution forecasts for urban applications, which emphasises the need to account for uncertainty.

The Nimrod system developed by the UK Met Office (Golding 1998) uses vector approaches to estimate motion for rainfall predictions up to an hour ahead, and for longer lead times combines the motion vectors and Numerical Weather Prediction (NWP) derived wind fields (Collier 2007). The prediction of convective rainfall fields, notably their genesis, has been identified as a problem, and recent finer 1-2km grid-scale models designed to better resolve convective processes have been developed (Roberts and Lean 2008). Fine resolution NWP have also been shown to better capture small scale orographic effects in comparison to coarser resolution models (Roberts et al. 2009). The STEPS algorithm combines the advantages of NWP with radar extrapolation to provide an ensemble forecast (Bowler et al. 2006), which has been applied to predict sewer system flows (Liguori et al. 2011); forecast skill was shown to decrease with time and at smaller spatial scales (Schellart et al. 2009). A difficulty identified with NWP, however, in the context of urban runoff prediction is that although the general pattern of rainfall may be produced, the displacement of heavy rainfall centres affects the ability to model the correct rainfall in small urban locations, with implications for uncertainty in sewer flow modelling (Rico-Ramirez et al. 2009).

Recently, more attention has been paid to the use of ensemble NWP as a means to propagate rainfall uncertainty into the hydrological and runoff models. Because of the sensitivity of meteorological models to physical representation and initial conditions, meteorological forecasting uncertainty is likely to dominate over longer time-periods (e.g. 2-3 days) than is suitable

for deterministic forecasting. Due to computational constraints in generating an ensemble, and the lower resolution that ensemble models have in comparison to deterministic models, it has been argued that ensembles should complement deterministic forecasts (Gouweleeuw et al. 2005). Further, computational time in generating the rainfall forecast to drive an UWWWS model will potentially take up time required to make the control decisions. At the same time, however, an understanding of the uncertainty may be important in determining the 'optimal' control setting, which over a long time-period in the face of driving condition uncertainty, may be one that minimises risk as opposed to that which produces the optimal deterministic control setting. Even in the case where quantitative predictions of rainfall are not exact, derived sewer control strategies (based on sewer model predictions) have been found to be more effective than using reactive control to measured system states (Verworn and Kramer 2005).

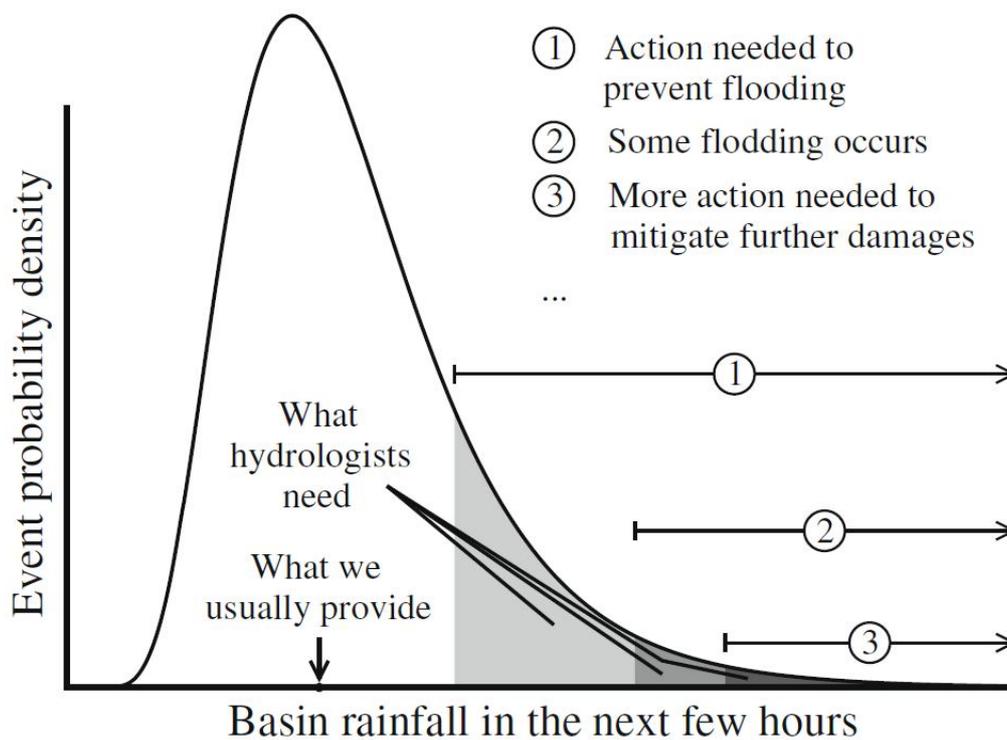


Figure 6. Schematic illustration of the probability density distribution of a rainfall forecast for a specific event (solid line), the information generally provided by radar-based QPF systems (the average expected rainfall, indicated with an arrow), and the information usually needed by basin managers (the probability that an event exceeds specific thresholds, shaded areas (the probability that an event exceeds specific thresholds, shaded areas; from Fabry and Seed 2009).

3.5.2 *Water Demand Forecasting*

Forecasting water demand in a given urban area is important over longer timescales for planning and system design (Froukh 2001), and over shorter timescales to assist managers in balancing the needs of different consumers. Short-term water demand forecasting (e.g. hourly and daily timescales) is a key component to facilitate real-time control decisions (Rao and Salomons 2007). Water demand is a complex function of climatic, social, economic and cultural drivers (Arbues et al. 2003; Dandy et al. 1997). A number of data driven techniques have been developed for predicting future water demand using past observations of measured water demand, climatic variables and economic factors.

A number of methods have been applied based purely on measured water demand and predict future water demand based on seasonal, weekly (e.g. weekend and weekday) and daily trends (Alvisi et al. 2007; Cutore et al. 2008; Nasser et al. 2010; Quevedo et al. 2010; Shang et al. 2006). For up to 3 day lead times, historic daily demand was found sufficient to predict future demand in Bangkok, however over medium term prediction with up to 6 month lead times, climatic as well as social variables (e.g. education status) were required to achieve optimal predictions (Babel and Shinde 2011). Climatic variables that have been used to predict water demand include temperature and rainfall (Asefa and Adams 2007; Babel and Shinde 2011; Gato et al. 2007; Herrera et al. 2010). Few studies have attempted to use socio-economic factors as data are likely more difficult to measure and update.

Methods applied for water demand prediction include Artificial Neural Networks (Babel and Shinde 2011; Cutore et al. 2008), Autoregressive models (asafa and adams; Molina barca), time-series models (Zhou et al. 2002), and more recently genetic programming (Nasser et al. 2011), random forests (Herrera et al. 2010) and more advanced regressive techniques including multivariate adaptive regression splines (Herrera et al. 2010). Jain et al (2001) 2001 found that an ANN model outperformed regression and time series models for predicting weekly water demand. Herrera et al (2010), however, found that ANN models performed poorly in comparison to support vector regression for predicting hourly water demand in an urban area of south-east Spain. In other studies good forecasting has been achieved using time-series models (zhou, et al 2002). The relative performance of different models is likely to depend on the specific environment, the variables of interest, and the forecast time period, which limits assertions about general applicability. However, as recommended by Hererra et al (2010), regular model updating

(calibration) as new information becomes available, is likely to improve the performance of all models.

Few studies have attempted to quantify the uncertainty associated with water demand predictions. Cutore et al (2008) applied the SCEM-UA algorithm to calibrate and quantify the uncertainty of an ANN model applied to predict water demand in Catania, Italy. A key limitation of many studies is that although demand may be predicted based on historic climatic variables, the forecasting situation requires separate predictions of those climatic variables, which as discussed in Section 3.5.1 contain considerable uncertainty. Based on uncertainties in forecast temperature and precipitation Zhang et al (2007) applied a perturbation method to temperature and a replacement method to rain days to generate an ensemble for 3 day ahead water demand predictions. Prediction bounds derived from such models may be propagated into a WDS model. Predictions of water demand are typically aggregated to a given user area; like rainfall predictions demand forecasts need to be downscaled to a level conducive to model application, which will add further uncertainty to WDS model predictions (Kang and Lansey 2009).

3.5.3 *Real-Time Models*

Given the short times available to make real time predictions, particularly when considering integrated systems, alternative modelling approaches have been employed that in the spectrum of model complexity, often have a smaller physical basis, and are more data driven in approach. Such models applied to UWWWS may take the form of simplified conceptual models, where for example flow in sewer systems is simulated with linear reservoirs (Butler and Schutze 2005). Such an approach has also been applied to simulate river systems, where simplified model parameters are calibrated against more detailed physical models of the system (e.g. 1D St. Venant Equation; Meirlaen et al. 2001). Such models have also been termed **Mechanistic Surrogate Models**, that although less similar to the real world, are computationally more feasible to apply in real-time (Vanrolleghem et al. 2005). During simplification, where model components are lumped, parameters in the new model often have little physical meaning. To overcome the difficulties imposed by the need for extensive data collection to constrain the surrogate models, more complex models have been applied to generate data for model calibration (Vanrolleghem et al. 2005). It should be noted however, that more complex and so called 'physically based' models, whilst often having lower structural uncertainty than simplified models, often contain a lot of parameters that are poorly constrained. The adequacy of using such a model

to constrain the structural uncertainty in a simpler model will be dependent also on parameter and input uncertainty.

Artificial Neural Networks (ANN) are a popular example of a **Data-Driven Model** structure that do not retain any model architecture relating to the system in question. Such models are essentially vector mapping functions where an input vector(s) is applied to the network, and an output vector is produced in response. The network architecture consists of an input layer that consists of processing elements that connect input variables to hidden layers in the model (Figure 7). The hidden layers receive weighted information from all nodes in the input layer, and provide separately weighted information to an output layer. A nonlinear transfer function can then be applied at the hidden layer node and output layer node to produce the output signal value (Jeong and Kim 2005; Zealand et al. 1999).

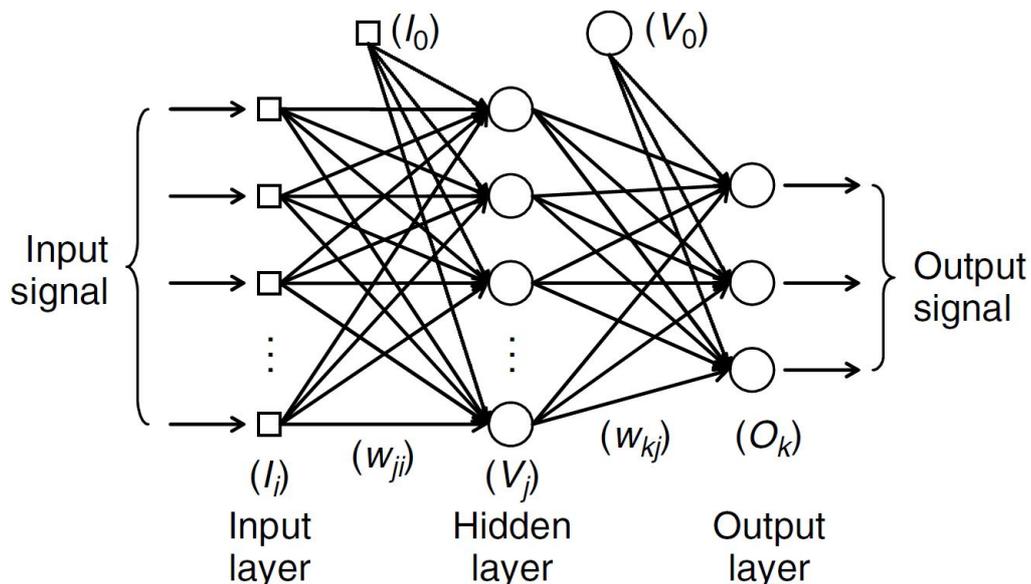


Figure 7. Architecture of a feed-forward ANN with one hidden layer (V_j); weights (w_{ij} and w_{kj}) control the strength of connection between the input layer (I_i) and the hidden layer, and between the hidden layer and the output layer (O_k) to produce the output signal (Jeong and Kim 2005).

The ANN is essentially a black-box model that requires calibration of weights and internal parameters to produce the output model response, which is then applied in a verification period to evaluate calibration performance of the ANN. An advantage of an ANN approach applied in calibration is that the statistical distribution of the output data need not be known a priori such that non-stationarities and non-gaussianities (reflecting structural uncertainties) are implicitly incorporated into the calibration procedure (Zealand et al. 1999). Further, the model is applicable in real-time due to fast computational

time compared to other modelling approaches (e.g. EPANET). In real-time forecasting ANN approaches have been applied to determine optimal control decisions based on current system states.

The energy cost minimisation system (ENCOMS), initially developed by POWADIMA, an EC Vth framework project (Jamieson et al. 2007), and later by Halcrow (Rao et al. 2007), couples a GA with an ANN to determine future optimal control conditions for the WDS considering minimum pressure requirements and optimising cost minimisation (timing of pumping schedules in low tariff periods). The ANN is trained on a series of EPANET simulations which use different decision variables (e.g. pump status on/off) to generate 24 hour ahead system forecasts (Rao and Salomons 2007). The ANN is then coupled with a GA, and using the current SCADA information and a demand forecasting model (Alvisi et al. 2007) the optimal operating conditions are determined using the GA-ANN. The optimal conditions are then applied to the WDS and the cycle repeated when the new SCADA data becomes available.

Significant cost savings were demonstrated in comparison to EPANET simulations of current practice when the ANN models were applied to a WDS in Haifa (Salomons et al. 2007), and also to a WDS in Valencia (Martinez et al. 2007). A clear limitation with this approach, as with other approaches that use existing, complex models for training/calibration, is the assumption that EPANET simulations provide adequate training of the ANN, and also adequate evaluation of GA-ANN performance. For real world application, errors associated with EPANET training simulations need to be constrained and propagated through the ANN to provide information on the reliability of optimised control strategies, which also need to be confronted with real world data. This is also the case when applying such models in real-time where the input data are also forecast from other models.

Modelling approaches applied in real time for system optimisation need to be ran a number of times to determine the best system operating procedure. When uncertainty is considered alongside optimisation a number of potential problems exist. First, if uncertainty is accounted for correctly, a number of control solutions may be produced that perform equally well; a form of equifinality that means it may be difficult to chose the best control procedure. Second, depending on model predictions of future system state, the time required to take appropriate actions (time horizon) may not be known a priori until the modelling is undertaken to predict future conditions. Third, propagating and reducing uncertainty in real-time will require an increased modelling time (Figure 7), which will require a longer input forecast at the start of the modelling period. Given that forecast accuracy declines with

increased lead time, accounting for this input uncertainty may conversely increase modelling uncertainty. A parameter that provided increased weight to near future events, as reported by Pleau et al. (2005), may help account for increased uncertainty with lead time. As noted by Cloke and Pappenberger (2009), an 'optimal framework' for dealing with uncertainty will inevitably blend more formal and informal treatments of uncertainty in the modelling cascade.

3.6 Summary

Real-time control in UWS is important to provide optimal system performance to meet a range of demands imposed by a number of stakeholders. Real-time modelling, facilitated by advances in data collection and communication techniques, has the potential to improve system performance in real-time by offering tools to forecast future system states and determine optimal control solutions. However, as discussed in this section, real-time modelling must deal with and propagate uncertainties associated with input data derived from both measurements and forecast models, and uncertainties associated with system modelling. The time available to make real time predictions is a key constraint on the detail with which a system may be represented, and the extent to which uncertainty may be dealt with adequately. Given the need for forecasts of demand for WDS modelling, and rainfall in UWWS modelling, an optimal framework will blend different methods for dealing with uncertainty to inform real-time control.

4 Data Assimilation and Real-Time Error Reduction Methods

4.1 Introduction

System Forecasts made in real time combine a representation of the current system state with forecasts of model forcing conditions (e.g. rainfall and water demand). As considered in Section 3, model forecasting conditions often contain significant errors. Forecasting errors are also related to model structural errors and uncertain parameters, and the errors in the system state at the time of forecasting. Sensitivity to initial state conditions has received wide attention in weather forecasting (Rabier et al. 1996), and can also lead to divergence between model predictions and observations.

Data Assimilation (DA) is a name provided to a class of methods that seek to combine uncertain models with uncertain data to provide the best estimate of the system state at a given point in time at which observations are available (Figure 8). The system model is propagated forwards in time from the initial states until the next set of observations become available (Figure 8b; Black Squares). At this point based on the observations, model states are updated (Figure 8b: vertical simulated discharge), and the model propagated forwards in time in response to the model forecasting conditions (e.g. observed or forecast rainfall) until the next set of observations becomes available. Data assimilation is closely related to State Estimation (SE), where measurements of the system are combined with numerical models to gain a global view of the system (Bargiela and Hainsworth 1989). The generic state-space formulation of a stochastic model may be represented as follows (see also Liu and Gupta 2007; Vrugt and Robinson 2007):

$$x_t = f(x_{t-1}, \theta, u_t) + \eta_t, \quad \eta_t \sim N[0, R_t^\eta] \quad (4.1)$$

$$z_t = H_t(x_t^*) + \varepsilon_t, \quad \varepsilon_t \sim N[0, R_t^\varepsilon] \quad (4.2)$$

Where x_t and x_{t-1} represent system state vectors at time t and $t-1$; x_t^* denotes the true model states; f represents model structure propagating the system from $t-1$ to t in response to the input vector u_t , and θ is a vector of time-invariant model parameters. The observation z_t is related to the model states through an observation mapping operator H_t ; η_t represents model error, and ε_t represents measurement error. Thus, in a DA procedure a model forecast is made and taking into account the relative errors between observation and

model, system state is updated through a combination of Equation 4.1 and Equation 4.2.

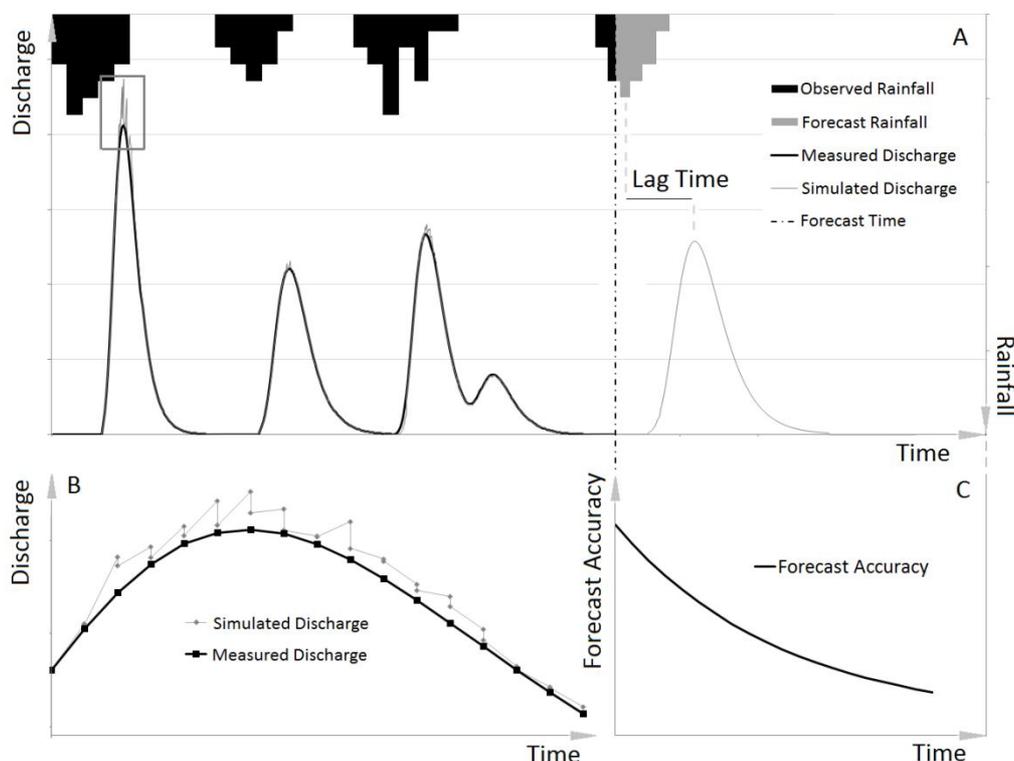


Figure 8. Schematic illustration of Data Assimilation. Measured and predicted discharge in response to rainfall for a data assimilation period, prior to discharge forecast (A); Enlarged version of the grey box in the left hand side of A showing simulation predictions and simulation updates, where vertical simulation lines show forecast innovation in response to measured discharge (B); Decline in Forecast Accuracy with increased time from the last assimilation step (C).

In the general implementation of DA the system states are updated and are then used as the initial conditions of a future deterministic system forecast. The future forecast skill is limited up to a time horizon where the initial conditions are washed out (Madsen and Skotner 2005); the forecast accuracy generally declines with increased distance from the forecast time (Figure 8C) due to model structural and parameter errors and errors propagated from the input conditions. In addition to DA approaches, error-correction procedures have also been developed for real time application that based on pre-determined understanding of forecast error (i.e. in the absence of real-time observations) correct future model predictions.

A number of methods have been developed and applied to reduce uncertainty in real-time model application, most widely in related scientific disciplines, including meteorology, hydrology and climatology (Evensen 2003; Matgen et al. 2010; van Leeuwen 2009). Such methods, despite having significant potential, have seen relatively little application in UWS modelling (Brdys et al. 2008; Kang and Lansey 2009). Section 4 will review existing DA and error-correction approaches for dealing with uncertainties in real-time modelling of UWS, considering computational issues related to these approaches and their implementation in UWS. Appendix A. contains a summary Table of some applications of the reviewed methods, which is an extension of the Table of uncertainty quantification methods presented in Hutton et al. (2011).

4.2 Kalman Filtering

The Kalman filter (KF) is a sequential filter method that utilises a model forecast step using observations from the previous time step to estimate current system state, and an update step utilising the observation at the current time step to refine the current forecast. The Kalman filter provides a solution provided that Equations 4.1 and 4.2 represent a Gaussian linear system (Evensen 1992; Evensen 2003):

$$x_t^a = x_t^f + K_t(z_t - H_t x_t^f) \quad (4.3)$$

K_t is the Kalman gain and is often expressed as:

$$K_t = P_t^f H_t^T (H_t P_t^f H_t^T + R_t^\varepsilon)^{-1} \quad (4.4)$$

where P is the error covariance matrix of the state variables; the superscripts f and a denote the state forecast (predicted) and analysed (corrected) estimate respectively; and T is the transpose. Equations 4.3 and 4.4 can be calculated as each new observation becomes available. The magnitude of the Kalman gain applied to the state estimate in Equation 4.3 depends on the relative magnitude of the observation error covariance (R_t^ε) and the state error covariance (P_t^f); a high measurement error covariance will result in a smaller update of the forecast vector. Although the Kalman filter equations explicitly account for observation and model errors, assimilation errors are sensitive to the information provided to inform the nature of the prior error models for both observations and the model. As discussed in Hutton et al. (2011), such data in many applications are limited. The KF algorithm has been applied in a WDS to estimate unknown roughness in a linear estimation problem (Todini

1999), to sewer flow forecasting (Gelfan et al. 1999), and water quality modelling (Schilling and Martens 1986). The KF performed well when applied to estimate nodal demand in a WDS in branched areas. However, the KF performed poorly due to strong nonlinearities in looped areas of the WDS network (Kang and Lansey 2009); the KF method is best suited to linear problems, which limits application in UWS models, as such systems are often designed to be looped to increase network resilience/guaranteed consumer demand (Walski et al. 2003).

The Extended Kalman Filter (EKF) was developed to work better in cases of system non-linearity; the model operator f and the observation operator H cannot be applied directly due to non-linearity, but are approximated with tangent linear operators (Jacobian) in equations (Evensen 2003). The EKF has been applied in modelling of sewer pollutant loads (Bechmann et al. 1999), WWTW optimisation in real-time (Brdys et al. 2008), water demand estimation (Nasseri et al. 2010), and more extensively in oceanography (Bertino et al. 2003), where for weakly nonlinear problems the EKF has been shown to perform well in comparison to the EnKF (Hoteit et al. 2005; Madsen and Canizares 1999). Assuming all other unknowns were previously calibrated (e.g. roughness), Shang et al. (2006) applied the EKF approach for real time calibration of water demand; a demand time-series model was applied to predict future demand, and corrected using measured node heads and pipe flows (Shang et al. 2006). Application of the EKF may be unstable in situations with large nonlinearities (Hoteit et al. 2005); unbounded error variance growth may occur due to a closure approximation which neglects higher order derivatives (Evensen 1994; Miller et al. 1994). Further, integrating the tangent linear model forwards to get the error covariance matrix comes at an expensive computational cost (Zhang and Pu 2010). Such computational costs have been alleviated to some extent through the application of reduced rank square root filters (Verlaan and Heemink 1997).

The Ensemble Kalman Filter (EnKF), introduced by Evensen (1994), was developed to overcome some of the problems associated with the EKF, by propagating an ensemble (n) of model states derived from Monte Carlo perturbations of input states u_t by adding a noise term $\delta_t \sim N[0, U_t]$. Equations 4.3 and 4.4 are applied to each ensemble member, and the error covariance is calculated from the ensemble mean (\bar{x}_t^f), which overcomes the problem of not knowing the true state when calculating the error covariance, and the computational costs associated with propagation of the error covariance matrix (Burgers et al. 1998; Liu and Gupta 2007):

$$P_t^f = \overline{(x_t^f - \bar{x}_t^f)(x_t^f - \bar{x}_t^f)^T} \quad (4.5)$$

A key feature of the EnKF is that an ensemble of observations is used in the update at each time step by adding noise to the measured variable. This step is essential to maintain variance in the updated ensemble but at the same time does not affect the prediction of the ensemble mean (Burgers et al. 1998; Evensen 2003). Perturbations, however, may increase the sampling error, and thus an alternative set of schemes apply square root filter methods (EnSRF) to avoid perturbing initial states (Anderson 2001; Tippett et al. 2003; Whitaker and Hamill 2002). The EnKF has been applied most widely in the fields of climatology, meteorology and oceanography (Annan et al. 2005; Evensen 2003; Hargreaves et al. 2004; Hoteit et al. 2005; Houtekamer et al. 2005; Zhou et al. 2006), and has been shown to work well in application to highly nonlinear problems (Evensen 1997). The EnKF has also been applied more recently to hydrological problems (Clark et al. 2008; Moradkhani et al. 2005b; Neal et al. 2007; Xie and Zhang 2010).

The EnKF has been considered computationally feasible for certain applications (Vrugt and Robinson 2007), however as argued by Madsen and Skotner (2005), to represent the covariance matrix properly 100 ensemble members have been required which may still be too expensive for operational real-time conditions. Other studies have required only 32 ensemble members (Weerts and el Serafy 2005); the required amount will depend on the specific system being simulated. Computational requirements may be reduced through the application of a local Kalman filter (Ott et al. 2004), and also modifications made by Pham (2001) who developed a second-order-exact EnKF, which ensures the ensemble members match the true covariance matrix. The method, which was applied with a forgetting factor that relies more on the observation in constructing the analysis vector from the gain matrix, reduced the number of ensemble members required (Pham 2001). To overcome recalculation of the error covariance matrix, the most costly part of the Kalman filter approach, Canizares et al (2001) applied the Kalman Filter with a constant weight matrix, calculated as the mean gain derived from an off-line ensemble filter computation.

Although the EnKF is less vulnerable to nonlinearity compared to EKF as the model uses the full nonlinear model, filter divergence could occur when nonlinearities result in a strongly non-gaussian distribution of the ensemble, as the filter only uses the first two moments of the ensemble in the correction step. In such cases the perturbed approach may better withstand non-linear error growth than deterministic approaches (Lawson and Hansen 2004). However, the EnSRF has outperformed the EnKF in application to rainfall-runoff modelling (Clark et al. 2008). Further methodological advances and

variants on the EnKF (see Evensen (2003) for a historical review) include the Retrospective Ensemble Kalman Filter (REnKF), which uses a measurement to update model states at all t-n time steps, where n represents the time steps in which the measurements are influenced by the model states to be updated (Pauwels and De Lannoy 2006). The REnKF may, for example, be applied to in the case where CSO volume measurements reflect system states at -n time-steps. Komma et al. (2008) presents an alternative approach that avoids calculating the Jacobian between states and observations, where once the analysed state of each ensemble is determined for downstream discharge, multiple realisations of each ensemble are propagated forwards from t-1, with random error added to each system state realisation. The states from the realisation which produce output discharge as close to the analysed state of the original ensemble member as possible are used in the update.

4.3 Sequential Monte Carlo Sampling (Particle Filtering)

Particle filtering (PF) is a technique for implementing a recursive Bayesian Filter by MCS (Arulampalam et al. 2002). The posterior density function is represented by a series of particles, each with associated weights (w_t^i):

$$P(x_t|z_t) \approx \sum_{i=1}^n w_t^i \delta(x_t - x_t^i) \quad (4.6)$$

where n is the number of particles and δ is the Dirac function. The weights are normalised such that: $\sum_i^n w_t^i = 1$. As the number of samples (n) becomes very large, the filter approaches the optimal posterior estimate. The method differs from EnKF as instead of updating particle state estimates, particle weights are updated using sequential importance sampling; such a procedure is adopted since sampling from the true posterior is usually impossible:

$$w_t^{i(*)} = \frac{P(x_t^i|z_t)}{q(x_t^i|z_t)} \quad (4.7)$$

where $*$ denotes the particle weight prior to normalization and $q()$ is the proposal distribution (importance density). Equation 4.7 can be re-arranged (see Arulampalam et al. 2002) to derive the sequential case where particle weights are updated recursively:

$$w_t^{i(*)} = w_{t-1}^{i(*)} \frac{P(z_t|x_t^i)P(x_t^i|x_{t-1}^i)}{q(x_t^i|x_{t-1}^i, z_t)} \quad (4.8)$$

The selection of the proposal distribution is an important decision (Arulampalam et al. 2002); generally the prior is chosen, which reduces the weighting update to:

$$w_t^{i(*)} = w_{t-1}^{i(*)} P(z_t | x_t^i) \quad (4.9)$$

Particle filtering therefore occurs recursively through the application of Equation 4.1, and particle weight updating using Equation 4.2 and Equation 4.9. A key advantage of particle filtering over KF is that the state-space model need not be linear or Gaussian. The Particle filtering method has been applied in climatology, meteorology and hydrological modelling (Moradkhani et al. 2005a; Pham 2001; Salamon and Feyen 2009; Smith et al. 2008; van Leeuwen 2009; Vossepoel and van Leeuwen 2007), and has been shown to outperform EnKF, but at increased computational cost (Pham 2001).

A key problem in the application of particle filtering is filter degeneracy; as only particle weights are updated, particles move away from high probability regions, which results in only a few samples actually representing the PDF. Computational time is wasted on calculating weights that are not contributing to understanding of the posterior distribution. Two broad methods for dealing with this problem may be considered (van Leeuwen 2009): Re-sampling the weights and moving the particles.

Re-sampling the weights may be achieved through sequential importance sampling (SIR) by abandoning particles with lower weights and generating multiple copies of particles with higher weights; the number of copies generating being proportional to particle weight (van Leeuwen 2009). Re-sampling may be inadequate, however for dealing with the divergence problem in large-scale applications as the ensemble size required for a successful filter performance scales exponentially with problem size (Snyder et al. 2008; van Leeuwen 2009). An alternative approach to resampling is to move the particles themselves closer to the observations by employing an EnKF, calculating the corrected particle weights, and finally applying a resampling procedure (van Leeuwen 2009). Van Leeuwen (2009) provides a detailed review of other methods for dealing with the degeneracy problem, including localization, the guided particle filter where particles are confronted with observations prior to the actual measurement time; the auxiliary particle filter which uses new weights to resample the initial ensemble; and the backtracking filter, which moves back in time and re-samples with a larger ensemble. As with the EnKF propagating multiple realisations of the model forward to account for uncertainty will take up computational time that may not be suitable for real-time application. A potential advantage of the PF over EnKF is that implementation within any

model structure may be easier given that the method does not update models states at each time step. Neither method, however, has been applied widely in UWS modelling, and their application to account for uncertainty in different system models remains to be evaluated.

4.4 Variational Data Assimilation

Variational Data Assimilation (VDA) is a method unlike the family of KF and PF real-time methods in that instead of being applied at a point at which data becomes available, VDA operates over a time-series of observation points; a method widely applied in weather forecasting (Li and Navon 2001). The formulation of VDA (specifically 4D-Var) can be defined as the minimisation of a cost function (J) that measures the weighted sum of squares between the background state and the observations over a given time interval $[t_0, t_n]$ (Ide et al. 1997):

$$J(x_{t_0}) = \frac{1}{2} [x_{t_0} - x_{t_0}^-] B_0^{-1} [x_{t_0} - x_{t_0}^-]^T + \frac{1}{2} \sum_{t=0}^n [z_i - H_i x_i] R_i^{-1} [z_i - H_i x_i]^T \quad (4.10)$$

Where the first term represents error in initial conditions and the second term represents error between model predictions and observations at all time points (where B_0^{-1} is error in prior estimate of state variables at t_0). The objective of the case presented in Equation 4.10 is to identify the best estimate of the initial state condition that minimises the cost function with respect to the initial conditions. The optimisation problem requires the calculation of the cost function, and the gradient, which can be computed from the adjoint technique in an iterative manner (Ide et al. 1997).

If the model dynamics are linear VDA is equivalent to KF with respect to model states at all time steps within the assimilation period. VDA does not model the dynamics of the second order component of state variables, and is therefore not amenable to probabilistic interpretation. Further, the method as presented in Equation 4.10 only considers observation error, as solving the full minimisation problem (that includes model and parameter error terms) in highly dimensional nonlinear problems is very difficult to solve (Li and Navon 2001; Liu and Gupta 2007). If the system is nonlinear, like EKF, the system is represented with the tangent linear mode of M and H ; therefore if nonlinearity is an important characteristic of the system, EnKF and PF are more appropriate (Liu and Gupta 2007). VDA methods are more suitable to complex problems as they are less demanding computationally, however they also do not provide an estimate of the predictive uncertainty. A combined approach where 4DVAR is coupled with EnKF was performed which was

shown to outperform both methods separately, but at a large computational cost (Hansen and Smith 2001).

4.5 Joint State and Parameter Estimation

In the approaches considered for real-time estimation, the majority of approaches do not consider parameter uncertainty and assume for real-time application that an optimal parameter set is already known (Shang et al. 2006). In general this is a realistic assumption and model parameters are often considered to vary more slowly in comparison to the evolution of system state in real time (Kang and Lansey 2009). However, as identified in Section 3, many approaches for parameter calibration do not adequately consider structural and measurement errors which affect the quality of conditioned parameter estimates. A number of methods have been developed which jointly deal with state and parameter estimation. These methods may be broadly classified into approaches that apply DA over a time-series used for calibration (static parameters), and approaches that consider parameters to be time-varying.

Vrugt et al. (2005) developed the Simultaneous (parameter) Optimisation and Data Assimilation (SODA) procedure. SODA implements the SCEM-UA parameter optimisation method in an outer loop, and for each sampled parameter set an inner EnKF loop is applied recursively to the time-series used to evaluate the performance of the parameter set. The method has been applied to hydrological modelling studies where recursive state estimation was only applied to a subset of model states, which improved model performance in both calibration and validation as the method attempted to account for state estimation errors (Vrugt et al. 2006). The method, however, still depends on the ability to specify a realistic estimate of model error, and does not consider the possibility that parameters may also be time varying.

Dual estimation has been developed where both parameters and model states are considered time-varying, where either EnKF (Moradkhani et al. 2005b) or PF (Moradkhani et al. 2005a) is applied not only to modify state variables during each update but also to sample parameter values. Therefore the PDF of model states and parameters is updated at each time-step, and provides good predictions to observed system response. A downside to joint state and parameter estimation is that the model response to a change in a given parameter value may not be observed until a number of time-steps ahead. Salomon and Feyen (2009) applied a PF to the joint state-parameter estimation problem over a time interval of observations reflecting concentration time for the hydrological problem considered (similar concerns may also affect urban

wastewater modelling employing such a procedure). A downside to the joint or dual estimation approach is that parameters are likely to vary more slowly than states; if both are updated each time step instabilities may occur (Liu and Gupta 2007). However, joint state and parameter estimation approaches applied to WDS models have allowed refinement of uncertainty bounds on for example pipe resistance during changing system conditions (Brdys and Chen 1995). Whether parameters require updating at each time step will depend on the specific model in question; In ANN models, for example, the separation of model structures and parameters is more difficult to define.

4.6 Forecasting Error-Correction

The assimilation methods considers thus far in Section 4 update model states based on observations. Therefore the accuracy of the system forecast beyond the last observation time is determined by the 'initial conditions' provided at the last observation. The forecast skill is limited to the time horizon over which the initial conditions are washed out (Madsen and Skotner 2005), and is reliant on the accuracy of the model structure and parameters determined a priori, which may be poor (Refsgaard 1997). An alternative group of error correction methods have been applied which update deterministic forecasts based on a correction model determined prior to the time of system forecast, when observations are available. Such models typically take advantage of the observation that errors are correlated in time, often as a function of the magnitude of the system state to be corrected.

ANN models, part of a class of data-driven models, as introduced in Section 3.5.3, have been applied to simulate model error time-series (Shamseldin and O'Connor 2001). Based on prior time series analysis, Abebe and Price (2003) constructed an ANN model using input nodes relating to past errors, past rainfall and antecedent conditions. The ANN model when trained on past error time-series improved on predictions derived from the conceptual rainfall-runoff model with lead times of 1-6 hours.

Other methods applied to error correction include Autoregressive Time Series models, based on past correlation of errors in the predicted and observed time-series (Lekkas 2008; Lekkas et al. 2001), Genetic Programming (Khu et al. 2001), and local models (Babel and Shinde 2011; Sannasiraj et al. 2005). Romano et al. (2010) applied a Group Method of Data Handling (GMDH) approach to predict future flow/pressure in WDS based on past measurements. The model is built as a network of polynomial functions, which unlike ANN and regression approaches does not require the network

architecture to be specified in advance, and uses all available data. The method could be applied to error-correction. Vojinovic et al. (2003) applied a hybrid modelling approach to simulate UWWS wet weather response which coupled MOUSE (MOUSE 2004) with a stochastic radial basis function neural network. The hybrid approach provided improved performance over the deterministic MOUSE predictions, and also in comparison to MOUSE was applied with an Autoregressive exogenous model (Vojinovic et al. 2003).

In a comparative study of linear, autoregressive and neural network based approaches, Goswami et al. (2005) found that whilst all 8 models considered produce reasonable predictions (Nash-Sutcliffe $R^2 > 90\%$) for 1 day ahead forecasts, the Non-Linear Auto-regressive eXogenous-Input Model (NARXM), the linear transfer function model, and a Neural Network Updating Model were found to be suitable for lead times of up to 6 days (Nash-Sutcliffe $R^2 > 90\%$). This Study, however, used known rainfall to evaluate the error-correction procedures. The ability of error correction procedures to account for strong input errors associated with input driver forecasts has yet to be evaluated, and will depend on the temporal nature of the errors in input drivers.

Error-correction methods have mainly focussed on correcting the model output time-series of interest. Madsen and Skotner (2005) applied a correction procedure that calibrated an error forecast model (AR1 and harmonic models applied) to a time series of innovations, and used pre-determined gain functions to update not only the measurement location but also the state variables. The Harmonic model in particular showed improved performance in comparison to the no update case for up to 24h lead times in a flood forecasting procedure (Madsen and Skotner 2005). Another method to distribute the prediction error from observation points into system states was recently presented by Mancarella et al. (2008), who employed a local model to estimate forecast error at measured locations in the computational domain, and correlations between model states to distribute these error corrections over the model domain. Such an approach is appealing as it represents a computationally efficient method for state updating, and because correlation between modelled states and other locations can be determined off-line prior to on-line application. Further, such an approach provides a means for optimising error correction measurement locations, prior to model application.

The error-correction procedures considered calibrate a model to the residuals derived prior to online forecasting. Such an approach can implicitly account for systematic errors in model structure, parameter values and errors in input values. However, the approach assumes that the output measurement values

contain no error when determining the residual time-series, and also that the fitted time-series model contains no errors. The error-correction when applied is therefore deterministic. An error-correction procedure following for example Asefa (2009) and also Zhang et al. (2009a), where multiple ANN (Bayesian Neural Networks) models are trained to the time series, may provide a more faithful representation of the fit of the error-correction model to the residual time series. When correcting the error in the deterministic model forecast, for each prediction time, an upper and lower bound of the corrected forecast may be given instead of a deterministic prediction. A related approach to account for uncertainty in future predictions was presented by Shrestha et al (2009). First, a process based model was applied with the GLUE methodology to reproduce prediction intervals over a calibration period derived from parameter sampling. Then two ANNs were trained, one to reproduce each prediction bound, based on model input data. The application of such an approach is computationally feasible in real-time to predict uncertainty bounds as each new observation becomes available; the approach is reliant on the accuracy of the underlying model and the robustness of the original prediction bounds. In theory, such an approach may be extended to methods such as that of Schoups and Vrugt (2010) that attempt to represent different forms of uncertainty in the model prediction bounds.

Error-correction procedures, because of their computational efficiency, are potentially the most suitable method to accounting for error in real-time model forecasting, particularly where computational time is limited by the need to identify the optimal management scenario for a system. However, in modifying control (actuator) settings the system in itself is modified at each model run. Therefore applying a priori error corrections to scenarios may not adequately reflect system error, as it is unlikely that data are available for all potential management scenarios to constrain an error correction procedure for different system control settings. Further, it may also be difficult to provide system state updates from errors determined at specific measuring locations, given that the relationship between the measurement site and other system states will be by definition modified by the control settings.

4.7 Summary

A number of real-time modelling approaches for quantifying and reducing uncertainty have been considered in Section 4. VDA has seen widest application in Meteorological studies, with explicit consideration of observational error and, relative to EnKF and PF, lower computational expense; however, solving for other forms of error in complex models is

difficult. Like EKF the method may not be suitable for highly nonlinear problems, and also does not provide any information on predictive uncertainty. EnKF and PF approaches are better suited to processing data that arrives in real time, and the ensemble nature of the methods makes them more amenable for dealing with model non-linearities. Both of these approaches attempt to represent model state error explicitly, however, as with model calibration approaches discussed in Hutton et al. (2011), information to constrain input data uncertainty (e.g. rainfall) and output data uncertainty (e.g. pertaining to system states or sewer CSO) alongside structural uncertainty, may be difficult to define. Unlike EnKF and PF approaches, computationally efficient error-correction procedures attempt to correct for modelling error in future system forecasts based on off-line residual calibration. Such methods have the potential to implicitly account for a number of error types provided they are manifest in the deterministic model's residual time-series.

Data availability and data quality are key factors governing the applicability of some of the methods considered in Section 4. Improved rainfall monitoring and sensor placement/performance, as briefly considered in Section 3, are other areas to be addressed in PREPARED work package 3 that will facilitate the application of the aforementioned methods for dealing with uncertainty. EnKF and PF methods have been applied successfully within model calibration frameworks to account for system uncertainty. Such application suggests a hierarchical approach for dealing with system uncertainty in real-time, combining calibration, state updating and error-correction methods applied at different temporal scales may provide an optimal framework to facilitate real-time control in UWS.

5 Conclusion

This reports fulfils the requirements of Deliverable 3.6.2 within work package 3.6 of the PREPARED Enabling change project (EC Seventh Framework Programme Theme 6), and has evaluated existing methods applied in a number of scientific fields for Data Assimilation and Error-correction to facilitate the application of numerical models in real-time. Real-time modelling in Urban Water Systems has been applied to help optimise the use of existing water supply and sanitation systems. However, such modelling approaches often do not consider inherent system uncertainty that may originate from a variety of sources and affect our ability to identify optimal operational solutions.

Data Assimilation approaches have been applied and developed most widely in related scientific disciplines for updating model predictions in real-time as new measurements become available. Kalman Filtering, notably Ensemble Kalman Filtering, and Particle Filtering methods provide approaches for propagating system uncertainty in real-time and accounting for data and model errors. Such methods provide the potential to account for uncertainty in real-time modelling introduced by model simplification and uncertainty associated with input forecasts. However, such methods are potentially demanding computationally for data assimilation, and like many of the methods reviewed in Hutton et al. (2011) are limited by the ability to specify the structure of model errors and data errors.

Error-correction methodologies are relatively simple to implement and in comparison to DA methods provide potential to reduce uncertainty in real-time system forecasts beyond the point of the last assimilation time step. Such methods can implicitly account for a range of uncertainties provided these uncertainties are manifest in the deterministic model residual time-series derived off-line prior to application. The error-correction methodologies reviewed mostly provide deterministic corrections to output time-series, despite the methods themselves containing uncertainties.

Joint state and parameter estimation approaches have also been applied in limited circumstances, where DA filters have been applied within calibration frameworks. Such methods suggest a hierarchical approach for dealing with model uncertainty, combining model calibration, data assimilation and error-correction applied at different temporal scales may provide an optimal framework to account for uncertainty in real-time modelling and control in Urban Water Systems. Although the methods presented here, as well as the techniques and methodologies that will be implemented in Task 3.6.2 can be

considered as generic, the final selection of the methodologies to be applied depends also on the specific requirements of the PREPARED cities selected for demonstration.

6 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.
- Achleitner, S., Fach, S., Einfalt, T., and Rauch, W. (2008). "Nowcasting of rainfall and of combined sewage flow in urban drainage systems." *11th International Conference on Urban Drainage* Edinburgh, Scotland, UK.
- Almeida, M. C., Bulter, D., and Frielder, E. (1999). "At-source domestic wastewater quality." *Urban Water*, 1, 49-55.
- Alvisi, S., Franchini, M., and Marinelli, A. (2007). "A short-term, pattern-based model for water-demand forecasting." *Journal of Hydroinformatics*, 9(1), 39-50.
- Anderson, J. L. (2001). "An ensemble adjustment Kalman filter for data assimilation." *Monthly Weather Review*, 129(12), 2884-2903.
- Annan, J. D., Lunt, D. J., Hargreaves, J. C., and Valdes, P. J. (2005). "Parameter estimation in an atmospheric GCM using the Ensemble Kalman Filter." *Nonlinear Processes in Geophysics*, 12(3), 363-371.
- Arbues, F., Garica-Valinas, M. A., and Martinez-Espineira, R. (2003). "Estimation of residential water demand: a state-of-the-art review." *Journal of Socio-Economics*, 32, 81-102.
- 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.
- Asefa, T. (2009). "Ensemble Streamflow Forecast: A GLUE-Based Neural Network Approach." *Journal of the American Water Resources Association*, 45(5), 1155-1163.
- Asefa, T., and Adams, A. (2007). "Short-term urban water demand forecast models in action: Challenged from model development to implementation to real-time operations." *World Environmental and Water Resources Congress 2007: Resoring Our National Habitat*, Tampa, Florida, U.S.A.
- 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.
- Austin, R. G., Waanders, B. V., McKenna, S., and Choi, C. Y. (2008). "Mixing at cross junctions in water distribution systems. II: Experimental study." *Journal of Water Resources Planning and Management-Asce*, 134(3), 295-302.
- Babel, M. S., and Shinde, V. R. (2011). "Identifying Prominent Explanatory Variables for Water Demand Prediction Using Artificial Neural Networks: A Case Study of Bangkok." *Water Resource Management*, 25, 1653-1676.

- 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, Cincinnati, Ohio, USA*.
- Bargiela, A., and Hainsworth, G. D. (1989). "Pressure and Flow Uncertainty in Water-Systems." *Journal of Water Resources Planning and Management-Asce*, 115(2), 212-229.
- Barillec, R., and Cornford, D. (2009). "Data assimilation for precipitation nowcasting using Bayesian inference." *Advances in Water Resources*, 32(7), 1050-1065.
- Barillon, B., Zenasni, A., Dauthuille, P., Cren-Olive, C., Jaffrezic-Renault, N., and Chappier, J. (2010). "On-line monitoring station for water quality assessment." *Integrating Water systems*.
- Bechmann, H., Nielsen, M. K., Madsen, H., and Puoulsen, N. K. (1999). "Grey-box modelling of pollutant loads from a sewer system." *Urban Water*, 71-78.
- Beck, M. B. (2005). "Vulnerability of water quality in intensively developing urban watersheds." *Environmental Modelling & Software*, 20(4), 379-380.
- 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.
- Berne, A., Delrieu, G., Creutin, J. D., and Obled, C. (2004). "Temporal and spatial resolution of rainfall measurements required for urban hydrology." *Journal of Hydrology*, 299(3-4), 166-179.
- Bertino, L., Evensen, G., and Wackernagel, H. (2003). "Sequential data assimilation techniques in oceanography." *International Statistical Review*, 71(2), 223-241.
- Bertrand-Krajewski, J. L., Bardin, J. P., Mourad, M., and Beranger, Y. (2003). "Accounting for sensor calibration, data validation, measurement and sampling uncertainties in monitoring urban drainage systems." *Water Science and Technology*, 47(2), 95-102.
- Beven, K., Heathwaite, L., Haygarth, P., Walling, D., Brazier, R. E., and Withers, P. (2005). "On the concept of delivery of sediment and nutrients to stream channels." *Hydrological Processes*, 19, 551-556.
- Bilotta, G. S., and Brazier, R. E. (2008). "Understanding the influence of suspended solids on water quality and aquatic biota." *Water Research*, 42(12), 2849-2861.
- Bilotta, G. S., Brazier, R. E., Haygarth, P. M., Macleod, C. J. A., Butler, R., Granger, S., Krueger, T., Freer, J., and Quinton, J. (2008). "Rethinking the contribution of drained and undrained grasslands to sediment-related water quality problems." *Journal of Environmental Quality*, 37(3), 906-914.
- Bloch, H. (1999). "The European Union Water Framework Directive, taking European water policy into the next millennium." *Water Science and Technology*, 40(10), 67-71.

- Blokker, E. J. M., Vreeburg, J. H. G., Buchberger, S. G., and van Dijk, J. C. (2008). "Importance of demand modelling in network water quality models: a review." *Drinking Water Engineering and Science*, 1, 27-38.
- Borchardt, D., and Statzner, B. (1990). "Ecological Impact of Urban Stormwater Runoff Studied in Experimental Flumes - Population Loss by Drift and Availability of Refugial Space." *Aquatic Sciences*, 52(4), 299-314.
- Boulos, P. F., Lansley, K. E., and Karney, B. W. (2004). *Comprehensive Water Distribution Systems Analysis Handbook For Engineers and Planners*, MWH Soft, Inc., Pasadena, California.
- Bowler, N. E., Pierce, C. E., and Seed, A. W. (2006). "STEPS: A probabilistic precipitation forecasting scheme which merges an extrapolation nowcast with downscales NWP." *Quarterly Journal of the Royal Meteorological Society*, 132, 2127-2155.
- 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.
- Brdys, M. A., and Chen, K. (1995). "Set membership estimation of state and parameters in quantity models of water supply and distribution systems." *Automatisierungstechnik*, 43(2), 77-84.
- Brdys, M. A., Grochowski, M., Giminski, T., Konarczak, K., and Drewa, M. (2008). "Hierarchical predictive control of integrated wastewater treatment systems." *Control Engineering Practice*, 16(6), 751-767.
- Broks, K., Geenen, A., Nelen, F., and Jacobsen, P. (1995). "The potential of real time control to reduce combined sewer overflow." *Nordic Hydrology*, 26, 223-236.
- Buchberger, S. G., and Wells, G. J. (1996). "Intensity, duration, and frequency of residential water demands." *Journal of Water Resources Planning and Management-Asce*, 122(1), 11-19.
- Burgers, G., van Leeuwen, P. J., and Evensen, G. (1998). "Analysis scheme in the ensemble Kalman filter." *Monthly Weather Review*, 126(6), 1719-1724.
- 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.
- Campisano, A., Schilling, W., and Modica, C. (2000). "Regulators' setup with application to the Roma-Cecchignola combined sewer system." *Urban Water*, 235-242.

- 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.
- Canizares, R., Madsen, H., Jensen, H. R., and Vested, H. J. (2001). "Developments in operational shelf sea modelling in Danish waters." *Estuarine Coastal and Shelf Science*, 53(4), 595-605.
- Casadio, A., Maglionico, M., Bolognesi, A., Artina, S., Salmoiraghi, G., and Locascio, A. (2008). "Toxicity and pollutant impact analysis into urban river due to CSO overflows." *11th International Conference on Urban Drainage*, Edinburgh, Scotland.
- Castro, A., Sanz, J. M., Ayala, I., Ayesa, E., Alferes, J., and Irizar, I. (2008). "AquaScan: design and implementation of an Internet-based service for the remote monitoring and management of decentralised WWTWs." *Water Science and Technology*, 57(10), 1495-1502.
- Chaubey, I., Haan, C. T., Grunwald, S., and Salisbury, J. M. (1999). "Uncertainty in the model parameters due to spatial variability of rainfall." *Journal of Hydrology*, 220(1-2), 48-61.
- Christensen, J. H., Hewitson, B., Busuioc, A., Chen, A., Gao, X., Held, I., Jones, R., Kolli, R. K., Kwon, W.-T., Laprise, R., Magaña Rueda, V., Mearns, L., Menéndez, C. G., Räisänen, J., Rinke, A., Sarr, A., and Whetton, P. (2007). "Regional Climate Projections." *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*. Solomon, D. Qin, M. Manning, Z. Chen, M. Marquis, K. B. Averyt, M. Tignor, and H. L. Miller, eds., Cambridge University Press, Cambridge, United Kingdom.
- Ciach, G. J. (2003). "Local random errors in tipping-bucket rain gauge measurements." *Journal of Atmospheric and Oceanic Technology*, 20(5), 752-759.
- Clark, M. P., Rupp, D. E., Woods, R. A., Zheng, X., Ibbitt, R. P., Slater, A. G., Schmidt, J., and Uddstrom, M. J. (2008). "Hydrological data assimilation with the ensemble Kalman filter: Use of streamflow observations to update states in a distributed hydrological model." *Advances in Water Resources*, 31(10), 1309-1324.
- Cloke, H. L., and Pappenberger, F. (2009). "Ensemble flood forecasting: A review." *Journal of Hydrology*, 375(3-4), 613-626.
- Collier, C. G. (2007). "Flash flood forecasting: What are the limits of predictability?" *Quarterly Journal of the Royal Meteorological Society*, 133(622), 3-23.
- Collier, C. G. (2009). "On the propagation of uncertainty in weather radar estimates of rainfall through hydrological models." *Meteorological Applications*, 16(1), 35-40.
- Cutore, P., Campisano, A., Kapelan, Z., Modica, C., and Savic, D. (2008). "Probabilistic prediction of urban water consumption using the SCEM-UA algorithm." *Urban Water Journal*, 5(2), 125-132.
- Dandy, G., Nguyen, T., and Davies, C. (1997). "Estimating residential water demand in the presence of free allowances." *Land Economics*, 73(1), 125-139.

- Darsono, S., and Labadie, J. W. (2007). "Neural-optimal control algorithm for real-time regulation of in-line storage in combined sewer systems." *Environmental Modelling & Software*, 22(9), 1349-1361.
- Davidson, J., Bouchart, F., Cavill, S., and Jowitt, P. (2005). "Real-time connectivity modeling of water distribution networks to predict contamination spread." *Journal of Computing in Civil Engineering*, 19(4), 377-386.
- 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 Clercq, B., Coen, F., Vanderhaegen, B., and Vanrolleghem, P. A. (1999). "Calibrating simple models for mixing and flow propagation in waste water treatment plants." *Water Science and Technology*, 39(4), 61-69.
- 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., Ashley, R., and Rest, D. (2000). "Modelling input of fine granular sediment into drainage systems via gully-pots." *Water Research*, 34(15), 3836-3844.
- Delleur, J. W. (2003). "The evolution of urban hydrology: Past, present, and future." *Journal of Hydraulic Engineering-Asce*, 129(8), 563-573.
- Devesa, F., Comas, J., Turon, C., Freixo, A., Carrasco, F., and Poch, M. (2009). "Scenario analysis for the role of sanitation infrastructures in integrated urban wastewater management." *Environmental Modelling & Software*, 24(3), 371-380.
- Entem, S., Lahoud, A., Yde, L., and Bendsen, B. (1998). "Real time control of the sewer system of Boulogne Billancourt - A contribution to improving the water quality of the Seine." *Water Science and Technology*, 37(1), 327-332.
- Evensen, G. (1992). "Using the Extended Kalman Filter with a Multilayer Quasi-Geostrophic Ocean Model." *Journal of Geophysical Research-Oceans*, 97(C11), 17905-17924.
- Evensen, G. (1994). "Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast the error statistics." *Journal of Geophysical Research*, 99, 10,143-10,162.
- Evensen, G. (1997). "Advanced data assimilation for strongly nonlinear dynamics." *Monthly Weather Review*, 125(6), 1342-1354.
- Evensen, G. (2003). "The Ensemble Kalman Filter: theoretical formulation and practical implementation." *Ocean Dynamics*, 53, 343-367.
- Fabry, F., and Seed, A. W. (2009). "Quantifying and predicting the accuracy of radar-based quantitative precipitation forecasts." *Advances in water Resources*, 32, 1043-1049.
- Faures, J. M., Goodrich, D. C., Woolhiser, D. A., and Sorooshian, S. (1995). "Impact of Small-Scale Spatial Rainfall Variability on Runoff Modeling." *Journal of Hydrology*, 173(1-4), 309-326.
- Fischer, A., Rouault, P., Kroll, S., Van Assel, J., and Pawlowsky-Reusing, E. (2009). "Possibilities of sewer model simplifications." *Urban Water Journal*, 6(6), 457-470.

- 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.
- Freni, G., Mannina, G., and Viviani, G. (2010). "Emission standards versus immission standards for assessing the impact of urban drainage on ephemeral receiving water bodies." *Water Science and Technology*, 61(6), 1617-1629.
- Friedler, E., Brown, D. M., and Butler, D. (1996). "A study of WC derived sewer solids." *Water Science and Technology*, 33(9), 17-24.
- Froukh, M. L. (2001). "Decision-support system for domestic water demand forecasting and management." *Water Resources Management*, 15(6), 363-382.
- Fu, G., Butler, D., and Khu, S.-T. (2008). "Comparison of control strategies for multi-objective control of urban wastewater systems." iEMSs 2008: International Congress on Environmental Modelling and Software: Integrated Sciences and Information Technology for Environmental Assessment and Decision Making, 4th Biennial Meeting, M. Sanchez-Marre, J. Bejar, J. Comas, A. Rizzoli, and G. Guariso, eds., Barcelona, Catalonia, 1347-1353.
- Fuchs, L., and Beeneken, T. (2004). "Development and Implementation of a Real-Time Control Strategy for the Sewer System of the Vienna City." International Conference on Urban Drainage Modelling, Dresden, 229-307.
- Fuchs, L., and Beeneken, T. (2005). "Development and implementation of a real-time control strategy for the sewer system of the city of Vienna." *Water Science and Technology*, 52(5), 187-194.
- Fulford, J. M., and Gonzalez-Castro, J. (2009). "Laser Water-Level Measurement Uncertainty." 33rd IAHR Congress: water Engineering for a Sustainable Environment.
- Gato, S., Jayasuriya, N., and Roberts, P. (2007). "Temperature and rainfall thresholds for base use urban water demand modelling." *Journal of Hydrology*, 337(3-4), 364-376.
- Gelfan, A., Hajda, P., and Novotny, V. (1999). "Recursive System Identification for Real-Time Sewer Flow Forecasting." *Journal of Hydrologic Engineering*, 4(3), 280-287.
- 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., Savic, D., and Kapelan, Z. (2008). "Pressure-driven demand and leakage simulation for water distribution networks." *Journal of Hydraulic Engineering-Asce*, 134(5), 626-635.
- Giustolisi, O., and Todini, E. (2009). "Pipe hydraulic resistance correction in WDS analysis." *Urban Water Journal*, 6(1), 39-52.
- Golding, B. W. (1998). "Nimrod: a system generating automatic very-short-range forecasts." *Meteorological Applications*, 5, 1-16.
- Golding, B. W. (2000). "Quantitative precipitation forecasting in the UK." *Journal of Hydrology*, 239(1-4), 286-305.

- Goovaerts, P. (2000). "Geostatistical approaches for incorporating elevation into the spatial interpolation of rainfall." *Journal of Hydrology*, 228(1-2), 113-129.
- Goulsby, B., and Samuels, P. (2005). "Language of Risk: Project Definitions." Report: T32-04-01, www.floodsite.net.
- Gouweleeuw, B. T., Thielen, J., Franchello, G., De Roo, A. P. J., and Buizza, R. (2005). "Flood forecasting using medium-range probabilistic weather prediction." *Hydrology and Earth System Sciences*, 9(4), 365-380.
- Groffman, P., Baron, J., Blett, T., Gold, A., Goodman, I., Gunderson, L., Levinson, B., Palmer, M., Paerl, H., Peterson, G., Poff, N., Rejeski, D., Reynolds, J., Turner, M., Weathers, K., and Wiens, J. (2006). "Ecological thresholds: The key to successful environmental management or an important concept with no practical application?" *Ecosystems*, 9(1), 1-13.
- Gruber, G., Bertrand-Krajewski, J. L., De Benedetis, J., Hochedlinger, M., and Lettl, W. (2005). "Practical aspects, experiences and strategies by using UV-VIS sensors for long-term sewer monitoring." *10th international conference on urban drainage*, Copenhagen, Denmark.
- Hall, J. W. (2003). "Handling uncertainty in the hydroinformatic process." *Journal of Hydroinformatics*, 5, 215-232.
- Hansen, J. A., and Smith, L. A. (2001). "Probabilistic noise reduction." *Tellus Series a-Dynamic Meteorology and Oceanography*, 53(5), 585-598.
- Hargreaves, J. C., Annan, J. D., Edwards, N. R., and Marsh, R. (2004). "An efficient climate forecasting method using an intermediate complexity Earth System Model and the ensemble Kalman filter." *Climate Dynamics*, 23(7-8), 745-760.
- Harremoes, P., and Rauch, W. (1999). "Optimal design and real time control of the integrated urban run-off system." *Hydrobiologia*, 410, 177-184.
- Helton, J. C., and Burmaster, D. E. (1996). "Guest editorial: Treatment of aleatory and epistemic uncertainty in performance assessments for complex systems." *Reliability Engineering & System Safety*, 54(2-3), 91-94.
- 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.
- Hirschi, R., Dupont, P., Avellan, F., Favre, J. N., Guelich, J. F., and Parkinson, E. (1998). "Centrifugal pump performance drop due to leading edge cavitation: Numerical predictions compared with model tests." *Journal of Fluids Engineering-Transactions of the Asme*, 120(4), 705-711.
- 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.
- Houtekamer, P. L., Mitchell, H. L., Pellerin, G., Buehner, M., Charron, M., Spacek, L., and Hansen, M. (2005). "Atmospheric data assimilation with an ensemble Kalman filter: Results with real observations." *Monthly Weather Review*, 133(3), 604-620.
- Hutton, C., Vamakeridou-Lyroudia, L. S., Kapelan, Z., and Savic, D. (2011). "Uncertainty Quantification and Reduction in Urban Water Systems"

- Modelling: Evaluation Report, PREPARED: Enabling Change, Work Package 3.6, European Commission 7th Framework Programme.", University of Exeter (UNEXE).
- Ide, K., Courtier, P., Ghil, M., and Lorenc, A. C. (1997). "Unified notation for data assimilation: Operational, sequential and variational." *Journal of the Meteorological Society of Japan*, 75(1B), 181-189.
- Ingeduld, P., and Turton, G. (2002). "Online analysis of water supply and water distribution systems." NZWWA Annual Modelling Conference, Auckland, New Zealand, .
- Jack, A. G., Petrie, M. M., and Ashley, R. M. (1996). "The diversity of sewer sediments and the consequences for sewer flow quality modelling." *Water Science and Technology*, 33(9), 207-214.
- Jain, A., Varshney, A. K., and Joshi, U. C. (2001). "Short-term Water Demand Forecasting Modelling at ITT Kanpur Using Artificial Neural Networks." *Water Resources Management*, 15, 299-321.
- Jamieson, D. G., Shamir, U., Martinez, F., and Franchini, M. (2007). "Conceptual design of a generic, real-time, near-optimal control system for water-distribution networks." *Journal of Hydroinformatics*, 9(1), 3-14.
- Jang, Y. (2010). "Topographic Factors in Precipitation Dynamics Affecting Water Resource Engineering in the Contiguous U.S.: Notes from Hydroclimatic Studies." *World Environmental and Water Resources Congress 2010: Challenges of Change: Proceeding of the World Environmental and Water Resources Congress. ASCE Conference Proceedings*, 371, 19.
- Jauregui, E., and Romales, E. (1996). "Urban effects on convective precipitation in Mexico city." *Atmospheric Environment*, 30(20), 3383-3389.
- Jeong, D. I., and Kim, Y. O. (2005). "Rainfall-runoff models using artificial neural networks for ensemble streamflow prediction." *Hydrological Processes*, 19(19), 3819-3835.
- Joannis, C., Ruban, G., Gromaire, M. C., Bertrand-Krajewski, J. L., and Chebbo, G. (2008). "Reproducibility and uncertainty of wastewater turbidity measurements." *Water Science and Technology*, 57(10), 1667-1673.
- 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.
- 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 Muttill, N. (2001). "Genetic programming and its application in real-time runoff forecasting." *Journal of the American Water Resources Association*, 37(2), 439-451.

- Komma, J., Bloschl, G., and Reszler, C. (2008). "Soil moisture updating by Ensemble Kalman Filtering in real-time flood forecasting." *Journal of Hydrology*, 357(3-4), 228-242.
- Korving, H., van Noordwijk, 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.
- 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.
- Larrarte, F. (2006). "Velocity fields within sewers: An experimental study." *Flow Measurement and Instrumentation*, 17(5), 282-290.
- Larrarte, F., Bardiaux, J. B., Battaglia, P., and Joannis, C. (2008). "Acoustic Doppler flow-meters: A proposal to characterize their technical parameters." *Flow Measurement and Instrumentation*, 19(5), 261-267.
- Larrarte, F., and Chanson, H. (2008). "Challenges in sewer hydrodynamics, sediment and measurements." *International meeting on measurments and hydraulics in sewers, Hydraulic model report no. CH70/08*, Brisbane, Australia.
- Lau, J., Butler, D., and Schutze, M. (2002). "Is combined sewer overflow spill frequency/volume a good indicator of recieving water quality impact." *Urban Water*, 4(2), 181-189.
- Lawson, W. G., and Hansen, J. A. (2004). "Implications of stochastic and deterministic filters as ensemble-based data assimilation methods in varying regimes of error growth." *Monthly Weather Review*, 132(8), 1966-1981.
- Lekkas, D. F. (2008). "Using complementary methods for improved flow forecasting." *Hydrological Sciences Journal-Journal Des Sciences Hydrologiques*, 53(4), 696-705.
- 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.
- Li, Z. J., and Navon, I. M. (2001). "Optimality of variational data assimilation and its relationship with the Kalman filter and smoother." *Quarterly Journal of the Royal Meteorological Society*, 127(572), 661-683.
- Liguori, S., Rico-Ramirez, M., Schellart, A., and Saul, A. (2011). "High-resolution ensemble rainfall forecasts in urban hydrological applications." *Geophysical Research Abstracts*, Vienna, Austria.
- 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.

- Liu, Y. Q., and Gupta, H. V. (2007). "Uncertainty in hydrologic modeling: Toward an integrated data assimilation framework." *Water Resources Research*, 43(7), -.
- Machell, J., Mounce, S. R., and Boxall, J. B. (2009). "Online modelling of water distribution systems: a UK case study." *Drinking Water Engineering and Science Discussions*, 2, 279-294.
- Madsen, H., and Canizares, R. (1999). "Comparison of extended and ensemble Kalman filters for data assimilation in coastal area modelling." *International Journal for Numerical Methods in Fluids*, 31(6), 961-981.
- 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.
- Mallick, K. N., Ahmed, I., Tickle, K. S., and Lansey, K. E. (2002). "Determining pipe groupings for water distribution networks." *Journal of Water Resources Planning and Management-Asce*, 128(2), 130-139.
- Mancarella, D., Babovic, V., Keijzer, M., and Simeone, V. (2008). "Data assimilation of forecasted errors in hydrodynamic models using inter-model correlations." *International Journal For Numerical Methods in Fluids*, 56, 597-605.
- Maribas, A., da Silva, M. D. L., Laurent, N., Loison, B., Battaglia, P., and Pons, M. N. (2008). "Monitoring of rain events with a submersible UV/VIS spectrophotometer." *Water Science and Technology*, 57(10), 1587-1593.
- Martinez, F., Hernandez, V., Alonso, J. M., Rao, Z. F., and Alvisi, S. (2007). "Optimizing the operation of the Valencia water-distribution network." *Journal of Hydroinformatics*, 9(1), 65-78.
- Matgen, P., Montanari, M., Hostache, R., Pfister, L., Hoffmann, L., Plaza, D., Pauwels, V. R. N., De Lannoy, G. J. M., De Keyser, R., and Savenije, H. H. G. (2010). "Towards the sequential assimilation of SAR-derived water stages into hydraulic models using the Particle Filter: proof of concept." *Hydrology and Earth System Sciences*, 14(9), 1773-1785.
- Meirlaen, J., Huyghebaert, B., Sforzi, F., Benedetti, L., and Vanrolleghem, P. (2001). "Fast, simultaneous simulation of the integrated urban wastewater system using mechanistic surrogate models." *Water Science and Technology*, 43(7), 301-309.
- Mejia, A. I., and Moglen, G. E. (2010). "Impact of the spatial distribution of imperviousness on the hydrologic response of an urbanizing basin." *Hydrological Processes*, 24(23), 3359-3373.
- 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.
- Messmer, A., Schutze, M., and Ogurek, M. (2008). "A Demonstration Software Tool For Real Time Control of Urban Drainage Systems." *11th International Conference on Urban Drainage*, Edinburgh, Scotland, UK.
- Miller, R. N., Ghil, M., and Gauthiez, F. (1994). "Advanced Data Assimilation in Strongly Nonlinear Dynamical-Systems." *Journal of the Atmospheric Sciences*, 51(8), 1037-1056.
- Moradkhani, H., Hsu, K. L., Gupta, H., and Sorooshian, S. (2005a). "Uncertainty assessment of hydrologic model states and parameters:

- Sequential data assimilation using the particle filter." *Water Resources Research*, 41(5), -.
- Moradkhani, H., Sorooshian, S., Gupta, H. V., and Houser, P. R. (2005b). "Dual state-parameter estimation of hydrological models using ensemble Kalman filter." *Advances in Water Resources*, 28(2), 135-147.
- Morin, E., Goodrich, D. C., Maddox, R. A., Gao, X. G., Gupta, H. V., and Sorooshian, S. (2006). "Spatial patterns in thunderstorm rainfall events and their coupling with watershed hydrological response." *Advances in Water Resources*, 29(6), 843-860.
- MOUSE. (2004). "MOUSE: Pipe Flow Reference Manual. DHI software 2004."
- Munavalli, G. R., and Kumar, M. S. M. (2005). "Water quality parameter estimation in a distribution system under dynamic state." *Water Research*, 39(18), 4287-4298.
- 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.
- Nasseri, M., Moeini, A., and Tabesh, M. (2011). "Forecasting monthly urban water demand using Extended Kalman filter and Genetic Programming." *Expert Systems with Applications*, 38, 7387-7395.
- Neal, J. C., Atkinson, P. M., and Hutton, C. W. (2007). "Flood inundation model updating using an ensemble Kalman filter and spatially distributed measurements." *Journal of Hydrology*, 336(3-4), 401-415.
- Nielsen, P. H., Raunkjaer, K., Norsker, N. H., Jensen, N. A., and Hvitvedjacobsen, T. (1992). "Transformation of Waste-Water in Sewer Systems - a Review." *Water Science and Technology*, 25(6), 17-31.
- Ocampo Martinez, C. A. (2007). "Model Predictive Control of Complex Systems including Fault Tolerance Capabilities: Application to Sewer Networks," Technical University of Catalonia, Terrassa.
- Ofwat. (2005). "Water Framework Directive - Economic Analysis of Water Industry Costs. Final Report." Ofwat, UK. Ove Arup and Partners Ltd, In association with Oxera Consulting Ltd.
- Ogden, F. L., and Julien, P. Y. (1994). "Runoff Model Sensitivity to Radar Rainfall Resolution." *Journal of Hydrology*, 158(1-2), 1-18.
- Ogden, S., and Watson, R. (1999). "Corporate performance and stakeholder management: Balancing shareholder and customer interests in the UK privatized water industry." *Academy of Management Journal*, 42(5), 526-538.
- Ott, E., Hunt, B. R., Szunyogh, I., Zimin, A. V., Kostelich, E. J., Corazza, M., Kalnay, E., Patil, D. J., and Yorke, J. A. (2004). "A local ensemble Kalman filter for atmospheric data assimilation." *Tellus Series a-Dynamic Meteorology and Oceanography*, 56(5), 415-428.
- Pahl-Wostl, C. (2005). "Information, public empowerment, and the management of urban watersheds." *Environmental Modelling & Software*, 20(4), 457-467.
- Pahl-Wostl, C. (2007). "The implications of complexity for integrated resources management." *Environmental Modelling & Software*, 22(5), 561-569.

- Pahl-Wostl, C., Sendzimir, J., Jeffrey, P., Aerts, J., Berkamp, G., and Cross, K. (2007). "Managing change toward adaptive water management through social learning." *Ecology and Society*, 12(2), -.
- Panguluri, S., Haught, R. C., Patterson, C. L., Krishnan, E. R., and Hall, J. (2005). "Real-time Remote Monitoring Of Drinking Water Quality." *Proceedings of the 2005 World Water and Environmental Resources Congress, May 15-19 2005, Anchorage, Alaska*.
- 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.
- Pasha, M. F. K., and Lansey, K. E. (2005). "Analysis of uncertainty on water distribution hydraulics and water quality." *Proceedings of the ASCE world Water and Environmental Resources Congress, May 15-19 2005, Anchorage, Alaska, USA*.
- Patterson, C. L., Haught, R. C., Goodrich, J. A., Impellitteri, C. A., and Moreno, H. E. (2005). "An Overview of U.S. EPA Research on Remote Monitoring and Control Technologies for Small Drinking Water Treatment Systems." *Proceedings of the 2005 World Water and Environmental Resources Congress, May 15-19, 2005., Anchorage, Alaska*.
- Pauwels, V. R. N., and De Lannoy, G. J. M. (2006). "Improvement of modeled soil wetness conditions and turbulent fluxes through the assimilation of observed discharge." *Journal of Hydrometeorology*, 7(3), 458-477.
- Penning-Rowsell, E. C., Tunstall, S. M., Tapsell, S. M., and Parker, D. J. (2000). "The benefits of flood warnings: Real but elusive, and politically significant." *Journal of the Chartered Institution of Water and Environmental Management*, 14(1), 7-14.
- Pham, D. T. (2001). "Stochastic methods for sequential data assimilation in strongly nonlinear systems." *Monthly Weather Review*, 129(5), 1194-1207.
- Pleau, M., Colas, H., Lavallee, P., Pelletier, G., and Bonin, R. (2005). "Global optimal real-time control of the Quebec urban drainage system." *Environmental Modelling & Software*, 20(4), 401-413.
- Quevedo, J., Puig, V., Cembrano, G., Blanch, J., Aguilar, J., Saporta, D., Benito, G., Hedo, M., and Molina, A. (2010). "Validation and reconstruction of flow meter data in the Barcelona water distribution network." *Control Engineering Practice*, 18(6), 640-651.
- Rabier, F., Klinker, E., Courtier, P., and Hollingsworth, A. (1996). "Sensitivity of forecast errors to initial conditions." *Quarterly Journal of the Royal Meteorological Society*, 122(529), 121-150.
- Rao, Z. F., and Salomons, E. (2007). "Development of a real-time, near-optimal control process for water-distribution networks." *Journal of Hydroinformatics*, 9(1), 25-37.
- Rao, Z. F., Wicks, J., and West, S. (2007). "Optimising water supply and distribution operations." *Proceedings of the Institution of Civil Engineers-Water Management*, 160(2), 95-101.

- 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., and Harremoes, P. (1999). "Genetic algorithms in real time control applied to minimize transient pollution from urban wastewater systems." *Water Research*, 33(5), 1265-1277.
- Rauch, W., Vanhooren, H., and Vanrolleghem, P. A. (1999). "A simplified mixed-culture biofilm model." *Water Research*, 33(9), 2148-2162.
- Refsgaard, J. C. (1997). "Validation and intercomparison of different updating procedures for real-time forecasting." *Nordic Hydrology*, 28(2), 65-84.
- Reichert, P., Borchardt, D., Henze, M., Rauch, W., Shanahan, P., Somlyódy, L., and Vanrolleghem, P. (2001). "River water quality model no. 1 (RWQM1): II. Biochemical process equations." *Water Science and Technology*, 43(5), 11-30.
- Reichert, P., and Vanrolleghem, P. (2001). "Identifiability and uncertainty analysis of the River Water Quality Model No. 1 (RWQM1)." *Water Science and Technology*, 43(7), 329-338.
- 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, -.
- Revelli, R., and Ridolfi, L. (2002). "Fuzzy approach for analysis of pipe networks." *Journal of Hydraulic Engineering-Asce*, 128(1), 93-101.
- Rico-Ramirez, M. A., Schellart, A. N. A., Liguori, S., and Saul, A. J. (2009). "Quantitative Precipitation Forecasting For a Small Urban Area: Use of a high-resolution numerical weather prediction model." *8th International Conference on Precipitation in Urban Areas, 10-13 December, St. Moritz, Switzerland*.
- Rieger, L., Langergraber, G., Kaelin, D., Siegrist, H., and Vanrolleghem, P. A. (2008). "Long-term evaluation of a spectral sensor for nitrite and nitrate." *Water Science and Technology*, 57(10), 1563-1569.
- Roberts, N. M., Cole, S. J., Forbes, R. M., Moore, R. J., and Boswell, D. (2009). "Use of high-resolution NWP rainfall and river flow forecasts for advance warning of the Carlisle flood, north-west England." *Meteorological Applications*, 16, 23-34.
- Roberts, N. M., and Lean, H. W. (2008). "Scale-selective verification of rainfall accumulations from high-resolution forecasts of convective events." *Monthly Weather Review*, 136(1), 78-97.
- 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.
- Romano, M., Kapelan, Z., and Savic, D. A. (2010). "Bayesian-based online burst detection in water distribution systems." *Integrating Water*

Systems: Proceeding of the Tenth International Conference on Computing and Control in the Water Industry.

- Romero-Gomez, P., Ho, C. K., and Choi, C. Y. (2008). "Mixing at cross junctions in water distribution systems. I: Numerical study." *Journal of Water Resources Planning and Management-Asce*, 134(3), 285-294.
- Rosenfeld, D. (2000). "Suppression of rain and snow by urban and industrial air pollution." *Science*, 287(5459), 1793-1796.
- Rossman, L. A. (2000). *EPANET2 Users Manual* National Risk Management Research Laboratory, U.S. Environmental Protection Agency, http://www.image.unipd.it/salandin/IngAmbientale/Progetto_2/EPANET/EN2manual.pdf, Cincinnati, OH 45268, USA.
- 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.
- Sadiq, R., Kleiner, Y., and Rajani, B. (2006). "Estimating risk of contaminant intrusion in water distribution networks using Dempster-Shafer theory of evidence." *Civil Engineering and Environmental Systems*, 23(3), 129-141.
- Salamon, P., and Feyen, L. (2009). "Assessing parameter, precipitation, and predictive uncertainty in a distributed hydrological model using sequential data assimilation with the particle filter." *Journal of Hydrology*, 376(3-4), 428-442.
- Salomons, E., Goryashko, A., Shamir, U., Rao, Z. F., and Alvisi, S. (2007). "Optimizing the operation of the Haifa-A water-distribution network." *Journal of Hydroinformatics*, 9(1), 51-64.
- Sannasiraj, S. A., Babovic, V., and Chan, E. S. (2005). "Local model approximation in the real time wave forecasting." *Coastal Engineering*, 52(3), 221-236.
- 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.
- Schaefli, B., Talamba, D. B., and Musy, A. (2007). "Quantifying hydrological modeling errors through a mixture of normal distributions." *Journal of Hydrology*, 332(3-4), 303-315.
- Schellart, A. N. A., Rico-Ramirez, M. A., Liguori, S., and Saul, A. J. (2009). "Quantitative Precipitation Forecasting For a Small Urban Area: Use of Radar Nowcasting." *8th International Workshop on Precipitation in Urban Areas, 10-13 December 2009, St. Moritz, Switzerland.*
- Schilling, W., and Martens, J. (1986). "Recursive state and parameter estimation with applications in water resources." *Applied Mathematical Modelling*, 10(6), 433-437.
- 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, -.

- Schroeder, K., and Pawlowsky-Reusing, E. (2005). "Current state and development of the real-time control of the Berlin sewage system." *Water Science and Technology*, 52(12), 181-187.
- Schutze, M., Campisano, A., Colas, H., Schilling, W., and Vanrolleghem, P. A. (2004). "Real time control of urban wastewater systems - where do we stand today?" *Journal of Hydrology*, 299(3-4), 335-348.
- Schutze, M., Erbe, V., Haas, U., Scheer, M., and Weyand, M. (2008). "Sewer system real-time control supported by the M180 guideline document." *Urban Water Journal*, 5(1), 67-76.
- Segond, M. L., Wheater, H. S., and Onof, C. (2007). "The significance of spatial rainfall representation for flood runoff estimation: A numerical evaluation based on the Lee catchment, UK." *Journal of Hydrology*, 347(1-2), 116-131.
- Seo, D. J., Koren, V., and Cajina, N. (2003). "Real-time variational assimilation of hydrologic and hydrometeorological data into operational hydrologic forecasting." *Journal of Hydrometeorology*, 4(3), 627-641.
- 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 Nesper, V. (1998). "Empirical and theoretical assessment of the wind induced error of rain measurement." *Water Science and Technology*, 37(11), 171-178.
- Sevruk, B., and Nevenic, M. (1998). "The geography and topography effects on the areal pattern of precipitation in a small prealpine basin." *Water Science and Technology*, 37(11), 163-170.
- Shamseldin, A. Y., and O'Connor, K. M. (2001). "A non-linear neural network technique for updating of river flow forecasts." *Hydrology and Earth System Sciences*, 5(4), 577-597.
- Shang, F., Uber, J., Murray, R., and Janke, R. (2008). "Model-based real-time detection of contamination events." *Proceedings of the 10th Annual Water Distribution Systems Analysis Conference*, Kruger National Park, South Africa.
- 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 Annual Water Distribution Systems Analysis Symposium*, Cincinnati, Ohio, USA.
- Shin, P. S., Song, Y. R., Choi, Y. J., and Park, Y. S. (2009). "Seoul(Korea) Online Water Quality Monitoring Of Drinking Water." *World Environmental and Water Resources Congress*, Kansas City, Missouri.
- 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 WWTW model applications: A

- critical discussion using an example from design." *Water Research*, 43(11), 2894-2906.
- Smith, P. J., Beven, K. J., and Tawn, J. A. (2008). "Detection of structural inadequacy in process-based hydrological models: A particle-filtering approach." *Water Resources Research*, 44(1), -.
- Snyder, C., Bengtsson, T., Bickel, P., and Anderson, J. (2008). "Obstacles to High-Dimensional Particle Filtering." *Monthly Weather Review*, 136(12), 4629-4640.
- Stinson, M. K., Field, R., Jacquet, G., and Villeneuve, E. (2000). "Optimal Real-Time Control of Combined Sewerage Systems: Two Case Studies." *Building Partnerships: Proceedings of Joint Conference on Water Resource Engineering and Water Resources Planning and Management*, doi:10.1061/40517(2000)41.
- Stoianov, I., Nachman, L., Whittle, A., Madden, S., and Lking, R. (2006). "Sensor networks for monitoring water supply and sewer systems: lessons from Boston." 8th Annual Water Distribution Systems Analysis Symposium, Cincinnati, Ohio, USA.
- 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.
- Stumwahrer, K., Matsche, N., and Winkler, S. (2003). "Influence of changes of the wastewater composition on the applicability of UV-absorption measurements at combined sewer overflows." *Water Science and Technology*, 47(2), 73-78.
- 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.
- Thorndahl, S., Beven, K. J., Jensen, J. B., and Schaarup-Jensen, K. (2008). "Event based uncertainty assessment in urban drainage modelling, applying the GLUE methodology." *Journal of Hydrology*, 357(3-4), 421-437.
- Thorndahl, S., Rasmussen, M. R., Grum, M., and Neve, S. L. (2009). "Radar based flow and water level forecasting in sewer systems - a danish case study." *8th International workshop on precipitation in urban areas, 10-13 December 2009, St Moritz, Switzerland*.
- Thorndahl, S., Rasmussen, M. R., Nielsen, J. E., and Larsen, J. B. (2010). "Uncertainty in nowcasting rainfall - a case study of the GLUE methodology." *ERAD 2010 - The Sixth european Conference on Radar in Meteorology and Hydrology*, Sibiu, Romania.
- 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.
- Todini, E. (1999). "Using a kalman filter approach for looped water distribution network calibration." *Water industry systems: modelling and optimization applications*, Volume 1., D. A. Savic and G. A. Walters, eds., Research Studies Press, Baldock, Hertfordshire, England, 327-336.
- Todini, E. (2004). "Role and treatment of uncertainty in real-time flood forecasting." *Hydrological Processes*, 18(14), 2743-2746.
- USEPA. (1995a). "United States Environmental Protection Agency, Combined Sewer Overflows - Guidance For Nine Minimum Controls. Report EPA 832-B-95-003."
- USEPA. (1995b). "United States Environmental Protection Agency. Combined Sewer Overflows: Guidance For Long-Term Control Plan. Report EPA 832-B-95-002."
- Vaes, G., and Berlamont, J. (1999). "Emission predictions with a multi-linear reservoir model." *Water Science and Technology*, 39(2), 9-16.
- Van Horne, M. P., Vivoni, E. R., Entekhabi, D., Hoffman, R. N., and Grassotti, C. (2006). "Evaluating the effects of image filtering in short-term radar rainfall forecasting for hydrological applications." *Meteorological Applications*, 13(3), 289-303.
- van Leeuwen, P. J. (2009). "Particle Filtering in Geophysical Systems." *Monthly Weather Review*, 137(12), 4089-4114.
- 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.
- Verlaan, M., and Heemink, A. W. (1997). "Tidal flow forecasting using reduced rank square root filters." *Stochastic Hydrology and Hydraulics*, 11(5), 349-368.
- Verworn, H. R., and Kramer, A. (2005). "Aspects and Effectiveness of real-time control in urban drainage systems combining radar rainfall forecasts, linear optimisation and hydrodynamic modelling." *Computers and Control in the Water Industry (CCWI 2005)*, Exeter, UK.
- 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.
- Vojinovic, Z., Kecman, V., and Babovic, V. (2003). "Hybrid approach for modeling wet weather response in wastewater systems." *Journal of Water Resources Planning and Management-Asce*, 129(6), 511-521.
- Vossepoel, F. C., and van Leeuwen, P. J. (2007). "Parameter estimation using a particle method: Inferring mixing coefficients from sea level observations." *Monthly Weather Review*, 135(3), 1006-1020.
- Vrugt, J. A., Gupta, H. V., and Nualain, B. O. (2006). "Real-time data assimilation for operational ensemble streamflow forecasting." *Journal of Hydrometeorology*, 7(3), 548-565.

- Vrugt, J. A., and Robinson, B. A. (2007). "Treatment of uncertainty using ensemble methods: Comparison of sequential data assimilation and Bayesian model averaging." *Water Resources Research*, 43(1), -.
- Wagener, T., McIntyre, N., Lees, M. J., Wheater, H. S., and Gupta, H. V. (2003). "Towards reduced uncertainty in conceptual rainfall-runoff modelling: Dynamic identifiability analysis." *Hydrological Processes*, 17(2), 455-476.
- Walski, T., Chase, D. V., Savic, D. A., Grayman, W. M., Beckwith, S., and Koelle, E. (2003). *Advanced Water Distribution Modeling and Management*, Haestad Methods Press, Waterbury, Connecticut.
- Weerts, A. H., and el Serafy, G. (2005). "Particle Filtering and Ensemble Kalman Filtering For Input Correction in Rainfall Runoff Modelling." *International Conference on Innovation Advances And Implementation of Flood Forecasting Technology*, Tromso, Norway.
- Weywand, M. (2002). "Real-time control in combined sewer systems in Germany- some case studies." *Urban Water*, 4, 347-354.
- Whitaker, J. S., and Hamill, T. M. (2002). "Ensemble data assimilation without perturbed observations." *Monthly Weather Review*, 130(7), 1913-1924.
- WHO. (2009). "Summary and policy implications Vision 2030: the resilience of water supply and sanitation in the face of climate change." WHO press, Geneva, Switzerland.
- Willems, P. (1999). "Stochastic generation of spatial rainfall for urban drainage areas." *Water Science and Technology*, 39(9), 23-30.
- Willems, P. (2008). "Quantification and relative comparison of different types of uncertainties in sewer water quality modeling." *Water Research*, 42(13), 3539-3551.
- Willems, P., and Berlamont, J. (1998). "Stochastic modelling of spatial rain cells." *Hydrology in a Changing Environment*, H. Wheather and C. Kirby, eds., John Wiley and sons, Chichester, England, 307-318.
- 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.
- Xie, X. H., and Zhang, D. X. (2010). "Data assimilation for distributed hydrological catchment modeling via ensemble Kalman filter." *Advances in Water Resources*, 33(6), 678-690.
- Yang, J., Reichert, P., Abbaspour, K. C., and Yang, H. (2007). "Hydrological modelling of the chaohe basin in china: Statistical model formulation and Bayesian inference." *Journal of Hydrology*, 340(3-4), 167-182.
- 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), -.
- Zealand, C. M., Burn, D. H., and Simonovic, S. P. (1999). "Short term streamflow forecasting using artificial neural networks." *Journal of Hydrology*, 214(1-4), 32-48.
- Zhang, H., and Pu, Z. (2010). "Beating the uncertainties: Ensemble forecasting and ensemble based data assimilation (review article)." *Advances in Meteorology*, 2010, Article ID 432160 10pp.

- Zhang, J., Song, R., Bhaskar, N. R., and French, M. N. (2007). "Ensemble Forecasting of Daily Water Demand." *World Environmental and Water Resources Congress 2007: Restoring our Natural Habitat*, Tampa, Florida, U.S.A.
- Zhang, X. S., Liang, F. M., Srinivasan, R., and Van Liew, M. (2009a). "Estimating uncertainty of streamflow simulation using Bayesian neural networks." *Water Resources Research*, 45, -.
- Zhang, X. S., Srinivasan, R., and Bosch, D. (2009b). "Calibration and uncertainty analysis of the SWAT model using Genetic Algorithms and Bayesian Model Averaging." *Journal of Hydrology*, 374(3-4), 307-317.
- Zhao, H., Wu, C., Gao, J., Wu, W., Liu, Y., Wang, Q., and Guan, W. (2005). "Study And Applications of on-line optimal operation in typical urban water supply system of China." *Computers and Control in the Water Industry (CCWI, 2005)*, Exeter , UK.
- Zhou, S. L., McMahon, T. A., Walton, A., and Lewis, J. (2002). "Forecasting operational demand for an urban water supply zone." *Journal of Hydrology*, 259(1-4), 189-202.
- Zhou, Y. H., McLaughlin, D., and Entekhabi, D. (2006). "Assessing the performance of the ensemble Kalman filter for land surface data assimilation." *Monthly Weather Review*, 134(8), 2128-2142.

7 Appendix A: Tabular Classification of Uncertainty Methodologies

The Table is an extended version of the methodologies reviewed in Hutton et al. (2011)

Reference	Sampling/ Optimisation Method/Model	Parameter Uncertainty	Structural Uncertainty	Input/Data Uncertainty	Output Uncertainty	State Uncertainty	Notes/Assumptions
Calibration Techniques							
Optimisation techniques (Savic et al. 2009)							
(Savic et al. 2009) review paper	GA; GN; GB; SA	Identify optimal parameter set.	-	-	minimised	-	Reduction of parameter uncertainty. No quantification of uncertainty.
FOSM							
(Kang and Lansey 2009; Lansey et al. 2001)	-	Mean and Variance	-	ND	PD from FOSM	-	Assumes linear approximation of model function and Gaussianity, and requires assumed posterior error model. ND data errors assumed
Formal Bayesian Approaches							
(Kapelan et al. 2007)	SCEM-UA	PPDF	EDF	EDF	PPDF	-	Assumed EDF error model.
(Freni and Mannina 2010)	MCS	PPDF	ND	ND	PPDF; PB	-	Assumed ND error model; Box-Cox transformation.
(Schaeffli et al. 2007)	M-H MCMC	PPDF	NMD	NMD	PB	-	AR model; NMD parameters calibrated; PD checks.
(Yang et al. 2007)	M-H MCMC	PPDF	ND	ND	PB	-	Assumed ND error model; Box-Cox transformation; AR Model; PD checks; calibrated error parameters.
(Willems 2008)	MCS	Separate	Inferred from	ERM	Total PB.	-	Parameters inferred separately from

		Calibration	VD				structural and input error.
(Schoups and Vrugt 2010)	DREAM-ZS	PPDF	SEP; BF; SD	SEP; BF; SD	PB	-	AR model; SD, BF and SEP parameters calibrated as function of flow magnitude and account implicitly for all errors; PD checks.
(Thyer et al. 2009)	MCMC	PPDF	Not explicit	DRM; ERM; HDM for output error	Parameter PB; Total PB.	-	PD checks.
(Renard et al. 2010)	MCMC		SP	DRM; HDM for output error.	Total PB.	-	PD checks. Difficulty of separating sources of error (input from structural) without sufficient prior information.
(Zhang et al. 2009b)	GA	Identify optimal parameter set.	BMA	-	BMA prediction bounds.	-	Assumes that different models cover all Structural error; no parameter uncertainty.
Informal 'Pseudo' Bayesian Approaches							
(Thorndahl et al. 2008)	MCS	PPDF	Implicit in IEDF	-	PB.	-	IEDF likelihood for parameter uncertainty; assumed likelihood function, behavioural threshold.
(Liu et al. 2009)	MCS	PPDF	Implicit	Implicit Input; Output RCEB.	PB	-	Structural and Input error inferred from non-stationary output; likelihood based on output RCEB.
Possibility Theory and Fuzzy Approaches							
(Revelli and Ridolfi 2002)	GN search of each α -cut	FMF	-	-	FMF	-	Output possibility distribution based on parameter uncertainty only.
(Branisavljevic et al. 2009)	GA search of each α -cut	FMF	-	-	FMF	-	Output possibility distribution based on parameter uncertainty only.
Evidence Theory							
(Sadiq et al. 2006)	DS	BPA	-	-	-	-	No formal model, but combination of evidence to produce Belief and Plausibility functions.

Real-Time approaches							
Kalman Filtering (Evensen 2003)							
(Todini 1999)	KF	Optimised by KF	-	ND	-	-	Assumed ND error model for demand observation errors. Model Linearity required.
(Kang and Lansey 2009)	MCS; LHS	PPDF	Implicit	ND errors in KF	PB	ND errors in KF	MCs used to evaluate LHS and FOSM; KF limited in nonlinear WDS networks.
(Shang et al. 2006)	EKF	A priori parameters	Implicit	Output errors MP;	% error of total	ND errors in EKF	Requires Linear adjoint model; Assumes Gaussianity of errors.
(Bechmann et al. 1999)	EKF; QN	Mean and Variance	Implicit	ND errors	R ² error	ND errors in EKF	Model Linearity assumed.
(Clark et al. 2008)	EnKF; EnSRF	A priori parameters	Implicit	Output errors MP; Input parameters	PB from ensemble	State perturbation parameters	Log transformation helped overcome nonlinearities.
(Xie and Zhang 2010)	EnKF	A priori parameters	Implicit	MP-ND input errors	RMSE	ND in EnKF	Single State vector for uncertain parameters
Particle Filtering (van Leeuwen 2009). See "Joint State and Parameter Estimation"							
Variational Data Assimilation							
(Li and Navon 2001)	VDA	A priori parameters	Implicit	ND errors	Output error minimised	-	No estimate of predictive uncertainty
(Seo et al. 2003)	QN	-	Implicit	ND input and output errors	Output error minimised	-	No estimate of predictive uncertainty
Joint State and Parameter Estimation							
(Vrugt et al. 2006)	Outer SCEM-UA; Inner EnKF.	PPDF	Implicit and then Inferred from VD	Input error Inferred from VD; output error MP	Parameter PB; Total PB.	EnKF	EnKF only applied to linear model states.
(Salamon and Feyen 2009)	PF	PPDF per time step (time-varying parameters)	Implicit	Input error Inferred from VD; ERM.	Total PB.	PF	Joint inference of parameter and states in State-space model

(Moradkhani et al. 2005b)	EnKF for Parameters; EnKF for States	PPDF per time step (time-varying parameters)	ND	MP-ND input and output error	Total PB.	EnKF	Separate EnKF for States and Parameters
Error-correction							
(Abebe and Price 2003)	ANN	Implicit	Implicit	Implicit	Output error minimised	-	ANN trained to residual error time-series. Output data error in training not considered.
(Mancarella et al. 2008)	Local Model	Implicit	Implicit	Implicit	Output error minimised	Minimised	State uncertainty minimised by extrapolating error correction from measured locations
(Shrestha et al. 2009)	GLUE MCS	PPDF	-	-	PB	-	ANN trained to reproduce PB of physically based model from input data

GA, genetic algorithm; **GN**, Gauss-Newton technique; **GB**, Gradient-Based optimisation; **SA**, Simulated Annealing; **FOSM**, First-Order Second-Moment; **SCEM-UA**, Shuffled Complex Evolution Metropolis algorithm; **PPDF**, Posterior Probability Distribution Function; **ND**, Normal Distribution; **EDF**, exponential power density function; **MCS**, Monte Carlo Simulation; **M-H**, Metropolis-Hastings; **MCMC**, Markov Chain Monte Carlo; **NMD**, Normal Mixture Distribution; **AR**, Autoregressive Model; **PD**, Posterior Diagnostics; **PB**, Prediction Bounds; **DREAM-ZS**, Differential Evolution Adaptive Metropolis Algorithm; **SEP**, Skewed Exponential Power Density; **BF**, Bias Parameter; **SD**, standard deviation; **DRM**, Daily Rainfall Multiplier; **ERM**, Event Rainfall Multiplier; **HDM**, Heteroscedastic Discharge; **SP**, Stochastic Parameters; **CL**, Confidence Limits; **VD**, Variance Decomposition; **BMA**, Bayesian Model Averaging; **IEDF**, Informal Exponential Density Function; **RCEB**, Rating Curve Error Bounds; **FMF**, Fuzzy Membership Function; **DS**, Dempster-Shafer rules of combination; **BPA**, Basic Probability Assignment; **KF**, Kalman filter; **LHS**, Latin Hypercube Sampling; **MP**, Magnitude proportional; **EKF**, Ensemble Kalman Filter; **QN**, Quasi-Newton method; **EnKF**, Ensemble Kalman Filter; **RMSE**, Root Mean Square Error; **PF**, Particle Filter; **EMC**, Evolutionary Monte Carlo.

8 Appendix B: Glossary of Terms

A glossary of terms has been included to facilitate understanding of the relevant report sections (The Glossary has been modified from Goulsby and Samuels 2005). The glossary builds on the terms defined in Hutton et al. (2011).

Accuracy - closeness to reality.

Adaptive capacity - Is the ability to plan, prepare for, facilitate, and implement adaptation options. Factors that determine cities' adaptive capacity include economic wealth, technology and infrastructure, knowledge and skills, the nature of its institutions, its commitment to equity, and its social capital.

Adaptive Strategy – Method for optimising/expanding existing systems to reduce risk and vulnerability to change (e.g. climate change).

Aims - The objectives of groups/individuals/organisations involved with a project. The aims are taken to include ethical and aesthetic considerations.

Aleatory uncertainty – *see Natural Variability*.

Basin (river) (see catchment area) - the area from which water runs off to a given river.

Calibration – *see Calibration parameters*.

Catchment area - the area from which water runs off to a river.

Bias - The disposition to distort the significance of the various pieces of information that have to be used.

Characterisation - The process of expressing the observed/predicted behaviour of a system and its components for optimal use in decision making.

Climate Change – changes in weather over > 30 year time-periods, notably in response to modern anthropogenic influence.

Combined Sewer Overflow (CSO) – Overflow discharge from combined sewer systems that bypasses the Wastewater Treatment Plant and enters directly into the receiving water body. CSO discharge typically occurs during rainfall events.

Concentration Time – *see Lag time*.

Conditional probability - The likelihood of occurrence of an event given the prior occurrence of another event.

Confidence interval - A measure of the degree of (un)certainty of an estimate, usually presented as a percentage. For example, a confidence level of 95% applied to an upper and lower bound of an estimate indicates there is a 95% chance the estimate lies between the specified bounds. Confidence limits can be calculated for some forms of uncertainty (see knowledge uncertainty), or estimated by an expert (see judgement).

Consequence - An impact such as economic, social or environmental damage/improvement that may result from a flood or UWS failure. May be expressed quantitatively (e.g. monetary value), by category (e.g. High, Medium, Low) or descriptively.

Coping capacity - The means by which people or organisations use available resources and abilities to face adverse consequences that could lead to a disaster.

Correlation - Between two random variables, the correlation is a measure of the extent to which a change in one tends to correspond to a change in the other. One measure of linear dependence is the correlation coefficient ρ . If variables are independent random variables then $\rho = 0$. Values of +1 and -1 correspond to full positive and negative dependence respectively. Note: the existence of some correlation need not imply that the link is one of cause and effect.

Decision uncertainty - The rational inability to choose between alternative options.

Design objective - The objective (put forward by a stakeholder), describing the desired performance of an intervention, once implemented.

Dependence - The extent to which one variable depends on another variable. Dependence affects the likelihood of two or more thresholds being exceeded simultaneously. When it is not known whether dependence exists between two variables or parameters, guidance on the importance of any assumption can be provided by assessing the fully dependent and independent cases (see also correlation).

Demand - Amount of water consumed/extracted by domestic and industrial users from the WDS (typically expressed in volumetric terms per unit time period).

Deterministic process / method - A method or process that adopts precise, single-values for all variables and input values, giving a single value output.

Discharge (stream, river, sewer pipe) - as measured by volume per unit of time.

Dry Weather Flow - Flow in the sewer system during dry weather that originates from domestic and industrial users.

Element - A component part of a system.

Epistemology - A theory of what we can know and why or how we can know it.

Error - Mistaken calculations (e.g. from a model) or measurements with quantifiable and predictable differences.

Expectation - the expected value of a variable refers to the mean value the variable takes. For example, in a 100 year period, a 1 in 100 year event is expected to be equalled or exceeded once. This can be defined mathematically.

Extrapolation - The inference of unknown data from known data, for instance future data from past data, by analysing trends and making assumptions. Applying a derived relationship from one time-period to conditions different from that in which the relationship was derived.

Failure - Inability to achieve a defined performance threshold (response given loading). "Catastrophic" failure describes the situation where the consequences are immediate and severe, whereas "prognostic" failure describes the situation where the consequences only grow to a significant level when additional loading has been applied and/or time has elapsed.

Forecast - Prediction of a future system state, typically derived from a numerical model.

Failure mode - Description of one of any number of ways in which a system may fail to meet a particular performance indicator.

Functional design - The design of an intervention with a clear understanding of the performance required of the intervention.

Governance - The processes of decision making and implementation

Harm - Disadvantageous consequences; economic, social or environmental (See *Consequence*).

Hazard - A physical event, phenomenon or human activity with the *potential* to result in harm. A hazard does not necessarily lead to harm.

Hazard mapping - The process of establishing the spatial extents of hazardous phenomena.

Hierarchy - A process where information cascades from a greater spatial or temporal scale to lesser scale and vice versa.

Human reliability - Probability that a person correctly performs a specified task.

Ignorance - Lack of knowledge.

Institutional uncertainty - inadequate collaboration and/or trust among institutions, potentially due to poor communication, lack of understanding, overall bureaucratic culture, conflicting sub-cultures, traditions and missions.

Integrated risk management- An approach to risk management that embraces all sources, pathways and receptors of risk and considers combinations of structural and non-structural solutions.

Integrated Water Resource Management - IWRM is a process which promotes the co-ordinated management and development of water, land and related resources, in order to maximise the resultant economic and social welfare in an equitable manner without compromising the sustainability of vital ecosystems.

Intervention - A planned activity designed to effect an improvement in an existing natural or engineered system (including social, organisation/defence systems).

Joint probability - The probability of specific values of one or more variables occurring simultaneously. For example, extreme water levels in estuaries may occur at times of high river flow, times of high sea level or times when both river flow and sea level are above average levels. When assessing the likelihood of occurrence of high estuarine water levels it is therefore necessary to consider the joint probability of high river flows and high sea levels.

Judgement - Decisions taken arising from the critical assessment of the relevant knowledge.

Knowledge - Spectrum of known relevant information.

Knowledge uncertainty - Uncertainty due to lack of knowledge of all the causes and effects in a physical or social system (also termed epistemic uncertainty). For example, a numerical model of the sewer system may not include an accurate mathematical description of all the relevant physical processes. The model is thus subject to a form of knowledge uncertainty. Various forms of knowledge uncertainty exist, including:

Process model uncertainty - All models are an abstraction of reality and can never be considered true. They are thus subject to process model uncertainty.

Measured data versus modelled data comparisons give an insight into the extent of model uncertainty but do not produce a complete picture.

Statistical inference uncertainty - Formal quantification of the uncertainty of estimating the population from a sample. The uncertainty is related to the extent of data and variability of the data that make up the sample.

Statistical model uncertainty - Uncertainty associated with the fitting of a statistical model. The statistical model is usually assumed to be correct. However, if two different models fit a set of data equally well but have different extrapolations/interpolations then this assumption is not valid and there is statistical model uncertainty.

Lag Time – The characteristic time for a response to an input at a given location in a system.

Likelihood - A general concept relating to the chance of an event occurring. Likelihood is generally expressed as a probability or a frequency (as a value between 0 = impossible; 1 = certain).

Marginal Probability – *see Probability.*

Mitigation – to moderate the force or impacts of an event.

Model Based predictive Control – the application of numerical models, often coupled with an optimisation procedure, to identify the optimal control decision of a system in response to future demands on that system.

Natural variability - Uncertainties that stem from the assumed inherent randomness and basic unpredictability in the natural world and are characterised by the variability in known or observable populations (also known as Aleatory uncertainty).

Nowcast – A forecast of the immediate state of a system, typically up to 6 hours.

Objectives – A goal, typically defined as the maximisation or minimisation of a given function. For example, minimise cost whilst maintain system performance.

Optimisation – Intervention that achieves the best performance of a system in reference to one or more (competing) objectives. In modelling, adjustment of system parameters to achieve objectives pertaining to the modelled system.

Operating Horizon – *See Time Horizon.*

Parameters - The parameters in a model are the constants chosen to represent the chosen context and scenario. In general the following types of parameters can be recognised:

Exact parameters - which are universal constants, such as the mathematical constant: Pi (3.14259...).

Fixed parameters - which are well determined by experiment and may be considered exact, such as the acceleration of gravity, g (approximately 9.81 m/s).

A-priori chosen parameters - which are parameters that may be difficult to identify by calibration and so are assigned certain values. However, the values of such parameters are associated with uncertainty that must be estimated on the basis of a-priori experience, for example detailed experimental or field measurements

Calibration parameters - which must be established to represent particular circumstances. They must be determined by calibration of model results for historical data on both input and outcome. The parameters are generally

chosen to minimise the difference between model outcomes and measured data on the same outcomes. It is unlikely that the set of parameters required to achieve a "satisfactory" calibration is unique, reflecting a state of equifinality.

Parameter Hypercube - Multi-dimensional mode space where each dimension consists of a range of potential values for a particular model parameter.

Performance - The degree to which a process or activity succeeds when evaluated against some stated aim or objective.

Performance indicator - The well-articulated and measurable objectives of a particular project or policy. These may be detailed engineering performance indicators, such as acceptable CSO volumes, minimum pressure in WDS, rock stability, or more generic indicators such as public satisfaction.

Possibility - The likelihood of a state or event occurring in the future. Possibility differs from probability. Possibility theory was developed in the face of uncertain and often subjective understanding of the propensity for future states with little information from the past to inform on future likelihood.

Precautionary Principle - Where there are threats of serious or irreversible damage, lack of full scientific certainty shall not be used as a reason for postponing cost-effective measures to prevent environmental degradation.

Precision - degree of exactness regardless of accuracy.

Preparedness - The ability to ensure effective response to the impact of hazards, including the issuance of timely and effective early warnings and the temporary evacuation of people and property from threatened locations.

Probability - A measure of our strength of belief that an event will occur. For events that occur repeatedly the probability of an event is estimated from the relative frequency of occurrence of that event, out of all possible events. In all cases the event in question has to be precisely defined, so, for example, for events that occur through time reference has to be made to the time period, for example, annual exceedance probability. Probability can be expressed as a fraction, % or decimal. For example the probability of obtaining a six with a shake of four dice is $1/6$, 16.7% or 0.167.

Probabilistic method - Method in which the variability of input values and the sensitivity of the results are taken into account to give results in the form of a range of probabilities for different outcomes.

Probability density function (distribution) - Function which describes the probability of different values across the whole range of a variable (for example across a parameter value in a particular model).

Probabilistic reliability methods - These methods attempt to define the proximity of a structure to fail through assessment of a response function. They are categorised as Level III, II or I, based on the degree of complexity and the simplifying assumptions made (Level III being the most complex).

Process model uncertainty - See *Knowledge uncertainty*.

Project Appraisal - The comparison of the identified courses of action in terms of their performance against some desired ends.

Progressive failure - Failure where, once a threshold is exceeded, significant (residual) resistance remains enabling the defence to maintain restricted performance. The immediate consequences of failure are not necessarily

dramatic but further, progressive, failures may result eventually leading to a complete loss of function.

Random events - Events which have no discernible pattern.

Receiving water body – A water body, typically a river, lake or sea that receives effluent from the Sewer system or WWTW.

Recovery time - The time taken for an element or system to return to its prior state after a perturbation or applied stress.

Reliability index - A probabilistic measure of the structural reliability with regard to any limit state.

Real-Time Control – The process by which control structures in a given system are modified in response to real-time information derived from in situ measurements and models.

Resilience - The ability of a system/community/society/defence to react to and recover from the damaging effect of realised hazards.

Resistance . The ability of a system to remain unchanged by external events.

Return period - The expected (mean) time (usually in years) between the exceedence of a particular extreme threshold. Return period is traditionally used to express the frequency of occurrence of an event, although it is often misunderstood as being a probability of occurrence.

Risk - Risk is a function of probability, exposure and vulnerability. Often, in practice, exposure is incorporated in the assessment of consequences, therefore risk can be considered as having two components: the probability that an event will occur and the impact (or *consequence*) associated with that event. See Section 4.3 above. Risk = Probability multiplied by consequence

Risk analysis - A methodology to objectively determine risk by analysing and combining probabilities and consequences.

Risk assessment - Comprises understanding, evaluating and interpreting the perceptions of risk and societal tolerances of risk to inform decisions and actions in the flood risk management process.

Risk communication (in context) - Any intentional exchange of information on environmental and/or health risks between interested parties.

Risk management - The complete process of risk analysis, risk assessment, options appraisal and implementation of risk management measures

Risk management measure - An action that is taken to reduce either the probability of flooding or the consequences of flooding or some combination of the two

Risk mapping - The process of establishing the spatial extent of risk (combining information on probability and consequences). Risk mapping requires combining maps of hazards and vulnerabilities. The results of these analyses are usually presented in the form of maps that show the magnitude and nature of the risk.

Risk mitigation - See *Risk reduction*.

Risk perception - Risk perception is the view of risk held by a person or group and reflects cultural and personal values, as well as experience.

Risk reduction - The reduction of the likelihood of harm, by either reduction in the probability of a flood occurring or a reduction in the exposure or vulnerability of the receptors.

Risk profile - The change in performance, and significance of the resulting consequences, under a range of loading conditions. In particular the

sensitivity to extreme loads and degree of uncertainty about future performance.

Risk register - An auditable record of the project risks, their consequences and significance, and proposed mitigation and management measures.

Risk significance (in context) - The separate consideration of the magnitude of consequences and the frequency of occurrence.

Robustness - Capability to cope with external stress. A decision is robust if the choice between the alternatives is unaffected by a wide range of possible future states of nature. Robust statistics are those whose validity does not depend on close approximation to a particular distribution function and/or the level of measurement achieved.

SCADA – Supervisory Control And Data Acquisition. Computer systems that monitor the state of a system, and allow control of devices within the system.

Scale - Difference in spatial extent or over time or in magnitude; critical determinant of vulnerability, resilience etc.

Scenario - A plausible description of a situation, based on a coherent and internally consistent set of assumptions. Scenarios are neither predictions nor forecasts. The results of scenarios (unlike forecasts) depend on the boundary conditions of the scenario.

Sensitivity - Refers to either: the resilience of a particular receptor to a given hazard. For example, frequent sea water flooding may have considerably greater impact on a fresh water habitat, than a brackish lagoon; or: the change in a result or conclusion arising from a specific perturbation in input values or assumptions.

Sensitivity Analysis - The identification of those parameters which critically affect the output of a model or process. Conducted to better understand system operation, and allocate resources to constrain model output.

Sewer System – Infrastructure of pipes and control structures that conveys sewerage and rainfall-runoff in urban areas from buildings and the roads to the wastewater treatment plant and receiving water body.

Skeletonisation – Removal of pipes not considered essential to the operation of a WDS model.

Source - The origin of a hazard (for example, heavy rainfall, strong winds, surge etc).

Stakeholders - Parties/persons with a direct interest (stake) in an issue.

Stakeholder Engagement - Process through which the stakeholders have power to influence the outcome of the decision. Critically, the extent and nature of the power given to the stakeholders varies between different forms of stakeholder engagement.

Statistic - A measurement of a variable of interest which is subject to random variation.

Strategy - A strategy is a combination of long-term goals, aims, specific targets, technical measures, policy instruments, and process which are continuously aligned with the societal context.

Strategic spatial planning - Process for developing plans explicitly containing strategic intentions referring to spatial development. Strategic plans typically exist at different spatial levels (local, regional etc).

Statistical inference uncertainty - *See Knowledge uncertainty*

Statistical model uncertainty - *See Knowledge uncertainty*

Sustainable Development - is development that meets the needs of the present without compromising the ability of future generations to meet their own needs

Susceptibility - The propensity of a particular receptor to experience harm.

System - An assembly of elements, and the interconnections between them, constituting a whole and generally characterised by its behaviour.

System state - The condition of a system at a point in time.

Tolerability - Refers to willingness to live with a risk to secure certain benefits and in the confidence that it is being properly controlled. To tolerate a risk means that we do not regard it as negligible, or something we might ignore, but rather as something we need to keep under review, and reduce still further if and as we can. Tolerability does not mean acceptability. For example, tolerance of CSO or sewer surcharge.

Time Horizon – Time required to initiate actions to mitigate against the impact of a forecasted event.

Uncertainty - A general concept that reflects our lack of sureness about someone or something, ranging from just short of complete sureness to an almost complete lack of conviction about an outcome.

Urban Wastewater System – triplet of components: Sewer System, Wastewater treatment plant and receiving water body designed to mitigate against flooding and provide sanitation.

Validation - is the process of comparing model output with observations of the 'real world'.

Variability - The change over time of the value or state of some parameter or system or element where this change may be systemic, cyclical or exhibit no apparent pattern.

Variable - A quantity which can be measured, predicted or forecast which is relevant to describing the state of the flooding system e.g. water level, discharge, velocity, wave height, distance, or time. A prediction or forecast of a variable will often rely on a simulation model which incorporates a set of parameters.

Vulnerability - Characteristic of a system that describes its potential to be harmed. This can be considered as a combination of susceptibility and value.

Wastewater Treatment Plant (WWTW) – Treatment plant for the removal of contaminants and nutrients from sewerage for entry as effluent into the receiving water body.

Water Distribution Network (WDS) – Network of pipes, pumps, nodes, tanks and valves that distributes drinking water to meet consumer demands.

