Uncertainty analysis in integrated catchment modelling

Moreno Rodenas, Antonio

DOI
10.4233/uuid:a8577854-a254-44a4-bdb2-b63218454828

Publication date
2019

Document Version
Final published version

Citation (APA)
Moreno Rodenas, A. (2019). *Uncertainty analysis in integrated catchment modelling*. [Dissertation (TU Delft), Delft University of Technology]. https://doi.org/10.4233/uuid:a8577854-a254-44a4-bdb2-b63218454828

UNCERTAINTY ANALYSIS IN INTEGRATED CATCHMENT MODELLING

Proefschrift

ter verkrijging van de graad van doctor
aan de Technische Universiteit Delft,
op gezag van de Rector Magnificus prof. dr. ir. T.H.J.J. van der Hagen, voorzitter van het College voor Promoties,
in het openbaar te verdedigen op mandag 8 juli 2019 om 15:00 uur

door

Antonio Manuel MORENO RODENAS

Ingeniero de Caminos, Canales y Puertos,
Universitat Politècnica de València, Spanje,
geboren te Valencia, Spanje.

Dit proefschrift is goedgekeurd door de

promotor: prof. dr. ir. F.H.L.R. Clemens
promotor: dr. ir. J.G. Langeveld

Samenstelling promotiecommissie:

Rector Magnificus,	voorzitter
Prof. dr. ir. F.H.L.R. Clemens	Technische Universiteit Delft
Dr. ir. J.G. Langeveld	Technische Universiteit Delft
Onafhankelijke leden:	
Prof. dr. Z. Kapelan	Technische Universiteit Delft
Prof. dr. P. Willems	Katholieke Universiteit Leuven
Prof. dr. G.B.M Heuvelink	Wageningen University and Research
Prof. dr. D. Muschalla	Graz University of Technology
Dr. W. Shepherd	University of Sheffield
Prof. dr. ir. M. Kok	Technische Universiteit Delft, reserve member

Keywords: Uncertainty analysis; Integrated urban water systems; Model emulation; Dissolved oxygen simulation

Printed by: proefschrift-aio.nl

Front & Back: Temperature-stratified dispersion of milk in water by A.M Moreno Rodenas.

ISBN 978-94-92801-89-0

An electronic version of this dissertation is available at
`http://repository.tudelft.nl/`.

CONTENTS

LIST OF TABLES

List of Figures

Summary

The adoption of increasingly restrictive water quality standards is directed to maintain natural ecosystems in a good status. Complying with such standards requires significant investments in water infrastructure and operations. Consequently, mathematical simulation is usually applied to assist in the decision-making process for such large-scale actuations. In particular, environmental models are proposed to represent the wastewater cycle in natural water bodies, such that the effect of different pollution mitigation alternatives can be estimated. Integrated catchment models (ICM) aim at simulating water quality dynamics by representing the link between urban drainage networks, wastewater treatment operations, rural hydrology and river physical-biochemical processes. However, these subsystems present dynamics across multiple spatiotemporal scales and many relevant processes are still not fully understood. System observations are scarce and often insufficient to identify most model representations. As a result, ICM studies often produce significant output uncertainties.

Uncertainty analysis (UA) aims at quantifying the degree of reliability of modelling outcomes and diagnosing simulation structures so that further data acquisition or model improvements can be directed. However, UA is still seldom applied in the field of urban drainage and environmental assessment studies, due to limitations in organisational and computational resources available for ICM practitioners. Consequently, model-end users are often not aware of the implications of many error sources, and there is an insufficient communication of model structural assumptions and its associated uncertainties. This thesis summarises efforts towards increasing the understanding of uncertainty analysis in water quality integrated catchment modelling studies.

This work presents methods to describe the effects of spatial and temporal characteristics of rainfall data for the simulation of dissolved oxygen depletion dynamics. Rainfall is one of the main driving forces for organic-pollution loads in many urbanised catchments, thus its appropriate consideration is of importance when simulating water quality impacts. The effect of using point and distributed rainfall data (e.g. rain gauges or radar) is discussed, highlighting the need for the careful consideration of spatial rainfall characteristics in low-land systems.

Uncertainties in combined sewer overflow (CSO) water quality influents also plays a major role when simulating receiving water quality impacts. Unfortunately, scientific understanding of in-sewer water quality processes is still poor, and thus such models are highly uncertain. In this thesis, a simple method to represent the observed probability distribution and correlation structure of water quality pollutant concentrations is presented. This is intended to draw attention on the possible underestimation of uncertainties when neglecting the correlation structure between CSO processes.

Nevertheless, formal quantification of uncertainties in ICMs is still not widely applied (neither in practice nor science). One of the current largest limitations in its applicability is the computational effort required. This thesis discusses the use of data-driven emulation schemes to accelerate model evaluation during inference or forward uncertainty quantification. A practical application is shown by emulating river flow and dissolved oxygen in a large-scale urban water system. The emulator allowed to estimate the effect of changes in physical and biochemical river parameters on the output of a computationally expensive simulator. This fast approximation was then used to infer parameter knowledge from local observations in the system. The application of model emulation techniques can facilitate the integration of real-world scale catchment systems into computationally demanding applications, such as sensitivity, uncertainty analysis or formal parametric inference.

On the other hand, data-driven emulators are still severely limited by the input and output spaces dimensionality. Classical applications involve emulating the effect of only a handful of static parameters towards a one or two dimensional output space. Consequently, time-dynamic inputs are seldom considered, thus limiting the range of applicability of emulation. This thesis presents a novel method for encoding rainfall dynamics and parametric information in a data-driven emulator for the physically-based simulation of overland flow hydrodynamics (e.g. 2D Shallow water equations). This implementation allows sampling arbitrarily long dynamic rainfall inputs (yet spatially homogeneous) and parametric variations at a fraction of the original simulator time. Therefore fast sampling applications as early-warning flooding schemes, statistical inference or real time control can be applied in physically-based overland flow propagation.

To conclude, this work presents the formal application of an uncertainty quantification and decomposition scheme to characterise the outcomes from a large-scale ICM for water quality assessment. A by-source uncertainty decomposition revealed that prior knowledge in river physical and biochemical parameters is responsible for most of the dissolved oxygen output variability in the Dommel river (The Netherlands). Local measurements in the system were used to update knowledge about the river submodel parameters. Rainfall uncertainties and CSO water quality submodel uncertainties remain the largest uncertainty sources in the system after the update of river parameters. Therefore, further efforts towards the reduction of epistemic uncertainties when modelling dissolved oxygen dynamics (in the Dommel) should be directed to better understand in-sewer water quality dynamics and rainfall spatio-temporal fields.

1

Introduction and scope

1.1. Water quality management

The EU Water Framework Directive (EC, 2000) enforces member states to reach a "good ecological and chemical status in inland and coastal water bodies". Currently, meeting the proposed environmental regulations is still a challenge for many European catchments. Complying with the environmental regulations often requires the adoption of intensive infrastructure development and regulatory plans. For instance, it is estimated that the extra cost needed to reach good surface water status in the Netherlands is 7.1 billion Euros (between 2007-2027, Ligtvoet et al. (2008)).

Water quality processes have a wide range of spatiotemporal characteristic scales associated (House et al., 1993), and mitigation or corrective measures largely differ depending on those. Table 1.1 shows some of the most relevant anthropogenic environmental impacts in surface and subsurface natural water bodies. For instance, the incorrect management of urban household sewage discharge is the direct cause for several of these water-pollution processes. Specially relevant (even in systems with fully connected sewage networks) are the discharge of wastewater through urban combined sewer overflow structures (CSOs) and insufficiently treated effluents from wastewater treatment facilities (WWTP). This discharge has the potential to impact the ecological and chemical status of receiving water bodies.

Mathematical models are widely used to optimise the effect of pollution corrective alternatives aiming towards a rational use of available resources. Integrated catchment modelling (ICM) which simulates water quality dynamics at the urban-rural scale is considered a key tool in the decision-making process for water management (Langeveld et al., 2013b; Rauch et al., 2002; Willems and Berlamont, 2002). ICMs often involve the joint simulation of sewer hydrodynamics, wastewater treatment processes, rural hydrology and river physical-biochemical dynamics, which renders highly complex modelling studies (Keupers, 2016; Muschalla et al., 2009; Solvi, 2006).

Environmental modelling outputs (e.g. ICMs derived water quality predictions) often have significant uncertainties associated. These uncertainties are related to the model abstraction process (i.e. highly complex systems represented by simplified and lumped mathematical descriptions), parameterisation (uncertain value of model parameters, which are system and time dependent), model forcing data (e.g. errors in measured rainfall maps, water temperature, solar radiation measurements etc.), model input data

Table 1.1: Common environmental impacts, factors and scales associated

Water quality process	Parameters (examples)	Main sources	Space and time scales
Oxygen depletion	COD BOD N_Kjendahl	CSOs WWTP	River-lake reach (100 m - 100 km) Hour - Day
Eutrophication	Total Phosphorus Total Nitrogen	CSOs WWTP Agriculture	River-lake reach (100 m - 100 km) Week - Month
Acute toxicity	NH3 Temperature pH	CSOs WWTP Industry	Local (10 m - 1000 m) Minute - Hour
Infective diseases	Faecal coliforms Salmonella Enteroviruses Vibrio cholera	CSOs Urban flooding WWTP	Local (10 m - 10 km) Day - Month
Aesthetic pollution	Macro-Solids Plastics	CSOs Industry	River - lake - coastal reach (100m - 100 km) Hours - Decade
Ecological toxicity	Endocrine disruptors Pesticides	CSOs WWTP Agriculture Industry	River - lake reach (100 m - 100 km) Hour - Decade (Trophic accumulation)
Accumulation of heavy metals	Mercury Arsenic Lead Zinc	Industry CSOs	Regional (10 km - 100 km) Year - Decade (Trophic and sediment accumulation)
Groundwater nitrate accumulation	NO3	Agriculture Farming	Regional (10 km - 100 km) Year - Decade

(e.g. digital elevation models, errors in underground infrastructure databases), model validation data (insufficient calibration-validation datasets), or model use (extrapolating system dynamics in the future). These uncertainty sources make that characterising the outcomes of ICMs as a deterministic process, might render incorrect system diagnosis and lead to inefficient decision-making (Schellart et al., 2010). Therefore, performing uncertainty analysis in ICMs is of foremost importance to increase reliability in model outcomes and to direct further monitoring and modelling efforts.

Comprehensive uncertainty analysis is mostly applied by academics and seldom in practice (Kleidorfer, 2010; Vanrolleghem et al., 2011). Furthermore, even academic examples for real-world integrated catchment modelling uncertainty analysis are very scarce (Tscheikner-Gratl et al., 2019). Thus, increasing our knowledge on the applicability of uncertainty analysis methodologies in ICMs is still necessary (Deletic et al., 2012).

1.2. THE EVOLUTION OF WATER QUALITY STANDARDS IN THE NETHERLANDS

THE progressive urbanisation experienced by most European countries from the early 1900s led to a spread of drinking water supply, and to the development of extensive sewer networks. Severe water quality issues were common due to the direct discharge of wastewater into surface waters. This steered the large-scale adoption of mechanical and biological treatment facilities for sewage, which occurred in the Netherlands by 1970-1980. The 1970 Pollution of Surface Waters Act (WVO Wet Verontreiniging Oppervlaktewater) was directed to the mitigation of countrywide severe oxygen depletion in river and lakes. This action focused on the reduction of discharged biodegradable matter and materialised as the adoption of standards for the biodegradable oxygen demand (BOD) and suspended solids (SS) concentration in WWTP effluents (Table 1.2). By 1978 new limitations were enforced to reduce ammonium levels. In the decade of the 80s, and to this day, increasingly rigid standards limit the discharge of nitrogenous and phosphorous nutrients in natural water surfaces, to reduce the risk of eutrophication.

Table 1.2: Evolution of Dutch wastewater treatment standards (adapted from Langeveld (2004))

Period	**Standard Type**	**Characteristics of the standard**
Before 1970	No standard	Approximately only 30% of the sewage production is treated. Mainly mechanical separation or trickling filters. Severe oxygen depletion issues in lakes and rivers across the country.
1970 - now	BOD	20 mgO2/l (10-day average)
	Suspended Solids	30 mgSS/l (10-day average)
1978 - now	Kjeldahl nitrogen	20 mgN/l (10-day average)
1990 - now	Total phosphorus	2 mgP/l (for plants of <100,000 p.e.)
		1 mgP/l (for plants of >100,000 p.e.)
1992 - now	Total Nitrogen	15 mgP/l (for plants of <20,000 p.e.) (Yearly average)
		10 mgP/l (for plants of >20,000 p.e.) (Yearly average)

In parallel to the adoption of wastewater treatment standards, limitations and design codes for the construction of urban drainage systems emerged as shown in Table 1.3. Early urban drainage networks were eminently combined systems (which today still represent roughly 60% of the Dutch drainage network), meaning that rain water and wastewater are collected and transported to treatment through the same pipe network. Under heavy storm conditions, storm water might exceed the transport and storage capacity of combined sewers. To prevent urban flooding, emergency discharge structures, also known as combined sewer overflows (CSO) discharge diluted wastewater to receiving water bodies, thus impacting on water quality if the pollutant load is too high or to frequent. Early guidelines appearing in the 1950s focused on limiting this overflow discharge frequency (to 3-10 CSO events per year). Excess pollution in surface waters, led to the redefinition of standards by the 1990s, which regulated the model-assisted design of sewer drainage systems. Firstly, the simulated annual volume of CSO discharges was limited and later it was extended to annual discharge of organic pollutants, as represented by chemical oxygen demand (COD).

By the late 1990s and 2000s, the scientific and regulatory community questioned further the validity of load-limiting based standards, which neglect the buffering capacity of receiving water bodies. By 2000s, there was a shift in environmental policies, especially after the approval of the EU Water Framework Directive (EC, 2000). The WFD defines an immision approach, which on top of the pre-existing load limiting standards, require all natural receiving water bodies to have a good ecological and chemical status. Good status is defined by evaluating the ecosystem's adequacy to sustain populations of indigenous fish species, micro-invertebrates and other selected flora and fauna. This

Table 1.3: Evolution of Dutch urban drainage environmental standards (adapted from Langeveld (2004))

Period	Standard Type	Characteristics
Before 1951	Wet weather flow dilution	Depending on the nature of the receiving water system, a dilution factor between 3-10 times should be guaranteed before CSOs are activated
1951 - 1992	CSO frequency limitations	Acceptable overflow frequency between 3-10 CSO discharges/year, depending on local characteristics
1992 - 2001	Limitation in CSO volume	Each designed combined sewer system should have an equivalent annual overflow volume as an idealized 'reference' system with an in-sewer storage capacity of 7 mm and a storm water settling tank of 2 mm with a pumping capacity of 0.7 mm/h + dry weather flow
2001 - now	Overflow Loads	The reference system is updated so it discharges the Chemical oxygen demand load equivalent to 50 kgCOD/ha/year. Although, a fix concentration of 250 mgCOD/l is used to compute the load, thus is still a volume based assessment
1992 - now	Receiving water quality assessment	Immision based approach. Evaluation of the discharges that cause a deterioration of the local receiving water body

approach requires not only attention to the urban drainage and WWTP systems, but also to the receiving water bodies.

Urban drainage and wastewater treatment systems are assets, which represent an important public investment. This infrastructure is often designed to be serviceable for long periods of time (30-60 years), yet change in technology, restrictions and legislative requirements occurs at shorter time-scales. Therefore, meeting the standards often requires retrofitting and upgrading pre-existing systems.

Figure 1.1 depicts the evolution of optimisation studies undertaken in the re-evaluation of Dutch urban drainage systems. Between 1990s and 2010s, mainly uncalibrated models and measurement studies were carried out to optimise the hydraulic design capacity in treatment works, and to re-evaluate compliance of emission regulations in sewer networks.

The call for model calibration and verification was early adopted from the 1990s by the hydraulic modelling community. Although its generalisation into water emission studies happens during the 2000s. This was eminently due to a lack of extensive monitoring data, and a slow transfer of scientific methods to practice. New modelling standards (e.g. Clemens (2001); WaPUG (2002)), led to the development of calibrated emission sewer and WWTP optimisation studies (2nd generation). Currently, inline with the legislative focus, optimisation studies account for the immision approach. Therefore, the link between urban water, WWTP and receiving water quality is increasingly being considered in new system studies. However, representing the full interaction of the water subsystems requires of large amounts of monitoring data (Langeveld et al., 2013a), which is still seldom available. This implies that the appropriate calibration and identification of simulated processes is still a challenge. This 3rd generation studies are carried out with a multiobjective perspective, trying to reduce water pollution impacts, allocate available resources optimally, reduce CO2 emissions and reduce the operational cost of the water infrastructure assets. The undertaking of such multi-scale and multi-objective assessments required the adaptation of simulation platforms to account for all relevant water system links, which led to the development of integrated catchment modelling studies.

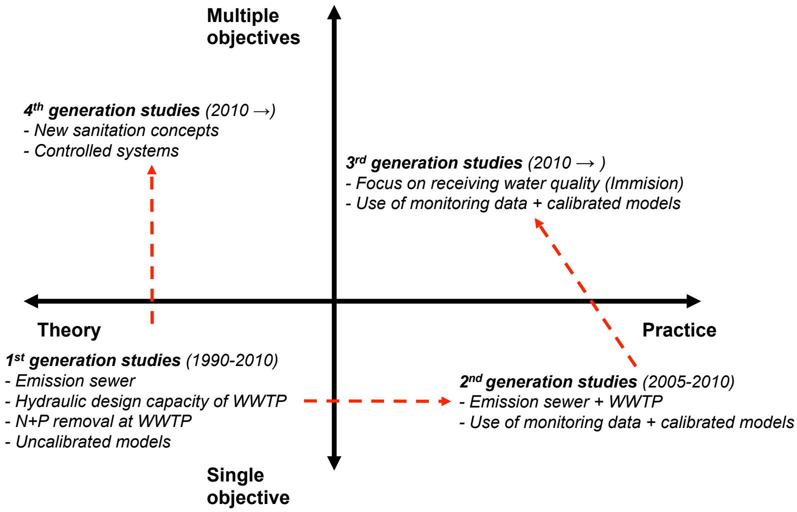


Figure 1.1: Generations of optimisation studies for urban water quality studies

Additionally, new sanitation concepts are emerging in parallel to the operation of conventional wastewater treatment facilities, focusing in resource recovery applications, distributed treatment (Larsen et al., 2016), and alternative sanitation technologies (Radhakrishnan et al., 2018). As these alternative technologies are adopted in practice, they will become part of future operational system studies.

1.3. INTEGRATED URBAN WATER QUALITY MODELLING

THE immision assessment paradigm required the adaptation of modelling platforms to account for the interaction between the different subsystems. Consequently, requiring to simulate the link between urban drainage water quality-quantity models, WWTP models, rural hydrology and receiving water models (Rauch et al., 2002). Figure 1.2 shows a time-line of the development of integrated urban water systems (IUWS) modelling as depicted by Bach et al. (2014). The 1st INTERURBA conference (Lijklema et al., 1993) marked the beginning of an international effort to consolidate theoretical ideas into real applications for integrated modelling. Several research studies started linking pre-existing submodel implementations in an input-to-output approach (see early examples in Leinweber et al. (1999) or Clifforde et al. (1999)), often transferring data in a manual way. During the 2000's there was a wide recognition of the need of deterministic integrated catchment modelling as a key tool in the assessment of urban drainage impacts in natural ecosystems, which was revisited in the 2nd INTERURBA conference (Harremöes, 2002). This led to a rapid growth of dedicated IUWS model software packages, such as SIMBA (IFAK, 2007) or WEST (Vanhooren et al., 2003; Solvi, 2006), among others. The popularisation of integrated catchment studies was followed by the creation of modelling guidelines and recommendations (Benedetti et al., 2013a; Muschalla et al., 2009).

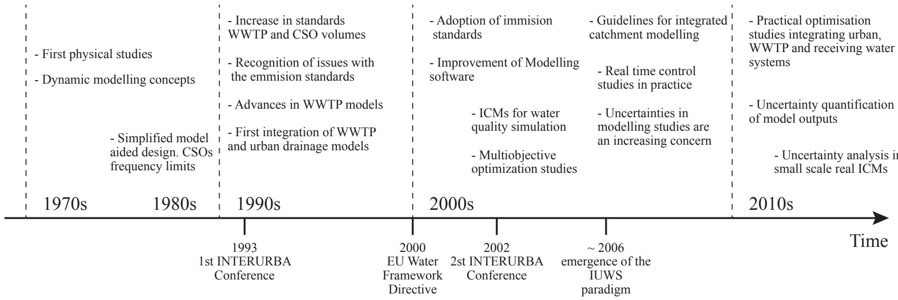


Figure 1.2: Time line of integrated urban catchment modelling development (adapted and expanded from Bach et al. (2014))

Deterministic ICMs are the chosen tool to assess compliance of the EU Water framework directive and from the 2010s, integrated catchment models are increasingly been applied in real-case studies as in Langeveld et al. (2013b), Andrés-Doménech et al. (2010) or Vezzaro et al. (2014b). These modelling platforms are used to justify significant investments in water systems. For instance, Benedetti et al. (2013b) discussed the selection of cost-effective solutions for the improvement of the water quality performance of the system of the Dommel river (The Netherlands). This made use of a deterministic ICM study to select measures with ranging capital investments between 40M€ to 160M€, aiming towards a reduction of dissolved oxygen and ammonium acute impacts in the river. Therefore, quantifying the degree of confidence in ICM outcomes, and transparently communicating its limitations is necessary for an educated decision making process.

1.4. UNCERTAINTIES IN ENVIRONMENTAL MODELLING

UNCERTAINTIES are intrinsic to any modelling activity. The mere act of modelling implies building a construct based on hypotheses on the phenomenological or mechanistic behaviour of a particular system. Given system observations, the selected mathematical model can be adjusted so it achieves a quantifiable degree of representativity. Then, this mathematical structure is often used to infer further knowledge, predict future system states or test the effect of hypothetical scenarios. Figure 1.3 shows a scheme of the model abstraction process from a given physical reality along with error sources.

Underlying assumptions in environmental models are often strong simplifications of complex phenomena. Also, gathering data about a particular system is hampered by logistic, budgetary or practical constraints, thus the complete physically-based system description is often utopical in most environmental modelling endeavours. Inputs used as boundary conditions in the system delimitation, often come from measurements or other modelling activities, which induce additional uncertainties. Furthermore, using the modelling platform to predict or test virtual scenarios implies extrapolating existing system mechanics. Thus, complex modelling studies for environmental applications often render highly uncertain outputs, to a level, which can sometimes preclude decision-making (Reichert and Borsuk, 2005). The identification and quantification of the effect

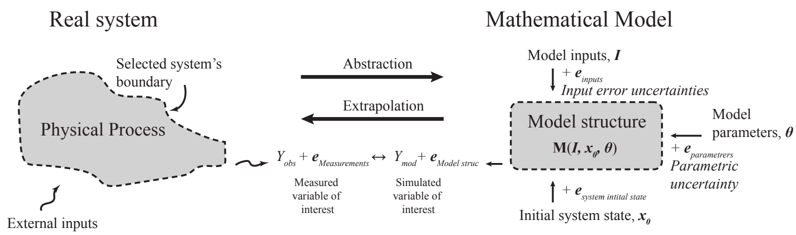


Figure 1.3: From the physical reality to model abstraction (adapted from Reichert (2014))

of those uncertainties is necessary for a transparent model implementation and an effective decision-making process (Pappenberger and Beven, 2006).

Modelling uncertainty is commonly classified in two broad categories (Slijkhuis et al., 1999):

- Aleatoric uncertainty: which is dominated by the intrinsic randomness of a certain process.
- Epistemic uncertainty: derived by a lack of knowledge on the underlying process.

This classification is rather subjective, since it depends on modeller's experience and means. In general, epistemic uncertainty is reserved for uncertainties which are related to the assumptions taken in the model abstraction phase and that can be reduced with further data acquisition, better calibration, or increase in process knowledge. Meanwhile, aleatoric uncertainty refers to the cases in which further efforts cannot be directed to reduce the uncertainty of a given process, and the modeller assumes a stochastic nature of the process. As an illustrative example, if challenged to model the trajectory of a cannon ball shot, one could take into account Newtonian mechanics, air drag, and wind direction to build a mathematical model of the expected trajectory. However, experimental observations may still render certain dispersion over the target. By taking into account even more processes and gathering additional knowledge (e.g. estimating the Coriolis acceleration or the projectile rotation and gyroscopic effect) one could further reduce the uncertainty of the predicted trajectory. However, measuring other influential physical variables as air density fluctuations in the projectile path, or propellant and projectile material heterogeneities, might be out of reach for today's observational capabilities. Thus, this residual error due to immeasurable variables might be considered as aleatoric uncertainty by the modeller. Uncertainty analysis is a scientific field, which deals with the quantification of modelling uncertainty sources with the aim of reducing epistemic uncertainties to aleatoric ones.

Refsgaard et al. (2007) proposed a classification of uncertainties in environmental modelling attending to: i) The location or source of uncertainty, ii) the type of uncertainty and iii) its nature (epistemic or stochastic). An analogous classification is adopted within the QUICS framework for uncertainty quantification and reporting in integrated catchment modelling studies (Tscheikner-Gratl et al., 2017). This structured classification is directed to systematically identify, report and quantify sources of uncertainty during the integrated model design and operation.

Uncertainty sources are classified as: a) Contextual, if the source of uncertainty is outside the boundaries of the system under study, b) input uncertainty, when it refers to forcing elements of the model simulation, c) parameter uncertainty, d) calibration, or uncertainties stemming from errors or insufficient data in the calibration measurement set and d) model structure, which is caused by an erroneous description of the system dynamics in the abstraction process. Nevertheless, several authors discuss that practical decomposition of parametric and structural uncertainties might not be possible, since they are fundamentally interlinked Bedford and Cooke (2001).

Attending to the degree of available knowledge, this classification identifies four types of uncertainty: a) Deterministic, b) statistical, c) scenario and d) deep uncertainties. This goes closely related to scale of determinacy from pure determinism to ignorance (Wynne, 1992). If the effect of a source of uncertainty cannot be quantified with the available data or knowledge, it is considered as a deep uncertainty source. This type of uncertainties might be larger than quantifiable ones in applications for environmental modelling (Willems, 2008), thus should be identified and reported, acknowledging the limits of extrapolation of the model under study. The evaluation of scenario uncertainty analysis is highly relevant when performing predictions from the modelling platform, and allows evaluating the effect of assumptions made in the abstraction process (e.g. revisiting assumptions of demographic growth when predicting an urban system development in time). On the other hand, studying statistical uncertainties plays a larger role in the diagnostic of model structures. This refers to the propagation of uncertainties in parameters, inputs and measured data to the output variables of the system. Statistical uncertainty evaluation can be used to direct further system observations and modelling efforts towards a reduction of epistemic sources of uncertainty.

1.5. STRATEGIES FOR THE QUANTIFICATION AND ANALYSIS OF MODELLING UNCERTAINTIES

A Large number of methods have been proposed to analyse uncertainties and their propagation in hydrological and water quality modelling (see a review in Beck (1987), Beven and Binley (1992) and Jakeman and Jakeman (2017)). Frameworks for their application in water quality and environmental modelling can be found in Tscheikner-Gratl et al. (2017), Deletic et al. (2012) or Refsgaard et al. (2006). However, the applicability of many of those strategies is limited by the characteristics and objectives of the model under study (Dotto et al., 2012; Sriwastava and Moreno-Rodenas, 2017).

Uncertainty analysis (UA) aims to identify and quantify the variability in modelling outcomes as a result of errors or uncertainties present in the simulation process (e.g. system abstraction, input forcing data errors, parameter uncertainties etc.). It can be understood, as evaluating the effect (in the modelling output) of the description of beliefs on the value of uncertain elements, when taking into account all inter-subjective scientific knowledge available to the modeller (Reichert et al., 2015), and its identification by source.

1.5.1. FORWARD PROPAGATION ANALYSIS

Forward uncertainty quantification is directed to estimate the degree of uncertainty in model outputs. A popular approach in environmental modelling is the use of Monte-Carlo (MC) simulation. MC has the advantage of being a purely numerical implementation, and highly parallelisable. MC is based in drawing random samples of inputs and parametric joint probability distributions and propagating them through the model structure. Each random sample serves to populate the probability distribution of the output, which can later be characterised by descriptive statistics.

Probability theory is used to describe variability in the value of model elements (often encoded as empirical or parametric probability density functions), which is derived from knowledge in physical constraints, measurements, expert elicitation (Garthwaite et al., 2005; O'Hagan, 1998) or inferred from observed data.

In order to accelerate the exploration of highly multidimensional spaces, samples can be drawn from structured samplings strategies (instead of drawing purely pseudo-random realisations), to this effect, Latin Hypercube (McKay et al., 1979) or orthogonal sampling (Owen, 1994) are often applied to reduce the number of required samples needed to describe the output distribution. Multi-fidelity modelling (Laloy et al., 2013)or model emulation (Conti and O'Hagan, 2010) can also be used when the simulator is too slow to sample from (yet limited to reduced parametric dimensionality). Also, when assessing the probability of rare events, MC convergence might require an impractically large amount of samples. To speed MC convergence, importance sampling is often applied (Bucklew, 2013).

On the other hand, when uncertain elements cannot be easily described in terms of formal probability distributions, scenario analysis may be applied. Thus, generating discrete simulations under changing conditions, which capture the variability of the underlying process and its influence in the modelling objectives (Börjeson et al., 2006; Urich and Rauch, 2014).

1.5.2. INVERSE UNCERTAINTY QUANTIFICATION

Model parametric inference represents a key approach in the description of modelling uncertainties. Prior knowledge about the uncertainty and characteristics of the model elements is updated in view of additional system observations. A Bayesian approach provides a formal manner to account for this transfer of observed information to a refined knowledge in model elements. In essence, the estimations of a model structure are contrasted against measured data:

$$Y_{obs} = M(x_0, I, \theta_M) + Z \tag{1.1}$$

being Y_{obs} a layout of observed system states, simulated through a mathematical model structure M, dependent on a set of initial conditions x_0, input forcing data I, uncertain parameters θ_m, and being Z an error generating process which captures model structural uncertainties and measurement errors. This approach requires the construction of Z, based in fundamental assumptions about the model and measurement error structure.

Prior knowledge in model parameters (or input latent variables) is encoded as a joint probability distribution of density, $P(\theta_m)$. Attending to the Bayes theorem:

$$P(\theta_m|Y_{obs}) = \frac{P(Y_{obs}|\theta_m) \cdot P(\theta_m)}{P(Y_{obs})} \tag{1.2}$$

the conditional (or posterior) distribution of the model parameters on the observed data, $P(\theta_m|Y_{obs})$ can be updated when knowing the value of the prior density $P(\theta_m)$, the probability of the data being generated by a combination of parameters (also known as likelihood distribution) $P(Y_{obs}|\theta_m)$, which is derived from the assumptions made in the stochastic model for the error generating process in Equation 1.1, and the total probability of the observed data, $P(Y_{obs})$. In practice, computing this last term is very challenging, since it constitutes an integral form covering all possible data realisations, thus it is often mathematically intractable from a practical perspective. Yet, treating $P(Y_{obs})$ as a scaling constant:

$$P(\theta_m|Y_{obs}) \propto P(Y_{obs}|\theta_m) \cdot P(\theta_m) \tag{1.3}$$

a sampling algorithm (typically a Markov-chain sequence) can be used to populate the posterior distribution $P(\theta_m|Y_{obs})$ (e.g. Metropolis et al. (1953) or Goodman and Weare (2010)) from random parameter vector realisations.

The error generating process Z, is encoded by the modeller's assumptions (Kuczera, 1983; Schoups and Vrugt, 2010), and errors induced by its incomplete description can be transmitted into the inferred probability distributions. Bias in model-measurement comparison is often present when complex physical processes with a strong system memory are simulated (e.g. flow dynamics or water quality processes), errors in the model structure can be identified by the use of time-dependent parametric descriptions (Reichert and Mieleitner, 2009), or by the description of model bias (Del Giudice et al., 2013). Nevertheless, observation errors and model structural uncertainties are challenging to decompose with this approach since they are fundamentally unidentifiable by the model-measurement comparison although, some authors have attempted to integrate its identification by using highly informative priors (Renard et al., 2010).

The adequate description of error generating processes in many cases is still a challenge, and although analysis of residuals in hydrological simulation have received a substantial amount of attention (Ammann et al., 2018; Evin et al., 2013a; Schoups and Vrugt, 2010), other variables, which present a more complex response (e.g. dissolved oxygen, ammonium and other non-conservative water quality processes) appear to be less explored in the literature, in Chapter 3 of this thesis, a parametric inversion scheme for flow and dissolved oxygen concentrations in a water quality simulator is discussed.

The difficulty to propose adequate descriptions of the likelihood function led to the development of non-formal analysis techniques as GLUE (Beven and Freer, 2001), which employs user-defined objective functions and thresholds to update the parametric probability distribution and that has been extensively used in environmental and urban drainage modelling (Dotto et al., 2012; Freni et al., 2008; Thorndahl et al., 2008). Although this method has been highly controversial due to the impossibility to test the validity of subjectively chosen criteria (Mantovan and Todini, 2006). Also, Clemens (2001) discussed the use of frequentist calibration in urban drainage hydrodynamic models and the extraction of local sensitivities and identifiability of calibrated parameters by principal component analysis of the Jacobian matrix at the optimal parameter vector. This calibration procedure can be used to direct the redefinition of the model structure or the measurement layout.

1.5.3. UNCERTAINTY DECOMPOSITION BY SOURCE

Quantifying modelling uncertainty levels in environmental modelling is key for gaining confidence in the model performance and to communicate it to end-users. However, of even more importance from the scientific and operational perspective is the diagnosis of the uncertainty propagation structure. This intends to classify uncertainty sources by relevance, aiming to direct modelling efforts and further data acquisition to improve current knowledge and reduce modelling epistemic uncertainties.

Sensitivity analysis (SA) schemes (or parameter screening) are often used to identify relevant parameters (Cosenza et al., 2013), or input-parameter combinations (Pianosi et al., 2016). Global sensitivity analysis schemes classify a selected parametric space based on the magnitude of their effect on a certain output. SA requires the definition of ranges for the parameters under study. In uncertainty analysis those parametric ranges are derived formally as a way of conveying the degree of certainty in their value. However, SA is often purely based on the model response, neglecting the effect of structural uncertainties.

Willems (2012) proposed a variance decomposition approach. The variance decomposition takes into account the structural mismatch at the model-observations comparison, and thus classifies the importance of uncertainty sources in a quantitative manner. This method is nevertheless subjected to several challenges, since it assumes independence of error sources (which might be violated when dealing with heavily interlinked systems, i.e. see Freni and Mannina (2010b)) and requires homoscedastic residuals (thus often relying on data-transformations).

Gupta et al. (2008) also discussed the need for developing a *diagnostic approach* in modelling endeavours. Current model inversion methods are considered to have poor capability to pin-down and identify errors in the model description phase at particular subprocesses of the modelling scheme. Likelihood-based methods fail to capture all available information, since they narrow down the residual structure to a zero-dimensional metric. This approach suggests the benefits of using multi-metric approximated Bayes computation (ABC) to identify individual processes erroneously described, which render a systematic mismatch with observations in the system, and thus, helping to guide in the model improvement phase (Vrugt and Sadegh, 2013; Kavetski et al., 2018). However, insufficient examples are still available and metrics for water quality processes have not been proposed or tested.

1.5.4. TECHNICAL CONSTRAINTS IN THE APPLICABILITY OF UNCERTAINTY ANALYSIS METHODOLOGIES

Despite the availability of many strategies, they are generally not applied in practice. Sriwastava and Moreno-Rodenas (2017) and Tscheikner-Gratl et al. (2019) discussed aspects of this phenomena in detail. In general, there is a lack of incentive in consultancy firms to acknowledge the inconsistency and limitation of modelling endeavours and there is still a culture in the regulatory body of *acceptance* of deterministic (and sometimes even uncalibrated/unvalidated) realisation of modelling studies, which does not force modellers to fully justify their assumptions.

There are only a few academic examples on full uncertainty analysis for integrated catchment modelling applications. Schellart et al. (2010) presented a full-integrated catchment model to which uncertainties were identified and partially propagated (due to largely computational constraints). Radwan et al. (2004a) presented a probabilistic assessment of a partial ICM (simulating only the river submodel) study for a relatively small system (500 ha, 12,000 p.e.), discussing that error in rainfall inputs and water quality sewer elements dominated receiving water uncertainties. A similar study was presented by Freni and Mannina (2010b) in which the uncertainty contribution by source was evaluated in a small integrated urban system (115 ha, 9,000 p.e.).

This lack of studies is probably due to technical constraints in setting up such type of models; insufficient monitoring data, massive labour effort needed to set a full integrated modelling study, and computational limitations. Recent frameworks for uncertainty quantification in urban drainage (e.g. Deletic et al. (2012) and Tscheikner-Gratl et al. (2017)), have provided a common nomenclature and understanding of uncertainty quantification in the field. Also, Chapter 4 of this thesis provides an example for a formal uncertainty analysis scheme applied to a large-scale integrated catchment modelling study (4400 ha, 750,000 p.e.). Further examples are still needed to complement current scientific literature.

One of the severe limitations in the implementations of formal uncertainty analysis methods is the fact that modelling structures are computationally heavy (hours-days per simulation). And thus, sampling the model often falls beyond the practical capabilities of most research and practical users. Chapter 3 of this thesis is dedicated to the use of data-driven emulation schemes to accelerate the sampling of computationally expensive modelling platforms in uncertainty analysis schemes.

1.6. UNCERTAINTY SOURCES IN INTEGRATED CATCHMENT MODELLING

INTEGRATED catchment modelling aims simulating the link between different water quantity and quality subsystems (Rauch et al., 2002). Tscheikner-Gratl et al. (2019) provides a comprehensive review of most commonly accounted subsystems in urban ICMs, which are classified as: Rainfall-runoff and wash-off submodels, urban drainage, rural runoff routing, river physical-biochemical and wastewater treatment process modelling. Each of these submodels induces uncertainties, which propagate towards the output variables. Identifying and classifying uncertainty sources is necessary for directing model structure improvements, data acquisition and communicating the limitations of the modelling study.

1.6.1. UNCERTAINTY SOURCES IN SUBMODEL COMPONENTS

1.6.1.1. URBAN DRAINAGE SUBMODEL

This submodel encompasses the simulation of rainfall-runoff, transport and generation of sewage occurring in urbanised areas. Often, several subprocesses are linked; rainfall-runoff dynamics (e.g. evaporation, infiltration, accumulation, and overland flow), surface pollution wash-off (e.g. mobilisation of surface particulate matter), flow routing in the transport system (e.g. water infiltration, hydrodynamics), and in-sewer water quality processes (e.g. sewage generation, transport and transformations).

Deletic et al. (2012) reviewed in detail the main sources of uncertainty in urban drainage modelling and proposed a guideline for its systematic consideration. Korving and Clemens (2005) presented an uncertainty analysis in the simulation of CSO volume, discussing the effect of database errors. Also, Sriwastava et al. (2018) reported a sensitivity analysis in CSO volume estimations of a hydrodynamic modelling study, concluding that runoff-coefficient, roughness and the weir crest level parameters to be the most relevant ones, yet neglecting the effect of rainfall uncertainties. Cristiano et al. (2017) and Ochoa-Rodriguez et al. (2015) addressed the effect spatial and temporal resolutions of rainfall in urban hydrology, yet many aspects of this interaction remain poorly understood. In the Section 2.2 of this thesis, the effect of rainfall input characteristics in receiving water body DO concentrations in an ICM study is presented (Moreno-Rodenas et al., 2017b).

In general, uncertainties associated to the water quality routine of urban drainage modelling are considered to be significantly larger than those contained in water quantity processes (Mannina and Viviani, 2010; Sandoval et al., 2018; Willems, 2006). Furthermore, uncertainties in CSO pollutant concentrations are reported to be highly relevant when modelling water quality impacts in the receiving water bodies (Radwan et al., 2004a).

1.6.1.2. WASTEWATER TREATMENT PLANT SUBMODEL

Modelling wastewater treatment processes has reached a certain degree of maturity, and it is a standard in practice (Gernaey et al., 2004). Common WWTP simulators link highly detailed biochemical process descriptions (e.g. ASM2d, Gernaey and Jørgensen (2004)) with low detail flow simulations (often well-stirred tanks). Nevertheless, current model structures often render highly over-parameterised implementations, requiring robust calibration practices to minimise uncertainties in the process (Rieger et al., 2012; Vanrolleghem et al., 2003). Sensitivity analysis is often proposed to identify relevant parameters in WWTP operation (Benedetti et al., 2012; Sin et al., 2011). Belia et al. (2009) identified the most relevant sources of uncertainty in WWTP simulations as arising from: influent flow, influent pollutant loads, solids retention time, overflow rates, aeration and denitrification rates. Yet, uncertainties associated with change in regulations or operations are often dominant when predicting WWTP scenarios (Dominguez and Gujer, 2006).

1.6.1.3. RURAL HYDROLOGY SUBMODEL

The rural catchment often represents a significant proportion of the baseflow contribution to river bodies, and also agricultural exploitations are distributed sources for several pollutant classes. Thus, estimating the surface and sub-surface flow propagation can be highly relevant for the river water quality dynamics, which are often intensively influenced by water depth and volume of the stretch (Moreno-Rodenas et al., 2017a). McMillan et al. (2012) provide a comprehensive benchmarking of observational uncertainties in hydrology, for rainfall, river discharge and water quality measurements. In hydrological flow estimations, the effect of input resolution uncertainties has been reported as relatively low (Camargos et al., 2018), yet they can become more relevant when predicting pollutant flows (Chaubey et al., 2005).

1.6.1.4. RIVER PHYSICAL AND BIOCHEMICAL SUBMODEL

Beck (1987) presented a detailed review on the uncertainty analysis in river water quality modelling studies. River water quality models render highly complex simulation studies, mixing physical, chemical and biological processes, which occur at a large range of time and space scales. The IWA group for river water quality modelling presented the RWQM1 (Shanahan et al., 2001; Reichert et al., 2001; Vanrolleghem et al., 2001). Yet this type of biokinetically focused models often render highly parameterised structures, with largely un-identifiable processes. Reichert and Vanrolleghem (2001) discussed the identifiability and uncertainty propagation in several forms of river water quality modelling when accessing to data of several measured river parameters.

Lindenschmidt et al. (2007) showed that for several examples, structural uncertainties are mostly dominant in river water quality modelling studies. Moreno-Rodenas et al. (2017a) compared the effect of two river hydrological routing conceptualisations (i.e. a varying parameter Muskingum-McCarthy routing and a tank-in-series schemes), showing the effect of river flow submodel structural uncertainties in the calibration of river biochemical processes, yet the interaction of the selected hydraulic model complexity and the water quality processes are poorly understood. Nevertheless, the implementation of river water quality modelling is highly dependent in local characteristics. Thus, detailed study of the relevant dynamics and uncertainty sources in each case is recommended. Chapter 3 of this thesis contains an example for the practical application of sensitivity analysis and parametric inversion in computationally expensive river models. Also, Chapter 4 of this thesis shows that literature-elicited parametric uncertainties for river biological processes dominate uncertainties in a highly urbanised river catchment for the simulation of dissolved oxygen processes, thus highlighting the relevance of inferring river biochemical parameter values from local measured data.

1.6.2. UNCERTAINTY SOURCES IN THE MODEL DEFINITION AND OPERATION

1.6.2.1. UNCERTAINTIES DUE TO BOUNDARY CONDITIONS

During the model construction process, the system of interest has to be clearly identified. As depicted in Figure 1.3, the system is delimited by a defined boundary region. The modelling process represents system dynamics inside this boundary, and relevant external influences are considered modelling forcing data or inputs. Identifying these elements and characterising the knowledge of their value is relevant to quantify their influence in the output variables of interest. The nature of such inputs in ICMs is often time and space dependent, and thus their characterisation often involves the construction of spatio-temporal stochastic models (see examples for rainfall forcing data stochastic descriptions in Del Giudice et al. (2016), or the merging of data error sources in rainfall spatiotemporal estimations in Cecinati et al. (2018)).

1.6.2.2. UNCERTAINTIES IN THE EXTRAPOLATION OF SYSTEM DYNAMICS

Verifying the accuracy of the model structure representation of the real system is usually performed comparing observed and simulated dynamics through different distance metrics. This is often regarded as calibration, validation and verification of model structures. Nevertheless, modelling platforms are often used to assist in the decision-making process to test alteratives (i.e. testing the effect of virtual system changes, e.g. Langeveld et al. (2013b)). However, the performance of the virtual system configuration cannot be formally evaluated without experimental replication (which in environmental modelling is often impractical). Therefore, the plausibility and validity of model structure assumptions during extrapolation should be discussed and reported.

1.6.2.3. UNCERTAINTIES INDUCED BY THE SUBMODEL INTEGRATION AND THE NUMERICAL IMPLEMENTATION

One challenge in the creation of integrated models is the representation of submodel links. Submodel conceptualisations often simulate different state variables (Benedetti et al., 2013a) which should be transformed to generate an output to input link basis (e.g. different pollutant representations in sewer water quality and WWTP simulators). Also, models often render processes at different spatial or temporal resolution, requiring upscaling or downscaling processes in the submodel boundary (Heuvelink, 1999; Torres-Matallana et al., 2018).

Solver selection and settings analysis should be carried out, aiming to minimise its influence in the simulation process (e.g. see Benedetti et al. (2012)). Also, errors in the implementation, and post-processing of simulation outcomes should be considered.

1.7. THE DOMMEL WATER SYSTEM

THIS thesis contains multiple references to the Dommel water system in which a systematic evaluation of modelling uncertainties was performed. This section introduces its main characteristics, and will be therein referred in the text.

The Dommel is a natural stream, which runs through the north of Belgium (approx. 35 km) and the south of the Netherlands (approx. 85 km), tributary to the Dieze and the Meuse, which drain into the North Sea. It has a discharge ranging from 2-30 m^3/s. The catchment covers an area of more than 800 km^2 composed predominantly of forestry and medium intensity agricultural areas, which are naturally drained and characterised by mild slopes. In the Dutch section, the Dommel receives the discharge of several combined municipal drainage systems scattered around the city of Eindhoven (approximately 4400 ha of connected urban area). There are nearly 200 Combined Sewer Overflow structures (CSOs) along the Dommel and its tributaries. Additionally, the system receives the effluent of a large Wastewater Treatment Plant (WWTP) of 750,000 p.e. design capacity located in the east of the city of Eindhoven. WWTP and CSO discharges of partially treated or untreated wastewater are relatively common under heavy rainfall conditions, leading to the deterioration of the river ecological and chemical conditions. In particular, the biological breakdown of high organic content loads (exacerbated by low DO WWTP effluent) generates oxygen depletion processes in the river, which can lead to fish death.

Figure 1.4 presents a scheme of the water system of the Dommel. The graphic depicts the wastewater transport system, which is actively controlled and mixes gravity and pressurised pipe networks. Also, the location of rain gauge measuring stations and the relative location of the catchment with respect to the two KNMI meteorological Radar stations are provided.

The Waterboard of the Dommel envisioned a series of substantial investments (Benedetti et al., 2013b) aiming improve the ecological status of the river Dommel. A full-integrated catchment model was developed (Langeveld et al., 2013b) as an output of the Kallisto project (Weijers et al., 2012), aiming towards a better understanding of pollution dynamics in the system. This ICM study aimed to simulate the link between the WWTP, all relevant urban and rural contributing areas and the receiving water body (the Dommel and its tributaries), thus modelling the dynamics of dissolved oxygen and ammonium impacts in the river.

The system can be divided in three main sections, which were modelled seamlessly under the same model platform:

1. The urban drainage system is composed of 29 urban drainage networks. Each municipality is connected to the main WWTP through a controlled pipe network mixing pressurised and gravity sections. The influent of the WWTP has three separate pipelines; Nuenen-Son, Riool Zuid and the Eindhoven collector. There are more than 200 CSOs along the river Dommel and its tributaries. Several of those structures have stormwater settling tanks (SST) to reduce solid discharges. A lumped model structure was used to simulate each local urban drainage (conceptualised as in Solvi (2006)). This represents a hydrological structure for the rainfall-runoff response (accounting for wetting losses) and a tank-in-series routing scheme to simulate gravity and pressurised sewer transport. Event mean concentration vectors were used to link water quantity sewer outputs with CSO discharge to the receiving water body.

2. A WWTP with three biological lines with a combined capacity of 26,250 m^3/h (depicted in Figure 1.5). Each line consists of; one primary settler, a biological treatment tank and four secondary settling tanks. Additionally the plant has a by-pass storm settling tank structure with an extra capacity of 8,750 m^3/h for storm water treatment. The WWTP was simulated with an ASM2d biokinetic model (Gernaey and Jørgensen, 2004). The calibration procedure is depicted in Langeveld et al. (2013b). The link between in-sewer water quality and the WWTP influent was represented by an empirical influent model (Langeveld et al., 2017).

3. The river Dommel and its local tributaries. With an approximate length of 120 km in the area of interest. The river system was modelled through the use of 65 tank-in-series well-stirred sections. This conceptualisation is a hydrological approximation of the flow propagation process. The physical and biochemical reactions relevant for dissolved oxygen dynamics were encoded following the Duflow dissolved oxygen routine. This includes oxidation of two organic matter fractions,

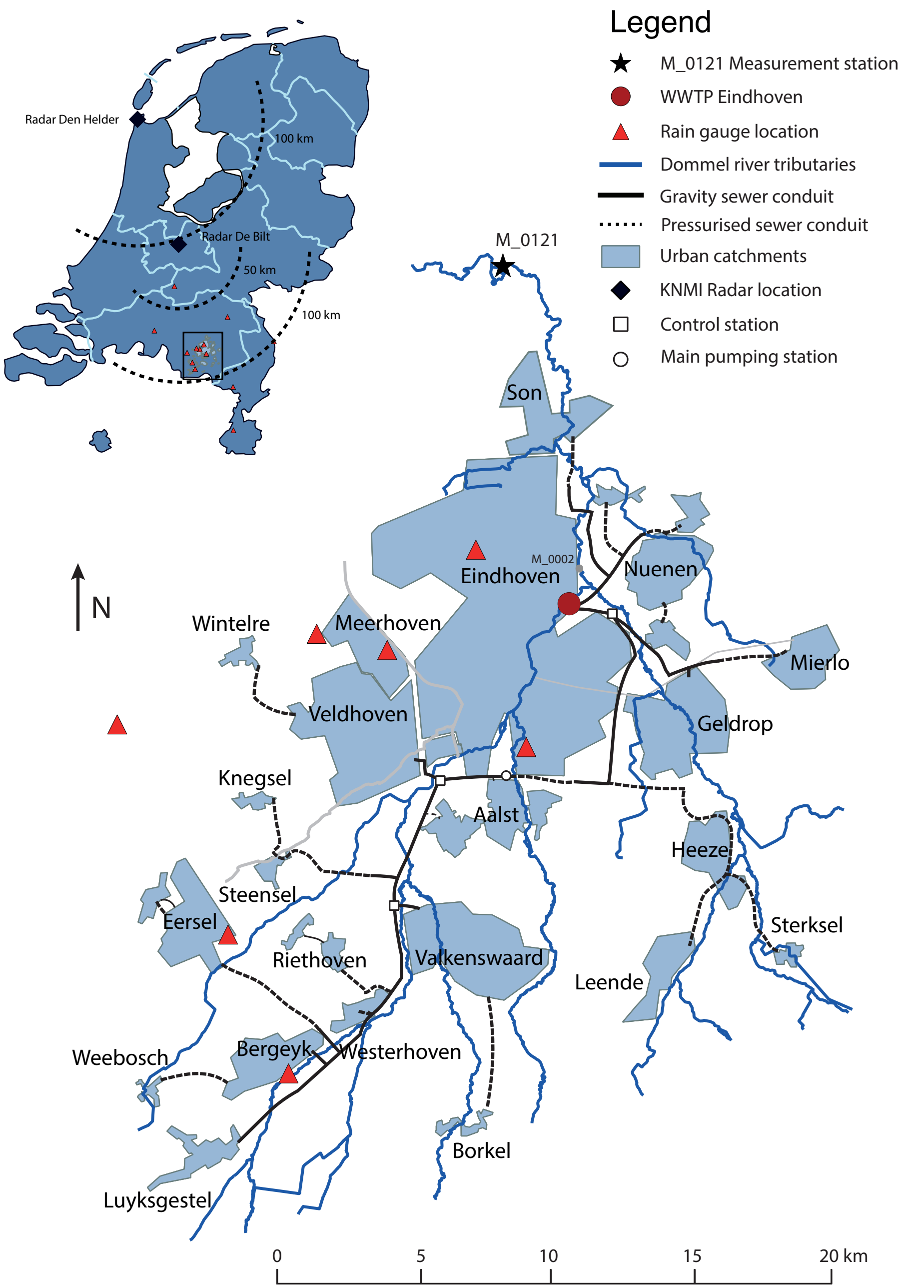


Figure 1.4: General scheme of the Dommel water system from the Belgium border to Sint-Oedenrode depicting rainfall measurement stations, the main municipal drainage and wastewater transport system

sedimentation, nitrification, reaeration and macrophyte oxygen exchange. The processes accounted in the river model can be seen at Appendix B.

The full-integrated model was built using WEST (DHI) simulation platform for ICMs. A detailed description of the integrated catchment model structure and its development and validation process can be found at Langeveld et al. (2013b). A description of the main spatial and temporal characteristic scales in the Dommel system was reported in Moreno-Rodenas et al. (2017b) and can be found in Chapter 2 of this Thesis.

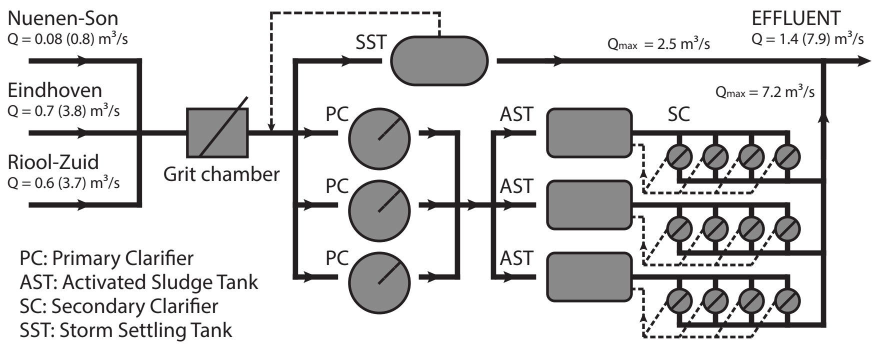


Figure 1.5: Scheme of the Eindhoven wastewater treatment plant. Flows are given by its median value and in parenthesis the 99% percentile from measured data series (2011-2013)

1.8. OBJECTIVES AND OUTLINE OF THE THESIS

THE formal consideration of uncertainties in real scale integrated water quality systems is still seldom performed in practice. There is still a lack of bibliographic sources covering uncertainty analysis studies and a lack of understanding on the interaction of many error sources in urban water quality studies. Consequently, the European Union financed the QUICS (Quantifying Uncertainties in Integrated Catchment Studies) project (FP7-ITN No.607000), which represents an effort to increase current knowledge about the nature and magnitude of error sources for the modelling of catchment scale water quality processes. This thesis, part of such endeavour, targets studying the applicability of uncertainty analysis strategies for large-scale water systems. The main goal is to identify relevant error sources, quantify their propagation across the integrated catchment model and identify the main contributors for the uncertainty of modelled dissolved oxygen dynamics in integrated water systems.

Chapter 2 provides the identification of the main sources of uncertainty for an ICM study of the river Dommel. This chapter contains the analysis of sub-model driven uncertainty sources and the detail description of two error sources. Firstly, the effect of spatial and temporal characteristics of urban rainfall forcing data is described for the simulation of dissolved oxygen dynamics in the river Dommel. Secondly, the effect of uncertainties in combined sewer overflow pollutant concentration in the receiving water quality simulations. Two stochastic processes were used to describe uncertain knowledge about sewer-to-river link variables. The impact of neglecting the correlation of the different variables is discussed, presenting a methodology to draw stochastic CSO pollutant variables respecting the observed correlation structure.

The numerical implementation is still one of the challenges of uncertainty analysis schemes in real-scale integrated catchment studies. This type of models often result in highly computationally expensive structures and sampling the simulator becomes impractical. The use of model emulation techniques to accelerate model sampling is discussed in Chapter 3. In a first example, the use of an emulator platform is described for the implementation of a global sensitivity analysis and a formal Bayesian inference to reduce parametric uncertainty in flow and dissolved oxygen dynamics for the river Dommel. Additionally, one of the main constraints of today's emulation state-of-the-art is the dimensionality curse. This forces the number of parameters used in the emulation process to be reduced, limiting the consideration of dynamic inputs. This is a severe limitation in cases in which an emulator platform is required to extrapolate to different dynamic conditions from the training scheme (e.g. an emulator to accelerate the simulation of flooding schemes under different storm conditions). In Chapter 3 a novel methodology to integrate the use of rainfall dynamics and static parameters for the emulation of 2-D flow propagation schemes is described.

Chapter 4 discusses the application of a statistical uncertainty analysis to a full-integrated catchment model. This involved quantifying the degree of uncertainty in modelling outcomes when simulating the current system status. It also provides a decomposition of the effect of several sources of uncertainty involved in the simulation of dissolved oxygen dynamics in the system. Knowledge on the relevance of the source of uncertainties is important to direct further monitoring and modelling efforts.

Chapter 5 presents the conclusions of this thesis, reflecting about the current problems of uncertainty quantification in integrated catchment modelling along with recommendations for future research.

2

Characterisation of uncertainty sources in integrated urban catchment modelling

2.1. Reporting modelling uncertainties [*]

This chapter briefly presents the application of the QUICS framework for reporting modelling uncertainties (Tscheikner-Gratl et al., 2017) to an ICM study for the simulation of dissolved oxygen dynamics in the Dommel water system. Furthermore, the detailed characterisation of two relevant uncertainty sources for the simulation of receiving water quality impacts is shown. Firstly, the effect of the spatio-temporal characteristics of urban rainfall input data in river dissolved oxygen dynamics is discussed. Secondly, the probabilistic description of combined sewer overflow pollutant concentrations is presented and the effect of neglecting the correlation structure observed in measured discharged pollutant fractions is shown.

2.1.1. Objectives of the modelling study

Clearly stating modelling objectives is key to effectively select the adequate structure, identifying necessary data and stating required tolerances in the simulation process (Refsgaard et al., 2007; Laniak et al., 2013)

The Dommel ICM study targets the simulation of dissolved oxygen dynamics at the receiving water body level. The model application involves the estimation of the effect of system changes (e.g. adding a storm water storage facilities, reducing particulate matter in CSOs etc.) on the frequency and magnitude of dissolved oxygen concentration water quality impacts. This implies that the simulator should closely represent the current system state (which can be assessed by comparing simulated and measured variables), yet

[*]This chapter section is partially adapted from: Tscheikner-Gratl, F., Lepot, M., Moreno-Rodenas, A., Schellart, A. (2017). QUICS report: A Framework for the application of uncertainty analysis. doi: 10.5281/zenodo.1240926

also should be valid to extrapolate system mechanistic relationships (e.g. virtual scenarios), which cannot be formally validated without physically changing the system.

The water quality status is assessed by concentration-duration-frequency (CDF) metrics (Appendix A), which limits are derived from a local ecological assessment (FWR, 2012). The performance of the simulator is also evaluated by comparison with measured time-series of flow, depth and dissolved oxygen at different locations of the system. Langeveld et al. (2013a) and Langeveld et al. (2013b) provide further detail in the construction of the modelling platform and the required data sources for its calibration and validation process.

2.1.2. Identification of uncertainty sources

During the model development phase, uncertainties at each stage should be reported and classified. Figure 2.1 provides the main sources of uncertainty in the Dommel ICM study. Uncertainties are reported following the classification depicted by Refsgaard et al. (2007), attending to the nature, type and source of modelling uncertainties.

Along with the identification of uncertainty sources, a pre-screening analysis is recommended to rank the most influential and uncertain factors, thus prioritising modelling efforts. This pre-screening is nevertheless subjective and based on the modeller's experience, since ranking formally many sources of uncertainty (e.g. deep uncertainties) is not straightforward, yet it is useful to report and discuss modelling abstraction assumptions. Tscheikner-Gratl et al. (2017) suggested the use of elicited uncertainty prioritisation graphical panels by showing an estimated sensitivity and uncertainty magnitude for each source. An example of the application of this prioritisation can be found in Figure 2.2, in which an elicited ranking was proposed for the main identified sources.

The estimated prioritisation was used to report the possible effect of the main sources and to discuss the actions taken to reduce uncertainties in the output variables inline with the model construction phase. Some of these efforts are hereby presented:

1. River upstream pollution loads (source 3): Water quality dynamics in the upstream river boundaries are not included in the modelling scheme, but are represented by a constant concentration pollutant load. This simplification was expected to impact the modelling of dissolved oxygen concentrations. Therefore the Waterboard of the Dommel carried out a dedicated monitoring campaign and verified that expected pollutant loads in the boundary have a reduced influence in the downstream sections.

2. Sediment evolution (dredging and transport) (source 9): The sediment bed represents a relevant oxygen sink term in the river Dommel. This effect is modelled through a constant sediment oxygen demand term and through a dynamic sedimentation and consumption of particulate BOD matter (Appendix B). However, dredging and transport of sediments are not appropriately modelled in the system, thus non-stationarities in the process might not be well captured, potentially rendering a model structural mismatch.

3. Errors at measured water quality data (source 10): Measurement conditions in natural water bodies are not ideal and constant maintenance is required. Errors in

reported series for water quality variables in the river are expected and this can influence calibration and validation processes. Detailed assessment of data quality is carried out to minimise this effect (see Schilperoort (2011)).

4. Urban drainage CSO mean pollutant concentration vectors (source 18): Sufficient data is not available for the calibration of local in-sewer water quality process models. Thus the model uses mean concentration pollutant vectors to estimate sewer loads into the receiving water body. This is expected to result in highly uncertain input loads, which effect should be quantified.

5. Rainfall data input characteristics (space-time resolution) (source 12): Selection of sources for rainfall data and the characteristics of the inputs are expected to influence the dynamics at the urban drainage processes, affecting timing and volume of urban discharges.

This chapter contains two examples in which the effect of rainfall data input characteristics and the impact of uncertainties in CSO mean pollution concentration vectors are explored. Characterising the effect of statistical uncertainty sources in the simulated dynamics of dissolved oxygen is also further discussed in Chapter 4, along with a formal ranking of the sensitivity of the model output. A dedicated study should be carried out when using the model under extrapolation conditions to evaluate the effect of scenario uncertainties.

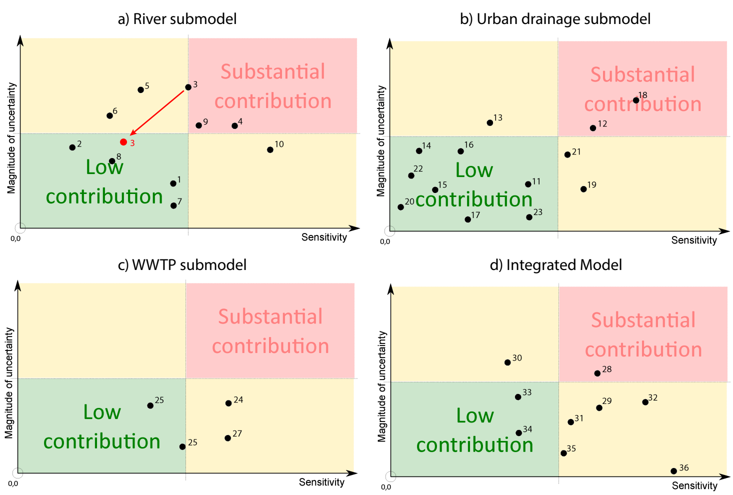


Figure 2.2: Prioritisation panels for the three main submodels and the full version of the Dommel ICM (Tscheikner-Gratl et al., 2017). Points refer to the uncertainty sources depicted in Figure 2.1.

System	id	Uncertainty sources	Source					Type				Nature		
			Context	Model Input	Parameter	Calibration	Model structure	Determinism	Statistical	Scenario	Deep Uncertainty	Epistemic	Stochastic	Ambiguity
River	1	Temperature River (measurement)		■					■				■	
	2	Luminosity River (measurement)		■					■				■	
	3	River upstream pollution (estimated)		■					■					■
	4	Baseflow hydrology					■		■			■		
	5	Pollution load rural catchment					■		■			■		
	6	River diversion/retention structures levels					■	■				■		
	7	River geometry		■					■			■		
	8	River energy losses/roughness			■				■			■		
	9	Sediment evolution (e.g. dredging and transport)					■				■	■		
	10	Errors in measured water quality data		■					■				■	
Urban drainage	11	Rainfall Data measurement errors		■					■				■	
	12	Rainfall input time-space resolution					■		■				■	
	13	Soil characteristics for infiltration			■				■				■	
	14	Water infiltration in the sewer				■			■				■	
	15	Evaporation potential		■					■				■	
	16	Daily/seasonal pattern urban pollution load		■					■				■	
	17	Population density	■							■		■		
	18	Urban drainage CSO pollution mean concentration			■				■				■	
	19	Pumping capacity-activation levels	■							■		■		
	20	Georeference of main CSO structures					■	■				■		
	21	CSO weir geometry					■	■				■		
	22	Layout of connected draining areas					■			■		■		
	23	Control set points in transport system				■				■		■		
Wastewater treatment plant	24	WWTP reactors dynamic state				■			■				■	
	25	Temperature water WWTP		■					■				■	
	26	Control WWTP			■			■				■		
	27	Water treatment chemical addition			■			■				■		
Integrated Model	28	Model structure extrapolation	■							■		■		
	29	Climatological scenarios	■							■		■		
	30	Urban to WWTP link state-variable transformations					■		■			■		
	31	WWTP to River link state-variable transformations					■		■			■		
	32	Changes of environmental criteria and legislation	■								■	■		
	33	Technological changes	■								■	■		
	34	Change in urbanisation structure	■								■	■		
	35	Change in land usage (agriculture-industrial)	■								■	■		
	36	Solver settings				■		■				■		

Figure 2.1: Classification of main sources of uncertainty for the Dommel ICM study (adapted from Tscheikner-Gratl et al. (2017)).

2.2. IMPACT OF SPATIOTEMPORAL CHARACTERISTICS OF RAINFALL INPUTS ON INTEGRATED CATCHMENT DISSOLVED OXYGEN SIMULATIONS *

2.2.1. INTRODUCTION

RAINFALL is one of the main driving forces for water pollution dynamics in urbanised catchments. Combined urban drainage systems often generate discharges from partially treated WWTP effluents and combined sewer overflow structures (CSO) under severe storm conditions. These discharges impact on the chemical and ecological status of the receiving surface waters. When modelling water quality processes, it is often necessary to represent urban pollution loads, which are heavily linked to rainfall-runoff dynamics. Several studies have addressed the effect of rainfall input errors in the behaviour of individual urban drainage systems for hydrodynamic modelling. For instance, Sun and Bertrand-Krajewski (2013) reported that the uncertainties in rainfall input data have a lower contribution to the total uncertainty compared with flow measurement errors for the simulation of discharges in a small urban drainage system (180 ha). Ochoa-Rodriguez et al. (2015) presented a comprehensive sensitivity analysis of urban drainage flow simulations (urban system with areas ranging between 200-800 ha) for varying rainfall spatio-temporal resolutions. This showed that the impact of rainfall errors rapidly decreases with catchment's area, and reported a strong link between temporal and spatial effects. The impact of rainfall dataset selection in the calibration of urban hydrodynamic models is recognised as significant (Tscheikner-Gratl et al., 2016).

Nevertheless, when modelling receiving water quality dynamics, urban drainage simulators are often conceptualised using spatially lumped model structures (Langeveld et al., 2013b; Vezzaro et al., 2014a). Lumped urban rainfall-runoff models reduce significantly the simulator's computational cost, which allow for long simulation periods and fast evaluation. Also, accuracy at modelled outflow variables is expected to be sufficient for the application at large-scale water quality modelling and real time control of drainage systems (van Daal-Rombouts et al., 2016). These lumped model structures further distort the effect of uncertainties in rainfall measurement data, especially those associated with the in-catchment micro spatial and temporal scales. Also, urban wastewater systems are often composed by several connected sewer networks across municipal areas (e.g. centralised wastewater treatment layouts), and the distances between those connected catchments are often beyond the de-correlation length of convective storm events (Bruni et al., 2015).

Literature covering the effect of rainfall input characteristics at integrated catchment scale and their impact on water quality dynamic simulation is scarce. The total contribution of rainfall errors in water quality modelling uncertainties is expected to be moderate (Willems, 2012), however it is important to characterise the desired properties of rainfall input datasets when modelling water quality processes at large scale urban systems.

This subchapter investigates the effect of the spatiotemporal variability of selected rainfall inputs on the simulation of dissolved oxygen dynamics in the Dommel urban wa-

*This section is an adapted version of: Moreno-Rodenas, A.M., Cecinati, F., Langeveld, J. and Clemens, F.H.L.R. (2017) Impact of Spatiotemporal Characteristics of Rainfall Inputs on Integrated Catchment Dissolved Oxygen Simulations. Water 9(12), 926. doi: 10.3390/w9120926 ter system. It aims to ascertain whether uncertainties induced by the misrepresentation of the rainfall field at the municipal scale are transferred to the simulation of dissolved oxygen dynamics at a sensitive river section. The performance of a large-scale conceptual integrated catchment model was compared against monitoring data. This provided insights on the optimal selection of rainfall data sources in large lowland urban water systems for dissolved oxygen environmental assessment.

2.2.2. Materials and Methods

2.2.2.1. System and data description

River data (water-depth, velocity and Dissolved Oxygen (DO) concentration) were obtained at a monitoring station (M0121, as shown in Figure 1.4) located at roughly 17 km downstream of the WWTP. DO measurements were taken by an online probe located in the river, with a sampling frequency of 10 minutes. The measuring station was periodically maintained and calibrated. Manual data inspection was performed to validate time series by the water authority of De Dommel. Time series for combined sewer overflow (CSO) discharges were estimated from water depth monitoring data at the main structures (although those estimations can contain significant level of uncertainties). The WWTP operator provided water quantity and quality data of the influent-effluent. Three years of data were obtained (2011-2013), more information about the data set can be found at (Langeveld et al., 2013a).

Measured rainfall data were obtained from three sources (Table 2.1): 1) An automatic rain gauge network from the Dutch Meteorological agency (KNMI). One of the stations was installed close to the civil airport of the city of Eindhoven (KNMI_370 in Figure 1.4), and six more were found within a radius of 70 km from the centre of catchment (those rain gauges add relevant information to constrain the interpolation scheme at the boundary areas). KNMI automatic rain gauges use a floating device and an electronic register, providing a high accuracy, measuring frequency and resolution (1% of the rainfall rate - 1 min - 0.02 mm/h). These stations are calibrated and maintained regularly. 2) A local network of tipping bucket rain gauges (resolution 0.15 mm). This is composed of six stations managed by the Waterboard de Dommel and the municipality of Eindhoven. Rain gauge stations are located within the urban area of several municipalities (Figure 1.4). 3) Single polarisation C-Band radar rainfall estimations from the composite of the KNMI. Rainfall estimated maps are provided at a resolution of 1 km^2 and 5-minute accumulation. The radar based estimates were also bias-corrected by the KNMI (from 3-hour and 1-day rain gauge accumulated data) in order to eliminate the systematic errors present in the raw radar data (Overeem et al., 2009).

The urban system connected to the studied section of the river Dommel is composed of 29 main contributing urbanised areas, which are characteristic of a lowland area with high in-sewer storage and low slopes. Catchment areas were extracted manually from a land-use GIS database. This catchment comprises a rural area of approximately 800 km^2, which contributes to the baseflow of the river and is mainly composed by forestry and low intensity agricultural areas. The system is characterised by mild slopes and it is naturally drained. Section 2.2.2.5 and 2.2.2.6 contain further description of the system under study, attending in detail to the relevant characteristic spatiotemporal scales.

Table 2.1: Rainfall data sources and measurement characteristics

Source	Sampling frequency	Spatial Resolution	Remarks
WB Dommel Rain gauges	5 min	Point data (6 gauges)	Tipping buckets, 0.1-0.15 mm resolution gauges
KNMI Rain gauges	10 min	Point data (7 gauges)	High quality maintenance, weighting gauges 0.01 mm resolution
KNMI Radar	5 min	1 km^2	KNMI areal Radar-based precipitation estimation C-Band single polarisation

2.2.2.2. MODEL STRUCTURE

An integrated catchment model (ICM) was developed with the purpose to simulate water quality dynamics (dissolved oxygen and ammonium) in the river Dommel. A previous version can be found at Langeveld et al. (2013b) along with a detailed description of the monitoring data needed for its development and the calibration process (in a data set from 2001-2010). The ICM consists of: a) An urban drainage routine, which was conceptualised as a set of lumped rainfall-runoff models representing the 29 urban systems of the Dommel area. A tanks in series scheme served to simplify the hydraulic routing network to the WWTP. A total of 192 CSO structures were reduced to 30 clusters representing the spatial locations of the most relevant discharge structures. b) A WWTP model, the characteristics and calibration of which are described in detail at Langeveld et al. (2013b). This model represents three biological lines and a storm-bypass section. The link of state variables at the boundary urban drainage – WWTP were produced using an stochastic generator (Langeveld et al., 2017). c) A river sub-model, which is composed by a flow propagation scheme, modelled through a hydrological storage-discharge model (Tank-In-Series scheme). This was calibrated based on river flow and depth measurements during 1 year (2012) using CSO and WWTP discharge measured data and hydrological derived base-flow as inputs. The calibration scheme produced a Nash-Sutcliffe efficiency at calibration/validation of 0.92/0.84 for river flow dynamic series. Further description of the calibration process can be found at Moreno-Rodenas et al. (2017a). The water quality routine was conceptualised assuming completely stirred reactor-like river sections (with section length between 800-3000 meters).

Figure 2.3 provides a scheme of the link between subsystems and the processes accounted for in the water quantity and quality routines. Within the river model section a graphical scheme summarises the water quality processes used to simulate dissolved oxygen dynamics. This included the balance between a three-phase layout (atmosphere, water volume and sediment), fractionation of Biological Oxygen Demand (BOD) in dissolved/particulate and fast/slow biodegradability, oxygen exchange from macrophyte biomass (Primary Producers) and nitrification-denitrification. The main source term in the DO dynamic balance is the reaeration process (KLT), which depends on water turbulence, depth and temperature. Meanwhile, consumption is generated by oxidation of fast and slow biodegradable matter in the sediment/suspended fraction (kd1, kd2 and kBODs) and from the nitrification process (knit). The sedimentation of organic matter was also considered (Vs1 and Vs2). Appendix B contains the full description of biological processes simulated in the river section.

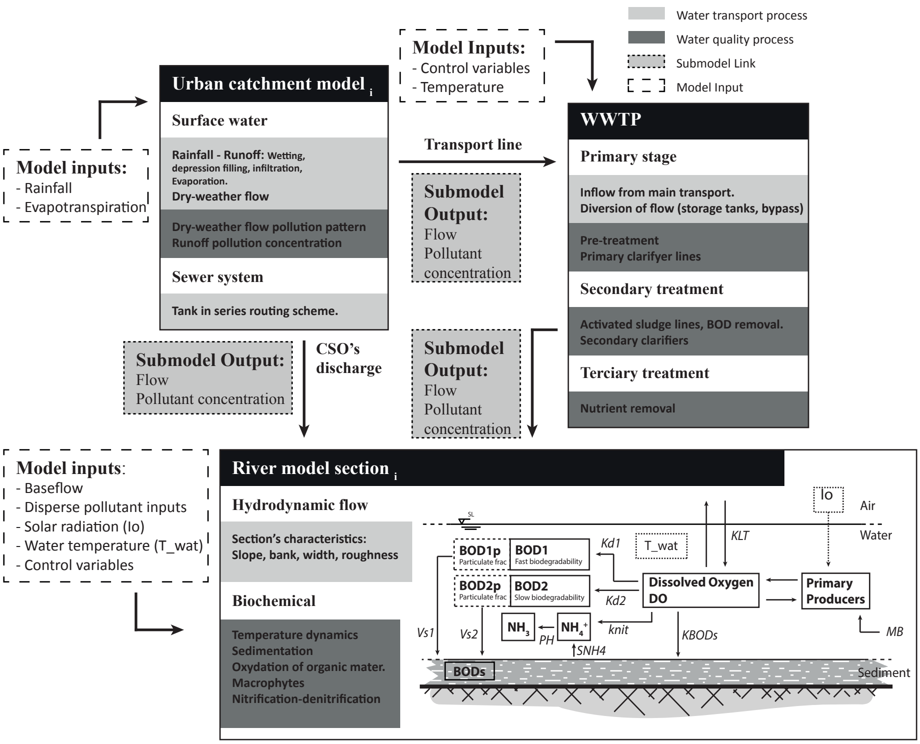


Figure 2.3: Scheme of submodel links and processes

2.2.2.3. STORM SELECTION

Three periods of summer were selected for a continuous modelling time-window: period 1) 10/08/2011 – 31/08/2011 (21 days), period 2) 05/07/2012 – 04/08/2012 (30 days) and period 3) 25/07/2013 – 19/08/2013 (25 days). Summer storm impacts are considered critical for the dissolved oxygen content in urban rivers since these events often couple several critical factors; high temperatures lead to a lower oxygen saturation point at the water mass (which consequently produces a lower initial DO concentration during the event), to faster biodegradability rates (in the suspended and sedimented fractions) and low base-flow level, which reduces the buffering capacity of the river.

As water quality-related processes exhibit a significant inertia (e.g. biological masses at the waste water treatment works can take stabilisation times up to weeks-months), the initial conditions for each simulated period were extracted from one-year of previous continuous simulation (using the rain gauge KNMI370 as an homogeneous urban rainfall input).

A total of 7 rainfall events were selected within the three periods (which induced oxygen levels at the river monitoring station lower than 3 mg/l). A summary of the storm characteristics can be found in Table 2.2. Rainfall maximum intensity and accumulated depth were calculated using five rain gauge stations located within the boundaries of the Dommel system (maximum distance-to-centre, 18 km). Rainfall variability and measuring errors induced dispersion in the chosen parameters, thus the mean value and one standard deviation are provided as $\hat{r}(\pm\sigma)$. Visual and quantitative assessment of rain gauge time series did not show systematic deviations in the measurements.

Table 2.2: Storm characteristics. Rainfall volume and maximum intensity computed from 5 rain gauges within the Eindhoven catchment. ADWP stands for antecedent dry weather period length. D refers to the duration of the rainfall in minutes. All variables were calculated from the 10 minutes accumulated time series

Event Code	Period	Storm time window	Rainfall depth (mm)	Max intensity (mm/h)	D (min)	ADWP (days)
1	1	18-08-2011 14:00 18-08-2011 21:00	7.9 (±5.4)	3.1 (±2.1)	170	4
2	1	22-08-2011 21:00 23-08-2011 14:00	30.8(±10.1)	17.8 (±6.9)	270	3
3	1	26-08-2011 06:00 26-08-2011 14:00	7.1 (±2.2)	5.8 (±2.1)	100	2.5
4	2	11-07-2012 10:00 11-07-2012 22:00	24.7 (±12.3)	7.6 (±3.8)	320	2
5	2	28-07-2012 17:00 28-07-2012 20:00	9.6 (±3.0)	7.9 (±2.0)	30	3.5
6	3	27-07-2013 05:00 28-07-2013 03:00	33.5 (±6.4)	17.6 (±5.4)	220	2
7	3	11-08-2013 11:00 11-08-2013 23:00	11.4 (±1.9)	6.5 (±0.8)	150	4

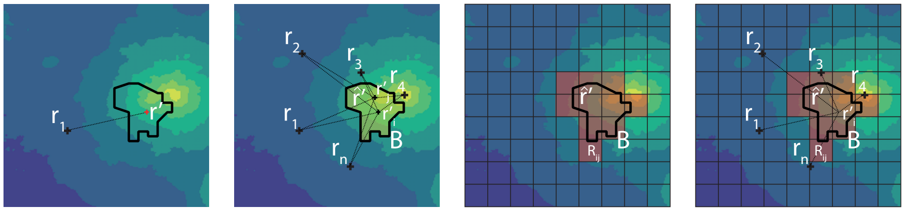
2.2.2.4. GENERATION OF RAINFALL ESTIMATIONS

Rain gauge and disdrometer networks are usually considered the most reliable data source for rainfall measurements at urban scale. However, those networks present low spatial densities, thus the representativity of point samples decreases when extrapolating measurements to ungauged locations. On the other hand, weather radar rainfall estimations can represent the spatial structure of storm processes (at the cost of lower accuracy). When possible, both data sources should be merged to provide rainfall measurements respecting both point rain gauge measurements and spatial structures contained in radar estimations (Nanding et al., 2015; Velasco-Forero et al., 2009).

The allocation of rainfall inputs from measured locations to the model of an individual draining area is generally done based on the modeller's expertise. The modeller defines a time-accumulation scale in which the rainfall estimations will be generated. This time step definition is constrained by data technical specifications (lower limit) and by the dynamics of the modelled process (upper limit). Manual/automatic quality checks are performed to assign reliable time-windows to each dataset. Then a selection of those measured points is used to generalise values to the rest of the catchment. This is often done by spatially interpolating sampled data by using Thiessen polygons, distance-averaging schemes or through geostatistical methods. A geostatistical interpolation scheme was used since it allows accounting for the spatial structure of the rainfall process and it facilitates the merging of different data sources.

When generating rainfall predictions, an often-neglected factor is the spatial support of the estimation. In this case, the studied ICM was composed by several individual lumped rainfall-runoff model structures. These models spatially aggregate the internal connected areas. Thus, the rainfall input should be representative for the full domain and not only of an arbitrary internal location. Although measurements provided by rain gauge networks are of point nature, rainfall predictions should be performed over a certain area. This is often referred to as change of block spatial support (Heuvelink, 1999).

Figure 2.4 depicts the strategies followed to generate spatial rainfall predictions by interpolating the different data sources: a) Direct application of a single rain gauge measured series to all subcatchments in the system. b) Rainfall estimations at area (B) block support of each subcatchment are rendered from a block kriging scheme from n point



(**a**) Single Rain-Gauge (**b**) Block kriging n RGs (**c**) Radar averaged (**d**) Radar + RGs, KED

Figure 2.4: Rainfall estimations under change of spatial support from different data sources; (a) Single rain gauge data source (direct application) (b) kriging estimation from n point measurements with change of block support for the area B, areal estimation (c) weighted average of radar map (R) for the area B (d) merged product between n rain gauges and the radar map through a kriging with external drift estimation under change of spatial support.

samples. c) The estimated rainfall is extracted from a weighted average of the radar pixels covered by the catchment. d) Radar measurements are used as an external covariate within the area support to refine an interpolation from n rain gauges (point-to-block kriging with external drift). The simplification of spatial rainfall description through point or block estimates will not always contribute significantly to the rainfall predictive accuracy. Its influence depends on the spatial variability of the observed field and the size of the support area. In other words, a highly correlated spatial field would be well approximated by a point estimate. However, spatial variability of rainfall may become increasingly relevant at large draining areas subjected to highly convective storm conditions.

Rainfall input time series were generated using four data sources:

1. Single rain gauge (*BK1_T*): Using only the most reliable rain gauge (KNMI_370 in Figure 1.4). This rain gauge was selected to generate spatially homogeneous rainfall fields, which simulates the case in which data is only present at a single location inside the catchment area.

2. Block kriging of all available rain gauges (*BKall_T*): Use of the full rain gauge network to generate predictions from an Ordinary Kriging model at the spatial support of each draining area.

3. Averaged radar quantitative estimation (*ARadar_T*): Spatially weighted average of the gridded radar-derived rainfall estimations (resolution 1 km^2) for each draining area.

4. Kriging with external drift (*UBK_T*): Use of the full rain gauge network and radar rainfall estimated field as a covariate. Predictions were rendered at spatial support of urban drainage systems.

Rainfall input time series were generated by varying the accumulated time step T (at 10, 30 and 60 minutes) for each of the four data sources, thus generating 12 different products for each of the 29 individual rainfall-runoff model catchments.

2.2.2.4.1. ORDINARY KRIGING WITH CHANGE OF SPATIAL SUPPORT

Kriging is a well-known geostatistical technique (Matheron, 1971) widely used to generate estimations of continuously spatially distributed variables at non-sampled locations, taking into account the spatial structure exhibited by the variable.

Let, $r(x_0)$ be the rainfall intensity at a certain non-sampled location x_0. This can be approximated by a linear combination of, $i = 1,...,n$ observed values at nearby locations, $r(x_i)$:

$$\hat{r}(x_0) = \sum_{i=1}^{n} [w_i(x_0) \cdot r(x_i)] \tag{2.1}$$

where w_i denotes a set of local weights. Ordinary Kriging stems from the assumption that the process is Gaussian and has a constant and unknown mean value or $E[r(x_0)] = E[r(x_i)] = m$. Forcing the estimation to be unbiased, $E[\epsilon(x_0)] = E[r(x_0) - \hat{r}(x_0)] = 0$ and minimising the prediction's variance leads to the derivation of the kriging system, which in matrix form is:

$$\begin{bmatrix} \mathbf{w} \\ \mu \end{bmatrix} = \begin{bmatrix} \mathbf{C}_{ij} & \mathbf{1} \\ \mathbf{1} & 0 \end{bmatrix}^{-1} \cdot \begin{bmatrix} \mathbf{C}_{i0} \\ 1 \end{bmatrix} \tag{2.2}$$

in which μ refers to the Lagrange multiplier (used on the minimisation of the variance), $\mathbf{C}_{ij} \in \mathbb{R}^{nxn}$ to the covariance matrix between data points, $\mathbf{C}_{i0} \in \mathbb{R}^{nx1}$ covariance vector between each of the measured points at the estimation's location and $\mathbf{1} \in \mathbb{R}^{nx1}$ an all-ones vector.

Block kriging usually denotes a variation of the kriging system in which the target is not a point estimate but a spatially averaged prediction within a certain area domain, B. This is computed as:

$$r(B) = \frac{1}{|B|} \sum_{j|j \in B} r_j \tag{2.3}$$

in this case, the covariance structure changes to the point-to-block covariance, or:

$$\tilde{C}_{iB} = cov(r_B, r_i) = E[r_B r_i] - E[r_B] \cdot E[r_i] = \frac{1}{|B|} \sum_{j|j \in B} C_{ij} \tag{2.4}$$

which represents the average covariance between the location i and all the possible locations within the block B. Thus the kriging system accounting for the estimation at block level is:

$$\begin{bmatrix} \mathbf{w}_B \\ \mu \end{bmatrix} = \begin{bmatrix} \mathbf{C}_{ij} & \mathbf{1} \\ \mathbf{1} & 0 \end{bmatrix}^{-1} \cdot \begin{bmatrix} \tilde{\mathbf{C}}_{iB} \\ 1 \end{bmatrix} \tag{2.5}$$

by solving $\mathbf{w}_B(x_B)$ and using it as the weights for the Equation 2.1, it is possible to generate estimations at the target area support.

2

2.2.2.4.2. KRIGING WITH EXTERNAL DRIFT (KED) AND CHANGE OF SPATIAL SUPPORT

Ordinary kriging is analogous to a Gaussian process model with an unknown but constant mean field. However, additional variables that are correlated to the targeted process (and extensively measured) can be used to further refine the spatial model. Regression kriging and Kriging with external drift formally account for this case. This involves assuming a spatial model, which has a deterministic mean ($m(x)$) and a stochastic component ($\epsilon(x)$):

$$r(x) = m(x) + \epsilon(x) \tag{2.6}$$

$$m(x) = \sum_{k=0}^{p} \beta_k \cdot q_k(x) \tag{2.7}$$

and in which the mean field is linearly related to a set of $k = 1, \ldots, p$ external variables $q_k(x)$ through a vector of weights β (also referred as regression coefficients). In Regression Kriging, the mean field is first fitted, and the field of stochastic residuals $\epsilon(x)$ is interpolated though Simple or Ordinary Kriging. On the other hand, Kriging with External Drift (KED) includes the regression process within the solution of kriging system. Thus, predictions on KED are also performed as a linear combination of the values at observed locations:

$$\hat{r}(x_0) = \sum_{i=1}^{n} w_i^{KED}(x_0) \cdot r(x_i) \tag{2.8}$$

in which the weight values are obtained by solving the KED equation system $\forall i \in 1, \ldots, n$:

$$\begin{gathered} \textstyle\sum_{j=1}^{n} w_i^{KED} \cdot C_{i,j} + \mu_0 + \sum_{k=1}^{p} \mu_k \cdot q_k(x_i) = C_{i0} \\ \textstyle\sum_{i=1}^{n} w_i^{KED} \cdot q_k(x_i) = q_k(x_0) \\ \textstyle\sum_{i=1}^{n} w_i^{KED} = 1 \end{gathered} \tag{2.9}$$

Radar-reflectivity derived rainfall maps were the only covariate used in this study. Therefore, the solution of the KED system in matrix notation with radar as an extra covariate could be expressed as:

$$\begin{bmatrix} \mathbf{w}^{KED} \\ \mu_0 \\ \mu_1 \end{bmatrix} = \begin{bmatrix} \mathbf{C}_{ij} & \mathbf{1} & \mathbf{R}_i \\ \mathbf{1} & 0 & 0 \\ \mathbf{R}_i^T & 0 & 0 \end{bmatrix}^{-1} \cdot \begin{bmatrix} \mathbf{C}_{i0} \\ 1 \\ R(x_0) \end{bmatrix} \tag{2.10}$$

in which μ_0 and μ_1 are the Lagrange multipliers, $\mathbf{R}_i^T = \{R(x_1), R(x_2), \ldots, R(x_n)\}$ a row vector of the radar values associated to the observed locations, $\mathbf{C}_{ij}$ the covariance matrix between sampled points and $\mathbf{C}_{i0} \in \mathbb{R}^{nx1}$ the covariance vector between each of the measured points and the targeted location.

As discussed in the case of Ordinary Kriging, the objective estimation should be done at an area block support. Thus the equations were adapted in an analogous manner:

$$\begin{bmatrix} \mathbf{w}_B^{KED} \\ \mu_0 \\ \mu_1 \end{bmatrix} = \begin{bmatrix} \mathbf{C}_{ij} & \mathbf{1} & \mathbf{R}_i \\ \mathbf{1} & 0 & 0 \\ \mathbf{R}_i^T & 0 & 0 \end{bmatrix}^{-1} \cdot \begin{bmatrix} \tilde{\mathbf{C}}_{i0} \\ 1 \\ \overline{R}(x_0) \end{bmatrix} \tag{2.11}$$

the covariance matrix should represent the point-to-block covariance ($\tilde{C}_{iB}$). And the radar covariate was averaged within the block area:

$$\overline{R}(x_B) = \frac{1}{|B|} \sum_{j|j\in B} R(x_j) \tag{2.12}$$

this step facilitates the implementation of the kriging system, and it is consistent since the deterministic mean (in Equation 2.6) is calculated through a linear application. Thus, averaging the covariate beforehand is equivalent to performing predictions at each location and later spatially averaging them.

2.2.2.4.3. RAINFALL SPATIAL CORRELATION

The covariance used in Equation 2.5 and Equation 2.11 represents the spatial structure exhibited by the measured process. This was expressed in terms of a semivariogram structure, which in this case is related to the covariance as:

$$\gamma(x_i, x_j) = C_0 - C(x_i, x_j) \tag{2.13}$$

being $C_0 = var(r(x))$ (total variance). An empirical semivariogram was computed from the rain gauge dataset. This should ideally be computed independently at each time step (as temporal autocorrelation was neglected). However, in most cases, the urban scale rain gauge network does not present enough spatial density to compute a reliable semivariogram structure per time step. Therefore the rain field was assumed to be stationary of second order, isotropic and was computed as a time-lumped process, obtaining an averaged structure for each storm event in the form:

$$\hat{\gamma}(x_i, x_j) = \frac{1}{2m} \sum_{t=1}^{m} \frac{1}{N(h) \cdot s_t^2(R)} \sum_{i=1}^{N(h)} [R(u_i, t) - R(u_i + h, t)]^2 \tag{2.14}$$

for which $N(h)$ is the number of sampled pairs located at distance $h = |x_i - x_j|$ (Euclidean distance). The semivariogram was time-averaged over each selected independent storm process and normalised by the variance of the rainfall variable at each time step $s_t^2(R)$.

The empirical semivariogram is a cloud of discrete points relating space $lag(h)$ and semivariance. In order to generate a continuous spatial correlation map, a valid functional representation (experimental semivariogram) was fitted. An exponential structure was used in this case:

$$\gamma(h) = \begin{cases} 0 & \text{for} \quad h = 0 \\ c_0 + c(1 - e^{\frac{-3|h|}{\phi}}) & \text{for} \quad h > 0 \end{cases} \tag{2.15}$$

where c is the partial sill, c_0 is the nugget effect and ϕ the effective range or decorrelation distance (distance at which the semivariogram value reaches the 95% of the total sill). No relevant nugget was found in a preliminary study and therefore it was excluded from the semivariogram model (which was reported also by van de Beek et al. (2012)).

2.2.2.5. TEMPORAL SCALES IN THE DOMMEL SYSTEM

A characterisation of the system's reaction time was extracted from observed datasets at the rain gauge network, CSO structures, WWTP and river. Table 2.3 provides a summary of the characteristic temporal scales. Additionally, the accumulated discharged volume from all CSO discharges and the minimum DO level in the river linked to each storm process are shown. This provides an insight on the impact mechanism of each event and on the reaction patterns of the system.

Table 2.3: Characteristic time scales. Calculated as the delay time between the main rainfall peak and the system's response for: 1) the sum of combined sewer overflows (CSO), 2) the time to reach maximum flow capacity at the treatment works (WWTP) and 3) the time of stabilisation of the dissolved oxygen minimum river level (DO, measured ~ $17km$ downstream of the WWTP, M_0121)

Event Code	Storm time window	CSO	WWTP	DO	CSO Volume (m^3)	Min DO level ($gO2/m^3$)
1	18-08-2011 14:00 18-08-2011 21:00	2 h	5 h	28 h	$119.6 \cdot 10^3$	1.3
2	22-08-2011 21:00 23-08-2011 14:00	1 h	2 h	13 h	$573.1 \cdot 10^3$	0.7
3	26-08-2011 06:00 26-08-2011 14:00	-	4 h	23 h	$3.7 \cdot 10^3$	0.8
4	11-07-2012 10:00 11-07-2012 22:00	3 h	3 h	23 h	$157.4 \cdot 10^3$	2.8
5	28-07-2012 17:00 28-07-2012 20:00	1 h	1 h	25 h	$20.8 \cdot 10^3$	0.8
6	27-07-2013 05:00 28-07-2013 03:00	2 h	3 h	18 h	$103.3 \cdot 10^3$	0.6
7	11-08-2013 11:00 11-08-2013 23:00	-	3 h	14 h	$0.45 \cdot 10^3$	1.6

2.2.2.6. SPATIAL SCALES IN THE DOMMEL SYSTEM

Figure 1.4 shows the geographical location of the different municipalities connected to the Dommel catchment. The system is composed by 29 individual drainage areas, which are scattered in an area of roughly $25x30km$. The distance from the different urban drainage systems along with their drained area morphology plays a role on the sensitivity to rainfall spatial scales (e.g. a highly clustered system could be sufficiently represented with a few sampled locations, while a sparse one may require a more detailed spatial description). Figure 2.5 introduces the main spatial dimensions of the Dommel system. The lower triangular matrix displays the distance between centroids of the catchment areas, the area A_i in km^2 is shown in the diagonal, while the upper triangular matrix introduces a ranking metric of importance between area couples (normalised $\sqrt{A_i} \cdot \sqrt{A_j}$). This ranking metric shows the clustering effect exerted by the biggest municipal area (the city of Eindhoven).

A global characteristic spatial scale was computed as a weighted average of the centroid-to-centroid distances plus the average area length scale, expressed as:

$$\hat{L}_c = \frac{\sum_{i>j}[(d_{ij} + \sqrt{A_i} + \sqrt{A_j}) \cdot \sqrt{A_i} \cdot \sqrt{A_j}]}{\sum_{i>j}[\sqrt{A_i} \cdot \sqrt{A_j}]} \approx 15km \tag{2.16}$$

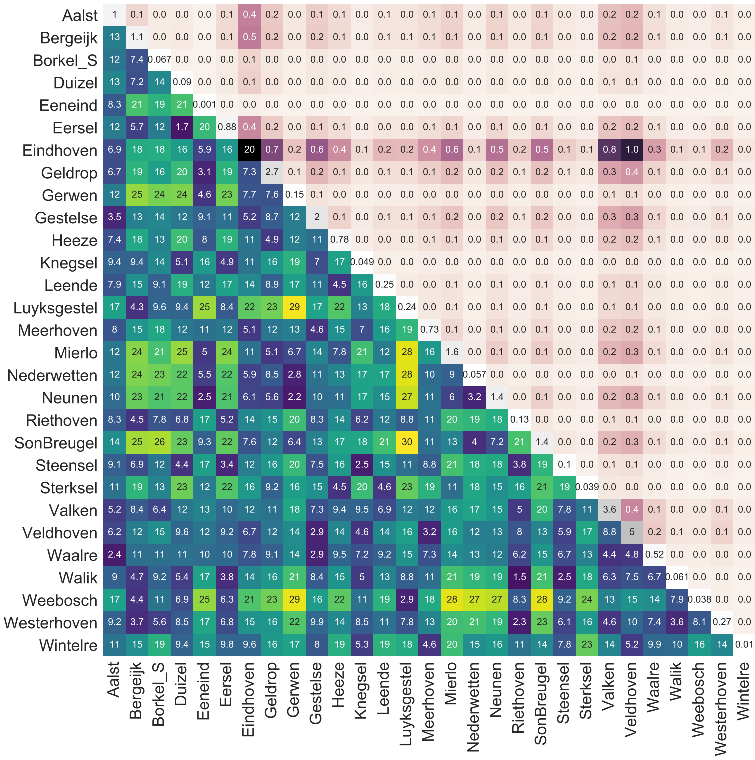


Figure 2.5: Length scales for each municipal urban drainage system. Distance between area centroids (km) is shown in the lower triangular. The diagonal displays the catchment's size (km^2). And the upper triangular shows a ranking metric of the relative importance between areas.

being A_i the area of the ith catchment, d_{ij} the distance between the centroids i and j. This metric aims at accounting for the distances between connected areas giving an increased importance to larger area couples, thus providing with an estimated catchment characteristic scale, which can be used to diagnose the system's behaviour against rainfall processes with different correlation scales.

2.2.3. RESULTS

2.2.3.1. URBAN DRAINAGE DYNAMICS

The effect of rainfall input source (space-time variability) in the urban drainage dynamics is illustrated by the comparison of the four most relevant internal variables: 1) Maximum estimated rainfall intensity in the catchment, 2) accumulated rainfall depth, 3) peak discharge at the CSO and 4) accumulated discharged volume. The estimated rainfall intensity and accumulated depth describes differences in the rainfall input at each catchment. Each of the selected inputs were propagated through the urban drainage submodel, and subsequently rendered the pollutant load in the river system. The urban drainage response is characterised by the CSO dynamics, described by the discharge peak and the accumulated discharged volume. Figure 2.6 shows the effect of rainfall input on the urban drainage system of Valkenswaard (Figure 1.4). This municipal system is an especially illustrative example since it is located approximately 12 km from the rain gauge KNMI_370 (source for the BK1 product) and represents a relevant contribution to the total discharged volume (second largest combined sewer connected area). This example allows the observation of the joint effect of time and space variability.

The selected time aggregation influences the estimated maximum rainfall intensity, which however does not propagate to the estimation of total rainfall volume and to the CSO simulation. A minor effect of time-aggregation at the interpolated rainfall products (BKall and UBK) can be observed. However, this difference is relatively small compared with the exhibited by the selected spatial data sources. The use of a single rain gauge input, BK1 creates an appreciable deviation in the computation of CSO dynamics in most of the rainfall processes studied when compared with monitoring data. Differences between the other three products (BKall, UBK and ARadar) are less apparent.

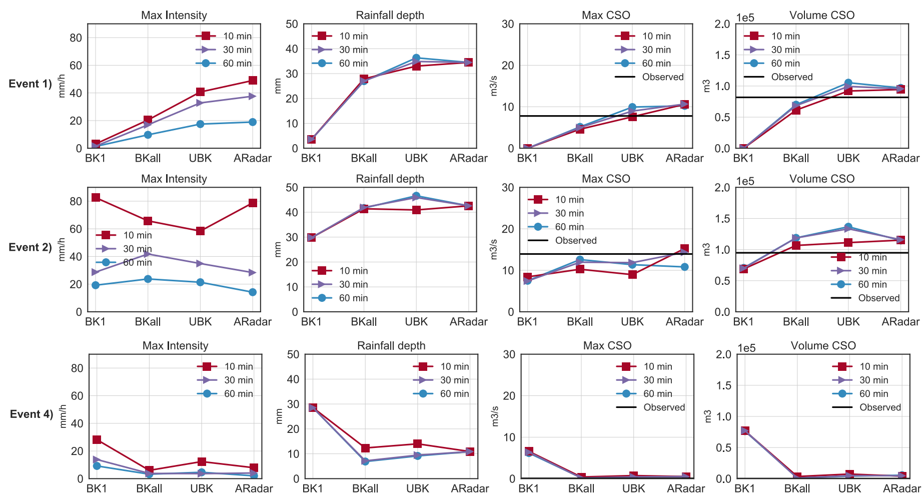


Figure 2.6: Rainfall input and urban drainage response for the municipality of Valkenswaard. Each graph depicts the effect of the four rainfall products rendered at different time accumulation steps. From left to right; maximum rainfall intensity, accumulated rainfall depth, maximum CSO flow and accumulated CSO volume.

The time-space rainfall input variations were decoupled to display the effect on all urban drainage catchments connected to the water system. Figure 2.7 shows the effect of variations of the spatial product (at fixed 60 minutes time aggregation) for each catchment (represented by their connected areas in abscises). The rainfall intensity and depth provided by the product BK1 produces a distortion of the rainfall map, which can be appreciated in events 1, 2 and 5 (Figure 2.7). Although in some cases like in Storm 6 or Storm 7, BK1 can be representative of the mean of the process, it generally induces a deviation into the rainfall input description with respect to the rest of the products generated. This error is not systematic and it is linked to the homogeneity of the rainfall process (if the central rain gauge is representative for the overall spatial domain or not). In consequence, those differences in the rainfall quantification due to spatial data variation affects the CSO simulated dynamics. Event 1 and 4 are a good example of this phenomena, for which BK1 underestimates first and overestimates later the total rain volume being consequently transferred to the accumulated CSO volume.

Figure 2.8 shows the effect of the selection of time-aggregation level at each urban drainage system (using the ARadar spatial data source as fixed input). Here it can be seen how the selection of a certain time aggregation (10-60 minutes) impacts on the estimation of the maximum rainfall intensity, however those differences do not propagate to the computation of accumulated rainfall volume and in consequence to the simulation of CSO discharges.

2.2.3.2. RAINFALL SPATIAL VARIABILITY

The spatial characteristics of the observed rainfall processes under study were estimated by the use of a time-averaged empirical semivariogram (as described in the methods section). Table 2.4 shows the fitted range and sill parameters for each storm process depending on the time-aggregation level. This describes the correlation structure present in the rainfall process. A large range indicates a spatially homogeneous rainfall process. The increase in accumulation time of the rainfall input produced a more spatially correlated map (seen by the increase on range). This was to be expected since aggregating on time acts as smoothing of the process. Additionally, Figure 2.9 shows the accumulated rainfall depth map at 24 hours from the start of each storm. This represents the storm main path, describing the total rainfall volume spatial distribution.

Table 2.4: Fitted parameters sill and range (ϕ, in km) for an averaged exponential semivariogram model.

	Event 1		**Event 2**		**Event 3**		**Event 4**		**Event 5**		**Event 6**		**Event 7**	
T	Sill	ϕ	Sill	ϕ	Sill	ϕ	Sill	ϕ	Sill	ϕ	Sill	ϕ	Sill	ϕ
10’	0.87	17.6	0.69	15.8	0.93	18.8	0.87	13.6	0.4	5.3	0.31	26.7	0.64	12.7
30’	0.86	24.7	0.88	24.5	0.87	25.2	0.87	25	0.83	23.4	0.9	26.4	0.91	26.4
60’	1.13	43.6	1.16	45.5	1.09	40.8	1.15	45.3	1.14	44.4	1.08	40.1	1.07	39.4

2.2.3.3. RAINFALL INPUT EFFECT ON DISSOLVED OXYGEN DYNAMICS

In order to assess the effect of rainfall input in dissolved oxygen simulations, modelled time-series were compared against existing monitoring data. Table 2.5 presents the residuals between the minimum oxygen level observed and modelled at all storm-input combinations. Minimum oxygen simulated level is highly relevant in practice since it captures the magnitude of impact at the receiving water body and is directly related to the pollution load in the system (which is driven by the rainfall process). Table 2.6 shows the root-mean-squared-error (rmse) between simulated-observed time series from the beginning of the storm event until 5 days later when the DO concentration is recovered. The performance indicators for dissolved oxygen show that the effect of time-accumulation of the rainfall input is negligible.

Figure 2.10 presents a graphical comparison between dissolved oxygen (DO) river observations and modelled series at each spatial data source selection (fixed 60 min-

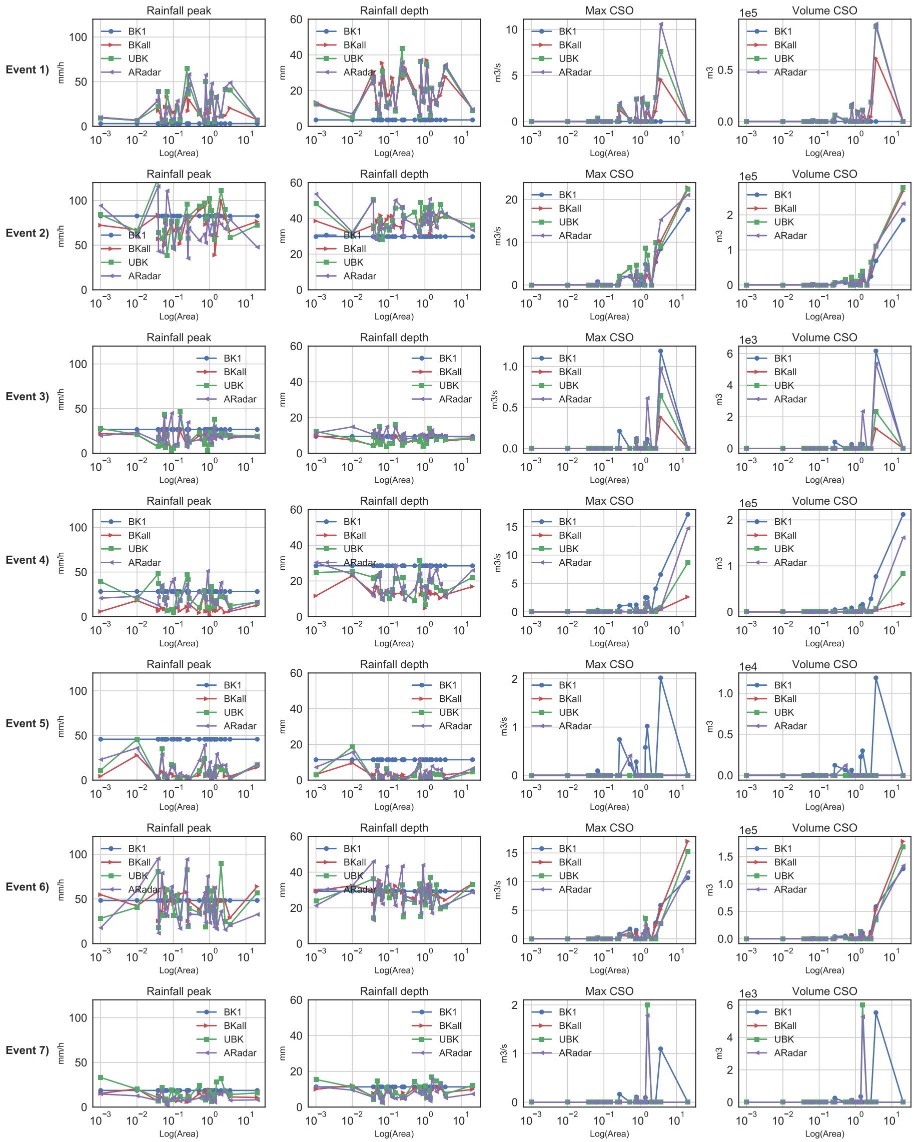


Figure 2.7: By-catchment response to rainfall spatial information for the 10 minutes temporal-accumulated product (catchments are represented in the x axis by their connected area in km^2).

utes time accumulation). The graph considers the three time-period simulations as described at Table 2.1.

This shows that the effect of rainfall selection had a varying impact depending on the storm process. This is linked to the mechanism of DO depletion of each event and on the characteristics of the rainfall process. For instance, storms from period 1 and 2 exhibit differences due to the spatial data selection in both the DO depletion depth and recovery duration. However, the events occurring in period 3 show low sensitivity to the

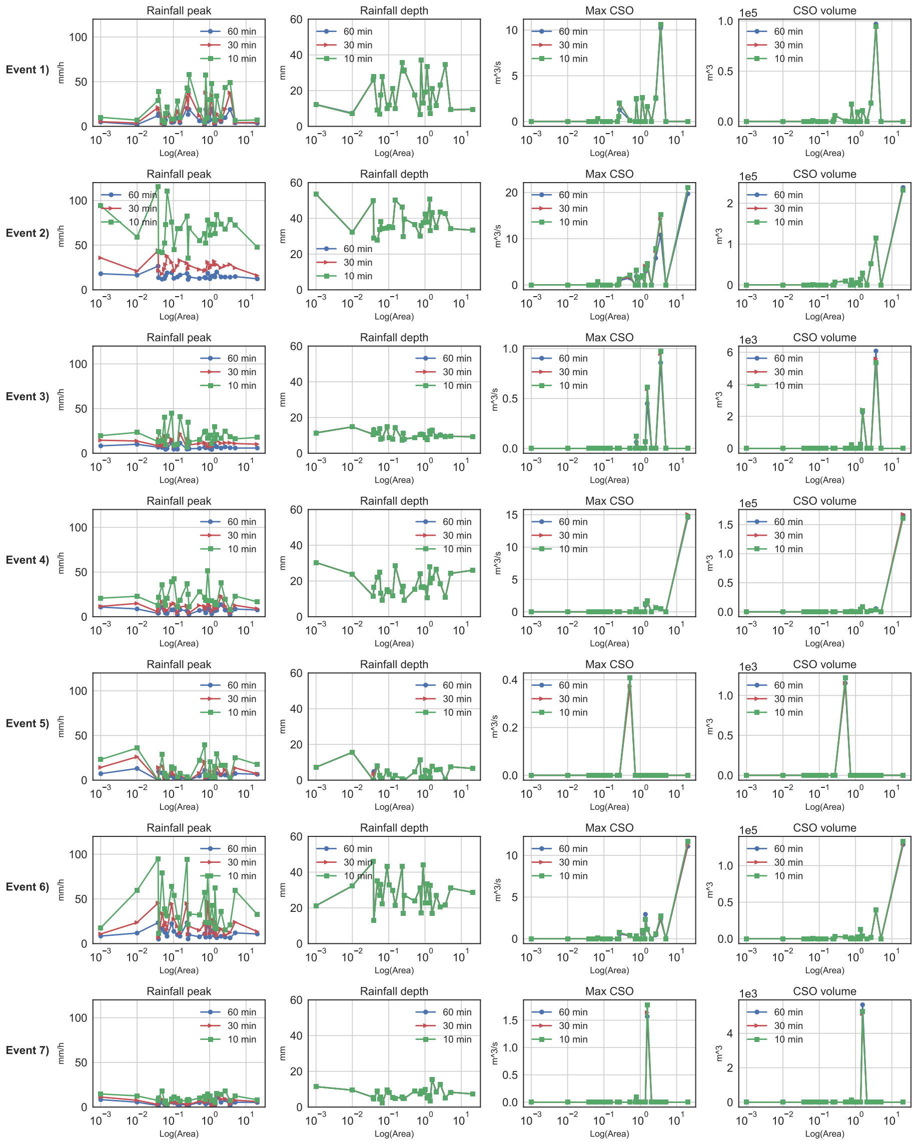


Figure 2.8: By-catchment response to rainfall temporal accumulation for the ARadar spatial product (catchments are represented in the x axis by their connected area in km^2).

selected rainfall input. In the case of event 7 this can be explained by the fact that CSO volumes were low (see Table 2.3), thus the DO depletion was likely not due to CSOs but rather due to the discharge from the WWTP (which filters the effect of time-space) or by unaccounted processes not included in the model structure. Event 6, on the other hand, presented a rainfall event for which the rain gauge KNMI_370 was a good representation of the main contributing urban areas.

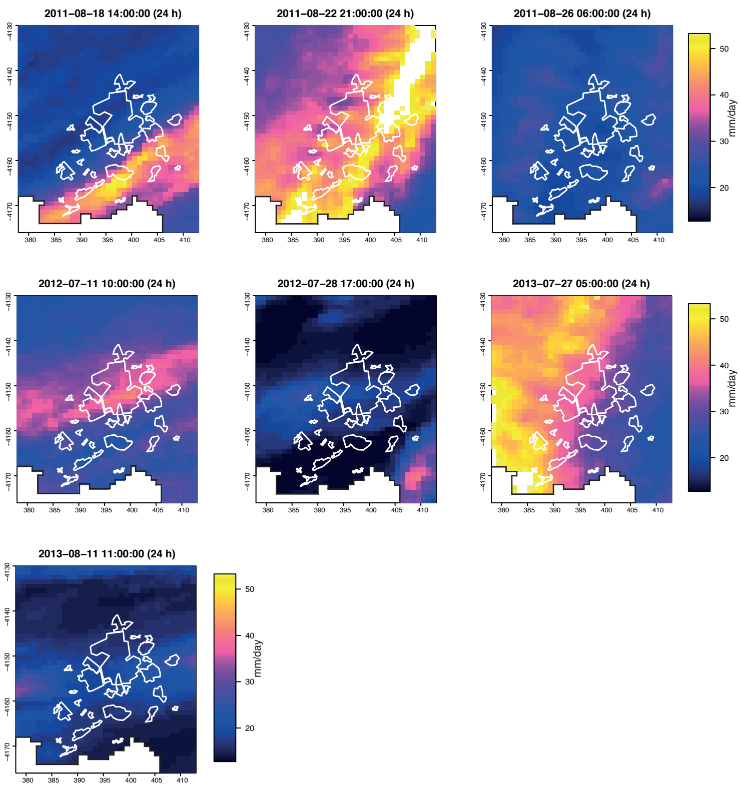


Figure 2.9: Spatial distribution of 24h-accumulated rainfall (over the urban areas of the Dommel system) covering the seven storm events (extracted from the radar product).

Table 2.5: Difference between observed and modelled minimum dissolved oxygen concentration ($gO2/m^3$).

Rainfall input	**Event 1**	**Event 2**	**Event 3**	**Event 4**	**Event 5**	**Event 6**	**Event 7**
BK1_10’	-3.9	-1	-1.4	1.2	-2.15	-0.49	-1.1
BK1_30’	-3.9	-1	-1.4	1.2	-2.15	-0.49	-1.2
BK1_60’	-3.9	-1	-1.4	1.2	-2.15	-0.49	-1
BKall_10’	-1.25	-0.76	-1.3	-0.05	-1.8	-0.38	-1.15
Bkall_30’	-1.25	-0.67	-1.21	-0.2	-1.5	-0.38	-1.1
Bkall_60’	-1.25	-0.58	-1.1	-0.2	-1.35	-0.38	-1.05
UBK_10’	-1.3	-0.49	-1.09	0.55	-1.74	-0.4	-1.2
UBK_30’	-1.28	-0.52	-1.09	0.65	-1.5	-0.4	-1.25
UBK_60’	-1.28	-0.51	-1.09	0.65	-1.4	-0.4	-1.2
ARadar_10’	-1.26	-0.71	-1.2	1.05	-0.9	-0.37	-1.4
ARadar_30’	-1.26	-0.72	-1.2	1.06	-0.9	-0.37	-1.4
ARadar_60’	-1.26	-0.72	-1.2	1.06	-0.9	-0.37	-1.3

2.2.4. DISCUSSION

Figure 2.6 and Figure 2.7 show the deviation generated by the homogenous rainfall source (BK1) for the rainfall estimation (depth and intensity), which was consequently trans-

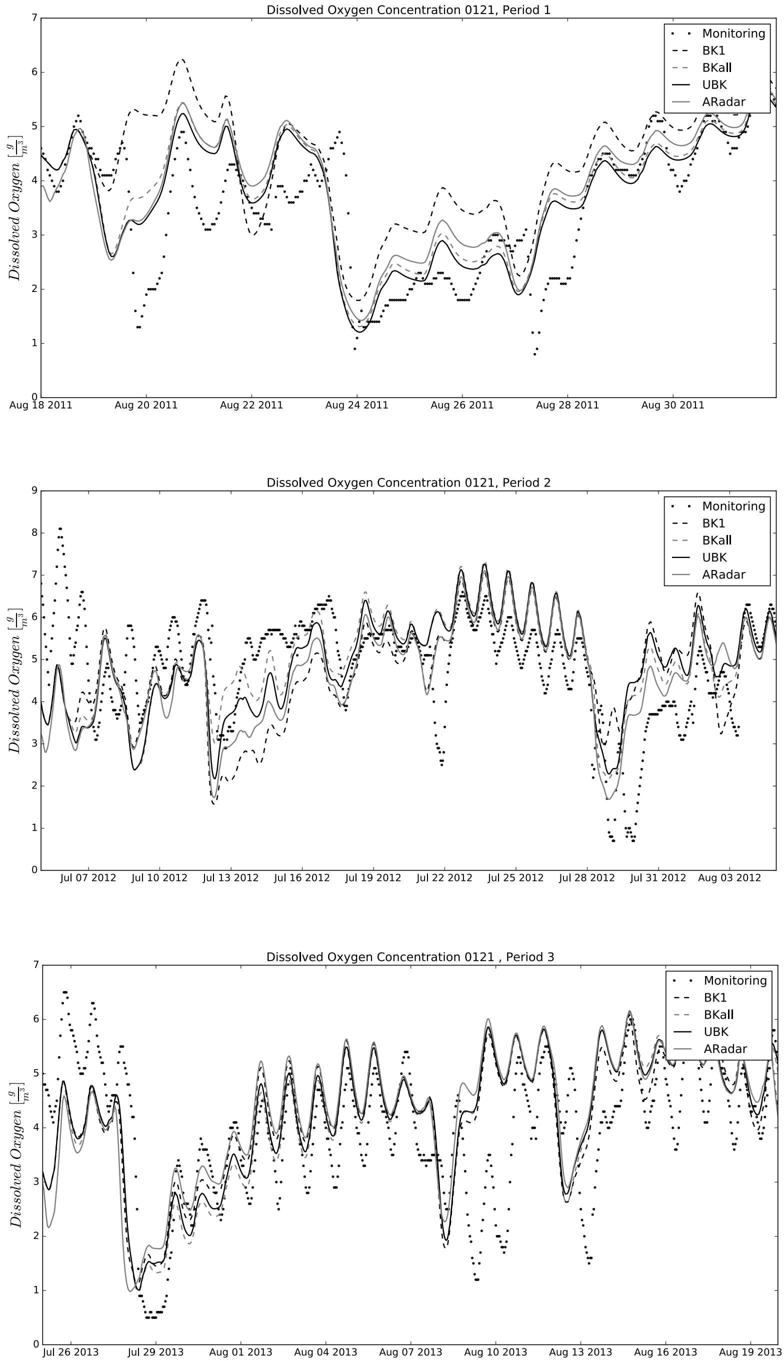


Figure 2.10: Graphical comparison of dissolved oxygen dynamics measured vs modelled by the different spatial products (60 min time-accumulated).

Table 2.6: RMSE between observed and modelled dissolved oxygen concentration (5 days).

Rainfall input	Event 1	Event 2	Event 3	Event 4	Event 5	Event 6	Event 7
BK1_10'	1.58	1.27	1.2	1.9	1.85	0.98	1.15
BK1_30'	1.58	1.27	1.2	1.92	1.87	0.99	1.15
BK1_60'	1.57	1.26	1.18	1.94	1.86	1	1.14
BKall_10'	1.16	0.99	0.54	0.88	1.76	0.96	1.2
Bkall_30'	1.12	0.95	0.44	0.87	1.58	0.98	1.18
Bkall_60'	1.12	0.91	0.31	0.89	1.5	0.98	1.17
UBK_10'	0.99	0.88	0.2	1.1	1.61	0.96	1.07
UBK_30'	1	0.87	0.15	1.15	1.81	0.95	1.1
UBK_60'	0.99	0.88	0.17	1.14	1.69	0.95	1.14
ARadar_10'	1.07	0.98	0.64	1.47	1.28	1.26	1.16
ARadar_30'	1.07	0.97	0.6	1.51	1.28	1.25	1.16
ARadar_60'	1.1	0.97	0.57	1.51	1.28	1.26	1.16

ferred to the CSO discharge patterns. There is a clear difference between predictions rendered by a homogenous rainfall data source (BK1) and the results stemming from the use of extra spatial information either the extended rain gauge network (BKall), radar spatial predictions (ARadar) or a merged product rain gauge-radar (UBK).

Varying the time aggregation of rainfall products (at 10, 30 or 60 minutes) influenced the maximum estimated rainfall peak. However, this difference was not transferred to the estimated total rainfall volume (depth) and to the CSO-DO dynamics. Figure 2.8 shows that the influence of time aggregation on CSO patterns is almost negligible when modelling discharged volumes and that this is generalizable to all catchments under study. This is explained by the nature of the urban drainage system, which as described, is characterised by low slopes and large in-sewer storage. Thus, the behaviour of CSO spills is dominated by the rainfall volume and less by its dynamic component.

A direct link between the rainfall characteristics (maximum intensity, total volume, average correlation range) and the sensitivity of dissolved oxygen to rainfall events could not be established. Figure 2.9 shows the accumulated rainfall volume spatial distribution of the rainfall events. It can be appreciated how the storm cell evolution can play an important role in the process since its variability occurs within the spatial scale of the catchment system. Thus, the spatial distribution of storm processes should be included in synthetic rain generators used for design and test purposes (as described in Willems (2001)). A slightly higher influence of the time-step is observed at both kriging inputs BKall and UBK (for modelling CSO discharge). This was due to the nature of the interpolation process which performance is known to decrease at short time scales (lower than 30', e.g. Nanding et al. (2015)).

The effect of time aggregation of the rainfall field in DO dynamics is negligible. This is shown at Table 2.5 in which the variation of the modelled-observed minimum oxygen level residual is practically insensitive to the variation in time. This effect is explained by the fact that DO dynamics in the system are mainly affected by the WWTP effluent and by the total CSO discharged volume. Both of them seem to be relatively insensitive to rainfall input time accumulation (between 10-60 minutes). Additionally, DO processes present a significant inertia with respect to the rainfall process (15-28 h, as shown in Table 2.3), acting like a high-frequency filter of the CSO dynamics.

This study intends to describe the interaction of a large-scale integrated urban catchment model and the rainfall input characteristics. It is relevant to mention that such integrated modelling studies are subjected to several sources of uncertainties (which are not independent). The water quality model applied has not been calibrated for the studied time series, due to the fact that the main goal of this study is the observation of changes due to input data source. Therefore several errors can be observed in the simulated time series. First, there is a systematic underestimation of DO depletion processes of 1-0.5 mg/l (which could be indicative of a wrong description of the urban-WWTP pollutant concentration load). Additionally the DO simulations reacted slightly faster (1-8 h) than observed in the measured data, which can be indicative of a wrong timing effect on CSO-WWTP effluent simulation (since hydraulics in the system were calibrated). Those errors however are not expected to have an influence on the conclusions of the study.

Partial studies on the effect of input errors in similar systems can be found in (Radwan et al., 2004a), which described the sensitivity of a river model to several sources of uncertainties, indicating that 60% of DO simulation variability is explained by the input data. However this considered variations in pollution load inputs from urban systems and not directly on rainfall. A full uncertainty quantification scheme was proposed for a small integrated system in Willems (2008), but only accounting for urban drainage and WWTP outputs (not receiving water), and reported that rainfall uncertainties had a contribution of 15-20% to uncertainties in BOD loads from a sewer and WWTP effluents. These studies report a relative sensitivity of DO or related variables to the input description. Thus reducing the errors in the rainfall input data set could influence the ability to simulate DO patterns. The findings of this study aim to describe the relevant characteristics that the input data source should have when modelling DO in a large urban integrated water system. Which can be used to direct further monitoring and modelling efforts in this system or in similar ones.

2.3. Accounting for correlation in the stochastic simulation of CSO pollutant loads *

2.3.1. Introduction

Pollutant concentrations at combined sewer overflows (CSO) exhibit dynamics that are challenging to represent. Current knowledge about the physical processes involved is still limited and sufficient data to calibrate models is often not available. It is acknowledged that deterministic results of sewer water quality models are generally poor (Dotto et al., 2014; Sandoval et al., 2018). For this reason, mean concentration pollutant vectors are often used as input when modelling sewer overflow impacts in scarcely monitored environments (Langeveld et al., 2013b; Llopart-Mascaró et al., 2015). This simplification can render uncertain model outputs. Nevertheless, these uncertainties can still be of the same order of magnitude when using uncalibrated sewer water quality models (Willems, 2006). Thus, assessing the propagation of uncertainties product of these assumptions is necessary when checking the reliability of the simulation platform.

Forward uncertainty propagation schemes rely on sampling parameters from probability ranges in order to evaluate its effect on the model output variables. Parametric or input probability distributions are often inferred, elicited from expert beliefs (O'Hagan, 1998) or when possible, derived from measured data. In this example, a water quality model utilises pollutant mean concentration values to transform flow simulations from urban drainage structures to receiving water pollutant loads (Langeveld et al., 2013b). The full-integrated model simulates water quality dynamics (i.e. dissolved oxygen impacts in the receiving water body) and integrates the simulation of rural-urban pollution sources, wastewater treatment processes and river biochemical and physical dynamics at the river Dommel.

Uncertainty ranges for water quality pollutants are derived from measured series, and propagated through the model platform, thus estimating the magnitude of their effect on dissolved oxygen concentrations in a downstream section of the receiving water body. Literature reports for CSO pollutant concentration across systems show that storm pollution loads present highly skewed distributions and strong pollutant correlations (Aarts et al., 2013).

A stochastic model is created to represent the CSO pollutant variability, from which samples are drawn in a Monte-Carlo forward uncertainty propagation. Nevertheless, parametrically representing a correlated non-Gaussian multivariate joint probability distribution is not straightforward, thus the correlation of uncertainty sources is often neglected. In this example, two stochastic models are compared, one that draws independent samples from the marginal distribution of pollutant loads (uncorrelated marginal) and one, which respects the correlation structure of observed pollution samples (copula sampler). The individual effect of the variability of CSO pollutant loads and the effect of neglecting pollutant correlation are shown.

*This section is based in the conference paper: Moreno-Rodenas, A. M., Langeveld, J. and Clemens, F. H. L. R. (2017). Accounting for correlation in uncertainty propagation. A copula approach for water quality modelling. Presented at the International conference in Urban drainage (ICUD 2017), Prage. Czech Republic

2.3.2. MATERIALS AND METHODS

The ICM study for the Dommel system (described in Section 1.7) was used to simulate water quality dissolved oxygen impacts in the receiving water body. To further minimise the effect of other uncertainty sources, CSO flow discharges were represented using estimated data-series from idealised weir discharge models and depth monitoring data (although these can still contain significant sources of uncertainty, see van Daal-Rombouts (2017)).

Distributions for each pollutant concentration were obtained from a monitoring campaign at several CSO structures (Moens et al., 2009). Figure 2.11 shows the marginal distribution from each pollutant variable extracted from the measurement data. A parametric probability distribution was fitted to each marginal distribution by selecting the minimum sum of square error fit. Additionally, the spearman correlation matrix was used to represent the structural correlation between pollutant variables. Dispersion of mean concentration values shown in the dataset were consistent with previously reported values in literature (e.g. see Aarts et al. (2013) for a compilation of CSO pollutant loads across European sewer systems).

A comprehensive review on generating correlated non-Gaussian multivariate samples can be found at Schoelzel and Friederichs (2008). In this study an elliptical Gaussian copula (GC) was applied as a way to generate a joint distribution, which reproduces the pollutant structural correlation and each marginal probability density.

In order to construct the GC, let $\mathbf{X}$ be a random vector, with components $[X_1,\dots,X_N]$ representing the set of pollutant concentrations with known marginal cumulative density functions (CDF) $[F_{X_1},\dots,F_{X_N}]$. By applying the CDF to the samples of each variable, $F_{X_i}(X_i) = U_i$ a uniformly distributed set can be obtained, $U_i \sim U(0,1)$, which can be transformed by $F^{-1}_{N(0,1)}(U_i) = Z_i$ (using the inverse CDF of a standard normal distribution). Thus the vector $\mathbf{Z}$ forms a Gaussian multivariate distribution $\sim N_k(0,\Sigma_{GC})$. This Gaussian multivariate density describes implicitly a Gaussian copula with a correlation structure given by the covariance matrix Σ_{GC}.

Samples from the joint probability density function defined by the copula are drawn following the inverse process, starting from a random realization of the copula density (drawing samples from $\mathbf{Z} \sim N_k(0,\Sigma_{GC})$, for which Σ_{GC} represents the observed covariance function) and back-transforming to the desired marginals by using $[F^{-1}_{X_1}(F_{N(0,1)}(Z_1)),$ $\dots, F^{-1}_{X_N}(F_{N(0,1)}(Z_N))]$, thus obtaining a vector of correlated samples which follow the individual marginal distribution and the correlation structure.

The model was evaluated using 500 samples from the GC distribution and from the uncorrelated marginals. The number of samples was selected to ensure a stable output variance. The variables used were concentrations of chemical oxygen demand (COD), biological oxygen demand (BOD) and ammonium concentrations (NH4). Predicted dissolved oxygen concentrations in a river section (M0121, Figure 1.4) were compared with monitoring data for a period of 6 months (01-04-2012 / 31-09-2012).

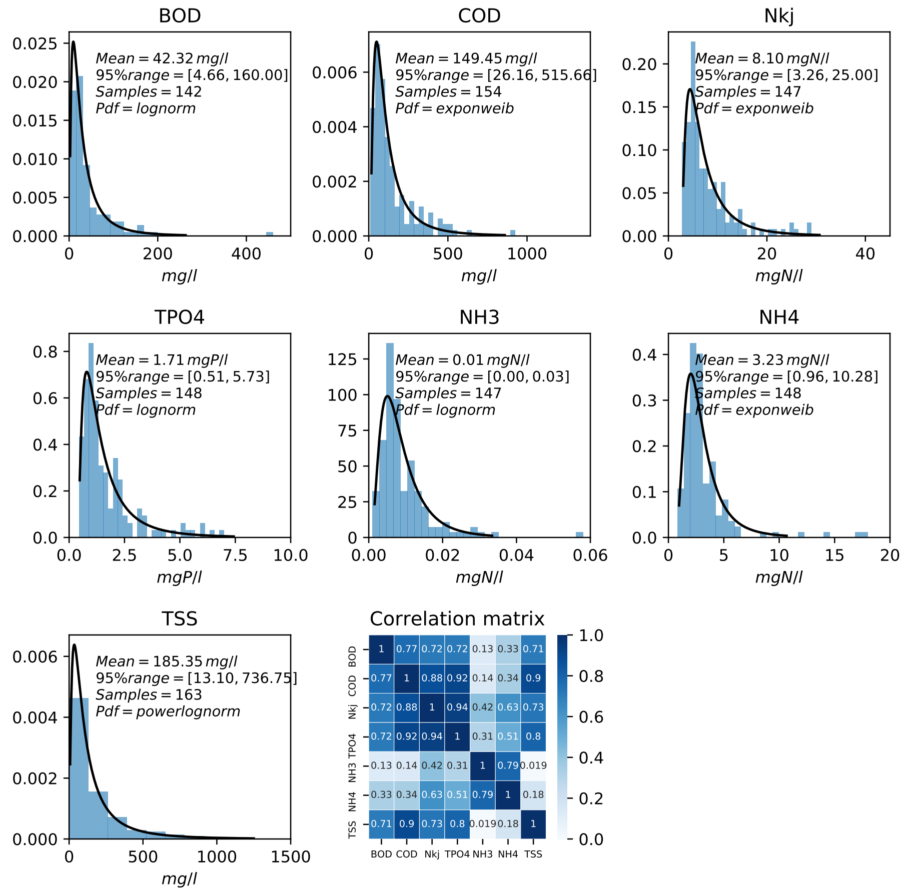


Figure 2.11: Histograms and correlation matrix (spearman rank correlation) of measured data, mean and 95% confidence intervals. Blue solid line; fitted probability density function

2.3.3. RESULTS AND DISCUSSION

A comparison of the rank (spearman's) correlation matrix for the measured data, along with 300 samples from the proposed Gaussian copula and random sampling from the marginal pdf's are shown at Figure 2.12. Also, a graphical scatter plot for the variables BOD and COD are provided in the three scenarios. The Gaussian copula formulation successfully captures the structural correlation of the dataset at the same time that realisations respect the marginal probability density functions for each pollutant variable.

The model was evaluated at 500 pseudo-random samples from the Gaussian copula and the uncorrelated marginal stochastic representations. Figure 2.13 depicts the comparison between the observed and simulated dissolved oxygen when using the Copula distribution for roughly 1.5 months occurring in summer of 2012. Storm events exhibit a differential sensitivity to CSO pollutant variability. This variable sensitivity is a good example of the complexity of the underlying process. Dissolved oxygen is depleted fol-

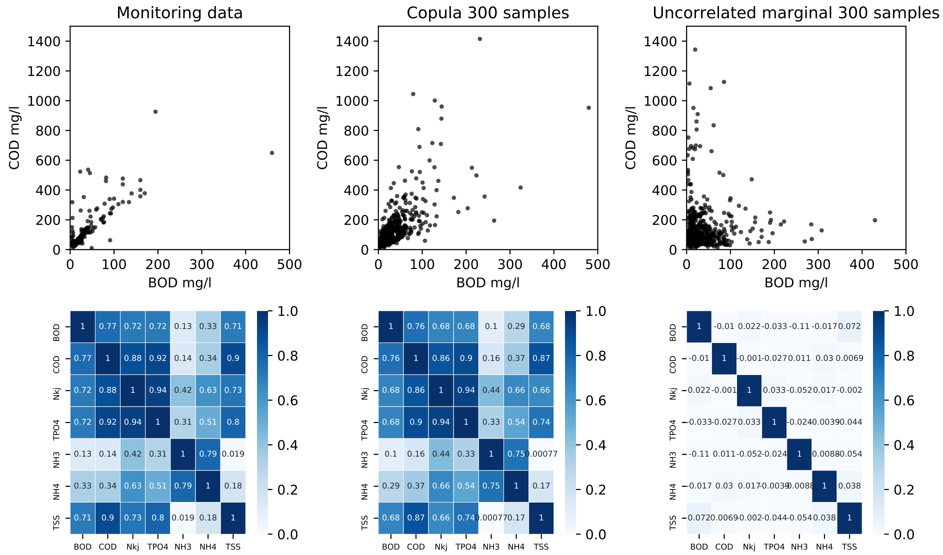


Figure 2.12: Comparison of measured data, Gaussian copula and random sampling from marginals, scatter plot of BOD-COD (above) and correlation matrix of measured and simulated concentrations (below).

lowing different mechanistic relationships depending on the dynamic state of the system. Some events as 14-Jul-2012 or 28-Aug-2012 show a high sensitivity to CSO pollutant concentrations, suggesting that those events were predominantly caused by urban discharges, whereas in the 18-Aug-2012, a high impact event occurs which shows a low sensitivity to CSO discharges, and hence likely caused by a WWTP overload.

Figure 2.14 presents the comparison of the river DO probability density when drawing samples from the CSO water quality copula representation or from the uncorrelated marginal distribution of pollutants at 8 characteristic events. The two-sample Kolmogorov-Smirnov and the k-sample Anderson-Darling tests were used to assess whether both samples (copula vs. random) produce an equivalent DO estimation. Comparison for all selected events in Figure 2.14 resulted in p-values < 1%, thus the null hypothesis (both samples originate from the same underlying probability distribution) could be rejected. This result suggest that neglecting the correlation structure shown by measured CSO water quality variables results in a significant underestimation of the effect in the receiving dissolved oxygen depletion simulation.

It should be noted that Gaussian copulas have zero tail dependence, meaning that extreme cases in the tails of the copula are asymptotically independent. This makes them relatively unsuitable for applications focusing in concomitant extreme value processes, where other copula formulations (e.g. Gumbel or Clayton) should be preferred (Malevergne and Sornette, 2003). In this case, the lack of a large measured CSO pollutant dataset makes the description of high-magnitude correlations unreliable, thus for simplicity a symmetric Gaussian copula was applied. Nevertheless, the selection of copula family should be considered based on the characteristics of the process to represent.

Additionally, the presented stochastic representation of pollutant concentrations is

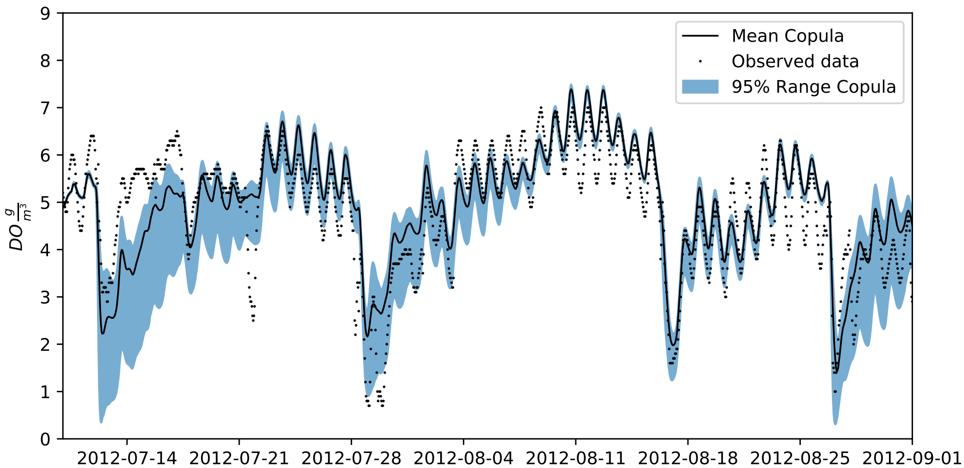


Figure 2.13: Dissolved oxygen concentration measured (black dots), simulated mean (black line) and 95% interval (blue range) from the propagation of the Gaussian Copula distribution at the river ($\sim 17km$ downstream of the WWTP, M0121).

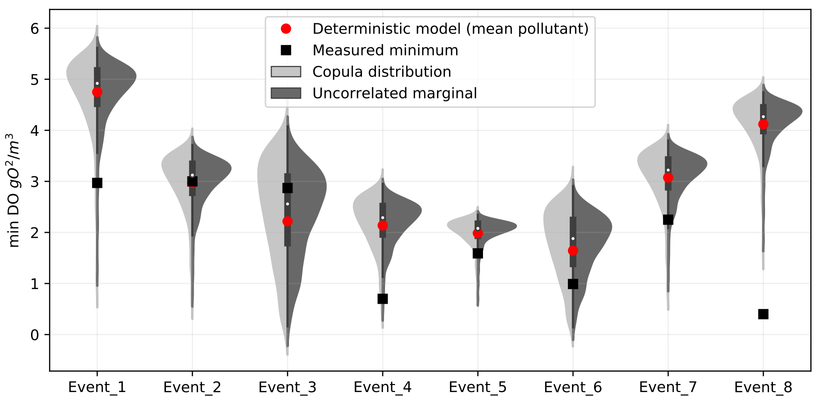


Figure 2.14: Minimum dissolved level for events in which river oxygen measured concentration fall below 3 mgO2/l (M0121 station). The graph shows the density of the predicted series from the forward uncertainty propagation (500 samples) using the Gaussian copula distribution and uncorrelated samples from the marginals. Additionally the observed minimum (black square) and the deterministic realization using the mean of the pollution concentrations (red dot) are shown.

highly simplified, since it does not account for the time-dependency of pollutant concentration, it assumes spatial independency across CSO systems and it does not consider the link between nearby events (e.g. availability of in-sewer sediment stocks). This simplified representation is due to the lack of long time-series for water quality in-sewer variables. Further knowledge in the process should be acquired to reduce epistemic uncertainties in the description of CSO pollutant loads.

2.4. SUMMARY AND CONCLUSIONS

THE identification and characterisation of uncertainty sources should be considered an elemental step of the integrated urban catchment modelling process (Tscheikner-Gratl et al., 2019). Existing uncertainty analysis frameworks (see Deletic et al. (2012), Refsgaard et al. (2007) or Tscheikner-Gratl et al. (2017)) provide a valuable guideline to report uncertainties to end-users in a structured manner. The a priori selection of the main sources of uncertainty sources depicted in this section (see Figure 2.2) should be complemented with a quantitative formal analysis scheme. An example of such analysis is further described in Chapter 4 of this thesis. Nevertheless, the exercise of eliciting the expected contribution of uncertainty sources a priori, allows discussing and communicating the effect of non-quantifiable sources of uncertainty (e.g. deep uncertainties). Such prior analysis should effectively communicate uncertainty sources so that the end-users are fully aware of the expected representativity and limitations of the modelling endeavour.

2.4.1. ON THE IMPACT OF RAINFALL SPATIO-TEMPORAL DESCRIPTION IN ESTIMATED DO DYNAMICS

The section 2.2 of this chapter addresses the influence of rainfall input spatial and temporal characteristics in the modelling of dissolved oxygen patterns in a large lowland integrated urban water system. The effect of varying the rainfall process description by using different accumulation time scales and different rainfall measurement sources was investigated. A total of 12 rainfall input products were tested for 7 storm events, which generated significant oxygen depletion events in the receiving water body. Rainfall inputs were generated at three time-accumulation levels (10, 30 and 60 minutes) and using four different rainfall sources: 1) a single rain gauge (BK1), 2) a block kriging interpolation from 13 rain gauges (BKall), 3) radar-derived rainfall composites from the KNMI (ARadar) and 4) the results of Kriging with External Drift merging of rain gauges and radar (UBK). All estimations were generated at the spatial support of the individual connected urban drainage networks.

The tested rainfall time-accumulation level did not influence the modelled dissolved oxygen patterns in the river. Although it influenced the maximum estimated rainfall intensity, the effect on rainfall-accumulated volumes was mostly limited. This effect was not propagated to simulated CSO dynamics at the urban system in most cases and therefore the effect on DO dynamics was negligible. It was found that the use of a single rainfall point measurement generates a non-systematic deviation in the simulated CSO series. This miss representation has the potential to affect DO dynamics.

The results obtained in this example can only be generalised to similar systems (i.e. low-land and volume driven) with an equivalent mechanistic relationship urban drainage-WWTP-river. Nevertheless, a detailed description is provided for: 1) Characterising the rainfall process and the system's response patterns, 2) generate rainfall inputs at the spatial support of the target catchment and 3) assess the sensitivity of the model to the characteristics of rainfall input. This process can be followed in other integrated catchment studies in order to identify the desired characteristics of rainfall input data sets for each selected system-variable.

2

2.4.2. THE EFFECTS OF NEGLECTING THE CORRELATION PATTERNS IN CSO POLLUTANT STOCHASTIC GENERATORS

Uncertainties in CSO pollution loads can impact significantly DO depletion patterns. The pollution load affects especially the intensity and the recovery of dissolved oxygen concentrations in the river Dommel, which are highly relevant parameters when characterizing river impacts through concentration-duration-frequency environmental impact tables (Appendix A).

This example presents a simple method to account for correlation in forward uncertainty propagation schemes for water quality applications. Not accounting for structural pollutant correlation in the uncertainty propagation of CSO uncertainties produced a certain underestimation of DO uncertainty bands. However, it has to be noted that the use of correlation did not lead to a systematic improvement of observed values coverage. Nevertheless, when knowledge about the correlation structure is available it is recommended to include it in the uncertainty propagation process.

3

ACCELERATING UNCERTAINTY QUANTIFICATION

3.1. EMULATION FOR WATER QUALITY AND QUANTITY SIMULATORS

ONE of the current challenges in the implementation of uncertainty quantification schemes in environmental modelling is the impracticability of sampling slow models (Reichert, 2012; Tscheikner-Gratl et al., 2019). This is specially applicable to integrated catchment modelling studies (Sriwastava and Moreno-Rodenas, 2017), where the computational burden increases rapidly when performing full-scale catchment simulations. A common approach to accelerate the simulation process is to simplify submodel representations by means of conceptual or parsimonious model structures, which is a common practice in urban drainage modelling (Solvi, 2006; van Daal-Rombouts, 2017) or flood modelling (Bermúdez et al., 2018). However, even conceptual-lumped approaches lead to slow simulators when scaled (e.g. real-scale integrated urban water systems), and this strategy does not respect the structure of physically based simulations, thus adding extra (structural) sources of uncertainty in the model operation process.

Data-driven model emulation can be used to accelerate the sampling of computationally expensive models. This aims to generate an interpolation map between a selected group of parameters and the solution space of some desired simulator variables (Castelletti et al., 2012). This has been successfully applied to several dynamic models (Carbajal et al., 2017; Conti and O'Hagan, 2010; Machac et al., 2016a; Yang et al., 2018).

This chapter comprises two sections; First an example is given in which a formal Bayesian inference scheme is applied to update prior knowledge on the river water quality and quantity submodel parameters of the Dommel system. The original model is computationally expensive and it should be evaluated during long time windows (months-years) to capture local and seasonal dynamics. Thus, the classical use of MCMC posterior sampling algorithms (which require on the order of tens of thousands samples) is beyond the available time and computational resources. Consequently, a dynamic emulator was created to link the river water quality and quantity parameters to the one-year (hourly frequency) output time-series of flow and dissolved oxygen concentration at the closing section of the Dommel system (station M_0121, Figure 1.4). Using this emulator, several likelihood structure hypotheses could be tested during the inference process, allowing to discuss the residual structure of the posterior inferred dynamics.

Secondly, classical model emulation techniques are severely limited to the use of a reduced number of parameters under a given dynamic period. This means that proposing an emulator for a certain hydraulic or hydrological process (for instance a rainfall-runoff model) is relatively feasible for a low number of static parameters (e.g. roughness or parameterised geometrical features). Yet, once trained, the emulator is only valid within that fixed interval of time. Thus, if the user wants to explore the effect of a different rainfall event (or dynamic process in the general case), the emulator should be retrained under the new conditions. This implies to sample again from the original computationally expensive simulator to build a new training database. In the second section of this chapter, a novel method is presented which allows including rainfall dynamic inputs of arbitrary length along with static parameters in an emulator structure for physically based flow simulation. After the training process, this emulator structure can be used to approximate the response of the model to new rainfall events and parametric scenarios in a fraction of the original simulator running time. This opens the applicability of physically based hydrodynamic simulation (e.g. 2D-SWE or other CFD based flood propagation implementations) to fast sampling schemes, as model-based flooding early warning schemes, real time control under uncertain data or propagation of uncertainties in rainfall and parameters for hydrodynamic simulators.

3.2. FACILITATING PARAMETRIC INFERENCE IN LARGE WATER QUALITY RIVER SYSTEMS *

3.2.1. INTRODUCTION

THE description of processes in environmental modelling is seldom purely physically based, which is due to an incomplete understanding of the real underlying dynamics, to the lack of field measurements or due to a need of simplification. This often leads to the use of non-physical parameters, which cannot be directly measured or that lump several processes. The numerical value of such parameters is determined by calibration to ensure that the model and reality have a quantifiable degree of resemblance. The transferability of parameters from one system to others is typically limited, yet the modeller often has some prior knowledge acquired by simulating similar cases, which could be used in the calibration process. This process is often approached from a Bayesian perspective, in which the modeller encodes its knowledge as a joint probability distribution of the parameters, which are updated in view of new information.

Integrated urban water modelling focuses on the joint simulation of processes affecting water dynamics through the urban-river system (Muschalla et al., 2009; Rauch et al., 2002). These models jointly evaluate wastewater treatment processes, urban drainage and river dynamics, which usually generates a rapid escalation of complexity (Benedetti et al., 2013a). The representation of all subsystems involved, produces highly parameterised conceptualisations, requiring a large amount of data in the identification-calibration process (Langeveld et al., 2013a). Additionally, the dynamics of interest often occur at very different time-space scales, for instance, urban combined sewer overflow (CSO) discharges have a characteristic time of minutes-hours whereas river dissolved oxygen dynamics are at hourly-monthly scale.

Quantifying and analysing uncertainties in integrated catchment modelling platforms is required to avoid over-confidence in modelling results and to guide further model improvements (Deletic et al., 2012; Tscheikner-Gratl et al., 2017). Most uncertainty analysis strategies rely on sampling intensive techniques (Dotto et al., 2012). For instance, Bayesian inversion schemes often require a large number of model evaluations (on the order of $10^4 - 10^5$) to reach convergence. This renders the inference in the original simulation platform impractical for most real-scale model platforms, which might take between minutes to hours per sample. Hence, the computational effort required is a severe limitation for the applicability of uncertainty analysis techniques in most real-scale integrated catchment modelling studies (Tscheikner-Gratl et al., 2019).

One approach to speed up convergence time is the use of optimised sampling schemes; as drawing samples in parallel (Goodman and Weare, 2010; Laloy and Vrugt, 2012) or with informed adaptive Markov chain Monte Carlo (Hoffman and Gelman, 2014). However, this still often requires a prohibitive number of model samples (for practical engineering standards). Another strategy is the use of data-driven or mechanistic model emulation, where a black-box (or grey-box) mathematical representation is used to replicate the link between a vector of parameters and the dynamic response of the simulator.

*This section is an adapted version of: Moreno-Rodenas, A., Langeveld, J., and Clemens, F. (2019). Parametric emulation and inference in computationally expensive integrated urban water quality simulators. Environmental Science and Pollution Research. doi: 10.1007/s11356-019-05620-1.

For instance, Laloy et al. (2013) proposed a two stage sampling scheme, first generating a rough estimate from a model surrogate and later from the simulator itself to perform parametric inference in a groundwater model. Carbajal et al. (2017) compared the performance of mechanistic vs. data-driven emulation for urban drainage simulators, concluding that in general a data-driven approach is to be preferred unless confronted with highly sparse training datasets. Yang et al. (2018) used a Gaussian process data-driven emulator to perform a parametric uncertainty analysis in a semi-distributed hydrological model. Those approaches are useful when the dimensionality of the problem allows for the fit of a data-driven emulator.

This subsection deals with a large-scale urbanised catchment model integrating wastewater treatment processes, urban drainage and receiving water dynamics for the estimation of dissolved oxygen depletion in a formal Bayesian inference scheme. Prior knowledge on the probability distribution of river water quality and quantity parameters is updated using flow and dissolved oxygen system observations. The simulator platform is too computationally expensive to be directly used in the sampling scheme. Thus two emulation platforms are created to represent a mapping for four water quantity and eight water quality process parameters for river flow and dissolved oxygen concentration. The training is performed under a given dynamic realisation of the system during the full year of 2012. The performance of the two emulators is validated under an independent dataset. The emulators are used to implement a global sensitivity analysis and an inference scheme under various likelihood function conceptualisations. The residual structure of the posterior mean sample is contrasted with the building hypotheses of the likelihood function. This example shows how the use of a dynamic emulator scheme can assist in gaining further knowledge in the sensitivity and probably distribution of process-driven model parameters in a large-scale water quality simulator.

3.2.2. MATERIALS AND METHODS

3.2.2.1. DYNAMIC EMULATION

Two polynomial chaos expansion (PCE) structures (Xiu, 2010) were used to emulate flow and dissolved oxygen dynamic series during one year at hourly frequency (01-01-2012 – 31-12-2012). The PCE emulator aims at creating an interpolation map between the parameters and model output multi-dimensional spaces, such that the effect on the model output of additional parameter combinations can be readily approximated.

The model (M) can be though of a coupled system of partial or ordinary differential equations, which represents the water quality processes of an integrated catchment system. An arbitrary time-dependent output state variable ($Y_{sim} \in \mathbb{R}^D$) can be computed by solving the model:

$$Y_{sim} = M(\mathbf{x}_0, \mathbf{x}(t), \theta_M, \theta_I) \tag{3.1}$$

given a set of m initial conditions ($\mathbf{x}_0 \in \mathbb{R}^m$), a set of dynamic inputs ($\mathbf{x}(t)$) and a number of global model parameters $[\theta_M, \theta_I]$. The set of global parameters is decomposed in two sub-groups, being $\theta_M \in \mathbb{R}^S$ a group of global parameters which are considered fixed (during the emulation), and $\theta_I \in \mathbb{R}^P$ a number of P global parameters which the modeller seeks to emulate.

The PCE is composed of a series of N orthogonal polynomials ($\Phi(\theta_I) \in \mathbb{R}^{Nx1}$), such that:

$$Y_{sim}(t,\theta_I) \approx \Phi(\theta_I)^T \cdot \mathbf{c} \tag{3.2}$$

where $\mathbf{c} \in \mathbb{R}^{NxD}$ is a matrix of coefficients which is fitted based on samples drawn from the simulator. Thus, Equation 3.2 forms a mapping between the parameter and output spaces ($\mathbb{R}^P \to \mathbb{R}^D$). The training dataset is pre-computed by evaluating the model response (Equation 3.1) at a number of K combinations of the parameter vector ($\theta_I = q_i for i = 1...K$). The training dataset is then used to calibrate the matrix of coefficients ($\mathbf{c}$):

$$\begin{bmatrix} Y_{sim}(t,\theta_I = q_1) \\ \vdots \\ Y_{sim}(t,\theta_I = q_N) \end{bmatrix} = \begin{bmatrix} \Phi_1(q_1) & \dots & \Phi_N(q_1) \\ \vdots & \ddots & \vdots \\ \Phi_1(q_K) & \dots & \Phi_N(q_K) \end{bmatrix} \cdot \begin{bmatrix} c_1(t) \\ \vdots \\ c_N(t) \end{bmatrix} \tag{3.3}$$

A least squares approach is used to approximate the value of the coefficients **c** based on known realizations of the model structure and the computed value of the polynomial series (only dependent on the emulated parameter vector). Then, Equation 3.2 can be used to approximate the output variable at new combinations o the emulated model parameters (θ_I). Further information about the fit of the polynomial expansion and the selection of orthogonal series can be found at Hadigol and Doostan (2018); Feinberg (2015); Xiu and Karniadakis (2002).

A training database of 200 samples was created by drawing realisations of flow series using a Latin hypercube sampling scheme (LHS) for four static parameters. Table 3.1 presents the river flow parameters along with the parametric range used to create the PCE database. Parametric ranges for the emulator training dataset were designed to fully cover the probability density distribution of the parameter prior-knowledge as extracted from expert elicitation, measurements and literature.

Table 3.1: River hydrology parameter PCE training ranges (emulation) and prior distributions (inference). ~U(a,b) refers to a uniform distribution between a and b.

Name	Units	Description	PCE training	Prior distribution
n	$s \cdot m^{-1/3}$	Manning's roughness	~ U(0.02, 0.15)	~ U(0.025, 0.12)
k_z	-	Embankment slope multiplier	~ U(0.3, 2)	~ U(0.7, 1.3)
k_W	-	River bed width multiplier	~ U(0.3, 2)	~ U(0.5, 1.5)
k_h	-	Rural flow input multiplier	~ U(0.3, 2)	~ U(0.7, 1.3)

A second emulator was built by simulating dissolved oxygen dynamics with the full-integrated model under the mean inferred flow parametric values. A LHS scheme was used to draw 200 samples from an eight-dimension parameter space of river water quality process parameters (Table 3.2). The process equations related to these parameters are shown at Appendix B. Both training samples were taken by fixing the rest of parameters and input stochastic processes of the integrated model to their mean value.

A Legendre orthogonal polynomial series truncated at third order was used in both emulators (Gautschi, 1994). Training was done by a non-intrusive least squares fitting of the polynomial coefficients (Hadigol and Doostan, 2018) using the implementation of Feinberg and Langtangen (2015). A detailed description of the polynomial expansion construction and fitting methodology used can be found at Moreno-Rodenas et al. (2018b). An independent validation set was created using a LHS of 100 samples for flow and 50 for dissolved oxygen. Nash-Sutcliffe efficiency was computed between the emulated and simulated one-year time series to validate the emulator's performance.

Table 3.2: River dissolved oxygen parameter PCE training ranges (emulation) and prior distributions (inference)

Name	**Units**	**Description**	**PCE training**	**Prior distribution**
Kd1	d^{-1}	Decay rate for BOD fast	~U(0.3, 1)	~U(0.3, 0.8)
Kd2	d^{-1}	Decay rate for BOD slow	~U(0.2, 1)	~U(0.2, 0.4)
Vs1	$m \cdot d^{-1}$	Sedimentation rate for BOD fast	~U(0.2, 40)	~U(0.5, 20)
Vs2	$m \cdot d^{-1}$	Sedimentation rate for BOD slow	~U(5, 100)	~U(10, 60)
TKd	-	Temperature coefficient for BOD oxidation	~U(1, 1.1)	~U(1, 1.1)
TKL	-	Temperature coefficient for reaeration	~U(1, 1.1)	~U(1, 1.03)
TSOD	-	Temperature coefficient for SOD	~U(1, 1.1)	~U(1, 1.1)
VKL	-	Velocity reaeration coefficient	~U(2, 8)	~U(2, 5)

3.2.2.2. PARAMETRIC INFERENCE

Prior parameter knowledge was encoded by means of uniform probability density functions. Ranges were defined based on literature values and expert criteria (non-formal elicitation). Table 3.1 and Table 3.2 show the prior probability density function selected for each parameter.

This prior knowledge was updated by using an observation layout $Y_{obs} \in \mathbb{R}^{1xn}$ for hourly measured flow and dissolved oxygen concentration at the outlet of the Dommel catchment during a period of approximately seven months (15-Jan-2012 – 05-Aug-2012). The basic model-observation layout was defined as:

$$Y_{obs} = M(\mathbf{x}_0, \mathbf{x}(t), \theta_M, \theta_I) + Z \tag{3.4}$$

where M refers to the system of differential equations representing the integrated model (urban drainage, WWTP and receiving water body), which depends on a set of initial conditions $\mathbf{x}_0$ (computed by a warming up simulation period of 1 year for the WWTP initial conditions and a dedicated general initialisation of the previous 15 days between 01-Jan-2012 until 15-Jan-2012 for the rest of the variables), which are fixed in the sampling scheme. A set of dynamic inputs $\mathbf{x}(t)$, containing distributed urban rainfall estimations, measured river and WWTP water temperature, river solar radiation, boundary conditions for river tributaries and control strategies. θ_M represents a vector of m global model parameters, which were fitted by submodel calibration and expert elicitation. Whereas θ_I represents the P model parameters used in the inference scheme for the river flow and dissolved oxygen dynamics. The term Z refers to the residual structure between the simulated ($M(\mathbf{x}_0, \mathbf{x}(t), \theta_M, \theta_I)$) and the observations ($Y_{obs}$). This error term lumps measurement and model errors together. During the inference scheme, a probabilistic description of the model-measurement residuals Z is assumed a priori, and is later validated based on the posterior computed residuals. A common initial guess is to assume that residuals are independent, identically and Gaussianly distributed. This assumption leads to the following log-likelihood structure:

$$\ell(\theta_I|Y_{obs}) \propto -\frac{1}{2}log(det(\Sigma)) - \frac{1}{2}(Y_{obs} - M(\theta_I))^T \cdot \Sigma^{-1} \cdot (Y_{obs} - M(\theta_I)) \tag{3.5}$$

where Σ represents the residual covariance structure. When assuming that residuals are uncorrelated and present a constant variance (homoscedasticity), the covariance takes the form:

$$\Sigma_Z = \sigma_1^2 \cdot \mathbf{I} \tag{3.6}$$

being $\mathbf{I} \in \mathbb{R}^{nxn}$ the identity matrix and σ_1^2 a constant variance of the residuals.

Flow dynamics are well known to render heteroscedastic error structures. This implies that residuals trend to be systematically larger when the discharge is larger. This is often encoded by assuming that the residual standard deviation follows a linear relationship with the simulated variable. This results in a log-likelihood function with the form described by Equation 3.5 and the following covariance matrix:

$$\Sigma_{Z\_het} = (\sigma_1 + \mathbf{Q} \cdot \sigma_2)^2 \cdot \mathbf{I} \tag{3.7}$$

where $\mathbf{I} \in \mathbb{R}^{nxn}$ is the identity matrix, $\mathbf{Q} \in \mathbb{R}^{1xn}$ is the computed variable vector, meanwhile σ_1 and σ_2 are the standard deviation hyperparameters of the error generating process.

Also, the inference of dynamic models often leads to autocorrelated residual structures. This is often taken into account by the use of a discrete autoregressive model of order p (Bates and Campbell, 2001), or as formulated by Schoups and Vrugt (2010):

$$\Phi_p(B) \cdot z_t \sim N(0, \sigma_1) \tag{3.8}$$

being $\Phi_B(B) = 1 - \sum_{i=1}^{p} p_i \cdot z_{t-i}$ an autoregressive polynomial of order p for the residual z_t, with Gaussian updates.

An equivalent formulation to account for a correlation structure was discussed by Honti et al. (2013), with the use of a bias stochastic process along with the error generating model (Z). If assuming a stationary continuous constant bias and heteroscedastic residuals, Equation 3.5 defines the log-likelihood function, with a covariance matrix defined as:

$$\Sigma_{(Z\_het+B)_{ij}} = (\sigma_1 + Q(t_i) \cdot \sigma_2)^2 \cdot \delta_{ij} + \sigma_3^2 \cdot e^{-|d_{i,j}| \cdot \tau^{-1}} \tag{3.9}$$

which are the i and j elements of the covariance matrix $\Sigma_{Z\_het+B} \in \mathbb{R}^{nxn}$, with δ_{ij}, the Kronecker's delta, $Q(t_i)$ the expected flow (at time t_i), $d_{i,j}$ the distance in hours between i and j elements, σ_3 the standard deviation of the stationary bias, and τ an extra hyperparameter which drives a correlation exponential decay. Del Giudice et al. (2013) discusses, that in practice the model effect and the bias descriptors can have a poor identifiability, thus inferring both, model parameters and bias hyperparameters, often requires assigning strong priors to the latter. Table 3.3 presents the prior distribution for the hyperparameters of the different likelihood distribution structures.

Table 3.3: Error model hyperparameters for the different hypotheses

Hyperparameter	Units	Description	Prior distribution
Flow i.i.d Gaussian			
σ_1	m^3/s	Std1, stationary standard deviation error	~U(0, 10)
Flow independent heteroscedastic Gaussian			
σ_1	m^3/s	Std1, stationary standard deviation error	~U(0, 10)
σ_2	m^3/s	Std2, stationary standard deviation error	~U(0, 10)
Flow AR(3) Gaussian updating			
σ_1	m^3/s	Std1, stationary standard deviation error	~U(0, 10)
$p_{1,2,3}$	-	Autocorrelation coefficients p1, p2, p3	~U(0, 1)
Flow Heteroscedastic normal error and exponentially correlated bias			
σ_1	m^3/s	Std1, linear intercept standard deviation error	~U(0, 10)
σ_2	m^3/s	Std2, linear slope standard deviation error	~U(0, 10)
σ_3	m^3/s	Std3, bias standard deviation	~U(0, 10)
τ	h	Tau, bias correlation exponential decay	~U(10, 80)
Dissolved oxygen i.i.d Gaussian			
σ_1	mgO_2/l	Std1, stationary error standard deviation	~U(0, 10)

Posterior samples were created using a Metropolis-Hasting algorithm (Hastings, 1970; Metropolis et al., 1953). The joint prior-probability distributions for the flow and dissolved oxygen river parameters were updated by drawing 50,000 samples from their posterior distribution by means of a Markov-chain Monte-Carlo sampling scheme (25,000 burn-in, 5 thinning). The Bayesian inference implementation was performed using the python probabilistic programming package pymc version 2.3.6 (Patil et al., 2010).

Evaluating the likelihood distribution when using the full bias-description term involves inverting a covariance matrix of size n. In this case this was prohibitively expensive when using the original measurement layout (n=4892). Thus, a shorter period was used to test the inference of the bias description (26-Jul-2012 – 14-Sep-2012). In this case, only 4,000 accepted samples were used (2,000 burn-in, 2 thinning).

3.2.3. RESULTS AND DISCUSSION

3.2.3.1. DYNAMIC EMULATION OF FLOW AND DISSOLVED OXYGEN CONCENTRATION

The performance of the trained emulator to represent the integrated catchment model outputs (at new parameter combinations) was tested using an independent dataset. Figure 3.1 shows the Nash-Sutcliffe efficiency (NSE) at 100 samples of the parametric space for the flow quantity simulation, which were used for validation of the emulator performance. Figure 3.2 shows the same test performed at 50 samples drawn from the dissolved oxygen concentration emulation scheme. The performance of both emulator implementations is consistent across the parameter ranges and varies between 0,99-1 NSE. The observed performance during validation was considered sufficient for the application of the inference scheme in this example. Figure 3.3 and Figure 3.4 provide a graphical comparison of the emulator and simulator responses for different random combination of the parametric space (independent form the training dataset).

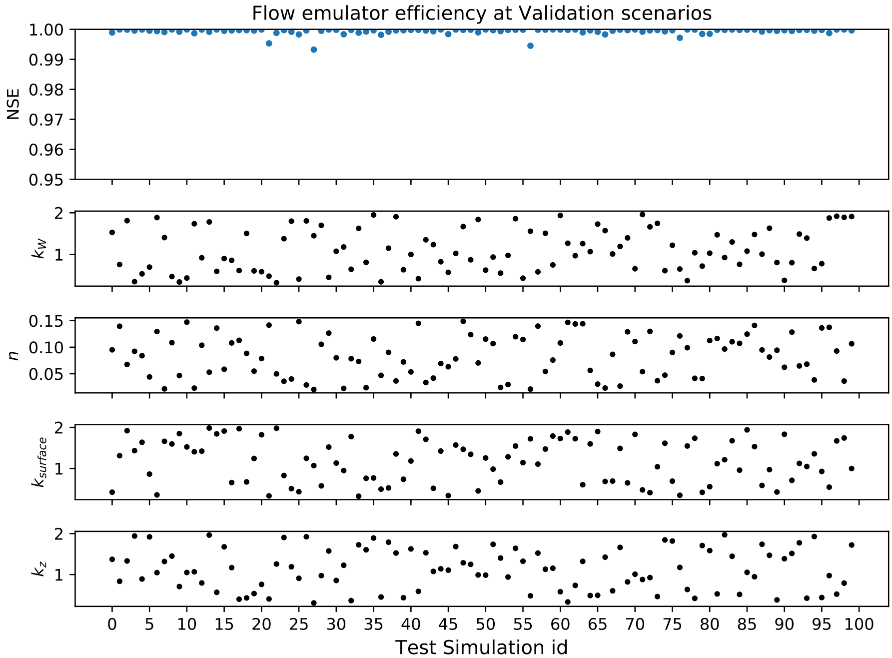


Figure 3.1: Nash-Sutcliffe efficiency at the flow emulator vs simulation for a four-dimensional parameter space under validation conditions.

When increasing the dimensionality of the parametric space in the emulation scheme, the number of required training samples grows rapidly (Xiu and Karniadakis, 2002). This causes that for some parametric space dimension, training the emulator becomes impractical, or to be a similar burden as directly sampling from the simulator. Table 3.4 presents the computational effort required to sample from the original simulator, training and operation of the emulator. In this case, those are the computed average timings when using a 2.2 GHz Intel Core i7 from mid 2014.

Table 3.4: Emulation vs. Model computational effort for one-year series (in seconds).

Sample	**Flow**	**DO**
Simulator sample	3300 s	3300 s
Training database (x200 simulator samples)	$660x10^3$ s	$660x10^3$ s
Training the emulator	14 s	61 s
Emulator sample	0.06 s	0.07 s

3.2.3.2. GLOBAL SENSITIVITY ANALYSIS OF PROCESS PARAMETERS

The emulator structure was used to estimate the sensitivity of the integrated catchment model outputs to variations of the river physical and biochemical parameters. Sam-

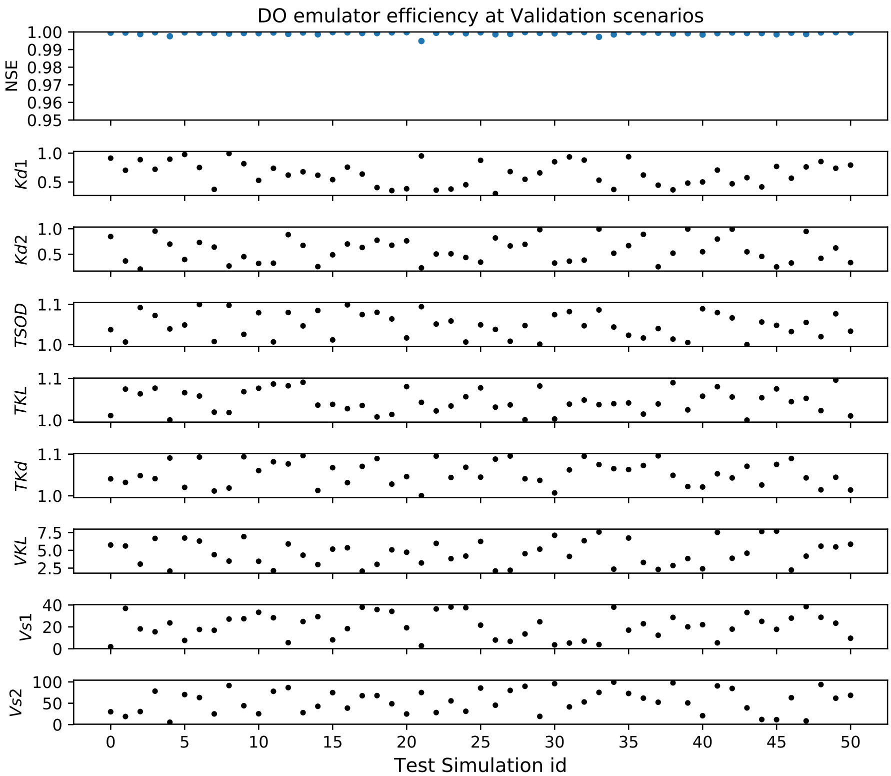


Figure 3.2: Nash-Sutcliffe Efficiency of dissolved oxygen emulator vs simulator for an eight-dimensional parameter space under validation conditions.

ples from the emulator were used to compute the Sobol global sensitive indexes (Sobol, 1993). Figure 3.5 and Figure 3.6 depict the first order indexes from the prior distribution of parameters at the river flow and DO dynamics. Figure 3.5 shows that the simulated flow level is highly sensitive to the parameter $k_{surface}$ during dry-weather periods, whereas the Manning's roughness (n) becomes relevant during the rising limb of the hydrographs. k_W (multiplier for the river bed width) shows a reduced influence to the overall dynamics, meanwhile k_z (a multiplier for the slope of the embankment) has a similar, yet less pronounced effect when compared to hydraulic roughness.

The results of the study of the sensitivity for the DO concentration simulation is shown in Figure 3.6. The parameter controlling the reaeration rate (VKL) dominates the dry-weather DO variability during summer times. This influence decreases during winter, where the temperature coefficient for the sediment oxygen demand (TSOD) becomes increasingly relevant. This has to do with the temperature inhibition model structure, which influences the oxidation rate of organic matter for temperatures differing from 20 degree Celsius. During oxygen recovery patterns, TSOD is also relatively relevant, since the main oxygen sink is the sediment layer. Sensitivity indexes do not show a consistent behaviour during acute oxygen depletion processes. Some depletion events,

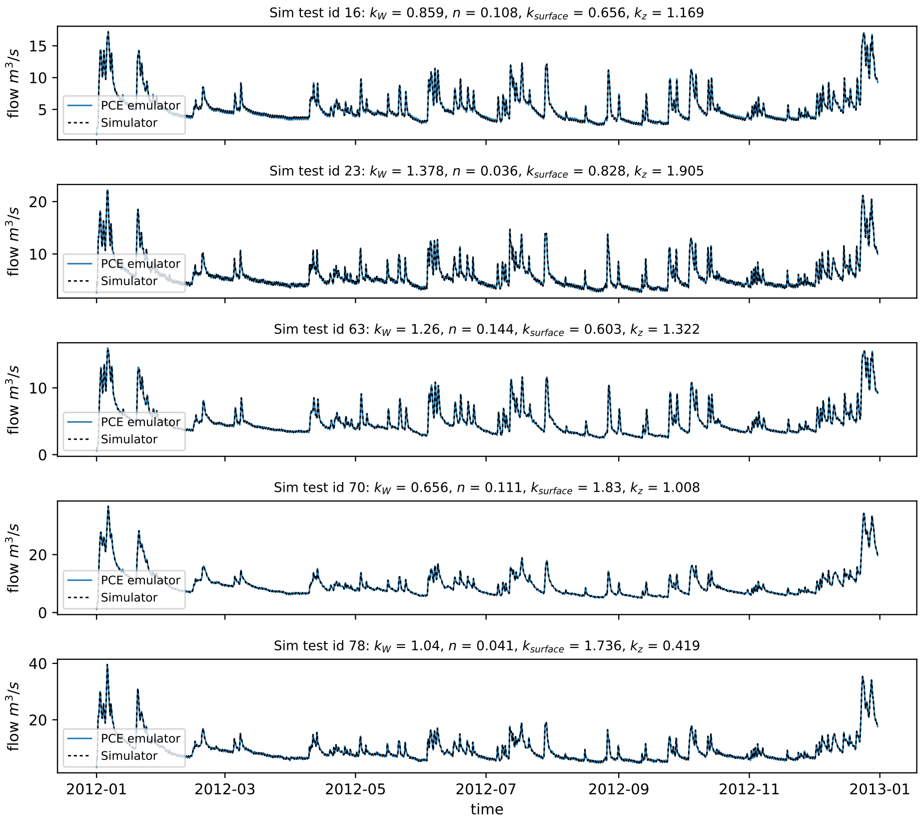


Figure 3.3: Emulator vs. Simulator river discharge graphical comparison for different test parameter combinations.

as the three occurring during July and September (also seen at Figure 3.6.b) present as dominant parameters kd1 and kd2 which are the oxidation rates for the two fractions of suspended BOD in the system. However, the events occurred in June and the three in October showed to be more sensitive to a different parameter combination as TKL or TSOD, which are related to temperature driven reaeration or oxygen consumption (also slightly to kd1). This is a good example of the complexity of the underlying process, in which interactions are highly dependent on the dynamic state of the system. For instance, if a storm event activates predominantly northern CSOs, which are closer to the outlet of the catchment, there is less time for the degradation of suspended matter to occur, thus variations in the consumption rate might have a higher influence the process. Also, events in which the WWTP is the main source of discharge (and not the CSOs), the settling facilities of the WWTP might lead to a lower sediment build-up in the river and thus increasing the relevance of suspended organic matter degradation.

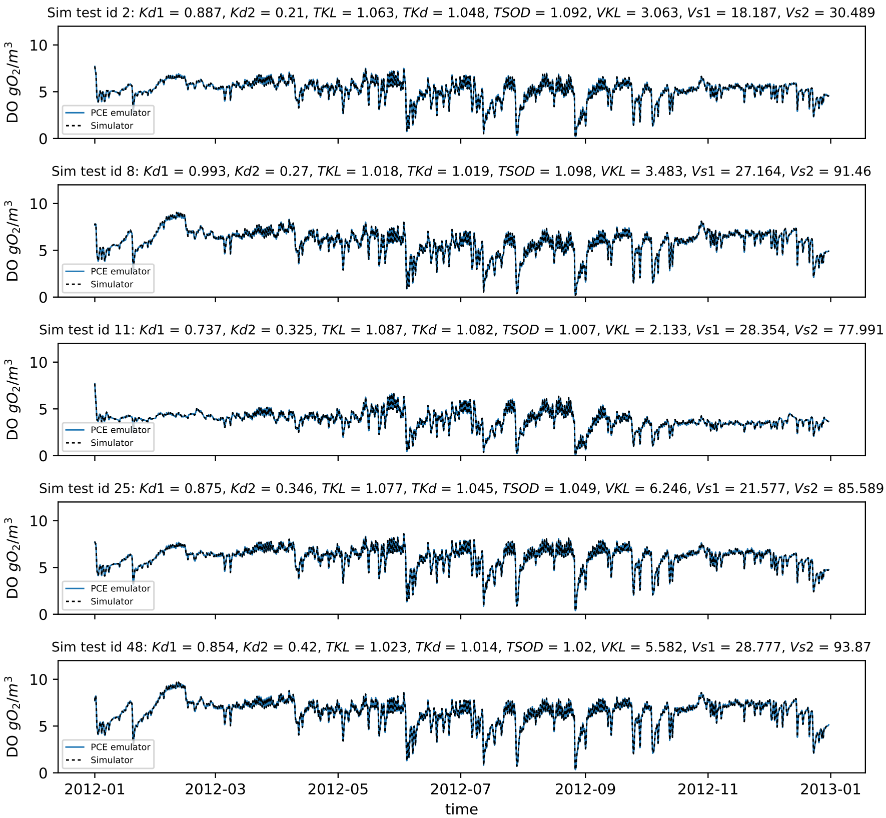


Figure 3.4: Emulator vs. Simulator dissolved oxygen time-series graphical comparison for different test parameter combinations.

3.2.3.3. PARAMETRIC INFERENCE

A data series of river flow and dissolved oxygen concentration (15-Jan-2012 until 05-Aug-2012) at M0121 station (the outlet of the catchment) was used to update prior knowledge (encoded as a probability distribution, Table 3.1 and Table 3.2). Figure 3.7 displays the comparison of measured flow (in black dashed line) and the posterior mean simulation (solid blue) plus its 95% uncertainty range. Also, a validation period (05-Aug-2012 until 31-Dec-2012) shows that the dynamics of the system are consistent. Flow series are reasonably well represented, although extremely dry periods induce a systematic overestimation of the flow as seen in July and in the beginning of September, this however is expected to have little influence in the water quality dynamics. The same comparison (measurement vs inferred and validation series) can be found in Figure 3.8 for the simulation of dissolved oxygen. The general dynamics of DO are captured, especially the depletion processes, daily and seasonal variation. Yet the uncertainty of the estimation is significantly large compared with the variable itself. This has certain implications for the applicability of the DO series in environmental studies. The use of DO deterministic estimations can lead to overconfidence on modelling results. During the 19th and the

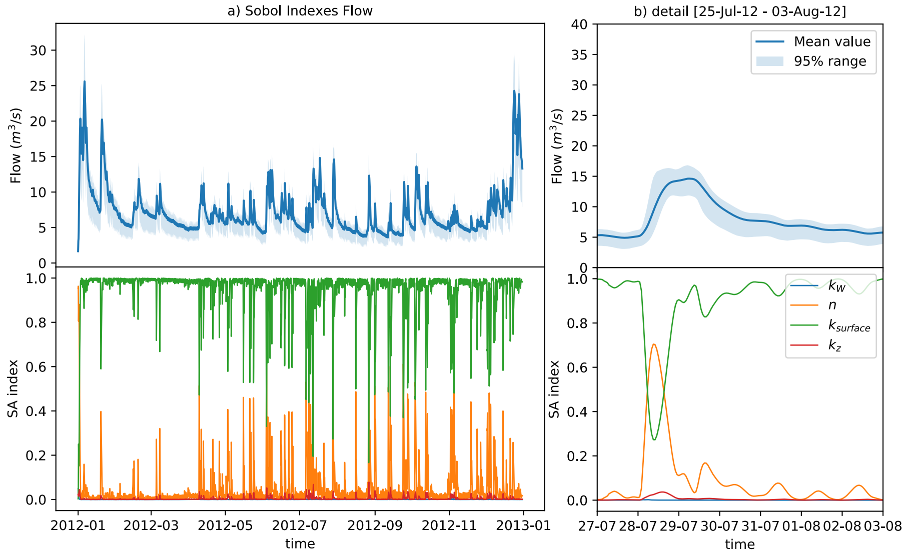


Figure 3.5: Sobol sensitivity indexes (first order) for the flow dynamics. Above, mean flow simulation and the 95% interval for the propagation of the parametric ranges. Below, sensitivity indexes for the four parameters. In the right (b) detail of the sensitivity during a medium-high intensity storm event.

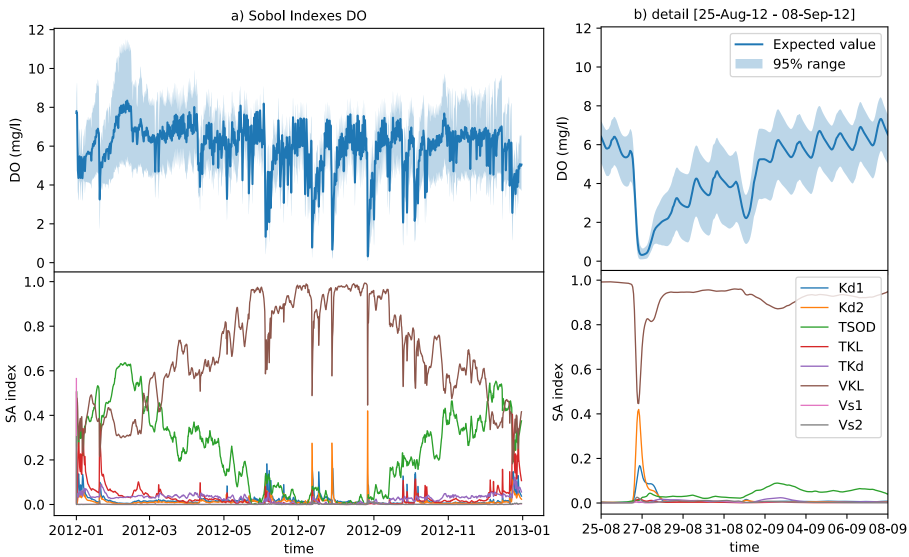


Figure 3.6: Sobol sensitivity indexes (first order) for the dissolved oxygen dynamics. Above, mean DO simulation and the 95% interval for the propagation of the parametric ranges. Below, sensitivity indexes for the eight parameters. In the right (b) detail of the sensitivity during a high intensity storm event.

29th of November 2012, Figure 3.8 shows a systematic mismatch of three acute depletion events, which are not captured by the model. This indicates that the model structure is apparently lacking some fundamental process, which has not been identified yet. The validated measuring data quality in this period is considered as good. This kind of depletion processes for the Dommel system should be further investigated in order to update the structure of the integrated catchment model in the future.

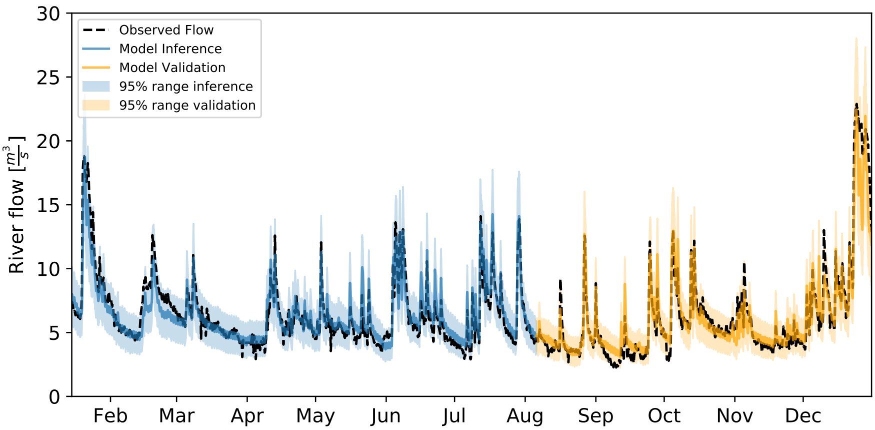


Figure 3.7: Posterior sample for the inferred flow dynamics between 15-Jan-2012 and 05-Aug-2012. In orange, the posterior distribution under validation conditions 05-Aug-2012 until 31-Dec-2012, in black observed flow at the station M0121.

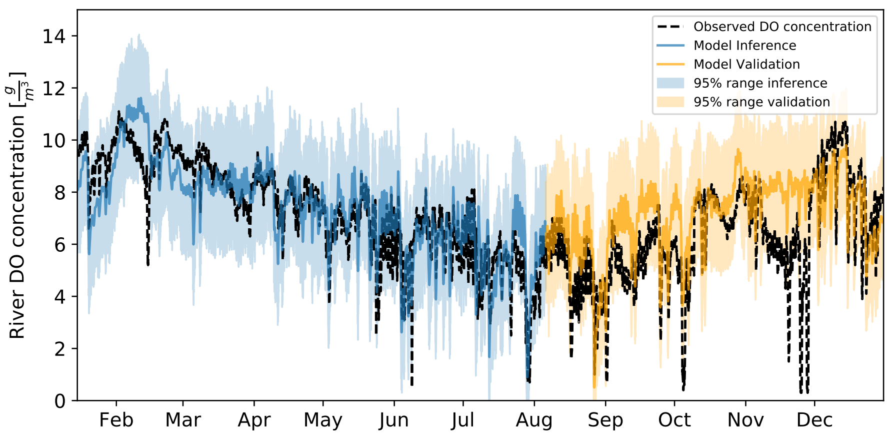


Figure 3.8: Posterior sample for the inferred dissolved oxygen dynamics between 15-Jan-2012 and 05-Aug-2012. In orange, the posterior distribution under validation conditions 05-Aug-2012 until 31-Dec-2012, in black observed flow at the station M0121.

The posterior probability density functions of the parameters for the water quantity and quality of the river section can be found at Figure 3.9 and Figure 3.10 respectively. The river parameter k_W is poorly identified, which is denoted by the wide range of the posterior distribution (Figure 3.9, k_W diagonal). This is also supported by the very low sensitivity of this parameter to the overall flow dynamics (Figure 3.5). The rest of the parameters are appropriately identified, and they appear to be mostly mutually independent with the exception of a strong negative correlation between k_z and n. This implies that both parameters interact in opposite directions, thus partly compensating each other. The joint inference/calibration of both elements is therefore not recommended, since there is not enough information content in the objective variable to identify them independently. Further use of this model should therefore prioritize fitting n, since this parameter has a much larger sensitivity than k_z. Water quality variables show a mostly independent joint posterior distribution with the exception of $kd1$ and $kd2$, which show a mild negative correlation. This is explained by the fact that both parameters influence the same process at two fractions of BOD for which DO measurements are probably insufficient to discriminate.

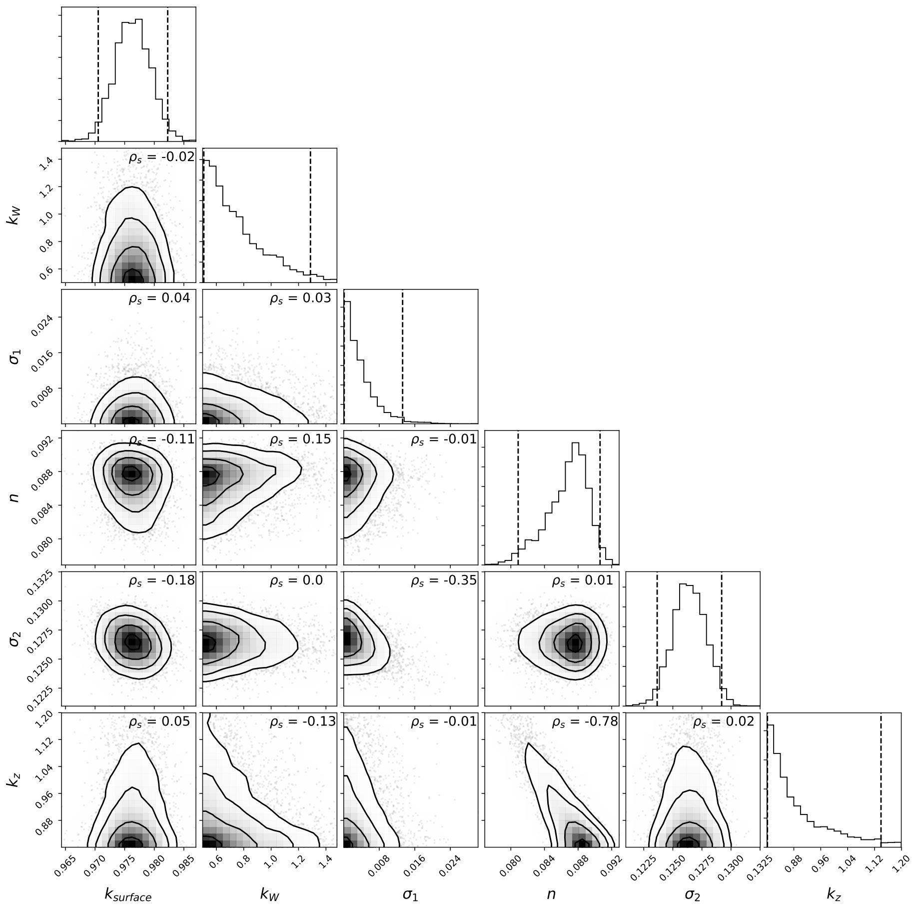


Figure 3.9: Posterior joint parametric distribution for the inference of the flow model parameters. σ_1 and $sigma_2$ are hyperparameters of the selected error generation process (heteroscedastic, independent Gaussian). The Spearman's correlation coefficient (ρ_s) is also shown at each parameter couple. The black dashed lines in the diagonal histogram plots represent the 95% data range.

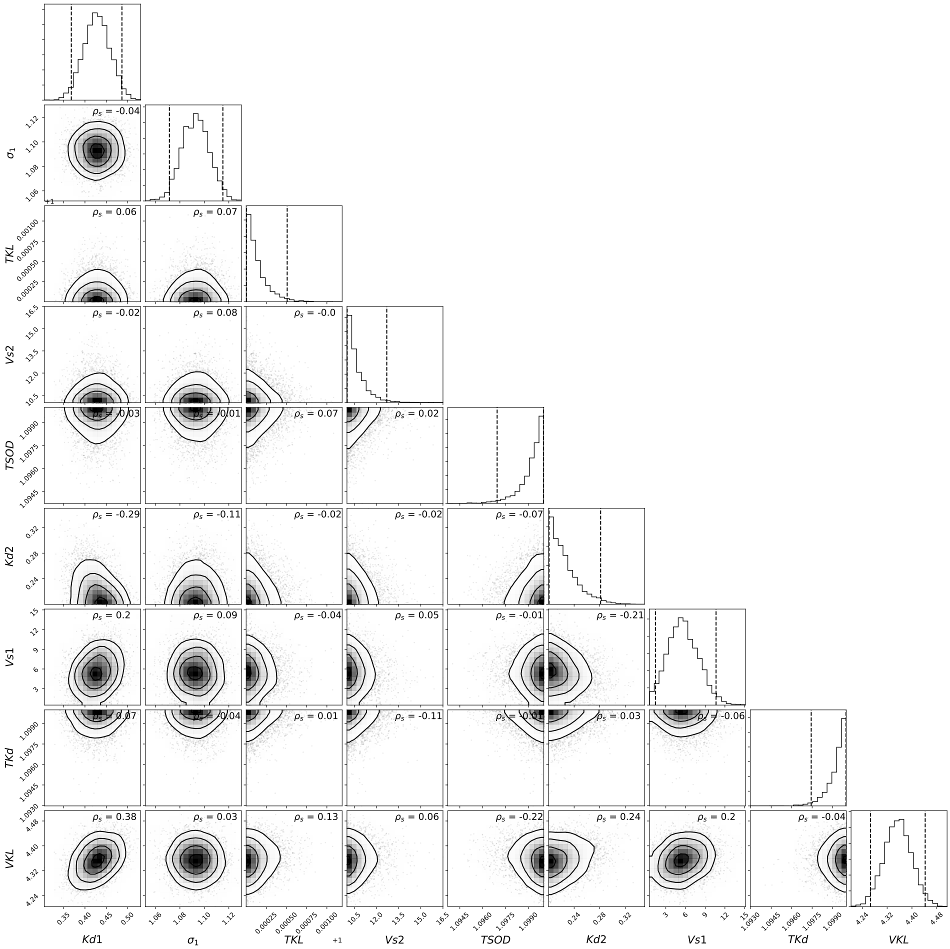


Figure 3.10: Posterior joint-parametric distribution for the inference of the water quality model parameters. σ_1 is the hyperparameter of the selected error generation process (independent, identically distributed Gaussian). The Spearman's correlation coefficient (ρ_s) is also shown at each parameter couple. The black dashed lines in the diagonal histogram plots represent the 95% data range.

3.2.3.4. ERROR GENERATING PROCESS AND LIKELIHOOD DESCRIPTION

Bayesian inference relies on the a priori definition of an error generating process (Equation 3.4), which constitutes the likelihood structure used to define the probability of a parameter sample being supported by the observations. The error generating process is selected based on a series of hypotheses, which can be encoded by expert guesses on the behaviour of the system. Yet those assumptions are still a subjective exercise and its validity should be checked once sampled from the posterior distribution. In this case, the initial error generation process for both flow and DO series was conceptualised as an independent, identically distributed Gaussian distribution. Once the first posterior samples for the flow inference are available it becomes apparent that the residual structure shows a mild heteroscedastic structure. This implies in this context that the residuals are positively correlated with the flow. This is well described in the hydrological literature (Sorooshian and Dracup, 1980) and is corrected by using a linear dependent variance structure in the error generating process for the flow inference. Figure 3.11 represents three relevant characteristics of the residual structure at the posterior samples of river flow, comparing the assumed error generating process (in black) and the sampled one (in blue). Figure 3.11.a shows the heteroscedasticity structure of the residuals, Figure 3.11.b shows the comparison of probability density functions and Figure 3.11.c shows the time-autocorrelation structure of the assumed and computed residuals. It is apparent that the residual independency assumption is violated, since computed residuals present a strong time autocorrelation structure. Figure 3.12 shows the comparison of the computed residuals and the assumed error generating process (independent, homoscedastic and Gaussian) for the dissolved oxygen in the river. The variance of the residuals is largely independent from the DO value. The proposed and computed probability densities have a reasonable match. Also, residuals present a clear autocorrelation structure, albeit shorter than that of the flow inference.

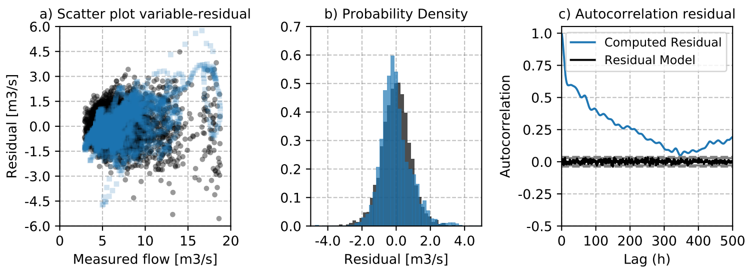


Figure 3.11: Residual structure at the flow posterior mean sample, a) scatter plot variable-residual showing the dependency of the variance, b) the residual probability density and c) the autocorrelation plot at different time-lag.

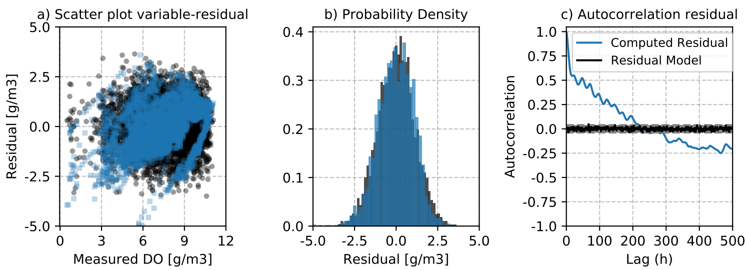


Figure 3.12: Residual structure at the dissolved oxygen posterior mean sample, a) scatter plot variable-residual showing the dependency of the variance, b) the residual probability density and c) the autocorrelation plot at different time-lag.

The time autocorrelation structure in hydrological inference has been discussed in several studies. For instance, Kuczera (1983) applied an ARMA (autoregressive moving-average) model to represent an autocorrelated likelihood structure in a hydrological model. Bates and Campbell (2001) argued that ARMA structures lead to local minima, and AR (autoregressive) models of order p are to be preferred. Schoups and Vrugt (2010) presented the use of an alternative likelihood structure, which addresses several common issues; like the non-normality of residuals (which was not relevant in our study), the variance non-stationarity (corrected by the use of a linear standard deviation dependency) and the temporal correlation of the residuals approached with an AR(p) model. Yet all these three studies simulated catchment hydrological flows at daily scales. In the examples presented in this thesis, the measurement layout has an hourly time-step, since relevant processes occur at those scales. It is expected that the autocorrelation structure becomes stronger when dealing with shorter time-scales. As seen in Figure 3.11 and Figure 3.12 the correlation is still around 0.5 in the order of 50-100 hours lag time. Honti et al. (2013) and Del Giudice et al. (2013) presented the direct encoding of a Bias description process within the error model. This was applied to urban drainage hydrodynamic simulation with time-steps of 1-2 minutes, which present strong autocorrelation structures. The Bias description can be conceptualised as function of different variables or inputs, yet in its basic form it constitutes a Gaussian multivariate distribution with an exponential covariance structure as in Equation 3.9.

The use of an AR(3) model in this case, rendered an almost negligible effect of the autoregressive parameters of higher order than one (<0.01), thus generating an equivalent AR(1) model. Measured series vs. inferred comparison and the residual structure can be seen at Figure 3.13. Although the autocorrelation of residuals is better represented, the fit of the mean sample did not improve, rather was degraded through accounting for the autocorrelation term. This was also discussed by Evin et al. (2013b), who showed that using AR(1) models for hydrological inference on the raw residuals can lead to strong interactions with the inferred parameters and degraded outcomes.

On the other hand, the use of a Bias description as in Del Giudice et al. (2013) becomes prohibitive for long time series. This implementation requires the inversion of a covariance matrix $\Sigma \in \mathbb{R}^{nxn}$, being n the size of the measurement layout. In this case, hourly sampling during 15-Jan-2012 until 05-Aug-2012 has a size of 4892. Expected values of the decay parameter τ are likely to produce a highly sparse covariance matrix, thus sparse inversion optimisation could be applied (Betancourt and Alvarado, 1986) yet intensive sampling for populating the posterior is still cumbersome. This large covariance matrix inversion renders the evaluation of the likelihood function computationally expensive, thus eliminating the benefits of the use of the dynamic emulator for the simulator response variables. A possible solution is to create a database of likelihood samples from the dynamic emulator which is used to build a second emulator linking the parameter space and the response of the likelihood function (Dietzel and Reichert, 2014). An illustrative example of a Bias description term is shown at Figure 3.14 in which the inference was performed in a shorter time-series (26-Jul-2012 – 14-Sep-2012).

The description of the autocorrelation structure in the residuals did not allow for a better description of the process, or a better understanding of the parametric uncertainty. Ammann et al. (2018) recently studied the representation of autocorrelated likelihood structures with the conventional error models as shown here. They discussed that the use of stationary autocorrelation models deteriorates the performance of the inferred model significantly (degrading even further when increasing the measurement layout frequency). They propose that the use of non-stationary autocorrelation schemes may overcome this problem, since hydrological models are expected to lose memory un-

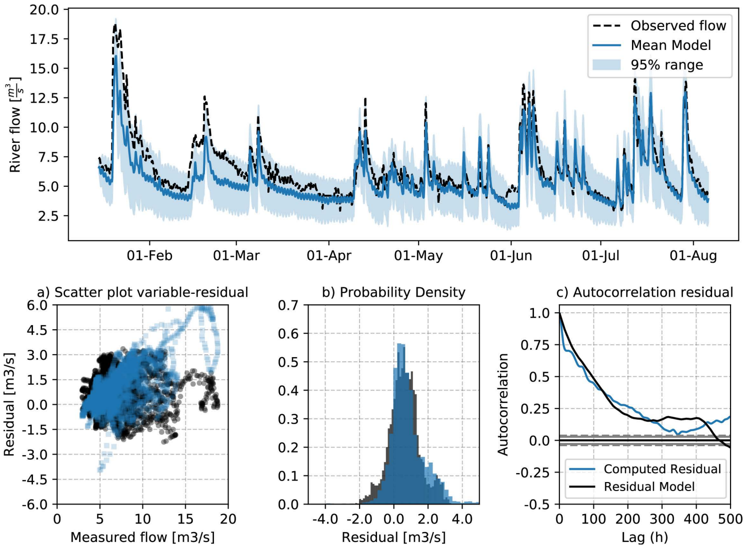


Figure 3.13: Autoregressive model order 3, comparison of measured and inferred dynamics and residual structure.

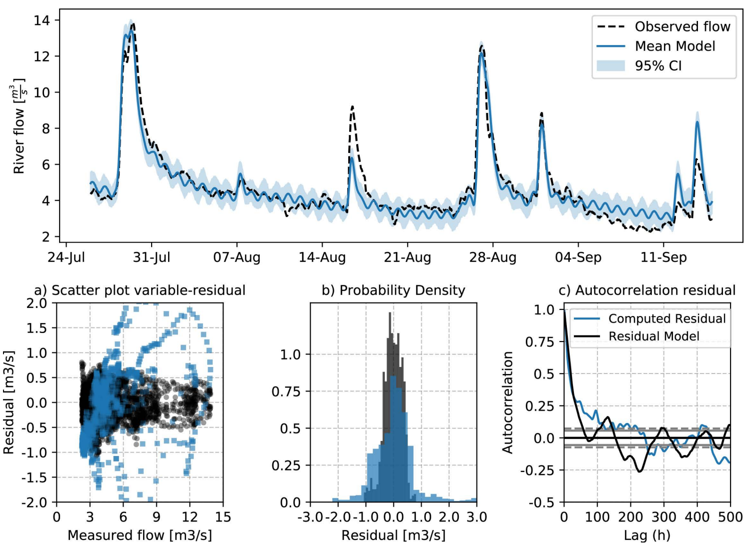


Figure 3.14: Bias description, comparison of measured and inferred dynamics and residual structure.

der storm events (thus dry-weather and wet-weather present different residual correlation patterns). This non-stationarity could not be found in this case, being the correlation structure in dry-weather and wet-weather for short scales (0-80 hours lag) equivalent and the long structures only slightly different (Figure 3.15). Also, this is not expected to be applicable for DO series, where residuals are even less structured as in hydrological flow.

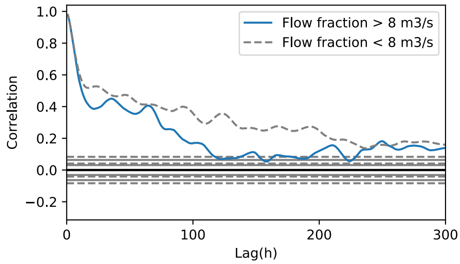


Figure 3.15: Autocorrelation structure for Flow residuals by magnitude.

A strong autocorrelation structure is expected due to the nature of the process and the measurement layout. Both flow and dissolved oxygen concentration present several dynamic modes, induced by storm events, daily fluctuation in the WWTP effluent and variation between dry-wet periods and temperature seasons. This happens well above the measurement frequency (hourly). Small temporal shifts are expected due to model structural misfit (e.g. incorrect CSO timing in the urban drainage scheme or misrepresentation errors in rainfall data). The temporal shift will likely render strongly correlated residuals in time. Yet these time-shifts are of little influence for the model application. The objective of the model is to represent dynamics of oxygen in a receiving water body for environmental policy assessment studies. These studies use metrics which lumps the time-dynamics, as frequency-duration-concentration tables (FWR, 2012), thus the exact timing of the oxygen depletion is not highly relevant, but rather the correct representation of the magnitude and duration of the event. Consequently, the stiff likelihood conditions required to construct formal inference schemes (as shown in this study) might not render the most adequate approach when dealing with the with long-term dissolved oxygen dynamic series (in which the system exhibit multitude of complex dynamic states). Approximated Bayesian computation (Toni et al., 2009) could be of interest by allowing defining metrics which attend to the relevant features of the dissolved oxygen output space (e.g. duration and magnitude of events, slopes of the depletion/recovery patterns, etc.). The selection of signature based metrics are being increasingly used for the diagnosis of hydrological modelling studies (Sadegh and Vrugt, 2014; Kavetski et al., 2018) and could also facilitate the identification and calibration of urban water quality dynamics.

3.3. EMULATING DYNAMIC INPUTS, INCORPORATING THE EFFECT OF TIME-DEPENDENT RAINFALL IN PHYSICALLY BASED 2D FLOW EMULATION *

3.3.1. INTRODUCTION

THE simulation of surface water flow dynamics in urban and rural catchment scale is of great importance in water management. During the past four decades, several distributed or semi-distributed models, having a physical basis, have been developed, such as MIKE-SHE, SWAT or TOPKAPI software (Abbott et al., 1986; Mazzetti, 2015; Neitsch et al., 2002) among others. These conceptualisations simulate several processes (infiltration, interception, snow melting, overland flow, groundwater flow, etc.) included in the water cycle using either differential or empirical equations. Usually, the overland flow propagation is approximated by simplified methods, such as the kinematic or dynamic wave equations and therefore cannot capture the non-linearities observed, especially in extreme events such as floods (Costabile et al., 2013; Liang, 2010; Singh et al., 2014).

In the last few years, there is an increasing trend to simulate overland flow dynamics at catchment scale using the full form of the two-dimensional Shallow Water Equations (2D-SWE). Fiedler and Ramirez (2000) and Esteves et al. (2000) represent early attempts to use numerical solvers based on the full form of 2D-SWE to simulate the rainfall-runoff process in micro-scale catchments. The rapid increase in computational power led to several examples of application to real catchments (Costabile et al., 2012, 2013; Liang, 2010; Nguyen et al., 2016; Singh et al., 2014).

Although the implementation of this kind of simulators is currently feasible in large-scale scenarios, the required computational time is a substantial limitation (in the order of hours-days). Therefore, their use in practice is still hampered (e.g. hydraulic design, real-time flood warning schemes, uncertainty analysis or parametric calibration). Accelerating the sampling of hydrodynamic models is an area of high interest. Several approaches focus on exploiting the benefits of high performance computing (Kalyanapu et al., 2011; Vacondio et al., 2014) or efficient grid definitions (Stelling, 2012). These strategies reduce the time required for computing events of interest, however the computational burden is still prohibitive for intensive model evaluation as required for uncertainty propagation of parametric inference.

The use of surrogate modelling is an alternative to approximate the behaviour of dynamic simulators at output variables of interest. Several authors have presented methods in which a large physically based model is substituted by a fast surrogate conceptual version (Bermúdez et al., 2018; Meert et al., 2018; van Daal-Rombouts, 2017). However, those simplifications do not respect the original model structure, and thus inference or calibration parameters do not have transferability. Another strategy focuses on learning directly the behaviour of the original simulator (building an interpolator between the parametric and output spaces). For instance, Wang et al. (2015) presented a data-driven emulator to facilitate uncertainty analysis in hydrologic modelling. A similar application for groundwater modelling can be found in Laloy et al. (2013). Carbajal et al. (2017)

*This section is an adapted version of: Moreno-Rodenas, A. M., Bellos, V., Langeveld, J. G., and Clemens, F. (2018). A dynamic emulator for physically based flow simulators under varying rainfall and parametric conditions. Water Research, 142, 512-527. doi: 10.1016/j.watres.2018.06.011

described a comparison of surrogate methods (mechanistic vs. data-driven) based on Gaussian processes to emulate the dynamics of urban drainage routing models. More applications can be found in urban hydrology (Machac et al., 2016a,b). These methods have the advantage of representing a mapping between model parameters and the simulator output space, which allow for a direct use in calibration, inference or uncertainty analysis schemes. However, they are generally hard-bounded to the rainfall time-series in which they were trained. This means that the model emulator is only valid for changes in the parametric space under a particular rainfall event and boundary conditions (outside those conditions a new sampling and training would be necessary). This has posed a strong limitation on the application of data-driven emulators in hydrology.

A potential approach to incorporate the rainfall variability in the data-driven emulator is to encode the rainfall process as a set of discrete parameters (representing rainfall intensity at each time step), which is added to the model parametric space to be emulated (Mahmood et al., 2018). However, common techniques used for model emulation as polynomial chaos expansion (PCE) or Gaussian processes (GP) are sensitive to the dimensionality of the parameter space, allowing only for a reduced number of parameters to be accounted for. This technique is limited to only short rainfall events, or very coarsely discretised, also time autocorrelation structures could pose further challenges. Another surrogate approach could be the use of recurrent neural networks (RNN/LSTM, e.g. long-short term memory configurations (Hochreiter and Schmidhuber, 1997)). These structures are designed to accommodate time-dependent processes. However, RNN implementations are expected to require a prohibitive number of simulator samples to guarantee that variations of time-dependent rainfall are sufficiently explored. Sajikumar and Thandeveswara (1999) provided an early example describing a rainfall-runoff process using a neural network by learning patterns from monitoring data. Although this could replicate the behaviour of a given status of the system, it did not allow for representing parameter variations or virtual changes in the system.

A way to cope with the approximation of physically-based flow simulators responses to new rainfall series is the hybrid combination of the hydrological Unit Hydrograph (UH) theory with hydrodynamic modelling (Bellos and Tsakiris, 2016). With this technique, an UH is derived using a physically based 2D hydrodynamic simulator. Then, the rainfall event is simulated abstracting from the precipitation losses due to infiltration, sewer drainage, interception etc., and composing the total flow response using proportionality and superposition of the UH. With this scheme a flood event can be readily evaluated. Nevertheless, the derivation of the UH is parameter dependent, and thus practical model use is still severely hampered (e.g. quantification of uncertainty, sensitivity analysis, calibration, etc.). Bellos et al. (2017) attempted to solve this issue by generalising the UH derivation with the use of a surrogate model for a uniform distribution of Manning coefficient values, thus including the parameter dependency. However, the solution of the 2D-SWE is often highly non-linear, which is strongly manifested when dealing with small catchments (characterised by fast response). This violates the assumption of linearity in the UH theory, thus rendering significant errors in the estimated hydrograph. In this work, we present a new methodology to emulate physically based hydrodynamic simulators. This is based on encoding the rainfall process as unitary responses through a dedicated sampling scheme and the use of a polynomial chaos expansion surrogate model which compensates for the effect of non-linearities due to the superposition and proportionality composition of the UH. The mismatch observed in the results derived by the hybrid technique and the full implementation of the physically based model in previous attempts Bellos and Tsakiris (2016); Bellos et al. (2017) is significantly reduced. This new surrogate technique allows the modeller to capture the behaviour of the real model to variations of time-dependent rainfall (spatially uniform) and parameter scenarios at a significant computational time reduction (in the order of milliseconds-seconds). In order to demonstrate the presented methodology, we compare the results derived by a simulator, the conventional use of UH linear theory and the new proposed surrogate model, in three case studies: a synthetic catchment which is represented by a simplified non-linear tank-in-series flow routing structure and two catchment geometries where the rainfall-runoff process is described by the full form of 2D-SWE.

3.3.2. METHODS AND MATERIALS

3.3.2.1. MODEL BASED UNIT HYDROGRAPH

The Unit Hydrograph (UH) theory (Dooge, 1959) is based on the assumption that there is a unique runoff response after a unitary rainfall depth, during a specific time interval, for every catchment. A commonly accepted unitary depth is $10mm$ $(l \cdot m^{-2})$ of rainfall, whereas the time interval varies and depends on the characteristics of the catchment, such as the size, time of concentration etc. With a given UH of a catchment, every runoff can be derived using as input data the rainfall event, based on the principles of the proportionality and superposition. The runoff response for rainfall depth in every time step is derived assuming that this rainfall depth is proportional to the corresponding unitary depth and therefore the responses should be proportional as well (principle of proportionality). The individual runoff responses are composed in time to represent the response for the given rainfall event (principle of the superposition). The UH theory has been often used to approximate the behaviour of catchments where limited data is available (ungauged basins). This is done using synthetic UHs, which are derived based on the catchment's characteristics (e.g. Snyder's US or the Soil Conservation Service UH, etc.). Alternatively, UHs of several durations can be derived from monitoring data when available. In this particular case, the objective is to represent the behaviour of a hydrodynamic simulator. Thus, a set of tailored simulations can be drawn from the model in order to capture its unitary reaction by using the principles of UH theory. From this set of samples, an emulator structure was applied to generate an interpolation map between new parameter and rainfall scenarios and the approximated response of the hydrodynamic simulator.

3.3.2.2. POLYNOMIAL CHAOS EXPANSION (PCE)

A weighted sum of orthogonal polynomials can be used to link a vector of random variables to space-time model outputs (Xiu, 2010; Xiu and Karniadakis, 2002) if the mapping is smooth. In this case, we denote a deterministic flow model M as:

$$Q(t) = M[x_0, I(t), \theta] \tag{3.10}$$

which has a set of initial conditions x_0, g time dependent inputs of length m, $I(t) \in$

$\mathbb{R}^{mxg}$ and a vector of model parameters $\theta \in \mathbb{R}^d$ and which solution renders the simulated flow dynamics at a certain location, $Q(t) \in \mathbb{R}^m$. Then, we aim to find a linear combination of polynomials (only dependent on the parametric space) which approximates a map between θ and $Q(t,\theta)$ as:

$$Q(t,\theta) \approx c(t)^T \cdot \Phi(\theta) \tag{3.11}$$

in which $\Phi(\theta) \in \mathbb{R}^{Nx1}$ is a series of N orthogonal polynomials constructed as basis of the joint probability density function of the parameter space θ and $c(t) \in \mathbb{R}^{Nx1}$ is a vector of coefficients $c_i(t) \in \mathbb{R}^m$ to be fitted to the particular model response.

The output space $Q(t,\theta)$ is interpolated through a weighted sum of polynomials. The fitting process is performed by deriving the model output at $i = 1,\ldots,K$ known parameter combinations ($\theta = q_i$) as $M[x_0, I(t)|\theta = q_i] = Q(t|q_i)$. The model samples and the polynomial approximation can be written as a system of equations:

$$\begin{bmatrix} Q(t|q_1) \\ \vdots \\ Q(t|q_K) \end{bmatrix} = \begin{bmatrix} \Phi_1(q_1) & \ldots & \Phi_N(q_1) \\ \vdots & \ddots & \vdots \\ \Phi_1(q_K) & \ldots & \Phi_N(q_K) \end{bmatrix} \cdot \begin{bmatrix} c_1(t) \\ \vdots \\ c_N(t) \end{bmatrix} \tag{3.12}$$

in which the only unknown factors are the coefficients $c_i(t)$. This system can be expressed in matrix form as:

$$\mathbf{Q}(t|q) = \mathbf{P}(q) \cdot \mathbf{c}(t) \tag{3.13}$$

for which $\mathbf{P} \in \mathbb{R}^{KxN}$ is the matrix of polynomials values corresponding to each parameter sample. By ensuring that the system is overdetermined ($K > N$), this system can be approximated by a least squares minimisation:

$$\mathbf{c}(t) = (\mathbf{P}(q)^T \cdot \mathbf{P}(q))^{-1} \cdot \mathbf{P}(q)^T \cdot \mathbf{Q}(t|q) \tag{3.14}$$

thus the values of $\mathbf{c}(t)$ are approximated so they map the selected polynomial basis $\Phi(\theta)$ and the sampled output series $\mathbf{Q}(t|q)$. Then the expression of Equation 3.11 can be used to interpolate the targeted output series to different values of the parameter set. This is known as a non-intrusive collocation method, other fitting methods and sampling strategies are discussed in Hadigol and Doostan (2018).

The orthogonal polynomials should be chosen as basis of the parameter probabilistic space. Xiu and Karniadakis (2002) provided a range of known polynomial basis and stochastic variable distributions. In this case, training parameter distributions were always considered independent and uniformly distributed. This has associated the Legendre polynomials. The creation of orthogonal polynomial series was done using the implementation of Gautschi (1994), which uses a three-term recursion relation for univariate polynomials:

$$\begin{gathered} \Phi_{n+1}(\theta_i) = \Phi_n(\theta_i)((\theta_i) - A_n) - \Phi_{n-1}(\theta_i) B_n \\ A_n = \frac{E[\theta_i \Phi_n^2]}{E[\Phi_n^2]}, B_n = \frac{E[\Phi_n^2]}{E[\Phi_{n-1}^2]}, \Phi_{-1} = 0, \Phi_0 = 1 \end{gathered} \tag{3.15}$$

where $\Phi_n(\theta_i)$ represents the n^{th} polynomial of the i^{th} parameter. As the parametric space is considered stochastically independent, the multivariate orthogonal expansion was obtained by multiplying the univariate ones:

$$\Phi_n(\theta) = \Phi_n(\theta_1)\ldots\Phi_n(\theta_d) \tag{3.16}$$

In practice, the polynomial series is truncated at a certain order p, which is selected such that a desired level of accuracy is achieved. Then, the number of polynomials (and thus of associated $c_i(t)$ to be approximated) is:

$$N = \frac{(p+d)!}{p!d!} \tag{3.17}$$

which is related to the polynomial truncation order p and to the dimensionality of the parameter vector d.

3.3.2.3. EMULATOR STRUCTURE

A polynomial chaos emulator can be used to represent an interpolation map between model parameters and the model unitary hydrograph response (Bellos et al., 2017). This is done by propagating a unitary rainfall event of $10mm$ at combinations of model parameter $\theta = q_i$ and fitting a PCE (as described previously) to the output unitary response, $UH_{10mm}(t,\theta)$:

$$UH_{10mm}(t,\theta) \approx c_{10mm}(t)^T \cdot \Phi_{10mm}(\theta) \tag{3.18}$$

The total hydrograph shape is then reconstructed for any given rainfall event and combination of parameters. This relies on the proportionality and superposition assumptions of the UH theory, which has proven insufficient to capture the real model dynamics in previous attempts (Bellos et al., 2017). The observed error in the comparison between the hydrodynamic simulator and the composition of the UH was separated in two sources: a) error due to the proportionality assumption; b) error due to the superposition assumption. According to the proportionality principle, the UH relates linearly to rainfall intensity. This assumption is not valid when dealing with a non-linear model structure. In order to correct for this error source, the behaviour of the unitary hydrograph to rainfall intensity was transferred to the PCE emulator. This was achieved by training the emulator through variations of the model parametric space and of the unitary rainfall intensity (R) (instead of using a fixed $10mm$ event to derivate the UH):

$$UH_P(t,\theta,R) \approx c_P(t)^T \cdot \Phi_P(\theta,R) \tag{3.19}$$

With this approach, the emulator UH_P included the proportional non-linear effect. This corrected for the proportionality error of the classical UH theory, which was tested by reproducing the behaviour of a non-linear tank-in-series model to variations of unitary rainfall intensity.

On the other hand, the superposition error is linked to the model dynamic state. According to UH theory, the flow response is unitary and independent in previous states and therefore fails to capture non-linear processes. In order to transfer the information of the system state to the emulator structure, we propose a new method to sample the real simulator and to build a unitary rainfall emulator (UH_{PS}). Figure 3.16 depicts the emulator development process. This structure was based on two simplification phases. First, a PCE emulator scheme was built following these steps: 1) a series of samples were drawn from the model parametric space $\theta = q_i$ and from two extra parameters R and R_P (which represented two consecutive unitary rainfall steps with duration Dt, selected following the UH theory). 2) Two samples were drawn from the original simulator using the same parameter sample (q_i) but using a rainfall time-series input of $\mathbf{R_T}(t) = [R_p, R, 0, 0, \ldots]_{Dt}$ and $\mathbf{R_p}(t) = [R_p, 0, 0, 0, \ldots]_{Dt}$ respectively (which are vectors of input rainfall intensity at a given Dt time step). The difference between the two output hydrographs was computed as $M(\theta = q_i, \mathbf{R_T}(t)) - M(\theta = q_i, \mathbf{R_P}(t))$, which provided an approximation of the state response of the model equations when a rainfall intensity R_p preceded a rainfall intensity R. This response contains the proportionality non-linear error information along with the superposition error (at one rainfall lag). The set of unitary responses was stored in a database. 3) The model unitary responses from the database were used to train a PCE emulator as:

$$UH_{PS}(t, \theta, R, R_p) \approx c_{PS}(t)^T \cdot \Phi_{PS}(\theta, R, R_p) \tag{3.20}$$

4) The performance of the unitary emulator was tested on several samples, which were not used to train the model. Once fitted the interpolator (PCE), a second simplification phase is performed; 5) the unitary responses from UHPS are used to reconstruct new rainfall-parametric scenarios; 6) Effective rainfall is computed (if required) and 7) superposing all unitary model responses obtained from Equation 3.20, which corresponds to the unitary response of the model under the parameter set θ and the given rainfall event. Additionally, initial boundary conditions can also be included in the emulation phase by conceptualising them as extra model parameters.

3.3.2.4. HYDROLOGICAL (CONCEPTUAL) FLOW MODEL

In order to illustrate the impact of non-linearities in the basic assumptions of the unit hydrograph theory (superposition and proportionality) a basic synthetic model was used (Figure 3.17). In this model the strength of the non-linear component can be varied allowing to test multiple conditions (at a reduced computational burden). This simple rainfall-runoff model structure was conceptualised as a chain of non-linear tanks:

$$\begin{gathered} \frac{dV_i}{dt} = Q_{in} - Q_{out} \\ Q_{out} = a \cdot V_i^b \end{gathered} \tag{3.21}$$

where V_i represents the volume stored at the i^{th} tank, Q_{in} the flow entering in the tank from the previous structure and Q_{out} represents the outflow, which follows a non-linear relationship with the current tank storage (Equation 3.21). The discharge is driven by a proportional coefficient a and an exponential parameter b. The runoff was calculated as $Q_{in\_tanks} = A \cdot R$ from a certain area (A), and was separated in two lines of five and three tanks, which converged at the outflow (Figure 3.17).

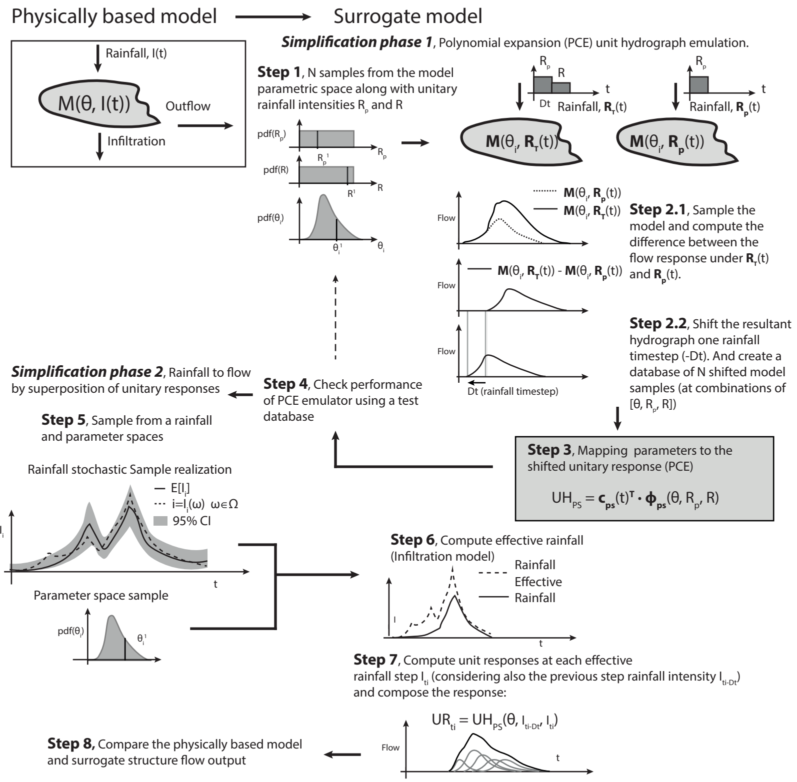


Figure 3.16: Emulator conceptual scheme for the correction of superposition and proportional non-linear errors (UH_{PS}).

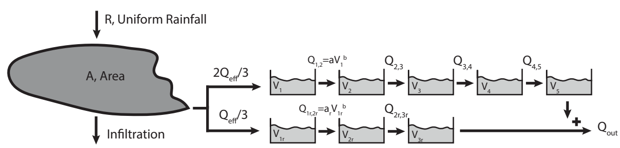


Figure 3.17: Scheme for a simplified/lumped non-linear flow model.

Both tank lines shared the exponential parameter b but had different proportional coefficients, namely a and a_r. Infiltration was set to 0. Thus the model was composed by a set of Ordinary Differential Equations (ODEs):

$$Q_{outflow} = M_{simplified}(a, a_r, b, A, R(t)) \tag{3.22}$$

represented by $M_{simplified}$ which depends on the set of tank-parameters, the catchment's area A and in a rainfall input $R(t)$. The parameters were manually selected ($a = 1.4, a_r = 2.2, b = 1.2 and A = 5$) in order to force a non-linear dynamic behaviour. Rainfall time-series were supplied with a frequency of 10 minutes. The effect of non-linearities and its correction through the use of the proposed emulator were computed on the simplified model at the fixed set of parameters.

Additionally an example is provided in which the full parameter space was used as an input for the emulator. The surrogate structure was trained using 3000 samples drawn from a Latin hypercube sampling (LHS) scheme of the full parametric space (a, a_r, b, A, R, R_p), where the first four elements referred to model parameters, and the last two encoded the rainfall process. Distribution and ranges used at the training dataset are described in Table 3.5.

Table 3.5: Parameter probability distributions used for training the simplified model emulator.

Parameter	a (-)	a_r (-)	b (-)	A (km^2)	R (mm)	R_p (mm)
Pdf	Uniform	Uniform	Uniform	Uniform	Uniform	Uniform
Range	[1, 1.6]	[1.8, 2.5]	[1.1, 1.2]	[4.5, 5.5]	[0.2, 20]	[0, 20]

An uncertainty analysis scheme was also implemented using the emulated structure from an elicited parameter probability distribution (Table 3.6). A total of 1000 samples were drawn from a Monte-Carlo sampling scheme assuming independency of parameters. This contained the four model parameters (a, a_r, A, b) and a rainfall multiplier $(K_{rainfall})$, which modified the rainfall input.

Table 3.6: Parameter marginal distributions (simplified model uncertainty propagation).

Parameter	a (-)	a_r (-)	A(km^2)	b (-)	$K_{rainfall}$ (-)
Pdf	truncNorm	truncNorm	truncNorm	Uniform	Uniform
Parameters	$\mu = 1.4$ $\sigma = 0.014$	$\mu = 2.2$ $\sigma = 0.088$	$\mu = 5$ $\sigma = 0.2$	-	-
Range	[1.2, 1.5]	[1.8, 2.5]	[4.5, 5.5]	[1.18, 1.95]	[0.95, 1.05]

3.3.2.5. PHYSICALLY BASED FLOW MODEL (2D SWE)

The applicability of the proposed emulation technique in physically based real scale models was tested by representing the dynamic response of the full solution of the 2D-SWE in two synthetic studies. A computational domain was created using a parabolic shape to represent a surface elevation model (Figure 3.18). The surface elevation model $z(x, y)$ (in meters) adopted the following form:

$$z(x, y) = (\alpha - \alpha_0 \cdot x) \cdot (y - y_0)^2 - \beta \cdot x + c \tag{3.23}$$

where $\alpha = 10^{-4}$, $\alpha_0 = 1.5x10^{-7}$, $y_0 = 250$, $\beta = 5x10^{-3}$ and $c = 10$. The computational domain was discretised in square-shaped cells of 10x10 meters, whereas a gate of 100 meters width was located at the outlet of the catchment where runoff was measured. The system had a surface of 0.26 km^2, which is representative of a small catchment characterised by fast response (Model SWE_parabola). Additionally, a set of five obstacles (square reflective boundaries) was located near the centre of the computational domain (Model SWE_urban).

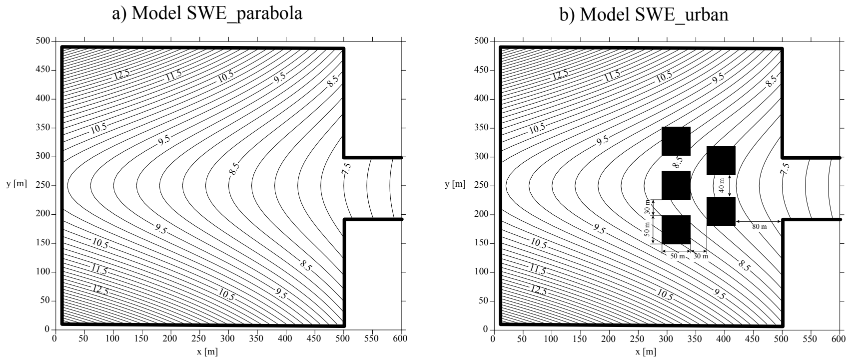


Figure 3.18: Surface elevation and boundary conditions of the two 2D-SWE simulators.

In its conservative form, the two dimensional shallow water equations can be written as:

$$
\begin{aligned}
\frac{\partial h}{\partial t} + \frac{\partial}{\partial x}(uh) + \frac{\partial}{\partial y}(vh) &= 0 \\
\frac{\partial (uh)}{\partial t} + \frac{\partial}{\partial x}(u^2 h + 0.5 \cdot gh^2) + \frac{\partial}{\partial y}(uvh) &= ghS_0^x - gR_H S_f^x \\
\frac{\partial (vh)}{\partial t} + \frac{\partial}{\partial x}(uvh) + \frac{\partial}{\partial y}(v^2 h + 0.5 \cdot gh^2) &= ghS_0^y - gR_H S_f^y
\end{aligned}
\tag{3.24}
$$

where the energy slopes are determined as:

$$
\begin{aligned}
S_f^x &= n^2 u(\sqrt{u^2 + v^2}) R_H^{-4/3} \\
S_f^x &= n^2 v(\sqrt{u^2 + v^2}) R_H^{-4/3}
\end{aligned}
\tag{3.25}
$$

for which h represents the free-surface depth, u and v the velocity components in the x and y axis respectively, R_H the hydraulic radius (which is approximated by h for unconstrained locations). The Manning's roughness coefficient is represented by n, S_0^x and S_0^y denote the bottom surface slope in each direction.

The numerical solver used for the flow dynamics simulation was the FLOW-R2D model, which solves the full form of 2D-SWE. The implementation consisted on the Finite Difference Method and a modified version of McCormack numerical scheme (Bellos and Tsakiris, 2014). The time step was $\Delta t = 0.01s$, whereas for the diffusion factor a typical value of $\omega = 0.999$ was selected. The Manning equation was used to represent friction. The threshold which distinguishes wet and dry cells was selected equal to $h_{dry} = 10^{-4}m$. For numerical reasons, an initial thin film of water equal to $1.1x10^{-5}m$ was set up in the catchment. This film was infiltrated in the first 15 min through the sink term of the mass equation of the 2D-SWE in a constant rate, in order to preserve mass balance between rainfall and runoff volumes. The boundaries of the catchment were set as walls through the reflection boundary technique in order to preserve mass balance as well, whereas the outlet of the catchment was set as open boundaries (Tsakiris and Bellos, 2014). A rainfall input series was used with a time step of 15 minutes.

The emulator scheme was implemented using variations of the surface Manning's roughness (n) as a model parameter (constant in the spatial domain). Table 3.7 presents the parametric distributions at training.

Table 3.7: Parameter probability distributions used for training the 2D-SWE model emulator.

Parameter	n ($s \cdot m^{-1/3}$)	R (mm)	R_p (mm)
Pdf	Uniform	Uniform	Uniform
Range	[0.03, 0.05]	[0.01, 20]	[0, 20]

A total of 300 samples were performed on the two simulators (drawn from a LHS). The 2D-SWE simulator generated numerical instabilities at very low rainfall intensities. Thus the sampling scheme was reduced to $R_p > 0.2$ mm and $R > 0.2$ mm. From the remaining dataset, 264 samples were used for training and 30 for validation of the PCE accuracy. A 5^{th} order Legendre polynomial set was used in both cases. Additionally, a validation set of rainfall and parametric scenarios (Table D.1, Appendix D) was used to compare the fit between the original simulators and the proposed surrogate models.

3.3.2.6. PERFORMANCE INDICATORS

The fit between model and emulator derived flow time series was assessed by the Nash-Sutcliffe Efficiency (NSE, Nash and Sutcliffe (1970)):

$$NSE = 1 - \frac{\sum_{t=0}^{T}(E_t - S_t)^2}{\sum_{t=0}^{T}(S_t - \hat{S})^2} \tag{3.26}$$

where E_t denoted the emulated and S_t the simulated values at time t, and $\hat{S}$ the mean value of the simulated series. This metric is normalised, thus allowing to compare between storm events of varying dynamics.

Besides, a Peak relative error (PRE) index was used in order to assess the accuracy of the flow maximum peak emulation, which follows:

$$PRE = \frac{S_{peak} - E_{peak}}{S_{peak}} \tag{3.27}$$

where S_{peak} and E_{peak} are the simulated and emulated values associated with the peak flow of the hydrograph.

3.3.3. RESULTS AND DISCUSSION

3.3.3.1. CORRECTING UNIT HYDROGRAPH ERRORS DUE TO NON-LINEARITIES

The simplified tank-in-series model was used to test the effect of superposition and proportionality assumptions in the reconstruction of hydrographs based in the UH linear theory. The model parameters were defined as described in the materials section to force a non-linear routing process (b = 1.2). Figure 3.19 shows the mismatch between a UH linear derived hydrograph reconstruction (from a 10 mm/10 minutes rainfall) and the original simulator. In the x-axis, a set of unitary (10 minutes) rainfall events is supplied with a ramping intensity (between 1-50 mm). The color-map represents hydrograph response. The normalised residuals at each time-step are shown below, which were computed as:

$$\tilde{\epsilon}_t = \frac{S_t - E_t}{S_t} \tag{3.28}$$

where S_t is the flow simulated at time t and E_t the corresponding emulated values. In the right (Figure 3.19) the simulated, approximated (by the UH theory) and residual values are shown for a 30 mm unitary rainfall. The residual structure indicates that for rainfall intensities above or below 10 mm the performance of the UH reconstruction rapidly decreases. This is due to the non-linear behaviour of the flow propagation, which explains why higher intensities are heavily underestimated and lower intensities overestimated through the linear representation of the UH.

Figure 3.20 shows the effect of proportionality when using the UH_P emulator. In this case the rainfall intensity is included in the surrogate structure, and thus the unitary response derivation represents the effect of non-linearities due to proportionality. The residual structure indicates a high similarity between the emulator-simulator outputs, when using only unitary rainfall at varying intensities.

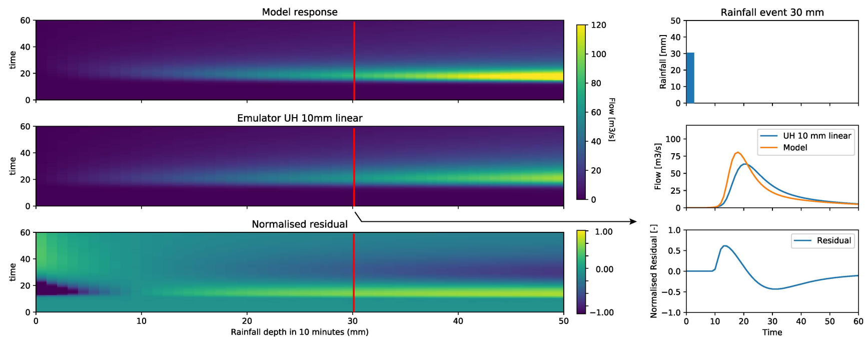


Figure 3.19: Effect of non-linearities in the proportionality assumption of the unit hydrograph theory at the simplified model.

Figure 3.21 depicts the errors due to the superposition of the unit hydrograph. In the x-axis a set of 100 rainfall events with varying duration and intensity were used as input for the simulator and for the UH_P emulator. The rainfall events had a random duration of between 10-50 minutes and a rainfall intensity ranging from 2-18 mm/10min. The x-axis was sorted by rainfall volume. In the right, the effect of a constant rainfall intensity of 12 mm/10min during 30 minutes is shown. The effect of non-linearities is still present in this case, although the proportional factor was corrected as shown in Figure 3.20. This indicates that the system state influences the response, and thus, should be accounted for in non-linear systems. Figure 3.22 shows the correction done by the use of the proposed emulator structure (UH_{PS}). This structure partially compensates for the proportional and the superposition effects. The residual plot in Figure 3.22 shows that these errors are significantly reduced in comparison with the previous scheme (UH_P).

The effect of the linear assumptions of the UH theory in hydrograph approximation

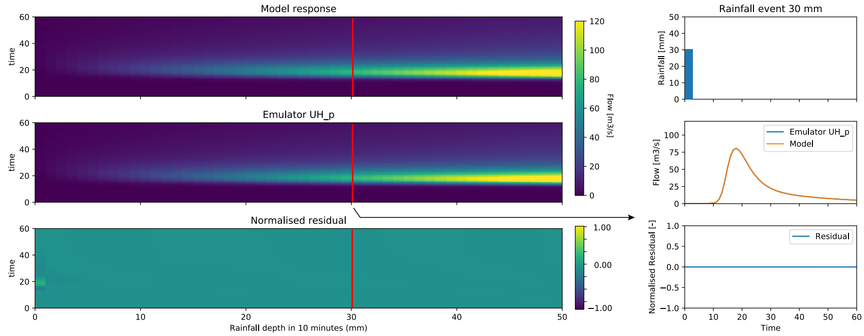


Figure 3.20: Correction of the proportionality error at the simplified model (UH_P).

have been reported in the literature for rural and urban catchments (Ding, 2011; Minshall, 1960; Musy, 1998). The example presented shows the separated effect of the superposition and proportionality simplifications of a unit hydrograph and a strategy to correct them in the emulator structure. With the use of the tank-in-series model it was possible to modify the level of non-linearity by changing the tank exponent (b). The reader is directed to the additional material (Figure C.1 and Figure C.2 in the Appendix C) where the same effect was computed for the case $b = 1$, which shows the applicability of the UH theory in linear cases. Nevertheless, flow propagation processes are seldom linear, and thus this effect is expected to be relevant in most applications.

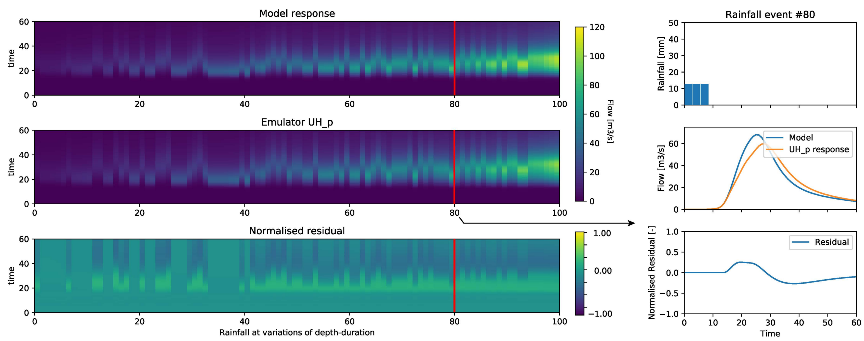


Figure 3.21: Superposition error after correction of the proportional error (UH_P).

3.3.3.2. EMULATION OF THE SIMPLIFIED FLOW MODEL

Once trained, the behaviour of the UH_{PS} emulator was compared with the simplified model response for combinations of rainfall and parametric scenarios. To that effect, 1000 samples were drawn from a Monte-Carlo sampling scheme of an elicited parameter distribution (Table 3.6). This contained the four model parameters (a, a_r, A, b) and a rainfall multiplier ($K_{rainfall}$). Figure 3.23 displays the distribution of the performance indicators PRE and NSE between UH_{PS} and the original simulator. NSE is close to one indicating an accurate fit between simulated-emulated time-series. PRE reports the relative error in the peak reproduction, which is around 1.2% of the maximum flow.

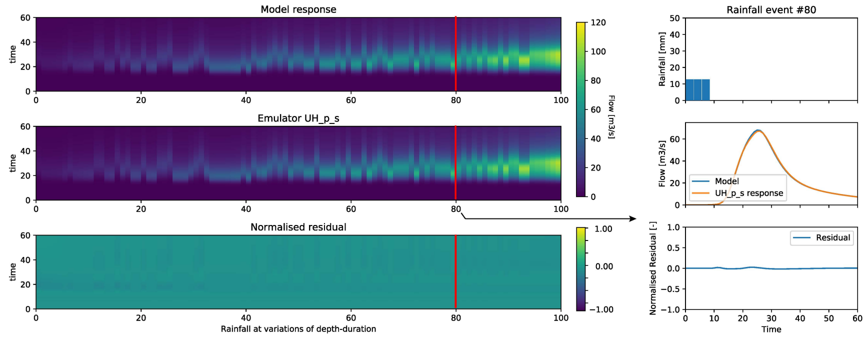


Figure 3.22: Superposition error after correction of the proportional and superposition error (UH_PS).

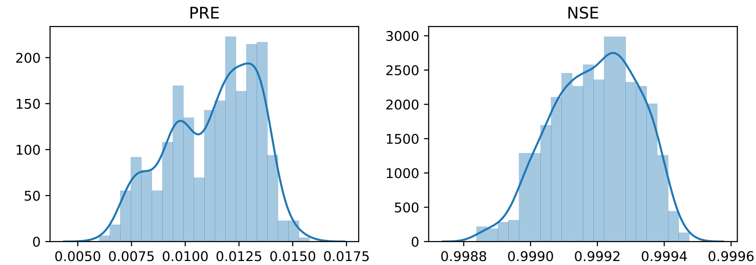


Figure 3.23: Performance indicators for the comparison simulator-UH_{PS} (1000 samples). Peak relative error (PRE) and Nash-Sutcliffe Efficiency (NSE) probability distributions

Figure 3.24 shows the graphical comparison of the UH_{PS} emulator and the simulator for the initial parameter sample at a particular rainfall series, along with the mean and 95% interval of the 1000 parameter samples performed on the emulator.

3.3.3.3. EMULATION OF THE 2-D SHALLOW WATER EQUATIONS

The dynamics of the 2D-SWE were emulated following the same process as in the simplified example. Figure 3.25 shows the NSE values for the emulated-simulated unitary responses at 30 combinations of parameters contained in the testing dataset. This served to validate the emulator of the unitary responses.

All methods were compared; the emulator UH_{PS} (which takes into account the error due to proportionality and superposition), the emulator UH_P (only the proportionality error), the conventional composition of the UH linear (using 10 mm unitary rainfall) and the results derived from the 2D-SWE hydrodynamic simulators (SWE_parabola and SWE_urban) in a set of 8 validation scenarios (Appendix D, Table D.1), for varying rainfall time series and roughness coefficient values. Table 3.8 presents the NSE and PRE values for all cases comparing with the hydrodynamic simulator. There is a clear increase in performance by the use of the proposed emulator structure (UH_{PS}) when approximating the flow dynamics. This increase is denoted by the high NSE values and a PRE one order of magnitude lower than the other emulator structures. In general, the use of the proportional correction (UH_P) is not sufficient, and although it reproduces the hydro-

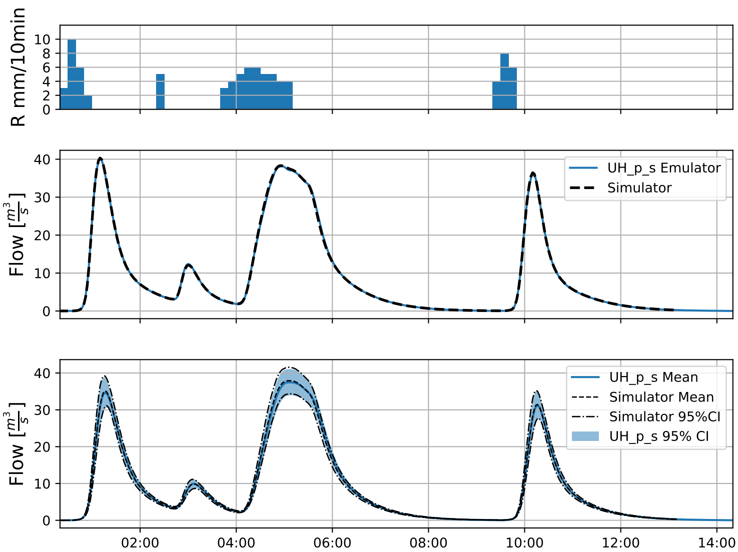


Figure 3.24: Model vs. emulator comparison; Mean rainfall series (above), emulator and simulator response at parameter values ($a = 1.4, a_s = 2.2, b = 1.2$ and $A = 5$) (middle) and comparison of forward uncertainty propagation from 1000 samples from the emulator and simulator (below).

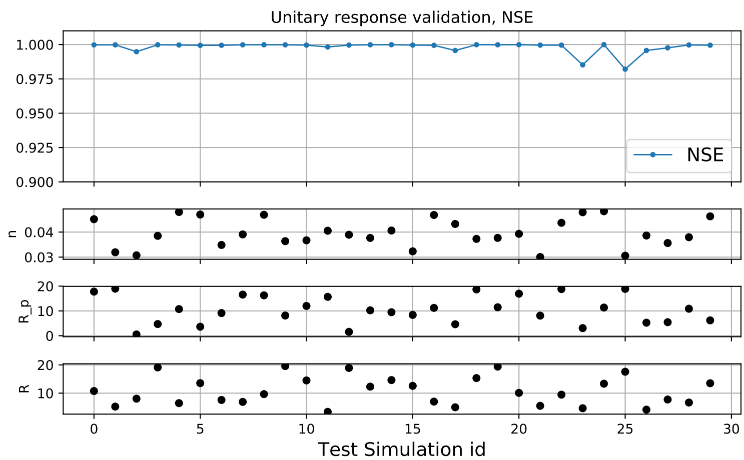


Figure 3.25: NSE and parameter values for the UH_{PS} emulator unitary responses at the test database (SWE_parabola).

graph peak better than the simple use of the linear UH structure, denoted by the lower PRE values (specially in rainfall events where the maximum intensity differs from 10 mm as in Validation events 0 to 3), it is heavily affected by the superposition error as seen in the Validation events 4 to 7.

The use of UH_{PS} improves the reconstruction of the simulated hydrograph and it

Table 3.8: Nash-Sutcliffe Efficiency and Peak relative error between simulated vs emulated (in the left, SWE_parabola, and in the right, shaded, SWE_urban).

	NSE						PRE					
	UH_{linear}		UH_P		UH_{PS}		UH_{linear}		UH_P		UH_{PS}	
	a	b	a	b	a	b	a	b	a	b	a	b
Validation_0 (n=0.035)	0.64	0.53	0.74	0.65	0.988	0.987	0.46	0.51	0.18	0.23	-0.04	-0.03
Validation_1 (n=0.042)	0.53	0.36	0.65	0.54	0.989	0.983	0.48	0.53	0.24	0.27	-0.043	-0.01
Validation_2 (n=0.037)	0.46	0.48	0.59	0.61	0.985	0.988	0.51	0.52	0.28	0.24	0.055	-0.02
Validation_3 (n=0.043)	0.39	0.34	0.52	0.53	0.987	0.981	0.52	0.54	0.3	0.28	0.032	0.007
Validation_4 (n=0.036)	0.62	0.6	0.48	0.37	0.996	0.988	0.42	0.44	0.43	0.49	0.026	-0.049
Validation_5 (n=0.040)	0.6	0.55	0.41	0.25	0.996	0.986	0.42	0.45	0.45	0.51	0.012	-0.03
Validation_6 (n=0.038)	0.57	0.52	0.49	0.38	0.996	0.996	0.46	0.45	0.45	0.45	0.02	-0.04
Validation_7 (n=0.041)	0.53	0.45	0.43	0.28	0.997	0.996	0.48	0.47	0.47	0.48	0.016	-0.03

captures the timing of the flow peak in all tested cases for both geometrical scenarios. Figure 3.26 shows a graphical comparison between the emulated and simulated values for the Validation_0 and Validation_1 scenarios for the simulator SWE_urban. UH_{PS} has a close fit to the simulator dynamics, with an error in the peak estimation of 1-5%. The comparison of SWE_parabola and SWE_urban are also shown in Figure 40 were the effect of geometrical obstacles of the second scheme generates a slightly delayed and attenuated flow response, which led to a decrease in performance of UH_{LINEAR} and UH_P. The reader is directed to the additional material for the graphical comparison at the 8 validation scenarios for the two simulators (Figure D.1 and Figure D.2, Appendix D).

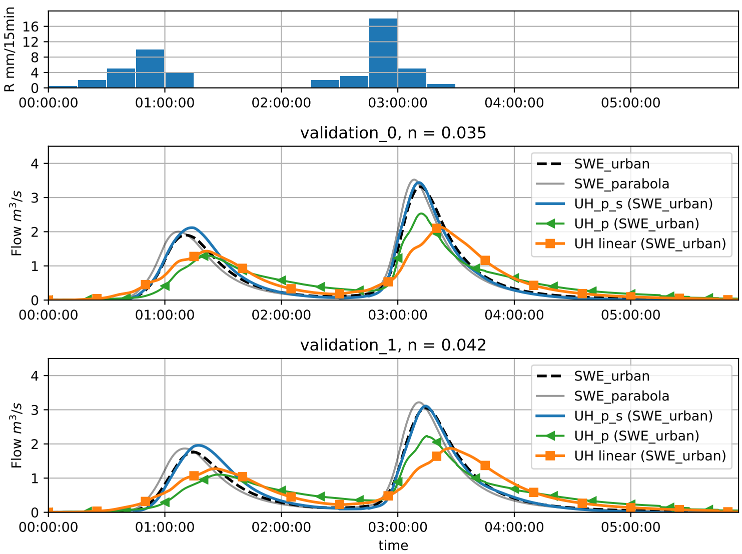


Figure 3.26: Validation of the SWE_urban emulator. Graphical comparison of the response of the real model under a rainfall event and two Manning values (0.035 and 0.042 respectively) against; use of the classical unit hydrograph theory from a 10 mm rainfall (UH_{linear}), the correction of the proportionality error only (UH_P) and the proposed emulator structure (UH_{PS}).

One of the advantages of the proposed methodology (UH_{PS} structure) is that the emulation errors can be decomposed directly in two sources: a) errors in the emulation of the unitary responses through a PCE and b) errors in the reconstruction of the hydrograph by composition of the unitary responses. The errors in phase a) can be described by comparing the PCE emulator and the simulator in a set of test parameter scenarios. This source is dependent on the capability of the emulation technique used to represent the link between the unitary response and the parameter space, which can be reduced in two ways: 1) by carefully selecting the emulator structure, which should be capable of representing a dynamic output, such as a PCE (Xiu and Karniadakis, 2002), by using a Gaussian Process as in Carbajal et al. (2017), Bayesian networks (Conti and O'Hagan, 2010) or recurrent neural networks (Gers et al., 2002). 2) By increasing the training dataset size (sampling from the simulator). Phase b) involves reconstructing the hydrograph from the unitary responses. The effect of the proportionality simplification is well represented by including the rainfall intensity in the emulator structure. However, the superposition composition has a dependency on the system dynamics and the unitary rainfall length. For highly dynamic systems, high-frequency rainfall data could be necessary, and thus, accounting for only one temporal step in the superposition scheme might not be sufficient.

A limitation of the proposed methodology is that its applicability is restricted to rainfall spatial homogeneity. The generalisation to spatio-temporal rainfall variability poses a challenge since it will increase significantly the degrees of freedom of the unitary response, which should be captured by the emulator phase, thus requiring a much larger number of samples from the real simulator.

Also, in practice the number of accounted parameters is limited in this approach. When using the PCE emulator, the number of polynomials follows Equation 3.17, which depends on the dimension of the parameter space (d) and the grade of the polynomial series (p). This relationship grows rapidly, forcing to increase the number of samples from the original model. Thus, for large dimensional spaces the number of required model samples could render the training impractical.

Table 3.9: Computational time of training vs. operation.

		Simulator (time)		Emulator (time)
	# training samples	sample	Validation scenario	Validation scenario
SWE_parabola	260	~ 14,400 s	~ 22,000 s	~ 0.040 s
SWE_urban	260	~ 29,000 s	~ 40,000 s	~ 0.040 s

For instance, the training of the SWE_parabola and SWE_urban emulators was done through 260 unitary responses requiring a total of 1040/2080 computation-hours (training was performed in the HPC platform of the Computational Centre of the National Technical University of Athens), details are provided in Table 3.9. The emulator sampling at the validation tests took ~40 ms per iteration (in a 2.2 GHz Intel Core i7) which represents a significant time reduction. The benefit of proposing an emulator is strongly case dependent and is linked to the number of required parameters and shape of the simulator response. Figure 3.27 provides a measure for the increase in performance (measured by the decrease in L2 norm) with the training data set size for the three tested emulators.

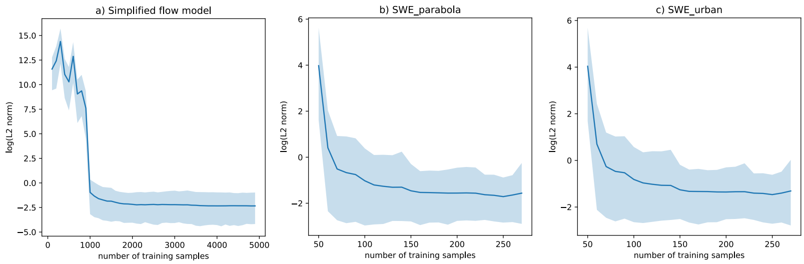


Figure 3.27: Emulator performance (mean and 95% quantile for L2 norm in logarithmic scale) and number of training samples. Number of parameters and polynomial degree in each case: a) 6-6, b) 3-5 and c) 3-5.

The stabilization of performance can be used to design a sequential sampling scheme, which can reduce the number of simulator samples.

3.4. CONCLUSIONS

3.4.1. PARAMETRIC INFERENCE AND EMULATION IN LARGE-SCALE INTEGRATED CATCHMENT MODELLING STUDIES

This chapter presents the example of an integrated catchment modelling study for evaluating water quality dynamics in a large-scale urbanised system in a formal Bayesian inference scheme. The model should capture the interaction between urban drainage processes, wastewater treatment plant and receiving water dynamics. This renders complex and slow modelling platforms. Using expert-literature elicited parametric ranges for river dynamics leads to highly uncertain dissolved oxygen simulations. Thus, as a means to reduce the output variability, a local data set could be used to update prior receiving water parametric knowledge using inference schemes. This inference requires large numbers of samples. However, the computationally expensive nature of integrated catchment simulators hampers the use of sensitivity analysis or formal inference.

Two dynamic emulators (polynomial orthogonal expansions) were fitted to represent flow and dissolved oxygen depletion for a one-year long time series under four and eight global parameters respectively. Only 200 model realisations sufficed to generate an acceptable interpolation in both cases. The emulators were validated using independent data, rendering a high quality mapping between the parametric space and the dynamic response. This technique still exhibits some severe limitations, like the impossibility to include large-parametric spaces, dynamic inputs (which are not parameterised) or parametric to output non-smooth relationships (e.g. bifurcation solution points). Nevertheless, these limitations did not apply to the present study. Both emulator schemes were used to compute the first order Sobol sensitivity indexes. Sobol indexes render a metric to assess the global sensitivity of model parameters to the output variability. This allowed increasing the understanding on the relevant processes involving flow propagation and dissolved oxygen depletion processes in the river Dommel. The sensitivity of oxygen depletion showed a non-stationary dependency across storm processes. Thus, the system does not always behave in the same manner depending on the dynamic state, the storm event and the season. In general terms, the reaeration rate showed to be the most relevant parameter under dry-weather flow. Depletion of fast-slow biodegradable matter is often the responsible for the sensitivity in the intensity of the oxygen dip (attending only to the river process parameters). Meanwhile oxygen recovery sensitivity is shared between the sediment oxygen demand and reaeration processes, with a strong curve dominated by seasonality (temperature driven). Several hypotheses were used to construct an inference error generating process. The use of an i.i.d Normally distributed error was sufficient to represent dissolved oxygen residuals. A heteroscedastic residual structure had to be fitted to the flow propagation process. However, both inferred residual series rendered a highly temporally correlated structure, which violates the assumption of independence. The residual autocorrelation is related to the measurement layout frequency (hourly) and the nature of the simulated processes such as; a strong memory effect, model structure-induced time shifts and input estimation-measurement errors. Various formulations to deal with the residual autocorrelation or structural bias were tested. However, the inferred dynamics either deteriorated or did not improve. Detailed investigation on the effects of neglecting the correlation structure in the dissolved oxygen residual structure is still missing. The use of alternative metrics for the inference of dissolved oxygen dynamics should be further studied.

The use of a dynamic emulation scheme allows gaining insights on the underlying mechanistic relationships of the integrated urban water quality system. This can be easily extended to similar environmental modelling studies, facilitating the application of sensitivity analysis, inference or calibration under long-time series and low-dimensionality parametric spaces.

3.4.2. EMULATION OF RAINFALL INPUTS AND PARAMETER SETS FOR PHYSICALLY-BASED FLOW DYNAMIC SIMULATIONS

In this section, a novel methodology to perform emulation of flow derived from hydrodynamic simulators is presented. This method allows approximating hydrographs generated by physically based hydrodynamic modelling under variations of model parameter values and rainfall series. This method is based on the combined use of the unit hydrograph theory (UH) and a polynomial chaos expansion (PCE) emulator. The use of the unit hydrograph theory is based upon the assumption that the underlying process is linear. This assumption is seldom met in physically based hydrodynamic simulators since the driving equations are fundamentally non-linear. In this new approach, information of the non-linear reaction of the model is encoded in the PCE structure by means of a dedicated sampling scheme.

The performance of the emulator was tested in three case studies; a) a simplified flow routing model based on a non-linear tank cascade, b-c) two synthetic computational domains simulated through the full form of the 2D Shallow Water Equations. In the three simulators the proposed emulator structure could approximate an output hydrograph under various combinations of rainfall time-series and model parameters. A comparison between the performance of the proposed emulator structure (UH_{PS}) and the use of the classical UH theory is also provided. This demonstrated that errors induced by the linear assumption of the UH theory can be significant when dealing with simulators based on the solution of the 2D-SWE. The method still has several limitations. For instance, the emulator can only accept spatially uniform rainfall time series. This can pose a severe constrain when dealing with large-scale catchments. Also, rainfall series inputs should have a constant sampling frequency. Errors in the emulator-simulator structure are expected to increase if the selected rainfall step-duration is much smaller than the catchment's concentration time for highly non-linear models, since the unit hydrograph superposition error will become dominant. The number of independent parameters is also limited in practice. This is a limitation of the emulator structure used (PCE) since the number of required model samples increases rapidly with the dimensionality of the parametric space. Nonetheless other data-driven surrogate strategies can be easily implemented. Further research should be conducted in these areas. The proposed methodology has the potential to significantly increase the range of applications of physically based flow propagation schemes, which is currently hampered by the computational burden.

4

Uncertainty analysis in large-scale integrated catchment modelling studies *

4.1. Introduction

Meeting the established environmental regulations (e.g. The Water Framework Directive of the European Union EC (2000)) is still a challenge in many densely urbanised catchments, as it often requires the implementation of intensive investment and regulatory plans (i.e. infrastructure construction, control systems or user limitations). Model-based decision-making is applied more frequently to explore and optimise the effect of different alternatives, aiming towards an efficient resource allocation. Therefore Integrated Catchment Modelling (ICM) has become an essential tool in the water quality management process over the last decades (Andrés-Doménech et al., 2010; Langeveld et al., 2013b; Willems and Berlamont, 2002). ICMs are, by definition, abstractions of highly complex water systems, usually constituted by the joint modelling of two or more subsystems of the urban water system (Keupers and Willems, 2017; Rauch et al., 2002). This often involves the joint simulation of sewer hydrodynamics, wastewater treatment processes, rural hydrology and river physical-biochemical dynamics (Benedetti et al., 2013a).

ICMs, like every other modelling process, contain various sources of uncertainty, due to the inherent system characteristics. Complex processes are represented with limited knowledge, relationships are calibrated with reduced data sets (which may lead to poorly identifiable parameterisations) and linked simulations are carried out over a wide range of spatiotemporal scales. Also, the stepwise process of abstraction from reality to model representation with its necessary simplifications and idealisations of the real systems includes the unavoidable occurrence of uncertainties (Muschalla et al., 2009).

These uncertainties encompass the errors introduced by model parameterisation, model-forcing data (e.g. precipitation), model input data (e.g. digital elevation model,

*This chapter is an adapted version of: Moreno-Rodenas, A. M., Tscheickner-Gratl, F., Langeveld, J., and Clemens, F.H.L.R. (2019). Uncertainty analysis in a large-scale water quality integrated catchment modelling study. Water Research, 158, 46-60 doi: 10.1016/j.watres.2019.04.016

soil or sewer conduit maps), model validation data (e.g. use of incorrect water level rating curves) and model structures (e.g. different mathematical representations).

The definition, recognition and consideration of these uncertainties is therefore of the utmost importance for the application of such models and for the interpretation of the hereby obtained results (Pappenberger and Beven, 2006; Schellart et al., 2010). At present however, a comprehensive uncertainty analysis is mainly applied in science and less in planning practice (Kleidorfer, 2010; Vanrolleghem et al., 2011).

Several frameworks have been proposed to facilitate the quantification and handling of these uncertainties in integrated urban water systems Deletic et al. (2012); Tscheikner-Gratl et al. (2017) or environmental modelling (Refsgaard et al., 2007). The quantification of modelling statistical uncertainties is often carried out by encoding system knowledge through a probabilistic description (Reichert et al., 2015) and sampling forward (i.e. Monte-Carlo sampling) to describe the variability at the targeted output variables. Additionally, the identification of the contribution of each uncertainty source is essential in the system analysis process, since it directs the modeller towards a rational reduction of epistemic uncertainties. For instance, Willems (2012) presented a variance decomposition methodology to quantify the partial contribution of uncertainty by source. Reichert and Mieleitner (2009) described the use of time-dependent parameters as a proxy to detect temporal windows of structural mismatch, pointing therefore at examining particular physical processes. Yang et al. (2018) used GLUE to extract the sensitivity from process-driven parameter structures. Inline with this, Gupta et al. (2008) discussed the need of diagnosis tools to guide in the model conceptualisation process.

Examples of uncertainty analysis applications in integrated catchment modelling for water quality estimation are still scarce (Tscheikner-Gratl et al., 2019). This is partially due to the significant amount of effort to monitor and set up large-scale modelling studies. Also, computational constraints have severely limited the applicability of proposed formal uncertainty analysis methodologies (see Schellart et al. (2010)). Therefore, only few examples are available in literature which deal with relatively small systems (Freni and Mannina, 2010a) or with individual sub-systems (Dotto et al., 2012; Radwan et al., 2004b).

Consequently, this chapter describes the application of a formal uncertainty analysis scheme to quantify dissolved oxygen modelling uncertainties in a large-scale (4400 ha of draining urban areas, a 750,000 p.e. WWTP and a sensitive receiving water body) water quality ICM study. The estimation of the relative contribution of different relevant uncertainty sources is also shown, which served towards directing further modelling and monitoring efforts in the system.

Initially, the forward propagation of all sources of uncertainty was proposed using the best available literature-expert-measurement derived parametric-input probability ranges. This will be therein referred as prior propagation. Upon analysis of the uncertainty contribution by source, river water quality and quantity parameters captured most of the variability of yearly dissolved oxygen (DO) dynamics. Thus, a dedicated inference scheme was proposed to update the river parametric distributions using local flow and DO measurements. The propagation of all parameter-input uncertainties and the updated parameter set for the river submodel is here presented (referred as posterior propagation). Comparison of the forward propagation from both prior-posterior uncertainty distributions with system observations, along with the current prioritisation of uncertainty sources is presented in this study. Complementarily, the impact of modelling uncertainties in the concentration-duration-frequency environmental assessment metrics is discussed, highlighting the possible implications of using such metrics in the selection of mitigation alternatives in environmental systems.

The discussion arising from this experience also serves to put into context the applicability of proposed uncertainty analysis techniques in real-world scale ICM studies.

4.2. MATERIALS AND METHODS

THE Dommel water system was used to describe the quantification and decomposition of uncertainty sources in the simulation of dissolved oxygen (described in Chapter 1). The model structure was revisited sequentially, identifying and analysing the major uncertainty sources for each submodel independently. This is reported in Tscheikner-Gratl et al. (2017), where the application of the Quantifying Uncertainty in Integrated Catchment Studies (QUICS) uncertainty framework was described for the Dommel case-study. The selection of the uncertainty sources was described in detail by subsystem attending to the modelling objectives.

4.2.1. PARAMETRIC UNCERTAINTY

Table 4.1 provides the characterised parameter probability distribution for the urban drainage water flow submodel. The most influential catchments were selected based on connected area and discharged volume contribution (Moreno-Rodenas et al., 2017b). Uncertainties in wetting losses (volume and availability), total connected area, in-sewer maximum storage, wastewater generation per inhabitant and reservoir constants (of the lumped sewer system conceptualisation) were considered. Those uncertainties were derived from expert knowledge in physical plausible constrained ranges.

Table 4.1: Urban drainage submodel water quantity parameters (~U uniform distribution (minimum, maximum), ~N normal distribution (mean, standard deviation))

Parameter name	**Units**	**Description**	**Uncertainty distribution**
MaxDepressionStoragec_24	mm	Wetting Losses storage (Eindhoven)	~U(1,5)
MaxDepressionStoragec_119	mm	Wetting Losses storage (Valkenswaard)	~U(1,5)
MaxDepressionStoragec_128	mm	Wetting Losses storage (Geldrop)	~U(1,5)
TotalArea_factc_24	-	Connected area multiplier (Eindhoven)	~U(0.9,1.1)
TotalArea_factc_119	-	Connected area multiplier (Valkenswaard)	~U(0.9,1.1)
TotalArea_factc_128	-	Connected area multiplier (Geldrop)	~U(0.9,1.1)
kc_24	min	Linear reservoir constant (Eindhoven)	~N(175,0.05x175)
kc_119	min	Linear reservoir constant (Valkenswaard)	~N(80,0.05x80)
kc_128	min	Linear reservoir constant (Geldrop)	~N(64,0.05x64)
V_Maxpipe_ES	m^3	Max in-sewer Storage volume (Eindhoven main)	~N(165,000,0.07x165,000)
V_MaxBT_119_1	m^3	Max in-sewer Storage volume (Valkenswaard_1)	~N(7,000, 0.07x7,000)
V_MaxBT_119_2	m^3	Max in-sewer Storage volume (Valkenswaard_2)	~N(5,000, 0.07x5,000)
V_MaxGB_119_3	m^3	Max in-sewer Storage volume (Valkenswaard_3)	~N(14,000, 0.07x14,000)
V_MaxGB_128	m^3	Max in-sewer Storage volume (Geldrop)	~N(27,800, 0.07x27,800)
V_MaxGB_127	m^3	Max in-sewer Storage volume (Mierlo)	~N(9,000, 0.07x9,000)
WastewaterPerIE	$m^3 d^{-1} pe^{-1}$	Waste-water production (All urban systems)	~N(0.19, 0.1x0.19)
YearlyEvaporation	$mm \cdot y^{-1}$	Average potential evaporation	~N(657, 0.2x657)

Table 4.2 presents the parameter distributions considered at the CSO water quality

generator. This contains four fractionation parameters, which were identified as a truncated Gaussian distribution (range [0, 1]) with a mean value provided by a non-formal expert elicitation and a standard deviation of 10% the mean value.

Table 4.2: Urban drainage submodel water quality parameters.

Parameter name	Units	Description	Uncertainty distribution
BOD_CSO	$g \cdot m^{-3}$	BOD concentration in CSO storm water	~Copula Model
COD_CSO	$g \cdot m^{-3}$	COD concentration in CSO storm water	~Copula Model
NH4_CSO	$g \cdot m^{-3}$	NH4 concentration in CSO storm water	~Copula Model
O2_CSO	$g \cdot m^{-3}$	O2 concentration in CSO storm water	~U(3,6)
fBOD1_BOD	-	fast BOD in total BOD (fraction)	~N(0.85,0.1x0.85)
fBOD1p_BODf	-	particulate BOD in fast BOD (fraction)	~N(0.3,0.1x0.3)
fBOD2_CODmBOD	-	slow BOD in COD-BOD (fraction)	~N(0.4,0.1x0.4)
fBOD2p_BODs	-	BOD particulate in the slow BOD (fraction)	~N(0.5,0.1x0.5)

Modelling in-sewer water quality dynamics is still a challenge (Willems, 2006) and sufficient data was not available for a reliable submodel calibration. The sewer transport system is characterised by long conduits (up to 20 km mixing gravity and pressurised sections) and it is heavily controlled, thus the measured water quality at the WWTP influent is expected to render a low representativity of the conditions at the CSOs, thus limiting the use of data-driven generators (Keupers and Willems, 2015). Thus a mean pollutant vector multiplier was used to approximate CSO loads from modelled flow dynamics. A monitoring campaign in the Dommel system reported concentration measurements for various water quality pollutants relevant for the integrated model (BOD, COD and NH4) at several CSO events (Moens et al., 2009). This allowed estimating pollutant probability distributions and correlation structures (Figure 4.1). A Gaussian copula stochastic model was proposed to generate random pollutant event concentrations (which respect the non-Gaussian marginal distributions and its correlation structure, Moreno-Rodenas et al. (2017c)). The stochastic model was based on the following hypotheses; a) Pollutant mean concentration remains fairly constant over the CSO event duration in the system of the Dommel (reported by van Daal-Rombouts (2017)), b) The studied CSO locations have comparable pollution dynamics over time (van Daal-Rombouts, 2017) and c) the instantaneous pollution concentration was assumed to be spatially independent since we could not establish a clear pattern from the available data. These assumptions are case study dependent. Also, the inclusion of a by-location correlation could be readily accounted for by the copula joint probability distribution, provided additional system measurements would prove its existence. A comparison of the measured and copula sampled pollutant concentrations can be seen in Chapter 2.

CSO dissolved oxygen concentrations were sampled from a uniform distribution U(3,6) gO2/m3. Since no onsite measurements were available, literature values were applied (Boomen and Icke, 2004; Diaz-Fierros et al., 2002).

Table E.3 and Table E.1 (Appendix E) show the uncertainty distributions associated with the WWTP submodel. Parameter distributions for the WWTP influent model (Table E1) were taken from Langeveld et al. (2017) which used several observed time-series to estimate the parameters of a WWTP empirical influent model using a Bayesian inference scheme. These parameters were described using a normally distributed density function (truncated between 0-1) with mean, the average of each posterior sample chain of the

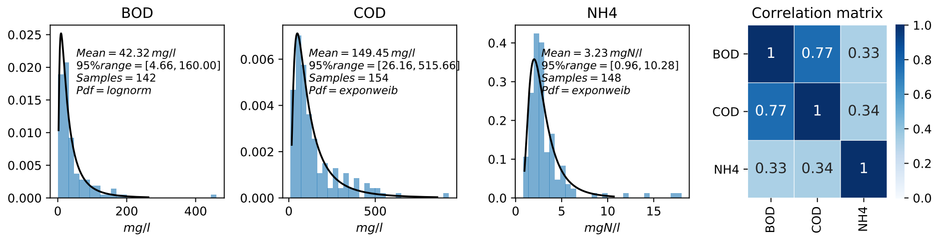


Figure 4.1: Measured distribution and spearman's correlation matrix of CSO pollutant concentrations at the Dommel system.

MCMC and standard deviation of 5% from the mean value.

Table 4.3 shows the prior parameter probability distributions assigned to the hydrologic flow and biochemical process models of the river. Prior distributions were defined from expert knowledge and literature. Sediment oxygen demand (SOD) was measured in the system by the water board De Dommel. The river model parameters were later inferred using a Bayesian inference scheme from flow and DO measurements at the closing section of the catchment (between 15-Jan-2012 until 04-Aug-2012) and validated with additional observations (05-Aug-2012 until 31-Dec-2012). A polynomial chaos expansion emulator was used to accelerate the sampling of the computationally expensive model during the inference process (Moreno-Rodenas et al., 2018a). This rendered an updated joint-parameter distribution set, which was later propagated through the full ICM by drawing correlated samples from the posterior parameter chains. The emulator construction and inference is described in detail at Chapter 3, Section 3.2 of this thesis.

Table 4.3: River hydrology and biochemical parameters prior distributions (* updated in the inference scheme).

Parameter name	Units	Description	Parameter prior
Kd1	d^{-1}	Decay rate for BOD fast	~U(0.5, 0.8)*
Kd2	d^{-1}	Decay rate for BOD slow	~U(0.2, 0.4)*
Vs1	$m \cdot d^{-1}$	Sedimentation rate for BOD fast	~U(0.2, 20)*
Vs2	$m \cdot d^{-1}$	Sedimentation rate for BOD slow	~U(15, 50)*
Knit	d^{-1}	Nitrification Rate	~U(0.15, 0.4)
TKd	-	Temperature coefficient for BOD oxidation	~U(1.03, 1.09)*
TKL	-	Temperature coefficient for reaeration	~U(1.01, 1.03)*
TSOD	-	Temperature coefficient for SOD	~U(1.045, 1.09)*
SOD	$g \cdot m^{-2} d^{-1}$	Sediment oxygen demand	~U(2.5, 3.5)
VKL	-	Velocity reaeration coefficient	~U(2, 3.5)*
MB	$g \cdot m^{-2}$	Macrophyte biomass	~U(40, 80)
n	$s \cdot m^{-1/3}$	Manning's roughness	~U(0.02, 0.12)*
K_z	-	Embankment slope multiplier	~U(0.7, 1.3)*
K_W_b	-	River estimated width multiplier	~U(0.7, 1.3)*
K_h	-	Rural flow input multiplier	~U(0.7, 1.3)*

Table 4.4 presents the parameter distributions used for the rural-to-river baseflow inflow pollutant loads. Values were estimated by expert elicitation since no measurements were available.

Table 4.4: Rural hydrology water quality inflow parameters.

Parameter name	Units	Description	Parameter distribution
NH4_in	$g \cdot m^{-3}$	NH4 inflow from rural	~U(0.05, 1.5)
BOD1_in	$g \cdot m^{-3}$	BOD1_in inflow from rural	~U(2, 4)
BOD1p_in	$g \cdot m^{-3}$	BOD1p_in inflow from rural	~U(0.01, 0.5)
BOD2_in	$g \cdot m^{-3}$	BOD2_in inflow from rural	~U(0.5, 1.5)
BOD2p_in	$g \cdot m^{-3}$	BOD2p_in inflow from rural	~U(0.01, 0.5)
BODs_in	$g \cdot m^{-3}$	BODs_in inflow from rural	~U(0.01, 0.5)
O2_in	$g \cdot m^{-3}$	O2_in inflow from rural	~U(5, 9)

4.2.2. DYNAMIC INPUT UNCERTAINTY

Errors in measured or estimated time-dependent inputs were represented as stochastic processes. Random sampling was applied to generate input realisation ensembles. Table E.3 presents the selected most relevant input processes in the system (indentified by Tscheikner-Gratl et al. (2017)). Rainfall uncertain realisations were sampled at the 9 largest catchment areas (which covered 3,930 ha of a total connected area of 4,400 ha, approx. 90%). Rainfall intensity at the spatial block-support of each individual catchment was extracted from the KNMI corrected radar measurements (1 km^2, 5 min) using hourly accumulation (Moreno-Rodenas et al., 2017b). An additive error model (dependent on rainfall intensity) was applied as proposed by Freni and Mannina (2010b).

$$R_i(t) = \begin{cases} R_{RAD,i}(t) \cdot (1 + \epsilon \cdot R_{RAD,i}(t)^a) & \text{if} \quad \int_{12h} |(R_{RAD,i}(t) - R_{RG,i})| dt \\ \sim U(0.9 \cdot min(R_{RAD,i}(t), R_{RG,i}(t)), 1.1 \cdot max(R_{RAD,i}(t), R_{RG,i}(t))) & \text{else} \end{cases} \tag{4.1}$$

in which the instantaneous estimated rainfall intensity, $R_{RAD,i}(t)$ at each location was corrected by a random normally-distributed error $\epsilon \sim N(\mu = 0, \sigma = 0.12)$, for $a = 0.2323$. This error model approximates a normal dispersion around the measured value, which parameters were fitted by comparing radar and rain gauge estimated rainfall series at the 13 locations in the system. The comparison between estimated rainfall from rain gauges and Radar sources showed large differences (especially during heavy convective storm processes, see bottom-left event at Figure 4.2). This measured difference could not be captured by the additive error structure alone. Thus a misdetection error structure was also applied. CSOs in the area are usually activated after rainfall volumes larger than 8-10 mm, and pumping capacities are designed to empty in-sewer storage in 10-12 hours. Therefore rainfall estimation differences were considered large when data sources (Radar and rain gauge inverse distance weighted interpolated values) had a cumulative difference larger than 10 mm within 12 hours. In such cases, the rainfall input ensemble is updated by sampling from a uniform distribution covering both estimations (interpolated rain gauge network – Radar, Equation 4.1). Figures 4.2 and Figure 4.3 show the rainfall input ensemble generated at four characteristic rainfall periods for the municipalities of Eindhoven (2000 ha) and Bergeijk (110 ha).

Biochemical processes at the river stretch are highly influenced by water temperature, which results in daily and seasonal variation. There were five temperature stations along the river section of interest with hourly measurements. A spatially homogeneous Gaussian Process (GP) was used to characterize the water temperature input uncertainty:

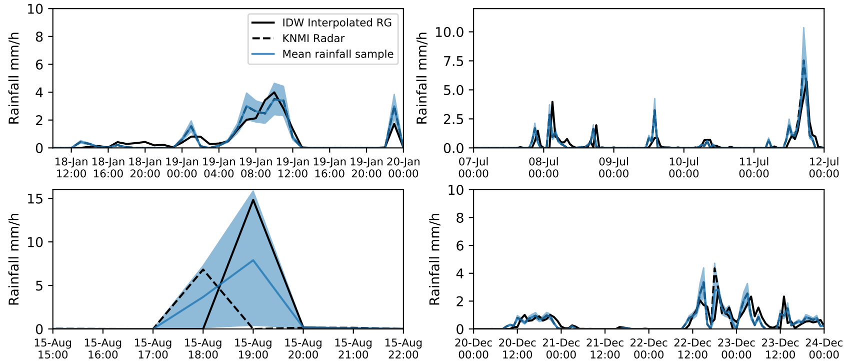


Figure 4.2: Example of the rainfall error model at four different periods. Estimated rainfall input mean (1000 samples) and 95% range (blue), KNMI radar estimate (dashed black) and interpolated (inverse-weighted-distance) value from the rain gauge network (black solid) at the urban system of Eindhoven.

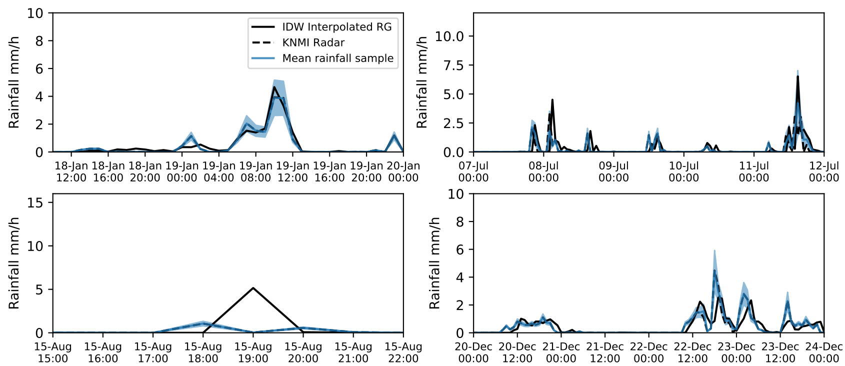


Figure 4.3: Example of the rainfall error model at four different periods. Estimated rainfall input mean (1000 samples) and 95% range (blue), KNMI radar estimate (dashed black) and interpolated (inverse-weighted-distance) value from the rain gauge network (black solid) at the urban system of Bergeijk.

$$T(t) \sim GP(\overline{T}(t), \Sigma_T) \tag{4.2}$$

with the average temperature between the five sensors $\overline{T}(t)$ as the mean of the process and covariance matrix described by a squared exponential structure:

$$\Sigma_{T,i,j} = \sigma_{ti} \cdot \sigma_{tj} \cdot e^{-0.5\cdot(\frac{t_d}{24})^2} \tag{4.3}$$

where σ_{ti} is the measured by-location temperature standard deviation at time i and j and t_d the time-lag.

The river solar radiation input ($Io_{measured}$) was measured at the KNMI meteorological station at the city of Eindhoven. This measured data were used as a spatially homogeneous time-dependent input for the entire river domain. An error model consisting in a normally distributed multiplier was used:

$$Io \sim Io_{measured} \cdot N_d(\mu = \mathbf{1}, \Sigma = \sigma^2_{Io} \cdot \mathbf{I}) \tag{4.4}$$

Water temperature at the wastewater treatment works was measured and used as input in the treatment process model. A multiplicative independent Gaussian error was implemented as:

$$TempWWTP \sim TempWWTP_{measured} \cdot N_d(\mu = \mathbf{1}, \Sigma = \sigma^2_{TempWWTP} \cdot \mathbf{I}) \tag{4.5}$$

where $\mathbf{1} \in \mathbb{R}^d$ is a vector of ones and $\mathbf{I} \in \mathbb{R}^{dxd}$ the identity matrix, being d the number of time-steps in the estimated series. The parameters $\sigma_{Io} = 0.15$ and $\sigma_T empWWTP = 0.05$ were assigned based on expected sensor representativity.

4.2.3. FORWARD UNCERTAINTY PROPAGATION

A Monte-Carlo (MC) based forward uncertainty propagation was performed using measured and expert-elicited parametric-input uncertainty probability distributions for the full year of 2012. Parametric samples were generated from a Latin hypercube sampler (LHS) to generate a low-discrepancy set. Dynamic inputs were sampled independently from the proposed stochastic processes. A first set of 600 samples was drawn from the full-integrated model (in parallel model instances) to characterize modelled river flow and dissolved oxygen concentration uncertainties at the closing section of the catchment (M_0121, Figure 1.4) using the best-available prior knowledge on the system. A Bayesian inference scheme was used to update the parametric distribution of several hydraulic and biochemical river model parameters (Table 4.3) given local observations in the system (Moreno-Rodenas et al., 2018a). The updated river submodel parameters were used to generate 600 additional samples from the model by drawing correlated samples from the inferred MCMC chains (Chapter 3, Section 3.2 of this thesis) and using LHS for the rest of the parameters. Both prior and posterior parameter distributions were compared with monitoring data available in the system.

4.2.4. UNCERTAINTY ANALYSIS BY VARIANCE DECOMPOSITION

A variance decomposition scheme was proposed following Willems (2012). This is based on defining independent groups of submodels or parameter-input clusters and analysing its contribution to the total residual error variance. Seven contributing groups were defined by selecting each submodel most relevant parameters-input sources (Table E.2, Appendix E). A period of ~ 2 months (05-Aug-2012 – 07-Oct-2012) was simulated which captured several summer oxygen depletion processes representative of the dynamics of interest. 15 time-points, which were temporally independent, and represented relevant dynamics, were selected to perform the decomposition of variance as:

$$\sigma^2_{E_Y - E_{Y_0}} = \sigma^2_{E_Y|rest\_var} + \sigma^2_{E_{Y_0}} + \sum_{i=1}^{7} \sigma^2_{E_Y|cluster_i} \tag{4.6}$$

for which $\sigma^2_{E_Y - E_{Y_0}}$ represent the variance of the model-observation residuals (which were computed during the forward propagation of all considered stochastic input and parametric uncertainty sources). The total residual variance is assumed to be composed of $\sigma^2_{E_{Y_0}}$ (variance of the measurement error), $\sigma^2_{E_Y|cluster_i}$ the partial variance of each selected parameter-input group (propagating the effect of the group meanwhile fixing the rest to their average value) and a term $\sigma^2_{E_Y|rest\_var}$ which represents the remaining variance not explained by measurement or input-parameters errors. $\sigma^2_{E_Y|rest\_var}$ was computed by estimating all other terms.

All variances should be homoscedastic (a Box-Cox transformation is often applied when data shows variance dependency). A very mild heteroscedasticity was found in the DO simulated residuals, which was corrected using a $\lambda = 0.9$ (Box-Cox). A dedicated data quality validation was performed to the DO measurement series in this period and the quality was expected to be high. Tolerances between 3-10% are often accepted in DO concentration measurements. Thus, a multiplicative random Gaussian error ($\sigma_{DO_t} = DO_t \cdot 0.05$) was applied to estimate $\sigma^2_{E_{Y_0}}$. Samples were drawn from a LHS (250 samples) varying every parameter-input cluster to characterize the partial variance ($\sigma^2_{E_Y|cluster_i}$). The reader is directed to Willems (2012) or Freni and Mannina (2010b) for further detail in the variance decomposition approach.

4

4.3. RESULTS AND DISCUSSION

4.3.1. FORWARD UNCERTAINTY QUANTIFICATION

UNCERTAINTY ranges for all relevant parameters and input sources were propagated through the full integrated model structure. Figure 4.4 presents the observed series of flow and DO at the final section of the system (M_0121, Figure 1.4) along with the mean of the simulated series and the 95% uncertainty range steming from the input-parametric variability when using all available prior knowledge in the system. The proposed model structure captures the flow dynamics in the river submodel, which are mainly driven by the baseflow inputs from the rural hydrology and the discharges from the WWTP (c.a. 40-50% of baseflow during summer) reasonably well (Nash-Sutcliffe efficiency of 0.82-0.86 at the inferred and validation river flow series respectively). The prior forward propagation of DO dynamics resulted in a relatively large dispersion; the 95% interval distance between model samples had a yearly average range of 3.7 $gO2/m^3$ ($\sigma = 0.76 gO2/m^3$). The 95% range was 3 $gO2/m^3$ for DO concentration lower than 2.5 $gO2/m^3$. Using the inferred water quality parameters the 95% range was reduced to 1.8 gDO/m^3 ($\sigma = 0.42 gDO/m^3$). Figure 4.5 provides the comparison of measured and modelled DO using the posterior forward propagation for a period of summer between July and October. Figure 4.6 and Figure 4.7 present the simulated rainfall, DO (measured-modelled) and the BOD dynamics at two sections of the river. Discharged BOD can reach high concentrations (up to 10-60 $gBOD/m^3$) at the receiving water body. Local measurements of BOD were not available at this period, but a monitoring campaign carried out during 2007-2008 (Moens et al. 2009) reported peaks of the same order of magnitude (up to 60-80 $gBOD/m^3$). Consumption of high BOD loads is the dominant process in the acute DO depletion. A fraction of the BOD load settles and degrades in the river bed, which dominates the speed of the DO recovery process (which can take between hours to days).

The depth and recovery pattern of most DO depletion processes are captured by the simulated and measured series (root-mean squared error of 1.25-1.6 $gO2/m^3$ during the inferred and validation time-periods respectively), yet a few events show insufficient depletion in the model. This can suggest that a certain process might still be missing in the model structure; i.e. insufficient WWTP loads under certain conditions or a stronger rural contribution. However, not enough system data was still available to further validate those hypotheses.

For instance, during the timespan between the 19th and the 29th of November 2012 three acute oxygen depletion processes occurred which were not captured by the model response (Figure 4.4). CSO discharges were not reported during this time. The same depletion process was captured by an upstream DO sensor located between the WWTP and the M_0121 sensor, but not by other sensors distributed throughout the system. During this period only minor inflow (approximately 12,000 m^3/h) was recorded at the influent of the WWTP (with 35,000 m^3/h maximum capacity). Thus, this oxygen depletion events could only be explained either by a major disturbance in the WWTP operations (which is not supported by effluent WWTP measurements of TSS, NH4 or COD) or a disturbance in the sediment bed of the river (i.e. dredging or mowing, which has been confirmed to have occurred by the waterboard De Dommel). Those processes were not described in the model structure and hence could not be captured by the simulation response. This type of events were not reported in monitoring data for previous years thus were neglected in this study. Yet this can become a relevant source of structural uncertainty and should be further investigated.

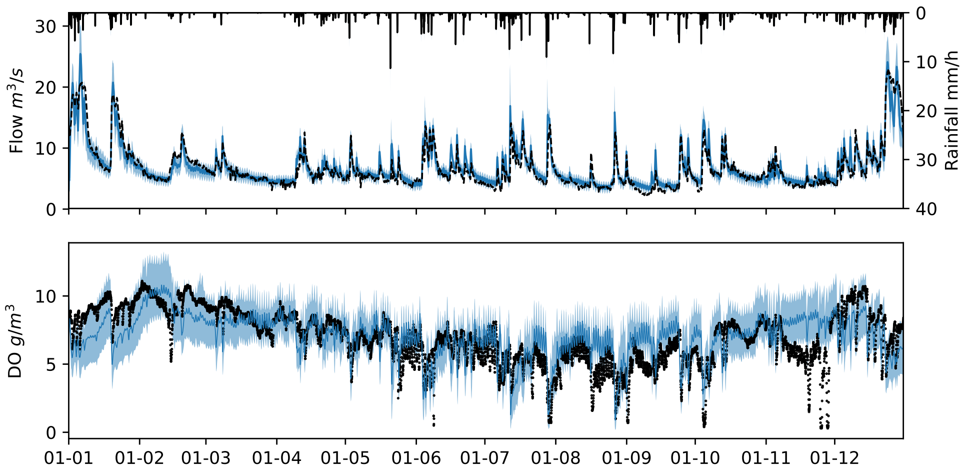


Figure 4.4: Full forward propagation of prior parameter-input distributions for flow, rainfall and dissolved oxygen at the closing section (M_0121) of the system, full 2012. Measured (black), simulated mean (solid) and 95% interval (band).

Figure 4.8 presents the cumulative probability density function of the yearly measured and modelled DO/flow series, along with the 95% ranges. As discussed previously the prior forward propagation rendered a large uncertainty range for the DO occurrence

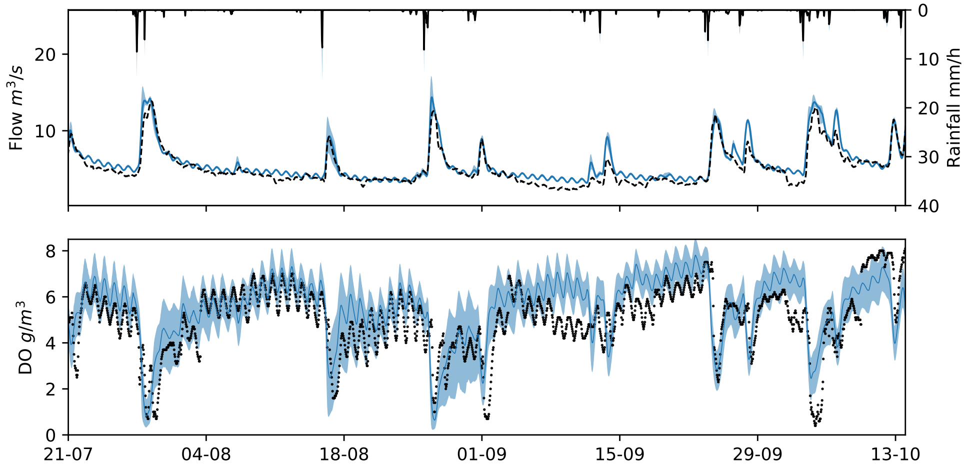


Figure 4.5: Full forward propagation of posterior parameter-input distributions for flow, rainfall and dissolved oxygen at the closing section (M_0121) of the system (21-Jul-2012 - 13-Oct-2012). Measured (black), simulated mean (solid) and 95% interval (band).

probability. When performing the posterior propagation (after inferring the parameter distributions for the river quantity and quality submodel parameters) the uncertainty band was significantly reduced; the average 95% uncertainty range was reduced from 3.7 $gO2/m^3$ to 1.8 $gO2/m^3$ and the two sample Kolmogorov-Smirnov (KS) test rendered that for 99.5% of the time-steps the null-hypothesis (both forward propagations render the same probability distribution) is rejected with p-value < 0.001 and mean K-S value of 0.27. This can be explained by the fact that the ranges attributed to the river parameters prior knowledge where narrowed down by the information provided in river DO-flow measurements. The mean simulated series for flow in both cases represented reasonably well the occurrence probability of high and medium hydrographs. Yet there was a systematic overestimation of low flows (river flow below 3-4 m^3/s), which is likely caused by an overestimation of the rural dry weather contribution. Yet, this is expected to have a limited influence in DO depletion dynamics which occur mainly during wet weather conditions.

The model exhibited a systematic underestimation of high dissolved oxygen concentration. This is seen during winter months (Figure 4.4). The seasonal DO variation in the model structure was captured by several factors; i.e. a constant sediment oxygen demand (SOD, see Table B.1, process number 5 Appendix B) and temperature driven inhibition coefficients for oxidation rates and reaeration patterns. SOD dynamics were estimated from system observations and we tried to respect the reported values. The inferred probability distributions for the biochemical river parameters could still not match well high DO concentration (Figure 4.8.b). This suggests that either the base DO inflows are underestimated in the current version of the model (e.g. too low WWTP DO effluent/Rural DO for which there were not reliable measurements) or that there is a structural process missing in the river conceptualisation (e.g. a stronger winter-summer sediment oxygen demand variability), a model structure improvement would require

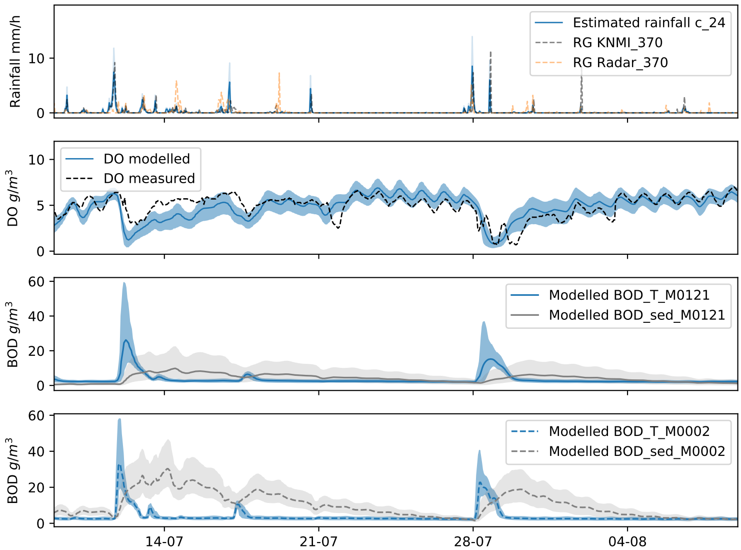


Figure 4.6: System dynamics detail comparing rainfall variability (KNMI rain gauge 370, KNMI Radar at the same location and the estimated intensity at the city of Eindhoven, c_24), measured-modelled dissolved oxygen, BOD_T (the sum of four fractions of BOD), BOD_sed (sediment BOD concentration) at two locations of the river M_0121, and M_0002 (Figure 1.4).

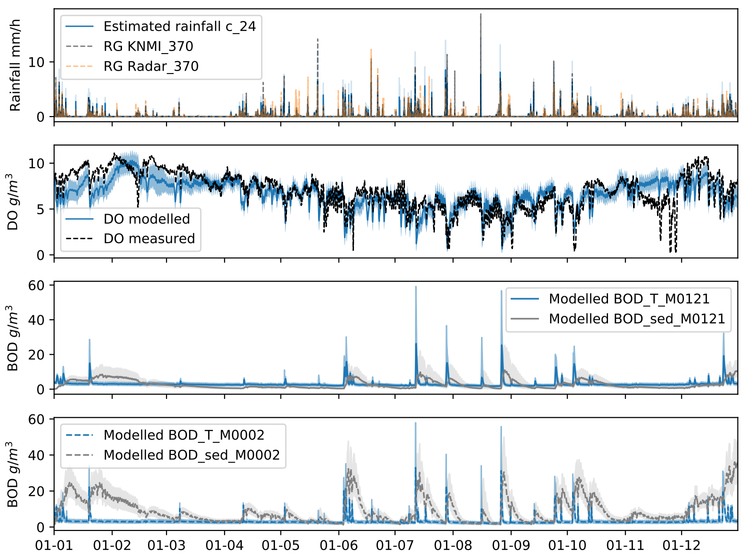


Figure 4.7: System dynamics during 2012 comparing rainfall variability (KNMI rain gauge 370, KNMI Radar at the same location and the estimated intensity at the city of Eindhoven, c_24), measured-modelled dissolved oxygen, BOD_T (the sum of four fractions of BOD), BOD_sed (sediment BOD concentration) at two locations of the river M_0121, and M_0002 (Figure 1.4).

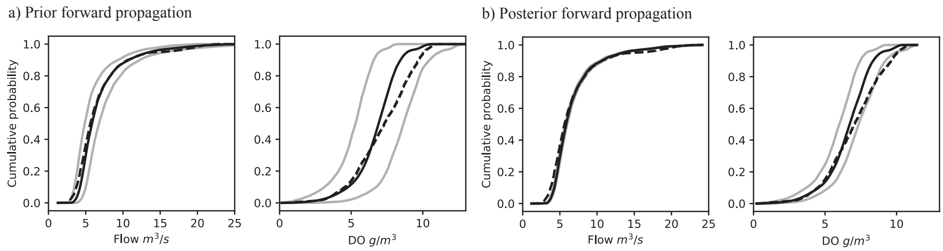


Figure 4.8: Cumulative probability density of Flow-DO measured (black dashed), simulated mean (black solid) and 2.5-97.5% percentiles (grey solid). a) Forward uncertainty propagation of all prior inputs-parameters. b) The resultant parametric-input uncertainty using the updated river water quality-quantity parameters.

further system observations. Nonetheless, the accurate simulation of high DO concentration is of little interest for the model application. On the other hand, low DO level yearly probability was better matched as can be seen in Figure 4.4.

The information from modelled-observed time-series in river water quality assessment studies is often compressed for system evaluation down to a low number of performance indicators. A common reporting method for river water quality status within the EU water framework directive compliance is the use of concentration-duration-frequency (CDF) tables (FWR, 2012). Limit levels are commonly extracted from an ecological assessment study, which defines survival conditions for critical species. Table A.1 (Appendix A) presents the environmental CDF tables for three levels of water quality compliance (Basic, Critical and Salmonid) in the river Dommel. Exceedance frequencies are computed and contrasted with the tolerated ones from which five status classes are derived:

1. Class 1: less than 0.5 times the tolerated frequency
2. Class 2: less than 1 time
3. Class 3: more than 1 time
4. Class 4: more than 1.2 times
5. Class 5: more than 2 times the tolerated frequency

Measured and modelled water quality status classes for three DO CDF tables are shown in Figure 4.9, where the black dot refers to the estimated class in the measurement dataset and the histogram shows the modelled status occurrence probability density. Low frequency (i.e. 0.1 and 1 times/year) would require a longer time series to be estimated reliably (3-10 years). Yet the current status of the river presents a fairly low quality status for low frequency-high magnitude oxygen depletion processes, which was captured in both measured-modelled series. Low magnitude-high frequency events with short duration had a higher degree of uncertainty. This is inline with the hypothesis that large DO events have often a lower degree uncertainty associated, since the system is oversaturated and despite the uncertainty sources, the model reproduces the strong DO depletion event. Yet low-medium events are often more sensitive to variability. Also, the information compression in the status classes for CDF metrics creates a

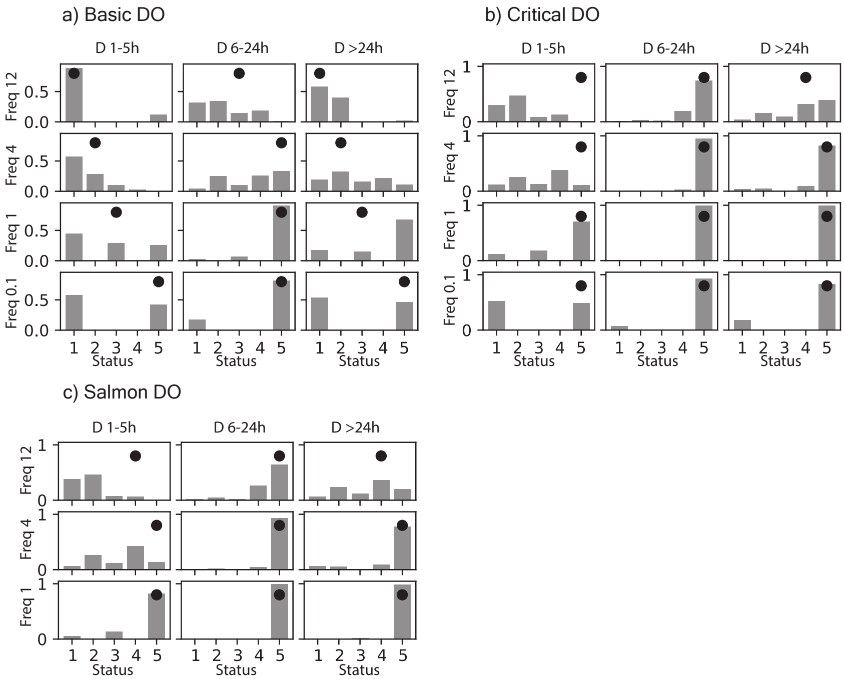


Figure 4.9: Water quality status assessments for the forward posterior propagation (histogram density) and measured data status (black circle). Basic, critical and salmonid tolerated dissolved oxygen depletion duration and yearly frequency (2012, excluded 19th-29th of November). The tolerated magnitude, duration and frequency of events in the system refer to the limits depicted in Table A.1 at Appendix A.

differential sensitivity. Class 5 captures a large range of system responses (>2 times the allowed frequency), meanwhile the occurrence of Class 1 to Class 4 has a sorter range (0-1.2 times the allowed frequency). Therefore systems with a poor environmental water quality status are easier to identify than those falling between good and medium status. This implies that CDF tables might not be the most appropriate metric to discriminate between the effects of different correction strategies (i.e. when selecting between a new real time control strategy or infrastructure investment to reduce DO depletion events). A modeller can discuss that the model performs well on the current system water quality status (poor environmental conditions, Class 5) but stochastic predictions of system improvements (Class 1-4) have the potential to result in a wider uncertainty range (probably beyond the marginal gain between alternatives).

4.3.2. UNCERTAINTY SOURCE ANALYSIS BY VARIANCE DECOMPOSITION

A variance decomposition scheme was applied to estimate the uncertainty contribution of different sources to the DO simulated series. The variance decomposition method provides a picture of the uncertainty contribution by source. However, it has certain limitations, which should be acknowledged and carefully addressed; first, it provides a lumped variance contribution, and thus identification of contribution at characteristic dynamic points proves difficult (limiting its diagnostic power). Secondly, it relies on certain hypotheses that might not be always met, as independence of error sources or homoscedasticity, which have the potential to distort the outcome, thus checking their influence is recommended. The design of this study tried to minimize the effect of these characteristics (following the recommendations of Willems (2012)).

Figure 4.10 provides the relationship between the estimated variance at DO and flow series in three representative points and the number of simulation samples. This shows that the number of samples selected in the uncertainty decomposition scheme (250 per cluster) was sufficient to provide a robust variance estimation. The decomposition of the prior parameters and input uncertainty (Figure 4.11) shows that the river flow and water quality parameters are the contributors of roughly 70% of the total uncertainty. Meanwhile rainfall uncertainty and CSO pollution parameters accounted for about 10% each. On the other hand, WWTP parameters and river inputs had a negligible effect on the global DO uncertainty. This shows that water quality and quantity river parameters captured most of the variability, which is caused by a high sensitivity of DO dynamics to the in-river biological processes and the relative poor knowledge on the actual parameter distributions. Therefore we proposed a parametric inference process for the river parameters, updating the prior assigned pdf's using local DO and flow measurements in the river. The inference process narrowed down the prior distribution of water quantity and quality parameters in the river, thus reducing the contribution of the parametric uncertainty in the river section to 16%. This reduction of parametric uncertainty was also transferred to the selected model-measurement error model in the likelihood formulation, which was captured in the variance decomposition scheme by the increase in the remaining uncertainty term (18%).

In the posterior variance decomposition, the contribution of rainfall uncertainty and the CSO pollution modules increased to 20-30% respectively. This shows that the DO dynamics are overall heavily influenced by rainfall driven discharges (WWTP and CSO) and especially by the water quality characteristics of the discharged volumes. Modelling CSO pollution concentrations is a challenge in urban drainage modelling and render highly uncertain results (Sandoval et al. 2018). Thus further monitoring and modelling efforts are still required in this area. Rainfall data uncertainties are shown to be relatively influential in the DO dynamics. The selected rainfall input error model (Equation 4.1) was motivated by the deviations found at rainfall estimations from the radar and the interpolated rain gauge network. Nevertheless, additional efforts should be directed to improve the quality of estimated rainfall inputs in the system.

The influence of the River submodel parameters is still not negligible (16%), which is mostly due to the river sediment oxygen demand parameter (SOD), for which system measurements show a temporal and spatial variabily (yet was here modelled as a global parameter). Urban drainage parameters had a comparatively low contribution to the DO variance (7%). This might be explained by the fact that UD parameters largely influence CSO discharge timing and shape, yet have relatively low effect on total discharged volume during large storm events. DO dynamics in the receiving water body are known to be intensively time-buffered by the nature of the process (Moreno-Rodenas et al., 2017b), thus pollutograph timing-shape errors are less dominant than errors in volume/mass estimation. The WWTP parameter set rendered a low contribution to the total variance (0.5%). WWTP is however a highly relevant process in the system dynamics and its proper conceptualisation and calibration is of foremost importance. Also, uncertainties associated with changes in regulations (Dominguez and Gujer, 2006) or operational changes can become dominant in certain scenarios and were here neglected. This study only focused on the effect of WWTP influent and fractionation parameters, which are reported to be some of the most relevant uncertainty sources (Belia et al., 2009) in WWTP outputs. Yet WWTP influent-fractionation parameters showed little influence in the modelling uncertainty for receiving water body DO concentration in the river Dommel.

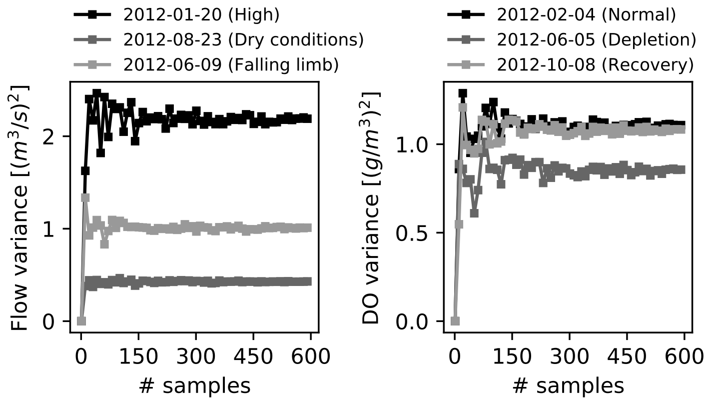
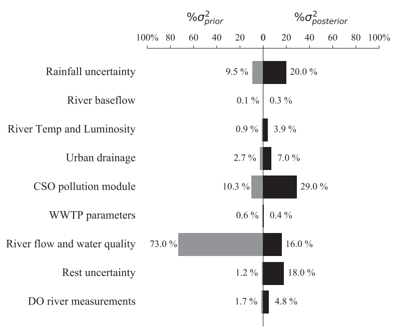


Figure 4.10: Flow and DO variance vs. number of model samples at three representative dynamic points.



Uncertainty sources	DO(Prior)	DO(Posterior)
Model input:		
Rainfall uncertainty	9.5	20
River baseflow input	0.1	0.3
River Temperature and Luminosity	0.9	3.9
Model structure and parameters:		
Urban drainage	2.7	7
CSO pollution module	10.3	29
WWTP parameters	0.6	0.4
River flow and water quality	73	16
Rest uncertainty	1.2	18
Measurement data:		
DO river measurements	1.7	4.8

Figure 4.11: Variance decomposition. Mean relative contribution [%] to the total model residual variance in DO at the receiving water body (location M_0121).

Few other studies exist that deal with uncertainty analysis of DO dynamic modelling in urbanised river catchments. For instance, Radwan et al. (2004b) presented a variance decomposition scheme for the modelling of DO in a river catchment in Belgium. Only the receiving water quality was modelled and CSO and rural pollution sources were considered as model inputs. They showed that roughly 30% of the variance of the process could be attributed to the river water quality parameters; whereas the input pollution loads explained 60% of it. Also, Freni and Mannina (2010b) and Freni and Mannina (2012) performed an uncertainty analysis of a small integrated system containing two urban drainage systems (~ 115 ha, 9,000 inhabitants), a WWTP and a river section in Sicily, Italy. The variance decomposition results showed that WWTP BOD discharge uncertainties were dominated by upstream submodels (e.g. sewer system or rainfall), whereas WWTP parameters had a lower influence. On the other hand, they reported that uncertainties in the water quality-quantity river parameters and the rest of the upstream submodels contributed 40% and 60% respectively to DO uncertainty at the receiving water body. Willems (2008) also presented a small-scale integrated catchment model (simulating urban drainage and WWTP but not receiving water) in which an uncertainty analysis scheme was proposed for several water quality variables (TSS, SS, BOD and NH4) reporting that rainfall uncertainty contribution represents 10-20% of the variance in all variables simulated at the outlet of the WWTP. Although uncertainty contribution is largely a case-dependent process, our results showed a similar structure as in previously reported studies; combined sewer overflow water quality characteristics and rainfall variability are the most relevant sources of uncertainty in DO simulation for the river Dommel in the studied period, provided that a detailed study is directed to identify and calibrate the river water quality dynamic processes.

The applicability of many uncertainty analysis techniques (Deletic et al., 2012; Jakeman and Jakeman, 2017) is limited when dealing with large-scale modelling applications. This is mainly due to insufficient observational data and the computational effort required. The example here provided shows that forward MC and variance decomposition schemes can be readily applied even for a computationally expensive systems (i.e. using code parallelisation and efficient sampling schemes). Performing Bayesian parameter inference is however prohibitive in most real-world cases, yet if carefully selecting a reduced number of parameters, model emulation can be used to accelerate the inference sampling (see emulators for hydrology; Machac et al. (2016a), Machac et al. (2018) hydrodynamic modelling; Moreno-Rodenas et al. (2018b), or for urban drainage inference; Wani et al. (2017)). In this study, the effect of updating prior knowledge in a set of highly influential parameters using the full ICM structure on the uncertainty contribution analysis was also shown.

Nevertheless, emulation strategies can only deal with low-dimensional parameter sets, thus inferring the full ICM parametric space falls beyond current data and computational capabilities for most real-world scale cases. In this line, Muschalla et al. (2009) discussed the process of abstracting ICM studies, reflecting about the need of partially calibrating (or inferring) sections of the full model based on intermediate measured state variables. Also, available datasets render many processes unidentifiable (e.g. in-sewer water quality processes), which should be accounted for during the uncertainty analysis scheme, yet identifiability still represent a major obstacle in the deployment of functional ICMs for environmental decision-making support.

4.4. CONCLUSIONS

THIS chapter presents an uncertainty analysis scheme of a large-scale integrated catchment model. The selected ICM accounts for urban drainage, WWTP and receiving water processes in an intensively urbanised catchment in the south of the Netherlands with the aim to simulate dissolved oxygen dynamics for ecological assessment. Stochastic error models were proposed for all relevant dynamic input sources, based either on measurement or process characteristics. Uncertainty probability distributions were assigned to the relevant parameters using all available knowledge on each process by means of expert elicitation or literature-measured ranges.

An uncertainty decomposition scheme showed that the river water quality and quantity parameters are responsible for ~70% of the uncertainty in DO river simulations when using expert elicited-literature based river parameter ranges. A first reduction of the uncertainty was achieved through inferring the river parametric vector using local measurement data (which reduced its contribution to ~16%). CSO water quality parameters were the most relevant source of uncertainty (~30%), indicating that monitoring and modelling efforts should be directed in that direction. Rainfall uncertainty accounted for roughly 20% of the variance in DO simulations, which was resultant of the relative disagreement of local rain gauge and radar rainfall estimations for the area. Our results showed similarities with previous reported studies for DO uncertainty analysis in integrated catchment systems. However, the generalisation of these results to other cases should be performed carefully since the structure and mechanistic relationship of system components may vary significantly and this has the potential to influence the uncertainty distribution by source.

The forward uncertainty propagation showed that the model structure lacks flexibility to accommodate seasonal winter DO levels. A suggested approach is to propose a time-dependent sediment oxygen demand process, yet this requires additional system observations. Also, the study of uncertainty propagation in concentration-duration-frequency ecological status tables showed that low water quality status systems are easier to identify (lower uncertainty associated) than good-moderate status. However, good and moderate water quality status are more sensitive to low DO variability and thus the uncertainty tends to be larger, limiting the identifiability of correction effects. This should be acknowledged when reporting modelling results and should be accounted for by decision-makers when dealing with simulation-based pollution mitigation measure selection.

5
CONCLUSIONS AND RECOMMENDATIONS

THIS thesis aims to enhance current knowledge on the practical applicability of uncertainty analysis frameworks in real-scale integrated catchment studies for the assessment of water quality impacts. In previous chapters the identification of individual sources of uncertainty when modelling dissolved oxygen dynamics in an urbanised catchment study was discussed. Additionally, examples are given to facilitate the application of uncertainty analysis methodologies when dealing with computationally expensive simulators. The decomposition of uncertainty contribution by source and the propagation of statistical uncertainties are presented for a large-scale integrated catchment simulator.

This chapter condenses the conclusions obtained through this work. Also recommendations for modellers, end users and further research in urban integrated catchment studies are provided.

5.1. CONCLUSIONS

THERE are still a limited number of examples of real-scale uncertainty analysis in the literature, which can be seen in Chapter 1 literature review. This hampers drawing general conclusions in a field, which is characterised by highly heterogeneous system configurations. Some of the reasons for this lack of studies are computational limitations for the applicability of formal methodologies to large-scale systems (discussed in Chapter 1 and developed in Chapter 3), the lack of incentives for practitioners to communicate uncertainties in their modelling endeavours, and the data and labour intensive process of setting up this type of simulators. Monitoring data availability is currently one of the most relevant limitations for the applicability of integrated catchment modelling in practice. Additionally, the following general conclusions can be drawn from the work presented in this thesis:

1 - Efforts in the description of rainfall inputs should focus on the spatial resolution

rather than the temporal one when simulating dissolved oxygen impacts in large-scale lowland systems. Chapter 4 shows that rainfall uncertainties contribute significantly to the variability of dissolved oxygen impact simulations in the river Dommel. Chapter 2 evaluates the relative influence of spatial and temporal scales contained in rainfall inputs from different data sources (i.e. different rain gauge networks and weather radar estimations). This showed that dissolved oxygen dynamics are eminently influenced by the total loading of the system, rather than the peaks in rainfall or flow. Lowland sewer systems are characterised by high in-sewer storage, and water quality impacts in the receiving water body present significant temporal smoothing processes. Therefore, errors in the rainfall total volume due to the spatial misrepresentation of the rainfall input dominate over time resolution (between 10-60 minutes).

2 - Data-driven model emulation can effectively facilitate the application of sampling intensive uncertainty analysis schemes in large-scale integrated catchment studies. As discussed in Chapter 1, formal application of uncertainty analysis schemes requires drawing a large number of samples to propagate uncertainties contained in the modelling process. However, real-world operational studies often render highly computationally expensive simulators, to which proposing sampling schemes fall beyond the computational budget of most practitioners and scientists. State of the art data-driven model emulation methods allow generating interpolation maps from a selected parametric space to the dynamic output space of the simulator. Chapter 3, section 3.2, shows a practical example in which year-long time series simulated outputs for flow and dissolved oxygen dynamics are emulated for a four and eight parametric dimensional space respectively. This allows proposing formal parametric inversion schemes, which facilitate the assimilation of observational data to update prior knowledge in river parametric uncertainties. These techniques have a direct transferability to many problems in integrated urban water quality modelling studies and its popularisation can stimulate the application of uncertainty propagation schemes in large-scale systems.

3 - The hybrid use of an orthogonal polynomial expansion and unit hydrograph theory can facilitate the assimilation of non-linearities present in flow propagation simulations to generate emulators which integrate parametric and rainfall temporal variability. One of the most relevant limitations in the application of data-driven model emulation is the dimensionality curse, which limits the dimensionality of the input space in emulation schemes. This forces generating interpolation maps, which deal with only a reduced number of global parameters, and yet under a fixed time-window. Thus, the emulation scheme is only valid under a given dynamic realisation, and its extrapolation in time is mostly hampered. In Chapter 3, section 3.3, a novel approach is presented for the assimilation of arbitrarily long (spatially homogeneous) rainfall series and parameter vectors in the emulation of 2D flow propagation schemes. The emulator strategy presented is shown to approximate closely the behaviour of purely physically based-simulations, yet at a fraction of the computational cost, which can facilitate the integration of physically based schemes in uncertainty analysis, early warning applications or fast simulation of overland flow in urban/rural areas.

4 - River water quality and quantity parametric uncertainty when using prior knowledge (extracted from literature-experts elicitation) dominates uncertainties in dissolved oxygen simulations. Provided inversion of river parameters, urban pollution loads (CSO water quality routine) and rainfall variability are the dominant sources of uncertainty. Chapter 4 deals with the decomposition of uncertainty by sources in a large-scale integrated catchment modelling study. River parametric uncertainty captured practically all output variability when using uncalibrated parameter ranges extracted from similar studies by expert elicitation and literature reported values. Thus, inferring or calibrating receiving water quality submodel parameters is highly relevant for a meaningful use of the modelling platform. Chapter 3, section 3.2 shows how to facilitate the inversion of parameter knowledge in river water quality and quality processes. After reducing the receiving water submodel parametric uncertainty, water quality load from combined sewer overflow discharges (CSOs) become dominant, and thus, modelling and monitoring efforts should be directed towards gaining further knowledge on in-sewer water quality processes. Rainfall spatial uncertainty also appeared as a relevant uncertainty source, thus appropriate description of this spatial process should receive detailed attention.

5 - Uncertainty quantification in concentration-duration-frequency environmental assessment metrics shows overconfidence in simulation results when deterministically validating systems under deteriorated status. Environmental metrics used to evaluate acute impact in water quality variables (e.g. dissolved oxygen, ammonia) are often based on Concentration-Duration-Frequency (CDF) tables, which assess the frequency of events and their magnitude. Tolerated values are extracted from ecological studies, which evaluate the viability of such conditions to sustain indigenous animal species populations. Often, integrated catchment studies are proposed when systems are deteriorated and corrective measures must be undertaken. Chapter 4 discusses the effect of statistical uncertainties in the status of the system based on CDF metrics for dissolved oxygen dynamics. The sensitivity of the metric to uncertainties in the modelling outcomes is expected to increase as the environmental status of the system improves. Implying that deterministically validating a poorly performing system based on CDF metrics alone can be relatively easy (since uncertainties are relatively low), yet the uncertainty in CDF metrics when the system is corrected, is expected to increase. Thus potentially precluding model-assisted decision-making based in CDF type metrics alone.

5.2. RECOMMENDATIONS AND FURTHER RESEARCH

THE current major bottleneck in integrated catchment modelling is the scarcity of monitoring data. Setting up extensive monitoring campaigns for long periods of time is costly in terms of material and labour costs. Yet integrated catchment modelling is severely hampered by the scarcity of observations in system dynamics. New techniques for distributed sensor networks are emerging in the field and the cost for data acquisition is expected to drop in the following years. Additional efforts to monitor urban drainage and river systems should be undertaken.

Further attention should be directed to combined sewer overflow water quality models. After detailed calibration of river water quality processes, pollutant loads from CSOs remain the largest source of uncertainty for the simulation of dissolved oxygen dynamics in the river Dommel, which is consistent with other literature examples. Current in-sewer modelling state-of-the-art is insufficient for a deterministic description of the pollutant loads during storm conditions. Further efforts should be undertaken to reduce epistemic uncertainties in water quality modelling for the sewer-to-river linkage. Especially needed are extensive datasets which allow exploring the link between easily measured variables (e.g. turbidity or conductivity) and costly but relevant ones (e.g. COD, BOD or NH4).

Additional work to accommodate input dynamics in data-driven emulator schemes is recommended. Data-driven emulation can facilitate the integration of uncertainty analysis schemes in large-scale computationally expensive simulators. However, this implementation is hampered by limitations in the dimensionality of the parameter vector under consideration. Further development is still needed to accommodate dynamic processes in the emulation scheme, thus increasing the horizon of applicability of such techniques.

Critical model evaluation should be encouraged in graduate (and post-graduate) programs for civil engineering. The culture of blindly relying in deterministic realisations of modelling endeavours should be changed. Especially relevant is to encourage such critical thinking within regulatory bodies as a mean to incentivise the consideration of model uncertainties in commissioned projects. Uncertainty analysis concepts should be transferred to future generations of civil engineers, which eventually will constitute the practitioner community for ICMs.

Additional examples of uncertainty analysis in real practical applications for integrated catchment modelling are still needed. The generalisation of results is hampered by the lack of studies in the field. Therefore complementing available knowledge would be beneficial to characterise the heterogeneity of water systems.

Alternative diagnostic methodologies to identify epistemic uncertainties in complex modelling applications have the potential to result in better directions towards improving model structures. The use of multi-metric techniques (e.g. approximate Bayesian computation) to diagnose and identify structural errors in water quality modelling platforms is promising. Adequate metrics directed to identify model structure failure mechanisms (e.g. missing CSO events, errors in WWTP effluents, mismatch in the recovery pattern of dissolved oxygen) should be defined and tested.

More robust environmental assessment metrics should be used in the model-based decision making process. Concentration-duration-frequency tables should be complemented with additional metrics to validate the performance of water quality simulation platforms and assist in a robust decision-making process.

The examples depicted in this thesis show efforts towards understanding and quantifying uncertainty sources when modelling dissolved oxygen dynamics. The simulation of DO is rather complex, since it is a non-conservative variable that depends on poorly understood processes expressed across highly heterogeneous scales. For instance, river water depth, turbulence or air velocity dominate the reaeration potential of flowing river bodies. The consumption of DO is driven by nitrification and oxidation processes. Organic and inorganic compounds which lead to oxygen consumption are present in dissolved and suspended form. Thus, they exhibit different mechanistic processes across the system (e.g changes in the availability of in-sewer stocks, river sediment dynamics etc.). The techniques here presented are nevertheless applicable to other variables of interest in the waste-water cycle process. Heavy metals, pharmaceutical compounds, ammonium or nitrate-phosphates are examples which present different physical and chemical process characteristics. Uncertainty analysis can help further describing the reliability of model predictions and direct efforts towards the improvement of model mathematical structures.

A

ENVIRONMENTAL ASSESSMENT METRICS, CONCENTRATION-DURATION-FREQUENCY TABLES

Concentration-duration-frequency tables for water quality environmental variables are estimated from ecological assessment studies. These studies assess the capability of selected fish and micro-invertebrate species to tolerate intermittent pollution conditions along all their life stages. Therefore a set of tolerated frequency is established in which the variable maintains a harmful concentration during a given time-period. Table A.1 and Table A.2 present the calculated tolerated concentrations for dissolved oxygen and ammonium in the river Dommel.

Table A.1: Environmental assessment concentration-duration-frequency (CDF) tables for dissolved oxygen concentration in the river Dommel (DO concentration in mg/l)

DO critical		Event duration			**DO basic**		Event duration			**DO salmon**		Event duration		
		1-5 h	**6-24 h**	**>24h**			**1-5 h**	**6-24 h**	**>24h**			**1-5 h**	**6-24 h**	**>24h**
Tolerated yearly frequency	**12**	5.5	6	7	Tolerated yearly frequency	**12**	3	3.5	4	Tolerated yearly frequency	**12**	5	5.5	6
	4	4	5.5	6		**4**	2.5	3	3.5		**4**	4.5	5	5.5
	1	3	4.5	5.5		**1**	2	2.5	3		**1**	4	4.5	5
	0.2	1.5	2	3		**0.2**	1	1.5	2					

Table A.2: Environmental assessment concentration-duration-frequency (CDF) tables for ammonium (NH4) concentration in the river Dommel (NH4 concentration in mg/l)

NH4 critical		Event duration		
		1-5 h	**6-24 h**	**>24h**
Tolerated yearly frequency	**12**	1.5	0.7	0.3
	4	2	1.2	0.5
	1	2.5	1.5	0.7
	0.2	4.5	3	1.5

Measured pollutant time series are converted to system failure events (following the

CDF conditions) and their yearly frequency is computed. By comparing the tolerated and measured-modelled failure frequencies a system status between 1-5 is given:

1. Class 1: less than 0.5 times the tolerated frequency
2. Class 2: less than 1 time
3. Class 3: more than 1 time
4. Class 4: more than 1.2 times
5. Class 5: more than 2 times the tolerated frequency

B

Process equations for the river model

Table A.1 presents the process matrix for the dissolved oxygen process routine at the river section. The model accounts for the transport and conversion rates of seven concentration state variables; BOD1 (Biological oxygen demand of fast biodegradable suspended fraction), BOD2 (Biological oxygen demand of slow biodegradable suspended fraction), BOD1p (Biological oxygen demand of fast biodegradable particulate fraction), BOD2p (Biological oxygen demand of slow biodegradable particulate fraction), BODs (Biological oxygen demand at the sediment section), NH4 (Ammonium) and DO (dissolved oxygen concentration).

Table B.1: Process matrix for the river water quality model structure

State Variable	**BOD1**	**BOD1p**	**BOD2**	**BOD2p**	**BODs**	**NH4**	**DO**	**rate**
	$gO2/m^3$	$gO2/m^3$	$gO2/m^3$	$gO2/m^3$	$gO2/m^3$	gN/m^3	$gO2/m^3$	
1a.Oxidation of fast-suspended fraction (BOD1)	-1						-1	$TKd^{T_{wat}-20} \cdot Kd1 \cdot BOD1 \cdot \frac{DO}{KO2+DO}$
1b.Oxidation of fast-particulate fraction (BOD1p)		-1					-1	$TKd^{T_{wat}-20} \cdot Kd1 \cdot BOD1p \cdot \frac{DO}{KO2+DO}$
2a.Oxidation of slow-suspended fraction (BOD2)			-1				-1	$TKd^{T_{wat}-20} \cdot Kd2 \cdot BOD2 \cdot \frac{DO}{KO2+DO}$
2b.Oxidation of slow-particulate fraction (BOD2p)				-1			-1	$TKd^{T_{wat}-20} \cdot Kd2 \cdot BOD2p \cdot \frac{DO}{KO2+DO}$
3a.Sedimentation of BOD1p		-1			+1			$Vs1 \cdot BOD1p$
3b.Sedimentation of BOD2p				-1	+1			$Vs2 \cdot BOD2p$
4.Oxidation of organic matter in the sediment					-1		-1	$TSOD^{T_{wat}-20} \cdot \frac{SOD}{d} \frac{DO}{KSO+DO}$
5. Constant sediment oxygen demand							-1	$TSOD^{T_{wat}-20} \cdot SOD \cdot \frac{BODs}{d} \frac{DO}{KSO+DO}$
6. Nitrification						-1	-4.57	$TKnit^{T_{wat}-20} \cdot Knit \cdot NH4 \cdot \frac{DO}{KNO2+DO}$
7.Photosynthesis macrophyte							+1	$TKp^{T_{wat}-20} \cdot kprodM \cdot Io \cdot \frac{MB}{d}$
8.Macrophyte oxygen consumption							-1	$TKp^{T_{wat}-20} \cdot kpcons \cdot \frac{MB}{d}$
9.Reaeration							+1	$TKL^{T_{wat}-20} \cdot VKL \cdot (CS - DO)$

The simplified river flow submodel was conceptualised as a tank in series lumped scheme, in which each tank is a well-stirred reactor in which the biochemical processes take place (Figure B.1). The mass balance equation is computed at each node:

$$\frac{dV_i}{dt} = Q_{in_i} - Q_{out_i} \tag{B.1}$$

where V_i is the volume at tank i and Q_{in_i} the sum of inflows from CSOs, upstream sections or rural hydrology. Q_{out_i} is the outflow from the tank at each instance and is computed as a free-surface flow following the Gauckler-Manning equation:

$$Q_{out_i} = A \cdot R_H^{2/3} \cdot slp^{1/2} \cdot n^{-1} \tag{B.2}$$

being A and R_H the area and hydraulic radius for a trapezoidal section with a given slope (slp) and bed Manning's roughness (n).

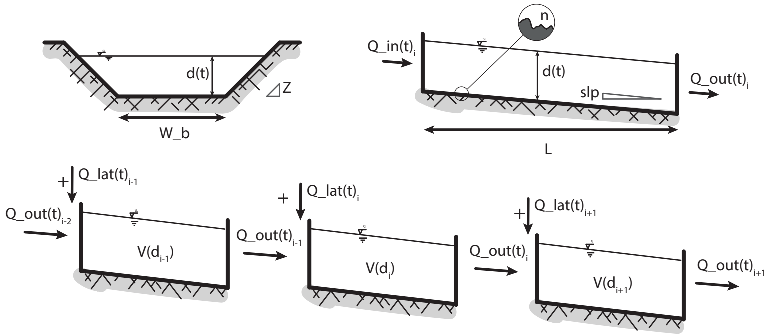


Figure B.1: Tank in series river flow scheme.

C

UNIT HYDROGRAPH SIMPLIFICATION UNDER LINEAR MODEL STRUCTURES

When using a linear flow propagation model the assumptions of proportionality and superposition of the unit hydrograph are congruent with the underlying dynamics. Figure C.1 shows the comparison between the simplified model described in the simplified routing scheme (Section 3.3.2.4, using $b = 1$, thus a set of linear tanks) and the use of a unit hydrograph derived from the response at 10 mm rainfall for an increasing unitary rainfall intensity from 1-50 mm/10min. Figure C.2 shows the same comparison for the joint proportionality and superposition assumptions for rainfall inputs of varying length (10-50 min) and intensity (1-20 mm/10min). In those cases the classical Unit Hydrograph theory can describe the model behaviour correctly, which is denoted by the almost null residual map

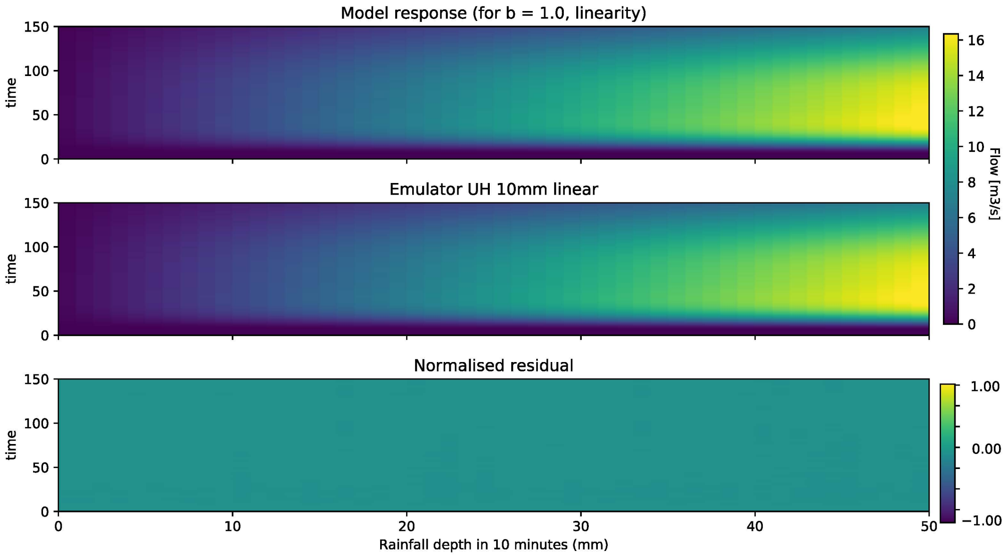


Figure C.1: Comparison of model vs linear unit hydrograph proportionality composition for a simplified linear model ($b = 1$).

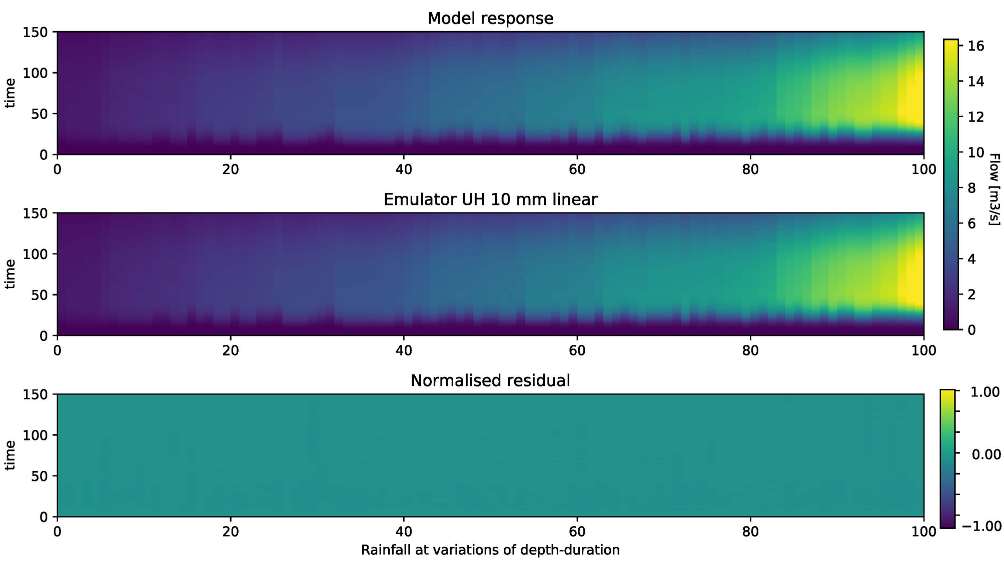


Figure C.2: Comparison of model vs linear unit hydrograph superposition and proportionality composition for a simplified linear model ($b = 1$).

D

RAINFALL VALIDATION SCENARIOS FOR THE 2D-SWE EMULATOR

Table D.1: Events (rainfall in mm and manning roughness) used to validate the 2D-SWE emulators.

	Val_0	**Val_1**	**Val_2**	**Val_3**	**Val_4**	**Val_5**	**Val_6**	**Val_7**
Parameter	n=0.035	n=0.042	n=0.037	n=0.043	n=0.036	n=0.040	n=0.038	n=0.041
time								
00:00:00	0.5	0.5	0.5	0.5	2	2	3	3
00:15:00	2	2	2	2	5	5	10	10
00:30:00	5	5	5	5	12	12	11	11
00:45:00	10	10	10	10	3	3	5	5
01:00:00	4	4	4	4	0	0	0	0
01:15:00	0	0	0	0	0	0	0	0
01:30:00	0	0	0	0	0	0	0	0
01:45:00	0	0	0	0	0	0	0	0
02:00:00	0	0	0	0	0	0	0	0
02:15:00	2	2	2	2	0	0	0	0
02:30:00	3	3	3	3	0	0	0	0
02:45:00	18	18	18	18	0	0	0	0
03:00:00	5	5	5	5	0	0	0	0
03:15:00	1	1	1	1	0	0	0	0
03:30:00	0	0	0	0	0	0	0	0
03:45:00	0	0	0	0	0	0	0	0
04:00:00	0	0	0	0	0	0	0	0

D

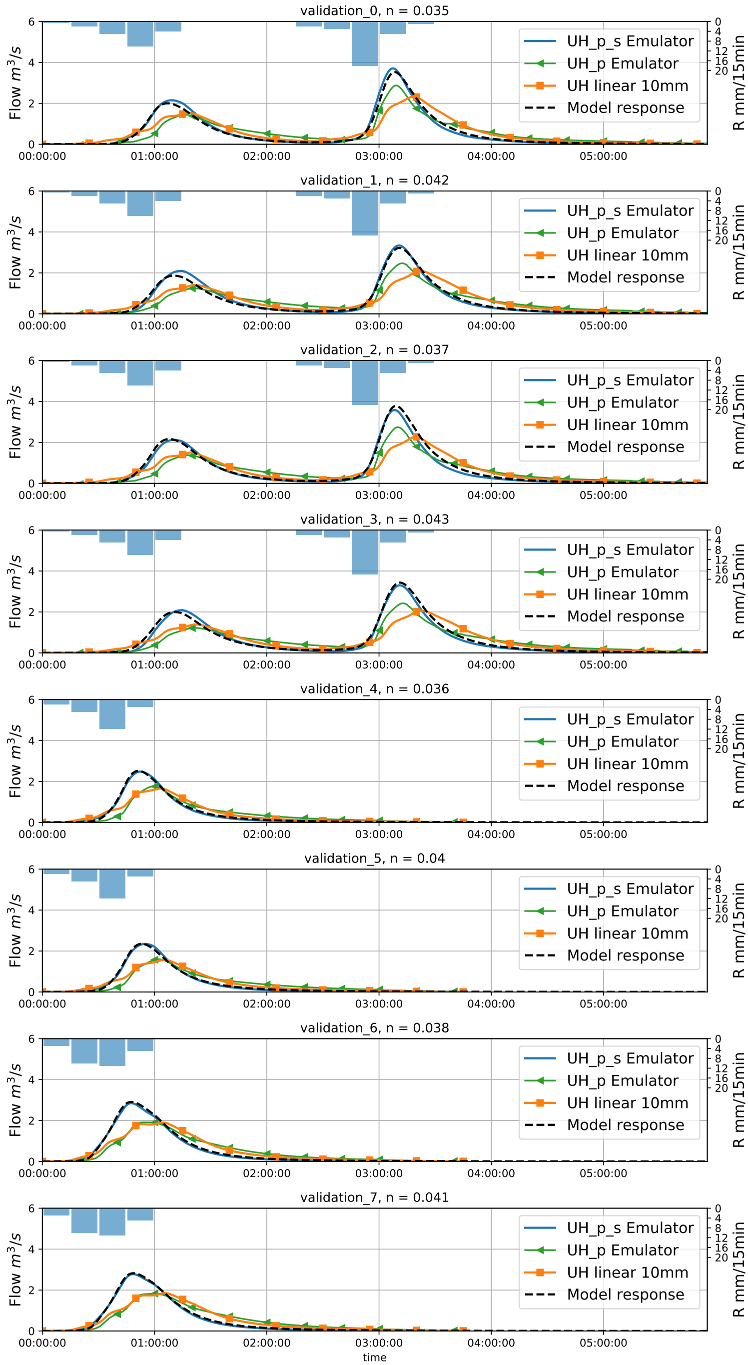


Figure D.1: Validation comparisons emulator vs SWE_parabola simulator.

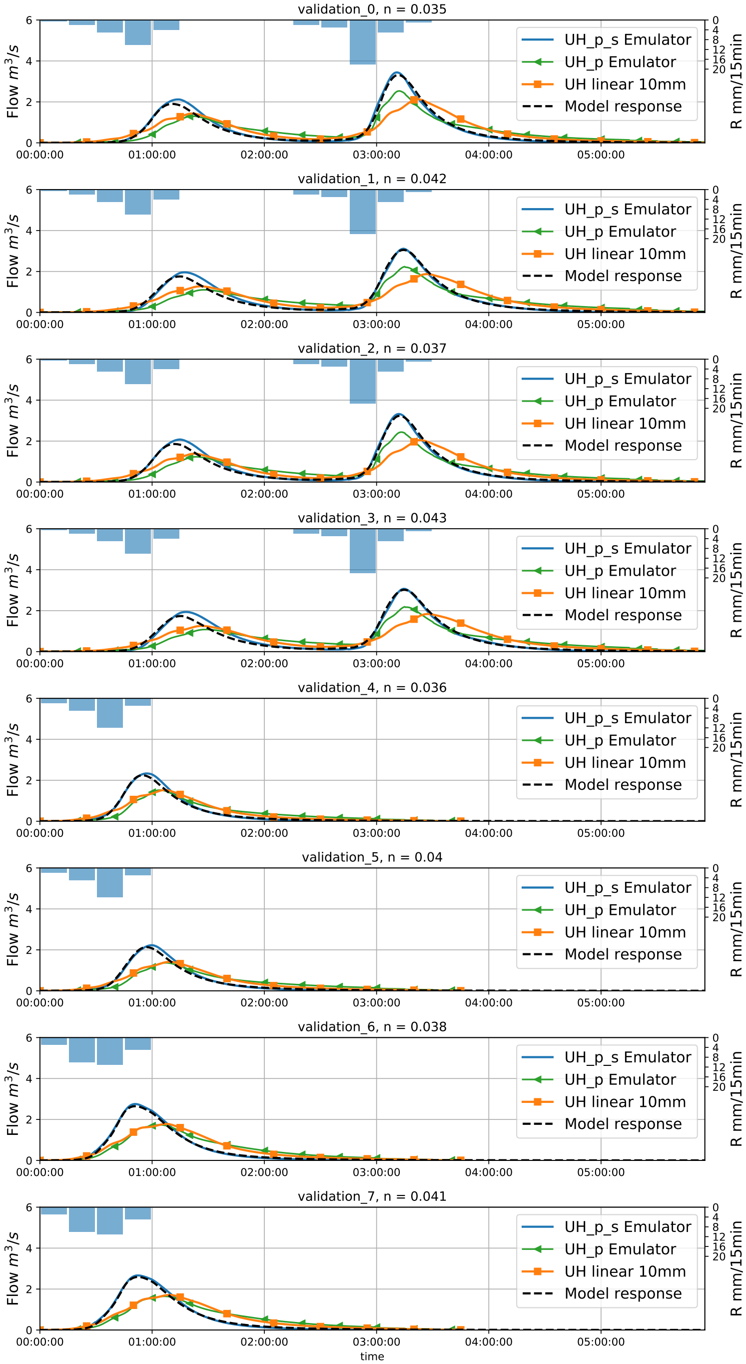


Figure D.2: Validation comparisons emulator vs SWE_urban simulator.

E

PARAMETRIC DISTRIBUTIONS FOR THE STATISTICAL UNCERTAINTY ANALYSIS OF THE DOMMEL ICM STUDY

This appendix contains additional information about the elicited distribution and input models used in the forward uncertainty propagation scheme depicted in Chapter 4. Table E.3 provides the parameter distributions for the empirical sewer-WWTP influent model extracted from Langeveld et al. (2017). Table E.1 shows the WWTP fractionation parametric distributions. Table E.2 shows the dynamic input groups and error models and Table E.4 presents the uncertainty sources clusters used in the variance decomposition scheme.

Table E.1: Uncertainty WWTP effluent fractionation model.

Name	Units	Description	Model
fBOD1p_BODf@WWTP2river	-	WWTP to river fractionation parameter	~U(0.05, 0.15)
fBOD2_BOD20@WWTP2river	-	WWTP to river fractionation parameter	~U(0.35, 0.45)
fBOD2p_BODs@WWTP2river	-	WWTP to river fractionation parameter	~U(0.25, 0.35)
f_S_F_w@fractionation	-	Fraction of fermentable readily biodegradable products in COD	~U(0.55, 0.65)
f_X_S_w@fractionation	-	Fraction of slowly biodegradable substrate products in COD	~U(0.4, 0.5)

Table E.2: Dynamic Input error models.

Name	Units	Description	Model
Rain_in_13	$mm \cdot h^{-1}$	Bergeik rainfall input	Equation 4.1
Rain_in_20	$mm \cdot h^{-1}$	Valkenswaard rainfall input	Equation 4.1
Rain_in_26	$mm \cdot h^{-1}$	Gestelse rainfall input	Equation 4.1
Rain_in_27	$mm \cdot h^{-1}$	Veldhoven rainfall input	Equation 4.1
Rain_in_33	$mm \cdot h^{-1}$	Geldrop rainfall input	Equation 4.1
Rain_in_34	$mm \cdot h^{-1}$	Mierlo rainfall input	Equation 4.1
Rain_in_36	$mm \cdot h^{-1}$	Eindhoven rainfall input	Equation 4.1
Rain_in_37	$mm \cdot h^{-1}$	Son and Breugel rainfall input	Equation 4.1
Rain_in_40	$mm \cdot h^{-1}$	Neunen rainfall input	Equation 4.1
Temp_WWTP	C^o	Water temperature at the WWTP_bio lines	Multiplicative ~N(1,0.03) [Equation 4.5]
I0_river	$W \cdot m^{-2}$	Solar radiation in the river surface	Multiplicative ~N(1,0.05) [Equation 4.4]
T_river	C^o	Water temperature in the river	$\sim GP(\overline{T}, \Sigma_T)$ [Equation 4.2]

Table E.3: Parametric Uncertainty WWTP influent model.

Name	Units	Description	Model
alpha_COD_av@ES_out	-	Influent water quality generator parameter*	~tN(0.63, 0.05x0.63)**
alpha_CODs_av@ES_out	-	Influent water quality generator parameter	~tN(0.95, 0.05x0.95)
alpha_NH4_av@ES_out	-	Influent water quality generator parameter	~tN(0.95, 0.05x0.95)
alpha_PO4_av@ES_out	-	Influent water quality generator parameter	~tN(0.63, 0.05x0.63)
alpha_TSS_av@ES_out	-	Influent water quality generator parameter	~tN(0.63, 0.05x0.63)
beta_COD_av@ES_out	-	Influent water quality generator parameter	~tN(0.47, 0.05x0.47)
beta_CODs_av@ES_out	-	Influent water quality generator parameter	~tN(0.82, 0.05x0.82)
beta_NH4_av@ES_out	-	Influent water quality generator parameter	~tN(0.82, 0.05x0.82)
beta_PO4_av@ES_out	-	Influent water quality generator parameter	~tN(0.47, 0.05x0.47)
beta_TSS_av@ES_out	-	Influent water quality generator parameter	~tN(0.47, 0.05x0.47)
alpha_COD_av@NS_out	-	Influent water quality generator parameter	~tN(0.63, 0.05x0.63)
alpha_CODs_av@NS_out	-	Influent water quality generator parameter	~tN(0.95, 0.05x0.95)
alpha_NH4_av@NS_out	-	Influent water quality generator parameter	~tN(0.95, 0.05x0.95)
alpha_PO4_av@NS_out	-	Influent water quality generator parameter	~tN(0.63, 0.05x0.63)
alpha_TSS_av@NS_out	-	Influent water quality generator parameter	~tN(0.63, 0.05x0.63)
beta_COD_av@NS_out	-	Influent water quality generator parameter	~tN(0.47, 0.05x0.47)
beta_CODs_av@NS_out	-	Influent water quality generator parameter	~tN(0.82, 0.05x0.82)
beta_NH4_av@NS_out	-	Influent water quality generator parameter	~tN(0.82, 0.05x0.82)
beta_PO4_av@NS_out	-	Influent water quality generator parameter	~tN(0.47, 0.05x0.47)
beta_TSS_av@NS_out	-	Influent water quality generator parameter	~tN(0.47, 0.05x0.47)
alpha_COD_av@RZ_out	-	Influent water quality generator parameter	~tN(0.49, 0.05x0.49)
alpha_CODs_av@RZ_out	-	Influent water quality generator parameter	~tN(0.95, 0.05x0.95)
alpha_NH4_av@RZ_out	-	Influent water quality generator parameter	~tN(0.95, 0.05x0.95)
alpha_PO4_av@RZ_out	-	Influent water quality generator parameter	~tN(0.49, 0.05x0.49)
alpha_TSS_av@RZ_out	-	Influent water quality generator parameter	~tN(0.49, 0.05x0.49)
beta_COD_av@RZ_out	-	Influent water quality generator parameter	~tN(0.47, 0.05x0.47)
beta_CODs_av@RZ_out	-	Influent water quality generator parameter	~tN(0.98, 0.05x0.98)
beta_NH4_av@RZ_out	-	Influent water quality generator parameter	~tN(0.98, 0.05x0.98)
beta_PO4_av@RZ_out	-	Influent water quality generator parameter	~tN(0.49, 0.05x0.49)
beta_TSS_av@RZ_out	-	Influent water quality generator parameter	~tN(0.49, 0.05x0.49)

*The influent parameters are equivalent to the proposed by Langeveld et al. (2017).
**~tN truncated ([0-1]) normal probability distribution.

Table E.4: Variance decomposition parameter-input groups.

Parameter-Input group	**Elements**
Rainfall uncertainty	Rain_in (all rainfall dynamic inputs in Table E.2)
River baseflow input	K_h (Rural input multiplier, Table 4.3)
River Temperature and luminosity	I0_river and T_river (Table E.2)
Urban drainage parameters	Table 4.1
CSO pollution module	BOD, NH4, COD and DO (x30 CSO locations, Table 4.2)
WWTP parameters	Table E.3 and Table E.1
River flow and water quality parameters	n, k_W_b, k_z, kd1, kd2, Vs1, Vs2, SOD, VKL, Knit, TKd and TKL (Table 4.3)

Acknowledgements

It seems like it was yesterday, that rainy and cold November I arrived to Delft. I remember well those days, hosted by Pablo Vizcaino in that high-up student flat in the outskirts of Delft. I remember those nights from the inflatable mattress, staring a the glowing greenhouses in the dark horizon and thinking about an uncertain future. It is nice to see how time gives you a different perspective. Four PhD years pass in the blink of an eye, yet many things happen during four years. It has been an intense journey, an emotional maze of fluctuating scientific excitement, sadness, stress, motivation and happiness. I want to thank all those that have been with me through them.

I would like to thank Prof. Francois Clemens, my promotor who managed to direct me without restricting my freedom, who constantly encouraged me to defy hierarchy and with whom I have shared the excitement of doing crazy experiments. I thank Dr. Jeroen Langeveld, my supervisor, for being the structure of my progress, for being there whenever needed.

I would like to thank my colleagues and organisators of the EU project QUICS with whom I have shared inestimable experiences. I would like to thank Francesca Cecinati for the crazy kriging time, also to Vivian Camacho and to Omar Wani and Jörg Rieckermann for their very attentive hosting which allowed me to learn and experience a different research environment.

I will always remember those transcendental discussions by the Aegean Sea with Franz Tscheickner-Gratl and Vasilis not so Bellos. Discussions which propagated through England, Netherlands and hopefully more places in the future.

I want to thank my colleagues at TUDelft; Nicola, Wouter, Johan and Petra. Of course also M. Spekkers, Matthijs, Elena, Bram, Eva, Kostas, Job and the magnificus A. Duinmeijer. I would like to thank Mathieu Lepot, for his tries to cheer his colleagues whenever needed.

I would like to thank Adithya, that indian with whom I began in the office and later shared surprising adventures like 1000 km by car with a french and a guinea pig towards la petite Bretagne, another 1000 km in Tamil lands and less than 1000 days below the same roof. I would like to thank other dearest inhabitants of Boterbrug, Mathilde and Magela. I thank Monicaco, Angel and Sotiria for having made Delft a bright and special place.

Although from the distance I still value and thank Raúl González for the inestimable friendship across space and time. Also Víctor Ferra, Edu Portillo and Vicente Mataix. I hope more adventures are still to come.

Agradezco enormemente el apoyo de mi familia, de los que están y los que si estuviesen (como dijo Pablo) estarian muy orgullosos. Mi madre y mi padre que me han dado lo mas valioso que tengo, una educación y los mejores valores que han podido transmitirme. Le doy las gracias a Pablo, mi hermano, que espero me siga necesitando para hacer tfgs, examenes de control y lo que quiera, porque yo se que siempre puedo contar con él.

For last, I thank from the bottom of my heart to Sophie Le Hesran. For being always by my side, whether supporting me, walking in the middle of the night through Delft because I don't want to sleep or making brighter every day of my life.

LIST OF PUBLICATIONS

Peer-reviewed journal publications

Moreno-Rodenas, A., Tscheickner-Gratl, F., Langeveld, J. and Clemens, F.H.L.R (Under review 2019). *Uncertainty analysis in a large-scale water quality integrated catchment modelling study*, Water Research **158**, 46-60, doi: 10.1016/j.watres.2019.04.016.

Moreno-Rodenas, A., Langeveld, J. and Clemens, F.H.L.R (Under review 2019). *Parametric emulation and inference in computationally expensive integrated urban water quality simulators*, Environmental Science and Pollution Research, doi: 10.1007/s11356-019-05620-1.

Moreno-Rodenas, A., Bellos, V., F., Langeveld, J. and Clemens, F.H.L.R (2018). *A dynamic emulator for physically based flow simulators under varying rainfall and parametric conditions*, Water Research **142**, 512-527, doi:10.1016/j.watres.2018.06.011.

Moreno-Rodenas, A., Cecinati, F., Langeveld, J. and Clemens, F.H.L.R (2017). *Impact of Spatiotemporal Characteristics of Rainfall Inputs on Integrated Catchment Dissolved Oxygen Simulations*, Water **9(12)**, 926, doi:10.3390/w9120926.

Duinmeijer A., Moreno-Rodenas A.M., Lepot M., van Nieuwenhuizen C., Meyer I. and Clemens F.H.L.R. (2019). *A Simple measuring set-up for the experimental determination of the dynamics of a large particle in the 3D velocity field around a free surface vortex*, Flow Measurement and Instrumentation **65**, 52 - 64, doi:10.1016/j.flowmeasinst.2018.10.007.

Cecinati, F., Moreno-Rodenas, A., Rico-Ramirez, M., ten Veldhuis, M.-c. and Langeveld, J. (2018). *Considering Rain Gauge Uncertainty Using Kriging for Uncertain Data*, Atmosphere **9(11)**, 446, doi:10.3390/atmos9110446..

Bellos, V., Kourtis, I., Moreno-Rodenas, A. and Tsihrintzis, V. (2017). *Quantifying Roughness Coefficient Uncertainty in Urban Flooding Simulations through a Simplified Methodology*, Water **9(12)**, 944, doi:10.3390/w9120944.

Tscheickner-Gratl, F., Bellos, V,. Schellart, A., Moreno-Rodenas A., Muthusamy, M,. Langeveld, J., Clemens, F.H.L.R, Benedeti, L., Rico-Ramirez, M.A., Fernandes de Carvalho, R., Breuer, L., Shucksmith, J., Heuvelink, G.B.M and Tait, S. (2019). *Recent insights on uncertainties present in integrated catchment water quality modelling*, Water Research **150**, 368 - 379, doi:10.1016/j.watres.2018.11.079.

International conference publications

Moreno-Rodenas, A. M., Langeveld, J. and Clemens, F.H.L.R. (2018). *Parametric inference in large water quality river systems*, In proceedings fo the 11th International Conference of Urban Drainage Modelling, UDM 2018. Palermo, Italy. (3rd best research paper Award)

Moreno-Rodenas, A. M., Bellos, V,. Langeveld, J., and Clemens, F.H.L.R. (2018). *Learning non-linearities from flow simulators*, In proceedings fo the 5th IAHR europe congress, IAHR 2018. Trento, Italy.

Moreno-Rodenas, A., Langeveld, J., and Clemens, F.H.L.R. (2017). *Accounting for correlation in uncertainty propagation, a copula approach for water quality modelling*, In Proceedings of the 14th IWA/IAHR International Conference on Urban Drainage. Prague, Czech Republic.

Moreno-Rodenas, A., Bellos, V., Langeveld, J. and Clemens, F.H.L.R. (2017). *Influence of river routing methods on integrated catchment water quality modelling*. Panta Rhei: Proceedings of the 10th World Congress of EWRA on Water Resources and Environment. Athens, Greece, 293–298.

Cecinati, F., Moreno-Rodenas, A. and Rico-Ramirez, M. *Integration of rain gauge errors in radar-rain gauge merging techniques*. Panta Rhei: Proceedings of the 10th World Congress of EWRA on Water Resources and Environment. Athens, Greece, 293–298.

Scientific reports

Sriwastava, A. and Moreno-Rodenas, A. (2018). *Report on uncertainty frameworks and potential improvements*. https://www.sheffield.ac.uk/quics/dissemination/reports, D1.1, D4.2

Tscheikner-Gratl, F., Lepot, M., Moreno-Rodenas and A. and Schellart, A. (2017). *A Framework for the application of uncertainty analysis*. https://www.sheffield.ac.uk/quics/dissemination, D6.7, doi:10.5281/zenodo.1240926.

Heuvelink, G.B.M., Cecinati, F., Lepot, M., Moreno-Rodenas, A.M., Sawicka, K. and Torres-Matallana, A. (2017). *Guidelines and manuals for end-users on uncertainty propagation analysis and sampling designs*. https://www.sheffield.ac.uk/quics/dissemination/reports, D6.2

About the Author

Antonio M. Moreno Rodenas was born in Valencia, Spain on August 29th 1990. He attended his elementary, primary and secondary education at the school *Jesús-María "Fernando el Católico"*. Immediately after, he began his studies of Civil Engineering at the School of *"Ingenieros de Caminos Canales y Puertos"* at the Polythecnic University of Valencia. At the end of his studies he visited during one year the University of Bristol (UK) where he investigated the the description of spatial and temporal storm processes. He graduated as a Civil Engineer with Honours in his final dissertation in 2014. In November of 2014 he moved to the Netherlands to start as a PhD candidate at the Delft University of Technology in the Water management department. During the last four years he has been investigating methods to facilitate uncertainty analysis in mathematical models for the simulation of water quality dynamics. His main target was to understand how different parts of catchment-scale water quality models contribute to uncertainties in their predictions. During his research he has been lucky enough to collaborate with research groups from Switzerland, United Kingdom and Greece. Also, he has attended to several international scientific conferences in the water quality and quantity sector. Next to his research, he has been involved in several projects which link personal and scientific interests. Antonio has assisted in the creation of a computer vision 3-D tracking system for his colleague's balls captured in a water vortex. He has also been pursuing the development of an autonomous microscopy system to extract behavioural patterns from small microinvertebrates. A significant part of the last four years of his work are summarised in this Thesis.

BIBLIOGRAPHY

Aarts, A., Langeveld, J., Ten Veldhuis, J., and Rietveld, L. (2013). Combined sewer overflows in urban areas-interpreting and comparing european cso monitoring results. In *Proc. 7th International Conference on Sewer Processes and Networks, Sheffield, UK, 28-30*.

Abbott, M., Bathurst, J., Cunge, J., O'Conell, P., and Rasmussen, J. (1986). An introduction to the european hydrological system – systeme hydrologique europeen, "she", 1: History and philosophy of a physically-based, distributed modelling system. *Journal of Hydrology*, 87:45–59.

Ammann, L., Reichert, P., and Fenicia, F. (2018). A framework for likelihood functions of deterministic hydrological models. *Hydrol. Earth Syst. Sci. Discuss.*, 2018:1–39.

Andrés-Doménech, I., Múnera, J. C., Francés, F., and Marco, J. B. (2010). Coupling urban event-based and catchment continuous modelling for combined sewer overflow river impact assessment. *Hydrology and Earth System Sciences*, 14(10):2057–2072.

Bach, P. M., Rauch, W., Mikkelsen, P. S., McCarthy, D. T., and Deletic, A. (2014). A critical review of integrated urban water modelling – urban drainage and beyond. *Environmental Modelling & Software*, 54:88–107.

Bates, B. C. and Campbell, E. P. (2001). A markov chain monte carlo scheme for parameter estimation and inference in conceptual rainfall-runoff modeling. *Water Resources Research*, 37(4):937–947.

Beck, M. B. (1987). Water quality modeling: a review of the analysis of uncertainty. *Water Resources Research*, 23(8):1393–1442.

Bedford, T. and Cooke, R. (2001). *Probabilistic Risk Analysis. Foundations and Methods.* Cambridge University Press, Cambridge.

Belia, E., Amerlinck, Y., Benedetti, L., Johnson, B., Sin, G., Vanrolleghem, P. A., Gernaey, K. V., Gillot, S., Neumann, M. B., Rieger, L., Shaw, A., and Villez, K. (2009). Wastewater treatment modelling: dealing with uncertainties. *Water Science and Technology*, 60(8):1929–1941.

Bellos, V., Kourtis, I., Moreno-Rodenas, A., and Tsihrintzis, V. (2017). Quantifying roughness coefficient uncertainty in urban flooding simulations through a simplified methodology. *Water*, 9(12):944.

Bellos, V. and Tsakiris, G. (2014). Comparing various methods of building representation for 2d flood modelling in built-up areas. *Water Resources Management*, 29(2):379–397.

Bellos, V. and Tsakiris, G. (2016). A hybrid method for flood simulation in small catchments combining hydrodynamic and hydrological techniques. *Journal of Hydrology*, 540:331–339.

Benedetti, L., Batstone, D. J., De Baets, B., Nopens, I., and Vanrolleghem, P. A. (2012). Uncertainty analysis of wwtp control strategies made feasible. *Water Quality Research Journal of Canada*, 47(1):14.

Benedetti, L., Langeveld, J., Comeau, A., Corominas, L., Daigger, G., Martin, C., Mikkelsen, P. S., Vezzaro, L., Weijers, S., and Vanrolleghem, P. A. (2013a). Modelling and monitoring of integrated urban wastewater systems: review on status and perspectives. *Water Science and Technology*, 68(6):1203–1215.

Benedetti, L., Langeveld, J., van Nieuwenhuijzen, A. F., de Jonge, J., de Klein, J., Flameling, T., Nopens, I., van Zanten, O., and Weijers, S. (2013b). Cost-effective solutions for water quality improvement in the dommel river supported by sewer–wwtp–river integrated modelling. *Water Science and Technology*, 68(5):965–973.

Bermúdez, M., Ntegeka, V., Wolfs, V., and Willems, P. (2018). Development and comparison of two fast surrogate models for urban pluvial flood simulations. *Water Resources Management*, 32(8):2801–2815.

Betancourt, R. and Alvarado, F. L. (1986). Parallel inversion of sparse matrices. *IEEE transactions on power systems*, 1(1):74–81.

Beven, K. and Binley, A. (1992). The future of distributed models: model calibration and uncertainty prediction. *Hydrological processes*, 6(3):279–298.

Beven, K. and Freer, J. (2001). Equifinality, data assimilation, and uncertainty estimation in mechanistic modelling of complex environmental systems using the glue methodology. *Journal of hydrology*, 249(1-4):11–29.

Boomen, R. v. d. and Icke, J. (2004). Tewor for duflow and sobek. homogeneization of water quality processes and coefficients (tewor voor duflow en sobek. uniformering waterkwaliteitsprocessen en coëfficiënten). Report, STOWA.

Bruni, G., Reinoso, R., van de Giesen, N. C., Clemens, F. H. L. R., and ten Veldhuis, J. A. E. (2015). On the sensitivity of urban hydrodynamic modelling to rainfall spatial and temporal resolution. *Hydrology and Earth System Sciences*, 19(2):691–709.

Bucklew, J. (2013). *Introduction to rare event simulation*. Springer Science & Business Media.

Börjeson, L., Höjer, M., Dreborg, K.-H., Ekvall, T., and Finnveden, G. (2006). Scenario types and techniques: towards a user's guide. *Futures*, 38(7):723–739.

Camargos, C., Julich, S., Houska, T., Bach, M., and Breuer, L. (2018). Effects of input data content on the uncertainty of simulating water resources. *Water*, 10(5):621.

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

Castelletti, A., Galelli, S., Restelli, M., and Soncini-Sessa, R. (2012). Data-driven dynamic emulation modelling for the optimal management of environmental systems. *Environmental Modelling & Software*, 34:30–43.

Cecinati, F., Moreno-Rodenas, A., Rico-Ramirez, M., ten Veldhuis, M.-c., and Langeveld, J. (2018). Considering rain gauge uncertainty using kriging for uncertain data. *Atmosphere*, 9(11):446.

Chaubey, I., Cotter, A. S., Costello, T. A., and Soerens, T. S. (2005). Effect of dem data resolution on swat output uncertainty. *Hydrological Processes*, 19(3):621–628.

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

Clifforde, I., Tomicic, B., and Mark, O. (1999). Integrated wastewater management-a european vision for the future. In *Proc. the Eighth International Conference on Urban Storm Drainage*, page 168.

Conti, S. and O'Hagan, A. (2010). Bayesian emulation of complex multi-output and dynamic computer models. *Journal of Statistical Planning and Inference*, 140(3):640–651.

Cosenza, A., Mannina, G., Vanrolleghem, P. A., and Neumann, M. B. (2013). Global sensitivity analysis in wastewater applications: A comprehensive comparison of different methods. *Environmental modelling & software*, 49:40–52.

Costabile, P., Costanzo, C., and Macchione, F. (2013). A storm event watershed model for surface runoff based on 2d fully dynamic wave equations. *Hydrological Processes*, 27(4):554–569.

Costabile, P., Costanzo, C., Macchione, F., and Mercogliano, P. (2012). Two-dimensional model for overland flow simulations: A case study. *European Water*, 38:13–23.

Cristiano, E., Veldhuis, M.-c. t., and Giesen, N. v. d. (2017). Spatial and temporal variability of rainfall and their effects on hydrological response in urban areas–a review. *Hydrology and Earth System Sciences*, 21(7):3859–3878.

Del Giudice, D., Albert, C., Rieckermann, J., and Reichert, P. (2016). Describing the catchment-averaged precipitation as a stochastic process improves parameter and input estimation. *Water Resources Research*, 52(4):3162–3186.

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

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

Diaz-Fierros, F., Puerta, J., Suarez, J., and Diaz-Fierros, F. (2002). Contaminant loads of csos at the wastewater treatment plant of a city in nw spain. *Urban Water*, 4(3):291–299.

Dietzel, A. and Reichert, P. (2014). Bayesian inference of a lake water quality model by emulating its posterior density. *Water Resources Research*, 50(10):7626–7647.

Ding, J. Y. (2011). A measure of watershed nonlinearity: interpreting a variable instantaneous unit hydrograph model on two vastly different sized watersheds. *Hydrology and Earth System Sciences*, 15(1):405–423.

Dominguez, D. and Gujer, W. (2006). Evolution of a wastewater treatment plant challenges traditional design concepts. *Water Res*, 40(7):1389–96.

Dooge, J. (1959). A general theory of the unit hydrograph. *Journal of Geophysical Research*, 64(2).

Dotto, C. B., Mannina, G., Kleidorfer, M., Vezzaro, L., Henrichs, M., McCarthy, D. T., Freni, G., Rauch, W., and Deletic, A. (2012). Comparison of different uncertainty techniques in urban stormwater quantity and quality modelling. *Water Res*, 46(8):2545–58.

Dotto, C. B. S., Kleidorfer, M., Deletic, A., Rauch, W., and McCarthy, D. T. (2014). Impacts of measured data uncertainty on urban stormwater models. *Journal of Hydrology*, 508:28–42.

EC (2000). Directive 2000/60. *EC of the Parliament and of the Council of the European Union*, 23(1).

Esteves, M., Faucher, X., Galle, S., and Vauclin, M. (2000). Overland flow and infiltration modelling for small plots during unsteady rain: numerical results versus observed values. *Journal of Hydrology*, 228:265–282.

Evin, G., Kavetski, D., Thyer, M., and Kuczera, G. (2013a). Pitfalls and improvements in the joint inference of heteroscedasticity and autocorrelation in hydrological model calibration. *Water Resources Research*, 49(7):4518–4524.

Evin, G., Kavetski, D., Thyer, M., and Kuczera, G. (2013b). Pitfalls and improvements in the joint inference of heteroscedasticity and autocorrelation in hydrological model calibration. *Water Resources Research*, 49(7):4518–4524.

Feinberg, J. (2015). Some improvements and applications of non-intrusive polynomial chaos expansions.

Feinberg, J. and Langtangen, H. P. (2015). Chaospy: An open source tool for designing methods of uncertainty quantification. *Journal of Computational Science*, 11:46–57.

Fiedler, F. R. and Ramirez, J. A. (2000). A numerical method for simulating discontinuous shallow flow over an infiltrating surface. *International journal for numerical methods in fluids*, 32:219–240.

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

Freni, G. and Mannina, G. (2010b). Uncertainty in water quality modelling: The applicability of variance decomposition approach. *Journal of Hydrology*, 394(3-4):324–333.

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

Freni, G., Mannina, G., and Viviani, G. (2008). Uncertainty in urban stormwater quality modelling: the effect of acceptability threshold in the glue methodology. *Water Res*, 42(8-9):2061–72.

FWR (2012). *Urban Pollution Management Manual*. Third edition edition.

Garthwaite, P. H., Kadane, J. B., and O'Hagan, A. (2005). Statistical methods for eliciting probability distributions. *Journal of the American Statistical Association*, 100(470):680–701.

Gautschi, W. (1994). Algorithm 726: Orthopol- a package of routines for generating orthogonal polynomials and gauss-type quadrature rules. *ACM Transactions on Mathematical Software (TOMS)*, 20(1):21–62.

Gernaey, K. V. and Jørgensen, S. B. (2004). Benchmarking combined biological phosphorus and nitrogen removal wastewater treatment processes. *Control Engineering Practice*, 12(3):357–373.

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

Gers, F. A., Eck, D., and Schmidhuber, J. (2002). Applying lstm to time series predictable through time-window approaches. In Tagliaferri, R. and Marinaro, M., editors, *Neural Nets WIRN Vietri-01*, pages 193–200. Springer London.

Goodman, J. and Weare, J. (2010). Ensemble samplers with affine invariance. *Communications in applied mathematics and computational science*, 5(1):65–80.

Gupta, H. V., Wagener, T., and Liu, Y. (2008). Reconciling theory with observations: elements of a diagnostic approach to model evaluation. *Hydrological Processes*, 22(18):3802–3813.

Hadigol, M. and Doostan, A. (2018). Least squares polynomial chaos expansion: A review of sampling strategies. *Computer Methods in Applied Mechanics and Engineering*, 332:382–407.

Harremöes, P. (2002). Integrated urban drainage, status and perspectives. *Water Science and Technology*, 45(3):1–10.

Hastings, W. K. (1970). Monte carlo sampling methods using markov chains and their applications. *Biometrika*, 57:97–109.

Heuvelink, G. Pebesma, J. (1999). Spatial aggregation and soil process modelling. *Geoderma*, 89(1-2):47–65.

Hochreiter, S. and Schmidhuber, J. (1997). Long short-term memory. *Neural Computation*, 9(8):1735–1780.

Hoffman, M. D. and Gelman, A. (2014). The no-u-turn sampler: adaptively setting path lengths in hamiltonian monte carlo. *Journal of Machine Learning Research*, 15(1):1593–1623.

Honti, M., Stamm, C., and Reichert, P. (2013). Integrated uncertainty assessment of discharge predictions with a statistical error model. *Water Resources Research*, 49(8):4866–4884.

House, M., Ellis, J., Herricks, E., Hvitved-Jacobsen, T., Seager, J., Lijklema, L., Aalderink, H., and Clifforde, I. (1993). Urban drainage-impacts on receiving water quality. *Water Science and Technology*, 27(12):117.

IFAK (2007). Simba, simulation of biological wastewater systems.

Jakeman, A. and Jakeman, J. (2017). Methods for uq in the water-environment-agriculture cross-sector.

Kalyanapu, A. J., Shankar, S., Pardyjak, E. R., Judi, D. R., and Burian, S. J. (2011). Assessment of gpu computational enhancement to a 2d flood model. *Environmental Modelling & Software*, 26(8):1009–1016.

Kavetski, D., Fenicia, F., Reichert, P., and Albert, C. (2018). Signature-domain calibration of hydrological models using approximate bayesian computation. theory and comparison to existing applications. *Water Resources Research*, 54(6):4059–4083.

Keupers, I. (2016). *Integrated river water quantity-quality modelling*. Doctoral thesis, KU Leuven.

Keupers, I. and Willems, P. (2015). Cso water quality generator based on calibration to wwtp influent data. In *Proceedings of the 10th International Conference on Urban Drainage Modelling. Québec*, pages 97–104.

Keupers, I. and Willems, P. (2017). Development and testing of a fast conceptual river water quality model. *Water Research*, 113:62–71.

Kleidorfer, M. (2010). *Uncertain calibration of urban drainage models*. Doctoral thesis, Innsbruck University.

Korving, H. and Clemens, F. (2005). Impact of dimension uncertainty and model calibration on sewer system assessment. *Water Science and Technology*, 52(5):35–42.

Kuczera, G. (1983). Improved parameter inference in catchment models: 1. evaluating parameter uncertainty. *Water Resources Research*, 19(5):1151–1162.

Laloy, E., Rogiers, B., Vrugt, J. A., Mallants, D., and Jacques, D. (2013). Efficient posterior exploration of a high-dimensional groundwater model from two-stage markov chain monte carlo simulation and polynomial chaos expansion. *Water Resources Research*, 49(5):2664–2682.

Laloy, E. and Vrugt, J. A. (2012). High-dimensional posterior exploration of hydrologic models using multiple-try dream(zs)and high-performance computing. *Water Resources Research*, 48(1).

Langeveld, J., Nopens, I., Schilperoort, R., Benedetti, L., de Klein, J., Amerlinck, Y., and Weijers, S. (2013a). On data requirements for calibration of integrated models for urban water systems. *Water Sci Technol*, 68(3):728–36.

Langeveld, J., Van Daal, P., Schilperoort, R., Nopens, I., Flameling, T., and Weijers, S. (2017). Empirical sewer water quality model for generating influent data for wwtp modelling. *Water*, 9(7):491.

Langeveld, J. G. (2004). *Interactions within wastewater systems.* Doctoral thesis, Delft University of Technology.

Langeveld, J. G., Benedetti, L., de Klein, J. J. M., Nopens, I., Amerlinck, Y., van Nieuwenhuijzen, A., Flameling, T., van Zanten, O., and Weijers, S. (2013b). Impact-based integrated real-time control for improvement of the dommel river water quality. *Urban Water Journal*, 10(5):312–329.

Laniak, G. F., Olchin, G., Goodall, J., Voinov, A., Hill, M., Glynn, P., Whelan, G., Geller, G., Quinn, N., Blind, M., Peckham, S., Reaney, S., Gaber, N., Kennedy, R., and Hughes, A. (2013). Integrated environmental modeling: A vision and roadmap for the future. *Environmental Modelling & Software*, 39:3–23.

Larsen, T. A., Hoffmann, S., Lüthi, C., Truffer, B., and Maurer, M. (2016). Emerging solutions to the water challenges of an urbanizing world. *Science*, 352(6288):928–933.

Leinweber, U., Hansen, J., Thomas, M., and Schmitt, T. (1999). Integrated design of sewerage system and wastewater treatment plant. In *Tagungsband des 11th European sewage and refuge symposium, im Rahmen der IFAT*, volume 99.

Liang, D. (2010). *Evaluation of runoff response to moving rainstorms.* Thesis, Marquette University Wisconsin, USA.

Ligtvoet, W., Beugelink, G., and Franken, R. (2008). Evaluation of the water framework directive in the netherlands; costs and benefits summary of pbl report 500140001. Report.

Lijklema, L., Tyson, J., and Lesouef, A. (1993). Interactions between sewers, treatment plants and receiving waters in urban areas: a summary of the interurba 1992 workshop conclusions. *Water science and technology*, 27(12):1.

Lindenschmidt, K.-E., Fleischbein, K., and Baborowski, M. (2007). Structural uncertainty in a river water quality modelling system. *Ecological Modelling*, 204(3-4):289–300.

Llopart-Mascaró, A., Farreny, R., Gabarrell, X., Rieradevall, J., Gil, A., Martínez, M., Puertas, J., Suárez, J., Río, H. d., and Paraira, M. (2015). Storm tank against combined sewer overflow: Operation strategies to minimise discharges impact to receiving waters. *Urban Water Journal*, 12(3):219–228.

Machac, D., Reichert, P., and Albert, C. (2016a). Emulation of dynamic simulators with application to hydrology. *Journal of Computational Physics*, 313:352–366.

Machac, D., Reichert, P., Rieckermann, J., and Albert, C. (2016b). Fast mechanism-based emulator of a slow urban hydrodynamic drainage simulator. *Environmental Modelling & Software*, 78:54–67.

Machac, D., Reichert, P., Rieckermann, J., Del Giudice, D., and Albert, C. (2018). Accelerating bayesian inference in hydrological modeling with a mechanistic emulator. *Environmental Modelling & Software*, 109:66–79.

Mahmood, M., Leopold, U., Torres-Matallana, J., Schutz, G., Clemens, F. H. L. R., and Bellos, V. (2018). Using a gaussian process emulator for data-driven surrogate modelling of a complex urban drainage simulator. In *AGU Fall Meeting*.

Malevergne, Y. and Sornette, D. (2003). Testing the gaussian copula hypothesis for financial assets dependences. *Quantitative Finance*, 3(4):231–250.

Mannina, G. and Viviani, G. (2010). An urban drainage stormwater quality model: Model development and uncertainty quantification. *Journal of Hydrology*, 381(3-4):248–265.

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

Matheron, G. (1971). *The theory of regionalized variables and its applications*, volume 5. Ecole nationale supérieure des mines de Paris.

Mazzetti, C. (2015). Topkapi, topographic kinematic approximation and integration.

McKay, M. D., Beckman, R. J., and Conover, W. J. (1979). Comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics*, 21(2):239–245.

McMillan, H., Krueger, T., and Freer, J. (2012). Benchmarking observational uncertainties for hydrology: rainfall, river discharge and water quality. *Hydrological Processes*, 26(26):4078–4111.

Meert, P., Pereira, F., and Willems, P. (2018). Surrogate modeling-based calibration of hydrodynamic river model parameters. *Journal of Hydro-environment Research*, 19:56–67.

Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., and Teller, E. (1953). Equation of state calculations by fast computing machines. *The journal of chemical physics*, 21(6):1087–1092.

Minshall, N. (1960). Predicting storm runoff on small experimental watersheds. *Journal of the Hydraulic Division*, 86(8):17–38.

Moens, M., Vroege, M., and Grendelman, R. (2009). Monitoring program for the river dommel through eindhoven, final report phase 1. Report, Arcadis Nederland b.v.

Moreno-Rodenas, A., Bellos, V., Langeveld, J., and Clemens, F. (2017a). Influence of river routing methods on integrated catchment water quality modelling. *Panta Rhei: Proceedings of the 10th World Congress of EWRA on Water Resources and Environment. Athens, Greece*, page 293–298.

Moreno-Rodenas, A., Cecinati, F., Langeveld, J., and Clemens, F. (2017b). Impact of spatiotemporal characteristics of rainfall inputs on integrated catchment dissolved oxygen simulations. *Water*, 9(12):926.

Moreno-Rodenas, A., Langeveld, J., and Clemens, F. (2017c). Accounting for correlation in uncertainty propagation. a copula approach for water quality modelling. In *ICUD 2017 - 14th IWA/IAHR International Conference on Urban Drainage*.

Moreno-Rodenas, A., Langeveld, J., and Clemens, F. (2018a). Parametric inference in large water quality river systems. In *International Conference of Urban Drainage Modelling*.

Moreno-Rodenas, A. M., Bellos, V., Langeveld, J. G., and Clemens, F. (2018b). A dynamic emulator for physically based flow simulators under varying rainfall and parametric conditions. *Water Res*, 142:512–527.

Muschalla, D., Schutze, M., Schroeder, K., Bach, M., Blumensaat, F., Gruber, G., Klepiszewski, K., Pabst, M., Pressl, A., Schindler, N., Solvi, A. M., and Wiese, J. (2009). The hsg procedure for modelling integrated urban wastewater systems. *Water Sci Technol*, 60(8):2065–75.

Musy, A. (1998). *Hydrologie Appliqu?Applied Hydrology)*. HGA, Bucharest.

Nanding, N., Rico-Ramirez, M. A., and Han, D. (2015). Comparison of different radar-raingauge rainfall merging techniques. *Journal of Hydroinformatics*, 17(3):422.

Nash, J. and Sutcliffe, J. (1970). River flow forecasting through conceptual models part i - a discussion of principles. *Journal of Hydrology*, 10:282–290.

Neitsch, S. L., Arnold, J. G., Kiniry, J. R., Williams, J. R., and King, K. W. (2002). Soil and water assessment tool, version 2000.

Nguyen, P., Thorstensen, A., Sorooshian, S., Hsu, K., AghaKouchak, A., Sanders, B., Koren, V., Cui, Z., and Smith, M. (2016). A high resolution coupled hydrologic–hydraulic model (hiresflood-uci) for flash flood modeling. *Journal of Hydrology*, 541:401–420.

Ochoa-Rodriguez, S., Wang, L.-P., Gires, A., Pina, R. D., Reinoso-Rondinel, R., Bruni, G., Ichiba, A., Gaitan, S., Cristiano, E., and van Assel, J. (2015). Impact of spatial and temporal resolution of rainfall inputs on urban hydrodynamic modelling outputs: A multi-catchment investigation. *Journal of Hydrology*, 531:389–407.

O'Hagan, A. (1998). Eliciting expert beliefs in substantial practical applications. *Journal of the Royal Statistical Society: Series D (The Statistician)*, 47(1):21–35.

Overeem, A., Holleman, I., and Buishand, A. (2009). Derivation of a 10-year radar-based climatology of rainfall. *Journal of Applied Meteorology and Climatology*, 48(7):1448–1463.

Owen, A. (1994). Lattice sampling revisited: Monte carlo variance of means over randomized orthogonal arrays. *The Annals of Statistics*, 22(2):930–945.

Pappenberger, F. and Beven, K. J. (2006). Ignorance is bliss: Or seven reasons not to use uncertainty analysis. *Water Resources Research*, 42(5):n/a–n/a.

Patil, A., Huard, D., and Fonnesbeck, C. J. (2010). Pymc: Bayesian stochastic modelling in python. *Journal of statistical software*, 35(4):1–81.

Pianosi, F., Beven, K., Freer, J., Hall, J. W., Rougier, J., Stephenson, D. B., and Wagener, T. (2016). Sensitivity analysis of environmental models: A systematic review with practical workflow. *Environmental Modelling & Software*, 79:214–232.

Radhakrishnan, A. T., van Lier, J., and Clemens, F. (2018). Rheological characterisation of concentrated domestic slurry. *Water research*, 141:235–250.

Radwan, M., Willems, P., and Berlamont, J. (2004a). Sensitivity and uncertainty analysis for river quality modelling. *Journal of Hydroinformatics*, 83(06.2).

Radwan, M., Willems, P., and Berlamont, J. (2004b). Sensitivity and uncertainty analysis for river quality modelling. *Journal of Hydroinformatics*, 6(2):83–99.

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

Refsgaard, J. C., van der Sluijs, J. P., Brown, J., and van der Keur, P. (2006). A framework for dealing with uncertainty due to model structure error. *Advances in Water Resources*, 29(11):1586–1597.

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

Reichert, P. (2012). Conceptual and practical aspects of quantifying uncertainty in environmental modelling and decision support. *Diss. International Environmental Modelling and Software Society (iEMSs)*.

Reichert, P. (2014). *Environmental Systems Analysis.* Zurich, Switzerland.

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 Borsuk, M. E. (2005). Does high forecast uncertainty preclude effective decision support? *Environmental Modelling & Software*, 20(8):991–1001.

Reichert, P., Langhans, S. D., Lienert, J., and Schuwirth, N. (2015). The conceptual foundation of environmental decision support. *Journal of environmental management*, 154:316–332.

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

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(5):n/a–n/a.

Rieger, L., Gillot, S., Langergraber, G., Ohtsuki, T., Shaw, A., Takacs, I., and Winkler, S. (2012). *Guidelines for using activated sludge models*. IWA publishing.

Sadegh, M. and Vrugt, J. A. (2014). Approximate bayesian computation using markov chain monte carlo simulation, dream (abc). *Water Resources Research*, 50(8):6767–6787.

Sajikumar, N. and Thandeveswara, B. (1999). A non-linear rainfall-runoff model using an artificial neural network. *Journal of Hydrology*, 216(1-2):32–55.

Sandoval, S., Vezzaro, L., and Bertrand-Krajewski, J.-L. (2018). Revisiting conceptual stormwater quality models by reconstructing virtual state variables. *Water Science and Technology*, 78(3):655–663.

Schellart, A. N., Tait, S. J., and Ashley, R. M. (2010). Towards quantification of uncertainty in predicting water quality failures in integrated catchment model studies. *Water Res*, 44(13):3893–904.

Schilperoort, R. (2011). *Monitoring as a tool for the assessment of wastewater quality dynamics*. Doctoral thesis, Technische Universiteit Delft.

Schoelzel, C. and Friederichs, P. (2008). Multivariate non-normally distributed random variables in climate research–introduction to the copula approach. *Nonlinear Processes in Geophysics*, 15(5):761–772.

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(10):n/a–n/a.

Shanahan, P., Borchardt, D., Henze, M., Rauch, W., Reichert, P., Somlyódy, L., and Vanrolleghem, P. (2001). River water quality model no. 1 (rwqm1): I. modelling approach. *Water Science and Technology*, 43(5):1–9.

Sin, G., Gernaey, K. V., Neumann, M. B., van Loosdrecht, M. C., and Gujer, W. (2011). Global sensitivity analysis in wastewater treatment plant model applications: prioritizing sources of uncertainty. *Water research*, 45(2):639–651.

Singh, J., Altinakar, M. S., and Ding, Y. (2014). Numerical modeling of rainfall-generated overland flow using nonlinear shallow-water equations. *Journal of Hydrologic Engineering*, 20(8).

Slijkhuis, K. A., Van Gelder, P., Vrijling, J., and Vrouwenvelder, A. (1999). On the lack of information in hydraulic engineering models. In *Safety and Reliability, Proceedings of ESREL '99-The Tenth European Conference on Safety and Reliability, Munich-Garching, Germany*, pages 713–718.

Sobol, I. M. (1993). Sensitivity estimates for nonlinear mathematical models. *Mathematical modelling and computational experiments*, 1(4):407–414.

Solvi, A.-M. (2006). *Modelling the sewer-treatment-urban river system in view of the EU Water Framework Directive*. Doctoral thesis, Ghent University.

Sorooshian, S. and Dracup, J. A. (1980). Stochastic parameter estimation procedures for hydrology rainfall-runoff models: Correlated and heteroscedastic error cases. *Water Resources Research*, 16(2):430–442.

Sriwastava, A. K. and Moreno-Rodenas, A. M. (2017). Report on uncertainty frameworks; quics - d.4.2 report on application of uncertainty frameworks, potential improvements. Report.

Sriwastava, A. K., Tait, S., Schellart, A., Kroll, S., Dorpe, M. V., Assel, J. V., and Shucksmith, J. (2018). Quantifying uncertainty in simulation of sewer overflow volume. *Journal of Environmental Engineering*, 144(7):04018050.

Stelling, G. S. (2012). Quadtree flood simulations with sub-grid digital elevation models. *Proceedings of the Institution of Civil Engineers - Water Management*, 165(10):567–580.

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

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):421–437.

Toni, T., Welch, D., Strelkowa, N., Ipsen, A., and Stumpf, M. P. (2009). Approximate bayesian computation scheme for parameter inference and model selection in dynamical systems. *Journal of the Royal Society Interface*, 6(31):187–202.

Torres-Matallana, J., Leopold, U., and Heuvelink, G. (2018). stupscales: An r-package for spatio-temporal uncertainty propagation across multiple scales with examples in urban water modelling. *Water*, 10(7):837.

Tsakiris, G. and Bellos, V. (2014). A numerical model for two-dimensional flood routing in complex terrains. *Water Resources Management*, 28(5):1277–1291.

Tscheikner-Gratl, F., Bellos, V., Schellart, A., Moreno-Rodenas, A., Muthusamy, M., Langeveld, J., Clemens, F., Benedetti, L., Rico-Ramirez, M. A., de Carvalho, R. F., Breuer, L., Shucksmith, J., Heuvelink, G. B., and Tait, S. (2019). Recent insights on uncertainties present in integrated catchment water quality modelling. *Water Research*, 150:368 – 379.

Tscheikner-Gratl, F., Lepot, M., Moreno-Rodenas, A. M., and Schellart, A. N. A. (2017). Quics - d6.7 a framework for the application of uncertainty analysis. Report.

Tscheikner-Gratl, F., Zeisl, P., Kinzel, C., Leimgruber, J., Ertl, T., Rauch, W., and Kleidorfer, M. (2016). Lost in calibration: why people still do not calibrate their models, and why they still should - a case study from urban drainage modelling. *Water Sci Technol*, 74(10):2337–2348.

Urich, C. and Rauch, W. (2014). Exploring critical pathways for urban water management to identify robust strategies under deep uncertainties. *Water research*, 66:374–389.

Vacondio, R., Dal Palù, A., and Mignosa, P. (2014). Gpu-enhanced finite volume shallow water solver for fast flood simulations. *Environmental Modelling & Software*, 57:60–75.

van Daal-Rombouts, P. (2017). *Performance evaluation of real time control in urban wastewater systems.* Doctoral thesis, Technische Universiteit Delft.

van Daal-Rombouts, P., Sun, S., Langeveld, J., Bertrand-Krajewski, J.-L., and Clemens, F. (2016). Design and performance evaluation of a simplified dynamic model for combined sewer overflows in pumped sewer systems. *Journal of Hydrology*, 538:609–624.

van de Beek, C. Z., Leijnse, H., Torfs, P. J. J. F., and Uijlenhoet, R. (2012). Seasonal semivariance of dutch rainfall at hourly to daily scales. *Advances in Water Resources*, 45:76–85.

Vanhooren, H., Meirlaen, J., Amerlinck, Y., Claeys, F., Vangheluwe, H., and Vanrolleghem, P. A. (2003). West: modelling biological wastewater treatment. *Journal of Hydroinformatics*, 5(1):27–50.

Vanrolleghem, P., Bertrand-Krajewski, J.-L., Brown, R., Croke, B. F. W., Kapelan, Z., Kleidorfer, M., Kuczera, G., McCarthy, D. T., Mikkelsen, P. S., Rauch, W., Refsgaard, J. C., and Deletic, A. (2011). Uncertainties in water system models, breaking down the water discipline silos. In *8th International IWA Symposium on Systems Analysis and Integrated Assessment in Water Management (WATERMATEX2011)*, pages 81–84.

Vanrolleghem, P., Borchardt, D., Henze, M., Rauch, W., Reichert, P., Shanahan, P., and Somlyódy, L. (2001). River water quality model no. 1 (rwqm1): Iii. biochemical submodel selection. *Water Science and Technology*, 43(5):31–40.

Vanrolleghem, P. A., Insel, G., Petersen, B., Sin, G., De Pauw, D., Nopens, I., Dovermann, H., Weijers, S., and Gernaey, K. (2003). A comprehensive model calibration procedure for activated sludge models. *Proceedings of the Water Environment Federation*, 2003(9):210–237.

Velasco-Forero, C. A., Sempere-Torres, D., Cassiraga, E. F., and Jaime Gómez-Hernández, J. (2009). A non-parametric automatic blending methodology to estimate rainfall fields from rain gauge and radar data. *Advances in Water Resources*, 32(7):986–1002.

Vezzaro, L., Benedetti, L., Gevaert, V., De Keyser, W., Verdonck, F., De Baets, B., Nopens, I., Cloutier, F., Vanrolleghem, P. A., and Mikkelsen, P. S. (2014a). A model library for dynamic transport and fate of micropollutants in integrated urban wastewater and stormwater systems. *Environmental Modelling & Software*, 53:98–111.

Vezzaro, L., Christensen, M. L., Thirsing, C., Grum, M., and Mikkelsen, P. S. (2014b). Water quality-based real time control of integrated urban drainage systems: A preliminary study from copenhagen, denmark. *Procedia Engineering*, 70:1707–1716.

Vrugt, J. A. and Sadegh, M. (2013). Toward diagnostic model calibration and evaluation: Approximate bayesian computation. *Water Resources Research*, 49(7):4335–4345.

Wang, S., Huang, G. H., Baetz, B. W., and Huang, W. (2015). A polynomial chaos ensemble hydrologic prediction system for efficient parameter inference and robust uncertainty assessment. *Journal of Hydrology*, 530:716–733.

Wani, O., Scheidegger, A., Carbajal, J. P., Rieckermann, J., and Blumensaat, F. (2017). Parameter estimation of hydrologic models using a likelihood function for censored and binary observations. *Water Research*, 121:290–301.

WaPUG, W. (2002). Code of practice for the hydraulic modelling of sewer systems.

Weijers, S., De Jonge, J., Van Zanten, O., Benedetti, L., Langeveld, J., Menkveld, H., and Van Nieuwenhuijzen, A. (2012). Kallisto: cost effective and integrated optimization of the urban wastewater system eindhoven. *Water Practice and Technology*, 7(2):wpt2012036.

Willems, P. (2001). A spatial rainfall generator for small spatial scales. *Journal of Hydrology*, 252:126–144.

Willems, P. (2006). Random number generator or sewer water quality model? *Water Science & Technology*, 54(6-7):387–394.

Willems, P. (2008). Quantification and relative comparison of different types of uncertainties in sewer water quality modeling. *Water Res*, 42(13):3539–51.

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

Willems, P. and Berlamont, J. (2002). Probabilistic emission and immission modelling: case-study of the combined sewer–wwtp–receiving water system at dessel (belgium). *Water Science and Technology*, 45(3):117–124.

Wynne, B. (1992). Uncertainty and environmental learning: reconceiving science and policy in the preventive paradigm. *Global environmental change*, 2(2):111–127.

Xiu, D. (2010). *Numerical methods for stochastic computations: a spectral method approach*. Princeton university press.

Xiu, D. and Karniadakis, G. E. (2002). The wiener-askey polynomial chaos for stochastic differential equations. *SIAM journal on scientific computing*, 24(2):619–644.

Yang, J., Jakeman, A., Fang, G., and Chen, X. (2018). Uncertainty analysis of a semi-distributed hydrologic model based on a gaussian process emulator. *Environmental Modelling & Software*, 101:289–300.