QUICS: Quantifying Uncertainty in Integrated Catchment Studies

D4.4 Good Practice Guidance: Incorporating Uncertainty in Integrated Catchment Studies

Lead Partner: CH2M

Revision: 31 May 2018

Report Details

Title: Good Practice Guidance: Incorporating Uncertainty in Integrated Catchment Studies

Deliverable Number (If applicable): 4.4

Author(s): Vasilis Bellos, Elliot Gill, Alma Schellart, Simon Tait

Dissemination Level: Public

Document History

Version	Date	Status	Submitted by	Checked by	Comment
1	31/01/2018	Draft	Vasilis Bellos	Simon Tait, Alma Schellart, Will Shepherd	
2	31/05/2018	Final	Vasilis Bellos	Elliot Gill, Alma Schellart	

Acronyms and Abbreviations

BOD	Biological Oxygen Demand
CFD	Computational Fluid Dynamics
COD	Chemical Oxygen Demand
CSO	Combined Sewer Overflow
DEM	Digital Elevation Model
DTM	Digital Terrain Model
GLUE	Generalized Likelihood Uncertainty Estimation
GPG	Good Practice Guidance
GUI	Graphical User Interface
ICM	Integrated Catchment Model
ICS	Integrated Catchment Study
PDE	Partial Differential Equations
R	River
RR	Rainfall-Runoff
RTC	Real Time Control
SWE	Shallow Water Equations
UD	Urban Drainage
WWTP	Waste Water Treatment Plant

Acknowledgements

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

Executive Summary

The aim of this report is to provide some guidance to practitioners seeking to understand uncertainty in modelling investigations used in urban drainage management. Urban drainage management is the activity addressed in integrated catchment studies; the interaction of rainfall, runoff, sewers, rivers and wastewater treatment

Model results are uncertain (more uncertain that generally perceived) and this should be taken into account when using models to plan improvements to urban drainage systems to achieve specified environmental outcomes. There are opportunities to use information about uncertainty to better understand trade-offs between risks and costs.

In the main part of this report a classification of available model structures in made for four types of model. Examples of uncertainty quantification studies are provided in appendices. This good practice guidance is presented as a primer for practitioners interested in applying uncertainty methods to real-world engineering problems.

CONTENTS

1 Introduction

The EU's Water Framework Directive (European Council, 2000) aims to deliver Good Ecological Status (or Good Ecological Potential) in all inland water bodies. In areas of heavy urbanization this is challenging because of wastewater treatment effluent, stormwater inputs and the operation of combined sewers overflows (CSOs) in wet weather. These inputs introduce polluting ammonium and oxygen depleting organic matter that are detrimental to aquatic life. Good Ecological Potential (or Status) is achieved if these inputs are managed to recognise river needs. Integrated Catchment Studies (ICS) are used to direct improvements to relevant parts of the urban drainage system using hydraulic and water quality models as the basis for decision making. These models are termed Integrated Catchment Models (ICM). Infrastructure improvements generally reduce the occurrence of combined sewer overflow and/or improve the quality of treated final effluent.

Models are highly uncertain, which can have two consequences in ICS. They may direct the practitioner to make insufficient improvements leaving an unacceptable environmental risk of not meeting desirable water quality standards. Or, they may direct the practitioner to over-invest in some element of the system resulting in an efficient use of resources. Worse still, a combination can occur whereby extra investment is made and environmental outcomes are still not met.

A simple example is illustrated in work done as part of the QUICS project alongside Belgian utility Aquafin (Sriwastava A., 2018). Here, modelling was used to calculate the storage required to contain the combined sewer overflow from a special storm in order to limit frequency to approximately 7 times per year. The 'normal' means of calculating storage results in a calculated storage requirement of 117 m^3, but uncertainty calculations demonstrate that there is a 50% probability that this is insufficient. It is as likely to be correct as incorrect, which presents an environmental risk. The analysis shows that risk of non-compliance can be reduced to only 10% by constructing a further 68 m^3. Uncertainty analysis allows trade-offs to be examined between cost and risk. In this case risk (of non compliance) is reduced by incurring additional costs in construction. Such trade-offs become more involved when water quality parameters further complicate analysis.

Through understanding, quantifying and working knowledgeably with model uncertainty in ICS, the practitioner can manage the aforementioned risks. Whilst there is a growing body of academic work in this field, there is little to guide the professionally working practitioner who is always constrained by time, budget and local rules, regulations and precedent governing ICS. This report serves as an introduction to the topic for practitioners and provides examples of how uncertainty can be managed in decision making. Its 'good practice guidance' aims to summarise some of the key research findings in this area, and organize them in a way that will provide the practitioner with more understanding of how uncertainty in introduced in the model predictions, and where to start looking to better understand and where possible reduce such uncertainties.

The structure of the report is composed of a main text in which the concept of Integrated Modelling is explained, and several types of common sub-models utilised for Integrated

Catchment Modelling (ICM) are presented, with their uncertainties. Finally a discussion about the applicability of a proper uncertainty analysis in these sub-models is made. The rest of the report is composed of appendices: the first one is a glossary, the next four are a brief presentation of each sub-model type within ICM (model structures, input data, model parameters, uncertainty quantification demonstrated by two examples). The final one is a brief presentation about linking several sub-models. The information in these appendices provides guidance to practitioners on the use of uncertainty analysis when applied in a variety of water modelling environments.

A guide to further reading on this topic and a *precis* of some notable publications is provided in the accompanying QUICS output D4.5.

2 Good Practice Guidance

2.1 Integrated Modelling

Integrated Catchment Modelling (ICM) is defined as the simulation of the linkage between the several sub-models, simulating processes of the water cycle (rural and urban) starting from the meteorological input (rainfall) until the final recipient, such as a river, the sea or a lake. For each process several sub-models are used.

The detail of each sub-model varies from a simplified empirical approach to a detailed numerical solver of a set of physically-based differential equations and depends on the process on which the modeller is focusing.

The models included in the context of ICM can be classified in four general categories:

1) Rainfall-Runoff (RR) models (for urban or rural catchments)

2) Urban Drainage (UD) models

3) Waste Water Treatment Plant (WWTP) models

4) River (R) models

The RR models are the hydrological models in which the water cycle is simulated, taking into account the several processes that occur in the spatio-temporal scale of the hydrological catchment (rainfall, infiltration, evaporation, snow melting, interception, overland flow, groundwater flow). The UD models are the hydraulic/hydrodynamic or CFD models with which the water flow inside the urban sewer system is simulated. They include scales from the flow inside the pipe network to flow into a manhole or a gully. The WWTP models simulate the process of contaminants removal in a WWTP. Finally, the R models are the hydraulic/hydrodynamic and water quality models with which the flow into the fluvial scale of a river is simulated.

There are several criteria in order to classify the structures of all the above types. The basic criteria is the nature of the variables output and whether are physically-based or based on empirical equations.

As mentioned previously, in the context of ICM, several types of sub-models are linked in order to simulate the propagation of the water behaviour through the water cycle. The most common types of links existing in practice are the following:

1) A RR model linked with an UD model

2) A RR model linked with a R model

3) An UD model linked with a WWTP model

4) An UD model linked with a R model

5) A WWTP model linked with a R model

6) A RR model linked with an UD, R and WWTP model.

2.2 Uncertainty quantification

Every model, since it is an abstract, simplification or interpretation of reality, cannot derive results with full accuracy. In order to be able to understand and quantify the uncertainty in model outputs, it is helpful to categorise uncertainty based on where it occurs in the model. Different academic definitions of uncertainty categorization exist, and some of these definitions are overlapping, but experience suggest that the following categorization is a helpful and pragmatic way for uncertainty quantification purposes:

1) Input data
2) Model parameters
3) Model structure

One additional uncertainty source in the context of ICM, is the linkage of the several sub-models, since they are simulating phenomena in different time and space scales.

The crucial factor for the uncertainty is to find a way in order to quantify it. For the input data uncertainty, it can be either measured or estimated. The uncertainty due to the models parameters is the most investigated source. Several methods exist, based on the Monte Carlo technique. In both cases, the quantification has the form of a probability distribution. A further extension of this technique is the Bayesian statistical inference method, in which *a priori* information of the parameters is used.

As far as the model parameters are concerned, the additional option of the sensitivity analysis can be implemented. In this kind of analysis, several techniques exist (such as screening methods) in order to define which parameter is affecting more the final output.

On the other hand, model structure uncertainty is the less studied and more difficulty to be investigated, due to its nature. Usually, implicit techniques are used to quantify model structure uncertainty, combining uncertainties due to the input data and the model parameters (Refsgaard et al. 2006).

One important factor for selecting an uncertainty analysis method, is whether observed data exist. The computational power needed for each model is one more important issue to be addressed: due to the fact that when an integrated model is implemented to real world case-studies, the required computational time is usually preventing uncertainty or sensitivity analysis, which need thousands of runs. Therefore, the only option to cope with uncertainty in these cases is to use surrogate modelling methods such as those suggested by Asher et al., 2015. These rely on simplification methods to speed simulation time that preserve key model relationships allowing uncertainties to be explored more realistically. Successful as these approaches can be, it is worth noting that evidence of compliance and presentation of results might be required (by the Regulator) to be presented in a certain model or using a specified process. In this regard, the requirements of Regulators often fall behind the capability of practitioners to present uncertainty information and discuss sensible cost-risk trade-offs. This is a significant barrier to the more widespread adoption of these techniques.

2.3 Applicability

In recent years, significant progress is made in the scientific field of uncertainty analysis, focusing on the theory, methodologies, algorithms, etc. In parallel, significant progress is observed in numerical modelling of the physical processes included in the context of ICM. Several in-house and commercial software products are available for hydrological, hydraulic, Computational Fluid Dynamics (CFD) and water quality simulations and they are widely used in practice and expected or required by Regulators. The focus of software development has been to include evermore complex and detailed representation of infrastructure.

However, there is a gulf between acamedic advances and the real-world tool used by practitioners. Several limitations prevent the widespread uptake of uncertainty analysis by practitioners. The most significant limitations are the following:

1) software automatisation

2) low-level software accessibility

3) computational cost

4) data availability

5) financing cost

1) *Software automatisation.* The most typical uncertainty analysis is based on Monte Carlo tecnhique. Using this technique, thousands or millions of simulations are required. The magnitude of this number makes non feasible the manual implementation by the modeller. Therefore, an automatisation of the software used, is needed. Since the majority of the end-users are using commercial software instead of in-house models, the modellers should find ways to automatise this software (probably with a batch file). In practice, it is observed for several software that this version is not available. In some cases is available, but the developer of the software considers this as and advanced version with an increased cost for the end-user, and in some cases this version needs no additional cost. In any case in which the automatisation is feasible, the most common limitation is the low level or negligible documentation, which has as a consecuence than only an experienced and advanced modellers with good programming skills could cope with automatisation.

2) *Low-level software accessibility*. It is observed that commercial software developers are not discussing the assumptions and the disadvantages of their product, although it is commonly accepted that since a numerical model is an abstract of reality, several approximations are made. The modeller of the end-user is assumed as a tool-user who just presses the 'magic button' of the model and receives the results. With this point of view, the way of modifying parameters used for the approximations adopted or modifying parameters which have to do with the numerical solution (e.g. time step, space step, tolerance in iteration loops) is not so clear for a common modeller, whereas sometimes, the user is advised by the documentation not to intrude in this level of the software. In the context of uncertainty analysis there might be a combination of model structure selected, input data and parametric values draw from a distribution, which has as a consequence of

the simulation (and hence the uncertainty analysis) crash. In order to cope with this software weakness, a more advanced access to the model parameters (as far as the approximations and the numerical solution is concerned) is required, and of course more experienced and familiar with numerical analysis modellers.

3) *Computational cost*. The computational cost is one of the more significant limitations for implementing an uncertainty analysis in real-world applications. In practice, even a simple software may require some minutes for one simulation, whereas more sophisticated software need hours or days for one simulation. Since uncertainty analysis is based on Monte Carlo techniques, this analysis is usually non-feasible and in the cases which is feasible, is a time consuming process. A strategy to tackle with this limitation is the use of surrogate models which are educated with results derived by the original detailed models. Needless to say that just non intrusive surrogate models can be used in practice. However, even the use of this type of surrogate models needs experienced and advanced modellers with good programming skills. Additionally, the High Performance Computing (HPC) techniques (clusters, parallel programming, etc.) can significantly improve the required simulation time, but still cannot solve the problem, whereas they increase the cost.

4) *Data availability*. It is generally accepted that there is lack of data in environmental engineering (especially in extreme conditions which is more important), for both quantity and quality variables. This lack is on one hand due to the structural weakness of measuring accurately natural phenomena in large scales, on the other hand due to the low level of investments in the field of monitoring real-world case studies. Additionally, the non-sharing policy of the available data, which unfortunately is common case in the scientific community, reproduces the limitations. Although the existence of data is not obligatory for uncertainty quantification (e.g. forward uncertainty propagation), however, for a complete uncertainty analysis in which an inference is required (e.g. Bayesian inference), observed data is needed.

5) *Financing cost.* There is a need to 'translate' all the above limitations to a common metric system, which in this case is rather the monetary system. The version of a software which supports automatisation, the development of an in-house model, the advanced level of experienced modellers with programming skills and able to cope with difficulties due to numerical analysis techniques, the computational infrastructure, the data acquisition or the monitoring of a case study, increase significantly the cost and require investments, which in the most of the cases are not efficient in the current situation. However, attention should be drawn, especially among the practitioners community, since a straightforward implementation of a software might have a consequence of over or under estimation of the design.

3 Appendix A: Glossary

Model

The term model refers to the mathematical description of a physical process. A model is based on a theoretical background and a set of equations (empirical or physically-based) which should be solved (numerically or analytically). For example, we classify urban water cycle models in four types: a) rainfall-runoff models; b) urban drainage models; c) waste water treatment plant models; d) river models. Every model has an input and an output.

Software

A software is an algorithm which solves the equations of a model, or combines several models. A software can be either commercial or in-house, open-source or closed-source, having Graphical User Interface (GUI) or not. The in-house software lot of times is referred as numerical model in the literature. An example of this term usage is: SWMM software combines a rainfall-runoff model and an urban drainage model.

Model parameters

The mathematical description of a model includes several parameters, which can be classified in three forms: a) black-box parameters; b) grey-box parameters; c) white-box parameters. The term black-box parameter means that the specific parameter has no physical meaning and should be calibrated, or derived from the literature or expert judgment. This type of parameters can be *ad hoc* parameters (local use) or global parameters (widely used). The term white-box parameter means that the specific parameter can be determined either with measurement or in a physically-based way. A grey-box parameter is in between black and white-box parameter. It should be noted that for most models, the required parameters are in the form of black or grey-box

Uncertainty

Since all models are abstractions of reality, several approximations and simplifications are made, which among others depend on the scale in which a phenomenon is examined. In modelling, there are three sources of uncertainty: input data, model parameters and model structure uncertainty. The input data (either in the form of measurements or in the form of output of another model) are characterised by uncertainties. Besides, since the majority of the required parameters for implementing a model are grey or black-box parameters, they are characterised by uncertainties as well. Finally, the level of simplification adopted, is the so called model structure uncertainty.

Monte Carlo technique

Monte Carlo technique is a numerical technique used for performing an uncertainty propagation analysis. The idea of Monte Carlo methodology is that we run thousands or millions model simulations, each time with inputs and parameters drawn from probability distributions that characterise their uncertainty.

Uncertainty propagation

Uncertainty propagation is the phenomenon that three sources of uncertainty (input data, model parameters and model structure) propagate to the model output.

Uncertainty quantification

Uncertainty quantification is the process in which the statistical characteristics of the uncertainty band of the derived results by a model (output), are calculated, such as the mean, the distribution and the confidence intervals.

Sensitivity analysis

In sensitivity analysis, the influence of model parameters in the derived results through a model structure (output), is investigated. In a complete sensitivity analysis, model parameters are ranked as far as the level of influence in the derived results is concerned.

Surrogate models

Surrogate models (or meta-models) are computationally cheap models which can be used after their training with results derived by the original detailed models, instead of more computationally expensive models. Surrogate models can be used in cases in which simulation time should be very small, such as: during the designing phase (optimising the dimensions of a structure or comparing several scenarios), in Real Time Control (RTC) schemes, uncertainty or sensitivity analysis, Decision Making schemes. There are three types of surrogate models: a) simplification or conceptualisation of the process; b) data-driven models using machine learning techniques (also known as emulators); c) intrusive model reduction, in which the modeller should intrude in the equations describe the phenomenon.

4 Appendix B: Rainfall-Runoff models

4.1 Model structures

The RR models are the hydrological models in which the water cycle is simulated, taking into account the several processes occur in the scale of the hydrological catchment (rainfall, infiltration, evaporation, snow melting, interception, ponding, wetting surface, overland flow, groundwater flow). Through this simulation, the meteorological variable which is the rainfall, is transformed to a hydraulic variable, such as the water flow. They can be characterised as generation models, in contrast with the two other types of models, which can be characterised as propagation models. There are several criteria in order to classify the structure of a hydrological model. The basic criterion is the nature of the variables output. According to this criterion, the models can be distinguished as:

1) Quantity models, in which the hydrological processes occur in the catchment are described mathematically through equations and output results are usually the water flow.

2) Quality models, in which the outputs are water quality parameters, either refer to soluble or solid parameters. Water quality parameters may be conservative, or subject to physical, chemical or biological transformations.

One more important criterion is the way with which the hydrological processes are simulated in the context of each model. According to this general criterion, the rainfall-runoff models can be classified as follows:

1) Data-driven models, in which all the physical processes are described by a set of mathematical transformation derived by input and output data, measured in each case study (*ad hoc* models).

2) Conceptual models, in which the physical processes are described by representing the catchment as a reservoir or a set of reservoirs. In this type of models, the majority of the processes are included in the context of the simulation, through empirical, semi-empirical or physically-based equations and using several parameters either obtained by the literature or black-box, *ad hoc* parameters.

3) Physically-based models, in which the physical processes are described by physically-based equations, usually integrated forms of more complex set of equations (Navier-Stokes equations, Richards equations, Transport equation, etc.).

The first category of the data-driven models can be sub-divided in the following categories:

1) Statistical models, in which the parameters used are derived by a statistical process of the input and output data.

2) Stochastic models, which are similar with the statistical models, however they incorporated stochastic terms.

3) Machine-learning models, such as the Gaussian Process or the Neural Networks models. This type of models is also mentioned as black-box models.

It should be noted that the boundaries between each category are fuzzy, whereas hybrid forms of models exist. Therefore, several similar classifications exist in the literature.

One other criterion is the way with which the spatial variability of the input data and the several required parameters spatial variability is taken into account. According to this criterion, the basic types are:

1) Lumped models

2) Distributed models

There are several hybrid forms in between these two approaches. According to how close to each type is the model, it can be classified as follows:

1) Semi-lumped models

2) Semi-distributed models

One other criterion is the time scale of the phenomenon simulated. Based on this criterion, the RR models can be classified as:

1) Continuous models, mostly used in water resources management and water quality assessments

2) Event-based models, mostly used in the design of hydraulic structures

Finally, one last criterion is the catchment's type. According to this criterion, the following types of rainfall-runoff models exist:

1) Rural scale models, in which the catchment characteristics are mainly rural

2) Urban scale models, in which the catchment characteristics are mainly urban

The quality models can be distinguished according to the nature of the output parameters as:

1) Soluble

2) Solid

One more criterion for the classification of the quality models is whether they incorporate chemical reactions or not. Therefore they can be distinguished as:

1) Physical quality models

2) Biochemical quality models

4.2 Input data

For both quantity and quality models, essential input data is the geometry representation of the catchment.

For the lumped models, each catchment is considered as one unit with specific topography characteristics.

For the distributed models, the required geometric data consist of a Digital Elevation Model (DEM) or a Digital Terrain Model (DTM).

In the semi-lumped or the semi-distributed models, the computational domain is divided in hydrological response units or sub-catchments. The level of the division indicates whether a model is semi-lumped or semi-distributed.

Except of the topography, the input is the rainfall, whereas boundary and initial conditions is required to be defined as inputs.

4.3 Model parameters

It is generally accepted that the rainfall-runoff models are considered as the more complex and present the larger variability, in comparison with the other types of models, due to the fact that they include in their context the greater amount of processes to be simulated. Therefore it is impossible to provide an entire list of the parameters incorporated in this type of models. However, for the quantity models, the parameters can be clustered in the following types:

1) Parameters which are related to the runoff of the rural or the urban catchment
2) Parameters which are related with the infiltration of the rural or the urban catchment
3) Parameters which are related with the interception of the of the urban catchment
4) Parameters which are related with the depression storage due to ponding, wetting surface and interception of the of the urban catchment
5) Parameters which are related with the evaporation of the of the rural or the urban catchment
6) Parameters which are related with the snow melting of the of the rural catchment
7) Parameters which are related with overland flow of the of the rural catchment
8) Parameters which are related with groundwater flow of the of the rural catchment

4.4 Uncertainty

4.4.1 Example B.1

Uncertainty source investigated:	Model parameters
Method used:	GLUE
Model(s) used:	Quantity model, Physically-based, Semi-distributed
	Quality model, Physically-based, Semi-distributed, solid
Case study:	Three Gorges Reservoir Region (China)

The material of this example is based on Shen et al. (2012). The case study is selected from an actual application; the Three Gorges Reservoir Region in China. The source of uncertainty which is examined is the uncertainty due to model parameters. The method used for the uncertainty analysis is the Generalized Likelihood Uncertainty Estimation (GLUE) method. In order to simulate the runoff process, the runoff curve number method was used, whereas for the infiltration phenomenon, the Green-Ampt method was implemented. For the sediment yield estimation, the Modified Universal Soil Loss Equation was used. The software used was the SWAT software, for 10,000 simulations. 20 parameters were chosen for the uncertainty analysis after a sensitivity analysis performed in the first step based on Morris screening method. The parameters which are investigated for the uncertainty to the output results are: the SCS runoff curve number for moisture condition II (ranges from −0.25 to 0.15), the base flow alpha factor (ranges from 0 to 1), the groundwater delay time (ranges from 1 to 45), the Manning's n value for overland flow (ranges from 0 to 0.5), the effective hydraulic conductivity in main channel alluvium (ranges from 0 to 150), the base flow alpha factor for bank storage (ranges from 0 to 1), the available water capacity factor (ranges from 0 to 1), the saturated hydraulic conductivity (ranges from -0.2 to 300), the soil bulk density (ranges from 0.1 to 0.6), the snowfall temperature (ranges from -5 to 5), the maximum amount of water that can be trapped in the canopy when it is fully developed (ranges from 0 to 100), the soil evaporation compensation factor (ranges from 0.01 to 1), the threshold water level in shallow aquifer for baseflow (ranges from 0 to 5000), the threshold water level in shallow aquifer (ranges from 0 to 500), the Universal Soil Loss Equation support practice factor (ranges from 0.1 to 1), the channel cover factor (ranges from 0 to 1), the channel erodibility factor ranges from 0 to 1), the channel sediment routing parameter (ranges from 0 to 0.05), the exponent parameter for calculating sediment re-entrained in channel (ranges from 1 to 1.5), the average slope length (ranges from -0.1 to 0.1). The input data consists of a real rainfall time series for the period 2004-2007. For the GLUE method, the likelihood function is the Nash-Sutcliffe coefficient, whereas for the sampling phase of the parameters, the Latin Hypercube Sampling was used. The likelihood function threshold which distinguishes the behavioral and non-behavioral set of parameters was set 0.5. The output results consist of a quantity variable (water flow) and a quality variable (sediment yield).In the following Figure B1, the 95% confidence interval is presented for each of the output variable, compared with observed and calibrated values. It seems that during the drier periods, the uncertainty band is relatively small (about 30 m^3/s), whereas in the peak periods the uncertainty band reaches more than 150 m^3/s. As far as the sediment yield is concerned, the uncertainty range is larger: during the dry periods the uncertainty band is about 50×10^4 tones, whereas in peak periods can reach about 600×10^4 tones of sediments.

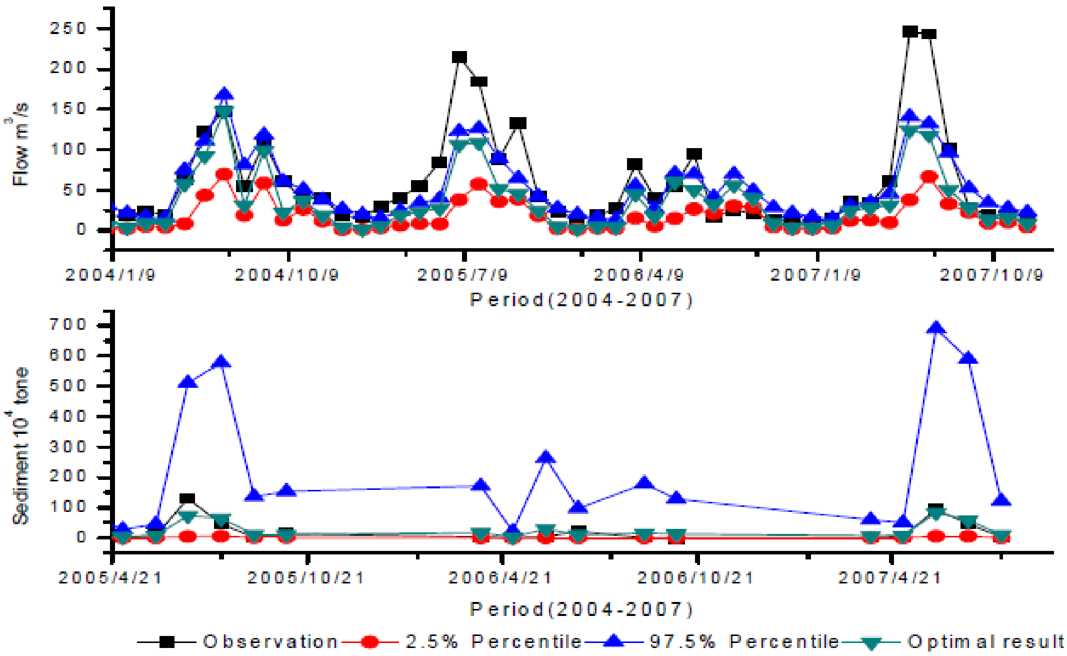

Figure B1. *95% confidence interval for the simulated flow and sediment yield using the GLUE method*

Notes: It is observed that several model parameters combinations are equifinal as far as the likelihood function value is concerned.

4.4.2 Example B.2

Uncertainty source investigated: Model parameters

Method used: GLUE

Model(s) used: Quantity model, Physically-based, Semi-distributed

Case study: Canadian Shield catchment (Canada)

The material of this example is based on Fu et al. (2015). The case study is selected from an actual application: the Canadian Shield catchment in Canada. The sources of uncertainty which are examined is the uncertainty due to model parameters. The method used for the uncertainty analysis is the GLUE method. The software used was the SWAT software, whereas two model structures were tested: SWAT and SWAT-CS (a version for Canadian catchments). For the uncertainty analysis, 12000 combinations of 22 parameters sampled by uniform distribution were implemented. The parameters are required to describe: a) interception; b) snowmelt; c) evapotranspiration; d) overland flow; e) river routing; f) infiltration; g) interflow; h) bedrock percolation; i) groundwater flow j) reservoir. The model first calibrated and validated against observed data. The input data consists of a real rainfall time series for the period 1978-1982. The output of the model is the Snow Water Equivalent and the Streamflow (quantity variables), in the sub-catchment HP4. For the GLUE method, the likelihood function is the Nash-Sutcliffe coefficient. The likelihood function threshold which distinguishes the behavioral and non-behavioral set of parameters was set at 0.45 for the Snow Water equivalent and 0.30 for the Streamflow. In the following Figure B2 and B3, the 95% confidence interval is presented for each of the output variable, compared with observed and calibrated values.

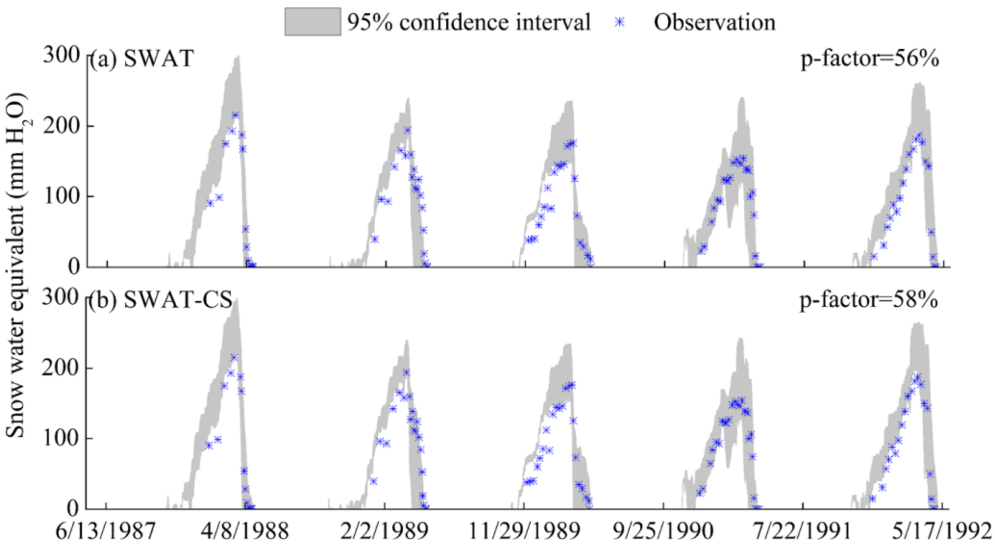

Figure B2. *95% confidence interval for simulated and observed Snow Water Equivalent using the GLUE method, for SWAT and SWAT-CS model structures.*

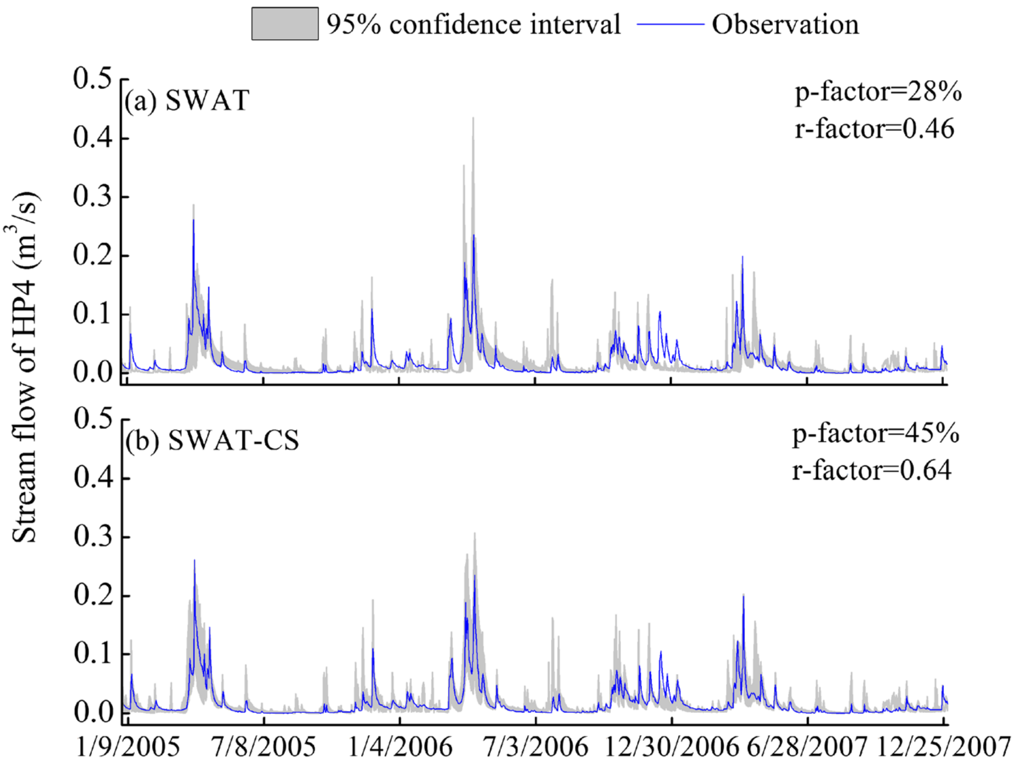

Figure B3. *95% confidence interval for simulated and observed Streamflow using the GLUE method, for SWAT and SWAT-CS model structures.*

5 Appendix C: Urban drainage models

5.1 Model structures

The UD models are models in which the flow, and sometimes the pollution, inside the urban sewer system is simulated. Several criteria exist for the classification of urban drainage models. One criterion is whether their desired output is flow quantity or water quality parameters:

1) Quantity models, in which the dynamics of the water flow is described mathematically through equations and output results are the water depth, the water velocity and the water volumetric flow rate.

2) Quality models, in which the outputs are water quality parameters, either refer to soluble or solid parameters. Water quality parameters may be conservative, or subject to physical, chemical or biological transformations.

The flow quantity models can be distinguished according to the level of detail in which the flow dynamics is described. Therefore, they can be classified as:

1) Simplified models (conceptual or empirical), in which the flow dynamics into the sewer system is described by approximations, often created by temporal or spatial averaging. The most common form of this type of models are the storage models. The kinematic wave approach or the diffusion wave approach for the flow inside the sewer can also be classified in this category.

2) Hydrodynamic models, in which the flow dynamics into the sewer system is simulated in more detail, using Partial Differential Equations (PDEs) based on the one-dimensional (1D) form of the Shallow Water Equations (1D-SWE), known also as 1D Saint-Venant equations, and which are the continuity plus the momentum equation. Only approximate numerical methods can be implemented for the solution of these PDE. The flow can be characterised either as under pressure or free surface flow.

3) Computational Fluid Dynamics (CFD) models, in which the flow dynamics into hydraulic structures such as manholes, gutters, gullies, weirs etc., is simulated. These models are based on the full form (3D) of the Navier-Stokes (NS) equations (PDE). Only approximate numerical methods can be implemented for the solution of these PDE.

The quality models can be distinguished according to the nature of the output parameters as:

1) Soluble

2) Solid

One more criterion for the classification of the quality models is whether they incorporate chemical reactions or not. Therefore the can be distinguished as:

1) Physical quality models

2) Biochemical quality models

Moreover, the physical quality models which refer to soluble parameters, can be classified according to the equation(s) solved:

1) River mixing

2) Transport

On the other hand, the physical quality models which refer to solid parameters, can be classified according to the equation(s) solved as well:

1) Sediment erosion

2) Transport

3) Deposition

Adopting the same criteria, the biochemical quality models which refer to soluble parameters can be classified:

1) Oxygen uptake

2) Degradation

As far as the SWE (moreover the Transport equation and all the PDEs) is concerned, there is no analytical solution and therefore they can be solved only using approximate numerical methods. The common numerical methods are:

1) The Finite Difference Method (FDM), which is based on Taylor series

2) The Finite Element Method (FEM), which is based on Galerkin method of weighted residuals

3) The Finite Volume Method (FVM), which is based on Divergence theorem

5.2 Input data

For quantity models, essential input data is the geometry representation of the case study. The geometry input for the simplified models consists of the representation of the urban drainage with a conceptual way.

For the hydrodynamic models consist of the representation of the sewer system (pipe characteristic length, junctions, etc.) with all the hydraulic constructions included in this system (manholes, weirs, etc.).

For the CFD models consist of the hydraulic structure under study, with a fine detail level. For the simplified models, as input data is considered the precipitation height of the catchment and the dry weather flow.

For the hydrodynamic models, as input data is considered the output of the rainfall-runoff model of the urban catchment and the dry weather flow as well.

For the CFD models and as far as the flow input (boundary condition), the input usually consists of the inflow into the structure, which is usually determined by a simplified or a hydrodynamic model, or less often by field measurements.

Apart from the geometry and inflow, initial conditions are required to be defined as inputs, usually with the form of the initial values of the variables calculated. Except of the above, the quality models use the output (hydrodynamic variables calculated) of the quantity models as an input as well.

5.3 Model parameters

Due to the combined use of several 'sub-models' into a model (for example the conceptual hydrological model for the surface runoff into a hydrodynamic model or the turbulence model used in the context of a CFD model, etc.), there are many model parameters incorporated in urban drainage models. Many are *ad hoc* parameters used for specific models and processes. In the context of this Appendix, only the parameters which are regularly used will be quoted. For the quantity models, these are the following parameters:

1) The friction coefficient, according to the selected friction model (for example Manning coefficient, Chezy coefficient, roughness height for the Darcy-Weisbach coefficient determination etc.).
2) Parameters which are related with the leak of the pipes.
3) Coefficients which are related with the several hydraulic structures of the case study.

As far as the physical quality models which refer to soluble parameters is concerned, the basic global parameter is:

1) The longitudinal dispersion coefficient.

As far as the physical quality models which refer to solid parameters is concerned, the basic global parameter is:

1) The parameters which characterise the sediments, such as the particle size distribution, the density, the fall velocity.

For the biochemical models which refer to soluble parameters correspondingly:

1) The Biological Oxygen Demand (BOD) or the Chemical Oxygen Demand (COD) as a proxy for BOD, due to the fact that the COD can be measured easier.
2) The Ammonium.
3) The Total Kjeldahl Nitrate.
4) The Phosphate.

5.4 Uncertainty

5.4.1 Example C.1

Uncertainty source investigated: Input data, Model parameters

Method used: Monte Carlo

Model(s) used: Quantity model, hydrodynamic model

Case study: Small catchment in Herent (Belgium)

The material of this example is based on Sriwastava A. (2018). The case study is selected from an actual application: a small urban catchment located in Herent (Belgium). The source of uncertainty which is examined is the input data and the model parameters, using the Monte Carlo technique. The model structure used is the InfoWorks-CS model. The input data which is examined is the weir crest level, whereas the parameters which are examined is the roughness height used in the Colebrook-White friction equation. For the Monte Carlo simulations, 1000 runs of the model were implemented. The weir crest level values randomly draw from a normal distribution, whereas the roughness height from a Log-logistic distribution and the fixed runoff coefficient from a truncated normal distribution correspondingly. The output variable is the Combined Sewer Overflow (CSO) discharge volume in one location of the catchment. In the Figure C1, it is shown the probability density of the CSO range as derived by the Monte Carlo analysis. It seems that the CSO volume ranges about 130 m^3 of water for the 90% confidence interval.

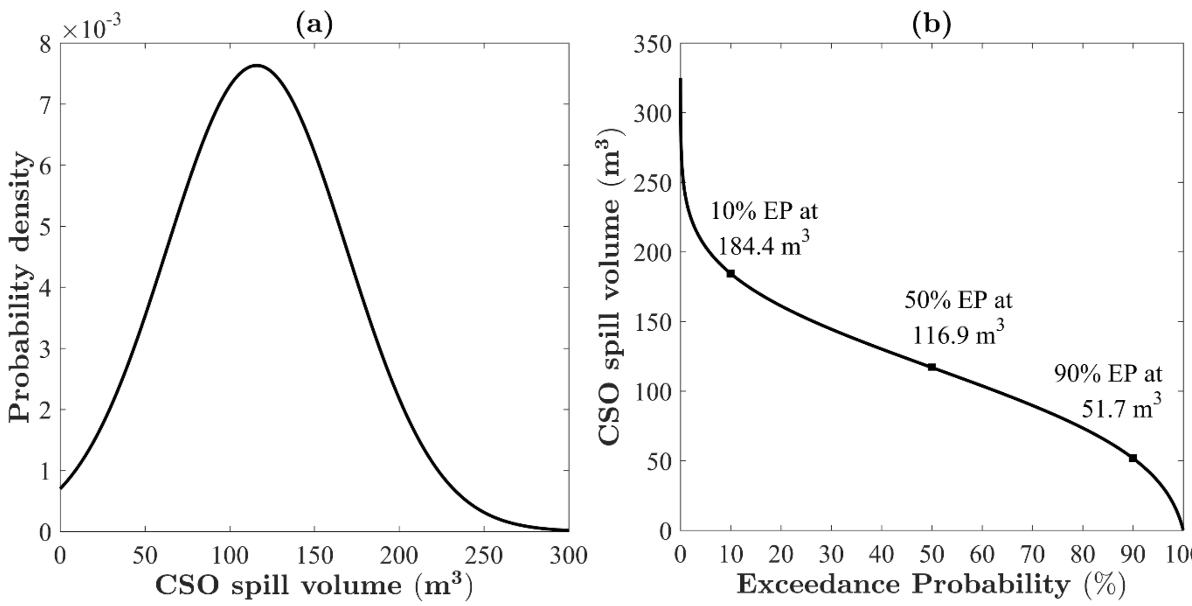

Figure C1. *CSO volume probability density (left) and exceedance probability (right) derived by Monte Carlo analysis*

Notes: -

5.4.2 Example C.2

Uncertainty source investigated: Model parameters

Method used: GLUE, SCEM-UA, AMALGAM, MICA

Model(s) used: Quantity, Quality, Simplified

Case study: Catchment in Melbourne (Australia)

The material of this example is based on Dotto et al. (2012). The case study is selected from an actual application: an urban catchment located in Melbourne (Australia). The source of uncertainty which is examined is the model parameters, using four uncertainty analysis techniques: the Generalized Likelihood Uncertainty Estimation (GLUE), The Shuffled Complex Evolution Metropolis algorithm (SCEM-UA), the multialgorithm, genetically adaptive multi-objective method (AMALGAM) and the classical Bayesian approach based on a Markov Chain Monte Carlo method and the Metropolise Hastings sampler (MICA). The model structure used is the simplified model SIMPLE KAREN. As far as the quantity part of the study, four parameters are examined: the Effective Impervious Factor, the time of concentration, the initial loss (li) and the evapotranspiration. As far as the water quality part of the study, two parameters are examined: the water quality scale coefficient (W) and the water quality shape coefficient (b). In Figure C2, the flow range derived with the four uncertainty techniques, is shown, compared with observed data. In Figure C3, the Total Suspended Solids concentration range derived with the four uncertainty techniques, is shown, compared with observed data, as well. It seems that during the more dry periods, the uncertainty interval is relatively small (about 0.1-0.2 m^3/s for all the uncertainty analysis methods), whereas in the peak periods the uncertainty band reaches about 0.7-1.0 m^3/s. As far as the Total Suspended Solids concentration range is concerned, the uncertainty band ranges from 20 to 100 mg/L for all the uncertainty analysis methods.

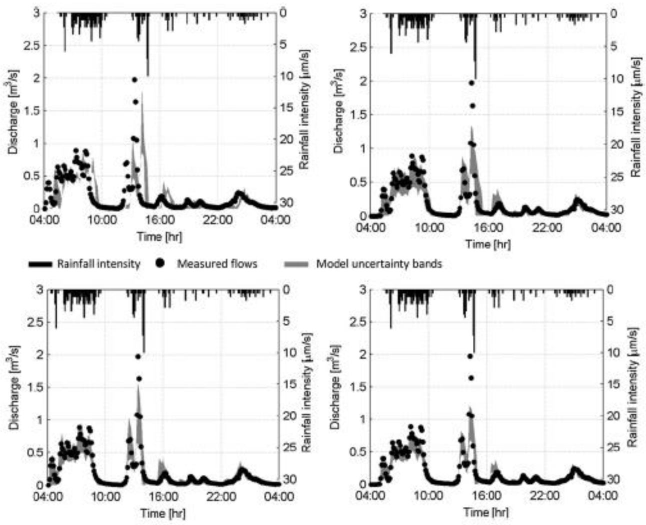

Figure C2. *Flow range derived by GLUE (top left), SCEM-UA (top right), AMALGAM (bottom left) and MICA (bottom right) during a rainfall event and compared with observed data*

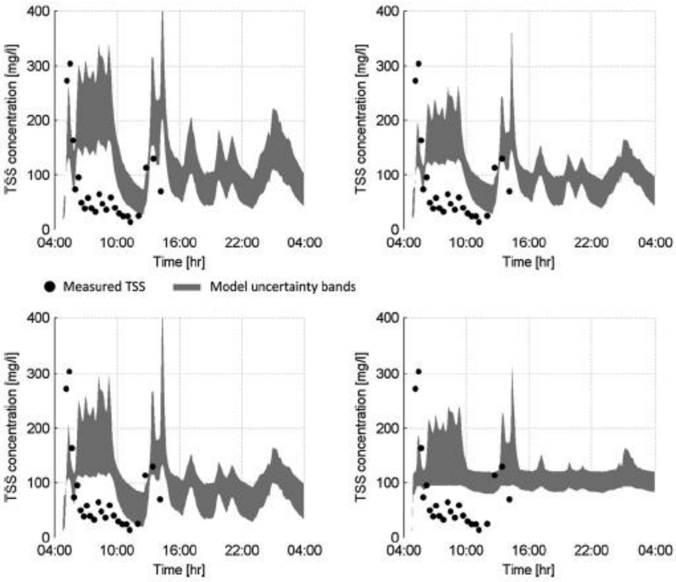

Figure C3. *Total Suspended Solids concentration derived by GLUE (top left), SCEM-UA (top right), AMALGAM (bottom left) and MICA (bottom right) during a rainfall event and compared with observed data*

6 Appendix D: Waste Water Treatment Plant models

6.1 Model structures

The WWTP models simulate the process of contaminants removal in a WWTP. Due to their nature, they are classified as coupled quantity/quality models.

In general, we can classify the WWTPs in two categories:

1) Municipal facilities

2) Industrial facilities

The industrial WWTPs are installed in big industries in order to treat their waste. Usually, municipal WWTPs have an extra influent from industrial WWTPs. This Appendix mainly focuses on municipal WWTPs, since they have great impact in ICM and besides, industrial WWTPs are highly specialised facilities, which would required a taylored study beyon the scope of this document.

In the context of the WWTPs, several biological, physical and chemical processes are occurring simultaneously. The linkage of these processes with several sub-models consist of a WWTP model. The complexity of this type of models justifies the reason which WWTP models should be considered as a separate category than UD models, although they are installed in a sewer system and hence they are linked only with UD models. The sub-models which can be linked in a WWTP model are the following:

1) Influent sub-model

2) Bio-chemical sub-model

3) Hydraulic sub-model

4) Process units sub-models

5) Control sub-models

The term process units refers to aeration, clarifiers, membranes, sludge treatment, Sequencing Batch Reactors, filters, etc. Finally, the sub-models included in a WWTP model can be classified as:

1) Empirical models

2) Physically-based models (e.g. kinetics-based models)

6.2 Input data

There are several input data required for a WWTP model. In general, the input data can be:

1) Inflow (influent) of the plant which has a quantity component (discharge) and quality component (pollutants concentration)

2) Temperature

3) Control operations

6.3 Model parameters

As long as the WWTP models are of great complexity and incorporate several sub-models, there is a great amount of parameters included in this type of models. These parameters can be classified in three big groups:

1) Operation and design parameters, which are mainly parameters related to the operation and the geometry of the plant

2) Water line parameters, which are parameters associated to the processes occuring in primary and secondary treatment

3) Sludge line parameters

6.4 Uncertainty

6.4.1 Example D.1

Uncertainty source investigated: Model parameters

Method used: Monte Carlo

Model(s) used: physically-based

Case study: Benchmark Simulation Model 2

The material of this example is based on Benedetti et al. (2012). The case study selected is the Benchmark Simulation Model 2, which is a protocol for evaluating the control strategy of WWTP model. The sources of uncertainty examined are the model parameters, using the Monte Carlo technique. The model structure used is the WEST software. The output of the model is the three evaluation criteria as determined in the Benchmark Simulation Model 2: a) Effluent Quality Index; b) Operation Cost Index; c) the period of time in which the effluent exceeds the limit of 4 mg NH4-N/L, expressed as a percentage of the whole evaluation period. First, a Global Sensitivity Analysis is performed in order to rank the parameters according to the influence on the output. In the following Figure D1, the Whisker box-plots for the three outputs, derived from the Monte Carlo simulations are presented.

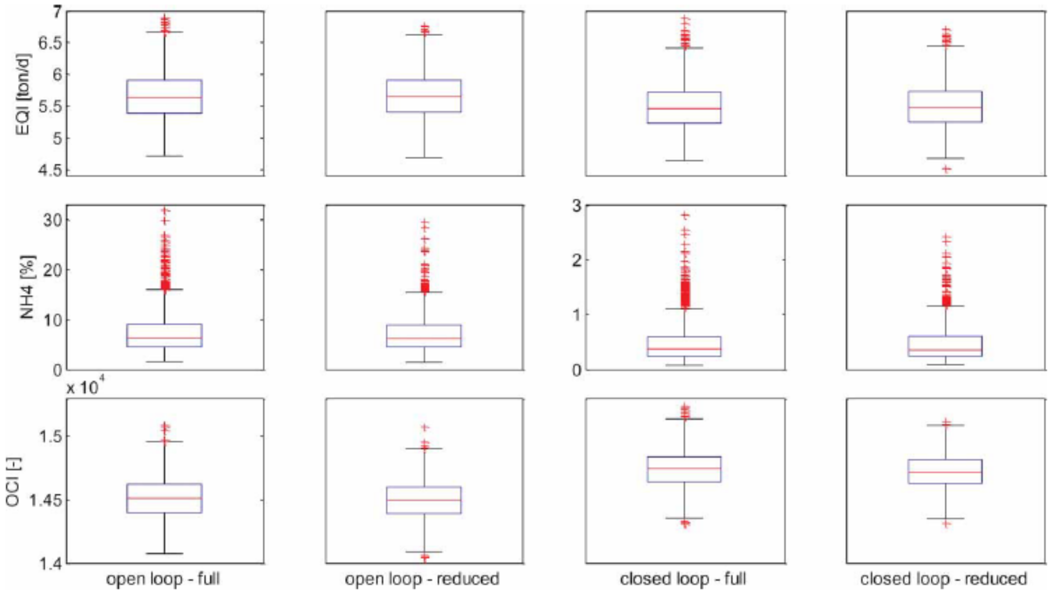

Figure D1. *Whisker box-plots of the three evaluation criteria of BSM2 examining parametric uncertainty using a reduced number of parameters*

Notes: -

6.4.2 Example D.2

Uncertainty source investigated: Model parameters

Method used: Monte Carlo

Model(s) used: physically-based

Case study: Eindhoven WWTP upgrade

The material of this example is based on Benedetti et al. (2013). The case study selected is the WWTP upgrade in Eindhoven. The sources of uncertainty which are examined are the model parameters, using the Monte Carlo technique. Specifically the parameters examined are: removal efficiency, certainty factor, peak factor. For the removal efficiency, the sample drawn by a uniform distribution, whereas for the certainty and peak factors from normal distribution. The model structure used is the WEST software. The output of the model is the NH_4 effluent. Before the uncertainty analysis, a sensitivity analysis is performed in order to select which parameters affect more the results. With this procedure, the parameters mentioned before, were selected for a further uncertainty analysis. In the following Figure D2, the uncertainty band of the simulated NH_4 effluent is presented against observed data.

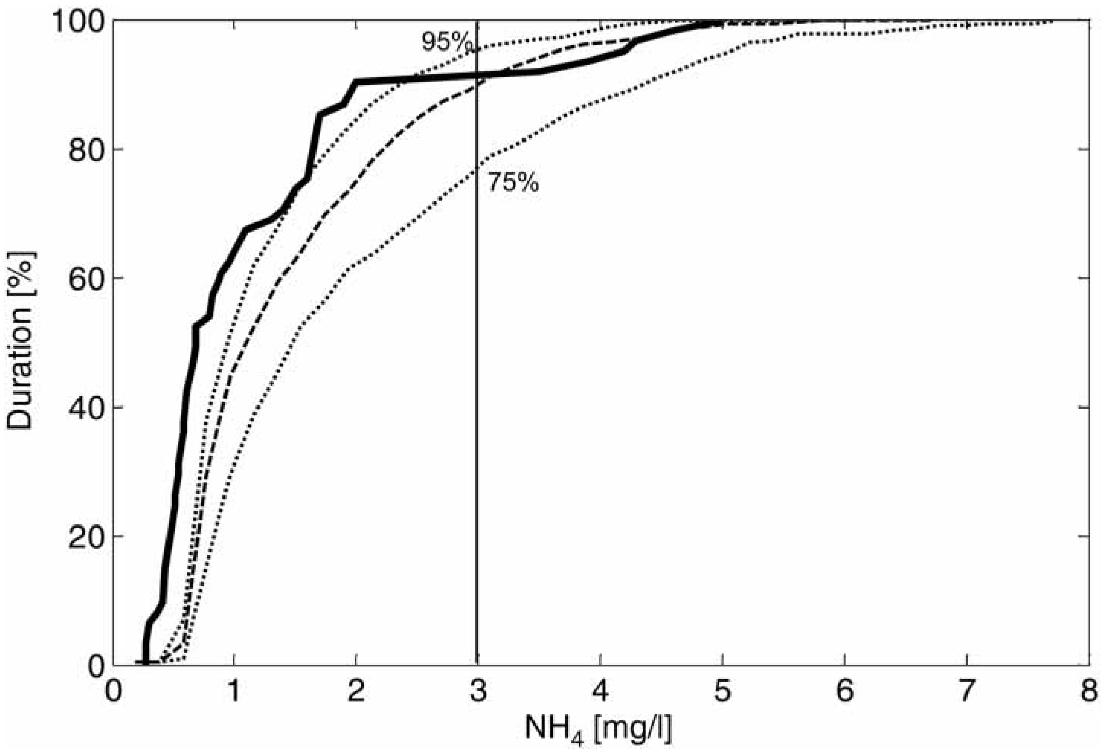

Figure D2. *Observed cumulative curves of NH_4 effluent (solid line) and simulated uncertainty band (5%, 50% and 95%) using Monte Carlo technique*

7 Appendix E: River models

7.1 Model structures

The R models are models in which the flow, and sometimes the pollution, in a river is simulated. Several criteria exist for the classification of river models. One criterion is whether their desired output is flow quantity or water quality parameters:

1) Quantity models, in which the dynamics of the water flow is described mathematically through equations and the output results are the water depth, the water velocity and the water volumetric flow rate.

2) Quality models, in which the outputs are water quality parameters. Models which simulate the dynamics of the pollutants in the water, the sediment dynamics or the morphodynamic changes in response to flows models are included in this category. Water quality parameters may be conservative, or subject to physical, chemical or biological transformations.

The flow quantity models can be distinguished according to the level of detail in which the flow dynamics is described. Therefore, they can be classified as:

1) Simplified models, in which the flow dynamics is described by approximations, often created by temporal or spatial averaging. Models such as those using Muskingum methods, kinematic and diffusion wave models, Manning equation, empirical water elevation - discharge relationships can be included in this category. Models based on integrated form of continuity and energy conservation are included in this category as well. Usually these models have analytical solutions.

2) Hydrodynamic models, in which the flow dynamics is simulated in more detail, using PDE based on the full form of the Navier-Stokes equations, which are the continuity plus the momentum equations. Only approximate numerical methods can be implemented for the solution of these PDE.

In fact, these simplified PDE are integrated forms of the Navier-Stokes equations in one or two dimensions, known also as the Shallow Water Equations (SWE) or the Saint-Venant Equations:

1) 1D models which are based on the 1D-SWE. These PDE are derived by integrating the Navier-Stokes equations in respect to the vertical and the transverse direction of the flow.

2) 1D+ models, in which the main river channel flow is simulated by solving the 1D-SWE, whereas the floodplains flow is simulated using empirical storage equations. These models are also known as pseudo or quasi-2D

3) 2D models which are based on the 2D-SWE. These PDEs are derived by integrating the Navier-Stokes equations in respect to the vertical direction of the flow.

4) 2D- models which are based on the 2D-SWE, but neglecting some terms.

The quality models can be distinguished according to the nature of the output parameters as:

1) Soluble

2) Solid

One more criterion for the classification of the quality models is whether they incorporate chemical reactions or not. Therefore the can be distinguished as:

1) Physical quality models

2) Biochemical quality models

Moreover, the physical quality models which refer to soluble parameters, can be classified according to the equation(s) solved:

1) River mixing

2) Transport

On the other hand, the physical quality models which refer to solid parameters, can be classified according to the equation(s) solved as well:

1) Sediment erosion

2) Transport

3) Deposition

Adopting the same criteria, the biochemical quality models which refer to soluble parameters can be classified:

1) Oxygen uptake

2) Degradation

As far as the SWE (moreover the Transport equation and all the PDE) is concerned, there is no analytical solution and therefore they can be solved only using approximate numerical methods. The common numerical methods are:

1) The Finite Difference Method (FDM), which is based on Taylor series

2) The Finite Element Method (FEM), which is based on Galerkin method of weighted residuals

3) The Finite Volume Method (FVM), which is based on Divergence theorem

Finally, there are several sub-models included in the river models, which characterise the structure of each model. Some of the most common sub-models are:

1) Turbulence models which are included as terms in the momentum equations.

2) Wet/dry models, with which the dry cells/elements are distinguished from the corresponding wet.

3) Urban environment models (the way with which the several obstacles exist in an urban environment, such as the buildings, are represented).

4) Infiltration models.

7.2 Input data

For quantity models, essential input data is the geometry representation of the river. In general, the geometry input for the 1D models consists of 'parallel' cross-sections, whereas for 2D models geometric data from a Digital Elevation Model (DEM) or a Digital Terrain Model (DTM) is used. Apart from the geometry, boundary and initial conditions is required to be defined as inputs.

For the quantity models the most common form of the upstream boundary conditions is the constant flow or flow time series (hydrograph) as input, which consists of output from a rural RR model, as well as potentially additional output from UD models. It should be noted that there are cases (such as when the final recipient is the sea) in which the input takes the form of the water level in respect to time should be defined as downstream boundary conditions. Apart from the boundary conditions, the initial condition of the river flow should also be defined, usually with the form of the initial values of the variables calculated.

As far as the water quality models is concerned, the input of the upstream boundary conditions has usually the form of the constant pollutant load or pollutant concentration in respect to time. Apart from this, the quality models use the output (hydrodynamic variables calculated) of the quantity models as an input as well. In the morphodynamic models, the granulometry of the river bottom is an input data as well.

7.3 Model parameters

Due to the complexity of the simulated phenomena there are many model parameters incorporated in river models. Many are *ad hoc* parameters used for specific models and processes. In the context of this Appendix, only the parameters which are regularly used will be quoted. For the quantity models, these are the following parameters:

1) The friction coefficient, according to the selected friction model (for example Manning coefficient, Chezy coefficient, roughness height for the Darcy-Weisbach coefficient determination etc.).

2) The grid/mesh resolution. It is noted that the grid is referred when the computational domain is represented by square shape cells (implementable with all the numerical methods) and mesh where the computational domain is represented by different than square shape cells, such as triangles or quadrilaterals (implementable just for the FEM and the FVM).

3) The energy slope required in the cases in which the normal depth boundary condition is set as an upstream boundary. It is noted that this slope is not equal to the surface slope.

Except the above, there are several common parameters which can be incorporated depending on the type of quantity model:

1) Parameters which are required in order to preserve the numerical stability.

2) Parameters which are used for the wet/dry modelling.

3) Parameters which are used in order to represent obstacles, such as buildings.

4) Parameters which are used in order to simulate the infiltration phenomenon.

As far as the physical quality models is concerned, the basic global parameters are:

1) The longitudinal dispersion coefficient for river mixing models (vertical and across the stream are much less common, mainly for 2D and 3D approaches, mostly interesting for lakes or very large rivers/estuaries)

2) The bed shear stress (e.g. the threshold of erosion models)

3) The shear velocity, as a measure of turbulence for both mixing and sediment transport models

4) The parameters which characterise the sediments, such as the particle size distribution, the density, the fall velocity.

For the biochemical models correspondingly:

1) The Biological Oxygen Demand (BOD) or the Chemical Oxygen Demand (COD) as a proxy for BOD, due to the fact that the COD can be measured easier

2) The Ammonium.

3) The Total Kjeldahl Nitrate.

4) The Phosphate.

7.4 Uncertainty

7.4.1 Example E.1

Uncertainty source investigated: Input data, model parameters

Method used: Monte Carlo

Model(s) used: Quantity models, Hydrodynamic, 1D; 2D-; 2D, FDM

Case study: Rafina stream (Greece)

The material of this example is based on Dimitriadis et al. (2016). The case study is selected from an actual application: Rafina stream, which is located north-east of Athens (Greece). The sources of uncertainty which are examined are the input data and the model parameters, using the Monte Carlo technique. The model structures used are the HEC-RAS software (1D model), the LISFLOOD software (2D-) and the FLO-2D software (2D) for 300 simulations each. The input data consists of a steady flow, for which the Monte Carlo simulations randomly draw from a uniform distribution with range 250 m^3/s to 1000 m^3/s. The examined parameter is the Manning's roughness coefficient, for which the Monte Carlo simulations randomly draw from a uniform distribution with range 0.01 $s/m^{1/3}$ to 0.1 $s/m^{1/3}$. The output variables are the water depths in the upstream and the downstream cross-sections correspondingly. In the following Figure E1, the empirical probability functions derived from the Monte Carlo simulations are presented. It is found that the distributions of the water depths approximate a normal distribution. It seems that the water depths range about 4 m using the HEC-RAS model, about 8-12 m using the LISFLOOD-FP model and about 2 m using the FLO-2D model, for the 95% confidence interval.

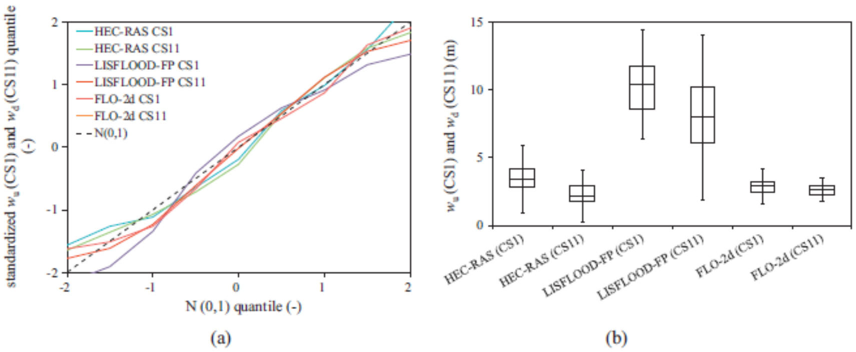

Figure E1. *Simulated water depths (w_u, upstream and w_d, downstream cross-section) derived from several model structures: a) probability functions and b) box plot*

Notes: It is noted that due to the computational burden for the 2D- and the 2D model (the magnitude for each simulation is hours), the total number of simulations is relatively low for a typical Monte Carlo based uncertainty analysis. It is also noted that the 2D- and the 2D model did not have the same grid resolution. (5 x 5 m for the 2D- model and 50 x 50 m for the 2D model).

7.4.2 Example E.2

Uncertainty source investigated: Input data, model parameters

Method used: Monte Carlo, GLUE

Model(s) used: Quantity, simplified, 1D

Case study: Tallahala Creek (USA)

The material of this example is based on Brandimarte and Woldeyes (2013). The case study is selected from an actual application: Tallahala Creek, near Waldrup, Mississippi, USA. The sources of uncertainty which are examined are the input data and the model parameters, using the Monte Carlo technique and the GLUE method. The combined uncertainty due to these two sources is also quantified as well. The model structures used is the 1D HEC-RAS software, using the steady flow mode. The input data consists of a steady flow, for which the 100 Monte Carlo simulations randomly draw from a normal distribution with a mean value equal to the flood peak with a return period T=100 years, estimated applying the Extreme Value distribution type I and a standard deviation equal to the standard error of estimate for the Extreme Value distribution type I distribution, evaluated using the Kite formula (Chow, 1988). The range of the water surface elevation due to the input data uncertainty in 12 stations is shown in the Figure E.2. The examined parameter is the Manning's roughness coefficient for the main channel and for the floodplains, using the GLUE method: 48 behavioral models were selected, using as a criterion that the Mean Absolute Error compared with the corresponding observed data should be less than 0.5 m. The range of the water surface elevation due to the model parameters uncertainty in 12 stations is shown in the Figure E.3. Finally, each of the 48 behavioral scenarios was run using as an input the 100 runs derived previously, in order to quantify the combined uncertainty due to the input data and the model parameters. The range of the water surface elevation due to the combined effect of input data and model parameters uncertainty in 12 stations is shown in the Figure E.4. It seems that the water surface elevation ranges from about 0.5 m in the worst case, to 0.2 m in the best case, for the 95% confidence interval, investigating the input data uncertainty. Investigating the parameters data uncertainty, these intervals are ranging from about 1.0 m to 0.3 m, whereas investigating both input data and parameter uncertainty, these intervals are ranging from about 1.5 m to 1.0 m correspondingly.

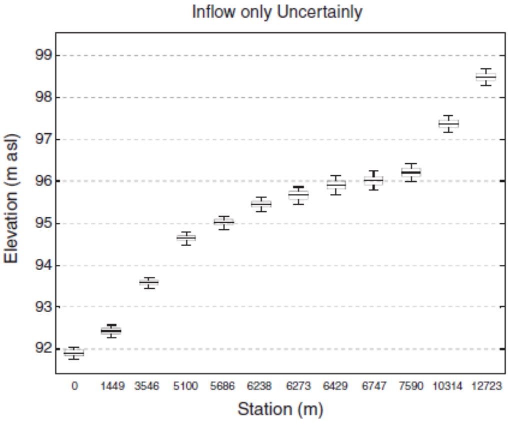

Figure E2. *Range of water surface elevation (Box plots with minimum and maximum values, first, second and third quartiles), due to the input data uncertainty*

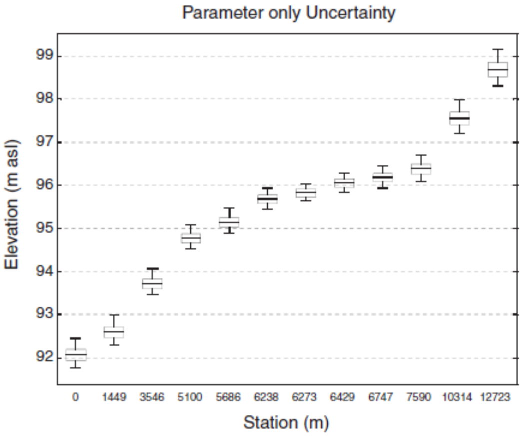

Figure E3. *Range of water surface elevation (Box plots with minimum and maximum values, first, second and third quartiles), due to the model parameters uncertainty*

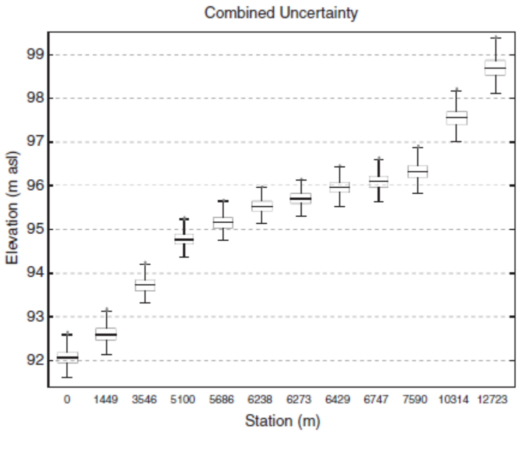

Figure E4. *Range of water surface elevation (Box plots with minimum and maximum values, first, second and third quartiles), due to the combined effect of input data and model parameters uncertainty*

Notes:-

8 Appendix F: Linkage

8.1 Linking sub-models

In the context of ICM, several sub-models which describe different processes of the water cycle (urban or rural) are linked in a way that the output of one model is the input for the other. The potential linkages included in the ICM are six:

1) *RR ~ UD*. A RR model has as an output quantity or quality variables in respect to time (e.g. flow hydrograph or pollutant concentration). These time series can be the input for UD models which simulate quantity and quality variables in respect to time, in the sewer system of an urban configuration.

2) *RR ~ R*. A RR model has as an output quantity or quality variables in respect to time (e.g. flow hydrograph or pollutant concentration). These time series can be the input for R models which simulate quantity and quality variables in respect to time, in the fluvial scale of a river located in urban or rural environments.

3) *UD ~ WWTP*. A UD model has as an output quantity or quality variables in respect to time (e.g. flow hydrograph or pollutant concentration). These time series can be the input (influent) for WWTP models which simulate quantity and quality variables in respect to time.

4) *WWTP ~ UD*. A WWTP model has as an output quantity or quality variables (effluent) in respect to time. These time series can be the input for UD models which simulate quantity and quality variables in respect to time.

5) *UD ~ R*. A UD model has as an output quantity or quality variables in respect to time (e.g. flow hydrograph or pollutant concentration). These time series can be the input for R models which simulate quantity and quality variables in respect to time, in the fluvial scale of a river located in urban environments.

6) *WWTP ~ R*. A WWTP model has as an output quantity or quality variables (effluent) in respect to time. These time series can be the input for R models which simulate quantity and quality variables in respect to time, in the case of an advanced waste water treatment level.

8.2 Uncertainty

The processes included in the internal structure of two linked sub-models usually are characterised by different spatial and temporal scales. This has as a consequence that several techniques should be followed for downscaling (disaggregation) or upscaling (aggregation) in space or time or both of them, for linking the several sub-models. Although in practice the downscaling or upscaling is made heuristically (with simple interpolation or extrapolation techniques), a consistent modelling approach requires more advanced techniques, such as statistical methods (e.g. Kriging methods). Especially for the spatial downscaling, the methods are also known as geostatistical methods. However, these methods are adding uncertainties to the output of the models.

The most common links in ICM which require downscaling and upscaling processes are the following:

1) *Input data - RR*. Rainfall records (input data) are in a larger temporal and spatial scale when they are obtained by radars (spatial scale of km^2 and temporal scale of min), than the scale of the common RR models. Therefore downscaling process should be followed.

2) *Input data - RR*. Rainfall records (input data) are in a smaller temporal and spatial scale when they are obtained by a network of gauges and then should be upscaled in order to be used by RR models.

3) *RR - UD or R*. The RR models are in different scale (catchment) than the UD or R models, whereas the output of RR models is in a larger temporal and spatial scale than a common UD or R model (spatial scale in m and temporal scale in s).

4) *UD - R*. Although the phenomena included in the water cycle in the urban configuration are more intense and rapid than the corresponding in rural configuration, the scales are the same (spatial scale in m and temporal scale in s). However, uncertainty due to downscaling should be considered in the cases in which a more rough, empirical UD model is linked with a more detailed R model.

8.2.1 Example F.1

Uncertainty source investigated: Input data

Method used: Monte Carlo

Model(s) used: -

Case study: Bradford (UK)

The material of this example is based on Muthusamy et al. (2017). The case study is selected from an actual application: a catchment located at Bradford, West Yorkshire, UK. Rainfall data collected from a cluster in an urban catchment are used in combination with spatial stochastic simulation to obtain optimal predictions of the spatially averaged rainfall intensity at any point in time within the urban catchment. The uncertainty in the prediction of catchment average rainfall intensity is obtained for multiple combinations of intensity ranges and temporal averaging intervals. Scarcity of measurement points is dealt with by pooling sample variograms of repeated rainfall measurements with similar characteristics. Normality of rainfall data is achieved through the use of normal score transformation. Geostatistical models in the form of variograms are derived for transformed rainfall intensity. Next, spatial stochastic simulation is applied to produce realisations of rainfall fields. These realisations in transformed space are first back-transformed and next spatially aggregated to derive a random sample of the spatially averaged rainfall intensity. Results show that the prediction uncertainty comes mainly from two sources: spatial variability of rainfall and measurement error. At smaller temporal averaging intervals both these effects are high, resulting in a relatively high uncertainty in prediction. With longer temporal averaging intervals the uncertainty becomes lower due to stronger spatial correlation of rainfall data and relatively smaller measurement error. Results also show that the measurement error increases with decreasing rainfall intensity resulting in a higher uncertainty at lower intensities. In Figure F1, the rainfall intensity with 95% uncertainty band is shown for a rainfall event.

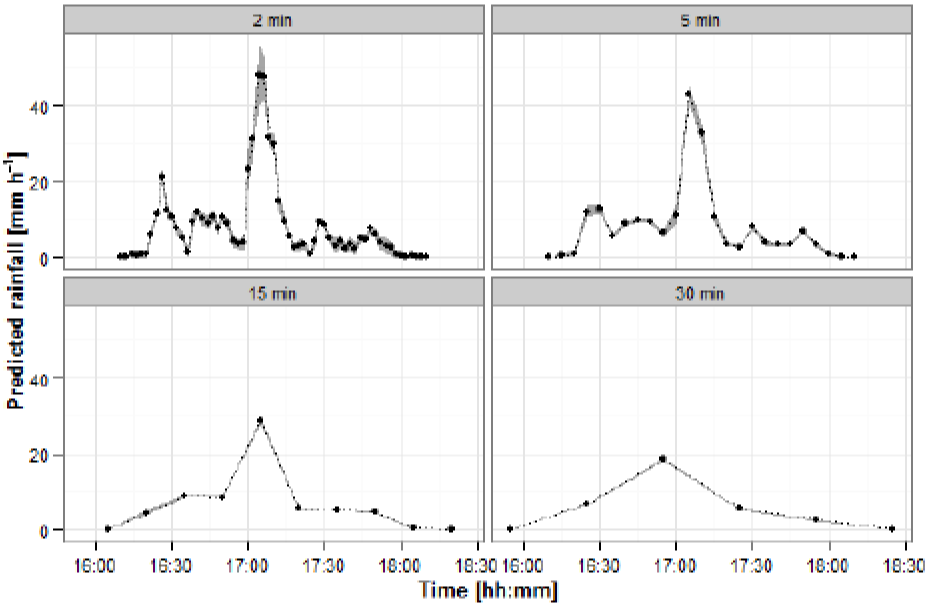

Figure F1. *Predictions of Areal Average Rainfall Intensity with 95% prediction intervals for a rainfall event, for different averaging intervals.*

Notes:-

9 References

Asher, M.J., Croke, B.F.W., Jakeman, A.J., Peeters, L.J.M., 2015. A review of surrogate models and their application to groundwater modeling. Water Resources Research, 5957-5973.

Benedetti, L., Batstone, J.D., De Baets, B., Nopens, I., Vanrolleghem, A.P., 2012. Uncertainty Analysis of WWTP control strategies made feasible. Water Quality Research Journal of Canada, 47(1), 14-29.

Benedetti, L., Belia, E., Cierkens, K., Flameling, T., De Baets, B., Nopens, I., Weijers, S., 2013. The incorporation of variability and uncertainty evaluations in WWTP design by means of stochastic dynamic modeling: the case of the Eindhoven WWTP upgrade. Water Science and Technology, 67(8), 1841-1850.

Brandimarte, L. and Woldeyes, M.K., 2013. Uncertainty in the estimation of backwater effects at bridge crossings. Hydrological Processes, 27, 1292-1300.

Chow, V.T., Maidment, D.R., Mays, L.W., 1988. Applied Hydrology, McGraw-Hill.

Deng, Z.Q., Singh, V.P., Bengtsson, L., 2001. Longitudinal dispersion coefficient in straight rivers. Journal of Hydraulic Engineering, 127, 919-927.

Dimitriadis, P., Tegos, A., Oikonomou, A., Pagana, V., Koukouvinos, A., Mamassis, N., Koutsoyiannis, D., Efstratiadis, A., 2016. Comparative evaluation of 1D and quasi-2D hydraulic models based on benchmark and real-world applications for uncertainty assessment in flood mapping. Journal of Hydrology, 534, 478-492.

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

European Council, 2000. Directive 2000/60/EC of the European Parliament and of the Council on establishing a framework for Community action in the field of water policy.

Fu, C., James, L.A., Yao, H., 2015. Investigations of uncertainty in SWAT hydrological simulations: a case study of a Canadian Shield catchment. Hydrological Processes, 29, 4000-4017.

Muthusamy, M., Schellart, A., Tait, S., Heuvelink, B.M.G., 2017. Geostatistical upscaling of rain gauge data to support uncertainty analysis of lumped urban hydrological models. Hydrology and Earth System Science, 1077-1091.

Refsgaard, C.J., van der Sluijs, P.J., Brown, J., van der Keur, P. A framework of dealing with uncertainty due to model structure error. Advances in Water Resources, 29, 1586-1597.

Shen, Z.Y., Chen, L., Chen, T., 2012. Analysis of parameter uncertainty in hydrological and sediment modelling using GLUE method: a case study of SWAT model applied to Three Gorges Reservoir Region, China. Hydrology and Earth System Science, 16, 121-132.

Sriwastava, A., Tait, S., Schellart, A., Kroll, S., Van Dorpe, M., Van Assel, J., Shucksmith, J., 2018. Quantifying uncertainty in the simulation of sewer overflow volume. Journal of Environmental Engineering, accepted for publication.