 [ISBN: 978-2-9105-4510-5]

Uncertainty propagation in urban hydrology water quality modelling

J.A. Torres-Matallana[*1,2], **U. Leopold**[1], **G.B.M. Heuvelink**[2]

[1]Luxembourg Institute of Science and Technology, Luxembourg
[2]Wageningen University, Netherlands

*Corresponding author: arturo.torres@list.lu

Abstract

Uncertainty is often ignored in urban hydrology modelling. Engineering practice typically ignores uncertainties and uncertainty propagation. This can have large impacts, such as the wrong dimensioning of urban drainage systems and the inaccurate estimation of pollution in the environment caused by combined sewer overflows. This paper presents an uncertainty propagation analysis in urban hydrology modelling. The case study was the Haute-Sûre catchment in Luxembourg for one yearly time series measured in 2010, and 10 individual rainfall events measured in 2011. The selection of model input variables for uncertainty quantification was based on their level of uncertainty and model sensitivity. Probability distribution functions were defined to represent the uncertainty of the input variables. We applied a Monte Carlo technique using a simplified model, EmiStatR, which simulates the volume and substance flows in urban drainage systems. We focus in loads and concentrations of chemical oxygen demand and ammonium, which are important variables for wastewater and surface water quality management.

Keywords: uncertainty propagation, urban hydrology modelling, EmiStatR, Monte Carlo

I INTRODUCTION

Uncertainty is often ignored in urban hydrology modelling (Mitchell et al., 2007; Bach et al., 2014). Engineering practice typically ignores uncertainties and uncertainty propagation, among others because of lack of user-friendly implementations (Schellart et al., 2010). This can have large impacts, such as the wrong dimensioning of urban drainage systems and the inaccurate estimation of pollution in the environment caused by combined sewer overflows (CSOs).

Six approaches to address uncertainty in the context of urban water systems are identified (Walker et al., 2003; Refsgaard et al., 2007; van Keur et al., 2008; Bach et al., 2014): 1) determinism; 2) statistical uncertainty; 3) scenario uncertainty; 4) qualitative uncertainty; 5) recognised ignorance; and 6) total (unrecognised) ignorance. Following van Keur et al. (2008), *determinism* applies when we have knowledge with absolute certainty about the system under analysis. The *statistical* approach is useful when it is possible to describe in statistical terms the uncertainty. This occurs when errors and uncertainties can be quantified by probability distribution functions (pdfs). The *scenario* approach, in contrast, applies when quantitative probabilities cannot be determined. It is used when possible outcomes can be listed without pretending that the list is exhaustive and without attaching probabilities to each possible outcome (Brown, 2004). The *qualitative* approach is used when *uncertainty cannot be characterised statistically, and not all outcomes are known* (Brown, 2004). Recognised ignorance occurs in a situation of awareness of lack of knowledge (van Keur et al., 2008). Finally, *total ignorance* is the state of *complete lack of awareness about possible outcomes* (van Keur et al., 2008). Amongst the approaches described above, in this paper we will limit ourselves to the statical approach to characterise and propagate uncertainties.

Sources of uncertainty in the context of the evaluation and design of urban water infrastructure are identified by Neumann (2007): model structure, model parameters, errors in input data, and in numerical and computational procedures. Our focus is on model input as a primary source of uncertainty.

Uncertainty propagation analysis can be done analytically by means of the Taylor series method, or numerically by means of Monte Carlo simulation. This paper presents an uncertainty propagation analysis in urban hydrology modelling. For this we apply a Monte Carlo analysis using the EmiStatR model, which simulates the volume and substance flows in urban drainage systems. We use a case study from the Haute-Sûre catchment in Luxembourg for one yearly time series measured in 2010, and 10 individual rainfall events measured in 2011. We focus on substances as chemical oxygen demand (COD) and ammonium (NH_4), which are important variables for wastewater and surface water quality management.

II MATERIALS AND METHODS

2.1 The EmiStatR model

The EmiStat model (Klepiszewski and Seiffert, 2013) is a Microsoft Excel based model that was developed for the Luxembourgish water authority as a tool for the evaluation of planning scenarios of sewer systems. EmiStatR is the R implementation of EmiStat. It provides a fast estimation of combined wastewater emissions by imposing a strongly simplified representation of the real-world system. It can aid the planning and design of hydraulic structures and pollutant handling, without the requirement of extensive simulation tools. The EmiStatR model includes six main components to simulate combined sewage discharges of a catchment (Figure 1).

The sewer system under investigation includes tank structures to store first flush pollutant peaks. After filling of the storage volume a combined sewage overflow structure discharges subsequent volume and pollutant inflows exceeding the structures outflow from the wastewater treatment plant (WWTP) to the receiving water.

In EmiStatR a simple volume balance taking into account inflow volume, present storage capacity and outflow to the WWTP is implemented to simulate the volume in the tank structure. In case of an overflow the pollutant concentrations in the CSO are equivalent to the combined sewage inflow concentrations of the structure.

The pollutants typically taken into account are total COD and NH_4. The variable COD is the standard used in the framework of the dimensioning of CSO structures in Luxembourg. NH_4 represents a diluted substance which can have a significant impact on surface water quality due to possible transformation to ammonia (NH_3).

At the CSO tank structure a simple volume balancing takes place: 1) substance and volume flows are stored and discharged to the WWTP if the storage volume is not completely filled up; 2) if the storage volume is completely filled up the proportion of the volume inflow which is not discharged to the WWTP goes to the CSO.

2.2 Sewer systems of the Haute-Sûre sub-catchments

The study area is a sub-catchment of the Haute-Sûre catchment in the north-west of Luxembourg. The combined sewer system of the sub-catchment drains eight villages: Büderscheid

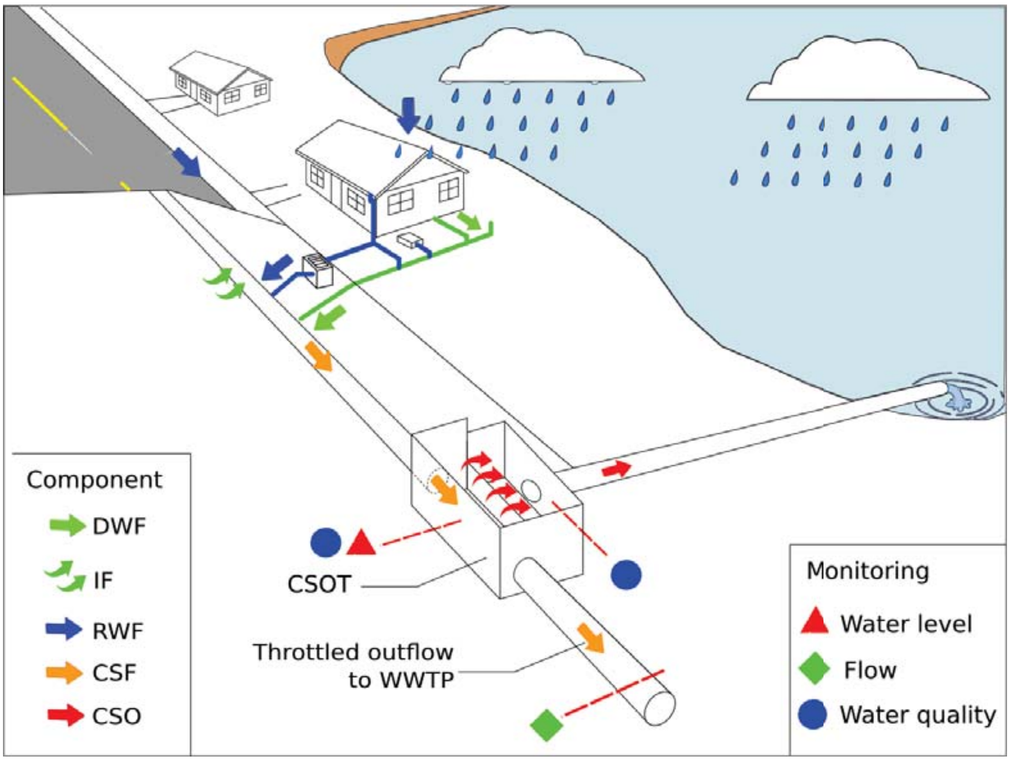

Figure 1: Main components of the EmiStatR model: 1) Dry Weather Flow (DWF) including Infiltration Flow (IF); 2) Pollution of DWF; 3) Rain Weather Flow (RWF); 4) Pollution of RWF; 5) Combined Sewage Flow (CSF) and pollution; and 6) Combined Sewer Overflow (CSO) and pollution. CSOT = CSO tank (Background adapted from: Sanitary-District (2015))

(BUD), Dahl (DAH), Eschdorf/Ost (ESO), Goesdorf (GOE), Heiderscheid (HEI), Kaundorf (KAU), Nocher (NOC) and Nocher-Route (NOR). The local sewer system downstream each village has a CSO tank to store pollutant peaks in the first flush of combined sewage flows. Table 2 shows the general characteristics of each CSO tank for each village. Figure 2 depicts their locations and the delineation of the catchment. The main land use types in the villages are residential, smaller industries and farms. Outside of the villages forest as well as agricultural arable and grassland are the dominating land uses. The receiving water bodies at CSO structures are tributaries of the river Sûre.

The general input data of the EmiStatR model for simulating the eight CSO structures is presented in Tables 1 and 2.

2.3 Model input uncertainty assessment

Following recommendations from Nol et al. (2010), not all model inputs were taken into account in the Monte Carlo uncertainty propagation analysis. Only those inputs that have a large uncertainty and to which the model is sensitive were included. The selection of model input for uncertainty quantification was based on a list of all model inputs, their level of uncertainty (low or high) and the level of model sensitivity (low or high). The level of uncertainty of the inputs was defined by expert judgement, literature research, measurements of different model inputs in the sewer system, and interviews with experts. The developers of the EmiStatR model, which have detailed knowledge about the processes in the system, were also consulted. The level of model sensitivity was derived by interpreting the model structure and components, interviews with experts, and model runs with EmiStatR. The main task to quantify input uncertainty is

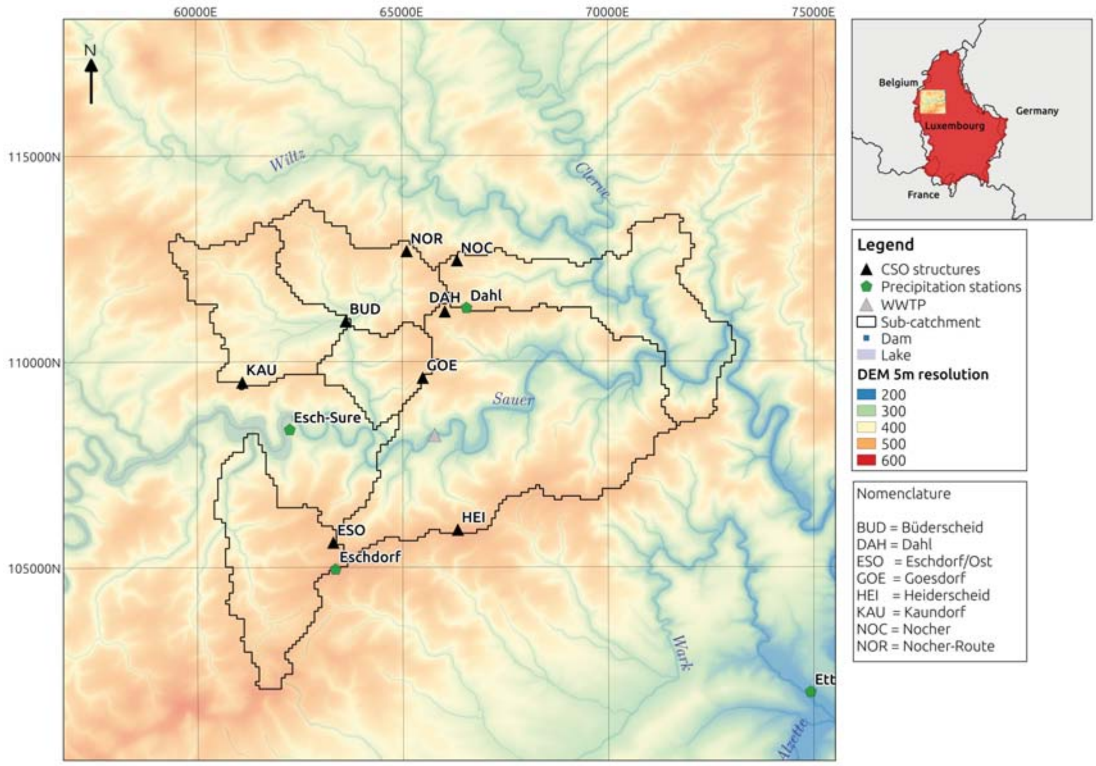

Figure 2: The Haute-Sûre catchment and villages of the sewer system. Location of CSO structures and precipitation measurement stations.

to define the probability distribution function (pdf) that represents the uncertainty of the variable chosen. The uncertainties of selected model inputs were characterized with pdfs following Heuvelink et al. (2007).

Measurement campaigns were done in Goesdorf from 28th April to 24th June 2011, in Kaundorf from 22nd June to 18th August in 2010 and from 20th July to 5th August in 2011, and in Nocher-Route from 18th November 2010 to 27th April 2011.

2.4 Uncertainty propagation

After the definition of model input uncertainties a Monte Carlo simulation was done to propagate input uncertainty to model output. The Monte Carlo method runs the EmiStatR model repeatedly, each time using different model input values, sampled from their probability distribution. The method thus consists of the following steps:

1. Repeat N times:
   (a) Generate a set of realisations of the uncertain model inputs
   (b) For this set of realisations, run the model and store the output
2. Compute and store sample statistics from the N model outputs.

Here, N is the number of Monte Carlo runs, i.e. the Monte Carlo sample size. Common sample statistics that measure the uncertainty are the standard deviation and the width of prediction intervals, which can be easily calculated from the N Monte Carlo outputs.

Variable	**Value**
Wastewater	
Water consumption, **qs**, [l/(PE · d]	150
Pollution COD, **CODs**, [g/(PE · d)]	104
Pollution NH_4, **NH4s**, [g/(PE · d)]	4.7
Infiltration water	
Inflow, **qf** ,[l/(s · ha)]	0.05
Pollution COD, **CODf**, [g/(PE · d)]	0
Pollution NH_4, **NH4f**, [g/(PE · d)]	0
Rainwater	
Precipitation time series, **P** [mm/min]	P1
Pollution COD, **CODr**, [mg/l]	80
Pollution NH4, **NH4r** [mg/l]	0
Storm water runoff	
Real flow time in the catchment, **tf** [min]	20

Table 1: General input data of the EmiStatR test model.

III RESULTS

3.1 Simulations of the CSO structures

Deterministic simulations of the EmiStatR model for the eight CSO structures were performed to compare these with the outcomes of the uncertainty analysis. The deterministic input variables of the simulations are given in Table 1. The individual deterministic input variables for each CSO structure are presented in Table 2. Nine output variables were evaluated. Three out of nine are related with water quantity: volume of filling of the CSO tank, $VTank$ [m^3]; overflow volume, V_{Ov} [m^3]; and overflow flow, Q_{Ov} [l/s]. Three other output variables are related with water quality in terms of COD: overflow COD load $B_{COD,Ov}$ [kg]; overflow COD average concentration, $C_{COD,Ov,Av}$ [mg/l]; overflow COD maximum concentration, $C_{COD,Ov,Max}$ [mg/l]. The remaining three output variables are also related with water quality but in terms of NH_4: overflow NH_4 load, $B_{NH4,Ov}$ [kg]; overflow NH_4 average concentration, $C_{NH4,Ov,Av}$ [mg/l]; and overflow NH_4 maximum concentration $C_{NH4,Ov,Max}$ [mg/l].

3.2 Model input uncertainty assessment

To assess the sensitivity of the model output to small changes in the model input, simulations with EmiStatR for each input variable with an increment and decrement of 10% with relation to the deterministic values (Tables 1 and 2) were performed. Basically, the variables $Aimp$, Qd, and V are the most sensitive variables of water quantity variables ($VTank$, V_{Ov} and Q_{Ov}). Regarding water quality in terms of COD the input variables $CODs$, $CODr$, $Aimp$, Qd, V, and P have greatest impact on output COD variables, whereas input variables qs, $NH4s$, qf, $NH4r$, $Aimp$, pe, Qd, V and P have the greatest impact on output NH_4 variables.

After evaluation of the model output sensitivity to input variables, and taking into account the degree of uncertainties of each input, we selected three input variables to be included in the uncertainty analysis: $CODs$, $NH4s$, and $CODr$. We leave the uncertainty propagation

Variable	**BUD**	**DAH**	**ESO**	**GOE**	**HEI**	**KAU**	**NOC**	**NOR**
Sub-catchment data								
Land use [-][1]	R/I	R/I	R/I	R/I	R/I	R/I	R/I	R/I
Total area [ha]	16.5	16.2	22.2	16.5	30.0	22.0	16.2	18.6
Impervious area [ha]	4.9	5.0	11.1	7.6	11.0	11.0	3.4	4.3
Population equivalents [PE]	289	460	705	611	676	358	260	326
Flow time structure [min]	10	10	7	10	6	10	10	10
Structure data								
Throttled outflow [l/s]	6	4	12	9	11	9	2	4
Volume [m^3]	90	270	330	190	220	180	166	157

[1]R = residential, I = industrial

Table 2: General input variables of the CSO structures of the EmiStatR test model.

analysis of $Aimp$ and P, which also score high on the sensitivity-uncertainty ranking, for future work.

The field measurements were the basis to characterise input uncertainty of $CODs$ and $NH4s$. Samples of COD and NH_4 in mg/l (91 in total for each variable) were analysed in the dry weather flow produced in the villages of Goesdorf, Kaundorf and Nocher-Route. An average wastewater amount was calculated for Goesdorf (153 l/(PE·d)), Kaundorf (112 l/(PE·d)) and Nocher-Route (94.3 l/(PE·d)). Table 3 presents the summary statistics of the dry weather flow measurements of COD and NH4 and the correspondent value of $CODs$ and $NH4s$. We use italic font for $CODs$ and $NH4s$ to emphasize that these are input variables of the model.

	COD [mg/l]	$CODs$ [g//(PE·d)]	**NH4** [mg/l]	$NH4s$ [g//(PE·d)]
Min	62	7	16.1	1.7
Mean	926	104	44.4	4.7
Max	3454	528	81.2	10.8
St. deviation	632	88	18.6	1.9

Table 3: Summary statistics of dry weather flow measurements for $CODs$ and $NH4s$ characterization.

Regarding $CODr$, no field measurements were available. Thus, expert judgement and values from the literature were used to characterise input uncertainty in $CODr$. A mean value of 80 mg/l and a standard deviation of 90 mg/l were used. For all three input variables, $CODs$, $NH4s$, $CODr$, we proposed a normal distribution to characterise input uncertainty with mean and standard deviation for each variable as defined above. In order to avoid negative values, we first transformed the variables by taking their natural logarithm.

3.3 Uncertainty propagation

We made a deterministic run of the model with values of the input variables as presented in Tables 1 and 2. Additionally, we performed 1,500 Monte Carlo simulations allowing $CODs$, $NH4s$,and $CODr$ as stochastic input variables with characteristics as defined in the previous section. In this way the total uncertainty of output variables due to input uncertainty was calculated. Figure 3 shows an example of model output and the total uncertainty band of 5 and 95 percentile for overflow concentration of COD (centre inset) and overflow concentration of NH_4

(bottom inset). In case of a rain event that produces a CSO, the uncertainty in the model output is quite large. Also, there is a systematic difference between the deterministic and the median run. The latter is always slightly above the deterministic run.

Next the contributions of each one of the input variables to the total uncertainty was calculated as well, by calculating the difference between the total uncertainty and the uncertainty obtained in the stochastic simulation of the other two variables. For instance, the uncertainty contribution of $CODs$ was calculated as the total uncertainty minus the uncertainty of the simulations running only $NH4s$ and $CODr$ in stochastic mode. Therefore, 4,500 additional Monte Carlo simulations were done to calculate the uncertainty contribution of the three input variables.

The results indicate that there is no uncertainty contribution of $CODs$, $NH4s$ and $CODr$ to output variables $VTank$, V_{Ov}, and Q_{Ov}. $CODs$ and $CODr$ contribute to uncertainty in $B_{COD,Ov}$ and $C_{COD,Ov,Av}$. $CODr$ has the most important uncertainty contribution to $B_{COD,Ov}$ (99.6%) and $C_{COD,Ov,Av}$ (90.4%). Finally, $NH4s$ contributes 100% of the uncertainty of $B_{NH4,Ov}$ and $C_{NH4,Ov,Av}$.

IV CONCLUSIONS AND FUTURE WORK

We applied an uncertainty analysis of the simplified urban drainage model EmiStatR. A characterisation of the input uncertainty of the three main input variables that control output uncertainty in water quality variables was done. We found that the uncertainty in the load of COD per capita per day in the sewage ($CODs$) and the concentration of COD in runoff ($CODr$) contribute to the uncertainty of the output variables: overflow COD load and overflow COD average concentration. $CODr$ has the most important uncertainty contribution in the load and concentration of COD. The load of NH_4 per capita per day in the sewage ($NH4s$) contributes totally in the uncertainty of overflow NH4 load and overflow COD average concentration.

Rainfall is one of the most important drivers in the definition of uncertainty of output variables as load and concentration of COD and NH_4. Therefore a detailed analysis should be done to quantify the contributions of rainfall input uncertainty in the total uncertainty. The rainfall uncertainty will be addressed in follow-up research as well as uncertainty due to impervious areas.

A further development of EmiStatR allows to analyse spatial model inputs. In this sense, a semi-distributed modelling framework can be implemented and several sub-catchments of the system modelled simultaneously, taking into account the inherent spatial variability of the inputs as rainfall and impervious areas and propagate the uncertainty of such inputs through the model. This also will be worked out in future research.

ACKNOWLEDGEMENTS

The work presented is part of the QUICS (Quantifying Uncertainty in Integrated Catchment Studies) project. This project has received funding from the European Unions Seventh Framework Programme for research, technological development and demonstration under grant agreement No. 607000 and the Luxembourg Institute of Science and Technology.

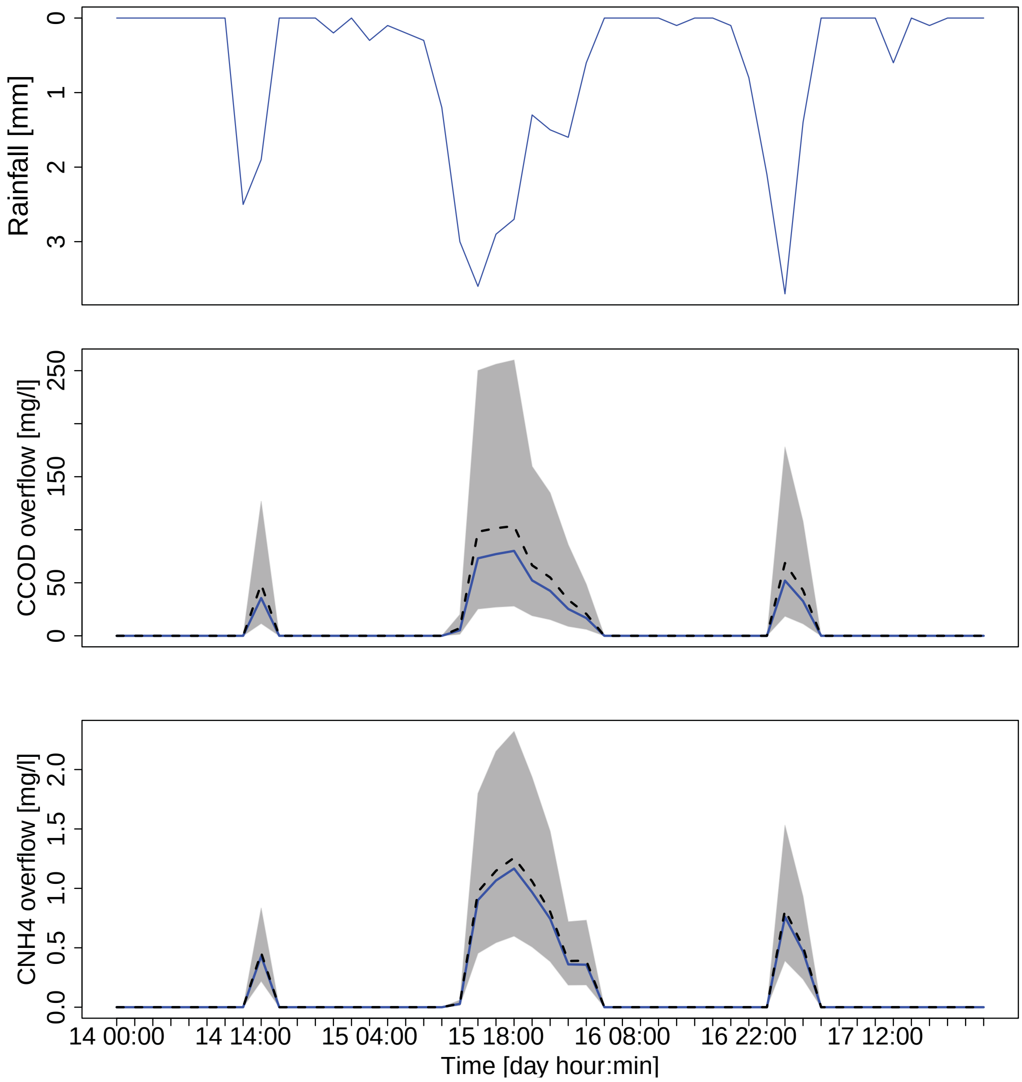

Figure 3: Precipitation, and overflow COD and NH_4 concentrations in deterministic and uncertainty outcomes. Location: Goesdorf. From 14 to 18 August 2010. Inset centre and bottoom: the blue line indicates the median, the black dotted line indicates the deterministic run, and the gray area indicates the 90 per cent uncertainty band.

References

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

Brown J. D. (2004). Knowledge, uncertainty and physical geography: Towards the development of methodologies for questioning belief. *Transactions of the Institute of British Geographers 29*(3), 367–381.

Deletic A., Dotto C., McCarthy D., Kleidorfer M., Freni G., Mannina G., Uhl M., Henrichs M., Fletcher T., Rauch W., Bertrand-Krajewski J., Tait S. (2012). Assessing uncertainties in urban drainage models. *Physics and Chemistry of the Earth 42-44*, 3–10.

Grigoriu M. (2012). *Stochastic Systems: Uncertainty Quantication and Propagation*. London: Springer-Verlag.

Heuvelink G. B. M., Brown J. D., van Loon E. E. (2007). A probabilistic framework for representing and simulating uncertain environmental variables. *International Journal of Geographical Information Science 21*(5), 497–513.

Klepiszewski K., Seiffert S. (2013). Statistische Erfassung von Entlastungsbauwerken der Mischwasserbehandlung im Einzugsgebiet der Chiers. MIGR EMISTAT-MW. Technical report, TUDOR Centre de Ressources des Technologies pour l'Environnement, Luxembourg.

Mckay M. D., Beckman R. J., Conover W. J. (1979). A Comparison of Three Methods for Selecting Values of

Input Variables in the Analysis of Output from a Computer Code. *American Statistical Association and the American Society for Quality*.

Minasny B., McBratney A. B. (2006). A conditioned Latin hypercube method for sampling in the presence of ancillary information. *Computers and Geosciences 32*(9), 1378–1388.

Mitchell V., Duncan H., Inman M., Rahilly M., Stewart J., Vieritz A., Holt P., Grant A., Fletcher T., Coleman J., Maheepala S., Sharma A., Deletic A., Breen P. (2007). State of the art review of integrated urban water models. In *Novatech 2007*, pp. 507–514.

Nash J. E., Sutcliffe J. E. (1970). River flow forecasting through conceptual models. Part 1 - a discussion of principles. *Journal of Hydrology (Amsterdam) 10*, 282–290.

Neumann M. B. (2007). *Uncertainty Analysis for Performance Evaluation and Design of Urban Water Infrastructure*. Ph. D. thesis, Swiss Federal Institute of Technology, ETH Zurich.

Nol L., Heuvelink G. B. M., Veldkamp a., de Vries W., Kros J. (2010). Uncertainty propagation analysis of an N2O emission model at the plot and landscape scale. *Geoderma 159*(1-2), 9–23.

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

Sanitary-District (2015, May). *Combined Sewer Overflow*. Internet.

Schellart A. N. A., Tait S. J., Ashley R. M. (2010). Towards quantification of uncertainty in predicting water quality failures in integrated catchment model studies. *Water Research 44*(13), 3893–3904.

van Keur P., Henriksen H. J., Refsgaard J. C., Brugnach M., Pahl-Wostl C., Dewulf A., Buiteveld H. (2008). Identification of major sources of uncertainty in current IWRM practice. Illustrated for the Rhine Basin. *Water Resources Management 22*(11), 1677–1708.

Walker W., Harremoës P., Rotmans J., van der Sluijs J., van Asselt M., Janssen P., Krayer von Krauss M. (2003, March). Defining Uncertainty: A Conceptual Basis for Uncertainty Management in Model-Based Decision Support. *Integrated Assessment 4*(1), 5–17.

Zoppou C. (2001). Review of urban storm water models. *Environmental Modelling and Software 16*(3), 195–231.