



Multivariate autoregressive modelling and conditional simulation for temporal uncertainty analysis of an urban water system in Luxembourg

Jairo Arturo Torres-Matallana^{1,2}, Ulrich Leopold², and Gerard B. M. Heuvelink¹

¹Soil Geography and Landscape Group, Wageningen University, Wageningen, the Netherlands

²Research Group for Sustainable Urban and Built Environment, Department for Environmental Research and Innovation, Luxembourg Institute of Science and Technology, Esch-sur-Alzette, Luxembourg

Correspondence: Jairo Arturo Torres-Matallana (arturo.torres@list.lu)

Received: 2 July 2020 – Discussion started: 10 July 2020

Revised: 19 October 2020 – Accepted: 10 November 2020 – Published: 13 January 2021

Abstract. Uncertainty is often ignored in urban water systems modelling. Commercial software used in engineering practice often ignores the uncertainties of input variables and their propagation because of a lack of user-friendly implementations. This can have serious consequences, such as the wrong dimensioning of urban drainage systems (UDSs) and the inaccurate estimation of pollution released to the environment. This paper introduces an uncertainty analysis in urban drainage modelling, built on existing methods and applied to a case study in the Haute-Sûre catchment in Luxembourg. The case study makes use of the EmiStatR model which simulates the volume and substance flows in UDS using simplified representations of the drainage system and processes. A Monte Carlo uncertainty propagation analysis showed that uncertainties in chemical oxygen demand (COD) and ammonium (NH₄) loads and concentrations can be large and have a high temporal variability. Furthermore, a stochastic sensitivity analysis that assesses the uncertainty contributions of input variables to the model output response showed that precipitation has the largest contribution to output uncertainty related with water quantity variables, such as volume in the chamber, overflow volume, and flow. Regarding the water quality variables, the input variable related to COD in wastewater has an important contribution to the uncertainty for the COD load (66 %) and COD concentration (62 %). Similarly, the input variable related to NH₄ in wastewater plays an important role in the contribution of total uncertainty for the NH₄ load (34 %) and NH₄ concentration (35 %). The Monte Carlo (MC) simulation procedure

used to propagate input uncertainty showed that, among the water quantity output variables, the overflow flow is the most uncertain output variable, with a coefficient of variation (cv) of 1.59. Among water quality variables, the annual average spill COD concentration and the average spill NH₄ concentration were the most uncertain model outputs (coefficients of variation of 0.99 and 0.82, respectively). Also, low standard errors for the coefficient of variation were obtained for all seven outputs. These were never greater than 0.05, which indicates that the selected MC replication size (1500 simulations) was sufficient. We also evaluated how the uncertainty propagation can more comprehensively explain the impact of water quality indicators for the receiving river. While the mean model water quality outputs for COD and NH₄ concentrations were slightly above the threshold, the 0.95 quantile was 2.7 times above the mean value for COD concentration and 2.4 times above the mean value for NH₄. This implies that there is a considerable probability that these concentrations in the spilled combined sewer overflow (CSO) are substantially larger than the threshold. However, COD and NH₄ concentration levels of the river water will likely stay below the water quality threshold, due to rapid dilution after CSO spill enters the river.

1 Introduction

Combined sewer systems are important components of urban water infrastructure. These systems are typically found in old

and large cities (Baker, 2009; Litrico and Fromion, 2009) and are designed to transport the water generated and accumulated in an urban catchment to the receiving water body. During normal conditions, all water is transported to the treatment facility before it is released to the environment. This is the so-called throttled outflow or pass-forward flow (Hager, 2010). However, during extreme conditions with heavy precipitation, the combined sewer overflow (CSO) discharges excess water directly to nearby streams, rivers, lakes, or other water bodies (Baker, 2009). The CSO contains polluted water and solid matter (Hager, 2010), which, when released to the environment, can have a damaging impact on the water quality status of the receiving waters (Bachmann-Machnik et al., 2018; Gasperi et al., 2012). CSO pollutant load emissions are of similar or greater magnitude than the emissions from wastewater treatment plants (Gasperi et al., 2012; Bachmann-Machnik et al., 2018). CSO discharge impacts are mainly high peak flows, high organic loads from single events, which can lead to oxygen depletion, and ecotoxic concentrations of ammonia (NH_3) (Miskewitz and Uchirin, 2013; Bachmann-Machnik et al., 2018). To reduce pollution in receiving waters it is important to minimise CSO load and concentration.

One of the main variables is the chemical oxygen demand (COD), which is an indicator of organic compounds in water. It is used to measure the effluent quality (Viana da Silva et al., 2011). High levels of COD are correlated with a decrease in the amount of dissolved oxygen (DO) available for aquatic organisms. A depletion in the DO concentration in the water column from near 9 mg L^{-1} (the maximum solubility of oxygen in estuarine water on an average summer day) to below 2 mg L^{-1} is referred to as hypoxia. If hypoxic conditions are reached, the health of the ecosystem is affected, causing physiological stress, and even death, to aquatic organisms (on Environmental and Natural Resources – CENR, 2003). Ammonium (NH_4) is another important variable and is an indicator of nitrogen compounds in water. Concentrations of NH_4 in water and wastewater are relevant because high levels of nitrogen in receiving waters can cause eutrophication and, therefore, excessive growth of algae and other micro-organisms, resulting in oxygen-dissolved depletion and fish toxicity (Huang et al., 2010).

To better assess environmental impacts, numerical models are applied in urban hydrology to simulate CSO emissions into the environment. It is recommended, however, that such modelling approaches consider the inherent uncertainty associated with the system representation and the approximation of the model to the reality (Hutton et al., 2011). Moreover, the model inputs are also not free of errors, and associated uncertainties will also propagate to the model output (Heuvelink, 1998).

A total of five approaches for representing the presence or absence of uncertainty and how it is represented in the context of urban water systems are often distinguished (Walker et al., 2003; Refsgaard et al., 2007; van der Keur et al., 2008;

Bach et al., 2014) as follows: (1) determinism, (2) statistical uncertainty, (3) scenario uncertainty, (4) recognised ignorance, and (5) total (unrecognised) ignorance. Following van der Keur et al. (2008), determinism applies when we have knowledge with absolute certainty about the system under analysis. This is the ideal world case, which is not realistic for urban hydrology systems. The statistical approach is useful when it is possible to describe uncertainty in statistical terms, i.e. when uncertainty can be characterised by probability distribution functions (pdf's). The scenario approach, in contrast, applies when quantitative probabilities cannot be determined and, instead, qualitative measures of uncertainty are used. It is used when possible outcomes of uncertain inputs are known but the probabilities of these outcomes are not (Brown, 2004). There is also no claim that the list of possible outcomes (scenarios) is exhaustive. Recognised ignorance occurs when there is an awareness of lack of knowledge but without any further possibility to process and address the recognised uncertainty. This is the case for very complex functional or inherently unidentifiable relationships, when, for example, predictions are infeasible due to the chaotic behaviour of the system or when our understanding of the system behaviour is too limited (van der Keur et al., 2008). This is common in social systems where the behaviour of humans and groups of humans may often be unpredictable. Finally, total ignorance is the state of complete lack of awareness about imperfect knowledge (van der Keur et al., 2008). It is the opposite of determinism and reflects a state in which we do not know that we do not know (Walker et al., 2003). Among the approaches described above, in this paper we will use the statistical approach to characterise and propagate uncertainties.

A total of three main sources of uncertainty in the context of performance evaluation analysis and design of urban water infrastructure and urban drainage modelling are identified (Walker et al., 2003; Neumann, 2007; Deletic et al., 2012). First, model input uncertainty is related to errors in input data, i.e. in driving forces such as precipitation. Second, parameter uncertainty is related to the uncertainty regarding the (calibrated) parameters of the model. Third, model structural uncertainty relates to uncertainty due to model conceptualisation and simplification. For instance, an urban drainage model might ignore certain sub-processes, such as evaporation or chemical transformation, or might simplify a non-linear relation between model variables to a linear relation. These types of uncertainties are not captured in model input and model parameter uncertainty and are represented by model structural uncertainty. The focus of this work is on the propagation of model input uncertainty.

Regarding methods for uncertainty propagation analysis, a distinction can be made between analytical methods, such as the Taylor series method (Heuvelink, 1998), and numerical techniques, such as Monte Carlo (MC) simulation. Numerical techniques are more flexible and hence more convenient for analysing uncertainty propagation with complex

models (Zoppou, 2001). MC simulations are computationally demanding, especially in the case of complex models, but they can still be used if there are sufficient computational resources (Bastin et al., 2013), among others, because it can greatly benefit from parallel computing.

Although uncertainty propagation analysis has been applied extensively in hydrologic modelling (e.g. Beven and Binley, 1992; Kuczera and Parent, 1998; Hutton et al., 2011; Vrugt et al., 2003a, b; Vrugt and Robinson, 2007; Renard et al., 2010; Datta, 2011), the number of applications of long-term simulations in urban drainage modelling is limited and typically does not consider the influence of the temporal and spatial correlation in the analysis of the propagation of input uncertainty. Temporal correlation occurs in uncertain dynamic variables, such as precipitation and COD of household wastewater, because values of these variables over short time lags will be more similar than over large time lags. The same concept applies to variables that are spatially distributed (Webster and Oliver, 2007). It is important to take the temporal (and spatial) correlation of uncertain inputs into account because this may have a major influence on the outcomes of an uncertainty analysis (Heuvelink, 1998). In this paper, we perform a temporal uncertainty propagation analysis in urban water modelling using a MC simulation. As a case study, we use the simplified model EmiStatR (Torres-Matallana et al., 2018b) to predict wastewater volume, COD, and NH_4 concentrations in CSOs for three urban–rural sub-catchments of the Haute-Sûre catchment in the northwest of Luxembourg.

The objectives of this study are to (1) select and characterise the main sources of input uncertainty accounting for the temporal auto- and cross-correlation within EmiStatR; (2) propagate input uncertainty through EmiStatR, taking into account the temporal auto- and cross-correlation of uncertain dynamic inputs; and (3) quantify and assess the contributions of each uncertainty source to model output uncertainty dynamically (over time) for the Luxembourg case study.

2 Materials and methods

2.1 The EmiStatR model

EmiStatR is used to simulate CSO flows and water quality concentrations. Details regarding the conceptual and mathematical model are provided in Torres-Matallana et al. (2018b). A list of the EmiStatR inputs and outputs is provided in Appendix A. The main components of the EmiStatR model are (1) dry weather flow (DWF), including infiltration flow (IF), (2) pollution of DWF, (3) rain weather flow (RWF), (4) pollution of RWF, (5) combined sewer flow (CSF) and pollution, and (6) combined sewer overflow (CSO) and pollution. Figure 1 illustrates the scheme of the sewer system analysed.

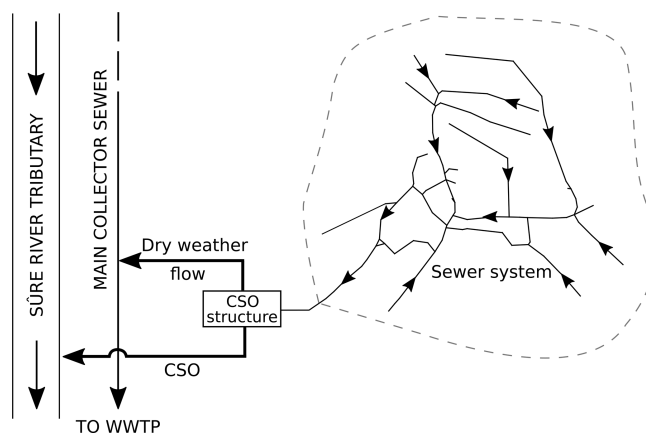


Figure 1. Scheme of the sewer system analysed. Adapted from Andrés-Doménech et al. (2010)

Basically, the total dry weather flow, Q_{DWF} (L s^{-1}), is calculated as follows:

$$Q_{\text{DWF}_t} = Q_{s_t} + Q_{f_t}, \quad (1)$$

where Q_{DWF_t} (L s^{-1}) is the dry weather flow at time t , and Q_{s_t} (L s^{-1}) is the dry weather flow of the residential sewage in the catchment at time t , calculated as $86\,400^{-1} \text{pe}_t \text{qs}_t$ (where $86\,400 = 24 \times 60 \times 60$ is a measurement unit conversion factor), with pe_t (PE) the population equivalents of the connected CSO structure at time t , and qs_t ($\text{L PE}^{-1} \text{d}^{-1}$) is the individual water consumption of households at time t . Q_{f_t} (L s^{-1}) is the infiltration flow at time t that enters the pipes from groundwater flow through cracks and joints, calculated as $A_{\text{imp}} \cdot q_{f_t}$, where A_{imp} (ha) is the impervious area of the catchment, and q_{f_t} ($\text{L s}^{-1} \text{ha}^{-1}$) is the infiltration water inflow flux (specific infiltration discharge from groundwater flow) at time t . Variables qs_t and pe_t are dynamic and can be defined as time series with daily, weekly, and seasonal patterns.

The contribution of rainwater to the combined sewage flow, Q_r ($\text{m}^3 \text{s}^{-1}$), is derived from precipitation as follows:

$$Q_{r_t} = \frac{1}{6} \cdot P_{t-t_{\text{RS}}} \cdot [C_{\text{imp}} \cdot A_{\text{imp}} + C_{\text{per}} \cdot (A_{\text{total}} - A_{\text{imp}})], \quad (2)$$

where $\frac{1}{6}$ is a factor for units conversion, $P_{t-t_{\text{RS}}}$ is precipitation at time $t - t_{\text{RS}}$ (mm min^{-1}), t_{RS} is a delay in time response related to flow time in the sewer system, A_{imp} is the impervious area of the catchment (ha), A_{total} is the total area of the catchment (ha), C_{imp} is the run-off coefficient for impervious areas (–), and C_{per} is the run-off coefficient for pervious areas (–). From Q_{r_t} , the CSO volume calculation is based on the exceeding volume stored in the combined sewer overflow chamber (CSOC). The CSO volume depends on four CSOC stages, namely (1) filling up, (2) CSO spill volume, (3) stagnation, and (4) emptying. The sum of the total dry weather

flow, Q_{DWF_t} , and the rain water flow, Q_{r_t} , is called combined sewer flow at time t , Q_{CSF_t} .

The COD load, $B_{COD, Sv}$ (g), in the spill overflow volume is calculated as a function of the spill overflow volume at time t , V_{Sv_t} (m^3), a combined sewer mixing ratio at time t , cs_{mr_t} (–), the mean dry weather pollutant concentration at time t , C_{COD_t} ($mg L^{-1}$), and the concentration due to rainwater pollution at time t , COD_{r_t} ($mg L^{-1}$) as follows:

$$B_{COD, Sv_t} = (cs_{mr_t} + 1)^{-1} V_{Sv_t} (cs_{mr_t} \cdot C_{COD_t} + COD_{r_t}). \quad (3)$$

The variable V_{Sv_t} depends directly on the water volume in the CSO chamber at time t , $V_{Chamber_t}$ (m^3). It is computed as follows:

$$V_{Sv_t} = \begin{cases} V_{r_t} + V_{dw_t} - V_{d_t}, & \text{if } V_{Chamber_t} = V, \\ V_{Chamber_t} - V & \text{if } V_{Chamber_t} > V, \\ \epsilon & \text{if } V_{Chamber_t} < V, \end{cases} \quad (4)$$

where V_{r_t} is the rain weather volume at time t accumulated during a time interval Δt (min), V_{dw_t} (m^3) is the total dry weather volume (amount of dry weather water in combined sewage flow) at time t , V_{d_t} is the volume of throttled outflow to the wastewater treatment plant (WWTP) at time t (m^3), V (m^3) is the CSOC volume, and ϵ is a numerical precision term set equal to 10^{-5} (m^3). While V_{Sv_t} , cs_{mr_t} , and C_{COD_t} are dynamic, COD_{r_t} can either be dynamic or assumed constant if the pollution concentration is assumed constant in time. C_{COD_t} ($mg L^{-1}$) is calculated as follows (Torres-Matallana et al., 2018b):

$$C_{COD_t} = \frac{10^3 \cdot pe_t \cdot C_{COD, s}}{qs_t \cdot pe_t + 86400 \cdot A_{imp} \cdot q_{f_t}}, \quad (5)$$

where $C_{COD, s}$ is the COD sewage pollution per capita (PE) load per day ($g PE^{-1} d^{-1}$). Similar equations, as above, apply to the second water pollution indicator NH_4 .

2.2 Sewer system in the Haute-Sûre catchment

The study area is composed of three sub-catchments of the Haute-Sûre catchment in the northwest of Luxembourg. The combined sewer system drains three villages, namely Goesdorf (GOE), Kaundorf (KAU), and Nocher-Route (NOR). The local sewer system downstream of each village has a CSO structure to store pollutant peaks in the first flush of combined sewage flows. Figure 2 depicts the location of the CSO structures and the delineation of the sub-catchments. The main land use types in the villages are residential, smaller industries, and farms. Outside the villages, forest and agricultural arable and grassland are the dominating land uses. The receiving water bodies of the CSO structures are tributaries of the river Sûre (or Sauer in German).

2.3 Input data

The input variables of the EmiStatR model are shown in Table S1 in the Supplement. Following Torres-Matallana et al. (2018b), seven input variables were calibrated, namely water consumption (qs_t), infiltration flow (q_{f_t}), flow time structure equivalent to the time of concentration to the combined sewer overflow structure (t_{fs}), run-off coefficient for impervious area (C_{imp}), run-off coefficient for pervious area (C_{per}), orifice coefficient of discharge (C_d), and the initial water level (Lev_{ini}). The main objective of the calibration process is to appropriately represent the water volume in the CSOC.

The observed precipitation (P_t) is a 1 year time series for 2010 at a 10 min time interval, measured at stations Esch-sur-Sûre and Dahl (Fig. 2). The variable water consumption (qs_t) is also dynamic and represented as a time series with a daily pattern according to factors proposed in the German ATV-DVWK-A 134 guideline (Evers et al., 2000).

The hydraulic variable measured is the water level in the CSOC, namely Lev (m). The temporal resolution of the measurements of Lev is 30 s. Regarding the wastewater quality (WWQ) characterisation, values of $C_{COD, s}$ and $C_{NH_4, s}$ in the wastewater were derived from DWF measurements at Goesdorf, Kaundorf, and Nocher-Route. A total of 91 2 h composite samples were taken and measured in the laboratory to determine the concentrations of COD ($mg L^{-1}$) and NH_4 ($mg L^{-1}$). This led to the identification of seven at Goesdorf on 4 May 2011, 48 between 19 June and 21 July 2010 at Kaundorf, and 36 between 9 March and 2 August 2011 at Nocher-Route. The variables COD_f and NH_{4f} were set to zero because the pollution contribution of the infiltration water is negligible in the study area. The contribution of ammonium from rainwater NH_{4r} was assumed constant and set to $2.00 mg L^{-1}$, while COD_r was equal to zero. Table S1 summarises the base values of the general input variables, and Table S2 presents the general characteristics of each CSO structure for each village and the base values of input variables. These base values were used when running EmiStatR in the deterministic mode (see Sect. 3.1). Some of the variables were calibrated based on observations in the CSOC to simulate water level and concentrations and loads of pollutants spilled in the CSO to the stream, river, or lake. These variables are water consumption (qs), infiltration flow (q_f), time flow (t_{fs}), run-off coefficient for the impervious area (C_{imp}), run-off coefficient for the pervious area (C_{per}), orifice coefficient of discharge (C_d), and initial water level (Lev_{ini}).

2.4 Selection of model input for uncertainty quantification

Following recommendations from Nol et al. (2010), not all model inputs were taken into account in the uncertainty propagation analysis. Only inputs that are very uncertain and to which the model output is very sensitive were included because these are the ones that have the largest contribution to

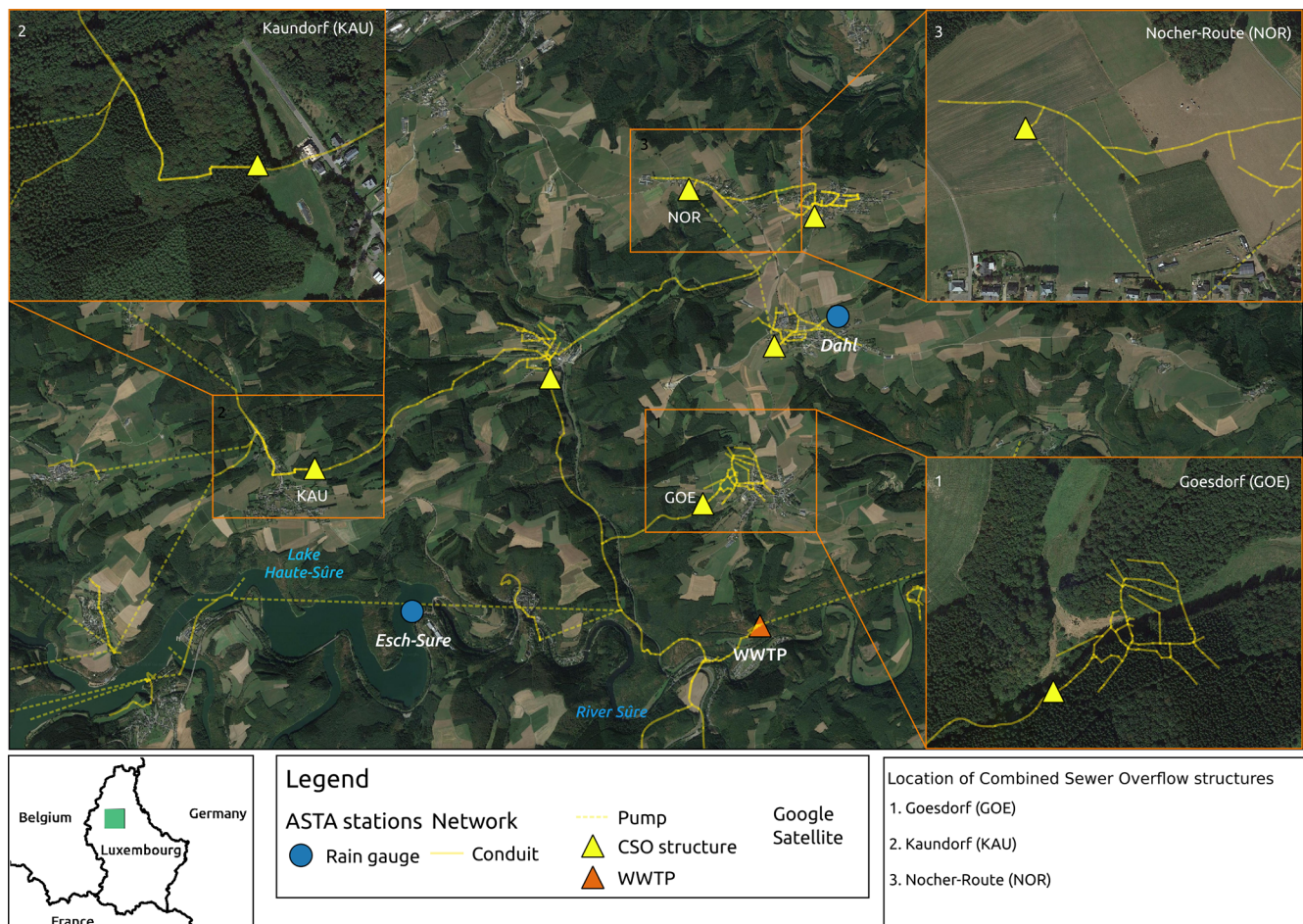


Figure 2. The three Haute-Sûre sub-catchments and locations of the combined sewer overflow (CSO) structures considered in this study. The background map is provided by © Google Maps.

output uncertainty (Heuvelink, 1998; Sect. 4.4). The level of uncertainty of the inputs was defined by expert judgement and similar case studies in the literature. A quick scan was used to determine the model sensitivity to each of the model inputs by running EmiStatR in deterministic mode with input base values given in Table S1. The level of model sensitivity was defined by analysing the mathematical model structure and components of the model, expert judgement, and simulations with EmiStatR. Inputs that rank high on both the level of uncertainty and on model sensitivity were selected and included in the uncertainty propagation analysis.

2.5 Uncertainty quantification of selected model input

Because we used a statistical approach, probability distribution functions (pdf's) are the basis for representing the uncertainties of the selected model inputs. This constitutes the most difficult step of an uncertainty propagation analysis and is done in different ways for constants and dynamic variables, as explained in the following sub-sections.

2.5.1 Uncertain constants

Following Heuvelink et al. (2007), an uncertain continuous numerical constant C can be characterised by its marginal (cumulative) pdf (mpdf) as follows:

$$F_C(c) = P(C \leq c). \quad (6)$$

Usually, a parametric approach can be taken, meaning that a common shape for F_C is chosen (e.g., normal, lognormal, exponential, or uniform) so that the mpdf is reduced to a number of parameters. In this study, the input variables that are in this category are water consumption (q_s), infiltration inflow (q_f), total area (A_{total}), impervious area (A_{imp}), the runoff coefficients for impervious area (C_{imp}) and pervious area (C_{per}), population equivalents (pe), flow time structure (t_{fs}), and initial water level (Lev_{ini}).

2.5.2 Univariate autoregressive modelling

Dynamic uncertain inputs may be temporally autocorrelated. This may dramatically influence the outcome of an uncertainty propagation analysis and must therefore be accounted for. One way of doing this is by assuming an autoregressive order one (AR(1)) model as follows:

$$\begin{aligned} y_t &= \mu + \phi(y_{t-1} - \mu) + w_t, \quad t = 1, 2, \dots, T, \\ y_0 &\sim \mathcal{N}(\mu, \sigma^2), \end{aligned} \quad (7)$$

where y_t is the uncertain input at time t , μ is its mean, ϕ is the autoregressive parameter ($0 \leq \phi < 1$), and w_t is a Gaussian white noise time series with mean zero and variance σ_w^2 . The initial value y_0 is taken from a normal distribution with mean μ and variance σ^2 . The parameters of the model can be estimated based on observations, or in the absence of observations, suitable values are taken based on expert judgement or literature reference values. Note that the effect of the initial condition usually fades out quickly and hence is not of great concern.

The implementation of the AR(1) model in R was done via the R function `arma.sim` of the R base package `stats` (R-Core-Team and contributors worldwide, 2017), both for model calibration and simulation.

2.5.3 Multivariate autoregressive modelling

In the case of multiple uncertain dynamic inputs, cross-correlation between these inputs may also need to be included. For example, $C_{\text{COD},s}$ and $C_{\text{NH}_4,s}$ and their uncertainties are likely correlated. This can be done using a multivariate AR(1) model (Luetkepohl, 2005), which is a natural extension of the univariate AR(1) model, as follows:

$$\begin{aligned} \mathbf{Y}(t+1) &= \boldsymbol{\mu} + \mathbf{A} \cdot [\mathbf{Y}(t) - \boldsymbol{\mu}] + \boldsymbol{\varepsilon}(t), \\ t &= 1, 2, \dots, T, \quad Y_0 \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Lambda}), \end{aligned} \quad (8)$$

where $\mathbf{Y}(t)$ is a vector of inputs at time t , \mathbf{A} is a square matrix with parameters that define how the variables at time $t+1$ depend on those at time t , $\boldsymbol{\mu}$ is now a vector of means, and $\boldsymbol{\varepsilon}(t)$ is a vector of zero-mean normally distributed white noise processes. We further assume that the variance–covariance matrix \mathbf{C} of $\boldsymbol{\varepsilon}(t)$ is time invariant. The initial value Y_0 is assumed normally distributed and uncorrelated ($\boldsymbol{\Lambda}$ is a diagonal matrix). In order to estimate the vector $\boldsymbol{\mu}$ and matrices \mathbf{A} and \mathbf{C} , a sample of the variables of interest is needed. Parameter estimation is done by means of the R package `mAr` (Barbosa, 2015).

2.5.4 Input precipitation model

In case precipitation is selected as an uncertain input to be included in the uncertainty analysis, it too must be charac-

terised by a pdf. Since precipitation, however, is not normally distributed and has many zeros, we cannot make the Gaussian assumption, and hence we cannot use the approach described in Sect. 2.5.2 to model its dynamic behaviour and uncertainty. In addition, we usually have precipitation measurements nearby so we need to condition the simulations to these measurements. Recall from Sect. 2.3 that, in the case study, precipitation data are recorded at stations Esch-sur-Sûre and Dahl.

Torres-Matallana et al. (2017) present a model to simulate precipitation inside a target catchment given a known precipitation time series in a nearby location outside the catchment, while accounting for the uncertainty that is introduced due to spatial variation in precipitation. The method used for input precipitation uncertainty characterisation is essentially the same as the application of a Kalman filter/smoothing (Kalman, 1960; Webster and Heuvelink, 2006). Calibration of the model requires precipitation time series at two locations near the catchment of interest. Once the model is calibrated, it is used to simulate precipitation inside the target catchment from a single precipitation time series nearby the catchment. Details of the calibration and conditional simulation are presented in Sect. S4.

2.6 Uncertainty analysis

We used a MC simulation (Hammersley and Handscomb, 1964; Kalos and Whitlock, 2008) to analyse how input uncertainty propagates through the EmiStatR model because it is flexible and straightforward to implement. It is also feasible in our case study because EmiStatR is a relatively simple model that does not involve a long computation time.

2.6.1 Monte Carlo simulation

The MC method runs the EmiStatR model repeatedly, each time using different model input values sampled from their pdf. The method thus consists of the following steps:

1. Repeat n times:
 - (a) Generate a set of realisations of the uncertain model inputs at 10 min resolution.
 - (b) Run the model at 10 min resolution and store output for this set of realisations. Later, in order to compute the summary statistics, a temporal aggregation of the model output to 1 h intervals is done.
2. Compute and store sample statistics from the n model outputs.

Here, n is the number of MC runs, i.e. the MC sample size. Common sample statistics that measure the uncertainty are the standard deviation and quantiles of the distribution of MC outputs, such as the difference between the 0.95 and 0.05 quantile, which can be easily calculated from the n MC outputs.

Sampling from the pdf of uncertain inputs was done using simple random sampling.

2.6.2 Monte Carlo output summary

Proper presentation of MC outputs is important for achieving the most from the experiment. Therefore, summary statistics are one important way to summarise the MC outputs. Commonly, a MC study yields n model outputs, which are stored in the MC result matrix \mathbf{X} in Boos and Osborne (2015). From this matrix, various statistics can be computed. Basic summary statistics include the mean μ_{MC} , the standard deviation (σ_{MC}), and the variance σ_{MC}^2 . From these, we can compute the coefficient of variation CV_{MC} (σ_{MC}/μ_{MC}), which is a dimensionless expression of relative uncertainty. The coefficient of variation is a standardised measure of the spread of a sampling distribution, which is useful because it allows a direct comparison of the variation in samples with different units or with very different means (Marwick and Krishnamoorthy, 2019). We computed estimates and standard errors for these statistics and also for the interquartile range (IQR_{MC}), 0.005 ($\zeta_{0.005}$) and 0.995 ($\zeta_{0.995}$) quantiles, and the 99 % width of the prediction band ($\zeta_{w, 0.99}$).

2.6.3 Bootstrap computation for Monte Carlo summary

Following Boos and Osborne (2015), who argue that “Good statistical practice dictates that summaries in MC studies should always be accompanied by standard errors,” we used the bootstrap method to compute the standard errors of all MC statistics. These tools are particularly relevant in a case without analytic solutions (Boos, 2003). According to Boos and Osborne (2015, p. 228) standard errors for MC output statistics are often not computed, which is an additional computational step on top of the overall analysis. Standard errors are straightforward to compute for simple statistics, such as the sample mean over the replications of the MC output, but are more difficult to compute for more complex statistics, such as medians, sample variances, and the classical Pearson measures of skewness and kurtosis. Therefore, to avoid burdensome computations, we opted to compute the standard errors by the bootstrap method. We briefly explain the bootstrap method below. For a detailed explanation, we refer to Efron (1979).

To compute the bootstrap variance of estimators, we follow the logic given by Boos and Osborne (2015). From a MC sample Y_1, \dots, Y_n , we draw a random sample of size n , with replacement Y_1^*, \dots, Y_n^* , and compute an estimator $\hat{\theta}$ of the MC statistic θ from this resample. We independently repeat this process B times, resulting in a sample of estimators

$\hat{\theta}_1, \dots, \hat{\theta}_B$. Then the bootstrap variance estimate, \hat{V}_B , is the sample variance of this sample of estimators as follows:

$$\hat{V}_B = \frac{1}{B-1} \sum_{i=1}^B (\hat{\theta}_i - \bar{\hat{\theta}})^2, \quad (9)$$

where $\bar{\hat{\theta}}$ is the mean of the sample of estimators. The MC standard error, se , is simply the square root of the bootstrap variance.

We implemented, in *stUPscales* (Torres-Matallana et al., 2019), specific routines for computing, by means of the bootstrap method, the MC estimators and their standard error for all MC statistics, where the variance of the model output is the most important. We compared our results with the results obtained using the *MonteCarlo.se* R package (Boos et al., 2019).

2.6.4 Contributions of input variables to total uncertainty

A number of $m+1$ MC analyses are needed to compute the contributions of input variables to total uncertainty, where m is the number of model input variables selected for uncertainty quantification. The first MC analysis, MC_{tot} , is done to compute the total output uncertainty by stochastically varying all input variables. The uncertainty associated with the first variable x_1 is quantified by a second MC analysis MC_1 in which only x_1 is equal to its deterministic value, while the other input variables vary stochastically. Similarly, the other MC simulations MC_2, MC_3, \dots , and MC_m are used to quantify the uncertainty for the variables x_2, x_3, \dots , and x_m .

To quantify the contributions of individual input variables to the total uncertainty of the model inputs, the stochastic sensitivity S_i for each uncertain input x_i is computed. It is defined as follows:

$$S_i = 1 - \frac{\text{Var}(MC_i)}{\text{Var}(MC_{tot})}. \quad (10)$$

The index represents the main effect contribution of each input factor to the total variance of the output. The larger the index, the more important the input uncertainty. We computed stochastic sensitivity indices per time step and aggregated contributions for the whole year. For plotting purpose, we aggregated the outputs from a 10 min time step to hourly time steps. The aggregation was done for each individual MC run before the contributions were computed.

2.7 Water quality impact

The results of the MC uncertainty propagation were also compared with the water quality standards. Standards are introduced to evaluate the impact of emissions of COD and NH_4 in CSOs in the receiving water. However, as Toffol

(2006) recognises, although there are European emission standards for wastewater treatment plant effluent, standards for combined sewer overflow are not so clear. According to Steinel and Margane (2011), the European Water Framework Directive (WFD) is mainly concerned with the natural state of waters. Therefore, emission standards for effluent discharge are not set. The European Union Council Directive 91/271/EEC (1991) sets standards for COD and total nitrogen; hence, similar values have been adopted in many European member states. For more details about guidelines and design procedures in Europe, see Blumensaat et al. (2012). We assessed the emissions according to the German guideline ATV-DVWK-A 128 (1992), which is the standard for the dimensioning and design of storm water structures in combined sewers and commonly used in Luxembourg. The Austrian ÖWAV-RB 19 (2007) is also taken into account because it provides key reference guidelines for the design of urban water infrastructure in central Europe. Three main indicators are taken into account, namely hydraulic impact, COD concentration, and acute ammonium toxicity.

2.7.1 Hydraulic impact

According to the Austrian guidelines, and as summarised by Kleidorfer and Rauch (2011), the evaluation of the hydraulic impact is given by the following:

$$Q_1 \leq f_h \cdot Q_{r1}, \quad (11)$$

where $0.1 \leq f_h \leq 0.5$, Q_1 (L s^{-1}) is the maximum sewer overflow discharge with return period of 1 year, and Q_{r1} (L s^{-1}) is the maximum water discharge in the river with return period of once per year. The factor f_h is taken as 0.1 in more sensitive streams, whereas it is 0.5 for streams with more stable bed and higher re-colonisation potential Toffol (2006). Time series of daily values recorded in 2006 to 2013 of the river Sûre at Heiderscheidgrund were used to compute the daily flow expected, with a return period of once per year (1.01 years), Q_{r1} .

We used the German guideline ATV-DVWK-A 128 (1992), which computes the throttle discharge at CSOs, $Q_{t, \text{CSO}}$ (L s^{-1}) as follows:

$$Q_{t, \text{CSO}} = f_t \cdot A_{\text{imp}}, \quad (12)$$

where $7.5 \leq f_t \leq 15$ and A_{imp} (ha) are the impervious areas connected to the combined sewer system. For the overflow flow MC output mean, the 0.95 and 0.995 quantile, we computed the exceedance percentage over the thresholds, which is calculated as the proportion of time steps exceeding the number of total time steps in the year (8759 time steps at 1 h).

2.7.2 COD concentration

Steinel and Margane (2011, Table 14) present the effluent standards for discharging into freshwater adopted in selected European countries. A COD concentration of 125 mg L^{-1} is reported for European Union countries. Austria has stricter rules, with a standard of 90 mg L^{-1} for populations between 50 and 500 inhabitants and 75 mg L^{-1} for populations greater than 500 inhabitants. The Goesdorf population was 1025 inhabitants by 2001 and was 1297 inhabitants by 2011 (Statec, 2020). For NH_4 a similar approach was used.

2.7.3 Acute ammonium toxicity

Following Kleidorfer and Rauch (2011), who argue that “the ammonia (NH_3) concentration depends on the ammonium (NH_4) concentration and on the dissociation equilibrium between NH_3 and NH_4 (which is influenced by temperature and pH-value)”. According to Kleidorfer and Rauch (2011), the Austrian guideline ÖWAV-RB 19 (2007) establishes a maximum value of 2.5 mg L^{-1} for the ammonium (NH_4) concentration calculated for a 1 h duration for salmonid streams. For cyprinid streams, a maximum value of 5.0 mg L^{-1} is recommended.

3 Results

3.1 Selection of model inputs for uncertainty quantification

In this section, we assess the degree of uncertainty and sensitivity for all input variables, following the procedure described in Sect. 2.4. We summarise the results in Tables 1 and 2. To better support our decisions, we also include a graphical assessment of the degree of uncertainty and sensitivity of each input, as in Tscheikner-Gratl et al. (2017, see Fig. S1).

3.1.1 Wastewater

Water consumption, q_s , is a fairly uncertain input variable, and the model output is sensitive to this variable. Volume and flow of CSO are sensitive to changes in q_s . Regarding water quality output, the total load of NH_4 is very sensitive to changes in q_s . Pollution of sewage as COD load per capita per day, $C_{\text{COD}, s}$, is the first selected input variable for propagation of uncertainty, due to the fact that it is both a very uncertain input variable and the model output (average and 99.9 percentile overflow COD concentration) is very sensitive to it. Pollution of sewage as NH_4 load per capita per day, $C_{\text{NH}_4, s}$, is also included in the uncertainty propagation analysis. It is a very uncertain input variable, and the model output (overflow load and concentrations of NH_4) is very sensitive to it. The variables $C_{\text{COD}, s}$ and $C_{\text{NH}_4, s}$ are very uncertain because these are correlated to the temporal and spatial

pattern of water consumption, which has a daily, weekly, and seasonal temporal variability.

3.1.2 Infiltration water

Inflow of infiltration water, q_f , is a very uncertain input variable because this inflow depends on the number of anomalies in the pipes (cracks or wrong connections) that allows infiltrations flowing into and out of the system. The distribution of these anomalies has a strong random component, and hence, q_f is very uncertain, and the model output is sensitive to it. Although this is a very uncertain input, the quick-scan analysis showed that model output sensitivity is not very high, as indicated in Table 1. For this reason, we did not include this variable in the uncertainty propagation analysis.

Pollution of infiltration water as COD load per capita per day, COD_f , and pollution of infiltration water as NH_4 load per capita per day, NH_{4f} , are not uncertain because in the Haute-Sûre study area the values of these variables are negligible.

3.1.3 Rainwater

Precipitation, P , is the main driving force of the model, and given the spatial variability of the rain fields, this input is considered very uncertain. The model output, additionally, is very sensitive to it. As a consequence, this input variable is treated as the third input variable in the uncertainty propagation analysis. Pollution of run-off as COD concentration, COD_r , is the fourth input variable considered in the uncertainty propagation, given that it is a very uncertain and very sensitive input variable, particularly to the load and concentration of COD in the overflow.

Pollution in rainwater as NH_4 concentration, NH_{4r} , is considered fairly uncertain. The model output (overflow load and concentration of NH_4) is very sensitive to it. Although model output is very sensitive to this model input variable, model input uncertainty is not very high, as indicated in Table 1. For this reason, it was not included in the uncertainty propagation analysis.

3.1.4 Sub-catchment

The model is very sensitive to the total area A_{total} and to the run-off coefficient for the pervious area (C_{per}) and sensitive to the impervious area A_{imp} and to the run-off coefficient for impervious area (C_{imp}). However, we did not include A_{total} in the uncertainty analysis because it can be fairly accurately derived from spatial databases, and hence, their uncertainty is not large. Although model output is very sensitive to the input variable C_{per} , the uncertainty about this variable is not very high, as indicated in Table 1. The reason behind this is that C_{per} can be derived fairly accurately from geographic information system (GIS) products, such as land use and soil type maps. Therefore, we did not include this variable in the uncertainty propagation analysis. The population equivalent

lents, pe , is a sensitive variable but not very uncertain. Hence, this variable was not included in the uncertainty analysis. The theoretical largest flow time in the catchment t_{fs} is not uncertain and not sensitive.

3.1.5 CSO structure

Although model output is very sensitive to maximum throttled outflow $Q_{d,max}$ and volume V , these are not included in the uncertainty analysis because their values are accurately known. The same is true for the variables curve level–volume $lev2vol$, orifice diameter D_d , and discharge coefficient C_d . These variables are accurately known and, therefore, not considered as being uncertain variables. The initial water level in the chamber Lev_{ini} is very uncertain, but the model output is not sensitive to this variable. Therefore, Lev_{ini} was not included in the uncertainty analysis.

3.2 Uncertainty quantification of selected model input

After ranking all inputs on levels of uncertainty and model sensitivity, we selected four input variables to be included in the uncertainty analysis. These are $C_{COD,S}$, $C_{NH_4,S}$, COD_r , and P (Table 2).

3.2.1 Sewage per capita COD and Ammonium

The fit of pdf's for the two uncertain inputs, $C_{COD,S}$ and $C_{NH_4,S}$, was based on measurements under dry weather flow conditions. Measurement campaigns were done in Goesdorf from 28 April to 24 June 2011, in Kaundorf from 22 June to 18 August in 2010 and from 20 July to 5 August in 2011, and in Nocher-Route from 18 November 2010 to 27 April 2011. Samples of COD and NH_4 in $mg\ L^{-1}$ (91 in total for each variable) were analysed. An average wastewater amount was calculated for Goesdorf ($153\ L\ PE^{-1}\ d^{-1}$), Kaundorf ($112\ L\ PE^{-1}\ d^{-1}$), and Nocher-Route ($94.3\ L\ PE^{-1}\ d^{-1}$). Table 3 presents summary statistics of the dry weather flow measurements of COD and NH_4 and the corresponding value of $C_{COD,S}$ and $C_{NH_4,S}$. COD is converted to $C_{COD,S}$ by means of a simple conversion from $mg\ L^{-1}$ to $g\ PE^{-1}\ d^{-1}$, by multiplying COD by the measured per capita flow ($112\ L\ PE^{-1}\ d^{-1}$) and dividing by 1000. NH_4 was converted to $C_{NH_4,S}$ in a similar way.

Closer inspection showed that $C_{COD,S}$ and $C_{NH_4,S}$ observations are best characterised by a lognormal distribution (Sect. S5). Since $C_{COD,S}$ and $C_{NH_4,S}$ are dynamic and cross-correlated, we calibrated a bivariate AR(1) model with state vector $Y = [\log(C_{COD,S})\ \log(C_{NH_4,S})]^T$. The estimated pa-

Table 1. Results of deterministic sensitivity analysis. Average percentage of change in model output caused by $\pm 10\%$ change in model inputs (qs, $C_{\text{COD},s}$, $C_{\text{NH}_4,s}$, COD_r , pe, and P as time series; VAR(1) model for $C_{\text{COD},s}$ and $C_{\text{NH}_4,s}$; and AR(1) model for COD_r and AR(1) conditioned for P ; see Appendix A for nomenclature definitions). Output change greater than 15 % is considered very high. Variable C_d (not shown in the table) leads to a percentage of change of less than 0.3 %, while variables t_{FS} and Lev_{ini} (not shown in the table) lead to no change in the output. Values greater than 15 are shown in bold.

Output variable	Input variable														
	qs	C _{COD, S}	C _{NH₄, S}	q _f	COD _r	NH _{4r}	A _{total}	A _{imp}	C _{imp}	C _{per}	pe	Q _{d, max}	V	D _d	P
V _{Chamber}	4.2	0.0	0.0	3.0	0.0	0.0	8.6	7.1	5.9	7.2	4.2	16.7	7.5	0.8	13.4
V _{Sv}	2.9	0.0	0.0	1.7	0.0	0.0	19.6	11.6	13.5	16.1	2.9	13.1	16.5	0.2	17.8
Q _{Sv}	0.6	0.0	0.0	0.2	0.0	0.0	2.7	1.5	1.1	1.8	0.6	13.1	14.8	0.6	12.4
B _{COD, Sv}	2.9	2.4	0.0	1.7	7.7	0.0	20.1	11.8	13.7	16.6	5.3	15.7	14.5	0.2	20.7
C _{COD, Sv, Av}	0.7	4.6	0.0	0.5	5.4	0.0	0.5	0.6	0.6	0.7	4.0	3.2	4.0	0.2	1.1
C _{COD, Sv, 99.9}	1.6	6.7	0.0	0.8	3.4	0.0	2.8	2.3	1.9	2.4	5.1	0.0	0.0	0.0	9.2
C _{COD, Sv, Max}	1.6	6.7	0.0	0.8	3.4	0.0	2.8	2.3	1.9	2.4	5.1	0.0	0.0	0.0	9.2
B _{NH₄, Sv}	3.1	0.0	3.3	1.8	0.0	6.7	20.4	12.0	13.8	16.8	6.4	17.0	13.4	0.3	22.1
C _{NH₄, Sv, Av}	0.9	0.0	5.8	0.6	0.0	4.2	0.6	0.8	0.9	0.9	5.3	4.3	5.4	0.2	1.5
C _{NH₄, Sv, 99.9}	1.6	0.0	7.6	0.8	0.0	2.4	3.5	2.6	2.3	2.9	6.1	0.0	0.0	0.0	11.3
C _{NH₄, Sv, Max}	1.6	0.0	7.6	0.8	0.0	2.4	3.5	2.6	2.3	2.9	6.1	0.0	0.0	0.0	11.3

rameters of the model, using the methodology described in Sect. 2.5.3, are as follows:

$$\mu = \begin{bmatrix} 4.40947 \\ 3.70411 \end{bmatrix} \quad \mathbf{A} = \begin{bmatrix} 0.99165 & -0.00319 \\ -0.00009 & 0.99455 \end{bmatrix}$$

$$\mathbf{C} = \begin{bmatrix} 0.00913 & 0.00224 \\ 0.00224 & 0.00185 \end{bmatrix}. \quad (13)$$

The defined multivariate autoregressive model also captures the dynamic behaviour, temporal correlation, and cross-correlation of the input variables, deriving the probability distributions of $C_{\text{COD},s}$ and $C_{\text{NH}_4,s}$ from measurements in the Haute-Sûre catchment, which agreed well with values reported in the literature (Katukiza et al., 2014; Heip et al., 1997).

3.2.2 Run-off COD concentration

Regarding COD_r , due to the fact that no field measurements were available, expert judgement and reference values from the literature were the basis for characterising the pdf of this input variable. The variable was assumed to be lognormally distributed with a mean value of 71 mg L^{-1} . Although, House et al. (1993) and Welker (2008) reported a higher value, namely 107 mg L^{-1} for COD_r , we selected a lower value due to the specific characteristics of the CSO system in the Haute-Sûre catchment. The value of 150 mg L^{-1} as standard deviation of COD_r leads to a coefficient of variation ($\text{SD} \cdot \text{mean}^{-1}$) equal to 2.11, which is greater than the coefficient of variation for $C_{\text{COD},s}$ (0.84). We allow the standard deviation of COD_r to be greater than the standard deviation of $C_{\text{COD},s}$ because COD measurements in rain water are very uncertain.

3.2.3 Input precipitation model

Precipitation and its associated uncertainty was modelled as an autoregressive model conditioned to the observed precipitation at a nearby measurement station. We assumed a multivariate lognormal distribution and included a temporal correlation of the simulated time series. Calibration of the precipitation model is done with the mAr package, as explained in Sect 2.5.4, and using a 10 min precipitation time series of stations Esch-sur-Sûre and Dahl for 2010. Upon calibration of the multivariate autoregressive model, we proceeded with the conditional simulation of Y_c (Sect. S4.2, Eq. 5). For this, we computed the parameters of the model, as shown in Eq. (14). The model parameters are given by (Torres-Matallana et al., 2017) the following:

$$\mu_1 = 2.85501 \quad \mu_\delta = 0.10194 \quad \mathbf{B} = \begin{bmatrix} 0.95650 & 0.03980 \\ 0.02429 & 0.88304 \end{bmatrix}$$

$$\sigma_1^2 = 0.07241 \quad \sigma_\delta^2 = 0.07951 \quad \rho_{1\delta} = -0.03876. \quad (14)$$

Next, we generated conditional simulations of the 10 min precipitation for 2010 for each subcatchment using the approach described in Sect. 2.5.4. Note that this involves simulating log-transformed precipitation, which can easily be transformed to precipitation data using the antilog. The simulation procedure was repeated as many times as simulated precipitation time series were required for the MC uncertainty propagation analysis.

The simulated precipitation time series captured the main statistics of the observed time series well. The reader can find evidence for this in the Supplement (Table S4 and Fig. S2). Despite the satisfactory performance of the pro-

Table 2. Input variables of the EmiStatR model and selection of inputs for uncertainty analysis based on input uncertainty level and model sensitivity level (Note: ++ is very uncertain/sensitive; – – is not uncertain/sensitive).

Input variable	Input uncertainty	Model sensitivity	Uncertainty analysis
Wastewater			
1. q_s	+	+	No
2. $C_{\text{COD}, S}$	++	++	Yes
3. $C_{\text{NH}_4, S}$	++	++	Yes
Infiltration water			
4. q_f	++	+	No
5. COD_f	– –	– –	No
6. NH_{4f}	– –	– –	No
Rainwater			
7. P	++	++	Yes
8. COD_r	++	++	Yes
9. NH_{4r}	+	++	No
Sub-catchment			
10. A_{total}	+	++	No
11. A_{imp}	+	+	No
12. C_{imp}	+	+	No
13. C_{per}	+	++	No
14. p_e	+	+	No
15. t_{FS}	–	– –	No
CSO structure			
16. $Q_{d, \text{max}}$	–	++	No
17. V	–	++	No
18. D_d	– –	– –	No
19. C_d	– –	– –	No
20. Lev_{ini}	++	– –	No

posed method, some cases showed an overestimation of the simulated precipitation, mainly due to high values of the ratio of the multiplicative factor $\delta(t)$. This behaviour was also recognised by McMillan et al. (2011), who stated that the multiplicative factor used in their study “does not capture the distribution tails, especially during heavy precipitation where input errors would have important consequences for run-off prediction”.

3.3 Uncertainty analysis

Model output sensitivity and the degree of uncertainties evaluation of each model input helped to define the four input variables included in the uncertainty analysis, namely $C_{\text{COD}, S}$, $C_{\text{NH}_4, S}$, COD_r , and P . In this section, we present the results of the uncertainty propagation for these four selected input variables to the model output, both for water quantity (volume in the combined sewer overflow chamber,

CSOC, and overflow volume and flow) and for water quality (loads and concentrations of chemical oxygen demand, COD, and ammonium, NH_4).

3.3.1 Monte Carlo output and uncertainty quantification

The seven output variables from EmiStatR were analysed by MC input uncertainty propagation. A detailed description of the Monte Carlo simulation size and timing is presented in the Supplement (Sect. S6).

Figure 3 illustrates the uncertainty propagation outcomes for the first Monte Carlo simulation, in which all input variables vary stochastically. The MC simulations were performed for the entire year of 2010 at 10 min time steps, which were aggregated to hourly time steps in the figure. The aggregation function was used for precipitation, CSO chamber volume, CSO spill volume, and loads was the sum, whereas for CSO spill flow and concentrations the aggregation function was the mean. The top of Fig. 3 shows input precipitation as main driving input. For illustration purposes, two events with a 2 d duration each are shown. The first event occurred in spring (May 2010) and the second in autumn (September 2010), and this shows that the uncertainty is high when there is a high precipitation event. The more intense the precipitation input, as seen in Fig. 3 in the top left (May event), the greater the uncertainty bandwidth for overflow flow and for COD and NH_4 loads (Fig. 3e–h) and concentrations (Fig. 3i–l). The MC-estimated statistics and the standard errors (se) are presented in Table 4. The table shows the uncertainty quantification of outputs obtained from the MC uncertainty propagation for the first MC simulation (all selected input variables are uncertain).

Table 4 shows the standard deviation (SD) and the coefficient of variation (cv) for the seven output variables considered in the uncertainty propagation. For the volume in the CSO chamber, V_{Chamber} , the annual mean standard deviation, σ_{MC} , (8.27 m^3) is lower than the mean, μ_{MC} , (92.51 m^3). This goes along with an annual mean coefficient of variation (CV_{MC}) of 0.100. A (CV_{MC}) greater than one means large uncertainty. The overflow spill volume, V_{SV} , had a coefficient of variation of 0.070, while it was 1.585 for the overflow flow, Q_{SV} . This shows that the relative uncertainty of the overflow flow is very large. Regarding the overflow COD load, the annual mean (1.18 kg) is similar as the annual mean standard deviation (1.27 kg). Similar behaviour was observed for the overflow COD concentration, which had an annual mean value of 170 mg L^{-1} and a standard deviation of 162 mg L^{-1} . For overflow NH_4 load and overflow NH_4 concentration, the annual mean also had the same order of magnitude as the annual mean standard deviation. Overflow COD and NH_4 loads had a coefficient of variation of 0.087 and 0.075, respectively, whereas the coefficient of variation for concentrations were 0.988 and 0.815, respectively. This suggests that overflow concentrations are more uncertain.

Table 3. Summary statistics of dry weather flow measurements for $C_{\text{COD},\text{S}}$ and $C_{\text{NH}_4,\text{S}}$ characterisation.

	COD (mg L ⁻¹)	$C_{\text{COD},\text{S}}$ (g PE ⁻¹ d ⁻¹)	$\log(C_{\text{COD},\text{S}})$ $\log(\text{g PE}^{-1} \text{d}^{-1})$	NH ₄ (mg L ⁻¹)	$C_{\text{NH}_4,\text{S}}$ (g PE ⁻¹ d ⁻¹)	$\log(C_{\text{NH}_4,\text{S}})$ $\log(\text{g PE}^{-1} \text{d}^{-1})$
Min	61.9	6.9	1.936	16.10	1.745	0.556
P5	216.8	23.8	3.167	20.55	2.102	3.018
Mean	925.5	104.2	4.378	44.38	4.733	1.473
P95	2032.0	236.8	5.466	79.00	7.684	2.039
Max	3454.0	528.5	6.270	81.20	10.771	2.377
SD	631.7	87.5	0.751	18.56	1.917	0.410

Table 4. Monte Carlo estimated statistics and standard errors (se) determined by bootstrapping for the MC simulations, where all selected input variables are uncertain R (model was run at 10 min time steps, and MC results were aggregated to 1 h averages over a 1 year period). See Appendix A for output variable nomenclature and units. Note: Interq. – interquartile; quant. – quantile; pbw – prediction band width; SD – standard deviation; var – variance; cv – coefficient of variation.

	Mean μ_{MC}	Interq. range IQR	0.005 quant. $\zeta_{0.005}$	0.995 quant. $\zeta_{0.995}$	99 % pbw $\zeta_{w, 0.99}$	0.05 quant. $\zeta_{0.05}$	0.95 quant. $\zeta_{0.95}$	90 % pbw $\zeta_{w, 0.90}$	SD σ_{MC}	var σ_{MC}^2	cv CV_{MC}
V_{Chamber} se	92.51 2.53	13.65 1.17	77.55 2.18	109.16 3.17	31.60 1.48	81.02 2.27	104.28 2.95	23.25 1.36	8.27 0.56	2984 482	0.100 0.001
V_{Sv} se	3.18 0.51	3.69 0.73	0.85 0.24	6.60 0.95	5.76 0.95	0.94 0.26	5.76 0.91	4.82 0.83	1.98 0.35	1100 259	0.070 0.012
Q_{Sv} se	55.49 5.67	76.77 10.26	0.37 0.13	267.1 13.75	266.7 14.54	1.24 0.22	165.3 10.82	164.1 10.89	64.50 4.57	7332 1,102	1.585 0.048
$B_{\text{COD},\text{Sv}}$ se	1.18 0.20	1.69 0.31	0.04 0.02	6.11 1.02	6.06 1.00	0.07 0.03	3.49 0.59	3.41 0.59	1.27 0.21	394 81.71	0.087 0.013
$B_{\text{NH}_4,\text{Sv}}$ se	0.052 0.009	0.077 0.014	0.004 0.002	0.174 0.029	0.170 0.028	0.006 0.002	0.125 0.022	0.120 0.021	0.045 0.008	0.546 0.115	0.075 0.013
$C_{\text{COD},\text{Sv},\text{av}}$ se	170.0 9.02	164.6 11.38	3.80 0.33	909.7 40.12	905.9 40.40	15.49 1.02	465.8 24.66	450.3 24.03	161.9 7.94	36 151 4,615	0.988 0.016
$C_{\text{NH}_4,\text{Sv},\text{av}}$ se	7.19 0.41	6.65 0.61	0.47 0.02	29.20 1.23	28.74 1.22	0.86 0.06	17.51 0.91	16.64 0.87	5.66 0.29	46.93 6.62	0.815 0.016

Low standard errors (se) for the coefficient of variation were obtained for all seven outputs. These were never greater than 0.05, which indicates that the selected MC replication size (1500 for mc1) is a suitable value. This holds for all output statistics because, in all cases, the standard error is small in comparison to the estimated value.

3.3.2 Contributions of input variables to total uncertainty

The contributions of input variables to the total uncertainty of the model inputs were also computed using the procedure described in Sect. 2.6.4. A total of four MC simulations with a total of 6000 runs were performed to estimate S_i (Eq. 10). Afterwards, four contributions were evaluated per time step and aggregated for the whole year. Following Eq. (10), the per time step contribution of input variables to output variables in terms of percentage of variance, stochastic sensitivity S_i of the input variables $C_{\text{COD},\text{S}}$, $C_{\text{NH}_4,\text{S}}$, COD_r , and P were calculated. An example of the contributions analysis per

time step is presented in Fig. 4. Here we remark that a high uncertainty over time is shown mainly for the spring event.

The aggregated-over-time contributions of input variables to output variables in terms of percentage of variance and stochastic sensitivity S_i of the input variables were also calculated (Table 5). Note that P is the only source of uncertainty for V_{Chamber} and V_{Sv} , while uncertainty in NH_4 inputs only propagates to NH_4 outputs, which is similar for COD (Fig. 4).

We found, as expected, that precipitation, P , is the only source of uncertainty from all uncertain input considered for water quantity output variables V_{Chamber} and V_{Sv} . Regarding average values for the whole year, for the water quality output variables $B_{\text{COD},\text{Sv}}$ and $C_{\text{COD},\text{Sv},\text{av}}$, $C_{\text{COD},\text{S}}$ has the largest contribution to the output variance, which is about 66 % for $B_{\text{COD},\text{Sv}}$ and about 62 % for $C_{\text{COD},\text{Sv},\text{av}}$. The second variable that contributes to the uncertainty of these COD output variables is P , with about 3 % for $B_{\text{COD},\text{Sv}}$ and 9 % for $C_{\text{COD},\text{Sv},\text{av}}$. Similarly, the input variable $C_{\text{NH}_4,\text{S}}$ plays an important role in the contribution of total uncertainty for

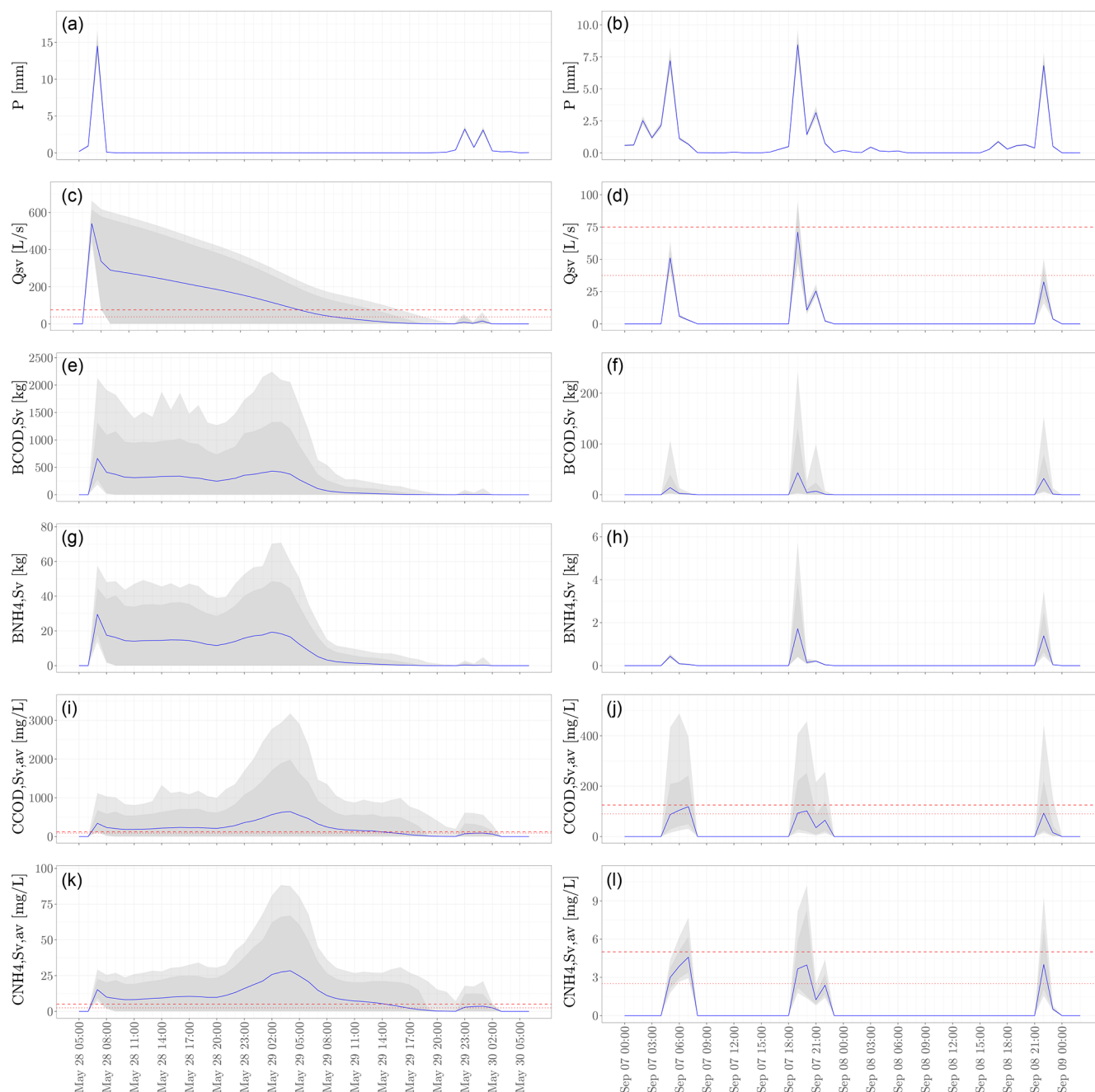


Figure 3. Uncertainty propagation outcomes for the first Monte Carlo simulation, where all input variables vary stochastically. The 99 % prediction interval is shown as a light grey shading, the 90 % prediction interval as a dark grey shading, and the mean value as a blue line. The MC simulations were performed for the entire year of 2010 at a 10 min time step, aggregated to hourly time steps in the figure. **(a, b)** Input precipitation. **(c, d)** In the overflow spill flow, the upper dashed red line indicates the 75 L s^{-1} threshold and the lower dotted red line the 37.5 L s^{-1} threshold. **(e, f)** Load of overflow COD. **(g, h)** Load of overflow NH_4 . **(i, j)** In the average spill COD concentration, the upper dashed red line indicates the 125 mg L^{-1} threshold, and the lower dotted red line indicates the 90 mg L^{-1} threshold. **(k, l)** In the average spill NH_4 concentration, the upper dashed red line indicates the 5.0 mg L^{-1} threshold and lower dotted red line the 2.5 mg L^{-1} threshold.

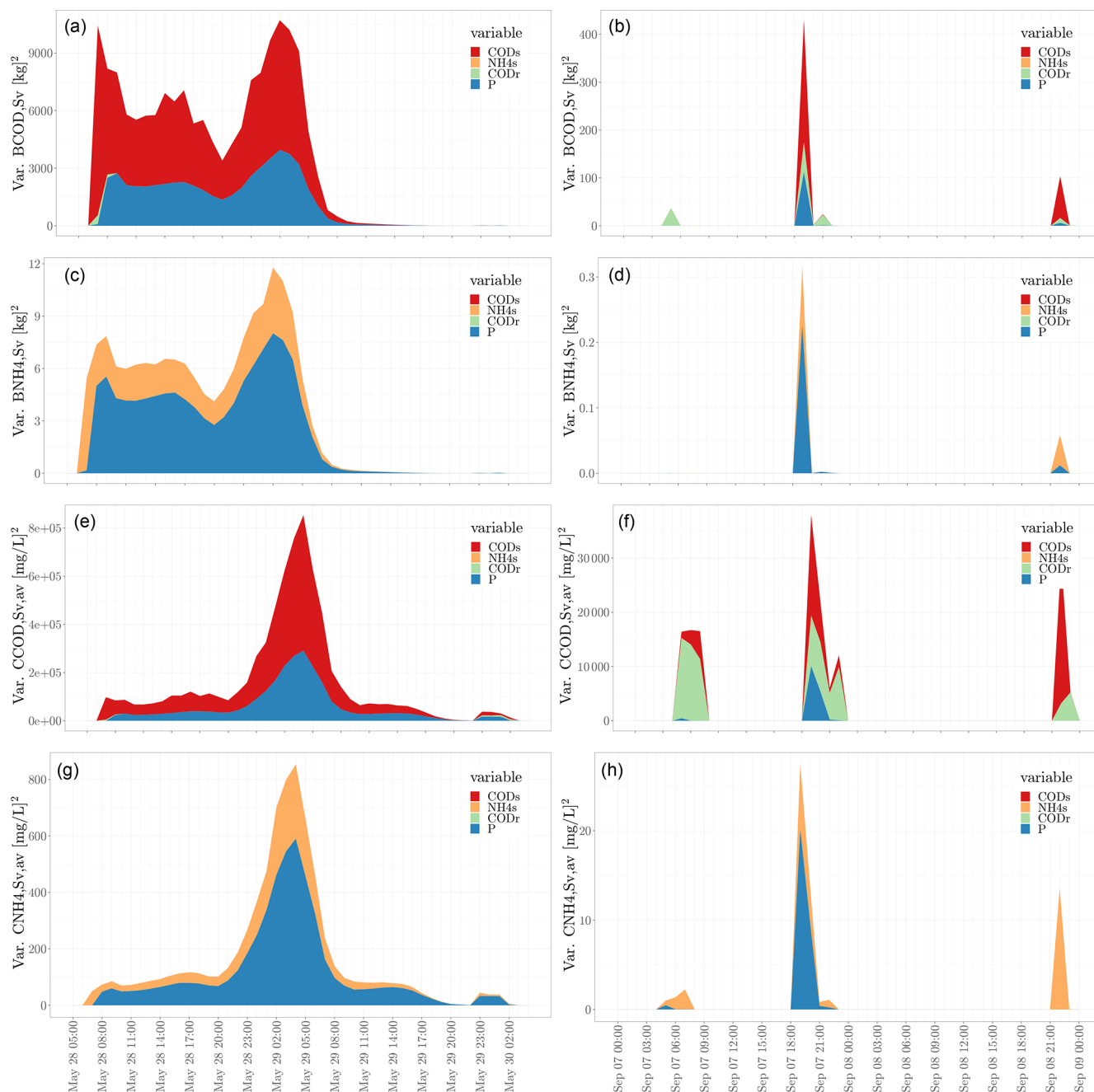


Figure 4. (a, b) Temporal contributions of input variables to the load of overflow COD, (c, d) load of overflow NH_4 , (e, f) concentration of overflow COD, and (g, h) concentration of overflow NH_4 in terms of variance. The MC simulations were performed for the entire year of 2010 at a 10 min time step, which were aggregated to hourly time steps. For illustration, two periods are shown from 28 to 30 May 2010 (left) and from 7 to 9 September 2010 (right). In the legend, CODs refers to $C_{\text{COD},S}$, and NH_4s refers to $C_{\text{NH}_4,S}$.

$B_{\text{NH}_4, Sv}$ (on average about 34 % of the variance for the whole year) and $C_{\text{NH}_4, Sv, av}$ (about 35 %). Equally contributing to the uncertainty of these NH_4 output variables is P with about 66 % for $B_{\text{NH}_4, Sv}$ and 65 % for $C_{\text{NH}_4, Sv, av}$. From these results, we can infer that precipitation is a main source of uncertainty for all six outputs considered.

3.4 Uncertainty and water quality impact

Quantification and assessment of the water quality impact is an important step after the uncertainty propagation. As described in Sect. 2.7, the assessment of water quality standards was done by taking into account the reference thresholds recommended in the European Union guidelines for COD and

Table 5. Aggregated-over-time contribution of input variables to output variables in terms of percentage of total variance.

Output variable	Stochastic sensitivity, S_i , of input variable (%)				
	Total	$C_{\text{COD}, S}$	$C_{\text{NH}_4, S}$	COD_r	P
V_{Chamber}	100.0	0.0	0.0	0.0	100.0
V_{Sv}	100.0	0.0	0.0	0.0	100.0
$B_{\text{COD}, \text{Sv}}$	100.0	65.7	0.0	2.9	31.4
$C_{\text{COD}, \text{Sv}, \text{av}}$	100.0	62.4	0.0	8.7	28.9
$B_{\text{NH}_4, \text{Sv}}$	100.0	0.0	34.4	0.0	65.6
$C_{\text{NH}_4, \text{Sv}, \text{av}}$	100.0	0.0	35.3	0.0	64.7

the German and Austrian guidelines for hydraulic impact and acute ammonium toxicity.

3.4.1 Hydraulic impact

From the time series of daily values for 2006 to 2013 of the river Sûre, a daily flow expected with a return period of once per year (1.01 years), Q_{r1} of $16 \text{ m}^3 \text{ s}^{-1}$ was computed at Heiderscheidergrund, which corresponds with the entire catchment area of the Haute-Sûre storm water system (182.1 ha). Therefore, we estimated the river daily flow in the Goesdorf CSO structure as a proportion to 30 ha, which is equal to $2.6 \text{ m}^3 \text{ s}^{-1}$. Following Eq. (11), the maximum sewer overflow discharge with return period of 1 year, Q_1 can have a value between 0.26 and $1.32 \text{ m}^3 \text{ s}^{-1}$. Accordingly, with the German guideline ATV-DVWK-A 128 (1992; Eq. 12), two additional thresholds are defined for the maximum sewer overflow discharge with return period of 1 year for the Goesdorf catchment ($A_{\text{imp}} = 5.0 \text{ ha}$). Q_1 is expected to vary between 37.5 and 75.0 L s^{-1} . We contrasted these values with those obtained from the uncertainty analysis. From Table 4, we obtained a 1 h mean value for the overflow spill flow, Q_{Sv} , of 55.5 L s^{-1} , 90 % prediction band width of 164.1 L s^{-1} , and standard deviation of 64.5 L s^{-1} . Figure 3a and b present the overflow spill flow for the two periods chosen for illustration. The upper dashed red line indicates the 75 L s^{-1} threshold, and the lower dotted red line indicates the 37.5 L s^{-1} threshold. Table 6 (top) shows the exceedance percentage of overflow spill flow over the 37.5 and 75.0 L s^{-1} thresholds for the mean, 0.95 quantile, and 0.995 quantile. We found a 0.49 % exceedance of the mean value over the 37.5 L s^{-1} threshold and about 1.7 % for the quantiles. As expected, slightly lower percentages were found for the 75.0 L s^{-1} threshold.

3.4.2 COD concentration

A reference COD concentration emission in CSOs was presented in Sect. 2.7.2. For the European Union, a value of 125 mg L^{-1} is used. We obtained a 1 h average spill COD concentration with a mean of 170 mg L^{-1} , standard deviation of 162 mg L^{-1} , and a 90 % prediction band width of 450 mg L^{-1} . Figure 3i and j present the average spill COD concentration. The upper dashed red line indicates

the 125 mg L^{-1} threshold and the lower dotted red line the 90 mg L^{-1} threshold. The mean COD concentration in the overflow volume was higher than the thresholds. However, note that when entering the river system it will quickly be diluted, suggesting that the negative impact on the environment will be dampened by the receiving water body.

Table 6 (centre row) shows the exceedance percentage of overflow COD concentration over the 90 and 125 mg L^{-1} thresholds for the mean, 0.95 quantile, and 0.995 quantile. We found a 1.62 % exceedance of the mean value over the 90 mg L^{-1} threshold and about 1.8 % for the quantiles. Slightly lower percentages were found for the 125 mg L^{-1} threshold for the mean value (1.03 %). For the quantiles equal values were found as for the 90 mg L^{-1} threshold.

3.4.3 Acute ammonium toxicity

We compared the acute ammonium toxicity reference values presented in Sect. 2.7.3 (2.5 mg L^{-1} for the ammonium concentration calculated for a 1 h duration for salmonid streams, and for cyprinid streams, a maximum value of 5.0 mg L^{-1}) was calculated with the values we found for ammonium. An average spill NH_4 concentration with a mean of 7.19 mg L^{-1} , standard deviation of 5.66 mg L^{-1} , and 90 % prediction band width of 16.64 mg L^{-1} was obtained. Figure 3k and l show the average spill NH_4 concentration for the two periods chosen for illustration. The ammonium (NH_4) concentrations in the overflow flow are higher than the reference values, which are given for concentrations in the river.

Table 6 (bottom row) shows the exceedance percentage of overflow NH_4 concentration over the 2.5 and 5.0 mg L^{-1} thresholds for the mean, 0.95 quantile, and 0.995 quantile. We found a 1.8 % exceedance of the mean and quantile values over the 2.5 and 5.0 mg L^{-1} thresholds. A slightly lower percentage (1.1 %) was found for the 5.0 mg L^{-1} threshold with regards to mean value.

4 Discussion

This study aimed to select and characterise the main sources of input uncertainty in urban water systems, while accounting for temporal auto- and cross-correlation of uncertain model inputs, by propagating input uncertainty through the EmiStatR model and quantifying and assessing the contributions of each uncertainty source to model output uncertainty dynamically (over time). In the following discussion, we start with the uncertainty and water quality impact of the model outputs to the environment, in relation to the uncertainty analysis. Next, we discuss the accuracy of Monte Carlo analysis, followed by a discussion of other sources of uncertainty. Finally, we highlight some limitations and possible solutions to the approach used in this work.

Table 6. Frequency (percentage) over time that environmental thresholds are exceeded for different statistics of the overflow spill flow, COD, and NH_4 concentration.

Output variable	Threshold	Statistic	Exceedance percentage
Q_{Sv} (L s^{-1})	37.5	Mean	0.49
	37.5	0.95 quantile	1.71
	37.5	0.995 quantile	1.74
	75.0	Mean	0.31
	75.0	0.95 quantile	1.51
	75.0	0.995 quantile	1.72
$C_{\text{COD, Sv, av}}$ (mg L^{-1})	90.0	Mean	1.62
	90.0	0.95 quantile	1.80
	90.0	0.995 quantile	1.82
	125.0	Mean	1.03
	125.0	0.95 quantile	1.80
	125.0	0.995 quantile	1.82
$C_{\text{NH}_4, \text{Sv, av}}$ (mg L^{-1})	2.5	Mean	1.78
	2.5	0.95 quantile	1.80
	2.5	0.995 quantile	1.82
	5.0	Mean	1.05
	5.0	0.95 quantile	1.78
	5.0	0.995 quantile	1.82

4.1 Uncertainty and water quality impact

Next, we discuss how the uncertainty propagation analysis done gives additional insight regarding hydraulics, COD concentration, and the acute ammonium toxicity impact on water quality over the river Sûre due to the CSO discharges under study. After doing the uncertainty propagation analysis, we not only have the predictions of model outputs, but we also know how uncertain these are. An added value arises when we take into account the uncertainty information. For the case of the overflow spill flow, the expected model output (mean of 55.5 L s^{-1}) is below the environmental threshold of 75 L s^{-1} , but the 0.95 quantile (164.1 L s^{-1}) is very much above the threshold. This indicates that there is a considerable chance of it being above the threshold.

Regarding water quality outputs, although the mean model output for COD and NH_4 concentrations is fairly above that of the thresholds, the 0.95 quantile is 2.7 times above the mean value for COD concentration and 2.4 times above the mean value for NH_4 . Also, here we can conclude that we are not certain that we are below the threshold because there is a considerable probability that the true values are above, even though the expected value is below the threshold.

We were able to compute the water quantity and quality at the CSO outlet to the river. We found that water quality (COD and NH_4) were sometimes above the environmental threshold. Even if the expected value was below the threshold, there could still be a considerable probability that the quality was above the threshold because of the large uncer-

tainty. Therefore, policy and decision-makers and water managers need to be aware of this because, whenever concentrations are above the threshold, this may harm the environment. Nevertheless it is worth noting that we computed the concentration in the outlet of the CSO. When this spilled water enters the river, it will quickly mix with the much cleaner river water and concentrations will drop quickly, so it is only a local problem. How local it is and how the river water quality is distributed in space and time is not an easy problem to solve and requires the use of hydrological and hydraulic river models, e.g. SIMBA (IFAK, 2007) or MIKE 11 (DHI, 2017). Those models have been well developed and, for some of them, uncertainty analyses have also been done (Beven and Binley, 1992; Refsgaard, 1997; Beven and Freer, 2001; Vrugt et al., 2003a; Vrugt et al., 2008; Beven et al., 2010; Andrés-Doménech et al., 2010; Beven, 2012; Jerves-Cobo et al., 2020; Yu et al., 2020), but obviously such uncertainty analyses can only be done if the inputs and the uncertainty associated with these inputs to these models are known. One of these inputs is the inlet from CSO. That is where our paper makes a very valuable contribution because our work has quantified water quantity and quality of CSO structures, including uncertainty, and that is exactly what these river models need to be able to do an uncertainty propagation analysis. We also recognise other attempts at quantity (e.g. Sriwastava et al., 2018) and quality, especially with measurements taken at CSOs, which demonstrate that the measured water quality at the WWTP influent is expected to render a low representativity of the conditions at the CSOs (e.g. Brombach et al., 2005; Diaz-Fierros T et al., 2002). We present some comparisons with these studies in the following lines.

Sriwastava et al. (2018) apply an uncertainty propagation to a complex hydrodynamic model to quantify uncertainty in sewer overflow volume. They used MC for the uncertainty propagation and Latin hypercube sampling (LHS) as an efficient sampling scheme. Although LHS ensures a full coverage of the sample space and provides a faster convergence than simple random sampling, the LHS application in the case of dynamic model inputs (e.g. precipitation, COD, and NH_4 inputs) is not trivial, and its implementation is more complex than in the case of sampling from static variables (i.e. uncertain constants). In our study, we sampled the time series of dynamic inputs using an implementation in stUP-scales (Torres-Matallana et al., 2019; Torres-Matallana et al., 2018a).

Diaz-Fierros T et al. (2002), in a study in the city of Santiago de Compostela (northwestern Spain; population about 100 000 inhabitants), where a combined sewer system feeds to a grossly undersized wastewater treatment plant, reported an event mean concentration (Diaz-Fierros T et al., 2002; Table 4) for the output variables $C_{\text{COD, Sv, av}}$ and $C_{\text{NH}_4, \text{Sv, av}}$ of 329.1 and 8.7 mg L^{-1} , respectively. These values are larger than those found by Brombach et al. (2005) and are more in agreement with our findings, especially for the case of $C_{\text{NH}_4, \text{Sv, av}}$. Diaz-Fierros T et al. (2002) reported values of

$C_{\text{COD}, \text{SV}, \text{av}}$ as high as 1073 mg L^{-1} , which agrees with the right-hand tail of the distribution obtained in our study (i.e. a 0.995 quantile of 909.7 mg L^{-1}). Similarly, for the case of $C_{\text{NH}_4, \text{SV}, \text{av}}$, Diaz-Fierros T et al. (2002) reported values as high as 32.5 mg L^{-1} , which is comparable with the 0.995 quantile (29.20 mg L^{-1}) found in our study.

It is worth noting that, regarding measurements taken at CSOs, the measured water quality at the WWTP influent is expected to render a low representativity of the conditions at the CSOs, as reported by Diaz-Fierros T et al. (2002) and Brombach et al. (2005). Thus, when comparing model outputs with independent measurements, one should bear in mind that discrepancies between measured and predicted values are not only caused by errors in model inputs, model parameters, and model structure but are also the result of errors in the water quality measurements.

4.2 Accuracy of Monte Carlo analysis

Regarding the Monte Carlo replication size for uncertainty propagation, we presented the results in Fig. S4 for three output variables and three replications size 250, 1000, and the selected 1500 (Nash–Sutcliffe efficiency, NSE, closer to 1.0 for most of the output variables). We compute replications for 50, 100, and from 250 to 2000 at steps of 250 replications and the comparison of two equal MC runs (MC1 and MC2) with different seed for the pseudo-random number generator. The results suggest that the output variables related to COD (load and concentration) have a larger dispersion when we compare MC1 and MC2 for the same replications size. This is also reflected in the larger standard errors reported in Table 4 for, for example, the overflow COD load. Nevertheless, 1500 runs are a feasible MC replication size for running a relatively simple and fast model as EmiStatR (7.29 min in average execution time, using parallel computing and 50 cores for a time series with 4464 time steps). For a more complex full hydrodynamic model with a high computational burden, 1500 replications repeated four times to compute contributions may be not possible. Therefore, we suggest checking the intermediate results of the MC convergence test. We found that for quantity variables as the spill overflow volume and quality variables as the overflow NH_4 load in 250 replications (7.10 min in average execution time using parallel computing and three cores for a time series with 4464 time steps) per individual MC execution was enough, which makes an execution of this kind of uncertainty propagation more feasible.

Figures 3 and 4 show that there is a large uncertainty for the early May event and smaller uncertainty for the September event. This is due mainly to the presence of a large dry period before the spill event in May, i.e. a shorter dry period preceding the spill flow leads to a lower uncertainty. This finding suggest also an interaction between the antecedent dry period and the concentration of pollutants.

4.3 Other sources of uncertainty

In this work, we only looked at input uncertainty and not at parameter and model structural uncertainty. Further research can be done on those topics. Neumann (2007) address how uncertainty ranges for parameters of full-scale systems are obtained and how model structure uncertainty manifests itself and can be quantified for a performance evaluation and the design of urban water infrastructure. Moreno-Rodenas et al. (2019) also studied and depicted how a model parameter is an important source of uncertainty. They emphasised that “still, uncertainty analysis is seldom applied in practice, and the relative contribution of the individual model elements is poorly understood.”. Also, they highlighted that, after inferring the river process parameters with system measurements of flow and dissolved oxygen, combined sewer overflow pollution loads became the dominant uncertainty source along with rainfall variability. These findings agreed with our results.

Bachmann-Machnik et al. (2018) recognised that the most important parameters causing uncertainties in the sewer system model are the connected area and the storm water run-off quality. Our analysis confirms these findings, specifically regarding the storm water run-off quality. In our study, the input variable run-off COD was an important source of uncertainty with relation to the annual mean overflow COD released from the CSO.

4.4 Limitations and possible improvements

Despite the extensive temporal uncertainty propagation analysis, the approach also has some limitations which we present hereafter, addressing possible solutions in future work.

1. *Incorporation of the spatial distribution of model inputs.* Specifically for precipitation, Breinholt et al. (2012) stated that, due to a poor representation of the spatial precipitation that is measured by point gauges and the complexity of the sewer systems, large output uncertainty can be expected. We also infer that we obtained a large output uncertainty due to neglecting the inherent spatial variability in precipitation. Therefore, we suggest that further research is needed to account for spatial variability in precipitation that can bring light an understanding of how this variability impacts the output uncertainty and thus quantify it properly. This issue should also be related to the problem of change of support. When modelling precipitation, we also ignored the support effect, i.e. we ignored that the sub-catchment area is much greater than a point. Future research may address this issue of change-of-support. Studies that tackled this issue are available, e.g. Leopold et al. (2006); Wadoux et al. (2017); Cecinati et al. (2018).

2. *Linkage of sub-models and uncertainty compensation effect.* Tscheikner-Gratl et al. (2019) addressed the question as to whether there is an increase in uncertainty by linking integrated models or whether a compensation effect could take place by which overall uncertainty in key water quality parameters decreases. Some further insight into this topic could be obtained by quantifying uncertainties at sub-model level and analysing whether the uncertainty at sub-model level is greater or smaller than at the overall level. With our implementation, this is not a difficult task because EmiStatR has a stringent modular design in which it is easy to analyse outputs and their uncertainties at sub-model level.
3. *Accounting for cross-correlation between the inputs precipitation and run-off COD concentration.* It is worth noting that we did not include correlation between COD_r and P . Including such a correlation would yield a more realistic model of the uncertainty because these variables are known to have a strong correlation. It is highly recommended that correlations between COD_r and precipitation be included because loads in chemical oxygen demand are correlated with the overland flow due to precipitation, which may transport distributed pollutants to the sewer system. Also, the inputs $C_{COD,S}$ and $C_{NH_4,S}$ can be related with a daily curve that reflects the pattern of consumption in the household, like the German ATV-DVWK-A 134 curve. We used the latest version of EmiStatR (version 1.2.2.0), which considers these kinds of patterns.
4. *Absence of high frequency water quality observations to compare with model outputs and uncertainty prediction bands.* In order to gain an understanding of the temporal dynamics of nutrients (nitrogen, N, and phosphorus, P), Yu et al. (2020) applied high-frequency monitoring in a groundwater-fed, low-lying urban polder in Amsterdam (the Netherlands). They argued that although spatial and temporal concentration patterns from discrete sampling campaigns of water quality parameters, such as *E. coli*, showed a clear dilution pattern, the temporal patterns of N and P were still poorly understood, given their reactive nature and more complex biogeochemistry. Therefore, high-frequency measurement is a key factor in understanding these temporal dynamics and patterns.
5. *Absence of a joint spatio-temporal uncertainty analysis.* According to Zhou et al. (2020), the limitations in algorithms for classic uncertainty estimates is the reason that only the uncertainty in one dimension (either temporal variability or spatial heterogeneity) is considered, whereas the variation in the other dimension is dismissed, resulting in an incomplete assessment of the uncertainties. Zhou et al. (2020) also showed that classic metrics underestimate the uncertainty through averaging, which means a loss of information in the variation across spatio-temporal scales. To handle this limitation, suitable methods are the 3D variance partitioning for a new uncertainty estimation in both spatio-temporal scales (Zhou et al., 2020) or spatio-temporal geostatistics (Gräler et al., 2016).
6. *Uncertainty analysis with complex models.* In this research, we were able to conduct a comprehensive Monte Carlo uncertainty propagation analysis, which required a large number of Monte Carlo runs. This was possible because we used a strongly simplified urban water system model, EmiStatR. For more complex models that take much more computing time, the application of a Monte Carlo uncertainty propagation analysis is more challenging. However, given sufficient resources it is possible because each model run can be run independently, and hence, the analysis is extremely suitable for parallelisation and cloud computing. In particular, the use of graphics processing units (GPUs) for heavy computation is promising. Some recent examples that demonstrate the potential of GPUs for this purpose are Eränen et al. (2014), Sten et al. (2016), and Sandric et al. (2019). Sriwastava et al. (2018) applied uncertainty propagation to a complex hydrodynamic model by selecting a small subset of dominant input/model parameters that explain most of the model output variance.
7. *Convolution of precipitation to run-off.* Equation (2) indicates that the run-off that enters the combined sewage in EmiStatR has the same shape as the precipitation, with an offset. Thus, for precipitation, no convolution is applied. In a new release of EmiStatR, we will incorporate convolution of precipitation when transforming precipitation to run-off.

5 Conclusions

In this final section, we conclude by highlighting the importance of temporal uncertainty propagation analysis and the selection and characterisation of uncertain model inputs impacting model sensitivity. We also point out that uncertainty propagation analysis helps to identify the sources that contribute the most and can provide better evidence for the impact assessment of pollutant release from sewer systems to the environment, in particular to the receiving waters.

1. *Uncertainty analysis is important because it quantifies the accuracy of model outputs and quantifies the uncertainty source contributions.* The latter provides essential information to help take informed decisions about how to improve the accuracy of the model output. But MC uncertainty analysis is only possible if it is computationally feasible. We used a simplified urban water system model with the capability to apply it for the minimisation of transient pollution from urban wastewater

systems in parallel mode, which minimises model running time, allowing for uncertainty propagation, long-term simulations, and the evaluation of complex scenarios. These capabilities are crucial also for, for example, real-time control applications, where simplified models of fast running times are desirable.

2. *Input variables that were very uncertain, and for which model output was very sensitive, were selected to be included in the uncertainty propagation analysis.* We found four main input variables to be analysed, namely (1) precipitation, P , (2) chemical oxygen demand sewage pollution per capita load per day, $C_{\text{COD},s}$, (3) ammonium pollution per capita load per day, $C_{\text{NH}_4,s}$, and (4) chemical oxygen demand COD_r concentration.
3. *Selected input variables for uncertainty propagation can be characterised in terms of the input uncertainty in four specific cases, depending on the type of input variable.* (i) Uncertain constant inputs, characterised by their marginal (cumulative) pdf, e.g. water consumption, infiltration flow, impervious area and run-off coefficients; (ii) temporally autocorrelated dynamic uncertain inputs, characterised by univariate time series autoregressive modelling, e.g. COD_r ; (iii) temporally cross-correlated multiple dynamic uncertain inputs, characterised by multivariate time series modelling, considering cross- and no-correlations among variables, e.g. $C_{\text{COD},s}$ and $C_{\text{NH}_4,s}$; and (iv) rain gauge input precipitation, characterised by autoregressive model conditioned to the observed precipitation (P).
4. *Model input uncertainty propagation through the simplified, combined sewer overflow model (EmiStatR) helped us to understand how uncertainty propagates and how large the uncertainty of EmiStatR outputs is in a case study.* Three output variables were considered for water quantity and four variables for water quality. The Monte Carlo uncertainty propagation analysis showed that, among the water quantity output variables, the overflow flow, Q_{Sv} , is the more uncertain output variable and has a large coefficient of variation (cv of 1.585). Among water quality variables, the annual average spill COD concentration, $C_{\text{COD},\text{Sv},\text{av}}$, and the average spill NH_4 concentration, $C_{\text{NH}_4,\text{Sv},\text{av}}$, were found to have large uncertainty (coefficients of variation of 0.988 and 0.815, respectively). Also, low standard errors (se) for the coefficient of variation were obtained for all seven outputs. They were never greater than 0.05, which indicated that the selected MC replication size (1500 simulations) was a suitable value.
5. *Main sources of uncertainty model outputs.* For water quantity outputs it was precipitation, while for COD, water quality outputs were P , $C_{\text{COD},s}$, and COD_r , and for NH_4 , outputs were P and $C_{\text{NH}_4,s}$.
6. *Uncertainty propagation analysis can explain the impact of water quality indicators to the receiving river for the Luxembourg case study more comprehensively.* Although the mean model water quality outputs for COD and NH_4 concentrations are fairly above the thresholds, the 0.95 quantile is 2.7 times above the mean value for COD concentration and 2.4 times above the mean value for NH_4 . We conclude that we are not certain that environmental thresholds are not exceeded because there is a considerable probability that values are above them, even though the expected value is below the thresholds. This is valid for concentrations in the spilled CSO; therefore, it is important to highlight that the results confirmed our hypothesis that annual mean COD and NH_4 river concentrations are lower than the released CSO concentrations due to dilution and are thus compliant with the water quality thresholds given by the guidelines consulted.

Appendix A

Table A1. List of variables of EmiStatR.

$csmr_t$	Combined sewer mixing ratio at time t	(–)
pe_t	Population equivalents of the connected CSO structure at time t	(PE)
q_{f_t}	Infiltration water inflow flux at time t	($L s^{-1} ha^{-1}$)
qs_t	Individual water consumption of households at time t	($L PE^{-1} d^{-1}$)
t_{fS}	Delay in time response related to flow time in the sewer system	(time step)
A_{imp}	Impervious area of the catchment	(ha)
A_{total}	Total area of the catchment	(ha)
$BCOD_{Sv}$	COD load in the spill overflow	(g)
$B_{NH_4, Sv}$	Spill overflow NH_4 load	(g)
C_d	Orifice coefficient of discharge in the CSOC structure	(–)
C_{imp}	Run-off coefficient for impervious areas	(–)
C_{per}	Run-off coefficient for pervious areas	(–)
C_{COD_t}	Mean dry weather pollutant (COD) concentration at time t	($mg L^{-1}$)
$C_{COD, S}$	COD sewage pollution per capita load per day	($g PE^{-1} d^{-1}$)
$C_{COD, Sv, 99.9}$	A 99.9th percentile spill COD concentration	($mg L^{-1}$)
$C_{COD, Sv, av}$	Average spill COD concentration	($mg L^{-1}$)
$C_{COD, Sv, max}$	Maximum overflow COD concentration	($mg L^{-1}$)
$C_{NH_4, S}$	NH_4 sewage pollution per capita load per day	($g PE^{-1} d^{-1}$)
$C_{NH_4, Sv, 99.9}$	A 99.9th percentile spill NH_4 concentration	($mg L^{-1}$)
$C_{NH_4, Sv, av}$	Average spill NH_4 concentration	($mg L^{-1}$)
$C_{NH_4, Sv, max}$	Maximum spill NH_4 concentration	($mg L^{-1}$)
COD_f	Infiltration pollution COD	($g PE^{-1} d^{-1}$)
COD_{r_t}	Concentration due to rainwater pollution (COD) at time t	($mg L^{-1}$)
D_d	Orifice diameter	(m)
Lev	Measured water level in the CSOC	(m)
Lev_{ini}	Initial water level in the CSOC	(m)
NH_{4f}	Infiltration pollution NH_4	($g PE^{-1} d^{-1}$)
NH_{4r}	Concentration due to rainwater pollution (NH_4)	($mg L^{-1}$)
P_t	Precipitation at time t	($mm min^{-1}$)
$Q_{d, max}$	Maximum throttled outflow	($L s^{-1}$)
Q_{f_t}	Infiltration flow at time t	($L s^{-1}$)
Q_{r_t}	Flow contribution of rainwater to the combined sewage flow at time t	($m^3 s^{-1}$)
Q_{s_t}	Dry weather flow of the residential sewage in the catchment at time t	($L s^{-1}$)
Q_{CSF_t}	Combined sewer flow at time t	($m^3 s^{-1}$)
Q_{DWF_t}	Total dry weather flow at time t	($L s^{-1}$)
Q_{Sv}	Overflow spill flow	($L s^{-1}$)
V	Volume of the CSOC structure	(m^3)
V_{d_t}	Volume of throttled outflow to the WWTP at time t	(m^3)
V_{dw_t}	Total dry weather volume (amount of dry weather water in combined sewage flow) at time t	(m^3)
V_{r_t}	Rain weather volume at time t	(m^3)
$V_{Chamber_t}$	Water volume in the CSO chamber at time t	(m^3)
V_{Sv_t}	Spill overflow volume at time t	(m^3)
ϵ	numerical precision term	(m^3)
Δt	Time interval	(min)

Code and data availability. The code scripts and data sets related to Figs. 3, 4, S3, and S4 are available from Zenodo (<https://doi.org/10.5281/zenodo.3928079>; Torres-Matallana, 2020) and GitHub (https://github.com/ArturoTorres/temporal_uncertainty_paper_reproducible.git, last access: 29 December 2020).

Supplement. The supplement related to this article is available online at: <https://doi.org/10.5194/hess-25-193-2021-supplement>.

Author contributions. JATM is the main author of the text in this article. UL and GBMH contributed to this article with their statistical, geostatistical, and programming knowledge and reviewed and edited the text. JATM developed all R code scripts for the computations and Monte Carlo simulations and performed the simulations and analysis in collaboration with UL and GBMH.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. We thank the Luxembourgish Administration des Services Techniques de l'Agriculture (ASTA) for the rainfall time series and the Observatory for Climate and Environment (OCE) of LIST for the technical support. We thank Kai Klepizewski for his advice and involvement during early stages of this research and the two anonymous reviewers for the constructive and valuable comments that helped us to improve this paper.

Financial support. This research received funding from the European Union's Seventh Framework Programme for research, technological development, and demonstration (Quantifying Uncertainty in Integrated Catchment Studies, QUICS, project; grant no. 607000), and the Luxembourg Institute of Science and Technology (LIST).

Review statement. This paper was edited by Nadav Peleg and reviewed by two anonymous referees.

References

- Andrés-Doménech, I., Múnera, J. C., Francés, F., and Marco, J. B.: Coupling urban event-based and catchment continuous modelling for combined sewer overflow river impact assessment, *Hydrol. Earth Syst. Sci.*, 14, 2057–2072, <https://doi.org/10.5194/hess-14-2057-2010>, 2010.
- Bach, P. M., Rauch, W., Mikkelsen, P. S., McCarthy, D. T., and Deletic, A.: A critical review of integrated urban water modelling – Urban drainage and beyond, *Environ. Model. Softw.*, 54, 88–107, 2014.
- Bachmann-Machnik, A., Meyer, D., Waldhoff, A., Fuchs, S., and Dittmer, U.: Integrating retention soil filters into urban hydrologic models – Relevant processes and important parameters, *J. Hydrol.*, 559, 442–453, <https://doi.org/10.1016/j.jhydrol.2018.02.046>, 2018.
- Baker, L. A. (Ed.): *The Water Environment of Cities*, Springer Science+Business Media, LLC, 233 Spring Street, New York, NY 10013, USA, 2009.
- Barbosa, S. M.: Package “mAr”: Multivariate AutoRegressive analysis, The Comprehensive R Archive Network, CRAN, 1.1-2 edn., 2015.
- Bastin, L., Cornford, D., Jones, R., Heuvelink, G. B. M., Pebesma, E., Stasch, C., Nativi, S., Mazzetti, P., and Williams, M.: Managing uncertainty in integrated environmental modelling: The UncertWeb framework, *Environ. Model. Softw.*, 39, 116–134, 2013.
- Beven, K. and Binley, A.: The future of distributed models: model calibration and uncertainty prediction, *Hydrol. Process.*, 6, 279–98, 1992.
- Beven, K. and Freer, J.: Equifinality, data assimilation, and uncertainty estimation in mechanistic modelling of complex environmental systems using the GLUE methodology, *J. Hydrol.*, 249, 11–29, [https://doi.org/10.1016/S0022-1694\(01\)00421-8](https://doi.org/10.1016/S0022-1694(01)00421-8), 2001.
- Beven, K., Leedal, D., and Alcock, R.: Uncertainty and Good Practice in Hydrological Prediction, *VATTEN*, 66, 159–163, 2010.
- Beven, K. J.: *Rainfall-Runoff Modelling: The Primer*, Wiley-Blackwell, second edn., Lancaster University, UK, 2012.
- Blumensaat, F., Staufer, P., Heusch, S., Reussner, F., Schütze, M., Seiffert, S., Gruber, G., Zawilski, M., and Rieckermann, J.: Water quality-based assessment of urban drainage impacts in Europe where do we stand today?, *Water Sci. Technol.*, 66, 304–313, <https://doi.org/10.2166/wst.2012.178>, 2012.
- Boos, D., Matthew, K., and Osborne, J.: MonteCarlo.se: Monte Carlo Standard Errors, available at: <https://CRAN.R-project.org/package=MonteCarlo.se> (last access: 29 December 2020), R package version 0.1.0, 2019.
- Boos, D. D.: Introduction to the Bootstrap and World, *Stat. Sci.*, 18, 168–174, 2003.
- Boos, D. D. and Osborne, J. A.: Assessing Variability of Complex Descriptive Statistics in Monte Carlo Studies Using Resampling Methods, *Int. Stat. Rev.*, 83, 228–238, <https://doi.org/10.1111/insr.12087>, 2015.
- Breinholt, A., Møller, J. K., Madsen, H., and Mikkelsen, P. S.: A formal statistical approach to representing uncertainty in rainfall-runoff modelling with focus on residual analysis and probabilistic output evaluation – Distinguishing simulation and prediction, *J. Hydrol.*, 472–473, 36–52, 2012.
- Brombach, H., Weiss, G., and Fuchs, S.: A new database on urban runoff pollution: comparison of separate and combined sewer systems, *Water Sci. Technol.*, 51, 119–128, <https://doi.org/10.2166/wst.2005.0039>, 2005.
- Brown, J. D.: Knowledge, uncertainty and physical geography: Towards the development of methodologies for questioning belief, *T. I. Brit. Geogr.*, 29, 367–381, <https://doi.org/10.1111/j.0020-2754.2004.00342.x>, 2004.
- Cecinati, F., Moreno-Ródenas, A. M., Rico-Ramirez, M. A., ten Veldhuis, M. C., and Langeveld, J. G.: Considering rain gauge uncertainty using kriging for uncertain data, *Atmosphere*, 9, 1–17, <https://doi.org/10.3390/atmos9110446>, 2018.
- Committee on Environmental and Natural Resources – CENR: An Assessment of Coastal Hypoxia and Eutrophication in U.S. Wa-

- ters, Tech. rep., NSTC, National Science and Technology Council, 725 17th Street, Washington, D.C., USA, 2003.
- Datta, A. R.: Evaluation of Implicit and Explicit Methods of Uncertainty Analysis on a Hydrological Modeling, Ph.D. thesis, University of Windsor, Canada, 2011.
- Deletic, A., Dotto, C., McCarthy, D., Kleidorfer, M., Freni, G., Mannina, G., Uhl, M., Henrichs, M., Fletcher, T., Rauch, W., Bertrand-Krajewski, J., and Tait, S.: Assessing uncertainties in urban drainage models, *Phys. Chem. Earth.*, 42–44, 3–10, <https://doi.org/10.1016/j.pce.2011.04.007>, 2012.
- DHI: MIKE11, A modeling system for rivers and channels, Reference Manual, DHI Water and Environment, Danish Hydraulic Institute, DHI, Hørsholm, Denmark, 2017.
- Díaz-Fierros, T., F., Puerta, J., Suarez, J., and Díaz-Fierros V., F.: Contaminant loads of CSOs at the wastewater treatment plant of a city in NW Spain, *Urban Water*, 4, 291–299, [https://doi.org/10.1016/S1462-0758\(02\)00020-1](https://doi.org/10.1016/S1462-0758(02)00020-1), 2002.
- Efron, B.: Bootstrap methods: Another look at the Jackknife, *Ann. Stat.*, 7, 1–26, 1979.
- Eränen, D., Oksanen, J., Westerholm, J., and Sarjakoski, T.: A full graphics processing unit implementation of uncertainty-aware drainage basin delineation, *Comput. Geosci.*, 73, 48–60, <https://doi.org/10.1016/j.cageo.2014.08.012>, 2014.
- Evers, P., Heinz, H., Hanitsch, P. H., Koch, G., Naupold, L., Tochtermann, W., Tornow, M., Zander, B., Mahret, H., and Warnow, D.: ATV-DVWK-A 134E: Planning and Construction of Wastewater Pumping Stations, Tech. rep., DWA, Germany, 2000.
- Gasperi, J., Zgheib, S., Cladière, M., Rocher, V., Moilleron, R., and Chebbo, G.: Priority pollutants in urban stormwater: Part 2 – Case of combined sewers, *Water Res.*, 46, 6693–6703, <https://doi.org/10.1016/j.watres.2011.09.041>, 2012.
- Gräler, B., Pebesma, E., and Heuvelink, G.: Spatio-Temporal Interpolation using {gstat}, *R J.*, 8, 204–218, 2016.
- Hager, W. H.: Wastewater hydraulics, second edn., <https://doi.org/10.1080/00221686.2011.614723>, Springer-Verlag, Berlin, Heidelberg, Germany, 2010.
- Hammersley, J. and Handscomb, D.: Monte Carlo Methods, Methuen & Co Ltd, London, 1964.
- Heip, L., Assel, J. V., and Swartentenbroekx, P.: Sewer flow quality modelling, *Water Sci. Technol.*, 36, 177–184, 1997.
- Heuvelink, G. B. M.: Error Propagation in Environmental Modelling with GIS, Reserach Monographs in GIS, CRC Press Taylor & Francis Group, Taylor & Francis Ltd. 11 New Fetter Lane, London EC4P 4EE, UK, 1998.
- Heuvelink, G. B. M., Brown, J. D., and van Loon, E. E.: A probabilistic framework for representing and simulating uncertain environmental variables, *Int. J. Geogr. Inf. Sci.*, 21, 497–513, <https://doi.org/10.1080/13658810601063951>, 2007.
- House, M. A., Ellis, J. B., Herricks, E. E., Hvitved-Jacobsen, T., Seager, J., Lijklema, L., Aalderink, H., and Clifford, I. T.: Urban Drainage – Impacts on Receiving Water Quality, *Water Sci. Technol.*, 27, 117–158, <https://doi.org/10.2166/wst.1993.0293>, 1993.
- Huang, H., Xiao, X., Yan, B., and Yang, L.: Ammonium removal from aqueous solutions by using natural Chinese (Chende) zeolite as adsorbent, *J. Hazard. Mater.*, 175, 247–252, <https://doi.org/10.1016/j.jhazmat.2009.09.156>, 2010.
- Hutton, C., Vamvakieridou-Lyroudia, L., Kapelan, Z., and Savic, D.: Uncertainty Quantification and Reduction in Urban Water Systems (UWS) Modelling: Evaluation Report, Tech. rep., The European Commission, 2011.
- IFAK: SIMBA (Simulation of Biological Wastewater Systems): Manual and Reference, Tech. rep., Institut für Automation und Kommunikation e. V., Magdeburg, Germany, 2007.
- Jerves-Cobo, R., Benedetti, L., Amerlinck, Y., Lock, K., De Mulder, C., Van Butsel, J., Cisneros, F., Goethals, P., and Nopens, I.: Integrated ecological modelling for evidence-based determination of water management interventions in urbanized river basins: Case study in the Cuenca River basin (Ecuador), *Sci. Total Environ.*, 709, 1–18, <https://doi.org/10.1016/j.scitotenv.2019.136067>, 2020.
- Kalman, R. E.: A new approach to linear filtering and prediction problems, *Transactions of the American Society of Mechanical Engineers: Journal of Basic Engineering*, 82D, 35–45, 1960.
- Kalos, M. H. and Whitlock, P. A.: Monte Carlo Methods, 2 edn., WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim, Germany, 2008.
- Katukiza, A. Y., Ronteltap, M., Niwagaba, C. B., Kansime, F., and Lens, P. N. L.: Grey water characterisation and pollutant loads in an urban slum, *Int. J. Environ. Sci. Technol.*, 12, 423436, <https://doi.org/10.1007/s13762-013-0451-5>, 2014.
- Kleidorfer, M. and Rauch, W.: An application of Austrian legal requirements for CSO emissions, *Water Sci. Technol.*, 64, 1081, <https://doi.org/10.2166/wst.2011.560>, 2011.
- Kuczera, G. and Parent, E.: Monte Carlo assessment of parameter uncertainty in conceptual catchment models: The Metropolis algorithm, *J. Hydrol.*, 211, 69–85, [https://doi.org/10.1016/S0022-1694\(98\)00198-X](https://doi.org/10.1016/S0022-1694(98)00198-X), 1998.
- Leopold, U., Heuvelink, G. B. M., Tiktak, A., Finke, P. A., and Schoumans, O.: Accounting for change of support in spatial accuracy assessment of modelled soil mineral phosphorous concentration, *Geoderma*, 130, 368–386, 2006.
- Litrico, X. and Fromion, V.: Modeling and Control of Hydroystems, Springer-Verlag, London, United Kingdom, <https://doi.org/10.1017/CBO9781107415324.004>, 2009.
- Luetkepohl, H.: New Introduction to Multiple Time Series Analysis, Springer-Verlag, Berlin, Heidelberg, Germany, 2005.
- Marwick, B. and Krishnamoorthy, K.: cvequality: Tests for the Equality of Coefficients of Variation from Multiple Groups, available at: <https://github.com/benmarwick/cvequality> (last access: 29 December 2020), r package version 0.2.0, 2019.
- McMillan, H., Jackson, B., Clark, M., Kavetski, D., and Woods, R.: Rainfall uncertainty in hydrological modelling: An evaluation of multiplicative error models, *J. Hydrol.*, 400, 8394, <https://doi.org/10.1016/j.jhydrol.2011.01.026>, 2011.
- Miskewitz, R. and Uchir, C.: In-Stream Dissolved Oxygen Impacts and Sediment Oxygen Demand Resulting from Combined Sewer Overflow Discharges, *J. Environ. Eng.*, 139, 1307–1313, [https://doi.org/10.1061/\(ASCE\)JEE.1943-7870.0000739](https://doi.org/10.1061/(ASCE)JEE.1943-7870.0000739), 2013.
- Moreno-Rodenas, A. M., Tschekner-Gratl, F., Langeveld, J. G., and Clemens, F. H.: Uncertainty analysis in a large-scale water quality integrated catchment modelling study, *Water Res.*, 158, 46–60, <https://doi.org/10.1016/j.watres.2019.04.016>, 2019.
- Neumann, M. B.: Uncertainty Analysis for Performance Evaluation and Design of Urban Water Infrastructure, Ph.D. thesis, Swiss Federal Institute of Technology, ETH Zurich, Switzerland, 2007.

- Nol, L., Heuvelink, G. B. M., A. Veldkamp, de Vries, W., and Kros, J.: Uncertainty propagation analysis of an N₂O emission model at the plot and landscape scale, *Geoderma*, 159, 9–23, 2010.
- R-Core-Team and contributors worldwide: The R Stats Package, The R Project for Statistical Computing, 3.5.0 edn., available at: <https://stat.ethz.ch/R-manual/R-devel/library/stats/html/00Index.html> (last access: 29 December 2020), 2017.
- Refsgaard, J. C.: Parameterisation, calibration and validation of distributed hydrological models, *J. Hydrol.*, 198, 69–97, [https://doi.org/10.1016/S0022-1694\(96\)03329-X](https://doi.org/10.1016/S0022-1694(96)03329-X), 1997.
- Refsgaard, J. C., van der Sluijs, J. P., Højberg, A. L., and Vanrolleghem, P. A.: Uncertainty in the environmental modelling process - A framework and guidance, *Environ. Model. Softw.*, 22, 1543–1556, <https://doi.org/10.1016/j.envsoft.2007.02.004>, 2007.
- Renard, B., Kavetski, D., Kuczera, G., Thyer, M., and Franks, S. W.: Understanding predictive uncertainty in hydrologic modeling: The challenge of identifying input and structural errors, *Water Resour. Res.*, 46, 1–22, <https://doi.org/10.1029/2009wr008328>, 2010.
- Sandric, I., Ionita, C., Chitu, Z., Dardala, M., Irimia, R., and Furtuna, F. T.: Using CUDA to accelerate uncertainty propagation modelling for landslide susceptibility assessment, *Environ. Model. Softw.*, 115, 176–186, <https://doi.org/10.1016/j.envsoft.2019.02.016>, 2019.
- Sriwastava, A. K., Tait, S., Schellart, A., Kroll, S., Dorpe, M. V., Assel, J. V., and Shucksmith, J.: Quantifying Uncertainty in Simulation of Sewer Overflow Volume, *J. Environ. Eng.*, 144, 04018050, [https://doi.org/10.1061/\(ASCE\)EE.1943-7870.0001392](https://doi.org/10.1061/(ASCE)EE.1943-7870.0001392), 2018.
- Statec: Statistics Portal of the Grans-Duchy of Luxembourg, available at: <https://statistiques.public.lu>, last access: 29 December 2020.
- Steinel, A. and Margane, A.: Best management practice guideline for wastewater facilities in karstic areas of Lebanon with special respect to the protection of ground- and surface waters, Tech. Rep. 2, Federal Ministry for Economic Cooperation and Development, Bundesanstalt für Geowissenschaften und Rohstoffe, BGR, Hannover, Germany, 2011.
- Sten, J., Lilja, H., Hyvältuoma, J., Westerholm, J., and Aspnäs, M.: Parallel flow accumulation algorithms for graphical processing units with application to RUSLE model, *Comput. Geosci.*, 89, 88–95, <https://doi.org/10.1016/j.cageo.2016.01.006>, 2016.
- Toffol, S. D.: Sewer system performance assessment – an indicators based methodology, Ph.D. thesis, Universität Innsbruck, Innsbruck, Austria, 2006.
- Torres-Matallana, J., Leopold, U., and Heuvelink, G.: stUP-scales: an R-package for spatio-temporal Uncertainty Propagation across multiple scales with examples in urban water modelling, *Water*, 10, 1–30, <https://doi.org/10.3390/w10070837>, 2018a.
- Torres-Matallana, J., Leopold, U., and Heuvelink, G.: stUP-scales: Spatio-Temporal Uncertainty Propagation Across Multiple Scales, available at: <https://CRAN.R-project.org/package=stUPscales> (last access: 29 December 2020), R package version 1.0.5.0, 2019.
- Torres-Matallana, J. A., Leopold, U., and Heuvelink, G. B. M.: Multivariate autoregressive modelling and conditional simulation of precipitation time series for urban water models, *European Water*, 57, 299–306, 2017.
- Torres-Matallana, J. A., Klepiszewski, K., Leopold, U., and Heuvelink, G.: EmiStatR: a simplified and scalable urban water quality model for simulation of combined sewer overflows, *Water*, 10, 1–24, <https://doi.org/10.3390/w10060782>, 2018b.
- Torres-Matallana, J. A.: R code and data to reproduce figures from the “Multivariate autoregressive modelling and conditional simulation for temporal uncertainty analysis of an urban water system in Luxembourg” paper, (Version v1.2) [Data set], Zenodo, <https://doi.org/10.5281/zenodo.3928079>, 2020.
- Tscheikner-Gratl, F., Lepot, M., Moreno-Rodenas, A., and Schellart, A.: QUICS Deliverable 6.7: A Framework for the application of uncertainty analysis, Tech. rep., Delft University of Technology, University of Sheffield and CH2M, Zenodo, <https://doi.org/10.5281/zenodo.1240926>, 2017.
- Tscheikner-Gratl, F., Bellos, V., Schellart, A., Moreno-Rodenas, A., Muthusamy, M., Langeveld, J., Clemens, F., Benedetti, L., Rico-Ramirez, M. A., de Carvalho, R. F., Breuer, L., Shucksmith, J., Heuvelink, G. B., and Tait, S.: Recent insights on uncertainties present in integrated catchment water quality modelling, *Water Res.*, 150, 368–379, <https://doi.org/10.1016/j.watres.2018.11.079>, 2019.
- van der Keur, P., Henriksen, H. J., Refsgaard, J. C., Brugnach, M., Pahl-Wostl, C., Dewulf, A., and Buiteveld, H.: Identification of major sources of uncertainty in current IWRM practice. Illustrated for the Rhine Basin, *Water Resour. Manage.*, 22, 1677–1708, <https://doi.org/10.1007/s11269-008-9248-6>, 2008.
- Viana da Silva, A. M. E., Bettencourt da Silva, R. J. N., and Camões, M. F. G. F. C.: Optimization of the determination of chemical oxygen demand in wastewaters, *Anal. Chim. Acta*, 699, 161–169, <https://doi.org/10.1016/j.aca.2011.05.026>, 2011.
- Vrugt, J. A. and Robinson, B. A.: Improved evolutionary optimization from genetically adaptive multimethod search, *P. Natl. Acad. Sci. USA*, 104, 708–711, <https://doi.org/10.1073/pnas.0610471104>, 2007.
- Vrugt, J. A., Gupta, H. V., Bastidas, L. A., Bouten, W., and Sorooshian, S.: Effective and efficient algorithm for multiobjective optimization of hydrologic models, *Water Resour. Res.*, 39, SWC 5-1–SWC 5-19, <https://doi.org/10.1029/2002wr001746>, 2003a.
- Vrugt, J. A., Gupta, H. V., Bouten, W., and Sorooshian, S.: A Shuffled Complex Evolution Metropolis algorithm for optimization and uncertainty assessment of hydrologic model parameters, *Water Resour. Res.*, 39, SWC 1-1–SWC 1-14, 2003b.
- Vrugt, J. A., ter Braak, C. J. F., Clark, M. P., Hyman, J. M., and Robinson, B. A.: Treatment of input uncertainty in hydrologic modeling: Doing hydrology backward with Markov chain Monte Carlo simulation, *Water Resour. Res.*, 44, 1–15, <https://doi.org/10.1029/2007WR006720>, 2008.
- Wadoux, A.-C., Brus, D., Rico-Ramirez, M., and Heuvelink, G.: Sampling design optimisation for rainfall prediction using a non-stationary geostatistical model, *Adv. Water Res.*, 107, 126–138, <https://doi.org/10.1016/j.advwatres.2017.06.005>, 2017.
- Walker, W., Harremoës, P., Rotmans, J., van der Sluijs, J., van Asselt, M., Janssen, P., and Krayen von Krauss, M.: Defining Uncertainty: A Conceptual Basis for Uncertainty Management in Model-Based Decision Support, *Integrated Assessment*, 4, 5–17, <https://doi.org/10.1076/iaij.4.1.5.16466>, 2003.

- Webster, R. and Heuvelink, G. B. M.: The Kalman filter for the pedologist's tool kit, *Eur. J. Soil Sci.*, 57, 758773, <https://doi.org/10.1111/j.1365-2389.2006.00879.x>, 2006.
- Webster, R. and Oliver, M.: *Geostatistics for Environmental Scientists*, 2nd edn., Wiley, Hoboken, New Jersey, USA, 2007.
- Welker, A.: Emissions of pollutant loads from combined sewer systems and separate sewer systems – Which sewer system is better, in: 11th International Conference on Urban Drainage, edited by: ICUD, Edinburgh, Scotland, UK, IAHR/IWA Joint Committee on Urban Drainage, 2008.
- Yu, L., Rozemeijer, J. C., Broers, H. P., van Breukelen, B. M., Middelburg, J. J., Ouboter, M., and van der Velde, Y.: Drivers of nitrogen and phosphorus dynamics in a groundwater-fed urban catchment revealed by high frequency monitoring, *Hydrol. Earth Syst. Sci. Discuss.*, <https://doi.org/10.5194/hess-2020-34>, in review, 2020.
- Zhou, X., Polcher, J., Yang, T., and Huang, C.-S.: A new uncertainty estimation approach with multiple datasets and implementation for various precipitation products, *Hydrol. Earth Syst. Sci.*, 24, 2061–2081, <https://doi.org/10.5194/hess-24-2061-2020>, 2020.
- Zoppou, C.: Review of urban storm water models, *Environ. Model. Softw.*, 16, 195–231, [https://doi.org/10.1016/S1364-8152\(00\)00084-0](https://doi.org/10.1016/S1364-8152(00)00084-0), 2001.