



# Identification of behavioural model input data sets for WWTP uncertainty analysis

E. Lindblom , U. Jeppsson  and G. Sin 


## ABSTRACT

Uncertainty analysis is important for wastewater treatment plant (WWTP) model applications. An important aspect of uncertainty analysis is the identification and proper quantification of sources of uncertainty. In this contribution, a methodology to identify an ensemble of behavioural model representations (combinations of input data, model structure and parameter values) is presented and evaluated. The outcome is a multivariate conditional distribution of input data that is used for generating samples of likely inputs (such as Monte Carlo input samples) to perform WWTP model uncertainty analysis. This article presents an approach to verify uncertainty distributions of input data (otherwise often assumed) by using historical observations and actual plant data.

**Key words** | BSM, calibration, influent data, Monte Carlo simulation, modelling

**E. Lindblom**  (corresponding author)  
**U. Jeppsson**   
 Division of Industrial Electrical Engineering and  
 Automation (IEA),  
 Lund University,  
 Lund,  
 Sweden  
 E-mail: erik.u.lindblom@ivl.se

**E. Lindblom**  
 IVL Swedish Environmental Research Institute,  
 Stockholm,  
 Sweden

**G. Sin**   
 Process and Systems Engineering Center  
 (PROSYS), Department of Chemical and  
 Biochemical Engineering,  
 Technical University of Denmark,  
 2800 Kgs. Lyngby,  
 Denmark

## INTRODUCTION

In this article, a set of dynamic wastewater treatment plant (WWTP) model input data is first identified and then used to run simulations, producing confidence intervals giving an estimate of the prediction uncertainties of further model applications. With prior knowledge of the input data and model parameter values to a dynamic WWTP model of the Henriksdal WWTP in Stockholm, Sweden, behavioural input data sets (BIDS) are identified and form model representations (i.e. combinations of input data, model structure and parameter values) that are: (1) consistent with measured data and (2) useful for scenario analyses by simulation. The concept of rejecting non-behavioural model representations has been adopted from the Generalized Likelihood Uncertainty Estimation (GLUE) methodology introduced by [Beven & Binley \(1992\)](#).

Dynamic WWTP models are today tools that are commonly used by both researchers and engineers. For most practical applications, an early goal in the modelling

procedure is calibration; to make simulation results match a set of measurements of target variables. A typical WWTP model contains several potential target variables, which should be considered simultaneously during model calibration ([Hauduc \*et al.\* 2015](#)). There are guidelines for selecting what variables to consider as well as acceptable uncertainty ranges for these, which among other things depend on the type of application ([Rieger \*et al.\* 2012](#)). *Ranges* are provided since a perfect match is not achievable; a set of model outputs  $\mathbf{y}$  will never exactly equal the corresponding measurements,  $\tilde{\mathbf{y}}$ , because of uncertainties in model structure,  $m$ , in applied parameter values,  $\theta$ , in model input data,  $\mathbf{I}$ , as well as experimental errors. An overview of the many uncertainties related to WWTP modelling is found in [Belia \*et al.\* \(2009\)](#).

Therefore, it is necessary to accept a model representation that is a sufficiently good approximation of reality, but it should then be recognized that there are several other possible representations leading (at least almost) to equally good results ([Beven & Freer 2001](#)). This is especially true if the assumptions of the calibration procedure are modified, e.g. if calibration data is added/removed or if an alternative objective function is used. In other words, the uncertainties are lumped into the calibration results and

This is an Open Access article distributed under the terms of the Creative Commons Attribution Licence (CC BY 4.0), which permits copying, adaptation and redistribution, provided the original work is properly cited (<http://creativecommons.org/licenses/by/4.0/>).

doi: 10.2166/wst.2019.427

then transformed into prediction uncertainty during model scenario analyses, i.e. simulations of a certain change/deviation from the normal operating point of the plant. A non-trivial question is then to what extent one can trust and apply the results of the scenario analyses.

The objective of this paper is to define and validate a method for quantifying the WWTP prediction uncertainty using a probabilistic/Monte Carlo sampling framework. The method, which is inspired by GLUE (Beven & Binley 1992), is based on grading and selecting behaviourally good model representations from a larger set, which is obtained from Monte Carlo simulations with input data sampled from a prior uncertainty distribution. In the literature review of Belia *et al.* (2013), it is found that the Monte Carlo method where the uncertainty of input parameters is propagated through the model without any feed-back is the most commonly used method of uncertainty analysis when evaluating WWTP designs. Examples of applications with this method are found in Benedetti *et al.* (2013) and Sin *et al.* (2009). Monte Carlo methods like the one presented in this study, where the prior uncertainty is updated by evaluating and comparing the simulations with measured data from the process, are, however, not as commonly applied, especially not for uncertainty analysis of full-scale WWTP models. Previously Sin *et al.* (2008) used a Monte Carlo-based calibration approach focusing on a limited number of kinetic activated sludge model (ASM) parameters and on aeration input data, and Mannina *et al.* (2012) used GLUE and varied 29 kinetic and stoichiometric ASM2d parameters to calibrate and assess the uncertainty of a full-scale WWTP.

Although it is possible to include an infinite number and various types of parameters in the uncertainty analysis the authors have, in this paper, chosen to focus on a limited amount of input data only. The choice was based on two criteria. (1) Identifiability analysis of WWTP models has indicated that only a fraction of WWTP kinetic and stoichiometric parameters are identifiable given typical data/information collected from full-scale plant operations (Brun *et al.* 2002; Vanrolleghem *et al.* 2003; Sin *et al.* 2008). (2) Previous studies on sensitivity analysis (Sin *et al.* 2011), as well as several calibration protocols (Hulsbeek *et al.* 2002; Melcer *et al.* 2003; Vanrolleghem *et al.* 2003), agree that the most significant parameters affecting uncertainty in WWTP plant models are the influent fractions and calibration of the solids balance.

Therefore, in the below case study application, the authors have selected the influent load, influent fractionation, influent temperature and parameters affecting the solids balance (and sludge retention time, SRT) as main input

sources of uncertainty. These variables are measured, which facilitates the quantification of the prior uncertainty ranges. The influent temperature was moreover chosen as an uncertain calibration parameter, because it explicitly affects the kinetic parameters (growth rate, decay rate, etc.) of the underlying microbiological processes (ASM models). The outcome from the method is a set of behaviourally good input data, BIDS, which are conditioned on a historical (calibration) data set that can be used in model-based applications, such as scenario analysis of plant operation.

## MATERIAL AND METHODS

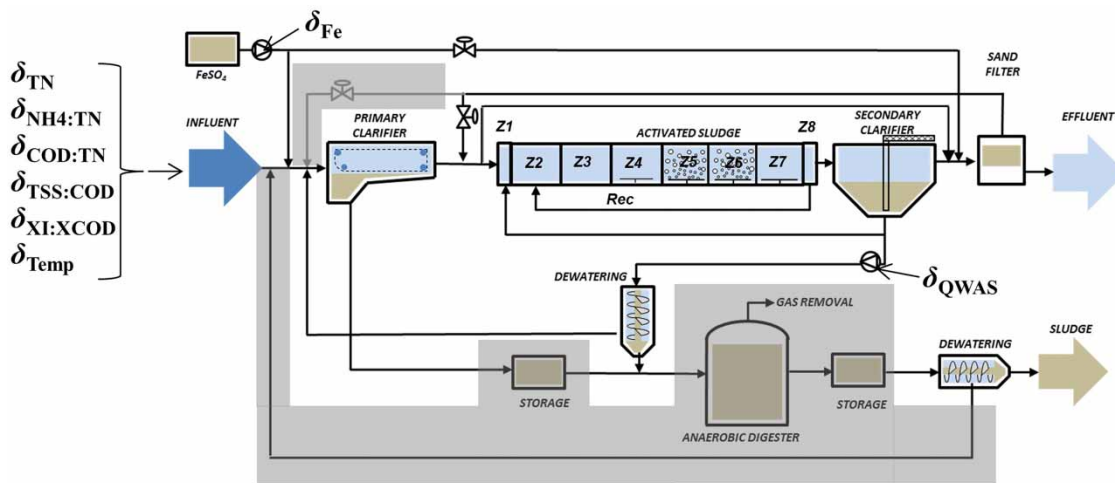
### WWTP and problem statement

Henriksdal WWTP (approximately 800,000 population equivalent) in Stockholm, Sweden, includes primary treatment, an activated sludge system, tertiary treatment in sand filters and anaerobic digestion for sludge stabilization and biogas production, see Figure 1. The plant is currently being operated at its maximum allowed sludge concentration in the bioreactors (mixed liquor suspended solids, MLSS), approximately  $2,500 \text{ g m}^{-3}$ .

With one out of seven biological treatment lines (14%) being out of operation due to reconstruction works, the nitrate recirculation ( $Q_{\text{Rec}}$ ) pump capacity and the number of available aerators at the studied WWTP have been reduced by 14%. Unavoidably, however, the influent wastewater flow rate and load to the plant increase all the time because of the population increase in Stockholm and densification of the wastewater catchment area. Still, the effluent requirement,  $10 \text{ gTN m}^{-3}$  as a yearly average for nitrogen, remains. To evaluate if the nitrogen removal process can be enhanced, either by speeding up the remaining recirculation pumps or by intensifying the aeration, the two scenarios were simulated and evaluated using a dynamic model of the WWTP as will be described in the following text.

### WWTP model under study

The Benchmark Simulation Model No. 2 (BSM2) (Gernaey *et al.* 2014) was a close to reality basis for modelling the plant. Computational speed, however, is an issue for Monte Carlo methods, and this was the main reason for excluding the anaerobic digester (AD) process model of BSM2 from this study. At the WWTP, and in the model, reject water from the AD is mixed with raw wastewater before the influent sampling point. A few minor additional



**Figure 1** | The layout of Henriksdal WWTP as implemented in the BSM2 framework. Components covered by shaded areas were not modelled. The  $\delta$  parameters show the projection of uncertainty.

modifications to the BSM2 layout were needed to adjust the model to the Henriksdal WWTP. (1) The total bioreactor volume is distributed into two smaller (Z1, Z8) and six equally sized zones (Z2–7), which is motivated by historical tracer tests and the placement of mixers (Z2–4) and controllable aeration (Z4–7) zones. (2) Dissolved ferrous sulphate ( $\text{FeSO}_4$ ) is added to the primary clarifier and to the sand filter inlets and is modelled by assuming an instant production of precipitates (inert suspended solids). (3) Inert suspended solids generation from the influent and iron dosage is modelled using the two particulate dummy state variables of BSM2. (4) Rapid two-media sand filter units are used to remove P and total suspended solids (TSS) from the secondary clarifier effluent and is modelled as an ideal thickener. In the model, the filter flush water is recirculated to the bioreactor, while in reality there is a possibility to recirculate that water to the inlet of the primary clarifier. However, the settings of the valves regulating the destiny of the flush water are not recorded at Henriksdal WWTP, thus adding to the model structure uncertainty.

### Adjustment of site-specific conceptual model parameters

The specific inorganic suspended solids (ISS) production was assigned a value of  $2.5 \text{ gISS gFe}^{-1}$  (ATV 2000). Then, by analysing 5 years of delivery data of the  $\text{FeSO}_4$  product (measured weight, iron content) and measured data on sludge leaving the plant with trucks (weight, iron content, ash content), it could be concluded that, to close the mass balance, the influent wastewater then contains approximately  $30 \text{ gISS m}^{-3}$ . This is conditional on the modelling assumption that ISS is not

formed by other processes and motivates the use of  $f_{\text{VSS:TSS}} = 0.9$  as the fixed parameter value for fractionating influent TSS to volatile suspended solids (VSS).

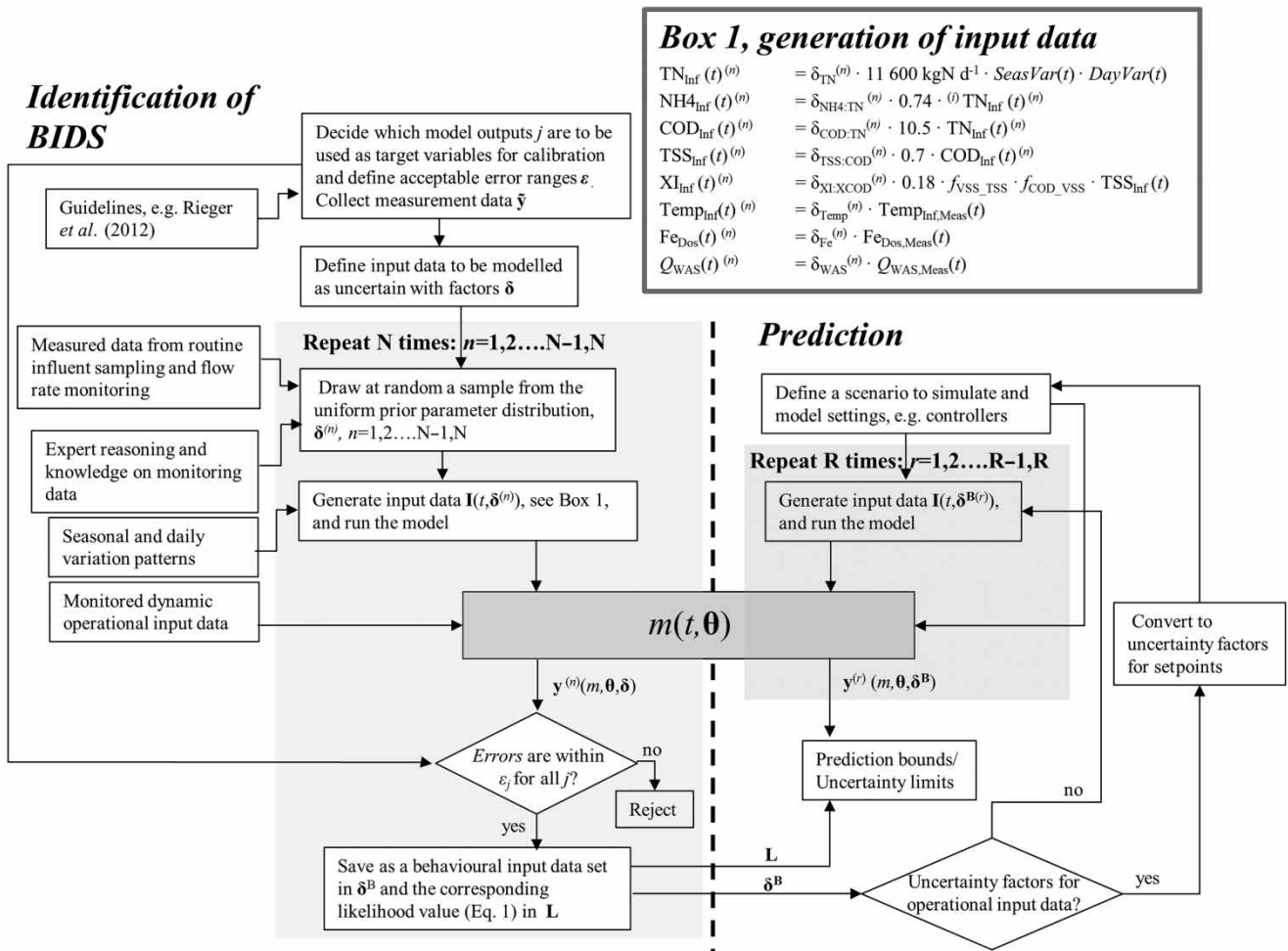
The removal efficiency of the conceptual primary clarifier model was modified to reduce approximately 60% of the particulate material by increasing the efficiency correction factor from 0.65 to 0.73. The removal of  $\text{Fe}^{2+}$  precipitates in the primary settler was reduced by decreasing the specific settleability factor from 1 to 0.75. This was motivated by measurements of total iron in the thickened waste activated sludge (WAS) flow, indicating the fraction of total added iron that ends up in the biological treatment due to the simultaneous precipitation.

The removal efficiency of the simplified sand filter model was set to 90% and together with the secondary settler model (using default BSM2 parameter values) the TSS in the effluent was on average  $4 \text{ mgTSS l}^{-1}$ , which is similar to measured values. Finally, the removal efficiency of the WAS thickener was set to 90% based on grab samples of TSS in the reject water and experience of the operational staff.

### Measurement data and acceptable uncertainty ranges

The first step in the process of identifying BIDS is to decide which model outputs  $j$  to consider and the acceptable uncertainty ranges  $\varepsilon = \{\varepsilon_j\}$ , see Figure 2. These were adopted from the ‘good modelling practice’ (GMP) protocol (Rieger et al. 2012) and the modelling application type ‘assessing the plant capacity for nitrogen removal’.

Measured process data  $\tilde{y}$  selected from a period of 200 days (May–November 2016) was selected. Weekly grab samples of MLSS ( $\tilde{V}_{\text{MLSS}}$ ) and the measured WAS



**Figure 2** | Schematic of the uncertainty analysis methodology. It is presumed that the problem statement has already been defined as this determines the data requirements and error ranges.

production ( $\tilde{\mathbf{y}}_{\text{WAS}}$ ) from weekly suspended solids (SS) grab samples and online flow measurements of the WAS were used to update the prior uncertainty with information related to the plant sludge production and SRT. Online measurements of  $\text{NH}_4$  ( $\tilde{\mathbf{y}}_{\text{NH}}$ ) and  $\text{NO}_3$  ( $\tilde{\mathbf{y}}_{\text{NO}}$ ) concentrations from the secondary clarifier effluent contributed information on the plant nitrogen removal capacity. ( $\tilde{\mathbf{y}}_{\text{NH}}$ ) and ( $\tilde{\mathbf{y}}_{\text{NO}}$ ) data were validated in the sense that the sets used for calibration originated from periods when duplicate sensor readings, providing similar results, were available.

The measured mean values  $S(\tilde{\mathbf{y}}_j)$  of the target variables are shown later in Table 1 as are the applied acceptable uncertainty ranges  $\epsilon$ . The acceptable uncertainty range for both  $\text{NO}_3\text{-N}$  and  $\text{NH}_4\text{-N}$  was  $0.9 \text{ mg l}^{-1}$ . As a criterion, the authors thus decided to accept all input data sets leading to almost complete nitrification, here defined as an average  $\text{NH}_4\text{-N}$  concentration in the aeration tank  $< 1.8 \text{ mgN l}^{-1}$ . For the WAS mass loading rate, a higher range ( $\pm 10\%$ )

than recommended ( $\pm 5\%$ ) was used. It is believed that 5% is a too narrow range since, in the applied model, uncertainties related to the TSS wasted with the filter flush water, as well as WAS being recirculated to the plant with reject water, are simulated as uncertainty in  $\mathbf{y}_{\text{WAS}}$ .

### Prior input uncertainty

During the identification phase with access to measured data the BSM framework with parameter values  $\theta$  was considered well established and kept fixed. Besides site-specific conceptual parameters, all ASM1 and settling parameters of the model were kept constant at their default values. The total uncertainty was projected on part of the model input data by the inclusion of eight uncertainty factors in the vector  $\delta = \{\delta_m\}$ , assumed to be independent and uniformly distributed ( $U$ ) with min/max ranges and referred to as the multivariate prior uncertainty distribution  $P$ .

**Table 1** | Properties of the applied prior uncertainty distribution and input data and the applied likelihood measure

| Model input data $\mathbf{I}$ with prior uncertainty $\delta$ |                                 |                    |                        | Measurement data $\mathbf{\bar{y}}$ |                    |                     |                    |
|---|---------------------------------|--------------------|------------------------|-------------------------------------|--------------------|---------------------|--------------------|
| $m$   | Unit                            | Mean $S(I_m)$      | [ $\pm\%$ ] $\delta_m$ | $J$                                 | Unit               | Mean $S(\bar{y}_j)$ | Range $\epsilon_j$ |
| TN  | kg d <sup>-1</sup>              | 11,600             | 20                     | MLSS                                | g m <sup>-3</sup>  | 2,300               | 230                |
| NH <sub>4</sub> :TN   | gN (gN) <sup>-1</sup>           | 0.74               | 5                      | WAS                                 | kg d <sup>-1</sup> | 22,900              | 2,290              |
| COD:TN  | gCOD (gN) <sup>-1</sup>         | 10.5               | 20                     | NH <sub>4</sub>                     | gN m <sup>-3</sup> | 0.9                 | 0.9                |
| TSS:COD   | gSS (gCOD) <sup>-1</sup>        | 0.7                | 20                     | NO <sub>3</sub>                     | gN m <sup>-3</sup> | 6.2                 | 0.9                |
| X <sub>I</sub> :X <sub>COD</sub>                              | gCOD (gCOD) <sup>-1</sup>       | 0.18               | 20                     |                                     |                    |                     |                    |
| Temp( $t$ )   | °C                              | 21 <sup>a</sup>    | 10                     |                                     |                    |                     |                    |
| Fe <sup>2+</sup> ( $t$ )                                      | gFe m <sup>-3</sup>             | 18 <sup>a</sup>    | 20                     |                                     |                    |                     |                    |
| Q <sub>WAS</sub> ( $t$ )                                      | m <sup>-3</sup> d <sup>-1</sup> | 3,260 <sup>a</sup> | 20                     |                                     |                    |                     |                    |

<sup>a</sup>Average value for the calibration data period.

The selection of input data to consider uncertain was concentrated on the influent wastewater loads and fractions, which are fed to the used influent wastewater generator model (Germaey et al. 2011) including uncertainty in particulate inert chemical oxygen demand (COD) ( $X_I$ ) content (determining the degradability of influent organic matter) and the wastewater temperature, see Box 1 in Figure 2. The influent temperature was chosen as an uncertain calibration parameter, because it explicitly affects the kinetic parameters (growth rate, decay rate, etc.) of the underlying microbiological processes (ASM models). By considering also the operational model inputs  $Q_{WAS}$  and iron dosage  $Fe_{DOS}$  as uncertain, the uncertainties in the settling tank model and the overall sludge balance were modelled, which affect key variables, such as actual sludge wastage from the plant and SRT (Ekama 2010).

Online input data sets considered to be known (hence not uncertain) and directly used as inputs to the model include the influent, return activated sludge (RAS) and nitrate recirculation (REC) flow rates and three sensor data readings of dissolved oxygen concentrations in bioreactors Z5,6,7. This choice was deliberately made to keep the model calibration complexity manageable and focus the discussion on analysis of the presented methodology.

Ranges (min/max) of  $P(\delta)$  were derived from online monitoring and influent sampling results as well as from expert reasoning (Sin et al. 2009), see Table 1.  $i = 1, \dots, N$  latin hypercube sampling of  $\delta$  from  $P$  were used to generate realizations of model Monte Carlo input data  $\mathbf{I}$ , see Figure 2. For example, the yearly average total nitrogen (TN) load was estimated to be 11,600 kgN d<sup>-1</sup> and to have an uncertainty of  $\pm 20\%$  (uniform distribution), thus  $\delta_{TN} \sim U(0.8, 1.2)$ , and the  $i$ th modelled influent TN load profile was generated by

multiplying the  $i$ th sample of  $\delta_{TN}$  with 11,600 and the estimated yearly and daily load patterns.

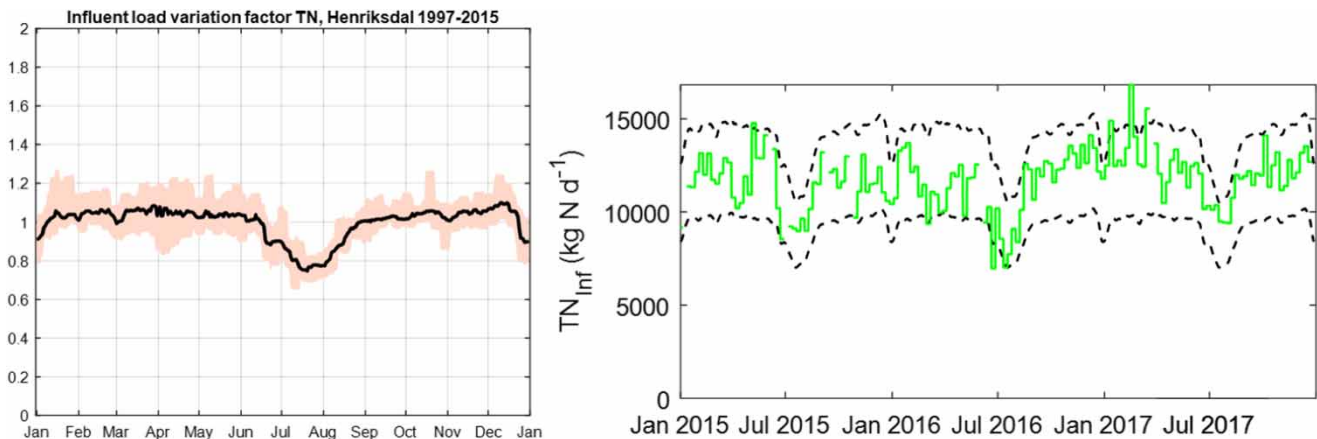
The dry weather daily load pattern ( $DayVar(t)$ ) was obtained by processing online measurements of influent flow rate and NH<sub>4</sub>-N concentrations (results not shown). The yearly pattern was derived as follows (see also Figure 3, left). Eighteen years of weekly flow proportional TN influent load data to the WWTP was arranged and scaled with the yearly average to produce 18 yearly load variation profiles, each having a mean value of one. The black line shows the average variation as a function of time, which was filtered and used as input data ( $SeasVar(t)$ ) to the simulations. The red region is delimited by the 5% and 95% percentiles of the 18 yearly patterns. In Figure 3 (right), the result of the anticipated min/max prior uncertainties for the TN load (dashed lines) are shown together with measurements from three years (green line).

The  $i$ th NH<sub>4</sub>, COD, TSS and  $X_I$  load profiles were then generated by multiplying the  $i$ th TN load with the  $i$ th samples of corresponding uncertainty parameters for wastewater composition:  $\delta_{NH_4:TN}$ ,  $\delta_{COD:TN}$ ,  $\delta_{TSS:COD}$  and  $\delta_{X_I:XCOD}$ . Thus, all pollutants were assumed to have the same diurnal and seasonal variation. Three uncertain input data sets were derived from measured time series data (influent temperature,  $Temp_{Inf}(t)$ ,  $Q_{WAS}(t)$  and ferrous iron dosage  $Fe_{DOS}(t)$ ) by directly multiplying these with their corresponding uncertainty factor.

### Selection and grading of BIDS

Having run  $N$  Monte Carlo simulations and after comparing the model results with measurements, realizations that differ





**Figure 3** | Left: yearly TN load variation curve,  $SeasVar(t)$ , used as input data. Right: measured TN load (green line) and min/max of prior uncertainty (dashed lines):  $TN_{inf}(t) = \delta_{TN} \cdot 11,600 \text{ kg N d}^{-1} \cdot SeasVar(t)$ . Please refer to the online version of this paper to see this figure in colour: <https://dx.doi.org/10.2166/wst.2019.427>.

more than the pre-defined uncertainty ranges,  $\varepsilon_j$ , for any target variable,  $j$ , are rejected. The subsets of  $\delta$  yielding results within the ranges are identified as behavioural (BIDS) and denoted  $\delta^B$ . An essential part of the presented methodology is the rejection (not only assigning a low likelihood) of non-behavioural models. This is because density-dependent sampling methods, such as Markov chain Monte Carlo, are not proposed here, and for the method to be applicable in practice, the number of ensemble simulations needed to be run for the prediction phase must be limited.

An input data set,  $i$ , was rejected if the resulting mean error (ME) of *any of the targets*  $j$  were outside the selected uncertainty ranges, e.g. if  $ME_j = |S(\mathbf{y}_j^{(i)}) - S(\tilde{\mathbf{y}}_j)| > \varepsilon_j$ , where  $S(\cdot)$  denotes the mean function. Then an accepted behavioural input data set must provide reasonably good results considering *all* target variables and thus, a bad prediction of, for example, MLSS *cannot* be compensated for by good WAS,  $NH_4$  and  $NO_3$  predictions. The choice of using the mean value as a statistic to distinguish between behavioural and non-behavioural model representations was based on the time averaging (month/year) proposed for the application type in the GMP protocol (Rieger *et al.* 2012). Also, according to the experience of the authors, a common and pragmatic first step during manual WWTP calibration is to compare results from steady state simulations with averaged results; the applied rejection methodology here is a way to automate this procedure.

Following the GLUE methodology each behavioural model representation is conditioned on past data and given a likelihood weight,  $L$ , according to the modeller's confidence in that particular model representation as a useful predictor of the future. The likelihoods are then normalized to unity and used to compute the

likelihood-weighted cumulative density function (cdf) for any model prediction. From the cdf the percentiles of the simulations with  $\delta^B$  are derived. These uncertainty limits reflect what the model can say about the system response after conditioning on the past calibration data.

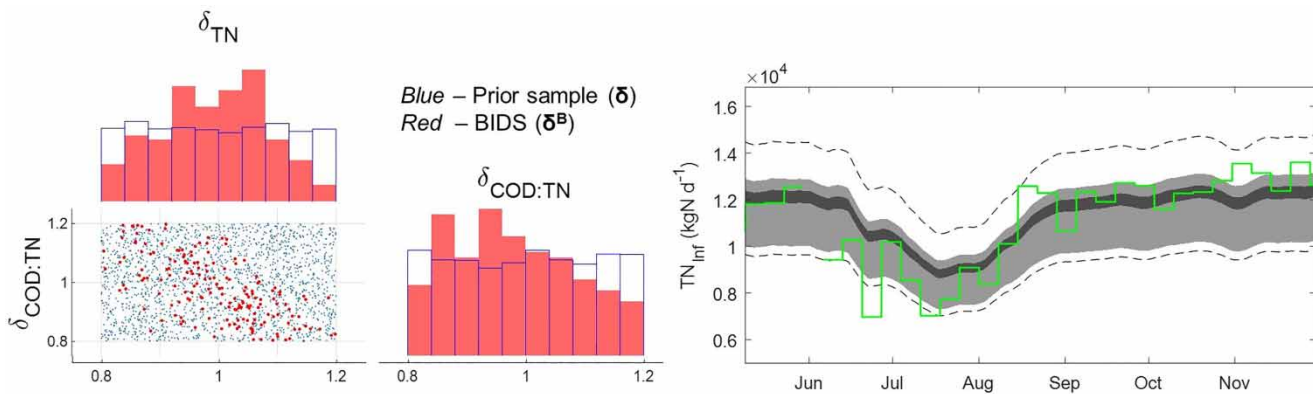
There are numerous ways to measure the efficiency of dynamic models (Hauduc *et al.* 2015) especially, as in the current study, when there are multiple target variables. The subjective choice of the likelihood measure in this paper was based on a product of four (one for each target variable  $j$ ) factors with exponential functions of the mean squared errors (MSE). For the  $r$ th behavioural input data set  $\delta^{B(r)}$ ,  $r = 1, \dots, R$ , it reads

$$L^{(r)} = \prod_j \exp\left(-MSE_j(\delta^{B(r)})/\varepsilon_j^2\right)$$

The uncertainty range,  $\varepsilon_j$ , is used to model the rate at which the likelihood decreases for an increase in MSEs (Lindblom *et al.* 2011). To exemplify: if a behavioural input data set  $\delta^{B(r)}$  yields MSEs of exactly the uncertainty ranges, the  $r$ th likelihood value becomes  $\exp(-1.4) = 0.018$ . The selection of likelihood function and statistic (resolution of data) is not obvious and an important topic for further research.

## RESULTS AND DISCUSSION

Figure 4 (left) shows histograms and a two-dimensional projection for two out of the eight uncertain input parameters ( $\delta_{TN}$  and  $\delta_{COD:TN}$ ).  $N = 2,000$  input data sets from the uniform prior uncertainty distribution are shown using blue



**Figure 4** | Left: histograms and a two-dimensional projection for two out of eight uncertain input parameters ( $\delta_{TN}$  and  $\delta_{COD:TN}$ ). Right: influent TN load. Grey regions – 95/5% and 75/25% likelihood-weighted uncertainty limits. Dashed black lines – min and max realizations of the prior uncertainty. Green line – weekly flow proportional measurements. Please refer to the online version of this paper to see this figure in colour: <https://dx.doi.org/10.2166/wst.2019.427>.

markers. With the assumptions made, 9% ( $R = 179$ ) of those were identified as BIDS forming, in combination with the model structure and default parameter values, plausible validated model representations, which were used further for scenario analyses. The BIDS are shown using red markers and for these two parameters there is a correlation. For high TN loads ( $\delta_{TN} > 1.1$ ) there are, for example, no BIDS with a high influent COD:TN ratio ( $\delta_{COD:TN} > 1.1$ ). This is because, at least partly, the influent COD load would then be too high and rejected by the WAS criterion (leading to extensive sludge production). A future sensitivity analysis (using Sobol indices or partial correlation coefficients) will be used to complement the selection and analysis of uncertain parameters.

In Figure 4 (right) the likelihood-weighted influent TN load is shown together with the initial prior uncertainty ranges (dashed lines) and measured data. For TN, as well as the other model inputs loads ( $NH_4$ , COD, TSS, 7-day biochemical oxygen demand), the anticipated prior uncertainty distribution results in simulated uncertainty ranges that cover the measurements. Moreover, after updating the distribution with process measurements the simulated ranges narrow.

Figure 5 shows in turquoise part of the process data,  $\check{y}$ , used to identify and grade the BIDS as well as the resulting uncertainty ranges. The WAS load was operated manually during this period and varies a lot because of limitations in the downstream sludge system. This results in varying MLSS concentrations which are well predicted by the simulations with BIDS. Considering the nitrogen target variables, the simulated uncertainty range of  $NH_4$  does not deviate significantly from the prior uncertainty. This is due to the fact that temperature was high during the calibration period and

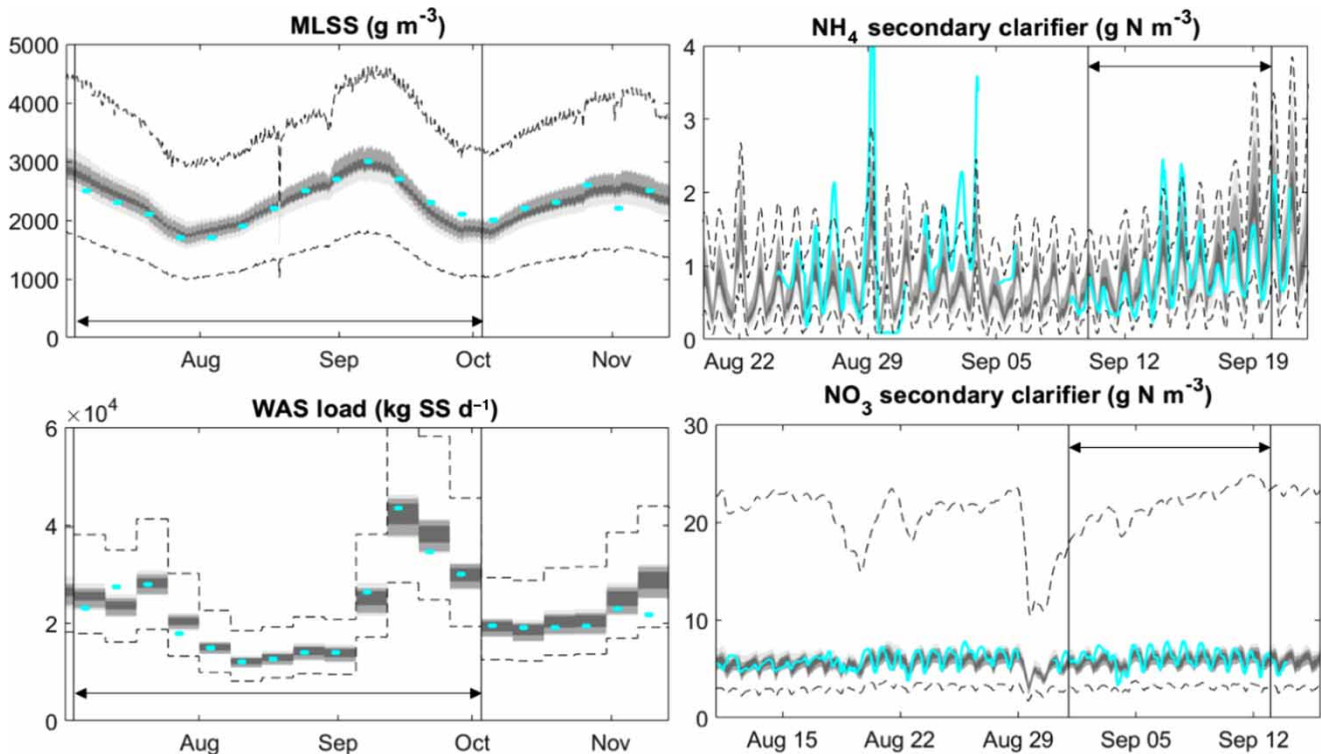
that the dissolved oxygen (DO) concentration inputs were considered fixed. Then nitrification was almost completely independent of the prior uncertainty. For  $NO_3$  on the other hand, the uncertainty is reduced significantly after identification and simulation with the BIDS.

### Scenario simulations

To test and validate the uncertainty analysis method, the identified 179 BIDS were used to simulate the two scenarios as outlined in the ‘WWTP and problem statement’ section. The initialization settings were dry weather conditions, an influent temperature of 14 °C and a constant DO setpoint in the last aerobic zone (Z7) of 0.5  $gO_2 m^{-3}$ . The DO setpoints in Z5–6 were adjusted (maximum 4  $gO_2 m^{-3}$ ) by an effluent ammonia-based aeration controller (ABAC) with setpoint 1.6  $gN m^{-3}$ .  $Q_{Rec}$  was kept constant and  $Q_{RAS}$  controlled to 50% of  $Q_{Inf}$  (Figure 6(a)). A constraint for the scenarios was that the MLSS concentrations of the bio-reactors were not allowed to increase above the average MLSS of the calibration period.

The loading and temperature input data series to the scenarios were obtained by replacing  $\delta$  with the behavioural uncertainty factors  $\delta^B$  in the equations of Box 1 in Figure 2 (running only the BIDS).

The uncertainty factors for operational input data needed to be translated to uncertainty factors for the setpoints of the controllers. Dosage of  $Fe^{2+}$  is modelled by a feed-forward controller proportional to the influent flow rate. For each of the 179 BIDS, the setpoint value ( $gFe m^{-3}$ ) is set as the average of the corresponding  $Fe^{2+}$  consumption during calibration. Thus, the BIDS that had a high/low chemical sludge production during calibration will have a



**Figure 5** | Turquoise dots and lines – part of the calibration data (indicated by arrows) used to identify and calculate the likelihood of the BIDS. Grey regions 95/5% and 75/25% likelihood-weighted uncertainty limits as simulated with the BIDS. Dashed black lines – min and max realizations of the prior uncertainty, which are used to contrast the output range of nitrate that corresponds to BIDS (much narrower). Please refer to the online version of this paper to see this figure in colour: <https://dx.doi.org/10.2166/wst.2019.427>.

similarly high/low production during the prediction. Regarding the uncertainty in  $Q_{WAS}$  it was noted that each set of the 179 BIDS defining the loads, temperature,  $Fe^{2+}$  dosage and  $Q_{WAS}$  was associated with a simulated MLSS concentration profile. Since the average MLSS of the calibration period was assumed to be the maximum allowed, the simulated average MLSS of each BIDS was used as the setpoint value for  $Q_{WAS}$  in the scenarios. Thus, for example, the food to microorganisms ratios of the BIDS were similar in the base scenarios.

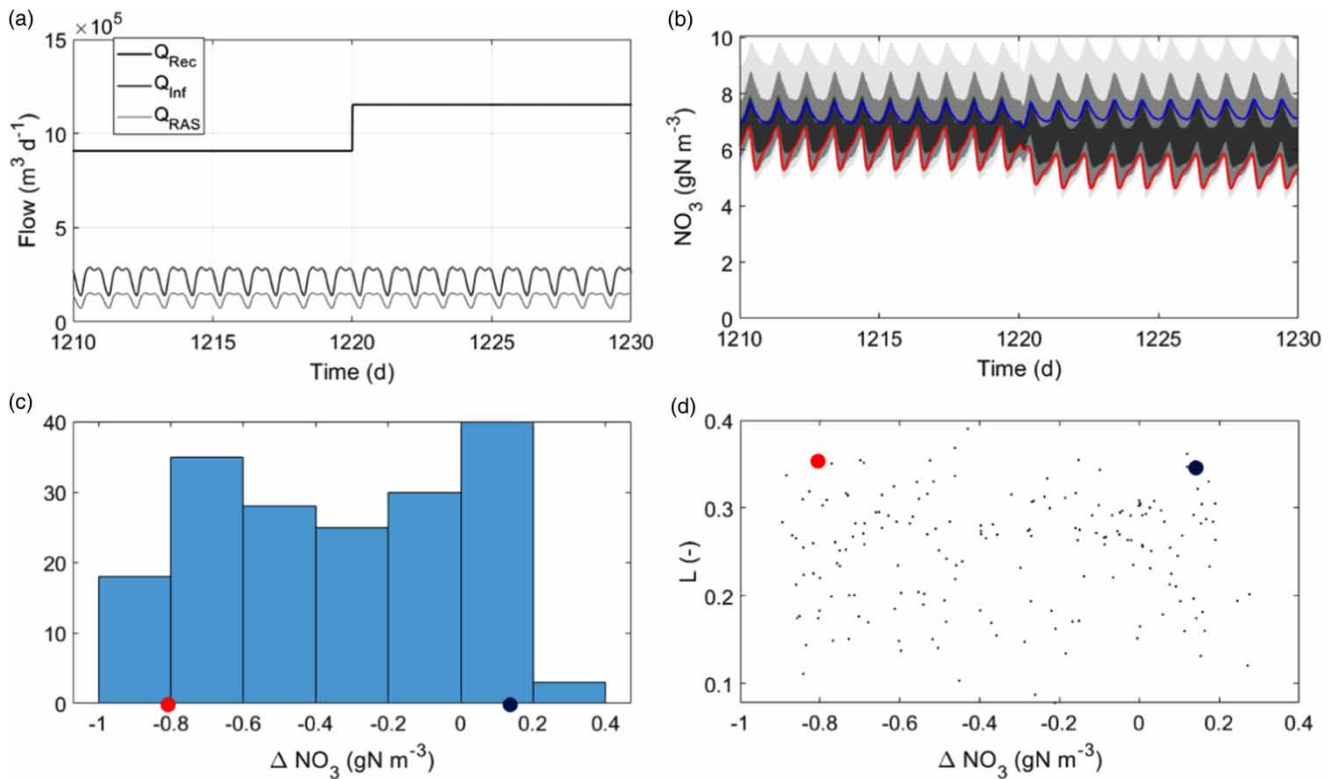
For each scenario, the results of two individual simulations, yielding different answers to the stated problem, are shown in Figure 6 using blue and red colouring. By doing so the importance of uncertainty analysis while drawing conclusions from dynamic WWTP model simulations is highlighted. According to the herein presented methodology, although most steps of the ‘good modelling practice’ are followed, by using and simulating only *one* calibrated model representation (combination of input data, model structure and parameter values), the resulting conclusion drawn could be either the ‘red’ one *or* the ‘blue’ one (or a third one). Recall that both simulations are based on realistic interpretations of the available input data (defined by the

prior distribution) and are judged to be sufficiently consistent with the measured process data (defined by the likelihood function).

### Scenario 1: Tuning of the engines of the remaining recirculation pumps, by increasing $Q_{Rec}$ 26% so that $Q_{Rec}/Q_{Inf}$ increases from 3.5 to 4.2 (Figure 6(a))

Figure 6(b)–6(d) show the simulated uncertainty limits of effluent  $NO_3$  with the realizations of two individual BIDS (coloured lines and dots) included for illustrative purpose. The differences in average effluent  $NO_3$  concentrations after and before the increase,  $\Delta NO_3$ , were studied. As shown in the histogram (Figure 6(c)), although most results indicate enhanced denitrification, there are also BIDS yielding deteriorated performance ( $\Delta NO_3 > 0$ ). Moreover, as shown in Figure 6(d), BIDS with similar likelihood values might lead to a decrease *or* an increase of the effluent  $NO_3$  concentration. Among the simulations yielding deteriorated denitrification, the  $NO_3$  in the anoxic zones is already high with the lower pumping rate, leading to the fact that the increased recirculation only means more oxygen being transported to the anoxic zone.



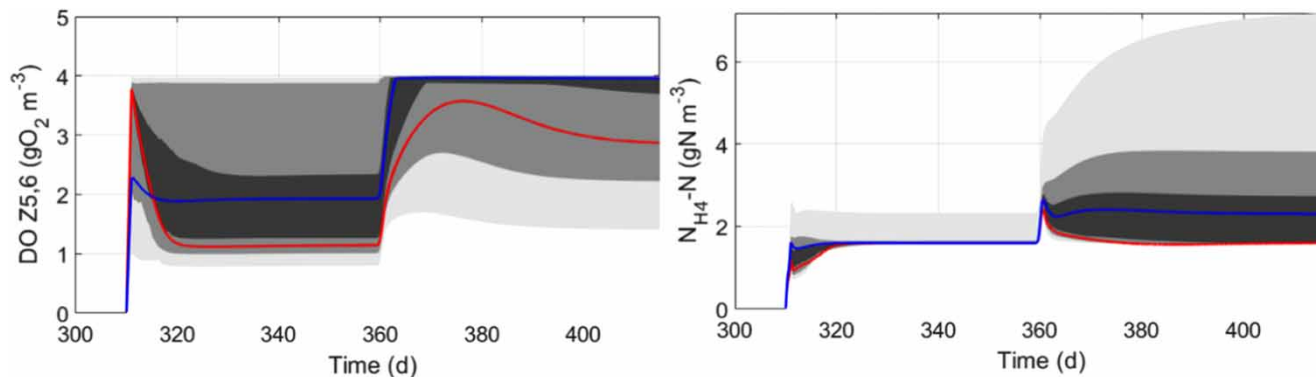


**Figure 6** | Example of a model application study (*Scenario 1*) evaluated with the BIDS. Grey regions – 100/0%, 95/5% and 75/25% likelihood-weighted uncertainty limits as simulated with the BIDS. See text for further details. Please refer to the online version of this paper to see this figure in colour: <https://dx.doi.org/10.2166/wst.2019.427>.

**Scenario 2: Increase in influent pollutant loads by increasing  $\delta_{\text{TN}}$  and  $\delta_{\text{Fe}}$  20% (all others, e.g. the COD and TSS loads, increase as well since generated from the behavioural subset of wastewater composition parameters (Table 1))**

To maintain the MLSS despite the load increase at  $t = 360$  d, the WAS flow rate increases in all simulations. The DO

concentrations in the aerobic zones (Figure 7, left) are increased by the ABAC controller to maintain the  $\text{NH}_4$  in the effluent at  $1.6 \text{ gN m}^{-3}$  although the aerobic SRT is decreased (Figure 7, right). By considering the likelihood-weighted 75/25% and 95/5% prediction limits, the results state that the DO will probably increase to above  $2 \text{ gO}_2 \text{ m}^{-3}$  and possible saturate at  $4 \text{ gO}_2 \text{ m}^{-3}$  while the plant can continue to nitrify (effluent ammonia  $< 4 \text{ gNH}_4\text{-N m}^{-3}$ ).



**Figure 7** | Example of a model application study (*Scenario 2*) evaluated with the BIDS. Grey regions – 100/0%, 95/5% and 75/25% likelihood-weighted uncertainty limits as simulated with the BIDS. See text for further details. Please refer to the online version of this paper to see this figure in colour: <https://dx.doi.org/10.2166/wst.2019.427>.

However, note that among the simulations with BIDS there are those indicating that nitrification cannot be expected to be complete following the load increase, see the 100/0% percentile. This highlights the fact that the result of a simulation depends on the assumptions made during calibration. The blue and red lines, for example, represent simulations with two of the 179 BIDS that slightly over- and underpredicted the measured MLSS concentration during calibration. The blue simulation is a result of a high anticipated value of the true WAS flow rate and thus a low SRT during calibration. To produce results within the acceptable uncertainty ranges for  $\text{NH}_4$  and  $\text{NO}_3$  (required as having been identified as BIDS) this set of input data is conditional on a high value of the influent temperature. The opposite holds for the red simulation (low  $Q_{\text{WAS}}$ , high SRT, low temperature). Obviously, the assumptions done for the 'blue' simulation during calibration involve a higher sensitivity to the increased load since the  $\text{NH}_4$  setpoint cannot be maintained.

An added value with applying the presented methodology, compared to more traditional methods based on, for example, expert ranges, is that the results are derived with actual and case-specific data and that the uncertainty ranges therefore will consider and be dependent on the amount and quality of accessible data. Another benefit is that the method is easy to use in practice. Once the BIDS have been identified they can be used for simulating the uncertainty of basically any scenario. Although the total uncertainty might not be the concern of, for example, a process engineer, it is also usually valuable to run at least a few alternative simulations while studying a problem. By sampling, for example, 5–10 input data sets from the BIDS and running the model several times, he/she will get much more insight compared to running one simulation only.

It should finally be noted that the results are dependent on several assumptions (choice of parameters to consider uncertain, the prior distribution, target variables, uncertainty ranges and likelihood function, etc.). The predicted uncertainty for scenarios that differ significantly from the calibration data period might therefore not be well described.

## CONCLUSIONS

Monte Carlo methods are powerful methods that offer a range of applications for WWTP process analysis, optimization and more. An important requirement of these methods is definition of an input uncertainty domain,

which directly affects the outcome and quality of the Monte Carlo analysis. The authors have presented a 'proof of concept' evaluation of a methodology to generate multivariate samples from the input uncertainty domain, which are conditioned on historical plant data and observations. The results show that the applied methodology is promising and useful to verify assumptions related to input data for Monte Carlo analysis.

The authors recognize that many assumptions are made in this study to develop and evaluate the proposed framework, yet they also believe it provides a pragmatic and useful method to improve credibility of Monte Carlo simulations. The latter require random samples of input data, which are often defined by expert knowledge without a systematic way to verify the validity of these 'expert' ranges. Rather than falsification or to inspire further mechanistic model development, the long-term goal of demonstrating the applicability of the methodology is to enhance the application of already available WWTP simulators combined with powerful Monte Carlo methods.

## REFERENCES

- ATV A-131 ATV-DVWK 2000 *A 131E. Dimensioning of Single-Stage Activated Sludge Plants*. German Association for Water, Wastewater and Waste. GFA Publishing Company of ATV-DVW, Hennef, Germany.
- Belia, E., Amerlinck, Y., Benedetti, L., Johnson, B., Sin, G., Vanrolleghem, P. A., Gernaey, K. V., Gillot, S., Neumann, M. B., Rieger, L., Shaw, A. & Villez, K. 2009 *Wastewater treatment modelling: dealing with uncertainties*. *Water Sci. Technol.* **60** (8), 1929–1941.
- Belia, E., Johnson, B., Benedetti, L., Bott, C. B., Martin, C., Murthy, S., Neumann, M. B., Rieger, L., Weijers, S. & Vanrolleghem, P. A. 2013 *Uncertainty evaluations in model based WRRF design for high level nutrient removal*. WERF Stock No. NUTR1R06q.
- Benedetti, L., Belia, E., Cierkens, K., Flameling, T., De Baets, B., Nopens, I. & Weijers, S. 2013 *The incorporation of variability and uncertainty evaluations in WWTP design by means of stochastic dynamic modeling: the case of the Eindhoven WWTP upgrade*. *Water Sci. Technol.* **67** (8), 1841–1850.
- Beven, K. & Binley, A. 1992 *Future of distributed models: model calibration and uncertainty prediction*. *Hydrol. Proc.* **6**, 279–298.
- Beven, K. J. & Freer, J. E. 2001 *Equifinality, data assimilation and uncertainty estimation in mechanistic modelling of complex environmental systems using the GLUE methodology*. *J. Hydrol.* **249**, 11–29.
- Brun, R., Kühni, M., Siegrist, H., Gujer, W. & Reichert, P. 2002 *Practical identifiability of ASM2d parameters – systematic*

- selection and tuning of parameter subsets. *Water Res.* **36** (16), 4113–4127.
- Ekama, G. A. 2010 The role and control of sludge age in biological nutrient removal activated sludge systems. *Water Sci. Technol.* **61** (7), 1645–1652.
- Gernaey, K. V., Flores-Alsina, X., Rosen, C., Benedetti, L. & Jeppsson, U. 2011 Dynamic influent pollutant disturbance scenario generation using a phenomenological modelling approach. *Environ. Model. Softw.* **26** (11), 1255–1267.
- Gernaey, K. V., Jeppsson, U., Vanrolleghem, P. A. & Copp, J. B. 2014 *Benchmarking of Control Strategies for Wastewater Treatment Plants*. IWA Scientific and Technical Report No. 23, IWA Publishing, London, UK.
- Hauduc, H., Neumann, M. B., Muschalla, D., Gamerith, V., Gillot, S. & Vanrolleghem, P. A. 2015 Efficiency criteria for environmental model quality assessment: a review and its application to wastewater treatment. *Environ. Model. Softw.* **68**, 196–204.
- Hulsbeek, J. J. W., Kruit, J., Roeleveld, P. J. & van Loosdrecht, M. C. M. 2002 A practical protocol for dynamic modelling of activated sludge systems. *Water Sci. Technol.* **45** (6), 127–136.
- Lindblom, E., Ahlman, S. & Mikkelsen, P. S. 2011 Uncertainty-based calibration and prediction of a stormwater surface accumulation-washoff model based on coverage of sampled Zn, Cu, Pb and Cd field data. *Water Res.* **45** (13), 3823–3835.
- Mannina, G., Cosenza, A. & Viviani, G. 2012 Uncertainty assessment of a model for biological nitrogen and phosphorus removal: application to a large wastewater treatment plant. *Phys. Chem. Earth* **42–44**, 61–69.
- Melcer, H., Dold, P. L., Jones, R. M., Bye, C. M., Takacs, I., Stensel, H. D., Wilson, A. W., Sun, P. & Bury, S. 2003 *Methods for Wastewater Characterisation in Activated Sludge Modelling*. Water Environment Research Foundation (WERF), Alexandria, VA, USA.
- Rieger, L., Gillot, S., Langergraber, G., Ohtsuki, T., Shaw, A., Takács, I. & Winkler, S. 2012 *Guidelines for Using Activated Sludge Models*. IWA Scientific and Technical Report No. 22, IWA Publishing, London, UK.
- Sin, G., De Pauw, D. J., Weijers, S. & Vanrolleghem, P. A. 2008 An efficient approach to automate the manual trial and error calibration of activated sludge models. *Biotechnol. Bioeng.* **100** (3), 516–528.
- Sin, G., Gernaey, K. V., Neumann, M. B., van Loosdrecht, M. C. M. & Gujer, W. 2009 Uncertainty analysis in WWTP model applications: a critical discussion using an example from design. *Water Res.* **43** (11), 2894–2906.
- Sin, G., Gernaey, K. V., Neumann, M. B., van Loosdrecht, M. C. M. & Gujer, W. 2011 Global sensitivity analysis in wastewater treatment plant model applications: prioritizing sources of uncertainty. *Water Res.* **45** (2), 639–651.
- Vanrolleghem, P. A., Insel, G., Petersen, B., Sin, G., De Pauw, D., Nopens, I., Weijers, S. & Gernaey, K. V. 2003 A comprehensive model calibration procedure for activated sludge models. *Proceedings of the 76th Annual WEF Technical Exhibition and Conference (WEFTEC2003)*, Los Angeles, CA, USA, October 11–15, 2003.

First received 2 October 2019; accepted in revised form 16 December 2019. Available online 26 December 2019