Evaluating the performance of a simple phenomenological model for online forecasting of ammonium concentrations

Vezzaro, Luca; Pedersen, Jonas Wied; Holm Larsen, Laura; Thirsing, Carsten; Bassø Duus, Lene; Breinholt, Anders; Mikkelsen, Peter Steen

Publication date:
2019

Document Version
Publisher's PDF, also known as Version of record

Citation (APA):
Vezzaro, L., Pedersen, J. W., Holm Larsen, L., Thirsing, C., Bassø Duus, L., Breinholt, A., & Mikkelsen, P. S. (2019). *Evaluating the performance of a simple phenomenological model for online forecasting of ammonium concentrations*. Paper presented at 9th International Conference on Sewer Processes & Networks, Aalborg, Denmark.

Evaluating the performance of a simple phenomenological model for online forecasting of ammonium concentrations

Luca Vezzaro[1,2], Jonas Wied Pedersen[2], Laura Holm Larsen[2], Carsten Thirsing[3], Lene Bassø Duus[4], Anders Breinholt[5], and Peter Steen Mikkelsen[2]

[1]Krüger A/S, Veolia Water Technologies, Gladsaxevej 363, 2860 Søborg, Denmark (Email: *lxv@kruger.dk*)
[2]Department of Environmental Engineering (DTU Environment), Technical University of Denmark, Building 115, 2800 Kongens Lyngby, Denmark (Email: *jowi@env.dtu.dk;psmi@env.dtu.dk*)
[3]BIOFOS A/S, Refshalevej 250, 1432 København K, Denmark (Email: *ct@biofos.dk*)
[4]Aarhus Vand A/S, Gunnar Clausens Vej 34, 8260 Viby J, Denmark (Email: *lba@aarhusvand.dk*)
[5]HOFOR. Hovedstadsområdets Forsyningsselskab, Ørestads Boulevard 35, 2300 København S, Denmark (Email: *anbr@hofor.dk*)

Abstract
A simple model for online forecasting of ammonium (NH_4^+) concentrations in sewer systems is proposed. The forecast model utilizes a simple representation of daily NH_4^+ profiles and the dilution approach combined with information from online NH_4^+ and flow sensors. The method utilizes an ensemble approach based on past observations to create model prediction bounds. The forecast model was tested against observations collected at the inlet of two WWTPs over an 11-month period. NH_4^+ data were collected with ion-selective sensors. The model performance evaluation focused on applications in relation to online control strategies. The results of the monitoring campaigns highlighted a high variability in daily NH_4^+ profiles, stressing the importance of an uncertainty-based modelling approach. Model performance was strongly affected by maintenance of the NH_4^+ sensors, which resulted in important variations of the sensor signal. The forecast model succeeded in providing outputs that potentially can be used for integrated control of wastewater systems. This study provides insights on full scale application of online water quality forecasting models in sewer systems. It also highlights several research gaps which - if further investigated - can lead to better forecasts and more effective real-time operations of sewer and WWTP systems.

Keywords
Online forecasts, Sensor maintenance, Uncertainty, Water Quality-based control

INTRODUCTION

Models for forecasting water quality in different parts of the integrated urban drainage-wastewater system (sewers, wastewater treatment plants - WWTP) can provide useful information for improving the operation of integrated urban drainage-wastewater systems (Yuan et al., 2019). These models can be used to quantify discharges from Combined Sewer Overflows (CSO), or in an online context as part of real-time control strategies aiming at optimizing WWTP operations (the so-called "software sensors" – e.g. Stentoft et al. (2017)). Further potential applications include Model Predictive Control approaches, allowing water quality based control of sewer systems (e.g. Vezzaro et al., 2013) or WWTPs (Stentoft et al., 2019) over different forecast horizons.

As pointed out in Langeveld et al. (2017), the increased availability of long-term times series of water quality parameters with high resolution in time (e.g. Schilperoort et al., 2012; Métadier and Bertrand-Krajewski, 2012; Alferes et al., 2013) allows the development of new models utilizing such information. There is a wide experience with the application of data-driven software sensors in WWTPs (Haimi et al., 2013; Newhart et al., 2019), and several studies on simulating WWTP influent quality (Martin and Vanrolleghem, 2014). Many of these studies employ empirical/phenomenological approaches (Langeveld et al., 2017; Gernaey et al., 2011), while Talebizadeh et al. (2016) proposed a stochastic influent generator to provide a more realistic description of the natural variability of WWTP influent. These examples are mostly based on offline simulation studies, using pre-validated data (e.g. Flores-Alsina et al., 2014), a condition that is not available under actual real-time conditions. The majority of research on forecasting of water quality indicators at WWTPs focuses on quantities within the process tanks or at the plant outlet. Few examples deal with predictions of the WWTP influent (e.g. Kusiak et al., 2013; Yu et al., 2018), despite its potential use in feed-forward control. Furthermore, model predictive power is often evaluated in terms of statistical assumptions regarding residuals, while online application requires more robust, *ad-hoc* metrics, focusing on the intended use of the model outputs.

This paper presents a simple phenomenological model specifically developed for online prediction of ammonium (NH_4^+) loads and concentrations along with their uncertainty. The model relies on a flow forecast and continuous online NH_4^+ measurements from an ISE (ion-selective electrode) sensor. In the evaluation of forecast skill in this study, the forecasts of ammonia loads are based on *ex-post flow forecasts*, i.e. measured flow values that are used "as if" they are forecasted values. In an operational setup, the measured flow should obviously be exchanged for real-time forecasts of flow. However, the *ex-post* setup in this paper ensures that the performance evaluation of the ammonium forecasts is independent of errors in the flow domain. The forecasts are tested at the inlet of two Danish WWTPs and the performance of the forecast model is evaluated over an 11-month period. The evaluation also aims at identifying further research gaps and improvements with specific focus on application in online control strategies.

MATERIAL AND METHODS

Water quality monitoring

Flow and ammonia measurements have been collected with a 2-minute frequency at the Viby WWTP (Aarhus, Denmark – since June 2018) and the Damhusaaen WWTP (Copenhagen, Denmark – since April 2018). The Viby catchment consists of 678 ha combined and 748 ha separate systems (Ahm et al., 2013). The majority of the system is gravity driven, while the flow from two minor subcatchments is pumped to the plant. The Damhusaaen WWTP receives wastewater from a 67 km^2 combined system, which is mainly gravity driven (flow from an adjacent catchment can be pumped in case of extreme events).

NH_4^+ is measured with ion-selective sensors. The sensors in Viby and in Damhusaaen are placed at different locations of the plants for logistical reasons (Figure 1): after the primary clarifier and a pumping station in Damhusaaen and after the grit removal in Viby (Table 1). Their sensor maintenance follows different schedules (Table 1). Data from the plants' SCADA systems are transferred to the AQUAVISTA™ cloud-based system (based on the STAR® system described in Nielsen and Onnerth, 1995), where they are automatically quality controlled based on simple rule methods (e.g. running variance, physical ranges, flat lines).

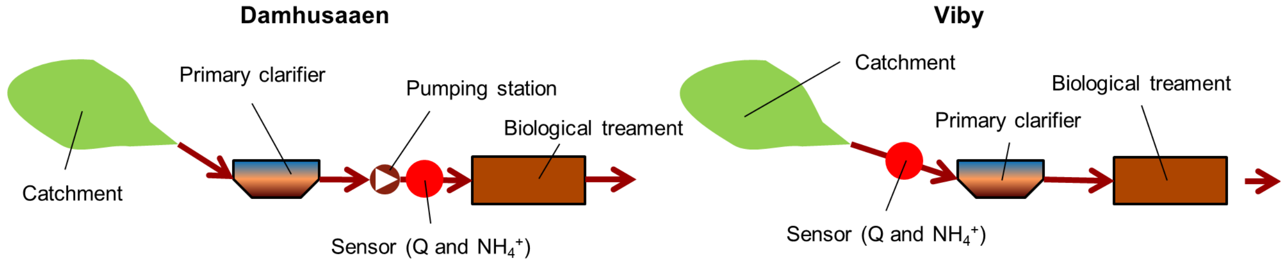

Figure 1. Schematic placement of the sensors in the two plant layouts.

Table 1. Relevant information for the two WWTPs included in the study.

	Damhusaaen	Viby WWTP
Sensor placement (Figure 1)	After primary clarifier and pumping station	At plant inlet (after grit removal)
Sensor cleaning and performance check	Weekly	Weekly
Sensor calibration	If sensor deviation from lab measurements >10%	If sensor deviation from lab measurements >5%
Start of monitoring campaign	April 2018	June 2018
Threshold used to define wet weather events (at WWTP inlet)	5000 m^3/hr	400 m^3/hr
Forecast model parameters	Six ($\alpha_0,\alpha_1,\alpha_2, \beta_1,\beta_{,2},V$)	Eight ($\alpha_0,\alpha_1,\alpha_2,\beta_1,\beta_{,2},\gamma_1,\gamma_2,\gamma_3$)

Forecast model

Ammonium Forecasts. The proposed model builds upon the widely applied concept of daily ammonium loads, i.e. the concept that NH_4^+ loads (i) only originates from domestic sources that can vary between weekdays and weekends, (ii) follows a typical diurnal profile over 24 hours, and (iii) is unaffected by wet weather events (e.g. Langeveld et al., 2017; Martin and Vanrolleghem, 2014). Forecasts of NH_4^+ loads during the *i-th* day are generated based on observations that have been collected in the most recent dry days (Figure 2a). Specifically, daily load profiles from *n* previous days are used to forecast NH_4^+ loads (i.e. the model uses a moving window of length *n* on previous observations). By combining the NH_4^+ loads with flow data it is then possible to estimate NH_4^+ concentrations by using a simple dilution approach. The operations of the forecast model for day *i* are schematized in Figure 2 and further explained in the following paragraphs:

Step 1: Every day, after midnight, flow measurements from day *i-1* are analysed for identifying potential wet weather events.

Step 2: If the previous day was a dry day, the parameters of the NH_4^+ daily profile for day *i-1* are estimated and stored in a database (Figure 2b).

Step 3: Optimal parameter sets from the previous *n* dry days of corresponding day type (weekday/weekend) are retrieved.

Step 4: Ammonium forecasts for day *i* are generated by running the model every 2-minutes. Forecast uncertainty is created as an ensemble with *n* members consisting of ammonia load profiles from the past *n* corresponding days (weekday/weeknd).

Model structure. The online model runs in two-minute time steps and uses a Fourier series to represent the diurnal variation of the NH_4^+ load over a day (as in Bechmann et al., 1999):

$$F_{NH4}(t) = \alpha_0 + \sum_{k=1}^{2}\left(\alpha_k sin(2\pi kt) + \beta_k cos(2\pi kt)\right) \quad (1)$$

where $\alpha_0, \alpha_1, \alpha_2$ and $\beta_1, \beta_{,2}$ are the Fourier coefficients, and the time t is expressed as fraction of a day. Ammonium concentrations S_{NH4} are then calculated by dividing the estimated loads by the flow $Q(t)$ [m^3/hr], as in Langeveld et al. (2017):

$$S_{NH4}(t) = F_{NH4}(t)/Q(t) \quad (2)$$

As mentioned earlier, the flow values $Q(t)$ that are used as input to the forecasting scheme are the values that were observed later in the day, and thus not real forecasts of the flow (also referred to as *ex-post* forecasts).

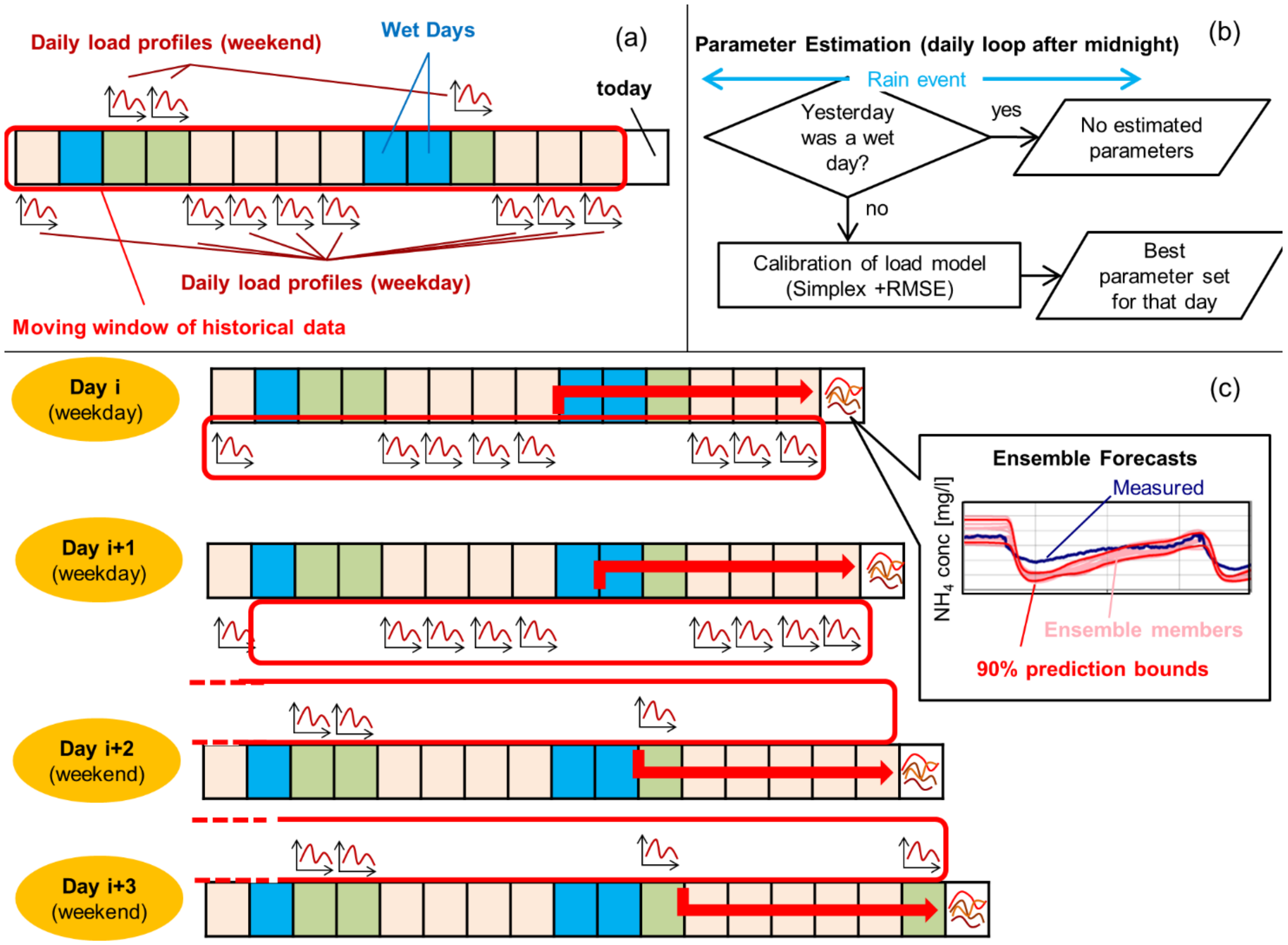

Figure 2. Conceptual schematization of the proposed forecast setup: (a) moving window of historical data (in this example with n = 8 days) and classification into weekdays, weekends, and wet days; (b) procedure for parameter estimation (run every day after midnight); (c) example showing how the forecast model is applied for a period from day i to day $i+1$.

Given the specific setup at the Damhusaaen site, the volume of the primary clarifiers (about 12,000 m^3) is included in the model as it leads to a time delay and to an attenuation of ammonium profiles between the inlet and the sensor location. To account for these issues, the estimated NH_4^+ loads at the inlet (from eq. 1) are routed through three cascading Continuously Stirred Tank Reactors.

The NH_4^+ concentration is then calculated by looking at the outlet of the last tank (eq. 3d). The mass balance of the three tanks is:

$$\frac{dM_{1,SNH4}}{dt} = F_{NH4} - \frac{M_{2,SNH4}}{V} Q \tag{3a}$$

$$\frac{dM_{2,SNH4}}{dt} = \left(M_{1,SNH4} - M_{2,SNH4}\right) \frac{Q}{V} \tag{3b}$$

$$\frac{dM_{3,SNH4}}{dt} = \left(M_{2,SNH4} - M_{3,SNH4}\right) \frac{Q}{V} \tag{3c}$$

$$\mathrm{S_{NH4}}'(\mathrm{t}) = \mathrm{M_{3,SNH4}}(\mathrm{t})/\mathrm{V} \tag{3d}$$

where V [m^3] is the volume of each tank (kept as a calibration parameter), and M_i [kg] are the mass of NH_4^+ as model states.

Furthermore, a preliminary analysis of the NH_4^+ loads measured at the Viby WWTP showed how the morning peak was characterized by a quite steep increase between 08:00 and 10:00. Since the Fourier series used by the model (eq. 1) encountered difficulties in representing such behaviour, an additional ammonium "pulse" was added. This is represented by an asymmetrical term:

$$F'_{NH4}(\mathrm{t}) = F_{NH4}(t) + \gamma_1 e^{\left(-\gamma_2\left(log_{10}(t)-log_{10}(\gamma_3)\right)\right)^2} \tag{4}$$

where γ_1 provides the magnitude of the additional peak (equivalent to a mass added to mimic the steep increase), γ_2 defines the duration of the peak, and γ_3 the timing of the extra peak, constrained to be between 07:00 and 11:00. Compared to a tabular description of the daily profile (as in Langeveld et al., 2017), this formulation was chosen to obtain a profile closer to the observations without significantly increasing the number of parameters.

Estimation of model parameters. The model parameters are estimated by using an optimization routine based on the Simplex method, minimizing the Root Mean Square Error (RMSE) between simulated and observed loads. The optimization is run once per day (just after midnight), using the data collected in the previous 24 hours. An optimal parameter set for the load model ($\theta_{opt,i}$) is obtained for each calendar day. It is assumed that rain-induced phenomena (e.g. first flush, WWTP inlet bypass) would affect the estimation of the NH_4^+ profiles. Therefore, the optimization procedure is not run for wet days. These are defined as the days when the measured flow exceeded the threshold for wet weather (Table 1) that is used to activate the wet weather controls at the plant. Small rain events, generating flows below the threshold, would therefore not be considered, as they would not lead to a change in the plant operations. In daily operation, the threshold value depends on the plant characteristics and status (based e.g. actual capacity of the biological treatment, settling conditions in the secondary clarifier). Once the flow falls back below the threshold, the following 12 hours (24 hours if the event volume is above 20,000 m3) are still characterized as wet periods . This is done because NH_4^+ concentrations are still affected by slow catchment runoff and concentrations are still below typical dry weather values. Days characterized by small rain events, generating flows below the threshold, are classified as dry periods, and thereby included in the calibration.

Evaluation of model performance

Experimental setup. The proposed forecast model was tested on the data collected from June 2018 to May 2019 (i.e. the performance evaluation covered 318 days for both locations). Online operations were mimicked by following the procedure described earlier (Figure 2c).

Performance evaluation. The model performance was calculated on a daily basis by comparing measured NH_4^+ concentrations against the output of the model for the specific *i-th* day. Two performance indicators were used: the Mean Absolute Relative Error (MARE) for evaluating the performance of the ensemble median forecast, and the coverage of observations to evaluate the skill of the ensemble spread:

$$MARE = \frac{1}{k}\sum_{i=1}^{k}\left|\frac{S_{NH4,sim,i}-S_{NH4,obs,i}}{S_{NH4,obs,i}}\right| \tag{5}$$

$$Coverage=\frac{1}{k}\sum_{i=1}^{k} I_i \; with \begin{cases} I_i = 0 \; for \; S_{NH4,sim05,i} > S_{NH4,obs,i} \; or \; S_{NH4,sim95,i} < S_{NH4,obs,i} \\ I_i = 1 \; for \; S_{NH4,sim05,i} < S_{NH4,obs,i} < S_{NH4,sim95,i} \end{cases} \tag{6}$$

Where k is the number of simulated values; $S_{NH4,obs,i}$ is the observations; $S_{NH4,sim,i}$ is the median of the simulated values; $S_{NH4,sim05,i}$ and $S_{NH4,sim95,i}$ are the 5% and the 95% percentile of the simulated values, respectively.

Among the potential applications of the online forecast model, the following options were hypothesized in order to investigate the performance during wet-weather events:

- Estimation of incoming ammonium loads, including potential first-flush peaks from the upstream catchment (as described by e.g. Krebs et al., 1999). Such forecasts can potentially be used to optimize the removal efficiency of the WWTP by using a Model Predictive Control (e.g. Stentoft et al., 2019).
- Estimation of ammonium dilution during a rain event. This information might open the possibility for water-quality based controls involving prioritizing bypass or diverting low pollution flows to natural waters, as described by e.g. Hoppe et al. (2011).

In the first case, the relative error was calculated on the ammonium load for the first 30 minutes of a rain event. In the second case, a contingency table (Bennett et al., 2013) was used to evaluate the ability of the forecast model to estimate dilution in the plant influent. Dilution is here defined a 10% drop of concentration below dry weather values (e.g. if dry weather concentration is 40 mg/l, a dilution event starts when the concentration drops below 36 mg/l). Since ammonium concentrations vary throughout the day and sensor measurements are affected by variability and outliers, the dry weather concentration threshold was defined as the 5th percentile concentration measured during the 2 hours before the start of the event.

RESULTS AND DISCUSSION

Measurement campaigns

The available datasets from the two plants are shown in Figure 4a-d. A total of 35 and 57 wet weather events were observed in Damhusaaen and Viby, respectively. The data at the Damhusaaen WWTP show the influence of the sensor's location within the plant, as the flow (light blue in Figure 4a) is affected by a pumping station, showing spikes often exceeding the wet-weather threshold. Therefore, flow measurements were smoothed by using a simple moving average with 60 steps (dark blue). The daily variations in NH_4^+ concentrations exhibit clear effects of attenuation by the volume of the primary clarifier (Figure 4c). Conversely, the data from the Viby WWTP show the characteristic daily patterns associated with dry-weather WWTP inlets.

Effects of sensor calibration in Damhusaaen is seen in the sudden changes in NH_4^+ concentrations, which in the most extreme cases can jump more than 20 mg/l before and after the calibration. This is consistent with the observations in Cecconi et al. (2019), who highlighted the (potentially negative) influence of the sensor calibration on the sensor readings. The effect of the different calibrations can be seen in the calculated ammonium load profiles (Figure 4e,g), which shows variations in the average daily level. Nevertheless, the daily load profiles measured at both Damhusaaen (Figure 4e,g) and Viby (Figure 4f,h) show a great inter-day variability. An approach based on e.g. only tabular values or fixed ammonium profiles, neglecting the natural variability of the simulated process, would be affected by important uncertainty. This stresses the importance of the proposed ensemble approach to generate model prediction bounds and allow for a more confident application of model forecast for online applications.

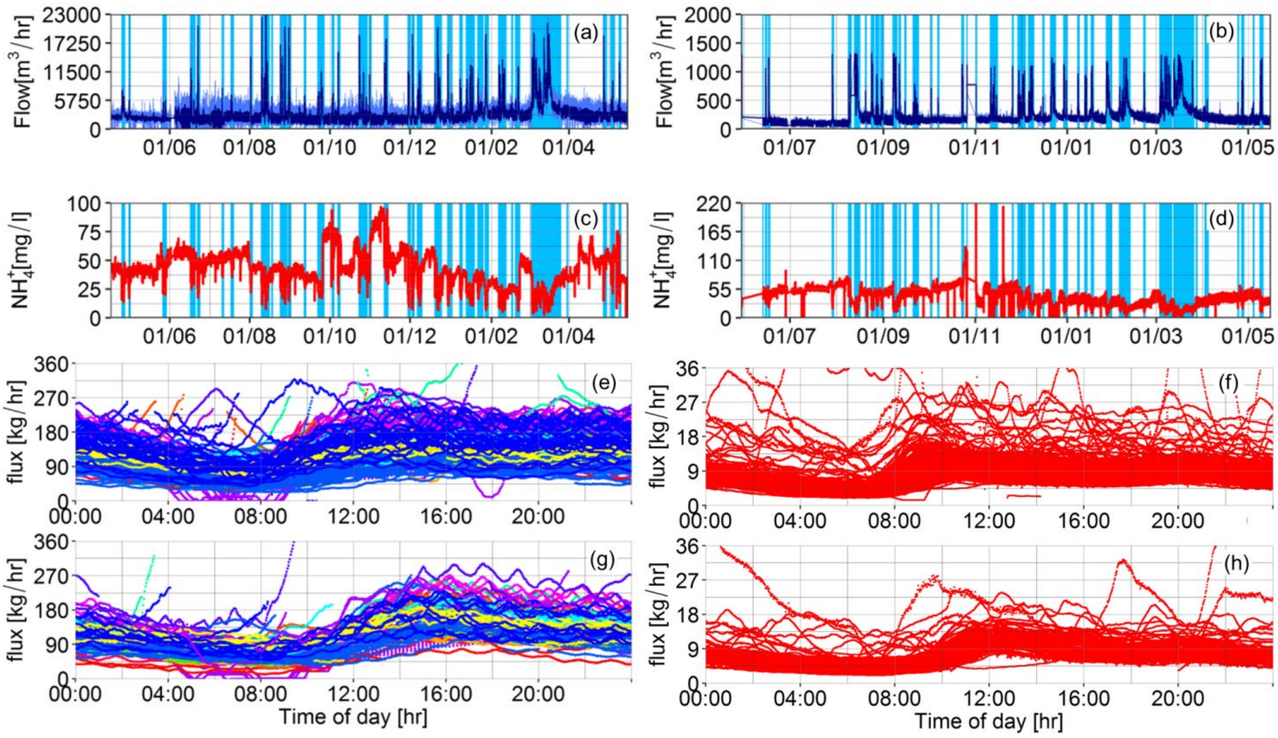

Figure 3. Overview of the measured flow (a,b) and ammonium (c,d) measurement at Damhusaen (left column) and Viby (right column). Blue line: raw flow data; dark blue line: filtered flow data. Wet weather events are shown with light blue background. Measured ammonium loads for weekdays (e,f) and weekends (g,h). Fluxes measured during different calibration periods in Damhusaaen (e,g) are shown by using different colour codes.

Long term performance

The performance of the forecast model over the whole analysed period is shown in Figure 4 where each red dot represents the average skill over a single day. Figure 4 shows how the sensor maintenance resulted in a deterioration of the forecast model performance after the signal correction. In fact, after calibration MARE tends to increase (Figure 4a), while the coverage drops (Figure 4b). This is explained by the fact the model uses a moving window of values preceding the forecast, which may include days where sensor calibration takes place. This result in predictions which are consistently over- or underestimating the NH_4^+ concentrations compared to the signal after the sensor is calibrated. Subsequently, prediction improves in the days following the calibration, since the window moves further, including an increasing number of days with the new sensor calibration and thereby discarding values from the “old” calibration.

Results from Viby (Figure 4c,d) show a deterioration of the forecast model performance following wet weather events. In the periods when the forecast model provided the best performance, MARE ranged 20-25% for both the plants, while the coverage was better in Damhusaaen (often achieving 100%). This can be explained by the difficulties in the proposed model structure to fully describe the specific daily ammonium pattern in Viby. The performance of the forecast model in terms of coverage could be improved by either modifying the model structure (i.e. identifying a better equation than eq. 4) or by increasing the length of the moving window.

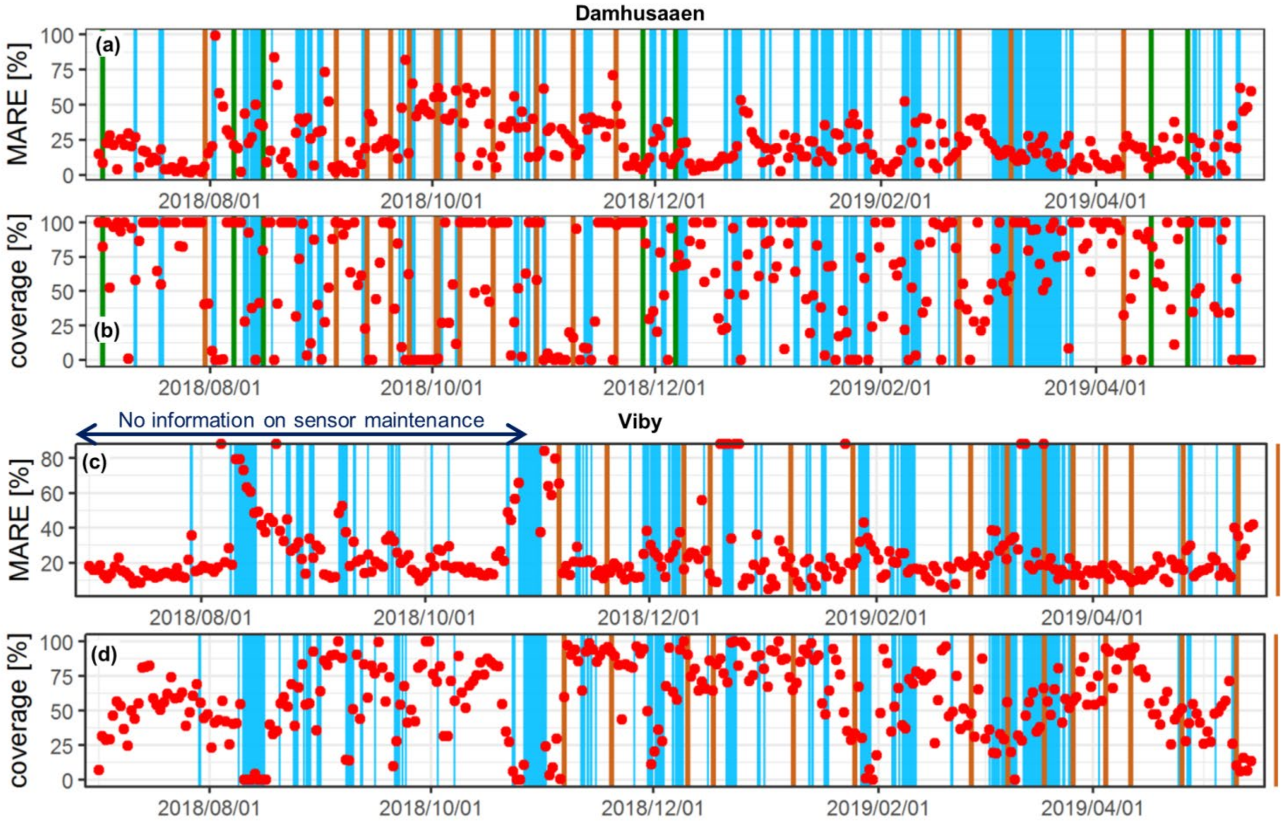

Figure 4. Performance of the forecast model during the simulation period for Damhusåen (a,b) and Viby (c,d). Each red dot represents the average performance for a given day, a blue background indicates wet periods, brown lines indicate sensor calibrations, and green lines indicate sensor cleaning (no calibration). Information on maintenance in Viby before November 2018 is missing.

Performance during wet weather events

Figure 5 provides an overview of the forecast model performance regarding its potential applications for online control applications (e.g. controlling aeration in case of first flush phenomena, or diverting low polluted flow to bypass structures). When looking at the prediction of NH_4^+ in the first phase of a rain event, the forecast model in Damhusaaen (Figure 5a) mostly remains in a ±40% range. In the majority of the events, the load was overestimated. Conversely, in Viby (Figure 5b), the forecast model consistently underestimated the initial load. It should be pointed out that such performance analysis is strongly affected by the sensor calibration as some of the over/underestimation might be explained by events taking place shortly after sensor maintenance. Generally, the data show that rain-induced peaks have longer duration than 30 minutes, i.e. the benefits of using such forecast for predictive control would be limited (after 30 min the wet-weather control would be fully operational). Figure 5c-d show the ability of predicting whether there are dilution effects from stormwater in each time step during an event. Here, the forecast model provided correct predictions (both correct positives and correct negatives) over 75% of the time for 22 events (63%) and 35 events (61%) in Damhusaaen and Viby, respectively. For some events the number of false positives (predicting a dilution while the concentration is still high, i.e. a prediction which might have negative consequences on the environment) was higher in Damhusaaen than in Viby. The number of events where the false positives exceeded 10% of the total event period was 13 (37%) in Damhusaaen and 5 (9%) in Viby. Figure 6 shows four examples of how the forecast model performed during different wet weather events. The examples suggest that the simple modelling approach based on dilution of ammonium loads is sufficient to grasp the dynamics at the beginning of a rain event, while it fails to represent the behaviour in the receding phase (Figure 6e).

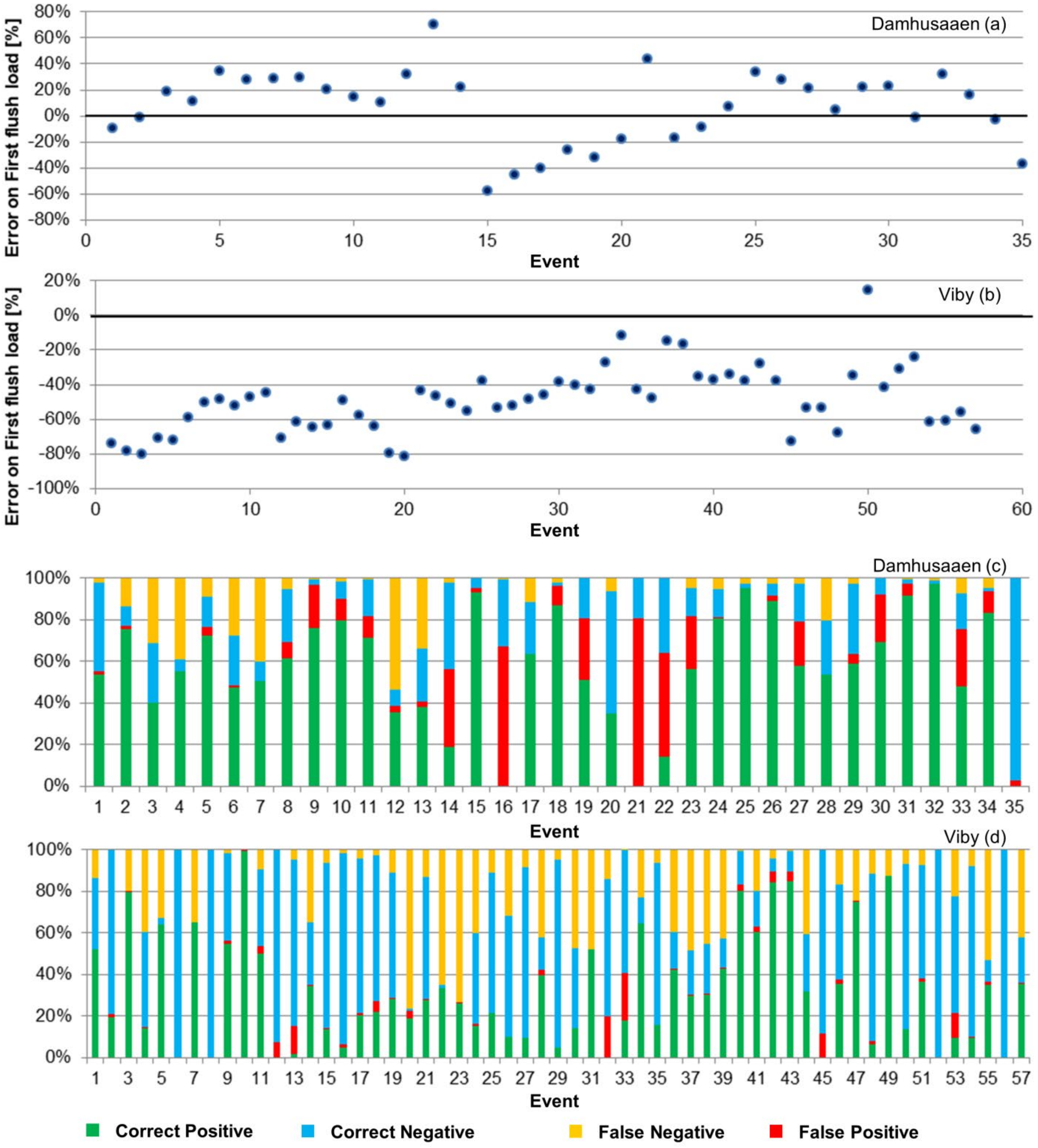

Figure 5. Forecast model performance during wet weather events in Damhusaaen (a,c) and Viby (b,d). (a,b) relative error in estimation of first flush load (first 30 min); graphical visualization of contingency table for dilution prediction (c,d).

Such behaviour is in line with the findings of Langeveld et al. (2017), who added an additional term to the model structure in order to obtain a better representation of the transition from the wet weather concentrations back to dry weather values. Considering that the majority of online applications for the proposed forecast model would focus on the initial phase of the event, structural shortcomings in the end of the events are not expected to affect its performance. Furthermore, this analysis illustrates the importance of evaluating the model performance in terms of its potential applications, instead of limiting the analysis to an evaluation of the model residuals.

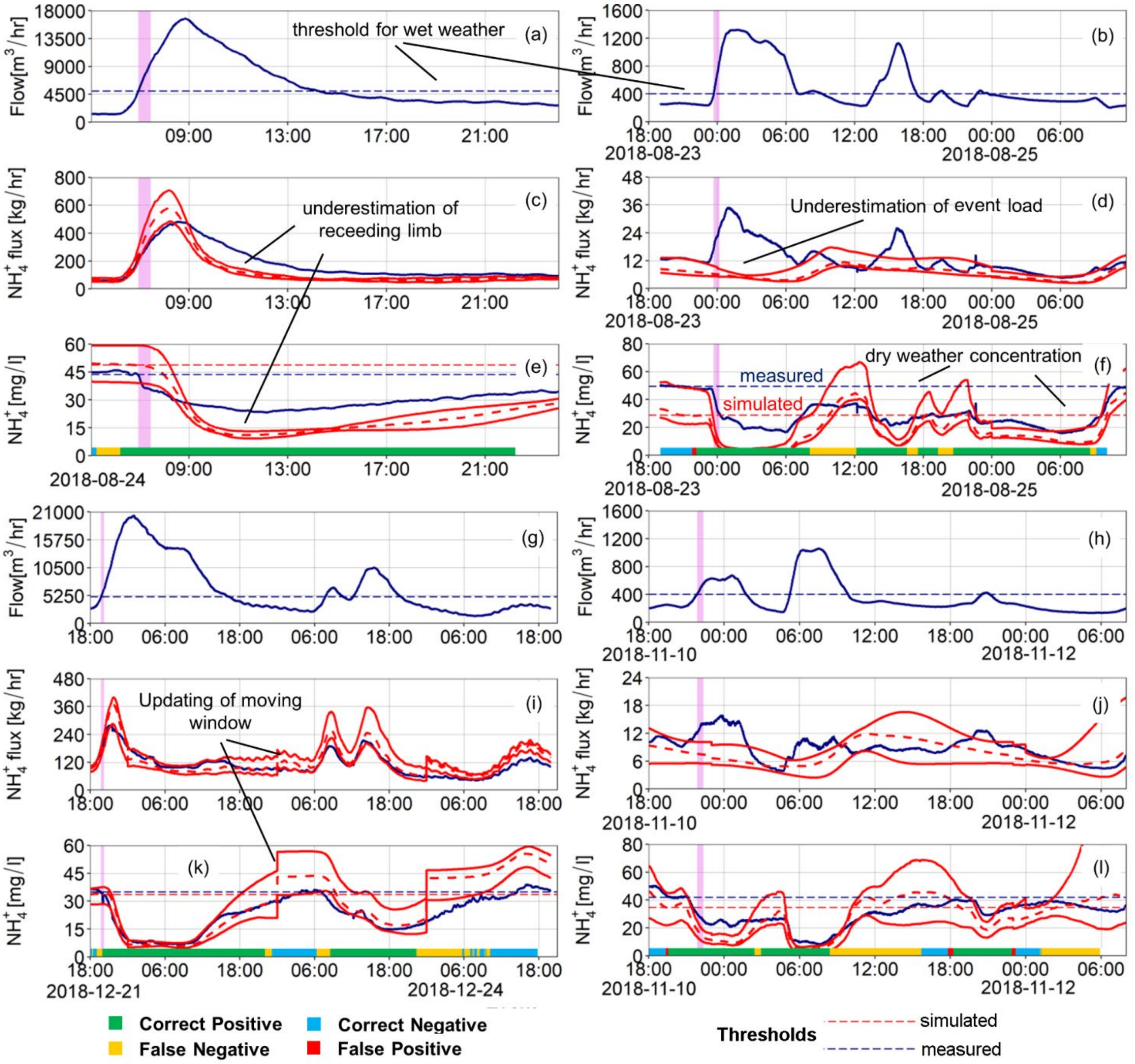

Figure 6. Examples of the forecast model predictions for selected events at the inlet of Damhusaaen (left column) and Viby (right column). Measured flow (a-b,g-h); measured and simulated NH_4^+ fluxes (c-d,i-j); and concentrations, along with dry weather NH_4^+ concentrations used to define dilution events (e-f,k-l). The values from the corresponding contingency table are shown at the bottom of the concentration graph (e-f,k-l). Simulated values are described as median values (dashed line) and 90% prediction bounds (solid line). Violet background identifies the first 30 min of the event used to evaluate first flush.

Figure 6(d,j) confirm the inadequacy of the model structure in Viby in representing the overall events. The available measurements suggest an increase in the ammonium loads, which might resemble the process described by Krebs et al. (1999). The model structure should therefore be adapted accordingly. Nevertheless, Figure 6(f,l) show how the model was still capable of detecting the dilution during wet weather events. Figure 6(i,k) also highlight an issue linked to the moving window approach: the updating of the window can in fact result in discrete jumps of the predicted values. For an ensemble-based approach as the one in this study, these variations could be reduced by increasing the size *n* of the window.

Research gaps and future developments

The available results show how the predictions of the proposed setup are strongly dependent on the signal provided by the ion-selective sensors and by the maintenance operations. As pointed out by Cecconi et al. (2019), excessive maintenance and/or improper sensor calibration might significantly decrease the reliability of ammonium measurements. Furthermore, the performance of the forecast model cannot be evaluated in a complete manner, since a major cause of poor indicator values is due to changes in the signal rather than in the model structure and/or parameters. Possible improvement of the proposed approach might include:

- Correction and transformation of the signal from the ion-selective sensor. Ideally, the raw signal from the sensor could provide a more reliable data source compared to the existing situation. A follow-up study on how to correct the raw ISE-based signal is currently being undertaken.
- Variable uncertainty description: the proposed ensemble approach equally weighs all the days in the calibration window. Approaches such as exponential filtering, weights on most recent days, etc. can be used to increase the importance of the most recent data for cases where this is desired.
- Use of stochastic models that combine a deterministic model with a stochastic term capable of describing the natural variability of the NH_4^+ concentrations. Possible techniques include the so-called grey-box models or the external bias description (Del Giudice et al., 2015).
- A thorough evaluation of the influence of the different model parameterizations on the resulting predictions. The effect of the length of the moving window, the number of ensemble members, the intervals used in the parameter estimation, etc. should be fully evaluated. This would provide robust guidelines for a wide application of the proposed method to other systems.
- Performance evaluation of the proposed approach at CSO structures, i.e. where the installation of a permanent online NH_4^+ sensor is less likely compared to WWTPs. Here the model would use historical data from monitoring campaigns of limited duration (e.g. an online sensor installed over a 2-3 months period) to forecast NH_4^+ concentrations. The duration of the historical dataset should be sufficient to confidently estimate representative daily profiles (and their variations).
- Performance evaluation using real flow forecasts, based on e.g. radar rainfall forecasts or numerical weather predictions. There is ample research on flow uncertainty estimation, and this comparison could show if the uncertainties discussed here are significant compared with those related to the rainfall forecasts, which are known to be very large.
- Performance evaluation based on event definitions that are not strictly linked to the plant operational settings. For example, a variable flow threshold, defined on the actual dry weather flow conditions rather than the used fixed value, could improve the understanding of the model behaviour in wet weather.
- Modification of the model structure by including additional terms, such as those included in the model from Langeveld et al. (2017). This would expand the applicability of the model to other applications outside WWTP control (e.g. for quantification of CSO and bypass load)

CONCLUSIONS

This work investigated the performance of simple model for online prediction of NH_4^+ concentrations intended for real time control applications. The analysis of the forecast model results showed that:

- The analysis of data from two Danish WWTPs showed high inter-diurnal variations in the incoming ammonium loads. This underlines the importance of using an uncertainty-based approach, which explicitly accounts for this natural variability.

- The simple structure of the model (based on Fourier series) should be adapted to the specific locations of the sensors and/or to the characteristics of the catchment.
- Calibration of the ammonium ISE sensors significantly affected the model performance. The proposed data-driven forecast model uses data from previous calibration periods, and its capability of matching the measured values drops just after the sensor is calibrated. This suggests a strong need for new approaches that can reduce the impact of the sensor calibration on the operation of online forecast models.
- The performance of the forecast model in relation to potential online control strategies provide satisfactory results. Specifically, the model provided good simulations of both the ammonium loads at the beginning of a rain event and the dilution induced by wet-weather events.

Overall, the proposed data-driven forecast model creates interesting opportunities for online forecasts of WWTP influent quality. Although further research is needed to improve the accuracy of the forecast model in terms of predicted concentrations, it can already open various possibilities for the implementation of online control strategies. The forecast model can also be applied for forecasting of incoming NH_4^+ loads and concentrations, creating new opportunities for Model Predictive Control of WWTPs.

ACKNOWLEDGEMENT

This research was financially supported by the Innovation Fund Denmark through the Water Smart Cities (WSC) project [grant number 5157-00009B]. The authors thank Tomas Trabjerg (Aarhus Vand A/S) for the information regarding the measurements campaign at the Viby WWTP.

REFERENCES

Ahm, M., Thorndahl, S., Rasmussen, M. R., Bassø, L. (2013) Estimating subcatchment runoff coefficients using weather radar and a downstream runoff sensor. Water Science and Technology, 68(6), 1293–1299.

Alferes, J., Tik, S., Copp, J., Vanrolleghem, P. A. (2013) Advanced monitoring of water systems using in situ measurement stations: data validation and fault detection. Water Science and Technology, 68(5), 1022–1030.

Bechmann, H., Nielsen, M. K., Madsen, H., Kjølstad Poulsen, N. (1999) Grey-box modelling of pollutant loads from a sewer system. Urban Water, 1(1), 71–78.

Bennett, N. D., Croke, B. F. W., Guariso, G., Guillaume, J. H. A., Hamilton, S. H., Jakeman, A. J., Marsili-Libelli, S., Newham, L. T. H., Norton, J. P., Perrin, C., Pierce, S. A., Robson, B., Seppelt, R., Voinov, A. A., Fath, B. D., Andreassian, V. (2013) Characterising performance of environmental models. Environmental Modelling and Software, 40, 1–20.

Cecconi, F., Reifsnyder, S., Ito, Y., Jimenez, M., Sobhani, R., Rosso, D. (2019) ISE-ammonium sensors in WRRFs: field assessment of their influencing factors. Environmental Science: Water Research & Technology, 5(4), 737–746.

Flores-Alsina, X., Saagi, R., Lindblom, E., Thirsing, C., Thornberg, D., Gernaey, K. V., Jeppsson, U. (2014) Calibration and validation of a phenomenological influent pollutant disturbance scenario generator using full-scale data. Water Research, 51(0), 172–185.

Gernaey, K. V., Flores-Alsina, X., Rosen, C., Benedetti, L., Jeppsson, U. (2011) Dynamic influent pollutant disturbance scenario generation using a phenomenological modelling approach. Environmental Modelling & Software, 26(11), 1255–1267.

Del Giudice, D., Löwe, R., Madsen, H., Mikkelsen, P. S., Rieckermann, J. (2015) Comparison of two stochastic techniques for reliable urban runoff prediction by modeling systematic errors. Water Resources Research, 51(7), 5004–5022.

Haimi, H., Mulas, M., Corona, F., Vahala, R. (2013) Data-derived soft-sensors for biological wastewater treatment plants: An overview. Environmental Modelling & Software, 47, 88–107.

Hoppe, H., Messmann, S., Giga, A., Gruening, H. (2011) A real-time control strategy for separation of highly polluted storm water based on UV-Vis online measurements - from theory to operation. Water Science and Technology, 63(10), 2287–2293.

Krebs, P., Holzer, P., Huisman, J. L., Rauch, W. (1999) First flush of dissolved compounds. Water Science and Technology, 39(9), 55–62.

Kusiak, A., Verma, A., Wei, X. (2013) A data-mining approach to predict influent quality. Environmental Monitoring and Assessment, 185(3), 2197–2210.

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

Martin, C., Vanrolleghem, P. A. (2014) Analysing, completing, and generating influent data for WWTP modelling: A critical review. Environmental Modelling & Software, 60, 188–201.

Métadier, M., Bertrand-Krajewski, J. L. (2012) The use of long-term on-line turbidity measurements for the calculation of urban stormwater pollutant concentrations, loads, pollutographs and intra-event fluxes. Water Research, 46(20), 6836–6856.

Newhart, K. B., Holloway, R. W., Hering, A. S., Cath, T. Y. (2019) Data-driven performance analyses of wastewater treatment plants: A review. Water Research, 157, 498–513.

Nielsen, M. K., Onnerth, T. B. (1995) Improvement of a recirculating plant by introducing star control. Water Science and Technology, 31(2), 171–180.

Schilperoort, R. P. S., Dirksen, J., Langeveld, J. G., Clemens, F. H. L. R. (2012) Assessing characteristic time and space scales of in-sewer processes by analysis of one year of continuous in-sewer monitoring data. Water Science and Technology, 66(8), 1614–1620.

Stentoft, P. A., Munk-Nielsen, T., Steen Mikkelsen, P., Madsen, H. (2017) "A Stochastic Method to Manage Delay and Missing Values for In-Situ Sensors in an Alternating Activated Sludge Process" in Proceedings of 12th IWA Specialized Conference on Instrumentation, Control and Automation, Quebec City, Canada, 11th-14th June 2017.

Stentoft, P. A., Munk-Nielsen, T., Vezzaro, L., Madsen, H., Mikkelsen, P. S., Møller, J. K. (2019) Towards model predictive control: online predictions of ammonium and nitrate removal by using a stochastic ASM. Water Science and Technology, 79(1), 51–62.

Talebizadeh, M., Belia, E., Vanrolleghem, P. A. (2016) Influent generator for probabilistic modeling of nutrient removal wastewater treatment plants. Environmental Modelling & Software, 77, 32–49.

Vezzaro, L., Lund Christensen, M., Thirsing, C., Grum, M., Mikkelsen, P. S. (2013) Water quality-based real time control of integrated urban drainage: a preliminary study from Copenhagen, Denmark. Procedia Engineering, 70, 1707–1716.

Yu, T., Yang, S., Bai, Y., Gao, X., and Li, C. (2018) Inlet water quality forecasting of wastewater treatment based on kernel principal component analysis and an extreme learning machine. Water (Switzerland), 10(7), 873.

Yuan, Z., Olsson, G., Cardell-Oliver, R., van Schagen, K., Marchi, A., Deletic, A., Urich, C., Rauch, W., Liu, Y., and Jiang, G. (2019) Sweating the assets – The role of instrumentation, control and automation in urban water systems. Water Research, 155, 381–402.