Journal of Water Process Engineering 69 (2025) 106821

Contents lists available at ScienceDirect

Journal of Water Process Engineering

journal homepage: www.elsevier.com/locate/jwpe

Short-term memory artificial neural network modelling to predict concrete corrosion in wastewater treatment plant inlet chambers using sulphide sensors

J. Mendizabal [a], D. Vernon [b], B. Martin [b], Y. Bajón-Fernández [a], A. Soares [a,*]

[a] Cranfield Water Science Institute, Cranfield University, Bedfordshire MK43 0AL, UK
[b] Thames Water, Reading STW, Reading RG2 0RP, UK

ARTICLE INFO

Editor: Guangtao Fu

Keywords:
Long short-term memory
Recurrent neural network
Monitoring
Rising main
Corrosion

ABSTRACT

Sulphide accumulation in lengthy rising mains can lead to significant concrete corrosion and odour issues at manholes and wastewater treatment plants (WWTPs). Monitoring dissolved sulphide, typically relies on auto-sampling or grab samples followed by laboratory analysis, remains underdeveloped. This study aimed to identify sources of concrete corrosion sources at a WWTP inlet chamber and develop a sulphide prediction model using artificial intelligence (AI). A dissolved sulphide sensor was installed at three rising mains (RM1 to RM3) and the combined inlet at a full-scale WWTP, providing a 5-minute resolution data that revealed a daily hydrogen sulphide (H_2S) pattern that inversely correlated with the flow rate. RM1 exhibited the highest sulphide load, peaking at 3.6 kg/d during cold months and 4.2 kg/d during warm months. RM3 and RM2 recorded loads of 2.96 kg/d and 0.98 kg/d, respectively, during cold months. A long short-term memory (LSTM) artificial neural network (ANN) model was developed to predict H_2S concentrations at RM1, using flow rate, temperature, and time of day as inputs. The model achieved a root mean square error (RMSE) of 0.34 and a Nash-Sutcliffe efficiency (NSE) of 0.57, accurately predicting the daily H_2S pattern. This study's main contributions include insights into sulphide dynamics from high-resolution sensor data, which could support corrosion management as part of a septicity warning system or feedforward control for sulphide treatment. Additionally, the AI-based prediction model offers potential for sensor repurposing, saving both capital and operational costs.

1. Introduction

Wastewater delivery to centralised wastewater treatment plants (WWTPs) can result in wastewater becoming septic, particularly at large sewerage networks. Septicity is characteristically identified by wastewater having a dark colour and rotten eggs smell. However, septicity is not a risk due to its colour or smell, it is mainly a risk due to the dissolved sulphide within the wastewater and its volatilisation to the air. The sulphide contained in septic wastewater can get stripped to the air in gravity sewers and high turbulence areas, such as drops [1]. The stripped hydrogen sulphide (H_2S) is known to cause corrosion to concrete and metal structures. Low concentrations of dissolved sulphide, even <0.5 mg S/L, have been found to cause moderate concrete corrosion [2]. The cost of microbial induced concrete corrosion (MICC) was found to be of £85 million per year in the UK and €450 million per year in Germany [3] and of €5 million per year in Flanders [4]. Increasing population growth in cities and urban areas will result in larger sewerage networks, which is likely to further increase septicity problems. Furthermore, application of water saving strategies [5] and the increase of global temperatures due to climate change [6] will further exacerbate the problem.

In sewerage networks, the first step to manage septicity is to identify septicity formation and H_2S release hotspots. This is usually done by using one of the following methods: i) Measurement of sulphide (either dissolved or H_2S in the gas phase) throughout the network or ii) Utilisation of a biochemical modelling tool such as WATS [7] or SEWEX [8]. The first method consists of directly measuring the dissolved sulphide by grab samples, which has the disadvantage of being labour intensive, not capturing the temporal variations and low precision due to the unstable nature of sulphide [9], data not instantly available (lab analysis time)

* Corresponding author.
E-mail addresses: j.mendizabal-bengoetxea@cranfield.ac.uk (J. Mendizabal), dejan.vernon@thameswater.co.uk (D. Vernon), ben.martin@thameswater.co.uk (B. Martin), y.bajonfernandez@cranfield.ac.uk (Y. Bajón-Fernández), a.soares@cranfield.ac.uk (A. Soares).

https://doi.org/10.1016/j.jwpe.2024.106821
Received 5 April 2024; Received in revised form 9 December 2024; Accepted 17 December 2024
Available online 26 December 2024
2214-7144/© 2025 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

and high costs. Hydrogen sulphide gas monitoring offers real-time monitoring opportunities utilising gas monitors. However, it has other disadvantages, such as not giving a direct measure of the septicity issue and high dependency on the positioning of the monitor [10]. The use of modelling tools also requires sulphide measurement throughout the network when setting up the model. Therefore, practical and precise ways of measuring dissolved sulphide in real-time are required.

Several dissolved sulphide sensors can be found in literature, but most of them rely on stripping the dissolved sulphide from the sample and measuring the H_2S gas and then calculating the dissolved sulphide based on Henry's law [11–13]. This type of sensors is usually large and expensive, as they require a stripping chamber [14]. There are only a few sensors that can measure sulphide in wastewater. The first use of a sensor to measured dissolved sulphide was reported by Sutherland-Stacey et al. 2008 [9], which relied on a UV/VIS spectrometer to measure the absorbance of the bisulphide ion and then combine it with a pH sensor to calculate the total dissolved sulphide. Recently, two more sensors have been developed, both based on electrochemical principles, measuring H_2S ions in wastewater. The sensor developed by Yavarinasab et al. [14] has been only reported to trialled in aqueous solution and in a lab set-up and it is not commercially available. The other is a commercially available SulfiLogger sensor [15], which has been applied in sewer management and optimisation of sulphide mitigation through chemical dosing [16].

Although the dissolved sulphide sensors offer high temporal resolution and precision, the cost of the sensors and maintenance required can incur high costs for users, with costs decreasing if these become mainstream and can be rented by water utilities. Nevertheless, costs can be reduced by the development of models to predict the dissolved sulphide concentration at a single point would offer great advantage, as they would allow for the sensors to be repurposed (moved to another location) once the model had been trained. Mechanistic models, such as WATS and SEWEX, offer great insights into the sewer processes and are widely applied for sulphide prediction. However, they require extensive sampling campaigns measuring many parameters. On the other hand, data-driven models can work with less parameters and still provide accurate predictions, as has been shown for sulphide formation prediction [17,18] and H_2S emissions at WWTP inlets [19]. Deep learning and particularly long short-term memory (LSTM) artificial neural network (ANN) has been shown to be great at sequential data prediction, as was the case for traffic prediction [20] and sewer flow prediction [21]. Machine learning approaches have been intensely investigated in the last few decades as the use of sensors is becoming more widely spread, making data available in tandem with computer processing power growing quickly and internet of things allowing for large sets of data transfers instantly. This approach is also attractive to make savings and optimise WWTPs for higher effluent quality, energy management, to maintenance needs. The optimisation of WWTPs is very complex with many independencies and traditional methods based on sampling and modelling only address objectives partially [22]. Recent reviews have been published indicating the application, advantages and limitations of machine learning tools in WWTPs. As explained by Fu et al. [23] machine learning algorithms like random forests, support vector machines, and long short-term memory (LSTM) networks are now can handle large and complex data sets and can be widely applied for urban wastewater management by dynamically adjusting processes based on real-time data. In the recent review by Xia et al. [24] highlights the growth and expansion of this field and breakthrough areas such as microbial diversity and micropollutant removal. Others demonstrate the impact of AI solutions around wastewater influent characterisation using deep-learning based detection for water quality [25], process optimisation [26] as well as microbial pathways [27]. Moreover, as the sector shifts towards sustainable wastewater management, machine learning plays a vital role in resource recovery, such as nutrient extraction thereby aligning with environmental goals [28]. Duarte et al. [29] have provided a detailed review of data-driven models and overall indicated the advantages of LSTM in WWTPs handling sequential data and capture long-term dependencies, making them superior to artificial neural networks (ANNs), support vector machines (SVMs), and autoregressive integrated moving average (ARIMA) for dynamic predictions. While hybrid models enhance accuracy, they are computationally intensive compared to standalone LSTM. As such long short-term memory (LSTM) artificial neural network (ANN) were selected for this study, which aims to i) Determine the source of concrete corrosion at a WWTP inlet chamber where several rising mains were discharging using dissolved sulphide sensor (SulfiLogger) used for the first time for this application, ii) Explore sulphide dynamics with high-resolution data aimed to reveal daily H_2S patterns in relation to flow rate and temperature, offering detailed insights into sulphide load variations across different mains, and iii) Create an artificial intelligence-based model, specifically a long short-term memory (LSTM) artificial neural network (ANN), to predict H_2S concentrations based on real-time data from the dissolved sulphide sensor.

Table 1
Characterisation of the rising mains monitored and wastewater.

Rising main	Diameter (m)	Length (m)	Dry weather flow (m^3/day)	Nitrate dosing (Y/N)
1	0.6	6800	6001	Y
2	0.6	5300	4908	N
3	0.38	3100	2960	N

Wastewater characterisation				
	pH	TSS (mg/L)	sCOD (mg/L)	Ammonia (mg/L)
Mean	7.78	450	200	40
STdev	0.27	103	39	3.6

2. Materials and methods

2.1. Site description

The study WWTP is in the UK and treats wastewater for around 99,000 population equivalent (PE), mainly from domestic sources. The wastewater is collected at an inlet chamber, with 12 rising mains of different diameters discharging into it. The inlet chamber presented signs of corrosion. The source of the corrosion was unknown due to the high number of rising mains discharging into the same chamber and the objective of the work was to identify the key contributors to the overall corrosion.

Out of the 12 rising mains, 3 were selected to be monitored, due to their size, colour of the wastewater, anecdotical evidence and a previous catchment wide septicity potential study. For this study, these have been labelled as rising main 1, 2 and 3 (RM1, RM2, RM3) and their characteristics defined in Table 1.

Rising main 1 (RM1) had calcium nitrate dosing at the pumping station wet well during the monitoring period. The calcium nitrate dosing was controlled based on temperature, wastewater flow and potentially more parameters (the dose was controlled by a contractor under a patented control philosophy). The average daily calcium nitrate dose was of 338.6 L/day. Grab samples were taken at the combined inlet and the wastewater characterised, presenting typical characteristics for the UK Table 1.

2.2. SulfiLogger sensor

SulfiLogger S1/X1-1020 (SulfiLogger A/S, Risskov, Denmark) were used in this study (hereby referred to as SulfiLogger). The SulfiLogger sensor electrochemically measures H_2S by detecting the current generated when H_2S passes through a silicone membrane and undergoes

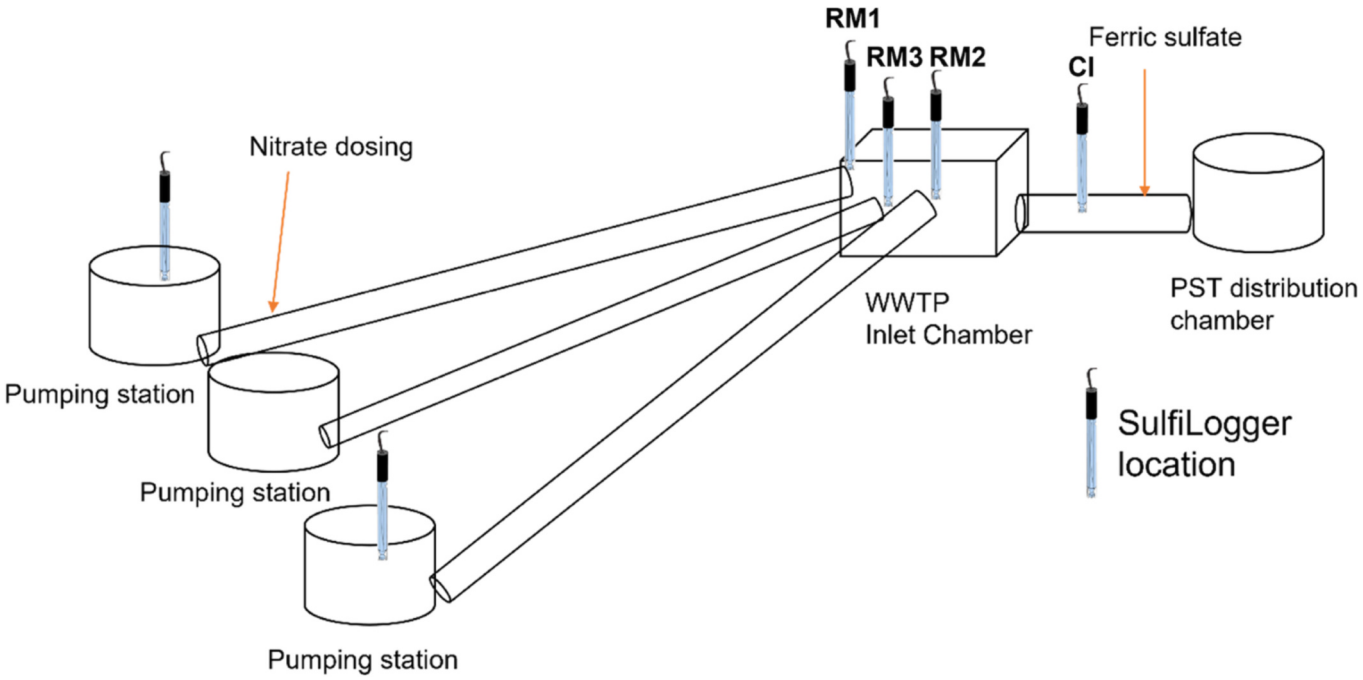


Fig. 1. SulfiLogger sensor locations.

oxidation. This current correlates with H_2S concentration in liquid or gas phases [16]. The sensor is made of stainless steel, it's lightweight (0.85 kg), compact and has a passive anti-fouling flush front [15]. The sensor has a measuring range of 0–5 mg H_2S/L and can be set-up at a measuring frequency as low as 1 min. The sensor was calibrated every 3 months, according to the manufacturer's instructions, by mounting the calibration cap to the sensor for a feed of calibration gas at 1000 mg/L H_2S [15]. Also important to consider is that the dissolved sulphide speciates differently depending on the pH of the wastewater (Eq. (1)) [30] as sulphide ions (S^{2-}), bisulfide ions (HS^-), and molecular hydrogen sulphide (H_2S). At lower pH values (more acidic conditions), hydrogen sulphide predominantly exists as H_2S. As pH increases (becoming more basic), H_2S dissociates into HS^-, and at even higher pH values, further dissociates into S^{2-}.

$$S^{2-} + 2H^+ \xleftarrow{K_{a2}\approx 0.8\bullet 10^{-17}} H^+ + HS^- \xleftarrow{K_{a1}\approx 1.0\bullet 10^{-7}} H_2S_{(aq)} \quad (1)$$

2.3. Sensor location

Six SulfiLogger sensors were acquired and installed, as seen in Fig. 1. Two were installed at two terminal pumping stations, for RM1 and RM2, respectively. They were installed hanging with a chain into the wet well and connected to a PowerCom Box, which provided power to the sensor and sent the data to the collection platform.

Three were installed at the outlet of the rising mains, within the inlet chamber. These were installed using a scaffolding and were introduced into the respective bellmouth. The last one was installed at the combined inlet after the inlet chamber.

2.4. Monitoring campaign

The sensors were installed and data was collected for 9 months (start of summer to next spring) at a measurement frequency of 5 min. The sensors data were further validated by grab sample analysis and comparison to the sensor measurement. The sensor measurements were stored in the PowerCom Box and were transmitted through a phone

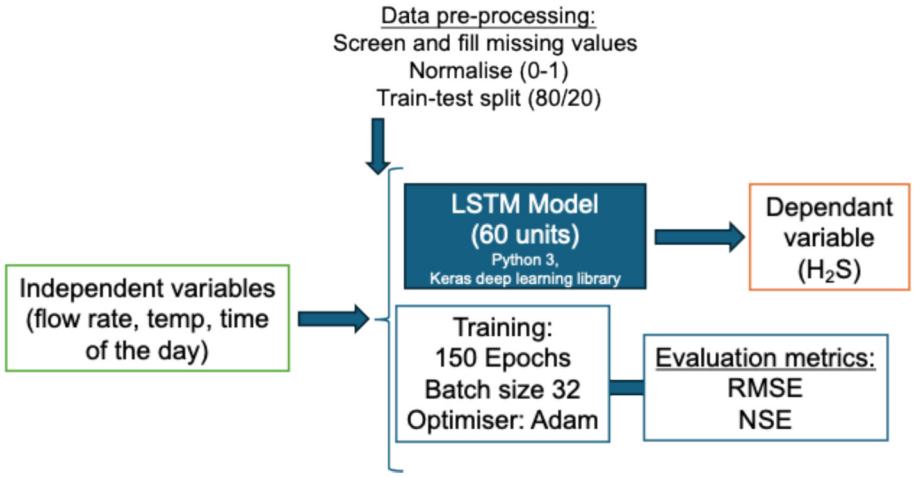


Fig. 2. Illustration of summarizing the LSTM model training and prediction process.

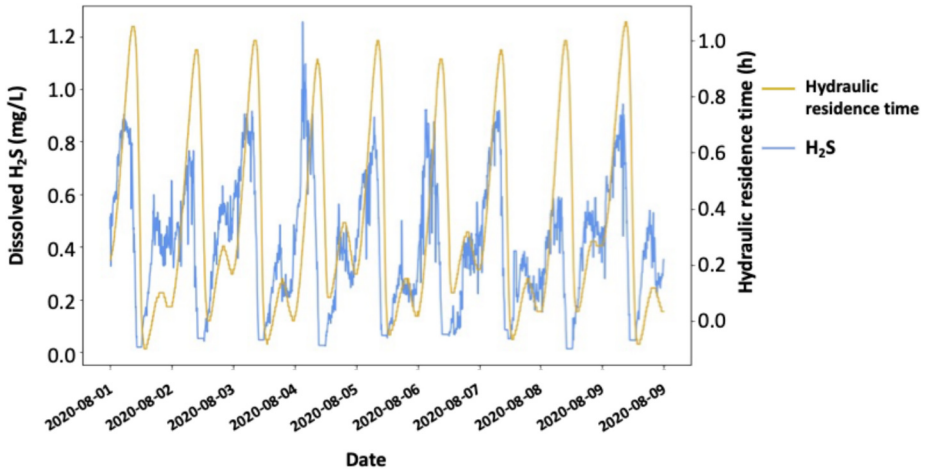


Fig. 3. Hydrogen sulphide (H_2S) concentration at the outlet of RM1 and the hydraulic residence time of the wastewater at RM1 between 01 and 08-2020 and 10-08-2020.

signal to the data collection webpage every half a day. The PowerCom Box batteries had to be replaced every 3 months, which was conveniently done when calibrating the sensors.

The monitoring campaign started on July 2020 with 3 sensors installed at the wet well of RM1, at the outlet of RM1 and at the combined inlet to the WWTP. After 3 months of monitoring, 3 more sensors were installed at the wet well of RM2, outlet of RM2 and outlet of RM3 (Fig. 1).

2.5. Analytical measurements

All the measurements follow methods described in the APHA [31]. Ammonia, total and soluble COD and dissolved sulphide were measured using Spectroquant cell tests, according to the manufacturer instructions (Merk, UK). Other parameters such as pH, ORP, DO and temperature were measured using hand held probes and instruments, in-situ at the WWTP using a high quality digital (HQD) meter with Intellical probes, according to the manufacturer instructions (PHC101, MTC101 and LDO101, respectively) (Hach, UK).

2.6. LSTM model set-up

A LSTM model was developed to predict the sulphide concentration at the outlet of RM1, based on readily available input parameters (Fig. 2). The input parameters selected were wastewater flowrate on the rising main, wastewater temperature and time of the day. The available data was cleaned by replacing missing values with the average between the two closest data points. Then, the data was normalized for each parameters using the min-max normalisation technique (Eq. (2)), which converted all indicators to values between 0 and 1.

$$x_{norm} = \frac{x_i - min(x)}{max(x) - min(x)} \tag{2}$$

The LSTM model was trained to predict H_2S concentration (the dependent variable) using flow rate, temperature, and time of day as independent variables. First, the data was pre-processed by filling missing values and normalizing all inputs to a 0–1 scale for efficient model convergence (Fig. 2). The dataset was then split, with 80 % used for training and 20 % for testing. The model architecture included one hidden LSTM layer with 60 units and a dense output layer, optimized using the Adam optimizer and mean squared error (MSE) loss. Training occurred over 150 epochs with a batch size of 32, employing back-propagation through time to minimize prediction errors. After each epoch, performance was validated on the test set using root mean square error (RMSE) and Nash-Sutcliffe efficiency (NSE). Hyperparameters were tuned to avoid overfitting, resulting in a single-layer LSTM as the optimal configuration. The LSTM model was developed in Python 3, using the Keras deep learning library. Other Python libraries were also used, such as NumPy, Scikit-learn and matplotlib. The model uses a single-step prediction horizon, forecasting the immediate next H_2S concentration based on recent input data. The methods help the model generalize effectively, reducing the risk of memorizing training data-specific patterns. To prevent overfitting, techniques like dropout regularization in LSTM layers (2014), can be applied [32].

2.7. Model performance comparison

The performance of the LSTM model was assessed by comparing it to a linear and a polynomial regression model. The polynomial regression was fitted to a sixth grade polynomial. The model performance criteria used in this study were the root mean square error (RMSE) and the Nash-Sutcliffe efficiency (NSE). The RMSE was calculated as Eq. (3).

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}\left(y_i^{obs} - y_i^{sim}\right)^2}{n}} \tag{3}$$

NSE is a normalized statistic that determines the relative magnitude of the residual variance ("noise") compared to the measured data variance ("information"). The NSE was calculated as Eq. (4).

$$NSE = 1 - \left[\frac{\sum_{i=1}^{n}\left(y_i^{obs} - y_i^{sim}\right)^2}{\sum_{i=1}^{n}\left(y_i^{obs} - y^{mean}\right)^2}\right] \tag{4}$$

3. Results and discussion

3.1. Monitoring results

The sensors at the wet wells of RM1 and RM2 showed that no dissolved sulphide was present, with concentrations below 0.01 mg/L H_2S being measured consistently within the 9 and 6 months monitoring, respectively. The H_2S concentrations at the outlets of RM1, RM2 and RM3, as well as the combined inlet (CI) concentrations, showed a very distinct daily pattern throughout the dry days of the monitoring period (Fig. 3). The H_2S concentration increased during the night time (00:00 to 8:00), was reduced during the morning peak (8:00 to 12:00) and was moderate during the afternoon and evening (12:00 to 00:00). The H_2S concentration profile was the inverse of the daily flow pattern to a typical WWTP [33] and therefore, followed quite well the pattern of the hydraulic residence time (HRT) in the rising main. This pattern aligns with known behaviour in wastewater systems, where longer HRT allows

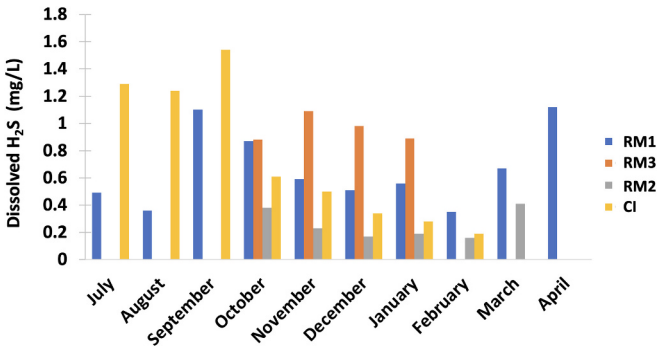


Fig. 4. Hydrogen sulphide (H_2S) monthly averages at RM1, RM2 and RM3 outlets and at the combined inlet of the WWTP.

more time for anaerobic conditions to develop, leading to increased microbial activity that produces H_2S as a byproduct. Conversely, shorter HRT during high-flow periods reduces the time available for these anaerobic processes, resulting in lower H_2S concentrations. The dependence of H_2S formation on HRT is well known, as had been shown by several authors when proposing empirical equations for sulphide prediction that included the HRT term [34–36]. The H_2S was always measured as 0 mg/L, i.e. below detection limit, in wet weather conditions, due to HRTs <4 h in the rising mains and possibly also due to the diluted wastewater, as COD has also been shown to impact the sulphide generation proportionally [37].

Interestingly, the H_2S pattern measured at the RM1 outlet was very similar to the pattern measured at the RM2 and RM3 outlets, which indicated that the nitrate dosing was not being as effective as intended. The nitrate dosing at the RM1 wet well was controlled based on wastewater temperature and RM1 flowrate. Therefore, it was expected to achieve a consistent H_2S concentration at the RM1 outlet, which was originally designed to keep H_2S below 0.1 mg/L. The fact that the H_2S pattern was similar to RM2 and RM3, which were not dosed, suggest the nitrate is being consumed faster than what was originally expected. This phenomenon has been observed and reported by [38] at a rising main in Australia and was linked with the growth of denitrifying biofilms within the rising main once the nitrate dosing was started.

There were several issues with some of the sensors during the monitoring campaign. The sensor at the RM3 outlet flatlined at 0 mg/L from February onwards and was not possible to fix the issue by calibrating the sensor. The RM2 outlet sensor failed the calibration procedure in April and the data collected was therefore considered as not precise. The combined inlet sensor also stopped measuring and the signal was flatlining. This was due to data transmission and calibration issues. Therefore, as can be seen in Fig. 4, the data for these periods was discounted.

The average H_2S concentrations for the RM1, RM2 and RM3 outlets and for the combined inlet of the WWTP are shown in Fig. 4. The corrosion potential of the rising mains was assessed based on the dissolved H_2S load, calculated as the average H_2S concentration times the dry weather flow (DWF). As H_2S usually shows seasonality, the assessment was divided to warm months (April – September) and cold months (November – January). In the cold months, the main source of corrosion potential was RM1, having a 3.6 kg/d H_2S load, 21.6 % higher than RM3, which had a 2.96 kg/d load and 266.8 % higher than RM2, which had a 0.98 kg/d load. During the warm months, it was only possible to calculate the load of RM1, which was of 4.2 kg/d and 16 % higher than in cold months. The slight increase on RM1 H_2S from cold months to warm months suggests the nitrate dosing was definitely having an effect. Even more when considering that July and August, the warmest months, with 21.05 °C and 24.77 °C temperature respectively, had lower concentrations than the cold months. However, the mechanism of the dissolved H_2S reduction could not be defined, as it was not possible to assess if the total dissolved sulphide was reduced or the pH of the wastewater was increased, as it has been proved that nitrate dosing results in higher wastewater pHs [39]. The increase in pH would justify the lower H_2S concentrations, as the dissolved sulphide speciates in different species depending on the pH of the wastewater (Eq. (1)) [40].

Similarly, the mass balances to assess the highest sources of sulphide to the WWTP were not possible to calculate. The pH of each rising main outlet would probably have been different and therefore, the total dissolved sulphide is unknown. When using dissolved sulphide sensors, which usually measure either the dissolved H_2S or the dissolved bisulphide, a pH probe is also required if the total dissolved sulphide is to be calculated. This has also been recently highlighted by [16] when comparing different dissolved sulphide sensors.

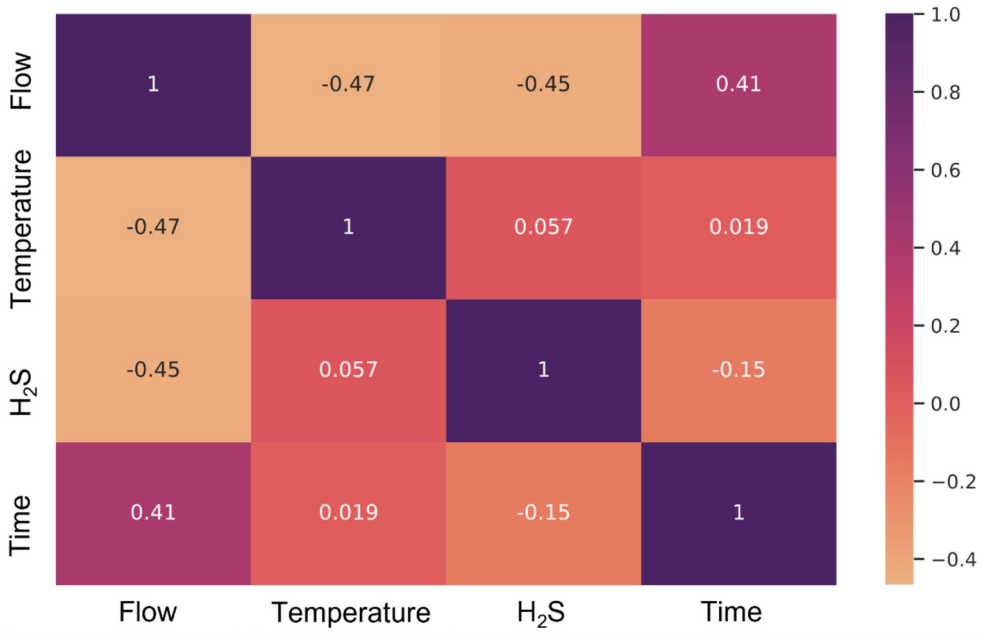


Fig. 5. Correlation matrix of the LSTM model variables. Flow in m^3/h; temperature in °C; time in day and H_2S in mg/L.

Table 2
Descriptive statistics of the LSTM model variables.

Parameter	Units	Average	STDEV	Min	Max
Flow	m^3/h	514.29	204.01	0	936
Temperature	°C	13.51	6.24	0	29.75
Time of the day	Day	0.5	0.29	0	1
H_2S	mg/L	0.65	0.53	0	5

STDEV - standard deviation.

3.2. LSTM model

As a proof of concept and to repurpose the SulfiLogger sensors to another location, a dissolved H_2S prediction model was developed for the RM1 outlet. RM1 was selected, as it had continuous data for 9 months covering cold and warm seasons. Instead of using a mechanistic model, such as the WATS or SEWEX, it was decided to follow a data-driven approach using the LSTM model. The decision was taken as the LSTM model required the monitoring and availability of significantly less parameters and was expected to achieve high accuracy.

As a first step, the input parameters selected, and the dissolved H_2S were analysed by using a correlation matrix (Fig. 5) The parameters were shown not to have strong correlations, all being below the 0.5 correlation parameter. This lack of correlation led to the development of a more complex machine learning model for better predictions in this study. The parameters used for the model development had values between the ones shown in Table 2. The RM1 flowrate ranged from 0 to 936 m^3/h, the temperature ranged from 0 to 29.75 °C, the time variable ranged from 0 to 1 days and the dissolved H_2S ranged from 0 to 5 mg/L, the 5 mg/L being the upper limit of the SulfiLogger measurement range. As the model is built on the training data, it can only be applied within the limits the model development data covers. This is one of the limitations of the model. When H_2S levels exceed the threshold of 5 mg/L, the sensor cannot provide accurate readings, which may lead to inaccuracies in the model's predictions during high-H_2S events. Combining the LSTM model with a mechanistic model can improve the handling of high-H_2S situations. Mechanistic models like WATS or SEWEX can simulate sulphide formation based on known sewer conditions and input parameters, providing estimates that can supplement the LSTM model's predictions during unmeasurable peaks. Nevertheless, if the aim of the model is focus on predicting and managing corrosion the sensor used can still be of use, as according to Hvitved-Jacobsen [2] a H_2S concentration >2 mg/L can be considered as potential to cause severe corrosion and therefore having accuracy of the concentrations >5 mg/L is not essential when assessing corrosion. Table 2 can serve as a framework for identifying future directions in applying ML models for wastewater treatment predictions.

The LSTM model was developed as described in Section 2.6, and the predicted against the measured dissolved H_2S can be observed in Fig. 6 for 4 days of the test set. As seen in the Fig. 6, the predicted H_2S by the LSTM model follows the measured curve quite well, with the three characteristic daily variations. However, some of the high peaks are underestimated by the LSTM model and suggest these cannot be well explained by the input parameters.

Also in Fig. 6, it can be seen that the polynomial regression follows very closely the prediction of the dissolved H_2S by the LSTM, whereas the linear regression has a worse performance. This is further confirmed by the model performance metrics in Table 3. The RMSE for the LSTM in the test set is the lowest at 0.34 and the NSE is the highest at 0.57, which indicates the LSTM model is the best performing model and achieves the overall closest predicted value to the measured dissolved H_2S (Table 3). The polynomial regression follows closely with a RMSE of 0.35 and a NSE of 0.54. On the other hand, the linear regression performs significantly worse with a RMSE of 0.46 and a NSE of 0.20. Adding additional LSTM layers to the LSTM model was trialled, up to 3 LSTM layers; however, although the train set RMSE was improved down to 0.30 compared to the 0.38 achieved by the final LSTM model, the test set RMSE was higher at 0.40 when using 3 LSTM layers. The phenomenon observed when increasing the LSTM layers is known as over-fitting and it is a common observation when developing data-driven models [31]. STM generally achieves a higher RMSE in sequential prediction tasks like H_2S concentration in wastewater treatment due to its ability to capture both short- and long-term dependencies [21]. This makes it preferable over simpler models (like linear regression and SVR) and even other DL architectures (such as RNN or CNN-LSTM) for this

Table 3
Model performance assessment for H_2S prediction at the outlet of RM1 using linear regression, polynomial regression and LSTM model.

Model	RMSE		NSE	
	Train	Test	Train	Test
LSTM	0.382	0.340	0.473	0.571
Linear	0.467	0.465	0.212	0.199
Polynomial	0.379	0.350	0.482	0.544

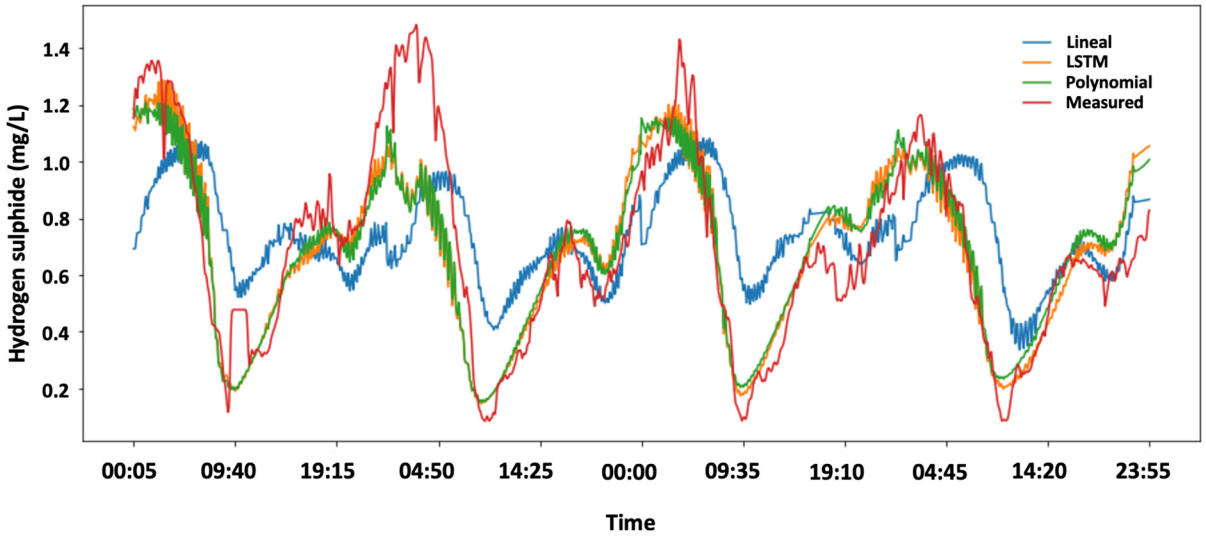


Fig. 6. Hydrogen sulphide (H_2S) concentration at the outlet of RM1 and the predicted concentrations using linear regression, polynomial regression and LSTM models.

Table 4
Summary of recent studies use of LSTM and similar models in WWTPs.

Model type	Application	Key findings	Further considerations	Ref.
LSTM	forecast wastewater treatment performance	Soft sensor outperformed multiple linear regression (MLR) models in forecasting key wastewater treatment parameters like COD, total nitrogen, NH_4^+, providing more accurate predictions based on lower-cost sensor data	Improve model scalability and integrate additional operational variables, such as aeration rates, to enhance prediction accuracy and generalize the model for real-world WWTP	[42]
ANN	Sulfide prediction in sewer systems	Showed ANN's ability to predict H_2S concentration with moderate accuracy	Insufficient handling of time-series dependencies; lack of real-time adaptability	[18]
NARX Neural Network	Emission predictions in WWTPs	Successfully predicted gaseous emissions	Validate in diverse scenarios, and improving predictive accuracy across various operational conditions.	[43]
LSTM	H_2S concentration in wastewater treatment	High R^2 achieved for real-time H_2S prediction using sequential dependencies	Further testing with additional influent variables (e.g., pH, nitrate levels) for enhanced accuracy	Proposed approach (this study)

application [41].

When comparing the model performance to other uses of the LSTM model, such as the one to predict flowrate to a WWTP developed by [21], achieved a much better performance with RMSE and NSE values of 0.04 and 0.94, respectively, for a 15 min interval data. The performance difference is expected to be related to the complexity of the system being modelled. The flowrate to a WWTP relies on few key factors that can be included in the prediction model. On the other hand, the H_2S formation is dependent on many variables, such as the rising main physical properties, the wastewater composition and the hydraulic conditions, and in this specific case, the nitrate dose could also have been an important factor not directly considered in the model development. Overall, the LSTM model performance was deemed acceptable for the potential uses of the predictions. The novelty of this study is development of an AI based model to accurately predict the daily H_2S pattern. This study's main contributions include insights into sulphide dynamics from high-resolution sensor data, which could support corrosion management as part of a septicity warning system or feedforward control for sulphide treatment. Other recent studies have documented the use of LSTM and similar models in WWTPs (Table 4). A ANN model was developed for sulphide prediction in sewer systems with moderate accuracy, but presented limitations on handling of time-series dependencies [18] (Table 4).

Other models have been successfully used for forecasting wastewater treatment performance and predict process emissions, as some examples of AI potential and growing field (Table 4).

Two main uses are suggested for the dissolved H_2S prediction in this study. The first one would be using the predicted dissolved H_2S as an input to a septicity warning system, such as the one described in [44]. The use of the warning system would be dependent on the specific WWTP, but more information is available in the paper itself. The second use would allow for a feedforward control of a sulphide treatment chemical, such as nitrate [45] or iron [46], at the prediction point, which has been shown in Australia to be more effective than dosing at the wet well of a rising main [47]. Additionally, the AI-based prediction model offers potential for sensor repurposing, saving both capital and operational costs.

Finally, the dissolved H_2S monitoring is required to be of a minimum of 8–9 months, and has to capture the warmest and coldest months of the year, as it is key to cover all the range of temperatures in the training set. Otherwise, the LSTM model would not be applicable on the extreme temperature months. As future work development of a nitrate dose controller based on a machine learning algorithm, using the dissolved sulphide measurements at the outlet of the rising main and trialling it at full-scale to measure the cost-benefits of the application of an advanced control mechanism.

4. Conclusions

In this study, the dissolved H_2S at 3 rising main outlets discharging at a WWTP inlet chamber were monitored using a novel dissolved sulphide sensor (SulfiLogger) for 9 months. This study successfully developed and validated LSTM model for H_2S concentration prediction for a rising main outlet using wastewater flowrate, temperature and the time of the day as input parameters, addressing critical concerns around concrete corrosion due to sulphide accumulation and has the potential to be applied in sulphide treatment chemical control systems. In contrast to previous models that rely on periodic sampling, this research contributes to knowledge by offering a robust, real-time perspective on sulphide dynamics. The model has limitations, including the detection limit of the sensors used, the narrow parameter scope that excludes important factors like pH and organic matter concentration, which could improve accuracy. The model was developed in a single wastewater treatment plant, offering potential for broader application with further validation across diverse facilities. While sensor maintenance and computational demands may pose challenges for smaller settings, these can be addressed with optimization. Expanding long-term validation will help ensure robust performance under varying conditions and operational changes. This study advances wastewater management by combining continuous high-resolution sensor data with a long short-term memory model to predict H_2S levels, which enhances infrastructure resilience. Additionally, the potential for sensor repurposing makes advanced corrosion management more accessible, underscoring the impact of machine learning on sewer infrastructure and WWTPs.

CRediT authorship contribution statement

J. Mendizabal: Writing – original draft, Methodology, Formal analysis, Data curation, Conceptualization. **D. Vernon:** Writing – review & editing, Methodology, Formal analysis, Data curation, Conceptualization. **B. Martin:** Writing – review & editing, Funding acquisition, Conceptualization. **Y. Bajón-Fernández:** Writing – review & editing, Supervision, Conceptualization. **A. Soares:** Writing – review & editing, Validation, Supervision, Project administration, Funding acquisition, Formal analysis, Data curation, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors gratefully acknowledge financial support from the Engineering and Physical Sciences Research Council (EPSRC) [grant number EP/R512515/1] through their funding of the STREAM Industrial Doctorate Centre, and the Industry project sponsor Thames Water. Data underlying this paper can be accessed at https://doi.org/10.57996/cran.ceres-2689.

Data availability

Data available on DOI.

References

[1] O. Lahav, A. Sagiv, E. Friedler, A different approach for predicting H_2S(g) emission rates in gravity sewers, Water Res. 40 (2) (2006) 259–266, https://doi.org/10.1016/j.watres.2005.10.026.

[2] T. Hvitved-Jacobsen, J. Vollertsen, A.H. Nielsen, Sewer Processes - Microbial and Chemical Process Engineering of Sewer Networks, 2nd ed., Taylor & Francis Group, 2013.

[3] C. Grengg, F. Mittermayr, N. Ukrainczyk, G. Koraimann, S. Kienesberger, M. Dietzel, Advances in concrete materials for sewer systems affected by microbial induced concrete corrosion: a review, Water Res. 134 (2018) 341–352, https://doi.org/10.1016/j.watres.2018.01.043.

[4] E. Vincke, Biogenic Sulfuric Acid Corrosion of Concrete: Microbial Interaction, Simulation and Prevention (Doctoral Dissertation), Universitei Gent, Gent, Belgium, 2008.

[5] J. Sun, S. Hu, K.R. Sharma, H. Bustamante, Z. Yuan, Impact of reduced water consumption on sulfide and methane production in rising main sewers, J. Environ. Manag. 154 (2015) 307–315, https://doi.org/10.1016/j.jenvman.2015.02.041.

[6] L. Cintra Campos, G. Darch, Adaptation of UK wastewater infrastructure to climate change, Infrastructure Asset Management 2 (3) (2015) 97–106, https://doi.org/10.1680/iasma.14.00037.

[7] J. Vollertsen, N. Revilla, T. Hvitved-Jacobsen, A.H. Nielsen, Modeling sulfides, pH and hydrogen sulfide gas in the sewers of San Francisco, Water Environ. Res. 87 (12) (2015) 1980–1989, https://doi.org/10.2175/106143015X14362865226752.

[8] K.R. Sharma, Z. Yuan, D. de Haas, G. Hamilton, S. Corrie, J. Keller, Dynamics and dynamic modelling of H_2S production in sewer systems, Water Res. 42 (10−11) (2008) 2527–2538, https://doi.org/10.1016/j.watres.2008.02.013.

[9] L. Sutherland-Stacey, S. Corrie, A. Neethling, I. Johnson, O. Gutierrez, R. Dexter, Z. Yuan, J. Keller, G. Hamilton, Continuous measurement of dissolved sulfide in sewer systems, Water Sci. Technol. 57 (3) (2008) 375–381, https://doi.org/10.2166/wst.2008.132.

[10] M. Pacheco Fernandez, D. Despot, M. Barjenbruch, Comparison of H_2S gas sensors: a sensor management procedure for sewer monitoring, Sustainability 13 (19) (2021) 1–16, https://doi.org/10.3390/su131910779.

[11] F.S. Abu-Hani, Y.E. Greish, S.T. Mahmoud, F. Awwad, A.I. Ayesh, Low-temperature and fast response H_2S gas sensor using semiconducting chitosan film, Sensors Actuators B Chem. 253 (2017) 677–684, https://doi.org/10.1016/j.snb.2017.06.169.

[12] K. Crowley, Fabrication of polyaniline-based gas sensors using piezoelectric inkjet and screen printing for the detection of hydrogen sulfide, IEEE Sensors J. 10 (9) (2010) 1419–1426, https://doi.org/10.1109/JSEN.2010.2046101.

[13] M.M. Montazeri, N. De Vries, A.D. Afantchao, A. O'Brien, P. Kadota, M. Hoorfar, Development of a sensing platform for nuisance sewer gas monitoring: hydrogen sulfide detection in aqueous versus gaseous samples, IEEE Sensors J. 18 (19) (2018) 7772–7778, https://doi.org/10.1109/JSEN.2018.2867522.

[14] A. Yavarinasab, S. Janfaza, M.M. Montazeri, N. Tasnim, A.D. Farahani, P. Kadota, P. Markin, A. Dalili, E. Taatizadeh, H. Tahmooressi, M. Hoorfar, A graphene-based chemical sensor for hydrogen sulphide measurement in water, in: IEEE Sensors Conference, 2020, https://doi.org/10.1109/SENSORS43011.2019.8956803.

[15] SulfiLogger, SulfiLogger H_2S sensor. https://sulfilogger.com/sensor-en/, 2020. (Accessed 28 November 2022).

[16] D. Despot, M. Pacheco Fernandez, M. Barjenbruch, Comparison of online sensors for liquid phase hydrogen sulphide monitoring in sewer systems, Water 13 (13) (2021) 1876, https://doi.org/10.3390/w13131876.

[17] A. El Brahmi, S. Abderafi, Artificial neural network analysis of sulfide production in a Moroccan sewerage network, Indian J. Sci. Technol. 6 (1) (2021) 193–204, https://doi.org/10.17509/ijost.v6i1.32322.

[18] A. El Brahmi, S. Abderafi, Hydrogen sulfide production assessment based on sewage physicochemical properties using artificial neural network, Materials Today: Proceedings 27 (2020) 3028–3032, https://doi.org/10.1016/j.matpr.2020.03.504.

[19] M. Zounemat-Kermani, D. Stephan, R. Hinkelmann, Multivariate NARX neural network in prediction gaseous emissions within the influent chamber of wastewater treatment plants, Atmos. Pollut. Res. 10 (6) (2019) 1812–1822, https://doi.org/10.1016/j.apr.2019.07.013.

[20] X. Ma, Z. Tao, Y. Wang, H. Yu, Y. Wang, Long short-term memory neural network for traffic speed prediction using remote microwave sensor data, Transportation Research Part C: Emerging Technologies 54 (2015) 187–197, https://doi.org/10.1016/j.trc.2015.03.014.

[21] D. Zhang, E.S. Holland, G. Lindholm, H. Ratnaweera, Hydraulic modeling and deep learning based flow forecasting for optimizing inter catchment wastewater transfer, J. Hydrol. 567 (2018) 792–802, https://doi.org/10.1016/j.jhydrol.2017.11.029.

[22] J. Wang, L. Li, J. Xia, W. Chen, Applications of machine learning techniques in water quality modeling and management: a review, Water Res. 188 (2021) 116490, https://doi.org/10.1016/j.watres.2020.116490.

[23] G. Fu, Y. Jin, S. Sun, Z. Yuan, D. Butler, The role of deep learning in urban water management: a critical review, Water Res. 223 (2022) 118973, https://doi.org/10.1016/j.watres.2022.118973.

[24] L. Xia, X. Hao, Y. Zhou, Machine learning-based global trends and the development prospects of wastewater treatment: a bibliometric analysis, Journal of Environmental Chemical Engineering 12 (3) (2024), https://doi.org/10.1016/j.jece.2024.112732.

[25] Y. Zhang, C. Li, H. Duan, K. Yan, J. Wang, W. Wang, Deep learning-based data-driven model for detecting time-delay water quality indicators of wastewater treatment plant influent, Chem. Eng. J. 467 (2023) 143483, https://doi.org/10.1016/j.cej.2023.143483.

[26] H. Sadoune, R. Rihani, F.S. Marra, DNN model development of biogas production from an anaerobic wastewater treatment plant using Bayesian hyperparameter optimization, Chem. Eng. J. 471 (2023) 144671, https://doi.org/10.1016/j.cej.2023.144671.

[27] J. Xue, Y. Li, S. Liu, Y. Bai, Q. Jiang, Y. Gao, Y. Yao, K. Shi, J. Xu, The important role of high sludge concentration in anaerobic biological systems with low COD/SO_4^{2-}-sulfate-containing wastewater predicted by machine learning: Insights from microbial community and metabolic pathways, Chem. Eng. J. 496 (2024) 154320, https://doi.org/10.1016/j.cej.2023.154320.

[28] A. Soo, L. Gao, H.K. Shon, Machine learning framework for wastewater circular economy—towards smarter nutrient recoveries, Desalination 592 (2024) 118092, https://doi.org/10.1016/j.desal.2024.118092.

[29] M.S. Duarte, G. Martins, P. Oliveira, B. Fernandes, E.C. Ferreira, M.M. Alves, F. Lopes, M.A. Pereira, P. Novais, A review of computational modeling in wastewater treatment processes, ACS ES&T Water 4 (3) (2023) 784–804, https://doi.org/10.1021/acsestwater.3c00117.

[30] C. Yongsiri, J. Vollertsen, M. Rasmussen, T. Hvitved-Jacobsen, Air-water transfer of hydrogen sulfide: an approach for application in sewer networks, Water Environ. Res. 76 (1) (2004) 81–88, https://doi.org/10.2175/106143004x141618.

[31] APHA, Standard Methods for the Examination of Water and Wastewater, 21st ed., American Public Health Association, Washington, DC, 2005.

[32] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, R. Salakhutdinov, Dropout: a simple way to prevent neural networks from overfitting, J. Mach. Learn. Res. 15 (1) (2014) 1929–1958.

[33] G. Tchobanoglous, H.D. Stensel, R. Tsuchihashi, F. Burton, M. Abu-Orf, G. Bowden, W. Pfrang, Wastewater Engineering: Treatment and Resource Recovery, 5th ed., McGraw-Hill Education, 2014.

[34] A.H. Nielsen, J. Vollertsen, H.S. Jensen, H.I. Madsen, T. Hvitved-Jacobsen, Aerobic and anaerobic transformations of sulfide in a sewer system - field study and model simulations, Water Environ. Res. 80 (1) (2008) 16–25, https://doi.org/10.2175/106143007x184537.

[35] D.K.B. Thistlethwayte, The Control of Sulphides in Sewerage Systems, Ann Arbor Science Publishers, 1972.

[36] R.D. Pomeroy, J.D. Parkhurst, The forecasting of sulfide buildup rates in sewers, Progress in Water Technology 9 (1977) 621–628, https://doi.org/10.1016/B978-0-08-020902-9.50083-3.

[37] E. Rudelle, J. Vollertsen, T. Hvitved-Jacobsen, A.H. Nielsen, Modeling anaerobic organic matter transformations in the wastewater phase of sewer networks, Water Sci. Technol. 66 (8) (2012) 1728–1734, https://doi.org/10.2166/wst.2012.378.

[38] G. Jiang, K.R. Sharma, Z. Yuan, Effects of nitrate dosing on methanogenic activity in a sulfide-producing sewer biofilm reactor, Water Res. 47 (5) (2013) 1783–1792, https://doi.org/10.1016/j.watres.2012.12.036.

[39] Z. Liang, D. Wu, G. Li, J. Sun, F. Jiang, Y. Li, Experimental and modelling investigations on the unexpected hydrogen sulphide rebound in a sewer receiving nitrate addition: mechanism and solution, J. Environ. Sci. 125 (2023) 630–640, https://doi.org/10.1016/j.jes.2021.12.038.

[40] C. Yongsiri, J. Vollertsen, M. Rasmussen, T. Hvitved-Jacobsen, Air-water transfer of hydrogen sulfide: an approach for application in sewer networks, Water Environ. Res. 76 (1) (2004) 81–88, https://doi.org/10.2175/106143004x141618.

[41] X. Shi, Z. Chen, H. Wang, D.Y. Yeung, W.K. Wong, W.C. Woo, Convolutional LSTM network: a machine learning approach for precipitation nowcasting, Adv. Neural Inf. Proces. Syst. 28 (2015).

[42] B. Xu, C.K. Pooi, K.M. Tan, S. Huang, X. Shi, H.Y. Ng, A novel long short-term memory artificial neural network (LSTM)-based soft-sensor to monitor and forecast wastewater treatment performance, Journal of Water Process Engineering 54 (2023) 104041, https://doi.org/10.1016/j.jwpe.2023.104041.

[43] N.B. Mihály, A.V. Luca, M. Simon-Várhelyi, V.M. Cristea, Improvement of air flowrate distribution in the nitrification reactor of the wastewater treatment plant by effluent quality, energy and greenhouse gas emissions optimization via artificial neural networks models, Journal of Water Process Engineering 54 (2023) 103935, https://doi.org/10.1016/j.jwpe.2023.103935.

[44] J. Mendizabal, D. Vernon, B. Martin, Y. Bajon-Fernandez, B. Jefferson, A. Soares, Development of a septicity scale for raw wastewater, Article in Preparation (2024).

[45] J. Mohanakrishnan, O. Gutierrez, K.R. Sharma, A. Guisasola, U. Werner, R. L. Meyer, J. Keller, Z. Yuan, Impact of nitrate addition on biofilm properties and activities in rising main sewers, Water Res. 43 (17) (2009) 4225–4237, https://doi.org/10.1016/j.watres.2009.06.021.

[46] B. Kiilerich, W. van de Ven, A.H. Nielsen, J. Vollertsen, Sulfide precipitation in wastewater at short timescales, Water 9 (9) (2017) 670–682, https://doi.org/10.3390/w9090670.

[47] R. Ganigue, O. Gutierrez, R. Rootsey, Z. Yuan, Chemical dosing for sulfide control in Australia: an industry survey, Water Res. 45 (20) (2011) 6564–6574, https://doi.org/10.1016/j.watres.2011.09.054.



https://dspace.org/

Short-term memory artificial neural network modelling to predict concrete corrosion in wastewater treatment plant inlet chambers using sulphide sensors

Mendizabal, J.; Vernon, Dejan; Martin, Ben; Bajón Fernández, Yadira; Soares, Ana

2025-01
Elsevier
https://doi.org/10.1016/j.jwpe.2024.106821
Mendizabal J, Vernon D, Martin B, et al., (2025) Short-term memory artificial neural network modelling to predict concrete corrosion in wastewater treatment plant inlet chambers using sulphide sensors. Journal of Water Process Engineering, Volume 69, January 2025, Article number 106821

Downloaded from a DSpace repository. https://dspace.org/