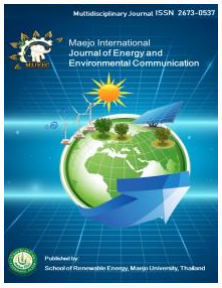
Maejo International Journal of Energy and Environmental Communication

Journal homepage: https://ph02.tci-thaijo.org/index.php/MIJEEC



ARTICLE

Prediction of MBR operating parameter using LSTM neural network

Vu Van Huynh[1], Minh Binh Nguyen[1], Tetsuro Ueyama[2], Satoshi Shirayanagi[3], Tetsuo Imai[4], Saenchan Somsri[1], Tomoaki Itayama[1,*]

[1]Graduate School of Engineering, Nagasaki University, 1-14 Bunkyo-machi, Nagasaki 852-8521, Japan
[2]R&D Division, Kyowakiden Industry Co., Ltd., 10-2 Kawaguchi-machi, Nagasaki, 852-8108, Japan
[3]F.C.C Co., Ltd., 7000-36 Nakagawa Hosoe-cho Kita-ku Hamamatsu Shizuoka 431-1394, Japan
[4]Faculty of Information Sciences, Hiroshima City University, 3-4-1, Ozuka-Higashi, Asaminami-ku, Hiroshima, 731-3194, Japan

ARTICLEINFO

Article history:
Received 29 September 2023
Received in revised form
21 October 2023
Accepted 29 October 2023

Keywords:
Wastewater treatment
MBR
Time series
Neural network
LSTM

ABSTRACT

This study investigated the forecasting ability of the long short-term memory neural network model (LSTM model), which is a type of recurrent neural network (RNN), for the dynamic character of membrane bioreactor (MBR). MBR is an advanced wastewater treatment system that combines activated sludge process with a membrane separation system. In this study, dissolved oxygen (DO), pH, trans membrane pressure (TMP), mixed liquor suspended solids (MLSS), and air flow rate of a bench-scale MBR were measured to obtain the time series data, and the time interval for each time series was unified to 1 hour. The training period of 640 hours was adopted for the LSTM model, and the remaining 160 hours were used as the testing period. The trained LSTM model predicted DO, pH, TMP, and MLSS one step ahead (one hour ahead), and multiple steps forecasts up to 6 hours ahead were also tested. The LSTM model succeeded in predicting MLSS one hour ahead with high accuracy. On the other hand, for DO and pH, the values predicted one hour ahead by the LSTM model reproduced their temporal fluctuation patterns to some extent. However, all of them tended to show predicted values that were lower than the actual values. The predicted values from the LSTM model did not reproduce the pattern of TMP changes well. In addition, the LSTM model was investigated the effect of forecasting horizons and look back period.

1. Introduction

Membrane bioreactors (MBRs) are wastewater treatment systems that combine membrane separation with activated sludge processes. They have the advantage of producing high-quality effluent with a compact system and are widely applied for both domestic and industrial wastewater treatment (Judd, 2016; Li et al., 2019). However, the MBRs still have some on-going challenges, including high energy consumption and membrane fouling. In general, aeration is the main source of energy consumption in MBRs, accounting for approximately 50% of the operating cost (Fenu et al., 2010; Wang et al., 2020). Therefore, minimizing aeration energy is a major challenge not only for MBRs, but also for conventional activated sludge systems, as it can

* Corresponding author.
E-mail address: itayama@nagasaki-u.ac.jp

reduce both operating costs and environmental impacts (Asadi et al., 2017; Gu et al., 2023).

Furthermore, as with other membrane separation process technologies, membrane fouling is also a significant issue that hinders the application of MBR system (Xiao et al., 2019). Since aeration has a cleaning effect on the membrane, reducing aeration may increase filtration pressure due to membrane fouling, which may increase operation cost (Du et al., 2020; Tang et al., 2022). Therefore, it is important for MBR technology to operate at the minimum amount of aeration that can avoid membrane fouling. Namely, an optimum control of the MBR system is one of the effective solutions that should be considered to balance between maintaining sufficient aeration and properly reducing energy consumption and prolong the membrane lifetime. Since the activated sludge system is principally complex nonlinear dynamical system consisting with tremendous number of microbes and their complex interactions, achieving optimal control is a challenge (Kaewpipat & Grady, 2002).

The operation of wastewater treatment systems in centralized and large-scale wastewater treatment plants (WWTPs) requires experienced engineers or technical staff, especially in developing countries. However, such human resources are often scarce or unavailable for decentralized and/or small-scale WWTPs. A possible solution is to implement a model based automatic predictive control system, which can adjust the system parameters according to the estimated values of operating variables (Chen et al., 2022). This approach can enhance the system performance and stability, as well as provide early warning for potential disturbances or anomalies (Zhang et al., 2018). Nevertheless, mechanistic models, which are commonly used for predictive control, involve many specific and sensitive parameters that are difficult and costly to identify and measure. Therefore, these models are not suitable for practical applications in real-time control systems.

Deep learning techniques using artificial neural networks (ANNs) show great potential in automatic learning and parameters predicting of dynamic systems (Dufera et al., 2022). An ANN consists of an input layer and an output layer, and the single or multiple hidden layers between the input and output layers are composed of artificial neuron units with nonlinear responses interconnected between each layer and the input and output layers (da Silva et al., 2017). Each interconnection between neurons has its own weight, which is updated sequentially through ANN training, and gradually learns the target image or time series. Especially, an ANNs that consist of many hidden layers and has an appropriate learning algorithm are generally called a deep learning ANNs. It is well known that deep learning ANNs have already been applied in many industrial fields. Research on wastewater treatment employing activated sludge, using ANNs is increasing (Schmitt et al., 2018).

For the goal of parameters prediction, multi-dimensional sensing data of MBR system will be used as input information for deep learning. Based on the characteristic of input information is in time-series form of data, the recurrent neural networks (RNNs), a substitute of ANNs, is a compatible technique for dynamic learning and prediction (Hewamalage et al., 2021). RNN is a sort of deep learning model, which uses previous experiments to predict upcoming events. However, during the operations, the RNN model has some disadvantages such as vanishing the gradient problems (leading the prediction to huge errors), storing information in short-term memories (ArunKumar et al., 2021; Bas et al., 2022). These issues are significant challenges for the RNN model in proceeding long sequential time-series dataset. Thus, LSTM, a sub-class of RNN, is a potential solution in dealing with long sequential sensing data (Karnam et al., 2022). Different from the basic RNN model, LSTM-RNN model has the capability of evaluating the importance of each information (Hewamalage et al., 2021). Thus, the model can store essential information and remove irrelevant information, which saves the model memory for operation of long-term prediction (Huang et al., 2022). LSTM-RNN has been applied to WWTP operation, such as effluent quality prediction, process control, and fault detection (Pisa et al., 2019; Toffanin et al., 2023; Wongburi & Park, 2023; Yoon et al., 2021; Zhong et al., 2022). However, the application of LSTM to MBR systems has not been extensively explored, to the best of our knowledge. In this study, we applied LSTM-RNN to multidimensional time-series data from various sensors installed on a bench-scale MBR and aimed to clarify its predictive ability for these time-series data.

2. Material and methods

2.1 Experimental set up of bench scale MBR system

This study employed data from a bench scale MBR that treated synthetic wastewater. The schematic diagram of the system is shown in Figure 1 The MBR system consisted of a reactor with an effective volume of 9.5 L and a flat-sheet membrane module (F.C.C. Co., Ltd, Japan) submerged inside the reactor. The membrane has a total surface area of 0.115 m^2 and a pore size of 0.08 μm. The reactor was aerated by an air diffuser located at the bottom of the reactor. The seed sludge for this bench scale MBR was collected from a pilot scale MBR system in Nagasaki University. The influent synthetic wastewater was formulated based on typical municipal wastewater characteristics to contain approximately 170 mg/L total organic carbon and 50 mg/L total nitrogen.

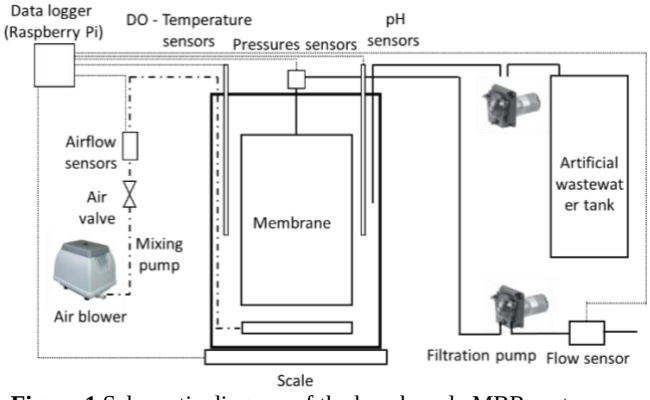


Figure 1 Schematic diagram of the bench scale MBR system

The MBR system was equipped with sensors to monitor the trans membrane pressure (TMP), dissolved oxygen (DO), pH, and airflow rate. The sensors were digital pressure gauges, DO sensors (FOM 1400, Automatic system Research Co., Ltd, Japan), pH sensors (PE -22, CEM Corporation, Japan), and airflow sensors (FSM2 series, CKD Corporation, Japan). The temperature data used was the value measured by the internal temperature sensor of the DO sensor. A Raspberry Pi 3 Model B+ (Raspberry Pi Foundation, UK) with an ADC Pi (AB Electronics, UK) acquired data from the sensor at 10 second intervals, while mixed liquor suspended solids (MLSS) data was

measured manually using a handy sensor (SS-10Z, Kasahara Chemical Instruments Corp, Japan) every 2-3 days.

In this study, the membrane flux was 14.89 LMH. The suction cycle of 8 min on and 2 min off was optimal for reducing membrane fouling and maintaining steady permeate flow based on preliminary experiments. The hydraulic retention time and solids retention time were 7 hours and 15 days, respectively. The air flow ranged from 8.36 to 10.34 L/min. The system was operated at a stable temperature of 25°C.

2.2 Data preparation

This study collected data from 800 hours of MBR operation, during which five parameters were monitored: DO, MLSS, TMP, pH, and airflow rate. The data collection frequency varied for different parameters. These parameters are related to the biological and physical processes that occur in the MBR system, such as oxygen transfer, membrane fouling, microbial activity, and sludge concentration. The MLSS data were obtained every 2-3 days, while the other parameters such as DO were measured every 10 seconds by the sensors installed at the MBR. To obtain a consistent and complete dataset, the MLSS data obtained were interpolated by third order polynomial to match the hourly time interval of the other parameters. The other sensor time series values were aggregated by taking the average value for each hour. Since each parameter had a different range of values, the Min-Max scaling method was applied to normalize the data based on the minimum and maximum values of each parameter. This method scales the data to a range of 0 to 1, where 0 represents the minimum value and 1 represents the maximum value. This method has been widely used in many studies on LSTM (Kuan et al., 2017; Qiao et al., 2021).

Then the dataset was then divided into target variables and explanatory variables for the LSTM model. The target variables were the four parameters that were to be predicted by the model: DO, TMP, pH, and MLSS. The explanatory variables were the airflow rate and the targets variables. Due to the operation under the constant temperature, we didn't use temperature data in this experiment. The dataset was also split into two subsets: a training subset (80% of the total data length, 640 hours of MBR operation) and a testing subset (20% of the total data length, 160 hours of MBR operation). The training subset was used to train the LSTM model, while the testing subset was used to evaluate its performance.

2.3 Long short-term memory model

Recurrent neural networks (RNNs) are a class of neural networks that allow previous outputs to be used as inputs while having hidden states. The structure of RNN is shown in Figure 2. For each timestep t, the input data x_t, the activation a_t and the output y_t are expressed as below:

$$a_t = g_1(W_{aa}a_{t-1} + W_{ax}x_t + b_a) \ldots . (1)$$

$$y_t = g_2(W_{ya}a_t + b_y) \ldots . (2)$$

where $W_{aa}, W_{ax}, W_{ya}, b_a, b_y$ are coefficients that are shared temporally and g_1, g_2 activation functions. In the training phase, the output y_t is comared to the target time series value. When simple one-dimensional time series data $\{ x_1, x_2, \ldots . x_t\}$ is trained for one-step prediction, the output y_t is compared to the next value x_{t+1}. That is, the weights W in the RNN is updated during the training process so that the square of this difference y_t-x_{t+1}, as an error function, is minimized.

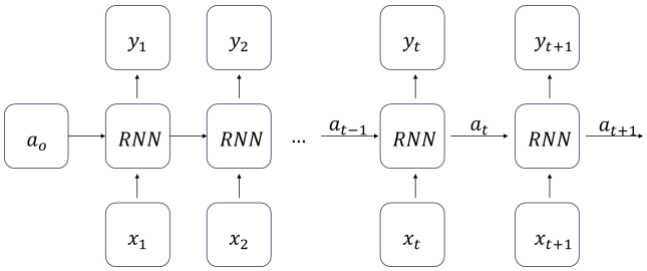


Figure 2 A typical structure of RNN for time series data

RNNs can learn from both long-term and short-term dependencies in the data and can adapt to changing patterns over time. However, during the operations, the RNN model has some disadvantages (ArunKumar et al., 2021; Bas et al., 2022). As shown in Figure 2, in an RNN with a single hidden layer, the hidden layer increases over time, corresponding to a neural network with a very large number of layers. Therefore, a vanishing gradient problem arises that leads to large errors in prediction. These issues are significant challenges for the RNN model in proceeding long sequential time-series dataset.

LSTM stands for Long Short-Term Memory, which is a type of RNN that can learn long-term dependencies in sequential data (Karnam et al., 2022). Through the additional memory unit, the long-term timing information is stored to capture the long-term dependencies in the data.

As shown in Figure 3, the memory unit of the LSTM neural network has three gates at each time step: the forget gate, the input gate, and the output gate. These gates enable the LSTM neural network to perform filtering and information storage functions. The forget gate decides which information to discard from the previous cell state, the input gate decides which information to add to the current cell state, and the output gate decides which information to output from the current cell state. These gates are composed of sigmoid activation functions and element-wise multiplication operations.

The forget gate combines the current input x_t with the previous hidden state h_{t-1}. All the input features are scaled through an activation function, and the scaling value is used to control the degree of forgetting of the previous cell state C_{t-1}. The equation of the forget gate is as follows:

$$f_t = \sigma(W_f . h_{t-1} + V_f . x_t + b_f) \ldots . (3)$$

At the input gate, new candidate values $\widetilde{C_t}$ are generated from the combination of x_t and h_{t-1} via a tanh activation function. Then, like the forget gate, the scaled value controls the degree to which the candidate values are updated. The equation for the input gate is as follows:

$$\widetilde{C_t} = tanh(W_C . h_{t-1} + V_C . x_t + b_C) \ldots . (4)$$

$$i_t = \sigma(W_i . h_{t-1} + V_i . x_t + b_i) \ldots . (5)$$

After that, the current cell state C_t is updated by the following equation:

$$C_t = i_t \circ \widetilde{C_t} + f_t \circ C_{t-1} . \ldots . (6)$$

The output gate determines what part of the information is output for C_t. The equation of the output gate is as follows:

$$o_t = \sigma(W_o . h_{t-1} + V_o . x_t + b_o) . \ldots . (7)$$

$$h_t = o_t \circ \tanh(C_t) \ldots . (8)$$

$$y_t = h_t \ldots . (9)$$

, where the element-wise product symbol $\circ$ is used.

In the formulas above, x_t and y_t is the input and output of LSTM model, respectively. i_t, o_t and f_t denote the input, output, and forget gate vector. h_t is a vector which denotes the hidden state of the cell while C_t is the cell state. At time step t, the $\widetilde{C_t}$ represents the cell state that is being determined to contain the important information or not, which will decide it can go to future processes or to be removed from the model memory. V_i, b_i, V_o, b_o, V_f, b_f and V_C, b_C are the weight matrix and bias vector corresponding to the current input, output, forget gate and cell state, respectively. W_i, W_o, W_f and W_C denote the weight matrices of the input gate, output gate, forget gate and cell state, respectively.

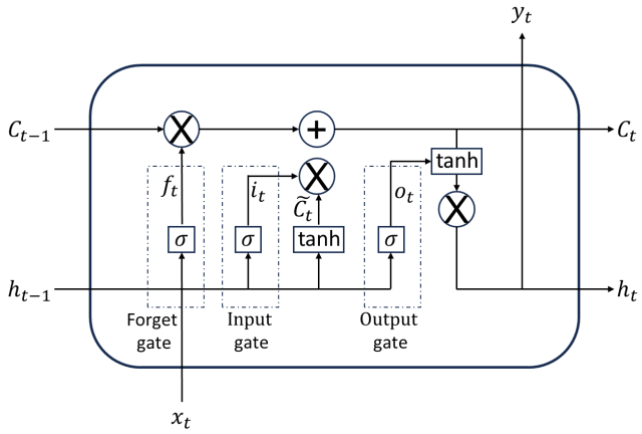


Figure 3 Illustration of a LSTM cell

The term "weight" in an LSTM model refers to the parameters that are learned during the training process. These parameters are used to transform the input data and hidden state into new values that are used to make predictions. In a typical LSTM model, there are three types of weights: input weights, recurrent weights, and bias weights. The input weights are used to transform the input data, while the recurrent weights are used to transform the hidden state from the previous time step. The bias weights are added to the transformed input and hidden state to produce the final output. The dimensions of these weight matrices depend on the number of input features, the number of hidden units, and the number of output features. The exact interpretation of these matrices can be complex, but they can be thought of as a set of linear transformations that map the input data and hidden state to new values. The process of fitting input data with known output data is called model training process. During this period, the model adjusts the weights to obtain minimum loss (the difference between predicted values and observed values in the known output data).

The performance and efficiency of an LSTM model depend largely on the choice of layer structures and hyperparameters, which are the parameters that are not learned by the model but are specified before the training process (Greff et al., 2017). In this study, the LSTM model adopted four layers structure (an input layer, two hidden layers, and an output layer). The activation function, which is a nonlinear function that transforms the input of a neuron into its output, was set to hyperbolic tangent for all layers. The batch size, which is the number of samples processed in each iteration, was set to one. The dropout rate, which is the probability of randomly dropping out neurons during training to prevent overfitting, was set to zero. The input time series data X_t consists of five sensor data DO, TMP, pH, MLSS, and airflow. Four of these time series DO, TMP, pH, and MLSS were chosen as the target variables to be predicted by the LSTM model, while the remaining parameter (airflow) was used as an explanatory variable. The LSTM model was trained and tested on four different scenarios, each with a different target variable, but the same set of explanatory variables was used (Figure 4). This way, the LSTM model could learn the complex interactions and dependencies among the different parameters in the MBR system.

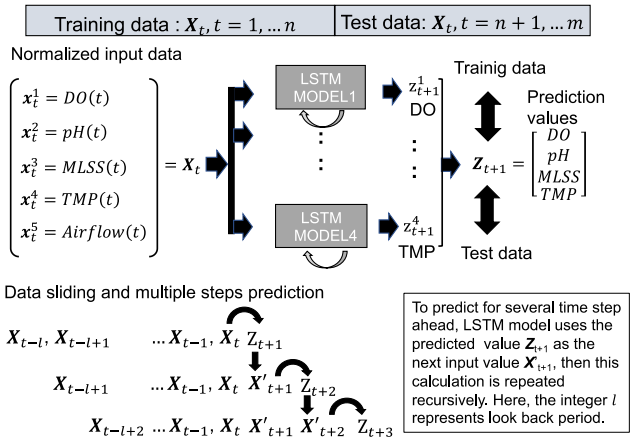


Figure 4 Time series prediction procedure using LSTM model

2.4 Model performance evaluation

To evaluate the predictive performance of the model, simulations were conducted for 15 trials, with results averaged to obtain robust metrics. Two key statistics were calculated to assess model accuracy - the root mean square error (RMSE) and Pearson's correlation coefficient (R). RMSE indicates the average deviation between predicted and observed values, with lower values showing less error. Pearson's R measures the strength of correlation, with values closer to 1 showing greater association between predictions and observations. Both metrics were computed using normalized data to enable standardized comparison across trials.

2.5 Software

The computational workflow for developing, training and evaluating the model utilized a computer workstation with an AMD RyzenTM 9 7900 CPU. Model implementation was performed in Python using TensorFlow and the Keras deep learning library. Additional Python packages for data handling and visualization included matplotlib (v3.7.1), seaborn (v0.12.2), pandas (v2.0.3), along with the core libraries of tensorflow (v2.10.1), keras (v2.10.0), and scipy (v1.10.1).

3. Result and Discussion

3.*1 One-hour prediction*

The bench scale MBR was operated for 800 hours to collect time

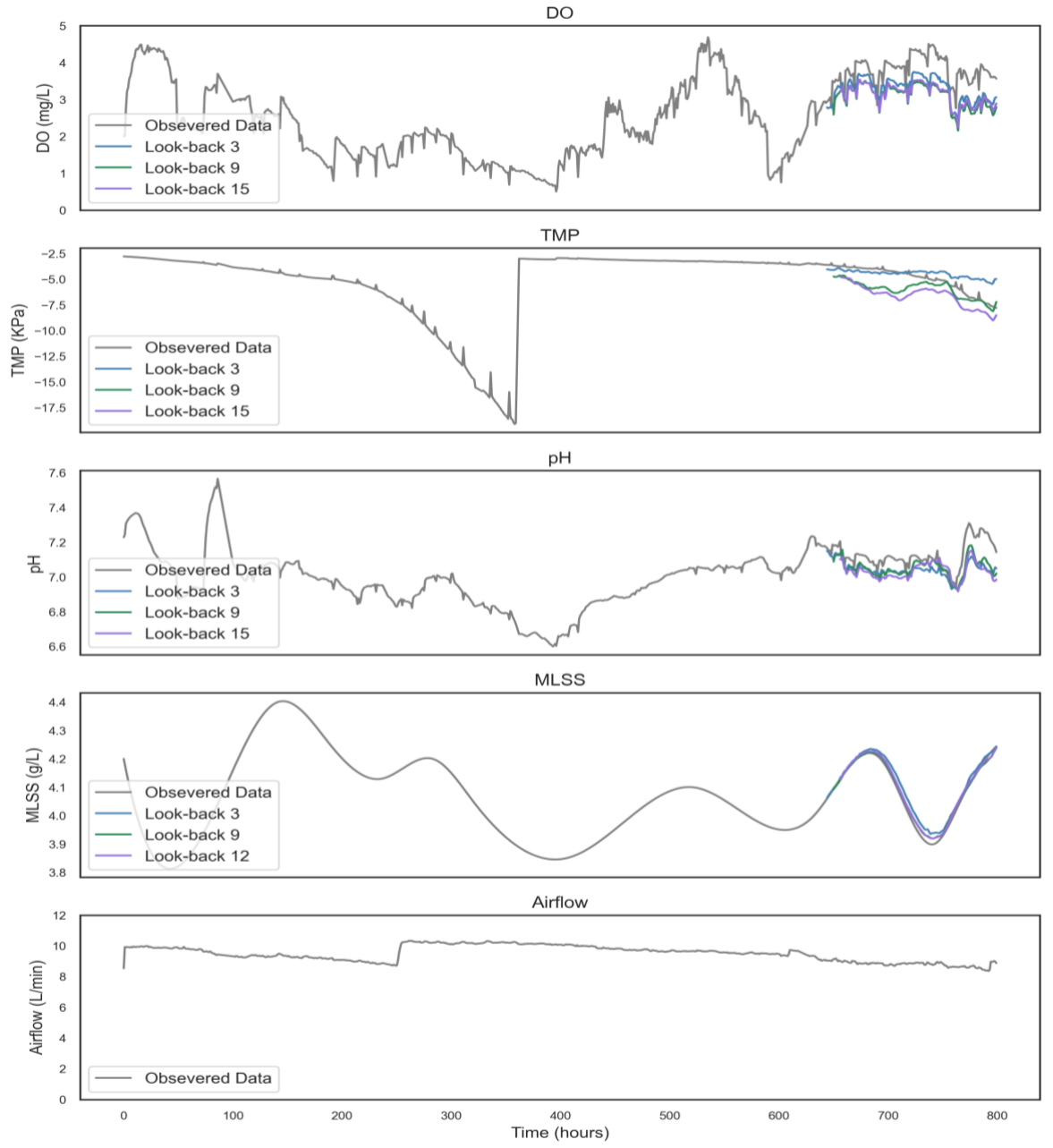


Figure 5 Observed data and one-hour prediction performance

series data of MBR operation status from the sensors (see Figure 4). The descriptive statistics (mean, standard deviation, min and max) of obtained data are shown in Table 1.

Table 1 Mean, standard deviation min and max of obtained data

Parameters	Mean	Standard deviation	Min	Max
DO (mg/L)	2.56	1.09	0.50	4.69
pH	7.01	0.16	6.60	7.56
MLSS (g/L)	4.06	0.14	3.81	4.40
TMP (KPa)	-5.21	3.48	-19.08	-2.78
Airflow (L/min)	9.50	0.50	8.36	10.34

LSTM model could predict the values one hour ahead in one step (refer to Figure 4). Figure 5 shows the original time series data and the time series of one hour ahead forecast values using different look back periods of 3, 9, and 15. It can be observed that the MLSS prediction curve well reproduced the observed MLSS changes for any look back period. For DO and pH, the LSTM model was also successful in predicting the tendency of rapid changes in observed values, but regardless of the look back period, both predicted values tended to be lower than the observed values. On the other hand, although the observed value of TMP showed a relatively simple change, the predicted curve did not sufficiently reproduce the trend of change in the observed value. In particular, as evident from Figure 5, the

prediction curves showed different trends depending on the value of the look back period. During this training period, there was only one case where fouling occurred and TMP increased (in a negative direction), so it is thought that LSTM was not able to fully learn the dynamic characteristics of TMP. On the contrary, other parameters such as pH had relatively sufficient fluctuations, and the model was able to learn their dynamic characteristics.

Look back period have a significant effect on the performance of the LSTM model (Chollet, 2021; Jeong et al., 2021). The optimal look back period depend on the characteristics of the data, such as the frequency, seasonality, trend, and noise. As shown in Figure 5, the prediction results varied depending on the look back period. Moreover, the look back period interacts with other parameters to affect the prediction accuracy. Generally, a larger look back period could capture more long-term dependencies and patterns in the data, and it also increased the complexity and computation time of the model. However, increasing the look back period too much could also reduce the effectiveness of the model by introducing irrelevant or redundant information (Koparanov et al., 2020). Figure 6 shows the RMSE values of the predictions for different look back periods. For DO and TMP, the RMSE was lowest with the look back period of 3 and increased as the look back period increased. For pH, which has low volatility and no clear trend, the optimal look back period is 9 hours, as shown by the lowest RMSE value among the different look back periods. However, the RMSE values do not vary significantly with the look back periods, suggesting that pH is not sensitive to the length of look back periods. For a smooth, moderately variable MLSS that the model could easily learn, a lookback period of 6 resulted in a low RMSE, and higher values did not make a significant difference in prediction accuracy. However, with the look back period of 3, the LSTM model may not have enough data to accurately predict MLSS, resulting in a higher RMSE than with the longer look back period. Therefore, to increase the accuracy and efficiency of the LSTM model, it seems important to select an appropriate look back period based on the characteristics of the data generation dynamics.

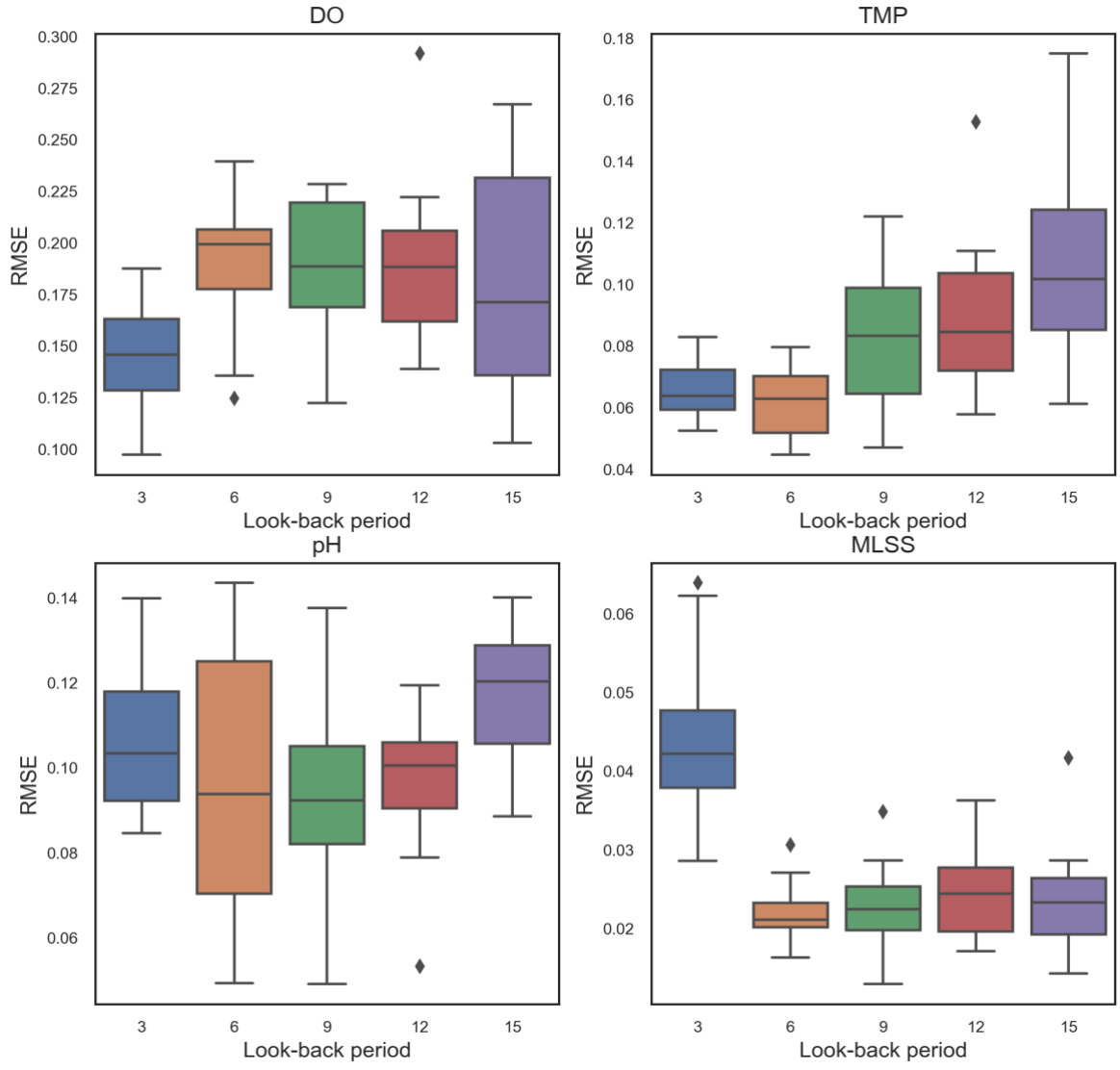


Figure 6 Performance of one-hour prediction at difference look back period

The above results indicate that three look back period provides good prediction performance except for MLSS. The correlation plot shown in Figure 7 compares the predicted values of the LSTM model with a lookback period of 3 to the observed data for MLSS, TMP, pH, and DO. The LSTM model predicted the MLSS with very high accuracy (RMSE = 0.044, R = 0.993), as shown by the correlation plot close to the Y=X line. This result was expected, since MLSS is related to the microbial growth in the MBR system, which does not vary significantly on an hourly basis, but rather on a daily scale. Therefore, the interpolation method used for MLSS was reasonable. However, the curved shape of the correlation plot suggests that the LSTM model learned the properties of polynomial interpolation, rather than the underlying dynamics of MLSS. On the other hand, the LSTM model could not predict TMP well (RMSE=0.066, R=0.913). As already mentioned, the LSTM model only had one chance to learn the characteristic TMP changes (see Figure 5), which was not enough to train the network of LSTM model. The pH and DO predictions were also inaccurate (pH: RMSE=0.106, R=0.631 and DO: RMSE=0.145, R=0.649), as indicated by the low correlation coefficients and the underestimation of these variables. However, Figure 5 shows that the LSTM model captured some of the short-term fluctuations of pH and DO, which are influenced by various factors such as wastewater composition, aeration rate, and membrane fouling. Previous research has also demonstrated LSTM's ability to capture these complex oscillations (Toffanin et al., 2023).

3.2 Multi-hours prediction

To extend the prediction range beyond 1 hour, the LSTM model adopted a recursive approach, as explained in Figure 4. The model can generate multi-step predicted values with this repeat. Obviously, predictions far into the future should be less accurate,

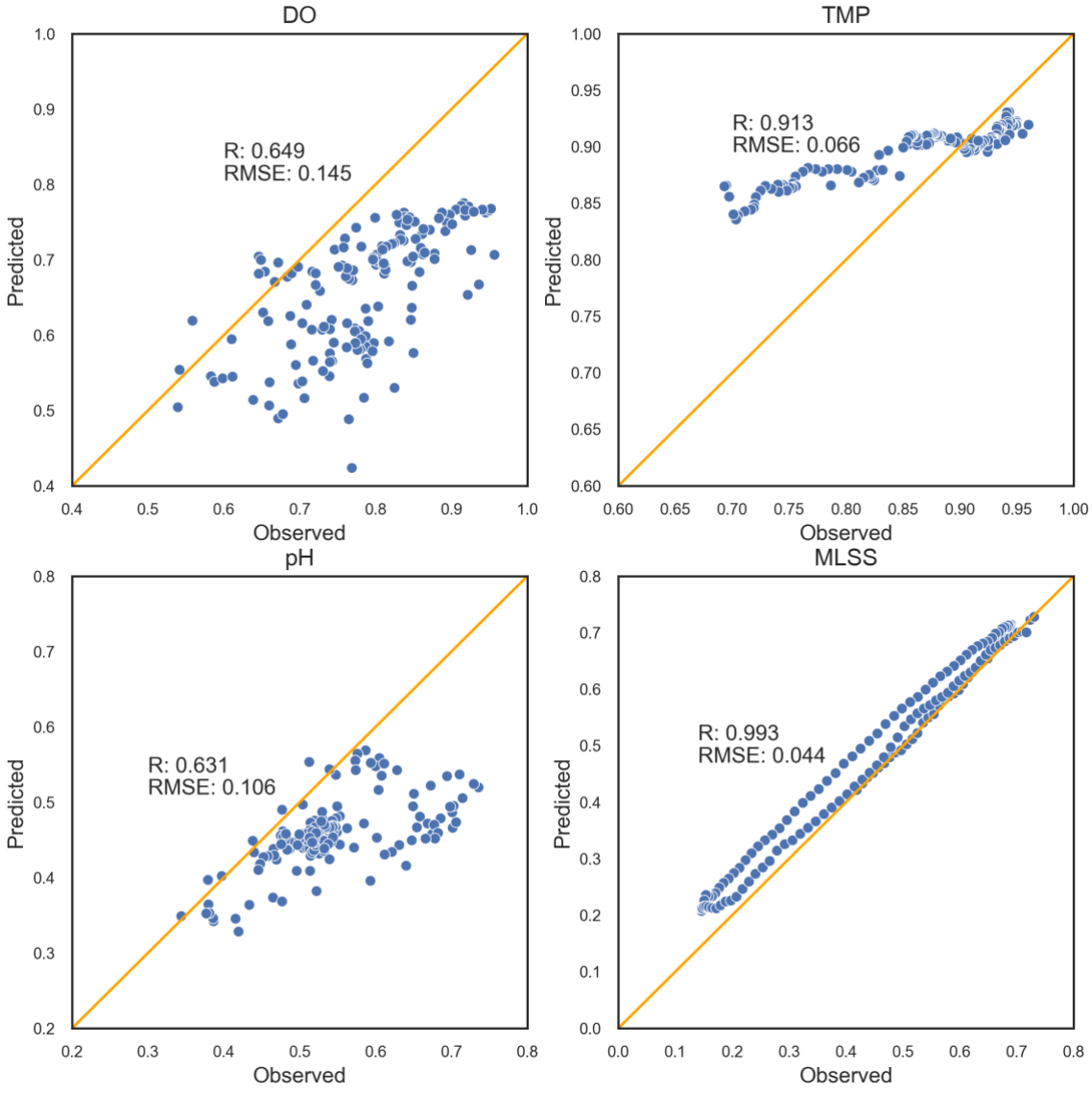


Figure 7 Correlation of observed data and one-hour prediction (look back period = 3, R is Pearson's correlation coefficient, RMSE is root mean square error)

and furthermore, the look back period of the LSTM model can influence the decrease in prediction accuracy after multiple steps, but their effects may not be always intuitive (Farhi et al., 2021). Figure 8 shows the change in RMSE of each predictor variable such as DO when the forecast horizon varies up to 6 hours (6 steps ahead) and the look back period varies up to 15 hours (15 steps backward).

For DO prediction, the LSTM model achieved the lowest RMSE at all time steps with the three hours look back period, and the RMSE basically increased as the prediction period progressed. However, when the model predicted DO 5 and 6 hours ahead, the 15 look back period had a lower RMSE. DO is influenced by the airflow rate and the microbial activity in the MBR system. As shown in Figure 5, air flow rate does not show large fluctuations like DO, so it is thought that microbial activity fluctuated greatly. The activated sludge system in MBR consists of a highly complex microbial ecosystem and is essentially a nonlinear system (Samsudin et al., 2014). Therefore, it is not surprising that DO exhibits chaotic behaviour. Predicting chaotic systems is not easy, even for short-term predictions, and the application of LSTM to predicting chaotic systems is at the forefront of research in the fields of nonlinear science (Chang et al., 2019; Fan et al., 2021; Sangiorgio & Dercole, 2020). For pH prediction, the optimal look back period was 9 hours, as discussed earlier. The RMSE also increased with the forecasting horizon as well as DO. However, the prediction performance was better than DO because the volatility of pH value was lower than DO.

For TMP forecasting, the best look back period was 6 hours for the 1-hour ahead forecast, while the 5-hour ahead forecast was best for a 3-hour look back period. This is a somewhat strange result, as it goes against the general idea that prediction accuracy decreases the further into the future. This result may be due to the LSTM

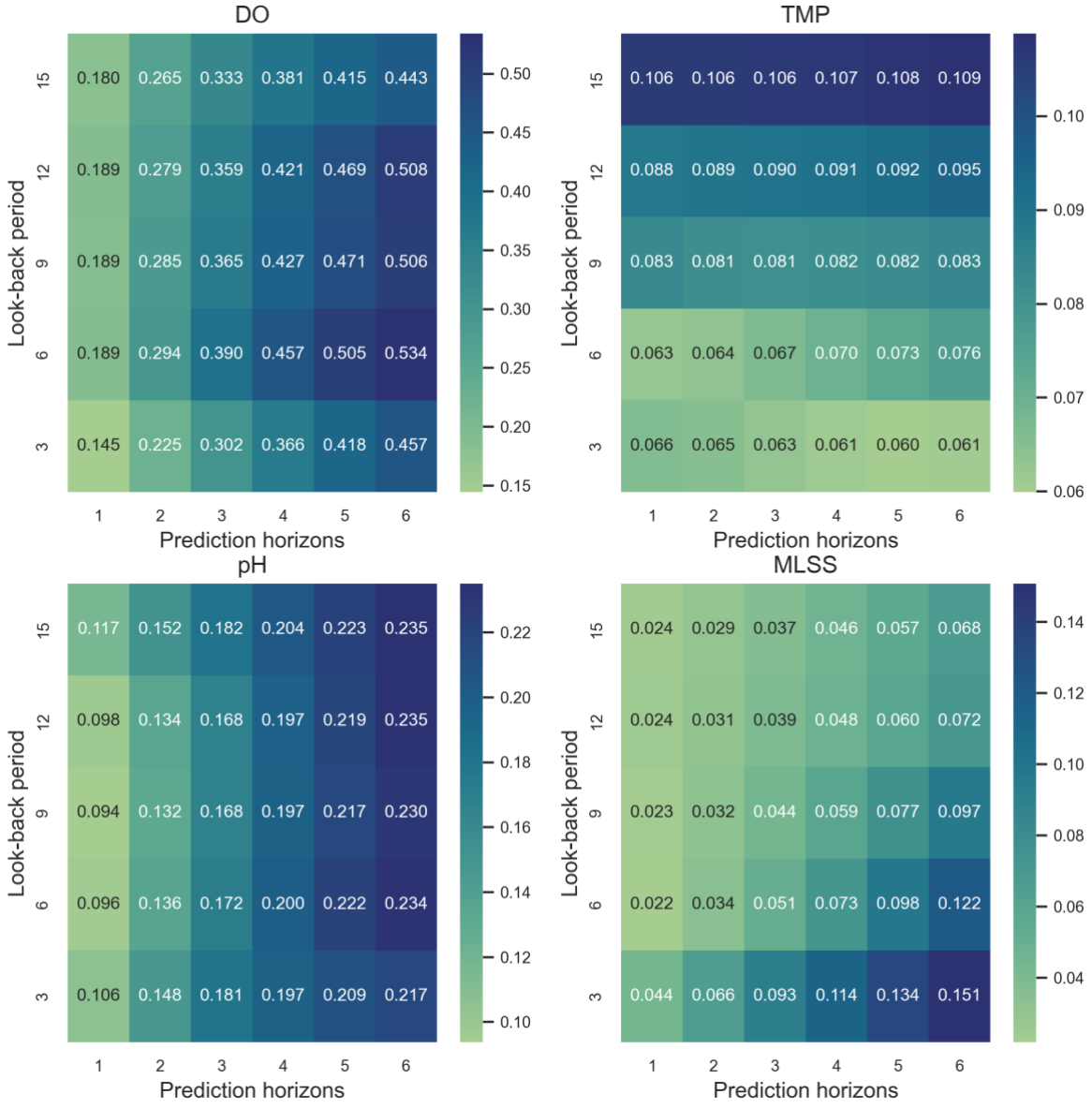


Figure 8 The performance of the model when using different look back period and time steps

model not learning the changes in TMP sufficiently.

For MLSS forecasting, the optimal look back period depended on the forecasting horizon. Previous study also mentioned a similar phenomenon (Farsi et al., 2021). If the forecasting horizon was 2 hours or less, the best performance was obtained with a 6-hour look back period. As the forecasting horizon was 3 hours or more, the performance improved with increasing look back period.

As already mentioned, changes in MLSS are correlated with microbial cell proliferation and cell lysis, and the time scale of the changes is on the order of days. For this LSTM model, the update time interval for the iterative calculations was 1 hour, which was sufficient to track changes on a daily time scale. Moreover, the MLSS smoothly changed due to the interpolation by 3rd order polynomial. It is easy for LSTM to learn the characteristics of this change. In actual sludge management, it is sufficient to be able to predict daily changes in MLSS. On the other hand, with DO, the time scale of fluctuation is shorter than that of MLSS, as there are cases where the DO drops suddenly after several hours. In actual MBR management, it is necessary to change the aeration amount in a short period of time to stabilize DO. In that case, forecasting one hour ahead is sufficient. As with DO, it is sufficient to predict pH on an hourly basis. However, to track such hourly changes, the update time interval for iterative calculations must be as short as 10 minutes. However, to track such changes over time, it is necessary to reduce the update interval between iterations of the LSTM model to approximately 10 minutes. Setting the time interval appropriately according to the time scale of the phenomenon is important in time series analysis models. Even in the LSTM model, since MLSS and DO or pH have different time scales, it is necessary to improve the calculation algorithm so that the update times can be set separately. Biological treatment systems such as activated sludge are complex nonlinear systems with many different time scales. The development of deep learning algorithms that can learn multidimensional time series based on these characteristics requires research that goes back to the basics, and close cooperation with the fields of data science and mathematical science is essential.

4. Conclusion

The LSTM model predicted MLSS one hour ahead with high accuracy, while it also reproduced the temporal fluctuation patterns of DO and pH to some extent. Nevertheless, the model tended to underestimate both DO and pH compared to the actual measurements. The model failed to reproduce the pattern of TMP changes well, despite having lower RMSE values than DO and pH. This might be because the membrane fouling of MBR, which affects TMP, occurred only once during the training period, and this single event was insufficient for training the LSTM model.

The optimal lookback period for the LSTM model varied depending on the parameter and the prediction horizon. For MLSS, a lookback period of 6 was optimal for forecasting 1 and 2 hours ahead, and longer lookback periods performed better for longer horizons. On the other hand, for the parameter DO, a lookback period of 3 was optimal for forecasting 1 to 3 hours ahead, while a lookback period of 15 was found to be better for forecasting 4 to 6 hours ahead. When it comes to pH, a lookback period of 9 gave the best results for forecasting 1 hour ahead, while a lookback period of 3 was optimal for forecasting 2 to 6 hours ahead. Finally, for the parameter TMP, a lookback period of 6 was optimal for forecasting 1 to 2 hours ahead, and a lookback period of 3 was better for forecasting 3 to 6 hours ahead.

Acknowledgement

This research was supported by the Project for Fostering of Proficient Researchers for the Establishment of the Research Center for Green Science (fellowship scholarship) and A-STEP Tryout (JPMJTM22EH) of Japanese Science and Technology Agency.

References

ArunKumar, K. E., Kalaga, D. V., Kumar, Ch. M. S., Kawaji, M., & Brenza, T. M. (2021). Forecasting of COVID-19 using deep layer Recurrent Neural Networks (RNNs) with Gated Recurrent Units (GRUs) and Long Short-Term Memory (LSTM) cells. Chaos, Solitons & Fractals, 146, 110861. https://doi.org/10.1016/j.chaos.2021.110861

Asadi, A., Verma, A., Yang, K., & Mejabi, B. (2017). Wastewater treatment aeration process optimization: A data mining approach. Journal of Environmental Management, 203, 630–639. https://doi.org/10.1016/j.jenvman.2016.07.047

Bas, E., Egrioglu, E., & Kolemen, E. (2022). Training simple recurrent deep artificial neural network for forecasting using particle swarm optimization. Granular Computing, 7(2), 411–420. https://doi.org/10.1007/s41066-021-00274-2

Chang, Z., Zhang, Y., & Chen, W. (2019). Electricity price prediction based on hybrid model of adam optimized LSTM neural network and wavelet transform. Energy, 187, 115804. https://doi.org/10.1016/j.energy.2019.07.134

Chen, Y., Zhang, H., Yin, Y., Zeng, F., & Cui, Z. (2022). Smart energy savings for aeration control in wastewater treatment. Energy Reports, 8, 1711–1721. https://doi.org/10.1016/j.egyr.2022.02.038

Chollet, F. (2021). Deep Learning with Python, Second Edition. Simon and Schuster.

da Silva, I. N., Hernane Spatti, D., Andrade Flauzino, R., Liboni, L. H. B., & dos Reis Alves, S. F. (2017). Artificial Neural Network Architectures and Training Processes. In I. N. da Silva, D. Hernane Spatti, R. Andrade Flauzino, L. H. B. Liboni, & S. F. dos Reis Alves (Eds.), Artificial Neural Networks: A Practical Course (pp. 21–28). Springer International Publishing. https://doi.org/10.1007/978-3-319-43162-8_2

Du, X., Shi, Y., Jegatheesan, V., & Haq, I. U. (2020). A Review on the Mechanism, Impacts and Control Methods of Membrane Fouling in MBR System. Membranes, 10(2), 24. https://doi.org/10.3390/membranes10020024

Dufera, T. T., Seboka, Y. C., & Fresneda Portillo, C. (2022). Parameter Estimation for Dynamical Systems Using a

Deep Neural Network. Applied Computational Intelligence and Soft Computing, 2022, e2014510. https://doi.org/10.1155/2022/2014510

Fan, D., Sun, H., Yao, J., Zhang, K., Yan, X., & Sun, Z. (2021). Well production forecasting based on ARIMA-LSTM model considering manual operations. Energy, 220, 119708. https://doi.org/10.1016/j.energy.2020.119708

Farsi, B., Amayri, M., Bouguila, N., & Eicker, U. (2021). On Short-Term Load Forecasting Using Machine Learning Techniques and a Novel Parallel Deep LSTM-CNN Approach. IEEE Access, 9, 31191–31212. https://doi.org/10.1109/ACCESS.2021.3060290

Fenu, A., Roels, J., Wambecq, T., De Gussem, K., Thoeye, C., De Gueldre, G., & Van De Steene, B. (2010). Energy audit of a full scale MBR system. Desalination, 262(1), 121–128. https://doi.org/10.1016/j.desal.2010.05.057

Greff, K., Srivastava, R. K., Koutník, J., Steunebrink, B. R., & Schmidhuber, J. (2017). LSTM: A Search Space Odyssey. IEEE Transactions on Neural Networks and Learning Systems, 28(10), 2222–2232. https://doi.org/10.1109/TNNLS.2016.2582924

Gu, Y., Li, Y., Yuan, F., & Yang, Q. (2023). Optimization and control strategies of aeration in WWTPs: A review. Journal of Cleaner Production, 418, 138008. https://doi.org/10.1016/j.jclepro.2023.138008

Hewamalage, H., Bergmeir, C., & Bandara, K. (2021). Recurrent Neural Networks for Time Series Forecasting: Current status and future directions. International Journal of Forecasting, 37(1), 388–427. https://doi.org/10.1016/j.ijforecast.2020.06.008

Huang, R., Wei, C., Wang, B., Yang, J., Xu, X., Wu, S., & Huang, S. (2022). Well performance prediction based on Long Short-Term Memory (LSTM) neural network. Journal of Petroleum Science and Engineering, 208, 109686. https://doi.org/10.1016/j.petrol.2021.109686

Jeong, K., Abbas, A., Shin, J., Son, M., Kim, Y. M., & Cho, K. H. (2021). Prediction of biogas production in anaerobic co-digestion of organic wastes using deep learning models. Water Research, 205, 117697. https://doi.org/10.1016/j.watres.2021.117697

Judd, S. J. (2016). The status of industrial and municipal effluent treatment with membrane bioreactor technology. Chemical Engineering Journal, 305, 37–45. https://doi.org/10.1016/j.cej.2015.08.141

Kaewpipat, K., & Grady, C. P. L. (2002). Microbial population dynamics in laboratory-scale activated sludge reactors. Water Science and Technology: A Journal of the International Association on Water Pollution Research, 46(1–2), 19–27.

Karnam, N. K., Dubey, S. R., Turlapaty, A. C., & Gokaraju, B. (2022). EMGHandNet: A hybrid CNN and Bi-LSTM architecture for hand activity classification using surface EMG signals. Biocybernetics and Biomedical Engineering, 42(1), 325–340. https://doi.org/10.1016/j.bbe.2022.02.005

Koparanov, K. A., Georgiev, K. K., & Shterev, V. A. (2020). Lookback Period, Epochs and Hidden States Effect on Time Series Prediction Using a LSTM based Neural Network. 2020 28th National Conference with International Participation (TEL ECOM), 61 – 64. https://doi.org/10.1109/ TEL ECOM50385.2020.9299551

Kuan, L., Zhenfu, B., Xin, W., Xiangrong, M., Honghai, L., Wenxue, S., Zijian, Z., & Zhimin, L. (2017). Short-term CHP heat load forecast method based on concatenated LSTMs. 2017 Chinese Automation Congress (CAC), 99–103. https://doi.org/10.1109/CAC.2017.8242744

Li, C., Deng, W., Gao, C., Xiang, X., Feng, X., Batchelor, B., & Li, Y. (2019). Membrane distillation coupled with a novel two-stage pretreatment process for petrochemical wastewater treatment and reuse. Separation and Purification Technology, 224, 23–32. https://doi.org/10.1016/j.seppur.2019.05.007

Pisa, I., Santín, I., Morell, A., Vicario, J. L., & Vilanova, R. (2019). LSTM-Based Wastewater Treatment Plants Operation Strategies for Effluent Quality Improvement. IEEE Access, 7, 159773–159786. https://doi.org/10.1109/ACCESS.2019.2950852

Qiao, D., Li, P., Ma, G., Qi, X., Yan, J., Ning, D., & Li, B. (2021). Realtime prediction of dynamic mooring lines responses with LSTM neural network model. Ocean Engineering, 219, 108368. https://doi.org/10.1016/j.oceaneng.2020.108368

Samsudin, S. I., Rahmat, M. F., Wahab, N. A., Razali, M. C., Gaya, M. S., & Salim, S. N. S. (2014). Improvement of Activated Sludge Process Using Enhanced Nonlinear PI Controller. Arabian Journal for Science and Engineering, 39(8), 6575–6586. https://doi.org/10.1007/s13369-014-1285-2

Sangiorgio, M., & Dercole, F. (2020). Robustness of LSTM neural networks for multi-step forecasting of chaotic time series. Chaos, Solitons & Fractals, 139, 110045. https://doi.org/10.1016/j.chaos.2020.110045

Schmitt, F., Banu, R., Yeom, I.-T., & Do, K.-U. (2018). Development of artificial neural networks to predict membrane fouling in an anoxic-aerobic membrane bioreactor treating domestic wastewater. Biochemical Engineering Journal, 133, 47–58. https://doi.org/10.1016/j.bej.2018.02.001

Tang, K., Xie, J., Pan, Y., Zou, X., Sun, F., Yu, Y., Xu, R., Jiang, W., & Chen, C. (2022). The optimization and regulation of energy consumption for MBR process: A critical review. Journal of Environmental Chemical Engineering, 10(5), 108406. https://doi.org/10.1016/j.jece.2022.108406

Toffanin, C., Di Palma, F., Iacono, F., & Magni, L. (2023). LSTM Network for the Oxygen Concentration Modeling of a Wastewater Treatment Plant. Applied Sciences, 13(13), 7461. https://doi.org/10.3390/app13137461

Wang, S., Zou, L., Li, H., Zheng, K., Wang, Y., Zheng, G., & Li, J. (2020). Full-scale membrane bioreactor process WWTPs in East Taihu basin: Wastewater characteristics, energy consumption and sustainability. Science of The Total Environment, 723, 137983. https://doi.org/10.1016/j.scitotenv.2020.137983

Wongburi, P., & Park, J. K. (2023). Prediction of Wastewater Treatment Plant Effluent Water Quality Using Recurrent Neural Network (RNN) Models. Water, 15(19), Article 19. https://doi.org/10.3390/w15193325

Xiao, K., Liang, S., Wang, X., Chen, C., & Huang, X. (2019).

Current state and challenges of full-scale membrane bioreactor applications: A critical review. Bioresource Technology, 271, 473–481. https://doi.org/10.1016/j.biortech.2018.09.061

Yoon, N., Kim, J., Lim, J.-L., Abbas, A., Jeong, K., & Cho, K. H. (2021). Dual-stage attention-based LSTM for simulating performance of brackish water treatment plant. Desalination, 512, 115107. https://doi.org/10.1016/j.desal.2021.115107

Zhang, J., Inamori, R., Suemura, T., Feng, C., Xu, K.-Q., & Inamori, Y. (2018). Advanced Wastewater Treatment and Power Reduction in a Multiple-Reactor Activated Sludge Process with Automatic Oxygen Supply Device System Installation. Japanese Journal of Water Treatment Biology, 54(1), 13–27. https://doi.org/10.2521/jswtb.54.13

Zhong, H., Yuan, Y., Luo, L., Ye, J., Chen, M., & Zhong, C. (2022). Water quality prediction of MBR based on machine learning: A novel dataset contribution analysis method. Journal of Water Process Engineering, 50, 103296. https://doi.org/10.1016/j.jwpe.2022.103296