

Prediction of Membrane Fouling based on SSA-LSTM Neural Network

Guobi Ling

College of Electrical and Information Engineering, Lanzhou University of Technology,
Lanzhou, 730050, China

*Corresponding author: guobi_ling@163.com

Abstract

In view of the difficulty in obtaining the membrane bioreactor (MBR) membrane flux in real time and stable control, a sparrow search algorithm (SSA) is proposed to optimize the SSA-LSTM prediction model of long short-term memory (LSTM) neural network. Firstly, principal component analysis (PCA) is used to realize the dimensionality reduction of auxiliary variables. Second, use the sparrow search algorithm to determine the relevant hyper-parameters of the LSTM neural network. Finally, the selected auxiliary variables are used as the input of the SSA-LSTM prediction model, and the membrane flux is used as the prediction output, the measured data is used as the sample for experimental verification. The accuracy rate reaches 94.31%, which is much higher than 63.63% of LSTM. The results show that the proposed membrane flux prediction model has higher prediction performance.

Keywords

MBR; Membrane Flux Prediction; LSTM; SSA-LSTM Model.

1. Introduction

In recent years, with the rapid economic growth and industrial development, water resource pollution deteriorates sharply, and sewage treatment is particularly important [1]. In the process of sewage treatment, improving sewage treatment efficiency, realizing green environmental protection, saving energy, reducing production costs, and improving economic benefits are crucial to the field of sewage treatment [2]. As an important means in sewage treatment engineering, MBR process has been rapidly favored by all countries in the world since the 21st century, with the progress of membrane separation technology, assembly structure and equipment manufacturing, as well as the tightening of sewage treatment and discharge standards [3]. MBR is a new type of wastewater treatment system combining membrane technology and biological treatment technology, which is mainly composed of membrane module and bioreactor. Compared with the traditional wastewater treatment process, MBR has the advantages of good and stable effluent quality, low sludge yield, compact equipment and no large space occupation [4]. At present, the study of MBR and the prediction of membrane fouling is one of the important topics in the field of sewage treatment. However, membrane fouling seriously affects the permeability and life of membrane, and membrane fouling will increase the operation cost of membrane bioreactor, which becomes a bottleneck problem limiting the wide application of MBR [5]. Therefore, timely and accurate prediction of membrane flux is the key to control membrane contamination.

The degree of membrane contamination is related to the value of membrane flux. In practice, the fouling state of membrane can be analyzed by predicting the value of membrane flux. In order to realize the detection of membrane permeability, researchers have done a lot of research. Al-Zoubi et Al. [6] obtained a number of process variables related to membrane

permeability through mechanism analysis, selected 29 of them as auxiliary variables, and established a soft sensing model of membrane permeability based on BPNN. The prediction accuracy reaches 70%, but the selection of too many auxiliary variables also leads to the poor anti-interference ability of the network. Liu et al. [7] A particle swarm optimization back propagation network (PSO-BP) model is studied, the results show that the optimized model has higher prediction accuracy, and the average error decreases from 2.35% to 0.83%. However, the combination of the optimization algorithm model and the parameter optimization process requires a large amount of time, which limits the scope of its application.

At present, SSA have been widely used in engineering problems. Liu et al. [8] established a prediction model based on sparrow search algorithm optimized support vector machine regression (SSA-SVR) to predict the subsidence of coal gangue roadbed of Shao Expressway in Hunan Province. The results show that the SSA-SVR prediction model has a high accuracy. Liu et al. [9] aiming at the instability of Extreme learning machine (ELM) model and inaccurate prediction results, a combined prediction model based on sparrow search algorithm (SSA) was proposed to optimize extreme learning machine and realize accurate prediction of wind power. meanwhile, LSTM neural network has been successfully applied in financial market price prediction [10], traffic flow prediction [11], ocean surface temperature prediction [12] and medical prediction [13].

However, the values of these hyper-parameters are often not optimal, which leads to the model not achieving the best prediction results. Therefore, this paper proposed SSA-LSTM algorithm to optimize and LSTM neural network through SSA, and applied it to membrane flux prediction in sewage treatment membrane pollution, and the simulation data of Seong-Hoon Yoon's spreadsheet model were used to verify the prediction model and apply it to the prediction of membrane flux value.

2. Long- and Short-Term Memory Network

The traditional neural network model will lose the remote information, and it is difficult to learn the long-distance dependent information. LSTM is an improvement of the recurrent neural network, which aims to overcome the defects of the recurrent neural network in processing long-term memory. The LSTM introduced the concept of cellular states, which determine which states should be preserved and which should be forgotten. The LSTM unit structure diagram is shown in Figure 1.

The LSTM unit consists of three parts: input gate i_t , forgetting gate f_t and output gate o_t . The input gate determines the data in the current input that needs to be kept in the current state. The forgetting gate determines the data to be forgotten in the status information at the last moment. The output gate determines the part of the information that can be used as the final output by the input at the current time, the input at the previous time, and the state combination. Formulas (1)- (5) describe the working principle of LSTM [14]:

$$i_t = \sigma(W_{xi}x_t + W_{hi}h_{t-1} + W_{ci}c_{t-1} + b_i) \quad (1)$$

$$f_t = \sigma(W_{xf}x_t + W_{hf}h_{t-1} + W_{cf}c_{t-1} + b_f) \quad (2)$$

$$c_t = f_t c_{t-1} + i_t \tanh(W_{xc}x_t + W_{hc}h_{t-1} + b_c) \quad (3)$$

$$o_t = \sigma(W_{xo}x_t + W_{ho}h_{t-1} + W_{co}c_{t-1} + b_o) \quad (4)$$

$$h_t = o_t \tanh(c_t) \tag{5}$$

In these formulas, i_t, f_t, o_t represents the output of the input gate, the forgetting gate and the output gate, $\sigma(\bullet)$ represent Sigmoid activation function, used to map the output of the LSTM gating unit to between $[0,1]$. W and b are the corresponding weight matrix and bias vector of the three control gates and memory cells, and the subscripts f, i and o correspond to the forgetting gate, input gate and output gate, respectively. x_t is the input of LSTM unit at time t , c_t and c_{t-1} are the values stored in the neuron at time t and $t-1$, respectively. h_t and h_{t-1} correspond to the output of the hidden layer at time t and $t-1$.

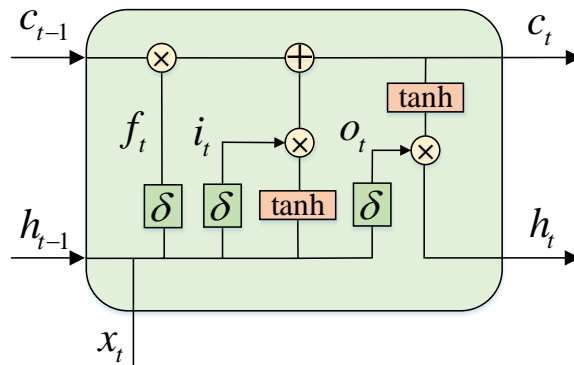


Figure 1. LSTM unit structure diagram.

3. Sparrow Search Algorithm

Sparrow search algorithm (SSA) is a new metaheuristic algorithm proposed by Xue and Shen in 2020 [15], inspired by the predatory and anti-predatory behaviors of sparrows in biology, it can be abstracted as an explorer-follower-warner model.

In SSA, if there are N sparrows in a D - dimensional search space, the location of each sparrow is shown in formula 2:

$$X = \begin{pmatrix} x_{1,1} & x_{1,2} & \cdots & x_{1,d} & \cdots & x_{1,D} \\ x_{2,1} & x_{2,2} & \vdots & x_{2,d} & \vdots & x_{2,D} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_{i,1} & x_{i,2} & \cdots & x_{i,d} & \cdots & x_{i,D} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_{N,1} & x_{N,2} & \cdots & x_{N,d} & \cdots & x_{N,D} \end{pmatrix} \tag{6}$$

In the formula, $i = 1, 2, \dots, N, d = 1, 2, \dots, D, x_{i,d}$ indicates the position of the i -th sparrow in the d -th dimension.

Since the explorer guides the movement of the whole sparrow population, and can find food anywhere, its location is updated as follows:

$$x_{i,d}^{t+1} = \begin{cases} x_{i,d}^t \cdot \exp\left(\frac{-i}{a \cdot iter_{max}}\right), R_2 < ST \\ x_{i,d}^t + Q \times L, R_2 \geq ST \end{cases} \tag{7}$$

In the formula, t represents the current number of iterations, $iter_{max}$ is the maximum number of iterations, a is a random number with an interval $(0,1]$, Q is a random number with a normal distribution, L represents a matrix of $1 \times d$, where each element is 1, $R_2 \in [0,1]$, represents the alert value; $ST \in [0.5,1]$, represents a security threshold. When $R_2 < ST$, it means there are no predators around, and the explorer will enter a wider search mode; if $R_2 \geq ST$, it means that some sparrows have found predators and all need to fly quickly to other security zones.

Followers follow the explorer in search of food and may compete with the explorer for food to increase their own predation rate, we can update formula to:

$$x_{i,d}^{t+1} = \begin{cases} Q \cdot \exp\left(\frac{xworst_d^t - x_{i,d}^t}{a \cdot iter_{max}}\right), & i > N/2 \\ xbest_d^{t+1} + \frac{1}{D} \sum_{d=1}^D (|x_{i,d}^t - xbest_d^{t+1}| \cdot rand\{-1,1\}), & i \leq N/2 \end{cases} \quad (8)$$

In the formula, $xworst_d^t$ represents the lowest overall position of sparrows in d -th dimension during t -th iteration, $xbest_d^t$ represents the lowest overall position of sparrows in d -th dimension during $t+1$ -th iteration. When $i > N/2$, the i -th follower with poor fitness was most likely to starve, otherwise, the i -th follower will randomly find a place near the best location for the explorer to feed.

Assuming that the early warning sparrow accounts for about 10%~20% of the sparrow population, its initial position is randomly determined, and the mathematical model can be expressed as:

$$x_{i,d}^{t+1} = \begin{cases} xbest_d^t + \beta \cdot |x_{i,d}^t - xbest_d^t|, & f_i \neq f_g \\ x_{i,d}^t + k \cdot \left(\frac{|x_{i,d}^t - xworst_d^t|}{|f_i - f_w| + \varepsilon}\right), & f_i = f_g \end{cases} \quad (9)$$

In the formula, β is a random normal distribution with a mean of 0 and a variance of 1, which is used as a step control parameter, k is a random number between $[-1, 1]$, f_i , f_g , f_w represents the fitness value of the current sparrow, the current global best fitness value, and the worst fitness value, respectively, ε is the minimum constant to avoid zero-point error. $f_i \neq f_g$ means sparrows are at the edge of the population, when $f_i = f_g$, sparrows in the middle of the population are aware of the danger and need to move elsewhere.

4. Membrane Fouling Prediction Model

4.1. Acquisition and Preprocessing of Experimental Data

The experimental data in this article come from a spreadsheet model of Seong-Hoon Yoon [16]. The subjects of study are all polyvinylidene fluoride (PVDF) hollow microfiltration(MF) components. The water inlet mode is external pressure, and the effective usage area of the membrane is 20m². The membrane module is in the MBR reaction tank because of the microorganism, the water quality and the water quality the components of pollutants are

complex, and the selection of influencing factors of membrane pollution largely determines the accuracy of membrane pollution prediction. Therefore, in order to improve the prediction accuracy of membrane pollution, the principal component analysis method is used to screen out the most influential factors of membrane pollution for experimental analysis. In order to ensure that the input and output data are in the same order of magnitude, the data is normalized first, then the input space is divided to facilitate the implementation of the algorithm.

4.2. Establishment of SSA-LSTM Prediction Model

Membrane flux is an important index parameter reflecting the degree of membrane fouling, therefore, we use it as the output of the model, and we use five factors after the principal component dimension reduction as input, finally, we establish a prediction model for membrane fouling based on SSA-LSTM as shown in figure 2.

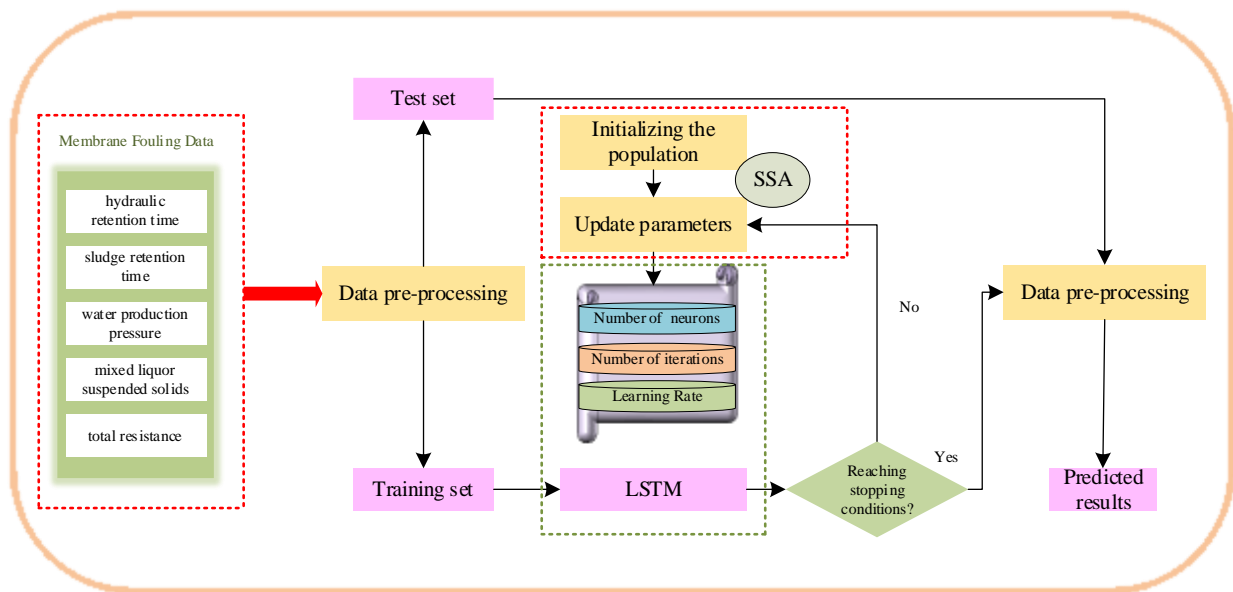


Figure 2. Membrane fouling prediction model.

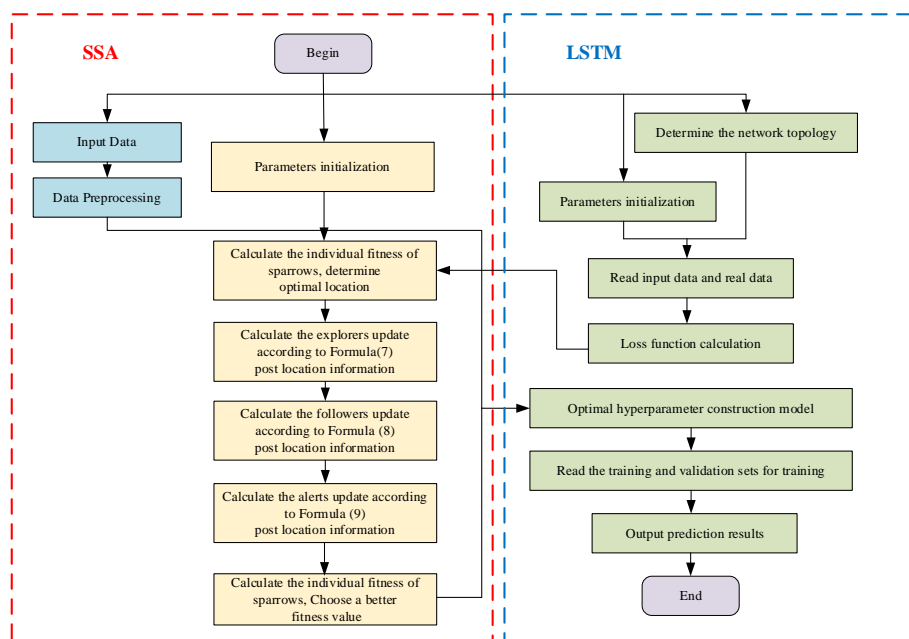


Figure 3. Flowchart of SSA-LSTM algorithm.

This article optimizes the learning parameters of LSTM network based on SSA, and then these optimized values are assigned to the network to get the optimized LSTM neural network. The performance of the optimized LSTM neural network is evaluated by experimental data. The algorithm flow chart is shown in Figure 3, where the left side of Figure 3 shows the algorithm flow of SSA and the right side shows the algorithm flow of LSTM model.

4.3. Evaluating Indicator

In order to make the output accuracy more intuitive, we introduce the following evaluation indicators.

1) Mean absolute percentage (MAPE)

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{y_i - y_i^*}{y_i^*} \right| \times 100\% \quad (10)$$

2) Root mean square error (RMSE)

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (y_i - y_i^*)^2}{n}} \quad (11)$$

3) Mean absolute error (MAE)

$$MAE = \frac{\sum_{i=1}^n |y_i - y_i^*|}{n} \quad (12)$$

In the formula, y_i is the true value, y_i^* is the model output value, n is the number of test samples.

5. Experiment and Simulation

5.1. Parameter Settings

SSA parameter is set to: number of sparrows were set up to 30, the maximum number of iterations is 100, safety threshold value of 0.6, the proportion of the entrant is 0.7, aware of the danger of the sparrow accounted for 0.2. LSTM network parameters: in order to ensure that the training speed and precision, the largest number of training for 200; The learning rate and the number of hidden layer units are determined by the optimization algorithm, in which the learning rate ranges within [0.001, 0.5] and the number of units ranges within [50,500].

5.2. Experimental Results and Analysis

In order to improve the prediction accuracy, we finally determined 200 sets of data for model simulation through PCA dimensionality reduction normalization, 160 sets of data for model training, and the remaining 40 sets of data for model testing. Compare the optimized network with the training results of traditional LSTM network are compared, as shown in figure 4, and error comparison diagram as shown in figure 5.

From the simulation curve in figure 4, it can be seen that the deviation between the unmodified LSTM network and the true value fluctuates greatly, and the improved prediction model has a smaller fluctuation than the true value. From the comparison of prediction errors in figure 5, it can be seen that the SSA-LSTM model proposed in this article has a greater improvement in

prediction errors than the standard algorithm, and the fluctuation range of errors is significantly reduced, which can better reflect the change of membrane flux.

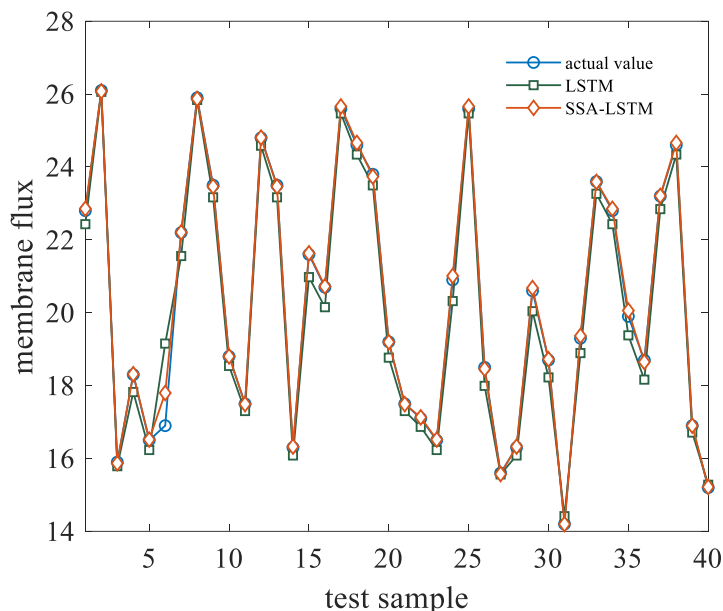


Figure 4. prediction contrast map.

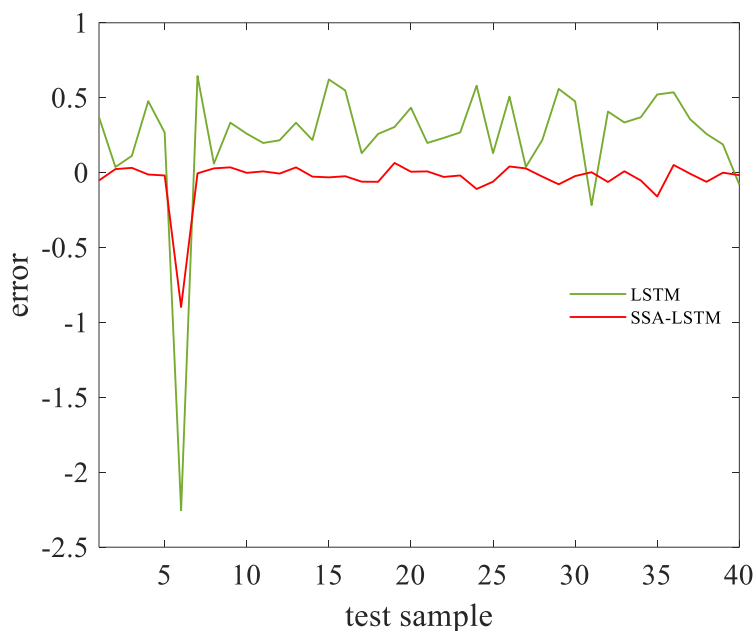


Figure 5. error comparison diagram.

According to the data in table 4, under the same experimental conditions, compared with the LSTM model before improvement, the improved soft sensing model reduces MAPE by 92.76%, RMSE by 96.78% and MAE by 93.61%. The prediction accuracy of the algorithm proposed in this article reaches 94.31%, which is much higher than 63.63% of LSTM. According to various indicators, the prediction effect of SSA-LSTM soft sensing model is better than that of the unmodified soft sensing model.

Table 1. Prediction error comparison.

Model	EVA		
	MAPE/%	RMSE	MAE
LSTM	0.0716	0.2917	0.3637
SSA-LSTM	0.0204	0.0224	0.0569

6. Conclusion

The intelligent property of LSTM neural network can predict MBR membrane fouling according to the characteristics of membrane flux. Since the prediction accuracy largely depends on the choice of network model parameters, the main factors of membrane fouling are analyzed in this chapter, and five main factors are obtained through dimensionality reduction of principal component data. Membrane fouling is characterized by membrane flux. The prediction model of LSTM membrane fouling is constructed, and the relevant parameters of LSTM network are optimized by sparrow search algorithm. The results show that the prediction accuracy of membrane fouling prediction model based on SSA-LSTM is significantly improved compared with that of single LSTM model, and the experiment reaches the expected goal.

Conflicts of Interest

The authors declare no conflict of interest.

References

- [1] Qin, X.; Gao, F.; Chen, G. Wastewater quality monitoring system using sensor fusion and machine learning techniques. *Water Res.* 2012, 46, 1133-1144.
- [2] Tsui, T.H.; Zhang, L.; Zhang, J.X.; Dai, Y.J.; Tong, Y.W. Engineering interface between bioenergy recovery and biogas desulfurization: Sustainability interplays of biochar application. *Renew sust energ rev.* 2022, 157,112053.
- [3] Han, H.G.; Zhang, S.; Qiao, J.F. Soft-sensor Method for Permeability of the Membrane Bio-Reactor Based on Recurrent Radial Basis Function Neural Network. *J. Beijing Univ. Technol.* 2017, 43, 1168-1174.
- [4] Deb, A.; Gurung, K.; Rumky, J.; Sillanpaa, M.; Manttari, M.; Kallioinen, M. Dynamics of microbial community and their effects on membrane fouling in an anoxic-oxic gravity-driven membrane bioreactor under varying solid retention time: A pilot-scale study. *Science of The Total Environment.* 2022, 807, 150878.
- [5] Shi, Y.K.; Wang Z.W.; Du, X.J.; Gong B.; Jegatheesan, V.; Haq UI. Recent advances in the prediction of fouling in membrane bioreactors. *Membranes.* 2021, 11, 381.
- [6] Al-Zoubi, H.; Hilal, N.; Darwish, N.A.; Mohammad, A.W. Rejection and modelling of sulphate and potassium salts by nanofiltration membranes: Neural network and Spiegler-Kedem model. *Desalination.* 2006, 206, 42-60.
- [7] Liu, Z.F.; Pan, D.; Wang, J.H.; Yang, S.X. The film pollution forecast of PSO-BP neural network in MBR technology. *Journal of Beijing University of Technology.* 2012, 38, 126-131.
- [8] Liu, Y.L. Study on prediction model of coal gangue subgrade settlement based on SSA-SVR. *Journal of Hebei University of Geosciences.* 2021, 44, 99-104.
- [9] Liu, D.; Wei, X.; Wang, W.Q.; Ye, J.H.; Ren, J. Short-term wind power prediction based on SSA-ELM. *Smart Power.* 2021, 49, 53-59.
- [10] T. Fischer and C. Krauss, "Deep learning with long short-term memory networks for financial market predictions," *Eur. J. Oper. Res.*, vol. 270,no. 2, pp. 654-669, Oct. 2018.

- [11] A. Alahi, K. Goel, V. Ramanathan, A. Robicquet, L. Fei-Fei, and S. Savarese, "Social LSTM: Human trajectory prediction in crowded spaces," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Jun. 2016, pp. 961–971.
- [12] Q. Zhang, H. Wang, J. Dong, G. Zhong, and X. Sun, "Prediction of sea surface temperature using Long Short-term memory," IEEE Geosci. Remote Sens. Lett., vol. 14, no. 10, pp. 1745–1749, Oct. 2017.
- [13] Y. Zhu, X. Fan, J. Wu, X. Liu, J. Shi, and C. Wang, "Predicting ICU mortality by supervised bidirectional LSTM networks," in Proc. 1st Joint Workshop AI Health, Organized Part Federated AI Meeting, 2018, pp. 49–60.
- [14] Zhu, J.; Liu, S.; Fan, N.; Shen, X.; Guo, X. A Short-term power load forecasting Method Based on LSTM Neural Network. China New Telecommun. 2021, 23, 167–168.
- [15] Xue, J.; Shen, B. A novel swarm intelligence optimization approach: sparrow search algorithm. Systems Science & Control Engineering an Open Access Journal. 2020, 8, 22-34.
- [16] Yoon, S.H. Membrane Bioreactor Processes (Advances in Water and Wastewater Transport and Treatment), 1st ed.; CRC Press: Florida, USA, 2020; pp. 55–67.