Received September 29, 2020, accepted October 11, 2020, date of publication October 20, 2020, date of current version October 30, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3032531

Energy and Materials-Saving Management via Deep Learning for Wastewater Treatment Plants

JIANHUI WANG[1,2,3], KEYI WAN[1], XU GAO[1,2], XUHONG CHENG[1], YU SHEN[1,3], ZHENG WEN[4], (Member, IEEE), USMAN TARIQ[5], AND MD. JALIL PIRAN[6], (Member, IEEE)

[1]National Research Base of Intelligent Manufacturing Service, Chongqing Technology and Business University, Chongqing 400067, China
[2]Chongqing Water Assets Management Company, Ltd., Chongqing 400069, China
[3]Chongqing South-to-Thais Environmental Protection Technology Research Institute Company, Ltd., Chongqing 400069, China
[4]School of Fundamental Science and Engineering, Waseda University, Tokyo 169-8050, Japan
[5]College of Computer Engineering and Sciences, Prince Sattam bin Abdulaziz University, Al-Kharj 11942, Saudi Arabia
[6]Computer Engineering Department, Sejong University, Seoul 209, South Korea

Corresponding authors: Yu Shen (shenyu@ctbu.edu.cn) and Md. Jalil Piran (piran@sejong.ac.kr)

This work was supported in part by the National Key Research & Development Program of China under Grant 2016YFE0205600, in part by the Postdoctoral Science Foundation under Grant 2019M653825XB, Grant cstc2019jcyj-bshX0061, and Grant 2019SWZC-bsh001, in part by the Research Project of Chongqing Technology and Business University under Grant ZDPTTD201917 and Grant 1952005, and in part by the Scientific and Technological Research Program of Chongqing Municipal Education Commission under Grant KJQN201800831 and Grant KJZD-M20200080.

ABSTRACT With the increasing public attention on sustainability, conservation of energy and materials has been a general demand for wastewater treatment plants (WWTPs). To meet the demand, efficient optimal management and decision mechanism are expected to reasonably configure resource of energy and materials.In recent years, advanced computational techniques such as neural networks and genetic algorithm provided data-driven solutions to overcome some industrial problems. They work from the perspective of statistical learning, mining invisible latent rules from massive data. This paper proposes energy and materials-saving management via deep learning for WWTPs, using real-world business data of a wastewater treatment plant located in Chongqing, China. Treatment processes are modeled through neural networks, and materials cost that satisfies single indexes can be estimated on this basis. Then, genetic algorithm is selected as the decision scheme to compute overall cost that is able to simultaneously satisfy all the indexes. Empirically, experimental results evaluate that with the proposed management method, total energy and materials cost can be reduced by 10%-15%.

INDEX TERMS Wastewater treatment, energy and material-saving, deep learning, optimal management, genetic algorithm.

I. INTRODUCTION

In recent years, the rapid development of society and economy have led to a proportional increase in wastewater production, bringing growing business pressure wastewater treatment plants (WWTPs) [1]. At the same time, the pollutant discharge standards of WWTPs have been raised for sustainable development [2], [3]. In this context, excessive chemicals are used during the treatment processes by many WWTPs, leading to meaningless waste of energy and materials [4], [5]. In fact, such high consumption of energy and materials go against the thought of sustainability to some extent. Therefore, it is of great significance to reasonably schedule energy consumption amount under the condition of meeting the discharge standard [6], [7].

However, the wastewater treatment process is a highly nonlinear industrial process system. The wastewater parameters of influent are uncertain, and the various biochemical reactions are associated with operating conditions. As a result, the wastewater treatment process is extremely complicated, suffering from lagging management, difficult operations and serious interference [8], [9]. Traditional mechanism-based models are reliable on physics process and biochemical reactions, which lead to difficulties during wastewater treatment process. Owing to the prevalence of sensors and Internet of Things (IoT), a huge amount of business data can be

The associate editor coordinating the review of this manuscript and approving it for publication was Sanjeevikumar Padmanaban.

collected from industrial processes, which can be used to guide industrial decision making from a novel insight [10]. To this end, this research manages to deal with optimal decision of WWTPs with respect to consumption amount of energy and materials, with the aid of data-driven methods [11]. Since useful information can be discovered from large amounts of historical data through robust analysis models [12], [13], they never need to deal with the complex physical process [14]–[16], biochemical reactions and mathematical equations. Therefore, they have been used in many typical industrial scenarios [17], [18].

In this research, data mining was introduced to optimize the specific WWTP, in order to minimize consumption of energy and materials, under the condition of discharge standard. This paper proposes energy and materials-saving management via deep learning for WWTPs. Particularly, a novel mechanism, named principal component analysis-convolutional neural network-long-short-term memory neural network (PCA-CNN-LSTM), was proposed for this purpose. Such hybrid neural network method determines the relationship between input and output, which is variable based on large amount of historical data. And it can be used to predict effluent parameters under new influent conditions. Then, it was reformed by feedback regulation and iterations to calculate optimized energy and materials consumption. Moreover, genetic algorithm (GA) was also introduced to optimize energy and materials consumption under multi-target of effluent parameters. Main contributions of this paper can be summarized as:

1) A new hybrid neural network (PCA-CNN-LSTM) model based on deep neural network was proposed. It was trained by masstive real-world historical data from a WWTP, and it can be used to predict the effluent parameters.
2) The PCA-CNN-LSTM was modified to optimize energy and materials consumption, and genetic algorithms was introduced to reduce the total cost of energy and materials under multi-target of effluent parameters.
3) The efficiency and stability of the proposed PCA-CNN-LSTM model on a real-world dataset was evaluated.

II. OVERVIEW

A. FRAMEWORK

This research put forward a new combined model to optimize the energy consumption and materials consumption of WWTPs. The technology road-map of prediction and optimization process is shown in Figure 1. The system includes prediction process, optimization process and inspection process.

1) The prediction process is to predict the effluent value according to the inflow conditions, energy consumption and materials consumption.

2) After prediction, data is transmitted to the optimization model, and the optimization process is responsible for optimizing the consumed energy and materials of the predicted effluent value.

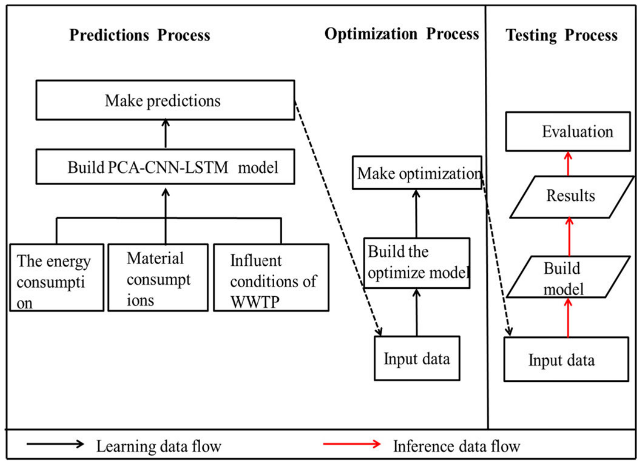


FIGURE 1. Technology road-map of prediction and optimization process.

3) Finally, in the testing process, the predicted results are compared with the real value, and the optimized drug consumption and energy consumption are compared with the data from the WWTP.

B. PROBLEM STATEMENT

This research takes the WWTP in Chongqing as the research object, and the research data content is the daily data of the WWTP from 2018 to 2019, including influent conditions, flow, effluent value, energy consumption, materials consumption, etc. However, due to the lack of separate energy consumption of pretreatment unit, biochemical treatment unit and sludge treatment unit, the data can't be optimized by specific formula. Take the energy consumption as an example, aeration energy is the main energy consumption in WWTPs, and the energy waste can not be negligible. If the aeration or total energy consumption can be optimized, the aeration system can be precisely controlled, and energy waste will be greatly reduced. Science the WWTP is a complex system, a hybrid neural network model is constructed to predict the effluent value. On this basis, the interaction between materials consumption, energy consumption and effluent value is used for subsequent algorithm operation.

C. PRELIMINARIES

The selected WWTP in Chongqing has been operated for 10 years, with an average daily wastewater treatment volume of more than 100,000 cubic meters. With advanced wastewater treatment equipment, the main process of the plant area adopts the Cyclic Activated Sludge Technology (CAST) treatment process, and the effluent water quality is Class 1 A standard according to the "Emission Standard of Pollutants for Urban Wastewater Treatment Plants" (GB18918-2002). The main pollutant indicators and values of WWTP effluent are shown in Table 1.

Firstly, the raw wastewater enters the coarse grid to remove suspended impurities and reduce the chance of blocking the lifting pump of the sump. Then, wastewater will enter the adjustment tank, which is lifted to the fine grid by the

TABLE 1. Pollutant discharge standards for urban WWTPs.

Indexes	COD	BOD	SS	NH_3-N	TN	TP	pH
Conditions of influent wastewater	350	120	90	30	45	2	6-9
Conditions of effluent wastewater	**50**	**10**	**10**	**8**	**15**	**0.5**	**6-9**

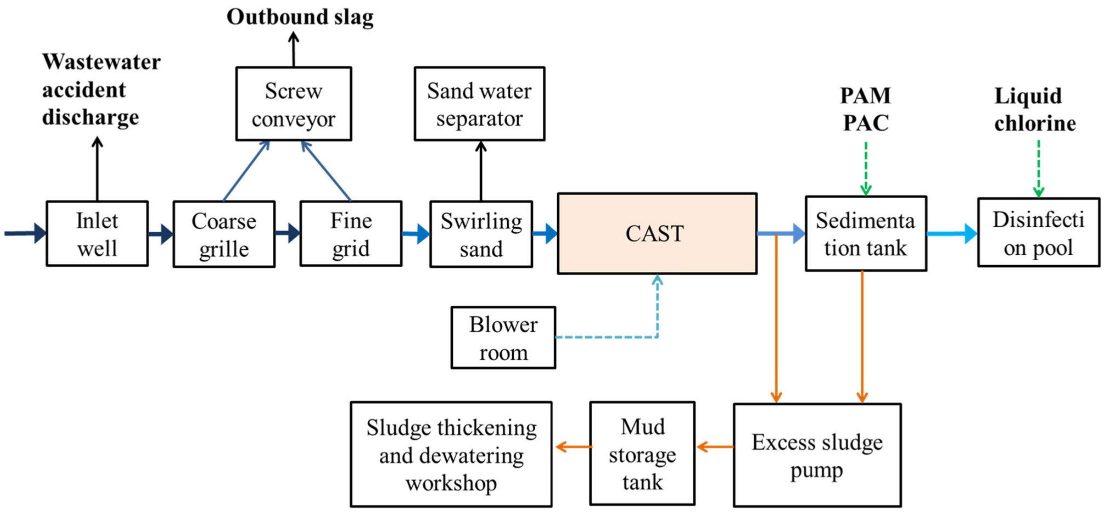


FIGURE 2. CAST process flow chart.

adjustment tank lift pump. After this process, the sand and water are separated.

In the following step, wastewater will get into the CAST biochemical reaction tank, where it undergoes anoxic and aerobic reaction to remove organic matter. Then, it will undergo a series of processes, including nitrification, denitrification treatment of ammonia nitrogen, phosphorus and other substances. Finally, the processes of wastewater treatment end with precipitation and drainage. In general, the CAST process includes 4 stages: wastewater inflow-aeration, wastewater inflow-precipitation, stop wastewater inflow-drainage, and wastewater inflow-idle. The specific process flow is shown in Figure 2.

Based on the principle of biological selection, the CAST process utilizes the release of phosphorus, denitrification, and rapid absorption of organic substrates in the influent water from the biological selection zone in the main reaction, which effectively inhibiting the growth and reproduction of filamentous bacteria and overcoming pollution. The mud expands, thereby enhancing the stability of the system operation. The variable volume operation also improves the adaptability of the system to changes in water quality and quantity and the flexibility of operation [19], [20].

III. METHODOLOGY

A. PCA-CNN-LSTM

1) PCA

The data taken is from the daily data of a WWTP for two years, including daily treatment volume, influent water quality and water volume, effluent water quality, E. coli, energy consumption, materials consumption (iron salt, aluminum salt, desliming flocculant, liquid chlorine, mud volume, etc.). Due to the large number of miscellaneous data, PCA statistical method is adopted in order to retain the influence characteristics of the original variable data and use as few variables as possible [21]. The main purpose of principal component analysis is to use fewer variables to explain the variation of the original variables, while converting many highly correlated variables into uncorrelated variables. Usually the number of selected variables is small, but the data that can fully explain this situation is called the principal component. This type of data processing method is to not only ensure less data loss, but also comprehensively simplify and optimize the data [22]. By reducing the dimensionality of data, PCA has been widely used in the wastewater treatment process to obtain basic information about wastewater treatment.

a: STANDARDIZATION

$$\widetilde{\aleph_{po}} = \frac{\aleph_{po} - \overline{\aleph_o}}{\mu_o} \tag{1}$$

Among them, there are p index variables for PCA, and there are a total of o evaluation indexes. $\overline{\aleph_o}$, μ_o are the sample mean and sample standard deviation of the o-th index.

b: CALCULATION OF THE CORRELATION COEFFICIENT MATRIX

$$R = \left(r_{po}\right)_{n*n} \tag{2}$$

$$r_{po} = \frac{\sum_{l=1}^{p} \widetilde{\aleph_{lp}} \bullet \widetilde{\aleph_{lp}}}{p-1} \tag{3}$$

Among them, $r_{pp} = 1$, and r_{po} is the correlation coefficient between the p-th index and the o-th index.

c: CALCULATION OF THE CHARACTERISTIC VALUE

According to the eigenvalues and eigenvectors of the correlation coefficient matrix in the second step, new index variables are formed according to the eigenvectors.

$$\begin{cases} \emptyset_1 = \cap_{11}\widetilde{p_1} + \cap_{21}\widetilde{p_2} + \ldots \cap_{a1}\widetilde{p_a} \\ \emptyset_2 = \cap_{12}\widetilde{p_1} + \cap_{22}\widetilde{p_2} + \ldots \cap_{a2}\widetilde{p_a} \\ \emptyset_b = \cap_{1b}\widetilde{p_1} + \cap_{2b}\widetilde{p_2} + \ldots \cap_{ab}\widetilde{p_a} \end{cases} \tag{4}$$

In the formula, $\emptyset_1$ is the first principal component, and $\emptyset_2$ is the second principal component.

d: CALCULATION OF THE COMPREHENSIVE EVALUATION VALUE

$$Z = \sum_{o}^{w} g_o y_o \tag{5}$$

$$g_o = \frac{\gamma_o}{\sum_{l=1}^{b} \gamma_l} \tag{6}$$

Among them, g_o is the information contribution rate of the o-th principal component.

2) CNN

CNN is one of the most representative neural networks in the field of deep learning technology. It is a feed forward neural network and has made many breakthroughs in the field of wastewater treatment [23]. It includes convolution calculation and has a deep structure, which can be used for supervised and unsupervised learning. Convolutional neural networks have three basic concepts: local receptive fields, shared weights, and pooling layers.

a: LOCAL RECEPTIVE FIELD

For a general deep neural network, each pixel of the input sequence is usually connected to every fully connected neuron, while the convolutional neural network connects each hidden node only to a certain local area of the sequence [24]. Therefore, the number of parameter training has been reduced.

b: WEIGHT SHARING

In the convolutional layer of CNN, the weights corresponding to neurons are the same, thus, the number of training parameters can be reduced. The shared weights and offsets are also called convolution kernels or filters.

$$a^1 = \sigma\left(b + \sum_{l=0}^{4}\sum_{m=0}^{4} wl, a_{j+l,k+m}\right) \tag{7}$$

In the formula, σ represents the activation function, b is the offset, w is the 5*5 shared weight matrix, and the matrix a represents the neurons in the input layer.

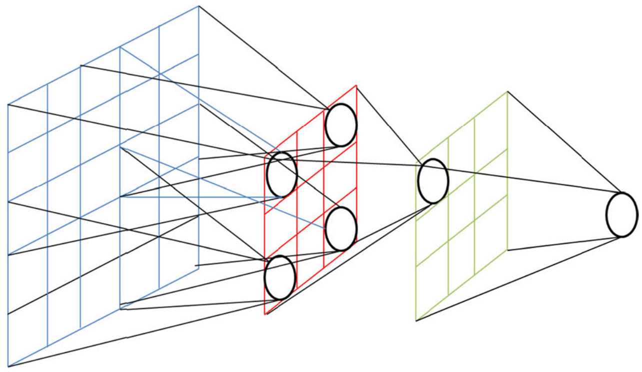


FIGURE 3. Concept map of local receptive field.

c: POOLING LAYER

The function of pooling layer is to reduce the feature map. The merge operation is independent for each depth slice, and the ratio is usually 2 * 2. The result of pooling layer will have a lower dimension, and it is not easy to generate the phenomenon of over fitting [25].

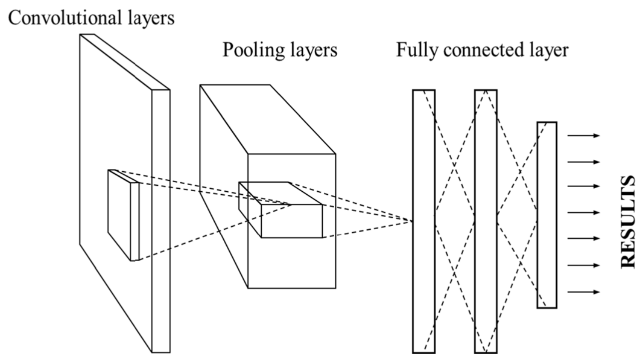


FIGURE 4. Network structure diagram of CNN.

3) LSTM

The LSTM neural network is an improvement of the recurrent neural network (RNN) [26], which is specially designed for the purpose of solving the problem all the time. Compared with ordinary RNN, LSTM performs better in longer sequences, being able to solve the problems that RNN cannot solve in longer training, namely gradient disappearance and gradient explosion. The main difference lies in the operation in the LSTM unit. It is allowed to keep or forget part of the information for the key cell state. The cell state of LSTM is similar to a conveyor belt. There are few linear intersections. Most of them run on a chain, so the information is not easy to change during the transportation. The core concept of LSTM is the unit state, which is actually a variety of gates in the LSTM structure. The unit state can send relevant information on the ''conveyor belt''. As the unit state is operating, information will be added or deleted through it. As for the unit state, the ''gate'' is another different neural network, and its function is to filter some information into the unit state [27], [28]. In the LSTM structure, it is impossible to add or delete information only by relying on the "conveyor belt". Only through the "gate" neural network, the various structures in the LSTM structure have different functions. However, they still interact, transmit and retain or delete information mutually.

In Figure 5, the network structure diagram of LSTM is shown. First, "forget gate" determines the information that should be discarded or retained. The output includes a vector through h_{t-1} and some information in x_t in the structure diagram. The range of this vector is (0, 1). The value of the vector indicates what information is retained or deleted by C_{t-1} in the cell state. Secondly, the "input gate layer" updates the unit state, sending the previous hidden state and current input to the sigmoid function and $\tanh$ function. The sigmoid function determines the value to be updated by converting the value from 0 to 1;$\tanh$ function compresses it into -1 and 1 for network adjustment, and then it multiplies the $\tanh$ output and the S-type output [29]. The sigmoid function outputs a certain message and it is important that $\tanh$ is needed to save production. In this way, there is enough information to calculate the original unit state and obtain the new unit state. Finally, the "output gate" determines the next hidden state. In the first step, the previously hidden function is sent to the Sigmoid function, followed by the current state input, while the new unit state is passed to the $\tanh$ function. The two can be multiplied to determine the relevant information that the unknown state should carry. Finally moving the new unit state and the new unit state to enter the next step [30].

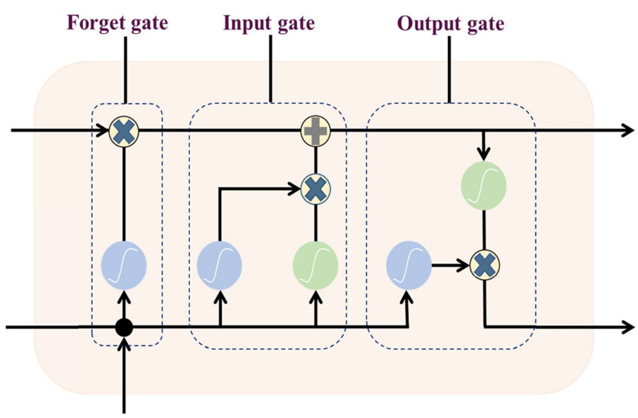


FIGURE 5. Network structure diagram of LSTM.

4) PCA-CNN-LSTM

CNN is essentially a multilayer perceptron. The key to CNN's success lies in its use of local connections and weight sharing, which means the weights of interconnections between neurons are shared by neurons mapped under the same feature. On the one hand, reducing the weight can make the model easier to be optimized; on the other hand, it also improves the operating effect of the model, while reducing the risk of over fitting. CNN's unique weighted fractional network structure makes it more similar to biological neural networks. Due to CNN's unique network structure, the complexity of the network model is reduced. As the number of weights is reduced, the running process is smoother [31]. Since this article studies a WWTP in a certain place, the information is relatively undisclosed, and the basic data is partially missing. CNN can avoid this situation. LSTM can solve long-term problems due to the particularity of the network structure, so it is often used to solve problems related to time series prediction [32]. It's suitable for the daily record data of WWTPs.

Combining CNN and LSTM neural network can process basic data with large defects, making the experiment more efficient. However, related studies have shown that when LSTM is training the model, it will change due to the time step and the number of neurons, and different parameters will affect the superiority of the model [33]. Therefore, the PCA method is introduced to preprocess the original data, simplify and delete the data, while retaining the original characteristics of the original data. First, the PCA method is taken for data preprocessing on the established database, then the CNN network model is established with the keras module to adjust the number and size of convolution kernels according to the actual situation. After that, LSTM recurrent neural network model is added, and the output dimensions are respectively 128, 64, 48. Finally, with two added hidden layers, the number of neurons is 32 and 1, and finally set the optimizer of the model with a learning rate of 0.0001. The training set and the test set are divided according to a ratio of 3:1. The network structure diagram of PCA-CNN-LSTM is shown in Figure 6.

B. OPTIMIZATION

The result of PCA-CNN-LSTM model operation is prediction result. During the specific optimization, it includes selecting the range of four indicators and continuously reducing one of the indicators. The neural network model is used to predict the wastewater output. If the water output meets the requirements, the variable will be continually reduced until the predicted value corresponding to the variable is infinitely close to the value corresponding to the national emission standard. The lower emission value is close to the national emission standard value, leaving extra space for the operational safety of the WWTP. COD is a suitable example.

Steps 1: The influent conditions (including wastewater parameters, energy and materials consumption) are input to the PCA-CNN-LSTM model, and the effluent COD value is predicted. The input energy consumption is marked as 'E'.

Steps 2: The maximum COD value allowed by the effluent discharge standard (COD <= 50 mg/l) is marked as "A". The range of effluent COD under optimal conditions is defined according to the required accuracy, such as effluent COD $\in$ [0.98*A, A].

Steps 3: The predicted effluent COD value is judged whether it belonging to the range or not. (1) If the effluent COD value is less than "0.98*A", the "E" will be reduced by 1% (this value is also choosed by the required accuracy), and the effluent COD value will be predicted and and judged

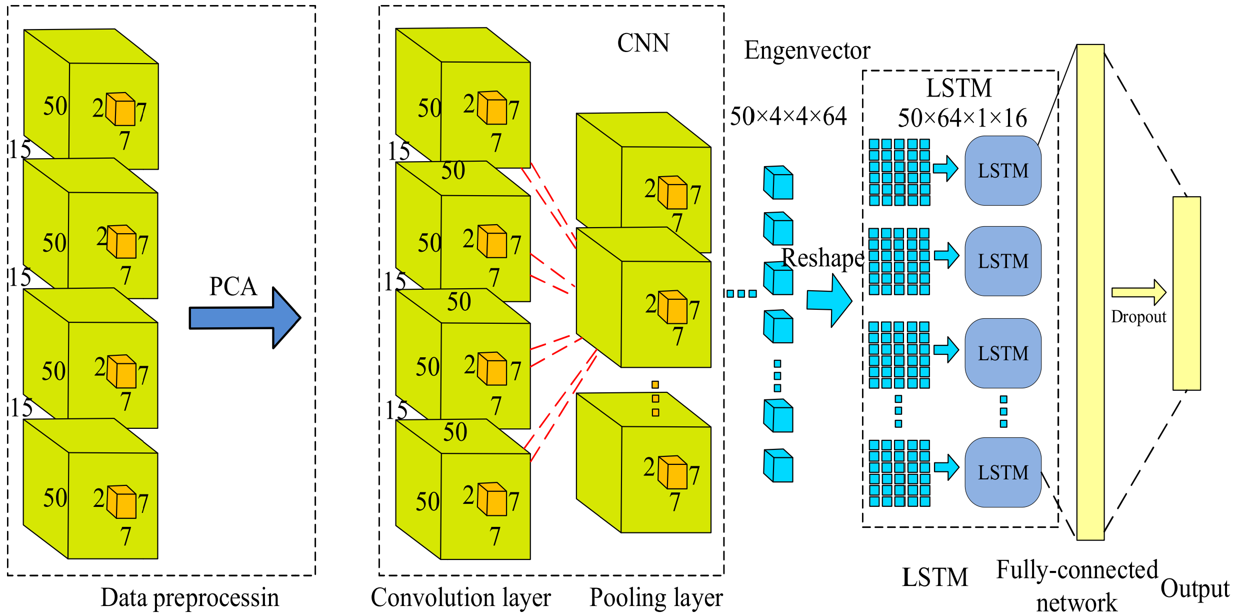


FIGURE 6. Network structure diagram of PCA-CNN-LSTM.

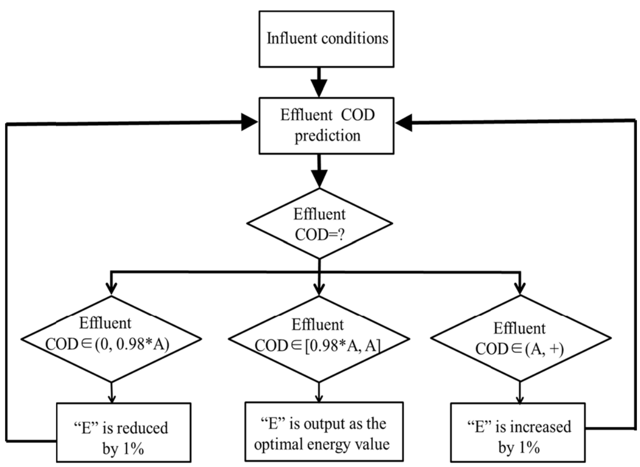


FIGURE 7. The optimization strategy of COD.

again; (2) If the effluent COD value is greater than "A", "E" will be increased by 1%, and the effluent COD value will be predicted and judged again. (3) If the effluent COD value belongs to the defined range [0.98*A, A], the "E" will be considered as the optimal energy value under current influent conditions.

Steps 4: Beginning the loop program to execute Step 3 until the the effluent COD value belongs to the defined range [0.98*A, A].

Steps 5: Return the calculated "E", the end.

The effluent indicators of WWTPs beyonds considered unilateral, therefore, this paper uses genetic algorithm to optimize the effluent indicators as a whole.

Steps 1: Different initial features are combined and discriminant values are calculated respectively.

Steps 2: The two combinations with the largest discriminant value are selected for hybridization, and some different features from each other to obtain a new feature combination.

Steps 3: The discriminant value of the new feature combination is calculated again, and is compared with the original data to select the combination with the largest discriminant value for hybridization.

Steps 4: The end. (1) Given a maximum genetic algebra, the algorithm iteration stops at 'MAXGEN'. (2) Given a lower bound calculation method, when the calculation reaches the required deviation, iteration is ended. (3) When the fitness of the algorithm can no longer be improved, the calculation is ended.

IV. EXPERIMENTS AND ANALYSIS

A. RESULTS OF PREDICTION

In order to comprehensively evaluate performance of the prediction model, the average absolute error (MAE), root mean square eXQrror (RMSE) and average absolute percentage error (MAPE) are used as accuracy indicators in the above training model.

$$MAE\left(\emptyset, \beth\right) = \frac{1}{\vartheta}\sum_{i=1}^{\vartheta}\left|\beth\left(\emptyset_i - \beth_i\right)\right| \tag{8}$$

$$RMSE = \sqrt{\frac{1}{\vartheta}\sum_{i=1}^{1}\left(\beth_i - \beth_n\right)^2} \tag{9}$$

$$MAPE = \frac{1}{\vartheta}\sum_{i=1}^{\vartheta}\frac{\left|\beth_i - \beth_n\right|}{\beth_i} \tag{10}$$

The results of prediction are shown in Table 2. The comparison of the true value and the predicted value of the test set are shown in Figure 9.

Algorithm 1 Optimal Decision Process

Initialization:

P = PCA-CNN-LSTM model
E = Input energy
M = Input material
A = The maximum COD value allowed by the effluent discharge standard (COD <= 50 mg/l)
NUM(S) = Influent wastewater parameters
COD = Effluent COD value

Start:

```
Calculate the COD by PCA-CNN-LSTM model
    P ⇐ E, M, NUM(S)
    COD
while |COD-0.99A| >0.01A
  if COD >= A
    E = E*1.01
    M = M*1.01
    P ⇐ E, M, NUM(S)
    COD
  else if COD <= 0.98*A
    E = E*0.99
    M = M*0.99
    P ⇐ E, M, NUM(S)
    COD
end if
end while
return E
```

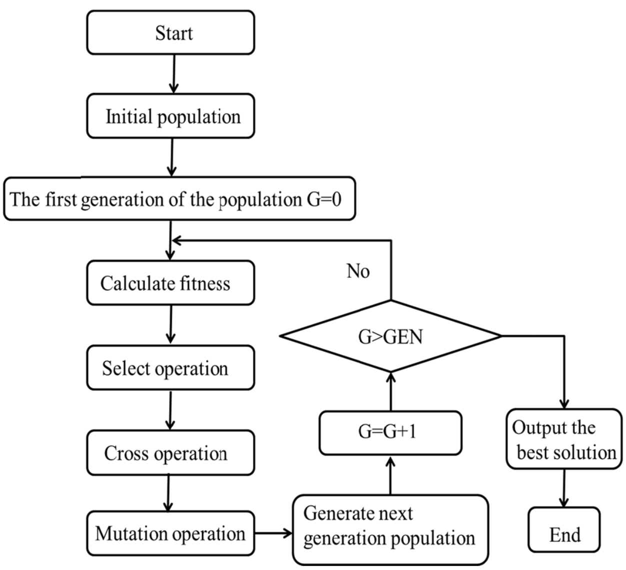


FIGURE 8. Process diagram of genetic algorithm.

B. RESULTS OF OPTIMIZATION

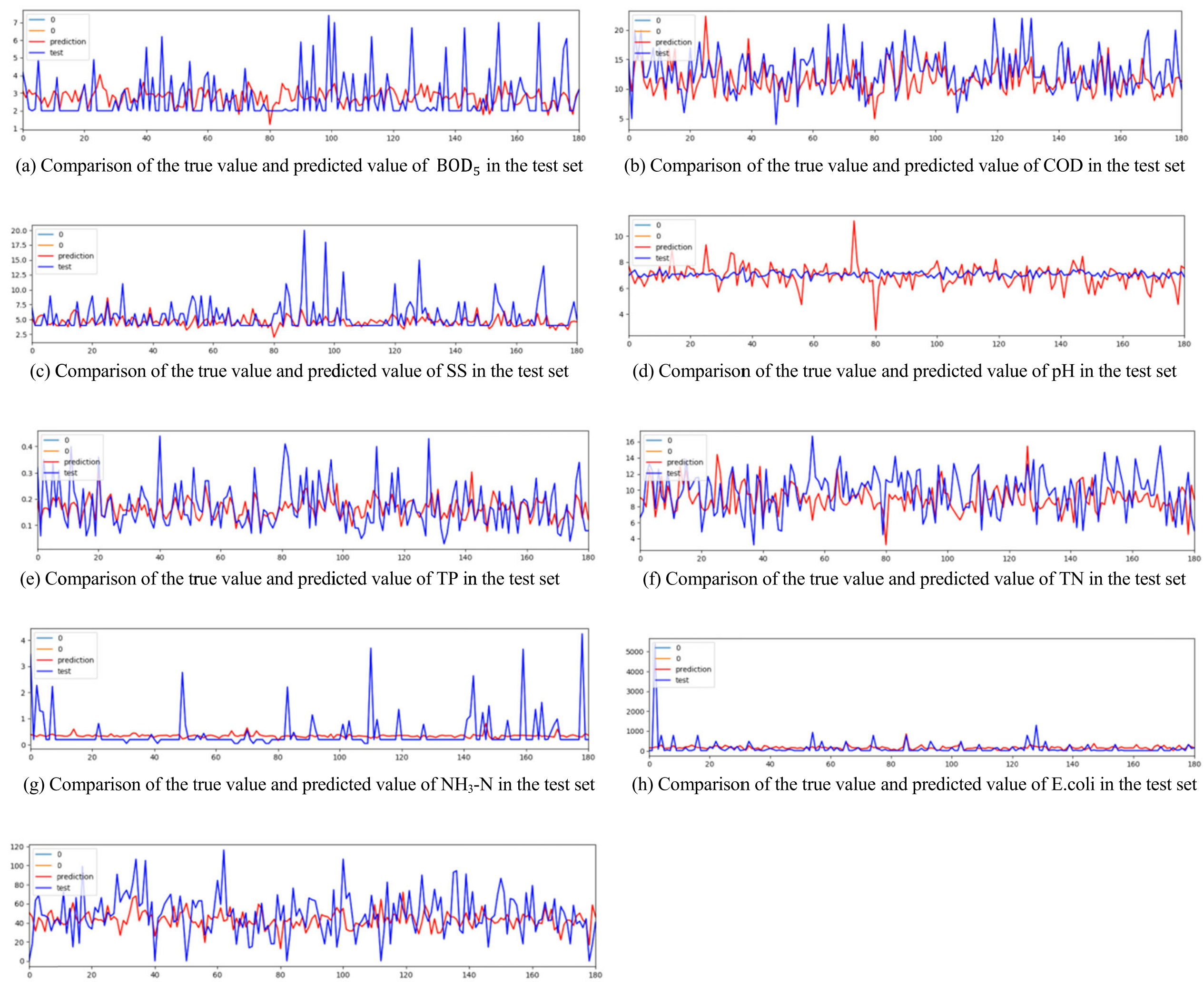
In order to optimize the daily energy consumption and reagents of the WWTP, to continuously reduce the parameter standards, and to ensure that the effluent water quality meets the standard, a single target can be expressed as a function of the control variable. According to the operating conditions, the optimization results can be obtained by the model, shown in Table 3.

TABLE 2. The results of predicion.

Index	MAE	RMSE	MAPE
Effluent BOD_5	0.9141	1.2984	0.3078
Effluent COD	2.7098	3.5454	0.2781
Effluent SS	1.5326	2.4698	0.2859
Effluent pH	0.5989	0.8889	0.0888
Effluent TP	0.0643	0.0829	0.5707
Effluent TN	2.3917	2.9816	0.2762
Effluent NH_3-N	0.3394	0.6784	0.9969
Escherichia coli	8.1697	2.1633	4.218
Mud volume	1.933	2.4207	/

Compared with the daily data of the WWTP, the electricity consumption and the medicament have all been reduced. Through data modeling, the optimal values of different wastewater quality of effluent are studied respectively. However, the actual operation of effluent from WWTP could not be viewed as a single index. Therefore, genetic algorithm (GA) is used to optimize it. Genetic algorithm is a method to search for the optimal solution by simulating the natural evolution process. The algorithm is simple and widely applied [34]. Because the overall search strategy and optimized search method of genetic algorithm do not need to rely on gradient information or other auxiliary knowledge when calculating, genetic algorithm provides a general framework for solving complex system problems. It does not depend on the specific area of the problem, and has strong robustness to problem. The main parameters of genetic algorithm are weights and thresholds. The parameters of this paper are set as follows: population size is 100, maximum genetic algebra is 100, learning rate is 0.1, and maximum training times is 100. The results of the overall optimization are shown in Table 4. The daily data of seven days from the WWTPW were randomly selected and the following comparisons between the electricity and the amount of chemicals were made. The results are shown in the Figure 10. The efficiency and stability of the proposed PCA-CNN-LSTM model on a real-world dataset was evaluated, and as shown in Figure 10. By comparing the differences of energy and material consumption before and after optimization, the experimental results can be estimated that the total energy and materials cost was reduced by 10%-15%.

As shown in the above figures, in the selected data, the optimized data is largely lower than the original data. This further shows that the electricity and chemical consumption have been optimized while ensuring that the WWTP's effluent meets the standard.



(a) Comparison of the true value and predicted value of BOD_5 in the test set

(b) Comparison of the true value and predicted value of COD in the test set

(c) Comparison of the true value and predicted value of SS in the test set

(d) Comparison of the true value and predicted value of pH in the test set

(e) Comparison of the true value and predicted value of TP in the test set

(f) Comparison of the true value and predicted value of TN in the test set

(g) Comparison of the true value and predicted value of NH_3-N in the test set

(h) Comparison of the true value and predicted value of E.coli in the test set

(i) Comparison of the true value and predicted value of Mud volume in the test set

FIGURE 9. Comparison of the true value and the predicted value of the test set.

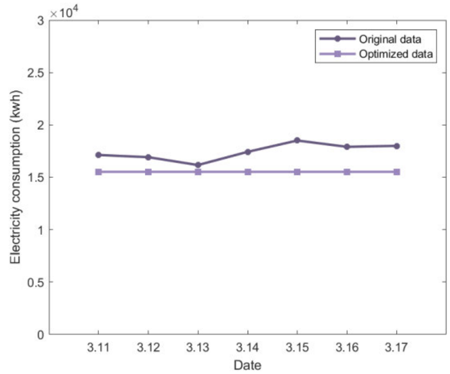
TABLE 3. Optimization result table.

	electricity consumption (O_1)KW*h	ferric salts (O_2)kg	Aluminium salt (O_3)kg	Desilt flocculant (O_4)kg	High efficiency flocculant (O_5)kg	Liquid chlorine (O_6)kg	sodium hypochlorite (O_7)kg
Effluent BOD_5=9.585mg/L	14900	2000	9980	49	30	398	1970
Effluent COD=49.362 mg/L	20650	2350	920	32	23	302	2350
Effluent SS=9.974 mg/L	9840	830	785	33	30	403	1150
Effluent pH=6.827	10820	970	700	15	28	250	1250
Effluent TP=0.497 mg/L	12730	600	980	18	28	370	1050
Effluent TN=14.902 mg/L	10360	1130	1020	18	44	340	1200
Effluent NH_3-N=7.789 mg/L	17690	1740	880	38	26	330	1550
Escherichia coli=987 /L	11600	1550	820	40	30	314	1020

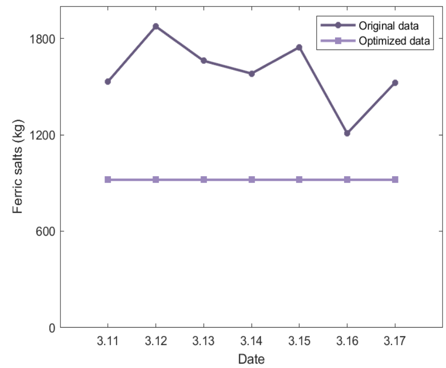
C. OPTIMAL OPERATION PLAN

The energy consumption of the WWTP is mainly used in the treatment promotion of wastewater, the feed and return of biological treatment, and the stabilization and treatment of sludge.

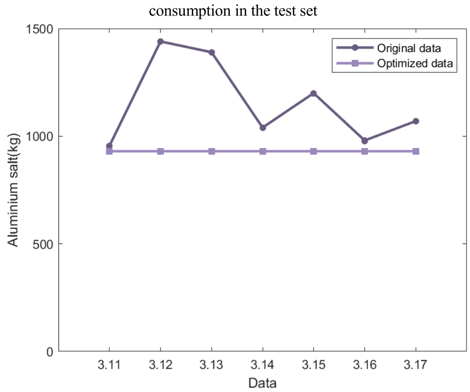
Through the optimization method, we obtained the electric quantity and required drug consumption while meeting the standard. According to the results, we put the drug into the wastewater to make it meet the standard. Therefore, the following points are proposed:



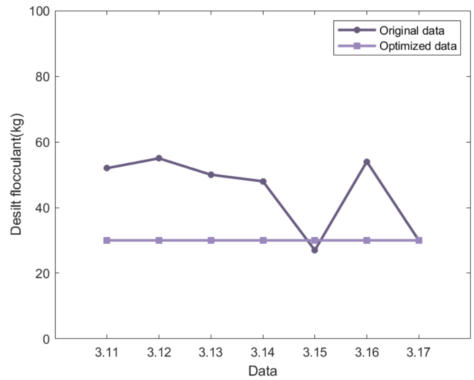
(a) Comparison of the true value and predicted value of electricity consumption in the test set



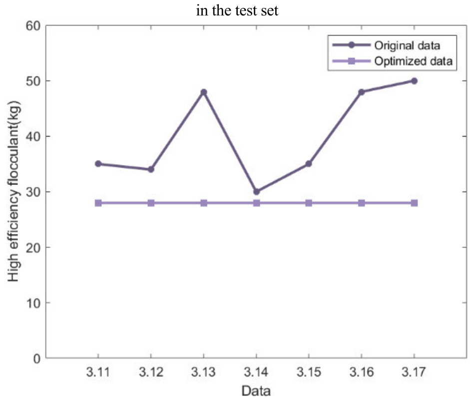
(b) Comparison of the true value and predicted value of ferric salts in the test set



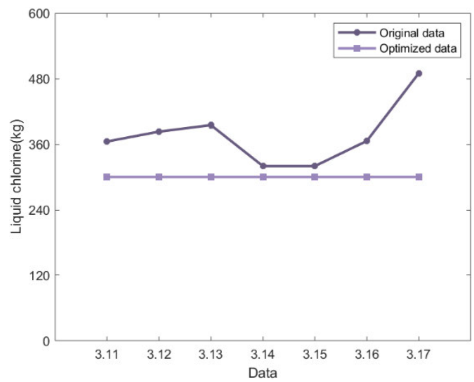
(c) Comparison of the true value and predicted value of aluminium salt in the test set



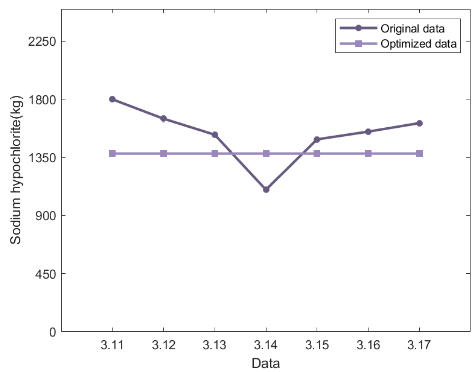
(d) Comparison of the true value and predicted value of desilt flocculant in the test set



(e) Comparison of the true value and predicted value of high efficiency flocculant in the test set



(f) Comparison of the true value and predicted value of liquid chlorine in the test set



(g) Comparison of the true value and predicted value of high efficiency flocculant in the test set

FIGURE 10. Comparison of the true value and the predicted value of the test set.

TABLE 4. Overall optimization result table.

	electricity consumption (O_1)KW*h	ferric salts (O_2)kg	Aluminium salt (O_3)kg	Desilt flocculant (O_4)kg	High efficiency flocculant (O_5)kg	Liquid chlorine (O_6)kg	sodium hypochlorite (O_7)kg
Effluent BOD_5=9.49 mg/L Effluent COD=49.3 mg/L Effluent SS=9.83 mg/L Effluent pH=6.24 Effluent TP=0.495 mg/L Effluent TN=14.89 mg/L Effluent NH_3-N=7.85 mg/L Escherichia coli=965 /L	15520	920	930	30	28	300	1380

(1) Dose of drugs should be accurate and multiple energy sources should complement each other. Moreover, refined management should be ensured and operating costs should be reduced.

(2) Efficiency of wastewater equipment should be improved and the energy consumption of the wastewater lifting system and the supply system should be reduced.

V. CONCLUSION

The wastewater treatment system is a highly non-linear industrial process control system, which is uncertain and time-varying. Meanwhile, the mathematical mechanism model of wastewater treatment process is difficult to be used accurately and duly. Current conservative operation mode in most WWTPs has high operation costs and wastes huge energy and material. In this paper, a new hybrid neural network (PCA-CNN-LSTM) model based on deep neural network was proposed. It was trained by massive real-world historical data from a WWTP, and it can be used to predict the effluent parameters. Then the PCA-CNN-LSTM was modified to optimize energy and materials consumption. Last genetic algorithms was introduced to reduce the total cost of energy and materials under multi-target of effluent parameters. The efficiency and stability of the proposed PCA-CNN-LSTM model on a real-world dataset was evaluated, and experimental results evaluate that the total energy and materials cost was reduced by 10%-15%. The proposed energy and materials-saving management method not only is a new solution for wastewater treatment, but also brings benefits to the development of the economy and society.

REFERENCES

[1] Y. Qiu, J. Li, X. Huang, and H. Shi, ''A feasible data-driven mining system to optimize wastewater treatment process design and operation,'' *Water*, vol. 10, no. 10, p. 1342, Sep. 2018.

[2] K. Smith and S. Liu, ''Energy for conventional water supply and wastewater treatment in urban China: A review,'' *Global Challenges*, vol. 1, no. 5, Aug. 2017, Art. no. 1600016.

[3] Y. Qiu, H.-C. Shi, and M. He, ''Nitrogen and phosphorous removal in municipal wastewater treatment plants in China: A review,'' *Int. J. Chem. Eng.*, vol. 2010, Jan. 2010, Art. no. 914110.

[4] D. Panepinto, S. Fiore, M. Zappone, G. Genon, and L. Meucci, ''Evaluation of the energy efficiency of a large wastewater treatment plant in italy,'' *Appl. Energy*, vol. 161, pp. 404–411, Jan. 2016.

[5] O. Nowak, P. Enderle, and P. Varbanov, ''Ways to optimize the energy balance of municipal wastewater systems: Lessons learned from austrian applications,'' *J. Cleaner Prod.*, vol. 88, pp. 125–131, Feb. 2015.

[6] K. Mizuta and M. Shimada, ''Benchmarking energy consumption in municipal wastewater treatment plants in Japan,'' *Water Sci. Technol.*, vol. 62, no. 10, pp. 2256–2262, Nov. 2010.

[7] K. Yu, M. Arifuzzaman, Z. Wen, D. Zhang, and T. Sato, ''A key management scheme for secure communications of information centric advanced metering infrastructure in smart grid,'' *IEEE Trans. Instrum. Meas.*, vol. 64, no. 8, pp. 2072–2085, Aug. 2015.

[8] M. Maktabifard, E. Zaborowska, and J. Makinia, ''Achieving energy neutrality in wastewater treatment plants through energy savings and enhancing renewable energy production,'' *Rev. Environ. Sci. Bio/Technol.*, vol. 17, no. 4, pp. 655–689, Dec. 2018.

[9] C. W. K. Chow, J. Liu, J. Li, N. Swain, K. Reid, and C. P. Saint, ''Development of smart data analytics tools to support wastewater treatment plant operation,'' *Chemometric Intell. Lab. Syst.*, vol. 177, pp. 140–150, Jun. 2018.

[10] A. M. Anter, D. Gupta, and O. Castillo, ''A novel parameter estimation in dynamic model via fuzzy swarm intelligence and chaos theory for faults in wastewater treatment plant,'' *Soft Comput.*, vol. 24, no. 1, pp. 111–129, Jan. 2020.

[11] W. Xiupeng, ''Modeling and optimization of wastewater treatment process with a data-driven approach,'' Ph.D. dissertation, Gradworks, 2013, vol. 17.

[12] F. Tao, Q. Qi, A. Liu, and A. Kusiak, ''Data-driven smart manufacturing,'' *J. Manuf. Syst.*, vol. 48, pp. 157–169, Jul. 2018.

[13] S. Arridge, P. Maass, O. Öktem, and C.-B. Schönlieb, ''Solving inverse problems using data-driven models,'' *Acta Numerica*, vol. 28, pp. 1–174, May 2019.

[14] H. David and N. Adams, ''Data mining,'' in *Wiley StatsRef, Statistics Reference Online*, 2014, pp. 1–7.

[15] A. Verma, X. Wei, and A. Kusiak, ''Predicting the total suspended solids in wastewater: A data-mining approach,'' *Eng. Appl. Artif. Intell.*, vol. 26, no. 4, pp. 1366–1372, Apr. 2013.

[16] A. Asadi, A. Verma, K. Yang, and B. Mejabi, ''Wastewater treatment aeration process optimization: A data mining approach,'' *J. Environ. Manage.*, vol. 203, pp. 630–639, Dec. 2017.

[17] Z. Zhang, A. Kusiak, Y. Zeng, and X. Wei, ''Modeling and optimization of a wastewater pumping system with data-mining methods,'' *Appl. Energy*, vol. 164, pp. 303–311, Feb. 2016.

[18] C. Fan, F. Xiao, Z. Li, and J. Wang, ''Unsupervised data analytics in mining big building operational data for energy efficiency enhancement: A review,'' *Energy Buildings*, vol. 159, pp. 296–308, Jan. 2018.

[19] D. Torregrossa, J. Hansen, F. Hernández-Sancho, A. Cornelissen, G. Schutz, and U. Leopold, ''A data-driven methodology to support pump performance analysis and energy efficiency optimization in waste water treatment plants,'' *Appl. Energy*, vol. 208, pp. 1430–1440, Dec. 2017.

[20] Y. Yonghong, "Design and implementation of CASS process automatic control system based on PLC," in *Proc. 5th Int. Symp. Comput.*, 2018, pp. 75–79.
[21] F. Gao and X. Zhang, "Simulating a cyclic activated sludge system by employing a modified ASM3 model for wastewater treatment," *Bioprocess Biosyst. Eng.*, vol. 40, no. 6, pp. 877–897, 2017.
[22] W. Piao, C. Kim, S. Cho, H. Kim, M. Kim, and Y. Kim, "Development of a protocol to optimize electric power consumption and life cycle environmental impacts for operation of wastewater treatment plant," *Environ. Sci. Pollut. Res.*, vol. 23, no. 24, pp. 25451–25466, 2016.
[23] S. Wold, K. Esbensen, and P. Geladi, "Principal component analysis," *Chemometrics Intell. Lab. Syst.*, vol. 2, nos. 1–3, pp. 37–52, 1987.
[24] X. Ouyang, K. Gu, and P. Zhou, "Spatial pyramid pooling mechanism in 3D convolutional network for sentence-level classification," *IEEE/ACM Trans. Audio, Speech, Language Process.*, vol. 26, no. 11, pp. 2167–2179, Nov. 2018.
[25] Z. Hu, Y. Zhang, Y. Zhao, M. Xie, J. Zhong, Z. Tu, and J. Liu, "A water quality prediction method based on the deep LSTM network considering correlation in smart mariculture," *Sensors*, vol. 19, no. 6, p. 1420, Mar. 2019.
[26] R. Girshick, J. Donahue, T. Darrell, and J. Malik, "Region-based convolutional networks for accurate object detection and segmentation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 38, no. 1, pp. 142–158, Jan. 2016.
[27] Z. Guo, Y. Shen, A. K. Bashir, M. Imran, N. Kumar, D. Zhang, and K. Yu, "Robust spammer detection using collaborative neural network in Internet of Thing applications," *Internet Things J.*, early access, Jun. 19, 2020, doi: 10.1109/JIOT.2020.3003802.
[28] F. Karim, S. Majumdar, H. Darabi, and S. Chen, "LSTM fully convolutional networks for time series classification," *IEEE Access*, vol. 6, pp. 1662–1669, 2018.
[29] K. Greff, R. K. Srivastava, J. Koutník, B. R. Steunebrink, and J. Schmidhuber, "LSTM: A search space odyssey," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 28, no. 10, pp. 2222–2232, Oct. 2017.
[30] T. Zhang, S. Song, S. Li, L. Ma, S. Pan, and L. Han, "Research on gas concentration prediction models based on LSTM multidimensional time series," *Energies*, vol. 12, no. 1, p. 161, Jan. 2019.
[31] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards real-time object detection with region proposal networks," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 6, pp. 1137–1149, Jun. 2017.
[32] B. Mamandipoor, M. Majd, S. Sheikhalishahi, C. Modena, and V. Osmani, "Monitoring and detecting faults in wastewater treatment plants using deep learning," *Environ. Monitor. Assessment*, vol. 192, no. 2, p. 148, Feb. 2020.
[33] Z. Zhuang, Z. Sun, Y. Cheng, R. Yao, and W. Zhang, "Modeling and optimization of paper-making wastewater treatment based on reinforcement learning," in *Proc. 37th Chin. Control Conf. (CCC)*, Jul. 2018, pp. 8342–8346.
[34] D. Whitley, "A genetic algorithm tutorial," *Statist. Comput.*, vol. 4, no. 2, pp. 65–85, Jun. 1994.

JIANHUI WANG was born in 1989. He received the Ph.D. degree in environmental science and engineering from the Faculty of Urban Construction and Environmental Engineering, Chongqing University, in 2017. He is currently holding a post-doctoral position with the Chongqing Water Asset Management Company, Ltd., and is also an Assistant Researcher with the National Research Base of Intelligent Manufacturing Service, Chongqing Technology and Business University. His main research areas include water pollution control technology and intelligent sewage treatment technology. In recent years, he has presided over one project of the China Postdoctoral Science Foundation; two projects/subjects of the Chongqing Municipal Science and Technology Bureau; two projects of the Chongqing Municipal Education Commission; and three scientific research projects of the Chongqing Technology and Business University. He has applied for more than 20 national invention patents, out of which five patents have been authorized and two software copyrights have been registered. He has published more than 10 academic papers in journals such as environmental pollution, analytical methods, environmental science, and pollution research.

KEYI WAN received the B.S. degree from the Chongqing Technology and Business University, in 2018. She is currently pursuing the master's degree in environmental engineering. She has participated in the application of seven patents and has also participated in one National Natural Science Foundation of China; one project of the Chongqing Science and Technology Bureau; and one scientific research project of the Chongqing Technology and Business University. Her research interest is the smart water management systems.

XU GAO received the Ph.D. degree in municipal engineering from Chongqing University, in 2002. He is currently a Professor with the Chongqing Technology and Business University, and a Doctoral Supervisor with the Chongqing Water Asset Management Company, Ltd. He is mainly engaged in the research and development of advanced environmental protection equipment and industrial water security technology. He has presided over eight national projects/tasks under the National Science and Technology Support Program and the National High-tech Research and Development Program, won three-second prizes at provincial and ministerial levels for scientific and technological progress, published more than fifty research papers, and has obtained more than twenty invention patents.

XUHONG CHENG received the B.E. degree from Chongqing Three Gorges University, in 2019. She is currently pursuing the M.E. degree with the Chongqing Technology and Business University. Her main research interests are related to sewage quality monitoring based on data analysis. She has participated in the application of seven patents, two projects of Chongqing Science and Technology, and one project of the Chongqing Municipal Education Commission.

YU SHEN received the B.S. degree from the Huazhong Agricultural University, in 2004, and the Ph.D. degree from the Dalian University of Technology, in 2009. He is a Researcher and Master's Tutor (management science and engineering, environmental science and engineering) with the Chongqing Technology and Business University and University of Chinese Academy of Sciences Graduate School. He once served as the Deputy Secretary-General of the Pharmaceutical Wastewater Treatment and Resource Innovation Industry Alliance of the Ministry of Environmental Protection (2014), and a member of the Chongqing Municipal Sewage Control Project Expert Group (2015). His main research interests are in environmental Big Data management, environmental equipment design, and operation management. In recent years, he has presided over nine projects/tasks at the provincial and ministerial levels or above, such as those sponsored by the Ministry of Science and Technology and Chongqing Science and Technology Bureau, with a cumulative research fund of more than 20 million Yuan. He has published more than forty papers in international journals and has obtained six authorized patents.

ZHENG WEN (Member, IEEE) received the B.E. degree in computer science and technology from Wuhan University, China, in 2009, and the M.Sc. and Ph.D. degrees from Waseda University, Tokyo, Japan, in 2015 and 2019, respectively. He became a Research Associate with Waseda University, in 2018, where he is currently an Assistant Professor (Lecturer) with the Department of Communication and Computer Engineering. His research interests include ICN/CCN for next-generation communication systems, AI, and the IoT.

USMAN TARIQ received the Ph.D. degree from Ajou University, South Korea. He is currently an Associate Professor with the College of Computer Engineering and Sciences, PSAU. He has led the design of a global data infrastructure simulator modeling, to evaluate the impact of competing architectures on the performance, availability, and reliability of the systems for Industrial IoT infrastructure. His international collaborations/collaborators include NYIT, Ajou University, PSU, University of Sherbrooke, COMSATS, NUST, UET, National Security Research Institute (NSR), Embry-Riddle Aeronautical University, Korea University, University of Bremen, and the Virginia Commonwealth University. His current interests are in applied cybersecurity, advanced topics in the Internet of Things, and health informatics. His research interests include large complex networks, which include network algorithms, stochastic networks, network information theory, and large-scale statistical inference. As a Network Security Theorist, his contributions towards addressing these challenges involve designing better wireless access networks and processing social data and scalable algorithms that can operate in a data center like facility.

MD. JALIL PIRAN (Member, IEEE) received the Ph.D. degree in electronics and radio engineering from Kyung Hee University, South Korea, in 2016. He was a Postdoctoral Research Fellow in resource management and quality of experience in 5G cellular networks and the Internet of Things (IoT) with the Networking Laboratory, Kyung Hee University. He is currently an Assistant Professor with the Department of Computer Science and Engineering, Sejong University, Seoul, South Korea. He has published a substantial number of technical papers in well-known international journals and conferences in research fields of resource allocation and management in 5G mobile and wireless communication, HetNet, IoT, multimedia communication, streaming, adaptation and QoE, as well as cognitive radio networks. In the worldwide communities, he has been an Active Delegate from South Korea in the Moving Picture Experts Group since 2013 and an Active Member of the International Association of Advanced Materials since 2017. He received the IAAM Scientist Medal of the year 2017 for notable and outstanding research in new age technology and innovation, Stockholm, Sweden. He has been recognized as the Outstanding Emerging Researcher by the Iranian Ministry of Science, Technology, and Research in 2017. His Ph.D. dissertation has been selected as the Dissertation of the Year 2016 by the Iranian Academic Center for Education, Culture, and Research in electrical and communications engineering.