Pol. J. Environ. Stud. Vol. 30, No. 5 (2021), 4267-4276
DOI: *10.15244/pjoes/132821* ONLINE PUBLICATION DATE: 2021-06-29

Original Research

Influent Quality and Quantity Prediction in Wastewater Treatment Plant: Model Construction and Evaluation

Rui Wang[1, 2, 3], Zhicheng Pan[4], Yangwu Chen[2, 3], Zhouliang Tan[2, 3]*, Jianqiang Zhang[1]

[1]Faculty of Geosciences and Environmental Engineering, Southwest Jiaotong University, 611756 Chengdu, China
[2]Key Laboratory of Environmental and Applied Microbiology, Chengdu Institute of Biology, Chinese Academy of Sciences, 610041 Chengdu, China
[3]Environmental Microbiology Key Laboratory of Sichuan Province, Chengdu Institute of Biology, Chinese Academy of Sciences, 610041 Chengdu, China
[4]School of Environment, Tsinghua University,100091 Beijing, China

Received: 3 December 2020
Accepted: 25 January 2021

Abstract

Influent quality and quantity were important factors that caused the abnormal operation of WWTP. In this study, the prediction models of influent quality and quantity were established based on four machine learning methods of Linear Regression, Ridge Regression, ElasticNet Regression and Lasso Regression. The meteorological conditions (precipitation and air temperature) and influent indicators (influent quantity, COD, and NH_3-N) were used as training data. The influent quantity prediction of the models were evaluated using the historical data obtained from a WWTP located in western China, and the results showed that the normal rates of influent quantity were ranged from 98.9%-100%. The highest accuracy was obtained with Ridge method which was 86.19% .

For influent quality (COD) prediction, Ridge method is relatively ideal, with 82% accuracy. For influent quality (NH_3-N) prediction, because of higher data normality rates, Lasso and ElasticNet method were more ideal, both with 74% accuracy. Further, in view of the reason of low prediction accuracy, this paper puts forward the idea of model improvement from the three directions of data fluctuation, correlation and amount. It is expected that this study will provide reference for similar research and provide a reference and thought for similar research.

Keywords: influent quality and quantity, machine learning, prediction models, air temperature, precipitation

*e-mail: 46896545@qq.com

Introduction

The activated sludge process was invented at Lawrence Sewage Test Station in Manchester in 1913, and it has become the most widely used method in municipal wastewater treatment plant (WWTP). WWTP is a complex biological system, and is usually regarded as a typical black box [1, 2]. It is known that influent fluctuation would exert an adverse effect on the performance of WWTP [3, 4]. According to 335 surveys conducted by Love N.G., a professor from Virginia Tech University, 90% of the sewage plants surveyed in the United States experienced abnormal biological treatment of wastewater. Seventy percent of wastewater treatment plants failed due to the influent fluctuation [5]. O 'Brien G J et al. pointed out that when influent and environmental conditions were suddenly changed (due to the impact of toxic substances or a sharp change in pH, etc.) the performance of wastewater treatment will be greatly affected [6,7].

For a long time, to meet the criterion for sewage discharge, operators rely mostly on their own experience and data (such as COD, ammonia, influent flow, etc.) obtained from online detection equipment to regulate the WWTP. However, during sudden abnormal influent, there is often hysteresis by means of manual controls, which makes it difficult to deal with the emergency in a timely manner [8, 9]. Faced with more stringent sewage discharge standards, this undoubtedly aggravates the potential risks and uncertainty in the sewage treatment operation process.

There is some regularity of influent quality and quantity from WWTP. For example, the concentration of influent pollutants is relatively high in winter and lower in summer, while the influent quantity is the opposite [10]. In order to grasp the dynamic changes of influent quality and quantity of WWTP in real time, the machine learning, mathematical modeling and other forecasting methods are increasingly being used to assist WWTP operation and management. For instance, influent prediction models were built with autoregressive integrated moving average (ARIMA), nonlinear autoregressive network (NAR) and support vector machine (SVM) regression time series algorithms, and the prediction effect was evaluated using the historical influent flow data of a sewage treatment plant (STP) as training data [11]. Hamid Zare Abyaneh studied the predictive effects of multiple linear regression (MLR) and artificial neural network (ANN) models on two major wastewater parameters for a sewage treatment plant. The correlation coefficient (R), root mean square error (RMSE) and bias values were used to evaluate the performance of the neural network model [12].

However, it should be pointed out that current researches were mainly focused on predicting either influent quantity or quality [13-16], while fewer studies concerned about influent quantity and quality prediction simultaneously, and there is still a lack of systematic research on dynamics and correlation between influent quantity and quality. And there was also no separation of rain and sewage in many countries, and influent quality and quantity are easily affected by precipitation, air temperature and other factors [17, 18]. Most studies only used historical WWTP influent quantity and (or) quality data for machine learning training data and did not consider the influence of external factors such as weather and temperature on influent [11, 19]. The models in previous studies have used historical influent quality and quantity as the only reference, ignoring the effects of meteorological factors such as precipitation, and have been built only based on present temporal fluctuations in influent quality and quantity with a higher accuracy, which is hard to predict the influent quality and quantity under the unconventional meteorological conditions such as heavy rainfall, continuous high temperature and drought weather, inducing to the limitations for actual applications.

Based on these issues, a more comprehensive model that takes into account more possible factors is needed to facilitate the auxiliary operation of sewage treatment facilities and effectively improve the warning ability of WWTP risk under special meteorological conditions. In this study, four linear regression methods (Linear Regression, Ridge Regression, ElasticNet Regression and Lasso Regression) were used to construct influent prediction models. To further improve the applicability of the models for influent quality and quantity prediction in WWTP, besides influent indicators (influent flow, COD, NH_3-N), local precipitation and air temperature that may affect the influent quality were also taken into account for model building and prediction evaluation. The results of this paper are more applicable to the actual operation management requirements of auxiliary operation of sewage treatment facilities, and also lay a foundation for the development of intelligent sewage treatment plants.

Materials and Methods

Determination of COD and NH_3-N

COD was determined by "Water quality-determination of the chemical oxygen demand-dichromate method" (HJ 828-2017) [20]. Ammonia nitrogen was determined by "Water quality – determination of ammonia nitrogen – Nessler's reagent spectrophotometry" (HJ 535-2009) [21].

Data Acquisition and Classification

This study selected the Gongxian WWTP (104°43′52.09″E,28°28′55.86″N), in Yibin City of Sichuan Province, China as the research area, with a capacity of 20000 m^3/d, used Anaerobic-Anoxic-Oxic (AAO) process. The five main record indicators include precipitation, temperature, influent flow, COD, and ammonia. Data were obtained for January 1,

2015 December 31, 2018. January 1 of the year was recorded as day 1, December 31 was day 366 (2016), and there were 365 days each year (2015, 2017, 2018). The WWTP location was shown in Fig. 1.

Analysis of Temperature and Precipitation

According to Fig. 2 and 3, there was little change on the local temperature and precipitation of Gongxian county from 2015 to 2018. higher temperature (>30ºC) days appeared in 2016 and 2017, while lower temperature (<10ºC) days happened in 2018. Meanwhile, there was no record of temperature below 0ºC in successive years. The local precipitation was mostly below 10 mm (24-hour precipitation), and more than 315 days (86%) were within this range during 2015-2018. Besides, the occurrence of rainstorm (the rainfall is equal to or greater than 50 mm in 24 hours), according to the Chinese meteorological department [22] is rare, with no more than seven days in consecutive years from 2015 to 2018.

Analysis of Influent Quality and Quantity

As shown in Figs 4-6, there was no drastic change for influent quantity during 2015-2018, and the average influent flow from 2015-2016 was slightly higher than that from 2017-2018, and the highest one-day inflow occurred in 2018. From 2015 to 2017, the average influent COD value was stable. However, it decreased by 16.6% in 2018. The highest value in 2018 was higher than any of the previous years, and the lowest COD value was much lower, indicating a significant fluctuation of influent indicators in 2018 compared with other years. The average values of influent NH_3-N were basically stable from 2015 to 2018, while the maximum increased. Compared with 2015, the maximum NH_3-N concentration in 2018 was three times of that of 2015,

Fig. 1. The location of Gongxian WWTP.

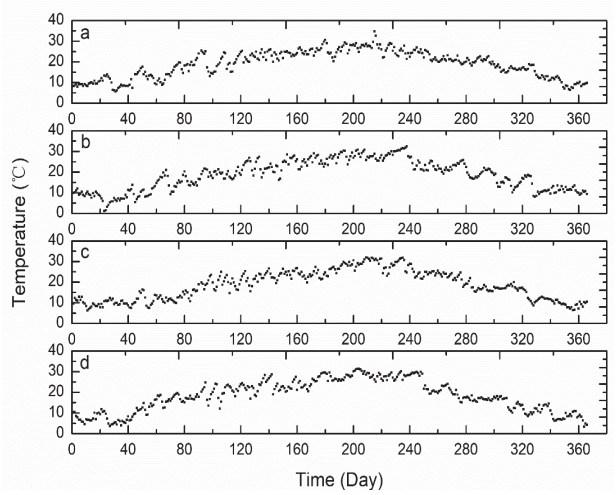

Fig. 2. Temperature variability chart of Gongxian County during 2015 a) -2018 d).

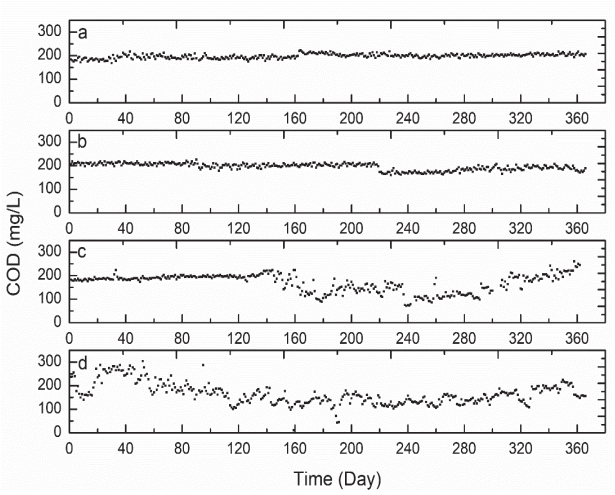

Fig. 5. The influent COD of Gongxian WWTP during 2015-2018: a) 2015, b) 2016, c) 2017, d) 2018.

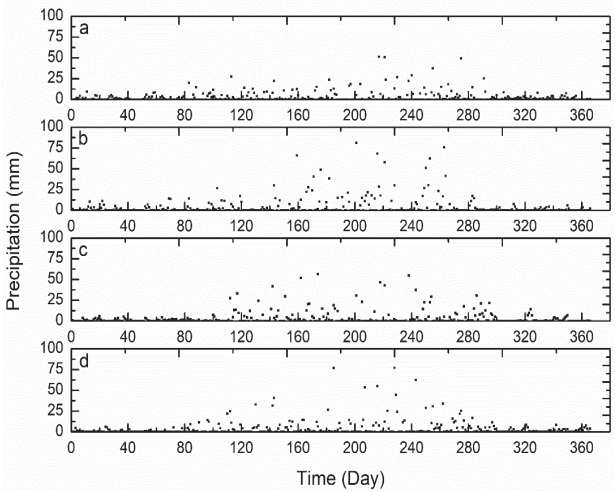

Fig. 3. The precipitation change map of Gongxian county during 2015-2018: a) 2015, b) 2016, c) 2017, d) 2018.

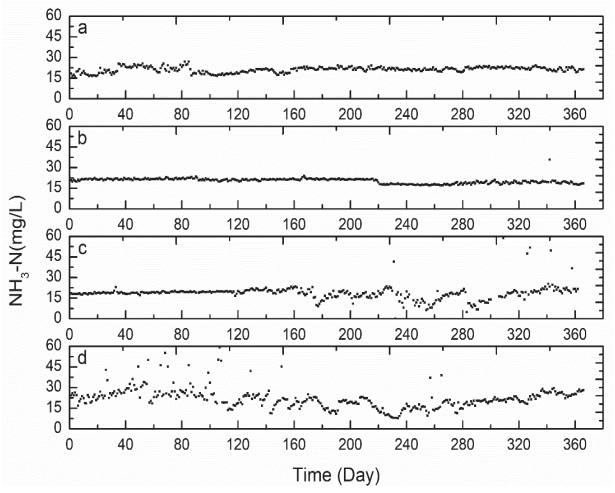

Fig. 6. The influent NH_3-N of Gongxian WWTP during 2015-2018: a) 2015, b) 2016, c) 2017, d) 2018.

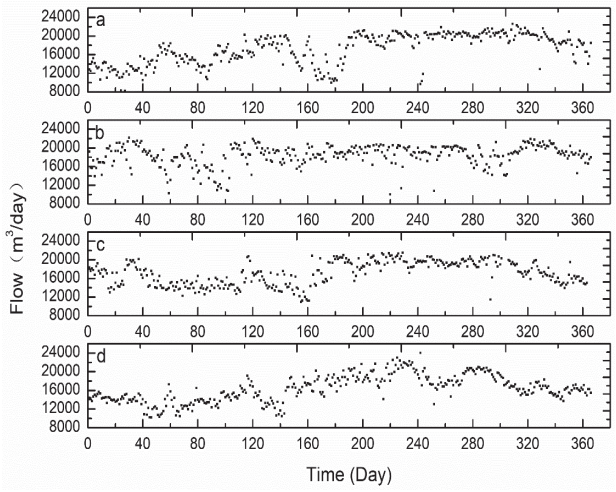

Fig. 4. Influent flow of Gongxian WWTP during 2015-2018: a) 2015, b) 2016, c) 2017, d) 2018.

and the minimum was only half of that in 2015, reflecting increased volatility in incoming NH_3-N.

Model Building

Linear Regression

The simplest linear regression is a straight line used to fit a series of points on a two-dimensional plane. The purpose of linear regression is to summarize the distribution rule or trend of all training samples, and finally to predict the new sample points. The general form of the equation for a line in a two-dimensional plane is expressed as y = Ax + B. After training the model using data from the training set, the optimal values of the two parameters A and B in the equation can be determined and used to predict newly observed samples to obtain the predicted value of y. In three

dimensions, the parameters of a two-dimensional plane should be determined, and so on. In n-dimensional space, the parameters of an 'n-1' dimensional hyperplane should be determined.

This method is called linear regression because the model is composed of linear combinations of all features, and its basic form is shown in Formula (1) [23]:

$$\hat{y} = h_\theta(x) = \theta_0 + \theta_1 x_1 + \theta_2 x_2 + \cdots + \theta_n x_n \quad (1)$$

$\hat{y}$: Represents the predicted value of a linear regression model

n: Represents the number of features

x_j: Represents the observation of the J_{th} feature

θ_j: Represents the value of the J_{th} parameter

With the training data and the model, it is also necessary to define the appropriate cost function, which quantifies the error between the predicted and observed values. After selecting an appropriate cost function, the training process identifies the minimum value. For the linear regression algorithm, the most commonly used cost function is the MSE function, as shown in Formula (2).

$$J_0 = J(\theta) = \frac{1}{m}\sum_{i=1}^{m}\left(\hat{y}^{(i)} - y^{(i)}\right)^2 \quad (2)$$

Then, the gradient descent method is used to train the model to obtain an optimal solution for the above mentioned cost function and the weight vector corresponding to Formula (3).

$$\overrightarrow{w^T} = (\theta_0, \theta_1, \theta_2, \ldots, \theta_n) \quad (3)$$

In the general Linear Regression mentioned above, the assumed function is a linear equation, which is expressed as a straight line in the two-dimensional plane. However, in many cases where the equation of the line does not fit the data well, polynomial regression may be an alternative. Higher powers (such as square or cubic terms) applied in polynomial regression means the increase of the model freedom, and it may be helpful to the capture of nonlinear changes in the data. It is known that adding higher-order terms also increase the complexity of the model. As model complexity increases, the capacity and the ability of the model to fit data also increased, which can further reduce the training error but increase the risk of overfitting.

In polynomial regression, the most important parameter is the degree of the highest power. If the degree of the highest power is n and there is only one characteristic, the polynomial regression equation can be expressed as Formula (4) [24].

$$\hat{y} = h_\theta(x) = \theta_0 + \theta_1 \cdot x + \theta_2 \cdot x^2 + \cdots + \theta_{n-1} \cdot x^{n-1} + \theta_n \cdot x^n \quad (4)$$

Ridge Regression

When using polynomial regression, if the highest degree term of polynomial is large, the model is prone to overfitting and regularization is commonly employed. Ridge Regression, also known as L2 regularization, is a method to prevent overfitting during linear regression. The only difference between ridge regression and polynomial regression is the cost function. The cost function of ridge regression is shown in Formula (5).

$$J(\theta) = J_0 + \lambda \sum_{j=1}^{n} {\theta_j}^2 \quad (5)$$

This adds a penalty term to the original cost function to make the weight of the higher order term close to zero. Ridge regression can be considered as long as the data is linearly dependent and the polynomial regression does not fit well, so that regularization is required.

Lasso Regression

Lasso regression is very similar to ridge regression in that it uses different regularization terms. Constraint parameters are established to prevent over fitting. However, there is another reason why Lasso is important: Lasso can train the parameters of some features with small functions to zero and obtain sparse solutions. In other words, dimensionality reduction (feature screening) is achieved in the training model with this method. The cost function of Lasso is shown in Formula (6).

$$J(\theta) = J_0 + \lambda \sum_{j=1}^{n} |\theta_j| \quad (6)$$

Elastic Net Regression

Elastic regression is used when too many features are sparse to zero and ridge regression is not regularized, or the regression coefficient attenuation is too slow. Therefore, ElasticNet regression can be used for synthesis to obtain better results. It integrates ridge regression and Lasso regression, and its loss function is shown in Formula (7), where R represents the proportion of the Lasso regression term.

$$J(\theta) = J_0 + \gamma\lambda \sum_{j=1}^{n} |\theta_j| + \frac{1-r}{2}\lambda \sum_{j=1}^{n} {\theta_j}^2 \quad (7)$$

For the purposes of this study, x_1 represents temperature, x_2 represents rainfall, and y represents the predictive variable. The simplest function expression is shown in Formula (8).

$$y = \theta_0 + \theta_1 \cdot x_1 + \theta_2 \cdot x_2 \quad (8)$$

After sample testing and a comparison of the above methods, we found that the highest order of

the eigenvalue was line 3 with the highest accuracy, namely the function expression shown in formula (9).

$$y = \theta_0 + \theta_1 \cdot x_1 + \theta_2 \cdot {x_1}^2 + \theta_3 \cdot {x_1}^3 + \theta_4 \cdot x_2 + \theta_5 \cdot {x_2}^2 + \theta_6 \cdot {x_2}^3 \quad (9)$$

Furthermore, considering the correlation between temperature and rainfall, the accuracy was maximized when the function expression is shown in formula (10).

$$y = \theta_0 + \theta_1 \cdot x_1 + \theta_2 \cdot {x_1}^2 + \theta_3 \cdot {x_1}^3 + \theta_4 \cdot x_2 + \theta_5 \cdot {x_2}^2 + \theta_6 \cdot {x_2}^3 + \theta_7 \cdot x_1 x_2 + \theta_8 \cdot {x_1}^2 x_2 + \theta_9 \cdot x_1 {x_2}^2 \quad (10)$$

In this study, polynomial regression in linear regression was first used as a machine learning algorithm, but if the highest polynomial degree term was large, the model was prone to overfitting. In this case we added a constraint on the parameter to the cost function called regularization, which is a common way to prevent overfitting. Through L1 and L2 regularization, Lasso regression and Ridge regression were obtained. L1 regularization reduces the complexity of the model by thinning (reducing the number of parameters), so that parameter values can be reduced to zero. L2 regularization reduces model complexity by reducing the numbers of parameters, so parameter values can only be reduced continuously but never to zero. ElasticNet regression, an algorithm considering both L1 and L2 regularization, is a good contraction method to handle multicollinearity and variable screening, with less precision loss.

Model Evaluation

Based on the platform of Python 3.7, the precipitation, temperature, influent flow, influent COD and ammonia were taken as an array to train the model. The time interval of the training data was from January 1, 2015 to December 31, 2017. Machine learning was conducted using the Linear Regression, Ridge, ElasticNet and Lasso methods.

Data from January 1, 2018 to December 31, 2018 were used as true values and compared with the values predicted by different methods. The prediction effect was evaluated by the following indicators.

Data Normal Rate

Based on common sense and relevant standards [25] negative values of flow rates, COD and ammonia that predicted by models were regarded as outliers by default, and those with COD values exceeding 500 mg/L were also considered as outliers (the design upper limit of COD in general WWTP is 500 mg/L).

$$S = \frac{Q_n}{Q} \quad (11)$$

S: Normal rate of data
Qn: Represents the number of normal predicted values obtained using different prediction methods, n = 1,2,3,4, respectively representing Linear regression, Ridge regression, ElasticNet regression and Lasso regression.
Q: Represents the number of observed values counted

Average Prediction Error

$$W = \frac{1}{n}\sum_{i=1}^{n}\frac{|K_P - K_T|}{K_T} \quad (12)$$

W: Average prediction error
K_P: Prediction values based on different methods
K_T: Statistical observations

Results and Discussion

In this study, four different machine learning methods were used to predict the influent quantity, COD and ammonia concentration in 2018 Gongxian WWTP. Results are shown in Figs 7-9 (FT represents the recorded inlet flow rate value, FP represents the predicted value; CT represents the measured value of influent COD, CP represents the predicted value; NT represents the measured value of influent ammonia, NF represents the predicted value).

Influent Quantity Prediction

The influent flow predicted by the four methods showed no significant difference on the data normal rate and average prediction error (Fig. 7, Table 1), and the data normal rates were greater than 98.9% and prediction accuracy ranged from 85.80% to 86.19%. It can be seen that the prediction effects of the four methods are relatively close. Hazhar Sufi Karimi et al.

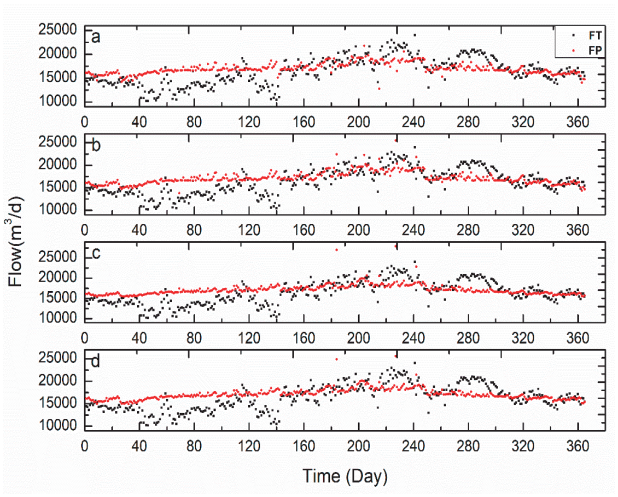

Fig. 7. Comparison of predicted and measured influent flow using Linear Regression a), Ridge Regression b), ElasticNet Regression c) and Lasso Regression d) methods.

Table 1. The data normality rate and average prediction error of influent flow predicted by different methods.

Method	Influent quantity				
	Total number of data sets	Number of normal data sets	Normal rate	Average prediction error	Accuracy
Linear Regression	365	361	98.9%	13.89%	86.11%
Ridge Regression	365	365	100%	13.81%	86.19%
Elastic Net Regression	365	365	100%	14.20%	85.80%
Lasso Regression	365	365	100%	14.02%	85.98%

have shown that Lasso has a relatively good predictive performance in traffic prediction [26]. In this study, Lasso also shows a better predictive performance, only slightly inferior to Ridge. Comprehensive analysis showed that Ridge had an ideal effect on influent flow prediction and can be used in WWTP.

Influent Quality Prediction

The influent COD values predicted by different methods are shown in Fig. 8 and Table 2. During Day 160 to Day 260, significant differences on COD values were occurred between predicted and measured values when Linear Regression method was applied, which was also reflected by average prediction errors of up to 37% (63% accuracy). Gao et al. used Multivariate Linear Regression method to construct the influent prediction model, which was effectively applied in general conditions [27-29], but also found that the prediction accuracy of the simple linear regression model was not the best, which was more consistent with the conclusion of this study.

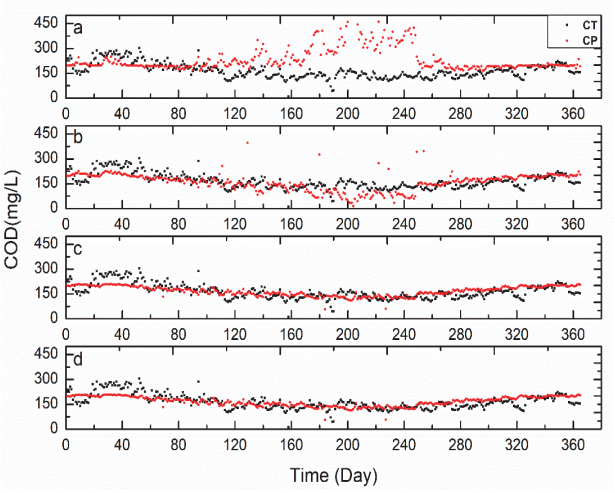

Fig. 8. Comparison of predicted and measured COD values using Linear Regression a), Ridge Regression b), ElasticNet Regression c) and Lasso Regression d) methods.

The prediction accuracies of three methods were accuracy of 78%, 78% and 82%, respectively. In comparison, Ridge had a slightly higher prediction accuracy than other two methods. In addition, the data normal rate predicted by the four methods were relatively high, all of them were greater than 96.16%. Overall, Ridge had the highest prediction accuracy with respect to influent COD values, while ElasticNet and Lasso had slight advantages on data normality rate. Influent NH_3-N predicted by different methods are shown in Fig. 9 and Table 3. Many outliers predicted by Linear Regression method were obtained in the range of Day140-Day250, while the prediction performance of Regression method was even worse (outliers were distributed in the range of day 100 to day 270). On the contrary, ElasticNet and Lasso methods maintained higher normal data rate (100%). In terms of prediction accuracy, Linear Regression and Ridge methods had a high accuracy of 82% and 81%, while ElasticNet and Lasso had higher prediction errors of 26% and 26% (74%, 74% accuracy), respectively. It should be noted that the lower prediction errors obtained by Linear Regression and Ridge methods were based on the elimination of many outliers (the predicted results

Table 2. The data normality rate and average prediction error of influent COD predicted by different methods.

Method	COD concentration of influent				
	Total number of data sets	Number of normal data sets	Normal rate	Average prediction error	Accuracy
Linear Regression	365	351	96.16%	37%	63%
Ridge Regression	365	355	97.26%	18%	82%
ElasticNet Regression	365	365	100%	22%	78%
Lasso Regression	365	365	100%	22%	78%

Table 3. The data normality rate and average prediction error of influent NH_3-N predicted by different methods.

Method	NH_3-N concentration of influent				
	Total number of data sets	Number of normal data sets	Normal rate	Average prediction error	Accuracy
Linear Regression	365	253	69.3%	18%	82%
Ridge Regression	365	192	52.6%	19%	81%
ElasticNet Regression	365	365	100%	26%	78%
Lasso Regression	365	365	100%	26%	78%

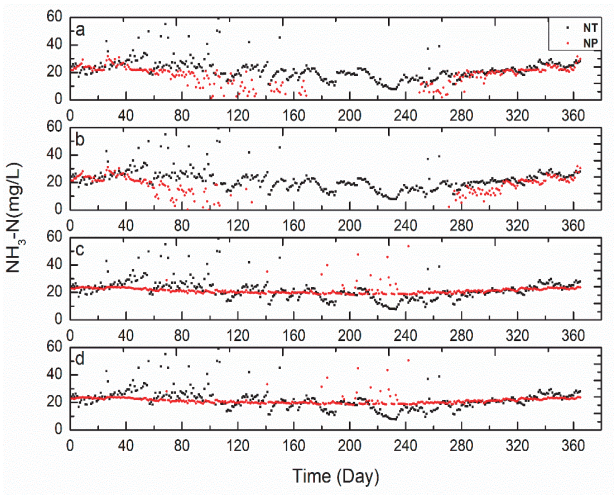

Fig. 9. Comparison of predicted and observed NH_3-N values using Linear Regression a), Ridge Regression b), ElasticNet Regression c) and Lasso Regression d) methods.

were negative), which as most likely the reason for their low prediction errors. Moreover, in terms of application and function of the prediction models, if the prediction accuracy difference is not very large, the predicted data normality rate should be a priority indicator. Therefore, ElasticNet and Lasso methods were more reasonable selections for influent NH_3-N prediction.

Error Analysis

Compared with traditional prediction studies, the methods adopted in this study showed no superior prediction accuracy. The possible reasons for this phenomenon are as follows: (a) **Data fluctuated**: the observed data of influent flow, COD and NH_3-N from 2015 to 2017 were used as training data to predict the data of 2018, and the observed value of 2018 was used as a control group to evaluate the prediction accuracy. It can be found that the average influent flow value from 2015 to 2018 was closed (16345 m^3/d-18328 m^3/d), with data fluctuation of 10.8%. However, the minimum value of influent flow data fluctuates greatly (1110 m^3/d-10176 m^3 /d), with data fluctuation of 89.1%. There is a fluctuation (165.85 mg/L-198.02 mg/L) in the average influent flow value of influent COD during 2015-2018, and the observed value in 2018 is significantly lower than that during 2015-2017. The mean value of the influent concentration of NH_3-N is relatively close (21.17 mg/L-22.39 mg/L), but the maximum influent concentration (27.06 mg/L-69.28 mg/L) fluctuates to 60.9%, and the minimum influent concentration (4.72 mg/L-16.89 mg/L) fluctuates to 72.1%. Theoretically, the more regular the data is, the higher the prediction accuracy will be, and the fluctuation of the data will lead to the decrease of the prediction accuracy [30]. (b) *Multifactorial influence*: The influent quality of the sewage plant was affected by weather, drainage system, population density, sewage collection, pipe network and other factors. However, in practical application, many factors can only be obtained in a long-term quantitative way, such as population density and pipe networks, etc., so it is difficult to obtain reliable dynamic data to improve prediction accuracy by machine learning. Using a few variables to predict a value that is jointly determined by many influencing factors with no obvious quantitative relationship will impact the prediction accuracy [31]. (c) **Data pool**: Data is the core of machine learning, and the size of the data pool has a great impact on the accuracy of prediction [32]. In the sewage treatment industry, the online monitoring system is adopted to obtain the data of influent quality and quantity, which is relatively easy to obtain a large number of training data, and lays a foundation for the sewage treatment industry to adopt the machine learning method to predict the influent quality and quantity. However, due to the late application and development of online monitoring system in China's environmental protection industry, the amount of online data stored in sewage plants is still small at present, only 3-5 years. This is far from enough for the training of a high-precision machine learning prediction model, which indirectly leads to the low prediction accuracy.

Conclusions

In most developing countries, rain and wastewater are still not fully separated. Ensuring the stable operation of WWTP is very important in the special weather conditions (such as heavy rain, high

temperature, etc.). Influent prediction and early warning system of WWTP can reduce the risk of abnormal operation. In this study, four machine learning methods of Linear Regression, Ridge, ElasticNet and Lasso were used to predict the influent quality and quantity. For influent quantity prediction, the model constructed by the all algorithms showed a high accuracy (85.80%-86.19%) and a high normal rate of data (98.99-100%). For influent quality (COD) prediction, the Ridge method (normal rate: 97.26%, accuracy: 82%) is relatively ideal. In terms of prediction for NH3-N, Lasso and ElasticNet (normal rate: 100%, accuracy: 78%) as ideal prediction method. The results proved that the machine learning prediction model of influent quantity and quality can be used as a warning module assisting the operation management of WWTP, and also be an important component of intelligent sewage treatment plants.

Acknowledgements

The authors would like to acknowledge the financial support from Instrument Developing Project of the Chinese Academy of Sciences (YJKYYQ20180002), Sichuan Key Point Research and Invention Program (2019YFS0501) and Youth Innovation Promotion Association CAS (2016331), respectively.

Conflict of Interest

The authors declare no conflict of interest.

References

1. KAPLEY A., PUROHIT H.J. Diagnosis of Treatment Efficiency in Industrial Wastewater Treatment Plants: A Case Study at a Refinery ETP. Environmental Science & Technology, **43** (10), 3789, **2009**.
2. HERRERO M., STUCKEY D.C. Bioaugmentation and its application in wastewater treatment: A review. Chemosphere, **140** (dec.), 119, **2015**.
3. KRZEMINSKI P., VAN DER GRAAF J.H.J.M., VAN LIER J.B. Impact of inflow conditions on activated sludge filterability and membrane bioreactor (MBR) operational performance. Desalination and Water Treatment, **56** (1), 1, **2015**.
4. CHEN Y.Z., LI B.K., YE L., PENG Y.Z. The combined effects of COD/N ratio and nitrate recycling ratio on nitrogen and phosphorus removal in anaerobic/anoxic/aerobic (A^2/O)-biological aerated filter (BAF) systems. Biochemical Engineering Journal, **93**, 235, **2015**.
5. LOVE N.G., BOTT C.B. A review and needs survey of upset early warning devices. Water Environment Research Foundation: Project 99-WWF-2, **2000**.
6. OBRIEN G.J., TEATHER E.W. A dynamic model for predicting effluent concentrations of organic priority pollutants from an industrial wastewater treatment plan. Water Environment Research, **67** (6), 935, **1995**.
7. XIAO Y.Y., DE ARAUJO C,. SZE C.C.,. STUCKEY D.C. Toxicity measurement in biological wastewater treatment processes: A review. Journal of Hazardous Materials, **286**, 15, **2015**.
8. BARONI P., BERTANZN G., COLLIVIGNARELLI C., ZAMBARDA V. Process improvement and energy saving in a full scale wastewater treatment plant: Air supply regulation by a fuzzy logic system. Environmental Technology, **27** (7), 733, **2006**.
9. NAGY-KISS A.M., SCHUTZ G. Estimation and diagnosis using multi-models with application to a wastewater treatment plant. Journal of Process Control, **23** (10), 1528, **2013**.
10. WANG Z.H. Impact analysis of seasonal and Process on Pollutant Removal and Costs of Wastewater Treatment Plant. Dalian University of Technology, **2014** [In Chinese].
11. ANSARI M., OTHMAN F., ABUNAMA T., El-SHAFIE A. Analysing the accuracy of machine learning techniques to develop an integrated influent time series model: case study of a sewage treatment plant, Malaysia. Environmental Science and Pollution Research, **25** (12), 12139, **2018**.
12. ABYANEH H.Z. Evaluation of multivariate linear regression and artificial neural networks in prediction of water quality parameters. Journal of Environmental Health, **12**, **2014**.
13. BOYD G., NA D., LI Z. SNOWLING S., ZHANG Q.Q., ZHOU P.X. Influent Forecasting for Wastewater Treatment Plants in North America. Sustainability, **11** (6), **2019**.
14. NAJAFZADEH M., ZEINOLABEDINI M. Prognostication of waste water treatment plant performance using efficient soft computing models: An environmental evaluation. Measurement, **138**, 690, **2019**.
15. SZELAG B., BARTKIEWICZ L., STUDZINSKI J., BARBUSINSKI K. Evaluation of the impact of explanatory variables on the accuracy of prediction of daily inflow to the sewage treatment plant by selected models nonlinear. Archives of Environmental Protection,**43** (3), 74, **2017**.
16. ZHOU P., LI Z., SNOWLING S., BAETZ B.W., NA D., BOYD G. A random forest model for inflow prediction at wastewater treatment plants. Stochastic Environmental Research and Risk Assessment, **33** (10), 1781, **2019**.
17. JIN Y., YOU X.Y., JI M. Process response of wastewater treatment plant under large rainfall influent flow. Environmental Engineering and Management Journal, **15** (11), 2357, **2016**.
18. RYU J., LEE J., OH J. Examination of the storage function of intercepting sewers using long-term flow monitoring data. Desalination and Water Treatment, **54** (4-5), 1299, **2015**.
19. KACZOR G., BUGAJSKI P. Impact of Snowmelt Inflow on Temperature of Sewage Discharged to Treatment Plants. Polish Journal of Environmental Studies, **21** (2), 381, **2012**.
20. Ministry of Ecology and Environment the People's Republic of China. Water quality-Determination of the chemical oxygen demand-Dichromate method, HJ 828-2017, **2017** [In Chinese].
21. Ministry of Ecology and Environment the People's Republic of China. Water quality – Determination of ammonia nitrogen – Nessler's reagent spectrophotometry, HJ 535-2009, **2009** [In Chinese].
22. China Weather. Available onlinehttp://www.weather.com.cn/zt/kpzt/314261.shtml (accessed on October 10, 2020) [In Chinese].

23. YANG Y.D. Development of the regional freight transportation demand prediction models based on the regression analysis methods. Neurocomputing, **158**, 42, **2015**.
24. SAYIN B., SEVGEN S., SAMLI R. Simulation of experimental parameters of RC beams by employing the polynomial regression method. Mechanics of Composite Materials, **52** (3), 379, **2016**.
25. General Administration of Quality Supervision, Inspection and Quarantine of the People's Republic of China. Wastewater quality standards for discharge to municipal sewers. GB/T 31962-2015, **2015** [In Chinese].
26. SUFI KARIMI H., NATARAJAN B., RAMSEY CL., HENSON J., TEDDER JL., KEMPER E. Comparison of learning-based wastewater flow prediction methodologies for smart sewer management. Journal of Hydrology, **577**, 123977, **2019**.
27. ABYANEH H.Z. Evaluation of multivariate linear regression and artificial neural networks in prediction of water quality parameters. Journal of Environmental Health Science and Engineering, **12**, 40, **2014**.
28. GAO X., BAI L. The Prediction of Indices at Infall of Confluent Flow Network of Wastewater with Multivariate Linear Regression. Proceedings of The 29th Chinese Control Conference, 5125, **2010** [In Chinese].
29. WANG X., KVAAL K., RATNAWEERA H. Explicit and interpretable nonlinear soft sensor models for influent surveillance at a full-scale wastewater treatment. plant Journal of Process Control, **77**, 1, **2019**.
30. KONG J., LEE H., KIM D., HAN SK., HA D., SHIN K., KIM S. Network-based machine learning in colorectal and bladder organoid models predicts anti-cancer drug efficacy in patients. Nature Communications, **11**, 5485, **2020**.
31. LIU G.M. Research on Two-Stage Feature Selection Methods in Machine Learning. Harbin, China: Harbin University of Science and Technology, **2015** [In Chinese].
32. RAO D.Y., JIN Y., TANG L.N., MIN S.J., YANG F., DING X., WU J. A method and system for automatic annotation of data based on intelligent pattern recognition. **2018** [In Chinese].