Open Journal of Modelling and Simulation, 2024, 12, 114-129
https://www.scirp.org/journal/ojmsi
ISSN Online: 2327-4026
ISSN Print: 2327-4018

Prediction and Analysis of Total Nitrogen in a Sewage Treatment Plant Effluent

Kemeng Xue

Beijing No.4 High School, Xicheng District, Beijing, China
Email: mollyxkm@126.com

How to cite this paper: Xue, K.M. (2024) Prediction and Analysis of Total Nitrogen in a Sewage Treatment Plant Effluent. *Open Journal of Modelling and Simulation*, **12**, 114-129.
https://doi.org/10.4236/ojmsi.2024.124008

Received: September 17, 2024
Accepted: October 18, 2024
Published: October 21, 2024

Abstract

Total nitrogen was an important indicator for characterizing eutrophication of polluted water. Although the use of water quality online monitoring instrument can monitor water quality changes in real time, the degree of intelligence was low, so it was urgent to predict the water quality and take precautions in advance. A predictive model for total nitrogen levels in a sewage treatment plant utilizing the Anaerobic-Anoxic-Oxic (AAO) process was investigated in this paper. This model demonstrated significant practical application value. Based on the ARIMA (Autoregressive Integrated Moving Average) model and taking into account the impact of Biochemical Oxygen Demand (BOD), a prediction model for effluent total nitrogen was developed. However, the initial results exhibited significant deviations. To address this issue, seasonal factors were further considered. Then, the dataset was divided into winter and Non-winter sub-samples, leading to a reconstruction of the prediction model. Additionally, in developing the Non-winter prediction model, life cycle considerations were incorporated, and consequently, a SARIMA (Seasonal Autoregressive Integrated Moving Average) model was established. The predicting deviation associated with both the winter and Non-winter forecasting models showed a significant reduction.

Keywords

Total Nitrogen, Sewage Treatment Plant, Prediction, ARIMA Model

1. Introduction

Total nitrogen was an important indicator for characterizing eutrophication of polluted water. At present, the AAO (Anaerobic-Anoxic-Oxic) process has been adopted in most municipal sewage treatment plant. Due to the complex process of sewage treatment process involved of physics, chemistry, and biology, the current operating procedures of sewage treatment plant were mainly controlled by personnel based on their own experience. However, with the widespread adoption of automatic monitoring instruments, improving the automation control level of sewage plants and ensuring stable and compliant effluent were currently key issues in the operation of sewage plants.

It was urgent to establish a predictive model for the effluent quality under a certain stable operating condition. At present, a big data modeling method based on statistical method and machine learning theory to industrial processes was becoming a popular research method [1]. The technical difficulty lied in how to construct models under the influence of multiple factors, and how to determine the interaction relationships between different factors [2]. Some papers [3] [4] have disclosed some prediction methods on effluent water quality. In these papers, Arima, BP Neural Network and support vector regression model has been reported [5].

A sewage treatment plant in Chengdu Shi, Sichuan Province was regarded as the research object in this paper. This plant collected urban residents' sewage, with a designed treatment capacity of 34,000 tons/day. The automation level of the control system and control efficiency in this plant was low due to the manual operation. This plant had automatic control instruments to collect water quality data, but there were not water quality prediction measures and feedback control system. Therefore, process adjustment was not timely and effluent quality was not stable.

2. Research Methods

2.1. Sewage Treatment Process

For urban sewage treatment plant, biological methods were mainly used, including activated sludge process, SBR process, AB process, anaerobic treatment process, biofilm process, etc. Among them, the AAO process based on the principle of activated sludge process was the main process currently.

The principle of AAO process was shown in Figure 1.

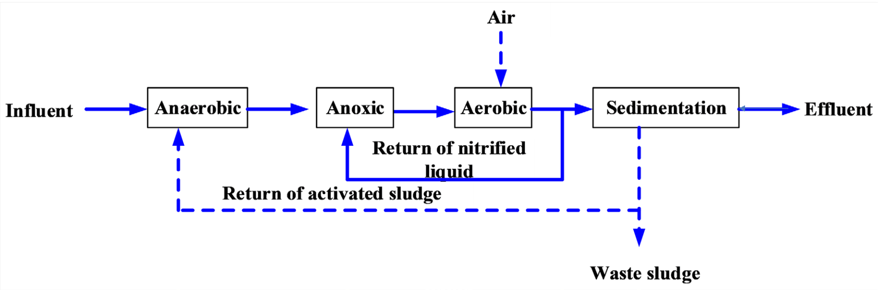

Figure 1. Process chart of a typical AAO.

As shown in Figure 1, the sewage influent flowed through the anaerobic tank, anoxic tank, and aerobic tank in sequence, and undergoes biochemical reactions with activated sludge to remove COD, nitrogen and phosphorus pollutants.

For the denitrification process of AAO, the total nitrogen in the effluent was not only related to the process parameters, such as temperature, pH, dissolved oxygen, etc., but also significantly correlated with parameters in the water quality, such as BOD (COD) and ammonia nitrogen. Biochemical oxygen demand (BOD) refers to the amount of dissolved oxygen consumed by microorganisms to decompose oxidizable substances, especially organic matter[1].

In the aerobic stage, ammonia nitrogen in the influent will be converted into nitrate nitrogen by the action of nitrite bacteria and nitrifying bacteria. Both microorganisms in this process were autotrophic only consumed oxygen and did not consume BOD, which indicates that nitrification process was independent of BOD. During the anoxic stage, nitrate nitrogen will be converted into nitrogen gas and then released from the water by denitrifying bacteria, under the condition of energy supply from BOD. Therefore, denitrification process was closely related to BOD concentration and the total nitrogen of the effluent was highly relevant to the BOD.

The influent and effluent water quality indicators and hydraulic retention time of key treatment units designed for this sewage treatment plant were shown in Table 1 and Table 2.

Table 1. Influent and effluent water quality in design documents.

Water quality indexes	pH (Dimensionless)	COD (Chemical oxygen demand, mg/L)	BOD (Biochemical oxygen demand, mg/L)	Total nitrogen, TN (mg/L)
Influent	6 - 9	<400	<200	<50
Effluent	6 - 9	<50	<10	<20

Table 2. HRT of different treatment units in an AAO plant.

Process units	Anaerobic tank	Anoxic tank	Aerobic tank	Sedimentation tank
HRT (h)	3.5	6	14	3

2.2. Water Quality Prediction Model

2.2.1. Introduction of ARIMA Model

The prediction of sewage treatment water quality mainly relied on studying historical water quality monitoring data, establishing model relationships to study the laws of future water quality changes, and exploring the development trends of future water quality. At present, the commonly used methods for water quality prediction internationally include time series analysis, regression analysis, grey system analysis and machine learning. Among them, the ARIMA model based on auto-regression has been widely used.

The ARIMA (p, d, q) model, also known as the Auto Regressive Integrated Moving Average Model, was a time series prediction method proposed by Box and Jenkins in the early 1970s. It can be used to predict the characteristics of random processes over time rather than being fixed, and the reason for the instability of time series was random rather than deterministic. The ARIMA model refereed to a

[1]Ammonia nitrogen and other factors do have an impact on TN removal, but due to data acquisition and other reasons, this study did not consider them with reasons descripted in 2.2.2.

model established by regressing only the lagged value of the dependent variable and the present value and lagged value of the random error term during the process of converting non-stationary time series into stationary time series.

In the ARIMA model, AR was auto-regressive and p was the corresponding auto-regressive term. D was the differential order used to obtain a stationary sequence. MA was the moving average, and q was the corresponding number of moving average terms [4].

The ARIMA model is a traditional model, and in this paper, it is applied to water quality prediction, divided into winter and summer based on local climate, and predictions are made using the ARIMA model for each season.

Establishing an ARIMA time series model involves three steps as follows:

1) Stability testing and differential processing of data

Plot the data and observe whether it was a stationary time series. If it was a stationary series, the ARIMA model can be directly established. For non-stationary time series, one or more (d-order) differentiating operations must be performed to convert it into a stationary time series before the model can be established.

2) Determine model parameters

The stationary time series obtained through differential operation was used to calculate the auto-correlation coefficient, ACF, and partial auto-correlation coefficient, PACF. By analyzing the auto-correlation graph and partial auto-correlation graph, the optimal auto-regressive order p and the q of the moving average were obtained. After obtaining the above parameters, the ARIMA model can be established.

3) Model validation

The next step was verifying whether the parameter estimation values of the fitted time series model were significant and verifying whether the residual sequence of the fitted time series model was a white noise sequence. If it does not match, go back to step 2.

The process of establishing the ARIMA model was shown in **Figure 2**.

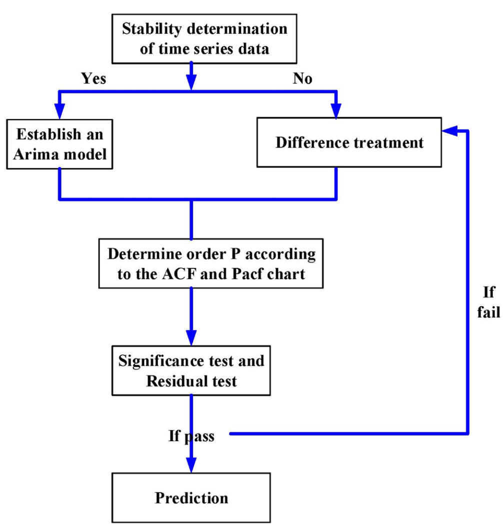

Figure 2. The process chart of modelling.

2.2.2. Determination of Input Variables

Based on the fundamental analysis of the principle of combined nitrification and denitrification for total nitrogen removal in AAO process, the factors such as ammonia nitrogen concentration, BOD concentration and COD concentration, dissolved oxygen concentration, temperature and pH that affect the denitrification effect of sewage treatment plant have been determined [2].

Firstly, although temperature and pH can affect the activity and reaction rate of bacteria, pH and T will generally not be adjusted due to the large water amount in the reality of sewage treatment. Therefore, pH and T were not considered as variables in this paper. However, considering the significant difference of water temperature in various months between winter and Non-winter, in this paper, the data were artificially divided into two periods, the winter period and the Non-winter period. Specifically, the data from Apr.1st to Nov.30th of each year were concluded as Non-winter period, and the data from Dec.1st to Mar.31st of next year were concluded as winter period. Furthermore, dissolved oxygen and ammonia nitrogen concentration were not considered as factors for investigation in this paper because there were no online monitoring probes of dissolved oxygen and ammonia nitrogen concentration.

Secondly, for domestic sewage in municipal treatment plant, there was a clear correlation between BOD and COD. Generally speaking, BOD was half of COD. Therefore, in this paper, only BOD was used as the parameter affecting the model of total nitrogen. In summary, this paper mainly focused on the relation between TN and BOD, respectively all months and distinguish between winter and Non-winter.

2.2.3. Data Preparation and Descriptive Statistical Analysis

1) Data preparation

The daily testing data of the sewage plant from Dec.1st, 2019 to Nov.30th, 2020 as the data sample were adopted in this paper. The monitoring frequency of water quality was once a day, so the time series had 365 sample points, which the water quality indicators included BOD and total nitrogen, with a total of 730 data values.

All data come from the self-control system and are exported in Excel format. After fitting a suitable ARIMA model based on this data, the predicted values were compared with the actual values to verify the predictive performance of the model.

The data modeling and processing process in this paper was carried out using Stata (Version 15.1) software. All variable was defined in **Table 3**.

Table 3. The table of variable definition.

Variable Name	Variable Definition	Unit
TN	Total nitrogen value in effluent water	mg/L
BOD	BOD value in effluent water	mg/L

2) Descriptive statistics of data

Table 4 showed the descriptive statistical results of the variables. As shown in

Table 4, descriptive statistical showed that there were 365 final samples, standard deviation was not large, indicating a smaller range of fluctuations.

Table 4. The table of variable descriptive statistics.

Var.	Obs.	Mean	Std. Dev.	Min	Max
TN	365	4.4529	0.8682	1.9924	7.032271
BOD	365	2.6049	0.7570	0.5000	5.2000

3. Results

3.1. Modeling without Considering Seasonal Changes

Due to the fact that the collected water quality indicator data was across the whole year, data modeling was carried out for the entire sample without considering seasonal change.

3.1.1. Stability Test

The step of stationarity test was shown as follows [6]. ADF unit root test was performed on time series data to verify whether the time series data was stationary. If the test statistic was less than the critical value and the p-value was less than the significance level, which was usually 0.05, and the null hypothesis was rejected. The data will not have unit roots and was considered stationary, and $d = 0$ was determined. If the p-value was greater than 0.05, it was considered that the data was non-stationary and the time series needed to be differenced. After being differenced, ADF unit root test was performed until the data was stationary, and the value of d was determined accordingly.

Using the Stata software command fuller, ADF unit root tests were performed on the time series of total nitrogen and BOD, and the results were shown in Table 5. It was evident that the p-values for both total nitrogen and BOD were below 0.05, indicating that these variables are stationary sequences and, therefore, do not require differentiation. Therefore, $d = 0$ was determined. Table 5 showed ADF test statistic.

Table 5. The table of Augmented Dickey-Fuller (ADF) test statistics.

Var.	ADF test statistic	Test statistic	Prob.*
TN (Total nitrogen)	ADF	−14.490	0.000
BOD	ADF	−9.993	0.000
Test critical values:	1% level	−3.451	
	5% level	−2.875	
	10% level	−2.570	

3.1.2. Autocorrelated Graph and Partial Autocorrelated Graph

After establishing the appropriate degree d, the parameters p and q in the ARIMA model were subsequently identified. In an ARIMA model, the PACF plot was used to identify the order p of the AR (autoregressive) component. If the PACF plot exhibited a significant drop-off in the correlation after a lag p, it could be assumed that the order p of the AR model was equal to that lag value. The ACF plot was mainly used to identify the order q of the MA model.

The shaded regions in the auto-correlation function (ACF) and partial auto-correlation function (PACF) plots typically denote statistical confidence intervals, which assisted in assessing the significance of the correlation coefficients. In the process of determining the p and q values for an ARIMA model, it was common practice to identify the last lag that significantly exceeded the shaded region in both PACF and ACF plots. Nevertheless, the interpretation of the graphical representations may not be straightforward. Therefore, AIC values for each model were typically aggregated to facilitate a comprehensive assessment. The model with the lowest AIC value was regarded as the optimal fitting model. When the sample size was small, or when the objective of model selection was to achieve accurate predictions, AIC was frequently employed due to its relatively lower penalty for model complexity, which permitted greater flexibility in fitting the model to the data. Conversely, when the sample size was large or when the aim of model selection was to provide a robust explanation of the data, BIC may be more appropriate because it imposes a higher penalty for model complexity, thereby favoring simpler models. Consequently, this study adopted the AIC value as the evaluation criterion.

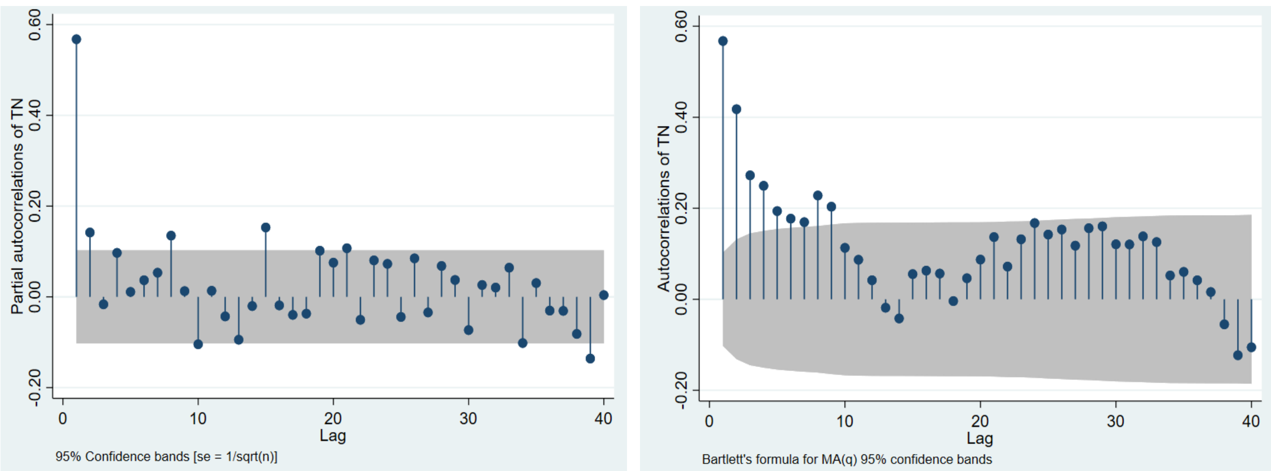

Figure 3. ACF diagram and PACF diagram of TN without considering seasonal changes.

The ACF diagram and PACF diagram of TN without considering seasonal changes was shown in **Figure 3**. As shown in **Figure 3**, it can be determined that the parameter p was 2, and the initial judgment of the parameter q was 3, 4, or 5.

3.1.3. Model Selection and White Noise Testing

Fitted the models and calculated the AIC and BIC values for each ARMA() specification.

Individually fitted the ARIMA(2,0,3), ARIMA(2,0,4), and ARIMA(2,0,5) models and calculated their AIC and BIC values. As presented in **Table 6**, the AIC and

BIC values for the ARIMA(2,0,3) model were the lowest; thus, it was preliminarily identified as the optimal fitting model.

Table 6. The AIC and BIC values for fitted model.

ARIMA(p,d,q) model	AIC	BIC
(2,0,3)	788.3416	819.5407
(2,0,4)	789.5414	824.6405
(2,0,5)	791.3816	830.3806

Predicted the residual r and conducted a white noise test on the residual r.

The Ljung-Box Q statistic served as a diagnostic tool to assess whether the residuals of a time series exhibited auto-correlation. A p-value exceeding 0.05 suggested that the residuals can be considered white noise, indicating an absence of correlation among observations in the series. Thus, it implied that the sequence was purely random. As presented in **Table 7**, with a p-value of 0.506 > 0.05, the residual series passed the white noise test, confirming that the ARIMA(2,0,3) model had effectively captured the dynamics of the time series. MSE 0.4

Table 7. Portmanteau test for white noise.

	Test statistic
Portmanteau (Q) statistic	55.6912
Prob > chi2(40)	0.0506

3.1.4. Model Prediction and Fitting

Using the ARIMA (2,0,3) model obtained above, the predict command was called to predict the values from Dec.1st, 2019 to Nov.30th, 2020.

Table 8 shows the MAE and MSE values for the model. MSE was 0.4853 and MAE was 0.5372. This indicates that the prediction error is relatively large.

Table 8. Evaluation metrics.

Evaluation Metrics	Obs.	Mean	Std. Dev.	Min	Max
MSE	365	0.4853	0.8611	0.0001	10.1270
MAE	365	0.5372	0.4441	0.0106	3.1823

The fitting results of the relationship diagram between actual and predicted values were shown in **Figure 4**, where the line with the label of the circle represented the actual value and the line with the label of the triangle represented the predicted value.

Subtracted the predicted value from the actual value to obtain the difference, and compared the difference with the actual value to obtain the deviation rate, as shown in **Figure 5**. From **Figure 5**, it can be seen that the deviation rate was mainly concentrated in (−40%, 30%). The results showed a few large prediction deviation without considering the seasonal change. Therefore, in order to decrease the prediction deviation, the seasonal change was considered at various months because of temperature change.

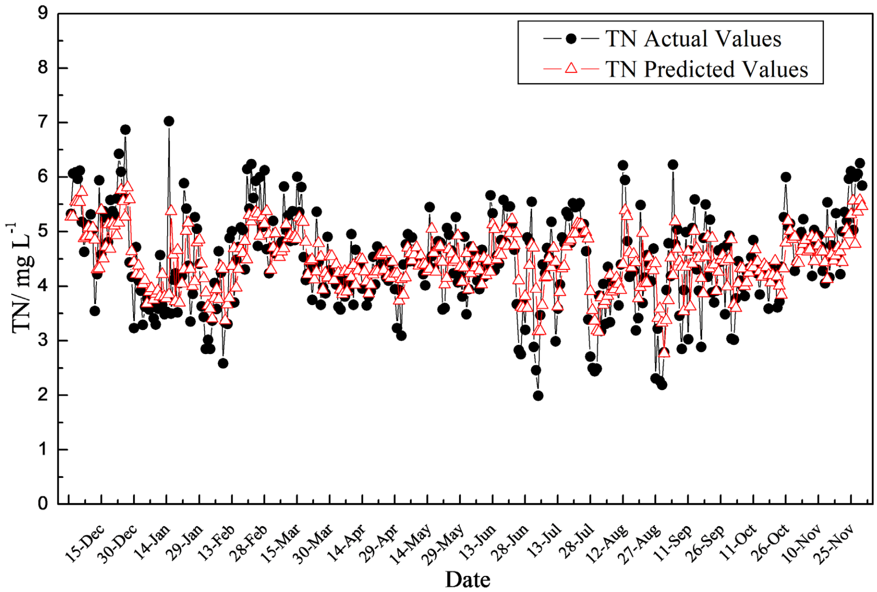

Figure 4. The relationship diagram between actual and predicted values developed using the ARIMA model without condisering seasonal changes.

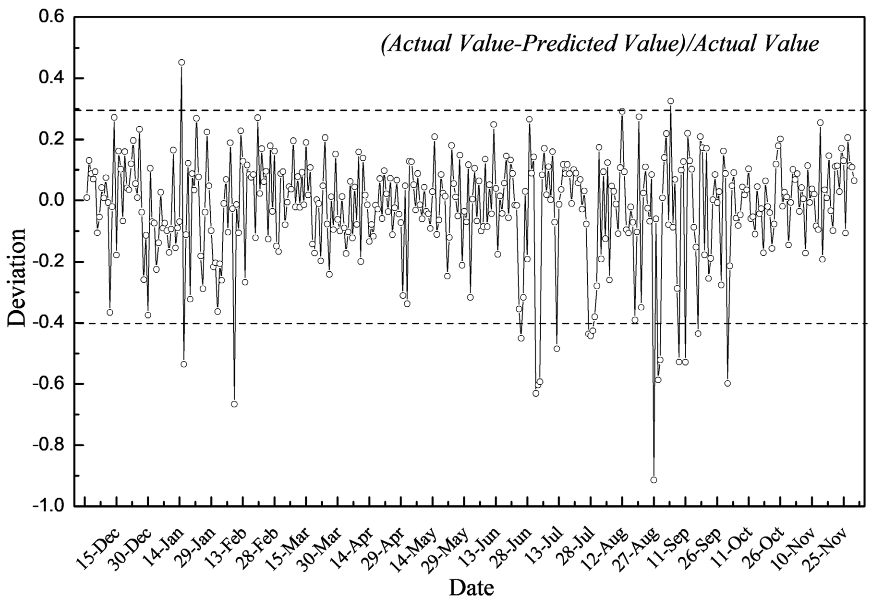

Figure 5. Deviation rate chart between predicted and actual values without considering seasonal changes.

3.2. Modeling Considering Seasonal Changes

The temperature in the Chengdu area exhibits seasonal variation throughout the year, with average temperatures as follows: 12˚C to 17˚C in spring, 20˚C to 28˚C in summer, 19˚C to 25˚C in autumn, and 6˚C to 12˚C in winter. Notably, winter temperatures consistently remain below 12˚C, a stipulation clearly outlined in China's water supply and drainage regulations. This is due to the significant impact that temperatures below 12˚C have on microbial activity. Consequently, all the samples will be categorized into two periods: the winter period (December through March of the following year) and the Non-winter period (April through November). The final number of winter samples was 121, and the number of Non-winter samples was 244.

3.2.1. Modeling in Winter

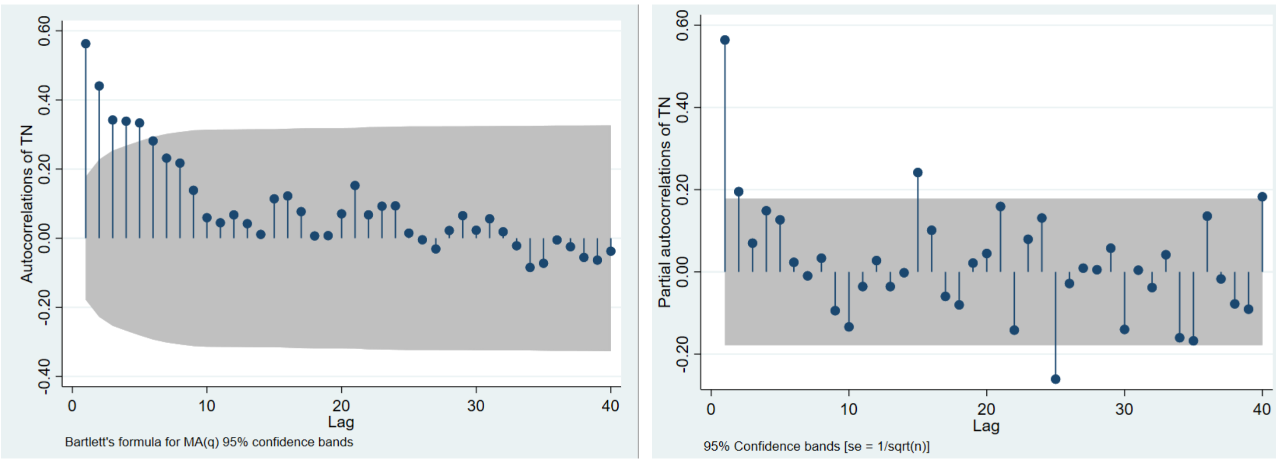

Figure 6. ACF diagram and PACF diagram of TN (In winter).

The ACF and PACF plots of TN in winter was shown in Figure 6. It can be determined that the parameter p was 1, and the initial judgment of the parameter q was 4 or 5 from Figure 6.

The AIC and BIC values for fitted model was shown in Table 9. It can be determined that the AIC value of the ARIMA(1,0,5) model was the smallest, and it was initially determined to be the optimal fitting model from Table 8.

Table 9. The AIC and BIC values for fitted model.

ARIMA(p,d,q) model	AIC	BIC
(1,0,4)	290.6683	313.0302
(1,0,5)	290.3118	315.4739

Predicted residuals, and conducted a white noise test on the residuals. It can be concluded that $p = 0.9784 > 0.05$, and the residual sequence passed the white noise test, indicating that the ARIMA(1,0,5) model has been successfully fitted.

Using the ARIMA (1,0,5) model obtained above, the predict command was called to predict the values from Dec.1st, 2019 to Mar.31st, 2020.

Table 10 shows the MAE and MSE values for the model. MSE was 0.5528 and MAE was 0.5753. Although the prediction error is still quite large, the degree of discreteness has been significantly reduced, which Standard deviation was 1.0581 and 0.4370 separately.

Table 10. Evaluation metrics.

Evaluation Metrics	Obs.	Mean	Std. Dev.	Min	Max
MSE	121	0.5528	1.0581	0.0000	9.5173
MAE	121	0.5753	0.4730	0.0055	3.0850

The fitting results were shown in Figure 7, where the line with the label of circle represented the actual value and the line with the label of triangle represented the predicted value.

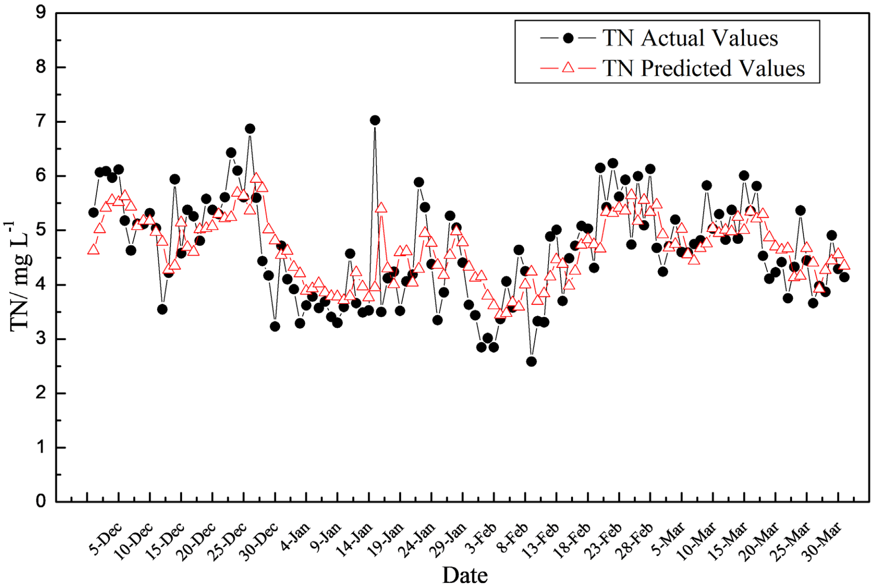

Figure 7. The relationship diagram between actual and predicted values developed using the ARIMA model in winter from 1 Dec.1st, 2019 to Mar.31st, 2020.

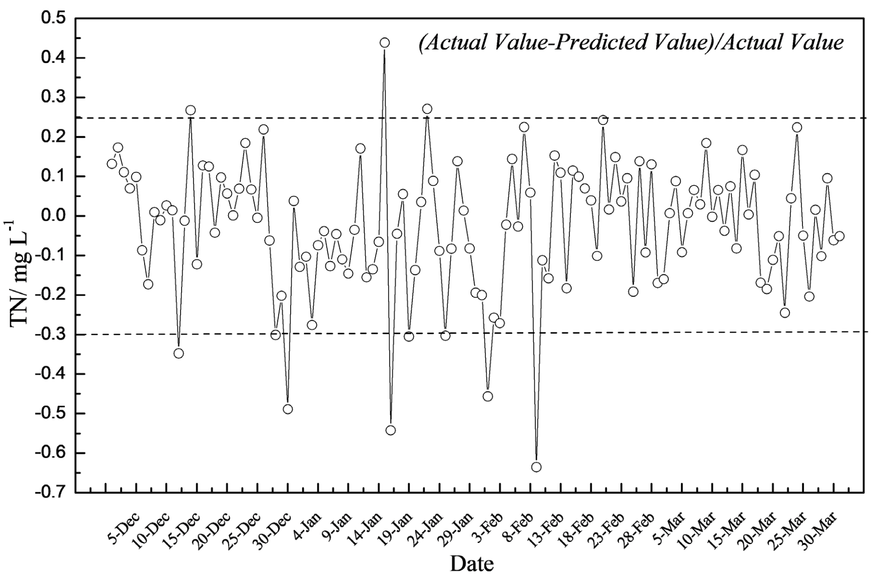

Figure 8. Deviation rate chart between predicted and actual values in winter from 1 Dec.1st, 2019 to Mar.31st, 2020.

From **Figure 8**, it can be seen that the deviation rate was mainly concentrated in (−30%, 20%), and only 12 of the predicted points were far off from the above range. The sample points with large deviation value accounted for 9.92 % of total sample. Especially, Most of the deviation values were negative, which can be beneficial for ensuring that the effluent meets the discharge standard.

3.2.2. Modeling in Non-Winter

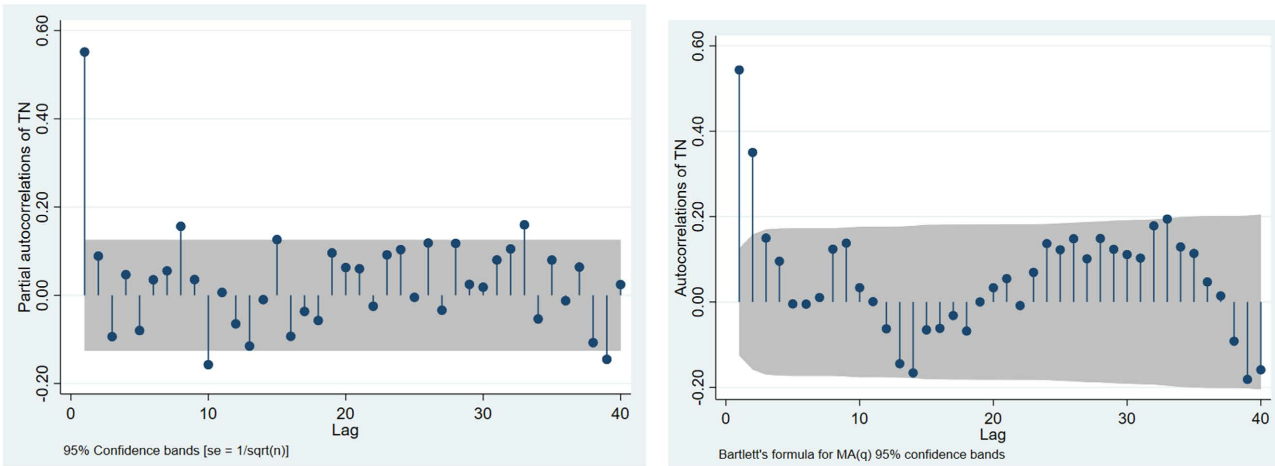

Figure 9. ACF diagram and PACF diagram of TN (In Non-winter).

The ACF and PACF plots of TN in Non-winter was shown in **Figure 9**. As shown in **Figure 9**, it can be determined that the parameter p was 1, and the initial judgment of the parameter q was 1 or 2.

The AIC and BIC values for fitted model was shown in **Table 11**. As shown in **Table 9**, the AIC value of the ARIMA(1,0,2) model was the smallest, and it was initially determined to be the optimal fitting model.

Table 11. The AIC and BIC values for fitted model.

ARIMA(p,d,q) Model	AIC	BIC
(1,0,1)	507.0278	524.5136
(1,0,2)	505.8516	526.8346

Based on ARIMA(1,0,1) and ARIMA(1,0,2) models, the residual series was predicted and the white noise test was conducted on the residual series, as shown in **Table 12**.

Table 12. Portmanteau test for white noise.

ARIMA(p,d,q) model		Test statistic
(1,0,1)	Portmanteau (Q) statistic	64.2148
	Prob.	0.0089
(1,0,2)	Portmanteau (Q) statistic	60.4527
	Prob > chi2 (40)	0.0199

Due to p-values all being less than 0.05 and the residual sequence was not white noise and has not passed the test, indicating that the ARIMA(1,0,2) model has been successfully fitted. The model needs to be modified.

Considering that the data obtained from sewage treatment plant may be closely linked to individuals' daily routines and work cycles, S = 5 or 7 was adjusted seasonally.

As shown in Table 13, the AIC value of SARIMA (1,0,2,7) was the smallest, and it was preliminarily determined that it was the optimal fitting model.

Table 13. The AIC and BIC values for fitted model.

SARIMA(p,d,q,s) model	AIC	BIC
(1,0,2,5)	508.1263	539.6008
(1,0,2,7)	505.7993	537.2738

Predicted residuals, and conducted a white noise test on the residuals. It can be concluded that $p = 0.1589 > 0.05$, and the residual sequence passed the white noise test, indicating that the SARIMA(1,0,2,7) model had been successfully fitted.

Table 14 shows the MAE and MSE values for the model. MSE was 0.4313 and MAE was 0.5131. Standard deviation was 0.6611 and 0.4108 separately. This indicates that the predictive error is relatively reduced, and the degree of dispersion is also decreasing.

Table 14. Evaluation metrics.

Evaluation Metrics	Obs.	Mean	Std. Dev.	Min	Max
MSE	244	0.4313	0.6611	0.0000	3.8424
MAE	244	0.5131	0.4108	0.0029	1.9602

The fitting results were shown in Figure 10, where the line with the label of circle represented the actual value and the line with the label of triangle represented the predicted value.

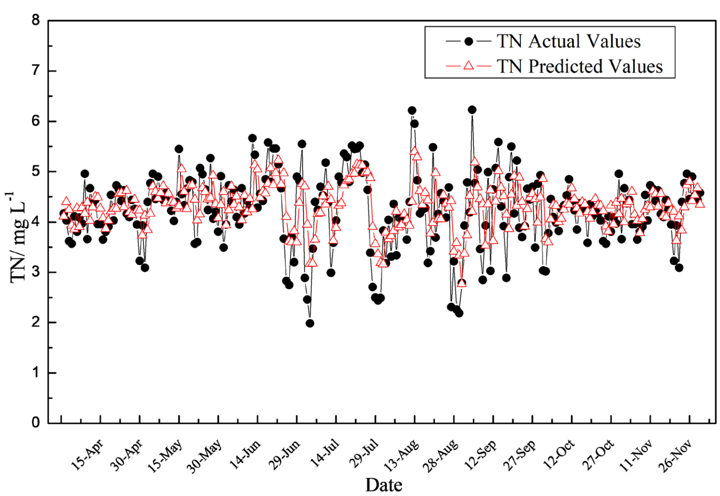

Figure 10. The relationship diagram between actual and predicted values developed using the ARIMA model in Non-winter from Apr.1st,2020 to Nov.30th, 2020.

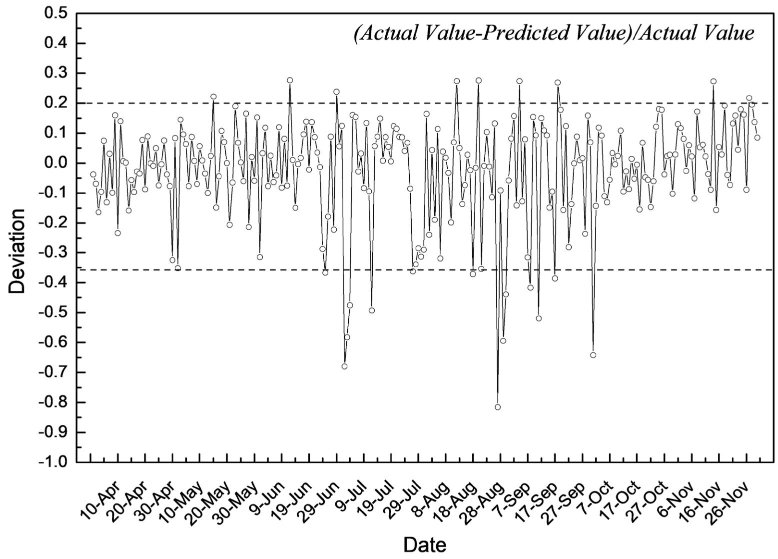

Figure 11. Deviation rate chart between predicted and actual values in Non-winter from 1 Apr.1st, 2020 to Nov.30th, 2020.

Figure 11 showed the deviation rate between predicted and actual values. From **Figure 11**, it can be seen that the deviation rate was mainly concentrated in (−35%, 20%). There were eight predicted value was in (20%, 30%), which accounted for 3.28% of total sample, and eleven predictions had large negative deviations of more than −35%. Firstly, Most of the deviation values were negative, which can be beneficial for ensuring that the effluent meets the discharge standard. Secondly, the days which has most of nagetive deviation was during from June to September, which the temperature were high and microbe had high activity.

4. Conclusions

Using the ARIMA model to predict TN effluent, a full sample analysis was firstly conducted without considering seasonal factors, and data from winter and Non-winter were analyzed separately. The simulation conclusions were as follows:

1) For the entire sample data considering seasonal changes, the deviation rate of the difference between predicted and actual values and the proportion of actual values was concentrated in (−40%, 30%), with a few were far off from the above range. The Deviation without considering the seasonal changes was high. As for considering the seasonal changes, for winter data, the deviation rate was concentrated at (−30%, 20%). For Non-winter data, the deviation rate was concentrated at (−35%, 20%). Both of the deviation decreased to a certain extent.

2) The ARIMA model can be used to predict the trend of TN data during the operation of sewage treatment plant, which was helpful for predicting operating conditions and adjusting operating strategies[2]. In particular, when a negative bias

[2]Because intervals with larger deviations tend to show negative deviations. Nevertheless, if our predicted values exceed the actual values, they may still hold utility for practical applications. Consequently, despite the considerable prediction bias, these predictions retain some degree of reference value in real-world scenarios.

exists—defined as our predicted values surpassing the actual values—these predictions may still possess utility for practical applications. Consequently, despite the significant prediction bias, they retain a degree of reference value in real-world contexts.

However, due to the use of linearization in the ARIMA model, the deviation was relatively large [7] [8]. Moreover, this paper only considers the influence of BOD on TN removal process. In fact, ammonia nitrogen concentration and dissolved oxygen concentration were important factors affecting TN removal efficiency. Consequently, future efforts will involve conducting multiple factor analyses and nonlinear modeling to enhance prediction accuracy.

5. Programming Code

```
sum TN
sum BOD
dfuller TN
dfuller BOD
ac TN
pac TN
arima TN BOD, ARIMA(p,d,q)
estat ic
predict r,res
wntestq r
arima TN BOD, ARIMA(p,d,q) SARIMA(p,d,q,s)
estat ic
```

Conflicts of Interest

The author declares no conflicts of interest regarding the publication of this paper.

References

[1] Newhart, K.B., Holloway, R.W., Hering, A.S. and Cath, T.Y. (2019) Data-Driven Performance Analyses of Wastewater Treatment Plants: A Review. *Water Research*, **157**, 498-513. https://doi.org/10.1016/j.watres.2019.03.030

[2] Liu, J., Li, T., Li, J. (2018) Prediction of Effluent Total Nitrogen Concentration in a Wastewater Treatment Plant Using a Particle Swarm Optimization-Support Vector Regression Model. *Chinese Journal of Environmental Engineering*, **12**, 119-126.

[3] Lin, J.M., Chen, J.L., Lin, J.J., *et al.* (2019) The Simulation and Prediction of TN in Sewage Treatment Effluent Using BP Neural Network and ARIMA Model. *Journal of Environmental Engineering Technology*, **9**, 573-578.

[4] Yao, Y.F., Jing, Y.S., Wang, L.Y., *et al.* (2023) Prediction Method of Total Nitrogen in Wastewater Treatment Plant Based on Integrated Model. *Industrial Water Treatment*, **43**, 187-194.

[5] Jin, Q.L. (2021) Study of Effluent Quality Prediction by ANN Model in WWTP. Master's Thesis, Tianjin University.

[6] Gu, J., Wang, J., Deng, J.H., *et al.* (2020) Water Quality Prediction Based on ARIMa

Model and BP Neural Network Algorithm. *Water Purification Technology*, **39**, 73-82.

[7] Li, C.W. (2018) Prediction of Effluent Total Nitrogen in Wastewater Treatment Plant Based on BP Neural Network Model. Master's Thesis, Huazhong University of Science and Technology.

[8] Wang, R. (2012) Prediction of Wastewater Treatment Based on Genetic Algorithm Optimization BP Artificial Neural Network. Master's Thesis, South China University of Technology.