Analysis of domestic sewage pollution based on weighted grey relational degree and GM(1,1) model

Lidong Liu[a], Yantao He[a,*], Rongfu Wang[a], Sijia Liu[b], Rong Zhang[a], Yongbin Luo[a], Gengshu Zhang[a], Ziyu Chen[a]

[a] Department of Computer Science, Guangdong University of Science and Technology, No. 2, Songshan Lake Section, 523668, China
[b] School of Mechanical Engineering, Hunan University of Science and Technology, Taoyuan Road, 411201, China

* Corresponding author: heyantao@gdust.edu.cn

Abstract. Domestic sewage has always been one of the main causes of river pollution. In this paper, a weighted grey correlation model is established to judge the influence of many factors, including domestic sewage, on river pollution. In the process of establishing the model, the AHP analytic hierarchy process is used to calculate the weight of each influencing factor, and the consistency test is used to judge the feasibility of the weight. This paper also establishes a prediction model of domestic sewage discharge based on GM(1,1) model, the average prediction error of this model is only 0.43%, which proves the superiority of the prediction model.

Keywords: Domestic sewage, weighted grey relational degree, GM (1,1) model

1 Introduction

Water pollution has many negative impacts on life and production. First of all, it poses a certain threat to human health. There are harmful chemicals, bacteria, and viruses in water pollution, which can cause various diseases such as intestinal, skin, and respiratory infections. Then, water pollution also has a serious impact on agricultural production. During the growth process of crops, clean water sources need to be injected. Once the water source is contaminated, it will not only allow the crops to absorb harmful substances but also lead to a decrease in yield and unavailability. It will also lead to a decrease in soil quality and further affect agricultural production. Moreover, water pollution also has an extreme impact on industrial production. Many industrial production processes require a large amount of water. If the water source is contaminated, the process of purifying the water source may require a significant amount of time and resources, which greatly increases production costs.

Among many reasons, the impact of domestic sewage accounts for the largest proportion. It mainly manifests in the following aspects: first, it may cause pathogen contamination, which mainly comes from urban domestic sewage, hospital sewage,

H. Bilgin et al. (eds.), *Proceedings of the 2023 5th International Conference on Civil Engineering, Environment Resources and Energy Materials (CCESEM 2023)*, Advances in Engineering Research 227,
https://doi.org/10.2991/978-94-6463-316-0_33

catering waste, etc. If these sewage containing pathogens enter the environment, it may cause various diseases and affect human health.

Secondly, domestic sewage can cause organic pollution. After organic matter enters the water body, it may consume oxygen in the water, posing a threat to the survival of aquatic organisms. At the same time, it contains a large amount of nutrients, such as nitrogen and phosphorus, which can promote eutrophication of the water body. When eutrophication occurs in the water body, it will lead to excessive reproduction of aquatic organisms such as algae and destroy the ecological balance. In general, domestic sewage poses a certain threat to the environment, economy and people's health. M Yasir uses metagenomic sequencing technology to detect microorganisms in domestic wastewater, discovering 87 pathogenic/opportunistic bacteria belonging to 47 distinct genera in the domestic sewage samples[1].With the help of grey relational analysis, Liang[2] designed an optimization model to determine the weight of decision makers, and further proposed a method of applying PFMCGDM, that is, to obtain the relative correlation degree of the alternatives, and rank them accordingly. In Ceylan's research, Autoregressive Integrated Moving Average (ARIMA), Support Vector Regression (SVR), GM(1,1) and Linear Regression (LR) analysis were used to estimate annual medical waste generation from 2018 to 2023 in Istanbul , so as to help decision-makers to take better measures and develop policies regarding waste management practices in the future[3]. Combining the improved grey correlation degree model with environmental Kuznets curve (EKC), Liu MH[4] quantitatively and qualitatively studied the relationship between water pollution and economic growth in Nansi Lake Basin (Jining, Zaozhuang, Heze) in Shandong Province. The results showed that the correlation degree between industrial wastewater and economic growth was Heze (0.652) > Zaozhuang (0.581) > Jining (0.538), and that between domestic sewage and economic growth was Jining (0.722) > Heze (0.721) > Zaozhuang (0.650). Pandey[5] uses grayscale association analysis (GRA) to evaluate the vulnerability of soil erosion in the Little Himalayan Basin. The method is to use information obtained by remote sensing and field data to evaluate the vulnerability of soil erosion in the sub -field in the watershed.

This paper predicts the discharge of domestic sewage in specific areas based on the weighted grey correlation degree and GM(1,1) model and calls on people to form the habit of saving water, treat domestic sewage correctly, and contribute to the protection of water resources.

2 Evaluation and analysis of multi-factor pollution based on weighted grey relational degree

In order to explore what social groups think is the most important factor affecting river pollution, this article distributes and collects more than a thousand valid questionnaires through the network, and obtains the survey results in figure 1. People regard domestic sewage as one of the main causes of river pollution.

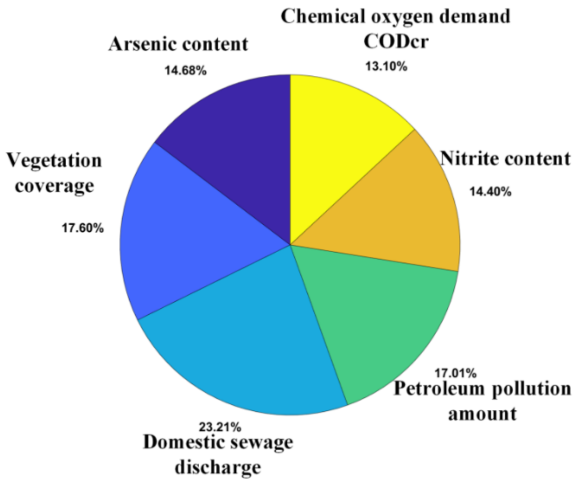

Fig. 1. Results of the questionnaire.

And then, the relevant data of Guangdong section of the Pearl River Basin from 2012 to 2017 are extracted and analyzed, and the data standardized by min-max data in Table 1 are obtained.

Table 1. The main factors affecting water quality.

	2012	2013	2014	2015	2016	2017
Petroleum pollution amount (ten thousand tons)	0.69	0.85	0.94	1.00	0.88	0.00
Arsenic content (ten thousand tons)	1.00	0.00	0.49	0.49	0.49	0.00
Vegetation coverage(%)	0.00	0.46	1.00	0.71	0.32	0.96
Domestic sewage discharge (100 million tons)	0.53	0.64	0.32	0.51	0.88	1.00
Nitrite content(ten thousand tons)	0.00	0.48	1.00	0.76	0.73	0.91
Chemical oxygen demand CODcr(mg/L)	0.00	0.03	0.31	0.65	0.47	1.00

AHP analytic hierarchy process is used to process the data and obtain the corresponding data weight. Let the weight be $k=\left[k_1,\cdots,k_6\right]$, where k_j is the weight corresponding to the j index. First of all, the average value of each analysis item is calculated, and the judgment matrix A is constructed as follows.

$$A=\begin{bmatrix} & \text{Petroleum} & \text{Arsenic} & \text{Vegetation} & \text{Domestic sewage} & \text{Nitrite} & \text{Chemical oxygen demand} \\ \text{Petroleum} & 1.00 & 1.76 & 1.26 & 1.12 & 1.12 & 1.77 \\ \text{Arsenic} & 0.57 & 1.00 & 0.72 & 0.64 & 0.64 & 1.00 \\ \text{Vegetation} & 0.79 & 1.38 & 1.00 & 0.89 & 0.89 & 1.40 \\ \text{Domestic sewage} & 0.89 & 1.57 & 1.13 & 1.00 & 1.00 & 1.58 \\ \text{Nitrite} & 0.83 & 1.52 & 1.12 & 0.98 & 0.98 & 1.56 \\ \text{Chemical oxygen demand} & 0.56 & 0.99 & 0.71 & 0.63 & 0.63 & 1.00 \end{bmatrix} \tag{1}$$

Using formula $CI=(Maximum\ eigenvalue-n)/(n-1)$ to check the consistency of the results, we get $CI=0, RI=1.26$, and bring it into formula $CR=CI/RI$ to get CR equal to 0, which proves that the selected weights are available.The operational formula (2) calculates the grey correlation coefficient:

$$\xi_t(j)=\frac{\min\limits_{1\le t\le 6}\min\limits_{1\le j\le 6}\left|a_0(j)-a_t(j)\right|+\rho \max\limits_{1\le t\le 6}\max\limits_{1\le j\le 6}\left|a_0(j)-a_t(j)\right|}{\left|a_0(j)-a_t(j)\right|+\rho \max\limits_{1\le t\le 6}\max\limits_{1\le j\le 6}\left|a_0(j)-a_t(j)\right|} \tag{2}$$

where $\xi_t(j)$ is the correlation coefficient of the comparison series a_t to the reference series a_0 on the j^{th} index, $\rho\in[0,1]$ is the resolution coefficient, $\min\limits_{1\le t\le 6}\min\limits_{1\le j\le 6}\left|a_0(j)-a_t(j)\right|$ and $\max\limits_{1\le t\le 6}\max\limits_{1\le j\le 6}\left|a_0(j)-a_s(j)\right|$ are the minimum difference and the maximum difference, respectively. If $\rho=0.50$ is set here, the larger the resolution coefficient ρ is, the greater the resolution is. Finally, the dimensionless processing mode is set to average, and the grey correlation coefficient is shown in Figure 2.

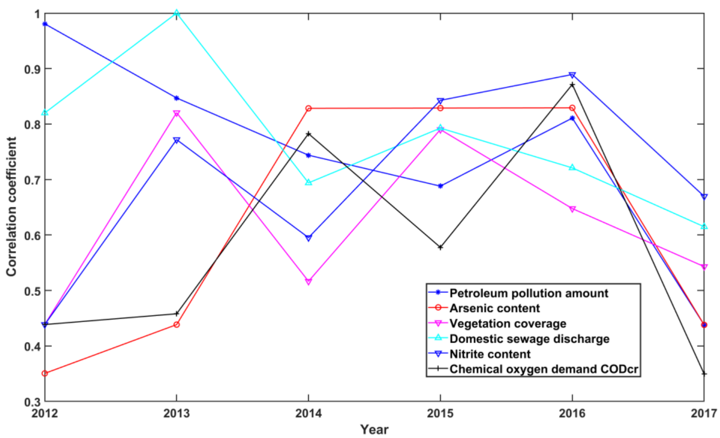

Fig. 2. Grey correlation coefficient.

The weighted grey correlation degree is obtained according to the following formula (3):

$$r_{\mathrm{j}} = \sum_{j=1} k_j \xi_t(j) \tag{3}$$

where r_j is the weighted grey correlation degree of the j^th evaluation index object, which is calculated by using the weight of AHP analytic hierarchy process. In order to avoid chance, it is necessary to accumulate the data of each selected year to find the mean. The final calculation results are shown in Table 2.

Table 2. The results of weighted grey correlation degree.

	2012	2013	2014	2015	2016	2017	$\overline{r_x} = \frac{1}{6}\sum_{t=1}^{6} r_{jt}$
r_1	0.19	0.17	0.15	0.14	0.16	0.08	0.148
r_2	0.04	0.05	0.10	0.10	0.10	0.05	0.073
r_3	0.07	0.13	0.09	0.13	0.11	0.09	0.103
r_4	0.16	0.2	0.12	0.16	0.14	0.12	0.150
r_5	0.08	0.14	0.10	0.15	0.15	0.12	0.123
r_6	0.05	0.05	0.09	0.07	0.10	0.04	0.066

Among them, r_{jt} represents the data of index j in the t^th year. According to the conclusion of weighted grey analysis, among the many influencing factors, the discharge of domestic sewage is the most important factor affecting water quality.

3 Prediction of domestic sewage discharge based on GM (1,1) model

Based on the above analysis results, and the domestic sewage discharge in Guangdong section of the Pearl River Basin is increasing year by year, this paper uses the domestic sewage discharge data from 2012 to 2017 (Table 3) to forecast GM (1,1) model.

Table 3. Total amount of domestic sewage discharged into the basin from 2012 to 2017 (100 million tons).

Year	2012	2013	2014	2015	2016	2017
Domestic sewage discharge	40.12	40.10	40.13	41.24	42.08	42.34

According to the class ratio test, the time series of domestic sewage discharge is established as follows, and the grade ratio $\lambda(t)$ is obtained.:

$$Y^{(0)}(t) = (Y^{(0)}(1), Y^{(0)}(2) \ldots Y^{(0)}(6)) = (40.12, \cdots, 42.34) \tag{4}$$

$$\lambda(t)=\frac{Y^{0}(t-1)}{Y^{0}(t)} \quad t=2\cdots 6 \tag{5}$$

For $\lambda(t)\in[0.75,1.33]$ $t=2\cdots 6$, $Y^{(0)}$ can be used to make a satisfactory GM(1,1) prediction. The prediction formula and the relative error test equation are obtained.

$$Y^{(1)}(t+1)=\left(Y^{(0)}(1)-\frac{b}{a}\right)e^{-at}+\frac{b}{a}=2475.19e^{-0.016t}-2435.07 \tag{6}$$

$$\delta(j)=\frac{\left|Y^{(0)}(t)-Y^{(1)}(t)-Y^{(1)}(t-1)\right|}{Y^{(0)}(t)} \tag{7}$$

the test results are calculated as shown in Table 4.

Table 4. Test conclusion.

Year	Original value	Predicted value	Residual error	Relative error (%)
2012	40.12	40.12	0.00	0.00
2013	40.10	39.90	0.20	0.49
2014	40.13	40.53	-0.40	0.99
2015	41.24	41.17	0.07	0.18
2016	42.08	41.82	0.26	0.62
2017	42.34	42.47	-0.13	0.32

It can be seen from the table that the average relative error is 0.43% and the maximum error is only 0.99%, which means that the prediction effect is good, the accuracy is high, and it has certain practicability. Figure 3 shows the prediction map of domestic sewage discharge in this basin in the next few years by using the prediction model proposed in this paper.

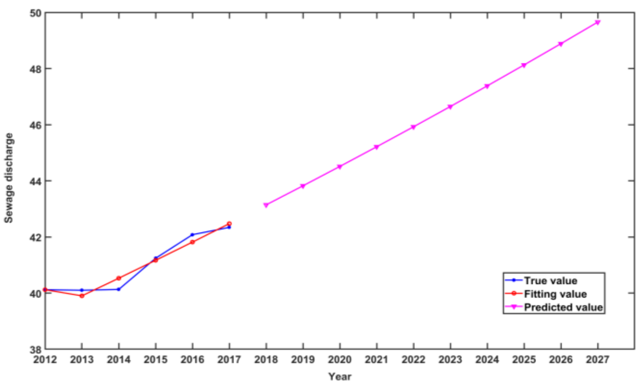

Fig. 3. Prediction of domestic sewage discharge.

4 Conclusions

In this paper, based on the water sample data of a branch river in the Guangdong section of the Pearl River Basin, a grey weighted correlation model is established to judge the influence of many factors, including domestic sewage, on river pollution. In the process of establishing the model, the AHP analytic hierarchy process is used to calculate the weight of each influencing factor, and the consistency test is used to judge the feasibility of the weight. The prediction model of domestic sewage discharge based on GM (1,1) is established. The average prediction error of the model is 0.43%, and the maximum prediction error is only 0.99%. This model can also be used for water sample analysis in other watersheds.

Acknowledgments

This research was funded in part by the Quality Engineering Project of Guangdong University of Science and Technology under Grant GKZLGC2022053, in part by the Dongguan Science and Technology of Social Development Program under Grant 20221800905782.

References

1. Yasir, M. (2020). Analysis of microbial communities and pathogen detection in domestic sewage using metagenomic sequencing. Diversity, 13(1), 6.
2. Liang, D., Darko, A. P., & Xu, Z. (2019). Pythagorean fuzzy partitioned geometric Bonferroni mean and its application to multi-criteria group decision making with grey relational analysis. International Journal of Fuzzy Systems, 21, 115-128..
3. Ceylan, Z., Bulkan, S., & Elevli, S. (2020). Prediction of medical waste generation using SVR, GM (1, 1) and ARIMA models: a case study for megacity Istanbul. Journal of Environmental Health Science and Engineering, 18, 687-697.
4. Liu, Y., Yang, L., & Jiang, W. (2020). Qualitative and quantitative analysis of the relationship between water pollution and economic growth: a case study in Nansi Lake catchment, China. Environmental Science and Pollution Research, 27(4), 4008-4020.
5. Pandey, S., Nautiyal, R., Kumar, P., Chandra, G., & Panwar, V. P. (2022). A grey relational model for soil erosion vulnerability assessment in subwatershed of lesser Himalayan region. Catena, 210, 105928.