Prediction of Industrial Hazardous Waste Production Based on Different Models

Shuangliu Liu, Liang Cheng, Peng Chen, Shunqing Xu, Jun Gao and Jing Lu*  
Chinese Academy for Environmental Planning, Beijing, China

*Corresponding author e-mail: lujing@caep.org.cn

Abstract. Based on national hazardous waste production data from 2008 to 2011, GM (1,1) grey model, unit industrial output value model and average annual growth rate model are used for prediction studies of which the prediction results was compared with the actual values of hazardous waste from 2012 to 2016. Under the selected background data, all the three widely-used models have large deviations. The model with the smallest deviation range is the annual growth rate model of which the prediction result is 1.25-2.05 times the actual value. The model with the second small deviation range is the unit output value model of which the prediction result is 1.36-3.63 times the actual value. GM (1,1) has the largest deviation range, of which the prediction result is 1.46-7.58 times the actual value. The causes of deviation are analyzed, and the results show that, due to the influence of different factors, the background data significantly rocketed in 2011 and 2016, resulting in the instability of the model prediction. Based on the application of the models, the applicability of different models is analyzed. It is suggested that the grey model should be preferred when there are a few data; and the easy-to-operate unit output value model and annual growth rate model shall be preferred when there are sufficient historical data. Meanwhile, the influence of industry and policy factors should be comprehensively considered to improve and optimize the prediction methods and parameters of the models, so as to provide reference for quantitative prediction of hazardous wastes.

1. Introduction

With the implementation of the Soil Pollution Prevention and Control Action Plan ("Soil Ten Measures"), the revision of the National Hazardous Waste List, and the release of the Soil Pollution Prevention and Control Law, the State requires stricter environmental supervision and risk prevention of hazardous wastes than ever before. In May 2018, the Ministry of Ecology and Environment kicked off a special initiative to crack down on violations of environmental laws concerning solid wastes (Waste Elimination Action), with the supervision and management of solid wastes regarded as the target of the main battlefield of environmental law enforcement. In the Opinions of the CPC Central Committee and the State Council on Comprehensively Strengthening Ecological Environmental Protection and Resolutely Promoting the Tough Fight for Prevention and Control of Pollution (Z.F. [2018] No. 17), the prevention and control of hazardous waste pollution is an essential part of “solidly advancing the battle for the protection of the pure land”. Scientific and reasonable prediction of hazardous waste production can provide data support and decision reference for government departments to conduct
hazardous waste management and guide the orderly development of hazardous waste pollution prevention and control. At present, there are mainly two methods for predicting the production of solid wastes[1-5]: one is to make prediction according to social and economic characteristics such as output value, population, slag ratio coefficient; the other is to make prediction based on mathematical statistical methods such as regression analysis, time series analysis, and grey model.

Currently, the existing prediction of hazardous waste production is mostly limited to a certain method. Few scholars compare and analyze several prediction methods or use known data to compare the prediction results with the actual values. Based on the actual value of industrial hazardous waste production in China and the data from 2008 to 2011, this study adopted three different models including grey model, unit gross industrial output value method and annual growth rate method, to predict the production in China and the data from 2008 to 2011. The prediction results are compared with the actual values. Based on the actual value of industrial hazardous waste production from 2012 to 2016. The prediction results are compared with the actual values, and the prediction accuracy and applicable conditions of different models are analyzed, so as to provide reference for quantitative prediction of hazardous wastes.

2. Modeling

2.1. GM (1,1) Grey Model
Grey model consists of high-order GM (n,1) model and multivariable GM (1,n) model. At a higher order, the calculation is more complicated, but the accuracy may be not high[6]. Thus, the first-order GM (1,1) model is the most widely used one [7-9].

The first-order differential equation corresponding to GM (1,1) is:

$$\frac{dx^{(1)}}{dt} + ax^{(1)} = u$$

Where, $x^{(1)}$ refers to a sequence obtained by one accumulated generating operation (AGO); $t$ refers to time; $a, u$ refers to undetermined coefficient.

The raw sequence is set as:

$$x^{(0)}(i) = \{x^{(0)}(1), x^{(0)}(2), ..., x^{(0)}(n)\}, i = 1, 2, ..., n$$

A new sequence generated by one accumulated generating operation:

$$x^{(1)}(i) = \sum_{m=1}^{i} x^{(0)}(m), i = 1, 2, ..., n$$

The coefficient is calculated through least square method, and the coefficient variable is recorded as:

$$\hat{a} = [a, u]^T = (B^TB)^{-1}B^TY$$

$$B = \begin{bmatrix}
-\frac{1}{2}[x^{(1)}(1) + x^{(1)}(2)] & 1 \\
-\frac{1}{2}[x^{(1)}(2) + x^{(1)}(3)] & 1 \\
\vdots & \vdots \\
-\frac{1}{2}[x^{(1)}(n-1) + x^{(1)}(n)] & 1
\end{bmatrix}$$

$$y_n = [x^{(0)}(2), x^{(0)}(3), ..., x^{(0)}(n)]^T$$

The differential equation is solved as follows:

$$\hat{x}^{(1)}(i + 1) = \left( x^{(0)}(1) - \frac{u}{a} \right) e^{-ai} + \frac{u}{a}$$

$$\hat{x}^{(0)}(1) = \hat{x}^{(1)}(1)$$

$$\hat{x}^{(0)}(i) = \hat{x}^{(i)}(1) - \hat{x}^{(i)}(i - 1), i = 1, 2, ..., n$$

The model accuracy test methods include posterior error test, relevance degree test and residual test. In this study, posterior error test is adopted.

The standard deviation of raw sequence $x^{(0)}(i)$ is:
The standard deviation of residual sequence $\epsilon(0)i$ is:

$$S_1 = \sqrt{S_1^2 / n - 1}$$

of which

$$S_1^2 = \sum_{i=1}^{n}[\epsilon(0)i - \bar{\epsilon}(0)]^2 = \frac{1}{n}\sum_{i=1}^{n}\epsilon(0)i$$

Thereby, the posterior error ratio is calculated:

$$C = \frac{S_1}{S_0}$$

Small error probability $P = \{ |\epsilon(0)i - \bar{\epsilon}(0)| < 0.6745 \cdot S_0 \}$

The grey model prediction accuracy classification standard is as described in Table 1:

| Small Error Probability (P) | Variance Ratio (C) | Prediction accuracy Grade |
|-----------------------------|-------------------|--------------------------|
| >0.95                       | <0.35             | Good                     |
| >0.80                       | <0.50             | Qualified                |
| >0.70                       | <0.65             | Barely qualified         |
| ≤0.70                       | ≥0.65             | Unqualified              |

2.2. Unit Gross Industrial Output Value Method

The unit gross industrial output value method is to calculate the predicted annual hazardous waste production after multiplying the predicted annual gross industrial output value by the production coefficient of hazardous wastes per unit of gross industrial output value[10]. The prediction model is as follows:

$$W_n = \epsilon_n \times M_n$$

(1)

$$\epsilon_n = \frac{W_0}{M_0} \times (1 - \delta_1)^n$$

(2)

$$M_n = M_0 \times (1 + \delta_2)^n$$

(3)

where, $W_n$ - predicted annual hazardous waste production, ten thousand tons;

$\epsilon_n$ - predicted annual emission intensity of hazardous wastes per unit of gross industrial output value, ten thousand tons / ten thousand Yuan;

$M_n$ - predicted annual gross industrial output value, ten thousand Yuan;

$\delta_1$ - attenuation coefficient, which mainly refers to the decline of hazardous waste emission intensity resulting from industrial transformation and chemical engineering regulation.

$\delta_2$ - annual average growth rate of gross industrial output value in the predicted period;

$n$ - period of prediction, years;

$W_0$ - base-year production of industrial hazardous wastes, ten thousand tons;

$M_0$ - base-year gross industrial output value, ten thousand Yuan.

2.3. Annual Average Growth Rate Method

In this method, the predicted annual hazardous waste production is calculated by multiplying the base-year industrial hazardous waste production by the annual average growth rate index of hazardous waste production in the predicted period. The prediction model is as follows:

$$W_n = W_0 \times (1 + \mu)^n$$

(4)

Of which, $W_n$ - predicted annual industrial hazardous waste production, ten thousand tons;
W₀ — base-year production of industrial hazardous wastes, ten thousand tons;
μ — annual average growth rate of industrial hazardous wastes in the predicted period.

3. Applied Research

3.1. Data Source
The data on industrial hazardous waste production were obtained from the *National Environmental Statistics Bulletin* over the years, and the data on gross industrial output value were obtained from *Statistical Bulletin on National Economic and Social Development* over the years. The basic data are shown in Table 2.

| Year | 2008 | 2009 | 2010 | 2011 | 2012 | 2013 | 2014 | 2015 | 2016 |
|------|------|------|------|------|------|------|------|------|------|
| Production of Hazardous Wastes (ten thousand tons) | 1357 | 1430 | 1587 | 3431 | 3465 | 3157 | 3633 | 3976 | 5347 |
| Industrial Output Value (ten thousand Yuan) | 146183 | 156958 | 186481 | 220592 | 235319 | 249684 | 271392 | 274278 | 296236 |

3.2. Application of Grey Model
Based on the original data on hazardous waste production from 2008 to 2011, a prediction model was built for the annual production of hazardous wastes from 2012 to 2016. The GM(1,1) prediction model is obtained by programming with matlab, as follows:

\[
\hat{x}^{(1)}(i+1) = 1559.44e^{0.52(i-1)} - 202.43
\]

The fitted values of the 2008-2011 model are 13.57, 10.64, 17.89 and 30.09 million tons respectively. The predicted values of hazardous waste production from 2012 to 2016 are shown in Table 3. After calculation, the small error probability P is 1, the variance ratio C is 0.29, and the relative residual Q is 0.12. According to Table 1, the prediction accuracy grade is good.

| Year | 2012 | 2013 | 2014 | 2015 | 2016 |
|------|------|------|------|------|------|
| Actual Value | 3465 | 3157 | 3633 | 3976 | 5347 |
| GM (1,1) | 5061 | 8513 | 14320 | 24086 | 40514 |
| Unit Output Value Model | 4711 | 6532 | 8170 | 12909 | 19397 |
| Annual Growth Rate Model | 4326 | 5456 | 6879 | 8675 | 10939 |

3.3. Application of Unit Output Value Model
Taking 2011 as the base year, the values of the parameters in Formulas (1) - (3) are assigned, among which W₀ refers to 34.31 million tons, M₀ refers to 22059.2 billion Yuan, δ₁ is assigned as -19.7%, the average reduction rate of the production coefficient of hazardous wastes per unit of output value from 2008 to 2011, δ₂ is assigned as 14.7%, the annual average growth rate of gross industrial output value from 2008 to 2011. The prediction results are shown in Table 3.

3.4. Application of Annual Average Growth Rate Model
Taking 2011 as the base year, the values of the parameters in Formula (4) are assigned, among which W₀ is 34.31 million tons, and μ is assigned as 26.1%, the annual average growth rate of industrial hazardous wastes from 2008 to 2011. The prediction results are shown in Table 3.
4. Analysis and Discussion

4.1. Comparison of Prediction results

By comparing the prediction results of the three models from 2012 to 2016 with the actual values, it is found that all of the prediction results have a very large deviation (Figure 1). The annual average growth rate model shows the smallest deviation range with its prediction result being 1.25-2.05 times the actual value. The unit output value model shows the second smallest deviation range with its prediction result being 1.36-3.63 times the actual value. The GM(1,1) shows the largest deviation range with its prediction result being 1.46-7.58 times the actual value.

In order to find out why the prediction results of those mathematical models widely used in the literature show such great deviation, the actual values of hazardous waste production are further analyzed. The production of hazardous wastes surged in 2011, from 15.87 million tons to 34.31 million tons, representing an increase of 116%. Then, the production of hazardous wastes grew by merely 3.8% on average from 2011 to 2015, and it surged again in 2016 with a growth rate of 34%. By analyzing the reasons for the two surges, consideration is given to the change in statistical scope in 2011 of which the hazardous waste production is extended from more than 10kg to more than 1kg. As for the surge in 2016, it may not be due to the growth in the actual production of hazardous wastes, but may be related to the large-scale implementation of the central environmental protection supervision since 2016. As the national supervision was strengthened, enterprises that concealed, underreported or secretly discharged hazardous wastes in the past reported and disposed of hazardous wastes in accordance with regulations, leading to a substantial increase in the production of hazardous wastes in the statistical scope. Prediction of the model is based on the hazardous waste production from 2008 to 2011 as historical data. However, due to the influence of statistical scope and policies, the actual waste production data from 2012 to 2016 are significantly different from the change trend of the selected base. Thus, prediction results of the three models show a large deviation. The GM(1,1) model is, essentially, an exponential function with $e$ as its base. As doubling of hazardous waste production in 2011 led to the steepened curve trend of fitting function, the GM(1,1) model shows the largest deviation.

![Fig 1. Comparison of prediction results of different models](image)

4.2. Model Applicability Analysis

Now grey GM (1,1) model is the most widely used in the prediction of hazardous waste production in the existing literature. The model is characterized by low demand for modeling data and high prediction accuracy. Only three or more data are needed to build the model. However, the building process of this
model is relatively complex, and it requires familiarity with programming software. In addition, under the condition of adequate historical data, the number of modeling data selected may have a great impact on the model function and accuracy. Taking the background data selected in this study as an example, and when modeling is based on the hazardous waste production from 2008 to 2011, the relative residual Q, variance ratio C, and small error probability P are respectively 0.12, 0.29 and 1 with which the prediction accuracy grade is good. When based on the hazardous waste production from 2008 to 2012, the three parameters are respectively 0.11, 0.39 and 0.8 with which the prediction accuracy is qualified. When based on the hazardous waste production from 2011 to 2015, the three parameters are respectively 0.04, 0.62 and 0.6 with which the prediction accuracy is unqualified. This shows that due to the difference in the laws of background data, not all regions’ hazardous waste production trends conform to the law of exponential growth. When selecting the prediction model of hazardous waste production, it is recommended to give priority to the grey model if there is a lack of background data.

The unit output value model and annual growth rate model are similar in essence, both of which embody the concept of compound growth rate. The unit output value model takes gross industrial output value as the main correlation factor affecting the production of hazardous wastes, and its building requires more parameters than the annual growth rate model. The advantages of the two models are simple in application, convenient in operation and small in deviation of prediction results. The disadvantage is that more historical data are needed to reflect the overall trend, and there are no specific parameters to measure when modeling the prediction accuracy. On the whole, in the absence of major industrial, population and policy changes, the growth trend of hazardous wastes is relatively stable, and when high availability is provided for historical data, the unit output value model or annual growth rate model is preferred, which is simple to calculate and can obtain relatively accurate prediction results.

5. Conclusion
a) Factors that affect the production of hazardous wastes within the statistical scope include population, gross industrial output value, industrial development planning, supervision, policy impacts, etc. In the case that certain factors cause great changes in hazardous waste production, the prediction results of the hazardous waste prediction model based on historical data will also be subject to large deviation.

b) To predict the production of hazardous wastes, the grey model is preferred with little historical data. When there are sufficient historical data, it is suggested that the unit output value model and annual growth rate model with simple operation should be given priority. Comprehensive consideration shall be given to the possible change factors including industries and policies to optimize and improve model prediction methods and parameters, so that the prediction results can fully support the decisions of the management department.

Acknowledgments
This work was financially supported by The Second National Census of Pollution Sources (22110399005) fund.

References
[1] GUO Ping, WANG Jinggang and ZHOU Bingyan. Prediction Research on Production of Industrial Hazardous Wastes in China [J]. Environmental Science & Technology, 2006, 29(2): 56-57.
[2] GUO Ping, ZHOU Bingyan and WANG Qi. Comparative Research on Production of Hazardous Wastes and Industrial Solid Wastes [J]. Environmental Pollution & Control, 2007(04): 320.
[3] Intharathirit R, Salam P A, Kumar S, et al. Forecasting of municipal solid waste quantity in a developing country using multivariate grey models[J]. Waste Manag, 2015, 39(Complete):3-14.
[4] Lin Y, Zhang J, Liu C. Generation Prediction and Countermeasure of Industrial solid waste in China[J]. Environmental Science & Management, 2008.
[5] ZHANG Qingxiang. Status Analysis and Countermeasure Research on Hazardous Waste
Pollution Prevention and Treatment in Shaanxi Province [D]. Chang’an University, 2014.

[6] ZHAO Rui, LIU Dan, LI Qibin et al. Prediction Research on Medical Waste Production [J]. Journal of Chengdu Medical College, 2008, 3(2): 115-119.

[7] Prediction Research on Production Status and Volume of Industrial Hazardous Wastes in Anhui Province [J]. Environmental Science and Management, 2016 (11).

[8] WANG Dongmei and HUI Yunmin. Prediction and Countermeasure Research on Production of Industrial Solid Wastes in Tianjin [J]. Environmental Science & Technology, 2010(s2): 655-657.

[9] HUI Yunmin and WANG Dongmei. Prediction and Countermeasure Research on Production of Industrial Solid Wastes in Binhai New Area [J]. Tianjin Science & Technology, 2012(2): 25-27.

[10] CUI Lingfeng, XU Bo, ZHU Huajun et al. Prediction of Hazardous Waste Production in One City of East China [J]. Chemical Industry Times, 2016, 30(6): 31-34.