Energy consumption prediction of cement production based on chaotic neural network-Markov chain

Run Yin¹, Chunxia Dou and Dong Yue
Institute of advanced technology, Nanjing University of Posts and Telecommunications, Nanjing 210023, China
¹ Email: 896603119@qq.com

Abstract. In order to improve the accuracy of energy consumption modeling and prediction in the cement production process, a cement production energy consumption prediction model based on chaotic neural network is proposed: 1. In the energy consumption modeling stage, chaotic neural network is used to reconstruct the phase of chaotic time series. Space, the chaotic neural network can still make high-precision predictions of the system even when the network input is incomplete or mutated, and the determination coefficient value is 0.019 higher than that of the RBF neural network. 2. In the energy consumption prediction stage, the introduction of Markov The residual error correction method is to correct the current forecast value based on the residual error between the historical predicted energy consumption value and the actual energy consumption value. The result shows that the relative residual error of the predicted value corrected by the Markov correction method decreases from -0.6% to -0.25 %, the predicted value of energy consumption is closer to the actual value. Based on the above description of the two stages of energy consumption modeling, the proposed cement production energy consumption prediction model has better prediction effects and higher prediction accuracy than traditional mechanism modeling in cement production energy consumption prediction.

1. Introduction
At present, cement production is an indispensable pillar industry for the development of my country's national economy, and at the same time as a high energy consumption industry, the demand for energy conservation and environmental protection in cement production is increasingly prominent. The energy consumption prediction of cement production is a prerequisite for reducing power consumption and improving the quality and efficiency of the cement production process [1]. Researchers at home and abroad have conducted research on the construction of energy consumption prediction models for cement furnaces and kilns, which can generally be divided into two types: mechanism-based modeling and data-based modeling [2]. In terms of mechanism modeling research, it is mainly from analysis Starting from the chemical and physical reactions of the calcination process in the kiln, based on the conservation of energy and mass[3], the precise mathematical expressions between the energy consumption and the main parameters of the cement kiln are studied and established. However, the calcination of the cement kiln involves many links, equipment and parameters. It has the characteristics of many variables and strong relevance. Mechanism modeling cannot accurately describe the relationship between the main parameters of the firing system and energy consumption. Therefore, data modeling is the current mainstream research direction, which mainly uses artificial intelligence methods such as expert systems, adaptive regression[4, 5], fuzzy systems, support vector...
machines[6, 7] and artificial neural networks[8] to study the energy consumption modeling of cement kilns. Because the complexity and randomness of energy consumption change factors are difficult to describe with mathematical models, this brings great difficulties to the accurate prediction of the model. Chaos is a common nonlinear dynamic behavior. The chaotic time series has internally determined regularity. The phase space of the chaotic attractor reconstructed by it has high-precision short-term predictability.

Markov process is often used to predict ship traffic flow, gas concentration, water demand, and stock trends. Lü Pengfei et al. [9] used BP neural network to establish a prediction model of ship traffic volume, and combined with Markov correction method to effectively improve the prediction accuracy. Han Tingting et al. [10] used the Markov correction method to modify the predicted value of the gray neural network model to make the predicted gas concentration change trend closer to the actual gas concentration change curve. Jing Yaping [11] combined the Markov correction method to establish a gray neural network urban water demand prediction model, and the test showed that it has obtained a prediction effect better than a single gray neural network prediction model. Wang et al. [12] established a fuzzy neural network prediction model based on Markov process, which accurately predicted the trend of stock indexes. This article can learn from the application of Markov correction in the above-mentioned fields to correct the predicted value of the network output.

In order to improve the accuracy of the prediction model, this paper establish a chaotic neural network prediction model. This model has very strong fault tolerance and associative memory functions, even when the network input is incomplete or mutated, it can not only accurately reproduce the non-linear dynamic behavior of the system, but also accurately predict the system in multiple steps within a certain range. When predicting energy consumption based on the model, the Markov correction method is used, that is, the energy consumption prediction value output by the network is corrected based on the residual of the historical energy consumption prediction value and the actual value, thereby improving the energy consumption prediction accuracy of cement production.

2. Energy consumption model of chaotic neural network

2.1. Reconstruction of the phase space of the chaotic attractor

This paper uses a simple and easy phase space reconstruction method [13]. In a nonlinear, large time delay system, the chaotic time series of univariate energy consumption values can be expressed as:

\[ x(t_j), j = 1, 2, 3, ..., n \]  

Where the time interval is \( \Delta t \). Since the structural characteristics of the chaotic attractor are included in the time series, in order to estimate the information of the system from the univariate time series, the energy consumption time series (1) is reconstructed into the m-dimensional phase space:

\[ Y(t_j) = (x(t_j), x(t_j + \tau), ..., x(t_j + (m-1)\tau)); j = 1, 2, ..., p \]  

\( m \) is the saturation embedding dimension of the phase space. There are \( p = \frac{n}{k} - (m-1) \) phase points in the m-dimensional phase space. When reconstructing the phase space, the saturation embedding dimension \( m \) and the delay time \( \tau \) need to be determined. Generally, \( \tau = \Delta t \) (time interval is \( \Delta t \)).

2.2. Obtaining the saturated embedding dimension \( m \) of reconstructed phase space

Take any \( m \) to construct the phase space. For all \( p \) phase points, if an arbitrarily small number \( \epsilon \) is given, calculate the Euclidean distance between all the phase points, and then compare the distance
\[ |Y(t_i) - Y(t_j)| < \varepsilon \] between how many point pairs there are. The ratio of the number of points with a distance less than \( \varepsilon \) in the total number \( p^2 \) of points is recorded as \( c(\varepsilon) \):

\[
c(\varepsilon) = \frac{1}{p^2} \sum_{i,j=1}^{p} \theta(\varepsilon - |Y(t_i) - Y(t_j)|)
\]

(3)

Among them, \( \theta(x) \) is the Heaviside function, in a certain area, when \( \varepsilon \) is sufficiently small, the following relationship \( c(\varepsilon) \) is satisfied:

\[
\ln c(\varepsilon) = \ln c(\varepsilon) + D \ln (\varepsilon)
\]

(4)

In the formula, \( C \) is a constant and \( D \) is a dimension: \( D = \lim_{\varepsilon \to 0} \frac{\ln c(\varepsilon)}{\ln (\varepsilon)} \), \( D \) is usually called the correlation dimension. Obtain \( \ln c(\varepsilon) \) and \( \ln (\varepsilon) \) curve in double logarithmic coordinates. The slope of the straight line segment of the curve is \( D \). As \( m \) increases, \( D \) gradually converges. When \( D \) converges, the corresponding dimension \( m \) is the saturated embedding dimension.

2.3. Obtaining the largest Lyapunov exponent and predictable scale

Set at time \( t_{i+1} = t_i + k\Delta t \), take the initial phase point \( Y(t_i) \) as the reference point in the \( m \)-dimensional phase space, according to the following formula

\[
L(t_i) = \min_{j \neq i} \left( |Y(t_i) - Y(t_j)| \right), i, j = 1, 2, \ldots, p
\]

(5)

and finally find the largest Lyapunov exponent, namely

\[
\lambda = \frac{1}{p-1} \sum_{i=1}^{p-1} \ln \frac{L(t_{i+1})}{L(t_i)}
\]

(6)

Obtain the largest Lyapunov exponent and calculate its reciprocal, which is called the predictable scale of the system. According to it, multi-step prediction of chaotic time series can be made.

2.4. Chaotic neural network prediction model

The chaotic time series has certain regularity inside, and the phase space of the chaotic attractor reconstructed by it has high-precision short-term predictability. That is, the predicted phase point \( Y(t_{p+1}) \) can be obtained by reconstructing the phase point \( Y(t_j)(j = 1, 2, \ldots, p) \) of the phase space, because the phase point contains only one unknown component \( x(t_{n+1}) \), so there is a mapping \( F \) such that

\[
x(t_{n+1}) = F(x(t_j)), j = 1, 2, \ldots, n
\]

(7)

In order to make the chaotic neural network model accurately reproduce the nonlinear dynamic behavior of the system, the number of model input variables \( M \) is set equal to the saturation embedding dimension \( m \) of the phase space, and the neural network is trained by the hierarchical genetic algorithm. The hierarchical genetic algorithm encodes the weight and structure (the selection of the number of neurons is \( N \)) at the same time. Through the global search, the weight and structure are optimized at the same time, and a strong fault tolerance and associative memory function is established. The neural network model can accurately construct the energy consumption model of
cement production even when the network input is incomplete or mutated, which provides a guarantee for the high-precision prediction of cement production energy consumption.

As long as it is within the predictable scale $\frac{1}{\lambda}$ of the system, multi-step prediction can be made according to the recurrence equation.

3. Markov correction of predicted value

The Markov chain prediction process describes the dynamic change process of a random time series. This process refers to the condition in the known moment $t_0$ that the state of the system or process at time $t$ ($t > t_0$) is only determined by the state of the time $t_0$ when the state of the time is known. It has nothing to do with the previous state [9-12]. The processing object of Markov process is discrete event data with large random volatility, which can be expressed in mathematical form as

$$P\{X_{k+1} = i_k | X_1 = i_1, X_2 = i_2, \ldots, X_k = i_k\} = P\{X_{k+1} = i_{k+1} | X_k = i_k\},$$  \hspace{1cm} (8)

Where $P$ is the conditional probability; $X_k$ is the sub-event; $i_k$ is the corresponding state.

The energy consumption value of the cement production process is sampled to obtain a set of discrete data in time series. This paper calculates the relative residual value of the actual energy consumption value $Y_1$ of each sample in the test sample set and the predicted value $Y_2$ of the network output, using Markov. The process establishes a residual correction model and corrects the energy consumption prediction value output by the model to make the current prediction value closer to the true value. The specific process is

Step 1: Compare the actual energy consumption value $Y_1$ of each sample in the test sample set with the predicted value $Y_2$ of the network output according to the time series, and find the relative residual $Z$ of the two as

$$Z = \frac{Y_1 - Y_2}{Y_1} \times 100\%$$  \hspace{1cm} (9)

Normalize the relative residual value to

$$Z^* = \frac{Z - Z_{\min}}{Z_{\max} - Z_{\min}}$$  \hspace{1cm} (10)

Step 2: According to the golden section method, the relative residual value is divided into n states $E_1, E_2, E_3, \ldots, E_n$ by size. The residual interval corresponding to n states is $Q_i = (a, b), i = 1, 2, \ldots, n$ ,

Step 3: Find the probability that the state $E_i$ transfers to a specific state $E_j$ only in one step, namely

$$P_{ij} = \frac{m_{ij}}{\sum_{j=1}^{n} m_{ij}}, i = 1, 2, \ldots, n ,$$  \hspace{1cm} (12)

Among them, $m_{ij}$ is the number of state $E_i$ transitions to state $E_j$ in the sequence. The one-step state transition probability matrix is formed by $P_{ij}$ , namely

$$A_1 = \begin{bmatrix} P_{i1} & P_{i2} & \cdots & P_{i n} \\ P_{i2} & P_{i2} & \cdots & P_{i n} \\ \vdots & \vdots & \ddots & \vdots \\ P_{in} & P_{in} & \cdots & P_{nn} \end{bmatrix}$$  \hspace{1cm} (13)

According to the C-K equation, the k-step state transition probability matrix is obtained as

$$A_k = \begin{bmatrix} P_{i1} & P_{i2} & \cdots & P_{i n} \\ P_{i2} & P_{i2} & \cdots & P_{i n} \\ \vdots & \vdots & \ddots & \vdots \\ P_{in} & P_{in} & \cdots & P_{nn} \end{bmatrix}^k = A^k ,$$  \hspace{1cm} (14)

Step 4: Establish a Markov chain prediction model as
\[ p_{k+1} = p_0 A^k, \quad (15) \]

Among them, \( p_0 \) is the probability distribution at the initial moment; \( p_{k+1} \) is the probability distribution at the \( k + 1 \) moment. The state and residual interval \( Q \in (Q_1, Q_2) \) corresponding to the time can be obtained from the probability distribution at the \( k + 1 \) moment, and the predicted value of the model is corrected according to equation (16).

\[ X_1 = \hat{X}/1 - Q_1; \quad X_2 = \hat{X}/1 - Q_2 \quad (16) \]

Among them, \( \hat{X} \) is the neural network prediction value; \( Q_1, Q_2 \) are the upper and lower limits of the corresponding residual interval respectively. The average value of the final \( X_1 \) and \( X_2 \) is the predicted value of the neural network after Markov correction.

4. Instance verification and data analysis

4.1. Energy consumption model of chaotic neural network

In order to establish an energy consumption model for cement production, it is necessary to conduct model training on a large number of sample data in the cement production process. Therefore, according to requirements, the sampling interval for each data set at the cement plant production site is 5 hours, and we randomly collect possible energy consumption dependent variable values (possible energy consumption dependent variables can be selected through simple mechanism analysis)[14-15], such as coal Material volume, CO volume percentage, raw meal flow rate, coolant explosion pipe pressure, air temperature entering the cooler, clinker flow rate and the corresponding electrical energy consumed in a certain period of time as a data set, a total of 2100 measurement data sets are used as sample data sets.

Using Matlab as the test platform, 100 data were randomly selected as the test set, and the remaining 2000 data were used as the modeling training set [16]. Create a chaotic neural network, with 8 energy-consuming dependent variables corresponding to 8 input nodes. According to the calculation formula of reference and some experience, the hidden layer of the neural network is generally 3 to 6 layers. The predicted value of energy consumption corresponds to an output node. And use hierarchical genetic algorithm to train neural network to achieve simultaneous optimization of weights and structure. Finally, the test was simulated, and the prediction result of the RBF neural network model was used as a comparison. The energy consumption prediction effect of the chaotic neural network is shown in Figure 1.

Figure 1. Energy consumption prediction results of chaotic neural network model.
The energy consumption prediction effect of the RBF neural network model is shown in Figure 2. The result of comparing the two is as follows:

Take the coefficient of determination as the evaluation index: \( R^2 = \frac{SSR}{SST} \), among them, \( SSR = \sum_{i=1}^{n} (\hat{y}_i - \bar{y})^2 \), \( \bar{y} \) is the mean value of all actual energy consumption values in the test set. \( SST = \sum_{i=1}^{n} (y_i - f_i)^2 \), \( f_i \) is the i-th predicted energy consumption value. The range of it is [0,1]. The closer the value is to 1, the greater the goodness of fit, that is, the higher the degree of interpretation of the input variable to the output, the higher the model accuracy.

The simulation results show that the coefficient of determination of the RBF model is 0.906, and the coefficient of determination of the chaotic neural network model is 0.92493, which is closer to 1, indicating that the model has good explanatory properties.

### 4.2 Markov correction of predicted value

**Table 1. Energy consumption prediction results of the integrated model.**

| sample | actual value /kWh | predictive value /kWh | relative residual /% | relative normalization | residual status |
|--------|-------------------|-----------------------|----------------------|------------------------|----------------|
| 1      | 14 033            | 14 220.799 03        | 1.3                  | 0.73                   | E3             |
| 2      | 15 784            | 15 371.366 10        | -2.6                 | 0.20                   | E1             |
| 3      | 14 189            | 14 656.310 36        | 3.3                  | 1.00                   | E3             |
| 4      | 13 077            | 13 432.284 19        | 2.7                  | 0.92                   | E3             |
| 5      | 16 003            | 15 569.137 32        | -2.7                 | 0.19                   | E1             |
| 6      | 14 981            | 14 369.311 54        | -4.1                 | 0                      | E1             |
| 7      | 15 273            | 15 394.603 02        | 0.7                  | 0.65                   | E2             |
| 8      | 12 549            | 12 822.296 76        | 2.1                  | 0.84                   | E3             |
| 9      | 14 667            | 14 854.968 56        | 1.2                  | 0.72                   | E2             |
| 10     | 14 367            | 14 545.776 58        | 1.2                  | 0.72                   | E2             |
| 11     | 15 398            | 15 084.425 57        | -2.0                 | 0.29                   | E1             |
| 12     | 16 112            | 16 009.663 94        | -0.6                 | 0.47                   | E2             |
| 13     | 14 783            | 14 888.319 29        | 0.7                  | 0.65                   | E2             |
| 14     | 14 769            | 14 547.065 94        | -1.5                 | 0.35                   | E1             |
| 15     | 16 011            | 15 922.499 27        | -0.6                 | 0.47                   | E2             |
In order to improve the prediction accuracy, the relative residual value of the actual value of the energy consumption of each sample in the test sample set and the predicted value of the network output is obtained according to the time series (Table 1), and normalized to [0, 1.00] (No. 5 columns), the average value of the relative residuals after normalization is 0.55. According to the rule of the golden section method (take $s = 1$), the energy consumption value is divided into three states $E_1$, $E_2$, $E_3$, the interval of $E_1$ is $[0, 0.47)$, and the interval of $E_2$ is $[0.47, 0.72]$, the interval of $E_3$ is $(0.72, 1.00]$. According to the description of the Markov correction process, The 14th energy consumption value is converted into the 15th energy consumption value in one step, and the 13th energy consumption value is converted to the 15th energy consumption value in the 2-step state transition probability matrix respectively:

$$A_1 = \begin{bmatrix} 0.2 & 0.6 & 0.2 \\ 0.4 & 0.4 & 0.2 \\ 0.5 & 0.25 & 0.25 \end{bmatrix}, \quad A_2 = \begin{bmatrix} 0.38 & 0.41 & 0.21 \\ 0.34 & 0.45 & 0.21 \\ 0.325 & 0.4625 & 0.2125 \end{bmatrix}.$$  

Similarly, the k-step state transition probability matrix can be obtained.

It can be seen from Table 1 that the 14th energy consumption value is in the state $E_1$, and its initial vector can be expressed as $P_0 = [1 \ 0 \ 0]$. After one step of transfer, the probability vector of the 15th energy consumption value is:

$$P_1 = P_0 \times A_1 = \begin{bmatrix} 0.2 & 0.6 & 0.2 \\ 0.4 & 0.4 & 0.2 \\ 0.5 & 0.25 & 0.25 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = [0.2 \ 0.6 \ 0.2].$$

Then the probability of the 15th energy consumption value being in the state $E_2$ is greater, and Markov correction is performed on the predicted value of the energy consumption model.

$$X_1 = \frac{\hat{X}}{1 - Q_1} = \frac{15922.5/1 + 0.006}{1} = 15827.5$$

$$X_2 = \frac{\hat{X}}{1 - Q_2} = \frac{15922.5/1 - 0.012}{1} = 16115.9$$

Among them, $\hat{X}$ is the predicted value of energy consumption for the 15th sample, which is 15 922.5; and $Q_1$, $Q_2$ is the lower limit and upper limit of the residual interval corresponding to the energy consumption value of the 15th sample, which are -0.6% and 1.2%. Finally, the average value of and is 15971.7, which is the revised predicted value, which is closer to the actual value of 16011. At this time, the relative residual of the two is reduced from the original -0.6% to -0.25%.

5. Conclusions

1) The chaotic neural network is used in the modeling, and the prediction coefficient is 0.92493. The result of the RBF neural network model is 0.906, indicating that the chaotic neural network energy consumption model has better interpretability and higher prediction accuracy than the RBF neural network model. The chaotic neural network model effectively realizes the accurate description of complex random energy consumption changes.

2) The energy consumption value predicted by the energy consumption model has been corrected by Markov and the relative residual error has decreased from -0.6% to -0.25%, indicating that the energy consumption forecast for cement production after Markov correction is closer to the actual energy consumption value.

3) In short, the use of a cement production energy consumption prediction model based on chaotic neural networks and Markov correction has a better prediction effect and higher prediction accuracy in the cement production energy consumption prediction, which is for the energy consumption supervision and energy saving of the cement production process. Provide a more accurate reference basis.
References

[1] Wang Junjie, Wei Liying, Wang Lan 2014 J Research on the evaluation system of energy saving and emission reduction technology in cement industry The silicate bulletin 33(5) 1268-1274

[2] Zhao Jie 2014 D Research on modeling and control of cement rotary kiln system Zhengzhou: Zhengzhou University

[3] Xie Sheng 2015 J Analyze the reasons for low coal consumption in cement kilns through heat balance calculations Cement 2015(1) 18-19

[4] Guo-Feng Fan, Li-Ling Peng, Wei-Chiang Hong and Fan Sun 2016 Electric load forecasting by the SVR model with differential empirical mode decomposition and autocorrelation Neurocomputing 173 958-970

[5] Zichen Zhang and Wei-Chiang Hong 2019 Electric load forecasting by complete ensemble empirical model decomposition adaptive noise and support vector regression with quantum-based dragonfly algorithm Nonlinear Dynamics 98 1107-1136

[6] Hu Qiao, He Zhengjia, Zhang Zhou, Suo Ziyang 2020 Sunsupport vector regression model hybridized with chaotic krill herd algorithm and empirical mode decomposition for regression task Neurocomputing 410 185-201

[7] Wei-Chiang Hong and Guo-Feng Fan 2019 Hybrid empirical mode decomposition with support vector regression model for short term load forecasting Energies 12(6) 1093

[8] Guo-feng Fan, Yan-Hui Guo, Jia-Mei Zheng and Wei-Chiang Hong 2020 A generalized regression model based on hybrid empirical mode decomposition and support vector regression with back propagation neural network for mid-short term load forecasting Journal of Forecasting 39(5) 737-756

[9] Lu Pengfei, Zhuang Yuan, Li Yang, et al 2017 J BP neural network-Markov forecasting model for ship traffic volume Journal of Shanghai Maritime University 38(2) 18-28

[10] Han Tingting, Wu Shiyue, Wang Pengjun 2014 J Gas concentration prediction based on Markov residual correction Industry and Mine Automation 40(3) 28-31

[11] Jing Yaping, Zhang Xin, Luo Yan 2011 J Combined forecast of urban water demand based on grey neural network and Markov chain Journal of Northwest Sci-Tech University of Agriculture and Forestry (Natural Science Edition) 39(7) 229-234

[12] WANG Yifan, CHENG SHihmin, SHU Meihua 2010 J Incorporating the Markov chain concept into fuzzy stochastic prediction of stock indexes Applied Soft Computing 10(2) 613-617

[13] Dong Jun, et al 1997 J Research progress and prospects of chaotic neural networks Information and Control 26(5) 360-367

[14] Zhao Zhengping 2018 D Research on thermal modeling of cement clinker firing system Changchun: Changchun University of Technology

[15] Hu Jingyang 2014 D Thermal balance analysis and system optimization research of cement rotary kiln Qingdao: Shandong University

[16] Gao Qian 2017 D Research on analysis methods of energy consumption characteristics based on big data Baoding: North China Electric Power University