**TELKOMNIKA Indonesian Journal of Electrical Engineering**
Vol. 12, No. 11, November 2014, pp. 7728 ~ 7737
DOI: 10.11591/telkomnika.v12i11.6278


# Research on Application of Sintering Basicity of Based on Various Intelligent Algorithms

**Song Qiang***, **Zhang Hai-Feng**
Mechanical Engineering Department, Anyang Institute of Technology, Anyang 455000, Henan
*Corresponding author, e-mail: songqiang01@126.com

***Abstract***

*Prediction of alkalinity in sintering process is difficult. Whether the level of the alkalinity of sintering process is successful or not is directly related to the quality of sinter. There is no good method, prediction of alkalinity by high complexity, the present nonlinear, strong coupling, high time delay, so the recent new technology, grey least square support vector machine have been introduced. In this paper, The weight of evaluation objectives has not given the corresponding consideration when solving the correlation degree by taking traditional grey relation analysis and it is with a lot of subjective factors, easily lead to mistakes in decision-making on program. What is more a kind of alkaline grey support vector machine model, enables us to develop new formulations and algorithms to predict the alkalinity. In the model, the data sequence of fluctuation is composed of grey theory and support vector machine is weakened, can deal with nonlinear adaptive information, combination and grey support vector machine these advantages. The results show that, the basicity of sinter, can accurately predict the small sample and reference information using the model. The experimental results show that, the grey support vector machine model is effective and with practical advantages of high precision, less samples, and simple calculation.*

***Keywords:*** *basicity in sintering process, grey relation analysis, grey least squares support vector machine, prediction, grey model*



## 1. Introduction

In the modern steel enterprises, the sintering process for blast furnace materials is one of the most important production processes. The sintering alkalinity has a direct effect on production and economic benefits of the whole steel enterprise [1]. Therefore, almost every steel factory is equipped with many instruments and automatic control systems in the sintering plant for the production process control. However, the complexity of the sintering process makes the process to be difficultly described by a set of mathematical models. Since this process often has large time delay and dynamic time variability, it is hard to perform control tasks of the whole sintering process using conventional control models.

Sintering process is a complex physical-chemistry process, which relates to a lot of characteristics including complicated mechanism, high nonlinearity, strong coupling, high time delay,and etc [1]. the mathematical model for whole sintering process cannot be established, thus we can only construct mathematical model for one of performance indicators, and the performance index of sintering process determines the policy of blending process. Because of the restriction of the detecting means, the chemical examination of sinter alkalinity generally needs 40 min. In the whole process, its time sometimes can even exceed 1 hour. Obviously, such a long time delay cannot meet the needs of actual productivity, and thus, the sinter alkalinity must be detected and a model for predicting the alkalinity should be established [1].

## 2. Grey Theory

### 2.1. Grey GM(1,1) Model

Grey theory is a method to study the small sample, poor information, uncertainty, with partial information known, part information unknown small sample, poor information uncertainty problem as the research object, the known information through data mining, to extract valuable information, a correct description of system behavior, evolution and effective monitoring and prediction of the position information system. Statistical prediction method has many

*Received May 15, 2014; Revised June 22, 2014; Accepted July 10, 2014*

advantages compared with the traditional method, it does not need to determine whether the forecast variables subject to normal distribution, don't need large sample statistics, that is to say the research object specially for the small sample, poor information uncertainty, do not need to change according to the change of input variables forecast model, produced by the grey sequence the grey system theory, think, although the objective system of complex data representation, scattered, but it always has a whole function, it must contain the inherent law of some. The key is how to choose the appropriate way to tap it and use it. All the grey sequence can be generated by a weakening the randomness, reveal the regularity. A differential equation is a unified model, differential equation model has higher prediction accuracy. The establishment of GM(1,1) model is basically a cumulative generation of the original data, so that the generated sequence has certain regularity, by establishing the differential equation model, obtain the fitting curve, and thus to predict the unknown parts of the system. The GM model is first of the original data, a accumulated, generate 1-AGO, accumulated data through data mining have certain regularity, the original data is not obvious regularity, and its development trend is swinging. If the original data were accumulated generating, its regularity is obvious

Assume that there is a time response sequence (which is called original time series),

$$x^{(0)} = \left(x^{(0)}(1), x^{(0)}(2), ....., x^{(0)}(n)\right) \tag{1}$$

Where x(0)(i) stand for the monitoring data an time I,i=0,1,2,….,n.

The forecast value $x^{(0)}(n+T), T = 1,2,\ldots, R.$ can be derived by the following three steps.

(1) Build up the first-order accumulating generator operator (AGO)

This step is to weaken the indeterminacy in the original time series and get a more regular time series. Let $x^{(1)}$ be the generated time series.

$$x^{(1)} = \left(x^{(1)}(1), x^{(1)}(2), ....., x^{(1)}(n)\right), x^{(1)}(k) = \sum_{i=0}^{k}\left(x^{(0)}(k)\right) \text{, k = 1, 2, …., n} \tag{2}$$

Where $x^{(1)}$ is the once Accumulated Generating Operation(1-AGO) sequence1-AGO

(2) Construct first-order linear differential equation, the whitenization differential equation can be obtained:

$$\frac{\mathrm{d}x^{(1)}}{\mathrm{d}t} + ax^{(1)} = u \tag{3}$$

Where, a is a developing coefficient, whose value reflects the variation relation of data; and u is grey action quantity, which is the most important difference between grey and common model.

Using the least squares estimation, a and u can be obtained:

$$\beta = \begin{bmatrix} a \\ u \end{bmatrix} = \left(B^{\mathrm{T}}B\right)^{-1}B^{\mathrm{T}}Y_N \tag{4}$$

Where,

$$B = \begin{bmatrix} -0.5\left(x^{(1)}(2) + x^{(1)}(1)\right) & 1 \\ -0.5\left(x^{(1)}(3) + x^{(1)}(2)\right) & 1 \\ \vdots & \vdots \\ -0.5\left(x^{(1)}(n) + x^{(1)}(n-1)\right) & 1 \end{bmatrix}$$

$$Y_N = \begin{bmatrix} x^{(0)}(2) \\ x^{(0)}(3) \\ \vdots \\ x^{(0)}(n) \end{bmatrix}$$

According to linear first-order differential equation, Equation (5) can be derived:

$$\hat{x}^{(1)}(k+1) = \left( x^{(0)}(1) - \frac{u}{a} \right) e^{-ak} + \frac{u}{a} \tag{5}$$

(3) Inverse the accumulation generation

Let $\hat{x}^{(0)}$ be the fitted and forecasted series,

$$\hat{x}^{(0)}(k) = \left\{ \hat{x}^{(0)}(1), \hat{x}^{(0)}(2), \ldots, \hat{x}^{(0)}(n), \ldots \right\} \tag{6}$$

Then, the predicted value can be calculated by Equation (7),

$$x^{(0)}(i) = \hat{x}^{(0)}(i) - \hat{x}^{(1)}(i-1), i = 1,2,\ldots,n \tag{7}$$

Where, $\hat{x}^{(0)}(1)$, $\hat{x}^{(0)}(2)$, ...., $\hat{x}^{(0)}(n)$ are fitted value of the original series; and $\hat{x}^{(0)}(n+1)$, $\hat{x}^{(0)}(n+2)$, ... are the forecast values.

(4) Error examination

The relative error can be calculated by Equation (8).

$$e(k) = \frac{x^{(0)}(k) - \hat{x}^{(0)}(k)}{x^{(0)}(k)} \times 100\ \% \tag{8}$$

Where, e is the error percentage.

**2.2. Residual Forecasting Model**

To evaluate modeling performance, we should do synthetic test of goodness:

$$C = \frac{S_2}{S_1} \tag{9}$$

Where $S^2_1 = \frac{1}{n} \sum_{k=1}^{n} (x^{(0)} - \bar{x}^{(0)})^2$; $S^2_2 = \frac{1}{n} \sum_{k=1}^{n} (\varepsilon(k) - \bar{\varepsilon})^2$.

Deviation between original data and estimating data:

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

$$P = P\{|\varepsilon(k) - \bar{\varepsilon}| < 0.6745 S_1\} \tag{10}$$

The precision grade of forecasting model can be seen in Table 1.

Finally, applying the inverse accumulated generation operation (AGO), we then have prediction values:

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

Table 1. Precision Grade of Forecasting Model

| Precision grade | *P* | C |
|---|---|---|
| Good | 0.95≤*p* | C≤0.35 |
| Qualified | 0.80≤*p*<0.95 | 0.35<C≤0.5 |
| Just | 0.70≤*p*<0.80 | 0.5<C≤0.65 |
| Unqualified | *p*<0.70 | 0.65<C |

### 2.3. Grey Relation Analysis

Grey relational is uncertain correlation between things, or uncertain correlation between the system factors, between the factors to principal act. The fundamental mission of grey relational analysis is geometry approach to the micro or macro basing on the sequence of behavioral factors, to analyze and determine the influence degree between each factors or the contribution measure of factor to the main behavior, and its fundamental idea is judge whether the geometry of the sequence of curves is closely linked according to the level of similarity of it, and the closer the curve, the greater the correlation of the corresponding sequences, conversely, the smaller[6]. The computational procedure of grey relational analysis is expressed below:

Assume,

$$X_i^0 = \left(x_i(1) - x_i(1), x_i(2) - x_i(1), x_i(3) - x_i(1), x_i(4) - x_i(1)\right) = \left(x_i^0(1), x_i^0(2), x_i^0(3), x_i^0(4)\right),$$

then:

$$\begin{bmatrix} X_0^0 \\ X_1^0 \\ X_2^0 \\ X_3^0 \\ X_4^0 \\ X_5^0 \\ X_6^0 \\ X_7^0 \\ X_8^0 \\ X_9^0 \end{bmatrix} = \begin{bmatrix} x_0^0(1) & x_0^0(2) & x_0^0(3) & x_0^0(4) \\ x_1^0(1) & x_1^0(2) & x_1^0(3) & x_1^0(4) \\ x_2^0(1) & x_2^0(2) & x_2^0(3) & x_2^0(4) \\ x_3^0(1) & x_3^0(2) & x_3^0(3) & x_0^0(4) \\ x_4^0(1) & x_4^0(2) & x_4^0(3) & x_4^0(4) \\ x_5^0(1) & x_5^0(2) & x_5^0(3) & x_5^0(4) \\ x_6^0(1) & x_6^0(2) & x_6^0(3) & x_6^0(4) \\ x_7^0(1) & x_7^0(2) & x_7^0(3) & x_7^0(4) \\ x_8^0(1) & x_8^0(2) & x_8^0(3) & x_8^0(4) \\ x_9^0(1) & x_9^0(2) & x_9^0(3) & x_9^0(4) \end{bmatrix} = \begin{bmatrix} 0 & -0.04 & 0.04 & 0.2 \\ 0 & -0.074 & -0.521 & -0.28 \\ 0 & -0.02 & -0.08 & 0.4 \\ 0 & -0.16 & -0.01 & -0.04 \\ 0 & 0.07 & 0.02 & 0 \\ 0 & 0.05 & -0.14 & -0.21 \\ 0 & 0 & 0.1 & 1 \\ 0 & 0.02 & 0.015 & 0.015 \\ 0 & -0.24 & 0.06 & 0.12 \\ 0 & -0.055 & -0.04 & 0.18 \end{bmatrix}$$

$$|s_i| = \left|\sum_{k=2}^{3} x_i^0(k) + \frac{1}{2}x_i^4(4)\right|$$ i=1,2,3,.....,9:

$s_0$=0.1;$s_1$=0.735;$s_2$=0.1;$s_3$=0.19;$s_4$=0.09;$s_5$=0.195;$s_6$=0.6;$s_7$=0.0325;$s_8$=0.12;$s_9$=0.805;

$$|s_i - s_0| = \left|\sum_{k=2}^{3}\left(x_i^0(k) - x_0^0(k)\right) + \frac{1}{2}\left(x_i^0(4) - x_0^0(4)\right)\right|, \text{ i=1, 2, .....9;}$$

Then we obtain:

$|s_1 - s_0|$=0.635; $|s_2 - s_0|$=0.2; $|s_3 - s_0|$=0.09; $|s_4 - s_0|$=0.19; $|s_5 - s_0|$=0.471;
$|s_6 - s_0|$=0.7; $|s_7 - s_0|$=0.1425; $|s_8 - s_0|$=0.02; $|s_9 - s_0|$=0.105;

$$\varepsilon_{0i} = \frac{1 + |s_0| + |s_i|}{1 + |s_0| + |s_i| + |s_i - s_0|}, \text{ i=1, 2, .....9;}$$

So we will obtain:

$\varepsilon_{01}$=0.7429; $\varepsilon_{02}$=0.8571; $\varepsilon_{03}$=0.9348;
$\varepsilon_{04}$=0.8633; $\varepsilon_{05}$=0.7333; $\varepsilon_{06}$=0.7081; $\varepsilon_{07}$=0.7717; $\varepsilon_{08}$=0.9839; $\varepsilon_{09}$=0.9478;

Then we may find
$\varepsilon_{08} > \varepsilon_{09} > \varepsilon_{03} > \varepsilon_{04} > \varepsilon_{02} > \varepsilon_{07} > \varepsilon_{01} > \varepsilon_{05} > \varepsilon_{06}$,
$X_8 > X_9 > X_3 > X_4 > X_2 > X_7 > X_1 > X_5 > X_6$

By the definition of the relational analysis, the allocation is sequenced by the relational coefficient. And the sequence of relational coefficient size is the order of ranking of the influence factor. Based on the above result, it is clear that $X_8$ = 0.8633 is maximum, and it represents that the relational degree is biggest between the first allocation center and the ideal allocation center. Therefore, the first allocation center is the most optimal choice.

$X_8$ is the optimal factors, $X_9$ ranked second, $X_6$ is the worst in all factors. That is to say, the CaO ratio of sinter basicity of pulverized coal ratio, basicity on sinter influence is relatively large, the thickness of the material layer and mixed FeO content in ore has very little effect on the basicity of sinter, might as well put the two operating variables from outside.

## 3. Least-Squares Support Vector Machines Algorithm Modeling

### 3.1. Least-squares Algorithm Support Vector Machine

Recently, least squares support vector machine (LS-SVM) has been applied to machine learning domain successfully. It is a promising technique owing to its successful application in classification and regression tasks. It is established based on the structural risk minimization principal rather than the minimized empirical error commonly implemented in the neural networks. LS-SVM achieves higher generalization performance than the neural networks in solving these machine learning problems. Another key property is that unlike the training of neural networks which requires nonlinear optimization with the danger of getting stuck into local minima, training LS-SVM is equivalent to solving a set of linear equation problem. Consequently, the solution of LS-SVM is always unique and globally optimal. In this study, the application of LS-SVM in the prediction of the alkalinity in sintering process was discussed [12-14].

Giving a training set $\{x_t, y_t\}_{t=1}^{N}$, with $x_t \in R^n$ and $y_t \in R$, where $x_t \in R^n$ is the input vector of the first *t* samples, $y_t \in R$ is the desired output value of the first *t* samples, and *N* is the number of samples, the problem of linear regression is to find a linear function *y*(*x*) ,which is equivalent to applying a fixed non-linear mapping of the initial data to a feature space.. In feature space, SVM models take the form:

$$y(x) = w^T \varphi(x) + b \tag{11}$$

Where, the nonlinear function mapping $\varphi(\square): R^n \to R^{n_h}$ maps the high-dimensional space into the feature space; and *w* is not a pre-specified dimensional, possibly infinite dimensional, b is a real constant.

The least squares approach prescribes choosing the parameters (*w*, *b*) to minimize the sum of the squared deviations of the data, and the square loss function is described as:

$$\min J(w,e) = \frac{1}{2} w^T w + \gamma \frac{1}{2} \sum_{t=1}^{N} e_t^2 \tag{12}$$

Where $\gamma$ is the trade-off parameter between a smoother solution, and training errors.

With constraints, $y(x) = w^{\mathrm{T}}\varphi(x_t) + b + e_t$, for $t$ = 1, …, $N$.

Important differences with standard SVM are the equality constraints and the squared error term, which greatly simplifies the problem.

Only equality constraints, and the optimization objective function is the error loss, which will simplify the problem solving.

To solve this optimization problem, one defines the following Lagrange function:

$$L(w,b,e,\alpha) = J(w,e) - \sum_{\mathrm{t}=1}^{\mathrm{N}} \alpha_{\mathrm{t}}\{w^{\mathrm{T}}\varphi(x_t) \tag{13}$$

Where, $\alpha_{\mathrm{t}}$ is an Lagrange multipliers. By Karush-Kuhn-Tucker (KKT) optimal conditions, the conditions for optimality are:

$$\begin{cases} \dfrac{\partial L}{\partial w} = 0 \to w = \sum_{t=1}^{N} \alpha_t \varphi(x_t) \\ \dfrac{\partial L}{\partial b} = 0 \to \sum_{t=1}^{N} \alpha_t = 0 \\ \dfrac{\partial L}{\partial e_t} = 0 \to \alpha_t = \gamma e_t \\ \dfrac{\partial L}{\partial \alpha_t} = 0 \to w^T \varphi_t + b + e_t - y_t = 0 \end{cases} \tag{14}$$

Where, $t$ = 1, …, $N$.

After elimination of $\mathrm{e}_{\mathrm{t}}$ and w, the solution is given by the following set of linear equations.

$$\begin{bmatrix} 0 & \boldsymbol{1}^T \\ \boldsymbol{1} & \varphi(x_t)^T\varphi(x_t) + \boldsymbol{D} \end{bmatrix} \begin{bmatrix} b \\ \boldsymbol{\alpha} \end{bmatrix} = \begin{bmatrix} 0 \\ \boldsymbol{y} \end{bmatrix} \tag{15}$$

Where, $\boldsymbol{y} = [y_1, \cdots, y_N]$, $1 = [1, \cdots, 1]$, $\boldsymbol{\alpha} = [\alpha_1, \ldots, \alpha_{\mathrm{N}}]$, $\boldsymbol{D} = diag[\gamma_1, \cdots, \gamma_{\mathrm{N}}]$. Select $\gamma$ >0, and guarantee matrix $\boldsymbol{\varphi} = \begin{bmatrix} 0 & \boldsymbol{1}^T \\ \boldsymbol{1} & \varphi(x_t)^T\varphi(x_t) + \boldsymbol{D} \end{bmatrix} \begin{bmatrix} \mathrm{b} \\ \boldsymbol{\alpha} \end{bmatrix} = \boldsymbol{\varphi}^{-1} \begin{bmatrix} 0 \\ \boldsymbol{y} \end{bmatrix}$.

### 3.2. Selection of the Kernel Function

By the KKT-optimal conditions, *w* is obtained, and thus the training sets of nonlinear approximation is obtained too.

$$y(x) = \sum_{t=1}^{N} \alpha_t k(x, x_t) + b \tag{16}$$

Where x, $x_i$ denote training point and support vector respectively,y is the output of network.. $\alpha_i$ and b are the solutions to Equation (13).

$$k(x_t, x_k) = \varphi(x_t)^T \varphi(x_k)\ t, k = 1, \cdots, N \tag{17}$$

The Selection of the kernel function $\varphi(\square): R^{\mathrm{n}} \to R^{n_{\mathrm{h}}}$ has several possibilities. It is arbitrary symmetric function which can the Mercer theorem. In this paper, the radial basis

function (RBF) is used as the kernel function of the LS-SVM, because RBF kernels tend to give good performance under general smoothness assumptions, since Gaussian RBF (Radial Basis Functiong, RBF) function is usually used as a kernel function [7],

$$K(x, x_t) = \exp\{-\| x - x_t \|_2^2 / 2\sigma^2\} \quad (18)$$

Where, $\sigma$ is a positive real constant.

By Equation (8) to Equation (10), the object nonlinear model is as follows:

$$y(x) = \sum_{t=1}^{N} \alpha_t \exp\{-\| x - x_t \|_2^2 / 2\sigma^2\} + b \quad (19)$$

The LS-SVM prediction involves two parameters to be optimized, which are $\sigma$ (the width of the Gaussian kernels which cover the input space) and $\gamma$ is viewed as regularization parameters,which controls the tradeoff bewteen complexity of the machine and the number of non-separable points [8].

LS-SVM is an efficient version of these improved SVM,Instesd of a quadratic programming problem in standard SVM,a set of linear equations based on KKT optimization condition are solved in LS-SVM,which can reduce the computational complexity and time for training to a certain extent [9, 11].

## 4. Grey Least Squares Support Vector Machines

The object of both grey forecast and SVM forecast study is small sample prediction. Although they build on the basis of different theories, there are some similarities between GM(1,1) and SVM. grey GM(1,l) by identifying the parameters of the model is actually based on the least squares linear regression, whereas support vector machine is evolved from the linear optimal surface. Both models have their own advantages and weaknesses, grey GM(l,l) is a model of the differential equation, which strengthens the regularity of raw data by cumulative generation; moreover, it is the fitting of exponential curve. Based on risk minimization, there must be over-fit, and support vector machine is a theory based on structural risk minimization, which has very good genaralization ability. If the two combine to enhance the regularity of raw data by accumulative generation, and to identify model parameters, at the same time to adopt structural risk minimization, which consolidates the advantage of the two models and can obtain better forecast accuracy.

Based on the above analysis, a new idea or algorithm is put forward. A grey support vector machine model is proposed to overcome these limitations on the basis of the forecasting models. The fluctuation of data sequence is weakened by the grey theory and the support vector machine is capable of processing nonlinear adaptable information, and the grey support vector machine is a combination of those advantages. Above all, grey theory is used to conduct a cumulative sequence of the raw data, and the least squares support vector machine is adopted for the process and prediction.

New algorithm design steps as follows:

(1) Firstly, the original sequence $X^{(0)} = \{x^{(0)}(1), x^{(0)}(2), \cdots, x^{(0)}(n)\}, x^{(0)}(i) \geq 0, i = 1,2, \cdots n$, and a sequence generated a cumulative production, as follows:

$$X^{(1)} = \{x^{(1)}(1), x^{(1)}(2), \cdots, x^{(1)}(n)\}, x^{(1)}(i) \geq 0, i = 1,2, \cdots n$$

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

(2) Secondly, select Kernel function $K(x_i, x)$;

(3) Solving optimization problems Eqn.(8) using support vector machine method.

(4) Build up regression function $y(x)=\sum_{i=1}^{N}\alpha_i k(x,x_i)+b$;

(5) Construct cumulative sequence and get $\widehat{X}^{(1)}$,where $\widehat{X}^{(1)}$ is the predictive value;

(6) $\widehat{X}^{(1)}$ by the cumulative reduction, by the original data sequence $X^{(0)}$ forecast model $\widehat{X}^{(0)}(k+1)=\widehat{X}^{(1)}(k+1)-\widehat{X}^{(1)}(k), k=n+1,n+2,\cdots$.

(7) Finally, model test.

## 5. Sinter Alkalinity Forecasts and Simulation Based on grey Least Squares Support Vector Machines

### 5.1. System Input Parameters Selection

The grey least support vector machine is used to predict the alkalinity, aiming at this important output index. In the whole process, the variables related to the alkalinity is synthesized and make sure ten important input variables as the input of grey neural network, such as the layer thickness, the trolley speed, adding water rate of the first mixture, the mixing temperature, the content of $SiO_2$ in the mineral, the content of CaO in the mineral, the content of FeO in the mineral, adding water rate of the second mixture, the proportion of CaO, and the proportion of coal.

### 5.2. Sample Data Processing

Because all the collected data is often not in the same order of magnitude, the collected data are normalized to [-1 1]; this will improve the training speed of neural network. We often uses the following formulo to cope with the initial data:

$$x_{ij}^{'}=\frac{x_{ij}-x_{j\min}}{x_{j\max}-x_{j\min}}\times 0.8+0.1 \tag{20}$$

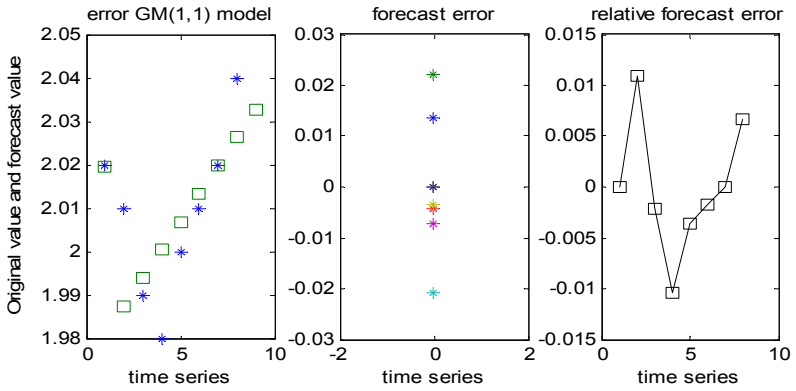
Where $x_{ij}^{'}$ and $x_{ij}$ were the old and new value of the variable for a sampling point respectively, $x_{j\max}$ and $x_{j\min}$ were the minimum and maximum value of the variable in the original dataset.

After computing using the neural network, the antinormalization processing is done to obtain the actual value of the output value forecasting.

Antinormalization of the following formula:

$$x_{ij}=(x_{j\max}-x_{j\min})\times(x_{ij}^{'}-0.1)+x_{j\min} \tag{21}$$

In the sampling dataset, there were inevitably some anomalies , and these data would give an impact on a certain model, and even lead to misleading. Therefore, the model data used in training samples dataset and test samples dataset was carefully selected.




文中

Figure 1. The Prediction of the Alkalinity Based on Grey GM(1,1)

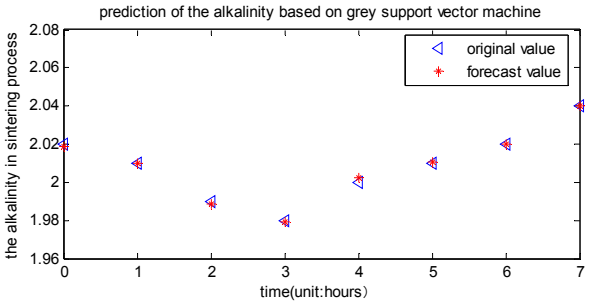



Figure 2. The Prediction of the Alkalinity Based on Grey Least Squares Support Vector Machines

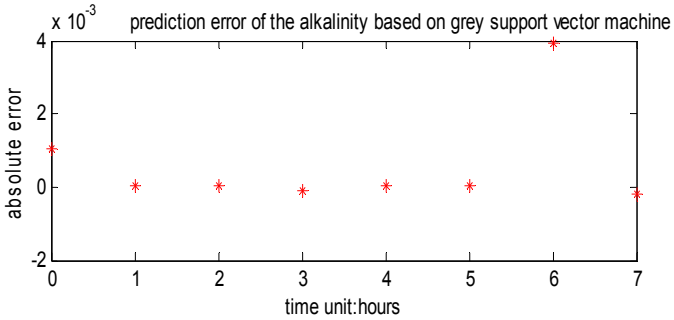



Figure 3. The Error Curve of the Alkalinity Based on Grey Least Support Vector Machine

Figure 1 illustrates the prediction error and relative prediction error of GM(1,1) model of the alkalinity, Figure 2 presents the fitting curve bases on grey least squares support vector machine. It can clearly seen that the predicted values are in good agreement with the desired ones in the whole ranges of time step, while Figure 3 show prediction error bases on grey least squares support vector machine.

From Table 1, the accuracy of the grey support vector machine reaches 0.273%, whereas the accuracy of GM(1,1) approach only is around -3.206%. It is not difficult to see that the forecast accuracy of grey least square support vector machine is higher than that of a single GM(1, l) model or the model of SVM, and has better robustness.Therefore, we can conclude that the grey least squares support vector machine exhibits excellent learning ability with fewer training data,the generalization capability of LS-SVM is greatly improved.

Table 2. The Comparison of the Alkalinity Based on Grey GM(1,1) and Grey Least Support Vector Machine

| Number | Original data | GM(1,1) model | | grey least support vector machine | |
|---|---|---|---|---|---|
| | | Model data | Relative error/% | Model data | Relative error/% |
| 1 | 2.02 | 2.02 | 0 | 2.0202 | -0.01 |
| 2 | 2.01 | 1.9878 | +1.1 | 2.009 | -0.05 |
| 3 | 1.99 | 1.9942 | -2.11 | 1.9887 | +0.065 |
| 4 | 1.98 | 2 | -1.01 | 1.9791 | +0.015 |
| 5 | 2 | 2.0070 | -0.35 | 1.9971 | +0.15 |
| 6 | 2.01 | 2.0135 | -0.174 | 2.0098 | +0.01 |
| 7 | 2.02 | 2.02 | 0 | 2.0188 | +0.059 |
| 8 | 2.04 | 2.0265 | -0.662 | 2.0393 | +0.034 |
| Average relative error/% | | -3.206 | | 0.273 | |

## 5. Conclusion

This paper has proposed a mathematics model of the alkalinity, which is realized via grey least squares support vector machine. This algorithm combines the advantages of GM(1,1) and LS-SVM. The new model fully makes use of the advantages of accumulation generation of GM(1,1) method, and weakens the effect of stochastic disturbing factor in original data series, and strengthens the regularity of raw data, and avoids the theoretical defects existing in the grey forecasting model. Besides, SVM's ability to handle high-dimension and incomplete data allows successful extraction of information even when part of the data records was missing or unreasonable owing to the problems of instrument malfunction or maintenance, calibration and climate influences, so LS-SVMs method is suitable to simulate the alkalinity in an efficient and stable way.

These results fully demonstrate the prediction accuracy of new model is superior to a single model, the theoretical analysis and simulation results are fully presented the validity of the forecast model. This shows that the grey least squares support vector machine is available for the modeling of the alkalinity, and can get better performance.

Although the proposed grey LS-SVM-based model may be superior to other modelling methods in some aspects, it has some potential drawbacks such as the underlying Gaussian assumptions related to a least squares cost function. Some researchers have made some efforts to overcome these by applying an adapted form called weighted LS-SVM. So we will intend to continue the studies on the application of the alkalinity in sintering process.

## References

[1] FAN Xiao-hui, WANG Hai-dong. Mathematical Model and Artificial intelligence of sintering process. Central South University. 2002.
[2] LIU Si-feng, DANG Yao-guo, FANG Zhi-geng. Grey Systems theory and appilication. China science press. 2004.
[3] LI Guo-zheng, WANG Meng, ZENG Hua-jun. An Introduction to support Vector machines and other Kernel-based Learning Methods. Publishing House of Electronics Industry. 2006.
[4] Evelio H, Yaman A. Control of Nonlinear Systems Using Polynomial ARMA Models. *AIChE Journal.* 1993; 39(3): 446.
[5] Narendra K S, Parthasarathy K. Identification and Control of Dynamical Systems Using Neural Networks. *IEEE Trans Neural Networks.* 1990; 1(1): 4-27.
[6] WANG Xu-dong, SHAO Hui-he. Neural Network Modeling and Soft-Chemical Measurement Technology. *Automation and Instrumentation.* 1996; 23(2): 28 (in Chinese).
[7] Cortes C, Vapnik V. Support Vector Machine. Machine Learning. 1995; 20(3): 273.
[8] WANG Yong, LIU Ji-zhen, LIU Xiang-jie et al. Modelling and Application of soft-sensor based on least squares support vector machines of oxygen-content in flue gases of plan. *Micro-computer information*. 2006; 10: 110-115.
[9] Kecman V. Learning and Soft Computing. Cambridge: The MIT Press. 2001.
[10] CHEN Xiao-fang,GUI Wei-hua,WANG Ya-lin etl. Soft-sensing model of sulfur content in agglomerate based on intelligent integrated strategy. *Control Theory & Application*. 2004; 21(1): 75～80.
[11] ZHANG Xue-gong. On the Statistical Learning Theory and Support Vector Machines. *Automation Journal*. 2000, 26(1): 1 (in Chinese).
[12] LI Guo-zheng, WANG Meng, ZENG Hua-jun. Support Vector Machine Introduction. Beijing: Electronics Industry Publish Press. 2004 (in Chinese).
[13] Mozer MC, Jordan MI, Petsche T, et al. Advances in Neural Information Processing Systems. 1997; 9(8): 134.
[14] Nello Cristianni, John Shawe-Taylor. An Introduction to Support Vector Machines and Other Kernel-Based Learning Method. Beijing: China Machine Press. 2005.