

## Article

# Mold Level Predict of Continuous Casting Using Hybrid EMD-SVR-GA Algorithm

Zhufeng Lei \*  and Wenbin Su

School of Mechanical Engineering, Xi'an Jiaotong University, 28 West Xianning Road, Xi'an 710049, China; wbsu@mail.xjtu.edu.cn

\* Correspondence: leizhufeng@stu.xjtu.edu.cn

Received: 22 February 2019; Accepted: 22 March 2019; Published: 26 March 2019



**Abstract:** The prediction of mold level is a basic and key problem of continuous casting production control. Many current techniques fail to predict the mold level because of mold level is non-linear, non-stationary and does not have a normal distribution. A hybrid model, based on empirical mode decomposition (EMD) and support vector regression (SVR), is proposed to solve the mold level in this paper. Firstly, the EMD algorithm, with adaptive decomposition, is used to decompose the original mold level signal to many intrinsic mode functions (IMFs). Then, the SVR model optimized by genetic algorithm (GA) is used to predict the IMFs and residual sequences. Finally, the equalization of the predict results is reconstructed to obtain the predict result. Several hybrid predicting methods such as EMD and autoregressive moving average model (ARMA), EMD and SVR, wavelet transform (WT) and ARMA, WT and SVR are discussed and compared in this paper. These methods are applied to mold level prediction, the experimental results show that the proposed hybrid method based on EMD and SVR is a powerful tool for solving complex time series prediction. In view of the excellent generalization ability of the EMD, it is believed that the hybrid algorithm of EMD and SVR is the best model for mold level predict among the six methods, providing a new idea for guiding continuous casting process improvement.

**Keywords:** empirical mode decomposition; support vector regression; genetic algorithm; mold level; continuous cast

## 1. Introduction

In the modern steel industry, high efficiency continuous casting technology has become the most internationally competitive core technology. The continuous casting process is a complex and continuous phase change process. There are many factors that affect the quality of slabs [1]. The research of the key technologies and cores in the high quality steel continuous casting process is mainly focused on mold level precision, the segment, and secondary cooling dynamic control [2].

Mold level is non-linear and non-stationary in terms of the time scale and does not satisfy Gaussian normal distribution. Therefore, the development and adoption of an effective signal processing method to predict the Mold level is hugely challenging. This brings great difficulties for the prediction of Mold level. The prediction of Mold level is crucial for improving the adaptive control of the continuous casting process. Therefore, an accurate prediction of mold level cannot only guarantee the quality of slab products, but also improve the automation level of the continuous casting manufacturing industry.

Precise mold level monitoring is regarded as the key to improving continuous casting production quality [3,4]. The mold level is an important reference for casting speed control, segment roll gap control, mold cooling water control, and stopper rod opening control. In the continuous casting production process, the fluctuation of mold levels will cause large amounts of slag in the mold to be involved in the molten steel, which will seriously affect the quality of the slab and may even lead

to accidents in the casting process, such as slab breakout and steel overflow at the top of the mold. Continuous high mold level operation will lead to overflow accidents, where the impurities float on the liquid surface, resulting in surface defects of strand and internal defect of the cast product, which in turn affects the surface and internal quality of the slab. Lowering the casting speed results in excessive fluctuations in the mold level, which affects the productivity and production rhythm. This impacts the quality of the slab and causes unplanned shutdown because of the stick and damage of the tundish slide-gate.

The mold level predict model can control and maintain the mold level according to the mold level historical data when the casting speed is disordered. As shown in Figure 1, after the transition of tundish, hot metal enters into mold through slide-gate, and the lower end of slide-gate is under the surface of mold. Hot metal transforms into solidified shell in mold and enters the root roll segment through the narrow surface of mold. When the casting speed is over-low, the hot metal level in mold is over-high, will cause overflow accident; when the casting speed is over-high, the hot metal level in mold is over-low, which leads to the bulging of root roll segments and cause results in a break-out accident. The stopper rod can control the flow rate of the hot metal when it flows into the mold.

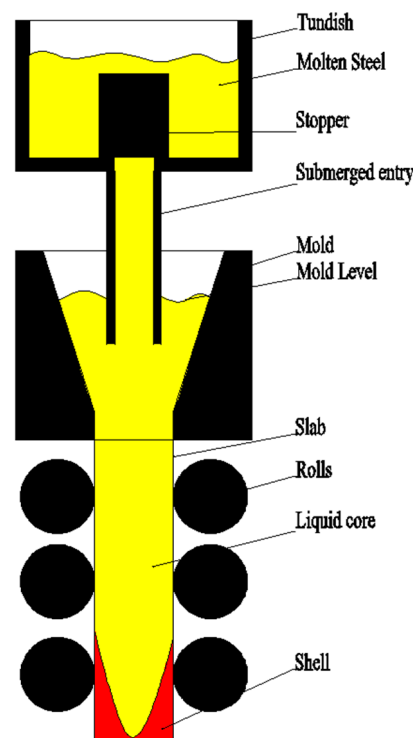


Figure 1. Mold level model.

Some researchers have adopted many methods for the prediction of time series. Numerical weather predict model is used for predicting future wind speed using mathematical models [5], multiple regression, exponential smoothing, autoregressive moving average model (ARMA), and many others for wind speed predict, power predict, and stock trend predict. Weron et al. [6] explained the complexities, strengths, and weaknesses of the available solutions for electricity price predicting, the opportunities and threats that predicting tools provide. However, the traditional time series predict methods, such as regression analysis and grey predict [7], have some shortcomings, where the prediction accuracy of signals with large fluctuations need to be improved [8].

In recent years, due to the rapid development of science and technology, artificial intelligence technology has been widely used and introduced into the prediction of time series [9]. Artificial neural networks (ANN) [10] and support vector regression (SVR) [11] methods are the main tools for dealing

with non-linear, non-stationary time series. ANN is an artificial intelligence method developed in the 1940s that can simulate human brain biological processes [9]. Wang et al. [12] proposed a new back propagation neural network algorithm to apply to a semi-distributed model. Fei He et al. [13] gives an advance artificial intelligent technology based on the genetic algorithm (GA) and the back propagation (BP) neural networks. SVR is a small sample machine learning method, based on statistical learning theory, Vapnik-Chervonenkis (VC) dimension theory, and minimum structural risk principles [10]. Based on limited sample information, it seeks the best compromise between model complexity and learning ability to achieve the best promotion effect [14,15]. Y. Liu established a method for online predictions of the silicon content in blast furnace ironmaking processes [16]. Silvano Cincotti et al. [17] used the SVM model to forecast the electricity spot-prices of the Italian power exchange (IPEX), which provided a better prediction accuracy, closely followed by econometric technique. Existing studies have shown that the ANN method takes a long time to calculate and is prone to localized minimization [3,18–20], leading to overfitting and poor predict results. SVR is more adaptable to overfitting than ANN because the parameters of SVR can be improved by means of global optimization.

Many studies show that ANN or SVR methods were employed for the prediction of time series but only a few pieces of literature have combined the two methods [21]. For that reason, we apply the combination of empirical mode decomposition (EMD) and SVR for prediction of time series. The concept of hybrid predict appears in numerical weather prediction (NWP)-based predicts, such as NWP-ANN/ARMA for solar radiation predict [20]. Ye and Liu [22] used EMD-SVR for short-term wind power predict and achieved good predict results.

In this paper, we present a novel predict method for time series of mold level, based on the combination of EMD and SVR with global optimization. The results of simulation experiments display their effective and competitive advantages by using the proposed hybrid algorithm. First, the original mold level signal is decomposed by EMD into several intrinsic mode functions (IMFs). Then the improved SVR model is optimized by GA and used to predict the subsequences. Finally, the predict sequence is reconstructed to obtain the predict result. The rest of this paper is organized as follows. In Section 2, the basic algorithms EMD and SVR are introduced. In Section 3, we present a novel predict method for time series of mold level based on the combination of EMD and SVR with global optimization. Section 4 consists of experimental results and analysis. Finally, we conclude our work in Section 5.

## 2. Basic Algorithm Research

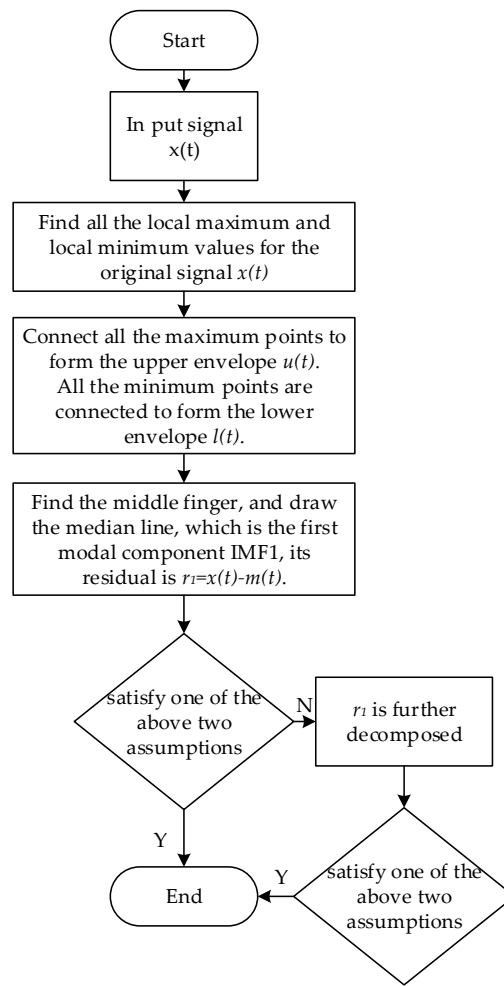
### 2.1. EMD Algorithm

EMD is an adaptive signal processing technique suitable for non-linear and non-stationary processes [23]. In 1998, Huang et al. [24] proposed the empirical mode decomposition technology, which has been widely used in biomedicine [25,26], speech recognition [27], system modeling [28–30], and process control [31,32]. Based on the time scales, EMD local features, such as local maxima, local minima, and zero-crossings, decompose the signal into several IMFs and a residual, the IMFs are orthogonal to each other. Modal decomposition is determined by the signal itself.

EMD satisfies the following basic assumptions:

- (1) In the entire data set, the number of extreme values and the number of zero crossings must be equal or at most have one point of difference.
- (2) At any point, the average is defined by the local maximum envelope, and the minimum envelope is zero.

EMD steps are shown in Figure 2.



**Figure 2.** The empirical mode decomposition (EMD) algorithm.

Finally, the original signal is decomposed into:

$$x(t) = \sum_{i=1}^N c_i + r_N \quad (1)$$

where  $x(t)$  is the original signal,  $c_i$  is the IMF,  $N$  is the number of IMFs and  $r_N$  is the residual.  $i = 1, 2, \dots, N$ .

## 2.2. SVR Algorithm

SVR does not solve only the classification problem, but also solves the regression problem. The basic model is the largest linear classifier as defined in the feature space [33]. SVR aims to achieve a distinction between the samples by constructing a hyperplane for classification, so that the sorting interval between the samples is maximized and the sample to the hyperplane distance is minimized.

Set a training data set for a feature space:

$D = \{(x_1, y_1), (x_2, y_2), \dots, (x_m, y_m)\}$ , where  $x_i$  is the  $i$ -th feature vector,  $y_i$  is the class tag of  $x_i$ .  
 $x_i \in \mathcal{X} = \mathbb{R}^n, y_i \in \mathcal{Y} = \{+1, -1\}, i = 1, 2, \dots, N$

The corresponding equation of the classification hyperplane is as follows:

$$h(x) = \omega \cdot x + b \quad (2)$$

where  $x$  is the input vector,  $\omega$  is the weight,  $b$  is the offset.

The classification decision function is as follows:

$$\text{Sign}(h(x)) \quad (3)$$

$$\begin{cases} h(x) > 0, y_i = 1 \\ h(x) < 0, y_i = -1 \end{cases} \quad (4)$$

The support vector machine is implemented to find the  $\omega$  and  $b$  when the interval between the separation hyperplane and the nearest sample point is maximized. When the training set is linearly separable, the sample points belonging to different classes can be separated by one or several straight lines with the largest interval. The maximum interval is solved by the following formula:

$$\max \gamma_i = y_i \left( \frac{\omega}{\|\omega\|} \cdot x_i + \frac{b}{\|\omega\|} \right) \quad (5)$$

$$s.t. y_i \left( \frac{\omega}{\|\omega\|} \cdot x_i + \frac{b}{\|\omega\|} \right) \geq \gamma, i = 1, 2, \dots, N \quad (6)$$

Thus, we can obtain the linear separable support vector machine optimization problem.

$$\min_{\omega, b} \frac{1}{2} \|\omega\|^2 \quad (7)$$

$$s.t. y_i(\omega \cdot x_i + b) - 1 \geq 0, i = 1, 2, \dots, N \quad (8)$$

In the actual data set, there are many specific points, making the data set linear inseparable; in order to solve this problem, we introduce a slack variable for each sample point.  $\xi_i \geq 0$  so that

$$y_i(\omega \cdot x_i + b) \geq 1 - \xi \quad (9)$$

For each slack variable  $\xi_i$ , pay a price  $\xi_i$ , and the optimization problem becomes:

$$\frac{1}{2} \|\omega\|^2 + C \sum_{i=1}^N \xi_i \quad (10)$$

where  $C > 0$  is the penalty factor.

Most of the data in the actual data are linearly inseparable. Therefore, these data should be mapped to a high-dimensional feature space through non-linear mapping, and the non-linear problem is transformed into a linear problem. The linear indivisible problem is transformed into a linear separable problem.

The kernel functions are introduced as follows:

$$K(x_i, x_j) = \varphi(x_i) \cdot \varphi(x_j) \quad (11)$$

where the value of the kernel equals the inner product of two vectors,  $x_i$  and  $x_j$ .

At this point, we obtain:

$$W(\alpha) = \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N \alpha_i \alpha_j y_i y_j K(x_i, x_j) - \sum_{i=1}^N \alpha_i \quad (12)$$

where  $\alpha_i > 0, i = 1, 2, \dots, N$  is the lagrangian multiplier and  $N$  is the number of samples.

In this paper, the radial basis function (RBF) is chosen as the support vector machine kernel function, and the expression is as follows:

$$K(x_i, x) = \exp\left(\frac{-\|x_i - x\|^2}{2g^2}\right) \quad (13)$$

where  $g$  is the kernel function coefficient.

At this point, the classification function becomes:

$$f(x) = \text{sign}\left[\sum_{i=1}^N \alpha_i y_i \exp\left(\frac{-\|x_i - x\|^2}{2g^2}\right) + b\right] \quad (14)$$

### 3. Hybrid Algorithm Research

The accuracy of prediction for the mold level is influenced by many factors. In order to improve the accuracy of prediction for the mold level, a predicting model, based on a Hybrid predict algorithm for the mold level, is proposed.

First, the original signal is subjected to data pre-processing to remove singular points. Then all data is marked in the range of 0 to 1 to improve computational efficiency. Finally, the hybrid model is used for data predict.

#### 3.1. Feasibility Analyses

EMD and SVR algorithm can be hybridized into an efficient hybrid algorithm. The reasons can be summarized as follows:

EMD is a decomposition algorithm that can decompose complex signals into simple signals. SVR is a regression algorithm. The fusion of EMD and SVR can help to predict complex signals, improve the ability of complex signal prediction, and it can obtain high-quality solutions. For this reason, the predict accuracy of hybrid algorithm is higher.

GA is a global optimization algorithm. The best model of SVR has a strong dependence on the kernel function parameters. GA can calculate the optimal SVR kernel function parameters faster and lay the foundation for the establishment of the best SVR model. Consequently, the efficiency of hybrid algorithm is higher.

EMD algorithm is an adaptive decomposition algorithm. There is no influence of human factors. The SVR model is more robust to nonlinear signals after optimization by the GA algorithm. Therefore, hybrid algorithm is more robust.

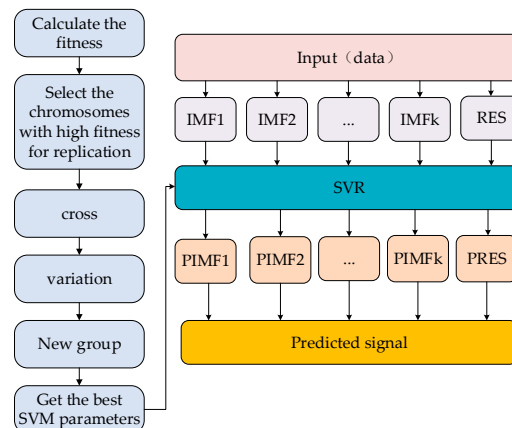
#### 3.2. The Full Procedure of EMD-SVR<sub>GA</sub>

EMD is an adaptive decomposition algorithm, based on the original signal, which can decompose complex signals into simple signals. Obviously, the prediction of simple signals is simpler than the prediction of complex signals, and the calculation cost is small. SVR is an excellent predictive algorithm, especially robust to nonlinear signals. GA algorithm is used to optimize parameters of the SVR kernel function, which further improves the SVR prediction accuracy. In order to make full use of the above algorithms, we present a hybrid algorithm, in which the GA is incorporated into the SVR. In EMD-SVR<sub>GA</sub>, SVR is used to predict simple signals, which is decomposed by EMD, and the computational cost is greatly reduced. The GA algorithm optimizes the parameters of the SVR kernel function, improves the SVR calculation efficiency, and makes the hybrid algorithm achieve faster convergence speed. In addition, the adaptive characteristics of EMD and the robustness of SVR to nonlinear, non-stationary, non-Gaussian distributed signals further improves the generalization ability of the hybrid algorithm. The EMD-SVR<sub>GA</sub> algorithm can be described by the following steps as shown in Table 1, and the corresponding flowchart is shown in Figure 3.

**Table 1.** The EMD-SVR<sub>GA</sub> algorithm.

Step 1. Decompose the mold level signal data into several IMFs and one residual by the EMD.
Step 2. Global optimization of C and g in SVR is performed using GA to determine the SVR model.
Step 3. Predicted IMFs (PIMF) obtained by SVR for each IMF.
Step 4. Predicted residual sequence obtained by SVR for residual sequence.
Step 5. Sum all predicted IMFs and residual sequences to obtain the predicted signal.

The flow chart of EMD-SVR is shown in Figure 3.



**Figure 3.** The flow chart of EMD-SVR. IMF: intrinsic mode functions; SVR: support vector regression; PIMF: predicted intrinsic mode functions.

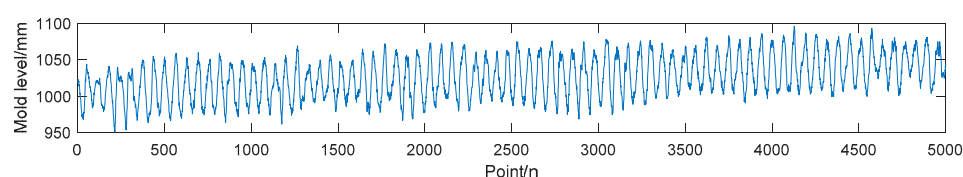
From the above, we can see that GA gives the optimal parameters of the SVR kernel function, improves the convergence speed of SVR, EMD decomposes the original signal of mold level, reduces the complexity of mold level signal, and further predicted by SVR. The roles of the three algorithms are different. In a word, EMD performs signal decomposition, and the GA performs SVR kernel function parameter optimization to make the SVR obtain the prediction result precisely and faster.

## 4. Experiments Studies

### 4.1. Problem Prescription

In order to clearly express the applicability, superiority, and generalization capability of the model applications, the mold level data of actual process parameters are used in this paper. These were collected from the continuous casting machines developed by the China National Heavy Machinery Research Institute Co., Ltd., Xi'an, China. We used an eddy current sensor to collect the mold-level signal at a steady cast speed. The cast speed is 0.9 m/min, and the tundish temperature is 1562 °C. Most of the disturbances are non-linear and non-stationary, and the long-term predict model is difficult to establish. This paper presents mold level predict model is important for mold level control to propose new ideas to improve the continuous casting automatic control.

A continuous casting production process data acquisition graph is presented in Figure 4. The time interval  $\Delta t = 1$  h, and the sampling frequency is 3 Hz.



**Figure 4.** Mold level. The unit of mold level is mm, n is the number of point.

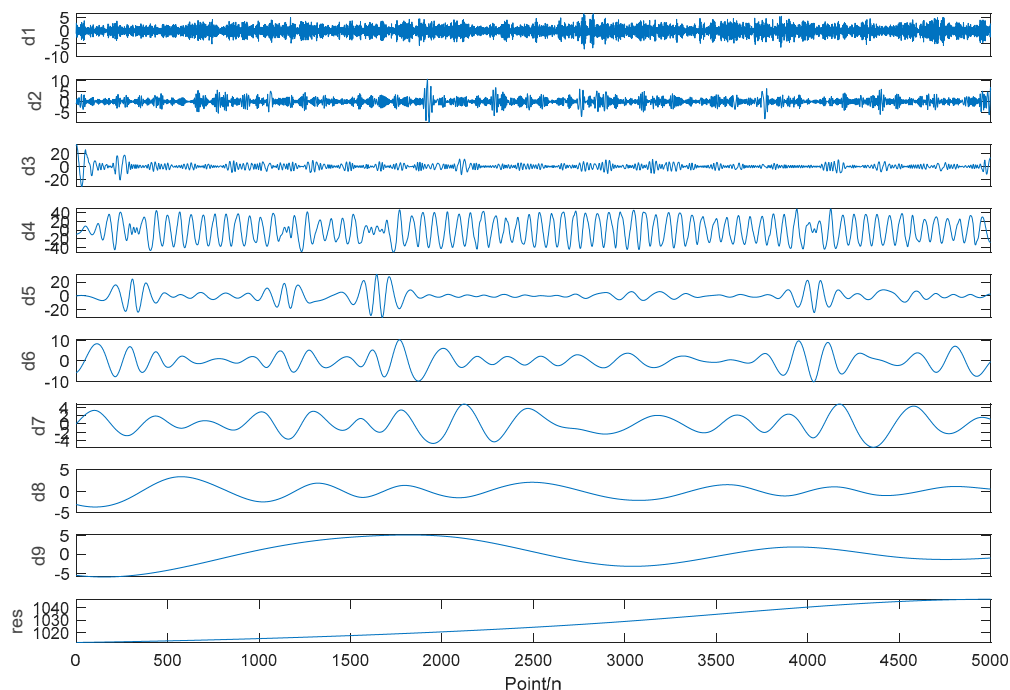
The main technical parameters of the continuous casting machine are shown in Table 2.

**Table 2.** Main technical parameters of the continuous casting machine.

Project	Specification
Continuous casting machine model	Curved continuous caster
Secondary cooling category	Aerosol cooling, dynamic water distribution
Gap control	Remote adjustment, dynamic soft reduction
Basic arc radius/mm	9500
Mold length/mm	900
Metallurgical length/mm	39,200
Mold vibration frequency/time/min	25–400
Mold vibration amplitude/mm	2–10
Slab width/mm	900–2150
Slab thickness/mm	230/250
Working speed/m/min	0.8–2.03

#### 4.2. EMD-SVR

During the continuous casting production process, the data was intercepted for one hour, the singularity points were removed according to the Layda criteria, the data from the first 40 min was used as the training set, and the last 20 min of data was used as the test set to verify the validity of the model. Descriptive statistics of mold level data are also given in Table 3. Mold-level data EMD results is shown in Figure 5, there is a trend term in Figure 5, which clearly shows that the mold level data is non-stationary. As shown in Figure 6,  $C$  is 95.5729 and  $g$  is 0.39511, through the global optimization of the GA. Then the SVR model was determined, each IMF and residual sequence is predicted by the SVR model. The final predict signal, as shown in Figures 7 and 8.



**Figure 5.** Mold-level data EMD results.  $d_i$  is the  $i$ -th IMF, the unit of  $d_i$  is mm,  $n$  is the number of Point, res is residual.



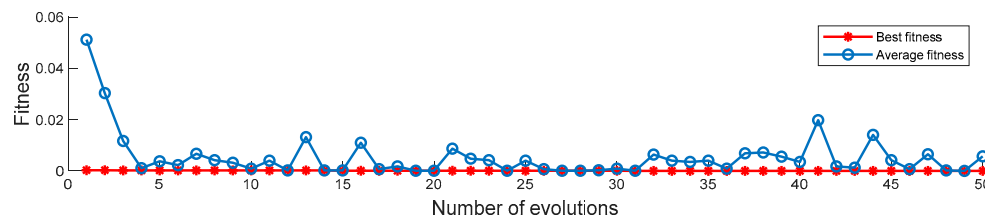


Figure 6. C and g optimization results.

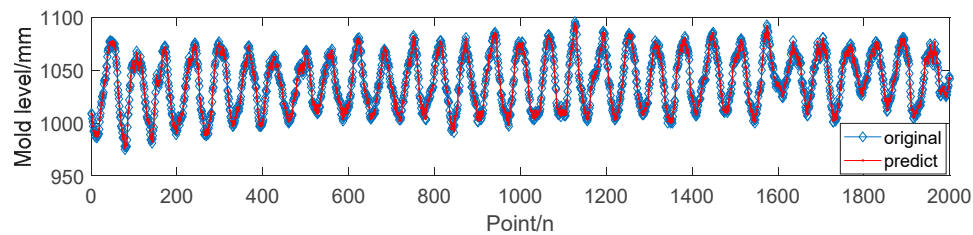


Figure 7. Comparison of EMD-SVR prediction results with original mold level data.

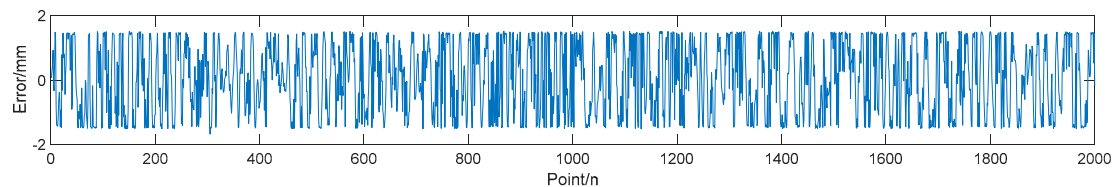


Figure 8. The EMD-SVR predict error.

Table 3. Descriptive statistics of mold level.

Project	Mold Level
Max-Min Values (mm)	951–1096
Mean (mm)	1026.7607
Standard Deviation (mm)	28.3495
Skewness	−0.02506
Kurtosis	2.1049

#### 4.3. Experimental Results and Analyses

We made a comparison between our proposed EMD-SVRGA method and the five kinds of algorithms, i.e., the algorithms include WT-ARMA, WT-SVR, WT-SVR<sub>GA</sub>, and EMD-ARMA, EMD-SVR<sub>GA</sub>. The computational results of the above methods are listed in Table 3 in detail. The parameter values of the model are displayed in Table 4.

Table 4. Parameter values of the model. ACF is Autocorrelation coefficient, PACF is Partial autocorrelation coefficient.

	C	g	ACF	PACF
WT-ARMA	-	-	2	2
WT-SVR	100	1	-	-
WT-SVR <sub>GA</sub>	16.3485	0.01773	-	-
EMD-ARMA	-	-	2	2
EMD-SVR	100	1	-	-
EMD-SVR <sub>GA</sub>	95.5729	0.39511	-	-

The performances of the three hybrid methods models are verified by four statistical indicators in this paper, and the best hybrid predict model that is suitable for continuous casting process parameters is selected.

Correlations between the original data and the predict data, which is characterized by correlation coefficients (CC)

$$R = \frac{Cov(P, A)}{\sqrt{Var(P) \cdot Var(A)}} \quad (15)$$

CC is defined as a statistical indicator used to reflect the close relationship between variables, the larger the CC, the better the algorithm performance.

Root-mean-square Error (RMSE)

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (P_i - A_i)^2}{n}} \quad (16)$$

RMSE is defined as reflect the degree of dispersion of a data set, measure the deviation between the observed value and the true value, the smaller the RMSE, the better the algorithm performance.

Mean Absolute Error (MAE)

$$MAE = \frac{\sum_{i=1}^n |P_i - A_i|}{n} \quad (17)$$

MAE is defined as average value of absolute error, reflect the actual situation of predict error better. The smaller the MAE, the better the algorithm performance.

Mean Absolute Percentage Error (MAPE)

$$MAPE = \frac{\sum_{i=1}^n \left| \frac{P_i - A_i}{A_i} \right|}{n} \times 100 \quad (18)$$

MAPE can be used to measure the outcome of a model predict. The smaller the MAPE, the better the algorithm performance.

Where  $P_i$  and  $A_i$  are the  $i$ -th predicted and actual values, respectively, and  $n$  is the total number of predict.

Predict model test results are shown in Table 5 and Figure 9.

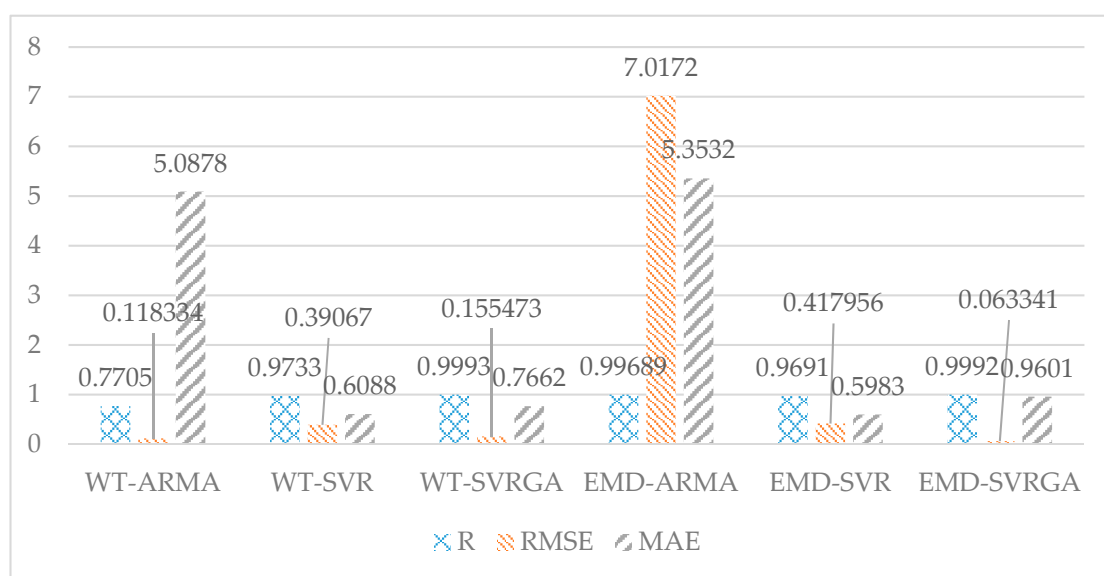


Figure 9. Comparison between EMD-SVR<sub>GA</sub> and other methods.

**Table 5.** Predict model test results.

	<b>R</b>	<b>RMSE</b>	<b>MAE</b>
WT-ARMA	0.7705	0.118334	5.0878
WT-SVR	0.9733	0.390670	0.6088
WT-SVR <sub>GA</sub>	0.9993	0.155473	0.7662
EMD-ARMA	0.99689	7.0172	5.3532
EMD-SVR	0.9691	0.417956	0.5983
EMD-SVR <sub>GA</sub>	0.9992	0.063341	0.9601

As shown in Table 5 and Figure 9, among the EMD based hybrid methods, the EMD-SVR had a better performance than EMD-ARMA, and EMD-SVR<sub>GA</sub> had better performance than EMD-SVR. Only the MAE is 0.9601, it is larger than the other methods. The other three indicators are all the best predictors of performance. Among the WT based hybrid methods, the WT-SVR had a better performance than WT-ARMA, and WT-SVR<sub>GA</sub> had better performance than WT-SVR, the performance of WT-SVR<sub>GA</sub> is the best, however, it is not stable. Only two indicators are better than the other two methods. RMSE is worse than WT-ARMA, and MAE is worse than WT-SVR.

The SVR based hybrid methods had better performance than ARMA based hybrid methods, due to the advantage of SVR as a non-linear data regression algorithm. The SVR based hybrid algorithm greatly improves the prediction accuracy compared to the traditional ARMA algorithm, and shows a strong generalization ability and robustness.

Among the SVR-based hybrid methods, R of the WT-SVR<sub>GA</sub> is close to the R of EMD-SVR<sub>GA</sub>. Only MAE of WT-SVR<sub>GA</sub> is 0.7662, it is better than EMD-SVR<sub>GA</sub>. Although the performance of WT-SVR<sub>GA</sub> predict is equivalent to that of EMD-SVR<sub>GA</sub>, the wavelet transform is a kind of transcendental non-adaptive transformation. The transform effect depends on the selection of the basis function and the order of transformation. There are many human factors. For instance, the prediction effect is not stable, adaptability cannot be guaranteed to all data, and the generalization ability is far less than EMD-SVR<sub>GA</sub>.

It is observed from the results that the R of SVR<sub>GA</sub>-based hybrid method is improved by 0.2287 compared with the ARMA based hybrid method, RMSE is reduced by 6.9538. R of EMD hybrid method is reduced by 0.0001, RMSE is reduced by 0.054993. The EMD-SVR<sub>GA</sub> hybrid method greatly improves the prediction accuracy compared to the traditional ARMA algorithm, and shows a strong generalization ability and robust.

#### 4.4. Predicting Accuracy Significance Tests

In order to verify the significant advantages of the proposed model in terms of prediction accuracy, some statistical tests are implemented in this section. Based on Dong et al. [34] and Fan et al. [35] suggested the Wilcoxon signed-rank test [36] and Friedman test [37], which are simultaneously applied in this paper.

The statistic  $W$  of the Wilcoxon signed-rank test:

$$W = \min\{r^+, r^-\} \quad (19)$$

The statistic  $F$  of the Friedman test:

$$F = \frac{12N}{k(k+1)} \left[ \sum_{j=1}^k R_j^2 - \frac{k(k+1)^2}{4} \right] \quad (20)$$

where  $N$  is the total number of predict results;  $k$  is the number of compared models;  $R_j$  is the average rank sum obtained in each predict value for each compared model as shown in Equation (21):

$$R_j = \frac{1}{N} \sum_{i=1}^N r_i^j \quad (21)$$

where  $r_i^j$  is the rank sum from 1 (the smallest predict error) to  $k$  (the worst predict error) for  $i$ -th predict result, for  $j$ -th compared model.

If the associated  $p$ -value of  $F$  meets the criterion of not acceptance, the null hypothesis, equal performance among all compared models, are also not held.

The results of Wilcoxon signed-rank test and Friedman test are listed in Table 6.

**Table 6.** Results of Wilcoxon signed-rank test and Friedman test.

Compared Models	Wilcoxon Signed-Rank Test				Friedman Test
	$\alpha = 0.025$		$\alpha = 0.05$		$\alpha = 0.05$
	<i>h</i> -Value	<i>p</i> -Value	<i>h</i> -Value	<i>p</i> -Value	
EMD-SVR <sub>GA</sub> vs. WT-ARMA	1	0	1	0	H <sub>0</sub> : The results of the six algorithms are equal $F = 35.8$ $p = 0$ (Reject H <sub>0</sub> )
EMD-SVR <sub>GA</sub> vs. WT-SVR	1	0	1	0	
EMD-SVR <sub>GA</sub> vs. WT-SVR <sub>GA</sub>	1	0	1	0	
EMD-SVR <sub>GA</sub> vs. EMD-ARMA	1	0	1	0	
EMD-SVR <sub>GA</sub> vs. EMD-SVR	1	0	1	0	

where  $\alpha$  is the level of significance,  $p$  returns whether the population producing the two independent samples is the same significant probability, and  $h$  returns the result of the hypothesis test. If the overall difference between  $x$  and  $y$  is not significant, then  $h$  is zero. If the overall difference between  $x$  and  $y$  is significant, then  $h$  is 1. If  $p$  is close to zero, then the null hypothesis can be questioned.  $H_0$  is the assumption of Friedman Test.

## 5. Conclusions

In this paper, a novel hybrid method, based on EMD-SVR<sub>GA</sub>, is proposed. The proposed method is applied to predict the mold level. In this method, the original mold level signal decomposed to several IMFs by the EMD algorithm. SVR is optimized by GA, predicted IMFs and predicted residual sequences are obtained by optimized SVR. The predicted result is reconstructed by EMD. The use of EMD-SVR<sub>GA</sub> model can achieve accurate predictions of non-linear, non-stationary data. This model uses GA to improve the global search capability of the parameters in the SVR models and avoid falling into local optimization, thereby optimizing SVR algorithm for better accuracy. In this paper, six predict algorithms were calibrated by four statistical indicators. The experimental results demonstrate the reliability and validity of the proposed EMD-SVR<sub>GA</sub> model in predict Mold level. The model has a strong generalization, capability and robustness.

The precise prediction of mold level provides a new idea for the continuous casting process improvement. Short-term mold level prediction can effectively avoid confusion of production rhythm caused by crystallizer level fluctuations, and long-term predictions can effectively avoid accidents, such as slab breakout and steel overflow at the mold top. Accurate predictions of mold levels has important practical significance.

**Author Contributions:** W.S. conceived and designed the research; Z.L. performed the experiment and wrote the manuscript.

**Funding:** This work was financially supported by the National Natural Science Foundation of China (NO. 51575429).

**Acknowledgments:** Q.G., X.L., H.Z., B.H., and Y.Z. are acknowledged for their valuable technical support.

**Conflicts of Interest:** The authors declare no conflicts of interest.

## Nomenclature

EMD	Empirical mode decomposition
IMF	Intrinsic mode function
TS	Time series
ARMA	Autoregressive moving average
ACF	Autocorrelation function
PACF	Partial autocorrelation function
SVR	Support vector regression
GA	Genetic Algorithm
WT	Wavelet Transform
CC	Correlation Coefficient
RMSE	Root-mean-square error
MAE	Mean Absolute Error

## References

- Ataka, M. Rolling Technology and Theory for the Last 100 Years: The Contribution of Theory to Innovation in Strip Rolling Technology. *ISIJ Int.* **2015**, *55*, 89–102. [\[CrossRef\]](#)
- Jin, X.; Chen, D.F.; Zhang, D.J.; Xie, X.; Bi, Y.Y. Water model study on fluid flow in slab continuous casting mould with solidified shell. *Ironmak. Steelmak.* **2011**, *38*, 155–159. [\[CrossRef\]](#)
- Shen, B.Z.; Shen, H.F.; Liu, B.C. Water modelling of level fluctuation in thin slab continuous casting mould. *Ironmak. Steelmak.* **2009**, *36*, 33–38. [\[CrossRef\]](#)
- Wang, X.; Zhang, S.; Yao, M.; Ma, H.; Zhang, X. Effect of casting process on mould friction during wide, thick slab continuous casting. *Ironmak. Steelmak.* **2014**, *41*, 464–473. [\[CrossRef\]](#)
- Lynch, P. The origins of computer weather prediction and climate modeling. *J. Comput. Phys.* **2008**, *227*, 3431–3444. [\[CrossRef\]](#)
- Weron, R. Electricity price predicting: A review of the state-of-the-art with a look into the future. *Int. J. Predict.* **2014**, *30*, 1030–1081. [\[CrossRef\]](#)
- Lee, W.-J.; Hong, J. A hybrid dynamic and fuzzy time series model for mid-term power load predicting. *Int. J. Electr. Power Energy Syst.* **2015**, *64*, 1057–1062. [\[CrossRef\]](#)
- Dai, S.; Niu, D.; Li, Y. Daily Peak Load Predicting Based on Complete Ensemble Empirical Mode Decomposition with Adaptive Noise and Support Vector Machine Optimized by Modified Grey Wolf Optimization Algorithm. *Energies* **2018**, *11*, 163. [\[CrossRef\]](#)
- Wang, W.-C.; Chau, K.-W.; Cheng, C.-T.; Qiu, L. A comparison of performance of several artificial intelligence methods for predicting monthly discharge time series. *J. Hydrol.* **2009**, *374*, 294–306. [\[CrossRef\]](#)
- Azad, H.B.; Mekhilef, S.; Ganapathy, V.G. Long-Term Wind Speed Predicting and General Pattern Recognition Using Neural Networks. *IEEE Trans. Sustain. Energy* **2014**, *5*, 546–553. [\[CrossRef\]](#)
- Gaudio, M.; Gorgone, E.; Labbe, M.; Rodriguez-Chia, A.M. Lagrangian relaxation for SVM feature selection. *Comput. Oper. Res.* **2017**, *87*, 137–145. [\[CrossRef\]](#)
- Wang, J.; Shi, P.; Jiang, P.; Hu, J.; Qu, S.; Chen, X.; Chen, Y.; Dai, Y.; Xiao, Z. Application of BP Neural Network Algorithm in Traditional Hydrological Model for Flood Predicting. *Water* **2017**, *9*, 48. [\[CrossRef\]](#)
- He, F.; Zhang, L. Mold breakout prediction in slab continuous casting based on combined method of GA-BP neural network and logic rules. *Int. J. Adv. Manuf. Technol.* **2018**, *95*, 4081–4089. [\[CrossRef\]](#)
- Fan, G.-F.; Peng, L.-L.; Hong, W.-C.; Sun, F. Electric load predicting by the SVR model with differential empirical mode decomposition and auto regression. *Neurocomputing* **2016**, *173*, 958–970. [\[CrossRef\]](#)
- Nie, H.; Liu, G.; Liu, X.; Wang, Y. Hybrid of ARIMA and SVMs for Short-Term Load Predicting. *Energy Procedia* **2012**, *16*, 1455–1460. [\[CrossRef\]](#)
- Liu, Y.; Gao, Z. Enhanced just-in-time modelling for online quality prediction in BF ironmaking. *Ironmak. Steelmak.* **2015**, *42*, 321–330. [\[CrossRef\]](#)
- Cincotti, S.; Gallo, G.; Ponta, L.; Raberto, M. Modeling and predicting of electricity spot-prices: Computational intelligence vs classical econometrics. *Ai Commun.* **2014**, *27*, 301–314. [\[CrossRef\]](#)

18. Ghosh, S.K.; Ganguly, S.; Chattopadhyay, P.P.; Datta, S. Effect of copper and microalloying (Ti, B) addition on tensile properties of HSLA steels predicted by ANN technique. *Ironmak. Steelmak.* **2009**, *36*, 125–132. [\[CrossRef\]](#)
19. Hong, W.-C. Chaotic particle swarm optimization algorithm in a support vector regression electric load predicting model. *Energy Convers. Manag.* **2009**, *50*, 105–117. [\[CrossRef\]](#)
20. Voyant, C.; Muselli, M.; Paoli, C.; Nivet, M.-L. Numerical weather prediction (NWP) and hybrid ARMA/ANN model to predict global radiation. *Energy* **2012**, *39*, 341–355. [\[CrossRef\]](#)
21. Ren, Y.; Suganthan, P.N.; Srikanth, N. A Comparative Study of Empirical Mode Decomposition-Based Short-Term Wind Speed Predicting Methods. *IEEE Trans. Sustain. Energy* **2015**, *6*, 236–244. [\[CrossRef\]](#)
22. Ye, L.; Liu, P. Combined Model Based on EMD-SVM for Short-term Wind Power Prediction. *Proc. CSEE* **2011**, *31*, 102–108.
23. Lei, Y.G.; Lin, J.; He, Z.J.; Zuo, M.J. A review on empirical mode decomposition in fault diagnosis of rotating machinery. *Mech. Syst. Signal Process.* **2013**, *35*, 108–126. [\[CrossRef\]](#)
24. Huang, N.E.; Shen, Z.; Long, S.R.; Wu, M.L.C.; Shih, H.H.; Zheng, Q.N.; Yen, N.C.; Tung, C.C.; Liu, H.H. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. *Proc. R. Soc. A Math. Phys. Eng. Sci.* **1998**, *454*, 903–995. [\[CrossRef\]](#)
25. Bajaj, V.; Pachori, R.B. Classification of Seizure and Nonseizure EEG Signals Using Empirical Mode Decomposition. *IEEE Trans. Inf. Technol. Biomed.* **2012**, *16*, 1135–1142. [\[CrossRef\]](#) [\[PubMed\]](#)
26. Priya, A.; Yadav, P.; Jain, S.; Bajaj, V. Efficient method for classification of alcoholic and normal EEG signals using EMD. *J. Eng.* **2018**, *2018*, 166–172. [\[CrossRef\]](#)
27. Prasanna Kumar, M.K.; Kumaraswamy, R. Single-channel speech separation using combined EMD and speech-specific information. *Int. J. Speech Technol.* **2017**, *20*, 1037–1047. [\[CrossRef\]](#)
28. Guo, Z.; Zhao, W.; Lu, H.; Wang, J. Multi-step predicting for wind speed using a modified EMD-based artificial neural network model. *Renew. Energy* **2012**, *37*, 241–249. [\[CrossRef\]](#)
29. Tang, J.; Zhao, L.; Yue, H.; Yu, W.; Chai, T. Vibration Analysis Based on Empirical Mode Decomposition and Partial Least Square. *Procedia Energy* **2011**, *16*, 646–652. [\[CrossRef\]](#)
30. Xu, G.; Tian, W.; Qian, L. EMD- and SVM-based temperature drift modeling and compensation for a dynamically tuned gyroscope (DTG). *Mech. Syst. Signal Process.* **2007**, *21*, 3182–3188. [\[CrossRef\]](#)
31. Luo, L.; Yan, Y.; Xie, P.; Sun, J.; Xu, Y.; Yuan, J. Hilbert-Huang transform, Hurst and chaotic analysis based flow regime identification methods for an airlift reactor. *Chem. Eng. J.* **2012**, *181*, 570–580. [\[CrossRef\]](#)
32. Srinivasan, R.; Rengaswamy, R.; Miller, R. A modified empirical mode decomposition (EMD) process for oscillation characterization in control loops. *Control Eng. Pract.* **2007**, *15*, 1135–1148. [\[CrossRef\]](#)
33. Zhang, X.Y.; Zhou, J.Z. Multi-fault diagnosis for rolling element bearings based on ensemble empirical mode decomposition and optimized support vector machines. *Mech. Syst. Signal Process.* **2013**, *41*, 127–140. [\[CrossRef\]](#)
34. Dong, Y.; Zhang, Z.; Hong, W.-C. A Hybrid Seasonal Mechanism with a Chaotic Cuckoo Search Algorithm with a Support Vector Regression Model for Electric Load Predicting. *Energies* **2018**, *11*, 1009. [\[CrossRef\]](#)
35. Fan, G.-F.; Peng, L.-L.; Hong, W.-C. Short term load predicting based on phase space reconstruction algorithm and bi-square kernel regression model. *Appl. Energy* **2018**, *224*, 13–33. [\[CrossRef\]](#)
36. Wilcoxon, F. Individual comparisons by ranking methods. *Biom. Bull.* **1945**, *1*, 80–83. [\[CrossRef\]](#)
37. Friedman, M. A comparison of alternative tests of significance for the problem of m rankings. *Ann. Math. Stat.* **1940**, *11*, 86–92. [\[CrossRef\]](#)

