

Article

# Grouting Power Prediction Using a Hybrid Model Based on Support Vector Regression Optimized by an Improved Jaya Algorithm

Linli Xue, Yushan Zhu, Tao Guan, Bingyu Ren \*, Dawei Tong and Binping Wu

State Key Laboratory of Hydraulic Engineering Simulation and Safety, Tianjin University, Tianjin 300350, China; xuelinli@tju.edu.cn (L.X.); zhuyushan3@tju.edu.cn (Y.Z.); tao.guan@tju.edu.cn (T.G.); tongdw@tju.edu.cn (D.T.); wubinping@tju.edu.cn (B.W.)

\* Correspondence: renby@tju.edu.cn; Tel.: +86-022-2789-0911

Received: 20 September 2020; Accepted: 15 October 2020; Published: 17 October 2020



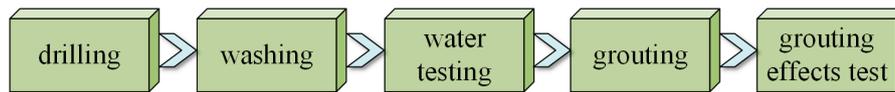
**Abstract:** Grouting power is a vital parameter that can be used as an indicator for simultaneously controlling grouting pressure and injection rate. Accurate grouting power prediction contributes to the real-time optimization of the grouting process, guaranteeing grouting safety and quality. However, the strong nonlinearity of the grouting power time series makes the forecasting task challenging. Hence, this paper proposes a novel hybrid model for accurate grouting power forecasting. First, empirical wavelet transform (EWT) is employed to decompose the original grouting series into several subseries and one residual adaptively. Second, partial autocorrelation function (PACF) is applied to identify the optimal input variables objectively. Then, support vector regression (SVR) is adopted to obtain prediction outcomes of each subseries, while an improved Jaya (IJaya) algorithm by coupling chaos theory and Lévy flights to improve the algorithm's accuracy performance is proposed to optimize the SVR hyperparameters. Finally, the prediction results of decomposed subseries are superimposed to produce the final results. A consolidation grouting project is taken as a case study and the computation results with the RMSE = 0.2672 MPa·L/min, MAE = 0.2165 MPa·L/min, MAPE = 3.85% and EC = 0.9815 demonstrate that the proposed model exhibits superior forecasting ability and can provide a viable reference for grouting construction.

**Keywords:** grouting power prediction; hybrid model; support vector regression; improved Jaya algorithm; hyperparameters optimization

## 1. Introduction

Grouting, as a method of foundation treatment, is a crucial step in guaranteeing a dam's safety and reliability [1]. Generally, the grouting procedures can be arranged in the following five sequences: drilling, washing, water testing, grouting and grouting effects test, which are depicted in Figure 1. Real-time optimization of grouting process contributes to guaranteeing grouting safety and quality. The grouting pressure and injection rate, which are two key grouting process parameters, are directly related to each other in practical projects, making the relationship between them of considerable importance [2], and conjoint adjustment of grouting pressure and injection rate reasonably contributes to controlling the grouting quality. Xia [3] proposed a new indicator, grouting power, which is the product of grouting pressure and injection rate, to link these two parameters. Because the grouting pressure and injection rate are not determined values but fluctuate within upper and lower limits in practical engineering, the grouting power has a certain range; in addition to being influenced by complex geographical conditions and construction conditions, it also exhibits highly nonlinear characteristics. If the grouting power is kept within a certain range, the grouting process is considered

to be normal and safe. Nevertheless, the adjustment of the grouting pressure and injection rate is conducted after the appearance of defective grouting conditions in practical engineering, which leads to poor control of the grouting process. Therefore, accurate and real-time prediction of grouting power is of great importance. First, accurate prediction can reflect whether defective grouting conditions will occur in advance. Second, based on the accurate prediction, construction workers can complete the reasonable regulation of the grouting pressure and injection rate to maintain the grouting power within a certain range and realize real-time optimization of the grouting process, guaranteeing grouting safety and quality.



**Figure 1.** The representation of grouting procedures.

With the rapid development of science and technology, intelligent algorithms have been introduced into the prediction research of grouting projects. Tran and Hoang [4] developed a groutability prediction model by utilizing support vector regression (SVR) and the differential evolution (DE) algorithm. Hoang et al. [5] combined SVR with differential flower pollination (DFP) for the prediction of groutability. Deng et al. [6] combined a support vector machine (SVM) with the hybrid grey wolf optimization algorithm (HGWO) to model groutability of fracture grouting. The back propagation neural network (BPNN) model and artificial neural network (ANN) model were established, respectively, to predict cement take by Yang [7] and Zhang et al. [8]. Fan et al. [9] applied SVR and multivariate regression analysis to predict cement take, and the superiority of SVR model was verified by a higher correlative coefficient. Öge [10] utilized an adaptive neuro fuzzy inference system (ANFIS) and multiple regression to predict cementitious grout take of mine shaft permeation. Li et al. [11] applied ANFIS to predict rock quality designation (RQD), permeability (Lugeon value), and fracture filled rate (FFR), then utilized these three predicted indicators to evaluate certain grouting efficiency. Most of the above research mainly focused on prediction of pre-grouting indicators such as groutability or post-grouting indicators such as cement take and grouting efficiency, and there is little attention paid on grouting power prediction. In addition, grouting power time series are often accompanied by highly nonlinear characteristics, making high-precision prediction a big challenge.

Time series prediction, which has been widely utilized in the field of streamflow [12], transportation [13] and dam deformation [14], among others, is the application of a model to predict future values based on historical data; it is also suitable for grouting power prediction. Currently, three main types of time series prediction models have been proposed: conventional statistical approaches, intelligent methods, and hybrid models [15]. Conventional statistical approaches, including autoregressive (AR) models [16] and autoregressive moving average (ARMA) models [17], are linear regression models, and a linear relationship is supposed between future values and historical values. However, considerable nonlinearity exists in the grouting power time series. Therefore, it is difficult to use conventional statistical approaches to capture the characteristics hidden in the grouting power time series. Artificial intelligence algorithms with good capabilities have emerged recently and are extensively used in time series prediction. ANNs [18] and SVMs [19] have flexible nonlinear modelling capability; however, they still have their own drawbacks. For example, NN models have problems of overfitting and are time consuming; meanwhile, sensitivity to parameter selection exists in both NN and SVM models. To overcome the inherent drawbacks of a single model, hybrid models based on the concept of “decomposition–prediction–integration” have been proposed recently and have proven to perform better than any single model alone in various fields, such as in landslide displacement [20] and hydrology [21].

Therefore, a novel hybrid model aiming at overcoming the limitations of single models and obtaining good prediction accuracy is proposed in this study for grouting power time series prediction. This prediction model integrates the merits of the empirical wavelet transform (EWT), the partial

autocorrelation function (PACF), SVR, and an improved (IJaya) algorithm and is supposed to achieve better accuracy.

The remainder of this paper is organized as follows: current studies related to the hybrid models based on “decomposition–prediction–integration” are presented in Section 2. The overall research framework is described in Section 3. Section 4 introduces the required individual methodologies, including the EWT, the PACF, SVR, and the IJaya algorithm. In Sections 5 and 6, a case study is performed, and the superiority of the proposed model is verified in comparison with four other competitors. Section 7 highlights the conclusions.

## 2. Literature Review

Hybrid models typically combine prediction algorithms with the data decomposition techniques and give full play to the advantages of each component to generate more accurate predictions. As an extension of SVMs, SVR was proposed by Drucker et al. [22] for regression analysis possessing superior capability in solving multidimensional, nonlinear problems with small samples [6], and has been proved to be effective in solving various forecasting problems [23,24]. Therefore, SVR-based models, which combine the SVR model with data decomposition techniques, are extensively used in time series prediction [25,26].

The EWT, which is a novel tool for data decomposition, has been successfully applied in many research areas [27–29] and was chosen as a data decomposition technique in this study to preprocess the original grouting power series, by which the original grouting power series is decomposed into a specific number of subseries and one residual series adaptively. When employing SVR models for the decomposed subseries prediction, the prediction performance relies greatly on the input variables and hyperparameters of forecasting components. Hence, identifying optimal input variables and hyperparameters is particularly necessary.

The PACF is used to determine the optimal input variables by calculating the correlation between data [30–32]. Meanwhile, identifying the optimal model hyperparameters is an optimization problem. Many studies use optimization algorithms to solve such optimization problems [33]. One of these methods employs the Jaya algorithm, which is a new population-based optimization method [34]. The main advantage of this algorithm is the fact that only two common parameters (e.g., the number of generations and the statistical population size) are required and no other specific parameters should be tuned. Unlike the Jaya algorithm, some other meta-heuristic algorithms require algorithm-specific parameters except for the two common parameters. For example, the DE algorithm needs the scaling factor and crossover probability, and the genetic algorithm requires mutation probability, crossover probability, and a selection operator. In addition, the Jaya algorithm only contains one update equation. Because of the above properties, this algorithm is easy to implement. Therefore, the Jaya algorithm has been extensively used in various practical optimization problems such as product design [35], parameter estimation [36], and image identification [37]. Although the Jaya algorithm has been proven to perform well, it still has the problem of premature convergence and becoming trapped in local optima frequently [38]. To overcome these problems, a new version of the Jaya algorithm is proposed herein. First, tent chaotic mapping has been employed in the Jaya algorithm to initialize the statistical population and preserve statistical population diversity. Moreover, we incorporate Lévy flights into the statistical population updating phase to enhance search capabilities and increase the probability of finding the optimal solution. This improved Jaya (IJaya) algorithm is then used to find the optimal hyperparameters of the SVR model.

Inspired by the concept of hybrid models, a novel hybrid model incorporating the EWT, PACF, SVR and IJaya algorithm is proposed for grouting power prediction based on “decomposition–prediction–integration”. The main aim of decomposition is to simplify the modelling complexity through decomposing the complex grouting power series into relatively simple subseries and one residual series, and denoise the original time series by discarding the residual series. Then the purpose of prediction is to decrease the prediction errors by using appropriate forecasting methods. Finally,

the purpose of integration is to yield the final result for the original series by aggregating prediction results of each subseries. The principal contributions of this paper can be summarized as follows:

- (1) The proposed model not only considers the characteristics of grouting power series but integrates the advantages of four algorithms to remedy the deficiencies of single models, and good prediction accuracy is obtained.
- (2) An IJaya algorithm is proposed to optimize the hyperparameters of SVR. In the IJaya algorithm, tent chaotic mapping has been employed in the Jaya algorithm to preserve statistical population diversity and prevent the algorithm from sticking on local optima. Also, Lévy flights is introduced to the statistical population updating phase of the basic Jaya algorithm to enhance search capabilities and increase the probability of finding the optimal solution.
- (3) The EWT is adopted as a data preprocessor to decompose the grouting power series into specific subseries and one residual series to extract meaningful information from the grouting power series. The residual is regarded as an uncorrelated white noise series and discarded to denoise the original grouting power series. The rest subseries are smoother and more predictable, contributing to obtain more precise prediction results.
- (4) The PACF is applied to calculate the partial correlation between the data in each subseries and identify the optimal input variables for predictors.

### 3. Research Framework

A hybrid model represented as EPIJaya-SVR is proposed for grouting power prediction in this study. It combines the merits of the EWT, the PACF, SVR, and the IJaya algorithm. The overall framework used is intuitively depicted in Figure 2. It mainly includes five steps:

- Step 1. Collect the original grouting power series based on the grouting real-time monitoring system [39]. Utilize the EWT to decompose the obtained time series into a specific number of subseries and one residual series adaptively to extract meaningful information from them and reduce modelling complexity. Denoise the original grouting power series by discarding the uncorrelated residual series.
- Step 2. Apply the PACF to analyze the correlation between the data in each subseries and determine the optimal input variables objectively.
- Step 3. Utilize the SVR model to forecast the decomposed subseries; employ the IJaya algorithm to optimize the hyperparameters of the built SVR models. The constructed IJaya-SVR model can thus achieve better prediction accuracy.
- Step 4. Sum the prediction results of the subseries to formulate an ensemble forecasting result for the original series.
- Step 5. Apply the hybrid model established by using the methodologies mentioned in steps 1–4 to an actual project. The grouting power variation range and trend can be predicted and analyzed. The applicability and advantages of the proposed hybrid model are discussed and verified in comparison with other prediction models.

### 4. Methodology

In this study, a novel hybrid model for grouting power forecasting with good prediction accuracy is presented. The main methodologies employed in the proposed hybrid model, consisting of the EWT, PACF, SVR, and IJaya algorithm, are introduced in the following.

#### 4.1. EWT

The EWT, which was proposed by Gilles [40], is a novel signal processing tool for identifying and extracting different intrinsic modes of time series. It can cope with the shortcomings of some conventional data decomposition techniques: for example, the wavelet transform (WT) [41]

lacks self-adaptive data processing capabilities and has difficulties in selecting a mother wavelet, while empirical mode decomposition (EMD) [42] and ensemble empirical mode decomposition (EEMD) [43] face the problem of insufficient mathematical definition.

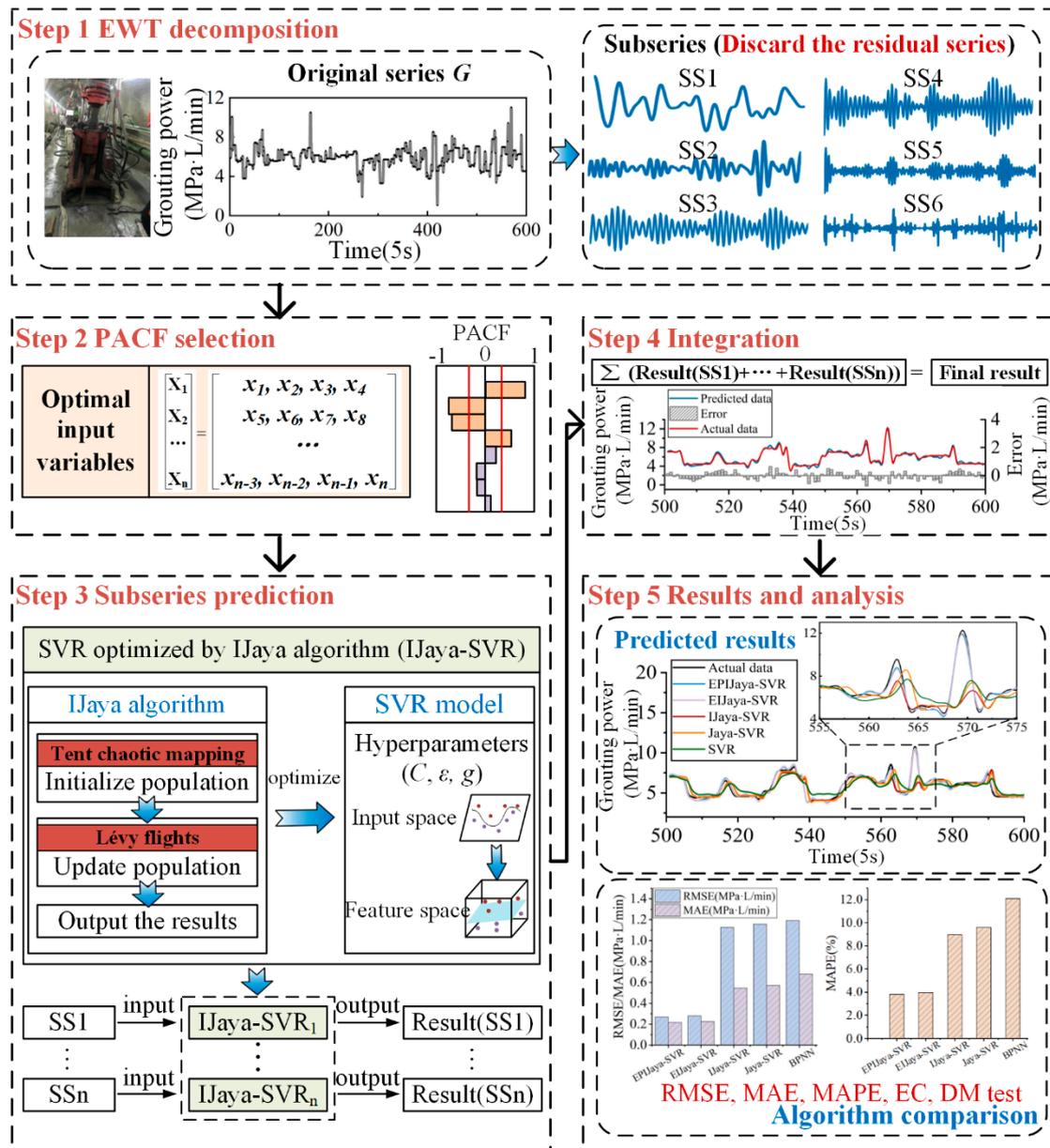


Figure 2. Research framework.

The EWT relies on powerful peak detection preprocessing to detect the local maximum value of the grouting power time series first, then performs spectrum segmentation, and, finally, generates a series of corresponding wavelet filters. By employing the EWT to decompose the grouting power into a limited number of subseries, good prediction accuracy can be obtained because the temporal variation characteristics in the subseries can be better captured by the prediction model.

First, divide the Fourier support  $(0, \pi)$  into  $N$  continuous segments, denoted as  $\Lambda_n = ()$ , where  $\omega_n$  are the limits of these segments ( $\omega_0 = 0, \omega_N = \pi$ ). Implement the EWT in the same way used as

in the classic WT. The detail coefficient  $w_f^\varepsilon(n, t)$  and approximation coefficient  $w_f^\varepsilon(0, t)$  can then be calculated as follows:

$$w_f^\varepsilon(n, t) = \langle f, \psi_n \rangle = \int f(\tau) \overline{\psi_n(\tau - t)} d\tau = \left( \hat{f}(\omega) \overline{\hat{\psi}_n(\omega)} \right)^\vee \tag{1}$$

$$w_f^\varepsilon(0, t) = \langle f, \phi_1 \rangle = \int f(\tau) \overline{\phi_1(\tau - t)} d\tau = \left( \hat{f}(\omega) \overline{\hat{\phi}_1(\omega)} \right)^\vee \tag{2}$$

where  $\hat{\Psi}_n(\omega)$  is the Fourier transform form of the empirical wavelet function  $\Psi_n(t)$ , and  $\hat{\Phi}_n(\omega)$  is the Fourier transform form of the empirical scale function  $\Phi_n(t)$ .

The original signal  $f(t)$  is decomposed into specific empirical modes as follows:

$$f_0(t) = w_f^\varepsilon(0, t) * \phi_1(t) \tag{3}$$

$$f_k(t) = w_f^\varepsilon(k, t) * \psi_k(t) \tag{4}$$

#### 4.2. PACF

The grouting power time series is decomposed by the EWT into several subseries, and the correlation between the data in each subseries is different. Therefore, in the design of the prediction model corresponding to different subseries, the correlation between the data in each subseries needs to be analyzed, and the optimal input variables of each subseries can be determined by analyzing the correlation. Suppose that  $x_t$  is the value needed to be predicted, and only  $x_{t-1}$  can be selected as the input variable if all the PACF values are within the 95% confidence interval  $(-1.96/\sqrt{N}, 1.96/\sqrt{N})$ . Meanwhile,  $x_{t-p}$  is taken as one of the input variables if the partial autocorrelation at lag  $p$  exceeds the 95% confidence interval. For details of the PACF process, see [44].

The PACF value at lag  $p$ , denoted by  $\alpha_{pp}$ , can be calculated as follows:

$$\begin{cases} \hat{\alpha}_{11} = \hat{\rho}_1 \\ \hat{\alpha}_{p+1,p+1} = \frac{\hat{\rho}_{p+1} - \sum_{j=1}^p \hat{\rho}_{p+1-j} \hat{\alpha}_{pj}}{1 - \sum_{j=1}^p \hat{\rho}_j \hat{\alpha}_{pj}} \\ \hat{\alpha}_{p+1,j} = \hat{\alpha}_{pj} - \hat{\alpha}_{p+1,p+1} \cdot \hat{\alpha}_{p,p-j+1} \end{cases} \tag{5}$$

where  $j = 1, 2, \dots, p$ ;  $p = 1, 2, \dots, M$ , where  $M$  is the maximum lag; and  $\rho_p$  is the autocorrelation function (ACF) at lag  $p$ .

#### 4.3. SVR Optimized by the IJaya Algorithm

SVR is an effective tool for nonlinear regression and is also suitable for grouting power forecasting. The detailed theory of SVR is presented elsewhere [45,46]. The radial basis function kernel, which is the most applicable kernel function [47], was adopted in this study. Through it, the original data space is mapped into an infinite-dimensional space theoretically. There are three user-determined parameters: the penalty parameter  $C$ , the insensitive loss function  $\varepsilon$ , and the kernel function parameter  $g$ ; these are critical to the performance of SVR and can be determined by optimization algorithm.

The new Jaya algorithm proposed to solve unconstrained and constrained optimization problems is a simple yet powerful optimization algorithm. Let  $O(X)$  be an objective function that needs to be optimized. Assume that the number of design variables, the statistical population size, and the maximum number of iterations are  $N_{var}$ ,  $N_{pop}$ , and  $N_{max}$ , respectively, and that  $i, j$ , and  $k$  are their

indices, where  $i \in (1, N_{var})$ ,  $j \in (1, N_{pop})$ , and  $k \in (1, N_{max})$ . If  $X_{i,j,k}$  stands for the value of the  $i_{th}$  design variable for the  $j_{th}$  candidate during the  $k_{th}$  iteration, then this value can be updated using:

$$X_{i,j,k}^{new} = X_{i,j,k} + r_1 \times (X_{i,best,k} - |X_{i,j,k}|) - r_2 \times (X_{i,worst,k} - |X_{i,j,k}|) \tag{6}$$

where  $X_{i,j,k}^{new}$  presents the upgraded value of  $X_{i,j,k}$ ;  $X_{i,best,k}$  and  $X_{i,worst,k}$  denote the best solution and the worst solution of the  $i_{th}$  variable during the  $k_{th}$  iteration, respectively;  $r_1$  and  $r_2$  stand for two random numbers with a range of (0,1). Item  $r_1 \times (X_{i,best,k} - |X_{i,j,k}|)$  indicates the tendency of the solution toward the best solution and item  $-r_2 \times (X_{i,worst,k} - |X_{i,j,k}|)$  indicates the tendency to keep away from the worst solution;  $X_{i,j,k}^{new}$  is accepted if an improved objective functional value corresponding it is achieved.

The following modifications are integrated into the Jaya algorithm to produce an IJaya algorithm and solve the problems of premature convergence and being trapped in local optima:

The proposed IJaya algorithm applies tent chaotic mapping to initialize the statistical population. The initial statistical population is usually generated randomly, which makes it difficult to ensure its diversity. Chaotic maps exhibit ergodicity, certainty, and the stochastic property, which can effectively preserve statistical population diversity and prevent the algorithm from being trapped in a local optimum [48]. In this study, tent chaotic mapping is selected to generate the initial statistical population, the steps of which are as follows:

- Generate the chaotic variable of  $X_k^i$ .

$$X_k^i = \frac{(X_k^i - X_{min}^i)}{(X_{max}^i - X_{min}^i)} \tag{7}$$

where  $k$  is the serial number of the  $k$ -th statistical population and  $X_{min}^i$  and  $X_{max}^i$  stand for two extreme values of the  $i$ -th variable, which are the minimum and the maximum, respectively.

- Apply tent chaotic mapping to generate the chaotic sequence:

$$X_{k+1}^i = \begin{cases} 2X_k^i & X_k^i < 0.5 \\ 2(1 - X_k^i) & X_k^i \geq 0.5 \end{cases} \tag{8}$$

- Map the chaotic sequence to the search space:

$$X_{k+1}^i = X_{min}^i + X_{k+1}^i (X_{max}^i - X_{min}^i) \tag{9}$$

We incorporate Lévy flights into the statistical population updating phase of the basic Jaya algorithm to enhance search capabilities. By employing Lévy flights to update the statistical population, variables can effectively approach their best value through taking short-distance jumps together with occasional long-distance jumps [49]. By substituting the  $X_{i,j,k}$  updated by Lévy flights, which is represented as  $Lévy(X_{i,j,k})$ , into formula (6), the algorithm's search ability is enhanced and the probability of finding the optimal solution is improved. For details of Lévy flights, see [50].

$$Lévy(X'_{i,j,k}) = X_{i,j,k} + \mu sign[rand - 0.5] \oplus Lévy(X_{i,j,k}) \tag{10}$$

$$X_{i,j,k}^{new} = Lévy(X'_{i,j,k}) + r_1 (X_{i,best,k} - |X_{i,j,k}|) - r_2 (X_{i,worst,k} - |X_{i,j,k}|) \tag{11}$$

where  $\mu$  represents the random number distributed uniformly,  $rand$  is a random number with a range of (0,1),  $\oplus$  stands for entry-wise multiplications.

SVR and the IJaya algorithm are combined to construct the prediction model for each subseries. The algorithm consists of the following steps, which are illustrated in Figure 3 [6]:

- Step 1. Parameter initialization Set the number of design variables,  $N_{var}$ , the statistical population size  $N_{pop}$ , the maximum number of iterations,  $N_{max}$ , and the lower and upper bounds of the design variables ( $lb$ ,  $ub$ ).
- Step 2. Hyperparameter ( $C$ ,  $\varepsilon$ ,  $g$ ) optimization Initialize the IJaya statistical population, then update the statistical population when the objective function value is better. When iteration  $k > N_{max}$ , exit the calculation and obtain the optimal ( $C$ ,  $\varepsilon$ ,  $g$ ).
- Step 3. Training and predicting Based on the hyperparameters obtained from step 2, the training and prediction models are constructed for grouting power prediction. Then the predicted grouting power, accuracy, and evaluation indices are calculated.

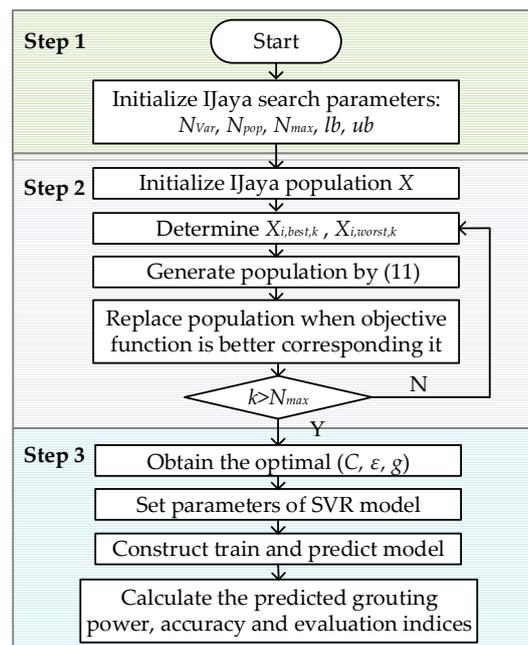


Figure 3. The flow chart of SVR optimized by IJaya.

## 5. Case Study

In this section, the proposed hybrid model is applied to predict the grouting power of hydropower project X in China, and its performance is analyzed and discussed. The data collection is described in Section 5.1. Section 5.2 provides the performance evaluation criteria. In Section 5.3, the prediction results of the EPIJaya-SVR model are reported. All experiments are conducted in MATLAB R2017b on Windows 10 with a 2.80 GHz Intel Core i7-1065G7 CPU, 64-bit and 16 GB RAM.

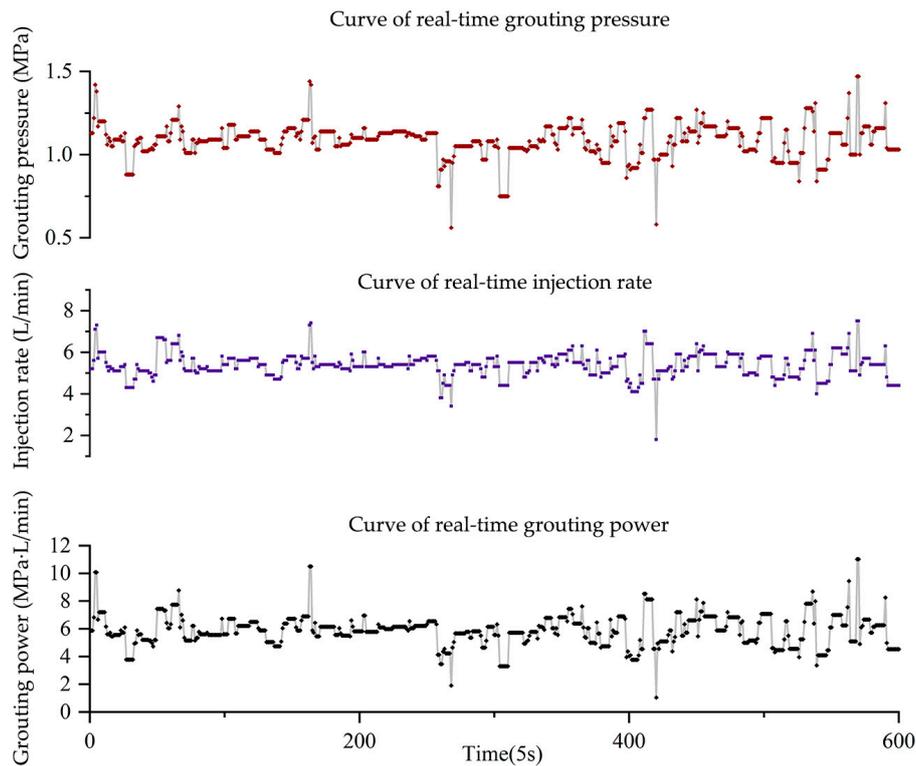
### 5.1. Data Collection

The proposed hybrid model is applied to a case study to forecast the grouting power of consolidation grouting in hydropower project X, which is located on the Yalong River in Sichuan Province in southwestern China. Hydropower project X is currently under construction; once completed, it will have a hydroelectric generating capacity of 1500 MW and the dam height will reach 155 m. The whole dam can be divided into 17 continuous segments. One grouting hole is selected randomly to serve as the research object. A grouting power dataset of the selected hole is collected with a time interval of 5 s based on a grouting real-time monitoring system. A total of 600 data collected during the grouting process were selected as a sample. Five-hundred data were selected to construct the training set for determining the optimal hyperparameters of the SVR model, while the remaining 100 data were selected as the testing set to evaluate the prediction performance of the obtained model. The statistical information i.e., maximum, minimum, mean, standard deviation and median of this grouting power

series are calculated and given in Table 1. Figure 4 shows the shapes of the grouting pressure, injection rate, and grouting power.

**Table 1.** Statistical descriptions of the grouting power series.

Data Set	Number	Maximum (MPa·L/min)	Minimum (MPa·L/min)	Mean (MPa·L/min)	Standard Deviation	Median (MPa·L/min)
All samples	600	11.025	1.044	5.890	1.126	5.880
Training set	500	10.512	1.044	5.901	1.043	5.880
Testing set	100	11.025	3.360	5.838	1.480	5.913



**Figure 4.** Original grouting power, grouting pressure, injection rate time series.

### 5.2. Evaluation Criteria

To compare the prediction performance of the proposed hybrid model EPIJaya-SVR with that of four selected competitors (EIJaya-SVR, IJaya-SVR, Jaya-SVR, and SVR), no single evaluation index can reflect the distributional characteristics of the errors. Therefore, four indices—the root mean square error (RMSE), mean absolute percentage error (MAPE), mean absolute error (MAE), and equal coefficient (EC)—are used as forecasting accuracy measures. The model with lower RMSE, MAE, and MAPE and higher EC (closer to 1) can predict grouting power more effectively.

In addition to these evaluation indices, the Diebold–Mariano (DM) test [51] is also adopted to verify whether the prediction performance of two models is significantly different. The null hypothesis of the DM test assumes that the prediction accuracy of the tested model is not as good as that of the reference model. To be specific, the prediction error  $e_{te,t} = x_t - \hat{x}_{te,t}$  of the tested model is greater than or equal to the prediction error  $e_{re,t} = x_t - \hat{x}_{re,t}$  of the reference model. The DM statistic can be expressed as:

$$S = \frac{\bar{\delta}}{(\hat{V}_g/N)^{1/2}} \sim N(0, 1) \tag{12}$$

where  $\bar{g}$  is the mean value of the loss function  $g_t (= e_{te,t}^2 - e_{re,t}^2)$ ;  $\hat{V}_g = \gamma_0 + 2 \sum_{t=1}^{\infty} \gamma_t$  ( $\gamma_t = \text{cov}(g_{t+1}, g_t)$ ), where  $\gamma_0$  is the variance of  $g_t$ ; and  $\hat{x}_{te,t}$  and  $\hat{x}_{re,t}$  denote the predicted values of  $x_t$  calculated by using the tested model  $te$  and reference model  $re$ , respectively.

The DM statistic obeys a standard normal distribution and  $\alpha$  is set to be the significance level.  $Z_{\alpha/2}$  is the critical value of the test. If  $|\text{DM}| \leq |Z_{\alpha/2}|$ , there is not enough evidence to suggest that the two models exist significant difference, and, if  $|\text{DM}| > |Z_{\alpha/2}|$ , the null hypothesis can be rejected.

### 5.3. Simulation

The grouting power time series was submitted into the proposed model to obtain the predicted results in the simulation. The whole flow of the simulation for the grouting power prediction is shown in the following subsections.

#### 5.3.1. Decomposition Process

As shown in Figure 4, the original grouting power series has large fluctuations, with values falling in the range (1.044, 11.025). Therefore, the EWT was employed to decompose the original grouting power series into six subseries and one residual series to denoise and reduce the modelling complexity of the grouting power series. Figure 5 illustrates the graphical representation of the decomposed results.

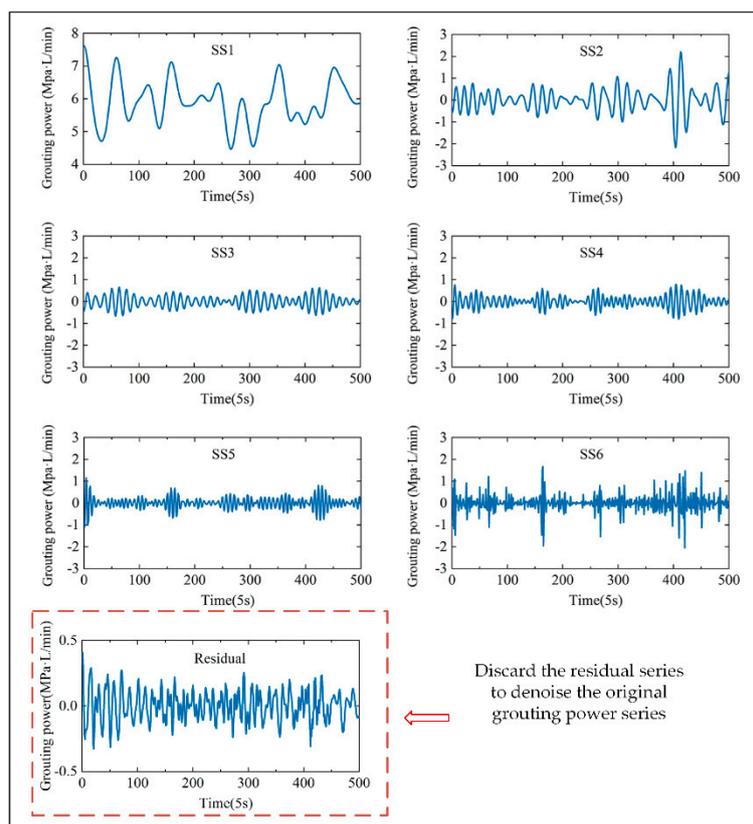
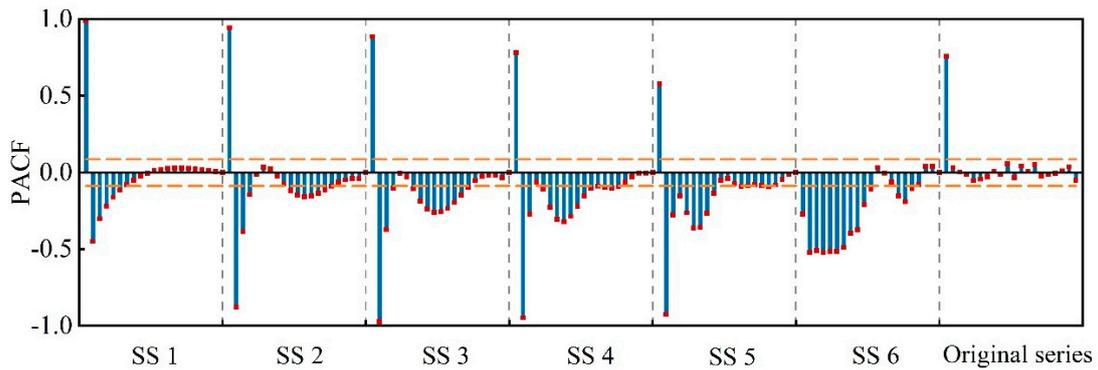


Figure 5. The decomposed subseries and residual series of the grouting power by EWT.

#### 5.3.2. PACF Results

According to the description of the PACF in Section 4.2, a correlation analysis was conducted for the data elements in each subseries. The length  $N$  of the grouting power series was 500, and its 95% confidence interval was (0.0877, +0.0877). The PACF values from lag 1 to lag 20 were calculated and are depicted in Figure 6.

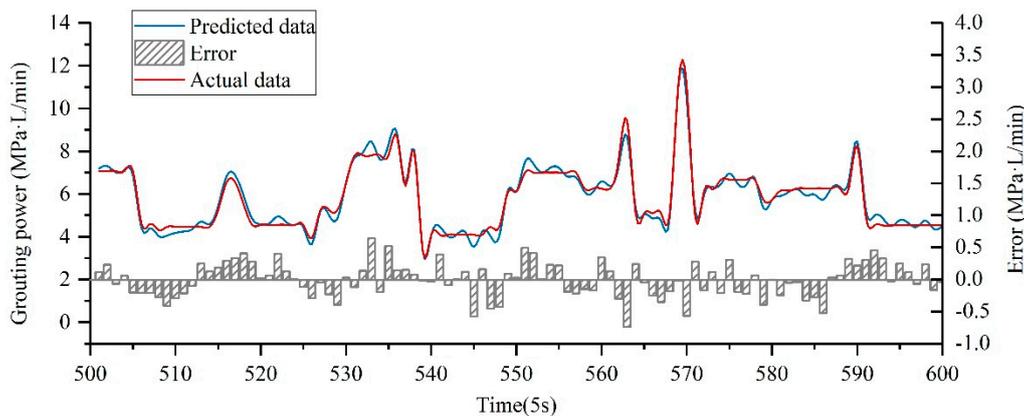


**Figure 6.** The PACFs of the original series and the subseries decomposed by EWT.

Figure 6 displays the characteristics of the partial autocorrelation of the original grouting power and the subseries. There is only one variable that can be selected to serve as the input variable for the original grouting power because of the large fluctuation. Nevertheless, more related variables can be selected for each subseries, because the subseries decomposed by the EWT have certain stability and regularity. Consequently, it is easy to establish a prediction model using the subseries than using the original grouting power series.

### 5.3.3. Prediction and Integration Processes

The SVR model optimized by using the IJaya algorithm was employed as an independent prediction tool to capture the nonlinear patterns of each subseries after the optimal input variables were identified by using the PACF. The predicted value of grouting power was then obtained by summing the predicted results of the subseries. The search ranges of the SVR algorithm were set to  $C = (1,1000)$ ,  $g = (0,10)$ , and  $\varepsilon = (0,0.2)$ . The statistical population size,  $N_{pop}$ , was set to 30 and the maximum number of iterations,  $N_{max}$ , was set to 100. The final prediction results are presented in Figure 7.



**Figure 7.** The PACFs of the original series and the subseries decomposed by EWT.

Based on accurate, real-time prediction of grouting power during the grouting process, construction workers can regulate the grouting pressure and injection rate reasonably to maintain the grouting power within a certain range, thereby avoiding the occurrence of defective conditions. From the curves of predicted data and actual data in Figure 7, it can be seen that the prediction error of grouting power based on the EPIJaya-SVR model is small (the average absolute error being 0.2165 MPa·L/min). The predicted and actual data have good consistency, which can meet actual engineering needs, and thus the model proposed here can provide practical assistance for guiding grouting construction.

## 6. Discussion

In this section, the influence of different statistical population size on IJaya algorithm and prediction performance of different models are discussed to verify the superior prediction ability of the proposed model.

### (1) Comparison of different statistical population size

Suitable statistical population size is beneficial to enhance the optimization ability of IJaya algorithm. In order to find the optimal statistical population size of IJaya algorithm in terms of grouting power prediction, different sizes which are 10, 20, 30, 40 are chosen. The comparison results are depicted in Figure 8.

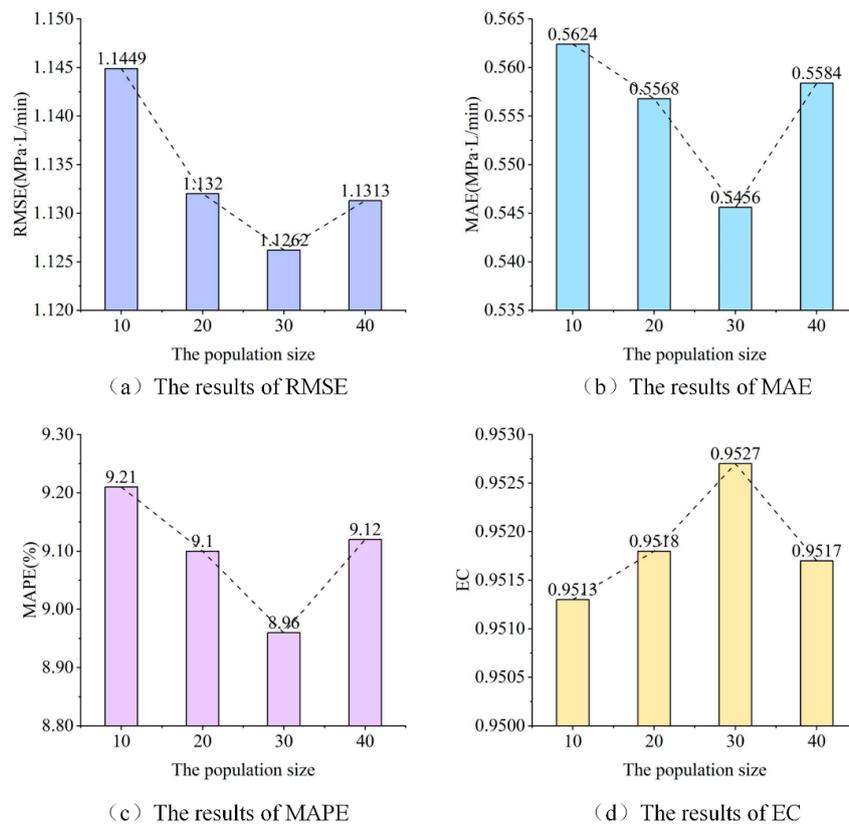


Figure 8. Selection of statistical population size of IJaya algorithm.

From Figure 8, a conclusion can be drawn that the optimization ability of IJaya algorithm is best when the statistical population size is 30. When decreasing the statistical population size of IJaya algorithm from 30, the performance of this algorithm gets worse. What’s more, when increasing the statistical population size beyond 30, the performance also gets worse. Therefore, the statistical population size is chosen as 30 in this paper.

### (2) Prediction performance of different models

It can be seen from above that the performance of EPIJaya-SVR in grouting power prediction is good, but it still needs to be compared with the other models (EIJaya-SVR, IJaya-SVR, Jaya-SVR, and SVR) considered here. EIJaya-SVR uses the historical values from lag 1 to lag 20 to construct input vectors for each subseries, and the other models all use the PACF to construct the optimal input vectors. The final prediction results of these five prediction models are depicted in Figure 9, and the results of the four performance indices and the DM test are listed in Table 2 and illustrated in Figure 10.

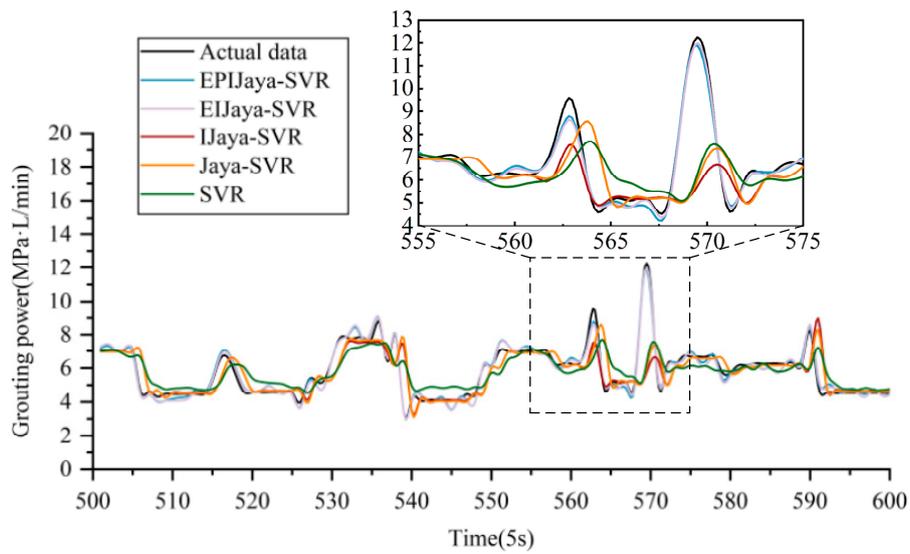


Figure 9. Actual grouting power and the prediction results of different models.

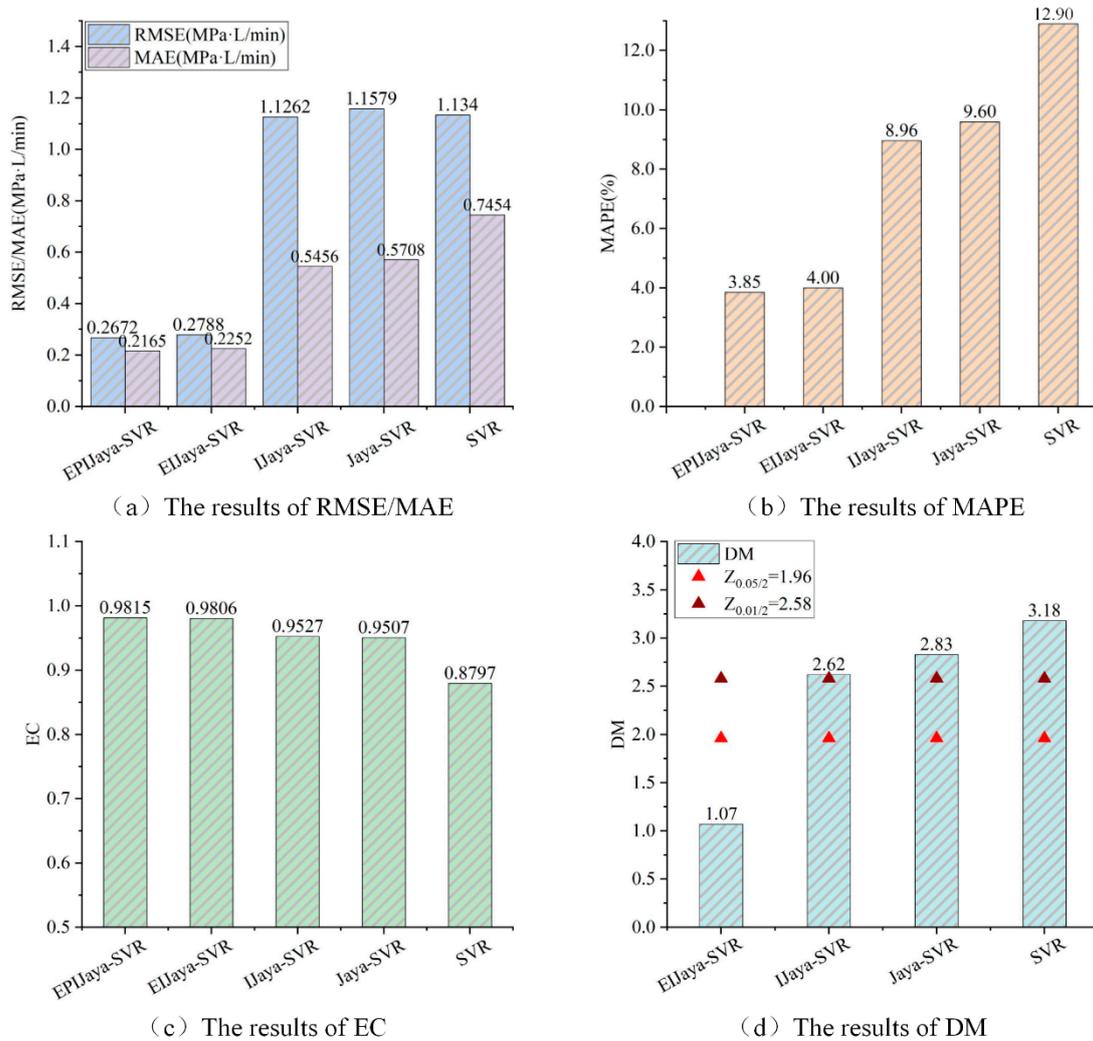


Figure 10. The results of performance indices and DM test of different models.

**Table 2.** The results of evaluation indices and DM test of different models.

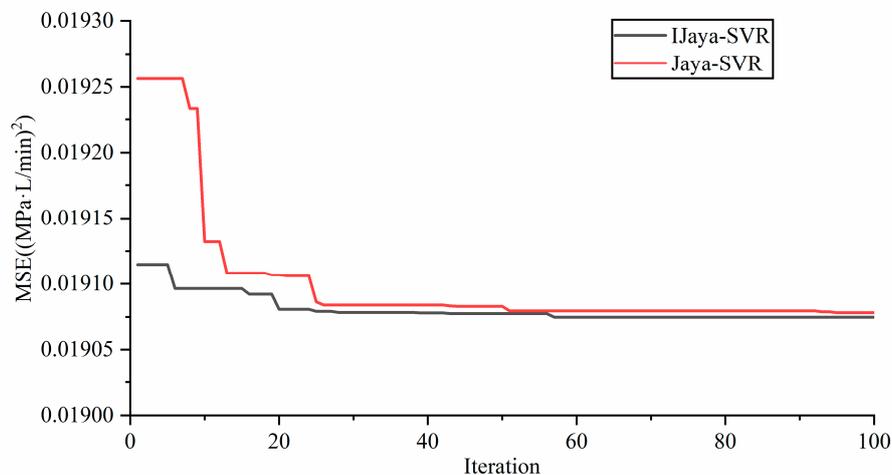
Model	RMSE (MPa·L/min)	MAE (MPa·L/min)	MAPE (%)	EC	DM
EPIJaya-SVR	0.2672	0.2165	3.85%	0.9815	-
EIJaya-SVR	0.2788	0.2252	4.00%	0.9806	1.07
IJaya-SVR	1.1262	0.5456	8.96%	0.9527	2.62
Jaya-SVR	1.1579	0.5708	9.60%	0.9507	2.83
SVR	1.1340	0.7454	12.90%	0.8797	3.18

From Table 2 and Figure 10, through the comparison of all the models, it is obvious that the EPIJaya-SVR model shows the best performance by obtaining the best RMSE, MAE, MAPE, and EC values of 0.2672 MPa·L/min, 0.2165 MPa·L/min, 3.85%, and 0.9815, respectively. Though the prediction performance of the EIJaya-SVR model with RMSE, MAE, MAPE, and EC values of 0.2788 MPa·L/min, 0.2252 MPa·L/min, 4.00%, and 0.9806, respectively, is little worse than that of the EPIJaya-SVR model, it still performs better than any other single model. This means that the EWT plays a significant role in improving prediction accuracy by denoising and transforming the nonlinear grouting power series into several relatively stable and regular subseries.

In general, the performance of forecasting models is supposed to be enhanced by a carefully selected set of input variables because the nonlinear features of grouting power can be effectively captured. However, the input variables for predictors are determined artificially, which leads to poor prediction accuracy. In this study, the PACF was employed to select variables objectively, thereby effectively avoiding the subjectivity and randomness caused by manual selection. As given in Table 2 and shown in Figure 10, the proposed hybrid model betters EIJaya-SVR with reductions of about 4.16%, 3.86%, and 3.75% in the values of RMSE, MAE, and MAPE and a 0.09% rise in the EC value. Hence, utilizing the PACF to select variables reasonably as the input vectors can improve prediction accuracy.

It also can be clearly observed from Table 2 and Figure 10 that IJaya-SVR and Jaya-SVR models outperform the SVR model with lower values of RMSE, MAE, MAPE and higher values of EC. This demonstrates the prediction ability of SVR model can be improved by adopting optimization algorithm to optimize hyperparameters. Besides, IJaya-SVR offers significant prediction superiority over Jaya-SVR with lower values of RMSE (1.1262 MPa·L/min), MAE (0.5456 MPa·L/min), and MAPE (8.96%) and a higher EC (0.9527). To some extent, these results come from the IJaya algorithm using tent chaotic mapping to generate the initial statistical population and introducing Lévy flights into the statistical population updating process in the basic Jaya algorithm to obtain the preferable hyperparameters ( $C, \epsilon, g$ ) of the prediction model; therefore, IJaya-SVR can achieve better precision.

The convergence curve is illustrated in Figure 11.



**Figure 11.** Convergence curves of different algorithms.

The DM test was also adopted to identify the statistical significance of the above models to further evaluate the prediction performance of the tested models. From the results in Table 2 and Figure 10, though the superiority of the proposed model over EIJaya-SVR is not obvious, the proposed model is superior to IJaya-SVR, Jaya-SVR, and SVR models at the 1% significance level, with the test results all above the critical value of the 1% significance level ( $Z_{0.01/2} = 2.58$ ). A conclusion obtained from the above analysis is that the hybrid model offers better performance than single models based on statistical thinking.

Actually, because the original grouting power is influenced by complex geological and construction conditions during the grouting process, it is too complex to be predicted accurately by employing a single model. Based on the above analysis, through the organic combination of the EWT, the PACF, SVR, and the IJaya algorithm, the proposed hybrid model has the best performance among all the models considered and can be effectively used for grouting power prediction. Based on the accurate prediction results, construction workers can identify whether the defective grouting conditions will occur in advance. If the predicted grouting power is out of the certain range, regulating the grouting pressure and injection rate to maintain it within the certain range to realize real-time optimization of the grouting process by which grouting safety and quality can be guaranteed.

## 7. Conclusions

For the sake of controlling the grouting process properly to guarantee grouting safety and quality, accurate and effective methods for grouting power prediction are urgently needed. However, grouting power time series exhibit highly nonlinear characteristics, making it difficult to predict grouting power accurately using a single model. The concept of “decomposition–prediction–integration” is applied to build a hybrid prediction model for grouting power. This model achieves high accuracy by linking the merits of the EWT, the PACF, SVR, and an IJaya algorithm. To be specific, the SVR model is used to predict the grouting power, and the IJaya algorithm, which introduces chaos theory and Lévy flights into the basic Jaya algorithm to improve the algorithm’s accuracy performance, is employed to optimize the hyperparameters of SVR. Moreover, prior to prediction, the EWT as a preprocessor is employed to decompose the original grouting power time series into several subseries and one residual series adaptively, by which meaningful information from the original series is extracted, and the modelling complexity is reduced sharply. The noise in the original grouting power series can be removed by discarding the residual series. Then, the PACF is utilized to determine the optimal input variables for the predictor of each subseries by calculating the correlation between the data to avoid the subjectivity caused by manual selection. Real-word simulations demonstrate that the proposed hybrid model has better prediction performance with higher accuracy compared with four other prediction models in terms of four evaluation indices and the DM test, and it can provide a feasible technical reference for guiding grouting construction.

In future research, the grouting power can be timely predicted by embedding the grouting power prediction model in a real-time grouting monitoring system. This would enable construction workers to regulate the grouting pressure and injection rate before the occurrence of defective grouting conditions, effectively guaranteeing grouting safety and quality.

**Author Contributions:** Conceptualization, L.X., Y.Z. and T.G.; methodology, L.X., Y.Z. and B.R.; formal analysis, Y.Z., T.G. and D.T.; writing—original draft preparation, L.X. and B.R.; writing—review and editing, L.X., Y.Z., T.G., B.R., D.T., and B.W. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by National Key R&D Program of China, grant number 2018YFC0406704; the National Natural Science Foundation of China, grant number 51839007; the Yalong River Joint Funds of the National Natural Science Foundation of China, grant number U1865204.

**Acknowledgments:** The authors are grateful to the editor and the reviewers of this paper, whose comments and suggestions significantly improved the quality of the paper.

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

## References

1. Zhu, Y.; Wang, X.; Deng, S.; Chen, W.; Shi, Z.; Xue, L.; Lv, M. Grouting process simulation based on 3D fracture network considering fluid-structure interaction. *Appl. Sci.* **2019**, *9*, 667. [[CrossRef](#)]
2. Warner, J. *Practical Handbook of Grouting: Soil, Rock, and Structures*; John Wiley & Sons: New York, NY, USA, 2004.
3. Xia, K. *Grouting Technology Collected Works*; China Water & Power Press: Beijing, China, 2016; pp. 52–57.
4. Tran, H.-H.; Hoang, N.-D. An Artificial Intelligence Approach for Groutability Estimation Based on Autotuning Support Vector Machine. *J. Constr. Eng.* **2014**, *2014*, 109184. [[CrossRef](#)]
5. Hoang, N.D.; Bui, D.T.; Liao, K.W. Groutability estimation of grouting processes with cement grouts using Differential Flower Pollination Optimized Support Vector Machine. *Appl. Soft Comput. J.* **2016**, *45*, 173–186. [[CrossRef](#)]
6. Deng, S.; Wang, X.; Zhu, Y.; Lv, F.; Wang, J. Hybrid Grey Wolf Optimization Algorithm-Based Support Vector Machine for Groutability Prediction of Fractured Rock Mass. *J. Comput. Civ. Eng.* **2019**, *33*, 04018065. [[CrossRef](#)]
7. Yang, C.P. Estimating cement take and grout efficiency on foundation improvement for Li-Yu-Tan dam. *Eng. Geol.* **2004**, *75*, 1–14. [[CrossRef](#)]
8. Zhang, L.; Li, Q.; Song, Y. Neural network-based experimental study on shaft water sealing by grouting. *Int. Geosci. Remote Sens. Symp.* **2007**, 3142–3145. [[CrossRef](#)]
9. Fan, G.; Zhong, D.; Wang, J.; Ren, B. Cement Take Evaluation and Prediction based on Empirical Relationships and Support Vector Regression. In Proceedings of the 2016 5th International Conference on Energy and Environmental Protection, Shenzhen, China, 17–18 September 2016; pp. 699–704.
10. Öge, İ.F. Prediction of cementitious grout take for a mine shaft permeation by adaptive neuro-fuzzy inference system and multiple regression. *Eng. Geol.* **2017**, *228*, 238–248. [[CrossRef](#)]
11. Li, X.; Zhong, D.; Ren, B.; Fan, G.; Cui, B. Prediction of curtain grouting efficiency based on ANFIS. *Bull. Eng. Geol. Environ.* **2019**, *78*, 281–309. [[CrossRef](#)]
12. Attar, N.F.; Pham, Q.B.; Nowbandegani, S.F.; Rezaie-Balf, M.; Fai, C.M.; Ahmed, A.N.; Pipelzadeh, S.; Dung, T.D.; Nhi, P.T.T.; Khoi, D.N.; et al. Enhancing the prediction accuracy of data-driven models for monthly streamflow in Urmia Lake Basin based upon the autoregressive conditionally heteroskedastic time-series model. *Appl. Sci.* **2020**, *10*, 571. [[CrossRef](#)]
13. Asadi, R.; Regan, A.C. A spatio-temporal decomposition based deep neural network for time series forecasting. *Appl. Soft Comput. J.* **2020**, *87*, 105963. [[CrossRef](#)]
14. Cao, E.; Bao, T.; Gu, C.; Li, H.; Liu, Y.; Hu, S. A Novel Hybrid Decomposition—Ensemble Prediction Model for Dam Deformation. *Appl. Sci.* **2020**, *10*, 5700. [[CrossRef](#)]
15. Yang, H.F.; Chen, Y.P.P. Hybrid deep learning and empirical mode decomposition model for time series applications. *Expert Syst. Appl.* **2019**, *120*, 128–138. [[CrossRef](#)]
16. Shakeel, A.; Tanaka, T.; Kitajo, K. Time-series prediction of the oscillatory phase of eeg signals using the least mean square algorithm-based ar model. *Appl. Sci.* **2020**, *10*, 3616. [[CrossRef](#)]
17. Xie, T.; Zhang, G.; Liu, H.; Liu, F.; Du, P. A hybrid forecasting method for solar output power based on variational mode decomposition, deep belief networks and auto-regressive moving average. *Appl. Sci.* **2018**, *8*, 1910. [[CrossRef](#)]
18. Azadeh, A.; Ghaderi, S.F.; Sohrabkhani, S. A simulated-based neural network algorithm for forecasting electrical energy consumption in Iran. *Energy Policy* **2008**, *36*, 2637–2644. [[CrossRef](#)]
19. Hong, W.C. Chaotic particle swarm optimization algorithm in a support vector regression electric load forecasting model. *Energy Convers. Manag.* **2009**, *50*, 105–117. [[CrossRef](#)]
20. Lian, C.; Zeng, Z.; Yao, W.; Tang, H. Displacement prediction model of landslide based on a modified ensemble empirical mode decomposition and extreme learning machine. *Nat. Hazards* **2013**, *66*, 759–771. [[CrossRef](#)]
21. Niu, W.-j.; Feng, Z.-k.; Zeng, M.; Feng, B.-f.; Min, Y.-w.; Cheng, C.-t.; Zhou, J.-z. Forecasting reservoir monthly runoff via ensemble empirical mode decomposition and extreme learning machine optimized by an improved gravitational search algorithm. *Appl. Soft Comput. J.* **2019**, *82*, 105589. [[CrossRef](#)]
22. Drucker, H.; Surges, C.J.C.; Kaufman, L.; Smola, A.; Vapnik, V. Support vector regression machines. *Adv. Neural Inf. Process. Syst.* **1997**, *1*, 155–161.

23. Bagheri, M.; Rezaei, H. Reservoir rock permeability prediction using SVR based on radial basis function kernel. *Carbonates Evaporites* **2019**, *34*, 699–707. [[CrossRef](#)]
24. Luo, X.; Yuan, X.; Zhu, S.; Xu, Z.; Meng, L.; Peng, J. A hybrid support vector regression framework for streamflow forecast. *J. Hydrol.* **2019**, *568*, 184–193. [[CrossRef](#)]
25. Liu, D.; Niu, D.; Wang, H.; Fan, L. Short-term wind speed forecasting using wavelet transform and support vector machines optimized by genetic algorithm. *Renew. Energy* **2014**, *62*, 592–597. [[CrossRef](#)]
26. Shiu, M.C.; Wei, L.Y.; Liu, J.W.; Huang, D.Y.; Tu, C.C.; Liao, K.H. A hybrid one-step-Ahead time series model based on GA-SVR and EMD for forecasting electricity loads. *J. Appl. Sci. Eng.* **2017**, *20*, 467–476. [[CrossRef](#)]
27. Cheng, H.; Ding, X.; Zhou, W.; Ding, R. A hybrid electricity price forecasting model with Bayesian optimization for German energy exchange. *Int. J. Electr. Power Energy Syst.* **2019**, *110*, 653–666. [[CrossRef](#)]
28. Dash, P.K.; Majumder, I.; Nayak, N.; Bisoi, R. Point and Interval Solar Power Forecasting Using Hybrid Empirical Wavelet Transform and Robust Wavelet Kernel Ridge Regression. *Nat. Resour. Res.* **2020**, *29*, 2813–2841. [[CrossRef](#)]
29. Zhang, Q.; Zhang, J. Short-Term Load Forecasting Method Based on EWT and IDBSCAN. *J. Electr. Eng. Technol.* **2020**, *15*, 635–644. [[CrossRef](#)]
30. Hu, J.; Wang, J. Short-term wind speed prediction using empirical wavelet transform and Gaussian process regression. *Energy* **2015**, *93*, 1456–1466. [[CrossRef](#)]
31. Huang, N.; Yuan, C.; Cai, G.; Xing, E. Hybrid short term wind speed forecasting using variational mode decomposition and a weighted regularized extreme learning machine. *Energies* **2016**, *9*. [[CrossRef](#)]
32. Xie, T.; Zhang, G.; Hou, J.; Xie, J.; Lv, M.; Liu, F. Hybrid forecasting model for non-stationary daily runoff series: A case study in the Han River Basin, China. *J. Hydrol.* **2019**, *577*, 123915. [[CrossRef](#)]
33. Wang, J.; Zhong, D.; Adeli, H.; Wang, D.; Liu, M. Smart bacteria-foraging algorithm-based customized kernel support vector regression and enhanced probabilistic neural network for compaction quality assessment and control of earth-rock dam. *Expert Syst.* **2018**, *35*, e12357. [[CrossRef](#)]
34. Venkata Rao, R. Jaya: A simple and new optimization algorithm for solving constrained and unconstrained optimization problems. *Int. J. Ind. Eng. Comput.* **2016**, *7*, 19–34. [[CrossRef](#)]
35. Sawulski, J.; Ławryńczuk, M. Optimization of control strategy for a low fuel consumption vehicle engine. *Inf. Sci.* **2019**, *493*, 192–216. [[CrossRef](#)]
36. Xu, S.; Wang, Y.; Wang, Z. Parameter estimation of proton exchange membrane fuel cells using eagle strategy based on JAYA algorithm and Nelder-Mead simplex method. *Energy* **2019**, *173*, 457–467. [[CrossRef](#)]
37. Suraj, S.; Sinha, R.K.; Ghosh, S. Jaya Based ANFIS for Monitoring of Two Class Motor Imagery Task. *IEEE Access.* **2016**, *4*, 9273–9282. [[CrossRef](#)]
38. Elattar, E.E.; ElSayed, S.K. Modified JAYA algorithm for optimal power flow incorporating renewable energy sources considering the cost, emission, power loss and voltage profile improvement. *Energy* **2019**, *178*, 598–609. [[CrossRef](#)]
39. Zhong, D.H.; Yan, F.G.; Li, M.C.; Huang, C.X.; Fan, K.; Tang, J.F. A Real-Time Analysis and Feedback System for Quality Control of Dam Foundation Grouting Engineering. *Rock Mech. Rock Eng.* **2015**, *48*, 1947–1968. [[CrossRef](#)]
40. Gilles, J. Empirical wavelet transform. *IEEE Trans. Signal. Process.* **2013**, *61*, 3999–4010. [[CrossRef](#)]
41. Xie, G.; Qian, Y.; Yang, H. Forecasting container throughput based on wavelet transforms within a decomposition-ensemble methodology: A case study of China. *Marit. Policy Manag.* **2019**, *46*, 178–200. [[CrossRef](#)]
42. Mi, X.; Liu, H.; Li, Y. Wind speed prediction model using singular spectrum analysis, empirical mode decomposition and convolutional support vector machine. *Energy Convers. Manag.* **2019**, *180*, 196–205. [[CrossRef](#)]
43. Devi, A.S.; Maragatham, G.; Boopathi, K.; Rangaraj, A.G. Hourly day-ahead wind power forecasting with the EEMD-CSO-LSTM-EFG deep learning technique. *Soft Comput.* **2020**, *7*. [[CrossRef](#)]
44. Sun, G.; Chen, T.; Wei, Z.; Sun, Y.; Zang, H.; Chen, S. A carbon price forecasting model based on variational mode decomposition and spiking neural networks. *Energies* **2016**, *9*, 54. [[CrossRef](#)]
45. Chou, J.S.; Yang, K.H.; Lin, J.Y. Peak Shear Strength of Discrete Fiber-Reinforced Soils Computed by Machine Learning and Metaensemble Methods. *J. Comput. Civ. Eng.* **2016**, *30*, 04016036. [[CrossRef](#)]
46. Wang, J.; Zhong, D.; Wu, B.; Shi, M. Evaluation of Compaction Quality Based on SVR with CFA: Case Study on Compaction Quality of Earth-Rock Dam. *J. Comput. Civ. Eng.* **2018**, *32*, 05018001. [[CrossRef](#)]

47. Smola, A.J.; Olkoph, B.S.C.H. A Tutorial On Support Vector Regression. *Stat. Comput.* **2004**, *14*, 199–222. [[CrossRef](#)]
48. Alatas, B. Chaotic bee colony algorithms for global numerical optimization. *Expert Syst. Appl.* **2010**, *37*, 5682–5687. [[CrossRef](#)]
49. Yu, J.T.; Kim, C.H.; Wadood, A.; Khurshaid, T.; Rhee, S.B. Jaya Algorithm with Self-Adaptive Multi-Population and Lévy Flights for Solving Economic Load Dispatch Problems. *IEEE Access* **2019**, *7*, 21372–21384. [[CrossRef](#)]
50. Viswanathan, G.M.; Afanasyev, V.; Buldyrev, S.V.; Murphyt, E.J.; Princt, P.A.; Stanley, H.E. Levy flight search patterns of wandering albatrosses. *Nature* **1996**, *381*, 413–415. [[CrossRef](#)]
51. Mariano, R.S. Comparing Predictive Accuracy AU—Diebold, Francis X. *J. Bus. Econ. Stat.* **2002**, *20*, 134–144. [[CrossRef](#)]

**Publisher’s Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).