

A combined prediction approach based on wavelet transform for crop water requirement

Zhongda Tian

ABSTRACT

The accurate prediction of crop water requirement is of great significance for the development of regional agriculture. Based on the wavelet transform, a combined prediction approach for crop water requirement is proposed. Firstly, the Mallat wavelet transform algorithm is used to decompose and reconstruct the crop water requirement series. The approximate and detail components of the original series can be obtained. The characteristics of approximate components and detail components are analyzed by Hurst index. Then, according to the different characteristics of the components, the particle swarm optimization algorithm optimized support vector machine is used to predict the approximate component, and the autoregressive moving average model is used to predict the detail components. Three-fold cross-validation is used to improve the generalization ability of the forecasting model. Finally, combined with the prediction value of each prediction model, the final prediction value of crop water requirement is obtained. The crop water requirement data from 1983 to 2018 in Liaoning Province of China are collected as the research object. The simulation results indicate that the proposed combined prediction approach has high prediction accuracy for crop water requirement. The comparison of performance indicators shows that the root mean square error of the proposed prediction approach reduced by 45.40% to 57.16%, mean absolute error reduced by 32.96% to 52.07%, mean absolute percentile error reduced by 33.02% to 52.37%, relative root mean square error reduced by 45.26% to 57.38%, square sum error reduced by 70.18% to 80.42%, and the Theil inequality coefficient reduced by 59.02% to 80.77%. R square increased by 16.46% to 54.77%, and the index of agreement increased by 3.82% to 23.37%. The results of Pearson's test and the DM test show that the association strength between the actual value and the prediction value of the crop water requirement is stronger. Moreover, the proposed prediction approach in this paper has higher reliability under the same confidence level. The effectiveness of the proposed prediction approach for crop water requirement is verified. The proposed prediction approach has great significance for the rational use of water resources, planning and management, promoting social and economic sustainable development.

Key words | autoregressive moving average, combined prediction, crop water requirement, particle swarm optimization, support vector machine, wavelet transform

Zhongda Tian

College of Information Science and Engineering,
Shenyang University of Technology,
Shenyang 110870,
China
E-mail: tianzhongda@126.com

INTRODUCTION

Water resource is the most important and essential of natural resources. As a global strategic economic resource, water resource is an evaluation component of comprehensive

strength of various countries and regions. However, water resources are not infinite renewable resources. If they cannot be reasonably used and protected, the water

resources on the earth may be exhausted one day. Many countries in the world are seriously short of water, which means that there is a serious shortage of agricultural water resources. The effective utilization of water resources is of strategic significance for the sustainable development of agriculture and even the whole national economy. However, with the increase of population and the improvement of people's living standards, the shortage of water resources is becoming more and more serious. Crop water consumption is an important part of world economic water consumption (Serrano-Coronel *et al.* 2018; Xue *et al.* 2019). Crop water consumption accounts for about 70%–80% of the world's total water consumption. Therefore, in order to realize the sustainable development of water resources, it is very important to use crop water resources reasonably. Crop water consumption has become an important indicator (Chlingaryan *et al.* 2018). Accurate prediction of crop water requirement is the basis of water resource comprehensive utilization planning. Therefore, reasonable prediction of crop water requirement and scientific long-term water supply and demand planning can provide a very necessary reference for strengthening water requirement management and economic development (Akhbari & Grigg 2014). According to different purposes and objects of water requirement prediction, the prediction of water requirement can be divided into short-term prediction, medium-term, and long-term prediction (Anele *et al.* 2017). The short-term prediction is generally a daily and hourly forecast for optimal control of water-use systems. Short-term prediction requires high precision and fast prediction speed. Long-term prediction generally refers to the annual forecast for the purpose of water resources planning, which requires a long forecast cycle and many factors to be considered. Based on the characteristics of various water requirement prediction results, the current agricultural water requirement prediction methods can be basically divided into four categories (Guo *et al.* 2020), as follows.

- (A) The first one is the water requirement prediction method based on statistical law. This method is based on the statistics of water consumption in the past or other statistics related to the past to analyze and study the regularity. According to the obtained regularity, the water requirement is calculated by interpolation and extension. The widely used methods are the comparison and extension method, empirical formula method, growth rate method and regression analysis method. Water requirement prediction methods based on statistical laws need long statistical data, and the trend of future water requirement growth should be consistent with the previous situation. In fact, it is impossible for the water requirement to increase infinitely along a past trend (St-Laurent *et al.* 2018).
- (B) The second is the water requirement prediction method based on water consumption mechanism. This method studies the growth law of the water requirement from internal mechanisms, and then forecasts the water requirement. Through research on the mechanism of water consumption and the supplement of each user, the relevant water balance relationship is established, and the water consumption law is mainly studied. Because of different users, the corresponding water consumption mechanism is also different. Therefore, this water consumption prediction method lacks generality (Alkhudhiri *et al.* 2019).
- (C) The third is the combination of the first and the second category, mainly the water quota prediction method. The water demand depends on the water consumption per unit index, that is, the quantity of total quota. It can be seen that the key to using the water quota prediction method lies in the prediction of net water quota and the prediction of the irrigation water utilization coefficient. To some extent, this kind of water requirement prediction is the prediction of the water quota. As a result, there is a lack of flexibility (He *et al.* 2018).
- (D) The fourth is the prediction method based on the model. The prediction method based on the model is to establish the water requirement model, determine the various boundary conditions of the water requirement model, so as to determine the relationship between the future water requirement and the current water requirement. Because of the complexity of a water system, it is impossible to establish a certain theoretical model to describe it, so the vast majority of model-based water requirement prediction models are based on the statistical analysis of historical data, and the only difference is the data processing method and application characteristics.

A number of factors affect crop requirement, such as society, economy, population, environment and water resource reserve (Wang & Zhang 2018). Some statistical data of the study indicate that crop water demand generally has a strong cyclical change. Due to the complexity of the water-using system, it is impossible to analyze and model it by theory, so the prediction method based on models is widely used and has achieved many results. At present, many scholars have put forward prediction methods or models for crop water requirement. In their study, Najafi & Asgari (2008) put forward a crop water requirement forecasting method based on evapotranspiration and the Hargreaves–Samani equation. But, the proposed prediction method needs many climate parameters. In their literature, Peng *et al.* (2019) use a back propagation (BP) neural network to predict crop water requirement. In the study, field crops in China are taken as research objects. However, the BP neural network is prone to fall into a local optimum, so it will affect the final prediction accuracy. The authors put forward the idea of producing fuzzy rules by genetic algorithms based on the Takagi–Surgeon fuzzy logic system from the dataset of multidimensional climate data and crop water requirements, and establishing the fuzzy model to predict crop water requirement (Zhang *et al.* 2008). The prediction model is tested and the result indicated that it is an effective method to forecast crop water requirements by a fuzzy rules model. However, the establishment of fuzzy rules relies too much on manual experience to adapt to the complex changes in the agricultural system. The authors use the least squares support vector machine (LSSVM) to predict and compute crop irrigation requirements (Guo *et al.* 2011). The drawback of their proposed prediction method is that LSSVM parameters are hard to determine. The authors also did not give a specific method to determine these parameters. In the literature (Ahmed 2018), multiple linear regression algorithms are used to forecast water requirement. Crop water requirement is a complex nonlinear system, thus, linear models are not suitable for water requirement forecasting. In the literature (Abishek *et al.* 2017), some important factors such as land evaporation transpiration rates, humidity, groundwater characteristics, and temperature have been considered to predict crop water needs in the Bijapur

district of Karnataka. However, the predictive models established by the authors are too complex and require too many parameters. Although these prediction methods of crop water requirement have achieved some results, they also have their own shortcomings and disadvantages. The most important problem is that these methods are single prediction methods or models. Each prediction model or method has its own shortcomings. Therefore, the combined prediction approach is an important research direction to enhance prediction accuracy.

The idea of the combined prediction approach is to use multiple models to predict the research objects, and a weighted average of the prediction results of multiple prediction models, so as to improve the prediction effect (Xiang *et al.* 2018; Samalot *et al.* 2019). In the field of water requirement prediction, some research results are presented. In the literature (Altunkaynak & Nigussie 2018), the combined first-order differencing–multilayer perceptron and combined linear detrending–multilayer perceptron are developed and compared with the stand-alone multilayer perceptron (MP) model for predicting the monthly water consumption of Istanbul. The results of the study show that the proposed method is a reliable pre-processing technique for monthly water demand prediction. In the study of Al-Zahrani & Abo-Monasar (2015), the combined technique of artificial neural networks and time series models is constructed based on the available daily water consumption. Results indicate that combined models give better prediction effect. Hu *et al.* (2019), proposed a hybrid model based on a convolutional neural network and bidirectional long-term and short-term memory networks for urban water demand prediction. The simulation results show that the proposed hybrid prediction model has smaller prediction errors. Many research results show that the combined prediction model has better prediction effect.

In this paper, the combined prediction approach is adopted to predict crop water requirement. The proposed combined prediction approach combines different prediction models, and improves prediction accuracy and reliability. Based on this consideration, this paper uses the wavelet transform to decompose and reconstruct the original crop water requirement, and obtains the approximate and detail components of the crop water requirement. According to the

Hurst index results for components, the approximate components have the characteristics of long correlation and nonlinearity. It is suitable to choose the support vector machine (SVM) as the prediction model. At the same time, particle swarm optimization (PSO) is used to optimize the parameters of the SVM model. The detail component has the characteristics of short correlation and stability, so autoregressive moving average (ARMA) is chosen as the prediction model. Finally, the predicted values of each component are added to get the final predicted values. Because each component is predicted by an independent and appropriate prediction model, the prediction approach has high prediction accuracy. The validity of the proposed prediction approach is verified by actual collected crop water requirement data.

The main novelty and contributions of this paper can be described as follows.

1. The wavelet transform is introduced to decompose and reconstruct the original crop water requirement sequence. The approximate and detail components are obtained.
2. The Hurst index is used to analyze the characteristics of components. According to different characteristics, each component selects the appropriate prediction model.
3. PSO algorithm optimized SVM is adopted to predict the approximate components. ARMA is adopted to predict the detail components.
4. The three-fold cross-validation method is used to improve the generalization ability of the SVM and ARMA model.
5. The final prediction value can be obtained by adding the predicted values of detail components and the predicted values of approximate components.

The content of this paper is as follows. The first section reviews the current research status of crop water requirement prediction. The second section outlines the characteristic analysis of crop water requirement based on wavelet transform. The third section introduces the preliminaries including the PSO algorithm, SVM and ARMA. The fourth section gives the implementation steps of the proposed prediction approach. The case study results are provided in the fifth section, and the validity of the proposed prediction approach is shown; and the sixth section summarizes the conclusions and future work.

THE CHARACTERISTIC ANALYSIS OF CROP WATER REQUIREMENT BASED ON WAVELET TRANSFORM

The wavelet transform uses an orthogonal basis to decompose the signal. The discrete wavelet transform consists of a series of parameters (Schimmack & Mercorelli 2018):

$$c_j(k) = \langle X, \varphi_{jk}(t) \rangle, d_j(k) = \langle X, \psi_{jk}(t) \rangle, j, k \in Z \quad (1)$$

where $\langle *, * \rangle$ is the inner product, $c_j(k)$ is the approximate component, and $d_j(k)$ is the detail component. The scale function $\varphi_{jk}(t)$ is derived from the translation and scale transformation of the mother wavelet $\varphi(t)$:

$$\varphi_{jk}(t) = 2^{-j/2} \varphi_{jk}(2^{-j}t - k) \quad (2)$$

$\varphi_j(t)$ is a low-pass filter, which can separate the low-frequency components of the input signal. The wavelet transform can decompose the signal into approximate components of large time-scale and detail components of different, small time-scale.

In this paper, the Mallat algorithm of fast discrete orthogonal wavelet transform is introduced to decompose and reconstruct the crop water requirement. Take a_j as a sequence to be decomposed, and the decomposition algorithm is as follows (Li & Chang 2015):

$$a_{j+1} = Ha_j, d_{j+1} = Gd_j, j = 0, 1, 2, \dots, N \quad (3)$$

where H is the high-pass filter and G is the low-pass filter. The Mallat algorithm decomposes the original crop water requirement sequence into approximate components of low frequency and detail components of high frequency. The approximate component can reflect the changing trend and characteristics of crop water requirement, while the detail component can reflect the dynamic factors such as disturbance. The Mallat algorithm can be used to reconstruct the decomposed crop water demand sequence, as follows:

$$a_{j-1} = a_j H^* + d_j G^*, j = 0, 1, 2, \dots, N \quad (4)$$

where H^* is the dual operator of H and G^* is the dual operator of G . The Mallat algorithm uses two interpolations to

compensate the zeros of each adjacent sequence in the input sequence, and keeps the length of the reconstructed sequence consistent after decomposition.

The Daubechies wavelet has good characteristics for non-stationary time series. However, the Daubechies wavelet with different N values has different processing effects. The larger the N value, the longer the calculation time. In this paper, the db3 wavelet is adopted through experiments and related literature (Moosavian *et al.* 2017; Parkale & Nalbalwar 2018). The number of decomposition layers is mainly related to the signal-to-noise ratio. When the signal-to-noise ratio is low, the input signal is dominated by noise. Therefore, the number of decomposition layers should be large, which is conducive to the separation of signal and noise. When the signal-to-noise ratio is high, the signal takes up a large proportion. At this time, if there are too few decomposition layers, the distortion of decomposition will be serious and the error will be large. In this paper, three-level decomposition and reconstruction is used to ensure the real-time accuracy of prediction (Özmen & Özşen 2018).

Data of crop water requirement from 1983 to 2018 in Liaoning Province of China were collected as the research object. Therefore, the total number of data is 36. These 36 groups of crop water requirement data are given in Table 1. These data can be downloaded at the website

(<http://slt.ln.gov.cn/>) of Liaoning Provincial Department of water resources. Liaoning Province is located in the south of northeast China. The distribution of rainfall is uneven, decreasing from east to west. The total amount of water resources in Liaoning Province is 34.179 billion m^3 , including 30.249 billion m^3 of surface water resources, 124.68 billion m^3 of groundwater resources, and 85.38 billion m^3 of surface water and groundwater repetition. The study area map is shown in Figure 1. There are ten large irrigation areas in Liaoning Province. The data collection method is based on irrigation area as the basic unit to calculate agricultural water requirement. Each irrigation area makes statistics on the data of rivers, canals, precipitation, etc. within its jurisdiction, and reports the data of agricultural water requirement to the management department in combination with its own geographic location and other information. The management department reviews the data and then publishes it to the public.

Thirty-six groups of crop water requirement data are decomposed by db3 wavelet. Figure 2 shows the original data of crop water requirement, the approximate component (Ca3), and detail components (Cd1, Cd2, Cd3) after decomposition and reconstruction.

It can be seen from Figure 2 that the approximate component of crop water requirement is close to the change trend of the original data, while the detail component becomes smooth with the increase of decomposition order. After wavelet transform, it is very necessary to choose the appropriate prediction model of approximate component and detail component. In this paper, the Hurst index of the time series is used to analyze the characteristics of approximate and detail components to determine the appropriate prediction model. The Hurst index (H) can be used to measure the similarity of time series (Deng *et al.* 2018). When H is 0.5, it shows that the time series is a random walk, and there is no correlation between events. When $H \in [0, 0.5)$, the sequence is anti-persistent. When $H \in (0.5, 1)$, the sequence is persistent, that is to say, the time series has self-similarity. The larger the H , the greater the self-similarity of the sequence is. Therefore, the larger H is, the higher the correlation degree of the time series is, and the higher the predictability. In this paper, the Hurst indexes of each component are calculated by using the R/S method (Xiao *et al.* 2019). The results are shown in Table 2.

Table 1 | Data of crop water requirement from 1983 to 2018 in Liaoning Province of China

Year	Crop water requirement ($10^8 m^3$)	Year	Crop water requirement ($10^8 m^3$)	Year	Crop water requirement ($10^8 m^3$)
1983	68.36	1995	85.20	2007	86.65
1984	69.37	1996	89.80	2008	89.21
1985	72.36	1997	92.57	2009	91.25
1986	74.27	1998	88.30	2010	90.31
1987	72.20	1999	90.70	2011	90.68
1988	76.59	2000	85.93	2012	89.30
1989	77.58	2001	82.00	2013	90.68
1990	80.10	2002	80.60	2014	91.27
1991	84.26	2003	80.40	2015	92.25
1992	85.27	2004	82.00	2016	92.67
1993	84.26	2005	83.95	2017	93.36
1994	86.80	2006	87.33	2018	92.28



Figure 1 | The study area map.

From the results in Table 2, the Hurst index of the approximate component is similar to the original crop water requirement, reflecting the strong nonlinear characteristics of the approximate component. Therefore, SVM with good prediction effect for nonlinear series is used to model and predict approximate components. Three detail components have larger Hurst indexes, showing a periodic stationary feature. The ARMA model can effectively analyze the correlation of periodic stationary data series, so this paper uses the ARMA model to predict the detail components with low frequency.

THE PRELIMINARIES

In this section, the preliminaries including ARMA, SVM, and the PSO algorithm used in this paper are introduced.

ARMA

The ARMA model is the most commonly used prediction model for stationary sequences (Lennon & Yuan 2019). Its expression is abbreviated as ARMA (p, q).

$$\Phi(B)x_t = \Theta(B)\varepsilon_t \quad (5)$$

where $\Phi(B) = 1 - \varphi_1 B - \dots - \varphi_p B^p$ is a p -order autoregressive coefficient polynomial; $\Theta(B) = 1 - \theta_1 B - \dots - \theta_q B^q$ is a q -order moving average coefficient polynomial; x_t , ($t = 1, 2, \dots, N$) is the time series; ε_t is a normal white-noise process with mean value of 0 and variance of σ_ε^2 ; φ_i , ($i = 1, 2, \dots, p$) and θ_i , ($i = 1, 2, \dots, q$) is the coefficient to be estimated; and B is a backward difference operator.

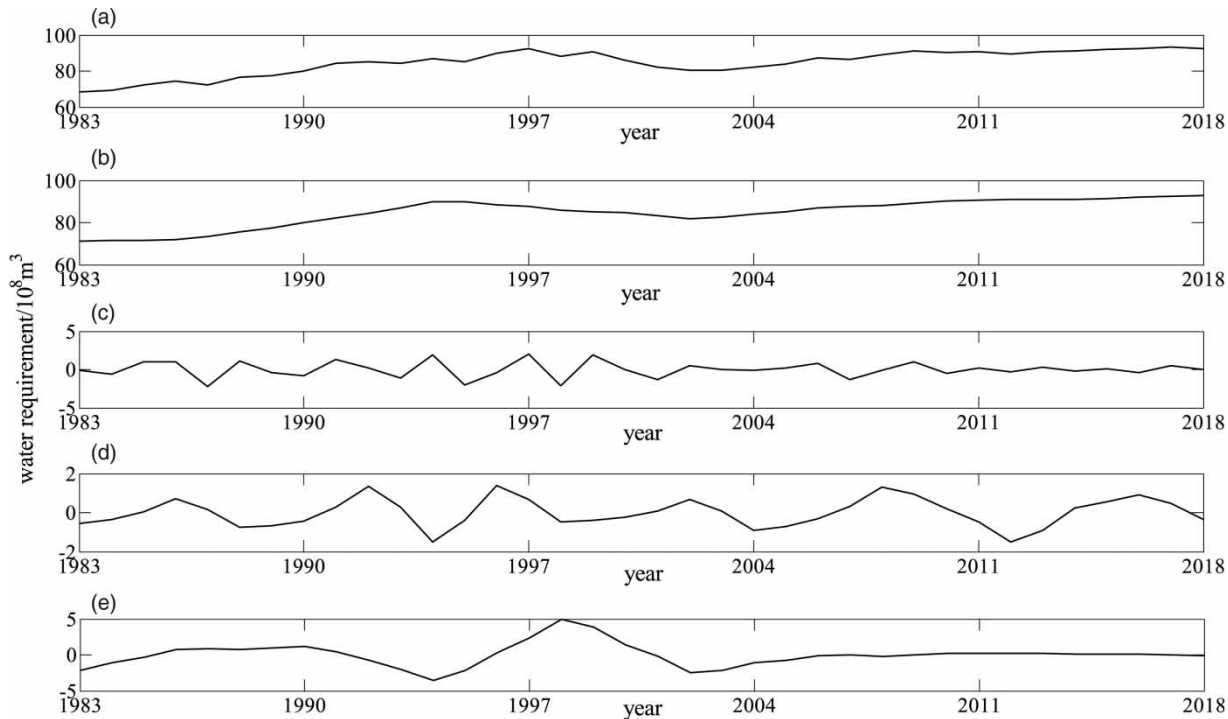


Figure 2 | Crop water requirement and its components after wavelet decomposition and reconstruction: (a) crop water requirement; (b) Ca3; (c) Cd1; (d) Cd2; (e) Cd3.

Table 2 | The Hurst index calculated by using the R/S method

Sequence	Hurst index
Crop water requirement	0.6126
Ca3	0.6352
Cd1	0.8113
Cd2	0.8762
Cd3	0.8808

In order to build the ARMA model, the order p and q of the model should be determined and fitted according to the truncation of the autocorrelation coefficient of the sample and the tailing of the autocorrelation coefficient of the error sequence. Then, the values of other unknown parameters in the model are determined by the Akaike information criterion (AIC) or least squares estimation. After all the parameters are determined, the target error sequence is fitted, and the fitting residual sequence is verified as the white-noise sample sequence. If the above conditions are met, the established model is appropriate, otherwise, the model needs to be modified or re-modeled until it passes the test.

SVM

SVM has good sparsity and generalization ability, and can solve practical problems such as those that are nonlinear, and of small sample, local minimum and high dimension. Some scholars point out that the performance of SVM is better than the traditional neural network algorithm.

Suppose that the nonlinear regression function can be expressed as:

$$f(\cdot) = \omega^T \Phi(\cdot) + b \tag{6}$$

Then the optimization problem can be shown to be:

$$\min Q(\omega, b, \xi, \xi^*) = \frac{1}{2} \|\omega\|^2 + C \sum_{i=1}^n (\xi_i + \xi_i^*) \tag{7}$$

Constraint conditions are:

$$\begin{cases} y_i - \omega^T \Phi(x_i) - b \leq \varepsilon + \xi_i \\ \omega^T \Phi(x_i) + b - y_i \leq \varepsilon + \xi_i^* \\ \xi_i, \xi_i^* \geq 0 \end{cases} \tag{8}$$

where C is the penalty coefficient and ξ and ξ^* are the non-negative slack variables; ξ and ξ^* are as follows:

$$\xi = \begin{cases} 0 & y - f(x) - \varepsilon \leq 0 \\ y - f(x) - \varepsilon & \text{others} \end{cases} \quad (9)$$

$$\xi^* = \begin{cases} 0 & \varepsilon - y - f(x) \leq 0 \\ \varepsilon - y - f(x) & \text{others} \end{cases} \quad (10)$$

where ε is an insensitive loss function.

The Lagrange function is introduced to obtain the following minimum value:

$$\begin{aligned} \min Q(\alpha_i, \alpha_i^*) = & \frac{1}{2} \sum_{i,j=1}^n (\alpha_i - \alpha_i^*)(\alpha_j - \alpha_j^*)K(x_i \cdot x_j) \\ & + \varepsilon \sum_{i=1}^n (\alpha_i + \alpha_i^*) - \sum_{i=1}^n (\alpha_i - \alpha_i^*)y_i \end{aligned} \quad (11)$$

Constraint conditions are:

$$\begin{cases} \sum_{i=1}^n (\alpha_i - \alpha_i^*) = 0, & i = 1, \dots, n \\ 0 \leq \alpha_i, \alpha_i^* \leq C \end{cases} \quad (12)$$

where $K(x_i, x_j)$ is a kernel function satisfying the Mercer conditions. The radial basis function (RBF) is a typical kernel function and can be expressed as the following:

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

where δ is the width of the RBF kernel function. The selection of the algorithm parameters has a great influence on the regression prediction performance of SVM. That means how to choose three parameters δ , ε and C is very important. In the next sub-section, the PSO algorithm is introduced to optimize these parameters.

PSO algorithm optimized SVM

The PSO algorithm is a search algorithm based on global parallelism. Its concept is simple, easy to realize and its convergence speed is fast. The PSO algorithm has been widely

used in many optimization problems (Tian *et al.* 2018; Tian *et al.* 2019a). The PSO algorithm is initialized to a group of random particles, and then the optimal solution is found by iteration. In each iteration, particles are updated according to two extreme values. The first is the optimal solution found by the particle itself, which is called the individual optimal value p_{best} . Another optimal solution is the optimal solution of the whole population called g_{best} .

In a D -dimensional target search space, N particles compose a population. The i th particle of the population is represented as an D -dimensional vector, which can be written as:

$$\mathbf{X}_i = (x_{i1}, x_{i2}, \dots, x_{iD}), \quad i = 1, 2, \dots, N \quad (14)$$

The i th particle's flying velocity is also an D -dimensional vector, which can be written as:

$$\mathbf{V}_i = (v_{i1}, v_{i2}, \dots, v_{iD}), \quad i = 1, 2, \dots, N \quad (15)$$

The i th particle's best search position by far is called the individual extremum, which can be written as:

$$p_{best} = (p_{i1}, p_{i2}, \dots, p_{iD}), \quad i = 1, 2, \dots, N \quad (16)$$

The whole population's best search position by far is called the global extremum, which can be written as:

$$g_{best} = (p_{g1}, p_{g2}, \dots, p_{gD}) \quad (17)$$

After two optimal values are found, the particles can update velocity and position according to Equations (18) and (19):

$$\begin{aligned} v_{id}^{k+1} = & \omega \times v_{id}^k + c_1 \times \text{rand}_1^k \times (P_{best\ id}^k - x_{id}^k) + c_2 \\ & \times \text{rand}_2^k \times (G_{best\ d}^k - x_{id}^k) \end{aligned} \quad (18)$$

$$x_{id}^{k+1} = x_{id}^k + v_{id}^{k+1} \quad (19)$$

where ω is inertia weight and c_1 and c_2 are learning factors. The steps of the PSO algorithm optimizing the parameters of the SVM model can be described as follows.

Step 1 The parameter initialization of the PSO algorithm includes inertia weight, learning factor, population size, maximum number of iterations, etc. The range of parameters (δ , ε and C) to be optimized is determined acting as particles.

Step 2 The training samples are predicted by the SVM model corresponding to each particle. The root mean square error (RMSE) between the predicted and actual values is taken as the fitness value of each particle. Then the current fitness value of each particle is compared with the best fitness value of the particle itself. If the fitness value of the current particle is better, select the current position of the particle as the best position of the particle.

Step 3 The optimal position of the particles is compared with the optimal position of the population. If the position of the particle is better, the optimal position of the particle is regarded as the optimal position of the population.

Step 4 According to Equations (18) and (19), the velocity and position of particles are updated.

Step 5 Determine whether the conditions for the end of optimization are met (maximum number of iterations or preset accuracy). If they are satisfied, the optimization process is completed and the optimal solution is obtained; otherwise, go to **Step 2** to continue the new round of search.

The detailed flow chart of the SVM model optimized by the PSO algorithm with upper and lower bound conditions is shown in Figure 3.

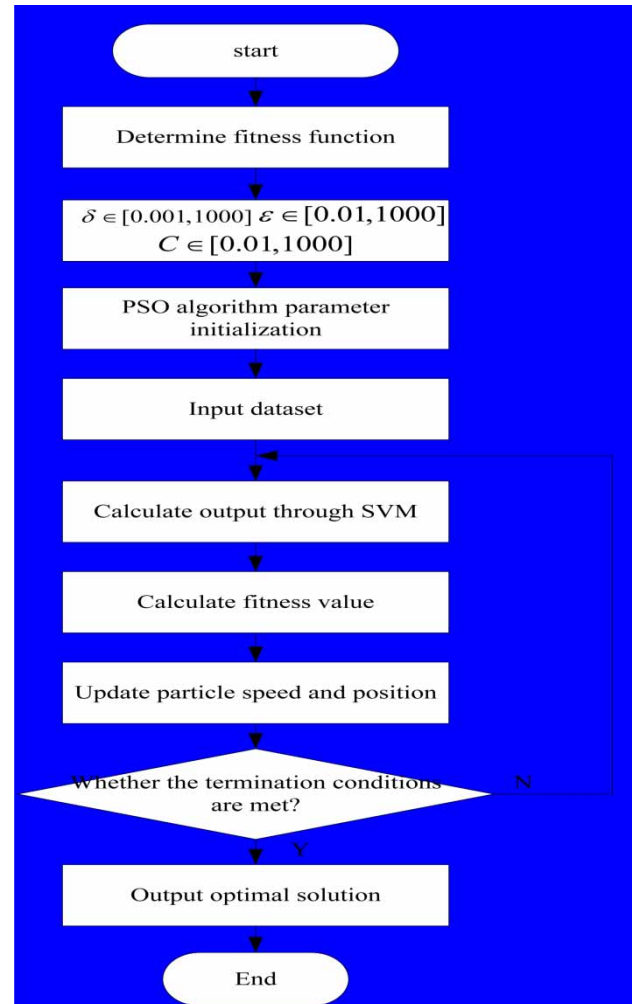


Figure 3 | The flow chart of the SVM model optimized by the PSO algorithm with upper and lower bound conditions.

THE IMPLEMENTATION STEPS OF THE PROPOSED PREDICTION APPROACH

The combined prediction approach of crop water requirement proposed in this paper is shown in Figure 4. It can be seen from Figure 4 that the implementation steps of the proposed prediction approach are as follows.

Step 1 The crop water requirement time series with length N is $c(1), c(2), \dots, c(N)$. The db3 wavelet is used for three-level decomposition and reconstruction, and an approximate component Ca3 and three detail components Cd1, Cd2 and Cd3 are obtained. The length of each component is still N . The approximate component is a long-

correlation nonlinear time series, while the detail component is a short-correlation stationary time series.

Step 2 SVM is used to train and model the approximate component CA3, and the PSO algorithm is used to optimize the parameters of the SVM model.

Step 3 The three detail components Cd1, Cd2 and Cd3 are trained with the ARMA model, and the parameters p and q of the three models are determined by AIC.

Step 4 After the establishment of the SVM and ARMA models, the crop water requirement sequence of the test samples is decomposed and reconstructed according to Step 1, and the future s groups crop water requirement is forecast in the completed training model. The final prediction value of crop water requirement is obtained by

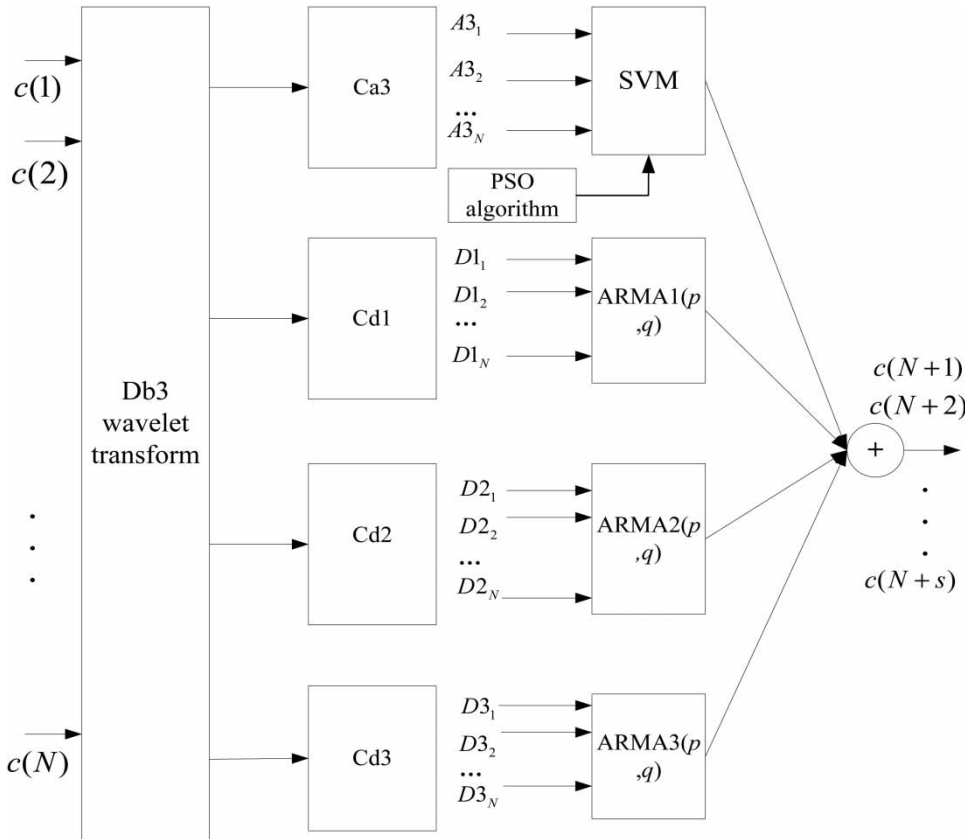


Figure 4 | The proposed combined prediction approach for crop water requirement.

adding the four prediction results. The final predictive values are $c(N + 1), c(N + 2), \dots, c(N + s)$.

CASE STUDY

The 36 groups of crop water requirement data in the section above on the characteristic analysis of crop water requirement based on wavelet transform are used for simulation. The previous 30 groups of data are used for model training and development. The latter six groups of data are used as test sets to verify the accuracy of the prediction approach. At the same time, in order to improve the generalization ability of the prediction model, the three-fold validation method is adopted. The 30 data in the training set are divided into three groups. Each time, the groups' 24 data are used for training and six data for validation. A total of three tests are conducted. RMSE is chosen as the evaluating

Table 3 | The model parameters and AIC values of three ARMA models

Detail component	Model	AIC
Cd1	ARMA(4, 3)	1.3572
Cd2	ARMA(3, 2)	1.0366
Cd3	ARMA(3, 1)	-0.9973

indicator. The minimum RMSE corresponding model is used as the final prediction model. The three approximate components of crop water requirement series are established by the ARMA prediction model, and the parameters of the model are confirmed by the AIC method. Three-fold validation is adopted. After modeling, the parameters and AIC values of the model corresponding to the minimum RMSE are as shown in Table 3.

For approximate component Ca3, SVM is used to model and predict. But the prediction effect of the SVM model is affected by parameters δ, ϵ and C . In this paper, the PSO

algorithm is used to determine the optimal parameters. The parameters of the PSO algorithm are: the number of population is 20, the maximum number of iterations is 100, the learning factor is 1.5 and 1.7, the inertia weight is 1, and the range of parameters to be optimized is $\delta \in [0.001, 1000]$, $\varepsilon \in [0.01, 1000]$, and $C \in [0.01, 1000]$. Three-fold validation is also adopted. The SVM model is trained and verified three times. The minimum RMSE corresponding model is the final SVM model. Figure 5 shows the fitness convergence curve of the PSO algorithm optimized SVM model corresponding to minimum RMSE. The optimized parameters are as follows: $\delta = 3.3281$, $\varepsilon = 0.3361$ and $C = 72.0126$.

After the parameters of the ARMA models for the detail components and the SVM prediction model for the approximate component are obtained, six groups of crop water requirement test data are predicted. Figure 6 is the prediction result of three detail components, and Figure 7 is the prediction result of the approximate component.

The final predicted value is obtained by combining the predicted value of the three detail components with the predicted value of the approximate component. Figure 8 is a comparison curve between the final predicted value by the proposed prediction approach and the actual value of crop water requirement. In order to indicate the validity of the prediction effect, this paper chooses another five state-of-

the-art prediction methods for comparison. These prediction methods are ARMA, SVM, LSSVM, BP neural network, and Elman neural network. The detailed parameters of these prediction methods are shown in Table 4. The ARMA model is determined by AIC. SVM and LSSVM are determined by the grid search method. The prediction results of these methods are also shown in Figure 9.

The actual value and prediction values of each prediction method are shown in Table 5. Table 6 gives the prediction error of each prediction method. From the results of Figure 7 and Tables 5 and 6, it can be seen that the prediction approach proposed in this paper can better predict crop water requirement. It also means the prediction error is small.

Figure 9 is the prediction error distribution histogram of the prediction methods mentioned in this study. From Figure 9, it can be known that because the prediction error is smaller and prediction error distribution is more uniform, the prediction effect and performance of the proposed approach in this paper are better.

In order to illustrate the effectiveness of the prediction method of crop water requirement, the following nine performance indicators are introduced to measure the prediction accuracy of the prediction method (Tian *et al.* 2017; Li & Han 2018).

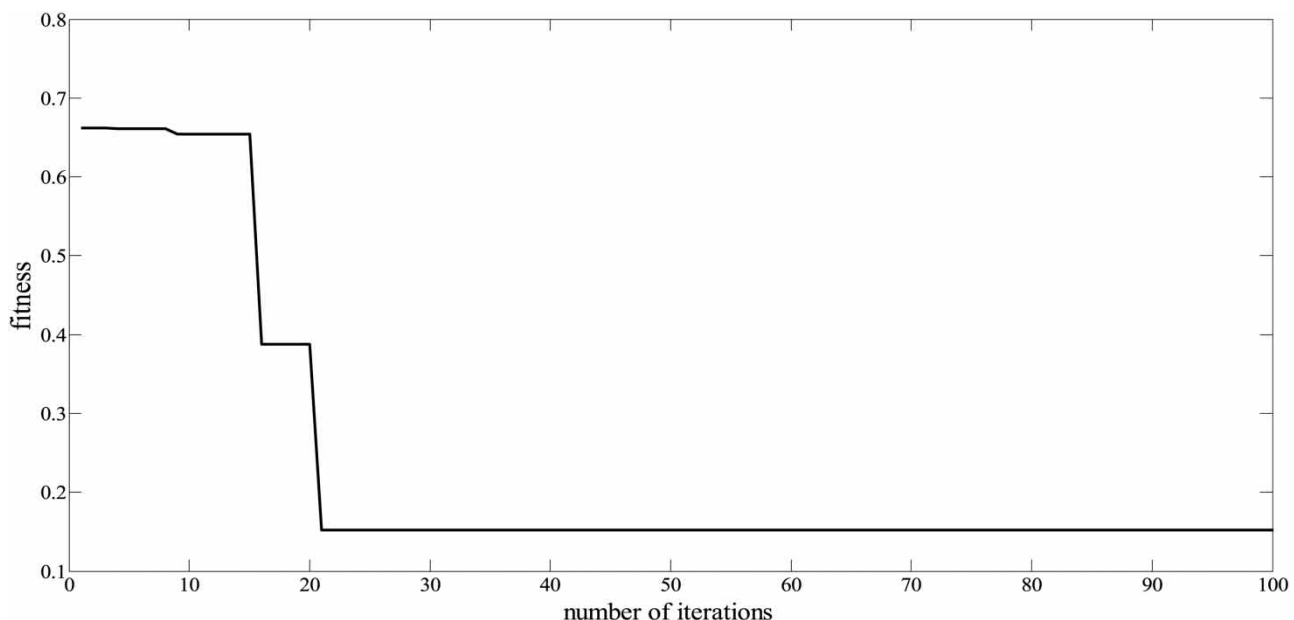


Figure 5 | The fitness curve of the PSO algorithm optimized SVM.

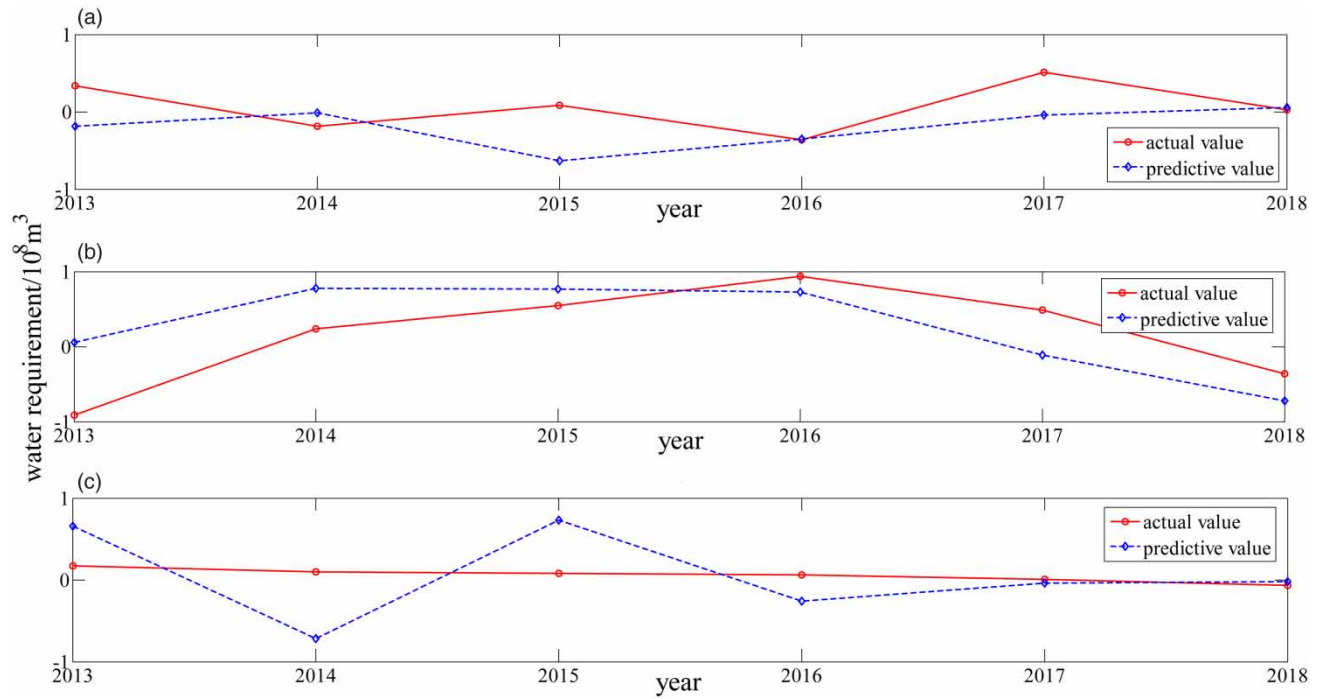


Figure 6 | The prediction result of three detail components: (a) Cd1; (b) Cd2; (c) Cd3.

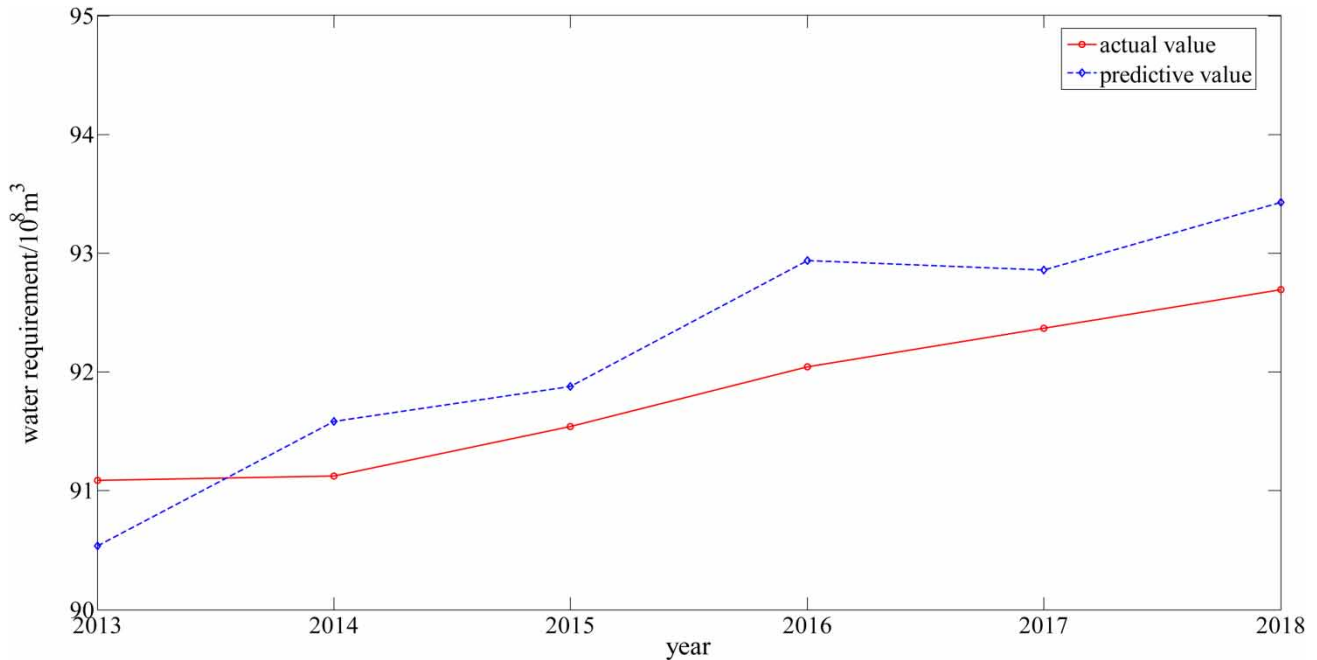


Figure 7 | The prediction result of the approximate component.

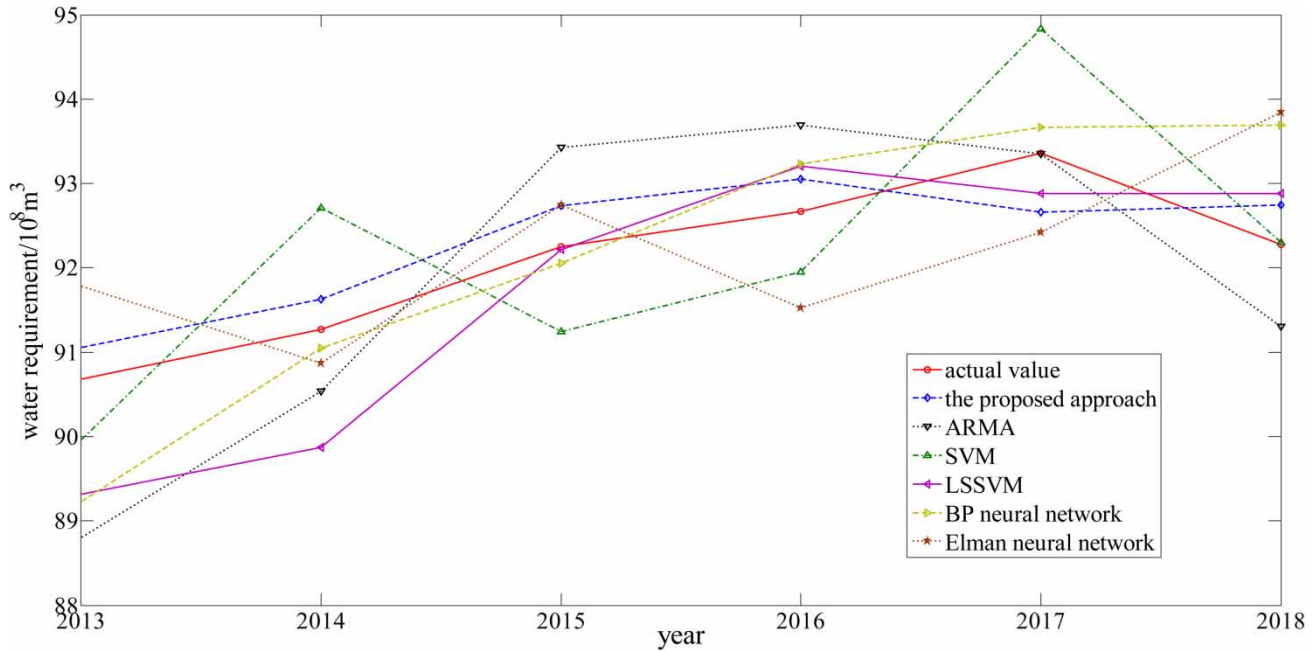


Figure 8 | The comparison of predicted results.

1. Root mean square error (RMSE):

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (c(i) - \hat{c}(i))^2} \tag{20}$$

2. Mean absolute error (MAE):

$$MAE = \frac{1}{N} \sum_{i=1}^N |c(i) - \hat{c}(i)| \tag{21}$$

3. Mean absolute percentile error (MAPE):

$$MAPE = \frac{1}{N} \sum_{i=1}^N |c(i) - \hat{c}(i)| \times 100 / c(i) \tag{22}$$

4. Relative root mean square error (RRMSE):

$$RRMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N \left(\frac{c(i) - \hat{c}(i)}{c(i)} \right)^2} \tag{23}$$

Table 4 | The parameters of the prediction methods

Method	Parameters
ARMA	p is 4, q is 2
SVM	δ is 6.2673, ϵ is 0.3659, C is 106.3680
LSSVM	σ^2 is 10.3625, γ is 77.3628
BP neural network	The input layer is 30, the middle layer is 20, the output layer is 6, the learning rate is 0.01, and the maximum number of iterations is 3,000
Elman neural network	The input layer is 30, the middle layer is 20, the output layer is 6, and the maximum number of iterations is 3,000

5. Square sum error (SSE):

$$SSE = \sum_{i=1}^N (c(i) - \hat{c}(i))^2 \tag{24}$$

6. R^2 (R square):

$$R^2 = \frac{\sum_{i=1}^N (c(i) - \bar{c})^2 - \sum_{i=1}^N (c(i) - \hat{c}(i))^2}{\sum_{i=1}^N (c(i) - \bar{c})^2} \tag{25}$$

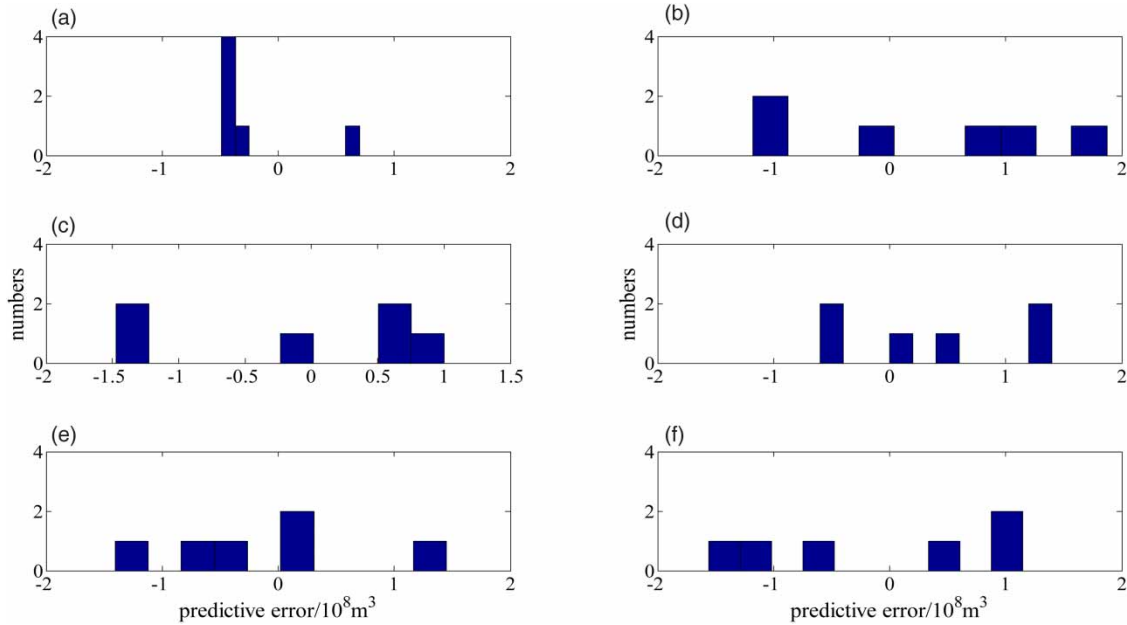


Figure 9 | The prediction error distribution histograms of the prediction methods.

Table 5 | The actual and detailed prediction values of each prediction method

Year	Actual value (10 ⁸ m ³)	Proposed approach (10 ⁸ m ³)	ARMA (10 ⁸ m ³)	SVM (10 ⁸ m ³)	LSSVM (10 ⁸ m ³)	BP neural network (10 ⁸ m ³)	Elman neural network (10 ⁸ m ³)
2013	90.68	91.0591	88.8073	89.9662	89.3185	89.2297	91.7798
2014	91.27	91.6285	90.5384	92.710	89.8733	91.0495	90.8673
2015	92.25	92.7386	93.4308	91.2480	92.2193	92.0541	92.7475
2016	92.67	93.0552	93.6887	91.9533	93.2091	93.2284	91.5229
2017	93.36	92.6590	93.3535	94.8312	92.8812	93.6670	92.4210
2018	92.28	92.7423	91.3004	92.2991	92.8772	93.6873	93.8416

Table 6 | The detailed prediction errors of each prediction method

Year	Proposed approach (10 ⁸ m ³)	ARMA (10 ⁸ m ³)	SVM (10 ⁸ m ³)	LSSVM (10 ⁸ m ³)	BP neural network (10 ⁸ m ³)	Elman neural network (10 ⁸ m ³)
2013	-0.3791	1.8727	0.7138	1.3615	1.4503	-1.0998
2014	-0.3585	0.7316	-1.4407	1.3967	0.2205	0.4027
2015	-0.4886	-1.1808	1.0020	0.0307	0.1959	-0.4975
2016	-0.3852	-1.0187	0.7167	-0.5391	-0.5584	1.1471
2017	0.7010	0.8065	-1.4712	0.4788	-0.3070	0.9390
2018	-0.4623	0.9796	-0.5191	-0.5972	-1.4073	-1.5616

7. Theil inequality coefficient (TIC):

$$TIC = \frac{\sqrt{(1/N)(c(i) - \hat{c}(i))^2}}{\sqrt{1/N \sum_{i=1}^N c(i)^2 + 1/N \sum_{i=1}^N \hat{c}(i)^2}} \quad (26)$$

8. The index of agreement (IA):

$$IA = 1 - \frac{\sum_{i=1}^N (c(i) - \hat{c}(i))^2}{\sum_{i=1}^N (|c(i) - \tilde{c}| + |\hat{c}(i) - \tilde{c}|)^2} \quad (27)$$

9. Reliability indicator:

$$R^{(1-a)} = \left[\frac{\xi^{(1-a)}}{N} - (1 - a) \right] \times 100\% \quad (28)$$

where N is the number of samples, $c(i)$ is the actual value of crop water requirement, $\hat{c}(i)$ is the prediction value of crop water requirement, \tilde{c} is the mean value of crop water requirement and $\xi^{(1-a)}$ is the number of actual values falling into the confidence interval under the confidence level $1 - \alpha$.

Furthermore, Pearson’s test and the Diebold–Mariano (DM) test are used to test prediction accuracy from the statistical perspective (Tian *et al.* 2019b). Pearson’s test can measure the association strength between the actual value and the predictive value. The results of correlation strength based on Pearson’s test are performed out to further verify the superiority of this prediction approach compared with other prediction methods. If Pearson’s correlation coefficient is equal to 1, it indicates that the actual value and the predictive value have a linear relationship. On the other hand, if Pearson’s correlation coefficient is equal to 0, there is no relationship between actual value and prediction value.

The definition of the DM test is as follows. The hypothesis tests are:

$$H_0: E(d_h) = 0, \forall n \quad (29)$$

$$H_1: E(d_h) \neq 0, \exists n \quad (30)$$

The DM test statistic value equals:

$$DM = \frac{\sum_{h=1}^k (L(\varepsilon_{t+h}^{(A)}) - L(\varepsilon_{t+h}^{(B)}))/k}{\sqrt{S^2/k}} s^2 \quad (31)$$

where ε_{t+h} is the prediction error, S^2 is an estimation value for the variance of $d_h = L(\varepsilon_{t+h}^{(A)}) - L(\varepsilon_{t+h}^{(B)})$, and L is the loss function, which is performed to measure the prediction accuracy. Two common versions of L widely used in the studies include absolute deviation error loss and square error loss. The test statistic DM is convergent to the standard normal distribution. The null hypothesis can be rejected if the following formula is satisfied:

$$|DM| > z_{\alpha/2} \quad (32)$$

where $z_{\alpha/2}$ is the critical z -value and α is the significance level. Under the assumption that the two models have the same predictive performance, the loss functions of the two models have equal unconditional expectations, that is:

$$E(d_h) = E(L(\varepsilon_1) - L(\varepsilon_2)) = 0 \quad (33)$$

If $E(d_h) > 0$, it means the predictive performance of model 2 is better than model 1. If $E(d_h) = 0$, it means the predictive performance of model 1 is the same as that of model 2. Otherwise, the predictive performance of model 1 is better than model 2.

Table 7 gives the RMSE, MAE, MAPE, RRMSE, SSE, R^2 , TIC and IA comparison of these prediction methods for crop water requirement. From the performance indicator comparison results in Table 7, RMSE, MAE, MAPE, RRMSE, SSE, and TIC of the proposed prediction approach are smaller than for the other prediction methods. In the meantime, the R^2 and IA values of the proposed prediction approach are closer to 1 than for the other prediction methods. The closer the values of R^2 and IA are to 1, the better the regression prediction performance of the model is. Therefore, the prediction accuracy of the proposed prediction approach for crop water requirement is better than that of the other prediction methods.

The results of Pearson’s test are given in Table 8. From Table 8, it can be observed that the results of Pearson’s

Table 7 | Comparison of performance indicators

Prediction methods	RMSE (10^8 m^3)	MAE (10^8 m^3)	MAPE (%)	RRMSE	SSE (10^8 m^3)	R^2	TIC	IA
Proposed approach	0.1686	0.4625	0.5013	0.0052	1.3645	0.7079	0.0026	0.9075
ARMA	0.3935	0.9650	1.0524	0.0122	7.4340	0.5914	0.0061	0.8230
SVM	0.3614	0.8939	0.9703	0.0111	6.2677	0.3417	0.0055	0.8062
LSSVM	0.3123	0.7340	0.8011	0.0097	4.6819	0.4123	0.0048	0.8656
BP neural network	0.3088	0.6899	0.7516	0.0095	4.5768	0.3202	0.0047	0.8728
Elman neural network	0.3610	0.9413	1.0215	0.0111	6.2553	0.3391	0.0055	0.6954

Table 8 | Results for Pearson's test of the prediction methods for crop water requirement

Prediction method	Pearson's correlation coefficient
Proposed approach	0.9014
ARMA	0.6883
SVM	0.8095
LSSVM	0.8114
BP neural network	0.7202
Elman neural network	0.7408

test of the proposed prediction approach are higher than those of the other prediction methods. The results show that the correlation strength between the actual value of crop water requirement and the prediction value of crop water requirement is greater than that of the other prediction methods.

The DM test values calculated by the square error loss function are listed in Table 9. From the result in Table 9, the DM test value between the proposed prediction approach and the other prediction models is greater than 0. This means that the proposed prediction approach significantly outperforms the other prediction models at different significance levels. Thus, it can reasonably be concluded

that the proposed prediction approach is superior to the other prediction methods.

Figure 10 is the reliability and confidence distribution of the prediction models mentioned in this paper. It can be observed from this figure that the proposed prediction approach has higher reliability under the same confidence level. The prediction value of crop water requirement can effectively track the variation of actual crop water requirement with higher confidence. The prediction value of crop water requirement of the proposed prediction approach falls more in the confidence interval of 90%, which shows the validity of the probability prediction results.

In conclusion, from the comparison results of the prediction value curve, prediction error distribution, performance indicators, Pearson's test and the DM test, the combined prediction approach proposed in this paper is superior to the other prediction methods. The main reason for the improvement of prediction accuracy is to reconstruct crop water requirement series by wavelet decomposition and obtain approximate and detail components. According to the different characteristics of the approximate component and the detail components, the suitable prediction model is used to improve prediction accuracy.

Table 9 | The DM test results between the proposed prediction approach and the other prediction models

DM (model 1, model 2)	1% significance level	5% significance level	10% significance level
DM (ARMA, proposed approach)	1.9277	1.1609	1.0671
DM (SVM, proposed approach)	2.3987	1.0729	1.0769
DM (LSSVM, proposed approach)	1.4157	1.0708	1.0446
DM (BP neural network, proposed approach)	1.2535	1.9303	1.0317
DM (Elman neural network, proposed approach)	2.3480	1.2260	1.0761

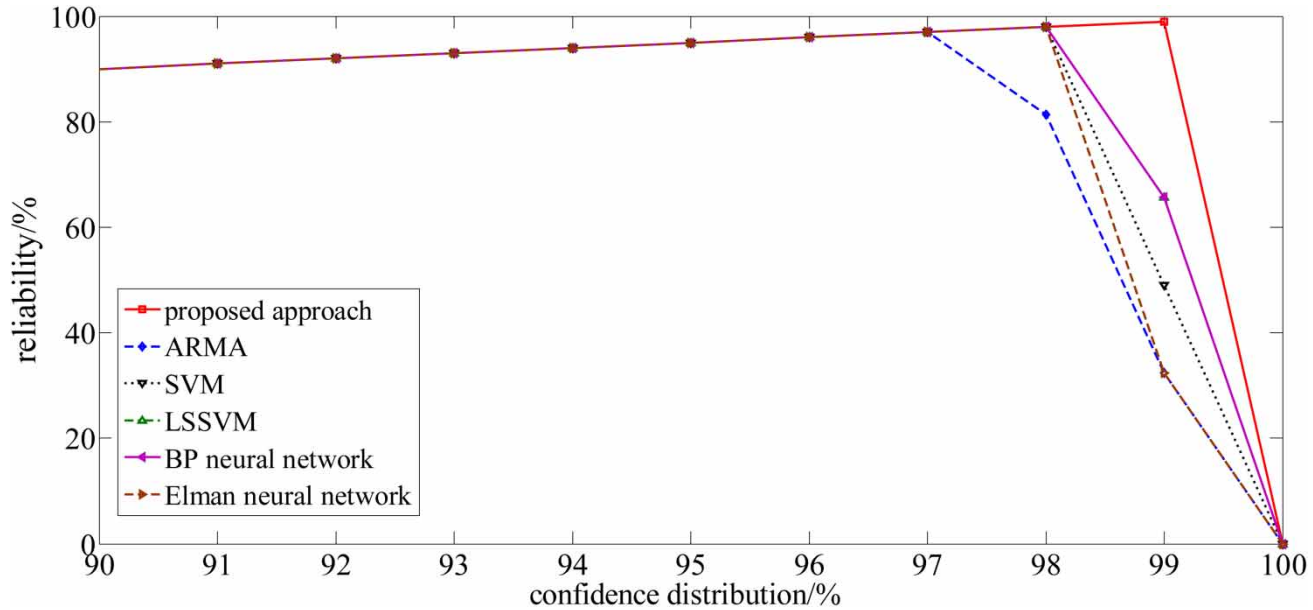


Figure 10 | Reliability and confidence distribution of the prediction models.

CONCLUSION

Crop water requirement prediction is the basis of the comprehensive utilization and planning of water resources. In order to meet the demand of sustainable development for water resources and make a scientific long-term supply-and-demand plan for water resources, it is necessary to make a reasonable prediction of crop water requirement. This paper proposed a combined prediction approach based on the wavelet transform for crop water requirement. Through the db3 wavelet transform, the crop water requirement time series with non-stationary characteristic is decomposed into a low-frequency approximate component and high-frequency detail components. The Hurst indexes of approximate components and detail components are calculated. According to the Hurst index, SVM is used to predict detail components. ARMA is used to predict detail components. Meanwhile, the PSO algorithm is introduced to optimize the parameters of SVM. The final predictive value of crop water requirement values can be obtained by the predictive value of detail components added to the predictive value of approximate components. The crop water requirement of Liaoning Province of China from 1983 to 2018 is collected as the simulation data. Compared with the state of the art, the simulation results

indicate that the proposed prediction approach for crop water requirement has higher prediction accuracy and better prediction effect. The comparison of performance indicators shows that the root mean square error of the proposed prediction approach reduced by 45.40% to 57.16%, mean absolute error reduced by 32.96% to 52.07%, mean absolute percentile error reduced by 33.02% to 52.37%, relative root mean square error reduced by 45.26% to 57.38%, square sum error reduced by 70.18% to 80.42%, and the Theil inequality coefficient reduced by 59.02% to 80.77%. R square increased by 16.46% to 54.77%, the index of agreement increased by 3.82% to 23.37%. The results of Pearson's test and the DM test show that the association strength between the actual value and the prediction value of the crop water requirement is stronger. Moreover, the proposed prediction approach in this paper has higher reliability under the same confidence level. The proposed prediction approach has great significance for the sustainable development of agriculture.

For any time series, according to different scales, we can use wavelet transform theory to decompose it into different frequency components, and obtain some behavior characteristics and structural information of the original time series. Because the decomposed signal has more of a single

frequency component than the original signal, a more stable time series can be obtained, and then appropriate models are used for prediction, which provides a powerful tool for modeling and prediction of complex time series. Future work for this study will use the proposed prediction approach to predict other complex time series to further verify the effectiveness of the method.

ACKNOWLEDGEMENTS

This paper is supported by the Science Research Project of Liaoning Education Department (No. LGD2016009) and the Natural Science Foundation of Liaoning Province of China (No. 20170540686).

CONFLICT OF INTEREST

The authors declare that they have no conflict of interest.

REFERENCES

- Abishek, B., Priyatharshini, R., Eswar, M. A. & Deepika, P. 2017 Prediction of effective rainfall and crop water needs using data mining techniques. In: *3rd IEEE International Conference on Technological Innovations in ICT for Agriculture and Rural Development*, IEEE, Piscataway, NJ, USA, pp. 231–235.
- Ahmed, F. 2018 An IoT-big data based machine learning technique for forecasting water requirement in irrigation field. In: *11th IFIP WG 8.9 Working Conference on Research and Practical Issues of Enterprise Information Systems* (A. M. Tjoa, L.-R. Zheng, Z. Zou, M. Raffai, L. D. Xu & N. M. Novak, eds), Springer, Cham, Switzerland, pp. 67–77.
- Akhbari, M. & Grigg, N. S. 2014 Water management trade-offs between agriculture and the environment: a multiobjective approach and application. *Journal of Irrigation and Drainage Engineering* **140** (8), 05014005.
- Alkudhiri, A., Bin Darwish, N. & Hilal, N. 2019 Analytical and forecasting study for wastewater treatment and water resources in Saudi Arabia. *Journal of Water Process Engineering* **32**, 100915.
- Altunkaynak, A. & Nigussie, T. A. 2018 Monthly water demand prediction using wavelet transform, first-order differencing and linear detrending techniques based on multilayer perceptron models. *Urban Water Journal* **15** (2), 177–181.
- Al-Zahrani, M. A. & Abo-Monasar, A. 2015 Urban residential water demand prediction based on artificial neural networks and time series models. *Water Resources Management* **29** (10), 3651–3662.
- Anele, A. O., Hamam, Y., Abu-Mahfouz, A. M. & Todini, E. 2017 Overview, comparative assessment and recommendations of forecasting models for short-term water demand prediction. *Water* **9** (11), 887.
- Chlingaryan, A., Sukkarieh, S. & Whelan, B. 2018 Machine learning approaches for crop yield prediction and nitrogen status estimation in precision agriculture: a review. *Computers and Electronics in Agriculture* **151**, 61–69.
- Deng, M. L., Lü, Q. F. & Zhu, W. Q. 2018 Stochastic averaging of quasi integrable and non-resonant Hamiltonian systems excited by fractional Gaussian noise with Hurst index $H \in (1/2, 1)$. *International Journal of Non-Linear Mechanics* **98**, 43–50.
- Guo, X. H., Sun, X. H. & Ma, J. J. 2011 Prediction of daily crop reference evapotranspiration (ET_0) values through a least-squares support vector machine model. *Hydrology Research* **42** (4), 268–274.
- Guo, W. Y., Liu, T., Dai, F. & Xu, P. 2020 An improved whale optimization algorithm for forecasting water resources demand. *Applied Soft Computing* **86**, 105925.
- He, Y. H., Yang, J., Chen, X. H., Lin, K. R., Zheng, Y. H. & Wang, Z. L. 2018 A two-stage approach to basin-scale water demand prediction. *Water Resources Management* **32** (2), 401–416.
- Hu, P., Tong, J., Wang, J. C., Yang, Y. & Turci, L. D. 2019 A hybrid model based on CNN and Bi-LSTM for urban water demand prediction. In: *IEEE Congress on Evolutionary Computation*, IEEE, Piscataway, NJ, USA, pp. 1088–1094.
- Lennon, H. & Yuan, J. S. 2019 Estimation of a digitised Gaussian ARMA model by Monte Carlo expectation maximisation. *Computational Statistics and Data Analysis* **133**, 277–284.
- Li, J. & Chang, L. P. 2015 A SAR image compression algorithm based on Mallat tower-type wavelet decomposition. *OPTIK* **126** (23), 3982–3986.
- Li, K. & Han, Y. 2018 Modelling for motor load torque with dynamic load changes of beam pumping units based on a serial hybrid model. *Transactions of the Institute of Measurement and Control* **40** (3), 903–917.
- Moosavian, A., Khazaei, M., Najafi, G., Khazaei, M., Sakhaei, B. & Mohammad, J. S. 2017 Wavelet denoising using different mother wavelets for fault diagnosis of engine spark plug. *Proceedings of the Institution of Mechanical Engineers, Part E: Journal of Process Mechanical Engineering* **231** (3), 359–370.
- Najafi, P. & Asgari, K. 2008 Forecasting crop water requirement by ET-HS model for arid and semi arid region of Iran. In: *2nd IFIP International Conference on Computer and Computing Technologies in Agriculture* (D. L. Li & C. J. Zhao, eds), Springer, New York, USA, pp. 529–536.
- Özmen, G. & Özşen, S. 2018 A new denoising method for fMRI based on weighted three-dimensional wavelet transform. *Neural Computing and Applications* **29** (8), 263–276.

- Parkale, Y. V. & Nalbalwar, S. L. 2018 Investigation on Daubechies wavelet-based compressed sensing matrices for ECG compression. In: *Computing, Communication and Signal Processing* (B. Iyer, S. L. Nalbalwar & N. P. Pathak, eds), Springer, Singapore, pp. 707–716.
- Peng, Y. Q., Xiao, Y. X., Fu, Z. T., Dong, Y. H., Zheng, Y. J., Yan, H. J. & Li, X. X. 2019 Precision irrigation perspectives on the sustainable water-saving of field crop production in China: water demand prediction and irrigation scheme optimization. *Journal of Cleaner Production* **230**, 365–377.
- Samalot, A., Astitha, M., Yang, J. M. & Galanis, G. 2019 Combined Kalman filter and universal Kriging to improve storm wind speed predictions for the northeastern United States. *Weather and Forecasting* **34** (3), 587–601.
- Schimmack, M. & Mercorelli, P. 2018 An on-line orthogonal wavelet denoising algorithm for high-resolution surface scans. *Journal of the Franklin Institute* **355** (18), 9245–9270.
- Serrano-Coronel, G., Chipana-Rivera, R., Moreno-Pérez, M. F. & Roldán-Cañas, J. 2018 Study of vertical water flows contribution to the crop water consumption in *suka kollus* using a mixed drainage system. *Agricultural Water Management* **206**, 86–94.
- St-Laurent, P. B., Gosselin, L. & Duchesne, C. 2018 Analysis and prediction of chilled water plant performance based on multivariate statistical methods and large historical data. *International Journal of Refrigeration* **90**, 132–144.
- Tian, Z. D., Li, S. J., Wang, Y. H., Wang, X. D. & Sha, Y. 2017 A prediction method based on wavelet transform and multiple models fusion for chaotic time series. *Chaos, Solitons and Fractals* **98**, 158–172.
- Tian, Z. D., Li, S. J., Wang, Y. H. & Wang, X. D. 2018 SVM predictive control for calcination zone temperature in lime rotary kiln with improved PSO algorithm. *Transactions of the Institute of Measurement and Control* **40** (10), 3134–3146.
- Tian, Z. D., Ren, Y. & Wang, G. 2019a Short-term wind speed prediction based on improved PSO algorithm optimized EM-ELM. *Energy Sources, Part A: Recovery, Utilization, and Environmental Effects* **41** (1), 26–46.
- Tian, Z. D., Li, S. J. & Wang, Y. H. 2019b A prediction approach using ensemble empirical mode decomposition-permutation entropy and regularized extreme learning machine for short-term wind speed. *Wind Energy* **23** (2), 177–206.
- Wang, L. N. & Zhang, H. H. 2018 An adaptive fuzzy hierarchical control for maintaining solar greenhouse temperature. *Computers and Electronics in Agriculture* **155**, 251–256.
- Xiang, Y., Gou, L., He, L. H., Xia, S. L. & Wang, W. Y. 2018 A SVR-ANN combined model based on ensemble EMD for rainfall prediction. *Applied Soft Computing* **73**, 874–883.
- Xiao, Z. K., Ding, W. L., Liu, J. S., Tian, M. Z., Yin, S., Zhou, X. H. & Gu, Y. 2019 A fracture identification method for low-permeability sandstone based on R/S analysis and the finite difference method: a case study from the Chang 6 reservoir in Huaqing oilfield, Ordos Basin. *Journal of Petroleum Science and Engineering* **174**, 1169–1178.
- Xue, J. Z., Khan, S., Sun, M., Anwar, S., Ren, A. X., Gao, Z. Q., Lin, W., Xue, J. F., Yang, Z. P. & Deng, Y. 2019 Effects of tillage practices on water consumption and grain yield of dryland winter wheat under different precipitation distribution in the loess plateau of China. *Soil and Tillage Research* **191**, 66–74.
- Zhang, J. B., Zhu, Y. P. & Chen, F. X. 2008 Forecast research of crop water requirements based on fuzzy rules. In: *Computer and Computing Technologies in Agriculture, Volume II* (D. L. Li, ed.), Springer, New York, USA, pp. 1267–1273.

First received 31 July 2019; accepted in revised form 2 February 2020. Available online 18 February 2020