Short-term prediction model for ammonia nitrogen in aquaculture pond water based on optimized LSSVM

Chen Yingyi[1,2,3*], Cheng Yanjun[1,2,3], Cheng Qianqian[1,2,3], Yu Huihui[1,2,3], Li Daoliang[1,2,3]

(1. *College of Information and Electrical Engineering, China Agricultural University, Beijing 100083, China*;
2. *Key Laboratory of Agricultural Information Acquisition Technology, Ministry of Agriculture, Beijing 100083, China*;
3. *Beijing Engineering and Technology Research Center for Internet of Things in Agriculture, Beijing, 100083, China*)

Abstract: Ammonia nitrogen is an important factor in aquaculture; it can poison aquatic nerves and cause economic loss. If the ammonia nitrogen content is too high, it can lead to serious losses in the fish population within a short period of time. However, ammonia nitrogen prediction is inaccurate due to the influence of many factors such as dissolved oxygen, water temperature, and pH. Thus, this paper presents an improved method for ammonia nitrogen prediction in aquaculture ponds. Principle component analysis is utilized to select the key factors from meteorological factors and water quality factors; wavelet threshold de-noising is used to process data; a least squares support vector regression (LSSVR) prediction model is established, and the key parameters are optimized by the adaptive mutation particle swarm optimization algorithm to obtain the optimal least squares support vector regression forecasting model.

Keywords: aquaculture, ammonia nitrogen prediction, least squares support vector regression

Citation: Chen, Y. Y., Y. J. Cheng, Q. Q. Cheng, H. H. Yu, and D. L. Li. 2017. Short-term prediction model for ammonia nitrogen in aquaculture pond water based on optimized LSSVM. International Agricultural Engineering Journal, 26(3): 416–427.

1 Introduction

Aquaculture ponds are a major part of freshwater aquaculture in China, accounting for 43.94% of freshwater aquaculture. Ammonia nitrogen, which is one of the key factors in aquaculture ponds, plays an important role during the process of aquatic growth. Excessive ammonia nitrogen even in a short time will affect aquatic development by poisoning its nerve center, even causing death. This seriously affects aquatic production, resulting in huge economic losses (Zhang and Zhu, 2012). As a result, accurate prediction of ammonia nitrogen in aquaculture ponds is extremely urgent.

The content of ammonia nitrogen in aquaculture ponds is influenced by other factors (e.g., water temperature, dissolved oxygen, and pH). These factors influence each other and are unstable. Thus, it is particularly important to filter out the key factors by optimizing the method. However, it is difficult to determine the representative variables that are independent of each other because of the limited knowledge regarding ammonia nitrogen in aquaculture ponds. Principal component analysis (PCA) is proposed for feature extraction and data dimension reduction (Pearson, 1901). It is effective to select the most representative key factors by using PCA because it considers the relationship among different factors (Singh et al., 2011, Combes and Azema, 2013). Thus, the key factors that affect the change in ammonia nitrogen were identified by adopting PCA in this paper.

The reality and accuracy of original data are interpreted as a condition necessary for further research (Singh et al., 2010). Nevertheless, the data arising from the monitoring stations and experiment might be polluted

Received date: 2017-06-26 **Accepted date:** 2017-08-17

* **Corresponding author: Chen Yingyi,** Ph.D., Associate Professor of College of Information and Electrical Engineering, China Agricultural University, Beijing 100083, China. Email: chenyingyi@cau.edu.cn. Tel: +86 10 62738489, Fax: +86 10 62737741.

by noise signals owing to systematic and random errors. This noisy data often made prediction relatively difficult (Najah et al., 2012). Therefore, it is necessary to remove such noise from the original data. Wavelet theory has been widely applied to signal processing since it was first proposed in the early 1980s by Grossman because of its ability of distinguish noise and useful signals (Grossmann and Morlet, 1984, Kmen and Aslan, 2013). Wavelet analysis is considered a useful tool to analyse detailed temporal patterns of water quality signals over different temporal scales (Liu et al., 2013). Therefore, we adopted wavelet analysis to de-noise and extract features of the original ammonia nitrogen data to improve forecast accuracy in this study.

There are many methods to forecast water quality, such as regression models, artificial neural network (ANN), and support vector machine (Maier et al., 2010, Singh et al., 2011, and Tan et al., 2012). The regression model is the most commonly used due to the rapidity and simplicity of prediction; however, it is difficult to guarantee the accuracy of prediction (Grbić et al., 2013, Wu et al., 2013). The ANN has been applied to many aspects but is not suitable for small sample data (Maier et al. 2010, Almonacid et al., 2013, Rouhani and Ravasan, 2013). The support vector machine is considered a better alternative to ANN due to its advantages in solving small samples, nonlinearity, high dimensions, local minimum points and other practical issues (Tan et al., 2012, Cortes and Vapnik, 1995, Dibike et al., 2001, and Chen and Li, 2014). The prediction model based on support vector machine was proved useful for predicting droughts and estimating uncertainty associated with drought predictions (Ganguli and Reddy, 2013). An unscented Kalman filter-based state-space vector regression approach was utilized to predict short-term wind speed and has been proved to have much better performance than ANN (Chen and Yu, 2014). An ASVM has proved to be a promising method for case adaptation in CBD systems (Qi et al., 2015). Thus, the support vector machine method was applied to the prediction of small sample data.

Least squares support vector regression (LSSVR) simplified the model standard SVR to a great extent by applying linear least squares criteria to the loss function instead of a traditional quadratic programming method, which greatly improves the calculation speed and accuracy (Suykens et al., 2002). LSSVR has been successfully applied to many prediction fields. Lin et al. attempted to use an LSSVR technique with monthly fuzzy weighted values to forecast revenue in uncertain economic conditions successfully (Lin et al., 2013). Utterance modelling with i-vectors, which was successfully applied to speaker recognition, has been used in conjunction with a WCCN and LSSVR to address speaker age estimation (Bahari et al., 2014). A PI-adaptive LSSVR controller was applied to a nonlinear inverted pendulum in the presence of disturbance (Naghash-Almasi and Khooban, 2016). Thus, the LSSVR was used to predict changes in ammonia nitrogen in this paper.

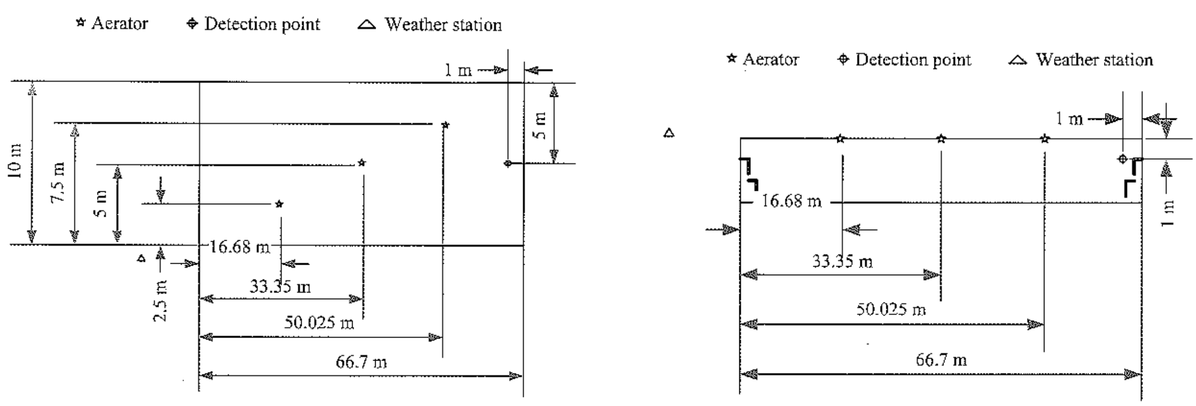
However, LSSVR performance heavily depends on the choice of kernel parameters and the regularization parameter, which are necessary to define the optimization problem and the final LSSVR model (Liu et al., 2013, Xie et al., 2013). Therefore, it is necessary to optimize these parameters. Particle Swarm Optimization (PSO) has been successfully applied to optimize the parameters of LSSVR. A PSO–SVM based on the association rules method was presented to diagnose erythemato-squamous diseases and was shown to be promising compared to the previously reported results (Abdi et al., 2013). Liu presented the dissolved oxygen prediction model based on LSSVR optimized by improved PSO and acquired a satisfactory forecasting result (Liu et al., 2013). Geng improved the forecasting precision of port throughput by applying the proposed simulated annealing particle swarm optimization (SAPSO) algorithm to the robust v-support vector regression model (RSVR) (Geng et al., 2015). The PSO was exploited in the SRITCSD method to serve as a multiclassifier for image texture features; meanwhile, the PSO, which can improve the performance of the SRITCSD method, was employed to optimize the LSSVR (Chang et al., 2016). However, PSO has many drawbacks, including premature convergence. The adaptive mutation particle swarm optimization algorithm (AMPSO) was proposed to solve this problem (Lu et al., 2005). Thus, the AMPSO was applied to optimize the parameters of the LSSVR in this study. The prediction model was tested and compared with other algorithms using the ammonia nitrogen data from a silver cod forming pond. The results show that the accuracy of prediction and the capability of generalization are greatly improved by our proposed approach.

This paper is organized as follows: Section 2 reports the construction of a hybrid forecasting model based on the wavelet analysis approach, least squares support vector regression and principal component analysis. Section 3 describes an application of the hybrid forecasting model. Finally, conclusions and future works are presented in Section 4.

2 Materials and methods

2.1 Data acquisition

The data used in this study were produced from the Tianxiang aquatic products Co., Ltd. in Ninghe County, Tianjin City, China. The experimental pond was approximately one acre, and the average water level was approximately 3 M. Three aerators were installed in the experimental pond. One point was chosen – where the distance from the forming pond was one meter and the depth was one meter – to collect the data (see Figure 1). A water-quality index, which was composed of ammonia nitrogen (AN), water temperature (WT), dissolved oxygen (DO) and pH, and a weather index, which was composed of rainfall (Ra), wind speed (WS), direction of wind (DW), solar radiation (SR), air temperature (AT), air humidity (AH) and atmospheric pressure (AP), were included in the data. The ammonia nitrogen was measured by a DZ-A type aquaculture water quality analyser every 4 hours starting from 0:00. Dissolved oxygen, water temperature and pH were detected by a HQ40d dual input multi-parameter digital analyser every 4 hours starting from 0:00. The weather index acquired from the small weather station was installed next to the silver cod forming pond.



a. The top view of the detection point

b. The sectional drawing of the detection point

Figure 1 Detection point of water quality data

2.2 Principal component analysis

Principal component analysis (PCA) was proposed for feature extraction and data dimension reduction based on multi-dimensional orthogonal linear transformation of statistical features (Pearson, 1901, Härdle and Simar, 2007). The core idea of PCA is to form several comprehensive indicators (principal component) from a linear combination of primitive variables by studying the original variation of the correlation matrix or covariance matrix of the internal structure of the relationship. The similarity between the variables should be taken into account in the first step. The correlation coefficient was applied to measure the similarity between variables; the formula was as follows Equation (1):

$$r_{jk}=\sum_{i=1}^{n}\frac{\sum_{i=1}^{n}(x_{ij}-\overline{x}_j)(x_{ik}-\overline{x_k})}{\left[\sum_{i=1}^{n}(x_{ij}-\overline{x}_j)^2\sum_{i=1}^{n}(x_{ik}-\overline{x_k})^2\right]^{\frac{1}{2}}}\quad (j=1,2,3,\ldots,m) \tag{1}$$

where, x_{ij} is the variable j in data I; x_{ik} is the variable k in data i; $\overline{x}_j$ is the average of variable j; $\overline{x_k}$ is the average of variable k; m is the total number of variables; n is the total number of each variable, and r_{jk} is the correlation coefficient between variable j and variable k.

Contribution rate and cumulative contribution rate are used to evaluate the principal component as follows Equations (2) and (3):

$$\tau_i = \frac{\lambda_i}{\sum_{k=1}^{m} \lambda_k} \quad (i = 1,2,..,m) \tag{2}$$

$$\eta_i = \frac{\sum_{k=1}^{i} \lambda_k}{\sum_{k=1}^{m} \lambda_k} \quad (i = 1,2,..,m) \tag{3}$$

where, λ_i and λ_k are the matrix eigenvalue; τ_i is the contribution rate of i th principal component, and η_i is the cumulative contribution rate of the first i principal component.

The load matrix of the principal component is used to screen key influencing factors as follows Equation (4):

$$\ell_{ij} = a_{ij}\sqrt{\lambda_i} \tag{4}$$

where, ℓ_{ij} is the correlation degree of principal components between the i th variable and the j th variable.

2.3 Wavelet analysis

Wavelet transform, which treats both the continuous and the discrete-time cases, has proven to be extremely valuable in signal processing (Daubechies, 1990, Rioul and Vetterli 1991). Unlike the Fourier transform, which can be utilized for a multi-scale analysis of a signal through dilation and translation, it can effectively extract the time-frequency features of a signal (Kisi and Cimen, 2011). For a continuous input signal, the time and scale parameters can be continuous, leading to the Continuous Wavelet Transform (Grossmann et al., 1989). Wavelet transform can be defined for discrete-time signals, leading to a Discrete Wavelet Transform (Daubechies, 1988, Rioul and Flandrin, 1992). The Discrete Wavelet Transform requires less computation time and is simpler to develop than the Continuous Wavelet Transform (Smith et al., 1998). Thus, the Discrete Wavelet Transform was applied to de-noise and extract features of the original ammonia nitrogen data.

For a discrete time series $f(t)$, the Discrete Wavelet Transform can be defined as the integration of a signal multiplied by a scaled and translated wavelet function at different time t, written as Equation (5):

$$W_\psi f(j,k) = \int f(t)\Psi^*_{j,k}(t)dt \tag{5}$$

The original signal $f(t)$ can be obtained by taking the inverse wavelet transform using the following Equation (6):

$$f(t) = \int\int W_\Psi f(j,k)\Psi_{j,k}(t)djdk \tag{6}$$

where, j, k are integer numbers; and $\Psi_{j,k}(t)$ is the wavelet basis function (Mallat 1989 Daubechies and Heil 1992), written as Equation (7):

$$\Psi_{j,k}(t) = 2^{-\frac{j}{2}}\Psi(2^{-j} - k) \tag{7}$$

2.4 Least squares Support Vector Regression

Giving a training set $\{x_i, y_i\}_{i=1}^{N}$, least squares support vector regression (Suykens et al. 2002) was defined as follows Equation (8):

$$\min_{w,b,e} J_p(w,e) = \frac{1}{2}w^T w + \gamma\frac{1}{2}\sum_{i=1}^{N} e_i^2 \tag{8}$$

where, w is the normal vector of the hyperplane; $\gamma \in R^+$ is the regularization parameter controlling model complexity and the overfitting phenomenon, $e=[e_1,e_2,\ldots,e_N]^T$, is the learning residual vector; $\varphi(\cdot)$ is the typical nonlinear mapping from the input space into the so-called feature space; and b is the bias.

Introduce the Lagrangian as Equation (9):

$$L(w,b,e;\alpha) = J_p(w,e) + \sum_{i}^{N}\alpha_i(y_i - w^T\varphi(x_i) - b - e_i) \tag{9}$$

where, α_i are the Lagrangian multipliers. The conditions for optimality are Equation (10):

$$\begin{cases} \frac{\partial L}{\partial w} = 0 \to w = \sum_{i=1}^{N}\alpha_i\varphi(x_i) \\ \frac{\partial L}{\partial b} = 0 \to \sum_{i=1}^{N}\alpha_i = 0 \\ \frac{\partial L}{\partial e_i} = 0 \to \alpha_i = \gamma e_i \\ \frac{\partial L}{\partial \alpha_i} = 0 \to y_i - w^T\varphi(x_i) - b - e_i = 0 \end{cases} \tag{10}$$

One obtains a set of linear equations after eliminating w and e in Equation (11)

$$\begin{bmatrix} 0 & 1^T \\ 1 & K \end{bmatrix}\begin{bmatrix} b \\ \alpha \end{bmatrix} = \begin{bmatrix} 0 \\ d \end{bmatrix} \tag{11}$$

where, $1=[1_1,1_2,\ldots,1_N]^{\mathrm{T}}$, $d=[d_1,d_2,\ldots,d_N]^T$, $K_{ij} = k(x_i, x_j) = \alpha(x_i)^T\alpha(x_j) + \frac{\delta_{ij}}{\gamma}$ is a Kernel function, $\delta_{ij} = \begin{cases} 1, i = j \\ 0, i \neq j \end{cases}$.

Hence, the regression model is found by solving Equation (9). The resulting least squares support vector regression then becomes Equation (12):

$$f(x)=\sum_{i=1}^{N}\alpha_i k(x_i,x)+B \quad (12)$$

2.5 AMPSO-based optimization of the LSSVR model

Particle swarm optimization algorithm (Eberhart et al., 2001, Trelea, 2003) which is carried out to find the optimal solution by evaluating the position, velocity and fitness of each particle – is a swarm intelligence optimization algorithm. The particle velocity and position update equations for the i th particle and d th dimension can be described as follows:

$$V_{id}^{k+1}=wV_{id}^{k}+c_1r_1(P_{id}^{k}-X_{id}^{k})+c_2r_2(P_{gd}^{k}-X_{id}^{k}) \quad (13)$$

$$X_{id}^{k+1}=X_{id}^{k}+\alpha V_{id}^{k+1} \quad (14)$$

where, w is the inertia weight; k is the iteration number; c_1 is the cognition learning factor; c_2 denotes the social learning factor; r_1 and r_2 are two independent uniformly distributed random variables with range [0,1], P_{id}^{k} denotes the best previous position encountered by the ith particle, and P_{gd}^{k} denotes the global best position of a particle; thus, X_{id} and V_{id} are the position and velocity of the particle, respectively.

An adaptive mutation particle swarm optimization algorithm is proposed in order to overcome the premature convergence of particle swarm optimization algorithm by increasing a random mutation operator (Lu et al., 2005). The particle swarm optimization algorithm can be continued in the case of premature convergence by initializing P_{gd}^{k} with a certain probability in Equations (15)-(17).

$$f_{avg}=\frac{1}{n}\sum_{i=1}^{n}f_i \quad (15)$$

$$f=\max\left\{1,\ \max_{1\le i\le n}|f_i-f_{avg}|\right\} \quad (16)$$

$$\sigma^2=\sum_{i=1}^{n}\left(\frac{f_i-f_{avg}}{f}\right)^2 \quad (17)$$

where, n is the particle swarm size; f_i is the fitness of particle i; f_{avg} is the average fitness of practices; σ^2 is the variance of population fitness, and f is the normalized scaling factor for limiting the size of σ^2.

P_{gd}^{k} is initialized by the following Equation (18) (Chen et al., 2006, Deyi et. al., 2011):

$$P_k=(P_{\max}-P_{\min})\left(\frac{\sigma_k^2}{n}\right)^2+(P_{\min}-P_{\max})\left(2\frac{\sigma_k^2}{n}\right)+P_{\max} \quad (18)$$

where, $P_{\max}$ is the maximum variation probability; $P_{\min}$ is the minimum variation probability; P_k is the group global optimization in k iteration, and σ_k^2 is the variance of population fitness in k iteration. Increasing random perturbation was adopted to mutate P_{gd}^{k} as follows Equation (19):

$$P_{gd}^{k}=P_{gd}^{k}(1+0.5\eta) \quad (19)$$

where, η is a random variable subject to Guass (0, 1).

The process of optimizing the LSSVR parameters with AMPSO is presented in Figure 2, which can be described as follows:

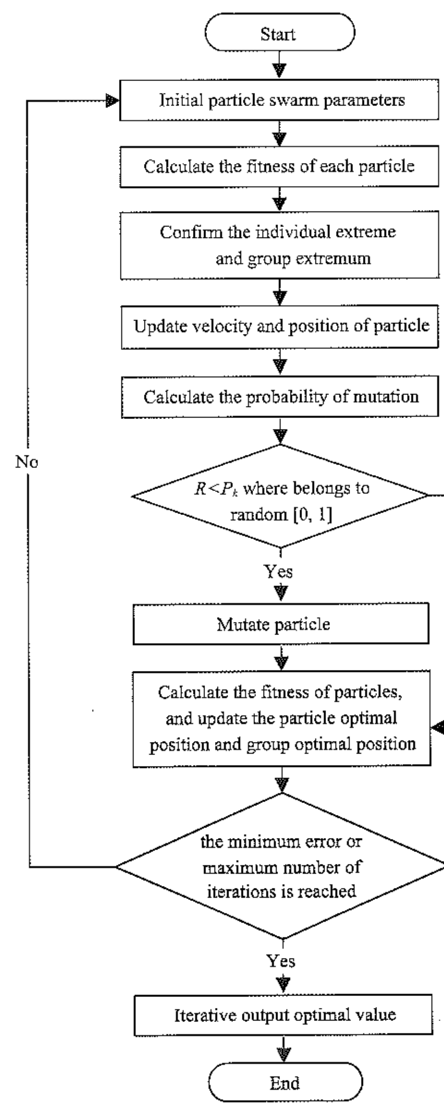


Figure 2 Process of optimizing the LSSVR parameters with AMPSO

Step1: Initial particle swarm parameters: inertia factor, acceleration constant, maximum number of iterations and the minimum allowable error of the algorithm. And initiate the velocity and position of the particles randomly.

Step2: Calculate the fitness of each particle according to the following formula Equation (20):

$$\text{itness} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(\hat{y}_i - y_i)^2} \tag{20}$$

where, n is the particle swarm size; y_i is the real value, and $\hat{y}_i$ is the forecast value.

Step3: Confirm the individual extreme and group extremum according to the fitness of each particle.

Step4: Update the velocity and position of the particle by using Equation (12), (13).

Step5: Calculate the probability of mutation by using Equation (14)-(17).

Step6: Mutate the particle by using Equation (18) while $r<P_k$, where r belongs to random [0, 1].

Step7: Calculate the fitness of particles, and update the particle optimal position and group optimal position.

Step8: Iteratively output the optimal value until the minimum error or maximum number of iterations is reached. Otherwise, go to step 4.

2.6 The overall structure of the proposed hybrid algorithm

The process for content prediction of ammonia nitrogen in aquaculture using optimized least square support vector regression is described as follows:

Step1: Collect the data used in this study, and then dispose error data and missing data.

Step2: Filter the key factors for the content prediction of ammonia nitrogen in aquaculture using principal component analysis.

Step3: De-noise and extract features of the key factors data and ammonia nitrogen using wavelet analysis.

Step4: Normalize the data and select the training sample and testing sample.

Step5: Build the model using the AMPSO-LSSVR algorithm.

Step6: Output the results, and evaluate the model.

3 Results and discussions

3.1 Principal component analysis for factors dimension reduction

The data which includes ammonia nitrogen (AN), water temperature (WT), dissolved oxygen (DO), pH, rainfall (Ra), wind speed (WS), direction of wind (DW), solar radiation (SR), air temperature (AT), air humidity (AH) and atmospheric pressure (AP) were acquired once every hour starting from 0:00 during the time period between October 2th and October 10th, 2015. Rainfall was deleted from the data set because there is no rain from October 2th to October 10th. The PCA was chosen to screen the key factors influencing the change of ammonia nitrogen according to the following steps:

Step1: Normalize the original data by the following Equations (21) and (22):

$$x_{ij}^* = \begin{cases} \dfrac{x_{ij} - \overline{x_j}}{S_j}, & S_j \neq 0 \\ 0, & S_j = 0 \end{cases} \quad \begin{pmatrix} i = 1,2,\dots,n \\ j = 1,2,\dots,m \end{pmatrix} \tag{21}$$

$$S_j = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_{ij} - \overline{x_j})^2} \tag{22}$$

where, x_{ij} is the variable j in data i; $\overline{x_j}$ is the average of variable j; m is the total number of variables; n is the total number of each variable; S_j is the standard deviation of variable j, and x_{ij}^* is the standard data.

Step2: Find the similarity between every pair of variables in the data set using the formula (1), and then list the correlation coefficient matrix.

Step3: Calculate the characteristic value and characteristic vector of the correlation coefficient matrix; the result is shown in Table 1.

Step4: Calculate the contribution rate and cumulative contribution rate using formula (2) and (3); the result is shown in Table 1.

Table 1 Factor the characteristic value and contribution rate of aquaculture

Principal fact	Characteristic value	Contribution rate, %	Cumulative contribution rate, %
1	2.9019	29.02	29.02
2	2.4923	24.92	53.94
3	1.4244	14.24	68.18
4	1.1512	11.51	79.69
5	0.8762	8.76	88.45
6	0.4517	4.52	92.97
7	0.3002	3.00	95.97
8	0.2195	2.20	98.17
9	0.1352	1.35	99.52
10	0.0475	0.48	10.00

5 factors were chosen as the key factors influencing the change in ammonia nitrogen because the cumulative contribution rate of the first five factors is greater than 85%, which we can see from Table 1.

Step5: Calculate the load matrix of the principal fact using Equation (4); the result is listed in Table 2.

Table 2 The load matrix of the principal fact

Fact	Principal fact 1	Principal fact 2	Principal fact 3	Principal fact 4	Principal fact 5
AN	−0.8414	−0.2664	−0.1803	0.1522	−0.0085
WT	0.7663	0.3250	−0.4508	−0.1490	0.1281
DO	−0.2576	0.6340	0.6548	−0.0796	0.0285
pH	0.1168	0.7607	0.5137	−0.1765	0.0112
WS	−0.7012	0.4886	−0.2305	−0.1454	0.0698
DW	0.1236	0.0250	0.1526	0.6828	0.7005
SR	−0.0334	0.5090	−0.2765	0.5948	−0.4280
AT	0.6568	0.6176	−0.3589	−0.0611	0.0795
AH	0.5473	−0.6535	0.2634	−0.2054	0.0354
AP	0.5364	−0.1682	0.3807	0.4258	−0.4153

From Table 2, we can see that the correlation coefficient of AN and WT on the first factor is larger than that of other factors; the contribution of pH to the second factor is the largest, the contribution of DO to the third factor is the largest; the contribution of DW to the fourth and fifth factors is the largest. Thus, the AN, WT, pH, DO and DW were chosen as the principal component index. This also corresponds to the selection of the key factors influencing the change in ammonia nitrogen.

3.2 Wavelet analysis for data de-noising

The Discrete Wavelet Transform was applied to de-noise and extract features of the key factor data selected in section 3.1 and ammonia nitrogen as according to the following steps:

Step1: Decompose the original signal into 3 scales with the appropriate wavelet basis function illustrated in Figure 3.

Step2: Process the wavelet coefficients with the appropriate wavelet thresholding function.

Step3: Obtain the de-noising signal by reconstructing the processed wavelet coefficients with soft thresholding.

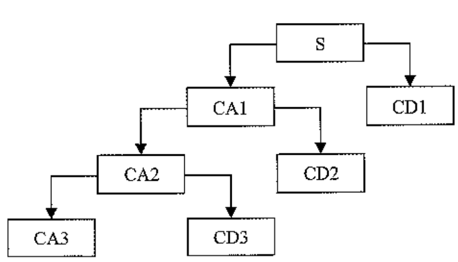


Figure 3 Tree structure of wavelet decomposition

It is difficult to find an ideal wavelet in signal analysis, so a compromise is made between performance and complexity. It is desirable to use few wavelet basis functions and wavelet thresholding functions from different families for the performance evaluation and selecting the wavelet that gives the best performance. The root mean square error (RMSE) and the signal-to-noise ratio (SNR) were used to measure performance. The RMSE and SNR can be illustrated as follows Equations (23) and (24):

$$\mathrm{RMSE}=\sqrt{\frac{1}{n}\sum[f(n)-\hat{f}(n)]^2} \tag{23}$$

$$\mathrm{SNR}=10\log_{10}\frac{\sum_n f^2(n)}{\sum_n[f(n)-\hat{f}(n)]^2} \tag{24}$$

where, $f(n)$ is the original signal, and $\hat{f}(n)$ is the signal after wavelet threshold de-noising.

The wavelet basis functions haar, dmey, dbN (N=2, 3, …, 10), symN (N=1, 2, … ,10), coifN (N=1, 2, …, 5) and wavelet thresholding function rigrsure were applied to process the ammonia nitrogen; the result is shown in Table 3.

Table 3 The RMSE and SNR of ammonia nitrogen with different wavelet basis functions

Wavelet basis function type	SNR	RMSE
Haar	23.3144	0.0586
dmey	23.0417	0.0605
db2	23.5433	0.0571
db3	23.8969	0.0548
db4	23.8567	0.0551
db5	24.1539	0.0532
db6	24.2507	0.0526
db7	22.9939	0.0608
db8	24.1971	0.053
db9	23.3543	0.0584
db10	23.984	0.0543
sym1	23.3144	0.0586
sym2	23.5433	0.0571
sym3	23.8969	0.0548
sym4	23.7257	0.0559
sym5	23.0542	0.0604
sym6	24.6398	0.0503
sym7	23.7485	0.0558
sym8	23.8564	0.0551
sym9	23.8572	0.0551
sym10	23.5851	0.0568
coif1	23.4919	0.0574
coif2	23.7919	0.0555
coif3	23.1602	0.0597
coif4	23.8815	0.0549
coif5	23.0811	0.0602

The wavelet basis function sym6 was selected as the best wavelet basis function due to the minimum RMSE (0.0503) and the maximum SNR (24.6398). The result of ammonia nitrogen noise reduction is presented in Figure 4. From Figure 4, it can be seen that the change curve of ammonia nitrogen becomes smooth, and the influence of noise and clutter on the ammonia nitrogen was eliminated.

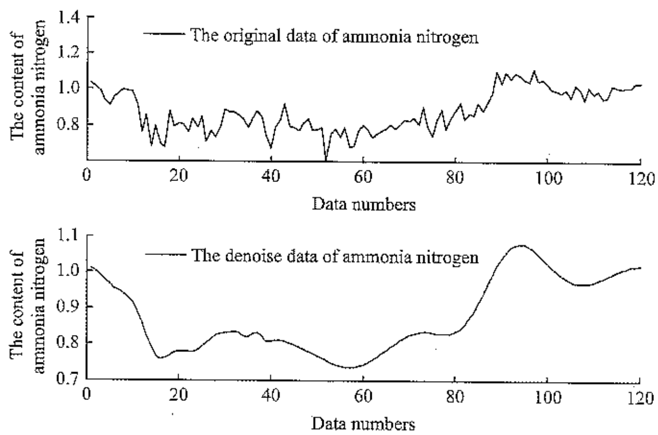


Figure 4 Result of ammonia nitrogen

The wavelet basis functions of water temperature (WT), dissolved oxygen (DO), pH, and direction of wind (DW) were chosen for the same method. Finally, the best wavelet basis function and wavelet thresholding function of each factor were selected as shown in Table 4.

Table 4 Best wavelet basis function of each factor

Factor	Wavelet basis function type	SNR	RMSE
AN	sym6	24.6398	0.0503
WT	coifl	29.9880	0.5977
DO	sym5	27.3846	0.4019
pH	sym3	38.3841	0.1017
DW	db6	47.4163	11.0376

3.3 Results evaluation

The proposed hybrid algorithm was implemented in the Matlab R2012a programming language. We obtained the optimal parameters γ and σ for the prediction model of ammonia nitrogen based on the AMPSO algorithm. The initial parameters of AMPSO were given as follows: the cognition learning factor c_1=1.5, the social learning factor c_2=1.7, the population size of swarm sizepop=30, the iteration number maxgen=300, the inertia weight wmax=1.2 and wmin=0.8; the fitness accuracy of the normalized samples are equal to 0.002. The changing trend of the fitness value is illuminated in Figure 5. From Figure 5, we can see that the fitness value tends to stabilize quickly; the AMPSO converges to the best solution quickly and is more appropriate for seeking the unknown parameters of the LSSVR. The optimal combination parameters were obtained, namely, γ=1000, σ=2.9144.

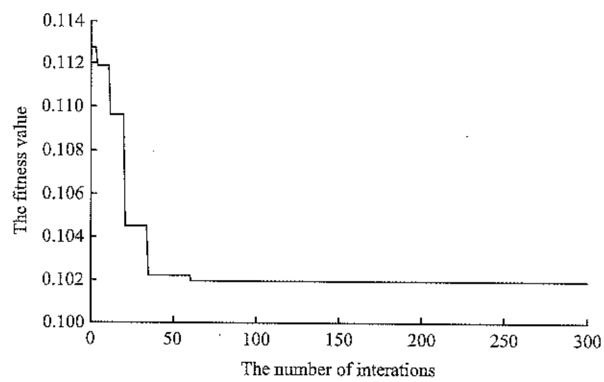


Figure 5 Trend of the fitness value

The optimal combination parameters were adopted to train the ammonia nitrogen prediction model. The training result of the prediction model was shown in Figure 6. From Figure 6, the real values and the predicted training values with AMPSO_LSSVR were not quite different. The test sample set was put into the trained model to predict the change in ammonia nitrogen. The ammonia nitrogen content prediction result is given in Figure 7. From Figure 7, it is shown that the ammonia nitrogen content prediction result is broadly in line with the real values.

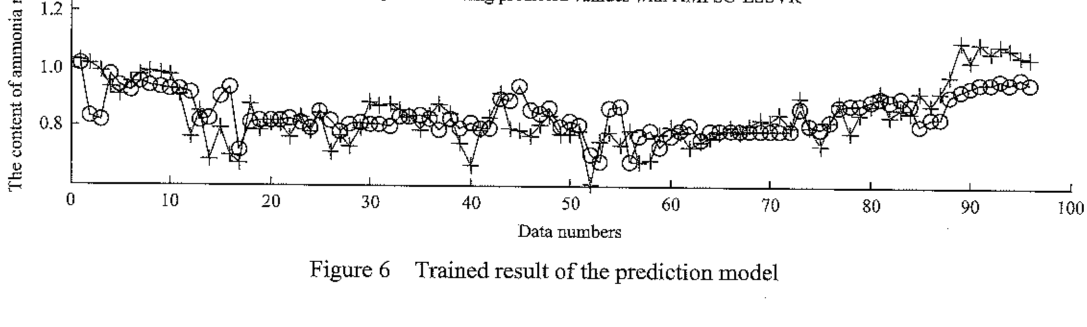


Figure 6 Trained result of the prediction model

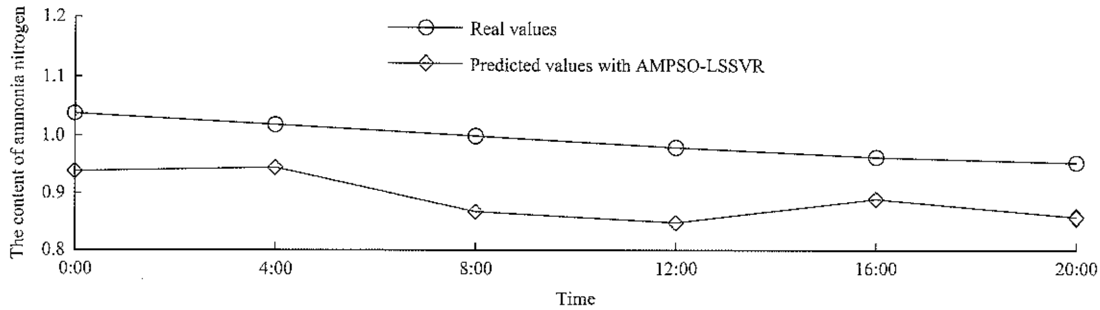


Figure 7 Ammonia nitrogen content prediction result

In addition, the standard least square support vector regression (LSSVR) was compared with the optimized least square support vector regression. The results of different methods are illustrated in Figure 8. The results of the optimized least square support vector regression are closer to the real values. Thus, the accuracy of ammonia nitrogen prediction in the pond aquaculture was improved.

The absolute errors of the standard least square support vector regression and the optimized least square support vector regression are represented in Figure 9. The absolute errors of the optimized least square support are smaller than those of the standard least square support vector regression.

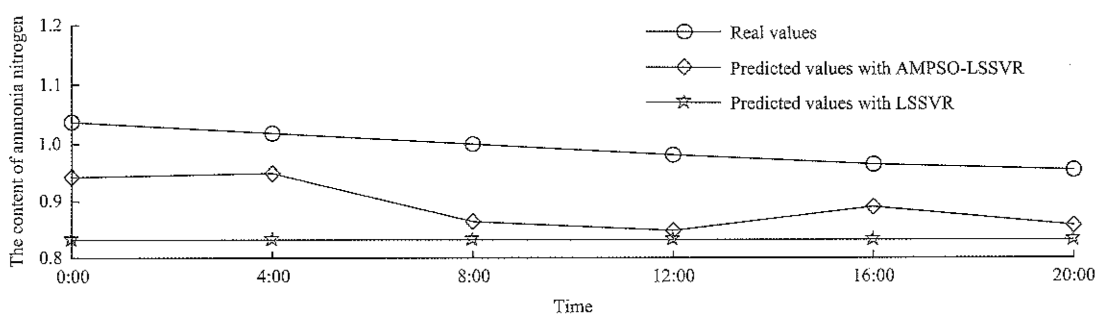


Figure 8 Predicted results of different methods

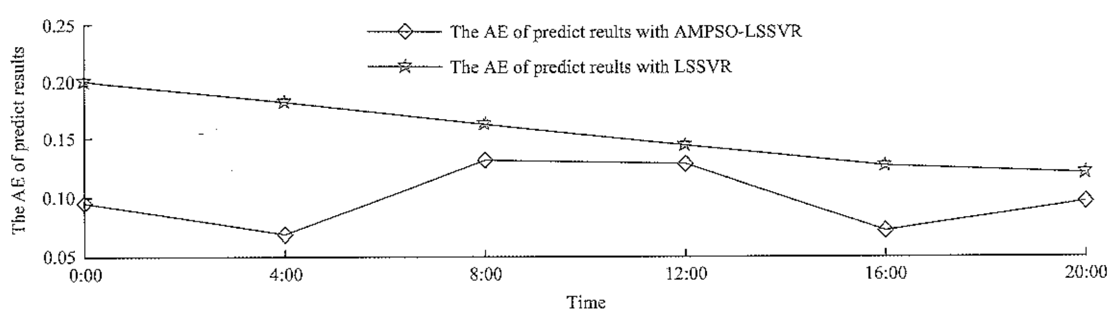


Figure 9 AE of different predicted results with different methods

Different standard statistical performance evaluation criteria, such as root mean square error (RMSE), the mean absolute percentage error (MAPE), the mean absolute error (MAE), the mean relative error (MRE), and the Nash–Sutcliffe efficiency coefficient (NSC), were used to evaluate the performance of various models. The RMSE, MAE and MRE were used to assess the prediction capability of the model proposed in this paper. The MAE and MRE can be illustrated as follows Equations (25) and (26):

$$MAE = \frac{1}{N}\sum_{i=1}^{N} | y_i - \hat{y}_i | \tag{25}$$

$$MRE = \frac{1}{N}\sum_{i=1}^{N} \left| \frac{y_i - \hat{y}_i}{y_i} \right| \tag{26}$$

where, y_i is the real value; $\hat{y}_i$ is the predicted value, and N is the number of samples.

The RMSE, MAE and MRE of different methods are listed in Table 5. The obtained results indicate that the optimized least square support vector regression is far superior to the standard least square support vector regression. The prediction RMSE of optimized least square support vector regression is reduced by 40.15% compared to that of standard least square support vector regression. The prediction MAE of optimized least square support vector regression is reduced by 41.82% compared to that of standard least square support vector regression. The prediction MRE of optimized least square support vector regression is reduced by 42.29% compared to that of standard least square support vector regression.

Table 5 RMSR and MAE of different methods

Method	RMSE	MAE	MRE
Optimized least square support vector regression	0.0945	0.0903	0.0887
Standard least square support vector regression	0.1579	0.1552	0.1537

The hybrid model proposed in this study is capable of searching for the parameter values of the LSSVR and RBF kernel function. The result of this study demonstrates that the prediction of ammonia nitrogen content is effective and feasible.

4 Conclusions

This paper proposed a hybrid forecasting model that combined principle component analysis, wavelet analysis, and least squares support vector regression and adaptive mutation particle swarm optimization algorithm. The results clearly show that compared with the standard least square support vector regression (LSSVR), the proposed hybrid method of optimized LSSVR has better prediction performance, as measured by RMSE. Further, the optimized LSSVR can effectively consider many dimensions and nonlinearity, non-stationary and finite samples and is a reliable forecasting tool for predicting ammonia nitrogen time series in modern intensive aquaculture.

There is room for further study and development. First, the change in ammonia nitrogen in the aquaculture pond will be different during different farming seasons and different growth periods. In future work, we plan to research the prediction of ammonia nitrogen in aquaculture ponds over longer time periods, taking into consideration different farming seasons and different growth periods to control the ammonia nitrogen in aquaculture ponds. Second, other optimization search algorithms, like genetic algorithm (GA) and ant colony optimization (ACO), can be compared with the adaptive mutation particle swarm optimization algorithm (AMPSO). These are all valuable problems for future research.

Acknowledgements

This paper was supported by the Beijing science and technology plan- Research on intelligent system and equipment for Large-scale freshwater fish breeding (No. Z171100001517016), the EU cooperation projects-Innovative model & demonstration-based water management for resource efficiency in integrated multitrophic aquaculture and horticulture systems (No. 619137), the International Science & Technology Cooperation Program of China - Key technology cooperating research on agricultural internet of things advanced sensors and intelligent processing (No.2013DFA11320), and the Shandong Province Key Research & Development Program - Research on aquaculture management and application platform technology based on big data (No. 2015GGC02066).

[References]

[1] Abdi, M. J., and D. Giveki. 2013. Automatic detection of erythemato-squamous diseases using PSO–SVM based on association rules. *Engineering Applications of Artificial Intelligence*, 26(26): 603–608.

[2] Almonacid, F., E. F. Fernández, P. Rodrigo, P. J. Pérez-Higueras, and C. Rus-Casas. 2013. Estimating the maximum power of a high concentrator photovoltaic (HCPV) module using an artificial neural network. *Energy*, 53(5): 165–172.

[3] Bahari, M. H., M. Mclaren, H. V. Hamme, and D. A. V. Leeuwen. 2014. Speaker age estimation using i-vectors. *Engineering Applications of Artificial Intelligence*, 34(3): 99–108.

[4] Chang, B., H. Tsai, and C. Yen. 2016. SVM-PSO based rotation-invariant image texture classification in SVD and DWT domains. *Engineering Applications of Artificial Intelligence*, 52(C): 96–107.

[5] Chen, G., J. Jia, and H. Qi. 2006. Study on the strategy of decreasing inertia weight in particle swarm optimization algorithm. *Journal of Xian Jiaotong University*, 40(1): 53–56.

[6] Chen, J. L., and G. S. Li. 2014. Evaluation of support vector machine for estimation of solar radiation from measured meteorological variables. *Theoretical & Applied Climatology*,

115(3-4): 627–638.

[7] Chen, K., and J. Yu. 2014. Short-term wind speed prediction using an unscented Kalman filter based state-space support vector regression approach. *Applied Energy,* 113(6): 690–705.

[8] Combes, C., and J. Azema. 2013. Clustering using principal component analysis applied to autonomy-disability of elderly people. *Decision Support Systems*, 55(2): 578–586.

[9] Cortes, C., and V. Vapnik. 1995. Support-vector networks. *Machine learning*, 20(3): 273–297.

[10] Daubechies, I. 1990. The wavelet transform, time-frequency localisation and signal analysis. *IEEE transactions on information theory*, 36(5): 961–1005.

[11] Daubechies, I., and C. Heil. 1993. Ten lectures on wavelets. *The Journal of the Acoustical Society of America*, 93(3): 1671–1671.

[12] Daubechies, I. 1988. Orthonormal bases of compactly supported wavelets. *Communications on Pure and Applied Mathematics*, 41(7): 909–996.

[13] Deyi ,Y. E., H. E. Zhengyou, and T. Zang. 2011. Siting and sizing of distributed generation planning based on adaptive mutation particle swarm optimization algorithm. *Power System Technology*, 35(6): 155–160.

[14] Dibike, Y. B., D. Solomatine, S. Velickov, and M. B. Abbott. 2001. Model induction with support vector machines: introduction and applications. *Journal of Computing in Civil Engineering*, 15(3): 208–216.

[15] Dökmen, F., and Z. Aslan. 2013. Evaluation of the parameters of water quality with wavelet techniques. *Water resources management*, 27(14): 4977–4988.

[16] Eberhart, R. C., Y. Shi, R. C. Eberhart, and Y. Shi. 2001. Particle swarm optimization: developments, applications, and resources. *In IEEE Congress on Evolutionary Computation (CEC 2001)*, 259–263. Seoul, South Korea, 27-30 May.

[17] Ganguli, P., and M. J. Reddy. 2013. Ensemble prediction of regional droughts using climate inputs and the SVM–copula approach. *Hydrological processes*, 28(19): 4989–5009.

[18] Geng, J., M. Li, Z. Dong, and Y. Liao. 2015. Port throughput forecasting by MARS-RSVR with chaotic simulated annealing particle swarm optimization algorithm. *Neurocomputing*, 147(1): 239–250.

[19] Grbić, R., D. Kurtagić, and D. Slišković. 2013. Stream water temperature prediction based on Gaussian process regression. *Expert systems with applications*, 40(18): 7407–7414.

[20] Grossmann, A., and J. Morlet. 1984. Decomposition of Hardy functions into square integrable wavelets of constant shape. *SIAM journal on mathematical analysis*, 15(4): 723–736.

[21] Grossmann, A., R. Kronland-Martinet, and J. Morlet. 1989. Reading and Understanding Continuous Wavelet Transforms. Germany: Springer Berlin Heidelberg.

[22] Härdle, W., and L. Simar. 2012. Applied multivariate statistical analysis. *Technometrics*, 31(2): 265–266.

[23] Kisi, O., and M. Cimen. 2011. A wavelet-support vector machine conjunction model for monthly streamflow forecasting. *Journal of Hydrology*, 399(1): 132–140.

[24] Lin, K., P. Pai, Y. Lu, and P. Chang. 2013. Revenue forecasting using a least-squares support vector regression model in a fuzzy environment. *Information Sciences*, 220(1): 196–209.

[25] Liu, S., L. Xu. Y. Jiang, D. Li, Y. Chen, and Z. Li. 2013. A hybrid WA-CPSO-LSSVR model for dissolved oxygen content prediction in crab culture. *Engineering Applications of Artificial Intelligence*, 29(3): 114–124.

[26] Lu, Z. S., Z. R. Hou, and J. Du. 2005. Particle swarm optimization with adaptive mutation. *Frontiers Engineering in China*, 1(1): 99–104.

[27] Maier, H. R., A. Jain, G. C. Dandy, and K. P. Sudheer. 2010. Methods used for the development of neural networks for the prediction of water resource variables in river systems: current status and future directions. *Environmental Modelling & Software*, 25(8): 891–909.

[28] Mallat, S. G. 1989. A theory for multiresolution signal decomposition: the wavelet representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11(7): 674–693.

[29] Naghash-Almasi, O., and M. H. Khooban. 2016. PI adaptive LS-SVR control scheme with disturbance rejection for a class of uncertain nonlinear systems. *Engineering Applications of Artificial Intelligence*, 52(C): 135–144.

[30] Najah, A. A., A. El-Shafie, O. A. Karim, and O. Jaafar. 2012. Water quality prediction model utilizing integrated wavelet-ANFIS model with cross-validation. *Neural Computing and Applications*, 21(5): 833–841.

[31] Pearson, K. 1901. On lines and planes of closest fit to systems of points in space. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 2(11): 559–572.

[32] Qi, J., J. Hu, and Y. Peng. 2015. Incorporating adaptability-related knowledge into support vector machine for case-based design adaptation. *Engineering Applications of Artificial Intelligence*, 37(9): 170–180.

[33] Rioul, O., and M. Vetterli. 1991. Wavelets and signal processing. *IEEE signal processing magazine*, 8(4): 14–38.

[34] Rioul, O., and P. Flandrin. 1992. Time-scale energy distributions: a general class extending wavelet transforms. *IEEE Transactions on Signal Processing*, 40(7): 1746–1757.

[35] Rouhani, S., and A. Z. Ravasan. 2013. ERP success prediction: An artificial neural network approach. *Scientia Iranica*, 20(3): 992–1001.

[36] Singh, C. B., R. Choudhary, D. S. Jayas, and J. Paliwal. 2010.

Wavelet analysis of signals in agriculture and food quality inspection. *Food and Bioprocess Technology*, 3(1): 2–12.

[37] Singh, K. P., N. Basant, and S. Gupta. 2011. Support vector machines in water quality management. *Analytica Chimica Acta*, 703(2): 152–162.

[38] Singh, V. H., M. Agrawal, G. C. Joshi, M. Sudershan, and A. K. Sinha. 2011. Elemental profile of agricultural soil by the EDXRF technique and use of the Principal Component Analysis (PCA) method to interpret the complex data. *Applied Radiation and Isotopes*, 69(7): 969–974.

[39] Smith, L. C., D. L. Turcotte, and B. L. Isacks. 1998. Stream flow characterization and feature detection using a discrete wavelet transform. *Hydrological Processes*, 12(2): 233–249.

[40] Suykens, J., T. Van Gestel, J. De Brabanter, B. De Moor, and J. Vandewalle. 2002. *Least Squares Support Vector Machines*. Singapore: World Scientific Publishing.

[41] Tan G., J. Yan, C. Gao, and S. Yang. 2012. Prediction of water quality time series data based on least squares support vector machine. *Procedia Engineering*, 31(16): 1194–1199.

[42] The Ministry of Agriculture, F. F. A., 2016. China fishery statistics yearbook 2016. China Agriculture Press: China.

[43] Trelea, I. C. 2003. The particle swarm optimization algorithm: convergence analysis and parameter selection. *Information Processing Letters*, 85(6): 317–325.

[44] Wu, J., J. Wang, H. Lu, Y. Dong, and X. Lu. 2013. Short term load forecasting technique based on the seasonal exponential adjustment method and the regression model. *Energy Conversion and Management*, 70(70): 1–9.

[45] Xie, G., S. Wang, Y. Zhao, and K. K. Lai. 2013. Hybrid approaches based on LSSVR model for container throughput forecasting: a comparative study. *Applied Soft Computing*, 13(5): 2232–2241.

[46] Zhang, W., and Y. Zhu. 2012. Advances on the research of the hazard of ammonia nitrogen in aquaculture water and its determination method. *Journal of Environmental Hygiene*, 2(6): 324–327.