Bacterial Foraging Optimization Combined with Relevance Vector Machine with an Improved Kernel for Pressure Fluctuation of Hydroelectric Units

Liying Wang[1, 2]
1. School of Mechanical, Electronic and Control Engineering, Beijing Jiaotong University, Beijing, 100044, China
2. School of Waterpower, Hebei University of Engineering, Handan 056038, China
Email: 2000wangly@163.com

Shaopu Yang
College of Mechanical Engineering, Shijiazhuang Tiedao University, Shijiazhuang 050043, China;

***Abstract*—The optimization of kernel parameters is an important step in the application of the Relevance Vector Machine (RVM) for many real-world problems. In this paper, firstly we have developed an improved anisotropic Gaussian kernel as the kernel function of the RVM model, whose parameters are optimized by Bacterial Foraging Optimization (BFO). Then the proposed method is applied to describing the pressure fluctuation characteristics of the draft tube of hydroelectric units of a hydropower station, through the comparison, the simulation results show the parameters of the improved anisotropic Gaussian kernel are well optimized using the BFO, and the acquired RVM model can precisely describe the pressure fluctuation characteristics of the draft tube, and the less training samples are required to establish the accurate RVM model implying that it is more sparse than its counterpart.**

***Index Terms*—Bacterial Foraging Optimization, Relevance Vector Machine, Hydroelectric units, Pressure Fluctuation**

I. INTRODUCTION

It is well known that pressure fluctuation in draft tube caused by vortex rope has a great influence on the stability of hydroelectric units. Pressure fluctuation of draft tube] is the main factors which effect the stable operation of turbine. In the operation process, all sorts of dynamic testing instruments are used to loot, record and analyze the pressure fluctuation signals, which can control the stable operation of turbine. In recent years, scientific and technological personals have discussed the pressure fluctuation of draft tube. Grasping the laws of pressure fluctuation has very important practical significance to contribute to the secure and stable running of hydro-generator units under different operating conditions [1-3]. Because the strong nonlinear characteristics of pressure fluctuation makes its expression and analysis difficult, so, how to find a more effective method to express the characteristics is becoming a burning question now [4, 5]. Recently there are many nonlinear approaches proposed are applied to the pressure fluctuation of the turbine, such as Artificial Neural Networks (ANN) [6-8], Support Vector Machine (SVM) [9] and so on. But they have many drawbacks, for example, the ANN can trap into local minimum and has inherent searching rate slowly when training[10] ,SVM is wasteful both of data and computation to determine the relative parameters through carrying on a cross-validation procedure, and it does not allow for the free use of an arbitrary kernel function [11-15].

RVM[16] is a probabilistic sparse kernel model identical in functional form to SVM, where a Bayesian approach to learning is adopted introducing a prior over the weights governed by a set of hyperparameters, its main advantages include its capability to obtain a generalization performance comparable to SVM but using dramatically fewer training samples, Furthermore, and suffering from none of the other limitations of SVM outlined above. Application of group foraging strategy of a swarm of E.coli bacteria in multi-optimal function optimization is the key idea of the new algorithm [17-19]. In this paper, an improved anisotropic Gaussian kernel as the kernel function of the RVM model is developed, whose parameters are optimized by the BFO. Then the proposed method is applied to describing the pressure fluctuation characteristics of the draft tube of hydroelectric units, the accuracy of results and sparseness of the acquired RVM model in the comparative simulation illustrate the superiority of the proposed method.

II. BACTERIAL FORAGING OPTIMIZATION

The bacterial foraging system consists of four principal mechanisms, namely swarming and tumbling, chemotaxis, reproduction, elimination-dispersal Below we briefly describe each of these processes [20-23]..

[1] Manuscript received Apr. 1, 2012; revised May 10, 2012; accepted May 18, 2012.

[2] Project number: 2012YJS103, E2010001026, 51075118.

doi:10.4304/jcp.8.5.1273-1278

Swarming and Tumbling. The flagellum is a left-handed helix configured so that as the base of the flagellum rotate counter clockwise from the free end of the flagellum looking toward the cell; it produces a force against the bacterium pushing the cell. This mode of motion is called swimming. Bacteria swim either for maximum number of steps N_s or less depending on the nutrition concentration and environment condition.

Chemotaxis. A chemotaxis step is the movement of an E.coli cell through swimming and tumbling via flagella. Biologically, an E.coli bacterium can move in two different ways. It can swim for a period in the same direction, or it may tumble, and alternate between these two modes of operation for the entire lifetime. Suppose $\theta^i(j,k,l)$ represents ith bacterium at jth chemotactic, kth reproductive and lth elimination dispersal step. $C(i)$ is the size of the step taken in the random direction specified by the tumble (run length unit). Then in computational chemotaxis the movement of the bacterium may be represented by

$$\theta^i(j+1,k,l)=\theta^i(j,k,l)+C(i)\frac{\Delta(i)}{\sqrt{\Delta^T(i)\Delta(i)}} \quad (1)$$

where $\triangle$ indicates a vector in the random direction whose elements lie in [–1, 1].

Reproduction. After N_c chemotactic steps, a reproduction step is taken. Let N_{re} be the number of reproduction steps to be taken. For convenience, we assume that S is a positive even integer. Let S_r=S/2 be the number of population members who have had sufficient nutrients so that they will reproduce (split in two) with no mutations. For reproduction, the population is sorted in order of ascending accumulated cost. The S_r least healthy bacteria die and the other S_r healthiest bacteria each split into two bacteria, which are placed at the same location.

Elimination-dispersal. Elimination event may occur for example when local significant increases in heat kills a population of bacteria that are currently in a region with a high concentration of nutrients. A sudden flow of water can disperse bacteria from one place to another. The effect of elimination and dispersal events is possibly destroying chemotactic progress, but they also have the effect of assisting in chemotaxis, since dispersal may place bacteria near good food sources.

III. Relevance Vector Machine in regression

Tipping [15, 16] proposed the Relevance Vector Machine in 2000. For a regression problem, given a training dataset $\{x_n,t_n\}_{n=1}^N$,

$$t_n=y(x_n,\mathbf{w})+\varepsilon_n \quad \mathbf{t}=\mathbf{y}+\boldsymbol{\varepsilon} \quad (2)$$

Where the errors $\boldsymbol{\varepsilon}=(\varepsilon_1,.....,\varepsilon_n)$ are modeled probabilistically as independent zero-mean Gaussian, with variance σ^2, so $p(\varepsilon)=\prod_{n=1}^N N(\varepsilon_n|0,\sigma^2)$, $\mathbf{w}=(w_1.....,w_M)$ is the parameter vector and $y(x_n,\mathbf{w})$ can be expressed as a linearly weighted sum of some basis functions $\phi(\mathbf{x})$:

$$y(\mathbf{x},\mathbf{w})=\sum_{m=1}^M w_m\phi_m(\mathbf{x})+w_0 \quad \mathbf{y}=\Phi\mathbf{w} \quad (3)$$

The likelihood of the complete dataset can be written as

$$p(\mathbf{t}|\mathbf{w},\sigma^2)=(2\pi\sigma^2)^{-N/2}\exp\left\{-\frac{1}{2\sigma^2}\|\mathbf{t}-\Phi\mathbf{w}\|^2\right\} \quad (4)$$

RVM adopts a Bayesian perspective, and constrains the parameter by defining an explicit prior probability distribution over them, the posterior distribution over the weights is thus given by:

$$p(\mathbf{w}|\boldsymbol{\alpha})=(2\pi)^{-M/2}\prod_{m=1}^M\alpha_m^{1/2}\exp(-\alpha_m w_m^2/2) \quad (5)$$

Given $\boldsymbol{\alpha}$, the posterior parameter distribution conditioned on the data is given by combining the likelihood and prior within Bayes's rules:

$$p(\mathbf{w}|\mathbf{t},\boldsymbol{\alpha},\sigma^2)=p(\mathbf{t}|\mathbf{w},\sigma^2)p(\mathbf{w}|\boldsymbol{\alpha})/p(\mathbf{t}|\boldsymbol{\alpha},\sigma^2) \quad (6)$$

Sparse Bayesian learning can be formulated as a type–II maximum likelihood procedure; that is, a most probable point estimate α_{MP} may be found throughout the maximization of the marginal likelihood with respect to the hyperparameters a_i

$$L(\boldsymbol{\alpha})=-\frac{1}{2}\left[N\log 2\pi+\log|\mathbf{C}|+\mathbf{t}^T\mathbf{C}^{-1}\mathbf{t}\right], \quad (7)$$

The predictive distribution for a new datum $\mathbf{x}_*$ is defined as follows:

$$p(t_*|\mathbf{t},\boldsymbol{\alpha}_{MP},\sigma^2_{MP})=\int p(t_*|\mathbf{w},\sigma^2_{MP})p(\mathbf{w}|\mathbf{t},\boldsymbol{\alpha}_{MP},\sigma^2_{MP})d\mathbf{w} \quad (8)$$

which is easily computed due to the fact that both integrated terms are Gaussian, resulting in a Gaussian too.

$p(t_*|\mathbf{t},\boldsymbol{\alpha}_{MP},\sigma^2_{MP})=N(t_*|y_*,\sigma^2_*)$,

with $$\sigma^2_*=\sigma^2_{MP}+\phi(\mathbf{x}_*)^T\Sigma\phi(\mathbf{x}_*) \quad (9)$$

IV. BFO for an Improved Anisotropic Kernel

It is widely acknowledged that a key factor that affects the performance of the RVM is the choice of the kernel. However, owing to the difficulty of appropriately refining the parameters some different types of kernels are restricted in practice [24]. We present here a technique that allows us to deal with a large number of kernel parameters and thus use more complex kernels.

A. An Improved Anisotropic Gaussian Kernel

Gaussian kernel function outperforms others when the lack of a prior knowledge in about the learning process, but Gaussian function is a local kernel function, the map characteristic of the Gaussian function for Y_i =0.3 is shown in Figure 1 according to equation (14), it can be seen that there exists larger kernel function value only near the test point Y_i =0.3, and the farther the point is from the test point, the smaller its kernel function value becomes, finally it approaches zero rapidly. So, the Gaussian kernel function only has an effect on samples near the neighborhood of the test point instead of that far from the test point [25]. Based on this, an improved Gaussian kernel function is proposed as follows [26],

$$K(\boldsymbol{x},\boldsymbol{y})=\exp\left(\frac{2\sigma^2}{(\boldsymbol{x}-\boldsymbol{y})^2+p}+q\right) \quad (10)$$

Where p is displacement factor ($p>0$), and q is fine adjustment factor which is usually set as 0. So, for it, there are two parameters to be determined including the width factor σ and displacement factor p. The map characteristic of the improved Gaussian function for Y_i =0.3 is shown in Figure 2 according to equation (10), it is clear that the improved Gaussian kernel function has both the local characteristic and the global characteristic, and the function value far from the test point decreases slower than that of the Gaussian kernel function.

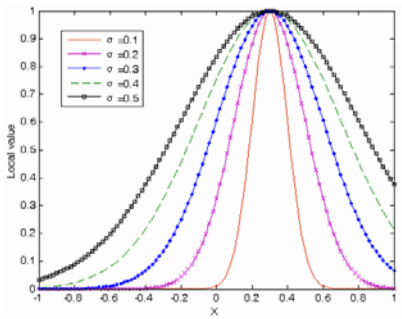


Figure 1. The map characteristic of Gaussian kernel function for Y_i =0.3

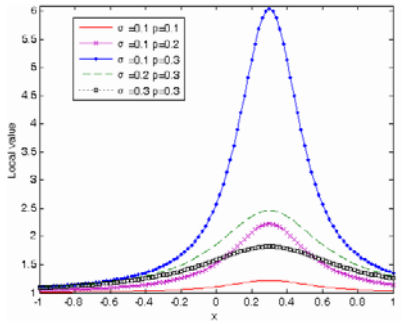


Figure 2.The map characteristic of improved kernel for Y_i =0.3i

Since many real-world databases contain characteristic attributes of very different natures, we consider sparse linear models whose kernel function is anisotropic kernel function, this is to say, we assign different values to the parameters of each kernel function for each input dimension, the anisotropic kernel function can lead to a significantly better performance than isotropic kernel function, so, in this paper, an improved anisotropic Gaussian kernel function is developed and is given by:

$$K(\mathbf{x},\mathbf{y}) = \exp\left(\sum_{i=1}^{d} \frac{2\sigma_i^2}{(x_i - y_i)^2 + p_i}\right) \qquad (11)$$

Where n equals the dimensionality of input vectors and for each width factor σ_i and displacement factor p_i: $\sigma_1 \neq \sigma_2 \neq \ldots \neq \sigma_d$, $p_1 \neq p_2 \neq \ldots \neq p_d$.

B. Kernel Parameters Optimization based on BFO

The determination of the kernel parameters need to meets a specific criterion. The following mean relative error is used as the performance criterion; meanwhile, it is also used as the fitness evaluation function of the BFO which is given by (12)

$$J_{fitness} = \frac{1}{N}\sum_{i=1}^{N} \frac{|y(i) - y^*(i)|}{y(i)} \qquad (12)$$

where $J_{fitness}$ is the fitness evaluation function, $y^*(i)$ is the prediction value of the ith sample, $y(i)$ is the actual value of the ith sample, and N is the number of samples.

Based on the aforementioned analyses, an improved anisotropic Gaussian kernel optimizing by BFO is proposed here, and its procedure is described as follows:

[Step 1] Initialize parameters S, N_c, N_s, N_{re}, N_{ed}, $C(i)$, P_{ed}, $C(i)(i=1,2,\ldots,S)$,φ^i, where
n: dimension of the search space,
S: the number of bacteria in the population,
N_c: number of chemotactic steps,
N_s: swimming length,
N_{re}: the number of reproduction steps,
N_{ed}: the number of elimination–dispersal events,
P_{ed}: elimination–dispersal with probability, and
$C(i)$: the size of the step taken in the random direction specified by the tumble.

[Step 2] Elimination-dispersal loop: $l = l+1$.

[Step 3] Reproduction loop: $k = k+1$.

[Step 4] Chemotaxis loop: $j = j+1$.

[a] For $i = 1,2,\cdots,N$, take a chemotactic step for bacterium i as follows.

[b] The fitness function $J(i,j,k,l)$ is calculated using equation (12).

Let $J(i,j,k,l) = J(i,j,k,l) + J_{cc}(\theta^i(j,k,l), p(j,k,l))$

(i.e., add on the cell-to-cell attractant–repellant profile to simulate the swarming behavior)

[c] Let $J_{last} = J(i,j,k,l)$ to save this value since we may find a better cost via a run.

[d] Tumble: generate a random vector $\Delta(i) \in R^n$ with each element $\Delta_m(i)$, a random number on $[-1,1]$.

[e] Move: let $\theta^i(j+1,k,l) = \theta^i(j,k,l) + C(i)\dfrac{\Delta(i)}{\sqrt{\Delta^T(i)\Delta(i)}}$

this results in a step of size $C(i)$ in the direction of the tumble for bacterium i.

[f]Compute $J(i,j+1,k,l)$ and let

$$J(i,j+1,k,l) = J(i,j,k,l) + J_{cc}(\theta^i(j+1,k,l), P(j+1,k,l))$$

[g] Swim.

i) Let m=0 (counter for swim length).

ii) While $m<N_s$ (if have not climbed down too long).

iii) Let $m=m+1$.

iv) If $J(i,j+1,k,l)<J_{last}$ (if doing better), let $J_{last} = J(i,j+1,k,l)$ and let

$$\theta^i(j+1,k,l) = \theta^i(j,k,l) + C(i)\frac{\Delta(i)}{\sqrt{\Delta^T(i)\Delta(i)}},$$

and use this $\theta^i(j+1,k,l)$ to compute the new $J(i,j+1,k,l)$ as we did in [f].

v) Else, let $m = N_s$.

[h] Go to next bacterium $(i+1)$ if $i \neq S$ (i.e., go to [b] to process the next bacterium).

[Step 5] If $j<N_c$, go to Step 4. In this case, continue chemotaxis since the life of the bacteria is not over.

[Step 6] Reproduction.

[a] For the given k and l, and for each $i=1,2,\cdots,S$, let

$$J_{health}^{i}=\sum_{j=1}^{N_c+1}J(i,j,k,l)$$

be the health of the bacterium i (a measure of how many nutrients it got over its lifetime and how successful it was at avoiding noxious substances). Sort bacteria and chemotactic parameters $C(i)$ in order of ascending cost J_{health}

[b] The S_r bacteria with the highest J_{health} die and the remaining S_r bacteria with the best values split.

[Step 7] If $k<N_{re}$, go to [Step 3]. In this case, we have not reached the number of specified reproduction steps, so we start the next generation of chemotactic loop.

[Step 8] Elimination-dispersal. For $i=1,2,\cdots,S$, with probability P_{ed}, eliminate and disperse each bacterium, and this result in keeping the number of bacteria in the population constant. To do this, if a bacterium is eliminated, simply disperse one to a random location on the optimization domain. If $l<N_{ed}$, then go to [Step 2]; otherwise end.

V. Experimental Results

The water heads are 56m, 63m, 68m and 73m respectively, each water head adopts 22 different operating condition points, the 88 samples corresponding to the characteristics of pressure fluctuation are used as a dataset, we select 2 samples from each water head randomly which are used as test samples, the remaining 80 samples are used as training samples. According to the field test data and the comprehensive characteristics, we establish the characteristics model of the pressure fluctuation among the operating water head, the output capacity and the dual amplitude based on RVM: $f=(H,N)$, where H is the operating water head, N is the output capacity, and f is the dual amplitude, it is clear that the input of the RVM model to be established is two dimensional, the kernel function corresponding to the RVM model should be as follows:

$$K(\mathbf{x},\mathbf{y})=\exp(\sum_{i=1}^{2}\frac{2\sigma_i^2}{(x_i-y_i)^2+p_i}) \quad (13)$$

So, the number of kernel parameters to be determined for the RVM model should be 4, they are σ_1, p_1, σ_2, p_2. which are optimized using the BFO, the parameter setup of the BFO is summarized in Table 1.

After the RVM model with the improved anisotropic Gaussian kernel is optimized by the BFO, the $J_{fitness}$ achieves 0.2162 and 17 relevance vectors are obtained when the optimal kernel parameters are σ_1= 29.3148, p_1,= 8.6663, σ_2= 66.3225, and p_2= 25.2493 respectively; finally the optimal kernel parameters are put into the RVM model to establish the pressure fluctuation characteristics of the turbine. Figure 3 illustrates the pressure fluctuation characteristics surface generated by the RVM with the improved anisotropic Gaussian kernel, the generated surface is complex and 17 out of 80 vectors are chosen as being relevant in this RVM model. Figure 4 epicts the pressure fluctuation curves under different water heads, all the relevance vectors are map out in each curve.

TABLE 1
PARAMETER SETUP FOR BFO

Parameter	S	N_c	N_s	N_{ed}	N_{re}	p_{ed}
BFO	50	100	12	4	16	0.25

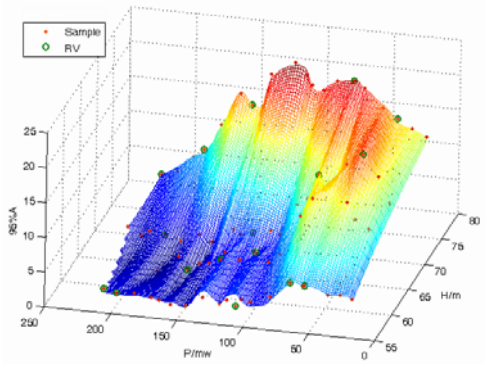
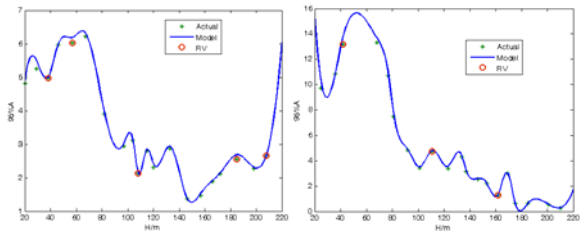


Figure 3. he fitting surface of operating points of RVM model


(a) Water head H=56 m (b) Water head H=63m

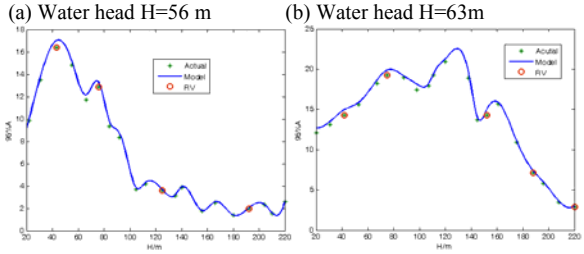

(c) Water head H=68m (b) Water head H=78m

Figure 4.The fitting surface of operating points of RVM model btained

For the comparison, the RVM model with the Gaussian kernel based on the BFO is also used to this experiment, the dataset and the parameters of BFO are the same as that of the proposed method, when the iteration is over the $J_{fitness}$ achieves 0.3651 and the optimal kernel parameters are σ= 29.3148, Figure 5 demonstrates the actual values versus the fitting values of two RVM model corresponding to the training sample points and test sample points respectively. It can be noticed that RVM with the improved kernel outperforms one with the Gaussian kernel, whether the training accuracy of training samples or the prediction accuracy of test samples. Table 2 lists the test relative error of 8 test sample points and the average relative error for the RVM models based on the two kernels under different water heads, it is obvious that the accuracy most of test sample points for the RVM model with the improved kernel is better than that of the

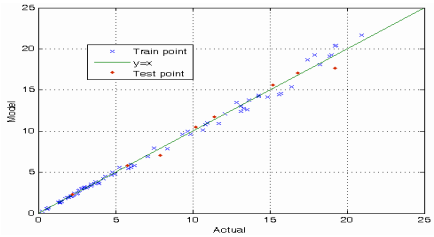
Gaussian kernel, and the average relative error for the improved kernel is 0.0219, that is lower than 0.0342 of the Gaussian kernel.

TABLE II
PREDICTING RESULTS OF TEST SAMPLES USING TWO RVM MODEL

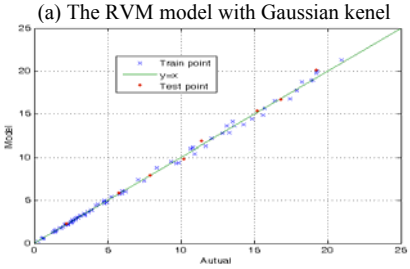
Water head /m	Output capacity /MW	Dual amplitude /95%A	Gaussian kernel /95%A	Improved kernel /95%A
56.00	105.37	2.20	2.2901	2.3280
56.00	42.24	5.82	5.8364	5.8311
63.00	83.15	7.91	7.0828	7.9854
63.00	74.62	10.24	10.5047	10.5254
68.00	68.30	11.41	11.6995	12.1398
68.00	33.72	15.23	15.5867	15.3536
78.00	56.21	16.81	17.0417	17.8336
78.00	110.26	19.22	17.6769	20.3552

Each training sample has an effect on the final solution to some degree, the weight of each train sample is used as the important degree which contributes to the final model Figure 6 shows the influence of each training ample for the two RVM model, the samples which do not influence on the RVM model are located in the x-axis, the samples which do not be in the x-axis would contribute to the final RVM model, for the RVM model with the improved kernel, the number of relevance vector is 17, for the RVM model with the Gaussian kernel, the number of relevance vector is 25, it shows the former is more sparse then the latter, which requiring less than a quarter of training samples can establish a good RVM model, and the training samples used as establishing the RVM model with the improved kernel, and the estimated variances values of the models are used to the RVM model with the Gaussian kernel too.

Table 3 summarizes the performance comparison for the RVM model corresponding to the two kernels, it presents not only the numbers of RV for the two RVM models, but only the MSE of the training samples and the test samples are listed, there are a great difference between the two model, to establish a good RVM model, the RVM model with the improved kernel requires only 20% of the training samples, is much less than 30% of the training samples required by the RVM model with the Gaussian kernel implying the great sparsity of the improved kernel, and then the MSE of training and test samples for the RVM model with the improved kernel are smaller than those for the RVM model with the Gaussian kernel.



(a) The RVM model with Gaussian kenel



(b) The RVM model with improved anisotropic Gaussian kernel

Figure 5 Actual values versus values of two RVM

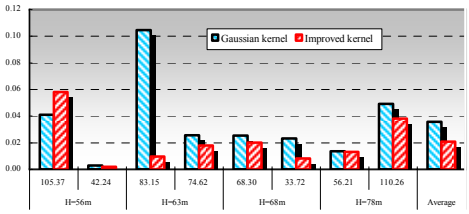


Figure 6.Test results for two RVM model

TABLE III
PERFORMANCE COMPARISON FOR THE RVM MODEL

Kernel method	number of RV	percent of RV	MSE of training set	MSE of test set
Gaussian kernel	25	30.125%	2.981%	3.416%
Improved kernel	17	20.125%	1.892%	2.361%

VI. CONCLUSIONS

In this paper, we have proposed an improved anisotropic Gaussian kernel used as the kernel function of the RVM model, and its parameters are optimized using the BFO, the acquired RVM model is applied to describing the pressure fluctuation characteristics of the draft tube of hydroelectric units of a hydropower station, the comparative simulation results show the proposed method can describe more accurately the pressure fluctuation characteristics of the draft tube, and then, it can give the right weight to the right characteristic which has more flexibility to reduce the number of relevance vector, it results in the great sparity for the good RVM model.

ACKNOWLEDGMENTS:

This work is supported by Chinese Universities Scientific Fund No. 2012YJS103, Natural Science Foundation of Hebei Province of China No.

E2010001026, and the Natural Science Foundation of China NO. 51075118.

REFERENCES

[1] HE cheng-lian, WANG zheng-wei, QIU hua. The experimental study of internal pressure fluctuation of draft tube. Journal of Machinery Industry, 2007, 38(11):62-65.

[2] Mural Yuichi, Yamamoto Fujio. Quadrant analysis of bubble induced velocity fluctuations in a transitional boundary layer. Journal of Hydrodynamics, 2009, 21(1):91-96.

[3] YUAN Shou-qi, NI Yong-yan, PAN Zhong-yong. Unsteady Turbulent Simulation and Pressure Fluctuation Analysis for Centrifugal Pumps. Chinese journal of mechanical engineering, 2009, 22(1): 64-69.

[4] JIAN hua, YAN Zhong-min. Hydraulic characteristics of bottom underlay-type pier for water-wing control [J]. Journal of Hydrodynamics, 2008, 20(6):65-73.

[5] WANG Xiao-ping,ZHEN Ling,HUANG Hai, CHENG Bo-chuan. Research on Pressure Fluctuations Time Serials by Deterministic Chaos Theory and Information En tropy.Chinese Journal of Scientific Instrument, 2007, 23(6):596-599.

[6] ZHAO Lin-ming, WEI De-hua, and HE Cheng-lian Neural network model for pressure fluctuation in draft tube of water turbine", Journal of Hydraulic Engineering, vol.36, No.11, pp1375-1378, 2005.

[7] Kwon, Y.-K., & Moon, B.-R.. A hybrid neurogenetic approach for stock forecasting. IEEE Transactions on Neural Networks, 18(3), 851-64,(2007).

[8] Kim, K. J., & Han, I.. Genetic algorithms approach to feature discretization in artificial neural networks for the prediction of stock price index. Expert Systems with Application, 19(2), 125–132 ,(2008)

[9] LI Zhijun, MENGWuzhou, MOU Kui. Modeling ofHydraulic Turbine Characteristic Based on Support Vector Machine, Modern Machinery,2010, 4:232-24.

[10] Hansen, J. V., & Nelson, R. D.. Neural networks and traditional time series Methods: A synergistic combination in state economic forecasts. IEEE Transactions on Neural Networks, 8(4), 863–873, (1997)

[11] Thissen U, Pepers M, U¨ stun B, Melssen WJ, Bydens LMC. Comparing support vector machines to PLS for spectral regression applications. Chemon. Intell. Lab. Syst. 2004; 73: 169-179.

[12] Belousov AI, Versakov SA, Von Frese J. A flexible classification approach with optimal generalization performance: support vector machines. Chemon. Intell. Lab. Syst. 2002; 64: 15-5.

[13] Belousov AI, Versakov SA, Von Frese J. Application aspects of supports vector machines. J. Chemom. 2002; 16: 482-489.

[14] Thissen U, van Brakel R, de Weijer AP, Melssen WJ, Bydens LMC. Using support vector machines for time series prediction. Chemon. Intell. Lab. Syst. 2003; 69: 35-49.

[15] Noslen Hernández, Isneri Talavera, Angel Dago, Rolando J. Biscay, Marcia M. Castro Ferreira, and Diana Porro, "Relevance vector machines for multivariate calibration purposes," Chemometrics, 2008(22):386-694.

[16] Tipping ME. The relevance vector machine[J]. Advances in Information Processing System, 2000, (2):652-658.

[17] Passino, K. M., "Biomimicry of bacterial foraging for distributed optimization and control," IEEE Control Systems Magazine, Vol. 22, No. 3, 52-67, 2002.

[18] Segall, J., S. Block, and H. Berg, "Temporal comparisons in bacterial chemotaxis," Proc. Nat. Acad. Sci., Vol. 83, 8987-8991, 1986.

[19] Liu, Y. and Passino, K. M.: "Biomimicry of Social Foraging Bacteria for Distributed Optimization: Models, Principles, and Emergent Behaviors," Journal of Optimization Theory And Applications: Vol. 115, No. 3, pp. 603–628, December 2002.

[20] V. Gazi, K.M. Passino, "Stability analysis of social foraging swarms", IEEE Transactions on Systems Man and Cybernetics Part B–Cybernetics, 34 (1) (2004) 539-57.

[21] Liu, Y. and Passino, K. M. Biomimicry of Social Foraging Bacteria for Distributed Optimization: Models, Principles, and Emergent Behaviors. Journal of Optimization Theory and Applications, 115:3, 2002, 603-628.

[22] Sambarta Dasgupta, Swagatam Das, Ajith Abraham, Arijit Biswas. Adaptive Computational Chemotaxis in Bacterial Foraging Optimization: An Analysis. IEEE Transactions on Evolutionary Computation, 13:4, 2009, 919-941.

[23] Wael Korani. Bacterial Foraging Oriented by Particle Swarm Optimization Strategy for PID Tuning. Proceedings of the 2008 GECCO conference companion on Genetic and evolutionary computation, 1823-1826, 2008.

[24] OLIVIER CHAPELLE, VLADIMIR VAPNIK, OLIVIER BOUSQUET, and SAYAN MUKHERJEE. Choosing Multiple Parameters for Support Vector Machines, Machine Learning, 46, 131-159, 2002.

[25] Yuan Xiaoyan, Liu Ailun.The Study and Application of Combinative kernel SVM, The Symposium of National Advanced Control Technology and Instrumentation Equipment Applications,56-61, 2007.

[26] YANG Liu, ZHANG Lei, ZHANG Shaoxun, andLIU Jianwei. Comparison Research of Single Kernel and Multi-kernel Relevance Vector Machine, Computer Engineering, 36(12):195-197, 2010.

Liying Wang was born in Shijiazhuang, Hebei Province, China, in 1978. She received the B.S. and the M.S. degrees both in Institute of Mechanism, Shijiazhuang Tiedao University, in 2000 and 2003 respectively, her current research interests include intelligent computing, intelligent control theory, and control systems engineering.