

Optimal parameters of the SVM for temperature prediction

XIAOYU SHI, QIANG HUANG, JIANXIA CHANG, YIMIN WANG,
JIANGQUN LEI & JING ZHAO

State Key Laboratory Base of Eco-Hydraulic Engineering in Arid Area, Xi'an University of Technology, Xi'an, 710048, P.R. China

Abstract This paper established three different optimization models in order to predict the Foping station temperature value. The dimension was reduced to change multivariate climate factors into a few variables by principal component analysis (PCA). And the parameters of support vector machine (SVM) were optimized with genetic algorithm (GA), particle swarm optimization (PSO) and developed genetic algorithm. The most suitable method was applied for parameter optimization by comparing the results of three different models. The results are as follows: The developed genetic algorithm optimization parameters of the predicted values were closest to the measured value after the analog trend, and it is the most fitting measured value trends, and its homing speed is relatively fast.

Key words temperature prediction; SVM; genetic algorithm; particle swarm optimization; developed genetic algorithm

1 INTRODUCTION

The influence of climate change on hydrology and water resources has become a topical hydrological issue. In recent years, global warming has become of great concern. So temperature prediction will be valuable for guiding public life and improving environment.

The principle of the support vector machine (SVM) method, which is an available method for small-scale samples, is structural minimization rather than empirical risk minimization (Gao *et al.*, 2013). Currently, scholars use the SVM method to solve a variety of support vector prediction problems, including short-term power load forecasting, human resource management forecasts, and predicting mineral resources. Its application to predict the temperature is still relatively small, and domestic, supporting vector regression in real-time forecasting applications (Feng *et al.*, 2005). In comparison to the SVM method, temperature prediction derived from the traditional support vector method was more accurate. However, parameter selection was a procrastination-inspiring and time-consuming process because it cannot be screened automatically. The input variables also cannot be selected effectively. Therefore, kernel function selection and parameter optimization was the research topic when SVM was applied to predict temperatures. Based on an optimization algorithm, the SVM method was set up. A grid search algorithm is the first parameter-selection method for SVM, which mainly finds the optimal by grid meshing. Although the use of a grid search could easily find the global optimal solution, it takes much more time for a larger-scale optimization. Using genetic algorithms, parameter optimization was perfect the first time, but the optimal values could only represent local optimization. Fewer GA evolutions are required for finding the optimal value range instead of the optimal values. The PSO optimization algorithm combined a speed and location search model. To find the optimal value, the speed and location of particles are constantly updated when they iterate by tracking the extreme, so their speed and position updated. However, it is easy to miss the optimal value because the convergence speed is becoming slower and swings in the vicinity during the late research period where the optimal value of the oscillation occurs. Thus, based on an improved genetic algorithm, the detailed grid mesh optimization was set up to find the optimal parameter range (Wang *et al.*, 2012). According to the developed genetic algorithm, three-dimensional parameters were optimised at a spatial angle to find the optimal parameter range, and it was verified using the grid method.

2 DATA SOURCES AND SVM METHOD

In this paper, the global re-analysis monthly mean data from the United States Prediction Center (NCEP) and the measured data within the Wei River basin from the National Weather site are used as large-scale climate data and regional climate data, respectively. Since the NCEP spatial

resolution of 2.5×2.5 and the latitude and longitude grid meteorological stations located 2×2 are different, we need to use statistical methods to establish the experience of the relationship between linear and nonlinear large-scale meteorological variables and regional meteorological variables through the establishment of a function based on a relationship between the feature amount for scale conversion function analysis. The monthly temperature data in Foping (1960–2010) and the 21 atmospheric circulation factor corresponding to the location of the weather station and the temperature, are closely related to an example of the temperature prediction analysis. These factors include Air850, Air500, Hgt850, Hgt500, Omega850, Omega500, Pres, Slp, Vwnd. A non-linear support vector machine was used to establish a statistical relationship model from 1960 to 2000 for temperature data modelling monthly and annually; temperature for the period 2001–2010 was predicted. Evaluation of the model is implemented by comparing the measured data of each period of time.

Support vector machine (SVM) is a newly developed classification technique, which aims at solving the classification problem by maximizing the margin between two opposite classes. One of the important tricks of SVM is the introduction of kernels, which enable SVM to have the ability of dealing with infinite or nonlinear features in a high dimensional feature space. Besides SVM, the kernel trick has also been widely used in many other machine learning algorithms, such as principal component analysis (PCA), linear discriminant analysis (LDA) and marginal Fisher analysis (MFA), resulting in many powerful kernel-based learning machines.

The SVM original form can expressed to the equation (1):

$$\text{Min} \frac{1}{2} \|w\|^2 + C \sum_{i=1}^i e_i \quad (1)$$

$$\text{Subject to: } y_i [(\omega x_i) + b] \geq 1 - \varepsilon_i \quad (i = 1, 2, 3 \dots, n), \varepsilon > 0$$

where C is a constant which is used to control punishment degree of misclassified samples. The larger the value of C , the greater loss of the objective function, indicating these samples are more significant.

The SVM kernel function in the input sample are values of low-dimensional space; the inner product value can be obtained under the high-dimensional space, so solving nonlinear problems. The RBF kernel function is most frequently applied kernel function. In this paper, SVM Gaussian kernel function was selected as RBF kernel function in the equation (2):

$$K(x_i, x_j) = \exp(-\gamma \|x_i - x_j\|^2), \gamma > 0 \quad (2)$$

Changing the kernel function parameters actually implicitly changes the mapping function, thereby changing the complexity of the spatial distribution of the sub-sample data. In SVM classification problem solving is also an important parameter. This article SVM parameter optimization, generally (c, γ) as the pending optimization variables.

3 THE SVM SETUP

3.1 PCA analysis

The PCA was employed for finding the major climate factors before using SVM for temperature prediction. PCA is a multivariate statistical method. In this paper, PCA on the dataset produced principal components (PCs) that explain over 90% of the total variances. In this short note, we demonstrate the use of principal components analysis (PCA) for one-class of support vector machine as a dimension reduction tool.

The selection of factors needs to follow the three requirements. First, the selected prediction indicators which are able to be accurately simulated by GCM, should have a strong correlation with predicted values. Secondly, screened factors must respond to the important physical processes and the rate of climate change at a large scale. Thirdly, no or little relationship exists between factors.

3.2 The SVM setup

The core concept of SVM method was mapping sample space into a high-dimensional space. An insensitive function was introduced in a featured space. Through defining the optimal hyperplane linear regression, an optimal value of a convex programming problem could be obtained (Chang *et al.*, 2008). The sample length was also considered when describing temperature changes by selecting indicators to guarantee enough samples participating in the model prediction. In this paper, the data of the former 40-years were applied for parameter set-up, while that of the last 10-years was available for forecasting. The predicted 10 years temperature from the SVM were compared with the actual air temperature. Different methods were carried out to find the optimal parameter. Then the relative optimal parameters were employed to build the model.

3.3 Evaluation of prediction model index

The temperature measurement and the absolute error of virtual series were analysed. The Euclidean distance parameter optimization results in different contrasts were used. The principle is the true distance between two points in space:

$$d = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + L(x_{10} - y_{10})^2} \quad (3)$$

where x_1, x_2, \dots, x_{10} represent predicted temperature under different methods and y_1, y_2, \dots, y_{10} represent monthly observed temperature data from the Foping site during 2001 to 2010.

4 PARAMETER OPTIMIZATION ALGORITHM OF THE SVM

When the support vector regression forecast was applied, the RBF kernel function, the loss coefficient (c) and epsilon value were set, defining c, g, p , respectively. The c was used to balance the training error and model complexity. The g determines the selection of the input space to feature space mapping method. Optimal parameters would be found after comparison between measured values of different parameters models (Wu *et al.*, 2011).

4.1 PSO parameter optimization algorithm

The particle swarm optimization was originally designed by Kennedy, it simulates the moving of social behaviour among individuals, through a multi-dimensional search space, each particle representing a single intersection of all dimensions. Particle swarm optimization algorithm simulates the behaviour of birds in their migration and cluster, and is a kind of evolutionary algorithm based on swarm intelligence (Mao *et al.*, 2013). Each member of the group is called one particle. With a certain speed, each particle has flight in a multi-dimensional search space. According to the particles' own experience, their flight speeds were dynamically adjusted, which indicated that each particle constantly revised its size and speed and so the single and group optimal values changed during the statistical iterations. Thus, a positive feedback mechanism of group's optimization is formed (Gao *et al.*, 2012).

The model is the combination of SVM in particle swarm optimization based on the monthly temperature prediction of Foping station. The nonlinear relationship between the forecasting is efficiently represented by the support vector machine, and the particle swarm optimization algorithm for support vector machine parameter selection in training.

In this paper, the optimal parameters were: $c = 10.3382$, $g = 0.01$, $p = 0.01$. The PSO method is stable, and better being less time-consuming.

4.2 GA parameter optimization algorithm

Genetic algorithms (GA) were proposed by Prof. Holland in the 1970s (Goldberg *et al.*, 1996). This method simulated the processes of selection, mutation, crossover and so on. GA is a global random search algorithm based on experience of the biological natural genetic mechanism where

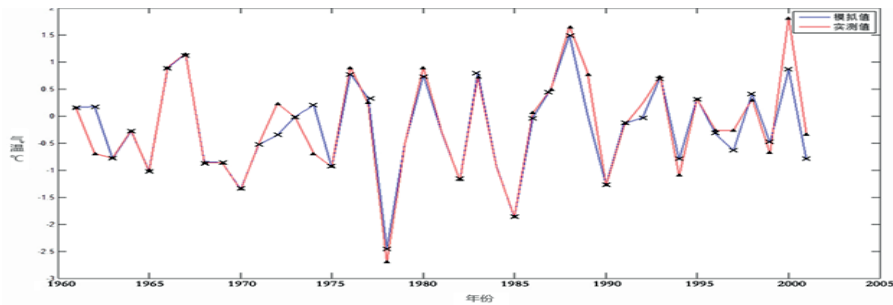


Fig. 1 Optimal temperature simulation based on PSO.

through constant reproduction the most suitable individual will be obtained (Leardi *et al.*, 1989). The main advantage of GA is that complex optimization problems can be solved by genetic operations. The initial population of GA is usually produced through genetic code, which is a potential solution set of optimization problem. The optimization process is a population genetically encoded by a certain number of individuals, with each individual chromosome an actual feature of entity. As the main carrier of genetic material on the chromosome, to facilitate the identification, you first need to realize the mapping from phenotype to genotype that is coded to work on behalf of the population after the original production, in accordance with the principle of survival of the fittest. The evolution of each generation to produce better approximate solutions, in every generation, is based on the individual fitness problem domain size of individual choice, and the help of Nature Genetics, combined genetic operators of crossover and mutation, resulting in a population representative of the new solution set (Hu *et al.*, 2011). This optimization process will make the same population as the natural evolution of an epigenetic generation of populations more than previous generations adapt to the environment, and the best individual in the last population is decoded.

The approximate optimal solution of the SVM model parameter optimization principle is initial operation of the climate impact of factors to calculate the value by adapting the objective function to obtain the most suitable parameters in the model to perform temperature forecasts. In this paper, the optimal parameters were: $c = 2.82259$, $g = 0.009537$, $p = 0.114848$. The GA parameter optimization method was more stable, but less time-consuming.

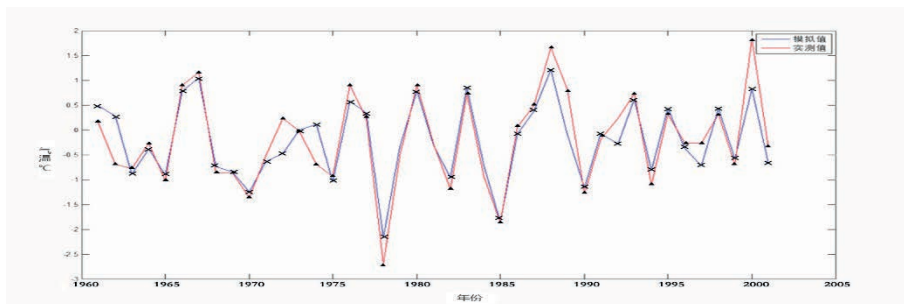


Fig. 2 Optimal temperature simulation based on GA.

4.3 Developed GA parameter optimization algorithm

In the GA developed to optimize the parameters of the process, the grid search and the GA were combined as an optimization strategy to optimize the three parameters of the SVM load model, and the more objective and effective cross-validation technique was introduced to train and assess the model. Generally, the grid search method in the optimization of the basic principle of parameters of SVM model is c and g is divided into a grid of transverse range (Yan *et al.*, 2012). The developed GA parameter optimization algorithm implementation process is the scope of the first parameter optimization through genetic algorithms to find the optimal parameters (Chatterjee *et al.*, 2010), and then use the detailed grid method to traverse all points within the grid value and find

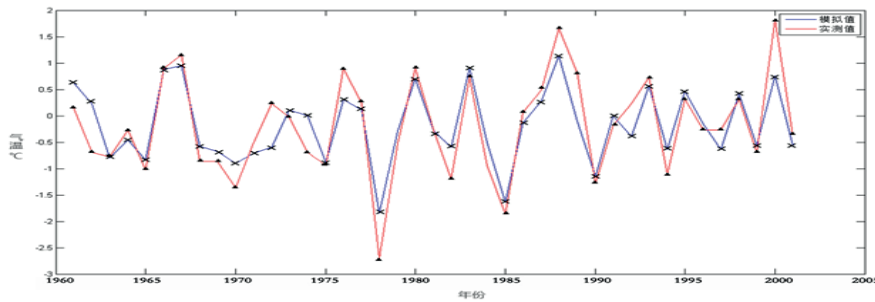


Fig. 3 Optimal temperature simulation based on developed GA.

the optimal parameter. In this paper, the optimal parameters were: $c = 5.1$, $g = 0.002$, $p = 0.187$. Although the developed genetic algorithm takes a relatively long time, the results obtained for the optimal parameters predicted by the model are very good.

4.4 Comparison among the three methods

Using the Euclidean distance calculation method we obtained:

$$d_{PSO} = 4.02933, d_{GA} = 3.383041, d_{GA+grid} = 2.67989, d_{PSO} > d_{GA} > d_{GA+grid} \quad (4)$$

Comparing predictions by the three parameter curves with the measured values (Fig. 4) shows the developed genetic algorithm is the best fit, followed by GA, and finally PSO.

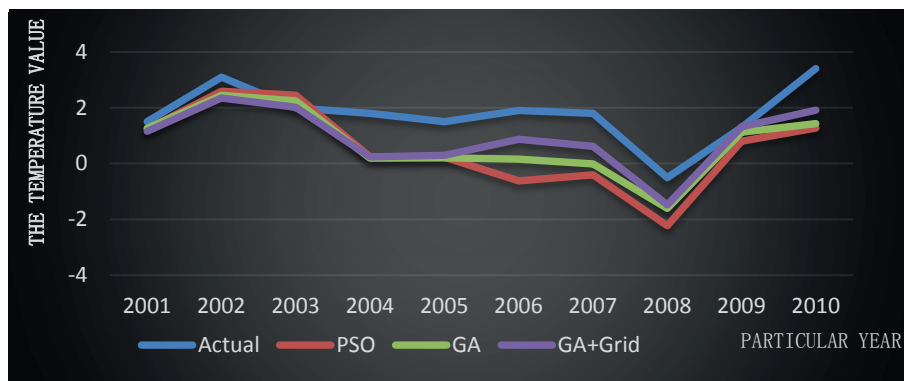


Fig. 4 Three kinds of optimization of the parameters of the trend line and method comparison chart.

In this paper, the SVM model is introduced into temperature prediction. Using multi angle analysis comparison, parameters optimized by the developed genetic algorithm gave predictive values closest to the actual values. The genetic algorithm and the particle swarm optimization algorithm were less successful. In contrast, PSO and GA showed that the genetic algorithm is faster, but the accuracy is poor, because the algorithm itself requires a large amount of data optimization constraints. The PSO algorithm is more stable than the original search capability of genetic algorithm, but also more time-consuming. The developed genetic algorithm and genetic algorithm are, in comparison, available, simple to use genetic algorithms for parameter optimization; with different sensitivities and selection of the initial value, the optimized model is different, it is easier to fall to a local minimum point. Although the developed genetic algorithm takes a relatively long time, the results obtained were optimal parameters and prediction by the model is very good. The developed genetic algorithm proposed for use in this paper.

5 CONCLUSION

The SVM model is introduced to predict temperature, and optimization of the parameters is essential and undoubtedly related to the prediction accuracy. This paper uses a genetic algorithm, particle swarm optimization and improved genetic algorithm parameter optimization algorithm.

Multi-angle analysis shows that using the improved genetic algorithm optimization parameters in the model values predicted closest to the measured values giving the best fit.

Acknowledgements This work was supported by the National Natural Science Foundation of China (grant nos 51190093, 51179149, 51179148), National Program on Key Basic Research Project (973 Program) in China (2011CB403306-3), Program for New Century Excellent Talents in University (NCET-10-0933), and governmental public industry research special funds for projects (201101043, 201101049).

REFERENCES

- Chatterjee, S., Bandopadhyay, S. and Machuca, D. (2010) Ore grade prediction using a genetic algorithm and clustering based ensemble neural network model. *Math. Geosci.* 42, 309–326.
- Gao Shang, Fucheng Liu and Hualong Yu (2013) Linear support vector machine regression models under different criteria. *Journal of Southeast University* 43, 44–47.
- Gao, C. *et al.* (2007) Price forecast in the competitive electricity market by support vector machine. *Physica A: Statistical Mechanics and its Applications* 382(1), 98–111.
- Goldberg, D.E. *Genetic Algorithms in Search, Optimization, and Machine Learning*. Addison-Wesley, Reading Menlo Park.
- Hanzhong Feng and Youyi Chen (2005) Support vector machine regression method in real-time business forecasting. *Journal of Applied Meteorology* 31, 41–44.
- Jia Wang (2011) Research on parameters optimization and application of mixed Kernel Support Vector Machine.
- Jianfeng Wang, *et al.* (2012) A parameter optimization method for an SVM based on improved grid search algorithm. *Applied Science and Technology* 39, 28–31.
- Leardi, R. (1996) *Genetic Algorithms in Feature Selection*. Academic Press, New York.
- Qi Wu (2011) A self-adaptive embedded chaotic particle swarm optimization for parameters selection of Wv-SVM. *Expert Systems with Applications* 38, 84–192.
- Radhika and Shashi (2009) Atmospheric temperature prediction using Support Vector Machines. *International Journal of Computer Theory and Engineering* 1, 1793–8201.
- Shaohua Ye (2010) Forecasting of water demand of city using improved Support Vector Machine. *Water Conservancy Science and Technology and Economy* 16, 495–497.
- Shutao Li and Mingkui Tan (2010) Tuning SVM parameters by using a hybrid CLPSO–BFGS algorithm. *Neurocomputing* 73, 2089–2096.
- Xioming Yan and Zhi Zheng (2011) Optimizing parameters of SVM based on combined bionic algorithm. *Journal of Guangxi Normal University* 29, 114–118.
- Zhongyi Hu (2011) Parameters optimization of SVMs based on memetic algorithms.
- Zongyu Mao, *et al.* (2013) Parameter optimization of SVM based on particle swarm optimization algorithm and GPU acceleration. *Journal of Xiamen University* 52, 609–612.