International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
IJISAE
ISSN:2147-6799 www.ijisae.org Original Research Paper

Application of SVM Network Model to Interpolate the Maximum and Minimum Ambient Temperature Parameters

Dinh Do Van

Submitted: 07/05/2023 **Revised**: 16/07/2023 **Accepted**: 05/08/2023

Abstract: In fact, knowing the weather information in a region is necessary for the planning, investment projects economic growth. However, not always regional meteorological observation stations, while nearby there are monitoring stations set. The problem posed by data measurement results of air monitoring stations located in neighborhoods that can be interpolated the results at the point of need to know. The problem of prediction parameters based on the results of further tests at the neighboring point is nonlinear interpolation problem, the results measured in the number of monitoring stations as input variables. Mathematical calculations provided for data processing field a lot of different interpolation algorithms [1, 2]. The learn to effectively apply these algorithms is also an important step in the process of data processing [3, 4]. Developing information technology for speeding up the adoption of more sophisticated algorithms, stronger to calculate, analyze and process data accurately [3]. Following the study forecasts the weather parameter interpolation method [5, 6], This paper proposes to use an artificial neural network SVM (Support Vector Machine) in model interpolation to predict the maximum and minimum temperature of the day on the measurement results of the monitoring stations nearby. The input data is the maximum value and minimum temperatures of the neighboring observation stations. The quality of the proposed solution is tested on real 2191 days data (from 01/01/2017 to 31/12/2022) in Hai Duong, Thai Binh, Bac Ninh and Quang Ninh province, Viet Nam.

Keywords: *monitoring stations, temperatures, interpolation, support vector machine.*

1. Introduction

Air temperature forecast is one of the main contents of the weather forecast, it has important implications for agriculture, industry and services,... to prevent and mitigate natural disasters, set design production plans, potential mining climate. The weather forecast data of the meteorological station or the central region has been compiled from the results of measurements of meteorological stations in the locality. Currently, due to the economic development - society in different areas, especially areas with high urbanization speed up the measurement result of a number of monitoring stations in the affected areas. In fact, in order to overcome this phenomenon may have to invest in building new monitoring stations in other locations are not affected by external factors or calibration specifications of monitoring stations

Alternatively, you can use interpolation methods of measuring results of monitoring stations nearby. Some traditional interpolation methods include Thiessen polygons, inverse distance interpolation, Kriging, splines, regression models and artificial neural network. According Journel and Huijbregts (1978) Kriging method [7] and Hornik et al., (1987) method of neural networks [8] is popular. However, Guhathakurta (2006) and Karmakar et al. (2008) have discovered that neural network method is rated better than other statistical methods [9]. There have been many neural network models used for parameters interpolation in weather problems. Rigol et al. (2001) describe the spatial interpolation of minimum air temperature daily using a neural network feed-forward back-propagation [10]. Simple network configuration was trained to predict the minimum temperature is used as input: (1) the date and topographical variations; (2) the temperature observed at a number of nearby places; (3) days, the terrain and variable ambient temperature observations. The article proposed application solutions artificial neural network (SVM - Support Vector Machine) modeling interpolation parameters based weather measurement results at the monitoring stations nearby. Following the model proposed using artificial neural network to interpolate weather parameters [11-17].

Sao Do University, Hai Duong-03500, Viet Nam
ORCID ID : 0000-0003-4425-2421
** Corresponding Author Email: dinh.dv@saodo.edu.vn*

2. Applications Model of Svm Networks for Interpolations Parameters Weather

2.1. Structure artificial neural network model SVM

Given a set of data including training samples which are input vectors (D dimensions) and the class code of each input vectors. We need to find a hyperplane $y = \mathbf{w} \cdot \mathbf{x} + b$ to classify the data set into two classes, with **w** is normal vector of hyperplane, used to adjust the direction of the hyperplane, value b is used to move the hyperplane parallel to itself.

We can use many hyperplanes to separate the data sets (Fig. 1) and there have been various algorithms to solve the problem, like Perceptron of Rosenblatt Algorithm [18], Linear Discriminant Algorithm of Fisher [19]. In SVM, however, optimal hyperplane is considered the hyperplane with the distance total to the nearest vectors of the two classes to be maximized. Besides, to guarantee a high generalibility, a slack variable is introduced to loose theconditions of class classification. The problem is led to solve the constrain optimization see equation (1).

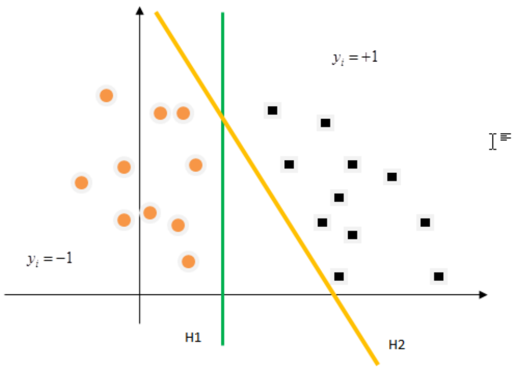

Fig. 1. The example two hyperplanes separating the two layers

$$\min_{w,b,\xi} \frac{1}{2} w^T \cdot w + C\sum_{i=1}^{N} \xi_i \qquad (1)$$

in order that

$$y_i(w^T \cdot x_i + b) + \xi_i - 1 \geq 0; \xi_i \geq 0, \forall i \in [1, N] \qquad (2)$$

with, C > 0 is the regularization parameter, ξ_i is slack variable. The problem (2) can be solved by SMO, where the method leads to solve the problem of Quadratic Programming:

$$\max_{\alpha} L(\alpha) = \sum_{i=1}^{N} \alpha_i - \frac{1}{2}\sum_{i,j} \alpha_i \cdot \alpha_j \cdot y_i \cdot y_j \cdot \Phi(x_i) \cdot \Phi(x_j) \qquad (3)$$

When achieving the α_i values from (3), we obtain the optimization **w** and *b* of hyperplane. The determined function of the subclasses has the form:

$$f(x) = \mathrm{sgn}\left(\alpha_i \cdot y_i \cdot \Phi(x_i) \cdot \Phi(x_j) + b^*\right) \qquad (4)$$

Supposing $K(x_i, x_j) = \Phi(x_i) \cdot \Phi(x_j)$ are the kernel function of the input space. The scalar product in the specializing space is equivalent to the kernel function $K(x_i, x_j)$ in the input space. Hence, instead of calculating directly the scalar product values, we indirectly perform via $K(x_i, x_j)$ for the next calculations.

2.2. Applications SVM networks modeling interpolation

The study model is constituted by Matlab®2010b, with SVM using the package LSSVMlabv1.8_R2009b_R2011a and designed accordingly the following steps: preparation data, constrituing forecast networks, choosing of methods and training the network, evaluate the dependability.

2.2.1. Preparing data

Ministry data for the study were obtained from the meteorological observation data from surface meteorological stations Hydrology River Delta region provided, which are:

- The monitoring parameters: Temperature maximum $(T_{\max})$, temperature minimum $(T_{\min})$,

- Monitoring duration: From 01/01/2017 to 31/12/2022,

- Monitoring regions: 4 provinces (Hai Duong, Thai Binh, Bac Ninh, Quang Ninh) in Viet Nam.

2.2.2. Constituing forecast networks

Assume that we need some parameters forecasting weather in Hai Duong (areas without monitoring stations), in fact to the construction forecast weather parameters we set the temperature measuring device, the humidity in a certain period of time from 6 months to 01 years of data will be used as the training data model (70% of the data) and used to test the model (30% of the data).

- Input data sets: Using data sets of 3 monitoring stations nearby Hai Duong province (Thai Binh, Bac Ninh, Quang Ninh);

- Output data sets: Use of monitoring data in Hai Duong province;

- Evaluating the dependability:

Mean Absolute Error: $MAE = \frac{1}{N}|y - d|$

Mean Relative Error: $MRE = \frac{1}{N}\left|\frac{y - d}{d}\right| \cdot 100\%$

Maximum Absolute Error: $MaxAE = \max|y - d|$

2.2.3. Model interpolation model used artificial neural network

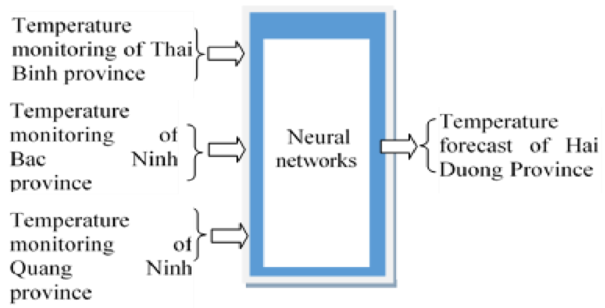

Fig. 2. The model interpolation using artificial neural network

In the experimental model the author used data dated 2191 and split into 02 sets of data: 1) The training set (70% of the data, the data corresponding date in 1481); 2) Set test (30% of data and 710 days of data respectively). Using the data observed in Hai Duong province to do the test model. Input data:

$$X_{input} = \begin{bmatrix} T_{ThaiBinh}(d-2191) & T_{BacNinh}(d-2191) & T_{QuangNinh}(d-2191) \\ T_{ThaiBinh}(d-2190) & T_{BacNinh}(d-2190) & T_{QuangNinh}(d-2190) \\ \vdots & \vdots & \vdots \\ T_{ThaiBinh}(d-2) & T_{BacNinh}(d-2) & T_{QuangNinh}(d-2) \\ T_{ThaiBinh}(d-1) & T_{BacNinh}(d-1) & T_{QuangNinh}(d-1) \end{bmatrix} \quad (5)$$

Output data:

$$Y_{output} = \begin{bmatrix} T_{HaiDuong}(d-2191) \\ T_{HaiDuong}(d-2190) \\ \vdots \\ T_{HaiDuong}(d-2) \\ T_{HaiDuong}(d-1) \end{bmatrix} \quad (6)$$

3. Results and Discussion

To evaluate the quality basis of the interpolation model, the article authors use polynomial interpolation, interpolation and interpolation arithmetic mean neurone network.

3.1. Polynomial interpolation

Assume the temperature in Hai Duong province and the temperature of the provinces of Thai Binh, Quang Ninh, Bac Ninh has ties to the equation (3).

$$T_{HaiDuong} \approx a \cdot T_{ThaiBinh} + b \cdot T_{BacNinh} + c \cdot T_{QuangNinh}$$

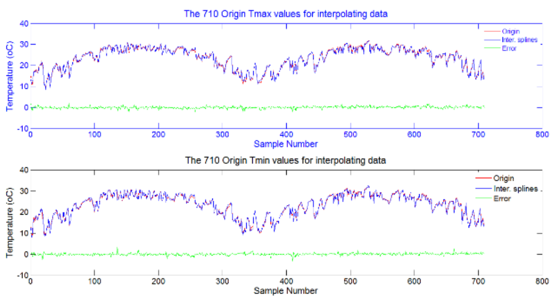

Fig. 3. Results of estimating the value T_{max}, T_{min} using polynomial interpolation methods

With data set for the maximum and minimum temperature, using the functions available in Matlab®R2010b us determine the coefficients a, b, c in the quations 7 corresponding T_{max} and T_{min}:

$$T_{max\_HaiDuong} \approx 0{,}3537 \cdot T_{max\_ThaiBinh} + 0{,}5637 \cdot T_{max\_BacNinh} + 0{,}0913 \cdot T_{max\_QuangNinh} \quad (8)$$

and

$$T_{min\_HaiDuong} \approx 0{,}3292 \cdot T_{min\_ThaiBinh} + 0{,}5774 \cdot T_{min\_BacNinh} + 0{,}1005 \cdot T_{min\_QuangNinh} \quad (9)$$

The quality evaluation interpolated using polynomial interpolation methods as Table 1. Using the polynomial interpolation method, the error for the testing process reached 1,81% when estimating T_{max} and 1,78% when estimating T_{min}.

Table 1. Results of errors when using polynomial interpolation method to estimate T_{max}, T_{min}

Temperature	MAE	MRE (%)	MaxAE
T_{max}	0,39	1,81	2,15
T_{min}	0,38	1,78	3,43

3.2. Interpolation method using average:

Temperatures in Hai Duong province with the average temperature of the neighbor provinces as the equation (10).

$$T_{HaiDuong} = \frac{T_{ThaiBinh} + T_{BacNinh} + T_{QuangNinh}}{3} \quad (10)$$

The quality evaluation interpolated using interpolation middle average methods as Table 2. With the interpolation method, the average error for the testing process is about more than 2,2% for T_{max} and T_{min}.

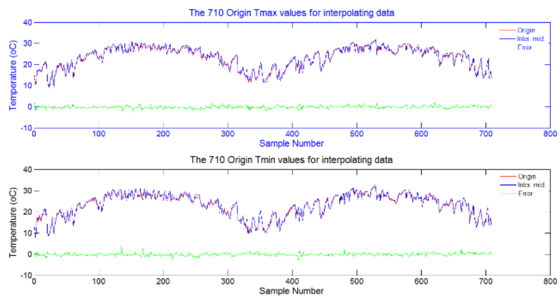

Fig. 4. Results estimate T_{max}, T_{min} interpolation method average

Table 2. Table Results of errors using the interpolation middle average method to estimate T_{max}, T_{min}

Temperature	MAE	MRE (%)	MaxAE
T_{max}	0,49	2,26	2,33
T_{min}	0,47	2,18	3,80

3.3. Interpolations using artificial neural networks

With the data set proposed in section 2.2.1. The author uses MLP, MLR, Elman and SVM neural networks to interpolate the predicted values.

Table 3. Results of errors using artifical neural network model interpolation maximum temperature

Neural networks	Error learning Tmax			Error testing Tmax		
	MAE	MRE(%)	MaxAE	MAE	MRE(%)	MaxAE
MLP	0,58	2,77	0,58	0,53	2,42	0,53
MRL	0,36	1,72	10,01	0,39	1,79	2,15
Elman	0,57	2,75	0,57	0,53	2,41	0,53
SVM	*0,36*	*1,73*	*9,95*	*0,39*	*1,79*	*2,1*

Table 4.

Table 5. Results of errors using artifical neural network model interpolation minimum temperature

Neural networks	Error learning T_{min}			Error testing T_{min}		
	MAE	MRE(%)	MaxAE	MAE	MRE(%)	MaxAE
MLP	0,52	2,58	0,52	0,54	2,48	0,54
MRL	0,35	1,75	9,14	0,38	1,78	3,43
Elman	0,54	2,66	0,54	0,54	2,48	0,54
SVM	*0,34*	*1,69*	*9,10*	*0,38*	*1,74*	*3,44*

Experimental whith interpolation models artificial neural networks different, such as MLP network, MLR, Elman and SVM; Results of the evaluation of the quality indicators as table 3 and table 4. Predicted results as table 3, table 4. See Table 3 and Table 4, model SVM networks for best results. Test error of SVM neural networks about 1,8%.

Fig. 5. Results of interpolation to estimate T_{max} whith set of data and testing data

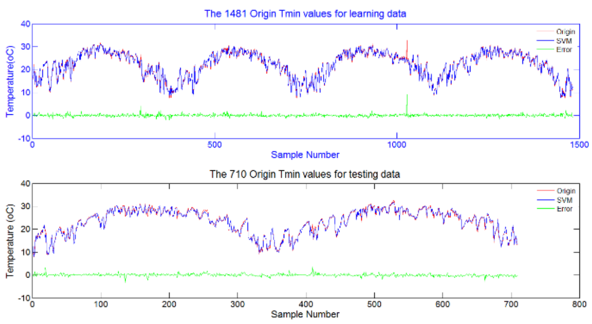

Fig. 6. Results of interpolation to estimate T_{min} whith set of data and testing data

Figures 5 and 6 are the results of the learning and testing process using the SVM neural network for interpolation to estimate the Tmax and Tmin values. Through the responses, the estimated value is always close to the actual measurement results at the monitoring stations, which shows that the proposed solution achieves good results.

4. Conclusions

The SVM neural network was used to determine the dependence funtions between temperatures of three monitoring stations (Thai Binh, Bac Ninh and Quang Ninh) and the measured temperature in Hai Duong with absolute average results are under 2,0% and without placing monitoring stations.

The application of neural solution is a proper research allows us to use the new technology, giving accurate results.

References

[1] Duong Thuy Vy, "Calculation method", *Publisher of Scie nce and Technology*. 2001

[2] Palovko A., Butusov P., "Calculation method", *Publisher BKV- Peterbua*. 2004.

[3] Sadanori Konishi, *Introduction to Multivariate Analysis: Linear and Nonlinear Modeling*, Texts in Statistical Science.

[4] Mathews J.H (2001), *Numerical methods Using Matlab*, Prentice Hall.

[5] R. Sluiter, "Interpolation methods for climate data", *KNMI, R&D Information and Observation Technology De Bilt*, 19 November 2008, Version 1.0.

[6] Alexandre Lanciani and Marta Salvati, "Spatial interpolation of surface weather observations in Alpine meteorological services", *Foralps Technical Report*, Università degli Studi di Trento, Dipartimento di Ingegneria Civile e Ambientale, Trento, Italy, 2008, pp. 40.

[7] Journel, A. G., Huijbregts, C. J., "Mining geostatistics", *Academic Press*, New York, 1978.

[8] Hornik, K., Stinchcombe, M., White, H., "Multilayer feedforward networks are universal approximators", *Neural Networks*, 1989, 2, 359-366.

[9] Guhathakurta, P., "Long Range Forecasting Indian Summer Monsoon Rainfall By Principle Component Neural Network Model", *J. Meteor. & Atomos. Phys*., 71, 255-266.

[10] Guhathakurta, P., Rajeevan, M., Thapliyal, V., "Long Range Forecasting Indian Summer Monsoon Rainfall by Hybrid Principal Component Neural Network Model", *Meteorology and Atmospheric Physics*, Springer-Verlag, Austria, 1999.

[11] Snell, S. E., “Spatial Interpolation of Surface Air Temperatures Using Artificial Neural Networks: Evaluating Their Use for Downscaling GCMs”, *Journal of Climate*, 2000, 13, 886-895

[12] Rigol, J.P., Jarvis, C.H., Stuart, N., “Artificial neural networks as a tool for spatial interpolation”, *International Journal of Geographical Information Science*, 2001, 15, 323-343

[13] Koike, K., Matsuda, S., Gu, B., “Evaluation of Interpolation Accuracy of Neural Kriging with Application to Temperature-Distribution Analysis”, Mathematical Geology, 2001, 33(4), 421-448.

[14] Kumar, S., “Neural Network Computer Engineering Serie”, *The McGraw-Hill*, New Delhi, 2007, 104-152

[15] Bryan, B.A., Adams, J. M., “Three-Dimensional Neurointerpolation of Neural Networkual Mean Precipitation and Temperature Surfaces for China”, *Geographical Analysis*, 2002, 34 (2), 93-111

[16] Silva, A. P., “Neural Networks Application to Spatial Interpolation of Climate Variables”, *Carried Out By*, STSM on the Framework of COST 719 ZAMG, 2003.

[17] Sivapragasam, C., Arun, V.M., Giridhar, D., “A simple approach for improving spatial interpolation of Rainfall using Neural network”, *Meteoro Atmos Phys*, Springer, 2010, 109, 1-7

[18] D.E. Rumelhart; G.E. Hinton and R.J. Williams, “Learning internal epresentations by error propagation”, *Parallel distributed processing: Explorations in the microstructure of cognition* (Cambridge MA.: MIT Press), 1986, 318-362.

[19] R.A. Fisher, “The Use of Multiple Measurements in Taxonomic Problems”, *in Annals of Eugenics*, 1936, No 7, pp. 102-106

[20] Yanamadni, V. R. ., Seetha, J. ., Kumar, T. S. ., Kannaiah, S. K. ., J, B. ., & Brahmaiah, M. . (2023). Computer-Aided Detection of Skin Cancer Detection from Lesion Images via Deep-Learning Techniques. International Journal on Recent and Innovation Trends in Computing and Communication, 11(2s), 293–302. https://doi.org/10.17762/ijritcc.v11i2s.6158

[21] Kamau, J., Ben-David, Y., Santos, M., Joo-young, L., & Tanaka, A. Predictive Analytics for Customer Churn in the Telecom Industry. Kuwait Journal of Machine Learning, 1(3). Retrieved from http://kuwaitjournals.com/index.php/kjml/article/view/130

[22] Shukla, A., Almal, S., Gupta, A., Jain, R., Mishra, R., & Dhabliya, D. (2022). DL based system for on-board image classification in real time, applied to disaster mitigation. Paper presented at the PDGC 2022 - 2022 7th International Conference on Parallel, Distributed and Grid Computing, 663-668. doi:10.1109/PDGC56933.2022.10053139 Retrieved from www.scopus.com