

Precipitation Estimation Using Support Vector Machine with Discrete Wavelet Transform

Mohamed Shenify¹ · Amir Seyed Danesh² ·
Milan Gocić³ · Ros Surya Taher⁴ ·
Ainuddin Wahid Abdul Wahab⁵ · Abdullah Gani⁵ ·
Shahaboddin Shamshirband⁵ · Dalibor Petković⁶

Received: 28 December 2014 / Accepted: 27 October 2015 /
Published online: 4 November 2015
© Springer Science+Business Media Dordrecht 2015

Abstract Precipitation prediction is of dispensable importance in many hydrological applications. In this study, monthly precipitation data sets from Serbia for the period 1946–2012 were used to estimate precipitation. To fulfil this objective, three mathematical techniques named artificial neural network (ANN), genetic programming (GP) and support vector machine with wavelet transform algorithm (WT-SVM) were applied. The mean absolute error (MAE), mean absolute percentage error (MAPE), root mean square error (RMSE), Pearson correlation coefficient (r) and coefficient of determination (R^2) were used to evaluate the performance of the WT-SVM, GP and ANN models. The achieved results demonstrate that the WT-SVM outperforms the GP and ANN models for estimating monthly precipitation.

Keywords Precipitation · Support vector machine · Discrete wavelet transform · Genetic programming · Artificial neural network

✉ Shahaboddin Shamshirband
shamshirband@um.edu.my

¹ Department of Computer Science, College of Computer Science and Information Technology, Albaha University, Albaha 65431, Kingdom of Saudi Arabia

² Department of Software Engineering, Faculty of Computer Science and Information Technology, University of Malaya, 50603 Kuala Lumpur, Malaysia

³ Faculty of Civil Engineering and Architecture, University of Niš, Aleksandra Medvedeva 14, 18000 Niš, Serbia

⁴ Faculty of Computer & Mathematical Sciences, Universiti Teknologi MARA, Shah Alam, Malaysia

⁵ Department of Computer System and Technology, Faculty of Computer Science and Information Technology, University of Malaya, 50603 Kuala Lumpur, Malaysia

⁶ Faculty of Mechanical Engineering, University of Niš, Aleksandra Medvedeva 14, 18000 Niš, Serbia

1 Introduction

Changes in precipitation patterns have an impact on water resources, agricultural production and global biodiversity. Thus, it is necessary to provide its accurate estimation and forecasting.

In recent years, soft computing methods have been used to estimate and forecast precipitation. Abbot and Marohasy (2014) applied artificial neural network (ANN) to forecast Queensland monthly rainfall as a continuous variable. It was found that the ANN forecasts were superior to Australian official forecasts. Wu et al. (2010) compared modular ANN (MANN) with ANN, K-nearest-neighbors (K-NN) and linear regression (LR), which was used in monthly rainfall series forecasting. The attained results indicated higher capability of MANN compared other selected models. Freiwani and Cigizoglu (2005) examined the precision of ANN technique to estimate monthly precipitation. Valverde Ramírez et al. (2005) utilized ANN to forecast daily rainfall. Their results demonstrated that ANN enjoys higher performance compared to the linear regression model. The similar results were provided by Mekanik et al. (2013), who compared the ANN and multiple regression analysis (MR). Moustiris et al. (2011) used ANN for forecasting the monthly maximum, minimum, mean and cumulative precipitation in Greece.

Genetic algorithm (GA) was utilized to model hydrological processes with no information on the exact relationship between their components. Nasserri et al. (2008) developed a model by combining the ANN with GA for simulating the rainfall field. The obtained results showed the coupled model provides more accuracy than the single ANN. Sedki et al. (2009) studied the efficiency of the GA for rainfall–runoff forecasting.

Support Vector Machine (SVM) (Vapnik 1995, 1998; Shamshirband et al. 2014) is based upon the structural risk minimization (SRM), which application improves the ability of the SVM for classification, pattern recognition and regression problems (Lu et al. 2011; Yang et al. 2014; Dileep and Sekhar 2014; Guo et al. 2014; Harris 2015). There are two main categories for SVMs namely Support Vector Classification (SVC) and Support Vector Regression (SVR). Recently, the SVM has obtained importance in forecasting and estimation problems of precipitation (Tripathi et al. 2006; Ortiz-García et al. 2014; Sánchez-Monedero et al. 2014). Each SVR model includes three parameters of gamma (γ), cost (C), and epsilon (ϵ).

Over the past years, SVM with wavelet (WT) has seen to made enormous interest in engineering applications (Peng and Chu 2004; Wang and Li 2011; Kalteh 2013, 2015; Du et al. 2013; Sahay and Srivastava 2014; Shiau and Huang 2014; Liang and Juang 2015; Sun et al. 2015; Yong et al. 2015; Mohammadi et al. 2015). Kisi and Cimen (2012) employed WT-SVM model to forecast daily precipitation. They concluded that the hybrid approach enhance the precision and offer more precision than the single SVM. Nourani et al. (2014) used precipitation satellite data in combination with the feed-forward neural network (FFNN) and the WT for modeling rainfall–runoff process on a daily and multi-step ahead time scales. They concluded that the application of the WT to the runoff data improve the capability of the FFNN rainfall–runoff models for predicting runoff peak values. Venkata Ramana et al. (2013) used Wavelet Neural Network (WNN) model, which is the combination of wavelet analysis and ANN to forecast monthly rainfall at Darjeeling rain gauge station, India. Feng et al. (2015) implemented wavelet analysis-support vector machine coupled (WA-SVM) model for monthly rainfall forecasting in arid regions.

In this study, an estimation model based on three soft computing methods namely WT-SVM, ANN and GP were developed. The precipitation data from Serbia was used as case study.

2 Material and Methods

2.1 Study Area and Data Collection

The study area was Serbia, which is located in the central part of the Balkan Peninsula between 41°53' and 46°11' latitude North and 18°49' and 23°00' longitude East with an area of 88,407 km². The climate of the country is temperate continental, with a gradual transition between the four seasons of the year. West Serbia is the wettest part of the country, while the north and south are the driest. The rivers in Serbia belong to the watersheds of the Black, Adriatic and Aegean seas. On more than 90 % of the Serbian territories there are rivers that join the Danube, thus going to the Black Sea.

Riverine and torrential floods are the most significant natural hazards on the territory of Serbia. Torrential flood represents sudden appearance of maximal discharge in river bed with high concentration of hard phase. Climate, specific characteristics of relief, distinctions of soil and vegetation cover, socio-economic conditions have done that the occurrence of torrential flood waves is one of the resulting forms of existing erosion processes. Thus, the accurate estimation of precipitation will help decision makers to plan water resources, agricultural production and land use in the region.

Series of monthly precipitation data were collected from 29 principal meteorological stations for the period of 1946–2012 (Fig. 1). The detail analysis of these data was done in Gocic and Trajkovic (2014a, b) using different multivariate analysis techniques such as R-mode principal component analysis, factor analysis and agglomerative hierarchical cluster analysis, as well as the more common descriptive statistical analysis. The mean annual precipitation varied between 549.7 and 954.9 mm, while the average precipitation for the whole country was 662.4 mm with the standard deviation of 132 mm. The most precipitation falls in June and May, while the least rainy month is February. The main amount of precipitation fell in the regions along the greatest rivers such as the Danube, the Sava and the Velika Morava.

According to the experiments, the input parameters (months and years) and output (precipitation data) are collected from 29 stations and averaged for Serbia and defined for the learning techniques. For the experiments, 60 % of the data was used to train samples and the subsequent 40 % served to test samples.

2.2 Support Vector Machine (SVM)

The SVM (Vapnik 1995, 1998) is based upon the principles of the statistical machine-learning process. One of the advantages of SVM is minimization of structural risks, which minimize the upper-bound generalization error rather than the local training error. SVM estimates the function as:

$$f(x) = w\varphi(x) + b \quad (1)$$

$$R_{SVMs}(C) = \frac{1}{2} \|w\|^2 + C \frac{1}{n} \sum_{i=1}^n L(x_i, d_i) \quad (2)$$

where $R = \{x_i, d_i\}_i^n$ is set of data points, x_i is the input vector, d_i is the target value, n is size of the data $\varphi(x)$ is the space feature, w is a normal vector, b is a scalar and

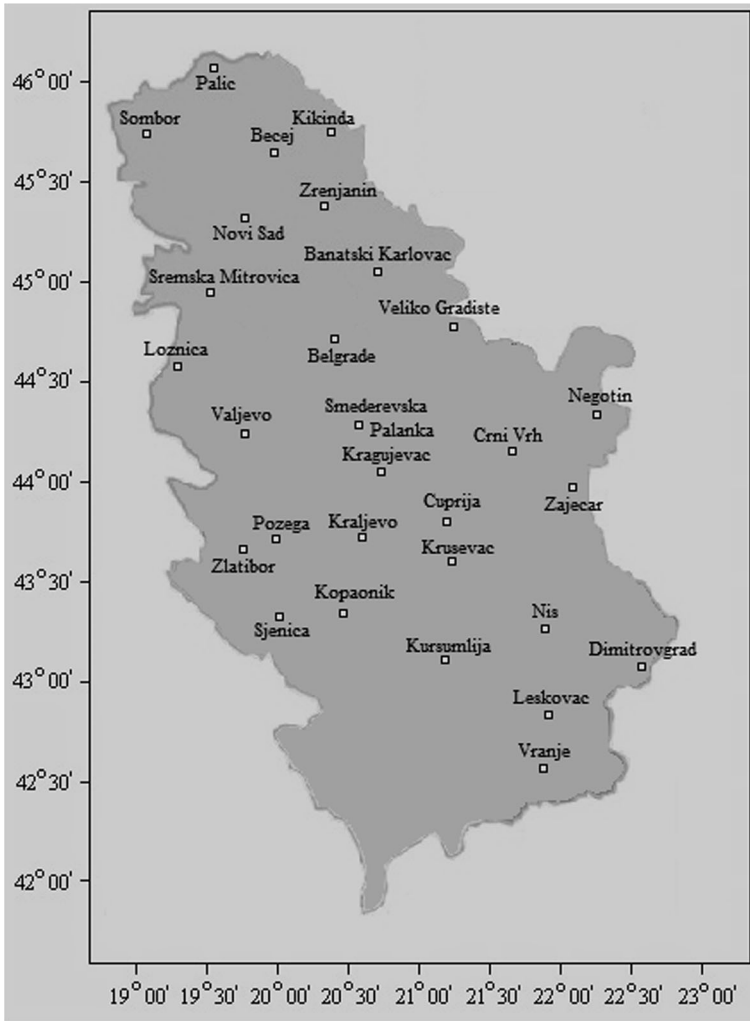


Fig. 1 Spatial distribution of the 29 principal meteorological stations in Serbia map

$C \frac{1}{n} \sum_{i=1}^n L(x_i, d_i)$ is the empirical error. The w and b parameters are estimated by minimizing the function Eq. (2) as:

Minimize

$$R_{SVMs}(w, \xi^*) = \frac{1}{2} \|w\|^2 + C \sum_{i=1}^n (\xi_i + \xi_i^*) \tag{3}$$

Subject to $\{d_i - w\varphi(x_i) + b_i \leq \varepsilon + \xi_i, w\varphi(x_i) + b_i - d_i \leq \varepsilon + \xi_i^*, \xi_i, \xi_i^* \geq 0, i = 1, \dots, l\}$ where ξ_i and ξ_i^* are positive slack variables, $\frac{1}{2} \|w\|^2$ is the regularization, C is penalty factor, ε is the loss function and l is the number of elements in the training dataset.

Equation (1) is estimated by Lagrange multiplier and optimality constraints, thereby obtaining the function given by Eq. (4).

$$f(x, a_i a_i^*) = \sum_{i=1}^n (a_i - a_i^*) K(x, x_i) + b \tag{4}$$

where $K(x, x_i) = \varphi(x_i)\varphi(x_i)$ is kernel function.

2.3 Discrete Wavelet Transform Algorithm

The wavelet transform (WT) is a mathematical expression for decomposing a time series data into several groups in order to better analysis of the components (Adamowski and Chan 2011; Jawerth and Sweldens 1994).

The discrete wavelet transform (DWT) is represented as:

$$W_x(m, n, \psi) = a_0^{-m/2} \int_{-\infty}^{\infty} f(t) \psi^* (a_0^{-m}t - nb_0) dt \tag{5}$$

where the parameters a and b are

$$a = a_0^m, \quad b = na_0^m b_0 \tag{6}$$

2.4 Artificial Neural Networks

The multi-layer network with a back propagation learning algorithm is one of the most popular neural network architectures. A neural network consists of three layers:

- (1) an input layer;
- (2) an output layer; and
- (3) an intermediate or hidden layer.

The input vectors are $D \in R^n$ and $D = (X_1, X_2, \dots, X_n)^T$; the outputs of q neurons in the hidden layer are $Z = (Z_1, Z_2, \dots, Z_n)^T$; and the outputs of the output layer are $Y \in R^m$, $Y = (Y_1, Y_2, \dots, Y_n)^T$. Assuming that the weight and the threshold between the input layer and the hidden layer are w_{ij} and y_j , respectively, and that the weight and the threshold between the hidden layer and output layer are w_{jk} and y_k , respectively, the outputs of each neuron in a hidden layer and output layer are:

$$Z_j = f \left(\sum_{i=1}^n w_{ij} X_i - \theta_j \right) \tag{7}$$

$$Y_k = f \left(\sum_{j=1}^q w_{kj} Z_j - \theta_k \right) \tag{8}$$

where $f()$ is an exchange capacity, which is the principle for mapping the neuron’s summed information to its yield, and, by a suitable decision, it is a method for presenting a non-linearity into the system’s outline. A standout among the most usually-utilized capacities is the sigmoid capacity, which is monotonically expanding and extends from zero to one.

2.5 Genetic Programming

Genetic programming (GP) is an evolutionary calculation focused around Darwinian hypotheses of common choice and survival to estimate the comparison, in typical structure, that best portrays how the yield identifies with the information variables. The calculation considers a beginning populace of haphazardly-created projects (comparisons), determined from the irregular blending of information variables, arbitrary numbers and capacities, which incorporate, for example, mathematical operators (+, −, ×, and ÷), mathematical capacities (sin, cos, exp, log), and sensible/examination capacities, which must be chosen properly focused on some understanding

of the methodology. Then, this populace of potential arrangements is subjected to an evolutionary procedure, and the “wellness” (a measure of how well they take care of the issue) of the advanced projects is assessed. Then, the individual projects that best fit the information are chosen from the starting populace. The projects that provide the best fit are chosen to trade a piece of the data between them to deliver better projects through “hybrid” and ‘transformation’, which copy the common world’s generation process. Trading the parts of the best projects with one another is called hybrid, and arbitrarily changing projects to make new projects is called change. The projects that fit the information less well are disposed of. This development procedure is rehashed over progressive eras and is determined towards discovering typical statements depicting the information, which can be deductively translated to determine learning about the methodology.

2.6 Evaluation Criteria

To assess the proficiency of the WT-SVM, GP and ANN models, the following statistical parameters were used:

- 1) mean absolute error (MAE)

$$MAE = \frac{1}{n} \sum_{i=1}^n |P_i - O_i|, \tag{9}$$

- 2) mean absolute percentage error (MAPE)

$$MAPE = \frac{100\%}{n} \sum_{i=1}^n \left| \frac{O_i - P_i}{O_i} \right|, \tag{10}$$

- 3) root mean square error (RMSE)

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (P_i - O_i)^2}{n}}, \tag{11}$$

- 4) Pearson correlation coefficient (r)

$$r = \frac{n \left(\sum_{i=1}^n O_i \cdot P_i \right) - \left(\sum_{i=1}^n O_i \right) \cdot \left(\sum_{i=1}^n P_i \right)}{\sqrt{\left(n \sum_{i=1}^n O_i^2 - \left(\sum_{i=1}^n O_i \right)^2 \right) \cdot \left(n \sum_{i=1}^n P_i^2 - \left(\sum_{i=1}^n P_i \right)^2 \right)}} \tag{12}$$

- 5) coefficient of determination (R²)

$$R^2 = \frac{\left[\sum_{i=1}^n (O_i - \bar{O}_i) \cdot (P_i - \bar{P}_i) \right]^2}{\sum_{i=1}^n (O_i - \bar{O}_i) \cdot \sum_{i=1}^n (P_i - \bar{P}_i)} \tag{13}$$

where P_i [mm] and O_i [mm] are the predicted and observed values of precipitation, respectively and n is the total number of test data.

3 Results and Discussion

3.1 Experimental Data

The soft computing models were trained with estimated precipitation data. The testing and training data are presented in Fig. 2. 70 % of data was used for training and 30 % of data was used for testing purpose of the soft computing models.

3.2 Performance Analysis

In this study, to perform the simulation, the Demuth Neural Network toolbox in MATLAB R2010a was utilized to develop the ANN model. Also, the libSVM (Chau and Wu 2010) in

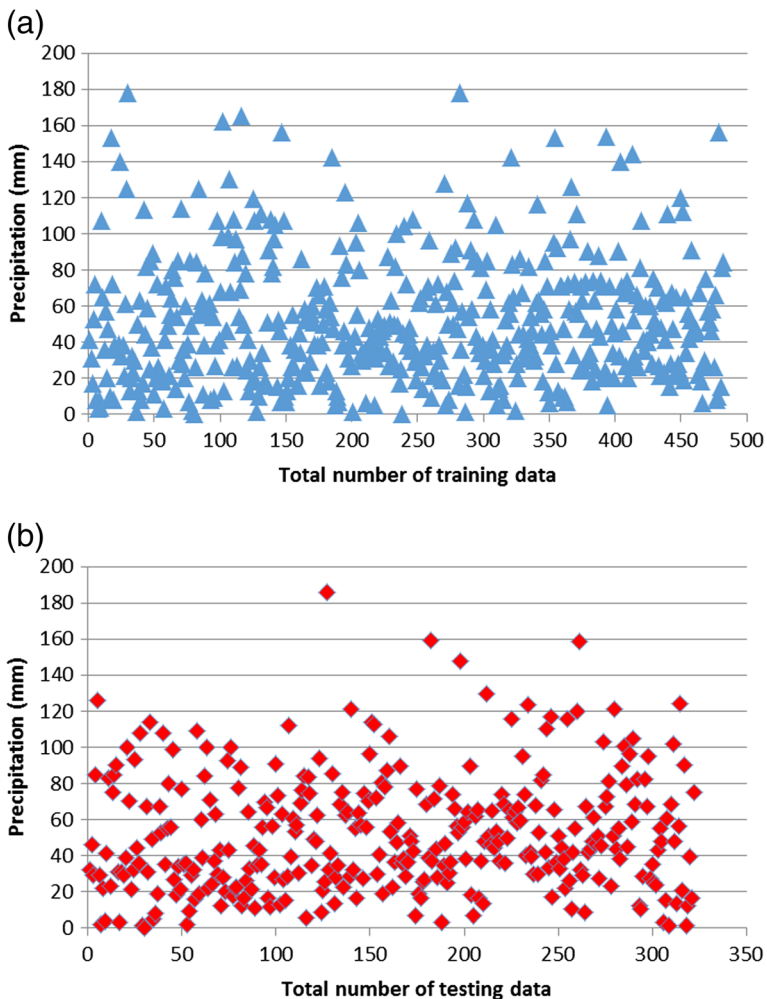


Fig. 2 Total number of **a** training data and **b** testing data for Serbia

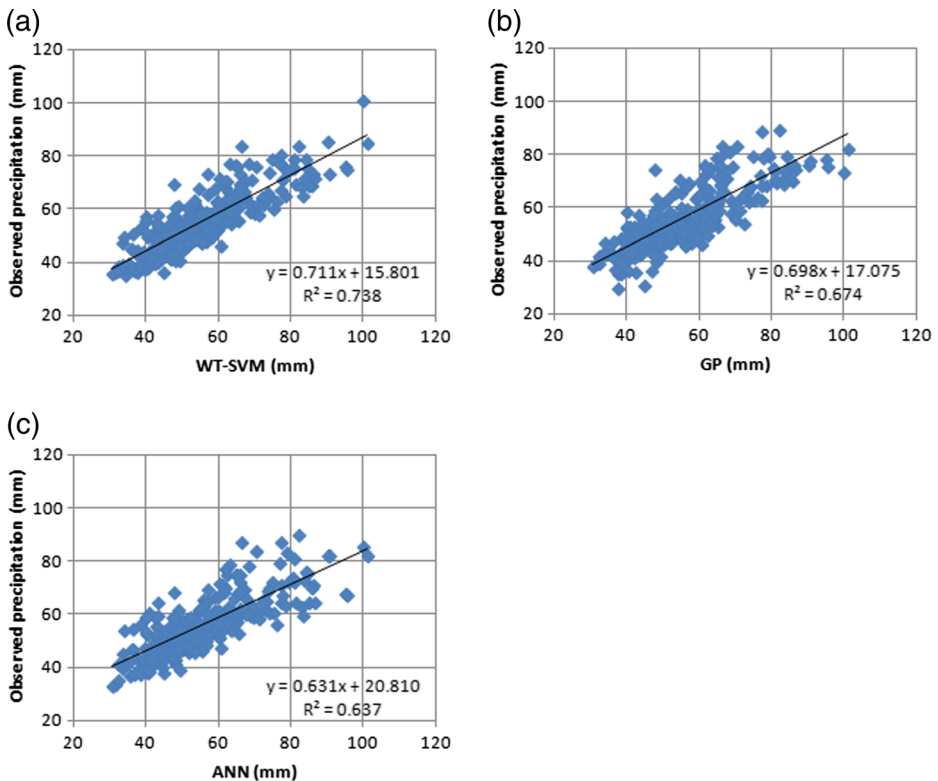
Table 1 Performance indices of three soft computing approaches for precipitation estimation in Serbia

Method	Statistical indicator				
	MAE	MAPE (%)	RMSE (mm)	r	R ²
GP	0.027	10.847	7.582	0.821	0.674
ANN	0.031	11.160	7.997	0.798	0.637
WT-SVM	0.015	9.314	6.778	0.859	0.738

MATLAB was used for developing the SVM model. For the SVM model, the three parameters of C , ε and γ were assigned as 2.1, 0.5 and 1.2, respectively.

The five statistical indicators i.e., MAE, MAPE, RMSE, r and R^2 were utilized to provide a comparison between the estimated and actual values of ANN and GP with WT-SVM. Table 1 summarizes the comparison results for precipitation estimation in Serbia. It is clearly found that the WT-SVM provides higher accuracy for precipitation estimation.

Figure 3 illustrates the scatter plots of the observed precipitation values against the estimated values by the WT-SVM, GP and ANN techniques. The results show that

**Fig. 3** Scatter plot of the a WT-SVM, b GP and c ANN precipitation estimation

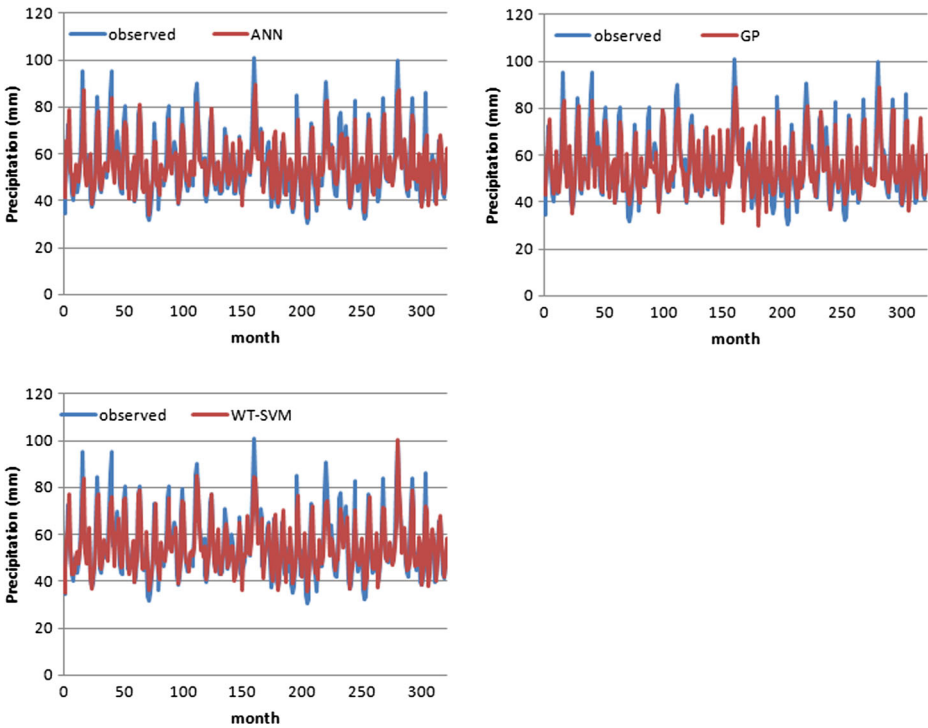


Fig. 4 Time series of monthly precipitation estimates of ANN, GP and WT-SVM in test period

there favorable agreements between the estimated values by WT-SVM and the observed values so that there are no points of significant overestimation or underestimation.

Figure 4 illustrates time series of precipitation estimates, while the Fig. 5 represents the residuals of the ANN, Gp and WT-SVM in test period. It can be noticed that the WT-SVM model offers higher accuracy compared to ANN and GP models. In addition, the WT-SVM predicts the maximum peak better than ANN and GP. The similar results have been represented in Kisi and Cimen (2012), Venkata Ramana et al. (2013) and Feng et al. (2015).

4 Conclusion

In this paper, three soft computing methods of WT-SVM, GP and ANN were used to estimate precipitation. The precipitation data were collected from 29 principal meteorological stations in Serbia for the period 1946–2012. The WT-SVM model was developed by combining the SVM with WT. The proposed method is unique; thus, it can boost the precision compared with the other earlier developed models. The performance of the simulation results were evaluated using five statistical parameters (MAE, MAPE, RMSE, r , R^2).

The results indicated that the WT-SVM model has a better accuracy in estimation of the precipitation compared to ANN and GP models. Therefore, the WT-SVM model can be assumed as an efficient and useful soft computing model to predict precipitation. The

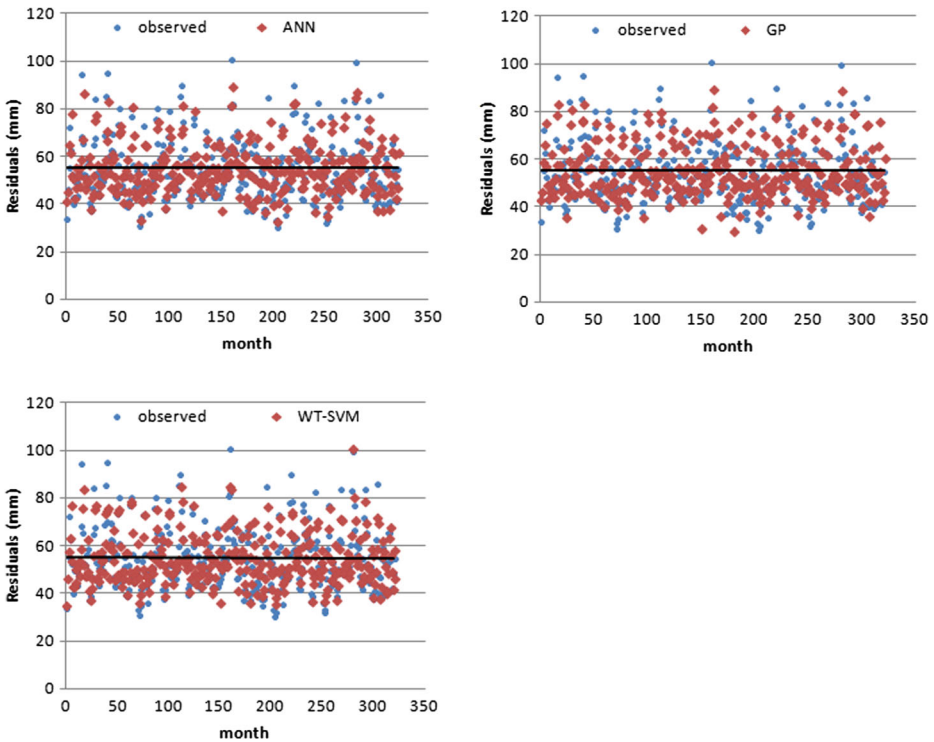


Fig. 5 Residuals of monthly precipitation estimates of ANN, GP and WT-SVM in test period

algorithm can be utilized also for other regions in future studies in order to prove the efficiency of the model for all weather conditions.

Acknowledgments This work is funded by the Malaysian Ministry of Higher Education under the University of Malaya High Impact Research Grant UM.C/625/1/HIR/MoHE/FCSIT/17, the Ministry of Education, Science and Technological Development, Republic of Serbia (Grant No. TR37003) and the ICT COST Action IC1408 Computationally-intensive methods for the robust analysis of non-standard data (CRoNoS).

References

- Abbot J, Marohasy J (2014) Input selection and optimisation for monthly rainfall forecasting in Queensland, Australia, using artificial neural networks. *Atmos Res* 138:166–178
- Adamowski J, Chan HF (2011) A wavelet neural network conjunction model for groundwater level forecasting. *J Hydrol* 407(1):28–40
- Chau K, Wu C (2010) A hybrid model coupled with singular spectrum analysis for daily rainfall prediction. *J Hydroinf* 12(4):458–473
- Dileep AD, Sekhar CC (2014) Class-specific GMM based intermediate matching kernel for classification of varying length patterns of long duration speech using support vector machines. *Speech Comm* 57:126–143
- Du S, Huang D, Lv J (2013) Recognition of concurrent control chart patterns using wavelet transform decomposition and multiclass support vector machines. *Comput Ind Eng* 66(4):683–695
- Feng Q, Wen X, Li J (2015) Wavelet analysis-support vector machine coupled models for monthly rainfall forecasting in arid regions. *Water Resour Manag* 29:1049–1065

- Freiwan M, Cigizoglu HK (2005) Prediction of total monthly rainfall in Jordan using feed forward backpropagation method. *Fresenius Environ Bull* 14(2):142–151
- Gocic M, Trajkovic S (2014a) Spatiotemporal characteristics of drought in Serbia. *J Hydrol* 510:110–123
- Gocic M, Trajkovic S (2014b) Spatio-temporal patterns of precipitation in Serbia. *Theor Appl Climatol* 117(3–4): 419–431
- Guo J, Yi P, Wang R, Ye Q, Zhao C (2014) Feature selection for least squares projection twin support vector machine. *Neurocomputing* 144:174–183
- Harris T (2015) Credit scoring using the clustered support vector machine. *Expert Syst Appl* 42(2): 741–750
- Jawerth B, Sweldens W (1994) An overview of wavelet based multiresolution analyses. *SIAM Rev* 36(3):377–412
- Kalteh AM (2013) Monthly river flow forecasting using artificial neural network and support vector regression models coupled with wavelet transform. *Comput Geosci* 54:1–8
- Kalteh AM (2015) Wavelet genetic algorithm-support vector regression (Wavelet GA-SVR) for monthly flow forecasting. *Water Resour Manag* 29(4):1283–1293
- Kisi O, Cimen M (2012) Precipitation forecasting by using wavelet-support vector machine conjunction model. *Eng Appl Artif Intel* 25(4):783–792
- Liang C-W, Juang C-F (2015) Moving object classification using local shape and HOG features in wavelet-transformed space with hierarchical SVM classifiers. *Appl Soft Comput* 28:483–497
- Lu C-J, Shao YE, Li P-H (2011) Mixture control chart patterns recognition using independent component analysis and support vector machine. *Neurocomputing* 74(11):1908–1914
- Mekanik F, Imteaz MA, Gato-Trinidad S, Elmahdi A (2013) Multiple regression and Artificial Neural Network for long-term rainfall forecasting using large scale climate modes. *J Hydrol* 503:11–21
- Mohammadi K, Shamsirband S, Tong CW, Arif M, Petkovic D, Ch S (2015) A new hybrid support vector machine-wavelet transform approach for estimation of horizontal global solar radiation. *Energy Convers Manag* 92:162–171
- Moustris KP, Larissi IK, Nastos PT, Paliatsos AG (2011) Precipitation forecast using artificial neural networks in specific regions of Greece. *Water Resour Manag* 25(8):1979–1993
- Nasseri M, Asghari K, Abedini MJ (2008) Optimized scenario for rainfall forecasting using genetic algorithm coupled with artificial neural network. *Expert Syst Appl* 35(3):1415–1421
- Nourani V, Hosseini Baghanam A, Adamowski J, Gebremichael M (2014) Using self-organizing maps and wavelet transforms for space-time pre-processing of satellite precipitation and runoff data in neural network based rainfall-runoff modeling. *J Hydrol* 476:228–243
- Ortiz-García EG, Salcedo-Sanz S, Casanova-Mateo C (2014) Accurate precipitation prediction with support vector classifiers: A study including novel predictive variables and observational data. *Atmos Res* 139:128–136
- Peng Z, Chu F (2004) Application of the wavelet transform in machine condition monitoring and fault diagnostics: a review with bibliography. *Mech Syst Signal Process* 18(2):199–221
- Sahay RR, Srivastava A (2014) Predicting monsoon floods in rivers embedding wavelet transform, genetic algorithm and neural network. *Water Resour Manag* 28:301–317
- Sánchez-Monedero J, Salcedo-Sanz S, Gutiérrez PA, Casanova-Mateo C, Hervás-Martínez C (2014) Simultaneous modelling of rainfall occurrence and amount using a hierarchical nominal-ordinal support vector classifier. *Eng Appl Artif Intel* 34:199–207
- Sedki A, Ouazar D, El Mazoudi E (2009) Evolving neural network using real coded genetic algorithm for daily rainfall-runoff forecasting. *Expert Syst Appl* 36(3):4523–4527
- Shamsirband S, Gocic M, Petković D, Saboohi H, Herawan T, Mat Kiah ML, Akib S (2014) Soft-Computing Methodologies for Precipitation Estimation: A Case Study. *IEEE J Sel Top Appl Earth Obs Remote Sens*. doi:10.1109/JSTARS.2014.2364075
- Shiau JT, Huang CY (2014) Detecting multi-purpose reservoir operation induced time-frequency alteration using wavelet transform. *Water Resour Manag* 28:3577–3590
- Sun Y, Leng B, Guan W (2015) A novel wavelet-SVM short-time passenger flow prediction in Beijing subway system. *Neurocomputing* 166:109–121
- Tripathi S, Srinivas VV, Nanjundiah RS (2006) Downscaling of precipitation for climate change scenarios: a support vector machine approach. *J Hydrol* 330(3–4):621–640
- Valverde Ramírez MC, Campos Velho HF, Ferreira NJ (2005) Artificial neural network technique for rainfall forecasting applied to the São Paulo region. *J Hydrol* 301(1–4):146–162
- Vapnik VN (1995) *The nature of statistical learning theory*. Springer, New York
- Vapnik VN (1998) *Statistical learning theory*. Wiley, New York
- Venkata Ramana R, Krishna B, Kumar SR, Pandey NG (2013) Monthly rainfall prediction using wavelet neural network analysis. *Water Resour Manag* 27:3697–3711

- Wang W, Li Y (2011) Wavelet transform method for synthetic generation of daily stream flow. *Water Resour Manag* 25:41–57
- Wu CL, Chau KW, Fan C (2010) Prediction of rainfall time series using modular artificial neural networks coupled with data-preprocessing techniques. *J Hydrol* 389(1–2):146–167
- Yang X, Tan L, He L (2014) A robust least squares support vector machine for regression and classification with noise. *Neurocomputing* 140:41–52
- Yong DD, Bhowmik S, Magnago F (2015) An effective power quality classifier using wavelet transform and support vector machines. *Expert Syst Appl* 42(15–16):6075–6081