University of Birjand

Water Harvesting Research Vol. 6, No. 1, Spring & Summer 2023, Original Paper p. 145-156

10.22077/JWHR.2024.7040.1120

Investigating the Effect of Wavelet Decomposition on the Performance of the Optimized Support Vector Regression in Precipitation Simulation

Majid Rastgary[a], Emad Mahjoobi[b*], Samad Emamgholizadeh[c]

[a]PhD student in Water Resources Management and Engineering, Faculty of Civil Engineering, Shahrood University of Technology, Shahrood, Iran.
[b]Assistant Professor, Department of Water and Environmental Engineering, Faculty of Civil Engineering, Shahrood University of Technology, Shahrood, Iran.
[c]Professor, Department of Water and Environmental Engineering, Faculty of Civil Engineering, Shahrood University of Technology, Shahrood, Iran.

[*]Corresponding Author, E-mail address: emahjoobi@shahroodut.ac.ir
Received: 13 October 2023/ **Revised**: 09 December 2023/ **Accepted**: 21 December 2023

Abstract
Considering the climatic changes and the increase of extreme values in recent years, in this study, the effect of time series decomposition based on wavelet transform in improving the performance of the optimized support vector regression model in the simulation of annual precipitation in Dashband and Tapik stations has been discussed and investigated in the Lake Urmia Basin in the period of 1971-2020. In this study, the Ant colony algorithm was used to optimize the parameters of the support vector regression model. Daubechies 4 wavelet with three decomposition levels 1, 2 and 3 was used to decomposition the time series of precipitation in the studied stations. The SVR model takes in annual precipitation data as input, while the decomposition-based models take in decomposed precipitation values. The results of investigation the error rate and efficiency of the 4 investigated models include optimized SVR, W1-SVR (optimized SVR based on level 1 decomposition), W2-SVR (optimized SVR based on level 2 decomposition) and W3-SVR (optimized SVR based on level 3 analysis) showed that the error rate of all 4 mentioned models is acceptable and the observed values are in the 95% confidence interval. The error rate of 5.20 and 6.68 mm in the simulation of precipitation in Dashband and Tapik stations using the optimized SVR model by time series decomposition based on wavelet theory in level 1 decomposition in the mentioned stations, 31 and 35 percent improvement has been found. The level 2 decomposition of the time series of precipitation obtained the lowest error among the different levels of decomposition, which was 3.42 and 3.26 mm in Dashband and Tapik stations, respectively. Considering the increase in simulation complexity with the involvement of wavelet theory, the error rate improvement and model performance are acceptable. The hybrid W-SVR model in this study provides reliable results for precipitation simulation. Analyzing the annual precipitation series makes it possible to develop the dimensions of the optimized SVR model.
Keywords: Ant Colony Algorithm, Daubechies, Decomposition, Wavelet Transform.

1. Introduction
Estimating and predicting rainfall and reaching the amount of runoff resulting from it, play a fundamental and effective role in the management and proper use of the basin, management of dams and reservoirs, minimizing damages caused by floods, droughts and water resources management. For this reason, it is of interest to hydrologists. The prediction of any event forms the basis of its crisis management, and this possibility is achieved when suitable forecasting models are available. Various methods are used to predict hydrological events, including rainfall. The results of using each of these methods are always accompanied by some error. Accurate forecasting of hydrological signals such as rainfall can provide useful information for the purpose of forecasting the amount of rainfall and managing water and soil resources in a basin. In addition, correct forecasting of hydrological signals plays an important role in

reducing the effects of drought on water resources systems (Venugopal and Foufoula-Georgiou, 1996; Kim et al., 2016; Karami and Dariane, 2017; Mendoza et al., 2019; Jamei et al., 2023).

Despite the highly stochastic nature of meteorological and hydrological processes, the development of models capable of describing such complex phenomena is a growing research area. Providing insight into the modeling of complex phenomena through a thorough literature review, current research, and expanding research horizons can increase the potential of accurate and well-designed models. Recently, integrated wavelet analysis has attracted increasing attention. Wavelet analysis has become a popular analytical technique due to its ability to simultaneously reveal spectral and temporal information in a signal (Wei et al., 2013). The use of wavelet analysis in meteorology and hydrology is a new topic and some researchers have started using it. Jayawardena et al. (2004) also used wavelet decomposition combined with Markov model to simulate the daily rainfall of Chao Phraya watershed in Thailand. Cannas et al. (2006) used wavelet transform to evaluate the effects of data processing in simulating the river flow using neural networks. They confirmed the accuracy and certainty of wavelet transform model.

Beecha and Chowdhury (2010) investigated rainfall characteristics and its distribution over the period 1925–2002 in Melbourne, Australia. They identified wavelet spectra of different rainfall intensities that were observed periodically in a 5 to 10 year cycle sequence. They stated that this phenomenon can be affected by the annual variability of El Niño Southern Oscillation. Asadi et al. (2013) presented a new combination of artificial neural network for rainfall-runoff modeling in Agh-Chai basin, Iran. The proposed model was a combination of data processing methods, genetic algorithms and Levenberg–Marquardt algorithm for neural network input training. The results showed that this method is more accurate than ANN and ANFIS in rainfall-runoff predicting.

Ouyang et al. (2016) used support vector regression and ensemble empirical mode decomposition to build a rainfall forecasting model. The proposed hybrid model was used to forecast monthly rainfall at a meteorological station in Changchun, China. The results of their research showed that the proposed hybrid model has the lowest NMSE and MAPE values of 0.10 and 14.90, respectively, and the highest NSE and CC values of 0.91 and 0.83, respectively, during the validation period. Zakhrouf et al. (2018) developed an approach based on wavelet transform and genetic algorithm to simulate the daily river flow in northern Algeria. Their results showed that the performance of mentioned approach was better than conventional models. Chong et al. (2020) predicted rainfall over the Langat river basin through the integration of wavelet transform and convolutional neural network. The results of their research showed that the proposed model can satisfactorily portray the rainfall time series patterns for both monthly rainfall forecast and daily rainfall forecast.

Wang et al. (2021) investigated the application of several wavelet decomposition-based forecasting models in annual rainfall forecasting, and proposed a new hybrid rainfall forecasting framework that couples extreme machine learning and wavelet packet decomposition. They stated that wavelet packet decomposition is used to decompose the main precipitation data into several sub-layers. They also stated that the extreme learning machine model, the autoregressive integrated moving average model, and the back-propagation neural network are used to realize forecasting calculations for the decomposed series. Also, the results of their research showed that the integrated model of extreme learning machine and wavelet packet decomposition works better than other models used in their research, and wavelet packet decomposition can significantly increase the performance of prediction models. Vivas et al. (2023) dealt with rainfall forecasting by proposing to use a method based on a combination of techniques in the framework of lagged regression model and wavelet decomposition principles. They implemented wavelet decomposition in a preprocessing step followed by the use of a long short-term memory network (LSTM) and proposed a forecast improvement step where the outputs are optimized by algorithms for monthly precipitation forecast corrections. The results of their research showed that the correction of rainfall forecasting biases provides the achievement of adjusted coefficients of determination greater than 0.76 and values of the normalized mean absolute error (NMAE) less than 0.31. Thomas et al. (2023) investigated the annual precipitation cycles in Kerala, India during the period of 1871-2016 using Fourier and wavelet techniques. They stated that the results of Fourier analysis and wavelet analysis are consistent. The wavelet power in two time series shows the common features in 8 to 16 years with different importance, which shows the relationship between them.

Investigation the literature riveiw shows that the use of values decomposed by the wavelet function can increase the accuracy of the simulations. Considering the climatic changes and the increase of the extreme values in recent years, this study tries to investigate the effect of decomposition the observation series in simulating the precipitation values. The main goal of this research is to evaluate the optimized support vector regression method in two modes of decomposition of observational values and no decomposition of observational values based on wavelet function in three levels 1, 2 and 3.

2. Materials and Methods

2.1. Study area

In this study, the precipitation values of Dashband and Tapik stations in Lake Urmia Basin located in the northwest of Iran have been used. The amounts of precipitation (mm) are on an annual scale in the period of 1971-2020. Figure 1 shows the study area and the location of selected stations. The statistical characteristics of the studied data were presented in Table 1.

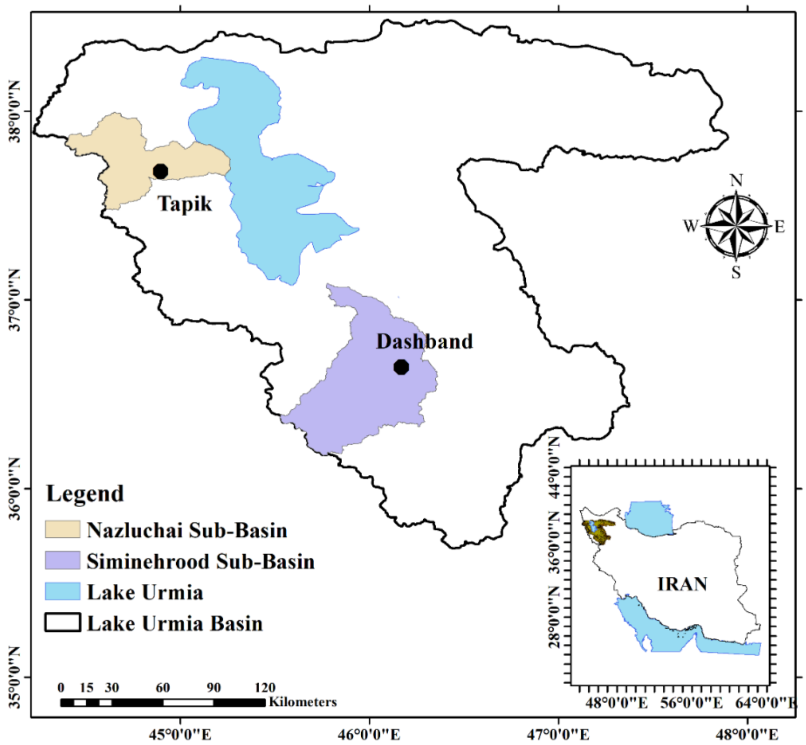


Fig. 1. The location of the studied sub-basins in Lake Urmia Basin and Iran.

Table 1. Statistical characteristics of the studied data

Station	Mean Annual (mm)	Max (mm)	Min (mm)	STD	Skew
Tapik	373.2	708.5	157.5	139.7	0.5
Dashband	394.3	920.4	152.3	156.8	1.3

2.2. Wavelet transform

Wavelet theory is a method in mathematics derived from Fourier's theory, which was proposed in the 19th century, but has been used for about a decade (Nazeri Tahroudi and Mirabbasi Najafabadi, 2023). Wavelet transform is an efficient mathematical transform in the field of signal processing. Wavelets are mathematical functions that represent the time scale shape of time series and their functions for analyzing time series that include non-stationary variables. Wavelet analysis provides long time intervals for low frequency information and shorter time intervals for higher frequency information. Wavelet analysis is able to show different aspects of data, breakpoints and discontinuities that other signal analysis methods may not be able to show. The wavelet function has two important characteristics of fluctuation and shortness. $\psi(x)$ is a wavelet function, if and only if the Fourier transform $\psi(x)$ satisfiesthe following condition (Mallat, 1989):

$$\int_{-\infty}^{+\infty} \frac{|\psi(x)|}{|\omega|^2} d\omega \quad < +\infty \tag{1}$$

This condition is known as the wavelet acceptance condition. The above equation can be considered equivalent to the following equation:

$$\psi(0) = \int_{-\infty}^{+\infty} \psi(x) dx = 0 \tag{2}$$

This characteristic of functions, which has a mean equal to zero, is not a strict filter, and many functions can be considered as wavelet functions based on it. $\psi(x)$ is the mother wavelet function is used in the analysis by two mathematical methods called translation and scale, which causes a change in the size and location of the analyzed signal.

$$\psi_{a,b}(x) = \frac{1}{\sqrt{a}} \psi\left(\frac{x-b}{a}\right) \tag{3}$$

Finally, the wavelet coefficient can be calculated at each signal point (b) and for each scale value (a) with the following equation (Mallat, 1989):

$$T(a,b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{+\infty} \psi\left(\frac{t-b}{a}\right) f(t) dt \tag{4}$$

In the above equation, *a* is equal to scale and *b* is equal to transformation. For the value of *T*, different values of *a* and *b* are obtained. Whenever *T* has the highest positive value, the most adjustment occurs. There is no adjustment for *T* equal to zero, and for a negative value of *T* the adjustment is reversed or the largest difference occurs. There are different types of wavelet functions and depending on their application, their accuracy is different. Depending on their application, wavelet functions have different types with different levels of accuracy. The most common wavelet functions are followed as Haar wavelet function. Harr wavelet function is simpler and one of the first wavelets. Daubechies wavelet function is one of the most efficient wavelet functions in detecting local discontinuities in signals. The Symlet wavelet function has properties similar to the Daubechies family. Sym6 and Sym4 functions are almost symmetrical and are used in damage detection. Other wavelets are Gaussian, Morlet, Meyer, Coif, Mexicanhat and Bior (Solgi et al., 2017; Nazeri Tahroudi and Mirabbasi, 2023a&b). One of the important key points in choosing mother wavelets is the nature and type of time series. Therefore, the patterns of mother wavelet functions, which can be geometrically adapted to the time series curve, will perform a better adaptation and the obtained results will be better.

2.3. Support vector regression

In the regression type of SVM model, a function related to the dependent variable *Y*, which is itself a function of several independent variables *x*, is estimated. Similar to other regression problems, it is assumed that the relationship between the independent and dependent variables is determined by an algebraic function such as *f(x)* plus a disturbance value (permissible error ε) (Eq.6) (Nazeri Tahroudi and Ramezani, 2020).

$$f(x) = W^T . \phi(x) + b \tag{5}$$

$$y = f(x) + noise \tag{6}$$

If *W* (vector of coefficients) and *b* (constant) are the characteristics of the regression function and ϕ is also the kernel function, then the goal is to find a functional form for *f(x)*. This is achieved by training the SVM model by a set of samples (training set). Therefore, to calculate *w* and *b*, it is necessary to optimize the error function in the *ε-SVM* model by considering the conditions listed in Eq.8.

$$\frac{1}{2} W^T W + C \sum_{i=1}^{N} \xi_i + C \sum_{i=1}^{N} \xi_i^* \tag{7}$$

$$W^T.\phi(x_i)+b-y_i \le \varepsilon+\xi_i^*$$
$$y_i - W^T.\phi(x_i)-b \le \varepsilon+\xi_i \quad (8)$$
$$\xi_i, \xi_i^* \ge 0 \quad , \quad i=1,...,N$$

In the above equations, C is a positive integer that determines the penalty when a model training error occurs. ϕ is Kernel function, N is the number of samples and two characteristics ξ_i and ξ_i^* are Slack Variable which determine the upper and lower bounds of the training error related to the allowed error value ε. In the problems, it is expected that the data will be placed within the boundary interval ε. Now, if a data is outside the range of ε, then there will be an equivalent error ξ_i and ξ_i^*. By introducing 2 Lagrange coefficients a_i and a_i^* the optimization problem will be solved by numerical maximization of the following quadratic function (Eslami et al., 2022).

$$\sum_{i=1}^{N} y_i(a_i+a_i^*)-\varepsilon\sum_{i=1}^{N}(a_i+a_i^*)- \quad (9)$$
$$0.5\sum_{i,j=1}^{N}(a_i+a_i^*)(a_j+a_j^*)\phi(x_i)^T\phi(x_j)$$

$$\sum_{i=1}^{N}(a_i+a_i^*)=0 \quad (10)$$
$$0\le a_i \le C, \ 0\le a_i^* \le C, \ i=1,2,...,N$$

After defining the Lagrange coefficients and the parameters w and b in the SVM regression model, it is calculated using the Karush–Kuhn–Tucker (KKT) conditions, in which $W=\sum_{j=1}^{N}(a_i+a_i^*)\phi(x_i)$. As a result, we will have a regression SVM model (Eslami et al., 2022):

$$W=\sum_{i=1}^{N}(a_i+a_i^*)\phi(x_i)^T\phi(\mathrm{x})+b \quad (11)$$

It should be noted that the Lagrange terms $(a_i+a_i^*)$ can be zero or non-zero. Therefore, only data sets whose coefficients are non-zero $\overline{a_i}$ are included in the final regression equation, and these data sets are known as support vectors. Simply, the support vectors are the data that help to build the regression function. Among the mentioned vectors, those whose value $|\overline{a_i}|$ is less than C are called Margin Support Vector. When the value $|\overline{a_i}|$ of the support vectors is equal to the value of C, it is known as Error Support Vector. Marginal support vectors are located at the edge of the insensitive boundary, while error support vectors are outside the interval. Finally, the regression SVM function can be rewritten in the following form:

$$f(x)=\sum_{i=1}^{N}a_i\phi(x_i)^T\phi(x_j)+b \quad (12)$$

In Eq.12, the calculation of $\phi(x)$ in its characteristic space may be very complicated. To solve this problem, the usual process in the regression SVM model is to choose a kernel function like $K(x_i,x)=\phi(x_i)^T\phi\sqrt{b^2-4ac}$. Different kernel functions can be used to build different types of *ε-SVM* model (Nazeri Tahroudi et al., 2018).

2.4. Ant colony algorithm (ACO)

Ant colony algorithm was suggested in 1991 by Colorni et al. (1991). One of the first applications of the ACO algorithm was to solve the traveling salesman problem. Since ACO algorithms depend on the type of use and the similarity of the ants' movement on the graph, using the traveling salesman problem to explain the basic principles of ant algorithms is very logical and was originally used as a type example to introduce this algorithm. For more explanations in this regard, see Colorni et al. (1991).

2.5. Model evaluation criteria

To select the best fitting model, root mean square error (RMSE) (Eq.13) and Nash-Sutcliffe efficiency (NSE) (Eq.14) were calculated. Any fitting model that has the highest value of NSE and the lowest RMSE is selected as the appropriate model.

$$RMSE=\sqrt{\frac{\sum_{i=1}^{n}(S_i-O_i)^2}{n-1}} \quad (13)$$

$$NSE = 1 - \frac{\sum_{i=1}^{n}(S_i - O_i)^2}{\sum_{i=1}^{n}(O_i - \bar{O}_i)^2} \quad (14)$$

In the above equations, n is the number of data, O_i is the observation values, $\bar{O}_i$ is the mean values of observation series and S_i is simulated value of precipitation (Nash and Sutcliffe, 1979; Pronoos Sedighi et al., 2023).

3. Results and Discussion

In this study, two models of optimized support vector regression (SVR) and optimized support vector regression based on wavelet theory (W-SVR) have been used to simulate precipitation in Tapik and Dashband stations located in the Lake Urmia Basin, northwest of Iran. At first, simulation of precipitation values was done using the optimized SVR model, and finally, by decomposition of the time series of precipitation by wavelet theory, simulation of precipitation was done using the W-SVR model.

3.1. The results of the simulation of precipitation of the studied stations using the SVR model and the ant colony optimizer algorithm

In this section, the optimization of parameters of SVR model was first investigated using the ant colony optimizer algorithm. The results of the investigation of the optimal parameters of the SVR model with 150 itterations presented as table 2.

Table 2. The results of optimal parameters of the SVR model in the simulation of precipitation in the study area

Station	ε	c	σ
Dashband	0.93	923.12	0.01
Tapik	0.98	927.23	0.01

In the next step, using the optimized parameters (table 2), the SVR model was implemented to simulate the precipitation values in the studied stations. The results of investigation and simulation of precipitation values in the studied stations were presented in figures 2 and 3. According to Figure 2-SVR, it can be seen that the confidence intervals of 95% of the simulation has completely covered the simulated values. The error value of 5.20 mm has also been obtained according to the optimized SVR model, which indicates an acceptable amount of error regarding the simulation of precipitation on an annual scale. The efficiency of the optimized SVR model in the simulation of precipitation amounts on an annual scale is also estimated to be 0.99, which shows an acceptable efficiency. According to Figure 3-SVR, it can also be seen that the amount of error in the simulation of precipitation values in Tapik station is more than that of Dashband station and is equal to 7.37mm. The 95% confidence interval also confirms the accuracy and efficiency of the optimized SVR model in precipitation simulation.

3.2. The results of simulation of precipitation at the studied stations using the optimized SVR model and wavelet theory

In order to investigate and evaluate the W-SVR model in the simulation of precipitation values in Dashband and Tapik stations in the period of 1971-2020, the wavelet function and the optimized SVR model were used. The simulations have been done using the Daubechies 4 wavelet, which is recommended by various researchers (Ahmadi and Maddah, 2021; Nekoeeyanet al., 2022; Nazeri Tahroudi and Mirabbasi Najafabad, 2023; Nazeri Tahroudi and Mirabbasi, 2023b). Three levels of decomposition were also considered regarding Daubechies 4 wavelet (level 1, level 2 and level 3) in the simulation of precipitation values in the studied stations. At first, the decomposition of the time series of precipitation was done in the studied stations at different levels, and the results were presented according to Figures 4 and 5 for Dashband and Tapik stations, respectively.

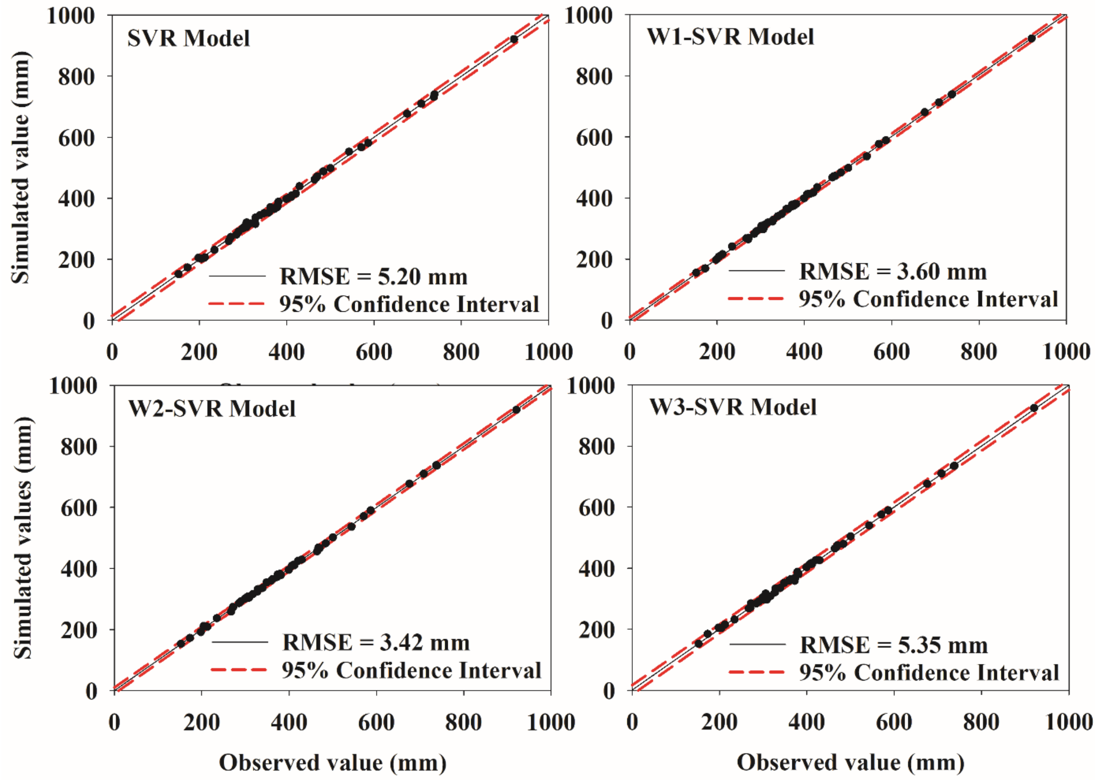


Fig. 2. The results of simulation of precipitation values in Dashband station in the period of 1971-2020 using SVR and W-SVR models.

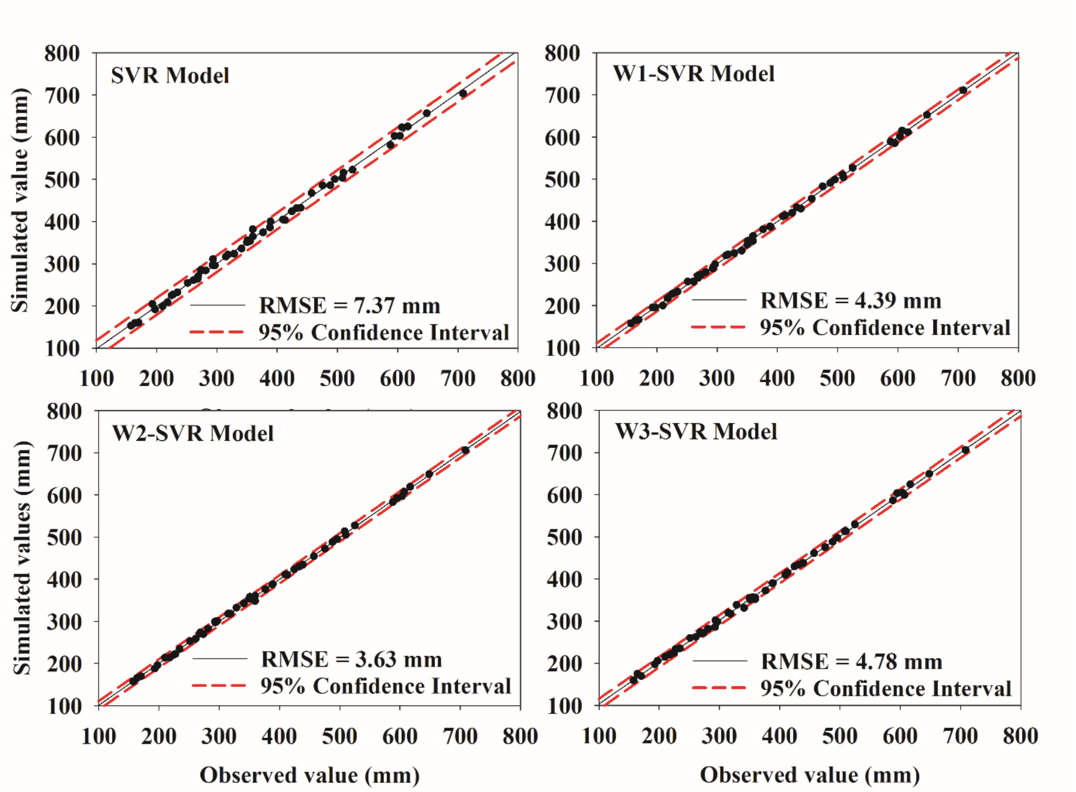
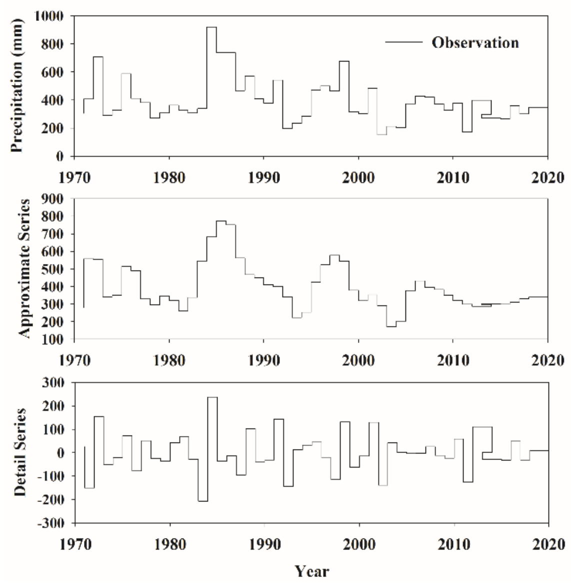
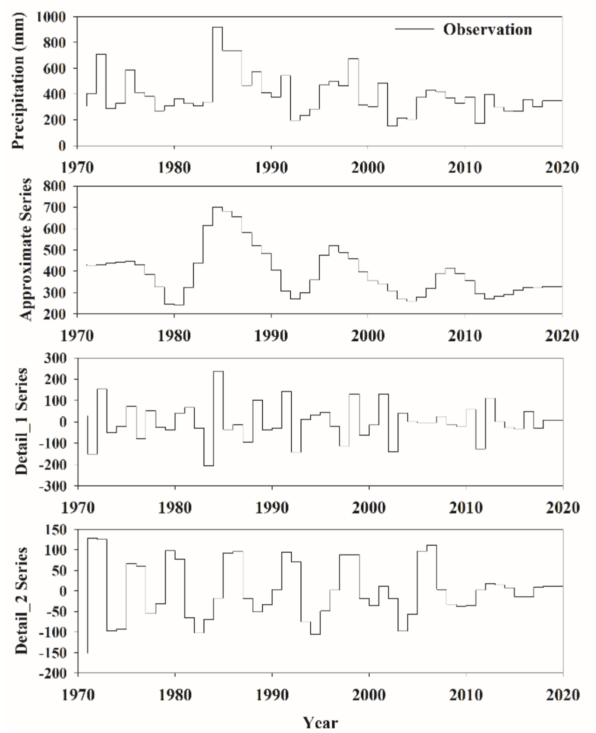


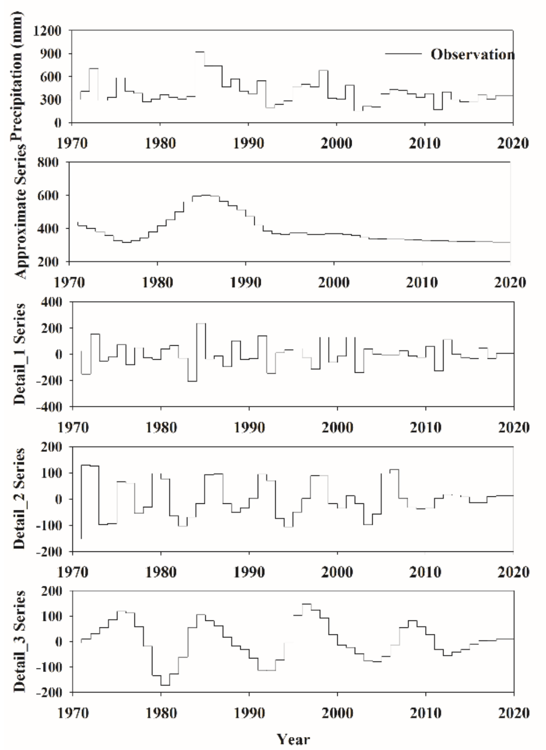
Fig. 3. The results of simulation of precipitation values in Tapik station in the period of 1971-2020 using SVR and W-SVR models.



Level 1 Decomposition

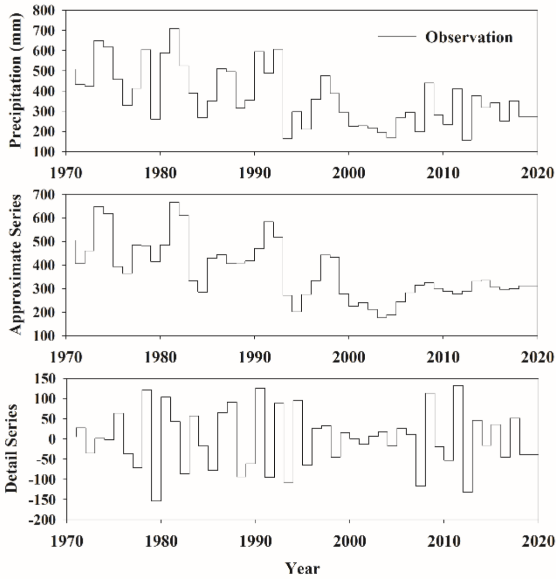


Level 2 Decomposition

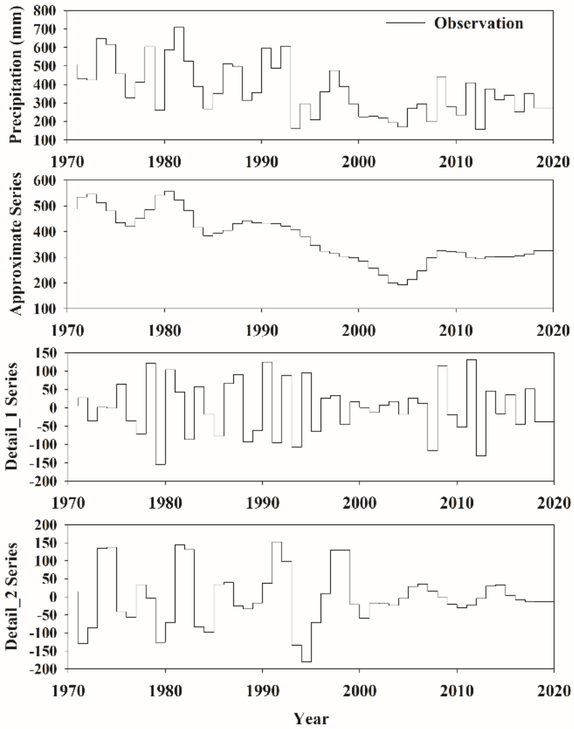


Level 3 Decomposition

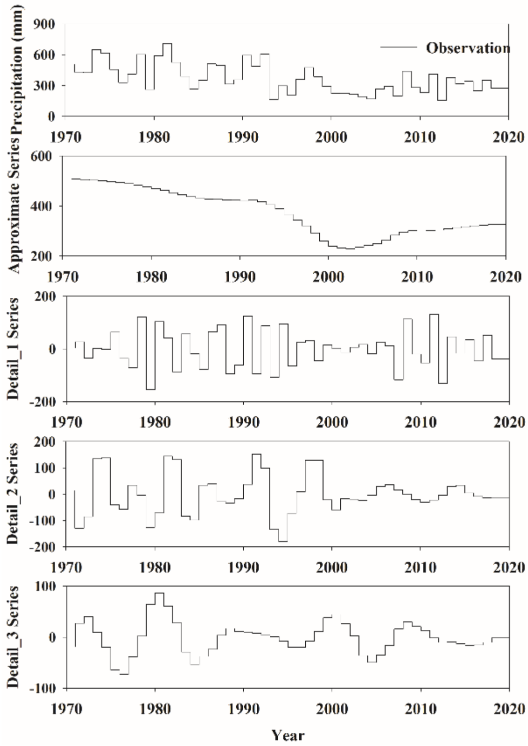
Fig. 4. The results of decomposition of precipitation values using wavelet function and different levels of decomposition at Dashband station



Level 1 Decomposition



Level 2 Decomposition



Level 3 Decomposition

Fig. 5. The results of decomposition of precipitation values using wavelet function and different levels of decomposition at Tapik station

The performance of the wavelet theory in this field is such that the sum of Approximate and Detail at each level of analysis will be equal to the observed values. As the level of decomposition increases, more details about the observed values appear. One of the most important applications of the wavelet theory in the decomposition of different levels of time series will be the increase in the dimension of simulations, which increases the possibility of using multivariable models, and as a result, multivariable models can be used, and also considered and examined multivariate structures. According to Figures 4 and 5, it can be seen that with the increase in the levels of decomposition, the dimensions of the simulations also increase. At this stage, the decomposed values in each level, which includes Approximate and Detail in level 1, Approximate and Detail 1 and 2 in level 2, and Approximate and Detail 1, 2 and 3 in level 3, are individually simulated by using optimized the SVR model and finally at each level of decomposition, the simulated values are compared with the observed values. The results of checking the error statistics and efficiency of the W-SVR model in simulating the precipitation values in the studied stations are presented in Figures 2 and 3. According to Figures 2 and 3, it can be seen that the simulated values of precipitation in both studied stations are within the confidence range of the simulation, which shows the proper performance of the W-SVR model in all three levels of decompositions: 1, 2 and 3. The efficiency of more than 99% is observed in the simulation of precipitation amounts at all levels of decompositions and also at both studied stations.

According to Figure 2-W1-SVR, it can be seen that level 1 decomposition of a wavelet theory using the optimized support vector regression method has been able to reduce the amount of precipitation simulation error in Dashband station by 1.6 mm compared to the optimized SVR model. The W2-SVR hybrid model, which shows the level 2 decomposition of the wavelet theory in the simulation of precipitation amounts at the Dashband station, also succeeded in reducing the RMSE statistic by about 1.78 mm compared to the optimized SVR model (W2-2 -SVR). According to Figure 2-W3-SVR, the simulated precipitation values at Dashband station in the period of 1971-2020 using the optimized support vector regression model and level 3 decomposition of the wavelet theory, compared to the observed values, the error rate is higher than the optimized SVR model without decomposition. In general, at Dashband station, the results of examining the error rate of the 4 studied models in the simulation of precipitation values on an annual scale showed that the level 3 decomposition of the precipitation time series has caused an increase in the error rate. Among the 4 studied models, the lowest amount of error is related to the optimized SVR model with the level 2 decomposition and the highest amount of error is related to the optimized SVR model with the level 3 decomposition. According to Figure 3, it can be seen that there is a good correlation between the observed and simulated values regarding the precipitation values in Tapik station. The optimized SVR-based decomposition model (W-SVR) in all three decomposition levels at Tapik station as well as Dashband station was able to provide high efficiency in precipitation simulation. According to Figure 3-W1-SVR, the amount of precipitation simulation error in level 1 decomposition is estimated to be 4.39 mm on an annual scale, which has reduced the amount of error by about 3 mm compared to the base model (optimized SVR). By increasing the level decomposition of the precipitation series to level 2, the error rate of the W2-SVR model in simulating the annual precipitation of Tapik station has decreased compared to the basic SVR model and also the W1-SVR model and has reached 3.63 mm.

Again, by increasing the decomposition level to level 3, the error rate of the W3-SVR model in simulating the annual precipitation of Tapik station has increased compared to levels 2 and 1 decompositions and has become equal to 4.78 mm. The results of checking the accuracy and effectiveness of the decomposition-based W-SVR model in simulating the precipitation at Tapik station showed that with the increase of decomposition level to level 3, the amount of simulation error increases compared to levels 1 and 2, but still compared to the optimized SVR model has a lower error rate. In general, according to the level of decomposition used, the Daubechies 4 wavelet function performed well in simulating precipitation in the study area, which is consistent with the studies of Ahmadi and Maddah (2021) and Nekoeeyan et al. (2022). Non-stationary behavior of time series, coupled with the implementation of conditional variance models, can enhance model accuracy and efficiency which can be considered in future studies (Kousali et al., 2022).

According to the presented figures, it is possible to see the better performance of the W-SVR model in the simulation of annual precipitation values. This model was able to increase the accuracy and efficiency of the simulation results by decomposing the time series and increasing the calculation dimension.

In addition to the RMSE statistic, in order to check the effectiveness of the studied models, the NSE statistic was also examined and presented in Table 3. The analysis of NSE values in simulating precipitation in the area indicates that the examined models are acceptable both with and without time series decomposition.

Table 3. The results of NSE statistics in the simulation of precipitation in the study area

Station	SVR	W1-SVR	W2-SVR	W3-SVR
Dashband	0.998	0.999	0.999	0.998
Tapik	0.997	0.999	0.999	0.998

4. Conclusion

In this study, in order to simulate annual rainfall in Dashband and Tapik stations located in Lake Urmia Basin, northwest of Iran, two optimized support vector regression models and hybrid W-SVR model (optimized support vector regression model based on wavelet function) was used. To implement the optimized support vector regression model, the Ant colony algorithm was used to optimize the parameters of the SVR model. By running the model with 150 itterations, the optimal parameters were estimated for the two studied stations. By simulating precipitation values on an annual scale using the optimized SVR model, the results showed that the optimized model in univariate mode was able to create a high correlation (R^2=0.99) between the observed and simulated values. In the next step, using level 1, 2 and 3 decompositions of observational values using wavelet theory, the time series of precipitation in the studied stations were decomposed into Approximate and Details values according to the level decomposition, which formed multivariate series. In level 1, 2 and 3 decomposition, simulations were performed in 3, 4 and 5 dimensions, respectively. By applying the optimized support vector regression method on decomposed data, W1-SVR, W2-SVR and W3-SVR models were created for level 1, 2 and 3 decomposition, respectively. The simulation results of the annual rainfall values in the studied stations showed that the W1-SVR model compared to the optimized SVR model in one-dimensional mode was able to reduce the error values in Tapik and Dashband stations by 35% and 31%, respectively. Compared to the optimized SVR model, the W2-SVR hybrid model was able to reduce the annual precipitation simulation error in Tapik and Dashband stations by 51% and 34%, respectively. A reduction of more than 50% in the annual precipitation simulation shows the positive effect of time series decomposition in the simulation. Also, the results of examining the error rate of W3-SVR hybrid model in the simulation of annual precipitation values in the studied stations showed that increasing the level of time series decomposition to more than 2 in the case of Dashband station was not satisfactory and increased the error rate. However, in the case of Tapik station, increasing the decomposition level to level 2 has also increased the performance and reduced the amount of precipitation simulation error compared to the optimized SVR model. However, the performance of the W3-SVR model in simulating the annual precipitation amounts in both stations is satisfactory. One of the factors that can cause the error to increase with the increase of the level of decomposition is the coefficient of variation of precipitation values in the studied stations ($CV_{Dashband}$=40% and CV_{Tapik}=36%) which increases the error resulting from the time series decomposition. But what can be mentioned as a general conclusion for this research is the superiority of level 2 analysis and W3-SVR hybrid model in simulating annual precipitation in the studied stations. The obtained results can be used in the existing decisions in the field of meteorology.

5. Disclosure Statement

No potential conflict of interest was reported by the authors.

6. References

Ahmadi, F., & Maddah, M. A. (2021). Development of Wavelet-Kstar Algorithm Hybrid

Model for the Monthly Precipitation Prediction (Case Study: Synoptic Station of Ahvaz). *Iranian Journal of Soil and Water Research*, *52*(2), 409-420.

Asadi, S., Shahrabi, J., Abbaszadeh, P., & Tabanmehr, S. (2013). A new hybrid artificial neural networks for rainfall–runoff process modeling. *Neurocomputing*, *121*, 470-480.

Cannas, B., Fanni, A., See, L., & Sias, G. (2006). Data preprocessing for river flow forecasting using neural networks: wavelet transforms and data partitioning. *Physics and Chemistry of the Earth, Parts A/B/C*, *31*(18), 1164-1171.

Colorni, A., Dorigo, M., & Maniezzo, V. (1991). Ant System: An autocatalytic optimizing process. Tech. Report 91-016, Politecnico di Milano, Italy.

Eslami, P., Nasirian, A., Akbarpour, A., & Nazeri Tahroudi, M. (2022). Groundwater estimation of Ghayen plain with regression-based and hybrid time series models. *Paddy and Water Environment*, *20*(3), 429-440.

Jamei, M., Ali, M., Malik, A., Karbasi, M., Rai, P., & Yaseen, Z. M. (2023). Development of a TVF-EMD-based multi-decomposition technique integrated with Encoder-Decoder-Bidirectional-LSTM for monthly rainfall forecasting. *Journal of Hydrology*, *617*, 129105.

Karami, F., & Dariane, A. B. (2017). Optimizing signal decomposition techniques in artificial neural network-based rainfall-runoff model. *International Journal of River Basin Management*, *15*(1), 1-8.

Kim, S., Kisi, O., Seo, Y., Singh, V. P., & Lee, C. J. (2016). Assessment of rainfall aggregation and disaggregation using data-driven models and wavelet decomposition. *Hydrology Research*, *48*(1), 99-116.

Kousali, M., Salarijazi, M., & Ghorbani, K. (2022). Estimation of non-stationary behavior in annual and seasonal surface freshwater volume discharged into the Gorgan Bay, Iran. *Natural Resources Research*, *31*(2), 835-847.

Mendoza, D. E., Samaniego, E. P., Mora, D. E., Espinoza, M. J., & Campozano, L. V. (2019). Finding teleconnections from decomposed rainfall signals using dynamic harmonic regressions: a Tropical Andean case study. *Climate Dynamics*, *52*, 4643-4670.

Nazeri Tahroudi, M., & Mirabbasi Najafabadi, R. (2023). Investigating the effect of precipitation series decomposition on the simulation of electrical conductivity of river flow (case study: Eskandari sub-basin). *Journal of Drought and Climate change Research*, *1*(1), 33-48. doi: 10.22077/jdcr.2023.5950.1005

Nazeri Tahroudi, M., & Mirabbasi, R. (2023a). Development of decomposition-based model using Copula-GARCH approach to simulate instantaneous peak discharge. *Applied Water Science*, *13*(9), 182.

Nazeri Tahroudi, M., & Mirabbasi, R. (2023b). Frequency decomposition associated with machine learning algorithms and copula modeling for river flow prediction. *Stochastic Environmental Research and Risk Assessment*, 1-22.

Nazeri Tahroudi, M., Amirabadyzadeh, M., & Zeynali, M. J. (2018). Evaluation of the Accuracy of Artificial Intelligence and Regression Models in Simulation the Daily Temperature. *Journal of Meteorology and Atmospheric Science*, *1*(1), 65-76.

Nazeri-Tahroudi, M., & Ramezani, Y. (2020). Estimation of dew point temperature in different climates of Iran using support vector regression. *IDŐJÁRÁS/Quarterly Journal of the Hungarian Meteorological Service*, *124*(4), 521-539.

Nekoeeyan, M. A., Radmanesh, F., & Ahmadi, F. (2022). Prediction of Monthly Streamflow Using Shannon Entropy and Wavelet Theory Approaches (Case study: Maroon River). *Water and Irrigation Management*, *12*(1), 15-31. doi: 10.22059/jwim.2022.335702.949. [In Persian].

Venugopal, V., & Foufoula-Georgiou, E. (1996). Energy decomposition of rainfall in the time-frequency-scale domain using wavelet packets. *Journal of Hydrology*, *187*(1-2), 3-27.

Wei, S., Yang, H., Song, J., Abbaspour, K., & Xu, Z. (2013). A wavelet-neural network hybrid modelling approach for estimating and predicting river monthly flows. *Hydrological Sciences Journal*, *58*(2), 374-389.

Zakhrouf, M., Bouchelkia, H., Stamboul, M., Kim, S., & Heddam, S. (2018). Time series forecasting of river flow using an integrated approach of wavelet multi-resolution analysis and evolutionary data-driven models. A case study: Sebaou River (Algeria). *Physical Geography*, *39*(6), 506-522.