BİLECİK ŞEYH EDEBALİ ÜNİVERSİTESİ

BŞEÜ Fen Bilimleri Dergisi
8(1), 228-240, 2021

BSEU Journal of Science
https://doi.org/10.35193/bseufbd.878624

e-ISSN:2458-7575 (https://dergipark.org.tr/tr/pub/bseufbd)

Araştırma Makalesi - Research Article

Utilization of Stochastic, Artificial Neural Network, and Wavelet Combined Models for Monthly Streamflow

Aylık Akış Tahmini için Stokastik, Yapay Sinir Ağı ve Dalgacık Bazlı Modellerin Kullanımı

Cenk Sezen[1], Turgay Partal[2*]

Geliş / Received: 19/02/2021 ***Revize / Revised: 27/05/2021*** ***Kabul / Accepted: 28/05/2021***

ABSTRACT

The development of various models to estimate hydrological variables, such as precipitation and runoff is significant regarding handling the water-related problems in the future. This study investigates the performances of Artificial Neural Network (ANN), Auto-Regressive Integrated Moving Average (ARIMA), Wavelet-ARIMA (WARIMA), and WARIMA-ANN models for monthly streamflow forecasting. These models were utilized in two stations of the Susurluk basin in Turkey. In this regard, first, the streamflow data were decomposed into components by wavelet transformation for the WARIMA and WARIMA-ANN models. After that, runoff predictions were performed for each model. As comparison criteria, Root Mean Square Error (RMSE), Kling Gupta Efficiency (KGE), and Nash Sutcliffe Efficiency (NSE) were taken into consideration. As a result, it was obtained that WARIMA and WARIMA-ANN models performed better than the ARIMA and ANN models, particularly. In addition, it was seen that wavelet transformation improved the performance of ARIMA and ARIMA-ANN models, obviously.

Keywords*- *ANN, ARIMA, Streamflow, Wavelet, Turkey

ÖZ

Yağış ve akış gibi hidrolojik verilerin tahmini için farklı modellerin geliştirilmesi gelecekte su ile ilgili problemlerle mücadele edebilmek açısından önemlidir. Bu çalışma, Yapay Sinir Ağı (ANN), Otoregresif Bütünleşik Hareketli Ortalama (ARIMA), Dalgacık-ARIMA (WARIMA) ve WARIMA-ANN modellerinin aylık akım tahmin performanslarını araştırmaktadır. Bu modeller, Türkiye'nin Susurluk havzasındaki iki istasyonda uygulanmıştır. Bu bağlamda, ilk olarak akış verileri WARIMA ve WARIMA-ANN modelleri için dalgacık dönüşümü ile bileşenlerine ayrılmıştır. Daha sonra, her bir model için akış tahminleri gerçekleştirilmiştir. Karşılaştırma ölçütü olarak, Hataların Ortalama Karakökü (RMSE), Kling-Gupta Verimliliği (KGE) ve Nash Sutcliffe Verimliliği (NSE) göz önünde bulundurulmuştur. Sonuç olarak, WARIMA ve WARIMA-ANN modellerinin, özellikle ARIMA ve ANN modellerine göre daha iyi performans gösterdiği tespit edilmiştir.Buna ek olarak, dalgacık dönüşümünün ARIMA ve ARIMA-ANN modellerinin performansını geliştirdiği belirgin şekilde görülmüştür.

Anahtar Kelimeler*- *ANN, ARIMA, Akış, Dalgacık, Türkiye

[1]İletişim: cenk.sezen@omu.edu.tr (https://orcid.org/0000-0003-1088-9360)
Department of Civil Engineering, Ondokuz Mayıs University, Kurupelit Campus, Samsun, Turkey
[2*] Sorumlu yazar iletişim: turgay.partal@omu.edu.tr (https://orcid.org/0000-0002-3779-441X)
Department of Civil Engineering, Ondokuz Mayıs University, Kurupelit Campus, Samsun, Turkey

I. INTRODUCTION

The examination or usage of different methods is significant for improving runoff forecasting. Furthermore, runoff prediction has become crucial concerning water resources planning and the determination of climate change impacts. Many studies have been carried out within this framework, and different methods have been put forward so far [1-3]. Aqil et al. [4] used the recurrent and Feed-Forward Neural Network structures by utilizing various training algorithms for river flow forecasting. They showed that the feed-forward neural network which is trained with the Levenberg-Marquardt algorithm, performed better than other model structures for the river flow prediction. Lin et al. [5] utilized the Support Vector Machine (SVM) models to estimate the hourly inflow in Taiwan for the typhoon warning periods. They pointed out that the SVM model had more advantages such as its ability to produce robust and efficient modelling results compared to the back propagation networks. Zadeh et al. [6] investigated the performance of ANN models by using different input combinations and activation functions (i.e. tangent sigmoid and logistic sigmoid functions) to forecast daily flow in Iran. They found that the selection of input data is significant and the increase of input data can adversely affect the model performance. In addition, they also stated that the ANN model with the tangent sigmoid function yielded better than the ANN model with the logistic sigmoid function for the flow prediction. Kurtulus and Razack [7] examined the performance of Adaptive Neuro-Fuzzy Interference System (ANFIS) and ANN models to predict daily flow in karstic aquifers in south-western France. They stated that the ANFIS model yielded a better performance than the ANN model regarding the forecasting of high flow. Furthermore, they emphasized that the utilization of different input data could affect the model performance. Goyal et al. [8] implemented various Support Vector Regression (SVR) models for the discharge forecasting in karst springs in Greece. They pointed out that the SVR models can be helpful concerning hydrological modelling. Shafaei et al. [9] utilized a hybrid model which consists of the wavelet transform, Seasonal Autoregressive Integrated Moving Average (SARIMA) and Artificial Neural Network (ANN) models for the estimation of precipitation. In this regard, first, they decomposed the precipitation data into the high and low-frequency wavelet components. Accordingly, the low-frequency component was forecasted by the SARIMA model, whereas the high-frequency components were estimated via the ANN model. They revealed that the wavelet-SARIMA-ANN model performed better than wavelet-ANN and wavelet-SARIMA models. Kaur et al. [10] used the wavelet transformation and Autoregressive Moving Average (ARMA) as a hybrid model, and they compared the performance of this hybrid model with ANN- Ensemble Kalman Filtering (EnKF) hybrid model. As a result, they found out that the wavelet-ARMA model outperformed the ANN-EnKF model to predict wind speed. Valipour et al. [11] utilized ARMA, Auto-Regressive Integrated Moving Average (ARIMA) and ANN models to forecast inflow in Iran. They stated that the ARIMA model outperformed the ARMA model. Furthermore, they also expressed that the ANN model could be ideal for inflow estimation compared to the ARMA and ARIMA models. Pektas and Cigizoglu [12] investigated the performance of ANN, Multiple Linear Regression (MLR),and ARIMA models for forecasting of the suspended sediment. Accordingly, they concluded that the ANN model performed slightly better than other models in their study. Liu et al. [13] searched for the impacts of climate change on the runoff regime in China by using the ARMA model. They found that precipitation is more influential than the temperature for the change in streamflow, and runoff had an increasing tendency due to climate change. Valipour [14] analysed the performances of ARIMA and SARIMA models to estimate runoff in the United States. Accordingly, it was suggested that the SARIMA method performed better than the ARIMA method for runoff forecasting. Lohani et al. [15] compared the performance of the ANN, Adaptive Neural Fuzzy Inference System (ANFIS), and AR models for reservoir inflow prediction. They revealed that the ANFIS model yielded better than the ANN and AR models for monthly inflow forecasting. Unes et al. [16] used the MLR, ANN, M5 Decision Tree (M5T), ANFIS, Mamdani-Fuzzy Logic (M-FL), and Simple Membership Functions and Fuzzy Rules Generation Technique (SMRGT) models for flow prediction in Stilwater River, USA. They found out that M-FL and SMRGT models yielded better flow estimation results than other models according to the correlation coefficient, Mean Square Error (MSE), and Mean Absolute Error (MAE) statistics. Fathian et al. [17] utilized the Self-Exciting Threshold Autoregressive (SETAR), Generalized Autoregressive Conditional Heteroscedasticity (GARCH), ANN, Multivariate Adaptive Regression Splines (MARS) and Random Forest (RF) models to estimate monthly runoff in Grand River, Canada. They also implemented hybrid models via coupling of ANN, MARS, and RF models with SETAR and GARCH models. They stated that hybrid models performed better than the stand-alone models concerning river flow prediction. Hussainand Khan [18] compared RF, Multilayer Perceptron (MLP), and SVR models for monthly flow forecasting in Hunza River, Pakistan. They revealed that the RF model yielded more accurate estimation results than the MLP and SVR models. Poonia and Tiwari [19] implemented the Feed-Forward Back Propagation (FFBP) and Radial Basis Function (RBF) neural networks for rainfall-runoff modelling in Narmada River, India. They put forward that ANN models can be useful for the estimation of hydrological variables in agricultural watersheds. These studies indicated that many different black-box and stochastic models have still been utilized to predict hydrological variables such as rainfall, runoff, and groundwater level. In this regard, the improvement of different combined models has importance concerning the estimation of hydrological variables to plan water resources in the future.

In this study, the performances of ANN, ARIMA, and Wavelet-Based ARIMA (WARIMA), and WARIMA-ANN models will be analyzed for the monthly streamflow forecasting in E03A002 Döllük and E03A016Yahyaköy stations of Turkey. For this purpose, Root Mean Square Error (RMSE), Nash Sutcliffe Efficiency (NSE), and Kling-Gupta Efficiency (KGE) error criteria were used for the comparison of model performance. This study is remarkable for comparing and combining stochastic and black-box models and observing the performance development of the ARIMA model thanks to wavelet transformation.

II. MATERIAL AND METHODS

A. Data

In this study, the streamflow data of the two stations, which are located in the northwestern part of Turkey, are used for runoff modeling. The location of the stations and the statistical data (standard deviation, mean, coefficient of skewness, minimum, maximum) which belong to these stations were indicated in Figure 1, Table 1 and Table 2, respectively. The streamflow data cover the period between January 1962 and September 2011 for each station. Accordingly, the streamflow data between January 1962 and October 2001 (80% of data) was used for the training period, whereas the streamflow data between November 2001 and September 2011 (20% of data) was used for the test period in WARIMA-ANN, WARIMA, ARIMA and ANN models.

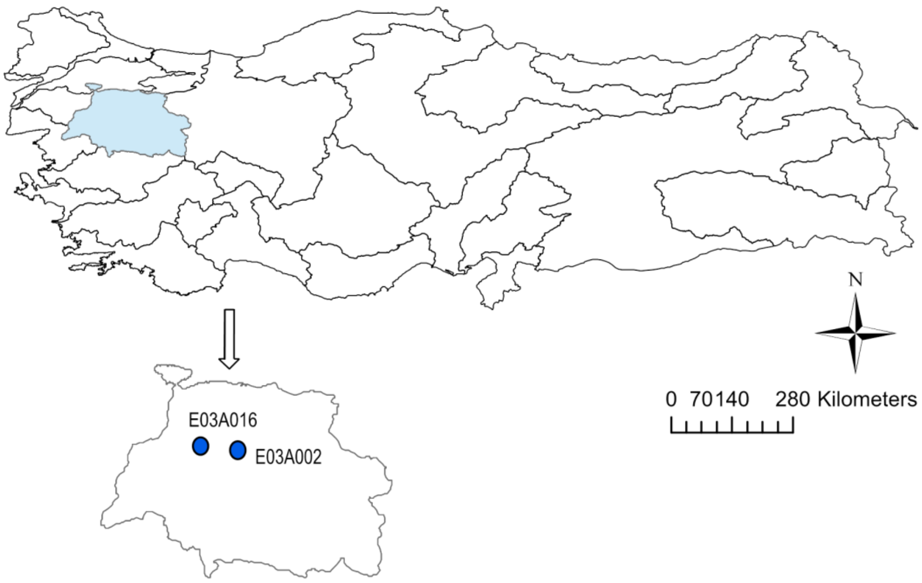

Figure 1. Location of stations in Susurluk basin, Turkey

Table 1. Spatial data of stations

Station No	Basin	Creek	Station Name	Elevation (m)	Lat.	Long.
E03A002	Susurluk	M. Kemalpaşa	Döllük	40	39°57′41″ N	28°30′58″ E
E03A016	Susurluk	Simay	Yahyaköy	30	39°59′10″ N	28°10′34″ E

BSEU Journal of Science
https://doi.org/10.35193/bseufbd.878624

BİLECİK ŞEYH EDEBALİ ÜNİVERSİTESİ

e-ISSN:2458-7575 (https://dergipark.org.tr/tr/pub/bseufbd)

Table 2. Statistical data of streamflow for each station

Station No	Station Name	Statistical Data				
		Mean (m^3/s)	**Std. Deviation (m^3/s)**	**Coefficient of Skewness**	**Minimum (m^3/s)**	**Maximum (m^3/s)**
E03A002	Döllük	52.7	54.5	2.0	5.3	350
E03A016	Yahyaköy	43.9	59.0	2.76	1.0	445.6

B. Methodology

1) Wavelet Transformation: The wavelet transformation is a procedure for fulfilling the time-frequency analysis [20]. Wavelet transform analysis, developed during the last two decades, appears to be a very effective tool in studying non-stationary time series. Wavelet function $\psi(\tau, s)$ for tϵ[-∞,∞] is obtained as shown in equation (1).

$$\psi(\tau, s) = s^{-1/2}\psi\left(\frac{t-\tau}{s}\right) \tag{1}$$

Here, *t* refers to time, τ stands for the time step in which the window function is iterated and,*s* for the wavelet scale [21]. According to this, the continuous wavelet transform of *x*(t) can be obtained as below.

$$W(\tau, s) = s^{-1/2}\int_{-\infty}^{\infty} x(t)\psi * \left(\frac{t-\tau}{s}\right) dt \tag{2}$$

Here, (*) refers to complex conjugate, $W(\tau,s)$ shows the two dimensional depiction of wavelet power. If one chooses scales and positions based on the powers of two (dyadic scales and positions), then the analysis will be much more efficient and accurate. This referred to discrete wavelet transformation [22]. Discrete wavelet transformation was demonstrated in equation (3).

$$\psi_{m,n}\left(\frac{t-\tau}{s}\right) = {s_0}^{-m/2}\psi\left(\frac{t-n\tau_0 s_0^m}{s_0^m}\right) \tag{3}$$

where, *m* and *n* are integers that control the scale and time, respectively,s_0 is a specified fixed dilation step greater than 1, and τ_0 f is the location parameter and must be greater than zero. Besides the translation step, $n\tau_0 s_0^m$ depends on the dilation s_0^m. The most general choice for the parameters s_0 and τ_0 is 2 and 1 (time steps), respectively. Discrete wavelet transformation is a very efficient way with regard to the practical aims [23]. The discrete wavelet transformation for a discrete-time series x_i, where x_i occurs at the discrete-time i (i.e., here integer time steps are used), was illustrated in equation (4).

$$W_{m,n} = 2^{-m/2}\sum_{i=0}^{N-1} x_i\, \psi(2^{-m}i - n) \tag{4}$$

In equation (4), $W_{m,n}$ is the coefficient of wavelet for $s = 2^m$and location$\tau = 2^m n$. Here, wavelet transformation was combined with the ARIMA and ANN models. Components obtained by discrete wavelet transformation were used as inputs for ARIMA and ANN models. The least asymmetric (la8) wavelet filter was utilized as a mother wavelet, and Multiresolution Analysis (MRA) introduced by [23] was used for WARIMA and WARIMA-ANN model. In this regard, data was decomposed into Detail (D_1, D_{2},..,D_j) components and Approximation (A_j) component, where j represents the decomposition level. In this study, streamflow data was decomposed by using discrete wavelet transformation.

2) ArtificialNeural Network: As an ANN training algorithm, the Levenberg-Marquardt backpropagation algorithm was performed to train the data. In this study, Streamflow Data preceding days (Q (t-1), Q (t-2) and Q (t-3)) were used for stream flow forecasting. In this context, different input combinations and architecture of ANN model were shown in Table 3 and Figure 2, respectively.

BSEU Journal of Science
https://doi.org/10.35193/bseufbd.878624

e-ISSN:2458-7575 (https://dergipark.org.tr/tr/pub/bseufbd)

Table 1. Input combinations for ANN model

Model	Input Combination No	Input Combination
ANN	1	Q(t-1)
	2	Q(t-1), Q(t-2)
	3	Q(t-1), Q(t-2), Q(t-3)

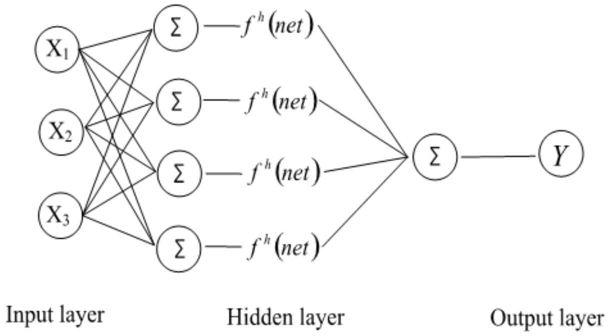

Figure 2. Structure of ANN Model

3) ARIMA Model: ARIMA (p, d, q) model consists of two parts. The first part is the differencing part (d) and the second part is ARMA (p,q) model. Differencing part is significant for the model concerning gaining stationarity[12]. ARMA model, which is the second part of the ARIMA model, is composed of the Autoregressive (AR) model and Moving Average (MA) model as indicated in equation (5) [24]. ARMA model can become stationary thanks to the differentiation part of the ARIMA model [12].

$$y_i = \varphi_1 y_{i-1} + \varphi_2 y_{i-2} + .. + \varepsilon_i - \theta_1 \varepsilon_{i-1} - \theta_2 \varepsilon_{i-2} - \theta_q \varepsilon_{i-q} \quad (5)$$

In this study, the ARIMA model was used for monthly streamflow forecasting, and it was also combined with wavelet transformation. Accordingly, high frequented wavelet components were simulated by using the ARIMA model. To obtain the appropriate ARIMA model for the simulation, Akeike Information Criterion (AIC) was taken into consideration, and then streamflow simulations were carried out. For this analysis, the 'forecast' package in R software was used [25-27]. Before the implementation of the ARIMA model, Box-Cox transformation was applied to data [28, 29].

4) WARIMA Based Model Structure: First, the streamflow data was decomposed into two wavelet components (D1 and D2) and an approximation component (A2) for WARIMA and WARIMA-ANN models. To decide the wavelet decomposition process, the formula (I=int[log (n)]) which was recommended by Wang and Ding [30] was utilized in this study. In that formula, *I* and *n* denote the level of decomposition and data length.

The architecture of WARIMA was illustrated in Figure 3. In WARIMA model, first Box-Cox transformation was implemented to D1 and D2 wavelet components, then D1 and D2 components were estimated by ARIMA model, whereas approximation component is not predicted. Finally, predicted D1 and D2 components, and A2 component were summed. As a result, streamflow forecasting values by using the WARIMA model were obtained.

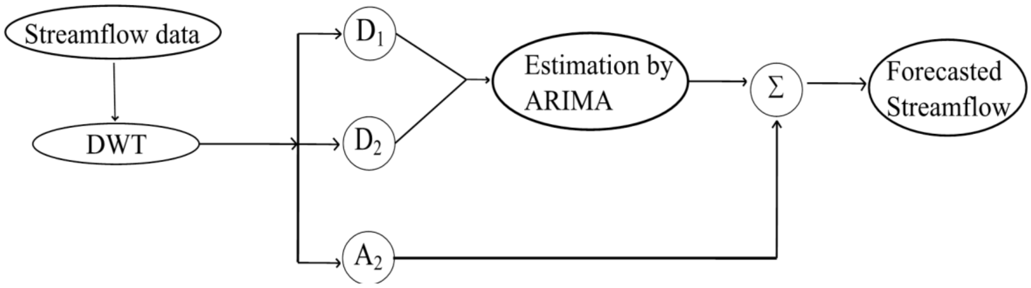

Figure 3. Structure of WARIMA Model

BSEU Journal of Science
https://doi.org/10.35193/bseufbd.878624
e-ISSN:2458-7575 (https://dergipark.org.tr/tr/pub/bseufbd)

In Figure 4, the structure of WARIMA-ANN was given. In this model, first Box-Cox transformation was implemented to D1 and D2 wavelet components, then D1 and D2 components were forecasted by ARIMA model while the A2 approximation component was estimated by ANN model. Then estimated components were summed, as seen in Figure 4.

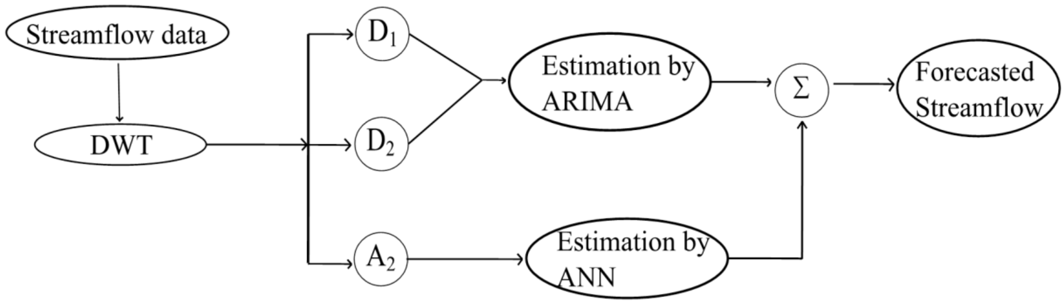

Figure 4. Structure of WARIMA-ANN Model

5)Evaluation of the Model Performance: The performance of the models used in this study was assessed according to RMSE, KGE, and NSE error criteria. The formulation of these criteria was indicated as follows:

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(Q_{sim,i} - Q_{obs,i}\right)^2} \quad (6)$$

$$NSE = 1 - \frac{\sum_{i=1}^{N}\left(Q_{obs,i} - Q_{sim,i}\right)^2}{\sum_{i=1}^{N}\left(Q_{obs,i} - \bar{Q}_{obs}\right)^2} \quad (7)$$

$$KGE = 1 - \sqrt{(r-1)^2 + (\alpha-1)^2 + (\beta-1)^2} \quad (8)$$

In equation (6) and equation (7), $Q_{obs,i,}$ refers to the observed flow*;* $Q_{sim,I}$ refers to the simulated flow for *i.th* time, and *N* refers to data length, respectively. $\overline{Q_{obs}}$ stands for the mean of observed values in equation (7). Furthermore, *r* stands for the correlation coefficient, α for the ratio of simulated mean flow to observed mean flow, and β the ratio of the standard deviation of the simulated flow to the standard deviation of the observed flow in equation (8).

III. RESULTS AND DISCUSSION

First, streamflow data was decomposed by using wavelet analysis. Wavelet components of streamflow data can be seen in Figure 5 and Figure 6 for Döllük and Yahyaköy stations, respectively. As stated in Methodology Section, for the ANN model different input combinations were tried, and input combination 3 gave the best results for the simulation compared with the other input combinations. Because of that, results for the ANN model were given only for the input combination 3. In the ANN model, the Levenberg-Marquardt backpropagation algorithm was used for the training part. The various hidden neurons were tried, and finally, the number of hidden neurons was selected as 3. Then, the selection of the appropriate ARIMA model was fulfilled as stated in Methodology Section. For the WARIMA and WARIMA-ANN models, wavelet components were obtained and then, the forecasting process was carried out as explained in Methodology Section. The ARIMA models which were prepared for the streamflow data of each station were indicated in Table 4 and Table 5, respectively. The frequency distribution of residuals in ARIMA models for both Döllük and Yahyaköy stations can be seen in Figure 7. It can be assumed that residuals of the ARIMA models follow the normal distribution (Figure 7).

BSEU Journal of Science
https://doi.org/10.35193/bseufbd.878624

e-ISSN:2458-7575 (https://dergipark.org.tr/tr/pub/bseufbd)

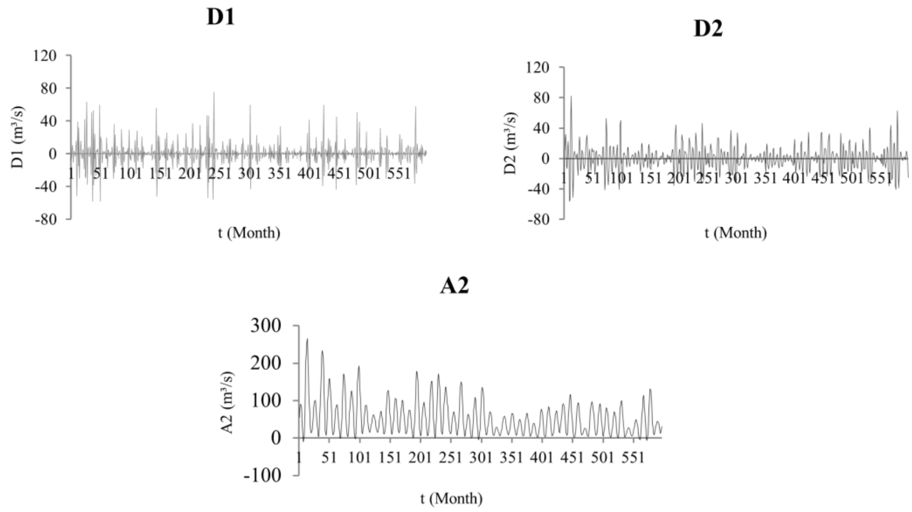

Figure 5. Wavelet Components of Streamflow Data for Döllük Station

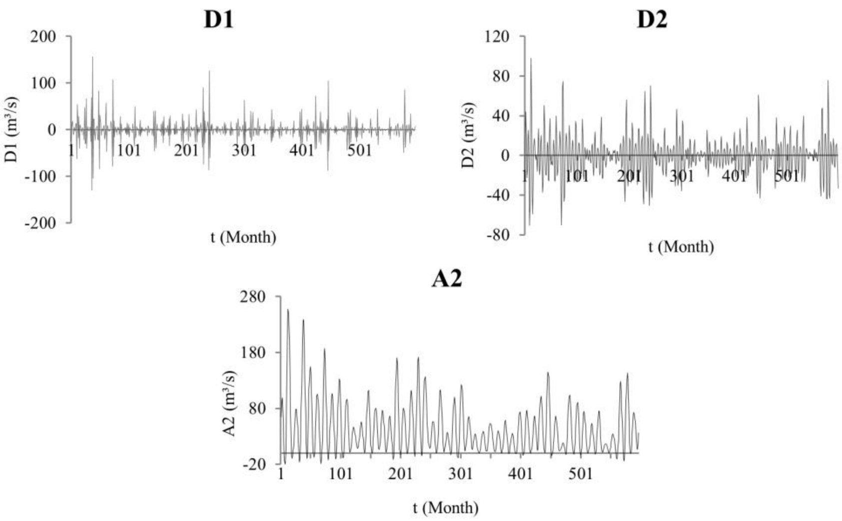

Figure 6. Wavelet Components of Streamflow Data for Yahyaköy Station

Table 4. ARIMA Models for the streamflow forecasting of Döllük station

Model Name	Input Data	ARIMA Models
ARIMA	Q	ARIMA (5,1,1)
WARIMA WARIMA-ANN	D1	ARIMA (1,0,5)
	D2	ARIMA (4,0,5)

Table 2. ARIMA Models for the streamflow forecasting of Yahyaköy station

Model Name	Input Data	ARIMA Models
ARIMA	Q	ARIMA (2,0,2)
WARIMA WARIMA-ANN	D1	ARIMA (1,0,5)
	D2	ARIMA (4,0,5)

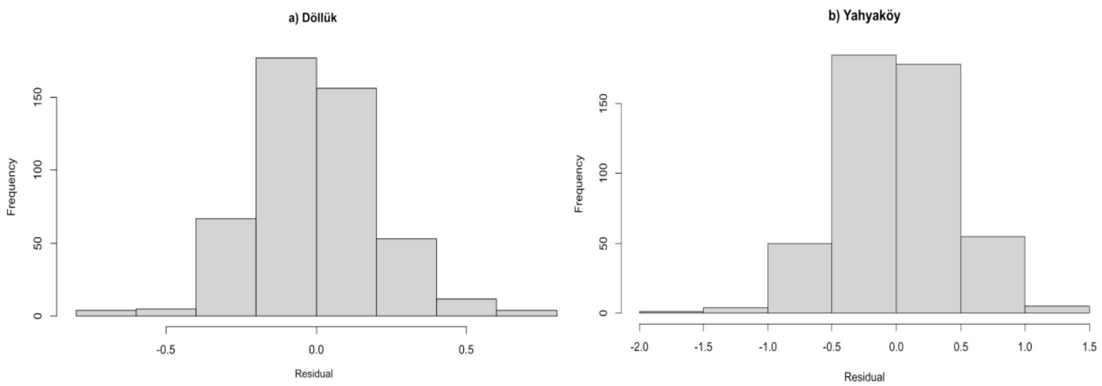

Figure 7. Histogram of residuals in ARIMA models for a) Döllükand b) Yahyaköy stations

After the models were prepared, streamflow forecasting was performed. The relationship between observed and predicted streamflow can be seen in Figure 8and Figure 9. According to Figure 8 and Figure 9, it can be understood that the ARIMA model yielded a poor performance in comparison with other models. However, the performance of the WARIMA and WARIMA-ANN models is better than the ANN model. It can also be inferred that wavelet analysis improved the performance of the ARIMA model significantly. Furthermore, coupling of ANN and ARIMA models in addition to wavelet analysis could enhance the forecasting performance of stand-alone models (i.e., ANN and ARIMA). This is also compatible with the findings of Sahay and Srivastava [22] that indicated the usefulness of hybrid models in hydrological modelling. Although WARIMA and WARIMA-ANN models had some shortages with regard to estimation of especially high flows, both models had a better performance than the ARIMA and ANN models. Scattering diagrams in Figure 10 and Figure 11 also pointed out that WARIMA and WARIMA-ANN models yielded better than the ARIMA and ANN models for runoff modelling. Even though scattering is more particularly for high flows in all models, it is much more in ARIMA and ANN than the WARIMA and WARIMA-ANN models. As stated by Fathian et al [17] stand-alone models could not exhibit sufficient performance for extreme runoff values compared to the hybrid models. In Figure 12 and Figure 13, the performance of models based on RMSE, NSE, and KGE was demonstrated. It can be seen that WARIMA-ANN performed slightly better than the WARIMA model (e.g.,NSE is 0.74 and 0.72; RMSE is 22.4 and 23.3 m^3/s and KGE is 0.73 and 0.73 for WARIMA-ANN and WARIMA models in Döllük station, respectively).In addition, the ARIMA model yielded the worst forecasting results in all models and ANN did not also perform more satisfactory results than the hybrid models (e.g., NSE is 0.47 and 0.15; RMSE is 35.6 and 45.2 and KGE is 0.5 and 0.24 for ANN and ARIMA models in Yahyaköy station, respectively). In this regard, it can be stated that decomposition of dataset via wavelet analysis and prediction of runoff by coupling of ANN and ARIMA models could be useful as shown by Shafaei et al [9].

BSEU Journal of Science
https://doi.org/10.35193/bseufbd.878624

e-ISSN:2458-7575 (https://dergipark.org.tr/tr/pub/bseufbd)

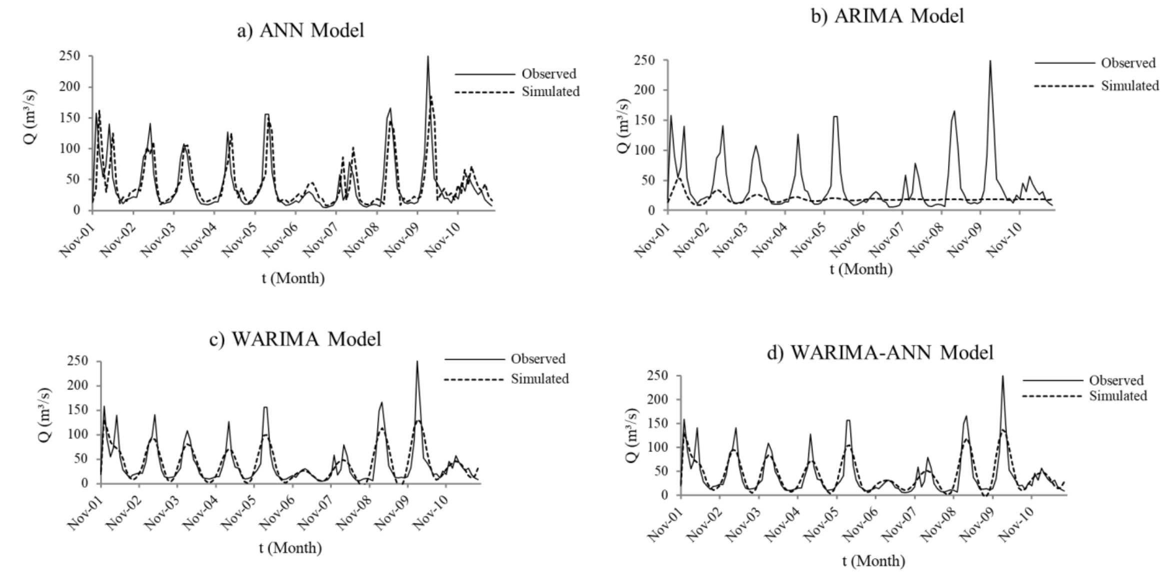

Figure 8. Relationship between forecasted and observed streamflow for a) ANN, b) ARIMA, c) WARIMA, d) WARIMA-ANN models in the Döllük station

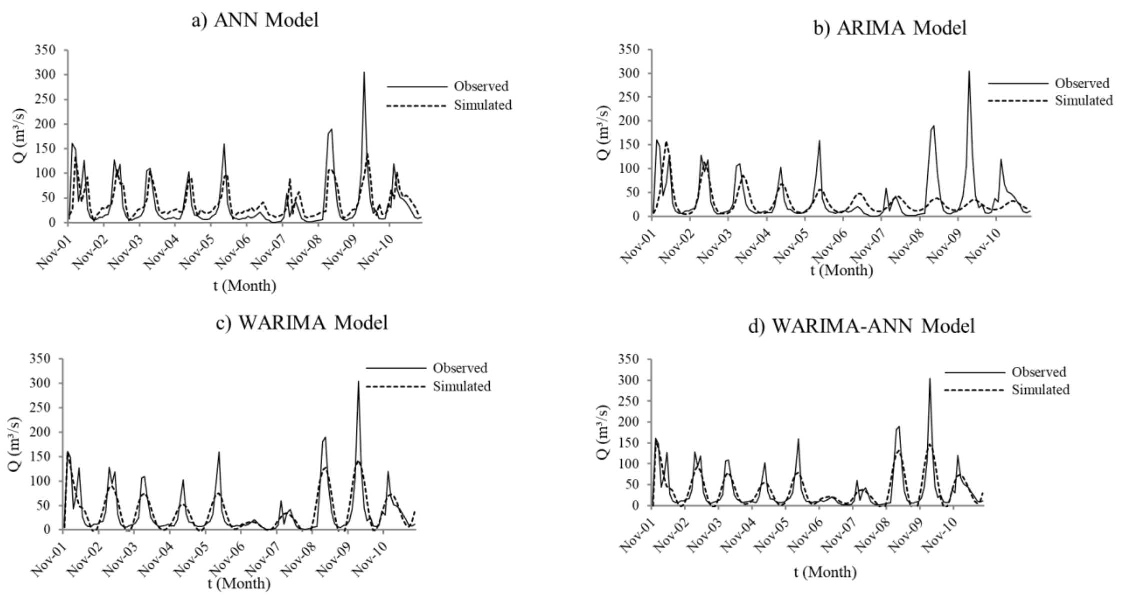

Figure 9. The relationship between forecasted and observed streamflow for a) ANN, b) ARIMA, c) WARIMA, d) WARIMA-ANN models in the Yahyaköy station

BŞEÜ Fen Bilimleri Dergisi
8(1), 228-240, 2021

BSEU Journal of Science
https://doi.org/10.35193/bseufbd.878624

e-ISSN:2458-7575 (https://dergipark.org.tr/tr/pub/bseufbd)

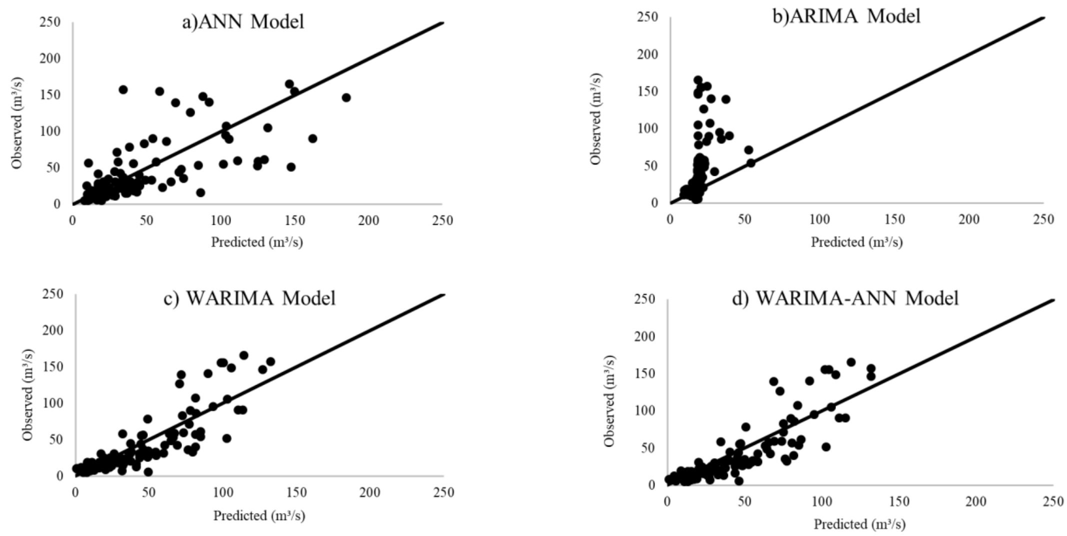

Figure 10. The scattering diagrams for a) ANN, b) ARIMA, c) WARIMA, d) WARIMA-ANN models in the Döllük station

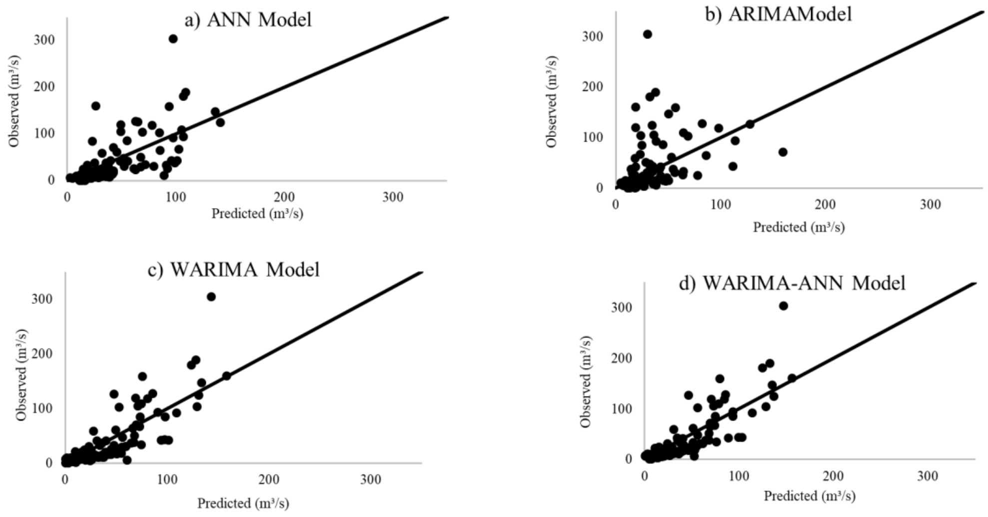

Figure 11. The scattering diagrams for a) ANN, b) ARIMA, c) WARIMA, d) WARIMA-ANN models in the Yahyaköy station

BŞEÜ Fen Bilimleri Dergisi
8(1), 228-240, 2021

BSEU Journal of Science
https://doi.org/10.35193/bseufbd.878624

e-ISSN:2458-7575 (https://dergipark.org.tr/tr/pub/bseufbd)

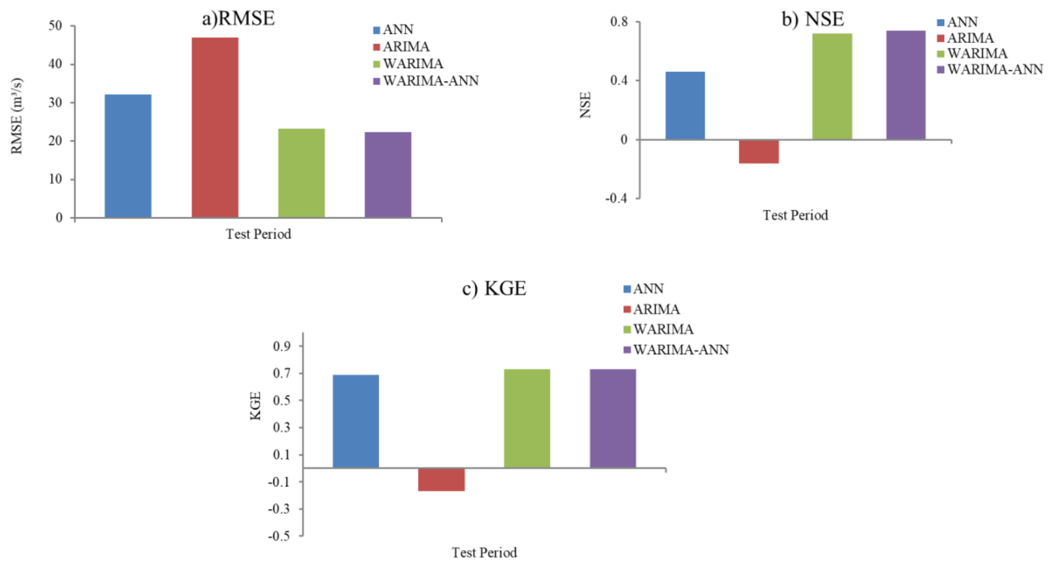

Figure 12. The performance of models a) RMSE, b) NSE and c) KGE for the Döllük station

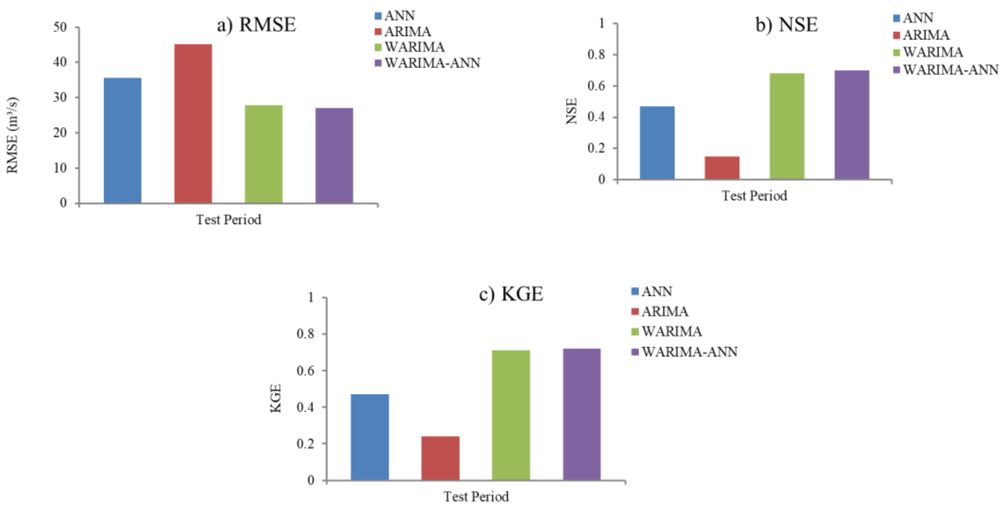

Figure 13. The performance of models (a) RMSE, (b) NSE and (c) KGE for the Yahyaköy station

IV. CONCLUSION

Due to the changing pattern in climate conditions, forecasting hydro meteorological variables has become more critical in extreme events, such as flash floods and droughts. In this regard, the utilization of different hybrid models to predict hydrological variables is very significant in water resource management in the future. The first target of study was the comparing various stochastic and black-box models for the monthly streamflow forecasting in Döllük and Yahyaköy stations situated in Susurluk basin of Turkey. Secondly, it was aimed to determine whether wavelet transformation enhanced the performance of the stochastic and black-box models. Within this framework, hybrid models such as WARIMA and WARIMA-ANN were used as well as the ARIMA and ANN models. As a result, first, it was obtained that the ANN model performed better than the ARIMA model. In addition, the WARIMA-ANN model performed slightly better than the WARIMA model. Both WARIMA and WARIMA-ANN models had superior performance in comparison with the ARIMA and ANN models significantly. In other words, wavelet transformation and coupling of models developed the performance of the

BSEU Journal of Science
https://doi.org/10.35193/bseufbd.878624

e-ISSN:2458-7575 (https://dergipark.org.tr/tr/pub/bseufbd)

ARIMA model considerably. Especially wavelet-based stochastic model and wavelet-based stochastic-black box yielded better performance than either the ARIMA or the ANN model for the monthly streamflow prediction. Although hybrid models yielded more satisfactory forecasting results than the ARIMA and ANN models, they also yielded unsatisfactory prediction performance with regard to extreme flow events. In this respect, further studies will be fulfilled to investigate different hybrid models to improve their performance for the prediction of particularly low and high flows in future studies.

ACKNOWLEDGEMENT

The authors are grateful to General Directorate of State Hydraulic Works for providing the data. This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors. Authors declare that they have no conflict of interest.

REFERENCES

[1] Jain, A. & Kumar A. M. (2007). Hybrid neural network models for hydrologic time series forecasting. *Applied Soft Computing*, 7(2), 585-592.

[2] Adamowski, J. & Chan, H. F. (2011). A wavelet neural network conjunction model for groundwater level forecasting. *Journal of Hydrology*, 407(1-4), 28-40.

[3] Wang, W.C, Chau, K.W., Xu, D.M. & Chen X.Y. (2015). Improving forecasting accuracy of annual runoff time series using ARIMA based on EEMD decomposition. *Water Resources Management*, 29(8), 2655-2675.

[4] Aqil, M, Kita, I., Yano, A. & Nishiyama, S. (2007). Neural networks for real time catchment flow modeling and prediction. *Water Resources Management*, 21(10), 1781-1796.

[5] Lin, G. F., Chen, G. R., Huang, P. Y. & Chou, Y. C. (2009). Support vector machine-based models for hourly reservoir inflow forecasting during typhoon-warning periods. *Journal of Hydrology*, 372(1-4), 17-29.

[6] Zadeh, M. R., Amin, S., Khalili, D. & Singh, V.P. (2010). Daily outflow prediction by multilayer perceptron with logistic sigmoid and tangent sigmoid activation functions. *Water resources management*, 24(11), 2673-2688.

[7] Kurtuluş, B. & Razack, M. (2010). Modeling daily discharge responses of a large karstic aquifer using soft computing methods: artificial neural network and neuro-fuzzy. *Journal of Hydrology*, 381(1-2), 101-111.

[8] Goyal, M. K., Sharma, A., Katsifarakis, K. L. (2017). Prediction of flow rate of karstic springs using support vector machines. *Hydrological Sciences Journal*, 62(13), 2175-2186.

[9] Shafaei, M., Adamowski, J., Fakheri-Fard, A., Dinpashoh, Y. & Adamowski, K. (2016). A wavelet-SARIMA-ANN hybrid model for precipitation forecasting. *Journal of Water and Land Development*, 28(1), 27-36.

[10] Kaur, D., Lie, T. T., Nair, N. K. & Vallès, B. (2015) Wind speed forecasting using hybrid wavelet transform-ARMA techniques. *AIMS Energy*, 3(1), 13-24.

[11] Valipour, M., Banihabib, M. E. & Behbahani, S. M. R. (2013). Comparison of the ARMA, ARIMA, and the autoregressive artificial neural network models in forecasting the monthly inflow of Dez dam reservoir, *Journal of Hydrology*, 476, 433-441.

[12] Pektas, A. O. & Cigizoglu, H.K, (2017). Long-range forecasting of suspended sediment, *Hydrological Sciences Journal*, 62(14), 2415-2425.

[13] Liu, Y., Wu, J., Liu, Y, Hu, B.X., Hao, Y, Huo, X., Fan, Y., Yeh, T. & Wang, Z.L. (2015). Analyzing effects of climate change on streamflow in a glacier mountain catchment using an ARMA model. *Quaternary International*, 358, 137-145.

[14] Valipour, M. (2015). Long-term runoff study using SARIMA and ARIMA models in the United States, *Meteorological Applications*, 22(3), 592-598.

[15] Lohani, A. K., Kumar, R. & Singh, R.D. (2012). Hydrological time series modeling: A comparison between adaptive neuro-fuzzy, neural network and autoregressive techniques. *Journal of Hydrology*, 442, 23-35.

[16] Unes, F., Demirci, M., Zelenakova, M., Calisici, M., Tasar, B., Vranay, F. & Kaya, Y. Z. (2020). River Flow Estimation Using Artificial Intelligence and Fuzzy Techniques. *Water*, 12(9), 2427.

[17] Fathian, F., Mehdizadeh, S., Sales, A. K. & Safari, M. J. S. (2019). Hybrid models to improve the monthly river flow prediction: Integrating artificial intelligence and non-linear time series models. *Journal of Hydrology*, 575, 1200-1213.

[18] Hussain, D. & Khan, A. A. (2020). Machine learning techniques for monthly river flow forecasting of Hunza River, Pakistan. *Earth Science Informatics*, 1-11.

[19] Poonia, V. & Tiwari, H. L. (2020). Rainfall-runoff modeling for the Hoshangabad Basin of Narmada River using artificial neural network. *Arabian Journal of Geosciences*, 13(18), 1-10.

[20] Partal, T. (2017). Wavelet regression and wavelet neural network models for forecasting monthly streamflow, *Journal of Water and Climate Change*, 8(1), 48-61.
[21] Meyer, Y. (1993). Wavelets algorithms and applications. Society for Industrial and Applied Mathematics, Philadelphia.
[22] Sahay, R. R. & Srivastava, A. (2014). Predicting monsoon floods in rivers embedding wavelet transform, genetic algorithm and neural network, *Water Resources Management*, 28(2), 301-317.
[23] Mallat, S. G. (1989). A theory for multiresolution signal decomposition: the wavelet representation. IEEE transactions on pattern analysis and machine intelligence, 11(7), 674-693
[24] Bayazıt, M. (1996). İnşaat mühendisliğinde olasılık yöntemleri. İstanbul Teknik Üniversitesi, İstanbul.
[25] Hyndman, R. J. & Khandakar, Y. (2008). Automatic time series forecasting: the forecast package for R. *Journal of Statistical Software*, 26(3), 1–22.
[26] Hyndman, R, Bergmeir, C., Caceres, G., Chhay, L., O'Hara-Wild M., Petropoulos, F., Razbash, S., Wang, E. & Yasmeen, F. (2018). Forecast: Forecasting functions for time series and linear models, R package version 8.3, http://pkg.robjhyndman.com/forecast (Erişimtarihi: 01.04.2018)
[27] R Core Team, R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/ (Erişimtarihi: 01.04.2018).
[28] Box, G. E. P. & Cox, D. R. (1964). An analysis of transformations. *Journal of the Royal Statistical Society: Series B (Methodological)*, 26, 211–246.
[29] Bickel, P. J. & Doksum K. A. (1981). An Analysis of Transformations Revisited. *Journal of the American Statistical Association*, 76, 296-311.
[30] Wang, W & Ding J. (2003). Wavelet network model and its application to the prediction of hydrology, *Nature and Science*, 1(1), 67-71.