Modelling of Bektas Creek Daily Streamflow with Generalized Regression Neural Network Method

Fatma AKCAKOCA[1] & Halit APAYDIN [2]

Research Scholar[1] and Professor[2]

[1-2]University of Ankara, Faculty of Agriculture

[1-2]Department of Agricultural Structures and Irrigation

Ankara, Turkey

ABSTRACT

The dependence of agricultural production on water resources is a known fact. Therefore, understanding hydrological processes and events in agricultural production form the basis of effective and reliable management of water resources. Many traditional methods used for the analysis of time-dependent variables in hydrology and meteorology assume linear relationships between these variables. However, the temporal changes of these parameters are quite complex and therefore cannot be easily modeled by conventional estimation methods. Artificial neural networks (ANNs), on the other hand, allow the analysis of nonlinear relationships or processes whose statistical or mathematical calculations cannot be determined in such systems. ANNs have been accepted as a successful model in multidimensional research involving dynamic processes in the field of hydrology.

This study, it is aimed at the modeling of stream flows in Bektas Creek. With the aim of modeling, daily meteorological parameters (precipitation, temperature, sunbathing time, relative humidity) measured in Kangal Region and one day delayed flows were used. Streamflow forecasts are simulated with the Generalized Regression Neural Network (GRNN). To reveal the difference of the GRNN model from other ANNs, the same data were also used in the feed-forward neural network (FFNN) model. Model performances were evaluated taking into account the Root Mean Square Error (RMSE), Mean Absolute Error (MAE) and Determination coefficient (R).

Keywords: Meteorological Parameters, Stream flow Prediction, Generalized Regression Neural Network, Feed-Forward Neural Network.

---

1. INTRODUCTION

Water resources management authorities, in the sustainable management and use of watershed water resources, agricultural irrigation needs; It states the necessity of planning the basin by hydrological, meteorological and agrological conditions. This may be possible by accurately determining the projection of water volume and distribution on a watershed basis, now and in the future.

In a watershed whose limits are determined, considerable information can be obtained about the nature of nature with hydrological and meteorological measurements based on time. However, for some parameters, measurements cannot be made completely or the process cannot be modeled completely with measurement data. Therefore; statistical models in hydrology are frequently used to predict future events of a certain magnitude. Knowing statistics is hydrology; The planning of water resources is very important in the design, optimization, and operation of water structures.

Traditional methods used for the analysis of time-dependent variables in hydrology and meteorology often assume that there are linear relationships between these variables [1]. But in reality, the temporal changes of these parameters are quite complex and therefore cannot be easily modeled by conventional estimation methods. Artificial neural networks (ANN) allow systems with such nonlinear relationships or processes to be analyzed without statistical or mathematical assumptions or calculations [2]. The ANN is basically of two different types: Feed Forward Neural Networks (FFNN) where inputs are transmitted only to the output layer and Recurrent Neural Networks (RNN) where inputs can advance in any direction [3]. Generalized Regression Neural Networks (GRNN), which was accepted as a type of FFNN, was proposed by Donald F. Specht in 1991. GRNNs are in the hidden layer; It uses gauss kernels that are successful in solving an infinite number of transactions that are impossible to calculate. According to Specht, GRNN is an effective technique in regression, prediction, and classification of dynamic systems created by nonparametric inputs [4].

Many studies are using ANNs to model the basic hydrological processes of the basins without the need for too much data. In literature, ANN showed a successful performance in precipitation-flow modeling studies comparing different types of ANN with traditional methods [5], [6], [8]. The management of ANN is quite complex, such as drought analysis, precipitation forecasts, sediment transport [7], [9], irrigation reservoir optimization [8] and groundwater modeling [10]. With its applications extending to its areas, the working area has become enriched gradually. Another frequently and successfully applied field of ANN methodology is the studies that are used with measured continuous flows data or for flow estimation of daily flows and meteorological parameters ([11], [12], [13], [14], [15].

In the sustainable management of water resources, ANN has an important place in the classification of nonlinear data, prospective estimates, and modeling of existing dynamic processes. Generalized Regression Neural Networks, one of the different network architectures developed in the field of ANN, is based on the standard regression statistics technique [16]. GRNNs estimate the variable with the least error as in the regression method, which is a powerful calculation tool. In the presented study, four different meteorological parameters (daily total precipitation, daily temperature, daily sunshine time and daily relative humidity) and one day delayed flow data are used to be in the daily flow forecast of Bektas Creek. For this purpose, streamflow estimates were simulated with generalized regression neural network and feed-forward artificial neural network.

2. MATERIAL and METHOD

2.1. Case Study

Turkey is divided into 26 major river watersheds for hydrological studies. Daily flow values measured between Jan 1993 and Sep 2018 were used in Akcakale Bektas Creek Station (D21A243) of the Sivas Region in the Fırat-Dicle Watershed (Figure 1). Precipitation area of the observation station is 101 km^2, the average annual flow is 0.478 m^3/s. The highest monthly average flow was in April with 3,871 m^3/s and the lowest monthly average flow was in August with 0.025 m^3/s.
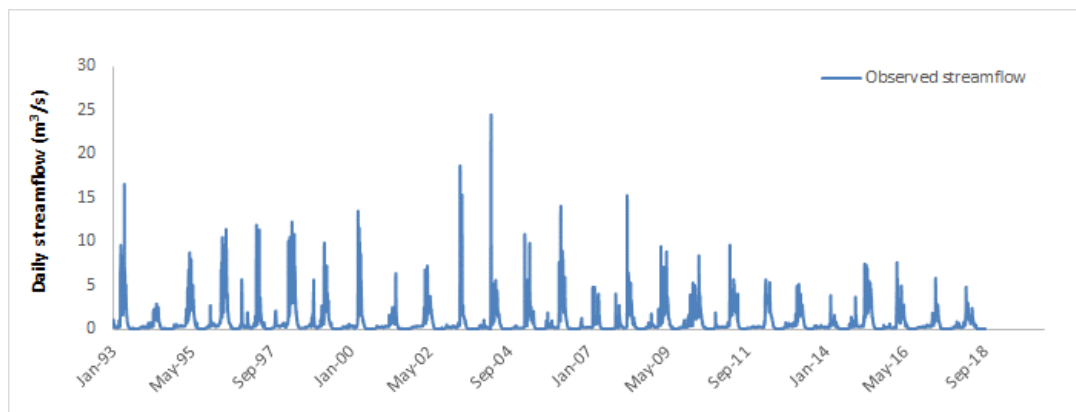

Figure 1. Daily streamflow data of Bektas Creek station (1993-2018)

The meteorological data used in the study were obtained from Kangal (17762) and Divrigi (17734) meteorology observation stations, which are closest and approximately elevation to the streamflow observation station, and the data of these two stations were used. The meteorological data set constitutes 9396 daily average temperature, sunshine duration relative humidity, total precipitation series between January 1993 and September 2018 (Figure 2). Average, standard deviation, minimum and maximum values of these data are given in Table 1.
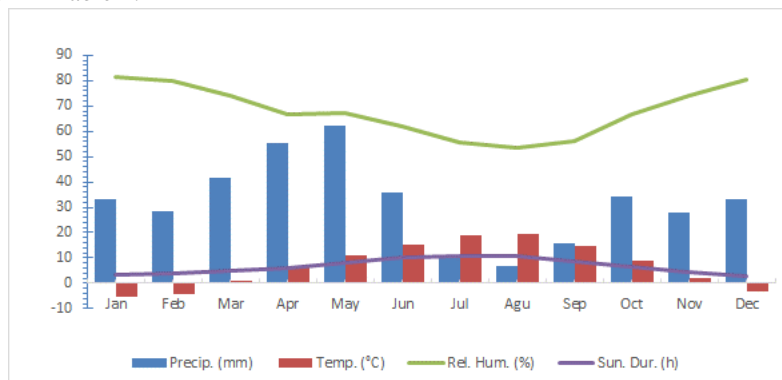

Figure 2. Daily meteorological data (1993-2018)
Table 1. Statistical features of the data set

Variable	Minimum	Maximum	Mean	Std. deviation
T_{mean} (°C)	-23,9	28,6	9,5	9,6
Sunshine duration (h)	0,0	14,0	7,0	4,0
Mean RH (%)	14,2	95,5	62,7	14,7
Precipitation (mm)	0,0	44,1	1,0	2,6
Streamflow (m³/s)	0,0	24,2	0,7	1,4

The location of the streamflow and meteorological observation stations on the map of the region is shown in Figure 3.

Figure 3. Location of streamflow-meteorological observation stations

2.2. Generalized Regression Neural Network (GRNN)

The generalized regression neural network consists of four layers: input, pattern, addition, and output layer (Figure 4). In the input layer, the number of input parameters is neuron: connected to the second layer by weights. GRNN uses normalized gauss nuclei as activation functions in hidden layers. Due to the effectiveness of Gauss nuclei on complex and infinite operations, it can memorize unique patterns between input and output parameters when network training is complete [4].
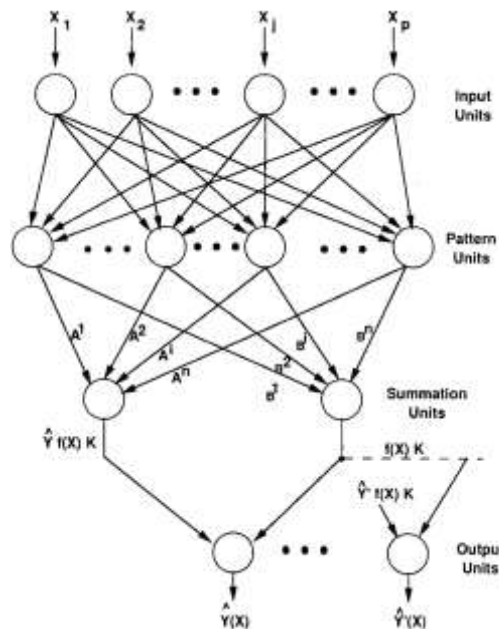

Figure 4. Generalized regression neural network structure [4]

The probability density function used as the GRNN algorithm is in the normal distribution (Equation 1). Each training example (X_i) is used as the average of the normal distribution. The Euclidean distance (D_i^2) between the training example and the prediction point is used as a measure of how well each training sample can represent the prediction distance (X) (Equation 2). It is called the standard deviation or smoothness parameter (σ), and as this value increases, the projection distance decreases and the decrease distance increases [4].

$$(X) = \frac{\sum_{i=0}^{n} Y^i exp\left(\frac{D_i^2}{2\sigma^2}\right)}{\sum_{i=0}^{n} exp\left(-\frac{D_i^2}{2\sigma^2}\right)} \qquad (1)$$

$$D_i^2 = (X - X^i)^T (X - X^i) \qquad (2)$$

2.3. Selecting Models

Correlations of precipitation, relative humidity, temperature, sunshine duration time, delayed streamflow, and daily flow values were used in the establishment of the models. According to the correlation results, the streamflow values on day t showed a high correlation with the flow on day t-1, while it had a relatively lower correlation with the precipitation and relative humidity values on day t. In addition, it is observed that the streamflow values on day t correlate negatively with the temperature and sun time on day t and are low (Table 2).

Table 2. Correlation matrix between flow and meteorological data

	Q_{t-1}	RH_t	T_t	n_t	P_t	Q_t
Q_{t-1}	1	0.13	-0.07	-0.1	0.13	0.89
RH_t	0.13	1	-0.67	-0.5	0.23	0.13
T_t	-0.07	-0.67	1	0.49	-0.06	-0.09
n_t	-0.1	-0.5	0.49	1	-0.18	-0.09
P_t	0.13	0.23	-0.06	-0.18	1	0.12
Q_t	0.89	0.13	-0.09	-0.09	0.12	1

Qt-1: one day delayed streamflow (m3/s), RHt: daily average relative humidity (%),
Tt: daily average temperature (°C), nt: daily average sunshine duration (h), Pt: daily
total precipitation (mm), Qt: daily streamflow (m3/s)

Different models were created according to the correlation results obtained (Table 3).

Table 3. Installed models

MODEL	Input Parameters	ANN Type		Output Parameter
model 1	Pt, Tt, RHt, nt, Qt-1	FFNN	GRNN	Qt
model 2	Pt, Tt, RHt, Qt-1	FFNN	GRNN	Qt
model 3	Pt, Tt, Qt-1	FFNN	GRNN	Qt
model 4	Pt, Qt-1	FFNN	GRNN	Qt
model 5	Pt	FFNN	GRNN	Qt
model 6	Qt-1	FFNN	GRNN	Qt

FFNN: Feed Forward Neural Network, GRNN: Generalized Regression Neural Network

GRNN has been used with the help of the toolbox on the MATLAB 8 version. The mesh structure consisting of the input layer, the pattern layer, and the output layer and the roughness coefficient (σ) 1.0 is used. FFNN, network architecture three-tier; In the input layer, the number of neurons up to the number of input parameters (n), a hidden layer of 10 neurons and the output layer [n, 10,1]. Levenberg-Marquart algorithm and tangent sigmoid activation function were used in the training of the network. Also, the training of the network: the momentum coefficient was 0.001, the learning rate was performed with 0.1 to 1000 repetitions. For both methods, the data set was used in 70% training, 15% validation and 15% testing.

2.4. Evaluation of Model Performance

Three different statistical methods, correlation coefficient (R), root mean square error (RMSE) and mean absolute error (MAE) values were taken into consideration in evaluating the performance of the model results (Equation 3, 4, 5). Model performance increases as the error results converge to 0, and the R-value approaches to 1.

$$R = \frac{\sum_{i=1}^{n}(X_i - \bar{X}) - \sum_{i=1}^{n}(X_i - Y_i)}{\sum_{i=1}^{n}(X_i - \bar{X})} \qquad (3)$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(X_i - Y_i)^2} \qquad (4)$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|X_i - Y_i| \qquad (5)$$

in equations:

Xi: observed value, $\bar{X}$: average observed value Yi: predicted value, n: number of values

3. RESULTS

The performance values of the results obtained from the models used in the study are given in Table 4. When the results are examined, GRNN performances for the six different models studied yielded higher results than FFNN. Model 1 results created from all meteorological parameters and previous flow inputs in both methods showed the highest performance compared to other models. The performance of model 5, which consists only of precipitation inputs, showed the lowest performance in both methods. In addition, in another model created by using the least input parameter, good performance was obtained with the FFNN method for model 6, which consists of one day delayed streamflow inputs (Table 4).

Table 4. Statistical evaluation results of model performances

ANN Type	FFNN			GRNN		
Perf. Statics	R	RMSE	MAE	R	RMSE	MAE
model 1	0,829	0,557	0,682	0,972**	0,236	0,631
model 2	0,825	0,568	0,706	0,922	0,370	0,592
model 3	0,819	0,564	0,694	0,841	0,489	0,564
model 4	0,828	0,555	0,683	0,791	0,545	0,456
model 5	0,037	0,286	0,553	0,036	0,255	0,513
model 6	0,828*	0,569	0,690	0,785	0,547	0,446

In daily streamflow estimation, the outputs produced by the most successful model 1 in both methods are given as scatter graph (Figure 4a, 4b) and time series (Figure 5a, 5b).
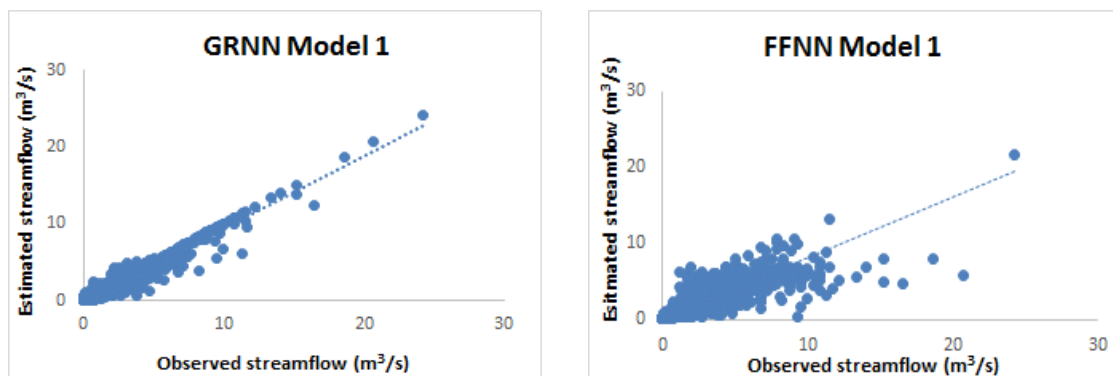

Figure 4. (a) Results of Model 1 in the GRNN method, (b) Results of Model 1 in the FFNN method
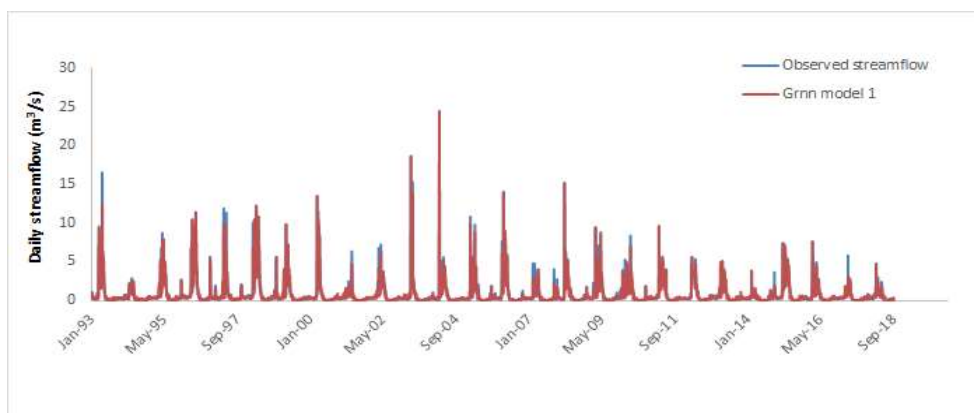

Figure 5. (a) Daily Streamflow Estimates of GRNN in Model 1 - Bektas Station
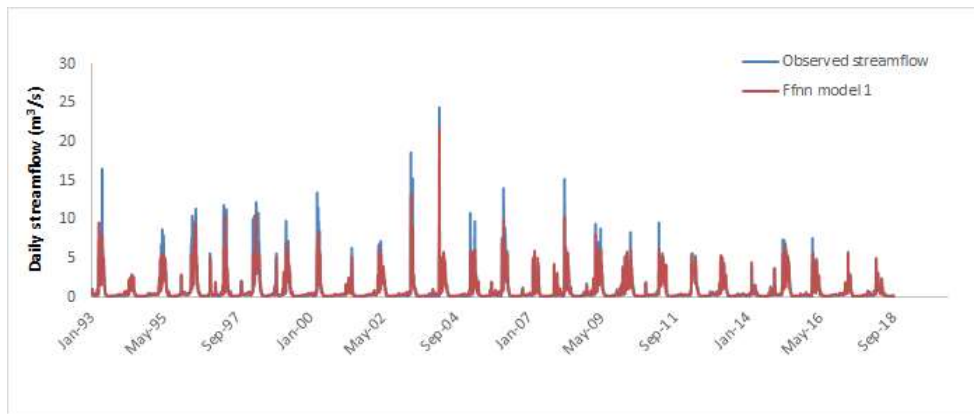
Figure 5. (b) Daily Streamflow Estimates of FFNN in Model 1 - Bektas Station

When Tables 3 and 4 are examined together, it is seen that Model 1 makes the best estimates, but needs five different input parameters. In some cases, it is acceptable to estimate within acceptable performance limits with a minimum of parameters. Outputs produced by model 6, which can be considered as an alternative for estimating daily flows with the least parameters, are given as scatter graphs (Figure 6a, 6b) and time-series graphics (Figure 7a, 7b).
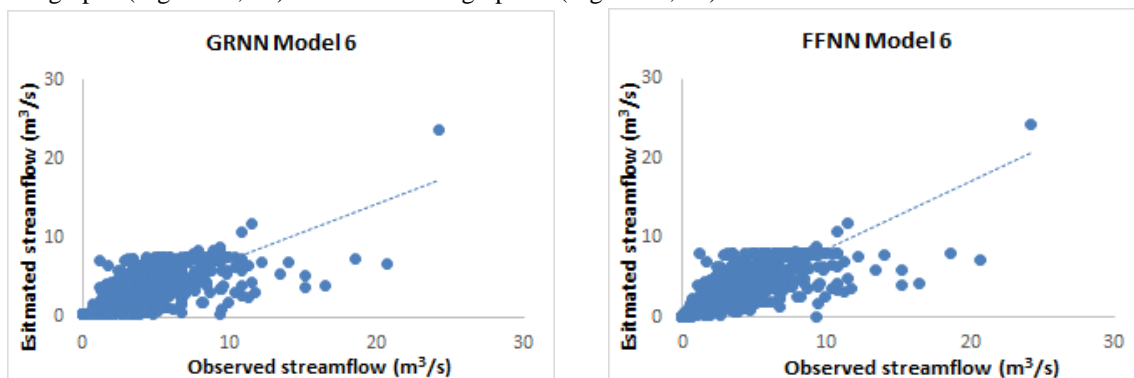

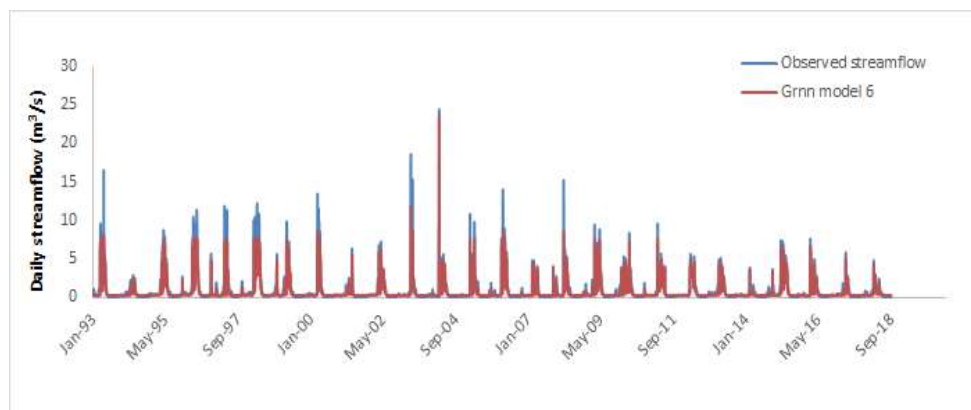**Figure 6. (a) Results of Model 6 in the GRNN method, (b) Results of Model 6 in the FFNN method**

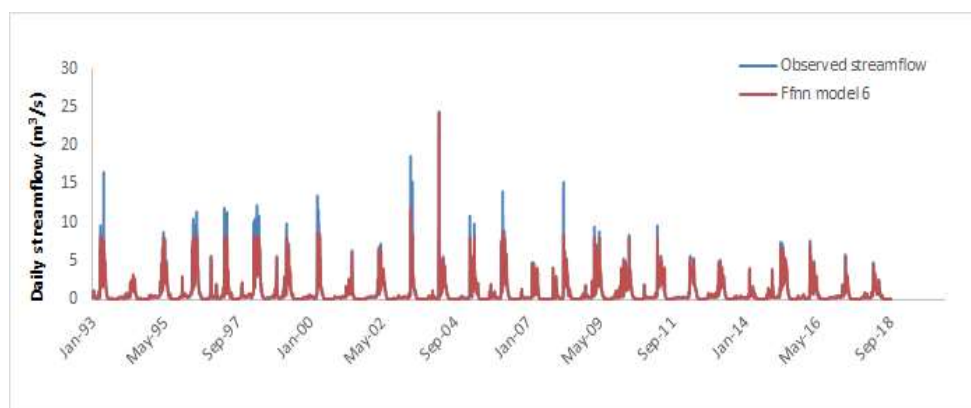**Figure 7. (a) Daily Streamflow Estimates of GRNN in Model 6 - Bektas Station**

Figure 7. (b) Daily Streamflow Estimates of FFNN in Model 6 - Bektas Station

4. CONCLUSIONS

In the study, daily flows of Sivas Akcakale-Bektas Creek observation station in Fırat-Dicle basin were modeled on the basis of the generalized regression neural network method. For this purpose, six different models were established in which one day delayed flows and meteorological parameters formed the inputs of the artificial neural network (ANN) method. The performances of all models in daily flow estimation obtained by the GRNN method were compared with the FFNN method.

When model performances are analyzed, it is seen that GRNN and FFNN methods give quite different results. The outputs of the 1st Model prepared with the GRNN method show a very close distribution to the y = x line and the harmony in estimates is remarkable. When the performance results are examined, the superiority of the GRNN method, which is thought to represent the daily flows of Akcakale Bektas Creek, has been proved over the FFNN method. In the 6th model, which was created with the input of only the previous flows, the FFNN method was found to be more successful than the GRNN method, but the homogeneous distribution could not be achieved as much as the results of the 1st model with the GRNN method. However, it was found to be usable in estimating the streamflow with the least parameters.

In addition, it is understood that meteorological parameters can be successfully modeled with the GRNN method as a function of daily flows. It seems that the GRNN approach can be used in modeling the flow characteristics of river watersheds.

REFERENCE

1. H.K. Cigizoglu." Estimation, forecasting and extrapolation of river flows by artificial neural networks." *Hydrological Sciences Journal*. 48(3), pp. 349–361, 2003. Available: DOI: 10.1623/hysj.48.3.349.45288
2. M. Yasar. "Completion of vacancies in daily flows by using artificial neural networks." Master's thesis, University of Pamukkale, 2004.
3. J. Al-Mahasneh, S. G. Anavatti and M. Garratt. "Review of applications of generalized regression neural networks in identification and control of dynamic systems." 2018.
4. 1.D.F. Specht (1991, Nov.). "A general neural network." *IEEE Transactions on Neural Networks*, 2(6), pp. 568-576. Available: DOI: 10.1109/72.97934
5. A.W. Minns and M.J. Hall "Artificial neural networks as rainfall runoff models" *Hydrological Sciences Journal*. 41(3), pp. 399-417, 1996.
6. D.A.K. Fernando and A.W. Jayawardena. "Runoff forecasting using RBF networks with OLS algorithm." *Journal of Hydrologic Engineering*. 3(3), pp. 203-209, 1998.
7. H.K. Cigizoglu. "Suspended sediment estimation and forecasting using artificial neural networks." *Turkish Journal of Engineering and Environmental Sciences*. 26, pp. 15-25, 2002.
8. N. Sajikumar and B. S. Thandaveswara. "A Non-linear Rainfall-Runoff Model Using an Artificial Neural Network." *Journal of Hydrology*. 216, pp. 32-55, 1999.
9. M. T. Sattari, H. Apaydin and F. Ozturk." Flow estimations for the Sohu Stream using artificial neural networks." *Environmental Earth Sciences*. 66(7), pp. 2031-2045, 2012. Available: DOI 10.1007/s12665-011-1428-7
10. S. Ranjithan, J.W. Eheart and J.H. Garrett. "Neural network-based screening for groundwater reclamation under uncertainty." *Water Resources Research*. 29(3), pp, 563-574, 1993.
11. S. Tokar and P. A. Johnson. "Rainfall-runoff modeling using artificial neural networks." *ASCE J. of Hydrol. Eng.* 4(3), pp. 232-239, 1999. Available: DOI 10.1061/(ASCE)1084-0699(1999)4:3(232)
12. H.K. Cigizoglu and M. Alp. "Rainfall-Runoff Modelling Using Three Neural Network Methods." *ICAISC*, 2004, pp. 166-171.
13. Ö. Kisi. "A Combined Generalized Regression Neural Network Wavelet Model for Monthly Streamflow Prediction." *KSCE Journal of Civil Engineering*. 15(8), pp. 1469-1479, 2011. Available: DOI 10.1007/s12205-011-1004-4
14. U. Okkan and A. Mollamahmutoglu. "Daily runoff modelling of Yigitler Stream by using artificial neural networks and regression analysis." *Dumlupinar University Journal of the Institute of Science*. 023, pp. 33-48, 2010.
15. M. T. Sattari, H. Apaydin and F. Ozturk." M5 model tree application in daily river flow forecasting in Sohu Stream, Turkey." *Water Resources*. 40(3), pp. 233-242, 2013. Available: DOI 10.1134/S0097807813030123
16. 2.M.M. Bauer. "General regression neural network for technical use." Master's thesis, University of Wisconsin-Madison, 1995.