The Eurasia Proceedings of Science, Technology, Engineering & Mathematics (EPSTEM)
ISSN: 2602-3199

The Eurasia Proceedings of Science, Technology, Engineering & Mathematics (EPSTEM)

Volume 1, Pages 372-378

ICONTES2017: International Conference on Technology, Engineering and Science

A WAVELET TRANSFORMATION-GENETIC ALGORITHM-ARTIFICIAL NEURAL NETWORK COMBINED MODEL FOR PRECIPITATION FORECASTING

Cenk Sezen
Ondokuz Mayıs University

Turgay Partal
Ondokuz Mayıs University

Abstract: Black box models are one of the most common hydrological models in order to make predictions of hydrological variables such as precipitation and stream flow. In this study, performance of a combined model which consists of wavelet transformation, genetic algorithm and artificial neural network (WGANN) were tested for prediction of monthly precipitation by using North Atlantic Oscillation (NAO) index, Southern Oscillation (SO) index and precipitation data as input in the model. The case study was carried out for Antalya which is located in Mediterranean region of Turkey. As a result, it was attained that WGANN model performed more successful than usual artificial neural network (ANN), multiple linear regression (MLR) and genetic algorithm-artificial neural network (GANN) models.

Keywords: Artificial neural network, genetic algorithm, precipitation, regression, wavelet transformation

Introduction

The forecasting of hydrological variables such as rainfall, runoff and evapotranspiration is very substantial to comprehend the hydrological process in the nature. For this purpose, many hydrological models have been put forward from past to present (Chen and Chang 2009; Zorn and Shamseldin 2015; Nasseri et al. 2008; Alp and Cigizoglu 2007). Daliakopoulos et al. (2005) predicted groundwater level by using artificial neural network (ANN) in Greece. Gao et al. (2010) utilized ANN model for predicting the stream flow in Huaihe River Basin of China. Moreover, there have been a lot of hybrid models in the literature that aim at improving the estimation of hydrological parameters (Kim and Valdés 2003; Partal and Cigizoglu 2008; Sahay and Srivastava 2014). Partal and Cigizoglu (2009) applied the hybrid model which composes of discrete wavelet transform (DWT) and ANN in order to forecast the daily precipitation and they obtained the well performance of this model as compared with usual ANN model or multi linear regression model. Shoaib et al. (2014) used the conjunction model of wavelet transformation and artificial neural networks in order for rainfall-runoff modelling and they acquired successful results for the application of discrete wavelet transformation with the Multilayer Perceptron Neural Network (MLPNN) and the Radial Basis Function Neural Network (RBFNN).

The target of this study is comparing the performances of multiple linear regression (MLP), ANN, the hybrid model of genetic algorithm-artificial neural network (GANN) and the hybrid model of wavelet transformation-genetic algorithm-artificial neural network (WGANN) for the projection of monthly precipitation in Antalya that is situated in Mediterranean region of Turkey.

- Selection and peer-review under responsibility of the Organizing Committee of the conference
*Corresponding author: Cenk Sezen-E-mail: cenk.sezen@omu.edu.tr

Methods

Wavelet Transformation

The wavelet transformation is a method for carrying out the time-frequency analysis in a time interval (Partal, 2017). If wavelets have finite dimensions in the time interval, it is called as discrete wavelet transformation (Sahay and Srivastava, 2014). In addition, for a continuous time interval tϵ[-□□□-□□□it is called as continuous wavelet transformation. For a continuous time domain, wavelet function □□□τ□ s□□ can be attained as indicated in equation (1).

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

In equation (1), *t* represents the time, τ stands for the time step in which the window function is iterated and *s* for the wavelet scale (Meyer, 1993). The continuous wavelet transform of *x(t)* is as illustrated in equation (2).

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

In equation (2), (*) refers the complex conjugate, *W*□□□,*s*□□represents the two dimensional illustration of wavelet power (Partal, 2017). Furthermore, discrete wavelet transformation was also shown in equation (3).

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

□□
□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

In equation (3), *m* and *n* are integers which control the wavelet scale and time, respectively, s_0 stands for a specific fixed expansion step greater than 1, and τ_0 for the location parameter and must be greater than zero (Partal, 2017). In addition, $n\tau_0 s_0^m$ is the translation step and it is based on expansion s_0^m. The most widespread option for s_0 is 2 and for τ_0 is 1.

Discrete wavelet transformation that exhibit the strength of two logarithmic scaling of the translations is very influential method in terms of practical aims (Mallat, 1989). The discrete wavelet transformation for a discrete time series *xi*, where *xi* takes place at discrete time *i* (i.e., here integer time steps are used), was illustrated in equation (4) (Partal, 2017).

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

In equation (4), $W_{m,n}$ is the coefficient of wavelet for the discrete wavelet of scale $s = 2^m$ and location $\tau = 2^m n$. In this study wavelet transformation was used with GANN (ANN model which is optimized by GA) model.

Artificial Neural Networks

The different kinds of artificial neural network models have been utilized in many studies of hydrology and water resources management (Hamed et al. 2004; Ramirez et al. 2005). In this study, Levenberg-Marquardt back propagation algorithm was utilized for training of input data in ANN model. The precipitation data were normalized between 0 and 1. The architecture of ANN model which was used in this study, is 5-3-1 (5 inputs, 3 hidden neurons and 1 output neuron).

Genetic Algorithm

Genetic algorithm is one of the optimization algorithms and there are a lot of applications of genetic algorithms with data-driven models in order to develop the performances of hybrid models (Sedki et al. 2009; Nasseri et al. 2008). In this study, genetic algorithms were used so as to optimize initial variables (weights and biases) of the ANN for the improving the performance of (WGANN) model. For genetic algorithm analysis, roulette wheel was used as selection, furthermore, population size was chosen as 100, mutation function was selected as uniform (mutation probability rate=0.05) and crossover ratio was selected as 0.9.

Multiple Linear Regression

Multiple linear regression (MLR) is a method which is based on researching the relationship between a dependent variable and independent variables as illustrated in equation (5).

$$y=\alpha_1 x_1 + \alpha_2 x_2 + \dots\dots + \alpha_i x_i + \beta \quad (5)$$

In equation (5), y stands for the dependent variable, xi for the independent variables, β for the intercept and αi for the coefficients.

Assessment of Model Performance

For the purpose of evaluating the performance of models, correlation coefficient (R) and root mean square error (RMSE) statistical values were calculated. The formula of RMSE were indicated in equation (6).

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

In equation (6), N represents the number of the dataset, Pi_{obs} stands for observed monthly precipitation and Pi_{sim} for the simulated monthly precipitation.

Case Study

The case study was implemented in Antalya that is situated in Mediterranean region of Turkey. The monthly precipitation total data of Antalya was provided by Turkish State Meteorological Service. North Atlantic Oscillation (NAO) index was obtained from that website:http://www.cpc.ncep.noaa.gov/products/precip/CWlink/pna/norm.nao.monthly.b5001.current.ascii.table (Climate Prediction Center, National Weather Service, NOAA) and Southern Oscillation (SO) index was acquired from that website: http://www.cpc.ncep.noaa.gov/data/indices/soi (Climate Prediction Center, National Weather Service, NOAA). In this study, the data cover the period between March 1960 and November 2005. The data length is totally 548 months, %75 of the data (411 months) was utilized for training period and the rest of the data (137 months) for test period. Some of the statistical data (mean, standard deviation, coefficient of skewness, minimum and maximum values) concerning the precipitation of Antalya were indicated in Table I. Furthermore, observed precipitation data for either training or test period were also shown in Figure 1 and Figure 2, respectively.

Table I. Precipitation statistics of Antalya

Period	Mean (mm)	Standard Deviation (mm)	Minimum (mm)	Maximum (mm)	Skewness
Training	87.3	122.3	0	797.8	2.3
Test	107.7	156.2	0	907.2	2.3

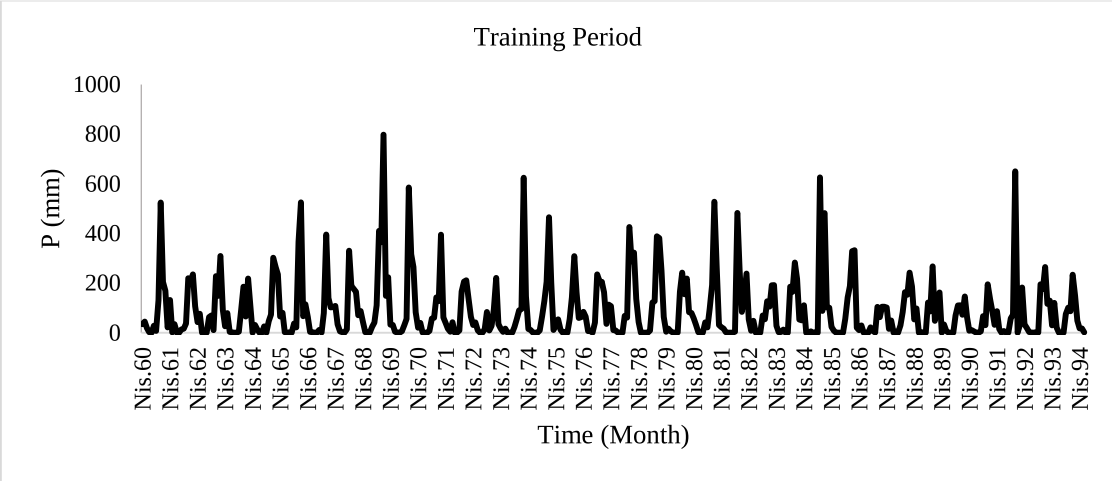


Figure 1. Observed monthly precipitation totals for training period

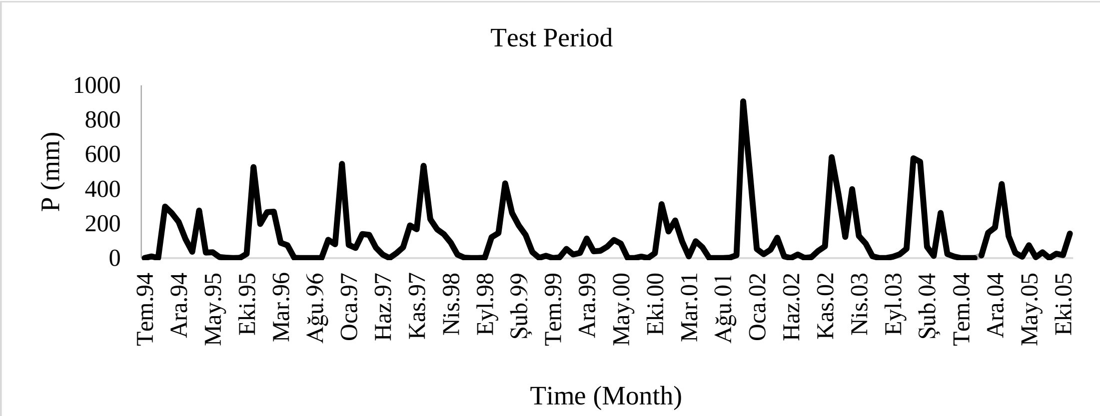
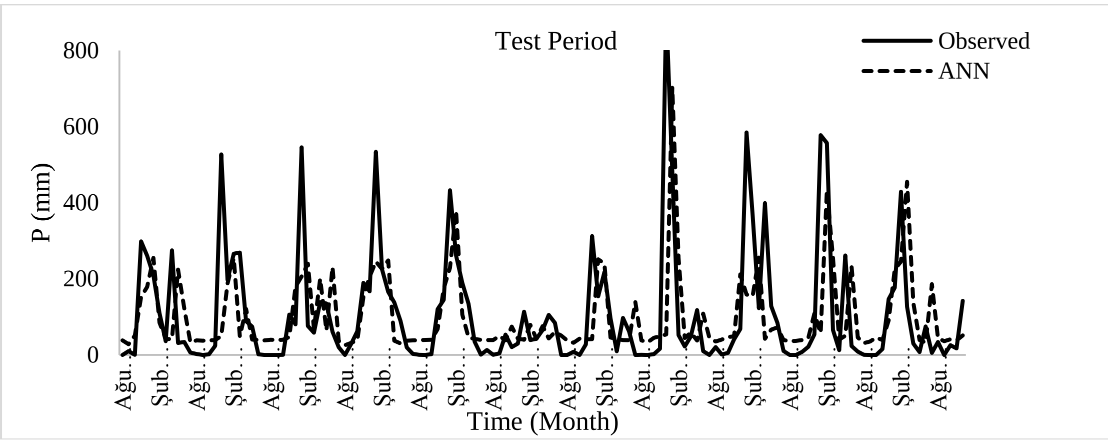


Figure 2. Observed monthly precipitation totals for test period

In order to make a prediction of the precipitation at that month (P(t)), NAO index preceding a month (NAO (t-1)), SO index preceding a month (SO (t-1)), the precipitation data preceding a month ((P (t-1)), the precipitation data preceding two months (P (t-2)) and the precipitation data preceding three months (P (t-3)) were used as input data for MLP, ANN, GANN and WGANN models. In other words, there are 5 inputs and 1 output in each model.

Results and Findings

The observed precipitation and conventional ANN model outcomes for test period were presented in Fig. 3.



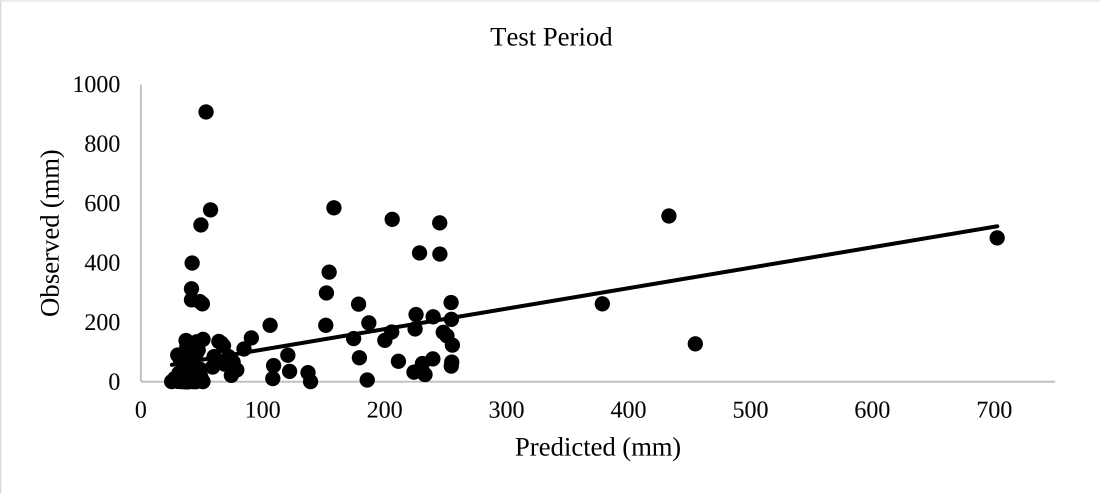
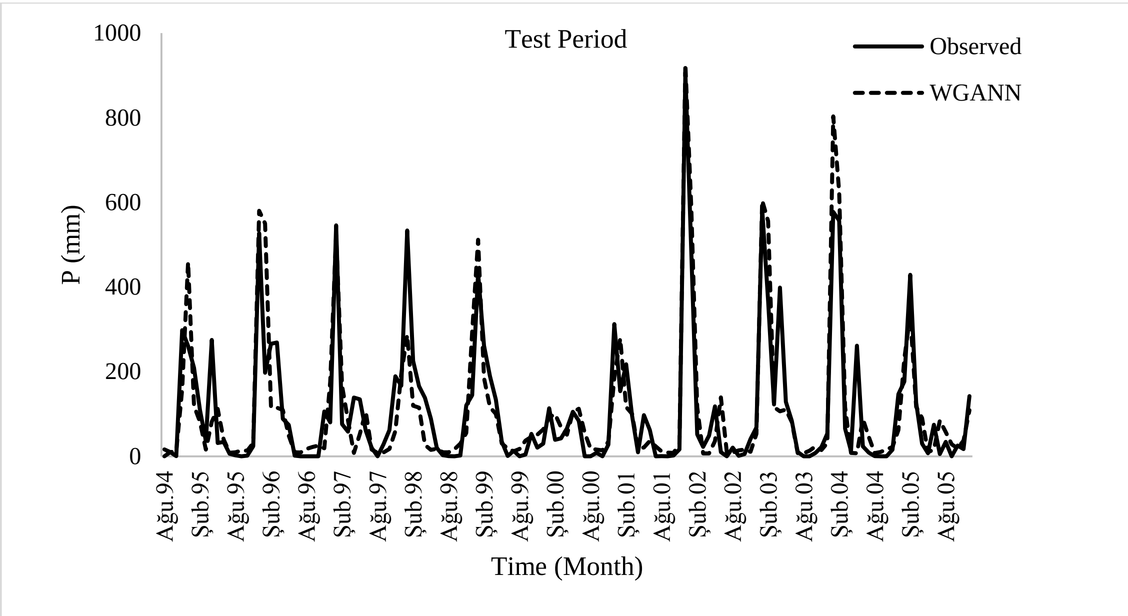


Figure 3. Comparison of observed and ANN model outcomes

The observed precipitation and WGANN model outcomes for test period were also presented in Figure 4. The observed precipitation data were decomposed into an approximation and seven details components using wavelet transformation. Then, the new precipitation data has constituted with the appropriate wavelet components. So, the new precipitation data were used as input for the genetic algorithm based neural network configuration.

The performances of all models (R and RMSE values) were demonstrated in Table II. According to Table II, it can be realized that MLR model is the least successful model among all of the models. Even though the performance of GANN model seems better than ANN model, their performances are close to each other. On the other hand as it can be also understood from Figure 4, WGANN model is far more accomplished than all other models. To illustrate, when Figure 3 and Figure 4 were taken into account, it can be seen that WGANN model is more overlapping with observed monthly precipitation than ANN model for test period. In addition, the R and RMSE values also proved the success of WGANN model as compared with ANN model.



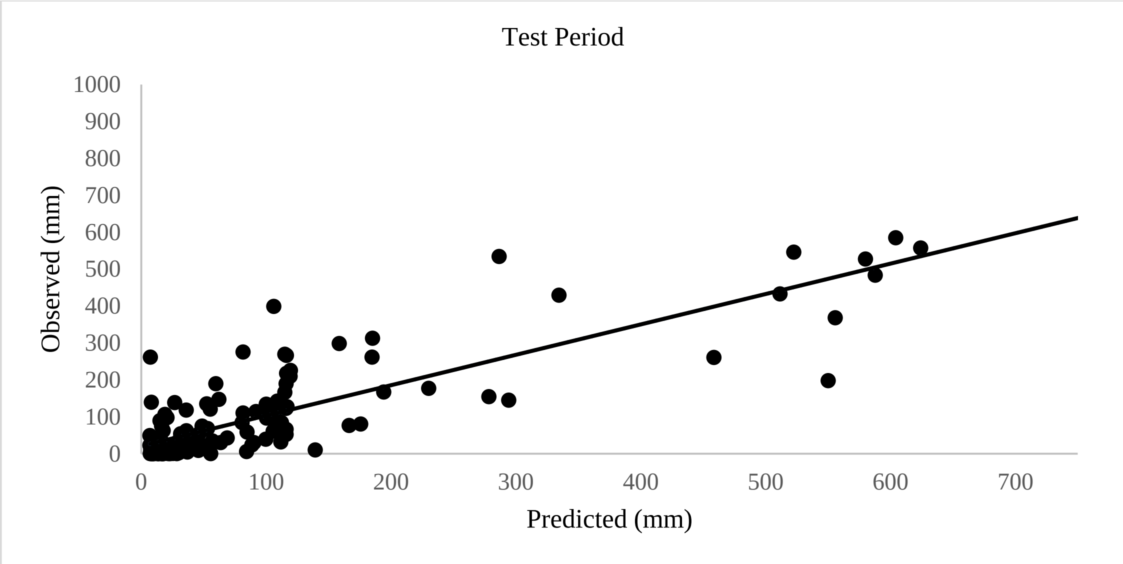


Figure 4. Comparison of observed and WGANN model outcomes

Table II. Performances of models for precipitation prediction

	MLR	**ANN**	**GANN**	**WGANN**
R	0.34	0.45	0.46	0.88
RMSE (mm)	149.5	143.3	141.3	80.7

Conclusion

In this study, monthly precipitation prediction of Antalya was carried out by using multiple linear regression (MLR), artificial neural network (ANN), the hybrid model of genetic algorithm- artificial neural network (GANN) and the hybrid model of wavelet transformation-genetic algorithm-artificial neural network (WGANN). In this context, as input data NAO index, SO index and precipitation data were used and performances of all models were compared. For comparison, correlation coefficient (R) and root mean square error (RMSE) were utilized. As a result of the analysis, it was obtained that WGANN model is the most successful model, whereas performance of MLR model is not quite accomplished for the estimation in comparison with other models. Besides, it was observed that GANN model performed a bit better than ANN model in this study. According to these results, it is concluded that the improvement of the hybrid models are very significant in terms of acquiring the more successful results concerning the estimation of hydrological variables. In this regard, further studies need to be done so as to observe the performances of distinctive models for the forecasting of hydrological parameters.

Recommendations

The authors would like to convey their thanks to Turkish State Meteorological Service for temperature and precipitation data.

References

Alp, M., & Cigizoglu, H. K. (2007). Suspended sediment load simulation by two artificial neural network methods using hydrometeorological data. Environmental Modelling & Software, 22(1), 2-13.

Chen, Y. H., & Chang, F. J. (2009). Evolutionary artificial neural networks for hydrological systems forecasting. Journal of Hydrology, 367(1), 125-137.

Daliakopoulos, I. N., Coulibaly, P., & Tsanis, I. K. (2005). Groundwater level forecasting using artificial neural networks. Journal of Hydrology, 309(1), 229-240.

Gao, C., Gemmer, M., Zeng, X., Liu, B., Su, B., & Wen, Y. (2010). Projected streamflow in the Huaihe River Basin (2010–2100) using artificial neural network. Stochastic Environmental Research and Risk Assessment, 24(5), 685-697.

Hamed, M. M., Khalafallah, M. G., & Hassanien, E. A. (2004). Prediction of wastewater treatment plant performance using artificial neural networks. Environmental Modelling & Software, 19(10), 919-928.

Kim, T. W., & Valdés, J. B. (2003). Nonlinear model for drought forecasting based on a conjunction of wavelet transforms and neural networks. Journal of Hydrologic Engineering, 8(6), 319-328.

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.

Meyer, Y., (1993). Wavelets algorithms & applications. Society for Industrial and Applied Mathematics, Philadelphia.

Nasseri, M., Asghari, K., & Abedini, M. J. (2008). Optimized scenario for rainfall forecasting using genetic algorithm coupled with artificial neural network. Expert Systems with Applications, 35(3), 1415-1421.

North Atlantic Oscillation (NAO) Index, Climate Prediction Center, National Weather Service, NOAA, (online). Retrieved from (10.06.2017) URL:http://www.cpc.ncep.noaa.gov/products/precip/CWlink/pna/norm.nao.monthly.b5001.current.ascii.table

Partal, T., & Cigizoglu, H. K. (2008). Estimation and forecasting of daily suspended sediment data using wavelet–neural networks. Journal of hydrology, 358(3), 317-331.

Partal, T., & Cigizoglu, H. K. (2009). Prediction of daily precipitation using wavelet—neural networks. Hydrological sciences journal, 54(2), 234-246.

Partal, T. (2017). Wavelet regression and wavelet neural network models for forecasting monthly streamflow. Journal of Water and Climate Change, 8(1), 48-61.

Ramirez, M. C. V., de Campos Velho, H. F., & Ferreira, N. J. (2005). Artificial neural network technique for rainfall forecasting applied to the Sao Paulo region. Journal of hydrology, 301(1), 146-162.

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.

Sedki, A., Ouazar, D., & El Mazoudi, E. (2009). Evolving neural network using real coded genetic algorithm for daily rainfall–runoff forecasting. Expert Systems with Applications, 36(3), 4523-4527.

Shoaib, M., Shamseldin, A. Y., & Melville, B. W. (2014). Comparative study of different wavelet based neural network models for rainfall–runoff modeling. Journal of hydrology, 515, 47-58.

Southern Oscillation (SO) Index, Climate Prediction Center, National Weather Service, NOAA, (online). Retrieved from (12.06.2017) URL: http://www.cpc.ncep.noaa.gov/data/indices/soi

Zorn, C. R., & Shamseldin, A. Y. (2015). Peak flood estimation using gene expression programming. Journal of Hydrology, 531, 1122-1128.