1. Introduction

The sustainable planning of water resources requires accurate estimation of hydrological data. The statistical methods are widely used in predicting hydrological data. Modern watershed management due to economic prosperity, increasing of population growing irrigation needs and industrial uses require more effective techniques or methods to predict future river flow correctly. Then, different methods such as artificial intelligence techniques have increasing trend in recent years (Awan and Bae, 2014; Xu et al., 2014; Valipour et al., 2013).

Hydrological data are generally expressed in time series form. One of the methods used in the analysis of time series is autoregressive (AR) models. The most widely used methodology for AR models in hydrology is Box-Jenkins methodology (Box et al., 1994, Keskin et al. 2006; Pektaş and Cigizoglu, 2013; Salas et al., 1980; Valipour, 2012; Valipour et al., 2013). Many
researchers also prefer artificial intelligence methods in the analysis of time series (Turan and Yurdusev, 2014; Yarar, 2014; Taylan, 2008; Keskin and Taylan, 2009; Chen et al., 2006, Tingsanchali and Gautam, 2000; Zadeh, 1965; See and Openshaw, 2000; Hundecha et al., 2001; Xiong et al., 2001).

Recently, hybrid models have begun to gain more and more attention. One of the hybrid models which used in hydrology is an integration of Artificial Neural Networks (ANN) and Fuzzy Logic (FL) modeling techniques. Hybrid models which has training and learning abilities of ANN models and writing rules facility (IF-THEN) of FL give more accurate results (Taylan, 2008; Chang and Chang, 2001). This hybrid model, called ANFIS (Adaptive Network Based Fuzzy Inference System), was firstly proposed by Jang (1992).

For flow prediction of Dalaman Stream in south west of Turkey different models were suggested. One of these models, developed by Taylan (2008), is AR-ANFIS model which its training sets were extended with synthetic series produced by autoregressive processes (AR). Taylan (2008) said that Dalaman stream flows were consistent with the AR (3) process and proposed that AR-ANFIS model which its prediction ability was better than single ANFIS model for Dalaman Stream.

In this study, it was proposed a hybrid model based on the ANFIS coupled with wavelet transform technique (Wavelet) for flow estimation of Dalaman Stream. The sub-series generated by wavelet transform techniques were used as input data of the ANFIS. It is seen that the new generated input data sets improves the results of models. Comparing the developed models, it was shown that the Wavelet-ANFIS models have the better prediction ability than the AR-ANFIS models. Consequently, Wavelet-ANFIS hybrid model could be used successfully in predicting of flow.

2. Scientific Literature Scanning

In hydrologic studies, data preprocessing techniques are also used to improve the performance of the developed artificial intelligence models. In this manner, many studies which combine the artificial intelligence methods with data preprocessing techniques, such as wavelet transform have been conducted in recent years. Wang et al. (2015) studied for predicting of runoff. They used decomposed annual runoff data. They showed that ensemble empirical mode decomposition (EEMD) can increase predicting ability and that the proposed EEMD-ARIMA (autoregressive integrated moving average) model could make better ARIMA modeling for runoff data predicting. Shafaei and Kişi (2015) studied on wavelet and ARMA, ANFIS and support vector regression (SVR) modeling techniques for predicting lake levels.

The results of the developed hybrid models by using wavelet were compared with single models. They showed that the hybrid models were good at forecasting lake levels. Badrzadeh et al. (2013) developed W-ANN and W-ANFIS models to show that data-preprocessing techniques increase the performance of flow forecasting models. Different combinations of models were developed for 1 to 5 days ahead flow forecasting in Harvey River located in the west of Australia. When the results of the hybrid models were compared with the original ANN and ANFIS models, it appeared that hybrid models showed much better performance. Kişi and Çimen (2011) developed a hybrid model which was combination of two methods, discrete wavelet transform and support vector machines (SVM). They compared this hybrid model with the single SVM model and they said that hybrid model increased the prediction ability of single SVM model. Nourani et al. (2009) investigated the effect of the combination of wavelet and ANN techniques for the estimation of rainfall in the Ligyanchai watershed located in Tabriz, Iran. Three different types of wavelet transforms were applied in the study and as a result, Haar wavelet gave lower performance when compared with db4 and Meyer wavelets.

In this study, a Wavelet-ANFIS hybrid model was suggested for flow estimation of Dalaman Stream, in the South west of Turkey. Therefore, study will provide support to literature.

3. Materials and Methods

3.1. Study Area and Data

The flow data of Dalaman Stream used in the development of the models was obtained from the General Directorate of Electrical Power Resources Survey Administration for 8-12 Akköprü Station in the south east of Turkey. The data used are monthly flow records from 1964 to 2003 years. Mean flow value of this station is 43.1m³/sn. The location of the Dalaman Stream and 8-12 Akköprü station was given in Fig. 1.
3.2. Suggested Methods

AR Models

Autoregressive (AR) models used in time series analysis have a very common usage. The influence of different parameters on the analysis of time series of hydrological records has great importance. Statistics and probabilities were used to determine the effect of these parameters.

The analysis of the time series is to build a model structure that fits the autoregressive process. The most widely used methodology for creating models is Box-Jenkins methodology (Box and Jenkins, 1970; Box et al., 1994). The aim in modeling is to develop a model that best represents the past and future of the time series.

The time series used for hydrology can be annual, seasonal, monthly or weekly. The structure for a time series in general is as follows:

\[ X_t = \sum a_i (X_{t-i}) + \epsilon_t \]  

Where \( t \) is the time, \( X \) is a variable, \( a \) is the coefficient indicating the relation of lagged values to the present value; and \( \epsilon \) is a residual part (DeLurgio, 1998).

The Adaptive Network-Based Fuzzy Inference System Model (ANFIS)

The ANFIS model, which consists of the use of fuzzy logic and artificial neural networks together, has the advantages of both artificial intelligence methods. That is, an ANFIS model combines the learning and computational power of neural networks with the logic structure of fuzzy control systems. In shortly, it has more transparency with the IF-THEN rule base and it is a self-adaptability modeling technique.

In fact, ANFIS is a Sugeno-type fuzzy model which has the ability of artificial neural network learning. An ANFIS structure consists of different layers and nodes (Tsoukalas and Uhrig, 1997)

Neuro-fuzzy control systems are based on Tagaki-Sugeno-Kang (TSK) fuzzy rules, which show a linear relationship between initial conditions and results. TSK fuzzy rules are as follows:

\[ R^1 : \text{IF } x_i \text{ is } A^i_1 \text{ AND } x_j \text{ is } A^j_1 \text{ THEN } y = f^1 = a^i_0 + a^i_1 x_1 + a^i_2 x_2 \]  

and

\[ R^2 : \text{IF } x_i \text{ is } A^i_2 \text{ AND } x_j \text{ is } A^j_2 \text{ THEN } y = f^2 = a^i_0 + a^i_1 x_1 + a^i_2 x_2 \]  

In a TSK fuzzy system, for given input values \( x_1 \) and \( x_2 \), the output \( y^* \) is calculated by the following equation:

\[ y^* = (\mu_i f^1 + \mu_j f^2) \mu_i + \mu_j \]  

where \( \mu_j \) is firing strengths of \( R_j \), \( j = 1, 2 \), given by the equation below,

\[ \mu_j = \mu_{\alpha_j}(x_i) + \mu_{\beta_j}(x_j) \]  

The layers and node functions are used in an ANFIS structure in which the product inference operation is used. In this ANN structure, each node in the first layer shows an input value. The nodes deliver the signals from the outside to the other layer. In the second layer, each node behaves like a membership function \( \mu_{A_i}(x_i) \), and it is selected as a bell-shaped curve with a maximum value of 1 and a minimum value of 0.

In third layer, every node is signed 1 and multiplies the incoming signals \( \mu_j = \mu_{\alpha_j}(x_i) + \mu_{\beta_j}(x_j) \) and sends the product out. Each node output indicates the firing strength of a rule. In the fourth layer, each node is named N and normalized by the firing strength of a rule. That is, jth node calculates the ratio of the firing strength of the jth rule to that of all the rules.

\[ \mu_j = \mu_j / \sum \mu_j \]  

in the fifth layer, each j-node gives the weighted consequent value

\[ \tilde{\mu}_j(a^0_i + a^i_1 x_1 + a^i_2 x_2) \]  

Where \( \tilde{\mu}_j \) is the output of fourth layer and \( (a^0_i + a^i_1 x_1 + a^i_2 x_2) \) is the parameter set to be edited. In the sixth layer, there is only one node and is demonstrated by \( \Sigma \), \( \Sigma \) shows the sum of the incoming signals to obtain the output of the entire system (Lin and Lee, 1995).

Discrete Wavelet Transform Technique (W)
Wavelet analysis is a multi-resolution analysis that depends on the number of repetitions and time. It has a big importance for Fourier transforms. The wave function \( \psi (t) \) is called the main wavelet and has shock characteristics and can be decreased to zero. It could be defined as \( \int_{-\infty}^{\infty} \psi (t) dt = 0 \). \( \Psi_{a,b} \) (t) could be obtained through compressing and expanding \( \psi (t) \):

\[
\Psi_{a,b} (t) = |a|^{-\frac{j}{2}} \psi \left( \frac{t-b}{a} \right) b \in R, a \in R, a \neq 0 \quad (9)
\]

Where \( \Psi_{a,b} (t) \) is successive wavelet; \( a \) is scale or frequency factor, \( b \) is time factor; \( R \) is the domain of real number.

If \( \Psi_{a,b} (t) \) satisfies equation (9), for the energy finite signal or time series \( f(t) \in L^2 (R) \), successive wavelet transform of \( f(t) \) is defined as

\[
W_\psi f (a, b) = \langle f, \Psi_{a,b} \rangle = \int_R f(t) \overline{\psi} \left( \frac{t-b}{a} \right) dt \quad (10)
\]

Where \( \overline{\psi} (t) \) is complex conjugate functions of \( \psi (t) \).

Equation (10) means that wavelet transform is the decomposition of \( f(t) \) under different resolution level (scale).

In fact, the consecutive wavelet is usually a discrete structure. If \( a = a_0^j, b = k b_0 a_0^j \), \( a_0 > 1 \), \( b_0 \in R \), \( k \) are integer number, discrete wavelet transform of \( f(t) \) is formed as follows:

\[
W_\psi f (j, k) = a_0^{-j/2} \int_R f(t) \overline{\psi} (a_0^{-j}t - kb_0) dt \quad (11)
\]

If \( a_0 = 2 \) and \( b_0 = 1 \), Equation (11) becomes binary wavelet transform:

\[
W_\psi f (j, k) = 2^{-j/2} \int_R f(t) \overline{\psi} (2^{-j}t - k) dt \quad (12)
\]

If \( W_\psi f (a, b) \) or \( W_\psi f (j, k) \) could represent the features of original time series in frequency (a or j) and time property (b or k) at the same time. If a or j becomes great, the frequency resolution of wavelet transform is high, but the time property resolution is low while if a or j is small, the frequency resolution of wavelet transform is low, but the time property resolution is high (Wang and Ding, 2003).

For a discrete time series \( f(t) \), when it consists at a different time \( t \), the discrete wavelet transform becomes

\[
W_f (m, n) = 2^{-m/2} \sum_{t=0}^{N-1} f(t) \psi^* (2^{-m} t - n) \quad (13)
\]

where \( W_f (m, n) \) is the wavelet parameter for the discrete wavelet of scale \( a = 2^m \) and location \( b = 2^n m, f(t) \) is a finite time series \( (t=0, 1, 2, ..., N-1) \), \( N \) is an integer power of 2 \( (N=2^m) \) and \( n \) is the time translation parameter, which varies in the range \( 0 < n < 2^M - 1 \), where \( 1 \leq m < M \), \( m \) and \( n \) are integers that run the wavelet scale/dilation and translation, respectively. At the greatest wavelet scale (i.e. \( 2^m \) where \( m=M \)) only one wavelet is required to cover the time interval and only one coefficient is produced. At the next scale \( (2^{m-1}) \), two wavelets cover the time interval, hence two coefficients are produced, and so on down to \( m=1 \). At \( m=1 \), the scale is \( 2^1 \), i.e. \( 2^{m-1} = 2^1 \) or \( N/2 \) coefficients are required to define the signal at this scale. The total number of wavelet parameters for a discrete time series of length \( N=2^m \) is then \( 1+2+4+8+...+2^m-1=N-1 \).

In addition to this, a signal smoothed component, \( \bar{W} \), is left, which is the signal mean. Thus, a time series of length \( N \) is broken into \( N \) components, i.e., with zero redundancy. The inverse discrete transform in a simple form is given by

\[
f(t) = \bar{W} + \sum_{m=1}^{M} W_m (t) \quad (14)
\]

Where \( \bar{W} \) is called the approximation sub-signal at level \( M \) and \( W_m (t) \) are detailed sub-signals at levels \( m=1, 2, ..., M \) (Tiwari and Chatterjee, 2011).

Discrete Wavelet Transform works on two sets of functions called scaling function (low pass filter) and the wavelet function (high pass filter). The signal/original time series data is passed through the low pass and high pass filters and subsequently separates the signal into approximation (\( A_m \)) and detail (\( D_i \)) components respectively. Then, this procedure is returned with successive approximations being decomposed in order, so that the signal is broken down into many lower resolution components as shown in Fig. 2. The low pass and high pass filters produce signals spanning only half the frequency band at all decomposition levels. This does the frequency resolution doubles because of fact that indefiniteness in frequency is reduced by half. The high pass filters are used to analyze the high frequencies while the low pass filters, on the other hand, are used to analyze the low frequency content of the signal (Shoaib et al., 2016).

![Figure 2](image)

**Figure 2.** n-level wavelet decomposition tree (Shoaib et al., 2016)

4. Results and Discussion
In this study, for flow prediction of Dalaman Stream, it was proposed an hybrid model (Wavelet-ANFIS) based on the ANFIS coupled with Wavelet transform technique (Wavelet). The results compared with AR-ANFIS model proposed by Taylan (2008). In order to estimate the flow of the Dalaman Stream, ANFIS models were formed using historical data and the synthetic series which was produced by autoregressive models by Taylan (2008). Taylan (2008) said that flow of Dalaman Stream was compatible with AR(3) and produced synthetic series with AR(3). Then, it was formed AR-ANFIS models by using historical data and synthetic series of Dalaman Stream. In modeling, the input layer was formed by using flows at times t-3, t-2, t-1 and periodicity components \( \cos(2\pi i/12), \sin(2\pi i/12) \) \( (i = 1, 2, ..., 12) \), and the output layer was selected a flow for time \( t \). For all of the data, the first 80\% were used during the training, while the remaining 20\% were used during the testing. The number of membership functions for each input of AR-ANFIS was investigated and it was selected as 3 for flows at time\( s\), t-2, t-1 inputs and 5 for \( \cos(2\pi i/12), \sin(2\pi i/12) \) inputs for optimum model structure. Membership function types were selected as Gaussian for inputs and as linear for output, respectively. The determination coefficient \( R^2 \) (Equation 15) and mean square error \( (\text{MSE}) \) (Equation 16) were calculated to evaluate the performance of the AR-ANFIS model. The \( R^2 \) and MSE values were found as 0.75 and 344 for training and 0.74 and 375 for testing, respectively (Taylan, 2008).

\[
R^2 = 1 - \frac{\sum_{i=1}^{N}(F_i(\text{measured}) - F_i(\text{model}))^2}{\sum_{i=1}^{N}(F_i(\text{measured}) - F_i(\text{mean}))^2}
\]

\[
\text{MSE} = \frac{1}{N} \sum_{i=1}^{N}(F_i(\text{measured}) - F_i(\text{model}))^2
\]

where \( N \) is the total number of data; \( F_i(\text{measured}) \) is measured flow value; \( F_i(\text{model}) \) is the result of the model; and \( F_i(\text{mean}) \) is the mean of measured flow value.

The scatter diagrams of the best AR-ANFIS model for the training and testing sets were given in Fig. 3. When the figures are analyzed, it can be seen that model results and measured values have a higher agreement in the AR-ANFIS model (Taylan, 2008).

Then, Wavelet-ANFIS models were developed. The discrete wavelet transform was applied to flow of Dalaman Stream. These flow values were decomposed into 8 detail components \( (2-4-8-16-32-64-128-256) \) \( (D1, D2, D3, D4, D5, D6, D7, D8) \) and one approximation component \( (A8) \). In the creation of sub-series, Haar which is the most commonly used wavelet in the discrete wavelet transform technique has been used. The correlations between the sub-series obtained by wavelet transform and output are calculated and given in Table 1.

**Table 1.** Correlation between sub-series and output

<table>
<thead>
<tr>
<th>Sub-series</th>
<th>t-3</th>
<th>t-2</th>
<th>t-1</th>
<th>( \cos(2\pi i/12) )</th>
<th>( \sin(2\pi i/12) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>D1</td>
<td>-0.046</td>
<td>-0.028</td>
<td>-0.044</td>
<td>0.005</td>
<td>0.165</td>
</tr>
<tr>
<td>D2</td>
<td>-0.128</td>
<td>-0.073</td>
<td>0.181</td>
<td>0.185</td>
<td>0.197</td>
</tr>
<tr>
<td>D3</td>
<td>0.066</td>
<td>0.374</td>
<td>0.528</td>
<td>0.202</td>
<td>0.454</td>
</tr>
<tr>
<td>D4</td>
<td>0.046</td>
<td>0.211</td>
<td>0.290</td>
<td>0.124</td>
<td>0.239</td>
</tr>
<tr>
<td>D5</td>
<td>0.129</td>
<td>0.159</td>
<td>0.183</td>
<td>0.052</td>
<td>0.049</td>
</tr>
<tr>
<td>D6</td>
<td>0.147</td>
<td>0.150</td>
<td>0.154</td>
<td>0.014</td>
<td>-0.100</td>
</tr>
<tr>
<td>D7</td>
<td>0.217</td>
<td>0.219</td>
<td>0.222</td>
<td>-0.005</td>
<td>0.055</td>
</tr>
<tr>
<td>D8</td>
<td>0.119</td>
<td>0.120</td>
<td>0.121</td>
<td>-0.062</td>
<td>-0.114</td>
</tr>
<tr>
<td>A8</td>
<td>0.230</td>
<td>0.230</td>
<td>0.230</td>
<td>0.230</td>
<td>0.230</td>
</tr>
</tbody>
</table>

The highest correlation values were found for the wavelet subseries D3, D4,D7 and A8 for t-1, t-2 input.
variables, for the wavelet subseries D6, D7 and A8 for t-3 input variable, for the wavelet subseries D2, D3, D4 and A8 for cos(2πi/12) and sin(2πi/12) input variables. Wavelet-ANFIS models have been developed by creating various combinations of these effective components. By using D2, D3, D4, D6, D7 and A8 wavelet sub-series, models with D3 (for inputs t-1, t-2, cos(2πi/12) and sin(2πi/12)) and D7 (for input t-3) sub-series generated the best correlation value for Wavelet-ANFIS models which were developed with 5 inputs. The results of the developed Wavelet-ANFIS models were given in the Table 2. In models, the inputs were formed by using flows at times t-3, t-2, t-1 and periodicity components cos(2πi/12), sin(2πi/12) (i = 1, 2, ..., 12), and the output was selected a flow for time t. But unlike the models in AR-ANFIS, for the first model, D3 sub-series were used instead of flows at time t-1. Similarly, for the third model, D7 sub-series were used instead of flows at time t-3. According to Table 2, different membership’s number was tried for each inputs of wavelet modeling. For all of the developed Wavelet-ANFIS models performed better than the model developed with AR-ANFIS model.

Table 2. The results of the Wavelet-ANFIS models

<table>
<thead>
<tr>
<th>Model</th>
<th>Number of membership functions</th>
<th>Training set</th>
<th>Testing set</th>
</tr>
</thead>
<tbody>
<tr>
<td>D3 for t-1</td>
<td>3</td>
<td>284</td>
<td>0.82</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>226</td>
<td>0.86</td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>213</td>
<td>0.87</td>
</tr>
<tr>
<td></td>
<td>9</td>
<td>150</td>
<td>0.91</td>
</tr>
<tr>
<td>D3 for t-2</td>
<td>3</td>
<td>288</td>
<td>0.79</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>216</td>
<td>0.85</td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>186</td>
<td>0.87</td>
</tr>
<tr>
<td></td>
<td>9</td>
<td>145</td>
<td>0.89</td>
</tr>
<tr>
<td>D7 for t-3</td>
<td>3</td>
<td>311</td>
<td>0.80</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>263</td>
<td>0.83</td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>232</td>
<td>0.85</td>
</tr>
<tr>
<td></td>
<td>9</td>
<td>156</td>
<td>0.90</td>
</tr>
<tr>
<td>D3 for cos(2πi/12)</td>
<td>3</td>
<td>463</td>
<td>0.71</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>379</td>
<td>0.76</td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>390</td>
<td>0.75</td>
</tr>
<tr>
<td></td>
<td>9</td>
<td>385</td>
<td>0.76</td>
</tr>
<tr>
<td>D3 for sin(2πi/12)</td>
<td>3</td>
<td>453</td>
<td>0.72</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>353</td>
<td>0.78</td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>350</td>
<td>0.78</td>
</tr>
<tr>
<td></td>
<td>9</td>
<td>350</td>
<td>0.78</td>
</tr>
</tbody>
</table>

Among these models, the highest determination coefficient was obtained for D3 wavelet of t-1 input model with 9 membership functions. The R² and MSE values were found as 0.90 and 156 for training and 0.92 and 84 for testing, respectively. The scatter diagrams of the best Wavelet-ANFIS model for the training and testing sets were given for D7 wavelet of t-3 inputs model in Figure 4. When the figures are analyzed, it can be seen that model results and measured values have a higher agreement in the W-ANFIS model than the AR-ANFIS model. Moreover, it can be said that Wavelet-ANFIS model has a better forecasting performance.

Figure 4. Scatter diagrams for the Wavelet-ANFIS model

5. Conclusion

In this study, flow estimation was performed using AR-ANFIS and Wavelet-ANFIS models for Dalaman Stream. For this, synthetic series, generated through autoregressive (AR) models, are used for training data sets of the ANFIS. Then, the sub-series, generated by wavelet models are used for training data sets of the ANFIS. The flow values of Dalaman Stream were decomposed into sub-series and the Wavelet-ANFIS models were developed using the efficient sub-series. When the Wavelet-ANFIS and AR-ANFIS models were compared, it was seen that the Wavelet-ANFIS model results have better agreement with real flow values.
Conflict of Interest
Any conflict of interest was not declared by author.

References


Valipour, M., 2012. Parameters estimate of autoregressive moving average and autoregressive integrated moving average models and compare


