Application of Soft Computing Methods in Predicting Evapotranspiration

Afshin Honarbakhsh¹, Mostafa Moradi Dashtpagerdi²*, Hassan Vagharfard³
¹Faculty of Natural Resources and Earth Sciences, University of Shahrekord, Shahrekord, Iran
²Graduated Watershed Management, Organization of Natural Resources and Watershed Management, Karoon Watershed Management Office (KWMO), Shahrekord, Iran
³Faculty of Natural Resources, Hormozgan University, Bandarabbas, Iran
Email: *moradi2763@yahoo.com

Received May 21, 2013; revised June 20, 2013; accepted June 28, 2013

ABSTRACT

Exact prediction of evapotranspiration is necessary for study, design and management of irrigation systems. In this research, the suitability of soft computing approaches namely, fuzzy rule base, fuzzy regression and artificial neural networks for estimation of daily evapotranspiration has been examined and the results are compared to real data measured by lysimeter on the basis of reference crop (grass). Using daily climatic data from Haji Abad station in Hormozgan, west of Iran, including maximum and minimum temperatures, maximum and minimum relative humidities, wind speed and sunny hours, evapotranspiration was predicted by soft computing methods. The predicted evapotranspiration values from fuzzy rule base, fuzzy linear regression and artificial neural networks show root mean square error (RMSE) of 0.75, 0.79 and 0.81 mm/day and coefficient of determination of (R²) of 0.90, 0.87 and 0.85, respectively. Therefore, fuzzy rule base approach was found to be the most appropriate method employed for estimating evapotranspiration.

Keywords: Evapotranspiration; Fuzzy Rule Base; Fuzzy Regression; Artificial Neural Network

1. Introduction

Evapotranspiration is one of the most important factors in agriculture and the hydrological cycle that can be influenced by global warming and climatic changes [1]. The process of evapotranspiration (ET) is an important part of the water cycle and exactly estimating the value of ET is necessary for designing irrigation systems and water resources management. Accurate estimation of ET is crucial in agriculture. This is due to the fact that its over-estimation causes waste of valuable water resources and its underestimation leads to the plant moisture stress and decrease in the crop yield. Prediction methods developed over the past few decades range from simple ones such as Blaney-Criddle to complex ones which use physical processes like Penman compound method [2]. Penman approach used parameters such as dynamic of evaporation, intensity of net radiations and surface aerodynamic characteristics. Montieth et al. (1965) later improved this method by considering the plant daily resistance and Penman-Montieth equation [3]. Several researchers studied validation of these equations [4,5]. Jenson et al. (1990) compared results of twenty of such methods with the results of lysimeters in 11 stations located in different parts of the world with various climates and concluded that in all climates the Penman-Montieth method gave the best results [6]. In recent years, soft computing methods including fuzzy rule base model (FRBM), artificial neural Networks (ANN) and also a combination of them have been employed for estimating evapotranspiration.

Burton et al. (2000) used ANNs and estimated daily evaporation from pan evaporation by 2044 data gathered from various places all over the world from 1992 to 1996 [7]. Input data were precipitation, temperature, relative humidity, solar irradiance and wind speed. Compared with multiple linear regressions methods such as the one proposed by Priestley-Taylor (1972), ANN provided the minimum error of 1.11 mm/day in ET estimation [8]. Odhiambo et al. (2001) compared the results from FRBM with those of Penman-Montieth [9] and Shayannejad et al. (2007) used Fuzzy linear regression (FLR) for ET estimation in Hamadan, Iran and demonstrated...
that FRL gave higher determination coefficient ($R^2$) with less errors than Penman-Montieth method [10]. Also, Hargreaves-Samani methods (1994, 1985) used two fuzzy rule base models, in which solar irradiance and relative humidity were the input data in the first model (FRBM-1) and wind speed was also added in the second model (FRBM-2) [11-13]. A Comparison with the lysimeter data showed that the standard errors for FRBM-1, FRBM-2 models, Penman-Montieth and Hargreaves-Samani were 0.73, 0.54, 0.50 and 0.66 mm per day respectively. It can be seen that FRBM-2 and Penman-Montieth yield similar errors despite the fact that the number of input parameters was less in FRBM-2.

2. Materials and Methods

The necessary climatic data for this research were provided from Haji Abad meteorological station, near Hormozgan, west of Iran. This station has longitude 55° and 55° North, and latitude 28° and 19° East, and elevation of 870 m above sea level. The climate can be described as arid and hot according to Kopen’s classification. Maximum and minimum daily air temperature is 49°C and 5°C respectively. The average annual rainfall during the period of 2000-2008 was 265 mm. A 1 m × 2.25 m × 1.2 m lysimeter equipped with drainage is used to measure ETP with grass reference crop. The soil characteristics could be described as: alkaline, deep, medium to heavy texture, electric conductivity of 0.55 to 0.85 deci siemens per meter, specific gravity of 1.63 - 1.91 gram per cm$^3$. A layer of 27 cm thickness gravel consisting of various sizes covered the slopped bottom of the lysimeter at the station and soil was added in separate horizontal layers. Daily ET was obtained using water balance model measuring water input and output and soil humidity.

In this study, three soft computing approaches namely, FRBM, FLR and ANN were used to estimate the potential evapotranspiration and they were evaluated using the lysimeter data.

2.1. Fuzzy Rule Base

Fuzzy rule-based models developed by Lotfizadeh (1965) for handling imprecise information, has found important application in various fields including water based systems in the last five decades [14]. Introduction of Linguistic Terms (LT) by Fontane et al. (1997) and application of complex mathematical models by Bárdossy et al. (1995), Pesti et al. (1996) have established this methodology as a reliable tool for predicting water resource parameters [15-17]. A FRBM contains membership functions of fuzzy sets constructed on the range of all the inputs to the model. The model matches the input and output, which also contains membership functions, with fuzzy rules. In this study, as suggested by Bárdossy and Duckstein (1995), following a local search on the four available membership functions of triangular, bell-shaped, dome-shaped and inverted cycloid, the triangular input membership function was selected based on the lowest root square mean error (RSME) of 0.75 and highest $R^2$ of 0.90 as shown in Table 1 [18].

2.2. FRBM Design

In the design of the FRBM, six inputs containing min. and max daily temperature ($T_{min}$, $T_{max}$), min. and max. Daily air relative humidity ($R_{h_{min}}$, $R_{h_{max}}$), daily wind speed (U), daily sunny hours (N), were considered and ET was the model output. In order to establish the rule-bases, 40 lines of the data containing inputs and outputs were selected randomly.

Five FRBM models (FRBM-1 to FRBM-5) were defined based on the quantity of linguistic terms and also, the type and number of input parameters mentioned above (see Table 2). Using 6 similar input parameters, FRBM-1, FRBM-2 and FRBM-3 have been defined with 2, 3 and 5 LT respectively, and as suggested by Figures 1-6, FRBM-3 with 5 LT showed the least RMSE of 0.75. FRBM-4 and FRBM-5 were hence defined using 5 LT but different types and number of input parameters. Based on the results demonstrated in Table 2, FRBM-3 with lowest RMSE, with input triangular membership function and 5 LT was selected as the best FRBM for this study.

2.3. Artificial Neural Network Method

ANNs are mathematical models consisting of highly interconnected processing nodes or elements (artificial neurons) under a pre-specified topology (sequence of layers or slabs with full or random connections between the layers). In 1950s Rosenblatt built many variations of a specific type of early neural computational models called perceptron network and developed associated learning rules which led to introduction of first practical

<table>
<thead>
<tr>
<th>Number</th>
<th>Membership Function Type</th>
<th>RMSE</th>
<th>$R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>TRI-MF</td>
<td>0.75</td>
<td>0.90</td>
</tr>
<tr>
<td>2</td>
<td>TRAP-MF</td>
<td>1.15</td>
<td>0.831</td>
</tr>
<tr>
<td>3</td>
<td>GBELL-MF</td>
<td>1.27</td>
<td>0.768</td>
</tr>
<tr>
<td>4</td>
<td>GAUSS1-MF</td>
<td>1.94</td>
<td>0.82</td>
</tr>
<tr>
<td>5</td>
<td>GAUSS2-MF</td>
<td>1.49</td>
<td>0.785</td>
</tr>
</tbody>
</table>

A. HONARBAKHSH ET AL. 399

Figure 1. Membership function, model FRBM-1, with 2 linguistic terms.

Figure 2. RMSE for model FRBM-1, with 2 linguistic terms.

Figure 3. Membership function, model FRBM-2, with 3 linguistic terms.

Figure 4. RMSE for model FRBM-2, with 3 linguistic terms.

Figure 5. Membership function, model FRBM-3, with 5 linguistic terms.

Figure 6. RMSE for model FRBM-3, with 5 linguistic terms.

application of ANN. They have been used extensively since 1980’s in a variety of diverse real world applications [19]. In this work, the multi-layer perceptron network has one input layer (with three processing elements), one hidden layer (with two processing elements) and one output layer (with one processing element).

2.4. Fuzzy Linear Regression

In regression analysis, the best mathematical expression describing the functional relationship between one response and one or more independent variables are obtained. Following the introduction of the fuzzy theory, by Lotfizadeh, fuzzy regression model (FRM) was developed by Tanaka et al. (1982) in which fuzzy uncertainties of dependent variables with the fuzziness of response functions were explained [20]. Based on the conditions of variables, there are 3 categories of FRM: 1) input and output data are both non-fuzzy numbers, 2) input data is non-fuzzy number but output data is fuzzy number, and 3) input and output are both non-fuzzy number [21,22] Estimation of fuzzy regression is the subject of continuous
research, is often carried out by two techniques, e.g.: fuzziness minimization by numerical method using linear programming (as suggested by Tanaka, 1982) and deviation minimization between the estimated and observed outputs, sometimes referred to as fuzzy least square method [23]. FLR has been used where response variable is in intervals. By taking mean or mode, interval value can be changed to crisp values but at a cost of losing useful information about the spread. Hence, no proper interpretation of the fuzzy regression interval can be made (Wang and Tsaur, 2000) Tanaka’s approach [24], referred to as possibilistic regression has also been criticized for both not being based on sound statistical principles [25], as well as creating computational difficulties when large number of data points is encountered [26]. Peters (1994) complains about Tanaka’s model being extremely sensitive to the outliers [27]. Kim et al. (1996) reported that fuzzy linear regression (FLR) may tend to become multicollinear as more independent variables are collected [28]. The drawback, on the other hand, with the fuzzy least square method is the spread of estimated response increases as the magnitude of explanatory response increases, even though the spread of observed responses are roughly constant or decreasing. To overcome this, Kao and Chyu (2002) proposed a “two-stage” approach for fitting fuzzy linear regression (FLR) through fuzzy least square approach and showed superiority over Diamond’s procedure [29]. This approach is discussed by Singh et al. (2007) and relevant nonlinear computer programs such as LINGO, have been developed to solve such cases [30]. As far as fuzzy nonlinear regression is concerned, Buckley and Feuring (2000) proposed “evolutionary algorithm solutions” in which for a given fuzzy data, algorithm searches from a library of fuzzy functions (including linear, polynomial, exponential and logarithmic) one which would fit the data [31]. In this study, using HYDROGENERATOR and LINGO softwares, a fuzzy possibilistic model was employed in which coefficients are fuzzy, while inputs and outputs are non-fuzzy observational. The model used may be represented by the following equation:

$$\tilde{y} = \tilde{A}_0 + \tilde{A}_1 x_1 + \tilde{A}_2 x_2 + \tilde{A}_3 x_3 + \cdots + \tilde{A}_n x_n$$

(1)

where, $\tilde{A}_0, \tilde{A}_1, \ldots, \tilde{A}_n$ are fuzzy coefficients and $x_1, x_2, x_3, \cdots, x_n$ are observational input variables which are normal numbers and $\tilde{y}$ is the fuzzy output for each variable $n$. Table 3 shows the object function and the restrictions used for the FLR in this work.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>FRBM-1</th>
<th>FRBM-2</th>
<th>FRBM-3</th>
<th>FRBM-4</th>
<th>FRBM-5</th>
</tr>
</thead>
<tbody>
<tr>
<td>minimum temperature</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Maximum temperature</td>
<td></td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>minimum humidity</td>
<td>*</td>
<td></td>
<td>*</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>maximum humidity</td>
<td>*</td>
<td></td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>wind speed</td>
<td>*</td>
<td></td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>sunny hour</td>
<td>*</td>
<td></td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>mean relative humidity</td>
<td></td>
<td></td>
<td>*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>mean temperature</td>
<td>*</td>
<td></td>
<td></td>
<td></td>
<td>*</td>
</tr>
</tbody>
</table>

RMSE mm/day

<table>
<thead>
<tr>
<th>Input Parameters</th>
<th>FRBM RMSE (mm/day)</th>
<th>ANN RMSE (mm/day)</th>
<th>FRM RMSE (mm/day)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$T_{\text{min}}$, $T_{\text{max}}$, $R_{\text{Hmin}}$, $R_{\text{Hmax}}$, $U$, $n$</td>
<td>0.75</td>
<td>0.81</td>
<td>0.79</td>
</tr>
<tr>
<td>$T_{\text{min}}$, $T_{\text{max}}$, $R_{\text{Hmin}}$, $R_{\text{Hmax}}$, $n$</td>
<td>0.97</td>
<td>0.81</td>
<td>0.96</td>
</tr>
<tr>
<td>$T_{\text{min}}$, $T_{\text{max}}$, $R_{\text{Hmin}}$, $R_{\text{Hmax}}$, $U$</td>
<td>0.88</td>
<td>0.89</td>
<td>0.97</td>
</tr>
<tr>
<td>$T_{\text{min}}$, $T_{\text{max}}$, $R_{\text{Hmin}}$, $U$, $n$</td>
<td>0.91</td>
<td>0.80</td>
<td>0.83</td>
</tr>
<tr>
<td>$T_{\text{min}}$, $T_{\text{max}}$, $R_{\text{Hmin}}$, $U$, $n$</td>
<td>1.08</td>
<td>0.91</td>
<td>0.95</td>
</tr>
<tr>
<td>$T_{\text{min}}$, $R_{\text{Hmin}}$, $R_{\text{Hmax}}$, $U$, $n$</td>
<td>1.23</td>
<td>1.43</td>
<td>0.98</td>
</tr>
<tr>
<td>$T_{\text{max}}$, $R_{\text{Hmin}}$, $R_{\text{Hmax}}$, $U$, $n$</td>
<td>1.64</td>
<td>1.22</td>
<td>1.09</td>
</tr>
</tbody>
</table>
Table 3: Linear programming model for solving linear regression with non-fuzzy observations.

**Fuzzy a linear regression:**

\[
\hat{y} = \tilde{A}_0 + \tilde{A}_1 x_1 + \tilde{A}_2 x_2 + \tilde{A}_3 x_3 + \cdots + \tilde{A}_n x_n
\]

**Function:**

\[
\text{Minimize: } mc_0 + \sum_{j=1}^{m} \sum_{i=1}^{n} c_i \left| y_i^j \right|
\]

(2)

\[
p_0 + \sum p_j x_j - (1-h) \left[ c_0 + \sum c_i x_i \right] \leq y_j
\]

(3)

\[
p_0 + \sum p_j x_j - (1-h) \left[ c_0 + \sum c_i x_i \right] \geq y_j
\]

(4)

3. Results and Discussion

For calculating ET in Penman-Mantis and Fuzzy regression methods, Excel and MATLAB softwares were applied, respectively. It is noted that RMSE and R² were used for validation and approval of the results.

**Sensitivity Analysis**

A sensitivity analysis was required to indicate which one of the input parameters has more important role on defining the ET in the models. This is carried out in two following methods: addition of input parameters and removal of input parameters. Accordingly, each parameter its addition or removal causes the most reduction in RMSE would be identified as the most sensitive parameter. In this work, using the latter approach, one of the six input parameters was removed at a time and the corresponding RMSE was calculated as shown in Table 3. Minimum temperature was therefore found to be the most sensitive parameter in all methods used while, the sunny hour showed the least sensitivity in FRBM, and maximum relative humidity was the least sensitive for ANN and FRM.

RMSE and R² were used to select the best method to determine ET amongst FRBM, ANN and FRM. As can be seen from Table 4, the results indicate that R² does not vary much (0.80 to 0.90), while RMSE alters more so that the least RMSE relates to FRBM model with five linguistic terms (FRBM-3), followed by ANN, FRM, FRBM-1, FRBM-2, FRBM-4 and FRBM-5, which showed higher RMSE (RMSE altered in the range of 0.75 to 1.18).

4. Conclusion

In this study, Soft Computing Methods in predicting evapotranspiration were reviewed in west of Iran. Considering Figures 7-9 in which the observed and estimated ET is demonstrated using the three models FRM, FRBM and ANN, fuzzy rule-based model proved to be the best method and is proposed to be used for ET estimation of the region. Iran loses 70% of annual precipitation by ET. It is obvious that in this country where there are many limitations to water resources management, increase in ET could lead to more problems [1]. Also, Water consumption was estimated 2200 m³ per person.
in 1990 in Iran, while it is predicted that in 2025 there will be 726 to 860 m³ [32]. Therefore, prediction of affecting factors such as evapotranspiration is necessary.

REFERENCES


http://dx.doi.org/10.1007/BF03326041