Science, Technology and Development 34 (2): 90-93, 2015
ISSN 0254-6418 / DOI: 10.3923/std.2015.90.93

Evaluation of Heuristic Models Capabilities for Estimating Reference Evapotranspiration (Case Study: Tabriz and Marageh Stations)

[1]Ali Firouzi, [2]Sepideh Karimi and [2]Jalal Shiri
[1]Sama Technical and Vocational Training College, Islamic Azad University, Tabriz Branch, Tabriz, Iran
[2]Department of Water Engineering, Faculty of Agriculture, University of Tabriz, Tabriz, Iran

Abstract: A modeling study is reported here for simulating daily reference evapotranspiration (ET_0) values in Tabriz and Marageh weather stations. Daily meteorological variables including air temperature, solar radiation, humidity and wind speed covering a period of 7 years are employed for developing and validation the ET_0 simulation models. The heuristic Gene Expression Programming (GEP) and Neuro-Fuzzy (NF) models were applied for simulating ET_0 values. The obtained results showed that radiation-based models produce the most accurate results among the applied input configurations.

Key words: Evapotranspiration, genetic programming, neuro-fuzzy, modeling

INTRODUCTION

Evapotranspiration (ET) is the combined processes of water loss from the bare soil and crop canopy; transpiration from the canopy to the atmosphere. In many arid and semi-arid areas where water resources are scarce and seriously endangered by overexploitation, the precise estimation of ET becomes imperative in the planning, management and scheduling of irrigation practices (Shiri and Kisi, 2011). Knowledge of the ET is also essential for analyzing water balances at the land surface, which is important to calculate the drainage requirements (De Ridder and Boonstra, 1994). Temperate areas are mainly concerned with the storage of water in dry seasons for the different uses over the year (Kisi, 2006). Also continuous simulation hydrological models generally requires ET amount for necessary analysis (Kay and Davies, 2008). Despite this significance, ET is one of the less understood components of the hydrological cycle (Brutsaert, 1982). Several attempts stated the needs for accurate estimation of ET in order to predict crop water requirement e.g., Hargreaves and Samani (1985), Makkink (1957) and Priestley and Taylor (1972). The term reference ET (ET_0) was introduced by the United Nations Food and Agriculture Organization (FAO) as a methodology for computing crop evapotranspiration (Doorenbos and Pruitt, 1977), because the interdependence of the factors affecting the ET makes the study of the evaporative demand of the atmosphere regardless of crop type, its stage of development and its management difficult. The reference evapotranspiration represents the evapotranspiration from a hypothesized reference crop (height 0.12 m, surface resistance 70 sec m^{-1} and albedo 0.23) (Allen *et al*., 1998). The adopted Penman-Monteith equation by FAO (which will be referred to as FAO56-PM model, in short) has two important advantages (Landeras *et al*., 2008): (i) It can be applied in a great variety of environments and climate scenarios without local calibration and (ii) It has been validated using lysimeters under a wide range of climatic conditions. However, the need for large number of climatic variables (e.g., air temperature, relative humidity, solar radiation and wind speed) is the major drawback of this model.

As an alternative, the heuristic models might be applied for mapping nonlinear inter-relationships of the weather parameters and ET_0. Among the others, the Gene Expression Programming (GEP) and Neuro-Fuzzy (NF) models have been applied for modeling ET_0 in different climatic contexts e.g. (Marti *et al*., 2015; Shiri *et al*., 2014a, b, c, 2015). The present study aims at ET_0 modeling in Tabriz and Marageh stations using GEP and NF techniques fed with similar input configurations of the well-known semi empirical approaches.

MATERIALS AND METHODS

FAO56-PM standard model: The Penman-Monteith equation modified by Allen *et al*. (1998) reads:

$$ET_0 = \frac{0.408\Delta(R_n - G) + \gamma \frac{900}{T_{mean} + 273} S_W (e_a - e_d)}{\Delta + \gamma(1 + 0.34U_2)} \quad (1)$$

where, ET_0 is the reference evapotranspiration (mm day^{-1}), Δ is the slope of the saturation vapor pressure function (kPa/°C), γ is the psychometric constant (kPa/°C), R_n is the net radiation (MJ m^{-2} day^{-1}), G is

Corresponding Author: Jalal Shiri, Department of Water Engineering, Faculty of Agriculture, University of Tabriz, Tabriz, Iran

the soil heat flux density (MJ m^{-2} day^{-1}), T_{mean} is the mean air temperature (°C), U_2 is the average 24 h wind speed at 2 m height (m sec^{-1}), e_a is the saturation vapor pressure (kPa) and e_d is the actual vapor pressure.

Hargreaves-samani equation: The Hargreaves Model (Hargreaves and Samani, 1985) reads:

$$ET_0 = 0.0023R_a\left(\frac{T_{max}+T_{min}}{2}+17.8\right)\sqrt{T_{max}-T_{min}} \qquad (2)$$

where, ET_0 is the reference evapotranspiration (mm day^{-1}), R_a is the terrestrial radiation (mm day^{-1}), T_{max} is the maximum air temperature (°C) and T_{min} is the minimum air temperature (°C).

Priestly-taylor equation: The Priestley-Taylor model (Priestley and Taylor, 1972) for computing ET_0 value is:

$$ET_0 = \frac{\alpha}{\lambda}\frac{\Delta}{\Delta+\gamma}(R_n - G) \qquad (3)$$

where, ET_0 is the reference evapotranspiration (mm day^{-1}), α is the 1.26 and λ is the latent heat of the evaporation (MJ kg^{-1}).

Gene expression programming: A generic programming procedure starts by random generation of chromosomes of the certain program (initial population), then the generated chromosomes are expressed and the fitness of each individual program is evaluated against a set of fitness cases (Ferreira, 2006). The programs are then selected according to their own fitnesses (their performance in that particular environment). The mentioned process is repeated until a good solution can be found for the studied phenomenon.

The application of GEP involves the next general steps:

- Selection of fitness function
- Choosing the set of terminals T and the set of functions F to create the chromosomes
- Choosing the chromosomal architecture
- Choosing the linking function
- Choosing the genetic operators

In the present work, the GeneXpro program was applied for modeling sea water levels. The Parsimony Pressure tool was applied to penalize the parse trees of each GEP model for condensing the model's expressions and avoiding from producing the nested functions.

Neuro-fuzzy systems: An Adaptive Neuro-Fuzzy Inference System (ANFIS) is a combination of an

Table 1: Statistical parameters of the applied data set

Parameters	C_{SX}	C_V	C_X	X_{mean}	X_{min}	X_{max}
Tabriz						
T_{mean} (°C)	-0.13	0/83	10/7	12/9	-12/2	34
T_{max} (°C)	-0.12	0/61	11/5	18/8	-6/4	41
T_{min} (°C)	-0.15	1/27	9/8	7/6	-16/8	27/6
H_R (%)	0.20	0/34	17/5	50/8	13/9	96/3
W_S (m sec^{-1})	1.05	0/44	1/5	3/5	0/3	13/3
S_R (MJ m^{-2} day^{-1})	-0.01	0/42	6/5	15/5	1/1	29/6
ET_0 (mm)	0.85	0/83	4/9	5/8	0/07	25/6
Marageh						
T_{mean} (°C)	-0/11	0/80	10/4	13	-13/5	32/8
T_{max} (°C)	-0/14	0/59	11/3	19/1	-7	40/4
T_{min} (°C)	-0/15	1/18	9/4	7/9	-20/6	26/6
R_H (%)	0/24	0/37	18/7	50/3	13/5	94/6
W_S (m sec^{-1})	0/66	0/46	1/6	3/6	0	10/8
S_R (MJ m^{-2} day^{-1})	-0/03	0/43	6/6	15/5	0	28/7
ET_0 (mm)	0/76	0/83	5/1	6/1	0	23/8

Adaptive Neural Network (ANN) and a fuzzy inference system. The parameters of the fuzzy inference system are determined by the neural network learning algorithms. Since this system is based on the fuzzy inference system, reflecting amazing knowledge, an important aspect is that the system should be always interpretable in terms of fuzzy IF-THEN rules. The ANFIS is capable of approximating any real continuous function on a compact set of parameters to any degree of accuracy (Jang, 1993). The ANFIS identifies a set of parameters through a hybrid learning rule combining back propagation gradient descent error digestion and a least squared error method.

Studied regions: Daily weather data from two weather stations, namely, Tabriz and Marageh were applied in the present study, covering a period of 7 years (2003-2009). The data comprises daily values of air temperature (T_A), relative humidity (H_R), wind speed (S_W) and total incoming radiation (S_R). Table 1 sums up the statistical parameters of the applied data. In the table, the X_{mean}, X_{max}, X_{min}, S_X, C_V and C_{SX} denote the mean, maximum, minimum, standard deviation, coefficient of variation and skewness coefficient, respectively.

Goodness of fit measures: Two statistical evaluation criteria were used to assess the model's performances: The correlation coefficient (R^2):

$$R^2 = \left(\frac{\sum_{i=1}^{n}(ET_0 - meanET_0)(ET_M - meanET_M)}{\sqrt{\sum_{i=1}^{n}(ET_0 - meanET_0)^2\sum_{i=1}^{n}(ET_M - meanET_M)^2}}\right)^2 \qquad (4)$$

And, the Root Mean Squared Error (RMSE):

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(ET_M - ET_0)^2}{n}} \qquad (5)$$

where, ET_M and ET_0 denote the values generated by different models and FAO56-PM ET_0 values, respectively. Also mean ET_0 and mean ET_M stand for the mean ET values, estimated by FAO56-PM and applied models, respectively.

RESULTS AND DISCUSSION

ET_0 values produced by FAO56-PM equation were considered as reference values for calibrating and evaluating all the applied models, which is a common process in these studies (Allen *et al*., 1998). In similarity to the introduced models, the input combinations of heuristic models (GEP and NF) evaluated in the current work is as follows:

- T_{mean}, T_{max}, T_{mean}, R_a
- T_{mean}, R_S
- T_{mean}, R_H

The input combinations (i)-(iii) were developed for valid comparison with the Hargreaves-Samani and Priestley-Taylor equations, respectively.

Models implementation: For a given input-output dataset [for instance, modeling ET_0 (as output) by using weather variables (as inputs)], various Sugeno's FIS models may be developed by using different identification methods (i.e., grid partitioning, subtractive clustering and Gustafson-Kessel clustering methods). However, since the selection of each identification method does not affect the obtained results significantly, the commonly used grid partitioning identification method was applied for constructing the Generalized Neuro-Fuzzy (GNF) models in the present study. Table 2 represents the statistical criteria for the training period of the applied models. The table states that the both methodologies have high accuracy for modeling ET_0 values.

Models testing: Table 3 sums up the testing statistics of the GEP and ANFIS models. From the table it is seen that the radiation-based heuristic models are superior with lower error magnitudes. Nevertheless, it is noted that the risk of over fitting is still remaining because of two block partitioning data management scenario, as well as the potential error sources due to the application of calculated targets because of lack of lysimetric data. As could be fore shadowed, the radiation-based models produced the most accurate results, while the temperature based models (combinations i and iii) offer higher scattering in simulation-emulation procedure. Owing that the studied region is located in a semi-arid environment with the potential effect of the latitude on the total received solar radiation, the effect of aerodynamic component on the simulated ET values would be minimal, although, a possible risk of omitting the aerodynamic part would alter the stochastic part's effect in the obtained results. A possible solution for this might be application of Kimberly-Penman model, which can distinguish the mentioned components quantitatively. However, by applying it, it would be questionable if the target values are closer to real magnitudes as discussed by Marti *et al*. (2015).

Table 2: Error statistics of the applied models-training period

	Tabriz		Marageh	
Parameters	R^2	RMSE (mm day^{-1})	R^2	RMSE (mm day^{-1})
ANFIS 1	0.941	1.144	0.941	1.200
ANFIS 2	0.913	1.379	0.924	1.356
ANFIS 3	0.929	1.251	0.938	1.233
GEP 1	0.934	1.230	0.919	1.362
GEP 2	0.913	1.359	0.931	1.291
GEP 3	0.930	1.236	0.916	1.372

Table 3: Error statistics of the applied models-testing period

	Tabriz		Marageh	
Parameters	R^2	RMSE (mm day^{-1})	R^2	RMSE (mm day^{-1})
ANFIS 1	0.935	1.354	0.941	1.431
ANFIS 2	0.909	1.590	0.926	1.583
ANFIS 3	0.928	1.384	0.938	1.406
GEP 1	0.937	1.355	0.920	1.632
GEP 2	0.911	1.581	0.938	1.496
GEP 3	0.929	1.349	0.921	1.646

Nevertheless, it might be more suitable to consider k-fold testing based GEP approaches, when there is a lack of observed variables in the train-test stages, being able to provide a model with high generalization ability. The performance fluctuations found out among soil samples highlight the need to assess the models performance through data set scanning procedures and not only considering a single data set assignment. Otherwise, the conclusions drawn up might be misleading.

ACKNOWLEDGMENTS

Due to a higher input-output mapping ability, the GEP models are found to be more accurate. It is important to acknowledge the mapping ability of the SGEP models, because they provide sufficiently accurate estimates, even though they are trained without considering patterns of the test stage. Hence, suitably fed the GEP algorithms are able to acquire knowledge from training data and use it satisfactorily for estimation elsewhere.

REFERENCES

Allen, R.G., L.S. Pereira, D. Raes and M. Smith, 1998. Crop Evapotranspiration: Guidelines for Computing Crop Water Requirements. UN-FAO, Rome, Italy, pp: 56.

Brutsaert, W., 1982. Evaporation into the Atmosphere: Theory, History and Applications. Vol. 1, Springer, Netherlands.

De Ridder, N.A. and J. Boonstra, 1994. Analysis of Water Balances. In: Drainage Principles and Applications, Second Edition, Ritzema, H.P. (Ed.)., International Institute for Land Reclamation and Improvement, Wageningen, The Netherlands, pp: 601-633.

Doorenbos, J. And W.O. Pruitt, 1977. Crop water requirements. FAO Irrigation and Drainage Paper No. 24, FAO., Rome.

Ferreira, C., 2006. Gene Expression Programming: Mathematical Modeling by an Artificial Intelligence. Springer, Berling, Heidelberg, Pages: 478.

Hargreaves, G.H. and Z.A. Samani, 1985. Reference crop evapotranspiration from temperature. Applied Eng. Agric., 1: 96-99.

Jang, J.S.R., 1993. ANFIS: Adaptive-network-based fuzzy inference system. IEEE Trans. Syst. Man Cybern., 23: 665-685.

Kay, A.L. and H.N. Davies, 2008. Calculating potential evaporation from climate model data: A source of uncertainty for hydrological climate change impacts. J. Hydrol., 358: 221-239.

Kisi, O., 2006. Daily pan evaporation modelling using a neuro-fuzzy computing technique. J. Hydrol., 329: 636-646.

Landeras, G., A. Ortiz-Barredo and J.J. Lopez, 2008. Comparison of artificial neural network models and empirical and semi-empirical equations for daily reference evapotranspiration estimation in the Basque Country (Northern Spain). Agric. Water Manage., 95: 553-565.

Makkink, G.F., 1957. Testing the Penman formula by means of lysimeters. J. Inst. Water Eng., 11: 277-288.

Marti, P., P. Gonzalez-Altozano, R. Lopez-Urrea, L.A. Mancha and J. Shiri, 2015. Modeling reference evapotranspiration with calculated targets: Assessment and implications. Agric. Water Manage., 149: 81-90.

Priestley, C.H.B. and R.J. Taylor, 1972. On the assessment of surface heat flux and evaporation using large-scale parameters. Mon. Weather Rev., 100: 81-92.

Shiri, J. and O. Kisi, 2011. Application of artificial intelligence to estimate daily pan evaporation using available and estimated climatic data in the khozestan province (South West Iran). J. Irr. Drain. Eng., 137: 412-425.

Shiri, J., A.A. Sadraddini, A.H. Nazemi, O. Kisi, G. Landeras, A.F. Fard and P. Marti, 2014a. Generalizability of gene expression programming-based approaches for estimating daily reference evapotranspiration in coastal stations of Iran. J. Hydrol., 508: 1-11.

Shiri, J., A.H. Nazemi, A.A. Sadraddini, G. Landeras, O. Kisi, A.F. Fard and P. Marti, 2014b. Comparison of heuristic and empirical approaches for estimating reference evapotranspiration from limited inputs in Iran. Comput. Electron. Agric., 108: 230-241.

Shiri, J., P. Marti and V.P. Singh, 2014c. Evaluation of gene expression programming approaches for estimating daily evaporation through spatial and temporal data scanning. Hydrol. Process., 28: 1215-1225.

Shiri, J., P. Marti, A.H. Nazemi, A.A. Sadraddini, O. Kisi, G. Landeras and A.F. Fard, 2015. Local vs. external training of neuro-fuzzy and neural networks models for estimating reference evapotranspiration assessed through k-fold testing. Hydrol. Res., 46: 72-88.