



ESTIMATION OF MONTHLY MEAN REFERENCE EVAPOTRANSPIRATION USING GENE EXPRESSION PROGRAMMING

Dr. Ali H. Al-Aboodi¹

¹ Department of Civil Engineering, College of Engineering, University of Basrah,
alialaboodi90@gmail.com

ABSTRACT

Evapotranspiration is a main component of the water cycle and is important in crop growth. Monthly mean reference evapotranspiration (ET_o) is estimated using gene expression programming (GEP) in Basrah City, south of Iraq. Various climatic data, such as air temperature, relative humidity, and wind speed are used as inputs of GEP model to estimate the values of reference evapotranspiration (ET_o) given by the FAO-56 (Penman-Monteith equation). Nine input combinations tested with GEP are coded as model No. (1-9). Root relative squared error (RRSE) is taken as fitness function in each of GEP models. GEP models with three climatic input variables (temperature, relative humidity, and wind speed) take the highest level in the performance. The GEP technique was successfully employed to estimate ET_o in the study area. The explicit formulas obtained can be used as powerful models for estimating the mean monthly ET_o in the irrigation practices with limited climatic data.

KEYWORDS: Evapotranspiration; ET_o ; GEP; Penman-Monteith; Models; Basrah; Iraq

تخمين المعدل الشهري لتبخّر-النّتح المرجعي باستخدام برمجة التعبير الجيني

د. علي العبودي

قسم الهندسة المدنية، كلية الهندسة، جامعة البصرة

الخلاصة

التبخّر-النّتح هو من المكونات الرئيسية لدورة المياه ومهم في نمو المحاصيل؛ تم تخمين المعدل الشهري لتبخّر-النّتح المرجعي (ET_o) باستخدام برمجة التعبير الجيني (GEP) في مدينة البصرة جنوب العراق. استخدمت بيانات مناخية مختلفة، مثل درجة حرارة الهواء، الرطوبة النسبية، وسرعة الرياح كمداخلات لنموذج برمجة التعبير الجيني لتخمين قيم (ET_o) المحسوبة بواسطة معادلة (FAO-56 , PENMAN-MONTEITH). اختبرت تسع نماذج مركبة باستخدام (GEP) بترميز متسلسل من نموذج 1 الى نموذج 9. استخدم خطأ الجذر التربيعي النسبي (RRSE) كدالة موائمة في نماذج (GEP).

احتلت نماذج (GEP) مع ثلاثة مدخلات مناخية (درجة الحرارة، الرطوبة النسبية، وسرعة الرياح) أعلى مستوى في الأداء. لقد بينت نتائج الدراسة انه يمكن استخدام (GEP) بنجاح لتخمين (ET_o) في منطقة الدراسة، وان الصيغ الصريحة التي تم الحصول عليها يمكن استخدامها كنماذج فعالة لتقدير المعدل الشهري لتبخّر-النّتح المرجعي في تطبيقات الري مع البيانات المناخية المحدودة.

1. INTRODUCTION

Evapotranspiration (ET) is the combination of soil evaporation and transpiration. It is defined as the sum of the volume of water used per unit area by the vegetative growth in transpiration and that evaporation from the soil, snow, or intercepted precipitation on a given area in any specified time (Al-Barrak, 1964). Evaporation and transpiration occur simultaneously, and there is no way of distinguishing between the two processes (Allen et al., 1998). The evaporation power of the atmosphere is expressed by the reference crop evapotranspiration (ET_0). ET_0 represents the evapotranspiration from a referenced vegetated surface; a large uniform grass field is considered as the reference surface. The concept of the reference evapotranspiration was introduced to study the evaporative demand of the atmosphere independently of crop type, crop development and management practices (Allen et al., 1998).

The FAO Penman-Monteith (PM) method is recommended as the sole method for estimating ET_0 . This method is closely represented for grass ET_0 because there is a relationship between transpiration of many plants and that from grass (Kisi et al., 2013). ET is an important part of the water cycle (Jason A. et al, 2010). ET is a complex phenomenon as it depends on several climatological factors, such as temperature, relative humidity, wind speed, radiation, and type and growth stages of the crop. ET can be directly measured by using lysimeter, but this method is time consuming and requires planning of accurate (Khoshhal, and Mokarram, 2012). Thus, indirect methods based on climatological data are suitable for ET_0 estimation (Kumar et al., 2002).

In the recent years, the artificial neural network (ANN) approach has been used to model reference evapotranspiration (Sudheer et al., 2003; Kumar et al., 2002; Kisi, 2006a, 2006b, 2007; Kumer et al., 2011; and Khoshhal, and Mokarram, 2012). Adaptive neuro-fuzzy computing technique is applied for estimating pan evaporation and evapotranspiration (Kisi and Ozturk, 2007). Dogan (2009) applied adaptive neuro-fuzzy inference system (ANFIS) for estimating ET_0 , and Cobaner (2011) used ANFIS models for estimating ET_0 with weather input data, solar radiation, and air temperature. Genetic programming and wavelet transform techniques have applied for estimating ET. Traore and Guven (2013) used the gene expression programming (GEP) for estimating ET_0 using routing weather data from the tropical seasonally dry regions of West Africa in Burkina Faso. Cobaner (2013) presented wavelet transform algorithm based on discrete wavelet transform and linear regression techniques for conversion of class A pan evaporation to ET_0 .

Inputs required for PM computations include several climate variables, such as air temperature, wind speed, relative humidity, and solar radiation that are not always available or reliable. Therefore, the use of GEP with reasonable results and with different weather input data it is a good idea. These models when calibrated for a specific area is a useful and benefit tool for estimating ET_0 with limited weather data and stay away from lengthy accounts of PM method. Therefore, the main objective of this research is using GEP models to estimate the monthly mean ET_0 given by the FAO-56 PM equation in Basrah City, south of Iraq. Various input combinations of weather data are investigated using GEP models; the best combinations of this data is selected according to the comparison criteria, which are, root mean squared error (RMSE), mean absolute error (MAE), and coefficient of correlation (R).

2. CASE STUDY

Basrah City is located at Shatt Al-Arab River in southern Iraq. It is located between longitude lines ($47^{\circ} 30'$ and $48^{\circ} 30'$) and latitude lines ($30^{\circ} 00'$ and $30^{\circ} 30'$) as shown in Fig. 1. Basrah has a hot desert. Rainfall usually begins in October and continuous till May; the maximum rainfall values may be attended during January while it may be vanished during Summer

Season. During Summer Months, from June to August, Basrah is considered one of the hottest cities on the Planet, with temperature exceeding 45 °C in July. Winter frost is not unknown. Nevertheless, the climate is considered healthy and agreeable. High relative humidity, sometime exceeding 90%, due to the location of Basrah City is closely to the Arabian Gulf. Since ancient times, Basrah has been an agricultural area where palm trees, fruit, and vegetables are planted. Basrah is also known for planting tomatoes in Safwan-Al Zubair area (south west of center city) in winter season, which supplies the tomatoes demands of other Iraqi Provinces. The climate information used in this study is obtained from the meteorological recording station in Basrah City for the period (1991-2012). Conservation of existing water supplies is the first importance in the water management. To achieve this need, more information about evapotranspiration and irrigation requirements for satisfactory crop production is necessary (Al-Barrak, 1964). For purposes of timely and efficient water application, agricultural managers have long relied on evapotranspiration measurements or estimations. Therefore, an accurate assessment of ET is perquisite to improve water management practices (Roula Bachour, 2013).

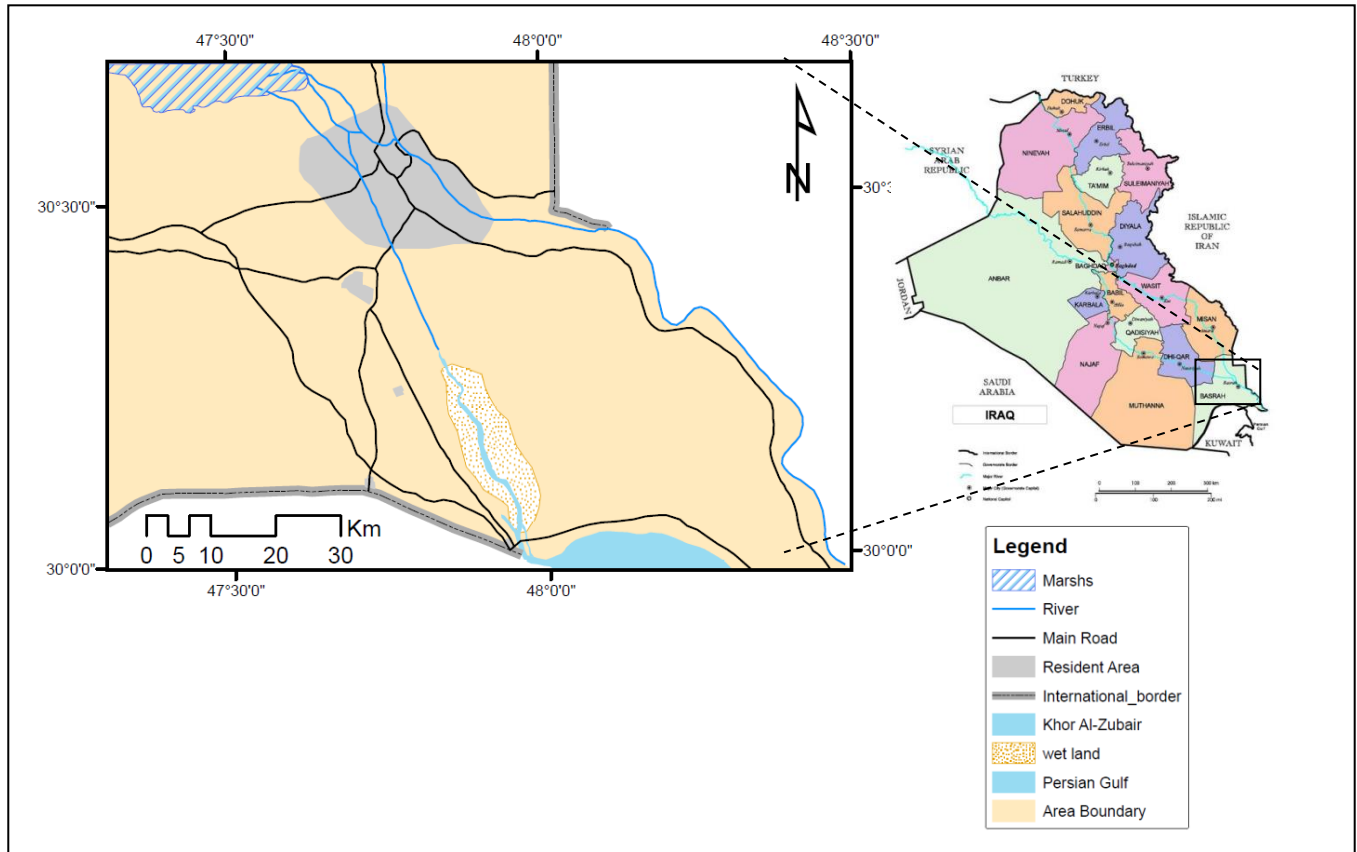


Fig. 1. Location of study area in reference to the map of Iraq.

3. CALCULATION OF REFERENCE EVAPOTRANSPIRATION ET_0

The FAO Penman-Monteith (PM) method to estimate ET_0 can be driven as (Allen et al., 1998).

$$ET_0 = \frac{0.408\Delta(R_n - G) + \frac{900\gamma u_2(e_s - e_a)}{T + 273}}{\Delta + \gamma(1 + 0.34u_2)} \quad (1)$$

Where:

ET_0 : Reference evapotranspiration [mm day^{-1}],

R_n : Net radiation at the crop surface [$\text{MJ m}^{-2} \text{day}^{-1}$],

G : Soil heat flux density [$\text{MJ m}^{-2} \text{day}^{-1}$],

T : Mean daily air temperature at 2 m height [$^{\circ}\text{C}$],

u_2 : Wind speed at 2 m height [m s^{-1}],

e_s : Saturation vapour pressure [kPa],

e_a : Actual vapour pressure [kPa],

$e_s - e_a$: Saturation vapour pressure deficit [kPa],

Δ : Slope vapour pressure curve [$\text{kPa } ^{\circ}\text{C}^{-1}$],

γ : Psychrometric constant [$\text{kPa } ^{\circ}\text{C}^{-1}$].

The slope of the relationship between saturation vapour pressure and temperature, Δ is given by

$$\Delta = \frac{4098 \left[0.6108 \exp \left(\frac{17.27T}{T + 237.3} \right) \right]}{(T + 237.3)^2} \quad (2)$$

Where:

Δ : Slope of saturation vapour pressure curve at air temperature T [$\text{kPa } ^{\circ}\text{C}^{-1}$],

T : Air temperature [$^{\circ}\text{C}$].

The atmospheric pressure, P , is the pressure exerted by the weight of the earth's atmosphere; this value is given by following equation

$$P = 101.3 \left(\frac{293 - 0.0065Z}{293} \right)^{5.26} \quad (3)$$

Where:

P : Atmospheric pressure [kPa],

Z : Elevation above sea level [m].

The psychrometric constant, γ , is given by:

$$\gamma = 0.6665 \times 10^{-3} P \quad (4)$$

Where:

P : Atmospheric pressure [kPa].

As saturation vapour pressure is related to air temperature, it can be calculated from the air temperature. The relationship is expressed by:

$$e^{\circ}(T) = 0.6108 \exp \left(\frac{17.27T}{T + 237.3} \right) \quad (5)$$

Where:

$e^{\circ}(T)$: Saturation vapour pressure at the air temperature T [kPa].

Due to the non-linearity of the above equation, the mean saturation vapour pressure for a day, week, decade or month should be computed as the mean between the saturation vapour

pressure at the mean daily maximum and minimum air temperatures for that period as shown in the following expression:

$$e_s = \frac{e^\circ(T_{max}) + e^\circ(T_{min})}{2} \quad (6)$$

Where:

$e^\circ(T_{min})$: Saturation vapour pressure at daily minimum temperature [kPa],

$e^\circ(T_{max})$: Saturation vapour pressure at daily maximum temperature [kPa].

The actual vapour pressure (e_a) can be calculated from the relative humidity (RH)

$$e_a = \frac{RH_{mean}}{100} \left[\frac{e^\circ(T_{max}) + e^\circ(T_{min})}{2} \right] \quad (7)$$

The extraterrestrial radiation, R_a , for each day of the year and for different latitudes can be estimated from the solar constant, the solar declination and the time of the year by:

$$R_a = \frac{24(60)}{\pi} G_{sc} dr [\omega s \sin(\phi) \sin(\delta) + \cos(\phi) \cos(\delta) \sin(\omega s)] \quad (8)$$

Where:

R_a : Extraterrestrial radiation [$\text{MJ m}^{-2} \text{ day}^{-1}$],

G_{sc} : Solar constant = $0.0820 \text{ MJ m}^{-2} \text{ min}^{-1}$,

Dr : Inverse relative distance Earth-Sun,

ωs : Sunset hour angle [rad],

ϕ : Latitude [rad],

δ : Solar declination [rad].

Solar radiation (R_s) is given by following formula

$$R_s = \left[a_s + b_s \frac{n}{N} \right] R_a \quad (9)$$

Where:

R_s : Solar or shortwave radiation [$\text{MJ m}^{-2} \text{ day}^{-1}$],

n : Actual duration of sunshine [hour],

N : Maximum possible duration of sunshine or daylight hours [hour],

n/N : Relative sunshine duration,

R_a : Extraterrestrial radiation [$\text{MJ m}^{-2} \text{ day}^{-1}$],

a_s : Regression constant, expressing the fraction of extraterrestrial radiation reaching the earth on overcast days ($n = 0$),

$a_s + b_s$: Fraction of extraterrestrial radiation reaching the earth on clear days ($n = N$).

Clear-sky solar radiation (R_{so}) is given by:

$$R_{so} = (0.75 + 2 \times 10^{-5}Z)R_a \quad (10)$$

Where:

Z: Station elevation above sea level [m].

The net shortwave radiation (R_{ns}) resulting from the balance between incoming and reflected solarradiation is given by:

$$R_{ns} = (1-\alpha)R_s \quad (11)$$

Where:

R_{ns} : Net solar or shortwave radiation [$\text{MJ m}^{-2} \text{ day}^{-1}$],

α : Albedo or canopy reflection coefficient, which is 0.23 for the hypothetical grass reference crop [dimensionless],

R_s : The incoming solar radiation [$\text{MJ m}^{-2} \text{ day}^{-1}$].

Net longwave radiation (R_{nl}) is given by the following formula:

$$R_{nl} = \left[\frac{(\sigma_{T_{max,K}} + \sigma_{T_{min,K}})}{2} \right] (0.34 - 0.14\sqrt{e_a}) \left(1.35 \frac{R_s}{R_{so}} - 0.35 \right) \quad (12)$$

Where:

R_{nl} : Net outgoing longwave radiation [$\text{MJ m}^{-2} \text{ day}^{-1}$],

σ : Stefan-Boltzmann constant [$4.903 \times 10^{-9} \text{ MJ K}^{-4} \text{ m}^{-2} \text{ day}^{-1}$],

$T_{max,K}$: Maximum absolute temperature during the 24-hour period [$\text{K} = ^\circ\text{C} + 273.16$],

$T_{min,K}$: Minimum absolute temperature during the 24-hour period [$\text{K} = ^\circ\text{C} + 273.16$],

e_a : Actual vapour pressure [kPa],

R_s/R_{so} : Relative shortwave radiation (limited to ≤ 1.0).

The net radiation (R_n) is the difference between the incoming net shortwave radiation (R_{ns}) and the outgoing net longwave radiation (R_{nl}) as followed:

$$R_n = R_{ns} - R_{nl} \quad (13)$$

Finally, soil heat flux density [$\text{MJ m}^{-2} \text{ day}^{-1}$] (G) for monthly period:

$$G_{month,i} = 0.14(T_{month,i} - T_{month,i-1}) \quad (14)$$

Where:

$T_{month,i}$: Mean air temperature of month i [$^\circ\text{C}$],

$T_{month,i-1}$: Mean air temperature of previous month [$^\circ\text{C}$].

4. GENE-EXPRESSION PROGRAMMING

Gene-expression programming (GEP) is a new evolutionary artificial intelligence technique developed by (Ferreira, 2001). GEP is the natural development of genetic algorithms (GAs) and genetic programming (GP). GEP uses the same kind of diagram representation of GP, but entities evolved by GEP are the expression of a genome (Ferreira, 2001). The genome or chromosome consists of a linear symbolic string of fixed length composed of one or more genes. The basic difference between these algorithms is the nature of the individuals. In GAs the individuals are symbolic string of fixed length (chromosomes), while in GP the individuals are non-linear entities of different size and shapes (parse trees), but in GEP the individuals are encoded as symbolic strings of fixed length (chromosomes) which are then expressed as non-linear entities of different size and shapes (expression trees). GEP genes are composed of a head and a tail. The head contains symbols that represent both functions and terminals, whereas the tail contains only terminals. The major types of terminal sets contain the independent variables of the problem. Table (1) shows terminal sets in the present research. After generation the initial population, each individual is then expressed, and its fitness is evaluated using one of the fitness function equations. According to fitness and the luck of the roulette, individuals are selected to be replicated. After selection process, the individuals are reproduced with some modifications performed by the genetic operators. Genetic operators, such as mutation, transposition and insertion sequence elements, and recombination are used for these modifications. The new individuals are then subjected to the same process of modification, with advances the search process, starting from the initial population toward the final population containing the desired optimal solution (Ferreira, 2001a, b).

Table 1. Input set for gene-expression programming models

Model No.	Input set	Model No.	Input set
1	MAX AIR TEMP, RELATIVE HUMIDITY , WIND SPEED	6	WIND SPEED, MIN AIR TEMP, RELATIVE HUMIDITY
2	MIN AIR TEMP, MAX AIR TEMP	7	MAX AIR TEMP, RELATIVE HUMIDITY
3	MEAN AIR TEMP, MAX AIR TEMP, RELATIVE HUMIDITY	8	WIND SPEED, RELATIVE HUMIDITY
4	MEAN AIR TEMP, MAX AIR TEMP, RELATIVE HUMIDITY , WIND SPEED	9	WIND SPEED, RELATIVE HUMIDITY, MIN AIR TEMP, MAX AIR TEMP
5	MIN AIR TEMP, WIND SPEED		

5. METHODOLOGY

In the present study, the data of the monthly mean air temperature (MEAN AIR TEMP), maximum air temperature (MAX AIR TEMP), minimum air temperature (MIN AIR TEMP), relative humidity (RELATIVE HUMIDITY), and wind speed (WIND SPEED) at 2 m above the ground surface in the Basrah City, south of Iraq, are used as inputs to GEP model to estimate the values of reference evapotranspiration (ET_o) given by the FAO-56 PM equation. DTREG (Predictive Modeling Software) (Phillip H. Sherrod, 2003) is used in the present

research. Nine input combinations tested in the study area with GEP are coded as Model No. (1-9) (see Table (1)). In the present research, the root relative squared error (RRSE) (Eq. 15) is taken as fitness function. This is based on the square root of the residual variance of the fitted model divided by the *initial variance*. *Initial variance* (Eq. 16) is the variance (Eq. 17) for the training data set using the mean value of the target variable as the predicted value for all rows. All fitness functions compute fitness scores that range from 0.0 to 1.0. A fitness of 0.0 means the model fits very poorly; it is worthless or not viable. A fitness score of 1.0 means the model fits the data perfectly. The first step in the GEP model building process is to create an initial population with a random set of functions and terminals. The population size in the present models is equal to 50; this is the number of chromosomes in the population being evolved. Usually a population size in the range of 30 to 80 chromosomes works well (Phillip H. Sherrod, 2003). Model building parameters used in this study were given in Table (2). For each model, root mean squared error (RMSE) (Eq. 18), Mean Absolute Error (MAE) (Eq. 19), Correlation between actual and predicted value (R) (Eq. 20) are used as evaluation criteria.

$$fitness = \frac{1}{1 + \sqrt{\frac{variance}{initialvariance}}} \quad (15)$$

$$initialvariance = \sum_{i=1}^N (T_i - \bar{T})^2 \quad (16)$$

$$variance = \sum_{i=1}^N (P_i - T_i)^2 \quad (17)$$

Where:

$\bar{T}$: The mean value of the target variable,

P_i : The predicted value for row i ,

T_i : The actual target value,

N : The number of rows in the training data set.

Table 2. Model building parameters of GEP

Parameter	Value	Parameter	Value
Population size	50	Max. generations	2000.0
Max. tries for initial population	10000	Gen. without improvement	1000.0
Gene per chromosome	5	Stop if fitness reaches	1.000
Gene head length	8	Fitness function	RRSE

$$RMSE = \left(\frac{\sum_{j=1}^n (Y_j - \hat{Y}_j)^2}{n} \right)^{\frac{1}{2}} \quad (18)$$

$$MAE = \frac{\sum_{j=1}^n |Y_j - \hat{Y}_j|}{n} \quad (19)$$

$$R = \frac{\sum_{j=1}^n [(Y_j - \bar{Y})(\hat{Y}_j - \bar{\hat{Y}})]}{\left[\sum_{j=1}^n (Y_j - \bar{Y})^2 \sum_{j=1}^n (\hat{Y}_j - \bar{\hat{Y}})^2 \right]^{1/2}} \quad (20)$$

Where:

Y and $\hat{Y}$: The observed and estimated values respectively,

n : The number of observations,

$\bar{Y}$ and $\bar{\hat{Y}}$: Mean of observed and estimated values respectively.

6. RESULTS AND DISCUSSION

Nine input combinations tested in the study area with GEP are coded as models from No. (1) to No. (9). The formula and statistical performance of each GEP model is shown in [Table \(3\)](#).

Table 3. Formulae and statistical performance of each GEP model

Model No.	Formulae	RMSE	MAE	R
1	$ET_o = (((9.4952 * \text{WIND SPEED}) / \text{RELATIVE HUMIDITY}) + (0.2896 * \text{MAX AIR TEMP})) - 7.1182 + \text{WIND SPEED}$	0.4766	0.3473	0.9942
2	$ET_o = (((-0.1669) * (\text{MAX AIR TEMP} + (10.6884))) + (\text{MAX AIR TEMP} / \text{Sqrt}(159.2888 / \text{MIN AIR TEMP}))) - 0.4782$	1.3250	0.9925	0.9548
3	$ET_o = (((152.6635 * ((\text{Sqrt}(\text{MEAN AIR TEMP}) + 14.5235) * (1.9101 / \text{MAX AIR TEMP}))) / \text{RELATIVE HUMIDITY}) - 6.8959) + (0.3826 * \text{MEAN AIR TEMP})$	1.1782	0.8452	0.9644
4	$ET_o = (\text{Sqrt}(\text{MEAN AIR TEMP} + 7.8913) - 7.915) + ((\text{MAX AIR TEMP} / \text{Sqrt}(\text{RELATIVE HUMIDITY})) + \text{WIND SPEED})$	0.4326	0.3349	0.9953
5	$ET_o = (0.33 * (\text{MIN AIR TEMP} - 1)) + (11.8244 / \text{MIN AIR TEMP}) + \text{WIND SPEED} + (0.1589 * (\text{MIN AIR TEMP} - 38.5756))$	0.7639	0.6182	0.9856
6	$ET_o = ((\text{WIND SPEED} + (\text{MIN AIR TEMP} / \text{Sqrt}(\text{RELATIVE HUMIDITY}))) - 3.4022) + (0.1891 * \text{MIN AIR TEMP})$	0.6064	0.4723	0.9914
7	$ET_o = (((-0.8818) * \text{Sqrt}(\text{MAX AIR TEMP})) / \text{RELATIVE HUMIDITY}) + 3.6097 + ((4.9292 * (\text{MAX AIR TEMP} + 2.4646)) / \text{RELATIVE HUMIDITY}) + ((-0.0698) * (51.3584 - \text{MAX AIR TEMP}))$	1.1532	0.8433	0.966
8	$ET_o = (\text{WIND SPEED} + (296.2335 / \text{RELATIVE HUMIDITY})) - 4.7581$	1.0599	0.8183	0.9713
9	$ET_o = ((0.6765 * \text{Sqrt}((0.7232 * \text{WIND SPEED}) / \text{RELATIVE HUMIDITY}) * \text{MIN AIR TEMP}) + \text{Sqrt}(5.0858 * \text{WIND SPEED}) + \text{Sqrt}(1.9498 * \text{MAX AIR TEMP})) - 8.8232$	0.3654	0.2688	0.9966

The input of the GEP model is based on available weather data in the study area; the performance of these models is evaluated by computing root mean squared error (RMSE),

mean absolute error (MAE), and coefficient of correlation (R). Different input sets may affect significantly the performance of model, for example, model No.2 has two inputs value, maximum and minimum air temperature; it underestimates the reference evapotranspiration given by the FAO-56 PM equation as shown in Fig. 2. Also, model No.5 has two weather inputs value, which are minimum air temperature and wind speed. The presence of wind speed with minimum air temperature increased the performance of this model as shown in Fig. 2. Addition of wind speed increases the model's performances by reducing RMSE by 42.35%, and increasing R^2 by 3.25% in comparison with model No.2. Model No.7 and model No.8 have two inputs (see Table 1); the performance of these models is slightly better than model No.2. The addition of relative humidity to temperature variable increases the performances of model No.7 by reducing RMSE 12.97% and increasing R^2 by 1.17% than model No.2 (see Fig. 2).

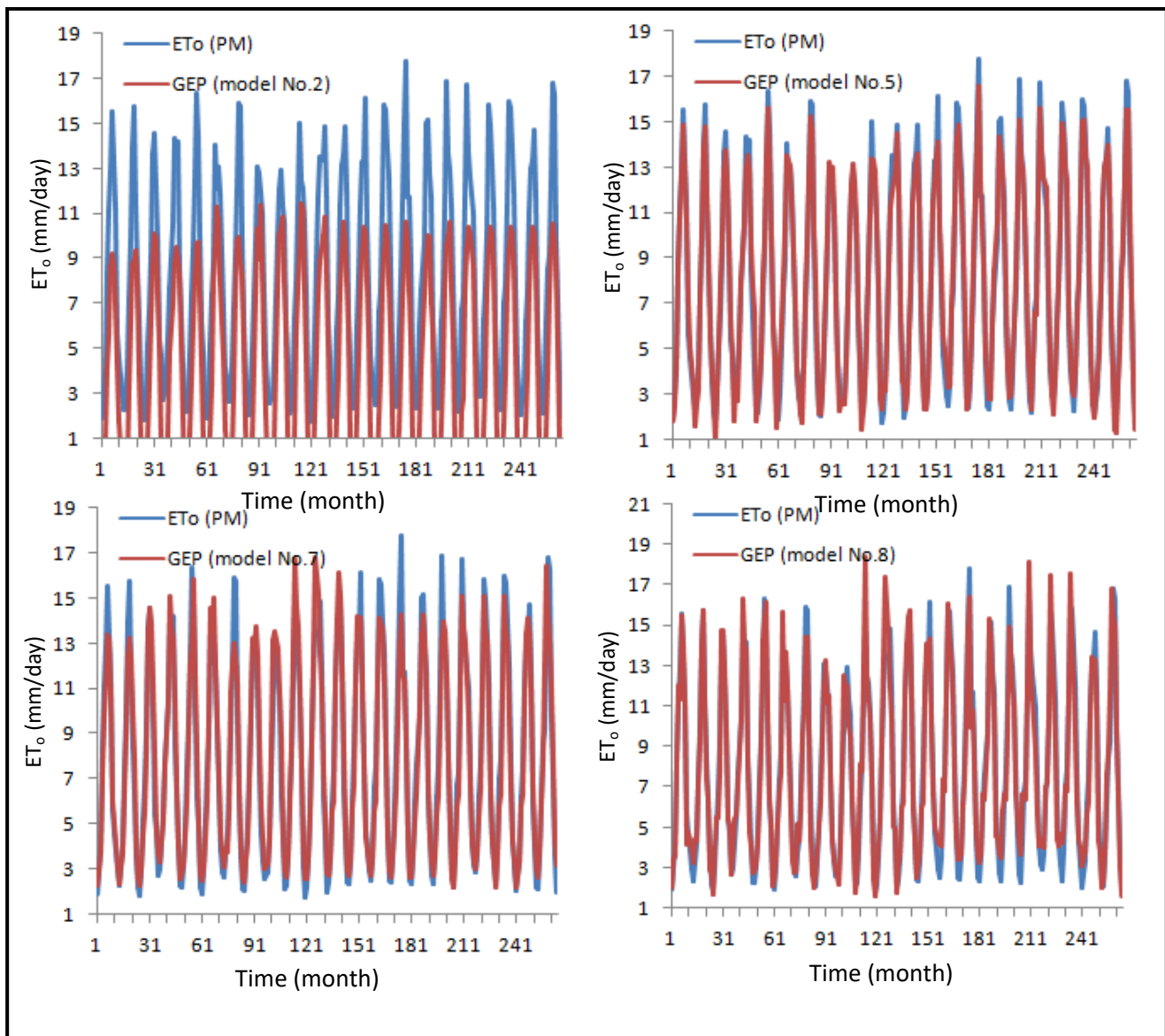


Fig. 2. Comparative plot of ET_0 given by the FAO-56 PM equation versus ET_0 estimated by GEP models (model No. 2, 5, 7, & 8).

The diversity of input variables increases the accuracy of predicted ET_o . The model with input variables (relative humidity and wind speed) in addition to temperature variable leads to a significant accuracy than model with only temperature variables as models (1, 3, and 6). Based on above results, two climatic variables are not enough for getting a very good estimation of ET_o . When increasing the number of input sets, such as wind speed, relative humidity, maximum and minimum air temperature in the estimation procedure, the GEP models' performance increases significantly (see Fig. 3).

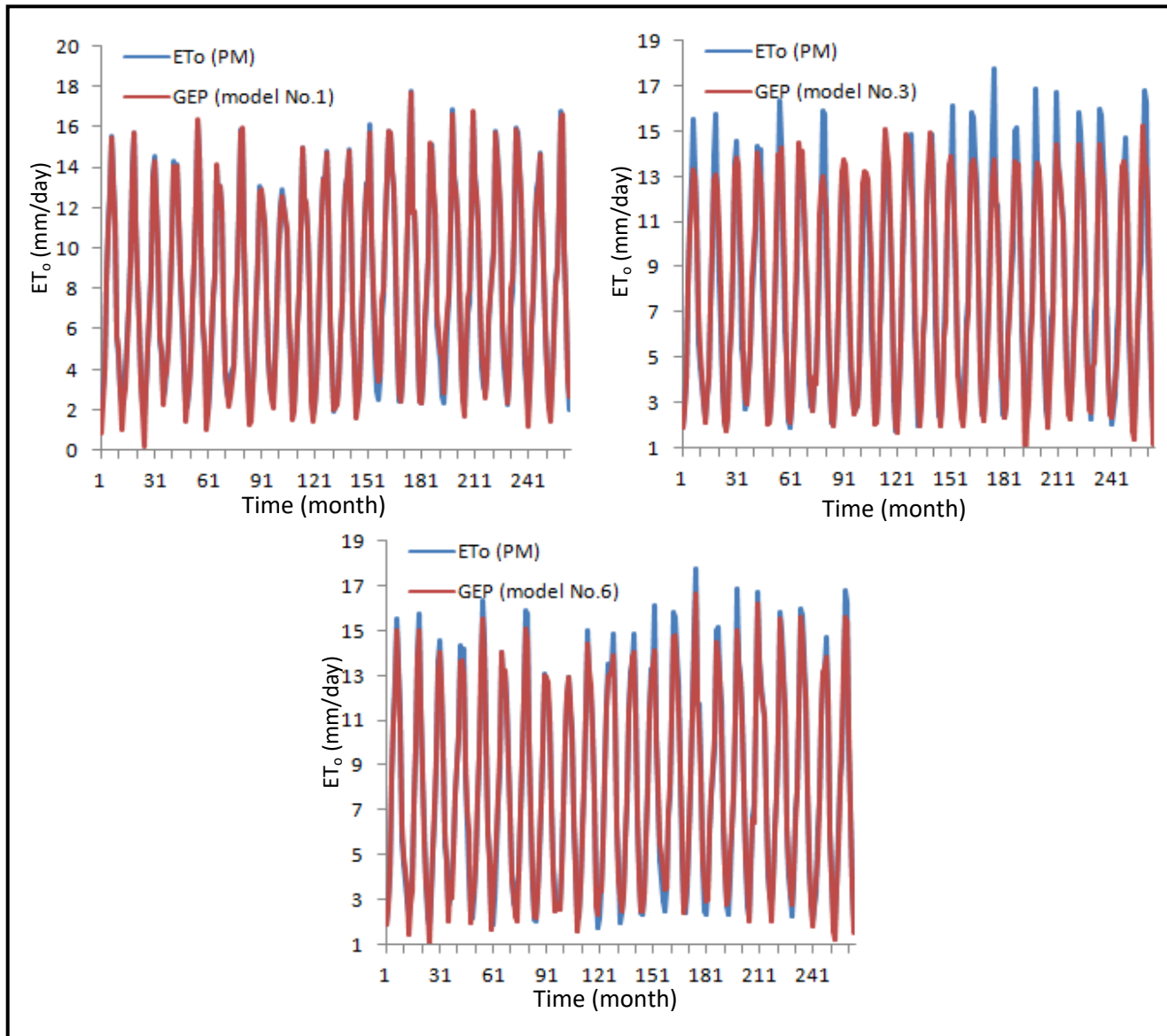


Fig. 3. Comparative plot of ET_o given by the FAO-56 PM equation versus ET_o estimated by GEP models (model No. 1, 3, & 6).

Obviously, as shown in the statistical performance of Table (3), the model No. 4 and model No.9 take The highest level in the performance with R equal to 0.9953 and 0.9966, respectively. The performance of model No.4 increases by reducing RMSE by 67.35%, and increasing R^2 by 4.24% than model No.2. Also, the performance of model No.9 increases by

reducing RMSE by 72.42% and increasing R^2 by 4.38% than model No.2. Fig. 4 shows the target and estimated monthly reference evapotranspiration for model No. 4 and 9.

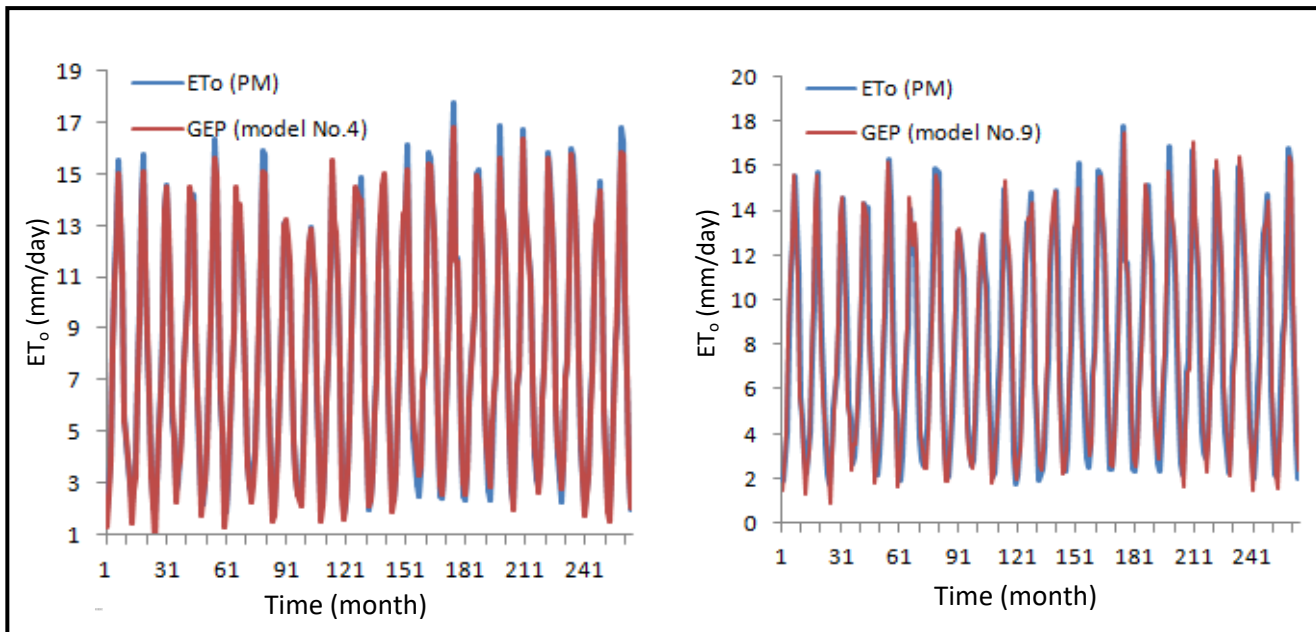


Fig. 4. Comparative plot of ET_0 given by the FAO-56 PM equation versus ET_0 estimated by GEP models (model No. 4 & 9).

7. CONCLUSION

In this study the performance of the gene expression programming technique to estimate reference evapotranspiration as a function of climatic variables in Basrah City is presented. Nine input combinations tested with GEP are coded as Model No. 1-9. The PM reference evapotranspiration was calculated based on 264 months of weather data from (1991-2012). Climatic variables in the input sets may affect the accuracy of GEP models. The diversity of input variables increases the accuracy of the ET_0 representation. The model with input variables (relative humidity and wind speed) leads to a significant accuracy than model with only input temperature variables. The GEP models with three climatic input variables (temperature, relative humidity, and wind speed) such as model No. 4 and model No.9 take the highest level in the performance with R equal to 0.9953 and 0.9966, respectively. The GEP technique was successfully employed to estimate ET_0 in the study area. The explicit formula obtained can be used as convenient and powerful model for estimating ET_0 in the irrigation practices with limited climatic data.

8. REFERENCES

- Al-Barrak, Ala H., 1964, Evaporation and potential evapotranspiration in central Iraq, Logan, Utah State Univ. Water Research Lab., Proress Report, Project WR 13, 203 p.
- Allen, R., G., Pereira, L. S., Raes, D., and Smith, M. (1998). "Crop evapotranspiration: guidelines for computing crop water requirements." FAO Irrigation and Drainage Paper No. 56, FAO, Rome, 300.

- Allen R., G., Pereira, Raes D, Smith M (1998) Crop evapotranspiration guideline for computing water requirements. FAO Irrigation Drainage Paper, No 56. FAO, Rome
- Burba.G., Jason.A.H., Pidwirny.M., Judith.S., Taub.D.R, 2010, "Evapotranspiration". In: Encyclopedia of Earth. Eds. Cutler J. Cleveland (Washington, D.C.: Environmental Information Coalition, National Council for Science and the Environment).
- Cobaner M (2011) Evapotranspiration estimation by two different neuro-fuzzy inference systems. *J Hydrol* 398(3–4):292–302
- Dogan, E. (2009) Reference evapotranspiration estimation using adaptive neuro-fuzzy inference systems. *Irrig Drain* 58:617–628
- Ferreira C (2001a) Gene expression programming in problem solving. In: Proceedings of the 6th online world conference on soft computing in industrial applications (invited tutorial)
- Ferreira C (2001b) Gene expression programming: a new adaptive algorithm for solving problems. *Complex Syst* 13(2):87–129
- Hernandez, S., Morales, L., and Sallis, P. (2011). "Estimation of reference evapotranspiration using limited climatic data and Bayesian model averaging." *ems*, pp.59-63, 2011 UK Sim 5th European Symposium on Computer Modeling and Simulation.
- Kisi O (2006a) Evapotranspiration estimation using feed-forward neural networks. *Nord Hydrol* 37(3):247–260
- Kisi O (2006b) Generalized regression neural networks for evapotranspiration modeling. *Hydrol Sci J* 51(6):1092–1105
- Kisi O (2007) Evapotranspiration modelling from climatic data using a neural computing technique. *Hydrol Process* 21:1925–1934.
- Kisi O, Ozturk O (2007) Adaptive neurofuzzy computing technique for evapotranspiration estimation. *J Irrig Drain Eng ASCE* 133(4):368–379
- Kisi O., Hatice Citakoglu , Murat Cobaner , Tefaruk Haktanir , , 2013, Estimation of Monthly Mean Reference Evapotranspiration in Turkey, *Water Resour Manage* (2014) 28:99–113.
- Khoshhal J., Mokarram M., 2012, Model for Prediction of Evapotranspiration Using MLP Neural Network. *INTERNATIONAL JOURNAL OF ENVIRONMENTAL SCIENCES* Volume 3, No 3.
- Kumar.M., Raghuwanshi.N.S., Singh.R., Wallender. W.W. and Pruitt.W.O, 2002, Estimating evapotranspiration using artificial neural network, *J Irrig Drain Eng ASCE* 128 (4): 224-233.
- Kumar M, Raghuwanshi NS, Singh R (2011) Artificial neural networks approach in evapotranspiration modeling: a review. *Irrig Sci* 29(1):11–25
- Phillip H. Sherrod (2003). DTREG Predictive Modeling Software. www.dtreg.com.
- Roula Bachour (2013) Modeling and forecasting evapotranspiration for better management of irrigation command areas, Ph.D thesis, Utah State University, Logan, Utah, 114p.
- Sudheer.K.P., Gosain.A.K., Ramasastri.K.S., 2003, Estimating actual evapotranspiration from limited climatic data using neural computing technique, *Journal of Irrigation and Drainage Engineering* 129 (3), 214–218.

Traore S, Guven A (2013) New algebraic formulations of evapotranspiration extracted from gene-expression programming in the tropical seasonally dry regions of West Africa. *Irrig Sci* 31:1–10