

Evaluating the generalizability of GEP models for estimating reference evapotranspiration in distant humid and arid locations

Hamed Kiafar¹ · Hossein Babazadeh¹ · Pau Martí² · Ozgur Kisi³ · Gorka Landeras⁴ · Sepideh Karimi⁵ · Jalal Shiri⁵

Received: 14 March 2016 / Accepted: 1 August 2016 / Published online: 10 August 2016
© Springer-Verlag Wien 2016

Abstract Evapotranspiration estimation is of crucial importance in arid and hyper-arid regions, which suffer from water shortage, increasing dryness and heat. A modeling study is reported here to cross-station assessment between hyper-arid and humid conditions. The derived equations estimate ET_0 values based on temperature-, radiation-, and mass transfer-based configurations. Using data from two meteorological stations in a hyper-arid region of Iran and two meteorological stations in a humid region of Spain, different local and cross-station approaches are applied for developing and validating the derived equations. The comparison of the gene expression programming (GEP)-based-derived equations with corresponding empirical-semi empirical ET_0 estimation equations reveals the superiority of new formulas in comparison with the corresponding empirical equations. Therefore, the derived models can be successfully applied in these hyper-arid and humid regions as well as similar climatic contexts especially in data-lack situations. The results also show that when relying on proper input configurations, cross-station might be a promising alternative for locally trained models for the stations with data scarcity.

1 Introduction

Evapotranspiration (ET) estimation is a key issue in hydrology, water resources management and planning, irrigation scheduling, ecological modeling and environmental studies, as well as for water allocating, especially in arid and semi-arid regions which suffer from non-uniform spatial-temporal distribution of total precipitation. The knowledge of differences between ET and precipitation (P) is important because it indicates whether the station (catchment) is water losing or water gaining. This has particular importance in arid and hyper-arid environments and would provide useful decision-making guidelines (Domingo et al. 2001). Alike to many points in the world, water resources are limited in Iran, especially in hyper-arid regions with non-uniform spatial and temporal distribution of precipitation throughout the year. The total annual precipitation in Iran is about 413 billion m^3 , from which a total 280 billion m^3 losses through evapotranspiration. Consequently, total renewable yearly water amount is less than 2000 m^3 for arid and hyper arid regions. Nonetheless, total water resource balance is sometimes negative leading to over-withdrawal of water in these regions. Limited total renewable water resources, increasing trend of population, disproportion of water consumption scheme, lower efficiencies of irrigation systems, and lack of a proper cope between the supplied water and its demand, make it necessary to accurately reset the demand amounts in these regions. So, accurate information about ET values in these regions is of crucial importance in water resources planning and management.

The term reference ET (ET_0) was introduced by the United Nations Food and Agriculture Organization (FAO) for computing crop evapotranspiration (Doorenbos and Pruitt 1977) and represents the evapotranspiration from a hypothesized reference crop (height 0.12 m, surface resistance 70 s/m and albedo 0.23) (Allen et al. 1998). So far, numerous attempts

✉ Jalal Shiri
j_shiri2005@yahoo.com; jalalshiri@tabrizu.ac.ir

¹ Science and Research Branch, Islamic Azad University, Tehran, Iran

² Departament de Biologia, Àrea d'Enginyeria Agroforestal, Universitat de les Illes Balears, Cra de Valldemossa km 7.5, Palma de Mallorca 07122, Spain

³ Civil Engineering Department, Architecture and Engineering Faculty, Canik Basari University, Samsun, Turkey

⁴ NEIKER, AB. Basque Country Research Institute for Agricultural Development, Alava, Basque Country, Spain

⁵ Water Engineering Department, Faculty of Agriculture, University of Tabriz, Tabriz, Iran

have been made to estimate ET_0 from measured meteorological parameters because experimental measurements cannot commonly be performed, due to absence of the required experimental equipment. The selection of one method usually depends on the amount of available inputs in practice. Years ago, the Penman-Monteith equation was adopted by the FAO [FAO56-PM] as the reference standard model for estimating ET_0 and evaluating other ET_0 equations (Allen et al. 1998). Despite being validated using a large variety of meteorological data worldwide, it presents the crucial drawback of depending from a large number of meteorological inputs for its application, which is especially dramatic in regions with a limited network of meteorological stations. Therefore, the development of alternative accurate enough models relying on fewer inputs is of great importance, especially in arid and semiarid regions with limited water and weather data availability.

In recent years, successful applications of data-driven heuristic approaches, among other artificial neural networks (ANNs), neuro-fuzzy inference systems (ANFIS), and gene expression programming (GEP), for ET_0 estimation have been performed worldwide. A complete review of such applications is beyond the scope of the present paper. In contrast to ANNs and ANFIS, GEP models can be translated into relatively simple expressions to be used by an end user.

Using the principles of genetic algorithms and genetic programming, GEP was developed by Ferreira (2001). The problems are encoded in linear chromosomes of fixed length as a computer program. GEP performs the symbolic regression using most of the genetic operators of genetic algorithm (GA). However, there are some differences between GEP and GA. Any mathematical expression defined as symbolic strings of fixed-length (chromosomes) in GA is represented to be nonlinear entities of different size and shapes (parse trees). But in GEP, it is encoded as simple strings of fixed-length which are subsequently expressed as expression trees of different size and shape (e.g., Muñoz 2005).

Among others, Parasuraman et al. (2007), Guven et al. (2008), Shiri et al. (2012), (2013a, 2014a, b) evaluated GP/GEP models for estimating ET_0 using meteorological inputs in different climatic contexts. Such studies have focused on local or regional ET_0 estimation using a single data set assignment. However, the cross-station assessment of GEP and other data-driven approaches is still very limited or even pending in ET_0 estimation. The development of new models relying on limited input, excluding variables like relative humidity, might be justified, looking for specific input-output pattern mapping under this climatic scenario. Further, the consideration of exogenous patterns from secondary ancillary stations for improving the performance of locally trained models has been assessed using ANNs and ANFIS, e.g., Kisi et al. (2012), where they developed generalized ANFIS model for pan evaporation modeling; Landaras et al. (2008), where they

developed regional ANN models for simulating ET_0 in a humid climate of Sapin; Martí and Gasque (2010) and Martí et al. (2010, 2011), where they used exogenous data for ET_0 simulations; and Shiri et al. (2013b), where a generalized ANFIS model was developed and evaluated to estimate ET_0 using data from distant stations. However, no GEP-based models have been applied in this regard. The present study aims at assessing some well-known ET_0 estimation models (three temperature-based models: Hargreaves-Samani, adjusted Hargreaves-Samani, Schendel; four radiation-based models: Priestley-Taylor, Makkink, Turc, Irmak; five mass transfer-based models: Dalton, Trabert, Meyer, WMO, Mahringer) based on limited inputs in hyper-arid and humid environments of Iran and Spain, respectively, including the development of the corresponding GEP-based models, i.e., fed with the same inputs. Further, the study also includes a cross-station scenario, assessing the performance differences of GEP models taking advantage of ancillary data from other stations.

2 Materials and methods

2.1 Study area and dataset

Data from four meteorological stations, namely Bam and Zahedan (located in the central and south-western Iran) and Derio and Igorre (located in the Basque Country, Northern Spain) were used in this study. Figure 1 shows the geographical positions of the stations. Daily maximum, minimum, and average air temperature ($T_{\max}$, $T_{\min}$, and T_{mean} , respectively); solar radiation (R_S); wind speed (W_S); and relative humidity (R_H) comprising a period of 9 years were used to estimate ET_0 . Two numerical indicators, namely the aridity index (I_A) (UNEP 1992), and the Currey continentality index (CI^{CU}) were calculated to better characterize the weather stations (see Table 1). In Table 1, it can be seen that the climatic characteristics of all stations attending to ET_0 are fluctuating during the study period. Nonetheless, the total annual precipitation and ET_0 values of Bam station corresponding to the year 2005 are clearly different from the average annual precipitation of the whole study period (320.1 vs. 61.7 mm for precipitation and 2218.79 vs. 1782.7 mm for ET_0). This could be based on the distance to the sea. In the case of Bam, this distance is lower, and as result of this, instead of being a hyper-arid site, it could present during some years more precipitation (2005) (anomalous year). However, such sharp variations are not observed for CI^{CU} values, which involve mild variations of temperature differences during the study period. The distance to the sea and altitude in Zahedan are also high. This might explain the lower $T_{\min}$ values (wider thermal ranges) in Zahedan in comparison to Bam. ET_0 indexes, i.e., maximum value, standard deviation and skewness coefficient

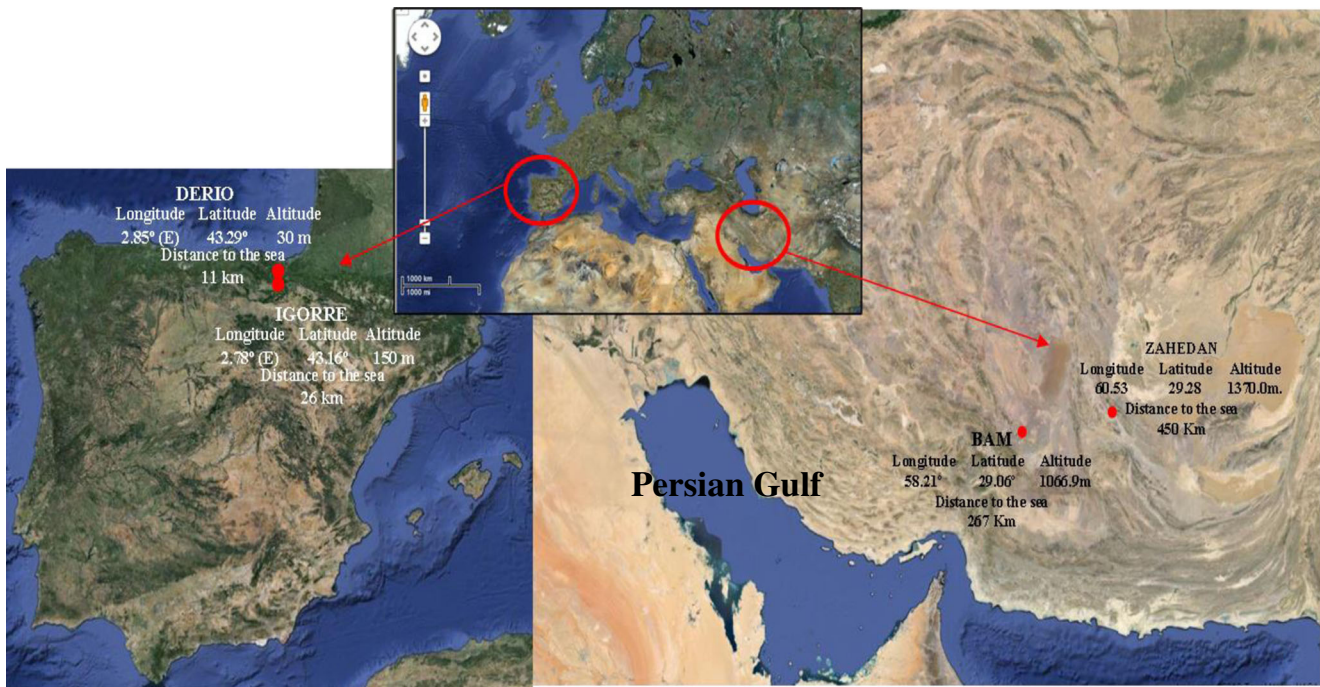


Fig. 1 Geographical positions of the studied stations

Table 1 Temporal variations of the climatologic characteristics of the studied stations during the study period (2000–2008)

	P (mm)	ET_0 (mm)	I_A^*	CI^{CU}	P (mm)	ET_0 (mm)	I_A^*	CI^{CU}
Iran								
	Bam				Zahedan			
2000	47.7	1520.45	0.031	3.031	40.7	1729.10	0.023	3.012
2001	21.1	1576.60	0.013	3.558	18.3	1758.44	0.010	3.308
2002	24	1575.34	0.015	3.096	34.2	1737.04	0.019	3.179
2003	24.3	1700.66	0.014	3.484	31.6	1674.27	0.018	3.447
2004	45.9	1699.89	0.027	3.059	64.9	1727.55	0.037	2.791
2005	320.1	2218.79	0.144	3.151	103.5	1638.36	0.063	2.970
2006	24	1903.24	0.019	3.419	41.1	1722.13	0.023	3.438
2007	33.9	1890.80	0.017	3.253	93.7	1638.85	0.057	3.105
2008	14.7	1958.63	0.007	3.844	101	1709.38	0.059	3.456
Total period	61.7	1782.70	0.031	3.840	58.7	1703.90	0.034	3.53
Spain								
	Derio				Igorre			
2000	1232.0	785.61	1.568	1.354	1612.3	745.82	2.161	1.526
2001	1223.2	754.21	1.621	1.259	1625.4	732.56	2.218	1.478
2002	1218.0	748.51	1.627	1.208	1611.4	722.12	2.231	1.321
2003	1210.3	816.60	1.482	1.930	1514.5	804.80	1.881	2.096
2004	1142.8	779.65	1.465	1.631	1410.6	781.75	1.804	1.699
2005	1294.1	820.84	1.576	1.689	1427.5	823.72	1.732	1.803
2006	1017.5	819.64	1.241	1.858	842.8	801.12	1.052	1.777
2007	1171.3	761.34	1.538	1.605	1371.5	710.19	1.931	1.829
2008	1486.6	785.96	1.891	1.429	1577.4	744.02	2.120	1.438
Total period	1221.75	785.81	1.556	2.073	1443.71	762.90	1.903	2.206

$$I_A = \frac{P}{ET_0}$$

$$CI^{CU} = \frac{M_i - m_i}{1 + \frac{\theta}{3}}$$

P total annual precipitation (mm), ET_0 total annual ET_0 (mm), I_A aridity index, CI^{CU} Currey continentality index, M_i maximum monthly average temperature ($^{\circ}\text{C}$), m_i minimum monthly average temperature ($^{\circ}\text{C}$), θ station latitude (degrees)

*According to the “World Atlas of Desertification” (UNEP 1992, 1997), dry lands have an aridity index of less than 0.65 and precipitation of less than 600 mm per year. Regions with I_A less than 0.05 are categorized as hyper-arid regions

(not presented here) in Bam are slightly higher in comparison to Zahedan, which could be related to the low R_H (minimum value) and high W_S (maximum value and standard deviation) values. So, the advection pikes in Bam might increase ET_0 and make more difficult to find out the model ET_0 response patterns. Attending to the humid stations, such sharp variations are observed for the precipitation values of Igorre Station, which presents, during some years, less precipitation (2006).

2.2 Input selection

Three categories of ET_0 estimation models, namely temperature-, radiation-, and mass transfer-based models were applied. As stated in Section 1, the input combinations were defined based on available known models relying on limited inputs, aiming at looking for specific GEP input-output relationships in the studied climatic context. Accordingly, the selected input combinations were based on the following existing approaches:

1. Temperature-based models. The Hargreaves-Samani (HS) model (Hargreaves and Samani 1985) is recommended by Allen et al. (1998) if only air temperature records are available. Although some researches recommend HS application for periods longer than 1 month (e.g., Shuttleworth 1993) the applicability of this model has been validated for daily time scales, too (e.g., Hargreaves and Allen 2003). There are various adaptations of the HS model in the literature for different climatic contexts, but in the present study, the original form of HS model (referred to as HS1) as well as the adjusted form of Trajkovic (2007) (HS2) will be used in the analysis (see Table 2 for expressions). Further, the Schendel model (Schendel 1967) was applied along with HS1 and HS2.
2. Radiation-based models. Four common radiation-based models, namely the Priestley-Taylor (PT) (1972), the Makkink (1957), the Turc (1961), and the Irmak (Irmak et al. 2003) models were used here. The PT model is the most widely used radiation-based ET_0 model as a revised form of the Penman model (1948). Makkink (1957) developed a model to estimate ET_0 for grassed lands under cool climatic conditions of the Netherlands. The Turc model is a simplification of the Makkink model and requires air temperature, solar radiation, and relative humidity as input variables. The Irmak model is a multi-regression-based equation, which has been calibrated and tested through Florida's data.
3. Mass transfer-based models. Using the Dalton's law, mass transfer-based models utilize the eddy motion transfer of water vapor from the evaporative surface into the surrounding atmosphere. These models are easier to use and generally demand air temperature, relative humidity, and wind speed measurements as inputs (Singh and Xu

1997). In the present study, the Dalton (1802), Trabert (1896), Meyer (1926), WMO (1966), and Mahringer (1970) models were considered.

The corresponding mathematical expressions of these models as well as the necessary meteorological inputs for their application are given in Table 2. These input combinations were used to feed the corresponding GEP-based models, too. Moreover, FAO56-PM ET_0 values were considered as targets for calibrating the applied models, which is an accepted and very common practice, given the absence of experimental measurements. Figure 2 illustrates the applied input configurations.

2.3 Gene expression programming

The application of the GEP procedure involves the following steps.

1. Determining the fitness function: the root mean square error (RMSE) fitness function is applied here according to Shiri et al. (2012).
2. Choosing the set of terminals T and the set of functions F : Here, the terminal set includes the meteorological variables. The appropriate functions for modeling ET_0 are $\{+, -, \times, \div, \sqrt[3]{}, \sqrt{}, \ln, e^x, x^2, x^3, \sin x, \cos x, \text{Arctgx}\}$ (Shiri et al. 2012).
3. Selecting the length of head (h) and genes per chromosome: Here, $h = 8$ and three genes per chromosome were employed according to Ferreira (2001).
4. Choosing of the linking function: Here, addition linking functions were applied according to Shiri et al. (2012).
5. Choosing the genetic operators: The parameters used per run are those used by Shiri et al. (2012).

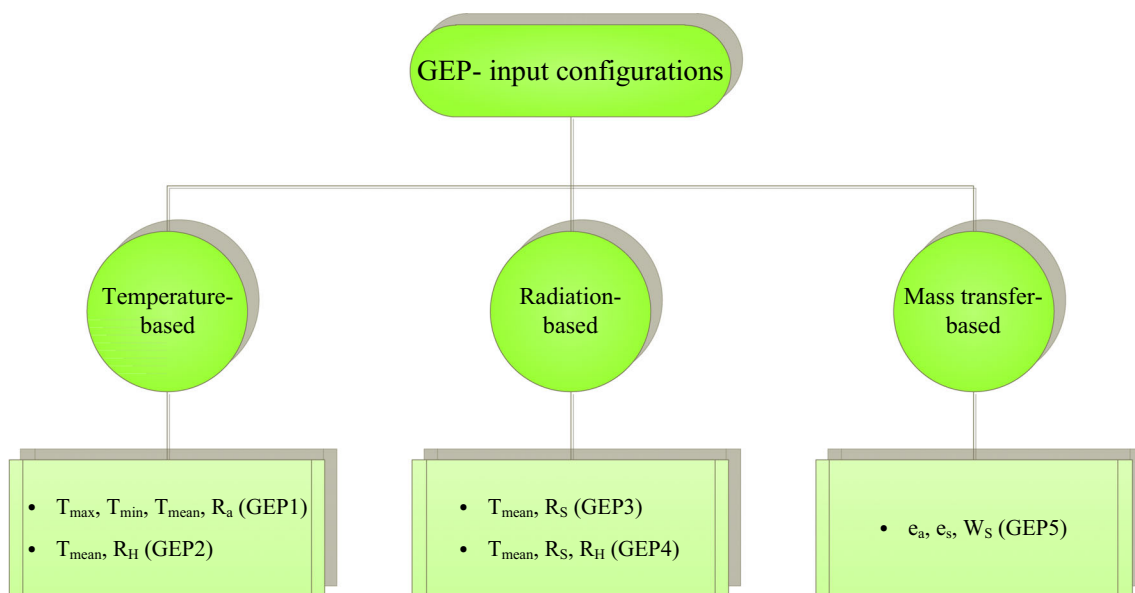
2.4 Study flowchart

Two methodological approaches were considered in this study for assessing the model performance: (a) local approach and (b) cross-station approach. In approach (a), the aforementioned models were applied individually per station using the complete local dataset (9-year daily parameters). The application of GEP-based models requires splitting the dataset into three subsets (training, testing, and validation). In approach (a), these subsets were defined chronologically. Accordingly, data from January 2000 to December 2004 (1827 patterns) were used for developing (training) the GEP models. Then, data from January 2005 to December 2006 (730 patterns) were used for testing the GEP models, and finally, the data from January 2007 to December 2008 (731 patterns) were reserved for an

Table 2 Mathematical expressions of applied ET_0 estimation equations

ET_0 models	Meteorological inputs	Expression
Standard ET_0 model FAO56-PM	$T_{\text{mean}}, R_S, W_S, R_H$	$ET_0 = \frac{0.408\Delta(R_n - G) + \gamma \frac{900}{T_{\text{mean}} + 273} W_S (e_s - e_a)}{\Delta + \gamma(1 + 0.34W_S)}$
Temperature-based ET_0 estimation models		
Hargreaves-Samani (HS1)	$T_{\text{mean}}, T_{\text{max}}, T_{\text{min}}, [Ra]$	$ET_0 = 0.0023 \frac{R_a}{\lambda} (T_{\text{mean}} + 17.8) \sqrt{T_{\text{max}} - T_{\text{min}}}$
HS2	$T_{\text{mean}}, T_{\text{max}}, T_{\text{min}}, [Ra]$	$ET_0 = 0.0023 \frac{R_a}{\lambda} (T_{\text{mean}} + 17.8) (T_{\text{max}} - T_{\text{min}})^{0.424}$
Schendel	T_{mean}, R_H	$ET_0 = 16 \frac{T_{\text{mean}}}{R_H}$
Radiation-based ET_0 estimation models		
Irmak	T_{mean}, R_S	$ET_0 = 0.149R_S + 0.079T_{\text{mean}} - 0.611$
Priestley-Taylor	T_{mean}, R_S	$ET_0 = \frac{\alpha}{\lambda} \frac{\Delta}{\Delta + \gamma} (R_n - G)$
Makkink	T_{mean}, R_S	$ET_0 = 0.61 \frac{\Delta}{\Delta + \gamma} \frac{R_S}{\lambda} - 0.12$
Turc	$T_{\text{mean}}, R_S, R_H$	$ET_0 = a_T 0.013 \frac{T_{\text{mean}}}{T_{\text{mean}} + 15} \frac{23.8856R_S + 50}{\lambda}$ $R_H \geq 50 \rightarrow a_T = 1$ $R_H < 50 \rightarrow a_T = 1 + \frac{50 - R_H}{70}$
Mass transfer-based ET_0 estimation models		
Dalton	e_a, e_s, W_S	$ET_0 = (0.3648 + 0.07223W_S)(e_s - e_a)$
Trabert	e_a, e_s, W_S	$ET_0 = 0.3075 \cdot \sqrt{W_S} (e_s - e_a)$
Meyer	e_a, e_s, W_S	$ET_0 = (0.375 + 0.0502W_S)(e_s - e_a)$
WMO	e_a, e_s, W_S	$ET_0 = (0.1298 + 0.0934W_S)(e_s - e_a)$
Mahringer	e_a, e_s, W_S	$ET_0 = 0.15072 \cdot \sqrt{3.6W_S} (e_s - e_a)$

In these equations, ET_0 reference evapotranspiration (mm/day), Δ slope of the saturation vapor pressure function (kPa/°C), γ psychrometric constant (kPa/°C), R_n net radiation (MJ/m²/day), G soil heat flux density (MJ/m²/day), T_{mean} mean air temperature (°C), W_S average 24 h wind speed at 2 m height (m/s), e_s saturation vapor pressure (kPa), e_a actual vapor pressure, α 1.26, λ latent heat of the evaporation (MJ/Kg), R_a extraterrestrial radiation (mm/day), R_S daily solar radiation (MJ/m²/day), R_H relative humidity (%), T_{max} maximum air temperature (°C), T_{min} minimum air temperature (°C)

**Fig. 2** Illustration of applied input configurations

independent validation of the GEP models. Cross-validating reduces the over-fitting risk and helps the user to efficiently assess the models' ability (Pour Ali Baba et al. 2013). In approach (b), two different applications were examined: (b1) external assessment of the GEP models and (b2) application with ancillary data. In approach (b1), the GEP model was trained using the complete dataset of one station and tested using the complete dataset of the second station. In approach (b2), the meteorological patterns from one station were used as inputs for estimating ET_0 at the second station. Figure 3 represents a schematic flowchart of the study. The different approaches mentioned above ((a), (b1), and (b2)) are represented in this figure. This flowchart is for arid stations. In the case of humid stations, the approach is the same.

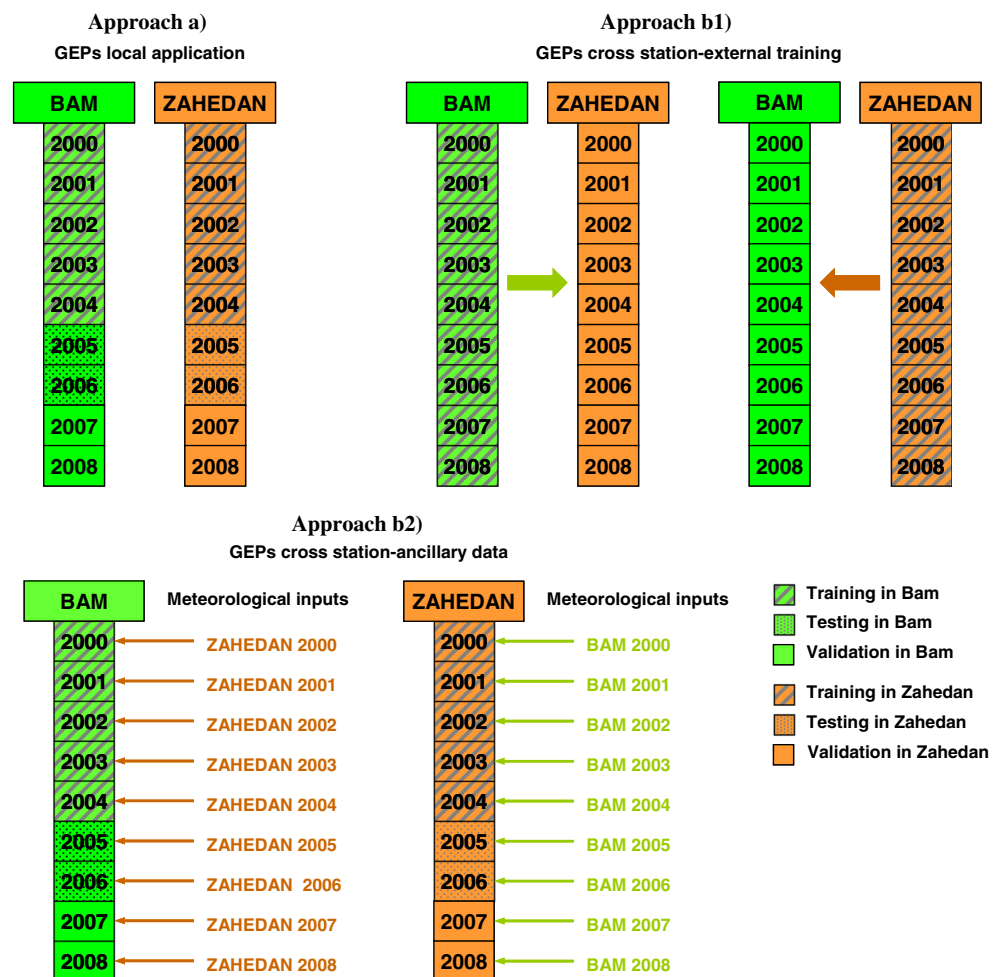
3 Results and discussions

3.1 Local derivation of ET_0 equations

The comparison of the monthly ET_0 values estimated with the conventional models in the arid and humid stations (not

presented here) shows that in case of temperature-based models, HS1 provides the most accurate results for the arid stations (Figs. 4 and 5). Comparing its performance with other temperature-based models, HS1 can estimate the ET_0 trend better than HS2 and Schendel models throughout the 12 months (dry as well as wet seasons). Also, three statistical parameters were used for assessing the models' performance, namely, the coefficient of determination (r^2), the root mean square error (RMSE), and the coefficient of residual mass (CRM) (Legates and McCabe 1999). The results presented in Table 3 confirm this statement. So, HS1 might be ranked as the most accurate temperature-based model in arid stations. On a RMSE basis, the HS1 performance in Bam and Zahedan is similar ($\Delta RMSE = 0.05$ mm/day), while HS2 and Schendel models show a different performance in both stations. Although the RMSE differences might be linked to the average ET_0 order of magnitude at each station, the CRM values (which are a weighted ET_0 differences) are quite different in both stations, suggesting a higher underestimation in Zahedan, especially for the Schendel model. This may be attributed to higher temperature range differences and lower mean wind speed values in Zahedan (Landaras et al. 2009). In

Fig. 3 Schematic study flowchart



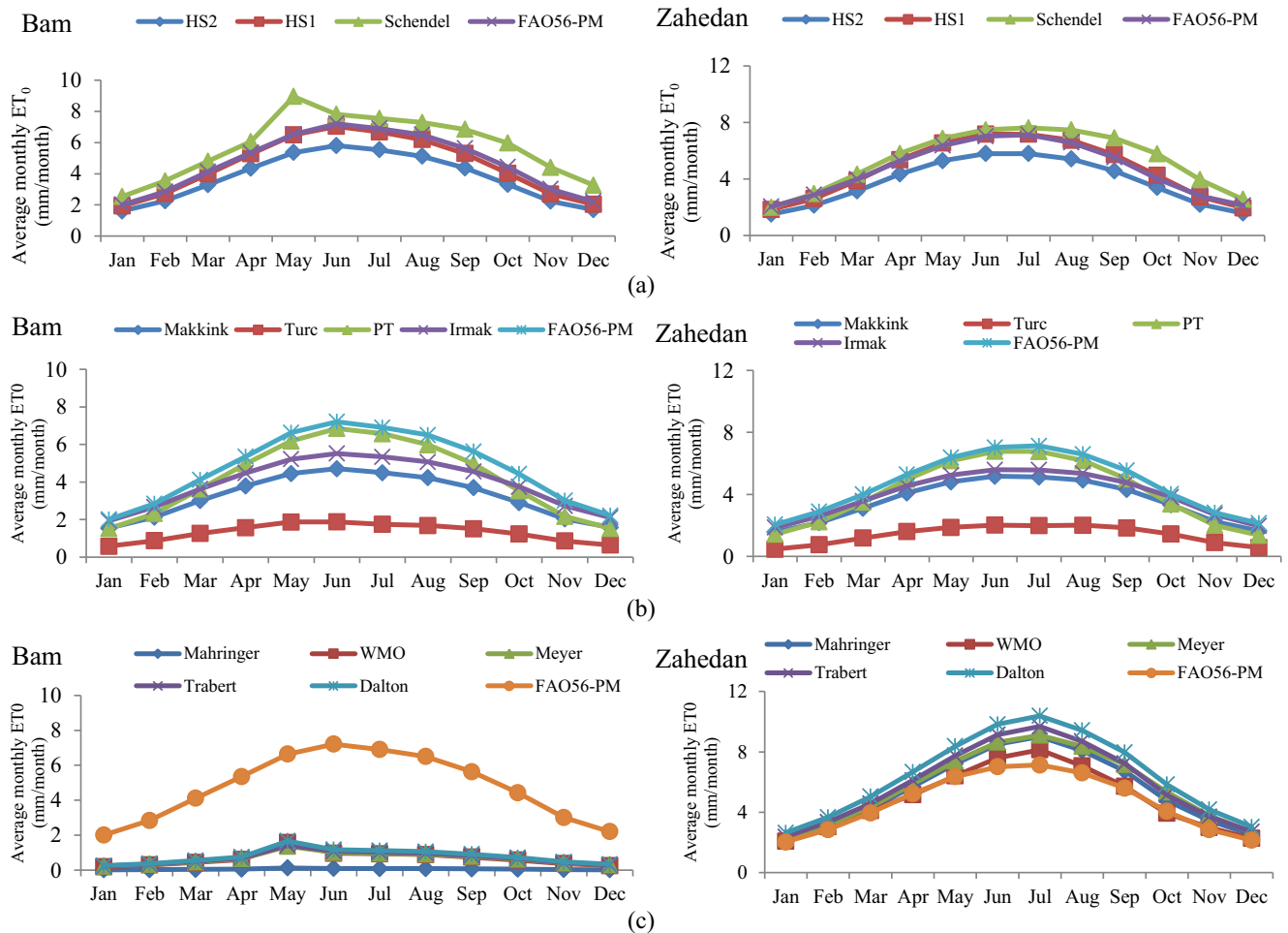


Fig. 4 Average monthly ET_0 values of the applied ET_0 estimation models during the study period (2000–2008) arid stations: **a** temperature-based models, **b** radiation-based models, and **c** mass transfer-based models

the case of humid stations, however, the HS2 model presents most accurate results as can be seen from Fig. 5 and Table 3. Comparing between two stations, the HS2 performance in Derio and Igorre is similar ($\Delta RMSE = 0.006$ mm/day; $\Delta CRM = 0.024$), while HS1 and Schendel models show different performance. Nonetheless, the CRM values of the HS2 model show overestimations trend in both stations, while the trend of Schendel and HS1 models are different in the stations.

Attending to the radiation-based models, all the applied models underestimate ET_0 throughout the months in the arid stations. The Turc model provides the worst estimates, while the PT model offers the most accurate estimations for Bam and Zahedan stations. On an RMSE basis, the overall accuracies of the radiation-based models in Zahedan are higher than those of Bam. This might be caused by the solar radiation-air temperature relationship characteristics, where the average air temperature in Bam is higher than in Zahedan (24.2 vs. 19.45 °C), while its average incoming solar radiation is lower than in Zahedan (17.61 vs. 20.45 MJ/m² day). Moreover, the inaccurate performance of the Turc model in both stations may be due to the high skewed nature of R_H in these stations.

Similar to the arid stations, in the humid locations, the applied radiation-based models underestimate ET_0 throughout the months, except the PT and Irmak models, which show overestimation trend in both Derio and Igorre stations, in some month (especially in warm season). Among the radiation-based models, the Irmak model has the most accurate performance in both Derio and Igorre, with the lowest performance difference between two stations ($\Delta RMSE = 0.005$ mm/day).

Finally, the mass transfer-based models provide inaccurate results in both arid and humid stations, although its simulations are more accurate for humid stations than those of arid stations. A reason for this might be the low aerodynamic effects in the studied arid regions, which make it difficult to estimate ET_0 from the available data using these models. In general, for the studied hyper-arid stations, the temperature-based models (i.e., HS1) might be ranked as the most accurate models, followed by radiation-based models (i.e., PT model) and mass transfer-based models (i.e., Mahringer). The performance of the models in Zahedan is more accurate than in Bam, which can be attributed to the high ET_0 values of Bam. A similar trend can be observed in the humid stations, where

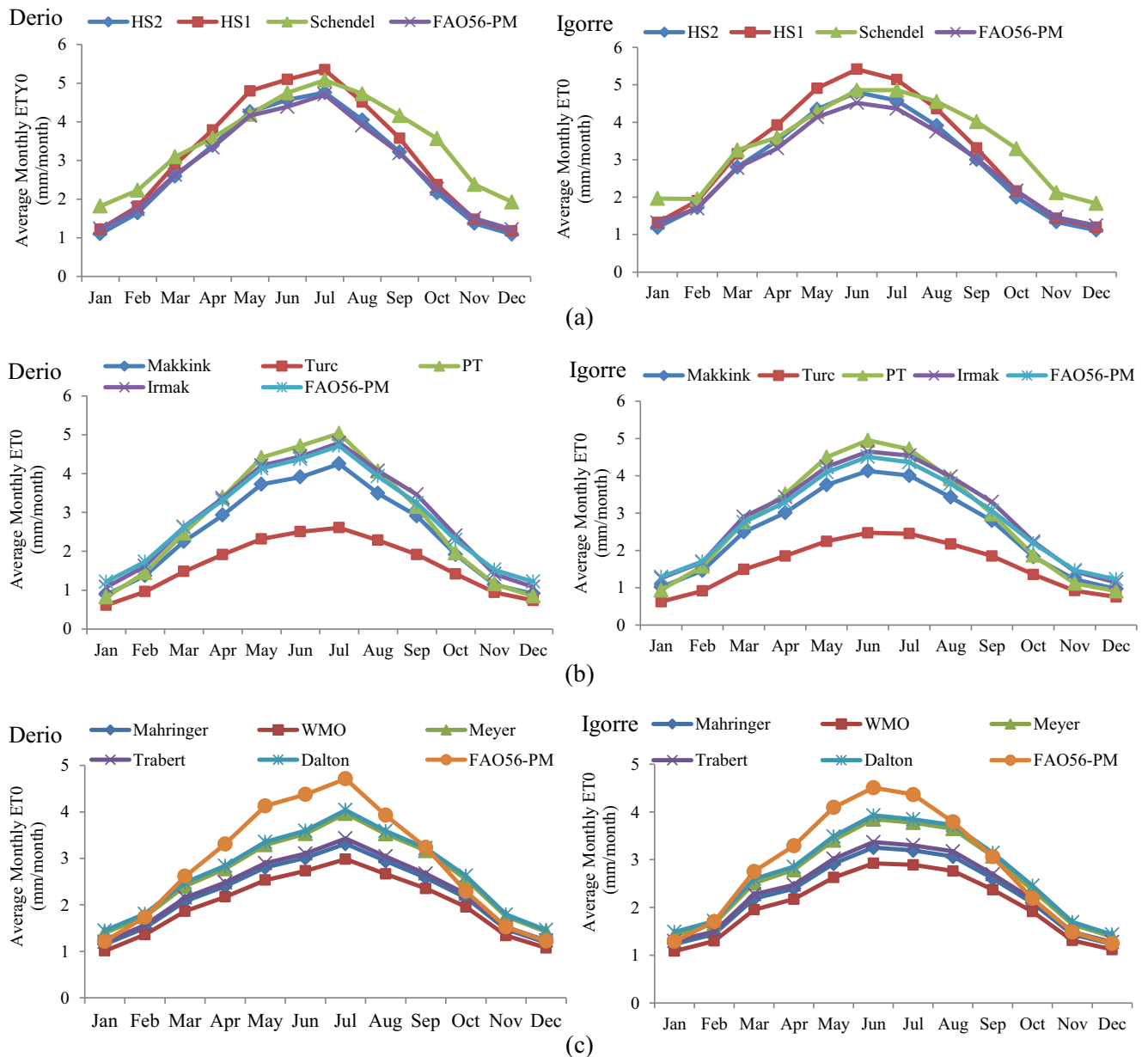


Fig. 5 Average monthly ET_0 values of the applied ET_0 estimation models during the study period (2000–2008) humid stations: **a** temperature-based models, **b** radiation-based models, and **c** mass transfer-based models

the Irmak model is ranked as the most accurate model, followed by HS2 model.

Tables 4 and 5 sum up the performance of the GEP-based models for, respectively, arid and humid stations during the testing and validation periods. Attending to the testing statistics, GEP5 (corresponding to mass transfer-based models) provides the most accurate results in both arid stations. According to Table 4, it seems that the relative difference between the GEP1 (temperature-based) and GEP4 (radiation-based) models is much larger in Bam than in Zahedan in the testing period. This suggests that R_s is much more effective over ET_0 in Bam than in Zahedan both in the testing and validation periods. The accuracy of the GEP3 and

GEP4 models is very similar in Bam, while the GEP4 model performs better than the GEP3 model in Zahedan attending to RMSE. This reveals that R_H might be more influential on ET_0 in Zahedan than in Bam. This can also be observed attending to the statistical characteristics of the R_H data, which have a high skewed distribution (not presented here) in Bam. The relative differences between the GEP5 and GEP4 models clearly show that the aerodynamic effect on ET_0 is higher in Bam than in Zahedan. It should be noted that the accuracy of the GEP models in the testing period (2005–2006) is lower than in the validation period (2007–2008). The reason for this might be the fact that Bam presents high ET_0 values in the test period (especially in 2005), as can be clearly observed in

Table 3 Statistics of the ET₀ estimation models during the study period (2000–2008)

	Arid stations (Iran)						Humid stations (Spain)					
	Bam			Zahedan			Derio			Igorre		
	r^2	RMSE (mm/day)	CRM	r^2	RMSE (mm/day)	CRM	r^2	RMSE (mm/day)	CRM	r^2	RMSE (mm/day)	CRM
Temperature-based models												
HS1	0.919	0.617	0.042	0.896	0.667	0.904	0.854	0.661	−0.190	0.882	0.688	0.185
HS2	0.917	1.216	0.208	0.909	1.062	0.748	0.854	0.428	0.002	0.868	0.434	0.026
Schendel	0.539	2.082	−0.214	0.836	1.150	0.716	0.504	1.110	−0.353	0.571	1.086	0.255
Radiation-based models												
Irmak	0.918	1.185	0.172	0.890	1.039	0.768	0.926	0.358	−0.013	0.936	0.353	−0.053
Priestly-Taylor	0.884	0.886	0.114	0.892	0.843	0.847	0.939	0.444	0.039	0.921	0.466	−0.001
Makkink	0.901	1.812	0.722	0.860	1.401	0.579	0.816	1.591	0.722	0.890	1.611	0.746
Turc	0.846	3.776	0.320	0.824	3.568	−1.729	0.919	0.540	0.227	0.919	0.482	0.184
Mass transfer-based models												
Dalton	0.313	7.210	−0.667	0.883	2.340	−0.183	0.405	0.922	0.104	0.326	1.092	0.073
Trabert	0.408	5.566	−0.512	0.877	1.875	0.245	0.368	1.095	0.336	0.302	1.167	0.293
Meyer	0.358	5.226	−0.429	0.903	1.504	0.514	0.413	0.915	0.140	0.332	1.069	0.115
WMO	0.224	7.440	−0.396	0.768	1.475	0.545	0.342	1.319	0.509	0.278	1.340	0.476
Mahringer	0.408	4.990	−0.406	0.877	1.480	0.529	0.368	1.142	0.382	0.303	1.197	0.342

Table 4 Testing and validation statistics of the GEP and best empirical models-arid stations

	Testing (2005–2006)			Validation (2007–2008)		
	r^2	RMSE (mm/day)	CRM	r^2	RMSE (mm/day)	CRM
Bam						
Temperature based						
GEP1	0.932	0.802	0.142	0.975	0.768	0.115
GEP2	0.708	1.257	0.121	0.951	0.939	0.140
HS1	0.908	0.808	0.105	0.972	0.762	0.121
Radiation based						
GEP3	0.916	0.847	0.120	0.964	0.908	0.146
GEP4	0.907	0.846	0.114	0.968	0.898	0.145
PT	0.885	1.119	0.176	0.932	1.147	0.186
Mass transfer based						
GEP5	0.902	0.667	0.043	0.950	0.593	0.050
Mahringer	0.180	9.821	−0.790	0.970	3.165	−0.477
Zahedan						
Temperature based						
GEP1	0.930	0.510	−0.032	0.937	0.505	−0.014
GEP2	0.900	0.622	−0.036	0.910	0.592	−0.011
HS1	0.909	0.640	0.918	0.912	0.616	0.927
Radiation based						
GEP3	0.932	0.531	−0.032	0.930	0.532	−0.017
GEP4	0.938	0.505	−0.028	0.935	0.501	−0.006
PT	0.901	0.779	0.879	0.903	0.796	0.879
Mass transfer based						
GEP5	0.936	0.495	−0.006	0.944	0.465	−0.006
Mahringer	0.876	1.376	0.623	0.882	1.398	0.628

Table 5 Testing and validation statistics of the GEP and best empirical models-humid stations

	Testing (2005–2006)			Validation (2007–2008)		
	r^2	RMSE (mm/day)	CRM	r^2	RMSE (mm/day)	CRM
Derio						
Temperature based						
GEP1	0.839	0.495	0.013	0.845	0.485	0.016
GEP2	0.677	0.711	−0.042	0.718	0.670	0.068
HS2	0.858	0.503	0.009	0.792	0.453	−0.001
Radiation based						
GEP3	0.937	0.309	0.010	0.938	0.305	0.001
GEP4	0.930	0.312	−0.013	0.940	0.301	0.005
Irmak	0.927	0.415	−0.007	0.927	0.858	−0.009
Mass transfer based						
GEP5	0.617	0.773	−0.055	0.579	0.804	0.023
Dalton	0.448	1.054	0.088	0.370	1.033	0.126
Igorre						
Temperature based						
GEP1	0.929	0.510	−0.030	0.930	0.490	−0.098
GEP2	0.840	0.735	−0.078	0.849	0.772	0.072
HS2	0.866	0.525	0.012	0.770	0.482	−0.071
Radiation based						
GEP3	0.967	0.350	0.006	0.977	0.286	−0.008
GEP4	0.934	0.490	−0.028	0.967	0.351	0.029
Irmak	0.925	0.445	−0.060	0.942	0.360	−0.013
Mass transfer based						
GEP5	0.729	0.938	−0.288	0.666	0.982	−0.332
Dalton	0.363	1.290	0.066	0.303	1.182	0.058

Table 1. It can be observed that the trend, at least for mass transfer-based models, is adverse to ET_0 equations and mass transfer-based GEP models perform better than temperature-based and radiation-based GEP models. This may be explained through aerodynamic component effects on ET_0 values. It seems that the mass transfer-based equations are not able to explain this effect in these dry regions suitably, but GEP models provide more accurate estimates with relatively lower error values. In Zahedan, wind speed presents the highest skewness in comparison to the other climatic parameters, and in both stations, the standard deviation values of this parameter take high values. This higher variation might cause inaccurate estimates when conventional approaches are used. Nevertheless, GEP allows for a more realistic mapping of this nonlinear complex process. Attending to the general estimation trends, the GEP-based models overestimate ET_0 values in Zahedan (negative CRM values) and underestimate it in Bam (positive values of CRM). In contrast to arid stations, the radiation-based GEP3 and GEP4 models perform more accurately than the other models in humid stations (see Table 5).

Table 6 Statistical criteria values of the GEP cross-station application: External training

	r^2	RMSE (mm/day)	CRM	r^2	RMSE (mm/day)	CRM
a				b		
GEP1	0.886	0.783	0.089	0.891	0.858	-0.104
GEP2	0.878	1.040	0.167	0.817	1.211	-0.170
GEP3	0.860	0.764	-0.044	0.871	1.233	-0.185
GEP4	0.898	0.707	-0.014	0.876	1.025	-0.149
GEP5	0.929	0.832	0.135	0.906	0.940	-0.138
c				d		
GEP1	0.808	0.589	-0.019	0.755	0.618	-0.026
GEP2	0.624	0.830	0.066	0.640	0.780	-0.080
GEP3	0.935	0.348	-0.033	0.931	0.326	0.012
GEP4	0.947	0.324	-0.034	0.920	0.350	0.013
GEP5	0.425	1.013	0.005	0.517	0.863	0.010
e				f		
GEP1	0.528	1.178	0.020	0.882	1.366	0.229
GEP2	0.624	4.578	0.987	0.520	2.151	0.335
GEP3	0.924	0.654	-0.183	0.916	1.075	0.245
GEP4	0.922	0.862	-0.318	0.876	0.941	0.124
GEP5	0.805	0.998	0.135	0.755	1.258	0.235

a GEP model trained using the whole data of Bam station and tested using the whole data of Zahedan station, b GEP model trained using the whole data of Zahedan station and tested using the whole data of Bam station, c GEP model trained using the whole data of Derio station and tested using the whole data of Igorre station, d GEP model trained using the whole data of Igorre station and tested using the whole data of Derio station, e GEP model trained using the whole data of arid stations and tested using the whole data of humid stations, f GEP model trained using the whole data of humid stations and tested using the whole data of arid stations

Kisi (2009) also indicated in his study that the R_H input has a significant effect on evaporation and adding this parameter into input combination significantly increases the models' accuracies in humid stations or climate. It is clear from Table 5 that the GEP models provide more accurate estimates for Derio Station than the Igorre. The reason of this may be the fact that the ET_0 data of Igorre Station have a wider range and a higher skewness than those of the Derio Station. In some cases, there are higher values of determination coefficient in the testing data set compared to the validation data set. This should be due to the fact that the characteristics of validation data set is more similar to the training (calibration) data set than the testing data set.

Table 7 Statistical criteria values of the GEP cross-station application: ancillary data application

	Test (2005–2006)			Validation (2007–2008)		
	r^2	RMSE (mm/day)	CRM	r^2	RMSE (mm/day)	CRM
First typology						
GEP1	0.882	0.674	-0.011	0.926	0.581	-0.033
GEP2	0.641	1.195	-0.029	0.904	0.664	-0.025
GEP3	0.727	1.028	-0.027	0.923	0.584	-0.042
GEP4	0.787	0.917	-0.016	0.926	0.563	-0.037
GEP5	0.646	1.202	-0.048	0.905	0.724	-0.075
Second typology						
GEP1	0.820	1.067	0.127	0.931	1.158	0.179
GEP2	0.694	1.275	0.121	0.822	1.370	0.180
GEP3	0.774	1.129	0.119	0.906	1.200	0.175
GEP4	0.813	1.074	0.125	0.932	1.147	0.180
GEP5	0.708	1.233	0.115	0.822	1.849	0.266
Third typology						
GEP1	0.729	0.943	0.070	0.621	1.092	-0.068
GEP2	0.597	1.144	-0.010	0.541	1.231	-0.028
GEP3	0.694	1.066	0.029	0.595	1.021	-0.041
GEP4	0.602	1.064	0.033	0.702	1.211	-0.041
GEP5	0.622	1.116	-0.017	0.654	1.244	-0.043
Fourth typology						
GEP1	0.606	0.793	0.010	0.579	0.990	0.004
GEP2	0.650	0.940	-0.040	0.596	0.936	0.054
GEP3	0.564	0.924	-0.012	0.508	1.013	0.060
GEP4	0.497	0.796	-0.021	0.550	0.980	0.018
GEP5	0.502	0.897	-0.006	0.564	0.891	0.053

First typology Meteorological parameters of Bam station are used as input variables to estimate ET_0 values of Zahedan station, Second typology Meteorological parameters of Zahedan station are used as input variables to estimate ET_0 values of Bam station, Third typology Meteorological parameters of Derio station are used as input variables to estimate ET_0 values of Igorre station, Fourth typology Meteorological parameters of Igorre station are used as input variables to estimate ET_0 values of Derio station

3.2 Cross-station derivation of ET_0 equations

As stated in Section 2, two different applications were considered in the cross-station approach: external assessment and ancillary data supply application. The r^2 , RMSE, and CRM results of the GEP models for the cross-station applications are given in Tables 6 and 7. In the case of the external assessment (a), the GEP4 (radiation-based) model provides the most accurate results, followed by the GEP3, GEP1, GEP5, and GEP2 models, respectively. In the case of the external assessment (b), however, the mass transfer-based GEP5 model performs better than the other models. The accuracy ranks of the other GEP models showing decreasing accuracy, are, respectively, GEP1, GEP4, GEP2, and GEP3. In the case of external assessment (c), similar to the case (a), the GEP4 (radiation-based) model has the best accuracy followed by the GEP3, GEP1, GEP2, and GEP5. In the case of external assessment (d), the GEP3 (radiation-based) model performs the best followed by the GEP4, GEP1, GEP2, and GEP5. In the case of external assessment (e), the radiation-based GEP3 model has the most accurate results followed by the GEP4, GEP5, GEP1, and GEP2. In the case of external assessment (f), the radiation-based GEP4 model gives the most accurate estimates followed by the GEP3, GEP5, GEP1, and GEP2. These different estimation trends of the GEP models in the external assessment may be due to different patterns in the trends of the climatic variables in four stations. Accordingly, the input-output relationships might differ in both arid and humid stations and might not be extrapolatable. It is clear from the Table 6 that the radiation-based GEP models (GEP3 and GEP4) generally perform better than the other models in the case of external training. Comparison of (a), (b) vs. (f) cases indicates that the use of data from arid station in training

considerably increases model accuracy in arid stations. This is also valid for the humid stations (see the cases of (c), (d) vs. (e) in Table 6). Similar to the previous application (local application), the GEP models generally gives better estimates in humid stations than the arid stations for the cross-station application with external training. GEP2 model (relying on T_{mean} and R_H) seems to be unable to estimate ET_0 with enough accuracy in both cases. According to Table 6, it seems preferable to use Bam meteorological data (in both cross applications) to estimate in Zahedan than to use Zahedan data to estimate Bam. This can be explained according to data statistics (not presented here), where ET_0 statistical indices presents a wider range in Bam than in Zahedan. So, there are more pike events of ET_0 in Bam than in Zahedan. The models calibrated using Zahedan's data may have difficulty in estimating ET_0 data (extrapolation difficulty) in Bam, because the training patterns cover a lower range than the corresponding test set. Another reason might be related with the skewed distributed ET_0 data in Bam, which makes it difficult to estimate through external training. Mass transfer-based GEP5 models seem to provide less-scattered estimates than the other GEP models. However, GEP5 models significantly underestimate ET_0 in Bam. Hence, based on the considered limited inputs used to feed the models, the relationships encountered might not be able to generalize properly out of the training station, especially if the range spectrum is very different within training and testing stations. The generalizability might be partially improved through the consideration of further inputs.

The test and validation performance accuracy of the GEP models considering ancillary input data is given in Table 7. The GEP4 (radiation-based) model provides the most accurate results in the validation period, followed, respectively, by the GEP1, GEP3, GEP2, and GEP5 models for the first and

Table 8 Mathematical expressions of the optimal GEP models

	Model	Expression
Optimal GEPs: Local application		
Zahedan	GEP5	$ET_0 = [\arctg(e_a - e_s) \cdot e_s]^4 + \arctg[2e_s - 1.865e_a] \cdot \sqrt[3]{W_s} + \left(W_s^2(e_a + e_s + W_s)^{-1}\right)^{0.166}$
Igorre	GEP3	$ET_0 = [R_s + T_{\text{mean}}R_s - 0.5656]0.25 + \arctg[\sin(-0.153R_s)] + \sin[\arctg(-41.68T_{\text{mean}}^{-1})]$
Optimal GEPs: Cross-station application: External training		
Scenario a	GEP4	$ET_0 = [2T_{\text{mean}} - R_s^2][\text{Exp}(0.225T_{\text{mean}})] + \text{Arctg}\left[(-3.878R_H)\left(\text{Exp}(0.17T_{\text{mean}})^{-1}\right)\right] + R_s^{0.66}$
Scenario b	GEP1	$ET_0 = [T_{\text{mean}}(0.0053(R_a - 5.3467))] + \arctg[T_{\text{mean}} + 6.8161R_a^{-1} - T_{\text{min}} - 0.71] + 0.0106T_{\text{max}} - 0.07$
Scenario c	GEP4	$ET_0 = \sin\left[\cos\sqrt{\text{Exp}(0.0014 + \sqrt[3]{R_s})}\right] + \sin\left[\text{Ln}\left((T_{\text{mean}} + 5.7106)^2 + 5.165R_s\right)\right] + \sqrt{\text{Exp}\sqrt[3]{R_s}}$
Scenario d	GEP3	$ET_0 = \cos[0.7475 + \sqrt[3]{T_{\text{mean}} + 6.082}] + (2T_{\text{mean}} + 1.577)^{0.16} \sqrt[3]{R_s} + \sin[\arctg\sqrt[3]{R_s} - 0.841\sqrt{R_s}]$

second topology. In the third topology, however, GEP3 (radiation-based) model has the best accuracy while the GEP5 model provides the best estimates in the fourth topology. In contrast to the local application, the mass transfer-based GEP5 model performs worse than the other models in the ancillary data application except for the fourth typology. A comparison with the results of the local application (Table 5) clearly reveals that the ancillary data application decreases the models' accuracies much more in Bam. The reasons given in the external training application seem to be also valid for this application (ancillary data application). Comparing with the cross-station application 1 (external training), the GEP models seem to be more accurate in this application, in which local meteorological data were used in calibration of the GEP models. The performance of the cross-station applications is quite accurate. So, in case of lack of meteorological data in hyper-arid and humid areas, the utilization of data of only one station could be interesting for the estimation of ET_0 in wide areas. Table 8 represents the optimal GEP mathematical expressions for applied scenarios. As can be clearly observed from the tables, the model expressions can be used by anyone not necessarily being familiar with GEP. The GEP model provides practical way for ET_0 estimation to obtain accurate results and encourages use of GEP in other aspects of water engineering studies.

Nevertheless, according to Santos et al. (2014), the worst results of the approaches considered in humid stations could be due to the necessity to include another independent variable as the NAO index, which might be subject to future studies.

4 Conclusions

This paper provides new expressions based on gene expression programming to estimate reference evapotranspiration from limited inputs in hyper-arid and humid environments. The performance of the new heuristic models is compared with the performance of the corresponding temperature-, radiation-, and mass transfer-based conventional approaches, considering a local and cross-station assessment in two hyper-arid and humid stations of Iran and Spain, respectively. The local prediction ability of the GEP models is higher than the performance of the conventional approaches. So, if enough local data series are available, the development of local GEP models can be a more accurate alternative to conventional existing approaches. The accuracy of the GEP models decreases outside by using cross-station scenario, because models relying on ancillary inputs might not be able to perform a suitable simulation. The results showed that in case of lack of meteorological data in hyper-arid and humid areas, the utilization of data of only one station could be interesting. In the present study, data from two hyper-arid and humid stations were applied for deriving the new expressions. Further studies

might be carried out for analyzing the ET_0 trends and deriving new equations using data from similar stations worldwide. These may be subjects for future studies.

References

- Allen RG, Pereira LS, Raes D, Smith M (1998) Crop evapotranspiration. Guidelines for computing crop evapotranspiration. FAO Irrigation and Drainage Paper no. 56, Rome, Italy
- Dalton J (1802) Experimental essays on the constitution of mixed gases; on the force of steam of vapour from waters and other liquids in different temperatures, both in a torricellian vacuum and in air on evaporation and on the expansion of gases by heat. *Mem Manch Lit Philos Soc* 5:535–602
- Domingo F, Villagarcía L, Boer MM, Alados-Arboledas L, Puigdefàbregas J (2001) Evaluating the long-term water balance of arid zone stream bed vegetation using evapotranspiration modeling and hill slope runoff measurements. *J Hydrol* 243:17–30
- Doorenbos J, Pruitt WO (1977) Crop water requirements. FAO Irrigation and Drainage Paper No. 24, Rome, Italy
- Ferreira C (2001) Gene expression programming: a new adaptive algorithm for solving problems. *Comp Syst* 13(2):87–129
- Güven A, Aytekin A, Yüce MI, Aksoy H (2008) Genetic programming-based empirical model for daily reference evapotranspiration estimation. *Clean Soil, Air, Water* 36(10–11):905–912
- Hargreaves GH, Allen RG (2003) History and evaluation of Hargreaves evapotranspiration eq. *J Irrig Drain Eng* 129(1):53–63
- Hargreaves GH, Samani ZA (1985) Reference crop evapotranspiration from temperature. *Appl Eng Agric* 1(2):96–99
- Irmak S, Irmak A, Allen RG, Jones JW (2003) Solar and net radiation-based equations to estimate reference evapotranspiration in humid climates. *J Irrig Drain Eng* 129(5):336–347
- Kisi O (2009) Daily pan evaporation modelling using multi-layer perceptrons and radial basis neural networks. *Hydrol Process* 23: 213–223
- Kisi O, Pour Ali Baba A, Shiri J (2012) Generalized neuro-fuzzy models for estimating daily pan evaporation values from weather data. *J Irrig Drain Eng* 138(4):349–362
- Landeras G, Ortiz-Barredo A, López JJ (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 Manag* 95:553–565
- Landeras G, Ortiz-Barredo A, López J (2009) Forecasting weekly evapotranspiration with ARIMA and artificial neural network models. *J Irrig Drain Eng* 135(3):323–334
- Legates DR, McCabe GJ (1999) Evaluating the use of “goodness-of-fit” measures in hydrologic and hydroclimatic validation. *Water Resour Res* 35(1):233–241
- Mahringer W (1970) Verdunstungsstudien am Neusiedler See. *Arch Met Geoph Biokl Ser B* 18:1–20
- Makkink GF (1957) Testing the Penman formula by means of lysimeters. *J Inst Water Eng* 11(3):277–288
- Martí P, Gasque M (2010) Ancillary data supply strategies for improvement of temperature-based ET_0 ANN models. *Agric Water Manag* 97(7):939–955
- Martí P, Royuela A, Manzano J, Palau G (2010) Generalization of ET_0 ANN models through data supplanting. *J Irrig Drain Eng* 136(3): 161–174
- Martí P, González-Altozano P, Gasque M (2011) Reference evapotranspiration estimation without local climatic data. *Irrig Sci* 29(6):469–495

- Muñoz DG (2005) Discovering unknown equations that describe large data sets using genetic programming techniques, Master Thesis, Electronic Systems, Linköping Institute of Technology, LITH-ISY-EX-05/3697
- Meyer A (1926) U"ber einige Zusammenh"nge zwischen Klima und Boden in Europa. *Chem Erde* 2:209–347
- Parasuraman K, Elshorbagy A, Carey SK (2007) Modelling the dynamics of the evapotranspiration process using genetic programming. *Hydrol Sci J* 52(3):563–578
- Penman WR (1948) Natural evaporation from open water, bare soil and grass. *Proc. R Soc Edinburg Sect A, Royal Society Publishing, London* 193(1032):120–145
- Pour Ali Baba A, Shiri J, Kisi O, Fakheri Fard A, Kim S, Amini R (2013) Estimating daily reference evapotranspiration using available and estimated climatic data by adaptive neuro-fuzzy inference system and artificial neural networks. *Hydrol Res* 44(1):131–146
- Priestley CHB, Taylor RJ (1972) On the assessment of surface heat flux and evaporation using large-scale parameters. *Mon Weather Rev* 100(2):81–92
- Santos JF, Portela MM, Pulido-Calvo I (2014) Spring drought prediction based on winter NAO and global SST in Portugal. *Hydrol Process* 28(3):1009–1024
- Schendel U (1967) Vegetationswasserverbrauch und -wasserbedarf. Habilitation, Kiel, p 137
- Singh VP, Xu CY (1997) Evaluation and generalization of 13 mass-transfer equations for determining free water evaporation. *Hydrol Process* 11:311–323
- Shiri J, Kisi O, Landeras G, Lopez JJ, Nazemi AH, Stuyt LCPM (2012) Daily reference evapotranspiration modeling by using genetic programming approach in the Basque Country (northern Spain). *J Hydrol* 414–415:302–316
- Shiri J, Sadraddini AA, Nazemi AH, Kisi O, Marti P, Fakheri Fard A, Landeras G (2013a) Evaluation of different data management scenarios for estimating daily reference evapotranspiration. *Hydrol Res* 44(6):1058–1070
- Shiri J, Nazemi AH, Sadraddini AA, Landeras G, Kisi O, Fakheri Fard A, Marti P (2013b) Global cross station assessment of neuro-fuzzy models for estimating daily reference evapotranspiration. *J Hydrol* 480:46–57
- Shiri J, Marti P, Singh VP (2014a) Evaluation of gene expression programming approaches for estimating daily pan evaporation through spatial and temporal data scanning. *Hydrol Process* 28(3):1215–1225
- Shiri J, Sadraddini AA, Nazemi AH, Kisi O, Landeras G, Fakheri Fard A, Marti P (2014b) Generalizability of gene expression programming-based approaches for estimating daily reference evapotranspiration in coastal stations of Iran. *J Hydrol* 508:1–11
- Shuttleworth WJ (1993) Evaporation. In: Maidment DR (ed) *Handbook of hydrology*. McGraw-Hill, New York
- Trabert W (1896) Neue Beobachtungen u"ber Verdampfungsgeschwindigkeiten. *Meteorol Z* 13:261–263
- Trajkovic S (2007) Hargreaves versus Penman-Monteith under humid conditions. *J Irrig Drain Eng* 133(1):38–42
- Turc L (1961) Evaluation des besoins en eau d'irrigation evapotranspiration potentielle. *Ann Agron* 12(1):13–49
- UNEP (United Nations Environment Programme) (1992) *World atlas of desertification*. Edward Arnold, London
- UNEP (United Nations Environmental Programme) (1997) *World atlas of desertification*. Editorial commentary by N. Middleton and D.S.G. Thomas. Edward Arnold, London
- WMO (1966) Measurement and estimation of evaporation and evapotranspiration. Tech. Pap. (CIMO-Rep) 83. Genf