Questioning of empirically derived and locally calibrated potential evapotranspiration equations for a lumped water balance model

Umut Okkan and Huseyin Kiymaz

ABSTRACT

One of the most essential inputs of water balance models is the part in which potential evapotranspiration (PET) is predicted. Especially in the conceptual-based lumped rainfall–runoff models, the steady runoff simulations can be made with acceptable PET predictions. The presented study is about exploring alternative PET equations that can be adapted to a parametric lumped model termed as the dynamic water balance model (*dynwbm*). Although the use of the Penman–Monteith equation often appears in the literature, a performance assessment was conducted on the *dynwbm* by using 21 PET equations. The implementation was performed on different river branches in the Gediz Basin, Turkey. The satisfactory PET equations have been selected by means of statistical techniques. As a result of the evaluation, it was observed that one of the radiation-based equations, McGuinness–Bordne, provided the most consistent performance. Alternatively, the presence of parsimonious equations requiring less meteorological variables has been questioned, thus locally calibrated temperature-based PET equations reflecting the PET estimations of McGuinness–Bordne have been proposed so as to be practically utilized in water balance modeling experiments for the basin.

Key words | empirical potential evapotranspiration, parsimonious PET equations, Penman-Monteith, radiation and temperature-based methods, water balance modeling

Umut Okkan (corresponding author)
Huseyin Kiymaz
Department of Civil Engineering, Hydraulic Division,
Balikesir University,
10145, Balikesir,
Turkey
E-mail: *umutokkan@balikesir.edu.tr*

INTRODUCTION

In the hydrological cycle, the loss of transpiration over the vegetation and the evaporation through the water bodies on the surface is called evapotranspiration. The maximum amount of this loss or its occurrence at the time of sufficient soil moisture refers to the potential evapotranspiration (PET), while actual evapotranspiration is limited to the current soil moisture content. Since many physical geographical factors (such as latitude, altitude, and vegetation) affect the PET, the mechanism of its occurrence varies depending on the region and can be associated with meteorological observations such as relative humidity (RH) and wind speed, primarily radiation and temperature (Xu & Singh 2002). In this context, however, direct measurement of the PET is not feasible, as in pan evaporation. Thus, the precise prediction of the PET has great importance especially in the determination of plant water consumption and irrigation water requirement. Accordingly, observations by lysimeters (direct method) or empirical formulas (indirect methods) can be taken as a basis. The question of which empirical method is more appropriate was addressed by a limited number of researchers up to the early 1980s (e.g., Thornthwaite 1948; Blaney & Criddle 1950; Makkink 1957; Hamon 1961; Jensen & Haise 1963; McGuinness & Bordne 1972; Priestley & Taylor 1972). Later, the FAO (Food and Agriculture Organization) report (FAO-56), edited by Allen *et al.* (1998), stated that

doi: 10.2166/wcc.2019.292

the Penman–Monteith (Pen-Mon) equation, which is a combination of the approaches of Penman (1948) and Monteith (1965), is a comprehensive indirect method. In some validation studies conducted in specific regions throughout the world with various climatic conditions, it has been emphasized that the Pen-Mon equation is highly compatible with lysimeter measurements (Itenfisu *et al.* 2003; Allen *et al.* 2005; Jain *et al.* 2008). It is therefore currently used as a reference formula (Pandey *et al.* 2016).

After the release of FAO-56, researchers tended to compare the Pen-Mon equation with other available methods. Moreover, re-calibration procedures were applied to update the coefficients of existing equations and the regression-based equations were also suggested (Xu & Singh 2002; Irmak *et al.* 2003; Xystrakis & Matzarakis 2011; Tabari *et al.* 2013; Bogawski & Bednorz 2014; Pandey *et al.* 2016). For example, in a study performed by Pandey *et al.* (2016), it was detected that Irmak (Irm), Makkink (Makk), Turc (Turc), and Blaney–Criddle (Bl-Cr) equations were more compatible with Pen-Mon for the north-eastern region of India. Xystrakis & Matzarakis (2011) applied 13 empirical equations on seven meteorological stations in southern parts of Greece and stated that the absolute biases obtained from the McGuinness–Bordne (McG-Bor) and Hamon (Ham) equations were relatively less. Xu & Singh (2002) applied five equations on the Changins station in Switzerland and suggested that if the Pen-Mon method was taken as the basis, the coefficients of Priestley–Taylor (Prs-Tyl) and Rohwer (Roh) equations should be recalibrated. Kellner (2001) also pointed out that the Prs-Tyl is a more appropriate equation for humid regions with high latitudes. Tabari *et al.* (2013) tested 31 equations at the Rasht station in the north of Iran, which has a humid climate. In their comparative study, it was argued that the Bl-Cr is more closely related to the Pen-Mon in terms of various statistical indices.

The main reason why the above-mentioned studies were performed is that the Pen-Mon equation of FAO-56 has more data requirements than other empirical ones. In particular, the solar radiation measurement network is not well-distributed as in temperature stations. Moreover, the regional studies will also shed light on which empirical equations can produce more consistent predictions. Even if these regional investigations are generally the subject of agricultural practices, a detailed analysis about PET is also needed for water balance modeling (Oudin *et al.* 2005a, 2005b). For instance, the lumped rainfall–runoff models also require PET input in addition to precipitation. In these models, the physical aspects of the transformation of rainfall into runoff are usually conceptualized by means of the parameters and the runoff series can be simulated by applying the continuity equation to the representative storage elements. Paturel *et al.* (1995) proved that runoff predictions derived from water balance models showed sensitivity to precipitation at a primary level and the PET also had a significant influence on model sensitivity. In the water balance modeling studies, the PET input can be defined as a regression-based function of mean temperature (T) and RH variables (e.g., $PET = aT^b(100-RH)$; $PET = aT^b$; $PET = ae^{bT}$) or any empirical PET equation can be used directly (Xu & Vandewiele 1995; Fowler *et al.* 2007; Okkan & Kirdemir 2018). On the other hand, Xu & Vandewiele (1995) stated that these extra free parameters embedded into modeling would make the calibration procedure more difficult and that geographic regionalization studies with empirical PET input would be more proper. Vandewiele *et al.* (1991) have confirmed that when an empirical PET is used, the parameters may be better associated with the physical properties of the basin.

Various rainfall–runoff model practices using diverse empirical PET inputs are found in the literature. However, the reason why the related empirical equation was preferred in those studies remained ungrounded (McKillop *et al.* 1999; Bárdossy & Das 2008; Caldwell *et al.* 2015). In a few studies, the impacts of varied PET inputs on the rainfall–runoff performances were investigated in depth. The most comprehensive among them was the study prepared by Oudin *et al.* (2005b). In their work, the predictions obtained from 27 empirical PET equations were separately evaluated as inputs in the rainfall–runoff modeling and the outputs derived from application were compared with observed runoff series. The study was carried out across a large area including the major basins in Australia, France, and the USA. In particular, the performance of the McG-Bor equation in the hydrological model was reported to be more reasonable than those obtained from other equations. In another study, Kannan *et al.* (2007) asserted that the Hargreaves method (Harg) yielded more reasonable results than those of the Pen-Mon in the performance evaluation of the SWAT (Soil and Water Assessment Tool)-based modeling study for Bedfordshire, UK. In a study conducted by Wang *et al.* (2006) for the northwestern Minnesota basin, unlike other studies, it was found that the responses of the hydrological model produced by Pen-Mon, Prs-Tyl, and Harg equations were close to each other.

As can be seen from the literature reviews, there is no general judgment about which empirical PET equation has a better impression on hydrological model outputs. In Turkey, while the Pen-Mon and Bl-Cr methods are generally used in the calculation of plant water consumption (e.g., Beyazgül *et al.* 2000; Okkan & Kirdemir 2018), there are no equations examined or proposed from the point of view of rainfall–runoff modeling. In addition, significant changes in meteorological variables have been observed in many regions in Turkey due to climatic change. The increase in population and industrialization capacity, as well as greenhouse gas emissions, which have been continuously increasing, have triggered a meaningful increase in temperature and the PET. In this context, some hydro-meteorological predictions of basins have been made under different scenarios in Turkey. However, comparative analyses on PETs are unfortunately not available in these studies (e.g., Ozkul 2009; Okkan & Fistikoglu 2014; Okkan & Kirdemir 2018).

The study, prepared on the basis of the various reasons mentioned above, deals with the evaluation of the PET equations which may constitute an input to a monthly water balance model arranged on the basin scale and the determination of alternative parsimonious equations in terms of meteorological variables. In this study, Gediz Basin, which represents an important reserve of agricultural activities in the Aegean region of Turkey, has been selected as the study area. It is thought that the study performed has a unique value in the context of the derived results and can be adapted by researchers to other neighboring regions. The remainder of the presented study is arranged as follows. The next section consists of the details about the study region and data, followed by the background of the PET equations employed. Then, we explain the fundamental mechanism of the implemented water balance model and its performance criteria. The results and discussions obtained from the employed modeling strategy are presented next and the final section provides brief conclusions.

STUDY AREA AND DATA

In the study, Gediz Basin, which is located in the western part of Turkey, was selected as the study area (Figure 1). The basin, which has a drainage area of nearly 17,000 km^2, is fed by the Gediz River as well as several streams. In the basin, dominated by the Mediterranean climate, the annual precipitation regime is around 550 mm for the reference climate period of 1981–2010. The annual mean temperature for the same period is around 15 °C throughout the basin. The prevalent economic activity in the basin is agriculture and the reservoirs in the basin are therefore generally operated for irrigational purposes.

There are 39 meteorological stations operated by the Turkish State Meteorological Service (MGM) and the General Directorate of State Hydraulic Works (DSI) in the basin. The locations of these stations are given in Figure 1. In all of these stations, precipitation observations are made, while only 20 of them have temperature data. Nine flow gauging stations (FGSs) representing the basin were also identified. The data of the FGSs for the 1981–2010 water year period were obtained from the DSI. The station named Muradiye Bridge has the largest drainage area on the main branch and is located in the western part of the basin, while the Borlu, Topuzdamlari, Derekoy, and Acisu stations feeding the Demirkopru reservoir are also major FGSs. In addition, Kayalioglu and Hacihidir in the northwestern part of the basin, Hacihaliller in the southwest, and Taytan Bridge in the southeastern part of the basin were included in the study. The locations of the FGSs are also given in Figure 1, and general information about them is summarized in Table 1. The weights of the precipitation stations representing the FGSs listed were obtained using Thiessen polygons, while the areal mean, maximum, and minimum temperature series were computed by arithmetic mean approach. Thus, the inputs to be used in the water balance modeling for the drainage area of each FGS were compiled.

In the study, while mean, maximum, and minimum temperatures (T, T_{max}, and T_{min}) were compiled from the meteorological stations which have a regular distribution over the basin, ERA-Interim data sets having $0.75^\circ \times 0.75^\circ$ resolution were used for other variables needed for various PET equations, denoted in Table 2. The compliance of these re-analysis data sets with the observations has been

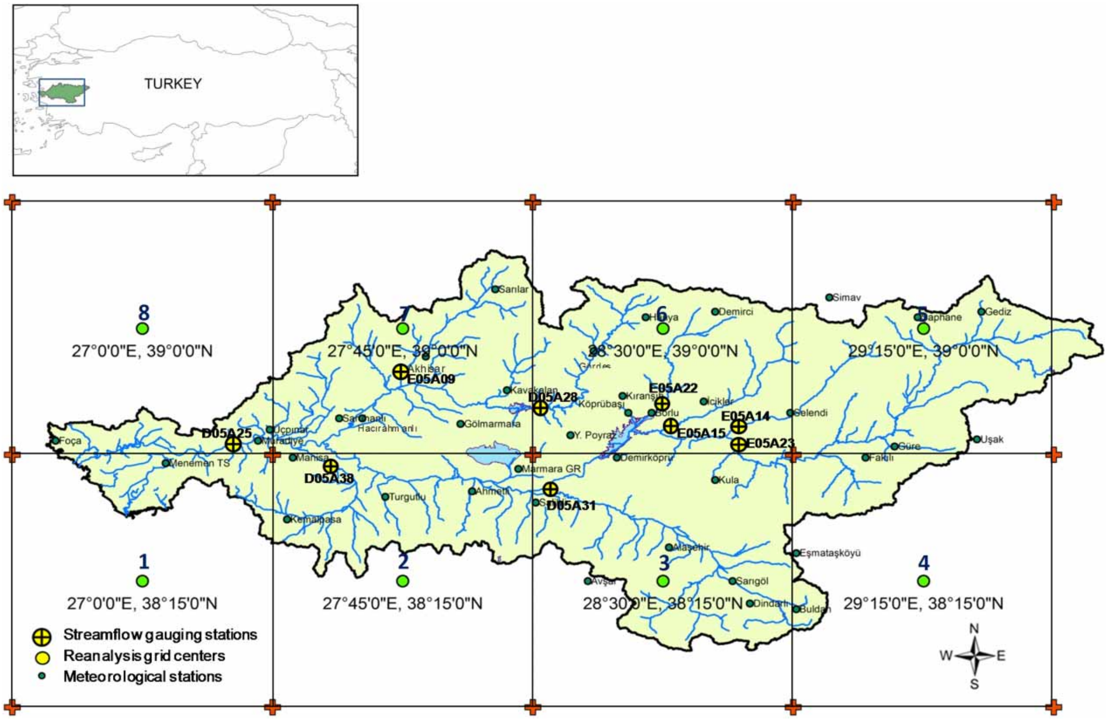

Figure 1 | Hydro-meteorological stations in the basin and eight ERA-Interim grids covering the study region.

verified by different researchers for various regions of the world (e.g., Dee *et al.* 2011; Bao & Zhang 2013). In Figure 1, eight ERA-Interim grids that almost uniformly comprise the study area and the center coordinates of the grids can be seen. These data sets are served by The European Center for Medium-Range Weather Forecasts for several categories. In addition, the grid numbers representing FGSs according to those of drainage areas are indicated in Table 1.

Before the application, the level of compatibility between ERA-Interim grid data and observations of meteorological stations was investigated. Since the temperature stations were regularly distributed in the basin, the average temperatures obtained from 20 stations and the arithmetic mean temperatures of the eight re-analysis grids were compared. Figure 2(a) demonstrates that the relationship is of good quality in terms of average temperatures. On

Table 1 | General information about FGSs used in the study

FGS	FGS code	Stream/Branch	Altitude (m)	Drainage area (km^2)	Representative grids
Borlu	E05A22	Demirci	245	818.8	Grid 6
Topuzdamlari	E05A15	Deliinis	381	739.6	Grid 6
Derekoy	E05A14	Selendi	345	689.6	Grid 6
Acisu	E05A23	Gediz	348	3,272.4	Grid 3 to 6
Taytan Bridge	D05A31	Alasehir	91	2,513.0	Grid 3
Kayalioglu	E05A09	Medar	77	901.6	Grid 7
Hacihaliller	D05A38	Nif	31	854.0	Grid 2
Muradiye Bridge	D05A25	Gediz	17	15,849.0	Grid 1 to 8
Hacihidir	D05A28	Gordes	305	808.2	Grid 6

Since the Muradiye Bridge and Acisu stations were represented by more than one grid, the arithmetic mean of the data of the relevant grids was used for them.

Table 2 | The empirical PET formulas used in the study

Method	Formula	Reference(s)
Thw	$PET = 16\ K_i(10T_i/J)^c$; $J = \Sigma(T_i/5)^{1.514}$ where I = January to December; $C = 0.000000675J^3 - 0.0000771J^2 + 0.01792J + 0.4924$. K constants can be provided from Ponce (1989)	Xu & Singh (2001), Pandey *et al.* (2016)
Rom	$PET = 0.0018(25 + T)^2(100\text{-}RH)$	Xu & Singh (1998)
Bl-Cr	$PET = kp(0.46T + 8.13)$. Based on Xu & Singh (2002), the coefficient of k was taken as 0.85 for April to September period and 0.45 for October to March period	Xu & Singh (2002)
Khr	$PET = 0.34pT^{1.3}$	Xu & Singh (2001)
Ham1	$PET = 0.6915N_m(D_L/12)^2 exp(0.062\ T)$	Xu & Singh (2001), Rosenberry *et al.* (2004)
Ham2	$PET = 0.1981N_m(D_L/12)exp(288.86e_s/(T + 273.3))$	Lu *et al.* (2005), Xystrakis & Matzarakis (2011)
Myr	$PET = A\ e_s(1 - RH/100)(1 + 0.225W_z/(z/8)^{0.15})$. (A can be taken as 11)	Singh & Xu (1997)
Pen	$PET = N_m 0.4655(1 + 0.24W_2)(e_s - e_a)$	Xu & Singh (1998)
Roh	$PET = N_m 0.44(1 + 0.27W_2)(e_s - e_a)$	Xu & Singh (2002)
Turc	$PET = N_m 0.013C_t\ (T/(15 + T))(23.8846R_s/N_m + 50)$ if $RH > 50\%$, C_t=1; if $RH \leq 50\%$, $C_t = 1 + (50 - RH)/70$	Xu & Singh (1998)
Harg	$PET = 0.0135(R_s/\lambda\rho)(T + 17.8)$	Xu & Singh (2000)
Mak	$PET = N_m(0.249(\Delta/(\Delta + \gamma))(R_s/N_m) - 0.12)$	Xystrakis & Matzarakis (2011)
Pen-Mon	$PET = \dfrac{N_m[(0.408\Delta\dfrac{(R_n - G_i)}{N_m} + 120\gamma W_2(e_s - e_a)/(T + 273)]}{\Delta + \gamma(1 + 0.34W_2)}$	Xu & Singh (2002), Allen *et al.* (1998)
Prs-Tyl	$PET = 0.514N_m(\Delta/(\Delta + \gamma))(R_n/N_m)$	Xu & Singh (2000)
Cpr	$PET = (6.1/10^6)(1000R_s)(1.8T + 1)$	Xystrakis & Matzarakis (2011)
J-H	$PET = (R_s/\lambda\rho)(0.025T + 0.08)$	Xystrakis & Matzarakis (2011)
Irm1	$PET = N_m(-0.611 + 0.149R_s/N_m + 0.079\ T)$	Irmak *et al.* (2003), Pandey *et al.* (2016)
Irm2	$PET = N_m(-0.642 + 0.174R_s/N_m + 0.0353\ T)$	Tabari *et al.* (2013), Pandey *et al.* (2016)
Irm3	$PET = N_m(-0.478 + 0.156R_s/N_m - 0.0112T_{max} + 0.0733T_{min})$	Tabari *et al.* (2013), Pandey *et al.* (2016)
McG-Bor	$PET = (0.0082(1.8T + 32) - 0.19)(23.8846R_s/1500)25.4$	Xu & Singh (2000)
Bai-Rob	$PET = N_m[0.0157T_{max} + 0.158(T_{max} - T_{min}) + 0.109R_a/N_m - 5.39]$	Pandey *et al.* (2016)

Where, N_m = total number of days in month m (since some methods give daily PETs in their original equations, the results are converted into monthly values with N_m); T = monthly mean temperature (°C); T_{max} = monthly maximum temperature (°C); T_{min} = monthly minimum temperature (°C); T_{dew} = monthly mean dew point temperature (°C); W_Z = monthly mean wind speed for z meter altitude (m/s); *Press* = monthly surface pressure (kPa); R_s = monthly solar radiation (MJ/m²); R_a = monthly extraterrestrial radiation (MJ/m²); R_n = monthly net radiation on the ground surface (MJ/m²); p = percentage of total daytime hours for the period used out of total daytime hours of the year; D_L = hours of daylight for a given month.
R_a, p, and D_L values can be provided from the FAO-56 report for different latitudes.

the other hand, surface pressure (*Press*), mean wind speed (W), and solar radiation (R_s) were measured at a rare station over the basin. Thus, an exemplary comparison was performed between the data of Akhisar meteorological station and the Grid 7 data for RH and R_s variables. Figure 2(b) and 2(c) support the fact that an adequate relationship can also be obtained with the data of RH derived from monthly mean temperature (T) and monthly mean dew point temperature (T_{dew}), and R_s served with ERA-Interim.

As predominant equations used in the study are particularly radiation based, it is essential to compile the radiation data (solar or net) at this stage. In the wake of extracting R_s

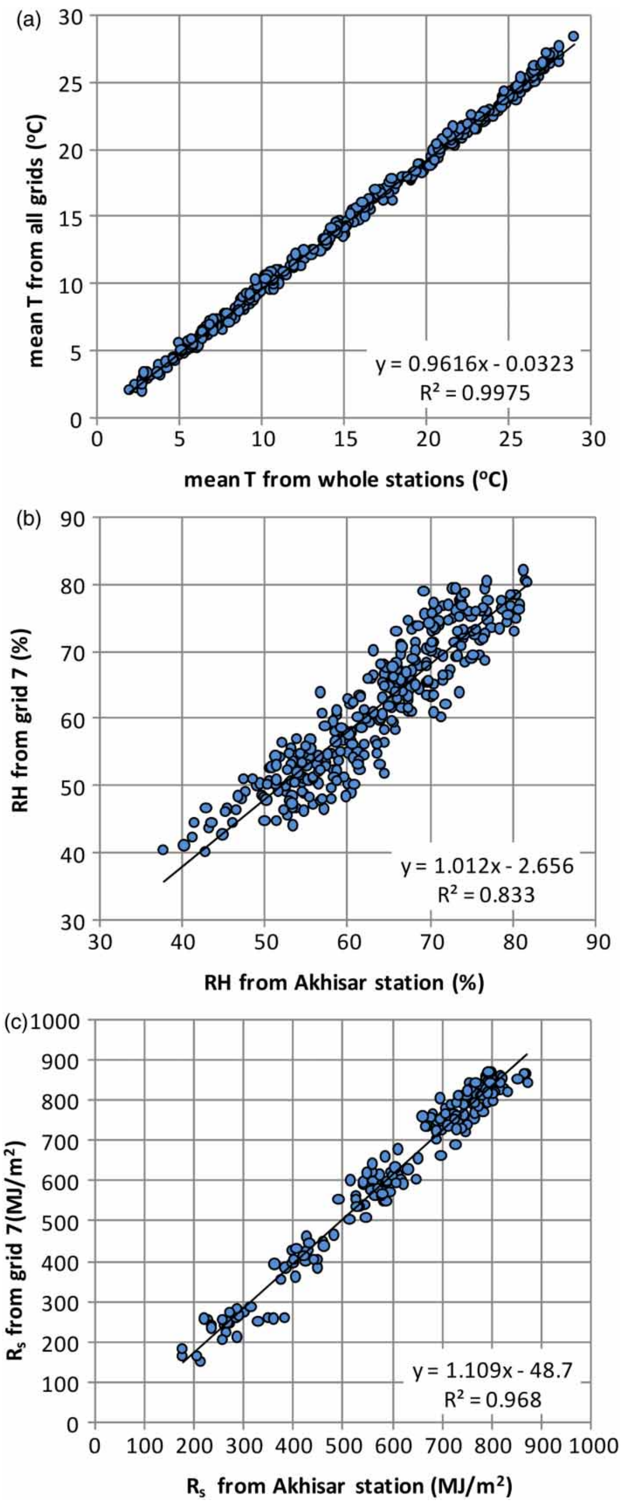

Figure 2 | Some validations by re-analysis data on the basis of station observations.

solar radiation data from the ERA-Interim database for the majority of radiation-based PET formulas, the difference between the R_{ns} net short-wave radiation and the R_{nl} net long-wave radiation, which gives the R_n net radiation on the ground surface, is also needed for other radiation-based equations such as Pen-Mon and Prs-Tyl. In this instance, it can be assumed that approximately 23% of R_s is reflected due to the albedo (Allen *et al.* 1998; Bogawski & Bednorz 2014), and the remaining part causes net short-wave radiation ($R_{ns} \approx 0.77R_s$). Here, R_{nl} values can also be estimated based upon the Stefan–Boltzmann law (see Bogawski & Bednorz 2014).

METHODOLOGY

Empirical PET equations used in the study

Various methods attributed to the prediction of PET can be found in the hydrology literature (Xu & Singh 2002; Xystrakis & Matzarakis 2011; Bogawski & Bednorz 2014; Pandey *et al.* 2016). These methods are generally evaluated in sub-categories as mass transfer, temperature-based, and radiation-based (Oudin *et al.* 2005a, 2005b). In this study, the effect of the PET predictions, which were separately generated from 21 equations, on the water balance modeling performances were investigated over the different FGSs in Gediz Basin, Turkey. Although many empirical PET equations and their modifications have been found in the literature, the focus of this study is on the equations that have been frequently cited. These equations are in different categories and they need several meteorological input sets. The employed equations such as Thornthwaite (Thw), Romanenko (Rom), Blaney–Criddle (Bl-Cr), Kharrufa (Khr), Hamon-1 (Ham1), and Hamon-2 (Ham2) can be evaluated in the category of temperature-based methods. Based on the research performed by Singh & Xu (1997), mass transfer methods such as Meyer (Myr), Rohwer (Roh), and Penman (Pen), which can be used in the estimation of surface evaporation, were also included in the study. The other utilized equations, such as Turc (Turc), Hargreaves (Harg), Makkink (Mak), Penman–Monteith (Pen-Mon), Priestley–Taylor (Prs-Tyl), Caprio (Cpr), Jensen–Haise (J-H), Irmak1 (Irm1), Irmak2 (Irm2), Irmak3 (Irm3), McGuinness–Bordne (McG-Bor), and Baier–Robertson (Bai-Rob), which demand more intensive data sets including solar, net, or extraterrestrial radiation, can be considered as radiation-based methods. Among them, as Pen-Mon and Prs-Tyl methods incorporate energy balance and aerodynamic water vapor mass transfer principles, they can also be known as combination methods (Oudin *et al.* 2005b). The formulas of empirical PET equations used in the study are given in Table 2. They were compiled from the references mentioned in the last column of Table 2. The weather data notations and their units are also indicated in this table. Additionally, Table 3 summarizes the information on several-supporter variables such as the temperature-based functions, wind speed scaling function, and some required constants.

Water balance model used in the study: *dynwbm*

Budyko (1958) argued that actual evapotranspiration (E_{act}) is a function of the precipitation (P) and the PET, and this relationship, which is derived for the annual time scale, is also referred to in the literature as the Budyko curve. Zhang *et al.* (2008) arranged this curve for monthly time scale data and integrated it into a water balance model termed dynamic water budget model (*dynwbm*), which is both conceptual and lumped. On the other hand, in the basins where there are dense agricultural activities, the groundwater table can be affected. In addition, the physical properties (land use or topographic patterns) may induce alteration in the groundwater level. Hence, an additional parameter, which is termed groundwater efficiency, was incorporated into the available *dynwbm* model by Okkan & Kirdemir (2018) in order to better conceptualize the groundwater storage process and to improve runoff prediction performances. Similarly, this five-parameter version was taken into consideration in the study.

In the model, the total amount of precipitation falling on the basin in any month is made up of two components, namely, direct runoff and catchment rainfall retention X, respectively. In this partition, the parameter α_1, which controls the first Budyko-type curve (see Figure 3, E1), plays an active role and a larger α_1 value results in more rainfall retention and less direct runoff. The model also has a sensitive parameter termed maximum soil moisture capacity (S_{max}), representing the soil and vegetation characteristics of the basin. On the other hand, the parameter α_2 controls evapotranspiration efficiency. In the case of an increase in this parameter value, an increment also occurs in the part of the water allocated to actual evapotranspiration E_{act} (Zhang *et al.* 2008; Li *et al.* 2016). The same parameter also controls the variable defined as evapotranspiration opportunity y, which is assumed to be composed of the sum of soil moisture content S and E_{act} (Sankarasubramanian & Vogel 2002). Thereby, y and E_{act} are both organized through the other Budyko-type curves (see E4 and E6 in Figure 3). During the month i, the available water content W_i may be expressed by the sum of the soil moisture remaining from the previous month (S_{i-1}) and X_i, as well as by the sum of the soil moisture content, actual evapotranspiration, and the amount of recharge (*Rec*) draining into groundwater storage. After *Rec*, S, and E_{act} are taken from the budget calculations, and baseflow prediction is made with the parameter d, the balance equation is then mounted for the groundwater storage G, which is postulated to represent linear reservoir behavior, using groundwater efficiency parameter e. The detailed description of the equations existing in the original model structure can be accessed from Zhang *et al.* (2008). The conceptual flowchart of the implemented version of the *dynwbm* model, its computation steps, and the related parameter definitions are given in Figure 3.

Table 3 | Some functions associated with temperature, wind speed scaling formula for 2 m altitude and basic constants (Allen *et al.* 1998; Xu & Singh 2001, 2002; Oudin *et al.* 2005b)

Variables	Functions/Constants
Saturation vapor pressure (mmHg)	$e_s = 4.5825 \exp[17.27\, T/(T + 237.3)]$
Actual vapor pressure (mmHg)	$e_a = 4.5825 \exp[17.27\, T_{dew}/(T_{dew} + 237.3)]$
Slope of vapor pressure curve (kPa/°C)	$\Delta = [546.4\, e_s]/[(T + 237.3)^2]$
Soil heat flux density (MJ/m²/month)	$G_i = 0.07(T_{i+1} - T_{i-1})$
Mean monthly relative humidity (%)	$RH = 100\, e_a/e_s$
Wind speed for 2 m altitude (m/s)	$W_2 = 4.87\, W_Z/(\ln(67.8z - 5.42))$
Water density	$\rho = 1{,}000$ kg/m³
Latent heat of vaporization	$\lambda \approx 2.45$ MJ/kg
Psychrometric constant (kPa/°C)	$\gamma = 0.00163$ (Press/λ)

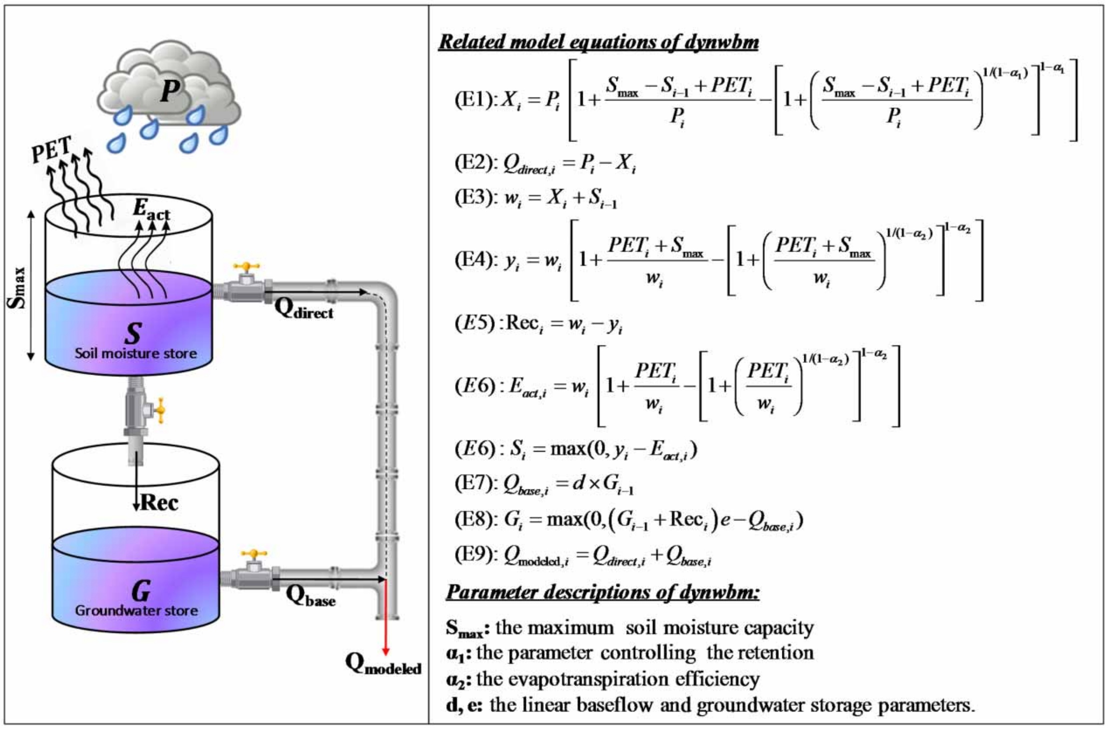

Figure 3 | Definitions and calculation steps for the *dynwbm* model.

Performance criteria for *dynwbm*

The numerical evaluation of *dynwbm* outputs is quite crucial to measuring the modeling performance. It is aimed at minimizing the root mean square error (*RMSE*) statistics in the calibration of the model. At this stage, the Levenberg–Marquardt (LM) algorithm was preferred since it is a highly qualified algorithm in the context of serial convergence and operation with only first order partial derivatives (Adeloye & Munari 2006). Additionally, several measures including Nash–Sutcliffe (*NS*) coefficient, and the RSR, which is the proportion of *RMSE* to standard deviation of observed data, were used to assess model performances, as recommended by Moriasi *et al.* (2007) and Okkan & Inan (2015). Formulations of the measures are given below.

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_{obs,i} - y_{m,i})^2} \quad (1)$$

$$NS = 1 - \frac{\sum_{i=1}^{n}(y_{obs,i} - y_{m,i})^2}{\sum_{i=1}^{n}(y_{obs,i} - y_{obs,mean})^2} \quad (2)$$

$$RSR = RMSE/S_{obs} \quad (3)$$

where n is the number of data during calibration or validation period, $y_{obs,i}$ and $y_{m,i}$ are the observed runoff data and modeled runoff in the i-th time period, respectively, $y_{obs,mean}$ and S_{obs} are the mean and standard deviation of observed runoff data during calibration or validation period, respectively.

RESULTS AND DISCUSSION

Sensitivity of hydrological model to different PET predictions

The long-term mean statistics collected from empirically derived PET predictions were examined prior to analyzing sensitivity of *dynwbm* to PETs. In this context, the dendrogram technique, which is a clustering analysis, was used to make an inference of long-term mean statistics of PETs throughout the basin. With this approach, the hierarchical relationship between the clusters represented in the mean statistics is summarized in Figure 4. The distance between the clusters was calculated with the Euclidean distance formula and Minitab package was used in the analysis. The lowest monthly average PET predictions were obtained from the Irm3, Bai-Rob, Ham1, Thw, and McG-Bor equations (67 mm/month, left-hand clustering in Figure 4), while Pen, Roh, and Rom were the highest predictions in the group (123 mm/month, right-hand clustering in Figure 4). In particular, it was considered that equations such as Pen and Roh, which were recommended for evaporations of water surfaces, were sensitive to the wind speed and therefore produced overestimation. While the long-term means range from 65 to 125 mm/month, the predictions of the Pen-Mon, the reference formula, are almost in the center of the two groups mentioned above (~95 mm/month).

After examining the long-term mean statistics, the effects of PET estimations on runoff modeling were assessed. In practice, data covering the 1981–1995 water year period were used for calibration, while data covering the 1996–2010 water year period were used in the validation phase. The optimization of the parameters of *dynwbm* during the calibration phase was based on the LM algorithm to minimize the RMSE. In the calibration, the optimization of the model with different initial conditions was repeated 30 times. Thus, we have tried to reduce the risk of getting the local minimums.

The model parameters determined for all FGSs and PET formula variations were evaluated at validation stage and the statistical performance metrics such as *RSR* and *NS* were calculated. For the example of Muradiye Bridge, *dynwbm* parameters calibrated under different PET inputs and the related validation performances are given in Figure 5. According to the coefficients of variation (C_v) regarding the calibrated parameters, the upper limit value of the soil moisture storage has a variability of up to 13% in the basin. As the drainage area of FGS grows, the uncertainty of the different PET equations over this parameter increases (especially Muradiye Bridge and Acisu). On the other hand, parameters *d* and *e* of the groundwater storage system were also subjected to significant variability depending on the PET method and flow regime. The fact that *d* and *e* were interdependent parameters in the same groundwater storage function made it difficult to generalize the PET-based uncertainties. In addition, the α_1 parameter, which allows precipitation to be converted to direct runoff, has no significant sensitivity (C_v < 3%). The increased inter-method variability in α_2 also increased the variability of E_{act}/PET ratio (15% < C_v < 25%).

In addition to the parameter-based interpretation, the main theme of the study is to examine the contribution of

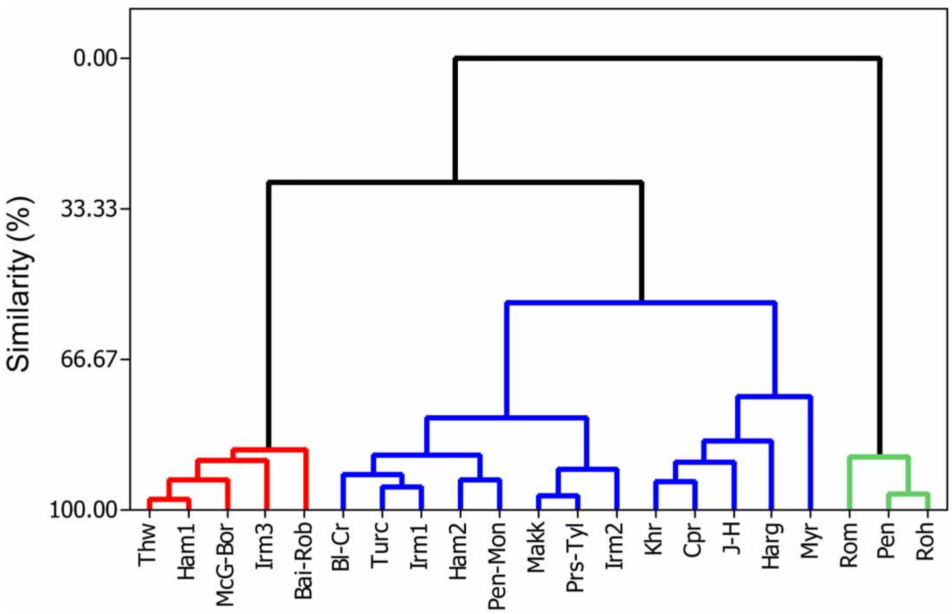

Figure 4 | Dendrogram analysis of the long-term mean PETs obtained throughout the basin.

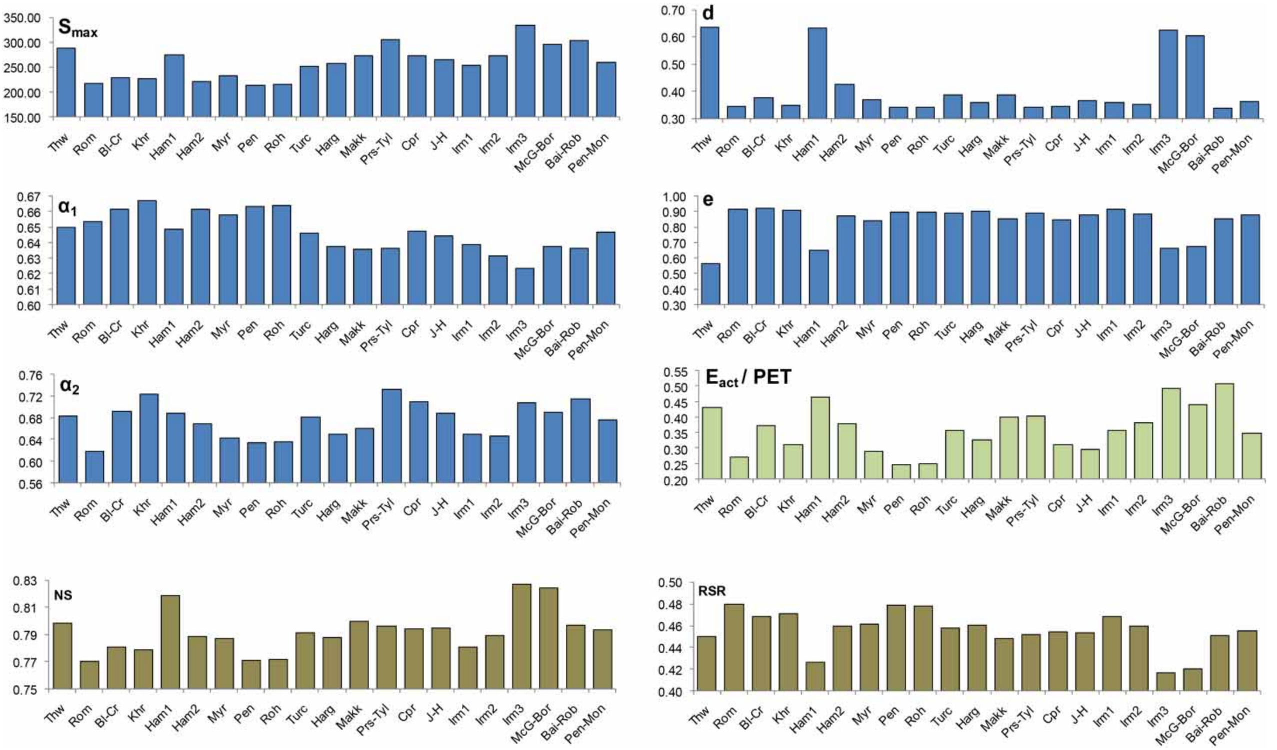

Figure 5 | Model parameters calibrated with predictions obtained under each PET equation and the NS and RSR performances pertaining to the validation period for Muradiye Bridge flow gauging station.

PET equations with different characteristics to the runoff generation capability. In this study, the question of which PET equations produced a higher quality runoff prediction was performed with the validation outputs of the *dynwbm*. In this context, only the NS results were handled because of their high correlation with RSR values (not shared in the paper). In order to scrutinize the results over the entire basin rather than an FGS-specific assessment, the NS performances obtained from nine FGSs were examined by means of a box-plot (Figure 6(a)). Moreover, it was thought that ranking of NS values (with ties) could provide an idea in evaluation (Figure 6(b)). As the small alterations among the NS values can bring about biases in the ranking, the values are rounded to two decimals. In Figure 6(a), it can be seen that both the medians and the inter-quartile ranges (IQRs) have significant dissimilarity depending on the method. In other words, the influence of the predictions obtained from the different PET equations on the water budget elements and hence the model performances cannot be denied. The median statistics varied from 0.77 to 0.81, and based on the criteria denoted by Moriasi *et al.* (2007), these values are attributed to 'very good' modeling. When the distances between IQRs and whiskers are also viewed, the wide range of indices extracted from the methods such as Bl-Cr, Rom, Ham2, Pen, Roh, Irm1, and Irm2 make their usability ineligible in the whole basin. It is quite apparent that Irm3, McG-Bor, and Prs-Tyl methods, which exhibit scattering above the threshold $NS = 0.75$, make a greater contribution to *dynwbm* for the basin. From the rank representations presented in Figure 6(b), the performance of the McG-Bor was found to be very close to that of the Irm3. In several respects, the notion is that McG-Bor is the most reasonable radiation-based formula for the scope of the study in terms of parsimony.

Investigation of parsimonious PET equations for hydrological model

In the results presented in the previous section, it was stated that the interaction of radiation-based McG-Bor with the

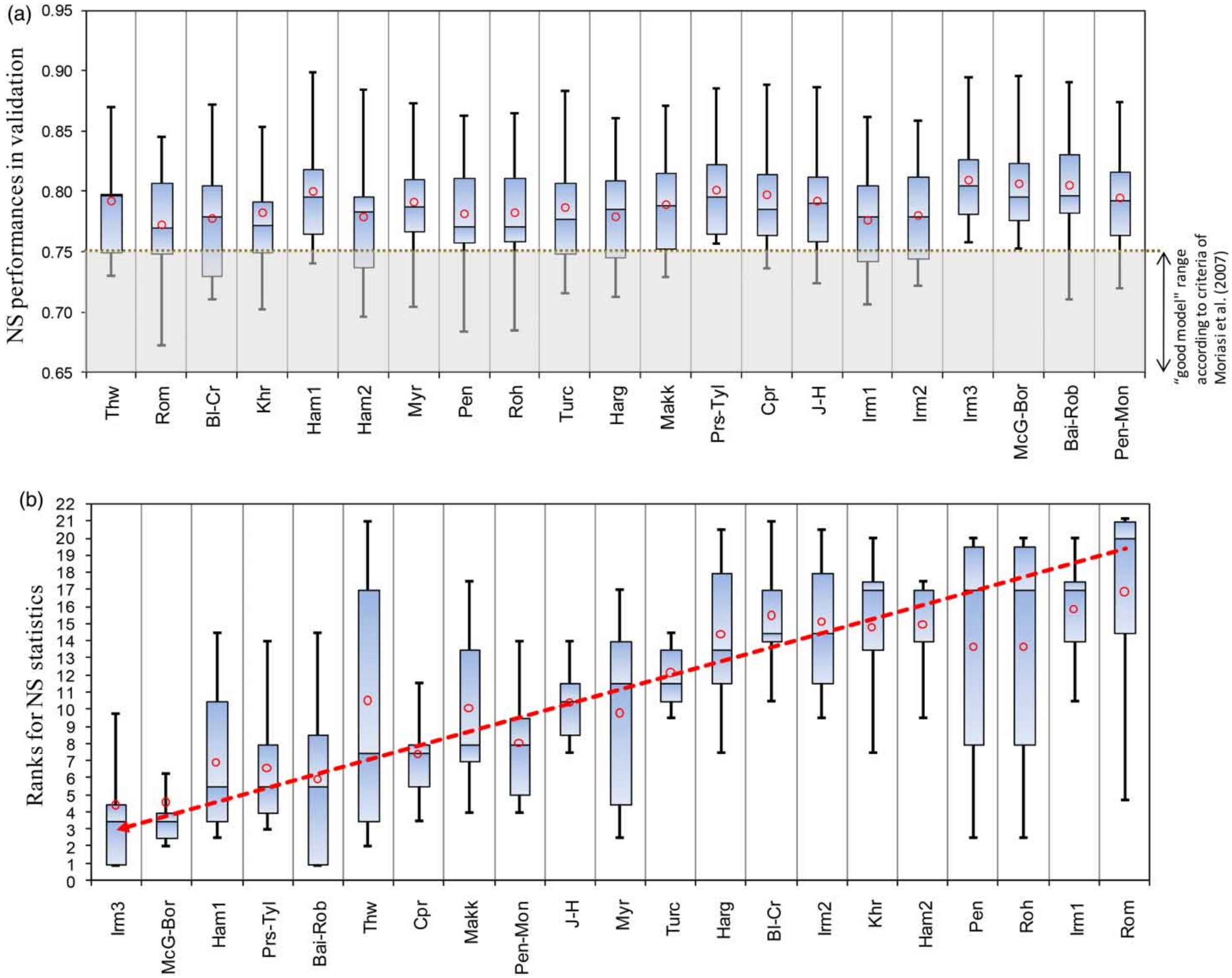

Figure 6 | Box diagram representation of (a) NS performances during the validation period of the *dynwbm* and (b) related ranking. The horizontal lines in the middle of the boxes and the circles represent the median and mean, respectively. The edges of the box are the 25th and 75th percentiles, and the whiskers extending to the extreme data points, not considered as outliers, are also indicated. The dashed line is the linear trend line drawn through the median of the boxes.

hydrological model in the basin was more consistent than the other types. However, the process of collecting or obtaining data related to solar radiation is not very practical. Therefore, it has been decided that the temperature-based Ham1 equation (first variant of Hamon equation) can be an alternative method since it is in the same dendrogram class as McG-Bor and it also displays relatively similar NS ranking score. Nevertheless, the results exhibited in the previous section (see Figure 6) have shown that it is required to recalibrate the defined constants involved in Ham1. When this empirical formula is taken as a basis, the general forms of various alternative equations (PET_{adj1}, PET_{adj2}, and PET_{adj3}), in which only the temperature data and day length are required, can be expressed with several parameters as follows:

$$PET_{adj1} = a\left(\frac{D_L}{12}\right)^b e^{cT} \tag{4}$$

$$PET_{adj2} = a\left(\frac{D_L}{12}\right)^b e^{cT} + d \tag{5}$$

$$PET_{adj3} = a\left(\frac{D_L}{12}\right)^b e^{cT} + d\,RH \tag{6}$$

In order to discover if further progression can be obtained, the parameters involved in the equations above have been recalibrated against the estimations of the McG-Bor method using an automatic optimization algorithm, Levenberg–Marquardt, as exerted in the *dynwbm* model as well. The cost function, CF, to be minimized can be stated as:

$$CF = \sqrt{\frac{\sum_{i=1}^{n} (PET_{McG-Bor,i} - PET_{adj,i})^2}{n}} = \text{minimum RMSE}, \quad (7)$$

where n is the number of data used, $PET_{McG\text{-}Bor}$ is the estimations computed by the McG-Bor, and PET_{adj} is the estimations derived from three other methods which consist of several parameters (a, b, and c in PET_{adj1}, a, b, c, and d in PET_{adj2} and PET_{adj3}).

The areal mean temperature data derived from all meteorological stations, which represent the total drainage area of the basin, were used to simulate the McG-Bor predictions by parsimonious temperature-based formulas denoted in Equations (4) to (6). In the calibration step, multiple initial values of parameters were tried to ensure that the global minimum of the RMSE was reached. As a result of the calibration, parameters a, b, and c defined in PET_{adj1} were 30.38, 2.49, and 0.0461, respectively (see Equation (8)). Even if this new equation reflects McG-Bor better than Ham1 (see Figure 7), it is questioned whether the systematic biases and the amount of RMSE can be reduced. In accordance with this purpose, the calibrated version of PET_{adj2}, having an additional constant term compared to the previous equation, is given in Equation (9). Ultimately, PET_{adj3} was offered, considering that RH values (in %) could be derived by T_{dew}, and the calibrated form is denoted in Equation (10). Based on the spatial mean temperature data in the basin, the scattering of the outputs produced by the operation of adjusted equations against those of

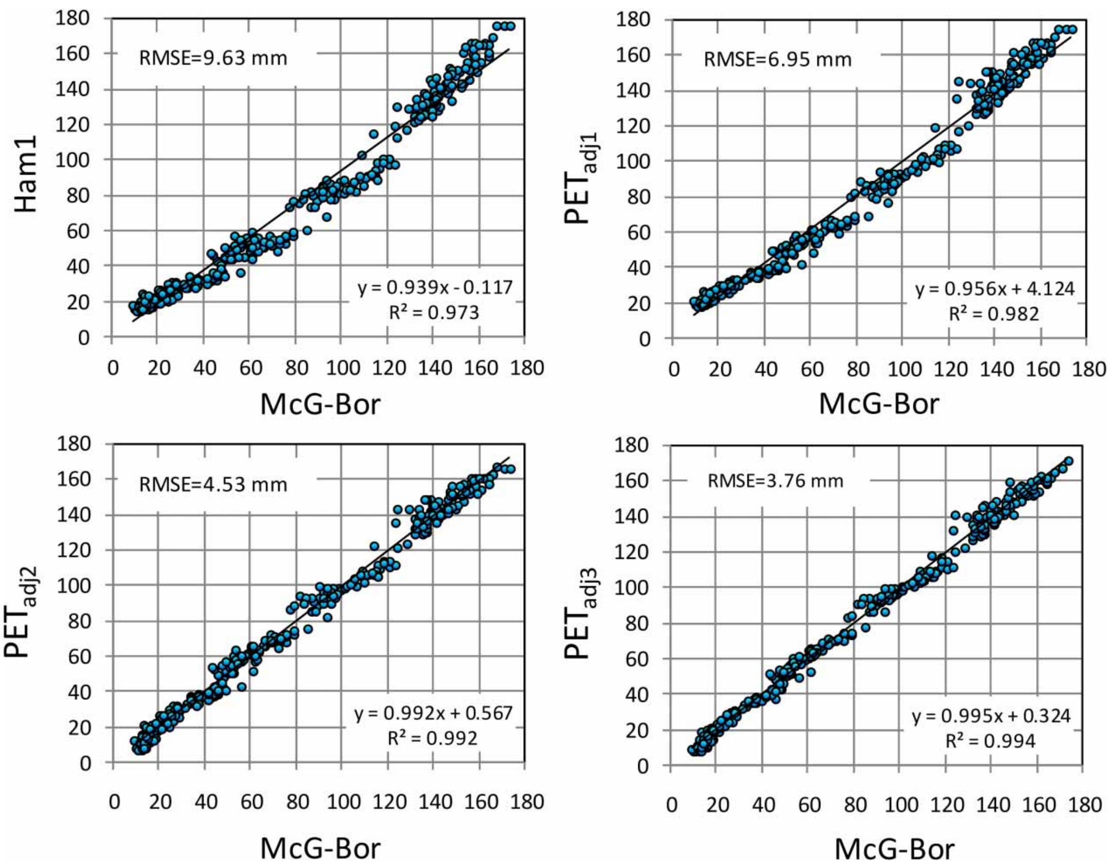

Figure 7 | Scattering of areal mean PET predictions produced through Ham1, PET_{adj1}, PET_{adj2}, and PET_{adj3} against those of McG-Bor.

McG-Bor are shown in Figure 7.

$$PET_{adj1} = 30.38\left(\frac{D_L}{12}\right)^{2.49} e^{0.0461T} \tag{8}$$

$$PET_{adj2} = 73.857\left(\frac{D_L}{12}\right)^{1.478} e^{0.0283T} - 47.293 \tag{9}$$

$$PET_{adj3} = 62.189\left(\frac{D_L}{12}\right)^{1.594} e^{0.0291T} - 0.490RH \tag{10}$$

As seen from Figure 7, the concordance of the predictions made by PET_{adj2} and PET_{adj3} with McG-Bor is more prominent compared to those of both Ham1 and PET_{adj1}. These equations were integrated into the hydrological model with the specified coefficients and *dynwbm*'s parameter optimization process on all FGSs was repeated. The validation performances obtained by the integration of the proposed PET equations into the hydrological model are given in Table 4 in comparison with the McG-Bor outputs. According to the results, PET_{adj3} showed almost identical validation performance with McG-Bor in the majority of stations (the small biases for Taytan Bridge, Kayalioglu, and Hacihaliller are negligible). Moreover, the usage of PET_{adj2} compared to McG-Bor formula is also functional and, at the same time, reasonable in case the *RH* data are not provided or computed. In other words, Table 4 has proved that the proposed equations requiring only temperature input are capable of competing with radiation-based equations.

Another idea is to write an intensive PET function containing four free parameters into the hydrological model considering the general structure defined in Equation (6). In this variation, which has been tried out but not comprehensively presented to the readers within the scope of the study, the current five-parameter *dynwbm* has become a model having nine free parameters to be calibrated. The common diagnosis monitored in those practices is the increase in the degree of freedom and a boosted performance in the calibration period, while a significant decrease in the validation performance is observed. An exemplary application supporting this finding is given in Figure 8 for Muradiye Bridge, which has the largest drainage area in the basin.

CONCLUSIONS

The performance of the conceptual water balance models depends on the input of PET in addition to precipitation. Thus, which PET equation should be used in the flow simulation stage is an essential issue, especially in arid basins. However, this problem has been discussed in relatively few studies (e.g., Oudin *et al.* 2005a, 2005b; Kannan *et al.* 2007). To contribute to the related literature, this study aimed to test various PET equations which can be presented as input to the *dynwbm* for Gediz Basin in Turkey, and to

Table 4 | Comparison of the impacts of McG-Bor and proposed equations on *dynwbm*

	NS			RSR		
FGSs	**McG-Bor**	PET_{adj2}	PET_{adj3}	**McG-Bor**	PET_{adj2}	PET_{adj3}
Borlu	0.9123	0.9134	0.9134	0.2954	0.2938	0.2931
Topuzdamlari	0.7848	0.7842	0.7968	0.4639	0.4646	0.4508
Derekoy	0.7571	0.7626	0.7722	0.4929	0.4873	0.4773
Acisu	0.7759	0.7718	0.7887	0.4734	0.4777	0.4597
Taytan Bridge	0.8183	0.8130	0.8162	0.4263	0.4324	0.4287
Kayalioglu	0.7526	0.7457	0.7465	0.4974	0.5043	0.5035
Hacihaliller	0.7954	0.7952	0.7948	0.4523	0.4526	0.4530
Muradiye Bridge	0.8240	0.8218	0.8242	0.4196	0.4221	0.4192
Hacihidir	0.8554	0.8541	0.8587	0.3802	0.3820	0.3759

The shaded values denote the best performance indices.

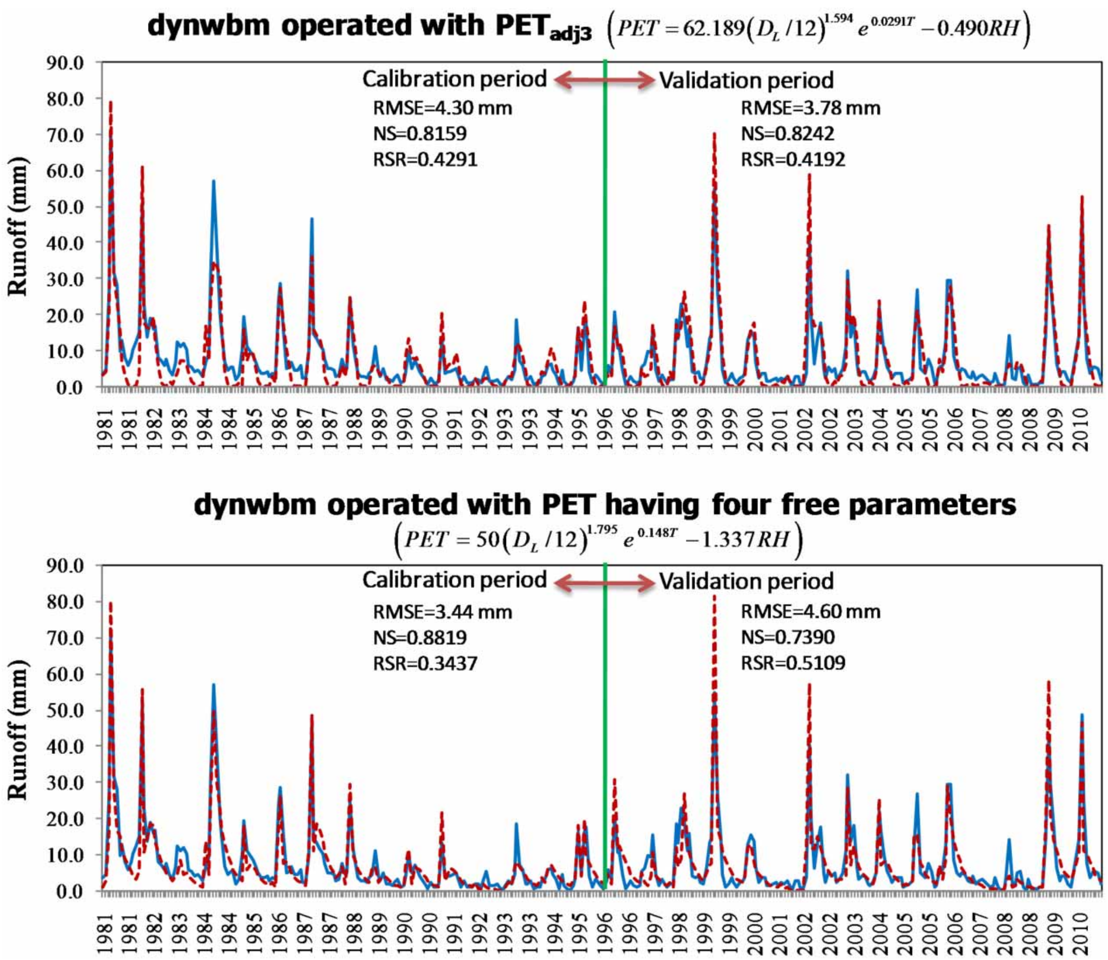

Figure 8 | Runoff outputs of water balance model that is run by the predictions of PET$_{adj3}$ and operated by PET function having four free parameters to be calibrated (the straight line and the dashed line denote the observed runoff and modeled runoff, respectively).

determine alternative equations that use less meteorological data. The accordance of *dynwbm* with the PET equations was questioned by means of the several indices in the validation period. In contrast to the idea defended in the study by Wang *et al.* (2006), it appears at first glance that the effects of different PET equations on the water balance model cannot be neglected. According to performance criteria, the radiation-based McG-Bor equation, which causes the hydrological model over the basin to behave in the 'very good' class, appears to be superior. According to the NS ranking assessment applied throughout the basin, it was found that the McG-Bor equation resulted in similar responses to the hydrological model as the temperature-based Ham1 equation. Since the Ham1 is an economical method in terms of data requirement and because of the difficulty of obtaining solar radiation in the basin, it is accepted as an alternative method. Another viewpoint is to recalibrate some of the constants defined in the Ham1 equation to simulate the McG-Bor predictions. In accordance with this alternative, PET$_{adj2}$ and PET$_{adj3}$ were proposed as a result of the various arrangements performed. When the PET predictions derived from these parsimonious equations were presented as inputs to the water balance model, and the validation outputs were evaluated following the calibration stage, it was detected that PET$_{adj3}$ in particular provided a rather similar performance to McG-Bor throughout the whole basin. Thus, it has been determined that the proposed equations have more practical usage compared to radiation-based equations.

Due to the limited studies conducted in the literature connected with the topic, it is thought that the presented study is an inspirational one in terms of the content and suggested methodology. Since all the deductions have been made for a semi-arid region in Turkey, the derived equations embedded into water balance modeling can also be adapted to problems such as missing flow completion, discharge simulation under climate change scenarios, and defining supply–demand relationships of a reservoir for the neighboring basins and regions dominated by Mediterranean climate characteristics. For future development, the aim is to widen the scope of the study by taking into account the climate change scenarios (representative concentration pathways) attributed to the Fifth Report of the Intergovernmental Panel on Climate Change. In addition, there will also be a focus on both model-based (e.g., downscaling model, PET model, rainfall–runoff model) and scenario-based uncertainties.

REFERENCES

Adeloye, A. J. & Munari, A. D. 2006 Artificial neural network based generalized storage–yield–reliability models using the Levenberg–Marquardt algorithm. *Journal of Hydrology* **326** (1–4), 215–230.

Allen, R. G., Pereira, L. S., Raes, D. & Smith, M. 1998 *Crop Evapotranspiration – Guidelines for Computing Crop Water Requirements*. FAO Irrigation and Drainage Paper, Paper No. 56, FAO, Rome, Italy, p. 300.

Allen, R. G., Walter, I. A., Elliot, R., Howell, T., Itenfisu, D. & Jensen, M. 2005 *The ASCE Standardized Reference Evapotranspiration Equation*. Final report. National Irrigation Symposium, Phoenix, AZ, USA.

Bao, X. & Zhang, F. 2013 Evaluation of NCEP–CFSR, NCEP–NCAR, ERA-Interim, and ERA-40 reanalysis datasets against independent sounding observations over the Tibetan Plateau. *Journal of Climate* **26** (1), 206–214.

Bárdossy, A. & Das, T. 2008 Influence of rainfall observation network on model calibration and application. *Hydrology and Earth System Sciences* **12** (1), 77–89.

Beyazgül, M., Kayam, Y. & Engelsman, F. 2000 Estimation methods for crop water requirements in the Gediz Basin of western Turkey. *Journal of Hydrology* **229** (1–2), 19–26.

Blaney, H. F. & Criddle, W. D. 1950 *Determining Water Requirements in Irrigated Areas From Climatological and Irrigation Data*. US Soil Conservation Service, Washington, DC, USA.

Bogawski, P. & Bednorz, E. 2014 Comparison and validation of selected evapotranspiration models for conditions in Poland (Central Europe). *Water Resources Management* **28** (14), 5021–5038.

Budyko, M. I. 1958 *The Heat Balance of the Earth's Surface*. US Department of Commerce, Washington, DC, USA.

Caldwell, P. V., Kennen, J. G., Sun, G., Kiang, J. E., Butcher, J. B., Eddy, M. C., Hay, L. E., LaFontaine, J. H., Hain, E. F., Nelson, S. A. C. & McNulty, S. G. 2015 A comparison of hydrologic models for ecological flows and water availability. *Ecohydrology* **8** (8), 1525–1546.

Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Alonso-Balmaseda, M., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A., van de Berg, L., Bidlot, J.-R., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A. J., Haimberger, L., Healy, S., Hersbach, H., Hólm, E. V., Isaksen, L., Kållberg, P. W., Köhler, M., Matricardi, M., McNally, A., Monge-Sanz, B. M., Morcrette, J.-J., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J.-N. & Vitart, F. 2011 The ERA-Interim reanalysis: configuration and performance of the data assimilation system. *Quarterly Journal of the Royal Meteorological Society* **137**, 553–597.

Fowler, H. J., Kilsby, C. G. & Stunell, J. 2007 Modelling the impacts of projected future climate change on water resources in north-west England. *Hydrology and Earth System Sciences* **11** (3), 1115–1126.

Hamon, W. R. 1961 Estimating potential evapotranspiration. *Journal of the Hydraulics Division, American Society of Civil Engineers* **87** (3), 107–120.

Irmak, S., Irmak, A., Allen, R. G. & Jones, J. W. 2003 Solar and net radiation-based equations to estimate reference evapotranspiration in humid climates. *Journal of Irrigation and Drainage Engineering* **129** (5), 336–347.

Itenfisu, D., Elliott, R. L., Allen, R. G. & Walter, I. A. 2003 Comparison of reference evapotranspiration calculations as part of the ASCE standardization effort. *Journal of Irrigation and Drainage Engineering* **129** (6), 440–448.

Jain, S. K., Nayak, P. C. & Sudheer, K. P. 2008 Models for estimating evapotranspiration using artificial neural networks, and their physical interpretation. *Hydrological Processes* **22** (13), 2225–2234.

Jensen, M. E. & Haise, H. R. 1963 Estimating evapotranspiration from solar radiation. Proceedings of the American Society of Civil Engineers. *Journal of the Irrigation and Drainage Division* **89**, 15–41.

Kannan, N., White, S., Worrall, F. & Whelan, M. 2007 Sensitivity analysis and identification of the best evapotranspiration and runoff options for hydrological modelling in SWAT-2000. *Journal of Hydrology* **332** (3–4), 456–466.

Kellner, E. 2001 Surface energy fluxes and control of evapotranspiration from a Swedish Sphagnum mire. *Agricultural and Forest Meteorology* **110** (2), 101–123.

Li, S., Xiong, L., Li, H., Leung, L. R. & Demissie, Y. 2016 Attributing runoff changes to climate variability and human activities: uncertainty analysis using four monthly water balance models. *Stochastic Environmental Research and Risk Assessment* **30** (1), 251–269.

Lu, J., Sun, G., McNulty, S. G. & Amatya, D. M. 2005 A comparison of six potential evapotranspiration methods for regional use in the Southeastern United States. *Journal of the American Water Resources Association* **41** (3), 621–633.

Makkink, G. F. 1957 Testing the Penman formula by means of lysimeters. *Journal of the Institution of Water Engineering* **11** (3), 277–288.

McGuinness, J. L. & Bordne, E. F. 1972 *A Comparison of Lysimeter Derived Potential Evapotranspiration with Computed Values*. Technical Bulletin 1452, Agricultural Research Service, US Department of Agriculture, Washington, DC, USA.

Mckillop, R., Kouwen, N. & Soulis, E. D. 1999 Modeling the rainfall-runoff response of a headwater wetland. *Water Resources Research* **35** (4), 1165–1177.

Monteith, J. L. 1965 Evaporation and the environment. In the state and movement of water in living organisms. In: *Proceedings of the 19th Symposium, Society of Experimental Biology*. Cambridge University Press, London, UK.

Moriasi, D. N., Arnold, J. G., Liew, M. W., Bingner, R. L., Harmel, R. D. & Veith, T. L. 2007 Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. *Transactions of the ASABE* **50** (3), 885–900.

Okkan, U. & Fistikoglu, O. 2014 Evaluating climate change effects on runoff by statistical downscaling and hydrological model GR2M. *Theoretical and Applied Climatology* **117** (1), 343–361.

Okkan, U. & Inan, G. 2015 Statistical downscaling of monthly reservoir inflows for Kemer watershed in Turkey: use of machine learning methods, multiple GCMs and emission scenarios. *International Journal of Climatology* **35** (11), 3274–3295.

Okkan, U. & Kirdemir, U. 2018 Investigation of the behavior of an agricultural-operated dam reservoir under RCP scenarios of AR5-IPCC. *Water Resources Management* **32** (8), 2847–2866.

Oudin, L., Michel, C. & Anctil, F. 2005a Which potential evapotranspiration input for a lumped rainfall-runoff model? Part 1 – Can rainfall-runoff models effectively handle detailed potential evapotranspiration inputs? *Journal of Hydrology* **303** (1–4), 275–289.

Oudin, L., Hervieu, F., Michel, C., Perrin, C., Andréassian, V., Anctil, F. & Loumagne, C. 2005b Which potential evapotranspiration input for a lumped rainfall–runoff model? part 2 – towards a simple and efficient potential evapotranspiration model for rainfall-runoff modelling. *Journal of Hydrology* **303** (1–4), 290–306.

Ozkul, S. 2009 Assessment of climate change effects in Aegean river basins: the case of Gediz and Buyuk Menderes Basins. *Climatic Change* **97** (1–2), 253–283.

Pandey, P. K., Dabral, P. P. & Pandey, V. 2016 Evaluation of reference evapotranspiration methods for the northeastern region of India. *International Soil and Water Conservation Research* **4** (1), 52–63.

Paturel, J., Servat, E. & Vassiliadis, A. 1995 Sensitivity of conceptual rainfall-runoff algorithms to errors in input data-case of the GR2M model. *Journal of Hydrology* **168** (1–4), 111–125.

Penman, H. L. 1948 Natural evaporation from open water, bare soil and grass. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences* **193**, 120–145.

Ponce, V. M. 1989 *Engineering Hydrology: Principles and Practices*. Pearson College Division, Englewood Cliffs, NJ, USA.

Priestley, C. H. B. & Taylor, R. J. 1972 On the assessment of surface heat fluxes and evaporation using large-scale parameters. *Monthly Weather Review* **100**, 81–92.

Rosenberry, D. O., Stannard, D. I., Winter, T. C. & Martinez, M. L. 2004 Comparison of 13 equations for determining evapotranspiration from a prairie wetland, Cottonwood Lake Area, North Dakota, USA. *Wetlands* **24** (3), 483–497.

Sankarasubramanian, A. & Vogel, R. M. 2002 Annual hydroclimatology of the United States. *Water Resources Research* **38** (6), 19-1–19-12.

Singh, V. P. & Xu, C. 1997 Evaluation and generalization of 13 mass-transfer equations for determining free water evaporation. *Hydrological Processes* **11** (3), 311–323.

Tabari, H., Grismer, M. E. & Trajkovic, S. 2013 Comparative analysis of 31 reference evapotranspiration methods under humid conditions. *Irrigation Science* **31** (2), 107–117.

Thornthwaite, C. W. 1948 An approach toward a rational classification of climate. *Geographical Review* **38** (1), 55–94.

Vandewiele, G. L., Xu, C. & Huybrechts, W. 1991 Regionalisation of physically-based water balance models in Belgium. application to ungauged catchments. *Water Resources Management* **5** (3–4), 199–208.

Wang, X., Melesse, A. M. & Yang, W. 2006 Influences of potential evapotranspiration estimation methods on SWAT's hydrologic simulation in a Northwestern Minnesota Watershed. *Transactions of the ASABE* **49** (6), 1755–1771.

Xu, C. & Singh, V. P. 1998 Dependence of evaporation on meteorological variables at different time-scales and intercomparison of estimation methods. *Hydrological Processes* **12** (3), 429–442.

Xu, C. & Singh, V. P. 2000 Evaluation and generalization of radiation-based methods for calculating evaporation. *Hydrological Processes* **14** (2), 339–349.

Xu, C. & Singh, V. P. 2001 Evaluation and generalization of temperature-based methods for calculating evaporation. *Hydrological Processes* **15** (2), 305–319.

Xu, C. & Singh, V. P. 2002 Cross comparison of empirical equations for calculating potential evapotranspiration with data from Switzerland. *Water Resources Management* **16** (3), 197–219.

Xu, C. & Vandewiele, G. 1995 Parsimonious monthly rainfall-runoff models for humid basins with different input requirements. *Advances in Water Resources* **18** (1), 39–48.

Xystrakis, F. & Matzarakis, A. 2011 Evaluation of 13 empirical reference potential evapotranspiration equations on the island of Crete in Southern Greece. *Journal of Irrigation and Drainage Engineering* **137** (4), 211–222.

Zhang, L., Potter, N., Hickel, K., Zhang, Y. & Shao, Q. 2008 Water balance modeling over variable time scales based on the Budyko framework – model development and testing. *Journal of Hydrology* **360** (1–4), 117–131.

First received 17 December 2018; accepted in revised form 9 April 2019. Available online 29 April 2019