Ecological Processes

RESEARCH

Open Access

Applicability and improvement of different potential evapotranspiration models in different climate zones of China

Zedong Li[1], Yiran Li[1], Xinxiao Yu[1,3*], Guodong Jia[1], Peng Chen[2], Pengfei Zheng[1], Yusong Wang[1] and Bingbing Ding[1]

Abstract

Background Accurate estimation of potential evapotranspiration (PET) is the key for studying land-air interaction hydrological processes. Several models are used to estimate the PET based on standardized meteorological data. Although combination-based models have the highest level performance estimation of PET, they require more meteorological data and may therefore be difficult to apply in areas lacking meteorological observation data.

Results The results showed significant differences in the spatial trends of PET calculated by different models in China, the Doorenbots–Pruitts model revealed the highest PET (1902.6 mm), and the Kuzmin model revealed the lowest PET (349.6 mm), with the largest difference being 5.5 times. The Romanenko and the Rohwer models were the recommended temperature-based and aerodynamic-based models. On the other hand, the Abtew model was more suitable for arid and semi-arid regions, while the Priestley–Taylor model was more suitable for humid regions. Combination-based models revealed ideal calculation accuracies, among which the Penman–Monteith model was the best option for PET calculation.

Conclusions The accuracy range of Romanenko, Rohwer, Abten, Priestley Taylor, and Penman Monteith models improved in MPZ and TCZ is higher than that improved in TMZ and SMZ. This does not mean that the improved models have higher accuracy in MPZ and TCZ than in TMZ and SMZ. On the contrary, the original model performed poorly in MPZ and TCZ, so the improved accuracy was relatively large. The unimproved model was already more suitable in TMZ and SMZ, so the improved accuracy was relatively small. Therefore, regional calibration of the PET models can improve the accuracy and applicability of PET calculation, providing a reference for studying hydrological processes in different climatic zones.

Keywords Potential evapotranspiration, Climatic zone, Temperature-based model, Aerodynamic-based model, Radiation-based model, Combination-based model, Model calibration

*Correspondence:
Xinxiao Yu
yuxinxiao2021@126.com
Full list of author information is available at the end of the article

Introduction

Evapotranspiration (ET) has long been recognized as a major component of the hydrologic cycle. In fact, it is the largest component of the hydrological budget behind precipitation (Gharbia et al. 2018), connecting the water, energy, and carbon cycles. In addition, evapotranspiration plays a key role in the land–atmosphere interaction system (Wang et al. 2011; Li et al. 2021) and has been widely used in extreme weather monitoring and sustainable water resource utilization evaluation (Sheffield et al. 2012; Duethmann and Blöschl 2018). However, accurate quantification of actual evapotranspiration is difficult due to the complex interaction between meteorological factors and underlying surface factors, making other variables often necessary to represent it, such as potential evapotranspiration (PET), reference crop evapotranspiration (ET_0), and pan evapotranspiration (Epan) (Zheng et al. 2017; Anabalón and Sharma 2017). PET has been widely used in watershed hydrological models to estimate runoff and aridity assessment (Douglas et al. 2009; Zhou et al. 2020). Moreover, long-term changes of PET have been used in climate-based prediction models to assess the impacts of climate change on ecosystems (Milly and Dunne 2016; Yang et al. 2021).

PET was first proposed by Thornthwaite (1948). It refers to the maximum amount of water loss through the evapotranspiration pathway under optimal conditions. However, for decades, scientists have defined PET in different ways and have even considered PET and ET_0 as the same concept. The use of "potential" by Hargreaves and Samani (1982) and "reference crop" by Hargreaves and Samani (1985) is a typical example. PET can be estimated by numerous models using observed meteorological data. These models are classified into temperature-based models, aerodynamic-based models, radiation-based models, and combination-based models according to different assumptions and input data requirements (Bormann 2011; Zhao et al. 2013; Xiang et al. 2020; Zhou et al. 2020; Yang et al. 2021). Temperature-based PET estimate was first proposed by Thornthwaite (1948) using the Thornthwaite model, then other temperature-based models such as the Baier–Robertson model, Blaney–Criddle model, Kharrufa model, and Oudin model were proposed (Blaney and Criddle 1950; Baier and Robertson 1965; Kharrufa 1985; Oudin et al. 2005). This model category has been extensively applied in PET estimation worldwide due to the fewer and simple data parameters involved and its reliable performance (Xiang et al. 2020; Yang et al. 2021). Aerodynamic-based PET models are derived from the Dalton model. However, several researchers have highlighted highly biased results of this model category as its underlying mechanism is significantly different from the current PET methods (Bormann 2011; Xiang et al. 2020), and PET can only be estimated after adjusting the parameters in the lack of meteorological data (Azhar and Perera 2011; Jakimavičius et al. 2013; Valipour 2014). Radiation-based models reveal the relationship between PET and radiation using empirical equations, such as the Priestley–Taylor equation, which is a simplified form of the Penman–Monteith equation (Jensen and Haise 1965). Although the Priestley–Tayor equation has been widely used in PET estimation worldwide, it may have a false basic premise that large-scale advection is not independent of surface energy balance (McNaughton and Spriggs 1988), resulting in significant variation in the accuracy of the estimate across regions. The combined category model considers both aerodynamics and energy balance, as ET pathway involve both evaporation and motion processes. Penman (1948) proposed the first typical PET combination category equation, considering the effect of temperature and other meteorological factors. In recent decades, numerous researchers have proposed Penman-based equations (Rijtema and Wageningen 1965; Van Bavel 1966; Wright and Jensen 1972; Thom and Oliver 1977), of which the most widely used is the Penman–Monteith equation, which includes a new coefficient representing crop surface roughness (*rs*) (Monteith 1965). The Food and Agriculture Organization (FAO) later recommended the Penman–Monteith equation as a standardized method for estimating ET_0 (Allen et al. 1998). However, its application is difficult and ambiguous as some scientists use it to estimate PET and others to estimate ET_0.

Numerous studies have assessed the reliability of PET models under different climates and underlying surface conditions. In fact, most studies have confirmed that the combined model has the best estimation accuracy of PET, followed, respectively, by radiation, aerodynamic and temperature-based models (Singh and Xu 1997; Douglas et al. 2009; Zuo et al. 2009; Bormann 2011; Valipour 2015; Zhou et al. 2020; Yang et al. 2021). However, almost all combined models require several climate variables, including net radiation, wind speed, and soil heat flux density, which are often difficult to obtain, especially in regions where extensive and reliable meteorological data are lacking, explaining the reason for using other model categories. Given the wide variety of PET models, several models have been developed for specific hydrological conditions and evaluated with limited measured data (Tegos et al. 2015), making it difficult to select the best model for estimating PET in a given area.

According to climatic characteristics, the climate in China is classified into five climatic zones, namely the mountain plateau zone (MPZ), temperate continental

zone (TCZ), temperate monsoon zone (TMZ), subtropical monsoon zone (SMZ), and tropical monsoon zone (PMZ) (Song et al. 2011). The PMZ was considered as adjacent SMZ, as reported by other studies, due to its small area (Fan et al. 2016; Yang et al. 2021). The SMZ, TMZ, MPZ, and TCZ are humid, semi-humid, semi-arid, and arid regions, respectively (Fan et al. 2016; Feng et al. 2018). Indeed, studies have revealed a significant variation in the PET estimated by different models among the four climatic zones (Zhou et al. 2020; Yang et al. 2021). Therefore, it is necessary to evaluate the reliability of PET models in different climatic zones.

Since PET is a theoretical concept that is difficult to be measured directly (Prudhomme and Williamson 2013), indirect estimation methods have been commonly used, including the lysimeter method (Herrnegger et al. 2012), scintillation method (Xu et al. 2013; Yee et al. 2015), Bowen ratio method (Douglas et al. 2009), eddy covariance method (Li et al. 2016; Zheng et al. 2017), and pan evaporation method (Xu and Singh 2001; Xu et al. 2016). At the watershed scale, some scientists have used hydrological models (Bai et al. 2016) and water balance methods (Lu et al. 2005) to validate the estimated PET. The pan evaporation (Epan) considers the effect of climatic factors in the PET determination, including temperature, radiation, and wind speed. This method is simple to operate and inexpensive. The FAO recommends using a class A Epan as a standard instrument for determining the PET, with a diameter and depth of 1.21 and 25.5 cm, respectively. This class has been commonly used worldwide in recent years and provided satisfactory results (Padmakumari et al. 2013; Azorin-Molina et al. 2015; Liu and Sun 2016). It should be pointed out that the amounts of Epan are not equal to PET rates, but in theory, the two variables should have a significant correlation (Liu et al. 2010). Epan integrates the effects of radiation, humidity, wind and air temperature on wet surface evaporation rates, and so provides measurements of evaporation that are concep-tually similar to PET rates (McVicar et al. 2007; Donohue et al. 2010). Therefore, the Epan was used in this study to validate the PET rate determined by several models. In order to use Epan in the PET model accuracy assessment, we assumed that Epan and PET are highly correlated. This study aims to: (1) use 30 PET models from four categories to estimate the PET values and trends during the 2011–2020 period in different climatic zones of China; (2) recommend the most suitable models of different categories with measured Epan values; (3) calibrate the recommended models of different climatic zones.

Materials and methods

Study area

China is located in the East Asia with a territory of more than 9.6 million square kilometers. Its topography is characterized by a three-trapezoidal distribution with highest mountains in the west and lowest plains in the east. The annual precipitation over China varies from 50 mm in the arid northwest to 2000 mm in the humid southeast. The climate is dominated by changes in winter and summer monsoons, which are significantly influenced by ENSO (Ding et al. 2014; Li et al. 2021). In summer, warm and humid south wind blows, however, in winter, mainly cold and dry northerly wind prevails (Ge et al. 2017).

Data collection

Daily meteorological data observed at 699 stations were obtained from the China Meteorological Administration (http://data.cma.cn/). The climate variables observed include precipitation (mm), air temperature (°C), relative humidity (%), atmospheric pressure (hPa), wind speed at 10 m height (m/s), and sunshine duration (h). The data used in this study were from 699 meteorological stations, including 87, 119, 195, and 298 in the MPZ, TCZ, TMZ, and SMZ, respectively (Fig. 1). The average elevation of meteorological stations, Epan evaporation, annual mean precipitation, air temperature, and aridity index obtained over the 2001–2020 period in the four climatic zones are reported in Additional file 1: Table S1). The missing data were interpolated based on linear regression relationships with those of neighboring stations. These meteorological stations were spatially well-distributed, thus reflecting the general characteristics of the regional climate in China. Meteorological data were transformed into raster data using Anusplin spatial interpolation software, covering entire China's area, with a resolution of 1 km. Radiation, vapor pressure, and

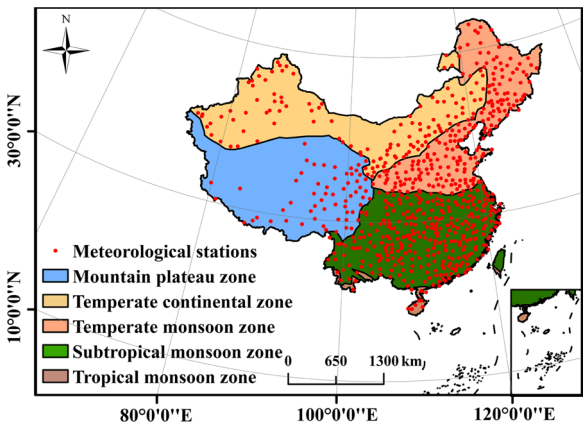

Fig. 1 Geographical distribution of meteorological stations in different climate zones of China

other meteorological variables, required in the PET calculation by the various methods, were estimated following the procedure described in FAO-56 (Allen et al. 1998).

The saturation vapor pressure and actual vapor pressure were calculated using the following equations:

$$e_s = 0.6108 \exp\left(\frac{17.27T}{T+237.3}\right) \quad (1)$$

$$e_a = (RH \times e_s)/100 \quad (2)$$

where e_s is the saturation vapor pressure (kPa); T is the air temperature (℃); e_a is the actual vapor pressure (kPa); RH is the relative humidity (%).

The radiations are given by:

$$R_a = \frac{24 \times 60}{\pi} \times G_{sc} \times d_r \times [\omega_s \times \sin(\varphi) \times \sin(\delta) + \cos(\varphi) \times \cos(\delta) \sin(\omega_s)] \quad (3)$$

$$R_s = \left(a_s + b_s \frac{n}{N}\right) R_a \quad (4)$$

$$R_{so} = (a_s + b_s) R_a \quad (5)$$

$$R_{ns} = (1-\alpha) R_s \quad (6)$$

$$R_{nl} = \sigma \left(\frac{T_{\max}^4 + T_{\min}^4}{2}\right) \times \left(0.34 - .014\sqrt{e_a}\right) \times \left(1.35 \frac{R_s}{R_{so}} - 0.35\right) \quad (7)$$

$$\mathrm{R_n} = \mathrm{R_{ns}} - \mathrm{R_{nl}} \quad (8)$$

where R_a is extraterrestrial radiation (MJ·m^{-2}·day^{-1}); R_s is solar or shortwave radiation; R_{so} is clear-sky solar radiation (MJ·m^{-2}·day^{-1}); R_{ns} is net solar radiation (MJ·m^{-2}·day^{-1}); R_{nl} is net longwave radiation (MJ·m^{-2}·day^{-1}); R_n is net radiation (MJ·m^{-2}·day^{-1}); G_{sc} is solar constant (0.0820 MJ·m^{-2}·min^{-1}); d_r is inverse relative distance Earth-Sun; ω_s is the sunset hour angle; δ is the solar declination; ϕ is the latitude; n is the actual duration of sunshine (hour); N is the maximum possible duration of sunshine or daylight (hour); α_s is the regression constant, expressing the fraction of extraterrestrial radiation reaching the earth on overcast days ($n=0$); $\alpha_s + b_s$ is the fraction of extraterrestrial radiation reaching the earth on sunny days ($n=N$), it is recommended that $\alpha_s = 0.25$, $b_s = 0.5$; α is the albedo (0.23); σ is the Stefan-Boltzmann constant (4.903×10^{-9} MJ·K^{-4}·m^{-2}·day^{-1}); $T_{\max}$ and $T_{\min}$ are the highest and lowest temperature within an hour (K), respectively.

Mann–Kendal trend test

We used the Mann–Kendal (MK) test to analyze the changing trend of PET calculated by 30 models from 2001 to 2020, the MK test is a non-parametric test method proposed by Mann and Kendall (Mann 1945; Kendall 1948). This test is currently recommended by the World Meteorological Organization (WMO) and has been widely used. The MK statistical test can be used for assessing linear and non-linear trends, with tolerance for outliers. Indeed, this test is suitable for hydrological data that does not meet the characteristics of normal distribution. The test statistic S in the MK test can be used to assess the hydrological or meteorological sequence trends (Zandt and Owens 1986) and can be calculated according to the following equations:

$$S = \sum_{k=1}^{n-1} \sum_{i=k+1}^{n} sign(x_i - x_k) \quad (9)$$

$$sign(xi - xk) = \begin{cases} +1, (x_i - x_k) > 0 \\ 0, (x_i - x_k) = 0 \\ -1, (x_i - x_k) < 0 \end{cases} \quad (10)$$

The significance of the trend is calculated using the test statistic Z, as follows:

$$Z = \begin{cases} \frac{S-1}{\sqrt{Var(S)}}, S > 0 \\ 0, S{=}0 \\ \frac{S+1}{\sqrt{Var(S)}}, S < 0 \end{cases} \quad (11)$$

where $Var(S)$ is the variance of S; $Z > 0$ and $Z < 0$ indicate increasing and decreasing trends, respectively. $|Z| \geq 1.96$ and $|Z| \geq 2.58$ are the critical values of the 95 and 99% significance levels, respectively.

Selected PET models

A total of 30 PET models were selected in this work, including 8 temperature-based methods, 7 aerodynamic methods, 11 radiation-based methods, and 4

Table 1 The 30 potential evapotranspiration (PET) models selected for this study

Category	No.	Methods	Equation	References
Temperature-based	1	Baier–Robertson	$PET = 0.157T_{\max} + 0.158(T_{\max} - T_{\min}) + 0.109R_a - 5.39$	Baier and Robertson (1965)
	2	Blaney–Criddle	$PET = kp(0.46T_a + 8.13)$	Blaney and Criddle (1950)
	3	Kharrufa	$PET = 0.34pT_a^{1.3}$	Kharrufa (1985)
	4	McCloud	$PET = 0.254 \times 1.07^{1.8T_a}$	Xiang et al. (2020)
	5	Oudin	$\begin{cases} PET=(\lambda\rho)^{-1}R_a((T_a+5)/100), T_a>-5°C \\ PET=0, T_a\leq -5°C \end{cases}$	Oudin et al. (2005)
	6	Romanenko	$PET = 4.5(1 + T_a/25)^2(1 - e_a/e_s)$	(Seiller and Anctil 2016)
	7	Schendel	$PET = (16T_a)/RH$	(Bormann 2011)
	8	Thornthwaite	$\begin{cases} PET = 0, T_a \leq 0°C \\ PET = 16\left(10T_a \Big/ \sum_{i=1}^{12}(T_a/5)^{1.514}\right)^A, T_a > 0°C \end{cases}$	Thornthwaite (1948)
Aerodynamic-based	9	Albrecht	$PET = (0.1005 + 0.297\mu_2)(e_s - e_a)$	Xiang et al. (2020)
	10	Brockamp–Wenner	$PET = 0.543\mu_2^{0.456}(e_s - e_a)$	Bormann (2011)
	11	Harbeck	$PET = 0.0578\mu_8(e_s - e_a) \times 25.4$	Singh and Xu (1997)
	12	Kuzmin	$PET = (1 + 0.21\mu_2)(e_s - e_a) \times 6$	Xiang et al. (2020)
	13	Mahringer	$PET = 2.86\mu_8^{0.5}(e_s - e_a)$	Mahringer (1970)
	14	Rohwer	$PET = 0.44(1 + 0.27\mu_2)(e_s - e_a)$	Zhao et al. (2013)
	15	Trabert	$PET = 0.3075\mu_8^{0.5}(e_s - e_a)$	Bormann (2011)
Radiation-based	16	Abtew	$PET = \alpha R_s/\lambda$	Zhao et al. (2013)
	17	Christiansen	$PET = 0.385R_s$	
	18	Doorenbos–Pruitt	$PET = \alpha\left(\frac{\Delta}{\Delta+\gamma}\frac{R_s}{\lambda}\right) + b$	Doorenbos and Pruitt (1975)
	19	Hargreaves	$PET = 0.0135(T_a + 17.8)\frac{R_s}{\lambda}$	Hargreaves (1975)
	20	Jensen–Haise	$PET = (0.014T_a - 0.37)(0.000673R_s) \times 25.4$	Zhao et al. (2013)
	21	Makkink	$PET = 0.61\frac{\Delta}{\Delta+\gamma}\frac{R_s}{\lambda} - 0.12$	Xu and Singh (2002)
	22	Milly–Dunne	$PET = 0.8\frac{(R_n-G)}{\lambda}$	Milly and Dunne (2016)
	23	Priestley–Taylor	$PET = 1.26\frac{\Delta}{\Delta+\gamma}\frac{(R_n-G)}{\lambda}$	Priestley and Taylor (1972)
	24	Stephens	$PET = (0.0158T_a + 0.09)R_s$	Stephens (1965)
	25	Stephens–Stewart	$PET = (0.0082T_a - 0.19)(R_s/1500) \times 25.4$	Zheng et al. (2017)
	26	Turc	$\begin{cases} PET=0.013(T_a/(T_a+15))\times(R_s+50)(1+(50-RH)/70), RH<50\% \\ PET=0.013(T_a/(T_a+15))\times(R_s+50), RH\geq 50\% \end{cases}$	Zhao et al. (2013)
Combination	27	Penman	$PET = \frac{\Delta(R_n-G)}{\lambda(\Delta+\gamma)} + \frac{\gamma}{\Delta+\gamma}\frac{6.43(1+0.536\mu_2)(e_s-e_a)}{\lambda}$	Yang et al. (2021)
	28	Penman–Monteith	$PET = \frac{0.408\Delta(R_n-G)+\gamma\frac{900}{T_a+273}\mu_2(e_s-e_a)}{\Delta+\gamma(1+0.34\mu_2)}$	(Zhao et al. 2013)
	29	Rijtema	$PET = \frac{\Delta(R_n-G)/\lambda+\gamma r\mu_2^{0.75}(e_s-e_a)}{(\Delta+\gamma)}$	Bormann (2011)
	30	Wright–Jensen	$PET = \frac{\Delta(R_n-G)}{\lambda(\Delta+\gamma)} + \frac{\gamma}{\Delta+\gamma}2.63(0.75 + 0.993\mu_2)(e_s - e_a)$	Allen (1986)

PET is the potential evapotranspiration (mm·day^{-1}); u_2 *and* u_8 are wind speed at 2 and 8 m height (m·s^{-1}), respectively; e_s *and* e_a are saturation vapor pressure and actual vapor pressure (kPa), respectively; T_a, $T_{\max}$ and $T_{\min}$ are average, maximum, and minimum daily air temperature (°C), respectively, °F unit for the Jensen–Haise and Stephens–Stewart equations; k is the monthly consumptive use coefficient; p is the percentage of total daytime hours for the period used (daily or monthly) out of total daytime hours of the year (365 × 12); RH is the relative humidity (%); A is a constant ($A = 6.75 \times 10^{-7}H^3 - 7.71 \times 10^{-5}H^2 + 1.792 \times 10^{-2}H + 0.49$); Δ is the slope of the vapor pressure curve (kPa·°C^{-1}); R_n, R_a, R_s are net, extraterrestrial, and incident solar radiation, respectively (MJ·m^{-2}·day^{-1}); G is the soil heat flux density (MJ·m^{-2}·day^{-1}), which can be neglected at a daily time step; r is the roughness coefficient; γ is a psychrometric constant (kPa·°C^{-1}); λ is the latent heat of vaporization (2.45 MJ·kg^{-1}); ρ is the water density (kg·m^{-3})

combination methods (Table 1). These models were selected not only because of the theoretical differences in their calculation formulas, but also because they represent the full range of key input meteorological variables (e.g., radiation, air temperature, relative humidity and wind speed) to be processed.

Evaluation criteria

Statistical indices were used for quantitative analysis of the PET modelling performance. The PET values calculated by the 30 models and were compared to Epan using a series of statistical criteria as follows:

$$R^2 = \left[\mathrm{cov}\left(PET, E_{pan}\right)/\sigma PET \sigma E_{pan}\right]^2 \tag{12}$$

$$MAE = \left(\sum_{i=1}^{n} \left|PET_i - E_{pan,i}\right|\right)/n \tag{13}$$

$$RMSE = \sqrt{\left(\sum_{i=1}^{n} \left|PET_i - E_{pani}\right|\right)^2/n} \tag{14}$$

$$NSE = 1 - \left(\sum_{i=1}^{n} \left|PET_i - E_{pani}\right|\right)^2 / \left(\sum_{i=1}^{n} \left|PET_i - \overline{E_{pani}}\right|\right)^2 \tag{15}$$

where R^2, MAE, RMSE and NSE are the coefficient of determination, mean absolute error (mm), root mean square error (mm), and Nash–Sutcliffe efficiency, respectively. n is the number of statistical days, and cov and σ are the covariance and standard deviation, respectively. The coefficient of determination R^2 was chosen as the representative evaluation index, and the other criteria included the NSE, MAE and RMSE were used as reference evaluation indices. The statistical analysis was conducted both for the overall time period and for each of the 12 months. All analyses were performed using R (4.1.0) software.

Model calibration

The best method for each climatic zone was modified to increase precision of estimating by calibration. The calibration method is similar to the study by Xu and Singh (2000), which changes the model that needs improvement to five forms with constants (Table 2), and then uses the PET benchmark (PETm) as a reference value to calibrate the model's constants.

$$\text{Romanenk}: \quad PET = a(1 + T_a/25)^2(1 - e_a/e_s) \tag{16}$$

$$\text{Rohwe}: \quad PET = a(1 + b\mu_2)(e_s - e_a) \tag{17}$$

$$\text{Abtew}: \quad PET = aR_s/\lambda \tag{18}$$

$$\text{Priestley} - \text{Taylor}: \quad PET = a\frac{\Delta}{\Delta + \gamma}\frac{(R_n - G)}{\lambda} \tag{19}$$

$$\text{Penman} - \text{Monteith}: \quad PET = \frac{a\Delta(R_n - G) + \gamma\frac{900}{T_a+273}\mu_2(e_s - e_a)}{\Delta + \gamma(1 + b\mu_2)} \tag{20}$$

a and b are variables to be estimated.

Results and discussion

Spatial and temporal variation of potential evapotranspiration in China calculated using 30 models

In total, 30 models were used to calculate the average PET in different climatic zones, using the interpolated raster data from the 699 weather stations observed over the 2011–2020 period, followed by MK test analysis to assess the PET trends. The results showed significant variations in the average PET between the PET models. In addition, considerable differences in temporal and spatial changes of the PET were observed between the climatic zones of China. The maximum PET values in all climatic zones of China were obtained using Doorenbos–Pruitt's model and were 1959.8 mm (MPZ), 1965.2 mm (TCZ), 1860.4 mm (SMZ), and 1825.2 mm (TMZ), while the lowest PET values were obtained using the Kuzmin model (375.8 and 339.4 mm in TCZ and SMZ, respectively), the Trabert model (304.1 mm in the MPZ), and the Turc model (317.4 mm in the TMZ). By considering the entire China area, most models revealed a PET range of 600–1200 mm. Kuzmin and Doorenbots–Pruitt models showed the minimum and maximum PET values of 349.6 and 1902.6 mm, respectively, indicating a difference in PET values between different models of up to 5.5 times. Based on the PET values obtained, the order of models was as follows: Doorenbots–Pruitts > Chirstiansen > Makkink > Abtew > Brockamp–Wenner > Albrecht > Romanenko > Milly–Dunne > Penman–Monteith > Hargreaves > Rohwer > Priestley–Taylor > Rijtema > Wright–Jensen > Penman > Schendel > Oudin > Baier–Robertson > Mahringer > Thornthwaite > Harbeck > McCloud > Jensen–Haise > Turc > Stephens–Stewart > Blaney–Criddle > Trabert > Kharrufa > Kuzmin (Fig. 2).

Table 2 Evaluation criteria between the original and calibrated potential evapotranspiration models against potential evapotranspiration reference in different climatic zones of China

Models	Climatic zones	Coefficients	*a*	*b*	R^2	MAE	RMSE	NSE
Romanenko	MPZ	Original	4.50	–	0.81	14.83	19.48	0.65
		Calibrated	3.90	–	0.81	12.29	18.62	0.68
	SMZ	Original	4.50	–	0.83	11.42	15.00	0.81
		Calibrated	4.30	–	0.83	10.90	14.18	0.98
	TCZ	Original	4.50	–	0.96	28.47	39.54	0.63
		Calibrated	3.50	–	0.96	11.37	15.16	0.95
	TMZ	Original	4.50	–	0.90	13.95	17.21	0.85
		Calibrated	4.20	–	0.90	12.40	15.25	0.88
Rohwer	MPZ	Original	0.44	0.27	0.80	23.15	26.76	0.33
		Calibrated	0.50	0.08	0.82	12.80	17.07	0.73
	SMZ	Original	0.44	0.27	0.88	19.14	21.79	0.59
		Calibrated	0.51	0.01	0.89	10.01	13.75	0.84
	TCZ	Original	0.44	0.27	0.95	15.06	25.39	0.85
		Calibrated	0.38	0.01	0.96	12.06	18.55	0.92
	TMZ	Original	0.44	0.27	0.91	13.30	17.44	0.85
		Calibrated	0.46	0.03	0.92	10.01	13.57	0.91
Abtew	MPZ	Original	0.53	–	0.81	35.12	37.18	0.29
		Calibrated	0.41	–	0.81	15.51	18.58	0.68
	SMZ	Original	0.53	–	0.93	11.32	13.65	0.84
		Calibrated	0.51	–	0.93	9.66	11.91	0.88
	TCZ	Original	0.53	–	0.91	24.28	29.17	0.80
		Calibrated	0.56	–	0.91	20.74	25.05	0.82
	TMZ	Original	0.53	–	0.92	19.65	22.57	0.74
		Calibrated	0.50	–	0.82	16.37	19.43	0.81
Priestley–Taylor	MPZ	Original	1.26	–	0.84	15.58	18.44	0.68
		Calibrated	1.32	–	0.84	14.75	18.05	0.70
	SMZ	Original	1.26	–	0.92	14.11	17.64	0.73
		Calibrated	1.18	–	0.92	11.25	13.86	0.83
	TCZ	Original	1.26	–	0.92	25.93	30.88	0.77
		Calibrated	1.58	–	0.92	18.40	22.24	0.88
	TMZ	Original	1.26	–	0.92	16.57	18.62	0.83
		Calibrated	1.23	–	0.92	16.14	17.42	0.84
Penman–Monteith	MPZ	Original	0.408	0.340	0.95	7.41	8.80	0.93
		Calibrated	0.410	0.368	0.96	6.913	8.667	0.94
	SMZ	Original	0.408	0.340	0.97	5.87	7.10	0.96
		Calibrated	0.390	0.331	0.98	5.01	6.58	0.96
	TCZ	Original	0.408	0.340	0.98	14.71	17.83	0.92
		Calibrated	0.456	0.320	0.98	6.58	9.01	0.98
	TMZ	Original	0.408	0.340	0.98	8.37	9.22	0.96
		Calibrated	0.405	0.304	0.98	4.59	6.68	0.98

The PET trends determined using the MK test over the 2001–2020 period for 30 models showed high spatial heterogeneity. Most models revealed a significant upward trend in PET in MPZ, while few radiation-based models showed a downward trend in the MPZ (Fig. 3). The models used showed considerable spatial differences in the PET trends across the climatic zones of China, suggesting that model evaluation is needed.

Comprehensive comparison of temperature-based models

In this study, the monthly mean PET values across China calculated by the temperature-based models, namely

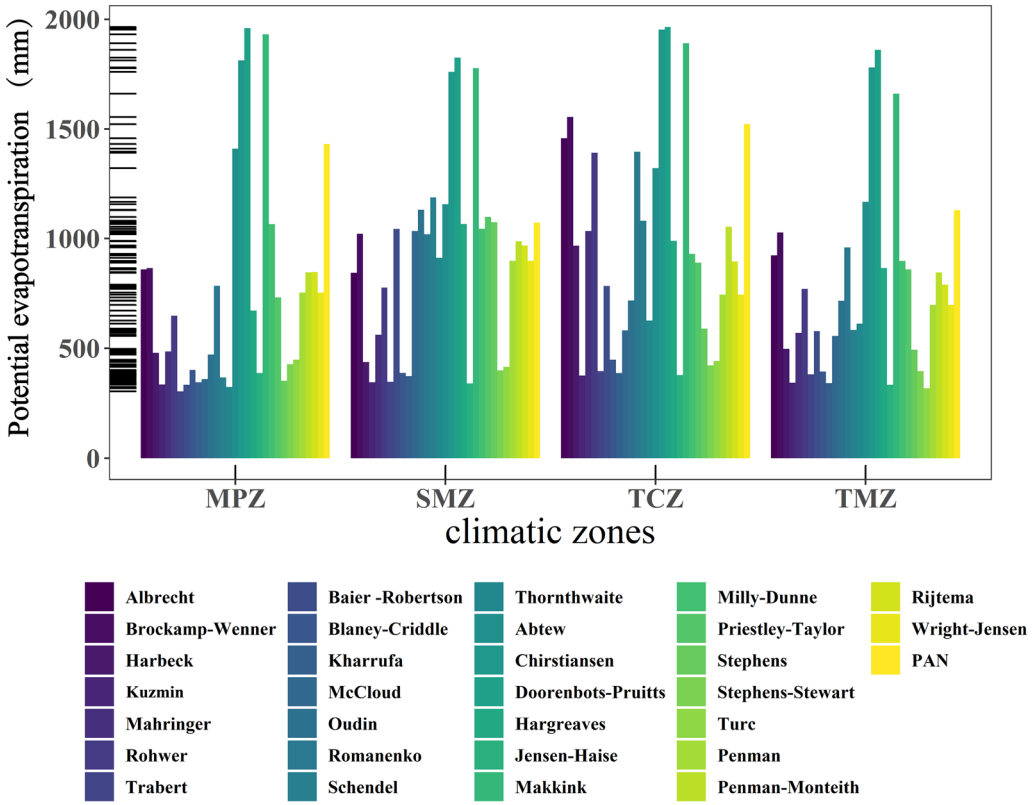

Fig. 2 Area-averaged annual potential evapotranspiration (2001–2020) calculated by 30 models and measured Epan evaporation in different climatic zones of China

Baier–Robertson, Blaney–Criddle, Kharrufa, McCloud, Oudin, Romanenko, Schendel, and Thornthwaite models were 83.71, 42.99, 38.85, 69.04, 88.88, 95.05, 91.04, and 70.34 mm, respectively, while the monthly Epan value was 100.76 mm. It should be noted that, except for the Baier–Robertson model, all other PET models revealed higher R^2 values with the Epan in the SMZ than those observed in other climatic zones (Fig. 4), suggesting better applicability of temperature-based models in humid regions.

On the other hand, the monthly PET rates of the 8 temperature-based models revealed different monthly and spatial trends in the R^2 values with Epan (Fig. 5). Indeed, significant R^2 values were obtained using the Oudin, Thornthwaite, Kharrufa, and McCloud models, it was smaller than the other models, and the variation in R^2 between months was irregular, indicating that these three models may not be suitable for the calculation of PET in China. The R^2 obtained using the Baier–Robertson, Romanenko, and Schendel models revealed similar trends. Moreover, the monthly PET and Epan observed in warm months showed higher R^2 than those observed in cold months, indicating that these three models are more appropriate for calculating the PET in warm seasons in China. Indeed, the results suggested that the Blaney–Criddle model is more suitable for calculating the PET in cold seasons in China. The daily average temperature is an important input variable for these models, and seasonal temperature differences may explain the variation in R^2 between months.

The monthly R^2 of the Romanenko model was higher without considerable variations between months as compared to those of other models (maximum R^2 of 0.90 in the TMZ). By considering the four climatic zones, the results showed a higher R^2 value of the Romanenko model than those of other models, with a maximum R^2 value observed in SMZ. Therefore, the results suggested that the Romanenko model may be the most suitable temperature-based model, especially in the humid regions of China. Zuo et al. (2009) used a temperature-based model to simulate PET in arid regions of Northwest China and found a high correlation coefficient ($r=0.97$) of the corrected Romanenko model, with a conversion factor value with the data measured by a small evaporating pan of 0.6, significantly better than those of other models. However, some studies may reach different conclusions due

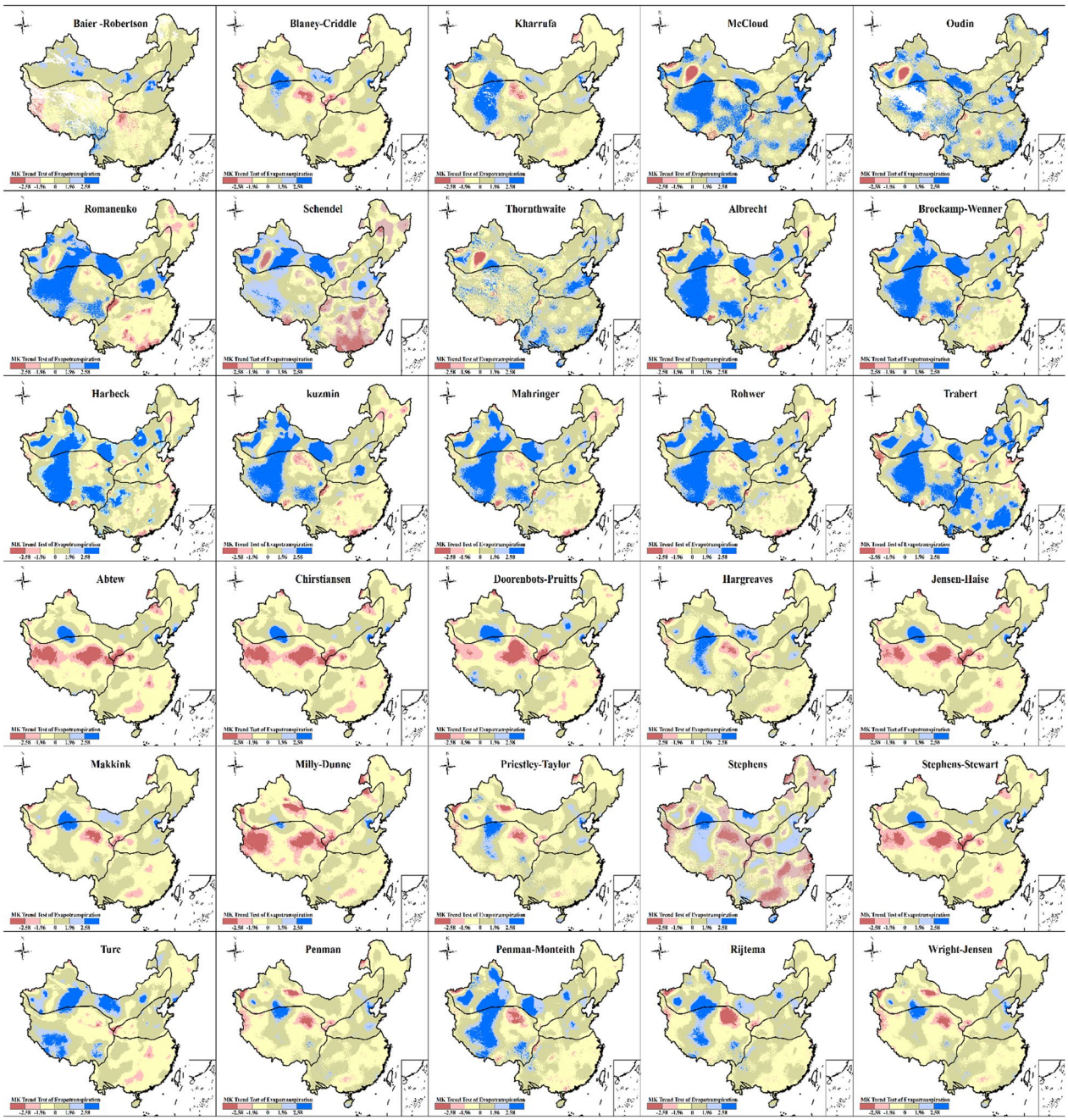

Fig. 3 MK trend results of average annual PET (2001–2020) calculated by different models. The blue and red colors indicate significant upward and downward trends in PET, respectively

to different methods and time scales used in evapotranspiration verification (Xu and Singh 2002; Lu et al. 2005). For example, Xu et al. (2002) compared five temperature-based and radiation-based methods in the Switzerland and recommended the Blaney–Criddle model as the best temperature-based method, which performed relatively poorly in this study.

Comprehensive comparison of aerodynamic-based models

The monthly average PET values across China calculated based on the aerodynamic Albrecht, Brockamp–Wenner, Harbeck, Kuzmin, Mahringer, Rohwer, and Trabert models were 85.24, 90.1, 44.87, 29.72, 54.34, 73.32, and 30.79 mm, respectively. The Rohwer model showed a higher R^2 value (0.73) than those of other models in all

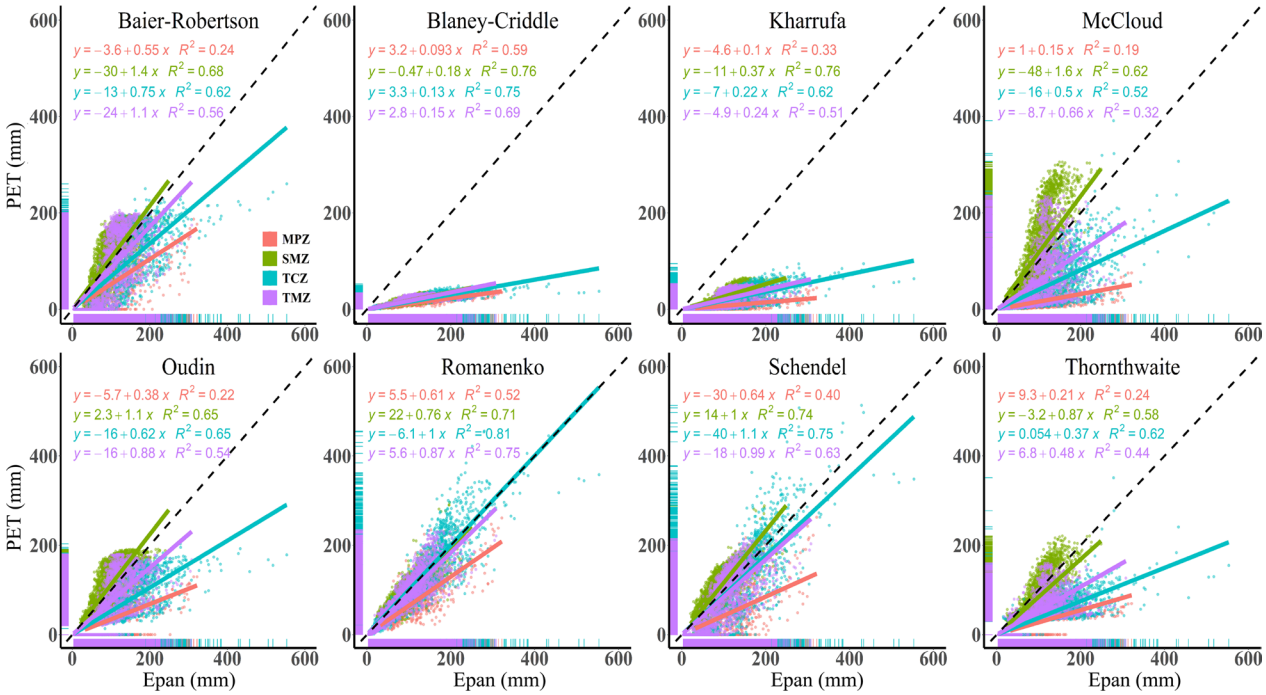

Fig. 4 Comparison of calculated and measured monthly PET (2001–2020) determined by 8 temperature-based models and Epan, respectively. Orange, green, blue, and purple line colors correspond to trend lines and PET scattered points in MPZ, SMZ, TCZ, and TMZ, respectively; the black line is the 1:1 line

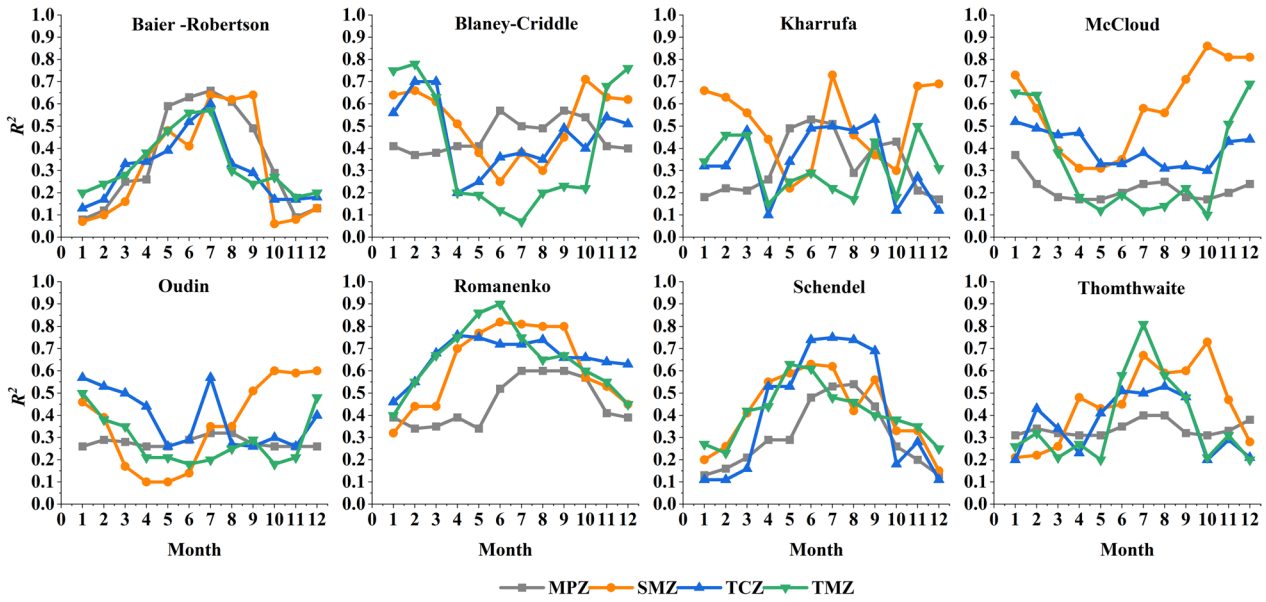

Fig. 5 Monthly variation of the mean ± standard deviation of R^2 (coefficient of determination) for 8 temperature-based models in the four climatic zones of China

climatic zones (Fig. 6). In addition, except for the Harbeck and Mahringer models, the R^2 values of aerodynamic-based models observed in TCZ were higher than those observed in other regions. On the other hand, the monthly PET of the aerodynamic models showed similar variations (Fig. 7), with high R^2 values in winter and

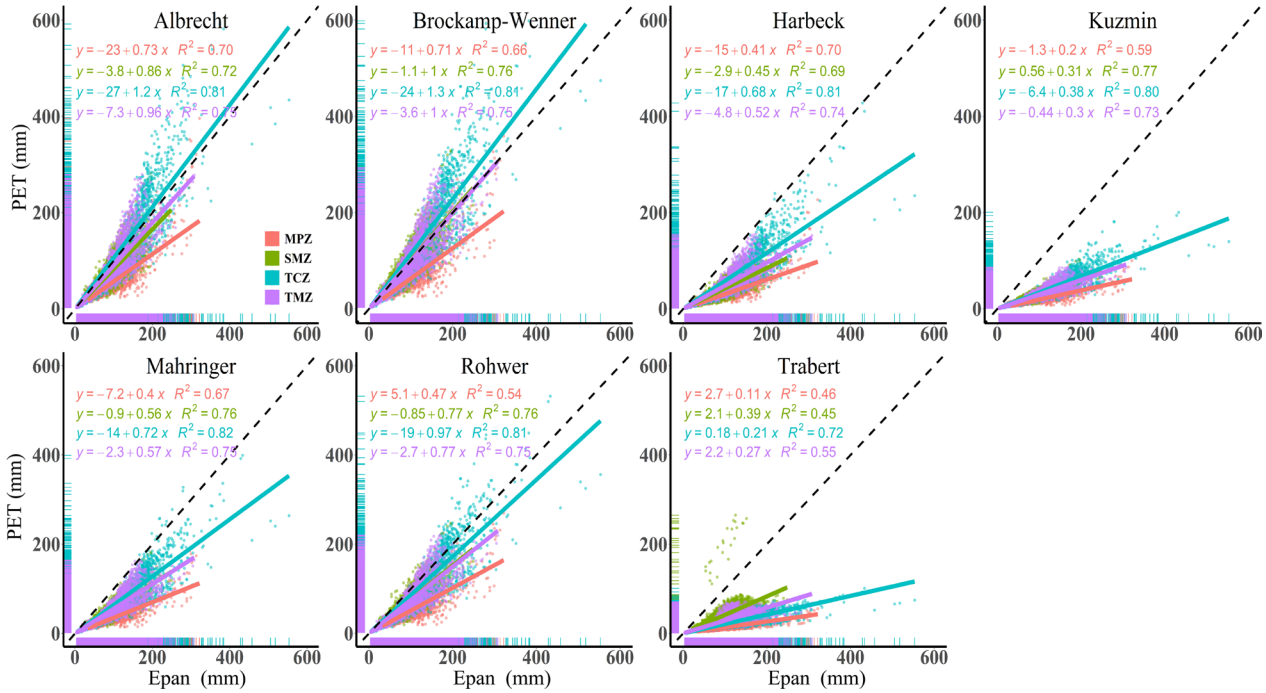

Fig. 6 Comparison of calculated and measured monthly PET (2001–2020) determined by 7 aerodynamic-based models and Epan, respectively

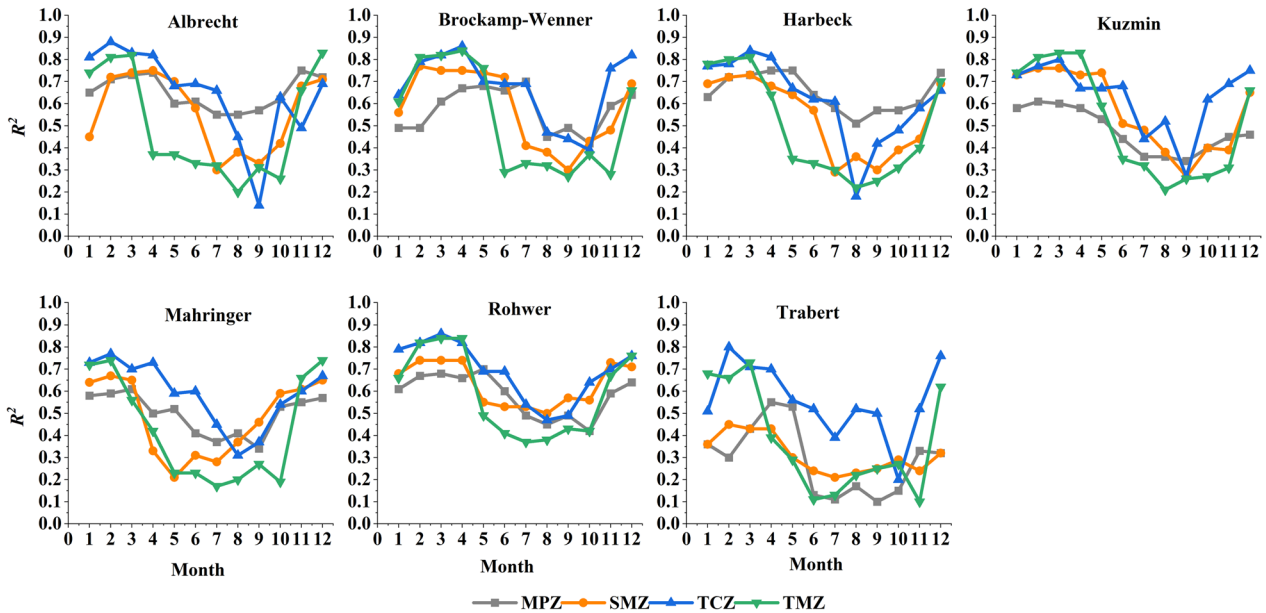

Fig. 7 Monthly variation of the mean ± standard deviation of R^2 (coefficient of determination) for 7 aerodynamic-based models in the four climatic zones of China

spring and low R^2 values in summer and autumn. Wind speed is an important input variable. In China, the wind speed is high in the spring season, with maximum and minimum values observed in April and August, respectively. The seasonal variation of the mean R^2 values may be due to the seasonal variation of wind speed. The results showed low monthly R^2 values of the Harbeck, Mahringer, and Trabert models (Fig. 7), the wind speed at 8 m, considered an input variable in these three models, is calculated using the wind speed profile relationship in

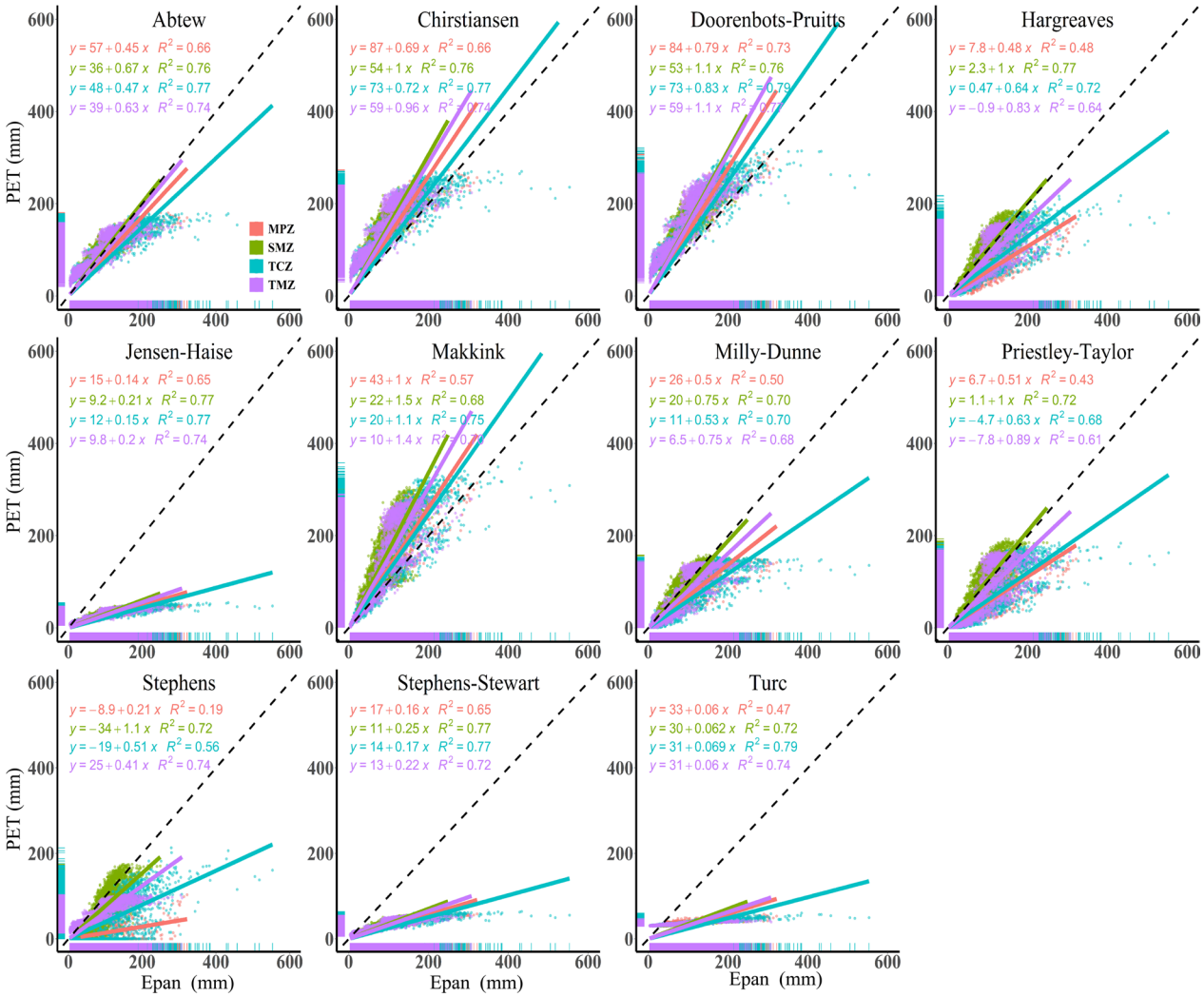

Fig. 8 Comparison of calculated and measured monthly PET (2001–2020) determined by 11 radiation-based models and Epan, respectively

FAO-56, which may lead to a certain calculation error. By taking all factors into consideration, it can be concluded that the Rohwer model is the most suitable aerodynamic model for PET calculations in China.

Comprehensive comparison of radiation-based models

The monthly average PET values across China calculated by the Abtew, Christiansen, Doorenbots–Pruitts, Hargreaves, Jensen–Haise, Makkink, Milly–Dunne, Priestley–Taylor, Stephens, Stephens–Stewart, and Truc models were 100.30, 151.73, 159.52, 83.38, 29.02, 153.93, 82.68, 82.96, 56.42, 34.39, and 36.97 mm. In addition, the Abtew and the Stephens–Stewart models revealed the highest R^2 values in TCZ, while the Priestley–Taylor model showed the highest R^2 values in SMZ, TCZ, and TMZ (Fig. 8). On the other hand, from a monthly perspective, the 11 radiation-based models showed similar seasonal changes, that is, showing an increase in spring, with maximum and minimum values in summer and winter, respectively. The seasonal changes in temperature and radiation explain the seasonal changes of R^2 values since these parameters are the principal input variables in radiation-based models (Fig. 9). Compared with models based on temperature, aerodynamic, and combination, none of the 11 radiation-based models outperformed the others (Fig. 9). Therefore, the following monthly comprehensive comparison was conducted (Fig. 10) to determine which model performed best in which month. The formula used by Abtew to calculate PET is similar to the Christiansen, Doorenbos–Pruitt, and Hargreens (Table 1), with the main influencing parameters being incident solar radiation and psychrometric constant, and Abtew was more suitable than the others. The selection of Stephens–Stewart model and Stephens model is also

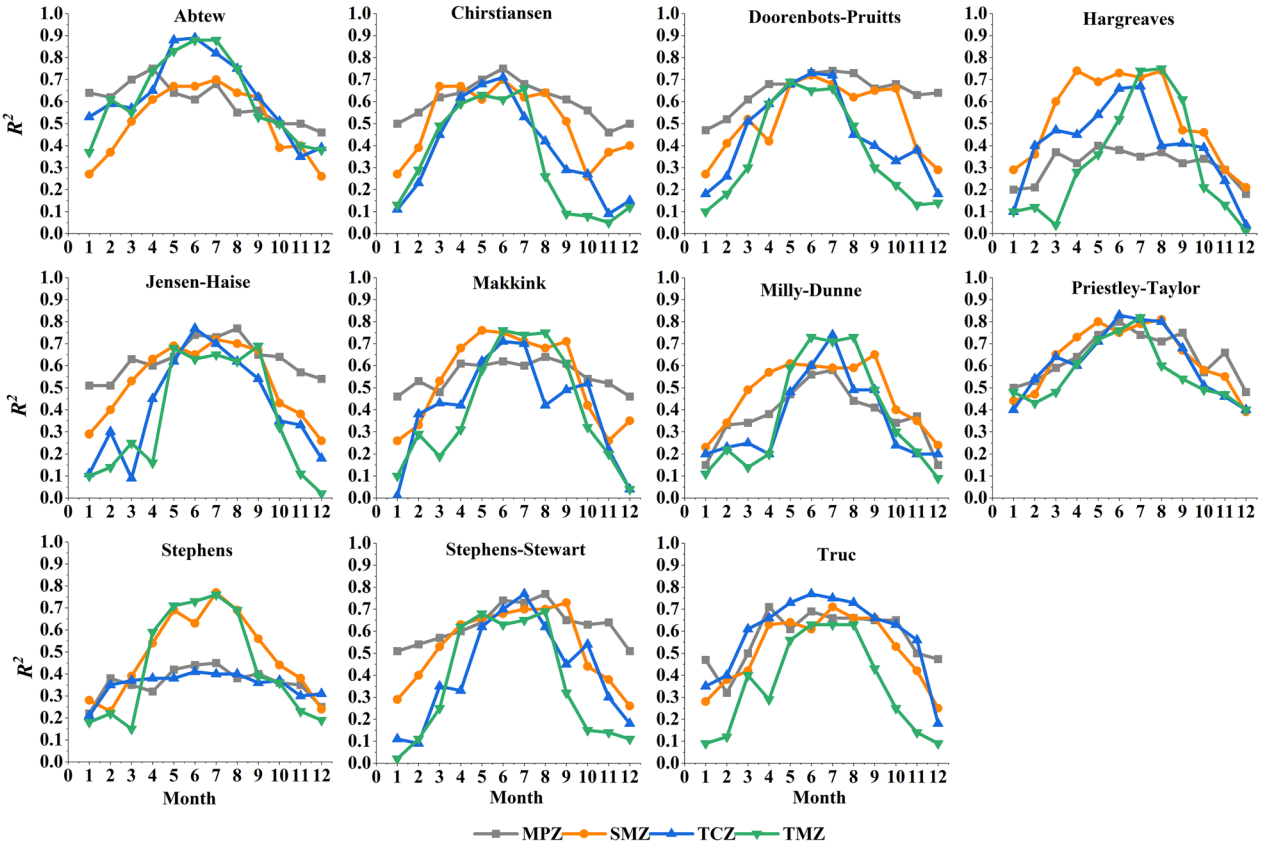

Fig. 9 Monthly variation of the mean ± standard deviation of R^2 (coefficient of determination) for the 11 radiation-based models in the four climatic zones of China

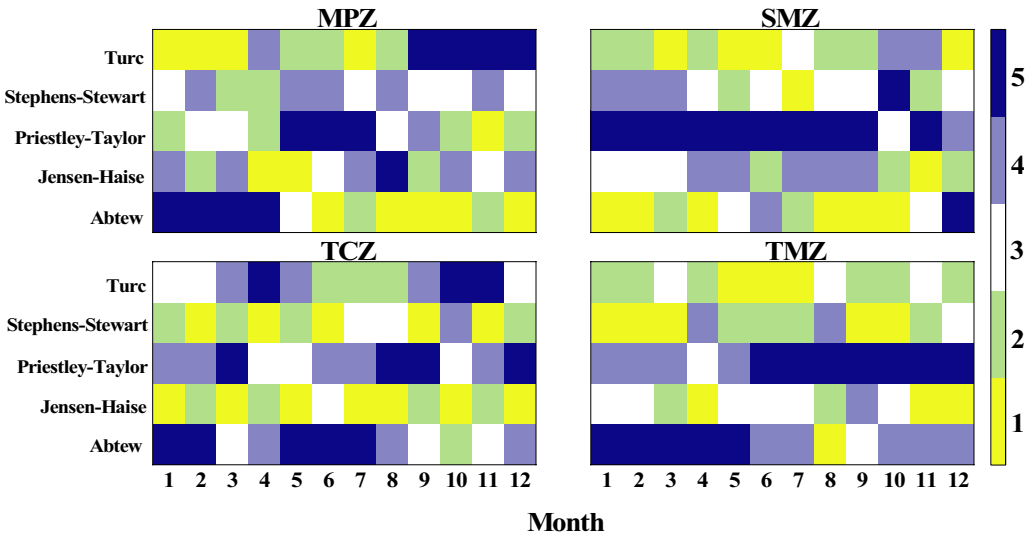

Fig. 10 Ranking of the mean of R^2 (coefficient of determination) over the 12 months for five radiation-based methods in four climatic zones of China

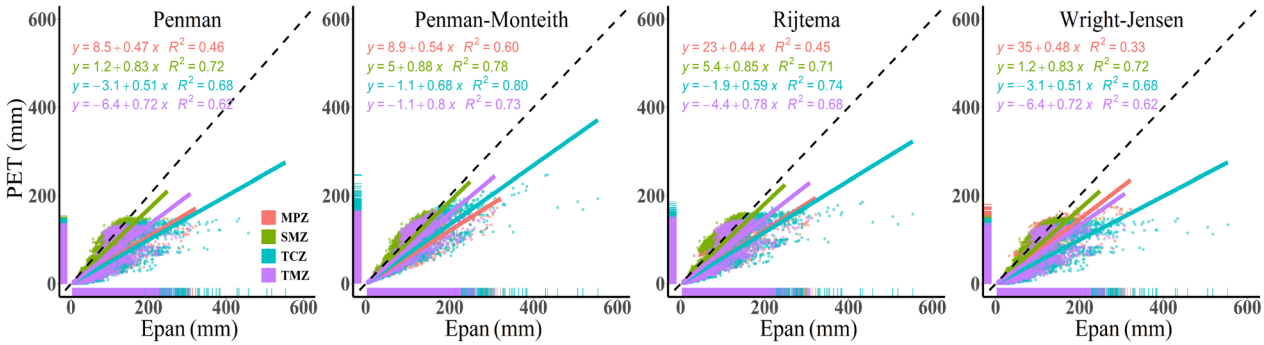

Fig. 11 Comparison of calculated and measured monthly PET (2001–2020) determined by the 4 combination-based models and Epan, respectively

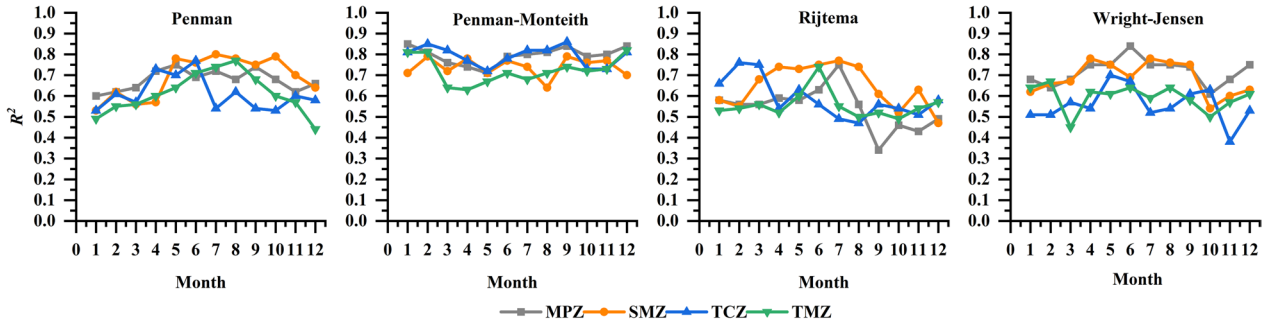

Fig. 12 Monthly variation of the mean ± standard deviation of R^2 (coefficient of determination) for the 4 combination-based models in the four climatic zones of China

for the same reason and the Milly–Dunne can be seen as a simplified version of the Priestley–Taylor (Table 1). Therefore, the Turc, Stephens–Stewart, Priestley–Taylor, Jensen–Haise, and Abtew models were selected for the following monthly comparisons (Fig. 10).

In MPZ, Abtew, Priestley–Taylor, and Jensen–Haise models revealed the highest R^2 values in January–April, May–July, August, respectively, while the Turc model showed the highest values in the remaining months. In SMZ, Stephens–Stewart and Abtew models showed the highest R^2 values in October and December, respectively, while the highest values of the Priestley–Taylor model were observed in the remaining months. The highest R^2 values observed in TCZ were revealed by the Turc model in April, October, and November, the Priestley–Taylor model in August and September, and the Abtew model in the remaining months, whereas in TMZ, the highest values were revealed by the Abtew model in the January–May period and December and the Priestley–Taylor model in the remaining months (Fig. 10). By comparing the calculated and measured PET values, we observed an underestimation and overestimation of the highest and lowest PET values, respectively, in the Stephens–Stewart and Truc model results (Fig. 8), which is consistent with the findings of Douglas et al. (2009) in Florida.

Although the results showed a lack of significant difference in the PET calculation accuracy of the 11 radiation-based models, it can be highlighted that the Abtew model is more suitable for arid regions (TCZ), while the Priestley–Taylor model is more suitable for humid and semi-humid regions (SMZ, TMZ). However, both PET models should be used with caution in colder months. Unlike the temperature- and aerodynamic-based models, radiation-based models have not shown a high calculation accuracy of any particular model. However, numerous studies have recommended the Abtew and Priestley–Taylor models, among radiation-based models, for calculating PET (Douglas et al. 2009; Valipour 2015), while other studies have reached different conclusions about the calculation accuracy of radiation-based models due to several reasons, including model selection, study area characteristics, year, variables used in the validation step, evaluation criteria, and evaluation scales (Xu and Singh 2002; Zuo et al. 2009; Bormann 2011; Yang et al. 2021).

Comprehensive comparison of combination-based models

The mean monthly PET values across China calculated by Penman, Penman–Monteith, Rijtema, and Wright–Jensen models were 68.22, 80.59, 76.17, and 70.72 mm, respectively. The calculation results of the Penman and

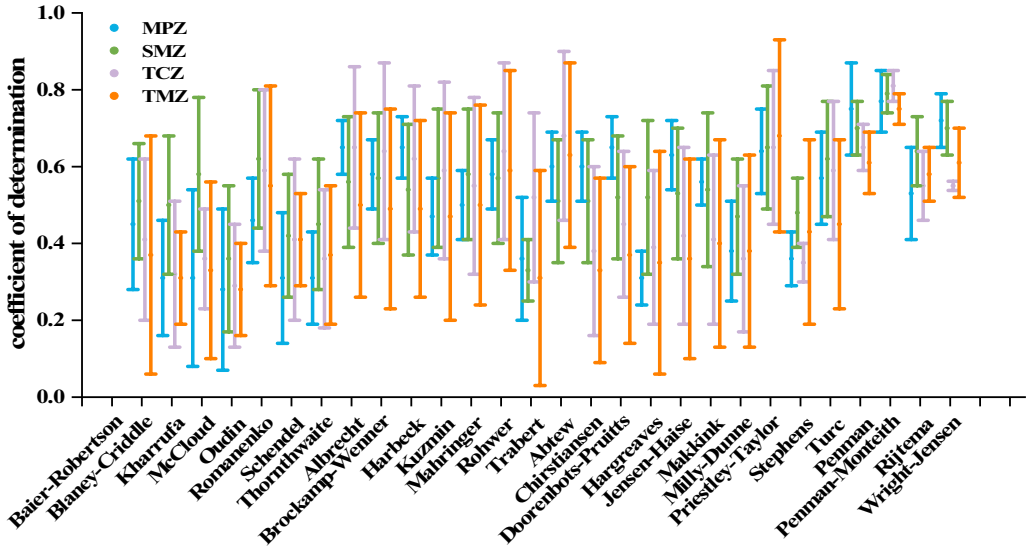

Fig. 13 Mean ± standard deviation of coefficient of determination for the 18 models in the four climatic zones of China. *MPZ* mountain plateau zone, *SMZ* subtropical monsoon zone, *TMZ* temperate monsoon zone, *TCZ* temperate continental zone

Penman–Monteith models were better than those of the Rijtema and Wright–Jensen models in entire China (Figs. 11, 12).

Unlike the temperature-, radiation-, and aerodynamic-based models, the trend change of the monthly average R^2 obtained using combination-based models did not show considerable seasonal changes, which may be due to the higher number of input meteorological variables as compared to other model category. However, the mean R^2 values of the Penman and Wright–Jensen models were slightly lower in cold months than in warm months (Fig. 12). The Rijtema model did not show a higher R^2 than those of other models for all months and climatic zones. In TCZ, the Penman–Monteith model showed higher R^2 values in all months and more than 6 months in other climatic zones of China. The Penman model revealed the highest R^2 values in July, August, and October in SMZ and in July and August in TMZ, while the Wright–Jensen model showed the highest R^2 in the March–June period in MPZ and in May in SMZ (Fig. 12).

The Penman–Monteith model is an improved version of the Penman model. Indeed, the results of the current study showed that the calculation accuracy of the Penman–Monteith model was higher than those of the Penman, Wright–Jensen, and Rijtema models. The latter revealed the lowest calculation accuracy. Many studies have indicated that the Penman–Monteith model has good accuracy in the PET calculation and even serves as a reference model for MODIS products (Chen and Liu 2020; Cheng et al. 2021).

Comprehensive comparison of the four PET model categories

The ranking results of the mean R^2 values demonstrated high calculation accuracies of the combination model category in the climatic zones of China, followed respectively by the radiation-, aerodynamic-, and temperature-based models. The average R^2 values of the combined model category were relatively close (between 0.62 and 0.88) compared with those of other model categories, indicating that this type of model has a relatively high level of stability. In addition, the stability of the model is also reflected in two other aspects. One is that the R^2 of the combination model has a small standard deviation, which is significantly lower than the standards of other types. Secondly, the R^2 of the combination model is also very close in different climatic zone (Fig. 13). During the overall time period, the four climatic zones ranked as follows in terms of the mean of R^2 between Epan and the PET calculated by the combination and aerodynamic methods: TCZ > TMZ > SMZ > MPZ. The corresponding rankings for the radiation-based were TCZ > SMZ > TMZ > MPZ and for the temperature-based methods were TCZ > TMZ > MPZ > SMZ.

PET is driven primarily by four key meteorological variables, including radiation, wind speed, temperature and vapor pressure (Donohue et al. 2010). The combination methods considered all four influencing factors. The results of the current study demonstrated that the calculation accuracies of the combined model category were higher than those of the other model categories,

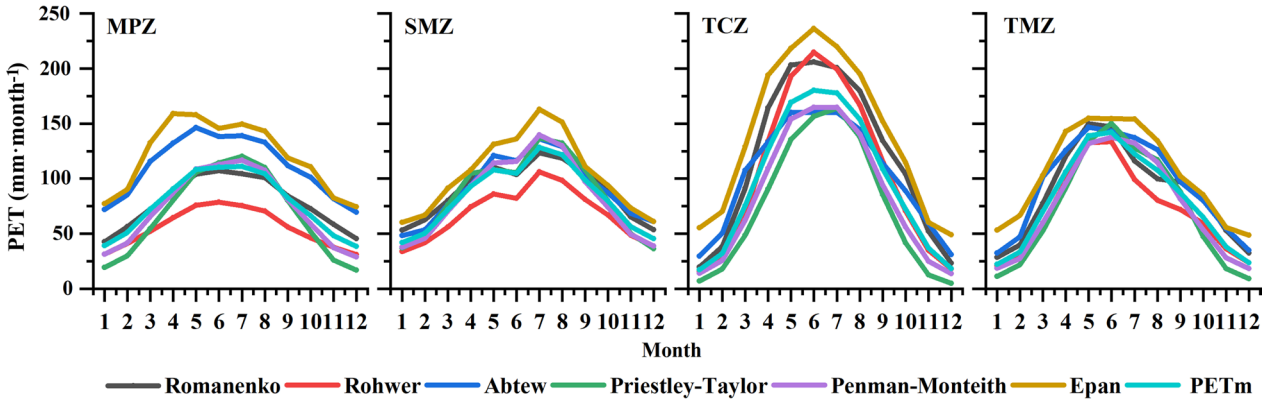

Fig. 14 Monthly variation of the calculated potential evapotranspiration by five PET models, PETm, and PET of Epan in the four climatic zones of China

followed by radiation-based models, which is consistent with results reported in several studies (Xu and Singh 2005; Donohue et al. 2010; Lakatos et al. 2020; Yang et al. 2021). This result may be due to the ignorance of one or multiple variable types in the PET calculation. In fact, the temperature- and radiation-based methods ignore the aerodynamic variables, while the aerodynamic-based models ignore the energy term (Shiri 2018; Shiri et al. 2019). In this study, the radiation based model outperformed the aerodynamic- and temperature based models, while the temperature-based model revealed the poorest calculation accuracies. Some authors showed similar results to those found in the current study (Xu and Singh 2001; Zuo et al. 2009; Valipour 2015; Yang et al. 2021), while others like Xu and Singh (2002) revealed inconsistent results, which may be due to the selection of the best model from each category for the cross-comparison and the difference between the chosen comparison method and the variables used for validation. In this study, except for the combined model category, the mean R^2 values of the other three model categories were not considerably different in different climatic zones. The differences in the mean R^2 values between PET models were found particularly within the same category of models. Therefore, the selection of different models of the same category for PET calculation may result in a biased evaluation of models, thus, further comparison of the best-selected models in each group is required to provide conclusive results, as reported in the next section.

Comprehensive comparison of the Romanenko, Rohwer, Abtew, Priestley–Taylor, and Penman–Monteith models

Potential evapotranspiration is a theoretical value that is difficult to quantify accurately, and thus comparing different model categories overcomes, to some extent, the limitations of certain model categories (Diks and Vrugt 2010; Chen and Liu 2020). According to the results presented in previous sections, the Romanenko temperature-based model, the Rohwer aerodynamic-based model, the Abtew radiation-based model, and the Priestley–Taylor and Penman–Monteith combined models were selected as the five optimal models for comparison. In the previous sections, we analyzed the correlations between the PET calculated by the models and the measured Epan, the actual values of the calculated PET presented in the subsequent part of this section.

The equally weighted mean values of the five selected models were taken as the PET benchmark (PETm) (Fig. 14). By considering the four climate zones, the PET values of the Romanenko model were higher than PETm in MPZ and TCZ and lower than PETm in SMZ and TMZ. The PET values of the Rohwer and Priestley–Taylor models were lower than PETm in MPZ, SMZ, and TMZ, and higher than PETm in TCZ. In addition, the PET values of the Penman–Monteith model were very close to PETm in all months, confirming the accuracy and applicability of the Penman–Monteith model in four climatic zones. It should be noted that the net radiation (clear-sky extraterrestrial radiation) was calculated in this study according to the FAO-56 method (As + bs, where As = 0.25, bs = 0.5) (Allen et al. 1998). Indeed, the MPZ is characterized by higher radiation levels as compared to other climatic zones in China due to the existence of high-altitude areas (e.g., the Qinghai-Tibet Plateau) (Chen et al. 2004; Zhang et al. 2010). In addition, the monthly average PET value calculated by the Abtew model was significantly higher than the PETm, suggesting an overestimation of the PET in alpine regions in actual operation by this model. While different model categories can be expected to exhibit different sensitivities due to the difference in the climatic variables used, the different sensitivities of PET models within the same category

obtained, considering the same climate variables, suggest that PET needs to be selected and calibrated carefully in PET studies in a climate change context.

Calibration of the Romanenko, Rohwer, Abtew, Priestley–Taylor, and Penman–Monteith models

The variables (Table 2) of Romanenko, Rohwer, Abtew, Priestley–Taylor, and Penman–Monteith models can be calibrated according to the four climatic zones if the mean PET value (PETm) of the five models is used as a benchmark. Table 2 shows the original and calibrated coefficients of the above five models in four climate zones. The PET values of models were closer to PETm values after the empirical coefficient correction (Fig. 15, Table 2), indicating higher PET accuracy (higher R^2 and NSE, lower MAE and RMSE). The Romanenko model, as the optimal temperature model, has significantly larger coefficient adjustments in the TCZ and MPZ than in the SMZ and TMZ, suggesting that the Romanenko model is not well suited for high latitudes and more arid regions. In addition, the adjustment of the empirical coefficients of the Rohwer model revealed the highest improvement of the PET values (Table 2 and Fig. 15), aerodynamic models were developed according to specific characteristics of regions. Indeed, the variation in the climate and underlying surface conditions may result in erroneous results. Thus, this model category can be applied in areas with the same conditions as the regions in which they were developed (Shiri 2018; Yang et al. 2021). Many studies have confirmed that the Abtew and Priestley–Taylor models are relatively accurate radiative category models (Singh and Xu 1997; Valipour 2015), and this viewpoint is also reflected in our study. Through the correction results, Abtew model has the largest adjustment amplitude in MPZ, while Priestley–Taylor model has the largest adjustment amplitude in TCZ, indicating that these two models have differences in their applicable spatial ranges, In the future, these two models can be selected according to the specific climate zone. The combination methods considered all meteorological factors and the original Penman model could calculate PET well, so the calibrated Penman model did not progress as much as other models. From the four climate zones, the adjustment amplitude of the five models in SMZ and TMZ is smaller than that in MPZ and TCZ, indicating that these humid areas have higher calculation accuracy in calculating PET through the models.

Integrated discussion

The PET for different climate zones in China over the past 20 years, calculated using 30 different types of models, exhibit significant spatial heterogeneity in both numerical values and trends. This indicates the necessity of assessing the applicability of different models in various climatic regions. Combination-based models, which comprehensively consider temperature, wind speed, radiation, and vapor pressure, demonstrate superior simulation performance compared to other types of models. This has been corroborated by numerous studies (Xu and Singh 2001; Zhou et al. 2020; Yang et al. 2021). Zuo et al. (2009) simulated PET in the arid region of Northwest China using a temperature-based model. They found that the corrected Romanenko model achieved a correlation coefficient of 0.97, with a conversion factor of 0.6 when compared to data from small evaporating pans. This performance was significantly superior to other temperature-based models, consistent with our research findings. In China, the spring season experiences higher wind speeds, peaking in April, while wind speeds reach their lowest point in August. As wind speed is a crucial input variable for aerodynamic models, the R^2 performance of aerodynamic models exhibits similar seasonal variations to wind speed. The Rohwer model is the only aerodynamic model we recommend based on our findings. Singh and Xu (1997) evaluated the calculation accuracy of 13 aerodynamic-based models at four climatic stations in Northwestern Ontario in Canada and revealed the highest accuracy of the Rohwer model for calculating PET, which is consistent with the results of this study. Radiation-based models did not show a clear preference for one specific model over others. This is in contrast to the temperature and aerodynamic models mentioned earlier, where a specific model demonstrated superior performance. Radiation-based models do not estimate accurately the effect of wind speed increasing on air resistance, thus resulting in underestimation of the PET. Indeed, the application of this model category is recommended in areas with low wind speed (Irmak and Irmak 2008; Xiang et al. 2020), In our study, based on climate zones, the Abtew model appears to be more suitable for arid and semi-arid regions (TCZ, MPZ), while the Priestley–Taylor model is better suited for relatively humid areas (SMZ, TMZ). However, caution is advised when using either model during cold months.

Calibrating the optimal model within each type significantly enhances the precision of PET simulations. Based on the calibration results, aerodynamic models require the most substantial adjustments. The calibrated Rohwer model performs exceptionally well in regions with higher wind speeds (TCZ), with an R^2 exceeding 0.9—the only region where this model achieves such high accuracy. Although both Abtew and Priestley–Taylor radiation models exhibit good performance in simulating PET, there are differences in their applicable spatial ranges. In the future, the choice between these two models may depend on specific climate conditions. The PM model,

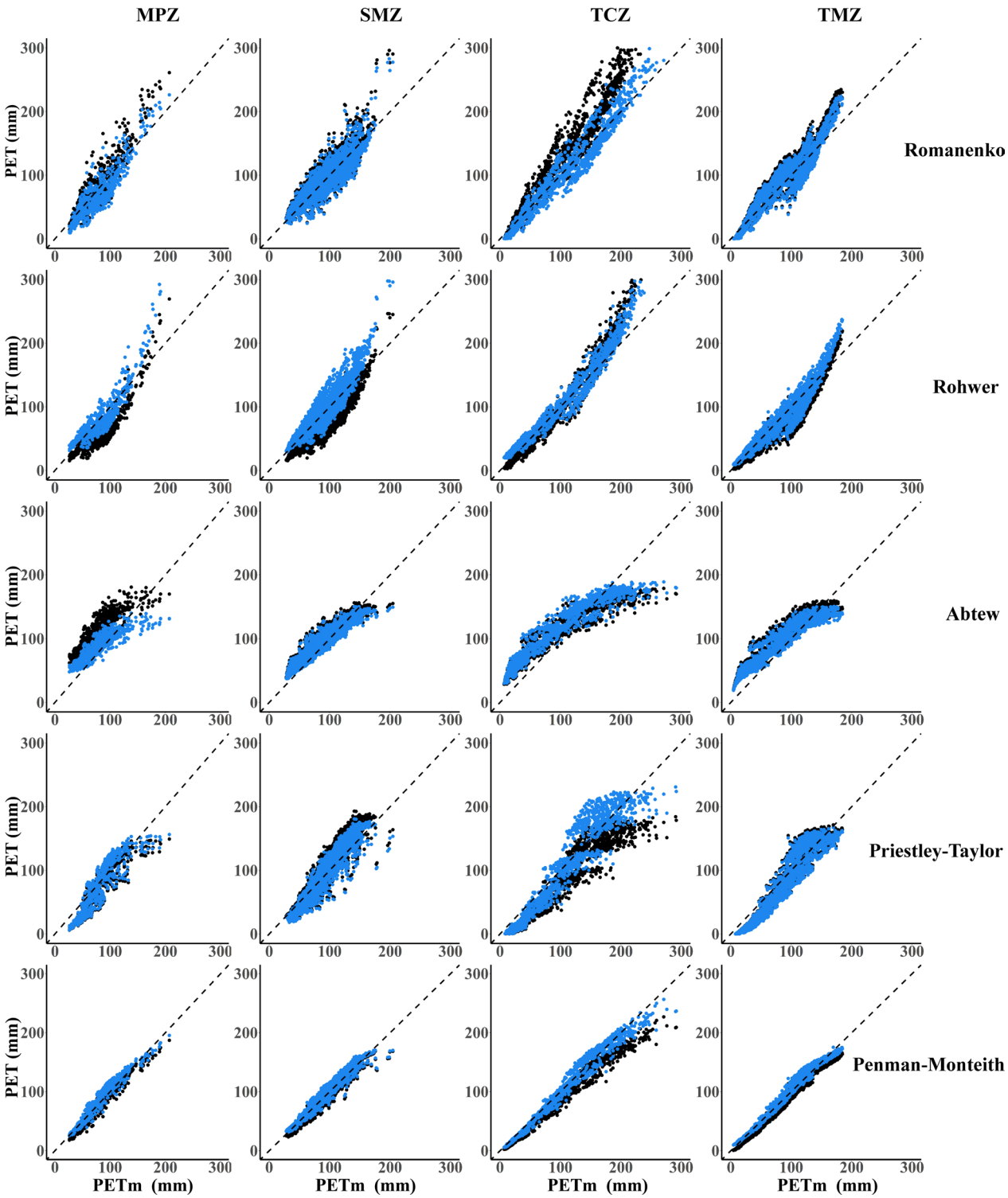

Fig. 15 Comparison of the monthly potential evapotranspiration (PET) calculated by the original (black circle) and calibrated (blue circle) potential evapotranspiration models against PETm in four climatic zones of China

as the optimal model among composite models and recommended by FAO, demonstrates superior performance among the 30 models studied and it is applicable to most regions in China.

Conclusions

After spatially interpolating meteorological data from 699 meteorological stations in four climatic zones of China from 2001 to 2020, 30 PET models were used to calculate PET in China. The PET values calculated varied significantly among the different PET models used, both across entire China and in different climatic zones. The Doorenbots–Pruitts model revealed the highest PET value (1902.6 mm), and the Kuzmin model revealed the lowest PET value (349.6 mm). In addition, the PET models used revealed considerable spatial heterogeneity.

The combined model category showed the highest calculation accuracy of the monthly mean PET in different climatic zones, followed respectively by radiation-, aerodynamic-, and temperature-based models. Based on the results obtained, the Romanenko and Rohwer were the recommended temperature- and aerodynamic-based models, respectively, while the remaining temperature- and aerodynamic-based models may not be suitable for PET calculations in the climatic zones of China. On the other hand, the Abtew model and the Priestley–Taylor radiation-based models outperformed other radiation models. Moreover, the Abtew model was more suitable for arid and semi-arid regions, while the Priestley–Taylor model was more suitable for humid regions. However, these radiation-based models should be used with caution in areas with low temperatures. In addition to the Rijtema model, the calculation accuracies of the combination-based models were relatively ideal, with the Penman–Monteith model being the best option for PET calculation.

The empirical coefficients of the optimal models of each category were calibrated using their average PET values. The results show an improvement in accuracy for all models, but there is a significant difference in the magnitude of the improvement over the different climatic zones, suggesting that the applicability of the different models varies considerably. In addition, from the perspective of the four climatic zones, the calculation accuracy range of Romanenko, Rohwer, Abten, Priestley Taylor, and Penman Monteith models improved in MPZ and TCZ is higher than that improved in TMZ and SMZ. This does not mean that the improved models have higher accuracy in MPZ and TCZ than in TMZ and SMZ. On the contrary, the original model performed poorly in MPZ and TCZ, so the improved accuracy was relatively large. The unimproved model was already more suitable in TMZ and SMZ, so the improved accuracy was relatively small. Therefore, regional calibration of the PET models can improve the accuracy and applicability of PET calculation, providing a reference for studying hydrological processes in different climatic zones.

Abbreviations

ET	Evapotranspiration
PET	Potential evapotranspiration
ET_0	Reference crop evapotranspiration
Epan	Pan evaporation
PETm	PET benchmark
MPZ	Mountain plateau zone
TCZ	Temperate continental zone
SMZ	Subtropical monsoon zone
TMZ	Temperate monsoon zone

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s13717-024-00488-7.

Additional file 1: Table S1. Meteorological stations and climate characteristics in the four climatic zones of China (2001–2020).

Acknowledgements

This work was supported by National Natural Science Foundation of China [U2243202], [42230714].

Author contributions

LZ wrote the paper and analyzed the data; LY selected the model; YX and JG constructed the thesis ideas; CP processed the latitude and longitude files; ZP, WY and DB discussed the details.

Funding

National Natural Science Foundation of China [U2243202], [42230714, 41977149].

Availability of data and materials

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

All authors approved the manuscript for publication in ecological processes.

Competing interests

No competing interest exists in this manuscript.

Author details

[1]Key Laboratory of Soil & Water Conservation and Desertification Combating of Ministry of Education, Beijing Forestry University, Beijing 100083, China. [2]State Key Laboratory of Earth Surface Processes and Resource Ecology, Faculty of Geographical Science, Beijing Normal University, Beijing 100875, China. [3]College of Horticulture and Forestry Sciences, Tarim University, Alar 841599, China.

Received: 22 November 2023 Accepted: 4 February 2024
Published online: 07 March 2024

References

Allen RG (1986) A Penman for all seasons. J Irrig Drain Eng 112:348–368. https://doi.org/10.1061/(ASCE)0733-9437(1986)112:4(348)

Allen RG, Pereira LS, Raes D, Smith M (1998) Crop evapotranspiration: guidelines for computing crop water requirements. Irrigation and Drainage Paper No. 56. Food and Agriculture Organization of the United Nations (FAO). Rome, Italy

Anabalón A, Sharma A (2017) On the divergence of potential and actual evapotranspiration trends: an assessment across alternate global datasets. Earth's Future 5:905–917. https://doi.org/10.1002/2016EF000499

Azhar AH, Perera BJC (2011) Evaluation of reference evapotranspiration estimation methods under Southeast Australian conditions. J Irrig Drain Eng 137:268–279. https://doi.org/10.1061/(ASCE)IR.1943-4774.0000297

Azorin-Molina C, Vicente-Serrano SM, Sanchez-Lorenzo A et al (2015) Atmospheric evaporative demand observations, estimates and driving factors in Spain (1961–2011). J Hydrol 523:262–277. https://doi.org/10.1016/j.jhydrol.2015.01.046

Bai P, Liu X, Yang T et al (2016) Assessment of the influences of different potential evapotranspiration inputs on the performance of monthly hydrological models under different climatic conditions. J Hydrometeorol 17:2259–2274. https://doi.org/10.1175/JHM-D-15-0202.1

Baier W, Robertson GW (1965) Estimation of latent evaporation from simple weather observations. Can J Plant Sci 45:276–284. https://doi.org/10.4141/cjps65-051

Blaney H, Criddle W (1950) Determining water requirements in irrigated areas from climatological and irrigation data, US Department of Agriculture, Soil Conservation Service, Technical Paper, No. 96

Bormann H (2011) Sensitivity analysis of 18 different potential evapotranspiration models to observed climatic change at German climate stations. Clim Change 104:729–753. https://doi.org/10.1007/s10584-010-9869-7

Chen JM, Liu J (2020) Evolution of evapotranspiration models using thermal and shortwave remote sensing data. Remote Sens Environ 237:111594. https://doi.org/10.1016/j.rse.2019.111594

Chen R, Ersi K, Yang J et al (2004) Validation of five global radiation models with measured daily data in China. Energy Convers Manage 45:1759–1769. https://doi.org/10.1016/j.enconman.2003.09.019

Cheng M, Jiao X, Jin X et al (2021) Satellite time series data reveal interannual and seasonal spatiotemporal evapotranspiration patterns in China in response to effect factors. Agric Water Manag 255:107046. https://doi.org/10.1016/j.agwat.2021.107046

Diks CGH, Vrugt JA (2010) Comparison of point forecast accuracy of model averaging methods in hydrologic applications. Stoch Environ Res Risk Assess 24:809–820. https://doi.org/10.1007/s00477-010-0378-z

Ding H, Greatbatch RJ, Park W et al (2014) The variability of the East Asian summer monsoon and its relationship to ENSO in a partially coupled climate model. Clim Dyn 42:367–379. https://doi.org/10.1007/s00382-012-1642-3

Donohue RJ, McVicar TR, Roderick ML (2010) Assessing the ability of potential evaporation formulations to capture the dynamics in evaporative demand within a changing climate. J Hydrol 386:186–197. https://doi.org/10.1016/j.jhydrol.2010.03.020

Doorenbos J, Pruitt WO (1975) Guidelines for predicting crop water requirements, irrigation and drainage paper no. 24, FAO–ONU, Rome. Italy

Douglas EM, Jacobs JM, Sumner DM, Ray RL (2009) A comparison of models for estimating potential evapotranspiration for Florida land cover types. J Hydrol 373:366–376. https://doi.org/10.1016/j.jhydrol.2009.04.029

Duethmann D, Blöschl G (2018) Why has catchment evaporation increased in the past 40 years? A data-based study in Austria. Hydrol Earth Syst Sci 22:5143–5158. https://doi.org/10.5194/hess-22-5143-2018

Fan J, Wu L, Zhang F et al (2016) Climate change effects on reference crop evapotranspiration across different climatic zones of China during 1956–2015. J Hydrol 542:923–937. https://doi.org/10.1016/j.jhydrol.2016.09.060

Feng Y, Jia Y, Zhang Q et al (2018) National-scale assessment of pan evaporation models across different climatic zones of China. J Hydrol 564:314–328. https://doi.org/10.1016/j.jhydrol.2018.07.013

Ge Q, Xue Z, Yao Z et al (2017) Anti-phase relationship between the East Asian winter monsoon and summer monsoon during the Holocene? J Ocean Univ China 16:175–183

Gharbia SS, Smullen T, Gill L et al (2018) Spatially distributed potential evapotranspiration modeling and climate projections. Sci Total Environ 633:571–592. https://doi.org/10.1016/j.scitotenv.2018.03.208

Hargreaves GH (1975) Moisture availability and crop production. Trans ASAE 18:980–984. https://doi.org/10.13031/2013.36722

Hargreaves GH, Samani ZA (1982) Estimating potential evapotranspiration. J Irrig Drain Eng Division 108:225–230. https://doi.org/10.1061/JRCEA4.0001390

Hargreaves GH, Samani ZA (1985) Reference crop evapotranspiration from temperature[J]. Appl Eng Agri 1(2):96–99. https://doi.org/10.13031/2013.26773

Herrnegger M, Nachtnebel H-P, Haiden T (2012) Evapotranspiration in high alpine catchments – an important part of the water balance! Hydrol Res 43:460–475. https://doi.org/10.2166/nh.2012.132

Irmak A, Irmak S (2008) Reference and crop evapotranspiration in South Central Nebraska. II: measurement and estimation of actual evapotranspiration for corn. J Irrig Drain Eng 134:700–715. https://doi.org/10.1061/(ASCE)0733-9437(2008)134:6(700)

Jakimavičius D, Kriaučiūnienė J, Gailiušis B, Šarauskienė D (2013) Assessment of uncertainty in estimating the evaporation from the Curonian Lagoon. Baltica 26:177–186. https://doi.org/10.5200/baltica.2013.26.18

Jensen ME, Haise HR (1965) Closure to "estimating evapotranspiration from solar radiation." J Irrig Drain Div 91:203–205. https://doi.org/10.1061/JRCEA4.0000342

Kendall MG (1948) Rank correlation methods. Griffin, Oxford, England

Kharrufa NS (1985) Simplified equation for evapotranspiration in arid regions. Hydrologie Sonderheft 5:39–47

Lakatos M, Weidinger T, Hoffmann L et al (2020) Computation of daily Penman-Monteith reference evapotranspiration in the Carpathian Region and comparison with Thornthwaite estimates. Adv Sci Res 16:251–259. https://doi.org/10.5194/asr-16-251-2020

Li S, Kang S, Zhang L et al (2016) Evaluation of six potential evapotranspiration models for estimating crop potential and actual evapotranspiration in arid regions. J Hydrol 543:450–461. https://doi.org/10.1016/j.jhydrol.2016.10.022

Li S, Wang G, Sun S et al (2021) Long-term changes in evapotranspiration over China and attribution to climatic drivers during 1980–2010. J Hydrol 595:126037. https://doi.org/10.1016/j.jhydrol.2021.126037

Liu W, Sun F (2016) Assessing estimates of evaporative demand in climate models using observed pan evaporation over China. J Geophys Res Atmos 121:8329–8349. https://doi.org/10.1002/2016JD025166

Liu Q, Yang Z, Xia X (2010) Trends for pan evaporation during 1959–2000 in China. Procedia Environ Sci 2:1934–1941. https://doi.org/10.1016/j.proenv.2010.10.206

Lu J, Sun G, McNulty SG, Amatya DM (2005) A comparison of six potential evapotranspiration methods for regional use in the southeastern united states. J Am Water Resources Assoc 41:621–633. https://doi.org/10.1111/j.1752-1688.2005.tb03759.x

Mahringer W (1970) Verdunstungsstudien am Neusiedler See. Arch Met Geoph Biokl B 18:1–20. https://doi.org/10.1007/BF02245865

Mann HB (1945) Nonparametric tests against trend. Econometrica 13:245. https://doi.org/10.2307/1907187

Mcnaughton K, Spriggs T (1988) An evaluation of the Priestley and Taylor equation and the complementary relationship using results from a mixed-layer model of the convective boundary layer. IAHS Publ, Wallingford, p 177

McVicar TR, Van Niel TG, Li L et al (2007) Spatially distributing monthly reference evapotranspiration and pan evaporation considering topographic influences. J Hydrol 338:196–220. https://doi.org/10.1016/j.jhydrol.2007.02.018

Milly PCD, Dunne KA (2016) Potential evapotranspiration and continental drying. Nature Clim Change 6:946–949. https://doi.org/10.1038/nclimate3046
Monteith JL (1965) Evaporation and environment. The stage and movement of water in living organisms. Symp Soc Exp Biol 19:205–234
Oudin L, Hervieu F, Michel C et al (2005) Which potential evapotranspiration input for a lumped rainfall–runoff model? J Hydrol 303:290–306. https://doi.org/10.1016/j.jhydrol.2004.08.026
Padmakumari B, Jaswal AK, Goswami BN (2013) Decrease in evaporation over the Indian monsoon region: implication on regional hydrological cycle. Clim Change 121:787–799. https://doi.org/10.1007/s10584-013-0957-3
Penman HL (1948) Natural evaporation from open water, bare soil and grass[J]. Math Phys Sci. 193(1032):120–145
Priestley CHB, Taylor RJ (1972) On the assessment of surface heat flux and evaporation using large-scale parameters. Mon Wea Rev 100:81–92. https://doi.org/10.1175/1520-0493(1972)100%3c0081:OTAOSH%3e2.3.CO;2
Prudhomme C, Williamson J (2013) Derivation of RCM-driven potential evapotranspiration for hydrological climate change impact analysis in Great Britain: a comparison of methods and associated uncertainty in future projections. Hydrol Earth Syst Sci 17:1365–1377. https://doi.org/10.5194/hess-17-1365-2013
Rijtema P, Wageningen C (1965) An analysis of actual evapotranspiration. SERBIULA (sistema Librum 20) 659
Seiller G, Anctil F (2016) How do potential evapotranspiration formulas influence hydrological projections? Hydrol Sci J 61:2249–2266. https://doi.org/10.1080/02626667.2015.1100302
Sheffield J, Wood EF, Roderick ML (2012) Little change in global drought over the past 60 years. Nature 491:435–438. https://doi.org/10.1038/nature11575
Shiri J (2018) Improving the performance of the mass transfer-based reference evapotranspiration estimation approaches through a coupled wavelet-random forest methodology. J Hydrol 561:737–750. https://doi.org/10.1016/j.jhydrol.2018.04.042
Shiri J, Nazemi AH, Sadraddini AA et al (2019) Alternative heuristics equations to the Priestley-Taylor approach: assessing reference evapotranspiration estimation. Theor Appl Climatol 138:831–848. https://doi.org/10.1007/s00704-019-02852-6
Singh VP, Xu C-Y (1997) Evaluation and generalization of 13 mass-transfer equations for determining free water evaporation. Hydrol Process 11:311–323. https://doi.org/10.1002/(SICI)1099-1085(19970315)11:3%3c311::AID-HYP446%3e3.0.CO;2-Y
Song Y, Achberger C, Linderholm HW (2011) Rain-season trends in precipitation and their effect in different climate regions of China during 1961–2008. Environ Res Lett 6:034025. https://doi.org/10.1088/1748-9326/6/3/034025
Stephens JC (1965) Discussion of "estimating evaporation from insolation." J Hydr Div 91:171–182. https://doi.org/10.1061/JYCEAJ.0001310
Tegos A, Malamos N, Koutsoyiannis D (2015) A parsimonious regional parametric evapotranspiration model based on a simplification of the Penman-Monteith formula. J Hydrol 524:708–717. https://doi.org/10.1016/j.jhydrol.2015.03.024
Thom AS, Oliver HR (1977) On Penman's equation for estimating regional evaporation. Q J Royal Meteo Soc 103:345–357. https://doi.org/10.1002/qj.49710343610
Thornthwaite CW (1948) An approach toward a rational classification of climate. Geogr Rev 38:55. https://doi.org/10.2307/210739
Valipour M (2014) Application of new mass transfer formulae for computation of evapotranspiration. J Appl Water Eng Res 2:33–46. https://doi.org/10.1080/23249676.2014.923790
Valipour M (2015) Evaluation of radiation methods to study potential evapotranspiration of 31 provinces. Meteorol Atmos Phys 127:289–303. https://doi.org/10.1007/s00703-014-0351-3
Van Bavel CHM (1966) Potential evaporation: the combination concept and its experimental verification. Water Resour Res 2:455–467. https://doi.org/10.1029/WR002i003p00455
Wang Y, Liu B, Su B et al (2011) Trends of calculated and simulated actual evaporation in the Yangtze River Basin. J Clim 24:4494–4507. https://doi.org/10.1175/2011JCLI3933.1
Wright JL, Jensen ME (1972) Peak water requirements of crops in Southern Idaho. J Irrig and Drain Div 98:193–201. https://doi.org/10.1061/JRCEA4.0013020
Xiang K, Li Y, Horton R, Feng H (2020) Similarity and difference of potential evapotranspiration and reference crop evapotranspiration – a review. Agric Water Manag 232:106043. https://doi.org/10.1016/j.agwat.2020.106043
Xu CY, Singh VP (2000) Evaluation and generalization of radiation-based methods for calculating evaporation[J]. Hydrol Proc 14(2):339–349. https://doi.org/10.1002/(SICI)1099-1085(20000215)14:2<339::AID-HYP928>3.0.CO;2-O
Xu C-Y, Singh VP (2001) Evaluation and generalization of temperature-based methods for calculating evaporation. Hydrol Process 15:305–319. https://doi.org/10.1002/hyp.119
Xu C-Y, Singh VP (2002) Cross comparison of empirical equations for calculating potential evapotranspiration with data from Switzerland. Water Resour Manage 16:197–219. https://doi.org/10.1023/A:1020282515975
Xu C-Y, Singh VP (2005) Evaluation of three complementary relationship evapotranspiration models by water balance approach to estimate actual regional evapotranspiration in different climatic regions. J Hydrol 308:105–121. https://doi.org/10.1016/j.jhydrol.2004.10.024
Xu Z, Liu S, Li X et al (2013) Intercomparison of surface energy flux measurement systems used during the HiWATER-MUSOEXE. J Geophys Res Atmos 118:13140–13157. https://doi.org/10.1002/2013JD020260
Xu X, Yu M, Lu J, Liu X (2016) Potential evapotranspiration estimation in the Upper Huaihe River Basin, China. Proc Eng 154:1018–1025. https://doi.org/10.1016/j.proeng.2016.07.591
Yang Y, Chen R, Han C, Liu Z (2021) Evaluation of 18 models for calculating potential evapotranspiration in different climatic zones of China. Agric Water Manag 244:106545. https://doi.org/10.1016/j.agwat.2020.106545
Yee MS, Pauwels VRN, Daly E et al (2015) A comparison of optical and microwave scintillometers with eddy covariance derived surface heat fluxes. Agric For Meteorol 213:226–239. https://doi.org/10.1016/j.agrformet.2015.07.004
Zandt G, Owens TJ (1986) Comparison of crustal velocity profiles determined by seismic refraction and teleseismic methods. Tectonophysics 128:155–161. https://doi.org/10.1016/0040-1951(86)90317-3
Zhang H, Liu G, Huang C (2010) Modeling all-sky global solar radiation using MODIS atmospheric products: a case study in Qinghai-Tibet Plateau. Chin Geogr Sci 20:513–521. https://doi.org/10.1007/s11769-010-0423-3
Zhao L, Xia J, Xu C et al (2013) Evapotranspiration estimation methods in hydrological models. J Geogr Sci 23:359–369. https://doi.org/10.1007/s11442-013-1015-9
Zheng H, Yu G, Wang Q et al (2017) Assessing the ability of potential evapotranspiration models in capturing dynamics of evaporative demand across various biomes and climatic regimes with ChinaFLUX measurements. J Hydrol 551:70–80. https://doi.org/10.1016/j.jhydrol.2017.05.056
Zhou J, Wang Y, Su B et al (2020) Choice of potential evapotranspiration formulas influences drought assessment: a case study in China. Atmos Res 242:104979. https://doi.org/10.1016/j.atmosres.2020.104979
Zuo D, Xu Z, Liu Z (2009) Comparison of the potential evapotranspiration estimation method based on air temperature in the arid areas of Northwest China. J Arid Land Resour Environ 23:123–130

Publisher's Note