Title: Evaluation of methods for estimating daily reference crop evapotranspiration at a site in the humid alpine meadow

Running title: Comparison of reference evapotranspiration with lysimeter measurements

Authors: Licong Dai [1,2,4], Ruiyu Fu[3], Xiaowei Guo[1,2*], Yangong Du[1,2], Fawei Zhang [1,2], Yikang Li[1,2], Guangmin Cao[1,2*]

Affiliations:

[1] Qinghai Provincial Key Laboratory of Restoration Ecology for Cold Region, Northwest Institute of Plateau Biology, Chinese Academy of Sciences, Xining 810001, China

[2] Key Laboratory of Adaptation and Evolution of Plateau Biota, Northwest Institute of Plateau Biology, Chinese Academy of Sciences, Xining, Qinghai 810001, China

[3] School of Geography and Planning, Nanning Normal University, Nanning 530001, China

[4]University of Chinese Academy of Science, Beijing 100049, China

* Corresponding author: Yangong Du and Guangmin Cao

E-mail: caogm@nwipb.cas.cn for Guangmin Cao and xwguo1206@163.com for XiaoweiGuo

Author contributions: L Dai and X Guo performed the research, analyzed data, and wrote the paper; F Zhang, R Fu, Y Li and Y Du analyzed data; G Cao conceived the study.

Conflict of interest: The authors declare no conflict of interest

Abstract

Evapotranspiration as the key component of the terrestrial water cycle, an accurate assessment of evapotranspiration is of great importance for water irrigation management. Although many applicable ET models have been developed, most models are mainly focused on low altitude regions, with little attention on alpine ecosystem worldwide. In this paper, we evaluate the performance of 14 evapotranspiration (ET_0) models by comparison with large weigh lysimeter measurements. Specifically, we use Bowen ratio-energy balance method, three combination models, seven radiation-based models and three temperature-based models driven with data from June 2017 to December 2018 in a humid alpine meadow, northeastern Qinghai-Tibetan Plateau. The daily actual evapotranspiration was obtained by large weighing lysimeters located in an alpine *Kobresia* meadow. We found that the performances of the 14 ET_0 models, ranked on the basis of their RMSE (root mean square error), decreased in the order: Bowen> Priestley-Taylor> DeBruin-Keijman> 1963Penman> FAO-24 Penman> FAO-56 Penman-Monteith > IRMAK1> Makkink(1957) > Makkink(1967)>Makkink> IRMAK2 > Hargreaves>Hargreaves1>Hargreaves2. For the combination models, FAO 24 Penman yielded the highest correlation, followed by Pen-63 and FAO-56 PM. For radiation-based models, PT and DK obtained the highest correlation, followed by Makkink, Makkink(1967) and Makkink (1957), IRMAK1 and IRMAK2. For temperature-based models, HAR, HAR1 and HAR2 obtained the same correlation. Overall, the Bowen performed best, with RMSEs 0.98, followed by combination models (ranged from 1.19 to 1.27 mm d^{-1} and averaged 1.22 mm d^{-1}), radiation-based models (ranged from 1.02 to 1.42 mm $d^{-1}$and averaged 1.27 mm d^{-1}) and temperature-based models (ranged from 1.47 to 1.48 mmd^{-1} and averaged 1.47 mm d^{-1}). Furthermore, all models tended to underestimate measured ET_a during periods of larger evaporative demand (i.e. growing season) and overestimate measured ET_a during lower evaporative demand (i.e. non-growing season). Our results could provide a new sight for the accurate assessment of evapotranspiration in an alpine ecosystem.

Keywords: alpine meadow; lysimeter measurement; Bowen ratio-energy balance method; combination models; radiation-based models; temperature-based models

1 Introduction

Evapotranspiration (ET) is one of the key parameters in the simultaneous processes of heat and water transfer to the atmosphere via transpiration and evaporation in the soil–plant–atmosphere system (Sentelhas, Gillespie, & Santos, 2010), thereby playing an important role in water balance calculations, water allocation and water irrigation management. Thus, an accurate assessment of evapotranspiration could improve water management strategies and promote the efficient use of water resources, especially in these regions suffering water shortages (Sun et al., 2011).

To-date, direct measurements of ET have been achieved by a variety of methods such as the Bowen Ratio Energy Balance System (A. Irmak & Irmak, 2008; S. Irmak, Allen, & Whitty, 2003; S. Irmak, Howell, Allen, Payero, & Martin, 2005; Si et al., 2005), lysimeters (Jia, Dukes, Jacobs, & Irmak, 2006; Valipour, 2015) and the eddy covariance technique (Novick et al., 2009; Zhang et al., 2018). Alternatively, ET can be indirectly assessed by applying various reference evapotranspiration equations. Several models of reference evapotranspiration have become widely used for the calculation of ET, and can be classified into three types: radiation-based models (Hargreaves & Samani, 1985), temperature-based models (Trajkovic, Gocic, Pongracz, & Bartholy, 2019), and combination models (Penman, 1963).While the development of these models has undoubtedly benefited the calculation of ET, it still difficult to choose the optimal one due to the availability of the observed data together with most models have not been evaluated against lysimeter measurements across a range of regions and climates (Kiefer, Andresen, Doubler, & Pollyea, 2019; Liu et al., 2017). To select the best performing models, many studies have been conducted to assess model performance under various climates. For instance, the Food and Agriculture Organization of the United Nations (FAO) recommend the Penman–Monteith FAO-56 combination equation (PM-56) as the standard equation for estimating reference evapotranspiration (Richard G Allen, Pereira, Raes, & Smith, 1998), and this has been widely used worldwide when compared with other equations (Cai, Liu, Lei, & Pereira, 2007). The advantages of the Penman–Monteith equation are that it is does not require any local calibration because it incorporates both physiological and aerodynamic parameters, and it has been well tested by a variety of lysimeters (Trajkovic, 2009). Although many models have been widely used to estimate ET, it should be noted that most previous models have only been evaluated with reference to FAO-56 PM (Cao, Li, Liu, Zhong, & Zhao, 2015; Liu et al., 2017), with few being tested against lysimeter measurements. Furthermore, the application of the PM-56 equation needs many meteorological inputs, such as wind speed, temperature, humidity and solar radiation, that are often not available in regions with harsh environments (Martel, Glenn, Wilson, & Kröbel, 2018; Tabari, Kisi, Ezani, & Talaee, 2012). Thus, it is essential to develop a relatively accurate reference evapotranspiration model that requires fewer meteorological parameters, to allow more simplified estimates of ET than those of PM-56, applicable across a range of climatic conditions (Tabari & Talaee, 2011). So far, many models have been developed. For example, Tabari (2010) assessed four reference evapotranspiration models in an arid climate, and found that the Turc model performed the best. Meanwhile, the Hargreaves equation performed best in semiarid regions (Sabziparvar & Tabari, 2010). Liu et al. (2017) compared 16 models for reference evapotranspiration against weighing lysimeter measurements, and found that the combination models performed best for estimating ET in semiarid regions. Overall, most previous studies have been conducted in low-humidity conditions at low altitude (i.e. arid and semiarid regions) (Liu et al., 2017; Sentelhas et al., 2010), with few studies in humid climates, particularly in alpine ecosystems.

The Qinghai-Tibetan Plateau (QTP),with an average altitude of 4000 m, is the world's highest alpine ecosystem and is also known as the "Asian tower", playing an important role in ensuring the safety of water resources in China and southeast Asia (Dai, Guo, Zhang, et al., 2019; Zou et al., 2017). The alpine meadow almost account for almost 60% of the plateau area (Dai, Ke, et al., 2019); therefore, accurate assessment of ET in an alpine ecosystem are not only provides new insights into the water cycle, but also benefit the formulation of water resource management strategies. Furthermore, given the uncertainty and confusion in the selection of ET equations across different regions and climates, it is critical to thoroughly understand the performance of the various models in a humid alpine meadow across worldwide (Zhang et al., 2018). The objectives of the study was to evaluate the performance of 14 evapotranspiration (ET_0) models by comparison with large weigh lysimeter measurements, with the aims of selecting the best fit model in applications in this climate zone worldwide to estimate the ET.

2 Materials and methods

2.1 Study area

The study was conducted at the Haibei National Field Research Station, Qinghai, China (37°37′ N, 101°19′E), which is situated on the Northeastern QTP at an elevation of 3200 m a.m.s.l (Fig.1a). This area is characterized by a plateau continental monsoon climate, with well-developed seasonally frozen ground. The average annual air temperature is –1.7°C, with the maximum monthly temperature in July (10.1°C) and minimum monthly temperature in January (–15.0°C). The annual precipitation is about 580 mm, of which 80% falls in the growing season (i.e. from May to September), lead to a high water content (close to field capacity) in the soil during growing season, thus the evapotranspiration during growing season can be considered as reference evapotranspiration (ET_0) according to FAO indications (Richard G Allen et al., 1998). The average annual pan evaporation is approximately 1191.4 mm (Zhang et al., 2018). The soil type around the lysimeter system is classified as Mat-Gryic Cambisol, which belongs to a clay loam, and has a thickness of approximately 60 – 80 cm (Dai, Guo, Du, et al., 2019), the basic soil property was show in Table 1. The grass crop is dominated by perennial sedge and graminoid species, including *Kobresia humilis*, *Stipa aliena* and *Elymus nutans*, which together constitute 60% - 80% of plant cover around the lysimeter system (Fig.1b).

2.2 ET measurement and data quality control

The actual ET was measured by large-scale weighing lysimeters (height 2 m, diameter 1 m, and resolution 1g) (Fig.1c), which was recorded with a data logger (CR1000, Campbell, USA). The soil in the two lysimeters was repacked soil, the time step of lysimeters measurement was 30 min. The grass crop was keep the consistent conditions to so that the measured data as representative as possible. To ensure data quality, all negative or abnormal ET values caused by falling soil particles were discarded; the abnormal ET values were these outlier that more than three times of average ET, this screening process yielded 393 days of data spanning June 2017 to December 2018. According to in-situ phenological observations of the foliage of dominant plants, we defined the growing season as that from May 1stto September 30th, while the period from October 1st to April 30th of the following year is defined as the non-growing season (Zhang et al., 2018).

2.3 Meteorological data collection

All meteorological variables needed to calculate ET_0 using the various models were obtained or estimated from the weather station at Haibei Station, and included relative humidity (RH), wind speed (Field, Behrenfeld, Randerson, & Falkowski), net radiation (R_n), total radiation (R_s), extraterrestrial solar radiation (R_a), soil heat flux (G), maximum air temperature (T_{max}),vapor pressure deficit (VPD), minimum air temperature (T_{min}) and mean air temperature (T). The radiation were measured by four radiometers (CNR4, Kipp&Zonen, Netherlands) at 1.5 m height; relative humidity, wind speed and mean air temperature were measured at 1.5 m height (HMP45C, Vaisala, Finland), and wind speed was converted to 2 m height for calculating ET_0. The soil heat flux (G) was measured by three heat flux plates (HFT-3, Campbell, USA), which were separately buried 5 cm beneath the surface. Half-hourly means of meteorological data were stored by a data logger (9210 XLITE, Sutron, USA). It should be note that the distance between weather station and lysimeters was less than 10 m.

2.4 Reference evapotranspiration models

A total of 14 often-used ET_0 models were selected for comparison, including Bowen ratio-energy balance method, three combination models (1963 Penman, FAO24 Penman and FAO-56 PM), seven radiation models (Priestley-Taylor, De Bruin-Keijman, Makkink, Makkink (1957), Makkink (1967), IRMAK1 and IRMAK2) and three temperature-based models (Hargreaves, Hargreaves1, Hargreaves2), to compare their performance by using lysimeter measurement. The specific equations and parameters of these models are listed in Table 2.

2.5 Evaluation criteria

Here, the ET_a measured by the large-scale weighing lysimeters was the actual ET, and the performances of the ET_0 equations were compared to these lysimeter system estimates on a daily time step. The pair-wise comparisons were conducted using general linear regression. For further comparison, the root mean squared error (RMSE), percentage error of estimate (PE), mean absolute error (MAE) and coefficient of determination (R^2) were used for the evaluation of reference evapotranspiration models. The RMSE, PE, MAE and R^2 are defined as:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(P_i - O_i)^2}{n}}$$

$$PE = \left|\frac{\overline{P} - \overline{O}}{\overline{O}}\right| \times 100\%$$

$$MAE = \frac{\sum_{i=1}^{n}(P_i - O_i)}{n}$$

$$R^2 = \frac{[\sum_{i=1}^{n}(P_i - \overline{P})(O_i - \overline{O})]^2}{\sum_{i=1}^{n}(P_i - \overline{P})^2 \sum_{i=1}^{n}(O_i - \overline{O})^2}$$

where P_i are the predicted values, O_i are observed values, $\overline{P}$ and $\overline{O}$ are the average of P_i and O_i, and n is the total number of data.

2.6 Statistical analysis

To achieve the best comparison between the models and measurements, we need to select the dominant meteorological factors affecting the measured ET. Given that it may not be appropriate to explore results based solely on the coefficient of independent variables in multiple regression analysis, owing to the strong collinearities and nonlinearities among meteorological factors, we adopted a boosted regression trees (BRT) model to quantitatively evaluate the relative influences of meteorological variables on measured ET. In the past, the BRT method has been widely used to improve the performance of a single model through by fitting a large number of models, ultimately yielding an overall prediction (Martínez-Rincón, Ortega-García, & Vaca-Rodríguez, 2012). Most importantly, the BRT can evaluate the relative influence of an independent variable on a dependent variable, without transformations, and can cope well with non-linear relationships. Furthermore, the BRT displays good performance in dealing with stronger collinearities and nonlinearities. Thus, the BRT was adopted to evaluate the individual influences of controlling factors on measured ET. All statistical analyses were conducted in R software version 3.03 (R Development Core Team, 2006), and all figures were plotted by Origin 9.0.

3 Results

3.1 Seasonal variation of measured ET and environmental variables

The measured ET showed a clear seasonal pattern, the growing season ET was significantly higher than that in the non-growing season (P<0.05) (Fig.2a). The average measured daily ET during the study period was 2.33±0.11 mm d^{-1}, with average daily measured ET of 4.14±0.13 mm and 0.65 ± 0.05 mmd^{-1} (here and throughout the remainder of the paper, values are expressed as mean ± 1 SD) during the growing season and non-growing season, respectively (Fig.2a). Environmental variables showed a similar seasonal pattern, with the maximum and minimum values in the growing season and non-growing season (Fig. 2).The average daily R_n, R_s, T_a, VPD and RH during the growing season were 9.53 ± 0.28 MJ m^{-2}, 18.50 ± 0.38 MJ m^{-2} and 10.15 ± 0.21 °C, 0.33 ± 0.01 kPa, 74.88 ± 0.62 %, respectively (Fig.2). The average daily R_n, R_s, T_a, VPD and RH during the non-growing season were 3.28 ± 0.15 MJ m^{-2}, 12.65 ± 0.25 MJ m^{-2} and -3.67 ± 0.42 °C, 0.23 ± 0.01 kPa, 57.16 ± 0.86 %, respectively.

3.2 Comparison of daily ET between reference evapotranspiration models and lysimeter measurements during the study period

The comparison of 14 evapotranspiration equations against the lysimeter measurements is presented in Fig.3 and Table 3, showing that the relationship between daily ET_0 calculated by the reference evapotranspiration equations and lysimeter measurements are each significantly

(P<0.01), with higher coefficients of determination(R^2) ranging from 0.59 to 0.86. For the combination models, FAO 24 Penman yielded the highest correlation, followed by FAO-56 PM and Pen-63. For radiation-based models, PT obtained the highest correlation, followed by DK, Makkink(1967) and Makkink (1957), IRMAK1, Makkink and IRMAK2. For temperature-based models, HAR obtained the highest correlation, HAR1 and HAR2 obtained the same correlation. The daily estimates of combination models generally underestimated the ET values measured by lysimeter, with MAE -0.26 mm d^{-1} to -0.01 mm d^{-1}. However, the radiation-based models (except PT and IRMAK1) and temperature-based models generally temperature-based models, with MAE ranging from -0.14 mm d^{-1} to 0.50 mm d^{-1} for radiation-based models, and ranging from 0.40 mm d^{-1} to 0.59 mm d^{-1} for temperature-based models.

During the whole study period, the RMSE of Bowen ratio-energy balance method was 0.98, the RMSE of combination models ranged from 1.19 to 1.27 mm d^{-1} and averaged 1.22 mm d^{-1}, Furthermore, the RMSE of FAO-56 PM increased from 1.22 to 1.29 mm d^{-1} as r_s changed from 20 to 60 s m^{-1}(Fig.4),. The RMSE for radiation-based models ranged from 1.02 to 1.42 mm $d^{-1}$and averaged 1.27 mm d^{-1}, and the RMSE for temperature-based models ranged from 1.47 to 1.48 mmd^{-1} and averaged 1.47 mm d^{-1}. Based on the RMSE, the performances of the 14 evapotranspiration models decreased in the order: Bowen>PT>DK>Pen-63>FAO-24 Pen>FAO-56 PM> IRMAK1>Makkink (1957) >Makkink (1967) > Makkink > IRMAK2>HAR>HAR1>HAR2. The best model (Good, Noone, & Bowen) was, respectively, 34% and 30% more accurate than the poorest (HAR2) and the commonly used FAO-56 PM equation. Furthermore, the Pen-63 and FAO-24 Pen demonstrated better performance than the commonly used FAO-56 PM equation. Overall, for the whole study period, the Bowen yielded the best performance, followed by the combination models, radiation-based models and temperature-based models.

3.3 Comparison of daily ET between reference evapotranspiration models and lysimeter measurements during the growing season

During the growing season, the daily ET_0 calculated by 14 evapotranspiration equations was significantly correlated with the lysimeter measurements (P<0.01), with R^2 ranging from 0.32 to 0.64 (Fig.5 and Table 4). Of the combination models, FAO-56 PM obtained the highest R^2, followed by FAO-24 Pen and Pen-63. Of the radiation-based models, PT obtained the highest R^2, followed by DK, IRMAK1, Makkink(1967), Makkink (1957), IRMAK2 and Makkink. It should be noted that Makkink(1967) and Makkink (1957) have the same R^2. Of the temperature-based models, HAR, HAR1 and HAR2 obtained the same R^2. Interestingly, all models (except for HAR1and HAR2) generally underestimated ET during the growing season, values of MAE ranged from −1.10 to -0.15 mm d^{-1}, with FAO-56 PM having the largest underestimate (by 26.59%) and HAR the minimum underestimate (by 3.56%) (Table 4).

The RMSE of Bowen was 1.31, the RMSE for combination models ranged from 1.38 to 1.58 mm d^{-1} and averaged 1.47 mm d^{-1}, the RMSE for radiation-based models ranged from 1.19 to 1.55 mm $d^{-1}$and averaged 1.40 mm d^{-1}, and the RMSE for temperature-based models ranged from 1.42 to 1.43 mm d^{-1} and averaged 1.42 mmd^{-1} (Table 4). Based on the RMSE values, the performances of the 14 evapotranspiration models follow the order: DK>PT> Makkink (1967)>Bowen>Pen-63>HAR1>HAR2>HAR>FAO-24Pen>Makkink>IRMAK1> Makkink(1957)> IRMAK2>FAO-56. Evidently, the best (DK) was 25% more accurate than the poorest (FAO-56). Overall, for the growing season period, the Bowen yielded the best performance, followed by the radiation-based models, temperature-based models and combination models.

3.4 Comparison of daily ET between reference evapotranspiration models and lysimeter measurements during the non-growing season

During the non-growing season, the daily ET_0 calculated by the 14 evapotranspiration equations was also significantly correlated with the lysimeter measurements (P<0.01), but with lower coefficients of determination (R^2) ranging from 0.17 to 0.64 (Fig.6 and Table 5). Of the combination models, FAO-24 Pen obtained the highest R^2, followed by Pen-63 and FAO-56 PM. Of the radiation-based models, PT and DK obtained the highest R^2, followed by Makkink(1967), Makkink (1957), Makkink, IRMAK1 and IRMAK2. Of the temperature-based models, HAR obtained the highest R^2. Interestingly, all model (except Bowen) generally overestimated the ET values measured by lysimeter during the non-growing season, with MBEs ranging from 0.40 to 1.20 mm d^{-1} and averaging 0.78 mm d^{-1}; Makkink (1967) yielded the largest underestimate (by 185.83 %) and PT the minimum underestimate (by 61.06%)(Table 5).

The RMSE of Bowen was 0.52, the RMSE for combination models ranged from 0.86 to 1.00 mm d^{-1} and averaged 0.92 mm d^{-1}, the RMSE for radiation-based models ranged from 0.80 to 1.44 mm $d^{-1}$and averaged 1.12 mm d^{-1}, and the RMSE for temperature-based models ranged from 1.47 to 1.54 mm d^{-1} and averaged 1.50 mm d^{-1}. Based on the RMSE, the performances of the reference evapotranspiration models decreased in the order: Bowen>PT>DK> FAO-24Pen> FAO-56>Pen-63>IRMAK1>Makkink(1957)>Makkink>IRMAK2>Makkink(1967)>HAR>HAR1>HAR2. Evidently, the best model was 66.24% more accurate than the poorest (HAR2). Overall, for the non-growing season period, the Bowen yielded the best performance, followed by the combination models, radiation-based models and temperature-based models.

3.5 Comparison of monthly averaged daily ET_0 between reference evapotranspiration

models and lysimeter measurements

The estimations of 14 evapotranspiration models were consistent with the pattern shown by lysimeter measurements (Fig. 7), with the peak in July. As already noted, the combination models and Bowen underestimated the measurements from May to September, and overestimated the measurements in the other months (Fig.7a and 7b). The radiation-based models underestimated the measurements from June to September, and overestimated the measurements in the other months (Fig. 7c). However, the temperature-based models generally overestimated the measured ET during most months (except July and August) (Fig. 7d). Overall, all models tended to underestimate the measured ET during the growing season (with larger evaporative demand), and overestimated ET during the non-growing season (with reduced evaporative demand).

3.6 Dominant factors affecting the seasonal variation in lysimeter ET measurements

BRT model indicated that the R_n was the dominant factor controlling the seasonal variation in measured ET during the whole study period, accounting for 69.02% of total variability, followed by VPD (7.13%), T (6.75%), RH (5.73%), R_a (5.15%), R_s (4.33%) and WS (1.85%) (Fig.8a). During the growing season, the R_n remained the main control of seasonal variation in measured ET (Fig.8b), accounting for 44.30% of total variability, followed by T (14.12%), R_s (12.56%), R_a (8.34%), RH (7.80%), VPD (6.87%) and WS (6.02%). However, the seasonal variation of measured ET in non-growing season was dominated by RH (Fig.8c), accounting for 27.99% of total variability, followed by R_n (20.99%), R_a (12.96%), VPD (12.56%), T (10.18%), R_s (9.39) and WS (5.94%) (Fig.8).

4 Discussion

4.1 The performance comparison of combination models against lysimeter measurements

Previous study have shown that the Penman family models are generally the most accurate when evaluating ET across various climate scenarios and regions (Liu et al., 2017). Of the penman models for ET_0, the Penman–Monteith FAO 56 has been considered as the standard equation for estimating evapotranspiration (Richard G Allen et al., 1998). For instance, Yoder et al. (2005) found that the Penman–Monteith FAO 56 display the best performance in the humid southeast United States. López-Urrea et al. (2006) test seven evapotranspiration equations using lysimeter observations in a semiarid climate, and found the FAO-56 Penman–Monteith equation was the most precise method compared with other evapotranspiration equations. Contrary to previous studies, the Penman–Monteith FAO 56 was not the best in our study, we found Pen-63 and FAO-24 Pen were more accurate (Pen-63 and FAO-24 Pen had smaller RMSE than Penman–Monteith FAO 56 during the whole study period). Such similar results have been reported in many other studies (Berengena & Gavilán, 2005; Martel et al., 2018). A more recent study also reported the poor performance of FAO-56 PM when compared with data from 20 FLUXNET towers (Ershadi, McCabe, Evans, Chaney, & Wood, 2014). Combining these results suggests that FAO-56 PM might not be the only standard model for evaluating ET_0, because it did not yield better accuracy than the other penman models. Given the better performance of Pen-63 and FAO-24 than FAO-56 PM in this study, we may apply old Penman family models to our study region, especially considering the Penman–Monteith FAO 56 requires many meteorological inputs, which limit its use in areas with sparse data, especially in harsh environment (Tabari et al., 2012). Overall, the poor performance of FAO-56 PM may attributed to the higher aerodynamic resistance (r_s), there is increasing evidence indicated that the underestimation of FAO-56 PM suggested that the fixed r_s= 70 s m^{-1} in the equation is probably too large (Liu et al., 2017), this result was also confirmed by our results that the values of RMSE increased from 1.22 to 1.29 mm d^{-1} as r_s changed from 20 to

60 s m^{-1}, and the RMSE was nearly unchanged when r_s varied between 0 and 20 s m^{-1}, the RMSE was range from 1.12 to 1.13 mm d^{-1} (Fig. 4). Therefore, reducing value of r_s from 70 s m^{-1} to 0–20 s m^{-1} can improve daily estimates of the FAO-56 PM. Other studies also found that the r_s should be a variable value rather than the fixed one (Richard G. Allen et al., 2006; Liu et al., 2017). For instance, the r_s should be smaller when it was being underestimated and should be larger when it was being overestimated (Ventura, Spano, Duce, & Snyder, 1999).

4.2 The performance comparison of radiation and temperature models against lysimeter measurements

For the performance comparison of radiation models against lysimeter measurements, we found that the PT models yielded the best performances of the radiation-based models, which was in line with the previous study conducted in humid areas that the PT method exert the good accuracy estimate for ET (Ershadi et al., 2014). There is increasing evidence indicated that the input parameters was the dominated factors affecting their performance (Lang et al., 2017), we thus conclude that the better performances of the PT models which might associated with the use of only the most important meteorological factors affecting ET such as net radiation (R_n), was supported by our results that the net radiation was the most important factor controlling measured ET (Fig.8). Compare with PT, the other radiation models just use the R_s as the mainly driving variable, the ET thus may overestimate because some R_s was reduced through reflecting into the atmosphere due to the high albedo in this region (Zhang et al., 2018). Furthermore, each model was developed from its specific underlying surface and climate conditions. For instance, the PT was established in a humid climate condition, which was suitable for our humid alpine meadow. Most importantly, the PT models required fewer meteorological inputs when compared with combination models. Combine those factors, we can prefer to recommend the PT and DK model for use in a humid alpine meadow on the northeastern Qinghai-Tibetan Plateau, especially when considering the difficulty in obtaining ET in this harsh climate.

For the performance comparison of temperature models against lysimeter measurements, previous study reported that the Hargreaves versions equation as one of the most simple empirical methods were widely used for the ET estimation due to its less meteorological data input, especially considering not all the meteorological data required in the standard PM-56 model (Jensen, Burman, & Allen, 1990). To further select the best Hargreaves versions equation, we compare the performance of original (HAR) and two modified versions (HAR1 and HAR2) of the Hargreaves equations, and found that the original HAR model had the lowest error (RMSE =1.47 mm d^{-1}, MAE=0.40 mm d^{-1} and PE=17.37 %), which was consistent with previous studies conducted on humid region (Tabari, 2010) but contrast to these study conducted in arid region that the modified Hargreaves equation display a more accurate estimation of evapotranspiration compared with the original Hargreaves equation (Ravazzani, Corbari, Morella, Gianoli, & Mancini, 2012). Overall, the temperature models display a poor performance compared with radiation models owing to the Hargreaves method was established in semiarid areas (Tabari, 2010), thus a local calibrations was required to improve the accuracy of Hargreaves method in other region.

4.3 The performance comparison of all the models

By comparing the four type models, we found that the Bowen yielded the best performance, followed by the combination models, radiation-based models and temperature-based models (Table 3). Overall, most radiation-based models generally underestimated the measured ET during the whole study period, whereas the temperature-based models tended to overestimate ET. This is consistent with previous studies where the Makkink and Priestley-Taylor models generally underestimated ET (Priestley & Taylor, 1972; Xu & Singh, 2002), while the Hargreaves equations often overestimate ET in cold-humid conditions and requires a local calibration (Berti, Tardivo, Chiaudani, Rech, & Borin, 2014). Given that the study region in our study belongs to humid alpine meadow, thus ET tended to be overestimated. An alternative explanation for the poor performance of the Hargreaves model in humid regions may owing to the Hargreaves method was established in semiarid areas (Tabari, 2010), and the R_a parameter used in the Hargreaves model, which is based on the maximum possible radiation value and does not take the atmospheric transmissivity into account. However, the atmosphere transmissivity in humid regions is affected by many factors, such as atmospheric moisture; thus, the solar radiation reaching the surface is significantly reduced due to the high atmospheric moisture content (Temesgen, Allen, & Jensen, 1999), resulting in the overestimation of solar radiation, ultimately leading to an overestimation of ET by the Hargreaves method.

Furthermore, there were also common features of all four groups of models. All the models tended to underestimate the measured ET during the growing season (with larger evaporative demand), and overestimated ET during the non-growing season (with reduced evaporative demand), which was consistent with a previous study conducted in a semi-arid region (Liu et al., 2017). Furthermore, we found that the measured ET and calculated ET_0 were less correlated during non-growing season than during growing season. These discrepancies may relate to the dominant component between transpiration and evaporation. The transpiration was the dominant during growing season, almost account for 75% of evapotranspiration, whereas the evaporation was the dominant component during non-growing season in the same study site (Zhang et al., 2018). Considering the evaporation process was much complex and affected by many environmental factors compared with transpiration process, ultimately lead to a poor correlation between measured ET and calculated ET_0 during non-growing season. Therefore, both Hargreave's equations and other models need further local or region calibration before being applied to a given region (Xu & Singh, 2002). Besides, it should be noted that the data used in this study just obtained from a year and a single weather station, which may insufficient to represent the whole humid climate or the alpine ecosystem but represent a specific site. Thus, a longer period and more lysimeter systems should be used in the alpine ecosystem in the future to obtain more accurate estimates of evapotranspiration over the northeastern Qinghai-Tibetan Plateau.

5 Conclusion

This study is the first to document information on the comparison of evapotranspiration models against lysimeter measurements in a humid alpine meadow. we found that the Bowen ratio-energy balance method performed the best, followed by combination models, radiation-based models and temperature-based models. In addition, the combination models tended to underestimate measured ET, whereas temperature-based models and most radiation-based models tended to overestimate measured ET during the whole study period. Specifically, all models tended to underestimate ET during the growing season and overestimate ET during the non-growing season, suggesting that these models should be calibrated or modified by local lysimeter data when extrapolated to other regions. Furthermore, the 1963 Penman and FAO-24 Penman models demonstrated better performances than the recommended FAO-56 Penman-Monteith (PM), suggesting that older Penman equations may be superior to the standard FAO-56 Penman-

Monteith model, especially considering the good performance of the 1963 Penman model in this study. Given the outstanding performance of Priestley-Taylor model, which require few meteorological inputs, we thus recommend that these two models can be used in other alpine meadows that have similar climates to that of the study region, to improve ET estimation.

Conflict of interest

The authors declare that there are no conflicts of interest associated in this manuscript.

Data Availability Statement

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

Acknowledgments

This work was supported by the National Natural Science Foundation of China (41730752, 31700395, 31770532 and 31400483), the comprehensive management and application demonstration of small watershed in Hainan prefecture (2019-SF-152).

References

Allen, R. G., Pereira, L. S., Raes, D., & Smith, M. (1998). Crop evapotranspiration-Guidelines for computing crop water requirements-FAO Irrigation and drainage paper 56. *Fao, Rome, 300*(9), D05109.

Allen, R. G., Pruitt, W. O., Wright, J. L., Howell, T. A., Ventura, F., Snyder, R., . . . Yrisarry, J. B. (2006). A recommendation on standardized surface resistance for hourly calculation of reference ETo by the FAO56 Penman-Monteith method. *Agricultural Water Management, 81*.

Berengena, J., & Gavilán, P. (2005). Reference evapotranspiration estimation in a highly advective semiarid environment. *Journal of Irrigation and Drainage Engineering, 131*(2), 147-163.

Berti, A., Tardivo, G., Chiaudani, A., Rech, F., & Borin, M. (2014). Assessing reference evapotranspiration by the Hargreaves method in north-eastern Italy. *Agricultural Water Management, 140*, 20-25.

Cai, J., Liu, Y., Lei, T., & Pereira, L. S. (2007). Estimating reference evapotranspiration with the FAO Penman–Monteith equation using daily weather forecast messages. *Agricultural and Forest Meteorology, 145*(1-2), 22-35.

Cao, J., Li, Y., Liu, X., Zhong, X., & Zhao, Y. (2015). Comparison of four combination methods for reference crop evapotranspiration. *Chinese Journal of Agrometeorology, 36*(4), 428-436.

Dai, L., Guo, X., Du, Y., Zhang, F., Ke, X., Cao, Y., . . . Cao, G. (2019). The Response of Shallow Groundwater Levels to Soil Freeze–Thaw Process on the Qinghai−Tibet Plateau. *Groundwater, 57*(4), 602-611.

Dai, L., Guo, X., Zhang, F., Du, Y., Ke, X., Li, Y., . . . Shu, K. (2019). Seasonal dynamics and controls of deep soil water infiltration in the seasonally-frozen region of the Qinghai-Tibet plateau. *Journal of Hydrology, 571*, 740-748.

Dai, L., Ke, X., Guo, X., Du, Y., Zhang, F., Li, Y., . . . Shu, K. (2019). Responses of biomass allocation across two vegetation types to climate fluctuations in the northern Qinghai–Tibet Plateau. *Ecology and evolution, 9*(10), 6105-6115.

De Bruin, H., & Stricker, J. (2000). Evaporation of grass under non-restricted soil moisture conditions. *Hydrological sciences journal, 45*(3), 391-406.

Droogers, P., & Allen, R. G. (2002). Estimating reference evapotranspiration under inaccurate data conditions. *Irrigation and drainage systems, 16*(1), 33-45.

Ershadi, A., McCabe, M., Evans, J. P., Chaney, N. W., & Wood, E. F. (2014). Multi-site evaluation of terrestrial evaporation models using FLUXNET data. *Agricultural and Forest Meteorology, 187*, 46-61.

FAO. (1998). Crop evapotranspiration Guidelines for computing crop water requirements. *Fao Irrigation & Drainage Paper, 56*.

Field, C. B., Behrenfeld, M. J., Randerson, J. T., & Falkowski, P. (1998). Primary production of the biosphere: integrating terrestrial and oceanic components. *Science, 281*(5374), 237-240.

Good, S. P., Noone, D., & Bowen, G. (2015). Hydrologic connectivity constrains partitioning of global terrestrial water fluxes. *Science, 349*(6244), 175-177.

Hansen, S. (1984). Estimation of Potential and Actual Evapotranspiration: Paper presented at the Nordic Hydrological Conference (Nyborg, Denmark, August-1984). *Hydrology Research, 15*(4-5), 205-212.

Hargreaves, G. H., & Samani, Z. A. (1985). Reference crop evapotranspiration from temperature. *Applied engineering in agriculture, 1*(2), 96-99.

Irmak, A., & Irmak, S. (2008). Reference and crop evapotranspiration in South Central Nebraska. II: Measurement and estimation of actual evapotranspiration for corn. *Journal of Irrigation and Drainage Engineering, 134*(6), 700-715.

Irmak, S., Allen, R., & Whitty, E. (2003). Daily grass and alfalfa-reference evapotranspiration estimates and alfalfa-to-grass evapotranspiration ratios in Florida. *Journal of Irrigation and Drainage Engineering, 129*(5), 360-370.

Irmak, S., Howell, T., Allen, R., Payero, J., & Martin, D. (2005). Standardized ASCE Penman-Monteith: Impact of sum-of-hourly vs. 24-hour timestep computations at reference weather station sites. *Transactions of the ASAE, 48*(3), 1063-1077.

Jensen, M. E., Burman, R. D., & Allen, R. G. (1990). *Evapotranspiration and irrigation water requirements*.

Jia, X., Dukes, M. D., Jacobs, J. M., & Irmak, S. (2006). Weighing lysimeters for evapotranspiration research in a humid environment. *Transactions of the ASABE, 49*(2), 401-412.

Kiefer, M. T., Andresen, J. A., Doubler, D., & Pollyea, A. (2019). Development of a gridded reference evapotranspiration dataset for the Great Lakes region. *Journal of Hydrology: Regional Studies, 24*, 100606.

Lang, D., Zheng, J., Shi, J., Liao, F., Ma, X., Wang, W., . . . Zhang, M. (2017). A Comparative Study of Potential Evapotranspiration Estimation by Eight Methods with FAO Penman–Monteith Method in Southwestern China. *Water, 9*(10), 734.

Liu, X., Xu, C., Zhong, X., Li, Y., Yuan, X., & Cao, J. (2017). Comparison of 16 models for reference crop evapotranspiration against weighing lysimeter measurement. *Agricultural Water Management, 184*, 145-155.

López-Urrea, R., de Santa Olalla, F. M., Fabeiro, C., & Moratalla, A. (2006). Testing evapotranspiration equations using lysimeter observations in a semiarid climate. *Agricultural Water Management, 85*(1-2), 15-26.

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

Martel, M., Glenn, A., Wilson, H., & Kröbel, R. (2018). Simulation of actual evapotranspiration from agricultural landscapes in the Canadian Prairies. *Journal of Hydrology: Regional Studies, 15*, 105-118.

Martínez-Rincón, R. O., Ortega-García, S., & Vaca-Rodríguez, J. G. (2012). Comparative performance of generalized additive models and boosted regression trees for statistical modeling of incidental catch of wahoo (Acanthocybium solandri) in the Mexican tuna purse-seine fishery. *Ecological Modelling, 233*, 20-25.

Nemani, R. R., Keeling, C. D., Hashimoto, H., Jolly, W. M., Piper, S. C., Tucker, C. J., . . . Running, S.

W. (2003). Climate-driven increases in global terrestrial net primary production from 1982 to 1999. *Science, 300*(5625), 1560-1563.

Novick, K., Oren, R., Stoy, P., Juang, J.-Y., Siqueira, M., & Katul, G. (2009). The relationship between reference canopy conductance and simplified hydraulic architecture. *Advances in Water Resources, 32*(6), 809-819.

Penman, H. (1963). Vegetation and Hydrology Tech. Comm. No 53. *Commonwealth Bureau of Soils, Harpenden, England, 125p*.

Poeplau, C., Don, A., Vesterdal, L., Leifeld, J., Van Wesemael, B., Schumacher, J., & Gensior, A. (2011). Temporal dynamics of soil organic carbon after land−use change in the temperate zone–carbon response functions as a model approach. *Global Change Biology, 17*(7), 2415-2427.

Priestley, C. H. B., & Taylor, R. (1972). On the assessment of surface heat flux and evaporation using large-scale parameters. *Monthly weather review, 100*(2), 81-92.

Ravazzani, G., Corbari, C., Morella, S., Gianoli, P., & Mancini, M. (2012). Modified Hargreaves-Samani equation for the assessment of reference evapotranspiration in Alpine river basins. *Journal of Irrigation and Drainage Engineering, 138*(7), 592-599.

Sabziparvar, A.-A., & Tabari, H. (2010). Regional estimation of reference evapotranspiration in arid and semiarid regions. *Journal of Irrigation and Drainage Engineering, 136*(10), 724-731.

Sentelhas, P. C., Gillespie, T. J., & Santos, E. A. (2010). Evaluation of FAO Penman–Monteith and alternative methods for estimating reference evapotranspiration with missing data in Southern Ontario, Canada. *Agricultural Water Management, 97*(5), 635-644.

Si, J., Feng, Q., Zhang, X., Liu, W., Su, Y., & Zhang, Y. (2005). Growing season evapotranspiration from Tamarix ramosissima stands under extreme arid conditions in northwest China.

Environmental Geology, 48(7), 861-870.

Sun, G., Alstad, K., Chen, J., Chen, S., Ford, C. R., Lin, G., . . . Miao, H. (2011). A general predictive model for estimating monthly ecosystem evapotranspiration. *Ecohydrology, 4*(2), 245-255.

Tabari, H. (2010). Evaluation of Reference Crop Evapotranspiration Equations in Various Climates. *Water Resources Management, 24*(10), 2311-2337.

Tabari, H., Kisi, O., Ezani, A., & Talaee, P. H. (2012). SVM, ANFIS, regression and climate based models for reference evapotranspiration modeling using limited climatic data in a semi-arid highland environment. *Journal of Hydrology, 444*, 78-89.

Tabari, H., & Talaee, P. H. (2011). Local Calibration of the Hargreaves and Priestley-Taylor Equations for Estimating Reference Evapotranspiration in Arid and Cold Climates of Iran Based on the Penman-Monteith Model. *Journal of Hydrologic Engineering, 16*(10), 837-845.

Temesgen, B., Allen, R., & Jensen, D. (1999). Adjusting temperature parameters to reflect well-watered conditions. *Journal of Irrigation and Drainage Engineering, 125*(1), 26-33.

Trajkovic, S. (2009). Comparison of radial basis function networks and empirical equations for converting from pan evaporation to reference evapotranspiration. *Hydrological Processes: An International Journal, 23*(6), 874-880.

Trajkovic, S., Gocic, M., Pongracz, R., & Bartholy, J. (2019). Adjustment of Thornthwaite equation for estimating evapotranspiration in Vojvodina. *Theoretical and Applied Climatology, 138*(3-4), 1231-1240.

Valipour, M. (2015). Retracted: Comparative Evaluation of Radiation-Based Methods for Estimation of Potential Evapotranspiration. *Journal of Hydrologic Engineering, 20*(5), 04014068.

Ventura, F., Spano, D., Duce, P., & Snyder, R. L. (1999). An evaluation of common evapotranspiration equations. *Irrigation ence, 18*(4), 163-170.

Xu, C.-Y., & Singh, V. (2002). Cross comparison of empirical equations for calculating potential evapotranspiration with data from Switzerland. *Water resources management, 16*(3), 197-219.

Yoder, R., Odhiambo, L. O., & Wright, W. C. (2005). Evaluation of methods for estimating daily reference crop evapotranspiration at a site in the humid southeast United States. *Applied engineering in agriculture, 21*(2), 197-202.

Zhang, F., Li, H., Wang, W., Li, Y., Lin, L., Guo, X., . . . Cao, G. (2018). Net radiation rather than surface moisture limits evapotranspiration over a humid alpine meadow on the northeastern Qinghai−Tibetan Plateau. *Ecohydrology, 11*(2), e1925.

Zou, D., Zhao, L., Yu, S., Chen, J., Hu, G., Wu, T., . . . Pang, Q. (2017). A new map of permafrost distribution on the Tibetan Plateau. *The Cryosphere, 11*(6), 2527.