Comparison of 13 models of reference evapotranspiration with large weighing lysimeter measurements in a humid alpine meadow, northeastern Qinghai-Tibetan Plateau

Licong Dai[1], Xiaowei Guo[1], Xun Ke[1], Yuting Lan[1], Yangong Du[1], Fawei Zhang[1], Yikang Li[1], and Guangmin Cao[1]

[1]Northwest Institute of Plateau Biology Chinese Academy of Sciences

April 28, 2020

Abstract

Accurate estimates of evapotranspiration (ET) are of great importance for water balance and energy exchange processes, as ET constitutes the key component of the terrestrial water cycle. Although many applicable reference evapotranspiration (ET0) models have been developed to estimate the ET, these are largely focused on low altitude regions, with little attention to alpine meadow. In this paper, we evaluate the performance of 13 ET0 models by comparison with large weigh lysimeter measurements. Specifically, we use three combination models, seven radiation-based models and three temperature-based models driven with data from 8 June 2017 to 18 September 2018 in a humid alpine meadow, northeastern Qinghai-Tibetan Plateau. The daily ET was also obtained by large weighing lysimeters located in an alpine *Kobresia* meadow. Results show that the performances of the 13 ET0 models, ranked on the basis of their RMSE (root mean square error), decreased in the order: DeBruin-Keijman>Priestley-Taylor> 1963 Penman> FAO-24 Penman>Hargreaves>Hargreaves2>Hargreaves1>IRMAK1>FAO-56Penman-Monteith>Makkink>Makkink (1967)>Makkink (1957)>IRMAK2. Overall, the radiation-based models performed best, with RMSEs ranging from1.03 to 1.47 mm d-1 and averaging 1.09 mm d-1, followed by the combination models (RMSE from 1.19 to 1.36 mm d-1 and averaging 1.26 mm d-1) and temperature models (RMSE from 1.28 to 1.32 mm d-1 and averaging 1.29 mm d-1). The best radiation-based model (DeBruin-Keijman) was more accurate than the best combination model (1963 Penman) and temperature model (Hargreaves) by 16.67% and 25.49%, respectively. The better performance of the radiation-based models over the other two types may be attributed to their inclusion of the dominant factors affecting ET, such as net radiation (Rn). All models tended to underestimate measured ET during periods of larger evaporative demand (i.e. growing season) and overestimate measured ET during lower evaporative demand (i.e. non-growing season). Our results could help in the selection of a suitable ET model for alpine ecosystems, thereby benefitting water irrigation management.

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, accurate estimates of evapotranspiration could improve water management strategies and promote the efficient use of water resources, especially in regions suffering water shortages (Ljungqvist et al., 2016; 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; Liu et al., 2017) and the eddy covariance technique (Gu et al., 2008; Novick et al., 2009). 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 (Doorenbos, 1975; Hargreaves and Samani, 1985), temperature-based models (Trajkovic, 2005; Valipour & Eslamian, 2014), and combination models (Monteith, 1965; 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 the together with large number of models also adds confusion and arbitrariness to model selection, because most models have not been evaluated against lysimeter measurements across a range of regions and climates. 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–MonteithFAO-56 combination equation (PM-56) as the standard equation for estimating reference evapotranspiration (Allen, Pereira, Raes, & Smith, 1998), and this has been widely used worldwide when compared with other equations (Cai, Liu, Lei, Pereira, & Luis, 2007). The advantages of the Penman–Monteith equationare 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, Yu-Zhong, Liu, Zhong, & Zhao, 2015; MartıNez-Cob & Tejero-Juste, 2004), with few being tested against lysimeter measurements (Jensen, Burman, & Allen, 1990; Ventura, Spano, Duce, & Snyder, 1999; Yoder, Odhiambo, & Wright, 2005). 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 (Hossein Tabari, Kisi, Ezani, & Talaee, 2012), especially in developing countries where have limited number of meteorological stations. Thus, it is essential to develop an relatively accurate reference evapotranspiration equation that requires fewer meteorological parameters, to allow more simplified estimates of ET than those of PM-56, applicable across a range of climatic conditions (Hossein Tabari, 2010; H. 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, Aeini, & Ghafouri, 2010; Hossein 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 and alpine grassland across the QTP 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 (Liu et al., 2017). In this study, we compared 13 reference evapotranspiration models against weighing lysimeter measurements on the northern Tibetan Plateau, with the aims of selecting the best fit model in applications in this region 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, 101deg19'E) (Fig.1a), which is situated on the Northeastern QTP at an elevation of 3200 m a.m.s.l. This area is characterized by a plateau continental monsoon climate, with well-developed seasonally frozen ground (Dai, Guo, Du, et al., 2019). The average annual air temperature is –1.7degC, with the maximum monthly temperature

2

in July (10.1degC) and minimum monthly temperature in January (–15.0degC). The annual precipitation is about 580 mm, of which 80% falls in the growing season (i.e. from May to September). The average annual pan evaporation is approximately 1191.4 mm (Zhang et al., 2018). The soil at the study site is classified as Mat-Gryic Cambisol, which belongs to a clay loam, and has a thickness of approximately 60 – 80 cm (Dai et al., 2019). The vegetation 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 daily actual ET was measured by large-scale weighing lysimeters (height 2 m, diameter 1 m, and resolution 0.01 mm) (Fig.1c), half-hourly averaged ET was recorded with a data logger (CR1000, Campbell, USA). 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 324 days of data spanning 8 June 2017 to 19 September 2018. According to in-situ phenological observations of the foliage of dominant plants, we defined the growing season as that from May 1st to 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 ET0 using the various models were obtained or estimated from the weather station at Haibei Station, and included relative humidity (RH), wind speed (WS), net radiation (Rn), total radiation (Rs), extraterrestrial solar radiation (Ra), soil heat flux (G), maximum air temperature (Tmax), vapor pressure deficit (VPD), minimum air temperature (Tmin) 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 ET0. 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).

2.4 Reference evapotranspiration models

A total of 13 often-used ET0 models were selected for comparison, including 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 1.

2.5 Evaluation criteria

Here, the ET measured by the large-scale weighing lysimeters was the actual ET, and the performances of theET0 equations were compared to these lysimeter system estimates on a daily time step. The pairwise comparisons were conducted using general linear regression. For further comparison, the root mean squared error (RMSE), percentage error of estimate (PE), mean absolute error (MBE) and coefficient of determination ($R2$) were used for the evaluation of reference evapotranspiration models. The RMSE, PE, MBE and$R2$ are defined as:

3

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

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

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

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

where Pi are the predicted values, Oi are observed values, and are the average of Pi and Oi, 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 2.62±0.12 mm d-1, with average daily measured ET of 4.17±0.13 mm and 0.65 ± 0.06 mm d-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, except for VPD, with the maximum and minimum values in the growing season and non-growing season. The average daily Rn, Rs and Ta during the growing season were 11.07 ± 0.27 MJ m-2, 20.80 ± 0.38 MJ m-2 and 10.18±0.23 , respectively (Fig.2b,c and d). The relative humidity peaked in August (80%), with an average of 62.91% over the whole study period (Fig.2f). The average daily VPD peaked in May (0.65 kPa), with an average of 0.45 kPa over the whole study period (Fig.2e).

3.2 Comparison of daily ET between reference evapotranspiration models and lysimeter mea-

surements during the study period

The comparison of 13 reference evapotranspiration equations against the lysimeter measurements is presented in Fig.3 and Table 2, showing that the relationship between daily ET0 calculated by the 13 reference evapotranspiration equations and lysimeter measurements are each significantly (P <0.01), with higher coefficients of determination (R 2) ranging from 0.61 to 0.81. 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. The daily estimates of combination models and radiation-based models (except Makkin and Makkin (1967)) generally underestimated the ET values measured by lysimeter: the mean difference in ET between lysimeter measurements and model estimates (i.e. MBE) ranged from -0.47 mm d-1 to -0.04 mm d-1 for combination models, and ranged from -0.28 mm d-1 to -0.02 mm d-1 for radiation-based models. However, the temperature-based models generally overestimated the ET values measured by lysimeter, with MBEs ranging from 0.10 mm d-1to 0.39 mm d-1. Overall, the combination models and radiation-based models, on average, underestimated ET by 12.79% and 5.93%, respectively, but the temperature-based models overestimated ET by an average of 7.99%.

During the whole study period, the RMSE of combination models ranged from 1.19 to 1.36 mm d-1 and averaged 1.26 mm d-1, the RMSE for radiation-based models ranged from 1.03 to 1.47mm d-1and averaged 1.09 mm d-1, and the RMSE for temperature-based models ranged from 1.28 to 1.32 mm d-1 and averaged 1.29 mm d-1. Based on the RMSE, the performances of the 13 reference evapotranspiration models decreased in the order: DK>PT>Pen-63>FAO-24 Pen>HAR>HAR2>HAR1>IRMAK1>FAO-56PM >Makkink>Makkink(1967)>Makkink(1957)>IRMAK2. The best model (DK) was, respectively, 49% and 33% more accurate than the poorest (BR) 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 radiation-based models yielded the best performance, followed by the combination 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 ET0 calculated by 13 reference evapotranspiration equations was significantly correlated with the lysimeter measurements (P <0.01), with$R$ 2 ranging from 0.32 to 0.60 (Fig. 4 and Table 3). 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, Makkink (1967), Makkink (1957), IRMAK1 and IRMAK2. It should be noted that the Makkink, Makkink (1967) and Makkink (1957) have the same$R$ 2. Of the temperature-based models, HAR and HAR2 obtained the highest R 2, followed by HAR1. Interestingly, all models (except for HAR1) generally underestimated ET during the growing season: values of MBE ranged from -0.88 to -0.03 mm d-1 and averaged -0.63 mm d-1 over all 13 models, with FAO-56 PM having the largest underestimate (by 39.75%) and HAR2 the minimum underestimate (by 0.68%) (Table 3).

The RMSE for combination models ranged from 1.31 to 1.67 mm d-1 and averaged 1.48 mm d-1, the RMSE for radiation-based models ranged from 1.19 to 1.56 mm d-1 and averaged 1.40 mm d-1, and the RMSE for temperature-based models ranged from 1.38 to 1.41 mm d-1 and averaged 1.40 mm d-1 (Table 3). Based on the RMSE values, the performances of the 13 reference evapotranspiration models follow the order: DK > PT > Makkink (1967) > IRMAK2 > Pen-63 > HAR2 > HAR > HAR1 > Makkink > FAO-24 Pen > IRMAK1 > FAO-56. Evidently, the best (DK) was 40% more accurate than the poorest (FAO-56). Overall, the radiation-based models demonstrated the best performance during the growing season, followed by temperature-based models and finally 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 ET0 calculated by the 13 reference evapotranspiration equations was also significantly correlated with the lysimeter measurements ($P<0.01$), but with lower coefficients of determination ($R2$) ranging from 0.16 to 0.43 (Fig. 5 and Table 4). Of the combination models, FAO-24 Pen obtained the highest $R2$, followed by FAO-56 PM and Pen-63. Of the radiation-based models, PT obtained the highest $R2$, followed by DK, Makkink (1967), Makkink (1957), Makkink, IRMAK1 and IRMAK2. Of the temperature-based models, HAR1 and HAR2 obtained the highest $R2$, then followed by HAR. Interestingly, all model generally overestimated the ET values measured by lysimeter during the non-growing season, with MBEs ranging from 0.30 to 1.20 mm d-1 and averaging 0.69 mm d-1 over the 13 models; Makkink(1967) yielded the largest underestimate (by 64.84%) and PT the minimum underestimate (by 31.37%) (Table 3).

The RMSE for combination models ranged from 0.78 to 1.02 mm d-1 and averaged 0.87 mm d-1, the RMSE for radiation-based models ranged from 0.69 to 1.43 mm d-1 and averaged 1.09 mm d-1, and the RMSE for temperature-based models ranged from 1.08 to 1.19 mm d-1 and averaged 1.13 mm d-1. Based on the RMSE, the performances of the reference evapotranspiration models decreased in the order: PT > DK > FAO-24Pen > FAO-56 > Pen-63 > IRMAK1 > HAR > Makkink (1957) > HAR2 > HAR1 > Makkink > IRMAK2 > Makkink (1967). Evidently, the best (PT) was 93.24% more accurate than the poorest (BR). Overall, the combination models yielded the best performance during the growing season, followed by radiation-based models and temperature-based models.

3.5 Comparison of monthly averaged daily ET0 between reference evapotranspiration models and lysimeter measurements

The estimations of monthly mean daily ET0 for combination models were consistent with the pattern shown by lysimeter measurements (Fig. 6a), with the peak in July being 4.08+-0.16 mm d-1, 3.76+-0.16 mm and 3.50+-0.14 mm d-1 for Pen-63, FAO-24Pen and FAO-56, respectively. As already noted, the combination models underestimated the measurements from May to October, and overestimated the measurements in the other months. Furthermore, Pen-63 yielded the best estimate of daily ET in the peak month, during which it underestimated the measurements by 1.50 mm d-1. The mean daily ET0 for temperature-based models also followed the pattern of lysimeter measurements (Fig. 6b), with the maximum in July reaching 4.60+-0.10 mm d-1, 4.87+-0.09 mm and 4.84+-.11 mm d-1 for HAR, HAR1 and HAR2, respectively. Moreover, the temperature-based models generally overestimated the measured ET during most months (Fig. 6c). HAR1 yielded the most accurate estimate of ET in the peak month, when it underestimated the measurements by only 0.13 mm d-1 on average. The monthly mean daily ET0 for the radiation-based models also peaked in July, at 4.26+-0.19 mm d-1, 4.31+-0.19 mm d-1, 4.04+-0.14 mm d-1, 3.79+-0.13 mm d-1, 4.49+-0.15 mm d-1, 3.75+-0.11 mm and 3.74+-0.13 mm d-1 for PT, DK, Makkink, Makkink (1959), Makkink (1967), IRMAK1 and IRMAK1, respectively. Again, the radiation-based models tended to underestimate the measured ET during the growing season, and overestimate the measured ET during the non-growing season.

3.6 Dominant factors affecting the seasonal variation in lysimeter ET measurements

BRT model indicated that the net radiance was the dominant factor controlling the seasonal variation in measured ET during the whole study period, accounting for 50.34% of total variability, followed by mean air temperature (25.63%), RH (4.96%), VPD (4.91) and Rs (4.33%) (Fig.7a). Factors Ra, G, TD and WS made little contribution to the seasonal variation in measured ET. During the growing season, the net radiance remained the main control of seasonal variation in measured ET (Fig. 7b), accounting for 39.85% of total variability, followed by VPD (14.57%), T (9.37%), Rs (7.91%) and RH (6.45%). However, the seasonal variation of measured ET in non-growing season was dominated by T (Fig. 7c), accounting for 29.35% of total variability, followed by RH (17.28%), Rn (16.33%) and Ra (9.72%).

4 Discussion

4.1 The performance comparison of combination models against lysimeter measurements

Many studies have shown that the Penman family models are generally the most accurate when evaluating ET across various climate scenarios and regions (Liu et al., 2017b; Qiu et al., 2002). Of the penman

models for ET0, the Penman–Monteith FAO 56 has been considered as the standard equation for estimating evapotranspiration (Allen et al., 1998). Contrary to what we expected, the Penman–Monteith FAO 56 was not the best in our study: here, we found that the combination models did not yield the best performance, and instead Pen-63 and FAO-24Pen were more accurate (Pen-63 and FAO-24 Pen had smaller RMSE and higher$R$ 2 than Penman–Monteith FAO 56 during both the growing season and non-growing season). Such similar results have been reported in many other studies (Berengena and Gavilan, 2005; Howell, 1999). For instance, Howell et al. (1998) operated a weighing lysimeter in Texas and found that the earlier 1948 Penman equation performed better than the FAO-56 PM (Berengena and Gavilan, 2005). 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, because it did not yield better accuracy than the other combination models or radiation-based models. The poor performance of FAO-56 PM may associate with measuring error of lysimeter and the lysimeter was not set in a reference environment. It should be noted that reference evapotranspiration calculated the FAO-56 PM was defined as crop height of 0.12 m, albedo of 0.23 and surface resistance of 70 s/m (Allen et al., 1998), such reference environment was often not satisfied in our study site. Therefore, the FAO-56 PM applied without local lysimeter data comparison lead to a great error. There is increasing evidence indicated that the Penman locally adjusted equation showed better performance than FAO-56 PM when comparing hourly and daily ET estimates against weighing lysimeter measurements (Berengena & Gavilan, 2005; Liu et al., 2017). Of course, 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 (Hossein Tabari et al., 2012).

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 and DK models yielded the best performances of the radiation-based models, and were respectively 44.11% and 42.71% more accurate than the poorest (IRMAK2) model, which was in line with the previous study conducted in humid areas that the PT method exert the good accuracy estimate for ET (Trajkovic & Kolakovic, 2009). 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 and DK models which might associated with the use of only the most important meteorological factors affecting ET such as net radiation (Rn), was supported by our results (Fig.6). Compare with PT and DK model, the other radiation models just use the Rs as the mainly driving variable, the ET thus may overestimate because some Rs 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 and DK 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 et al., 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 =1.28 mm d-1, MBE=0.17 mm d-1 and PE=5.92%), which was consistent with previous studies conducted on humid region (Hossein 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, 2011). Overall, the

temperature models display a poor performance compared with radiation models owing to the Hargreaves method was established in semiarid areas (Hossein 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 three type models, we found that the radiation-based models display better performance(small RMSE) when compared with the Penman family models and temperature-based models, average RMSE is 1.09 mm d-1, 1.26 mm d-1 and 1.29 mm d-1 for radiation-based models, combination models and temperature-based models, respectively (Table 2). The better performance of the radiation-based models compared with other two type models might be attributed to the incorporation of important input parameter such as net radiation. Although the combination models also include net radiation parameter, the combination models incorporate much other meteorological factors that not mainly controlling ET except for net radiation, which lead to a poor performance for combination models. In addition, the radiation-based models require less meteorological input compared with combination models, and the net radiation was the dominated factor that controlling the ET for this humid region. 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 (Fooladmand & Haghighat, 2010; 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 also relate to the Ra parameter used in the Hargreaves model (Fontenot, 2004), 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 three 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., 2017b). Furthermore, we found that the measured ET and calculated ET0 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 ET0 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 single lysimeter and a single weather station, which may insufficient to represent the whole humid climate or the alpine ecosystem but represent a specific site. Thus, 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 13 reference crop evapotranspiration models against lysimeter measurements in a humid alpine meadow, northeastern Qinghai-Tibetan Plateau, we found that the radiation-based models performed better than combination models and temperature-based models. In addition, the combination models and most radiation-based models tended to underestimate measured ET, whereas temperature-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-56Penman-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 the DeBruin-Keijman and Priestley-Taylor models, 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.

Acknowledgments

Conflict of interest

The authors declare no conflict of interest.

Date availability statement

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

Hosted file

Figure.docx available at https://authorea.com/users/285507/articles/409120-comparison-of-13-models-of-reference-evapotranspiration-with-large-weighing-lysimeter-measurements-in-a-humid-alpine-meadow-northeastern-qinghai-tibetan-plateau

Hosted file

Table1.docx available at https://authorea.com/users/285507/articles/409120-comparison-of-13-models-of-reference-evapotranspiration-with-large-weighing-lysimeter-measurements-in-a-humid-alpine-meadow-northeastern-qinghai-tibetan-plateau

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.

Berengena, J., & Gavilan, 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. U., Lei, T., Pereira, S., & Luis. (2007). Estimating reference evapotranspiration with the FAO Penman-Monteith equation using daily weather forecast messages. Agricultural & Forest Meteorology, 145(1), 22-35.

Cao, J. F., Yu-Zhong, L. I., Liu, X. Y., Zhong, X. L., & Zhao, Y. M. (2015). Comparison of Four Combination Methods for Reference Crop Evapotranspiration. Chinese Journal of Agrometeorology.

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.

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

DROOGERS, Peter, ALLEN, & Richard, G. (2002). Estimating reference evapotranspiration under inaccurate data conditions. Irrigation & Drainage Systems, 16(1), 33-45.

Ershadi, A., Mccabe, M. F., Evans, J. P., Chaney, N. W., & Wood, E. F. (2014). Multi-site evaluation of terrestrial evaporation models using FLUXNET data. Agricultural & Forest Meteorology, 187(8), 46-61.

Fontenot, R. L. (2004). An evaluation of reference evapotranspiration models in Louisiana.

Fooladmand, H. R., & Haghighat, M. (2010). Spatial and temporal calibration of Hargreaves equation for calculating monthly ETo based on Penman-Monteith method. Irrigation & Drainage, 56(4), 439-449.

Gu, S., Tang, Y., Cui, X., Du, M., Zhao, L., Li, Y., . . . Qi, P. (2008). Characterizing evapotranspiration over a meadow ecosystem on the Qinghai-Tibetan Plateau. Journal of Geophysical Research: Atmospheres, 113(D8).

Hansen, S. (1984). Estimation of Potential and Actual EvapotranspirationPaper 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.

Howell, T. A. (1999). Adjusting Temperature Parameters to Reflect Well-Watered Conditions. Journal of Irrigation & Drainage Engineering, 126(5), 26-33.

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 & 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. A., Allen, R. G., Payero, J. O., & Martin, D. L. (2005). Standardized ASCE Penman-Monteith: impact of sum-of-hourly vs. 24-hour timestep computations at reference weather station sites. Transactions of the American Society of Agricultural Engineers, 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 lysimeter for evapotranspiration research in a humid environment. Transactions of the Asabe, 49(2), 401-412.

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.

Ljungqvist, F. C., Krusic, P. J., Sundqvist, H. S., Zorita, E., Brattstrom, G., & Frank, D. (2016). Northern Hemisphere hydroclimate variability over the past twelve centuries. Nature, 532(7597), 94-98.

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

MartıNez-Cob, A., & Tejero-Juste, M. (2004). A wind-based qualitative calibration of the Hargreaves ET 0 estimation equation in semiarid regions. Agricultural Water Management, 64(3), 251-264.

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(2), 20-25.

Monteith, J. L. (1965). Evaporation and environment. Symp Soc Exp Biol, 19(19), 205-234.

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.

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.

QIU, G. Y., Miyamoto, K., SASE, S., GAO, Y., SHI, P., & YANO, T. (2002). Comparison of the three-temperature model and conventional models for estimating transpiration. Japan Agricultural Research Quarterly: JARQ, 36(2), 73-82.

Ravazzani, G., Corbari, C., Morella, S., Gianoli, P., & Mancini, M. (2011). 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., Aeini, A., & Ghafouri, M. (2010). Evaluation of Class A Pan Coefficient Models for Estimation of Reference Crop Evapotranspiration in Cold Semi-Arid and Warm Arid Climates. Water Resources Management, 24(5), 909-920.

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, 444-445(10), 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. (2005). Temperature-based approaches for estimating reference evapotranspiration. Journal of Irrigation and Drainage Engineering, 131(4), 316-323.

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., & Kolakovic, S. (2009). Evaluation of reference evapotranspiration equations under humid conditions. Water Resources Management, 23(14), 3057.

Valipour, M. (2014). Retracted: Comparative Evaluation of Radiation-Based Methods for Estimation of Potential Evapotranspiration. Journal of Hydrologic Engineering, 20(5), 04014068.

Valipour, M. (2015). Evaluation of radiation methods to study potential evapotranspiration of 31 provinces. Meteorology and Atmospheric Physics, 127(3), 289-303.

Valipour, M., & Eslamian, S. (2014). Analysis of potential evapotranspiration using 11 modified temperature-based models. International Journal of Hydrology Science and Technology, 4(3), 192-207.

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

Xu, C. Y., & Singh, V. P. (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.