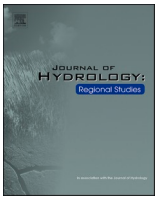
Journal of Hydrology: Regional Studies 50 (2023) 101573

Contents lists available at ScienceDirect

Journal of Hydrology: Regional Studies

journal homepage: www.elsevier.com/locate/ejrh



Performance of potential evapotranspiration models across different climatic stations in New South Wales, Australia

Lijie Shi [a], Bin Wang [b,*], De Li Liu [b,c,**], Puyu Feng [d], James Cleverly [e], Linchao Li [f], Gengxi Zhang [a], Qiang Yu [g,h,***]

[a] College of Hydraulic Science and Engineering, Yangzhou University, Yangzhou, Jiangsu 225009, China
[b] NSW Department of Primary Industries, Wagga Wagga Agricultural Institute, Wagga Wagga, NSW 2650, Australia
[c] Climate Change Research Centre and ARC Centre of Excellence for Climate System Science, University of New South Wales, Sydney, NSW 2052, Australia
[d] College of Land Science and Technology, China Agricultural University, Beijing 100193, China
[e] School of Life Sciences, Faculty of Science, University of Technology Sydney, P.O. Box 123, Broadway, Sydney, NSW 2007, Australia
[f] College of Natural Resources and Environment, Northwest A&F University, Yangling 712100, China
[g] State Key Laboratory of Soil Erosion and Dryland Farming on the Loess Plateau, Northwest A&F University, Yangling, Shaanxi 712100, China
[h] College of Resources and Environment, University of Chinese Academy of Science, Beijing 100049, China

ARTICLE INFO

Keywords:
Potential evapotranspiration
Penman
Model evaluation
Temporal trends
Periodicity

ABSTRACT

Study region: New South Wales, southeast Australia
Study focus: Estimating potential evapotranspiration (ETp) rates, detecting its temporal trends and analysing its interannual oscillation are critical for long-term assessment of water availability and regional drought. This study aimed to evaluate the comprehensive performance of 12 simplified models in characterising ETp against the benchmark Penman model across different climate sites in southeast Australia. This study used Taylor skill score (S), normalised root mean square error (nRMSE) and relative mean bias error (rMBE) to estimate models' capability in estimating ETp rates. Then, this study adopted Mann-Kendall test and continuous wavelet transform (CWT) to test temporal trends and periodicity of ETp estimated by all models.
New hydrology insights for the region: Jensen-Haise model was capable to produce fair (nRMSE $\leq$ 30%) estimates of daily ETp across all stations. Models except Mak1 were generally able to produce reasonable estimates of ETp at larger time scale. In addition, we found that the 12 alternative ETp models generally agreed with the Penman model on the primary (9.6–12.4-year) and quasi (2.6–3.9-year) periods of ETp, but they did not show matchable ability in detecting ETp temporal trends. The comprehensive investigation on models' performance will shed light on models' selection in estimation of drought and hydrological cycle.

* Corresponding author.
** Corresponding author at: NSW Department of Primary Industries, Wagga Wagga Agricultural Institute, Wagga Wagga, NSW 2650, Australia.
*** Corresponding author at: State Key Laboratory of Soil Erosion and Dryland Farming on the Loess Plateau, Northwest A&F University, Yangling, Shaanxi 712100, China.
E-mail addresses: bin.a.wang@dpi.nsw.gov.au (B. Wang), de.li.liu@dpi.nsw.gov.au (D.L. Liu), yuq@nwafu.edu.cn (Q. Yu).

https://doi.org/10.1016/j.ejrh.2023.101573
Received 18 July 2023; Received in revised form 26 October 2023; Accepted 12 November 2023
Available online 17 November 2023

1. Introduction

Evapotranspiration (ET) is a key nexus to water, energy, and nutrients cycles in ecosystems (Allen et al., 1998; Jung et al., 2010; McMahon et al., 2016; Peng et al., 2018). As an explanation, evapotranspiration accounts for 60% (up to 95% in arid regions) of water loss in the hydrological budget (Jung et al., 2010; Kool et al., 2014) and consumes more than 50% of the solar energy absorbed by land surfaces (Trenberth et al., 2009). Actual evapotranspiration (ETa), potential evapotranspiration (ETp), and reference evapotranspiration (ET0) are three related yet slightly different terms in evapotranspiration estimation for different use. In specific, actual evapotranspiration is the total amount of water evaporating from land surface to atmosphere, having significant influence on regional or global water availability (Xiong et al., 2023). It can be estimated by water balance method at regional or global scale or directly measured by lysimeters, eddy covariance flux towers, or satellite-based remote sensing (Xiong et al., 2023; Zhang et al., 2020). Conceptually, ETp is the maximum evapotranspiration from a surface which is not limited by water availability. In other words, it represents for atmospheric evaporative demand under certain climatic conditions and is the maximum possible value of ETa for a evaporative surface (Donohue et al., 2010; Kumar et al., 1987). ETp is an important input to various hydrological models (Oudin et al., 2005; Peng et al., 2016) and drought indexes (Scheff and Frierson, 2015; Sheffield et al., 2012; Xu et al., 2015). Penman model is the most classical model in ETp estimation (Donohue et al., 2010; Shuttleworth, 1993). Similar to ETp, ET0 is the maximum evapotranspiration from a well-watered hypothetical crop surface, which is with an assumed height of 0.12 m, a surface resistance of 70 s m^{-1}, and an albedo of 0.23. ET0 is widely used in calculation of crop water requirement and agricultural irrigation management (Li et al., 2015; Zhang et al., 2010). The most classical model to estimate ET0 is Penman-Monteith model recommended by FAO (Jensen and Allen, 1990). Our study aims to shed light on the option of ETp models as an input for hydrological models or drought indexes. Thus, this study focused on ETp models' assessment.

Due to difficulty in ET measurement, it is commonly to estimate ET by different models, as mentioned above water balance model for ETa estimation, Penman-Monteith model for ET0 estimation, and the classical Penman model for ETp estimation. The accuracy and superiority of Penman model in estimating ETp and capturing its temporal trends have been demonstrated across a broad range of climatic conditions (Donohue et al., 2010; Yang et al., 2019). However, it may be limited to use due to its intensive requirement of climatic data. In this context, various simplified models based on statistical functions between meteorological parameters and ETp were developed to estimate ETp. Contrary to the sound performance of Penman model across various climatic conditions, the

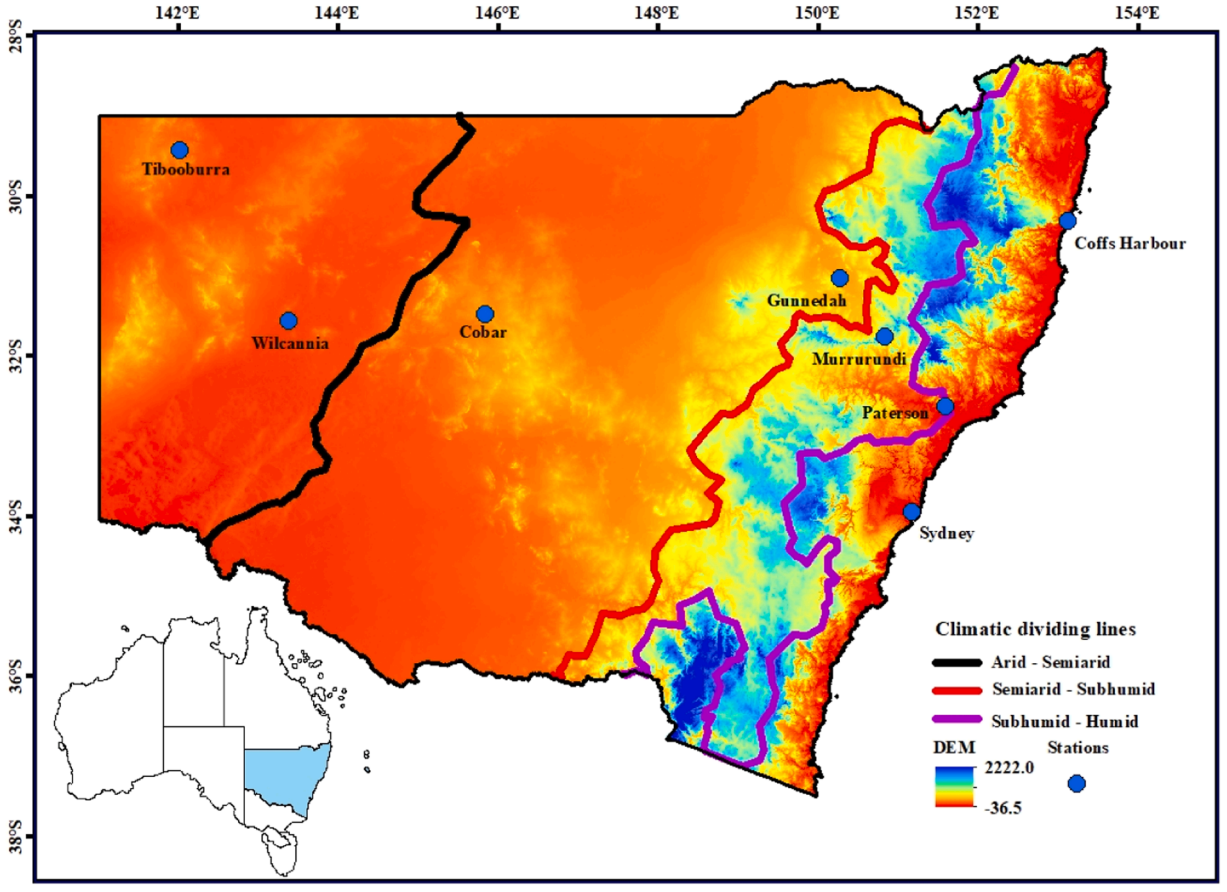


Fig. 1. Spatial distribution of eight weather stations across different climate regimes in NSW.

performance of simplified ETp models is mostly location-depended. Therefore, assessments of their accuracy beyond the conditions where they were originally developed is necessary (Almorox et al., 2015; Tabari, 2009). For instance, Lu et al. (2005) assessed the performance of six ETp models against ETa calculated by water balance equation in the southeast United States. They found that ETp values were significantly different among these models. Based on the research, they recommended to use Priestley-Taylor, Turc, and Hamon models to estimate ETp in that region. Based on three sites in southeast Australia, Azhar and Perera (2011) compared performance of ten ET0 models against ETa measured from standard-grass weighting lysimeters. They reported that combination methods generally produced the most accurate ET estimates. At global scale, Almorox et al. (2015) assessed the performance of 11 temperature-based models for estimating ETp or ET0 across various climatic conditions with Penman-Monteith as the benchmark. They concluded that the Hargreaves-Samani model produced the most accurate global average performance independence of the climate conditions.

However, it is not comprehensive enough to evaluate models' performance based on estimation of ETp rates alone. As well known by researchers, both increasing and decreasing trends of ETp have been reported under a changing climate and influence of human activities on hydrological cycle (Irmak et al., 2012; Peng et al., 2020; Roderick et al., 2007; Xiong et al., 2022). For instance, Xiong et al. (2022) divided ET into five components and found significant increase in global ET driven by climatic factors or human interventions. Similarly, Peng et al. (2020) found that vegetation greenness in China lead to an increase in ET. Ahmadi and Javanbakht (2020) found that ET0 would increase with climate getting drier and warmer in southwest Iran. Furthermore, climatic anomalies (such as extreme temperature or precipitation) often occurred periodically. For example, Li et al. (2019) investigated the spatiotemporal evolution of extreme temperature events across China. They found that besides the increasing/decreasing trends six extreme temperature indices showed periodically oscillation both in historical and future climate scenarios. Given the strong influence of climatic factors and human activities on ETp, it is reasonable to conjecture that the dynamic temporal trends and the interannual oscillation of ETp may continue going on (Liang et al., 2010; Wang et al., 2017). In this context, the ability in capturing ETp trends and the possible periodically oscillation is also expected from ETp models. However, few studies have investigated the periodicity of ETp. Especially, the intercomparison of ability of different ETp models on detecting periodicity of ETp has rarely reported.

In summary, evapotranspiration rates, temporal trends, and interannual oscillation are all important characteristics of ETp (Liang et al., 2010). Comprehensive investigation on models' performance in these perspectives will fill the existing research gap and shed light on models' selection in estimation of drought and hydrological cycle. Thus, we aimed to evaluate the ability of 12 ETp models in the above-mentioned perspectives against the Penman model across eight climatic stations in NSW, Australia.

2. Materials and methods

2.1. Study areas and climate data sets

New South Wales (NSW) is the most populous state in Australia. It is located in southeast Australia (Fig. 1), accounting for 10.4% of the Australian land area (8.1×10^5 km^2). NSW is divided into four distinct geographical sections by natural features: the east coast, the mountains (the Great Dividing Range), the central plains, and the western plains. These diverse geographical features create a varied climate across NSW, namely arid climate at the westernmost area, semi-arid climate at the midwestern plains, sub-humid climate at the mideastern area, and the humid climate along the eastern coastal area. In specific, average annual rainfall increases from 50 mm year^{-1} for the westernmost area to 1500 mm year^{-1} for the eastern coast. Eight stations across NSW were selected as they have complete set of climate data (Fig. 1). The detailed information of these stations was displayed in Table 1.

Daily meteorological data including maximum temperature (T_{max}), minimum temperature (T_{min}), maximum and minimum relative humidity (RH_{max} and RH_{min}, respectively), solar radiation (R_s), and rainfall (P) were obtained from the Scientific Information for Land Owners (https://www.longpaddock.qld.gov.au/silo/datadrill/index.php). The measured wind speed data was obtained from the Bureau of Meteorology (http://www.bom.gov.au/). All these meteorological factors have complete observations in the research period (1970–2014) with the exception of wind speed observation missing less than 5%. For days with missing wind data, we adopted long-term local daily average wind speed for that day to drive Penman model and other wind-involved models. With these data as input,

Table 1
Multi-year (1970–2014) average of air temperature (T), solar radiation (Rs), relative humidity (RH), wind speed (u_2), vapour pressure deficit (VPD), rainfall (P), potential evapotranspiration (ETp), and aridity index (AI) of eight stations in New South Wales and their geographical information. Data used to calculate the averaged meteorological factors, ETp, and AI is from 1970 to 2014.

Stations	Lon (degree)	Lat (degree)	DEM (m)	T (°C)	Rs (MJ m^{-2} day^{-1})	RH (%)	u_2 (m/s)	VPD (kpa)	P (mm)	ETp (mm)	AI (P/ETp†)
Tibooburra (TBR)	142.0	-29.4	183	20.9	20.7	48.2	1.9	1.8	261	2009	0.16
Wilcannia (WCN)	143.4	-31.6	75	19.5	19.7	55.3	2.7	1.6	292	2070	0.17
Cobar (CBR)	145.8	-31.5	260	19.0	19.4	54.6	2.0	1.5	407	1847	0.26
Gunnedah (GND)	150.3	-31.0	307	18.5	18.6	63.2	1.8	1.2	640	1650	0.47
Murrurundi (MRD)	150.8	-31.8	466	15.5	17.5	71.8	1.5	0.9	865	1410	0.74
Paterson (PTS)	151.6	-32.6	30	18.0	16.9	71.5	2.3	0.9	926	1535	0.73
Sydney (SYD)	151.2	-34.0	6	18.2	16.4	67.6	3.2	0.8	1083	1586	0.82
Coffs Harbour (CHB)	153.1	-30.3	5	18.9	17.5	72.3	2.7	0.8	1678	1557	1.28

†P = rainfall; AI = aridity index.

daily ETp was firstly calculated. Then, monthly, seasonal, and annual ETp were the accumulation of daily ETp.

2.2. Estimation of potential evapotranspiration

ETp measured with Pan, lysimeters, or eddy covariance flux is the most accuracy estimation/observation of ETp and supposed to be benchmark for empirical models' assessment (Liu et al., 2017; Martel et al., 2018). However, observation of ETp by these methods is generally expensive, labour-intensive, and high maintenance, thus not always available (Azhar and Perera, 2011; Martel et al., 2018). Among different models, Penman is the most classical one and widely used as benchmark to assess other models' performance. In this study, we also assess the 12 models' (Table 1) ability in ETp estimation, capturing its temporal trends and periodically oscillation.

2.2.1. Penman model

Combining radiative component with aerodynamic transfer, the Penman model is the classic for ETp estimation (Penman, 1948; Valiantzas, 2006). Among numerous variations of Penman model, the one given by Shuttleworth (1993) is simple, accurate, and has been widely used for ETp estimation from water surface (Donohue et al., 2010; Valiantzas, 2006; Zhou et al., 2006), as shown in Eq. (1):

$$ET_{p,Penman} = ET_{pR} + ET_{pA} = \frac{0.408\Delta}{\Delta+\gamma}R_n + \frac{\gamma}{\Delta+\gamma}\frac{6.43(1+0.536u_2)(e_s-e_a)}{\lambda} \quad (1)$$

where ET_{Penman} (mm day^{-1}) is ETp from open water; ET_{pR} (mm day^{-1}) is the radiative component and E_{pA} (mm day^{-1}) is the aerodynamic component. R_n (MJ m^{-2} day^{-1}) is net radiation, which is the algebraic sum of the net short and long wave radiation and could be calculated based on the process of Allen et al. (1998); T (°C) is mean daily air temperature at 2 m height; u_2 (m s^{-1}) is wind speed at 2 m height; e_s (kPa) is saturation vapour pressure; e_a (kPa) is actual vapour pressure; (e_s-e_a) (kPa) is saturation vapour pressure deficit; Δ(kPa °C^{-1}) is the slope of the vapour pressure curve; γ (kPa °C^{-1}) is psychrometric constant; λ is the latent heat of vaporisation of water (=2.45 MJ kg^{-1} at 20 ℃). More details about this equation can be referred to Shuttleworth (1993).

2.2.2. Temperature-based ETp models

Generally, temperature-based models rely on the reliable assumption that temperature is an indicator of the evaporative power of the atmosphere (McKenney and Rosenberg, 1993). Good performance of temperature-based models has been reported in literature. For instance, Tabari (2009) claimed that Hargreaves (HS) was able to estimate ETp accurately in various climates except humid climate. Meanwhile, obvious underestimation of ETp by HS under dry and windy regions was also common (Hargreaves and Allen, 2003). Therefore, we also investigated the performance of temperature-based models including Hargreaves (HS) (Droogers and Allen, 2002), Schendel (Sc) (Djaman et al., 2015; Schendel, 1967), and Ivanov (Iv) (Valipour et al., 2017) in this study.

2.2.3. Radiation-based ETp models

Radiation-based models adopt solar radiation companied with air temperature to estimate ETp based on energy balance (Muniandy et al., 2016; Xu and Singh, 2000). The commonly used radiation-based models including Jensen-Haise (JH) (Jensen and Haise, 1963), Priestley-Taylor (PT) (Priestley and Taylor, 1972), Makkink (Mak1) (Makkink, 1957), modified Makkink (Mak), Abtew (Ab) (Abtew, 1996), and Turc (Tu) (Turc, 1961) were adopted in this study. Among them, PT, Mak1 and Mak are simplifications of Penman model. PT was originally developed to calculate ETp from a saturated land surface or an open water surface under conditions of minimal advection (Priestley and Taylor, 1972). Mak1 model was developed under temperate humid conditions. The difference between Mak1 and PT is that Mak1 requires the incoming solar radiation whereas PT requires net radiation as input. Jensen-Haise (JH) model was developed based on numerous evapotranspiration observations by soil sampling (Jensen and Haise, 1963; Zhang et al., 2018). Turc (Tu) was developed under general climatic conditions of western Europe (Xu and Singh, 2000).

2.2.4. Mass transfer-based models

The mass transfer-based models are generally developed based on Dalton's gas Law (Tabari et al., 2013) to estimate evaporation from free water surface. They adopt the aerodynamic concept of water vapour movement from the evaporating surface to the air to estimate ETp (Muniandy et al., 2016). Three commonly used mass transfer-based models, WMO (Valipour et al., 2017), Mahringer (Mah) (Mahringer, 1970), and Trabert (Tra) (Valipour et al., 2017) were studied in this research.

2.3. Models' performance in estimating ETp rates

This study adopted Taylor skill score (S), normalised root mean square error (nRMSE), and relative mean bias error (rMBE) to assess the performance of 12 alternative ETp models against Penman model. The S combines correlation coefficient (R) and standard deviation (σ) into one index to comprehensively evaluate model performance (Taylor, 2001; Wang et al., 2015). The nRMSE is a powerful index to measure the relative difference of ETp calculated by alternative models versus Penman-calculated ETp. Performance of models is considered excellent if nRMSE is lower than 10%; good when it is higher than 10% but lower than 20%; fair with nRMSE between 20% and 30%; poor with nRMSE higher than 30% (Dettori et al., 2011; Nouri and Homaee, 2018). The rMBE is a useful index to evaluate model's bias and systematic error (Nouri and Homaee, 2018). The positive (negative) values of rMBE represent the model's tendency to overestimate (underestimate) ETp relative to Penman model. A high-performing model will have high values of S with

values of rMBE and nRMBE close to 0%.

The mathematical equations of these statistical indexes are as following:

$$S = \frac{4(1+R)^2}{\left(\frac{\sigma_M}{\sigma_{Penman}} + \frac{\sigma_{Penman}}{\sigma_M}\right)^2 (1+R_0)^2}, \tag{2}$$

where S is the Taylor skill score; R is the correlation coefficient between an alternative model and Penman model; R_0 is the maximum correlation coefficient attainable (0.999 is used in this study). $\sigma_M$and σ_{Penman} are the standard deviations of ETp for an alternative model and Penman model, respectively;

$$nRMSE = \frac{100}{\overline{ET_{Penman}}} \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(ET_{M,i} - ET_{Penman,i}\right)^2}, \tag{3}$$

$$rMBE = \left(\frac{100}{\overline{ET_{Penman}}}\right) \frac{1}{n}\sum_{i=1}^{n}\left(ET_{M,i} - ET_{Penman,i}\right) \tag{4}$$

where $ET_{M,i}$ and $ET_{Penman,i}$ are ETp calculated with an alternative model and Penman model, respectively. $\overline{ET_{Penman}}$ is the average Penman-calculated ETp and n is the number of the samples.

2.4. Temporal trends and periodic analysis of ETp

The Mann-Kendall test (MK test) is widely used in trend detection. One of the most outstanding advantages of MK test is that it is rank-based (non-parametric), which means it is not affected by the actual distribution of the original data even when significantly skewed with some outliers (Hu et al., 2020). Various studies showed that MK test is a powerful tool in analysing the seasonal and annual trends in climate data (such as rainfall and temperature). Companied with Sen's estimator (also non-parametric), both trends and slope magnitudes of rainfall, temperature, runoff, evapotranspiration, and other hydrological data were widely documented in literature (Douglas et al., 2000; Fan et al., 2016; Peng et al., 2017; Tabari et al., 2011). In fact, MK test is highly recommended by the World Meteorological Organization to identify trends in hydro-meteorological time series. Thus, this study also adopted MK test to identify temporal trends of climatic factors and ETp.

The temporal trend of data series can be detected by the standardised test statistic (Z) in M-K test. Generally, three different confidence intervals were commonly used in MK test to identify whether the temporal trend is significant or not. That is $|Z| > 1.65$ at 90% confidence level, $|Z| > 1.96$ at the 95% confidence level, and $|Z| > 2.58$ at 99% confidence level. A positive Z-value shows an increasing trend, while a negative Z-value indicates a decreasing trend. To capture the temporal trends of climatic factors and ETp as thorough as possible, this study adopted all three confidence intervals to test their trends significance.

The statistical value, S_v, and the standardised test statistic, Z, are defined by equations from (5) to (8) (Han et al., 2018; Pohlert, 2016):

$$S_v = \sum_{k=1}^{n-1}\sum_{j=k+1}^{n} \mathrm{sgn}\left(x_j - x_k\right), \tag{5}$$

with

$$\mathrm{sgn}(x) = \begin{cases} 1 & if \ \ x > 0 \\ 0 & if \ \ x = 0 \\ -1 & if \ \ x < 0 \end{cases}, \tag{6}$$

with the variance of S_v

$$\sigma^2 = \left\{ n(n-1)(2n+5) - \sum_{j=1}^{p} t_j\left(t_j - 1\right)\left(2t_j + 5\right) \right\} \Big/ 18, \tag{7}$$

$$Z = \begin{cases} \dfrac{S_v - 1}{\sigma} & if \quad S_v > 0 \\ 0 & if \quad S_v = 0 \\ \dfrac{S_v + 1}{\sigma} & if \quad S_v < 0 \end{cases}, \tag{8}$$

where x_j and x_k are two sequential values of the variable, n is the length of the data sequence, p is the number of tied groups, t_j is the number of data values in the jth group. In addition to the M-K test, Sen's slope estimator test was applied to calculate the magnitude of the ETp trend. The slope β (Gao et al., 2017; Pohlert, 2016) is calculated as Eq. (9):

$$\beta = Median\left(\frac{x_j - x_k}{j-k}\right), \qquad 1 < k < j, \tag{9}$$

Continuous wavelet transform (CWT) is a powerful tool for analysing the periodic oscillations of climate anomalies (e.g. temperature) (da Silveira and Pezzi, 2014; Torrence and Compo, 1998) and natural hazards (e.g. droughts) (Özger et al., 2009; Wang et al., 2017). In this method, signals were decomposed into wavelet coefficients, which localised in both time and frequency due to dilation and translation of a mother wavelet. We used the Morlet wavelet as the mother wavelet in this study. The Morlet wavelet is a plane wave modified by a Gaussian, having a zero mean and providing a balance between time and frequency localisations (Zhang et al., 2007). It is defined as:

$$\psi_0(t) = \pi^{-1/4} e^{i\omega_0 t} e^{-t^2/2}, \tag{10}$$

where ω_0, the dimensionless angular frequency, was set to 6 to render the Morlet wavelet analytical (Angi and Harald, 2014; Farge, 1992).

2.5. Sensitivity analysis of ETp to meteorological factors

To detect which meteorological factor have the greatest influence on ETp, this study calculated the sensitivity coefficient of T, Rn, RH, and wind speed (u_2) with the method proposed by McCuen (1974), as shown in Eq. (11). This method was based on partial derivation and widely used in sensitivity analysis of ETp to meteorological factors (She et al., 2017; Yang et al., 2019).

$$SC_{v_i} = \lim_{v_i \to 0}\left(\frac{\Delta ET_p / ET_p}{\Delta v_i / v_i}\right) = \frac{\partial ET_p}{\partial v_i} \cdot \frac{v_i}{ET_p}, \tag{11}$$

	TBR			WCN			CBR			GND			
	S	nRMSE	rMBE	S	nRMSE	rMBE	S	nRMSE	rMBE	S	nRMSE	rMBE	
JH	0.90	24	-2	0.89	29	-14	0.94	20	-7	0.90	25	-4	JH
Ab	0.87	30	-19	0.83	38	-27	0.92	30	-23	0.88	32	-21	Ab
Tu	0.89	26	-12	0.82	36	-25	0.92	26	-17	0.82	31	-19	Tu
PT	0.78	42	-34	0.72	48	-39	0.83	39	-32	0.85	34	-24	PT
Mak	0.73	35	-24	0.67	44	-31	0.79	34	-24	0.78	33	-20	Mak
Mak1	0.64	46	-36	0.57	53	-42	0.70	45	-36	0.70	43	-33	Mak1
HS	0.81	33	-20	0.79	36	-22	0.88	28	-20	0.83	29	-13	HS
IV	0.70	55	31	0.81	36	4	0.79	40	16	0.81	33	-1	IV
WMO	0.68	46	4	0.66	53	9	0.74	42	-3	0.78	44	-22	WMO
Mah	0.59	62	25	0.61	64	26	0.63	58	18	0.74	44	-12	Mah
Tra	0.54	73	35	0.56	75	35	0.58	67	27	0.69	48	-5	Tra
Sc	0.30	129	70	0.57	68	25	0.43	94	42	0.77	40	13	Sc
JH	0.94	20	-9	0.85	30	-9	0.90	26	-14	0.91	20	-6	JH
Ab	0.94	25	-19	0.82	38	-25	0.84	41	-35	0.88	33	-27	Ab
Tu	0.89	27	-20	0.74	36	-22	0.81	34	-26	0.84	27	-20	Tu
PT	0.93	26	-19	0.77	41	-26	0.84	40	-31	0.89	30	-22	PT
Mak	0.88	26	-16	0.73	38	-23	0.79	37	-28	0.85	28	-21	Mak
Mak1	0.81	37	-30	0.65	48	-36	0.71	47	-40	0.78	40	-34	Mak1
HS	0.92	21	0	0.76	36	-12	0.73	41	-30	0.72	34	-24	HS
IV	0.77	38	-25	0.72	40	-24	0.77	34	-16	0.66	40	-25	IV
WMO	0.83	49	-39	0.77	48	-25	0.79	37	-16	0.80	41	-29	WMO
Mah	0.84	40	-25	0.79	41	-14	0.77	36	-3	0.79	34	-16	Mah
Tra	0.82	39	-19	0.75	43	-8	0.74	40	4	0.77	33	-10	Tra
Sc	0.80	31	-7	0.71	34	-1	0.78	30	4	0.66	28	0	Sc
	S	nRMSE	rMBE	S	nRMSE	rMBE	S	nRMSE	rMBE	S	nRMSE	rMBE	
	MRD			PTS			SYD			CHB			

Fig. 2. The Taylor skill score (S), nRMSE (%), and rMBE (%) between daily ETp estimated by tested models and Penman in the research period (1970–2014) for eight stations in NSW. The pink colour filled panel of S represents for S ≥ 0.60; The pink colour filled panel of nRMSE represents for nRMSE ≤ 10%; The yellow colour filled panel of nRMSE represents for 10% < nRMSE ≤ 20%; The green colour filled panel of nRMSE represents for 20% < nRMSE ≤ 30%. The pink colour filled panel of rMBE represents for absolute rMBE ≤ 10%; The yellow colour filled panel of rMBE represents for 10% < absolute rMBE ≤ 20%; The green colour filled of rMBE represents for 20% < absolute rMBE≤ 30%.

where SC_{vi} is the sensitivity coefficient of the ith meteorological factor v. It is dimensionless, thus convenient for the comparison among different factors. When SC_{vi} is positive, it indicates that ETp increases with the increase of meteorological factor v_i. On the contrary, a negative SC_{vi} means that ETp decreases with the increase of the variable. Meanwhile, the larger absolute SC_{vi} is, the greater influence of meteorological factor has on ETp. Given that only Penman model includes all the four meteorological factors, thus this study only did sensitivity analysis based on Penman model.

3. Results

3.1. Performance of models in estimating ETp rates at various time scales

This study assessed models' performance in the estimation of ETp at daily (Fig. 2), monthly (Figs. S1–S4), seasonal (Figs. 3 & 5), and annual (Figs. 4 & 6) scales. The results showed that though S of all models were larger than 0.60 (indicating high correlation with Penman model), only model JH, whose S was larger than 0.90, nRMSE ranged from 20% to 30%, and rMBE ranged from − 2% to − 14%, was able to produce good or fair estimation in daily ETp across all eight stations (Fig. 2). Followed model JH, models including Tu, Ab, PT, Mak, HS, and Sc were able to produce fair estimation of daily ETp at two to four stations out of eight stations. For instance, model Sc greatly overestimated daily ETp at arid stations and got its best performance at two humid stations (with nRMSE slightly lower than 30%). On the contrary, model HS generally performed better at semi-arid and sub-humid stations than it did at other stations.

Most models' performance was improved at monthly (Figs. S1–S4) and seasonal scales (Figs. 3 & 5). As seasonal ETp was the accumulation of monthly ETp in that season, models' performance at monthly scale generally unified with that in the corresponding seasons. Thus, the results focused on the description of models' performance at seasonal scale. Generally speaking, all radiation-based models (JH, Ab, Tu, Mak, and Mak1) and temperature-based HS (temperature-based) tended to underestimate monthly and seasonal ETp whereas mass transfer-based models (WMO, Mah, and Tra) and temperature-based IV and Sc generally overestimated seasonal ETp at arid and semi-arid stations but underestimated it at sub-humid and humid stations. Another common pattern shared by radiation-based models and HS was that their performance was better in estimation of summer and autumn ETp than they did in winter and spring whereas mass transfer-based models did not show unified seasonal pattern in their performance (Figs. 3 & 5). Radiation-based models and HS were developed based on temperature and radiation (Rs or Rn), which were generally simplified as linear

		TBR			WCB			CBR			GND			MRD			PTS			SYD			CHB		
		S	nRMSE	rMBE	S	nRMSE	rMBE	S	nRMSE	rMBE	S	nRMSE	rMBE	S	nRMSE	rMBE	S	nRMSE	rMBE	S	nRMSE	rMBE	S	nRMSE	rMBE
Spring	JH	0.63	11	-8	0.57	22	-20	0.79	14	-13	0.88	9	-8	0.79	16	-15	0.53	18	-16	0.53	20	-19	0.76	11	-11
	Ab	0.59	24	-23	0.57	31	-31	0.72	27	-26	0.87	23	-23	0.85	23	-22	0.56	29	-28	0.42	38	-37	0.74	30	-30
	Tu	0.71	17	-15	0.66	30	-29	0.91	20	-19	0.85	21	-21	0.61	24	-23	0.40	28	-27	0.41	29	-28	0.58	24	-23
	PT	0.09	35	-33	0.09	39	-38	0.13	32	-31	0.34	24	-22	0.36	20	-18	0.18	28	-26	0.14	31	-29	0.46	23	-22
	Mak	0.36	28	-27	0.32	35	-34	0.46	28	-26	0.62	22	-21	0.57	20	-19	0.33	28	-27	0.27	31	-30	0.57	24	-24
	Mak1	0.30	39	-38	0.26	45	-44	0.38	39	-38	0.52	34	-34	0.48	33	-32	0.27	40	-39	0.21	42	-41	0.49	36	-36
	HS	0.30	24	-22	0.36	26	-25	0.40	23	-22	0.65	16	-15	0.67	8	-4	0.42	17	-15	0.21	33	-31	0.49	26	-26
	IV	0.50	24	19	0.68	13	-8	0.68	12	5	0.64	15	-11	0.90	33	-33	0.81	31	-30	0.90	23	-23	0.67	31	-30
	WMO	0.54	14	0	0.57	13	3	0.57	17	-10	0.68	29	-27	0.77	41	-40	0.68	26	-24	0.64	22	-18	0.75	30	-29
	Mah	0.51	24	19	0.56	22	17	0.45	22	10	0.59	21	-17	0.72	29	-27	0.69	18	-14	0.64	13	-7	0.72	18	-17
	Tra	0.46	33	28	0.51	30	26	0.41	28	18	0.54	17	-11	0.66	24	-22	0.64	14	-8	0.59	13	0	0.67	13	-11
	Sc	0.18	66	58	0.40	20	9	0.31	42	32	0.56	13	2	0.85	21	-20	0.73	16	-14	0.91	8	-7	0.64	13	-12
Summer	JH	0.72	7	6	0.49	9	-6	0.71	6	0	0.90	5	3	0.89	3	1	0.82	5	1	0.62	9	-7	0.73	4	2
	Ab	0.68	16	-15	0.46	23	-22	0.64	19	-18	0.88	17	-16	0.88	14	-13	0.79	18	-17	0.50	31	-30	0.71	25	-24
	Tu	0.72	12	-11	0.52	24	-23	0.75	18	-17	0.76	22	-22	0.75	20	-20	0.56	22	-22	0.43	27	-27	0.57	21	-21
	PT	0.16	30	-30	0.08	35	-34	0.17	29	-28	0.53	19	-18	0.67	12	-12	0.47	15	-14	0.35	21	-20	0.51	12	-11
	Mak	0.47	29	-28	0.22	35	-34	0.36	29	-28	0.65	24	-24	0.74	19	-19	0.54	23	-23	0.40	29	-28	0.57	22	-22
	Mak1	0.39	39	-39	0.18	45	-44	0.29	40	-39	0.56	36	-35	0.66	32	-32	0.45	35	-35	0.33	40	-39	0.49	34	-34
	HS	0.41	19	-18	0.34	21	-20	0.41	19	-18	0.61	11	-10	0.74	6	4	0.63	6	-3	0.22	28	-27	0.49	22	-22
	IV	0.35	39	37	0.49	16	11	0.61	21	19	0.57	10	-4	0.81	29	-29	0.90	32	-32	0.82	29	-29	0.81	39	-39
	WMO	0.40	23	19	0.38	28	23	0.38	19	9	0.60	20	-18	0.67	39	-38	0.62	32	-31	0.56	23	-21	0.60	36	-36
	Mah	0.33	46	44	0.36	46	43	0.29	39	33	0.52	12	-6	0.64	25	-23	0.52	22	-19	0.60	13	-10	0.58	26	-25
	Tra	0.30	57	55	0.32	57	54	0.26	49	43	0.48	11	1	0.59	20	-18	0.47	18	-13	0.55	11	-3	0.53	21	-20
	Sc	0.09	98	93	0.21	45	39	0.27	58	53	0.53	14	10	0.78	11	-10	0.84	8	-7	0.76	9	-8	0.75	12	-11
Autumn	JH	0.75	5	0	0.31	16	-13	0.60	10	-6	0.76	6	-3	0.83	8	-7	0.71	6	-2	0.37	15	-10	0.70	5	-1
	Ab	0.76	18	-18	0.40	27	-26	0.56	24	-22	0.73	21	-21	0.88	17	-17	0.73	21	-20	0.25	34	-32	0.69	25	-25
	Tu	0.76	12	-11	0.40	25	-24	0.60	18	-16	0.66	18	-17	0.71	15	-14	0.53	15	-14	0.25	23	-21	0.58	16	-15
	PT	0.07	36	-35	0.01	43	-41	0.05	36	-35	0.25	30	-29	0.46	23	-22	0.30	25	-24	0.08	35	-33	0.38	24	-23
	Mak	0.60	21	-20	0.21	30	-28	0.40	23	-21	0.59	19	-18	0.76	13	-13	0.56	17	-16	0.24	27	-24	0.63	18	-18
	Mak1	0.52	33	-33	0.17	41	-40	0.34	35	-34	0.51	32	-32	0.69	28	-28	0.48	31	-31	0.19	39	-37	0.55	32	-32
	HS	0.50	20	-19	0.29	23	-22	0.35	22	-20	0.40	17	-15	0.73	5	1	0.53	11	-9	0.07	32	-29	0.47	23	-23
	IV	0.27	47	44	0.60	18	13	0.56	30	27	0.58	15	11	0.85	17	-16	0.82	15	-14	0.82	9	-5	0.74	19	-18
	WMO	0.51	10	-1	0.58	14	5	0.56	14	-5	0.65	20	-18	0.82	38	-38	0.75	31	-30	0.65	19	-14	0.76	28	-28
	Mah	0.39	24	20	0.51	27	22	0.39	25	16	0.49	14	-7	0.76	24	-23	0.62	20	-18	0.59	15	0	0.65	16	-14
	Tra	0.35	32	29	0.46	36	31	0.35	33	25	0.44	13	0	0.71	19	-18	0.56	16	-11	0.54	18	7	0.60	11	-7
	Sc	0.13	76	71	0.51	33	30	0.42	51	47	0.62	30	29	0.59	11	9	0.68	22	22	0.68	26	25	0.41	19	18
Winter	JH	0.62	17	-16	0.28	30	-28	0.59	23	-21	0.69	17	-17	0.66	28	-28	0.12	32	-29	0.19	32	-29	0.37	19	-18
	Ab	0.67	27	-26	0.38	36	-35	0.60	33	-32	0.73	28	-28	0.81	30	-30	0.15	41	-39	0.12	45	-43	0.41	32	-32
	Tu	0.66	14	-13	0.24	27	-25	0.60	18	-16	0.67	13	-12	0.63	21	-20	0.11	28	-25	0.16	29	-25	0.29	18	-17
	PT	0.02	44	-43	0.01	51	-50	0.03	44	-42	0.11	38	-37	0.13	37	-36	0.01	54	-52	0.01	59	-56	0.05	44	-43
	Mak	0.48	13	-12	0.19	25	-22	0.46	15	-12	0.71	8	-6	0.73	9	-7	0.11	28	-25	0.09	31	-27	0.32	19	-17
	Mak1	0.40	28	-27	0.15	38	-37	0.38	30	-28	0.66	24	-24	0.67	26	-25	0.09	41	-39	0.07	44	-41	0.26	33	-33
	HS	0.53	20	-20	0.29	25	-23	0.37	22	-20	0.61	15	-14	0.73	8	-7	0.09	32	-29	0.04	40	-37	0.29	26	-26
	IV	0.40	26	22	0.71	10	-4	0.70	16	12	0.41	15	9	0.79	10	-8	0.63	11	-5	0.81	16	12	0.61	7	1
	WMO	0.72	20	-19	0.77	15	-13	0.81	24	-23	0.87	31	-31	0.93	37	-37	0.79	12	-9	0.75	12	-2	0.78	17	-16
	Mah	0.57	11	-4	0.74	9	0	0.61	15	-6	0.65	25	-24	0.68	25	-24	0.82	8	-2	0.69	19	13	0.66	8	0
	Tra	0.52	11	3	0.69	13	8	0.56	16	1	0.60	20	-18	0.63	20	-19	0.77	11	6	0.64	27	21	0.60	12	8
	Sc	0.36	36	32	0.67	12	8	0.74	26	24	0.47	24	22	0.62	7	1	0.41	17	12	0.80	28	26	0.47	23	22

Fig. 3. The Taylor skill score (S), nRMSE (%), and rMBE (%) between seasonal ETp estimated by tested models and Penman in the research period (1970–2014) for eight stations in NSW. The filled colour panels of S, nRMSE, and rMBE have the same meaning with that in Fig. 2.

	TBR			WCN			CBR			GND			
	S	nRMSE	rMBE	S	nRMSE	rMBE	S	nRMSE	rMBE	S	nRMSE	rMBE	
JH	0.60	6	-2	0.36	16	-14	0.62	9	-7	0.81	5	-4	JH
Ab	0.57	20	-19	0.38	28	-27	0.54	24	-23	0.78	21	-21	Ab
Tu	0.67	13	-12	0.44	26	-25	0.72	18	-17	0.73	20	-19	Tu
PT	0.07	34	-34	0.03	40	-39	0.07	33	-32	0.34	25	-24	PT
Mak	0.37	25	-24	0.19	32	-31	0.32	25	-24	0.60	20	-20	Mak
Mak1	0.30	36	-36	0.15	43	-42	0.26	37	-36	0.51	33	-33	Mak1
HS	0.32	20	-20	0.26	23	-22	0.30	21	-20	0.51	14	-13	HS
IV	0.36	33	31	0.60	10	4	0.69	18	16	0.52	8	-1	IV
WMO	0.50	10	4	0.53	14	9	0.51	13	-3	0.66	23	-22	WMO
Mah	0.43	28	25	0.52	28	26	0.38	25	18	0.57	14	-12	Mah
Tra	0.39	37	35	0.47	38	35	0.34	33	27	0.52	10	-5	Tra
Sc	0.13	74	70	0.35	29	25	0.39	45	42	0.49	16	13	Sc
JH	0.77	10	-9	0.55	11	-9	0.32	17	-14	0.58	7	-6	JH
Ab	0.80	19	-19	0.55	25	-25	0.19	36	-35	0.57	28	-27	Ab
Tu	0.60	20	-20	0.39	23	-22	0.20	27	-26	0.39	20	-20	Tu
PT	0.38	19	-19	0.20	26	-26	0.06	32	-31	0.25	23	-22	PT
Mak	0.60	17	-16	0.35	24	-23	0.13	29	-28	0.39	21	-21	Mak
Mak1	0.52	30	-30	0.29	36	-36	0.10	41	-40	0.33	34	-34	Mak1
HS	0.62	4	0	0.37	13	-12	0.07	32	-30	0.42	24	-24	HS
IV	0.84	25	-25	0.74	24	-24	0.82	16	-16	0.63	26	-25	IV
WMO	0.72	39	-39	0.67	26	-25	0.61	19	-16	0.68	29	-29	WMO
Mah	0.70	25	-25	0.59	17	-14	0.61	11	-3	0.63	17	-16	Mah
Tra	0.65	20	-19	0.54	12	-8	0.56	13	4	0.58	12	-10	Tra
Sc	0.78	8	-7	0.66	5	-1	0.81	7	4	0.51	4	0	Sc
	S	nRMSE	rMBE	S	nRMSE	rMBE	S	nRMSE	rMBE	S	nRMSE	rMBE	
	MRD			PTS			SYD			CHB			

Fig. 4. The Taylor skill score (S), nRMSE (%), and rMBE (%) between annual ETp estimated by tested models and Penman in the research period (1970–2014) for eight stations in NSW. The filled colour panels of S, nRMSE, and rMBE have the same meaning with that in Fig. 2.

correlation with ETp (Table 2). Thus, high radiation and temperature in summer, which resulted in relatively high ETp and lower underestimation, might be the main cause result in their better performance. Temperature and radiation in autumn were comparable with that in spring and slight lower than that in summer but the high humidity in autumn result in smaller ETp, thus better performance of the above-mentioned models in autumn.

In specific, model JH produced excellent (nRMSE $\leq$ 10%, -7% < rMBE < 6%, pink colour filled in Fig. 3) estimation of summer and autumn ETp across all stations except for the good (10% < nRMSE $\leq$ 20%, yellow colour filled in Fig. 3) estimation of autumn ETp at WCB (rMBE = -13%) and SYD (rMBE = -10%). For spring and winter, JH generally produced good (10% < nRMSE $\leq$ 20%) or fair (20% < nRMSE $\leq$ 30%) estimation of ETp and the performance in spring was generally better than that in winter. Following JH, Tu produced good or fair estimation of seasonal ETp across all stations. The third and fourth well or fairly performed radiation-based model was Ab and Mak, respectively. Model PT was more suitable for seasonal ETp estimation at sub-humid and humid stations. Performance of HS at semi-arid and sub-humid stations was better than that at arid and humid stations. Performance of IV and Sc showed large variation among stations. Meanwhile, ETp estimated by them showed larger inter-annual fluctuation especially at arid and semi-arid stations (Fig. 5). For mass transfer-based models, WMO outperformed Mah and Tra. In specific, WMO was able to produce good or fair estimation of seasonal ETp at most stations whereas Mah and Tra tended to seriously overestimate summer and autumn ETp at arid and semi-arid stations. Their performance also showed large inter-annual fluctuation and variation among stations.

JH followed by Tu still outperformed all other models in annual ETp estimation (Figs. 4 & 6). In specific, JH produced excellent or good estimation of annual ETp across all stations and Tu produced good or fair estimation of annual ETp across all stations. Ab and Mak produced good or fair estimation of ETp at seven out of eight stations. Performance of PT became good or fair with climate getting wetter whereas annual ETp estimated by Mak1 was generally poor (nRMSE > 30%) for most stations. Temperature-based models HS and IV also produced good or fair estimation of annual ETp for most stations. Performance of HS was better at semi-arid and sub-humid stations than it did at arid and humid stations. On the contrary, IV performed better at arid and semi-arid stations than it did at the wet

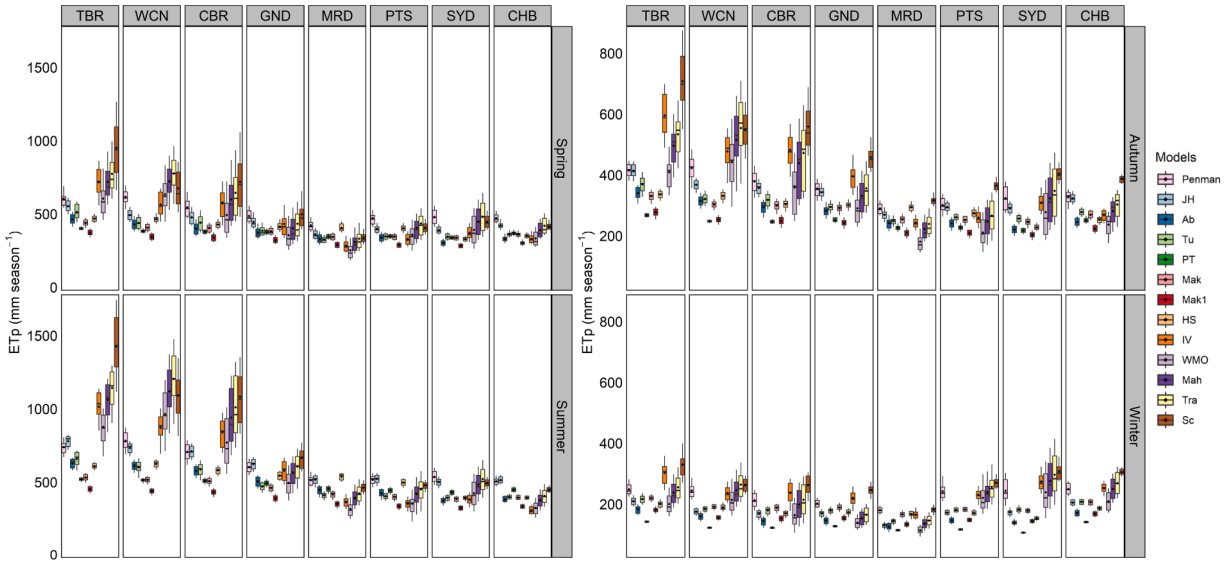


Fig. 5. Boxplots of seasonal ETp in the research period (1970–2014) for eight stations in NSW. Lower and upper box boundaries indicate the 25th and 75th percentiles, respectively. The black line and dot inside each box indicate the median and mean, respectively. The lower and upper whiskers indicate the 10th and 90th percentiles, respectively.

stations at annual scale. Performance of Sc showed large difference among stations, that is, excellent at sub-humid and humid stations but poor or barely fair/good at arid and semi-arid stations. Mass transfer-based models WMO and Mah produced good or fair estimation of annual ETp for most stations whereas performance of Tra was not acceptable at arid stations but good at sub-humid and humid stations.

3.2. Ability of alternative models to capture temporal trends in ETp

Fig. 6 displayed that ETp showed inter-annual variation during the research period. Compared with Penman model, mass transfer-based models and temperature-based models except HS generally exaggerated the magnitude of the variation, especially at arid and semi-arid stations. On the contrary, the inter-annual variation detected by radiation-based models was smaller than that detected by Penman. The unmatched temporal trends of ETp were also confirmed by the result of MK test (Fig. 7). In specific, Penman model only detected significant increase of ETp at semi-arid CBR and the two humid stations (SYD & CHB) with increasing rates less than 10 mm year^{-1} at SYD and 4 mm year^{-1} at the other two stations. However, the increasing rates detected by them were around two times than that detected by Penman model. On the contrary, most radiation-based models noticed significant increase of ETp at CBR, SYD, and PTS with smaller increase rates. As to temperature-based models, IV and Sc detected increase in ETp for most stations but the increase detected by HS was not significant for most stations. In summary, the change trends of ETp detected by the 12 tested models were not so consistent with that detected by Penman model.

The change trends of ETp were generally driven by changes of climatic factors. Thus, this study also analysed the change trends of climatic factors including Rn, u_2, RH, T and difference between Tmax and Tmin (DT), as shown in Fig. S5. It showed that Rn at TBR and GND showed significant decrease; u_2 showed significant decrease at arid stations, significant increase at humid stations, but no uniform change pattern for the semi-arid and sub-humid stations. RH generally showed decrease trend at all stations and the decrease was significant at arid and humid stations. On the contrary, T showed significant increase at all stations except WCN. DT generally showed an increasing trend except the significant decrease at SYD. It indicated that the increase rate of Tmin at SYD was larger than that of Tmax at SYD.

3.3. Ability of alternative models to analyze the periodicity in ETp

The results of wavelet analysis showed that all models were capable to detect the primary period (pp, characterised with a maximum vibration intensity) and the quasi period (qp, characterised with secondary maximum vibration intensity) of ETp (Fig. 8) in spite of slight difference among 13 ETp models. In addition, the multi model averaged pp and qp at eight stations were similar, ranging from 9.6 year to 12.4 year (Figs. S6–S13), and from 2.6 year to 3.9 year (Figs. S6–S13), respectively. Thus, we only displayed the results of Tibooburra as an example (Fig. 8). Figures of other stations were offered in the supplementary material (Figs. S6–S13).

In specific, WMO produced the same pp with Penman model at all stations except Sydney, where Penman-calculated pp and qp were 3.9 year and 8.3 year, respectively (Fig. S12). The Penman-calculated pp and qp at the other seven stations ranged from 8.3 year to 12.4 year and from 2.6 year to 4.6 year, respectively. Second to WMO, Mah and Tra produced the same pp with Penman at five out of eight stations. As to radiation-based and temperature-based models, they generally produced slightly longer pp and qp than Penman

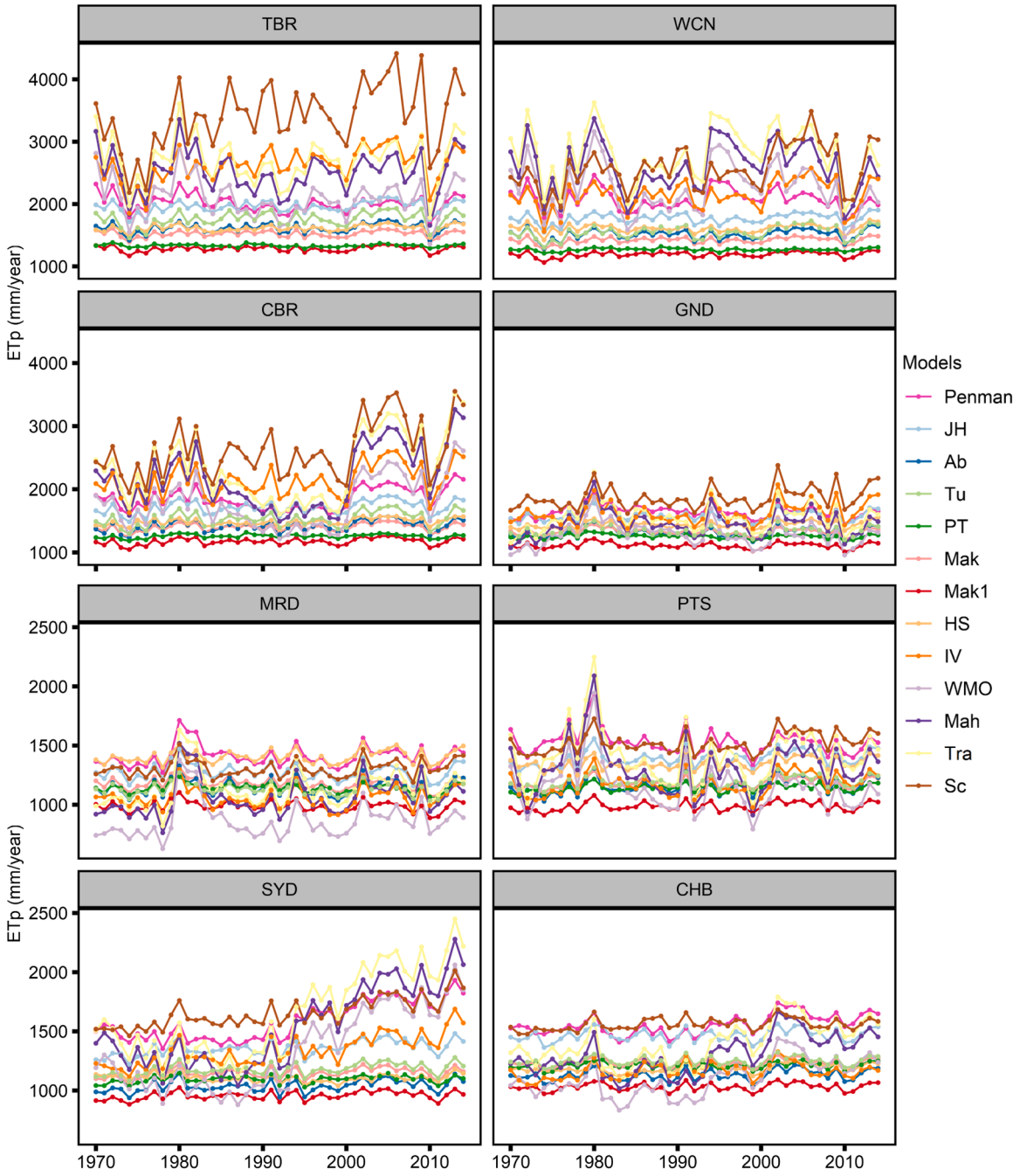


Fig. 6. The annual ETp estimated by the 12 tested models and Penman at eight stations across different climate regimes in NSW, eastern Australia.

did.

3.4. Sensitivity of ETp to meteorological factors

Without surprising, T, Rn, and u_2 showed positive influence on ETp whereas RH showed negative influence in dependence of time scales, and SC of these factors showed difference (Figs. 9, 10 & S14). At annual scale, Rn followed by RH showed greatest influence on ETp at arid and semi-arid stations, with SC ranging from 0.48 to 0.59 and − 0.39 to − 0.56, respectively (Fig. 9). On the contrary, RH followed by Rn showed greatest influence on ETp at humid and sub-humid stations, with SC ranging from − 0.64 to − 0.83 and 0.54–0.63, respectively (Fig. 9). As to T and u_2, T showed slightly greater influence on ETp than u_2 at all stations besides WCN. Another noticeable finding was that the difference of SC between dominant and secondary factors at wetter stations was larger than that at drier

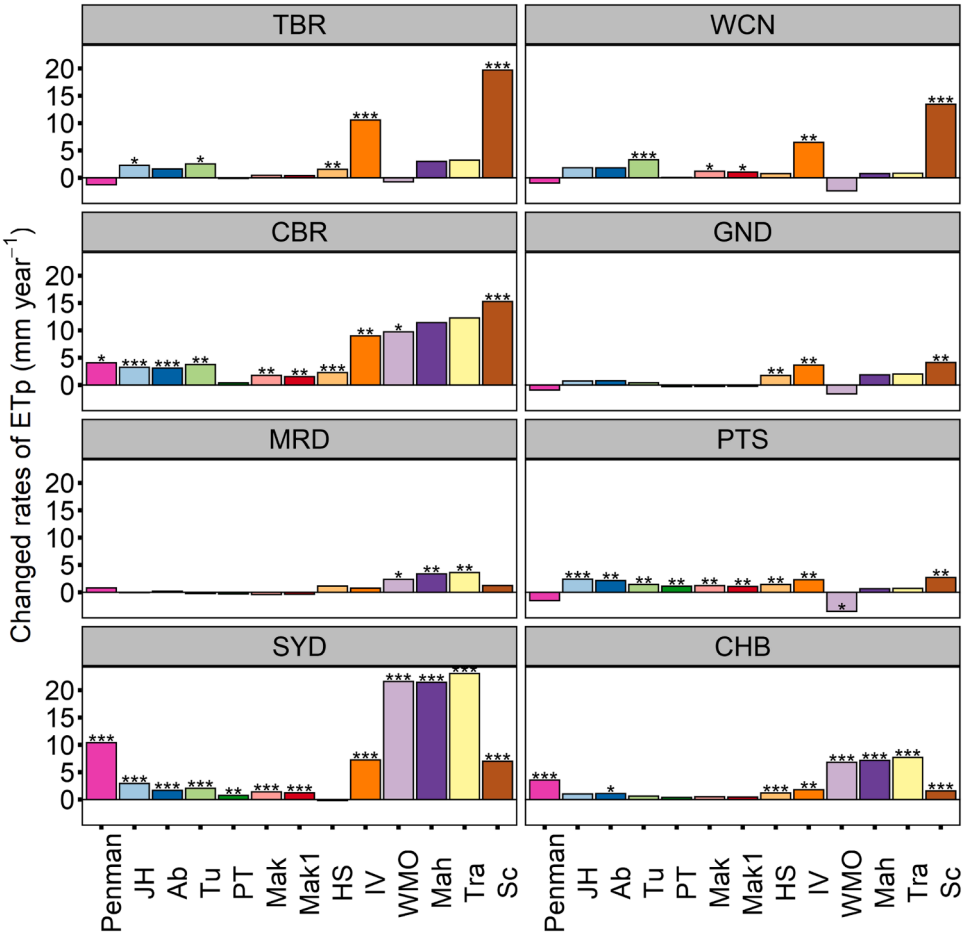


Fig. 7. The change rates and long-term trends of ETp estimated by the 12 tested models and Penman with Mann-Kendall test and Sen's estimator. The positive values indicated upward trend of ETp whereas negative values mean downward trend. *** indicated significant trend at 99% confidence level, ** indicated significant trend at 95% confidence level, and * indicated significant trend at 90% confidence level.

stations. For instance, SC of the dominant (Rn) and secondary (u_2) factors at TBR (arid station) was 0.52 and 0.25 (about two times difference), respectively, indicating 10% increase of Rn with 5.2% increase of ETp and 10% increase of u_2 with 2.5% increase of ETp. However, SC of the dominant (RH) and secondary (u_2) factors at CHB (humid station) was − 0.83 and 0.22 (nearly four times difference), respectively, indicating 10% increase of RH with 8.3% decrease of ETp and 10% increase of u_2 with 2.2% increase of ETp.

Rn and RH were still the dominant factors influencing seasonal ETp (Fig. 10). In specific, SC of Rn was generally larger than the absolute SC of RH in spring and summer but the difference between them became smaller at wetter stations or the influence of RH was even larger than that of Rn. On the contrary, it was RH followed by Rn that had the largest influence on ETp in autumn and winter, especially for wetter stations. Similar to that at annual scale, T was generally the third importance factor influencing ETp except that u_2 took its place in winter for most stations. In addition, SC of RH (followed by Rn) showed the largest variation among seasons for a given station. That is, SC of RH generally became larger from spring to winter but had the lowest value in summer (on the contrary for Rn). For instance, SC of RH (Rn) at TBR was − 0.31 (0.54) in spring, − 0.24 (0.56) in summer, − 0.45 (0.51) in Autumn, and − 0.74 (0.46) in winter. The variation in SC of T and u_2 was smaller among seasons and stations. In general, SC of T and u_2 ranged from 0.20–0.37 and 0.13–0.38, respectively. Changes of SC at monthly scale (Fig. S14) were generally unified with that at seasonal scale. In other words, Rn followed by RH was the dominant factor influencing ETp in spring and summer months across all stations except the humid ones whereas RH followed by Rn was the first important factor influencing ETp in autumn and winter months.

4. Discussion

Our results showed that only the radiation-based JH was able to produce fair estimation of daily ETp across all eight stations (Fig. 2). At larger time scales (i.e., monthly, seasonal, and annual scale), the top two models were JH and Tu, which produced acceptable ETp estimation across all stations. The rest radiation-based models except Mak1 (Ab, PT & Mak), temperature-based models (HS, IV & Sc), and mass transfer-based models (WMO, Mah & Tra) were able to produce reasonable ETp estimation (nRMSE $\leq$ 30%) for

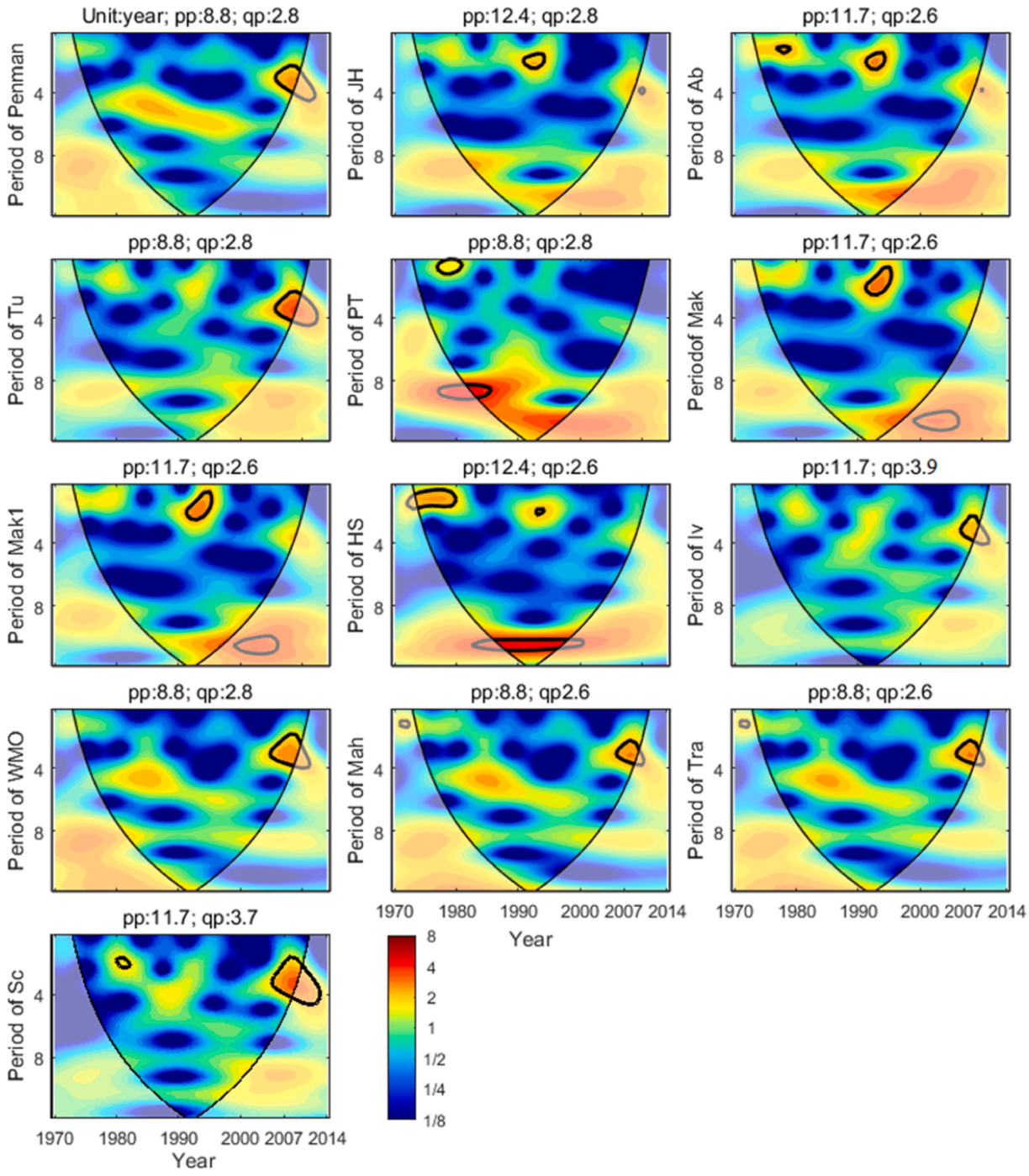


Fig. 8. The wavelet-spectra and variances of annual ETp estimated by 12 tested models and Penman at Tibooburra. The thin solid lines denote the cones of influence, and the thick solid lines show the 95% confidence levels. The colour bar means the vibration intensity of the periods at different timescales.

most stations in spite of the difference among stations and seasons (Figs. 3–6 & S1–S4). For instance, performance of PT became better at sub-humid and humid stations. The poor performance of PT model was reflected by the significant underestimation of ETp at arid and semi-arid stations. This may be explained by the fact that the PT model was developed for saturated land and open water surfaced where advection effects were negligible (Priestley and Taylor, 1972). Thus, the original coefficient of 1.26 embedded in PT may lead to significant underestimation of ETp at arid and semi-arid stations (Li et al., 2017; Tongwane et al., 2017). In Liu and Yang (2021), they found that PT model was more suitable for regions where temperature was a dominant factor influencing ETp. Our study found that Rn

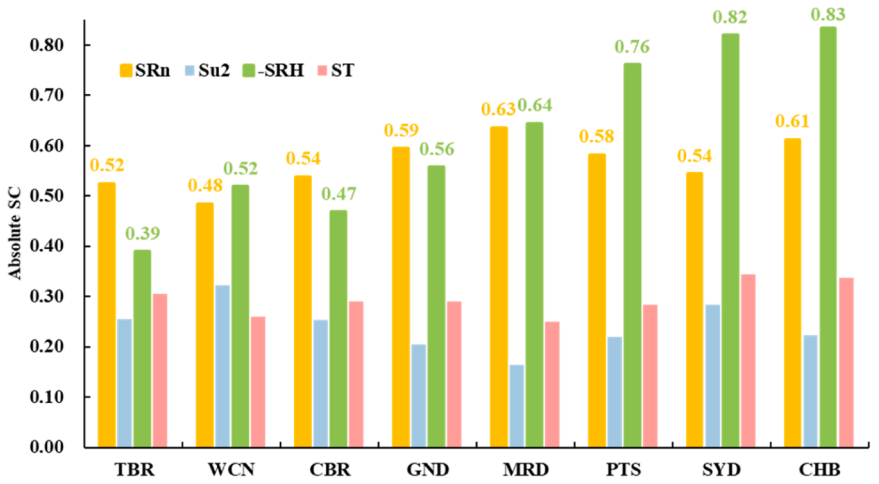


Fig. 9. The averaged absolute sensitivity coefficients of Rn (SRn), u_2 (Su2), RH (-SRH), and T (ST) at annual scale during the study period from 1970 to 2014. Only RH showed negative influence on ETp. Only RH showed negative influence on ETp. To compare the value more conveniently, the absolute value of SRH was adopted.

	SRn	Su2	-SRH	ST	SRn	Su2	-SRH	ST	
	Spring				Summer				
TBR	0.54	0.27	0.31	0.34	0.56	0.24	0.24	0.37	TBR
WCN	0.50	0.33	0.44	0.29	0.53	0.30	0.32	0.34	WCN
CBR	0.56	0.25	0.38	0.32	0.58	0.24	0.30	0.36	CBR
GND	0.62	0.20	0.48	0.31	0.65	0.18	0.41	0.33	GND
MRD	0.66	0.17	0.55	0.26	0.71	0.13	0.49	0.28	MRD
PTS	0.59	0.23	0.64	0.28	0.69	0.16	0.60	0.30	PTS
SYD	0.57	0.29	0.68	0.34	0.64	0.24	0.73	0.37	SYD
CHB	0.62	0.23	0.70	0.34	0.70	0.18	0.77	0.36	CHB
TBR	0.51	0.23	0.45	0.32	0.46	0.25	0.74	0.26	TBR
WCN	0.47	0.31	0.61	0.27	0.40	0.33	0.99	0.22	WCN
CBR	0.52	0.24	0.55	0.31	0.47	0.26	0.92	0.24	CBR
GND	0.56	0.21	0.64	0.30	0.50	0.22	0.93	0.25	GND
MRD	0.62	0.16	0.75	0.25	0.51	0.22	1.04	0.20	MRD
PTS	0.60	0.19	0.88	0.29	0.39	0.33	1.09	0.25	PTS
SYD	0.53	0.28	0.98	0.35	0.36	0.38	1.06	0.30	SYD
CHB	0.61	0.22	0.99	0.34	0.45	0.30	1.02	0.28	CHB
	Autumn				Winter				
	SRn	Su2	-SRH	ST	SRn	Su2	-SRH	ST	

Fig. 10. The averaged absolute sensitivity coefficients of Rn (SRn), u_2 (Su2), RH (-SRH), and T (ST) at seasonal scale during the study period from 1970 to 2014. Only RH showed negative influence on ETp. To compare the value more conveniently, the absolute value of SRH was adopted. The left upper panel is for spring; the right upper panel is for summer; the left bottom panel is for autumn; and the right bottom panel is for winter.

and RH were the dominant factors influencing ETp. The unmatchable role of temperature played in structure of PT and influence on ETp in the study region may also partly explain the relatively poor performance of PT. Similarly, the poor performance of Sc and the overestimation by mass transfer-based models at arid/semi-arid stations was also mainly due to models' structure (Valipour et al., 2017). High temperature, low relative humidity, and large vapour pressure deficit were common at arid/semi-arid stations, thus leading to overestimation of ETp by Sc.

The simplified models only agreed with Penman model on the temporal trends of ETp at stations where Penman detected a significant trend in ETp change (Fig. 7). This finding agreed with Bormann (2010), which claimed that simplified ETp models only reacted in the same direction of ETp change with Penman model at stations exhibiting strong climate trends. Similarly, Donohue et al. (2010) found that the ability of simplified ETp models in capturing the temporal change of ETp in Australia was not comparable with Penman model. As to models' ability in capturing periodicity of ETp, mass transfer-based models produced the same pp of ETp with Penman at 60% stations whereas radiation-based and temperature-based models produced slightly longer periods than Penman did. The disparities in capturing dynamics of ETp among models are mainly caused by models' structure and underlying physics (Liu et al., 2023; Liu et al., 2022; Liu and Yang, 2021; Vishwakarma et al., 2022). In this study, RH and Rn were demonstrated as the dominant factors influencing ETp (Figs. S5, 9, and 10). The significant decrease of RH combined with significant increase of wind speed and temperature lead to the significant increase of ETp at SYD and CHB (Guo et al., 2017; Liu et al., 2023). However, the radiation-based models and temperature-based HS were generally based on Rs (an incarnation of Rn with no significant trends at Sydney and Coffs) and temperature with little manifestation of RH and wind speed in their equations (Table 2). On the contrary, mass transfer-based models simplify the complicated relationship between ETp and VPD (an incarnation of RH) as a linear function and linear or square root function with wind speed (Table 2). The difference in model's structure can partly explain why radiation-based models generally underestimate whereas mass transfer-based models overestimate the temporal trends of ETp. In addition, Liu et al. (2022) found that PM-based models tended to overestimate the sensitivity of ETp to temperature. Thus, there is a possibility that the significant increase of ETp at SYD and CHB captured by Penman model was partly overestimated due to the significant increase in temperature.

Even though the simplified models adopt empirical/experimental coefficients to weaken the influence caused by models' structure and assumption in evaporation process, the application of these models beyond where they were developed companied with uncertainty (Abhishek et al., 2021; Cristea et al., 2013; Nouri and Homaee, 2018; Xiong et al., 2023). For instance, model HS had better performance at semi-arid and sub-humid stations than it did at other stations (Figs. 3–5). However, its better performance still varied among these stations in terms of levels of performance (excellent, good, or fair). Thus, it's with high confidence to conclude that uncertainty exists when it comes to upscale models' performance to larger scale or other study regions even if the climate condition is similar. This kind of uncertainty was mainly caused by the spatiotemporal heterogeneity of climate and widely reported in literatures (McMahon et al., 2016; Xiang et al., 2020; Xiong et al., 2023). Another factor leading to the disparities in models' performance is that these models were developed based on different evaporative surface (Table 2). The vagueness of evaporative surface is embedded in the definition of potential evapotranspiration (McMahon et al., 2016; McMahon et al., 2013). Similar to estimation of crop evapotranspiration based on ETp, crop coefficients (Kc) can be used to adjust ETp caused by difference in evaporative surface. Meanwhile, it is likely that a location-specific calibration might improve model performance (Cristea et al., 2013; Shiri, 2017; Sumner and Jacobs, 2005; Xu and Singh, 2002). However, the re-adjustment of models' coefficients could be site-specific (Ravazzani et al., 2012) and greatly depend on the study period (Raziei and Pereira, 2013; Tabari and Talaee, 2011). In other words, the possibility of misleading to decision-makers companies with the readjustment of original coefficients embedded in empirical modes (Nouri and Homaee, 2018). Thus, we only evaluated models' performance based on the original coefficients.

In addition to mathematical models, researchers had more choice in ETp estimation and observation, such as mass conservation method, reanalysis and satellite-based remote sensing products, in-situ measurements, machine learning method and so on (Mehdizadeh, 2018; Xiong et al., 2023; Zhang et al., 2019). On the one hand, various methods in ETp estimation and observation open up new ways to understand ETp process and hydrological cycle better, as suggested by Xiong et al. (2023) 'different types of datasets have their unique values'. For instance, it is based on ETp observation with eddy covariance measurements that Liu et al. (2022) demonstrated the PM-based model overestimate the sensitivity of ETp to temperature. Similarly, Pimentel et al. (2023) adopted global evapotranspiration data from MODIS16 to investigate performance of JH, PT, and HS in hydrological modelling. On the other hand, it may also bring more uncertainties and disparities in ETp estimation due to the difference in methods, spatial resolution, and model assumptions, thus leading to confusion in the selection of the most appropriate product or method to use (Liu et al., 2016; Pascolini-Campbell et al., 2020; Xiong et al., 2023). In this context, it is important to leverage the multi-source of ETp or related product to get comprehensive information in hydrological studies. Meanwhile, measures such as adopting ensemble mean of various ETp products can be taken to minimise uncertainties in future studies (Abhishek et al., 2021; Xiong et al., 2022).

5. Conclusions

This study assessed the performance of 12 simplified models in estimating potential evapotranspiration (ETp) rates, capturing its temporal trends, and interannual oscillation with Penman model as the benchmark at eight stations in New South Wales, southeast Australia. All models were likely to underestimate ETp with the exception that mass transfer-based (WMO, Mah, and Tra) and temperature-based IV and Sc overestimated it at arid and semi-arid stations. For daily ETp estimation, only model JH was recommendable across New South Wales when Penman was limited to use; HS was recommended for semi-arid and sub-humid climate stations. Other models' performance showed large variation among stations, thus not recommendable for daily ETp estimation. Models' performance except Mak1 was generally improved in estimation of seasonal and annual ETp. In specific, JH and Tu were the most recommendable models for seasonal and annual ETp estimation considering their stable and firm performance across all stations.

Table 2
Information of 13 models adopted in this study to calculate potential evapotranspiration.

Models	References	Formula	Notes
Penman	Donohue et al. (2010); Milly and Dunne (2016)	$ET_p = \frac{0.408\Delta}{\Delta+\gamma}(R_n - G) + \frac{\gamma}{\Delta+\gamma}\frac{6.43(1+0.536u_2)(e_s-e_a)}{\lambda}$	Open water evaporation, often called as Penman potential evaporation
Jensen-Haise (JH)	Jensen and Haise (1963)	$ET_p = 0.0102(T+3)R_s$	ETp from an alfalfa surface
Abtew (Ab)	Abtew (1996)	$ET_p = 0.01786\frac{R_s T_{\max}}{\lambda}$	ETp from a grass surface
Turc (Tu)	Turc (1961)	$ET_p = (0.3107R_s + 0.65)\frac{Ta_t}{T+15} \quad a_t = \begin{cases} 1 & RH \geq 50\% \\ 1+\frac{50-RH}{70} & RH < 50\% \end{cases}$	ETp from a grass surface
Priestley-Taylor (PT)	Priestley and Taylor (1972)	$ET_p = 1.26\left[\frac{\Delta}{\Delta+\gamma}\frac{R_n}{\lambda} - \frac{G}{\lambda}\right]$	ETp from a pasture surface and E from an open water surface
Modified Makkink (Mak)	Hansen (1984)	$ET_p = 0.7\frac{\Delta}{\Delta+\gamma}\frac{R_s}{\lambda}$	ETp from a grass surface
Makkink (Mak1)	Makkink (1957)	$ET_p = 0.61\frac{\Delta}{\Delta+\gamma}\frac{R_s}{\lambda} - 0.12$	ETp from a grass surface
Hargreaves (HS)	Hargreaves et al. (1985)	$ET_p = 0.0023 \times 0.408R_a(T_{\max} - T_{\min})^{0.5}(T+17.8)$	ETp from a grass surface
Ivanov (Iv)	Valipour et al. (2017)	$ET_p = 0.00006(25+T)^2(100-RH)$	ETp from an open water surface
Schendel (Sc)	Djaman et al. (2015)	$ET_p = 16\frac{T}{RH}$	ETp from an open water surface
WMO	Valipour et al. (2017)	$ET_p = (1.298 + 0.934u)(e_s - e_a)$	ETp from an open water surface
Mahringer (Mah)	Mahringer (1970)	$ET_p = 2.86u^{0.5}(e_s - e_a)$	ETp from an open water surface
Trabert (Tra)	Valipour et al. (2017)	$ET_p = 3.075u^{0.5}(e_s - e_a)$	ETp from an open water surface

Then, Ab and Mak were also choices for seasonal (especially summer and autumn) and annual ETp estimation nearly for all stations. By contrast, PT was effective at sub-humid and humid stations but poor performed at arid stations. Temperature-based models HS and IV were also reliable at most stations with better performance at semi-arid and sub-humid stations, but Sc was only recommended for humid/sub-humid stations. Mass transfer-based model WMO performed generally good with better performance at arid and semi-arid stations while Tra and Mah could be used to estimate ETp in humid/sub-humid climates. The 12 alternative models only showed agreement with Penman model at stations where Penman detected significant change trends of ETp, but the increasing rates detected by them were smaller (radiation-based models) or larger (mass transfer-based models) than Penman did.

Though the meteorological stations used in this study varies from arid, semi-arid, sub-humid, to humid climate conditions, the limited number of available stations (eight in total) may bring limitations and uncertainties when it comes to upscale the findings to larger spatial scale or study regions due to heterogeneity in climate. This study assessed models' performance only based on Penman model. However, there are more and more ETp methods and products available with the development of machine learning, remote sensing, and satellite technology. In this context, it is necessary for future study to leverage the multi-source datasets to assess models' performance and adopt multi-source ensemble mean value of ETp to minimise the potential uncertainties.

Our study uncovers the most appropriate models in ETp estimation for different time scales in New South Wales, southeast Australia, which is susceptible to drought and flood hazards. Findings of this study can be used in model's selection in the regional drought assessment and runoff simulation, especially under future climate scenarios, where the observed ETp by satellite remote sensing or traditional methods and full set of climatic factors from general circulation models are not always available for regional study.

CRediT authorship contribution statement

Linchao Li: Software. **Gengxi Zhang:** Writing – review & editing. **Qiang Yu:** Funding acquisition, Supervision. **Bin Wang:** Formal analysis, Methodology, Writing – review & editing. **De Li Liu:** Conceptualization, Project administration. **Puyu Feng:** Methodology, Software. **James Cleverly:** Conceptualization. **Lijie Shi:** Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Software, Writing – original draft, Writing – review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgements

The first author appreciated the scholarship from the Chinese Scholarship Council (12818579). We appreciated the funding (NO. 137012561 & NO. 137013033) Yangzhou University and Yangzhou City offered for our research. We thank the Scientific Information for Land Owners and Bureau of Meteorology for offering the climatic data used in this work. We also acknowledge the support from Natural Science Foundation of China (NO. 41961124006).

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.ejrh.2023.101573.

References

Abhishek, Kinouchi, T., Sayama, T., 2021. A comprehensive assessment of water storage dynamics and hydroclimatic extremes in the Chao Phraya River Basin during 2002–2020. J. Hydrol. 603, 126868 https://doi.org/10.1016/j.jhydrol.2021.126868.
Abtew, W., 1996. Evapotranspiration measurements and modeling for three wetland systems in south florida. JAWRA J. Am. Water Resour. Assoc. 32 (3), 465–473.
Ahmadi, S.H., Javanbakht, Z., 2020. Assessing the physical and empirical reference evapotranspiration (ETo) models and time series analyses of the influencing weather variables on ETo in a semi-arid area. J. Environ. Manag. 276, 111278 https://doi.org/10.1016/j.jenvman.2020.111278.
Allen, R.G., Pereira, L.S., Raes, D., Smith, M., 1998. Crop Evapotranspiration - Guidelines for Computing Crop Water Requirements. FAO Irrigation & Drainage Paper 56, Rome, Italy.
Almorox, J., Quej, V.H., Martí, P., 2015. Global performance ranking of temperature-based approaches for evapotranspiration estimation considering Köppen climate classes. J. Hydrol. 528, 514–522. https://doi.org/10.1016/j.jhydrol.2015.06.057.
Angi, R., Harald, S., 2014. WaveletComp: Computational Wavelet Analysis. R package version 1.0.
Azhar, A.H., Perera, B.J.C., 2011. Evaluation of reference evapotranspiration estimation methods under southeast Australian conditions. J. Irrig. Drain. Eng. 137 (5), 268–279. https://doi.org/10.1061/(ASCE)IR.1943-4774.0000297.
Bormann, H., 2010. Sensitivity analysis of 18 different potential evapotranspiration models to observed climatic change at German climate stations. Clim. Change 104 (3–4), 729–753. https://doi.org/10.1007/s10584-010-9869-7.
Cristea, N.C., Kampf, S.K., Burges, S.J., 2013. Revised coefficients for Priestley-Taylor and Makkink-Hansen equations for estimating daily reference evapotranspiration. J. Hydrol. Eng. 18 (10), 1289–1300. https://doi.org/10.1061/(ASCE)HE.1943-5584.0000679.
Dettori, M., Cesaraccio, C., Motroni, A., Spano, D., Duce, P., 2011. Using CERES-wheat to simulate durum wheat production and phenology in Southern Sardinia, Italy. Field Crops Res. 120 (1), 179–188. https://doi.org/10.1016/j.fcr.2010.09.008.
Djaman, K., et al., 2015. Evaluation of sixteen reference evapotranspiration methods under sahelian conditions in the Senegal River Valley. J. Hydrol. Reg. Stud. 3, 139–159.
Donohue, R.J., McVicar, T.R., Roderick, M.L., 2010. Assessing the ability of potential evaporation formulations to capture the dynamics in evaporative demand within a changing climate. J. Hydrol. 386 (1–4), 186–197. https://doi.org/10.1016/j.jhydrol.2010.03.020.
Douglas, E.M., Vogel, R.M., Kroll, C.N., 2000. Trends in floods and low flows in the United States: impact of spatial correlation. J. Hydrol. 240 (1), 90–105. https://doi.org/10.1016/S0022-1694(00)00336-X.
Droogers, P., Allen, R.G., 2002. Estimating reference evapotranspiration under inaccurate data conditions. Irrig. Drain. Syst. 16 (1), 33–45.
Fan, J., Wu, L., Zhang, F., Xiang, Y., Zheng, J., 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.
Farge, M., 1992. Wavelet transforms and their applications to turbulence. Annu. Rev. Fluid Mech. 24 (1), 395–458.
Gao, Z., He, J., Dong, K., Li, X., 2017. Trends in reference evapotranspiration and their causative factors in the West Liao River basin, China. Agric. For. Meteorol. 232, 106–117. https://doi.org/10.1016/j.agrformet.2016.08.006.
Guo, D., Westra, S., Maier, H.R., 2017. Sensitivity of potential evapotranspiration to changes in climate variables for different Australian climatic zones. Hydrol. Earth Syst. Sci. 21 (4), 2107–2126. https://doi.org/10.5194/hess-21-2107-2017.
Han, J., et al., 2018. Spatio-temporal variation of potential evapotranspiration and climatic drivers in the Jing-Jin-Ji region, North China. Agric. For. Meteorol. 256–257, 75–83. https://doi.org/10.1016/j.agrformet.2018.03.002.
Hansen, S., 1984. Estimation of potential and actual evapotranspiration. Paper presented at the Nordic Hydrological Conference (Nyborg, Denmark, August - 1984), 15(4–5), pp. 205–212.
Hargreaves, G.H., Allen, R.G., 2003. History and evaluation of hargreaves evapotranspiration equation. J. Irrig. Drain. Eng. 129 (1), 53–63. https://doi.org/10.1061/(ASCE)0733-9437(2003)129:1(53).
Hargreaves, G.L., Hargreaves, G.H., Riley, J.P., 1985. Irrigation water requirements for Senegal River basin. J. Irrig. Drain. Eng. 111 (3), 265–275. https://doi.org/10.1061/(ASCE)0733-9437(1985)111:3(265).
Hu, Z., Liu, S., Zhong, G., Lin, H., Zhou, Z., 2020. Modified Mann-Kendall trend test for hydrological time series under the scaling hypothesis and its application. Hydrol. Sci. J. 65 (14), 2419–2438. https://doi.org/10.1080/02626667.2020.1810253.
Irmak, S., Kabenge, I., Skaggs, K.E., Mutiibwa, D., 2012. Trend and magnitude of changes in climate variables and reference evapotranspiration over 116-yr period in the Platte River Basin, central Nebraska–USA. J. Hydrol. 420–421, 228–244. https://doi.org/10.1016/j.jhydrol.2011.12.006.
Jensen, M.E., Allen, R.G., 1990. Evaporation, evapotranspiration, and irrigation water requirements. Evaporation, Evapotranspiration, and Irrigation Water Requirements. https://doi.org/10.1061/9780784414057.
Jensen, M.E., Haise, H.R., 1963. Estimating evapotranspiration from solar radiation. Proc. Am. Soc. Civ. Eng. J. Irrig. Drain. Div. 89, 15–41.
Jung, M., et al., 2010. Recent decline in the global land evapotranspiration trend due to limited moisture supply. Nature 467 (7318), 951.
Kool, D., et al., 2014. A review of approaches for evapotranspiration partitioning. Agric. For. Meteorol. 184, 56–70. https://doi.org/10.1016/j.agrformet.2013.09.003.
Kumar, K.K., Kumar, K.R., Rakhecha, P., 1987. Comparison of Penman and Thornthwaite methods of estimating potential evapotranspiration for Indian conditions. Theor. Appl. Climatol. 38 (3), 140–146.
Li, S., et al., 2015. Comparison of several surface resistance models for estimating crop evapotranspiration over the entire growing season in arid regions. Agric. For. Meteorol. 208, 1–15. https://doi.org/10.1016/j.agrformet.2015.04.002.

Li, Y., Yao, N., Chau, H.W., 2017. Influences of removing linear and nonlinear trends from climatic variables on temporal variations of annual reference crop evapotranspiration in Xinjiang, China. Sci. Total Environ. 592, 680–692. https://doi.org/10.1016/j.scitotenv.2017.02.196.
Liang, L., Li, L., Liu, Q., 2010. Temporal variation of reference evapotranspiration during 1961–2005 in the Taoer River basin of Northeast China. Agric. For. Meteorol. 150 (2), 298–306. https://doi.org/10.1016/j.agrformet.2009.11.014.
Liu, S., Xie, Y., Fang, H., Xu, P., Du, H., 2023. A method for identifying the dominant meteorological factors of atmospheric evaporative demand in mid-long term. Water Resour. Res. 59 (7), e2022WR033321 https://doi.org/10.1029/2022WR033321.
Liu, W., et al., 2016. A worldwide evaluation of basin-scale evapotranspiration estimates against the water balance method. J. Hydrol. 538, 82–95. https://doi.org/10.1016/j.jhydrol.2016.04.006.
Liu, X., et al., 2017. Comparison of 16 models for reference crop evapotranspiration against weighing lysimeter measurement. Agric. Water Manag. 184, 145–155. https://doi.org/10.1016/j.agwat.2017.01.017.
Liu, Z., Yang, H., 2021. Estimation of water surface energy partitioning with a conceptual atmospheric boundary layer model. Geophys. Res. Lett. 48 (9), e2021GL092643 https://doi.org/10.1029/2021GL092643.
Liu, Z., Han, J., Yang, H., 2022. Assessing the ability of potential evaporation models to capture the sensitivity to temperature. Agric. For. Meteorol. 317, 108886 https://doi.org/10.1016/j.agrformet.2022.108886.
Lu, J., Sun, G., McNulty, S.G., Amatya, D.M., 2005. A comparison of six potential evapotranspiration methods for regional use in the Southeastern United States. JAWRA J. Am. Water Resour. Assoc. 41 (3), 621–633. https://doi.org/10.1111/j.1752-1688.2005.tb03759.x.
Mahringer, W., 1970. Verdunstungsstudien am neusiedler See. Arch. Meteorol. Geophys. Bioklimatol. Ser. B 18 (1), 1–20.
Makkink, G., 1957. Testing the Penman formula by means of lysimeters. J. Inst. Water Eng. 11, 277–288.
Martel, M., Glenn, A., Wilson, H., Kröbel, R., 2018. Simulation of actual evapotranspiration from agricultural landscapes in the Canadian Prairies. J. Hydrol. Reg. Stud. 15, 105–118. https://doi.org/10.1016/j.ejrh.2017.11.010.
McCuen, R.H., 1974. A sensitivity and error analysis cf procedures used for estimating evaporation1. JAWRA J. Am. Water Resour. Assoc. 10 (3), 486–497. https://doi.org/10.1111/j.1752-1688.1974.tb00590.x.
McKenney, M.S., Rosenberg, N.J., 1993. Sensitivity of some potential evapotranspiration estimation methods to climate change. Agric. For. Meteorol. 64 (1–2), 81–110.
McMahon, T.A., Peel, M.C., Lowe, L., Srikanthan, R., McVicar, T.R., 2013. Estimating actual, potential, reference crop and pan evaporation using standard meteorological data: a pragmatic synthesis. Hydrol. Earth Syst. Sci. 17 (4), 1331–1363. https://doi.org/10.5194/hess-17-1331-2013.
McMahon, T.A., Finlayson, B.L., Peel, M.C., 2016. Historical developments of models for estimating evaporation using standard meteorological data. Wiley Interdiscip. Rev. Water 3 (6), 788–818.
Mehdizadeh, S., 2018. Estimation of daily reference evapotranspiration (ETo) using artificial intelligence methods: Offering a new approach for lagged ETo data-based modeling. J. Hydrol. 559, 794–812. https://doi.org/10.1016/j.jhydrol.2018.02.060.
Milly, P.C.D., Dunne, K.A., 2016. Potential evapotranspiration and continental drying. Nat. Clim. Change 6, 946. https://doi.org/10.1038/nclimate3046. https://www.nature.com/articles/nclimate3046#supplementary-information.
Muniandy, J.M., Yusop, Z., Askari, M., 2016. Evaluation of reference evapotranspiration models and determination of crop coefficient for Momordica charantia and Capsicum annuum. Agric. Water Manag. 169, 77–89. https://doi.org/10.1016/j.agwat.2016.02.019.
Nouri, M., Homaee, M., 2018. On modeling reference crop evapotranspiration under lack of reliable data over Iran. J. Hydrol. 566, 705–718. https://doi.org/10.1016/j.jhydrol.2018.09.037.
Oudin, L., et al., 2005. Which potential evapotranspiration input for a lumped rainfall–runoff model? J. Hydrol. 303 (1–4), 290–306. https://doi.org/10.1016/j.jhydrol.2004.08.026.
Özger, M., Mishra, A.K., Singh, V.P., 2009. Low frequency drought variability associated with climate indices. J. Hydrol. 364 (1), 152–162. https://doi.org/10.1016/j.jhydrol.2008.10.018.
Pascolini-Campbell, M.A., Reager, J.T., Fisher, J.B., 2020. GRACE-based mass conservation as a validation target for basin-scale evapotranspiration in the contiguous United States. Water Resour. Res. 56 (2), e2019WR026594 https://doi.org/10.1029/2019WR026594.
Peng, B., 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 (8), 2259–2274. https://doi.org/10.1175/JHM-D-15-0202.1.
Peng, B., Liu, X., Zhang, Y., Liu, C., 2020. Assessing the impacts of vegetation greenness change on evapotranspiration and water yield in China. Water Resour. Res. 56 (10), e2019WR027019 https://doi.org/10.1029/2019WR027019.
Peng, L., Li, D., Sheffield, J., 2018. Drivers of variability in atmospheric evaporative demand: multiscale spectral analysis based on observations and physically based modeling. Water Resour. Res.
Peng, S., et al., 2017. Spatiotemporal change and trend analysis of potential evapotranspiration over the Loess Plateau of China during 2011–2100. Agric. For. Meteorol. 233, 183–194. https://doi.org/10.1016/j.agrformet.2016.11.129.
Penman, H.L., 1948. Natural evaporation from open water, bare soil and grass. Proc. R. Soc. Lond. Ser. A Math. Phys. Sci. 193 (1032), 120–145.
Pimentel, R., et al., 2023. Which potential evapotranspiration formula to use in hydrological modeling world-wide? Water Resour. Res. 59 (5), e2022WR033447 https://doi.org/10.1029/2022WR033447.
Pohlert, T., 2016. Non-parametric Trend Tests and Change-point Detection. CC BY-ND, p. 4.
Priestley, C.H.B., Taylor, R.J., 1972. On the assessment of surface heat flux and evaporation using large-scale parameters. Mon. Weather Rev. 100 (2), 81–92. https://doi.org/10.1175/1520-0493(1972)100<0081:Otaosh>2.3.Co;2.
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. J. Irrig. Drain. Eng. 138 (7), 592–599. https://doi.org/10.1061/(ASCE)IR.1943-4774.0000453.
Raziei, T., Pereira, L.S., 2013. Estimation of ETo with Hargreaves–Samani and FAO-PM temperature methods for a wide range of climates in Iran. Agric. Water Manag. 121, 1–18. https://doi.org/10.1016/j.agwat.2012.12.019.
Roderick, M.L., Rotstayn, L.D., Farquhar, G.D., Hobbins, M.T., 2007. On the attribution of changing pan evaporation. Geophys. Res. Lett. 34 (17) https://doi.org/10.1029/2007GL031166.
Scheff, J., Frierson, D.M.W., 2015. Terrestrial aridity and its response to greenhouse warming across CMIP5 climate models. J. Clim. 28 (14), 5583–5600. https://doi.org/10.1175/jcli-d-14-00480.1.
Schendel, U., 1967. Vegetationswasserverbrauch und-wasserbedarf. Habilitation, Kiel, p. 137.
She, D., Xia, J., Zhang, Y., 2017. Changes in reference evapotranspiration and its driving factors in the middle reaches of Yellow River Basin, China. Sci. Total Environ. 607–608, 1151–1162. https://doi.org/10.1016/j.scitotenv.2017.07.007.
Sheffield, J., Wood, E.F., Roderick, M.L., 2012. Little change in global drought over the past 60 years. Nature 491 (7424), 435.
Shiri, J., 2017. Evaluation of FAO56-PM, empirical, semi-empirical and gene expression programming approaches for estimating daily reference evapotranspiration in hyper-arid regions of Iran. Agric. Water Manag. 188, 101–114. https://doi.org/10.1016/j.agwat.2017.04.009.
Shuttleworth, W.J., 1993. Evaporation. In: Maidment, D.R. (Ed.), Handbook of Hydrology. McGraw-Hill, New York, pp. 4.1–4.53.
da Silveira, I.P., Pezzi, L.P., 2014. Sea surface temperature anomalies driven by oceanic local forcing in the Brazil-Malvinas Confluence. Ocean Dyn. 64 (3), 347–360. https://doi.org/10.1007/s10236-014-0699-4.
Sumner, D.M., Jacobs, J.M., 2005. Utility of Penman–Monteith, Priestley–Taylor, reference evapotranspiration, and pan evaporation methods to estimate pasture evapotranspiration. J. Hydrol. 308 (1), 81–104. https://doi.org/10.1016/j.jhydrol.2004.10.023.
Tabari, H., 2009. Evaluation of reference crop evapotranspiration equations in various climates. Water Resour. Manag. 24 (10), 2311–2337. https://doi.org/10.1007/s11269-009-9553-8.
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. J. Hydrol. Eng. 16 (10), 837–845. https://doi.org/10.1061/(ASCE)HE.1943-5584.0000366.

Tabari, H., Marofi, S., Aeini, A., Talaee, P.H., Mohammadi, K., 2011. Trend analysis of reference evapotranspiration in the western half of Iran. Agric. For. Meteorol. 151 (2), 128–136. https://doi.org/10.1016/j.agrformet.2010.09.009.
Tabari, H., Grismer, M.E., Trajkovic, S., 2013. Comparative analysis of 31 reference evapotranspiration methods under humid conditions. Irrig. Sci. 31 (2), 107–117. https://doi.org/10.1007/s00271-011-0295-z.
Taylor, K.E., 2001. Summarizing multiple aspects of model performance in a single diagram. J. Geophys. Res. Atmos. 106 (D7), 7183–7192.
Tongwane, M.I., Savage, M.J., Tsubo, M., Moeletsi, M.E., 2017. Seasonal variation of reference evapotranspiration and Priestley-Taylor coefficient in the eastern Free State, South Africa. Agric. Water Manag. 187, 122–130. https://doi.org/10.1016/j.agwat.2017.03.013.
Torrence, C., Compo, G.P., 1998. A practical guide to wavelet analysis. Bull. Am. Meteorol. Soc. 79 (1), 61–78.
Trenberth, K.E., Fasullo, J.T., Kiehl, J., 2009. Earth's global energy budget. Bull. Am. Meteorol. Soc. 90 (3), 311–324. https://doi.org/10.1175/2008bams2634.1.
Turc, L., 1961. Estimation of Irrigation Water Requirements, Potential Evapotranspiration: A Simple Climatic Formula Evolved up to Date, 12, pp. 13–49.
Valiantzas, J.D., 2006. Simplified versions for the Penman evaporation equation using routine weather data. J. Hydrol. 331 (3), 690–702. https://doi.org/10.1016/j.jhydrol.2006.06.012.
Valipour, M., Gholami Sefidkouhi, M.A., Raeini-Sarjaz, M., 2017. Selecting the best model to estimate potential evapotranspiration with respect to climate change and magnitudes of extreme events. Agric. Water Manag. 180, 50–60. https://doi.org/10.1016/j.agwat.2016.08.025.
Vishwakarma, D.K., et al., 2022. Methods to estimate evapotranspiration in humid and subtropical climate conditions. Agric. Water Manag. 261, 107378 https://doi.org/10.1016/j.agwat.2021.107378.
Wang, B., Liu, D.L., Asseng, S., Macadam, I., Yu, Q., 2015. Impact of climate change on wheat flowering time in eastern Australia. Agric. For. Meteorol. 209–210, 11–21. https://doi.org/10.1016/j.agrformet.2015.04.028.
Wang, Z., et al., 2017. Spatiotemporal variability of reference evapotranspiration and contributing climatic factors in China during 1961–2013. J. Hydrol. 544, 97–108. https://doi.org/10.1016/j.jhydrol.2016.11.021.
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.
Xiong, J., et al., 2022. Annual runoff coefficient variation in a changing environment: a global perspective. Environ. Res. Lett. 17.
Xiong, J., et al., 2023. ET-WB: water balance-based estimations of terrestrial evaporation over global land and major global basins. Earth Syst. Sci. Data Discuss. 2023, 1–47. https://doi.org/10.5194/essd-2023-188.
Xu, C.-Y., Singh, V.P., 2000. Evaluation and generalization of radiation-based methods for calculating evaporation. Hydrol. Process. 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, V.P., 2002. Cross comparison of empirical equations for calculating potential evapotranspiration with data from Switzerland. Water Resour. Manag. 16 (3), 197–219. https://doi.org/10.1023/a:1020282515975.
Xu, K., et al., 2015. Spatio-temporal variation of drought in China during 1961–2012: a climatic perspective. J. Hydrol. 526, 253–264. https://doi.org/10.1016/j.jhydrol.2014.09.047.
Yang, Y., et al., 2019. Sensitivity of potential evapotranspiration to meteorological factors and their elevational gradients in the Qilian Mountains, northwestern China. J. Hydrol. 568, 147–159. https://doi.org/10.1016/j.jhydrol.2018.10.069.
Zhang, Q., Xu, C.-y, Jiang, T., Wu, Y., 2007. Possible influence of ENSO on annual maximum streamflow of the Yangtze River, China. J. Hydrol. 333 (2–4), 265–274.
Zhang, Q., Cui, N., Feng, Y., Gong, D., Hu, X., 2018. Improvement of Makkink model for reference evapotranspiration estimation using temperature data in Northwest China. J. Hydrol. 566, 264–273. https://doi.org/10.1016/j.jhydrol.2018.09.021.
Zhang, X., Kang, S., Zhang, L., Liu, J., 2010. Spatial variation of climatology monthly crop reference evapotranspiration and sensitivity coefficients in Shiyang river basin of northwest China. Agric. Water Manag. 97 (10), 1506–1516. https://doi.org/10.1016/j.agwat.2010.05.004.
Zhang, Y., et al., 2019. Coupled estimation of 500 m and 8-day resolution global evapotranspiration and gross primary production in 2002–2017. Remote Sens. Environ. 222, 165–182. https://doi.org/10.1016/j.rse.2018.12.031.
Zhang, Y., et al., 2020. Can remotely sensed actual evapotranspiration facilitate hydrological prediction in ungauged regions without runoff calibration? Water Resour. Res. 56 (1), e2019WR026236 https://doi.org/10.1029/2019wr026236.
Zhou, M.C., et al., 2006. Estimating potential evapotranspiration using Shuttleworth–Wallace model and NOAA-AVHRR NDVI data to feed a distributed hydrological model over the Mekong River basin. J. Hydrol. 327 (1), 151–173. https://doi.org/10.1016/j.jhydrol.2005.11.013.