Environmental Modelling & Software 78 (2016) 216–224

Contents lists available at ScienceDirect

Environmental Modelling & Software

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

An R package for modelling actual, potential and reference evapotranspiration

Danlu Guo*, Seth Westra, Holger R. Maier

School of Civil, Environmental & Mining Engineering, The University of Adelaide, Adelaide, 5005, Australia

ARTICLE INFO

Article history:
Received 1 June 2015
Received in revised form
31 December 2015
Accepted 31 December 2015
Available online 21 January 2016

Keywords:
Evapotranspiration (ET)
Evaporation
Ensemble modelling
R package
Evapotranspiration software

ABSTRACT

Evapotranspiration (ET) is a vital component of the hydrological cycle and there are a large number of alternative models for representing ET processes. However, implementing ET models in a consistent manner is difficult due to the significant diversity in process representations, assumptions, nomenclature, terminology, units and data requirements. An R package is therefore introduced to estimate actual, potential and reference ET using 17 well-known models. Data input is flexible, and customized data checking and pre-processing methods are provided. Results are presented as summary text and plots. Comparisons of alternative ET estimates can be visualized for multiple models, and alternative input data sets. The ET estimates also can be exported for further analysis, and used as input to rainfall-runoff models.

1. Introduction

Evapotranspiration (ET) is defined as the transfer of liquid water to the atmosphere as water vapor from bare soil and water bodies such as rivers and lakes (evaporation), as well as vegetated surfaces through plants' leaves (transpiration) (Allen et al., 1998; Dingman, 2015). ET is often one of the largest fluxes of water from catchments (Baumgartner et al., 1975), so that estimating its magnitude is critical for many applications. Factors that influence ET include: 1) the state of climate variables, such as temperature, relative humidity, wind speed and solar radiation, which influences the potential ET rate; 2) the water availability, which determines if actual evapotranspiration (AET) occurs at its potential rate (potential evapotranspiration, or PET) where sufficient water is present, or whether it occurs at a lower rate due to moisture limitations; and 3) the evaporative surface, with commonly modelled surfaces including natural catchments, 'reference' crops (ET_0), and open water bodies.

Understanding the dominant ET processes and quantifying ET rates provide useful information for diverse applications. For example, catchment management makes use of information on AET over the land surface, reservoir management requires information on open-water evaporation (e.g. McJannet et al., 2008), rainfall-runoff modelling often requires estimates of catchment-averaged PET (e.g. Andréassian et al., 2004; Oudin et al., 2005a,b), and agricultural studies often require information on ET_0 (e.g. Doorenbos, 1977; Shuttleworth and Wallace, 2009). However, obtaining observations of these specific ET rates can be challenging. This is because the measurement of AET is difficult, typically involving sophisticated spatial and temporal scaling techniques from sap flow observations to represent the entire canopy, or using expensive micrometeorological eddy flux instrumentation that is generally not available for most practical applications; furthermore, PET is a conceptual quantity that cannot be 'measured' directly (Gasca-Tucker et al., 2007; Fisher et al., 2011). Therefore, these rates are usually estimated using models, so that the selection and implementation of ET process models becomes critical.

There are multiple models available for estimating ET rates. According to McMahon et al. (2013), alternative ET models can represent the same ET processes differently by: (1) placing emphasis on different sub-processes, such as mass transfer and energy balance processes; (2) focussing on the dominant processes that occur in different environments, including humid and arid climates; (3) having different requirements for inputting climate data and different interpretations of the constants' values; and (4) conforming to different hierarchies for handling missing data and adjusting biased estimates.

In order to provide better information on the selection of an

* Corresponding author.
E-mail address: Danlu.Guo@adelaide.edu.au (D. Guo).

http://dx.doi.org/10.1016/j.envsoft.2015.12.019

appropriate model, guidance on ET model formulation and related issues was provided by McMahon et al. (2013). However, the implementation of these and other formulations is complicated by the significant diversity in process representations, assumptions, nomenclature, terminology, units and data requirements, which can make it difficult to implement the mathematical representations of these ET models, and can lead to coding inconsistencies and errors. This has a number of potentially negative implications on ET modelling studies, such as reducing confidence in the results presented, and providing difficulties for objectively comparing the results from different studies.

A practical aspect that can benefit from a more standardized approach to ET model implementation is the use of ensemble ET models. Applications of ensemble modelling can lead to a better understanding of ET model structural uncertainty (e.g. Beven and Freer, 2001; Duan et al., 2007; Kavetski and Fenicia, 2011; Velázquez et al., 2012), by:

1) assessing the impact of multiple ET models based on historical climate assumptions, to quantify PET and AET uncertainty (Xu and Singh, 2000, 2002; Tabari et al., 2013), and determine the effect of ET estimates on hydrologic modelling, water resource assessments (Yin and Brook, 1992; Oudin et al., 2005a,b; Kannan et al., 2007; Rosenberry et al., 2007; Horváth et al., 2010), ecological and agricultural studies (Nichols et al., 2004; Gasca-Tucker et al., 2007; Fisher et al., 2011), and;
2) assessing the impact of using multiple ET models under a changing climate, considering potential changes in both the ET-related processes and climate variables (McKenney and Rosenberg, 1993; Kay and Davies, 2008; Kingston et al., 2009; Donohue et al., 2010; Bormann, 2011; Prudhomme and Williamson, 2013; Thompson et al., 2014).

To further support a range of ET modelling studies, there is a need to facilitate the implementation of different ET models in a convenient, consistent and efficient manner. There are some existing software packages focussing on specific ET modelling needs and aspects: such as the 'ET_0 Calculator' (Raes and Munoz, 2008) to calculate ET_0 using the FAO-56 Penman-Monteith model, the Fortran code 'Morton WREVAP' (McMahon et al., 2013) to implement the Morton ET models, and the R package 'SPEI' (Beguería et al., 2013), which includes multiple ET models and several drought indices to estimate the Standardized Precipitation-Evapotranspiration Index (SPEI). However, to our knowledge, there has not been a freely available tool which enables the implementation of a large number of alternative ET models in a consistent manner.

This paper presents an R software package to estimate ET from 17 alternative models: fifteen of the models are based on those summarize in McMahon et al. (2013), as well as the Jensen-Haise and the McGuinness-Bordne models, sourced from Prudhomme and Williamson (2013). These estimate a range of ET quantities (AET, PET and ET_0), take a range of climate processes and variables into account, and run at daily or monthly time-steps. Data input is flexible and data checking and pre-processing options are included. The availability of such a consistent software framework for implementing modelling approaches is important from the perspective of ensemble modelling, comparison among different models and data sets (for examples see Dawson et al., 2007; Galelli et al., 2014), as well as analysis of model and input uncertainty (Leavesley et al., 2006; Clark et al., 2008; Andrews et al., 2011).

The remainder of this paper is organized as follows. The package is described in Section 2, including the evapotranspiration models included, as well as the package structure and core functions. In Section 3, two different Australian catchments are used to demonstrate various features of the package including: (1) data pre-processing; (2) estimation of ET and producing summaries and plots of results; and (3) comparison of estimates with ensemble ET models and input data sets. In Section 4, some potential further analyses with the package and limitations are discussed, which are followed by the conclusions in Section 5.

2. The evapotranspiration package

2.1. Evapotranspiration models

The R package *Evapotranspiration* includes 17 models, which use one or several climate variables to estimate PET, AET and ET_0 at a single location using input data at sub-daily, daily and monthly resolutions. Although the models consist of different process representations, they are all based on the two fundamental components that drive ET:

1) Energy balance, which determines the latent heat of vaporization; and
2) Mass transfer, which influences the rate of movement of water vapor away from the evaporating surface.

The latent heat can be estimated considering the energy balance as:

$$\lambda E = R - H - G + A_d \tag{1}$$

where λ is the latent heat of vaporization, E is the rate of evapotranspiration, R is the net incoming radiation received at the soil/plant surfaces (which is determined by the total incoming solar radiation R_s), H is the sensible heat exchange with the atmosphere through convection (which is determined by the air temperature T), G is the heat exchange with the ground, and A_d is the net input of water advected energy, such as water inflow to a lake, which only applies for open-water bodies.

The mass transfer of water vapor is influenced by the vapor gradient (i.e. the difference between saturated and actual vapor pressure, which is related to relative humidity RH and temperature T) and wind speed u_z. Next to the evaporative surface, a thin non-turbulent layer of air provides resistance to evaporation flux, known as the aerodynamic resistance (Penman, 1948). For plant leaves, surface resistance is also important, as transpiration is regulated by the degree of stomatal opening in leaves (Monteith, 1991). Combining the energy balance and mass transfer components, the four key climate variables related to ET are T, RH, R_s and u_z

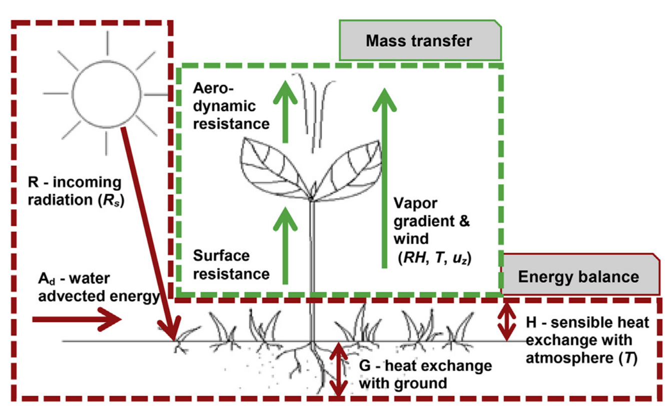

Fig. 1. ET-related processes accounted for by the mass transfer and energy balance, with the relevant atmospheric variables in brackets: T = air temperature, R_s = incoming solar radiation, RH = relative humidity, u_z = wind speed.

(as illustrated in Fig. 1).

Over the past decades, a large number of ET models have been developed by representing these processes in different ways. In this package, 17 of these models are included, which are based on different relationships among the ET processes and the four climate variables, and thus having different data requirements of climate variables and corresponding units (which are detailed in Table 1).

The various models included in the package *Evapotranspiration* are detailed in Table 1. The PET and ET_0 models consider different sets of ET sub-processes and associated climate variables, including incoming radiation, vapor gradient, the heat exchanges with the atmosphere and the ground, advection processes and the surface resistance of vegetation (see the references in Table 1 for further details). The five AET models (i.e. Brutsaert-Strickler, Granger-Gray, Szilagyi-Jozsa, Morton CRAE and Morton CRWE) are all based on an observed complementary relationship (CR, first raised by Bouchet, 1963) between PET and AET, which states that as the evaporating surface dries, the decrease in AET is complemented by an equal increase in PET. The two Morton models (Morton, 1983b, 1983a) can estimate both the PET and AET explicitly at the equilibrium temperature (i.e. the temperature at the evaporating surface), by following the energy-balance and vapor transfer equations, respectively. Alternatively, the Brutsaert-Strickler, Granger-Gray and Szilagyi-Jozsa methods estimate AET by integrating the Penman and Priestley-Taylor models within the CR framework in different ways (Brutsaert and Stricker, 1979; Granger and Gray, 1989; Szilagyi, 2007). Note that these quantities are equivalent under special conditions: technically, when sufficient water is present, the rate of PET and AET are equivalent to each other, and for a defined vegetated surface, the rate of PET and ET_0 are equivalent.

The equations for 15 ET models included in the package (all models except for Jensen-Haise and McGuinness-Bordne) are sourced from McMahon et al. (2013), which have all been verified with examples presented in their original paper. The availability of reliable verification is the key reason that we select the majority of ET models within this package from McMahon et al. (2013). For the other two structurally simple models, Jensen-Haise and McGuinness-Bordne which are sourced from Prudhomme and Williamson (2013), there are no published examples of implementation available for verification. We have ensured that the equations are correct by verifying their formulae in a number of alternative references including Jensen and Haise (1963), Xu and Singh (2000) and Oudin et al. (2005a).

2.2. Structure and core functions

The functions, data inputs and outputs, and graphical features of the package are summarized in Fig. 2. The data pre-processing function *ReadInputs()* is developed for loading and processing sub-daily and daily raw climate data. The processed data are then ready to feed into the generic function *ET…()*, where each of the 17 different methods can be called by substituting the '…' by the function name (e.g. '*ET.Penman ()*' to call the Penman model). The function performs calculations for the relevant ET model and generates a calculation summary.

Having calculated the ET quantity, the function *ETPlot()* can then be called to plot the original estimates, as well as aggregations and averages at different time scales. Function *ETComparison()* facilitates comparison of results and visualization of uncertainties from using different models and/or different input data. Finally, *ETForcing()* enables the association between estimated ET and different climate variables to be plotted.

Table 1
Data requirements for different models. D = daily, M = monthly.

ET model name and corresponding function name in package	Time step	Climate input data required[a]					Quantity estimated		
		T_{max}, T_{min}	RH_{max}, RH_{min}	R_s	U_z	T_{dew}	*PET*	*ET_0*	*AET*
Penman 1948 (Penman, 1948) and Penman 1956 (Penman, 1956) *ET.Penman*	D	✓	✓	✓	✓		✓		✓ (open water)
Penman-Monteith FAO-56 (Allen et al., 1998) and ASCE-EWRI (Allen et al., 2005) *ET.PenmanMonteith*	D	✓	✓	✓	✓			✓ (short crop)	
Matt-Shuttleworth (Shuttleworth and Wallace, 2009) *ET.MattShuttleworth*	D	✓	✓	✓	✓			✓ (well-watered)	
Priestley-Taylor (Priestley and Taylor, 1972) *ET.PriestleyTaylor*	D	✓	✓	✓			✓ (advection-free)		
PenPan[b] (Rotstayn et al., 2006) *ET.PenPan*	D	✓	✓	✓	✓		✓		
Brutsaert-Strickler (Brutsaert and Stricker, 1979) *ET.BrutsaertStrickler*	D	✓	✓	✓	✓				✓ (areal)
Granger-Gray (Granger and Gray, 1989) *ET.GrangerGray*	D	✓	✓	✓	✓				✓ (areal)
Szilagyi-Jozsa (Szilagyi, 2007) *ET.SzilagyiJozsa*	D	✓	✓	✓	✓				✓
Makkink (De Bruin, 1981)*ET.Makkink*	D	✓		✓				✓	
Blaney-Criddle (Allen and Pruitt, 1986) *ET.BlaneyCriddle*	D	✓	✓	✓	✓			✓ (well-watered)	
Turc (Turc, 1961) *ET.Turc*	D	✓		✓				✓	
Hargreaves-Samani (Hargreaves and Samani, 1985) *ET.HargreavesSamani*	D	✓						✓	
Chapman Australian[c] (Chapman, 2001) *ET.ChapmanAustralian*	D	✓	✓	✓	✓		✓		
Jensen-Haise (Jensen and Haise, 1963; Xu and Singh, 2000; Prudhomme and Williamson, 2013) *ET.JensenHaise*	D	✓		✓			✓		
McGuiness-Bordne (Oudin et al., 2005a; Prudhomme and Williamson, 2013) *ET.McGuinnessBordne*	D	✓					✓		
Morton CRAE (Morton, 1983a) *ET.MortonCRAE*	M	✓		✓		✓	✓		✓
Morton CRWE (Morton, 1983b) *ET.MortonCRWE*	M	✓		✓		✓	✓		✓ (shallow lake)

[a] T_{max}/T_{min} = maximum/minimum temperature (°C), R_s = incoming solar radiation ($MJ.m^{-2}$), RH_{max}/RH_{min} = maximum/minimum relative humidity (%), u_z = wind speed ($m.s^{-1}$), T_{dew} = dew point temperature (°C).
[b] The original PenPan model estimates the actual evaporation from a Class-A Pan (i.e. a circular pan with diameter of 1.2 m and depth of 0.25 m, which is constructed of galvanised iron and supported on a wooden frame at 30–50 mm above the ground). This rate of evaporation is closely related to the PET, so that is it possible to approximate PET from pan evaporation by adjustment using a pan coefficient (McMahon et al., 2013).
[c] The original Chapman model (Chapman, 2001) uses only the measurements of Class-A Pan evaporation and is therefore fully empirical. However, in the *Evapotranspiration* package, it has been adapted to utilize the outputs of the PenPan model so it can be considered to capture the same set of ET sub-processes as the PenPan model.

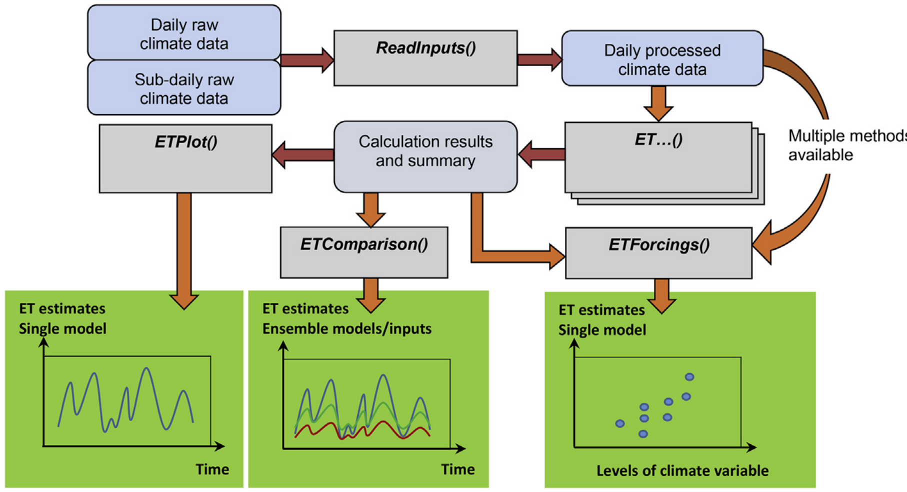

Fig. 2. Schematic diagram of the features of the package *Evapotranspiration:* the blue boxes represent data or results that are produced and/or processed by the functions, represented in the grey boxes; the green boxes represent expected results. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Function *ReadInputs()* is designed for checking data availability, and identifies missing entries and errors from the input sub-daily or daily raw climate data. The availability of the date data (i.e. year, month and date) is checked first, since these data are compulsory for the function to read the time-series-like climate data. *ReadInputs*() then reads through the raw climate data presented, and reports all input variables that are available to use. Specific data requirements for the individual models (see Table 1) are checked prior to performing the calculations in function *ET…()*. A specific format of the input data is required in terms of variable names and units, as well as the input data file format, which is different for daily and sub-daily raw data (see Section 1.1 of the supplementary material, within which Table 1 provides the detailed format requirements for the raw climate data). To assist users with preparing the raw input data, a summary of the relevant unit conversions is also provided in Table 2 of the supplementary material.

Next, *ReadInputs()* checks for missing entries in each of the available climate variables, and the quality of the data is assessed against two user-defined threshold values for: (1) the maximum acceptable percentage of missing data; (2) the maximum acceptable duration of continuous missing data as a percentage of total data duration. If the data quality is not acceptable (i.e. either of the percentage and/or duration of missing data has exceeded the user-defined threshold values), the program will be terminated with a warning message.

For data with acceptable quality but still containing some missing values, a warning is given with a default of assigning 'NA' for the missing values (which leads to 'NA's in the output estimates if they are used in *ET…()*). The user can also use the in-built gap-filling routine to interpolate for the missing values, with four alternative gap-filling methods (see Table 3 in the supplementary material for details) including:

1) Replacement with same-month average (adapted from Narapusetty et al., 2009);
2) Replacement with same-season average (adapted from Narapusetty et al., 2009);
3) Replacement with same day-of-the-year average (Narapusetty et al., 2009);
4) Interpolation between the two bounding values, which is only suitable for missing time increments in which values are available at adjacent increments (McMahon et al., 2013). When there is more than one consecutive missing entry, this interpolation fails, with a warning given.

The function also includes simple primary checks for abnormal values in each climate variable: for example, any temperature data greater than 100 °C are considered as abnormal. Warnings are issued for the abnormal values detected, and again, the users can choose if the abnormal values will be corrected in the function, using one of the four interpolation methods mentioned previously. Details of the four interpolation methods and definitions of abnormal values for each climate variable are presented in Tables 4 and 5 in the supplementary material.

After completing the quality checks, all sub-daily raw data are aggregated to a daily time-step, as required by most ET models; such temporal aggregation is not performed for raw climate data that are already available at a daily time-step.

As already discussed, *ET…()* is a generic function, which includes 17 different specific methods that are all named following the format of *ET.methodname()*, as detailed in Table 1. If a specific ET model is selected by user, the function first performs a specific check for the data requirement, which is different for each of the 17 models in *ET…()* (see Table 1 for details). If a certain input variable required by the ET model is not available, the function will search for whether there are alternative ways to estimate the missing variable from other available variables; however, if no alternative data or methods are available, the function will be terminated with a warning. The available methods to estimate missing input variables are summarized in Table 3 in the supplementary material.

In the case where a specific ET model is not specified (i.e. the generic function *ET()* is called directly instead of *ET.methodname()*), the first task *ET…()* performs is to estimate as many missing climate variables as possible. Then a default method to estimate ET is selected based on all the available input variables, which include both the original input variables presented, as well as the variables estimated from them. Wherever data are available, the ET model that has the highest data requirements (and thus provides the most detailed physically-based process representation) is selected as default. The detailed selection of default models for different data availability is in Table 6 of the supplementary material.

Besides the input climate data, a list of constants is also required in *ET…()*. The definitions and suggested values of all constants are summarized in Table 7 in the supplementary material. A number of arguments are included in each ET model to allow additional user decisions in modelling ET. A common argument in all models is the choice of time-step for the output. The default time-step of the output ET estimates is daily for all models running at a daily time-step (i.e. all models except for Morton, as shown in Table 1), however, monthly and annual outputs can also be produced when specified; the two Morton models by default produce monthly output, while the user can also choose to obtain annual output. For models with multiple versions (e.g. the Penman, 1948, 1956 models, which have different wind functions) and requiring additional user decisions (e.g. calculation options, assumptions) there are additional individual arguments to enable flexible choices among different pathways. The complete details on the use of constants and available arguments for different ET models are presented in Table 8 in the supplementary material.

Once being called with sufficient data, and provided all constants and arguments have been specified, function *ET...()* performs calculations for individual ET models. A user-friendly summary of the results is printed on the screen, which confirms the choice of model and sub-model, along with the corresponding versions, the quantities calculated, as well as options for alternative calculations and assumptions. A basic statistical summary of the entire output time-series is also presented (as illustrated with an example in Fig. 3). The full results are stored as an R list file, as well as a csv file, which is automatically saved to the working directory. It contains both the calculation summary and the entire time series of the output, in which the ET estimates are organized in rows for different time increments.

A number of plotting tools are available to analyse the outputs. Function *ETPlot()* uses the estimated daily ET from individual ET models to generate aggregation plots and average plots at daily, monthly and annual time steps. Function *ETComparison()* produces comparison plots of different sets of ET estimates, to compare the outputs from (1) different ET models; (2) different versions of the same ET model (e.g. the 1948 and 1956 versions of the Penman model); (3) the same ET model with different calculation options, such as alternative approaches for data infilling and/or; (4) different sets of input climate data. For each quantity, three types of plots, including time series plots, non-exceedance probability plots and box plots, can be produced. Plots of uncertainty ranges can also be produced for daily estimates, monthly and annual aggregates and monthly and annual averages. Finally, the function *ETForcing()* is an additional plotting tool for visualizing the association between estimated ET and different climate variables within existing data.

3. Case studies

Two case studies have been used to demonstrate the core utilities of the package *Evapotranspiration*, using sub-daily climate data from meteorological sites at Adelaide (34.9290° S, 138.6010° E) and Alice Springs (23.7000° S, 133.8700° E) in Australia for the common period from 01/01/1989 to 30/03/2005.

3.1. Basic features: pre-processing input data, calculating and visualizing estimates

ReadInputs() is called first with the raw sub-daily data from the Adelaide case study, and the maximum percentage of acceptable number and duration of missing data set to 10% and 3%, respectively. The function displays a summary of data quality when checking through each input variable. The raw sub-daily data are then aggregated to a daily timescale. The missing values and abnormal values in each input variable are corrected with the corresponding averages from the same days of the year (i.e. day-of-the-year average). The processed data are then ready for the ET models to use.

The Penman open-water ET is estimated for the Adelaide case study using function *ET.Penman()*. The arguments are set so that (1) the time-step for calculation is daily; (2) the actual sunshine hours are used for calculating solar radiation; (3) the actual wind data are used; (4) the Penman 1948 wind function (Penman, 1948) is used to estimate the mass transfer component in the Penman model; and (5) the evaporative surface is open water (albedo = 0.08, roughness height = 0.001 m). The calculated time series of Penman ET from *ET.Penman()* has been saved in an R data list, while output is printed to the screen, which confirms the choice of model and the selection of alternative calculation options, and also gives a basic statistical summary of the entire time-series of ET estimates.

Fig. 3 is a screenshot of data processing and ET estimation with *ReadInputs()* and *ET.Penman()* for this case study.

The plots of estimated daily ET and monthly averaged daily ET have been produced for the Adelaide case study using function *ETPlot()* (Fig. 4). Although it is difficult to detect any trend from the highly fluctuating daily estimates (Fig. 4a), there is a very strong seasonal pattern, displayed in the monthly average plot (Fig. 4b). The ET peaks during the summer months, as would be expected due to the higher temperature and solar radiation during this time of the year.

```
> data <- ReadInputs(climatedata,constants,stopmissing=c(10,3), timestep = "subdaily",
+                    interp_missing = T, interp_abnormal = T,
+                    missing_method = "DoY average", abnormal_method = "DoY average")
The maximum acceptable percentage of missing data is 10 %
The maximum acceptable percentage of continuous missing data is 3 %
Warning: missing data of 'Tmax.daily'(daily maximum temperature), calculated from subdaily 'Temp.subdaily'
Warning: missing data of 'Tmin.daily'(daily minimum temperature), calculated from subdaily 'Temp.subdaily'
Warning: missing data of 'u2.subdaily', calculated from 'uz.subdaily
Number of missing values in uz.subdaily:  3
% missing data:  0.03 %
Maximum duration of missing data as percentage of total duration:  0.02 %
Warning: missing data of 'RHmax.daily'(daily maximum relative humidity), calculated from subdaily 'RH.subdaily'
Warning: missing data of 'RHmin.daily'(daily minimum relative humidity), calculated from subdaily 'RH.subdaily'
> Results <- ET.Penman(data, constants, solar="sunshine hours", wind="yes", windfunction_ver = "1948", alpha = 0.08, z0 = 0.001)
Penman Open-water Evaporation
Evaporative surface: water, albedo = 0.08 ; roughness height = 0.001 m
Sunshine hour data have been used for calculating incoming solar radiation
Wind data have been used for calculating the Penman evaporation. Penman 1948 wind function has been used.
Timestep: daily
Units: mm
1280 ET estimates obtained
Time duration: 2001-03-01 to 2004-08-31
Basic stats
Mean: 4.86
Max: 12.78
Min: 1.03
```

Fig. 3. Example of a typical session of data processing with *ReadInputs()* and ET estimation with *ET.Penman()* for the Adelaide case study.

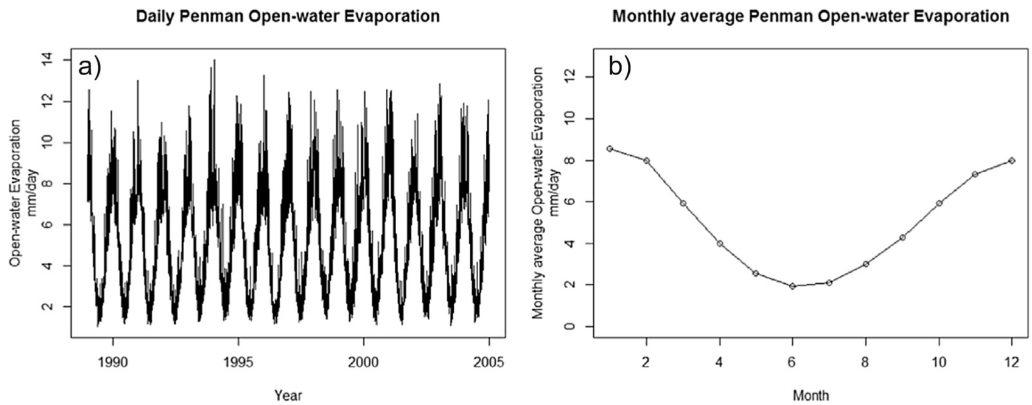

Fig. 4. a) Daily estimates of Penman open-water ET (left panel); b) Monthly averaged daily Penman open-water ET (right panel) for the Adelaide case study, generated by *ETPlot()*.

3.2. Advanced features: analyses with ensemble models and different input data sets

The features of function *ETComparison()* are demonstrated for both the Adelaide and Alice Springs case studies. First, plots of the time series and the non-exceedance probabilities for monthly ET estimates have been produced to compare estimates from the Penman-Monteith FAO56 and Priestley-Taylor models during 1989–1991 (Fig. 5). From Fig. 5a and b we observed that:

1) When comparing across the two case study sites, the inter-model differences in estimates are greater at Alice Springs, with the Priestley-Taylor model producing consistently lower estimates than the Penman-Monteith, and;
2) When comparing the seasonal patterns, the inter-model differences in estimates are most significant for the peak estimates, which occur in every summer (for example, at the start of 1990).

Fig. 5c shows the distribution of the monthly estimates within the period and from the different models, which is consistent with previous observations: the ET estimates from the Priestley-Taylor model are consistently lower compared with those obtained using the Penman-Monteith model, with the greatest difference of approximately 100 mm for the peak estimates at Alice Springs. These results reflect the structural differences in the two models, as the Penman-Monteith model explicitly takes the mass transfer for evapotranspiration into account, which is higher during summer periods and for arid and windy conditions (as reported in McKenney and Rosenberg, 1993; Yin et al., 2010), such as those experienced in Alice Springs.

Another application of *ETComparison()* is demonstrated in Fig. 6, in which the effect of uncertainties in input climate data under climate changes are shown for the Adelaide case study, together with the model uncertainty. To maintain the simplicity and clarity of the example, we focus only on the potential uncertainties in the future temperature due to climate change, without considering the probability of individual changes or potential variations in other climate variables. We perturb the existing temperature data within a range of 0 to +8 °C, which is considered to encompass all plausible future changes in temperature in Australia by 2100 (Stocker et al., 2013). Within this range, 500 random samples have been drawn and the corresponding perturbations are applied to the historical time series of T_{max} and T_{min}, resulting in 500 sets of input climate data. These 500 sets were then fed into *ETComparison()* to generate the corresponding ET outputs using both the Penman-Monteith and Priestley-Taylor models. The resulting ranges of the monthly ET estimates for the period between 1989 and 1999 are shown in Fig. 6. As can be seen, there is a greater range in ET estimates from the Penman-Monteith model. Since both ET models use the same temperature data as inputs, this indicates that temperature has a greater impact on the ET estimates obtained using the Penman-Monteith model—a pattern also observed in McKenney and Rosenberg (1993). This difference can be due to the structural differences between the two models: as the Penman-Monteith explicitly considers the mass transfer processes that are related to temperature, the importance of temperature is higher in the Penman-Monteith model.

It is worth mentioning that all climate variables other than temperature (i.e. *RH*, R_s and u_z) are kept at their current levels in this example, which is unrealistic under future conditions. Therefore, the results should only be considered as illustrative of the key feature of function *ETComparison()*, as a tool to compare ET estimates from multiple input data sets and ET models. In a formal assessment of the impact of climate-related input uncertainty, it is necessary to consider the potential uncertainty in the full set of climate variables that influence ET (Goyal, 2004; Whateley et al., 2014).

4. Discussion

4.1. Further analyses with other software packages

The output from this package is formatted as time-series-like data in the *zoo* format (in which every data point is linked to a specific time point, see Zeileis et al., 2015), so it can be easily extracted and used as an input to other R-based software packages for a range of further analyses. Examples include using ET estimates as input to hydrologic models in *hydromad* (Andrews et al., 2011), and to investigate the sensitivity of ET estimates to changes in the input climate data using *sensitivity* (Pujol et al., 2014). Since the output is also saved to a csv file (as detailed in Section 2.2), it can also be imported to external software packages.

4.2. Limitations

Although the package provides features for checking missing values and errors in the input climate data, as well as interpolation methods for these problematic data, caution is required to minimize the risk of misuse. In developing this package, we have tested the data processing tools with our own test data sets, as well as a number of user-provided data sets, and we have ensured that the package runs free of errors with these existing data sets. However, since every data set is different, it is recommended that users should exercise their own quality-control procedure prior to using

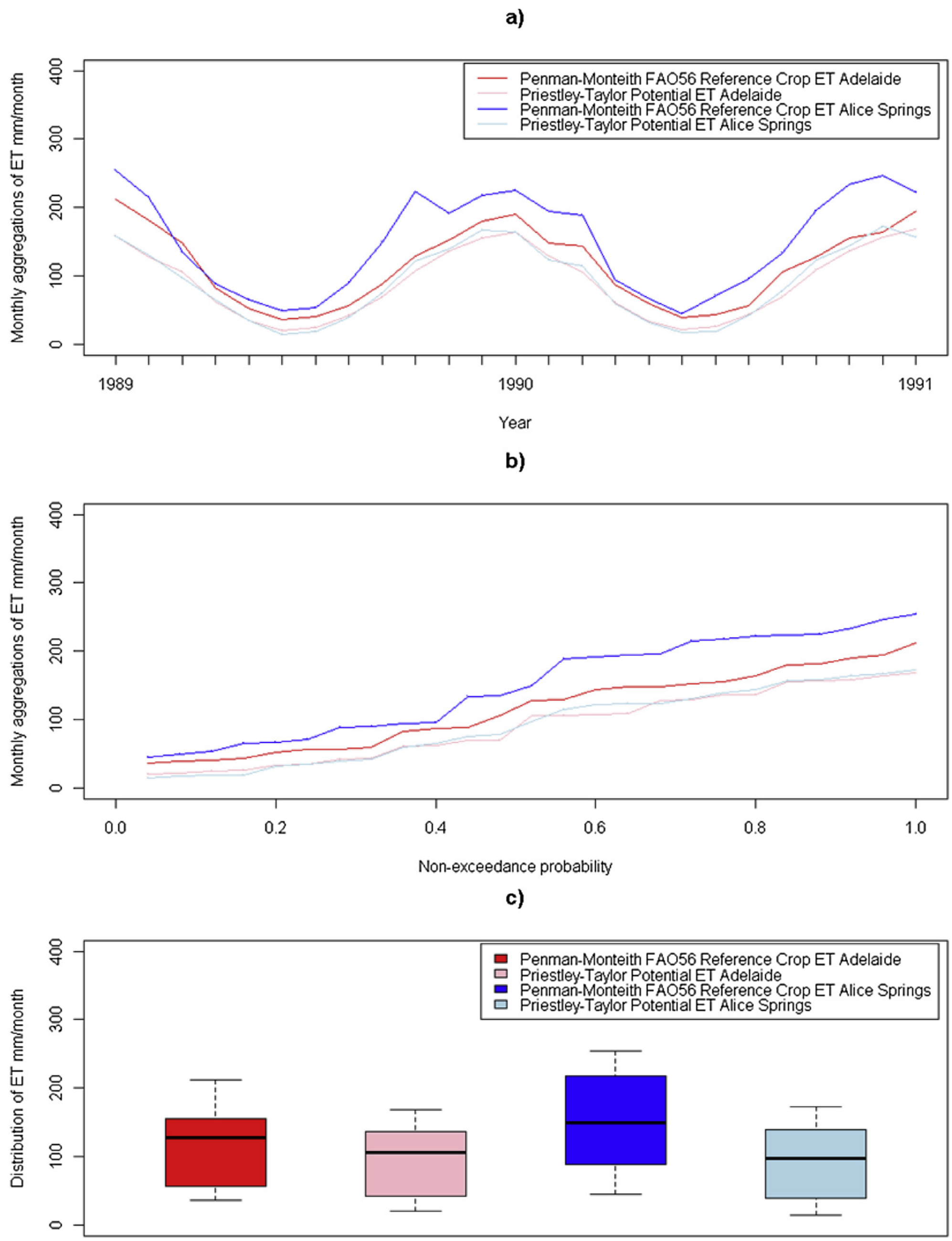

Fig. 5. Comparison of monthly ET estimates from two models (Penman-Monteith FAO56 and Priestley-Taylor) and two locations (Adelaide and Alice Springs) using *ETComparison()* for a) time-series; b) non-exceedance probability, and; c) distribution.

the package, to ensure that best-quality data are provided for ET estimation and the impact of data quality on the estimates is minimized.

Users should also be aware of the full assumptions and limitations prior to using any ET model in this package. Almost every ET model contains assumptions relating to the specific climate conditions under which the models apply. For example, some models assume that sub-processes related to ET are negligible, while other models are only calibrated to the climate of a specific region (a full list of assumptions and limitations for each individual model is given in Table 9 in the supplementary material, which is summarized from the existing literature). These assumptions limit the models' ability to generalize to a wider range of climate zones, leading to varying performance of ET models under different climate settings (Rosenberry et al., 2007; Tabari et al., 2013). A further problem arises if the models are to be applied to estimate ET under climate change conditions, which can mean that existing ET processes and related climate variables are likely to be different to those for which the models are best suited, potentially causing deteriorating model performance (Prudhomme and Williamson, 2013; Thompson et al., 2014).

5. Summary and conclusions

This paper presents an R package *Evapotranspiration* for the estimation of actual, potential and reference crop ET using 17 models in a consistent, convenient and efficient manner. The pre-processing tool provides flexible methods for checking and

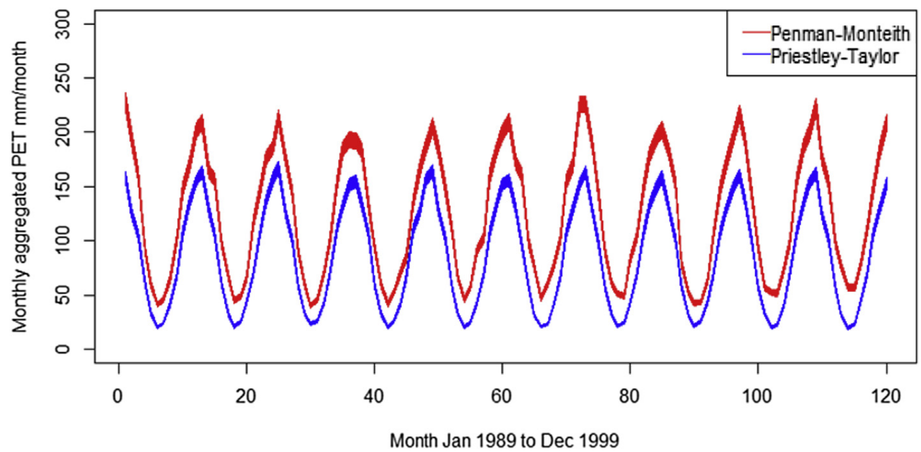

Fig. 6. Uncertainties in monthly ET estimates from two models (Penman-Monteith FAO56 and Priestley-Taylor) at Adelaide, each executed for 500 sets of input data sampled with 0 to +8 °C uncertainty in temperature, generated by *ETComparison()*.

processing raw input climate data, which are then fed into user-selected ET models. The presentation of results is in the form of both summary text and plots. Comparison between multiple ET models and input data sets is also supported. Estimates from the package can be conveniently extracted for further analysis, such as rainfall-runoff modelling and sensitivity analyses. It is hoped that this package will increase consistency in the results presented in ET studies, and increase our ability to investigate the impact of structural uncertainty in ET model formulations via the use of ensemble modelling.

Acknowledgements

The authors wish to thank Murray Peel and Thomas McMahon for their valuable comments on delineating the scope of the paper, and to Joseph Guillaume and an anonymous reviewer for their thoughtful comments on the manuscript.

Software availability

Description: Package *Evapotranspiration*.
Developers: Danlu Guo, Seth Westra.
Year First Available: 2014.
E-mail: Danlu.guo@adelaide.edu.au.
Website:http://cran.r-project.org/web/packages/Evapotranspiration/index.html.
Hardware Requirement: General-purpose computer.
Software Requirement: R version 2.10 or later.
Programming Language: R.

Licensing

This software is made freely available under the terms and conditions of the GNU General Public License.

Appendix A. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.envsoft.2015.12.019.

References

Allen, R., Pruitt, W., 1986. Rational use of the FAO Blaney-Criddle formula. J. Irrigation Drainage Eng. 112 (2), 139–155.
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, vol. 300. FAO, Rome, p. 6541.
Allen, R.G., Walter, I.A., Elliott, R.L., Howell, T.A., Itenfisu, D., Jensen, M.E., Snyder, R.L., 2005. The ASCE Standardized Reference Evapotranspiration Equation. ASCE Publications.
Andréassian, V., Perrin, C., Michel, C., 2004. Impact of imperfect potential evapotranspiration knowledge on the efficiency and parameters of watershed models. J. Hydrology 286 (1–4), 19–35.
Andrews, F., Croke, B., Jakeman, A., 2011. An open software environment for hydrological model assessment and development. Environ. Model. Softw. 26 (10), 1171–1185.
Baumgartner, A., Reichel, E., Lee, R., 1975. The World Water Balance: Mean Annual Global, Continental and Maritime Precipitation, Evaporation and Run-off. Elsevier Scientific Publishing Company.
Beguería, S., Vicente-Serrano, S.M., Beguería, M.S., 2013. Package 'SPEI'.
Beven, K., Freer, J., 2001. Equifinality, data assimilation, and uncertainty estimation in mechanistic modelling of complex environmental systems using the GLUE methodology. J. Hydrology 249 (1–4), 11–29.
Bormann, H., 2011. Sensitivity analysis of 18 different potential evapotranspiration models to observed climatic change at German climate stations. Clim. Change 104 (3–4), 729–753.
Bouchet, R.J., 1963. Evapotranspiration réelle et potentielle, signification climatique. In: Paper Presented at General Assembly Berkeley, Gentbrugge, Belgium.
Brutsaert, W., Stricker, H., 1979. An advection-aridity approach to estimate actual regional evapotranspiration. Water Resour. Res. 15 (2), 443–450.
Chapman, T., 2001. Estimation of evaporation in rainfall-runoff models. In: Ghassemi, F., Post, D., Sivapalan, M., Vertessy, R. (Eds.), MODSIM2001: Integrating Models for Natural Resources Management across Disciplines, Issues and Scales, MSSANZ, vol. 1, pp. 293–298.
Clark, M.P., Slater, A.G., Rupp, D.E., Woods, R.A., Vrugt, J.A., Gupta, H.V., Wagener, T., Hay, L.E., 2008. Framework for Understanding Structural Errors (FUSE): a modular framework to diagnose differences between hydrological models. Water Resour. Res. 44 (12).
Dawson, C.W., Abrahart, R.J., See, L.M., 2007. HydroTest: a web-based toolbox of evaluation metrics for the standardised assessment of hydrological forecasts. Environ. Model. Softw. 22 (7), 1034–1052.
De Bruin, H., 1981. The determination of (reference crop) evapotranspiration from routine weather data. Evaporation Relat. Hydrology 25–37.
Dingman, S.L., 2015. Physical Hydrology. Waveland Press.
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. Hydrology 386 (1), 186–197.
Doorenbos, J., 1977. Guidelines for predicting crop water requirements. FAO Irrigation Drainage Pap. 24, 15–29.
Duan, Q., Ajami, N.K., Gao, X., Sorooshian, S., 2007. 'Multi-model ensemble hydrologic prediction using Bayesian model averaging. Adv. Water Resour. 30 (5), 1371–1386.
Fisher, J.B., Whittaker, R.J., Malhi, Y., 2011. ET come home: potential evapotranspiration in geographical ecology. Glob. Ecol. Biogeogr. 20 (1), 1–18.
Galelli, S., Humphrey, G.B., Maier, H.R., Castelletti, A., Dandy, G.C., Gibbs, M.S., 2014. An evaluation framework for input variable selection algorithms for environmental data-driven models. Environ. Model. Softw. 62 (0), 33–51.
Gasca-Tucker, D., Acreman, M., Agnew, C., Thompson, J., 2007. Estimating evaporation from a wet grassland. Hydrology Earth Syst. Sci. 11 (1).
Goyal, R.K., 2004. Sensitivity of evapotranspiration to global warming: a case study of arid zone of Rajasthan (India). Agric. Water Manag. 69 (1), 1–11.
Granger, R.J., Gray, D., 1989. 'Evaporation from natural nonsaturated surfaces. J. Hydrology 111 (1), 21–29.

Hargreaves, G.H., Samani, Z.A., 1985. 'Reference crop evapotranspiration from ambient air temperature. Am. Soc. Agric. Eng. no. fiche no. 85-2517.
Horváth, S., Szép, I.J., Makra, L., Mika, J., Pajtók-Tari, I., Utasi, Z., 2010. Effect of evapotranspiration parameterisation on the Palmer drought severity Index. Phys. Chem. Earth, Parts A/B/C 35 (1–2), 11–18.
Jensen, M.E., Haise, H.R., 1963. 'Estimating evapotranspiration from solar radiation', Proceedings of the American Society of Civil Engineers. J. Irrig. and Drain. Div. 89, 15–41.
Kannan, N., White, S.M., Worrall, F., Whelan, M.J., 2007. Sensitivity analysis and identification of the best evapotranspiration and runoff options for hydrological modelling in SWAT-2000. J. Hydrology 332 (3–4), 456–466.
Kavetski, D., Fenicia, F., 2011. Elements of a flexible approach for conceptual hydrological modeling: 2. Application and experimental insights. Water Resour. Res. 47 (11), W11511.
Kay, A.L., Davies, H.N., 2008. Calculating potential evaporation from climate model data: a source of uncertainty for hydrological climate change impacts. J. Hydrology 358 (3–4), 221–239.
Kingston, D.G., Todd, M.C., Taylor, R.G., Thompson, J.R., Arnell, N.W., 2009. Uncertainty in the estimation of potential evapotranspiration under climate change. Geophys. Res. Lett. 36 (20), L20403.
Leavesley, G.H., Markstrom, S.L., Viger, R.J., 2006. USGS modular modeling system (MMS)-precipitation-runoff modeling system (PRMS). Watershed models 159–177.
McJannet, D., Webster, I., Stenson, M., Sherman, B., 2008. Estimating Open Water Evaporation for the Murray-Darling Basin: a Report to the Australian Government from the CSIRO Murray-Darling Basin Sustainable Yields Project.
McKenney, M.S., Rosenberg, N.J., 1993. Sensitivity of some potential evapotranspiration estimation methods to climate change. Agric. For. Meteorology 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.
Monteith, J., 1991. 'Weather and water in the Sudano-Sahelian zone'. In: Sivakumar, M.V.K., Wallace, J.S., Renard, C., Giroux, C. (Eds.), Soil Water Balance in the Sudano-Sahelian Zone, Proceedings of the Niamey Workshop, Niamey, Niger, February 1991, Niamey, Niger, pp. 11–29. IAHS Publ. no. 199.
Morton, F.I., 1983a. Operational estimates of areal evapotranspiration and their significance to the science and practice of hydrology. J. Hydrology 66 (1–4), 1–76.
Morton, F.I., 1983b. Operational estimates of lake evaporation. J. Hydrology 66 (1–4), 77–100.
Narapusetty, B., DelSole, T., Tippett, M.K., 2009. Optimal estimation of the climatological mean. J. Clim. 22 (18), 4845–4859.
Nichols, J., Eichinger, W., Cooper, D., Prueger, J., Hipps, L., Neale, C., Bawazir, A., 2004. Comparison of Evaporation Estimation Methods for a Riparian Area. Final report. UHR Technical Report, no. 436.
Oudin, L., Hervieu, F., Michel, C., Perrin, C., Andréassian, V., Anctil, F., Loumagne, C., 2005a. Which potential evapotranspiration input for a lumped rainfall–runoff model?: Part 2—Towards a simple and efficient potential evapotranspiration model for rainfall–runoff modelling. J. Hydrology 303 (1–4), 290–306.
Oudin, L., Michel, C., Anctil, F., 2005b. Which potential evapotranspiration input for a lumped rainfall-runoff model?: Part 1—Can rainfall-runoff models effectively handle detailed potential evapotranspiration inputs? J. Hydrology 303 (1–4), 275–289.
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.
Penman, H.L., 1956. Evaporation: an introductory survey. Neth. J. Agric. Sci. 4, 9–29.
Priestley, C., Taylor, R., 1972. On the assessment of surface heat flux and evaporation using large-scale parameters. Mon. Weather Rev. 100 (2), 81–92.
Prudhomme, C., Williamson, J., 2013. Derivation of RCM-driven potential evapotranspiration for hydrological climate change impact analysis in Great Britain: a comparison of methods and associated uncertainty in future projections. Hydrology Earth Syst. Sci. 17 (4), 1365–1377.
Pujol, G., Iooss, B., Iooss, M.B., 2014. Package ‘sensitivity’.
Raes, D., Munoz, G., 2008. The ETo calculator. Land water digital media Ser.
Rosenberry, D.O., Winter, T.C., Buso, D.C., Likens, G.E., 2007. Comparison of 15 evaporation methods applied to a small mountain lake in the northeastern USA. J. Hydrology 340 (3), 149–166.
Rotstayn, L.D., Roderick, M.L., Farquhar, G.D., 2006. A simple pan-evaporation model for analysis of climate simulations: evaluation over Australia. Geophys. Res. Lett. 33 (17).
Shuttleworth, W., Wallace, J., 2009. Calculating the water requirements of irrigated crops in Australia using the Matt-Shuttleworth approach. Trans. ASABE 52 (6), 1895–1906.
Stocker, T.F., Qin, D., Plattner, G.-K., Tignor, M., Allen, S.K., Boschung, J., Nauels, A., Xia, Y., Bex, V., Midgley, P.M., 2013. Climate change 2013: the physical science basis. In: Intergovernmental Panel on Climate Change, Working Group I Contribution to the IPCC Fifth Assessment Report (AR5). Cambridge Univ Press, New York.
Szilagyi, J., 2007. On the inherent asymmetric nature of the complementary relationship of evaporation. Geophys. Res. Lett. 34 (2), L02405.
Tabari, H., Grismer, M.E., Trajkovic, S., 2013. Comparative analysis of 31 reference evapotranspiration methods under humid conditions. Irrigation Sci. 31 (2), 107–117.
Thompson, J.R., Green, A.J., Kingston, D.G., 2014. Potential evapotranspiration-related uncertainty in climate change impacts on river flow: an assessment for the Mekong River basin. J. Hydrology 510 (0), 259–279.
Turc, L., 1961. Estimation of irrigation water requirements, potential evapotranspiration: a simple climatic formula evolved up to date. Ann. Agron. 12 (1), 13–49.
Velázquez, J.A., Schmid, J., Ricard, S., Muerth, M.J., Gauvin St-Denis, B., Minville, M., Chaumont, D., Caya, D., Ludwig, R., Turcotte, R., 2012. An ensemble approach to assess hydrological models' contribution to uncertainties in the analysis of climate change impact on water resources. Hydrology Earth Syst. Sci. 17 (2013), 565–578.
Whateley, S., Steinschneider, S., Brown, C., 2014. A climate change range-based method for estimating robustness for water resources supply. Water Resour. Res. 50 (11), 8944–8961.
Xu, C.Y., Singh, V.P., 2000. Evaluation and generalization of radiation-based methods for calculating evaporation. Hydrol. Process. 14 (2), 339–349.
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.
Yin, Y., Wu, S., Chen, G., Dai, E., 2010. Attribution analyses of potential evapotranspiration changes in China since the 1960s. Theor. Appl. Climatol. 101 (1–2), 19–28.
Yin, Z.-Y., Brook, G.A., 1992. Evapotranspiration in the Okefenokee Swamp watershed: a comparison of temperature-based and water balance methods. J. Hydrology 131 (1), 293–312.
Zeileis, A., Grothendieck, G., Ryan, J.A., Andrews, F., Zeileis, M.A., 2015. Package ‘zoo’.