http://doi.org/10.1590/1678-992X-2021-0229

ISSN 1678-992X

SCIENTIA AGRICOLA

Research Article

Improvement of reference crop evapotranspiration estimates using limited data for the Brazilian Cerrado

Daniel Althoff[1]*, Lineu Neiva Rodrigues[2]

[1]Universida Federal de Viçosa – Depto. de Engenharia Agrícola, Av. Peter Henry Rolfs, s/n – 36570-900 – Viçosa, MG – Brasil.
[2]Embrapa Cerrados, BR-020, km 18 – 73310-970 – Planaltina, DF – Brasil.
*Corresponding author <daniel_althoff@hotmail.com>

Edited by: Paulo Cesar Sentelhas / Thiago Libório Romanelli

Received September 10, 2021
Accepted March 28, 2022

ABSTRACT: Evapotranspiration (ET) is a key component of the hydrological cycle. Therefore, adequately estimating it is crucial to improving water resource planning and management. One of the most affordable methods of estimating ET is first to estimate reference crop evapotranspiration (ETo) and later associate it to crop and soil coefficients. The FAO Kc-ETo approach can be used only when ETo is computed with the FAO Penman-Monteith equation. However, low data availability may restict the equations used to estimate ETo. In this study, we assess and calibrate common methods used to estimate ETo under such conditions of limited data availability. Based on the annual calibration, the Makkink (NSE = 0.85) outperformed the Priestley-Taylor (NSE = 0.73), Hargreaves-Samani (NSE = 0.56), and Penman-Monteith temperature approach (NSE = 0.58). The seasonal calibration of parameters showed no significant improvement to the methods assessed ($\Delta NSE \leq 0.01$), except for the Priestley-Taylor ($\Delta NSE = 0.06$). The performance of temperature-based equations was particularly limited due to the performance of the equation adopted to estimate global solar radiation. Thus, improving the representation of global solar radiation for limited data availability can also play a key role in improving ETo prediction.

Keywords: Penman-Monteith temperature approach, tropical savannah, estimating dew-point temperature, average wind speed

Introduction

In Brazil, disputes over water use seem to be more widespread in the northeastern and southeastern regions, where rainfall is highly variable. This being the case, special attention should be given to the savannah biome (Cerrado), which represents ~24 % of Brazil's territory, and is the most important and the last agricultural frontier in Brazil and one of the few areas with the potential to increase food production sustainably (Klink, 2014; Wendt et al., 2015).

In this context, crop evapotranspiration (ETc) is a crucial component of the hydrological cycle and water balance in the soil. Thus, obtaining more accurate evapotranspiration (ET) estimates is crucial to the development of more efficient water source management, irrigation practices, and trustworthy climate studies. Because of the high costs associated with obtaining direct ET observations, it is commonly estimated by physics-based or empirical equations. ET can be estimated from meteorological data in the case of crops by first calculating a reference crop evapotranspiration (ETo) as defined in the Food and Agriculture Organization Irrigation and Drainage paper 56 (FAO56, Allen et al., 1998). Crop evapotranspiration results from adjusting ETo to water-stress conditions and specific crop coefficients.

The Penman-Monteith (PM) equation parametrized in FAO56 (Allen et al., 1998) has long been used as the standard method for estimating daily ETo, which requires data on solar radiation (or sunshine hours), air temperature, relative humidity, and wind speed. However, many regions worldwide lack one or more meteorological inputs required by the PM equation (Zanetti et al., 2019).

Several empirical equations have been developed to provide ETo estimates under conditions of limited meteorological data availability, such as the Hargreaves and Samani (1985) (HS), Priestley and Taylor (1972) (PT), and Makkink (1957) (MK) equations. These equations were developed considering specific climatic conditions and may require calibration to adequately estimate ETo in other locations and under different conditions (Gavilán et al., 2006).

Given the importance of both providing more accurate estimates for environmental and hydrological studies, and developing sustainable irrigation management practices, the objectives of the present study were to calibrate new coefficients for the PT, MK, HS, and the PM temperature approach (PMT) methods, on both the annual and seasonal scales for the entire Cerrado region. Although solar radiation data is less commonly available, temperature-radiation-based methods were included in this study to compare with temperature-based methods that might have errors from estimating solar radiation prior to the ETo calculation.

Materials and Methods

Study area

The Cerrado is Brazilian's second-largest biome, covering 23 % of the territory and extending into 11 states. The geographic coordinates are from 60°28′22″

CC BY

W to 41°16′48″ W and from 2°19′48″ S to 24°40′55″ S (Figure 1), and its altitude ranges from 0 m to 2854 m above sea level (mean altitude: 396 m). Tropical grasslands and savannah characterize vegetation in the Cerrado. The region's climate is predominantly classified as tropical with dry winter (Alvares et al., 2013). Thus, the region is characterized by two well-defined seasons: a wet season from Oct to Apr, when ~85 % to 90 % of the annual rainfall occurs, and a dry season from May to Sept.

Data

Data from 2000 to 2020 featuring 82 automatic weather stations (AWS) and 50 conventional weather stations (Figure 1) were used to assess reference crop evapotranspiration equations for the Cerrado. Data were acquired from the Brazilian National Institute of Meteorology (INMET) Meteorological Database. The INMET's weather stations were installed and are maintained according to the World Meteorological Organization guide to meteorological instruments and observation methods (WMO, 2008). Although there is no information pertaining to the stations' management of the grass surface, we adopted these as reference sites. The stations are listed in Appendix I together with their geographic coordinates, elevation, and number of observations during the study period.

INMET's CWS records daily data consisted of daily mean, maximum and minimum air temperature (T_{avg}, T_{max}, and T_{min}, °C); average relative humidity (RH, %), wind speed at 10 m above ground (Ws, m s^{-1}); and sunshine hours (SH, h). SH and Ws were converted into global solar radiation (Rs, MJ m^{-2} d^{-1}) and wind speed at 2 m above ground (u_2, m s^{-1}), respectively, adopting the FAO56 methodology (Allen et al., 1998). The AWS records hourly data for the same variables but instead of sunshine hours, records solar radiation directly. Although hourly computations of ETo provide more accurate predictions, hourly records from the AWS dataset were summarised into daily values to render CWS and AWS equivalent, e.g., the minimum and maximum daily values from hourly temperatures were denominated as T_{min} and T_{max}. Temperature, relative humidity, and wind speed were averaged, while solar radiation was summed over daily periods.

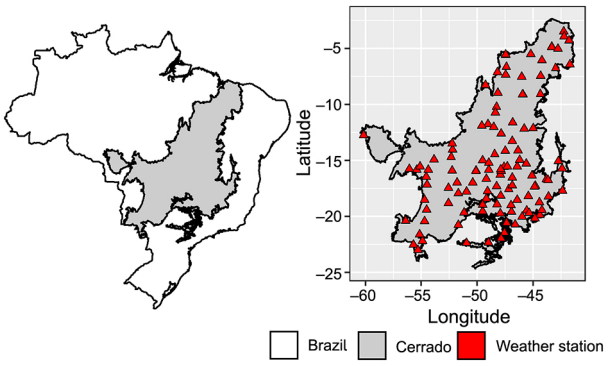

Figure 1 – The Cerrado biome in relation to Brazil and the weather stations used in this study.

The time series for all weather stations were carefully screened to eliminate flawed daily observations, e.g., infinite values, negative wind speed values, T_{min} higher than T_{max}, SH or Rs above the maximum physically possible, and RH outside the range of 5 to 100 %. The Rs time series was compared to the clear-sky solar radiation time series according to Allen (1996), and periods with shifts concerning clear-sky solar radiation were removed. Data were also visually screened for homogeneity (Althoff et al., 2020; Xavier et al., 2016). Weather stations with less than ten years of observations were excluded from the analysis.

Reference crop evapotranspiration

This section describes all ETo equations considered in this study. The Penman-Monteith equation using daily input was selected in this study as the reference ETo equation for the evaluation of ETo methods which can be adopted under conditions of limited data availability. The Penman-Monteith equation has been shown to be the most accurate equation for estimating the ETo measured by lysimeters in regions with the same climate classification as the predominant one in the Cerrado (Barros et al., 2009; Mendonça et al., 2003). Thus, we first describe the Penman-Monteith equation (following section), which requires that all meteorological variables described in the previous section (Data) are available. Next, we describe the equations used to estimate ETo under limited data availability, including temperature-based and temperature-radiation-based methods. These equations are presented in a generic form, i.e., with coefficients instead of the originally proposed values.

FAO Penman-Monteith

From a diverse number of models developed in the literature for estimating ETo, the Penman-Monteith (PM) standardized by FAO 56 (Allen et al., 1998) has been widely used as a reference for calibrating a diverse number of methods (Althoff et al., 2018; Feng et al., 2017; Ferreira et al., 2019; Shiri et al., 2014). Penman-Monteith's ETo is calculated as follows:

$$ETo = \frac{0.408(Rn - G) + \gamma \frac{900}{T_{avg} + 273} u_2 (e_s - e_a)}{\Delta + \gamma(1 + 0.34u_2)} \quad (1)$$

where ETo is the reference crop evapotranspiration (mm d^{-1}), Rn the net radiation (MJ m^{-2} d^{-1}), G the soil heat flux (MJ m^{-2} d^{-1}), T_{avg} the mean air temperature (°C), u_2 the wind speed at 2 m above ground (m s^{-1}), e_s

the mean saturation vapor pressure for a day, kPa; e_a = actual vapor pressure, kPa; Δ = slope of saturation vapor pressure curve at air temperature (kPa °C^{-1}); and, γ = psychrometric constant (kPa °C^{-1}). Rn, e_s, e_a Δ, and γ were calculated according to FAO56 (Allen et al., 1998), while G can be assumed to equal zero as the soil heat flux beneath the reference grass is very limited on a daily time scale.

Penman-Monteith temperature approach

Since the acquisition of all meteorological data needed to calculate PM can be expensive, it is a widespread practice to work with reduced data. To cope with fewer variables, several alternatives were proposed by Allen et al. (1998). In the Penman-Monteith temperature approach (PMT), we assume that only the temperature daily range (T_{max} and T_{min}) is available. Alternative methods were used to estimate missing variables, i.e., relative humidity, solar radiation, and wind speed.

When relative humidity is unavailable, Allen et al. (1998) recommend computing actual vapor pressure, e_a, considering dew point temperature (T_{dew}) to be equal to T_{min} (Eq. 2). Although T_{min} can exceed T_{dew} in semi-arid regions, the opposite can also be true in humid regions (Paredes and Pereira, 2019). The high seasonality in tropical regions may also affect this relationship between wet (Oct-Apr) and dry seasons (May-Sept). It is common to assume $T_{min} = T_{dew}$, however, a temperature correction factor (a_T) can be used to improve the model. Once ea is calculated using T_{max}, T_{min}, and RH (see Allen et al., 1998), the equations below can be used to derive a_T:

$$e_a = e^o(T_{dew}) = 0.6108 \exp\left[\frac{17.27 T_{dew}}{T_{dew} + 237.3}\right] \quad (2)$$

$$T_{dew} = T_{min} - a_T \quad (3)$$

where a_T is the dew temperature correction coefficient (°C).

In the absence of sunshine hours or solar radiation, Allen et al. (1998) suggest that Rs can be estimated using the equation proposed by Hargreaves and Samani (1985), which expresses Rs as a function of temperature range:

$$Rs = k_{Rs} (T_{max} - T_{min}) 0.5\, Ra \quad (4)$$

where k_{Rs} is an empirical radiation adjustment coefficient (°C$^{-0.5}$) and Ra the extraterrestrial solar radiation (MJ m^{-2} d^{-1}). Ra is calculated considering the station latitude and day of the year (Allen et al., 1998). A k_{Rs} equal to 0.16 is suggested for inland sites.

When wind speed data is not available, Allen et al. (1998) recommend using the world average wind speed value of 2.0 m s^{-1}. An average wind speed for the Cerrado and its wet and dry seasons were also evaluated instead of using the world average for model improvement. Finally, the Penman-Monteith temperature approach (PMT) can be calculated using Eq. (1).

Hargreaves-Samani

The Hargreaves-Samani (HS) model (Eq. 5) is a temperature-based method also recommended by Allen et al. (1998) for occasions when solar radiation, relative humidity, and wind speed data are missing. The HS equation has shown reliable results with global validity (Allen et al., 1998) and is calculated as follows:

$$ETo = 0.0135\ k_{Rs}(T_{max} - T_{min})^{0.5})(T_{avg} + offset)\frac{Ra}{\lambda} \quad (5)$$

where 0.0135 is a factor in the conversion of units, offset an empirical coefficient, and λ is the latent heat of vaporization, commonly taken to be 2.45 MJ kg^{-1}. Hargreaves and Samani (1985) suggest the offset = 17.8.

Priestley-Taylor

The Priestley-Taylor (PT) is a temperature-radiation model (Eq. 6) that was originally proposed for estimating evaporation from saturated land and open water sites in advection-free scenarios. As the method does not consider vapor deficit and aerodynamic influence on evapotranspiration, it does not require wind speed and relative humidity data. The PT equation can be described as follows:

$$Eto = \alpha \frac{\Delta}{\Delta + \gamma} \frac{(Rn - G)}{\lambda} \quad (6)$$

where α = empirical coefficient, and α is recommended to be equal to 1.26 by Priestley and Taylor (1972).

Makkink

The Makkink (MK) model is also a temperature-radiation-based method (Eq. 7), whose coefficients have already been parameterized by many authors. However, its original form is written as follows:

$$Eto = c_1 \frac{\Delta}{\Delta + \gamma} \frac{(Rs)}{\lambda} + c_2 \quad (7)$$

where c_1 and c_2 = empirical coefficients. The values suggested for the empirical coefficients c_1 and c_2 are 0.61 and –0.12 (Makkink, 1957), respectively.

Model improvement

The meteorological dataset was divided into two subsets: one set ranging from 2000 to 2014 (approximately 70 %), termed the calibration set, and a second from 2015 to 2020 (approximately 30 %), termed the validation set. The equations were assessed for the validation set, considering both the original empirical coefficients and coefficients calibrated for the calibration set. This study differs from previous studies in that it considers the whole Cerrado biome and uses a single coefficient parametrization for its entirety. The empirical coefficients were calibrated for the Cerrado using the calibration set

considering the (i) whole period (annual) and (ii) wet and dry seasons (seasonal). The empirical coefficients were calibrated by minimizing the sum of squares with the Levenberg-Marquardt algorithm (Moré, 1978).

The performance of the equations using their original and calibrated coefficients (annual and seasonal) was evaluated for the validation set considering the (i) entire validation set, (ii) monthly periods, and (iii) spatially (station-wise). Figure 2 presents a flowchart of the study structure.

Evaluation criteria

The statistical indices adopted as performance criteria for the ETo equations were the percentage bias (PBIAS), the absolute mean error (MAE), the root mean square error (RMSE), and the Nash and Sutcliffe (1970) efficiency (NSE):

$$PBIAS = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{X_i - Y_i}{\bar{Y}}\right) \quad (8)$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|X_i - Y_i| \quad (9)$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(X_i - Y_i)^2} \quad (10)$$

$$NSE = 1 - \frac{\sum_{i=1}^{n}(X_i - Y_i)^2}{\sum_{i=1}^{n}(X_i - \bar{X})^2} \quad (11)$$

where X_i and Y_i denote ETo values estimated by Penman-Monteith and remaining methods, respectively. $\bar{X}$ and $\bar{Y}$ represent the corresponding mean *ETo* values; i refers to the i^{th} value of *ETo* and n is the number of data observations.

Percentage bias indicates if a model generally under or overestimates PM-ETo. Mean absolute error indicates the average error magnitude, while RMSE assigns more weight to larger errors. Nash-Sutcliffe efficiency ranges from $-\infty$ to 1, where 1 indicates a perfect model and negative values are negligible and indicate that the model is as good as the average of the observations (Schaefli and Gupta, 2007). Uncalibrated models are used here as benchmarks to evaluate if there is any improvement obtained by calibrating the models' coefficients or if the model must, nevertheless, be considered unacceptable (NSE < 0).

Results and Discussion

Data overview

The daily observations from weather stations gathered for the Cerrado from 2000 to 2020 are displayed in Figure 3. The number of observations for each station ranged from 10.1 to 20.5 years. AWSs yield more data from recent periods, while CWSs yield more data from the early 2000s, although some have been decommissioned more recently. To overcome the discontinuity of observation data sets in

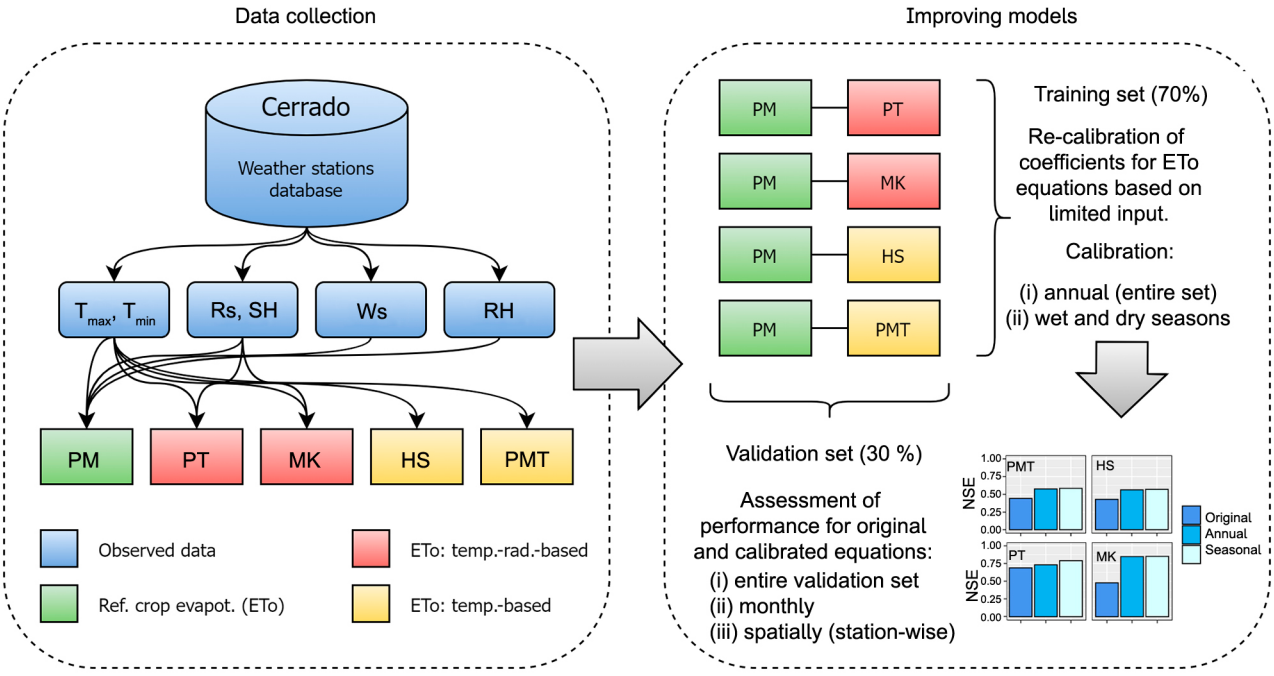

Figure 2 – Flowchart describing data collection, model improvement strategies, and performance evaluation. ETo = reference crop evapotranspiration; T_{max} = daily maximum temperature; T_{min} = daily minimum temperature; Rs = solar radiation; SH = sunshine hours; Ws = wind speed; RH = relative humidity; PM = Penman-Monteith; PT = Priestley-Taylor; MK = Makkink; PMT = Penman-Monteith temperature approach; NSE = Nash-Sutcliffe efficiency.

space and time, satellite data and reanalysis products may be considered a possible alternative, though they might require bias correction or adjustment to the well-watered conditions of reference sites.

The meteorological variables show a clear distinction between wet and dry seasons. For example, relative humidity decreases throughout the dry season until Oct, when the rainy season begins. The dry season is also associated with less cloud cover due to lower relative humidity. This results in a lower spread of solar radiation distribution during these months. In the dry season (winter), minimum daily temperatures are lower, while wind speed is generally higher.

Daily mean air temperature values varied from 4.9 to 34.6 °C (avg = 23.3 °C) in the dry season and from 9.2 to 35.6 °C (avg = 24.9 °C) in the wet season. Similarly, wind speed ranged from 0 to approximately 8.0 m s^{-1} in both seasons, with averages equal to 1.15 and 1.29 for the wet and dry seasons, respectively. Average relative humidity for the dry and wet seasons were 59.9 % and 73.6 %, respectively. The average solar radiation was similar for both wet (19.3 MJ m^{-2} d^{-1}) and dry (18.5 MJ m^{-2} d^{-1}) seasons. However, it registered a higher variance in the wet season due to rainfall and the increased cloudiness.

Reference evapotranspiration generally follows the same behavior of maximum daily air temperature and solar radiation. Solar radiation and wind speed generally increase at the end of the dry season. or the mainpart the high solar radiation results fom low relative humidity and cloudiness. Thus, low relative humidity associated with high solar radiation and wind speed produced the highest average ETo in Sept (4.97 mm d^{-1}). In contrast, lower temperatures and solar radiation resulted in the lowest average ETo in June (3.30 mm d^{-1}).

In the wet season, Rs showed the highest linear correlation to ETo (cor = 0.94), followed by RH (–0.75) and T_{max} (0.68). The variables least correlated to ETo in the wet season were u_2 (0.25) and T_{min} (0.09). In contrast, Rs (0.81), RH (–0.67), and T_{max} (0.66) correlation to ETo declined in the dry season, while for u_2 (0.43) and T_{min} (0.46) it increased.

Improved parameters for the Cerrado

The alternative methods for computing ETo were calibrated using the calibration set considering the annual and season periods. The parameters, i.e., empirical coefficients and correction factors, are displayed in Table 1. The dew point temperature correction factor, a_T,

Table 1 – Summary of calibrated parameters.

Parameters	Original	Annual	Wet	Dry
a_T	0.00	–0.42	–0.93	0.57
u_2	2.00	1.23	1.17	1.31
k_{Rs}	0.160	0.159	0.155	0.165
offset	17.8	14.9	14.8	15.1
α	1.26	1.18	1.13	1.30
c_1	0.61	0.70	0.67	0.76
c_2	–0.12	0.10	0.28	–0.26

a_T = temperature correction factor (°C); u_2 = mean wind speed 2 meters above ground (m s^{-1}); k_{Rs} = empirical radiation adjustment coefficient (°$C^{-0.5}$); offset = empirical coefficient of the Hargreaves-Samani model; α = empirical coefficient of the Priestley-Taylor model; c_1 and c_2 = empirical coefficients of the Makkink model.

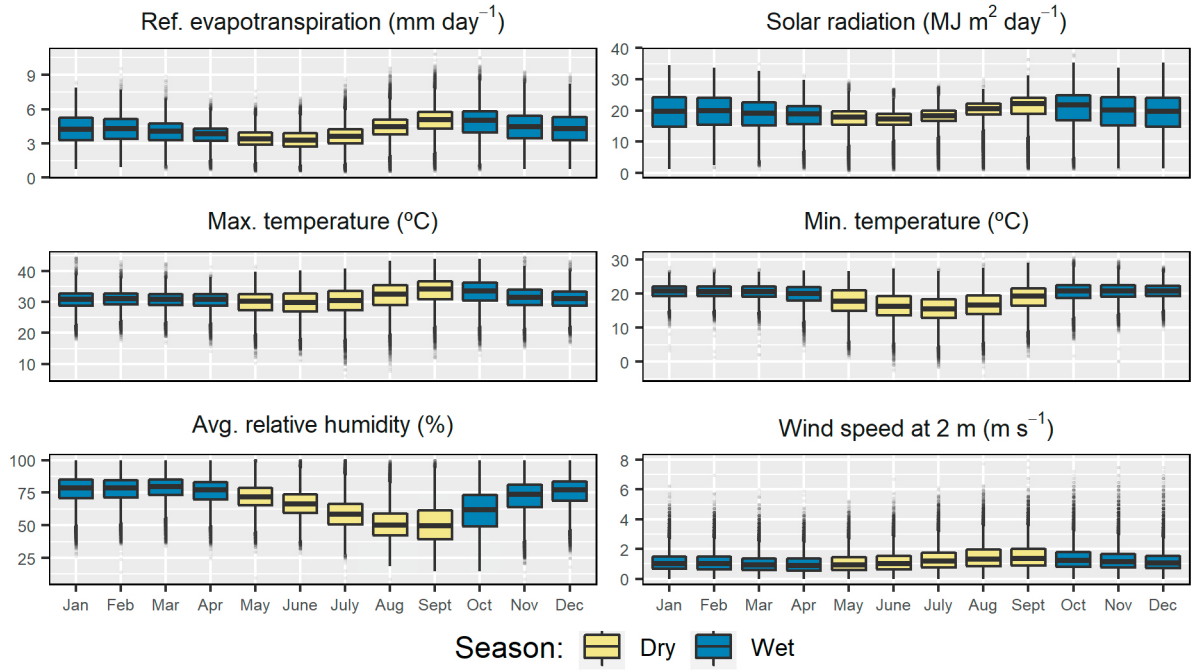

Figure 3 – Daily observations of climatic data in the Cerrado.

was negative (–0.42) for the annual period, implying that T_{dew} is generally higher than T_{min}. Negative a_T values are generally seen in humid regions (Paredes and Pereira, 2019), similar to here, i.e., negative [positive] value calibrated for the wet [dry] season. T_{dew} is commonly higher than T_{min} when daytime absolute humidity measurements exceed nighttime measurements due to evaporative conditions and humidification from daytime evapotranspiration (Allen, 1996). In contrast, despite the dry season's lower relative humidity and higher wind speeds, a_T was lower than 1 °C. In both cases, a_T was lower than 1 °C because the temperature (T_{min}) will approach T_{dew} at reference sites provided the wind speed is relatively calm (Allen, 1996). The average wind speed (1.23 m s^{-1}) in the Cerrado is lower than the global average, with lower averages seen in the wet season, as discussed in the previous section. The k_{Rs} factor resulted in values very close to that suggested for the inland region. Although higher values were generally observed near the ocean, they were recorded in the dry season (0.165) instead of the wet season (0.155).

The offset used in the Hargreaves-Samani equation was lower for the Cerrado (14.9), with a slight variance between wet and dry seasons. The empirical coefficient of the PT equation was lower in the wet season (1.13) and higher in the dry season (1.30), when advection is likely higher due to increased wind speed and lower relative humidity. The MK equation calibrated a more significant slope (c_1) for the dry season (0.76) as opposed to the wet season (0.67), while the intercept (c_2) was higher during the wet season (0.28) and lower in the dry season (–0.26).

Performances of alternative reference evapotranspiration methods

The performance achieved by the alternative ETo models considering the original parametrization and the improved model for the Cerrado is displayed in Figure 4. The original parametrization for HS, PMT, and PT revealed a an overestimated ETo, while MK generally underestimated it. After improving the parameters, all the methods resulted in negligible PBIAS (< 3 %). Before calibration, the MAE was equal to 0.74, 0.75, 0.56, and 0.74 mm d^{-1} for PMT, HS, PT, and MK, respectively. After calibrating the parameters for the annual period, MAE declined by 17 %, 16 %, 18 %, and 55 % for PMT, HS, PT, and MK, respectively. The MK had the most significant improvement since it showed the greatest bias before calibration. PMT and HS had minor improvements, likely due to limitation of using only temperature data.

The methods showed slight improvement by considering a calibration on a seasonal versus an annual basis. For example, the parameters calibrated on an annual or seasonal basis were very similar for PMT and HS. The temperature correction factor, average wind speed, and k_{Rs} presented minor differences between

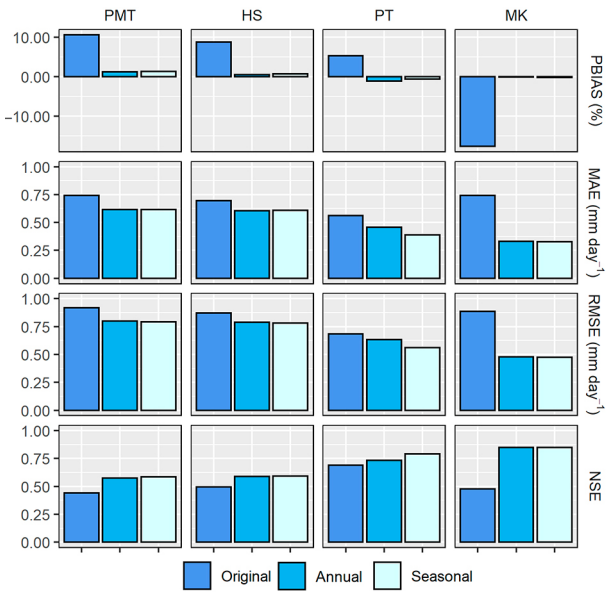

Figure 4 – Performance achieved by the alternative reference evapotranspiration methods computed in their original form, and with parameters improved for the Cerrado on an annual and seasonal basis. PBIAS = percentual bias; MAE = mean absolute error; RMSE = root mean square error; NSE = Nash-Sutcliffe Efficiency; PMT = Penman-Monteith temperature approach; HS = Hargreaves-Samani; PT = Priestley-Taylor; MK = Makkink.

the seasons (Table 1). Only the PT equation showed significant performance improvement by considering the seasons, with MAE declining up to 31 % in this case. The MK method seemed to have already improved sufficiently for the annual period by considering the slope (c_1) and intercept in calibration (c_2), while PT relied entirely on the empirical coefficient to scale its estimates in different seasons.

The small or insignificant improvements for the methods on a seasonal basis are, despite the better-adjusted parameters, the result of their lack of information on variables such as wind speed and relative humidity. For instance, other studies also concluded that only minor improvements could be achieved for the HS equation in the dry season for a border region between the Cerrado, Caatinga, and Atlantic rainforest with climate classified as tropical, semi-arid, and humid subtropical for the weather stations studied (Althoff et al., 2019). As wind speed correlation to ETo increases in the dry season, and relative humidity also has a relatively high correlation to ETo, it is difficult for the models to capture the ETo dynamics in this period without them.

Overall, the best performing method after calibration was MK, with NSE = 0.85 for both seasonal and annual parametrizations, followed by PT (seasonal: NSE = 0.79; annual NSE = 0.73), PMT (seasonal: NSE = 0.59; annual: NSE = 0.58), and HS (seasonal: NSE = 0.57; annual: NSE = 0.56). For instance, annual parametrization resulted in an RMSE for MK equal to

0.48 mm d^{-1}, which is lower by 40 %, 41 %, and 25 % when compared to PMT (0.80 mm d^{-1}), HS (0.79 mm d^{-1}), and PT (0.63 mm d^{-1}), respectively.

Other alternative methods based on machine learning have been reported with good performance in different studies. For instance, Ferreira et al. (2019) used artificial neural networks and a single temperature-based model for Brazil and achieved performance similar to our results (RMSE = 0.81 mm d^{-1}). The authors demonstrated that this was an improvement compared to the HS and PMT equations calibrated for the entire Brazilian territory. However, distributing an equation is still more practical than sharing a machine learning model with stakeholders. Following this idea, using a genetic algorithm, Valle Júnior et al. (2020) generated new equations for a Cerrado region under a tropical savanna climate with dry winter similar to ours. However, these equations' performance was still short when compared to the calibrated version of the equations used here with similar input.

Clustering weather stations under similar climates has improved the final performance when modeling ETo for Brazil (Ferreira et al., 2019). However, the clustering algorithm used on a national scale for Brazil grouped stations in regions similar to that delimited by the Brazilian biomes Amazon rainforest, Cerrado, Caatinga, and Atlantic rainforest (Ferreira et al., 2019). This is because, under different climate types, the chances are that a single parametrization would be biased towards the climate type more often seen by weather stations. Additionally, it has also been shown that clustering stations for a region with a similar climate by the Cerrado border did not improve the calibration performance (Althoff et al., 2019). Since the Cerrado climate is predominantly characterized by a tropical savanna climate with strong seasonality, parametrization of coefficients for the entire biome is more practical for stakeholders than station-wise or sub-regional parametrizations. This further reinforces that choosing a single set of coefficients calibrated for the Cerrado is practical and the recommended approach.

The scatterplots in Figure 5 reinforce the lower level of sensitivity shown by methods based solely on temperature, where the regression slopes did not show any significant improvement with calibration. This shows that most of the improvements in errors came from reducing bias, but the methods still result in relatively large errors on a daily time scale. Despite both MK and PT being temperature-radiation based, MK was the method that returned the best overall performance, and slope, intercept, and coefficient of determination (R^2) closer to the desired values.

However, even when these methods are adopted for water management purposes, e.g., irrigation, the under and/or overestimation in consecutive periods can balance each other out. Thus, to conclude that the methods systematically under or overestimated ETo in specific months, the performance was assessed for each month separately. Figure 6 presents the performance criteria for the annual parametrization, while Figure 7 presents the performance criteria for the seasonal parametrization. Annual parametrization (Figure

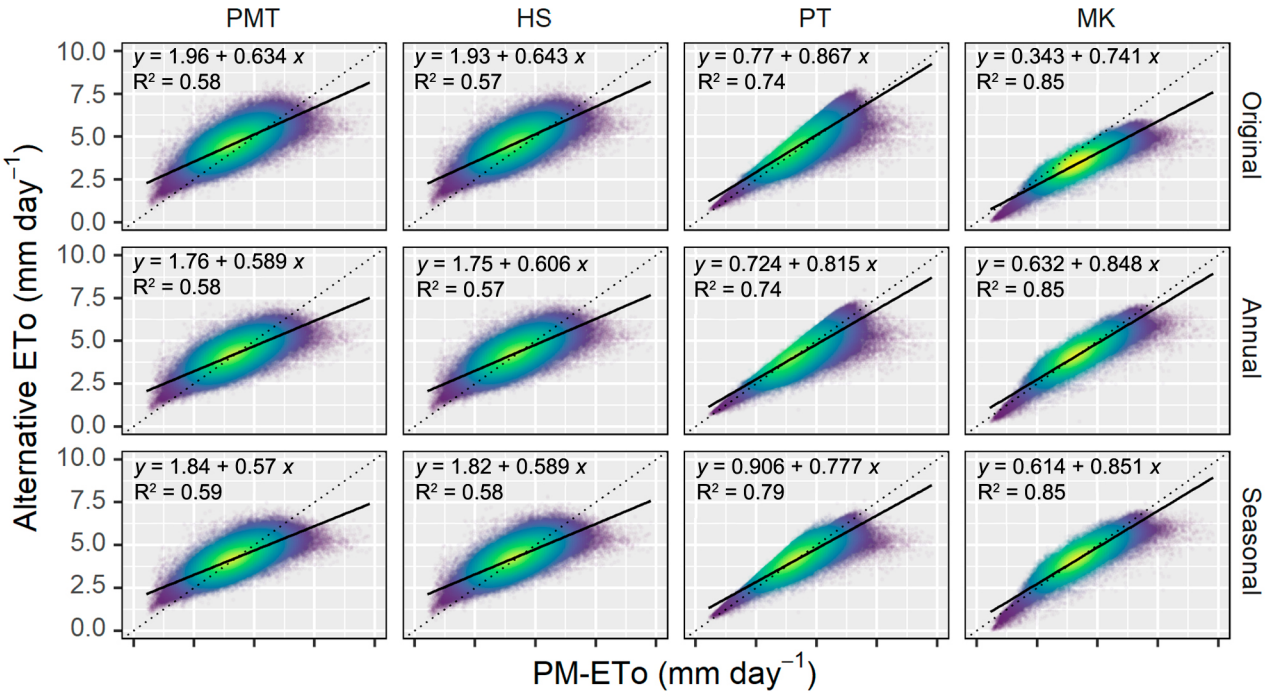

Figure 5 – Reference evapotranspiration values observed (Penman-Monteith) and estimated by alternative models. PMT = Penman-Monteith temperature approach; HS = Hargreaves-Samani; PT = Priestley-Taylor; MK = Makkink.

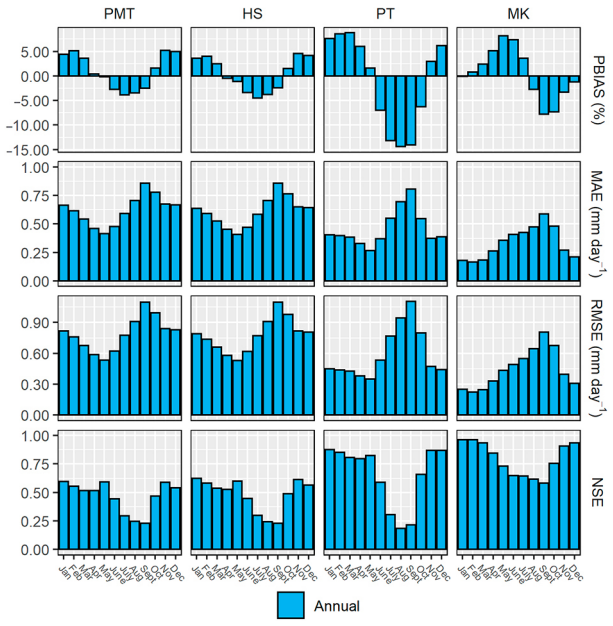

Figure 6 – Performance achieved by the alternative reference evapotranspiration methods computed with parameters improved for the Cerrado on an annual basis. PBIAS = percentual bias; MAE = mean absolute error; RMSE = root mean square error; NSE = Nash-Sutcliffe Efficiency; PMT = Penman-Monteith temperature approach; HS = Hargreaves-Samani; PT = Priestley-Taylor; MK = Makkink.

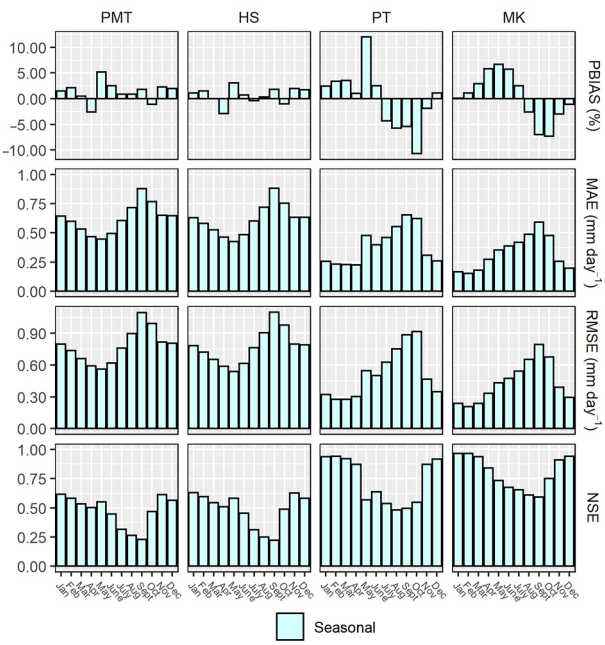

Figure 7 – Performance achieved by the alternative reference evapotranspiration methods computed with parameters improved for the Cerrado on a seasonal basis. PBIAS = percentual bias; MAE = mean absolute error; RMSE = root mean square error; NSE = Nash-Sutcliffe Efficiency; PMT = Penman-Monteith temperature approach; HS = Hargreaves-Samani; PT = Priestley-Taylor; MK = Makkink.

6) presents a strong seasonal presence of bias in predictions. For example, temperature-based methods tend to overestimate ETo during the wet season and underestimate it during the dry season. PT shows similar biased behavior. In contrast, MK tends to overestimate ETo from Feb to July, then underestimates it from Aug to Dec. This shows that MK generally overestimates ETo in the period where it is supposedly lower and underestimates it in the period when it is higher.

For parametrization on a seasonal basis, the overall bias across the months was reduced. Interestingly, despite showing sparse performance improvement, the biases for temperature-based methods were better balanced across all months after the seasonal parametrization. PMT and HS presented PBIAS below 5 % for all months except May. This signifies that over more extended periods, e.g., weeks or month-long periods, the estimates made by more straightforward methods can result in better estimates than expected. In contrast, the biases for MK and PT across the months maintained a certain seasonality, resulting in PBIAS values exceeding 5 % in several months.

The alternative ETo methods are also assessed for each station to provide a regional assessment of their performance (Figure 8). Central Cerrado showed a somewhat lower performance. For the temperature-based methods PMT and HS, a few stations resulted in negative NSE, which highlights their inadequacy for predicting ETo. The MK method outperformed the temperature-based methods for all stations and PT in 79 % of the stations. Even when PT outperformed MK, the difference was negligible. Thus, the MK method is recommended regardless of the site within the Cerrado. As for the temperature-based equations, the HS method outperformed the PMT in 72 % of the stations.

Our main recommendation for future research is that a more concentrated effort should be channeled into estimating solar radiation using limited data. Solar radiation is the variable with the highest correlation to ETo, and it is fundamental to any endeavor to improve the estimation process when data is missing. Using the suggested equation (Eq. 4), a closer look at solar radiation estimates reveals poor overall performance. For example, even when seasonal calibration was included, Eq. (4) resulted in an MAE = 2.88 MJ m^{-2} d^{-1}, RMSE = 3.63 MJ m^{-2} d^{-1}, and NSE = 0.46. Several alternative methods (Fan et al., 2018) have been proposed and assessed with better outcomes than the traditional formulation proposed by Hargreaves and Samani (1985). Thus, there seems to be room for significant improvement for PMT and HS, especially when their balanced biases across all months are taken into account.

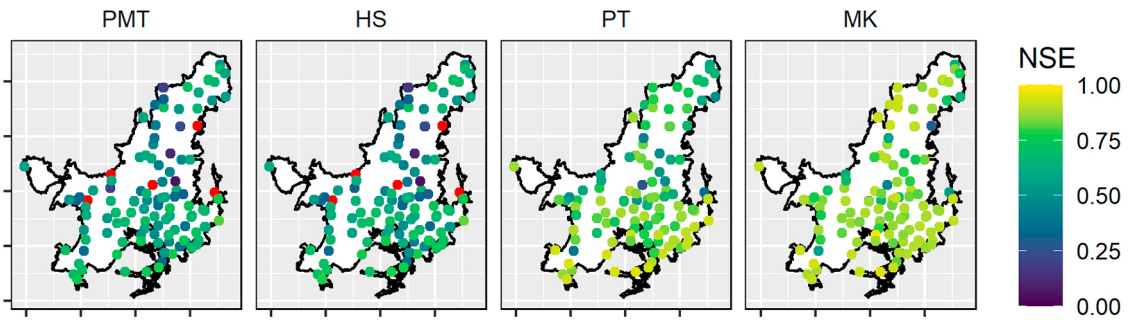

Figure 8 – Performance achieved by the alternative reference evapotranspiration methods for each station in the Cerrado with parameters improved for the Cerrado on a seasonal basis. Red points are stations where NSE scored below 0. NSE = Nash-Sutcliffe Efficiency; PMT = Penman-Monteith temperature approach; HS = Hargreaves-Samani; PT = Priestley-Taylor; MK = Makkink.

Conclusions

In this study, the most popular alternatives to compute reference evapotranspiration were assessed concerning the Cerrado biome. These methods are the Penman-Monteith Temperature approach (PMT), Hargreaves-Samani (HS), Priestley-Taylor (PT), and Makkink (MK), which are particularly useful when only limited data is available. The parametrization of these methods was also improved by considering an annual and seasonal basis for the region. Improved parametrization for the Cerrado reduced the mean absolute error between 16 % to 55 % for these methods for the annual parametrization, except for the PT method. Overall, the best performing temperature-radiation-based method was the MK, with an NSE of 0.85 for both annual and seasonal parametrization. The best temperature-based method was the PMT, with an NSE of 0.58 and 0.59 for seasonal and annual parametrization, respectively.

Despite the seasonal parametrization failing to improve performance criteria, it was seen to reduce systematic biases on a monthly scale, especially in the case of the temperature-based methods. Future research should focus efforts in the main part on improving solar radiation estimates from limited data. Solar radiation showed the highest correlation with reference evapotranspiration, and the traditional method proposed for its estimation performed poorly in this study. This likely constrained the opportunity for temperature-based methods to estimate reference evapotranspiration adequately.

Acknowledgments

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) – Finance code 001, and by the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) – Grant number 142273/2019-8.

Authors' Contributions

Conceptualization: Althoff, D.; Rodrigues, L.N. **Data acquisition**: Althoff, D. **Design of methodology:** Althoff, D.; Rodrigues, L.N. **Software development:** Althoff, D. **Writing and editing**: Althoff, D.; Rodrigues, L.N.

References

Allen, R.G. 1996. Assessing integrity of weather data for reference evapotranspiration estimation. Journal of Irrigation and Drainage Engineering 122: 97-106. https://doi.org/10.1061/(ASCE)0733-9437(1996)122:2(97)

Allen, R.G.; Pereira, L.S.; Raes, D.; Smith, M. 1998. Crop Evapotranspiration: Guidelines for Computing Crop Water Requirements. 9ed. FAO, Rome, Italy (Irrigation and Drainage Paper, 56).

Althoff, D.; Bazame, H.C.; Filgueiras, R.; Dias, S.H.B. 2018. Heuristic methods applied in reference evapotranspiration modeling. Ciência e Agrotecnologia 42: 314-324. https://doi.org/10.1590/1413-70542018423006818

Althoff, D.; Santos, R.A.; Bazame, H.C.; Cunha, F.F.; Filgueiras, R. 2019. Improvement of Hargreaves-Samani reference evapotranspiration estimates with local calibration. Water 11: 2272. https://doi.org/10.3390/w11112272

Althoff, D.; Dias, S.H.B.; Filgueiras, R.; Rodrigues, L.N. 2020. ETo-Brazil: a daily gridded reference evapotranspiration data set for Brazil (2000-2018). Water Resource Research 56: e2020WR027562. https://doi.org/10.1029/2020WR027562

Alvares, C.A.; Stape, J.L.; Sentelhas, P.C.; Gonçalves, J.L.M.; Sparovek, G. 2013. Köppen's climate classification map for Brazil. Meteorologische Zeitschrift 22: 711-728. https://doi.org/10.1127/0941-2948/2013/0507

Barros, V.R.; Souza, A.P.; Fonseca, D.C.; Silva, L.B. 2009. Reference evapotranspiration in the region of Seropédica-RJ using weighting lysimeter and mathematical models. Revista Brasileira de Ciências Agrárias 4: 198-203 (in Portuguese, with abstract in English). https://doi.org/10.5039/agraria.v4i2a13

Fan, J.; Chen, B.; Wu, L.; Zhang, F.; Lu, X.; Xiang, Y. 2018. Evaluation and development of temperature-based empirical models for estimating daily global solar radiation in humid regions. Energy 144: 903-914. https://doi.org/10.1016/j.energy.2017.12.091

Feng, Y.; Cui, N.; Gong, D.; Zhang, Q.; Zhao, L. 2017. Evaluation of random forests and generalized regression neural networks for daily reference evapotranspiration modelling. Agricultural Water Management 193: 163-173. https://doi.org/10.1016/j.agwat.2017.08.003

Ferreira, L.B.; Cunha, F.F.; Oliveira, R.A.; Fernandes Filho, E.I. 2019. Estimation of reference evapotranspiration in Brazil with limited meteorological data using ANN and SVM: a new approach. Journal of Hydrology 572: 556-570. https://doi.org/10.1016/j.jhydrol.2019.03.028

Gavilán, P.; Lorite, I.J.; Tornero, S.; Berengena, J. 2006. Regional calibration of Hargreaves equation for estimating reference ET in a semiarid environment. Agricultural Water Management 81: 257-281. https://doi.org/10.1016/j.agwat.2005.05.001

Hargreaves, G.H.; Samani, Z.A. 1985. Reference crop evapotranspiration from temperature. Applied Engineering in Agriculture 1: 96-99. https://doi.org/10.13031/2013.26773

Klink, C.A. 2014. Policy Intervention in the Cerrado Savannas of Brazil: Changes in the Land Use and Effects on Conservation. p. 1-21. In: Consorte-McCrea, A.G.; Santos, E.F., eds. Ecology and conservation of the Maned Wolf: multidisciplinary perspectives. CRC Press, Boca Raton, FL, USA.

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

Mendonça, J.C.; Sousa, E.F.; Bernardo, S.; Dias, G.P.; Grippa, S. 2003. Comparison of estimation methods of reference crop evapotranspiration (ETo) for northern Region of Rio de Janeiro State, Brazil. Revista Brasileira de Engenharia Agrícola e Ambiental 7: 275-279 (in Portuguese, with abstract in English). https://doi.org/10.1590/S1415-43662003000200015

Moré, J.J. 1978. The Levenberg-Marquardt algorithm: implementation and theory. p. 105-116. In: Numerical analysis: lecture notes in mathematics. Springer, Berlin, Germany. https://doi.org/10.1007/BFb0067700

Nash, J.E.; Sutcliffe, J.V. 1970. River flow forecasting through conceptual models. Part I. A discussion of principles. Journal of Hydrology 10: 282-290. https://doi.org/10.1016/0022-1694(70)90255-6

Paredes, P.; Pereira, L.S. 2019. Computing FAO56 reference grass evapotranspiration PM-ET_o from temperature with focus on solar radiation. Agricultural Water Management 215: 86-102. https://doi.org/10.1016/j.agwat.2018.12.014

Priestley, C.H.B.; Taylor, R.J. 1972. On the assessment of surface heat flux and evaporation using large-scale parameters. Monthly Weather Review 100: 81-92. https://doi.org/10.1175/1520-0493(1972)100<0081:OTAOSH>2.3.CO;2

Schaefli, B.; Gupta, H.V. 2007. Do Nash values have value? Hydrology Process 21: 2075-2080. https://doi.org/10.1002/hyp.6825

Shiri, J.; Nazemi, A.H.; Sadraddini, A.A.; Landeras, G.; Kisi, O.; Fard, A.F.; Marti, P. 2014. Comparison of heuristic and empirical approaches for estimating reference evapotranspiration from limited inputs in Iran. Computers and Electronics in Agriculture 108: 230-241. https://doi.org/10.1016/j.compag.2014.08.007

Valle Júnior, L.C.G.; Ventura, T.M.; Gomes, R.S.R.; Nogueira, J.S.; Lobo, F.A.; Vourlitis, G.L.; Rodrigues, T.R. 2020. Comparative assessment of modelled and empirical reference evapotranspiration methods for a brazilian savanna. Agricultual Water Management 232: e106040. https://doi.org/10.1016/j.agwat.2020.106040

Wendt, D.E.; Rodrigues, L.N.; Dijksma, R.; Dam, J.C.V. 2015. Assessing groundwater potential use for expanding irrigation in the Buriti Vermelho watershed. Irriga & Inovagri 1: 81-94. https://doi.org/10.15809/irriga.2015v1n2p81

World Meteoerolocal Organization [WMO]. 2008. Guide to Meteorological Instruments and Methods of Observation. 7ed. World Meteorological Organization, Geneva, Switzerland. (WMO, 8)

Xavier, A.C.; King, C.W.; Scanlon, B.R. 2016. Daily gridded meteorological variables in Brazil (1980-2013). International Journal of Climatology 36: 2644-2659. https://doi.org/10.1002/joc.4518

Zanetti, S.S.; Dohler, R.E.; Cecílio, R.A.; Pezzopane, J.E.M.; Xavier, A.C. 2019. Proposal for the use of daily thermal amplitude for the calibration of the Hargreaves-Samani equation. Journal of Hydrology 571: 193-201. https://doi.org/10.1016/j.jhydrol.2019.01.049

Appendix I – Weather stations' geographic coordinates, elevation, and the number of observations in the study period. Stations beginning with "A" are automatic weather stations.

INMET ID	Observations	Latitude	Longitude	Elevation
82296	5408	–3.47	–42.27	44.1
82298	4692	–3.90	–42.25	87.1
82474	3915	–4.28	–41.80	160.0
82476	4861	–4.87	–43.36	96.7
82480	7010	–4.28	–41.79	157.9
82564	5156	–5.54	–47.48	126.3
82571	6238	–5.51	–45.24	154.2
82578	4394	–5.03	–42.80	75.7
82659	7462	–7.10	–48.20	231.9
82676	4507	–6.03	–44.23	175.6
82678	6707	–6.76	–43.00	126.6

Continue...

Appendix I - Continuation.

82768	6233	–7.53	–46.05	263.5
82861	6945	–8.26	–49.26	179.0
82863	6715	–8.97	–48.18	189.5
82970	6157	–9.11	–45.95	285.1
82975	5934	–9.07	–44.37	330.6
83033	7313	–10.19	–48.30	291.7
83064	5278	–10.71	–48.41	243.3
83228	7368	–12.02	–48.54	252.2
83270	6426	–13.47	–52.27	430.0
83319	6807	–14.70	–52.35	305.3
83332	3866	–14.09	–46.37	830.4
83334	4900	–14.95	–46.24	854.6
83358	6068	–15.83	–54.40	374.4
83363	4582	–15.82	–55.42	787.0

Continue...

Appendix I - Continuation.

83364	7092	–15.78	–56.07	140.0
83368	7065	–15.90	–52.25	327.0
83373	5620	–15.93	–47.88	1100.6
83374	4469	–15.91	–50.13	512.2
83376	7473	–15.85	–48.97	766.9
83377	7416	–15.79	–47.93	1161.4
83379	7339	–15.55	–47.34	938.7
83388	4487	–15.08	–42.75	625.0
83437	5407	–16.69	–43.84	645.9
83452	5737	–16.78	–43.67	655.6
83464	6463	–17.92	–51.72	669.8
83470	6077	–17.79	–50.96	780.0
83479	7219	–17.24	–46.88	711.4
83483	5443	–17.35	–44.91	505.2
83531	6052	–18.51	–46.43	940.3
83533	5396	–19.71	–45.36	695.0
83538	5165	–18.23	–43.64	1296.1
83565	4474	–19.66	–51.19	429.6
83577	5954	–19.73	–47.95	737.0
83582	6042	–20.03	–46.00	661.3
83586	5279	–19.46	–44.25	732.0
83630	7365	–20.58	–47.38	1003.6
83635	7128	–20.17	–44.87	787.4
83669	6167	–21.46	–47.58	620.0
83726	5292	–21.98	–47.88	859.8
A001	6135	–15.79	–47.93	1161.0
A002	6107	–16.64	–49.22	727.3
A003	4412	–17.75	–49.10	751.1
A004	4452	–14.45	–48.45	583.0
A009	5173	–10.19	–48.30	291.9
A010	3751	–12.62	–47.87	285.0
A012	4908	–16.26	–47.97	1000.8
A013	4176	–15.90	–52.25	327.4
A014	4293	–15.94	–50.14	512.9
A015	4178	–14.98	–49.54	551.2
A016	4325	–17.92	–51.72	670.1
A017	4436	–14.09	–46.37	830.0
A018	4085	–12.02	–48.54	250.9
A019	4720	–11.75	–49.05	279.1
A020	4670	–8.97	–48.18	189.7
A021	3915	–7.10	–48.20	230.8
A022	3863	–15.22	–48.99	667.0
A023	4649	–16.97	–51.82	740.1
A024	3884	–14.13	–47.52	1264.7
A025	3914	–17.79	–50.96	780.1
A026	3784	–17.45	–52.60	862.0
A027	4257	–16.96	–50.43	678.7
A032	3874	–13.25	–46.89	551.3
A033	4268	–17.30	–48.28	757.3
A034	4265	–18.15	–47.93	900.7
A035	3934	–18.41	–49.19	491.2
A038	4197	–11.59	–46.85	727.9
A039	3726	–11.89	–49.61	215.2
A045	4104	–15.60	–47.63	1030.4
A205	3836	–7.34	–47.46	182.9

Continue...

Appendix I - Continuation.

A221	3696	–5.51	–45.24	154.2
A223	4276	–9.11	–45.93	283.7
A224	3825	–6.65	–47.42	183.0
A225	4374	–5.56	–47.46	118.0
A237	3746	–4.82	–43.34	84.9
A241	3885	–8.30	–49.28	175.7
A311	4800	–6.76	–43.00	126.4
A312	3871	–5.03	–42.80	75.2
A346	4238	–7.44	–44.35	398.8
A363	3928	–6.40	–41.74	313.4
A402	5003	–12.12	–45.03	474.2
A404	5379	–12.15	–45.83	760.7
A505	5614	–19.61	–46.95	1018.3
A506	5903	–16.69	–43.84	645.9
A507	5552	–18.92	–48.26	874.8
A512	4633	–18.95	–49.53	540.1
A516	4164	–20.75	–46.63	781.7
A519	4526	–19.54	–49.52	559.1
A523	4867	–19.00	–46.99	978.1
A528	4532	–18.20	–45.46	931.0
A535	4512	–19.89	–44.42	753.5
A536	4354	–19.48	–45.59	721.1
A537	4791	–18.23	–43.65	1359.3
A538	4950	–18.75	–44.45	669.5
A541	4508	–17.71	–42.39	932.1
A542	4643	–16.55	–46.88	640.9
A544	4176	–15.52	–46.44	894.0
A545	4283	–17.26	–44.84	505.3
A546	4601	–17.56	–47.20	997.0
A547	4550	–16.36	–45.12	490.3
A548	4222	–15.30	–45.62	873.2
A551	4547	–15.72	–42.44	850.1
A702	5310	–20.45	–54.72	528.5
A703	4909	–22.55	–55.72	668.0
A704	4532	–20.78	–51.71	328.9
A705	4558	–22.36	–49.03	636.2
A708	5597	–20.58	–47.38	1002.7
A711	4166	–21.98	–47.88	859.3
A718	4135	–22.37	–50.97	398.8
A720	3943	–18.51	–54.74	251.5
A721	4036	–22.19	–54.91	463.3
A722	3727	–20.40	–56.43	132.5
A730	4016	–18.80	–52.60	820.8
A731	3750	–21.61	–55.18	388.9
A732	3706	–19.42	–54.55	646.0
A750	4036	–23.00	–55.33	433.6
A907	4721	–16.46	–54.58	289.9
A908	3796	–14.02	–52.21	440.0
A912	4169	–15.53	–55.14	748.3
A931	4066	–14.93	–53.88	664.2
A933	3684	–17.18	–54.50	593.1
A938	3805	–12.73	–60.16	583.3