CROP SCIENCE

Generalizability of machine learning models and empirical equations for the estimation of reference evapotranspiration from temperature in a semiarid region

LUCAS B. FERREIRA, FERNANDO F. DA CUNHA, GUSTAVO H. DA SILVA,
FLAVIO B. CAMPOS, SANTOS H.B. DIAS & JANNAYTON E.O. SANTOS

Abstract: The Penman-Monteith equation is recommended for the estimation of reference evapotranspiration (ET_o). However, it requires meteorological data that are commonly unavailable. Thus, this study evaluates artificial neural network (ANN), multivariate adaptive regression splines (MARS), and the original and calibrated Hargreaves-Samani (HS) and Penman-Monteith temperature (PMT) equations for the estimation of daily ET_o using temperature. Two scenarios were considered: (i) local, models were calibrated/developed and evaluated using data from individual weather stations; (ii) regional, models were calibrated/developed using pooled data from several stations and evaluated independently in each one. Local models were also evaluated outside the calibration/training station. Data from 9 stations were used. The original PMT outperformed the original HS, but after local or regional calibrations, they performed similarly. The locally calibrated equations and the local machine learning models exhibited higher performances than their regional versions. However, the regional models had higher generalization capacity, with a more stable performance between stations. The machine learning models performed better than the equations evaluated. When comparing the ANN models with the HS equation, mean RMSE reduced from 0.96 to 0.87 and from 0.84 to 0.73, in regional and local scenarios, respectively. ANN and MARS performed similarly, with a slight advantage for ANN.

Key words: ANN, cross-station, external validation, MARS, regional models.

INTRODUCTION

Quantification of evapotranspiration is of vital importance for irrigation scheduling. The FAO-56 Penman-Monteith (FAO-PM) equation is widely recommended for the estimation of reference evapotranspiration (ET_o) (Allen et al. 1998). However, it requires meteorological variables that are commonly unavailable or unreliable (Almorox et al. 2018, Pinheiro et al. 2019). Thus, equations that require only air temperature can be used as an alternative way since temperature is commonly measured.

The Hargreaves-Samani (HS) equation can be used when only air temperature data are available (Allen et al. 1998, Zanetti et al. 2019). In addition, several studies have shown that the FAO-56 Penman-Monteith equation using only measured data on temperature, commonly named Penman-Monteith temperature (PMT), can also be used (Raziei & Pereira 2013, Alencar et al. 2015, Almorox et al. 2018). However, both the HS equation and the PMT equation have their performance varying according to the climatic conditions of the place where they are

used. Thus, the calibration of these equations is extremely important (Zanetti et al. 2019).

In recent years, machine learning methods, such as artificial neural network (ANN), support vector machine (SVM) and gene expression programming (GEP), have been used to estimate environmental, hydrologic and climatological parameters (Ferreira et al. 2019a, Mehdizadeh et al. 2017, Ozoegwu 2019, Saggi & Jain 2019). These methods are known for their abilities in working with complex problems. Thus, they become powerful tools for ET_o modeling.

Among machine learning models, ANN has been used for the estimation of ETo by several authors (Antonopoulos & Antonopoulos 2017, Kumar et al. 2011, Ferreira et al. 2019a). Wang et al. (2011), using ANN to estimate ET_o in arid regions of Africa, reported that this technique outperformed empirical equations. Ferreira et al. (2019a), evaluating temperature-based ANN in several places of Brazil, reported better results of this technique over empirical equations. Kumar et al. (2011) evaluated several studies and concluded that ANN is superior to conventional methods.

Another promising technique for the estimation of ET_o is multivariate adaptive regression splines (MARS). This is a nonparametric regression analysis used to study nonlinear relations between a response variable and a set of predictor variables (Koc & Bozdogan 2015). Mehdizadeh et al. (2017), working with several data availability scenarios, found that MARS was more efficient to estimate ET_o than empirical equations, SVM and GEP. Ferreira et al. (2019b) reported better results for MARS in relation to empirical equations in several climate types and data availability scenarios. In contrast with ANN, the use of a MARS model, after its development, occurs through an algebraic equation, which may facilitate the use of the final model. Despite its potential, there are limited studies using MARS

for the estimation of ET_o (Ferreira et al. 2019b, Mehdizadeh et al. 2017).

ET_o models can be calibrated/developed with local or regional data. The first case is the most common approach in the literature. However, a local model can show good performance in the station where it was developed and show poor performance in other stations, which can limit its real applicability or even make it useless (Kiafar et al. 2017, Reis et al. 2019). Thus, it is important to evaluate the generalization capacity of the models, assessing their performance outside the calibration/training station. On the other hand, regional models (i.e., models calibrated/developed with pooled data from several weather stations) can be key options in places without data for calibration or development of local models. In contrast with local models, regional models are developed to be used at any place of a particular region. In Brazil, studies addressing the development of regional models are scarce (Ferreira et al. 2019a, Reis et al. 2019, Zanetti et al. 2019).

In northern Minas Gerais, Brazil, a semiarid climate prevails. In this region, in addition to a large number of farms, there are public irrigation perimeters, where irrigation plays a fundamental role in the existence of a profitable agriculture. Thus, the development of studies that can contribute to a better irrigation and water resources management is of essential importance. In this context, this study evaluated the performance of ANN and MARS and the original and calibrated HS and PMT equations to estimate daily ET_o in a semiarid region of Minas Gerais, Brazil, considering two scenarios: (i) local, models were calibrated/developed and evaluated using data from individual weather stations; and (ii) regional, models were calibrated/developed using pooled data from several weather stations and evaluated independently in each one (leave-one-out

cross-validation). The local models were also evaluated considering a cross-station approach.

MATERIALS AND METHODS

Database and study area

Daily data from nine weather stations (2002-2016), obtained from the Meteorological Database for Teaching and Research (BDMEP) of the Brazilian National Institute of Meteorology (INMET), were used. The stations are located in northern Minas Gerais, Brazil, as shown in Figure 1. The main meteorological characteristics of the stations are presented in Table I.

Maximum and minimum air temperature, relative humidity, sunshine duration and wind speed were used. Wind speed, measured at 10 m height, was converted to 2 m and solar radiation was estimated based on sunshine duration, according to Allen et al. (1998). Days with missing data were removed. The dataset was divided into training set (2002-2011) and test set (2012-2016),

which were used to develop/calibrate the models and to test them, respectively. The mean numbers of samples (for each weather station) contained in the training and test sets were 3186 and 1312, respectively.

Methods for the estimation of ETo

To calibrate the PMT and HS equations and to develop the ANN and MARS models, as well as to evaluate these models, daily ET_o estimated using the FAO-PM equation with all required data (Equation 1) was adopted as reference.

$$ET_{oFAO\text{-}PM} = \frac{0.408\Delta\,(R_n\text{-}G)+\gamma\dfrac{900}{T+273}u_2\,(e_s\text{-}e_a)}{\Delta+\gamma\,(1+0.34u_2)} \tag{1}$$

where: $ET_{oFAO\text{-}PM}$ - reference evapotranspiration calculated by Penman-Monteith, mm d^{-1}; R_n - net solar radiation, MJ m^{-2} d^{-1}; G - soil heat flux, MJ m^{-2} d^{-1} (considered as null for daily estimates); T - daily mean air temperature, °C; u_2 - wind speed at 2 m height, m s^{-1}; e_s - saturation vapour pressure, kPa; e_a - actual vapour pressure, kPa;
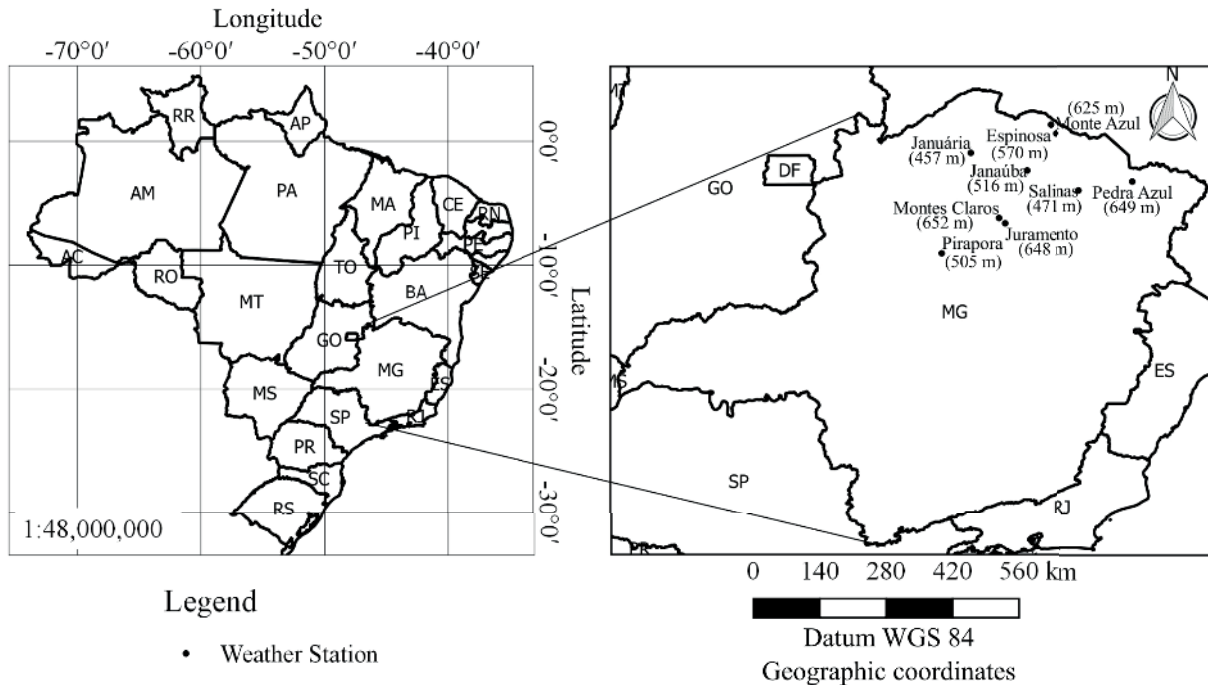
Figure 1. Geographic location and altitude of the weather stations, as well as political divisions of Brazil, highlighting the state of Minas Gerais (MG).

Table I. Daily mean values and standard deviations of meteorological variables for the weather stations used (2002-2016).

Station	T_x (°C)	T_n (°C)	n (h)	RH (%)	U_2 (m s^{-1})	ET_o (mm d^{-1})
Espinosa	31.1 (±2.7)	19.8 (±2.6)	8.3 (±3.1)	60.8 (±14.6)	2.1 (±1.1)	5.0 (±1.5)
Janaúba	32.1 (±2.8)	19.2 (±2.5)	8.1 (±2.9)	57.5 (±13.9)	1.1 (±1.1)	4.4 (±1.2)
Januária	32.2 (±2.8)	18.6 (±3.1)	8.3 (±3.3)	63.5 (±15)	1.5 (±0.6)	4.6 (±1.3)
Juramento	30.1 (±2.9)	16.8 (±3.3)	7.5 (±3.1)	67.5 (±13.2)	1.1 (±0.5)	4.0 (±1.1)
Monte Azul	31.3 (±2.8)	20.4 (±2.1)	8.1 (±3)	56.6 (±14)	1.8 (±1.2)	4.9 (±1.4)
Montes Claros	30.1 (±2.8)	18.0 (±2.9)	7.7 (±3.4)	62.2 (±14.8)	1.4 (±0.5)	4.2 (±1.2)
Pedra Azul	28.8 (±3.3)	17.8 (±2.8)	6.7 (±3.4)	70.6 (±10.6)	1.0 (±0.5)	3.7 (±1.2)
Pirapora	31.7 (±2.9)	19.1 (±3.3)	7.7 (±3.1)	66.9 (±13.5)	1.2 (±0.5)	4.3 (±1.3)
Salinas	31.0 (±3.2)	18.6 (±3.2)	6.5 (±3.4)	68.9 (±13)	1.2 (±0.5)	4.0 (±1.3)

T_x - maximum air temperature; T_n - minimum air temperature; n - sunshine duration; RH - relative humidity; U_2 - wind speed at 2 m height; ET_o - reference evapotranspiration estimated by the Penman-Monteith equation.

Δ - slope of the saturation vapour pressure function, kPa °C^{-1}; γ - psychometric constant, kPa °C^{-1}.

Two data management scenarios were used in this study: local scenario: models were calibrated/developed and evaluated using data from each weather station individually; and regional scenario: models were calibrated/developed using pooled data from all the weather stations, except the station in which the model was evaluated, performing a 9-fold cross-validation (leave-one-out cross-validation) (Figure 2). The local models were also evaluated considering a cross-station evaluation, evaluating them outside the calibration/training station (Figure 2). All the models studied were calibrated/developed using data from 2002 to 2011 (ten years) and evaluated using data from 2012 to 2016 (five years). Both cross-station evaluation for local models and leave-one-out cross-validation for regional models are important strategies to assess the performance of the models outside the training station, allowing a more robust evaluation.

Regional scenario leave-one-out cross-validation   Local scenario cross-station evaluation

■ Training station (2002-2011)   ■ Test station (2012-2016)
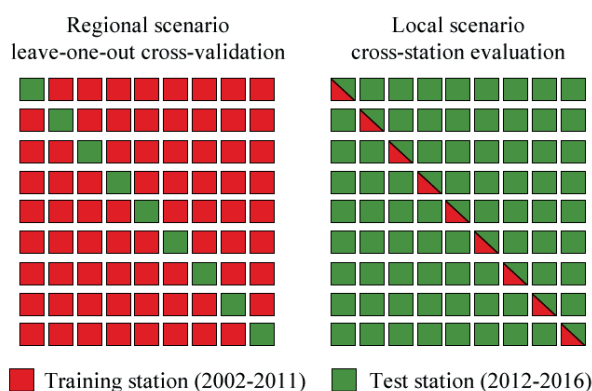
Figure 2. Data management scenarios used in the study.

To estimate ET_o with the PMT equation, Equation 1 was used, with actual vapour pressure and solar radiation estimated using Equations 2 and 3, respectively, and wind speed was set at 2 m s^{-1}, as recommended by Allen et al. (1998).

$$e_a = 0.611 \exp\left[\frac{17.27 T_{min}}{T_{min} + 237.3}\right] \qquad (2)$$

where: e_a - actual vapour pressure, kPa; T_{min} - minimum air temperature, °C.

$$Rs = 0.16\,Ra\,(T_{max} - T_{min})^{0.5} \qquad (3)$$

where: Rs - solar radiation, MJ m^{-2} d^{-1}; R_a - extraterrestrial radiation, MJ m^{-2} d^{-1}; T_{max} - maximum air temperature, °C; T_{min} - minimum air temperature, °C.

To estimate ET$_o$ using the HS equation, the following equation was implemented:

$$ET_{oHS} = 0.0023 R_a\,(T + 17.8)\,(T_{max} - T_{min})^{0.5} \qquad (4)$$

where: ET$_{oHS}$ - reference evapotranspiration calculated by Hargreaves-Samani, mm d^{-1}; R_a - extraterrestrial radiation, mm d^{-1}; T_{max} - maximum air temperature, °C; T_{min} - minimum air temperature, °C; T - mean air temperature, °C.

The calibrations of the PMT and HS equations were performed by simple linear regression, as suggested by Allen et al. (1998). For this, a linear regression was fitted with ET$_o$ values estimated using the FAO-PM equation as the dependent variable and those estimated using the equation under evaluation as the independent variable. The obtained intercept (a) and slope (b) were used as calibration parameters, according to Equation 5.

$$ET_{ocal} = a + b(ET_o) \qquad (5)$$

where: ET$_{ocal}$ - calibrated reference evapotranspiration, mm d^{-1}; a and b - calibration parameters; ET$_o$ - reference evapotranspiration estimated by the original equation (equation under study), mm d^{-1}.

Regarding the machine learning methods, ANN and MARS were developed considering maximum temperature, minimum temperature and extraterrestrial radiation as input variables.

ANN is a supervised machine learning model inspired by the human brain that can be used for classification and regression tasks. It typically consists of layers of neurons, with weights representing the connections between neurons. Further details regarding ANN and its usage for ET$_o$ modeling can be seen in Kumar et al. (2011).

ANNs of the feed-forward multilayer perceptron type with stochastic gradient descent training algorithm optimized with momentum term were used. The ANNs architecture (i.e., number of layers and neurons), momentum term and learning rate were defined by trial and error. Thus, the ANNs developed were composed of an input layer, one hidden layer and an output layer. The input layer was composed of three variables, the hidden layer was composed of ten neurons, and the output layer was composed of one neuron, as shown in Figure 3. Hyperbolic tangent function was used as activation function in the hidden layer and identity function was used in the output layer. Learning rate and momentum term were set to 0.001 and 0.9, respectively. The number of training epochs was 500 in local scenario and 400 in regional scenario. ANN models were implemented using the TensorFlow and Keras libraries for the Python programming language.

Before ANN training, to avoid convergence problems, input and output data were standardized according to the following equation. The mean (μ) and standard deviation (σ) were calculated with data from the training set (2002-2011).

$$x_{ni} = \frac{x_i - \mu}{\sigma} \qquad (6)$$

where: x_{ni} - standardized value; x_i - observed value; μ - mean; σ - standard deviation.

Multivariate adaptive regression splines (MARS) is a regression technique initially proposed by Friedman (1991). This technique is able to model nonlinearities and interactions and automatically choose the input variables that are really important. A MARS model is composed of base functions, which are set at
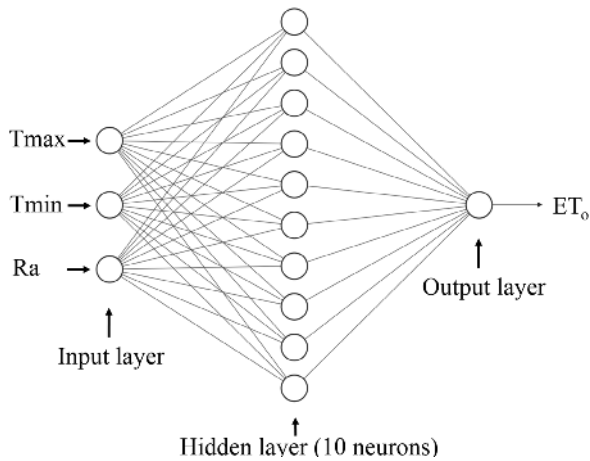
Figure 3. Artificial neural network architecture used in the study.

$y = 1.26\max(0, x - 30.1) + 1.24\max(0, 30.1 - x)$

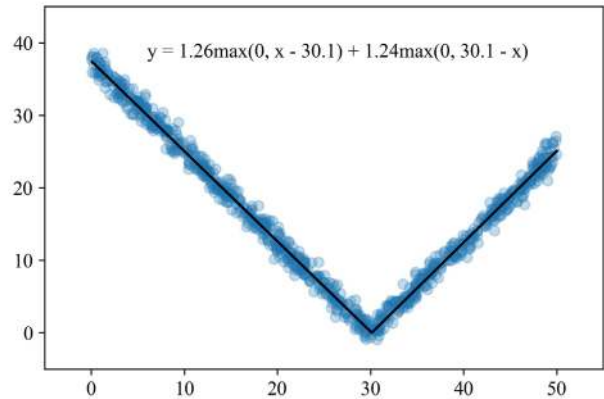**Figure 4.** One-dimensional MARS model example.

different intervals of the independent variables. Base functions work according to the following equations:

$$y = \max(0, x - c) \tag{7}$$

$$y = \max(0, c - x) \tag{8}$$

where: c - constant called knot; x - input variable; y - output variable.

To build a MARS model, two steps are required, the forward and backward steps. In the first one, an over-fitted model is built, with a large number of knots; in the backward step, a pruning technique is used to remove redundant knots (Kisi 2015). More details regarding MARS can be seen in Cheng & Cao (2014). As an example, a one-dimensional model is illustrated in Figure 4. MARS models were implemented using the py-earth library for the Python programming language. Hyperparameter tuning was done by grid search with k-fold cross-validation (k=3). The following hyperparameters and their respective values were assessed: penalty (3.0, 5.0, 10.0, 20.0, 30.0), endspan_alpha (0.01, 0.05, 0.1, 0.5), and minspan_alpha (0.01, 0.05, 0.1, 0.5). The order of interaction (max_degree) was limited to four to avoid extremely complex equations.

Performance comparison criteria

The performance of the models was evaluated for each weather station, with data from the test set, using root mean square error (RMSE), coefficient of determination (R^2) and mean bias error (MBE), according to the following equations.

$$RMSE = \sqrt{\frac{1}{n}\sum(Pi - Oi)^2} \tag{9}$$

$$R^2 = \left[\frac{\sum(P_i - \bar{P})(O_i - \bar{O})}{\sqrt{(\sum(P_i - \bar{P})^2)(\sum(O_i - \bar{O})^2)}}\right]^2 \tag{10}$$

$$MBE = \frac{1}{n}\sum(Pi - Oi) \tag{11}$$

where: RMSE - root mean square error, mm d^{-1}; R^2 - coefficient of determination; MBE - mean bias error, mm d^{-1}; Pi - value predicted by the model, mm d^{-1}; Oi - observed value, mm d^{-1}; - mean of values predicted by the model, mm d^{-1}; - mean of observed values, mm d^{-1}; n - number of data pairs.

RESULTS AND DISCUSSION

Empirical equations

The original PMT and HS equations had a wide performance variation between the weather stations, with RMSE ranging from 0.64 to 1.45 and from 0.70 to 1.29 mm d^{-1} and MBE ranging from -0.99 to 0.63 and from -0.68 to 0.92 mm d^{-1}, respectively for the PMT and HS equations (Figure 5). For R^2, a variation from 0.34 to 0.77 was observed for both equations.

According to Raziei & Pereira (2013), empirical equations have their performance affected according to the climatic conditions of the place where they are applied, reinforcing the need for calibration. Empirical equations typically show poorer performance in conditions different from those where they were developed. Sentelhas et al. (2010) also reported a wide performance variation for the PMT and HS equations, with RMSE ranging from 0.90 to 1.40 and from 0.75 to 1.95 mm d^{-1}, respectively.

Analyzing MBE behavior for the PMT and HS equations, both equations obtained negative values only in Espinosa and Monte Azul stations. This is possibly explained by the higher mean wind speed and lower mean relative humidity found in these sites (Table I). According to Allen et al. (1998), wind has a great effect on ET$_o$ in dry and hot environments due to the greater removal of water vapour stored in the air. In addition, Gavilán et al. (2006) found that the HS equation underestimated ET$_o$ in cases in which wind speed exceeded 1.5 m s^{-1}.

By evaluating R^2 results, the lowest R^2 values were observed at Espinosa, Janaúba and Monte Azul weather stations, where there are the highest standard deviations of wind speed, 1.1, 1.1 and 1.2 m s^{-1}, respectively (Table I). This is probably due to the difficulty of the PMT and HS equations in capturing the effect of large wind speed oscillations since it promotes ET$_o$

fluctuations that are not directly captured by these equations. Shiri (2017), working in a hyper-arid region, concluded that wind speed is one of the variables that most affects ET$_o$.

Comparing the original PMT and HS equations, the PMT equation outperformed the HS equation at almost all weather stations, except at Espinosa and Monte Azul stations, where the PMT equation had slightly higher RMSE values. Alencar et al. (2015), working with several stations in the state of Minas Gerais, also reported better performance for the PMT equation over the HS equation. Similarly, Ferreira et al. (2019a) found better performance of the PMT equation over the HS equation. This behavior is probably associated to the physical basis of the FAO-PM equation, which is partially conserved when considering the PMT equation.

After local calibration, performance improvements were observed for the PMT equation and the HS equation when they were evaluated in the stations where the calibrations were performed, with lower RMSE and MBE absolute values (Figure 5). R^2 is not affected by calibration based on linear regression. Although the PMT equation outperformed the HS equation before local calibration, the performance of both equations became very close after calibration, with similar RMSE, MBE and R^2 values. Calibration process incorporates local climatic characteristics into the model, making ET$_o$ estimates closer to the reference values. Several authors also reported improvements after local calibration of empirical equations (Kisi & Zounemat-Kermani 2014, Shiri 2017).

In the cross-station evaluation, an unstable behavior was observed after local calibration, with gains and losses of performance in relation to the original equations. For the PMT and HS equations, the models calibrated at Januária, Montes Claros, Pedra Azul and Pirapora stations had a relatively good generalization capacity,
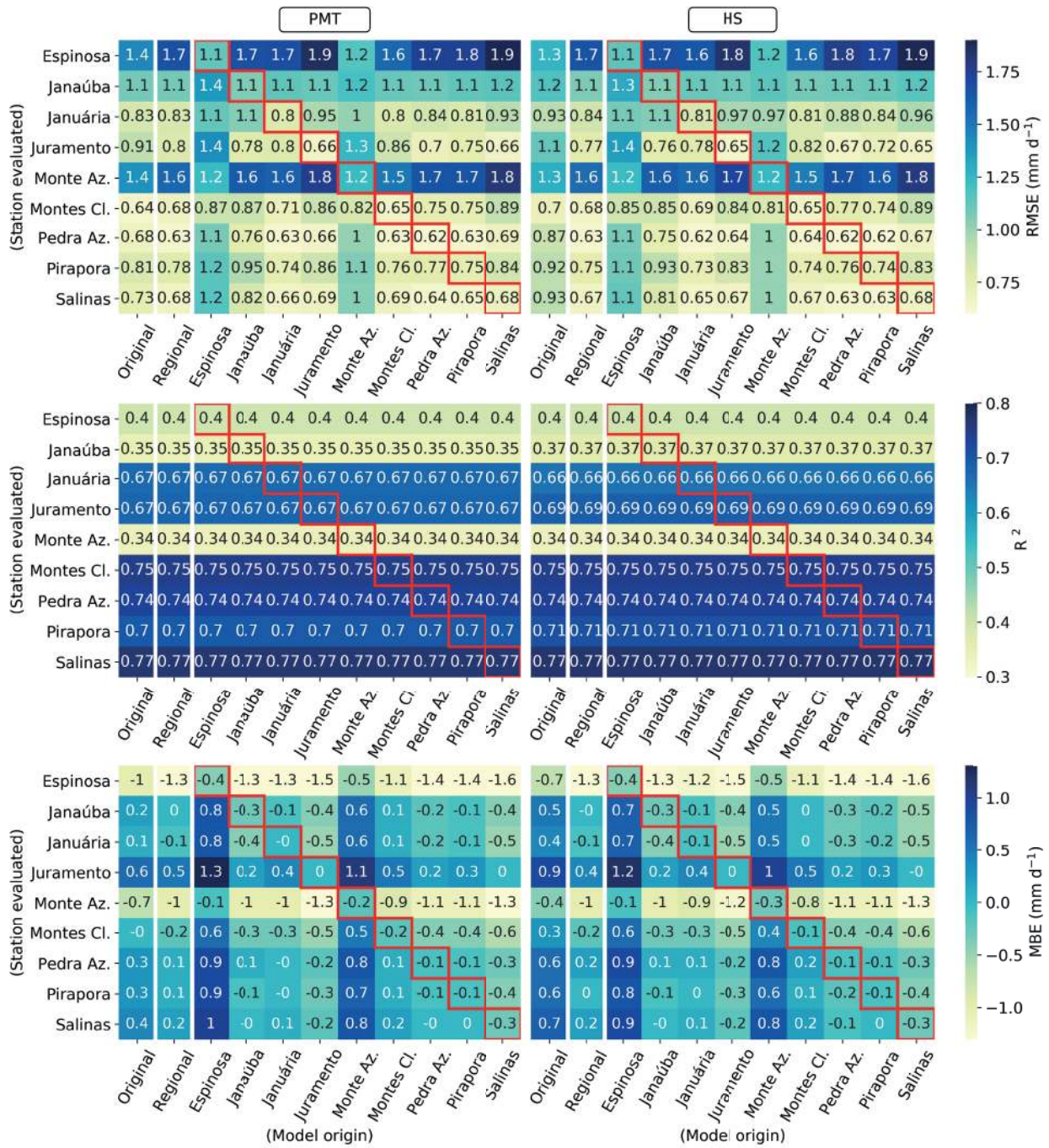
Figure 5. Statistical indices for the original and calibrated (local and regional) Penman-Monteith temperature (PMT) and Hargreaves-Samani (HS) equations.

showing performance improvements outside the calibration stations. These models exhibited RMSE values lower or close to those obtained for the original equations in most stations, however, they performed worse at Espinosa and Monte Azul stations. The models calibrated at the other stations showed performance improvements over the original equation only for some stations.

Regarding regional calibrations, the regional HS showed more expressive performance gains over its original version than the regional PMT over its original version. The regional HS only did not have lower RMSE values at Espinosa and Monte Azul stations, reducing RMSE for all other stations. Mean RMSE over the stations was reduced from 1.02 to 0.96 (6%) and median RMSE reduced from 0.93 to 0.77 (17%). For the PMT equation, although regional calibration reduced RMSE for some stations, it increased RMSE for Espinosa, Monte Azul and Montes Claros stations. Mean RMSE over the stations was increased from 0.95 to 0.97 (2%). However, median RMSE decreased from 0.83 to 0.80 (4%). Comparing the regional HS and PMT, they generally had the same performance.

Machine learning methods

The ANN and MARS models obtained similar performances in local and regional scenarios, but the ANN models performed a little better, with slightly lower RMSE values and slightly higher R^2 values (Figure 6). In the cross-station evaluation, as reported for the empirical equations, there was an unstable behavior. The models with the best results were those developed at Januária, Montes Claros and Pirapora stations. The models developed at Espinosa, Janaúba and Monte Azul had the worst results outside the training stations. On the other hand, the regional models had a more stable performance, with RMSE values higher than those obtained with

the local models, but lower than some of the values observed in the cross-station evaluation.

Overall evaluation

It is important to highlight that Espinosa, Janaúba and Monte Azul stations had the worst performances for all models studied. This is probably because there are larger oscillations of wind speed (greater standard deviations) in these sites (Table I), and all the models evaluated do not directly capture these oscillations since they do not use wind speed as input. In addition, the models calibrated/developed in these stations had the worst generalization capacities, not showing good performances outside the calibration/training stations.

The empirical equations, ANN and MARS models developed with local data outperformed the models developed with regional data when they were evaluated in the same station that they were calibrated/developed (Figure 7). Shiri et al. (2014) also obtained superior performance for models developed with local data. However, despite the higher performance of local models, they are commonly required in places where there are no data available to calibrate/develop them. Thus, a local model should be applied in places with climatic characteristics similar to the place where it was calibrated/developed, which limits its use. If this requirement is not met, the calibrated model can perform even worse than the original one.

In this study, it was observed that, in some cases, the use of a model developed or calibrated at a station more distant can provide better results than a model from a nearer station. For example, all the models developed at Januária station performed better at Montes Claros station than those developed at Juramento station, which is much closer to Montes Claros station. On the other hand, models calibrated/developed on a regional
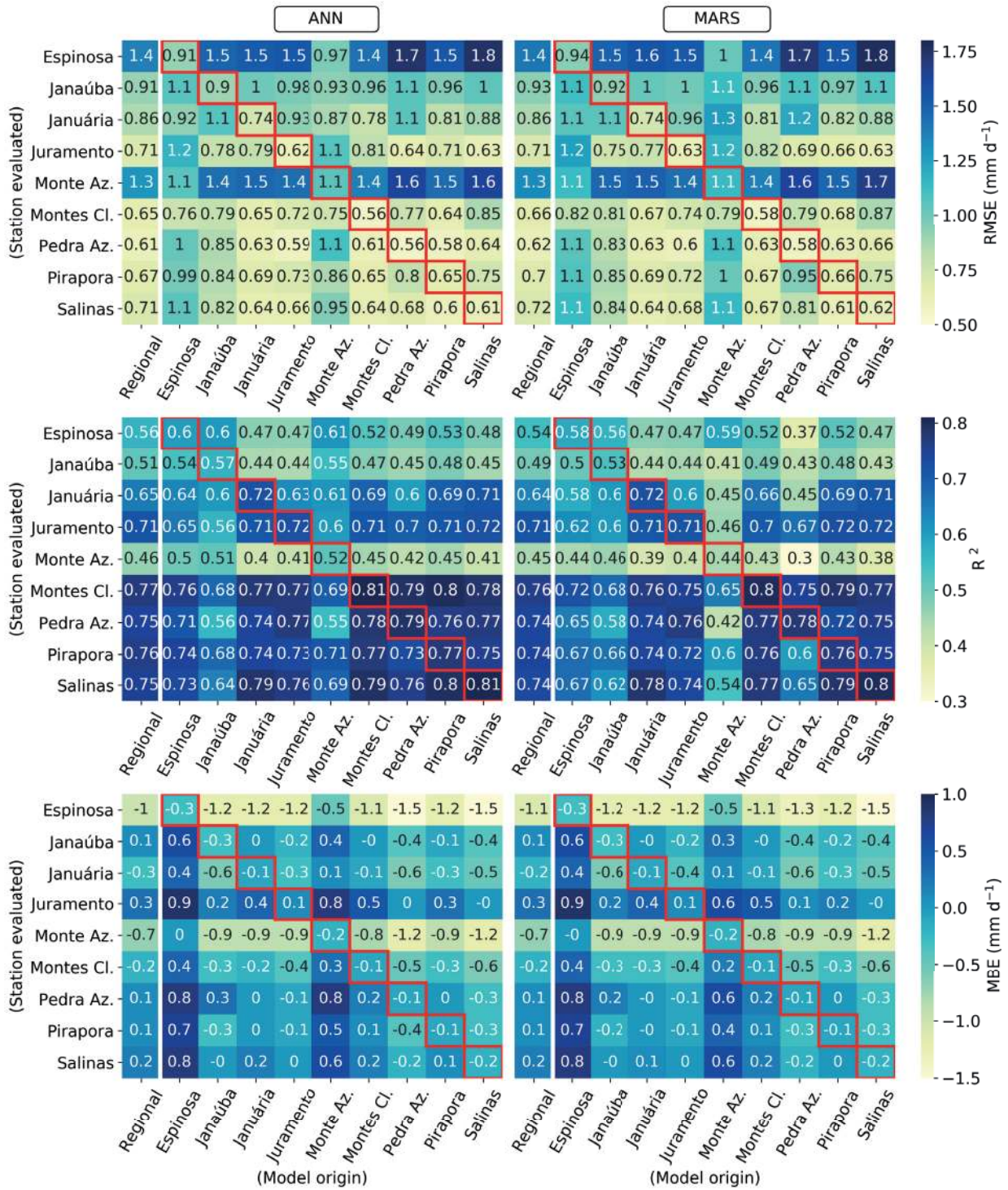
Figure 6. Statistical indices for the artificial neural network (ANN) and multivariate adaptive regression splines (MARS) models developed in local and regional scenarios.

scale can be a more flexible approach, allowing to use a single model in an entire region and avoiding problems with highly site-specific models. Therefore, regional models can be an interesting approach, especially for places without data for calibration/development of local models. In addition, according to Pereira et al. (2015), machine learning models remain empirical and may not translate well in time and space. Thus, since regional models are developed with a larger amount of data, they can be more stable in time and space than local models.

When comparing the machine learning models and the empirical equations, the first ones showed better performances in both regional and local scenarios (Figure 7). When comparing the ANN models with the HS equation, mean RMSE reduced from 0.96 to 0.87 (9%) and from 0.84 to 0.73 (13%), in regional and local scenarios, respectively. It should be noted that at Espinosa, Janaúba and Montes Azul stations, there was a high increase in R^2 values, mainly for the local models, indicating the higher capacity of machine learning models

to capture complex relations between input variables and ET_o (Figures 5 and 6). This behavior reaffirms the superiority of machine learning models over traditional equations, reported by Kumar et al. (2011) and Mehdizadeh et al. (2017), among others.

Comparing the performance of the regional ANN and MARS with the PMT and HS equations, it was noted that the machine learning models perform better than the original and regionally calibrated versions of the mentioned equations (Figure 7) and, at Janaúba and Pirapora stations, even better than the locally calibrated equations. Shiri et al. (2014) and Feng et al. (2017) also reported superior performance of machine learning methods developed with regional data in relation to empirical equations. Thus, the regional ANN and MARS are good options to estimate ET_o in the study region, outperforming traditional equations. Future studies should focus on the development of regional models with higher performances, trying to get even closer to the performance of local models.

Although the ANN models performed slightly better than the MARS models, both models
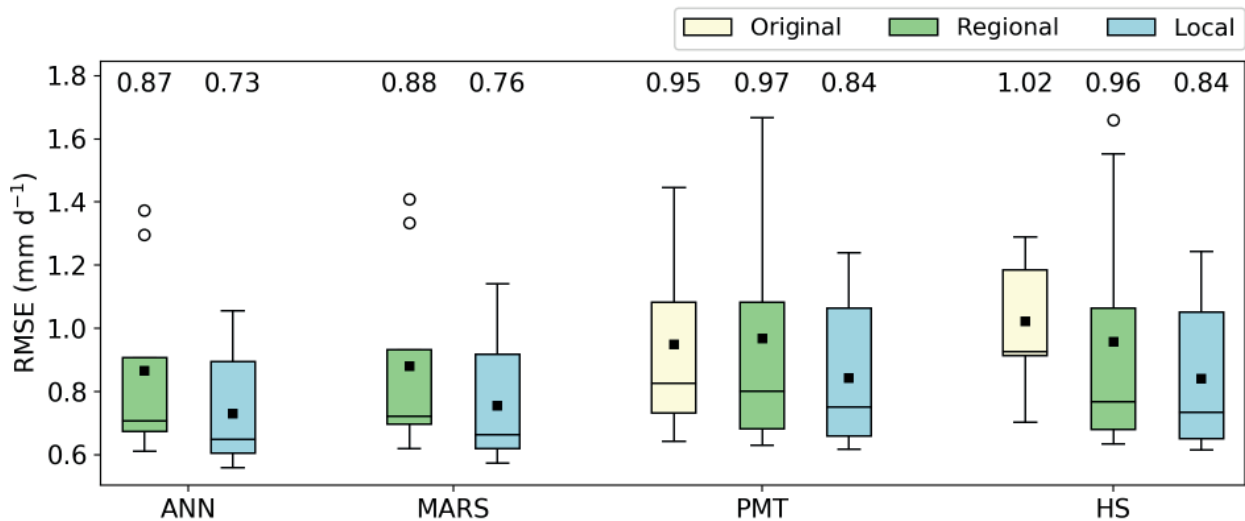
Figure 7. Boxplots and mean values of RMSE for the artificial neural network (ANN) and multivariate adaptive regression splines (MARS) models, as well as the Penman-Monteith temperature (PMT) and Hargreaves-Samani (HS) equations developed/calibrated in local and regional scenarios.

Table II. Local and regional calibration parameters for the PMT and HS equations.

Model origin	PMT		HS	
	a	b	a	b
Espinosa	0.1832	1.0938	0.2927	0.9945
Janaúba	1.5807	0.5783	1.5588	0.5470
Januária	-1.0889	1.1779	-0.8543	1.0616
Juramento	-0.1206	0.9032	-0.0707	0.8399
Monte Azul	0.9553	0.8916	1.1128	0.7932
Montes Claros	-0.7102	1.1271	-0.5210	1.0150
Pedra Azul	-0.5525	1.0350	-0.5061	0.9514
Pirapora	-1.0871	1.1576	-0.8631	1.0394
Salinas	-0.7581	1.0309	-0.7361	0.9608
Regional	-0.0951	0.9762	-0.0390	0.9015

Table III. Local and regional MARS models obtained in the study.

Model origin	MARS model
Espinosa	$-41.423 + 2.7196e^{-1}.R_a^2 - 5.0390e^{-3}.R_a^2.h(-T_n + 25.0) - 5.2382e^{-3}.R_a.T_x.T_n + 1.0254e^{-1}.T_x^2 + 3.9435e^{-3}.R_a.T_x^2 - 2.9581e^{-3}.T_x^3 + 1.0398e^{-3}.R_a^2.h(-T_x + 38.0) - 3.7558e^{-1}.R_a.T_n + 6.0016.T_n - 3.5143e^{-4}.R_a.T_n^2 - 2.6005e^{-3}.R_a^3 + 1.0428e^{-2}.R_a.T_x.h(T_n - 25.0) - 6.3394e^{-3}.R_a.T_x.h(-T_n + 25.0)$
Janaúba	$-325.57 + 29.956.T_x - 8.8447e^{-1}.R_a.T_x + 10.995.R_a - 3.4319e^{-2}.R_a^2 - 9.9053e^{-1}.T_x^2 + 1.0720e^{-2}.T_x^3 + 1.6784e^{-2}.R_a.T_x^2 - 2.9399e^{-4}.R_a.T_x^2.h(T_x - 39.2) + 3.0681e^{-4}.R_a.T_x^2.h(-T_x + 39.2) - 8.3129e^{-4}.R_a^2.T_n - 4.694e^{-5}.R_a^2.T_n^2 + 3.0124e^{-5}.R_a^3.T_n - 2.6418e^{-5}.T_x^3.T_n + 4.3985e^{-5}.R_a.T_x.T_n^2 + 8.2571e^{-4}.T_x^2.T_n$
Januária	$-28.559 - 6.6592e^{-2}.R_a.T_x - 2.2803e^{-3}.R_a.T_x.T_n + 1.8621.R_a + 2.1953e^{-2}.R_a.h(T_n - 24.4) + 1.6362e^{-2}.R_a.h(-T_n + 24.4) + 4.6810e^{-2}.T_x^2 - 1.7234e^{-3}.T_x^3 - 1.2138e^{-3}.R_a^2.h(-T_n + 24.4) + 3.5961e^{-4}.R_a.T_n.h(-T_n + 24.4) + 2.0414e^{-3}.R_a.T_x^2 + 1.5642e^{-3}.T_x^2.T_n$
Juramento	$9.0233 - 1.0103e^{-1}.R_a.T_x + 1.0755.R_a - 1.0932e^{-2}.R_a.T_n - 6.1662e^{-3}.R_a^2 + 3.1937e^{-4}.R_a.T_x.T_n + 1.0627e^{-1}.T_x^2 - 1.6991e^{-3}.T_x^3 + 1.0227e^{-3}.R_a.T_x^2 - 5.2177e^{-2}.T_x.h(-R_a + 41.4047)$
Monte Azul	$-13.936 + 4.6869e^{-2}.T_n.h(-T_x + 37.6) - 1.8987e^{-1}.R_a.h(-T_n + 24.4) + 7.291e^{-2}.R_a.h(-T_x + 37.6) + 1.0629e^{-2}.R_a^2.h(-T_n + 24.4) + 5.1516e^{-1}.R_a + 1.6835e^{-4}.T_x^2.T_n.h(-T_x + 37.6) - 2.1195e^{-4}.R_a.T_x.T_n.h(-T_x + 37.6) + 2.115e^{-3}.R_a^2.h(-T_x + 37.6) + 5.3726e^{-4}.R_a^2.h(T_x - 37.6).h(-T_n + 24.4) - 1.4488e^{-4}.R_a^2.h(-T_x + 37.6).h(-T_n + 24.4) - 1.5856e^{-4}.R_a^3.h(-T_n + 24.4) - 5.4953e^{-3}.R_a.T_n.h(-T_x + 37.6) + 8.2211e^{-4}.R_a.T_n.h(-T_n + 24.4)$
Montes Claros	$19.826 - 3.5816.T_x + 1.9231e^{-2}.R_a.T_x - 2.3466e^{-5}.R_a^2.T_x.T_n + 3.0637e^{-2}.R_a.T_n - 1.0877e^{-3}.R_a.T_n^2 + 1.5385e^{-1}.T_x^2 + 7.5849e^{-5}.R_a.T_x.T_n.h(T_x - 38.1) - 7.5083e^{-5}.R_a.T_x.T_n.h(-T_x + 38.1) - 2.4163e^{-3}.T_x^3 + 2.4051e^{-5}.R_a^2.T_n^2 + 5.168e^{-1}.T_n$
Pedra Azul	$3.6091 - 1.855.T_x - 7.4244e^{-2}.R_a.T_x + 1.0663.R_a - 1.0333e^{-2}.R_a^2 + 3.2964e^{-4}.R_a^2.T_n + 9.8423e^{-2}.T_x^2 - 2.0805e^{-3}.T_x^3 + 1.5451e^{-3}.R_a.T_x^2 + 8.9347e^{-4}.T_x^2.T_n + 1.2795e^{-3}.R_a.T_x.h(-T_n + 23.0)$
Pirapora	$-6.6135 - 2.8623.T_x - 8.5331e^{-2}.R_a.T_x + 2.7956.R_a - 2.4393e^{-2}.R_a^2 + 8.8371e^{-4}.R_a^2.T_n + 1.4499e^{-1}.T_x^2 - 2.4422e^{-3}.T_x^3 + 1.6963e^{-3}.R_a.T_x^2 - 5.5077e^{-2}.R_a.T_n + 4.5415e^{-4}.T_x^2.T_n - 3.1871e^{-4}.R_a.T_n^2 + 6.6413e^{-1}.T_n$
Salinas	$7.8813 - 1.3735.T_x + 8.8493e^{-3}.R_a.T_x + 3.063e^{-5}.R_a.T_x.T_n.h(T_x - 39.0) - 3.7837e^{-5}.R_a.T_x.T_n.h(-T_x + 39.0) - 7.5465e^{-6}.R_a^2.T_x.T_n + 6.0612e^{-2}.T_x^2 - 9.4987e^{-4}.T_x^3 + 1.9254e^{-2}.R_a.T_n - 4.7952e^{-6}.R_a.T_x.T_n^2$
Regional	$-62.27 + 4.0581.T_x - 2.1050e^{-1}.R_a.T_x + 2.9354.R_a - 8.3421e^{-4}.R_a^2.T_n + 6.6552e^{-4}.R_a.T_x.T_n - 8.4280e^{-2}.T_x^2 + 6.4573e^{-3}.R_a.T_x^2 - 4.9517e^{-5}.R_a.T_x^3 + 1.3600e^{-5}.R_a^3.T_n - 7.7994e^{-6}.R_a^2.T_n^2 - 5.1786e^{-4}.R_a^2.T_x$

T_x – maximum air temperature, °C; T_n – minimum air temperature, °C; R_a – extraterrestrial radiation, MJ m^{-2} d^{-1}; $h(x) = \max(0, x)$.

presented similar performances in regional and local scenarios (Figure 7). These results indicate that, in addition to ANN, which has already been considered by several authors as an efficient method (Kumar et al. 2011, Yassin et al. 2016, Antonopoulos & Antonopoulos 2017, Ferreira et al. 2019a), MARS models can also be used for the estimation of ET_o in cases where only air temperature data are available. Mehdizadeh et al. (2017), analyzing MARS, empirical equations, SVM and GEP, concluded that MARS was the most efficient technique in several data availability scenarios, including the one used in this study. Ferreira et al. (2019b), comparing MARS and empirical equations in several climate types and data availability scenarios, also reported superior performance for MARS. It is also important to remember that MARS can be used in the form of an algebraic equation, which can make it simpler to use by an end user.

To make the models obtained in this study available for future studies or practical applications, the local and regional calibration parameters of the PMT and HS equations, as well as the regional and local MARS models, are presented in Tables II and III, respectively.

CONCLUSIONS

ANN, MARS, and empirical equations (PMT and HS) in their original and calibrated forms were evaluated for the estimation of daily ET_o based on temperature data. Two scenarios were considered: (i) local, models were calibrated/developed and evaluated using data from individual weather stations; (ii) regional, models were calibrated/developed using pooled data from several weather stations and evaluated independently in each station (leave-one-out cross-validation). The local models were also evaluated considering a cross-station approach.

The original PMT equation exhibited better performance than the original HS equation, however, after local or regional calibrations, these had similar performances.

The local calibration of the PMT and HS equations promoted higher performance gains than those obtained with regional calibration. Similarly, the local ANN and MARS had better performance than their regional versions. However, the regional empirical equations, ANN and MARS models had higher generalization capacity, showing a more stable performance between the stations evaluated.

The machine learning techniques studied had better performance than the PMT and HS equations in their original and calibrated forms in local and regional scenarios. The ANN and MARS models showed similar performances, however, the ANN models performed slightly better. On the other hand, MARS has the advantage that it can be used in the form of algebraic expression.

Acknowledgments

REFERENCES

ALENCAR LP, SEDIYAMA GC & MANTOVANI EC. 2015. Estimation of reference evapotranspiration (ETo) under FAO standards with missing climatic data in Minas Gerais, Brazil. Eng Agríc 35: 39-50. doi:10.1590/1809-4430-Eng. Agric.v35n1p39-50/2015.

ALLEN RG, PEREIRA LS, RAES D & SMITH M. 1998. Crop evapotranspiration: Guidelines for computing crop water requirements. FAO Irrigation and Drainage Paper 56. FAO, Rome, 300 p.

ALMOROX J, SENATORE A, QUEJ VH & MENDICINO G. 2018. Worldwide assessment of the Penman–Monteith temperature approach for the estimation of monthly reference evapotranspiration. Theor Appl Climatol 131: 693-703. doi:10.1007/s00704-016-1996-2.

ANTONOPOULOS VZ & ANTONOPOULOS AV. 2017. Daily reference evapotranspiration estimates by artificial neural networks technique and empirical equations using limited input climate variables. Comput Electron Agric 132: 86-96. doi:10.1016/j.compag.2016.11.011.

CHENG MY & CAO M. 2014. Accurately predicting building energy performance using evolutionary multivariate adaptive regression splines. Appl Soft Comput 22: 178-188. doi:10.1016/j.asoc.2014.05.015.

FENG Y, PENG Y, CUI N, GONG D & ZHANG K. 2017. Modeling reference evapotranspiration using extreme learning machine and generalized regression neural network only with temperature data. Comput Electron Agric 136: 71-78. doi:10.1016/j.compag.2017.01.027.

FERREIRA LB, CUNHA FF, OLIVEIRA RA & FERNANDES FILHO EI. 2019a. Estimation of reference evapotranspiration in Brazil with limited meteorological data using ANN and SVM – a new approach. J Hydrol 572: 556-570. doi:10.1016/j.jhydrol.2019.03.028.

FERREIRA LB, DUARTE AB, CUNHA FF & FERNANDES FILHO EI. 2019b. Multivariate adaptive regression splines (MARS) applied to daily reference evapotranspiration modeling with limited weather data. Acta Sci Agron 41: e39880. doi:10.4025/actasciagron.v41i1.39880.

FRIEDMAN JH. 1991. Multivariate Adaptive Regression Splines. Ann Stat 19: 1-67. doi:10.1214/aos/1176347963.

GAVILÁN P, LORITE IJ, TORNERO S & BERENGENA J. 2006. Regional calibration of Hargreaves equation for estimating reference et in a semiarid environment. Agric Water Manag 81: 257-281. doi:10.1016/j.agwat.2005.05.001.

KIAFAR H, BABAZADEH H, MARTI P, KISI O, LANDERAS G, KARIMI S & SHIRI J. 2017. Evaluating the generalizability of GEP models for estimating reference evapotranspiration in distant humid and arid locations. Theor Appl Climatol 130: 377-389. doi:10.1007/s00704-016-1888-5.

KISI O. 2015. Pan evaporation modeling using least square support vector machine, multivariate adaptive regression splines and M5 model tree. J Hydrol 528: 312-320. doi:10.1016/j.jhydrol.2015.06.052.

KISI O & ZOUNEMAT-KERMANI M. 2014. Comparison of Two Different Adaptive Neuro-Fuzzy Inference Systems in Modelling Daily Reference Evapotranspiration.

Water Resour Manag 28: 2655-2675. doi:10.1007/s11269-014-0632-0.

KOC EK & BOZDOGAN H. 2015. Model selection in multivariate adaptive regression splines (MARS) using information complexity as the fitness function. Mach Learn 101: 35-58. doi:10.1007/s10994-014-5440-5.

KUMAR M, RAGHUWANSHI NS & SINGH R. 2011. Artificial neural networks approach in evapotranspiration modeling: A review. Irrig Sci 29: 11-25. doi:10.1007/s00271-010-0230-8.

MEHDIZADEH S, BEHMANESH J & KHALILI K. 2017. Using MARS, SVM, GEP and empirical equations for estimation of monthly mean reference evapotranspiration. Comput Electron Agric 139: 103-114. doi:10.1016/j.compag.2017.05.002.

OZOEGWU CG. 2019. Artificial neural network forecast of monthly mean daily global solar radiation of selected locations based on time series and month number. J Clean Prod 216: 1-13. doi:10.1016/j.jclepro.2019.01.096.

PEREIRA LS, ALLEN RG, SMITH M & RAES D. 2015. Crop evapotranspiration estimation with FAO56: Past and future. Agric Water Manag 147: 4-20. doi:10.1016/j.agwat.2014.07.031.

PINHEIRO MAB, OLIVEIRA ALM, BORGES JÚNIOR JCF, OLIVEIRA ECD & CARVALHO LGD. 2019. Reference evapotranspiration based on temperature in Minas Gerais state, Brazil. Ciênc Agrotec 43: e004219. doi:10.1590/1413-7054201943004219.

RAZIEI T & PEREIRA LS. 2013. Estimation of ETo with Hargreaves-Samani and FAO-PM temperature methods for a wide range of climates in Iran. Agric Water Manag 121: 1-18. doi:10.1016/j.agwat.2012.12.019.

REIS MM, DA SILVA AJ, JUNIOR JZ, SANTOS LDT, AZEVEDO AM & LOPES EMG. 2019. Empirical and learning machine approaches to estimating reference evapotranspiration based on temperature data. Comput Electron Agric 165: 104937. doi:10.1016/j.compag.2019.104937.

SAGGI MK & JAIN S. 2019. Reference evapotranspiration estimation and modeling of the Punjab Northern India using deep learning. Comput Electron Agric 156: 387-398. doi:10.1016/j.compag.2018.11.031.

SENTELHAS PC, GILLESPIE TJ & SANTOS EA. 2010. Evaluation of FAO Penman-Monteith and alternative methods for estimating reference evapotranspiration with missing data in Southern Ontario, Canada. Agric Water Manag 97: 635-644. doi:10.1016/j.agwat.2009.12.001.

SHIRI J. 2017. Evaluation of FAO56-PM, empirical, semi-empirical and gene expression programming approaches for estimating daily reference evapotranspiration in

hyper-arid regions of Iran. Agric Water Manag 188: 101-114. doi:10.1016/j.agwat.2017.04.009.

SHIRI J, NAZEMI AH, SADRADDINI AA, LANDERAS G, KISI O, FAKHERI FARD A & MARTI P. 2014. Comparison of heuristic and empirical approaches for estimating reference evapotranspiration from limited inputs in Iran. Comput Electron Agric 108: 230-241. doi:10.1016/j.compag.2014.08.007.

WANG YM, TRAORE S, KERH T & LEU JM. 2011. Modelling reference evapotranspiration using feed forward backpropagation algorithm in arid regions of Africa. Irrig Drain 60: 404-417. doi:10.1002/ird.589.

YASSIN MA, ALAZBA AA & MATTAR MA. 2016. Artificial neural networks versus gene expression programming for estimating reference evapotranspiration in arid climate. Agric Water Manag 163: 110-124. doi:10.1016/j.agwat.2015.09.009.

ZANETTI SS, DOHLER RE, CECÍLIO RA, PEZZOPANE JEM & XAVIER AC. 2019. Proposal for the use of daily thermal amplitude for the calibration of the Hargreaves-Samani equation. J Hydrol 571: 193-201. doi:10.1016/j.jhydrol.2019.01.049.

LUCAS B. FERREIRA[1]
https://orcid.org/0000-0001-6838-3114

FERNANDO F. DA CUNHA[1]
https://orcid.org/0000-0002-1671-1021

GUSTAVO H. DA SILVA[1]
https://orcid.org/0000-0002-1046-7646

FLAVIO B. CAMPOS[1]
https://orcid.org/0000-0002-1235-2447

SANTOS H.B. DIAS[2]
https://orcid.org/0000-0001-7800-5416

JANNAYTON E.O. SANTOS[1]
https://orcid.org/0000-0001-5301-8309

[1]Universidade Federal de Viçosa, Departamento de Engenharia Agrícola, Av. Peter Henry Rolfs, s/n, 36570-000 Viçosa, MG, Brazil

[2]Universidade Estadual de Ponta Grossa, Departamento de Solos e Engenharia Agrícola, Av. General Carlos Cavalcanti, 4748, Campus Uvaranas, 84030-900, Ponta Grossa, PR, Brazil

Correspondence to: **Lucas Borges Ferreira**
E-mail: contato.lucasbf@gmail.com

Author contributions

Lucas B. Ferreira: Conceptualization, Methodology, Software, Formal analysis, Writing - original draft. Fernando F. da Cunha: Conceptualization, Writing - review & editing, Formal analysis, Supervision. Gustavo H. da Silva and Jannaylton E.O. Santos: Writing - review & editing, Formal analysis. Flavio B. Campos: Writing - review & editing. Santos H.B. Dias: Writing - review & editing, Formal analysis, Visualization.