Article

Assessing the Robustness of Pan Evaporation Models for Estimating Reference Crop Evapotranspiration during Recalibration at Local Conditions

Konstantinos Babakos [1], Dimitris Papamichail [1], Panagiotis Tziachris [2], Vassilios Pisinaras [2], Kleoniki Demertzi [3] and Vassilis Aschonitis [2],*

[1] Department of Hydraulics, Soil Science and Agricultural Engineering, Faculty of Agriculture, Forestry and Natural Environment, Aristotle University of Thessaloniki, 54124 Thessaloniki, Greece; kbabakos@gmail.com (K.B.); papamich@agro.auth.gr (D.P.)

[2] Soil and Water Resources Institute, Hellenic Agricultural Organization-Demeter, 57001 Thermi-Thessaloniki, Greece; p.tziachris@swri.gr (P.T.); v.pisinaras@swri.gr (V.P.)

[3] Goulandris Natural History Museum, Greek Biotope/Wetland Centre, 57001 Thermi-Thessaloniki, Greece; kldemertzi@ekby.gr

* Correspondence: v.aschonitis@swri.gr; Tel.: +30-2310-473429

Abstract: A classic method for assessing the reference crop evapotranspiration (ET_o) is the pan evaporation (E_{pan}) method that uses E_{pan} measurements and pan coefficient (k_p) models, which can be functions of relative humidity (RH), wind speed (u_2), and temperature (T). The aim of this study is to present a methodology for evaluating the robustness of regression coefficients associated to climate parameters (RH, u_2, and T) in pan method models during recalibration at local conditions. Two years of daily data from April to October (warm season) of meteorological parameters, E_{pan} measurements from class A pan evaporimeter and ET_o estimated by ASCE-standardized method for the climatic conditions of Thessaloniki (Greece, semi-arid environment), were used. The regression coefficients of six general nonlinear (NLR) regression E_{pan} models were analyzed through recalibration using a technique called "random cross-validation nonlinear regression RCV-NLR" that produced 1000 random splits of the initial dataset into calibration and validation sets using a constant proportion (70% and 30%, respectively). The variance of the regression coefficients was analyzed based on the 95% interval of the highest posterior density distribution. NLR models that included coefficients with a 95% HPD interval that fluctuates in both positive and negative values were considered nonrobust. The machine-learning technique of random forests (RF) was also used to build a RF model that includes E_{pan}, u_2, RH, and T parameters. This model was used as a benchmark for evaluating the predictive accuracy of NLR models but, also, for assessing the relative importance of the predictor climate variables if they were all included in one NLR model. The findings of this study indicated that locally calibrated NLR functions that use only the E_{pan} parameter presented better results, while the inclusion of additional climate parameters was redundant and led to underfitting.

Keywords: class A pan evaporation; reference evapotranspiration; pan models' recalibration at local conditions; effects of climate parameters in pan evaporation models

1. Introduction

The parameter of reference crop evapotranspiration ET_o describes the maximum water losses that can be achieved by evaporation and transpiration from a field covered by a reference crop (e.g., grass) under no water restrictions, and it is used for water balance analysis and irrigation planning [1,2]. Various methods have been proposed for ET_o assessment [1–11], with the most popular being the

FAO-56 Penman-Monteith [1] from Food and Agriculture Organization (FAO) of the United Nations, and its updated version, the ASCE-standardized method [2], which reflects the current state-of-the-art, providing ET_o estimations for two reference crops (a short and a tall reference crop, which correspond to clipped grass and alfalfa, respectively). The ASCE-standardized method has been proposed by the Task Committee of Environmental and Water Resources Institute (EWRI) of the American Society of Civil Engineers (ASCE) as the most precise method and requires a complete set of the four climatic parameters of temperature, solar radiation, relative humidity, and wind speed, which, in many cases, are not all available.

The problem of data availability can be confronted by other methods of reduced parameters [3–11], among which, there is the pan evaporation method that differs from the others, because it requires direct evaporation measurements (E_{pan}) from an evaporimeter. In this method, ET_o is commonly calculated as the multiplication product of E_{pan} and a pan coefficient (k_p), which depends on the local conditions of the surrounding environment [1]. There are various types of evaporimeters, but the most commonly used type is the class A pan. The class A evaporation pan is circular, 120.7 cm in diameter and 25 cm deep. It is made of galvanized iron (22 gauge) or Monel metal (0.8 mm). The pan is mounted on a wooden open frame platform, which is 15 cm above ground level. The pan must be level. It is filled with water to 5 cm below the rim, and the water level should not be allowed to drop to more than 7.5 cm below the rim. The water should be regularly renewed, at least weekly, to eliminate extreme turbidity. The pan, if galvanized, should be painted annually with aluminum paint [1].

With respect to class A pan evaporation, various regression k_p models have been developed, which are functions of various parameters such as relative humidity, wind speed, temperature, and the fetch distance between the pan evaporation and boundary of grass covered or fallow soil surface [12–21]. The specific regression k_p models have been tested in many environments [19–60], while, in many cases, their regression coefficients have been recalibrated or their form has been changed (e.g., using machine-learning models) to describe better the local conditions [19,26,33,38,40,41,47,49].

There are four main issues during k_p models' recalibration or formation that should be considered. The first issue is that the predefined forms of k_p models are subjected to nonlinear regression analysis (nonlinear regression also includes machine-learning techniques) without investigating if the climate parameters inside a k_p model (e.g., RH, u_2, and T) have a robust effect.

The second issue is that the regression analysis is usually performed considering the k_p as dependent parameter. This approach does not allow a comparative analysis between the effects of E_{pan} and the other climate parameters (e.g., u_2, RH, and T), because E_{pan} measurements are embodied in the "measured" k_p values. This approach also creates other problems that are associated to the regression analysis. During a regression analysis, an optimization algorithm aims at minimizing the errors between observed and predicted values of the dependent variable, and because of this procedure, the observed k_p with larger values have a greater effect on the error measures [61] and, consequently, a greater effect on regulating the regression coefficients. For example, if the measured k_p values of colder days (lower ET_o) are larger from the respective values of warmer days (higher ET_o), which is not unusual [29,36,49], then the final calibrated k_p model will perform better at colder days with lower ET_o values. It is also worth mentioning that there are uncertainties in the validity of ET_o during the cold periods [7] and thus, the k_p values of those periods may not be accurate.

The third issue is that the estimated ET_o by a complete formula (e.g., ASCE-standardized method) and the measured E_{pan} evaporation are both affected by the same climate parameters, and someone would expect that solely E_{pan} could explain the largest portion of variance in estimated ET_o values, especially when there are no observed extreme deviations of meteorological parameters from normal conditions. That means that a model that considers only E_{pan} without the inclusion of other parameters could provide an adequate prediction accuracy (examples of such models are given in [25,35,43,46,47,60], with R^2 ranging between ~0.7–0.95), while the other climate parameters may be redundant. The reasons that could explain a low accuracy regression model between ET_o and E_{pan} are (a) the selection of a nonappropriate form of the regression model and (b) the introduction of experimental bias during ET_o

and E_{pan} measurements and, especially, during the measurement of real ET_o in lysimeter conditions where additional effects from the soil and the grass crop may occur [62]. It is also worth mentioning that there are very few cases of calibrated k_p models using measured ET_o from lysimeters [13,15,16,22,24,33], while, in most cases, ET_o is calculated using a benchmark method (e.g., ASCE-standardized method) that already uses a full set of climate parameters. Calibrating k_p models that include the same climate parameters used for calculating ET_o values may also be questionable from a statistical/mathematical point of view.

The fourth issue is that the relationship between ET_o and E_{pan} is usually defined a priori by $ET_o = k_p \cdot E_{pan}$ without investigating if these two parameters are related in a different way.

Considering the above issues, the objective of this study is to propose an integrated methodology for evaluating the robustness of regression coefficients and, consequently, their associated climate parameters that participate in general forms of pan evaporation models during recalibration at local conditions. In this study, the general forms of existing pan evaporation models are used as examples, and they are recalibrated at local conditions based on estimated daily ET_o values by the ASCE-standardized method for short grass (equivalent to the FAO-56 method). The proposed methodology that is going to be presented in this study allows to select the important climate parameters (e.g., RH, u_2, and T) in combination with E_{pan} measurements for building a robust model for the estimation of ET_o at local conditions.

2. Materials and Methods

2.1. Data

Daily meteorological data of temperature (T), relative humidity (RH), solar radiation (R_s), wind speed at 2 m above ground (u_2), and precipitation (P) covering the warm-dry cultivation periods (May to September) of the years 2008 and 2009 were obtained by the meteorological station of the Aristotle University Farm (40°32′08″ N, 22°59′18″ E, ~1 m a.s.l.) in Thessaloniki, Greece (Table 1). The daily values of climate parameters were calculated as mean values of hourly observations of a 24-h period.

The climate data were used for estimating daily ET_o based on the ASCE-standardized method [2]. Additionally, daily E_{pan} measurements were obtained during the same period using a class A pan evaporimeter of Monel metal with fetch distance F = 1 m for green upwind fetch (Case A) (Table 1). The data describe adequately the meteorological conditions of the warm-dry season in the Thessaloniki Plain in Greece, where the climate is considered a semi-arid Mediterranean environment. The period of May–September is the period for cultivating summer crops and the period of interest for calculating ET_o for irrigation purposes. The records of rainfall days ($p > 0$) during May–September were also excluded from the analysis, leading to a final number of 212 daily records of meteorological and E_{pan} data.

2.2. Reference Crop Evapotranspiration

ET_o was estimated in this study using the ASCE-standardized method at a daily step for short reference crops (clipped grass of 12 cm) by the following function [2]:

$$ET_o = [0.408 \cdot \Delta(R_n - G) + C_n \cdot \gamma \cdot u_2 (e_s - e_a)/(T + 273.16)]/[\Delta + \gamma(1 + C_d \cdot u_2)] \tag{1}$$

where ET_o is the reference crop evapotranspiration (mm d^{-1}), R_n is the net radiation at the crop surface (MJ m^{-2} d^{-1}), u_2 is the mean daily wind speed at 2-m height (m s^{-1}), T is the mean daily air temperature (°C), G is the soil heat flux density at the soil surface (MJ m^{-2} d^{-1}), e_s is the mean daily vapor pressure (kPa), e_a is the mean daily actual vapor pressure (kPa), Δ is the slope of the saturation vapor pressure-temperature curve (kPa °C^{-1}), γ is the psychrometric constant (kPa °C^{-1}), C_n and C_d are constants, which vary according to the time step, the reference crop type (bulk surface resistance

and aerodynamic roughness of the surface), and daytime/nighttime ratio. For short reference crop calculations at daily steps using Equation (1), the following values 900, 0.34, and 0 are considered for C_n, C_d, and G, respectively, as suggested by Allen et al. [1,2], providing equivalent ET_o results to the FAO-56 Penman-Monteith equation.

Table 1. Minimum, maximum, and average values of daily temperature (T), wind speed at 2 m above ground surface (u_2), relative humidity (RH), incident solar radiation (R_s), measured Class A pan evaporation (E_{pan}), and estimated reference evapotranspiration according to the ASCE-standardized method (Equation (1)) after excluding rainfall days.

Year	Month	T (°C)			u_2 (m s^{-1})			RH (%)		
		min	max	average	min	max	average	min	max	average
	May	20.6	26.1	22.7	1.3	1.8	1.5	50.8	72.7	58.7
	June	26.1	30.3	26.1	1.3	2.0	1.5	37.5	63.7	53.2
2008	July	26.9	29.0	26.9	1.2	3.3	1.7	31.7	65.4	48.0
	August	27.9	30.6	27.9	1.1	2.9	1.4	31.6	65.0	47.2
	September	16.7	27.0	23.6	1.0	3.0	1.3	38.9	68.9	56.0
	May	20.2	25.6	22.9	1.3	2.3	1.6	52.0	72.5	60.2
	June	21.4	27.9	24.3	1.2	3.3	1.5	34.5	74.2	58.7
2009	July	24.3	31.0	27.6	1.2	3.3	1.6	36.7	67.7	51.9
	August	22.4	29.4	26.3	1.0	1.5	1.2	45.9	79.5	61.9
	September	18.6	26.8	21.8	0.8	1.9	1.1	55.4	81.1	63.6

Year	Month	R_s (w m^{-2})			E_{pan} (mm day^{-1})			ET_o (mm day^{-1})		
		min	max	average	min	max	average	min	max	average
	May	294.8	330.6	312.8	6.0	9.6	7.6	5.0	6.8	5.8
	June	178.2	336.3	306.3	5.8	13.4	9.3	3.4	7.8	6.2
2008	July	193.3	336.7	305.8	5.3	14.9	9.9	3.9	8.9	6.4
	August	172.4	305.1	266.4	5.3	11.4	8.5	3.4	7.7	5.7
	September	121.0	248.6	216.4	2.5	7.4	5.7	2.6	4.7	4.1
	May	229.2	315.7	283.7	5.2	9.9	8.0	3.8	5.4	5.4
	June	216.5	347.7	303.8	5.2	13.0	8.7	3.8	8.4	5.8
2009	July	263.3	343.0	314.1	6.6	13.3	9.5	4.8	8.0	6.4
	August	194.7	297.4	254.3	4.6	9.3	7.3	3.2	6.1	4.9
	September	85.8	252.3	190.5	2.6	7.3	4.9	1.8	4.7	3.4

2.3. Pan Evaporation Method and General Forms for Local Conditions

The general equation of the pan evaporation method for estimating reference crop evapotranspiration ET_o is the following [63]:

$$ET_o = k_p \cdot E_{pan} \tag{2}$$

where E_{pan} is the measured evaporation from the evaporimeter (mm day^{-1}), and k_p is the pan evaporation coefficient. The range of k_p values is approximately between 0.35 to 1.1 for various pan types and various local conditions of climates and surrounding environments [14,63].

For the purposes of this study, the k_p models of Cuenca [13], Allen and Pruitt [14], Snyder [15], Pereira et al. [16], and Orang [17] were used. These models have been developed for estimating short reference crops based on measurements from class A pan evaporimeters and are given, respectively, by the following equations:

$$\begin{aligned}k_p = 0.475 - 2.4 \times 10^{-4}u_2 + 5.16 \times 10^{-3}RH + 1.18 \times 10^{-3}F - 1.6 \times 10^{-5}RH^2 \\ - 1.01 \times 10^{-6}F^2 - 8 \times 10^{-9}RH^2u_2 - 1 \times 10^{-8}RH^2F\end{aligned} \tag{3}$$

$$k_p = 0.108 - 3.31 \times 10^{-4}u_2 + 4.22 \times 10^{-2}\ln(F) + 14.34 \times 10^{-2}\ln(RH)$$
$$- 6.31 \times 10^{-4}[\ln(F)]^2\ln(RH) \tag{4}$$

$$k_p = 0.482 - 3.76 \times 10^{-4}u_2 + 2.4 \times 10^{-2}\ln(F) + 4.5 \times 10^{-3}RH \tag{5}$$

$$k_p = 0.85(\Delta + \gamma)/[\Delta + \gamma(1 + 0.33u_2)] \tag{6}$$

$$k_p = 0.51206 - 3.21 \times 10^{-4}u_2 + 31.886 \times 10^{-3}\ln(F) + 28.89 \times 10^{-4}RH \tag{7}$$

where RH is the mean daily relative humidity (%), u_2 is the mean daily wind speed at the height of 2 m above the ground (in km d^{-1} for Equations (3)–(5) and (7) and in m sec^{-1} for Equation (6)), and F is the upwind distance from the pan to the border of grass-covered land (m). A difference exists in Equation (6) in comparison to the other k_p models, because it is a function of u_2 and temperature T (T is included in the Δ parameter), and it is also a ratio without an intercept, where all coefficients are related to u_2 and T. An additional important k_p model is the one of Raghuwanshi and Wallender [18], but it is not used in this study, because it consists of categorical factors where their classes could not be adequately covered by the dataset of this study for allowing a proper recalibration of all its compartments.

On the other hand, Snyder et al. [25] proposed a sine wave equation for estimating ET_o using only E_{pan} adjusted for F = 100 m, which was further modified by Ghare et al. [60], who added an intercept as follows:

$$ET_o = 8.889\sin[(E_{pan}/19.2)\cdot(\pi/2)] + 0.6373 \tag{8}$$

Equation (8) is a quite flexible formula and provides an alternative relation between ET_o and E_{pan} compared to the commonly used Equation (2).

The five k_p models (Equations (3)–(7)) were transformed to their initial general form in order to analyze the robustness of their regression coefficients and, consequently, the effects of their associated climate parameters (e.g., RH, u_2, and T) through regression analysis. In Equations (3)–(7), the fetch distance was F = 1 m, which led to the following simplified general forms, respectively:

$$k_p = a + b\cdot u_2 + c\cdot RH + d\cdot RH^2 + e\cdot RH^2 u_2 \text{ (from Equation (3))} \tag{9}$$

$$k_p = a + b\cdot u_2 + c\cdot\ln(RH) \text{ (from Equation (4))} \tag{10}$$

$$k_p = a + b\cdot u_2 + c\cdot RH \text{ (from Equations (5) and (7))} \tag{11}$$

$$k_p = a(\Delta + \gamma)/[\Delta + \gamma(b + c\cdot u_2)] \text{ (from Equation (6))} \tag{12}$$

where a, b, c, d, and e are regression coefficients.

Similarly, the general form of Equation (8), with an intercept according to Ghare et al. [61] for estimating ET_o using only E_{pan}, was defined as:

$$ET_o = a\cdot\sin[(E_{pan}/b)\cdot(\pi/2)] + c \tag{13}$$

Finally, the simple regression tool of STATGRAPHICS Centurion XV software (StatPoint Technologies Inc.) was used in order to assess the optimum linear formula with transformed variables between ET_o and E_{pan}. The specific tool analyzes automatically 27 linear formulas with various transformations of the variables (Table S1 in Supplementary Materials), providing the option to the user to choose the best one based on performance metrics. For the dataset of this study, the best transformation for both variables (ET_o and E_{pan}) was found to be the following:

$$\ln(ET_o) = a + b\cdot\ln(E_{pan}) \rightarrow ET_o = \exp[a + b\cdot\ln(E_{pan})] \tag{14}$$

In this study, the second nonlinear form of Equation (14) was used in the comparative analysis of the models.

2.4. Methods of Analysis

2.4.1. Nonlinear Regression with Random Cross-Validation

The robustness of the regression coefficients of the six models (Equations (9)–(14)) was investigated using a random cross-validation nonlinear regression (RCV-NLR) analysis considering ET_o as the dependent variable (i.e., the four general nonlinear forms of k_p, Equations (9)–(12), were included in Equation (2) for estimating ET_o). The RCV-NLR analysis is based on a random splitting of the initial dataset into a calibration set (70% of the records) and a validation set (30% of the records). This random splitting was performed 1000 times, leading to a respective number of calibration and validation pairs of datasets. The RCV-NLR analysis was performed for each one of the six models based on the calibration sets, providing 1000 estimations of their respective regression coefficients. The estimated coefficients of each calibration set were used to validate each model based on the respective validation set. The RCV-NLR was built in R software using the "nls.lm" function of the {minpack.lm} package [64], which includes the Levenberg-Marquardt nonlinear least-squares algorithm.

The range of 1000 solutions of each regression coefficient of calibration and validation procedures was defined by the 95% confidence interval of the highest posterior density (HPD) distribution. The 2.5% and 97.5% thresholds (HPD thresholds), which contain the central 95% interval of the HPD distribution, were estimated in R software using the "p.interval" function of the {LaplacesDemon} package [65], which returns unimodal or multimodal HPD intervals, depending on the form of the probability distributions. Based on the RCV-NLR results, a model was considered robust according to the following rule: "a model is robust only when the 95% HPD intervals of all its regression coefficients preserve a constant sign (+ or -)" (i.e., a 95% HPD that contains positive and negative values of a regression coefficient indicates a nonrobust coefficient and, consequently, a nonrobust model).

2.4.2. Multiple Random Forests

Random forests (RF) [66] is a machine-learning method that is an extension and improvement of the Classification and Regression Trees (CART) method, also called decision trees. RF uses a modification of the bootstrap aggregating technique (bagging), by which a large collection of decorrelated, noisy, approximately unbiased trees are built and averaged in order to reduce the model variance and instability problems [67]. RF is an ensemble method where the aggregation of multiple trees improves the overall prediction accuracy, with both low bias and low variance results [66,68]. Some of the advantages of RF are the ability of modeling high-dimensional nonlinear relationships, resistance to overfitting, relative robustness, estimation of variable importance, and few user-defined parameters [66–69].

The hyperparameters of a RF model play a significant role in a model's performance, so the determination of their values is crucial. The hyperparameters that were considered were the number of the regression trees (num.trees), the number of candidate predictors that were randomly sampled (mtry), the proportion of train set that was used for building the model (sample.fraction), and the minimum number of points in the terminal nodes of the regression trees (min.node.size). A number of combinations of various values of hyperparameters was constructed, and by execution of the "ranger" package [70], their optimal set of values was determined (mtry = 3, num.trees = 1000, sample.fraction = 0.7, and min.node.size = 3). The validation set was 30% of the initial dataset, as in the case of RCV-NLR. RF also estimates the predictor variables' importance. The importance of a predictor variable is calculated as the mean—across all trees—decrease (%) of accuracy, expressed by the % change in Mean Squared Error - MSE (%), on the Out-of-Bag (OOB) sample when the variable is not taken into account by permuting its values randomly and maintaining the others as they were. Using constant optimal values of hyperparameters, the internal random procedures in RF lead every time to different solutions. For this reason, 1000 RF iterations (multiple random forests—MRF) were made in order to make comparisons with the NLR models. The MRF analysis was performed using all the predictor variables (E_{pan}, u_2, RH, and T) that participate in Equations (9)–(14). RF is not restricted by the limitations of a predefined nonlinear form and can be used as a benchmark model for evaluating

the predictive accuracy of NLR models (e.g., Equations (9)–(14)) but, also, for assessing the relative importance of the predictor climate variables if they were all included in one NLR model.

2.4.3. Models' Metrics

The metrics of R^2 and the root mean square error (RMSE) were estimated for each calibration and validation dataset of the RCV-NLR analysis (Equations (9)–(14)) and OOB and validation dataset of the MRF analysis, leading to 1000 respective estimations of their values for each model. Moreover, 1000 estimations of the slope and intercept of the trend line in the 1:1 plot of observed (ET$_o$-ASCE) vs. predicted (RCV-NLR or MRF) models only using the validation sets were also made as complementary metrics. The 1000 estimations of the metrics R^2, RMSE, slope, and intercept were also analyzed through the computation of HPD intervals.

3. Results

3.1. RCV-NLR Approach

The minimum, mean, maximum, and 2.5% and 97.5% HPD thresholds of the regression coefficients of Equations (9)–(14) derived by the RCV-NLR analysis during calibration are provided in Table S2 (Supplementary Materials). The HPD distributions of the regression coefficients based on their 1000 values are given in Figure 1.
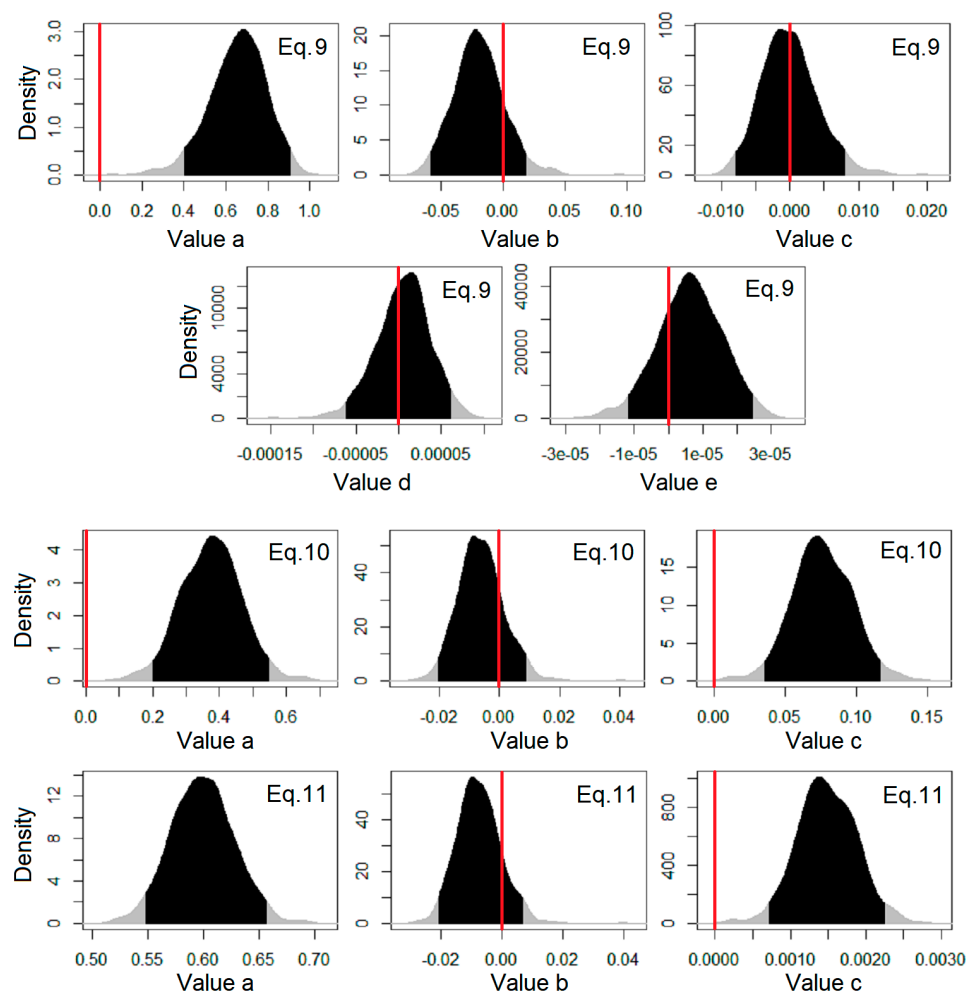
Figure 1. *Cont.*
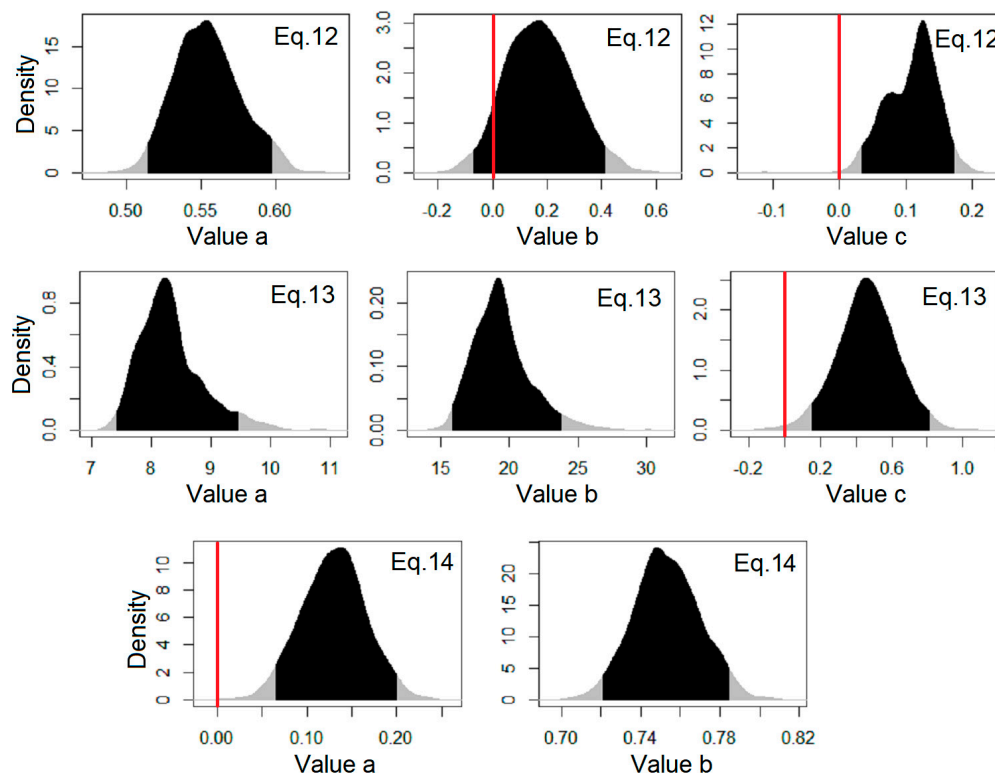
Figure 1. High posterior density (HPD) distributions of the 1000 values of each regression coefficient of Equations (9)–(14) during the random cross-validation nonlinear regression (RCV-NLR) calibration procedure (when the red line is in the black area, the coefficient and the respective model is not robust).

Considering Table S2 and Figure 1, the following were observed:

- For Equation (9), four out of five regression coefficients (b, c, d, and e coefficients) could not preserve a stable sign inside their 95% HPD interval. These coefficients are associated to RH and u_2 based on the specific k_p model.
- For both Equations (10) and (11), one out of three regression coefficients (b coefficient for both cases) could not preserve a stable sign inside its 95% HPD interval. This coefficient is associated to u_2 in both k_p models.
- For Equation (12), one out of three regression coefficients (b coefficient) could not preserve a stable sign inside its 95% HPD interval. Due to the complex form of Equation (12), this coefficient is associated to all climate parameters that participate in the specific k_p model.
- For Equations (13) and (14), all their coefficients presented a stable sign inside their 95% HPD intervals.

Considering the results of Table S2 and Figure 1, the rule of a constant sign in the 95% HPD interval of all regression coefficients is violated in all models except those that use only the E_{pan} parameter (Equations (13) and (14)). Based on this rule, these two models are expected to be more robust than the other four models (Equations (9)–(12)). This suggestion is verified by the results of R^2 and RMSE in the calibration and validation procedures (Table S3 and Figure 2). Figure 2 shows an alternative way to present the HPD distributions using modified box-whisker plots with boxes adjusted to the 95% HPD interval. These plots allow an easier visual comparison of the models' metrics. The results of Figure 2 show that the models of Equations (13) and (14) present higher R^2 and lower RMSE values than the other four models (Equations (9)–(12)) in both calibration (Figure 2a,c) and validation (Figure 2b,d) procedures. The results of Equations (13) and (14) are quite similar, with a slight better performance of Equation (13) (mean $R^2 = 0.872$ and mean RMSE = 0.429 mm d^{-1} during

calibration and mean $R^2 = 0.866$ and mean RMSE = 0.443 mm d^{-1} during validation), because it has three regression coefficients instead of two, as in the case of Equation (14).
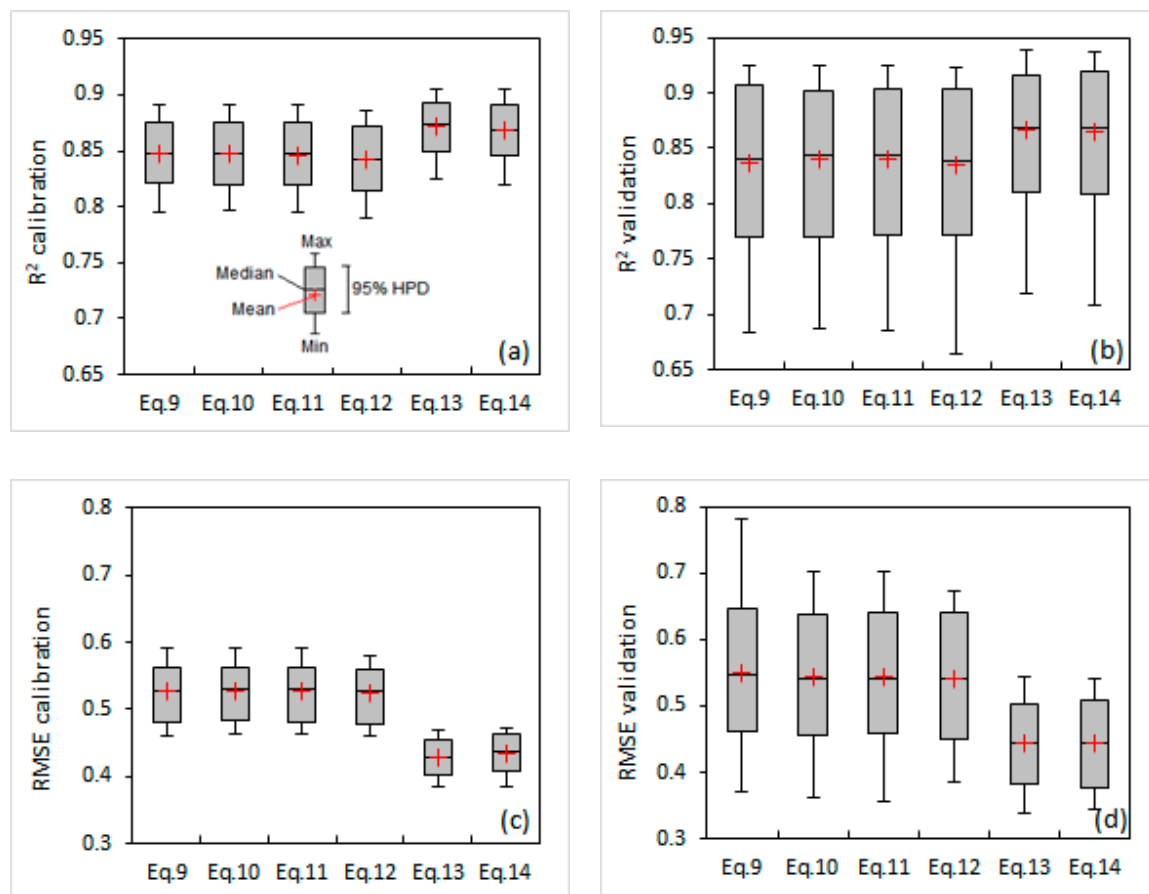

Figure 2. Range of (**a,b**) R^2 and (**c,d**) root mean square error (RMSE) during calibration and validation of the six models (Equations (9)–(14)) presented as modified box-whisker plots, with boxes adjusted to the 95% HPD interval.

3.2. MRF Approach

The minimum, mean, maximum, and 2.5% and 97.5% HPD thresholds of R^2 and RMSE during the OOB and validation procedures of the MRF approach are provided in the modified box-whisker plots of Figure 3a,b, respectively. Their respective values are provided in Table S4 of the Supplementary Materials. The performance of the MRF modeling approach that uses the full set of variables (i.e., E_{pan}, u_2, RH, and T) was generally better than the RCV-NLR models, according to the R^2 and RMSE metrics (mean $R^2 = 0.891$ and mean RMSE = 0.398 mm d^{-1} during the OOB procedure and mean $R^2 = 0.894$ and mean RMSE = 0.398 mm d^{-1} during validation of the MRF).

The respective modified box-whisker plots of the variables' importance are also provided in Figure 4 and their respective values in Table S5 of the Supplementary Materials. Based on Figure 4 and Table S5, the importance of the variables follows the order $E_{pan} > u_2 > RH > T$, with the MSE% values of E_{pan} being at least five times greater than the rest of the parameters.
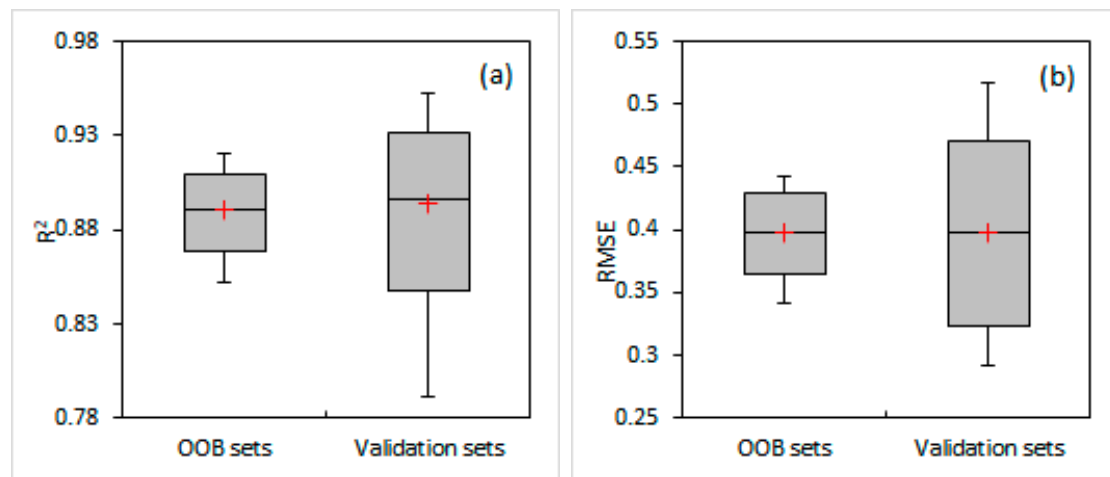
Figure 3. Range of (**a**) R^2 and (**b**) RMSE of the Out-of-Bag (OOB) and validation sets of the multiple random forests (MRF) analysis presented as modified box-whisker plots, with boxes adjusted to the 95% HPD interval.
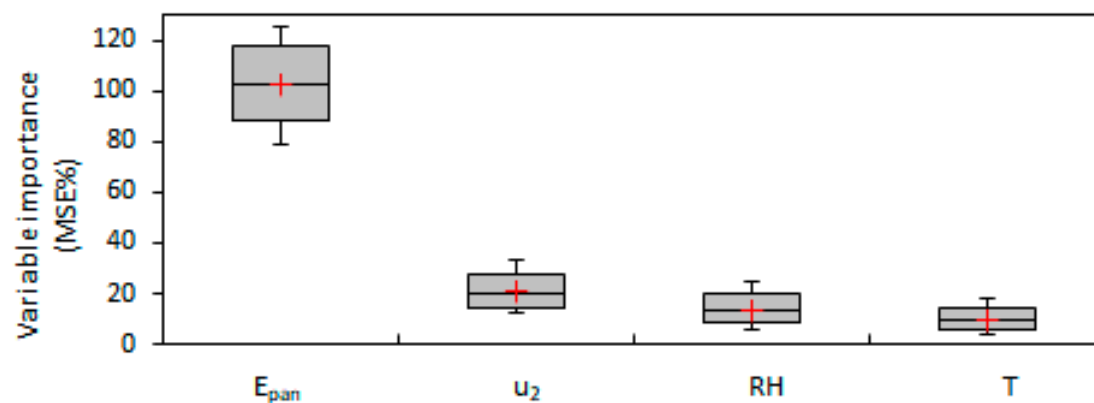

Figure 4. Range of the variables' importance in the MRF analysis presented as modified box-whisker plots, with boxes adjusted to the 95% HPD interval (Epan: daily pan evaporation (mm d^{-1}), u_2: daily wind speed at 2 m above ground surface (m s^{-1}), RH: daily relative humidity (%), T: daily temperature (°C).

3.3. Trend Line (Slope and Intercept) of Observed vs. Predicted ET$_o$ Values from 1:1 Plots Based on the Validation Datasets of All Models

The minimum, mean, maximum, and 2.5% and 97.5% HPD thresholds of the slopes and intercepts of the trend lines of the observed vs. predicted ET$_o$ values from 1:1 plots were estimated for all the validation cases from the RCV-NLR analysis of Equations (9)–(14) and MRF analysis (Figure 5a,b, Table S6) (in all cases, the validation data are the 30% of the initial dataset). The mean values and the 95% HPD intervals of the slopes and intercepts (Figure 5) showed that the ET$_o$ predictions of Equations (13) and (14) are better distributed along the perfect fit line of 45 degrees in comparison to the rest RCV-NLR models and the MRF model.
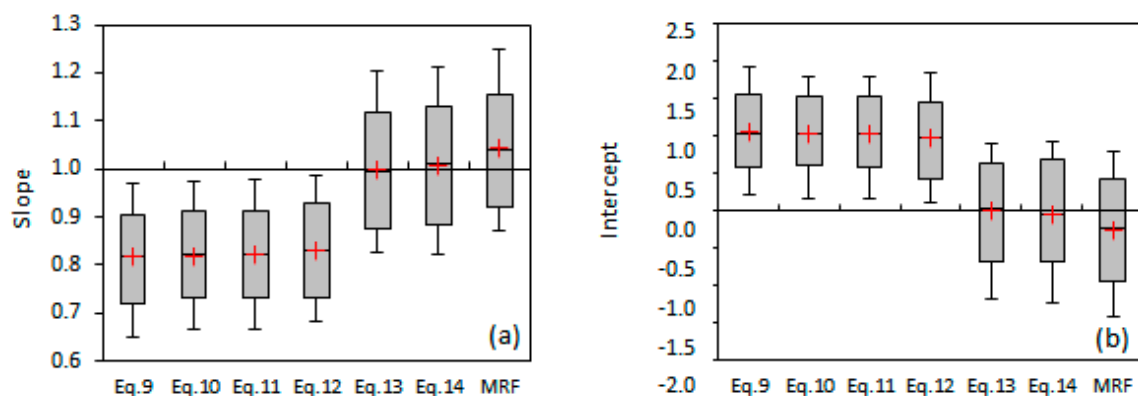
Figure 5. Range of (**a**) slopes and (**b**) intercepts of the trend lines of observed vs. predicted ET_o values from 1:1 plots using all the validation cases from the RCV-NLR analysis of Equations (9)–(14) and MRF presented as modified box-whisker plots, with boxes adjusted to the 95% HPD interval (optimum slope equal to 1, and optimum intercept equal to 0; 1:1 plots were constructed with observed values at the vertical and predicted values at the horizontal axis).

4. Discussion

4.1. Testing the Inclusion of Meteorological Parameters as Independent Variables in the E_{pan} Methodology

The results of this study showed that the decision to include additional meteorological parameters in the E_{pan} methodology for assessing ET_o or to use existing k_p models that already use meteorological parameters should be tested with more than one methodology before their use. The advantages and disadvantages of each methodology used in this study can be summarized as follows:

- The RCV-NLR method with various existing models has the advantage of analyzing the HPD interval of the coefficients, allowing to derive conclusions about the effects of their associated parameters (e.g., RH, u_2, and T). On the other hand, this method has the following disadvantages: (a) it is based on predefined nonlinear models that are formed by the user and may not be efficient enough to capture the responses of the dependent variable, and (b) in some cases, the form of the nonlinear models is complex, and the effect of the regression coefficient cannot be easily interpreted (e.g., coefficients of Equation (12)).
- The MRF method has the following advantages: (a) it is a machine-learning technique that is not restricted by predefined nonlinear forms by the user, (b) its predictions can be used as a measure of the predictive accuracy of typical NLR models, and (c) it provides the relative importance of the parameters. On the other hand, this method has the following disadvantages: (a) an RF model cannot be visualized as a function that is easily applicable by others, and (b) it is impossible to derive information that could explain whether an independent variable positively or negatively affects the dependent variable.

Considering the above, it is crucial to combine the two methodologies in order to derive an integrated conclusion that will allow the selection of the optimum NLR or RF model.

4.2. Effect of the Nonlinear Form on the Importance of Independent Variables

The results of this study showed that different forms of a nonlinear model affect differently the robustness of a regression coefficient and, consequently, the robustness of its associated independent variable. For example, the regression coefficient c in Equations (10) and (11) is associated to the RH and is robust according to the 95% HPD interval rule, while, in Equation (9), all the regression coefficients that are associated to the RH (i.e., the c, d, and e coefficients) are not robust (Figure 1). As regards u_2, its respective coefficient b is not robust in Equations (9)–(11) due to the positive and negative values in

the respective 95% HPD intervals. Considering the above, someone could agree that RH presents a more clear and robust effect on ET_o in the majority of NLR cases than u_2, according to the RCV-NLR analysis, but this suggestion is opposite to the results of the MRF analysis, since the importance of the u_2 variable is greater than the one of RH (Figure 4). These results highlight the limitations of NLR models to capture correctly the effect of an independent variable.

4.3. Is It Really Needed to Include Climate Parameters in E_{pan} Methodology for Estimating ET_o?

The variables' importance provided by MRF analysis and the small difference in statistical metrics (R^2 and RMSE) between the MRF and Equations (13) and (14) showed that simple NLR models that use only E_{pan} can reach a very satisfactory performance when they are calibrated for local conditions. Moreover, it was observed that the inclusion of additional climate parameters in NLR models may lead to underfitting, since Equations (9)–(12) showed poorer performance than Equations (13) and (14). The MRF analysis also showed the dominance of the E_{pan} variable over the other three parameters (u_2, RH, and T), even though the dependent variable (ET_o of the ASCE-standardized method) was estimated directly by them. This result can be easily justified by the fact that E_{pan} includes/combines all the effects of the rest variables, indicating that their inclusion may be redundant. Our opinion is that the general forms of Equations (13) and (14) or some of the alternative general equations provided in Table S1 can describe satisfactorily the relation between ET_o and E_{pan} when they are calibrated for local conditions. The mean values and 95% HPD intervals of the slopes and intercepts that were derived from the 1:1 plots of the observed vs. predicted ET_o values of the validation datasets (Figure 5) also showed that the ET_o predictions of Equations (13) and (14) are better distributed along the perfect fit line of 45 degrees in comparison to the MRF model and the rest of the RCV-NLR models. This indicates that MRF may be the most powerful model to minimize the overall error (in terms of RMSE and R^2), but it is not so robust at larger and smaller ET_o values, probably due to bias from the inclusion of the climate parameters of u_2, RH, and T.

4.4. Future Studies for Different Fetch and Different Climate Types

The precalibrated models of Equations (3)–(9) were given in this study in order to provide the source of different NLR forms, while the general forms that occurred (Equations (9)–(14)) were free from F, because they had to be recalibrated for a specific environment with constant fetch conditions. On the other hand, in the case where recalibration should be performed using data that include different fetch conditions, then the same analysis should be performed with the inclusion of F in the general forms of Equations (9)–(14). Such an analysis would be interesting, allowing also to assess the robustness of F parameter through analysis on its associated regression coefficients. The result of such an analysis regarding the robustness of the F parameter is unknown, and future studies should also focus on this issue.

The analysis of this study is based only on two years during the period May–September due to the temporary installation of a class A pan in the context of a parallel evapotranspiration experiment with lysimeters cropped, with flooded rice during the same period [71,72]. We believe that the selected dataset fulfills the requirements for supporting the results and conclusions of this study, despite the fact that it covers a small period and despite the fact that it describes the conditions of one climate type (semi-arid environment). Our opinion is based on the high variability of the daily records of the parameters (T, u_2, RH, E_{pan}, and ET_o) that were used in this study, which are provided in the form of histograms (Figure S1 in the Supplementary Materials). The large coverage and the distribution of these parameters is adequate for supporting robust conclusions about the effects of the T, u_2, and RH parameters in pan method models for the specific climate type, but future studies are also required to prove these findings for other climate types.

5. Conclusions

This study presented an integrated methodology for evaluating the robustness of regression coefficients associated to climate parameters (RH, u_2, and T) in pan method models during recalibration at local conditions. The results of this study showed that:

(a) the typical formula of Equation (2), with the use of a predefined form of a k_p model with locally calibrated coefficients, may not be the optimum solution for estimating ET_o. In the case of using such models, the calibration should be performed using as a dependent variable the ET_o and not the "measured" k_p (defined as the ratio ET_o/E_{pan}).

(b) the inclusion of climate parameters (e.g., u_2, RH, and T) in pan method models (NLR, MRF, etc.) for estimating ET_o may lead to underfitting and can be considered questionably from a statistical/mathematical point of view when ET_o is not measured but computed by a formula that is based on the same climate parameters.

(c) locally calibrated nonlinear regression functions for estimating ET_o, which use only the E_{pan} parameter, can have high predictive accuracy without the inclusion of additional climate parameters and can provide more balanced predictions along the perfect fit line of 45 degrees in 1:1 plots of observed vs. predicted ET_o.

The findings of the specific study should also be investigated for different fetch and different climate types in future studies.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Allen, R.G.; Pereira, L.S.; Raes, D.; Smith, M. *Crop Evapotranspiration: Guidelines for Computing Crop Water Requirements. Irrigation and Drainage Paper 56*; Food and Agriculture Organization of the United Nations: Rome, Italy, 1998.

2. Allen, R.G.; Walter, I.A.; Elliott, R.; Howell, T.; Itenfisu, D.; Jensen, M. *The ASCE Standardized Reference Evapotranspiration Equation*; Final Report (ASCE–EWRI); Task Committee on Standardization of Reference Evapotranspiration, Environmental and Water Resources Institute of the American Society of Civil Engineers: Reston, VA, USA, 2005.

3. Priestley, C.H.B.; Taylor, R.J. On the assessment of surface heat flux and evaporation using large–scale parameters. *Mon. Weather Rev.* **1972**, *100*, 81–92. [CrossRef]

4. Hargreaves, G.H.; Samani, Z.A. Estimating potential evapotranspiration. *J. Irrig. Drain Eng.* **1982**, *108*, 223–230.

5.　Itenfisu, D.; Elliott, R.L.; Allen, R.G.; Walter, I.A. Comparison of reference evapotranspiration calculations as part of the ASCE standardization effort. *J. Irrig. Drain Eng.* **2003**, *129*, 440–448. [CrossRef]

6.　Alexandris, S.; Kerkides, P.; Liakatas, A. Daily reference evapotranspiration estimates by the "Copais" approach. *Agric. Water Manage.* **2006**, *82*, 371–386. [CrossRef]

7.　Aschonitis, V.G.; Papamichail, D.; Demertzi, K.; Colombani, N.; Mastrocicco, M.; Ghirardini, A.; Castaldelli, G.; Fano, E.-A. High–resolution global grids of revised Priestley–Taylor and Hargreaves–Samani coefficients for assessing ASCE–standardized reference crop evapotranspiration and solar radiation. *Earth Syst. Sci. Data* **2017**, *9*, 615–638. [CrossRef]

8.　Wang, K.; Dickinson, R.E. A review of global terrestrial evapotranspiration: Observation, modeling, climatology, and climatic variability. *Rev. Geophys.* **2012**, *50*, 1–54. [CrossRef]

9.　McMahon, T.A.; Peel, M.C.; Lowe, L.; Srikanthan, R.; McVicar, T.R. Estimating actual, potential, reference crop and pan evaporation using standard meteorological data: A pragmatic synthesis. *Hydrol. Earth Syst. Sci.* **2013**, *17*, 1331–1363. [CrossRef]

10.　Valipour, M. Analysis of potential evapotranspiration using limited weather data. *Appl. Water Sci.* **2017**, *7*, 187–197. [CrossRef]

11.　Valiantzas, J.D. Temperature- and humidity-ased simplified Penman's ET0 formulae. Comparisons with temperature-based Hargreaves-Samani and other methodologies. *Agric. Water Manage.* **2018**, *208*, 326–334. [CrossRef]

12.　Frevert, D.K.; Hill, R.W.; Braaten, B.C. Estimation of FAO evapotranspiration coefficients. *J. Irrig. Drain. Eng.* **1983**, *109*, 265–270. [CrossRef]

13.　Cuenca, R.H. *Irrigation System Design: An Engineering Approach*; Prentice Hall: Englewood Cliffs, NJ, USA, 1989; p. 552.

14.　Allen, R.G.; Pruitt, W.O. FAO-24 reference evapotranspiration factors. *J. Irrig. Drain. Eng.* **1991**, *117*, 758–773. [CrossRef]

15.　Snyder, R.L. Equation for evaporation pan to evapotranspiration conversions. *J. Irrig. Drain. Eng.* **1992**, *118*, 977–980. [CrossRef]

16.　Pereira, A.R.; Villanova, N.; Pereira, A.S.; Baebieri, V.A. A model for the class—A pan coefficient. *Agric. For. Meteorol.* **1995**, *76*, 75–82. [CrossRef]

17.　Orang, M. *Potential Accuracy of the Popular Non-Linear Regression Equations for Estimating Pan Coefficient Values in the Original and FAO-24 Tables. Unpublished Report*; California Department of Water Resources: Sacramento, CA, USA, 1998.

18.　Raghuwanshi, N.S.; Wallender, W.W. Converting from pan evaporation to evapotranspiration. *J. Irrig. Drain. Eng.* **1998**, *124*, 275–277. [CrossRef]

19.　Grismer, M.E.; Orang, M.; Snyder, R.; Matyac, R. Pan evaporation to reference evapotranspiration conversion methods. *J. Irrig. Drain. Eng.* **2002**, *128*, 180–184. [CrossRef]

20.　Wahed, M.H.A.; Snyder, R.L. Simple equation to estimate reference evapotranspiration from evaporation pans surrounded by fallow soil. *J. Irrig. Drain. Eng.* **2008**, *134*, 425–429. [CrossRef]

21.　Tabari, H.; Mark, E.; Trajkovic, S. Comparative analysis of 31 reference evapotranspiration methods under humid conditions. *Irrig. Sci.* **2013**, *31*, 107–117. [CrossRef]

22.　Katul, G.G.; Cuenca, R.H.; Grebet, P.; Wright, J.L.; Pruitt, W.O. Analysis of evaporative flux data for various climates. *J. Irrig. Drain. Eng.* **1992**, *118*, 601–618. [CrossRef]

23.　Irmak, S.; Haman, D.Z.; Jones, J.W. Evaluation of Class A pan coefficients for estimating reference evapotranspiration in humid location. *J. Irrig. Drain. Eng.* **2002**, *128*, 153–159. [CrossRef]

24.　Sentelhas, P.C.; Folegatti, M.V. Class A pan coefficients (Kp) to estimate daily reference evapotranspiration (ETo). *Rev. Bras. Eng. Agric. Amb.* **2003**, *7*, 111–115. [CrossRef]

25.　Snyder, R.L.; Orang, M.; Matyac, S.; Grismer, M.E. Simplified estimation of reference. Evapotranspiration from pan evaporation data in california. *J. Irrig. Drain. Eng.* **2005**, *131*, 249–253. [CrossRef]

26.　Khoob, A.R.; Behbahani, S.; Nazarifar, M.H. Developing pan evaporation to grass reference evapotranspiration conversion model a case study in Khuzestan Province. *J. Appl. Sci.* **2007**, *7*, 936–941. [CrossRef]

27.　Gundekar, H.G.; Khodke, U.M.; Sarkar, S.; Rai, R.K. Evaluation of pan coefficient for reference crop evapotranspiration for semi-arid region. *Irrig. Sci.* **2008**, *26*, 169–175. [CrossRef]

28. Xing, Z.; Chow, L.; Meng, F.-R.; Rees, H.W.; Monteith, J.; Lionel, S. Testing reference evapotranspiration estimation methods using evaporation pan and modeling in maritime region of Canada. *J. Irrig. Drain. Eng.* **2008**, *134*, 417–424. [CrossRef]

29. Ghare, A.D.; Porey, P.D. Estimation of reference evapotranspiration of Nagpur region using simplified approach. In Proceedings of the 1st International Conference on Emerging Trends in Engineering and Technology, ICETET 2008, Nagpur, India, 16–18 July 2008; pp. 1022–1028.

30. Rahimikhoob, A. An evaluation of common pan coefficient equations to estimate reference evapotranspiration in a subtropical climate (north of Iran). *Irrig. Sci.* **2009**, *27*, 289–296. [CrossRef]

31. Sabziparvar, A.-A.; Tabari, H.; Aeini, A.; Ghafouri, M. Evaluation of class a pan coefficient models for estimation of reference crop evapotranspiration in cold semi-arid and warm arid climates. *Water Resour. Manag.* **2010**, *24*, 909–920. [CrossRef]

32. Esteves, B.S.; Mendonça, J.C.; Sousa, E.F.; Bernardo, S. Evaluation of "Class A" pan coefficients to estimate the referential evapotranspiration in Campos dos Goytacazes, RJ. *Rev. Brasil. Engenh. Agric. Amb.* **2010**, *14*, 274–278. [CrossRef]

33. Trajkovic, S.; Kolakovic, S. Comparison of simplified pan-based equations for estimating reference evapotranspiration. *J. Irrig. Drain. Eng.* **2010**, *136*, 137–140. [CrossRef]

34. Ditthakit, P.; Chinnarasri, C. Estimation of pan evaporation coefficient using neuro-genetic approach. *Am. J. Environ. Sci.* **2011**, *7*, 397–401. [CrossRef]

35. Cunha, A.R. Calculation of the Class—A pan coefficient in greenhouse and field by different methods [Coeficiente do tanque Classe A obtido por diferentes métodos em ambiente protegido e no campo]. *Semin. Cienc. Agrar.* **2011**, *32*, 451–464. [CrossRef]

36. Aschonitis, V.G.; Antonopoulos, V.Z.; Papamichail, D. Evaluation of pan coefficient equations in a semi-arid mediterranean environment using the ASCE—Standardized penman-monteith method. *Agric. Sci.* **2012**, *3*, 58–65. [CrossRef]

37. Mohammadi, M.; Ghahraman, B.; Davary, K.; Liaghat, A.M.; Bannayan, M. Pan coefficient (Kp) estimation under uncertainty on fetch. *Meteorol. Atmos. Phys.* **2012**, *117*, 73–83. [CrossRef]

38. Sabziparvar, A.A.; Shadmani, M. Evaluation of pan coefficients from ANN, ANFIS, and empirical methods, for estimation of daily reference evapotranspiration. *J. Earth Space Phys.* **2012**, *38*, 229–240.

39. Kaya, S.; Evren, S.; Dasci, E.; Bakir, H.; Adiguzel, M.C.; Yilmaz, H. Evaluation of pan coefficient for reference crop evapotranspiration for Igdir region of Turkey. *J. Food Agric. Environ.* **2012**, *10*, 987–991.

40. Duan, C.; Miao, Q.; Cao, W.; Wang, Y. Estimation of reference crop evapotranspiration by Chinese pan evaporation in Northwest China. *Trans. Chin. Soc. Agric. Eng.* **2012**, *28*, 94–99. [CrossRef]

41. Ditthakit, P.; Chinnarasri, C. Estimation of pan coefficient using M5 model tree. *Am. J. Environ. Sci.* **2012**, *8*, 95–103. [CrossRef]

42. Sahoo, B.; Walling, I.; Deka, B.C.; Bhatt, B.P. Standardization of reference evapotranspiration models for a subhumid valley rangeland in the Eastern Himalaya. *J. Irrig. Drain. Eng.* **2012**, *138*, 880–895. [CrossRef]

43. Pradhan, S.; Sehgal, V.K.; Das, D.K.; Bandyopadhyay, K.K.; Singh, R. Evaluation of pan coefficient methods for estimating FAO-56 reference crop evapotranspiration in a semi-arid environment. *J. Agrometeorol.* **2013**, *15*, 1–4.

44. da Cunha, P.C.R.; do Nascimento, J.L.; da Silveira, P.M.; Júnior, J.A. Efficiency of methods for calculating class A pan coefficients to estimate evapotranspiration reference [Eficiência de métodos para o cálculo de coeficientes do tanque classe A na estimativa da evapotranspiração de referência]. *Pesqui. Agropecu. Trop.* **2013**, *43*, 114–122. [CrossRef]

45. Rahimikhoob, A.; Asadi, M.; Mashal, M. A comparison between conventional and M5 Model Tree methods for converting pan evaporation to reference evapotranspiration for semi-arid region. *Water Resour. Manag.* **2013**, *27*, 4815–4826. [CrossRef]

46. Heydari, M.M.; Heydari, M. Evaluation of pan coefficient equations for estimating reference crop evapotranspiration in the arid region. *Arch. Agron. Soil. Sci.* **2014**, *60*, 715–731. [CrossRef]

47. Zhao, L.W.; Zhao, W.Z. Evapotranspiration of an oasis-desert transition zone in the middle stream of Heihe River, Northwest China. *J. Arid Land* **2014**, *6*, 529–539. [CrossRef]

48. Pereira, P.C.; Da Silva, T.G.F.; Silva, S.M.S.E.; Da Neto, J.F.C.; De Morais, J.E.F. Evaluation and applicability of the class A pan coefficient in the middle Pajeu, Pernambuco [Avaliação e aplicabilidade do coeficiente do tanque classe "a" no médio pajeú, pernambuco]. *Rev. Caatinga* **2014**, *27*, 131–140.

49. Bhartiya, K.M.; Ghare, A.D. Relative humidity based model for estimation of reference evapotranspiration for western plateau and hills region of India. *Water Resour. Manag.* **2014**, *28*, 3355–3364. [CrossRef]

50. Peixoto, T.D.C.; Levien, S.L.A.; Bezerra, A.H.F.; Sobrinho, J.E. Evaluation of different methodologies of class a pan ETo calculation in Mossoró, RN, Brazil [Avaliação de diferentes metodologias de estimativa da eto baseadas no tanque classe a, em mossoró, RN-1]. *Rev. Caatinga* **2014**, *27*, 58–65.

51. Bayatvarkeshi, M.; Abyaneh, H.Z. Validating pan coefficient equations to estimate reference evapotranspiration in comparing with lysimeter data of grass crop. *Glob. J. Adv. Pure Appl. Sci.* **2014**, *2*, 9–18.

52. Bogawski, P.; Bednorz, E. Comparison and validation of selected evapotranspiration models for conditions in Poland (Central Europe). *Water Resour. Manag.* **2014**, *28*, 5021–5038. [CrossRef]

53. Meshram, D.T.; Gorantiwar, S.D. Evaluation of pan coefficient for estimating reference crop evapotranspiration in solapur station, Maharashtra. *Mausam* **2015**, *66*, 205–210.

54. Andrade, A.D.; Miranda, W.L.; de Carvalho, L.G.; Figueiredo, P.H.F.; da Silva, T.B.S. Performance of methods for calculating the pan coefficient for estimating reference evapotranspiration [Desempenho de métodos de cálculo do coeficiente de tanque para estimativa da evapotranspiração de referência]. *IRRIGA* **2016**, *21*, 119–130. [CrossRef]

55. SreeMaheswari, C.H.; Jyothy, S.A. Evaluation of Class A pan coefficient models for estimation of reference evapotranspiration using penman-monteith method. *Int. J. Sci. Techn. Eng.* **2017**, *3*, 90–93.

56. Lennartz, F.; Kloss, S. Evaluating class a pan-based estimates of daily reference evapotranspiration with respect to irrigation scheduling on sandy soils in a hot arid environment. *J. Irrig. Drain. Eng.* **2018**, *144*, 1–9. [CrossRef]

57. Poddar, A.; Gupta, P.; Kumar, N.; Shankar, V.; Ojha, C.S.P. Evaluation of reference evapotranspiration methods and sensitivity analysis of climatic parameters for sub-humid sub-tropical locations in western Himalayas (India). *ISH J. Hydraul. Eng.* **2018**, in press. [CrossRef]

58. Khobragade, S.D.; Semwal, P.; Kumar, A.R.S.; Nainwal, H.C. Pan coefficients for estimating open-water surface evaporation for a humid tropical monsoon climate region in India. *J. Earth Syst. Sci.* **2019**, *128*, 1–14. [CrossRef]

59. Kumar, N.; Shankar, V.; Poddar, A. Investigating the effect of limited climatic data on evapotranspiration-based numerical modeling of soil moisture dynamics in the unsaturated root zone: A case study for potato crop. *Model. Earth Syst. Environ.* **2020**, in press. [CrossRef]

60. Ghare, A.D.; Porey, P.D.; Ingle, R.N. Discussion of "Simplified estimation of reference evapotranspiration from pan evaporation data in California" by Richard L. Snyder, Morteza Orang, Scott Matyac, and Mark E. Grismer. *J. Irrig. Drain. Eng.* **2006**, *132*, 519–520. [CrossRef]

61. Aschonitis, V.G.; Lekakis, E.; Tziachris, P.; Doulgeris, C.; Papadopoulos, F.; Papadopoulos, A.; Papamichail, D. A ranking system for comparing models' performance combining multiple statistical criteria and scenarios: The case of reference evapotranspiration models. *Environ. Modell. Softw.* **2019**, *114*, 98–111. [CrossRef]

62. Clark, C. Measurements of actual and pan evaporation in the upper Brue catchment UK: The first 25 years. *Weather* **2013**, *68*, 200–208. [CrossRef]

63. Doorenbos, J.; Pruitt, W.O. *Guidelines for Predicting Crop Water Requirements. Irrigation and Drainage Paper No. 24*, 2nd ed.; Food and Agriculture Organization of the United Nations: Rome, Italy, 1977.

64. Elzhov, T.V.; Mullen, K.M.; Spiess, A.N.; Bolker, B. Package Minpack.lm: R Interface to the Levenberg-Marquardt Nonlinear Least-Squares Algorithm Found in Minpack, Plus Support for Bounds. 2016. Available online: https://cran.r-project.org/web/packages/minpack.lm/minpack.lm.pdf (accessed on 1 May 2020).

65. Bernardo, J.M. Intrinsic credible regions: An objective Bayesian approach to interval estimation. *Soc. Estad. Investig. Operat.* **2005**, *14*, 317–384. [CrossRef]

66. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [CrossRef]

67. Hastie, T.; Tibshirani, R.; Friedman, J. *The Elements of Statistical Learning*, 2nd ed.; Springer: New York, NY, USA, 2009.

68. Díaz-Uriarte, R.; De Andres, S.A. Gene selection and classification of microarray data using random forest. *BMC Bioinform.* **2006**, *7*, 1–13. [CrossRef]

69. Strobl, C.; Malley, J.; Tutz, G. An introduction to recursive partitioning: Rationale, application, and characteristics of classification and regression trees, bagging, and random forests. *Psychol. Methods* **2009**, *14*, 323–348. [CrossRef] [PubMed]

70. Wright, M.N.; Wager, S.; Probst, P. Package 'Ranger': A Fast Implementation of Random Forests. 2020. Available online: https://cran.r-project.org/web/packages/ranger/ranger.pdf (accessed on 1 May 2020).

71. Aschonitis, V.G.; Papamichail, D.M.; Lithourgidis, A.; Fano, E.A. Estimation of leaf area index and foliage area index of rice using an indirect gravimetric method. *Commun. Soil Sci. Plant Analys.* **2014**, *45*, 1726–1740. [CrossRef]

72. Aschonitis, V.; Diamantopoulou, M.; Papamichail, D. Modeling plant density and ponding water effects on flooded rice evapotranspiration and crop coefficients: Critical discussion about the concepts used in current methods. *Theor. Appl. Climatol.* **2018**, *132*, 1165–1186. [CrossRef]