Article

Artificial Neural Network Model of Soil Heat Flux over Multiple Land Covers in South America

Bruno César Comini de Andrade [1,*], Olavo Correa Pedrollo [1], Anderson Ruhoff [1], Adriana Aparecida Moreira [1], Leonardo Laipelt [1], Rafael Bloedow Kayser [1], Marcelo Sacardi Biudes [2], Carlos Antonio Costa dos Santos [3], Debora Regina Roberti [4], Nadja Gomes Machado [5], Higo Jose Dalmagro [6], Antonio Celso Dantas Antonino [7], José Romualdo de Sousa Lima [8], Eduardo Soares de Souza [9] and Rodolfo Souza [10]

1   Institute of Hydraulic Research, Universidade Federal do Rio Grande do Sul, Porto Alegre 91501-970, Brazil; olavo.pedrollo@ufrgs.br (O.C.P.); anderson.ruhoff@ufrgs.br (A.R.); adriana.moreira@ufrgs.br (A.A.M.); leonardo.laipelt@ufrgs.br (L.L.); rafael.kayser@ufrgs.br (R.B.K.)
2   Physics Institute, Universidade Federal do Mato Grosso, Cuiabá 78060-900, Brazil; marcelo@fisica.ufmt.br
3   Department of Atmospheric Sciences, Universidade Federal de Campina Grande, Campina Grande 58429-900, Brazil; carlos.santos@ufcg.edu.br
4   Physics Institute, Universidade Federal de Santa Maria, Santa Maria 97105-900, Brazil; debora@ufsm.br
5   Instituto Federal de Mato Grosso, Cuiabá 78060-900, Brazil; nadja.machado@blv.ifmt.edu.br
6   Postgraduate Program in Environmental Science, Universidade de Cuiabá, Cuiabá 78005-300, Brazil; higo.dalmagro@kroton.com.br
7   Department of Nuclear Energy, Universidade Federal de Pernambuco, Recife 50740-540, Brazil; antonio.antonino@ufpe.br
8   Academic Unit of Garanhuns, Federal University of Agreste of Pernambuco, Garanhuns 55292-278, Brazil; romualdo.lima@ufape.edu.br
9   Academic Unit of Serra Talhada, Universidade Federal Rural de Pernambuco, Serra Talhada 56900-000, Brazil; eduardo.ssouza@ufrpe.br
10 Department of Biological & Agricultural Engineering, Texas A&M University, College Station, TX 77843, USA; Rodolfo.souza@tamu.edu
*   Correspondence: bruno.comini@ufrgs.br

Abstract: Soil heat flux (G) is an important component for the closure of the surface energy balance (SEB) and the estimation of evapotranspiration (ET) by remote sensing algorithms. Over the last decades, efforts have been focused on parameterizing empirical models for G prediction, based on biophysical parameters estimated by remote sensing. However, due to the existing models' empirical nature and the restricted conditions in which they were developed, using these models in large-scale applications may lead to significant errors. Thus, the objective of this study was to assess the ability of the artificial neural network (ANN) to predict mid-morning G using extensive remote sensing and meteorological reanalysis data over a broad range of climates and land covers in South America. Surface temperature (T_s), albedo (α), and enhanced vegetation index (EVI), obtained from a moderate resolution imaging spectroradiometer (MODIS), and net radiation (R_n) from the global land data assimilation system 2.1 (GLDAS 2.1) product, were used as inputs. The ANN's predictions were validated against measurements obtained by 23 flux towers over multiple land cover types in South America, and their performance was compared to that of existing and commonly used models. The Jackson et al. (1987) and Bastiaanssen (1995) G prediction models were calibrated using the flux tower data for quadratic errors minimization. The ANN outperformed existing models, with mean absolute error (MAE) reductions of 43% and 36%, respectively. Additionally, the inclusion of land cover information as an input in the ANN reduced MAE by 22%. This study indicates that the ANN's structure is more suited for large-scale G prediction than existing models, which can potentially refine SEB fluxes and ET estimates in South America.

Keywords: remote sensing; surface energy balance; deep learning; MODIS; GLDAS 2.1; flux towers

1. Introduction

Soil heat flux (G) is one of the main components of the surface energy balance (SEB) and accounts for the energy transferred to and from the land surface and deeper layers of the ground [1]. The SEB determines the energy and mass exchanges of the soil-plant-atmosphere continuum, surface temperature (T_s), the amount of energy available for evaporation and, ultimately, SEB defines local climate conditions and water availability [2]. Most energy that is transferred into the soil during sunlight hours is transferred back at night. Thus, G in vegetated surfaces is generally the smallest of the SEB fluxes and has been neglected in daily considerations of the SEB with some success [3,4]. However, significant energy transfers occur both during the day and night, and G's exclusion from instantaneous SEB estimations can lead to substantial errors, e.g., mid-morning G (usually between 9 am and 12 pm) can be as large as sensible heat flux (H) for some land covers [5,6].

Several remote sensing products provide valuable knowledge at different spatiotemporal scales to assess land surface and meteorological conditions. Motivated by this availability, many models were developed to calculate SEB fluxes from the integration of remote sensing data, mainly focused on actual evapotranspiration (ET) estimation [7–18]. Among them, the SEBAL (surface energy balance algorithm for land) is one of the most used models around the world [11,12,19–23], with relatively accurate predictions of ET [24–26].

However, the empirical nature of the developed G prediction models usually implicates that their application to conditions other than the ones for which they were developed (e.g., land cover, climate) may result in large uncertainties [27], suggesting a need for local calibration for accurate estimations [28,29]. Moreover, errors in G can profoundly decrease the accuracy of SEB models [1,23,29–32]. To address this limitation, a local calibration can be applied [21,33], but overall this may be impractical for operational purposes, such as SEB mapping over large areas. Comparative studies revealed a wide range in models' performance, especially across different land covers [27,33], and seasonal and regional biases have been identified, indicating potential issues with model comparisons at the local scale [27].

To overcome the issue of regional biases when performing a local scale model calibration/validation, we built a network of flux towers from regional initiatives to represent the wide-ranging climate and ecosystem diversity of South America [34]. In addition to the most used flux tower network data compilation [35] from the large-scale biosphere-atmosphere experiment in the Amazon (LBA-ECO) [36], we also included data from the SULFLUX (South Brazilian network of surface fluxes and climate change), ONDACBC (National Observatory of Water and Carbon Dynamics in the Caatinga Biome) [37], and the PELD Pantanal (long-term ecological research in Pantanal) [38], in conjunction with flux measurements supported by Brazilian universities, including the UFMT (Federal University of Mato Grosso) [39] and the USP (University of Sao Paulo) [40], funded by national and regional research agencies. This network encompasses most land covers in South America, corresponding to over 80% of its area (except for urban, barren, and high altitude forest land covers) [41], and has been used as validation data for SEB flux estimations via remote sensing [23,32,33].

An alternative to the existing methods of modeling G is the use of artificial neural networks (ANNs), which are universal approximators [42] and have been successfully tested (used or applied) in several hydrological processes and modeling attempts and have shown better performance than existing conceptual and empirical models [42–48]. ANNs are computational models analogous to the brain's biological behavior, simulating its capabilities of learning and memorizing. The ANN relates input data to specified outputs through a series of intertwined layers of neurons. The neurons connect and transform input variables via activation functions. An ANN can be trained by pairing historical data and calibrating the connections' synaptic weights to obtain the best relationship between inputs and outputs [42].

One of the main disadvantages of ANNs is that their structure does not describe the underlying processes that explain the target variable, thus being labeled as black-

box models [49]. On the other hand, the empirical origin of the existing models for instantaneous G prediction also presents the same limitation. With this in mind, ANNs become an attractive methodology for G prediction. In fact, previous applications of ANNs for G prediction at the local scale [46,47] have yielded promising results. Considering the availability of the long time series of measured G, representative of the South American landscape [34,41], and the readily available remote sensing data, the usage of the ANN approach for large-scale applications is encouraged as a next step to assess its efficiency in predicting G.

Therefore, given the diversity of G estimation methods and their limited application conditions, this paper focuses on assessing the ability of ANNs to predict G over a wide range of ecosystems in South America, using long-term remote sensing and meteorological time series, and comparing the ANN's performance to that of commonly used G models. With this study, we hope to assist future SEB closure studies and ET estimates with more accurate estimates of G and consequently H and LE.

2. Materials and Methods

2.1. Study Sites

In this study, data from 23 flux towers were used to calibrate and validate the ANNs, located in South America. The area surrounding each flux tower was considered one study site. Table 1 lists the flux towers surrounding land covers and the data acquisition periods, while Figure 1 displays the location and land cover of the study sites. For simplification and repeatability, the land covers were grouped into one of five types: forest, irrigated cropland, cropland, grassland, and savanna.

Table 1. Description of the flux towers used in this study.

Name	Land Cover	Main Land Cover	Available Data Period	Source
AFL	Rainforest	Forest	February 2003–March 2004	[39]
BAN	Seasonally flooded forest	Forest	October 2003–December 2006	[50]
BDP	Seasonally flooded forest	Forest	June 2011–August 2015	[39]
BRA	Cerrado	Savanna	January 2011–December 2011	[51]
CAL	Cropland (soy)	Cropland	January 2009–September 2014	[52,53]
CAM	Seasonally flooded forest	Forest	January 2007–January 2009	[38]
CDS	Cropland (flooded rice)	Irrigated cropland	January 2011–December 2014	[54,55]
CMT	Irrigated cropland	Irrigated cropland	September 2015–February 2017	[56]
CSJ	Caatinga	Savanna	April 2019–December 2019	[57]
CST	Caatinga	Savanna	February 2014–March 2020	[55,58]
FEX	Pasture	Grassland	September 2006–March 2010	[39]
FMI	Cerrado	Savanna	April 2009–May2013	[39]
FSN	Pasture	Grassland	March 2002–July 2003	[59]
M34	Tropical forest	Forest	January 2000–September 2006	[60]
NPW	Seasonally flooded forest	Forest	January 2015–July 2017	[61]
PAL	Grassland	Grassland	September 2013–September 2016	[62]
PAN	Cerrado	Savanna	January 2001–December 2006	[63]
PDG	Cerrado	Savanna	January 2001–December 2006	[64]
PDS	Cropland (flooded rice)	Irrigated cropland	October 2003–March 2004	[65,66]
SIN	Transitional forest (cerradão)	Forest	May2005–October 2008	[39]
S77	Pasture/agriculture	Grassland	August 2000–November 2005	[67]
S83	Tropical forest	Forest	June 2000–March 2004	[68]
SMA	Grassland	Grassland	January 2013–December 2016	[62]
USE	Cropland (sugarcane)	Cropland	January 2001–December 2002	[69]
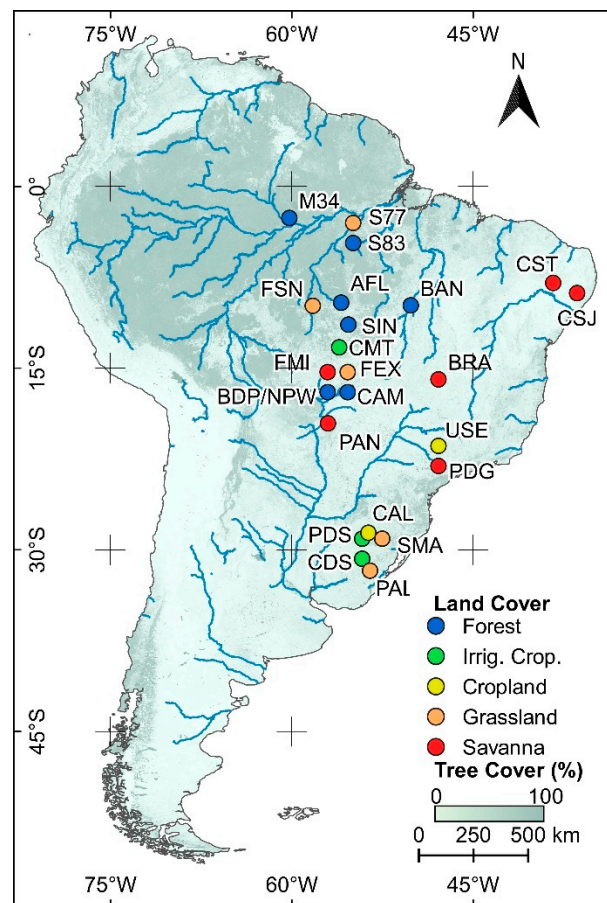
Figure 1. Flux tower locations and main land cover types used in this research.

2.2. Input Data

Table 2 shows the input data used in the development of the ANN. Surface temperature (T_s), surface albedo (α), and enhanced vegetation index (EVI) data were obtained from a moderate resolution imaging spectroradiometer (MODIS) product at daily time intervals. Given that the training process of ANN requires a large amount of data, the MODIS products were selected due to their high temporal resolution. MODIS's image acquisition (on-board of the Terra satellite) occurs at around 10:30 AM (local time). The net radiation (R_n) dataset was acquired from the global land data assimilation system version 2 (GLDAS 2.1) reanalysis product at a 3 h time intervals, averaged for the 9 AM to 12 PM window. Both satellite and reanalysis products were available from the Google Earth Engine platform (https://earthengine.google.com/ [Accessed date: 30 June 2020]).

Table 2. Input data description.

Data	Product	Spatiotemporal Resolution	Source
Net radiation (Rn)	GLDAS 2.1	0.25°/3 h	[70]
Land surface temperature (Ts)	MOD11A2.V6	1000 m/1 day	[71]
Albedo (α)	MCD43A3.V6	500 m/1 day	[72]
Enhanced vegetation index (EVI)	MOD13Q1.V6	250 m/16 days	[73]

Flux tower data were obtained at hourly and sub-hourly time intervals and were also resampled and averaged to the 9 AM to 12 PM local time window. A search was executed on each flux tower dataset to remove incomplete (gap-filling) data. R_n values 10% lower than the five days moving average were discarded to remove cloudy sky data. Finally, G

values beyond the range of -50% to 50% of R_n were also removed from the analysis [74–76]. After this filtering procedure, a total of 3778 samples remained available.

Figure 2 displays the correlation plot between the observed G from the flux towers and the R_n, T_s, EVI, and α datasets. The correlation plots and r values are distinguished by land cover. Correlation among the input datasets is generally weak, with $|r| < 0.50$ for most land covers, where $|r|$ is the absolute value of r. Exceptions ($|r| > 0.50$) are mainly found for T_s and R_n on most land covers. The correlation of observed G to input datasets is also low in general. The correlation plots are useful to identify potentially redundant inputs of the ANN. Due the low correlation among the variables, it can be assumed that they may contribute independently to the ANN predictions.
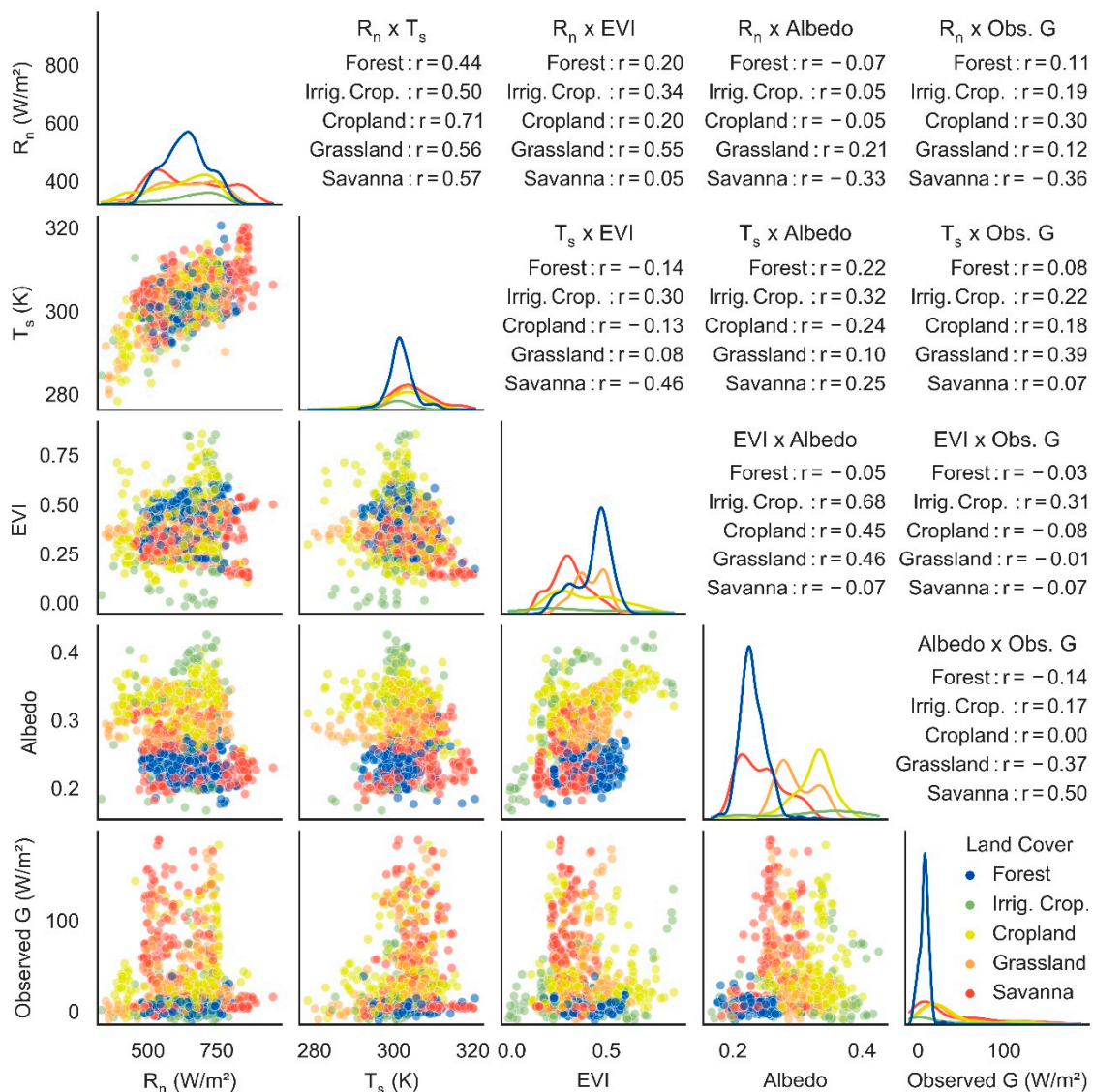

Figure 2. Correlation plot and r values of the observed G and the input datasets for all flux towers, distinguished by land cover.

2.3. ANN's Structure

Figure 3 displays a generic scheme of the ANNs that were built. The input data are combined into a series of processing units in the hidden layer, called neurons. These "hidden" neurons are then connected to an output neuron that converts the hidden layer's information into the desired output. In order to transform the data, each neuron passes the received information through an activation function, such as the bipolar and unipolar sigmoid functions. The connections between neurons are established by synaptic weights

(or simply weights), and these, together with the activation functions, determine the relationship between the input data and the corresponding outputs. The weight values are defined during the training process. The number of neurons in the hidden layer (m) determines the ANN complexity and is assessed based on a search process for a network with the least complexity that still obtains the same performance, with a validation sample, as an initial network, that is purposefully over-sized [77]. The ANNs were developed with the most common inputs used in models for G computation [43–47]. This allows for the assessment of the benefits of using ANNs compared to preexisting models.

The study sites were grouped into different land cover types to consider characteristics that are difficult to capture from remote sensing data, such as subcanopy vegetation density, vegetation height, and soil skin temperature. Land cover data were inputted as numerical data in the ANNs, as follows: 1 for cropland; 2 for irrigated cropland; 3 for forest; 4 for grassland; and 5 for savanna. The impact of the inclusion of land cover data as an input was assessed using a comparison of the ANNs with and without the land cover information. Thus, two networks were generated. From this point on, the two networks will be referred to as: (i) Gc: an ANN that considers land cover as an input; and (ii) Gnc: an ANN that does not consider land cover as an input.
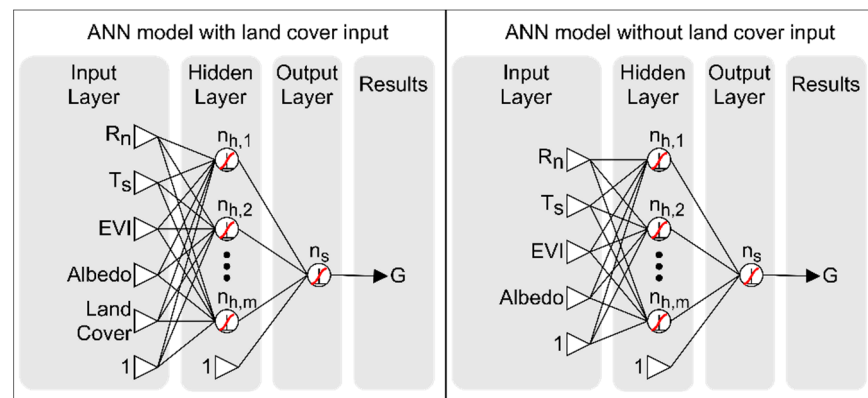
Figure 3. Structure of the two ANNs used for this study. One includes the site's land cover as an input (Gc) and another is without land cover information (Gnc).

In order to avoid preferential treatment of inputs with different magnitudes, input and output data were scaled using a normalization procedure, shown by Equation (1):

$$X' = \frac{(X - X_{bottom})}{X_{top} - X_{bottom}} \tag{1}$$

where X' is the scaled variable; X is the raw variable; and X_{bottom} and X_{top} are the data bottom and top limits, respectively. The top and bottom limits were chosen based on observations of the 23 study sites' available series and are shown in Table 3. These values also represent the valid range for applications of the trained ANNs.

Table 3. Data limits used for the scaling procedures presented in Equation (1).

Data	Bottom Limit	Top Limit
Soil heat flux (G)	-20.0 W/m^2	200.0 W/m^2
Net radiation (R$_n$)	150.0 W/m^2	1000.0 W/m^2
Land surface temperature (T$_s$)	275.0 K	325.0 K
Albedo (α)	0.1	0.5
Enhanced vegetation index (EVI)	-0.1	1.0
Land cover	0.0	6.0

The developed ANNs are generically represented by Equation (2):

$$O' = f_o \left\{ \sum_{i=1}^{m} w_{o,i} f_h \left[\sum_{j=1}^{n} \left(w_{h,i,j} X'_j \right) + b_{h,i} \right] + b_o \right\} \quad (2)$$

where O' is the ANN's output, equivalent to G in its scaled form; n is the number of input datasets; m is the number of neurons in the hidden layer; w_h, b_h, f_h, w_o, b_o, and f_o are the synaptic weights (w), biases (b), and activation functions (f) of the hidden (h) and output (o) layers, respectively.

The unipolar sigmoid function was used as the activation function for both the hidden and output layers.

2.4. ANN's Training Process

The ANN's training consisted of adjusting of the synaptic weights and biases by minimization of the output error using the backpropagation algorithm [78]. w and b values were updated using the method presented in [79] and the learning rate was calculated according to the training acceleration techniques proposed by [80].

The cross-validation method was used to control overfitting. This method consists of dividing the datasets into three series of similar sizes: a training series, a validation series, and a verification series. Weight and bias updates occur to reduce the total quadratic error in the training dataset (S_t) with each iteration, but the quadratic error in the validation series (S_v) is also observed. The final values of w and b are the ones that provide the minimal S_v value. The ANN performance is then assessed for the verification series.

The number of iterations (or cycles) must be sufficient to identify the best iteration (minimal error). If the best cycle is too close to the end of the network training, it is possible, and probable, that a better solution could be found if there were more iterations. Thus, a method for increasing the total number of cycles was applied to the ANN training, as follows:

A starting minimum of 200,000 cycles was defined;

If the best iteration occurred before the 100,000th cycle, the training was considered complete. If it happened after, the total number of cycles was doubled, up to 400,000 cycles.

This second step was repeated either until the best cycle was found in the first half of the training process or until a limit of five duplications was reached, reaching up to 6.4 million iterations. This limit was applied to reduce computational running time and to avoid infinite loops.

In order to assess the ANN's complexity, the training process was repeated with the progressive reduction in the number of neurons in the hidden layer, from a maximum of 12 neurons, until a decrease in its generalization capacity was observed, measured by the quadratic error in the validation series. The smallest number of neurons that did not yield a decrease in the network's generalization capacity was then chosen.

After assessing the ANN's complexity, the data series division was evaluated, as recommended by [81]. One-third of the total data was reserved for the verification series. The remaining data were split into calibration and validation series. The tested split ratio range was larger than that recommended by [81], because it was noticed that better predictions of G happened for smaller training datasets. Therefore, a broader range of split ratios needed to be tested. A sequence of tests between 10/90 and 90/10 was executed to determine the optimal calibration/validation split ratio. The quadratic error present in the validation series was used to determine the optimal data division.

Due to the ANN's empirical nature and the randomness of the initial conditions, training runs could prematurely stop at a local minimum on the objective function surface. To avoid this, after each randomization and partitioning of the data series, the cross-validation process was repeated ten times. Since data randomization was performed 40 times, ANN training was repeated in parallel 400 times for each of the two networks

created. The trained ANNs with the best performance in the validation series were then used to predict G in the verification data series.

2.5. Performance Assessment and Input Data Contribution Analysis

The ANN's predictions were assessed against data measured in the flux towers via a set of statistical indexes, calculated for the validation data series of each of the 40 data randomizations. The Nash–Sutcliffe coefficient (NS) [82] value in the validation dataset was used as the criterion for selecting the optimal ANN.

For comparison purposes, G was also estimated by the original and adjusted Jackson et al. (1987) [83] and Bastiaanssen (1995) [84] equations. Reparameterization was performed using the minimum quadratic error as the goal function. Table 4 presents the originals and adjusted equations (for results, see Section 3.2. Models' performance evaluation), where T_s is the surface temperature (°C); α is the surface albedo; and NDVI is the normalized difference vegetation index.

Table 4. Original Jackson et al. (1987) and Bastiaanssen (1995) equations and adjusted equations calibrated for minimum quadratic error.

Original Equations	Adjusted Equations
Jackson et al. (1987):	Jackson et al. (1987):
$G = 0.583 R_n \exp(-2.13 \cdot NDVI)$	$G = 0.150 R_n \exp(-0.89 \cdot NDVI) - 34.3$
Bastiaanssen (1995):	Bastiaanssen (1995):
$G = T_s R_n (0.0038 + 0.0074\alpha)\left(1 - 0.98 NDVI^4\right)$	$G = T_s R_n (-0.0019 + 0.0106\alpha)\left(1 - 1.10 NDVI^4\right) + 10.4$

The performances of the Gc and Gnc ANNs, and of the Jackson et al. (1987) and Bastiaanssen (1995) models, were evaluated via the comparison of NS, mean absolute error (MAE), root-mean-square error (RMSE) [85], bias [86], the slope coefficient (A_0) [87], and the linear correlation coefficient (r) [88], calculated for the verification series. The MAE, RMSE, and bias metrics were calculated as absolute values and as a percentage of the mean of the observed verification data series.

The input data sets contribution to the ANNs was calculated by the connection weights method proposed by [89] and modified by [90]. In this method the relative importance of the input variable is entirely derived from the synaptic weights of the ANN post-training.

3. Results

3.1. ANN Training Process

The complexity analysis of the ANNs indicated that its performance improves initially with higher complexity but is stable for complexities greater than four neurons. Therefore, for security reasons, seven neurons for Gc and five neurons for Gnc were chosen.

After assessing the ANN's complexity, the optimal dataset training/validation split ratio was verified, resulting in 748 samples for the training series, 1745 samples for the validation series, and 1285 samples for the verification series, for both the Gc and the Gnc ANNs. The synaptic weights of the trained ANNs, as well as the relative importance of each input variable to each ANN, are available as Supplementary Information (Tables S1–S3).

3.2. Model Performance Evaluation

Performance metrics of the Gc and Gnc ANNs, and the Jackson et al. (1987) and Bastiaanssen (1995) original and adjusted models, are displayed in Table 5. The adjusted versions of the Jackson et al. (1987) and Bastiaanssen (1995) models yielded significantly better performance metrics in relation to the original versions. Given this difference in performance, the original versions were not further assessed. The ANNs performed better than the adjusted models, with lower values of MAE and RMSE, and higher NS, A_0, and r values. The Gc ANN yielded the best NS, MAE, RMSE, A_0, and r values.

Table 5. Performance metrics of the Gc and Gnc ANNs, and of the Jackson et al. (1987) and Bastiaanssen (1995) original and adjusted models.

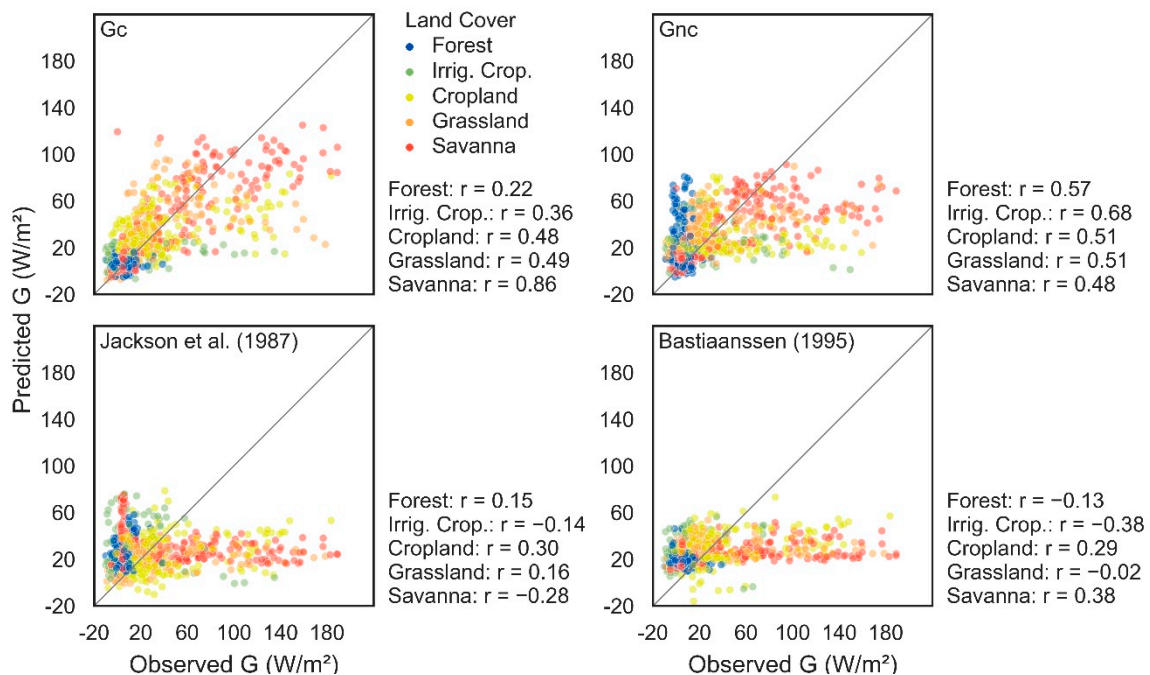| Metric	Gc	Gnc	Jackson et al. (1987)		Bastiaanssen (1995)	
			Original	Adjusted	Original	Adjusted
NS	0.53	0.27	−23.11	−0.17	−7.16	0.02
MAE (W/m^2)	14.04	18.06	145.66	24.43	84.69	21.78
MAE (%)	56%	70%	603%	101%	351%	90%
RMSE (W/m^2)	23.95	28.90	159.15	35.00	92.61	32.00
RMSE (%)	95%	113%	659%	145%	384%	133%
Bias (W/m^2)	−1.87	−0.82	145.36	1.07	81.73	0.99
Bias (%)	−7%	−3%	602%	4%	339%	4%
A_0	0.52	0.30	0.08	0.01	0.12	0.08
r	0.73	0.53	0.04	0.02	0.12	0.21

3.3. Model Performance by Land Cover

Figure 4 compares the seasonal G in each land cover, given by the flux tower observations and by the Gc and Gnc ANNs, and the Jackson et al. (1987) and Bastiaanssen (1995) adjusted models. While forest land cover displays a relatively constant observed G throughout the year, other land covers present a broader seasonal range and larger monthly standard deviations (given by shaded areas). Overall, the Gc ANN yielded the best adherence to the observed G series, showing similar seasonal patterns and standard deviations, especially in the savanna land cover. On the other hand, the Gnc ANN and the adjusted models yielded a smaller seasonal range and standard deviation, not adhering to the observed G series as well as the Gc ANN.
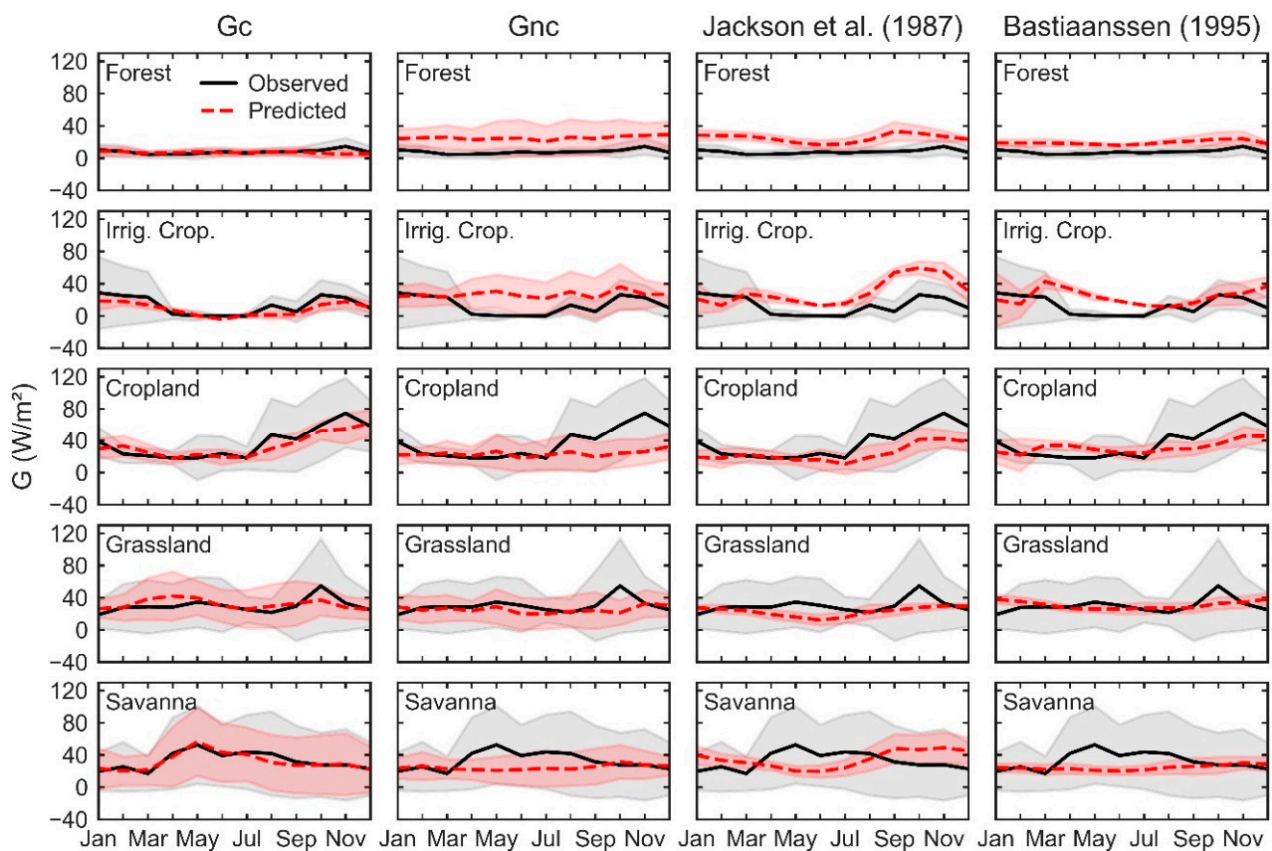
Figure 4. Observed (black, continuous) and predicted (red, dashed) seasonal monthly values of G, for each land cover. Shaded areas represent each series monthly standard deviations, while the lines represent the monthly averages.

Figure 5 presents the correlation plots of observed G in the flux towers to the predicted G by the Gc and Gnc ANNs, and by the Jackson et al. (1987) and Bastiaanssen (1995) adjusted models, distinguished by land cover. The Gc ANN presents the closest adherence to the 1:1 relationship line (black line in each plot), as the other methods underestimate higher values of G. H. A higher correlation was observed for the Gc ANN over savanna (r = 0.86), followed by grassland (r = 0.49). In contrast, the lowest correlation values were found over forest (r = 0.22) and irrigated cropland (r = 0.36). The Gnc model yielded a similar correlation for all land covers, ranging from r = 0.48 over savanna to r = 0.68 over irrigated cropland. The adjusted Jackson et al. (1987) best and worst performances were observed for grassland (r = 0.30) and savanna (r = −0.28), respectively. The Bastiaanssen (1995) adjusted model performed better over savanna (r = 0.38) and cropland (r = 0.29), and worse over irrigated cropland (r = −0.38) and forest (r = −0.13).

Figure 5. Correlation plots and linear correlation coefficient (r) values of observed G (horizontal axis) to predictions (vertical axis) by Gc and Gnc ANNs, as well as adjusted Jackson et al. (1987) and adjusted Bastiaanssen (1995) models, distinguished by color for each land cover. The diagonal black line represents the perfect 1:1 correlation between observed and predicted G.

Figure 6 displays other performance metrics of the neural networks and adjusted models for each land cover. G predictions over forest yielded generally low NS and A_0 values, except for Gnc (NS = 0.70). On the other hand, errors (MAE, RMSE, and bias) were also lower for forest in comparison to other land covers. Except Gc bias (−0.19 W/m²), the positive bias yielded by Gnc and the adjusted models indicates an overestimation of G for forests by models that do not consider land cover as an input. The Gnc ANN yielded the lowest errors of all models. The models and ANNs performed similarly over irrigated cropland, cropland, and grassland, being slightly better for grassland and worse for irrigated cropland. NS and A_0 values were low, ranging between −0.89 and 0.41 and between −0.23 and 0.35, respectively. The MAE ranged between 19.08 and 33.41 W/m²; RMSE ranged from 27.67 to 42.47 W/m²; and bias between −13.60 and 14.10 W/m². Over savanna, the performance of the ANNs and adjusted models differ the most. Gc predictions yielded NS = 0.72, while the others ranged between −0.33 and 0.01. Errors (MAE and RMSE) of the Gc and Gnc ANNs were nearly 50% lower than errors from the Jackson et al. (1987) and Bastiaanssen (1995) adjusted models. A_0 values were better for the ANNs as well, being 0.66 for Gc and 0.41 for Gnc, while the adjusted models ranged between −0.10 and 0.06.

Some contradictions can be identified between Figures 4 and 5 and Figure 6. Although higher correlation values are yielded by Gnc in comparison to the Gc ANN, Gc presents systematic biases of seasonal G values, as well as higher errors (MAE and RMSE) and lower values of NS and A_0. This discrepancy indicates that the correlation coefficient alone can be misleading as an accuracy metric.
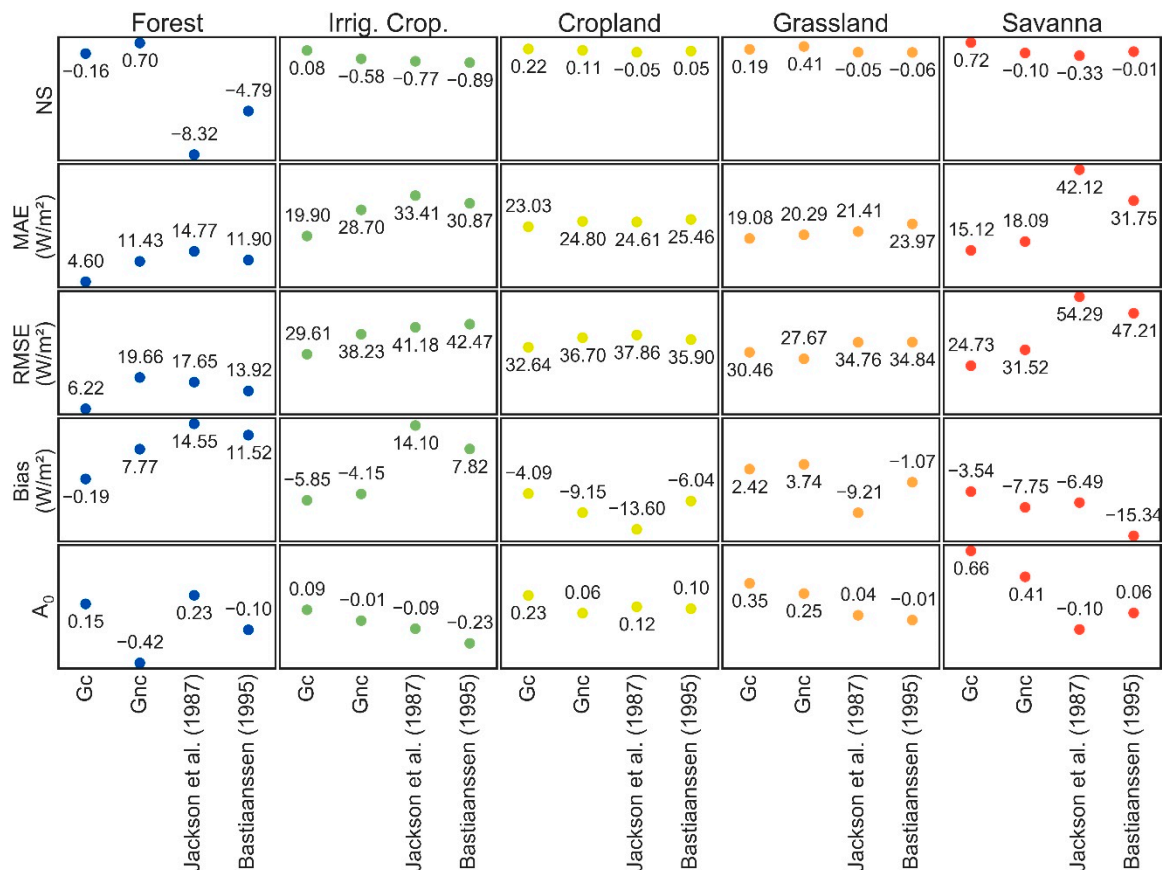
Figure 6. Other performance metrics (NS, MAE, RMSE, bias, and A_0) of the Gc, Gnc, adjusted Jackson et al. (1987), and adjusted Bastiaanssen (1995), distinguished by land cover.

4. Discussion

Uncertainties in the estimation of SEB fluxes from remote sensing data have motivated the use of ANNs as an alternative to existing models [91–93], but reference material regarding the validation of G estimates is still scarce, compared to other SEB fluxes. Table 6 presents the performance metrics of the G predictions via remote sensing presented by some recent studies. These works used either airborne [14,46,94] or orbiting [23,28,32,33,48,95] platforms associated with meteorological data to estimate G. Validation occurred via comparison to flux tower data in one or more land covers. In most studies, the number of samples was low, ranging between 12 and 698. An exception is the work performed by [27], with over 100,000 samples. The lowest MAE and RMSE values were found by [14] in cotton crops, while the highest values were obtained by [32] over forest. Generally, the highest correlation and lowest error values were obtained with airborne remote sensing data, which have higher spatial resolutions and are used over more homogenous land surfaces. This indicates that using remote sensing data with a finer spatial resolution may better correlate to G measurements.

Table 6. Compilation of other validation studies of G estimation from remote sensing data.

Source	Land Cover	Number of Samples	MAE (W/m^2)	RMSE (W/m^2)	r
[14]	Cropland, shrubland, and grassland	620	4.5–41.5	5.4–46.3	0.96–0.97
[23]	Savanna and cropland	56	35.0–41.0	-	0.66–0.75
[27]	Various	100,234	-	26.9–42.1	0.32–0.46
[32]	Forest	12	108.3–108.9	109.3–110.0	-
[33]	Croplands and forest	131	-	29.2–45.3	0.06–0.76
[46]	Croplands	50	-	-	0.73–0.96
[48]	Grasslands	67	40.6	43.0	0.70
[94]	Grasslands	698	-	23.0–28.0	0.57–0.77
[95]	Irrigated cropland and savanna	18	-	13.3	0.90

Compared to the metrics presented in Table 6, the Gc and Gnc ANNs and the Jackson et al. (1987) and Bastiaanssen (1995) models performed within expectations. However, unlike other studies, the models' adjustments and the ANN's training were performed for various land covers and a larger data series. This way, correlation values may be lower, however the resulting models and ANNs provide a higher generalization potential. Overall, the MAE and RMSE values of the ANNs and adjusted models in Table 5 reside in the lower portion of the ranges presented in Table 6.

As in our study, G predictions of the studies from Table 6 showed a lower correlation in densely forested areas. One possible reason for this is the low variance of observed G in forests. Another possible reason is that G is dependent on the soil skin temperature, while remote sensing products provide canopy surface temperature, which can be uncorrelated in dense vegetation. In fact, [96] verified that, despite the generally high correlation between soil and canopy temperatures, this relationship is weaker in cloud-free images, when radiative heating leads to fast changes in canopy temperature while soil temperature changes slowly. This limitation hinders the improvement of G prediction with orbiting remote sensors on forests. However, the Gc ANN yielded visibly lower errors (MAE, RMSE, and bias), indicating the superiority of this method.

The Bastiaanssen (1995) equation provided better G predictions than the Jackson et al. (1987) model, following observations of [33] and the assumption by [27] that models that consider T_s outperform those that do not. When compared to the existing models, both ANNs yielded better performance metrics, with Gc yielding the best predictions and closest seasonal patterns to those observed in the flux towers. Therefore, the introduction of land cover as a numerical input has also proved to be effective.

In this study, the optimal training/validation split ratio contrasts with that recommended by [81] (30/70 against 50/50 to 70/30). It is supposed that this difference is a result of the data set magnitude. As observed by [81], ANN training with large data sets is less sensitive to partitioning, since there is enough data to support pattern recognition. Thus, a smaller part of the data can be reserved for the training set.

The Gnc ANN yielded better performance than that of the Bastiaanssen (1995) model, despite both models using the same input data. This indicates that this ANN's structure is more suited for large-scale G estimation. In addition, ANNs can incorporate many more inputs, potentially achieving better performance. However, including too many inputs requires expanding the complexity of the network, which causes longer computational times. The low correlation among the considered input variables (R_n, T_s, EVI, and α) provides a low redundancy level to the developed ANNs, and each one is essential to the accurate prediction of G, as also indicated by [46]. The addition of land cover information complemented the input data series and improved performance without significantly impacting computational times. In Figure 4, systematic biases were identified for Gnc, Jackson et al. (1987), and Bastiaanssen (1995). This indicates that existing models' datasets do not fully comprehend the dynamics of G, further demonstrating the importance of land cover as a key factor in G prediction. Thus, additional datasets should be considered in the future development of models for G prediction. Figure 4 shows that the average value

and seasonality ranges of G vary significantly between the different land covers. It also shows that, even though this flux is of small magnitude compared to others of the SEB, it is not negligible. To the best of our knowledge, this work is one of the first large-scale model developments for G prediction in South America, comprehending multiple land covers and climates representative of most South American ecosystems. This work can potentially improve ET estimates in South America and similar regions worldwide.

5. Conclusions

In this study, ANNs were developed to calculate G via the integration of satellite remote sensing and meteorological reanalysis data. Compared to that of existing models, the performance of the generated ANNs was better overall, yielding lower errors and higher correlation values. This indicates that the ANN's structure can better approximate the behavior of G over various land covers and climate conditions.

The inclusion of land cover information into the ANNs as an input improved the accuracy of the G predictions. The superior performance of the ANNs with land cover as an input indicates that the commonly used remote sensing data may be insufficient to fully capture the differences among the surfaces and appropriately predict G. However, it is recommended to include additional remote sensing datasets in such models, especially those used in image land cover classification, instead of land cover data. This would remedy possible issues with the lack of standardization of land cover classification systems. On the other hand, the addition of input data sets increases the complexity of the ANNs and may even reduce accuracy. Thus, parsimony is recommended when selecting predictor data sets.

These findings demonstrate that the developed ANNs can predict G spatiotemporal variability more accurately than existing models. Despite the limitation of the distribution of the available flux towers, the wide variety of land covers considered, encompassing most of South America, and the length of the time series used in the ANN's training mean that the developed ANNs also yielded a higher generalization ability than the existing models. However, the ANN's accuracy over high altitude and meridional land covers should also be assessed for greater reliability.

For future studies, we suggest the mapping of G over the whole South America using the ANNs and a comparison of this to existing global products. The investigation of the effects of ANN-based G estimates on error reduction of surface energy balance fluxes and evapotranspiration modelling is also recommended.

Author Contributions: Conceptualization, B.C.C.d.A., O.C.P. and A.R.; Data curation, M.S.B., C.A.C.d.S., D.R.R., N.G.M., H.J.D., A.C.D.A., J.R.d.S.L., E.S.d.S., R.S., A.A.M., L.L. and R.B.K.; Funding acquisition, A.R.; Methodology, B.C.C.d.A. and O.C.P.; Project administration, A.R.; Supervision, O.C.P. and A.R.; Writing—original draft, B.C.C.d.A.; Writing—review & editing, B.C.C.d.A., O.C.P., A.R., M.S.B., C.A.C.d.S., D.R.R., A.C.D.A., J.R.d.S.L., E.S.d.S. and R.S. All authors have read and agreed to the published version of the manuscript.

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

References

1. Kustas, W.P.; Norman, J.M. Evaluation of soil and vegetation heat flux predictions using a simple two-source model with radiometric temperatures for partial canopy cover. *Agric. For. Meteorol.* **1999**, *94*, 13–29. [CrossRef]
2. Trenberth, K.E.; Fasullo, J.T.; Kiehl, J. Earth's Global Energy Budget. *Bull. Am. Meteorol. Soc.* **2009**, *90*, 311–324. [CrossRef]
3. Senay, G.B.; Bohms, S.; Singh, R.K.; Gowda, P.H.; Velpuri, N.M.; Alemu, H.; Verdin, J.P. Operational Evapotranspiration Mapping Using Remote Sensing and Weather Datasets: A New Parameterization for the SSEB Approach. *JAWRA J. Am. Water Resour. Assoc.* **2013**, *49*, 577–591. [CrossRef]
4. Senay, G.B.; Schauer, M.; Friedrichs, M.; Velpuri, N.M.; Singh, R.K. Satellite-based water use dynamics using historical Landsat data (1984–2014) in the southwestern United States. *Remote Sens. Environ.* **2017**, *202*, 98–112. [CrossRef]
5. Sauer, T.J.; Horton, R. Soil Heat Flux. In *Micrometeorology in Agricultural Systems*; Agronomy Monograph; American Society of Agronomy, Inc., Crop Science Society of America, Inc., Soil Science Society of America, Inc.: Madison, WI, USA, 2005; Volume 47, pp. 131–154.
6. Heusinkveld, B.G.; Jacobs, A.F.G.; Holtslag, A.A.M.; Berkowicz, S.M. Surface energy balance closure in an arid region: Role of soil heat flux. *Agric. For. Meteorol.* **2004**, *122*, 21–37. [CrossRef]
7. Kalma, J.D.; Jupp, D.L.B. Estimating evaporation from pasture using infrared thermometry: Evaluation of a one-layer resistance model. *Agric. For. Meteorol.* **1990**, *51*, 223–246. [CrossRef]
8. Menenti, M.; Choudhury, B.J. Parameterization of land surface evaporation by means of location dependent potential evaporation and surface temperature range. In *Exchange Processes at the Land Surface for a Range of Space and Time Scales, Proceedings of the Yokohama Symposium, Yokohama, Japan, 13–16 July 1993*; International Association of Hydrological Sciences: Wallingford, UK, 1993; pp. 561–568.
9. Norman, J.M.; Kustas, W.P.; Humes, K.S. Source approach for estimating soil and vegetation energy fluxes in observations of directional radiometric surface temperature. *Agric. For. Meteorol.* **1995**, *77*, 263–293. [CrossRef]
10. Anderson, M.C.; Norman, J.M.; Diak, G.R.; Kustas, W.P.; Mecikalski, J.R. A two-source time-integrated model for estimating surface fluxes using thermal infrared remote sensing. *Remote Sens. Environ.* **1997**, *60*, 195–216. [CrossRef]
11. Bastiaanssen, W.G.M.; Menenti, M.; Feddes, R.A.; Holtslag, A.A.M. A remote sensing surface energy balance algorithm for land (SEBAL) 1. Formulation. *J. Hydrol.* **1998**, *212–213*, 198–212. [CrossRef]
12. Bastiaanssen, W.G.M.; Pelgrum, H.; Wang, J.; Ma, Y.; Moreno, J.F.; Roerink, G.J.; van der Wal, T. A remote sensing surface energy balance algorithm for land (SEBAL) 2. Validation. *J. Hydrol.* **1998**, *212–213*, 213–229. [CrossRef]
13. Roerink, G.J.; Su, Z.; Menenti, M. S-SEBI: A simple remote sensing algorithm to estimate the surface energy balance. *Phys. Chem. Earth Part B Hydrol.* **2000**, *25*, 147–157. [CrossRef]
14. Su, Z. The Surface Energy Balance System (SEBS) for estimation of turbulent heat fluxes. *Hydrol. Earth Syst. Sci.* **2002**, *6*, 85–100. [CrossRef]
15. Allen, R.G.; Tasumi, M.; Trezza, R. Satellite-Based Energy Balance for Mapping Evapotranspiration with Internalized Calibration (METRIC)—Model. *J. Irrig. Drain. Eng.* **2007**, *133*, 380–394. [CrossRef]
16. Senay, G.; Budde, M.; Verdin, J.; Melesse, A. A Coupled Remote Sensing and Simplified Surface Energy Balance Approach to Estimate Actual Evapotranspiration from Irrigated Fields. *Sensors* **2007**, *7*, 979–1000. [CrossRef]
17. Anderson, M.C.; Kustas, W.P.; Norman, J.M.; Hain, C.R.; Mecikalski, J.R.; Schultz, L.; González-Dugo, M.P.; Cammalleri, C.; d'Urso, G.; Pimstein, A.; et al. Mapping daily evapotranspiration at field to continental scales using geostationary and polar orbiting satellite imagery. *Hydrol. Earth Syst. Sci.* **2011**, *15*, 223–239. [CrossRef]
18. Senay, G.B.; Kagone, S.; Velpuri, N.M. Operational Global Actual Evapotranspiration: Development, Evaluation, and Dissemination. *Sensors* **2020**, *20*, 1915. [CrossRef]
19. Bastiaanssen, W.G.M.; Noordman, E.J.M.; Pelgrum, H.; Davids, G.; Thoreson, B.P.; Allen, R.G. SEBAL Model with Remotely Sensed Data to Improve Water-Resources Management under Actual Field Conditions. *J. Irrig. Drain. Eng.* **2005**, *131*, 85–93. [CrossRef]
20. Bastiaanssen, W.; Thoreson, B.; Clark, B.; Davids, G. Discussion of "Application of SEBAL Model for Mapping Evapotranspiration and Estimating Surface Energy Fluxes in South-Central Nebraska" by Ramesh, K. Singh, Ayse Irmak, Suat Irmak, and Derrel, L. Martin. *J. Irrig. Drain. Eng.* **2010**, *136*, 282–283. [CrossRef]
21. Singh, R.K.; Irmak, A.; Irmak, S.; Martin, D.L. Application of SEBAL model for mapping evapotranspiration and estimating surface energy fluxes in south-central Nebraska. *J. Irrig. Drain. Eng.* **2008**, *134*, 273–285. [CrossRef]
22. Zhang, S.W.; Lei, Y.P.; Li, H.J.; Wang, Z. Temporal-spatial variation in crop evapotranspiration in Hebei Plain, China. *J. Food Agric. Environ.* **2010**, *8*, 672–677.
23. Ruhoff, A.L.; Paz, A.R.; Collischonn, W.; Aragao, L.E.O.C.; Da Rocha, H.R.; Malhi, Y.S. A MODIS-Based Energy Balance to Estimate Evapotranspiration for Clear-Sky Days in Brazilian Tropical Savannas. *Remote Sens.* **2012**, *4*, 703–725. [CrossRef]
24. Liou, Y.-A.; Kar, S. Evapotranspiration Estimation with Remote Sensing and Various Surface Energy Balance Algorithms—A Review. *Energies* **2014**, *7*, 2821–2849. [CrossRef]
25. Bhattarai, N.; Shaw, S.B.; Quackenbush, L.J.; Im, J.; Niraula, R. Evaluating five remote sensing based single-source surface energy balance models for estimating daily evapotranspiration in a humid subtropical climate. *Int. J. Appl. Earth Obs. Geoinf.* **2016**, *49*, 75–86. [CrossRef]

26. Wagle, P.; Bhattarai, N.; Gowda, P.H.; Kakani, V.G. Performance of five surface energy balance models for estimating daily evapotranspiration in high biomass sorghum. *ISPRS J. Photogramm. Remote Sens.* **2017**, *128*, 192–203. [CrossRef]
27. Purdy, A.J.; Fisher, J.B.; Goulden, M.L.; Famiglietti, J.S. Ground heat flux: An analytical review of 6 models evaluated at 88 sites and globally. *J. Geophys. Res. Biogeosci.* **2016**, *121*, 3045–3059. [CrossRef]
28. Cammalleri, C.; La Loggia, G.; Loggia, A.; Maltese, A. Critical analysis of empirical ground heat flux equations on a cereal field using micrometeorological data. In Proceedings of the SPIE—The International Society for Optical Engineering, Berlin, Germany, 1–3 September 2009. [CrossRef]
29. Kilic, A.; Singh, R.K.; Walter-Shea, E.A.; Verma, S.; Suyker, A.E. Comparison and Analysis of Empirical Equations for Soil Heat Flux for Different Cropping Systems and Irrigation Methods. *Pap. Nat. Resour.* **2011**, *334*, 67–80.
30. Russell, E.S.; Liu, H.; Gao, Z.; Finn, D.; Lamb, B. Impacts of soil heat flux calculation methods on the surface energy balance closure. *Agric. For. Meteorol.* **2015**, *214–215*, 189–200. [CrossRef]
31. Dhungel, S.; Barber, M. Estimating Calibration Variability in Evapotranspiration Derived from a Satellite-Based Energy Balance Model. *Remote Sens.* **2018**, *10*, 1695. [CrossRef]
32. Laipelt, L.; Ruhoff, A.L.; Fleischmann, A.S.; Kayser, R.H.B.; Kich, E.; da Rocha, H.R.; Neale, C.M.U. Assessment of an Automated Calibration of the SEBAL Algorithm to Estimate Dry-Season Surface-Energy Partitioning in a Forest–Savanna Transition in Brazil. *Remote Sens.* **2020**, *12*, 1108. [CrossRef]
33. Danelichen, V.H.M.; Biudes, M.S.; Souza, M.C.; Machado, N.G.; Silva, B.B.; Nogueira, J.S. Estimation of soil heat flux in a neotropical Wetland region using remote sensing techniques. *Rev. Bras. Meteorol.* **2014**, *29*, 469–482. [CrossRef]
34. Villarreal, S.; Vargas, R. Representativeness of FLUXNET sites across Latin America. *J. Geophys. Res. Biogeosci.* **2021**, *126*, e2020JG006090. [CrossRef]
35. Saleska, S.R.; Da Rocha, H.R.; Huete, A.R.; Nobre, A.D.; Artaxo, P.E.; Shimabukuro, Y.E. LBA-ECO CD-32 Flux Tower Network Data Compilation, Brazilian Amazon: 1999–2006. ORNL Distributed Active Archive Center 2013. Available online: http://daac.ornl.gov/cgi-bin/dsviewer.pl?ds_id=1174 (accessed on 30 June 2020).
36. Davidson, E.A.; Artaxo, P. Globally significant changes in biological processes of the Amazon Basin: Results of the Large-scale Biosphere–Atmosphere Experiment. *Glob. Chang. Biol.* **2004**, *10*, 519–529. [CrossRef]
37. Borges, C.K.; dos Santos, C.A.C.; Carneiro, R.G.; da Silva, L.L.; de Oliveira, G.; Mariano, D.; Silva, M.T.; da Silva, B.B.; Bezerra, B.G.; Perez-Marin, A.M.; et al. Seasonal variation of surface radiation and energy balances over two contrasting areas of the seasonally dry tropical forest (Caatinga) in the Brazilian semi-arid. *Environ. Monit. Assess.* **2020**, *192*, 524. [CrossRef]
38. Tabarelli, M.; da Rocha, C.F.D.; Romanowski, H.P.; Rocha, O.; de Lacerda, L.D. *PELD—CNPq: Dez Anos do Programa de Pesquisas Ecológicas de Longa Duração do Brasil: Achados, Lições e Perspectivas*; Universitária da UFPE: Recife, Brazil, 2013; p. 446. ISBN 978-85-415-0329-7.
39. Biudes, M.S.; Vourlitis, G.L.; Machado, N.G.; de Arruda, P.H.Z.; Neves, G.A.R.; de Almeida Lobo, F.; Neale, C.M.U.; de Souza Nogueira, J. Patterns of energy exchange for tropical ecosystems across a climate gradient in Mato Grosso, Brazil. *Agric. For. Meteorol.* **2015**, *202*, 112–124. [CrossRef]
40. Da Rocha, H.R.; Manzi, A.O.; Cabral, O.M.; Miller, S.D.; Goulden, M.L.; Saleska, S.R.; Coupe, N.; Wofsy, S.C.; Borma, L.S.; Artaxo, P.; et al. Patterns of water and heat flux across a biome gradient from tropical forest to savanna in Brazil. *J. Geophys. Res.* **2009**, *114*. [CrossRef]
41. Eva, H.D.; Belward, A.S.; De Miranda, E.E.; Di Bella, C.M.; Gond, V.; Huber, O.; Jones, S.; Sgrenzaroli, M.; Fritz, S. A land cover map of South America. *Glob. Chang. Biol.* **2004**, *10*, 731–744. [CrossRef]
42. Hornik, K.; Stinchcombe, M.; White, H. Multilayer feedforward networks are universal approximators. *Neural Netw.* **1989**, *2*, 359–368. [CrossRef]
43. Silverman, D.; Dracup, J.A. Artificial Neural Networks and Long-Range Precipitation Prediction in California. *J. Appl. Meteorol.* **2000**, *39*, 57–66. [CrossRef]
44. Zanetti, S.S.; Sousa, E.F.; Oliveira, V.P.; Almeida, F.T.; Bernardo, S. Estimating Evapotranspiration Using Artificial Neural Network and Minimum Climatological Data. *J. Irrig. Drain. Eng.* **2007**, *133*, 83–89. [CrossRef]
45. Tabari, H.; Sabziparvar, A.-A.; Ahmadi, M. Comparison of artificial neural network and multivariate linear regression methods for estimation of daily soil temperature in an arid region. *Meteorol. Atmos. Phys.* **2010**, *110*, 135–142. [CrossRef]
46. Canelón, D.J.; Chávez, J.L. Soil Heat Flux Modeling Using Artificial Neural Networks and Multispectral Airborne Remote Sensing Imagery. *Remote Sens.* **2011**, *3*, 1627–1643. [CrossRef]
47. Jimeno-Sáez, P.; Senent-Aparicio, J.; Pérez-Sánchez, J.; Pulido-Velazquez, D. A Comparison of SWAT and ANN Models for Daily Runoff Simulation in Different Climatic Zones of Peninsular Spain. *Water* **2018**, *10*, 192. [CrossRef]
48. Käfer, P.; Souza da Rocha, N.; Ribeiro Diaz, L.; Kaiser, E.; Santos, D.; Veeck, G.; Robérti, D.; Rolim, S.; Oliveira, G. Artificial neural networks model based on remote sensing to retrieve evapotranspiration over the Brazilian Pampa. *J. Appl. Remote Sens.* **2020**, *14*, 038504. [CrossRef]
49. Adadi, A.; Berrada, M. Peeking inside the black-box: A survey on Explainable Artificial Intelligence (XAI). *IEEE Access* **2018**, 52138–52160. [CrossRef]
50. Borma, L.S.; Da Rocha, H.R.; Cabral, O.M.; von Randow, C.; Collicchio, E.; Kurzatkowski, D.; Brugger, P.J.; Freitas, H.; Tannus, R.; Oliveira, L.; et al. Atmosphere and hydrological controls of the evapotranspiration over a floodplain forest in the Bananal Island region, Amazonia. *J. Geophys. Res.* **2009**, *114*. [CrossRef]

51. Santos, A.J.B.; Silva, G.T.D.A.; Miranda, H.S.; Miranda, A.C.; Lloyd, J. Effects of fire on surface carbon, energy and water vapour fluxes over campo sujo savanna in central Brazil. *Funct. Ecol.* **2003**, *17*, 711–719. [CrossRef]

52. Aguiar, L.J.G.A. *Fluxos de Massa e Energia Para a Cultura de Milho (Zea mays L.) no Rio Grande do Sul*; Universidade Federal de Viçosa: Viçosa, Brazil, 2011.

53. Moreira, V.S. *Balanço de Água no Ciclo da Cultura se Soja: Representação no Modelo de Vegetação Dinâmica Agro-Ibis*; Universidade Federal de Santa Maria: Santa Maria, Brazil, 2012.

54. Carneiro, J.V. *Estimativas das Trocas Líquidas de Carbono em Duas Áreas de Cultivo de Arroz Irrigado na Região Central do RS*; Universidade Federal de Santa Maria: Santa Maria, Brazil, 2012.

55. Souza, R.; Feng, X.; Antonino, A.; Montenegro, S.; Souza, E.; Porporato, A. Vegetation response to rainfall seasonality and interannual variability in tropical dry forests. *Hydrol. Process.* **2016**, *30*, 3583–3595. [CrossRef]

56. Lathuillière, M.J.; Dalmagro, H.J.; Black, T.A.; Arruda, P.H.Z.d.; Hawthorne, I.; Couto, E.G.; Johnson, M.S. Rain-fed and irrigated cropland-atmosphere water fluxes and their implications for agricultural production in Southern Amazonia. *Agric. For. Meteorol.* **2018**, *256–257*, 407–419. [CrossRef]

57. Brito, T.; Lima, J.; Oliveira, C.; Souza, R.; Antonino, A.; Medeiros, É.; Souza, E.; Alves, E. Mudanças no Uso da Terra e Efeito nos Componentes do Balanço Hídrico no Agreste Pernambucano. *Rev. Bras. Geogr. Fís.* **2020**, *13*, 870–886. [CrossRef]

58. da Silva, P.F.; de Sousa Lima, J.R.; Antonino, A.C.D.; Souza, R.; Souza, E.S.; de Silva, J.R.I.; Alves, E.M. Seasonal patterns of carbon dioxide, water and energy fluxes over the Caatinga and grassland in the semi-arid region of Brazil. *J. Arid. Environ.* **2017**, *147*, 71–82. [CrossRef]

59. Hasler, N.; Avissar, R. What Controls Evapotranspiration in the Amazon Basin? *J. Hydrometeorol.* **2007**, *8*, 380–395. [CrossRef]

60. Araújo, A.C. Comparative measurements of carbon dioxide fluxes from two nearby towers in a central Amazonian rainforest: The Manaus LBA site. *J. Geophys. Res.* **2002**, *107*. [CrossRef]

61. Dalmagro, H.J.; de Arruda, P.H.Z.; Vourlitis, G.L.; Lathuillière, M.J.; Nogueira, J.d.S.; Couto, E.G.; Johnson, M.S. Radiative forcing of methane fluxes offsets net carbon dioxide uptake for a tropical flooded forest. *Glob. Chang. Biol.* **2019**, *25*. [CrossRef] [PubMed]

62. Rubert, G.C.R.; Roberti, D.R.; Diaz, M.B.; de Moraes, O.L.L. Estimativa da evapotranspiração em área de pastagem em Santa Maria—RS. *Ciênc. Nat.* **2016**, *38*, 300–304. [CrossRef]

63. Oliveira, M.B.L.; Von Randow, C.; Manzi, A.O.; Alvala, R.; As, L.D.A.; Souza, A. Fluxos Turbulentos de Energia Sobre o Pantanal Sul Matogrossense. *Rev. Bras. Meteorol.* **2006**, *21*, 371–377.

64. Da Rocha, H.R.; Freitas, H.C.; Rosolem, R.; Juárez, R.I.N.; Tannus, R.N.; Ligo, M.A.; Cabral, O.M.R.; Dias, M.A.F.S. Measurements of CO$_2$ exchange over a woodland savanna (Cerrado Sensu stricto) in southeast Brasil. *Biota Neotrop.* **2002**, *2*, 1–11. [CrossRef]

65. Timm, A.U.; Roberti, D.R.; Streck, N.A.; de Gonçalves, L.G.G.; Acevedo, O.C.; Moraes, O.L.; Moreira, V.S.; Degrazia, G.A.; Ferlan, M.; Toll, D.L. Energy partitioning and evapotranspiration over a rice paddy in Southern Brazil. *J. Hydrometeorol.* **2014**, *15*, 1975–1988. [CrossRef]

66. Zimmer, T.; Stefanello, M.B.; Moreira, V.S.; Diaz, M.B.; Souza, V.D.A.; Roberti, D.R.; Maria, S. Estimating the soil thermal conductivity using experimental soil heat flux in a rice paddy area. *Amer. J. Environ. Eng.* **2016**, *6*, 103–108.

67. Sakai, R.K.; Fitzjarrald, D.R.; Moraes, O.L.L.; Staebler, R.M.; Acevedo, O.C.; Czikowsky, M.J.; Da Silva, R.; Brait, E.; Miranda, V. Land-use change effects on local energy, water, and carbon balances in an Amazonian agricultural field. *Glob. Chang. Biol.* **2004**, *10*, 895–907. [CrossRef]

68. Goulden, M.L.; Miller, S.D.; Da Rocha, H.R.; Menton, M.C.; De Freitas, H.C.; Michela, A.; Figueira, S.; Cleilim, A.; Dias, A.; Sousa, D. Diel and seasonal patterns of tropical forest CO2 exchange. *Ecol. Appl.* **2004**, *14*, 42–54. [CrossRef]

69. Cabral, O.M.R.; Da Rocha, H.R.; Ligo, M.A.V.; Brunini, O.; Dias, M.A.S. Fluxos turbulentos de calor sensivel, vapor de água e CO2 sobre plantaçao de cana-de-açúcar (*Saccharum* sp.) em Sertãozinho-SP. *Rev. Bras. Meteorol.* **2003**, *18*, 61–70.

70. Rodell, M.; Houser, P.R.; Jambor, U.; Gottschalck, J.; Mitchell, K.; Meng, C.-J.; Arsenault, K.; Cosgrove, B.; Radakovich, J.; Bosilovich, M.; et al. The Global Land Data Assimilation System, Bull. *Amer. Meteor. Soc.* **2004**, *85*, 381–394. [CrossRef]

71. Wan, Z.; Hook, S.; Hulley, G. MOD11A2 MODIS/Terra Land Surface Temperature/Emissivity 8-Day L3 Global 1 km SIN Grid V006 Data Set. *NASA EOSDIS Land Process. DAAC* **2015**. [CrossRef]

72. Schaaf, C.; Wang, Z. MCD43A3 MODIS/Terra+Aqua BRDF/Albedo Daily L3 Global—500 m V006 Data Set. *NASA EOSDIS Land Process. DAAC* **2015**. [CrossRef]

73. Didan, K. MOD13Q1 MODIS/Terra Vegetation Indices 16-Day L3 Global 250m SIN Grid V006 [Data Set]. *NASA EOSDIS Land Process. DAAC* **2015**. [CrossRef]

74. Idso, S.B.; Aase, J.K.; Jackson, R.D. Net radiation—soil heat flux relations as influenced by soil water content variations. *Bound. Layer Meteorol.* **1975**, *9*, 113–122. [CrossRef]

75. Kustas, W.P.; Daughtry, C.S.T. Estimation of the Soil Heat Flux/Net Radiation from Spectral Data. *Agric. For. Meteorol.* **1990**, *49*, 205–223. [CrossRef]

76. Kustas, W.P.; Daughtry, C.S.T.; Van Oevelen, P.J. Analytical Treatment of the Relationships between Soil Heat Flux/Net Radiation and Vegetation Indices. *Remote Sens. Environ.* **1993**, *46*, 319–330. [CrossRef]

77. Lucchese, L.V.; Oliveira, G.G.d.; Pedrollo, O.C. Attribute selection using correlations and principal components for artificial neural networks employment for landslide susceptibility assessment. *Environ. Monit. Assess.* **2020**, *192*, 129. [CrossRef] [PubMed]

78. Rumelhart, D.E.; Hinton, G.E.; Williams, R.J. Learning representations by back-propagating errors. *Nature* **1986**, *323*, 533–536. [CrossRef]

79. Widrow, B.; Hoff, M.E. Adaptive Switching Circuits. In *1960 IRE WESCON Convention Record*, 4th ed.; Part 4; IRE: New York, NY, USA, 1960; pp. 96–104.

80. Vogl, T.P.; Mangis, J.K.; Rigler, A.K.; Zink, W.T.; Alkon, D.L. Accelerating the convergence of the backpropagation method. *Biol. Cybern.* **1988**, *59*, 257–263. [CrossRef]

81. Crowther, P.S.; Cox, R.J. A Method for Optimal Division of Data Sets for Use in Neural Networks. *Knowl. Based Intell. Inf. Eng. Syst.* **2005**, 1–7. [CrossRef]

82. Nash, J.E.; Sutcliffe, J.V. River flow forecasting through conceptual models part I—A discussion of principles. *J. Hydrol.* **1970**, *10*, 282–290. [CrossRef]

83. Jackson, R.D.; Moran, M.S.; Gay, L.W.; Raymond, L.H. Evaluating evaporation from field crops using airborne radiometry and ground-based meteorological data. *Irrig. Sci.* **1987**, *8*, 81–90. [CrossRef]

84. Bastiaanssen, W.G.M. Regionalization of Surface Flux Densities and Moisture Indicators in Composite Terrain. Ph.D. Thesis, Wageningem Agricultural University, Wageningen, The Netherlands, November 1995; p. 273.

85. Willmott, C.J.; Matsuura, K. Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. *Clim. Res.* **2005**, *30*, 79–82. [CrossRef]

86. Boylan, J.W.; Russell, A.G. PM and light extinction model performance metrics, goals, and criteria for three-dimensional air quality models. *Atmos. Environ.* **2006**, *40*, 4946–4959. [CrossRef]

87. Sen, P.K. Estimates of the regression coefficient based on Kendall's tau. *J. Am. Stat. Assoc.* **1968**, *63*, 1379–1389. [CrossRef]

88. Wilks, D.S. *Statistical Methods in the Atmospheric Sciences*; Academic Press: London, UK, 2011; p. 676.

89. Garson, G.D. Interpreting neural-network connection weights. *Artif. Intell. Expert* **1991**, *6*, 47–51.

90. Goh, A.T.C. Back-propagation neural networks for modeling complex systems. *Artif. Intell. Eng.* **1995**, *9*, 143–151. [CrossRef]

91. Alemohammad, S.H.; Fang, B.; Konings, A.G.; Aires, F.; Green, J.K.; Kolassa, J.; Miralles, D.; Prigent, C.; Gentine, P. Water, Energy, and Carbon with Artificial Neural Networks (WECANN): A statistically based estimate of global surface turbulent fluxes and gross primary productivity using solar-induced fluorescence. *Biogeosciences* **2017**, *14*, 4101–4124. [CrossRef] [PubMed]

92. Feng, J.; Wang, W.; Xu, F.; Sun, S. Estimating surface heat and water vapor fluxes by combining two-source energy balance model and backpropagation neural network. *Sci. Total. Environ.* **2020**, *729*, 138724. [CrossRef] [PubMed]

93. Walls, S.; Binns, A.D.; Levison, J.; MacRitchie, S. Prediction of actual evapotranspiration by artificial neural network models using data from a Bowen ratio energy balance station. *Neural Comput. Appl.* **2020**. [CrossRef]

94. Anderson, M.; Norman, J.; Kustas, W.; Houborg, R.; Starks, P.; Agam, N. A thermal-based remote sensing technique for routine mapping of land-surface carbon, water and energy fluxes from field to regional scales. *Remote Sens. Environ.* **2008**, *112*, 4227–4241. [CrossRef]

95. de Teixeira, A.H.; Bastiaanssen, W.G.M.; Ahmad, M.D.; Bos, M.G. Reviewing SEBAL input parameters for assessing evapotranspiration and water productivity for the Low-Middle São Francisco River basin, Brazil. *Agric. For. Meteorol.* **2009**, *149*, 462–476. [CrossRef]

96. Kim, Y.; Still, C.J.; Hanson, C.V.; Kwon, H.; Greer, B.T.; Law, B.E. Canopy skin temperature variations in relation to climate, soil temperature, and carbon flux at a ponderosa pine forest in central Oregon. *Agric. For. Meteorol.* **2016**, *226–227*, 161–173. [CrossRef]