Article

MODIS-Based Estimation of Terrestrial Latent Heat Flux over North America Using Three Machine Learning Algorithms

Xuanyu Wang [1,2] iD **, Yunjun Yao** [1,2,*] iD **, Shaohua Zhao** [3] **, Kun Jia** [1,2] iD **, Xiaotong Zhang** [1,2] **, Yuhu Zhang** [4] **, Lilin Zhang** [1,2] **, Jia Xu** [1,2] **and Xiaowei Chen** [1,2]

1   State Key Laboratory of Remote Sensing Science, Faculty of Geographical Science,
    Beijing Normal University, Beijing 100875, China; skywalkerwxy@163.com (X.W.); jiakun@bnu.edu.cn (K.J.);
    xtngzhang@bnu.edu.cn (X.Z.); 201521170059@mail.bnu.edu.cn (L.Z.); diamond1201@sina.cn (J.X.);
    18368091857@163.com (X.C.)
2   Beijing Engineering Research Center for Global Land Remote Sensing Products, Institute of Remote Sensing
    Science and Engineering, Faculty of Geographical Science, Beijing Normal University, Beijing 100875, China
3   Satellite Environment Center, Ministry of Environmental Protection, Beijing 100094, China; zshyytt@126.com
4   College of Resource Environment and Tourism, Capital Normal University, Beijing 100048, China;
    zhang_yuhu@163.com
*   Correspondence: boyyunjun@163.com; Tel.: +86-10-5880-3002

Abstract: Terrestrial latent heat flux (LE) is a key component of the global terrestrial water, energy, and carbon exchanges. Accurate estimation of LE from moderate resolution imaging spectroradiometer (MODIS) data remains a major challenge. In this study, we estimated the daily LE for different plant functional types (PFTs) across North America using three machine learning algorithms: artificial neural network (ANN); support vector machines (SVM); and, multivariate adaptive regression spline (MARS) driven by MODIS and Modern Era Retrospective Analysis for Research and Applications (MERRA) meteorology data. These three predictive algorithms, which were trained and validated using observed LE over the period 2000–2007, all proved to be accurate. However, ANN outperformed the other two algorithms for the majority of the tested configurations for most PFTs and was the only method that arrived at 80% precision for LE estimation. We also applied three machine learning algorithms for MODIS data and MERRA meteorology to map the average annual terrestrial LE of North America during 2002–2004 using a spatial resolution of 0.05°, which proved to be useful for estimating the long-term LE over North America.

Keywords: latent heat; machine learning algorithms; plant functional type

1. Introduction

Terrestrial latent heat flux (LE) plays an important role in the global water cycle, carbon, and surface energy exchanges [1–3]. LE can be measured directly by using lysimeter and eddy covariance (EC) flux towers, but the former method is time-consuming, expensive, and needs carefully planned experiments [4]. Although the FLUXNET project has provided EC flux towers to measure LE, it is inherently difficult to accurately estimate LE, especially at large spatial scales due to the heterogeneity in the terrestrial landscape. Moreover, the sparse observations from EC flux towers limit the accurate characterization of spatiotemporal LE patterns over large spatial scales.

Remote sensing, especially from the Moderate Resolution Imaging Spectroradiometer (MODIS) satellite, can provide temporally and spatially continuous information for estimating LE, such as: land surface temperature (LST); the fraction of absorbed photosynthetically active radiation (FPAR);

the normalized difference vegetation index (NDVI); the enhanced vegetation index (EVI); leaf area index (LAI); and, albedo [5–7]. Currently, there are numerous satellite-based algorithms available to estimate LE, e.g., surface energy balance (SEB) models [8,9] use remotely sensed products and meteorological data driving a residual SEB equation to estimate LE. Penman-Monteith (PM) and Priestley-Taylor (PT) approaches used PM and PT equations to calculate LE, respectively [10–13]. Statistical and empirical methods calculated LE using empirical equations, which include vegetation parameters and LE observations [6,14–17]. Data assimilate methods improved LE estimation by assimilating satellite land-surface variables into land surface models [18,19]. However, these algorithms show substantial differences in LE estimation [20]. Additionally, the existing global LE products, such as MOD16 and the EUMETSAT Satellite Application Facility on Land Surface Analysis (LSA-SAF) product (LSA-SAF MSG), have a 5 km spatial resolution and daily temporal resolution [7], but validation results have revealed that these products often overestimate LE at most AsiaFlux sites [21].

Satellite-based empirical and semi-empirical LE algorithms have been scaled up from site to regional scales by relating observed LE to satellite-based vegetation parameters and other key meteorological variables [14]. As the most viable empirical algorithms, machine learning algorithms have the potential to generate robust relationships, and once trained, they are very fast to apply [22]. These algorithms can provide accurate LE estimation as long as enough training datasets are representative of all the behaviors in the systems. However, there is still some controversy about which machine learning algorithm is best in LE estimates. Different input variables with different plant functional types (PFTs) over different study regions may lead to different result. For example, Bruton et al. [23] employed precipitation (P), air temperature (Ta), solar radiation (Rs), wind speed (Ws), and relative humidity (RH) as input data for the artificial neural network (ANN) algorithms to estimate LE. The validation results showed that ANN algorithm was able to perform reliable LE estimation with an R^2 of 0.717 and a root mean square error of 1.11 mm. Compared with multiple linear regression and the Priestley-Taylor method, the ANN method estimated the most accurate LE with all of the available variables. El-Shafie et al. [24] used ANN models for prediction evapotranspiration in Iran. The results demonstrated that an ANN model predicted daily E with a significant level of accuracy using only the maximum and minimum temperatures successfully. Shrestha and Shukla [25] used SVM for predicting generic crop coefficient (Kc) and crop evapotranspiration (ETc), using a uniquely large dataset from lysimeters for multiple crop-season combinations under plastic mulch conditions. The results showed that the SVM model was superior to the Artificial Neural Network and Relevance Vector Machine models, two data-driven models used in hydrology. Deo et al. [26] used three machine learning algorithms, including Relevance Vector Machine (RVM), Extreme Learning Machine (ELM), and MARS for the prediction of monthly evaporative loss using limited meteorological data from the Amberley weather station in Australia. All three machine learning models performed well, with RVM proving to be the best. They also found that incident solar radiation and air temperatures are the two most influential factors in determining the performance of machine learning algorithms.

The observation data from EC flux towers are more adequate and accurate over North America, so there already exist some studies of machine learning algorithms for LE estimation over North America. For example, Yang et al. [27] used SVM combined with ground-measured ET from 25 AmeriFlux sites and remotely sensed inputs (land surface temperature, enhanced vegetation index, and land cover) to predict ET over the United States (U.S.) with an average test error of 0.62 mm/day and an R^2 of 0.75. When compared to neural network and multiple regressions, SVM performed better. Adnan et al. [28] developed a model to estimate evapotranspiration with meteorological parameters (temperature, relative humidity, wind speed, and precipitation) using different machine learning techniques. They found that ANN performed better than LLSVM, MARS, and M5Tree models and gave the nearest values as compared with the actual value. Therefore, substantial differences still exist in simulating LE for different machine learning algorithms. Moreover, there is a lack of validating and evaluating different machine learning algorithms for LE estimation at different PFTs over North America.

In this study, we evaluated and applied three machine learning algorithms to estimate terrestrial LE over North America. These machine learning algorithms included artificial neural network (ANN), support vector machines (SVM), and multivariate adaptive regression spline (MARS), which were using air temperature (T), relative humidity (RH), wind speed (Ws), and MODIS-based NDVI as input data. The objectives of this study were as follows: (1) to assess the performance of the three machine learning algorithms for different PFTs based on a series of cross-validations using long-term Ameriflux EC observations from 2000 to 2007, and (2) to generate a daily LE product during 2002–2004 over North America using three machine learning algorithms with well-quantified accuracy based on MODIS data and Modern Era Retrospective Analysis for Research and Applications (MERRA) meteorological data.

2. Materials and Methods

2.1. Machine Learning Algorithms

2.1.1. Artificial Neural Network

An Artificial neural network (ANN) consists of a set of processing elements, including input, hidden, and output layers. Each layer includes an array of processing elements. There is a connection between each of the neurons in any given layer with each of the neurons in the next layer [4]. This connection between neurons can be described as following function:

$$y_i = f_i(\sum_{j=1}^{n} w_{ij}x_j - \theta_i) \tag{1}$$

where y_i is the output of the neuron i, f_i is nonlinear, such as Gaussian function, etc. w_{ij} is the connection weight between neuron i and j. x_j is the jth input to the neuron and θ_i is the bias of the neuron.

The processing steps require aggregating the weighted inputs and the activation function, such as the Sigmoid function and Tanh function. The activation function limits the output values between two asymptotic values, typically 0 and 1 or −1 and +1. In feed-forward artificial neural networks (FFNN), each neuron at input and inner layers receives input data then processes and passes to the next layer. The output result is carried out layer by layer in the forward direction. The weights are adjusted through reverse pass, which calculated error signals propagate backward through the network. The weights of the output neuron layer are adjusted first because the target value of output neuron layer can guide the adjustment of the associated weights. The numbers of input and output parameters determine the numbers of neurons in the input layer and the output layer, respectively [29]. The model of FFNN is shown in Figure 1.
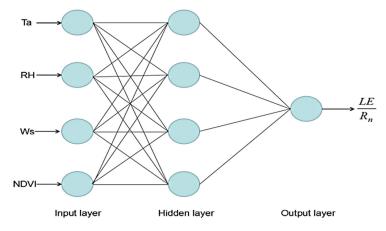

Figure 1. Architecture of the neural network model used in this study.

2.1.2. Support Vector Machine

Support vector machine (SVM) were developed by Cortes and Vapnik (1995) to separate the sample data through some nonlinear mapping from the input vectors into the high dimensional feature space, and can find an optimal separating hyper plane between data points of different classes in a high dimension space [27,30,31]. When using SVM classification, we seek the optimal classification plane to maximize the distance between the two classes, with the points falling within the boundary, called the support vector [30]. Mapping technology is implemented through the kernel function. We used the radial basis function (RBF) kernel function type in this study because former studies shows that the RBF kernel performs better than any other kernel functions [32], it can be expressed as:

$$K(x_i, x_j) = \exp(-\frac{1}{2\sigma^2}\|x_i - x_j\|^2) \qquad (2)$$

Here, σ is variance. The complexity of SVM does not depend on the dimensions of the feature space, but rather on the number of support vectors (Figure 2).
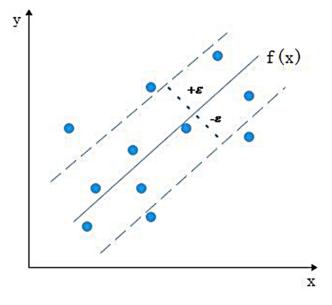

Figure 2. One-dimensional linear regression with ε-insensitive band for the support vector machine (SVM) algorithm.

The regression estimation with support vector is to estimate a function according to a given data set $\{(x_i, y_i)\}_n$. Herein, the input vectors (x_i) refer to T, RH, Ws, and NDVI, whereas the target values (y_i) refer to the ratio of LE values as determined by the ground-measured data and surface net radiation (R_n), and n is the total number of data sets.

The linear regression function uses the following function:

$$f(x)' = w \cdot f(x) + b \qquad (3)$$

Here, $f(x)$ is a nonlinear function, by which x is mapped into a feature space, b is a weight vector, and w denotes a coefficient that has to be estimated from the data. A nonlinear regression function is given by minimizing the sum of the empirical risk and a complexity term, as in the following expression:

$$f(x) = \sum_{i=1}^{n} (a_i^* - a_i)K(x_i, x_j) + b \qquad (4)$$

With $a_i a_i^* \geq 0$, $(a_i$ and a_i^* are Lagrange multipliers). $i = 1, \ldots, N$, $K(x_i, x_j)$ is the kernel function and b can be calculated as following function:

$$b = \frac{1}{N_{SV}} \left\{ \sum_{0 < a_i < C} \left[y_i - \sum_{x_j \in SV} (a_j - a_j^*) K(x_j, x_i) - \varepsilon \right] + \sum_{0 < a_i^* < C} \left[y_i - \sum_{x_j \in SV} (a_j - a_j^*) K(x_j, x_i) + \varepsilon \right] \right\} \quad (5)$$

where N_{SV} is the number of support vectors.

2.1.3. Multivariate Adaptive Regression Spline

Multivariate adaptive regression spline (MARS) is a flexible nonparametric modeling tool that was popularized by Friedman (1991) for solving regression-type problems [33]. It does not require an assumption about the particular type of relationship between outcome variable and predictor variables [34,35]. Instead, MARS use coefficients and basic functions that are decided entirely by regression data to construct this relation. The MARS model can be described as follows:

$$f(x) = \beta_0 + \sum_{i=1}^{k} c_i B_i(x) \quad (6)$$

Here, the summation $f(x)$ is over the k non-constant terms in the model. β_0 is the intercept parameter. Each c_i is a constant coefficient, $B_i(x)$ is a basic function, which takes one of following forms: a constant 1, a hinge function or two or more hinge functions. A hinge function (Figure 3) has the form $\max(0, x - c)$ or $\max(0, c - x)$, c is a constant, called the knot.

The first step to build a MARS model is partitioning the training data set into several splines in equivalent interval basis [33]. Most importantly, the model ensures that there is adequate data in a subgroup in order to avoid over-fitting using the shortest distance between neighboring knots [36,37]. The forward pass and the backward pass are two important phases in building a MARS model. To find the best subset, we use generalized cross validation (GCV) to compare the performance of sub-models. Better sub-model has lower GCV values. GCV can be calculated as follows:

$$GCV(N) = \left(\frac{1}{n} \frac{\sum_{m=1}^{N} \left(y_i - \hat{y}_i \right)^2}{\left(1 - \frac{C(M)}{n} \right)^2} \right) \quad (7)$$

Here, the n was the number of observations, y_i was the response value of object i, $\hat{y}_i$ was the predicted response value of object i, and $C(M)$ was the penalty factor.
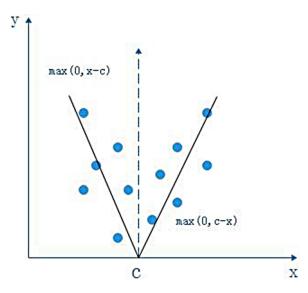

Figure 3. The hinge functions and knot location in the multivariate adaptive regression spline (MARS) model.

2.2. Experimental and Simulated Data

2.2.1. Eddy Covariance Observations

The machine learning algorithms were validated and evaluated using ground-measured flux data. All of the turbulent flux observations were measured by the EC method and the data cover the period from 2000 to 2007. The data were collected from 85 EC flux tower sites and were provided by AmeriFlux. These flux towers are located mainly in North America (50°–170°W, 10°–70°N, Figure 4). The flux tower sites cover five major global land-surface biomes: cropland (CRO; 7 sites), deciduous broadleaf forest (DBF; 11 sites), evergreen needleleaf forest (ENF; 41 sites), grass and other types (GRA; 17 sites), and shrubland (SHR; 9 sites). These data sets include half-hourly or hourly ground-measured incident Rs, RH, Ta, diurnal air temperature range (DT), Ws, vapor pressure (e), sensible heat flux (H), surface net radiation (R_n), ground heat flux (G), and LE. We acquired daily data from aggregated half-hourly or hourly data without using additional quality control [38–40] and removed the zero values. For the unclosed energy problem, we used the following method that was developed by Twine et al. [41] to correct the LE for all flux towers:

$$LE = (R_n - G)/(LE_{ori} + H_{ori}) \times LE_{ori} \tag{8}$$

where LE is the corrected latent heat, H_{ori} and LE_{ori} are the uncorrected sensible heat flux and latent heat, respectively.
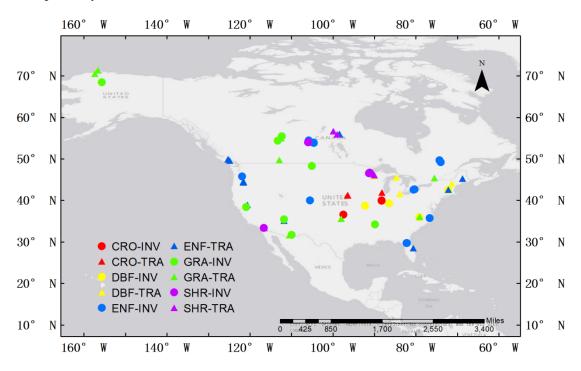

Figure 4. Location of the 85 eddy covariance flux towers used in this study. INV means the data of this site were used to inverse, TRA means that the data of this site were used for training.

2.2.2. MODIS and MERRA Data

We used the NDVI values as the input variables and the daily NDVI values were temporally interpolated from the 16 day MODIS NDVI (MOD13A2) product [42] at 1-km spatial resolution using linear interpolation [43]. When 16 day NDVI data are missing or not available, we used the closest reliable 16 day values to replace the missing or original data.

We used the daily R_n, RH, Ta, and Ws products with a spatial resolution of $1/2° \times 2/3°$ from MERRA data provided by National Aeronautics and Space Administration (NASA) to evaluate the

performance of all the LE algorithms for all the flux tower sites in this study. To match MODIS pixels, we used the method proposed by Zhao et al. [43] which is a spatial interpolation method using a cosine function to interpolate coarse-resolution MERRA data to 1 km^2 pixels. Theoretically, this method uses the four MERRA cells surrounding a given pixel to remove sharp changes from one side of a MERRA boundary to the other to improve the accuracy of MERRA data for each 1 km pixel.

To map the final LE product with 0.05° spatial resolution using three machine learning algorithms, we used Collection 5 MODIS NDVI (MOD13C1: CMG, 0.05°) [44] and the daily 0.05° MERRA data interpolated using the method proposed by Zhao et al. [43] from the original MERRA data. To evaluate the performance of three machine learning algorithms, we used MODIS LE product (MOD16A2) [7] with 0.05° spatial resolution.

2.2.3. Criteria of Evaluation

We used three different statistical criteria to evaluate the performance of the machine learning algorithms that were used in this study: coefficient of determination (R^2); the average deviation of ground-measured LE value and estimated LE value (Bias); and, root mean square error (RMSE) [45,46].

R^2 can be defined as the square of correlation coefficient R. The difference between it and the correlation coefficient is to remove $|R| = 0$ and 1, as it can prevent an exaggerated interpretation of the correlation coefficient. The closer that R^2 is to 1, the more relevant the observations and estimates. The following equation was used to calculated coefficient of determination:

$$R^2 = \left(\frac{\sum_{i=1}^n (X_i - \overline{X})(Y_i - \overline{Y})}{\sqrt{\sum_{i=1}^n (X_i - \overline{X})^2 (Y_i - \overline{Y})^2}} \right)^2 \tag{9}$$

where X_i and Y_i are the observed and estimated values, respectively, $\overline{X}$ and $\overline{Y}$ are the average of X_i and Y_i, and n is the total number of the data.

Bias is the average value of the absolute of differences between the predicted and observed LE values. A low bias indicates a good model performance. The bias can be calculated from the following equation:

$$Bias = \frac{\sum_{i=1}^n (X_i - Y_i)}{n} \tag{10}$$

where X_i and Y_i are the observed and estimated values, respectively, and n is the total number of the data.

RMSE can be defined as the square root of the average value of the differences between the predicted and observed LE values. A perfect match between the two values would yield RMSE = 0, and it implies the best performance model. The larger RMSE is, the lower model performance. RMSE can be calculated from the following function:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (X_i - Y_i)^2} \tag{11}$$

where X_i and Y_i are the observed and estimated values, respectively, and n is the total number of the data.

2.2.4. Experimental Setup

For each PFT, we chose half of the sites to train, the remaining sites were used to reverse. As for PFT CRO, three sites were for training and four for reversing. For DBF, five sites were for training and six for reversing. ENF/GRA/SHR has 21/9/5 sites for training and 20/8/4 for reversing, respectively. We put the same PFT data together for the training and validation of three machine learning algorithms. For the training step, we chose Ta, RH, Ws, and NDVI as input variables and the ratio of ground-measured LE and R_n as output variables, and built the ANN/SVM/MARS model. Then, we put the input variables

of testing data to the model and got the predicted LE/R_n. To get the predicted LE, we multiplied the results by the R_n of the testing data. Finally, we evaluated the predicted LE with the ground-measured LE of the testing data. Meteorology data including Ta, RH, Ws, and R_n were acquired from EC flux tower sites and MERRA data, respectively.

For the ANN algorithm, we trained, validated, and tested the data by using feed-forward artificial neural networks (FFNN), which is the most commonly used ANN model. Levenberg-Marquardt algorithms were applied to train the network and a Sigmoid function was used since it gave the best results of the model. The number of hidden layers and neurons in hidden layers was determined by trial and error, and we finally decided that the number of hidden layers was 10. For the SVM algorithm, we implement the SVM on R platform using R language with package e1071, which provided an interface connects libsvm that was based on the C++ language, developed by Chih Chung-Chang and Chih-Jen Lin [47,48]. The process of modeling training and inversion was similar to the ANN process. The present study built the MARS model and estimated testing data LE using ARESlab (Adaptive Regression Splines toolbox) in MATLAB, and set the maxfuncs and maxinteractions to 15 and 4, respectively.

3. Results

3.1. Algorithms Evaluation Based on Specific Site Data

At the flux tower site scale, the three machine learning algorithms exhibited substantial differences in LE modeling for different PFTs. The statistical parameters of training and validation of each model are provided in Table 1.

Table 1. Training and validation statistics for the different algorithms based on specific site. All of the coefficient of determination (R^2) values are significant with a 99% confidence.

PFT	Algorithms	R^2 Train	R^2 Test	Bias (W/m^2) Train	Bias (W/m^2) Test	RMSE (W/m^2) Train	RMSE (W/m^2) Test
CRO	ANN	0.83	0.81	0.4	0.6	16.47	21.86
	SVM	0.83	0.8	−0.54	1.1	17.92	24.37
	MARS	0.81	0.78	1.15	2.7	19.23	25.38
DBF	ANN	0.89	0.88	−0.08	0.7	14.48	17.17
	SVM	0.89	0.84	1.19	1.6	17.25	17.91
	MARS	0.88	0.85	−0.8	−1.6	17.15	18.19
ENF	ANN	0.82	0.77	0.55	1.05	12.15	17.74
	SVM	0.83	0.72	−1.35	−3.47	13.69	20.02
	MARS	0.81	0.72	−1.88	−2.36	14.63	20.05
GRA	ANN	0.83	0.8	1.1	3.55	15.05	16.14
	SVM	0.84	0.62	5.3	8.37	15.81	24.76
	MARS	0.83	0.6	4.12	6.35	16.45	25.39
SHR	ANN	0.72	0.7	0.13	0.25	13.05	13.35
	SVM	0.7	0.64	−0.08	−0.17	14.24	15.22
	MARS	0.67	0.6	0.2	−0.71	14.78	15.95

Figure 5 shows that the ANN algorithm performed best among the three algorithms for each PFT, while SVM performed better than MARS for CRO, GRA, and SHR, and made a poor performance for DBF land cover type. For CRO, the validation result showed that the R^2 was very close among the three algorithms, the highest R^2 of ANN was 0.81, while the lowest R^2 was 0.03 lower than that and each of the three algorithms had a large RMSE greater than 20 W/m^2. Meanwhile, ANN had a lower bias than SVM and MARS. ANN for DBF had the best performance with the highest R^2 (0.89/0.88) ($p < 0.01$) and lower RMSE (14.48/17.17 W/m^2) in comparison to other PFTs for training and validation, respectively.

This may be caused by the characteristic of the DBF. Although the R^2 of SHR was the lowest among the five PFTs for validation results, the bias and the RMSE of SHR were lower than any other PFTs, with the lowest bias and RMSE reaches of -0.17 and 13.35 W/m^2. For ENF and GRA sites, the lowest RMSE of the estimated LE versus ground observations was approximately 17.74 and 16.14 W/m^2, respectively, and the R^2 was approximately less than 0.8. Especially, for GRA sites, the ANN algorithm that was driven by tower-specific meteorological variables and satellite-based NDVI had an absolutely higher R^2 and a lower bias and RMSE than the other machine learning algorithms tested.

Figure 5. Bar graphs of the training and validation statistics (R^2, Bias and root mean square error (RMSE)) of three algorithms driven by tower-specific meteorology for five PFTs at the 85 flux tower site. All R^2 values are significant with a 99% confidence.

Figure 6 shows that the three algorithms exhibited most features of measured LE seasonality in the observed inversion data for different PFTs. We randomly selected an inversion site that has at least three years of complete ground-measured data for each PFT, and then chose three years of complete ground-measured data and the corresponding estimated LE data using three machine learning algorithms. We plotted the eight-day average LE to replace daily LE data. Five sites were US-Ne3 (CRO), US-WCr (DBF), CA-Qcu (ENF), US-FPe (GRA), and US-SO4 (SHR). We compared the estimated LE with the observed LE in 2002–2004 except at US-SO4, for which we used the period 2004–2006. In comparison to the other two algorithms, the ANN algorithm produced seasonal LE variations that were closest to the observed LE. Estimated LE showed great consistency, with observed LE in lower values for all of the PFTs in Figure 6, but for GRA land cover type, the ANN and the SVM kept the same trend, while the results of the MARS were absolutely lower than result of the ANN, the SVM and observed LE in high value area. We also found some abnormal observed LE value in 2004 for SHR, which may lead to the inaccuracy of estimation results.
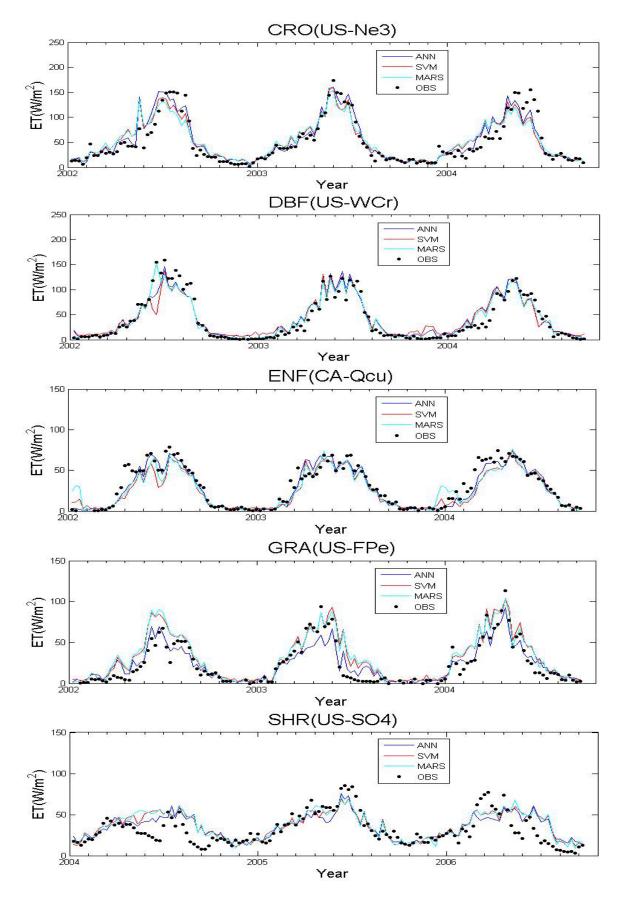
Figure 6. Examples of the eight-day terrestrial latent heat flux (LE) average as measured and estimated using different machine learning algorithms for the different PFTs.

3.2. Algorithms Evaluation Based on MERRA Data

Table 2 shows the statistical parameters of training and validation of each model that is driven by MERRA meteorology. When compared to the results that are driven by tower-specific meteorology, results driven by MERRA meteorology showed poor performance in all PFTs. This may be related to the uncertainty of MERRA data.

Figure 7 shows the training and validation statistics R^2, bias, and RMSE of the three algorithms that are driven by MERRA meteorology for five PFTs. The value of R^2 was lower and RMSE was higher when compared to the results that are driven by tower-specific meteorology. With similar results being shown in Figure 5, the ANN demonstrated the best performance among the three machine learning algorithms at both the training and validation stages. The biggest difference, however, was that the ANN performed absolutely higher R^2 than any of the other algorithms in Figure 7, while the R^2 is very close among the three algorithms in Figure 5 for CRO and DBF. Since the only difference in input was meteorology data, the most likely reason that the ANN performed better than than SVM and MARS when using inaccurate input variables (besides MERRA meteorology data being more inaccurate than the tower-specific meteorology data) has to do with the ANN's own characteristics. As for bias, ANN performed much better than the others, while SVM showed the poorest performance for most PFTs. As for RMSE, results of SVM may be a little better than results of MARS for some PFTs, but on the whole, their performances were similar.
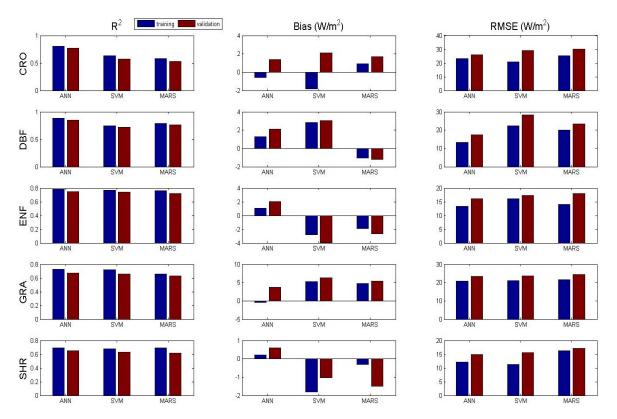

Figure 7. Bar graphs of the training and validation statistics (R^2, Bias and RMSE) of three algorithms driven by MERRA meteorology for five PFTs at the 85 flux tower sites. All of the R^2 values are significant with a 99% confidence.

Table 2. Training and validation statistics for the different algorithms based on MERRA data. All R^2 values are significant with a 99% confidence.

PFT	Algorithms	R^2 Train	R^2 Test	Bias (W/m^2) Train	Bias (W/m^2) Test	RMSE (W/m^2) Train	RMSE (W/m^2) Test
	ANN	0.81	0.77	−0.6	1.4	23.2	25.9
CRO	*SVM*	0.63	0.57	−1.8	2.1	20.83	29.22
	MARS	0.58	0.53	0.9	1.7	25.35	30.33
	ANN	0.88	0.85	1.3	2.1	13.25	17.45
DBF	*SVM*	0.75	0.72	2.82	3.06	22.42	28.39
	MARS	0.79	0.76	−1.03	−1.19	19.92	23.44
	ANN	0.78	0.75	1.06	2.05	13.34	16.18
ENF	*SVM*	0.77	0.74	−2.81	−3.95	16.13	17.31
	MARS	0.76	0.72	−1.9	−2.65	14.12	18
	ANN	0.73	0.67	−0.35	3.75	20.85	23.3
GRA	*SVM*	0.72	0.66	5.27	6.37	21.01	23.68
	MARS	0.66	0.63	4.8	5.35	21.68	24.43
	ANN	0.69	0.65	0.22	0.61	12.08	14.85
SHR	*SVM*	0.68	0.63	−1.81	−1.04	11.22	15.63
	MARS	0.69	0.62	−0.31	−1.5	16.21	17.1

3.3. Mapping of Terrestrial LE Using Three Machine Learning Algorithms

We applied the ANN algorithm, the SVM algorithm, and the MARS algorithm with MERRA meteorological data and MODIS product to estimate annual LE at a 0.05° spatial resolution from 2002–2004 over North America. Figure 8 shows maps of annual terrestrial LE averaged for 2002–2004 over North America.
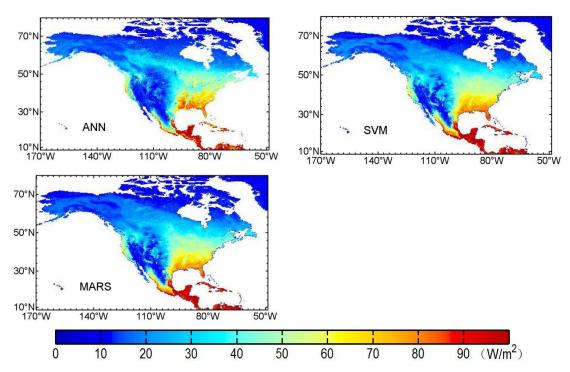

Figure 8. The map of mean annual terrestrial LE from 2002 to 2004 at a spatial resolution of 0.05° using three machine learning algorithms driven by MERRA meteorology over North America.

All of the algorithms yielded high annual LE over the east and west coasts of the United States. In the low latitudes, the highest annual LE was approximately larger than 90 W/m^2, appearing on the area with latitudes lower than 20°N. The biggest difference between the three algorithms was the

ANN estimated higher LE values on the east coast of the United States and lower LE values in areas of the western United States when compared to the SVM and the MARS, which returned similar results.

4. Discussion

4.1. Performance of the Machine Learning Algorithms

Our study, based on the estimated daily LE at 85 EC flux tower sites using different machine learning algorithms that were driven by tower-specific and MERRA meteorology, illustrated that the ANN algorithm yielded the best LE estimate. The ANN had the highest R^2 (0.81 and 0.70) ($p < 0.01$) and lowest RMSE (17.33 and 20.22 W/m^2) in comparison to the SVM and MARS algorithms for tower-specific and MERRA meteorological data, respectively (Figures 9 and 10). The results were consistent with the findings of Lu and Zhuang, 2010, which used remote sensing data from the MODIS, meteorological and eddy flux data, and an ANN technique to develop a daily product for the period of 2004–2005 for the conterminous U.S. [49]. They found that the ANN predicted daily LE well (R^2 = 0.52–0.86).
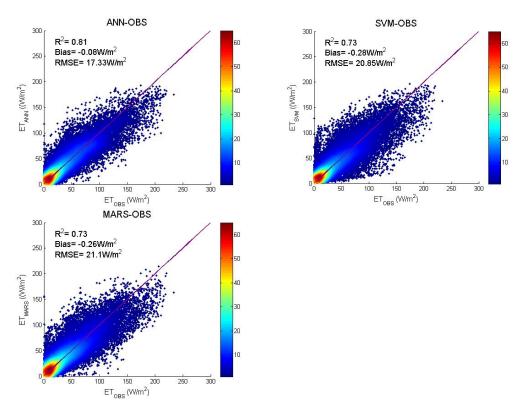

Figure 9. Comparison of daily LE observations for all 85 flux tower sites and LE estimates using different machine learning algorithms driven by tower-specific meteorology.

Validation for 85 EC flux tower sites also indicated that the three machine learning algorithms used were reliable and robust for major land cover types in North America. Figures 9 and 10 both show that the ANN algorithm had no significant LE bias and yielded the closest LE to tower flux data relative to the SVM and MARS algorithms. We also found large inter-biome differences, with better performance at DBF and CRO sites. Several studies have revealed that some algorithms that exhibit strong seasonality for vegetation indices, such as NDVI for accurate capture of information on seasonal changes in vegetation can yield considerably better LE estimates for seasonal vegetation types such as DBF [10,50–54]. Thus, NDVI as an input variable determined the accuracy of LE quantification. All of the machine learning algorithms yielded poor LE estimation for GRA and SHR sites, which may be attributable in part to the fact that seasonal GRA and SHR variation was less evident when

satellite-derived vegetation indices (e.g., NDVI) saturate asymptotically. Performance differences among machine learning algorithms may be caused by the different scope for each algorithm, as well as the quality and quantity of input variables. The SVM and MARS highlight global rather than local optima, while the ANN may ensure local optimization and leads to a better performance when compared to SVM and MARS [24,27,30,55]. When using MERRA meteorology data as input variables, SVM and MARS showed even worse results, while ANN was slightly worse as compared with results using tower-specific meteorology data as input variables. On the other hand, some studies have indicated that ANN can be considered more suitable to serve as a tool to estimate LE when input meteorological variables are insufficient [29]. Thus, we may conclude that ANN performs better than SVM and MARS, especially when there are insufficient input variables or the accuracy of them was not very high.
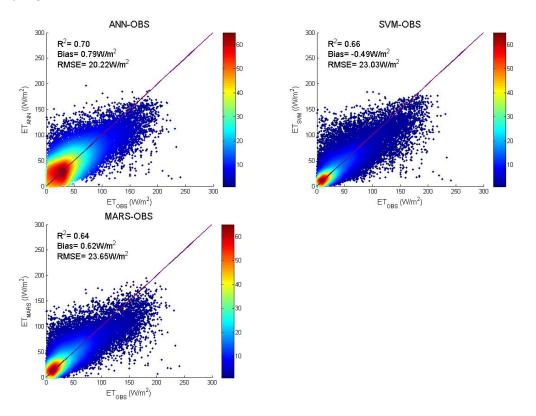

Figure 10. Comparison of daily LE observations for all 85 flux tower sites and LE estimates using different machine learning algorithms driven by MERRA meteorology.

Several previous studies have shown that spatial scale mismatch among different data sources, model input errors, and the limitations of the machine learning algorithms itself all affect the accuracy of LE estimation [7,10,56]. We used MERRA products with a spatial resolution of $1/2° \times 1/3°$ and MODIS products with a resolution of 1 km, their resolution being greater than the footprint for field measurements, which is usually several hundred meters [57,58]. This is one of the reasons that estimated LE using tower meteorology as input data performs better than the estimated LE using MERRA meteorology as input data. Such representation of the field measurement footprint may also lead to bias in the machine learning algorithms. In addition, the accuracy of model input variables can also influence the accuracy of the LE estimates. Many studies have demonstrated that there are large errors in the MERRA meteorology and MERRA data tend to underestimate R_n at high values when compared with ground measurements [49,58]. Recent studies have also revealed errors in MODIS NDVI when compared with ground measurements [49,57].

4.2. Comparison between Different LE Products

When compared with the SVM and the MARS algorithms, the ANN algorithm yielded lower annual terrestrial LE in central and southern North America and higher LE in the northern and eastern portions of this study area (Figure 11a,b). The LE values that were estimated using MARS algorithm were lower than the LE values estimated using the SVM algorithm in most areas, but overall, the difference between them was small and less than 10 W/m^2 (Figure 11c). The relevant differences between the flux estimations at the Mexican highlands, which are surrounded by Sierra Madre, may be mostly caused by the land surface heterogeneity and the different plant functional types. The PFT of the highlands is mainly grassland, while coastal and southeastern parts are covered with rainforest. The spatial differences among three machine learning algorithms may be mostly caused by different PFTs. The ANN showed higher LE for ENF and DBF, which are mostly located at the coasts of North America. In the center and west of the study area, due to the relatively high elevation and its own geographic and geomorphic conditions, the PFTs mostly contain SHR or GRA, so the ANN yielded lower LE than SVM and MARS. However, the results that are presented by SVM and MARS show no obvious difference for different PFTs.

Figure 12 shows spatial differences in annual terrestrial LE (2002–2004) between the MODIS LE product and the LE product using three machine learning algorithms. Relative to MODIS LE product, the ANN, SVM, and MARS results yielded lower LE in the north and central study area, and higher LE on the east coast and in the northeast of North America. This may be caused by the precision and quantity of the test data sample, as well as the characteristics of the machine learning algorithms, so the simulation value of machine learning LE is lower than the value of the MODIS LE product over a high LE area. This finding is consistent with other studies that showed that the MODIS LE product overestimated LE at high LE areas [23]. The difference between them was less than 20 W/m^2 in most areas. Given the accuracy of MERRA meteorology data and the MODIS LE product, we can conclude that machine learning algorithms are applicable for terrestrial LE mapping and the inversion results have a relatively small gap when compared to the MODIS LE product.
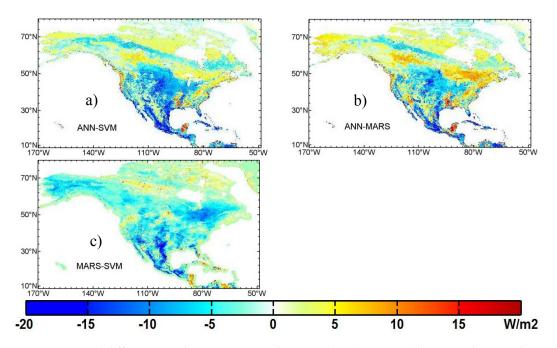

Figure 11. Spatial differences in the average annual terrestrial LE (2002–2004) between three machine learning algorithms. .
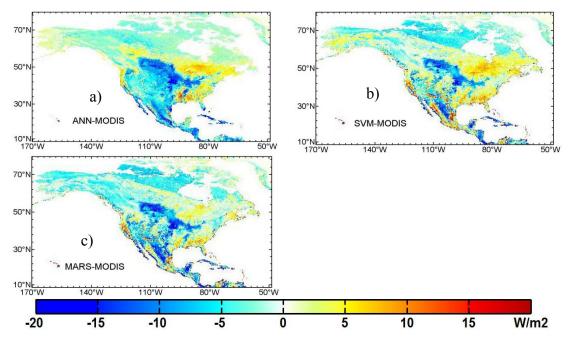
Figure 12. Spatial differences in the average annual terrestrial LE (2002–2004) between MODIS LE product and LE product using three machine learning algorithms.

4.3. Limitations and Recommendations for Future Research

Although the ANN leads to better performance when compared to other machine learning algorithms, it has several limitations. First, ANN requires a relatively long processing time to train a model, especially for ENF and GRA, the input data of which are inaccurate, and thus require more learning cycles. Second, the ANN model may trap in local optimization, leading to worse generalization performance. Further, all machine learning algorithms behave relatively unpredictably when used with input ground-measured LE deviating from those presented during the training stage [24,59]. Most importantly, we evaluated and validated LE performance using the same ground-measure LE data from EC sites. The machine learning algorithms do not possess the useful information to directly deliver additional confidence LE maps, and we cannot ensure the accuracy of these ground-measure LE data [60].

To make the training samples more applicable over North America, we should add samples from other PFTs, such as evergreen broadleaf forest and deciduous needleleaf forest [59]. The EC data of these specific PFTs is few over North America, however. We may develop the machine learning algorithms coupled with semi-empirical and physical methods that do not require training samples in the future to improve the terrestrial LE at more different PFTs.

5. Conclusions

The primary objective of this paper was to demonstrate the usage of three machine learning algorithms to estimate terrestrial LE for different PFTs over North America. We used tower-specific and MERRA meteorology data and MODIS NDVI product as input data and the ratio of observed LE and meteorology parameter R_n as output data to build training models. All three machine learning algorithms proved to be reliable and robust for major land cover types in North America but the ANN algorithm performed better than the SVM and MARS algorithms based on specific site and MERRA data.

The results revealed that the ANN algorithm was more efficient in accounting for the nonlinear relationship between climatic variables, vegetation indices, and the corresponding LE. We found that ANN performed better in most land cover types, for except ENF and GRA, when using MERRA

meteorology data as input data in the study area. The validation results for machine learning algorithms driven by tower-specific meteorology and MERRA meteorology showed that the results of SVM were similar to the results of MARS. For DBF sites, the MARS has a higher R^2 and a lower RMSE than SVM, for other sites, SVM perform a little better than MARS.

We concluded that it is reasonable to use machine learning algorithms to estimate LE by building relationships between input variables and outputs (LE). The accuracy of input variables, such as meteorology data and MODIS NDVI, can also influence the accuracy of the LE estimates. It is also applicable to map the terrestrial LE in different regions for different PFTs using any of these three machine learning algorithms; although, the performance of ANN was the best, the differences among them were small. The inversion results have a relatively small gap compared with MODIS LE product, with the difference between them being less than 20 W/m^2 in most areas, and machine learning algorithms performing better at the area where MODIS LE product overestimated LE.

Author Contributions: Xuanyu Wang and Yunjun Yao prepared the manuscript. Shaohua Zhao, Kun Jia, Xiaotong Zhang, Yuhu Zhang, Lilin Zhang, Jia Xu and Xiaowei Chen contributed to the discussion.

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 No 56; Food and Agriculture Organization of the United Nations (FAO): Rome, Italy, 1998.

2. Liang, S.; Wang, K.; Zhang, X.; Wild, M. Review on estimation of land surface radiation and energy budgets from ground measurement, remote sensing and model simulations. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2010**, *3*, 225–240. [CrossRef]

3. Wang, K.; Dickinson, R.E. A review of global terrestrial evapotranspiration: Observation, modeling, climatology, and climatic variability. *Rev. Geophys.* **2012**, *50*, RG2005. [CrossRef]

4. Kumar, M.; Raghuwanshi, N.S.; Singh, R.; Wallender, W.W.; Pruitt, W.O. Estimating evapotranspiration using artificial neural network. *J. Irrig. Drain. Eng.* **2002**, *128*, 224–233. [CrossRef]

5. Los, S.O.; Collatz, G.J.; Sellers, P.J.; Malmström, C.M.; Pollack, N.H.; Defries, R.S.; Bounoua, L.; Parris, M.T.; Tucker, C.J.; Dazlich, D.A. A global 9-yr biophysical land surface dataset from noaa avhrr data. *J. Hydrometeorol.* **2000**, *1*, 183–199. [CrossRef]

6. Jin, Y.; Randerson, J.T.; Goulden, M.L. Continental-scale net radiation and evapotranspiration estimated using modis satellite observations. *Remote Sens. Environ.* **2011**, *115*, 2302–2319. [CrossRef]

7. Mu, Q.; Zhao, M.; Running, S.W. Improvements to a modis global terrestrial evapotranspiration algorithm. *Remote Sens. Environ.* **2011**, *115*, 1781–1800. [CrossRef]

8. 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]

9. Kustas, W.P.; Norman, J.M. Use of remote sensing for evapotranspiration monitoring over land surfaces. *Hydrol. Sci. J.* **1996**, *41*, 495–516. [CrossRef]

10. Mu, Q.; Heinsch, F.A.; Zhao, M.; Running, S.W. Development of a global evapotranspiration algorithm based on modis and global meteorology data. *Remote Sens. Environ.* **2007**, *111*, 519–536. [CrossRef]

11. Monteith, J.L. Evaporation and environment. *Symp. Soc. Exp. Biol.* **1965**, *19*, 205. [PubMed]

12. Priestley, C.H.B.; Taylor, R.J. On the assessment of surface heat flux and evaporation using large-scale parameters. *Mon. Weather Rev.* **2009**, *100*, 81–92. [CrossRef]

13. Cleugh, H.A.; Leuning, R.; Mu, Q.; Running, S.W. Regional evaporation estimates from flux tower and modis satellite data. *Remote Sens. Environ.* **2007**, *106*, 285–304. [CrossRef]

14. Wang, K.; Wang, P.; Li, Z.; Cribb, M.; Sparrow, M. A simple method to estimate actual evapotranspiration from a combination of net radiation, vegetation index, and temperature. *J. Geophys. Res. Atmos.* **2007**, *112*, D15107. [CrossRef]

15. Jung, M.; Reichstein, M.; Ciais, P.; Seneviratne, S.I.; Sheffield, J.; Goulden, M.L.; Bonan, G.; Cescatti, A.; Chen, J.Q.; Jeu, R.D. Recent decline in the global land evapotranspiration trend due to limited moisture supply. *Nature* **2010**, *467*, 951–954. [CrossRef] [PubMed]

16. Yao, Y.; Liang, S.; Qin, Q.; Wang, K.; Zhao, S. Monitoring global land surface drought based on a hybrid evapotranspiration model. *Int. J. Appl. Earth Obs. Geoinf.* **2011**, *13*, 447–457. [CrossRef]

17. Qin, Q. Simple method to determine the priestley-taylor parameter for evapotranspiration estimation using albedo-vi triangular space from modis data. *J. Appl. Remote Sens.* **2011**, *5*, 3505.

18. Xu, T.; Liang, S.; Liu, S. Estimating turbulent fluxes through assimilation of geostationary operational environmental satellites data using ensemble kalman filter. *J. Geophys. Res. Atmos.* **2011**, *116*, 644. [CrossRef]

19. Xu, T.; Liu, S.; Liang, S.; Qin, J. Improving predictions of water and heat fluxes by assimilating modis land surface temperature products into the common land model. *J. Hydrometeorol.* **2010**, *12*, 227–244. [CrossRef]

20. Jiménez, C.; Prigent, C.; Mueller, B.; Seneviratne, S.I.; Mccabe, M.F.; Wood, E.F.; Rossow, W.B.; Balsamo, G.; Betts, A.K.; Dirmeyer, P.A. Global intercomparison of 12 land surface heat flux estimates. *J. Geophys. Res. Atmos.* **2011**, *116*, 3–25. [CrossRef]

21. Kim, H.W.; Hwang, K.; Mu, Q.; Lee, S.O.; Choi, M. Validation of modis 16 global terrestrial evapotranspiration products in various climates and land cover types in asia. *KSCE J. Civ. Eng.* **2012**, *16*, 229–238. [CrossRef]

22. Sheffield, J.; Wood, E.F. Projected changes in drought occurrence under future global warming from multi-model, multi-scenario, ipcc ar4 simulations. *Clim. Dyn.* **2008**, *31*, 79–105. [CrossRef]

23. Bruton, J.M.; Mcclendon, R.W.; Hoogenboom, G. Estimating daily pan evaporation with artificial neural networks. *Am. Soc. Agric. Biol. Eng.* **2000**, *43*, 491–496. [CrossRef]

24. El-Shafie, A.; Alsulami, H.M.; Jahanbani, H.; Najah, A. Multi-lead ahead prediction model of reference evapotranspiration utilizing ann with ensemble procedure. *Stoch. Environ. Res. Risk Assess.* **2013**, *27*, 1423–1440. [CrossRef]

25. Shrestha, N.K.; Shukla, S. Support vector machine based modeling of evapotranspiration using hydro-climatic variables in a sub-tropical environment. *Bioresour. Technol.* **2013**, *128*, 351. [CrossRef]

26. Deo, R.C.; Samui, P.; Kim, D. Estimation of monthly evaporative loss using relevance vector machine, extreme learning machine and multivariate adaptive regression spline models. *Stoch. Environ. Res. Risk Assess.* **2016**, *30*, 1769–1784. [CrossRef]

27. Yang, F.; White, M.A.; Michaelis, A.R.; Ichii, K.; Hashimoto, H.; Votava, P.; Zhu, A.X.; Nemani, R.R. Prediction of continental-scale evapotranspiration by combining modis and ameriflux data through support vector machine. *IEEE Trans. Geosci. Remote Sens.* **2006**, *44*, 3452–3461. [CrossRef]

28. Adnan, M.; Latif, M.A.; Abaid-ur-Rehman; Nazir, M. Estimating evapotranspiration using machine learning techniques. *Int. J. Adv. Comput. Sci. Appl.* **2017**, *8*, 108–113.

29. Tabari, H.; Marofi, S.; Sabziparvar, A.A. Estimation of daily pan evaporation using artificial neural network and multivariate non-linear regression. *Irrig. Sci.* **2010**, *28*, 399–406. [CrossRef]

30. Vapnik, V.N. *The Nature of Statistical Learning Theory*; Springer: New York, NY, USA, 1995; pp. 988–999.

31. Nurmemet, I.; Ghulam, A.; Tiyip, T.; Elkadiri, R.; Ding, J.L.; Maimaitiyiming, M.; Abliz, A.; Sawut, M.; Zhang, F.; Abliz, A. Monitoring soil salinization in keriya river basin, northwestern China using passive reflective and active microwave remote sensing data. *Remote Sens.* **2015**, *7*, 8803–8829. [CrossRef]

32. Dibike, Y.B.; Velickov, S.; Solomatine, D.; Abbott, M.B. Model induction with support vector machines: Introduction and applications. *J. Comput. Civ. Eng.* **2001**, *15*, 208–216. [CrossRef]

33. Friedman, J.H. Multivariate adaptive regression splines. *Ann. Stat.* **1991**, *19*, 1–67. [CrossRef]

34. Butte, N.F.; Wong, W.W.; Adolph, A.L.; Puyau, M.R.; Vohra, F.A.; Zakeri, I.F. Validation of cross-sectional time series and multivariate adaptive regression splines models for the prediction of energy expenditure in children and adolescents using doubly labeled water. *J. Nutr.* **2010**, *140*, 1516. [CrossRef] [PubMed]

35. Veaux, R.D.D.; Ungar, L.H. *Multicollinearity: A Tale of Two Nonparametric Regressions*; Springer: New York, NY, USA, 2007; pp. 393–402.

36. Adamowski, J.; Chan, H.; Prasher, S.; Ozga-Zielinski, B.; Sliusarieva, A. Comparison of multiple linear and nonlinear regression, autoregressive integrated moving average, artificial neural network, and wavelet artificial neural network methods for urban water demand forecasting in montreal, Canada. *Water Resour. Res.* **2012**, *48*, 1528. [CrossRef]

37. Sephton, P. *Forecasting Recessions: Can We Do Better on Mars?* Federal Reserve Bank of St. Louis: Washington, DC, USA, 2001; Volume 83.

38. Liu, S.M.; Xu, Z.W.; Wang, W.Z.; Jia, Z.Z.; Zhu, M.J.; Bai, J.; Wang, J.M. A comparison of eddy-covariance and large aperture scintillometer measurements with respect to the energy balance closure problem. *Hydrol. Earth Syst. Sci.* **2011**, *15*, 1291–1306. [CrossRef]

39. Jia, Z.; Liu, S.; Xu, Z.; Chen, Y.; Zhu, M. Validation of remotely sensed evapotranspiration over the hai river basin, China. *J. Geophys. Res. Atmos.* **2012**, *117*, 13113. [CrossRef]

40. Xu, Z.; Liu, S.; Li, X.; Shi, S.; Wang, J.; Zhu, Z.; Xu, T.; Wang, W.; Ma, M. Intercomparison of surface energy flux measurement systems used during the HiWATER-MUSOEXE. *J. Geophys. Res. Atmos.* **2013**, *118*, 13140–13157. [CrossRef]

41. Twine, T.E.; Kustas, W.P.; Norman, J.M.; Cook, D.R.; Houser, P.R.; Teyers, T.P.; Prueger, J.H.; Starks, P.J.; Wesely, M.L. Correcting eddy-covariance flux underestimates over a grassland. *Agric. For. Meteorol.* **2000**, *103*, 279–300. [CrossRef]

42. Huete, A.; Didan, K.; Miura, T.; Rodriguez, E.P.; Gao, X.; Ferreira, L.G. Overview of the radiometric and biophysical performance of the modis vegetation indices. *Remote Sens. Environ.* **2002**, *83*, 195–213. [CrossRef]

43. Zhao, M.; Heinsch, F.A.; Nemani, R.R.; Running, S.W. Improvements of the modis terrestrial gross and net primary production global data set. *Remote Sens. Environ.* **2005**, *95*, 164–176. [CrossRef]

44. Friedl, M.A.; Mciver, D.K.; Hodges, J.C.F.; Zhang, X.Y.; Muchoney, D.; Strahler, A.H.; Woodcock, C.E.; Gopal, S.; Schneider, A.; Cooper, A. Global land cover mapping from modis: Algorithms and early results. *Remote Sens. Environ.* **2002**, *83*, 287–302. [CrossRef]

45. Gandomi, A.H.; Roke, D.A. Intelligent formulation of structural engineering systems. In Proceedings of the Seventh MIT Conference on Computational Fluid and Solid Mechanics—Focus: Multipgysics & Multiscale, Cambridge, MA, USA, 12–14 June 2013.

46. Cristau, H.-J.; Cellier, P.P.; Spindler, J.F.; Taillefer, M. Highly efficient and mild copper-catalyzed N- and C-arylations with aryl bromides and iodides. *Chemistry* **2004**, *10*, 5607–5622. [CrossRef] [PubMed]

47. Chang, C.C.; Lin, C.J. *Libsvm: A Library for Support Vector Machines*; ACM: New York, NY, USA, 2011; pp. 1–27.

48. Meyer, D. Support vector machines the interface to libsvm in package e1071. *R News* **2013**, *1*, 1–3.

49. Lu, X.L.; Zhuang, Q.L. Evaluating evapotranspiration and water-use efficiency of terrestrial ecosystems in the conterminous united states using modis and ameriflux data. *Remote Sens. Environ.* **2010**, *114*, 1924–1939. [CrossRef]

50. Fisher, J.B.; Tu, K.P.; Baldocchi, D.D. Global estimates of the land–atmosphere water flux based on monthly avhrr and islscp-ii data, validated at 16 fluxnet sites. *Remote Sens. Environ.* **2008**, *112*, 901–919. [CrossRef]

51. Vinukollu, R.K.; Wood, E.F.; Ferguson, C.R.; Fisher, J.B. Global estimates of evapotranspiration for climate studies using multi-sensor remote sensing data: Evaluation of three process-based approaches. *Remote Sens. Environ.* **2011**, *115*, 801–823. [CrossRef]

52. Vinukollu, R.K.; Meynadier, R.; Sheffield, J.; Wood, E.F. Multi-model, multi-sensor estimates of global evapotranspiration: Climatology, uncertainties and trends. *Hydrol. Process.* **2011**, *25*, 3993–4010. [CrossRef]

53. Yebra, M.; Dijk, A.V.; Leuning, R.; Huete, A.; Guerschman, J.P. Evaluation of optical remote sensing to estimate actual evapotranspiration and canopy conductance. *Remote Sens. Environ.* **2013**, *129*, 250–261. [CrossRef]

54. Demarty, J.; Chevallier, F.; Friend, A.; Viovy, N.; Piao, S.; Ciais, P. Assimilation of Global Modis Leaf Area Index Retrievals within a Terrestrial Biosphere Model. *Geophys. Res. Lett.* **2007**, *34*. [CrossRef]

55. Specht, D.F. A general regression neural network. *IEEE Trans. Neural Netw.* **1991**, *2*, 568. [CrossRef] [PubMed]

56. Zhang, L.; Yao, Y.; Wang, Z.; Jia, K.; Zhang, X.; Zhang, Y.; Wang, X.; Xu, J.; Chen, X. Satellite-derived spatiotemporal variations in evapotranspiration over northeast China during 1982–2010. *Remote Sens.* **2017**, *9*, 1140. [CrossRef]

57. Zhao, M.; Running, S.W.; Nemani, R.R. Sensitivity of moderate resolution imaging spectroradiometer (modis) terrestrial primary production to the accuracy of meteorological reanalyses. *J. Geophys. Res. Biogeosci.* **2006**, *111*, 338–356. [CrossRef]

58. Baldocchi, D. Breathing of the terrestrial biosphere: Lessons learned from a global network of carbon dioxide flux measurement systems. *Aust. J. Bot.* **2008**, *56*, 1–26. [CrossRef]

59. Yao, Y.; Liang, S.; Li, X.; Chen, J.; Liu, S.; Jia, K.; Zhang, X.; Xiao, Z.; Fisher, J.B.; Mu, Q. Improving global terrestrial evapotranspiration estimation using support vector machine by integrating three process-based algorithms. *Agric. For. Meteorol.* **2017**, *242*, 55–74. [CrossRef]

60. Yao, Y.; Liang, S.; Yu, J.; Zhao, S.; Lin, Y.; Jia, K.; Zhang, X.; Cheng, J.; Xie, X.; Sun, L. Differences in estimating terrestrial water flux from three satellite-based priestley-taylor algorithms. *Int. J. Appl. Earth Obs. Geoinf.* **2017**, *56*, 1–12. [CrossRef]