Estimating Surface Downward Longwave Radiation Using Machine Learning Methods

Chunjie Feng [1,2], Xiaotong Zhang [1,2,*], Yu Wei [1,2], Weiyu Zhang [1,2], Ning Hou [1,2], Jiawen Xu [1,2], Kun Jia [1,2], Yunjun Yao [1,2], Xianhong Xie [1,2], Bo Jiang [1,2], Jie Cheng [1,2] and Xiang Zhao [1,2]

[1] State Key Laboratory of Remote Sensing Science, Jointly Sponsored by Beijing Normal University and Institute of Remote Sensing and Digital Earth of Chinese Academy of Sciences, Beijing 100875, China; fcj20190901@mail.bnu.edu.cn (C.F.); weiyu@mail.bnu.edu.cn (Y.W.); zhangweiyu@mail.bnu.edu.cn (W.Z.); houning_0110@mail.bnu.edu.cn (N.H.); Xujw@mail.bnu.edu.cn (J.X.); jiakun@bnu.edu.cn (K.J.); yaoyunjun@bnu.edu.cn (Y.Y.); xianhong@bnu.edu.cn (X.X.); bojiang@bnu.edu.cn (B.J.); Jie_Cheng@bnu.edu.cn (J.C.); zhaoxiang@bnu.edu.cn (X.Z.)

[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

*   Correspondence: xtngzhang@bnu.edu.cn

check for updates

Abstract: The downward longwave radiation (L_d, 4–100 μm) is a major component of research for the surface radiation energy budget and balance. In this study, we applied five machine learning methods, namely artificial neural network (ANN), support vector regression (SVR), gradient boosting regression tree (GBRT), random forest (RF), and multivariate adaptive regression spline (MARS), to estimate L_d using ground measurements collected from 27 Baseline Surface Radiation Network (BSRN) stations. L_d measurements in situ were used to validate the accuracy of L_d estimation models on daily and monthly time scales. A comparison of the results demonstrated that the estimates on the basis of the GBRT method had the highest accuracy, with an overall root-mean-square error (RMSE) of 17.50 W m^{-2} and an R value of 0.96 for the test dataset on a daily time scale. These values were 11.19 W m^{-2} and 0.98, respectively, on a monthly time scale. The effects of land cover and elevation were further studied to comprehensively evaluate the performance of each machine learning method. All machine learning methods achieved better results over the grass land cover type but relatively worse results over the tundra. GBRT, RF, and MARS methods were found to show good performance at both the high- and low-altitude sites.

Keywords: downward longwave radiation; machine learning; GBRT; energy budget; random forest

1. Introduction

It is important for knowledge of weather and climate phenomena to study the Earth's surface radiation budget. The downward longwave radiation (L_d, 4–100 μm) is mainly emitted by H_2O, CO_2, O_3 molecules, and cloud droplets in the atmosphere near the Earth's surface [1] and is an indispensable part of the surface radiation energy balance and budget of the Earth's climate system [2]. Developing an accurate, global, and long-term L_d database is critical for studying changes in the Earth's climate. However, the spatial coverage of L_d observation sites is sparse and even entirely absent due to the high cost of L_d measurements. Accuracy estimation using denser stations has great importance for studying the distribution of L_d on global and regional scales.

Various models have been developed for L_d estimation over stations under clear sky conditions. The methods used for estimating L_d under clear sky conditions can be split into two types: the

parameterization methods and physical-based methods. Numerous parameterization methods usually estimate L_d from air temperature (Ta) and water vapor pressure under clear sky conditions [3–7]. Wu et al. [3] utilized eight meteorological parameterization methods to estimate instantaneous L_d using the Moderate Resolution Imaging Spectroradiometer (MODIS) atmospheric products under clear-sky conditions. The evaluation results displayed that the root-mean-square error (RMSE) values of estimated instantaneous L_d based on the selected eight methods were all less than 35 W m^{-2} at 17 sites collected from FLUXNET, AmeriFlux, BSRN, AsiaFlux, and SURFRAD networks. Although much attention has been paid to parameterization methods, most of the existing parameterization methods are only valid for local atmospheric conditions and geographic coverage. The physical-based methods estimate L_d using retrievals of Ta, humidity profiles, and top of atmosphere (TOA) radiance from satellite observations based on radiative transfer model [8,9]. Duguay et al. [8] utilized radiative transfer model to estimate L_d using satellite-retrieved and digital terrain data. The physical-based methods depend strongly on the satellite-retrieved temperature and humidity profile data for the lower atmosphere, which are not always available and are associated with great uncertainties [9–11].

Quantifying cloud effects for L_d estimates is important under cloudy sky conditions. Cloud cover fraction is usually used to measure the impact of the cloud on L_d and correct L_d under cloudy sky conditions, which is proportional to the total cloudiness [9,12,13]. Reliable cloud cover fraction measurements can be obtained from meteorological observations and satellite cloud detection products [14]. For cloudy sky situation, cloud cover fraction can also be estimated from the proportion of the measured horizontal global solar radiation to the horizontal global solar radiance [15,16]. The parameterization method of cloud effect on L_d proposed by the Crawford and Duchon [15] is considered as one of the best methods [17,18], and it has been widely used to estimate cloudy-sky L_d. For instance, Wang et al. [19] used this method [15] to explain the impact of the cloud on L_d. The evaluation results displayed that the instantaneous L_d estimates had an average bias value of 2 W m^{-2} (0.6%) under all-sky conditions at 36 ground observation sites. Since the downward shortwave radiation (DSR, 0.3–3 μm) data are more available than data for L_d, which is not a generally conventional observation parameter, Yang et al. [20] estimated L_d over the Tibetan Plateau by calculating the cloud cover fraction using DSR. The evaluation results showed an error of less than 30 W m^{-2} for four Chinese Meteorological Administration (CMA) stations in the Tibetan Plateau. The accuracy for cloudy-sky L_d estimates is lower than that in the condition of clear sky. Most previous studies that estimated L_d under cloudy sky conditions over China only used measurements collected at the Heihe River Basin [21–23]. However the accuracy of parameterization methods under cloudy sky conditions relies strongly on the calibration data, and such methods do not fully consider the impact of cloud characteristics on the model, so they may have a large bias outside their local calibration parameter range [11].

In addition to the methods mentioned above, machine learning methods are increasingly used to retrieve both surface shortwave radiation and longwave radiation [24–31]. Wei et al. [24] utilized four machine learning methods to estimate DSR using Advanced Very High Resolution Radiometer data. The gradient boosting regression tree (GBRT) method performed best on both daily and monthly time scales. Wang et al. [25] estimated all-sky surface shortwave net radiation (SSNR) using four machine learning methods with the Landsat Thematic Mapper (TM)/Enhanced Thematic Mapper Plus (ETM+) TOA reflectance and other ancillary information including clearness index and water vapor. The L_d estimates on the basis of the GBRT method similarly performed best, with an RMSE of 73.23 W m^{-2} and 18.76 W m^{-2}, a basis of 0.64 W m^{-2} and −1.74 W m^{-2}, and a coefficient of determination value (R^2) of 0.88 and 0.94 on the instantaneous and daily time scales, respectively. These results indicate that applying machine learning methods to estimate surface radiation is reliable. Moreover, Cheng et al. [30] developed linear and dynamic learning neural network models to estimate the global clear-sky surface upwelling longwave radiation using the MODIS TOA radiance. Compared with the empirical methods [32–36], physics-based methods [8,37,38], parameterization methods [5,10,22,39–42], and hybrid methods [43–48] for L_d estimation, the machine learning methods automatically learn the

relationship between input parameters and surface radiation and extract complex features from simple features. All in all, it is promising to estimate L_d on the basis of the machine learning methods.

Therefore, the objective of this study was to compare the accuracy and precision of five machine learning methods, namely artificial neural network (ANN), GBRT, random forest (RF), multivariate adaptive regression spline (MARS), and support vector regression (SVR) for L_d estimation. This study is organized as follows: The ground measurements and the five machine learning methods are introduced in Section 2. Section 3 provides the validation and comparison results of five machine learning methods for L_d estimation. Finally, the discussion and conclusion are presented in Sections 4 and 5, respectively.

2. Data and Methods

2.1. BSRN Datasets

In this study, the ground observations collected at Baseline Surface Radiation Network (BSRN) stations from 2000 to 2015 were utilized to train and validate the models on the basis of machine learning methods, including L_d (W m^{-2}), downward shortwave radiation (DSR, W m^{-2}), air temperature (Ta, °C) at 2 m height, relative humidity at 2 m height (RH, %), and elevation (m). BSRN is a project established by the World Climate Research Program (WCRP) aiming to survey changes in the Earth's surface radiation and provide support in the validation of qualities for satellite measurements and climate models [49].

The first four BSRN stations to be established began data recording in January 1992. Since then, four sites have been operating continuously ever since. Presently, 52 BSRN stations exist on seven continents: 14 in North America, 10 in Europe, 8 in Asia, 7 in Oceania, 5 in South America, 4 in Africa, and 4 in Antarctica [50]. Recent reports indicated that the BSRN measurements have the highest possible accuracy at high temporal resolution in various climate zones and with uncertainties of approximately ±5 W m^{-2} [51]. Data collected at 27 sites from 2000 to 2015 were used to train and evaluate the accuracy of the estimation model after excluding sites with questionable or discontinuous observations. Figure 1 shows the spatial distribution for 27 BSRN stations. Table 1 lists the detailed information for these BSRN stations, including geographic location, elevation, and land cover.
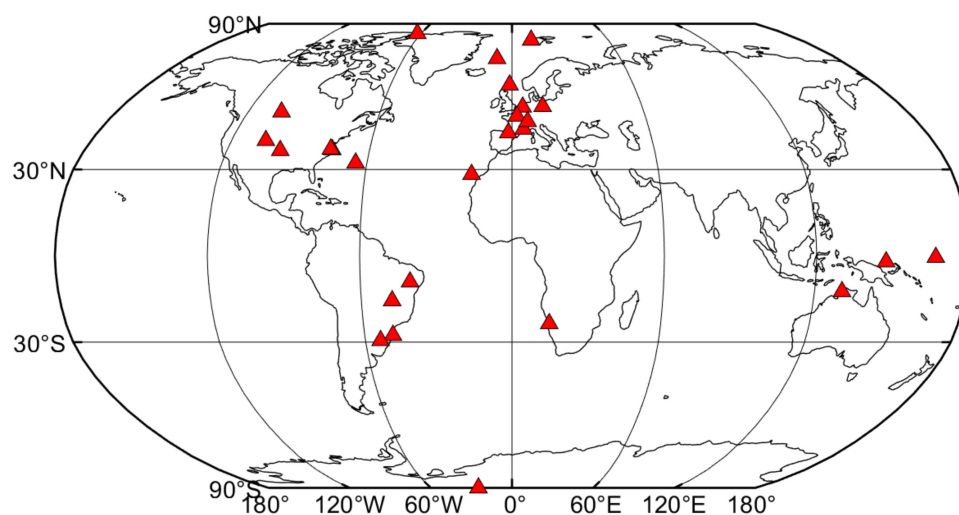

Figure 1. Spatial distribution of the Baseline Surface Radiation Network (BSRN) stations (27 sites represented by red triangles).

Table 1. Detailed information of 27 BSRN stations.

No.	Station	Full Name	Time Period	Latitude (°)	Longitude (°)	Elevation (m)	Land Cover
01	ALE	Alert	2004–2014	82.4900	−62.4200	127.0	Tundra
02	BER	Bermuda	2000–2013	32.2670	−64.6670	8.0	Ocean
03	BIL	Billings	2000–2012	36.6050	−97.5160	317.0	Grass
04	BOU	Boulder	2000–2015	40.1250	−105.2370	1577.0	Grass
05	BRB	Brasilia	2008–2015	−15.6010	−47.7130	1023.0	Concrete
06	CAB	Cabauw	2005–2015	51.9711	4.9267	0	Grass
07	CAR	Carpentras	2000–2015	44.0830	5.0590	100.0	Cultivated
08	CLH	Chesapeake Light	2000–2015	36.9050	−75.7130	37.0	Ocean
09	CNR	Cener	2009–2015	42.8160	−1.6010	471.0	Asphalt
10	DAR	Darwin	2002–2015	−12.4250	130.8910	30.0	Grass
11	E13	S. Great Plains	2000–2012	36.6050	−97.4850	318.0	Grass
12	FLO	Florianopolis	2000–2015	−27.6047	−48.5227	11.0	Concrete
13	GOB	Gobabeb	2012–2015	−23.5614	15.0420	407.0	Desert
14	GVN	Georg von Neumayer	2000–2015	70.6500	−8.2500	42.0	Ice
15	IZA	Izana	2009–2015	28.3093	−16.4993	2372.9	Rock
16	LER	Lerwick	2007	60.1389	−1.1847	80.0	Grass
17	LIN	Lindenberg	2001–2007	52.2100	14.1220	125.0	Cultivated
18	LRC	Langley Research Center	2014–2015	37.1038	−76.3872	3.0	Grass
19	MAN	Momote	2000–2013	−2.0580	147.4250	6.0	Grass
20	NAU	Nauru Island	2000–2013	−0.5210	166.9167	7.0	Rock
21	NYA	Ny-Ålesund	2000–2013	78.9250	11.9300	11.0	Tundra
22	PAL	Palaiseau Cedex	2003–2013	48.7130	2.2080	156.0	Concrete
23	PAY	Payerne	2000–2011	46.8150	6.9440	491.0	Cultivated
24	PTR	Petrolina	2008–2015	−9.0680	−40.3190	387.0	Concrete
25	REG	Regina	2000–2011	50.2050	−104.7130	578.0	Cultivated
26	SMS	São Martinho da Serra	2009–2015	−29.4428	−53.8231	489.0	Concrete
27	SPO	South Pole	2000–2015	−89.9830	−24.7990	2800.0	Glacier

2.2. Machine Learning Methods

Machine learning is dedicated to the study of how to use experience to improve the performance of the system itself through computational means, so as to generate a "model" from the data on the computer to make judgments about new situations [52,53]. These methods can learn rules from a large amount of historical data to intelligently identify new samples or predict the future. Five machine learning methods, namely artificial neural network (ANN), support vector regression (SVR), gradient boosting regression tree (GBRT), random forest (RF), and multivariate adaptive regression spline (MARS), were used to estimate L_d using ground measurements collected from 27 BSRN stations. Four variables extracted at BSRN sites were selected as predictor variables: elevation, Ta, RH, and DSR. The ground-measured daily L_d was used as a target variable. In order to acquire the best model, the dataset was randomly grouped into two parts: 80% (61,108 samples for daily) as a training dataset and remaining 20% (13,015 samples for daily) as a test dataset. By looping in the threshold of each parameter, the optimal parameters with the lowest average RMSE were selected through 5-fold cross-validation in the training data for each method.

2.2.1. Support Vector Regression

The support vector machine (SVM) is a maximum-margin classifier based on structural risk minimization in statistical learning theory developed by Vapnik [54]. The support vector regression (SVR) is an application of the SVM algorithm for regression problems [55]. It uses some nonlinear mapping to separate the sample data from the input vectors into a high-dimensional feature space and can be used to find the optimal separation hyperplane between data points of different categories in the

high-dimensional space. The SVR algorithm can overcome overfitting by introducing regularization parameters to finitely constrain samples. However, the SVR method is more suitable for small-sample data due to the high computational complexity of the algorithm.

The accuracy of the SVR method is highly influenced by the kernel. Considering that previous studies have shown that the performance of the radial basis function (RBF) kernel is better than any other kernel function in the case of regression, we used a RBF kernel in this study [56]. When training an SVR model with the RBF kernel, two parameters must be considered: C and gamma. The C parameter represents the penalty factor, which controls the trade-off between the complexity of the function and the frequency that allows the error to fall outside the SVR regression tube. Gamma represents the influence of a single training example. A larger gamma means that other examples must be influenced more closely. The range of arguments in the SVR model was predefined as (C $\in$ [0.01, 1, 10, 100], gamma $\in$ [0.01, 0.1, 10, 100]). The 5-fold cross-validation results showed that an SVR model with C parameter of 1 and gamma parameter of 0.1 was optimal for estimating the L_d.

2.2.2. Gradient Boosting Regression Tree

The gradient boosting regression tree (GBRT) algorithm can be considered as a collection of weak prediction models (typically decision trees) to enhance boosting. Similar to other boosting methods, this algorithm builds models in a stepwise fashion and generalizes these methods by allowing optimization of any differentiable loss function. The GBRT algorithm generally reduces both the variance and bias, which has been shown to be a very accurate method. This method can be used to automatically find nonlinear interactions through decision tree learning [57].

The GBRT model is influenced by four parameters: the number of iterations, the learning rate, the depth of the tree, and the sampling rate. The range of parameters was predefined (iterative number $\in$ [10, 60], learning rate $\in$ [2, 4], tree depth $\in$ [2, 10], sampling rate $\in$ [2, 6]). The validation results showed that a GBRT model with an iteration number of 150, learning rate parameter of 0.1, tree depth of 5, and sampling rate of 0.8 was optimal for estimating the L_d.

2.2.3. Random Forest

The random forest (RF) is an ensemble of multiple decision trees and was first proposed by Breiman for classification and regression problems [58]. Unlike other decision tree algorithms, the generated tree can grow to the maximum possible depth in the new training dataset by utilizing a combination of different features. The trees in the RF method are not pruned, further reducing the computational load [59]. This is the key advantage of the RF algorithm compared to other decision trees algorithms [60]. The RF algorithm can be considered as an improved version of bagging, which is based on training many classifiers on bootstrapped samples from a training set. Bagging can reduce the variance of the classification. In addition, the RF algorithm is relatively robust to outliers and noise and faster than bagging or boosting. The main factors influencing the RF model are the min-sample-split, n-estimators, min-samples-leaf, and max-features parameters. The extent of parameters was predefined (n-estimators $\in$ [10, 60], max features $\in$ [2, 4], min-samples-split $\in$ [2, 10], min-samples-leaf $\in$ [2, 6]). The 5-fold cross-validation results displayed that it is optimal to estimate the L_d for the RF model with a n-estimators parameter of 30, max features parameter of 4, min-samples-split parameter of 3, and min-samples-leaf parameter of 2.

2.2.4. Multivariate Adaptive Regression Splines

The multivariate adaptive regression splines (MARS), developed by Friedman [61], is a nonparametric regression method that can be used to build relationships between explanatory and response variables. One of the advantages of the MARS method is that the input space can be divided into multiple regions, and each region has its own regression equation. This makes the MARS method particularly suitable for problems with higher input dimensions.

The MARS method was implemented using the py-earth package in the Python platform. The estimation accuracy of the MARS model is mostly affected by the max-terms, max-degree, and endpan-alpha parameters. The extent of parameters was predefined (max-terms $\in$ [100, 500], max-degree $\in$ [1, 7], endspan-alpha $\in$ [0.3, 0.9]). The 5-fold cross-validation results proved that a max-terms parameter of 100, max-degree parameter of 2, and an endspan-alpha parameter of 0.3 were most suitable parameters for the MARS model to estimate the L_d.

2.2.5. Artificial Neural Network

The artificial neural network (ANN) is a network formed by the interconnection of many processing nodes or neurons in a certain order. There can be one or more nonlinear layers that contain several neurons between input and output layers. By repeated learning and training with known information, the network achieves the purpose of processing information and simulating the relationship between input and output through gradually adjusting the weight of neuron connections [62]. The ANN method does not need to know the exact relationship between input and output and does not require many parameters. Therefore, compared with traditional data processing methods, the ANN method has obvious advantages in processing fuzzy data, random data, and nonlinear data and is especially suitable for data with large scale and complex structure.

In this study, the ANN method was implemented on the Python platform using the Neural Network module in the scikit-learn toolbox [63]. The ANN model is influenced by six parameters: activation, hidden-layer-sizes, learning-rate, batch-size, learning-rate-init, and max-iter. The extent of parameters was predefined as (activation $\in$ ['relu', 'identity', 'logistic', 'tanh', 'sigmoid'], hidden-layer-sizes $\in$ [5, 30], learning-rate $\in$ ['invscaling', 'constant', 'invscaling', 'adaptive'], batch-size $\in$ [50, 250], learning-rate-init $\in$ [0.0001, 0.001, 0.001, 0.1], max-iter $\in$ [80, 180]). The 5-fold cross-validation results proved that an activation of 'relu', hidden-layer-size of 10, learning-rate of 'invscaling', batch-size of 250, learning-rate-init of 0.001, and max-iter of 120 were most suitable parameters for the ANN model to estimate the L_d.

3. Results

3.1. Validation against Ground Measurements

After confirming the optimal parameters, machine learning methods were utilized to estimate L_d on the training and test datasets. Ground observations for daily L_d collected at 27 BSRN stations were used to validate the accuracy of each machine learning method. Table 2 and Figures 2–5 show statistical measures for the training and test datasets of each method on both daily and monthly time scales. The root-mean-square error (RMSE), bias, and correlation coefficient (R) between BSRN observation L_d and L_d estimates based on different methods are used to measure the performance of each method. Moreover, the fitted linear regression equations are also provided.

Table 2. Statistical parameters for the training and test datasets of five methods on both daily and monthly time scales.

Time Scale	Dataset	Method	R	RMSE (W m^{-2})	Bias (W m^{-2})	Fitted Linear Regression Equation
Daily time scale	Training Dataset	GBRT	0.99	13.43	0.00	$y = 0.97 * x + 9.06$
		RF	1.00	7.64	−0.03	$y = 0.98 * x + 4.85$
		MARS	0.98	16.9	0.00	$y = 0.96 * x + 13.42$
		SVR	0.98	26.25	−8.87	$y = 0.74 * x + 75.58$
		ANN	0.98	16.84	0.28	$y = 0.96 * x + 13.42$
	Test Dataset	GBRT	0.96	17.5	−2.77	$y = 0.95 * x + 12.26$
		RF	0.96	18.96	−3.91	$y = 0.95 * x + 13.02$
		MARS	0.97	16.56	−2.87	$y = 0.97 * x + 5.97$
		SVR	0.96	18.18	4.49	$y = 0.98 * x + 12.09$
		ANN	0.98	17.27	−5.6	$y = 0.91 * x + 21.9$

Time Scale	Dataset	Method	R	RMSE (W m^{-2})	Bias (W m^{-2})	Fitted Linear Regression Equation
Monthly time scale	Training Dataset	GBRT	1.00	7.95	0.04	$y = 0.99 * x + 2.57$
		RF	1.00	4.3	0.05	$y = 0.99 * x + 1.48$
		MARS	0.99	11.82	0.38	$y = 0.98 * x + 5.46$
		SVR	0.99	23.2	−8.01	$y = 0.75 * x + 70.54$
		ANN	0.99	11.26	0.18	$y = 0.99 * x + 4.53$
	Test Dataset	GBRT	0.98	11.19	−1.62	$y = 1.00 * x − 6.98$
		RF	0.98	11.81	−2.99	$y = 1.00 * x − 3.47$
		MARS	0.99	10.21	−1.92	$y = 1.02 * x − 8.52$
		SVR	0.99	11.56	5.33	$y = 1.04 * x − 6.98$
		ANN	0.99	11.29	−5.25	$y = 0.95 * x + 11.68$

Figures 2 and 3 show a comparison between the daily L_d estimates and the ground measurements for the training and test datasets, respectively. On the whole, the performance of each model on the test dataset did not differ much compared to that on the training dataset. The RMSE and bias on the training dataset ranged from 7.64 to 26.25 W m^{-2} and −8.87 to 0.28 W m^{-2}, respectively; these values on the test dataset ranged from 16.56 to 18.96 W m^{-2} and −3.91 to 4.49 W m^{-2}, respectively. The accuracy of the L_d estimates on the basis of the RF method is best when compared to the daily ground measurements on the training dataset, obtaining overall RMSE, bias, and R of 7.64 W m^{-2}, −0.03 W m^{-2}, and 1, respectively. The fitted line of the RF method, with a slope and intercept of 0.98 and 4.85, respectively, indicates that the L_d estimates were very close to the ground observation values on the training dataset. However, the RF model displayed overfitting of L_d compared to the ground measurements, as it displayed a lower RMSE and intercept on the training set but higher values on the test set. The overall RMSE, bias, and R values on the test dataset were 18.96 W m^{-2}, −3.91 W m^{-2}, and 0.91 between the L_d estimates on the basis of the RF model and the ground measurements, respectively. The other four models did not exhibit the same evident overfitting problem as found in the RF model. The overfitting problem may be relevant to a specific shortcoming of the RF method. When samples are concentrated in a specified extent, the result of the RF model displays the phenomenon of overfitting. Compared with the other models, the SVR model was the most inaccurate for the training dataset. When the SVR model was used to estimate daily L_d, the overall RMSE, bias, and R values for the training dataset were 26.25 W m^{-2}, −8.87 W m^{-2}, and 0.98, respectively. For the SVR method, the slope and intercept of the fitted linear regression equation on the training dataset were 0.74 and 75.58, respectively, which indicated that the daily L_d estimates seriously deviate from the 1:1 line when compared to the other four machine learning methods. The SVR method tended to overestimate L_d for low values and to underestimate L_d for high values. The overall RMSE value of the SVR model on the test dataset decreased to 18.18 W m^{-2}; however, the value was still higher than those of the other models, except the RF model. The best correlation was seen between the L_d estimates on the basis of the MARS method and the ground measurements, which achieved an overall RMSE, bias, and R values on the test dataset of 16.56 W m^{-2}, −2.87 W m^{-2}, and 0.97, respectively. The slope and intercept of the fitted linear regression equation on the test dataset were 0.97 and 5.97, respectively, which was the closest to the 1:1 line when compared to the other four machine learning methods. The L_d estimates on the basis of the GBRT and ANN methods also correlated well with the ground measurements. The RMSE, bias, and R values for the GBRT model were 17.5 W m^{-2}, −2.77 W m^{-2}, and 0.96, respectively; these values for the ANN model were 17.27 W m^{-2}, −5.6 W m^{-2}, and 0.98, respectively. The similarity between the ANN and MARS models was that the differences between their RMSEs on the training and test datasets were small, less than 1 W m^{-2}. The RMSEs on the training dataset were 16.90 W m^{-2} and 17.27 W m^{-2} for the MARS and ANN models, respectively. Compared with the MARS model, the GBRT model achieved lower a RMSE of 13.43 W m^{-2} on the training dataset, and the RMSE on the test dataset differed by less than 1 W m^{-2}. Therefore, it is reasonable and accurate to estimate the daily L_d on the basis of the GBRT model.
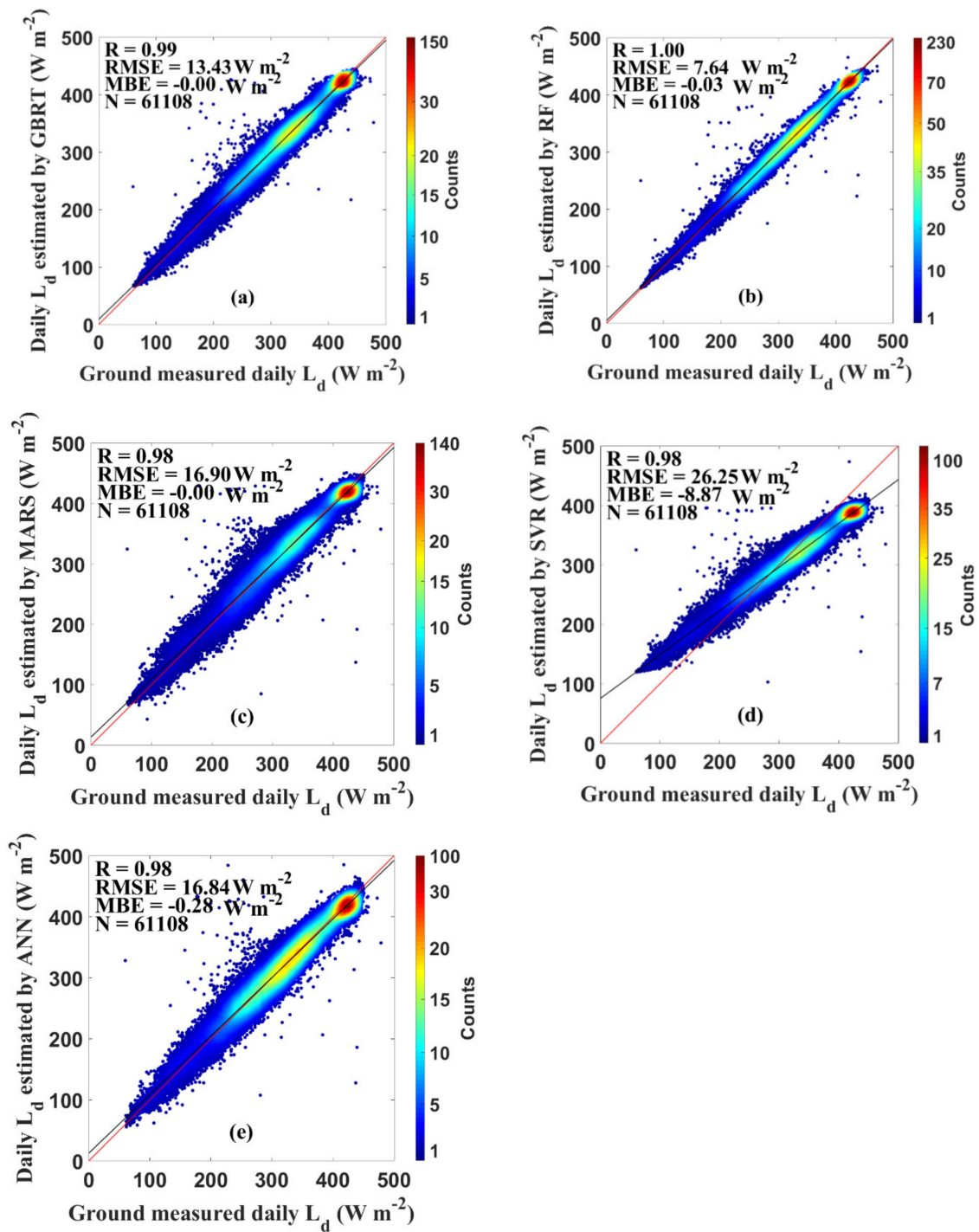
Figure 2. Evaluation results of the daily L_d estimates on the basis of the (**a**) gradient boosting regression tree (GBRT), (**b**) random forest (RF), (**c**) multivariate adaptive regression spline (MARS), (**d**) support vector regression (SVR), and (**e**) artificial neural network (ANN) methods against the ground measurements on the training dataset. N is the number of data records. The red and black lines represent the 1:1 and fitted lines, respectively.
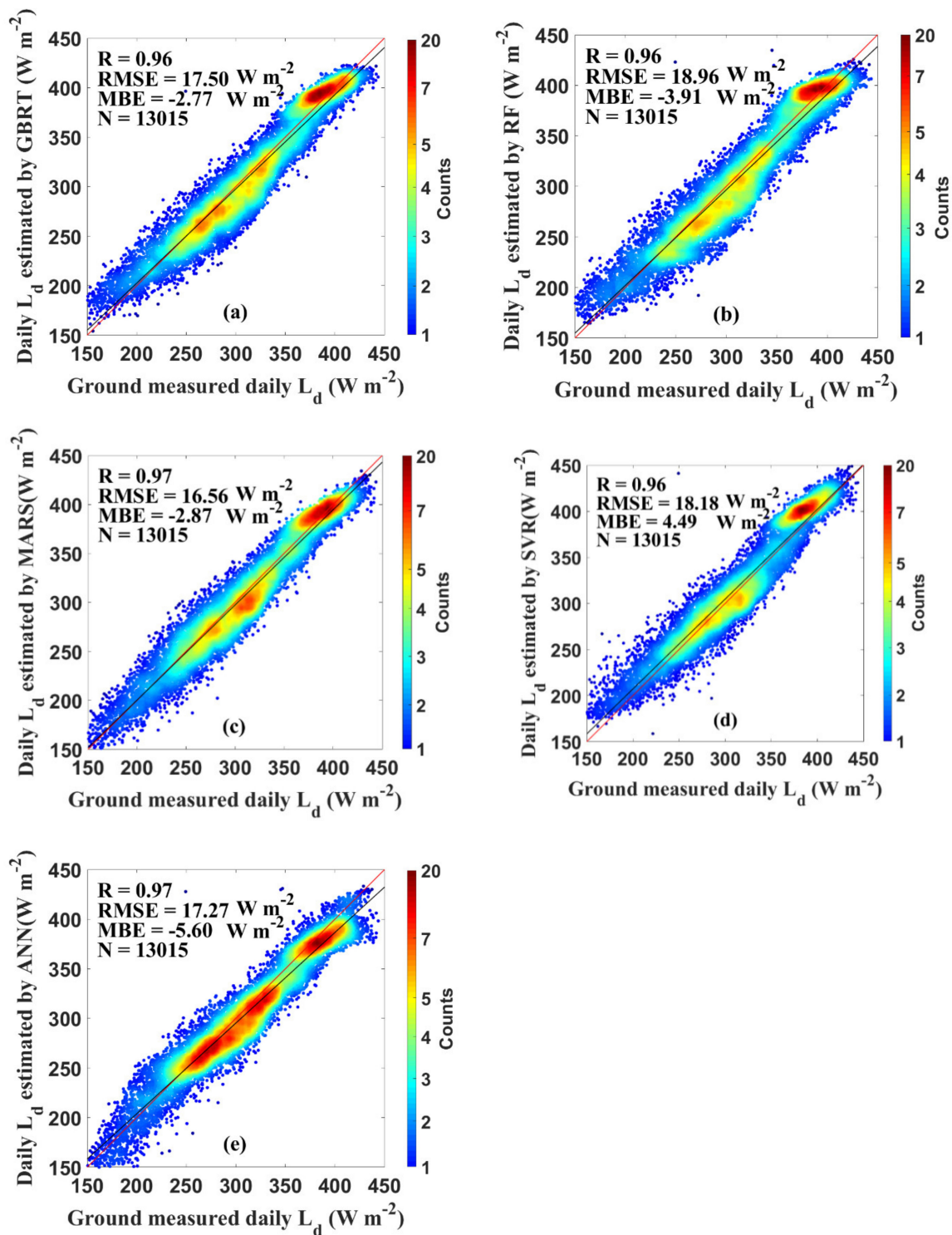
Figure 3. Evaluation results of the daily L_d estimates on the basis of the (**a**) GBRT, (**b**) RF, (**c**) MARS, (**d**) SVR, and (**e**) ANN methods against the ground measurements on the test dataset. N is the number of data records. The red and black lines represent the 1:1 and fitted lines, respectively.

Monthly L_d estimates were obtained by averaging the daily values over a month. If there were more than nine missing daily L_d estimates in a single month at a station, the data for that month at that station were deleted. The calculated average monthly L_d estimates were utilized to validate the consistency against the BSRN monthly L_d measurements on a monthly time scale. Figures 4 and 5 display the evaluation results for the L_d estimates on the basis of each method on a monthly time scale.

On the whole, the performance of each model for the test dataset did not differ much compared to that for the training dataset on a monthly time scale. The RMSE and bias on the training dataset ranged from 4.3 to 23.2 W m^{-2} and −8.01 to 0.38 W m^{-2}, respectively; these values on the test dataset ranged from 10.21 to 11.81 W m^{-2} and −2.99 to 5.33 W m^{-2}, respectively. When the RF method was used to estimate L_d, the overall RMSE, bias, and R value for the training dataset were 4.3 W m^{-2}, 0.05 W m^{-2}, and 1, respectively, showing the best correlation. These values for the training dataset were 7.95 W m^{-2}, 0.04 W m^{-2}, and 1, respectively, between the monthly L_d estimates on the basis of the GBRT method and the ground measurements, indicating that the GBRT method was also suitable for predicting L_d. Compared with the other methods, higher RMSE and bias for the training dataset existed between the ground measurements and the monthly L_d estimates on the basis of the SVR method, with values of 23.2 W m^{-2} and −8.01 W m^{-2}, respectively. As displayed in Figure 5, the MARS model obtained the lowest RMSE when compared to other models, achieving the overall RMSE, bias, and R value of 10.21 W m^{-2}, −1.92 W m^{-2}, and 0.99 on the test dataset, respectively. The same conclusion can be drawn from the test results on a daily time scale. The overall RMSE, bias, and R values of the GBRT method were 11.19 W m^{-2}, −1.62 W m^{-2}, and 0.98 on the test dataset, respectively. Although the RMSE of the MARS method was lower than that of the GBRT method, the difference between them was less than 1 W m^{-2}, on a monthly time scale. The bias of the GBRT method was lower than that of the MARS method on both daily and monthly time scales. The RMSE, bias, and R values of the ANN method were 11.29 W m^{-2}, 5.25 W m^{-2}, and 0.99, respectively. The RF model performed worst on a monthly time scale, with an overall RMSE value increasing to 11.81 W m^{-2}, for the test dataset. Therefore, it is also rational and accurate to conduct monthly L_d estimation on the basis of the GBRT model.

According to the evaluation results for the five machine learning methods, the GBRT, RF, ANN, and MARS methods have a significant advantage in accuracy over the SVR method. Compared with other machine methods, the GBRT method showed a relatively high precision on both daily and monthly time scales, obtaining RMSE on the test dataset of 17.50 W m^{-2} and 11.19 W m^{-2}, respectively. In summary, the GBRT method is relatively accurate for L_d estimation using elevation, Ta, RH, and DSR among five machine learning methods studied.
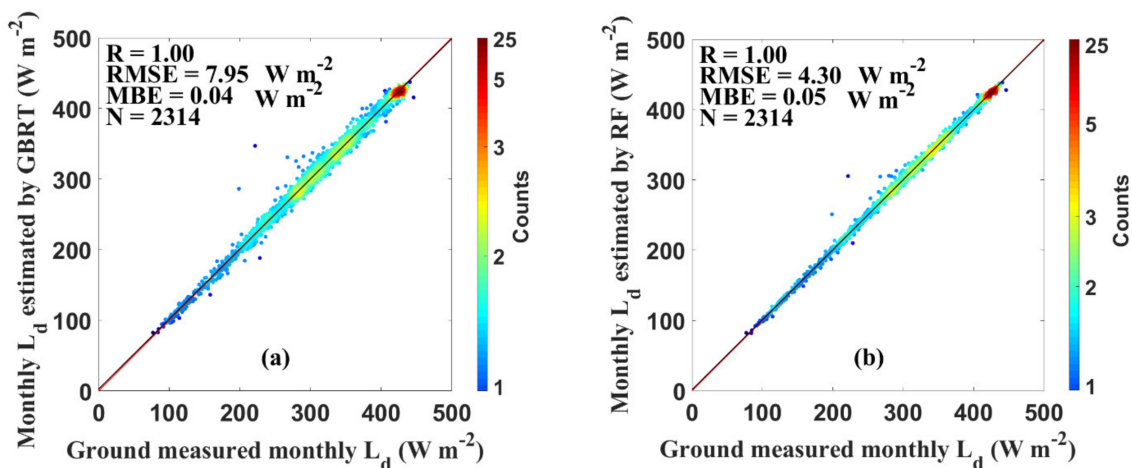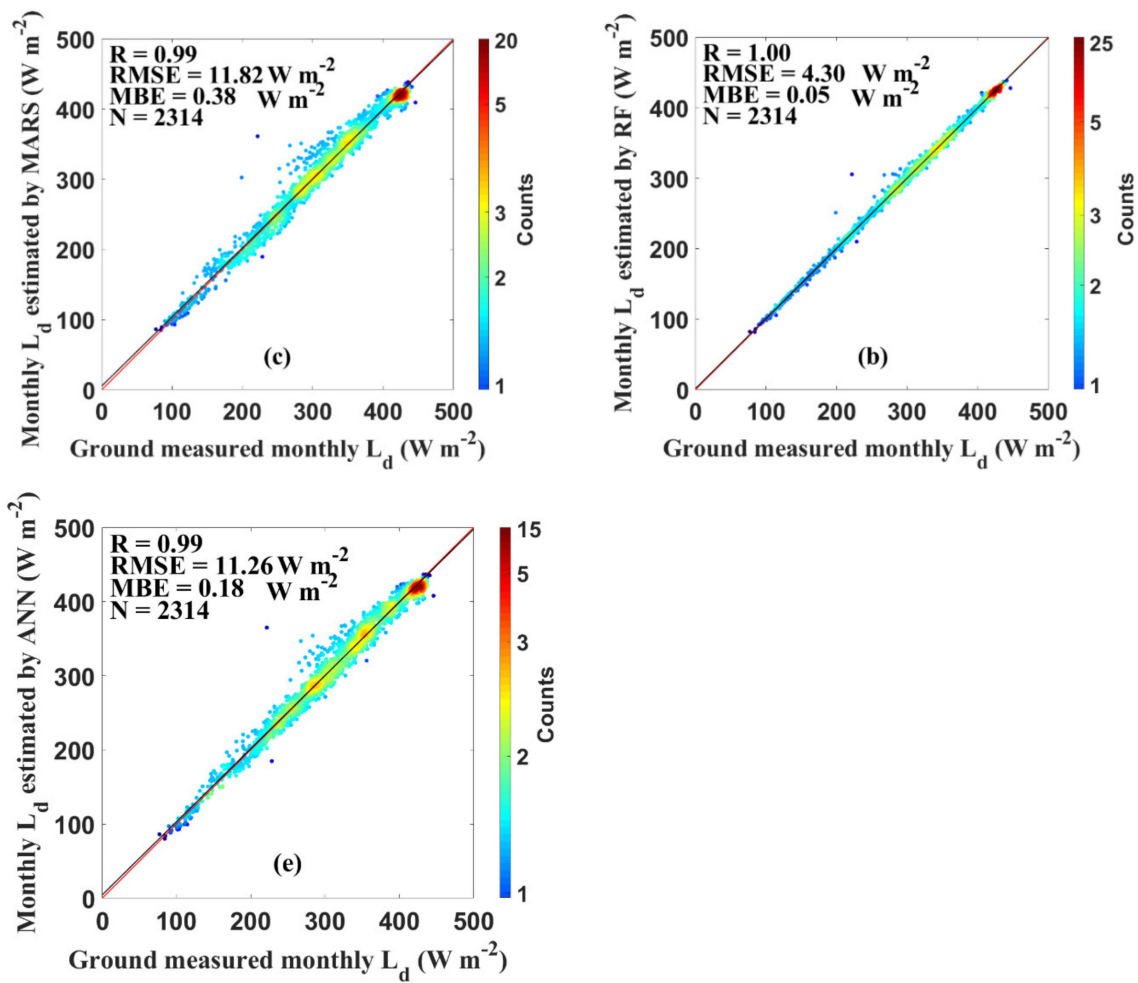

Figure 4. *Cont.*

Figure 4. Evaluation results of the monthly L_d estimates on the basis of the (**a**) GBRT, (**b**) RF, (**c**) MARS, (**d**) SVR, and (**e**) ANN methods against the ground measurements on the training dataset. N is the number of data records. The red and black lines represent the 1:1 and fitted lines, respectively.
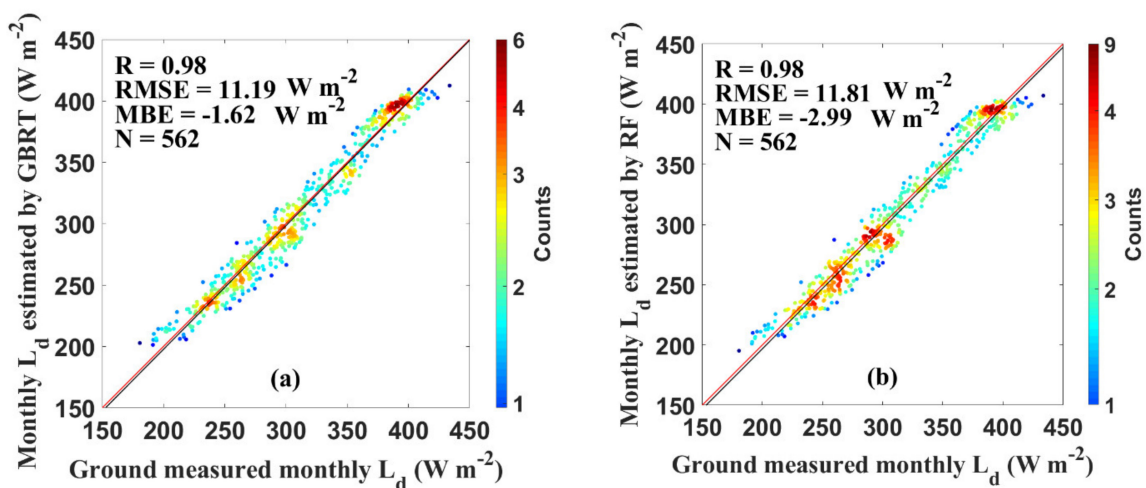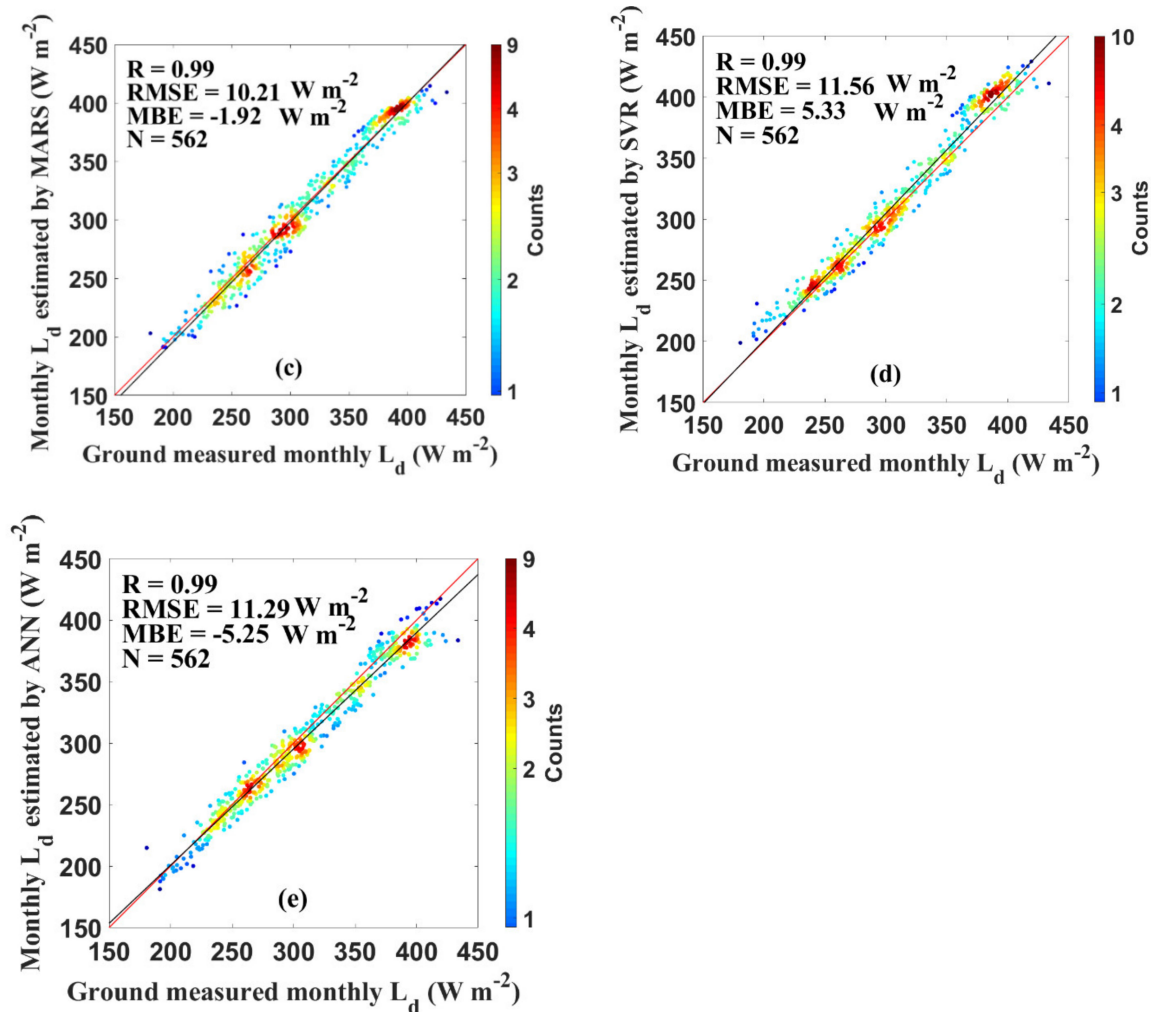

Figure 5. *Cont.*

Figure 5. Evaluation results of the monthly L_d estimates on the basis of the (**a**) GBRT, (**b**) RF, (**c**) MARS, (**d**) SVR, and (**e**) ANN methods against the ground measurements on the test dataset. N is the number of data records. The red and black lines represent the 1:1 and fitted lines, respectively.

3.2. The Influence of Land Cover and Elevation on L_d Estimation

Land cover and elevation can influence the interaction between the earth and atmosphere [64]. In this section, we present the results of our attempt to investigate the influence of land cover and elevation on the L_d estimation based on the above five machine learning methods using both training data (61,108 samples for daily) and test data (13,015 samples for daily) at all sites. We divided the sites into seven primary land cover types: desert, bare land, cropland, grass, glacier, ocean, and tundra. Among them, asphalt, concrete, and rock land covers are considered as bare land category and ice land cover is considered as glacier category. Then, the accuracy of each method over different land cover types was evaluated. The statistical results for five machine learning methods over each land cover type and surface elevation are provided in Tables 3 and 4 and Figures 6 and 7. As shown in Figures 6 and 7, the box plots can visually display the outliers, data dispersion, and bias of a dataset. The box part represents the middle 50% of the observed value, and the upper boundary, middle line, and lower boundary of the box represent the upper quartile, median, and lower quartile, respectively. The dotted lines extending from the edge of the box represent the maximum and minimum values, and the scattered red asterisks represent the outliers of the dataset.

Although the selected machine learning methods have diverse performance over different land covers on the training dataset, all methods can generally achieve better results over the grass land cover type but show relatively poorer results over the tundra. Moreover, the five methods showed

better performance on the training dataset than on the test dataset. On the training dataset, the RMSE ranged from 6.02 to 23.11 W m^{-2} and 6.64 to 24.13 W m^{-2} over the desert and grass land cover types, respectively; the mean RMSE of the five methods was 13.28 W m^{-2} and 13.75 W m^{-2} over desert and grass land cover types, respectively. Compared with the grass land cover type, the larger bias values over the desert on the training dataset were −6.77 W m^{-2}, −1.81 W m^{-2}, −1.49 W m^{-2}, −1.15 W m^{-2}, and −16.53 W m^{-2} for the ANN, GBRT, MARS, RF, and SVR methods, respectively. On the test dataset, the RMSE of all methods over the desert and grass land cover types ranged from 10.66 to 21.89 W m^{-2} and 11.85 to 24.39 W m^{-2}, respectively; the mean RMSE was 15.31 W m^{-2} and 15.52 W m^{-2} over the desert and grass land cover types, respectively. Meanwhile, the RMSE over the tundra ranged from 9.16 to 24.84 W m^{-2} and 20.09 to 22.91 W m^{-2} and the mean RMSE of five methods was 19.25 W m^{-2} and 21.12 W m^{-2} on the training and test datasets, respectively. Over the cropland, the RMSE and bias of all methods on the training dataset varied from 8.06 to 19.24 W m^{-2} and −5.82 to 0.84 W m^{-2}, respectively, which indicated that all models underestimated the L_d, except the ANN method. The RMSE and bias of all methods on the test dataset varied from 14.76 to 19.13 W m^{-2} and −5.32 to 1.47 W m^{-2}, respectively. Over the bare land cover, the GBRT method on the test dataset obtained the lowest RMSE, 13.05 W m^{-2}, but this value was still higher than those obtained over the desert and grass land cover types. The RMSE of all methods over the ocean was similar with that obtained over the glacier on both training and test datasets, with differences of less than 1 W m^{-2}, except for the SVR method.

Among five machine learning methods, the RF method over all land cover types performed best on the training dataset; its RMSE and bias ranged from 6.02 to 9.16 W m^{-2} and −1.15 to 0.49 W m^{-2}, respectively. However, it achieved worse results on the test dataset, with RMSE and bias ranging from 15.70 to 22.91 W m^{-2} and −5.07 to 1.05 W m^{-2}, respectively. This may be related to the overfitting issue for the RF method. The SVR method showed larger RMSE and bias, with values ranging from 19.13 to 28.63 W m^{-2} and −15.47 to 18.53 W m^{-2}, respectively, on the test dataset. When the SVR method was used to estimate L_d, the performance was the best over the cropland and worst over the glacier, which was different from the other methods. For the MARS method, the RMSE and bias over different land cover types ranged from 10.58 to 20.96 W m^{-2} and −4.52 to 0.49 W m^{-2}, respectively, on the test dataset. Among seven land cover types, the MARS method performed worst over the tundra, obtaining overall RMSE and bias values of 20.96 W m^{-2} and −4.52 W m^{-2}, respectively, on the test dataset, whereas this method performed best over the desert, with RMSE and bias values decreasing to 10.58 W m^{-2} and 0.32 W m^{-2}, respectively, on the test dataset. The underestimation of L_d on the test dataset occurred over the desert, grass, and tundra land cover types for all methods. The overestimation of L_d on the test dataset occurred over the cropland and ocean land cover types, except with the RF and SVR methods. According the validation dataset on the test dataset, the distribution of RMSE values over different land cover types was found to be uniform between the GBRT and MARS methods, both ranking from smallest to largest for the desert, grass, bare land, glacier, ocean, cropland, and tundra. However, the RMSE values of the GBRT method were lower than those of the MARS method on the test dataset, except for the desert and cropland land cover types. The performance of the GBRT method was second only to the RF method on the training dataset and was the best over most land cover types on the test dataset. For the GBRT method, the difference between RMSE on the training and test datasets was small, less than 1 W m^{-2} over different land cover types, which indicated that this method did not have overfitting phenomenon and was relatively most suitable to estimate L_d.

Similar to the method of analyzing the effect of land cover types, we divided the site elevation (H, in m) into four ranges, namely H < 500 m, 500 m < H < 1000 m, 1000 m < H < 2000 m, and H > 2000 m, to evaluate the influence of site elevation. The validation results are shown in Figures 6 and 7 and Tables 3 and 4. According to the mean RMSE of five methods at different elevations, it can be concluded that all models can achieve relatively better results at the 1000 m < H < 2000 m elevation sites regardless of considering the training or test dataset. The mean RMSE of five methods at the 1000 m < H < 2000 m elevation sites was 12.00 W m^{-2} and 13.54 W m^{-2}, respectively, on the training and test datasets. For the H > 2000 m elevation sites, the performance of five methods on the

training dataset was quite different, and the RF and SVR methods obtained lowest and highest RMSE, with values of 6.86 W m^{-2} and 33.74 W m^{-2}, respectively, on the training dataset. For the H < 500 m elevation sites, all methods on the test dataset underestimated L_d, except the RF method whose bias was zero. Compared with the H < 500 m elevation sites, using the SVR method to estimate L_d at the 500 m < H < 1000 m elevation sites had better fitting result, obtaining lower RMSE and bias regardless of considering the training or test dataset. When the ANN, GBRT, MARS, and RF methods were used to estimate L_d, the RMSE at the 500 m < H < 1000 m elevation sites was higher than that obtained at the H < 500 m elevation sites, with value of 19.94 W m^{-2}, 17.48 W m^{-2}, 17.64 W m^{-2}, and 17.79 W m^{-2} respectively. The L_d overestimation of all methods on the test dataset occurred at the 500 m < H < 1000 m elevation sites, with bias value of 1.03 W m^{-2}, 1.04 W m^{-2}, 1.37 W m^{-2}, 4.84 W m^{-2}, and 1.35 W m^{-2} for the GBRT, MARS, RF, ANN, and SVR methods, respectively.

According to Tables 3 and 4 and Figures 6 and 7, it can be concluded that the RF method on the training dataset performed best at all elevation ranges, with RMSE and bias ranging from 6.22 to 9.70 W m^{-2} and –0.44 to 0.37 W m^{-2}, respectively, whereas its performance on the test dataset was worse than those of the GBRT and MARS methods. The difference between RMSEs on the training and test datasets was larger (>6 W m^{-2}). The SVR method had a larger RMSE, with value ranging from 16.70 to 33.74 W m^{-2} and 16.76 to 32.88 W m^{-2}, respectively, on the training and test datasets. The performance of the GBRT and MARS methods was similar, both receiving the lowest RMSE at 1000 < H < 2000 m sites and the highest RMSE at 500 m < H < 1000 m sites on both training and test datasets. Their bias values were close to zero on the test dataset, with difference less than 1 W m^{-2} at the different elevation ranges. When the ANN method was used to estimate L_d, the performance was the best at 1000 m < H < 2000 m sites and worst at 500 m < H < 1000 m sites, with RMSE of 13.98 W m^{-2} and 20.07 W m^{-2} on the test dataset, respectively. All in all, the GBRT method not only lacked an overfitting phenomenon but also performed best on the test dataset.

Table 3. Averaged RMSE and bias (W m^{-2}) of sites with same land cover type or elevation on the training dataset.

Class		No. of Samples	RMSE/Bias	ANN	GBRT	MARS	RF	SVR	Mean
Land cover	Bare land	12,759	RMSE	17.27	13.01	13.37	7.01	23.59	14.85
			Bias	1.85	0.18	0.56	0.39	−13.68	−2.14
	Cropland	11,742	RMSE	17.05	14.81	14.81	8.06	19.24	14.79
			Bias	0.84	−0.15	−0.12	−0.15	−5.82	−1.08
	Desert	909	RMSE	15.37	11.11	10.8	6.02	23.11	13.28
			Bias	−6.77	−1.81	−1.49	−1.15	−16.53	−5.55
	Glacier	5452	RMSE	18.66	14.14	14.1	8.70	30.36	17.19
			Bias	−6.19	−0.26	−0.37	−0.20	19.68	2.53
	Grass	17,618	RMSE	13.81	12.07	12.09	6.64	24.13	13.75
			Bias	−2.38	−0.36	−0.19	−0.34	−14.97	−3.65
	Ocean	6768	RMSE	18.11	15.04	14.97	8.46	21.35	15.59
			Bias	5.92	0.45	0.49	0.49	−11.01	−0.73
	Tundra	5860	RMSE	20.34	20.71	21.18	9.16	24.84	19.25
			Bias	−1.53	−3.89	−4.95	0.19	6.57	−0.72
elevation	0–500	58,813	RMSE	16.84	14.53	14.69	7.71	23.67	15.49
			Bias	−0.47	−0.59	−0.54	0	−10.28	−2.38
	500–1000	4158	RMSE	20.07	17.63	17.7	9.7	20.24	17.07
			Bias	3.67	−0.08	−0.15	0.37	0.4	0.84
	1000–2000	7259	RMSE	13.98	11.59	11.52	6.22	16.7	12.00
			Bias	−1.84	0.08	0	−0.44	−4.99	−1.44
	>2000	3893	RMSE	17.9	11.61	11.6	6.86	33.74	16.34
			Bias	1.22	−0.08	−0.03	0.17	23.29	4.91

Table 4. Averaged RMSE and bias (W m^{-2}) of sites with same land cover type or elevation on the test dataset.

| Class		No. of Samples	RMSE/Bias	ANN	GBRT	MARS	RF	SVR	Mean
Land cover	Bare land	2736	RMSE	17.36	13.05	13.48	15.70	23.71	16.66
			Bias	1.78	−0.03	0.32	−1.18	−13.73	−2.57
	Cropland	2465	RMSE	17.15	14.86	14.76	18.07	19.13	16.79
			Bias	1.47	0.40	0.49	−3.21	−5.32	−1.23
	Desert	206	RMSE	14.86	10.66	10.58	18.55	21.89	15.31
			Bias	−6.88	−1.84	−1.51	−5.07	−15.47	−6.15
	Glacier	1145	RMSE	17.63	13.76	13.79	16.57	28.63	18.08
			Bias	−6.65	−0.22	−0.38	1.05	18.53	2.47
	Grass	3794	RMSE	14.07	11.85	11.89	15.42	24.39	15.52
			Bias	−2.66	−0.46	−0.37	−3.31	−15.51	−4.46
	Ocean	1417	RMSE	17.97	14.70	14.74	16.82	24.39	17.72
			Bias	6.07	0.38	0.35	−1.84	−10.98	−1.20
	Tundra	1252	RMSE	20.09	20.50	20.96	22.91	21.16	21.12
			Bias	−0.82	−3.32	−4.52	−4.93	−10.85	−4.89
elevation	0–500	10,421	RMSE	16.88	14.34	14.53	14.2	23.66	16.72
			Bias	−0.50	−0.59	−0.59	0.10	−10.27	−2.37
	500–1000	685	RMSE	19.94	17.48	17.64	17.79	19.77	18.52
			Bias	4.84	1.03	1.04	1.37	1.35	1.93
	1000–2000	1271	RMSE	14.4	11.81	11.88	12.83	16.76	13.54
			Bias	−1.66	0.29	0.19	−0.50	−4.90	−1.32
	>2000	638	RMSE	15.76	11.57	11.50	12.50	32.88	16.84
			Bias	10.77	−0.08	−0.03	0.17	23.29	6.82
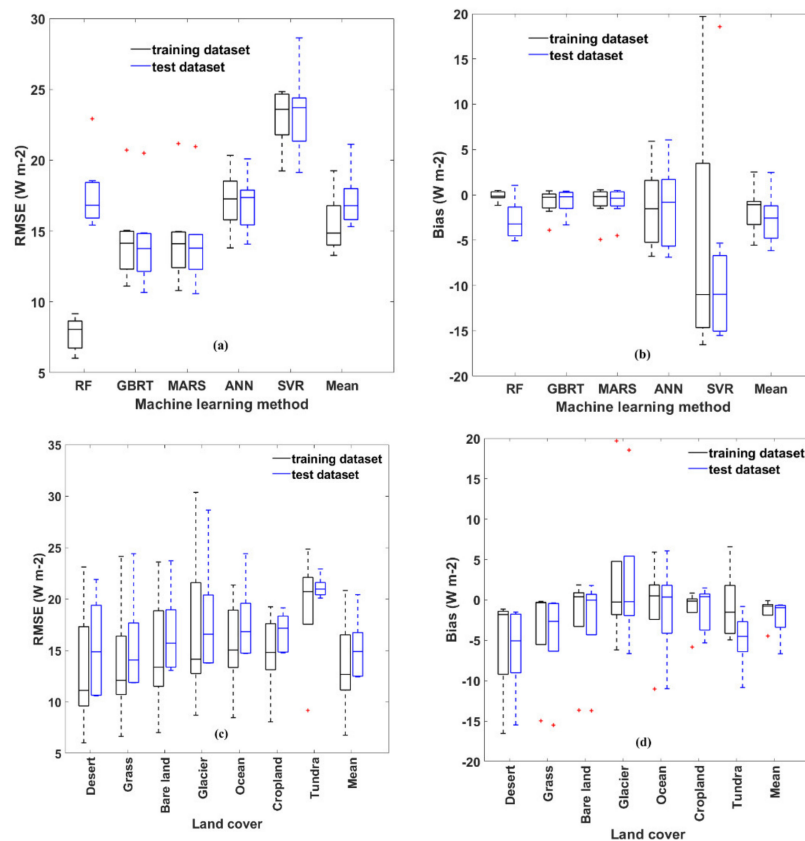
Figure 6. Box plots of the (**a**) RMSE and (**b**) bias of different land cover types for each method; the (**c**) RMSE and (**d**) bias of different machine learning methods over each land cover type. The middle line within the box represents the median value. The red asterisks represent the outliers.
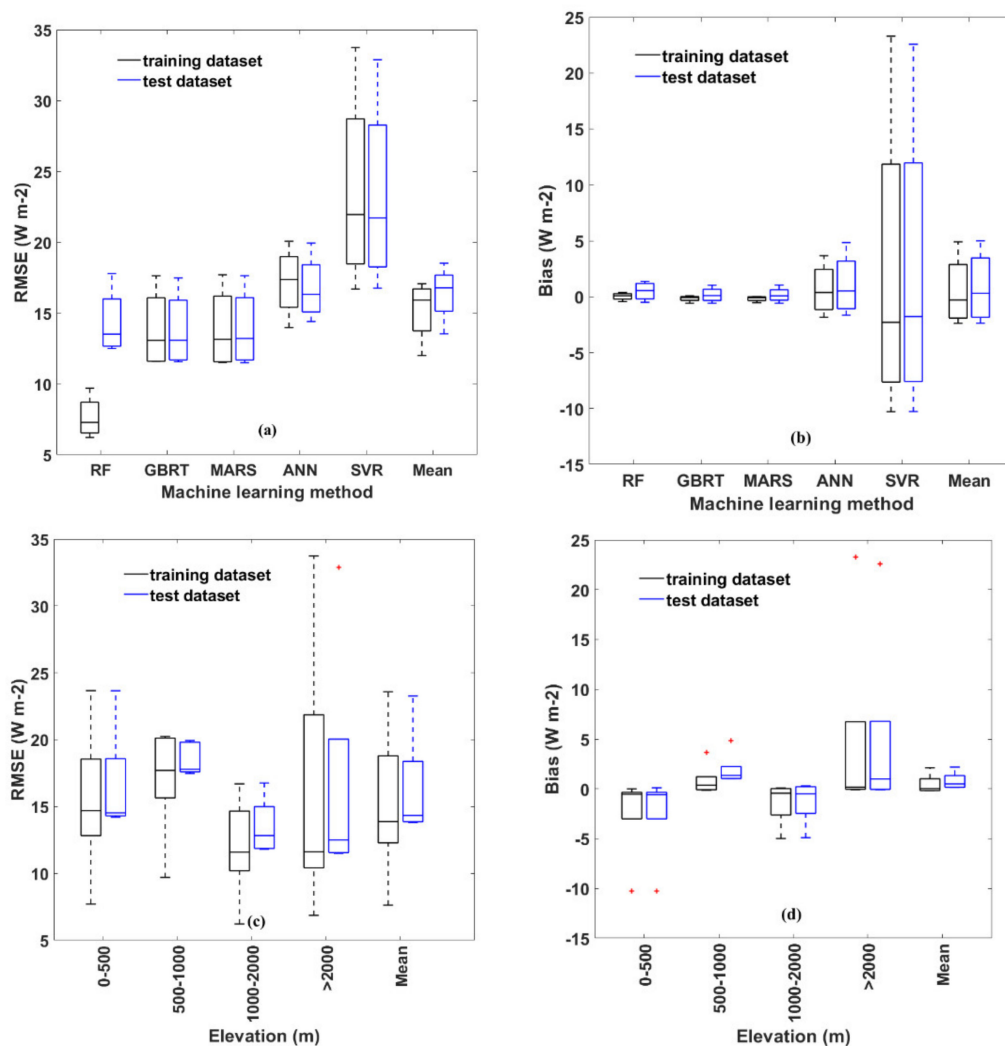
Figure 7. Box plots of the (**a**) RMSE and (**b**) bias of different elevation ranges for each method; the (**c**) RMSE and (**d**) bias of different machine learning methods at each elevation range. The middle line within the box represents the median value. The red asterisks represent the outliers.

4. Discussion

In this study, we applied five machine learning methods, namely artificial neural network (ANN), support vector regression (SVR), gradient boosting regression tree (GBRT), random forest (RF), and multivariate adaptive regression spline (MARS), to estimate L_d using the ground measurements collected at 27 Baseline Surface Radiation Network (BSRN) stations from 2000 to 2015. The evaluation results indicated that different methods have different performances on the training or test dataset, on the daily or monthly time scale, over the different land cover types and at different elevation ranges. On the whole, for the training dataset, the RMSEs of five methods were consistent on daily or monthly time scale, ranked from lowest to highest as follows: RF, GBRT, ANN, MARS, and SVR. The lowest RMSEs were 7.64 W m^{-2} and 4.3 W m^{-2} and the highest RMSEs were 26.25 W m^{-2} and 23.2 W m^{-2} on the daily and monthly time scales, respectively, for the training dataset. For the test dataset, although the sorting of RMSE between five methods was diverse on daily and monthly time scales, the differences between RMSE values for different methods were small, less than 3 W m^{-2} on both daily and monthly time scales. It is worth noting that for the training dataset, the fitted regression line of the SVR method seriously deviated from the 1:1 line, whose coefficients were 0.74 and 0.75 and intercepts were 75.58 and 70.54 on the daily and monthly time scales, respectively. Although the RF method obtained the lowest RMSE on the training dataset, it performed worse than the GBRT method

on the test dataset. The performance of the GBRT method was second only to the RF method on the training dataset and was the best for most land cover types and elevation ranges on the test dataset. Furthermore, the difference between RMSEs on the training and test datasets was small, less than 1 W m^{-2}, regardless of land cover type or elevation range, which indicated that this method did not have overfitting phenomenon and was relatively the most suitable method for estimating L_d.

Although machine learning methods can reliably estimate L_d using elevation, Ta, RH, and DSR, there are still some disadvantages. The GBRT method does not have an explicit regularization, and its regression tree learner is regarded as a black box, which may not be valid [57]. The RF method has an obvious overfitting issue, especially when samples are concentrated in a specified extent [65]. Moreover, the ground measurements themselves are flawed, which is also one of the potential sources of errors in the model [66].

5. Conclusions

L_d is one of the key components for calculating the surface energy budget. Accurate estimation of L_d has great importance for knowledge of the Earth's radiation budget on global and regional scales. The primary goal of this paper was to demonstrate the usage of five machine learning methods, i.e., ANN, GBRT, RF, MARS and SVR, to estimate L_d using the ground measurements collected at 27 BSRN stations from 2000 to 2015. Validation was carried out using the L_d ground observations on both daily and monthly time scales. According to the comparison of the evaluation results, the GBRT method showed the highest accuracy for L_d estimation; its overall RMSE and R on the test dataset were 17.50 W m^{-2} and 0.96, respectively, on a daily time scale. These values were 11.19 W m^{-2} and 0.98, respectively, on a monthly time scale. The experimental operating environment was a Windows 10 Intel Core i7-8700 CPU, 3.2 GHz, 3.19 GB RAM processor. The mean calculation time of the GBRT method was within 15 s. Therefore, the GBRT method is relatively the most suitable for estimating L_d regardless of considering the calculation time or prediction accuracy.

The influence of land cover and elevation on the five machine learning methods was also discussed. The accuracy of each machine learning method was found to be better over the grass land cover type but relatively worse over the tundra. The evaluation result indicated that all models could achieve better results at the 1000 m < H < 2000 m elevation sites. The GBRT method performed better at all different elevation ranges, with RMSE and bias ranging from 11.57 to 17.48 W m^{-2} and −0.59 to 1.03 W m^{-2}, respectively, on the test dataset, whereas the SVR method obtained larger RMSE and bias, with values ranging from 16.76 to 32.88 W m^{-2} and −10.27 to 22.57 W m^{-2}, respectively, on the test dataset.

We conclude that it is reasonable to apply machine learning methods to estimate L_d by building relationships between input variables and L_d. Although the GBRT method showed the best performance, the differences among the different machine learning methods used were small. Analysis of bias showed that the L_d estimates on the basis of five machine learning methods were approximately underestimated or overestimated by 2 to 8 W m^{-2}. Reducing bias and overfitting by optimizing the selection of parameters and variables requires further research.

Author Contributions: Conceptualization, X.Z. (Xiaotong Zhang); data curation, C.F. and Y.W.; supervision, X.Z. (Xiaotong Zhang), Y.W., W.Z., N.H., and J.X.; writing—original draft, C.F.; writing—review & editing, X.Z. (Xiaotong Zhang), Y.W., C.F., K.J., Y.Y., X.X., B.J., J.C., and X.Z. (Xiang Zhao). All authors have read and agreed to the published version of the manuscript.

References

1. Guo, Y.M.; Cheng, J.; Liang, S.L. Comprehensive assessment of parameterization methods for estimating clear-sky surface downward longwave radiation. *Theor. Appl. Climatol.* **2019**, *135*, 1045–1058. [CrossRef]
2. Oke, T.R. *Boundary Layer Climates*, 2nd ed.; Routledge: London, UK, 1996; p. 435.
3. Wu, H.R.; Zhang, X.T.; Liang, S.L.; Yang, H.; Zhou, G.Q. Estimation of clear-sky land surface longwave radiation from MODIS data products by merging multiple models. *J. Geophys. Res. Atmos.* **2012**, *117*. [CrossRef]
4. Idso, S.B. A set of equations for full spectrum and 8-Mu-M to 14-Mu-M and 10.5-Mu-M to 12.5-Mu-M thermal-radiation from cloudless skies. *Water Resour. Res.* **1981**, *17*, 295–304. [CrossRef]
5. Brutsaert, W. Derivable formula for long-wave radiation from clear skies. *Water Resour. Res.* **1975**, *11*, 742–744. [CrossRef]
6. Swinbank, W.C. Long-wave radiation from clear skies. *Q. J. R. Meteorol. Soc.* **1963**, *89*, 339–348. [CrossRef]
7. Brunt, D. Notes on radiation in the atmosphere. *Q. J. R. Meteorol. Soc.* **1932**, *58*, 389–420. [CrossRef]
8. Duguay, C.R. An Approach To the Estimation of Surface Net-Radiation in Mountain Areas Using Remote-Sensing and Digital Terrain Data. *Theor. Appl. Climatol.* **1995**, *52*, 55–68. [CrossRef]
9. Ellingson, R.G. Surface longwave fluxes from satellite-observations—A critical-review. *Remote Sens. Environ.* **1995**, *51*, 89–97. [CrossRef]
10. Duarte, H.F.; Dias, N.L.; Maggiotto, S.R. Assessing daytime downward longwave radiation estimates for clear and cloudy skies in Southern Brazil. *Agric. For. Meteorol.* **2006**, *139*, 171–181. [CrossRef]
11. Wang, K.C.; Dickinson, R.E. Global atmospheric downward longwave radiation at the surface from ground-based observations, satellite retrievals, and reanalyses. *Rev. Geophys.* **2013**, *51*, 150–185. [CrossRef]
12. Bilbao, J.; De Miguel, A.H. Estimation of daylight downward longwave atmospheric irradiance under clear-sky and all-sky conditions. *J. Appl. Meteorol. Clim.* **2007**, *46*, 878–889. [CrossRef]
13. Niemela, S.; Raisanen, P.; Savijarvi, H. Comparison of surface radiative flux parameterizations—Part I: Longwave radiation. *Atmos. Res.* **2001**, *58*, 1–18. [CrossRef]
14. Ackerman, S.A.; Holz, R.E.; Frey, R.; Eloranta, E.W.; Maddux, B.C.; McGill, M. Cloud detection with MODIS. Part II: Validation. *J. Atmos. Ocean. Technol.* **2008**, *25*, 1073–1086. [CrossRef]
15. Crawford, T.M.; Duchon, C.E. An improved parameterization for estimating effective atmospheric emissivity for use in calculating daytime downwelling longwave radiation. *J. Appl. Meteorol.* **1999**, *38*, 474–480. [CrossRef]
16. Hollands, K.G.T.; Suehrcke, H. A three-state model for the probability distribution of instantaneous solar radiation, with applications. *Sol. Energy* **2013**, *96*, 103–112. [CrossRef]
17. Choi, M.H.; Jacobs, J.M.; Kustas, W.P. Assessment of clear and cloudy sky parameterizations for daily downwelling longwave radiation over different land surfaces in Florida, USA. *Geophys. Res. Lett.* **2008**, *35*. [CrossRef]
18. Kjaersgaard, J.H.; Plauborg, F.L.; Hansen, S. Comparison of models for calculating daytime long-wave irradiance using long term data set. *Agric. For. Meteorol.* **2007**, *143*, 49–63. [CrossRef]
19. Wang, K.C.; Liang, S.L. Global atmospheric downward longwave radiation over land surface under all-sky conditions from 1973 to 2008. *J. Geophys. Res. Atmos.* **2009**, *114*. [CrossRef]
20. Yang, K.; He, J.; Tang, W.J.; Qin, J.; Cheng, C.C.K. On downward shortwave and longwave radiations over high altitude regions: Observation and modeling in the Tibetan Plateau. *Agric. For. Meteorol.* **2010**, *150*, 38–46. [CrossRef]
21. Li, W.-B.; Wu, J.-B.; Wang, A.-Z.; Guan, D.-X.; Jin, C.-J. Applicability of daytime downward longwave radiation parameterized models in Changbai Mountains, Northeast China. *Ying Yong Sheng Tai Xue Bao J. Appl. Ecol.* **2015**, *26*, 497–504.
22. Yu, S.; Xin, X.; Liu, Q. Comparison of Atmospheric Downward Longwave Radiation Parameterizations. *Adv. Earth Sci.* **2011**, *26*, 751–762.
23. Yu, S.S.; Xin, X.Z.; Liu, Q.H.; Zhang, H.L. Estimation of Clear-Sky Downward Longwave Radiation from Satellite Data in Heihe River Basin of Northwest China. *Int. Geosci. Remote Sens.* **2011**, 269–272. [CrossRef]
24. Wei, Y.; Zhang, X.T.; Hou, N.; Zhang, W.Y.; Jia, K.; Yao, Y.J. Estimation of surface downward shortwave radiation over China from AVHRR data based on four machine learning methods. *Sol. Energy* **2019**, *177*, 32–46. [CrossRef]

25. Wang, Y.Z.; Jiang, B.; Liang, S.L.; Wang, D.D.; He, T.; Wang, Q.; Zhao, X.; Xu, J.L. Surface Shortwave Net Radiation Estimation from Landsat TM/ETM plus Data Using Four Machine Learning Algorithms. *Remote Sens.* **2019**, *11*, 2847. [CrossRef]

26. Wang, T.X.; Yan, G.J.; Chen, L. Consistent retrieval methods to estimate land surface shortwave and longwave radiative flux components under clear-sky conditions. *Remote Sens. Environ.* **2012**, *124*, 61–71. [CrossRef]

27. Ryu, Y.; Jiang, C.; Kobayashi, H.; Detto, M. MODIS-derived global land products of shortwave radiation and diffuse and total photosynthetically active radiation at 5 km resolution from 2000. *Remote Sens. Environ.* **2018**, *204*, 812–825. [CrossRef]

28. Hou, N.; Zhang, X.T.; Zhang, W.Y.; Wei, Y.; Jia, K.; Yao, Y.J.; Jiang, B.; Cheng, J. Estimation of Surface Downward Shortwave Radiation over China from Himawari-8 AHI Data Based on Random Forest. *Remote Sens.* **2020**, *12*, 181. [CrossRef]

29. Hao, D.L.; Asrar, G.R.; Zeng, Y.L.; Zhu, Q.; Wen, J.G.; Xiao, Q.; Chen, M. Estimating hourly land surface downward shortwave and photosynthetically active radiation from DSCOVR/EPIC observations. *Remote Sens. Environ.* **2019**, *232*. [CrossRef]

30. Cheng, J.; Liang, S.L. Global Estimates for High-Spatial-Resolution Clear-Sky Land Surface Upwelling Longwave Radiation From MODIS Data. *IEEE Trans. Geosci. Remote* **2016**, *54*, 4115–4129. [CrossRef]

31. Chen, J.; He, T.; Jiang, B.; Liang, S.L. Estimation of all-sky all-wave daily net radiation at high latitudes from MODIS data. *Remote Sens. Environ.* **2020**, *245*. [CrossRef]

32. Bristow, K.L.; Campbell, G.S. On the Relationship Between Incoming Solar-Radiation and Daily Maximum and Minimum Temperature. *Agric. For. Meteorol.* **1984**, *31*, 159–166. [CrossRef]

33. Dumas, A.; Andrisani, A.; Bonnici, M.; Graditi, G.; Leanza, G.; Madonia, M.; Trancossi, M. A new correlation between global solar energy radiation and daily temperature variations. *Sol. Energy* **2015**, *116*, 117–124. [CrossRef]

34. Hargreaves, G.H.; Samani, Z.A. Estimating potential evapo-transpiration. *J. Irrig. Drain. Div. ASCE* **1982**, *108*, 225–230.

35. Klink, J.C.; Dollhopf, K.J. An Evaluation of Satellite-Based Insolation Estimates for Ohio. *J. Clim. Appl. Meteorol.* **1986**, *25*, 1741–1751. [CrossRef]

36. Tarpley, J.D. Estimating Incident Solar-Radiation at the Surface from Geostationary Satellite Data. *J. Appl. Meteorol.* **1979**, *18*, 1172–1181. [CrossRef]

37. Dedieu, G.; Deschamps, P.Y.; Kerr, Y.H. Satellite Estimation of Solar Irradiance at the Surface of the Earth and of Surface Albedo Using a Physical Model Applied To Meteosat Data. *J. Clim. Appl. Meteorol.* **1987**, *26*, 79–87. [CrossRef]

38. Pinker, R.T.; Ewing, J.A. Modeling Surface Solar-Radiation—Model Formulation and Validation. *J. Clim. Appl. Meteorol.* **1985**, *24*, 389–401. [CrossRef]

39. Sridhar, V.; Elliott, R.L. On the development of a simple downwelling longwave radiation scheme. *Agric. For. Meteorol.* **2002**, *112*, 237–243. [CrossRef]

40. Yu, S.S.; Xin, X.Z.; Liu, Q.H.; Zhang, H.L.; Li, L. Comparison of Cloudy-Sky Downward Longwave Radiation Algorithms Using Synthetic Data, Ground-Based Data, and Satellite Data. *J. Geophys. Res. Atmos.* **2018**, *123*, 5397–5415. [CrossRef]

41. Zhou, Y.P.; Cess, R.D. Algorithm development strategies for retrieving the downwelling longwave flux at the Earth's surface. *J. Geophys. Res. Atmos.* **2001**, *106*, 12477–12488. [CrossRef]

42. Zhou, Y.P.; Kratz, D.P.; Wilber, A.C.; Gupta, S.K.; Cess, R.D. An improved algorithm for retrieving surface downwelling longwave radiation from satellite measurements. *J. Geophys. Res. Atmos.* **2007**, *112*. [CrossRef]

43. Lee, H.T.; Ellingson, R.G. Development of a nonlinear statistical method for estimating the downward longwave radiation at the surface from satellite observations. *J. Atmos. Ocean. Technol.* **2002**, *19*, 1500–1515. [CrossRef]

44. Tang, B.; Li, Z.L. Estimation of instantaneous net surface longwave radiation from MODIS cloud-free data. *Remote Sens. Environ.* **2008**, *112*, 3482–3492. [CrossRef]

45. Wang, J.; Tang, B.H.; Zhang, X.Y.; Wu, H.; Li, Z.L. Estimation of Surface Longwave Radiation over the Tibetan Plateau Region Using MODIS Data for Cloud-Free Skies. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2014**, *7*, 3695–3703. [CrossRef]

46. Wang, W.H.; Liang, S.L. Estimation of high-spatial resolution clear-sky longwave downward and net radiation over land surfaces from MODIS data. *Remote Sens. Environ.* **2009**, *113*, 745–754. [CrossRef]

47. Wang, W.H.; Liang, S.L. A Method for Estimating Clear-Sky Instantaneous Land-Surface Longwave Radiation With GOES Sounder and GOES-R ABI Data. *IEEE Geosci. Remote Sens.* **2010**, *7*, 708–712. [CrossRef]

48. Wang, W.H.; Liang, S.L.; Augustine, J.A. Estimating High Spatial Resolution Clear-Sky Land Surface Upwelling Longwave Radiation From MODIS Data. *IEEE Trans. Geosci. Remote* **2009**, *47*, 1559–1570. [CrossRef]

49. Ohmura, A.; Dutton, E.G.; Forgan, B.; Frohlich, C.; Gilgen, H.; Hegner, H.; Heimo, A.; Konig-Langlo, G.; McArthur, B.; Muller, G.; et al. Baseline Surface Radiation Network (BSRN/WCRP): New precision radiometry for climate research. *Bull. Am. Meteorol. Soc.* **1998**, *79*, 2115–2136. [CrossRef]

50. Zhang, T.P.; Stackhouse, P.W.; Gupta, S.K.; Cox, S.J.; Mikovitz, J.C.; Hinkelman, L.M. The validation of the GEWEX SRB surface shortwave flux data products using BSRN measurements: A systematic quality control, production and application approach. *J. Quant. Spectrosc. Radiat. Transf.* **2013**, *122*, 127–140. [CrossRef]

51. Liang, S.L.; Wang, K.C.; Zhang, X.T.; 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]

52. Mitchell, T.M. *Machine Learning*; McGraw-Hill: New York, NY, USA, 1997.

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

54. Vapnik, V.N. *The Nature of Statistical Learning Theory*; Springer: New York, NY, USA, 1995.

55. Smola, A.J.; Scholkopf, B. A tutorial on support vector regression. *Stat. Comput.* **2004**, *14*, 199–222. [CrossRef]

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

57. Johnson, R.; Zhang, T. Learning Nonlinear Functions Using Regularized Greedy Forest. *IEEE Trans. Pattern Anal. Mach. Intell.* **2014**, *36*, 942–954. [CrossRef]

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

59. Pal, M. Random forest classifier for remote sensing classification. *Int. J. Remote Sens.* **2005**, *26*, 217–222. [CrossRef]

60. Ksiazek, W.; Abdar, M.; Acharya, U.R.; Plawiak, P. A novel machine learning approach for early detection of hepatocellular carcinoma patients. *Cogn. Syst. Res.* **2019**, *54*, 116–127. [CrossRef]

61. Friedman, J.H. Multivariate Adaptive Regression Splines. *Ann. Stat.* **1991**, *19*, 1–67. [CrossRef]

62. Da Silva, I.N.; Hernane Spatti, D.; Andrade Flauzino, R.; Liboni, L.H.B.; Franco, D.R.A.S. *Artificial Neural Networks: A Practical Course*; Springer International Publishing: Cham, Switzerland, 2017.

63. Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; et al. Scikit-learn: Machine Learning in Python. *J. Mach. Learn. Res.* **2011**, *12*, 2825–2830.

64. Rizou, M.; Nnadi, F. Land use feedback on clear sky downward longwave radiation: A land use adapted model. *Int. J. Climatol.* **2007**, *27*, 1479–1496. [CrossRef]

65. Salles, T.; Goncalves, M.; Rodrigues, V.; Rocha, L. Improving random forests by neighborhood projection for effective text classification. *Inform. Syst.* **2018**, *77*, 1–21. [CrossRef]

66. Yang, F.; Cheng, J. A framework for estimating cloudy sky surface downward longwave radiation from the derived active and passive cloud property parameters. *Remote Sens. Environ.* **2020**, *248*. [CrossRef]