



Article

Estimating Rice Leaf Nitrogen Concentration: Influence of Regression Algorithms Based on Passive and Active Leaf Reflectance

Jia Sun ¹ , Jian Yang ^{2,*}, Shuo Shi ^{1,3,*}, Biwu Chen ¹ , Lin Du ², Wei Gong ^{1,3} and Shalei Song ⁴

¹ State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan 430079, Hubei, China; sunjia@whu.edu.cn (J.S.); cbw_think@whu.edu.cn (B.C.); liesmars_LiDAR@foxmail.com (W.G.)

² Faculty of Information Engineering, China University of Geosciences, Wuhan 430074, Hubei, China; linyufocus@foxmail.com

³ Collaborative Innovation Center of Geospatial Technology, Wuhan 430079, Hubei, China

⁴ Wuhan Institute of Physics and Mathematics, Chinese Academy of Sciences, Wuhan 430071, Hubei, China; songshalei@gmail.com

* Correspondence: wind_yang@whu.edu.cn (J.Y.); shishuo@whu.edu.cn (S.S.); Tel.: +86-130-1806-8732 (S.S.)

Received: 13 July 2017; Accepted: 12 September 2017; Published: 13 September 2017

Abstract: Nitrogen (N) is important for the growth of crops. Estimating leaf nitrogen concentration (LNC) accurately and nondestructively is important for precision agriculture, reduces environmental pollution, and helps model global carbon and N cycles. Leaf reflectance, especially in the visible and near-infrared regions, has been identified as a useful indicator of LNC. Except reflectance passively acquired by spectrometers, the newly developed multispectral LiDAR and hyperspectral LiDAR provide possibilities for measuring leaf spectra actively. The regression relationship between leaf reflectance spectra and rice (*Oryza sativa*) LNC relies greatly on the algorithm adopted. It would be preferable to find one algorithm that performs well with respect to passive and active leaf spectra. Thus, this study assesses the influence of six popular linear and nonlinear methods on rice LNC retrieval, namely, partial least-square regression, least squares boosting, bagging, random forest, back-propagation neural network (BPNN), and support vector regression of different types/kernels/parameter values. The R^2 , root mean square error and relative error in rice LNC estimation using these different methods were compared through the passive and active spectral measurements of rice leaves of different varieties at different locations and time (Yongyou 4949, Suizhou, 2014, Yangliangyou 6, Wuhan, 2015). Results demonstrate that BPNN provided generally satisfactory performance in estimating rice LNC using the three kinds of passive and active reflectance spectra.

Keywords: leaf nitrogen concentration; hyperspectral LiDAR; multispectral LiDAR; regression; machine learning

1. Introduction

Plants require nitrogen (N) nutrition for many important compounds, such as amino and nucleic acids [1,2]. Over a wide range, plant growth is linearly dependent on N supply [3]. Above that range, higher rates of fertilization do not necessarily improve yield and can cause serious water pollution [4]. Additionally, leaf nitrogen concentration (LNC) is associated with the photosynthetic capacity of leaves and thus allows global-scale CO₂ assimilation to be modelled based on specific leaf area and LNC [5,6]. There is a need for the N condition of plants to be determined efficiently and accurately for fertilization guidance and for understanding carbon and N cycles.

Traditional hand-held or chemically-determined LNC can be accurate but destructive or unrealistic for large-area needs. Meanwhile, various remote sensing techniques have been proposed as convenient and precise methods for detecting the physiological condition of plants [7,8]. The visible and red-edge region has been recognized as an important indicator of chlorophyll and N [9–11]. Apart from the traditional passive (using sunlight as the light source) reflectance acquired by spectrometers, laser intensity can also reflect the target's properties in an active (light source is provided by the sensor) way. Both techniques have their respective advantages and disadvantages. Passive spectrometers often cover a wide spectral range with high spectral resolution. However, they tend to be influenced by environmental factors, such as weather and sunlight conditions, and have a large field of view ($FOV > 50\text{ cm}^2$), depending at the range to target. By contrast, multispectral LiDAR (MSL) and hyperspectral LiDAR (HSL) strengthen the spectral detecting abilities of traditional single-wavelength LiDAR and simultaneously provide three-dimensional information and spectral properties. However, they are often limited to a few to dozens of channels [12–15]. The laser-induced fluorescence technology is also promising in detecting the biochemical conditions of plants [16].

Both passive and active reflectance spectra have been successfully applied in estimating the LNC of rice [17,18], which is an important crop. However, the regression result can be strongly influenced by the adopted machine learning algorithm [19,20]. Moreover, no algorithm presents a universally good performance, and the “best” regression method varies with the characteristics of the problem and the data set provided. Though deep learning has demonstrated outstanding abilities in studying internal data mechanisms [21,22], its demand for training data volume and computation hardware is high. Training samples can be expensive, especially for remote sensing applications.

The aim of this study is to investigate the performance of six popular machine learning algorithms in the context of detecting rice LNC with passive and active spectral data (i.e., ASD FieldSpec Pro FR (Analytical Spectral Devices, field spectroradiometer, full-range, Inc., Boulder, CO, USA), MSL, and HSL) and provide reference to find a comparatively “universal” regression algorithm for similar applications. This study also endeavors to offer guidance for choosing classification and regression algorithms for the development MSL/HSL systems. As the exact relationship between LNC and leaf reflectance remains unclear, both linear and nonlinear regression models were applied in this study, including partial least-square regression (PLSR), least squares boosting (LSBoost), bagging (Bag), random forest (RF), back-propagation neural network (BPNN), and support vector regression (SVR) of different types/kernels/parameter values. The R^2 , root mean square error (RMSE) and relative error (RE) in rice LNC estimation using these different methods were compared, using the passive and active measurements of rice of different varieties and at different locations and time, namely, Yongyou 4949 in Suizhou, 2014, and Yangliangyou 6 in Wuhan, 2015.

2. Materials and Methods

2.1. Study Sites Description

Yongyou 4949 and Yangliangyou 6 were the rice cultivars used in this study. They were cultivated at Junchuan County, Suizhou, and the experimental area of Huazhong Agricultural University in Wuhan, in 2014 and 2015, respectively (Figure 1). These two areas lie in Hubei Province ($108^{\circ}21'42''$ – $116^{\circ}07'50''$ E, $29^{\circ}01'53''$ – $33^{\circ}6'47''$ N), in the center of China. The climate of this area is subtropical monsoon humid, which is suitable for the growth of paddy rice.

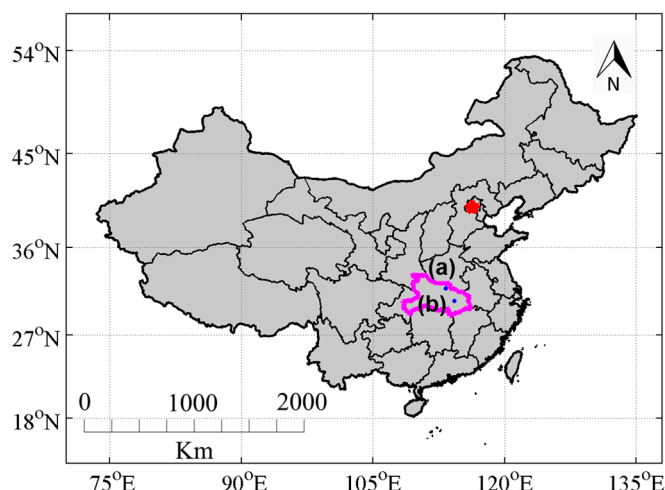


Figure 1. Locations of the two experimental fields: (a) Yongyou 4949, Junchuan County in Sui Zhou; (b) Yangliangyou 6, Huazhong Agricultural University in Wuhan.

For obtaining rice samples with a variety of LNCs, different levels of urea fertilizer treatments were applied in separate fields while controlling other cultivation conditions identically. In 2014, the crops were seeded on 27 April and transplanted on 1 June. Altogether six levels of urea fertilizer (0, 189, 229.5, 270, 310.5, and 351 kg/ha) were implemented, with 30% at seeding, 20% at tillering, 25% at shooting, and 25% at booting stage. In 2015, the crops were seeded on 30 April, and transplanted on 27 May. Four levels of urea fertilizer (0, 120, 180, 240 kg/ha) were implemented, with 60% at seeding, 20% at tillering and 20% at shooting. Three replicate fields were utilized for every urea fertilizer level. In each plot, at least six fully expanded second leaves from the top were selected randomly, obtaining a total of 220 leaves. The fresh leaf samples were sealed in plastic bags, kept in ice chests, and then transported to the laboratory. The LNC of the rice samples were determined through the Kjeldahl method [23] at the Wuhan Academy of Agricultural Science and Technology.

2.2. Passive and Active Spectra Measurement

The ASD FieldSpec Pro FR (Analytical Spectral Devices, Inc., Boulder, CO, USA) is a popular detector that can measure reflectance spectra from 400 to 2500 nm with a resampled 1 nm resolution. It has been widely utilized in remote sensing because of its convenience as a field spectrometer and high spectral resolution [24,25]. A strong correlation between foliar N and chlorophyll has been found in various plant species using different spectral indices [26–28]. The sensitivity of various biochemical parameters to spectra according to the PROSPECT model shows that chlorophyll content and leaf structure index prevail from 400 to 1000 nm [29]. Thus, the 400 to 1000 nm range was chosen for the ASD data in the analysis.

The two active sensors investigated were the MSL and the HSL system. The MSL system emits and detects light at four wavelengths: 556, 670, 700, and 780 nm. The HSL system has a white continuum light source and 32 channels that receive light from 538 to 910 nm, with a spectral resolution of 12 nm. Both systems were developed by Wuhan University and are laboratory-based prototypes. Further detailed description of the two systems are referred to in the works of Wei et al. [30] and Du et al. [31]. The spectral characteristics of rice samples were acquired by MSL, ASD, and HSL. For each leaf, three positions were randomly selected, and the average reflectance was later seen as a proxy for the reflectance of the whole leaf. A reference panel (Spectralon, Labsphere, Inc., North Sutton, NH, USA, reflectance nearly 99%) was measured before and after the spectral measurements by the three sensors in the laboratory under the same lighting conditions as the detected rice samples (lighted by a halogenated lamp of 50 W in the case of ASD and then by the MSL and HSL system respectively).

The return intensity of the reflected light must be calibrated against the measurements of a standard white reference panel. The reflectance at each wavelength can be obtained by dividing the raw return intensity value of the target by the averaged value of the reference panel. Thus, some systematic and random errors can be eliminated. The Savitzky–Golay smoothing filter [32] with a third-order polynomial function and a bandwidth of 25 nm was applied to the ASD data within the range of 400–1000 nm.

2.3. Analytical Methods

Principal Component Analysis: The ASD provides spectral data with hundreds of wavebands that provide high dimensionality that is often much higher than the number of available training samples. Additionally, significant redundancy is observed in high-dimensional spectral information. Principal component analysis (PCA) is a useful tool for reducing the correlation among high-dimensional data and can produce several principal components (PCs) that contain nearly as much information as the whole dataset [33]. Considering the noise and the redundancy of information from many wavelengths, the accuracy with hyperspectral data is not necessarily higher than that with the first few PCs.

In the process of reducing the dimensionality of ASD spectra, the eigen vectors and eigen values of the covariance matrix composed of multi-dimensional data were first calculated. The data vectors were mapped from the original space to a new orthogonal space using the PCs corresponding to the first few eigen values. In this study, the first three PCs and the hyperspectral data collected by ASD were used as the regression model input (explaining > 99% variance). PCA transformation was not performed on the active spectra because the spectral dimensions of MSL and HSL were not as large as the samples.

Partial least-square regression: As a popular linear model, PLSR has been applied in a wide range of fields, including remote sensing. PLSR is a good alternative to many traditional multiple linear regression and PC regression methods because of its robustness [34]. Similar to PCs regression, a PLSR model tries to find the multidimensional direction in the independent space that explains the maximum multidimensional variance direction in the dependent space (y) in an iterative manner [35]. The final model that predicts the p dependent variables y (in this case, p equals 1, LNC) has the following form:

$$y_j = c_{j1}x_1 + c_{j2}x_2 + \dots + c_{jm}x_m, \quad j = 1, 2, \dots, p, \quad (1)$$

where x_1 to x_m are the standardized PCs calculated on mean-centered variables.

Least squares boosting and Bagging: The exact relationship between LNC and reflectance spectra remains unclear, and is possibly nonlinear. With this consideration, more nonlinear regression algorithms were employed in this study.

LSBoost was developed by Friedman [36] based on the squared loss function. It is a kind of boosting regression method. Boosting creates an ensemble of models that collectively make a prediction to improve performance in the case of the rice LNC. The ensemble is composed of base learners. Each base learner minimizes the residual error of the previous one iteratively with weighted samples. The LSBoost algorithm was chosen because it has a solid mathematical foundation compared with other boosting algorithms. LSBoost begins with an initial guess f_0 and then fits a sequence of M weighted models of T_1 to T_M (decision tree as the base learners in this study) [37]. The final model has the following form:

$$f_M(x) = f_0(x) + v \times \sum_{m=1}^M \rho_m T_m(x), \quad (2)$$

where ρ_m denotes the weight for base learner m and v denotes the learning rate. The final output is the weighted average output of all the base learners.

Both boosting and bagging [38] build a set of base learners that are combined by voting. Bagging builds a set of base learners by generating replicated bootstrap data, and boosting does so by adjusting the weights of training samples [39]. Bagging and boosting have several differences. First, the training

samples of bagging are selected randomly and independently, whereas those of boosting are related to the result of the previous learning. Second, each of the base learners in boosting has its respective weight, whereas in bagging the weights are equal. Third, the base learners of bagging can be generated in a parallel manner, instead of sequentially as that in boosting.

In this study, the LSBoost and Bag regressions were executed with the “fitensemble” function in MATLAB R2015b (Mathworks Inc., Natick, MA, USA), for fitted ensemble for regression. Decision trees were used as the individual models that form the ensemble, as is often adopted [37]. Decision trees select important input dimensions in its calibration process. Different numbers of regression trees were investigated to establish the LSBoost and Bag regression models. Each setting was repeated 50 times, and the average was taken for model performance assessment.

Random forest: Proposed by Breiman [40], RF added an additional layer of randomness to bagging. In addition to constructing each tree using a different bootstrap sample of the data, RF changes how regression trees are constructed. In standard trees, each node is split using the best split among all variables. In contrast, each node in RF is split using the best among a subset of randomly selected predictors at that node [41]. This strategy improves the model’s generalization performance and its training efficiency. RF has three parameters: the number of regression trees, the number of predictors to select randomly for each decision split, and the number of predictor splits summed over all trees [42]. The defaults of the latter two parameters in the routines were accepted in the analysis, and the number of trees was tested. Each setting was repeated 50 times, and the average was calculated for model performance assessment.

Back-propagation neural network: As a kind of feedforward network, BPNN has been widely applied in solving various nonlinear problems [43,44]. It has the advantage of self-learning and self-adaption with generally good performance [45]. BPNN was applied in this study to model rice LNC and reflectance. BPNN is based on the strategy of gradient descent and adjusts the parameters toward the negative gradient direction.

A neural network consists of two or more layers, each with a number of neurons. Each hidden and output neuron in BPNN processes its inputs by multiplying each by the respective weight, summing the product, and passing the sum through a nonlinear transfer function for the final result. The weights of the neurons are modified in response to the errors between the actual and target output values. Training is performed by repeatedly updating the weights at the end of each cycle until the average sum squared error over all the training samples are minimized and within the tolerance specified [46]. More details can be referred to in [47,48]. The regression performance was investigated with different training functions and hidden layer sizes. Each setting was repeated 50 times and the average was obtained for model performance assessment.

Support vector regression: SVR can construct both linear and nonlinear regressions using different kernels. Different from an artificial neural network (ANN), SVR has an excellent generalization performance with a strong theoretical foundation in statistical learning theory [49]. SVR is insensitive to the dimension number of training samples and requires a small number of training samples [50].

SVR fits linear regression in the high-dimension feature space and attempts to reduce model complexity by minimizing the empirical risk. SVR has two types: ε -SVR and ν -SVR. ε -SVR estimates an unknown continuous-valued function based on a finite set of noisy samples [51]. The optimization problem of SVR has the following form:

$$\min_{w,b,\xi,\xi^*} \frac{1}{2} w^T w + C \times \frac{1}{l} \sum_{i=1}^l (\xi_i + \xi_i^*), \quad (3)$$

where C is the constant that determines the trade-off, ξ and ξ^* are the slack variables, l is the loss function, and $w^T w$ indicates the model complexity.

In v-SVR, the size of accuracy parameter ε is traded off against model complexity and slack variables via a constant v [52]. Its constrained optimization problem is as follows:

$$\min_{w,b,\xi,\xi^*} \frac{1}{2} w^T w + C \times (v\varepsilon + \frac{1}{l} \sum_{i=1}^l (\xi_i + \xi_i^*)), \quad (4)$$

In SVR, four kinds of kernels can be chosen: linear, polynomial, RBF, and sigmoid. The latter three kernels establish nonlinear relationships. In this work, all data preprocessing, regression, and evaluation was executed in MATLAB R2015b (Mathworks Inc., Natick, MA, USA), where many regression toolbox and LIBSVM [53] are available. The type of transfer function/kernel function and the associated parameters have to be determined in advance for BPNN and SVR. Thus, the optimization process using different kernels or transfer functions and main parameter with varying values were presented, while other parameters were settled through five-fold cross validation and are not presented below.

2.4. Statistical Parameters

A total of 220 rice samples were collected in 2014 and 2015. They were divided randomly into two datasets: 80% (176) as the training dataset and the remaining 20% (44) as the validation dataset for predicting LNC. The coefficient of determination (R^2), root mean square error (RMSE), and relative error (RE) were calculated as follows to evaluate the performance of the estimation models:

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^n (\hat{y}_i - y_i)^2}{n}}, \quad (5)$$

$$\text{RE} = \frac{100}{\bar{y}} \text{RMSE}, \quad (6)$$

where \hat{y}_i , y_i , and \bar{y} are the estimated, observed, and average observed rice LNC, respectively, and n is the number of samples.

3. Results

3.1. Principal Components Analysis of ASD Spectra

ASD acquires the reflected signals of the targets in a wide range with high spectral resolution. Figure 2 shows that the first three PCs of ASD reflectance contained most of the spectral characteristics. Among them, PC1 was dominated by the red-edge (around 700 nm) and the near-infrared region, with a small slope in the green light domain around 530 nm. PC2 and PC3 contained several complementary spectral features to PC1. PC2 is generally significant in the visible region, covering the absorption spectrum of chlorophylls and carotenoids [54]. PC3 represents the derivative of the reflectance spectra. The first three PCs extracted the most important information in the reflectance spectra (cumulative variance > 99%). In the regression process, both hyperspectral data and PC1–3 were applied in every regression algorithm except for PLSR, because the process of extracting the PCs was originally incorporated in the method. The results showed that for every investigated algorithm except for LSBoost and SVR, the first three PCs showed a stronger correlation with foliar N than the hyperspectral data. This is because LSBoost and SVR have an advantage in dealing with high-dimensional data [50,55]. Hence, the results using the ASD hyperspectral data instead of PCs for these two algorithms are shown in Section 3.3.

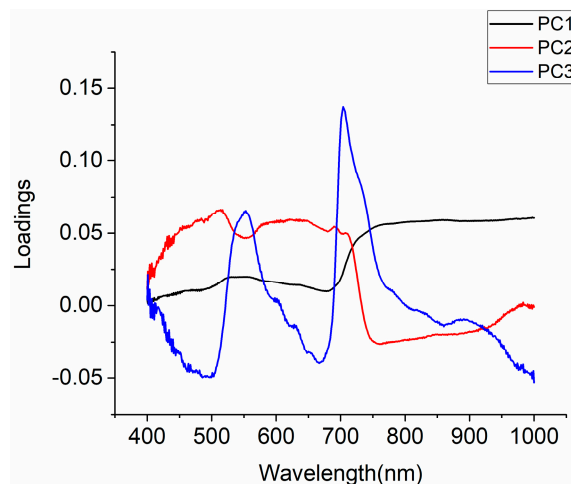


Figure 2. Loading weights of the first three principal components (PCs) of the ASD reflectance spectra.

3.2. Linear Regression Algorithms

PLSR and SVR with a linear kernel are the two linear regression algorithms investigated in this study. In the optimization process, when the contribution of factors in PLSR was above 90%, the influence of this parameter on the result was marginal. For SVR regression, v-SVR showed a superior ability to ϵ -SVR, regardless of the spectra used (Figure 3).

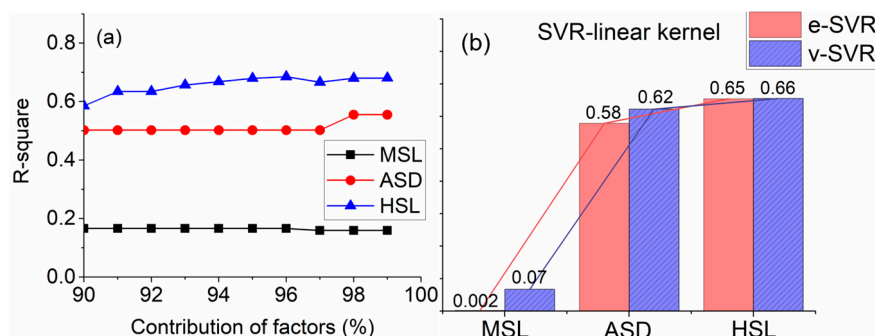


Figure 3. The estimation R^2 of LNC using reflectance spectra from MSL, ASD, and HSL with two linear regression algorithms—(a) partial-least square regression (PLSR) with varying contribution of factors and (b) support vector regression (SVR) with a linear kernel (with different types of SVR). Other parameters were settled through 5-fold validation.

Figure 3 shows the performance of PLSR with different factor contributions, and ϵ -SVR and v-SVR with a linear kernel. Figure 4 is scatters of observed and predicted rice LNC using the two linear algorithms with optimal parameters. Figures 3 and 4 show that the MSL data showed a weak correlation with foliar N using PLSR ($R^2 = 0.17$, RMSE = 5.08). Linear regression is not enough to characterize the relationship between rice LNC and MSL, which has only four wavelengths. This was confirmed by the results of SVR with a linear kernel. Meanwhile, the relationships of ASD and HSL with LNC were stronger than that of MSL. Our results also show that the estimation precision of LNC based on ASD was comparable with that of HSL using SVR with a linear kernel ($R^2 = 0.62$, 0.66 , RMSE = 3.56, 3.46, respectively).

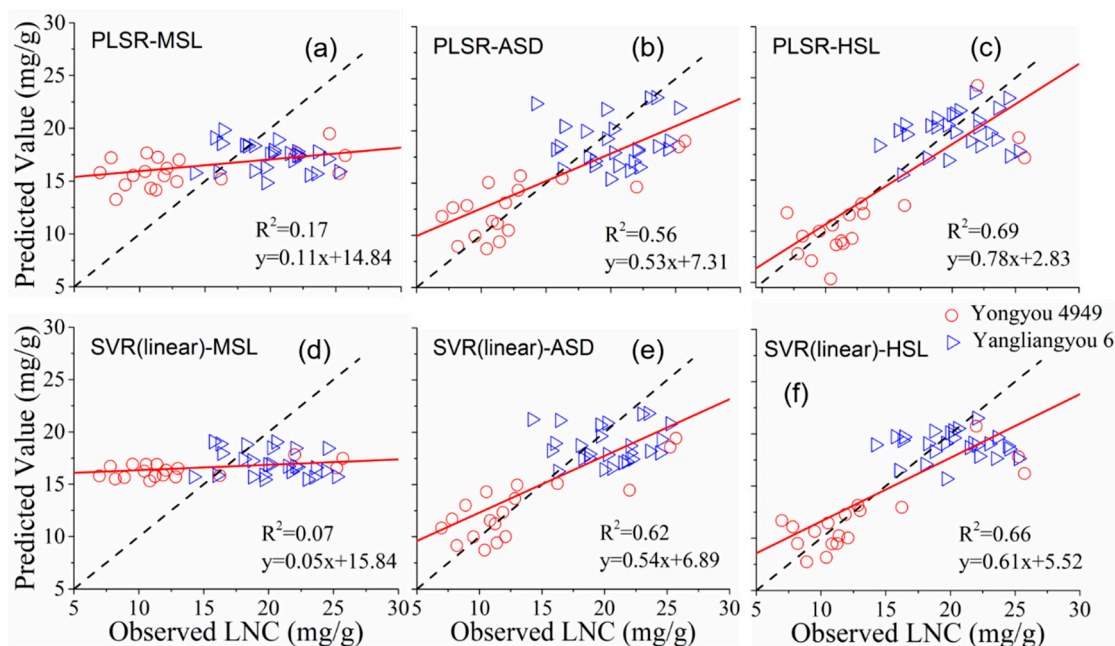


Figure 4. Relationship between the observed leaf nitrogen concentration (LNC) and the predicted LNC based on passive and active detectors with PLSR ((a–c) are the results using spectra from MSL, ASD, HSL, respectively) and SVR (linear kernel, (d–f) are the results using spectra from MSL, ASD, HSL, respectively) (the dashed line represents 1:1 line) based on validation dataset ($n = 44$).

3.3. Nonlinear Regression Algorithms

Figure 5 shows the process of parameter, transfer function, and kernel optimization using different nonlinear algorithms (LSBoost, Bag, RF, BPNN, and SVR).

Optimization process of algorithms of least squares boosting (LSBoost), bagging (Bag), random forest (RF), and SVR using R^2 as a measure (ASD, MSL, and HSL). Other parameters were settled through 5-fold validation.

As an ensemble learning algorithm, LSBoost uses several base learners (decision tree in this case) to improve the performance of weak learners. HSL requires comparatively fewer decision trees to provide a stable estimation of rice LNC, whereas 170 trees are optimally required by MSL and ASD. Above this threshold, more decision trees did not provide a more accurate LNC estimation. Using the LSBoost algorithm, passive and active spectra showed a stronger relationship to LNC than PLSR and SVR with a linear kernel. Hence, this initially suggests that the relationship between spectral data and foliar N is nonlinear rather than linear.

Fewer base learners were required for Bag and RF to provide a satisfactory LNC estimation than the order of magnitude in LSBoost. Moreover, the fold lines of R^2 with different base learner numbers were rather flat. The results show that the relationship for reflectance spectra and foliar N using Bag was comparable with that using RF.

BPNN has 13 different varieties of transfer functions: trainbfg, trainbr, traincgb, traincgf, traincgp, traingd, traingda, traingdm, traingdx, trainlm, trainoss, trainrp, and trainscg. Until now, no settled criterion for selecting a transfer function exists because the choice should be based on specific circumstances. The performance of different transfer functions deviated greatly with the use of the spectral data from different sensors (Figure 5). However, transfer functions traingd and traingdm produced the worst LNC estimation for the reflectance spectra from every detector, with a lower R^2 than any other transfer function ($R^2 < 0.5$). Except for these two transfer functions, the performance of the different transfer functions was similar for ASD regardless of the number of hidden layers. The circumstance for HSL data was similar, except for trainlm, whose performance degraded with the

HSL with SVR, SVR, and BPNN, respectively. In summary, BPNN exhibited significant abilities in building the relationship between passive and active reflectance spectra and rice LNC.

Table 1. Regression performance of the six algorithms with optimal parameters (PLSR, LSBoost, Bag, RF, BPNN, and SVR; unit of RMSE: mg/g).

Sensor	PLSR			LSBoost			Bag		
	R ²	RMSE	RE	R ²	RMSE	RE	R ²	RMSE	RE
MSL	0.17	5.08	29.37	0.45	4.16	24.07	0.58	3.66	21.17
ASD	0.56	3.76	21.75	0.61	3.47	20.03	0.72	3.21	18.53
HSL	0.69	3.27	18.89	0.71	3.27	18.88	0.78	2.75	15.89
Sensor	Random Forest			BPNN			SVR		
MSL	0.59	3.64	21.04	0.69	3.48	20.09	0.70	3.35	19.38
ASD	0.72	3.21	18.54	0.77	3.09	17.88	0.73	2.94	16.98
HSL	0.79	2.74	15.86	0.78	2.68	15.50	0.73	2.92	16.88

4. Discussion

The in vivo specific absorption coefficient of N is hard to get, meaning that LNC can only be detected in an indirect way. Considering the close relations between leaf chlorophylls and N concentration, LNC can be detected through the reflective characteristics of chlorophylls in visible and NIR regions in leaf reflectance [58,59]. Thus, by using machine learning algorithms to study the relationship between leaf spectra and LNC, ideally, the scatter points of predicted versus observed LNC would fall along the 1:1 line. However, apart from chlorophylls, a small portion of N also resides in material such as protein, nucleic acid, and auxin etc. [3]. Systematic and random noise have an influence on the spectra. Also, the performance of different regression algorithms and their ability to analyze the relationship between leaf spectra and LNC determines the regression results. Therefore, the fitting line of the scatter deviates from the 1:1 line, and to different extents with respect to different detectors and regression algorithms. The nearer to the 1:1 line indicates a better regression result.

Boosting and bagging have a sound theoretical base as classifiers and regression methods. In the experiments of Quinlan [39] with various datasets, boosting appeared to be a more effective method in classifying than bagging. However, in this study, bagging generally presented more accurate predictions of LNC than LSBoost, regardless of the dataset used. This could be caused by the difference in the sensitivity to noise of the two methods. Tzeng et al. [60] found that boosting tends to perform poorly in terms of accuracy when there is noise in the data. In contrast, bagging is not very sensitive to noise in the data.

The RF algorithm has become increasingly popular in many different applications [61]. RF is an improvement of bagging. However, it performed similarly with Bag in this study in terms of accuracy, and its performance was more stable with a varying number of base learners.

The feed-forward BPNN is the type of neural network that is most commonly used in remote sensing [45]. BPNN provided the best overall performance in terms of accuracy estimating rice LNC among all the investigated algorithms in this study, that is, highest for ASD and second highest for MSL and HSL compared with other methods. Neural approaches have been shown to be more accurate than other techniques, and are insensitive to incomplete or noisy data [46]. However, overfitting is also often encountered because of its powerful representation ability. Early stopping and regularization help address this problem [62]. Ensuring that the global minimum instead of a local minimum is found by means of simulated annealing and genetic algorithms is important [63,64].

SVR presented rice LNC estimation using MSL data with the highest precision, and ASD as the second highest compared with other algorithms. With a clear theoretical foundation, SVM is favored in remote sensing classification and regression [65,66]. However, the regression relationship between HSL and LNC established by v-SVR using kernel RBF was not as good as those of Bag, RF, and BPNN. Theoretically, SVR has a more powerful transferability than BPNN because it aims at minimizing the

structured risk and has a global optimum. A dataset that covers a wider range is required to test whether the good performance of BPNN over SVR stands in terms of predicting vegetation LNC.

Table 1 and Figures 2 and 5 show that the spectra measured by HSL generally presented a stronger relationship to rice LNC than MSL and is stronger or comparable with ASD. Fusing passive images and LiDAR data introduce sophisticated problems of registration [67]. With more bands in the system, the spectral characteristics of targets can be accurately characterized and the full potential of the backscattered intensity of laser light can be gained. Nevalainen et al. [68] investigated several vegetation indexes measured by a HSL system in reflecting nitrogen in oat samples ($R^2 = 0.757$, RMSE = 0.37 mg/g). The exploration was promising but tentative because of limited samples. A dual-wavelength system was employed in assessing crop foliar nitrogen with better results than a single green laser wavelength ($R^2 = 0.47$ versus 0.72) [13] and can be improved with more laser wavelengths. The detector ASD has more bands (601 in this study versus 32) with a finer spectral resolution than HSL. In contrast, LiDAR has a smaller FOV, and the laser return allows different types of surfaces to be resolved in the vertical dimension, resulting in a more accurate detection, while the radiance measured using ASD is a function of the total vertical column [67].

More samples of different rice varieties across locations are required to test the transferability of BPNN in the accurate estimation of rice LNC. Several researchers have proposed using multiple kernels learning in SVR to deal with problems where single-kernel may be insufficient [69,70]. Our next step is to extend the experience with the MSL and HSL systems in leaf spectra toward canopy spectra, where the circumstance is much more sophisticated. Many factors have an influence on the canopy reflectance that have to be considered, such as the leaf area index, leaf inclination distribution, relative leaf size, and soil background, etc. [71]. In addition, it should be noted that it is not always possible to retrieve the actual leaf constituents from leaf or canopy reflectance. This is because very similar reflectance spectra can be obtained from very different combinations of input parameters (leaf constituents, viewing angles, canopy architecture and soil background, etc.), leading to the well-known “ill-posed” problem [29,72]. Though the inverse problem is by nature ill-posed, some measures can be taken to help alleviate it, such as restriction of ranges of input parameters and using prior information [24,73]. One of the future directions of MSL/HSL is to make it more compact and mounted on vehicles and UAVs for wider application.

5. Conclusions

This study investigated the performance of six popular linear and nonlinear regression algorithms to predict paddy rice LNC using the spectra from three different passive and active sensors: ASD, MSL, and HSL. Furthermore, the influence of parameter values, transfer functions, kernels, and regression types on the estimation results at a leaf scale from the reflectance spectra was verified.

The results demonstrate that BPNN has a greater ability to characterize the relationship between reflectance spectra and rice LNC than other investigate algorithms and provided generally satisfactory results on all passive and active spectra (R^2 : 0.69–0.78). The HSL constructed a regression relationship with higher R^2 compared with MSL and higher or comparable results to ASD regardless of the algorithm adopted.

The results also show that HSL is capable of estimating vegetation biochemical properties. The advantage of the HSL over the MSL is obvious. Further work will be performed to extend the results of different datasets and locations and investigate the canopy scale.

Acknowledgments: This study was supported by the National Natural Science Foundation of China (Grant No. 41601360; 41611130114), Wuhan Morning Light Plan of Youth Science and Technology (2017050304010308), Natural Science Foundation of Hubei Province (Grant No. 2015CFA002), the Fundamental Research Funds for the Central Universities (Grant No. 2042016kf0008), the Fundamental Research Funds for the Central Universities, China University of Geosciences (Wuhan) (Grant No. CUG170661), and the Open Research Fund of State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing (Grant No. 15R01). We gratefully acknowledge the open access to the LIBSVM.

Author Contributions: W.G. designed the study and supervised the project. S.S. selected the sampling sites. J.S., S.S., J.Y., L.D., and B.W.C. carried out the sampling and performed the experimental work. J.S. analyzed the data and wrote the manuscript. S.S., J.Y., and S.L.S. edited the manuscript. All authors discussed the results and contributed to the manuscript.

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

References

1. Lam, H.-M.; Coschigano, K.; Oliveira, I.; Melo-Oliveira, R.; Coruzzi, G. The molecular-genetics of nitrogen assimilation into amino acids in higher plants. *Annu. Rev. Plant Biol.* **1996**, *47*, 569–593. [[CrossRef](#)] [[PubMed](#)]
2. Egli, M.; Saenger, W. *Principles of Nucleic Acid Structure*; Springer Science & Business Media: Berlin, Germany, 2013.
3. Springer. Plant ecology. *Plant Ecol.* **2012**, *213*, 175–183.
4. Carpenter, S.R.; Caraco, N.F.; Correll, D.L.; Howarth, R.W.; Sharpley, A.N.; Smith, V.H. Nonpoint pollution of surface waters with phosphorus and nitrogen. *Ecol. Appl.* **1998**, *8*, 559–568. [[CrossRef](#)]
5. Poorter, H.; Evans, J.R. Photosynthetic nitrogen-use efficiency of species that differ inherently in specific leaf area. *Oecologia* **1998**, *116*, 26–37. [[CrossRef](#)] [[PubMed](#)]
6. Wong, S.-C.; Cowan, I.R.; Farquhar, G.D. Leaf conductance in relation to rate of CO₂ assimilation I. Influence of nitrogen nutrition, phosphorus nutrition, photon flux density, and ambient partial pressure of CO₂ during ontogeny. *Plant Physiol.* **1985**, *78*, 821–825. [[CrossRef](#)] [[PubMed](#)]
7. Zhu, X.; Wang, T.; Darvishzadeh, R.; Skidmore, A.K.; Niemann, K.O. 3d leaf water content mapping using terrestrial laser scanner backscatter intensity with radiometric correction. *ISPRS J. Photogramm.* **2015**, *110*, 14–23. [[CrossRef](#)]
8. Wang, Q.; Li, P. Hyperspectral indices for estimating leaf biochemical properties in temperate deciduous forests: Comparison of simulated and measured reflectance data sets. *Ecol. Indic.* **2012**, *14*, 56–65. [[CrossRef](#)]
9. Yu, K.; Lenz-Wiedemann, V.; Chen, X.; Bareth, G. Estimating leaf chlorophyll of barley at different growth stages using spectral indices to reduce soil background and canopy structure effects. *ISPRS J. Photogramm.* **2014**, *97*, 58–77. [[CrossRef](#)]
10. Hunt, E.R.; Doraiswamy, P.C.; McMurtrey, J.E.; Daughtry, C.S.; Perry, E.M.; Akhmedov, B. A visible band index for remote sensing leaf chlorophyll content at the canopy scale. *Int. J. Appl. Earth Obs. Geoinf.* **2013**, *21*, 103–112. [[CrossRef](#)]
11. Ramoelo, A.; Skidmore, A.K.; Cho, M.A.; Schlerf, M.; Mathieu, R.; Heitkönig, I.M. Regional estimation of savanna grass nitrogen using the red-edge band of the spaceborne rapideye sensor. *Int. J. Appl. Earth Obs. Geoinf.* **2012**, *19*, 151–162. [[CrossRef](#)]
12. Gong, W.; Sun, J.; Shi, S.; Yang, J.; Du, L.; Zhu, B.; Song, S. Investigating the potential of using the spatial and spectral information of multispectral LiDAR for object classification. *Sensors* **2015**, *15*, 21989–22002. [[CrossRef](#)] [[PubMed](#)]
13. Eitel, J.U.; Magney, T.S.; Vierling, L.A.; Dittmar, G. Assessment of crop foliar nitrogen using a novel dual-wavelength laser system and implications for conducting laser-based plant physiology. *ISPRS J. Photogramm.* **2014**, *97*, 229–240. [[CrossRef](#)]
14. Nevalainen, O.; Hakala, T.; Suomalainen, J.; Mäkipää, R.; Peltoniemi, M.; Krooks, A.; Kaasalainen, S. Fast and nondestructive method for leaf level chlorophyll estimation using hyperspectral LiDAR. *Agric. For. Meteorol.* **2014**, *198*, 250–258. [[CrossRef](#)]
15. Chen, B.; Shi, S.; Gong, W.; Zhang, Q.; Yang, J.; Du, L.; Sun, J.; Zhang, Z.; Song, S. Multispectral LiDAR point cloud classification: A two-step approach. *Remote Sens.* **2017**, *9*, 373. [[CrossRef](#)]
16. Yang, J.; Du, L.; Sun, J.; Zhang, Z.; Chen, B.; Shi, S.; Gong, W.; Song, S. Estimating the leaf nitrogen content of paddy rice by using the combined reflectance and laser-induced fluorescence spectra. *Opt. Express* **2016**, *24*, 19354–19365. [[CrossRef](#)] [[PubMed](#)]
17. Sun, J.; Shi, S.; Gong, W.; Yang, J.; Du, L.; Song, S.; Chen, B.; Zhang, Z. Evaluation of hyperspectral LiDAR for monitoring rice leaf nitrogen by comparison with multispectral LiDAR and passive spectrometer. *Sci. Rep.* **2017**, *7*, 40362. [[CrossRef](#)] [[PubMed](#)]

18. Kokaly, R.F.; Clark, R.N. Spectroscopic determination of leaf biochemistry using band-depth analysis of absorption features and stepwise multiple linear regression. *Remote Sens. Environ.* **1999**, *67*, 267–287. [[CrossRef](#)]
19. Yang, J.; Sun, J.; Du, L.; Chen, B.; Zhang, Z.; Shi, S.; Gong, W. Effect of fluorescence characteristics and different algorithms on the estimation of leaf nitrogen content based on laser-induced fluorescence LiDAR in paddy rice. *Opt. Express* **2017**, *25*, 3743–3755. [[CrossRef](#)] [[PubMed](#)]
20. Du, L.; Shi, S.; Yang, J.; Sun, J.; Gong, W. Using different regression methods to estimate leaf nitrogen content in rice by fusing hyperspectral LiDAR data and laser-induced chlorophyll fluorescence data. *Remote Sens.* **2016**, *8*, 526. [[CrossRef](#)]
21. LeCun, Y.; Bengio, Y.; Hinton, G. Deep learning. *Nature* **2015**, *521*, 436–444. [[CrossRef](#)] [[PubMed](#)]
22. Schmidhuber, J. Deep learning in neural networks: An overview. *Neural Netw.* **2015**, *61*, 85–117. [[CrossRef](#)] [[PubMed](#)]
23. Lang, C.A. Simple microdetermination of kjeldahl nitrogen in biological materials. *Anal. Chem.* **1958**, *30*, 1692–1694. [[CrossRef](#)]
24. Ali, A.M.; Darvishzadeh, R.; Skidmore, A.K.; van Duren, I.; Heiden, U.; Heurich, M. Estimating leaf functional traits by inversion of prospect: Assessing leaf dry matter content and specific leaf area in mixed mountainous forest. *Int. J. Appl. Earth Obs. Geoinf.* **2016**, *45*, 66–76. [[CrossRef](#)]
25. Zhao, F.; Guo, Y.; Huang, Y.; Reddy, K.N.; Lee, M.A.; Fletcher, R.S.; Thomson, S.J. Early detection of crop injury from herbicide glyphosate by leaf biochemical parameter inversion. *Int. J. Appl. Earth Obs. Geoinf.* **2014**, *31*, 78–85. [[CrossRef](#)]
26. Clevers, J.G.; Gitelson, A.A. Remote estimation of crop and grass chlorophyll and nitrogen content using red-edge bands on Sentinel-2 and-3. *Int. J. Appl. Earth Obs. Geoinf.* **2013**, *23*, 344–351. [[CrossRef](#)]
27. Baret, F.; Houlès, V.; Guérif, M. Quantification of plant stress using remote sensing observations and crop models: The case of nitrogen management. *J. Exp. Bot.* **2007**, *58*, 869–880. [[CrossRef](#)] [[PubMed](#)]
28. Cammarano, D.; Fitzgerald, G.J.; Casa, R.; Basso, B. Assessing the robustness of vegetation indices to estimate wheat N in mediterranean environments. *Remote Sens.* **2014**, *6*, 2827–2844. [[CrossRef](#)]
29. Li, P.; Wang, Q. Retrieval of leaf biochemical parameters using prospect inversion: A new approach for alleviating ill-posed problems. *IEEE Trans. Geosci. Remote Sens.* **2011**, *49*, 2499–2506.
30. Wei, G.; Shalei, S.; Bo, Z.; Shuo, S.; Faquan, L.; Xuewu, C. Multi-wavelength canopy LiDAR for remote sensing of vegetation: Design and system performance. *ISPRS J. Photogramm.* **2012**, *69*, 1–9. [[CrossRef](#)]
31. Du, L.; Gong, W.; Shi, S.; Yang, J.; Sun, J.; Zhu, B.; Song, S. Estimation of rice leaf nitrogen contents based on hyperspectral LiDAR. *Int. J. Appl. Earth Obs. Geoinf.* **2016**, *44*, 136–143. [[CrossRef](#)]
32. Savitzky, A.; Golay, M.J. Smoothing and differentiation of data by simplified least squares procedures. *Anal. Chem.* **1964**, *36*, 1627–1639. [[CrossRef](#)]
33. Jolliffe, I. *Principal Component Analysis*; Wiley Online Library: Hoboken, NJ, USA, 2002.
34. Geladi, P.; Kowalski, B.R. Partial least-squares regression: A tutorial. *Anal. Chim. Acta* **1986**, *185*, 1–17. [[CrossRef](#)]
35. Hansen, P.; Schjoerring, J. Reflectance measurement of canopy biomass and nitrogen status in wheat crops using normalized difference vegetation indices and partial least squares regression. *Remote Sens. Environ.* **2003**, *86*, 542–553. [[CrossRef](#)]
36. Friedman, J.H. Greedy function approximation: A gradient boosting machine. *Ann. Stat.* **2001**, *29*, 1189–1232. [[CrossRef](#)]
37. Heijst, D.V.; Potharst, R.; Wezel, M.V. A support system for predicting ebay end prices. *Decis. Support Syst.* **2008**, *44*, 970–982. [[CrossRef](#)]
38. Breiman, L. *Bagging Predictors*; Kluwer Academic Publishers: Dordrecht, The Netherlands, 1996; pp. 123–140.
39. Quinlan, J.R. Bagging, boosting, and c4.5. In Proceedings of the Thirteenth National Conference on Artificial Intelligence and Eighth Innovative Applications of Artificial Intelligence Conference, Portland, OR, USA, 4–8 August 1996; pp. 725–730.
40. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [[CrossRef](#)]
41. Liaw, A.; Wiener, M. Classification and regression by randomforest. *R News* **2002**, *2*, 18–22.
42. Mutanga, O.; Adam, E.; Cho, M.A. High density biomass estimation for wetland vegetation using worldview-2 imagery and random forest regression algorithm. *Int. J. Appl. Earth Obs. Geoinf.* **2012**, *18*, 399–406. [[CrossRef](#)]

43. Pradhan, B.; Lee, S. Regional landslide susceptibility analysis using back-propagation neural network model at cameron highland, Malaysia. *Landslides* **2010**, *7*, 13–30. [[CrossRef](#)]
44. Heermann, P.D.; Khazenie, N. Classification of multispectral remote sensing data using a back-propagation neural network. *IEEE Trans. Geosci. Remote Sens.* **1992**, *30*, 81–88. [[CrossRef](#)]
45. Atkinson, P.M.; Tatnall, A.R.L. Introduction neural networks in remote sensing. *Int. J. Remote Sens.* **1997**, *18*, 699–709. [[CrossRef](#)]
46. Goh, A. Back-propagation neural networks for modeling complex systems. *Artif. Intell. Eng.* **1995**, *9*, 143–151. [[CrossRef](#)]
47. Caudill, M.; Butler, C. *Naturally Intelligent Systems*; MIT Press: Cambridge, MA, USA, 1992.
48. Eberhart, R.C. *Neural Network Pc Tools: A Practical Guide*; Academic Press: Salt Lake City, UT, USA, 2014.
49. Wu, X.; Kumar, V.; Quinlan, J.R.; Ghosh, J.; Yang, Q.; Motoda, H.; McLachlan, G.J.; Ng, A.; Liu, B.; Philip, S.Y. Top 10 algorithms in data mining. *Knowl. Inf. Syst.* **2008**, *14*, 1–37. [[CrossRef](#)]
50. Joachims, T. *Text Categorization with Support Vector Machines: Learning with Many Relevant Features*; Springer: Berlin, Germany, 1998.
51. Smola, A.J.; Schölkopf, B. A tutorial on support vector regression. *Stat. Comput.* **2004**, *14*, 199–222. [[CrossRef](#)]
52. Schölkopf, B.; Smola, A.J.; Williamson, R.C.; Bartlett, P.L. New support vector algorithms. *Neural Comput.* **2000**, *12*, 1207–1245. [[CrossRef](#)] [[PubMed](#)]
53. Chang, C.-C.; Lin, C.-J. Libsvm: A library for support vector machines. *ACM Trans. Intell. Syst. Technol. (TIST)* **2011**, *2*, 27. [[CrossRef](#)]
54. Gitelson, A.A.; Gritz, Y.; Merzlyak, M.N. Relationships between leaf chlorophyll content and spectral reflectance and algorithms for non-destructive chlorophyll assessment in higher plant leaves. *J. Plant Physiol.* **2003**, *160*, 271–282. [[CrossRef](#)] [[PubMed](#)]
55. Lutz, R.W.; Bühlmann, P. Boosting for high-multivariate responses in high-dimensional linear regression. *Stat. Sin.* **2006**, *16*, 471–494.
56. Crisp, D.J.; Burges, C.J.C. A geometric interpretation of v-svm classifiers. In *Advances in Neural Information Processing Systems*; MIT Press: Cambridge, MA, USA, 1999; pp. 244–250.
57. Yao, X.; Huang, Y.; Shang, G.; Zhou, C.; Cheng, T.; Tian, Y.; Cao, W.; Zhu, Y. Evaluation of six algorithms to monitor wheat leaf nitrogen concentration. *Remote Sens.* **2015**, *7*, 14939–14966. [[CrossRef](#)]
58. Kokaly, R.F. Investigating a physical basis for spectroscopic estimates of leaf nitrogen concentration. *Remote Sens. Environ.* **2001**, *75*, 153–161. [[CrossRef](#)]
59. Chen, P.; Haboudane, D.; Tremblay, N.; Wang, J.; Vigneault, P.; Li, B. New spectral indicator assessing the efficiency of crop nitrogen treatment in corn and wheat. *Remote Sens. Environ.* **2010**, *114*, 1987–1997. [[CrossRef](#)]
60. Tzeng, Y.C.; Chiu, S.H.; Chen, K.S. Improvement of remote sensing image classification accuracy by using a multiple classifiers system with modified bagging and boosting algorithms. In *Proceedings of the IEEE International Conference on Geoscience and Remote Sensing Symposium*, Denver, CO, USA, 31 July–4 August 2006; pp. 2769–2772.
61. Genuer, R.; Poggi, J.-M.; Tuleau-Malot, C. Variable selection using random forests. *Pattern Recognit. Lett.* **2010**, *31*, 2225–2236. [[CrossRef](#)]
62. Girosi, F.; Jones, M.; Poggio, T. *Regularization Theory and Neural Networks Architectures*; MIT Press: Cambridge, MA, USA, 1995; pp. 219–269.
63. Goldberg, D.E. *Genetic Algorithms in Search, Optimization and Machine Learning*; Addison-Wesley Longman Publishing Co., Inc.: Boston, MA, USA, 1988; Volume xiii, pp. 2104–2116.
64. Aarts, E.; Korst, J. Simulated annealing and boltzmann machines: A stochastic approach to combinatorial optimization and neural computing. *Siam Rev.* **1989**, *xii*, 323.
65. Ma, Y.; Gong, W.; Wang, P.; Hu, X. New dust aerosol identification method for spaceborne LiDAR measurements. *J. Quant. Spectrosc. Radiat. Transf.* **2011**, *112*, 338–345. [[CrossRef](#)]
66. Yang, J.; Gong, W.; Shi, S.; Du, L.; Sun, J.; Ma, Y.; Song, S. Accurate identification of nitrogen fertilizer application of paddy rice using laser-induced fluorescence combined with support vector machine. *Plant Soil Environ.* **2015**, *61*, 501–506. [[CrossRef](#)]
67. Eitel, J.U.; Höfle, B.; Vierling, L.A.; Abellán, A.; Asner, G.P.; Deems, J.S.; Glennie, C.L.; Joerg, P.C.; LeWinter, A.L.; Magney, T.S. Beyond 3-d: The new spectrum of LiDAR applications for earth and ecological sciences. *Remote Sens. Environ.* **2016**, *186*, 372–392. [[CrossRef](#)]

68. Nevalainen, O.; Hakala, T.; Suomalainen, J.; Kaasalainen, S. Nitrogen concentration estimation with hyperspectral LiDAR. *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci.* **2013**, *2*, 205–210. [[CrossRef](#)]
69. Gehler, P.; Nowozin, S. On feature combination for multiclass object classification. In Proceedings of the 2009 IEEE 12th International Conference on Computer Vision, Kyoto, Japan, 29 September–2 October 2009; pp. 221–228.
70. Yeh, C.-Y.; Huang, C.-W.; Lee, S.-J. A multiple-kernel support vector regression approach for stock market price forecasting. *Expert Syst. Appl.* **2011**, *38*, 2177–2186. [[CrossRef](#)]
71. Jacquemoud, S.; Verhoef, W.; Baret, F.; Bacour, C.; Zarco-Tejada, P.J.; Asner, G.P.; François, C.; Ustin, S.L. Prospect+ sail models: A review of use for vegetation characterization. *Remote Sens. Environ.* **2009**, *113*, S56–S66. [[CrossRef](#)]
72. Combal, B.; Baret, F.; Weiss, M.; Trubuil, A.; Mace, D.; Pragnere, A.; Myneni, R.; Knyazikhin, Y.; Wang, L. Retrieval of canopy biophysical variables from bidirectional reflectance: Using prior information to solve the ill-posed inverse problem. *Remote Sens. Environ.* **2003**, *84*, 1–15. [[CrossRef](#)]
73. Wang, Y.; Li, X.; Nashed, Z.; Zhao, F.; Yang, H.; Guan, Y.; Zhang, H. Regularized kernel-based brdf model inversion method for ill-posed land surface parameter retrieval. *Remote Sens. Environ.* **2007**, *111*, 36–50. [[CrossRef](#)]



© 2017 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).