

Article

# Estimating the Biomass of Maize with Hyperspectral and LiDAR Data

Cheng Wang <sup>1,2</sup>, Sheng Nie <sup>2,3,\*</sup>, Xiaohuan Xi <sup>2</sup>, Shezhou Luo <sup>2</sup> and Xiaofeng Sun <sup>3</sup>

<sup>1</sup> Beijing Advanced Innovation Center for Imaging Technology, Capital Normal University, Beijing 100094, China; wangcheng@radi.ac.cn

<sup>2</sup> Key Laboratory of Digital Earth Science, Institute of Remote Sensing and Digital Earth, Chinese Academy of Sciences, Beijing 100094, China; xixh@radi.ac.cn (X.X.); luosz@radi.ac.cn (S.L.)

<sup>3</sup> University of Chinese Academy of Sciences, Beijing 100049, China; sxf0709@163.com

\* Correspondence: niesheng@radi.ac.cn; Tel.: +86-108-217-8120

Academic Editors: Jixian Zhang, Xiangguo Lin, Nicolas Baghdadi, Richard Gloaguen and Prasad S. Thenkabail  
Received: 5 September 2016; Accepted: 21 December 2016; Published: 27 December 2016

**Abstract:** The accurate estimation of crop biomass during the growing season is very important for crop growth monitoring and yield estimation. The objective of this paper was to explore the potential of hyperspectral and light detection and ranging (LiDAR) data for better estimation of the biomass of maize. First, we investigated the relationship between field-observed biomass with each metric, including vegetation indices (VIs) derived from hyperspectral data and LiDAR-derived metrics. Second, the partial least squares (PLS) regression was used to estimate the biomass of maize using VIs (only) and LiDAR-derived metrics (only), respectively. Third, the fusion of hyperspectral and LiDAR data was evaluated in estimating the biomass of maize. Finally, the biomass estimates were validated by a leave-one-out cross-validation (LOOCV) method. Results indicated that all VIs showed weak correlation with field-observed biomass and the highest correlation occurred when using the red edge-modified simple ratio index (ReMSR). Among all LiDAR-derived metrics, the strongest relationship was observed between coefficient of variation ( $H_{CV}$ ) of digital terrain model (DTM) normalized point elevations with field-observed biomass. The combination of VIs through PLS regression could not improve the biomass estimation accuracy of maize due to the high correlation between VIs. In contrast, the  $H_{CV}$  combined with  $H_{mean}$  performed better than one LiDAR-derived metric alone in biomass estimation ( $R^2 = 0.835$ ,  $RMSE = 374.655 \text{ g/m}^2$ ,  $RMSE_{CV} = 393.573 \text{ g/m}^2$ ). Additionally, our findings indicated that the fusion of hyperspectral and LiDAR data can provide better biomass estimates of maize ( $R^2 = 0.883$ ,  $RMSE = 321.092 \text{ g/m}^2$ ,  $RMSE_{CV} = 337.653 \text{ g/m}^2$ ) compared with LiDAR or hyperspectral data alone.

**Keywords:** airborne LiDAR; hyperspectral data; biomass; maize; partial least squares (PLS) regression

## 1. Introduction

Crop biomass serves as one of the critical inputs in growth monitoring and yield estimation models, and plays an important role in agricultural, ecological, and meteorological applications [1–5]. It is, therefore, essential to accurately estimate crop biomass. However, accurate estimation of crop biomass remains a challenging task. Field measurement techniques, such as destructive sampling, can accurately obtain crop biomass. However, they are labor-intensive and time-consuming, and not suitable for large-scale spatial and temporal measurements [6].

Unlike field measurement techniques, remote sensing is an effective technology for estimating crop biomass at the regional or global scale [1,7,8]. Many previous studies have successfully estimated crop biomass using traditional optical remote sensing data [2,6,9,10]. Both physical and statistical models were developed to estimate crop biomass. The physical models are complex, thus, only a few previous

studies conducted crop biomass estimation using the inversion of physical models [11,12]. In contrast, spectral vegetation indices (VIs) were commonly used to estimate crop biophysical variables due to the simplicity and practicality [13–17]. However, the estimation accuracy of crop biomass is usually affected by background soil reflectance, atmospheric and water absorption using broadband VIs. The emergence of hyperspectral data overcame the limitation by adopting distinctive narrow bands [18]. Hyperspectral sensors can provide fine spectral resolution bands and abundant information regarding the biochemical and biophysical composition of vegetation canopy [19–22]. In fact, hyperspectral data has been successfully employed in crop biomass estimation by many previous studies. However, the estimation accuracy of crop biomass is generally limited due to the asymptotic saturation problem, which is common with passive optical sensors [9,18,22,23]. Additionally, passive optical sensors record electromagnetic energy which is mostly reflected or absorbed in the uppermost canopy layers [24,25]. In this case, optical remote sensing sensors provide limited information of vegetation vertical structure, the estimation accuracy of crop biomass may, therefore, be affected.

Different from optical sensors, Light Detection and Ranging (LiDAR) is an active remote sensing technology capable of providing three-dimensional information on vegetation canopy structure [26,27]. Moreover, it is able to overcome the saturation problem of optical remote sensing. Therefore, LiDAR can provide effective data for estimating vegetation structure parameters. LiDAR data has been successfully used to estimate forest biomass in many previous studies. Both height metrics and laser penetration indices, derived from LiDAR returns, are commonly employed in forest biomass estimation [3,28–30]. However, the application of airborne discrete-return LiDAR data is generally limited in environments with short, dense vegetation maybe due to the following two reasons [20,31,32]. First, dense vegetation limits laser penetration to the ground, which may be lower the estimation accuracy of vegetation biophysical variables. Second, the discrete-return LiDAR data provides limited information about the vertical vegetation structure because it only records one return for each emitted pulse in areas with short vegetation [31–33]. Finally, airborne LiDAR data cannot provide spectral characteristics of vegetation canopy because they work with a single-wavelength.

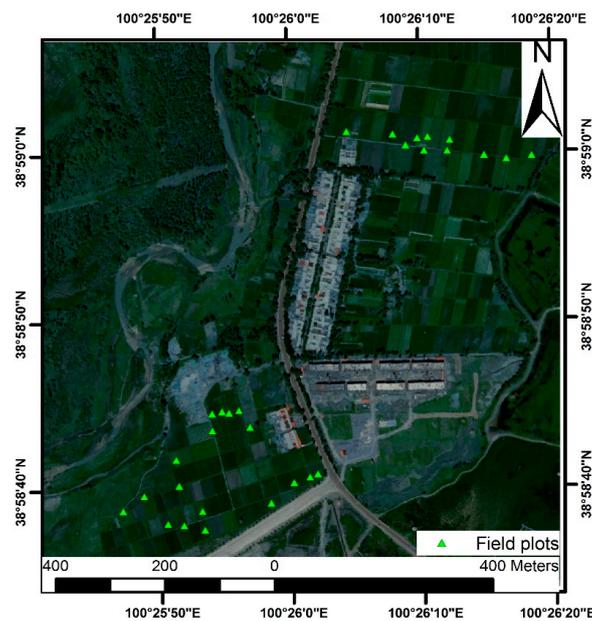
In summary, airborne LiDAR data provides detailed vertical structural information of vegetation canopy and hyperspectral data reflects abundant spectral information of the vegetation canopy. In this case, the fusion of LiDAR and hyperspectral data appears to be an interesting option for better estimating vegetation biomass. The previous studies have demonstrated the capability of fused hyperspectral and LiDAR data for providing robust vegetation biomass estimates in forest ecosystems [25,29,34–36], even in herbaceous environments [22,37]. However, there is no literature that has focused on the study of crop biomass estimation through the fusion of hyperspectral and LiDAR data. This study can, therefore, serve as a new attempt to evaluate the fusion of LiDAR and hyperspectral data in estimating crop biomass.

The objective of this paper is to explore the potential of hyperspectral and LiDAR data in better estimating biomass of maize. More specifically, our goals were: (1) to examine the relationship between field-observed biomass with VIs derived from hyperspectral data; (2) to evaluate the utility of LiDAR-derived metrics for biomass estimation; and (3) to determine whether the fusion of LiDAR and hyperspectral data will improve the biomass estimation accuracy of maize.

## 2. Materials

### 2.1. Study Area

The experiment was carried out in Heihe River Basin, Gansu province of China (100°20'E and 100°30'E, 38°50'N, and 39°03'N, Figure 1). The average annual precipitation and temperature are 140 mm and 7 °C in the middle reaches of the Heihe River, respectively. The terrain in this experiment site is relatively flat, and the average elevation is about 1403 m above sea level. The dominated vegetation is cropland with maize being the major crop, occupying 95% of whole study area.



**Figure 1.** Location of the study area and filed plots.

## 2.2. Airborne LiDAR Data

Airborne LiDAR data were acquired in July 2012 using a Leica Airborne Laser Scanner (ALS70) system (Leica, Lucerne, Switzerland). The average flying height was 1300 m above the ground level with an average point density of 7.4 points/m<sup>2</sup>. The LiDAR data was post processed to calculate LiDAR-derived metrics. It mainly consists of three separate stages: outlier removal, point cloud classification, and digital terrain model (DTM) generation. Outliers in the original LiDAR datasets were first removed using a statistical method based on the elevation frequency histogram of point clouds [38]. Then LiDAR point clouds were classified as canopy and ground returns using the adaptive triangulation network (TIN) filtering algorithm developed by Axelsson embedded in the TerraScan software (TerraSolid, Ltd., Helsinki, Finland) [39]. A DTM was finally generated from LiDAR ground returns with a grid cell size of 1 m. We removed the influence of topography through using the generated DTM and obtained the DTM normalized point clouds. The LiDAR-derived metrics can, therefore, be calculated from DTM normalized point clouds.

## 2.3. Airborne Hyperspectral Data

Compact airborne spectrographic imager (CASI) imagery was acquired in July 2012, and provided by “Heihe Watershed Allied Telemetry Experimental Research” (HiWATER) [40]. CASI data were collected in “Hyperspectral Mission” mode, with 72 channels of approximately 7.5 nm bandwidth across a wavelength range of 400–940 nm. The final spatial resolution was at 1 m after radiometric correction and ortho-rectification based on a LiDAR-derived DTM. Although the radiation calibration and geometric correction of CASI imagery have been carried out by the data provider, atmospheric correction must be conducted on original hyperspectral bands to accurately calculate VIs. In this study, atmospheric correction of the CASI imagery was performed using the Fast Line-of-Sight Atmospheric Analysis of Spectral Hypercubes (FLAASH) algorithm [41], which is based on a MODTRAN4 approach for path scattered radiance, absorption, and adjacency effects [42]. In addition, a list of other preprocessing steps, such as removal of aircraft motion effects, and flat field adjustments of surface reflectance spectra, were also conducted using ENVI software in this study.

## 2.4. Field Measurement

The field measurements were carried out from 11 July 2012 to 13 July 2012. There are a total of 38 sample plots with a size of 4.0 m × 4.0 m square across the whole study area. However, only 29 sample plots were analyzed in this study due to the bad quality of hyperspectral data. For each plot, the center coordinate was accurately determined using a 1 cm level real-time kinematic Global Positioning System (RTK-GPS) (Trimble Navigation Ltd., Sunnyvale, CA, USA). Additionally, both the height and LAI of maize were also directly measured in this study (see Table 1). Field LAIs were measured using an LAI-2200 (LI-Cor Inc., Lincoln, NE, USA). The heights of all maize plant within the same plot were measured using a measuring tape, and the average height was regarded as reference height of maize. The field-observed biomass of maize can be calculated using Equation (1), which was obtained by long-term observation of the entire growing season of maize according to the study of Gao et al., [10].

$$\text{Biomass} = 1.01 \times \text{height} \times \exp\left(1.05 \times \text{LAI}\right), \left(R^2 = 0.88\right) \quad (1)$$

where height is the mean height, LAI is the leaf area index, and Biomass is the field-observed biomass of maize.

**Table 1.** Summary statistics of field-measured parameters in field plots.

	Min	Max	Mean	SD
Height (cm)	121.82	309.09	228.62	0.52
LAI	1.59	4.69	2.88	0.9
Biomass (g/m <sup>2</sup> )	126.12	3428.78	1078.61	912.08

## 3. Methods

### 3.1. Deriving Metrics from Hyperspectral Data

Many previous studies have investigated a list of VIs derived from hyperspectral data for crop biomass estimation [1,7,17,43]. Based on the findings of these studies, six VIs were selected to estimate biomass of maize in this study (Table 2): Simple Ratio Vegetation Index (SR) [44], Normalized Difference Vegetation Index (NDVI) [45], Modified Simple Ratio Index (MSR) [24], Modified Soil-Adjusted Vegetation Index (MSAVI) [46], Red Edge NDVI (ReNDVI) [47] and Red Edge Modified Simple Ratio Index (ReMSR) [15]. For each plot, the VIs of each pixel was calculated from the CASI imager, and all VIs within the same plot were averaged.

**Table 2.** Definitions and formulas of VIs in this study.

Index	Abbr.	Formula
Simple Ratio Vegetation Index	SR	$SR = \frac{R_{800}}{R_{670}}$
Normalized Difference Vegetation Index	NDVI	$NDVI = \frac{R_{800} - R_{670}}{R_{800} + R_{670}}$
Modified Simple Ratio Index	MSR	$MSR = \left(\frac{R_{800}}{R_{670}} - 1\right) / \sqrt{\left(\frac{R_{800}}{R_{670}} + 1\right)}$
Modified Soil-Adjusted Vegetation Index	MSAVI	$MSAVI = 0.5 \times [2R_{800} + 1 - \sqrt{(2R_{800} + 1)^2 - 8(R_{800} - R_{670})}]$
Red Edge NDVI	ReNDVI	$ReNDVI = \frac{R_{750} - R_{705}}{R_{750} + R_{705}}$
Red Edge Modified Simple Ratio Index	ReMSR	$ReMSR = \left(\frac{R_{750}}{R_{705}} - 1\right) / \sqrt{\left(\frac{R_{750}}{R_{705}} + 1\right)}$

Note that there is no band centered at 705 nm for the CASI imager; two adjacent bands were chosen to calculate the reflectance  $R_{705}$ . R is the abbreviation of reflectance.

### 3.2. Deriving Metrics from LiDAR

A list of LiDAR-derived metrics, including height metrics and laser penetration indices, were adopted to estimate the biomass of maize in this study according to the previous studies [3,6,25,32]. These metrics consists of: the mean ( $H_{mean}$ ) and maximum ( $H_{max}$ ) height of LiDAR returns, the standard deviation ( $H_{SD}$ ) and coefficient of variation ( $H_{CV}$ ) of normalized DTM points' elevations, and canopy cover (CC) derived from airborne LiDAR data. The CC is calculated as the ratio of the number of canopy returns and all returns. The definition and explanation of these LiDAR-derived metrics were listed in Table 3.

**Table 3.** Definitions and formulas of LiDAR-derived metrics in this study.

Index	Abbr.	Formula
Max height	$H_{max}$	$H_{max} = \max(h_i), 1 \leq i \leq N$
Mean height	$H_{mean}$	$H_{mean} = \frac{1}{N} \sum_{i=1}^N h_i$
Standard deviation	$H_{SD}$	$H_{SD} = \sqrt{\frac{\sum_{i=1}^N (h_i - H_{mean})^2}{N - 1}}$
Coefficient of variation	$H_{CV}$	$H_{CV} = \frac{H_{SD}}{H_{mean}}$
Canopy cover	CC	$CC = \frac{N_{canopy}}{N_{all}}$

Note:  $h_i$  is the height of  $i$ -th point, and  $N$  is the total number of LiDAR points in the plot.

### 3.3. Regression Analysis

In this study, the relationship between field-observed biomass with each metric was first examined using both the linear and exponential regression analyses according to the previous studies [17,48]. These metrics included VIs derived from hyperspectral data and LiDAR-derived metrics.

Additionally, partial least squares (PLS) regression was implemented with VIs (only), LiDAR-derived metrics (only) showed in Table 3, and all metrics, respectively. Different from multiple linear regression, PLS regression is a multivariate statistical method which can effectively deal with the problem of multicollinearity among predictor variables [25,49]. It has been successfully employed in vegetation biomass estimation. In the process of PLS regression, variance inflation factors (VIFs) were simultaneously calculated to evaluate importance of individual predictors. Only predictor variables with VIF values greater than 1.0 were considered according to the previous study [25].

For all biomass estimation models, both the coefficient of determination ( $R^2$ ) and root-mean-square error (RMSE) were calculated to assess the established models. In addition, the leave-one-out cross-validation (LOOCV) is used to evaluate the generalization capability of regression models because there is no additional data for model validation [50,51]. Specifically, the root mean square error from the cross validation analysis ( $RMSE_{CV}$ ), calculated from the predicted residual sum of squares (PRESS statistic) and the number of observations, was used to validate the accuracies of biomass estimation models.

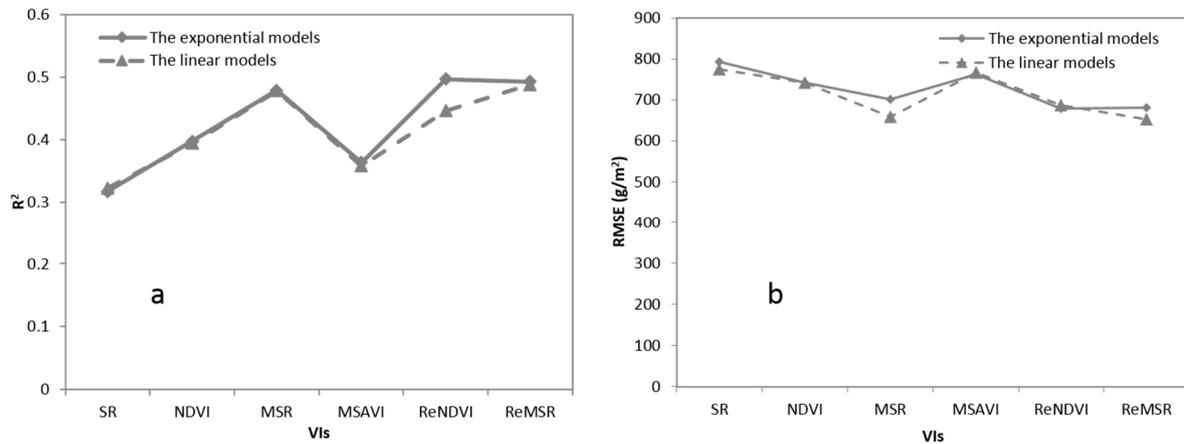
## 4. Results

### 4.1. VIs as Predictors of Vegetation Biomass

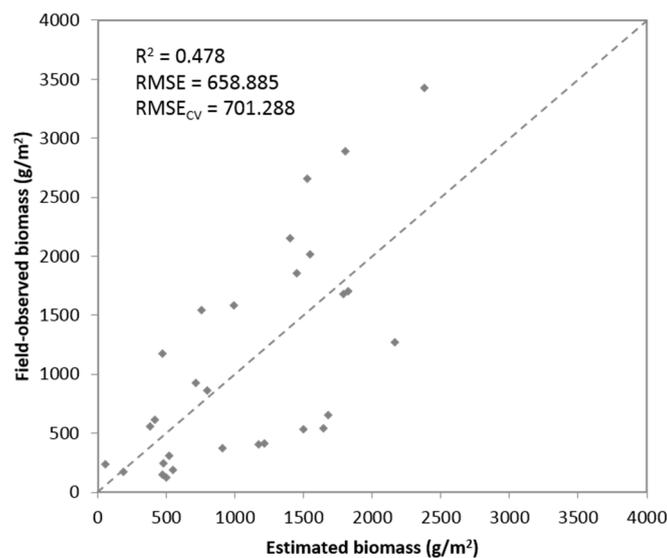
To examine the relationship of biomass against VIs, both the linear and exponential regression analysis were carried out in this study. The  $R^2$  and RMSE of biomass estimation models using different VIs are showed in Figure 2.

From Figure 2, all VIs showed a weak relationship with field-observed biomass due to high RMSE and low  $R^2$ . Among all VIs, ReMSR had the lowest RMSE, and the highest  $R^2$  for biomass estimation of maize regardless of linear or exponential models.

Since VIs showed a high degree of multi-collinearity (see Table 4), when they were used in PLS regression, they did not improve the predictability considerably. Therefore, the best biomass estimation model using VIs is built based on ReMSR alone. Figure 3 showed the scatterplot of field-observed biomass versus estimated biomass using VIs alone.



**Figure 2.** The determination coefficient  $R^2$  (a) and root-mean-squared error (RMSE) (b) of biomass estimation models using different VIs.



**Figure 3.** The scatterplot of field-observed biomass versus estimated biomass using VIs (only).

**Table 4.** Pearson’s correlation between VIs themselves.

VIs	SR	NDVI	MSR	MSAVI	ReNDVI	ReMSR
SR	1.000	0.986	0.893	0.996	0.896	0.900
NDVI	0.986	1.000	0.953	0.996	0.951	0.959
MSR	0.893	0.953	1.000	0.927	0.969	0.995
MSAVI	0.996	0.996	0.927	1.000	0.925	0.933
ReNDVI	0.896	0.951	0.969	0.925	1.000	0.989
ReMSR	0.900	0.959	0.995	0.933	0.989	1.000

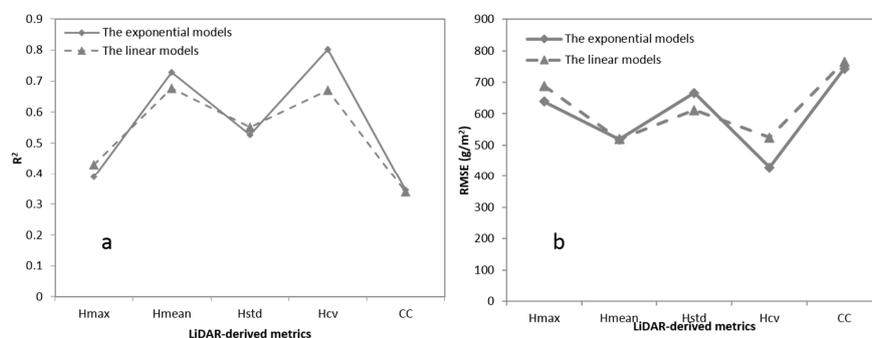
#### 4.2. LiDAR-Derived Metrics as Predictors of Vegetation Biomass

The vertical accuracies of DTM greatly affected the calculation of LiDAR-derived metrics, and then limited the biomass estimation accuracy of maize. Thus, it is essential to assess the accuracy of LiDAR-derived DTM. The elevation differences were first calculated by subtracting the elevations of GPS points from the corresponding DTM, then the mean error (ME) and RMSE were calculated to assess the vertical accuracy of LiDAR-derived DTM. Results showed that the LiDAR-derived DTM in our experiment site has a high accuracy with the ME and RMSE of 0.07 m and 0.09 m, respectively.

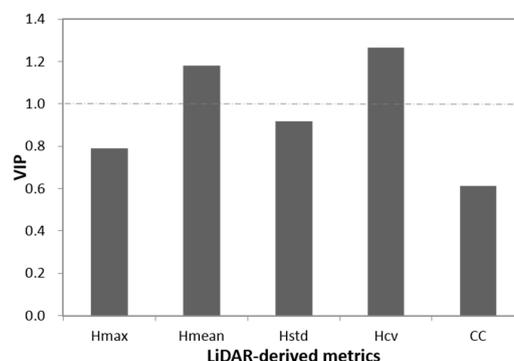
Additionally, we performed linear and exponential regression analyses of field-observed biomass against each LiDAR-derived metric in this study. The variation of  $R^2$  and RMSE using different LiDAR-derived metrics were shown in Figure 4. Among all of the LiDAR-derived metrics,  $H_{CV}$  provided the best agreement with field-observed biomass in an exponential model and explained 80.1% of the variance in vegetation biomass (RMSE = 429.144 g/m<sup>2</sup>). This was followed by  $H_{mean}$  ( $R^2 = 0.729$ , RMSE = 518.509 g/m<sup>2</sup>) with an exponential model. However, the CC was not sensitive to field-observed biomass regardless of the linear or exponential regression model due to the high RMSE and low  $R^2$ .

Similarly, the PLS regression method was also used to estimate the biomass of maize using LiDAR-derived metrics. The VIP values of all LiDAR-derived metrics were calculated using the PLS regression (as shown in Figure 5). Only the LiDAR-derived metrics with VIP values greater than 1.0 were considered in this study. In this case, the best estimation model is constructed based on  $H_{CV}$  and  $H_{mean}$ . Figure 6 showed the scatterplot of field-observed biomass versus estimated biomass using LiDAR-derived metrics alone. The combined use of LiDAR-derived metrics performed better than one LiDAR-derived metric alone in estimating biomass of maize.

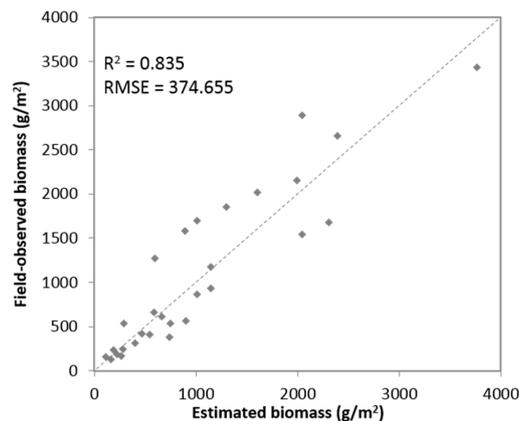
From Figures 3 and 6, we can concluded that LiDAR data had a stronger capability than hyperspectral data in estimating the biomass of maize due to the higher  $R^2$  and lower RMSE.



**Figure 4.** The determination coefficient  $R^2$  (a) and root mean squared error (RMSE) (b) of biomass estimation using different LiDAR-derived metrics.



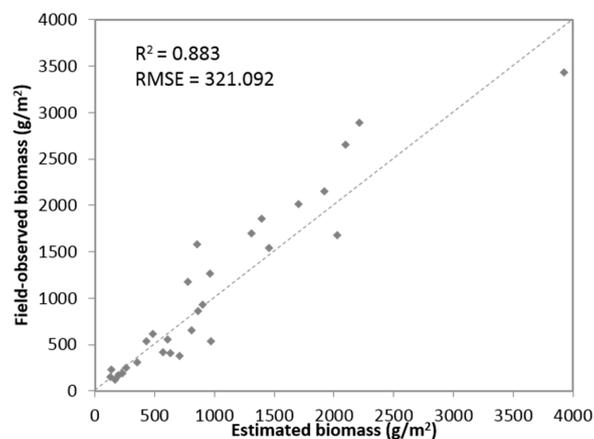
**Figure 5.** VIP values of all LiDAR-derived metrics for biomass estimation.



**Figure 6.** The scatterplot of field-observed biomass versus estimated biomass using LiDAR-derived metrics (only).

#### 4.3. Fusion of LiDAR and Hyperspectral Data for Vegetation Biophysical Variables Predictions

To evaluate the fusion of LiDAR and hyperspectral data in estimating the biomass of maize, the PLS regression analysis of field-observed biomass against all metrics including VIs and LiDAR-derived metrics was conducted in this study. Figure 7 showed the scatterplot of field-observed biomass versus estimated biomass using all metrics. From Figure 7, the synergistic use of hyperspectral and LiDAR data explained 88.3% of variance (RMSE = 321.092) in biomass measurements, whereas the best models using LiDAR and hyperspectral data alone explained 83.5% (RMSE = 374.655) and 47.8% (RMSE = 658.885) of variance in biomass measurements, respectively. The biomass estimation accuracy of maize could be improved by the fusion of LiDAR and hyperspectral data.



**Figure 7.** The scatterplot of field-observed biomass versus estimated biomass using the fusion of hyperspectral and LiDAR data.

## 5. Discussion

Many previous studies have successfully estimated crop biomass using airborne hyperspectral data [1,17]. However, our study found that VIs derived from CASI data had limited capability in estimating the biomass of maize. Although narrowband VIs were more sensitive to vegetation biomass than broadband Vis, according to the previous study [49], they still saturated over medium to high biomass area. Thus, VIs derived from hyperspectral data could not provide accurate biomass estimates because most of the biomass values are high in this experiment site. Additionally, results indicated that the combined use of VIs did not improve the predictability considerably due to the high correlation between VIs themselves.

The accuracy of LiDAR-derived DTM is a key factor for accurately calculating LiDAR-derived metrics and, thereafter, greatly affected the biomass estimation of maize. Findings of previous studies indicated that the accuracy of LiDAR-derived DTM is largely determined by: (1) vegetation canopy structure characteristics [33]; (2) terrain, mainly including slope and terrain irregularities; and (3) sensor characteristics, such as laser point density. Thus, the accuracy of LiDAR-derived DTM generally showed significant difference in different environments [31,37]. Our findings revealed that the accuracy of LiDAR-derived DTM is high in our study area. The main reason is that the terrain in this experimental site is relatively flat. Although there may be no ground points over extremely dense areas. The LiDAR-derived DTM can still be accurately obtained by the interpolation method.

Several LiDAR-derived metrics have been employed in crop biomass estimation in this analysis. Results indicated that  $H_{CV}$  showed highest correlation with the biomass of maize among all metrics.  $H_{CV}$  represents the relative dispersion of the vertical laser point height distribution, which has been demonstrated to be a good predictor of LAI in the study of Nie et al., (2016). Simultaneously, the biomass of maize is greatly dependent on the LAI according to Equation (1). Thus, the biomass of maize can be accurately estimated by  $H_{CV}$ . Although the study of Li et al., found a strong relationship between the canopy cover CC with the biomass of maize [6], low correlation was observed between CC and the biomass of maize in this study. This is maybe due to the two following reasons. On one hand, a thick canopy structure greatly reduces the probability of laser penetration to the ground. On the other hand, LiDAR cannot discriminate between the first and subsequent returns and only a single return is recorded for short vegetation. In this case, there may not be enough ground returns in some plots with extremely high canopy cover, and then the CC cannot be accurately calculated. Thus, the accuracy of biomass estimates is low using CC. Additionally, results also indicated that the combination of  $H_{mean}$  and  $H_{CV}$  can provide more accurate biomass estimates of maize. This may be owing to the fact that biomass was mainly determined by LAI and the height of maize. The mean height  $H_{mean}$  showed a significant relationship with the height of maize, and  $H_{CV}$  is closely related to the LAI of maize. Thus, the combined use of  $H_{mean}$  and  $H_{CV}$  could improve the biomass estimation accuracy.

Many previous studies indicated that the fusion of hyperspectral and LiDAR data can improve the estimation accuracy of forest biomass compared with either data alone [25,29,34,36]. However, we did not find literature regarding sensor data fusion for crop biomass estimation. Results in this study indicated that the fusion of LiDAR and hyperspectral data yielded the highest biomass estimation accuracy of maize using PLS regression. This may be due to the fact that LiDAR data can provide abundant information about vegetation canopy structure, while hyperspectral data can provide complementary spectral information regarding the biochemical and biophysical composition of vegetation canopy. However, the data fusion approach only reported a slight improvement for the uncertainty of biomass compared with the LiDAR-derived metrics alone, with RMSE values decreasing from 374.655 g/m<sup>2</sup> to 321.092 g/m<sup>2</sup>. Similar results were also reported in previous studies, which indicated the fusion of hyperspectral data could not significantly improve the estimation accuracy of forest biomass [25,49]. The main reason is that hyperspectral data could not provide accurate biomass estimates due to the saturation problem and LiDAR-derived metrics generally showed a strong relationship with field-observed biomass, thus, fused hyperspectral and LiDAR data only made a small improvement in biomass estimation accuracy.

To summarize, the biomass estimation accuracy of maize is low using hyperspectral data because of the problem of saturation over medium to high biomass areas. In contrast, LiDAR data can overcome the saturation problem and obtained more accurate biomass estimates of maize. Thus, LiDAR is considered as a more effective technology for biomass estimation. The fusion of LiDAR and hyperspectral data could obtain a higher biomass estimation accuracy of maize compared with LiDAR data alone. Our study made a new attempt to estimate crop biomass using fused LiDAR and hyperspectral data by PLS regression, which will provide valuable guidance for both data fusion and crop biomass estimation. However, this study also has several limitations. Field-observed biomass was indirectly calculated using the allometric equation. More accurate measurement of field-observed

biomass should be carried out in the future. Additionally, future works should also focus on the data fusion method for better combining of hyperspectral data with LiDAR data.

## 6. Conclusions

Our study aimed to explore the potential of hyperspectral and LiDAR data in estimating the biomass of maize. Each metric from either VIs or LiDAR-derived metrics was used to estimate the biomass of maize. The estimation models were also developed using PLS regression of biomass against VIs (only), LiDAR-derived metrics (only), and all metrics, respectively. The results of this study revealed the following conclusions. Moderate relationships existed between the biomass of maize and VIs. In contrast, LiDAR-derived metrics showed strong correlation with field-observed biomass of maize. In addition, the combined use of LiDAR-derived metrics resulted in a better biomass estimates ( $R^2 = 0.835$ ,  $RMSE = 374.655 \text{ g/m}^2$ ,  $RMSE_{CV} = 393.573 \text{ g/m}^2$ ), and the fusion of the hyperspectral and LiDAR data also led to a great improvement in biomass estimation accuracy of maize ( $R^2 = 0.883$ ,  $RMSE = 321.092 \text{ g/m}^2$ ,  $RMSE_{CV} = 337.653 \text{ g/m}^2$ ).

**Acknowledgments:** This work was supported by the National Natural Science Foundation of China under Grant 41271428 and Grant 41371350.

**Author Contributions:** C.W. and S.N. designed the experiments, analyzed the data and wrote the paper; X.X. and X.S. revised the paper; S.L. collected the field data.

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

## References

1. Gitelson, A.; Merzlyak, M.N. Spectral reflectance changes associated with autumn senescence of *Aesculus hippocastanum* L. and *Acer platanoides* L. Leaves. Spectral features and relation to chlorophyll estimation. *J. Plant Physiol.* **1994**, *143*, 286–292. [[CrossRef](#)]
2. Liu, J.; Pattey, E.; Miller, J.R.; McNairn, H.; Smith, A.; Hu, B. Estimating crop stresses, aboveground dry biomass and yield of corn using multi-temporal optical data combined with a radiation use efficiency model. *Remote Sens. Environ.* **2010**, *114*, 1167–1177. [[CrossRef](#)]
3. Chen, Q. Lidar remote sensing of vegetation biomass. *Remote Sens. Nat. Resour.* **2013**, *399*, 399–420.
4. Pereira Coltri, P.; Zullo, J.; Ribeiro do Valle Goncalves, R.; Romani, L.A.S.; Pinto, H.S. Coffee crop's biomass and carbon stock estimation with usage of high resolution satellites images. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2013**, *6*, 1786–1795. [[CrossRef](#)]
5. Eitel, J.U.H.; Magney, T.S.; Vierling, L.A.; Brown, T.T.; Huggins, D.R. Lidar based biomass and crop nitrogen estimates for rapid, non-destructive assessment of wheat nitrogen status. *Field Crop. Res.* **2014**, *159*, 21–32. [[CrossRef](#)]
6. Li, W.; Niu, Z.; Huang, N.; Wang, C.; Gao, S.; Wu, C. Airborne lidar technique for estimating biomass components of maize: A case study in Zhangye City, Northwest China. *Ecol Indic.* **2015**, *57*, 486–496. [[CrossRef](#)]
7. Haboudane, D.; Miller, J.R.; Pattey, E.; Zarco-Tejada, P.J.; Strachan, I.B. Hyperspectral vegetation indices and novel algorithms for predicting green LAI of crop canopies: Modeling and validation in the context of precision agriculture. *Remote Sens. Environ.* **2004**, *90*, 337–352. [[CrossRef](#)]
8. Tilly, N.; Hoffmeister, D.; Cao, Q.; Huang, S.; Lenz-Wiedemann, V.; Miao, Y.; Bareth, G. Multitemporal crop surface models: Accurate plant height measurement and biomass estimation with terrestrial laser scanning in paddy rice. *J. Appl. Remote Sens.* **2014**, *8*. [[CrossRef](#)]
9. Becker-Reshef, I.; Vermote, E.; Lindeman, M.; Justice, C. A generalized regression-based model for forecasting winter wheat yields in Kansas and Ukraine using MODIS data. *Remote Sens. Environ.* **2010**, *114*, 1312–1323. [[CrossRef](#)]
10. Gao, S.; Niu, Z.; Huang, N.; Hou, X. Estimating the leaf area index, height and biomass of maize using HJ-1 and RADARSAT-2. *Int. J. Appl. Earth Obs. Geoinf.* **2013**, *24*, 1–8. [[CrossRef](#)]
11. Jacquemoud, S.; Bacour, C.; Poilve, H.; Frangi, J.P. Comparison of four radiative transfer models to simulate plant canopies reflectance: Direct and inverse mode. *Remote Sens. Environ.* **2000**, *74*, 471–481. [[CrossRef](#)]

12. Tang, S.; Chen, J.; Zhu, Q.; Li, X.; Chen, M.; Sun, R.; Zhou, Y.; Deng, F.; Xie, D. LAI inversion algorithm based on directional reflectance kernels. *J. Environ. Manag.* **2007**, *85*, 638–648. [[CrossRef](#)] [[PubMed](#)]
13. Boegh, E.; Soegaard, H.; Broge, N.; Hasager, C.; Jensen, N.; Schelde, K.; Thomsen, A. Airborne multispectral data for quantifying leaf area index, nitrogen concentration, and photosynthetic efficiency in agriculture. *Remote Sens. Environ.* **2002**, *81*, 179–193. [[CrossRef](#)]
14. Broge, N.H.; Mortensen, J.V. Deriving green crop area index and canopy chlorophyll density of winter wheat from spectral reflectance data. *Remote Sens. Environ.* **2002**, *81*, 45–57. [[CrossRef](#)]
15. Wu, C.; Niu, Z.; Tang, Q.; Huang, W. Estimating chlorophyll content from hyperspectral vegetation indices: Modeling and validation. *Agric. For. Meteorol.* **2008**, *148*, 1230–1241. [[CrossRef](#)]
16. Nguy-Robertson, A.L.; Peng, Y.; Gitelson, A.A.; Arkebauer, T.J.; Pimstein, A.; Herrmann, I.; Karnieli, A.; Rundquist, D.C.; Bonfil, D.J. Estimating green LAI in four crops: Potential of determining optimal spectral bands for a universal algorithm. *Agric. For. Meteorol.* **2014**, *192–193*, 140–148. [[CrossRef](#)]
17. Xie, Q.; Huang, W.; Liang, D.; Chen, P.; Wu, C.; Yang, G.; Zhang, J.; Huang, L.; Zhang, D. Leaf area index estimation using vegetation indices derived from airborne hyperspectral images in winter wheat. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2014**, *7*, 3586–3594. [[CrossRef](#)]
18. Gong, P.; Pu, R.; Biging, G.S.; Larrieu, M.R. Estimation of forest leaf area index using vegetation indices derived from hyperion hyperspectral data. *IEEE Trans. Geosci. Remote Sens.* **2003**, *41*, 1355–1362. [[CrossRef](#)]
19. Koetz, B.; Sun, G.; Morsdorf, F.; Ranson, K.J.; Kneubuehler, M.; Itten, K.; Allgoewer, B. Fusion of imaging spectrometer and lidar data over combined radiative transfer models for forest canopy characterization. *Remote Sens. Environ.* **2007**, *106*, 449–459. [[CrossRef](#)]
20. Gilmore, M.S.; Wilson, E.H.; Barrett, N.; Civco, D.L.; Prisloe, S.; Hurd, J.D.; Chadwick, C. Integrating multi-temporal spectral and structural information to map wetland vegetation in a lower connecticut river tidal marsh. *Remote Sens. Environ.* **2008**, *112*, 4048–4060. [[CrossRef](#)]
21. Delegido, J.; Alonso, L.; González, G.; Moreno, J. Estimating chlorophyll content of crops from hyperspectral data using a normalized area over reflectance curve (NAOC). *Int. J. Appl. Earth Obs. Geoinf.* **2010**, *12*, 165–174. [[CrossRef](#)]
22. Hladik, C.; Schalles, J.; Alber, M. Salt marsh elevation and habitat mapping using hyperspectral and LiDAR data. *Remote Sens. Environ.* **2013**, *139*, 318–330. [[CrossRef](#)]
23. Zhao, K.; Valle, D.; Popescu, S.; Zhang, X.; Mallick, B. Hyperspectral remote sensing of plant biochemistry using bayesian model averaging with variable and band selection. *Remote Sens. Environ.* **2013**, *132*, 102–119. [[CrossRef](#)]
24. Chen, J.M.; Cihlar, J. Retrieving leaf area index of boreal conifer forests using Landsat TM images. *Remote Sens. Environ.* **1996**, *55*, 153–162. [[CrossRef](#)]
25. Vaglio Laurin, G.; Chen, Q.; Lindsell, J.A.; Coomes, D.A.; Frate, F.D.; Guerriero, L.; Pirotti, F.; Valentini, R. Above ground biomass estimation in an African tropical forest with LiDAR and hyperspectral data. *ISPRS J. Photogramm.* **2014**, *89*, 49–58. [[CrossRef](#)]
26. Cui, Y.; Zhao, K.; Fan, W.; Xu, X. Retrieving crop fractional cover and LAI based on airborne LiDAR data. *J. Remote Sens.* **2011**, *15*, 1276–1288.
27. Hofle, B. Radiometric correction of terrestrial LiDAR point cloud data for individual maize plant detection. *IEEE Geosci. Remote Sens. Lett.* **2014**, *11*, 94–98. [[CrossRef](#)]
28. Zhao, K.; Popescu, S.; Nelson, R. LiDAR remote sensing of forest biomass: A scale-invariant estimation approach using airborne lasers. *Remote Sens. Environ.* **2009**, *113*, 182–196. [[CrossRef](#)]
29. Clark, M.L.; Roberts, D.A.; Ewel, J.J.; Clark, D.B. Estimation of tropical rain forest aboveground biomass with small-footprint LiDAR and hyperspectral sensors. *Remote Sens. Environ.* **2011**, *115*, 2931–2942. [[CrossRef](#)]
30. Swatantran, A.; Dubayah, R.; Roberts, D.; Hofton, M.; Blair, J.B. Mapping biomass and stress in the Sierra Nevada using LiDAR and hyperspectral data fusion. *Remote Sens. Environ.* **2011**, *115*, 2917–2930. [[CrossRef](#)]
31. Luo, S.; Wang, C.; Pan, F.; Xi, X.; Li, G.; Nie, S.; Xia, S. Estimation of wetland vegetation height and leaf area index using airborne laser scanning data. *Ecol. Indic.* **2015**, *48*, 550–559. [[CrossRef](#)]
32. Nie, S.; Wang, C.; Dong, P.; Xi, X. Estimating leaf area index of maize using airborne full-waveform lidar data. *Remote Sens. Lett.* **2016**, *7*, 111–120. [[CrossRef](#)]
33. Luo, S.; Wang, C.; Xi, X.; Pan, F. Estimating FPAR of maize canopy using airborne discrete-return lidar data. *Opt. Express* **2014**, *22*, 5106–5117. [[CrossRef](#)] [[PubMed](#)]

34. Asner, G.P.; Knapp, D.E.; Kennedy-Bowdoin, T.; Jones, M.O.; Martin, R.E.; Boardman, J.; Field, C.B. Carnegie airborne observatory: In-flight fusion of hyperspectral imaging and waveform light detection and ranging for three-dimensional studies of ecosystems. *J. Appl. Remote Sens.* **2007**, *1*. [[CrossRef](#)]
35. Anderson, J.; Plourde, L.; Martin, M.; Braswell, B.; Smith, M.; Dubayah, R.; Hofton, M.; Blair, J. Integrating waveform LiDAR with hyperspectral imagery for inventory of a northern temperate forest. *Remote Sens. Environ.* **2008**, *112*, 1856–1870. [[CrossRef](#)]
36. Latifi, H.; Fassnacht, F.; Koch, B. Forest structure modeling with combined airborne hyperspectral and LiDAR data. *Remote Sens. Environ.* **2012**, *121*, 10–25. [[CrossRef](#)]
37. Kulawardhana, R.W.; Popescu, S.C.; Feagin, R.A. Fusion of LiDAR and multispectral data to quantify salt marsh carbon stocks. *Remote Sens. Environ.* **2014**, *154*, 345–357. [[CrossRef](#)]
38. Lin, X.; Zhang, J. Segmentation-based filtering of airborne LiDAR point clouds by progressive densification of terrain segments. *Remote Sens.* **2014**, *6*, 1294–1326. [[CrossRef](#)]
39. Axelsson, P. DEM generation from laser scanner data using adaptive tin models. *Int. Arch. Photogramm. Remote Sens.* **2000**, *33*, 111–118.
40. Li, X.; Cheng, G.; Liu, S.; Xiao, Q.; Ma, M.; Jin, R.; Che, T.; Liu, Q.; Wang, W.; Qi, Y. Heihe watershed allied telemetry experimental research (HiWATER): Scientific objectives and experimental design. *Bull. Am. Meteorol. Soc.* **2013**, *94*, 1145–1160. [[CrossRef](#)]
41. Matthew, M.W.; Adler-Golden, S.M.; Berk, A.; Felde, G.; Anderson, G.P.; Gorodetzky, D.; Paswaters, S.; Shippert, M. Atmospheric correction of spectral imagery: Evaluation of the flaash algorithm with AVIRIS data. In Proceedings of the IEEE Applied Imagery Pattern Recognition Workshop, Washington, DC, USA, 16–18 October 2002; pp. 157–163.
42. Berk, A.; Bernstein, L.; Anderson, G.; Acharya, P.; Robertson, D.; Chetwynd, J.; Adler-Golden, S. Modtran cloud and multiple scattering upgrades with application to AVIRIS. *Remote Sens. Environ.* **1998**, *65*, 367–375. [[CrossRef](#)]
43. Thomas, V.; Finch, D.A.; McCaughey, J.H.; Noland, T.; Rich, L.; Treitz, P. Spatial modelling of the fraction of photosynthetically active radiation absorbed by a boreal mixedwood forest using a LiDAR-hyperspectral approach. *Agric. For. Meteorol.* **2006**, *140*, 287–307. [[CrossRef](#)]
44. Jordan, C.F. Derivation of leaf-area index from quality of light on the forest floor. *Ecology* **1969**, *50*, 663–666. [[CrossRef](#)]
45. Rouse, J., Jr. *Monitoring the Vernal Advancement and Retrogradation (Green Wave Effect) of Natural Vegetation*; NASA Technical Reports Server: Hampton, VA, USA, 1974.
46. Qi, J.; Chehbouni, A.; Huete, A.; Kerr, Y.; Sorooshian, S. A modified soil adjusted vegetation index. *Remote Sens. Environ.* **1994**, *48*, 119–126. [[CrossRef](#)]
47. Gitelson, A.A.; Viña, A.; Arkebauer, T.J.; Rundquist, D.C.; Keydan, G.; Leavitt, B. Remote estimation of leaf area index and green leaf biomass in maize canopies. *Geophys. Res. Lett.* **2003**, *30*. [[CrossRef](#)]
48. Jin, X.; Yang, G.; Xu, X.; Yang, H.; Feng, H.; Li, Z.; Shen, J.; Lan, Y.; Zhao, C. Combined multi-temporal optical and radar parameters for estimating LAI and biomass in winter wheat using HJ and RADARSAR-2 data. *Remote Sens.* **2015**, *7*, 13251–13272. [[CrossRef](#)]
49. Luo, S.; Wang, C.; Xi, X.; Pan, F.; Peng, D.; Zou, J.; Nie, S.; Qin, H. Fusion of airborne LiDAR data and hyperspectral imagery for aboveground and belowground forest biomass estimation. *Ecol. Indic.* **2017**, *73*, 378–387. [[CrossRef](#)]
50. Brovelli, M.A.; Crespi, M.; Fratarcangeli, F.; Giannone, F.; Realini, E. Accuracy assessment of high resolution satellite imagery orientation by leave-one-out method. *ISPRS J. Photogramm. Remote Sens.* **2008**, *63*, 427–440. [[CrossRef](#)]
51. Peduzzi, A.; Wynne, R.H.; Thomas, V.A.; Nelson, R.F.; Reis, J.J.; Sanford, M. Combined use of airborne LiDAR and DBinSAR data to estimate LAI in temperate mixed forests. *Remote Sens.* **2012**, *4*, 1758–1780. [[CrossRef](#)]

