energies

MDPI

Article

Estimating Field-Level Perennial Bioenergy Grass Biomass Yields Using the Normalized Difference Red-Edge Index and Linear Regression Analysis for Central Virginia, USA

Yuki Hamada *, Colleen R. Zumpf, John J. Quinn and Maria Cristina Negri

Environmental Science Division, Argonne National Laboratory, 9700 S. Cass Ave., Lemont, IL 60439, USA
* Correspondence: yhamada@anl.gov; Tel.: +1-630-252-0087

Abstract: We investigated the indicative power of the normalized difference red-edge index (NDRE) for estimating field-level perennial bioenergy grass biomass yields utilizing Sentinel-2 imagery and a linear regression model as a rapid, cost-effective method for biomass yield estimations for bioenergy. We used 2019 data from three study sites containing mature perennial bioenergy grass stands in central Virginia, USA. Of the simulated daily NDRE values based on the temporally weighted averaging of two temporal neighbors, we found the strongest index–yield correlation on 11 August (R = 0.85). We estimated the perennial bioenergy grass biomass yields for (1) all sites using the data pooled from the three sites (all-site estimation) and (2) each site using the data pooled from the other two sites (cross-site estimation). The estimated field-level perennial bioenergy grass biomass yields strongly correlated with the recorded yields (average R^2 = 0.76), with a root mean square error (RMSE) of 1.5 Mg/ha and a mean absolute error (MAE) of 1.2 Mg/ha for the all-site estimation. For the cross-site estimation, the site with diverse perennial grass types had the weakest correlation (R^2 = 0.44) of the sites, indicating a difficulty in accounting for heterogeneous index–yield relationships in a single model. In addition to identifying a strong indicative power of the NDRE for estimating the overall perennial bioenergy grass biomass yields at a field level, the findings from this study call for an analysis across multiple perennial grasses and a comparison using multiple sites to understand (1) if the indicative power of the index shifts from the biomass of the specific perennial bioenergy grass type to the overall biomass during the growing season and (2) the level of perennial bioenergy grass heterogeneity that may hinder the remotely sensed biomass yield estimation using a single model.

Keywords: remote sensing; Sentinel-2; vegetation index; switchgrass; spectral response; renewable resource; clean energy

check for updates

Citation: Hamada, Y.; Zumpf, C.R.; Quinn, J.J.; Negri, M.C. Estimating Field-Level Perennial Bioenergy Grass Biomass Yields Using the Normalized Difference Red-Edge Index and Linear Regression Analysis for Central Virginia, USA. *Energies* **2023**, *16*, 7397. https://doi.org/10.3390/en16217397

Academic Editors: João Fernando Pereira Gomes and Toufik Boushaki

Received: 1 August 2023
Revised: 27 October 2023
Accepted: 29 October 2023
Published: 2 November 2023

Copyright: © 2023 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 (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Perennial grasses offer many ecosystem services, such as supporting wildlife habitats [1,2], moderating soil erosion [3–5], conserving water [6], and improving the soil and water quality [7,8]. Their abilities to produce a high biomass yield with a low fertilizer input and grow across many regions make perennial grasses an ideal bioenergy feedstock [9], playing a key role in the clean energy transition. Strategically incorporating these perennial bioenergy grasses within current row crop production systems could optimize the environmental benefits while improving the livelihoods of farmers [10,11] and mitigating conflict with food production [12,13]. One strategy is to utilize areas with historically low yields, nutrient leaching, and/or soil erosion, defined as marginal lands, which are often scattered across the agricultural landscape [12,13]. Growing high-yielding perennial bioenergy grass with low inputs on marginal lands across multiple regions can be an ideal strategy for replacing traditional fuels. Timely estimation of biomass supplies for bioenergy could accelerate this strategy by providing valuable information to refineries to prepare for biomass

processing [14,15], using agencies to track and evaluate the success of such a strategy, and using researchers to identify more advantageous grass varieties or cultivars [14,16].

1.1. Use of Remote Sensing for Estimating Perennial Bioenergy Grass Biomass Yields

Remote sensing is widely used for estimating the vegetation abundance, and it is particularly useful for rapid biomass estimations across large and/or remote areas [17,18]. Spectral vegetation indices have been extensively utilized for estimating aboveground biomass for a range of plant types using both proximal and remote sensing techniques [19–21]. A spectral vegetation index is a numerical value calculated using two or more spectral reflectance values; it correlates with one or more plant properties, such as the leaf nitrogen content, canopy moisture content, and biomass volume in the field [22]. Numerous spectral vegetation indices are widely used for plant biomass estimation; however, few studies have investigated their utility for lignocellulosic perennial bioenergy grasses.

Four studies pertaining to the estimation of perennial bioenergy grass biomass yields [23–26] suggested the potential of three vegetation indices: the normalized difference vegetation index (NDVI [27]), the green NDVI (GNDVI [28]), and the normalized difference red-edge index (NDRE [29]). Kanemasu et al. [23] visually presented a strong agreement between the estimated and measured dry biomasses, indicating the potential of using the NDVI for estimating the perennial grass dry biomass of big bluestem (*Andropogon gerardii*) and Indiangrass (*Sorghastrum nutans*) in Kansas. Hamada et al. [24] estimated the plot-level perennial bioenergy grass biomass yields of multiple switchgrass (*Panicum virgatum)* cultivars across five fields in the US Midwest with a coefficient of determination (R^2) as high as 0.88 using the midsummer GNDVI with a linear regression model. Using optical images collected using an unmanned aerial platform, Namoi et al. [25] estimated the plot-level *Miscanthus* biomass yields using the NDRE from midsummer with a stronger correlation (R^2 = 0.82–0.97) than the NDVI and GNDVI. Foster et al. [26] found that the red-edge spectral range was the key determinant for estimating the biomass yields of switchgrass and mixed grasses, which supports the effectiveness of vegetation indices using the red-edge band, such as the NDRE.

1.2. Relevance of the NDRE for Estimating Perennial Bioenergy Grass Biomass Yields

Existing studies [25,26,30,31] indicate the great relevance of vegetation indices utilizing the red-edge spectral bands, such as the NDRE, for estimating the vegetation biomass. Naguib and Daliman [30] described that while the NDVI highly correlates with vegetation cover, the NDRE corresponds to the plant chlorophyll and nitrogen contents. Because the chlorophyll and nitrogen contents indicate the accumulation of biomass more strongly than the vegetation cover, the NDRE is a suitable choice for biomass yield estimation. Chang and Shoshany [31] found that indices using the red-edge band could improve the sensitivity to the biomass quantity for moderate and dense vegetation growth without reaching the maximum index value, also known as saturation. All of these properties are useful for estimating the perennial bioenergy grass biomass.

With the scaling of remotely sensed estimation for the perennial bioenergy grass biomass yield in mind, we previously examined the indicative power of 30 spectral vegetation indices derived from 20 m resolution Sentinel-2 satellite imagery [32]. The study [32] aimed to estimate the field-level perennial bioenergy grass biomass yields for four study sites in central Virginia, USA. We modeled the perennial bioenergy grass biomass yield for each site independently using a linear regression model. We found that the NDRE showed the strongest and most consistent correlation with the recorded biomass yields across the sites with an average R^2 of 0.75, root mean square error (RMSE) of 1.3 Mg/ha, and mean absolute error (MAE) of 1.0 Mg/ha. This suggested the potential utility of the index for rapid, reliable estimations of perennial bioenergy grass biomass yields. However, varying dates for image availability across the study sites prevented us from modeling the perennial bioenergy grass biomass yields in an integrated manner using multiple sites.

Thus, we could not gain insights regarding the applicability of the method for estimating perennial bioenergy grass biomass yields beyond a single site.

Based on the consideration of the previous studies and the benefit of NDRE outlined earlier in this introduction, the goal of the current study was to advance our previous understanding of the indicative power of the NDRE for estimating field-level perennial bioenergy grass biomass yields using a common linear regression model that integrates datasets from multiple study sites. To accomplish this goal, our objectives were to enable the estimation of perennial bioenergy grass biomass yields by (1) integrating data from all study sites by simulating daily NDRE values, (2) estimating the perennial bioenergy grass biomass yields using two distinct approaches, and (3) evaluating the efficacy of the promising approach.

2. Materials and Methods

2.1. Study Area

The analysis initially considered four study sites of mature perennial bioenergy grasses located in central Virginia (Figure 1). Each site consisted of multiple fields of perennial grasses, ranging from 13 to 24 fields (Table 1). However, some of the fields at Sites B, C, and D were too small, limiting remotely sensed imagery sampling, and therefore, we excluded those fields from the modeling analysis (Table 1).

Table 1. Summary of the study site characteristics for the 2019 season and Sentinel-2 satellite image availability.

	Site A	**Site B**	**Site C**	**Site D**
Elevation (m)	305	116	168	244
Total field count	13	18	15	24
Field count available for modeling	13	16	14	20
Total area [1] (ha)	243.4	134.3	63.8	100.2
Average field size (ha) [standard deviation]	18.7 [20.6]	8.4 [7.9]	4.3 [3.4]	5.0 [2.6]
Average field-level perennial bioenergy grass yield [2] (Mg/ha) [standard deviation]	7.9 [2.4]	5.7 [1.8]	2.7 [0.6]	7.2 [2.0]
Perennial bioenergy grass variety	Switchgrass 'Alamo'	Switchgrass 'Kanlow', *Miscanthus*, mixed grass (big bluestem, Indiangrass, switchgrass)	Mixed grass (big bluestem, Indiangrass, switchgrass)	Switchgrass 'Blackwell'
Harvest activities for the 2019 season				
Mowing period	9/18–10/6	8/13–8/29	9/3–9/5	7/30–8/4
Raking period	9/19–10/7	8/25–8/31	9/4–9/6	7/30–8/12
Baling begins	9/19–10/7	8/27–8/31	9/4–9/6	8/4–8/21
Dates of Sentinel-2 image availability for the 2019 season				
April	—	11, 16	4	4
May	24	19, 21, 24	14, 19, 24	4, 19, 24
June	3, 23, 28	3, 15	3, 23	3, 13, 28
July	3, 13, 28	10, 13, 20, 30	3, 8, 18, 28	3, 13, 28
August	12	9, 19, 29	22	12, 22
September	11, 16, 21	8, 21	1, 6, 11, 16, 21	1, 11, 16, 21, 26
October	1, 11, 21	1, 18, 28	1, 11	1, 11
Total images considered	14	19	18	19

[1] Areas only account for the available fields utilized for modeling. [2] The recorded field-level perennial bioenergy grass biomass yields were obtained from baling.

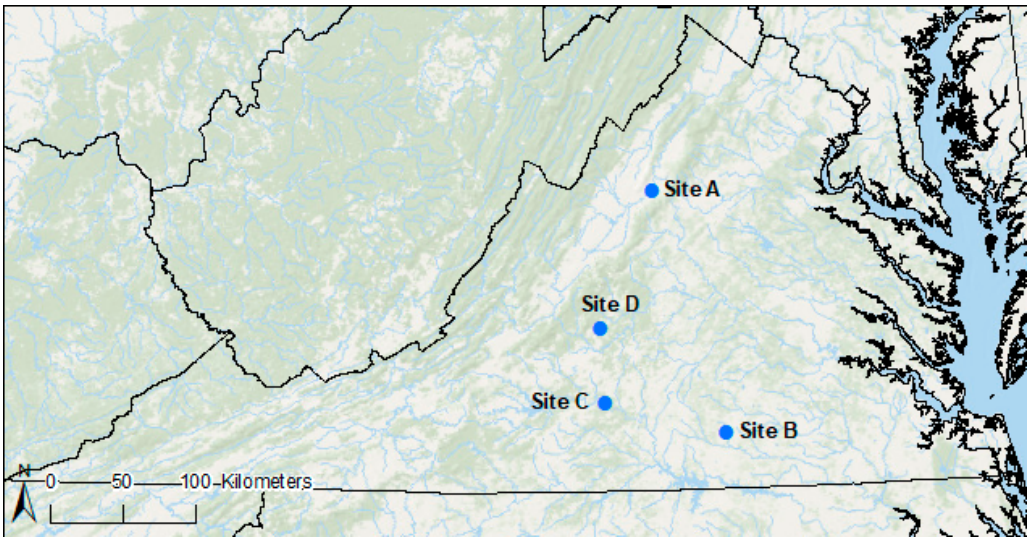


Figure 1. Study site locations in the state of Virginia, USA.

2.2. Data

The primary data used for modeling perennial bioenergy grass biomass yields were 20 m resolution Sentinel-2 satellite multispectral imagery available via the European Space Agency (ESA; https://sentinel.esa.int/web/sentinel/missions/sentinel-2; accessed on 15 January 2023) and perennial bioenergy grass biomass yields recorded at the time of harvest in 2019 (Table 1). Details of the harvesting activities, such as mowing, raking, and baling, were used to describe patterns in perennial bioenergy grass biomass yield estimations. The 20 m resolution Sentinel-2 satellite imagery consists of visible spectral bands (blue, green, and red), red-edge bands, near-infrared bands (NIR), and shortwave infrared (SWIR). The red-edge bands located between the visible and NIR spectral regions are sensitive to the plant chlorophyll and/or nitrogen contents [22,33,34]. The NIR bands located approximately in the 850–1100 nm spectral region are indicative of the leaf and canopy structure [22]. The SWIR bands located approximately in the 1100–2500 nm range are known to correlate with the plant and soil moisture contents [22,35–37] and lignocellulose in plants. In our previous study [32], field-level NDRE values were calculated using cloud-free, shadow-free Sentinel-2 imagery, along with another 29 spectral vegetation indices using 70 available images. Our previous study [32] provides details of the image processing procedures. Several features of Sentinel-2 imagery, including the image accessibility, cost-effectiveness, useful spectral bands for vegetation studies, and standardized image collection and processing methods, make the dataset highly suitable for scaling the plant biomass estimation across regions.

A total of 63 fields (or data points) were available for analysis, consisting of 13, 16, 14, and 20 fields from Sites A, B, C, and D, respectively (Table 1). The average perennial bioenergy grass biomass yields (ranging from 2.7 Mg/ha [Site C] to 7.9 Mg/ha [Site A]), and standard deviations (ranging from 0.6 Mg/ha [Site C] to 2.4 Mg/ha [Site A]) varied across the sites. Of all sites, Site B was the most heterogeneous of the perennial grass types grown across the fields (*Miscanthus*, switchgrass, and a grass mixture) (Table 1). Sites A and D grew a single switchgrass cultivar across all fields and yielded higher biomasses per unit area than Sites B and C. Site C grew a grass mixture across all fields and had the lowest biomass yield of all sites. The timing of the harvesting activities varied greatly across the sites. For example, the first day of mowing ranged from 30 July for Site D to 18 September for Site A (51 days apart).

2.3. Analysis

Using the NDRE layers generated from the available Sentinel-2 imagery in the previous study [32], we simulated the daily NDRE values for each of the study sites based on the temporally weighted averaging of two temporal neighbors. For each data gap, we calculated an average daily change in the NDRE value based on the difference between two temporal neighbors divided by the number of days in the gap. The average daily change was positive when the NDRE showed an increasing trend (e.g., early growing season); the change was negative when the index showed a decreasing trend (e.g., senescence). We added the value of the average daily change to the value from the previous day across the gap to fill each data gap. We pooled these simulated daily NDRE values from the four study sites and the recorded yields and calculated the correlation coefficients between the recorded yields and the NDRE values across the season. The date with the strongest correlation was selected as the optimal date for NDRE for estimating the perennial bioenergy grass biomass yields for subsequent analysis.

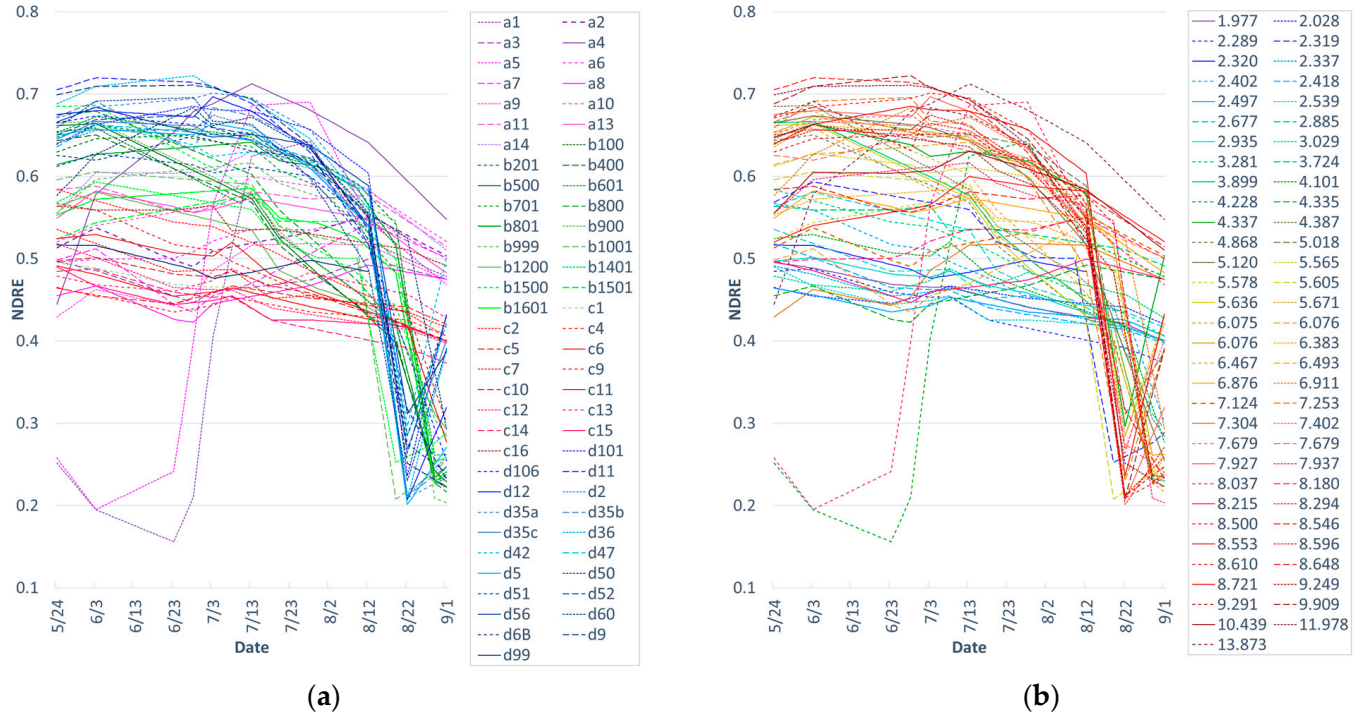
Using the simulated NDRE values on the optimal date, we estimated the field-level biomass yields for perennial bioenergy grasses at the study sites using two approaches. First, we estimated the field-level biomass yields across the sites based on a linear regression model using the pooled dataset from all sites (all-site estimation). For the all-site estimation, we employed five-fold cross-validation. We randomly divided the dataset into five subsets of almost equal size. We developed a model using four subsets and validated the model using the one remaining subset. We repeated the steps five times with different subset combinations. Second, we estimated the field-level biomass yields for each of the study sites based on a linear regression model developed using the data pooled from the other sites (cross-site estimation). For both methods, we evaluated the accuracy of the estimated perennial bioenergy grass biomass yields based on the R^2, MAE, and RMSE.

3. Results

3.1. Seasonal Trajectory of the NDRE

Figure 2 shows seasonal trajectories of the simulated field-level NDRE values across the perennial bioenergy grass fields for the 2019 season. The NDRE values generally reached their peak greenness in June and July (Figure 2). For Sites B and D, the NDRE values sharply decreased from mid-August, which was caused by senescence and/or biomass removal by harvesting. In each field, the NDRE trajectories were generally consistent across the fields with a few exceptions. While Sites B, C, and D showed very uniform trajectories across the fields, Fields 1 and 5 in Site A (Figure 2a, a1 and a5) showed late starts compared with the rest of the fields. The NDRE values in these two Site A fields almost caught up with those for the other fields in five to 10 days. Field 4 in Site A (Figure 2a, a4), which had the highest perennial bioenergy grass biomass yield (13.9 Mg/ha) of all fields, began with one of the lowest NDRE values, and the index values consistently increased at a higher rate than those in other fields until peaking around early to mid-July.

When comparing the NDRE seasonal trajectories across the yields (Figure 2b), the higher and lower index values generally corresponded to the lower and higher perennial bioenergy grass biomass yields, respectively. There were some inconsistencies, for example, some sites with high biomass yields, such as Fields 3 and 6 at Site A (Figure 2a, a3 and a6), which had yields of greater than 8 Mg/ha, had similar NDRE values to fields with yields of less than 3 Mg/ha (Figure 2b).



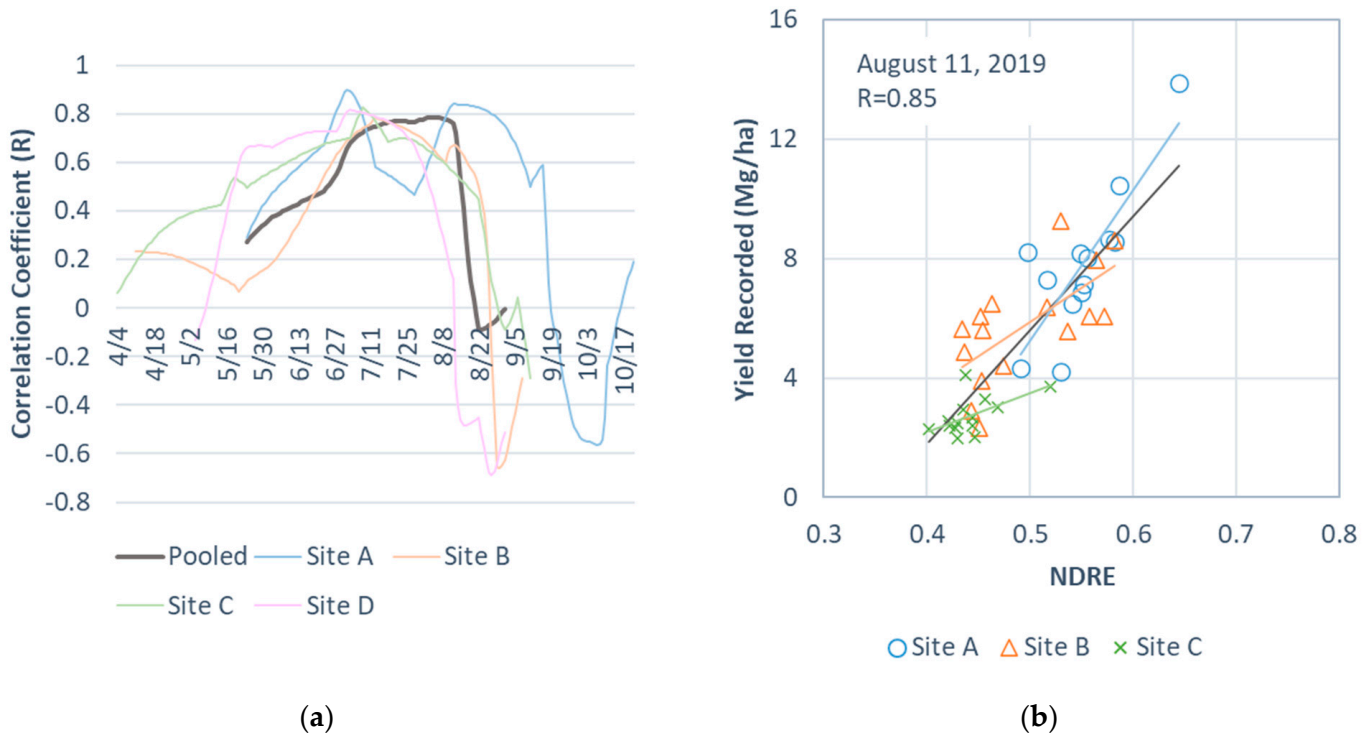
(**a**) (**b**)

Figure 2. Seasonal trajectories of the simulated field-level NDRE values from 24 May to 1 September 2019, across the fields (**a**) by site and (**b**) by yield in Mg/ha. (**a**) Site IDs and field IDs are combined in the legend (e.g., a1 = Site A Field 1; d60 = Site D Field 60). (**b**) Lower yields are shown in cooler colors (e.g., purples and blues), and higher yields are shown in warmer colors (e.g., oranges and reds).

3.2. Optimal Date of NDRE for Estimating Perennial Bioenergy Grass Biomass Yields

Figure 3 shows (a) the seasonal trajectories of the correlation coefficient (R) between the recorded perennial bioenergy grass biomass yields at the field level and the simulated NDRE values across the study season and (b) a scatter plot of the simulated NDRE values against the recorded yields on the selected date. The trajectories showed strong correlations between the NDRE and the recorded yields around mid-July for Sites B, C, and D, peaking on 8, 13, and 7 July, respectively (Figure 3a). Site A showed two strong NDRE-yield peaks, one on 2 July and the other on 12 August. Varying timings of the onsets of senescence and harvest were also apparent in the correlation trajectory (Figure 3a). Harvest at Site D began approximately 50 days earlier than at Site A. Compared to lowland varieties for other sites ('Kanlow' and 'Alamo'), the short growing season of the upland switchgrass variety ('Blackwell') grown on Site D likely determined the earlier harvest for the site.

When the data from all study sites were pooled, the NDRE–yield correlations exceeded 0.78 for 11 days, from 31 July to 10 August, with the strongest correlation occurring on 5 August (R = 0.79; Figure 3a). However, there was great uncertainty in the simulated NDRE values for Site D on 5 August because the biomass for the site was completely mowed by that date. Because of a lack of images for Site D during that period (Table 1), we could not manually exclude the mowed fields from the dataset. An alternative date for estimating the perennial bioenergy grass biomass yields for the four sites was 30 July (R = 0.77). The correlation indicated that the NDRE could explain approximately 59% of the yield variation. When excluding Site D, we found the strongest NDRE–yield correlation on 11 August (R = 0.85; Figure 3b), which was prior to mowing for the three remaining sites. Thus, we selected 11 August and opted to exclude Site D from the subsequent perennial bioenergy grass biomass yield estimation. This resulted in a total of 43 data points or fields for analysis, which allowed three subsets containing nine data points and two subsets containing eight data points for five-fold cross validation for the all-site estimation.



(**a**) (**b**)

Figure 3. Correlations between the recorded perennial bioenergy grass biomass yields and the NDRE for the study sites during the study season. (**a**) Seasonal trajectories of the correlation coefficient (R) between the recorded perennial bioenergy grass biomass yields at the field level and the simulated NDRE values across the study season and (**b**) a scatter plot of NDRE values on 11 August 2019 against the recorded yields.

3.3. Accuracy of Estimated Field-Level Perennial Bioenergy Grass Biomass Yields Using a Linear NDRE Regression Model

Figure 4 shows scatter plots of the recorded field-level perennial bioenergy grass biomass yields against the estimated yields using a linear NDRE regression model. For the all-site estimation, the estimated field-level perennial grass biomass yields showed a strong correlation with the recorded biomass yields (median R^2 = 0.75, Figure 4a). The estimated biomass yields were most consistent with the recorded yields of all biomass yield estimations performed in this study, which is indicated by a slope of 0.86 (close to 1) and an intercept of 0.26 (close to 0) (Figure 4a).

For the cross-site estimation, correlations between the estimated and recorded yields varied across the sites, ranging from R^2 = 0.33 to R^2 = 0.68 (Figure 4b,c). The highest biomass yield estimated for Site A was 9.7 Mg/ha, despite the site's highest recorded yield being 13.9 Mg/ha. The biomass yields for Site B had the most consistent relationship with the recorded yields (slope = 0.85, intercept = 0.31) of all three sites in the cross-site estimation, despite their moderate correlation (R^2 = 0.44; Figure 4c). The estimated yield for Site C was minimally correlated with the recorded yields at R^2 = 0.33 (Figure 4d).

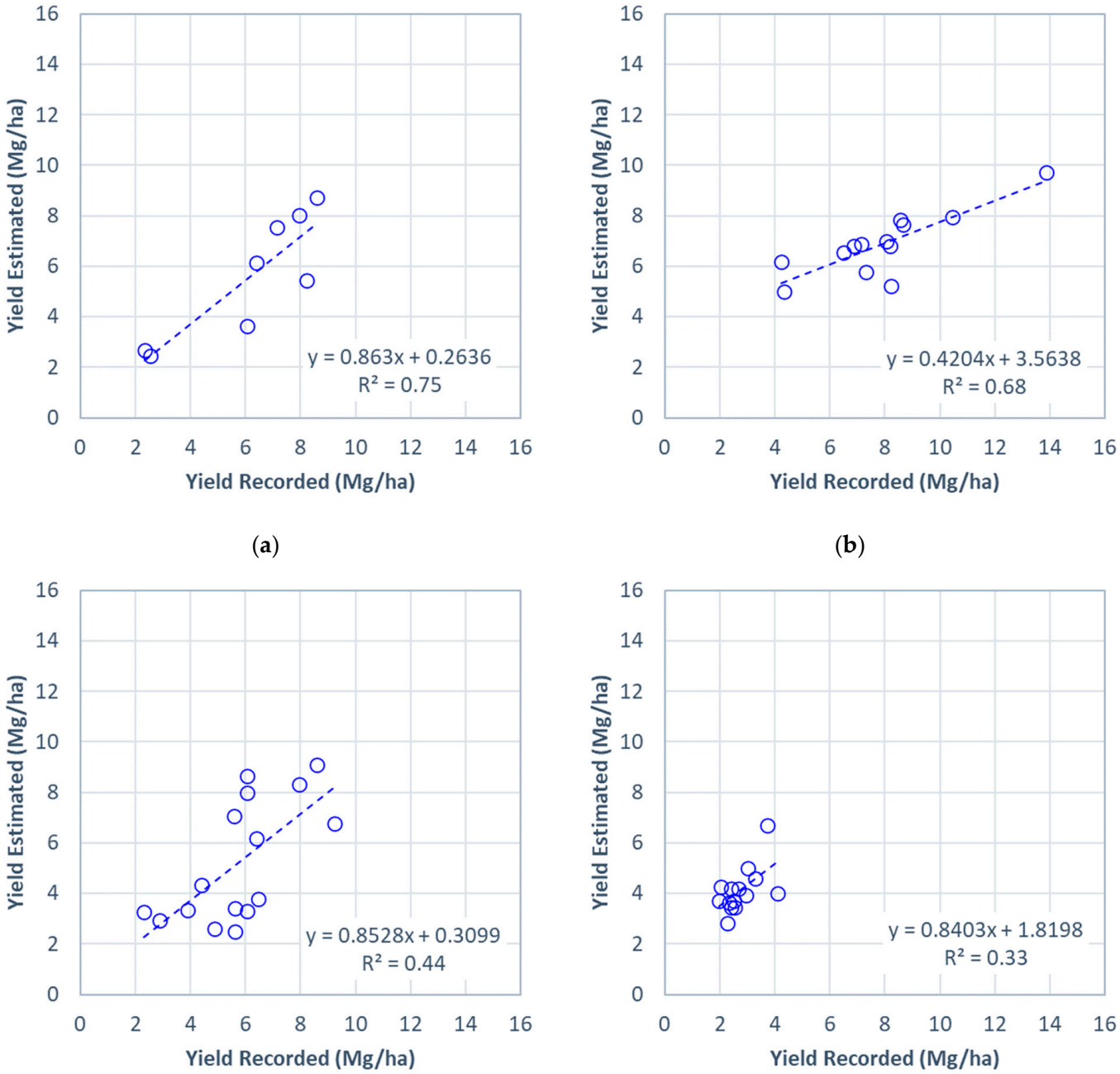


Figure 4. Scatter plots of the estimated and recorded perennial bioenergy grass biomass yields at the field level for (**a**) all sites: (**b**) Site A, (**c**) Site B, and (**d**) Site C. The biomass yields were estimated using a linear regression model with the simulated NDRE for 11 August 2019, pooled from (**a**) all three sites (all-site estimation) and (**b**–**d**) the other two sites (cross-site estimation). The scatter plot for the all-site estimation (**a**) represents the model iteration resulting in the median R^2.

3.4. Statistical Significance of the NDRE for Estimating the Perennial Bioenergy Grass Biomass Yields Using a Linear Regression Model

Table 2 summarizes the statistics of the estimated perennial bioenergy grass biomass yields using a linear NDRE regression model for the all-site and cross-site estimations. For the cross-site estimation, Site B, growing the most diverse perennial bioenergy grasses (Table 1), had the largest error (RMSE = 1.9 Mg/ha, MAE = 1.5 Mg/ha; Table 2) of all three sites. Site C, having the lowest average biomass yields (2.7 Mg/ha; Table 1), had the smallest error (RMSE = 1.6 Mg/ha, MAE = 1.4 Mg/ha; Table 2). For the all-site estimation, the correlation between the estimated and recorded biomass yields had a smaller error (RMSE = 1.5 Mg/ha, MAE = 1.2 Mg/ha) than any of the cross-site estimations (Table 2).

Table 2. Accuracy of the estimated perennial bioenergy grass biomass yields based on a linear regression model using the simulated NDRE for 11 August 2019.

Site	R^2	RMSE (Mg/ha)	MAE (Mg/ha)
All	0.76	1.5	1.2
A	0.69	1.8	1.4
B	0.44	1.9	1.5
C	0.33	1.6	1.4

The perennial bioenergy grass biomass yields for Sites A, B, and C were estimated using the data pooled from the other two sites. The biomass yields for all sites were estimated using the data pooled from all three sites. R^2 = coefficient of determination, RMSE = root mean square error, MAE = mean absolute error.

As the statistics show, the pooled data from all sites would be most successful for estimating the perennial bioenergy grass biomass yields. To examine the efficacy of the model with the all-site approach, we performed an F-test based on ten iterations of the linear regression analysis using 75% of the randomly selected data points (i.e., 31 fields) from the entire dataset for each iteration. The R^2 values for those iterations ranged from 0.70 to 0.80. The F values ranged from 67.6 to 114.3, all of which were outside the 99% confidence interval of between 0 and 7.56. The F values were statistically significant ($p < 0.01$) for all iterations. These results suggest that the NDRE was significantly effective for estimating the perennial bioenergy grass biomass yields using a linear regression model with data pooled from all sites.

3.5. Bias and Consistency in Estimating the Field-Level Perennial Bioenergy Grass Biomass Yields Using a Linear NDRE Regression Model

The plots in Figure 5 show errors in the estimated field-level perennial bioenergy grass biomass yields using a linear NDRE regression model based on the all-site and cross-site estimation approaches. The estimation error for the all-site estimation was approximately ±0.5 Mg/ha across the recorded yields, except for two fields with high yields, which did not establish a clear bias or inconsistency in the biomass yield estimation (Figure 5a). Each of the three sites showed distinct patterns in the cross-site estimation (Figure 5b–d). The model increasingly underestimated the biomass yields as the recorded yields increased, indicating inconsistent estimation across the yields within the site (Figure 5b). The magnitude of the estimation errors for Site B were evenly distributed across the recorded yields, indicating a consistent estimation with a minimal bias in the yield estimation (Figure 5c). For Site C, the model consistently overestimated the biomass yields across the yields, indicating a positive bias in the estimation (Figure 5d). Inadequate representation of low yields in the data from Sites A and B likely caused the bias (Table 1).

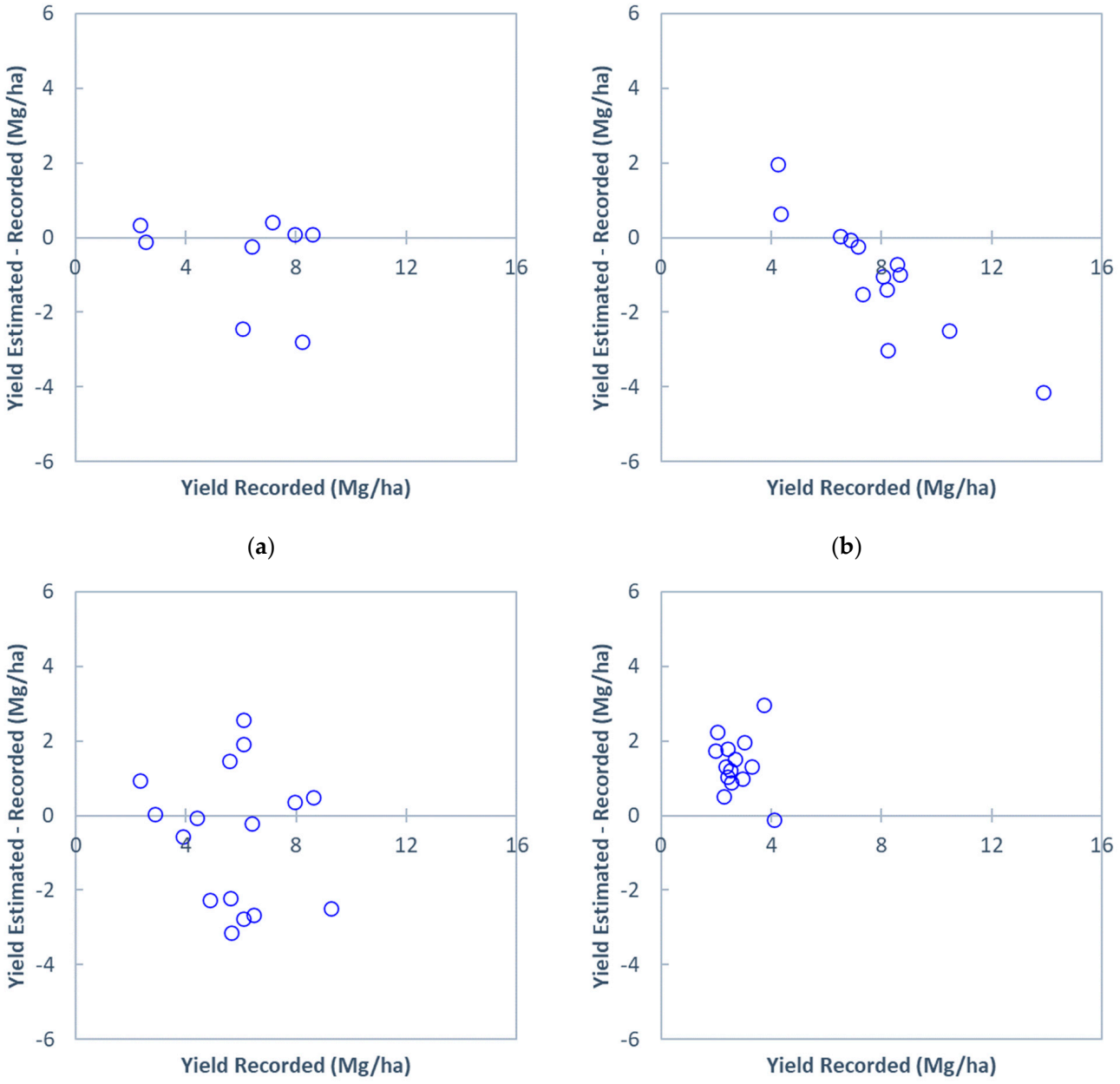


Figure 5. Plots of errors in the estimated field-level perennial bioenergy grass biomass yields at (**a**) all sites, (**b**) Site A, (**c**) Site B, and (**d**) Site C. The biomass yields were estimated using a linear regression model with the simulated NDRE for 11 August 2019, pooled from (**a**) all three sites (all-site estimation) and (**b**–**d**) the other two sites (cross-site estimation). The scatter plot for the all-site estimation (**a**) represents the model iteration resulting in the median R^2.

4. Discussion

4.1. Impact of the Harvest Date on Determining a Common Optimal Date for Multiple Sites

Determining the optimal date for estimating the perennial bioenergy grass biomass yields across multiple sites was not straightforward, despite the availability of daily NDRE values for all sites from the simulation. Based on our previous study [32], we expected to find a common optimal date for all sites if index values were available from all dates. However, in the current study, a notably early harvest for Site D (i.e., end of July through to the beginning of August) prevented us from using the date with the strongest index–yield

correlation (i.e., 5 August). Because grass types and the regional climate strongly influence the timing of the harvest, estimating biomass yields using a single date for multiple sites or across regions is challenging, especially if the optimal date approaches the late growing season. This is one anticipated limitation for scaling our current remote sensing method for estimating the perennial bioenergy grass biomass yields.

4.2. Possible Transition of the NDRE Indication from the Early to the Late Growing Season

The optimal late summer date for estimating the perennial bioenergy grass biomass yields suggested a possible change in the indication of the NDRE during the growing season. We observed that the NDRE correlates with specific plant types from the early season to mid-season, while the index correlates with the overall biomass in the late season. Figure 6 presents side-by-side comparisons of accuracy measures (i.e., R^2, MAE, and RMSE) (a) between the all-site estimation in the current study and the site-specific estimation in our previous study [32] and (b) between the all-site and cross-site estimations in the current study. The site-specific estimation using the site-specific optimal date in mid-summer [32] had more accurate estimated biomass yields than the all-site estimation using the common optimal date in late summer (i.e., 11 August), except for Site B (Figure 6a). This observation is consistent with existing studies [23–25], suggesting that plant- or species-specific relationships between the index and biomass are often strongest at or near the peak greenness.

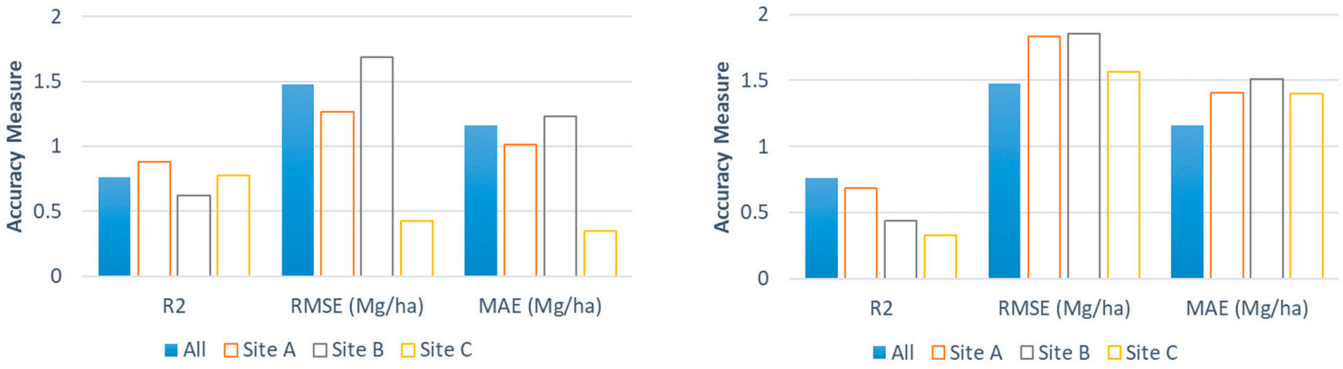


Figure 6. Comparison of the model performance for the field-level perennial grass biomass yield estimation for the (**a**) all-site estimation using 11 August data in the current study vs. the site-specific estimation on site-specific optimal dates in July from the previous study [32] and (**b**) all-site vs. cross-site estimations using 11 August data in the current study.

Unlike the previous comparison, when using a date in late summer (i.e., 11 August), the all-site estimation had more accurate estimated biomass yields than those from the cross-site estimation (Figure 6b). These contrasting observations suggest that, from early summer to mid-summer, the NDRE values are strongly correlated with the biomass yields of each perennial bioenergy grass type with distinct characteristics (e.g., color, structure) or a combination of those grasses when the biomass accumulation is at or near its peak. In the late season, the correlations of NDRE values changed to the overall biomass as those distinct characteristics weakened through senescence. Examining relationships between the recorded yields and index values by perennial bioenergy grass types across multiple dates during a season would allow for the testing of this hypothesis.

4.3. Challenge in Estimating Biomass Yields for Sites with Hetergeneous Perennial Bioenergy Grass Production

Regardless of the timing of the optimal date, estimating the perennial bioenergy grass biomass yields was challenging for the site growing heterogeneous grass types

across the fields. The accuracy of the estimated biomass yields for Site B was consistently lower than that of the other sites across the measures, regardless of the estimation (Figure 6). The site had various perennial bioenergy grass types growing across the fields, including switchgrass, *Miscanthus*, and mixed grass (Table 1). This heterogeneity created varying NDRE–yield relationships within the site, resulting in a moderate correlation between the estimated and recorded biomass yields, even on the date with the strongest NDRE–yield relationship for the site (13 July, R^2 = 0.61 [32]). Stratifying the Site B fields by grass types would enable analyses of type-specific NDRE–yield relationships with those of Site A (switchgrass) and Site C (mixed grass). These analyses could reveal optimal dates by perennial grass instead of by site, which would influence the strategy for scaling the estimation of the perennial bioenergy grass biomass yield across large areas.

For this heterogeneous site, the all-site estimation using 11 August data had more accurate estimated biomass yields than both cross-site and site-specific estimations (Figure 6). This may support the hypothesis of the attenuation of the species-specific index–yield relationship later in the season in addition to the observations discussed in the previous section.

4.4. Limitations Associated with Inconsistent Yield Ranges across the Study Sites

The primary reason for larger discrepancies between the estimated and recorded perennial bioenergy grass biomass yields for the cross-site estimations (Table 2, Figure 6a) than those for the all-site estimations (Table 2, Figure 6a) was the inconsistent yield ranges across the study sites. For example, low yields for Site C were not represented in the model developed using the data from Sites A and B, which had moderate to higher yields (Table 1); thus, the model did not have the ability to accurately estimate low yields for Site C. Similarly, higher yields for Site A were not well explained in the model by the data from Sites B and C, which had lower to moderate yields; thus, the model could not accurately estimate high yields for Site A. Properly evaluating the cross-site approach would require yield ranges comparable across the sites or those with a great overlap.

5. Conclusions

We investigated the indicative power of NDRE for estimating field-level perennial bioenergy grass biomass yields for study sites with mature stands in central Virginia, USA using a linear regression model. To compensate for the varying image availability across the sites, we simulated daily NDRE values across the 2019 season based on the temporally weighted averaging of two temporal neighbors in the index calculated from the available images. Using the simulated NDRE values on the date with the strongest index–yield correlation, we estimated the field-level perennial bioenergy grass biomass yields for the three sites using data pooled from all of those sites (all-site estimation) and for each site using the data pooled from the other two sites (cross-site estimation). We compared the correlations and discrepancies between the estimated and recorded biomass yields for the two estimation approaches as well as with those from the previous site-specific study [32]. Overall, the results showed a strong potential for utilizing a linear NDRE regression model for estimating field-level perennial bioenergy grass biomass yields when using the data pooled from all study sites. In addition to this finding, there were key takeaways from this study for scaling the remote sensing method tested in this study:

1. Inconsistent with the existing studies [24,25,32], the optimal date for estimating field-level perennial bioenergy grass biomass yields for the central Virginia sites using the pooled data was in late summer (11 August). We hypothesized that the index–biomass relationships specific to a perennial bioenergy grass type are the strongest in early summer to mid-summer [24,32] and become more indicative of the overall biomass than the biomass specific to a certain grass type as crops mature and senesce later in summer. This could be an important consideration for scaling the perennial bioenergy grass biomass yield estimation for a large number of sites. Examining relationships between the recorded yields and index values by perennial bioenergy grass types from early summer to right before harvest would allow the testing of this hypothesis.

2. We should expect an earlier harvest for sites with a short growing season or for growing perennial grasses with earlier peak biomass yields than other sites. The optimal date may overlap with, or occur after, harvest for those sites. The sites with early harvest may require a separate model because of their potentially distinct yield–index relationships. This would be a critical limitation for scaling the current method for estimating perennial bioenergy grass biomass yields over large areas.
3. Regardless of the approach (i.e., all-site, cross-site, or site-specific estimations), estimating the biomass yields for a site with heterogeneous perennial bioenergy grasses is more difficult than for sites with a monoculture or uniform mixed grass for the fields within a site. Stratifying the fields by grass types would allow type-specific index–yield relationships to be compared for fields with comparable crop types. This comparison could determine the optimal date for yield estimation by grass type. It could also identify grasses that may hinder the remotely sensed biomass yield estimation using a single model due to their unique spectral-response characteristics.

Based on the above takeaways, we recommend an investigation of the index–yield relationships across perennial bioenergy grasses using sites with comparable yield ranges to understand (1) the specificity of the index–yield relationship by the perennial bioenergy grass type, (2) the change in its strength across seasons, and (3) the possibility of generating a universal model for the perennial bioenergy grass biomass yield estimation across large areas. Such a study for multiple seasons could further the development of a scalable biomass yield estimation approach for perennial bioenergy grasses using remote sensing.

Author Contributions: Conceptualization, Y.H.; methodology, Y.H. and C.R.Z.; validation, Y.H.; formal analysis, Y.H.; investigation, Y.H.; resources, J.J.Q.; data curation, C.R.Z., Y.H. and J.J.Q.; writing—original draft preparation, Y.H.; writing—review and editing, C.R.Z., J.J.Q., Y.H. and M.C.N.; visualization, Y.H.; supervision, J.J.Q.; project administration, J.J.Q.; funding acquisition, M.C.N. and J.J.Q. All authors have read and agreed to the published version of the manuscript.

Funding: The US Department of Energy (DOE), Office of Energy Efficiency and Renewable Energy, Bioenergy Technologies Office, under Award No. DE-EE0022598. Argonne National Laboratory is managed by UChicago Argonne, LLC, for the US DOE under contract DE-AC02-06CH11357.

Data Availability Statement: The data used in the study are confidential.

Acknowledgments: The authors extend their appreciation to Antares and FDC Enterprises for sharing their datasets and to Tim Rooney, Bill Belden, and Kevin Comer for their helpful discussions. The authors also thank Katherine Szoldatits for generating a figure in the manuscript and the anonymous reviewers for their suggestions and constructive criticism for improving the quality of this manuscript.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

References

1. Werling, B.P.; Dickson, T.L.; Isaacs, R.; Gaines, H.; Gratton, C.; Gross, K.L.; Liere, H.; Malmstrom, C.M.; Meehan, T.D.; Ruan, L.; et al. Perennial grasslands enhance biodiversity and multiple ecosystem services in bioenergy landscapes. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 1652–1657. [CrossRef] [PubMed]
2. Schulte, L.A.; Niemi, J.; Helmers, M.J.; Liebman, M.; Arbuckle, J.G.; James, D.E.; Kolka, R.K.; O'neal, M.E.; Tomer, M.D.; Tyndall, J.C.; et al. Prairie strips improve biodiversity and the delivery of multiple ecosystem services from corn–soybean croplands. *Proc. Natl. Acad. Sci. USA* **2017**, *114*, 11247–11252. [CrossRef] [PubMed]
3. Acharya, B.S.; Blanco-Canqui, H. Lignocellulosic-based bioenergy and water quality parameters: A review. *GCB Bioenergy* **2018**, *10*, 504–533. [CrossRef]
4. Wang, E.; Cruse, R.M.; Sharma-Acharya, B.; Herzmann, D.E.; Gelder, B.K.; James, D.E.; Flanagan, D.C.; Blanco-Canqui, H.; Mitchell, R.B.; Laird, D.A. Strategic switchgrass (*Panicum virgatum*) production within row cropping systems: Regional-scale assessment of soil erosion loss and water runoff impacts. *GCB Bioenergy* **2020**, *12*, 955–967. [CrossRef]

5. Zaimes, G.N.; Schultz, R.C. Assessing Riparian Conservation Land Management Practice Impacts on Gully Erosion in Iowa. *Environ. Manag.* **2012**, *49*, 1009–1021. [CrossRef]
6. Yimam, Y.T.; Ochsner, T.E.; Fox, G.A. Hydrologic cost-effectiveness ratio favors switchgrass production on marginal croplands over existing grasslands. *PLoS ONE* **2017**, *12*, e0181924. [CrossRef]
7. Blanco-Canqui, H. Energy Crops and Their Implications on Soil and Environment. *Agron. J.* **2010**, *102*, 403–419. [CrossRef]
8. Zhou, X.; Helmers, M.J.; Asbjornsen, H.; Kolka, R.; Tomer, M.D.; Cruse, R.M. Nutrient removal by prairie filter strips in agricultural landscapes. *J. Soil Water Conserv.* **2014**, *69*, 54–64. [CrossRef]
9. de Freitas, E.N.; Khatri, V.; Contin, D.R.; de Oliveira, T.B.; Contato, A.G.; Peralta, R.M.; dos Santos, W.D.; Martinez, C.A.; Saddler, J.N.; Polizeli, M.d.L.T. Climate change affects cell-wall structure and hydrolytic performance of a perennial grass as an energy crop. *Biofuels Bioprod. Biorefining* **2022**, *16*, 471–487. [CrossRef]
10. Jager, H.I.; Hilliard, M.R.; Langholtz, M.H.; Efroymson, R.A.; Brandt, C.C.; Nair, S.S.; Kreig, J.A. Ecosystem service benefits to water users from perennial biomass production. *Sci. Total. Environ.* **2022**, *834*, 155255. [CrossRef]
11. Brandes, E.; McNunn, G.S.; A Schulte, L.; Bonner, I.J.; Muth, D.J.; A Babcock, B.; Sharma, B.; A Heaton, E. Subfield profitability analysis reveals an economic case for cropland diversification. *Environ. Res. Lett.* **2016**, *11*, 014009. [CrossRef]
12. Ssegane, H.; Negri, M.C.; Quinn, J.; Urgun-Demirtas, M. Multifunctional landscapes: Site characterization and field-scale design to incorporate biomass production into an agricultural system. *Biomass-Bioenergy* **2015**, *80*, 179–190. [CrossRef]
13. Ssegane, H.; Negri, M.C. An Integrated Landscape Designed for Commodity and Bioenergy Crops for a Tile-Drained Agricultural Watershed. *J. Environ. Qual.* **2016**, *45*, 1588–1596. [CrossRef] [PubMed]
14. Ahamed, T.; Tian, L.; Zhang, Y.; Ting, K. A review of remote sensing methods for biomass feedstock production. *Biomass-Bioenergy* **2011**, *35*, 2455–2469. [CrossRef]
15. Schmer, M.R.; Mitchell, R.B.; Vogel, K.P.; Schacht, W.H.; Marx, D.B. Spatial and Temporal Effects on Switchgrass Stands and Yield in the Great Plains. *BioEnergy Res.* **2010**, *3*, 159–171. [CrossRef]
16. Cacho, J.F.; Feinstein, J.; Zumpf, C.R.; Hamada, Y.; Lee, D.J.; Namoi, N.L.; Lee, D.; Boersma, N.N.; Heaton, E.A.; Quinn, J.J.; et al. Predicting Biomass Yields of Advanced Switchgrass Cultivars for Bioenergy and Ecosystem Services Using Machine Learning. *Energies* **2023**, *16*, 416. [CrossRef]
17. Chen, Y.; Guerschman, J.; Shendryk, Y.; Henry, D.; Harrison, M.T. Estimating Pasture Biomass Using Sentinel-2 Imagery and Machine Learning. *Remote Sens.* **2021**, *13*, 603. [CrossRef]
18. Reinermann, S.; Asam, S.; Kuenzer, C. Remote Sensing of Grassland Production and Management—A Review. *Remote Sens.* **2020**, *12*, 1949. [CrossRef]
19. Ahmad, A.; Gilani, H.; Ahmad, S.R. Forest Aboveground Biomass Estimation and Mapping through High-Resolution Optical Satellite Imagery—A Literature Review. *Forests* **2021**, *12*, 914. [CrossRef]
20. Chao, Z.; Liu, N.; Zhang, P.; Ying, T.; Song, K. Estimation methods developing with remote sensing information for energy crop biomass: A comparative review. *Biomass-Bioenergy* **2019**, *122*, 414–425. [CrossRef]
21. Giovos, R.; Tassopoulos, D.; Kalivas, D.; Lougkos, N.; Priovolou, A. Remote Sensing Vegetation Indices in Viticulture: A Critical Review. *Agriculture* **2021**, *11*, 457. [CrossRef]
22. Jensen, J.R. *Remote Sensing of the Environment: An Earth Resource Perspective*; Prentice Hall: Upper Saddle River, NJ, USA, 2003.
23. Kanemasu, E.T.; Demetriades-Shah, T.H.; Su, H.; Lang, A.R.G. Estimating Grassland Biomass Using Remotely Sensed Data. *Appl. Remote Sens. Agric.* **1990**, 185–199. [CrossRef]
24. Hamada, Y.; Zumpf, C.R.; Cacho, J.F.; Lee, D.; Lin, C.-H.; Boe, A.; Heaton, E.; Mitchell, R.; Negri, M.C. Remote Sensing-Based Estimation of Advanced Perennial Grass Biomass Yields for Bioenergy. *Land* **2021**, *10*, 1221. [CrossRef]
25. Namoi, N.; Jang, C.; Robins, Z.; Lin, C.-H.; Lim, S.-H.; Voigt, T.; Lee, D. Aerial Imagery Can Detect Nitrogen Fertilizer Effects on Biomass and Stand Health of *Miscanthus* × *giganteus*. *Remote Sens.* **2022**, *14*, 1435. [CrossRef]
26. Foster, A.J.; Kakani, V.G.; Mosali, J. Estimation of bioenergy crop yield and N status by hyperspectral canopy reflectance and partial least square regression. *Precis. Agric.* **2016**, *18*, 192–209. [CrossRef]
27. Tucker, C.J. Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sens. Environ.* **1979**, *8*, 127–150. [CrossRef]
28. 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]
29. 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]
30. Naguib, N.S.; Daliman, S. Analysis of NDVI and NDRE Indices Using Satellite Images for Crop Identification at Kelantan. *IOP Conf. Ser. Earth Environ. Sci.* **2022**, *1102*, 012054. [CrossRef]
31. Chang, J.; Shoshany, M. Red-edge ratio normalized vegetation index for remote estimation of green biomass. In Proceedings of the 2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Beijing, China, 10–15 July 2016; pp. 1337–1339. [CrossRef]
32. Hamada, Y.; Zumpf, C.R.; Quinn, J.J. *Predicting Switchgrass Biomass Yields Using a Spectral Vegetation Index Derived from Multispectral Satellite Imagery*; Technical Report, ANL/EVS-23/20; Argonne National Laboratory (ANL): Argonne, IL, USA, 2023; 44p.

33. Datt, B. A New Reflectance Index for Remote Sensing of Chlorophyll Content in Higher Plants: Tests using Eucalyptus Leaves. *J. Plant Physiol.* **1999**, *154*, 30–36. [CrossRef]
34. Rouse, J.W.; Haas, R.H.; Schell, J.A.; Deering, D.W. Monitoring Vegetation Systems in the Great Plains with ERTS. *NASA Spec. Publ.* **1974**, *351*, 309.
35. Jiang, Z.; Huete, A.R.; Didan, K.; Miura, T. Development of a two-band enhanced vegetation index without a blue band. *Remote Sens. Environ.* **2008**, *112*, 3833–3845. [CrossRef]
36. Jin, S.; Sader, S.A. Comparison of time series tasseled cap wetness and the normalized difference moisture index in detecting forest disturbances. *Remote Sens. Environ.* **2005**, *94*, 364–372. [CrossRef]
37. Klemas, V.; Smart, R. The Influence of Soil Salinity, Growth Form and Leaf Moisture on the Spectral Radiance of Spartina alterniflora Canopies. *Photogramm. Eng. Remote Sens.* **1983**, *49*, 77–83.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.