# remote sensing

*Article*

# Classification of C3 and C4 Vegetation Types Using MODIS and ETM+ Blended High Spatio-Temporal Resolution Data

**Xiaolong Liu [1,2,3], Yanchen Bo [1,2,\*], Jian Zhang [1,2] and Yaqian He [1,2,4]**

[1] State Key Laboratory of Remote Sensing Science, Research Center for Remote Sensing and GIS, and School of Geography, Beijing Normal University, Beijing 100875, China;
E-Mails: liuxl@mail.bnu.edu.cn (X.L.); janechang_910330@163.com (J.Z.);
yahe@mix.wvu.edu (Y.H.)
[2] Beijing Key Laboratory for Remote Sensing of Environment and Digital Cities, Beijing 100875, China
[3] College of Tourism & Geography Science, Yunnan Normal University, Kunming 650500, China
[4] Department of Geology and Geography, West Virginia University, Morgantown, WV 26506, USA

**\*** Author to whom correspondence should be addressed; E-Mail: boyc@bnu.edu.cn;
Tel.: +86-10-5880-2062; Fax: +86-10-5880-5274.

**Abstract:** The distribution of C3 and C4 vegetation plays an important role in the global carbon cycle and climate change. Knowledge of the distribution of C3 and C4 vegetation at a high spatial resolution over local or regional scales helps us to understand their ecological functions and climate dependencies. In this study, we classified C3 and C4 vegetation at a high resolution for spatially heterogeneous landscapes. First, we generated a high spatial and temporal land surface reflectance dataset by blending MODIS (Moderate Resolution Imaging Spectroradiometer) and ETM+ (Enhanced Thematic Mapper Plus) data. The blended data exhibited a high correlation ($R^2$ = 0.88) with the satellite derived ETM+ data. The time-series NDVI (Normalized Difference Vegetation Index) data were then generated using the blended high spatio-temporal resolution data to capture the phenological differences between the C3 and C4 vegetation. The time-series NDVI revealed that the C3 vegetation turns green earlier in spring than the C4 vegetation, and senesces later in autumn than the C4 vegetation. C4 vegetation has a higher NDVI value than the C3 vegetation during summer time. Based on the distinguished characteristics, the time-series NDVI was used to extract the C3 and C4 classification features. Five features were selected from the 18 classification features

according to the ground investigation data, and subsequently used for the C3 and C4 classification. The overall accuracy of the C3 and C4 vegetation classification was 85.75% with a kappa of 0.725 in our study area.

## 1. Introduction

Research on biogeochemical cycling, the global carbon cycle and climate change have shown that the spatial distribution of C3 and C4 vegetation is relevant to atmospheric $CO_2$ and temperature changes [1–4]. C4 plants tend to favor warmer environments (warm season plants), and C3 plants thrive in areas with lower temperatures (cool season plants) in the mid-latitudes [3,5]. The balance between C3 and C4 plants changes with the atmospheric $CO_2$ content variation [6,7]. Therefore, mapping of C3 and C4 plants is important for the study of regional climate change and carbon cycle.

Previous studies have attempted to map C3 and C4 plants using remote sensing data. Hyperspectral data have proven to be effective in mapping C3 and C4 plants. Irisarri claimed that the hyperspectral data could discriminate C3 and C4 plants inside a laboratory [8]. Hyperspectral remote sensing data with bands centered at 470 nm, 530 nm, 600 nm, 660 nm, 700 nm, 720 nm, 820 nm, 1540 nm, 2060 nm, 2280 nm, 2300 nm, 2450 nm and 2470 nm showed potential for classifying C3 and C4 plants, and features selected from these bands were used to classify C3 and C4 plants [9]. The chlorophyll fluorescence derived from hyperspectral remote sensing data was used for discriminating C3 and C4 plants [10–12]. A later study showed that chlorophyll fluorescence could be detected from space [13], which indicated that it is possible to classify C3 and C4 vegetation using space borne remote sensing data. However, satellite borne sensors, such as the Thermal And Near-infrared Sensor for carbon Observation-Fourier Transform Spectrometer (TANSO-FTS) on the Japanese Greenhouse gases Observing SATellite (GOSAT), the MEdium Resolution Imaging Spectrometer (MERIS) aboard the European Space Agency's (ESA's) ENVIronmental SATellite (ENVISAT), and the TOMS (Total Ozone Mapping Spectrometer) aboard the Orbiting Carbon Observatory-2 (OCO-2) launched by NASA have coarse spatial resolutions (300 m for ENVISAT, 10.5 km for GOSAT and $1.29 \times 2.25$ km for OCO-2). The coarse spatial resolution of satellite data leads to a large number of mixed pixels, especially for fragmented landscapes. This implies that the chlorophyll fluorescence derived from satellite remote sensing data is not yet suitable for high spatial resolution C3 and C4 plant classification over the regional scale.

Physiologically, C3 and C4 lifeforms are distinguished by their different photosynthetic pathways through which carbon is fixed into carbohydrates. Vegetation utilizing C3 and C4 photosynthetic pathways exhibit physiological and morphological differences that result in dissimilar responses to environmental conditions, such as light saturation, maximum rate of net photosynthesis, optimum temperature for net photosynthesis, transpiration rate and growth rate [14,15]. Although, with their sensitivity to varying environmental disturbances, C3 and C4 species exhibit markedly different seasonal activity cycles [16]. Compared to C4 species, C3 species green up earlier and are more active under the cooler temperature conditions of spring and early fall. In contrast, C4 species green up later in the growing season and are more active under the warmer, drier conditions of mid to late summer [17]. These contrasts in phenological

characteristics make the C3 and C4 vegetation types exhibiting asynchronous seasonality in their pattern of greenness.

Because of the seasonal differences between C3 and C4 plants, measurement of vegetation greenness (e.g., NDVI) derived from time series of remote sensing data have the potential to discriminate C3 and C4 plants. Studies have been conducted to determine the temporal offsets of photosynthetic activity for these two types of vegetation [5,17–19]. Foody *et al.* proved the possibility of mapping C3 and C4 vegetation composition in South Dakota, US by capturing the asynchronous seasonal profile from time series of MTCI (MERIS terrestrial chlorophyll index) data product [19]. Wang *et al.* classified C3 and C4 type of grasses using time-series MODIS derived phenology metrics in the U.S. Great Plains [5]. Compared to the satellite remote sensing data with many mixed pixels and the airborne data that are economically-costly for the regional scale C3 and C4 distribution mapping, the high temporal resolution data products that can be used as time-series data are more suitable. The high temporal resolution data products, including AVHRR, MODIS and MTCI MERIS, *etc.* are usually used to capture the phenological asynchronicity of C3 and C4 vegetation. However, the existing high temporal resolution data, such as MODIS and AVHRR data, are not suitable to be used for mapping C3 and C4 plants in regions with fragmented landscapes due to their coarse spatial resolutions. Remote sensing data with finer spatial resolutions, such as Landsat TM/ETM+, however, could not capture the fine dynamics of the vegetation due to their long revisiting cycles.

Data with both high spatial and temporal resolutions are still not available to extract the time-series features for C3 and C4 classification at a regional scale. To generate time-series satellite data with high spatial and temporal resolutions, several data fusion models have been proposed and have been proven to be practicable. Gao *et al.* developed a Spatial and Temporal Adaptive Reflectance Fusion Model (STARFM) for blending the Landsat and MODIS data to generate daily surface reflectance data at a 30 m spatial resolution [20]. To overcome the shortcoming of STARFM in inaccurate prediction of the surface reflectance over heterogeneous landscapes, an Enhanced Spatial and Temporal Adaptive Reflectance Fusion Model (ESTARFM) has been proposed to generate more accurate land surface reflectance in heterogeneous regions [21].

The aim of this study is to propose a framework for mapping C3 and C4 vegetation types in spatially heterogeneous landscapes using high spatio-temporal resolution remote sensing data. The high spatial resolution time-series data was derived from fusing daily MODIS land surface reflectance data and Landsat ETM+ data. The study area is the middle reaches of the Heihe River Basin, Gansu Province, China, where the landscape is fragmented and the vegetation distribution is complicated. The proposed framework includes: (1) to generate daily 30m resolution land surface reflectance data by fusing MODIS and Landsat ETM+ data using the ESFARFM and to evaluate the accuracy of the fused data; (2) to generate the NDVI time-series data using the fused land surface reflectance data, and to extract classification features from the NDVI time-series data for C3 and C4 plant classification; and (3) to classify the C3 and C4 plants in the middle reaches of Heihe Watershed, China, using the SVM (Support Vector Machine) classifier with selected classification features.

This paper is organized as follows. In Section 2, we describe the study area and the dataset used in the analysis. In Section 3, we present our methods for the data pre-processing, data fusion accuracy assessment, C3 and C4 vegetation seasonality feature extraction and the C3 and C4 vegetation classification.

The results are presented and analyzed in Section 4. A summary of the whole study and the conclusions are provided in Section 5.

## 2. Study Area and Data

### 2.1. Study Area

The study area is located in the middle reaches of the Heihe River basin (north corner: 40°19′31″N, 99°9′34″E; east corner: 38°39′23″N, 101°45′34″E; south corner: 101°4′26″N, 38°6′24″E; west corner: 39°33′16″N, 98°21′33″E), Gansu Province, northwest of China (Figure 1). The study area is characterized by its fragmented landscapes: sparse natural vegetation and artificial oasis across the area, such as trees, shrubs, grassland, irrigated crops, riparian ecosystem, and desert. The field investigation revealed that there are both natural and planted C3 and C4 plants in the study area (List 1).
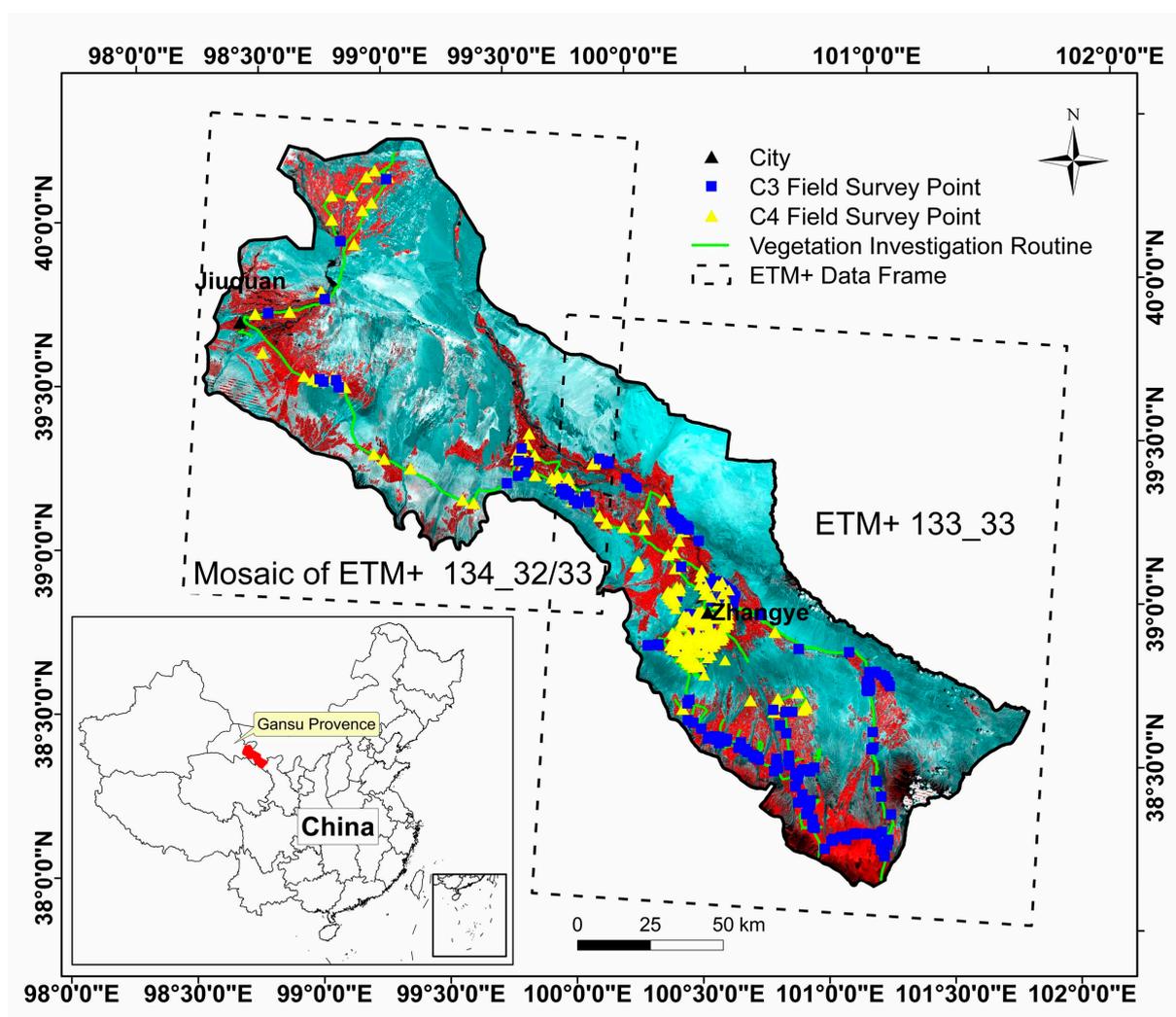


**Figure 1.** Location of the study area. The image is the false color composition of Landsat ETM+ data (R: band 4, G: band 3, B: band 2). The yellow triangles represent the field investigated C4 vegetation and the blue squares represent the C3 vegetation. The bright green line is the field vegetation survey routine.

**List 1.** C3 and C4 plants in the study area.

| C3 Plants | C4 Plants |
|---|---|
| Bulrush (*Phragmites australis* L.), Rice (*Oryza sativa* L.), | Sacsaoul (*Haloxylon Aammodendron* L.), |
| Castor (*Ricinus communis* L.), Sugar Beet (*Beta vulgaris* L.), | Sorghum (*Sorghum bicolor* L.), Maize (*Zea mays* L.), |
| Sunflower (*Helianthus annuus* L.), Soybean (*Glycine max* L.), | Calligonium (*Calligonum arborescens* L.), |
| Wheat (*Triticum aestivum* L.), Alfalfa (*Medicago sativa* L.), | SalsolaCollina (*Salsola collina Pall* L.), |
| White Poplar (*Populus alba* L.), Potato (*Solanum tuberosum* L.), | Pteridophyta (*Suaeda dendroides* (*C. A. Mey.*) *Moq.* L.) |
| Rape (*Brassica napus* L.), Rape (*Hippophae rhamnoides* L.), | |
| Barley (*Hordeum vulgare* L.) | |

*2.2. Data*

The MODIS and ETM+ reflectance data have long been acknowledged as being of good quality and continuity. The MODIS data we used was the Surface Reflectance product (MOD09GQ) of the fifth version (downloaded from: http://reverb.echo.nasa.gov/reverb). The MOD09GQ is daily land surface reflectance data with red (R) and near infrared (NIR) spectral bands, and its spatial resolution is 250 m. In comparison with the eight-day or 16-day products, the daily reflectance data can better capture the phenological differences between the C3 and C4 plants, especially during the critical growth and senescence stages in early spring and fall. Plants in this study area begin to green-up during the middle of April and senesce in the middle of October. The data used in this study was from 13 April 2012 (103rd day of the year) to 31 October 2012 (304th day of the year), which covered the entire C3 and C4 vegetation growth period of the study area.

The 30m resolution Landsat 7 ETM+ data were downloaded from the USGS Land Processes Distributed Active Archive Center. There are three scenes of ETM+ data to cover the study area. They are 134_32, 134_33, and 133_33. The 134_32 and 134_33 were acquired on the same day because they are on the same orbit from north to south. After the removal of the cloud-covered images, eight scenes of ETM+ data (DOY 112 to DOY 304) were remaining for 133_33, and 10 scenes of ETM+ data for for 134_32/134_33 (Table 1). We used the Landsat ecosystem disturbance adaptive processing system (LEDAPS) to create the Landsat-based surface reflectance data. Through the LEDAPS system, the Landsat data were calibrated, converted to Top-of-Atmosphere (TOA) reflectance, and atmospherically corrected using the 6S model [22]. The geometric correction to the ETM+ and MODIS data was based on the GCPs (Ground Control Points) collected by the HiWATER (Heihe Water Allied Telemetry Experimental Research) project [23], using the rational polynomial model.

The temperature decreases due to high elevation will affect the distribution of C4 vegetation, and therefore must be considered. The GDEM (Global Digital Elevation Model) data (the 2nd version) with 30 m spatial resolution and 20m vertical resolution was used to present the elevation. The GDEM data was downloaded from http://gdem.ersdac.jspacesystems.or.jp/.

The ground investigation was performed from 8 July to 9 August 2012 by the HiWATER team. The field investigation route and the investigation points are shown in Figure 1. The field investigation was intensified in areas where the distribution of vegetation was fragmented, and each of the vegetation types were covered in the field survey. The vegetation being investigated includes natural and artificial plant for their species, height, and acreage of crop stands. We then categorized each vegetation type into

C3 or C4 vegetation. Only patches larger than 25 ha (equivalent to approximately four 250 m pixels in area) in size were chosen as the C3 and C4 "ground truth" data to be used in the classification and accuracy assessment process. The chosen ground truth investigation points include 682 points for C3 and 499 points for C4.

**Table 1.** Remote sensing data used for generating the high spatial and temporal resolution Normalized Difference Vegetation Index (NDVI) data.

| Data | MODIS | ETM+ | |
|---|---|---|---|
| | h25_v04/h25_v05 | 134_32/134_33 | 133_33 |
| **Spatial Resolution** | **250 m** | **30 m** | **30 m** |
| **DOY** | 103–304 (daily) | 103 | -- |
| | | 119 | 112 |
| | | 135 | 176 |
| | | 151 | 192 |
| | | 167 | 224 |
| | | 215 | 240 |
| | | 231 | 272 |
| | | 247 | 288 |
| | | 279 | 304 |
| | | 295 | -- |

## 3. Methodology

Two main steps were performed for the C3 and C4 vegetation classification: generating the time-series NDVI data at a 30m resolution, the classification of the C3 and C4 vegetation based on the time-series NDVI data and its accuracy assessment (Figure 2).
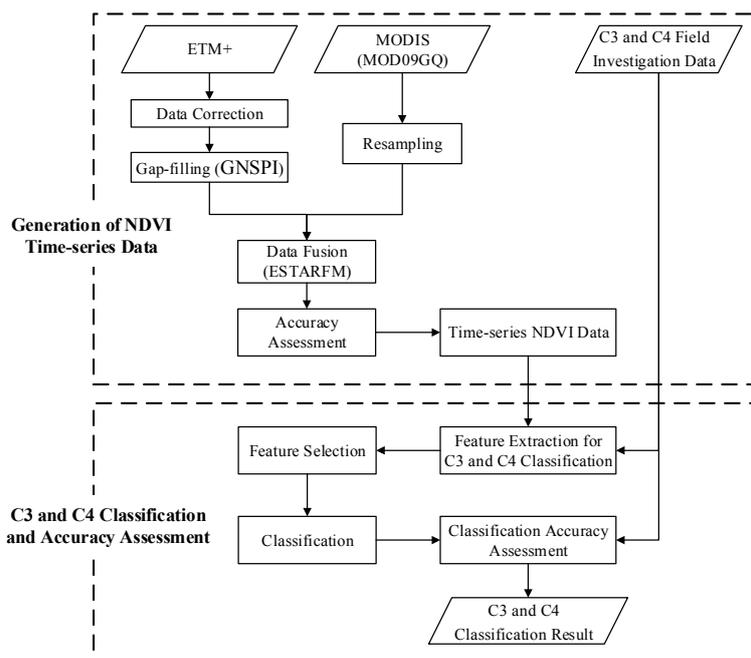


**Figure 2.** Flowchart of the C3 and C4 vegetation classification process.

*3.1. High Temporal and Spatial Resolution NDVI Data Generation*

We employed the ESTARFM data fusion algorithm to derive the high spatial and temporal resolution data. The ESTARFM was proven to be able to accurately predict the surface reflectance and preserve the details in high resolution, especially for heterogeneous landscapes[21]. To predict the daily 30 m spatial resolution land surface reflectance, we employed the MODIS and Landsat ETM+ data for the antecedent and subsequent date.

However, the ETM+ data used included SLC-off (Scan Line Corrector-off) images as only SLC-off images were available for the Landsat series data during 2012. The un-scanned pixels roughly occupy 22% of an ETM+ image, limiting the application of the ETM+ data [24]. Fortunately, a few algorithms were presented to solve this problem [25–27]. Here, we used GNSPI (Geostatistical Neighborhood Similar Pixel Interpolator), an algorithm based on the geostatistical theory and NSPI (Neighborhood Similar Pixel Interpolator) [28], to fill the gaps of the ETM+ SLC-off images before data fusion.

Prior to implementing the ESTARFM data fusion algorithm, we used MODIS Reprojection Tools (MRT) to reproject and resample the MODIS data to the spatial resolution of the ETM+ imagery. The clouded images were excluded according to the QC (Quality Control) data along with the MOD09GQ data. A bilinear algorithm was used in the resampling process to reduce the effect of the georeferencing error. The ESTARFM requires at least two pairs of fine- and coarse-resolution images that were acquired on the same date and a set of coarse-resolution images for desired prediction dates. There are 10 scenes of cloud-free ETM+ data for 134_32/134_33 and 8 scenes for 133_33, as listed in Table 1. To minimize the uncertainty caused by human activity or environmental changes, the temporally closest available data were set as a pair. Hence, there were nine pairs for 134_32/134_33 and 7 pairs for 133_33. In the following data fusion process, the study area was divided into two sub-areas according to the ETM+ data coverage, as is shown in Figure 1.

To derive the phenological parameters for C3 and C4 vegetation classification, the widely used Normalized Difference Vegetation Index (NDVI) was employed [29–31]. The time-series NDVI were calculated using the predicted time-series surface reflectance data of red (R) and infrared (NIR) spectral bands at 30 m resolution.

$$NDVI = \frac{NIR - R}{NIR + R} \tag{1}$$

The time-series NDVI profile of vegetation over the growing season is shown in Figure 3. There are abrupt shifts in the raw time-series NDVI profile, which may be caused by climate and atmospheric variability, bi-directional of reflectance, and sun zenith angle changes that occur all around a year [32–35]. The removal of noise and disturbances are critical for the extraction of the C3 and C4 vegetation phenological features [30,31]. To remove noises in the time-series data, there are many algorithms available [36–39].

In this study, three different noise removal methods, symmetric Gaussian functions, double logistic functions and Savitzky-Golay filtering, were tested with the original time-series data using the TIMESAT program [36]. As shown in Figure 3a,c, the asymmetric Gaussian function and double logistic functions changed the NDVI value unexpectedly before DOY180. This result agreed with the previous research in which the asymmetric Gaussian and double logistic functions were problematic for application to the irregular VI time-series [29,40]. Although the Savitzky-Golay filtering performed better, there was still

some undesirable noise in the curve (Figure 3b). Thus, we performed a second Savitzky-Golay filtering to the first Savitzky-Golay filtered result, and we named it as "Double Savitzky-Golay filtering (Double S-G)". The Double S-G result was improved compared with the other three results in our test, even though it fits to the mean of the NDVI data rather than to the upper envelope.
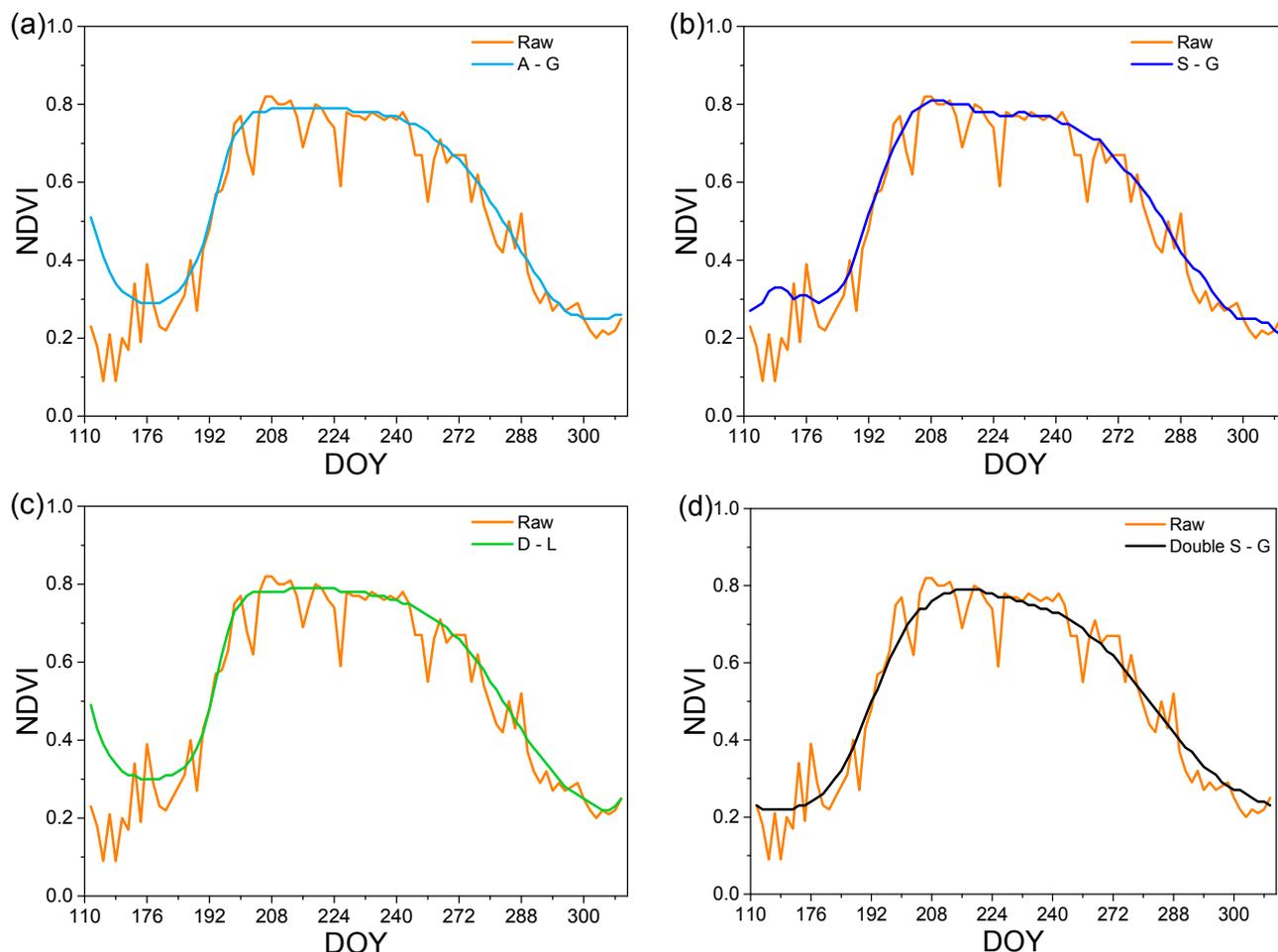


**Figure 3.** Removal of noise in time-series NDVI profile. (**a**) asymmetric Gaussian function; (**b**) Savitzky-Golay filtering; (**c**) double logistic functions; (**d**) Double Savitzky-Golay filtering.

### 3.2. Feature Extraction for C3 and C4 Vegetation Classification

As shown in Figure 4, C3 and C4 plants have distinguishable spatial distribution trends according to altitude. Based on the field investigation points, we found that there were hardly any C4 plants in areas with altitude higher than 2000 m above sea level. Thus, we divided the study area into two areas: the area above 2000 m a.b.s.l. and the area below 2000 m a.b.s.l. (Figure 4). Vegetation in the area above 2000 m a.b.s.l. was classified as C3 functional plant type.

Each C3 or C4 field survey point was plotted as a time-series NDVI profile (Figure 5). Figure 5a,b revealed that within the heterogeneous geographical environment, C3 and C4 vegetation have similar seasonality characteristics. Most of the C3 and C4 NDVI values are similar during the entire growing season. Uncertainties exist in both the C3 and C4 time-series NDVI profiles, and the values increased in summer when the plants were thriving. However, the time-series profile of mean NDVI (solid line in

Figure 5a,b) shows that the NDVI values of the C4 plants were higher than those of C3 plants during the summer. For the green-up and the senescence phases of a season, the NDVI values of the C3 plants were higher than those of C4.
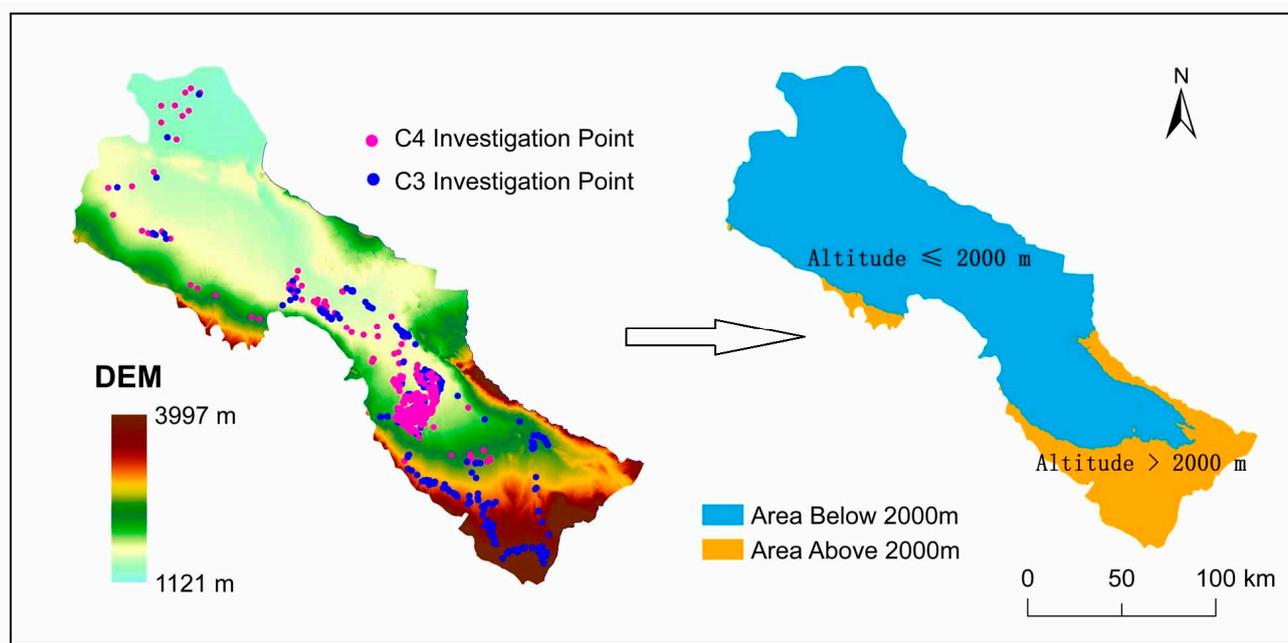


**Figure 4.** C3 and C4 vegetation distribution as a function of the elevation in the study area.
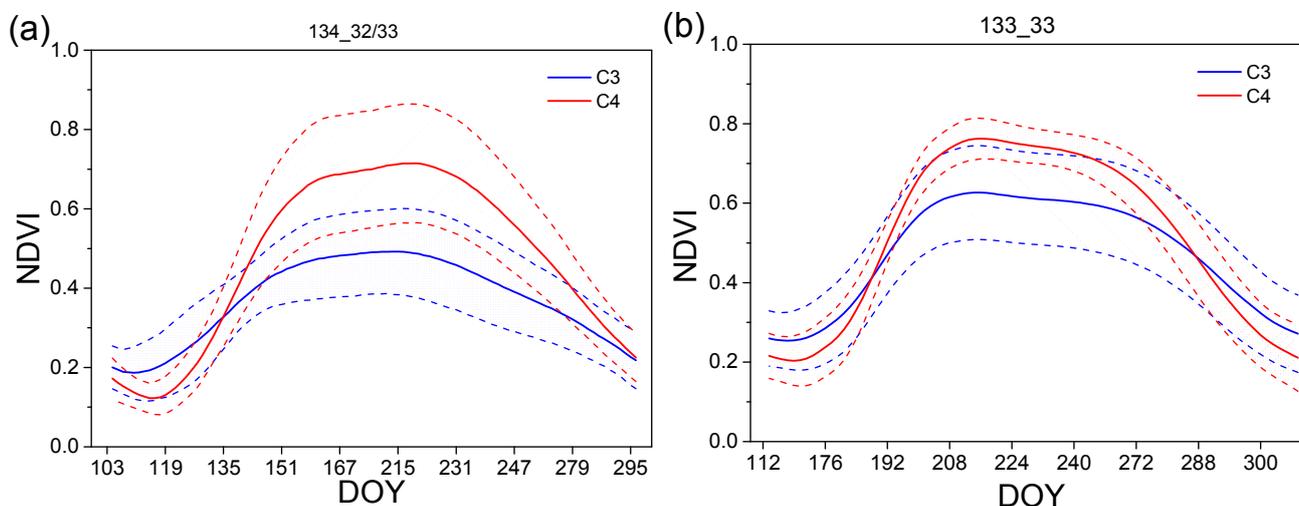


**Figure 5.** Time-series NDVI profile plotted according to field investigation points. Shaded areas indicate the variance of the C3 and C4 NDVI. (**a**) time-series NDVI profile of 134_32/33 sub-area; (**b**) time-series NDVI profile of 133_33 sub-area.

According to the seasonality differences presented in the time-series NDVI profile, 18 features were extracted (Table 2). These features were used to characterize the differences between the C3 and C4 plants. Some of the features are depicted in Figure 6.

**Table 2.** The 18 features that were extracted for the classification of the C3 and C4 vegetation.

| No. | Classification Features | Value Range | Description |
|---|---|---|---|
| 1 | Max NDVI value | −1, 1 | Maximum NDVI value |
| 2 | Min NDVI value | −1, 1 | Minimum NDVI value |
| 3 | Max NDVI value date | 0, 86 | DOY when maximum NDVI value achieved |
| 4 | Min NDVI value date | 0, 86 | DOY when minimum NDVI value achieved |
| 5 | Integral D35-D45 | 0, 10 | Integral between DOY35 and DOY45 |
| 6 | Integral of NDVI | 0, 86 | Integral of NDVI time-series curve |
| 7 | Max_NDVI/Integral | −∞, +∞ | Ratio of maximum NDVI to NDVI curve integral |
| 8 | Max derivative | −∞, +∞ | Maximum derivative value of NDVI profile |
| 9 | Min derivative | −∞, +∞ | Minimum derivative value of NDVI profile |
| 10 | Min_derivative NDVI value | −1, 1 | NDVI value when Minimum derivative achieved |
| 11 | Integral between max and min derivatives | 0, 86 | Integral between maximum and minimum derivatives |
| 12 | Integral between D_SOS and D_max | 0, 86 | Integral between the start of the season and maximum NDVI |
| 13 | Number of local max NDVI value | 0, 86 | Number of local maximum NDVI values |
| 14 | Integral between max and first local max | 0, 86 | Integral between maximum and first local maximum NDVI |
| 15 | D_SOS | 0, 86 | Date of the start of the season |
| 16 | D_EOS | 0, 86 | Date of the end of the season |
| 17 | EOS NDVI value | −1, 1 | NDVI value at the end of the season |
| 18 | SOS NDVI value | −1, 1 | NDVI value at the start of the season |
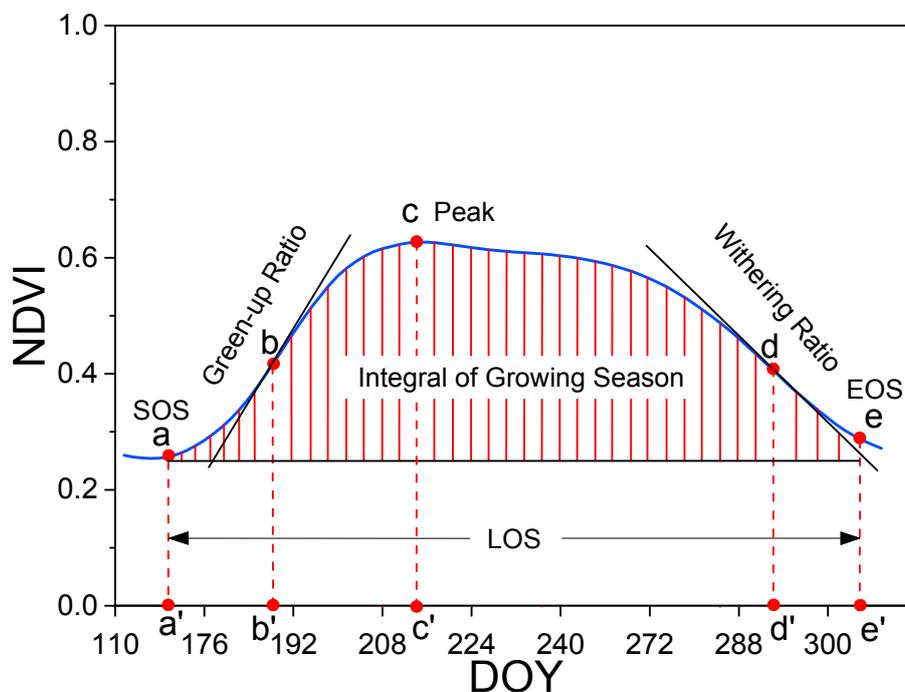


**Figure 6.** C3 and C4 time-series NDVI curve and extracted features. Time-series NDVI curve features: Start of Season (SOS), green-up ratio, peak, withering ratio, End of Season (EOS), Length of Season (LOS), integral of growing season, and their corresponding dates.

It is our concern that whether all features were needed to be used in the classification process, if not, how should a subset be chosen that minimizes any loss of information essential to the C3 and C4 vegetation classification? That indicates that a feature selection process for features listed in Table 2 should be conducted to discard those features that are not effective in C3 and C4 vegetation discrimination [41]. Feature selection was conducted in two steps: (1) feature values were plotted in a box plot to compare their separability, referring to the mean value and the variance of the C3 and C4 classes; (2) the Jeffries-Matusita (J-M) distance statistic [41], which could quantify the separability between two classes effectively, was employed. The J-M distance between a pair of class specific probability functions is given by:

$$JM_{ij} = \left\{ \int_x \left[ \sqrt{p(x|\omega_i)} - \sqrt{p(x|\omega_j)} \right]^2 dx \right\}^{1/2} \tag{2}$$

In this study, $x$ denote the values of the selected C3 and C4 classification features, and $\omega_i$ and $\omega_j$ denote the C3 and C4 classes, respectively. Under normal conditions, Equation (2) reduces to:

$$JM_{ij} = 2 \left( 1 - e^{-B_{ij}} \right) \tag{3}$$

where

$$B_{ij} = \frac{1}{8} (\mu_i - \mu_j)^T \left( \frac{\Sigma_i + \Sigma_j}{2} \right)^{-1} (\mu_i - \mu_j) + \frac{1}{2} \ln \left( \frac{1}{2} \frac{|\Sigma_i + \Sigma_j|}{\sqrt{|\Sigma_i||\Sigma_j|}} \right) \tag{4}$$

In Equation (4), $\mu_i$ and $\mu_j$ correspond to the C3 and C4 NDVI mean values, and $\Sigma_i$ and $\Sigma_j$ are unbiased covariance matrices of C3 and C4. The J-M distance, which ranges between 0 and 2, provides a general measure of the separability between two classes based on the average distance between their class density functions [42].

The SVM classifier, which is available in the software ENVI 5.0 (ITT-Visual Information Solutions, USA) was employed in the classification process. SVM is based on the statistical machine learning theory and determines the location of decision boundaries that produce an optimal separation of classes. ENVI's implementation of SVM uses the pairwise classification strategy for multiclass classification. For the training of the SVM classifier, 2/3 of the field investigation points were randomly chosen, and the remaining 1/3 of the points were used as the "ground truth" to assess the classification accuracies.

## 4. Results and Discussion

### 4.1. NDVI Prediction Accuracy Assessment

4.1.1. NDVI Prediction Accuracy within the Growing Season

The predicted NDVI was calculated using the blended reflectance at red (R) and near infrared (NIR) bands. The surface reflectance blending accuracies are shown in Figures 7 and 8, which were similar to results of [20,21]. The blended red bands had higher correlation with the observed ETM+ data. Lower blending accuracy may be caused by farming activities at the end of the growing season. For the 134_32/33 sub-area, the blending accuracies of the DOY 215 and DOY 231 were lower because there

were clouded areas in these ETM+ scenes (although not very large). The $R^2$ of the blended and ETM+ reflectance were mostly higher than 0.73 for red band, and higher than 0.47 for near infrared band, respectively in the sub-area of 133_33. In the area of 134_32/33, the $R^2$ of the blended and ETM+ reflectance were mostly higher than 0.7 for red bands, and higher than 0.5 for the near infrared bands.

A comparison between the predicted NDVI and the observed ETM+ NDVI is provided in Figures 9 and 10. The accuracy assessment shows most of the predicted data are closer to the 1–1 line, ($R^2 > 0.74$). The higher $R^2$ were achieved during the middle of the growing season, whereas smaller $R^2$ appeared during the withering phase. The minimum $R^2$ were 0.53 and 0.74 at the end of the 133_33 and 143_32/33 sub-areas, respectively, which may be due to the seasonal farming activities in the crop area.

The NDVI derived from the blended reflectance data were more closely matched with data calculated from the ETM+ reflectance during the summer. The average $R^2$ between the predicted NDVI and the NDVI from ETM+ was larger than 0.88 for 133_33 and larger than 0.76 for 134_32/33 sub-area in summer.
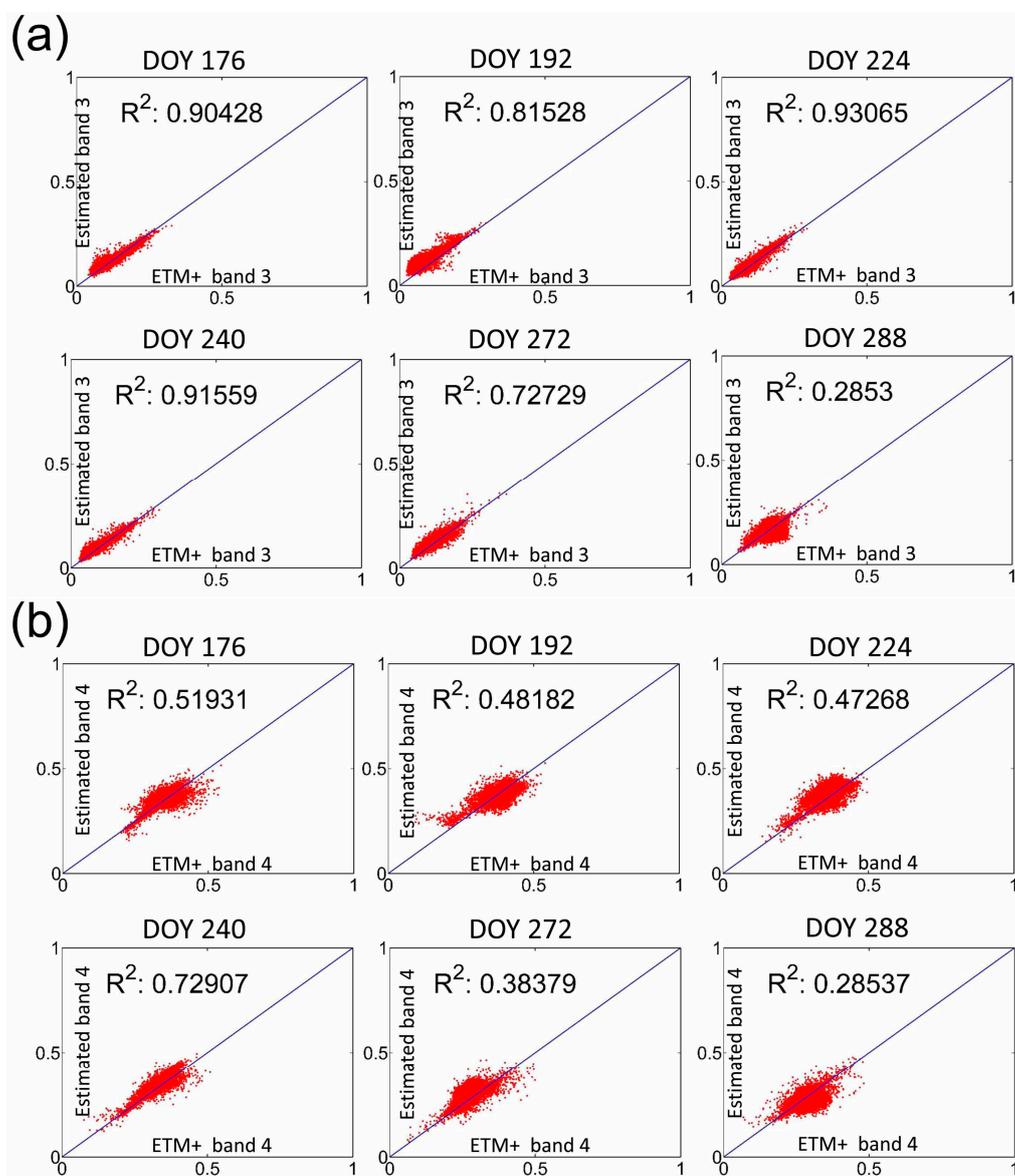


**Figure 7.** Scatterplots of the blended and the ETM+ data for the 133_33 sub-area. (**a**) scatterplots of the red band; (**b**) scatterplots of the near infrared band.
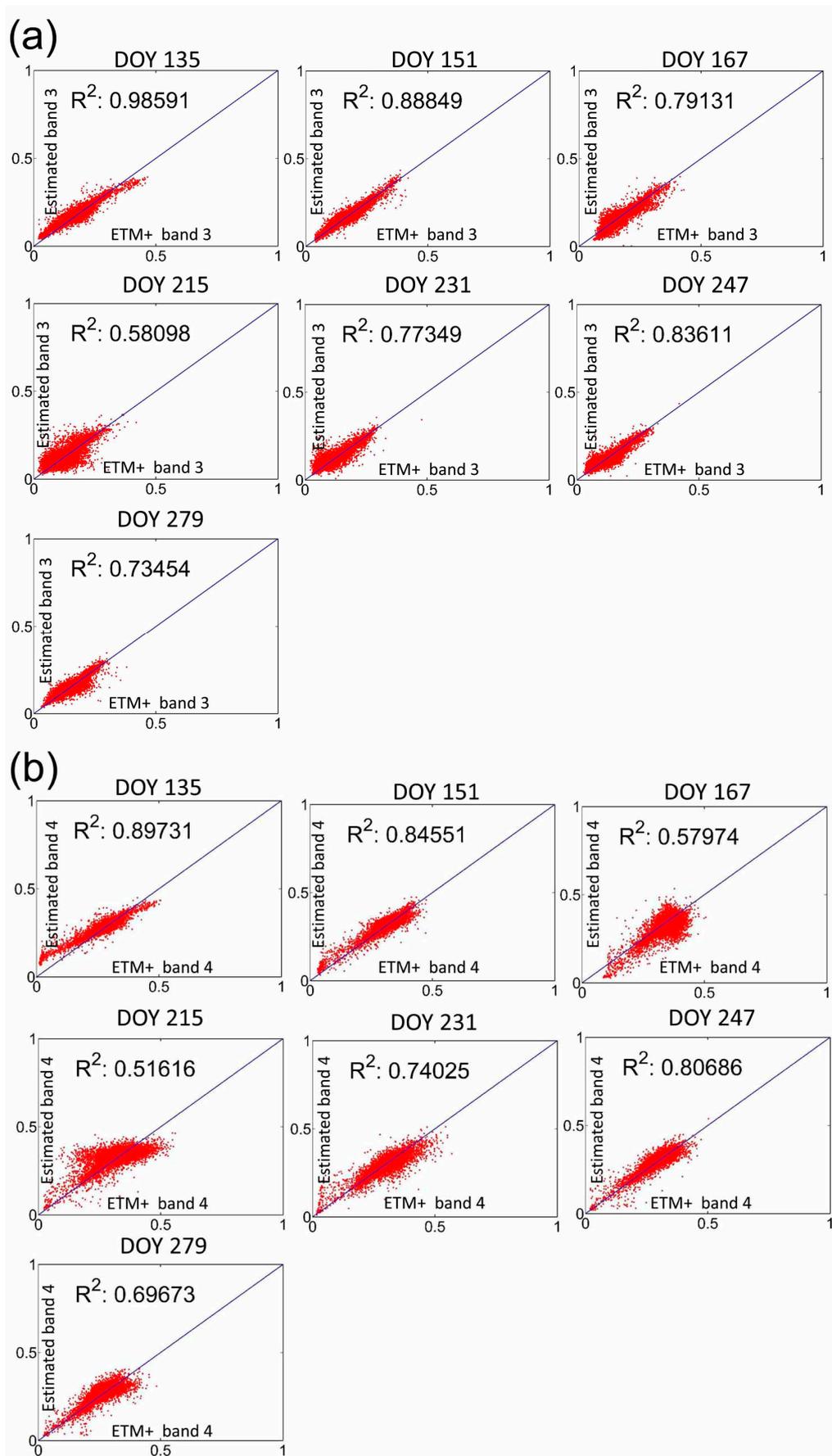
**Figure 8.** Scatterplots of the blended and the ETM+ data for the 134_32/33 sub-area. (**a**) scatterplots of the red band; (**b**) scatterplots of the near infrared band.
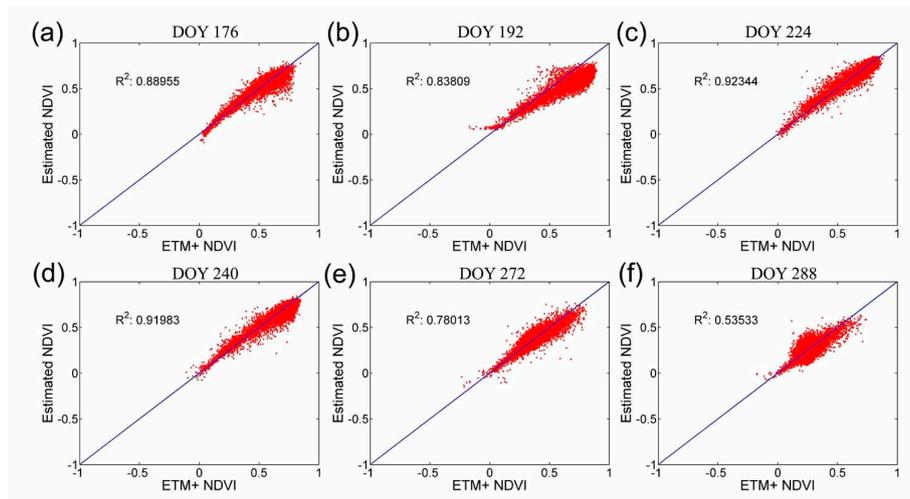
**Figure 9.** Scatterplots of the predicted NDVI and the actual ETM+ NDVI for the 133_33 sub-area. (Predicted NDVI by blending MODIS and ETM+ data of: (**a**) DOY 112, DOY 192; (**b**) DOY 176, DOY 224; (**c**) DOY 192, DOY 240; (**d**) DOY 224, DOY 272; (**e**) DOY 240, DOY 288; (**f**) DOY 272, DOY 304.
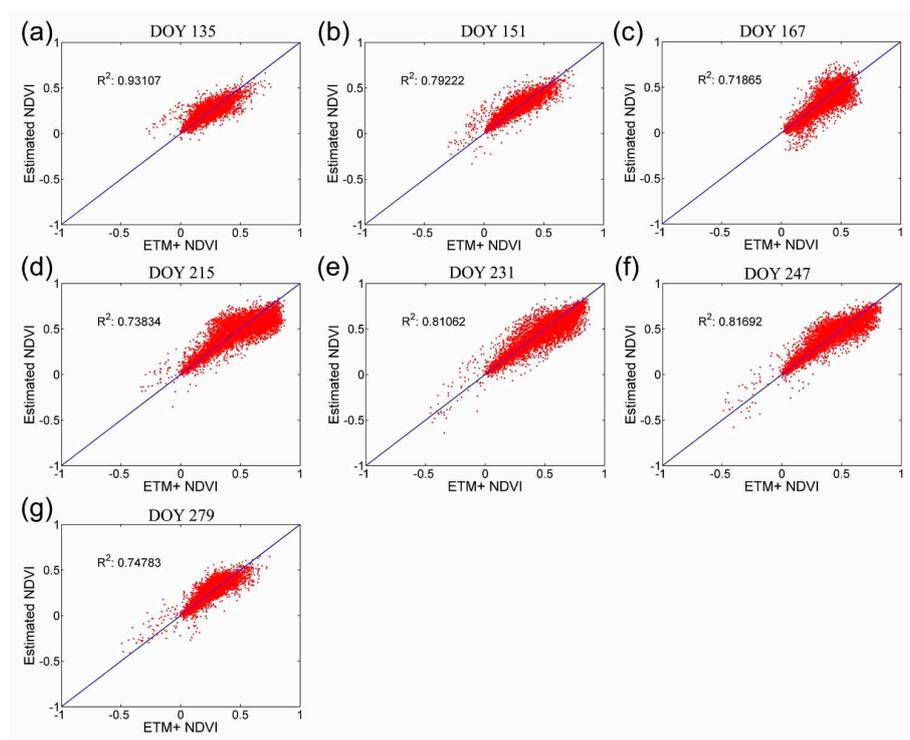


**Figure 10.** Scatterplots of the predicted NDVI and the actual ETM+ NDVI for the 134_32/33 sub-area. (Predicted NDVI by blending MODIS and ETM+ data of: (**a**) DOY 119, DOY 151; (**b**) DOY 135, DOY 167; (**c**) DOY 151, DOY 215; (**d**) DOY 167, DOY 231; (**e**) DOY 215, DOY 247; (**f**) DOY 231, DOY 279; (**g**) DOY 247, DOY 295.

The NDVI prediction accuracies for the ETM+ 134_32/33 sub-area are a bit lower than for the 133_33 sub-area. The ESTARFM algorithm was less accurate for the prediction of the near-infrared data, and the field investigation revealed that there was a larger area of desert and Gobi in the 143_32/33 sub-area.

Additionally, there were more frequent farming activities occurring in the 143_32/33 sub-area and more vegetable cultivation fields. Fortunately, the C4 crops in both sub-areas have similar farming seasonality, which means that the lower NDVI prediction accuracy over the 143_32/33 area will not significantly affect the C3 and C4 classification.

4.1.2. Time Interval Effect to the NDVI Blending Accuracy

In the ESTARFM data fusion algorithm, the time interval has a considerable effect on data fusion accuracy [21] because vegetation types or sun zenith angle changes will occur during the long time span. Thus, a long time interval for an input data pair will cause the NDVI prediction accuracy to decrease.

The results in Figure 11 show the NDVI blending accuracy at different time intervals of the input data. The results indicated the accuracies of blended NDVI will decrease as the time interval increases. An $R^2$ of 0.73 between the blended NDVI and the NDVI from ETM+ data was achieved at the time interval of 96 days over homogeneous farming land.
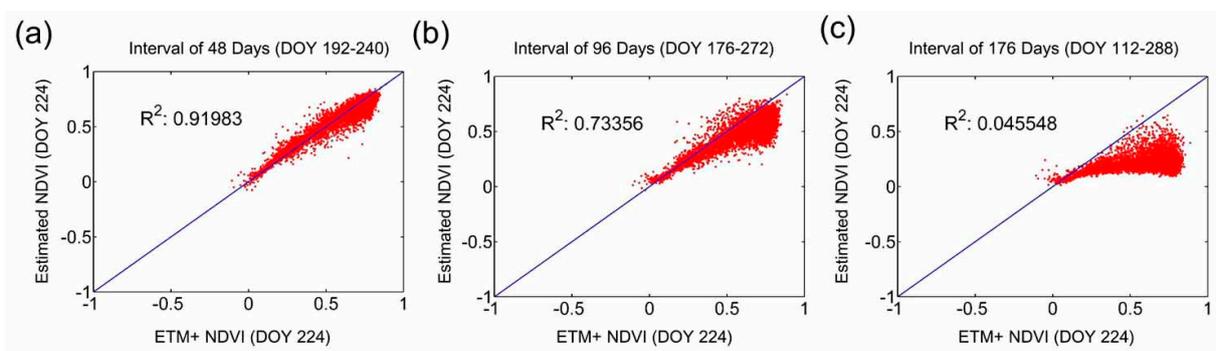


**Figure 11.** Input data time interval and the NDVI blending accuracy. Blending accuracy with a time interval of: (**a**) 48 days; (**b**) 96 days; and (**c**) 176 days.

However, the time intervals for the data fusion in Figures 9 and 10 are at least twice as long as our conducted data fusion. For example, for the NDVI prediction of DOY 176 in the 133_33 sub-area, the input data pairs include the MODIS and ETM+ data in DOY 112 and DOY 192. For the data between DOY 112 and DOY 176 to be predicted, we used the MODIS and ETM+ data in DOY 112 and DOY 176. The time intervals were 80 days (192–112) and 64 days (176–112) for the accuracy assessment and the actual data fusion, respectively. This indicated that the NDVI data prediction accuracy may be higher than the accuracy given above.

*4.2. Feature Selection for C3 and C4 Vegetation Classification*

Figure 12 are the boxplots of the 18 C3 and C4 classification features that were extracted from the time-series NDVI data. The boxplots were generated using the data derived from the 18 features according to the field investigation points, which illustrate the C3 and C4 data distribution of each feature, including the maximum/minimum values, the 75th percentile, 50th percentile (median), mean (in circle), and 25th percentile. From these statistical data, it is easy to determine the five distinguishable features of C3 and C4 plants: Max NDVI value, Min NDVI value, Integral D35–D45, Integral of NDVI, and Max_NDVI/Integral (Table 3). Figure 13 shows the spatial distribution of the selected 5 features listed in Table 3.
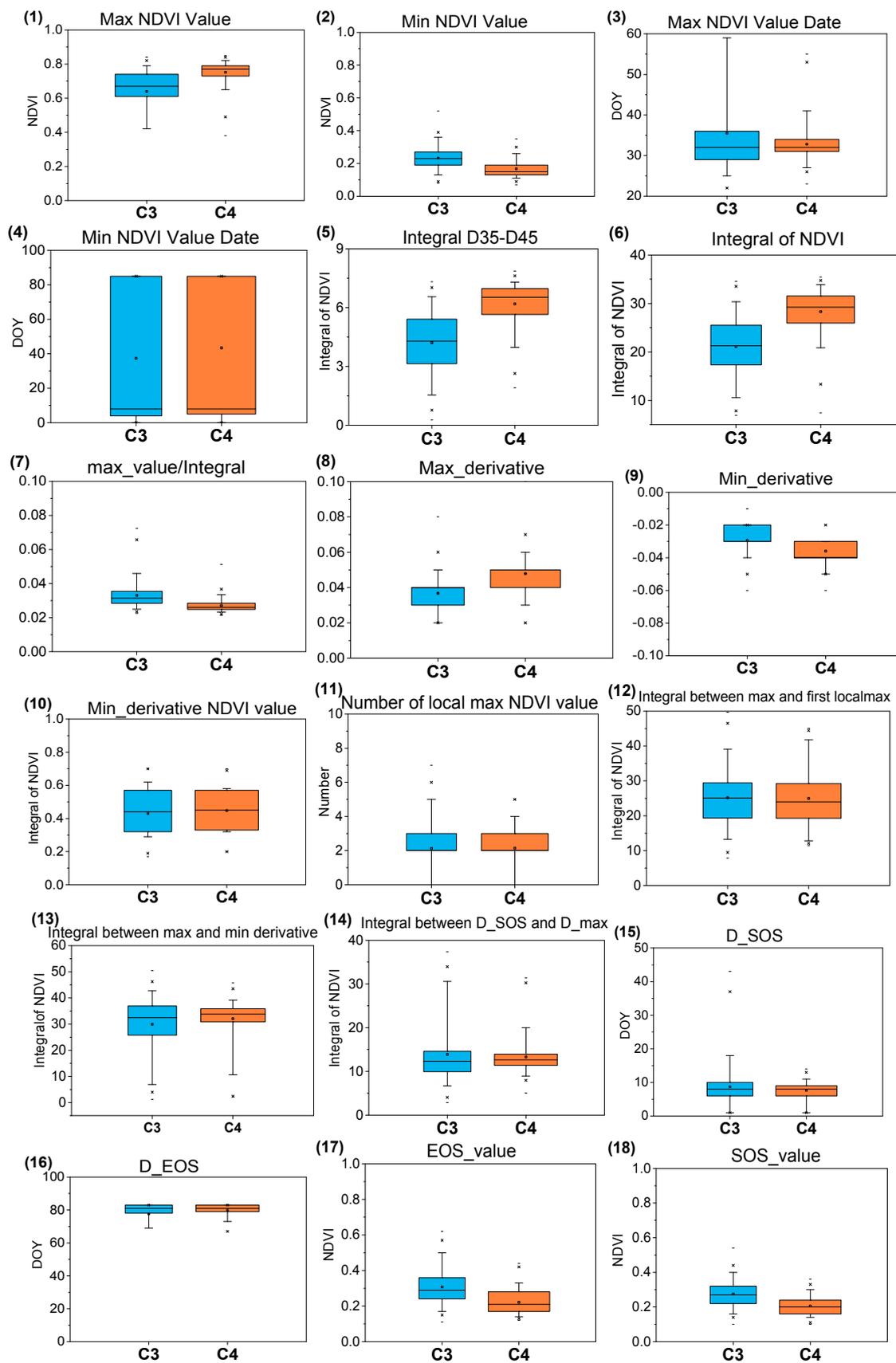
**Figure 12.** Boxplots of the 18 C3 and C4 classification features. (**1**–**18**) are boxplot of every classification feature listed in Table 2.
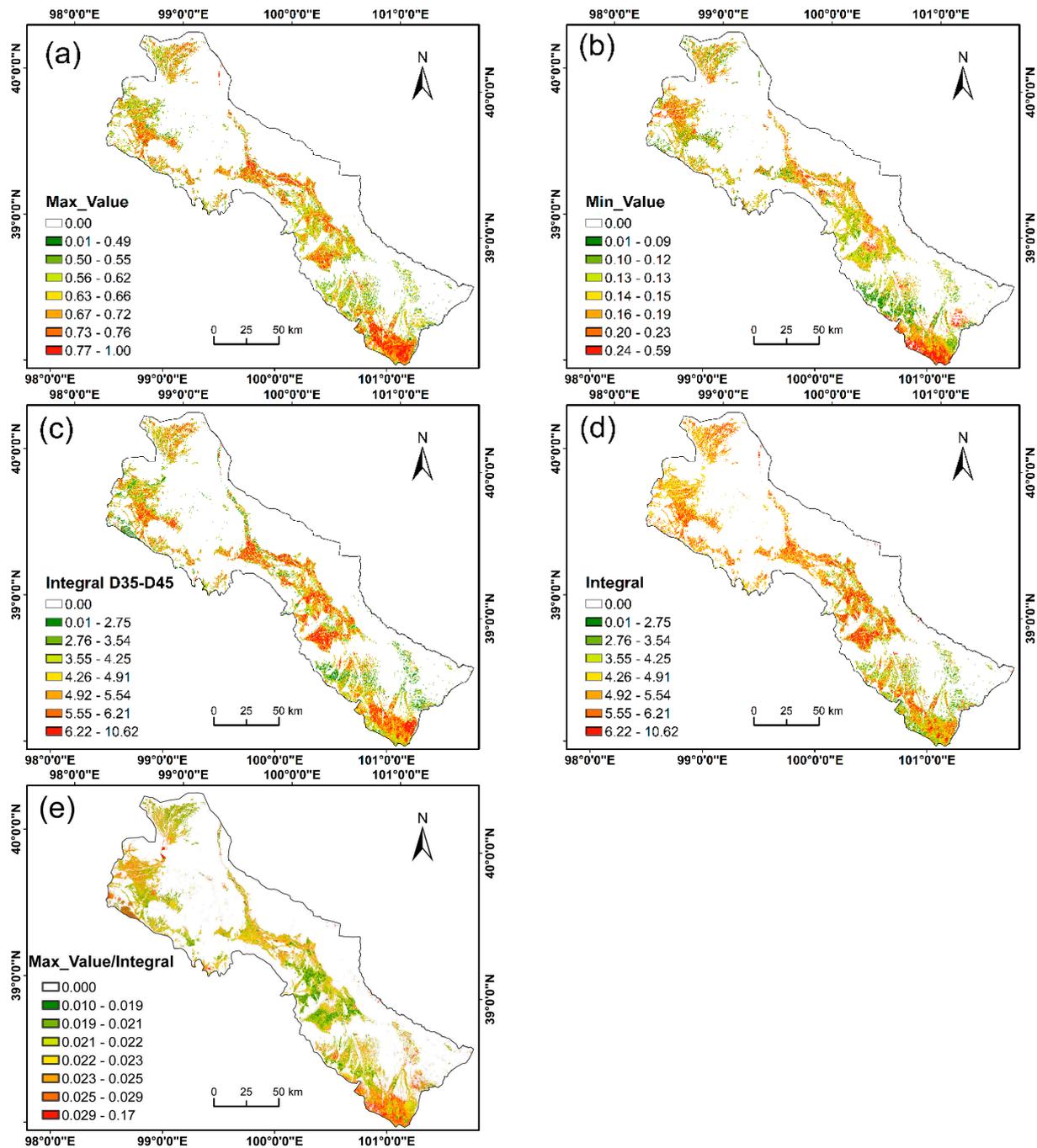
**Figure 13.** The maps of the selected features for the C3 and C4 vegetation classification. (**a**) the maximum NDVI value of the growing season; (**b**) the minimum NDVI value of the growing season; (**c**) the integral of NDVI between DOY 35 and DOY 45; (**d**) integral of NDVI for the growing season; and (**e**) the ratio between maximum NDVI and the integral of NDVI within the growing season.

To test the separability of the selected features for the C3 and C4 vegetation classification, the J–M distance was employed. The J-M distance between the C3 and C4 classes was based on each feature and combination of features to evaluate the overall separability of the selected features (Table 4). As shown in Table 4, larger J-M distances were found for more combination of more features. The combination between any two features of the five selected features selected was lower than three,

and the combination of four features was better than three, *etc.* The J-M distance was 1.93 when all the five features were all used to classify the C3 and C4 plants, which indicates that the C3 and C4 plants could most distinguishable by using the combination of the five selected features.

**Table 3.** The 5 features selected from the 18 features that were extracted from the time-series NDVI data.

| No. | Classification Features | Valve Range | Description |
|-----|------------------------|-------------|-------------|
| 1 | Max NDVI value | −1, 1 | Maximum NDVI value |
| 2 | Min NDVI value | −1, 1 | Minimum NDVI value |
| 3 | Integral D35–D45 | 0, 10 | Integral between DOY35 and DOY45 |
| 4 | Integral of NDVI | 0, 86 | Integral of NDVI time-series curve |
| 5 | Max_NDVI/Integral | $-\infty, +\infty$ | Ratio of maximum NDVI to NDVI curve integral |

**Table 4.** The J-M distances between different feature combinations (1: Max NDVI value, 2: Min NDVI value, 3: Integral D35-D45, 4: Integral of NDVI, 5: Max_NDVI/Integral).

| Feature Combination | J-M Distance | Feature Combination | J-M Distance | Feature Combination | J-M Distance | Feature Combination | J-M Distance |
|---------------------|--------------|---------------------|--------------|---------------------|--------------|---------------------|--------------|
| 1,2 | 1.55 | 1,2,3 | 1.82 | 1,2,3,4 | 1.78 | 1,2,3,4,5 | 1.93 |
| 1,3 | 1.6 | 1,2,4 | 1.77 | 1,2,3,5 | 1.89 | -- | -- |
| 1,4 | 1.6 | 1,2,5 | 1.75 | 1,3,4,5 | 1.87 | -- | -- |
| 1,5 | 1.6 | 1,3,4 | 1.74 | 2,3,4,5 | 1.84 | -- | -- |
| 2,3 | 1.6 | 1,3,5 | 1.81 | -- | -- | -- | -- |
| 2,4 | 1.6 | 1,4,5 | 1.72 | -- | -- | -- | -- |
| 2,5 | 1.6 | 2,3,4 | 1.74 | -- | -- | -- | -- |
| 3,4 | 1.29 | 2,3,5 | 1.76 | -- | -- | -- | -- |
| 3,5 | 1.6 | 3,4,5 | 1.74 | -- | -- | -- | -- |
| 4,5 | 1.6 | -- | -- | -- | -- | -- | -- |

The two-sample Kolmogorov-Smirnov test was employed to test whether the five selected features for C3 and C4 vegetation classification are significantly different. The two-sample Kolmogorov-Smirnov test is one of the most widely used nonparametric statistical test methods for comparing two independent samples with no assumption made concerning the distribution of the variables, and is sensitive to differences in both location and shape of the empirical cumulative distribution functions of the two samples [43]. The two-sample Kolmogorov–Smirnov statistic is given by:

$$D_{n,n'} = \sup_{x} \left| F_{1,n}(x) - F_{2,n'}(x) \right| \tag{5}$$

where $F_{1,n}$ and $F_{2,n'}$ are the empirical distribution functions of the first and the second sample respectively, $n$ and $n'$ are the numbers of samples for the first and the second sample, and *sup* is the supremum function. The null hypothesis is rejected at level α if:

$$D_{n,n'} > D_{crit} \tag{6}$$

where

$$D_{crit} = c(\alpha)\sqrt{\frac{n+n'}{nn'}} \qquad (7)$$

The value of $c(\alpha)$ in Equation (7) for each level of α is given in [44]. The $c(\alpha)$ is 1.95 at the α = 0.001 level. We used the MATLAB software to conduct the two-sample Kolmogorov-Smirnov test of every chosen classification feature, where the *p*-value was used to represent the probability of accepting the $H_0$ hypothesis at level α There are 682 C3 samples and 499 C4 samples, respectively. The null hypothesis at the α = 0.001 level is that:

$H_0$: samples derived from the C3 and C4 classification feature are not significantly different;
$H_1$: samples derived from the C3 and C4 classification feature are significantly different.

The Kolmogorov-Smirnov test results for the five selected C3 and C4 vegetation classification features are shown in Table 5. For each feature, the $D_{n,n'}$ is much larger than $D_{crit}$ (α = 0.001), which means that each of the K-S test result rejects the $H_0$ hypothesis at the 99.9% confidence level. These results indicate that the distributions of the selected features for C3 and C4 vegetation classification are significantly different at the 99.9% confidence level.

**Table 5.** Kolmogorov-Smirnov tests for C3 and C4 classification features.

| Features | $D_{n,n'}$ | $D_{crit}$ (α = 0.001) | *p*-Value |
|---|---|---|---|
| Integral 35–45 | 0.6433 | 0.1149 | $1.00 \times 10^{-7}$ |
| Integral | 0.6104 | 0.1149 | $9.97 \times 10^{-64}$ |
| Ratio | 0.3660 | 0.1149 | $3.50 \times 10^{-23}$ |
| Max | 0.6177 | 0.1149 | $3.07 \times 10^{-65}$ |
| Min | 0.3239 | 0.1149 | $3.03 \times 10^{-18}$ |

*4.3. Accuracy Assessment of C3 and C4 Vegetation Classification*

To compare the C3 and C4 plant classification accuracy at different spatial resolutions, we classified the C3 and C4 in both the time-series MODIS data at 250m resolution and the predicted data at 30 m resolution with the same selected five features. The classification results are shown in Figure 14. Figure 14a is the classification result based on the time-series MODIS NDVI data (250 m resolution), and Figure 14b is the classification result using the blended time-series NDVI data (30 m resolution). The classification accuracy assessments were conducted using the ground investigation data (205 points for C3 and 224 points for C4).

The classification accuracy measurements of the time-series MODIS NDVI and blended time-series NDVI data are shown in Tables 6 and 7, respectively. The classification accuracy of the blended high spatial resolution data is noticeably higher than that of the MODIS data. The overall accuracy and kappa coefficient of the former are 85.75% and 0.7235 respectively, whereas the overall classification accuracy of the time-series MODIS NDVI data is 69.65% with a kappa coefficient of 0.4. According to the MODIS data classification confusion matrix, many C3 plants were classified as C4. The possible reason is that the C3 vegetation having a more fragmented distribution than the C4 vegetation, and MODIS data at coarser resolution contained more mixed-pixels that to be easily misclassified. As shown in Figure 14, there is much difference between the classification results from MODIS and the blended data.

Because the major difference between the MODIS and the blended data is that they have different spatial resolutions, the differences between the classifications based on MODIS and the blended data are largely due to the spatial heterogeneity of the vegetation type distribution. The ground investigation had revealed there was a large number of small parcels of vegetable and maize with a fragmented distribution around Jiuquan city, northwest of the study area.
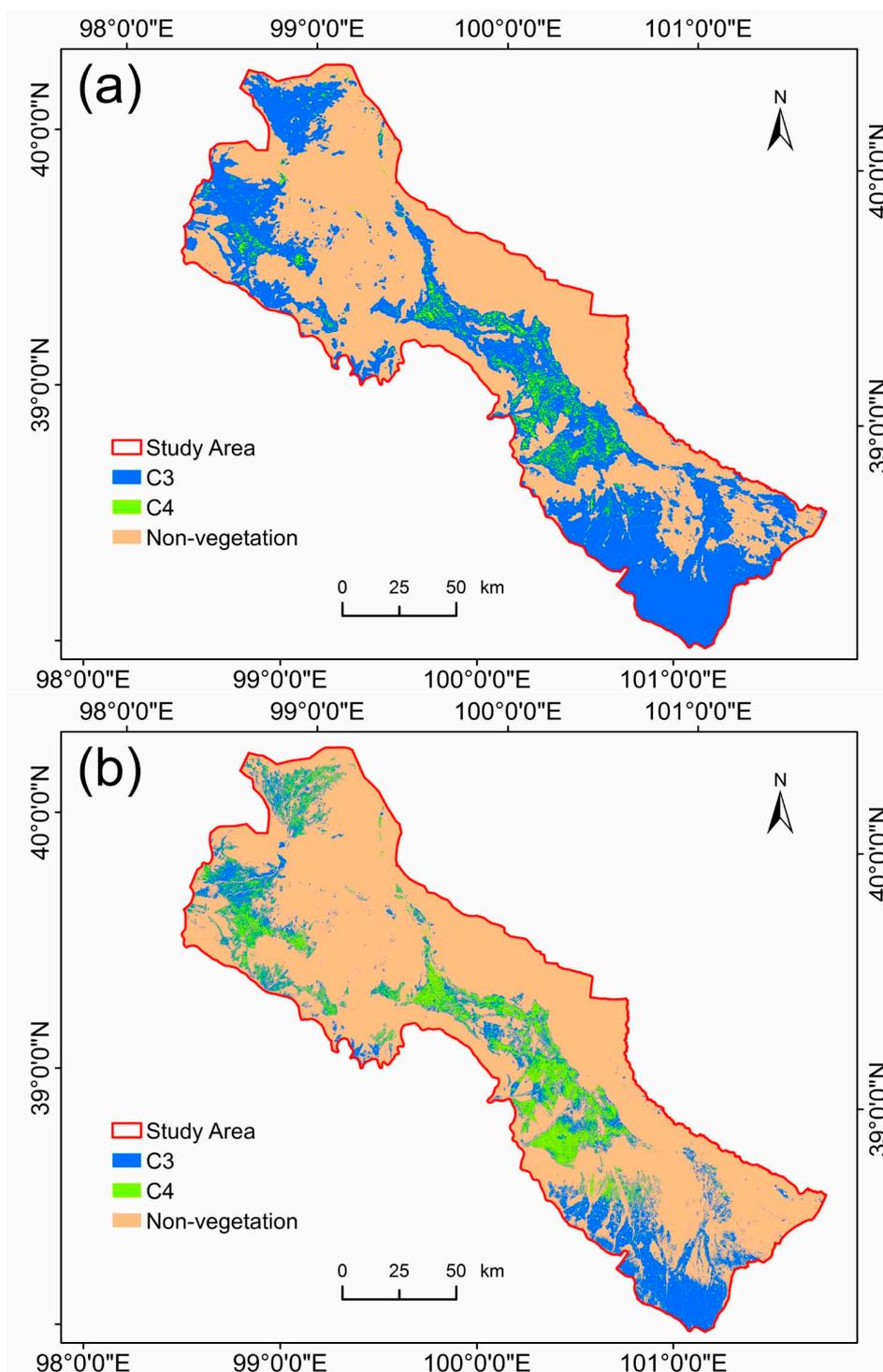


**Figure 14.** The C3 and C4 vegetation classification results over the study area. (**a**) the C3 and C4 vegetation classification result of MODIS data; (**b**) the C3 and C4 vegetation classification result of MODIS and ETM+ blended data.

**Table 6.** C3 and C4 classification accuracy of moderate resolution imaging spectroradiometer (MODIS) time-series data.

| -- | | **Ground Truth** | | | |
|---|---|---|---|---|---|
| | Class | C3 | C4 | Total | User's Accuracy |
| | C3 | 183 | 110 | 293 | 62.46% |
| | C4 | 19 | 113 | 132 | 85.61% |
| Classification Results | Total | 202 | 223 | 425 | -- |
| | Producer's Accuracy | 90.59% | 50.67% | -- | -- |
| | Overall Accuracy: 69.65% | | | | |
| | Kappa: 0.4041 | | | | |

**Table 7.** C3 and C4 classification accuracy of blended time-series data.

| -- | | **Ground Truth** | | | |
|---|---|---|---|---|---|
| | Class | C3 | C4 | Total | User's Accuracy |
| | C3 | 179 | 31 | 210 | 85.24% |
| | C4 | 17 | 188 | 205 | 91.71% |
| Classification Results | Total | 205 | 223 | 428 | -- |
| | Producer's Accuracy | 87.32% | 84.3% | -- | -- |
| | Overall Accuracy: 85.75% | | | | |
| | Kappa: 0.7235 | | | | |

## 5. Conclusions

We presented a framework for high spatial resolution C3 and C4 vegetation classification in regions with fragmented landscapes. In this framework, daily land surface reflectance data at 30 m spatial resolution was generated by fusing MODIS and Landsat ETM+ data using the ESFARFM algorithm. Based on the time-series NDVI data generated from the fused land surface reflectance data, features for C3 and C4 vegetation classification were extracted and selected. C3 and C4 vegetation was classified using the selected features and the nonparametric machine learning classifier, SVM.

The C3 and C4 classification framework was tested in the middle reaches of Heihe Watershed that locates in Gausu Province, China. The result indicated that the average $R^2$ between the predicted NDVI and the ETM+ derived NDVI was more than 0.88. The combination of 5 selected classification features (minimum/maximum NDVI value, integral of time-series NDVI, the ratio between maximum NDVI value, and the integral of time-series NDVI) could better capture the differences between C3 and C4 vegetation. Compared to the C3 and C4 vegetation classification using the time-series MODIS data with 250m spatial resolution, the fused time-series data with 30m spatial resolution achieved a higher C3 and C4 vegetation classification accuracy (16% higher than those of MODIS C3 and C4 classification accuracy). The fused time-series NDVI data could map C3 and C4 vegetation distribution better over regions with fragmented landscapes.

Compared to the previous study of C3 and C4 grasses classification in the U.S. Great Plains [5,19], the classification results in this study by using blended finer resolution time-series remote sensing data shows more spatial details of C3 and C4 vegetation distribution. This is a critical advantage for C3 and C4 vegetation mapping in regions with spatially heterogeneous landscape.

Long time interval between Landsat TM/ETM+ data may introduce large uncertainties in the blended data. To achieve accurate classification of C3 and C4 using the methodology presented in this paper, one should collect time series Landsat TM/ETM+ data with as short time interval as possible.

C3 and C4 vegetation within the same climate zone show markedly different seasonal activity cycles. Thus, we suggest our method to be used in the same climate zone. The methodology presented in this paper also has the potential to map land cover types with a high spatial resolution time-series remote sensing data.

## Acknowledgments

## Author Contributions

Yanchen Bo conceived and designed the study, and finalized the manuscript; Xiaolong Liu analyzed all the data, conducted the experiments and wrote the first draft of manuscript; Jian Zhang and Yaqian He collected and pre-processed the original data and revised the manuscript.

## Conflicts of Interest

The authors declare no conflict of interest.

## References

1    Wittmer, M.H.; Auerswald, K.; Bai, Y.; Schaeufele, R.; Schnyder, H. Changes in the abundance of C3/C4 species of Inner Mongolia grassland: Evidence from isotopic composition of soil and vegetation. *Glob. Chang. Biol.* **2010**, *16*, 605–616.

2    Smith, T.T.M.; Shugart, H.H.; Woodward, F.I. *Plant Functional Types: Their Relevance to Ecosystem Properties and Global Change*; Cambridge University Press: Cambridge, UK, 1997; Volume 1.

3    Davidson, A. A comparison of three approaches for predicting C4 species cover of northern mixed grass prairie. *Remote Sens. Environ.* **2003**, *86*, 70–82.

4    Ustin, S.L.; Gamon, J.A. Remote sensing of plant functional types. *New Phytol.* **2010**, *186*, 795–816.

5    Wang, C.; Hunt, E.R., Jr.; Zhang, L.; Guo, H. Phenology-assisted classification of C3 and C4 grasses in the U.S. Great Plains and their climate dependency with MODIS time series. *Remote Sens. Environ.* **2013**, *138*, 90–101.

6    Collatz, G.J.; Berry, J.A.; Clark, J.S. Effects of climate and atmospheric $CO_2$ partial pressure on the global distribution of C4 grasses: Present, past, and future. *Oecologia* **1998**, *114*, 441–454.

7    Morgan, J.A.; Milchunas, D.G.; LeCain, D.R.; West, M.; Mosier, A.R. Carbon dioxide enrichment alters plant community structure and accelerates shrub growth in the shortgrass steppe. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 14724–14729.

8    Irisarri, J.G.N.; Oesterheld, M.; Verón, S.R.; Paruelo, J.M. Grass species differentiation through canopy hyperspectral reflectance. *Int. J. Remote Sens*. **2009**, *30*, 5959–5975.

9    Adjorlolo, C.; Mutanga, O.; Cho, M.A.; Ismail, R. Spectral resampling based on user-defined inter-band correlation filter: C3 and C4 grass species classification. *Int. J. Appl. Earth Obs. Geoinf.* **2013**, *21*, 535–544.

10   Amoros-Lopez, J.; Gomez-Chova, L.; Vila-Frances, J.; Alonso, L.; Calpe, J.; Moreno, J.; del Valle-Tascon, S. Evaluation of remote sensing of vegetation fluorescence by the analysis of diurnal cycles. *Int. J. Remote Sens.* **2008**, *29*, 5423–5436.

11   Zarco-Tejada, P.J.; Miller, J.R.; Mohammed, G.H.; Noland, T.L.; Sampson, P.H. Estimation of chlorophyll fluorescence under natural illumination from hyperspectral data. *Int. J. Appl. Earth Obs. Geoinf.* **2001**, *3*, 321–327.

12   Liu, L.; Cheng, Z. Mapping C3 and C4 plant functional types using separated solar-induced chlorophyll fluorescence from hyperspectral data. *Int. J. Remote Sens.* **2011**, *32*, 9171–9183.

13   Joiner, J.; Yoshida, Y.; Vasilkov, A.P.; Yoshida, Y.; Corp, L.A.; Middleton, E.M. First observations of global and seasonal terrestrial chlorophyll fluorescence from space. *Biogeosciences* **2011**, *8*, 637–651.

14   Risser, P.G. *The True Prairie Ecosystem*; Academic Press: Waltham, MA, USA, 1981.

15   Nobel, P.S. *Physicochemical and Environmental Plant Physiology*; Academic Press: Waltham, MA, USA, 1999.

16   Ode, D.J.; Tieszen, L.L.; Lerman, J.C. The seasonal contribution of C3 and C4 plant species to primary production in a mixed prairie. *Ecology* **1980**, *61*, 1304–1311.

17   Goodin, D.G.; Henebryf, G.M. A technique for monitoring ecological disturbance in tallgrass prairie using seasonal NDVI trajectories and a discriminant function mixture model. *Remote Sens. Environ.* **1997**, *62*, 270–278.

18   Ricotta, C.; Reed, B.C.; Tieszen, L.T. The role of C3 and C4 grasses to interannual variability in remotely sensed ecosystem performance over the U.S. Great Plains. *Int. J. Remote Sens.* **2003**, *24*, 4421–4431.

19   Foody, G.M.; Dash, J. Discriminating and mapping the C3 and C4 composition of grasslands in the Northern Great Plains, USA. *Ecol. Inform.* **2007**, *2*, 89–93.

20   Gao, F.; Masek, J.; Schwaller, M.; Hall, F. On the blending of the Landsat and MODIS surface reflectance: Predicting daily Landsat surface reflectance. *IEEE Trans. Geosci. Remote Sens.* **2006**, *44*, 2207–2218.

21   Zhu, X.; Chen, J.; Gao, F.; Chen, X.; Masek, J.G. An enhanced spatial and temporal adaptive reflectance fusion model for complex heterogeneous regions. *Remote Sens. Environ.* **2010**, *114*, 2610–2623.

22   Masek, J.G.; Vermote, E.F.; Saleous, N.E.; Wolfe, R.; Hall, F.G.; Huemmrich, K.F.; Gao, F.; Kutler, J.; Lim, T.K. A Landsat surface reflectance dataset for North America, 1990–2000. *IEEE Geosci. Remote Sens. Lett.* **2006**, *3*, 68–72.

23  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.

24  Arvldson, T.; Goward, S.; Gasch, J.; Williams, D. Landsat-7 long-term acquisition plan: Development and validation. *Photogramm. Eng. Remote Sens.* **2006**, *72*, 1137–1146.

25  Bédard, F.; Reichert, G.; Dobbins, R.; Trépanier, I. Evaluation of segment-based gap-filled Landsat ETM+ SLC-off satellite data for land cover classification in southern Saskatchewan, Canada. *Int. J. Remote Sens.* **2008**, *29*, 2041–2054.

26  Chen, J.; Zhu, X.; Vogelmann, J.E.; Gao, F.; Jin, S. A simple and effective method for filling gaps in Landsat ETM+ SLC-off images. *Remote Sens. Environ.* **2011**, *115*, 1053–1064.

27  Roy, D.P.; Ju, J.; Lewis, P.; Schaaf, C.; Gao, F.; Hansen, M.; Lindquist, E. Multi-temporal MODIS–Landsat data fusion for relative radiometric normalization, gap filling, and prediction of Landsat data. *Remote Sens. Environ.* **2008**, *112*, 3112–3130.

28  Zhu, X.; Liu, D.; Chen, J. A new geostatistical approach for filling gaps in Landsat ETM+ SLC-off images. *Remote Sens. Environ.* **2012**, *124*, 49–60.

29  Pan, Z.; Huang, J.; Zhou, Q.; Wang, L.; Cheng, Y.; Zhang, H.; Blackburn, G.A.; Yan, J.; Liu, J. Mapping crop phenology using NDVI time-series derived from HJ-1 A/B data. *Int. J. Appl. Earth Obs. Geoinform.* **2015**, *34*, 188–197.

30  Chen, J.; Jönsson, P.; Tamura, M.; Gu, Z.; Matsushita, B.; Eklundh, L. A simple method for reconstructing a high-quality NDVI time-series data set based on the Savitzky-Golay filter. *Remote Sens. Environ.* **2004**, *91*, 332–344.

31  Viovy, N.; Arino, O.; Belward, A. The Best Index Slope Extraction (BISE): A method for reducing noise in NDVI time-series. *Int. J. Remote Sens.* **1992**, *13*, 1585–1590.

32  Brown, M.E.; de Beurs, K.M. Evaluation of multi-sensor semi-arid crop season parameters based on NDVI and rainfall. *Remote Sens. Environ.* **2008**, *112*, 2261–2271.

33  Linderholm, H.W. Growing season changes in the last century. *Agric. For. Meteorol.* **2006**, *137*, 1–14.

34  Sakamoto, T.; Wardlow, B.D.; Gitelson, A.A.; Verma, S.B.; Suyker, A.E.; Arkebauer, T.J. A two-step filtering approach for detecting maize and soybean phenology with time-series MODIS data. *Remote Sens. Environ.* **2010**, *114*, 2146–2159.

35  Hird, J.N.; McDermid, G.J. Noise reduction of NDVI time series: An empirical comparison of selected techniques. *Remote Sens. Environ.* **2009**, *113*, 248–258.

36  Jönsson, P.; Eklundh, L. Timesat—A program for analyzing time-series of satellite sensor data. *Comput. Geosci.* **2004**, *30*, 833–845.

37  Jonsson, P.; Eklundh, L. Seasonality extraction by function fitting to time-series of satellite sensor data. *IEEE Trans. Geosci. Remote Sens.* **2002**, *40*, 1824–1832.

38  Jakubauskas, M.E.; Legates, D.R.; Kastens, J.H. Crop identification using harmonic analysis of time-series AVHRR NDVI data. *Comput. Electron. Agric.* **2002**, *37*, 127–139.

39  Sakamoto, T.; Yokozawa, M.; Toritani, H.; Shibayama, M.; Ishitsuka, N.; Ohno, H. A crop phenology detection method using time-series MODIS data. *Remote Sens. Environ.* **2005**, *96*, 366–374.

40    Cong, N.; Piao, S.; Chen, A.; Wang, X.; Lin, X.; Chen, S.; Han, S.; Zhou, G.; Zhang, X. Spring vegetation green-up date in China inferred from SPOT NDVI data: A multiple model analysis. *Agric. For. Meteorol.* **2012**, *165*, 104–113.

41    Richards, J.A.; Richards, J. *Remote Sensing Digital Image Analysis*; Springer: Berlin, Gemany, 1999; Volume 3.

42    Wardlow, B.; Egbert, S.; Kastens, J. Analysis of time-series MODIS 250 m vegetation index data for crop classification in the U.S. Central Great Plains. *Remote Sens. Environ.* **2007**, *108*, 290–310.

43    Massey, F.J., Jr. The kolmogorov-smirnov test for goodness of fit. *J. Am. Stat. Assoc.* **1951**, *46*, 68–78.

44    Rohlf, F.J.; Sokal, R.R. *Statistical Tables*; Macmillan: London, UK, 1995.