A Novel Efficient Method for Land Cover Classification in Fragmented Agricultural Landscapes Using Sentinel Satellite Imagery

Xinyi Li 1, Chen Sun 2, Huimin Meng 1, Xin Ma 3, Guanhua Huang 1 and Xu Xu 1,*

1 Chinese-Israeli International Center for Research and Training in Agriculture, College of Water Resources & Civil Engineering, China Agricultural University, Beijing 100083, China; xinyi_li@cau.edu.cn (X.L.); s20193091624@cau.edu.cn (H.M.); ghuang@cau.edu.cn (G.H.)
2 Institute of Environment and Sustainable Development in Agriculture, Chinese Academy of Agricultural Sciences, Beijing 100081, China; sunchen@caas.cn
3 Water Resources Research Institute of Inner Mongolia, Hohhot 010051, China; nmgmaxin@126.com
* Correspondence: xushengwu@cau.edu.cn

Abstract: Updated and accurate land cover maps are essential and crucial for sustainable crop production and efficient land management. However, accurate and efficient land cover mapping is still a challenge for agricultural regions with complicated landscapes. This study proposed a novel spectral-phenological based land cover classification (SPLC) method to identify the land cover for fragmented agricultural landscapes, with less requirement of ground truth data. The SPLC method integrated a pixel-based support vector machine (SVM) algorithm for cropland and various non-cropland classification, and a phenology-based automatic decision tree algorithm for identification of various crop types. It was then tested and applied in two typical case areas (i.e., Jiyuan in the upstream and Yonglian in the downstream) of Hetao Irrigation District (Hetao) in the upper Yellow River basin (YRB), northwest China. The field survey sampling data and the regional visual interpretation maps were jointly used to evaluate the accuracy of land cover classification. Results indicated that stable phenological rules can be established for crop identification even with complex planting patterns, and the SPLC method performed well in land cover mapping in case areas. Four high-accuracy land cover maps were produced for Jiyuan in 2020 and 2021, Yonglian in 2021, and Hetao in 2021. The overall accuracies (OA) can reach 0.90–0.94 based on evaluation with abundant ground truth data, and land cover maps agreed well with the visual interpretation maps in space. Overall, the case application validated the applicability and efficiency of the SPLC method in land cover mapping for regions with fragmented agricultural landscapes, and also implied the potential use in other similar regions.

Keywords: land cover classification; crop mapping; phenology; fragmented agricultural landscapes; Sentinel-2 satellite

1. Introduction

Most of the urgent challenges humanity is facing today are directly or indirectly related to agricultural production [1]. With global population growth and dietary changes, agricultural production needs to be increased and improved to ensure global food security. The increased food production may undermine many ecosystem services on a global scale [2]. Therefore, to avoid the loss of ecosystem services, it is an essential prerequisite to keep the balance of all services provided by the ecosystem, which demands reliable and accurate agricultural monitoring to investigate land management practices and ensure sustainable crop production [3]. For this purpose, detailed spatial information about the land cover in agricultural areas is needed for decision-makers, scientists, and local managers [4]. The remote sensing technology has been widely and successfully applied in producing land
cover maps with reasonable precision and high efficiency [5–9]. However, since agricultural land is strongly affected by the spatial and temporal dynamics within and between each growing season, crop mapping is still a very challenging task, especially for regions with fragmented agricultural landscapes [10]. It is noteworthy that the recent emergence of remote sensing satellites with high spatial and temporal resolutions may have the potential to capture the agricultural land use dynamics and improve crop mapping accuracy.

The supervised classification method is the predominant approach in land cover identification, with remarkable performance in remote sensing image classification. Traditional supervised classification is based on information on spectrum, texture, and typical physiological features (e.g., vegetation index) [11–13]. Each pixel is distinguished and classified as a specified type of land cover. This method requires the training data from sampling, and its accuracy is strongly dependent on sampling data richness. Moreover, due to image- or time-specific limitations, this type of method is more suitable for classifying non-cropland land coverages with stable surface characteristics [14]. Whereas croplands usually vary strongly within short time intervals and crop planting pattern varies as well in different years. Thus, reference data collection and training datasets must be re-prepared when applied in different periods, which is costly and inefficient. Therefore, developing an automated algorithm, which can provide stable classification rules without re-training and re-validation for different seasons/years, is very attractive for land cover classification, particularly in long-term applications.

The phenological changes occurring in the growing period are crop-type specific, and the life cycles of beginnings and endings differ depending on the crop types [10]. The use of multi-temporal image-based phenology has played an important role in crop classification [15–17]. Some researchers extract phenological information by analyzing multi-temporal profiles, and then study the seasonal dynamic differences of different crop types to select the optimum imaging date for classification [18–21]. Other studies explicitly derive quantitative phenological metrics and build classification rules based on crop calendars and crop conditions at each phase of the growing period [22–24]. Phenological metrics have the advantage of strong physiological meaning. The rules based on phenological metrics are thus more robust than those using statistical thresholds, and the latter is more dependent on specific datasets with more uncertainties, along with also being harder to generalize [25]. However, it is still a challenge to establish the applicable/general rules for fragmented agricultural landscapes with diverse crop types. The use of satellite remote sensing time series-based phenological information may be capable of accelerating the development of rules on a large regional scale and improving the stability of rules over time [14].

In recent years, different satellite data sources, such as the Moderate Resolution Imaging Spectroradiometer (MODIS) [22,26], Landsat TM/ETM+/OLI [27,28], and SPOT [18], have been widely used in land cover classification. However, satellite images with the coarse spatial resolution are usually not preferable for crop identification in heterogeneous agricultural fields, which may lead to mixed pixel problems (e.g., MODIS) [22]. On the other hand, the lower temporal resolution may result in lower accuracy due to the difficulties in capturing the seasonal dynamics of cropland and extracting phenological metrics (e.g., with Landsat or SPOT) [29]. Since 2015, the Sentinel series satellites have been launched in succession with relatively high resolutions in both space and time. The regular revisiting intervals and relatively high spatial resolution (10 m) offer a unique opportunity to systematically monitor crops at an approximately weekly repeat cycle [30]. Previous studies have indicated that the vegetation index (VI) series derived from optical Sentinel-2A/2B could present a complete temporal profile of different land cover types, and the retrieved phenological metrics are comparable with local agro-metrological observations. Some studies have adopted the metrics of Sentinel-2A/2B data to identify the land cover types [1,13,24]. However, to our knowledge, the successful use in heterogeneous and fragmented agricultural landscape mapping is still rarely reported.
Taking into account the considerations above, the assumption is made that: (1) using sufficiently high spatial-temporal resolution remotely sensed data from satellites might be able to identify certain distinct phenological metrics for different crops, even with complex and diverse planting patterns; and (2) the stable and long-lasting classification rules may be able to be established based on phenological metrics, even without field-survey data. Therefore, the objectives of this study were: (1) to introduce a novel spectral-phenological based land cover classification (SPLC) method to realize the efficient regional-scale land cover identification, particularly for complicated fragmented agricultural landscapes; and (2) to further conduct a case test with Sentinel-2 time series data in a typical agricultural region in the upper Yellow River basin (YRB) of northwest China. A pixel-based algorithm is firstly developed to identify cropland and non-cropland land cover types based on optimum spectrum, texture, and Land Surface Water Index (LSWI) threshold time series dataset. Then, an automatic decision tree phenology-based method is developed to realize the classification of various crop types. The proposed SPLC method that does not need reference (field surveys) data could fully implement the knowledge of land cover characteristics and crop phenology into improving classification performance.

2. Methodology

2.1. Study Area and Datasets

2.1.1. Hetao and Two Case Areas

The Hetao Irrigation District (Hetao), located in the upper Yellow River basin (YRB) of northwest China, is selected as a typical study area with fragmented agricultural landscapes (Figure 1). It is one of the most important agricultural production bases in China, covering an area of 1.12 Mha. Similar to other agricultural areas in northwest China, Hetao has an arid to semi-arid continental climate, with sufficient light and heat resources in the plant growing season but a relatively short frost-free period [31]. A variety of crops are thus planted and grown with high yield/quality and economic values, but with only single-season planting per agricultural year; meanwhile, there is often a large, scatted distribution of natural vegetation patches, due to secondary soil salinity or desertification problems [32]. Irrigation is necessary for crop growth, with water mainly diverted from the Yellow River. In addition to the small farmland size, the landscapes have a typical feature of the fragmentary distribution of various farmlands and natural patches [32]. This pattern of complicated fragmented agricultural landscapes is also a representative of most agricultural areas in northwest China [9,33,34].

In Hetao, the main land use types are cropland and natural land, which account for 51% and 19% of the total area, respectively. The natural land here is a general term for grassland, woods, shrubs, swamps, and abandoned farmlands. It is distributed in flakes within or surrounding the cropland. The remaining parts are distributed with dune (13%), residential area (13%), and water body (4%, including a tail water lake) [35]. The cropland has been divided into small fields and distributed to different smaller holders, due to the special “farmland use rights”. The crop pattern presents a complicated and fragmented distribution in space, with wheat, maize, sunflower, and vegetables (e.g., watermelon, tomato, and pepper) grown. In addition, the crops planted in a certain field will also be changed every 1–3 years to improve soil fertility and help control pests and diseases. As a result, the landscapes are highly fragmented and composited by a mosaic pattern of small mixed croplands and natural vegetable patches [32].

In this study, two case areas in Hetao were selected: (1) one case area named Jiyuan is located in the upstream (western) area of Hetao (40°45′–40°52′N, 106°59′–107°07′E), which covers an area of 8490 ha; and (2) the other case area named Yonglian is located in the downstream (eastern) area of Hetao (41°01′–41°08′N, 107°59′–108°02′E), covering an area of 2940 ha. Sunflower and maize are both the major crops in the two case areas, with wheat or vegetables mainly grown in other cropland areas. Jiyuan can be seen as the representative area in the up-and midstream Hetao where the irrigation water allocation is more sufficient and convenient. The planting structure is more complex and diverse
(with more salt-intolerant crops) than the downstream region, due to less soil salinization. Yonglian could be a representative area in the downstream Hetao with less and later irrigation water allocation and later planting dates. It has a relatively simple cropping pattern (with more salt-tolerant sunflower planted) because of severe soil salinization over a long time [32,36].

Figure 1. Location of the study area (Hetao Irrigation District) and the two case areas (Jiyuan and Yonglian).

2.1.2. Datasets

(1) Satellite data and preprocessing

The surface reflectance data from Sentinel-2 satellite were adopted in this study, being available from Sentinel-2 Level-2A product developed and distributed by ESA (https://scihub.copernicus.eu, accessed on 28 March 2022). Sentinel-2 satellites carry an optical sensor called the Multi-Spectral Instrument (MSI) which acquires data with 13 spectral bands of variable spatial resolution (10–60 m). The combination uses of optical satellite Sentinel-2A and Sentinel-2B could provide image data at an interval of 3–5 days. In this study, all of the images during the period from 100 to 280 day of the year (DOY,
Julian day number) were selected, which covered the growing seasons of all crop types. All bands were resampled to 10 m spatial resolution.

Spectral indices that were sensitive to vegetable greenness and water status were used to capture the physiological differences of land cover types and characterize the growth profile of individual crop types. The Normalized Difference Vegetable Index (NDVI, Equation (1)) [37] and Land Surface Water Index (LSWI, Equation (2)) [38] are two widely used vegetation indices (VIs) and were adopted in this study. They were calculated from the red, near-infrared (NIR), and shortwave infrared (SWIR) spectral bands based on the time series surface reflectance data from Sentinel-2 (Equations (1) and (2)).

\[
\text{NDVI} = \frac{\rho_{\text{NIR}} - \rho_{\text{Red}}}{\rho_{\text{NIR}} + \rho_{\text{Red}}} \tag{1}
\]

\[
\text{LSWI} = \frac{\rho_{\text{NIR}} - \rho_{\text{SWIR}}}{\rho_{\text{NIR}} + \rho_{\text{SWIR}}} \tag{2}
\]

The NDVI reflects the degree of vegetable greenness, and LSWI is sensitive to the land surface moisture from both vegetables and soils. We reconstructed the time series VIs data at a 5-day interval by temporal resampling (using linear interpolation). The temporal resampling should be effective and has no significant impact on the classification accuracy, according to the results of Pelletier et al. [39]. For phenological analysis, the Savitzky–Golay (S–G) filter proposed by Fischer et al. [40] was used to further smooth the NDVI time series to eliminate some small fluctuations. LSWI time series data were not smoothed for phenological analysis, because LSWI is sensitive to vegetable water and soil moisture affected by rainfall and snowfall [24].

(2) Sampling and verification data

Field surveys were carried out to collect ground samples during the two years 2020 and 2021 (Figure 2). The location and the crop type of each sample were recorded using the Global Positioning System (GPS) in the field along the route. The samples involved all land cover types in the study area of Hetao. As a result, 189 samples were collected in Jiyuan in 2020; and 251, 146, and 163 samples were obtained in Jiyuan, Yonglian, and Hetao in 2021, respectively (Table 1). After the field surveys, all the ground samples were visually checked using high-resolution images in Google Earth and Sentinel-2 RGB images. The Planet RGB (R: Red, G: Green, and B: Blue) satellite images (with 3 m spatial resolution) were obtained for Jiyuan on August 5th and Yonglian on July 30th. The true land cover maps were made by visual interpretation of the Planet RGB data, with the combination of field surveys and experiences. They could be used to further test and verify the reliability of the SPLC method.

2.2. Description of SPLC Method

The main principle of the SPLC method and its workflow for land cover classification are provided in Figure 3. The SPLC method mainly consisted of two calculation parts: (1) the support vector machine (SVM) algorithm was used to identify the cropland and each of the non-cropland land cover types, on the basis of optimal spectrum, texture, and LSWI time series data. This step could well eliminate the interference of non-croplands on the crop type classification; and (2) the decision tree algorithm was adopted to identify various crop types with establishing the stable phenological rules, which were based on the phenological metrics extracted from NDVI time series data. A detailed description of the SPLC method was provided in the following sections.

2.2.1. Maps of Non-Cropland Land Cover Types as Masks in the Algorithm

The land cover types could be grouped into five categories in Hetao: residential area, natural land, water body, dune, and cropland. Different land cover types had different spectral and texture characteristics, and the differences were especially obvious in August when crops often grew vigorously. Therefore, the Sentinel-2 data on August 9th was filtered as the optimal bands (blue, green, red, red edge, NIR, and SWIR), and used to calculate
the texture information for classification. Meanwhile, the seasonal dynamics of LSWI data from Sentinel-2 were systematically studied for the cropland (i.e., integration of wheat, maize, sunflower, and vegetables), and various non-croplands (i.e., residential area, natural land, water body, and dune) in two case areas (Figure 4). The LSWI = 0.2 was seen as the threshold to divide the cropland and the non-cropland because it was found that LSWI was larger than 0.2 for crops during the main growing period in the study area. The LSWI dataset was reclassified with the pixel value equal to 1 (if LSWI > 0.2) and equal to 0 (if LSWI ≤ 0.2), respectively. This treatment could better reflect the differences of LSWI between croplands and non-croplands. Finally, the optimal spectrum and texture data and the LSWI time series data were combined and used as the input dataset (i.e., dataset #1) for classification (Figure 3).

Figure 2. The distribution of the ground truth samples in two case areas and Hetao: Jiyuan in 2020 (a) and 2021 (b), Yonglian in 2021 (c), and Hetao in 2021 (d).

The SVM algorithm was then adopted to identify the cropland and each of the non-cropland cover types (i.e., water body, dune, residential area, and natural land), using dataset #1. The high-quality regions of interest (ROIs) training samples were obtained through visual interpretation of the Worldview-3 images that were provided by Google Earth (Table 2). Note that the ROIs could be readily available by visual interpretation without any field surveys. The pixels, which were identified as croplands, were to be extracted and used for further crop classification.
Table 1. Overview of ground truth data in two case areas and Hetao.

| Year          | Water Body | Dune | Natural Land | Residential Area | Wheat | Maize | Sunflower | Vegetable |
|---------------|------------|------|--------------|------------------|-------|-------|-----------|-----------|
| Jiyuan (2021) | Polygon (ROI) | 15   | 10           | 25               | 25    | 20    | 54        | 65        | 37        |
|               | pixel      | 156  | 85           | 225              | 219   | 161   | 575       | 591       | 320       |
| Yonglian (2021)| Polygon (ROI) | 3    | /            | 25               | 22    | /     | 32        | 43        | 21        |
|               | pixel      | 21   | /            | 225              | 198   | /     | 264       | 333       | 193       |
| Jiyuan (2020) | Polygon (ROI) | 10   | 10           | 15               | 20    | 19    | 46        | 39        | 30        |
|               | pixel      | 113  | 85           | 135              | 174   | 156   | 409       | 435       | 252       |
| Hetao (2021)  | Polygon (ROI) | 21   | 25           | 16               | 34    | 16    | 16        | 16        | 19        |
|               | pixel      | 186  | 287          | 144              | 1016  | 144   | 144       | 144       | 171       |

Figure 3. The workflow for identifying and mapping land cover by time series Sentinel-2 data. It includes satellite data preprocessing, reference data collecting, non-cropland mapping, cropland mapping, and accuracy assessment.
Figure 4. Seasonal dynamics of Sentinel-2-based LSWI (a) for cropland, water body, natural land, dune, and residential area. (b–f) show the WorldView-3 images for cropland, water body, natural land, dune, and residential area in September 2021, respectively.

Table 2. Overview of the training samples for classification of the croplands and various non-croplands in the first step of the SPLC method.

| Year  | Water Body | Dune | Natural Land | Residential Area | Cropland |
|-------|------------|------|--------------|------------------|----------|
| Jiyuan (2021) | Polygon (ROI) pixel | 5 | 8 | 15 | 17 | 69 |
| Yonglian (2021) | Polygon (ROI) pixel | 3 | / | 25 | 30 | 30 |
| Jiyuan (2020) | Polygon (ROI) pixel | 5 | 8 | 15 | 17 | 88 |
| Hetao (2021) | Polygon (ROI) pixel | 213 | 117 | 118 | 500 | 515 |

2.2.2. Crop-Classification Based on Phenological and Pixel Patterns

The NDVI temporal profile analysis, widely used to identify the cropping phenological characteristics [41,42], was conducted using the TIMESAT toolbox [43]. NDVI increases from the green-up stage, reaches the maximum during the peak growing stage, and reduces gradually until crop maturity or harvest [44]. Due to the large uncertainty in one-year time series data, replicated time series NDVI data spanning three years were used in this study to extract seasonal parameters. With the TIMESAT toolbox, the NDVI time series data were...
first fitted to a S-G function, and then phenological metrics were extracted to make up the dataset #2 (Figures 3 and 5). The three phenological metrics included the start of the season (SOS, or green-up date), the end of the season (EOS, or senescent date), and the growing-season length (GSL). These features were defined by using the commonly used algorithms and NDVI time series data [43]. The pixels that were clearly identified as croplands in the first step were to be further classified into different crop types. The classification principle was based on the phenological metrics which were generally crop-specific and different for different crop types. The stable classification rules were established with signature analysis of phenology for each crop. The decision tree algorithm was then developed for crop classification based on the obtained phenology-based rules. The typical case of this algorithm for Hetao is presented in Figure 6.

**Figure 5.** An example of NDVI time series (circle symbol) and the fitted curve (dashed line) for a cropland field in the case area. Phenological metrics are labeled.

**Table 3.** Parameters of the decision tree and their calibrated values in two case areas.

| Threshold Parameters (Day) | Description                                                                 | Initial (by Experiences and Surveys) | Calibrated (Jiyuan) | Calibrated (Yonglian) |
|----------------------------|-----------------------------------------------------------------------------|--------------------------------------|---------------------|-----------------------|
| $L_{w_1}$                  | The growing-season length to identify the first type of wheat               | 140                                  | 130                 | 130                   |
| $S_{w_2}$                  | The start of the season to identify the second type of wheat                | 120                                  | 130                 | 130                   |
| $S_{V_1}$                  | The start of the season to identify the first type of vegetable             | 180                                  | 190                 | 200                   |
| $S_{V_2}$                  | The start of the season to identify the second type of vegetable            | 150                                  | 150                 | 160                   |
| $E_{V_3}$                  | The end of the season to identify the third type of vegetable               | 270                                  | 260                 | 260                   |
| $L_{M_S}$                  | The growing-season length to identify maize and sunflower                   | 110                                  | 100                 | 95                    |
In Hetao, the crops can be divided into four major types: wheat, maize, sunflower, and vegetable. The phenological characteristics for these crop types are provided in Figure 7, on the basis of systematic field surveys. The initial values of the threshold in the rules were determined with field surveys. The ground sample data of crops were then overlaid with the three phenological metric layers to carry out a signature analysis. The SOS of wheat occurred mostly in late April to early May. Meanwhile, due to the intercropping of wheat–maize or wheat–sunflower in the study area, the growing-season length (GSL) of this type of wheat was larger than other crop types. The SOS of maize was mostly in June. The SOS of sunflower was slightly later than maize, while the GSL was shorter than that of maize. There were many types of vegetables planted in Hetao, so the phenological characteristics were more complicated than those for other crop types. Generally speaking, it can be divided into two main categories. The first kind of vegetables (e.g., watermelon and tomato) was sown in later June and harvested in late August, so the SOS of this kind of vegetables was late, and the GSL was relatively short. Another kind of vegetables (e.g., squash (Cucurbita pepo L.) and pepper) had similar phenological characteristics to sunflower, which was sown in mid-June and harvested in mid-to-late September (Figure 8). Based on the above systematic field surveys and signature analysis of phenological metrics, the thresholds for the decision tree algorithm were finally defined (Table 3).

![Decision Tree Diagram](image)

Figure 6. The decision tree rules for crop classification. The meaning of the parameters and their calibrated values are shown in Table 3.

![Crop Calendars](image)

Figure 7. Crop calendars for the major crops in Hetao. This figure shows the crop phenology stages which start in April and end in September within a year.
Figure 8. Signature analysis of crops (wheat, maize, sunflower, and vegetable) at each phenology metric layer using the histogram statistics method (a–c). The training samples used in this analysis are shown in Figure 2. (d–g) show the field photos for wheat, maize, sunflower, and vegetable, respectively, taken in July and August 2021.

2.3. Assessment of Classifier Performance

The ground truth sample data (Table 1) collected in 2020 and 2021 was used for accuracy evaluation. A confusion matrix was created to evaluate the accuracy of the classification by using indicators of producer’s accuracy (PA), user’s accuracy (UA), and overall accuracy (OA). Then, the identified land cover maps by the SPLC method were further assessed by comparing them with the true land cover maps by visual interpretation in space, both for Jiyuan and Yonglian (in 2021). In addition, the total area for each land cover type was calculated for the SPLC method and compared to those for the visual interpretation method. After detailed testing in two case study areas, the SPLC method was further validated in the entire Hetao region using the ground truth sample data in 2021 (Table 1).

3. Results and Discussion

3.1. Overall Classification Accuracy

3.1.1. Accuracy Evaluation Using Reference Data

Four land cover maps (i.e., Jiyuan in 2020 and 2021, Yonglian in 2021, and Hetao in 2021) were produced by the SPLC method. The calibrated values of threshold parameters for phenological rules were obtained by slight adjustments of the initial ones (Table 3). There were little differences in parameter values between Jiyuan (upstream Hetao) and Yonglian (downstream Hetao), with a maximum difference of 10 days (e.g., $S_{v_1}$). Two case testing indicated that the phenological metrics were stable enough to be used for crop classification. The evaluation results of the confusion matrix are presented in Table 4.
Table 4. The confusion matrix of land cover identification accuracy (PA: producer’s accuracy; UA: user’s accuracy).

| Accuracy | Water Body | Dune | Natural Land | Residential Area | Wheat | Maize | Sunflower | Vegetable |
|----------|------------|------|--------------|------------------|-------|-------|-----------|-----------|
| Jiyuan (2021) | PA 0.82 | 0.91 | 0.98 | 0.89 | 0.98 | 0.99 | 0.98 | 0.78 |
|          | UA 1.00 | 0.91 | 0.95 | 0.95 | 0.97 | 0.96 | 0.92 | 0.94 |
|          | Overall Accuracy: 0.94 |
| Yonglian (2021) | PA 0.86 | / | 0.87 | 0.87 | / | 0.93 | 0.76 | 0.99 |
|          | UA 1.00 | / | 0.87 | 0.88 | / | 0.94 | 0.99 | 0.74 |
|          | Overall Accuracy: 0.90 |
| Jiyuan (2020) | PA 0.86 | 0.98 | 0.98 | 0.89 | 0.95 | 0.94 | 0.82 | 0.78 |
|          | UA 1.00 | 0.93 | 0.96 | 0.97 | 0.78 | 0.85 | 0.93 | 0.85 |
|          | Overall Accuracy: 0.91 |
| Hetao (2021) | PA 1.00 | 0.93 | 0.89 | 0.98 | 0.92 | 0.89 | 0.95 | 0.90 |
|          | UA 1.00 | 0.84 | 1.00 | 0.97 | 0.92 | 0.96 | 0.88 | 0.94 |
|          | Overall Accuracy: 0.94 |

In the two case areas, results showed that the various non-cropland cover types (i.e., water body, dune, residential area, and natural land) could be effectively identified in the first step of the SPLC method. The OA values reached from 0.90 to 0.94 for the three land cover maps, and meanwhile, the PA and UA of non-cropland land cover types were both over 0.82. These indicated good performance in the classification of these land cover types. For crop identification, the maize was accurately identified with all values of UA and PA being greater than 0.9 except for the UA in 2020 of Jiyuan (0.85). This resulted from the phenological characteristics of maize being more obvious than those of other crop types. Sunflower and wheat had relatively lower accuracies than maize, except for the UA of wheat in Jiyuan in 2020 and the PA of sunflower in Yonglian in 2021, which were both lower than 0.8. The other UA and PA values were over 0.8 for three land cover maps. Vegetables had the lowest inversion accuracy, with about half of accuracy evaluation values lower than 0.8. This can be attributed to the phenological metrics of vegetables being relatively not so determined compared with those of other crops. Various varieties and different irrigation and field management measures further made the phenological metrics more undetermined for vegetables.

Furthermore, the evaluation results of the confusion matrix (Table 4) indicated the good performance in the land cover classification for the entire Hetao region. The OA value reached 0.94, and the PA and UA for all land cover types were greater than 0.84 in Hetao. These further verified the reliability of the proposed SPLC method on large regional scales.

3.1.2. Performance of the Classifier at the Spatial Scale

The comparison of land cover maps by the SPLC method and that by visual interpretation are shown in Figures 9 and 10 for Jiyuan and Yonglian, respectively. It was obvious that the results of the SPLC method had very high similarity with the visual interpretation maps in space. Furthermore, four typical regions in Jiyuan and three regions in Yonglian were respectively selected to evaluate the accuracy of the SPLC method for different land cover types. Comparison results indicated that different non-cropland land cover types were identified with high accuracy, especially for water body areas (Figures 9f and 10c).

It is worth noting that accurate identification of non-cropland in the first step effectively removed the negative influences on the further classification of various crop types. The spectral characteristics of natural lands with high soil salinity levels were similar to those of residential areas, which led to the misclassification between natural lands and residential areas. In addition, some narrow roads, drainage ditches, and canals were still not yet well identified, because the spatial resolution of the Sentinel-2 images (10 m) was lower than that of the Planet images (3 m) for visual interpretation. For the crop maps, most sunflower,
maize, and wheat fields were accurately identified; and meanwhile, the accuracy for vegetables was relatively low due to less clear phenological characteristics for vegetables. In Yonglian, some sunflower fields in the visual interpretation map were misclassified as other types by the SPLC method. This was mainly because soil salinization, which was often in sunflower fields or their local parts, would result in low emergence and survival rates. Therefore, the characteristics in the internal pixels of some sunflower fields may be similar to other land cover types (e.g., natural land and vegetables). This resulted in the problem of mixed classification. Moreover, many small and scattered fields could be accurately identified with the SPLC method (Figures 9 and 10), compared to the visual interpretation method, because it may be not possible or realistic to identify these scattered fields one by one using manual visual interpretation.

Figure 9. Land cover classification maps for Jiyuan in 2021: (a) for visual interpretation, and (b) for SPLC method. Four sample areas (c–f) were selected to show the zoom-in views.
The statistical areas of all land cover types for the SPLC method were presented and compared with those for visual interpretation in 2021 (Figure 11). The statistical areas calculated by the two methods agreed with each other. There was a significant linear relationship between the results of the two methods, producing a coefficient of determination ($R^2$) of 0.97 and a coefficient of regression ($b$) of 1.0 in Jiyuan (2021), and $R^2 = 0.85$, $b = 0.84$ in Yonglian (2021). This further implied that the SPLC method had high accuracy in two case areas, and thus had good regional applicability in Hetao. In addition, the total area of sunflower (946 ha) in Yonglian was about 25% lower than that estimated from the visual interpretation map (1262 ha) (Figure 11), mainly due to the high similarity in phenological metrics between several salted-affected sunflower fields and natural lands as explained above. In addition, the image comparisons of the local area in Hetao also showed that the accuracy of the classification results was reliable (Figure 12).

**Figure 10.** Land cover classification maps for Yonglian in 2021: (a) for visual interpretation, and (b) for SPLC method. Three sample areas (c–e) were selected to show the zoom-in views.

**Figure 11.** A comparison of the statistical area for different land cover types by SPLC method and visual interpretation method for Jiyuan and Yongliang in 2021.
Our method could achieve better or similar accuracies compared with other similar studies [10, 22, 45]. Although some studies had acquired higher accuracies, their classification targets were usually easier to differentiate due to significantly different characteristics, or the relatively simple and uniform crop planting structure in the study area [24, 46]. Furthermore, compared with other studies in Hetao [42, 47, 48], the SPLC method may have several improvements or advantages: (1) compared with the study that classified all land cover types together [47], the non-cropland land cover types and crop types were identified by the SPLC method in two steps. Therefore, the negative impacts of other non-cropland land covers on crop identification (especially the lush natural land grown with salt-tolerant vegetation) were avoided as much as possible; (2) the classification rules established based on the three phenological metrics (SOS, EOS, and GSL) were more effective for crop classification. Compared with the classification methods which used other phenological metrics [42, 48], the mixed classification was significantly reduced and the accuracy of crop classification was improved; (3) the land cover classification with more crop types was considered and realized in our study; (4) compared with other methods, which were based on ground truth training sampling [42, 47], the ground sampling work was not required when using the SPLC method. This could save time and labor costs, and meanwhile avoid the possible negative impacts of the quality of training points on the classification accuracy.
3.2. Analysis of Land Cover Mapping Results

The spatial distribution of land covers in Jiyuan and Yonglian (during 2021) are presented in Figures 9b and 10b, respectively. For Jiyuan, the cropland (wheat, maize, sunflower, and vegetables) accounted for about 73.9% of the total area, followed by residential area (14%), natural land (10.3%), dune (1.1%), and water body areas (0.7%). In Yonglian, the cropland was the dominant land cover type, accounting for about 60% of the total area. Natural land reached 566.5 ha, which was close to the area of residential land (563.2 ha). Water bodies only occupied 0.1% of the total area. Note that large areas of natural land with salt-tolerant vegetation were both distributed in the northern parts of the two case areas. This was due to the relatively high soil salinity levels caused by less irrigation and lower drainage effects in northern downstream regions.

Sunflower was the main crop type, accounting for about 48% and 58% of cropland area in Jiyuan and Yonglian, respectively. The planting area of maize was the second largest, accounting for about 30% of the total cropland area both for Jiyuan and Yonglian. In Jiyuan, the area of wheat was 620.2 ha, which was only 9% higher than that of vegetables (569.0 ha). These two crop types both occupied a small proportion of croplands (i.e., 10% for wheat and 9% for vegetables), and most of them had a scattered distribution in the area. There was no wheat planted in Yonglian, and a small part (about 11%) of the area was grown with vegetables. Wheat and vegetables were mainly planted in Jiyuan, while the vast majority of cropland in Yonglian was planted with sunflower. This was because the soil salinity was higher in the downstream Hetao (e.g., Yonglian), and lower in up-and midstream Hetao (e.g., Jiyuan) [49]. The spatial distribution of crops was almost consistent with soil salinity distribution. The croplands with serious salinity problems were usually planted with more salt-tolerant sunflower which had satisfactory economic benefits in salt-affected croplands. The less salt-tolerant crops (maize, wheat, and vegetables) were mainly distributed in non-salted or slightly salt-affected lands, with timely and adequate irrigation.

In addition, the land cover spatial distribution in Jiyuan was more fragmented than that of Yonglian, with different land covers intertwined (Figures 9 and 10). This is because farmers can grow flexibly and diversely with more irrigation allocation and better water conveyance services in Jiyuan upstream of Hetao. Besides, the local managers had carried out a lot of land transfer and integration work in recent years in Yonglian, due to less population/production and stronger demand for intensive planting in the downstream Hetao. Several large demonstration farms were also built in the case area of Yonglian, and thus the planting pattern in Yonglian was relatively simple and uniform in some local places.

The land cover map of Hetao was produced for the year 2021, as shown in Figure 12. The cropland (wheat, maize, sunflower, and vegetables) accounted for about 54.8% of the total area in Hetao, followed by residential area (25.8%), dune (10.1%), natural land (4.9%), and water body (4.4%). Note that large areas of dune were distributed in the western Hetao which was adjacent to the eastern part of the Wulanhuhe desert (Figure 1). Most of the water bodies were sporadically distributed in Hetao, while the largest one (i.e., the Wuliangsu Lake) was located in the east region as a tail water lake. Results indicated that maize was the main crop type in Hetao, accounting for about 44% of the total cropland area. Meanwhile, the planting area of sunflower was the second largest area, being about 35% of the total cropland area. The areas of vegetables and wheat both occupied a small proportion of croplands (i.e., 8% for wheat and 13% for vegetables), and most of them had a scattered distribution in Hetao. Meanwhile, wheat was mainly planted in the upstream regions. In addition, results of the entire Hetao region further indicated that the spatial distribution of land cover was more fragmented in the upstream area than that in the downstream area (Figure 12). Land cover distribution and fragmentation degree were largely affected by the irrigation and salt salinity conditions as discussed above, resulting in significant spatial differences between the upstream and downstream areas of Hetao.
3.3. Implications and Improvements

Case testing showed that using the SPLC method could efficiently identify the fragmented land covers with various crop types. In particular, once the phenological rules are established, the SPLC method can be used to make land cover maps in other seasons/years, without the need for ground truth sampling. Based on the optimal spectrum, texture, and LSWI time series data, the SVM algorithm could be used to identify the cropland and various non-cropland land cover types. This algorithm uses the ROIs (obtained from Google Earth) as training samples, and thus saves economic and time costs. The thresholds used for the LSWI time series data have been well applied in other regions (e.g., [23,44]), which may imply that our cropland and non-cropland land cover types classification algorithm may have good regional applicability. Moreover, the high accuracy classification of non-croplands in the first step can efficiently remove the possible negative effects of non-croplands on further crop classification.

The crop classification rules are developed by selecting the unique phenological metrics of different crop types. For a given area, the environmental factors of crop planting, such as topography, climate, planting, and management activities, are relatively stable and consistent over normal years. The SPLC method adopts an automated decision tree algorithm with stable classification rules for crop mapping. Therefore, it is promising to use the SPLC method to map the crop classification distribution in other seasons/years, without the requirement of ground truth sampling data and re-training and re-validation. This could save time and labor costs, and avoid the possible negative impacts of the quality of training points on the classification accuracy. With respect to the crop mapping in other regions and crop types, the SPLC method could be applied by properly adjusting the classification rules and the thresholds of phenological metrics with local sampling/survey data or known metric information. It seems to be feasible, since the potential differences in the phenological metrics of different crops under different environmental conditions can be addressed. This has also been supported by some previous studies on using phenology-based approaches to map various dominant crop types (e.g., paddy rice, oat, maize, and soybean) at 10-m or coarser spatial resolutions [10,14,50]. Currently, various machine learning and deep learning models are explored for crop classification [27,51]. This crop classification framework can be updated by integrating the phenological metrics and the machine learning models to map crop types in the follow-up investigations.

In addition, the study area has an arid climate with infrequent cloud cover, so adequate numbers of good observations are available by Sentinel-2 image data. However, the images from single sensors are not enough in some regions with more frequent cloud cover, so the harmonization of the images from different sensors is thus necessary [52,53]. With the development of the fusion methods, the blending of time series images with high spatial and temporal resolution could be obtained, and the crop mapping algorithms will be developed at the field scale over large spatial domains.

4. Conclusions

A novel spectral–phenological land cover classification (SPLC) method that was proposed for efficient land cover identification in fragmented agricultural landscapes was presented in this paper. The SPLC method was developed based on the integration of the time series optical images (Sentinel-2A/B), support vector machine (SVM), and decision tree algorithm. It was well tested and applied in two case areas (i.e., Jiyuan in upstream and Yonglian in downstream) of Hetao Irrigation District (Hetao) of the upper Yellow River basin, Northwest China. Case testing indicated that the first step of the SPLC method was able to identify the cropland and various non-cropland land cover types, using the SVM algorithm and the input data of optimal spectrum, textures, and LSWI time series. This step could efficiently remove the possible negative effects of non-croplands on the subsequent crop classification. For crop identification, analyses showed that stable classification rules could be established by using the phenological metrics retrieved from the NDVI time series. In addition, the evaluation in the entire Hetao further validated the reliability of the SPLC.
method in land cover classification for fragmented agricultural landscapes. Finally, four land cover maps at 10-m spatial resolution for Jiyuan in 2020 and 2021, Yonglian in 2021, and Hetao in 2021 were produced with high accuracy. The main advantages of the SPLC method can be concluded as: (1) it could be used to map crop classification distribution in other seasons/years, without requirements of ground truth sampling data, once the classification rules are established; and (2) it was capable of avoiding mixed classification problems due to its strong physiological meaning. Overall, the potential of high temporal and spatial resolution satellite remote sensing data in extracting phenological metrics and realizing crop classification was fully presented in this study, especially for regions with complex agricultural landscapes. The proposed SPLC method could also be applied for crop mapping in other agricultural regions with complicated land covers, by properly adjusting the classification rules and the thresholds of phenological metrics.

Author Contributions: Investigation, X.L., H.M., X.M. and X.X.; Methodology, X.L., C.S. and X.X.; Supervision, C.S., X.X. and G.H.; Validation, X.L., H.M. and X.M.; Writing—original draft, X.L.; Writing—review & editing, C.S. and X.X. All authors have read and agreed to the published version of the manuscript.

Funding: This research was jointly supported by the National Natural Science Foundation of China (grant numbers: 52022108 and U2243217), the project “Research on efficient operation and guarantee technology for drip irrigation system and farmland irrigation/drainage system in Hetao Irrigation District” (2021-2023) funded by Water Resources Department of Inner Mongolia Autonomous Region, and the 13th Five-year National Key Research and Development Program of the Chinese Ministry of Science and Technology (grant numbers: 2107YFD0300403).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References
1. Griffiths, P.; Nendel, C.; Hostert, P. Intra-annual reflectance composites from Sentinel-2 and Landsat for national-scale crop and land cover mapping. Remote Sens. 2019, 220, 135–151. [CrossRef]
2. Foley, J.A.; Defries, R.; Asner, G.P.; Barford, C.; Bonan, G.; Carpenter, S.R.; Chapin, F.S.; Coe, M.T.; Daily, G.C.; Gibbs, H.K.; et al. Global consequences of land use. Science 2005, 309, 570–574. [CrossRef] [PubMed]
3. Löw, F.; Duveiller, G. Defining the spatial resolution requirements for crop identification using optical remote sensing. Remote Sens. 2014, 6, 9034–9063. [CrossRef]
4. Feng, X.; Fu, B.; Yang, X.; Lü, Y. Remote sensing of ecosystem services: An opportunity for spatially explicit assessment. Chin. Geogr. Sci. 2010, 20, 522–535. [CrossRef]
5. Alcantara, C.; Kuemmerle, T.; Prischepov, A.V.; Radeloff, V.C. Mapping abandoned agriculture with multi-temporal MODIS satellite data. Remote Sens. Environ. 2012, 124, 334–347. [CrossRef]
6. Jia, K.; Liang, S.; Wei, X.; Yao, Y.; Su, Y.; Jiang, B.; Wang, X. Land cover classification of Landsat data with phenological features extracted from time series MODIS NDVI data. Remote Sens. 2014, 6, 11518–11532. [CrossRef]
7. Qin, Y.; Xiao, X.; Dong, J.; Chen, B.; Liu, F.; Zhang, G.; Zhang, Y.; Wang, J.; Wu, X. Quantifying annual changes in built-up area in complex urban-rural landscapes from analyses of PALSAR and Landsat images. ISPRS J. Photogramm. Remote Sens. 2017, 124, 89–105. [CrossRef]
8. Villa, P.; Pinardi, M.; Bolpagni, R.; Gillier, J.-M.; Zinke, P.; Nedelcuţ, F.; Bresciani, M. Assessing macrophyte seasonal dynamics using dense time series of medium resolution satellite data. Remote Sens. Environ. 2018, 216, 230–244. [CrossRef]
9. Zhong, B.; Ma, P.; Nie, A.; Yang, A.; Yao, Y.; Lü, W.; Zhang, H.; Liu, Q. Land cover mapping using time series HJ-1/CCD data. Sci. China Earth Sci. 2014, 57, 1790–1799. [CrossRef]
10. Bargiel, D. A new method for crop classification combining time series of radar images and crop phenology information. Remote Sens. Environ. 2017, 198, 369–383. [CrossRef]
11. Shao, Y.; Taff, G.N.; Ren, J.; Campbell, J.B. Characterizing major agricultural land change trends in the Western Corn Belt. ISPRS J. Photogramm. Remote Sens. 2016, 122, 116–125. [CrossRef]
38. Xiao, X.; Boles, S.; Liu, J.; Zhuang, D.; Frolking, S.; Li, C.; Salas, W.; Moore, B. Mapping paddy rice agriculture in southern China using multi-temporal MODIS images. Remote Sens. Environ. 2005, 95, 480–492. [CrossRef]

39. Pelletier, C.; Valero, S.; Ingliada, J.; Champion, N.; Dedieu, G. Assessing the robustness of Random Forests to map land cover with high resolution satellite image time series over large areas. Remote Sens. Environ. 2016, 187, 156–168. [CrossRef]

40. Fischer, G.; Shah, M.; Van Velthuizen, H.; Nachtergaele, F.O. Global Agro-Ecological Assessment for Agriculture in the 21st Century; IIASA: Laxenburg, Austria, 2001.

41. Jönsson, P.; Eklundh, L. Seasonality extraction by function fitting to time-series of satellite sensor data. IEEE Trans. Geosci. Remote Sens. 2002, 40, 1824–1832. [CrossRef]

42. Wen, Y.; Shang, S.; Rahman, K. Pre-constrained machine learning method for multi-year mapping of three major crops in a large irrigation district. Remote Sens. 2019, 11, 242. [CrossRef]

43. Jönsson, P.; Eklundh, L. TIMESAT—A program for analyzing time-series of satellite sensor data. Comput. Geosci. 2004, 30, 833–845. [CrossRef]

44. Chen, B.; Xiao, X.; Ye, H.; Ma, J.; Doughty, R.; Li, X.; Zhao, B.; Wu, Z.; Sun, R.; Dong, J.; et al. Mapping forest and their spatial–temporal changes from 2007 to 2015 in tropical Hainan Island by integrating ALOS/ALOS-2 L-band SAR and Landsat optical images. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 2018, 11, 852–867. [CrossRef]

45. Arvor, D.; Jonathan, M.; Meirelles, M.S.P.; Dubreuil, V.; Durieux, L. Classification of MODIS EVI time series for crop mapping in the state of Mato Grosso, Brazil. Int. J. Remote Sens. 2011, 32, 7847–7871. [CrossRef]

46. Zhong, L.; Gong, P.; Biging, G.S. Efficient corn and soybean mapping with temporal extendability: A multi-year experiment using Landsat imagery. Remote Sens. Environ. 2014, 140, 1–13. [CrossRef]

47. Feng, Z.; Huang, G.; Chi, D. Classification of the complex agricultural planting structure with a Semi-Supervised Extreme Learning Machine framework. Remote Sens. 2020, 12, 3708. [CrossRef]

48. Yu, B.; Shang, S. Multi-Year Mapping of maize and sunflower in Hetao Irrigation District of China with high spatial and temporal resolution vegetation index series. Remote Sens. 2017, 9, 855. [CrossRef]

49. Huang, Q.; Xu, X.; Lü, L.; Ren, D.; Ke, J.; Xiong, Y.; Huo, Z.; Huang, G. Soil salinity distribution based on remote sensing and its effect on crop growth in Hetao Irrigation District. Trans. Chin. Soc. Agric. Eng. 2018, 34, 102–109.

50. Massey, R.; Sankey, T.T.; Congalton, R.G.; Yadav, K.; Thenkabail, P.S.; Ozdogan, M.; Sánchez Meador, A.J. MODIS phenology-derived, multi-year distribution of conterminous U.S. crop types. Remote Sens. Environ. 2017, 198, 490–503. [CrossRef]

51. Luciano, A.C.d.S.; Picoli, M.C.A.; Rocha, J.V.; Duft, D.G.; Lamparelli, R.A.C.; Leal, M.R.L.V.; Le Maire, G. A generalized space-time OBIA classification scheme to map sugarcane areas at regional scale, using Landsat images time-series and the random forest algorithm. Int. J. Appl. Earth Obs. Geoinf. 2019, 80, 127–136. [CrossRef]

52. Claverie, M.; Ju, J.; Masek, J.G.; Dungan, J.L.; Vermote, E.F.; Roger, J.-C.; Skakun, S.V.; Justice, C. The Harmonized Landsat and Sentinel-2 surface reflectance data set. Remote Sens. Environ. 2018, 219, 145–161. [CrossRef]

53. Reiche, J.; Verbesselt, J.; Hoekman, D.; Herold, M. Fusing Landsat and SAR time series to detect deforestation in the tropics. Remote Sens. Environ. 2015, 156, 276–293. [CrossRef]