agriculture

MDPI

Article

Advancing County-Level Potato Cultivation Area Extraction: A Novel Approach Utilizing Multi-Source Remote Sensing Imagery and the Shapley Additive Explanations–Sequential Forward Selection–Random Forest Model

Qiao Li, Xueliang Fu *, Honghui Li and Hao Zhou

School of Computer and Information Engineering, Inner Mongolia Agricultural University, Hohhot 010011, China; liq@emails.imau.edu.cn (Q.L.); lihh@imau.edu.cn (H.L.); zhouhh@emails.imau.edu.cn (H.Z.)
* Correspondence: fuxl@imau.edu.cn; Tel.: +86-159-4761-7598

Abstract: Potato, a vital food and cash crop, necessitates precise identification and area estimation for effective planting planning, market regulation, and yield forecasting. However, extracting large-scale crop areas using satellite remote sensing is fraught with challenges, such as low spatial resolution, cloud interference, and revisit cycle limitations, impeding the creation of high-quality time–series datasets. In this study, we developed a high-resolution vegetation index time–series by calculating coordination coefficients and integrating reflectance data from Landsat-8, Landsat-9, and Sentinel-2 satellites. The vegetation index time–series were enhanced through using linear interpolation and Savitzky–Golay (S-G) filtering to reconstruct high-quality data. We employed the harmonic analysis of NDVI time–series (HANTS) method to extract features from the time–series and evaluated the classification accuracy across five feature sets: vegetation index time–series features, band means, vegetation index means, texture features, and color space features. The Random Forest (RF) model, utilizing the full feature set, emerged as the most accurate, achieving a precision rate of 0.97 and a kappa value of 0.94. We further refined the feature subset using the SHAP-SFS feature selection method, leading to the SHAP-SFS-RF classification approach for differentiating potato from non-potato crops. This approach enhanced accuracy by approximately 0.1 and kappa value by around 0.2 compared to the RF model, with the extracted areas closely aligning with statistical yearbook data. Our study successfully achieved the accurate extraction of potato planting areas at the county level, offering novel insights and methodologies for related research fields.

Keywords: potato; planting area extraction; remote sensing; feature extraction; machine learning

Academic Editor: Tarendra Lakhankar

Received: 14 November 2024
Revised: 28 December 2024
Accepted: 29 December 2024
Published: 3 January 2025

Citation: Li, Q.; Fu, X.; Li, H.; Zhou, H. Advancing County-Level Potato Cultivation Area Extraction: A Novel Approach Utilizing Multi-Source Remote Sensing Imagery and the Shapley Additive Explanations–Sequential Forward Selection–Random Forest Model. *Agriculture* **2025**, *15*, 92. https://doi.org/10.3390/agriculture15010092

1. Introduction

As the fourth most significant food crop globally and the only tuber crop used as a staple, the potato not only has significant economic value, but also demonstrates excellent adaptability [1,2]. Drawing from the latest statistics, the global potato planting area has exceeded 17.8 million hectares, with an annual output of 374 million tons, while China leads globally in terms of the cultivation area and output [3]. Given its strong adaptability, high productivity, and rich nutritional value, the potato is essential for guaranteeing food security and promoting sustainable agricultural development in a context faced with global challenges such as population explosion, climate change, and frequent natural

disasters [4]. China possesses a rich historical background in potato cultivation, with a widespread distribution across the country, divided into four major agro-ecological regions based on natural conditions, with each region complementing the other through seasonal differences to form a large-scale production pattern. Wuchuan County in Inner Mongolia, as a representative of the northern one-season crop area, has become one of the main advantageous potato production areas in the country by virtue of its suitable climatic conditions for potato growth. Therefore, precise identification of potato-growing areas is of great significance for market regulation, policy formulation, and food security guarantee at the national and regional levels.

Traditional crop acreage statistics usually use remote sensing images combined with spatial sampling techniques to select survey samples, and grass-roots investigators measure the crops in the sample plots on the ground and then synthesize the data to estimate the acreage of major food crops [5]. However, this traditional statistical method, which relies on manual labor and reporting level by level, encounters the dual problems of timeliness and accuracy in the implementation process, posing difficulties in ensuring the validity and reliability of the data. Satellite remote sensing technology, with its wide coverage, strong real-time monitoring capability, and rich information, can quickly and conveniently obtain surface observation information on a regional scale, providing an efficient means of extracting potato planting area in a large region [6]. Utilizing satellite remote sensing technology for identifying and extracting crop planting areas represents a key aspect of precision agriculture, which to a certain extent promotes the development of intelligent agriculture 4.0. At the same time, fast and accurate large-scale extraction of potato planting area is fundamentally important in monitoring the growth condition of potato and predicting the yield. Using a combination of bands 5, 6, and 4 of the Landsat-8 satellite, Li, D. et al. [7] conducted a comprehensive analysis of five periods of imagery during the period from June to August 2017, and accurately extracted and analyzed potato cultivation in Keshan Farm, Heilongjiang Province; the calculated cultivated area was highly consistent with the actual data, demonstrating a high extraction accuracy. Ashourloo, D. et al. [8] identified four unique spectral properties of potato, including near-infrared reflectance values at cultivation and harvest dates, changes in near-infrared reflectance at peak greenness, and the ratio of near-infrared reflectance values to red reflectance values at peak greenness, which were used to differentiate potato from other crops; the overall accuracy of the method in tests conducted at four research sites in Iran and the U.S. was over 90% and the kappa coefficient was also higher than 0.8, showing excellent differentiation ability. In 2023, Zhao, L.H. et al. [9] utilized Google Earth Engine (GEE) and Sentinel-2 (S2) remote sensing data to extract spectral, index, and textural features of potato fields in Shandong Province, China. Through their extraction and monitoring efforts, they achieved a classification accuracy of 92.5%, and the kappa coefficient, a measure of inter-rater reliability, reached an impressive 0.916. Although these studies have achieved remarkable results, the application of satellite remote sensing technology in large-scale potato crop extraction still faces challenges. The spatial and temporal resolution of satellite remote sensing data limits the ability to obtain detailed information when planning large-scale crop cultivation and resource allocation. In addition, uncertainties in the acquisition and resolution of satellite remote sensing data and in feature extraction may affect the market's judgment and regulation of crop supply and demand. The relationship between satellite remote sensing data and yield exhibits variability across years and regions, so exploring stable large-area potato crop extraction models across regions and years is critical to improving the accuracy and stability of yield estimation models.

The methods employed for recognizing crops and measuring their cultivation areas via satellite remote sensing are mainly categorized into three types: based on single-temporal-

phase images, based on multi-temporal-phase images, and based on time–series images. Single-time-phase images refer to remote sensing images captured at a single specific moment, and researchers analyze the spectral information and image features of these images to identify crops and extract planting areas. However, due to the cyclical nature of crops, single-time-phase data cannot fully capture the growth dynamics of crops and are susceptible to factors such as light, growth stage, and shooting angle, which limits the accuracy of identification results. Multi-temporal remote sensing imagery involves sequences of crop images collected at different points in time, and this method can capture the growth cycle and morphological changes of crops, providing richer information for crop identification [10]. Time–series imagery, on the other hand, focuses on analyzing continuous changes in crop remote sensing images over time, and is more focused on revealing the evolutionary trend of a variable over time than multi-temporal data [11]. In the field of agriculture, the sequential Normalized Difference Vegetation Index (NDVI) data are extensively utilized because they provide a robust indication of the growth cycle and climatic characteristics of crops [12]. Nevertheless, the construction of satellite remote sensing-based time–series datasets faces multiple challenges, including frequent cloud cover; the limitations of remote sensing technology are often evident in the realms of temporal, spatial, and spectral resolution. All these factors may have an influence on the quality of the dataset and conclusions drawn from the analysis [13].

Remote sensing satellites that provide high-temporal-resolution image data are often accompanied by coarse spatial resolution. An example is the Moderate-Resolution Imaging Spectroradiometer (MODIS) instrument on the Terra and Aqua satellite platforms; this offers a revisit interval ranging from 1 to 2 days, but the spatial resolution is only between 250 and 1000 m. The Landsat series of satellites, on the other hand, provide a spatial resolution of 30 m, but have an Earth-visit cycle of up to 16 days. High-resolution multispectral imaging satellites such as S2, featuring a 5-day revisit period, are capable of spatial-resolution-based monitoring at 10, 20, and 60 m. In order to enhance the temporal continuity of clear-resolution images, researchers often fuse remote sensing data from multiple satellites. The combination of Landsat and S2 sensors offers an extensive array of multispectral data with medium–high resolution, finding significant utility across various domains. The commonality in the design of these sensors further enhances the compatibility of their data. Li et al. [14] achieved a mean global revisit frequency of 2.9 days by combining Landsat-8 (L8) and S2 satellite data. Army et al. [15] improved the accuracy of vegetation phenology monitoring by coordinating reflectance and fusing vegetation index (VI) data using L8, S2A, and GF-1 satellite data. Landsat-9 (L9) was successfully deployed on 27 September 2021; this provided additional opportunities for processing time–series of multispectral moderate resolution images. However, due to the differences in orbit, illumination, and viewing angle during data acquisition of different satellite sensors, the acquired surface reflectance values may differ; therefore, the uniform calibration of these differences is required. Trevisiol, F. et al. [16] introduced the ideal linear transformation factors for the European region, achieving compatibility between L9 and L8 as well as S2 across Europe, which provides an important reference for cross-sensor data fusion.

It has been demonstrated that machine learning (ML) and deep learning (DL) methodologies are efficacious for enabling crop classification, acreage, and the retrieval of pivotal data from remote sensing sources, including unmanned aerial vehicles (UAVs) and satellites [17]. The performance of ML models relies heavily on extracting, constructing, and filtering useful features from raw data. And DL models can automatically capture deep features in the data through their multi-level neural network structures. Nevertheless, DL model training process necessitates an extensive dataset, and it is a challenge to collect multi-year wide-area crop data in the face of large-scale areas, complex cropping structures,

and variable terrain; these factors limit the application of DL models in these scenarios. Therefore, ML methods have been more widely applied in large-scale crop area extraction from satellite remote sensing [18–20]. For feature extraction, ML models rely on spectral and spatial features extracted from remotely sensed data, which can be either original or derived, with the aim of enhancing the separability between targets. Spectral feature extraction includes methods such as Principal Component Transform, Minimum Noise Separation, Tassel Hat Transform, and VI, while spatial feature extraction involves texture, shape, and spatial relationship features. These features are essential in enhancing the precision of remote sensing image categorization.

In the domain of advancing remote sensing technology, there is a substantial generation of remote sensing data [21], which puts higher requirements on downloading and processing [22]. To solve these problems, several online processing platforms for remote sensing data have emerged, among which GEE is the most prominent platform, which delivers powerful cloud computing capability, massive data sources, efficient data processing and multiple programming language support, and finds broad application in a spectrum of research fields, encompassing land cover [23], crop classification [24] and mapping spatial distribution of crops [25] and other research fields.

The majority of prior research has concentrated on the regions where crops like maize, rice, and wheat are cultivated; however, the importance of potato as a key crop in some specific regions, such as cold zones at high latitudes, should not be ignored, and studies on the extraction and distribution of potato crop area are still in the developmental stages. In this study, Wuchuan County in Hohhot City, Inner Mongolia, the main production area of potato, was taken as the study area. The aims were as follows: (1) Calculate the best linear transformation coefficient to harmonize the time–series satellite imagery from L8, L9, and S2 over the research region. (2) Calculate and reconstruct the vegetation index time–series (VI time–series) data; time–series features were extracted based on harmonic analysis. (3) Calculate the multi-temporal band reflectance means, multi-temporal vegetation index means (VI time means), texture features, and color space features as additional features. (4) Evaluate the accuracy of ML models and feature extraction algorithms for potato crop classification.

2. Materials and Methods

2.1. Study Area

Wuchuan County, situated within the heart of the Inner Mongolia Autonomous Region, falling under Hohhot City's governance, is the core area of agricultural production and ecological functions in the northern part of the city. The county is situated at the northern base of the Yinshan Mountains, with geographic coordinates ranging from 40°47′~41°23′ north latitude to 110°31′~111°53′ east longitude, and altitudes ranging from 1219 m to 2259 m (Figure 1). The climate features a mid-temperate continental monsoon type, marked by frequent sunny conditions, abundant sunshine, and obvious seasonal changes, with a yearly average temperature that hovers around 4.2 °C. Yearly precipitation amounts fall within the range of 360 to 366 mm, which is predominantly focused between June and August, constituting over 63% of the yearly precipitation totals. The duration of the frost-free season spans between 90 and 120 days, with the predominant soil type being sandy chestnut–calcic.

Thanks to its unique geographical and climatic conditions, Wuchuan County has become a major producer of highland specialty crops, especially potatoes. Potato cultivation occupies a central position in the county's agriculture and has a significant impact on promoting the rural revitalization strategy, facilitating economic development and improving people's livelihoods. Therefore, accurate identification of potato-growing areas in

Wuchuan County is important for optimizing agricultural production management, policy formulation, and food security assessment.

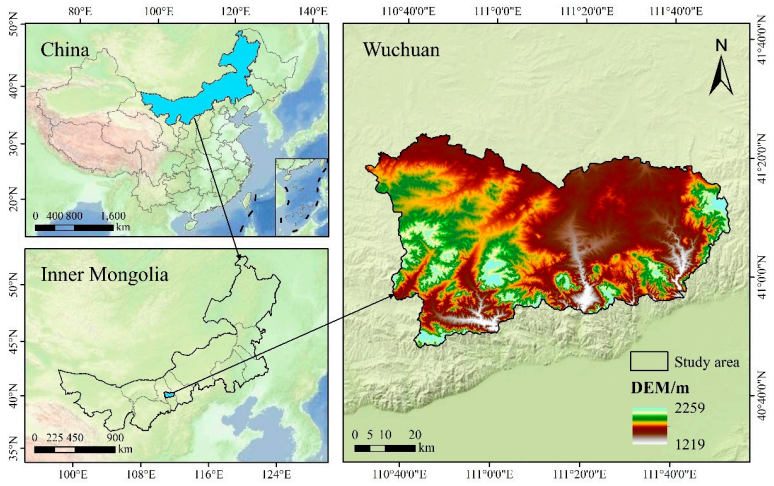

Figure 1. Geographical positioning of the research region.

2.2. Data

2.2.1. Multi-Source Satellite Remote Sensing Image Data

S2 is divided into two satellites, S2A and S2B, each carrying a multispectral imager (MSI). S2A and S2B have an orbital altitude of 786 km and are capable of capturing 13 spectral bands (including 3 red-edge bands) with a width of 290 km. The spatial resolution on the ground measures 10 m, 20 m, and 60 m. Each satellite has an individual revisit cycle of 10 days; yet, when their coverage is combined, the effective revisit cycle is reduced to 5 days, and the atmospherically and radiometrically corrected S2B dataset can be obtained by accessing the "COPERNICUS/S2_SR_HARMONIZED" dataset on the GEE platform. L8 is equipped with a pair of sensors: the Operational Land Imager (OLI) and the Thermal Infrared Sensor (TIRS), which are labeled in the GEE as "LANDSAT/LC08/C02/T2_L2". Launched in September 2021, L9 represents the latest in the Landsat series, equipped with a second-generation Operational Land Imager (OLI-2) and Thermal Infrared Sensor (TIRS-2). Both Landsat satellites have an orbital altitude of 705 km and a revisit period of 16 days, which can be shortened to 8 days if they are used jointly, and the data are identified in GEE as "LANDSAT/LC09/C02/T1_L2". Table 1 provides a compilation of the spectral bands for the MSI, OLI, and OLI-2 sensors. The comparison shows that these imagers overlap in some spectral bands, which provides a possibility for data integration [26,27].

Table 1. Introduction of the three spectral bands of the imager.

L8 (OLI)/L9 (OLI-2)			S2 (MSI)		
Bands	**Wavelength (μm)**	**Resolution (m)**	**Bands**	**Wavelength (μm)**	**Resolution (m)**
Coastal (Band1)	0.430–0.450	30	Coastal (Band1)	0.433–0.453	60
Blue (Band2)	0.450–0.510	30	Blue (Band2)	0.458–0.523	10
Green (Band3)	0.533–0.590	30	Green (Band3)	0.543–0.578	10
Red (Band4)	0.636–0.673	30	Red (Band4)	0.650–0.680	10
	-		Red Edge (Band5)	0.698–0.713	20
	-		Red Edge (Band6)	0.733–0.748	20
	-		Red Edge (Band7)	0.773–0.793	20

Table 1. *Cont.*

L8 (OLI)/L9 (OLI-2)			S2 (MSI)		
Bands	**Wavelength (μm)**	**Resolution (m)**	**Bands**	**Wavelength (μm)**	**Resolution (m)**
NIR (Band5)	0.851–0.879	30	NIR (Band8)	0.785–0.900	10
	-		Red Edge (Band8a)	0.855–0.875	20
	-		Water Vapor (Band9)	0.935–0.955	60
Cirrus (Band9)	1.360–1.380	30	Cirrus (Band10)	1.360–1.390	60
SWIR-1 (Band6)	1.566–1.651	30	SWIR-1 (Band11)	1.565–1.655	20
SWIR-2 (Band7)	2.107–2.294	30	SWIR-2 (Band12)	2.100–2.280	20
PAN (Band8)	0.500–0.680	15		-	

2.2.2. Sample Data

The dataset employed in this research was divided into two main parts: data from field surveys and data obtained after visual interpretation and labeling based on high-resolution Google historical images. The field survey data originated from a field survey in Wuchuan County throughout the period of potato cultivation, starting May–September 2023. With the aim of complying with the principles of accuracy and scientificity of the collected samples, the study selected large areas (greater than 50 acres) with uniform growth of each type of crop in individually planted areas and tried to cover the whole county as much as possible. The collected sample data predominantly documented the types of crops, along with the GPS-derived latitude and longitude coordinates of the central points of the extensive fields. The total number of sampling points for various crops was 593 (including 268 for potatoes and 325 for other crops). Field collection images of field potato data in June, July, and August are shown in Figure 2.

Figure 2. Wuchuan County field potatoes in June, July, and August.

2.2.3. Mask Data

For the purposes of this research, the foundational land cover map for Wuchuan County was sourced from the outcomes generated by the research group led by Professor Yang Jie and Huang Xin at Wuhan University. The data download can be found at https:

//zenodo.org/records/12779975, accessed on 10 September 2024. The results utilized 335,709 views of Landsat data on the GEE platform to produce the China Land Cover Dataset (CLCD), which started in 1985 and has been updated to 2023. The CLCD contains nine different land use types, including farmland, shrub, forest, watershed, grassland, snow and ice, bare ground, impervious surface, and wetland. The snow and ice land cover category is likely to be extremely limited in this region, given the high temperatures and low snowfall potential in Wuchuan County during the potato-growing season. This study used 2023 30 m resolution land cover raster data with non-farmland types within Wuchuan County, removed through a masking process.

2.3. Methods

The research introduces an approach aimed at potato crop cultivation area extraction based on ML model and multi-source medium- and high-resolution satellite images. The approach is structured into three primary stages (Figure 3): The first stage is to select three satellite remote sensing images obtained during the potato fertility period from May 2023 to October 2023 for preprocessing, respectively; then, the linear coefficients are calculated to unify and harmonize the surface reflectance values of different satellite remote sensing data to construct a time–series dataset of remote sensing images at 10-day intervals. The second stage is grounded in the feature extraction of the time–series dataset, and five different input feature sets are constructed for assessing the impact of varying feature combinations on classification precision. The third stage is the assessment of potato and non-potato crop classification accuracies based on the feature extraction algorithm and machine learning classification model. The fourth stage is to conduct the pixel potato planting area calculation to obtain an distribution map of the potato planting area in Wuchuan County.

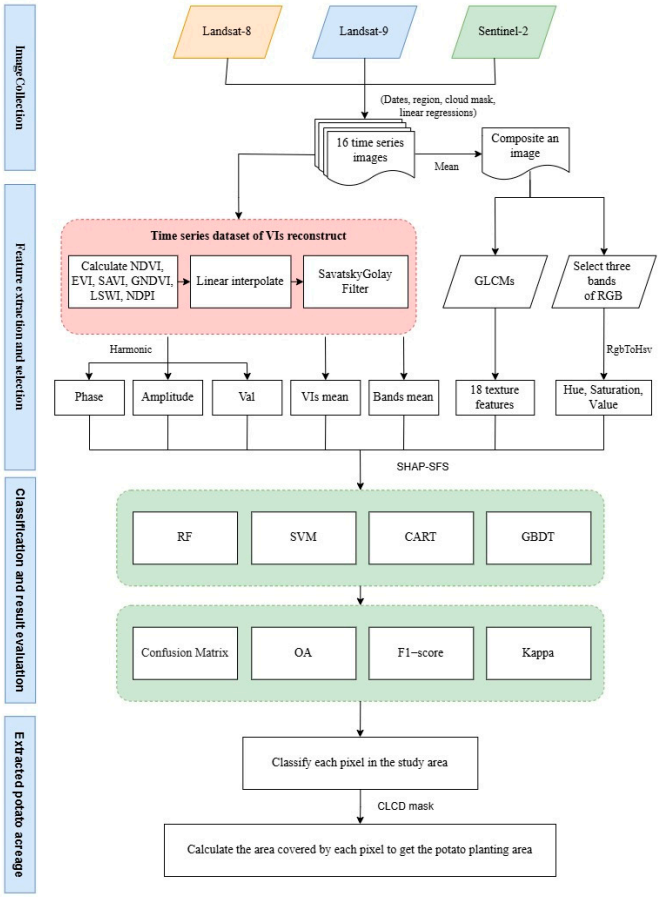

Figure 3. The technical roadmap of this study.

2.3.1. Calculation of the L9 Coordination Factor with S2 and L8

Between 1 May 2023 and 30 September 2023, covering the entire study area, 3-view images from the three satellites with the closest temporal separation were selected and stacked together, and GEE performed nearest-neighbor resampling by default during reprojection. Comparative cross-sensor analysis was performed by randomly sampling a large number of pixels within the boundaries of the research locale. The sampling approach was grounded in a stratified sampling method for the eight land use types in the 2023 CLCD data, except for the snow and ice category, to obtain a sample that is characteristic of each category. A cap of 5000 was established for the maximum random sampling of pixel points within each category. Subsequent filtration removed any null samples, culminating in a final count of 29,755 valid pixels in the sample set. Six similar band reflectance values were extracted for each pixel. The linear regression equations involving S2 and L9, and L8 and L9, were derived using ordinary least squares (OLS) methodology, with subsequent calculations of the equations' slopes and intercepts as detailed in Table 2.

Table 2. Slope and intercept of linear regression equations for similar bands of MSI to OLI-2 and OLI to OLI-2.

Bands	S2 (MSI) and L9 (OLI-2)				L8 (OLI) and L9 (OLI-2)			
	Slope	Intercept	R^2	RMSE	Slope	Intercept	R^2	RMSE
Blue	0.6864	−0.0007	0.8348	0.0145	1.0444	−0.0014	0.9008	0.0112
Green	0.8441	0.0025	0.8861	0.0159	1.0492	−0.0038	0.9387	0.0117
Red	0.8637	0.0044	0.9179	0.0187	1.032	−0.0003	0.9503	0.0146
NIR	0.9133	0.0131	0.8862	0.0356	0.9595	0.0132	0.943	0.0252
SWIR-1	0.923	−0.0055	0.9	0.0302	1.0084	0.0049	0.9333	0.0245
SWIR-2	0.8611	−0.0011	0.9245	0.0268	1.0164	0.0047	0.9377	0.0243

When comparing the data from the OLI and OLI-2 sensors, a high degree of similarity and small differences are clearly observed. Specifically, the model's fitting efficacy is particularly pronounced within the red spectral band. All band-specific fitting coefficients surpass the threshold of 0.9, while the root mean square error (RMSE) is maintained below 0.252, indicating a good fit. Despite the relatively large differences between MSI and OLI-2, the fitting coefficients of the two remain above 0.8, showing a strong correlation. In the fitting of S2 and L9, the linear regression of the SWIR-2 band showed the highest significance, which may imply that this band has an important reference value in the comparative analysis of data between the two sensors.

2.3.2. Preprocessing of Satellite Remote Sensing Images

For the remote sensing image analysis, the image data covering the study area were first screened by setting a time range (1 May 2023–30 September 2023). For S2 images, images with more than 80% cloud coverage were excluded; for L8 and L9 images, images with more than 50% cloud coverage were excluded. Subsequently, clouds and shadows within the imagery were detected and obscured to enhance the precision of derived surface data. Similar bands were selected and renamed for ease of analysis. Next, the reflectance data of the L8 and S2 images were uniformly converted to the L9 scale using the previously computed conversion coefficients (slope and intercept) to enable comparative analyses between the different datasets. On the GEE platform, most of the image data have already undergone official preprocessing, including steps such as atmospheric correction, radiometric correction and topographic correction. On this basis, further de-clouded preprocessing was carried out on the L8, L9, and S2 images; finally, 125 views of high-quality remote sensing images were obtained. To address the issue pertaining to missing image data po-

tentially arising from the de-clouding process, the image synthesis technique was adopted to reconstruct the complete image in a specific time period. By comparing the 10-day and 15-day synthesis intervals, the 10-day interval was chosen, which not only ensured the completeness of the image data, but also maintained the high-density characteristics of the time–series. Eventually, we constructed a time–series dataset containing 16 views of synthesized images, which provides a solid data base for subsequent feature extraction and classification.

2.3.3. Reconstruction of VI Time–Series

Time–series analysis of VIs is important for revealing crop growth cycles and phenological characteristics [28]. In previous studies, the NDVI is extensively applied within area extraction studies for major grain crops such as rice, maize, and wheat because of their robustness [29–32]. However, NDVI tends to reach saturation at higher vegetation densities, resulting in a decrease in its sensitivity [33]. Since different VIs have their own advantages in practical applications and are affected by factors such as month, sampling site environment, climatic conditions, and field management, single reliance on a particular VI may lead to confusing analysis results. Therefore, in this study, six different VIs were extracted from 16-scene remote sensing images and the corresponding time–series data were constructed. There were significant correlations between these indices and crop physiological parameters, and they covered different spectral bands and influencing factors. The name of each VI, its calculation formula, and its references are detailed in Table 3.

Table 3. The research employed a set of spectral indices.

Abridge	Name	Calculation Formula	Bibliography
NDVI	Normalized Difference Vegetation Index	$\frac{\rho_{NIR}-\rho_{Red}}{\rho_{NIR}+\rho_{Red}}$	[34]
EVI	Enhanced Vegetation Index	$\frac{2.5\times(\rho_{NIR}-\rho_{Red})}{\rho_{NIR}+6\times\rho_{Red}-7.5\times\rho_{Blue}+1}$	[35]
SAVI	Soil-Adjusted Vegetation Index	$\frac{1.5\times(\rho_{NIR}-\rho_{Red})}{\rho_{NIR}+\rho_{Red}+0.5}$	[36]
GNDVI	Green Normalized Difference Vegetation Index	$\frac{\rho_{NIR}-\rho_{Green}}{\rho_{NIR}+\rho_{Green}}$	[37]
LSWI	Land Surface Water Index	$\frac{\rho_{NIR}-\rho_{SWIR}}{\rho_{NIR}+\rho_{SWIR}}$	[38]
NDPI	Normalized Difference Phenology Index	$\frac{\rho_{NIR}-(0.74\times\rho_{Red}+0.26\times\rho_{SWIR})}{\rho_{NIR}+(0.74\times\rho_{Red}+0.26\times\rho_{SWIR})}$	[39]

Although the remotely sensed image dataset integrates images from three satellites, the sequential data of VIs continue to possess discontinuities in time due to the limitations of the revisit period and cloud shading. In addition, influenced by sensor performance, cloud cover and shading, atmospheric sols, and other environmental factors, remotely sensed time–series data often contain noise, which may affect the accuracy of the VIs and thus interfere with data analysis [40]. To solve these problems, Chen et al. [41] devised an innovative approach for reconstructing high-quality Landsat–MODIS NDVI time–series data, which is referred to as the Gap Filling and Savitzky–Golay filtering method (GF-SG). This study borrows from this technique to generate high-quality VI data series using linear interpolation and S-G filtering. The linear interpolation window is set to 60 days, which means that each image will look for images within 60 days from the target image; for the blank pixels in the image, their values will be interpolated using the average of the before and after images. The specific interpolation formula is as follows (1):

$$y = y_1\ (y_2 - y_1\ ) \times ((t - t_1)/(t_2 - t_1)) \quad (1)$$

where y: interpolated image; y_1: pre-interpolation image; y_2: post-interpolation image; t: interpolation date; t_1: pre-interpolation image date; t_2: post-interpolation image date.

Among many time–series data reconstruction techniques, the S-G filtering algorithm is widely favored for its excellent stability and efficiency [42]. The key parameters of this algorithm include the window half-width (m) and polynomial coefficients (d), which have a decisive impact on the filtering performance. If m is assigned a value that is too minimal, it could result in the model becoming overly fitted to the data, thus failing to capture the long-term trend effectively; meanwhile, too large a value of m may ignore critical changes in the time–series. The polynomial coefficient d generally ranges between 2 and 4. Lower values of d tend to produce smoother outputs but may introduce some bias; conversely, higher values of d, while reducing bias, may overfit the data and lead to noisier results. To balance smoothness and accuracy, we chose to set the window half-width m to 4 and the polynomial coefficient d to 3. The reconstruction roadmap is depicted in Figure 4.

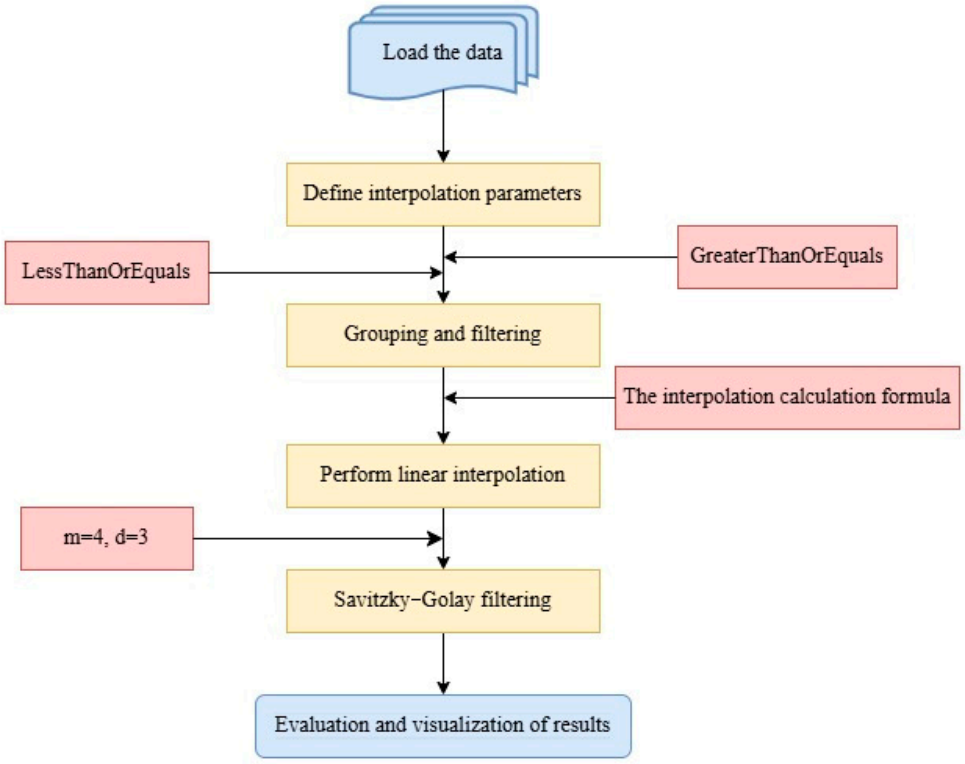

Figure 4. Time–series reconstruction roadmap.

2.3.4. Extraction of VI Time–Series Features Based on HANTS

The importance of deriving potent phenological data regarding potatoes from intricate time–series satellite imagery cannot be overstated [43]. Past studies have analyzed features from both time and frequency domain perspectives. The time domain focuses on the relationship of the signal over time and reveals the trend of crop cover by fitting a linear function. Frequency domain analysis focuses on the frequency characteristics of the signal and analyzes the seasonal phenology of crops by decomposing the time–series, which helps in crop classification and area estimation. The time–series and frequency spectrum are interconverted by Fourier transform. In this research, HANTS was used to decompose VIs time–series data into multiple harmonic components to achieve the extraction of meaningful periodic and trend information from complex time–series [44]. The time–series is decomposed into infinite sine and cosine waves of different frequencies by Equation (2).

$$y(t) = a_0 + \sum_{i=1}^{n} [a_i \cos(2\pi f_i t) + b_i \sin(2\pi f_i t)], t = 1, 2, \cdots, N \quad (2)$$

where a_0 is the average value of the VI time–series; a_i and b_i denote the sine and cosine parameters, respectively; t denotes the number of sampling points; N denotes the length of the time–series; and n denotes the number of harmonics, $2n + 1 \leq N$.

Time variables were added to the computed images of the six VIs, and cosine and sine harmonic terms were computed from the time variables. Multiple linear regression was performed using the harmonic terms (e.g., constant, t, cos, sin) as independent variables and the different VIs as dependent variables, and the phase, amplitude, and harmonic fitted mean of the time–series of the six VIs were calculated using the built-in reducer (Reducer) on the GEE.

2.3.5. Color Space Feature and Texture Feature Extraction

In addition to spectral features, spatial features should not be neglected; color space and texture features are two important concepts in image processing and computer vision. Because converting the color space model and performing texture feature extraction both operate on a single image, the preprocessed set of time–series remote sensing images is synthesized into a single image by calculating the mean value. The color space of an image is the mathematical model or coordinate system that describes the colors in the image. Common color spaces include RGB, HSV, and HSI. Conversion from RGB to HSV is performed by selecting the three bands of RGB from the synthesized image data from the loaded image data. In GEE, this is usually performed by image.select('red_mean', 'green_mean', 'blue_mean'), where "red_mean" stands for the red band average of the time–series image, "green_mean" stands for the green band average, and "blue_mean" stands for the blue band mean. Using the ee.Image.rgbToHsv() method of the GEE platform to convert the RGB image to HSV image, three new bands were generated: hue, saturation, and luminance. The values of these bands were located in the floating-point type [0, 1] range, which were used as the features for the subsequent studies.

If the Gray-Level Co-occurrence Matrix (GLCM) and its corresponding texture features are individually calculated for each band within an image on the GEE platform, the computation will produce huge texture data bands, which are beyond the computational memory of GEE. To avoid this phenomenon, the grayscale layer is computed before the texture features are computed, and the grayscale maps are synthesized according to Equation (3).

$$\text{gray} = (0.3 \times \text{Red}) + (0.59 \times \text{Green}) + (0.11 \times \text{Blue}) \quad (3)$$

After calculating the grayscale layer, you can directly call the .glcmTexture() function, and the dimension of the textural window is configured as 4; then, you obtain an image containing 18 texture feature bands.

In this study, five types of input feature sets were constructed to assess the importance of VI temporal features, texture features, and color spatial components for potato crop classification: (1) phase, amplitude, and mean values after harmonic fitting of 6 VIs, totaling 18 features (VIsTC); (2) VI mean values of 16 views of remotely sensed imagery (VIsmean); (3) band mean values of 16 views of remotely sensed imagery (Bandsmean); (4) 18 texture features based on the synthetic imagery (TFs); (5) 3 color spatial component features (HSV). There is then a total of 51 input features.

2.3.6. Selection of Models and Assessment of Their Performance

The training and test datasets are first extracted from the final image (final_img) containing all the features. The sampleRegions() function is used to extract the sample data from the specified bands; then, the null sample data is filtered out; subsequently, the dataset is divided into training and validation subsets with an 8:2 proportion. Next, a new category label is set, and each feature is buffered by 20 m, which is equivalent to expanding two

pixels outward; this both enlarges the sample area and ensures that, in the training and test datasets, it captures sufficient spatial information, thereby enhancing the model's predictive accuracy. Finally, the split data are merged into the global training and test set feature sets. After constructing the training and test datasets, four classifiers, namely Random Forest (RF), Support Vector Machine (SVM), Classification and Regression Tree (CART), and Gradient-Boosted Decision Tree (GBDT), were selected to classify potato and non-potato crops. On the GEE platform, the "ee.Classifier.smile" function can be called directly to create different classifier objects. RF, CART, and GBDT are all based on decision trees, so the number of decision trees is the most important parameter in this type of algorithm, and it can be debugged through several tests. The remaining parameters within the classification model are typically assigned their default settings, with the comprehensive feature set being utilized for iterative testing within the model. The number of RF decision trees is determined to be 110, the maximum depth of decision trees in CART is 50, and the number of GBDT decision trees is 60. SVM is also a powerful classification method that is especially suitable for classification problems in high-dimensional spaces. Applying the Radial Basis Function (RBF) to the SVM classifier, the gamma is set, where the dataset feature number is 51.

2.3.7. Feature Screening

The process of feature selection holds significant importance within the realm of ML. High-dimensional raw feature datasets often contain redundant and irrelevant features that reduce computational efficiency and model accuracy. SHAP (SHapley Additive exPlanations) is a method grounded in game theory designed to interpret the outputs of any ML model [45]. In this study, SHAP values are calculated for measuring how features affect the dependent variable, which can be used to calculate the marginal contribution value of each feature and measure the importance of the features in the feature selection task. The 51 features were exported from GEE and implemented on Python 3.9 software to construct a binary classification model, and the ML test was used to obtain the feature contribution using SHAP to interpret the model, because SHAP can only provide the contribution ranking but cannot determine the optimal feature subset dimensions, so the optimal feature subset was determined based on this and combined with the Sequential Forward Selection (SFS) method. The first feature is selected as the independently optimal feature, and the second feature is ranked second in terms of feature contribution in combination with the first feature; one feature is added at a time, and a curve is generated to determine the optimal feature subset dimension based on the overall accuracy change.

3. Results and Analysis

3.1. Reconstructed Time–Series of VIs and Analysis of Potato Phenological Period

Two satellite images were required to achieve full coverage of the Wuchuan County area. S2, L8, and L9 satellite images with less than 50% cloudiness were screened between May and October 2023 to synthesize complete Wuchuan County images. A total of 21 complete S2 images, 7 L8 images, and 8 L9 images were retrieved, with few individual satellite remote sensing image data and some of them had severe cloud pollution. After analyzing the sensors carried by the three satellites, S2, L8 and L9, it was found that they have similar bands, and that the band fitting coefficients are above 8.3; the images from the three satellites are integrated. This strategy not only enhances the spatial and temporal resolution of the data, but also minimizes the impact of cloud pollution on the data quality, thus improving the accuracy and reliability of the monitoring results. After the de-clouded preprocessing of the integrated images, which resulted in a serious lack of pixel information, we used linear interpolation and filtering techniques for the reconstruction of the VI

time–series. As shown in Figure 5, taking the image with severe missing pixel information after preprocessing on 30 June as an example, the effect was significantly improved after the reconstruction process. By this method, we successfully filled in the blank pixel values in the image and reduced the noise interference, which not only improved the overall quality of the image but also increased the information content. These improvements laid the foundation for the construction of VIs time–series, ensuring the accuracy and reliability of the subsequent analysis.

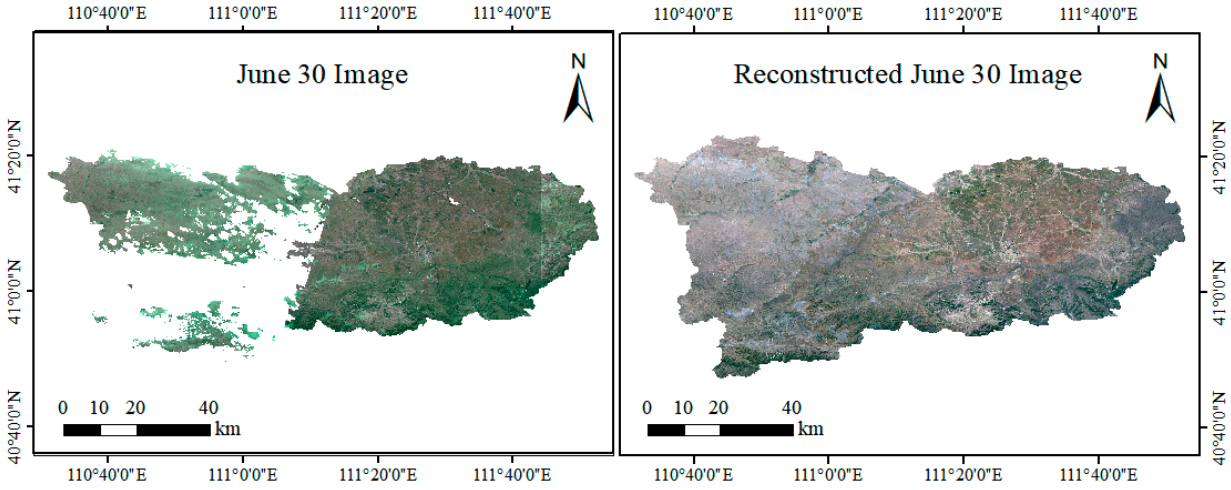

Figure 5. Comparison of satellite image reconstruction before and after (for example, on 30 June).

Utilizing the spatial coordinates from the potato sampling locales, VI data were derived from both the pristine imagery and from each perspective image post the application of linear interpolation and S-G filtering. This process resulted in the generation of six time–series curves of VIs, illustrating the temporal dynamics of these indices throughout the potato cultivation period, as presented in Figure 6.

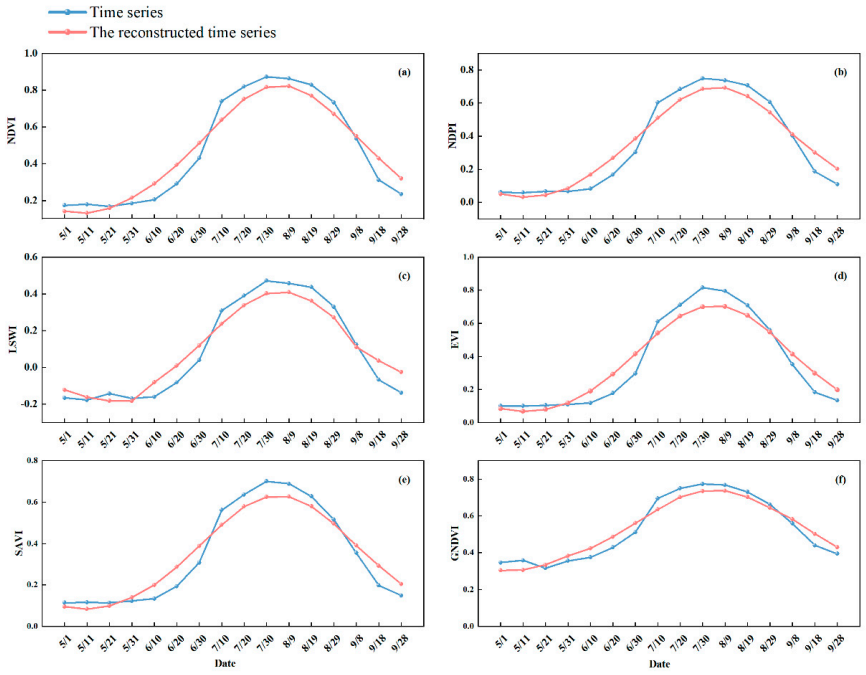

Figure 6. Time–series curves of the original and reconstructed VIs. (**a**) NDVI time–series; (**b**) EVI time–series; (**c**) SAVI time–series; (**d**) RVI time–series; (**e**) MSAVI time–series; (**f**) GNDVI time–series.

The pattern of the curves' average values corresponded to the phenological trends observed for potatoes within the research region. Combined with the analysis of the general spectral characteristics of green vegetation, in early May, the time node of the minimum value of NDVI represented the sowing period of potato. Potato in Wuchuan County began to be sown around May 11, and was in the seedling stage in early June. With the increase in vegetation cover, NDVI continued to rise, and from July to August, potato reached a peak of 0.822 in the NDVI value about 90 days after planting. This indicated that, from late August to early September, the potatoes entered the ripening period, and the NDVI values gradually decreased. The potatoes were harvested in most areas of Wuchuan County from mid–late-September. The trend of the NDPI was similar to that of the NDVI, but the peak appeared a little later, which might reflect some kind of correlation or hysteresis regarding the relationship between short-wave infrared reflectance and vegetative growth. The values of LSWI were negative from May to June, and the absolute value increases gradually, which may indicate that there are fewer surface water bodies or that the reflectance characteristics of the water bodies are weaker in this period. The trend of EVI is similar to that of NDVI, but the value is relatively higher, which may be due to the fact that EVI takes into account the factors of soil background and atmospheric influence, which makes it more accurate in assessing the vegetation cover. The trend of SAVI is similar to that of NDVI and EVI, but the value is slightly lower, which may be due to the fact that SAVI takes into account the soil brightness and atmospheric influence, which may be the result of a delay in the vegetation growth. Due to the fact that SAVI accounts for the impact of soil luminance on the VI, the trend of GNDVI is similar to those of the other VIs, but the value is relatively higher. This is probably due to the fact that GNDVI focuses more on the assessment of green vegetation. The imputed dataset addressed gaps within the original data, resulting in a trend of the VI that closely mirrored the original dataset's pattern, with the locations of the peaks and valleys largely unchanged. The variability and standard deviation for the regenerated NDVI time–series were approximately 0.1 and 0.19 less than those of the initial time–series, whereas the correlation coefficient with the original series reached 0.95. The findings suggest that the image reconstruction technique employed in this research successfully addresses image gaps, substantially mitigating noise interference and enhancing data integrity.

3.2. Optimal Classification Model Determination and Classification Accuracy Assessment for Different Input Feature Sets

The full-featured data were input into the four classification models to obtain the confusion matrix and two evaluation metrics (Figure 7). It is easy to see from the metrics evaluation that the RF model has the best performance with OA above 0.9 and the highest kappa coefficient being found for both RF and CART. GBDT performance is followed by the SVM model with the worst performance. The next different feature datasets inputs and feature selection are based on the best model RF. RF, CART, and GBDT are all built on a decision tree foundation, and their model performance is usually superior compared to SVM. In comparison with CART and GBDT, RF utilizes an integrated learning strategy that significantly reduces the risk of overfitting and enhances the generalization ability of the model by constructing and fusing the prediction results from multiple decision trees. In the crop classification task, RF exhibits high accuracy in line with the existing findings in the literature [46,47]. In addition, RF can effectively handle common problems such as high-dimensional features and missing values, and shows strong robustness to noise and irrelevant features.

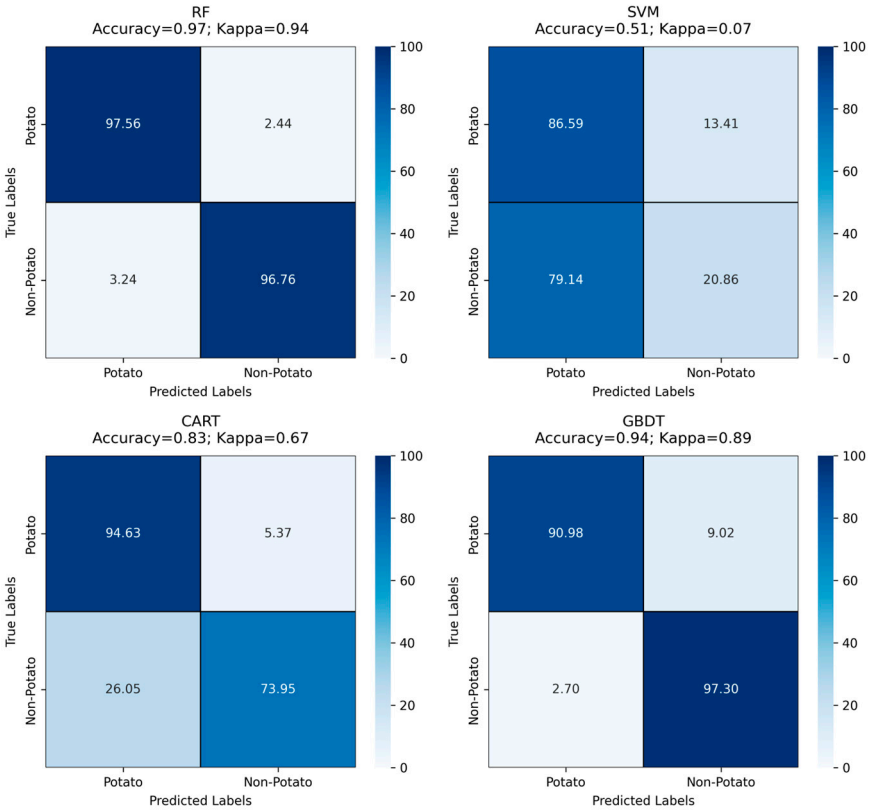

Figure 7. Confusion matrix for four models based on full features.

Within the scope of this research, a comparison of the five feature sets' accuracy was conducted utilizing the RF model (Figure 8), where the overall accuracy OA of VIsCT was 0.958. This is a very high accuracy, indicating that the VI time–series feature performs very well in potato crop classification. This is followed by the band reflectance mean and VI mean. Spectral features are the most effective features in the classification task compared to image features.

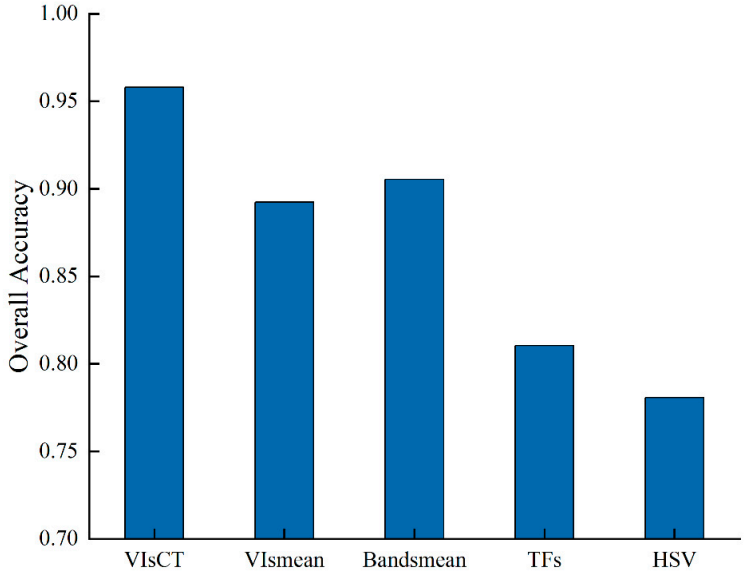

Figure 8. The overall accuracy of the 5 input feature sets.

3.3. Ranking of Feature Contribution and Determination of Optimal Feature Subset Based on SHAP Value

Through the model evaluation, the RF model performs the best, and the method of SHAP-SFS feature selection is combined with the RF classifier to filter the optimal feature subset, which is ranked based on the SHAP value for the feature contribution (Figure 9).

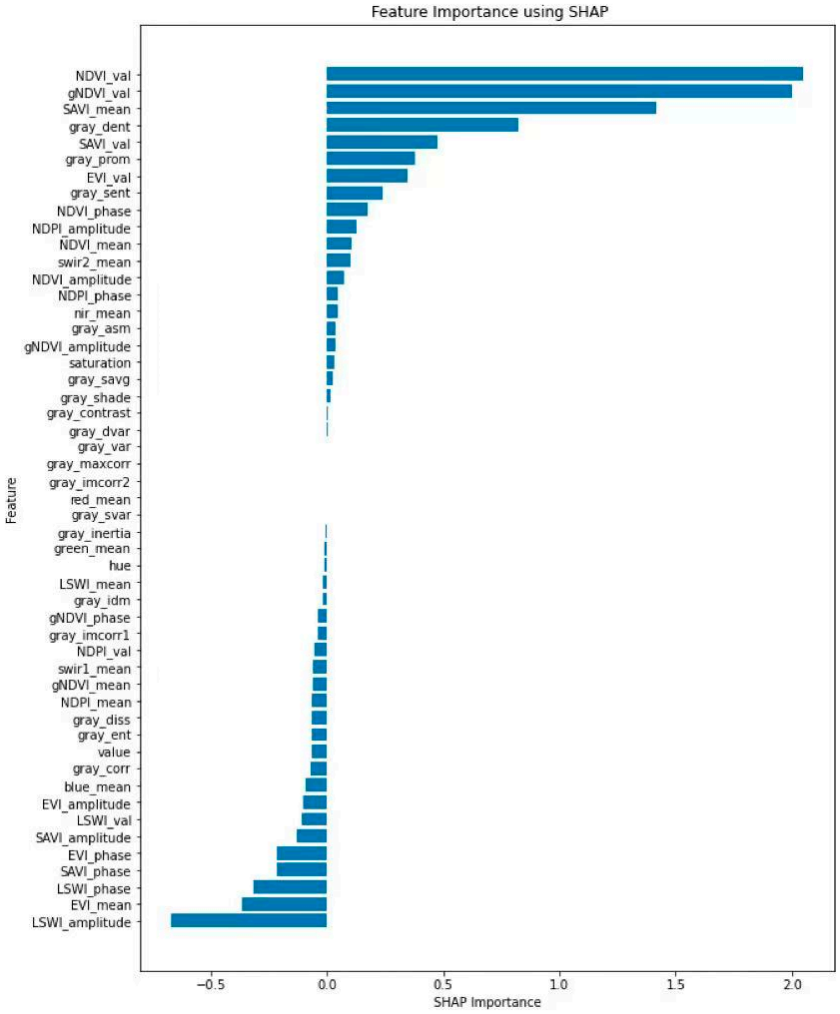

Figure 9. Ranking of feature importance based on SHAP values (SHAP values are shown on the scale of 1×10^{-15}).

As shown in Figure 9, the top five features in terms of importance are the mean of the NDVI harmonic fit, the mean of the GNDVI harmonic fit, the mean of SAVI, the differential entropy, and the mean of the SAVI harmonic fit, respectively. Only the SHAP of saturation is positive in the color spatial features. The contribution of the LSWI time–series features in the VIs was relatively low in potato classification, indicating that the terrestrial water body indices of potato and other crops did not differ much. The NDVI time–series features were the VIs with the highest contributions. It can be proved that NDVI time–series can recognize potato crops well. All other VIs make contributions. This provides a good index for crop classification. The top five texture feature contributions are differential entropy, cluster prominence, sum entropy, angular second moments, and sum mean. Based on the feature importance ranking, the optimal feature subset dimension was determined by combining SFS; according to the curve describing the overall accuracy change (Figure 10), the results indicated that the SHAP-RF model achieved peak accuracy with a feature dimension of 41. By inputting the optimal feature subset into RF, OA, kappa coefficient, and F1-score were improved; so, in this study, the SHAP-SFS-RF model was selected to extract potato planting

areas in Wuchuan County. The SHAP value, as a model prediction interpretation tool, fuses global and local feature importance assessment to accurately identify the features that have a significant impact on the prediction results. Applying the SHAP-SFS technique enables the retention of attributes that are critical in modeling prediction by recursively screening out those features that have less impact. This process not only optimizes the size of the feature set but also excels in maintaining or improving model accuracy.

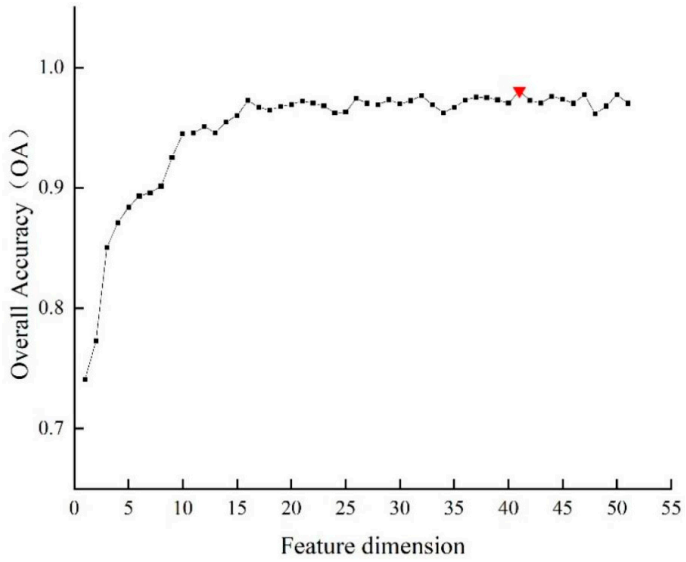

Figure 10. The correlation linking the model's classification precision to the quantity of input features (red triangles represent the best feature dimensions).

As shown in Table 4, the accuracy of the original RF model is 0.97, indicating that the model performs well in correctly classifying the samples. The accuracy of the model (SHAP-SFS-RF) after SHAP-SFS feature selection was improved to 0.98, showing the positive effect of feature selection in improving the model performance. The F1 score of the SHAP-SFS-RF model was improved to 0.98, which further proves the optimizing effect of the feature selection on the performance of the model. The increase in the kappa coefficient from 0.94 to 0.96 may mean that the model becomes more accurate and reliable in distinguishing different categories, especially in the extraction of potato planting areas. Improvement could mean that the model became more accurate and reliable in distinguishing between potato and non-potato crops, especially in the extraction of potato-growing areas. This may be due to the fact that the feature selection process helped the model to reduce the effect of noise and irrelevant information, thus improving the accuracy of the classification.

Table 4. Model evaluation before and after feature selection.

Indicators	RF	SHAP-SFS-RF
Accuracy	0.97	0.98
F1-score	0.97	0.98
Kappa	0.94	0.96

3.4. Extracted Potato Area Compared to Statistics

Initially, each pixel in the synthetic image was categorized into two groups: potato and non-potato crops. This resulted in the creation of mask images for both categories; the area covered by each pixel (in square kilometers) was calculated, and the vector map of Wuchuan County was used for masking; then, the CLCD dataset was used to mask the non-agricultural data in the image to generate the highlighted area. The highlighted area was merged regionally. Masking was performed to generate highlighted areas, and area

merging was performed on the area images of the highlighted areas. The areas of potato and non-potato crops were calculated using the summation reducer ee.Reducer.sum, possessing a spatial resolution of 10 m. The maximum number of pixels processed was 1×10^{13}, and the tile scaling ratio was 16 times, to obtain the spatial distribution map of potatoes in Wuchuan County (Figure 11). The distribution of potato cultivation in Wuchuan County is greatly influenced by topography. The eastern, southern, and western parts of Wuchuan County are surrounded by mountain ranges on three sides; here, the mountainous sector makes up 47% of the total area, and the terrain is gradually low and slow from the south to the north. The soil is fertile and well-drained, which provides high-quality soil conditions for the growth of potatoes. From Figure 11, we can also see that potato cultivation is also mainly concentrated in the northern lower altitude areas. In addition to topographic factors, crop distribution is also affected by temperature, precipitation, irrigation, and other factors. Follow-up studies can analyze other factors affecting potato distribution.

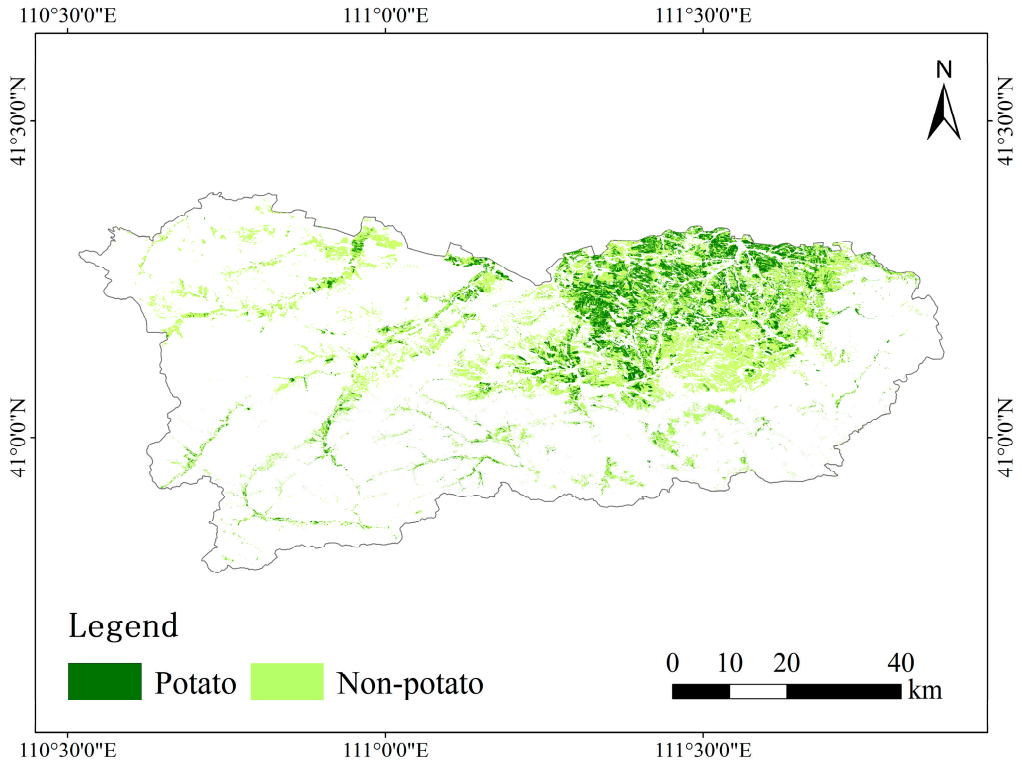

Figure 11. Spatial distribution map of potatoes in Wuchuan County.

To validate the accuracy of the identified potato planting area in Wuchuan County, the calculated potato planting area was compared with the data in the 2023 Hohhot City Statistical Yearbook. The categorized area of potato was 244.043 km^2, i.e., 24,404.3 ha, and the statistical area was 23,302 ha. The results of the comparison with the statistical area showed high consistency with the statistical area. The relative error was less than 5%, and the classification results, compared with the statistical data, demonstrated reliable accuracy, while the extraction results were consistent and trustworthy. To confirm the model's adaptability, the area extraction was conducted in the principal potato cultivation region of Chahar Right Rear Banner in Ulanqab City. Without the need for sampling, the model determined the area to be 111.286 km^2, equivalent to 11,128.6 ha, with the statistical data showing 10,677 ha. This yields a relative error of 4.6%. When juxtaposed with the Ulanqab City Statistical Yearbook, the model's extraction outcomes demonstrate a dependable level of precision, affirming the model's transferability.

4. Discussion

The revisit intervals of various satellites, the spatial resolution of onboard sensors, and cloud cover in the study area all influence the development of time–series data. In this paper, data from different satellites are converted to the same reflectance standard to facilitate the construction of time–series data. This approach improves the availability and analysis of remote sensing time–series data; satellite fusion also uses mathematical algorithms to merge data from different satellites so as to increase the resolution, coverage, or continuity of the data or time–series, in order to obtain higher quality remote sensing images. However, satellite fusion requires a large amount of computational resources and time, which may increase the cost of data processing. The uniform reflectance method is simple to operate and facilitates rapid data processing and analysis, but in some extreme cases, such as cloud cover or complex terrain areas, uniform reflectance may miss some important details and fail to provide sufficient information.

It is especially important to extract effective time–series features about different crops from inside the time–series data with complex features. Previous researchers have explored analyzing the features from different perspectives of time domain and frequency domain. The time domain describes the mathematical function or physical signal versus time and reveals trends in crop cover by analyzing the fitted linear function. The frequency domain is a coordinate system that describes the signal in terms of frequency characteristics. Decomposition of the time–series using spectral analysis helps to analyze the different phenological characteristics of different crops due to seasonality, so as to carry out crop classification and area extraction. The time domain and frequency domain are converted to each other by Fourier transform and inverse Fourier transform. Current time domain feature extraction algorithms are prone to overfitting problems at the peak of crop growth. In this study, the temporal features were also extracted from the frequency domain perspective, demonstrating the applicability of frequency domain temporal features.

The GEE platform is highly portable in many aspects such as data, code, platform, results, and collaboration, enabling users to conduct geospatial data analysis and research more flexibly and efficiently. However, it also has some limitations. Firstly, many processing algorithms of GEE are not transparent, so it is difficult for users to find the root cause of the problem and the solution when they encounter problems. Secondly, the GEE cloud platform itself does not support deep learning model training, requiring locally trained models to be uploaded into GEE Assets.

By integrating multi-source satellite data, the spatial and temporal resolution of the data can be effectively improved. In this way, more detailed surface coverage information can be obtained, so as to accurately capture the subtle dynamic changes in the potato planting area. In the process of data processing, the fusion of spectral, texture, and color multi-dimensional features, and in-depth feature selection is conducted; this will significantly improve the accuracy of the classification model, so that the extraction of potato planting areas is more accurate. In addition, the combination of meteorological data, soil data, and field management data lays a solid foundation for the construction of a large-scale potato yield prediction model and market forecasting tools; this helps in making scientific and reasonable decisions in agricultural production and market planning. Collecting crop sample data in the field is often time-consuming and labor-intensive, and future research trends may shift towards using smaller datasets to enable crop area estimation. However, variability in crop rotation systems and satellite imagery may make it difficult for models to overcome the effects of inter-annual variability. Future work will utilize transfer learning to validate how the model performs in different regions in different years.

5. Conclusions

This study addresses the critical need for precise potato crop area estimation by developing a high-resolution VI time–series from L8, L9, and S2 satellites. We enhanced the data through linear interpolation and S-G filtering, and employed the HANTS method to extract multi-dimensional features, leading to a robust classification system. The RF model, augmented with SHAP-SFS feature selection, achieved significant accuracy improvements, with a precision rate of 0.97 and a kappa value of 0.94, outperforming previous models. This methodology not only advances potato crop area extraction but also provides a detailed spatial distribution map for Wuchuan County, offering valuable insights for agricultural planning and research.

While this research marks a significant step forward, there is scope for further enhancement. Future work will explore additional machine learning algorithms to refine crop classification accuracy and incorporate more environmental factors and crop growth models for a deeper understanding of agricultural dynamics. The approach will also be generalized to other crops and regions to assess its broader applicability, ensuring the methods developed here can contribute to a wider array of agricultural monitoring and research initiatives.

Author Contributions: Conceptualization, Q.L.; methodology, Q.L.; software, Q.L.; validation, Q.L.; formal analysis, Q.L.; data curation, Q.L.; writing—original draft preparation, Q.L.; writing—review and editing, H.Z. and X.F.; visualization, Q.L.; supervision, X.F. and H.L.; project administration, X.F.; funding acquisition, X.F. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Natural Science Foundation of China (61962047), the Inner Mongolia Autonomous Region Science and Technology Major Special Project (2021ZD0005), the Inner Mongolia Autonomous Region Natural Science Foundation (2024MS06002), the Inner Mongolia Autonomous Region Universities and Colleges Innovative Research Team Program (NMGIRT2313), the Basic Research Business Fund for Inner Mongolia Autonomous Region Directly Affiliated Universities (BR22-14-05), and the Collaborative Innovation Projects between Universities and Institutions in Hohhot (XTCX2023-20, XTCX2023-24).

Institutional Review Board Statement: Not applicable.

Data Availability Statement: If data are needed, interested parties may contact the corresponding author.

Acknowledgments: The authors would like to express their sincere gratitude to the research team led by Professor Yang Jie and Professor Huang Xin from Wuhan University for generously providing the land cover data base map of Wuchuan County, which served as a foundational component of this study. Their valuable contributions and expertise have significantly enhanced the quality and depth of our research. The authors gratefully acknowledge the support provided by the Google Earth Engine platform for data processing and analysis.

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

References

1. Zhang, H.; Xu, F.; Wu, Y.; Hu, H.H.; Dai, X.F. Progress of potato staple food research and industry development in China. *J. Integr. Agric.* **2017**, *16*, 2924–2932. [CrossRef]
2. Zierer, W.; Rüscher, D.; Sonnewald, U.; Sonnewald, S. Tuber and Tuberous Root Development. In *Annual Review of Plant Biology*; Merchant, S.S., Ed.; Annual Reviews: San Mateo, CA, USA, 2021; Volume 72, pp. 551–580.
3. Wang, N.; Reidsma, P.; Pronk, A.A.; de Wit, A.J.W.; van Ittersum, M.K. Can potato add to China's food self-sufficiency? The scope for increasing potato production in China. *Eur. J. Agron.* **2018**, *101*, 20–29. [CrossRef]
4. Devaux, A.; Goffart, J.P.; Kromann, P.; Andrade-Piedra, J.; Polar, V.; Hareau, G. The Potato of the Future: Opportunities and Challenges in Sustainable Agri-food Systems. *Potato Res.* **2021**, *64*, 681–720. [CrossRef]

5. Kuang, X.; Guo, J.; Bai, J.; Geng, H.; Wang, H. Crop-Planting Area Prediction from Multi-Source Gaofen Satellite Images Using a Novel Deep Learning Model: A Case Study of Yangling District. *Remote Sens.* **2023**, *15*, 3792. [CrossRef]
6. Pan, Y.; Li, L.; Zhang, J.; Liang, S.; Zhu, X.; Sulla-Menashe, D. Winter wheat area estimation from MODIS-EVI time series data using the Crop Proportion Phenology Index. *Remote Sens. Environ.* **2012**, *119*, 232–242. [CrossRef]
7. Li, D. Research on remote sensing monitoring of potato planting in reclamation area based on LANDSAT8TM images. *Mod. Agric.* **2018**, *7*, 68–69.
8. Ashourloo, D.; Shahrabi, H.S.; Azadbakht, M.; Rad, A.M.; Aghighi, H.; Radiom, S. A novel method for automatic potato mapping using time series of Sentinel-2 images. *Comput. Electron. Agric.* **2020**, *175*, 105583. [CrossRef]
9. Zhao, L.; Zhang, S.; Sun, Y. Research on area extraction of potato planting area in Shandong Province based on Google Earth engine. *Mod. Agric. Technol.* **2023**, *17*, 72–75, 89.
10. Feng, Y.; Chen, B.; Liu, W.; Xue, X.; Liu, T.; Zhu, L.; Xing, H. Winter Wheat Mapping in Shandong Province of China with Multi-Temporal Sentinel-2 Images. *Appl. Sci.* **2024**, *14*, 3940. [CrossRef]
11. Song, W.; Wang, C.; Dong, T.; Wang, Z.; Wang, C.; Mu, X.; Zhang, H. Hierarchical extraction of cropland boundaries using Sentinel-2 time-series data in fragmented agricultural landscapes. *Comput. Electron. Agric.* **2023**, *212*, 108097. [CrossRef]
12. Li, S.; Xu, L.; Jing, Y.; Yin, H.; Li, X.; Guan, X. High-quality vegetation index product generation: A review of NDVI time series reconstruction techniques. *Int. J. Appl. Earth Obs. Geoinf.* **2021**, *105*, 102640. [CrossRef]
13. Rajadel-Lambistos, C.; Izquierdo-Verdiguier, E.; Moreno-Martínez, A.; Maneta, M.P.; Begueria, S.; Kimball, J.S.; Clinton, N.; Atzberger, C.; Camps-Valls, G.; Running, S.W. Within-season crop monitoring at continental scale utilizing new gap-filled Landsat temporal series. *Int. J. Digit. Earth* **2024**, *17*, 2359577. [CrossRef]
14. Li, J.; Roy, D.P. A Global Analysis of Sentinel-2A, Sentinel-2B and Landsat-8 Data Revisit Intervals and Implications for Terrestrial Monitoring. *Remote Sens.* **2017**, *9*, 902. [CrossRef]
15. Lu, J.; He, T.; Song, D.-X.; Wang, C.-Q. Land Surface Phenology Retrieval through Spectral and Angular Harmonization of Landsat-8, Sentinel-2 and Gaofen-1 Data. *Remote Sens.* **2022**, *14*, 1296. [CrossRef]
16. Trevisiol, F.; Mandanici, E.; Pagliarani, A.; Bitelli, G. Evaluation of Landsat-9 interoperability with Sentinel-2 and Landsat-8 over Europe and local comparison with field surveys. *ISPRS J. Photogramm. Remote Sens.* **2024**, *210*, 55–68. [CrossRef]
17. Tian, Y.; Yang, C.; Huang, W.; Tang, J.; Li, X.; Zhang, Q. Machine learning-based crop recognition from aerial remote sensing imagery. *Front. Earth Sci.* **2021**, *15*, 54–69. [CrossRef]
18. She, B.; Hu, J.; Huang, L.; Zhu, M.; Yin, Q. Mapping Soybean Planting Areas in Regions with Complex Planting Structures Using Machine Learning Models and Chinese GF-6 WFV Data. *Agriculture* **2024**, *14*, 231. [CrossRef]
19. Wei, M.F.; Qiao, B.J.; Zhao, J.H.; Zuo, X.Y. The area extraction of winter wheat in mixed planting area based on Sentinel-2 a remote sensing satellite images. *Int. J. Parallel Emergent Distrib. Syst.* **2020**, *35*, 297–308. [CrossRef]
20. Chen, Y.; Hou, J.; Huang, C.; Zhang, Y.; Li, X. Mapping Maize Area in Heterogeneous Agricultural Landscape with Multi-Temporal Sentinel-1 and Sentinel-2 Images Based on Random Forest. *Remote Sens.* **2021**, *13*, 2988. [CrossRef]
21. Lai, J.; Kang, X.; Lu, X.; Li, S. A review of land observation satellite remote sensing application technology with new generation artificial intelligence. *Natl. Remote Sens. Bull.* **2022**, *26*, 1530–1546. [CrossRef]
22. Saralioglu, E.; Gungor, O. Crowdsourcing in Remote Sensing: A Review of Applications and Future Directions. *Ieee Geosci. Remote Sens. Mag.* **2020**, *8*, 89–110. [CrossRef]
23. Kumar, L.; Mutanga, O. Google Earth Engine Applications Since Inception: Usage, Trends, and Potential. *Remote Sens.* **2018**, *10*, 1509. [CrossRef]
24. Amini, S.; Saber, M.; Rabiei-Dastjerdi, H.; Homayouni, S. Urban Land Use and Land Cover Change Analysis Using Random Forest Classification of Landsat Time Series. *Remote Sens.* **2022**, *14*, 2654. [CrossRef]
25. Zhang, C.; Zhang, H.; Tian, S. Phenology-assisted supervised paddy rice mapping with the Landsat imagery on Google Earth Engine: Experiments in Heilongjiang Province of China from 1990 to 2020. *Comput. Electron. Agric.* **2023**, *212*, 108105. [CrossRef]
26. Vermote, E.; Justice, C.; Claverie, M.; Franch, B. Preliminary analysis of the performance of the Landsat 8/OLI land surface reflectance product. *Remote Sens. Environ.* **2016**, *185*, 46–56. [CrossRef]
27. Main-Knorn, M.; Pflug, B.; Louis, J.M.B.; Debaecker, V.; Müller-Wilm, U.; Gascon, F. Sen2Cor for Sentinel-2. In *Proceedings Volume 10427, Image and Signal Processing for Remote Sensing XXIII*; SPIE: Warsaw, Poland, 2017; Volume 1042704. [CrossRef]
28. Fan, D.; Zhao, X.; Zhu, W.; Zheng, Z. A Review of Factors Influencing the Accuracy of Remote Sensing for Plant Climate Monitoring. *Prog. Geosci.* **2016**, *35*, 304–319. [CrossRef]
29. Yao, Y.; Wu, T.; Li, Y.; Li, Z.; Qian, X.; Zhang, L.; Xing, S.; Zhang, H. Rice field identification based on rice weather parameters and object-oriented algorithms. *J. Agric. Eng.* **2024**, *40*, 150–158.
30. Yu, R.; Liu, X.; Yang, X.; Zhao, B. Research on winter wheat information extraction and irrigation information recognition method based on time series. *China Rural. Water Conserv. Hydropower* **2024**, 68–82. [CrossRef]
31. Yao, F.; Feng, L.; Zhang, J. Corn Area Extraction by the Integration of MODIS-EVI Time Series Data and China's Environment Satellite (HJ-1) Data. *J. Indian Soc. Remote Sens.* **2014**, *42*, 859–867. [CrossRef]

32. Zhang, J.; Feng, L.; Yao, F. Improved maize cultivated area estimation over a large scale combining MODIS–EVI time series data and crop phenological information. *ISPRS J. Photogramm. Remote Sens.* **2014**, *94*, 102–113. [CrossRef]
33. Gu, Y.; Wylie, B.K.; Howard, D.M.; Phuyal, K.P.; Ji, L. NDVI saturation adjustment: A new approach for improving cropland performance estimates in the Greater Platte River Basin, USA. *Ecol. Indic.* **2013**, *30*, 1–6. [CrossRef]
34. Rouse, J.W.; Haas, R.H.; Schell, J.A.; Deering, D.W. Monitoring vegetation systems in the Great Plains with ERTS. *NASA Spec. Publ.* **1974**, *351*, 309.
35. Huete, A.; Didan, K.; Miura, T.; Rodriguez, E.P.; Gao, X.; Ferreira, L.G. Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote Sens. Environ.* **2002**, *83*, 195–213. [CrossRef]
36. Huete, A.R. A soil-adjusted vegetation index (SAVI). *Remote Sens. Environ.* **1988**, *25*, 295–309. [CrossRef]
37. Gitelson, A.A.; Kaufman, Y.J.; Merzlyak, M.N. Use of a green channel in remote sensing of global vegetation from EOS-MODIS. *Remote Sens. Environ.* **1996**, *58*, 289–298. [CrossRef]
38. Jurgens, C. The modified normalized difference vegetation index (mNDVI) a new index to determine frost damages in agriculture based on Landsat TM data. *Int. J. Remote Sens.* **1997**, *18*, 3583–3594. [CrossRef]
39. Xu, D.; Wang, C.; Chen, J.; Shen, M.; Shen, B.; Yan, R.; Li, Z.; Karnieli, A.; Chen, J.; Yan, Y.; et al. The superiority of the normalized difference phenology index (NDPI) for estimating grassland aboveground fresh biomass. *Remote Sens. Environ.* **2021**, *264*, 112578. [CrossRef]
40. Li, R.; Zhang, X.; Liu, B.; Zhang, B. A Review on the Development of Filter Reconstruction Algorithms for Remote Sensing Time Series Data. *J. Remote Sens.* **2009**, *13*, 335–341.
41. Chen, Y.; Cao, R.; Chen, J.; Liu, L.; Matsushita, B. A practical approach to reconstruct high-quality Landsat NDVI time-series data by gap filling and the Savitzky–Golay filter. *ISPRS J. Photogramm. Remote Sens.* **2021**, *180*, 174–190. [CrossRef]
42. Cao, R.; Chen, Y.; Shen, M.; Chen, J.; Zhou, J.; Wang, C.; Yang, W. A simple method to improve the quality of NDVI time-series data by integrating spatiotemporal information with the Savitzky-Golay filter. *Remote Sens. Environ.* **2018**, *217*, 244–257. [CrossRef]
43. Zhong, L.; Hu, L.; Zhou, H. Deep learning based multi-temporal crop classification. *Remote Sens. Environ.* **2019**, *221*, 430–443. [CrossRef]
44. Zhou, J.; Jia, L.; Menenti, M. Reconstruction of global MODIS NDVI time series: Performance of Harmonic Analysis of Time Series (HANTS). *Remote Sens. Environ.* **2015**, *163*, 217–228. [CrossRef]
45. Lundberg, S.M.; Lee, S.-I. A unified approach to interpreting model predictions. In Proceedings of the 31st International Conference on Neural Information Processing Systems, Long Beach, CA, USA, 4–9 December 2017; Curran Associates Inc.: Red Hook, NY, USA, 2017; pp. 4768–4777.
46. Luan, W.; Shen, X.; Fu, Y.; Li, W.; Liu, Q.; Wang, T.; Ma, D. Research on Maize Acreage Extraction and Growth Monitoring Based on a Machine Learning Algorithm and Multi-Source Remote Sensing Data. *Sustainability* **2023**, *15*, 16343. [CrossRef]
47. Mou, H.; Li, H.; Zhou, Y.; Dong, R. Response of Different Band Combinations in Gaofen-6 WFV for Estimating of Regional Maize Straw Resources Based on Random Forest Classification. *Sustainability* **2021**, *13*, 4603. [CrossRef]