Extending the Pairwise Separability Index for Multicrop Identification Using Time-Series MODIS Images

Qiong Hu, Wenbin Wu, Qian Song, Qiangyi Yu, Miao Lu, Peng Yang, Huajun Tang, and Yuqiao Long

***Abstract*—The pairwise separability index (SI) has been demonstrated as an effective indicator for capturing crucial phenological differences between two plant species. However, its application to crop types, which have more obvious phenological characteristics than natural vegetation, has received less attention, and extending the pairwise SI to multiple crops for feature selection still remains a challenge. This paper presented two SI extension approaches ($\mathbf{SI_{ave}}$ and $\mathbf{SI_{min}}$) to select the optimal spectro-temporal features for multiple crops, and investigated their classification performance using Heilongjiang Province, China, as a study area. Feature interpretability and classification accuracy of different crops were evaluated for the two approaches. The results showed that the $\mathbf{SI_{ave}}$ approach generally has relatively high feature interpretability due to its better description of crucial phenological characteristics of different crops. Those crops with high separability are insensitive to the extension approach and have similar classification accuracy for the two approaches, whereas those crops with poor separability show good performance with the $\mathbf{SI_{min}}$ method. Due to the higher temporal autocorrelation, the optimal features for crop classification that are selected by the $\mathbf{SI_{ave}}$ approach exhibit greater information redundancy across the time domain than those that are selected by the $\mathbf{SI_{min}}$ approach, which largely explains the relatively low classification accuracy achieved using the $\mathbf{SI_{ave}}$ approach. These comparison results between $\mathbf{SI_{min}}$ and $\mathbf{SI_{ave}}$ approaches also indicate that time-series images with high temporal resolution do not necessarily produce high classification accuracy, regardless of their ability to describe the seasonal characteristics of crops.**

***Index Terms*—Crop identification, extension approaches, feature interpretability, moderate resolution imaging spectroradiometer (MODIS), separability index (SI).**

I. Introduction

TIMELY and accurate large-scale information about the spatial distributions of crops is important for crop growth monitoring, acreage surveys, yield estimation, and water management [1]–[4]. In addition to being time-consuming and labor intensive, ground observations or visits may generate errors and discrepancies in declarations because of their subjective nature [5], [6]. Newly emerging remote sensing technology has proved to be a more efficient and reliable method to quickly derive large-scale crop distribution information with good temporal consistency [7], [8].

The remote sensing approach for crop identification largely depends on the spectral discrimination capability. However, the intensive management of croplands, together with the frequent modification of farming strategies, increases the variability of the spectral signatures of crops [9], [10]. Moreover, the spectral reflectance of crops is strongly correlated with leaf pigment, nutrients, and structural properties and varies with the growing season [11], [12]. Consequently, different crops in the same region may share similar spectral characteristics, whereas the same crop in different locations may exhibit substantially different spectral signatures, which is challenging for crop identification using remotely sensed images [1], [5]. It is thus critical to better understand the spectral separability of different crop types and how it varies over time to facilitate determining the optimal time period and appropriate spectral features for crop identification [10], [13], [14].

Many studies have investigated the spectral characteristics of vegetation. Somers and Asner [13] proposed the separability index (SI) based on hyperspectral Hyperion data and defined it as the ratio of the interclass endmember variability to the intraclass endmember variability, which was already used to evaluate the spectro-temporal separability of plant species and to select the optimal features for plant species classification. Their SI-based separability analysis determined the optimal features and eventually yielded high classification accuracy with a relatively low cost and low time investment. The successful tests of SI-based feature selection for plant species classification [13], [15], [16] inspired our interest in applying the SI for crop identification, which is currently rare. Furthermore, similarly to the well-known Jeffries–Matusita (JM) distance, the SI is a pairwise measure that is naturally suitable for two-class classification (i.e., native and invasive species). However, how to extend the pairwise SI measure to a global measure of multiple crops (such as rice, corn, soybeans and wheat) remains a key issue. In general, there are two extension strategies that have been widely used for extending JM measure to determine the optimal feature subsets when multiple classes must be considered [17], [18]. The first method is to calculate the average distance and choose the features with the largest average distance as the optimal features [17], [19]. The second is to choose the feature subsets that allow the largest separation between the

Manuscript received February 3, 2016; revised April 21, 2016; accepted May 12, 2016. Date of publication July 28, 2016; date of current version September 27, 2016. This work was supported by the National Administration of Surveying, Mapping and Geoinformation of China under Grant 201512028.

Q. Hu, W. Wu, Q. Yu, M. Lu, P. Yang, H. Tang, and Y. Long are with the Key Laboratory of Agri-informatics, Ministry of Agriculture/Institute of Agricultural Resources and Regional Planning, Chinese Academy of Agricultural Sciences, Beijing 100081, China (e-mail: wuwenbin@caas.cn).

Q. Song is with the Key Laboratory of Agri-informatics, Ministry of Agriculture/Institute of Agricultural Resources and Regional Planning, Chinese Academy of Agricultural Sciences, Beijing 100081, China, and also with the Remote Sensing Technology Center, Heilongjiang Academy of Agriculture Sciences, Harbin 150086, China.

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TGRS.2016.2581210

least separable pair of classes [17], [20]. These two extension strategies should theoretically be suitable for extending the pairwise SI in the context of multiple crops classification for feature selection. However, their performances in feature selection for crop identification have not been explored. It is thus necessary to have a systematic and quantitative evaluation of these two SI extension approaches so as to provide valuable insights for regional multicrop classification.

The objective of this paper is to evaluate the performances of these two extension strategies (for simplicity, these two methods are referred to as $\mathrm{SI_{ave}}$ and $\mathrm{SI_{min}}$) in feature selection for crop identification using time-series MODIS images. The interpretability of the features that are selected using these two methods and the classification accuracy that are achieved using the selected features were used for evaluation. Specifically, the spectro-temporal $\mathrm{SI_{ave}}$ and $\mathrm{SI_{min}}$ matrices were calculated based on each extension strategy and were presented visually to qualitatively examine the feature interpretability. Four feature subsets of different sizes were then selected for each crop type based on the estimated spectro-temporal $\mathrm{SI_{ave}}$ and $\mathrm{SI_{min}}$ charts, and the associated classification accuracy were assessed for these two extension strategies. Finally, the temporal autocorrelation of each time-series vegetation index (VI) that most likely caused the difference in the classification accuracy of these two extension strategies was investigated.

II. Two Approaches for Extending the Pairwise SI to Multiclass Cases

A. Pairwise SI

The pairwise SI was originally used to investigate the spectral variability within (Δintra) and between (Δinter) plant species using hyperspectral images [13]. The SI between a pair of class-specific separations is defined as follows [13], [16]:

$$\mathrm{SI_{ij}}(m,n) = \frac{\Delta\mathrm{inter}(i,j)}{\Delta\mathrm{intra}(i,j)} = \frac{|\overline{u_i} - \overline{u_j}|}{1.96*(\sigma_i + \sigma_j)} \quad (m = \text{Spectrum } 1, \text{Spectrum } 2, \ldots, \text{Spectrum } m,\ n = \text{Time } 1, \text{Time } 2, \ldots, \text{Time } n) \tag{1}$$

where $\overline{u_i}$ and $\overline{u_j}$ are the mean spectral values of spectral image m at time period n for endmember class i (e.g., rice) and endmember class j (e.g., corn), respectively, and σ_i and σ_j are the standard deviations of classes i and j, respectively. $|\overline{u_i} - \overline{u_j}|$ reflects the interclass variability, and $(\sigma_i + \sigma_j)$ represents the intraclass variability. A high interclass variability and low intraclass variability result in a large SI value, which is expected to generate good classification results. The pairwise SI can be applied to each spectro-temporal feature $F(m,n)$, which is composed of the m spectral image (spectral band or VI) $\{\text{Spectrum } 1, \text{Spectrum } 2, \ldots, \text{Spectrum } m\}$ measured at n time period $\{\text{Time } 1, \text{Time } 2, \ldots, \text{Time } n\}$. The SI need not assume a normal distribution for targeted measurements and has low limitations for its dynamic range, which allows it to be more broadly suitable than the JM distance [14] and more suitable to rank the separability of spectral and temporal features for a thorough understanding of whether a feature is or is not important in identifying a given land cover class.

B. SI_{ave} and SI_{min} Extension Approaches

As discussed earlier, the two extension strategies that have been widely used to extend the JM distance are theoretically suitable for extending the pairwise SI for multiclass cases. $\mathrm{SI_{ave}}$ and $\mathrm{SI_{min}}$ can be calculated using the following, respectively:

$$\mathrm{SI_{ave}}(m,n) = 2\sum_{i=1}^{C'}\sum_{j>i}^{C'} \rho_i * \rho_j * \mathrm{SI_{ij}} \quad (m = \text{Spectrum } 1, \text{Spectrum } 2, \ldots, \text{Spectrum } m;\ n = \text{Time } 1, \text{Time } 2, \ldots, \text{Time } n) \tag{2}$$

$$\mathrm{SI_{min}}(m,n) = \min_{i=1\&j>1}^{C'} \mathrm{SI_{ij}} \quad (m = \text{Spectrum } 1, \text{Spectrum } 2, \ldots, \text{Spectrum } m;\ n = \text{Time } 1, \text{Time } 2, \ldots, \text{Time } n). \tag{3}$$

$\mathrm{SI_{ave}}$ calculates the average of all two-class SIs, whereas $\mathrm{SI_{min}}$ considers the minimum value over all pairwise SIs, which allows the best separation between the least separable pair of classes. m is the spectral image, n is the time period, $\mathrm{SI_{ij}}$ is the pairwise SI that is calculated using (1), C' is the number of classes that are considered, and ρ_i and ρ_j are the *a priori* class probabilities [17]. Similarly to the pairwise SI, $\mathrm{SI_{ave}}$ and $\mathrm{SI_{min}}$ can be applied to all features $F(m,n)$ that are composed of the m spectral image and measured at n time period.

III. Experimental Design

This paper is focused on assessing the two SI extension approaches ($\mathrm{SI_{ave}}$ and $\mathrm{SI_{min}}$) for multiple-crop identification. To do so, time-series MODIS images were selected for a case study in Heilongjiang Province, China. Two important evaluation criteria were used: evaluation in terms of the feature interpretability by $\mathrm{SI_{ave}}$ and $\mathrm{SI_{min}}$ and evaluation in terms of the classification accuracy of multiple crops using the selected features. The former considers whether the features that are derived from the $\mathrm{SI_{ave}}$ and $\mathrm{SI_{min}}$ methods capture the characteristics of the crop growing status, whereas the latter considers whether the crop classification using the selected features generates satisfactory classification accuracy. The flow diagram in Fig. 1 outlines the main processing and analysis steps, which are described in detail in the following.

A. Study Area

Heilongjiang Province, which is located in Northeastern China, was selected as the study area (see Fig. 2). Heilongjiang is an agriculturally dominated province with a cropland area of 11.87 million ha (approximately 67% of its total land area) that accounts for approximately 22% of the nation's grain production. It has a long cold winter and a short summer, with an average annual temperature of -4 °C to 4 °C and 2200–2900 hours of sunshine. Annual precipitation averages to 380–600 mm, which is mostly concentrated in June to August [21]. The major

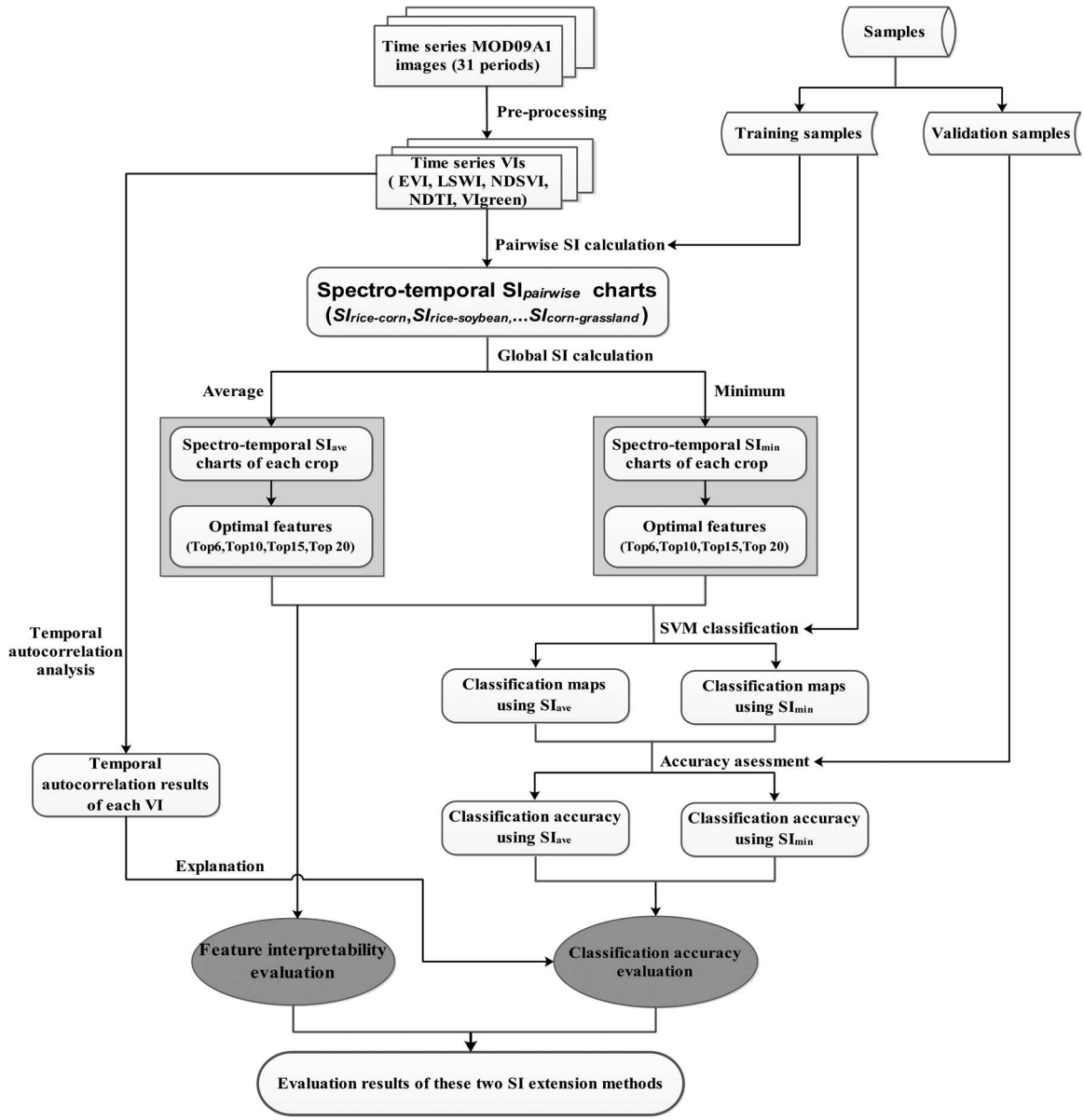


Fig. 1. Flow diagram of the processing and analysis steps.

crop types are rice, corn, soybeans, and wheat in this region, which are harvested once each year due to the limited sunshine hours and accumulated heat [22]. Their growing calendars, including key phenological events, are shown in Fig. 3. In general, rice is sown in early April and harvested in late September, experiencing a relatively long growing period. Wheat is sown earliest and harvested earliest of the four crops. Both corn and soybeans are cultivated under dry farming conditions and have similar growing periods. The phenological differences of the crops are closely related to the differences in leaf pigment, leaf water content, and canopy structure, which can be reflected by the variance in the spectral reflectance across the growing period.

B. Data Sources

1) Time-Series MODIS Data: This paper used time series of the eight-day composite MODIS Surface Reflectance Product (MOD09A1) with spatial resolutions of 250 m (Bands 1 and 2) and 500 m (Bands 3–7). The MOD09A1 product was geometrically, atmospherically, and bidirectional reflectance distribution fraction corrected and was deemed appropriate for reflecting the necessary spectral and temporal information for land-cover mapping [23], [24]. Four tiles (i.e., h25v03, h26v03, h26v04, and h27v04) that covered the entire study area and spanned the key growing season (from DOY 65 to DOY 305; 31 composite periods) in 2011 were acquired from the USGS EROS Data Center (http://edc.usgs.gov/). These tiled images were then mosaicked, reprojected from sinusoidal to WGS84 using the MODIS Reprojection Tool (https://lpdaac.usgs.gov/tools/modis_reprojection_tool), and clipped to the Heilongjiang boundary through an Interactive Data Language batch processing program.

Five VIs that were derived from different spectrum channel combinations, (i.e., the visible–visible, visible–NIR, visible–SWIR, NIR–SWIR, and SWIR–SWIR regions; Table I), were calculated for each composite period (a total of $31 \times 5 = 155$ VI

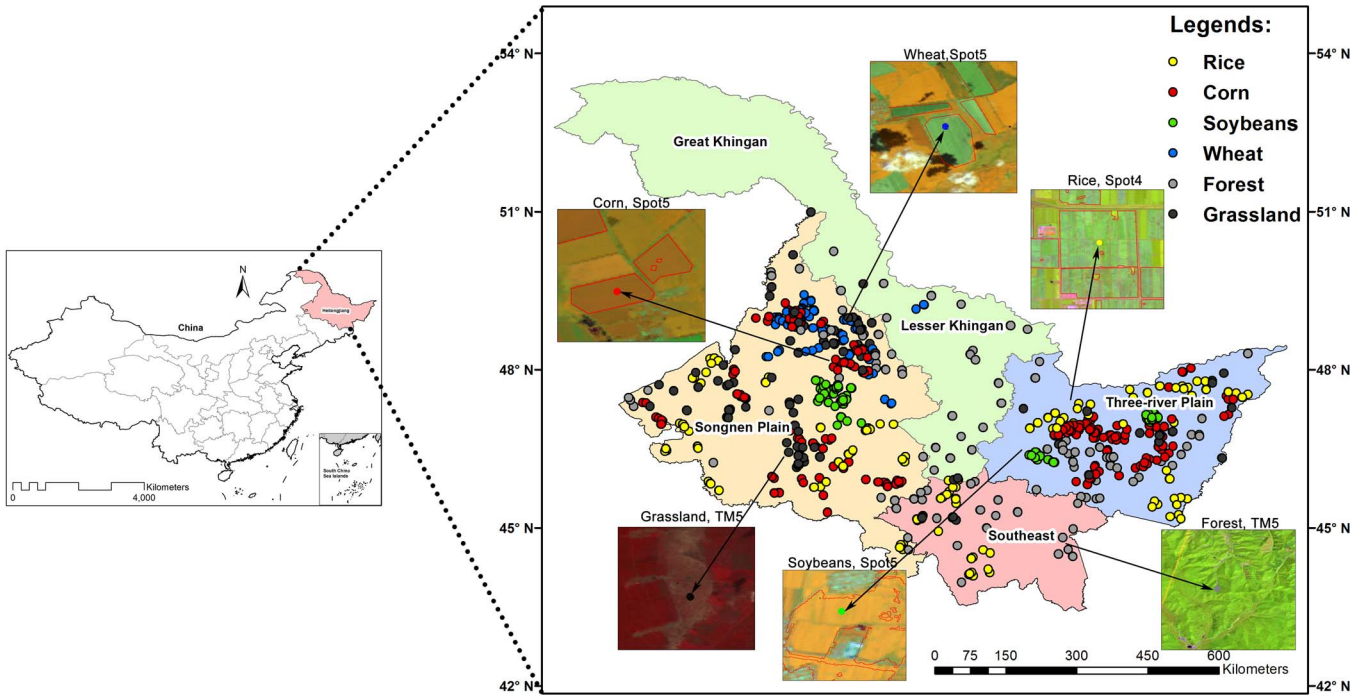


Fig. 2. Location of the study area and sampling points.

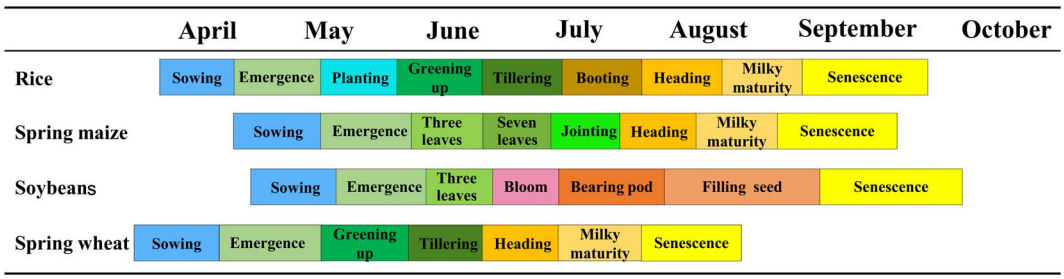


Fig. 3. Crop calendars for the four major crops in Heilongjiang Province.

images). These VIs were selected because they have been widely used in crop identification studies and have shown great potential for the detection of certain crop-field characteristics, such as leaf pigments, water content, and residue cover. The use of time-series VIs from different spectral channels is expected to adequately capture the phenological differences between different crop types and natural vegetation and consequently improve the crop classification accuracy.

2) Sampling Point Data: A total of 1920 sampling points, including 1135 crop (rice, corn, soybeans, and wheat) samples and 785 natural vegetation (mainly forest and grasslands) samples, were collected through visual interpretation of high spatial resolution images (i.e., SPOT4, SPOT5, and Landsat TM+) and ground visits to targeted locations (see Fig. 2). The size of the sampling plots exceeded 25 ha (500 m $\times$ 500 m) to guarantee relatively homogeneous samples for the subsequent SI-based separability analysis and classifier learning. Moreover, the sampling sites were spatially distributed in major crop areas of Heilongjiang Province to ensure an adequate number of crop samples and the representativeness of intraclass variations. In this paper, 100 samples of rice, corn, soybeans, wheat, forest, and grassland were randomly selected and used to conduct separability analysis. A total of 790 samples (415 crops and 375 noncrops) served as training samples for support vector machine (SVM) classifications, and the remaining 1130 samples were reserved for the accuracy assessment.

C. Performance Assessment and Comparison

1) Computation of SI_{ave} and SI_{min} and Evaluation of Feature Interpretability: The pairwise SI was calculated first using (1), where m are the VIs that are used in this paper, (i.e., EVI, LSWI, NDSVI, NDTI, and VIgreen), and n are the time periods, which ranges from 1 to 31. The SI of the two-class pair was calculated for individual VIs at each time period and was then used to examine how the separability between the two classes changes over the spectral and temporal domains (five VIs and 31 time periods, respectively). SI_{ave} and SI_{min} were then calculated for all 5 $\times$ 31 features $F(m,n)$ using (2) and (3), respectively, in which C' was equal to 6, and ρ_i and ρ_j were 0.17 because the six targeted classes had the same number of training samples for the SI calculation. We investigated both individual crop types and multiple crops to examine how the separability changes as the number of targeted classes increases. When the targeted object is one individual crop (e.g., rice), only the SI_{ij} values that are linked to the targeted

TABLE I
LIST OF VIS USED IN THIS PAPER

Spectral region	Vegetation index (VI)	VI adapted to MOD09A1	Commonly related to	Adapted from reference
Visible-Visible	Green vegetation index (VIgreen)	$\frac{B_4 - B_1}{B_4 + B_1}$	Leaf pigments, vegetation status	[5]
Visible-NIR	Enhanced vegetation index (EVI)	$2.5 * \frac{B_2 - B_1}{B_2 + 6 * B_1 - 7.5 * B_3 + 1}$	Vegetation status, canopy structure	[38, 39]
NIR-SWIR	Land surface water index (LSWI)	$\frac{B_2 - B_6}{B_2 + B_6}$	Water content, residue cover	[40]
Visible-SWIR	Normalized difference senescent vegetation index (NDSVI)	$\frac{B_6 - B_1}{B_6 + B_1}$	Vegetation status, water content, residue cover	[41]
SWIR-SWIR	Normalized difference tillage index (NDTI)	$\frac{B_6 - B_7}{B_6 + B_7}$	Non-photosynthetic components, residue cover	[41, 42]

Note: B1, B2, B3, B4, B5, B6 and B7 are the MODIS surface reflectance of red (Band1, 620–670 nm), near infrared (Band 2, 841–876 nm), blue (Band 3, 459–479 nm), green (Band 4, 545–565 nm), shortwave infrared (Band5, 1230–1250 nm), shortwave infrared (Band 6, 1628–1652 nm), and shortwave infrared (Band7, 2105–2155 nm) bands respectively.

class (i.e., rice–corn, rice–wheat, rice–soybean, rice–forest, and rice–grassland) were used to calculate $\mathrm{SI}_{\mathrm{ave}}$ and $\mathrm{SI}_{\mathrm{min}}$. When multiple crops are considered, all $\mathrm{SI}_{\mathrm{ij}}$ values, except $\mathrm{SI}_{\mathrm{forest-grassland}}$ that has little or no effect on crop identification, were used to calculate $\mathrm{SI}_{\mathrm{ave}}$ and $\mathrm{SI}_{\mathrm{min}}$. The resulting $\mathrm{SI}_{\mathrm{ave}}$ and $\mathrm{SI}_{\mathrm{min}}$ were illustrated in a matrix of five VIs by 31 time periods. These derived spectro-temporal charts of $\mathrm{SI}_{\mathrm{ave}}$ and $\mathrm{SI}_{\mathrm{min}}$ intuitively reflect the overall separability of each feature $F(m, n)$ for a given crop, which is further used to determine the optimal spectro-temporal features for classifying multiple crops. To qualitatively assess the feature interpretability that is determined using different extension approaches, a visual evaluation with the reference of crop calendar in Heilongjiang was conducted to examine whether the derived feature prioritization can respond to the crop growth status and how well the derived optimal features with high separability express the key phenological characteristics.

2) SVM Classification and Evaluation of Classification Accuracy: Although assessing feature interpretability can describe how well the selected features that are obtained by the $\mathrm{SI}_{\mathrm{ave}}$ and $\mathrm{SI}_{\mathrm{min}}$ methods capture the characteristics of crop growth, it fails to depict the difference in the classification accuracy of multiple crops between these two methods. To do this, for each extension approach, a certain number of important features with higher $\mathrm{SI}_{\mathrm{ave}}$ or $\mathrm{SI}_{\mathrm{min}}$ values were selected from the charts and used for crop classification by an SVM classifier. SVMs are widely used supervised classifiers that are based on a structural risk-minimization strategy to determine the location of decision boundaries that produce the optimal separation of classes in the training data [25], [26]. An SVM's ability to handle nonlinear separation boundaries, to generalize well from a limited number of training samples, and to support high-dimensional feature inputs makes it particularly efficient in discriminating complex hierarchical land-cover patterns [27]–[29]. SVM maximizes the margin between different classes to identify an optimal hyperplane using a small number of training samples. For nonlinearly separable input spaces, a kernel function is employed to transform the training data into a higher dimensional feature space; thus, linear class separation is possible [30]. In this paper, the Gaussian radial basis function kernel was selected because of its superior performance compared with other kernel functions [31]. This function requires two parameters: the regularization parameter C and the kernel bandwidth Υ. We estimated these parameters using a procedure that involves cross-validation and grid searching.

We intentionally selected the top 6, 10, 15, and 20 features to investigate the impacts of feature size on classification accuracy. All of the SVM classifications were conducted with the same 790 training samples and were implemented with the MATLAB-based LibSVM [32], which offers iterative determination of the most important parameters. The classification accuracy assessment was performed using the same 1130 validation samples, and commonly used accuracy metrics (producers' accuracy and users' accuracy) were used to evaluate the classification performances of different sets of features that were selected by the $\mathrm{SI}_{\mathrm{ave}}$ and $\mathrm{SI}_{\mathrm{min}}$ methods.

3) Temporal Autocorrelation Analysis: The information redundancy of input variables is a limiting factor for image classification accuracy [33]. We further investigated whether the temporal autocorrelation of selected features can explain the difference in classification accuracy between the $\mathrm{SI}_{\mathrm{ave}}$ and $\mathrm{SI}_{\mathrm{min}}$ methods. Because the five chosen VIs are a nonlinear combination of different reflectance bands at different spectrum channels and are therefore weakly correlated across the spectral domain, we specifically focused on the temporal autocorrelation of the time-series VIs because the eighth-day temporal interval of the MODIS data may contain information redundancy between neighboring images in the temporal domain. When the $\mathrm{SI}_{\mathrm{ave}}$ and $\mathrm{SI}_{\mathrm{min}}$ methods were used to determine the optimal features, a different level of temporal autocorrelation remained, which may partially account for the difference in classification accuracy between these two extension approaches. In this paper, the correlation coefficients between an image at a given time and other images that were acquired at different times were computed for individual VIs using the following:

$$R^2 = \frac{\mathrm{Cov}(\mathrm{VI}_{n1}, \mathrm{VI}_{n2})^2}{\mathrm{Var}(\mathrm{VI}_{n1}) * \mathrm{Var}(\mathrm{VI}_{n2})} \tag{4}$$

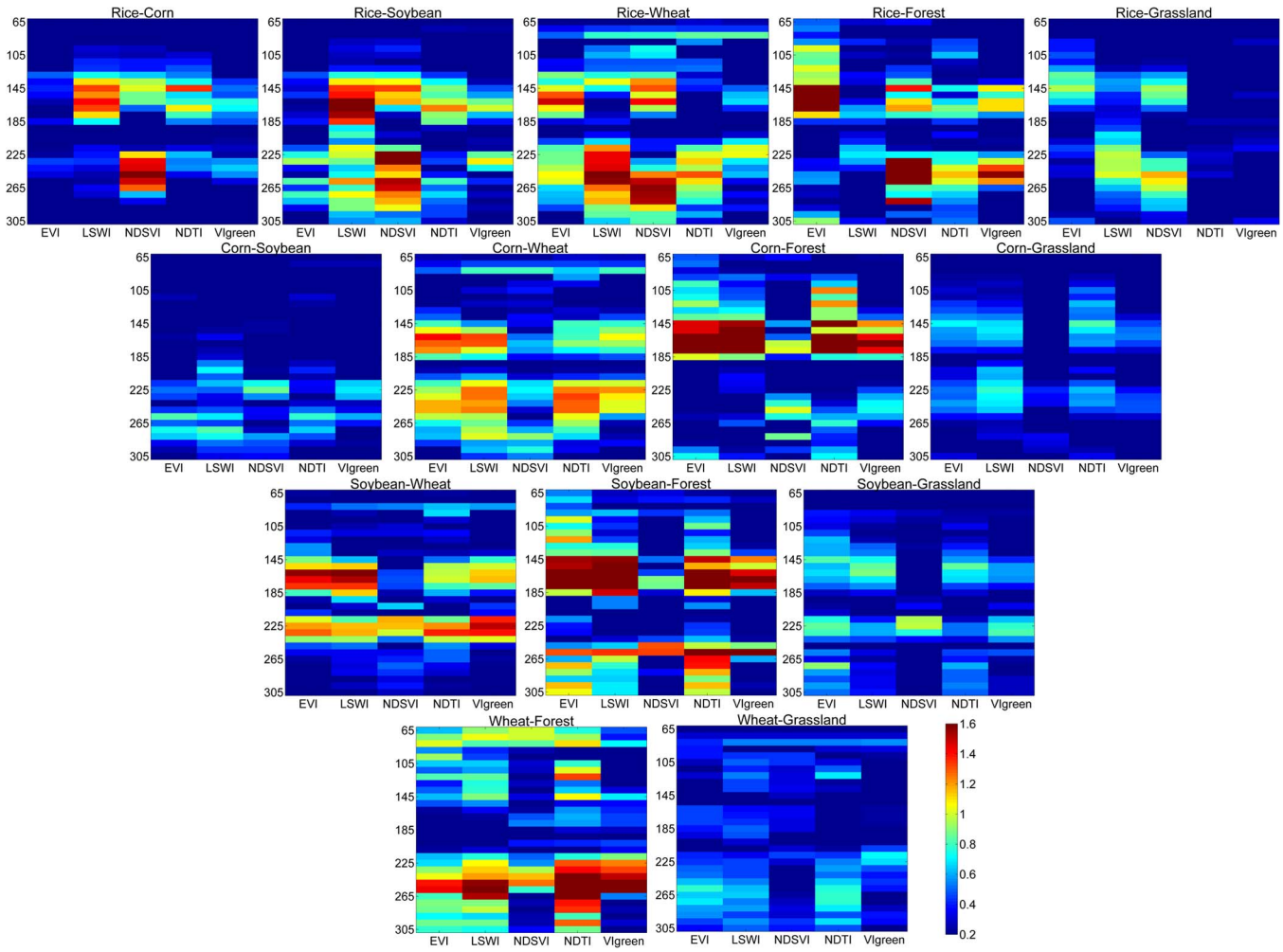


Fig. 4. Spectro-temporal SI_{ij} charts for the six classes of interest. The horizontal and vertical axes of the SI charts represent the VI and the time scale, respectively. The value in each grid represents the pairwise SI value of the corresponding feature $F(m, n)$ where m is the individual VI $\{\text{EVI}, \text{LSWI}, \ldots, \text{VIgreen}\}$ and n is the individual date $\{65, 73, 81, \ldots\ldots, 305\}$. Note that the $SI_{\text{forest}-\text{grassland}}$ that has little or no effect on crop identification was eliminated here. The redder the grid cell is, the more separate the two classes for a given feature $F(m, n)$ becomes.

where VI_{n1} and VI_{n2} are the variances of the VI value of the VI image acquired at one time period n_1 and at the other time period n_2, respectively, and $\text{Cov}(VI_{n1}, VI_{n2})$ is the covariance between these VI values. The R^2 calculation was performed to each VI image (i.e., EVI, LSWI, NDSVI, NDTI and VIgreen), and the resulting R^2 was illustrated in a matrix of 31 by 31 time periods.

IV. Results

A. Feature Interpretability of SI_{ave} and SI_{min}

Fig. 4 presents the spectro-temporal SI_{ij} charts for the pairwise classes. The horizontal and vertical axes of the SI_{ij} charts represent the VI (m) and time (n) domains, respectively. As the SI_{ij} value increases, the two classes for a given feature $F(m, n)$ become more separable, and the corresponding grid in the chart becomes increasingly red. The SI_{ij} charts show how the separability of two classes changes over the five VIs and 31 time points. These dynamic changes in the SIs mirror the occurrence of phenological events across the crop growing season, which agrees with the conclusions of Somers and Asner [13], [16]. For instance, during days 130–155 (late May) of the year, rice in Heilongjiang undergoes transplanting, during which the water content of the paddy field are significantly higher than those of corn fields at the stage of emergence. Thus, the pairwise "rice–corn" chart in this period is expected to have a relatively high pairwise SI for LSWI that is sensitive to leaf water and soil moisture [24], [34]. Fig. 4 also shows that different pairwise classes exhibit obvious discrepancies in separability for the given features of the spectral VIs and time periods, which highlights the merit of using different time-series VIs for multiclass identification because they can detect the subtle phenological differences between these classes.

The spectro-temporal charts that were derived using the two extension methods, (i.e., SI_{ave} and SI_{min}) are shown in Fig. 5, where the value in each grid is the overall separation measure, SI_{global}, which describes the overall separability between a specific class and other classes. In general, the charts that are derived from SI_{ave} exhibit subtler color variations than those from SI_{min}. Thus, feature prioritization based on the SI_{ave} approach shows higher coherence across the spectral and time scales, whereas the SI_{min} approach shows relatively high variations in feature prioritization, particularly for features with high SI_{global} values. When the number of pairwise classes increases,

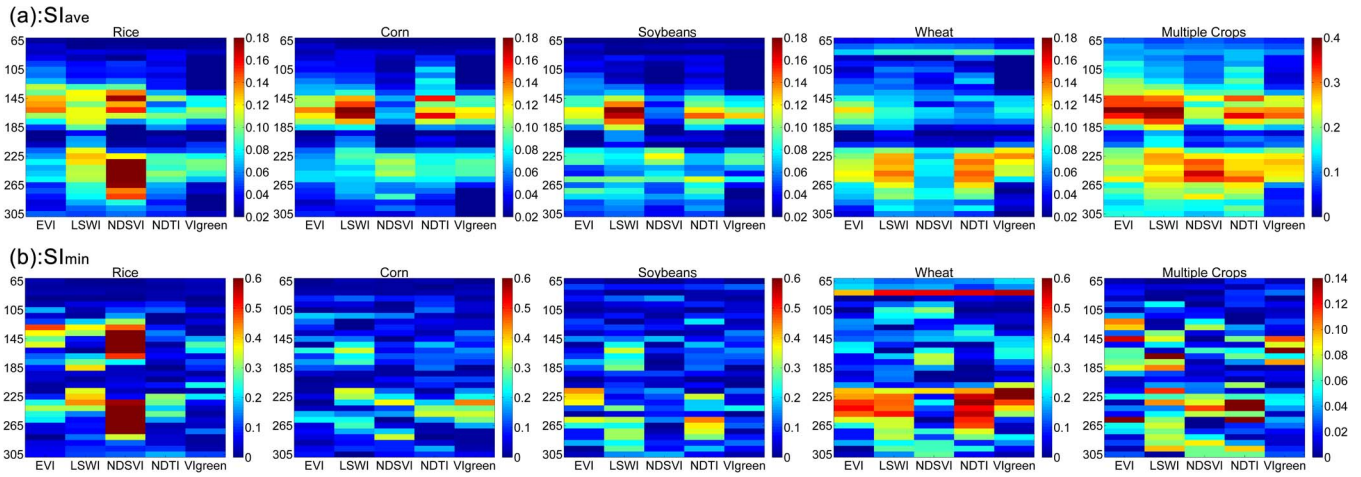


Fig. 5. Spectro-temporal $\mathrm{SI_{ave}}$ and $\mathrm{SI_{min}}$ charts for the individual crop types and multiple crops. The horizontal and vertical axes represent the VI and the time scale, respectively. (a) and (b) correspond to the $\mathrm{SI_{ave}}$ and $\mathrm{SI_{min}}$ extension approaches, respectively. "Rice," "Corn," "Soybeans," and "Wheat" indicate that the targeted class is an individual crop type, and the five associated $\mathrm{SI_{ij}}$ values in Fig. 4 were integrated to calculate the $\mathrm{SI_{ave}}$ and $\mathrm{SI_{min}}$ matrices. "Multiple Crops" indicates that the targeted classes include all four crops, and all 14 $\mathrm{SI_{ij}}$ values in Fig. 4 were integrated to calculate the $\mathrm{SI_{ave}}$ and $\mathrm{SI_{min}}$ matrices. The redder the grid cell is, the higher the overall separation between a specific class and other classes for a given feature $F(m, n)$ becomes.

the amplitude of the variation in feature prioritization increases dramatically in the $\mathrm{SI_{min}}$-based "multiple crops" chart, whereas that in the $\mathrm{SI_{ave}}$-based "Multiple crops" chart changes more slowly across the spectral and temporal scales. Fig. 5 also shows that some important phenological events are not expressed well by the $\mathrm{SI_{min}}$ approach. For instance, the transplanting stage, together with LSWI, has been demonstrated to be an effective feature to differentiate rice from other crops; this effectiveness can be seen in the $\mathrm{SI_{ave}}$ approach but is masked in the $\mathrm{SI_{min}}$ approach. Corn and soybeans can be distinguished from rice, wheat, and even grasslands using LSWI during the trefoil period due to the relatively low surface water and sparse vegetation cover (see Fig. 4). This phenological characteristic is expressed by $F(\mathrm{LSWI}, 140{-}170)$ in the $\mathrm{SI_{ave}}$ approach but fails in the $\mathrm{SI_{min}}$ approach. Therefore, $\mathrm{SI_{global}}$ that is derived from $\mathrm{SI_{ave}}$ tends to correlate better with the growth status of the crops. From the perspective of feature selection, the $\mathrm{SI_{ave}}$ approach better reflects the key phenological characteristics of crops, and the derived feature prioritization shows better interpretability than the $\mathrm{SI_{min}}$ approach.

When comparing the four individual crops, more red grids are found in the rice and wheat charts that are derived from the two extension methods, which suggests that rice and wheat have relatively strong separability compared with soybeans and corn. This observation can be explained by the fact that rice is irrigated and has unique phenological phases of flooding and transplanting when the field is a mixture of abundant surface water and a few rice plants; therefore, this crop presents a higher $\mathrm{SI_{global}}$ value in the feature $F(\mathrm{NDSVI}\ \&\mathrm{LSWI}\&\mathrm{EVI}, 140{-}165)$. When rice enters its milky maturity to senescence stages, the yellowing of the rice leaves and the low leaf water content result in lower NDSVI values than those of other crops and natural vegetation, which produces relatively high separability for the feature $F(\mathrm{NDSVI}, 241{-}265)$. Wheat in Heilongjiang is harvested earliest (late July to early August), whereas other crops are still in the milky maturity or seed filling stages and the natural vegetation is in a period of vigorous growth. Once wheat is harvested, its surface water and vegetation coverage are significantly lower than those of other classes, which can be expressed by the related VIs (NDTI, LSWI, and VIgreen) and the time periods (220 to 250 days) that show higher $\mathrm{SI_{global}}$ values in the charts.

The feature prioritization in Fig. 5 shows that both the $\mathrm{SI_{ave}}$ and $\mathrm{SI_{min}}$ methods have good similarity for rice and wheat but apparent differences for corn and soybeans. This finding is not surprising for rice and wheat because their pairwise $\mathrm{SI_{ij}}$ remains high (see Fig. 4), and exhibits a relatively low variance, which results in a stronger separability for rice and wheat (see Fig. 5). However, soybeans and corn have high $\mathrm{SI_{ij}}$ values for the pairwise classes of rice–soybean, soybean–forest, and corn–forest but have relatively low pairwise $\mathrm{SI_{ij}}$ values for corn–soybean, corn–grassland, and soybean–grassland. The high variances in pairwise $\mathrm{SI_{ij}}$ cause significant differences in feature ranking and prioritization derived from the $\mathrm{SI_{ave}}$ and $\mathrm{SI_{min}}$ extension methods. Thus, crops with low separability tend to be more sensitive to the extension approach for feature selection.

B. *Classification Accuracy of* $\mathrm{SI_{ave}}$ *and* $\mathrm{SI_{min}}$

Both charts that are derived from the $\mathrm{SI_{ave}}$ and $\mathrm{SI_{min}}$ approaches (see Fig. 5) accurately depict the ranking of feature importance for both individual crops and multiple crops. To further assess the performances of the two approaches on crop identification, the most important 6, 10, 15, and 20 features were selected from these charts and are shown in Table II, where the case of multiple crops is illustrated as an example. The optimal features that are selected using the $\mathrm{SI_{ave}}$ and $\mathrm{SI_{min}}$ approaches show substantial differences in the selected VIs and time points.

These selected features were used for the classification of individual crops and multiple crops. The classification maps of multiple crops, which are derived from the $\mathrm{SI_{ave}}$ and $\mathrm{SI_{min}}$ methods using the top 20 features, are shown in Fig. 6 because of their relatively high classification accuracy. Fig. 7 shows the producers' and users' accuracy of all targeted classes,

TABLE II
OPTIMAL 6, 10, 15, AND 20 FEATURES FOR THE IDENTIFICATION OF MULTIPLE CROPS BASED ON SI_{ave} AND SI_{min}, RESPECTIVELY

Features	SI_{ave}	SI_{min}
Top_6	(EVI, 161&169) (LSWI, 161&169) (NDSVI, 249) (NDTI, 169)	(EVI, 257) (LSWI, 169) (NDTI, 233&241&257) (Vigreen, 161)
Top_10	(EVI, 153&161&169) (LSWI, 153&161 &169&177) (NDSVI, 249&257) (NDTI,169)	(EVI,145&257) (LSWI, 169&217) (NDTI, 233&241&257) (NDSVI, 241) (VIgreen, 145&161)
Top_15	(EVI, 145&153&161&169) (LSWI, 153&161 &169&177) (NDSVI, 249&257&233) (NDTI, 145&169&257) (VIgreen, 169)	(EVI,121&145&257) (LSWI, 169&185&217&233&297) (NDSVI, 241) (NDTI, 233&241&257) (VIgreen,145&161&177)
Top_20	(EVI, 145&153&161&169) (LSWI, 145&153&161 &169&177&225) (NDSVI, 233&241&249&257) (NDTI, 145&161&169&257) (VIgreen,161&169)	(EVI,121&129&145&185&257) (LSWI, 169&185&217&233&297) (NDTI, 233&241&257&297) (NDSVI, 233&241&289) (VIgreen, 145&161&177)

Note: The features in gray are the common features of the two approaches. Multiple Crops indicates that the targeted classes include all four crops, and all fourteen SI_{ij} values in Fig. 4 were integrated to calculate the SI_{ave} and SI_{min} matrixes. The "&" represents the parallel relationship between two time periods. For instance, (EVI, 161 & 169) equals to *F(EVI, 161)* and *F(EVI, 169)*.

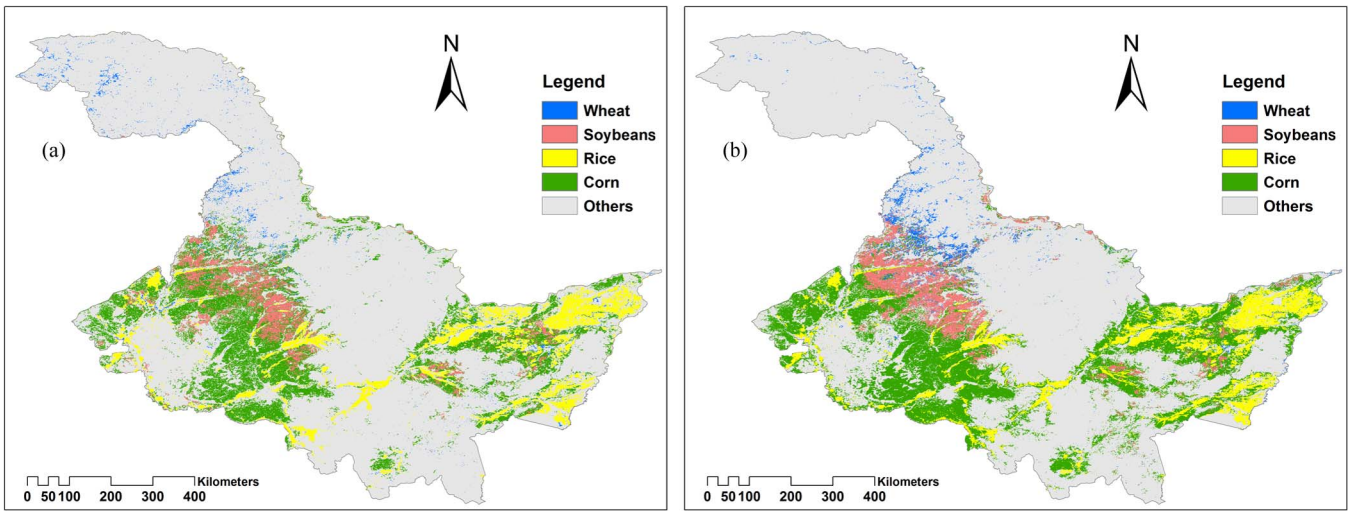


Fig. 6. Classification maps of multiple crops based on the top 20 features derived from (a) the SI_{ave} approach and (b) the SI_{min} approach.

among which the producers' accuracy is a measure of omission error and indicates the probability of actual areas being correctly classified, whereas the users' accuracy is a measure of commission error and indicative of the probability that a category classified on the map actually represents that category on the ground. When compared with the other three crops, rice generally shows the highest overall accuracy and a small accuracy difference between the SI_{ave} and SI_{min} approaches, except that the users' accuracy of the "Top_6" feature subspace that is derived from SI_{min} is slightly less than that of SI_{ave}. The high classification accuracy and accuracy consistency of rice are attributed to its strong separability and good agreement between the SI_{ave}- and SI_{min}-derived SI_{global} prioritizations, as was shown in Section IV-A. As the feature quantities increase further, the classification accuracy of rice is expected to increase accordingly, and the difference between these two approaches will decrease gradually due to the increasing number of common optimal features that are shared by the two approaches. For corn and soybeans that present relatively poor class separability (see Fig. 5), their classification accuracy using the features that are selected by SI_{min} are much higher than that of SI_{ave}. This outcome may occur because the SI_{min} method tends to select the minimum SI_{ij} over all pairwise SI_{ij} to determine SI_{global}, thus giving a higher priority to the least separable pair of classes, which normally has a large impact on the overall accuracy in the context of classes that are significantly confused. Thus, the selected optimal features with the highest minimum SI_{ij} from the SI_{min} method will result in relatively good accuracy compared with the SI_{ave} method. Moreover, the classification accuracy of soybeans is lower than that of corn, and the difference in accuracy between the two extension approaches is distinctly larger. This observation is due to the relatively heterogeneous nature of soybeans fields in Heilongjiang, which results in more mixed pixels than corn when using MODIS images for classification. In addition, although soybeans and corn experience similar phenological periods, the big discrepancy in biochemical and structural properties at specific phenological stages can result in different selection of optimal spectro-temporal features and thus difference in their classification accuracy. The users' accuracy for wheat using the SI_{ave} method was significantly higher than that using SI_{min}, but no clear pattern between these two

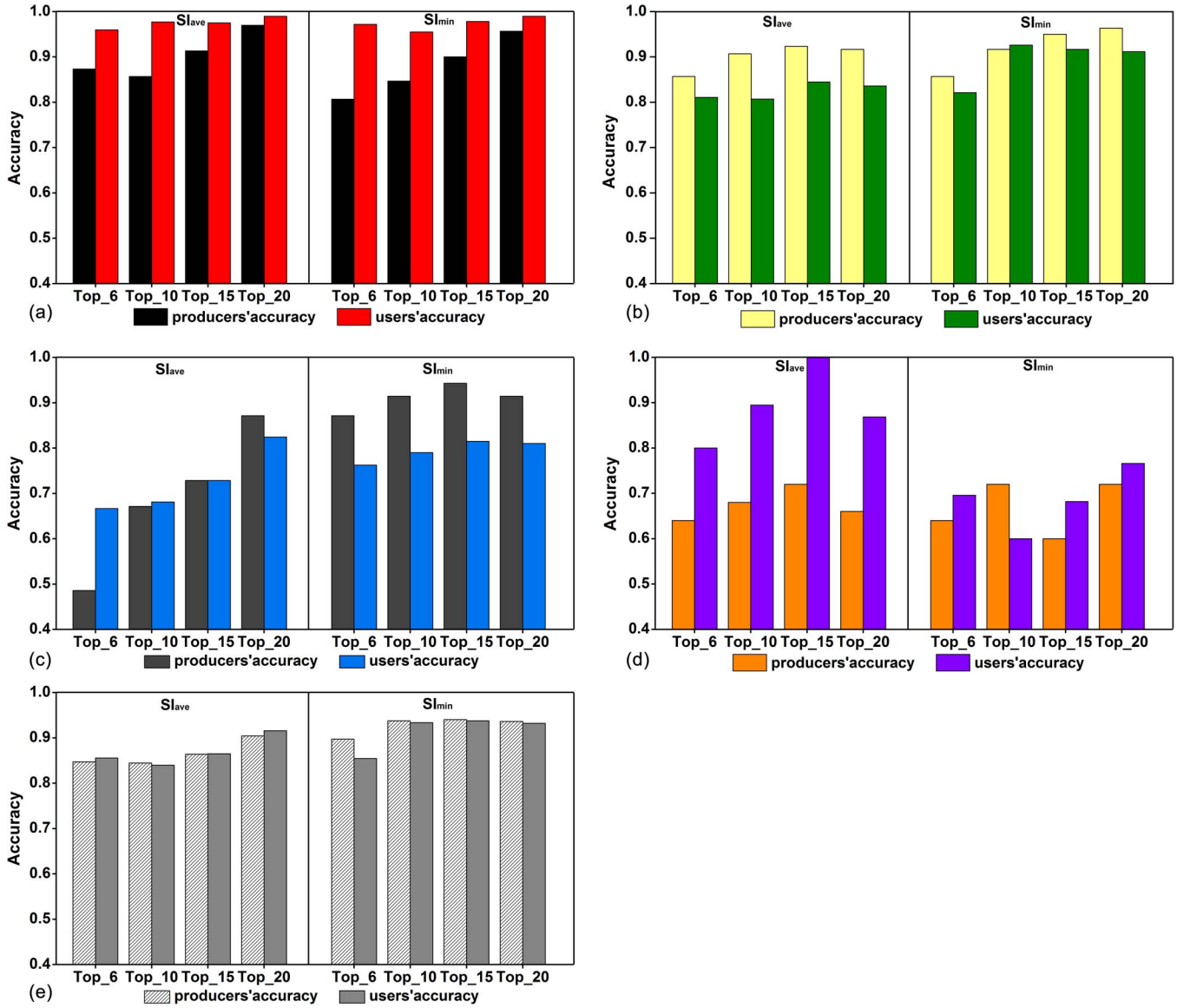


Fig. 7. Producers' accuracy and users' accuracy for targeted crops using features that are derived from the SI_{ave} and SI_{min} approaches. (a), (b), (c), (d) and (e) represent the results for rice, corn, soybeans, wheat and multiple crops, respectively, and Top_6, Top_10, Top_15 and Top_20 represent different numbers of optimal features as shown in Fig. 5.

approaches was found for the producers' accuracy for wheat. However, the classification accuracy fluctuates substantially as the feature dimensions increase for both SI_{ave} and SI_{min}. The low proportion of wheat areas in Heilongjiang Province (see Section III-A) may be responsible for the relatively low classification accuracy and the large variance in accuracy among the different feature subspaces. Finally, for multiple crops, the SI_{min} method also shows higher classification accuracy than SI_{ave}, particularly for the "Top_10" and "Top_15" feature subspaces.

The results of the comparison of crop classification accuracy illustrate that SI_{min} generally performs better in crop identification than SI_{ave}. Crops with high separability (such as rice) are less sensitive to the extension approach, and their classification accuracy remains stable. However, the crops with poor separability, particularly when they are associated with fragmented field sizes, have good performance with the SI_{min} method because this method prioritizes the classes with poor separability and thus generates relatively high classification accuracy for these crops. Low proportions of cultivated area increase the volatility of the classification accuracy between the different feature subspaces and consequently add further uncertainty to the selection of the optimal extension approach.

C. Reasons for the Different Accuracy of These Two Methods

The results that were presented in Section IV-B show that the SI_{min} approach generally has better crop identification accuracy than the SI_{ave} approach. A temporal autocorrelation analysis was performed to explain the difference in accuracy between these two extension approaches from the perspective of information redundancy. Fig. 8 shows the temporal autocorrelation of five time-series VIs at a temporal interval of eight days. The value of each grid represents the correlation coefficient

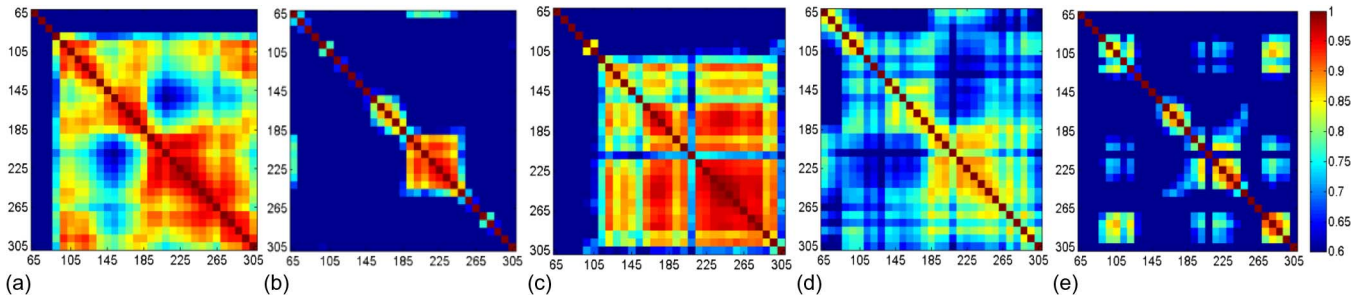


Fig. 8. Temporal autocorrelation of the time-series VIs. (a) EVI, (b) LSWI, (c) NDSVI, (d) NDTI, and (e) VIgreen. Both the horizontal and vertical axes of the charts represent the time scale. The redder the grid cell is, the higher the correlation between two VI images acquired on a different date is.

of determination R^2 between one VI image and another image acquired on a different date. As R^2 increases, the information in the two images becomes increasingly redundant. As expected, high levels of redundancy occurred between temporally neighboring images. Because of the high continuity across time, the optimal features that were selected by the $\mathrm{SI}_{\mathrm{ave}}$ approach for crop classification have greater information redundancy than those selected by the $\mathrm{SI}_{\mathrm{min}}$ approach. This largely explains the difference in accuracy between the two approaches. For example, the 20 selected optimal features include $F(\mathrm{EVI}, 145{-}169)$, $F(\mathrm{LSWI}, 145{-}177\ \&225)$ and $F(\mathrm{NDSVI}, 233{-}257)$ for multiple crops (Table II), and they all have higher R^2 values on the corresponding dates on the EVI, LSWI, and NDSVI charts in Fig. 8. In addition, the top 6, 10, and 15 features that were derived from $\mathrm{SI}_{\mathrm{ave}}$ generally include more time-adjacent features that contain greater information correlation than $\mathrm{SI}_{\mathrm{min}}$, which could account for the relatively low classification accuracy of the $\mathrm{SI}_{\mathrm{ave}}$ method compared with $\mathrm{SI}_{\mathrm{min}}$ (see Section IV-B). The optimal features of the individual soybeans and corn crops that were selected by the $\mathrm{SI}_{\mathrm{ave}}$ approach, such as $F(\mathrm{EVI}, 161\text{–}177)$ and $F(\mathrm{LSWI}, 161\text{–}177)$, largely coincide with the red grids that are associated with high information redundancy in Fig. 8 and thus result in lower classification accuracy. As described in Section IV-A, the derived feature prioritization by $\mathrm{SI}_{\mathrm{ave}}$ changes slightly with time relative to $\mathrm{SI}_{\mathrm{min}}$ and has good consistency with the real crop growth processes in Heilongjiang. However, its relatively high temporal autocorrelation of the derived optimal features causes more information redundancy than that of $\mathrm{SI}_{\mathrm{min}}$, which consequently reduces the classification accuracy.

The temporal autocorrelation of selected features can have a negative effect on the classification accuracy because the generated information redundancy can reduce the usefulness of information for specific classes compared with uncorrelated features of the same size. This observation suggests that crops with high separability only need to achieve a certain magnitude of optimal features for classification and that more information with high redundancy neither adds value to the classification nor substantially reduces its classification accuracy. This is why rice has relatively high classification accuracy and a relatively small difference in accuracy between the two approaches, despite the temporal autocorrelation between its optimal features. This temporal autocorrelation of the MODIS images with the eight-day time interval indicates that time-series images with high temporal resolution do not necessarily produce high classification accuracy because high temporal resolution leads to high information redundancy. Furthermore, Fig. 8 reveals that different VIs have different degrees of temporal autocorrelation. The time-series EVI and NDSVI contain more temporal autocorrelation than the other VIs, and the period from days 225 to 255, when most of the crops in Heilongjiang are in the stage of reproductive growth, was surprisingly found to be associated with information redundancy, which should draw attention to the selection of optimal features for crop identification.

V. Discussion

Although there are many prominent feature selection methods for land cover classification [25], [35], [36], the SI-based method in this study was demonstrated to be particularly suitable for selecting the optimal spectro-temporal features for crop identification due to its ease of use, efficiency, and most importantly its transparent feature importance ranking that well reflects the crucial phenological characteristics of crops. This paper exemplifies the possibility of extending the pairwise $\mathrm{SI}_{\mathrm{ij}}$ to determine the optimal features for multiclass classification.

In this paper, feature interpretability and classification accuracy were taken as major indicators to evaluate the performance of the two SI extension approaches ($\mathrm{SI}_{\mathrm{ave}}$ and $\mathrm{SI}_{\mathrm{min}}$). Classification accuracy is widely used as one of the most important criteria because accurate classification maps are useful for supporting various monitoring applications and decision making. However, the ability to clearly interpret the results of selected optimal features must be taken into consideration [36]. Easy interpretability generally implies that the selected features can well describe the intrinsic and unique physical properties of the targeted classes that are significantly different from other classes. Therefore, the selected features with high interpretability tend to be relatively robust when they are applied to another year or study area. This paper shows that the $\mathrm{SI}_{\mathrm{ave}}$ approach has better feature interpretability but relatively lower classification accuracy than the $\mathrm{SI}_{\mathrm{min}}$ approach because of its high temporal autocorrelation of time-series VIs. This finding suggests that to some extent achieving high feature interpretability may require sacrificing classification accuracy. When a strategy to reduce temporal autocorrelation is performed to improve the classification accuracy of $\mathrm{SI}_{\mathrm{ave}}$, its feature interpretability could be somehow decreased because the natural dependencies between certain features across the time and spectral domains are destroyed.

The SI_{ave} approach is easily affected by an atypically large pairwise SI_{ij}, which can leverage the mean distance and mask the overall classification difficulty, whereas the SI_{min} approach is particularly sensitive to very small pairwise SI_{ij} values and may overstate the overall classification difficulty and neglect the effective influence of large pairwise SI_{ij} values. Theoretically, when the variance of all targeted pairwise SI_{ij} values is sufficiently small, the difference in classification accuracy and interpretability between these two extension approaches should also be small. However, in practice, the real population distribution characteristics of the targeted classes are unknown, and the separation principle of SI_{global} does not exactly coincide with the intrinsic mechanism of feature selection of the used classifier. Therefore, the extent to which SI_{global} value has the ability to separate two classes remains unknown; thus, it is difficult to determine which approach is better simply based on the variance of limited samples. Therefore, in this paper, we did not evaluate the suitability of these two approaches directly based on the variance characteristic of all associated pairwise SI_{ij} but instead focused on the classification accuracy using optimal features derived from these two methods.

Four feature subspaces with different sizes were used in this paper because feature quantity is another important factor for classification accuracy in addition to feature quality [3], [37], and classifications with a large number of features may blur the effects of feature quality on accuracy. Therefore, to objectively assess the classification performances of these two approaches, the number of optimal features that are used for SVM classification should be the same. In addition, to examine the stability of the accuracy difference and exploit the maximum accuracy possible, various feature space dimensions were tested for these two approaches. The results show that when the number of selected features reaches a certain level, the classification accuracy becomes stable or decreases, which is the well-known "Hughes effect" [3], [25]. As illustrated in Section IV-B, the overall classification accuracy increases as the number of features increases. The SI_{min} approach seems to easily produce the "Hughes effect" because of its higher classification accuracy than the SI_{ave} approach. In particular, when the number of features increases from 15 to 20, the classification accuracy of corn, soybeans, and multiple crops with SI_{min} remains stable or decreases slightly. Because the objective of this paper was to conduct an evaluation and comparison between the two SI extension approaches rather than to propose a classification strategy to maximize the absolute classification accuracy, the accuracy for the different crop types in this paper may not be the highest achievable accuracy. We conclude that the SI-based analysis that automatically identifies the individual optimal spectro-temporal features, combined with a strategy of reducing information redundancy to determine the most appropriate size of the feature subspace, is a good option for feature selection and may yield a high accuracy and low computing time.

Although this paper aims to investigate the potential of these two extension approaches in selecting the optimal time period and appropriate spectral features for multiple crop types, the evaluation results offered other valuable insights into crop classification. First, the derived SI_{ave} and SI_{min} charts highlight the clear superiority of combining multispectral and multitemporal images for multiple crop identification because they are capable of depicting the dynamic characteristics of agricultural land use systems and capturing the subtle differences between crop types, particularly for those crops that have different growth periods or share similar phenological phrases but present significantly different biochemical and structural properties. Additionally, individual spectral or temporal features do not contribute equally to crop identification, and there is feature prioritization, which can facilitate the appropriate selection of imported features for regions with similar planting patterns. For example, another spectral feature (e.g., LSWI) can compensate for the absence of certain temporal images of some VIs (e.g., EVI) that may be destroyed by cloud cover. Second, the existence of temporal autocorrelation emphasizes that high-temporal-resolution time-series images do not necessarily guarantee high classification accuracy due to the occurrence of information redundancy compared with images with relatively low temporal resolution. However, the identification of multiple crop types with substantial intraclass variability and interclass similarity requires a relatively high temporal resolution. Therefore, the strategy of selecting the appropriate temporal resolution and making trade-offs between feature interpretability, classification accuracy, and computing efficiency should be further studied in the future.

It needs to be noted that, although the major data source in this paper is time-series MODIS, the proposed extended methods for feature selection do not have strict requirements for satellite images according to the algorithm principles. This may inspire more future attempts of these techniques to time-series high and medium spatial resolution images such as Landsat 8 and sentinel-2A for multiple crops classification as these images are capable of characterizing spectral and texture dynamics of crops particularly in complex agriculture landscapes such as South China. Moreover, it is also necessary to test the proposed method in different regions particularly those with multiple cropping practices and to compare their performance in the future since crop variety, crop calendar, and agricultural management generally vary with regions.

VI. Conclusion

The SI has been demonstrated to be particularly useful for characterizing the seasonal dynamics of land-cover classes, which inspired our interest in investigating its potential for selecting the optimal time period and appropriate spectral features for multiple-crop identification. From the perspective of feature selection, previous SI studies have focused on two-class cases (i.e., native and invasive species) and have not addressed multiclass cases. This paper has extended this pairwise SI to multiple classes using the SI_{ave} and SI_{min} approaches and evaluated their performance in the feature selection of different crop types using multiple time-series MODIS VI images in Heilongjiang Province, China. The feature interpretability and classification accuracy that were derived from the SVM classifier were assessed. The evaluation results of feature interpretability illustrated that the SI_{ave} approach has higher feature interpretability because it is superior at expressing the key

phenological characteristics of crops, and the derived prioritizations correspond better to the growth status of the crops. The evaluation of the classification accuracy illustrated that the crops with high separability (e.g., rice) are insensitive to the different extension approaches and have stable classification accuracy for the two approaches, whereas the SI_{min} approach is more suitable than SI_{ave} for crops with poor separability (e.g., corn and soybeans), particularly when they are associated with fragmented field sizes. The relatively low classification accuracy of SI_{ave} is largely explained by its relatively significant temporal autocorrelation because its selected optimal features tend to obtain higher temporal adjacency, which results in greater information redundancy than in the SI_{min} approach.

Both the SI_{ave} and SI_{min} charts present clear feature prioritizations, which can provide guidance for the imagery collection schedules in regions with similar agricultural systems. Because high temporal resolution may cause high information redundancy between time-adjacent images, high-temporal-resolution images do not necessarily result in high classification accuracy for crops, despite their good ability to express dynamic seasonal characteristics. These results establish the basis for an operational crop monitoring program in Northeast China using time-series MODIS images. However, balancing the feature interpretability, classification accuracy and classification efficiency should be considered further in the future.

ACKNOWLEDGMENT

The authors would like to thank the anonymous reviewers for their comments and suggestions, which considerably improved this paper and the fellows at the Remote Sensing Technology Center, Heilongjiang Academy of Agriculture Sciences, for providing the essential crop field samples.

REFERENCES

[1] B. D. Wardlow and S. L. Egbert, "Large-area crop mapping using time-series MODIS 250 m NDVI data: An assessment for the US Central Great Plains," *Remote Sens. Environ.*, vol. 112, no. 3, pp. 1096–1116, 2008.

[2] M. Ozdogan and G. Gutman, "A new methodology to map irrigated areas using multi-temporal MODIS and ancillary data: An application example in the continental US," *Remote Sens. Environ.*, vol. 112, no. 9, pp. 3520–3537, Sep. 2008.

[3] F. Loew, U. Michel, S. Dech, and C. Conrad, "Impact of feature selection on the accuracy and spatial uncertainty of per-field crop classification using support vector machines," *ISPRS J. Photogramm.*, vol. 85, pp. 102–119, Nov. 2013.

[4] W. Wu, Q. Yu, V. H. Peter, L. You, P. Yang, and H. Tang, "How could agricultural land systems contribute to raise food production under global change?" *J. Integr. Agriculture*, vol. 13, no. 7, pp. 1432–1442, 2014.

[5] J. M. Peña-Barragán, M. K. Ngugi, R. E. Plant, and J. Six, "Object-based crop identification using multiple vegetation indices, textural features and crop phenology," *Remote Sens. Environ.*, vol. 115, no. 6, pp. 1301–1316, Jun. 2011.

[6] L. Zhong, T. Hawkins, G. Biging, and P. Gong, "A phenology-based approach to map crop types in the San Joaquin Valley, California," *Int. J. Remote Sens.*, vol. 32, no. 22, pp. 7777–7804, 2011.

[7] Q. Hu *et al.*, "Exploring the use of google earth imagery and object-based methods in land use/cover mapping," *Remote Sens.-Basel*, vol. 5, no. 11, pp. 6026–6042, Nov. 2013.

[8] M. W. Liu, M. Ozdogan, and X. Zhu, "Crop type classification by simultaneous use of satellite images of different resolutions," *IEEE Trans. Geosci. Remote*, vol. 52, no. 6, pp. 3637–3649, Jun. 2014.

[9] T. G. Van Niel and T. R. McVicar, "Determining temporal windows for crop discrimination with remote sensing: A case study in south-eastern Australia," *Comput. Electron. Agr.*, vol. 45, no. 1, pp. 91–108, 2004.

[10] B. D. Wardlow, S. L. Egbert, and J. H. Kastens, "Analysis of time-series MODIS 250 m vegetation index data for crop classification in the US Central Great Plains," *Remote Sens. Environ.*, vol. 108, no. 3, pp. 290–310, 2007.

[11] D. K. Bolton and M. A. Friedl, "Forecasting crop yield using remotely sensed vegetation indices and crop phenology metrics," *Agr. Forest Meteorol.*, vol. 173, pp. 74–84, May 2013.

[12] T. Sakamoto, B. D. Wardlow, and A. A. Gitelson, "Detecting spatiotemporal changes of corn developmental stages in the U.S. corn belt using MODIS WDRVI data," *IEEE Trans. Geosci. Remote*, vol. 49, no. 6, pp. 1926–1936, Jun. 2011.

[13] B. Somers and G. P. Asner, "Multi-temporal hyperspectral mixture analysis and feature selection for invasive species mapping in rainforests," *Remote Sens. Environ.*, vol. 136, pp. 14–27, 2013.

[14] J. Zhang, B. Rivard, A. Sánchez-Azofeifa, and K. Castro-Esau, "Intra- and inter-class spectral variability of tropical tree species at La Selva, Costa Rica: Implications for species identification using HYDICE imagery," *Remote Sens. Environ.*, vol. 105, no. 2, pp. 129–141, 2006.

[15] B. Somers and G. P. Asner, "Tree species mapping in tropical forests using multi-temporal imaging spectroscopy: Wavelength adaptive spectral mixture analysis," *Int. J. Appl. Earth Observ.*, vol. 31, pp. 57–66, 2014.

[16] B. Somers and G. P. Asner, "Hyperspectral time series analysis of native and invasive species in hawaiian rainforests," *Remote Sens.-Basel*, vol. 4, no. 12, pp. 2510–2529, 2012.

[17] L. Bruzzone, F. Roli, and S. B. Serpico, "An extension of the Jeffreys-Matusita distance to multiclass cases for feature selection," *IEEE Trans. Geosci. Remote*, vol. 33, no. 6, pp. 1318–1321, Nov. 1995.

[18] P. Hao *et al.*, "The potential of time series merged from Landsat-5 TM and HJ-1 CCD for crop classification: A case study for Bole and Manas Counties in Xinjiang, China," *Remote Sens.-Basel*, vol. 6, no. 8, pp. 7610–7631, 2014.

[19] H. Carrão, P. Gonçalves, and M. Caetano, "Contribution of multispectral and multitemporal information from MODIS images to land cover classification," *Remote Sens. Environ.*, vol. 112, no. 3, pp. 986–997, 2008.

[20] M. Herold, M. E. Gardner, and D. A. Roberts, "Spectral resolution requirements for mapping urban areas," *IEEE Trans. Geosci. Remote*, vol. 41, no. 9, pp. 1907–1919, Sep. 2003.

[21] S. Shi *et al.*, "Influence of climate and socio-economic factors on the spatio-temporal variability of soil organic matter: A case study of Central Heilongjiang Province, China," *J. Integr. Agr.*, vol. 13, no. 7, pp. 1486–1500, 2014.

[22] J. Sun, W. Wu, H. Tang, and J. Liu, "Spatiotemporal patterns of non-genetically modified crops in the era of expansion of genetically modified food," *Sci. Rep.-U.K.*, vol. 5, pp. 1–7, Sep. 2015.

[23] J. Chang, M. C. Hansen, K. Pittman, M. Carroll, and C. DiMiceli, "Corn and soybean mapping in the United States using MODIS time-series data sets," *Agron. J.*, vol. 99, no. 6, pp. 1654–1664, 2007.

[24] X. Xiao *et al.*, "Mapping paddy rice agriculture in South and Southeast Asia using multi-temporal MODIS images," *Remote Sens. Environ.*, vol. 100, no. 1, pp. 95–113, Jan. 2006.

[25] M. Pal and F. Giles, "Feature selection for classification of hyperspectral data by SVM," *IEEE Trans. Geosci. Remote*, vol. 45, no. 5, pp. 2297–2307, May 2010.

[26] A. Mathur and G. M. Foody, "Multiclass and binary SVM classification: Implications for training and classification users," *IEEE Geosci. Remote Sens. Lett.*, vol. 5, no. 2, pp. 241–245, Apr. 2008.

[27] G. Mountrakis, J. Im, and C. Ogole, "Support vector machines in remote sensing: A review," *ISPRS J. Photogramm.*, vol. 66, no. 3, pp. 247–259, 2011.

[28] J. Tigges, T. Lakes, and P. Hostert, "Urban vegetation classification: Benefits of multitemporal RapidEye satellite data," *Remote Sens. Environ.*, vol. 136, pp. 66–75, 2013.

[29] R. Piiroinen, J. Heiskanen, M. Mõttus, and P. Pellikka, "Classification of crops across heterogeneous agricultural landscape in Kenya using AisaEAGLE imaging spectroscopy data," *Int. J. Appl. Earth Observ.*, vol. 39, pp. 1–8, 2015.

[30] C. Huang, L. S. Davis, and J. R. G. Townshend, "An assessment of support vector machines for land cover classification," *Int. J. Remote Sens.*, vol. 23, no. 4, pp. 725–749, 2002.

[31] L. Xu, J. Li, and A. Brenning, "A comparative study of different classification techniques for marine oil spill identification using RADARSAT-1 imagery," *Remote Sens. Environ.*, vol. 141, pp. 14–23, 2014.

[32] C. Chang and C. Lin, "LIBSVM: A library for support vector machines," *ACM Trans. Intell. Syst. Tech.*, vol. 2, no. 3, pp. 1–27, 2011.

[33] E. L. Hestir *et al.*, "Identification of invasive vegetation using hyperspectral remote sensing in the California Delta ecosystem," *Remote Sens. Environ.*, vol. 112, no. 11, pp. 4034–4047, 2008.

[34] G. Zhang *et al.*, "Mapping paddy rice planting areas through time series analysis of MODIS land surface temperature and vegetation index data," *ISPRS J. Photogramm.*, vol. 106, pp. 157–171, 2015.
[35] F. M. B. Van Coillie, L. P. C. Verbeke, and R. R. De Wulf, "Feature selection by genetic algorithms in object-based classification of IKONOS imagery for forest mapping in Flanders, Belgium," *Remote Sens. Environ.*, vol. 110, no. 4, pp. 476–487, Oct. 2007.
[36] A. S. Laliberte, D. M. Browning, and A. Rango, "A comparison of three feature selection methods for object-based classification of sub-decimeter resolution UltraCam-L imagery," *Int. J. Appl. Earth Obs.*, vol. 15, pp. 70–78, Apr. 2012.
[37] F. Zhou, A. Zhang, and L. Townley-Smith, "A data mining approach for evaluation of optimal time-series of MODIS data for land cover mapping at a regional level," *ISPRS J. Photogramm.*, vol. 84, pp. 114–129, 2013.
[38] Y. Pan, L. Li, J. Zhang, S. Liang, X. Zhu, and D. Sulla-Menashe, "Winter wheat area estimation from MODIS-EVI time series data using the Crop Proportion Phenology Index," *Remote Sens. Environ.*, vol. 119, pp. 232–242, Apr. 2012.
[39] J. Zhang, L. Feng, and F. Yao, "Improved maize cultivated area estimation over a large scale combining MODIS-EVI time series data and crop phenological information," *ISPRS J. Photogramm.*, vol. 94, pp. 102–113, 2014.
[40] X. Xiao *et al.*, "Mapping paddy rice agriculture in southern China using multi-temporal MODIS images," *Remote Sens. Environ.*, vol. 95, no. 4, pp. 480–492, 2005.
[41] L. Zhong, P. Gong, and G. S. Biging, "Efficient corn and soybean mapping with temporal extendability: A multi-year experiment using Landsat imagery," *Remote Sens. Environ.*, vol. 140, pp. 1–13, 2014.
[42] A. P. VanDeventer, A. D. Ward, P. H. Gowda, and J. G. Lyon, "Using thematic mapper data to identify contrasting soil plains and tillage practices," *Photogramm. Eng. Remote Sens.*, vol. 63, no. 1, pp. 87–93, Jan. 1997.

Qiong Hu received the B.S. degree from Huazhong Agricultural University, Wuhan, China, in 2012. She is currently working toward the Ph.D. degree with the Key Laboratory of Agri-informatics, Ministry of Agriculture/Institute of Agricultural Resources and Regional Planning, Chinese Academy of Agricultural Sciences, Beijing, China.

Her research interests include the crop mapping using time-series Moderate Resolution Imaging Spectroradiometer images, global cropland change research based on Globeland30 data, and spatiotemporal dynamics of crop planting patterns.

Wenbin Wu received the B.S. degree in geography from Huazhong Normal University, Hubei, China, in 1998; the M.Eng. degree in environmental engineering from the Chinese Academy of Agricultural Sciences, Beijing, China, in 2005; and the Ph.D. degree in agricultural spatial informatics from the University of Tokyo, Tokyo, Japan, in 2008.

He is currently a Research Professor with the Key Laboratory of Agri-informatics, Ministry of Agriculture/Institute of Agricultural Resources and Regional Planning, Chinese Academy of Agricultural Sciences. His research interests include agricultural remote sensing, monitoring and modeling of changes in agricultural land systems by integrating remote sensing, geographic information systems, and statistical methodologies.

Qian Song received the B.S. degree in geographic information systems and the M.Eng. degree in forestry management from Northeast Forestry University, Harbin, China, in 2010. She is currently working toward the Ph.D. degree with the Key Laboratory of Agri-informatics, Ministry of Agriculture/Institute of Agricultural Resources and Regional Planning, Chinese Academy of Agriculture Sciences, Beijing, China.

She is also with the Remote Sensing Technology Center, Heilongjiang Academy of Agriculture Sciences, Harbin. Her research interests include the cropping pattern mapping based on multiseason GF-1 images, Africa's cropland change research based on Globeland30 data, and spatiotemporal dynamics analysis of crop planting intensity.

Qiangyi Yu received the M.Sc. and Ph.D. degrees in agricultural resources and environment from the Chinese Academy of Agricultural Sciences, Beijing, China, in 2010 and 2013, respectively.

He is currently an Assistant Research Fellow with the Key Laboratory of Agri-informatics, Ministry of Agriculture/Institute of Agricultural Resources and Regional Planning, Chinese Academy of Agricultural Sciences, and a visiting Postdoctoral Researcher with the Environmental Geography Group, Vrije Universiteit Amsterdam, Amsterdam, The Netherlands. His research interests include land system mapping and modeling.

Miao Lu received the B.S. degree in photogrammetry and remote sensing from the PLA Information Engineering University, Henan, China, in 2007 and the Ph.D. degree in photogrammetry and remote sensing from Wuhan University, Hubei, China in 2014.

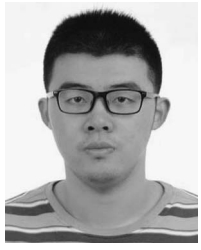
She is currently a Postdoctoral Researcher with the Key Laboratory of Agri-Informatics, Ministry of Agriculture/Institute of Agricultural Resources and Regional Planning, Chinese Academy of Agricultural Sciences. Her research interests include land-cover mapping and change detection based on remote sensing images.

Peng Yang received the B.S. degree in environment science from Wuhan University, Hubei, China, in 1996; the M.Eng. degree in environmental engineering from the Chinese Academy of Agriculture Sciences, Beijing, China, in 2000; and the Ph.D. degree in civil engineering from the University of Tokyo, Tokyo, Japan, in 2005.

In 2000, he joined the Institute of Agricultural Resources and Regional Planning, Chinese Academy of Agricultural Science, Beijing. In 2013, he was promoted to Professor with the Key Laboratory of Agri-informatics, Ministry of Agriculture/Institute of Agricultural Resources and Regional Planning. His research interests include remote sensing applications in the field of agriculture and global change.

Huajun Tang received the M.Sc. and Ph.D. degrees in land sciences from Ghent University, Ghent, Belgium, in 1987 and 1991, respectively.

From 1991 to 1993, he was a Postdoctoral Researcher with Ghent University and was involved in the study of quantitative land evaluation and geographic information system applications. In 1993, he joined the Institute of Agricultural Resources and Regional Planning, Chinese Academy of Agricultural Sciences. In 1996, he was promoted to the General Director of this institute. He is currently a Research Professor, the Deputy Secretary of the Party Leadership Group, and the Vice-President of Chinese Academy of Agricultural Sciences. His research interests include better utilization of agricultural land resources, spatial distribution of crops, and changes in agricultural cropping structures.



Yuqiao Long received the B.S. degree from the Chengdu University of Information Technology, Chengdu, China, in 2015. He is currently working toward the Master's degree with the Chinese Academy of Agriculture Sciences, Beijing, China.

His research interests include the agricultural landscape and cropland use/change based on remote sensing and geographic information systems.