Crop Mapping Based on Temporal and Spatial Sample Migrations: A Case Study Over Three Counties in Heilongjiang Province, Northeast China

Hao-Nan Zuo, Pei Leng, *Senior Member, IEEE*, Yu-Xuan Li, Qian Song, and Zhao-Liang Li, *Fellow, IEEE*

***Abstract*—Crop mapping is crucial for agricultural management and yield prediction. Currently, remote sensing-based crop mapping over a large region is still challenging due to the requirement of sufficient in-season crop samples, which is commonly costly and time-consuming. To address this challenge, a spatial and temporal sample migration method was proposed and evaluated in three typical agricultural counties in Heilongjiang province, Northeast China. On one hand, ground crop samples collected from the previous two years (2020 and 2021) in Nenjiang County were temporally migrated to the target year (2022). On the other hand, ground crop samples collected from the year 2022 in Fujin County were spatially migrated to the adjacent Tongjiang County for mapping crops for the same year. This enabled crop mapping in the absence of current crop samples. In the proposed method, the Sentinel-2 data were primarily used to obtain target curves and reference curves for crop samples. In addition, by balancing the quantity and quality of migrated samples, an optimal migration rule was designed to obtain migrated samples using the dynamic time-warping algorithm over the study area. Finally, the migrated crop samples were used to mapping crop distribution for the target year and region. The results indicated that the overall accuracy can reach 95.7% with the temporal migration in Nenjiang, whereas the spatial migration from Fujin to Tongjiang was approximately 75.6%. The proposed approach reveals significant potential for crop mapping without the knowledge of in-season crop samples, especially for the use of historical crop samples for temporal migration.**

***Index Terms*—Crop mapping, historical crop samples, Sentinel-2, spatiotemporal migration.**

I. Introduction

WITH the implementation of numerous remote sensing monitoring projects, agricultural remote sensing has been greatly promoted, leading to significant advancements in the extraction of spatial distribution information of crops in terms of theory, methodology, and applications [1], [2], [3]. Compared to traditional manual field survey, the use of remote sensing technology was considered an effective tool for mapping crop spatial distribution and estimating crop area [4]. Nowadays, with the development and application of machine learning, an increasing number of researchers were combining machine learning with remote sensing for crops mapping, achieving significant progress in various disciplines, and providing unprecedented capabilities for understanding and managing the Earth's surface. The integration of machine learning can automatically generate high-resolution maps and perform complex land cover classification tasks faster and more accurately than traditional manual methods [5], [6]. Machine learning algorithms, particularly deep learning models such as convolutional neural networks, have shown significant performance in extracting features from remote sensing images and distinguishing different land cover categories. By analyzing multispectral and hyperspectral data, these algorithms can detect subtle patterns and spectral features that may not be discernible to the human eye, thereby creating detailed and rich information maps [7], [8]. In addition, many scholars integrated auxiliary data sources such as terrain information, climate data, or socioeconomic indicators into machine learning models, which can better consider spatial variability and temporal dynamics, thereby improving the accuracy and robustness of land cover classification [9]. However, the sole use of machine learning combined with remote sensing data requires high demands on remote sensing data, necessitating high-quality datasets as training support [10].

Due to the discrepancy in plant pigments, water content, leaf cell structure, and varying growth seasons or regional planting systems, the spectral characteristics of crops were evidently different. This often led to the use of spectral and vegetation index temporal features for extracting spatial distribution information of crops [11]. In cases where planting structures were complex or imaging data were missing, crop classification, and mapping become more difficult, especially for crops with similar spectral characteristics, efficient utilization of historical data becomes essential to compensate for data deficiencies and aid in crop classification mapping. Many scholars employed various approaches to utilize temporal data to increase data utilization rates and combined algorithms to achieve accurate crop mapping. Some scholars focused on advanced data processing [12], [13], [14], or fitting time series of remote sensing data over long periods and combining them with machine learning algorithms for

Manuscript received 28 April 2024; revised 14 June 2024; accepted 5 July 2024. Date of publication 11 July 2024; date of current version 26 August 2024. This work was supported in part by the National Natural Science Foundation of China under Grant 42271384 and in part by Beijing Natural Science Foundation under Grant 6242030. *(Corresponding author: Pei Leng.)*

The authors are with the State Key Laboratory of Efficient Utilization of Arid and Semi-arid Arable Land in Northern China/Institute of Agricultural Resources and Regional Planning, Chinese Academy of Agricultural Sciences, Beijing 100081, China (e-mail: 82101232133@caas.cn; lengpei@caas.cn; 82101222292@caas.cn; songqian01@caas.cn; lizhaoliang@caas.cn).

Digital Object Identifier 10.1109/JSTARS.2024.3426671

precise classification [15]. In addition, another group of scholars concentrated on improving algorithms [16], [17]. However, most scholars only considered the temporal attributes of current data, making it difficult to achieve crop mapping when there was a lack of accurate data for the current year.

Transfer learning aimed to improve the performance of a target learner in a target domain by transferring knowledge from different but related source domains, which was an effective way for fully using historical or spatially adjacent samples [18]. However, most of the existing literatures on transfer learning mainly concerned land cover mapping. In contrast, due to the effects of human activities on croplands, transfer learning-based crop mapping received less attention. Currently, the Euclidean distance, spectral angle distance, and dynamic time warping (DTW) have been frequently used in transfer learning to obtain target samples from historical information. Specifically, Compared to traditional Euclidean distance methods, DTW could match data that were inconsistent in time or length and reduced the effects of noise such as cloud masking and weather nudging [19], [20]. Meanwhile, the commonly used transfer features for crop mapping mostly utilize vegetation indices, and the transferred features included spectral, vegetation indices, texture, and terrain [21], [22]. Through temporal migration of historical samples, many scholars transferred invariant training samples from a reference year to their specified target year, then crop mapping with satisfactory accuracy for the target year can be obtained by the migrated crop samples, making crop mapping no longer rely on in-season crop samples. For instance, by using the time-weighted dynamic time warping method to measure the spectral-temporal discrepancies between labeled samples from the source area and unlabeled samples from the target area, labeled crop datasets can be automatically generated over the target area. This can effectively solve the problem of low inter-class similarity for crop mapping, making significant progress in the field of crop migration mapping [23]. However, compared to the temporal migration of crop samples, most scholars have conducted limited exploration into spatial sample migration. A portion of scholars have used specific models, such as generating and updating corn and soybean distribution maps in different regions through the image segmentation models [24]. Nevertheless, most scholars have focused on improving certain deep-learning methods. For instance, they have enhanced convolutional neural networks to depict agricultural fields in different regions using multimodal remote sensing data [25], or improved deep learning models to study one [26], [27], or two crops [28] by extracting specific features of these crops for spatial migration, achieving relatively accurate mapping results.

So we defined crops as having not only temporal attributes but also unique spatial attributes, especially for the regions with similar planting patterns. The spatial and temporal migration of crop samples provided a richer perspective for observing and interpreting the dynamic nature of crops. Specifically, compared to the temporal sample migration that normally occurs in a fixed area, spatial sample migration was more expected because it may provide an alternative for crop mapping across adjacent regions with similar planting patterns, which can significantly decrease the demand for crop mapping with sufficient local crop samples. To this end, the present study will aim to investigate both temporal and spatial sample migrations for mapping crop without the knowledge of current crop samples, which is especially feasible for regions with similar planting patterns. Section II will present the study area and data. Section III will describe the methods used in this study. The results and discussion will be presented in Sections IV and V, respectively, and Section VI concludes this article.

II. Study Area and Data

A. Study Region

In this article, three typical agricultural counties of Nenjiang, Fujin, and Tongjiang in Heilongjiang Province, were selected as the study area (see Fig. 1). Nenjiang County was geographically located in the west of Heilongjiang Province, whereas Fujin and Tongjiang were two adjacent counties located in the east of Heilongjiang Province. The primary reason for selecting the three counties as study areas was that they were important commercial grain bases of the Heilongjiang Province. Specifically, the major corps in Nenjiang county were corn and soybean, whereas rice, corn, and soybean were dominated in both Fujin county and Tongjiang county. All three major crops were characterized by an annual ripening system, and the growing seasons of the three crops in the study areas were nearly the same. Concretely, corn and soybean were generally sown at the middle to end of April, and the harvest period was from late September to early October, whereas rice was generally sown in the early to middle April, and the harvest period was about the early October. Hence, either for the temporal sample migration in Nenjiang County or the spatial sample migration from Fujin County to Tongjiang County, the planting patterns were nearly the same, which was also an important reason for the selection of the study area.

B. Sentinel-2 Images

In order to conduct real-time monitoring of land, ocean, natural disasters, and other natural phenomena, the European Space Agency launched the Sentinel-2A and Sentinel-2B satellites on 23 June 2015 and 7 March 2017, respectively. Both Sentinel-2A and Sentinel-2B were equipped with a multispectral instrument consisting of 13 channels for multispectral imaging. Each satellite had a time resolution of ten days and was revisited every five days. We used the level-2 product of Sentinel-2 data on the Google Earth engine (GEE) platform, which contained atmospherically corrected surface reflectance in the visible, near-infrared, red-edge, water vapor, cirrus, and short-wave infrared bands, with resolutions of 10, 20, and 60 m in each band, respectively. The spectral curves of reflected light from plants and other surfaces differ significantly in the visible and near-infrared wavelengths, providing the basis for remote sensing to identify and extract unique vegetation features. Compared to using single-band information, features derived from multiple bands were beneficial for crop identification [29]. Furthermore, to avoid the influence of low-resolution bands on crop sample migration, we discarded bands in the water vapor, cirrus cloud, and visible light with a relatively coarse spatial resolution of 60

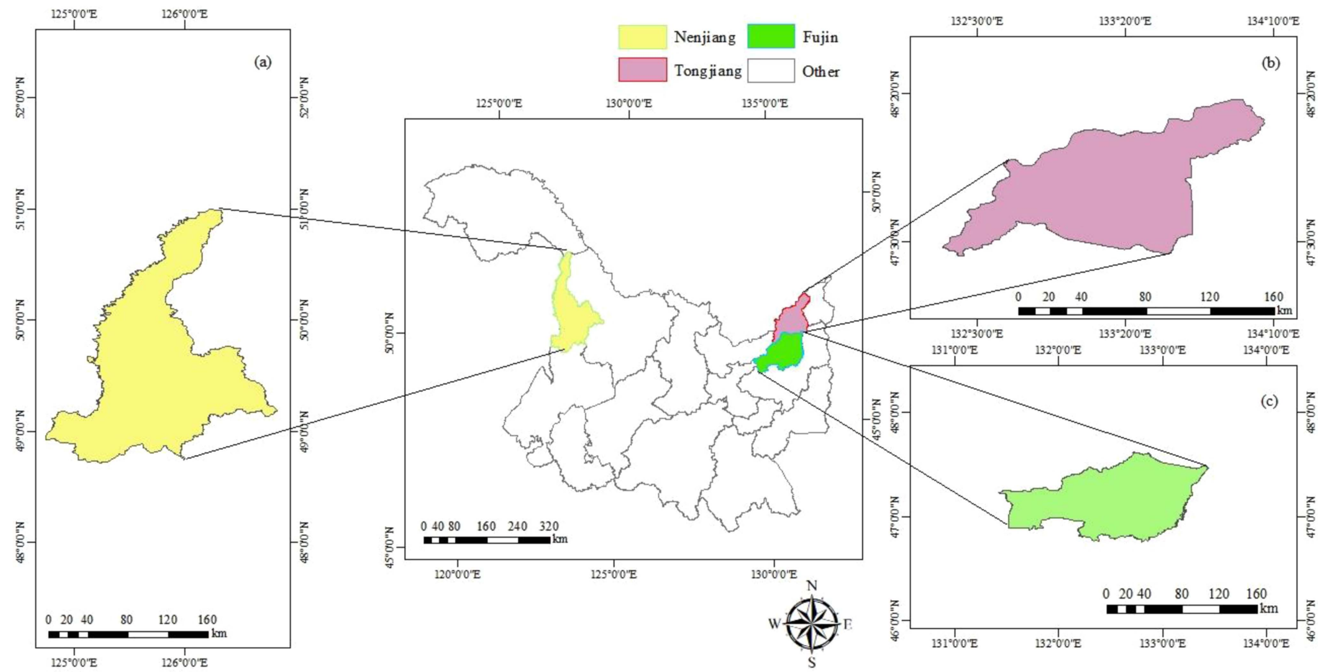

Fig. 1. Study area of three typical agricultural counties in Heilongjiang Province. (a) Nenjiang. (b) Tongjiang. (c) Fujin.

TABLE I
DESCRIPTION OF SPECTRAL CHARACTERISTICS OF THE SENTINEL-2 DATA

Band	Name	Central wavelength (nm)	Spatial resolution (m)
Band2	BLUE	496.6(S2A) / 492.1 (S2B)	10
Band3	GREEN	560 (S2A) / 559 (S2B)	10
Band4	RED	664.5 (S2A) / 665 (S2B)	10
Band8	NIR	835.1 (S2A) / 833 (S2B)	10
Band11	SWIR1	1613.7 (S2A) / 1610.4 (S2B)	20
Band12	SWIR2	2202.4 (S2A) / 2185.7 (S2B)	20

m. Finally, we selected six bands-Band2 (Blue), Band3 (Green), Band 4 (Red), Band 8 (NIR), Band 11 (SWIR1), and Band 12 (SWIR2)-with higher spatial resolution closely related to crops in the present study, as shown in Table I.

C. Ground Survey Data

Ground field surveys were conducted in a period of three years from 2020 to 2022. We recorded the crop types by taking photos with the camera and obtained specific coordinate locations through OvitalMap software. The investigated ground features included rice, corn, soybean, and others. It noted that various other features were grouped into "Other" in the crop mapping results because the concerned crops (rice, corn, and soybean) account for a vast majority of the study region. Another reason was that the present study mainly focused on the temporal and spatial sample migration for crop mapping with the three major crops. In the process of selecting the samples for each crop, we tried to collect as many as possible in all ranges based on the principles of randomness and uniformity. In Nenjiang County (see Table II), there were 311 soybean samples and 357 corn samples in 2020, with an additional 20 samples for Other. In 2021, 358 soybean samples, 300 corn samples, and 21 other samples were collected. Specifically, these samples will be used for temporal sample migration. As for the Fujin County (see Table III), 335 rice samples, 147 soybean samples, 145 corn samples, and 28 other samples were collected in 2022. For Tongjiang County, we utilized the randompoints function in the

TABLE II
GROUND SAMPLES COLLECTED IN NENJIANG

Year	Soybean	Corn	Other	Total
2020	311	357	20	688
2021	358	300	21	679

TABLE III
GROUND SAMPLES COLLECTED IN FUJIN

Year	Soybean	Corn	Rice	Other	Total
2022	147	145	335	28	655

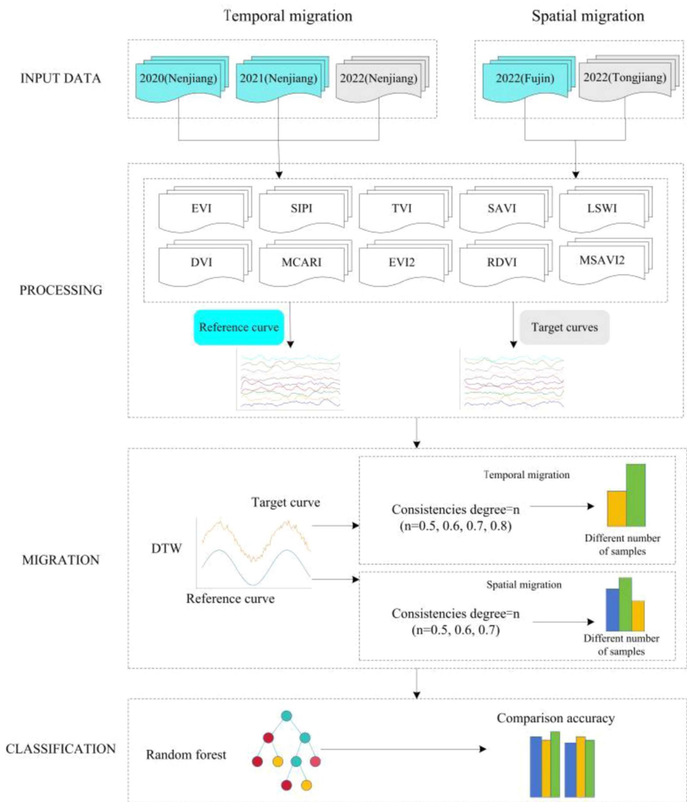

Fig. 2. Flowchart for crop mapping with temporal and spatial sample migration.

GEE platform to randomly generate 1000 unknown samples for investigating spatial sample migration.

III. Methodology

A. Temporal and Spatial Sample Migrations

By using the temporal and spatial sample migration, the method proposed in this article was expected to mapping crops without the knowledge of current samples. Fig. 2 depicts the flowchart of the present study. In general, the entire method consisted of four steps in two parts. The first part was temporal sample migration, in which ground crop samples collected from the previous two years (2020 and 2021) in Nenjiang County were primarily used to obtain reference curves for crops. Then, the distance between the target curves of the migrated samples and the reference curves was calculated through DTW, with minimum distance regarded as indicative of the similarity between the two curves. Furthermore, to obtain migrated crop samples for crop mapping in the target year (2022), a threshold value was set by balancing the number of samples and accuracy for crop mapping. Specifically, we considered the minimum cumulative distance as favoring a particular crop type, and the final successfully migrated crop samples were determined by the determination of the consistency degree defined in the following section. Finally, the successfully migrated crop samples were used for crop mapping in the target year via random forest (RF) algorithm. As for the second part of spatial migration, following the similar procedure, ground crop samples collected from the year 2022 in Fujin County were spatially migrated to the adjacent Tongjiang County for mapping crops for the same year. Specifically, the Sentinel-2 sample data constructed in 2022 in Fujin County and the randomly generated sample data in Tongjiang County were primarily conducted to obtain feature sequences. Further, the distance between the two feature sequences was calculated using the DTW algorithm for label matching. Similar to temporal migration, a consistency degree was determined to obtain the finally migrated crop samples from Fujin County to Tongjiang County. The following sections will illustrate the key steps of the proposed approach in detail.

B. Determination of Crop Features

In the proposed approach, optimal crop features were required to distinguish various crops. To this end, we selected ten features from multiple spectral bands sensitive to the growth response of major crops using the advantage of high spatial and temporal resolution and a large number of spectral bands of the Sentinel-2 data, which was beneficial to minimize the effects of soil background, atmospheric, and solar targeting sensors. Specifically, some of these features were recognized as helpful and advantageous for the migration identification of maize, soybean, and rice [30], [31], [32], [33], [34], [35], [36]. The ten features were enhanced vegetation index (EVI), renormalized difference vegetation index, triangular vegetation index (TVI), land surface water index (LSWI), soil-adjusted vegetation index (SAVI), modified chlorophyll absorption ratio index (MCARI), structure insensitive pigment index (SIPI), difference vegetation index (DVI), enhanced vegetation index 2 (EVI2), and modified soil-adjusted vegetation index 2 (MSAVI2), which can be written as follows:

$$\mathrm{EVI} = 2.5 \times \frac{\rho_{\mathrm{nir}} - \rho_{\mathrm{red}}}{\rho_{\mathrm{nir}} + 6 \times \rho_{\mathrm{red}} - 7.5 \times \rho_{\mathrm{blue}} + 1} \tag{1}$$

$$\mathrm{RDVI} = \frac{\rho_{\mathrm{nir}} - \rho_{\mathrm{red}}}{\sqrt{\rho_{\mathrm{nir}} + \rho_{\mathrm{red}}}} \tag{2}$$

$$\mathrm{TVI} = 0.5 \times [120 \times (\rho_{\mathrm{nir}} - \rho_{\mathrm{green}})] - 200 \times (\rho_{\mathrm{red}} - \rho_{\mathrm{green}}) \tag{3}$$

$$\mathrm{LSWI} = \frac{\rho_{\mathrm{nir}} - \rho_{\mathrm{swir}}}{\rho_{\mathrm{nir}} + \rho_{\mathrm{swir}}} \tag{4}$$

$$\mathrm{SAVI} = 1.5 \times \frac{\rho_{\mathrm{nir}} - \rho_{\mathrm{red}}}{\rho_{\mathrm{nir}} + \rho_{\mathrm{red}} + 0.5} \tag{5}$$

$$\mathrm{MCARI} = [(\rho_{\mathrm{nir}} - \rho_{\mathrm{red}}) - 0.2 \times (\rho_{\mathrm{nir}} - \rho_{\mathrm{green}})] \times \frac{\rho_{\mathrm{nir}}}{\rho_{\mathrm{red}}} \tag{6}$$

$$\mathrm{SIPI} = \frac{\rho_{\mathrm{nir}} - \rho_{\mathrm{blue}}}{\rho_{\mathrm{nir}} - \rho_{\mathrm{red}}} \tag{7}$$

$$\mathrm{DVI} = \rho_{\mathrm{nir}} - \rho_{\mathrm{red}} \tag{8}$$

$$\mathrm{EVI2} = 2.5 \times \frac{\rho_{\mathrm{nir}} - \rho_{\mathrm{red}}}{\rho_{\mathrm{nir}} + 2.4 \times \rho_{\mathrm{red}} + 1} \tag{9}$$

$$\mathrm{MSAVI2} = \frac{2 \times \rho_{\mathrm{nir}} + 1 - \sqrt{(2 \times \rho_{\mathrm{nir}} + 1)^2 - 8 \times (\rho_{\mathrm{nir}} - \rho_{\mathrm{red}})}}{2} \tag{10}$$

where ρ_{blue}, ρ_{green}, ρ_{red}, ρ_{nir}, and ρ_{swir} are reflectance of the blue, green, red, near-infrared, and shortwave-infrared bands, respectively.

C. Dynamic Time-Regularized Migration Decisions

The DTW was a powerful algorithm widely used for aligning and comparing reference and target data, especially in cases where traditional distance measurement methods failed to meet requirements. Its core competency lies in handling time shifts, speed changes, or nonlinear distortions that may occur within sequences. The working principle of DTW involved dynamically adjusting local time warping to find the optimal alignment path between two feature sequences. The algorithm entailed constructing a cost matrix to capture dissimilarities between individual data points. Utilizing dynamic programming, it computed the cumulative cost matrix and determined the best warping path by minimizing the total accumulated cost. Local transformations such as insertion, deletion, and substitution enabled DTW to flexibly align sequences and captured subtle temporal variations. In addition to conventional uses, DTW was also applied for pixel-based or logic-based crop mapping, and it held particular significance in temporal and spatial transfer analysis. Whether in ecological studies or geographical trajectory analysis, the adaptability of this algorithm to migration made it an indispensable tool for capturing dynamic patterns in time-varying sequences. In the present study, we applied DTW to remote sensing data to investigate its effectiveness in temporal and spatial sample migrations.

For temporal sample migration, we first obtained the reference curves of known samples (including soybean, corn, and other) from 2020 and 2021 in Nenjiang County. Further, the DTW distance between the curves of unknown samples in 2022 and the known samples were calculated, in which the minimum DTW distance value was assigned to the corresponding crop type. For spatial sample migration, the reference curves were determined from Fujin County with the samples collected in 2022. Then, the DTW distance between the curves of samples randomly generated in Tongjiang County and those of the Fujin County was calculated. Similarly, as in temporal sample migration, the randomly generated sample was assigned a crop type when the minimum DTW distance occurred. In the present study, the consistency degree (Ω) was defined to represent the agreement between all the reference curves and migrated curves. Specifically, by calculating the distance between each feature curve and the standard curve, we can determine which crop type each feature in the sample was more inclined toward. Then, the consistency degree can be obtained by taking the maximum value of the proportion with the judgment for each feature. For example, in a sample from Nenjiang, if seven out of the ten features judged it as soybean and three features judged it as corn, we considered its consistency degree to be 0.7. The consistency degree was written as follows:

$$\Omega = \left\{ \frac{\sum_{i=1}^{n} \Phi_{\mathrm{i}}(x_j)}{n} \right\}_{\max} \tag{11}$$

where n is the total number of features selected, j indicates different crop categories, and $\Phi(x)$ denotes the crop categories judged by the particular feature.

D. Accuracy Assessment

In this study, the metrics of user accuracy (UA), producer accuracy (PA), and overall accuracy (OA) were conducted to assess the accuracy of the proposed approach for crop mapping over the study area. Among these metrics, OA can adequately reflect the comprehensive accuracy of the results, while PA and UA can be used to evaluate the classification accuracy of specific crop types. The three metrics can be written as follows:

$$\mathrm{PA} = \frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FN}} \times 100\% \tag{12}$$

$$\mathrm{UA} = \frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FP}} \times 100\% \tag{13}$$

$$\mathrm{OA} = \frac{\mathrm{TP}+\mathrm{TN}}{\mathrm{TP}+\mathrm{FP}+\mathrm{TN}+\mathrm{FN}} \times 100\% \tag{14}$$

where TP and TN represent true positive and true negative, respectively; FP and FN represent false positive and false negative, respectively.

IV. Results

A. Correctly Migrated Samples With Varying Consistency Degree

Due to the relatively similar growth cycles of corn and soybean, and to differentiate them from other crops such as rice, we utilized the DTW algorithm to quantify the similarities and dissimilarities in migration patterns between different crop types and consistency degrees. Another motivation was that the consistency degree was commonly subjective, and it could affect the number of correctly migrated samples and the accuracy of crop mapping. This enabled a detailed understanding of the migration results of soybean, corn, rice, and other crops at different consistency degrees.

In temporal sample migration, as shown in Fig. 3, it can be observed that at a lower consistency degree of 0.5, the migration quantity for all crop types was relatively high, whereas with increasing consistency degree, the migration quantity for all crop types decreased. When the consistency degree was at a higher level of 0.8, the migration sample quantities for both corn and soybean fell into the single digits, indicating a very low migration quantity, and with further an increase in consistency degree, no migration samples meeting the requirements were available. Similar phenomena can also be found in spatial sample migration, suggesting that certain consistency degrees should be set for balancing the number of correctly migrated samples and the accuracy for crop identification, because fewer samples can commonly lead to decreased accuracy of crop mapping [37]. As for the spatial sample migration, at a lower consistency degree of 0.5, rice revealed the highest correct migration, followed by Other, corn, and soybean. This indicates significant success rate of rice samples can be achieved across the two adjacent counties of Fujin and Tongjiang. When the consistency degree was set

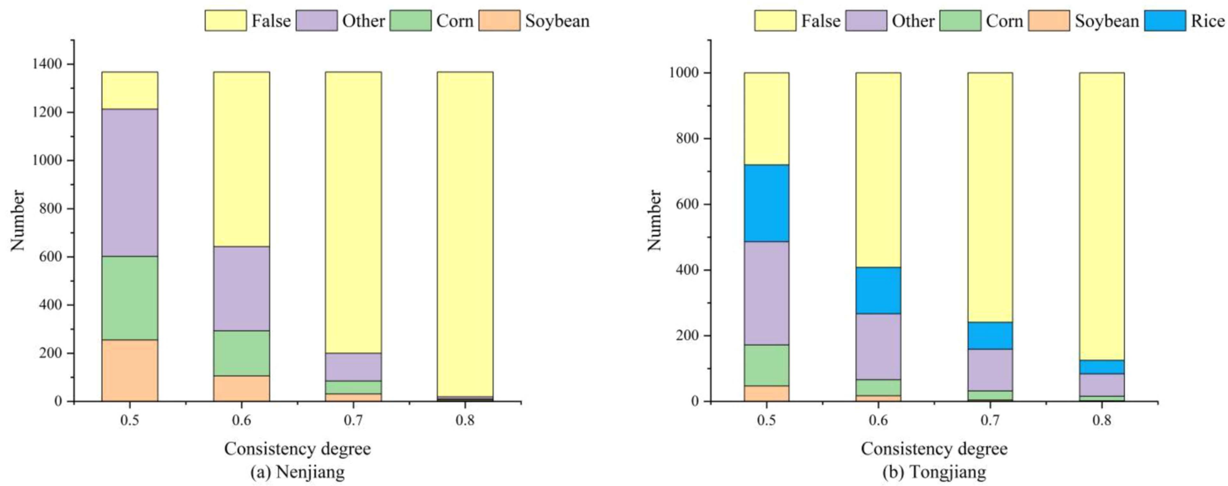

Fig. 3. Temporal and spatial migration in the number of crops migrating under different consistency degree. (a) Temporal migration. (b) Spatial migration.

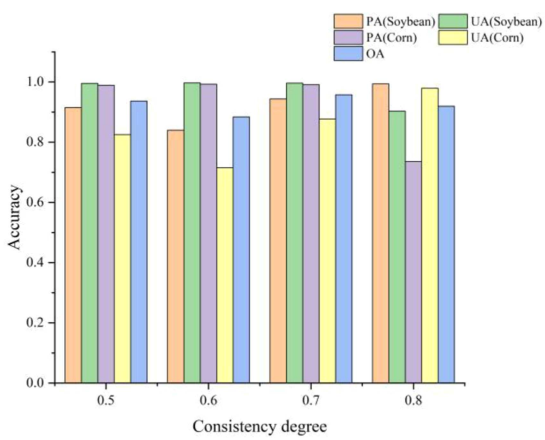

Fig. 4. Overall precision of temporal migration and UA and PA for each crop.

at a high level of 0.7, both corn and soybean experienced a significant decrease in the quantity of correctly migrated samples, indicating a higher failure rate under these circumstances.

B. Crop Mapping With Temporal Sample Migration

Changes in consistency degree led to differences in the number of successfully migrated samples, which also affected the subsequent crop mapping and accuracy assessment. After obtaining migrated samples at different consistency degrees, we used the RF classifier to obtain crop distribution over the target year and region. Based on the migration quantity obtained earlier, we considered four consistency degrees (0.5, 0.6, 0.7, and 0.8) in temporal sample migration. Fig. 4 depicts the OA, PA, and UA for migrated samples with the four consistency degrees. It was evident that the OA ranges from about 88.4% to 95.7%, with the highest OA occurring at a consistency degree of 0.7. This indicates that the accuracy of migrated samples reached a relatively high degree, with a considerable proportion of instances correctly migrated and identified. In addition, the PA and UA varied from approximately 84.0% to 99.4% and from 71.5% to 99.5%, respectively. Based on these results, it was determined that a consistency degree of 0.7 was optimal in the study area for temporal sample migration. In this situation, the OA, PA, and UA for corn and soybean were generally better than those with other consistency degrees.

With the determination of the consistency degree, crop mapping in the target year of 2022 in Nenjiang County was further conducted with the RF algorithm. Fig. 5 depicts the spatial distribution of corn, soybean, and other using temporal migrated samples with different consistency degrees. It was obvious that the result with a consistency degree of 0.8 was very different from the results obtained with other consistency degrees. The PA was with a low level of 73.6%, resulting in a noticeably lower accuracy. Furthermore, for the cases with consistency degrees of 0.5 and 0.6, it can be observed that soybean was significantly less than that with a consistency degree of 0.7. The PA for soybean in both Fig. 5(a) and (b) was lower than that in Fig. 5(c), indicating that (a) and (b) have a higher number of misclassifications of soybean as corn compared to (c). The lower OA in Fig. 5(b) was mainly due to the higher number of misclassifications of soybean as corn compared to other situations, with a PA of 84.0% for soybean, whereas in other cases, the PA of soybean was significantly higher than 90%. With the increase of consistency degree, it was found that the improvement of soybean identification was more significant than that of corn. This may be attributed to the more correctly migrated soybean samples than the corn samples. Combining the results of OA and field inspection, a consistency degree of 0.7 in temporal sample migration over the study area appears to be an optimal one in the present study. Fig. 6 depicts the invariable rate of migrated samples while changing the consistency degree from one value to another. Taking a presupposed consistency degree of 0.5 as an example, changing the consistency degree to higher values (0.6, 0.7, or 0.8) implies more stringent requirements for determining successfully migrated crop samples. This can logically result in a decrease in the quantity of migrated samples. Overall, we found that above 80% migrated soybean and corn samples remain stable when the consistency degree improved

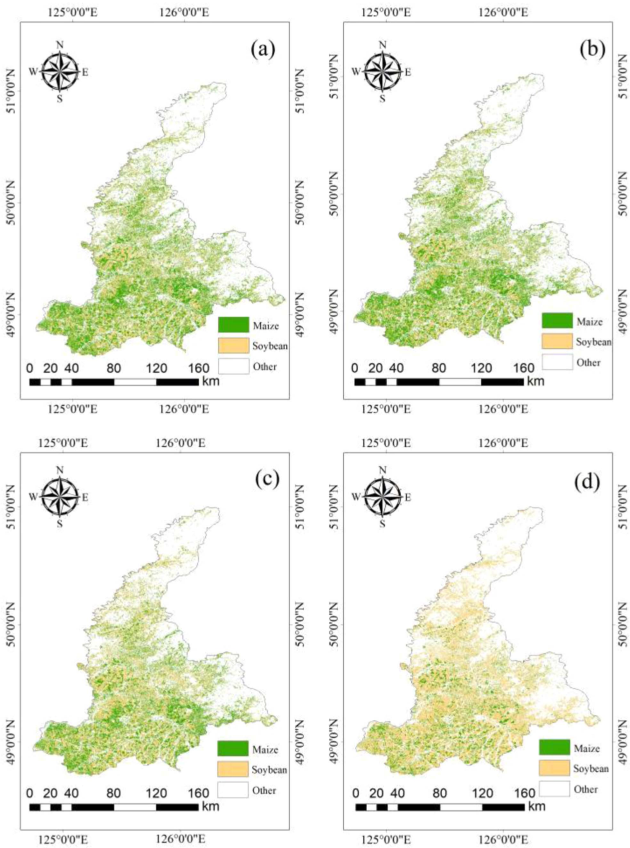

Fig. 5. Crop mapping using temporal sample migration with different consistency degree. (a) Consistency degree = 0.5. (b) Consistency degree = 0.6. (c) Consistency degree = 0.7. (d) Consistency degree = 0.8.

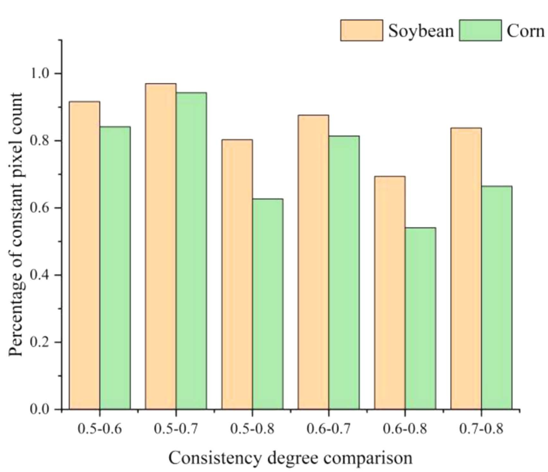

Fig. 6. Percentage of the constant pixels when changing consistency in temporal migration.

to 0.7 from 0.5. When the consistency degree was set as a high value of 0.8, the significant rate of approximately 30.7% for corn samples would change to another land cover type, which was most likely to affect the stability of crop identification. Hence, the determination of the consistency degree not only influenced the quantity of migrated samples but also affected the accuracy of crop mapping. These results further indicated that a consistency degree of 0.7 was optimal for the temporal sample migration over the study area.

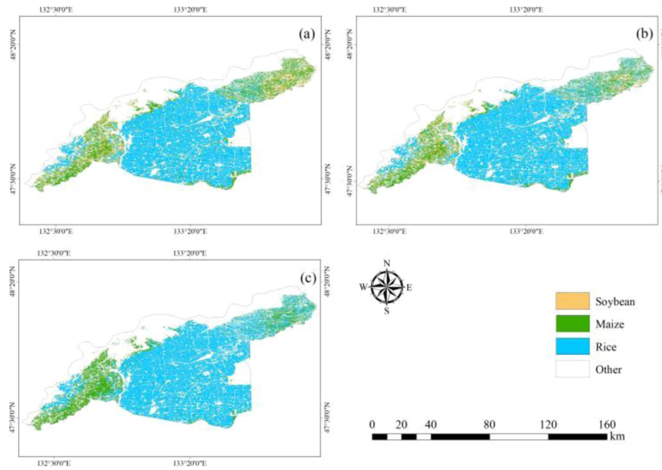

Fig. 7. Crop mapping using spatial sample migration with different consistency degree. (a) Consistency degree ≥0.5. (b) Consistency degree ≥0.6. (c) Consistency degree ≥0.7.

C. Crop Mapping With Spatial Sample Migration

Spatial migration was similar to temporal migration, and we used OA as the evaluation metric. Based on the migration amounts obtained earlier, we considered three degrees of consistency (0.5, 0.6, and 0.7) in spatial sample migration. In spatial migration, when the consistency degree was 0.5, the OA was approximately 64.8%, while at a consistency degree of 0.6, the highest OA was achieved at around 75.6%. However, after consistency increased to 0.7, the OA was less than 50%. This indicated that at a consistency of 0.6, the accuracy of migrated samples reached a relatively good level, with a sufficient proportion of instances being correctly migrated and identified. Therefore, we determined a consistency degree of 0.6 as the optimal value for spatial sample migration in the study area. Under this circumstance, corn, soybean, and rice can be effectively identified through spatial sample migration.

After the determination of the consistency degree, the RF algorithm was further used to map the crops in Tongjiang County in 2022. Fig. 7 depicts the spatial distribution of corn, soybean, rice, and other using spatially migrated samples with different consistency degrees. It was evident that the results at a consistency degree of 0.7 differ significantly from those at other consistency degrees. The classification accuracy of soybean was notably lower because many soybean samples have not been successfully migrated over the study area, and some soybean samples were incorrectly identified as rice. In addition, based on the results of different consistency degrees in Fig. 7(a)–(c), it can be observed that rice in the central part of the study region was relatively stable compared to other areas, showing condensable accuracy with the migrated samples under different conditions. Considering the OA and field investigation results, a consistency degree of 0.6 for spatial sample migration in the study seemed to be optimal.

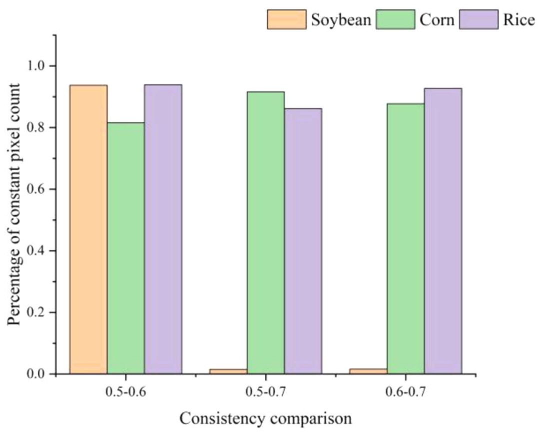

Fig. 8. Percentage of the constant pixels when changing consistency in spatial migration.

Furthermore, we visualized the changes in pixel counts of corn, soybean, and rice under different consistency conditions. Fig. 8 depicts the invariable rate of migrated samples when change the consistency degree from one value to another. Taking the three crops at a consistency degree of 0.5 as an example, when changed the consistency degree to 0.7, although corn and rice samples maintained a relatively stable status, it appeared a dramatic decrease in the migrated soybean samples, indicating that soybean samples would migrate to another crop type under such conditions. By contrast, when enhancing the consistency degree from 0.5 to 0.6, over 80% samples of corn and 85% samples of soybean and rice remain unchanged, which can guarantee the stability of crop identification. These results further confirmed that a consistency degree of 0.6 was optimal for the spatial sample migration over the study area.

V. Discussion

A. Advantages and Challenges

To date, most of the previous literature on sample migration primarily focused on mapping land cover categories such as wetlands and forests, rather than on crop mapping. One of the main reasons was that intensive managements commonly happen to croplands. Moreover, most of the existing methods mainly aimed at temporal migration to utilize historical crop samples for in-season crop mapping. The present study investigated both spatial and temporal crop sample migration over the three typical agricultural counties with similar planting patterns under the same climate conditions. In terms of temporal migration, the present study revealed considerable accuracy with OA of 95.7%, which was generally comparable or better than most of the existing methods for mapping in-season crops with the information from historical samples [38], [39]. Meanwhile, although a decreased accuracy was obtained for spatial sample migration, the in-season maps of corn, soybean, and rice over Tongjiang County were expected to reach an OA over 75% without the knowledge of local crop samples, indicating significant potential for crop mapping when local crop samples were insufficient. Compared to scholars using the DTW method for sample migration mapping, the temporal migration accuracy in our study is comparable or even better. What stands out is our increased emphasis on spatial research, discussing the feasibility of spatial migration mapping for various different crops, and achieving comparatively satisfactory results. From another perspective of the consistency degree, it was evident that the optimal consistency degree for spatial migration was lower than that in temporal migration, indicating that spatial sample migration was more complicated. With the increase of consistency degree, the quantity of successfully migrated samples significantly decreased in theory, leading to imbalanced sample proportions between crops. This was more obvious for spatial migration, in which soybean samples were more likely to be confused with other crops. Following the results at different consistency degrees, it was found that OA would not consequentially increase with a higher consistency degree; instead, an optimal consistency degree existed for balancing the quantity and quality of migrated samples. Specifically, we maximized the utilization of limited crop samples by matching reference curves of various features and decided on an adequate consistency degree for migration. Nevertheless, the determination of the consistency degree seemed a challenging task for migrating samples, because it depended on specific conditions. Another challenge of the proposed approach was the feasibility over extended regions. It was evident that the two adjacent counties of Fujin and Tongjiang were not only characterized with the same climate patterns but also with similar planting patterns where rice, corn, and soybean were the dominant crops. Specifically, the planting areas for corn and soybean were close, and the planting area for rice was significantly more than the other two crops in both Fujin and Tongjiang. When encountering complicated planting patterns and natural conditions (e.g., more crop types, different cropping intensity, fragmented parcels, and complex terrain), the proposed approach may suffer from more challenges because various factors should be considered, rather than the features of crops. As a consequence, the determination of a proper consistency degree is expected to be more difficult under such circumstances, especially for the spatial sample migration. Nevertheless, future work can investigate more in-depth for the proposed approach with varying planting patterns and climate conditions.

B. Uncertainty Analysis

Several sources of errors would constrain the performance of the proposed method, especially for spatial sample migration. First, migrated crop samples inherently faced issues such as misclassification or decreased quantity, fostering internal uncertainty and skewing classification outcomes [40]. This was reasonable because no in-season ground crop sample truths were collected in the target year for temporal migration and region for spatial migration. However, this was also the primary motivation of the present study. As for the decreased quantity of migrated crop samples, it was inevitable because it followed the rule for determining a successfully migrated sample based on the consistency degree. However, increasing the quantity of historical (or adjacent region) crop samples for temporal (or spatial) migration could be an adequate solution to this issue. Second, data quality would significantly influence the performance of crop mapping. A well-acknowledged issue in crop mapping includes sensor noise, atmospheric effect, and cloud cover, especially for optical data [41], [42]. Even with satellite images captured under low cloud cover conditions, noise remained inevitable, particularly when covering extensive geographic regions [43]. Typically, remote sensing images captured during critical crop phenology stages (such as July and August for migration studies) effectively distinguished between crop types [44]. Moreover, the sensor and cloud noise also hindered obtaining continuous and valid time series data, potentially compromising subsequent image fusion quality and the consistency of migrated results. Third, spectral differences, stemming from varied climatic periods of crop varieties, can also impact migration outcomes [45], [46]. For the spatial sample migration, the reduced quantity of migrated soybean samples when changing the consistency degree from 0.5 to 0.6 or from 0.6 to 0.7 would probably due to the fact that they were easily confused with corn [47]. In addition, the existence of mixed pixels and decentralized small land parcels further complicated the effectiveness of crop mapping, because multiple features inevitably appear under such circumstances [48], [49].

VI. Conclusion

In this research, we have investigated crop mapping without the knowledge of current crop samples, which was particularly suitable for crop mapping when in-season or local crop samples were insufficient. The method circumvented traditional annual field crop sample collection, significantly reducing the manual and time costs associated with continuous crop mapping. Specifically, the consistency degree was defined to evaluate the migrated crop samples with respect to both quality and quantity. When we altered the sample size based on the consistency degree, the quality of the samples also changed. Normally, the contradiction between the increased consistence degree and decreased correctly immigrated samples exists, and we have explored the optimal consistency degree under various conditions to ensure the accuracy of the mapping. Evaluation of the proposed approach in three typical agricultural counties in Heilongjiang Province, Northeast China showed satisfactory accuracy, especially in Nenjiang County with an OA of 95.7% for the temporal sample migration. This implied that historical crop samples can be adequately used to obtain in-season crop distribution with the proposed method. Although the results with spatial sample migration revealed a decreased accuracy compared to those with temporal sample migration, it was still promising since the proposed method no longer relies on in-season crop samples. Moreover, it was noted that the proposed method was expected to be feasible with similar planting patterns and climate conditions, which was also the basic hypothesis for the proposed approach. Although that did not necessarily mean the temporal and spatial sample migration method was only feasible under such circumstances, it was predictable that more difficulties existed for the proposed method across varying planting patterns and climate conditions.

References

[1] M. Weiss, F. Jacob, and G. Duveiller, “Remote sensing for agricultural applications: A meta-review,” *Remote Sens. Environ.*, vol. 236, 2020, Art. no. 111402.

[2] R. Sishodia, P. Ram, L. Ray, and S. Singh, “Applications of remote sensing in precision agriculture: A review,” *Remote Sens.*, vol. 12, 2020, Art. no. 3136.

[3] J. Segarra, M. Buchaillot, J. Araus, and S. Kefauver, “Remote sensing for precision agriculture: Sentinel-2 improved features and applications,” *Agronomy*, vol. 10, 2020, Art. no. 641.

[4] X. Song, W. Huang, M. Hansen, and P. Potapov, “An evaluation of Landsat, Sentinel-2, Sentinel-1 and Modis data for crop type mapping,” *Sci. Remote Sens.*, vol. 3, 2021, Art. no. 100018.

[5] M. Pal, “Random forest classifier for remote sensing classification,” *Int. J. Remote Sens.*, vol. 26, pp. 217–222, 2005.

[6] Z. Al-Ali, M. Abdullah, N. Asadalla, and M. Gholoum, “A comparative study of remote sensing classification methods for monitoring and assessing desert vegetation using a UAV-based multispectral sensor,” *Environ. Monit. Assessment*, vol. 192, 2020, Art. no. 389.

[7] J. Xu et al., “Towards interpreting multi-temporal deep learning models in crop mapping,” *Remote Sens. Environ.*, vol. 264, 2021, Art. no. 112599.

[8] B. Zhang, L. Zhao, and X. Zhang, “Three-dimensional convolutional neural network model for tree species classification using airborne hyperspectral images,” *Remote Sens. Environ.*, vol. 247, 2020, Art. no. 111938.

[9] S. Zhang, J. Yang, P. Leng, Y. Ma, H. Wang, and Q. Song, “Crop type mapping with temporal sample migration,” *Int. J. Remote Sens.*, pp. 1–19, 2023, doi: 10.1080/01431161.2023.2192881.

[10] H. Huang, J. Wang, C. Liu, L. Liang, C. Li, and P. Gong, “The migration of training samples towards dynamic global land cover mapping,” *ISPRS J. Photogramm. Remote Sens.*, vol. 161, pp. 27–36, 2020.

[11] H. Li, J. Wan, S. Liu, H. Sheng, and M. Xu, “Wetland vegetation classification through multi-dimensional feature time series remote sensing images using mahalanobis distance-based dynamic time warping,” *Remote Sens.*, vol. 14, 2022, Art. no. 501.

[12] M. Belgiu and O. Csillik, “Sentinel-2 cropland mapping using pixel-based and object-based time-weighted dynamic time warping analysis,” *Remote Sens. Environ.*, vol. 204, pp. 509–523, 2018.

[13] V. Maus, G. Camara, R. Cartaxo, A. Sanchez, F. M. Ramos, and G. de Queiroz, “A time-weighted dynamic time warping method for land-use and land-cover mapping,” *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 9, no. 8, pp. 3729–3739, Aug. 2016.

[14] Q. Peng et al., “A twenty-year dataset of high-resolution maize distribution in China,” *Sci. Data*, vol. 10, 2023, Art. no. 658.

[15] S. Lv, X. Xia, and Y. Pan, “Optimization of characteristic phenological periods for winter wheat extraction using remote sensing in plateau valley agricultural areas in Hualong, China,” *Remote Sens.*, vol. 15, 2023, Art. no. 28.

[16] J. Xu, H. Zhao, P. Yin, D. Jia, and G. Li, “Remote sensing classification method of vegetation dynamics based on time series landsat image: A case of opencast mining area in China,” *Eurasip J. Image Video Process.*, 2018, Art. no. 113.

[17] F. Petitjean, G. Forestier, G. Webb, A. Nicholson, Y. Chen, and E. Keogh, “Dynamic time warping averaging of time series allows faster and more accurate classification,” in *Proc. 14th IEEE Int. Conf. Data Mining*, 2014, pp. 470–479.

[18] F. Zhuang et al., “A comprehensive survey on transfer learning,” *Proc. IEEE*, vol. 109, no. 1, pp. 43–76, Jan. 2021.

[19] E. Fekri, H. Latifi, M. Amani, and A. Zobeidinezhad, “A training sample migration method for wetland mapping and monitoring using Sentinel data in Google Earth engine,” *Remote Sens.*, vol. 13, 2021, Art. no. 4169.

[20] R. Kate, “Using dynamic time warping distances as features for improved time series classification,” *Data Mining Knowl. Discov.*, vol. 30, pp. 283–312, 2016.

[21] Q. Zhang et al., “A method for measuring similarity of time series based on series decomposition and dynamic time warping,” *Appl. Intell.*, vol. 53, pp. 6448–6463, 2023.

[22] R. Shen, J. Dong, W. Yuan, W. Han, T. Ye, and W. Zhao, “A 30 m resolution distribution map of maize for China based on landsat and Sentinel images,” *J. Remote Sens.*, vol. 2022, 2022, Art. no. 9846712.

[23] M. Belgiu, W. Bijker, O. Csillik, and A. Stein, “Phenology-based sample generation for supervised crop type classification,” *Int. J. Appl. Earth Observ. Geoinf.*, vol. 95, 2021, Art. no. 102264.

[24] S. A. Zaheer, Y. Ryu, J. Lee, Z. Zhong, and K. Lee, "In-season wall-to-wall crop-type mapping using ensemble of image segmentation models," *IEEE Trans. Geosci.*, vol. 61, 2023, Art. no. 4411311.
[25] Z. Cai et al., "Improving agricultural field parcel delineation with a dual branch spatiotemporal fusion network by integrating multimodal satellite data," *ISPRS J. Photogramm. Remote Sens.*, vol. 205, pp. 34–49, 2023.
[26] L. Yang et al., "Inter-continental transfer of pre-trained deep learning rice mapping model and its generalization ability," *Remote Sens.*, vol. 15, 2023, Art. no. 2443.
[27] P. Wei, D. Chai, T. Lin, C. Tang, M. Du, and J. Huang, "Large-scale rice mapping under different years based on time-series Sentinel-1 images using deep semantic segmentation model," *ISPRS J. Photogramm. Remote Sens.*, vol. 174, pp. 198–214, 2021.
[28] Y. Wang, L. Feng, Z. Zhang, and F. Tian, "An unsupervised domain adaptation deep learning method for spatial and temporal transferable crop type mapping using Sentinel-2 imagery," *ISPRS J. Photogramm. Remote Sens.*, vol. 199, pp. 102–117, 2023.
[29] O. Csillik, M. Belgiu, G. Asner, and M. Kelly, "Object-based time-constrained dynamic time warping classification of crops using Sentinel-2," *Remote Sens.*, vol. 11, 2019, Art. no. 1257.
[30] F. Lupo, M. Linderman, V. Vanacker, E. Bartholome, and E. Lambin, "Categorization of land-cover change processes based on phenological indicators extracted from time series of vegetation index data," *Int. J. Remote Sens.*, vol. 28, pp. 2469–2483, 2007.
[31] M. Kawarkhe and V. Musande, "Performance analysis of possibilistic fuzzy clustering and support vector machine in cotton crop classification," in *Proc. 3rd Int. Conf. Adv. Comput., Commun. Inf.*, 2014, pp. 961–967.
[32] Y. Zhang, J. Liu, L. Wan, and S. Qi, "Land cover/use classification based on feature selection," *J. Coastal Res.*, vol. 73, pp. 380–385, 2015.
[33] V. Musande, A. Kumar, and K. Kale, "Cotton crop discrimination using fuzzy classification approach," *J. Indian Soc. Remote Sens.*, vol. 40, pp. 589–597, 2012.
[34] Z. Mao, L. Deng, F. Duan, X. Li, and D. Qiao, "Angle effects of vegetation indices and the influence on prediction of spad values in soybean and maize," *Int. J. Appl. Earth Observ. Geoinf.*, vol. 93, 2020, Art. no. 102198.
[35] K. Sivabalan and E. Ramaraj, "Shortwave infrared-based phenology index method for satellite image land cover classification," in *Proc. 8th Int. Conf. Soft Comput. Problem Solving*, 2018, pp. 877–889.
[36] W. Fang, H. Zhu, S. Li, H. Ding, and R. Bi, "Rapid identification of main vegetation types in the Lingkong mountain nature reserve based on multi-temporal modified vegetation indices," *Sensors*, vol. 23, 2023, Art. no. 659.
[37] T. Weitkamp and K. Poolad, "Evaluating the effect of training data size and composition on the accuracy of smallholder irrigated agriculture mapping in Mozambique using remote sensing and machine learning algorithms," *Remote Sens.*, vol. 15, 2023, Art. no. 3017.
[38] N. You, J. Dong, J. Li, J. Huang, and Z. Jin, "Rapid early-season maize mapping without crop labels," *Remote Sens. Environ.*, vol. 290, 2023, Art. no. 113496.
[39] C. Zhang et al., "Rapid in-season mapping of corn and soybeans using machine-learned trusted pixels from cropland data layer," *Int. J. Appl. Earth Observ. Geoinf.*, vol. 102, 2021, Art. no. 102374.
[40] S. Morell-Monzo, M. Sebastia-Frasquet, J. Estornell, and E. Molto, "Detecting abandoned citrus crops using Sentinel-2 time series. A case study in the comunitat valenciana," *ISPRS J. Photogramm. Remote Sens.*, vol. 201, pp. 54–66, 2023.
[41] L. Meng et al., "Large-scale and high-resolution paddy rice intensity mapping using downscaling and phenology-based algorithms on Google Earth engine," *Int. J. Appl. Earth Observ. Geoinf.*, vol. 128, 2024, Art. no. 103725.
[42] T. Schroeder, M. Schaale, J. Lovell, and D. Blondeau-Patissier, "An ensemble neural network atmospheric correction for Sentinel-3 OLCI over coastal waters providing inherent model uncertainty estimation and sensor noise propagation," *Remote Sens. Environ.*, vol. 270, 2022, Art. no. 112848.
[43] J. Ojo, "Geo-spatial distribution of cloud cover and influence of cloud induced attenuation and noise temperature on satellite signal propagation over Nigeria," *Adv. Space Res.*, vol. 59, pp. 2611–2622, 2017.
[44] V. Konduri, J. Kumar, W. Hargrove, F. Hoffman, and A. R. Ganguly, "Mapping crops within the growing season across the United States," *Remote Sens. Environ.*, vol. 251, 2020, Art. no. 112048.
[45] L. Yin, N. You, G. Zhang, J. Huang, and J. Dong, "Optimizing feature selection of individual crop types for improved crop mapping," *Remote Sens.*, vol. 12, 2020, Art. no. 162.
[46] H. Wei et al., "Spatiotemporal expansion and methane emissions of rice-crayfish farming systems in Jianghan plain, China," *Agricultural Forest Meteorol.*, vol. 347, 2024, Art. no. 109908.
[47] X. Liu et al., "Spatial-temporal patterns of features selected using random forests: A case study of corn and soybeans mapping in the US," *Int. J. Remote Sens.*, vol. 40, pp. 269–283, 2019.
[48] D. Jiang, S. Chen, J. Useya, L. Cao, and T. Lu, "Crop mapping using the historical crop data layer and deep neural networks: A case study in Jilin Province China," *Sensors*, vol. 22, 2022, Art. no. 5853.
[49] X. Zhao, T. Wu, S. Wang, K. Liu, and J. Yang, "Cropland abandonment mapping at sub-pixel scales using crop phenological information and MODIS time-series images," *Comput. Electron. Agriculture*, vol. 208, 2023, Art. no. 107763.

Hao-Nan Zuo received the B.Mgt. degree in information management and information systems from the School of Economics and Management, China University of Petroleum, Beijing, China, in 2023.

He is currently working with State Key Laboratory of Efficient Utilization of Arid and Semi-arid Arable Land in Northern China/Institute of Agricultural Resources and Regional Planning, Chinese Academy of Agricultural Sciences, Beijing. His research interest focuses on agricultural remote sensing.

Pei Leng (Senior Member, IEEE) received the Ph.D. degree in cartography and geographical information system from the University of Chinese Academy of Sciences, Beijing, China, in 2015.

He is currently an Associate Professor with the State Key Laboratory of Efficient Utilization of Arid and Semi-Arid Arable Land in Northern China, Institute of Agricultural Resources and Regional Planning, Chinese Academy of Agricultural Sciences, Beijing. His research interests include methodological developments of satellite soil moisture and energy fluxes and agricultural remote sensing.

Yu-Xuan Li received the Bachelor of Science degree in physics from the Department of Physics, Taiyuan Normal University, Jinzhong, China, in 2019.

She is currently working with the State Key Laboratory of Efficient Utilization of Arid and Semi-arid Arable Land in Northern China/Institute of Agricultural Resources and Regional Planning, Chinese Academy of Agricultural Sciences, Beijing, China. Her research interest focuses on agricultural remote sensing.

Qian Song received the Ph.D. degrees in agricultural resources and environment from the Chinese Academy of Agricultural Sciences, Beijing, China, in 2016.

She is currently an Associate Professor with the State Key Laboratory of Efficient Utilization of Arid and Semi-arid Arable Land in Northern China/Institute of Agricultural Resources and Regional Planning, Chinese Academy of Agricultural Sciences. Her research interests include crop mapping and spatio-temporal dynamics of crop planting patterns based on remote sensing images.

Zhao-Liang Li (Fellow, IEEE) received the Ph.D. degree in geophysics and remote sensing from the University of Strasbourg, Strasbourg, France, in 1990.

Since 1992, he has been a Research Scientist with CNRS, Illkirch, France. He joined the Institute of Agricultural Resources and Regional Planning, Beijing, China, in 2013. He has participated in many national and international projects, such as NASA-funded MODIS, EC-funded program EAGLE, ESA-funded program SPECTRA, etc. His research interests include thermal infrared radiometry, parameterization of land surface processes at large scale, and the assimilation of satellite data to land surface models.