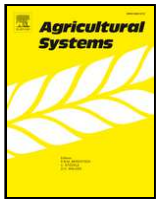
Agricultural Systems xxx (xxxx) 103462

Contents lists available at ScienceDirect

Agricultural Systems

journal homepage: www.elsevier.com/locate/agsy



Towards automation of in-season crop type mapping using spatiotemporal crop information and remote sensing data

Chen Zhang [a], Liping Di [a,*], Li Lin [a], Hui Li [a], Liying Guo [a], Zhengwei Yang [b], Eugene G. Yu [a], Yahui Di [a], Anna Yang [a]

[a] Center for Spatial Information Science and Systems, George Mason University, Fairfax, VA 22030, USA
[b] U.S. Department of Agriculture National Agricultural Statistics Service, Washington, DC 20250, USA

ARTICLE INFO

Editor: Kairsty Topp

Keywords:
Crop mapping
Agriculture 4.0
Remote sensing
Cropland Data Layer
Sentinel-2
Google Earth Engine

ABSTRACT

CONTEXT: Mapping uncertain crop types from satellite images is a promising application in agricultural systems. However, it is still a challenge to automate in-season crop type mapping over a large area because of the insufficiency of ground truth and issues of scalability, reusability, and accessibility of the classification model. This study introduces a framework for automatic crop type mapping using spatiotemporal crop information and Sentinel-2 data based on Google Earth Engine (GEE). The main advantage of the framework is using the trusted pixels extracted from the historical Cropland Data Layer (CDL) to replace ground truth and label training samples in satellite images.

OBJECTIVE: This paper will carry out the following three objectives: (1) assessing spatiotemporal crop information derived from the historical crop cover maps; (2) mapping uncertain crop cover, mainly crop fields without regular historical crop rotation patterns, from remote sensing data using supervised learning classification and validating mapping results; and (3) automating in-season crop mapping and exploring the scalability of the framework.

METHODS: The proposed crop mapping workflow consists of four stages. The data preparation stage preprocesses CDL and Sentinel-2 data into the required structure. The spatiotemporal crop information sampling stage extracts trusted pixels from the historical CDL time series and labels Sentinel-2 data. Then a crop type classification model can be trained using the supervised learning classifier in the model training stage. In the mapping/validation stage, an in-season crop cover map over the full Sentinel-2 tile will be produced using the trained model and the classification performance will be validated using CDL or other ground truth data.

RESULTS AND CONCLUSIONS: We systematically perform a group of experiments for in-season mapping of five major crop types (corn, cotton, rice, soybeans, and soybeans-wheat double cropping) over the Mississippi Delta region. The result indicates that the crop cover map of the study area is expected to reach 80%–90% agreement with CDL within the growing season. To further facilitate the use of the framework, we also develop a GEE-enabled online prototype, In-season Crop Mapping Kit, and explore its scalability over agricultural fields in various ecoregions including California, Idaho, Kansas, and Illinois.

SIGNIFICANCE: The mapping-without-ground-truth approach described in this paper can significantly simplify the sampling process and save substantial human intervention and financial resources. The findings and outputs will benefit the agriculture community and other agricultural sectors ranging from government, academia, and companies.

1. Introduction

In the 2030 Agenda for Sustainable Development (UN, 2015), the United Nations (UN) lists zero hunger as one of 17 Sustainable Development Goals (SDGs). According to the Global Report on Food Crises 2019, >113 million people across 53 countries experienced acute hunger in 2018 (FSIN, 2019). Meanwhile, the World Population Prospects 2019 shows the average estimate of the world population expected to reach 8.54 billion by 2030, 9.73 billion by 2050, and 10.87 billion by 2100 (UN, 2019). The analysis of the Food and Agriculture

* Corresponding author.
E-mail address: ldi@gmu.edu (L. Di).

https://doi.org/10.1016/j.agsy.2022.103462
Received 26 April 2022; Received in revised form 27 June 2022; Accepted 12 July 2022
0308-521/© 20XX

Organization (FAO) of UN indicates such a population growth rate would lead to a substantial rise in global food demand and potentially jeopardize the sustainability of agriculture systems in the 21st century (FAO, 2019). Therefore, it is a great challenge to make the agricultural system efficient and sustainable to balance the increasing demand for foods with natural resources and climate change.

Agriculture has been the most fundamental activity to produce food as well as other living essentials such as fiber, fuel, timber, and medicine from thousands of years ago in human society. With the rapid development of information science and technology in the past few decades, a new paradigm, Agriculture 4.0, has been raised to achieve greater agricultural production efficiency (Rose and Chilvers, 2018). This new paradigm aims to integrate agriculture with state-of-the-art digital technologies, such as smart farming, big data, digital twins, Artificial Intelligence (AI), Internet of Things (IoT), and Information and Communications Technology (ICT) (Santos Valle and Kienzle, 2020; Zhai et al., 2020; Latino et al., 2021; Verdouw et al., 2021; Osinga et al., 2022). With the rapid volume growth of Earth Observation (EO) data during the past few decades, remote sensing has become a key technology in Agriculture 4.0 to collecting, managing, and understanding agro-geoinformation data (Di and Yang, 2014). Remote sensing data acquired by Earth resources satellite sensors, such as Terra/Aqua Moderate Resolution Imaging Spectroradiometer (MODIS), Landsat Operational Land Imager (OLI)/Thematic Mapper (TM), and Sentinel-2 Multispectral Instrument (MSI), are widely applied in agriculture-related studies, including agricultural sustainability, food security, environmental health, bioenergy, natural resource conservation, land use management, carbon accounting, and agricultural commodity trading (Brown and Pervez, 2014; Khanal et al., 2017; Piedelobo et al., 2019; Gao and Zhang, 2021). Furthermore, the emergence of high-performance cloud computing platforms and infrastructures for EO data, such as NASA's Earth Observing System Data and Information System (EOSDIS), Google Earth Engine (GEE), Sentinel Hub, Open Data Cube, and Euro Data Cube, greatly simplified the process for integrating a plethora of remote-sensing-based data products (e.g., land use land cover, soil moisture, vegetation index, evapotranspiration, and land surface temperature) into agricultural systems and applications (Li, 2020; Charvat et al., 2021; Killough et al., 2021; Wagemann et al., 2021; Zhang et al., 2021b).

As one of the critical applications in remote sensing, crop mapping aims to identify the specific crop types and delineate their extent within the field from satellite images (NRCan, 2014). In particular, the remote-sensing-based crop mapping data will provide essential information to support crop yield prediction, crop condition monitoring, natural hazard assessment, and many other socio-economic activities. For example, the Cropland Data Layer (CDL) product of the U.S. Department of Agriculture (USDA) National Agricultural Statistics Service (NASS) is produced using remote sensing data from multiple satellite sensors including Landsat and Sentinel-2 (USDA NASS, 2022). The Annual Crop Inventory (ACI) of Agriculture and Agri-Food Canada (AAFC) is produced using satellite images from Landsat-8, Sentinel-2, and RADARSAT-2 sensors (Fisette et al., 2013). Although with accurate and detailed crop cover information, CDL and ACI are the post-season crop mapping data products usually released to the public in the early next year. Thus, the remote-sensing-based in-season and early-season crop type mapping has become a key research topic in the land use and land cover (LULC) community. Among various agricultural regions, the Conterminous United States (CONUS), the largest producer of corn and soybeans in the world, is one of the most ideal study areas to investigate and test new mapping methods in recent years. For example, Hao et al. (2015) explored the early-season crop type classification using MODIS data in the state of Kansas. Varmaghani and Eichinger (2016) investigated the early-season classification of corn and soybean using Landsat data in Iowa. Cai et al. (2018) developed a machine-learning-based approach for in-season corn-soybean classification using Landsat data within the state of Illinois. Wang et al. (2019) introduced an unsupervised clustering method for crop mapping over Corn Belt. Zhong et al. (2019) developed a deep learning framework to classify summer crops in California state. Zhang et al. (2019) predicted the spatial distribution of annual crop cover from the historical crop planting information in the U.S. Corn Belt. Konduri et al. (2020) used MODIS-derived vegetation index time series to map major crop types across CONUS before harvest. Johnson and Mueller (2021) used archival land cover information for mapping corn, soybeans, and winter wheat over Corn Belt and Great Plains region. Xu et al. (2021) compared several deep learning models in crop mapping using multi-temporal Landsat data at study areas in Iowa and Indiana.

However, producing the field-level in-season crop cover maps for the CONUS is still challenging. There are three main problems. First, producing a high-quality in-season crop cover map requires a large number of ground truths. However, surveying ground truth over a large geographic area is labor-intensive and time-critical. It requires a substantial investment of human and financial resources unaffordable by individual researchers, small research groups, or even large organizations. Second, although many crop mapping methods can reach excellent results for a specific area at the early growing season, they are difficult to be scaled up to a large geographic area because of the regional difference in geography, crop types, growing season, and temporal collection of remote sensing images. Third, it is difficult for many users to reuse the classification models, especially the complex deep learning models trained with large-scale computing resources. Therefore, a reliable, scalable, and reusable framework for automatic in-season crop type mapping is highly demanded for agricultural and environmental modeling.

In response to the above challenge, we conceive, design, and implement a framework for automation of in-season crop mapping using spatiotemporal crop information and Sentinel-2 data. The research involves the state-of-the-art methods, technologies, and tools in Agriculture 4.0, ranging from remote sensing, geographic information science and systems, to machine learning and geospatial cyberinfrastructure. This paper will carry out the following three objectives: (1) assessing spatiotemporal crop information derived from the historical crop cover maps; (2) mapping uncertain crop cover, mainly crop fields without regular historical crop rotation patterns, from remote sensing data using supervised learning classification and validating mapping results; and (3) automating in-season crop mapping and exploring the scalability of the framework.

2. Materials and methods

2.1. Data

Table 1 summarizes the main datasets used in this study. The proposed framework adopts the Sentinel-2 Level-2A product as the remote sensing dataset for the crop type mapping. The Copernicus Sentinel-2 mission is operated by the European Space Agency (ESA), which consists of two twin polar-orbiting satellites (Sentinel-2A and Sentinel-2B) with global coverage. The Sentinel-2A and Sentinel-2B satellite were

Table 1
Summary of datasets used in the proposed crop mapping framework.

	Sentinel-2 MSI Level-2A	Cropland Data Layer
Spatial Resolution	10 m–60 m	30 m
Temporal Resolution	5-day	Annual (normally released in the early of the following year)
Coverage	March 2018 - present (Europe) December 2018 - present (Global)	1997–2007 (Some states) 2008 - present (CONUS)

launched in June 2015 and March 2017, respectively. The Sentinel-2 satellites revisit every five days under the same viewing angles with a spatial resolution of 10–60 m. The Sentinel-2 satellites equip with the MSI sensors, which covers 13 spectral bands ranging from visible and near-infrared to shortwave infrared wavelengths. Compared with the Landsat data, Sentinel-2 data have a higher spatial and temporal resolution, which is more suitable for the in-season mapping and near-real-time agricultural monitoring. We also used CDL as the reference dataset to extract spatiotemporal crop information and validate the mapping results. It is currently one of the most widely-used national-scale field-level crop mapping data annually produced by USDA NASS. It covers the entire CONUS at 30 m spatial resolution from 2008 to the present and some states from 1997 to 2007. Over 200 land use classes are provided in the CDL, and the accuracy for major crop types is close to 95% (Boryan et al., 2011).

2.2. Study area

This study demonstrates the capability of the proposed crop mapping workflow over the Mississippi Delta region. This study area is located within the Mississippi Alluvial Plain Ecoregion and Mississippi Valley Loess Plains Ecoregion, which contains most parts of eastern Arkansas, western Mississippi, western Tennessee, and a small part of southeastern Missouri and southwestern Kentucky. Compared with the U.S. Corn Belt which contains only two dominant crop types (i.e., corn and soybeans), the Mississippi Delta region has a more diverse cropping environment, where the major crop types include corn, cotton, rice, soybeans, and soybeans-wheat double cropping (plant/harvest soybeans then plant/harvest winter wheat in the same field throughout one year). Fig. 1 shows the crop cover map of the study area and highlights the covered tiles of Sentinel-2 data.

According to the Usual Planting and Harvesting Dates for U.S. Field Crops (USDA NASS, 2010), the planting dates of major crops within the Mississippi Delta region are usually active from April to May in the study area. The active harvesting dates of these crops are from September to October. Winter wheat is typically planted from October to November and harvested in June of the following year. Fig. 2 shows the timeline of usual planting and harvesting dates for major crop types in the four states within the study area.

2.3. Design of automatic crop mapping framework

Fig. 3 illustrates the proposed automatic in-season crop mapping framework. The framework is implemented on the GEE platform (Gorelick et al., 2017), which has been widely used in agricultural systems and applications with EO data (Rembold et al., 2019; Filippi et al., 2020; Luo et al., 2021; Suchi et al., 2021; Yang et al., 2021; Zurqani et al., 2021). Users can access the framework described in this study via the application programming interface (API) and graphical user interface (GUI) with the GEE Code Editor. For each classification task, users need to assign the Sentinel-2 tile as the area of interest (AOI), the start/end date of Sentinel-2 data, and several optional inputs. Then an in-

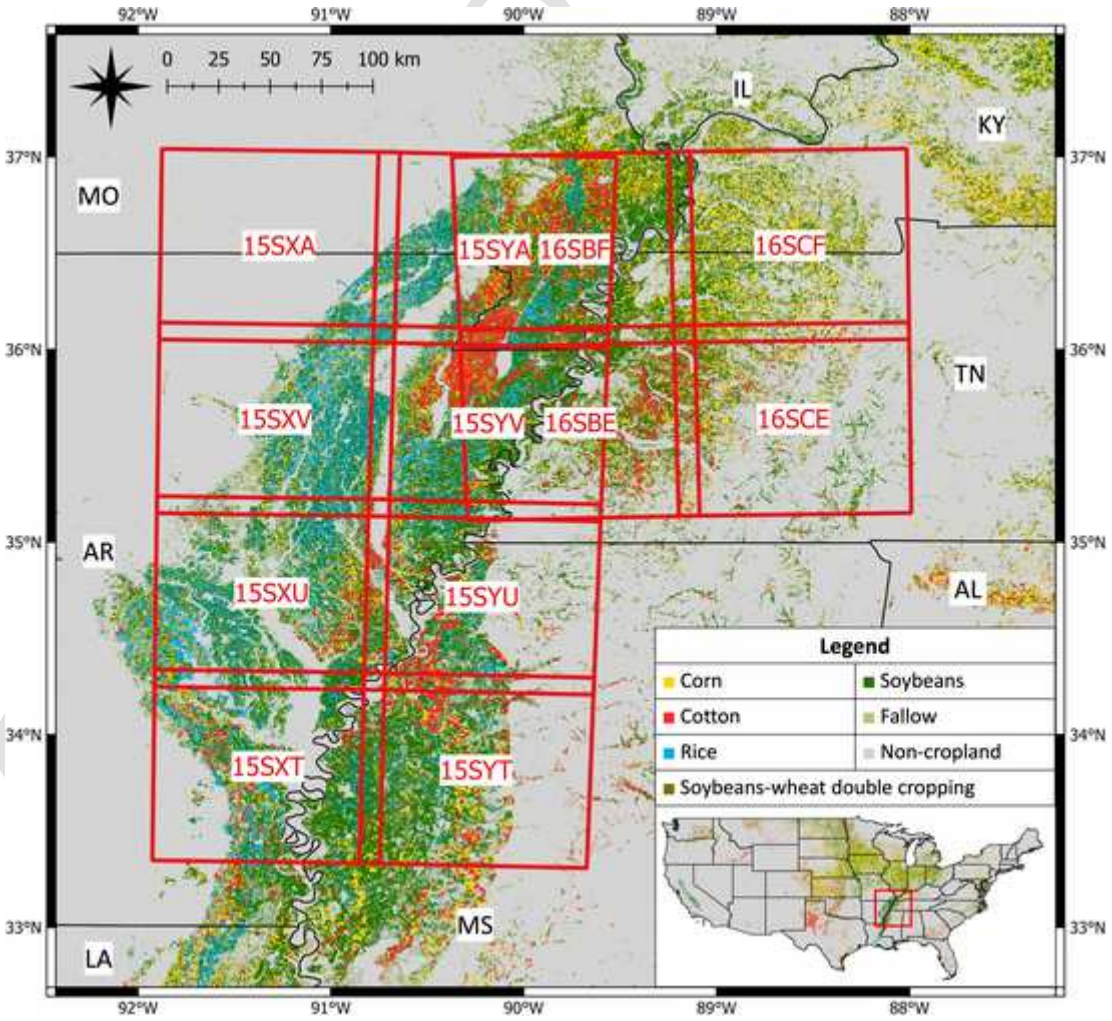


Fig. 1. Map of the Mississippi Delta region (data from 2020 Cropland Data Layer). The red boundaries and annotations refer to the Sentinel-2 image tiles investigated in this study. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

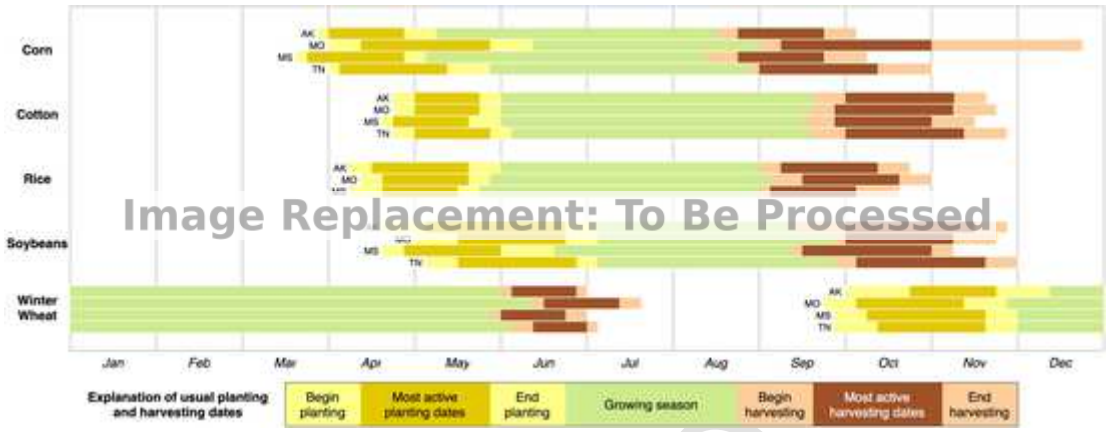


Fig. 2. Usual planting and harvesting dates for major crop types in the different states in Mississippi Delta region (data from USDA NASS (2010)).

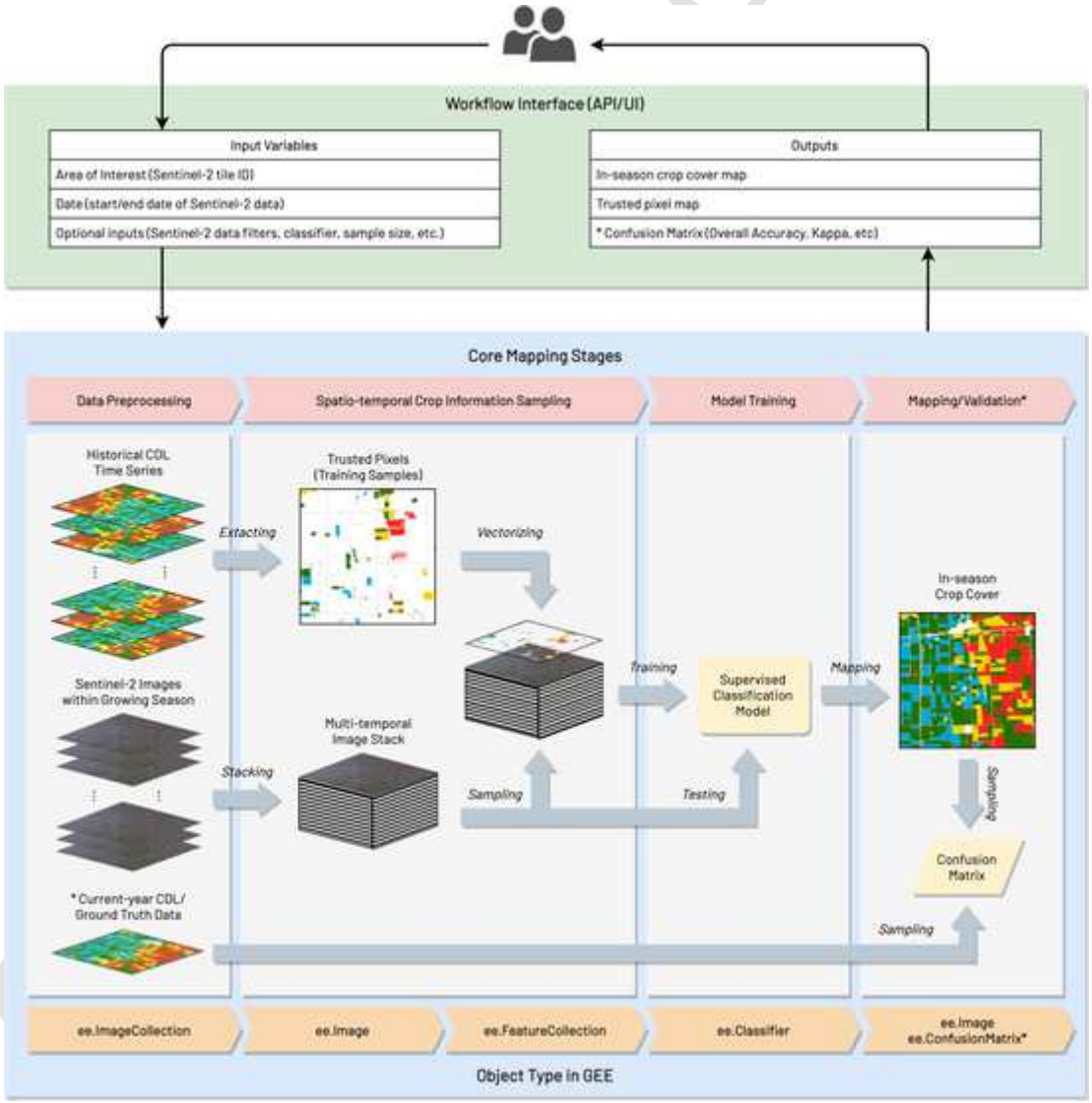


Fig. 3. The framework of automatic crop type mapping based on GEE. (* The confusion matrix will be calculated for validation only if the current-year CDL or ground truth data is applicable.)

season crop cover map along with the trusted pixel map and confusion matrix (if reference data applicable) will be automatically generated. The core mapping workflow consists of the following stages: data preprocessing, spatiotemporal crop information sampling, crop type classification, and mapping/validation. The data preparation stage preprocesses CDL and Sentinel-2 data into the required structure. The spatiotemporal crop information sampling stage extracts trusted pixels from the historical CDL time series and labels Sentinel-2 data. Then a

crop type classification model can be trained using the supervised learning classifier in the model training stage. In the mapping/validation stage, an in-season crop cover map over the full Sentinel-2 tile will be produced using the trained model. The classification performance will be validated using CDL or other ground truth data. This section will present the details of each mapping stage in the proposed framework.

2.3.1. Data preparation

The data preparation is the first stage of the framework, which aims to filter and preprocess the geospatial data used for crop type mapping. Since the complete volume of Sentinel-2 data and CDL data are already available in GEE, these data can be imported from GEE's public data catalog without retrieving from the other providers. Both CDL and Sentinel-2 datasets are the "ee.ImageCollection" object type in GEE, which contains a collection of images with the same metadata scheme. The CDL image collection is straightforward, which consists of a list of annual images (one image per year). We can directly import the CDL image collection without further processing. It should be noted that we only use CDL data after 2008 because CDL does not cover the entire CONUS before 2008. On the other hand, the data quality of the early-year CDL in many regions is unsatisfactory.

The image collection of Sentinel-2 L2A data includes a large volume of images with abundant image properties. When preparing the Sentinel-2 data for the crop type mapping, we filter the qualified image with several conditions. First, since the tiling grid splits the Sentinel-2 data into 100 × 100 km squares, some images contain a part of "no data pixels" that are overlaid with the diagonal stripes from the orbital track of the satellite. To make sure the spectral and temporal information are valid among all pixels in the target tile, we only use the images without "no data pixels" by limiting the "NODATA_PIXEL_PERCENTAGE" property of the Sentinel-2 L2A image. To further improve the quality of the image collection, we also restrict the "CLOUDY_PIXEL_PERCENTAGE" property to filter out images with cloud coverage exceeding 30%. By applying the above filters, a collection of high-quality Sentinel-2 images for crop type mapping will be prepared. According to our observation and experiment, 30% is a fair condition to balance the image quantity and quality. Under such conditions, there would be > 20 Sentinel-2 images of the growing season available for most tiles. Since the model learns the spectral profile from a large volume of trusted pixels, the impact of cloudy pixels in a single image would be limited. Fig. 4 lists the qualified Sentinel-2 images acquired between the usual planting and harvesting date over the study area.

2.3.2. Spatiotemporal crop information sampling

The key of the proposed in-season crop mapping method is using trusted pixels as training samples to derive uncertain crop cover from remote sensing images. Many studies have shown crop rotation is a common and important practice in agricultural systems around the world (Peltonen-Sainio and Jauhiainen, 2019; Wieme et al., 2020; Sietz et al., 2021; Nowak et al., 2022). In our previous study, we have used machine learning to recognize corn-soybean rotation patterns from the historical CDL time series (Zhang et al., 2021a, 2021c). It was found that alternate corn-soybean rotation patterns and monocropping patterns are widely distributed across the U.S. agricultural regions. These two crop rotation patterns are applied for spatiotemporal crop information sampling in the implementation of this framework.

Specifically, we use the alternate cropping patterns to pick the pixels with the same crop type category every other year in the CDL time series. For example, if a pixel is categorized as "corn" in the odd-numbered year while it is tagged as a "non-corn" type in the even-numbered year, this pixel will be counted as a trusted "corn" pixel in the odd year. Alternatively, if a pixel is constantly categorized as "corn" in the CDL time series, it is also counted as a trusted "corn" pixel. Fig. 5 shows an example of extracting spatiotemporal trusted pixels of major crop types in 2020 from the historical CDL time series (2008–2019) based on alternating cropping and monocropping patterns.

The trusted pixels will be used to label multi-temporal Sentinel-2 data prepared in the previous stage. Before labeling training samples, we need to stack all bands from each qualified Sentinel-2 image over the target tile into one multi-temporal image stack. For each qualified Sentinel-2 image, we stack three visible spectrum bands (blue, green, red), one near-infrared (NIR) band, and two shortwave infrared (SWIR) bands. These spectral bands are found effective for classifying different crop types in remote sensing data (Cai et al., 2018; Song et al., 2021; Wang et al., 2021). On the other hand, the Normalized Difference Vegetation Index (NDVI) and Normalized Difference Water Index (NDWI) have been proven to be sensitive features for distinguishing crop types in the early growing season (Hao et al., 2015). NDVI measures the vegetation index by normalizing the difference between NIR reflectance and red light reflectance (Tucker, 1979). Similarly, NDWI enhances the presence of water features by normalizing the difference between green light reflectance and NIR reflectance (McFeeters, 1996). Therefore, we also derived NDVI and NDWI from the spectral bands of each Sentinel-2 image and add these two index features to the multi-temporal image stack. Table 2 lists the selected spectral features for model training. In this stage, the object type of the multi-temporal Sentinel-2 image stack will be converted to "ee.Image", which is a single multi-band image but

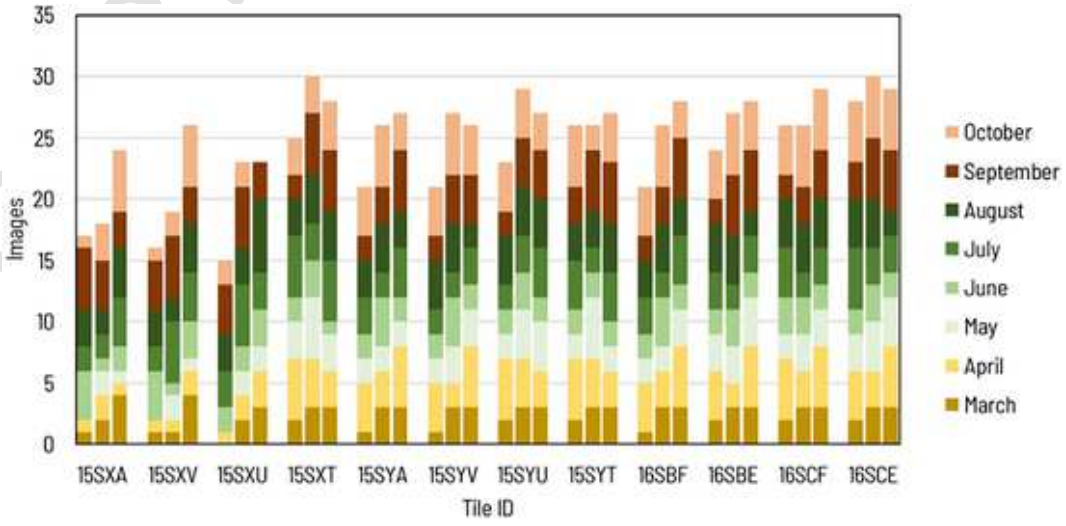


Fig. 4. Summary of qualified Sentinel-2 images in the study area from 2019 to 2021 (left to right in each column).

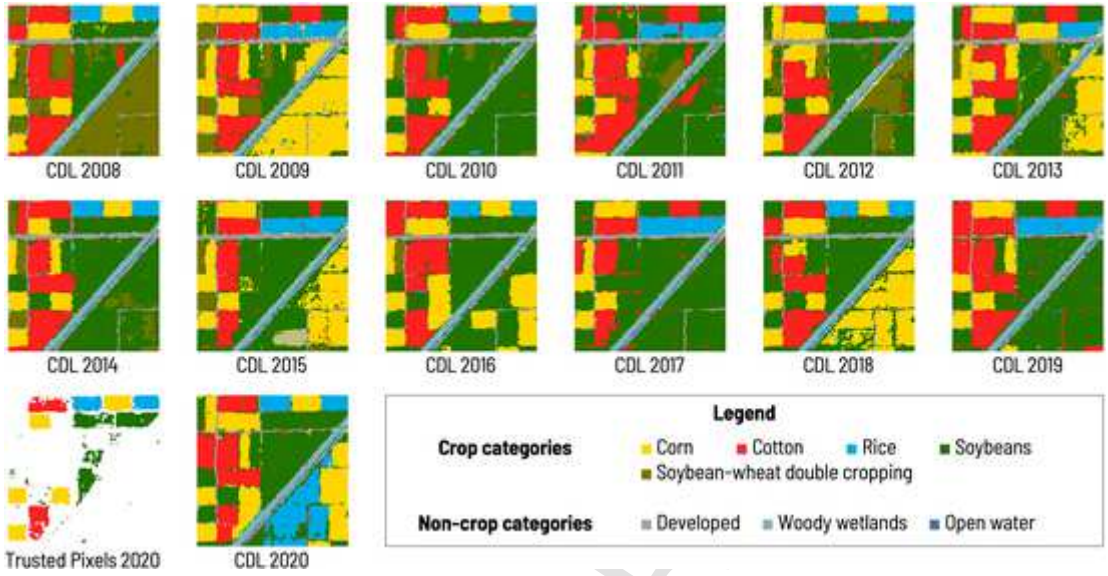


Fig. 5. Example of extracting spatiotemporal trusted pixels of 2020 from historical CDL time series (2008–2019) based on monocropping patterns and alternate cropping patterns (area selected from the northern Mississippi Delta region).

Table 2
Selected spectral features of Sentinel-2 image for model training.

Name	Band	Central wavelength	Resolution
Blue	Band 2	496.6 nm (S2A) / 492.1 nm (S2B)	10 m
Green	Band 3	560 nm (S2A) / 559 nm (S2B)	10 m
Red	Band 4	664.5 nm (S2A) / 665 nm (S2B)	10 m
NIR	Band 8	835.1 nm (S2A) / 833 nm (S2B)	10 m
SWIR-1	Band 11	1613.7 nm (S2A) / 1610.4 nm (S2B)	20 m
SWIR-2	Band 12	2202.4 nm (S2A) / 2185.7 nm (S2B)	20 m
NDVI	$\frac{Band\ 8-Band\ 4}{Band\ 8+Band\ 4}$	N/A	10 m
NDWI	$\frac{Band\ 3-Band\ 8}{Band\ 3+Band\ 8}$	N/A	10 m

with abundant spectral and temporal features within the growing season.

2.3.3. Model training

The training of crop type classification model is a typical supervised classification process. We use trusted pixels to label spectral features in multi-temporal Sentinel-2 image stack, then feed these training samples to the supervised-learning classifier for model training. GEE offers a suite of ready-to-use classifier APIs for supervised classification model training. Our framework chooses Random Forest (Breiman, 2001) as the default classifier, which has been widely applied in remote-sensing-based LULC classification (Tatsumi et al., 2015; Pelletier et al., 2016; Steinhausen et al., 2018; Hao et al., 2020). The classifier uses 128 trees by default, which is an optimal tree number in the forest to obtain a good balance between performance and processing time (Oshiro et al., 2012). Besides, we provide an option for users to choose other classifiers, such as Classification and Regression Trees (CART) (Breiman, 1984) and Support Vector Machine (SVM) (Burges, 1998).

2.3.4. Mapping and validation

After the classification model training is completed, the in-season crop cover map will be produced by applying the well-trained classification model to the full Sentinel-2 image stack. To be consistent with CDL, the spatial resolution of in-season maps is resampled from 10 m to 30 m before validation. For each tile of mapping result, we measure the classification accuracy using CDL as reference data and calculate the Overall Accuracy (OA) and Kappa coefficient based on the confusion matrix (Stehman, 1997). The OA reflects the ratio of correctly classified test samples in the total number of samples. The Kappa coefficient (Cohen, 1960) measures the interrater reliability of the classification result. We also use precision, recall, and F1-score to measure the classification performance of each crop type. As the growing season progresses, the monthly in-season crop cover maps are produced and validated by the end of April, May, June, July, and August.

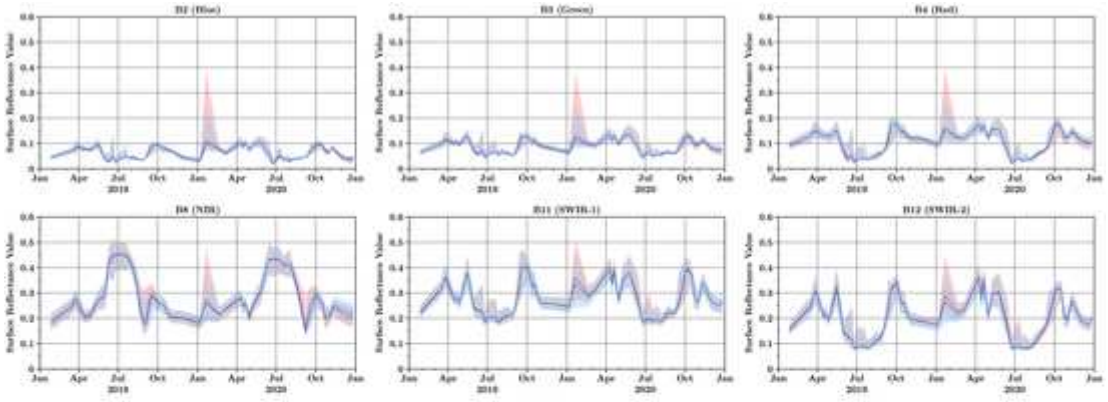
3. Experiments and results

This section presents four groups of experiments to test the feasibility of in-season crop type mapping with the proposed framework. Section 3.1 analyzes the spectral and temporal features of trusted pixels for five major crop types. Section 3.2 illustrates the spectral and index profiles as well as the performance of the classification model trained with the spectral and temporal features. Additionally, an online prototype for the automatic in-season crop mapping based on the proposed framework is demonstrated in Section 3.4.

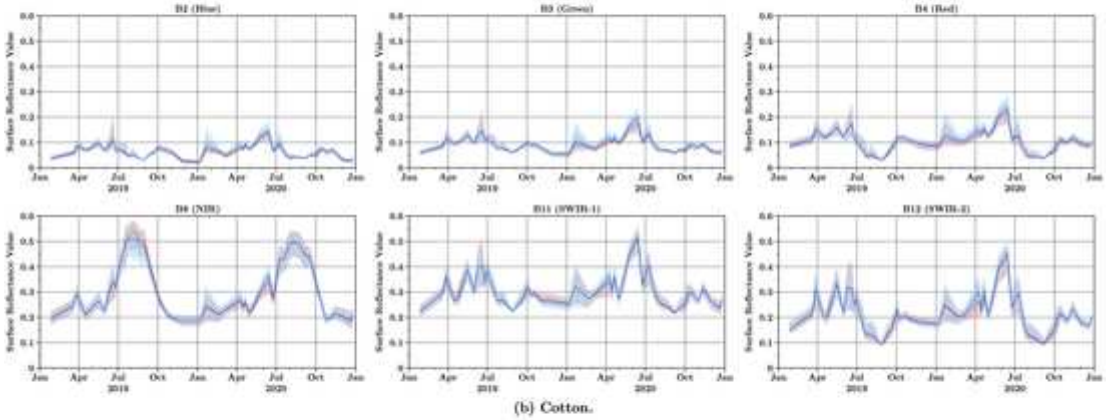
3.1. Spectral and temporal analysis of trusted crop information

The first experiment aims to test whether the spatiotemporal crop information extracted from historical CDL is trusted. In the remote-sensing-based supervised classification, the classifier automatically learns the spectral and temporal features in the training data. If a large number of training samples are mislabeled, the classification performance will be significantly affected. In this case, the classification model is trained using samples in the multi-temporal Sentinel-2 image stack. Before training the classification model, we need to validate the spectral and temporal profiles for trusted pixels.

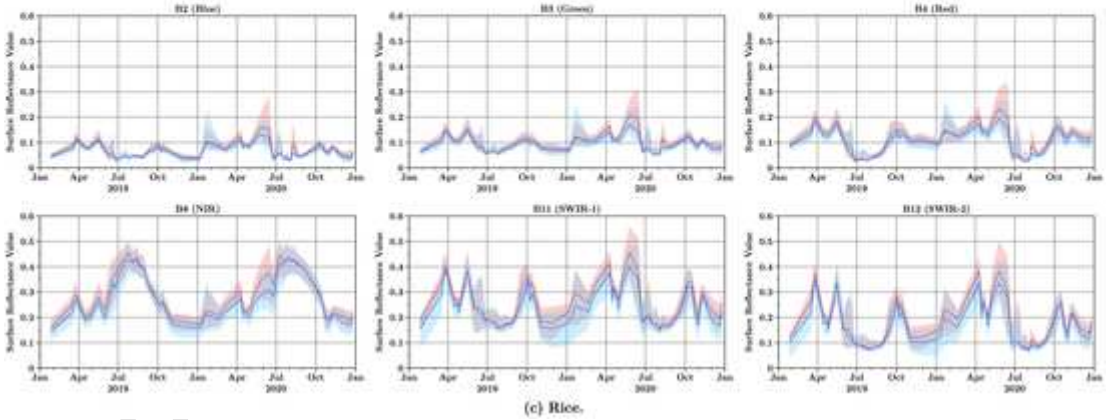
Fig. 6 compares the time-series profiles of six Sentinel-2 bands for each crop type labeled with trusted pixels and CDL. The result respectively displays the profiles from six Sentinel-2 spectral bands (blue, green, red, NIR, SWIR-1, and SWIR-2) for corn (Fig. 6a), cotton (Fig. 6b), rice (Fig. 6c), soybeans (Fig. 6d), and soybeans-wheat double cropping (Fig. 6e). Each profile ranges from 2019 to 2020 and derives from the Sentinel-2 tile "15SYA" where the five major crop types are widely distributed. We aggregate the surface reflectance value over time by specifying the value of the lower quartile (Q1), median (Q2), and upper quartile (Q3) among all pixels of the tile. It can be seen from the result that trusted pixels and CDL pixels of most observed bands have consistent spectral and temporal features over time.



(a) Corn.



(b) Cotton.



(c) Rice.

(*caption on next page*)

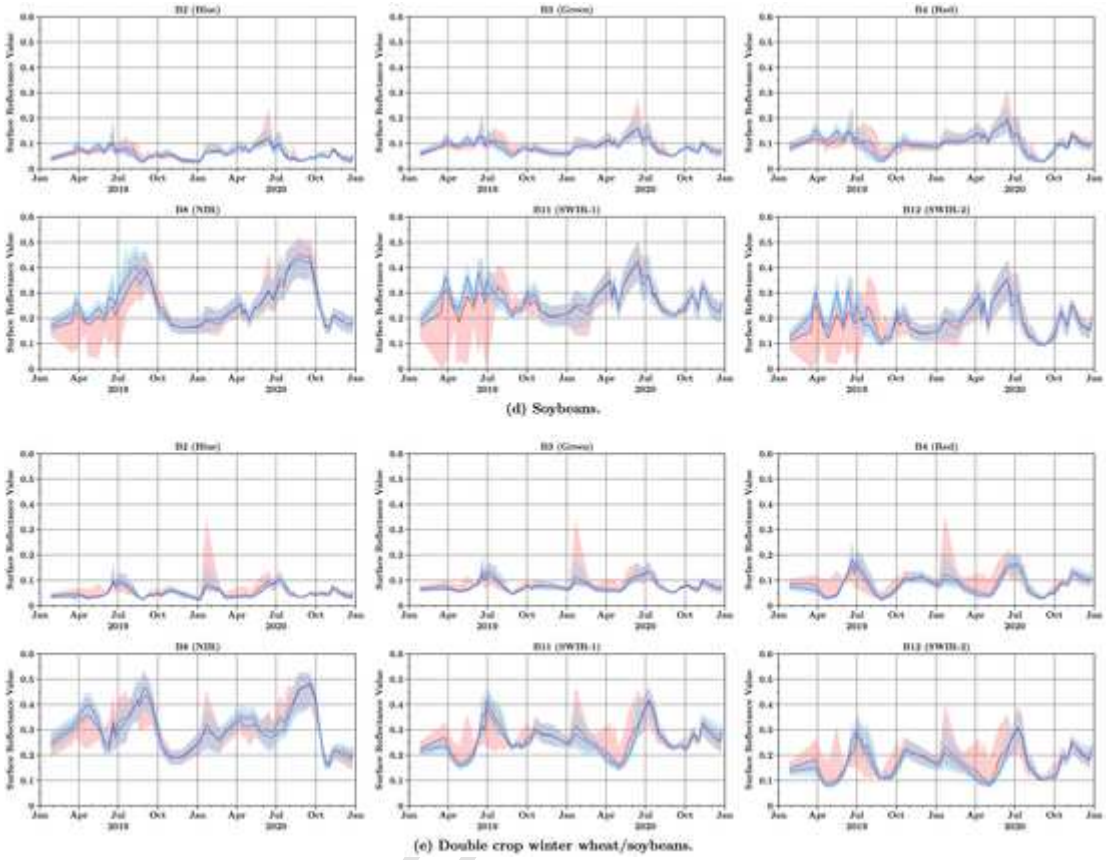


Fig. 6. Comparison of Sentinel-2 spectral profile for crop types labeled with trusted pixels and CDL. Each profile reflects the median value (solid line) and lower/upper quartile value (solid colour) of all trusted pixels (red) or CDL pixels (blue). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Based on the spectral profile, we calculate the R^2 coefficient between the reference data (CDL) and the predicted result (trusted pixels). Fig. 7 illustrates the validation result of the correlation between the trusted pixel's surface reflectance value (red lines in Fig. 6) and CDL's surface reflectance value (blue lines in Fig. 6). The comparison result suggests the spectral profile of corn pixels (Fig. 7a), cotton pixels (Fig. 7b), rice pixels (Fig. 7c), and soybeans-wheat double-cropping pixels (Fig. 7e) labeled with the trusted pixel is highly consistent ($R^2 > 0.9$) with pixels labeled with CDL. The correlation for soybeans pixels (Fig. 7d), especially the SWIR-1 band ($R^2 = 0.76$) and SWIR-2 band ($R^2 = 0.78$), is relatively lower. This error is caused by the offset between spectral profiles of trusted pixels and CDL in 2019, which is noticeable in Fig. 6d.

3.2. In-season crop mapping using supervised classification

This section investigates the capability of in-season crop type mapping using spatiotemporal crop information with supervised classification. As described in Section 2.3.3, GEE provides several common supervised classifier options. The crop type classification presented in this study results from the random forest model. It is well known that the spectral characteristics of different crop types will be differentiated in phenological stages. Thus, we need to examine whether these characteristics can be reflected in the remote sensing data.

Fig. 8 shows the intercomparison of spectral and index profiles for the five crop types. Each value in the profile represents the median value among all trusted pixels of each crop type. It can be interpreted from the comparison result that spectral profiles of corn, cotton, rice, and soybeans start to differentiate at the beginning of the growing season. At around DOY 150, the visible light bands and SWIR bands of corn first began to drop. The profile of rice has a similar trend with corn, but its reflectance value range is wider, especially for the visible light bands. Cotton and soybeans also have similar spectral profiles, but the NIR and SWIR bands of cotton are significantly higher between DOY 180–240 and DOY 130–180. Soybeans-wheat double cropping has the unique profile curves before DOY 180 because the growing season of winter wheat starts from December of the previous year to May of the current year in Mississippi Delta. It can be easily observed from the NDVI profiles that the crop green-up sequence in the study area is corn, rice, cotton, and soybeans, while the NDWI profiles change oppositely in the same sequence.

When training the crop type classification model, the supervised classifier will learn the features in the above spectral and temporal profiles for each crop type. Then the trained model will be able to classify crop types on all pixels over the tile. Fig. 9 demonstrates the details of in-season crop mapping results in 2019 and 2020. In this experiment, we use the 2019 and 2020 CDL as reference data (Fig. 9a) and trusted pixels, which are extracted from 2008 to 2018 CDL and 2008–2019 CDL respectively, as labels of the training dataset (Fig. 9b). Then we perform the crop type classification and generate the in-season cover maps using the multi-temporal Sentinel-2 images from April to the end of April, May, June, July, and August (Fig. 9c). In addition, the difference maps (Fig. 9d) show the improvements in the agreement between CDL and the in-season crop cover maps over time. It is also found from the results that the tiling effect may take place in the early season when available images are significantly inconsistent between adjacent tiles. With the progress of the growing season, more images will be used in the model training (~20 images by July) as well as the spectral profiles between tiles will be getting consistent. Although it is still a challenge to

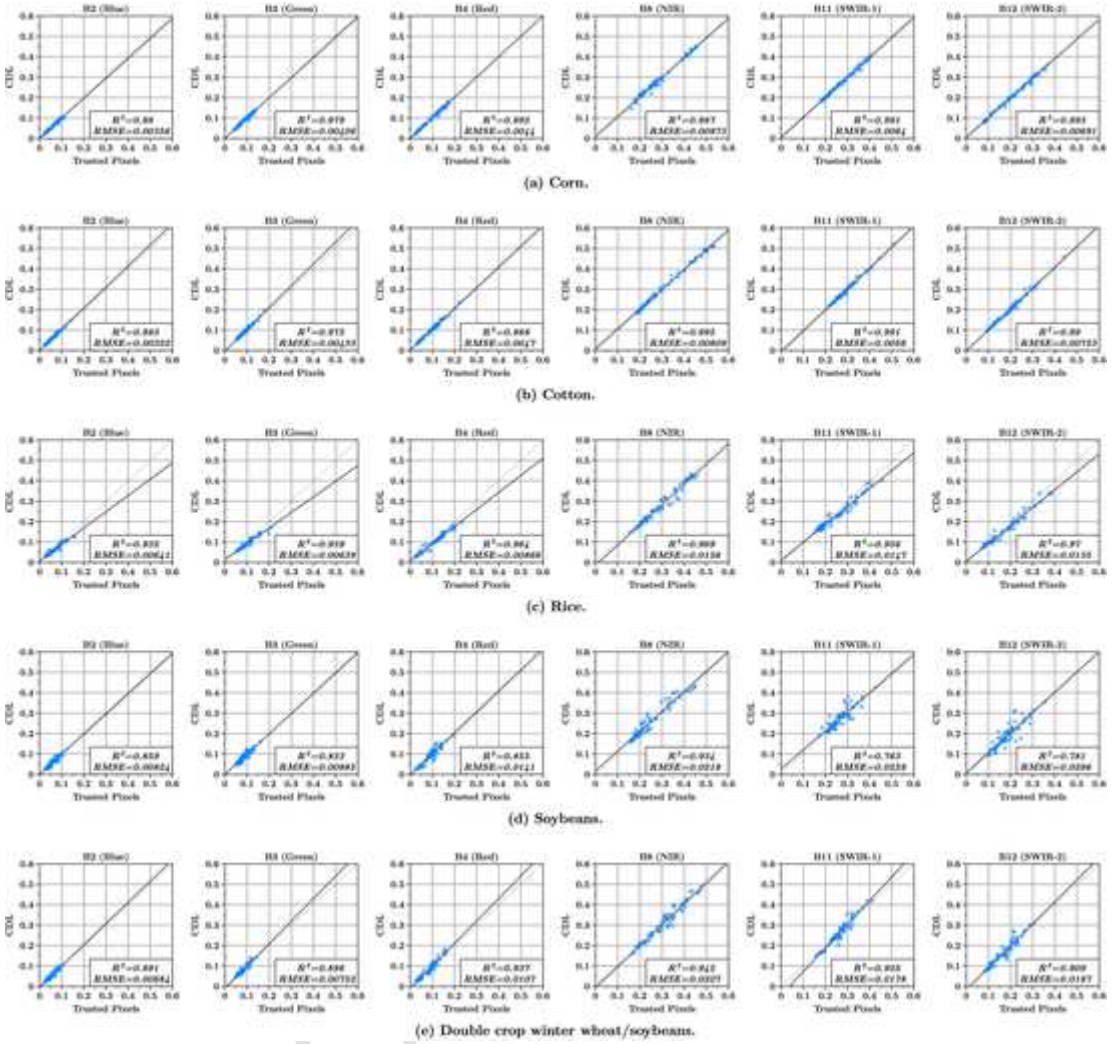


Fig. 7. Validation of surface reflectance value correlation between trusted pixels and CDL. Each subfigure contains in total of 59 sample points, which are taken from the Sentinel-2 time series of 29 images in 2019 and 30 images in 2020 respectively.

ultimately eliminate the tiling effect through the whole season, its impact on the mosaiced crop cover map is limited after July.

3.3. Classification performance evaluation

We use CDL as reference data to measure the classification performance for the crop cover maps created by the proposed framework. Fig. 10 shows the trend of precision, recall, and F1-score for major crop types within the growing season. Among the five types of crops, cotton (red line) and rice (blue line) have the best overall performance (F1-score > 0.9). In comparison, the performance of corn (yellow line), soybeans (green line), and soybeans-wheat double cropping (brown line) has varied between 2019 and 2020. In 2019, the precision of soybeans is significantly high, but the recall is the lowest among all crops during the entire observation period. Meanwhile, the precision and recall rates of corn and soybeans-wheat double cropping are relatively lower than in 2020. Such a high-precision and low-recall result for soybeans indicates more soybeans pixels were omitted and misclassified to other crops, primarily corn and soybeans-wheat double cropping, in 2019. This classification bias mainly results from the offset of spectral profiles for trusted pixels of soybeans in 2019 as indicated in Section 3.1. To sum up, the overall crop type classification performance will be improving with the increase of Sentinel-2 images used between DOY 120 to DOY 210. All crop types will reach the highest classification performance after DOY 220 in all observation years.

To further test the performance of the classification model, we apply the framework to all 12 Sentinel-2 tiles in the study area. Fig. 11 shows the change in crop type classification performance (OA and Kappa) for all tiles within the growing season. Like the classification result of each crop type as above, both OA and Kappa will continue to increase and reach the highest around DOY 220 for most test tiles. It can be concluded that the in-season crop type mapping result is expected to reach 80%–90% agreement with CDL by August in the Mississippi Delta region. According to the historical crop planting and harvesting dates reported by USDA, the most active planting dates for major crop types in Mississippi Delta usually last longer than in the U.S. Corn Belt. For example, in the leading production states of corn and soybeans, such as Iowa and Illinois, the planting dates of major crops typically range from mid-April to early June. But in the Arkansas state, which covers most of the study area, corn is usually planted from Apr 1 to May 15, while soybeans are planted from May 5 to Jun 22. Thus, the best observation dates for in-season mapping may vary among different regions in different years and depend on data availability.

3.4. A prototype on earth engine apps platform

Another main objective of this paper is automating crop type mapping with the proposed framework and exploring its potential scalability. To demonstrate the automation capability, we developed and published an online prototype, In-season Crop Mapping Kit (https://czhang

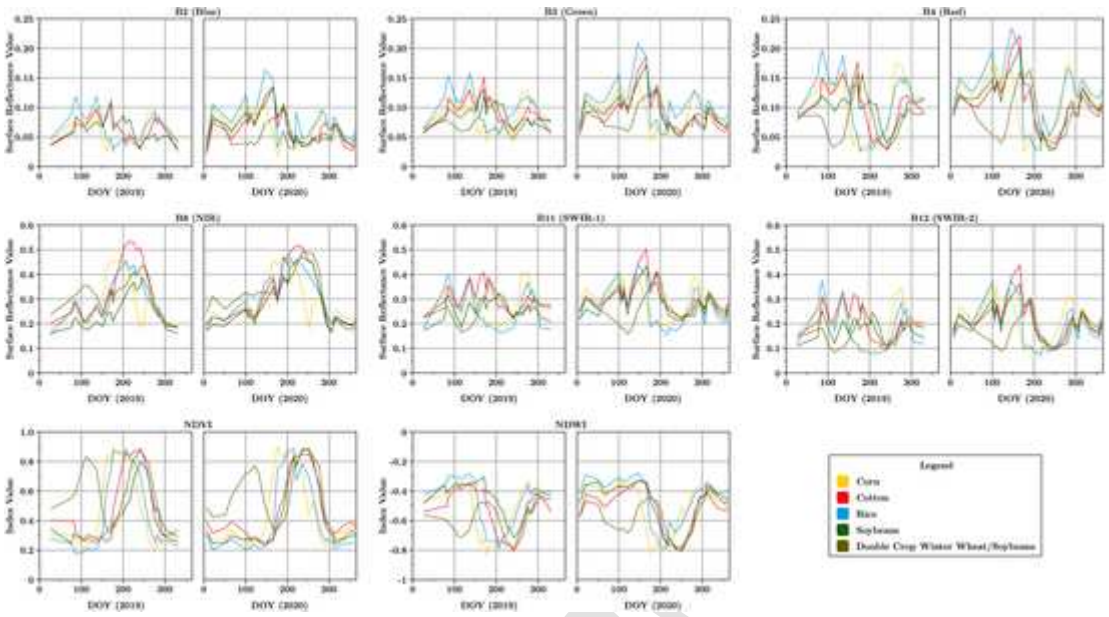


Fig. 8. Intercomparison of spectral and index profile for multi-class trusted pixels in 2019 (left) and 2020 (right).

11.users.earthengine.app/view/agkit4ee-inseason), on the platform of Earth Engine Apps. The prototype is implemented as an extension module of the AgKit4EE toolkit (Zhang et al., 2020), mainly providing remote-sensing-based crop mapping functionalities.

Fig. 12 shows the GUI of the prototype, which is composed of a configuration panel and a map explorer. Before starting crop type mapping, users can specify the target tile with several filters for the Sentinel-2 images (e.g., cloud percentage, year, start/end DOY). The prototype provides additional options allowing users to select classifiers as well customize band combinations for the classification. Based on the given conditions, an on-demand crop cover map will be automatically created and then displayed on the map explorer. To demonstrate the scalability of the framework, we also perform the in-season crop mapping results in other ecoregions with different crop type distributions through the prototype. The 2021 crop cover maps over four Sentinel-2 tiles across the CONUS are illustrated. Specifically, the tile "11SKA" mainly covers the Central Valley Area of California. This area provides diverse crop types, including grains, vegetables, fruits, and nuts. The tile "12TVP" is lying in Eastern Idaho, which is one of the most productive potato-growing areas in the world along with other crops, such as spring wheat, barley, and alfalfa. The tile "14SPG" is over South Central Kansas, where the major crops include corn, soybeans, winter wheat, and soybeans-wheat double cropping. The tile "16TCK" is located within Central Illinois, a typical Corn Belt area covered by corn and soybeans.

4. Discussion

4.1. Applicability and scalability of the approach

This study has addressed the three main problems in remote-sensing-based crop mapping coined at the beginning of the paper. First, the insufficiency of ground truth in the U.S. at the early growing stage has been solved. It is well-known that supervised crop type classification models require plenty of training labels. The main novelty of our framework is using the spatiotemporal crop information extracted from the historical CDL to replace ground truth and label training samples in satellite images. It is also potentially applicable to regions outside the U.S. where historical crop cover maps are available (e.g., Canada with ACI as reference data). This mapping-without-ground-truth approach will significantly simplify the sampling process and save substantial human intervention and financial resources.

The second problem that we have largely settled in this study is the scalability issue in remote-sensing-based crop mapping. Different from those mapping methods that heavily rely on the classification model and algorithm, we designed a general supervised-learning-based framework that is not limited to the classifier, crop type, and geographical conditions. Considering many previous studies have investigated the remote-sensing-based in-season and early-season mapping of corn and soybeans for the U.S. Corn Belt, we choose the Mississippi Delta as the study area, which has a more complex and diverse crop planting environment. We also demonstrated the capability of in-season crop mapping over agricultural fields in California, Idaho, Kansas, and Illinois. Since the trusted pixels are widely distributed across CONUS, it is easy to scale the framework to a large experimental area.

4.2. Accessibility and reusability with Google earth engine

This study demonstrated the capability of GEE for facilitating accessibility and reusability of the automatic in-season crop type mapping. The implementation of our GEE-enabled prototype, In-season Crop Mapping Kit, has significantly simplified the access to in-season crop cover information and the use of the mapping framework. Besides the web application, the source code and APIs are also available through GEE. These open-source materials will substantially accelerate the reuse of the proposed framework by not only the agricultural sectors but also academia, governments, and companies.

Some preliminary results have shown great potential transferability of the machine learning model trained by training samples in the U.S. (Hao et al., 2020). Although it is still challenging to be directly applied in agricultural regions outside the CONUS, the accessibility and reusability feature provided by GEE will enable the user community to expand the mapping-without-ground-truth workflow and explore the transferability of trusted pixels for the global-scale in-season mapping.

4.3. Limitations and potential solutions

The proposed framework labeled satellite images with trusted spatiotemporal crop information, which are derived from historical CDL using monocropping and alternate cropping patterns. There are two potential issues with this approach. On the one hand, the crop rotation

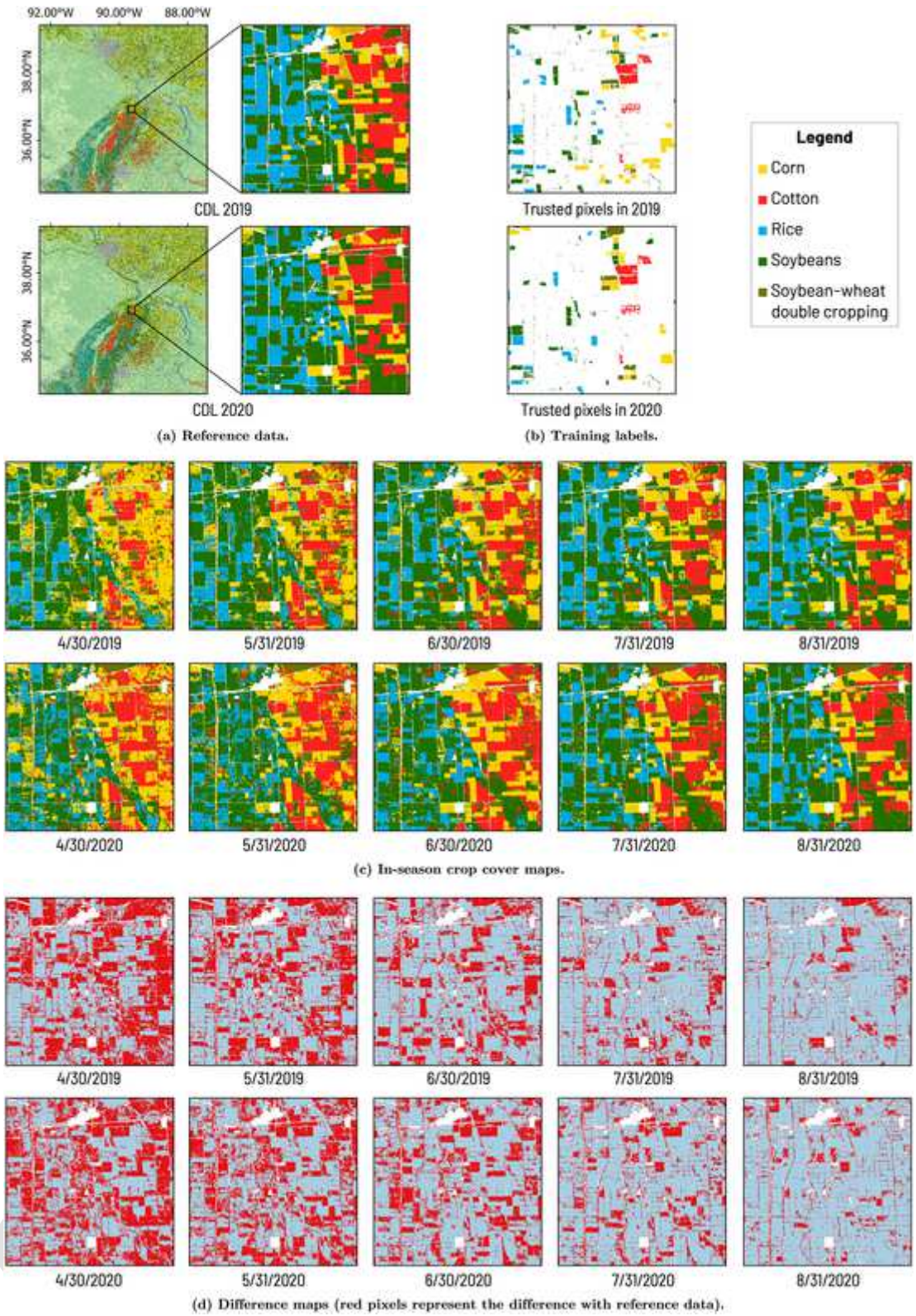


Fig. 9. Examples of 2019 (upper) and 2020 (lower) in-season crop mapping. The reference data are the original CDL data. The training labels are trusted pixels extracted from the CDL time series (2019 and 2020 trusted pixels are respectively extracted from 2009 to 2018 CDL and 2010–2019 CDL). The in-season crop cover maps are the monthly crop type classification results as the growing season progresses. The difference maps highlight the disagreement between each in-season map and CDL.

Fig. 10. Classification performance for all crop types of 2019 (upper) and 2020 (lower). According to the planting stage, the entire season is divided into the planting season (DOY 120–150), the growing season (DOY 150–240), and the harvesting season (DOY 240–270). The solid lines in all charts represent the metric evolution of each crop type as the season progresses.

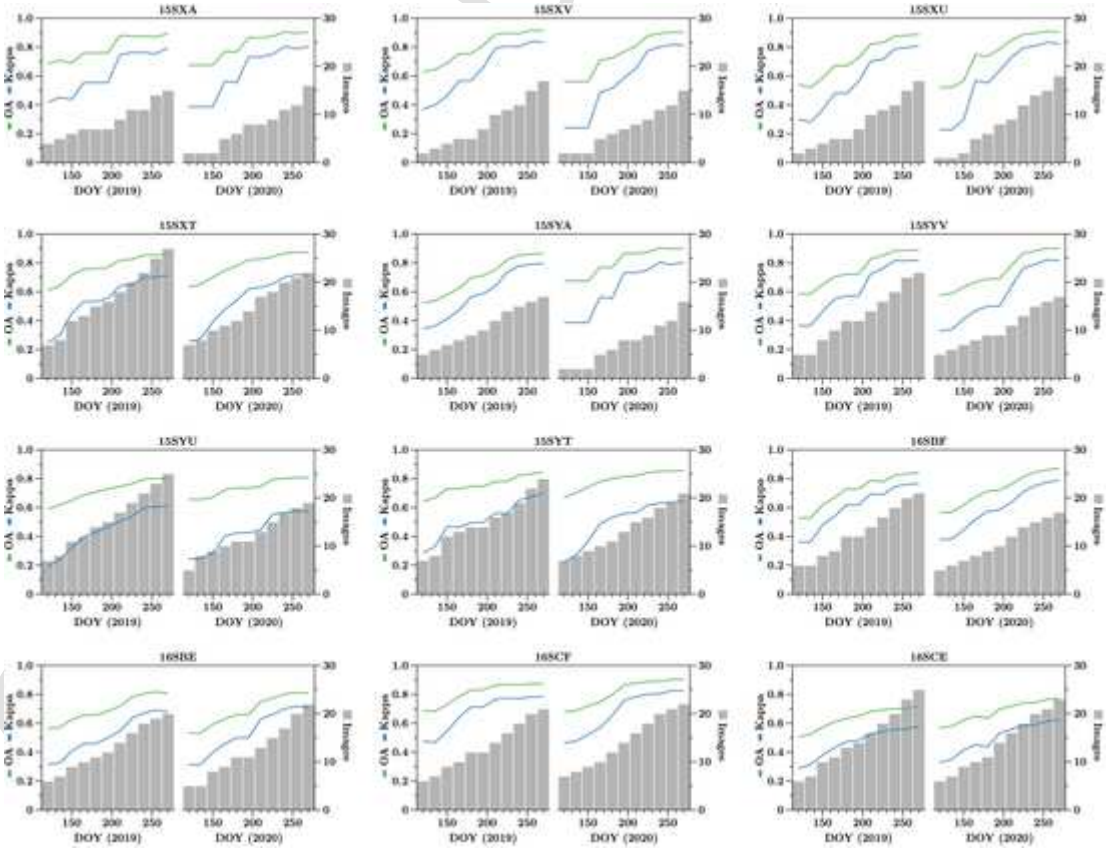


Fig. 11. Overall in-season crop type classification performance for all tiles in 2019 (left) and 2020 (right). Each profile reflects the OA value (green line) and Kappa value (blue line) with the number of used Sentinel-2 images (gray column) through the growing season. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

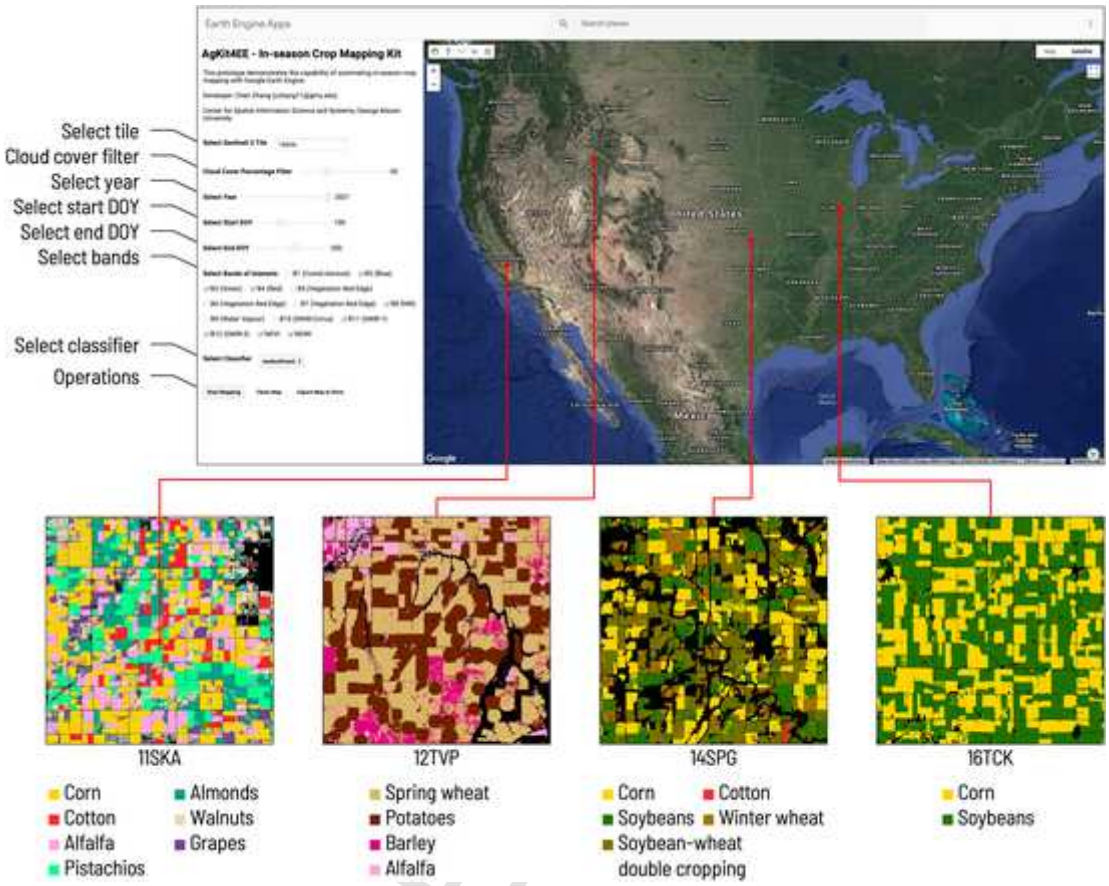


Fig. 12. Graphical user interface of In-season Crop Mapping Kit on Earth Engine Apps (https://czhang11.users.earthengine.app/view/agkit4ee-inseason) and examples of in-season crop cover maps across CONUS in 2021.

patterns of some pixels could be changed due to natural hazards, climate change effects, market situation, loss of soil fertility, government policy, and other dynamic or uncertain factors. From the historical CDL, we found a small number of fallow lands randomly occur every year over the study area. It is still a challenge to correctly predict these fallow lands without ground truth. On the other hand, the framework currently relies on the CDL data for training and validation. However, CDL also contains a certain number of misclassified pixels, mixed pixels, and noisy pixels (Lin et al., 2022). According to the accuracy assessment of 2020 Arkansas CDL by USDA NASS (USDA NASS, 2021), the user's accuracy of major crop types (corn, cotton, rice, soybeans) is between 90% to 97%. The classification performance of the proposed workflow can be further validated using the surveyed ground truth data. The errors in CDL may also affect the accuracy of trusted pixels then impact classification performance. A potential way to improve the classification performance is to combine the CDL confidence layer with CDL (Lark et al., 2021), which may reduce the number of trusted pixels but would further improve the classification accuracy.

The cloud coverage in the remote sensing image is another potential issue that may affect the performance of crop type classification. We set several conditions to filter cloudless Sentinel-2 images in the Sentinel-2 data archive. Although we did not encounter the insufficiency of qualified images in our experiment, it cannot be guaranteed that sufficient qualified Sentinel-2 images are available across the entire CONUS in the following years. Since the Landsat data are also archived in the GEE's data catalog, we will integrate the Landsat 8 and the upcoming Landsat 9 data into the automation framework in the next phase.

5. Conclusion

This study addressed three major challenges for in-season crop type classification raised at the beginning of the paper. First, to solve the insufficiency of training samples in the early growing season, we developed a novel mapping-without-ground-truth approach for in-season crop type classification. By applying crop rotation patterns, a large number of trusted pixels with abundant spatiotemporal crop information were extracted from the historical CDL, which can replace ground truth while labeling training samples in remote sensing images. Second, this mapping-without-ground-truth approach solved the applicability and scalability issue of in-season crop mapping over the CONUS. As a general crop mapping framework, it is also potentially applicable in any agricultural region where the historical crop cover data are available. Third, the release of the associated software significantly improved the accessibility and reusability of the in-season crop mapping framework described in this study. The GEE-enabled online prototype, AgKit4EE - In-season Crop Mapping Kit, and its open-source code/APIs suite can be easily integrated and reused in other agricultural systems and applications.

Specifically, we demonstrated a case study and applied the proposed framework to classify five major crop types (corn, cotton, rice, soybeans, and soybeans-wheat double cropping) of the Mississippi Delta ecoregion. It is found that the spectral and temporal profiles of corn, cotton, rice, and soybeans-wheat double-cropping pixels that are labeled with the trusted pixels are highly correlated ($R^2 > 0.9$) with pixels labeled with CDL. The correlation for soybeans pixels is relatively lower, especially in 2019, but still acceptable. Using these trusted pixels to label training samples with the random forest classifier, the performance of crop type classification is expected to reach 80%–90% agree-

ment with CDL by August within the study area. Meanwhile, the capability of in-season crop mapping across different ecoregions was demonstrated using our online prototype on the Earth Engine Apps platform.

In the future, we will apply the framework to the entire CONUS and produce a CDL-like in-season crop cover data product. Other satellite data sources, such as Landsat 8 and Landsat 9 data, will be tested. Moreover, this study used the default hyperparameter values of the classifier offered by GEE while training the classification model. The impact of hyperparameters with multiple classifiers will be investigated. In the next stage, we will develop a spatiotemporal learning strategy based on the current implementation to maximally transfer trusted pixels and trained models from the U.S. to other study areas.

Data availability

The data described in this work are available through the Earth Engine Apps platform: https://czhang11.users.earthengine.app/view/agkit4ee-inseason. The AgKit4EE - In-season Crop Mapping Kit is open-source software and its source code can be accessed at: https://code.earthengine.google.com/?accept_repo=users/czhang11/agkit4ee-inseason.

CRediT authorship contribution statement

Chen Zhang : Conceptualization, Methodology, Software, Formal analysis, Writing – original draft. **Liping Di :** Conceptualization, Supervision, Funding acquisition, Writing – review & editing. **Li Lin :** Software, Validation. **Hui Li :** Software, Validation. **Liying Guo :** Investigation, Funding acquisition. **Zhengwei Yang :** Investigation, Funding acquisition. **Eugene G. Yu :** Investigation, Funding acquisition. **Yahui Di :** Validation. **Anna Yang :** Validation.

Declaration of Competing Interest

The authors declare that there is no conflict of interest regarding the publication of this article.

Acknowledgments

This research is supported by grants from USDA NIFA (Grant #: 2021-67021-34151, PI: Dr. Liping Di) and National Science Foundation INFEWS program (Grant #: CNS-1739705, PI: Dr. Liping Di).

References

Boryan, C., Yang, Z., Mueller, R., Craig, M., 2011. Monitoring US agriculture: the US Department of Agriculture, National Agricultural Statistics Service, cropland data layer program. Geocarto Int. 26, 341–358. https://doi.org/10.1080/10106049.2011.562309.

Breiman, L., 1984. Classification and Regression Trees. Routledge. https://doi.org/10.1201/9781315139470.

Breiman, L., 2001. Random forests. Mach. Learn. 45, 5–32. https://doi.org/10.1023/A:1010933404324.

Brown, J.F., Pervez, M.S., 2014. Merging remote sensing data and national agricultural statistics to model change in irrigated agriculture. Agric. Syst. 127, 28–40. https://doi.org/10.1016/j.agsy.2014.01.004.

Burges, C.J.C., 1998. A tutorial on support vector machines for pattern recognition. Data Min. Knowl. Disc. 2, 121–167. https://doi.org/10.1023/A:1009715923555.

Cai, Y., Guan, K., Peng, J., Wang, S., Seifert, C., Wardlow, B., Li, Z., 2018. A high-performance and in-season classification system of field-level crop types using time-series Landsat data and a machine learning approach. Remote Sens. Environ. 210, 35–47. https://doi.org/10.1016/j.rse.2018.02.045.

Charvat, K., Horakova, Sarka, Druml, S., Mayer, W., Šafář, V., Kubíčková, H., Kolitzus, D., Lopez, J., Esbrí Palomares, M., Catucci, A., Lucau, C., Mildorf, T., Rybokienė, S., Oort, V., 2021. D5.6 White Paper on Earth Observation Data in Agriculture [WWW Document]. EO4AGRI Consortium. URL. https://eo4agri.eu/.

Cohen, J., 1960. A coefficient of agreement for nominal scales. Educ. Psychol. Meas. 20, 37–46. https://doi.org/10.1177/001316446002000104.

Di, L., Yang, Z., 2014. Foreword to the special issue on agro-Geoinformatics—the applications of geoinformatics in agriculture. IEEE J. Sel. Top. Appl. Earth Obs. Rem. Sens. 7, 4315–4316. https://doi.org/10.1109/JSTARS.2014.2382411.

FAO, 2019. The Future of Food and Agriculture: Alternative Pathways to 2050.

Filippi, P., Whelan, B.M., Vervoort, R.W., Bishop, T.F.A., 2020. Mid-season empirical cotton yield forecasts at fine resolutions using large yield mapping datasets and diverse spatial covariates. Agric. Syst. 184, 102894. https://doi.org/10.1016/j.agsy.2020.102894.

Fisette, T., Rollin, P., Aly, Z., Campbell, L., Daneshfar, B., Filyer, P., Smith, A., Davidson, A., Shang, J., Jarvis, I., 2013. AAFC annual crop inventory. In: 2013 Second International Conference on Agro-Geoinformatics (Agro-Geoinformatics). Presented at the 2013 Second International Conference on Agro-Geoinformatics (Agro-Geoinformatics). pp. 270–274. https://doi.org/10.1109/Argo-Geoinformatics.2013.6621920.

FSIN, 2019. Global Report on Food Crises.

Gao, F., Zhang, X., 2021. Mapping crop phenology in near real-time using satellite remote sensing: challenges and opportunities. J. Rem. Sens. 2021. https://doi.org/10.34133/2021/8379391.

Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., Moore, R., 2017. Google earth engine: planetary-scale geospatial analysis for everyone. Remote Sens. Environ. 202, 18–27. https://doi.org/10.1016/j.rse.2017.06.031.

Hao, P., Zhan, Y., Wang, L., Niu, Z., Shakir, M., 2015. Feature selection of time series MODIS data for early crop classification using random Forest: a case study in Kansas, USA. Remote Sens. 7, 5347–5369. https://doi.org/10.3390/rs70505347.

Hao, P., Di, L., Zhang, C., Guo, L., 2020. Transfer learning for crop classification with cropland data layer data (CDL) as training samples. Sci. Total Environ. 733, 138869. https://doi.org/10.1016/j.scitotenv.2020.138869.

Johnson, D.M., Mueller, R., 2021. Pre- and within-season crop type classification trained with archival land cover information. Remote Sens. Environ. 264, 112576. https://doi.org/10.1016/j.rse.2021.112576.

Khanal, S., Fulton, J., Shearer, S., 2017. An overview of current and potential applications of thermal remote sensing in precision agriculture. Comput. Electron. Agric. 139, 22–32. https://doi.org/10.1016/j.compag.2017.05.001.

Killough, B., Rizvi, S., Lubawy, A., 2021. Advancements in the open data cube and the use of analysis ready data in the cloud. In: 2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS. Presented at the 2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS. pp. 1793–1795. https://doi.org/10.1109/IGARSS47720.2021.9553063.

Konduri, V.S., Kumar, J., Hargrove, W.W., Hoffman, F.M., Ganguly, A.R., 2020. Mapping crops within the growing season across the United States. Remote Sens. Environ. 251, 112048. https://doi.org/10.1016/j.rse.2020.112048.

Lark, T.J., Schelly, I.H., Gibbs, H.K., 2021. Accuracy, Bias, and improvements in mapping crops and cropland across the United States using the USDA cropland data layer. Remote Sens. 13, 968. https://doi.org/10.3390/rs13050968.

Latino, M.E., Corallo, A., Menegoli, M., Nuzzo, B., 2021. Agriculture 4.0 as enabler of sustainable Agri-food: a proposed taxonomy. IEEE Trans. Eng. Manag. 1–20. https://doi.org/10.1109/TEM.2021.3101548.

Li, Z., 2020. Geospatial big data handling with high performance computing: current approaches and future directions. In: Tang, W., Wang, S. (Eds.), High Performance Computing for Geospatial Applications, Geotechnologies and the Environment. Springer International Publishing, Cham, pp. 53–76. https://doi.org/10.1007/978-3-030-47998-5_4.

Lin, L., Di, L., Zhang, C., Guo, L., Di, Y., Li, H., Yang, A., 2022. Validation and refinement of cropland data layer using a spatial-temporal decision tree algorithm. Sci. Data 9, 63. https://doi.org/10.1038/s41597-022-01169-w.

Luo, C., Liu, H., Lu, L., Liu, Z., Kong, F., Zhang, X., 2021. Monthly composites from Sentinel-1 and Sentinel-2 images for regional major crop mapping with Google earth engine. J. Integr. Agric. 20, 1944–1957. https://doi.org/10.1016/S2095-3119(20)63329-9.

McFeeters, S.K., 1996. The use of the normalized difference water index (NDWI) in the delineation of open water features. Int. J. Remote Sens. 17, 1425–1432. https://doi.org/10.1080/01431169608948714.

Nowak, B., Michaud, A., Marliac, G., 2022. Assessment of the diversity of crop rotations based on network analysis indicators. Agric. Syst. 199, 103402. https://doi.org/10.1016/j.agsy.2022.103402.

NRCan, 2014. Fundamentals of Remote Sensing: A Canada Centre for Remote Sensing Remote Sensing Tutorial.

Oshiro, T.M., Perez, P.S., Baranauskas, J.A., 2012. How many trees in a random Forest? In: Perner, P. (Ed.), Machine Learning and Data Mining in Pattern Recognition. Lecture Notes in Computer Science. Springer, Berlin, Heidelberg, pp. 154–168. https://doi.org/10.1007/978-3-642-31537-4_13.

Osinga, S.A., Paudel, D., Mouzakitis, S.A., Athanasiadis, I.N., 2022. Big data in agriculture: between opportunity and solution. Agric. Syst. 195, 103298. https://doi.org/10.1016/j.agsy.2021.103298.

Pelletier, C., Valero, S., Inglada, J., Champion, N., Dedieu, G., 2016. Assessing the robustness of random forests to map land cover with high resolution satellite image time series over large areas. Remote Sens. Environ. 187, 156–168. https://doi.org/10.1016/j.rse.2016.10.010.

Peltonen-Sainio, P., Jauhiainen, L., 2019. Unexploited potential to diversify monotonous crop sequencing at high latitudes. Agric. Syst. 174, 73–82. https://doi.org/10.1016/j.agsy.2019.04.011.

Piedelobo, L., Hernández-López, D., Ballesteros, R., Chakhar, A., Del Pozo, S., González-Aguilera, D., Moreno, M.A., 2019. Scalable pixel-based crop classification combining Sentinel-2 and Landsat-8 data time series: case study of the Duero river basin. Agric. Syst. 171, 36–50. https://doi.org/10.1016/j.agsy.2019.01.005.

Rembold, F., Meroni, M., Urbano, F., Csak, G., Kerdiles, H., Perez-Hoyos, A., Lemoine, G., Leo, O., Negre, T., 2019. ASAP: a new global early warning system to detect anomaly hot spots of agricultural production for food security analysis. Agric. Syst. 168, 247–257. https://doi.org/10.1016/j.agsy.2018.07.002.

Rose, D.C., Chilvers, J., 2018. Agriculture 4.0: broadening responsible innovation in an era of smart farming. Front. Sustain. Food Syst. 2, 87. https://doi.org/10.3389/fsufs.2018.00087.

Santos Valle, S., Kienzle, J., 2020. Agriculture 4.0 – agricultural robotics and automated equipment for sustainable crop production. Integr. Crop Manage. 24.

Sietz, D., Conradt, T., Krysanova, V., Hattermann, F.F., Wechsung, F., 2021. The crop generator: implementing crop rotations to effectively advance eco-hydrological modelling. Agric. Syst. 193, 103183. https://doi.org/10.1016/j.agsy.2021.103183.

Song, X.-P., Huang, W., Hansen, M.C., Potapov, P., 2021. An evaluation of Landsat, Sentinel-2, Sentinel-1 and MODIS data for crop type mapping. Sci. Rem. Sens. 3, 100018. https://doi.org/10.1016/j.srs.2021.100018.

Stehman, S.V., 1997. Selecting and interpreting measures of thematic classification accuracy. Remote Sens. Environ. 62, 77–89. https://doi.org/10.1016/S0034-4257(97)00083-7.

Steinhausen, M.J., Wagner, P.D., Narasimhan, B., Waske, B., 2018. Combining Sentinel-1 and Sentinel-2 data for improved land use and land cover mapping of monsoon regions. Int. J. Appl. Earth Obs. Geoinf. 73, 595–604. https://doi.org/10.1016/j.jag.2018.08.011.

Suchi, S.D., Menon, A., Malik, A., Hu, J., Gao, J., 2021. Crop identification based on remote sensing data using machine learning approaches for Fresno County, California. In: 2021 IEEE Seventh International Conference on Big Data Computing Service and Applications (BigDataService). Presented at the 2021 IEEE Seventh International Conference on Big Data Computing Service and Applications (BigDataService). pp. 115–124. https://doi.org/10.1109/BigDataService52369.2021.00019.

Tatsumi, K., Yamashiki, Y., Canales Torres, M.A., Taipe, C.L.R., 2015. Crop classification of upland fields using random forest of time-series Landsat 7 ETM+ data. Comput. Electron. Agric. 115, 171–179. https://doi.org/10.1016/j.compag.2015.05.001.

Tucker, C.J., 1979. Red and photographic infrared linear combinations for monitoring vegetation. Remote Sens. Environ. 8, 127–150. https://doi.org/10.1016/0034-4257(79)90013-0.

UN, 2015. Transforming our World: The 2030 Agenda for Sustainable Development.

UN, 2019. World Population Prospects 2019. United Nations, Department of Economic and Social Affairs, Population Division.

USDA NASS, 2010. Usual Planting and Harvesting Dates for U.S. Field Crops. USDA NASS.

USDA NASS, 2021. 2020 Arkansas Cropland Data Layer (data.raster Digital Data). USDA, NASS, Spatial Analysis Research Section.

USDA-NASS, 2019. CropScape and Cropland Data Layers - FAQs [WWW Document]. URL. https://www.nass.usda.gov/Research_and_Science/Cropland/sarsfaqs2.phphttps://www.nass.usda.gov/Research_and_Science/Cropland/sarsfaqs2.php. (accessed 2.15.19).

Varmaghani, A., Eichinger, W.E., 2016. Early-season classification of corn and soybean using Bayesian discriminant analysis on satellite images. Agron. J. 108, 1880–1889. https://doi.org/10.2134/agronj2015.0454.

Verdouw, C., Tekinerdogan, B., Beulens, A., Wolfert, S., 2021. Digital twins in smart farming. Agric. Syst. 189, 103046. https://doi.org/10.1016/j.agsy.2020.103046.

Wagemann, J., Siemen, S., Seeger, B., Bendix, J., 2021. A user perspective on future cloud-based services for big earth data. Int. J. Digital Earth 14, 1758–1774. https://doi.org/10.1080/17538947.2021.1982031.

Wang, S., Azzari, G., Lobell, D.B., 2019. Crop type mapping without field-level labels: random forest transfer and unsupervised clustering techniques. Remote Sens. Environ. 222, 303–317. https://doi.org/10.1016/j.rse.2018.12.026.

Wang, Y., Zhang, Z., Feng, L., Ma, Y., Du, Q., 2021. A new attention-based CNN approach for crop mapping using time series Sentinel-2 images. Comput. Electron. Agric. 184, 106090. https://doi.org/10.1016/j.compag.2021.106090.

Wieme, R.A., Carpenter-Boggs, L.A., Crowder, D.W., Murphy, K.M., Reganold, J.P., 2020. Agronomic and economic performance of organic forage, quinoa, and grain crop rotations in the Palouse region of the Pacific northwest, USA. Agric. Syst. 177, 102709. https://doi.org/10.1016/j.agsy.2019.102709.

Xu, J., Yang, J., Xiong, X., Li, H., Huang, J., Ting, K.C., Ying, Y., Lin, T., 2021. Towards interpreting multi-temporal deep learning models in crop mapping. Remote Sens. Environ. 264, 112599. https://doi.org/10.1016/j.rse.2021.112599.

Yang, G., Yu, W., Yao, X., Zheng, H., Cao, Q., Zhu, Y., Cao, W., Cheng, T., 2021. AGTOC: a novel approach to winter wheat mapping by automatic generation of training samples and one-class classification on Google earth engine. Int. J. Appl. Earth Obs. Geoinf. 102, 102446. https://doi.org/10.1016/j.jag.2021.102446.

Zhai, Z., Martínez, J.F., Beltran, V., Martínez, N.L., 2020. Decision support systems for agriculture 4.0: survey and challenges. Comput. Electron. Agric. 170, 105256. https://doi.org/10.1016/j.compag.2020.105256.

Zhang, C., Di, L., Lin, L., Guo, L., 2019. Machine-learned prediction of annual crop planting in the U.S. Corn Belt based on historical crop planting maps. Comput. Electron. Agric. 166, 104989. https://doi.org/10.1016/j.compag.2019.104989.

Zhang, C., Di, L., Yang, Z., Lin, L., Hao, P., 2020. AgKit4EE: a toolkit for agricultural land use modeling of the conterminous United States based on Google earth engine. Environ. Model. Softw. 129, 104694. https://doi.org/10.1016/j.envsoft.2020.104694.

Zhang, C., Di, L., Hao, P., Yang, Z., Lin, L., Zhao, H., Guo, L., 2021a. Rapid in-season mapping of corn and soybeans using machine-learned trusted pixels from cropland data layer. Int. J. Appl. Earth Obs. Geoinf. 102, 102374. https://doi.org/10.1016/j.jag.2021.102374.

Zhang, C., Di, L., Yang, Z., Lin, L., Zhao, H., Yu, E.G., 2021b. An overview of agriculture cyberinformatics tools to support USDA NASS decision making. In: 2021 9th International Conference on Agro-Geoinformatics. Presented at the 2021 9th International Conference on Agro-Geoinformatics. pp. 1–6. https://doi.org/10.1109/Agro-Geoinformatics50104.2021.9530327.

Zhang, C., Yang, Z., Di, L., Lin, L., Hao, P., Guo, L., 2021c. Applying machine learning to cropland data layer for agro-Geoinformation discovery. In: 2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS. Presented at the 2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS. pp. 1149–1152. https://doi.org/10.1109/IGARSS47720.2021.9554628.

Zhong, L., Hu, L., Zhou, H., 2019. Deep learning based multi-temporal crop classification. Remote Sens. Environ. 221, 430–443. https://doi.org/10.1016/j.rse.2018.11.032.

Zurqani, H.A., Allen, J.S., Post, C.J., Pellett, C.A., Walker, T.C., 2021. Mapping and quantifying agricultural irrigation in heterogeneous landscapes using Google earth engine. Rem. Sens. Appl. 23, 100590. https://doi.org/10.1016/j.rsase.2021.100590.