

Automated 10-m Resolution In-season Crop-type Data Layer Mapping for Contiguous United States

Received: 31 October 2025

Accepted: 19 March 2026

Cite this article as: Li, H., Di, L., Zhang, C. *et al.* Automated 10-m Resolution In-season Crop-type Data Layer Mapping for Contiguous United States. *Sci Data* (2026). <https://doi.org/10.1038/s41597-026-07099-1>

Hui Li, Liping Di, Chen Zhang, Liying Guo, Eugene G. Yu, Bosen Shao, Ziao Liu & Hanxi Li

We are providing an unedited version of this manuscript to give early access to its findings. Before final publication, the manuscript will undergo further editing. Please note there may be errors present which affect the content, and all legal disclaimers apply.

If this paper is publishing under a Transparent Peer Review model then Peer Review reports will publish with the final article.

Automated 10-m Resolution In-season Crop-type Data Layer Mapping for Contiguous United States

Hui Li¹, Liping Di^{1*}, Chen Zhang¹, Liying Guo¹, Eugene G. Yu¹, Bosen Shao¹, Ziao Liu¹, Hanxi Li¹

¹Center for Spatial Information Science and Systems, George Mason University, Fairfax, VA, 22030, USA.

*e-mail: ldi@gmu.edu

Nationwide in-season crop planting data is critical for timely agricultural decision-making and application development in the U.S. Currently, the primary source of crop planting data is the Cropland Data Layer (CDL), an annual product from the U.S. Department of Agriculture (USDA) that is available to public in February of following year, mainly supporting post-season applications. To address the need for high-resolution, in-season crop planting information, we developed an automated crop-type mapping workflow to produce a new data product: 10m resolution In-season Crop-type Data Layer (ICDL) maps for June, July, and August of current, available publicly with a delay of only 5 days. The workflow extracts training labels from historical CDL data and incorporates Sentinel-2 and Landsat 8/9 observations to conduct supervised time-series classifications. The outputs are assembled using a multilevel mosaicking process to produce the Contiguous U.S. ICDL. Validation of the ICDL product demonstrated its high accuracy. Training labels accuracies ranged from 0.825 to 0.937, while classification accuracies improved from 0.807 in June to 0.984 in August, consistently outperforming the annual CDL. Moreover, ICDL-based acreage estimates for major crops showed close agreement with official USDA National Agricultural Statistics Service (NASS) statistics. The ICDL datasets are publicly available on the CropSmart web portal, providing timely, high-resolution crop-type information that can directly support national-scale agricultural monitoring, management, and decision-making.

Background & Summary

U.S. crop production is not only critical for global food supply but also an important source of raw materials for renewable energy and other industries, thereby deeply influencing global food, energy, and economic security¹. Many crop management and associated decision-makings have to be made while the crops are still growing. In this context, having a high-resolution, in-season, national-scale crop-type map is indispensable—it offers timely insights into where, what and how many acres crops are planted across the growing season, supporting decision makings in the entire crop production chain, such as agricultural planning, yield forecasting, biofuel logistics, commodity pricing, and food security assessment. The USDA's Cropland Data Layer (CDL) has long served as a cornerstone dataset for supporting cropping decision making. It offers annual, 10-30 meter resolution, crop-specific land cover information across the Contiguous United States (CONUS) through the NASS program^{2,3}. Although its overall crop classification accuracy typically ranges from 85% to 95% for major crops, the CDL's post-season release—often in the early months of the following year—limits its value for time-sensitive, in-season applications³⁻¹¹.

The effectiveness of some timely crop mapping studies depends heavily on the availability of reliable crop-type ground survey data, which is crucial for generating accurate training data to train models/algorithms for classifying satellite imagery. This data directly influence model performance by guiding the classification process. This survey data is often collected through field surveys and observation

to build the training dataset^{12,13}. Rußwurm et al. 2023¹⁴ introduced ELECTS, an end-to-end recurrent neural network designed for early classification of time-series data for in-season crop type mapping. By combining multi-source remote sensing imagery with official crop records, the model successfully performed in-season crop-type mapping across four diverse regions with strong accuracy. Nevertheless, in-season ground surveys, reported data and official crop records are usually inaccessible by public, and the manual label collection is highly labor- and cost-intensive. Researchers have also turned to historical cropland products as a source of crop training label. Konduri et al. 2020 produced growing season crop maps across the U.S., by using CDL pixels and MOD13Q1 NDVI data to build the Mapcurves model¹⁵. The map reached 90% of accuracy at the end of August for dominant crops within test years (2015–2018). The map lacks sharpness along fine field boundaries as reported. Tran et al. 2022 extracted training labels from good quality CDL pixels, combined with purified Sentinel-2 pixels to training the random forest classifier, and produced post-season 10m crop type map in 2019 for South Dakota and California¹⁶. The study reported overall accuracy of 94% and 83% for these two states. Johnson et al. 2021 evaluated four data integration strategies—CDL, Landsat 7/8, Sentinel-2, and their combination—within a random forest framework to generate land cover maps for the U.S. Corn Belt¹⁷. Their results showed that combining CDL with full-season Landsat 7/8 and Sentinel-2 imagery provided the highest classification accuracy for corn and soybean.

The above-mentioned studies are either for a small geographic area or for producing the crop-type map at the end of a growing season or post season. One of the biggest challenges to generate in-season crop-type maps for large geographic areas, e.g., the CONUS, is the unavailability of ground truth data evenly distributed over the large mapping area. Several studies have addressed these challenges using a “Trusted Pixels” approach, which derive reliable early-season training labels from historical CDL^{6-8,17-20}. Building upon this approach, our previous study implemented the production of 30m In-season Crop-Type Data Layer (ICDL) maps covering Contiguous U.S. at the end of May, June, and July in 2022 without relying on current ground survey data⁶. The results are highly accurate: in Nebraska and Iowa, 30m ICDL achieved F1 scores of 0.911 and 0.845 for corn, and 0.959 and 0.969 for soybean by late July 2022. In this study, we improved our previous production workflow to enable automatically generate Contiguous United States (CONUS) monthly 10m ICDL maps in the beginning of July, August and September from 2022 onward. Mapping performance was rigorously evaluated using both ground survey data and official crop acreage statistics for 2022 and 2023. The improvements in data production include increasing spatial resolution from 30 m to 10 m, adjusting the production period from May–July to June–August to better align with the peak growing season for increasing the mapping accuracy, updating training label sampling strategy to avoid training bias, and reaching higher overall accuracy.

Methods

Study area and growing season period. The CONUS is study area in this paper, which encompasses 48 States and the District of Columbia, as shown in Fig. 1. All U.S. major crop types, such as corn, soybean, wheat, cotton, and rice, were traditionally cultivated in the CONUS.

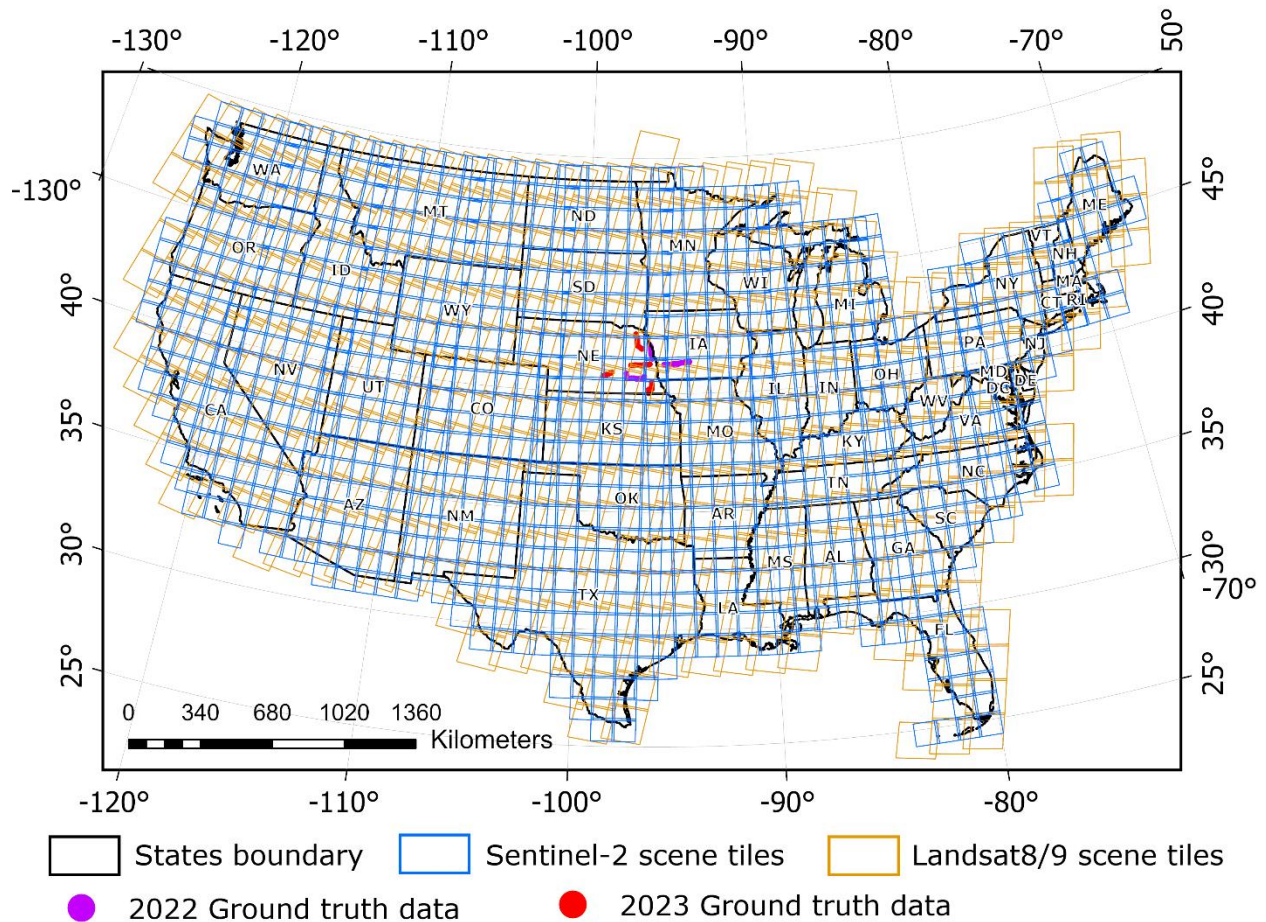


Fig. 1. Study area of CONUS, Sentinel-2 and Landsat 8/9 scenes covering CONUS

This study produced tens of crop and land cover types in-season maps and focused on the mapping performance of major crop: corn, soybean and cotton. Corn and soybean, as most significant types of row crops in the U.S., primarily grow in Corn Belt region, including 13 midwestern States, including Illinois, Indiana, Iowa, Kansas, Kentucky, Michigan, Minnesota, Missouri, Nebraska, North Dakota, Ohio, South Dakota, and Wisconsin²¹. Cotton is mainly cultivated in the southern states of the U.S., like Georgia and Texas. These three crops' growth season is typically divided into three phases: planting as early as in April-May, mid-season in July and August, and harvesting as late as in November. This study concentrates on the summer season (June-August) to produce monthly ICDL for the CONUS. Fig.2 illustrates the growth session durations of three major crops.

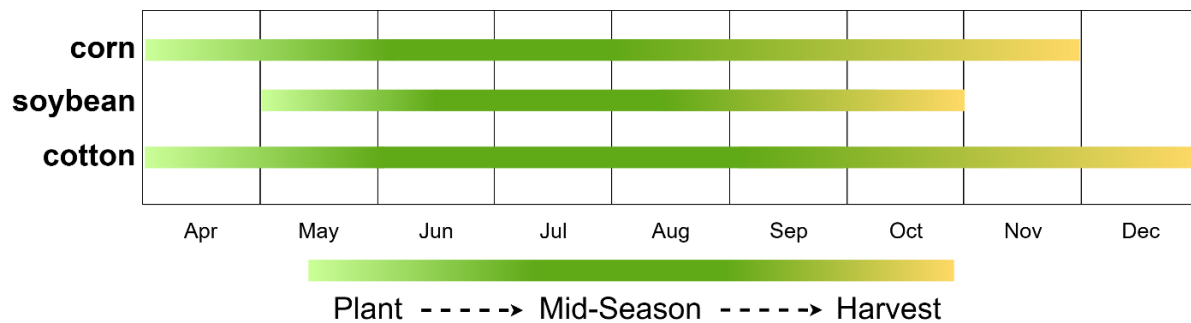


Fig. 2. Major crops growth season^{6,22}

Data collection. This study utilized multispectral surface reflectance data from Sentinel-2²³ and Landsat 8/9²⁴. Sentinel-2 provides 13 spectral bands at up to 10 m resolution with a 5-day revisit cycle²³, while Landsat 8/9 offers seven comparable bands at 30 m resolution with a 16-day revisit cycle²⁴. Both datasets include Blue, Green, Red, NIR, SWIR-1, and SWIR-2 bands, though wavelength ranges differ (Table 1). Tiles intersecting the CONUS boundary (Fig. 1) were selected, yielding 990 Sentinel-2 and 459 Landsat 8/9 tiles. For each tile, time-series images with <10% cloud cover and <10% missing data were collected for May–June, May–July, and May–August 2022 to capture key crop growth stages. CDL²⁵ data from 2016–2021 and 2017–2022 were collected to generate training labels for 2022 and 2023. The U.S. TIGER 2018 state boundary shapefile²⁶ were used to define the mapping area. All data collection was performed in Google Earth Engine (GEE) Colab²⁷ using data paths: “COPERNICUS/S2_SR_HARMONIZED”, “LANDSAT/LC08/C02/T1_L2”, “LANDSAT/LC09/C02/T1_L2”, “USDA/NASS/CDL”, and “TIGER/2018/States”, with all processing codes available in the GitHub repository described in the Code Availability section. Ground truth data were collected in Nebraska and Iowa in 2022 and in Nebraska in 2023, comprising 313 field polygons in 2022 and 205 polygons in 2023 for pixel-level validation. NASS’s reported acreages (NRA)²⁸ were collected to evaluate the state-level major crop-type acreage estimation.

Table 1. Sentinel-2 and Landsat 8/9 imagery properties

Spectrum channel	Sentinel-2A		Landsat-8/9	
	Resolution(meter)	Wavelength (μm)	Resolution(meter)	Wavelength (μm)
Blue	10 m	0.492 -0. 497	30 m	0.45 - 0.51
Green	10 m	0.559 - 0.560	30 m	0.53 - 0.59
Red	10 m	0.665 - 0.665	30 m	0.64 - 0.67
NIR	10 m	0.833-0.8351	30 m	0.85 - 0.88
SWIR 1	20 m	1.610 -1.614	30 m	1.57 - 1.65
SWIR 2	20 m	2.186 - 2.202	30 m	2.11 - 2.29

Automated workflow for in-season mapping. This study developed an automated workflow to generate 10 m ICDL over CONUS using a time-series classification framework (Fig. 3). Training labels derived from historical CDL data were paired with May–June, May–July, and May–August multispectral temporal data within each satellite tile to train random forest classifiers, producing crop-type maps for June, July, and August. Sentinel-2 and Landsat 8/9 classifications were conducted separately, and tile-level outputs were mosaicked into CONUS-wide maps. Final monthly 10 m ICDL products were generated by integrating Landsat 8/9 and Sentinel-2 results followed by post-processing. Classification was implemented in Google Colab²⁷, while mosaicking and refinement were performed locally using GDAL²⁹ with multiprocessing. The workflow runs automatically at the beginning of July, August, and September as new satellite data become available.

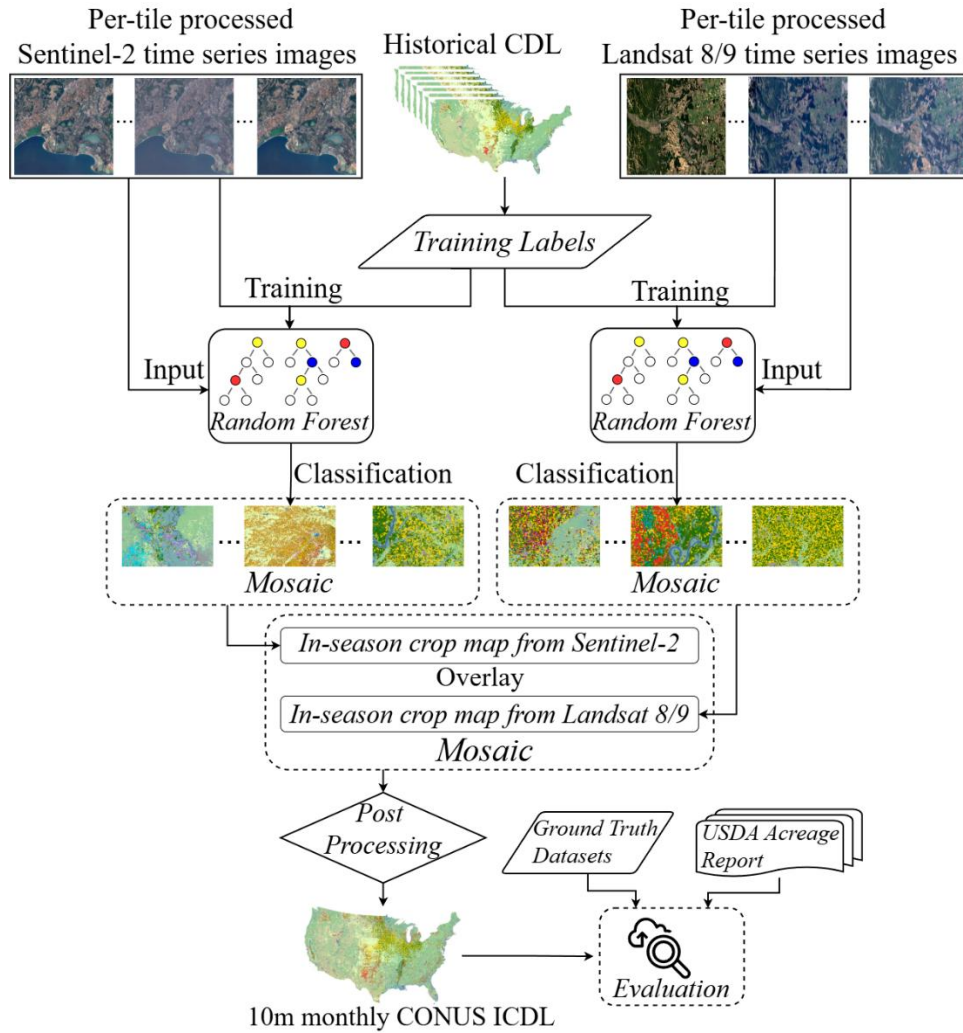


Fig. 3. Automated in-season crop-type mapping workflow

Extraction of training labels. Training labels were derived using stable crop rotation patterns to predict current crop types. Rotations were classified as consecutive, where the same crop is continuously cultivated over six years (e.g., corn–corn–corn...), or alternative, where a crop is grown every other year (e.g., corn–soybean–corn...). Pixels exhibiting these patterns in historical CDL are highly predictive of current-year crops and were extracted as training labels. This approach has been validated and widely applied in previous studies^{6–8,19}. In this study, 2022 and 2023 training labels were extracted from 2016–2021 and 2017–2022 CDL, respectively.

Satellite multispectral data processing. To ensure consistency, Sentinel-2 SWIR1 and SWIR2 bands (originally 20 m) and Landsat 8/9 bands (30 m) were resampled to 10 m using bilinear interpolation. This study used Blue, Green, Red, NIR, SWIR1, SWIR2, Normalized Difference Vegetation Index (NDVI), and Normalized Difference Water Index (NDWI) as classification indices. NDVI³⁰ and NDWI³¹ were calculated as:

$$NDVI = \frac{NIR - RED}{NIR + RED} \quad (1)$$

$$NDWI = \frac{GREEN - NIR}{GREEN + NIR} \quad (2)$$

Filtered images were then sorted by acquisition time via the `system:time_end` attribute to assemble image stacks for each observation tile.

Classifier training and classification. Training labels were integrated with Sentinel-2 and Landsat 8/9 time-series data per tile to construct training datasets. To reduce bias across crop types, a stratified sampling strategy was applied, extracting 1,000 random samples per land-cover type per tile for each classification month. These datasets were used to train random forest classifiers with 100 decision trees for every tile, due to their robustness to outliers and noise, reduced risk of overfitting, and strong predictive performance even with noisy data^{6,32–36}. Sentinel-2 and Landsat 8/9 classifications were performed separately; for example, a Sentinel-2 tile’s June crop map was generated using the corresponding May–June imagery and trained classifier.

Mosaic and post-processing. Two monthly ICDL maps—one from Sentinel-2 and one from Landsat 8/9—were generated by mosaicking all tile-level classifications. Only high-quality imagery was used, resulting in occasional data gaps. The final monthly ICDL map was produced by prioritizing Sentinel-2 over Landsat 8/9 to reduce gaps, with pixels lacking valid observations in both datasets marked as ‘NoData.’ An eight-neighborhood majority filter was then applied to remove isolated spurious pixels.

Evaluation metrics. Mapping performance was evaluated using two approaches: pixel-level accuracy and state-level acreage comparison. Ground-surveyed crop polygons served as reference data for pixel-level validation, from which User Accuracy, Producer Accuracy, Overall Accuracy, and F1 (Equation 3) were derived. At the state level, major crop (corn, soybean, cotton) acreage from the August ICDL maps was compared with USDA NASS³⁷ statistics, using percentage difference (Δ), coefficient of determination (R^2), and mean squared error (MSE) as metrics. CDL products were included for comparison in both validation steps.

$$F1 = \frac{2 * UserAccuracy * ProducerAccuracy}{UserAccuracy + ProducerAccuracy} \quad (3)$$

Data Records

The dataset is available at Zenodo repositories^{38–40}. These repositories contain validated ICDL maps for 2022 and 2023, organized into six files: `Inseason2022June10m.tif`, `Inseason2022July10m.tif`, `Inseason2022Aug10m.tif`, `Inseason2023June10m.tif`, `Inseason2023July10m.tif`, and `Inseason2023Aug10m.tif`. All files are provided in GeoTIFF format with EPSG:5070 projection at 10 m spatial resolution covering the contiguous United States (CONUS). Each file contains a single band in which pixel values represent land cover classes consistent with the CDL classification scheme⁴¹.

Data Overview

In 2022–2023, Sentinel-2 and Landsat 8/9 images for May–June, May–July, and May–August were automatically processed on July 1, August 1, and September 1, generating CONUS-wide 10 m ICDL maps by July 5, August 5, and September 5. Maps use CDL land-cover codes and colors to show near–real-time crop type distributions. Mapping performance improved from June to August (Fig. 4), with clearer field boundaries and more homogeneous patterns. This trend is consistent with crop phenology: early-season emergence in June results in greater spectral mixing and lower accuracy, whereas peak canopy development in July and August enhances spectral separability and reduces within-field variability. Mapping performance is substantially improved during the mid- to late-growing season due to richer spectral information and accumulated temporal observations.

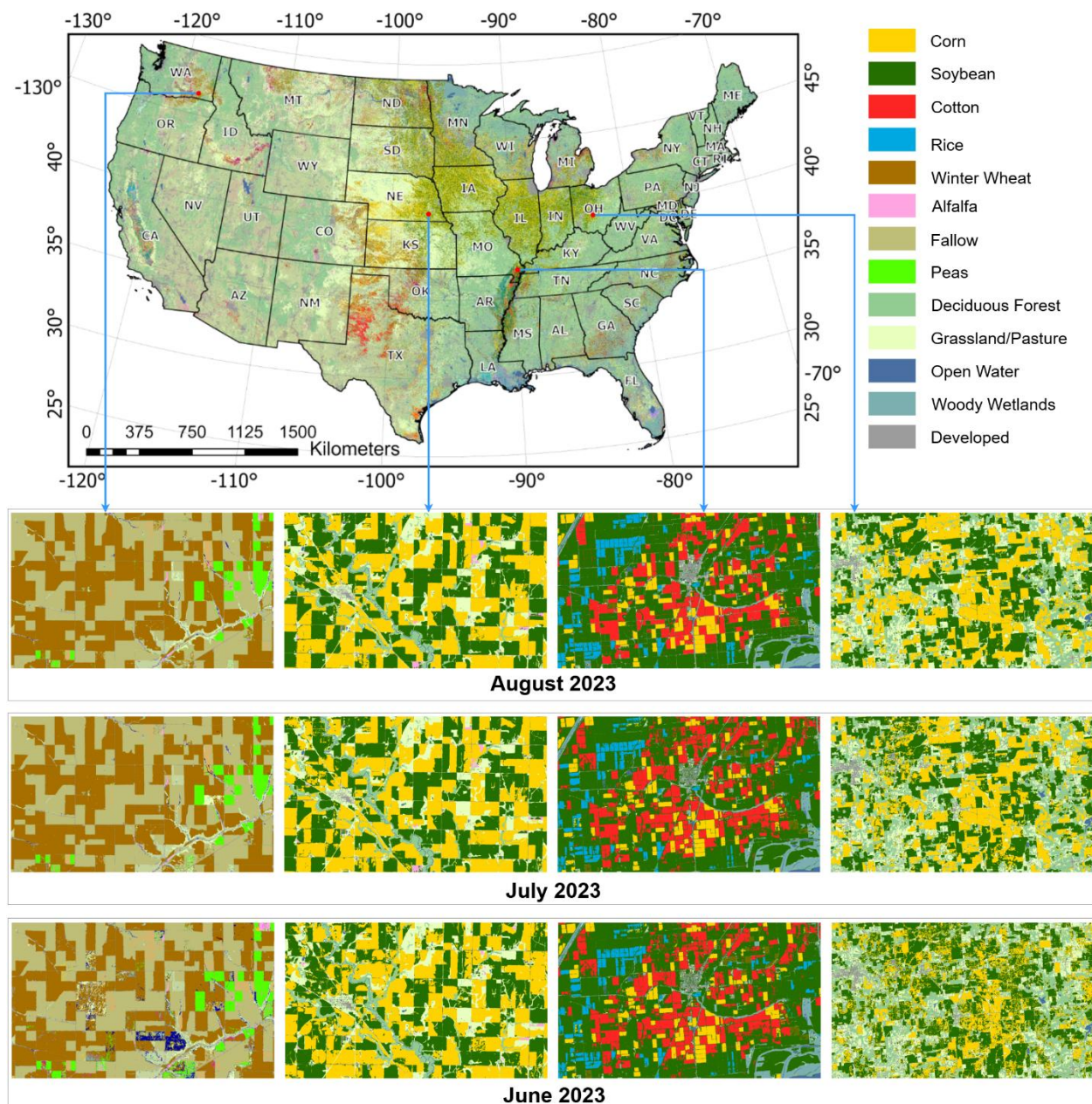


Fig. 4. 2023 CONUS 10m in-season crop-type maps

Technical Validation

We conducted both pixel-level and state-level validations to assess the quality of the ICDL maps. Ground-truth data from Nebraska and Iowa for two years were used for pixel-level validation of training label and classification accuracy. Validation of dominant crops, such as corn and soybean, within the U.S. Corn Belt provides a strong indicator of map quality in terms of both accuracy and precision. Nebraska and Iowa together account for approximately one-third of total corn acreage and one-fifth of total soybean acreage in the Corn Belt²⁸, a region characterized by relatively homogeneous cultivation conditions, making these two states statistically and agronomically representative of the region. Randomly collected ground-truth data from two states ensure unbiased validation under representative production conditions, while accounting for practical limitations in large-scale ground-truth data collection. The user accuracy, producer

accuracy, overall accuracy, and F1 score between ground truth data and produced data were calculated to analyze the data quality. State-level validation was further conducted by comparing ICDL-estimated crop acreages with official acreage statistics²⁸ across multiple states and three major crop types over two years. CDL was included in these two validations for quality comparison.

Validation of training labels. Fig. 5 illustrates the spatial distribution of training label layer in Nebraska in 2023. For assessment, corn and soybean field polygons were intersected with the training label layer, from which 500 validation points were randomly selected per state to calculate accuracy metrics. The result is shown in Table 2, illustrating that in 2022, the overall accuracy in Nebraska (0.825) is lower than Iowa (0.920), and Nebraska's soybean F1(0.784) is lower than corn (0.854). Nebraska training label accuracy increases to high level in 2023. These results indicate that training labels maintain medium-to-high accuracy overall, but exhibit uneven distribution across region, time, and crop variety, reflecting slight discrepancies in rotation patterns across geographical regions and time periods.

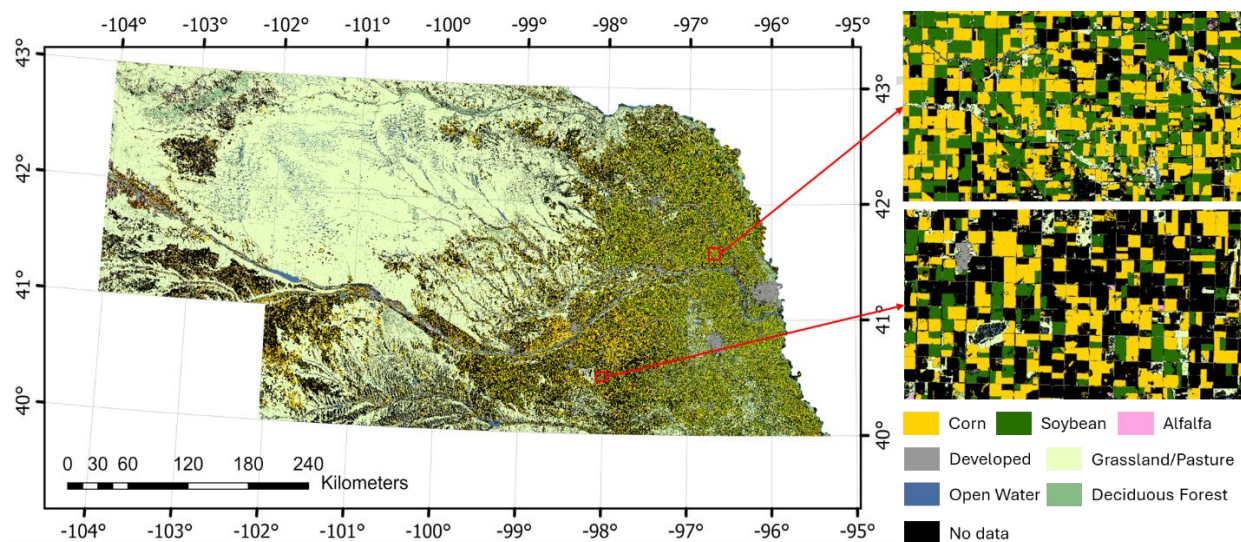


Fig. 5. Training labels of Nebraska in 2023

Table 2. Training labels accuracy assessment

Year	State	Crop	User Accuracy	Producer Accuracy	Overall Accuracy	F1
2022	Nebraska	corn	0.811	0.902	0.825	0.854
		soybean	0.856	0.724		0.784
	Iowa	corn	0.939	0.876	0.920	0.906
		Soybean	0.881	0.941		0.910
2023	Nebraska	corn	0.990	0.931	0.937	0.960
		Soybean	0.853	0.953		0.900

Validation of classification accuracy. Accuracy metrics between ground truth and monthly ICDL maps were calculated from confusion matrices in 2022 and 2023 across Nebraska and Iowa, with the same metrics computed for CDL for the corresponding years and states (Tables 3, 4, and 5). Evaluation focused on the dominant regional crop types—corn and soybean; all other land cover classes were assigned a pixel value of zero as category "other".

In 2022, Nebraska ICDL maps show an increasing accuracy trend from June to August. Overall accuracy starts at 0.807 in June and climbs to 0.927 in August, exceeding the 2022 CDL overall accuracy of 0.916. Iowa shows higher mapping performance, with overall accuracy increasing slightly from 0.949 to 0.978, marginally above CDL's 0.976. In 2023, Nebraska's monthly ICDL accuracy improved from 0.906

in June to 0.984 in August, surpassing CDL accuracy of 0.976. The monthly improvement pattern—with a sharp increase between June and July and stabilization in August—is consistent across both years and states, indicating that classification quality is dependent on the availability of time-series spectral data as the growing season progresses. The discrepancy in mapping performance between Nebraska and Iowa is consistent with the observed differences in training label quality: in 2022, Nebraska's lower training label accuracy corresponds to reduced mapping performance relative to Iowa, and year-to-year and crop-type F1 score comparisons further confirm this relationship.

Table 3. 2022 Nebraska ICDL validation

State	Data	Month	Crop	User Accuracy	Producer Accuracy	Overall Accuracy	F1
Nebraska	2022 ICDL	June	corn	0.673	0.952	0.807	0.789
			soybean	0.861	0.709		0.778
			other	0.996	0.760		0.862
		July	corn	0.800	0.970	0.856	0.877
			soybean	0.811	0.799		0.805
			other	0.996	0.799		0.887
		August	corn	0.887	0.967	0.927	0.925
			soybean	0.909	0.925		0.917
			other	0.997	0.889		0.940
	2022 CDL	corn		0.949	0.997	0.916	0.972
		soybean		0.833	0.976		0.899
		other		0.996	0.775		0.872

Table 4. 2022 Iowa ICDL validation

State	Data	Month	Crop	User Accuracy	Producer Accuracy	Overall Accuracy	F1
Iowa	2022 ICDL	June	corn	0.952	0.961	0.949	0.956
			soybean	0.925	0.961		0.943
			other	0.972	0.925		0.948
		July	corn	0.988	0.979	0.972	0.983
			soybean	0.950	0.979		0.964
			other	0.979	0.958		0.968
		August	corn	0.985	0.982	0.978	0.983
			soybean	0.970	0.979		0.974
			other	0.979	0.973		0.976
	2022 CDL	corn		0.970	0.976	0.976	0.973
		soybean		0.991	0.976		0.983
		other		0.967	0.976		0.971

Table 5. 2023 Nebraska ICDL validation

State	Data	Month	Crop	User Accuracy	Producer Accuracy	Overall Accuracy	F1
Nebraska	2023 ICDL	June	corn	0.850	0.937	0.906	0.891
			soybean	0.881	0.841		0.861
			other	0.997	0.940		0.968
		July	corn	0.994	0.988	0.983	0.991
			soybean	0.962	0.994		0.978
			other	0.994	0.967		0.980

	August	corn	0.997	0.991	0.984	0.994
		soybean	0.960	0.997		0.978
		other	0.997	0.964		0.980
	2023 CDL	corn	0.979	0.988	0.976	0.983
		soybean	0.965	0.994		0.979
		other	0.984	0.946		0.965

Validation of major crop-type acreage estimates. Corn, soybean, and cotton pixels in specific states from August ICDL and annual CDL maps were converted to acreage, comparing with NRA. The acreage estimates derived exclusively from August ICDL maps, given their superior mapping accuracy.

2022 ICDL acreage estimates. As shown in Table 6, August 2022 ICDL corn acreage estimates closely match NRA across most states, with percentage differences within $\pm 5\%$ for the majority of states; only Kansas shows a larger deviation (-9.5%). The R^2 between ICDL and NRA is comparable to CDL, and total ICDL corn acreage differs from NRA by only -0.3% , compared to 1.3% for CDL. ICDL also achieves a lower MSE (-2) than CDL (7). Overall, ICDL estimates are highly consistent with NRA and demonstrate lower error, achieving near-perfect total corn acreage estimation across the Corn Belt in 2022.

Table 6. Corn planted acreage comparison among ICDL, CDL, and NRA at the state-level in 2022. Δ ICDL and Δ CDL are relative differences derived as a percentage compared to the NRA. The unit for crop acres is 10^4 acres.

State	NRA	ICDL	Δ ICDL	CDL	Δ CDL
Kentucky	144	148	3.0%	146	1.2%
Kansas	550	497	-9.5%	556	1.0%
Iowa	1290	1317	2.1%	1269	-1.6%
Illinois	1080	1115	3.2%	1059	-2.0%
Indiana	525	547	4.2%	525	0.1%
Michigan	235	237	0.7%	259	10.3%
Minnesota	800	738	-7.7%	822	2.7%
Nebraska	960	1016	5.8%	1004	4.6%
North Dakota	295	304	3.2%	292	-0.9%
Ohio	340	348	2.2%	338	-0.6%
South Dakota	575	539	-6.3%	597	3.9%
Missouri	335	313	-6.6%	327	-2.3%
Wisconsin	395	381	-3.5%	427	8.0%
Total	7524	7499	-0.3%	7621	1.3%
R^2		0.992		0.997	
MSE		-2		7	

For soybean (Table 7), most state-level ICDL estimates are within $\pm 5\%$ of NRA, with the exception of Kansas (-18.6%). CDL showcases more differences in Kentucky (-16.6%), Kansas (-12.2%), and Michigan (10.8%). The coefficient of determination (R^2) between ICDL and NRA is comparable to CDL. At the regional scale, ICDL-derived total soybean acreage differs from NRA by -1.2% , slightly larger than the -0.2% difference observed for CDL. Similarly, ICDL estimates showcase a higher MSE (-7) than CDL (-1). Meanwhile, soybean acreage estimates from 2022 ICDL exhibits slightly lower accuracy than corn.

Table 7. Soybean planted acreage comparison among ICDL, CDL, and NRA at the state-level in 2022.

State	NRA	ICDL	Δ ICDL	CDL	Δ CDL
Kentucky	195	203	4.3%	163	-16.6%

Kansas	505	411	-18.6%	443	-12.2%
Iowa	1010	1024	1.4%	1005	-0.5%
Illinois	1080	1069	-1.0%	1046	-3.2%
Indiana	585	623	6.5%	590	0.9%
Michigan	225	217	-3.8%	249	10.8%
Minnesota	745	746	0.1%	773	3.7%
Nebraska	575	532	-7.5%	581	1.0%
North Dakota	570	561	-1.6%	597	4.7%
Ohio	510	536	5.1%	523	2.5%
South Dakota	510	502	-1.7%	532	4.4%
Missouri	610	597	-2.1%	588	-3.6%
Wisconsin	216	229	6.2%	233	7.9%
Total	7336	7250	-1.2%	7323	-0.2%
R2		0.985		0.989	
MSE		-7		-1	

Across Texas and Georgia, ICDL cotton acreage estimates exhibit smaller and more balanced deviations than CDL, underestimating Texas by -10.7% and slightly overestimating Georgia by 4.5% , compared with CDL's larger overestimation (Table 8). Regionally, ICDL underestimates total acreage by -8.6% , with a lower MSE (-39) than CDL (70).

Table 8. Cotton planted acreage comparison among ICDL, CDL, and NRA at the state-level in 2022.

State	NRA	ICDL	Δ ICDL	CDL	Δ CDL
Texas	785	701	-10.7%	881	12.2%
Georgia	129	135	4.5%	172	33.5%
Total	914	836	-8.6%	1053	15.2%
MSE		-39		70	

2023 ICDL acreage estimates. Corn acreage estimates across 13 Corn Belt states show deviations within $\pm 4.4\%$ in 11 states, as shown in Table 9. Slightly higher deviation is observed in Minnesota (-9.4%) and Missouri (-7.4%). Both ICDL and CDL exhibit extremely high coefficient with NRA at the state level, indicating strong linear agreement. At the regional scale, ICDL provides a closer approximation to the NRA (-0.8%), while CDL shows a slight positive bias (1.1%). ICDL has a slight negative bias (-5), whereas CDL shows a larger one (7). These results suggest that 2023 ICDL, similar to 2022, provides more accurate and less biased corn acreage estimates across the Corn Belt.

Table 9. Corn planted acreage comparison among ICDL, CDL, and NRA at the state-level in 2023.

State	NRA	ICDL	Δ ICDL	CDL	Δ CDL
Kentucky	160	158	-1.5%	165	3.3%
Kansas	575	568	-1.2%	565	-1.7%
Iowa	1310	1323	1.0%	1281	-2.2%
Illinois	1120	1165	4.0%	1114	-0.5%
Indiana	545	563	3.3%	555	1.9%
Michigan	240	242	0.9%	270	12.4%
Minnesota	860	779	-9.4%	875	1.8%
Nebraska	995	1039	4.4%	1040	4.6%

North Dakota	405	392	-3.2%	393	-2.9%
Ohio	360	350	-2.8%	353	-1.8%
South Dakota	630	602	-4.4%	647	2.7%
Missouri	385	357	-7.4%	375	-2.5%
Wisconsin	400	388	-2.9%	436	9.1%
Total	7985	7925	-0.8%	8071	1.1%
R ²		0.993		0.996	
MSE		-5		7	

August ICDL soybean estimates closely match NRA acreages as well, as shown in Table 10, with 11 states showing deviations within $\pm 6.1\%$. Moderate deviation is observed in North Dakota (-15.1%), and slight overestimation occurs in Wisconsin (11.6%). CDL exhibits larger deviation in several states, including substantial underestimation in Kentucky (-23.8%) and Kansas (-18.5%) and overestimation in Michigan (20.1%). ICDL slightly outperforms CDL in capturing state-level variability, as indicated by the higher R², and demonstrates more consistent performance across states. Consistent with 2022, 2023 ICDL provides reliable soybean acreage estimates across the Corn Belt, despite a slightly higher regional MSE.

Table 10. Corn planted acreage comparison among ICDL, CDL, and NRA at the state-level in 2023

State	NRA	ICDL	Δ ICDL	CDL	Δ CDL
Kentucky	183	188	2.5%	139	-23.8%
Kansas	443	416	-6.1%	361	-18.5%
Iowa	995	1009	1.4%	985	-1.0%
Illinois	1035	995	-3.9%	970	-6.3%
Indiana	550	570	3.7%	553	0.6%
Michigan	204	204	0.2%	245	20.1%
Minnesota	735	733	-0.2%	765	4.1%
Nebraska	525	505	-3.8%	529	0.7%
North Dakota	620	526	-15.1%	677	9.1%
Ohio	475	504	6.0%	495	4.1%
South Dakota	510	493	-3.3%	535	4.9%
Missouri	560	554	-1.2%	523	-6.6%
Wisconsin	211	235	11.6%	224	6.0%
Total	7046	6932	-1.6%	7001	-0.6%
R ²		0.985		0.975	
MSE		-9		-3	

For the cotton acreage estimates in Texas and Georgia (Table 11), ICDL overestimates acreage, particularly in Texas (15.2%), while Georgia shows only minor overestimation (1.7%). CDL outputs much higher deviation in Georgia (21.9%). Regionally, ICDL's total acreage overestimation (13.0%) is slightly higher than CDL's (11.6%), with comparable MSE values, indicating similar overall error despite ICDL's larger deviation in Texas.

Table 11. Cotton planted acreage comparison among ICDL, CDL, and NRA at the state-level in 2023

State	NRA	ICDL	Δ ICDL	CDL	Δ CDL
Texas	612	705	15.2%	670	9.6%
Georgia	120	122	1.7%	146	21.9%
Total	732	827	13.0%	817	11.6%

MSE		47		42	
-----	--	----	--	----	--

Limitations. ICDL data exhibit inconsistencies in mapping accuracy across spatial, temporal, and crop-type dimensions, attributable to training label quality and input imagery quality. Training labels were extracted using fixed rotation patterns across CONUS, leading to uneven label quality and corresponding variability in mapping accuracy. This is evident in the validation results: in 2022, lower training label accuracy in Nebraska relative to Iowa corresponds directly to reduced mapping performance, and year-to-year variation in label accuracy in Nebraska is reflected in map quality. Corn training labels consistently achieve higher F1 scores than soybean in Nebraska, contributing to higher classification accuracy for corn. Similar effects likely occur for other crop types—the lower cotton acreage estimates are consistent with reduced training label quality for that crop. Additionally, variability in satellite imagery quality constrains classification performance in some states, such as Wisconsin and North Dakota, where modest deviations from NRA statistics are observed. These limitations are important for users to consider when applying the data, particularly in regions or for crop types where training label accuracy may be lower.

Data Availability

The ICDL 2022 and 2023 monthly datasets validated in this study are openly available in three Zenodo repositories^{38–40}: <https://zenodo.org/records/17456018>, <https://zenodo.org/records/17494692>, and <https://zenodo.org/records/17457566>. Each Zenodo repository contains data in GeoTIFF format, which can be visualized using ArcGIS or QGIS with *Colormap* or *Unique Values* rendering styles. The annual maps can also be accessed through the online system <http://cloud.csiss.gmu.edu/cropsmart>.

Code Availability

The scripts used to generate the 10m ICDL dataset are available in this GitHub repository: <https://github.com/huiliterry/AutomatedMapping>.

References

1. *An Econometric Analysis of U.S. Crop Yield and Cropland Acreage: Implications for the Impact of Climate Change*. (2010). doi:10.22004/ag.econ.61527.
2. USDA. NASS. USDA - National Agricultural Statistics Service- Research and Science - Cropland Data Layer Releases. https://www.nass.usda.gov/Research_and_Science/Cropland/Release/index.php (2025).
3. Boryan, C., Yang, Z., Mueller, R. & Craig, M. Monitoring US agriculture: the US Department of Agriculture, National Agricultural Statistics Service, Cropland Data Layer Program. *Geocarto International* **26**, 341–358 (2011).
4. CropScape. CropScape - NASS CDL Program. <https://nassgeodata.gmu.edu/CropScape/> (2025).
5. USDA.NASS. USDA - NASS- Research and Science - Cropland Data Layer Releases. https://www.nass.usda.gov/Research_and_Science/Cropland/Release/index.php (2025).

6. Li, H. *et al.* Automated In-season Crop-type Data Layer Mapping without Ground Truth for the Conterminous United States based on Multisource Satellite Imagery. *IEEE Transactions on Geoscience and Remote Sensing* 1–1 (2024) doi:10.1109/TGRS.2024.3361895.
7. Zhang, C. *et al.* Towards automation of in-season crop type mapping using spatiotemporal crop information and remote sensing data. *Agricultural Systems* **201**, 103462 (2022).
8. Li, H., Di, L., Zhang, C., Lin, L. & Guo, L. Improvement Of In-season Crop Mapping For Illinois Cropland Using Multiple Machine Learning Classifiers. in *2022 10th International Conference on Agro-geoinformatics (Agro-Geoinformatics)* 1–6 (2022). doi:10.1109/Agro-Geoinformatics55649.2022.9859153.
9. Lin, L. *et al.* Validation and refinement of cropland data layer using a spatial-temporal decision tree algorithm. *Sci Data* **9**, 63 (2022).
10. Liu, Z. *et al.* In-season crop yield prediction: State of the art and future research direction. *International Journal of Applied Earth Observation and Geoinformation* **146**, 105129 (2026).
11. Li, H. *et al.* In-Season Mapping of Sugarcane Planting Based on Sentinel-2 Imagery. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* **18**, 1410–1421 (2025).
12. Gallo, I., Ranghetti, L., Landro, N., La Grassa, R. & Boschetti, M. In-season and dynamic crop mapping using 3D convolution neural networks and sentinel-2 time series. *ISPRS Journal of Photogrammetry and Remote Sensing* **195**, 335–352 (2023).
13. Song, X.-P. *et al.* National-scale soybean mapping and area estimation in the United States using medium resolution satellite imagery and field survey. *Remote Sensing of Environment* **190**, 383–395 (2017).
14. Rußwurm, M. *et al.* End-to-end learned early classification of time series for in-season crop type mapping. *ISPRS Journal of Photogrammetry and Remote Sensing* **196**, 445–456 (2023).
15. Konduri, V. S., Kumar, J., Hargrove, W. W., Hoffman, F. M. & Ganguly, A. R. Mapping crops within the growing season across the United States. *Remote Sensing of Environment* **251**, 112048 (2020).
16. Tran, K. H., Zhang, H. K., McMaine, J. T., Zhang, X. & Luo, D. 10 m crop type mapping using Sentinel-2 reflectance and 30 m cropland data layer product. *International Journal of Applied Earth Observation and Geoinformation* **107**, 102692 (2022).

17. Johnson, D. M. & Mueller, R. Pre- and within-season crop type classification trained with archival land cover information. *Remote Sensing of Environment* **264**, 112576 (2021).
18. Rahman, M. S., Di, L., Yu, E., Zhang, C. & Mohiuddin, H. In-Season Major Crop-Type Identification for US Cropland from Landsat Images Using Crop-Rotation Pattern and Progressive Data Classification. *Agriculture* **9**, 17 (2019).
19. Zhang, C. *et al.* Rapid in-season mapping of corn and soybeans using machine-learned trusted pixels from Cropland Data Layer. *International Journal of Applied Earth Observation and Geoinformation* **102**, 102374 (2021).
20. Zhang, C., Di, L., Lin, L. & Guo, L. Extracting Trusted Pixels from Historical Cropland Data Layer Using Crop Rotation Patterns: A Case Study in Nebraska, USA. in *2019 8th International Conference on Agro-Geoinformatics (Agro-Geoinformatics)* 1–6 (2019). doi:10.1109/Agro-Geoinformatics.2019.8820236.
21. USDA.ERS. USDA ERS - Cotton & Wool. <https://www.ers.usda.gov/topics/crops/cotton-wool/> (2022).
22. USDA. FAS. Crop Calendar Charts. <https://ipad.fas.usda.gov/ogamaps/cropcalendar.aspx> (2025).
23. Copernicus Sentinel-2 (processed by ESA). MSI Level-1C TOA Reflectance Product. Collection 1. European Space Agency,. https://doi.org/10.5270/S2_-742ikth (2021).
24. Earth Resources Observation And Science (EROS) Center. Landsat 8-9 Operational Land Imager / Thermal Infrared Sensor Level-2, Collection 2 [dataset]. U.S. Geological Survey. <https://doi.org/10.5066/P9OGBGM6> (2020).
25. United States Department of Agriculture (USDA) National Agricultural Statistics Service (NASS). Cropland Data Layer: USDA NASS, USDA NASS Marketing and Information Services Office, Washington, D.C. <https://croplandcros.scinet.usda.gov/> (20240131).
26. U.S. Census Bureau. 2018 TIGER/Line Shapefiles. U.S. Department of Commerce. <https://www.census.gov/geographies/mapping-files/time-series/geo/tiger-line-file.html> (2018).
27. Gorelick, N. *et al.* Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sensing of Environment* **202**, 18–27 (2017).

28. USDA. NASS. Acreage | Economics, Statistics, and Market Information System. *Acreage* <https://esmis.nal.usda.gov/publication/acreage> (2025).
29. Rouault, E. *et al.* GDAL. Zenodo <https://doi.org/10.5281/zenodo.17978407> (2025).
30. Rouse, J. W., Haas, R. H., Deering, D. W., Schell, J. A. & Harlan, J. C. *Monitoring the Vernal Advancement and Retrogradation (Green Wave Effect) of Natural Vegetation*. <https://ntrs.nasa.gov/citations/19750020419> (1974).
31. McFEETERS, S. K. The use of the Normalized Difference Water Index (NDWI) in the delineation of open water features. *International Journal of Remote Sensing* **17**, 1425–1432 (1996).
32. Pflugmacher, D., Rabe, A., Peters, M. & Hostert, P. Mapping pan-European land cover using Landsat spectral-temporal metrics and the European LUCAS survey. *Remote Sensing of Environment* **221**, 583–595 (2019).
33. Sabaghy, S., Abuzar, M., Crawford, D., McAllister, A. & Sheffield, K. Remote sensing for land cover mapping across Victoria, Australia – a machine learning application. *Scientific Data* **12**, 566 (2025).
34. Hermosilla, T., Wulder, M. A., White, J. C. & Coops, N. C. Land cover classification in an era of big and open data: Optimizing localized implementation and training data selection to improve mapping outcomes. *Remote Sensing of Environment* **268**, 112780 (2022).
35. Li, H. *et al.* In-season sugarcane mapping in the U.S. and Brazil using time-invariant phenological features. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 1–20 (2026) doi:10.1109/JSTARS.2026.3657381.
36. Li, H. *et al.* In-season crop-type mapping in Kenya using Sentinel-2 imagery. in *2024 12th International Conference on Agro-Geoinformatics (Agro-Geoinformatics)* 1–6 (2024). doi:10.1109/Agro-Geoinformatics262780.2024.10660971.
37. USDA. NASS. USDA - National Agricultural Statistics Service - Publications - Reports by Title or Release Day. https://www.nass.usda.gov/Publications/Reports_by_Release_Day/index.php.
38. Li, H. & Di, L. Contiguous United States 10m-resolution In-season Crop-type Data Layer 2023 June-August. Zenodo <https://doi.org/10.5281/zenodo.17457566> (2025).

39. Li, H. & Di, L. Contiguous United States 10m-resolution In-season Crop-type Data Layer 2022 June. Zenodo <https://doi.org/10.5281/zenodo.17494692> (2025).
40. Li, H. & Di, L. Contiguous United States 10m-resolution In-season Crop-type Data Layers 2022 July-August. Zenodo <https://doi.org/10.5281/zenodo.17456018> (2025).
41. USDA.NASS. USDA - National Agricultural Statistics Service - Research and Science - Cropland Data Layers. *Cropland Data Layers - FAQs* https://www.nass.usda.gov/Research_and_Science/Cropland/sarsfaqs2.php.

Author Contributions

Hui Li and Liping Di designed the research. Hui Li implemented data production, data quality validation, and manuscript drafting. Liping Di secured the research funding, validation data collection, manuscript review. Chen Zhang provided manuscript review. Liying Guo and Eugene G. Yu provided advice on data analysis. Bosen Shao drafted partial literature review. Ziao Liu, and Hanxi Li contributed data processing.

Competing Interests

The authors declare no competing interests.

Funding

This study is supported by grants from NSF (grant# 2228000, 2236137, 2345039, PI: Dr. Liping Di) and USDA NIFA (grant # 2021-67021-34151, PI: Dr. Liping Di).