Computers and Electronics in Agriculture 213 (2023) 108199

Contents lists available at ScienceDirect

Computers and Electronics in Agriculture

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

Cyberinformatics tool for in-season crop-specific land cover monitoring: Design, implementation, and applications of iCrop

Chen Zhang [a], Liping Di [a,*], Li Lin [a], Haoteng Zhao [a], Hui Li [a], Anna Yang [a], Liying Guo [a], Zhengwei Yang [b]

[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

A R T I C L E I N F O

Keywords:
Cyberinformatics
GIS
FAIR data principle
Agro-geoinformation
Land use land cover
Crop monitoring

A B S T R A C T

Cyberinformatics tools have supported decision makings in agriculture through cutting-edge big data, artificial intelligence/machine learning (AI/ML), and high-performance computing technologies. An open and easy-to-use agricultural cyberinformatics tool based on the findable, accessible, interoperable, reusable (FAIR) data principle is essential for the efficient distribution of crop-specific land cover information. This paper introduces iCrop, a new cyberinformatics tool to enable in-season crop type monitoring for the Conterminous United States (CONUS). As a web-based geographic information system (GIS), iCrop not only delivers three sets of new ML-based field-level crop-specific land cover geospatial data, including pre-season crop cover maps, in-season crop cover maps, and Refined Cropland Data Layer (R-CDL), but also provides a suite of mapping and geo-processing functionalities through the FAIR geospatial data standards, such as Web Map Service (WMS), Web Coverage Service (WCS), and Web Processing Service (WPS). Meanwhile, we outline several use cases to highlight iCrop's applications under various agricultural operation scenarios, its functionality for land use change analysis, and its interoperability with generic web-based and desktop GIS software (e.g., GeoPlatform and QGIS). Our experimental results show that the new cyberinformatics tool can provide timely and unique crop-specific land cover information through the geoprocessing functionalities to facilitate U.S. agricultural information management and decision support. Moreover, this paper can be used as a systematic guidance for the design and implementation of the cyberinformatics tool to disseminate agro-geoinformation based on the FAIR data principle.

1. Introduction

Cyberinformatics is a critical component in modern computer software and geographic information systems (GIS), which leverages the capabilities of cyberinfrastructure, big data analytics, artificial intelligence (AI), machine learning (ML), Internet of Things (IoT) to facilitate the collection, management, analysis, visualization, and dissemination of geoinformation data over the high-speed network. Within the agricultural sector, cyberinformatics tools have demonstrated great potential to facilitate the process and analysis of agro-geoinformation, thereby supporting the management of spatio-temporal variability in crop production (Di and Üstündağ, 2021; White et al., 2021). In recent years, the U.S. Department of Agriculture (USDA) National Institute of Food and Agriculture (NIFA) has continuously invested in Agriculture and Food Research Initiative (AFRI) to investigate the food and agriculture cyberinformatics tools (USDA NIFA, 2020, 2021). The results and findings from these research initiatives, spanning various agricultural applications like evapotranspiration estimation (Armstrong et al., 2022), droplet detection for agricultural spraying systems (Acharya et al., 2022), food environment assessment (Mulrooney et al., 2021), soil carbon management (Sanderman et al., 2021), real-time decision making in plant phenotyping (Singh et al., 2021), crop yield forecast (Medina et al., 2021), have indicated that cyberinformatics can efficiently aid agricultural information management and decision support.

As the pioneer of applying cyberinformatics tools as the primary GIS software to serve agricultural research, USDA National Agricultural Statistics Service (NASS) has been at the forefront of this effort, consistently using web-based geospatial data service systems to access,

* Corresponding author.
E-mail addresses: czhang11@gmu.edu (C. Zhang), ldi@gmu.edu (L. Di), llin2@gmu.edu (L. Lin), hzhao22@gmu.edu (H. Zhao), hli47@gmu.edu (H. Li), ayang24@gmu.edu (A. Yang), lguo2@gmu.edu (L. Guo), zhengwei.yang@usda.gov (Z. Yang).

https://doi.org/10.1016/j.compag.2023.108199
Received 3 July 2023; Received in revised form 11 August 2023; Accepted 28 August 2023
Available online 12 September 2023
0168-1699/© 2023 Elsevier B.V. All rights reserved.

manage, process, analyze, and disseminate agro-geoinformation based on the Findable, Accessible, Interoperable, and Reusable (FAIR) data principle (Wilkinson et al., 2016; Barton et al., 2022; Top et al., 2022; Hu et al., 2023; Wolfert et al., 2023). For example, CropScape is a well-known agricultural geospatial data service system to serve Cropland Land Data (CDL) data products through web geoprocessing services (Han et al., 2012). VegScape is a similar web-based data service system to disseminate a variety of near-real-time vegetation index (VI) data products derived from Moderate Resolution Imaging Spectroradiometer (MODIS) data (Zhang et al., 2022b). Crop Condition and Soil Moisture Analytics (Crop-CASMA) is a new tool to monitor and analyze U.S. soil moisture conditions using remotely sensed soil moisture data derived from NASA's Soil Moisture Active Passive (SMAP) mission (Yang et al., 2021; Zhang et al., 2022c). In 2021, USDA NASS announced CroplandCROS as an alternative web application to host CDL data with a new user interface (USDA NASS, 2021). These cyberinformatics tools have helped users worldwide to better understand U.S. agriculture and significantly facilitated decision makings by providing diverse geoinformation through the easy-to-use web interface.

Among the existing geospatial data in agro-geoinformation, USDA NASS's CDL is by far the most downloaded crop-specific land cover data product. As of Dec 2022, the CDL and its derivative data products (e.g., crop frequency layers, crop mask layer) have been distributed to over 330,000 users worldwide through CropScape. However, the current-year CDL data is usually released to the public until early next year (e.g., the 2022 CDL was released in February 2023), which means CropScape cannot be applied to provide timely agricultural land cover information within the growing season. To tackle this issue and offset the lag of CDL, many AI/ML approaches and algorithms for early- and in-season crop type classification are implemented (Hao et al., 2015; Yaramasu et al., 2020; Johnson and Mueller, 2021; Zhang et al., 2021; Abernethy et al., 2023). Therefore, developing a CropScape-like cyberinformatics tool that provides user-experience-centric sharing and analysis of these in-season crop-specific land cover classification results based on the FAIR data principle is critical for the timely agricultural decision support and distribution of agro-geoinformation.

In this paper, we introduce a new cyberinformatics tool, iCrop, for monitoring crop-specific land cover in support of the U.S agriculture. The main objective of the proposed software are: (1) being the first cyberinformatics tool to provide pre-, and in-season crop-specific land cover data for the Conterminous United States (CONUS); (2) offering a suite of easy-to-use functionalities, such as zonal statistics and land use land cover (LULC) change detection, to deliver timely and abundant field-level crop type information before and within the growing season; and (3) implementing the latest open geospatial standards and facilitating the agricultural decision making with the unique geoprocessing capabilities and national-scale field-level products based on the FAIR data principle. Furthermore, this paper serves not only as an introduction to the design and implementation of the new cyberinformatics tool, but also as a user guide for the iCrop application and its associated web services. In the rest of the paper, Section 2 presents the general context, system design, data production methods, and the implementation of map and geoprocessing services of the proposed tool. Section 3 gives a group of use cases to demonstrate iCrop's geoprocessing capability, LULC analysis functionality, and interoperability with other generic GIS software. Section 4 discusses the contribution of this study, the advantages of the system architecture, the limitations of the current implementation, potential improvements, and future research recommendations. The conclusions are summarized in Section 5.

2. Methods

2.1. General context

The cyberinformatics tool is a core component in an Earth observation system, which serves as a crucial link between the collection and analysis of geoinformation data and the users who need this information to make informed decisions (Lin et al., 2019; Bayat et al., 2021; Graf et al., 2022; Tan et al., 2022). In the context of crop monitoring system as illustrated in Fig. 1, it is essential for users to have the ability to process, manage, and deliver geospatial data in a timely manner to provide valuable agro-geoinformation. An efficient cyberinformatics tool is therefore critical to unlocking the full potential of agro-geoinformation data acquired by various sensors and enhancing decision-making processes in agriculture. To make the cycle more effective and enables timely access to early- and in-season crop-specific land cover information, the cyberinformatics tool must be capable of serving up-to-date geospatial data to users and providing geoprocessing functionalities using GIS over high-speed internet.

The implementation of a cyberinformatics tool for pre- and in-season crop-specific land cover data sharing and analysis requires tackling three essential questions: (1) how to leverage AI/ML techniques to facilitate national-scale pre- and in-season crop mapping; (2) how to facilitate timely decision makings in agriculture through web geoprocessing; and (3) how to make the new geospatial data products FAIR? This study will investigate the above questions by demonstrating the system design (Section 2.2), data production (Section 2.3), and implementation (Sections 2.4 and 2.5) of the new cyberinformatics tool.

2.2. System design

The new cyberinformatics tool introduced in this paper is called iCrop web-based data service system (hereafter called "iCrop" or "iCrop system" for simplicity). It inherits the service-oriented architecture of CropScape and comes with three significant improvements. First, we develop, validate, and mature the ML algorithms to derive a collection of new crop cover data products for the Conterminous United States (CONUS) automatically. Second, the new data products are implemented and distributed through standardized web-based geoprocessing and map services. Third, the entire system is migrated to a cloud environment to enhance user experience and support.

As shown in Fig. 2, the iCrop system architecture comprises four main components: a data production module, a web service system module, front-end clients, and the cloud infrastructure. The data production module handles the generation of all ML-based crop-specific land cover geospatial data in the system. In this system, each data product is produced on the geospatial cloud computing platforms and GIS software, such as Google Earth Engine (GEE) and ArcGIS. The web service system module is the core of the iCrop system, where all geospatial data are processed into interoperable maps and disseminated through standardized OGC web geoprocessing and map service interfaces. These web services are fully interoperable with the iCrop web application as well as other OGC-compliant GIS software/libraries and third-party applications. Meanwhile, the entire iCrop system, including all data, web services, and the web client, is hosted as Platform-as-a-Service (PaaS) on the GeoBrain Cloud (https://cloud.csiss.gmu.edu/) to guarantee the scalability and timeliness of each geoprocessing request. The details of the implementation of the data production module, web service system, and the iCrop client will be described in the following sections.

2.3. Data production

Table 1 summarizes the spatial and temporal information of the new crop-specific land cover data available on the iCrop system. Consistent with the official CDL, the three new geospatial data products, including pre-season crop cover, in-season crop cover, and Refined Cropland Data Layer (R-CDL), cover the entire CONUS at 30 m resolution. Each data product is produced using the specific ML model and released at the different crop growth stages.

The pre-season crop cover map is the prediction of U.S. crop cover using an ML-based crop sequence prediction model proposed by Zhang

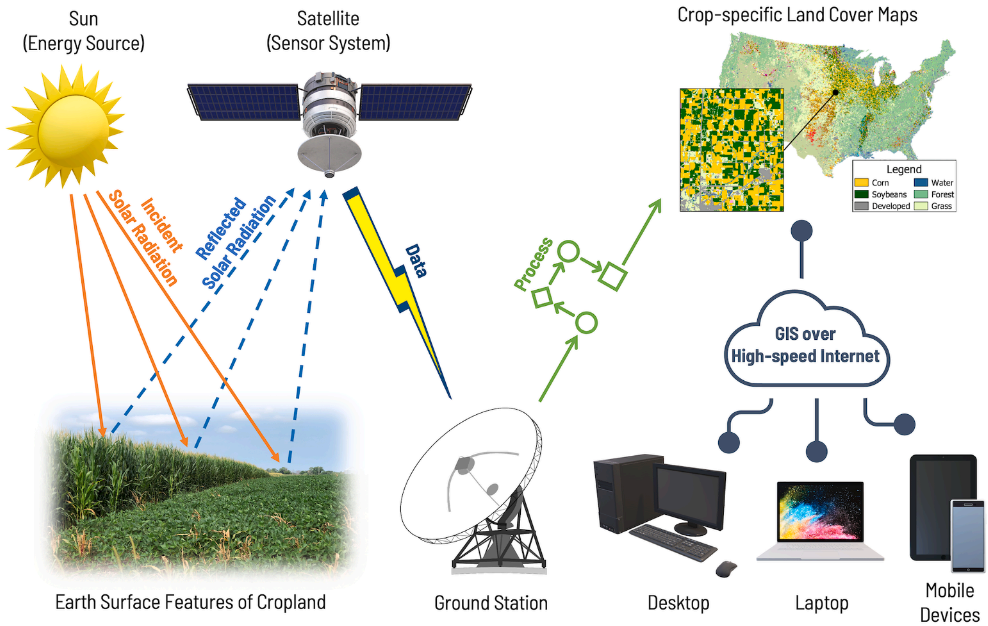


Fig. 1. The cycle of collection, processing, management, and dissemination of agro-geoinformation from Earth observation data in the crop monitoring system.

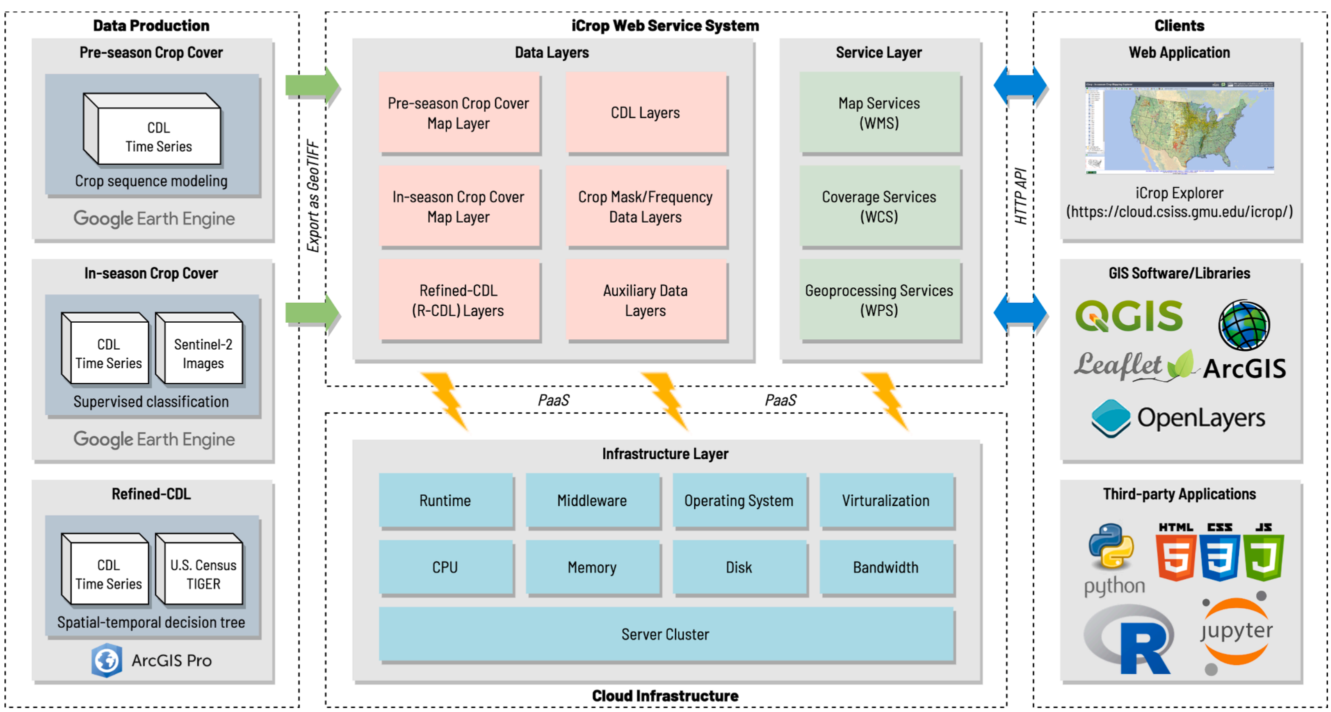


Fig. 2. Design of the iCrop system architecture.

Table 1
Summary of land cover data available on the iCrop system.

Data	Temporal Coverage	Spatial Coverage	Spatial Resolution	Release Date	ML Model
Pre-season crop cover	2022 – Current year (single annual map)	CONUS	30 *m*	Before growing season starts *	Artificial neural network (Zhang et al., 2019)
In-season crop cover	2022 – Current year (annual maps by May, June, and July)	CONUS	30 *m*	Within growing season (updated monthly between June and August)	Random forest (Zhang et al., 2022a)
R-CDL	2017 – Previous year (single annual map)	CONUS	30 *m*	After growing season ends *	Decision tree (Lin et al., 2022)

* Usually available in February after the new CDL data is released

et al. (2019). The model uses the artificial neural network to automatically learn crop rotation patterns from the 10-year CDL time series. Then the spatial distribution of pre-season crop cover will be predicted based on the historical crop sequence in CDL with the trained ML model. The pre-season data will be released before the growing season starts. For example, the production of the 2023 pre-season crop cover map is based on the 2013–2022 CDL, which began after the release of the 2022 CDL on February 1st and finished in late February 2023.

The in-season crop cover map is a remote-sensing-based crop-specific land cover data product derived from CDL and satellite images. The production of in-season crop mapping adopts a novel mapping-without-ground-truth workflow developed by Zhang et al. (2022a). The main advantage of this method is using the trusted pixels, the high-confidence pixels that are predicted from CDL time series using regular crop rotation patterns (e.g., monocropping, alternate cropping), to replace ground truth and label training samples in satellite images for crop type classification, which can significantly reduce the ground-truthing process and save substantial resource needs and labor costs. Implemented on the GEE platform, the mapping-without-ground-truth workflow employs the random forest classifier, which is widely used in remote-sensing-based crop mapping (Wang et al., 2019; Hao et al., 2020; Pierre Pott et al., 2022; Rufin et al., 2022; Soltanikazemi et al., 2022). As an in-season data product, the crop cover map by May/June/July will be updated monthly within the growing season. For example, the 2023 in-season crop cover map by May will be released in early June of 2023. With the progress of the growing season, the classification accuracy will be improved because of more abundant spectral features captured in the satellite images.

The R-CDL, on the other hand, is the refinement data product of CDL. The official CDL contains a certain number of misclassified pixels, mixed pixels, and noisy pixels. These errors may lead to biased results in the following studies and analyses. To further refine the CDL data, Lin et al. (2022) proposed a decision tree method to find and validate anomalous pixels. This method first uses a decision tree to identify the potentially misclassified pixels in the single crop cover map, then adjust these pixels using spatial–temporal information from the long-term CDL time series. Same as the official CDL, the R-CDL is an annual data product. The R-CDL is produced using ArcGIS and currently available from 2017 to 2022. Once the new CDL is released, the corresponding R-CDL will be immediately generated.

2.4. Implementation of geospatial web services

Similar to many cyberinformatics tools based on the FAIR data principle (Zhang et al., 2020; Wu et al., 2021; Erazo Ramirez et al., 2022; Redhead et al., 2022; Menegon et al., 2023), iCrop's geospatial data is disseminated through OGC web service standard interfaces. When implementing the service layer of iCrop, we used MapServer (McKenna and MapServer PSC, 2021) as the back-end server to power the web mapping capabilities via standardized OGC specifications, such as WMS (OGC, 2006) and WCS (OGC, 2012). We also used ArcGIS as an alternative back-end WMS/WCS server to host the new geospatial data products described in this study through ArcGIS REST APIs. These standard interfaces facilitated interoperability with the front-end client as well as other WMS/WCS-compliant GIS software and applications.

Specifically, WMS standardized the operations for requesting the dynamic map of a geospatial image via the HTTP protocol. Table 2 summarizes the operations and request examples of iCrop WMS. The *GetCapabilities* operation enables users to request the metadata of all crop cover map layers available on iCrop. The *DescribeLayer* operation allows users to request the description of a specific map layer. The *GetMap* operation request image of a specific map layer based on mapping parameters such as projection, bounding box, image format, and width/height of the map. In addition, the legend graphics of a specific map layer can be requested via the *GetLegendGraphic* operation.

Compared to the map image provided by WMS, WCS returns coverage data files with metadata and supports a wider range of formats that are typically required for further modeling and analysis. Table 3 summarizes the WCS operations and provides examples of requests used in iCrop. The functions and usage of WCS operations are similar to those of WMS. The *GetCapabilities* operation requests metadata for all available coverage layers on iCrop. The *DescribeCoverage* operation requests a description of a specific coverage layer. The *GetCoverage* operation requests the raster source of a specific crop cover layer based on coverage parameters, such as projection, bounding box, scale, and output data format.

In order to cater to the users' geospatial data analysis requirements, the iCrop system incorporates various widely used geoprocessing functionalities. Fig. 3 shows the diagram of the geoprocessing workflow with the iCrop system. For instance, users can initiate an on-demand data download process by selecting the target data layer by year and the area of interest (AOI) by region, state, Agricultural Statistics Districts (ASD), county, or user-defined area. Besides, the iCrop system integrates a variety of common geoprocessing capabilities, such as zonal statistics, map generation, and change analysis. The zonal statistics function calculates the pixel value distribution over a specific region. The map production processes create the map of specific data products with legend and boundary information in PDF format.

2.5. Graphical user interface of web client

We implemented a web-based application, In-season Crop Mapping Explorer, as the front-end client for the iCrop system. Fig. 4 shows the graphical user interface (GUI) of the web client. The layout of this client inherits the well-known CropScape and VegScape application, which consists of a main interactive map explorer along with a menu bar and a sidebar. The main interactive map explorer displays geospatial data layers and provides basic map visualization functionalities. The main

Table 2
WMS operations and request examples of the iCrop system.

Operation	Description	Request Example
GetCapabilities	Get metadata of all available map layers	Metadata of iCrop WMS: https://cloud.csiss.gmu.edu/icrop-service/wms_icrop.cgi?SERVICE=WMS&VERSION=1.1.1&REQUEST=GetCapabilities
DescribeLayer	Describe information about the requested map layer	Description of in-season crop cover map for July 2023: https://cloud.csiss.gmu.edu/icrop-service/wms_icrop.cgi?SERVICE=WMS&VERSION=1.1.1&REQUEST=DescribeLayer&LAYERS=cdl_2023_inseason_july
GetMap	Get image for a map layer of a geospatial data layer with specific parameters (e.g., layer name, projection, bounding box, and width/height of map)	Image of in-season crop cover map for July 2023: https://cloud.csiss.gmu.edu/icrop-service/wms_icrop.cgi?SERVICE=WMS&VERSION=1.1.1&REQUEST=GetMap&LAYERS=cdl_2023_inseason_july&TRANSPARENT=true&SRS=EPSG:5070&BBOX=-2354935.721,311822.402,2256319.225,3165592.366&FORMAT=image/png&WIDTH=600&HEIGHT=400
GetLegendGraphic	Get legend graphics image of a specific map layer	Legend image of in-season crop cover map: https://cloud.csiss.gmu.edu/icrop-service/wms_icrop.cgi?VERSION=1.1.1&SERVICE=WMS&REQUEST=getlegendgraphic&LAYER=cdl_2023_inseason_july&FORMAT=image/png

Table 3
WCS operations and request examples of the iCrop system.

Operation	Description	Request Example
GetCapabilities	Get metadata about available coverage layers	Metadata of iCrop WCS: https://cloud.csiss.gmu.edu/icrop-service/wms_icrop.cgi?SERVICE=wcs&VERSION=2.0.1&REQUEST=GetCapabilities
DescribeCoverage	Describe the information about the requested coverage	Coverage description of in-season crop cover map for July 2023: https://cloud.csiss.gmu.edu/icrop-service/wms_icrop.cgi?SERVICE=wcs&VERSION=2.0.1&REQUEST=DescribeCoverage&COVERAGEID=cdl_2023_inseason_july
GetCoverage	Get raster source of a geospatial data layer with specific parameters (e.g., coverage layer name, projection, bounding box, scale, output data format)	Raster data of in-season crop cover map for July 2023: https://cloud.csiss.gmu.edu/icrop-service/wms_icrop.cgi?SERVICE=wcs&VERSION=2.0.1&REQUEST=GetCoverage&COVERAGEID=cdl_2023_inseason_july&FORMAT=image/tiff&SUBSET=x(310785,330045)&SUBSET=y(2145855,2165115)&SUBSETTINGCRS=http://www.opengis.net/def/crs/EPSG/0/5070

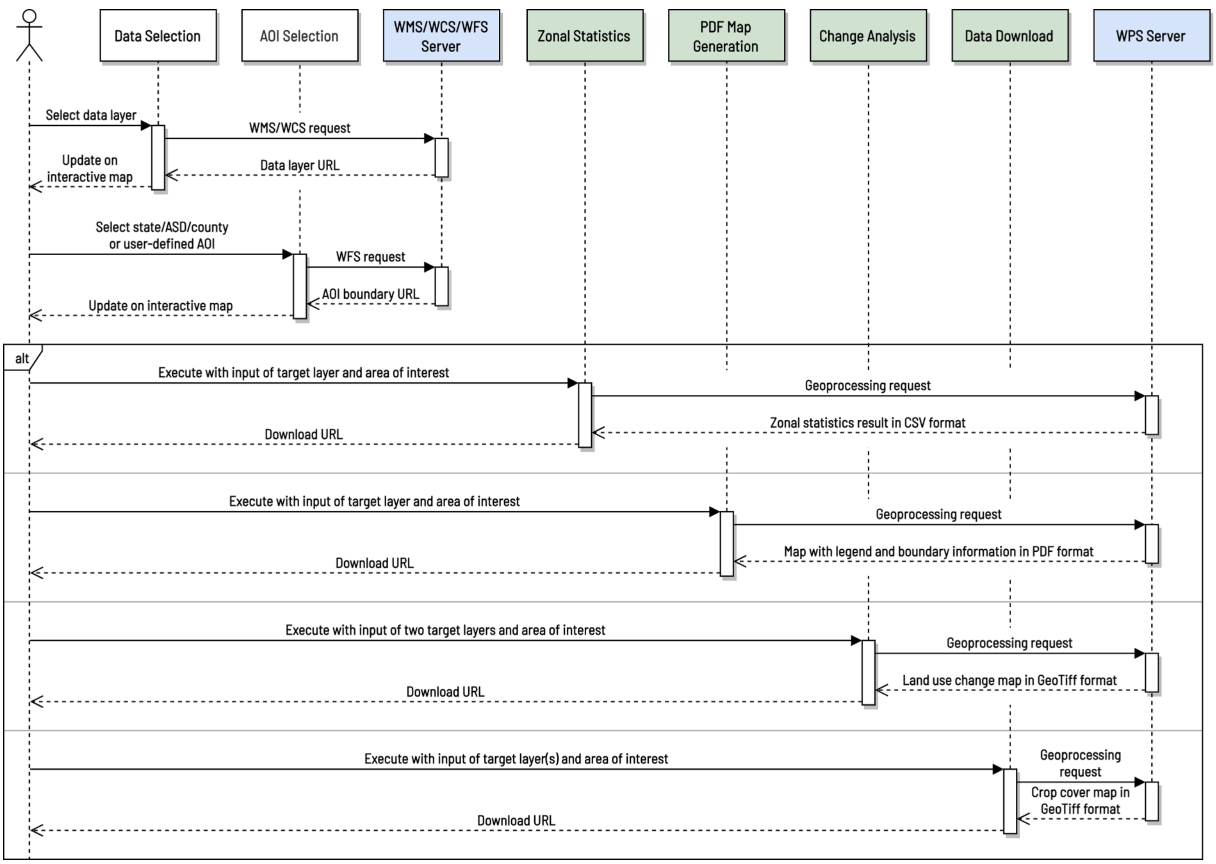


Fig. 3. Diagram of geoprocessing workflow with the iCrop system. The "alt" fragments represent the alternative scenarios of the geoprocessing workflow.

menu contains a toolbox bar with all available geoprocessing functions and links for user guides and documentation. The sidebar has a layers tab, a legend tab, and an overview mini-map. The layers tab lists all types of geospatial data, including all crop cover data products, background layer, crop mask layer, crop frequency layer, and auxiliary layers (i.e., boundaries, water, road). The legend tab shows the color table for crop type categories in the map. The overview mini-map displays the current extent of the explorer over the CONUS.

3. Use cases and results

3.1. Geoprocessing capabilities

The iCrop system presented in this study was initially developed to offer early-season decision support and applications through a suite of user-experience-centric geoprocessing functionalities. CropScape is being widely applied to produce maps and statistics data in white papers, business plans, technical reports, and other documents for multiple operational uses, which has successfully demonstrated the efficiency of cyberinformatics tools in agriculture. With the addition of three new field-level crop cover data products covering the CONUS, iCrop has extended the geoprocessing capabilities of CropScape, enabling more time-sensitive agricultural applications, and enhancing decision-making processes.

Fig. 5 showcases a group of examples of how common iCrop geoprocessing capabilities can be applied for operational uses. As demonstrated in this example, an AOI needs to be defined by region, state, ASD, and county before any processes. This example illustrates the file download, zonal statistics, and PDF map generation capabilities based on the 2022 in-season crop cover data for Pemiscot County, Missouri, of which corn, cotton, rice, soybeans, and winter wheat are the dominant crop types. The file download function will create a direct download link for the data layer of the selected years in GeoTIFF or KML format. The zonal statistics function counts the pixel and calculates the acreage for

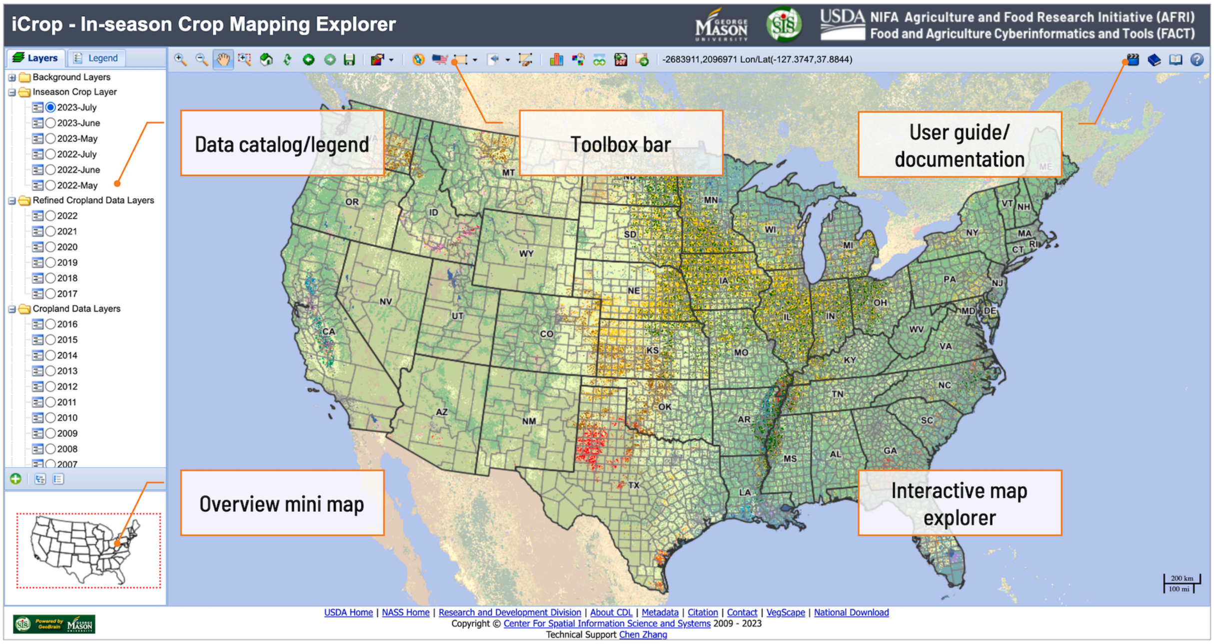


Fig. 4. GUI of the web-based client for iCrop (https://cloud.csiss.gmu.edu/icrop).

each crop type or land cover category within the AOI. Then the statistics results can be displayed as pie/column charts in the web application, or downloaded the raw data in CSV format. The PDF map function generates a print-ready map in PDF format with customizable titles, paper sizes, boundary layers, and land cover categories.

3.2. LULC change analysis

This use case demonstrates the change analysis capability of iCrop for crop rotation detection. The change analysis function of iCrop compares the LULC change between two years and generates the map for the changing area of the specific crop rotation pattern. Fig. 6 shows an example of detecting corn-soybeans rotation for Kossuth County of Iowa between 2022 and 2021 using the change analysis function of iCrop.

Located in the central U.S. Corn Belt, corn and soybeans are two dominant crop types in Kossuth County as shown in Fig. 6 (a). In 2021, Kossuth County led all Iowa counties in total soybean production with 13.4 million bushels and corn for grain production with 57.7 million bushels produced (USDA NASS, 2022a, 2022b). It is well-known that crop rotation is a common practice in U.S. agriculture. The spatial distribution of corn-soybeans and soybeans-corn rotation over the study area is illustrated in Fig. 6 (b). From the change detection results between 2021 and 2022, we can observe that 195,396.6 of total 280,996.5 acres of corn cropland followed soybeans-corn rotation, as well as 197,184.7 of total 219,710.8 acres of soybeans cropland followed corn-soybeans rotation.

3.3. Interoperability with GIS software

The map and coverage services of iCrop are fully interoperable with third-party WMS/WCS-compliant clients and applications. Thus, iCrop can be coupled to generic GIS software with OGC standards support, such as QGIS and ArcGIS. Fig. 7 demonstrates how a WMS layer of the 2023 in-season crop cover was stacked with the 2022 crop mask layer, county boundaries, and ASD boundaries in QGIS. In this example, the crop cover map is requested through the map layer "cdl_2023_inseason_july" of iCrop WMS.

We also tested the interoperability of iCrop on the GeoPlatform (https://www.geoplatform.gov/). Developed by Federal Geographic Data Committee (FGDC), the GeoPlatform program provides a portfolio of FAIR geospatial data, services, and applications under open licenses. It is mainly oriented to federal agencies, governments, private sectors, academia, and the general public. For example, the FAIR capabilities of the GeoPlatform have supported many application scenarios in the OGC Disaster Pilot programs, such as typhoon monitoring (Hu et al., 2022), flood detection (Lin et al., 2021), and climate services (Lieberman et al., 2022). Therefore, we choose GeoPlatform as another WMS/WCS-compliant platform to test the interoperability of the iCrop web services. Fig. 8 shows an example of integrating the iCrop WMS into GeoPlatform. The online prototype can be accessed via the GeoPlatform ArcGIS Online at https://geoplatform.maps.arcgis.com/apps/mapviewer/index.html?webmap=84a5aedb51e944ee924348f529a49cb5.

4. Discussion

4.1. Cyberinformatics tool based on FAIR data principle

The iCrop system significantly contributes to the advancement of cyberinformatics in U.S. agriculture through its provision of a specialized tool that combines FAIR data capabilities, flexible GIS interoperability, and user-friendly functionalities. As the first geospatial data service system offering pre- and in-season crop-specific land cover data for the CONUS, iCrop provides a comprehensive suite of easy-to-use functionalities, including zonal statistics and LULC change detection. These features enable the timely and abundant availability of field-level crop type information before and during the growing season.

In the realm of geospatial analysis, the rapid development of geospatial cloud computing platforms like GEE and Microsoft's Planetary Computer has substantially enhanced the capabilities and flexibility of geospatial analysis. However, the iCrop system distinguishes itself by emphasizing FAIR data principles. It focuses on the distribution of crop-specific land cover data through open geospatial data interfaces (e.g., WMS, WCS, WPS), which have been widely adopted by web-based GIS software and applications in agriculture (Nash et al., 2009). In contrast, commercial cloud platforms like GEE, while offering extensive functionalities and rich datasets, often lacks inherent adherence to FAIR data principles and restricts data usage to their own platform. By enabling easy interoperability with widely used GIS software like ArcGIS and QGIS, iCrop's data and geoprocessing services offer a more inclusive approach. Our data and geoprocessing services, on the other hand, can

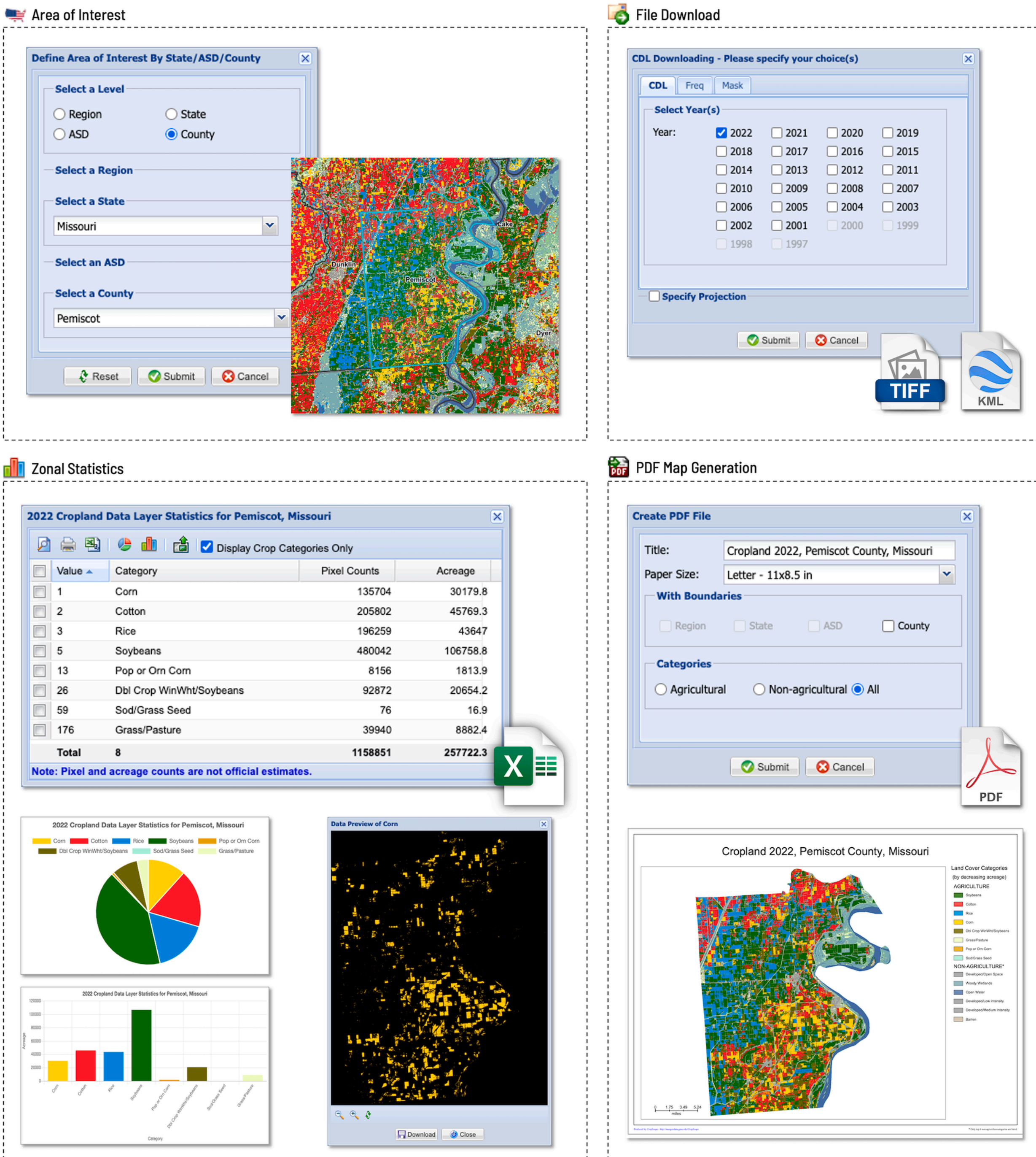


Fig. 5. Examples of the iCrop geoprocessing capabilities. The statistics and mapping results are from the 2022 crop cover for Pemiscot, Missouri.

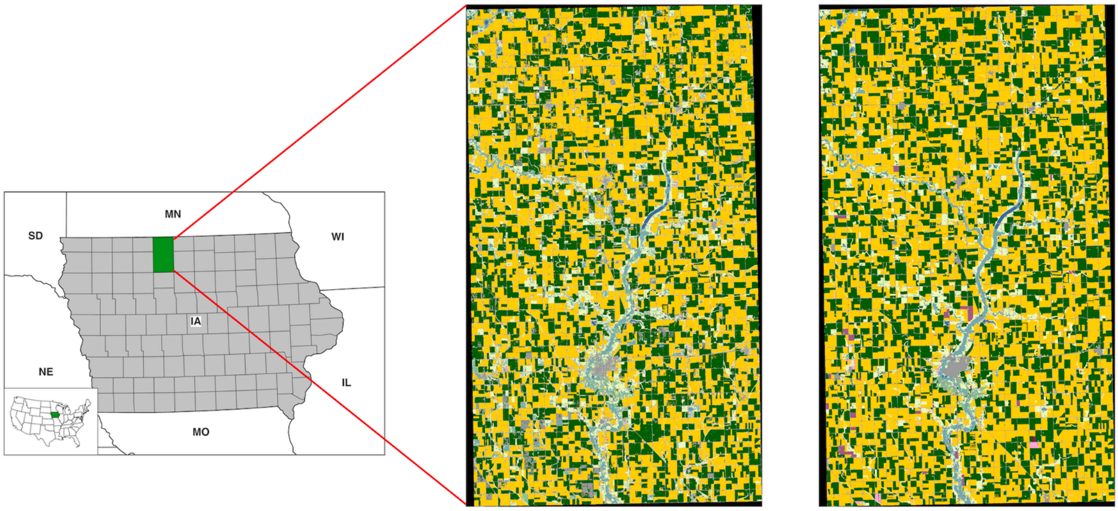
be easily interoperable with common GIS software, such as ArcGIS and QGIS.

It is important to note that the iCrop system is not designed to compete with geospatial cloud computing platforms. Instead, it serves as a complementing tool that addresses a distinct set of needs. While geospatial cloud computing platforms excel in their functionalities, iCrop focuses on providing specialized FAIR data capabilities and interoperability with common GIS software. This ensures that the iCrop system can be seamlessly integrated into existing geospatial workflows and extends its usability to a broader user base.

Another pivotal advantage of iCrop lies in its user-friendliness. While geospatial cloud computing platforms such as GEE may require coding skills and entail a steep learning curve, iCrop strives to offer a comprehensive suite of on-demand geoprocessing functions, such as zonal statistics and LULC change analysis, through a user-friendly interface. This approach empowers not only software engineers but also end-users and stakeholders like farmers and government agencies, who may possess limited technical or programming skills, to derive maximum benefits from the system. With its ability to deliver timely crop-specific land cover data and functionalities to diverse user groups, iCrop demonstrates great potential to enhance U.S. agricultural information management and decision support.

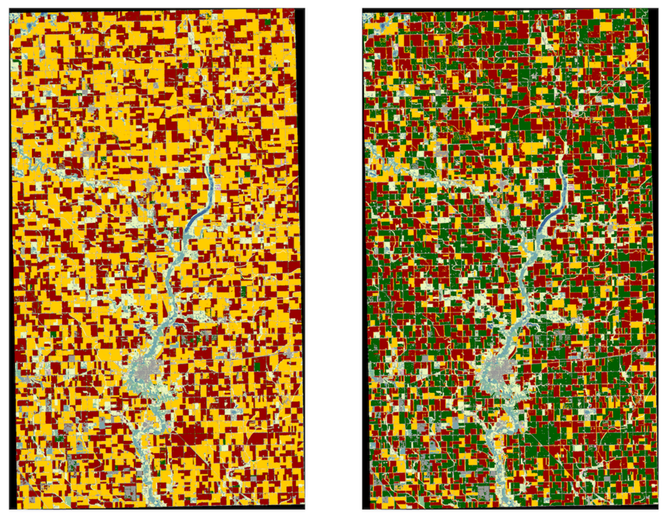
4.2. *Advantages of new crop-specific land cover data*

This study has demonstrated how ML-based crop cover data can be



(a) Comparison of in-season crop cover map by July 2022 (left) and R-CDL 2021 (right).

Crop Type	Pixels	Acreage
Corn (soybeans-corn rotation)	878603	195396.6
Corn (total)	1263504	280996.5
Soybeans (corn-soybeans rotation)	886643	197184.7
Soybeans (total)	987932	219710.8
Total cropland	2493254	554486.3



(b) Spatial distribution of corn-soybeans rotation (left) and soybeans-corn rotation (right) in 2022.

Fig. 6. Example of corn-soybeans rotation for Kossuth County of Iowa in 2022 using the change analysis function of iCrop. The yellow and green pixels represent corn and soybeans. The red pixels in the crop rotation maps represent the changing area.

easily shared and analyzed through a cyberinformatics tool that adheres to the FAIR data principles. When compared to CDL data, the ML-based crop cover data products on iCrop offer several advantages. First, the artificial neural network can automatically and reliably recognize crop sequence patterns from the historical CDL time series. With the pre-season crop mapping, the field-level spatial distribution and acreage of crop cover can be estimated before the growing season.

Second, the remote-sensing-based crop cover map for CONUS is significantly accelerated using the new mapping-without-ground-truth approach. By combining spectral information in satellite images with spatiotemporal information from historical CDL, the CDL-like in-season crop cover data for CONUS is available on iCrop as early as May. This is as much as eight months ahead of the official release of CDL data.

In contrast, the R-CDL method utilizes a decision tree approach to refine CDL data using spatial and temporal crop information. Validation results reveal enhanced classification accuracy, and a strong correlation between R-CDL and NASS crop acreage estimations at both county and state levels. This indirect validation highlights the effectiveness and efficiency of AI/ML techniques in further improving the accuracy of CDL data.

4.3. Potential improvements on the current tool

While the ML model effectively generates pre-season crop cover maps by learning crop sequence features in the historical CDL time series, predicting crop types for croplands that deviate from established patterns remains a challenge due to dynamic and uncertain factors. To address this, we plan to incorporate additional features, such as agricultural commodity prices and weather information, into the prediction model's training. Recent studies have demonstrated the potential of combining multisource geoinformation data and model to enhance crop and LULC modeling (Abunnasr and Mhawej, 2022; Baum et al., 2023; Leo et al., 2023; Massigoge et al., 2023; Meki et al., 2023).

To further enhance the accuracy of the current crop mapping results, the impact of other agro-geoinformation data and geospatial land cover data products, such as National Land Cover Database (NLCD) (Dewitz, 2021) and Land Change Monitoring, Assessment, and Projection (LCMAP) (Pengra et al., 2021), will be investigated and integrated into the ML models of the data production module. Additionally, more geoprocessing services, such as advanced raster calculation, time series analysis, and confidence layer comparison, will be integrated with the iCrop system in the next phase of development.

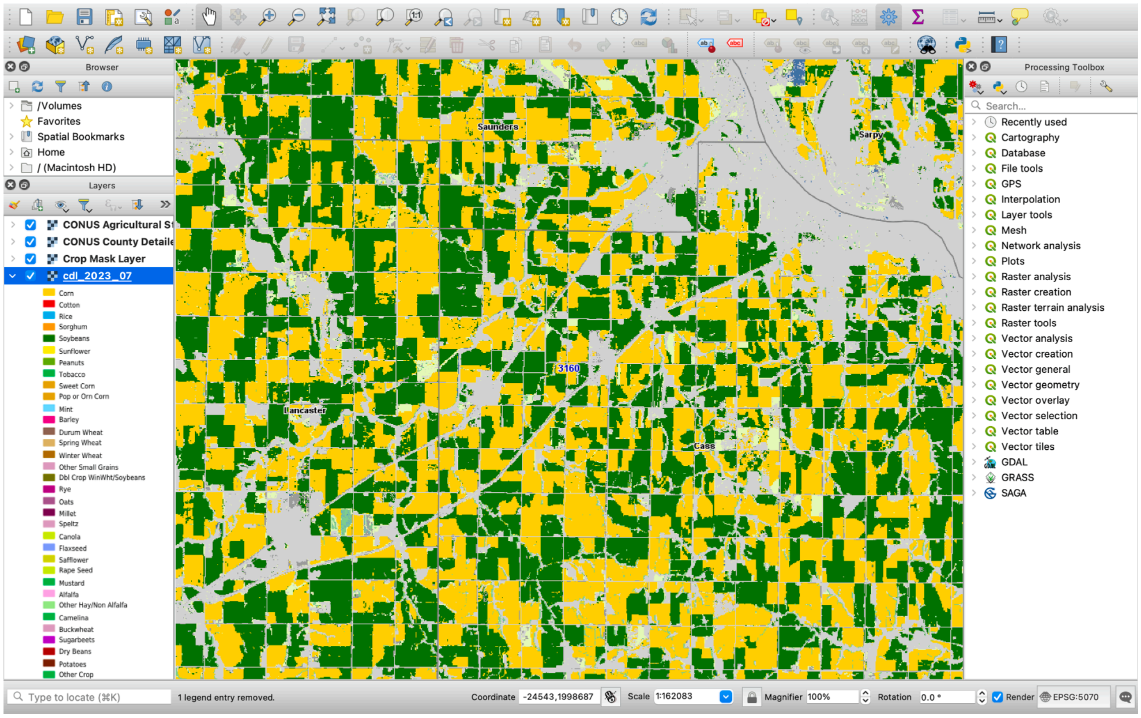


Fig. 7. Exploring crop cover via the iCrop WMS using QGIS. This example shows the 2023 in-season crop cover map with the 2022 crop mask layer and county/ASD boundaries.

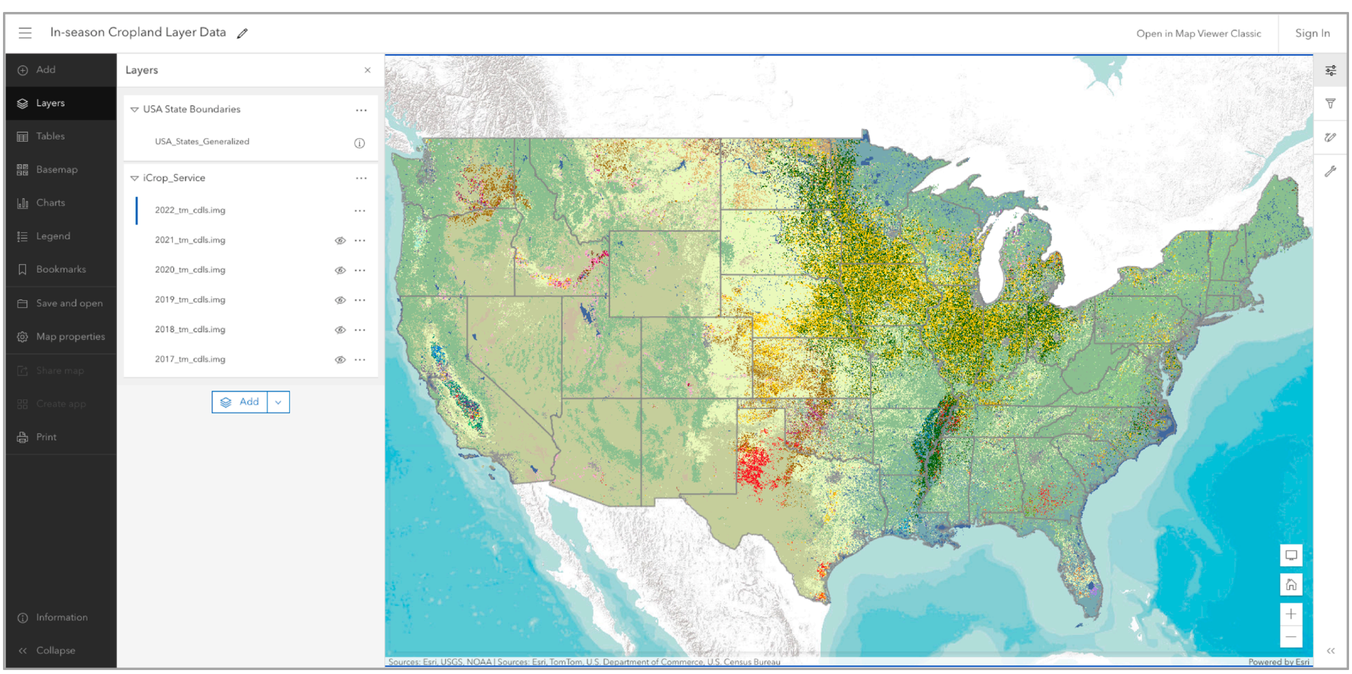


Fig. 8. Interoperability of the iCrop WMS with FGDC's GeoPlatform based on ArcGIS Online.

4.4. Future research recommendation

The subsequent developments could be conducted based on the new cyberinformatics tool proposed in this study. For example, Yue et al. (2022) recently proposed a FAIR training data model specifically serving Earth observation AI/ML tasks. It is a promising GIS application to incorporate the training data model into iCrop's geoprocessing services via the FAIR interface, which can potentially automate the crop-specific training sample selection and labeling process for the follow-up AI/ML model development in Earth observation research. Tekinerdogan and Verdouw (2020) and Verdouw et al. (2021) depict the conceptual frameworks for designing and implementing digital twins in agriculture. The iCrop's data and web services have great potential to be integrated into such digital twin frameworks to advance smart agriculture. The another potential future research recommendation is to explore the possibility of integrating the in-season land cover data of iCrop with the geoscience model workflow through OGC standard interfaces to enable the capabilities of crop damage assessment.

Many research topics in geoinformation science can be significantly facilitated using the new geospatial data of iCrop. Due to the insufficiency of ground truth data during the early growing season, iCrop's national-scale crop cover data products can be used as reference data sets in various agro-geoinformation studies. For example, the national-scale crop yield can be progressively estimated using the pre- and in-season data within the growing season. The detailed crop loss due to natural disasters, such as floods, drought, and hurricanes, can be dynamically assessed using the monthly in-season crop cover maps.

Moreover, we will optimize the crop type classification algorithm and test more AI/ML strategies to improve the mapping results. The zonal statistics function estimates the crop acreage by counting pixels in crop cover maps. This result could be biased because of the known data issues discussed above. Therefore, we will explore the feasibility of integrating the advanced crop yield estimation models and frameworks (Jeffries et al., 2020; Lin et al., 2020; Li et al., 2021; Andrade et al., 2022; Eamen et al., 2022; Nóia Júnior et al., 2022) into the system to optimize the statistics results.

5. Conclusion

This paper introduced the design, implementation, and application of iCrop for the exploration of crop-specific land cover data in support of U.S. agriculture. Three types of new ML-based crop cover data products, pre-season crop cover, in-season crop cover, and R-CDL, were disseminated through this new cyberinformatics tool based on FAIR data principles. Improved from the service-oriented architecture of well-known CropScape, iCrop enabled interoperable mapping and geoprocessing capability of timely observation of crop-specific land cover through the easy-to-use web application and OGC standards-compliant GIS software. We demonstrated several common geoprocessing functionalities, including on-demand file download, zonal statistics, and PDF map production, using the iCrop web client. Specifically, the change analysis function was applied to analyze the LULC change for the Kossuth County of Iowa between 2021 and 2022. Also, we tested the interoperability of the iCrop web service using generic GIS software, such as QGIS, and web-based GIS platforms, such as FGDC's GeoPlatform based on ArcGIS Online. The result suggested that the proposed cyberinformatics tool can efficiently and effectively provide timely agricultural LULC information to facilitate U.S. agricultural information management and decision support.

Software availability

Software Name: iCrop

Developer: Center for Spatial Information Science and Systems, George Mason University

iCrop Web Client: https://cloud.csiss.gmu.edu/icrop/.

iCrop WMS: https://cloud.csiss.gmu.edu/icrop-service/wms_icrop.cgi?SERVICE=WMS&VERSION=1.1.1&REQUEST=GetCapabilities.

iCrop WCS: https://cloud.csiss.gmu.edu/icrop-service/wms_icrop.cgi?SERVICE = wcs&VERSION = 2.0.1&REQUEST = GetCapabilitieshttps://cloud.csiss.gmu.edu/icrop-service/wms_icrop.cgi?SERVICE=wcs&VERSION=2.0.1&REQUEST=GetCapabilities.

ArcGIS REST API: https://crop.csiss.gmu.edu/arcgis/rest/services/iCrop_Service/MapServer

CRediT authorship contribution statement

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

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgment

This research is supported by a grant from USDA NIFA (Grant #2021-67021-34151, PI: Dr. Liping Di) and a grant from NSF (Grant #2236137, PI: Dr. Liping Di). The authors would like to thank two anonymous reviewers for their valuable comments and suggestions.

References

Abernethy, J., Beeson, P., Boryan, C., Hunt, K., Sartore, L., 2023. Preseason crop type prediction using crop sequence boundaries. Comput. Electron. Agric. 208, 107768 https://doi.org/10.1016/j.compag.2023.107768.

Abunnasr, Y., Mhawej, M., 2022. Towards a combined Landsat-8 and Sentinel-2 for 10-m land surface temperature products: The Google Earth Engine monthly Ten-ST-GEE system. Environmental Modelling & Software 155, 105456. https://doi.org/10.1016/j.envsoft.2022.105456.

Acharya, P., Burgers, T., Nguyen, K.-D., 2022. AI-enabled droplet detection and tracking for agricultural spraying systems. Comput. Electron. Agric. 202, 107325 https://doi.org/10.1016/j.compag.2022.107325.

Andrade, J.F., Mourtzinis, S., Rattalino Edreira, J.I., Conley, S.P., Gaska, J., Kandel, H.J., Lindsey, L.E., Naeve, S., Nelson, S., Singh, M.P., Thompson, L., Specht, J.E., Grassini, P., 2022. Field validation of a farmer supplied data approach to close soybean yield gaps in the US North Central region. Agr. Syst. 200, 103434 https://doi.org/10.1016/j.agsy.2022.103434.

Armstrong, S., Khandelwal, P., Padalia, D., Senay, G., Schulte, D., Andales, A., Breidt, F. J., Pallickara, S., Pallickara, S.L., 2022. Attention-based convolutional capsules for evapotranspiration estimation at scale. Environ. Model. Softw. 152, 105366 https://doi.org/10.1016/j.envsoft.2022.105366.

Barton, C.M., Ames, D., Chen, M., Frank, K., Jagers, H.R.A., Lee, A., Reis, S., Swantek, L., 2022. Making modeling and software FAIR. Environ. Model. Softw. 156, 105496 https://doi.org/10.1016/j.envsoft.2022.105496.

Baum, M.E., Sawyer, J.E., Nafziger, E.D., Huber, I., Thorburn, P.J., Castellano, M.J., Archontoulis, S.V., 2023. Evaluating and improving APSIM's capacity in simulating long-term corn yield response to nitrogen in continuous- and rotated-corn systems. Agr. Syst. 207, 103629 https://doi.org/10.1016/j.agsy.2023.103629.

Bayat, B., Camacho, F., Nickeson, J., Cosh, M., Bolten, J., Vereecken, H., Montzka, C., 2021. Toward operational validation systems for global satellite-based terrestrial essential climate variables. Int. J. Appl. Earth Obs. Geoinf. 95, 102240 https://doi.org/10.1016/j.jag.2020.102240.

Dewitz, J., 2021. National Land Cover Database (NLCD) 2019 Products. 10.5066/P9KZCM54.

Di, L., Üstündağ, B. (Eds.), 2021. Agro-geoinformatics: Theory and Practice, Springer Remote Sensing/Photogrammetry. Springer International Publishing, Cham. 10.1007/978-3-030-66387-2.

Eamen, L., Brouwer, R., Razavi, S., 2022. Comparing the applicability of hydro-economic modelling approaches for large-scale decision-making in multi-sectoral and multi-regional river basins. Environ. Model. Softw. 152, 105385 https://doi.org/10.1016/j.envsoft.2022.105385.

Erazo Ramirez, C., Sermet, Y., Molkenthin, F., Demir, I., 2022. HydroLang: An open-source web-based programming framework for hydrological sciences. Environ. Model. Softw. 157, 105525 https://doi.org/10.1016/j.envsoft.2022.105525.

Graf, L.V., Perich, G., Aasen, H., 2022. EOdal: An open-source Python package for large-scale agroecological research using Earth Observation and gridded environmental data. Comput. Electron. Agric. 203, 107487 https://doi.org/10.1016/j.compag.2022.107487.

Han, W., Yang, Z., Di, L., Mueller, R., 2012. CropScape: A Web service based application for exploring and disseminating US conterminous geospatial cropland data products for decision support. Comput. Electron. Agric. 84, 111–123. https://doi.org/10.1016/j.compag.2012.03.005.

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. (Basel) 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. Science of The Total Environment 733, 138869. https://doi.org/10.1016/j.scitotenv.2020.138869.

Hu, L., Fang, Z., Zhang, M., Jiang, L., Yue, P., 2022. Facilitating Typhoon-Triggered Flood Disaster-Ready Information Delivery Using SDI Services Approach—A Case Study in Hainan. Remote Sens. (Basel) 14, 1832. https://doi.org/10.3390/rs14081832.

Hu, L., Zhang, C., Zhang, M., Shi, Y., Lu, J., Fang, Z., 2023. Enhancing FAIR Data Services in Agricultural Disaster: A Review. Remote Sensing 15, 2024. https://doi.org/10.3390/rs15082024.

Jeffries, G.R., Griffin, T.S., Fleisher, D.H., Naumova, E.N., Koch, M., Wardlow, B.D., 2020. Mapping sub-field maize yields in Nebraska, USA by combining remote sensing imagery, crop simulation models, and machine learning. Precis. Agric. 21, 678–694. https://doi.org/10.1007/s11119-019-09689-z.

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.

Leo, S., De Antoni Migliorati, M., Nguyen, T.H., Grace, P.R., 2023. Combining remote sensing-derived management zones and an auto-calibrated crop simulation model to determine optimal nitrogen fertilizer rates. Agr. Syst. 205, 103559 https://doi.org/10.1016/j.agsy.2022.103559.

Li, R., Wei, C., Afroz, M.D., Lyu, J., Chen, G., 2021. A GIS-based framework for local agricultural decision-making and regional crop yield simulation. Agr. Syst. 193, 103213 https://doi.org/10.1016/j.agsy.2021.103213.

Lieberman, J., Hempelmann, N., Stephens, A., Ehbrecht, C., Smith, T., Landry, T., Wilson, C., Pechorro, E., 2022. FAIR building blocks for climate resilience information systems (No. EGU22-6593). Presented at the EGU22, Copernicus Meetings. 10.5194/egusphere-egu22-6593.

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.

Lin, L., Di, L., Zhang, C., Guo, L., Tang, J., Yu, E.G., Rahman, Md.S., Zhao, H., Yu, Z., Sun, Z., Gaigalas, J., 2019. Building Near-Real-Time MODIS Data Fusion Workflow to Support Agricultural Decision-making Applications, in: 2019 8th International Conference on Agro-Geoinformatics (Agro-Geoinformatics). Presented at the 2019 8th International Conference on Agro-Geoinformatics (Agro-Geoinformatics), pp. 1–5. 10.1109/Agro-Geoinformatics.2019.8820229.

Lin, L., Di, L., Yu, E., Zhang, C., Zhao, H., 2021. Disaster Information Dissemination During Emergency Event: An Experiment in OGC Disaster Resilience Pilot, in: 2021 9th International Conference on Agro-Geoinformatics (Agro-Geoinformatics). Presented at the 2021 9th International Conference on Agro-Geoinformatics (Agro-Geoinformatics), pp. 1–4. 10.1109/Agro-Geoinformatics50104.2021.9530289.

Lin, T., Zhong, R., Wang, Y., Xu, J., Jiang, H., Xu, J., Ying, Y., Rodriguez, L., Ting, K.C., Li, H., 2020. DeepCropNet: a deep spatial-temporal learning framework for county-level corn yield estimation. Environ. Res. Lett. 15, 034016 https://doi.org/10.1088/1748-9326/ab66cb.

Massigoge, I., Carcedo, A., de Borja Reis, A.F., Mitchell, C., Day, S., Oliverio, J., Truong, S.H., McCormick, R.F., Rotundo, J., Lira, S., Ciampitti, I., Messina, C.D., 2023. Exploring avenues for agricultural intensification: A case study for maize-soybean in the Southern US region. Agr. Syst. 204, 103539 https://doi.org/10.1016/j.agsy.2022.103539.

McKenna, J., MapServer, P.S.C., 2021. MapServer. https://doi.org/10.5281/zenodo.5842013.

Medina, H., Tian, D., Abebe, A., 2021. On optimizing a MODIS-based framework for in-season corn yield forecast. Int. J. Appl. Earth Obs. Geoinf. 95, 102258 https://doi.org/10.1016/j.jag.2020.102258.

Meki, M.N., Osorio-Leyton, J., Steglich, E.M., Kiniry, J.R., Propato, M., Winchell, M., Rathjens, H., Angerer, J.P., Norfleet, L.M., 2023. Plant parameterization and APEXgraze model calibration and validation for US land resource region H grazing lands. Agr. Syst. 207, 103631 https://doi.org/10.1016/j.agsy.2023.103631.

Menegon, S., Fadini, A., Perini, L., Sarretta, A., Depellegrin, D., De Maio, E., Farella, G., Landini, M., Fava, C., Ferrarin, C., Ghezzo, M., Manea, E., Porporato, E.M.D., Pastres, R., Sedioli, O., Politikos, D.V., Maina, I., Kavadas, S., Matarrese, R., Barbanti, A., 2023. A geoportal of data and tools for supporting Maritime Spatial Planning in the Adriatic-Ionian Region. Environ. Model. Softw. 160, 105585 https://doi.org/10.1016/j.envsoft.2022.105585.

Mulrooney, T., Foster, R., Jha, M., Beni, L.H., Kurkalova, L., Liang, C.L., Miao, H., Monty, G., 2021. Using geospatial networking tools to optimize source locations as applied to the study of food availability: A study in Guilford County. North Carolina. Applied Geography 128, 102415. https://doi.org/10.1016/j.apgeog.2021.102415.

Nash, E., Korduan, P., Bill, R., 2009. Applications of open geospatial web services in precision agriculture: a review. Precis. Agric. 10, 546–560. https://doi.org/10.1007/s11119-009-9134-0.

USDA NASS, 2021. CroplandCROS web app updates existing geospatial data product for agricultural commodities in the United States [WWW Document]. URL https://www.nass.usda.gov/Newsroom/archive/2021/10-28-2021.php (accessed 11.6.22).

USDA NASS, 2022a. Iowa Ag News - 2021 Corn County Estimates.

USDA NIFA, 2021. NIFA Invests $15.5M in Food and Agriculture Cyberinformatics Tools to Boost Agricultural Production [WWW Document]. National Institute of Food and Agriculture. URL http://www.nifa.usda.gov/about-nifa/press-releases/nifa-invests-155m-food-agriculture-cyberinformatics-tools-boost (accessed 10.11.22).

Nóia Júnior, R. de S., Fraisse, C.W., Bashyal, M., Mulvaney, M.J., Seepaul, R., Zientarski Karrei, M.A., Iboyi, J.E., Perondi, D., Cerbaro, V.A., Boote, K.J., 2022. Brassica carinata as an off-season crop in the southeastern USA: Determining optimum sowing dates based on climate risks and potential effects on summer crop yield. Agricultural Systems 196. https://doi.org/10.1016/j.agsy.2021.103344.

OGC, 2006. OpenGIS® Web Map Server Implementation Specification [WWW Document]. URL https://portal.ogc.org/files/?artifact_id=14416.

OGC, 2012. OGC® WCS 2.0 Interface Standard [WWW Document]. URL https://portal.ogc.org/files/09-110r4.

Pengra, B., Stehman, S.V., Horton, J.A., Wellington, D.F., 2021. Land Change Monitoring, Assessment, and Projection (LCMAP) Collection 1.1 Annual Land Cover and Land Cover Change Validation Tables (1985-2018) for the Conterminous United States. 10.5066/P9MLPFOH.

Pierre Pott, L., Jorge Carneiro Amado, T., Augusto Schwalbert, R., Mateus Corassa, G., Antonio Ciampitti, I., 2022. Crop type classification in Southern Brazil: Integrating remote sensing, crop modeling and machine learning. Computers and Electronics in Agriculture 201. https://doi.org/10.1016/j.compag.2022.107320.

Redhead, J.W., Burkmar, R., Brown, M., Pywell, R.F., 2022. E-Planner: A web-based tool for planning environmental enhancement on British agricultural land. Environ. Model. Softw. 155, 105437 https://doi.org/10.1016/j.envsoft.2022.105437.

Rufin, P., Bey, A., Picoli, M., Meyfroidt, P., 2022. Large-area mapping of active cropland and short-term fallows in smallholder landscapes using PlanetScope data. International Journal of Applied Earth Observation and Geoinformation 112, 102937. https://doi.org/10.1016/j.jag.2022.102937.

Sanderman, J., Savage, K., Dangal, S.R.S., Duran, G., Rivard, C., Cavigelli, M.A., Gollany, H.T., Jin, V.L., Liebig, M.A., Omondi, E.C., Rui, Y., Stewart, C., 2021. Can Agricultural Management Induced Changes in Soil Organic Carbon Be Detected Using Mid-Infrared Spectroscopy? Remote Sens. (Basel) 13, 2265. https://doi.org/10.3390/rs13122265.

Singh, A., Jones, S., Ganapathysubramanian, B., Sarkar, S., Mueller, D., Sandhu, K., Nagasubramanian, K., 2021. Challenges and Opportunities in Machine-Augmented Plant Stress Phenotyping. Trends in Plant Science 26, 53–69. https://doi.org/10.1016/j.tplants.2020.07.010.

Soltanikazemi, M., Minaei, S., Shafizadeh-Moghadam, H., Mahdavian, A., 2022. Field-scale estimation of sugarcane leaf nitrogen content using vegetation indices and spectral bands of Sentinel-2: Application of random forest and support vector regression. Comput. Electron. Agric. 200, 107130 https://doi.org/10.1016/j.compag.2022.107130.

Tan, X., Jiao, J., Zhong, Y., Ma, A., Xu, Y., Sha, Z., Huang, F., Wan, Y., Hu, W., 2022. The CNRIEEEMC: A communication-navigation-remote sensing-integrated ecological environment emergency monitoring chain for tailings areas. Int. J. Appl. Earth Obs. Geoinf. 108, 102710 https://doi.org/10.1016/j.jag.2022.102710.

Tekinerdogan, B., Verdouw, C., 2020. Systems Architecture Design Pattern Catalog for Developing Digital Twins. Sensors 20, 5103. https://doi.org/10.3390/s20185103.

Top, J., Janssen, S., Boogaard, H., Knapen, R., Şimşek-Şenel, G., 2022. Cultivating FAIR principles for agri-food data. Comput. Electron. Agric. 196, 106909 https://doi.org/10.1016/j.compag.2022.106909.

USDA NASS, 2022b. Iowa Ag News - 2021 Soybeans County Estimates.

USDA NIFA, 2020. NIFA Invests Over $7 Million in Big Data, Artificial Intelligence, and Other Cyberinformatics Research [WWW Document]. National Institute of Food and Agriculture. URL http://www.nifa.usda.gov/about-nifa/press-releases/nifa-invests-over-7-million-big-data-artificial-intelligence-other (accessed 10.11.22).

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

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.

White, E.L., Thomasson, J.A., Auvermann, B., Kitchen, N.R., Pierson, L.S., Porter, D., Baillie, C., Hamann, H., Hoogenboom, G., Janzen, T., Khosla, R., Lowenberg-DeBoer, J., McIntosh, M., Murray, S., Osborn, D., Shetty, A., Stevenson, C., Tevis, J., Werner, F., 2021. Report from the conference, 'identifying obstacles to applying big data in agriculture'. Precis. Agric. 22, 306–315. https://doi.org/10.1007/s11119-020-09738-y.

Wilkinson, M.D., Dumontier, M., Aalbersberg, Ij.J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.-W., da Silva Santos, L.B., Bourne, P.E., Bouwman, J., Brookes, A.J., Clark, T., Crosas, M., Dillo, I., Dumon, O., Edmunds, S., Evelo, C.T., Finkers, R., Gonzalez-Beltran, A., Gray, A.J.G., Groth, P., Goble, C., Grethe, J.S., Heringa, J., 't Hoen, P.A.C., Hooft, R., Kuhn, T., Kok, R., Kok, J., Lusher, S.J., Martone, M.E., Mons, A., Packer, A.L., Persson, B., Rocca-Serra, P., Roos, M., van Schaik, R., Sansone, S.-A., Schultes, E., Sengstag, T., Slater, T., Strawn, G., Swertz, M. A., Thompson, M., van der Lei, J., van Mulligen, E., Velterop, J., Waagmeester, A., Wittenburg, P., Wolstencroft, K., Zhao, J., Mons, B., 2016. The FAIR Guiding Principles for scientific data management and stewardship. Sci Data 3, 160018. 10.1038/sdata.2016.18.

Wolfert, S., Verdouw, C., van Wassenaer, L., Dolfsma, W., Klerkx, L., 2023. Digital innovation ecosystems in agri-food: design principles and organizational framework. Agr. Syst. 204, 103558 https://doi.org/10.1016/j.agsy.2022.103558.

Wu, F., Wu, B., Zhu, W., Yan, N., Ma, Z., Wang, L., Lu, Y., Xu, J., 2021. ETWatch cloud: APIs for regional actual evapotranspiration data generation. Environ. Model. Softw. 145, 105174 https://doi.org/10.1016/j.envsoft.2021.105174.

Yang, Z., Zhang, C., Zhao, H., Sun, Z., Bindlish, R., Liu, P.-W., Colliander, A., Mueller, R., Di, L., Crow, W., Reichle, R.H., 2021. Crop-CASMA - A Web GIS Tool for Cropland Soil Moisture Monitoring and Assessment Based on SMAP Data, in: 2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS. Presented at the 2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS, pp. 6315–6318. 10.1109/IGARSS47720.2021.9554526.

Yaramasu, R., Bandaru, V., Pnvr, K., 2020. Pre-season crop type mapping using deep neural networks. Comput. Electron. Agric. 176, 105664 https://doi.org/10.1016/j.compag.2020.105664.

Yue, P., Shangguan, B., Hu, L., Jiang, L., Zhang, C., Cao, Z., Pan, Y., 2022. Towards a training data model for artificial intelligence in earth observation. Int. J. Geogr. Inf. Sci. 36, 2113–2137. https://doi.org/10.1080/13658816.2022.2087223.

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., Hao, P., Yang, Z., Lin, L., Zhao, H., Guo, L., 2021. 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., Lin, L., Li, H., Guo, L., Yang, Z., Yu, E.G., Di, Y., Yang, A., 2022a. Towards automation of in-season crop type mapping using spatiotemporal crop information and remote sensing data. Agr. Syst. 201, 103462 https://doi.org/10.1016/j.agsy.2022.103462.

Zhang, M., Jiang, L., Yue, P., Gong, J., 2020. Interoperable web sharing of environmental models using OGC web processing service and Open Modeling Interface (OpenMI). Environ. Model. Softw. 133, 104838 https://doi.org/10.1016/j.envsoft.2020.104838.

Zhang, C., Yang, Z., Di, L., Yu, E.G., Zhang, B., Han, W., Lin, L., Guo, L., 2022b. Near-real-time MODIS-derived vegetation index data products and online services for CONUS based on NASA LANCE. Sci. Data 9, 477. https://doi.org/10.1038/s41597-022-01565-2.

Zhang, C., Yang, Z., Zhao, H., Sun, Z., Di, L., Bindlish, R., Liu, P.-W., Colliander, A., Mueller, R., Crow, W., Reichle, R.H., Bolten, J., Yueh, S.H., 2022c. Crop-CASMA: A web geoprocessing and map service based architecture and implementation for serving soil moisture and crop vegetation condition data over U.S. Cropland. International Journal of Applied Earth Observation and Geoinformation 112, 102902. https://doi.org/10.1016/j.jag.2022.102902.