



Integrating Multiscale Geospatial Analysis for Monitoring Crop Growth, Nutrient Distribution, and Hydrological Dynamics in Large-Scale Agricultural Systems

Olatunde D. Akanbi^{1,2,3} · Deepa C. Bhuvanagiri^{2,3,4} · Erika I. Barcelos^{1,2,3} · Arafath Nihar^{2,4} · Brian Gonzalez Hernandez^{2,4} · Jeffrey M. Yarus^{1,2,3}  · Roger H. French^{1,2,3,4}

Accepted: 4 November 2023
© The Author(s) 2024

Abstract

Monitoring crop growth, soil conditions, and hydrological dynamics are imperative for sustainable agriculture and reduced environmental impacts. This interdisciplinary study integrates remote sensing, digital soil mapping, and hydrological data to elucidate intricate connections between these factors in the state of Ohio, USA. Advanced spatiotemporal analysis techniques were applied to key datasets, including the MODIS sensor satellite imagery, USDA crop data, soil datasets, Aster GDEM, and USGS stream gauge measurements. Vegetation indices derived from MODIS characterized crop-specific phenology and productivity patterns. Exploratory spatial data analysis show relationships of vegetation dynamics and soil properties, uncovering links between plant vigor, edaphic fertility, and nutrient distributions. Correlation analysis quantified these relationships and their seasonal evolution. Examination of stream gauge data revealed insights into spatiotemporal relationships of nutrient pollution and stream discharge. By synthesizing diverse geospatial data through cutting-edge data analytics, this work illuminated complex interactions between crop health, soil nutrients, and water quality in Ohio. The methodology and findings provide actionable perspectives to inform sustainable agricultural management and environmental policy. This study demonstrates the significant potential of open geospatial resources when integrated using a robust spatiotemporal framework. Integrating additional measurements and high-resolution data sources through advanced analytics and interactive visualizations could strengthen these insights.

Keywords MODIS NDVI · Soil-crop monitoring · Data integration · Nutrient distribution · Exploratory geospatial analysis

Introduction

The increasing availability of spatial data from diverse sources, including satellite imagery and sensors, has provided the academic community with unprecedented opportunities to gain valuable insights into the environment, make informed decisions, mitigate nutrient loss, and promote environmental safety through hazard prevention. Spatial data have found diverse applications, such as in weather Ahmad and Zeeshan (2022); Peng et al. (2016), soil minerals Demattê et al. (2018); Palmer et al. (2021), monitoring floods Fisher et al. (2016), preventing fire hazards Navarro et al. (2017), informing population decisions Azar et al. (2013), water, building, land cover, and information management Yaman

et al. (2021). These examples highlight the wealth of information that holds extensive potential for monitoring and analyzing agricultural environments.

In recent years, researchers have used geospatial techniques using satellite imagery to monitor vegetation growth Jackson et al. (2004). The widely employed Normalized Difference Vegetation Index (NDVI), derived from data sources such as MODIS, Landsat, or Sentinel Rouse et al. (1974a, b); Waring et al. (2006), has been instrumental. Aside from NDVI metrics, Leaf Area Index (LAI) has been used in canopies and transpiration Fang and Liang (2014); Chen (2018), and Soil Adjusted Vegetation Index (SAVI) has been used to minimize the influences of soil on canopy Huete (1988). Through examining the spectral reflectance of vegetation, NDVI provides a valuable indicator of plant health and vigor Huang et al. (2021); DeFries and Townshend (1994); Pettorelli et al. (2005). Other indices include Enhanced Vegetation Index (EVI) Jiang et al. (2008), Normalized Difference

✉ Jeffrey M. Yarus
jmy41@case.edu

Extended author information available on the last page of the article

Water Index (NDWI) Gao (1996), Green Normalized Difference Vegetation Index (GNDVI) Shaver et al. (2006), Chlorophyll Index (CI) Shibayama and Akiyama (1986); Shrestha et al. (2012), and Water Index (WI) Peñuelas et al. (1997). Most of these indices have NDVI as the baseline.

While progress has been made in crop monitoring using satellite data from the NDVI, challenges persist in relating vegetation patterns to key *soil* factors impacting plant growth and vigor. Most of these challenges hinge on integrating diverse multi-modal data. Digital soil mapping and analysis of soil surveys have been valuable for understanding soil nutrients, texture, fertility, and more Subburayalu et al. (2014); Arshad et al. (1997); Mallah et al. (2022). In addition, hydrological data such as stream-flow measurements enables the characterization of water quality issues such as agricultural runoff and nutrient pollution Michalak et al. (2013); Smith et al. (2015). In particular, the flow of excess contaminants, including nitrate and nitrite, into water streams has contributed to serious environmental and health problems Schlossberg (2017); Sinha et al. (2017).

In addition, increased flow of nitrate and nitrite into waterways has encouraged eutrophication, a depletion of oxygen in water, leading to the suffocation of marine life Singh et al. (2022). High concentrations of nitrate and nitrite found in drinking water have caused methemoglobinemia (“blue baby syndrome”) in infants Manassaram et al. (2010); Coffman et al. (2021). Because of its impact on the health and well-being of people and the environment, it is of interest to determine how to model and predict the behavior of nitrate and nitrite, besides other contaminants, in watersheds.

Many features can affect the presence of nitrate and nitrite in watersheds, such as soil characteristics and nutrients, agricultural practices, and discharge Dubrovsky and Hamilton (2010). Discharge, or stream-flow, is defined as the speed a volume of water has as it crosses a specific reference point. Linking water contamination spatiotemporally to land management practices can inform conservation efforts.

Integrating data on crops, soils, and water can provide a holistic perspective not attainable through isolated data sources. A spatiotemporal approach leveraging these interconnected domains has immense potential for novel and interdisciplinary insights into complex agricultural and environmental systems Chen et al. (2004); Wang et al. (2003). However, significant challenges persist in effectively tracking crop health over time, capturing spatial variability within fields, and linking vegetation dynamics to key soil factors that impact plant growth, yield, and vigor. While significant progress has been made in crop monitoring using geospatial data, challenges persist in effectively integrating diverse data sources of varying spatial and temporal resolution and conducting Exploratory Spatial Data Analysis (ESDA) simultaneously in space and time Hamdi et al. (2022).

The integration and correlation of multi-modal data from different sources and databases are crucial for understanding the intricate relationships between these features. However, platforms that offer comprehensive data management often lack capabilities for flexible exploratory analysis.

Although platforms like IBM’s Physical Analytics Integrated Data Repository and Services (PAIRS) Lu et al. (2016); Lu and Hamann (2021); Klein et al. (2015) offer comprehensive data management and analytic capabilities, challenges remain in effectively integrating and correlating diverse data and conducting ESDA. While IBM offers curated data spanning multiple petabytes, effectively harnessing its potential for exploratory analysis remains a complex task. To address the ongoing challenges in data integration and ESDA, we leveraged the capabilities of the IBM Environmental Intelligence Suite (EIS). EIS is a suite of tools and services built on IBM PAIRS, tailored for environmental monitoring and analysis.

Our aim is to integrate data from multiple sources in different formats with different spatial and temporal resolutions, including MODIS satellite imagery MODIS (2019), USDA crop planting data United States Department of Agriculture (2019), soil databases (SSURGO and WoSIS) Staff (2019); Batjes et al. (2019), Aster GDEM elevation data Tachikawa et al. (2011a,b), and USGS stream-flow measurements Survey (2019), to conduct a comprehensive analysis of crop growth, health, and water/nutrient flow in Ohio during 2019.

Effective crop monitoring through integrating data sources can support improved yield forecasting, targeted field-specific interventions, and optimized inputs to manage plant growth. In addition, a greater understanding of water and nutrient behavior can aid in reducing the effects of pollution from contributing factors, such as agricultural land use.

Ohio was selected as a baseline state and region of interest for this analysis given its significance in the Center for Advancing Sustainable and Distributed Fertilizer Production (CASFER), a National Science Foundation (NSF) Engineering Research Center (ERC) Botte et al. (2023); Ai et al. (2023). This state has major importance and serves as a baseline location where CASFER technologies will be implemented and validated for creating a nitrogen-circular economy. This study aimed to leverage the wealth of data and expertise in this state before extending the analysis to other states across the United States.

Advanced analyses are needed to fully harness the wealth of insights that can be derived from integrating multi-modal geospatial data. By integrating data spanning satellite imagery, soil surveys, land use, and hydrological data, researchers can gain a multifaceted understanding of agricultural ecosystems.

Most prior work has focused on analyzing one data type in isolation. An integrated spatiotemporal analysis approach could capture intricate connections between crop health, edaphic

factors, and water dynamics. One of the key challenges in geospatial analysis is the storage and scalability of large amounts of data. With the Common Research Analytics and Data Lifecycle Environment (CRADLE), we can overcome this challenge. CRADLE, our Distributed and High-Performance Computer (D/HPC) integrated Hadoop Cluster Hu et al. (2013); Khalilnejad et al. (2020), provides the infrastructure to handle and analyze big geospatial data effectively and efficiently. By leveraging the capabilities of CRADLE, we successfully downloaded and ingested a vast amount and variety of geospatial data from IBM EIS and other sources, ensuring efficient storage and processing. With this infrastructure, we aimed to explore the correlations between vegetation patterns, soil properties, and nutrient distribution.

In this study, we show the potential for leveraging open geospatial data and exploratory analysis techniques to achieve a more holistic perspective on agricultural and environmental management Lnenicka and Nikiforova (2021); Coughlan (2020). Our primary aim is to integrate multi-modal data from multiple sources in different formats with different spatial and temporal resolutions, including satellite imagery, soil databases, elevation data, and stream-flow measurements, to conduct a comprehensive analysis of crop growth, health, and water/nutrient flow in Ohio during 2019. By doing so, we seek to improve crop monitoring, enhance yield forecasting, target field-specific interventions, and optimize inputs to manage plant growth effectively. We shed light on the current state of geospatiotemporal analysis in agriculture, emphasize the significance of integrated approaches,

and contribute to ongoing efforts to promote sustainable agricultural practices and environmental management. This research will explore data science methodologies, present comprehensive analysis findings, and discuss the implications derived from our study, providing valuable insights into the field of agricultural and environmental science.

The advanced spatiotemporal analysis techniques provide unique holistic insights into agricultural and environmental systems not attainable from isolated data sources. The computational framework leverages high-performance computing capabilities to handle massive datasets efficiently. The approach links crop growth dynamics, soil properties, and nutrient transport in streams at broad geospatiotemporal scales.

The following sections of this paper will provide an in-depth exploration of the data science methodologies employed, present the analysis findings comprehensively, and discuss the implications derived from our study.

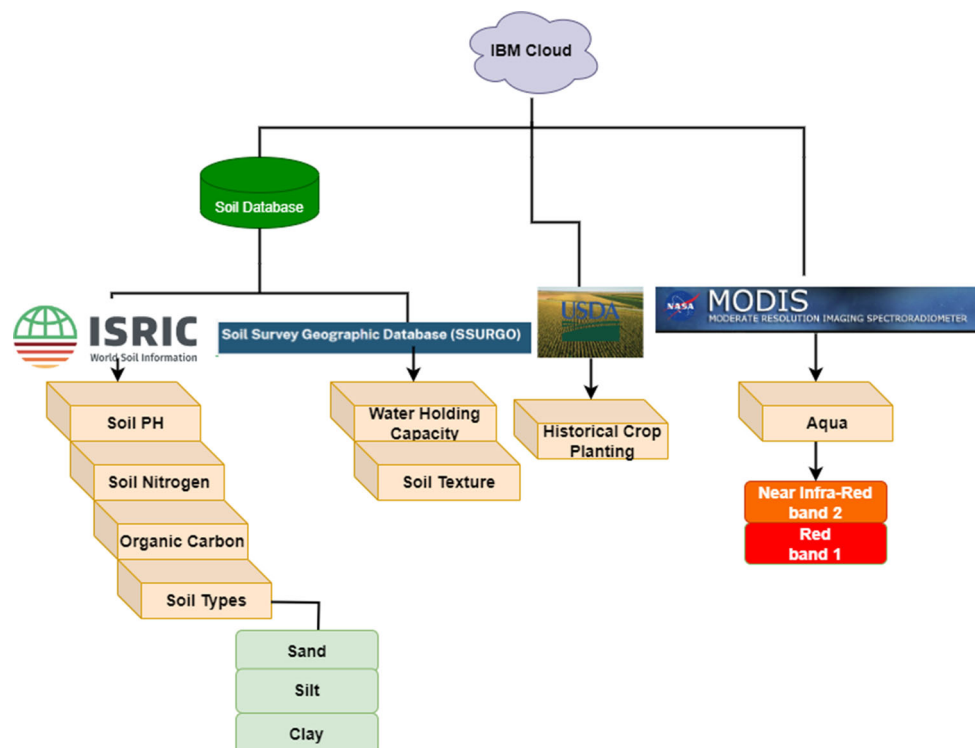
While this manuscript focuses on presenting the methodology and results for Ohio, supplementary results for Texas and Florida were analyzed using the same framework and are provided in the Appendices for reference.

Datasets

Soil and Crop Datasets

Figure 1 shows the data used in this work as obtained from IBM EIS. These data encompass a diverse range of spatial

Fig. 1 Overview of the key data used in this research, including MODIS Aqua satellite imagery, soil data, and USDA crop data



and temporal resolutions, which include daily MODIS Aqua 250m resolution imagery, USDA historical crop planting data at 30m resolution, and soil databases such as SSURGO and WoSIS. Both soil data have a spatial resolution of 250 m. All these data were collected for the year 2019.

- **MODIS Aqua 250m Resolution Imagery**
The MODIS Aqua data comprises remotely sensed imagery captured by the Moderate Resolution Imaging Spectroradiometer (MODIS) instrument aboard the Aqua satellite. The data were collected at a spatial resolution of 250m and provide daily observations. For this research, we used the red and near-infrared (NIR) bands for the analysis.
- **USDA Historical Crop Planting Data (30m Resolution)**
The USDA historical crop planting dataset provides detailed information on crop types and their spatial distribution across the study area. The data has a finer spatial resolution of 30m.
- **Soil Databases (SSURGO and WoSIS) at 250m Resolution**
The soil databases used in our study include the Soil Survey Geographic Database (SSURGO) and the World Soil Information Service (WoSIS). These databases provide comprehensive information about soil properties and characteristics. Both SSURGO and WoSIS have a spatial resolution of 250m, matching the resolution of the MODIS Aqua imagery. WoSIS provides data on soil pH, soil nitrogen, organic carbon, and soil types, while SSURGO offers data on water-holding capacity and soil texture. These soil properties are crucial for understanding the soil's composition and suitability for various applications.

As shown in Fig. 1, by combining the MODIS Aqua imagery, USDA historical crop planting data, and the soil databases (SSURGO and WoSIS), we can gain insights into the dynamic interplay between agricultural practices, soil conditions, and environmental factors within the study area.

Water and Elevation Datasets

To evaluate hydrologic properties and how their behavior varies across space and time, three different data sets were considered. The data collected include Global Digital Elevation Model (GDEM Version 3) data collected from the Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) METI and NASA (2019), the United States Geological Survey National Water Information System (USGS NWIS) Survey (2019) stream flow data, and nutrient information got from The Water Quality Portal (WQP) Survey et al. (2019).

- **ASTER GDEM**

The ASTER GDEM dataset provides elevation information obtained from the Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) instrument aboard Terra, a NASA satellite. These data contain elevation information for over 99% of the Earth's surface, with each GeoTIFF image covering 1 x 1 degree (111 km x 111 km) surface area. The spatial resolution is 30 m and there are about 12,967,201 pixels or estimated elevation values per GeoTIFF. In this study, Version 3 was used, which contains data collected in 2019.

- **USGS NWIS**
The USGS NWIS data contains in situ stream water gauge measurements. Over 22,000 features were measured. For this study, nitrate and nitrite concentrations and discharges were analyzed. Data was collected every 15 min for 2019.
- **WQP**
The WQP provides nutrient content information from stream water gauges collected by the EPA and USGS. The data was collected daily for 2019. *In situ*, water gauge measurements, both on water characteristics such as discharge and nutrient content (e.g., nitrates and nitrites), provide important information on the state of water quality at a time. This, in combination with elevation data, can provide insights into the behavior of these water features in relation to the location of river networks over time.

In the following sections, we will discuss the data preprocessing steps, integration techniques, and analytical methods employed to extract valuable information and derive meaningful conclusions from these data.

Methodology

This research was divided into two distinct sections, each focusing on different aspects of geospatial analysis. The first section was dedicated to the analysis of soil-related factors, including soil properties and crop growth, while the second section centered on the analysis of water-related factors.

Crop and Soil

In this study, we conducted a geospatial analysis of soil properties and crop growth patterns to gain insights into the relationships between soil attributes, vegetation dynamics, and nutrient distribution. The methodology involved the following steps:

Data Acquisition and Integration

We acquired geospatial data relevant to land and crop analysis from IBM EIS, including MODIS satellite imagery,

USDA crop planting data, and soil databases (SSURGO and WoSIS). The data were downloaded as zip files, with each file containing layered GeoTIFFs Ritter and Ruth (1997); OGC (2023). To ensure compatibility and consistency across various sources, these GeoTIFFs were subjected to preprocessing steps, including unzipping, stacking, masking, cropping, and resampling. The data were integrated by aligning spatial and temporal dimensions, addressing inconsistencies, and standardizing coordinate systems.

To integrate the multi-modal data analyzed in this work, including the MODIS, soil, and crop data layers, resampling and aggregation techniques were used to align the data spatially and statistically. For instance, the 250m resolution of the soil data and MODIS imagery were upscaled to the 30m resolution of the crop planting maps. This integration process helped ensure the data layers matched spatially for coherent analysis while preserving the fundamental integrity of the data patterns. As is typical of such up-scaling and aggregation approaches, the integrated data maintained the overall statistical distributions and relationships critical for the correlation analysis while aligning spatially across sources.

Geospatial Analysis of Soil and Crop Parameters

To analyze soil properties, we leveraged IBM Environmental Intelligent Suite to retrieve publicly available national soil databases and extract relevant parameters, including soil classification, organic matter content, pH, and nitrogen levels.

For vegetation dynamics, we used the Normalized Difference Vegetation Index (NDVI) derived from the MODIS Aqua satellite data. The NDVI quantifies vegetation health and vigor by contrasting red and near-infrared reflectance

using the following equation:

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

NIR is the near-infrared reflectance, and Red is the visible red reflectance. NDVI values range from -1 to 1, but noise and cloud contamination are removed by scaling to -0.2 to 1 Zhu et al. (2013). The NDVI time series enabled the characterization of vegetation phenology. Spatial analysis of NDVI patterns was conducted in relation to soil parameters in order to examine correlations between crop growth, soil nutrients, and related factors across the study area.

Acquisition of Water Data

Streamwater features were collected and extracted from three different sources. Elevation information used to calculate slope and stream networks was obtained from ASTER GDEM as satellite imagery provided as a grid of pixels. *In situ* water gauge measurements were collected from USGS NWIS, measuring over 22,000 parameters, including discharge, temperature, pH, etc., every 15 min.

Lastly, nutrient content information from water gauges was obtained from the Water Quality Portal (WQP), which serves as a portal to access data collected from the Environmental Protection Agency (EPA), and USGS. Water gauge data was collected in tabular form.

Geospatial Analysis of Water Parameters

From ASTER GDEM data, hydrographic features and stream networks can be extracted. A Python package, pysheds, was

Fig. 2 Step 1: Pipeline to extract stream networks from digital elevation model data

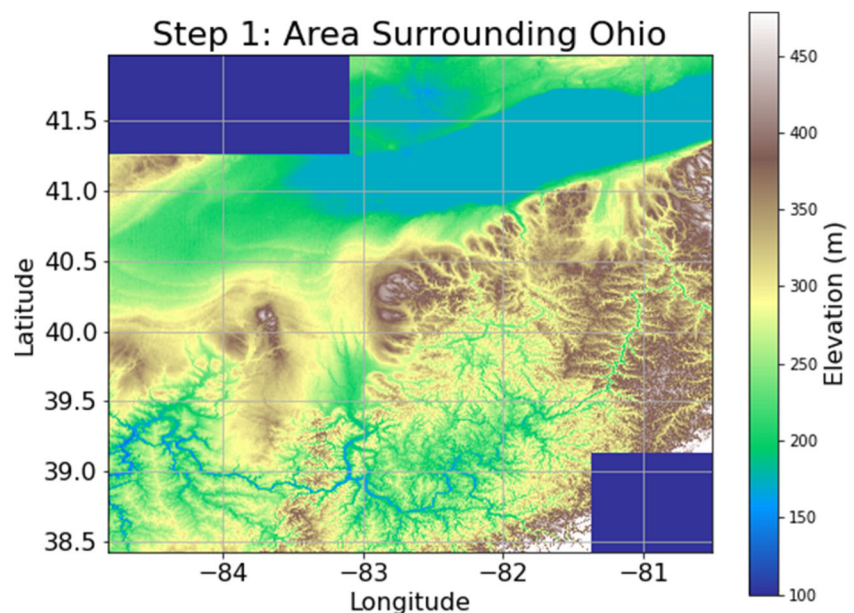
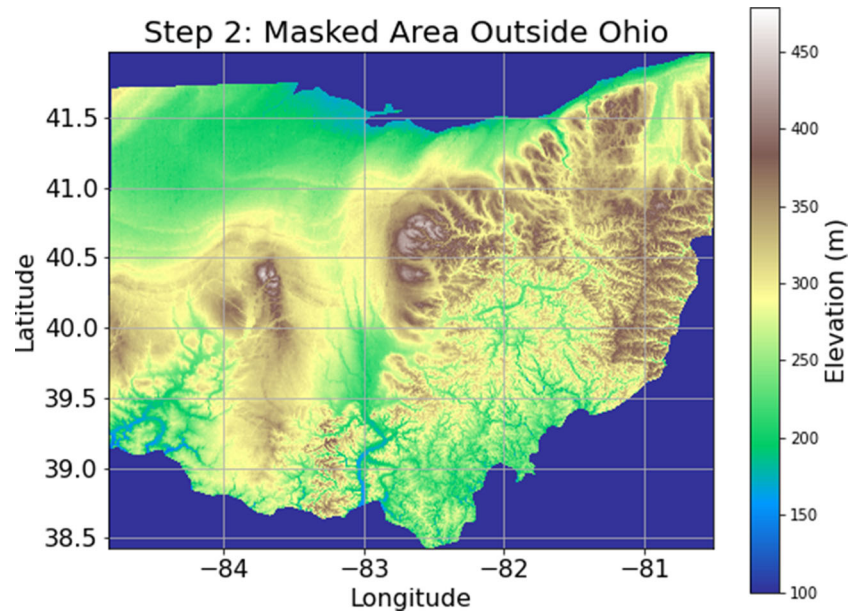


Fig. 3 Step 2: Pipeline to extract stream networks from digital elevation model data



used to implement the pipeline used for extracting stream networks Bartos (2023). This pipeline involves, first, determining the direction of water flow on land based on elevation using the D8 algorithm Jones (2023). This algorithm determines the steepest descent for each cell based on the elevation difference between the current cell and its neighboring cells. Once the direction of flow is determined, cell accumulation can be calculated. Cell accumulation allows one to see how many cells will accumulate or flow into a particular cell based on flow direction. This allows one to see where the streams, rivers, lakes, and oceans are located, as these are areas where water will accumulate or flow. From here, stream networks

(or D8 channels) can be extracted, allowing one to see both the shape and location of these streams or bodies of water.

For this analysis, digital elevation model GeoTIFFs covering the region of Ohio were selected. The area outside of Ohio was masked, and stream networks were extracted following the methodology, as seen in the steps shown in Figs. 2, 3, 4, 5, 6, and 7. The pipeline to extract stream networks from GDEMs can be applied to new versions of GDEMs, allowing the model to update as new information is obtained.

Besides elevation, in situ water gauges offer detailed spatial and temporal insights into the behavior of water features. Spatial and temporal analyses were conducted to understand

Fig. 4 Step 3: Pipeline to extract stream networks from digital elevation model data

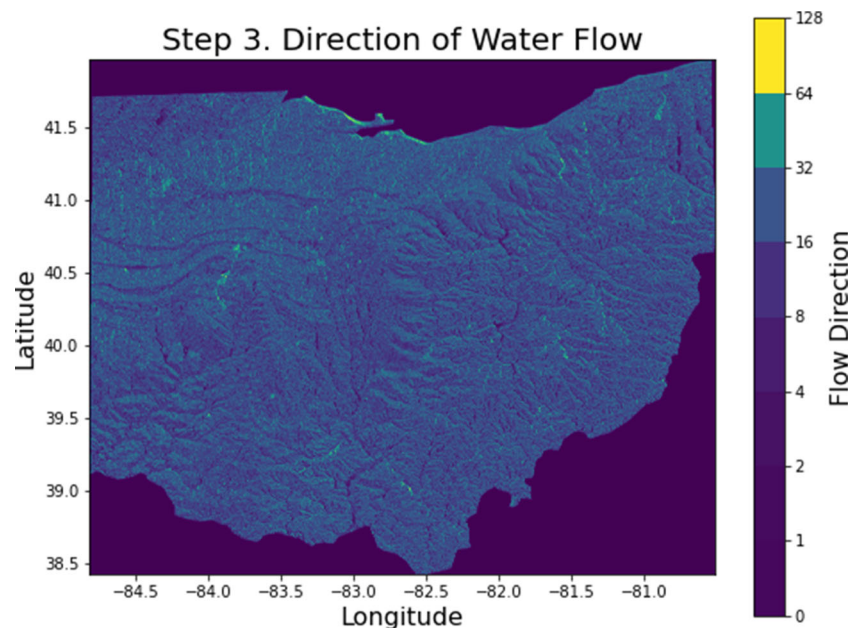


Fig. 5 Step 4: Pipeline to extract stream networks from digital elevation model data

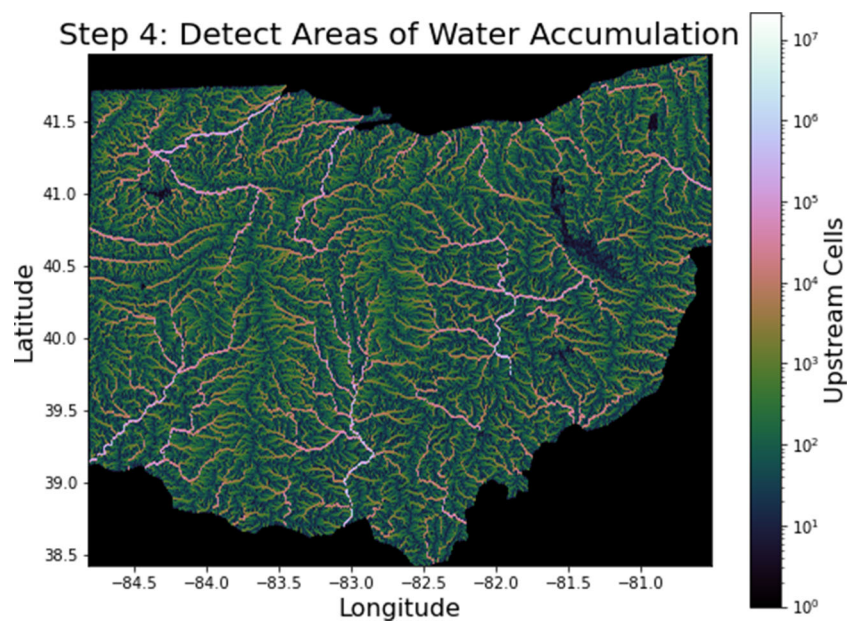


Fig. 6 Step 5: Pipeline to extract stream networks from digital elevation model data

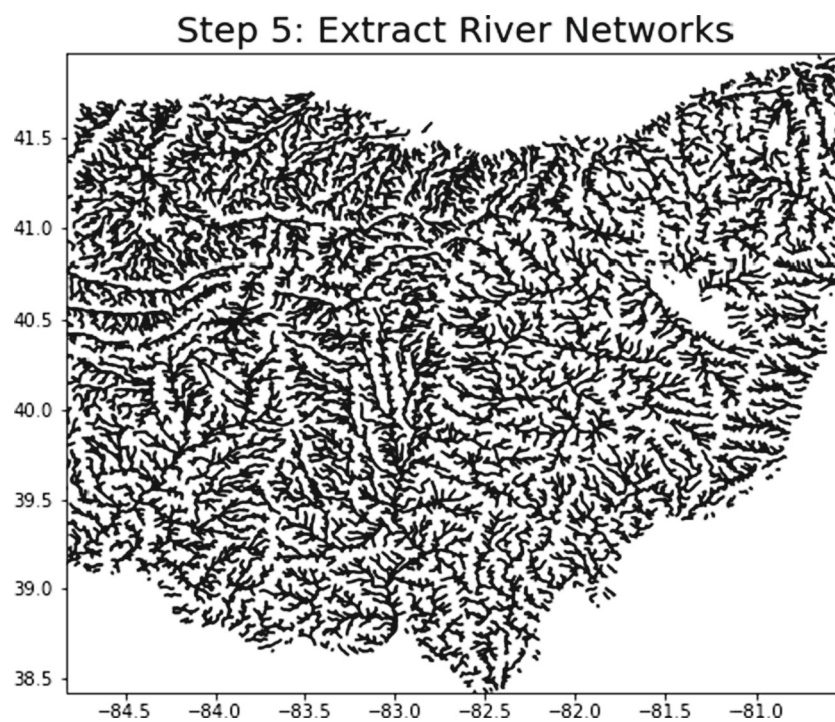
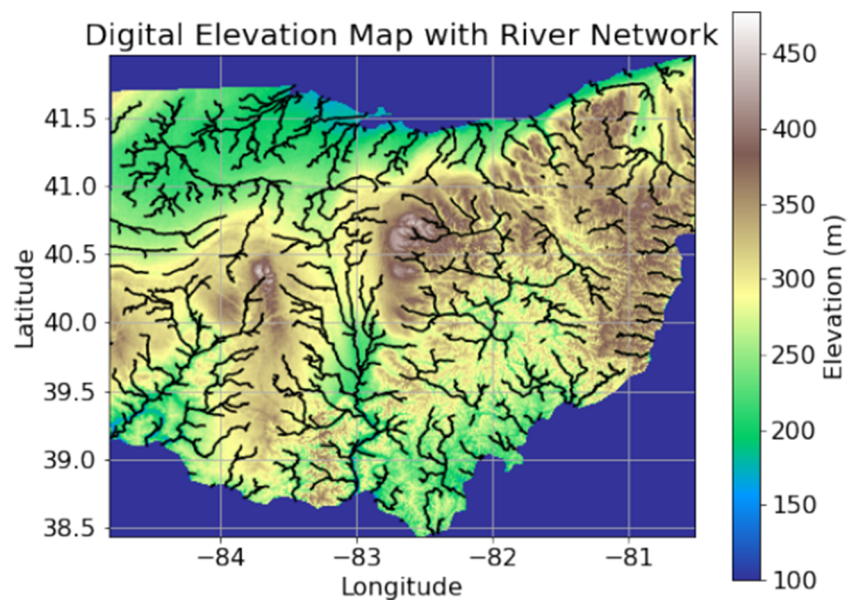


Fig. 7 Result: Pipeline to extract stream networks from digital elevation model data



the behavior of nutrient and water flow in streams over various time periods.

Improved Methodology

The methodology presented here improves on data integration approaches. It uses the powerful capabilities of the R programming language R Core Team (2023) and specialized geospatial analysis packages, specifically the `terra` and `raster` packages Hijmans et al. (2023a,b), to efficiently process and analyze satellite imagery and soil data. The raw images were obtained from the IBM EIS as GeoTIFF layers embedded in zip files representing distinct data. For a given data, each individual GeoTIFF layer corresponds to a distinct point in time. As an example, the daily MODIS red band data contains 365 layers for the year 2019, with each layer representing measurements for a single day.

By leveraging R and these geospatial tools, the multivariate spatiotemporal satellite and soil data were analyzed to draw insightful conclusions. While the `terra` and `raster` packages offer similar functionalities for raster analysis, in this context, they are used complementarily. Specifically, the `terra` library is employed for the efficient processing of multi-temporal satellite data series, including reading, cropping, masking, and stacking. On the other hand, the `raster` library is used for conducting geospatial multi-layer correlation analysis across both the soil and satellite raster images. The integrated use of both packages allows for streamlined data processing through `terra` along with customized analysis workflows via `raster`.

In the spatial correlation analysis, we employed the `corLocal()` function from the `raster` package. This function

determines correlations for each pixel between two raster layers. While Pearson, Kendall, and Spearman correlation coefficients are all viable choices based on data distribution, we opted for the Pearson coefficient for this study. We used the Pearson correlation coefficient for this analysis because the data was normally distributed and a linear relationship exists between the variables. Pearson is appropriate when these assumptions hold, as it assesses the linear dependence between two continuous variables. It was preferable to the Spearman or Kendall correlations, which are non-parametric measures not requiring normally distributed data. Using Pearson allowed us to quantify the strength and direction of the linear relationship between the soil and satellite raster layers.

Additionally, the `magick` package was used to create GIFs for better data visualization Ooms [aut and cre (2023).

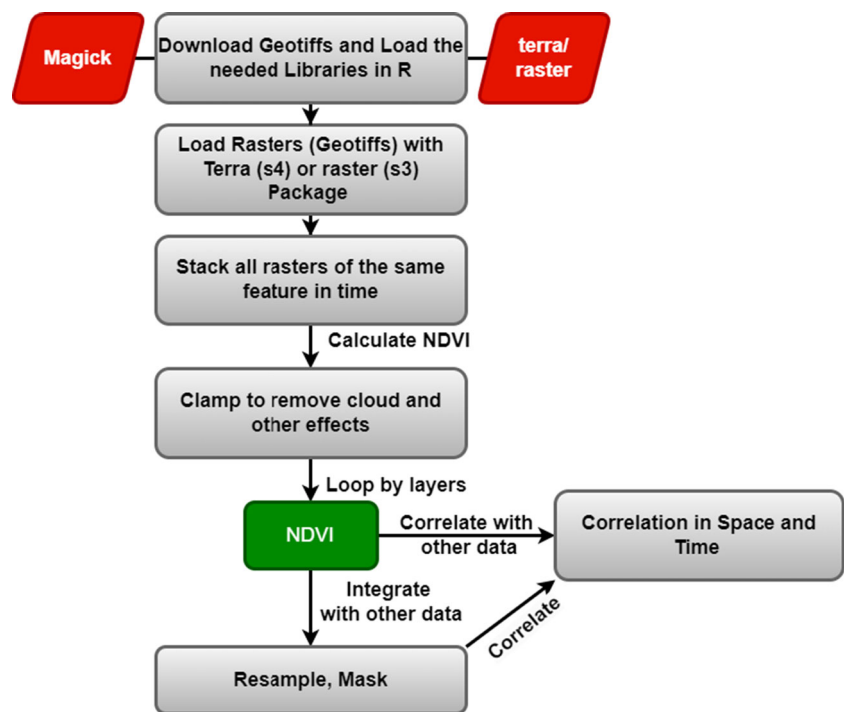
Figure 8 presents a visual representation of the key steps involved in integrating and analyzing crop and soil data spatially and temporarily.

Temporal and Spatial Integration The data have been synchronized based on their temporal and spatial dimensions to facilitate a comprehensive through-time analysis. By sequentially overlaying the data based on their distinct features, this study provides a uniform assessment of non-stationary attributes.

Cloud and Noise Removal To enhance the reliability of the data, a “clamp” process is applied to mitigate the impact of cloud cover and other potential noise effects present in the satellite imagery.

Iterative Analysis The methodology adopts an iterative approach, analyzing data layer by layer to extract crucial

Fig. 8 Schematic Overview of the methodology for integrating and analyzing satellite imagery and soil data across space and time using R packages



information from each piece of data while ensuring the overall completeness of the study.

Integration with Additional Data Emphasizing a holistic approach, the geospatial data is seamlessly integrated with other relevant data. This integration involves employing resampling and masking procedures to ensure spatial alignment and coherent correlation analysis.

Correlation Analysis The methodology conducts spatial and temporal correlation analysis to explore the intricate relationships between soil properties, crop growth patterns, and nutrient distribution. Established indices like the Normalized Difference Vegetation Index (NDVI) are used to assess vegetation dynamics and their connection to soil attributes.

Robustness The study employs distributed and High-Performance Computing (HPC) to store and analyze large amounts of data, ensuring efficiency and reliability in the computational process. Specifically, the added parallel processing capability facilitates working with data-intensive time series analysis.

Limitations While the **raster** or **terra** package enables efficient analysis of geospatial data, some limitations exist when working with extensive raster time series, as with the daily MODIS NDVI data. Conducting computations on the full 365 daily NDVI layers requires substantial memory to store and analyze the data. Even with HPC resources, mem-

ory constraints may necessitate subsetting or aggregating the data for certain processing steps. Additionally, processing the massive time series is computationally intensive, often taking hours to run certain workflows. Visualizing and exploring the outputs also poses challenges because of the large file sizes. This highlights the computational and visualization challenges of big data analysis using R packages on a local computer.

The improved methodology establishes a clear and rigorous framework for conducting geospatiotemporal analysis, integrating advanced tools and approaches to derive valuable insights from satellite and soil data. The transparency and attention to detail in the methodology contribute to the credibility and reproducibility of the research outcomes.

Work Flow

The research workflow, as depicted in Fig. 9, guided the entire process from data query and downloads to the final results for the soil-related data. The initial steps involved querying and downloading geospatial data from IBM EIS to our D/HPC CRADLE, encompassing data for Ohio during the year 2019. The subsequent steps included data preprocessing, vegetation classification, and crop healthiness mapping using the calculated NDVI. Correlation analysis with soil properties was also conducted. The final outputs comprise GIFs that provide valuable insights into vegetation patterns,

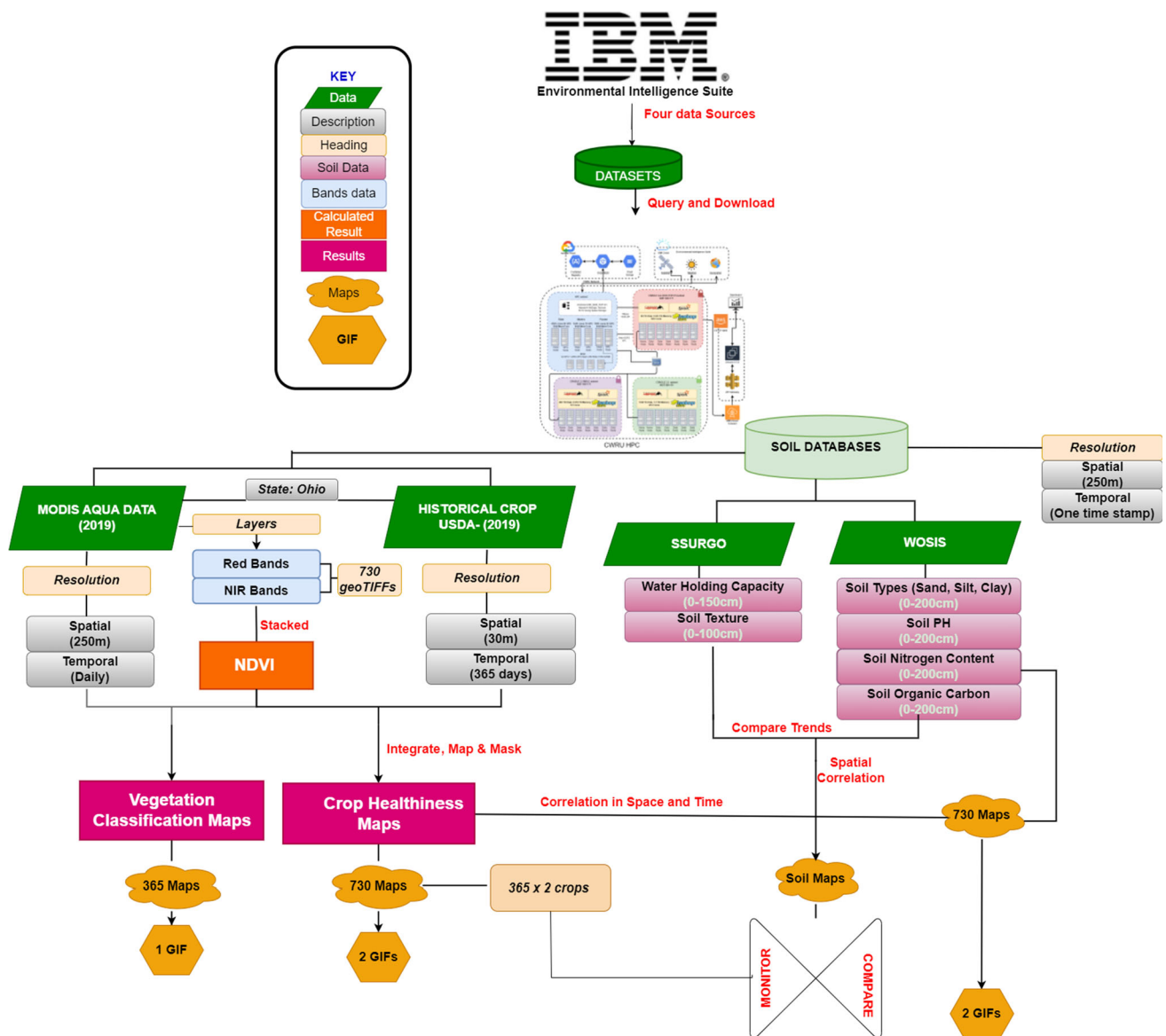


Fig. 9 Detailed workflow depicting the pipeline from initial data query and download to generating final results

crop healthiness, and correlations with soil attributes across the studied region and during time periods.

For the water analysis, the initial steps involved querying tabular and raster data from USGS, WQP, and NASA/METI to be stored in CRADLE D/HPC. Subsequent steps included data preprocessing and extracting features of interest, including stream networks, discharge, and nitrate and nitrite information. Nitrate and nitrite information was classified into levels to better understand its effect on the environment. Plots were created to help further understand the behavior of discharge nitrate and nitrite.

Results

Vegetation Classification

Figures 10, 11, 12, and 13 display seasonal NDVI maps for Ohio, selected from 365 daily NDVI maps generated in 2019. The maps showcase the seasonal variations in vegetation greenness across the state, with each map showing winter, spring, summer, and fall. The seasonal NDVI maps reveal a distinct shift in vegetation density between seasons.

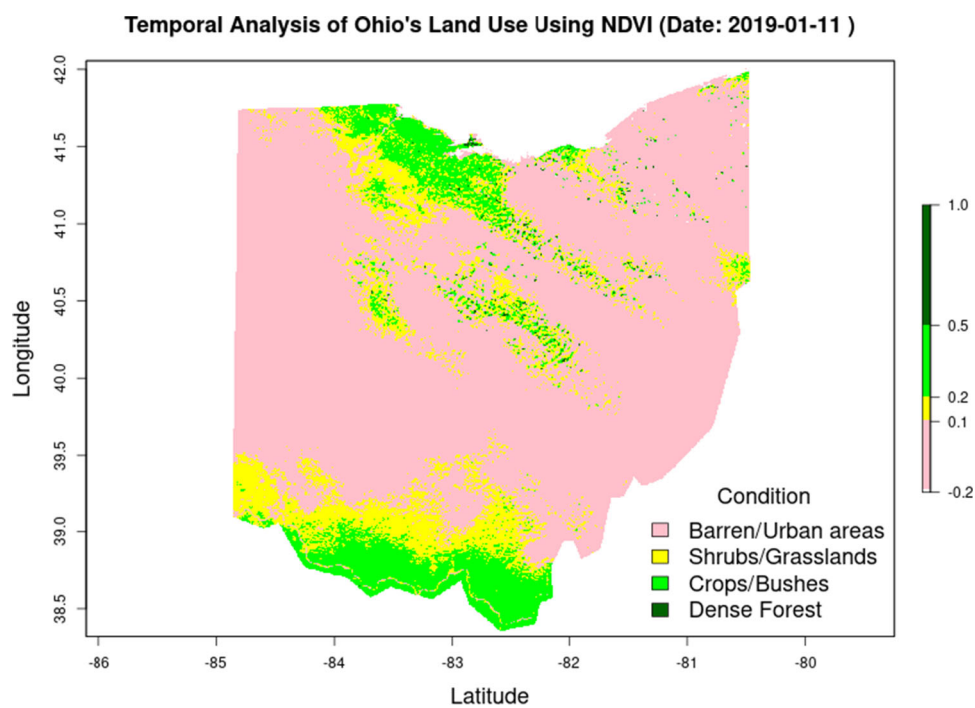


Fig. 10 Winter: Seasonal vegetation density classification map based on MODIS NDVI analysis for Ohio in 2019

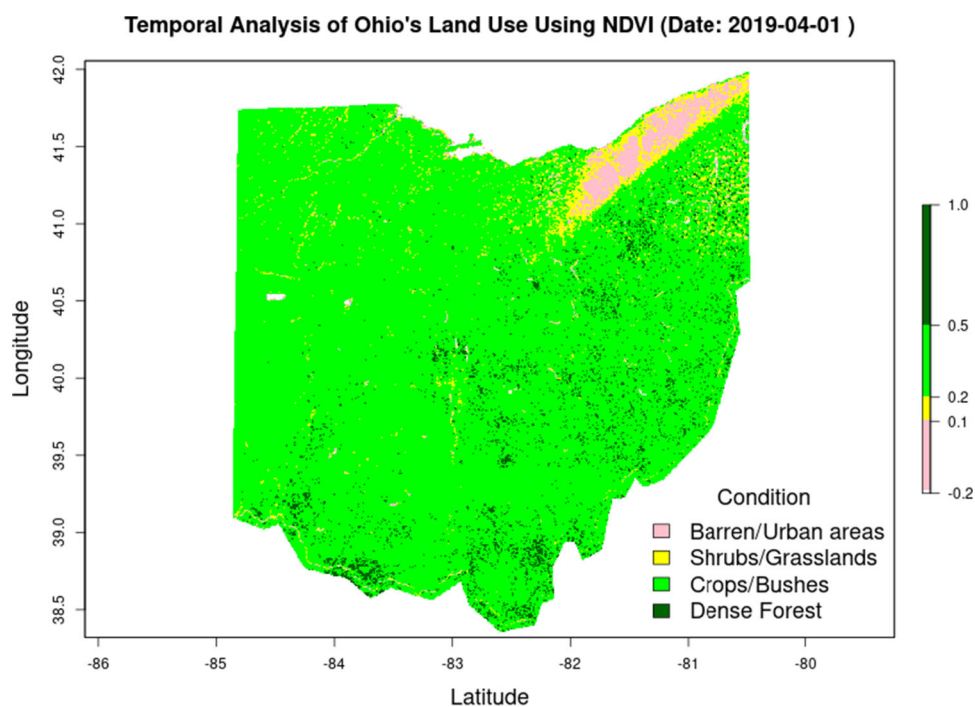


Fig. 11 Spring: Seasonal vegetation density classification map based on MODIS NDVI analysis for Ohio in 2019

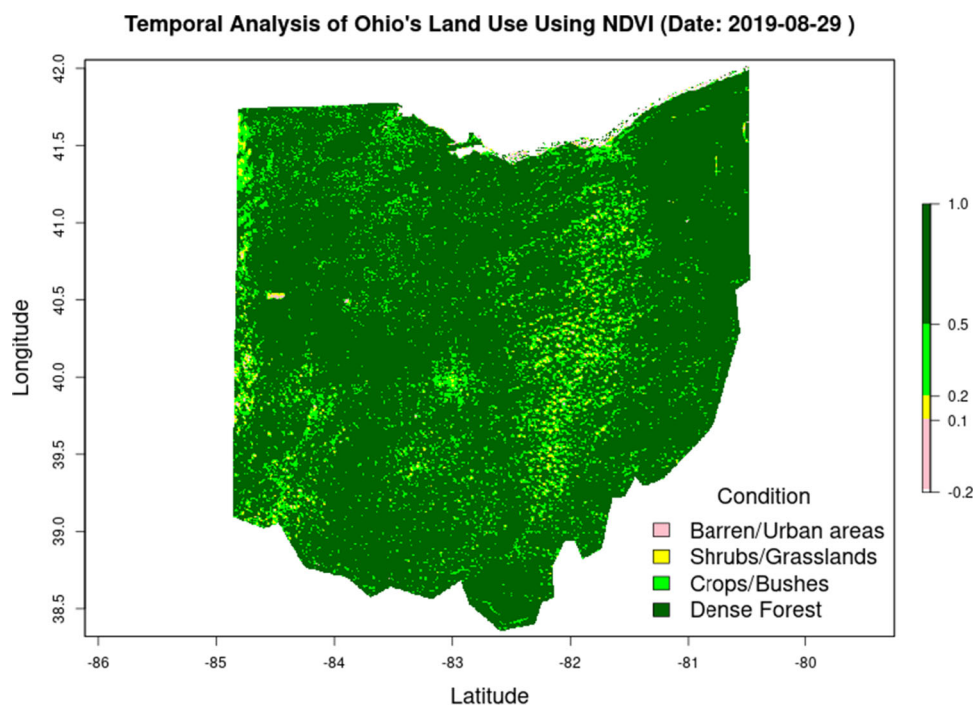


Fig. 12 Summer: Seasonal vegetation density classification map based on MODIS NDVI analysis for Ohio in 2019

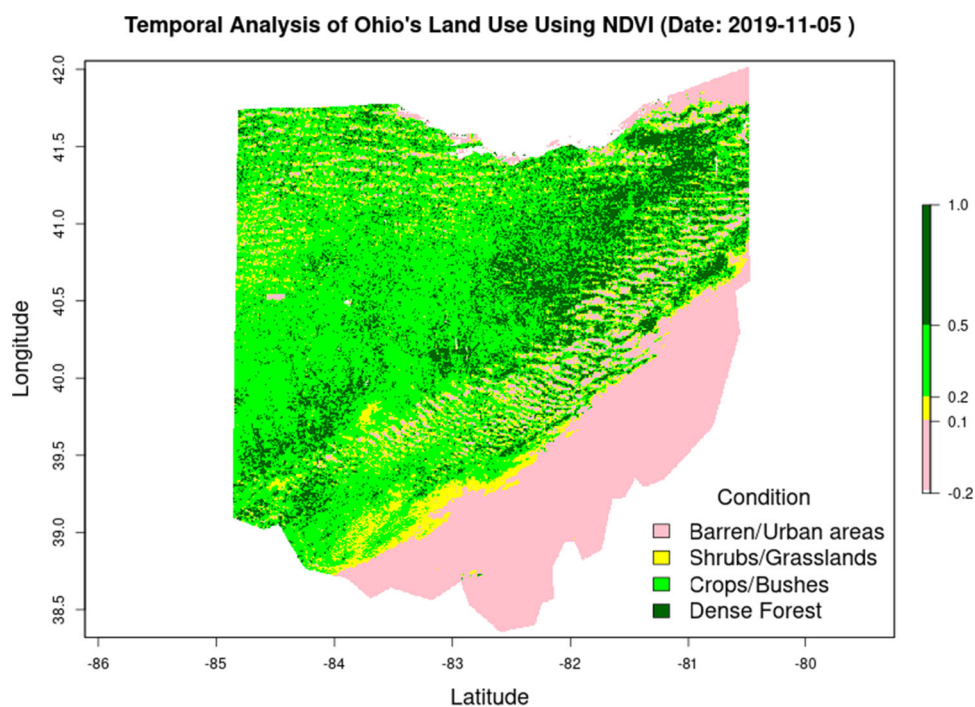


Fig. 13 Fall: Seasonal vegetation density classification map based on MODIS NDVI analysis for Ohio in 2019

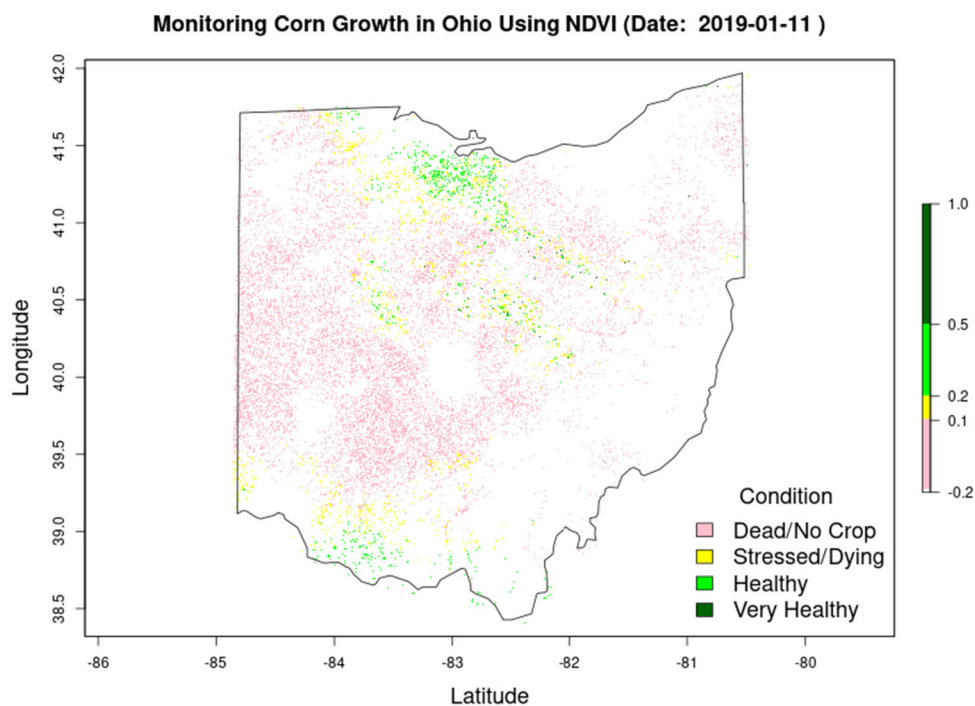


Fig. 14 Winter: Crop-specific monitoring of corn growth in Ohio during 2019 using MODIS NDVI time series data

Crop Monitoring

Figures 14, 15, 16, 17, 18, 19, 20, 21, and 22 demonstrate the ability of MODIS NDVI time series data to distinguish

the seasonal growth patterns of corn and soybeans in Ohio. The NDVI maps shown for each crop are selected from 365 daily NDVI maps generated in 2019, with one map displayed

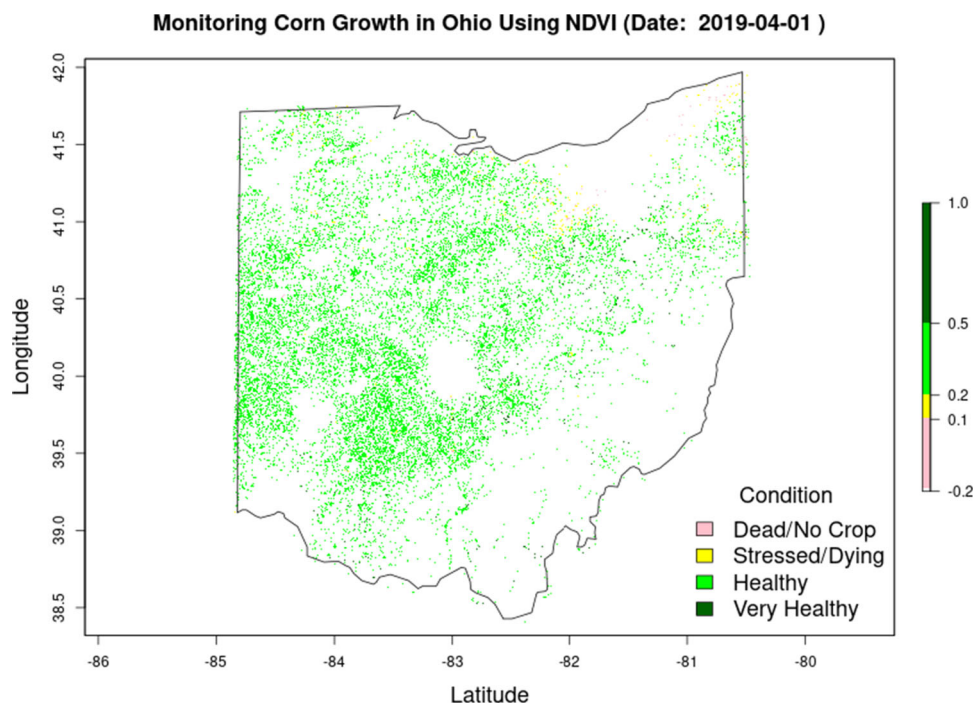


Fig. 15 Spring: Crop-specific monitoring of corn growth in Ohio during 2019 using MODIS NDVI time series data

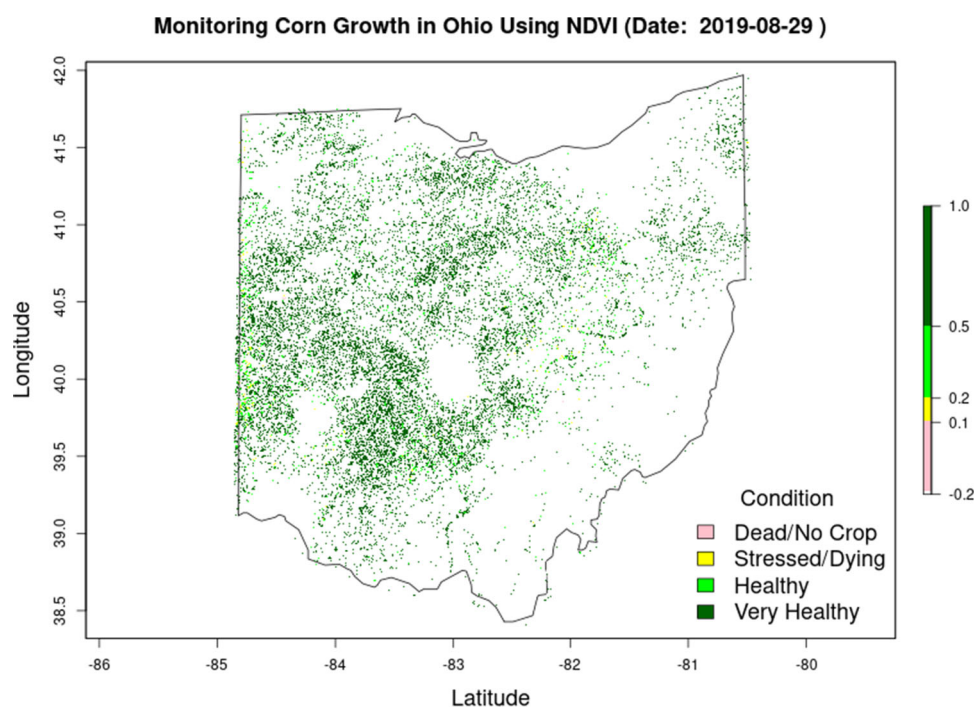


Fig. 16 Summer: Crop-specific monitoring of corn growth in Ohio during 2019 using MODIS NDVI time series data

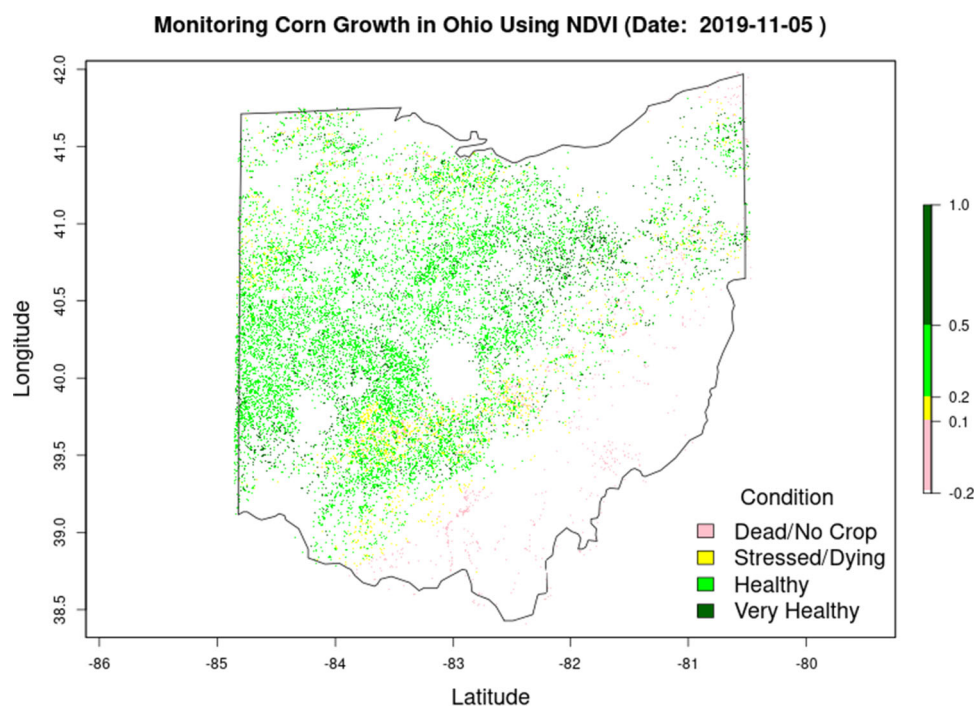


Fig. 17 Fall: Crop-specific monitoring of corn growth in Ohio during 2019 using MODIS NDVI time series data

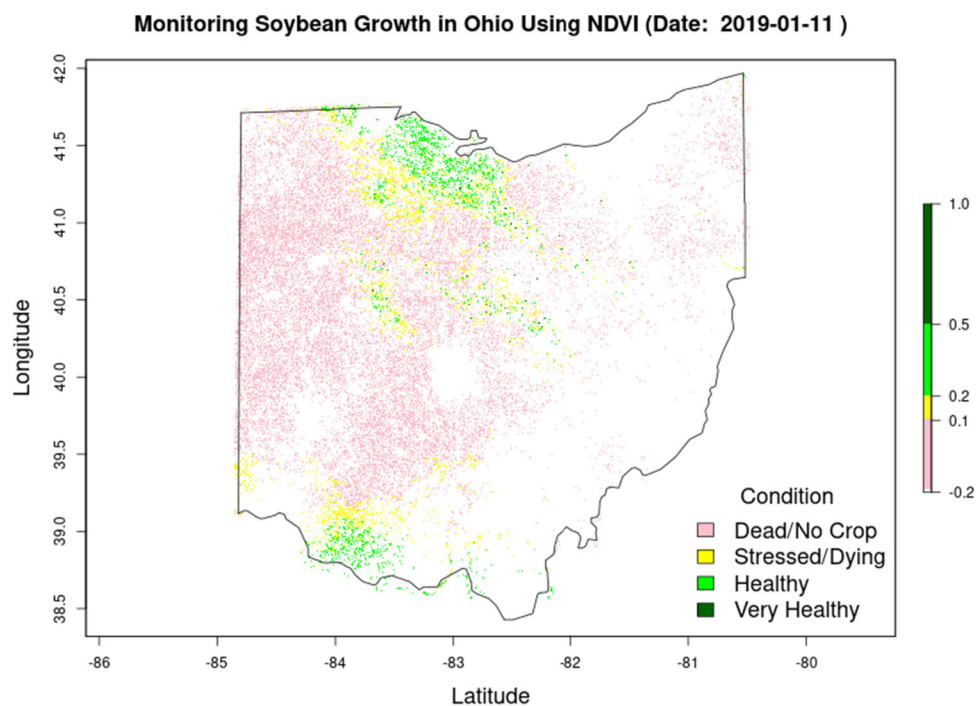


Fig. 18 Winter: Crop-specific monitoring of soybean growth in Ohio during 2019 using MODIS NDVI time series data

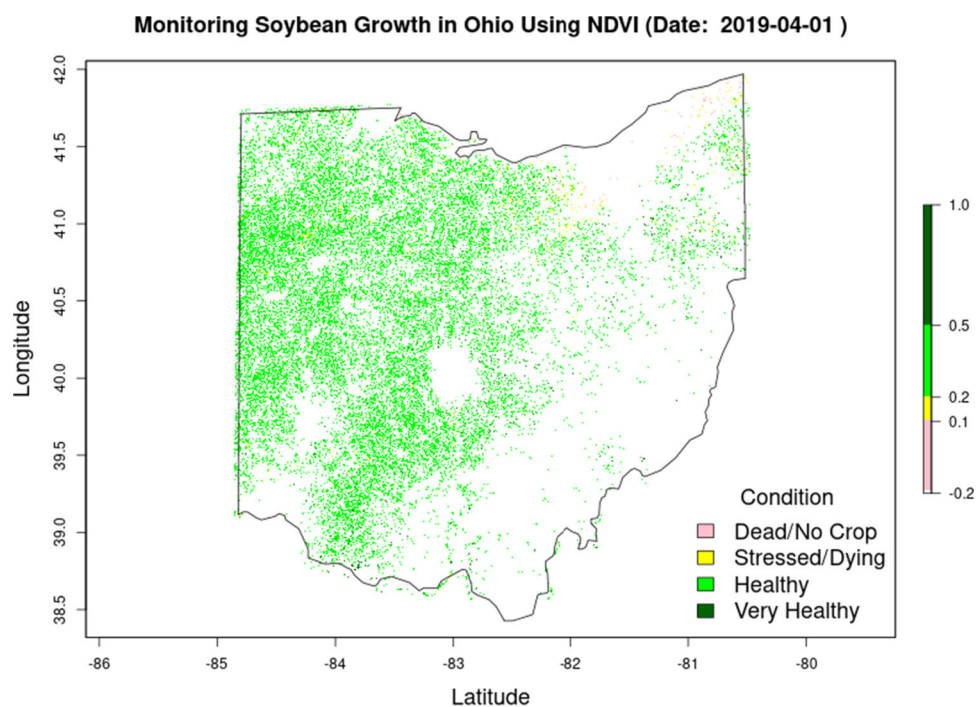


Fig. 19 Spring: Crop-specific monitoring of soybean growth in Ohio during 2019 using MODIS NDVI time series data

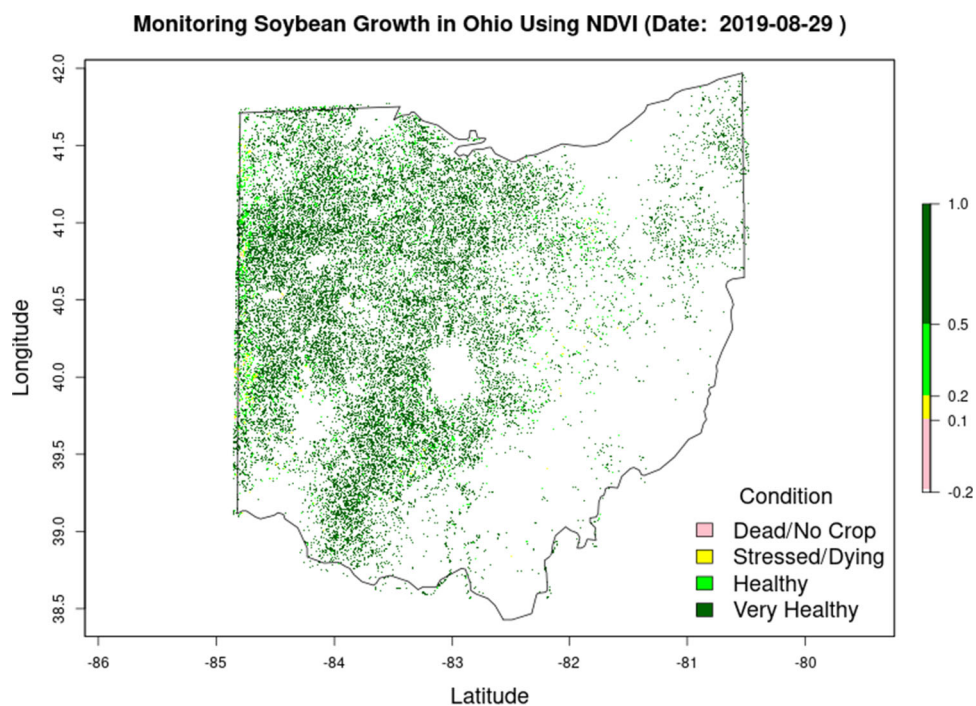


Fig. 20 Summer: Crop-specific monitoring of soybean growth in Ohio during 2019 using MODIS NDVI time series data

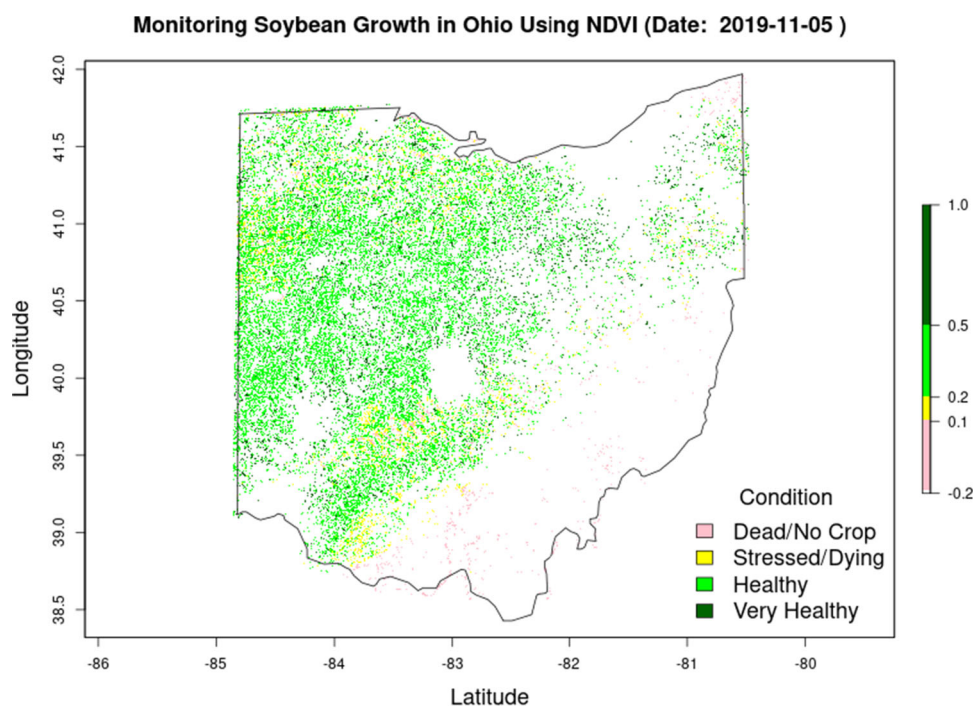
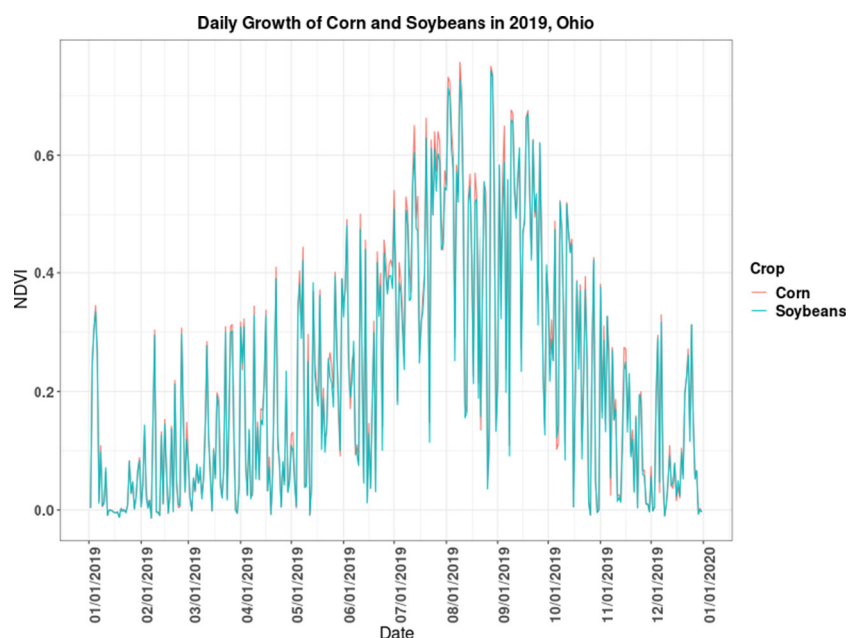


Fig. 21 Fall: Crop-specific monitoring of soybean growth in Ohio during 2019 using MODIS NDVI time series data

Fig. 22 Average seasonal growth profiles of corn and soybean in Ohio during 2019 derived from MODIS NDVI

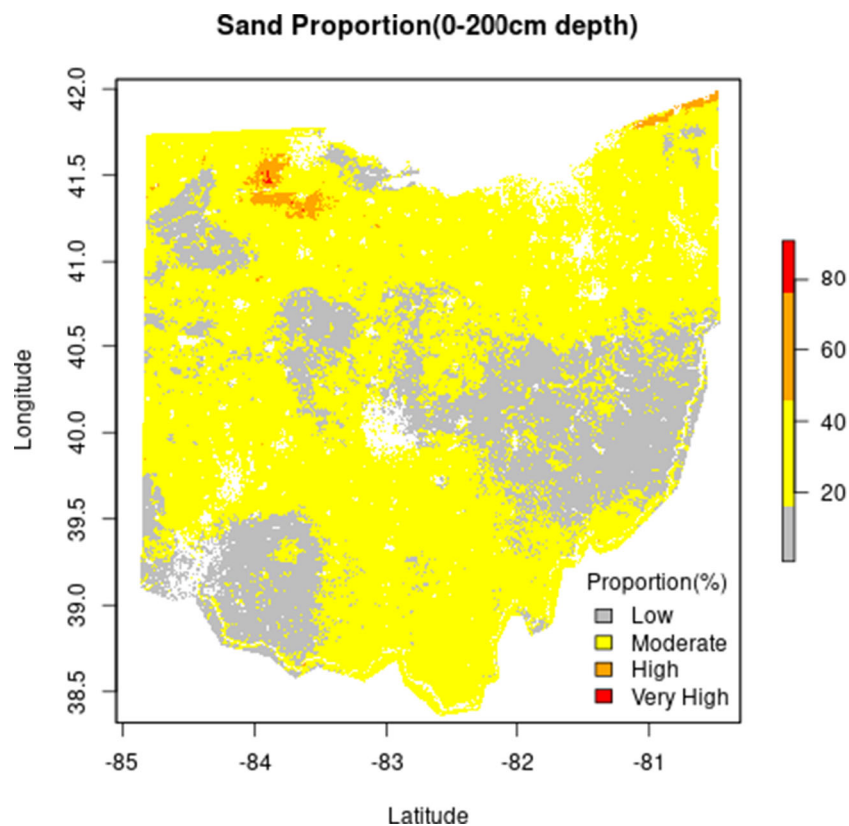


per season. The profiles reveal similar growth trajectories for corn and soybeans, peaking in the summer.

Soil Analysis

- Soil Types

Fig. 23 Sand proportion in Ohio at 0–200 cm depth



Figures 23, 24, 25, and 26 verify higher silt/clay soil content in northwest Ohio compared to predominantly sandy fractions in other parts of the state, as revealed by the soil type classification maps.

- Soil Physical Properties

Fig. 24 Silt proportion in Ohio at 0–200 cm depth

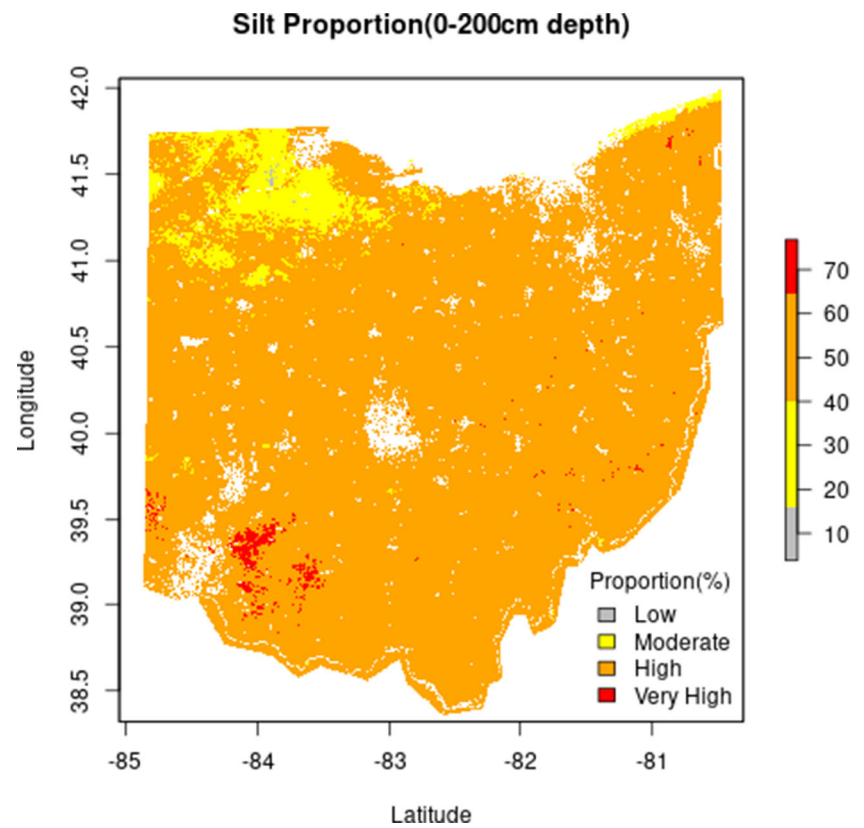


Fig. 25 Clay proportion in Ohio at 0–200 cm depth

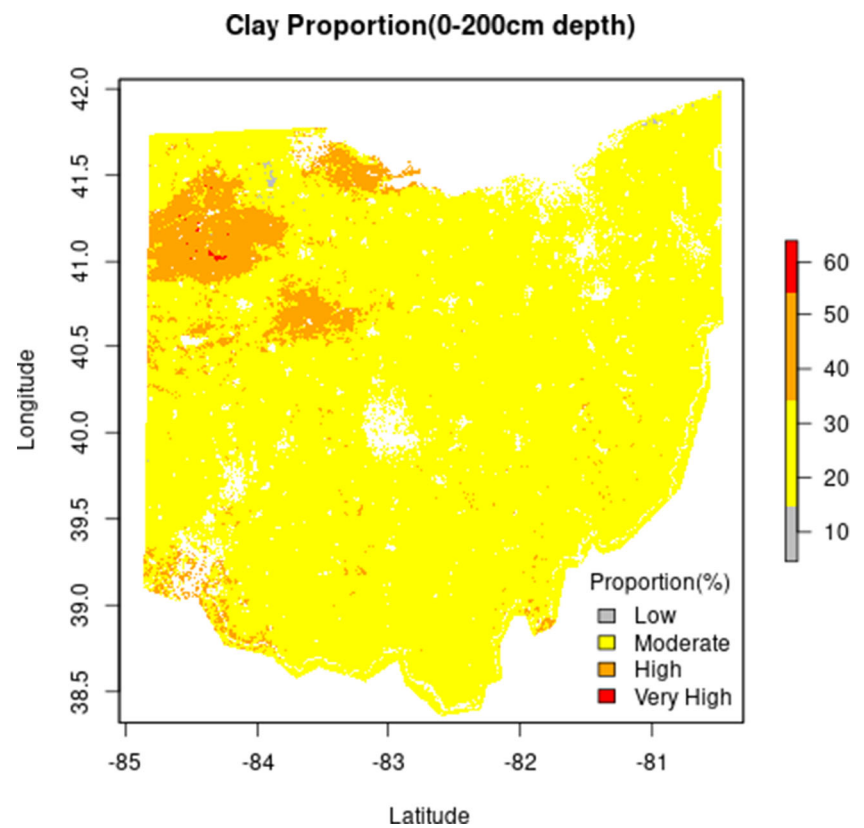
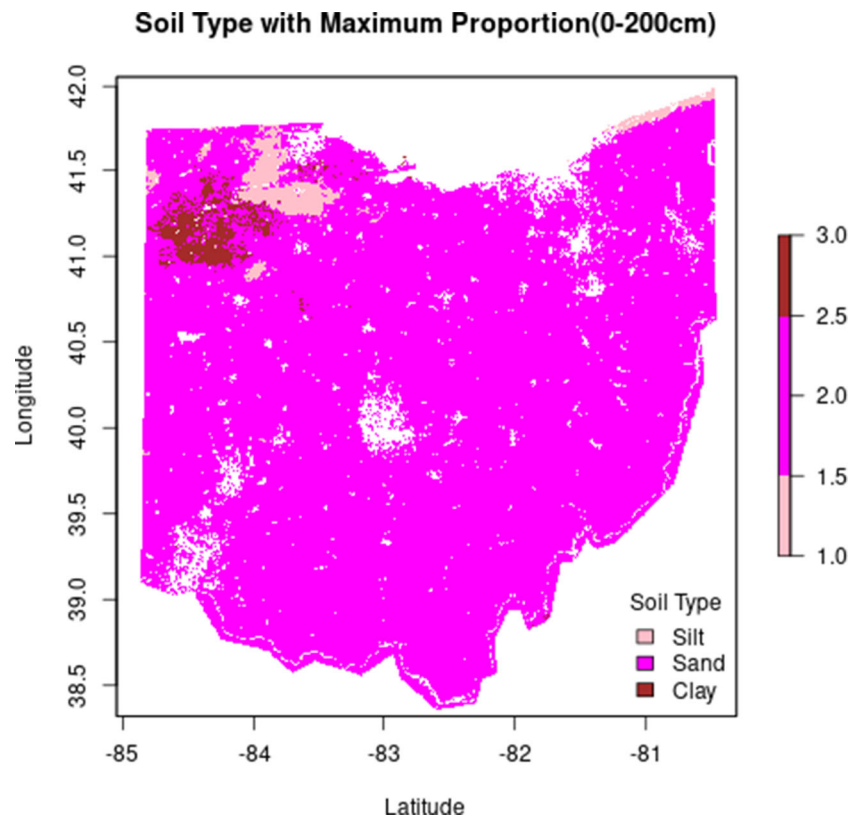


Fig. 26 Maximum occurrence of sand, silt, or clay in Ohio at 0–200 cm depth



The maps in Figs. 27 and 28 showcase the soil physical properties of water holding capacity and texture in Ohio. Lower moisture storage is evident in the northwestern region, while moderate capacity is seen elsewhere. Soil texture appears finer towards Lake Erie and coarser in southern Ohio.

- **Soil Chemical Properties**

Figures 29, 31, and 30 show the maps of soil pH, organic carbon, and nitrogen levels in Ohio, revealing distinct spatial patterns. Soil pH is higher in northwestern Ohio and lower in the southeast. Organic carbon content is elevated in eastern Ohio. Soil nitrogen distributions match vegetation density, with higher levels in northwestern Ohio near Lake Erie and the Maumee watershed.

Spatial Correlation

- **Temporal Correlation**

In Figs. 32 and 33, the spatiotemporal correlation maps between soil nitrogen and crop NDVI are presented, with one map selected per crop from 365 daily NDVI correlation maps generated in 2019. The maps showcase positive relationships between soil nitrogen content and crop health and greenness for corn and soybeans in Ohio.

- **One Time-Stamp Correlation**

In Figs. 34 and 35, spatial correlation maps illustrate the relationships between soil nitrogen and factors, including

organic carbon and pH levels in Ohio at a single point in time. The organic carbon correlation analysis shows the expected positive association with soil nitrogen content, while the pH correlation is negative.

Water Feature Analysis

Figures 36, 37, 38, 39, 40, 41, 42, and 43 display the seasonal variation of different water features in Ohio in 2019. More specifically, nitrate, nitrite, and discharge were analyzed in relation to the stream networks extracted from the digital elevation model discussed in Section 2.2.

Concentrations of nitrate and nitrite were grouped into four categories: less than 4 mg/L representing the baseline amount of nitrate and nitrite expected in water (normal), 4–7 mg/L representing above-normal concentrations (high), 7–10 mg/L representing concentrations approaching toxic levels and may need close monitoring (warning), and greater than 10 mg/L representing toxic levels of nitrate and nitrite that can lead to health problems (toxic). A numerical summary of the nitrate and nitrite concentrations by season can be seen in Tables 1, 2, 3, and 4. Toxic levels of nitrate and nitrite were seen in winter, spring, and fall, with the most sites in the toxic range present in the fall in the lower west. Fall also had the highest concentrations of nitrate and nitrite at 22.99 mg/L (Table 4).

Fig. 27 Soil water holding capacity in Ohio at 0–150 cm depth

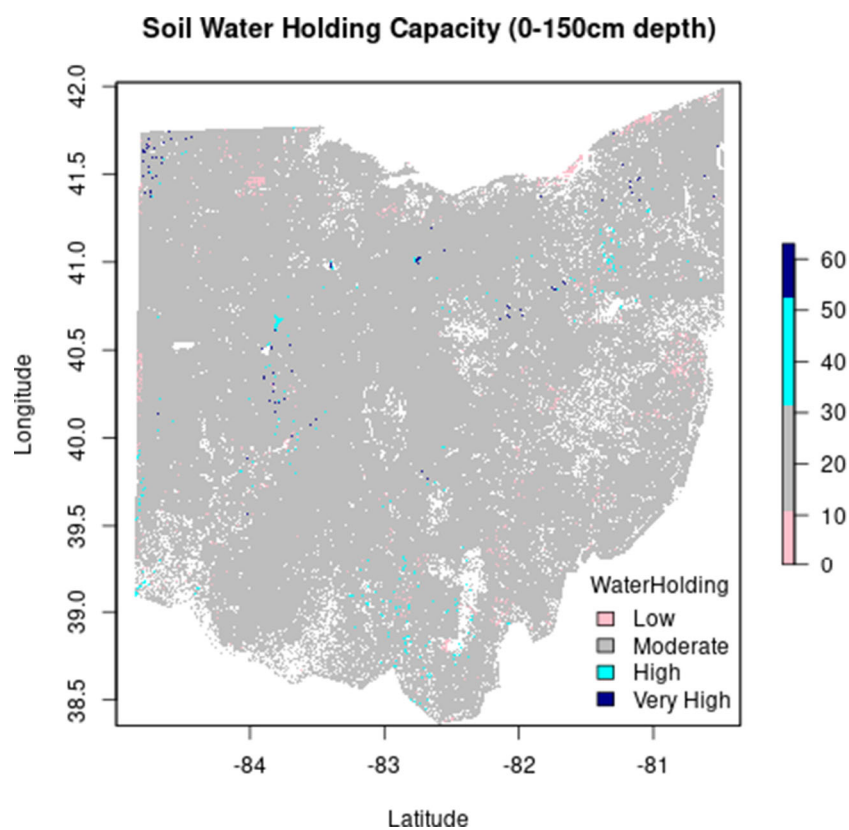


Fig. 28 Soil texture in Ohio at 0–100 cm depth

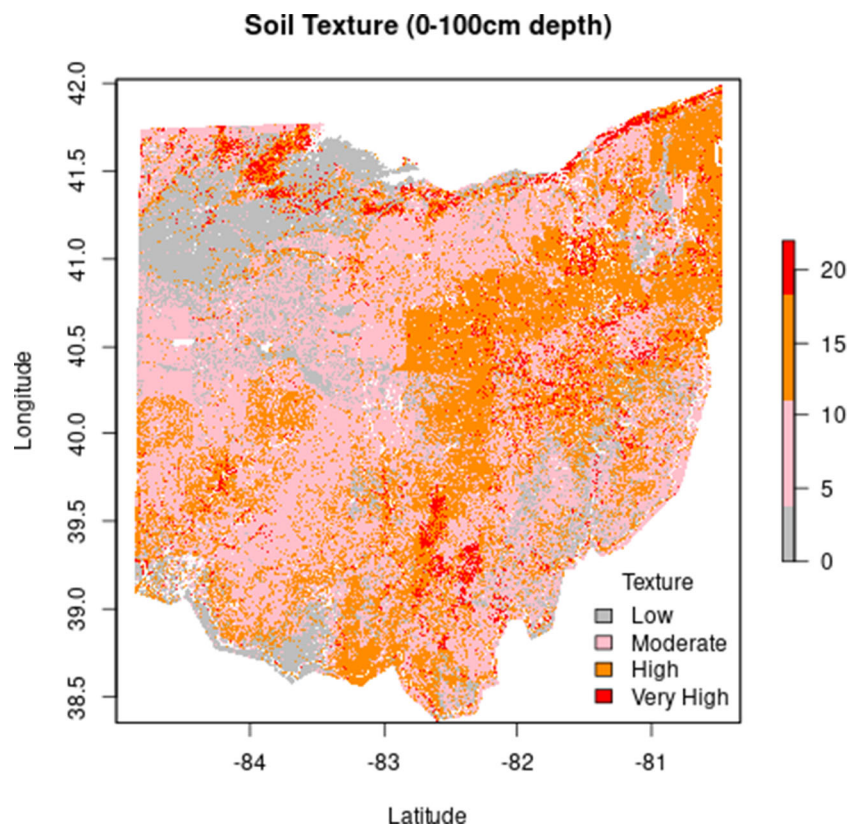


Fig. 29 Soil pH in Ohio at 0–200 cm depth

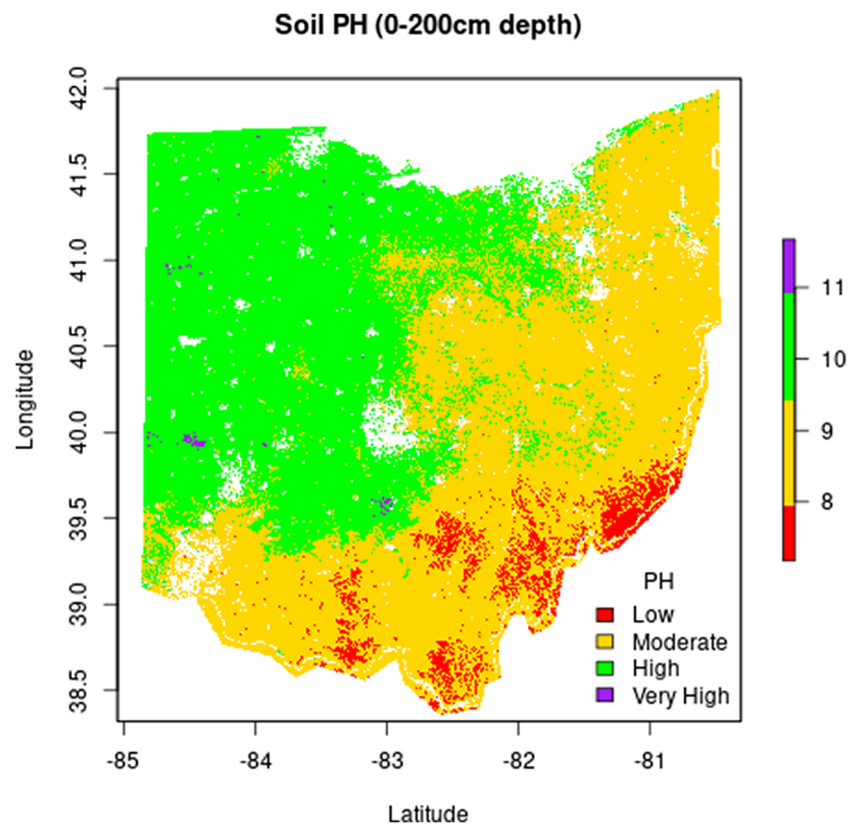


Fig. 30 Soil organic carbon content in Ohio at 0–200 cm depth

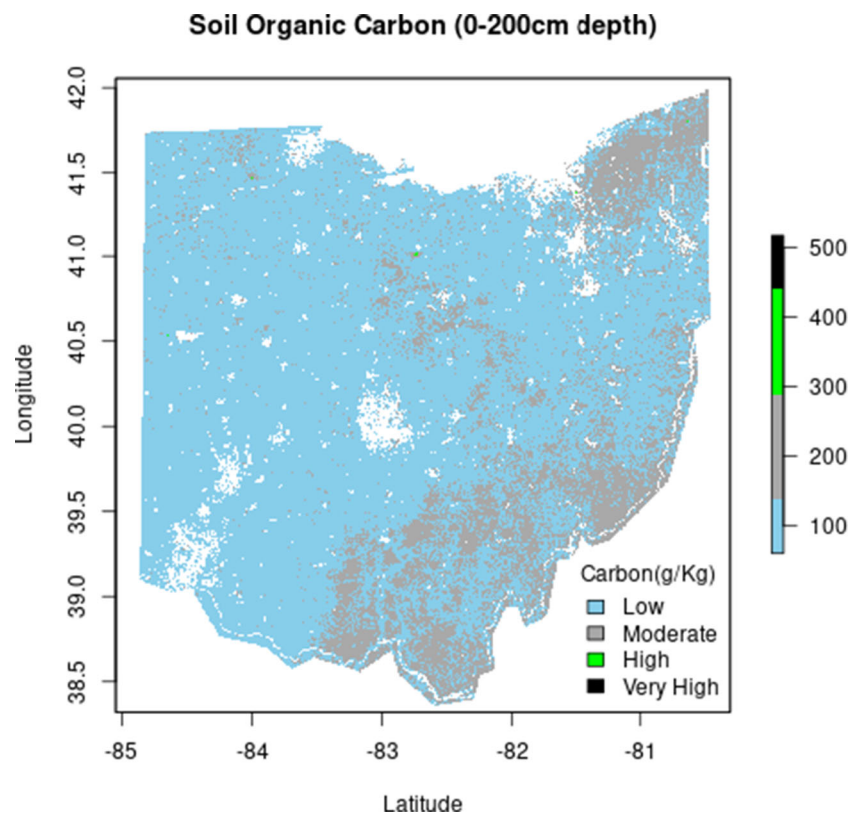


Fig. 31 Soil nitrogen content in Ohio at 0–200 cm depth

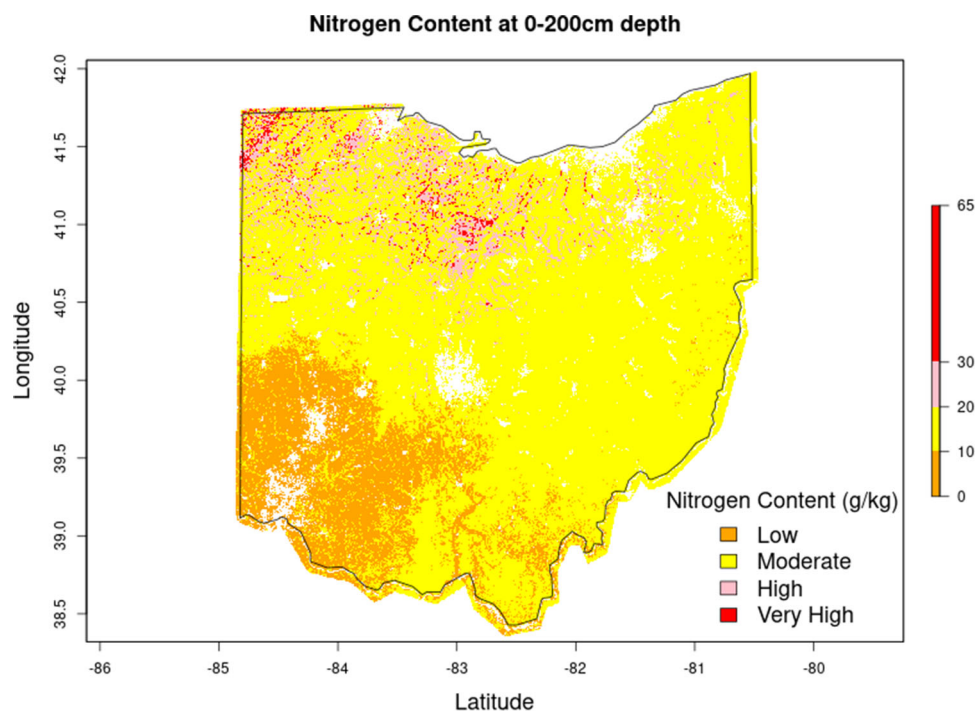


Fig. 32 Correlation of Soybeans NDVI with soil nitrogen in Ohio

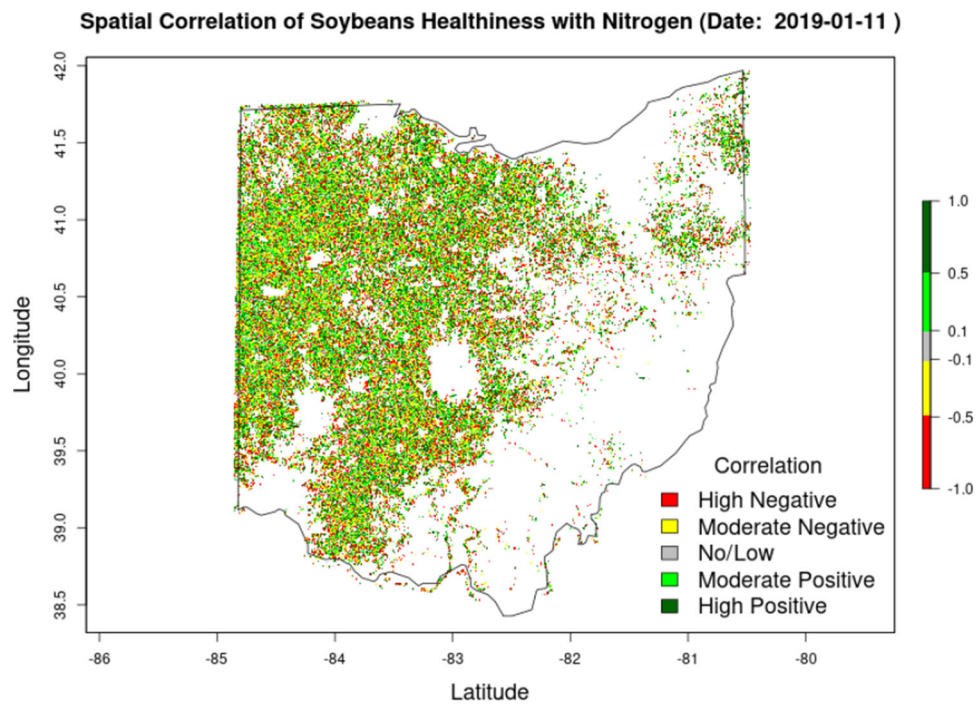


Fig. 33 Correlation of Corn NDVI with soil nitrogen in Ohio

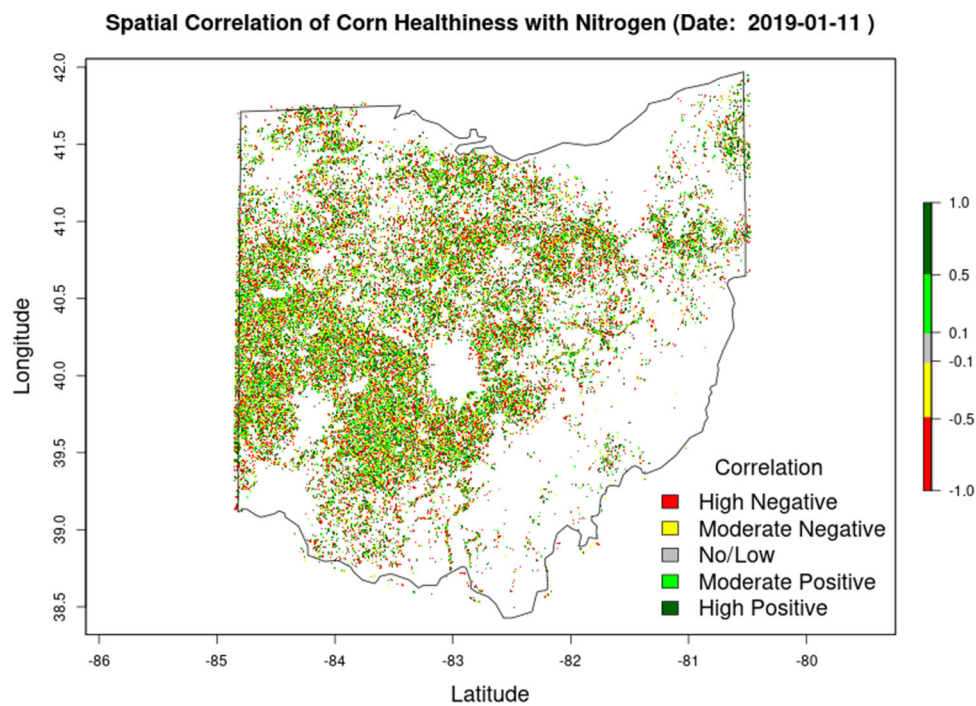


Fig. 34 Spatial correlation of soil organic carbon to soil nitrogen content in Ohio

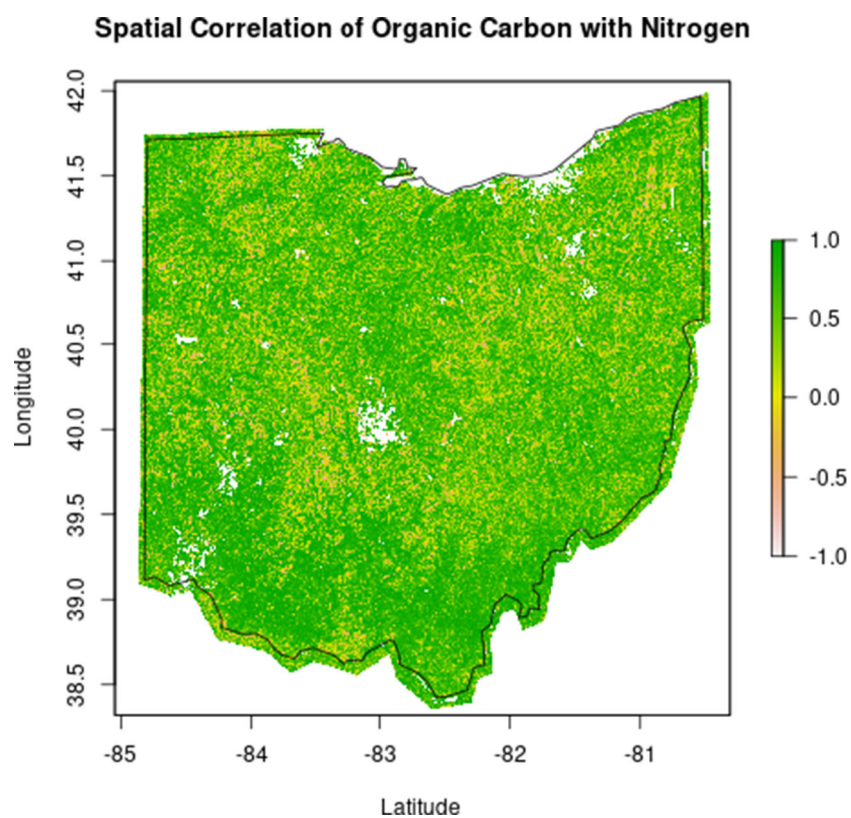


Fig. 35 Spatial correlation of soil pH to soil nitrogen content in Ohio

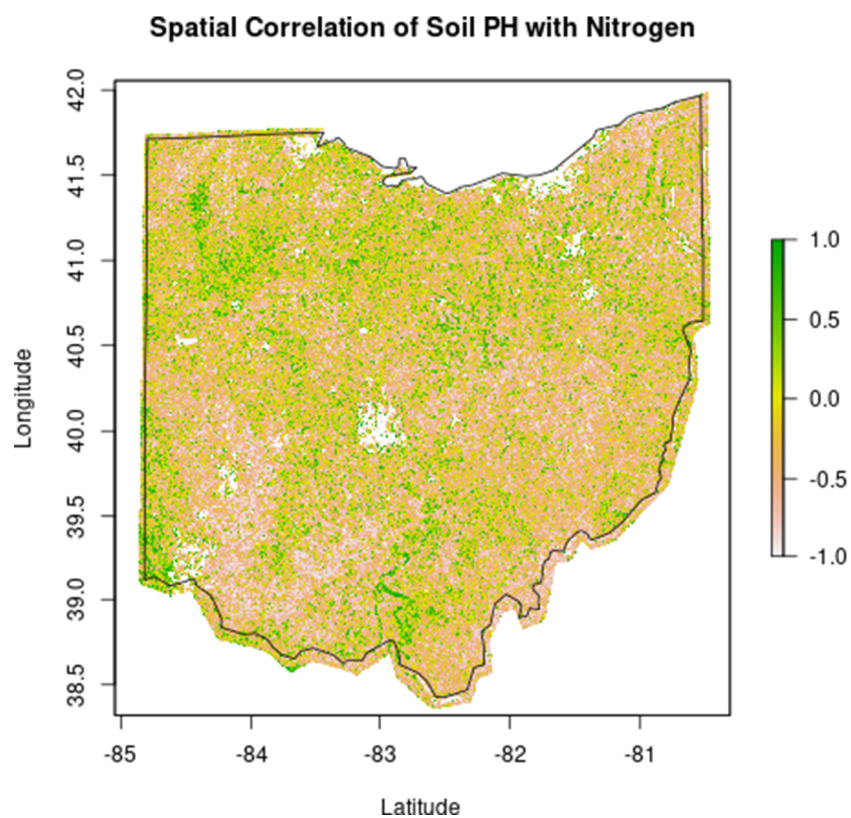


Fig. 36 Winter: Seasonal variation of nitrate and nitrite across Ohio in 2019

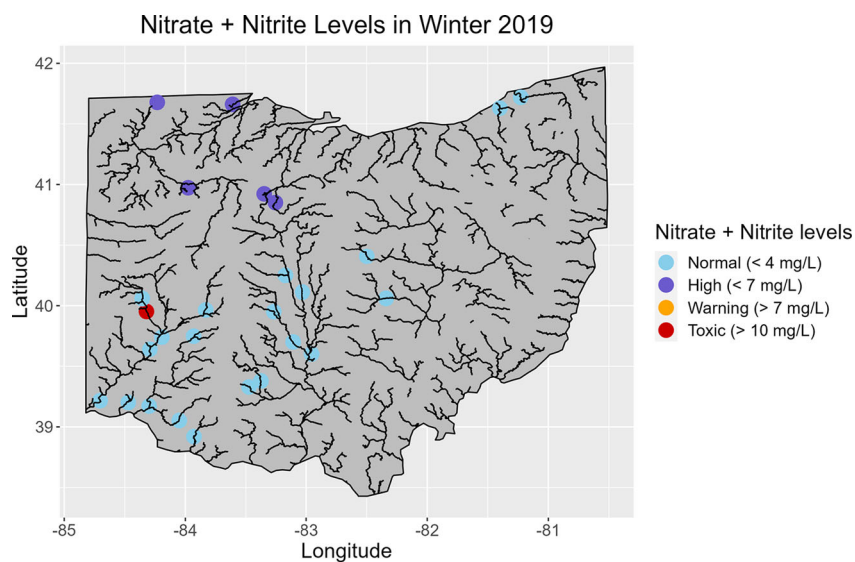


Fig. 37 Spring: Seasonal variation of nitrate and nitrite across Ohio in 2019

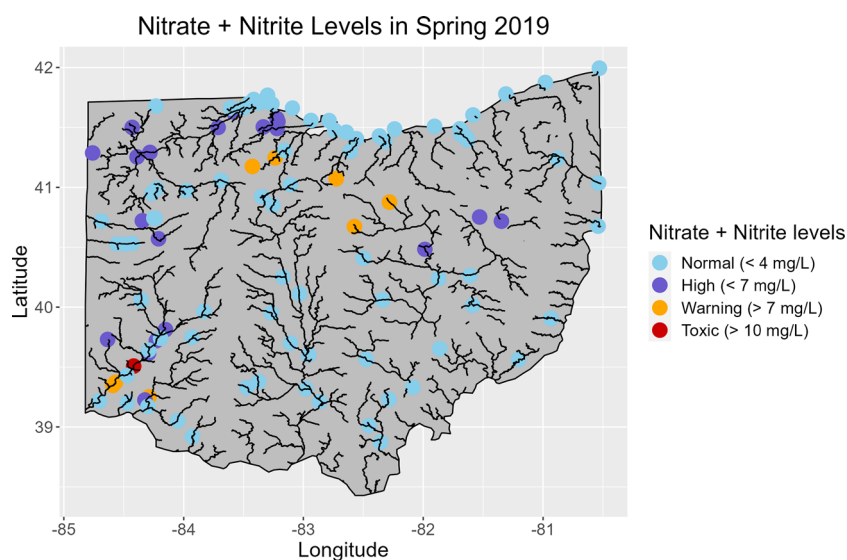


Fig. 38 Summer: Seasonal variation of nitrate and nitrite across Ohio in 2019

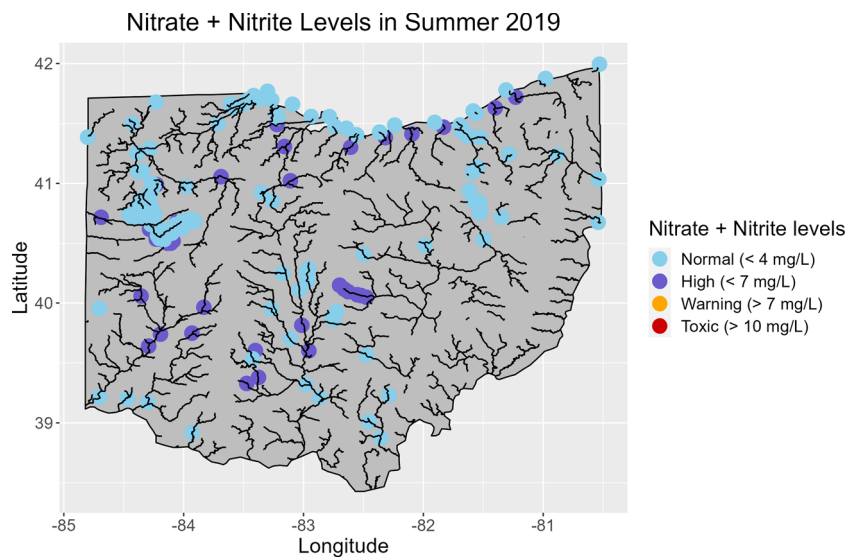


Fig. 39 Fall: Seasonal variation of nitrate and nitrite across Ohio in 2019

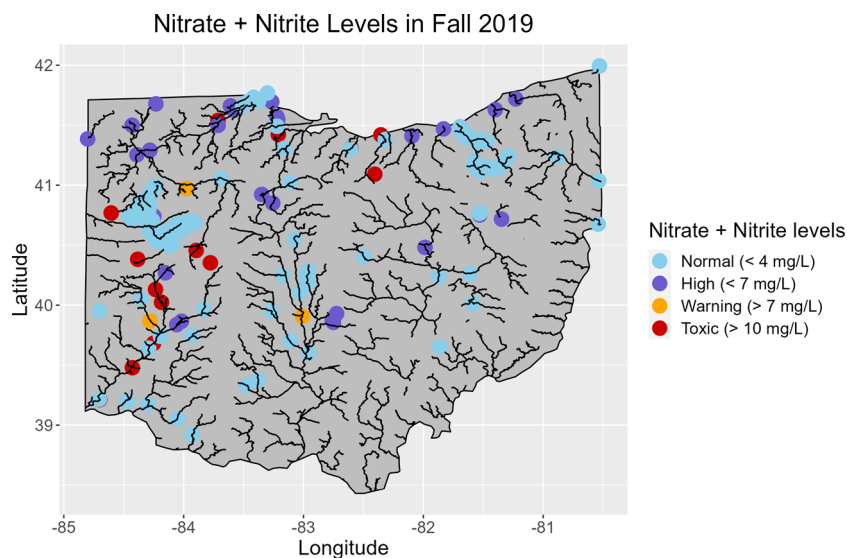


Fig. 40 Winter: Seasonal variation of discharge across Ohio in 2019

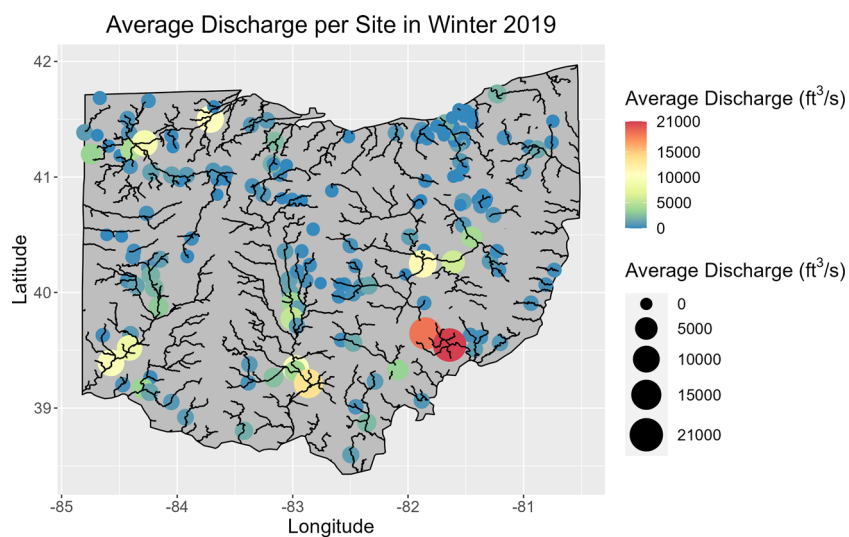


Fig. 41 Spring: Seasonal variation of discharge across Ohio in 2019

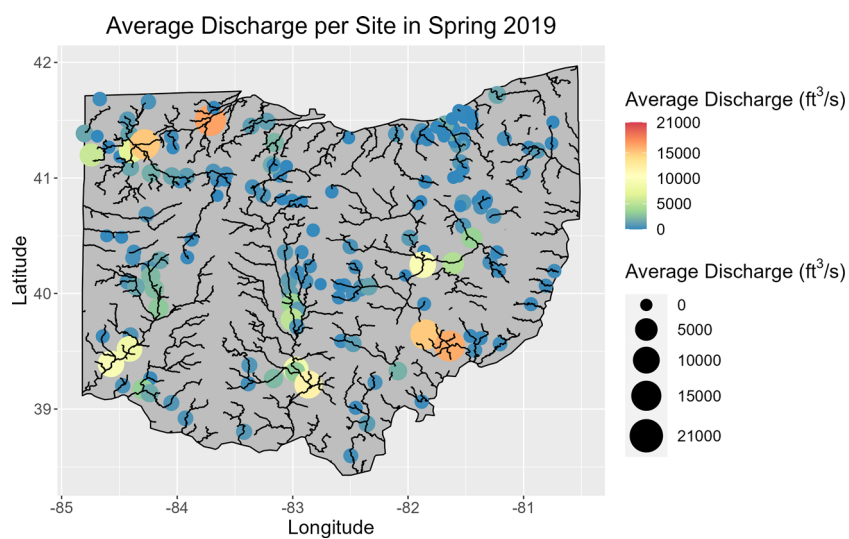


Fig. 42 Summer: Seasonal variation of discharge across Ohio in 2019

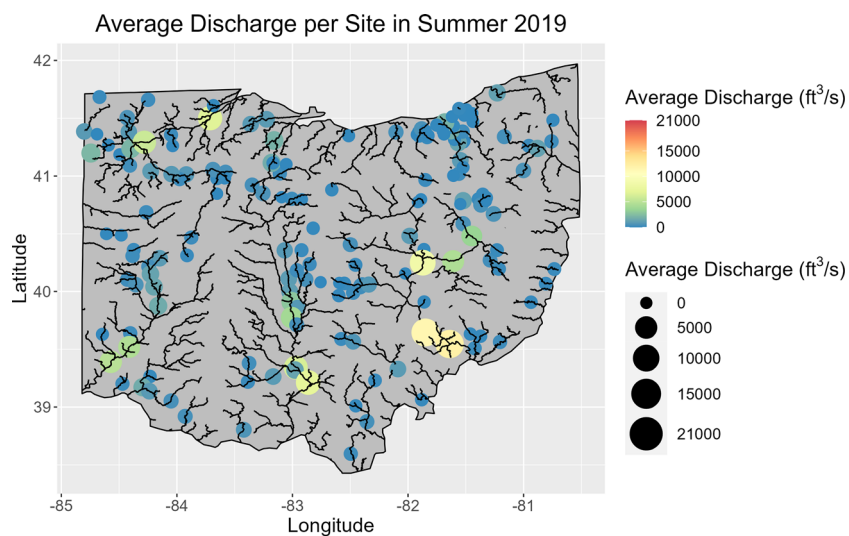


Fig. 43 Fall: Seasonal variation of discharge across Ohio in 2019

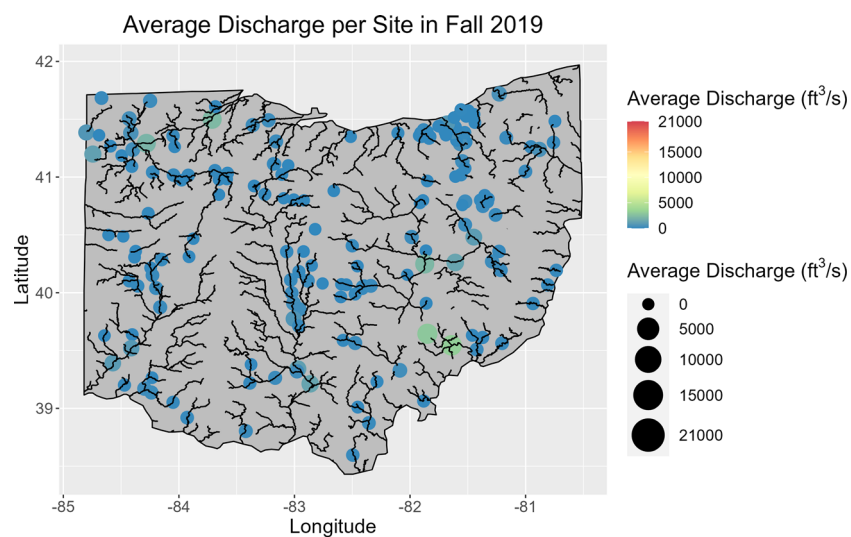


Fig. 44 Average discharge per day for Ohio in 2019, trend line displayed in blue

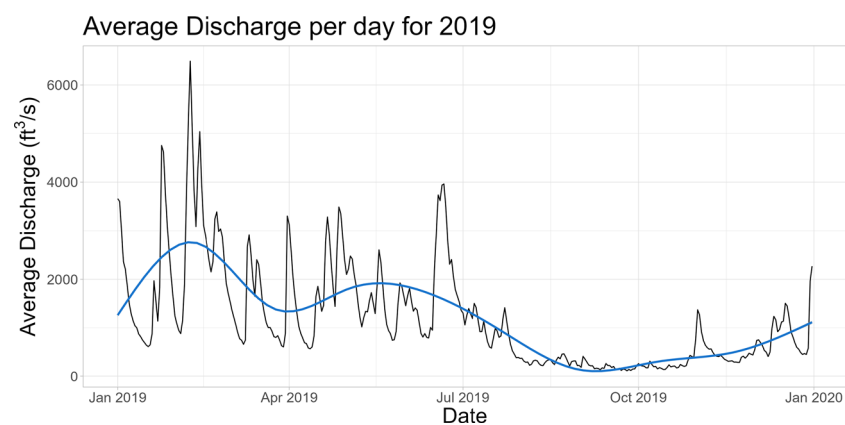


Fig. 45 Average nitrate and nitrite per day for Ohio in 2019, trend line displayed in green

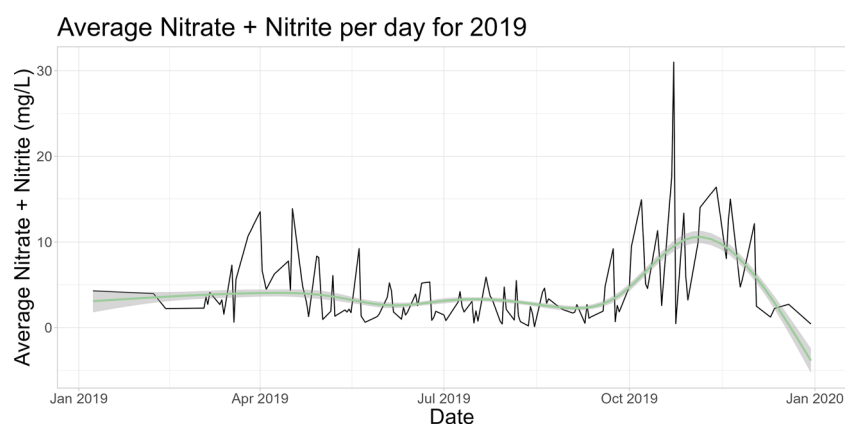


Table 1 Summary of nitrate and nitrite concentrations in Winter 2019

Level	Minimum	Mean	Maximum	n
Toxic (> 10 mg/L)	12.100	12.100	12.100	1
Warning (> 7 mg/L)	7.920	8.019	8.069	3
High (< 7 mg/L)	4.134	4.134	4.134	5
Normal (< 4 mg/L)	0.512	1.922	3.262	21

Table 2 Summary of nitrate and nitrite concentrations in Spring 2019

Level	Minimum	Mean	Maximum	n
Toxic (> 10 mg/L)	13.500	13.500	13.500	1
Warning (> 7 mg/L)	7.539	8.506	9.743	9
High (< 7 mg/L)	4.112	5.052	6.320	20
Normal (< 4 mg/L)	0.954	2.306	3.646	80

Table 3 Summary of nitrate and nitrite concentrations in Summer 2019

Level	Minimum	Mean	Maximum	n
Toxic (> 10 mg/L)	n/a	n/a	n/a	0
Warning (> 7 mg/L)	n/a	n/a	n/a	0
High (< 7 mg/L)	4.183	4.758	5.496	43
Normal (< 4 mg/L)	0.823	2.603	3.965	145

Table 4 Summary of nitrate and nitrite concentrations in Fall 2019

Level	Minimum	Mean	Maximum	n
Toxic (> 10 mg/L)	11.280	13.710	22.990	14
Warning (> 7 mg/L)	7.920	8.019	8.069	3
High (< 7 mg/L)	4.574	4.835	5.850	28
Normal (< 4 mg/L)	0.512	1.922	3.762	132

Table 5 Summary of discharge per season in Ohio 2019

Season	Minimum	Mean	Maximum	n
Winter	0	1307.170	20992.170	177
Spring	0.001	797.720	12086.871	177
Summer	0.001	797.522	12086.871	177
Fall	0	189.114	3176.768	117

Figures 36, 37, 38, 39, 40, 41, 42, and 43 also display discharge per season. Visually and based on Table 5, winter had the highest discharge, while fall had the lowest amount of discharge on average. Between winter, summer, and fall, the locations with the greatest discharge were in relatively the same locations, such as the southwestern and northeastern parts of Ohio.

Discharge had a bimodal distribution early in the year, with the highest peak occurring in February and the second peak occurring in May, according to Fig. 44, while nitrate and nitrite concentrations were stable during that period (Fig. 45). Nitrate and nitrite concentrations peaked later in the year, around late October, while discharge was stable during this time.

Discussion

Interpretation of Vegetation Classification Results

The vegetation classification results in Figs. 10, 11, 12, and 13 provide an informative baseline characterization of seasonal variations in vegetation density across Ohio in 2019 using MODIS NDVI. As noted by Huang et al. (2021), NDVI is widely used to examine vegetation patterns and dynamics across large regions. Aligning with previous studies DeFries and Townshend (1994); Pettorelli et al. (2005), our results show NDVI's capability to capture phenological shifts corresponding to the state's seasonal weather and temperature patterns.

The lower winter NDVI values in Ohio likely reflect colder temperatures inhibiting plant growth, consistent with prior work showing sparser canopy cover in temperate regions during cold periods Chen et al. (2004); Jenkins et al. (2007). Overall, these baseline vegetation maps validate the use of MODIS NDVI time series for delineating regional and seasonal distinctions in vegetation density and phenology, consistent with previous studies DeFries and Townshend (1994); Dragoni and Rahman (2012). The results establish an informative foundation for future crop monitoring, climate change, or land cover change analysis across these major agricultural states.

Implications of Crop Monitoring Results

The crop-specific temporal monitoring results in Sect. 4.2 and Figs. 14, 15, 16, 17, 18, 19, 20, 21, and 22 provide valuable insights into the growth profiles and phenological patterns of key crops in Ohio during 2019. The distinct seasonal NDVI signatures exhibited by corn and soybean in Ohio (Figs. 14, 15, 16, 17, 18, 19, 20, and 21) demonstrate MODIS data's potential for capturing crop-specific growth stages and phenology, an essential application as noted by prior research

Dragoni et al. (2011); Lu and Zhuang (2010). Corn displayed vigorous NDVI growth from June through September, aligning with its typical summer cultivation period Suyker and Verma (2012).

Soybeans also showed more moderate NDVI increases during June–September, corresponding to their growth cycle Setiyono et al. (2007); Bradley et al. (2007). Studies have validated the use of MODIS NDVI for monitoring corn and soybean development across the Midwest, including Ohio Sakamoto et al. (2010); Zhong et al. (2016); Wardlow et al. (2007). The higher winter crop NDVI in northern Ohio likely reflects the concentrated summer growing period enabled by a shorter temperate season, while the warmer climate in southern Ohio allows for a longer cultivation period and summer crop planting Hatfield et al. (2014). Regional variations in climate and growing season length underpin these geographic crop patterns within the state.

In conclusion, the white regions evident in the crop monitoring maps show locations where those particular crops are not grown, aligning with known regional agricultural distributions. The crop monitoring results across Ohio underscore MODIS NDVI data's capability to capture regional variations in crop phenology and development patterns Ko et al. (2009); Piccinni et al. (2009); Abedinpour (2015). The distinct temporal signatures enable differentiation between crops within a state based on their geography-specific growing seasons and climate suitability.

These findings highlight the potential of MODIS time series for scalable vegetation monitoring and classification in agricultural regions. Applications could aid crop type mapping, growth stage modeling, and yield prediction at state and national scales to support food security assessments. However, validation with higher-resolution data would be required to translate these methods to field-level precision agriculture.

Significance of Soil Analysis Findings

• Soil Types

The soil type classification maps depicted in Figs. 23, 24, 25, and 26 validate and further explain the established patterns of soil texture variations throughout Ohio Subburayalu et al. (2014). Northwestern Ohio, with its pronounced silt and clay content, has fine-textured soils that trace back to their origins as glacial lacustrine deposits Conrey (1941); Easterly (1964).

Conversely, sandy soils dominate a significant portion of the state, a reflection of the coarser alluvial, glaciofluvial, and loess (a type of very fine-grain, windblown sediment) parent materials present in these regions. Prior work has mapped such geographic distinctions in soil texture at the state scale using digital soil data. Previous studies

have delineated these distinctions in soil texture at the state's scale using digital soil data Arshad et al. (1997); Mallah et al. (2022) and have underscored the pivotal role of soil composition in agricultural land management Lindbo et al. (2012). These maps provide an updated baseline understanding of soil textural patterns relevant to crop planning and other applications.

• Soil Physical Properties

The reduced water retention observed in sections of northwestern Ohio, as illustrated in Figs. 27 and 28, is because of the presence of coarser sandy soils. These soils are associated with the glacial beach ridge deposits found along Lake Erie Conrey (1941); Easterly (1964). Research shows that coarse-textured soils tend to have reduced moisture retention capabilities Subburayalu et al. (2014). Concurrently, the higher surface soil texture in this area is associated with a greater silt-clay content. This composition provides structural stability and fertility, which are crucial for agricultural practices Arshad et al. (1997); Extension (2023); Saxton and Rawls (2006). However, the medium-textured soils across much of Ohio provide favorable physical conditions for crop cultivation, as underscored by the state's ranking among national leaders in corn and soybean production USDA (2022).

• Soil Chemical Properties

The observed geographic patterns in key soil chemical properties, including nitrogen, organic carbon, and pH (Figs. 29, 30, and 31) agree with documented spatial distributions attributed to environmental factors. Soil nitrogen levels often correspond to fertilizer inputs and land management, with croplands showing elevated concentrations Jarecki and Lal (2005). Soil organic carbon relates to vegetation, parent material, and climate Jung and Lal (2011). Regional soil pH is influenced by native vegetation, drainage, and other soil-forming factors Beery and Wilding (1971).

Notably, northwestern Ohio exhibits high soil nitrogen content that implicates multiple interconnected factors. The intensive corn and soybean agriculture provides abundant nitrogen inputs, while flat topography and prevalent tile drainage facilitate nitrogen transport from fields Michalak et al. (2013); Smith et al. (2015). Coarser sandy soils with lower organic matter offer reduced natural nitrogen retention Heathwaite et al. (2000); Mohammadpour and Grady (2023). Freeze-thaw cycles may speed up nitrogen release and nitrate leaching Fouli et al. (2013). Climate patterns driving Lake Erie's hypoxia relate to soil nitrogen accumulation Bosch et al. (2013). The heightened levels of nitrogen observed in the Maumee River watershed hold a critical implication. They hint at the presence of significant nitrogen reservoirs that possess the potential to be mobilized, ultimately finding their way into the intricate network of rivers and,

ultimately, into Lake Erie Scavia et al. (2014). The relevance of this observation gains clarity when considering two crucial aspects of soil composition and behavior. First, the indication of lower organic carbon content carries the implication of diminished capacity for nutrient retention. In this context, it implies that soils with lower organic carbon content are less effective at holding onto essential elements like nitrogen, thus facilitating their movement through the landscape. Second, the interactions delineated within soil pH relationships elucidate another dimension: the potential for nitrogen volatilization losses Malhi and Nyborg (1991). This signifies that variations in soil pH can trigger the release of nitrogen compounds into the air, a process that subsequently contributes to the departure of nitrogen from the soil ecosystem. Unveiling this nexus provides a crucial link between soil dynamics and the larger ecological context, particularly in relation to Lake Erie.

The amalgamation of heightened nitrogen concentrations diminished retention capacity due to lower organic carbon, and the potential for nitrogen loss through volatilization points to a pathway through which the elevated nitrogen in the Maumee River watershed could impact Lake Erie's ecosystem Ai et al. (2023). Relating the soil data to hydrologic pathways could help quantify nitrogen fluxes from vulnerable fields into Lake Erie. This could inform conservation practices targeting key areas.

In summary, these soil chemical maps establish an informative baseline understanding of nutrient status and environmental impacts. They offer insights into spatial factors influencing nitrogen availability and transport risks.

Discussion of Spatial Correlation Results

The spatial correlation analysis in Sect. 4.4 offers valuable insights into the relationships between soil factors and crop growth by harnessing the multi-source integrated geospatial data. The spatiotemporal crop NDVI-nitrogen correlation maps reveal positive Pearson correlation coefficients ranging from 0.3 to 0.8, showing locations where higher NDVI is associated with higher soil nitrogen content. The organic carbon-nitrogen correlation analysis shows strong positive coefficients between 0.6 and 0.9, reflecting the link between soil organic matter and nitrogen supply. Meanwhile, the pH-nitrogen correlation exhibits negative coefficients from -0.5 to -0.8 , corresponding to reduced nitrogen availability in acidic soils.

The correlation analysis provides valuable insights into the complex relationships between soil fertility, crop growth, and associated water quality impacts. The crop-nitrogen correlations suggest field-specific interventions could optimize

plant nutrition based on soil nitrogen content mapped across space and time. For instance, variable-rate nitrogen fertilizer applications could target areas with lower soil nitrogen availability within fields to promote vigorous crop growth.

Temporal Correlation Analysis

The spatiotemporal crop NDVI-nitrogen correlation maps (Figs. 32 and 33), provide valuable insights into the evolution of soil nitrogen-crop relationships over the growing season. The positive correlation values show locations where higher NDVI is associated with higher soil nitrogen content, representing a positive relationship between crop greenness and soil fertility. Meanwhile, negative values correspond to areas where higher NDVI coincides with lower soil nitrogen content, indicative of an inverse relationship between the two layers.

The correlation coefficients denote the strength of these relationships, with values closer to -1 or 1 representing stronger correlations. As documented in prior studies, adequate soil nitrogen availability promotes vegetation growth and crop yields Johnson and Raun (2003). The positive correlations observed across the season validate this fundamental connection between soil nitrogen stores and crop vigor. However, the strength of the correlations varied over time, with the highest correlations evident during the peak summer months. This aligns with the crop growth stages, as corn and soybeans have the greatest nitrogen requirements later in the season when rapid biomass accumulation occurs Tremblay et al. (2012); Russell (1973); Shahandeh et al. (2005). Capturing this temporal variability provides nuanced insights into the nonlinear, dynamic nature of crop-soil nitrogen linkages over the cropping cycle.

One Time-Stamp Correlation Analysis

The spatial correlation maps at a single point in time (Figs. 34 and 35) further explain relationships among key soil properties. As expected, soil organic carbon exhibited a strong positive correlation with total nitrogen, as soil organic matter contains nitrogen that is mineralized into plant-available forms through microbial decomposition Liptzin et al. (2022).

Notably, higher soil carbon in eastern Ohio (Fig. 30), likely increased organic nitrogen reserves that slowly mineralize to support crop productivity Wang et al. (2017). This aligns with the positive carbon-nitrogen correlation, suggesting a stable soil nitrogen supply in high-carbon areas. In contrast, the negative pH-nitrogen correlation corresponds to chemical losses in acidic soils. Lower pH in northwestern Ohio (Fig. 29), promotes nitrogen leaching and volatilization, reducing availability Russell (1973); Ding et al. (2014). This suggests manure pH change could optimize nitrogen retention in acidic soils Guo et al. (2016). This implies

greater nitrogen loss risks that could impact downstream water quality. Adjusting manure pH could optimize retention in sensitive areas.

Meanwhile, the soil carbon-nitrogen links show manure or compost additions to low-carbon soils could improve organic nitrogen retention and prevent losses. The soil pH patterns underscore the need for better management of acidic soils to reduce nitrogen leaching and volatilization losses. Relating these findings to hydrologic pathways could help model downstream nitrogen delivery and prioritize conservation efforts in critical source areas.

Overall, these spatial relationships provide a nuanced biogeochemical-carbon perspective that could inform agricultural best practices. The integrated methodology could be applied to additional regions to model soil-crop interactions and forecast water quality outcomes under various climate and land management scenarios. This systems-level understanding will become increasingly valuable for tackling the complex sustainability challenges associated with food production.

Discussion of Water Feature Analysis

Figures 36, 37, 38, 39, 40, 41, 42, and 43 show the spatial relationship of nitrate and nitrite and discharge in 2019 across four seasons in Ohio. The highest concentrations of nitrate and nitrite were seen in the fall, with 7% of the gauges reporting toxic levels ($>10\text{ mg/L}$) coming from southwest Ohio. Prime farmlands are in the west of Ohio. The watershed draining into the river can carry water from these farmlands and also the urban area adjacent to the Great Miami River. Given this, the observed increase in nitrate and nitrite could be because of sewage and fertilizer, the two main causes of the increase in nitrate and nitrite levels in water Service et al. (1997); Dubrovsky and Hamilton (2010).

Plant growth during the spring and summer months involves increased uptake of nutrients, which are later released back into the environment once plants reach the end of their life cycle. In addition, water runoff from the post-harvested field may also contribute to an increase in nitrate and nitrite levels Gale et al. (2006). Wastewater treatment plants (WWTP) can contribute to higher nutrient concentrations Zouboulis and Tolkou (2015).

A preliminary investigation into the surroundings of some gauges that reported toxic and non-toxic levels of nitrate and nitrite shows toxic gauges are almost always near WWTPs. A more systematic investigation is required to quantify their contribution to nitrate and nitrite concentrations. Lastly, some gauges displaying high nitrate and nitrite concentrations are along the same river, showing that a specific watershed in the southwest may need closer inspection.

Discharge was found to increase in the winter, spring, and summer months, with the highest values found in the winter.

This could be attributed to snowmelt and precipitation during these months Van Metre et al. (2016). It is also noted that the discharge was higher in some sites irrespective of the seasons (Figs. 36, 37, 38, 39, 40, 41, 42, and 43). This could be attributed to the presence of dams Liu et al. (2014).

Figures 44 and 45 highlight the relationship between discharge and nitrate and nitrite during 2019. A notable pattern is present where lower discharge corresponds to higher nitrate and nitrite concentrations. This is a known behavior reflecting the settlement of nutrients due to lower flow rates in streams.

The spatiotemporal patterns within and between nitrate and nitrite and discharge reflect potential avenues to explore how other additional features, such as agricultural land use and soil types, can be integrated to further understand the behavior of nutrient flow.

Conclusion

This research shows the significant potential of leveraging publicly available geospatial data through an integrated spatiotemporal analysis framework. Synthesizing multi-source satellite, crop, soil, and hydrological measurements revealed intricate connections between agricultural productivity, soil fertility dynamics, and nutrient loss risks. The multi-faceted perspective attained via advanced data fusion and analytic techniques provided actionable insights to guide sustainable agricultural management practices and environmental policy decisions. Quantifying complex relationships through correlation analysis offered valuable perspectives.

While moderate-resolution data sufficiently characterizes regional crop growth and soil patterns, high-resolution data could enable field-scale monitoring and personalized interventions. Expanding the geographic scope across more agriculturally significant areas would provide broader contextual insights. Incorporating longer-term records alongside weather, climate, management, and socioeconomic data would further inform the analysis of seasonal variability and drivers. Transitioning such integrated platforms into operational use for stakeholders would multiply their impact. Enhanced computing capabilities, along with interactive data visualization and modeling tools, could empower dynamic exploratory analysis. Machine learning methods could extract deeper insights from multiplying data streams.

This work makes key contributions through the novel integration of multi-modal geospatial data and advanced spatiotemporal techniques. The methodology provides a foundation to guide sustainable agricultural practices and environmental monitoring. If transitioned into predictive modeling and tools for stakeholders, the approach could support critical decision-making. The framework offers a unique holistic perspective by linking crop, soil, and water dynam-

ics using advance spatiotemporal integration approach and Distributed High-Performance Computing.

As data availability and analytical capabilities continue advancing, such holistic approaches will become increasingly vital for ensuring productivity, sustainability, and resilience for agricultural systems.

Supplementary Information The GIFs resulting from our analysis and simulations are available for viewing. These GIFs visually depict the multi-scale geospatial analysis conducted for monitoring crop growth, land use/ vegetation classification, and correlations for 2019 in Ohio, Texas, and Florida Akanbi et al. (2023).

Appendices A, B, C, and D

This analysis was also conducted for Texas and Florida using the same methodology and data sources described for Ohio. The same analytical techniques were applied to characterize regional crop patterns and phenology in these major agricultural states based on MODIS NDVI time series data. These appendices provide the key results for vegetation classification and crop monitoring in Texas and Florida to complement the primary Ohio findings.

The supplementary results for these additional states are presented here for reference.

Appendix A: Texas Vegetation Classification

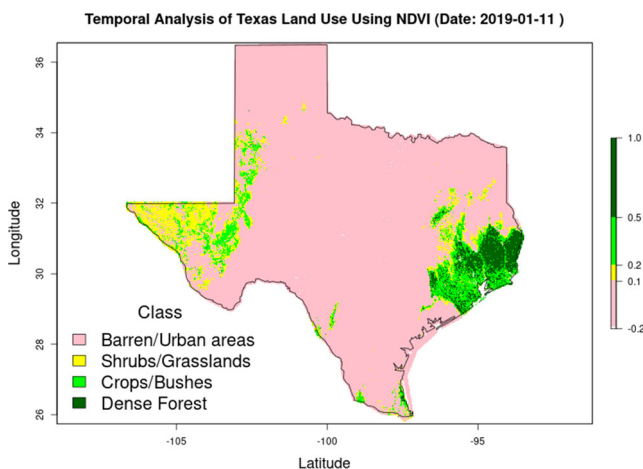


Fig. 46 Winter: Seasonal vegetation density classification map for Texas in 2019 based on MODIS NDVI analysis

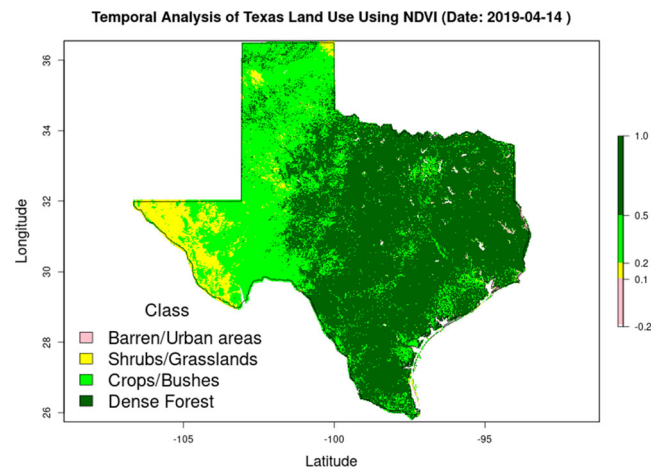


Fig. 47 Spring: Seasonal vegetation density classification map for Texas in 2019 based on MODIS NDVI analysis

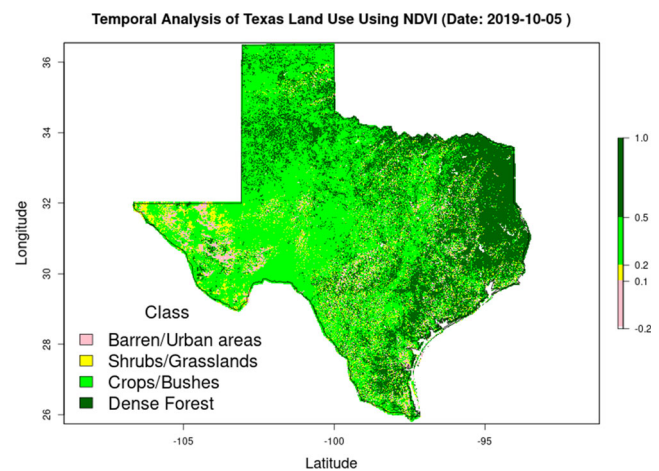


Fig. 48 Summer: Seasonal vegetation density classification map for Texas in 2019 based on MODIS NDVI analysis

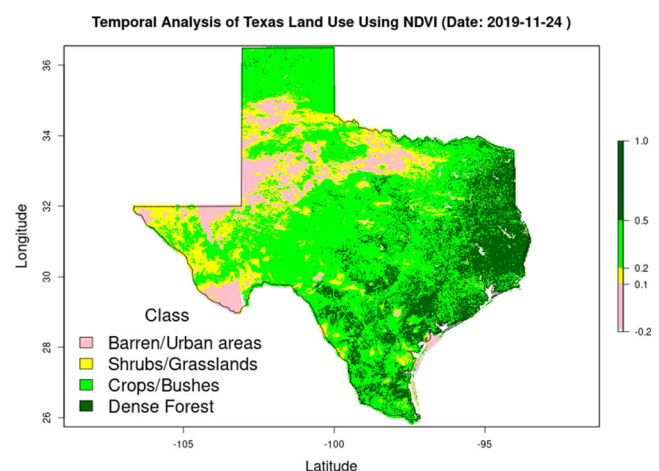


Fig. 49 Fall: Seasonal vegetation density classification map for Texas in 2019 based on MODIS NDVI analysis

Appendix B: Texas Crop Monitoring

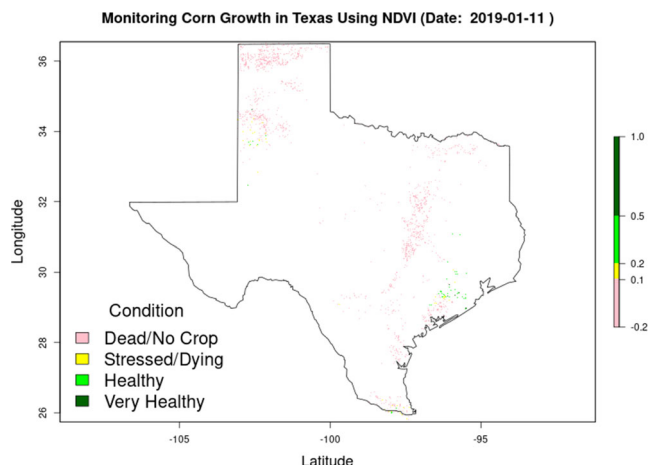


Fig. 50 Winter: Crop-specific monitoring of corn growth in Texas during 2019 using MODIS NDVI time series data

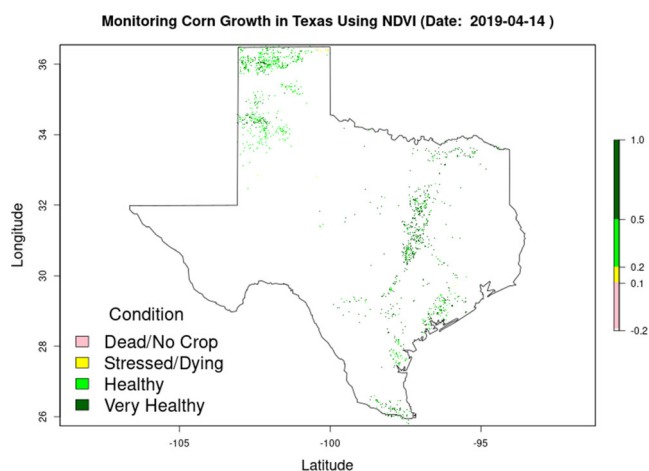


Fig. 51 Spring: Crop-specific monitoring of corn growth in Texas during 2019 using MODIS NDVI time series data

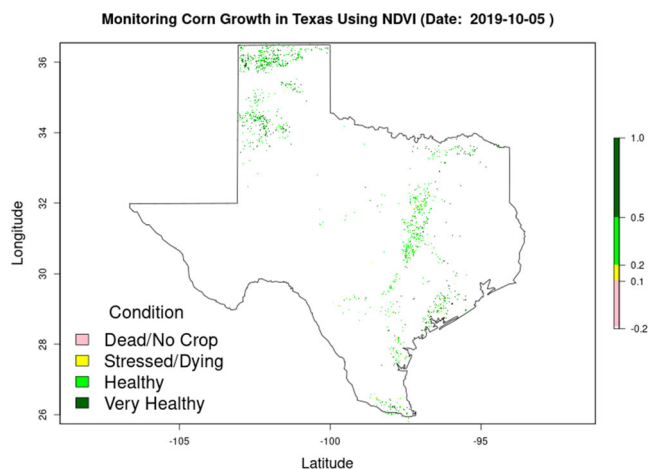


Fig. 52 Summer: Crop-specific monitoring of corn growth in Texas during 2019 using MODIS NDVI time series data

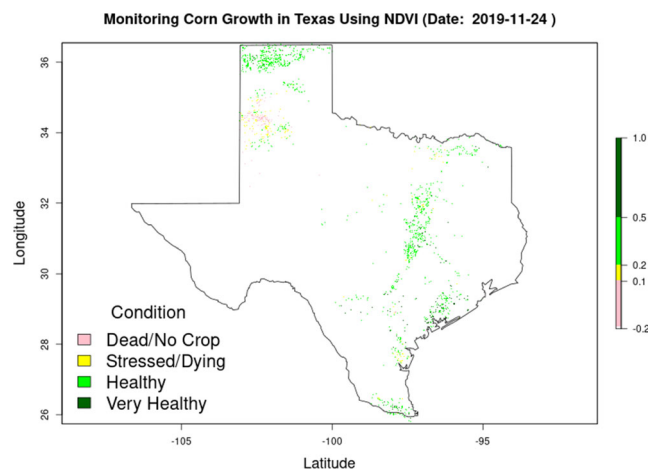


Fig. 53 Fall: Crop-specific monitoring of cotton growth in Texas during 2019 using MODIS NDVI time series data

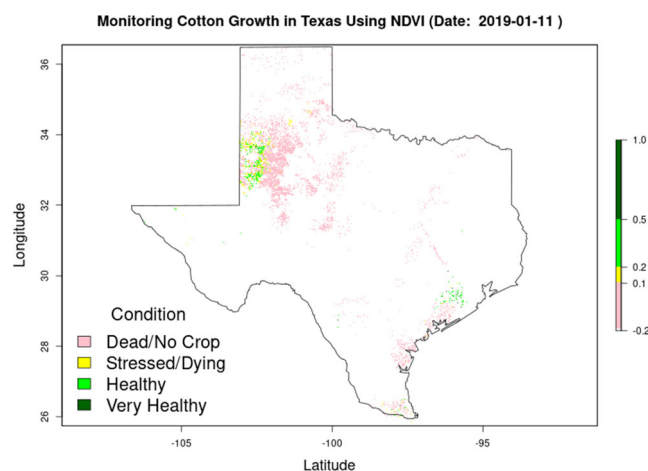


Fig. 54 Winter: Crop-specific monitoring of cotton growth in Texas during 2019 using MODIS NDVI time series data

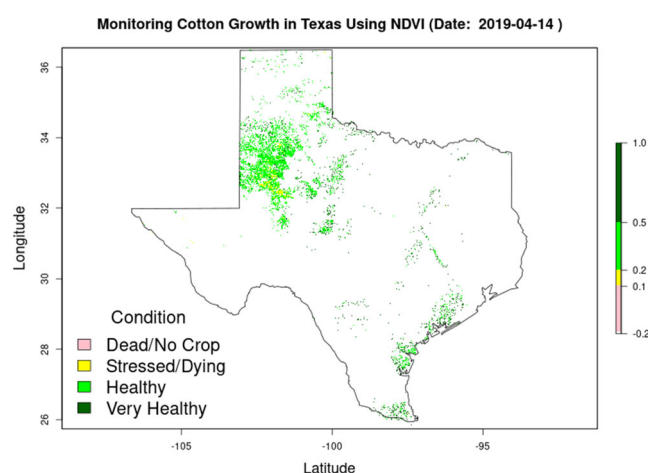


Fig. 55 Spring: Crop-specific monitoring of cotton growth in Texas during 2019 using MODIS NDVI time series data

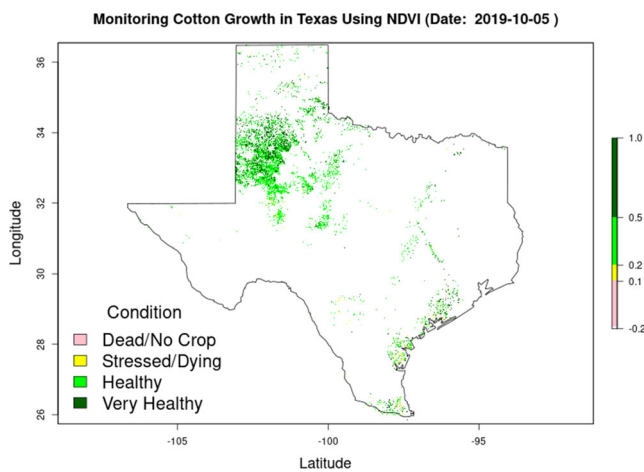


Fig. 56 Summer: Crop-specific monitoring of cotton growth in Texas during 2019 using MODIS NDVI time series data

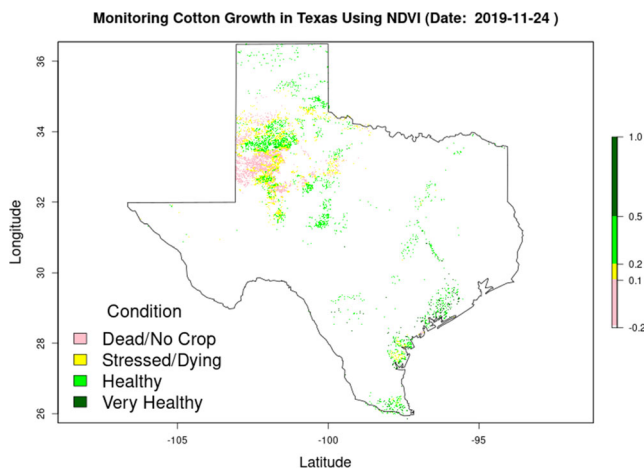


Fig. 57 Fall: Crop-specific monitoring of cotton growth in Texas during 2019 using MODIS NDVI time series data

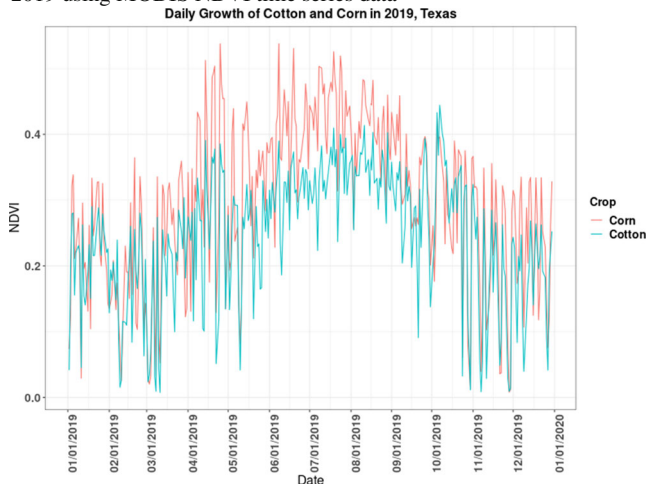


Fig. 58 Average seasonal growth profiles of cotton and corn in Texas during 2019 derived from MODIS NDVI

Appendix C: Florida Vegetation Classification

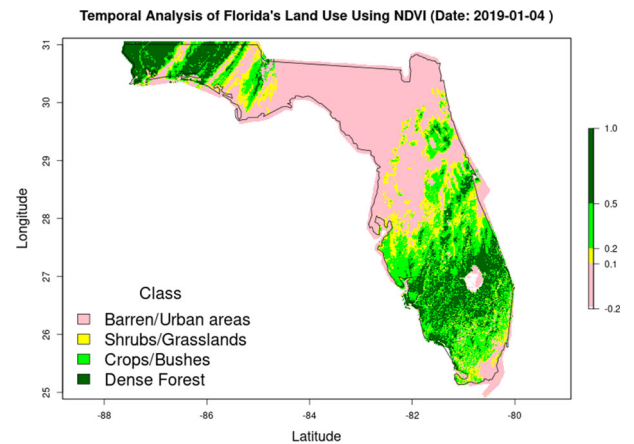


Fig. 59 Winter: Seasonal vegetation density classification map for Florida in 2019 based on MODIS NDVI analysis

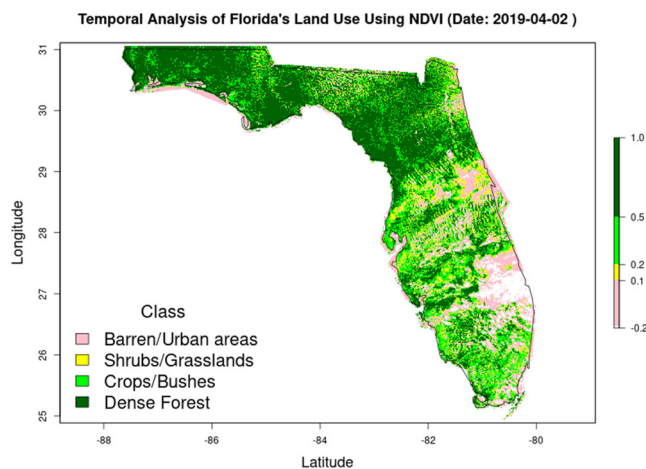


Fig. 60 Spring: Seasonal vegetation density classification map for Florida in 2019 based on MODIS NDVI analysis

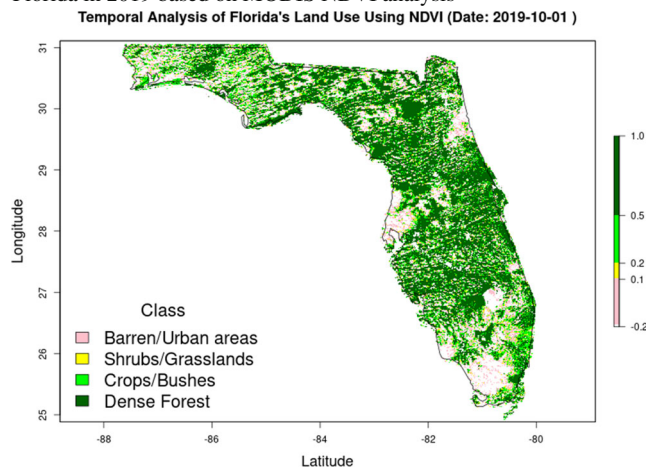


Fig. 61 Summer: Seasonal vegetation density classification map for Florida in 2019 based on MODIS NDVI analysis

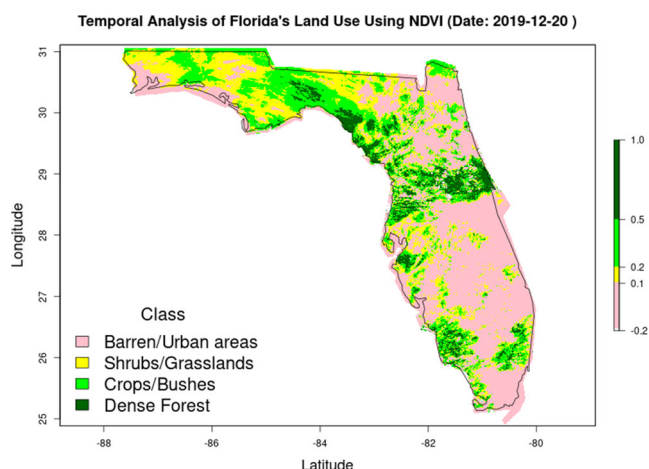


Fig. 62 Fall: Seasonal vegetation density classification map for Florida in 2019 based on MODIS NDVI analysis

Appendix D: Florida Crop Monitoring

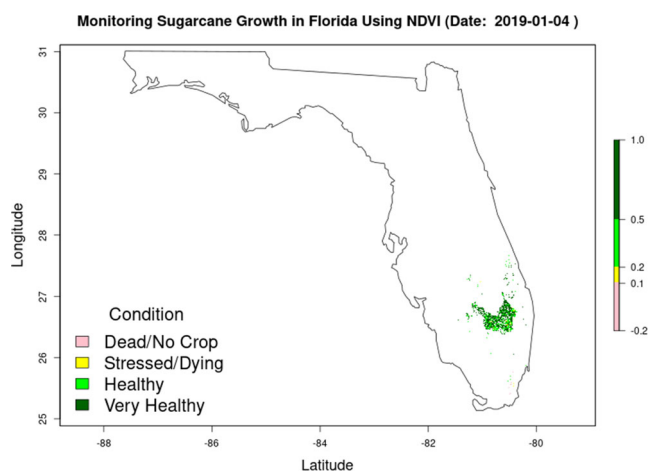


Fig. 63 Winter: Crop-specific monitoring of sugarcane growth in Florida during 2019 using MODIS NDVI time series data

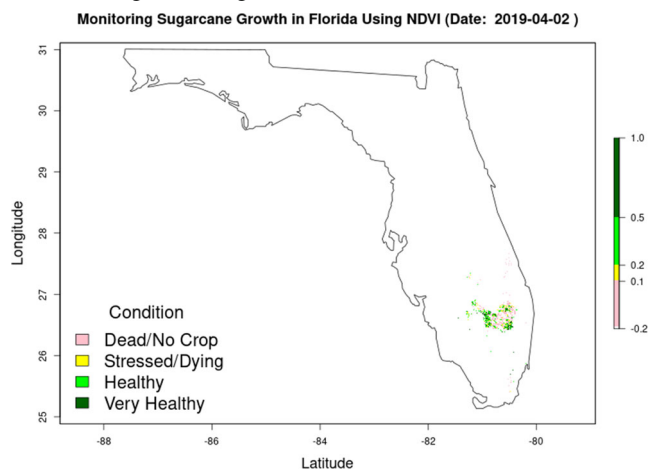


Fig. 64 Spring: Crop-specific monitoring of sugarcane growth in Florida during 2019 using MODIS NDVI time series data

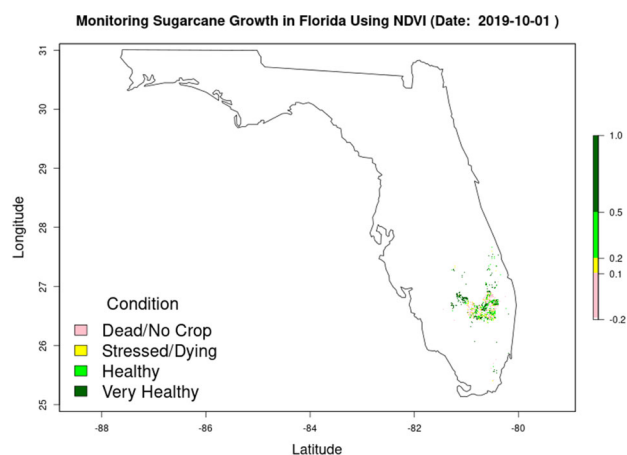


Fig. 65 Summer: Crop-specific monitoring of sugarcane growth in Florida during 2019 using MODIS NDVI time series data

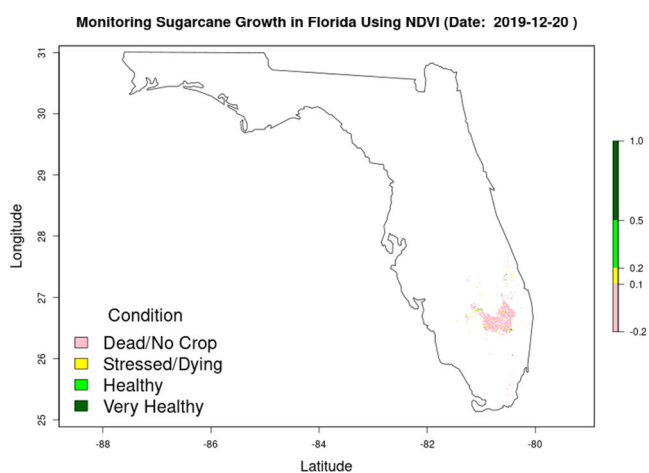


Fig. 66 Fall: Crop-specific monitoring of sugarcane growth in Florida during 2019 using MODIS NDVI time series data

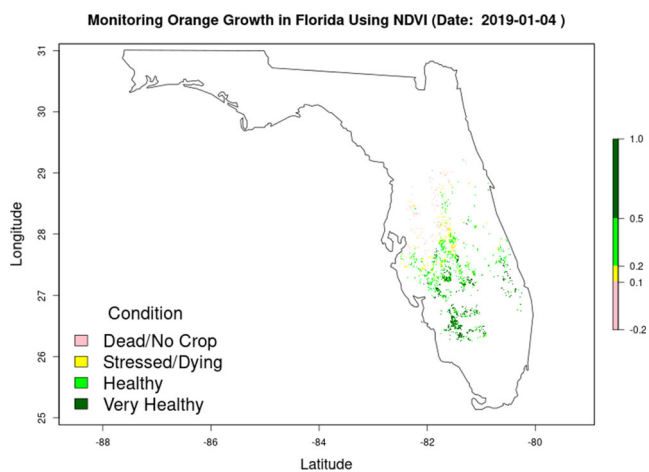


Fig. 67 Winter: Crop-specific monitoring of orange growth in Florida during 2019 using MODIS NDVI time series data

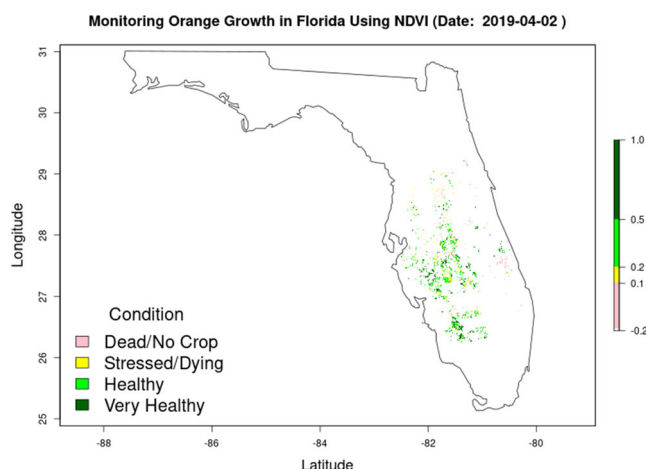


Fig. 68 Spring: Crop-specific monitoring of orange growth in Florida during 2019 using MODIS NDVI time series data

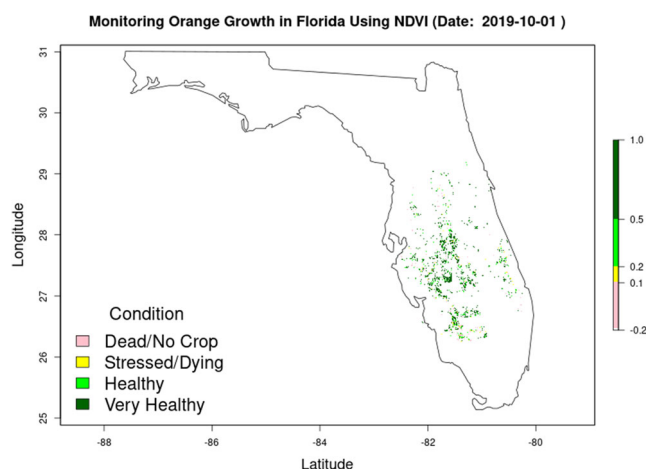


Fig. 69 Summer: Crop-specific monitoring of orange growth in Florida during 2019 using MODIS NDVI time series data

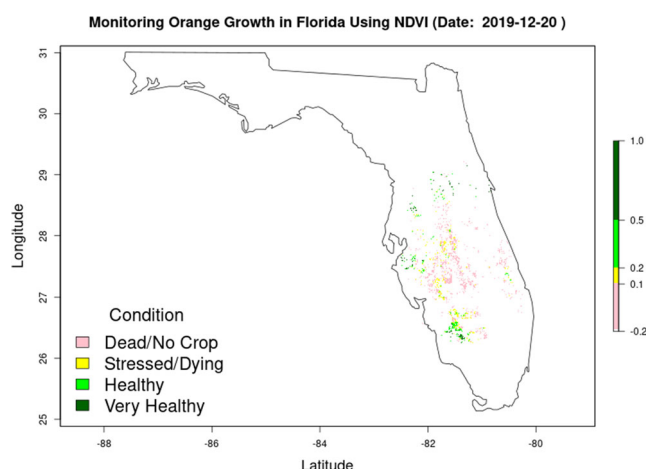


Fig. 70 Fall: Crop-specific monitoring of orange growth in Florida during 2019 using MODIS NDVI time series data

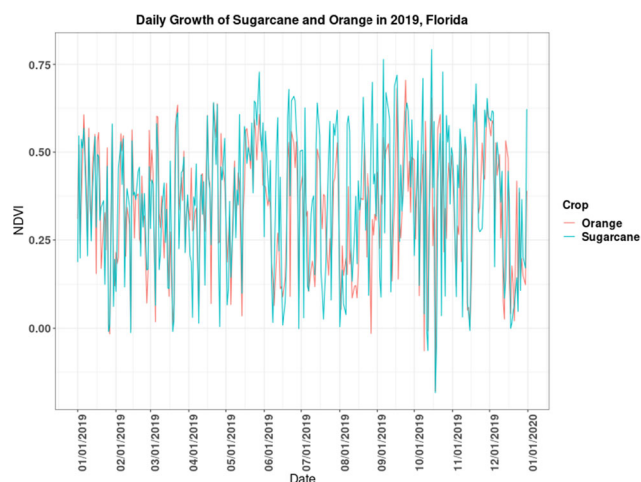


Fig. 71 Average seasonal growth profiles of sugarcane and oranges in Florida during 2019 derived from MODIS NDVI

Acknowledgements This study was conducted in the SDLE Research Center at Case Western Reserve University. This work made use of the High-Performance Computing Resource in the Core Facility for Advanced Research Computing at Case Western Reserve University. The authors acknowledge Hendrik Hamann and IBM Environmental Intelligence Suite for some of the data integrated for this study.

Funding This work is supported in part by the National Science Foundation under Grant No. 2133576.

Data Availability The data used in this project are sourced exclusively from open-access repositories. All data outputs generated during the study are publicly available and properly referenced in this paper.

Declarations

Compliance with Ethical Standards The authors were compliant with the ethical standards.

Ethics Approval All authors certify that no ethical approval was required for the subject matter or materials discussed in this manuscript. All authors certify that the research subject of this manuscript does not raise any ethical concerns.

Informed Consent All authors provided informed consent. They were fully informed about the study's purpose, procedures, potential risks, and benefits, and they voluntarily agreed to participate with a clear understanding of their rights and the confidentiality of their data.

Competing Interests The authors declare no competing interests.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your

intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References


- Abedinpour M (2015) Evaluation of growth-stage-specific crop coefficients of maize using weighing lysimeter. *Soil Water Res* 10(2):99–104
- Ahmad M, Zeeshan M (2022) Validation of weather reanalysis datasets and geospatial and techno-economic viability and potential assessment of concentrated solar power plants. *Energy Convers Manage* 256:115366
- Ai H, Zhang K, Sun J et al (2023) Short-term Lake Erie algal bloom prediction by classification and regression models. *Water Res* 232:119710. <https://doi.org/10.1016/j.watres.2023.119710>
- Akanbi O, Bhuvanagiri D, Barcelos E et al (2023) Integrating multiscale geospatial analysis for monitoring crop growth, nutrient distribution, and hydrological dynamics in large-scale agricultural systems. <https://osf.io/z6cvh/>. publisher: OSF
- Arshad M, Lowery B, Grossman B (1997) Physical tests for monitoring soil quality. *Methods Assess Soil Qual* 49:123–141
- Azar D, Engstrom R, Graesser J et al (2013) Generation of fine-scale population layers using multi-resolution satellite imagery and geospatial data. *Remote Sens Environ* 130:219–232
- Bartos M (2023) Mdbartos pysheds: Earth Americas simple and fast watershed delineation in python. <https://github.com/mdbartos/pysheds>
- Batjes NH, Ribeiro E, van Oostrum AJM (2019) Standardised soil profile data for the world (Wosis snapshot - September 2019). Accessed 23 Jun 2023
- Beery M, Wilding LP (1971) The relationship between soil pH and base-saturation percentage for surface and subsoil horizons of selected mollisols, alfisols, and ultisols in Ohio. *Ohio J Sci* 71(1):43–55. <https://kb.osu.edu/handle/1811/5590>
- Bosch NS, Allan JD, Selegue JP et al (2013) Scenario-testing of agricultural best management practices in Lake Erie Watersheds. *J Great Lakes Res* 39(3):429–436
- Botte G, French RH, Hatzell M et al (2023) Center for advancing sustainable and distributed fertilizer production (CASFER), and NSF engineering research center. <https://www.casfer.us>
- Bradley BA, Jacob RW, Hermance JF et al (2007) A curve fitting procedure to derive inter-annual phenologies from time series of noisy satellite NDVI data. *Remote Sens Environ* 106(2):137–145
- Chen J (2018) Remote sensing of leaf area index and clumping index. In: *Comprehensive Remote Sensing*. Elsevier, Oxford, pp 53–77. <https://doi.org/10.1016/B978-0-12-409548-9.10540-8>
- Chen J, Jönsson P, Tamura M et al (2004) A simple method for reconstructing a high-quality NDVI time-series data set based on the Uavitzky-Golay filter. *Remote Sens Environ* 91(3–4):332–344
- Coffman VR, Jensen AS, Trabjerg BB et al (2021) Prenatal exposure to nitrate from drinking water and markers of fetal growth restriction: a population-based study of nearly one million Danish-born children. *Environ Health Perspect* 129(2):027002. <https://doi.org/10.1289/EHP7331>
- Conrey GW (1941) The origin of Ohio soils. *Ohio J Sci* 41(3):201–206
- Coughlan T (2020) The use of open data as a material for learning. *Education Tech Research Dev* 68(1):383–411. <https://doi.org/10.1007/s11423-019-09706-y>
- DeFries RS, Townshend J (1994) NDVI-derived land cover classifications at a global scale. *Int J Remote Sens* 15(17):3567–3586
- Demattê JAM, Fongaro CT, Rizzo R et al (2018) Geospatial soil sensing system (geos3): a powerful data mining procedure to retrieve soil spectral reflectance from satellite images. *Remote Sens Environ* 212:161–175
- Ding LJ, An XL, Li S et al (2014) Nitrogen loss through anaerobic ammonium oxidation coupled to iron reduction from paddy soils in a chronosequence. *Environ Sci Technol* 48(18):10641–10647
- Dragoni D, Rahman AF (2012) Trends in fall phenology across the deciduous forests of the Eastern USA. *Agric For Meteorol* 157:96–105
- Dragoni D, Schmid HP, Wayson CA et al (2011) Evidence of increased net ecosystem productivity associated with a longer vegetated season in a deciduous forest in South-Central Indiana, USA. *Glob Change Biol* 17(2):886–897
- Dubrovsky NM, Hamilton PA (2010) Nutrients in the nation's streams and groundwater: national findings and implications. <https://pubs.usgs.gov/fs/2010/3078/>
- Easterly NW (1964) Distribution patterns of Ohio cruciferae. *Castanea* 29(3):164–173
- Extension OSU (2023) Soil type and history. <https://soilhealth.osu.edu/soil-health-assessment/soil-type-history>. Accessed 23 Jun 2023
- Fang H, Liang S (2014) Leaf area index models. In: *Reference Module in Earth Systems and Environmental Sciences*. Elsevier. <https://doi.org/10.1016/B978-0-12-409548-9.09076-X>
- Fisher A, Flood N, Danaher T (2016) Comparing Landsat water index methods for automated water classification in Eastern Australia. *Remote Sens Environ* 175:167–182
- Fouli Y, Cade-Menun BJ, Cutforth HW (2013) Freeze-thaw cycles and soil water content effects on infiltration rate of three Saskatchewan soils. *Can J Soil Sci* 93(4):485–496
- Gale ES, Sullivan DM, Cogger CG et al (2006) Estimating plant-available nitrogen release from manures, composts, and specialty products. *J Environ Qual* 35(6):2321–2332. <https://doi.org/10.2134/jeq2006.0062>
- Gao BC (1996) NDWI-a normalized difference water index for remote sensing of vegetation liquid water from space. *Remote Sens Environ* 58(3):257–266
- Guo L, Wu G, Li Y et al (2016) Effects of cattle manure compost combined with chemical fertilizer on topsoil organic matter, bulk density and earthworm activity in a wheat-maize rotation system in Eastern China. *Soil Tillage Res* 156:140–147
- Hamdi A, Shaban K, Erradi A et al (2022) Spatiotemporal data mining: a survey on challenges and open problems. *Artif Intell Rev* 1–48
- Hatfield JL, Takle GR, Grotjahn R et al (2014) Agriculture. Climate change impacts in the United States: the third national climate assessment. US Global Change Research Program, Washington, DC, pp 150–174
- Heathwaite L, Sharpley A, Gburek W (2000) A conceptual approach for integrating phosphorus and nitrogen management at watershed scales. *J Environ Qual* 29(1):158–166
- Hijmans RJ, Bivand R, Pebesma E et al (2023a) Terra: Spatial data analysis
- Hijmans RJ, van Etten J, Sumner M et al (2023b) Raster: Geographic data analysis and modeling
- Huang S, Tang L, Hupy JP et al (2021) A commentary review on the use of normalized difference vegetation index (NDVI) in the era of popular remote sensing. *J For Res* 32(1):1–6
- Huete AR (1988) A soil-adjusted vegetation index (SAVI). *Remote Sens Environ* 25(3):295–309
- Hu Y, Hosain MA, Jain T et al (2013) Global SunFarm data acquisition network, energy CRADLE, and time series analysis. In: *2013 IEEE Energytech*. IEEE, pp 1–5
- Jackson TJ, Chen D, Cosh M et al (2004) Vegetation water content mapping using Landsat data derived normalized difference water index for corn and soybeans. *Remote Sens Environ* 92(4):475–482
- Jarecki MK, Lal R (2005) Soil organic carbon sequestration rates in two long-term no-till experiments in Ohio. *Soil Sci* 170(4):280–291

- Jenkins J, Richardson A, Braswell B et al (2007) Refining light-use efficiency calculations for a deciduous forest canopy using simultaneous tower-based carbon flux and radiometric measurements. *Agric For Meteorol* 143(1–2):64–79
- Jiang Z, Huete AR, Didan K et al (2008) Development of a two-band enhanced vegetation index without a blue band. *Remote Sens Environ* 112(10):3833–3845
- Johnson G, Raun W (2003) Nitrogen response index as a guide to fertilizer management. *J Plant Nutr* 26(2):249–262
- Jones R (2023) Algorithms for using a dem for mapping catchment areas of stream sediment samples. https://www.sciencedirect.com/science/article/pii/S0098300402000225?casa_token=bOqFqWzA1mgAAAAA:NvV3YqvaxTkFjUfXUICp7wCLCgmLpu3ZN8SqWkQoPkMmY5GcW3f193p3W-r22LP5rkwRDH-iXbe
- Jung JY, Lal R (2011) Impacts of nitrogen fertilization on biomass production of switchgrass (*panicum virgatum* L.) and changes in soil organic carbon in Ohio. *Geoderma* 166(1):145–152
- Khalilnejad A, Karimi AM, Kamath S et al (2020) Automated pipeline framework for processing of large-scale building energy time series data. *PLoS One* 15(12):e0240461. <https://doi.org/10.1371/journal.pone.0240461>
- Klein LJ, Marianno FJ, Albrecht CM et al (2015) Pairs: a scalable geo-spatial data analytics platform. In: 2015 IEEE International Conference on Big Data (Big Data). IEEE, pp 1290–1298
- Ko J, Piccinini G, Marek T et al (2009) Determination of growth-stage-specific crop coefficients (kc) of cotton and wheat. *Agric Water Manag* 96(12):1691–1697
- Lindbo DL, Kozlowski DA, Robinson C et al (2012) Know soil know life. *Soil Sci Soc Am*. <https://doi.org/10.2136/2012.knowsoil.c1>, <https://cir.nii.ac.jp/crid/1130282269040517248>
- Liptzin D, Norris C, Cappellazzi S et al (2022) An evaluation of carbon indicators of soil health in long-term agricultural experiments. *Soil Biol Biochem* 172:108708. <https://doi.org/10.1016/j.soilbio.2022.108708>
- Liu Y, Yang W, Yu Z et al (2014) Assessing effects of small dams on stream flow and water quality in an agricultural watershed. *J Hydrol Eng* 19(10):05014015. [https://doi.org/10.1061/\(ASCE\)HE.1943-5584.0001005](https://doi.org/10.1061/(ASCE)HE.1943-5584.0001005)
- Lnenicka M, Nikiforova A (2021) Transparency-by-design: what is the role of open data portals? *Telematics Inform* 61. <https://doi.org/10.1016/j.tele.2021.101605>
- Lu X, Zhuang Q (2010) Evaluating evapotranspiration and water-use efficiency of terrestrial ecosystems in the conterminous United States using MODIS and AmeriFlux data. *Remote Sens Environ* 114(9):1924–1939
- Lu S, Hamann HF (2021) IBM pairs: scalable big geospatial-temporal data and analytics as-a-service. In: *Handbook of Big Geospatial Data*. Springer International Publishing, Cham, pp 3–34. https://doi.org/10.1007/978-3-030-55462-0_1
- Lu S, Shao X, Freitag M et al (2016) IBM pairs curated big data service for accelerated geospatial data analytics and discovery. In: 2016 IEEE International Conference on Big Data (Big Data). IEEE, pp 2672–2675
- Malhi S, Nyborg M (1991) Recovery of 15 n-labelled urea: influence of zero tillage, and time and method of application. *Fertilizer Res* 28:263–269
- Mallah S, Delsouz Khaki B, Davatgar N et al (2022) Predicting soil textural classes using random forest models: learning from imbalanced dataset. *Agronomy* 12(11):2613
- Manassaram DM, Backer LC, Messing R et al (2010) Nitrates in drinking water and methemoglobin levels in pregnancy: a longitudinal study. *Environ Health* 9(1):60. <https://doi.org/10.1186/1476-069X-9-60>
- METI, NASA (2019) ASTER GDEM and ASTER water body dataset (ASTWBD). <https://www.jspacesystems.or.jp/ersdac/GDEM/E/>
- Michalak AM, Anderson EJ, Beletsky D et al (2013) Record-setting algal bloom in Lake Erie caused by agricultural and meteorological trends consistent with expected future conditions. *Proc Natl Acad Sci* 110(16):6448–6452
- MODIS N (2019) Modis aqua data. <https://ladsweb.modaps.eosdis.nasa.gov/>. Accessed 23 Jun 2023
- Mohammadpour P, Grady C (2023) Regional analysis of nitrogen flow within the Chesapeake bay watershed food production chain inclusive of trade. *Environ Sci Technol* 57(11):4619–4631
- Navarro G, Caballero I, Silva G et al (2017) Evaluation of forest fire on madeira island using Sentinel-2A MSI imagery. *Int J Appl Earth Obs Geoinf* 58:97–106
- OGC (2023) OGC GeoTIFF standard
- Ooms J, J, cre (2023) Magick: advanced graphics and image-processing in R
- Palmer MJ, Jamieson HE, Radková AB et al (2021) Mineralogical, geospatial, and statistical methods combined to estimate geochemical background of arsenic in soils for an area impacted by legacy mining pollution. *Sci Total Environ* 776:145926
- Peng YF, Tang JH, Yc Fu et al (2016) Analyzing personal happiness from global survey and weather data: a geospatial approach. *PLoS One* 11(4):e0153638
- Peñuelas J, Pinol J, Ogaya R et al (1997) Estimation of plant water concentration by the reflectance water index WI (r900/r970). *Int J Remote Sens* 18(13):2869–2875
- Pettorelli N, Vik JO, Mysterud A et al (2005) Using the satellite-derived NDVI to assess ecological responses to environmental change. *Trends Ecol Evol* 20(9):503–510
- Piccinini G, Ko J, Marek T et al (2009) Determination of growth-stage-specific crop coefficients (kc) of Maize and Sorghum. *Agric Water Manag* 96(12):1698–1704
- R Core Team (2023) R: the R project for statistical computing
- Ritter N, Ruth M (1997) The GeoTiff data interchange standard for raster geographic images. *Int J Remote Sens* 18(7):1637–1647. <https://doi.org/10.1080/014311697218340>
- Rouse JW, Haas RH, Deering DW et al (1974a) Monitoring the vernal advancement and retrogradation (Green Wave Effect) of natural vegetation. Technical Report E75-10354, NASA. <https://ntrs.nasa.gov/citations/19750020419>
- Rouse Jr JW, Haas RH, Schell JA et al (1974b) Monitoring vegetation systems in the great plains with ERTS. In: *Third Earth Resources Technology Satellite-1 Symposium: The Proceedings of a Symposium Held by Goddard Space Flight Center at Washington, DC on December 10-14, 1973: Prepared at Goddard Space Flight Center, vol. 351. Scientific and Technical Information Office, National Aeronautics and Space ...*, Goddard, Washington DC, USA, pp 309
- Russell E (1973) Plant nutrition and crop production. *Plant Nutrition and Crop Production*. pp 11–28
- Sakamoto T, Wardlow BD, Gitelson AA et al (2010) A two-step filtering approach for detecting maize and soybean phenology with time-series MODIS data. *Remote Sens Environ* 114(10):2146–2159
- Saxton KE, Rawls WJ (2006) Soil water characteristic estimates by texture and organic matter for hydrologic solutions. *Soil Sci Soc Am J* 70(5):1569–1578
- Scavia D, Allan JD, Arend KK et al (2014) Assessing and addressing the re-eutrophication of Lake Erie: central basin hypoxia. *J Great Lakes Res* 40(2):226–246
- Schlossberg T (2017) Fertilizers, a boon to agriculture. *Pose Growing Threat to U.S. Waterways*, The New York Times
- Service NRC, Cartography N, Center GIS (1997) Prime farmland distribution. <https://www.loc.gov/resource/g4081j.ct011814/>
- Setiyono T, Weiss A, Specht J et al (2007) Understanding and modeling the effect of temperature and daylength on soybean phenology under high-yield conditions. *Field Crop Res* 100(2–3):257–271

- Shahandeh H, Wright A, Hons F et al (2005) Spatial and temporal variation of soil nitrogen parameters related to soil texture and corn yield. *Agron J* 97(3):772–782
- Shaver T, Khosla R, Westfall D (2006) Utilizing green normalized difference vegetation indices (GNDVI) for production level management zone delineation in irrigated corn. In: *The 18th World Congress of Soil Science*
- Shibayama M, Akiyama T (1986) A spectroradiometer for field use: Vi. radiometric estimation for chlorophyll index of rice canopy. *Jap J Crop Sci* 55(4):433–438
- Shrestha S, Brueck H, Asch F (2012) Chlorophyll index, photochemical reflectance index and chlorophyll fluorescence measurements of rice leaves supplied with different n levels. *J Photochem Photobiol B* 113:7–13
- Singh S, Anil AG, Kumar V et al (2022) Nitrates in the environment: a critical review of their distribution, sensing techniques, ecological effects and remediation. *Chemosphere* 287. <https://doi.org/10.1016/j.chemosphere.2021.131996>
- Sinha E, Michalak AM, Balaji V (2017) Eutrophication will increase during the 21st century as a result of precipitation changes. *Science* 357(6349):405–408. <https://doi.org/10.1126/science.aan2409>
- Smith DR, King KW, Johnson L et al (2015) Surface runoff and tile drainage transport of phosphorus in the Midwestern United States. *J Environ Qual* 44(2):495–502
- Staff SS (2019) Web soil survey. <https://websoilsurvey.nrcs.usda.gov/>. Accessed 23 Jun 2023
- Subburayalu S, Jenhani I, Slater B (2014) Disaggregation of component soil series on an Ohio county soil survey map using possibilistic decision trees. *Geoderma* 213:334–345
- Survey USG, Agency EP, Council NWQM (2019) Water quality data home. <https://www.waterqualitydata.us/#advanced=true>
- Survey USG (2019) USGS water data for the nation. <https://waterdata.usgs.gov/>. Accessed 23 Jun 2023
- Suyker AE, Verma SB (2012) Gross primary production and ecosystem respiration of irrigated and rainfed maize-soybean cropping systems over 8 years. *Agric For Meteorol* 165:12–24
- Tachikawa T, Hato M, Kaku M et al (2011a) Aster global digital elevation model version 2. <https://www.jspacesystems.or.jp/ersdac/GDEM/E/1.html>. Accessed 23 Jun 2023
- Tachikawa T, Hato M, Kaku M et al (2011b) The characteristics of ASTER GDEM version 2. In: *IGARSS*
- Tremblay N, Bouroubi YM, Bélec C et al (2012) Corn response to nitrogen is influenced by soil texture and weather. *Agron J* 104(6):1658–1671
- United States Department of Agriculture (2019) Historical crop planting data. https://www.nass.usda.gov/Quick_Stats/. Accessed 23 Jun 2023
- USDA (2022) Ohio - state agriculture overview - 2022. https://www.nass.usda.gov/Quick_Stats/. Accessed 23 Jun 2023
- Van Metre PC, Frey JW, Musgrove M et al (2016) High nitrate concentrations in some Midwest United States streams in 2013 after the 2012 drought. *J Environ Qual* 45(5):1696–1704. <https://doi.org/10.2134/jeq2015.12.0591>
- Wang J, Rich PM, Price KP (2003) Temporal responses of NDVI to precipitation and temperature in the Central Great Plains, USA. *Int J Remote Sens* 24(11):2345–2364
- Wang T, Kang F, Cheng X et al (2017) Spatial variability of organic carbon and total nitrogen in the soils of a subalpine forested catchment at Mt. Taiyue, China. *Catena* 155:41–52
- Wardlow BD, Egbert SL, Kastens JH (2007) Analysis of time-series MODIS 250 m vegetation index data for crop classification in the US Central Great Plains. *Remote Sens Environ* 108(3):290–310
- Waring R, Coops N, Fan W et al (2006) Modis enhanced vegetation index predicts tree species richness across forested ecoregions in the contiguous USA. *Remote Sens Environ* 103(2):218–226
- Yaman B, Thompson K, Brennan R (2021) A SKOS taxonomy of the UN global geospatial information management data theme. In: *4th International Workshop on Geospatial Linked Data at ESWC 2021*. CEUR-WS, Herssonissos, Greece(Online). <https://ceur-ws.org/Vol-2977/paper11.pdf>
- Zhong L, Hu L, Yu L et al (2016) Automated mapping of soybean and corn using phenology. *ISPRS J Photogramm Remote Sens* 119:151–164
- Zhu J, Miller A, Lindsay C et al (2013) Modis NDVI products and metrics user manual. Geographic Information Network of Alaska, UAF
- Zouboulis A, Tolkou A (2015) Effect of climate change in wastewater treatment plants: reviewing the problems and solutions. In: Shrestha S, Anal AK, Salam PA, et al (eds) *Managing Water Resources under Climate Uncertainty: Examples from Asia, Europe, Latin America, and Australia*. Springer Water, Springer International Publishing, Cham, pp 197–220. https://doi.org/10.1007/978-3-319-10467-6_10

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Authors and Affiliations

Olatunde D. Akanbi^{1,2,3} · Deepa C. Bhuvanagiri^{2,3,4} · Erika I. Barcelos^{1,2,3} · Arafath Nihar^{2,4} · Brian Gonzalez Hernandez^{2,4} · Jeffrey M. Yarus^{1,2,3}  · Roger H. French^{1,2,3,4}

Olatunde D. Akanbi
oda10@case.edu

Deepa C. Bhuvanagiri
dcb117@case.edu

Erika I. Barcelos
eib14@case.edu

Arafath Nihar
axn392@case.edu

Brian Gonzalez Hernandez
brg62@case.edu

Roger H. French
rxf131@case.edu

¹ Department of Material Science and Engineering, Case Western Reserve University, Cleveland 44106, OH, USA

² SDLE Research Center, Case Western Reserve University, Cleveland 44106, OH, USA

³ Center for Advancing Sustainable and Distributed Fertilizer Production, Case Western Reserve University, Cleveland 44106, OH, USA

⁴ Department of Computer and Data Sciences, Case Western Reserve University, Cleveland 44106, OH, USA