



Proceeding Paper

Integration of Object-Oriented Remote Sensing and Machine Learning to Create Field Model for Optimized Regional Agricultural Management

Emine Senkardesler ¹

¹ Department of Geodetic and Geographic Information Technologies, Middle East Technical University; emine.senkardesler.0@gmail.com

Abstract: In an era marked by tools like Artificial Intelligence (AI), Machine Learning (ML) and remote sensing (RS), agriculture is a primary beneficiary. These technologies help to optimize agricultural productivity, by improving resource usage and increasing yield. They not only optimize resource use but also adapt to climate change, necessitating the management of risks associated with agricultural practices. Vegetation Indices (VI) such as Normalized Difference Vegetation Index (NDVI) are relatively simple yet useful algorithms that can be used to implement precision agriculture (PA). Optical satellite images can sense the reflected lights coming from leaves which can provide various crop development information used to implement PA. Agriculture sector is important for regional economy. If managed properly, many problems related to this sector like climate change, environmental problems, and economic development, can be eliminated. PA applications can be used to establish regional management policies. This study involves monitoring agricultural production both seasonally and daily using Sentinel-2 multi-spectral time-series data. Time-series images from 2017 to 2022 are analyzed to estimate phenological dates of crops. To understand these stages, a combination of MSAVI (Modified Soil Adjusted Vegetation Index) and NDVI is used. First, mean MSAVI is calculated by the year, depending on thresholds, NDVI values are replaced with MSAVI values for certain dates, and phenological dates are determined according to the merged mean Vegetation Index (VI) values. The results are compared with Crop Progress Report (CPR) published by United States Department of Agriculture (USDA) with Root Mean Square Error (RMSE). After finding the stages, the field is mosaicked for each stage for each year. For the bare soil dates, Normalized Difference Salinity Index (NDSI) is calculated to understand the change of soil salinity. For the dates of emergence and silking MSAVI is used. For the dough, dent, mature and harvest stages NDVI is used. To understand daily changes, object-oriented and pixel-based methods (land segmentation) for field model are used to detect trends in the field. Standard deviation of every pixel is calculated and clusters are created with k-means clustering algorithm. The field model includes the characteristics of the field. In PA, site specific solutions are extremely important to get the optimum results. Since agricultural events have a great effect on agricultural applications, using meteorological data is the main milestone to improve this study. Overall, this research aims to contribute to regional agricultural production and management modules by using remote sensing and machine learning technology.



Citation: Lastname, F.; Lastname, F.; Lastname, F. Title. *Environ. Sci. Proc.* **2023**, *1*, 0. <https://doi.org/>

Published:



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: remote sensing; precision agriculture; machine learning; image processing; vegetation index; regional agricultural management, phenology, NDSI; MSAVI; NDVI

1. Introduction

In order to meet the nutritional needs of the increasing population, there is a need to implement agricultural practices in smarter and more strategic ways. These advanced but careful methods are needed not only to meet demand but also to optimize costs while considering sustainability. One of the main information resource to understand a field is

looking its time series history [11]. Agricultural RS, which mainly uses surface reflectance information of visible, near - infrared and shortwave - infrared regions of electromagnetic spectrum, are used to handle site - specific solutions to implement PA. GeoAI (Geospatial Artificial Intelligence) is a sub-field of Artificial Intelligence (AI) which focuses on the applications of AI to geospatial data and problems. GeoAI methods can analyze remote sensing data to make better predictions for implementing PA.

It has been known that NDSI is used at the very beginning of planting to evaluate the soil conditions. It measures the salt content [15] in the soil which can impact the growth and yield of crops. MSAVI can be used to monitor early growth because it eliminates the background effects in areas where soil is not completely covered by vegetation [2]. NDVI is used when the crops have grown to a stage where the fields are fully covered by vegetation.

Phenological stages of crops refer to the distinct phases in the life cycle of a crop. Understanding these stages is crucial for managing crop growth as they determine the optimal timing for various agricultural activities like irrigation, application of fertilizers and pesticides, and harvest. Timing is crucial for agricultural practices so that the in - season interventions can be done on time and future seasons can be planned for optimum production.

The motivation for this study emerged from the potential that technology holds for revolutionizing agriculture. Vegetation Indices (VI) such as NDVI, MSAVI, NDSI can be used to implement precision agriculture (PA), which optimizes resource use, increases yield, handles risk management and so on. Additionally, the usage of ML algorithms for predicting phenological shifts in crops enables optimal agricultural planning for future seasons. Thus, the combination of remote sensing and data-driven approach can create a transformative impact on regional agricultural practices and policies.

There are some novel methods to estimate phenologic dates [3], [6], however these works mostly focus on regional based and use satellite data with low temporal and spatial resolution. In this study, phenologic dates are tried to estimate field level with high spatial and temporal resolution satellite images by using MSAVI and NDVI. For bare soil stages, NDSI used to understand soil salinity. Also, after deciding the dates, phenological stage rasters are mosaicked in between to understand field dynamics in time series manner. These mosaicked rasters are studied to create meaningful results and clustered according to the temporal data.

2. Materials and Methods

2.1. Study Site

Illinois is the top farming state in the United States and is in a region known as the Corn Belt. The main crop grown in the state is corn and soybeans. In this study, one field located at Illinois is selected. Corn is the only product grown in this field between the years 2017 and 2022. This information is taken from crop type maps published by United States Department of Agriculture - National Agricultural Service (NASS). The field is located over 40.128° to 40.142° N latitude and 88.335°W to 88.35°W longitude and covers an area of about 240 ha.

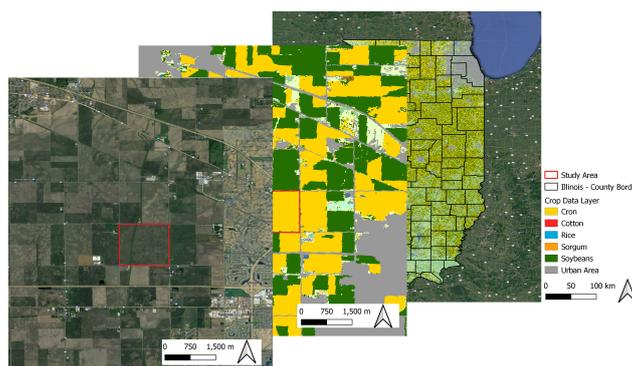


Figure 1. Study Site

2.2. Remote Sensing Data

Sentinel satellites are part of European Space Agency’s (ESA) Copernicus Program. In this study Sentinel 2 bands are used between the years 2017 and 2022.

2.3. Crop Progress Report

United States Department of Agriculture publishes Crop Progress Report (CPR) which gives information about phenological dates of crops. For this study, CPR data downloaded, interpolated, and masked by selected field area. According to the mean pixel values for every week, a table is created and crop phenological stages are written to the corresponding values. in CPR report, corn phenological stages are declared as planted, emerged, silking, dough, dent, mature and harvest (Table 1).

Table 1. Description of phenological stages for corn

Stage	Explanation ¹
Planted	When crops are planted
Emerged	When the plants can be seen above the soil
Silking	When thread-like filaments appear from the tips of the ears
Dough	When about half the kernels have indents and all kernels have a doughy substance
Dent	When every kernel is fully indented, and the ear feels solid with mostly no liquid inside the kernels
Mature	The plant is considered frost-resistant, corn is nearly ready for harvesting, the outer layers are open, and no green leaves are there
Harvest	The plant is collected from the field, either by cutting, threshing, or other means

¹ USDA - NASS

2.4. Phenology Estimation Model

In this study for detection of phenology stages, combination of threshold based and slope-based method is developed. The first step of this method is to find MSAVI and NDVI time series values. Satellite images can contain some misinformation because of atmospheric noises, clouds, and shadows. Pre-Processing steps are handled for new merged time series vegetation index data to eliminate these kinds of noises and misinformation and to smooth the time series data (Figure 2). In this study, median filter [7] and Savitzky-Golay (SG) filter [16] are used. First, median filter is applied, aiming to smooth the time series by removing observations that deviate from the local trend. The process is done by replacing each value with the median value during its five - day temporal moving window. Then, SG filter is implemented to the filtered time series data. In SG filter, there are two parameters, window length and polynomial order. The window length specifies how many neighboring points are used for the polynomial fit. The polynomial order defines complexity of the fitting polynomial.

Figure 2. Original MSAVI, original NDVI and Merged Vegetation Index Data (MSAVI bigger than 0.6 is replaced with NDVI values) (First Graph), Pre-Processing steps applied to Merged Vegetation Index (Second Graph)

After that, the dates and values of MSAVI having mean values bigger than 0.6 are replaced with NDVI values, and a new data set is created. VI thresholds are determined to estimate emerged, silking, and dough stages. With the combined VI dataset, slope based method is studied to estimate dent, mature and harvest stages. Planted stage is not estimated, instead bare soil dates are extracted from MSAVI graph to measure soil salinity.

2.5. Field Characteristics Model

Based on the crop stage dates determined with phenology estimation model, raster data belonging to those dates are mosaicked. For bare soil dates, NDSI images are merged. For emergence and silking stages, MSAVI rasters are merged, and for dough, dent, mature and harvest NDVI rasters are merged. Mosaicked are created first as year based, then the cluster information coming from the years mosaicked again. At the end, one raster layer for every stage is created and standard deviation is calculated for the pixels. The time series standard deviation layer gives information for pixel variability over time [13]. This information is important to understand field on site specific. After extracting standard deviation raster, with k-means algorithms, clusters are created. These clusters show how the field needs to be cared (Figure 4).

3. Results

After running threshold and slope based phenological detection algorithm, the results are compared to CPR dates. In Table 2, estimated and CPR based phenological dates for the year 2021 can be seen. Same table is created for the other years. For the entire study years, depending on week of the year (WOY), the phenological dates are marked with CPR dates (Figure 3).

Table 2. 2021 - Estimated dates are compared with CPR dates

2021	Threshold Based	CPR Based
Planted	-	9 April
Emerged	26 March	30 April
Silking	30 May	4 Jun
Dough	3 August	16 July
Dent	23 August	13 Augustus
Mature	7 September	27 Augustus
Harvest	27 September	17 September

The accuracy of the combined estimation method is evaluated using the RMSE (Root Mean Square Error) method. To calculate RMSE, start dates of phenological stages are turned to Day of Year (DOY). The overall RMSE for estimating corn phenological stages is calculated as 7.44 days. A stage-specific analysis reveals that the highest error occurs in the Emergence stage with an RMSE of 5.58 days, while the lowest error is found in the Silking stage with an RMSE of 1.67 days. Other stages like Harvest, Dough, Mature, and Dent have RMSE values of 2.58, 2.97, 3.51, and 3.54 days respectively. When the Emergence stage is excluded from the dataset, the overall RMSE improves to 4.61 days.

Figure 3. Phenological stages found by threshold and slope based algorithm and from CPR are marked for week of the year

For the bare soil dates which are the range between 4 February 2021 to 16 March 2021, NDSI rasters are mosaicked and clustered to understand the bare soil trends. Same process is applied for the other years (Figure 4).

Clusters are created depending on the pixel level standard deviation values. These clusters are named as object - based field characteristics for decision makers to understand field more precise (Figure 4).

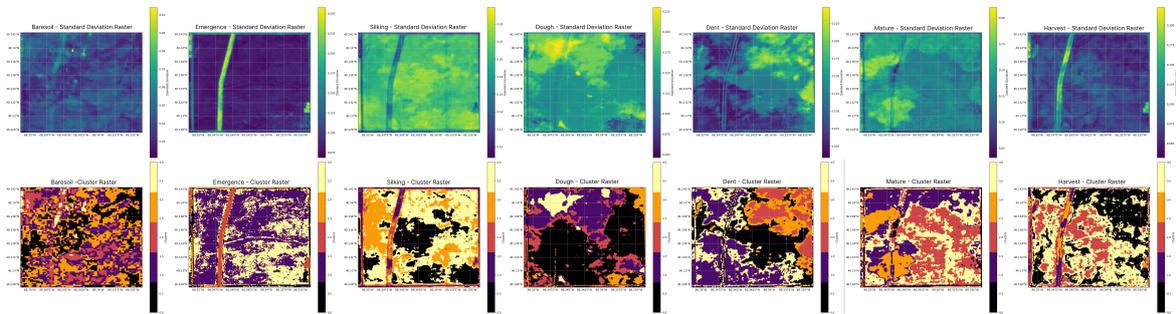


Figure 4. Standard deviations of pixel values in phenological stage rasters and results from the K-means clustering algorithm. Cluster with high number is calculated depending on high standard deviation pixels, cluster with low number is calculated depending on low standard deviation pixels.

If crop fields are well-managed, one would anticipate high variability in conditions. On the other hand, low variability suggests that the land is not being actively managed, or the land is showing a trend that the area possibly has a characteristic problem. indicating that the area should be managed differently than entire field [13]. Therefore, the clusters with high class are showing the problematic areas for different stages. Management decisions should be based on these classes.

4. Discussion

The model currently investigates corn field in Illinois, and its applicability to other crops and regions needs further study. Also, there’s room for improvement in the early stage "Emergence" predictions. Further ground observation work is needed to understand and correct the field characteristics in the field model. Future work can also focus on incorporating other factors like weather conditions and soil nutrient content to improve the model’s robustness.

5. Conclusions

This study exemplifies how the integration of GeoAI, remote sensing, and machine learning can improve precision agriculture. It not only offers a robust model for understanding and predicting crop phenological stages but also opens roads for more focused and site-specific agricultural practices. In precision agriculture, right source, right rate, right time and right place is important. In this study, the right timing is estimated through phenology analysis. Field model, as related to phenology estimation, tries to create solutions for right source, right place and right rate.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The CPR Data is downloaded from USDA - NASS site https://www.nass.usda.gov/Research_and_Science/Crop_Progress_Gridded_Layers/index.php. Sentinel 2 satellite images are downloaded form Sentinel Hub <https://www.sentinel-hub.com>.

Conflicts of Interest: The author declares no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

MSAVI	Modified Soil Adjusted Vegetation Index
NDVI	Normalized Difference Vegetation Index
VI	Vegetation Index
PA	Precision Agriculture
AI	Artificial Intelligence
GeoAI	Geospatial Artificial Intelligence
RMSE	Root Mean Square Error
NASS	National Agriculture Statistics Service
USDA	United States Department of Agriculture

References

1. Yang, Z.; Diao, C.; Gao, F. Towards Scalable Within-Season Crop Mapping With Phenology Normalization and Deep Learning. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* **2023**, *16*, 1390–1402, doi:10.1109/JSTARS.2023.3237500.
2. Guo, Bing; Yang, Fei; Fan, Yewen; Han, Baomin; Chen, Shuting; Yang, Wenna. Dynamic monitoring of soil salinization in Yellow River Delta utilizing MSAVI–SI feature space models with Landsat images. *Environmental Earth Sciences* **2019**, *78*, 308.
3. Diao, C.; Yang, Z.; Gao, F.; Zhang, X.; Yang, Z. Hybrid phenology matching model for robust crop phenological retrieval. *ISPRS Journal of Photogrammetry and Remote Sensing* **2021**, *181*, 308–326, doi:10.1016/j.isprsjprs.2021.09.011.
4. Diao, C.; Li, G. Near-Surface and High-Resolution Satellite Time Series for Detecting Crop Phenology. *Remote Sensing* **2022**, *14*(9), 1957, doi:10.3390/rs14091957.
5. Beck, P.S.A.; Atzberger, C.; Høgda, K.A.; Johansen, B.; Skidmore, A.K. Improved monitoring of vegetation dynamics at very high latitudes: A new method using MODIS NDVI. *Remote Sensing of Environment* **2006**, *100*(3), 321–334, doi:10.1016/j.rse.2005.10.021.
6. Diao, C. Remote sensing phenological monitoring framework to characterize corn and soybean physiological growing stages. *Remote Sensing of Environment* **2020**, *248*, 111960, doi:10.1016/j.rse.2020.111960.
7. Migliavacca, M.; Galvagno, M.; Cremonese, E.; Rossini, M.; Meroni, M.; Sonnentag, O.; Cogliati, S.; Manca, G.; Diotri, F.; Busetto, L.; Cescatti, A.; Colombo, R.; Fava, F.; Morra Di Cella, U.; Pari, E.; Siniscalco, C.; Richardson, A.D. Using digital repeat photography and eddy covariance data to model grassland phenology and photosynthetic CO₂ uptake. *Agricultural and Forest Meteorology* **2011**, *151*, 1325–1337, doi:10.1016/j.agrformet.2011.05.012.
8. You, X.; Meng, J.; Zhang, M.; Dong, T. Remote Sensing Based Detection of Crop Phenology for Agricultural Zones in China Using a New Threshold Method. *Remote Sensing* **2013**, *5*(7), 3190–3211, doi:10.3390/rs5073190.
9. Shen, Y.; Wu, L.; Di, L.; Yu, G.; Tang, H.; Yu, G.; Shao, Y. Hidden Markov Models for Real-Time Estimation of Corn Progress Stages Using MODIS and Meteorological Data. *Remote Sensing* **2013**, *5*(4), 1734–1753, doi:10.3390/rs5041734.
10. Cong, N.; Piao, S.; Chen, A.; Wang, X.; Lin, X.; Chen, S.; Han, S.; Zhou, G.; Zhang, X. Spring vegetation green-up date in China inferred from SPOT NDVI data: A multiple model analysis. *Agricultural and Forest Meteorology* **2012**, *165*, 104–113, doi:10.1016/j.agrformet.2012.06.009.
11. Thenkabail, P.S.; Lyon, J.G.; Huete, A., Eds. *Advanced Applications in Remote Sensing of Agricultural Crops and Natural Vegetation: Hyperspectral Remote Sensing of Vegetation Second Edition Volume IV*, 2nd ed.; CRC Press: 2018; ISBN 978-0-429-43116-6; doi:10.1201/9780429431166.
12. Gao, Feng; Zhang, Xiaoyang. Mapping Crop Phenology in Near Real-Time Using Satellite Remote Sensing: Challenges and Opportunities. *Journal of Remote Sensing* **2021**, *2021*/8379391.
13. Kanjir, Urška; Đurić, Nataša; Veljanovski, Tatjana. Sentinel-2 Based Temporal Detection of Agricultural Land Use Anomalies in Support of Common Agricultural Policy Monitoring. *ISPRS International Journal of Geo-Information* **2018**, *405*.
14. Yang, Yanjun; Ren, Wei; Tao, Bo; Ji, Lei; Liang, Liang; Ruane, Alex C.; Fisher, Joshua B.; Liu, Jianguo; Sama, Michael; Li, Zhe; Tian, Qingjiu. Characterizing spatiotemporal patterns of crop phenology across North America during 2000–2016 using satellite imagery and agricultural survey data. *ISPRS Journal of Photogrammetry and Remote Sensing* **2020**, *170*, 156–173.
15. Khan, Nasir M.; Rastokuev, Victor V.; Sato, Y.; Shiozawa, S. Assessment of hydrosaline land degradation by using a simple approach of remote sensing indicators. *Agricultural Water Management* **2005**, *77*, 96–109.
16. Chen, Jin; Jönsson, Per.; Tamura, Masayuki; Gu, Zhihui; Matsushita, Bunkei; Eklundh, Lars. A simple method for reconstructing a high-quality NDVI time-series data set based on the Savitzky–Golay filter. *Remote Sensing of Environment* **2004**, *91*, 332–344.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.