A Geospatial AI Framework for Farmer-Centered Sustainable Agricultural Zone Management

Patricia Leigh Cabañes Alegria[1], Leyla Chavez*, Madi Arnold*, Yeshi Sherpa†, Cullen McGinnis†, Landon Dierkes†, Dr. Papia Rozario*, Dr. Rahul Gomes†, Dr. Pedro Sanchez‡, Dr. Debankur Sanyal§,

[1]University of Maryland, Mathematics and Social Data Science, *University of Wisconsin-Eau Claire, Geography and Anthropology,
†University of Wisconsin-Eau Claire, Computer Science, ‡University of Arizona, Biosystems Engineering, §University of Arizona, Environmental Science

Introduction

Sustainable agriculture increasingly relies on data-driven approaches to characterize within-field variability and support efficient management decisions. This project proposes a geospatial artificial intelligence (GeoAI) framework for farmer-centered sustainable agricultural zone management through the development and evaluation of alternative Management Zone Delineation (MZD) strategies.

- The partition of fields into Management Zones (MZ) is essential for targeted fertilizer application to avoid runoff, leaching, and other channels of nonpoint source pollution.
- This is especially pertinent in arid regions constrained by extreme temperatures and moisture scarcity.

Ultimately, this work delivers an intuitive, farmer-centric decision-support tool that drives data-backed, sustainable agriculture.

Study Area

The field sites for this study are located in Yuma and Maricopa counties in Southwest Arizona.

Southwest Arizona's Geography

- The study area is within the desert basins.
- The climate is dry and arid (Köppen class: BWh). Annual precipitation is around 4.09 inches, and it sits at an elevation of 351 feet above sea level.
- This region is renowned for its vast expanse of crop farms despite being classified as a semi-arid environment not naturally suited for crop growth.

Field 1 has data for wheat and field 2 has data for lettuce. Primary spectral data is collected using an in-field optical spectrometer embedded in a probe as shown in figure 1.

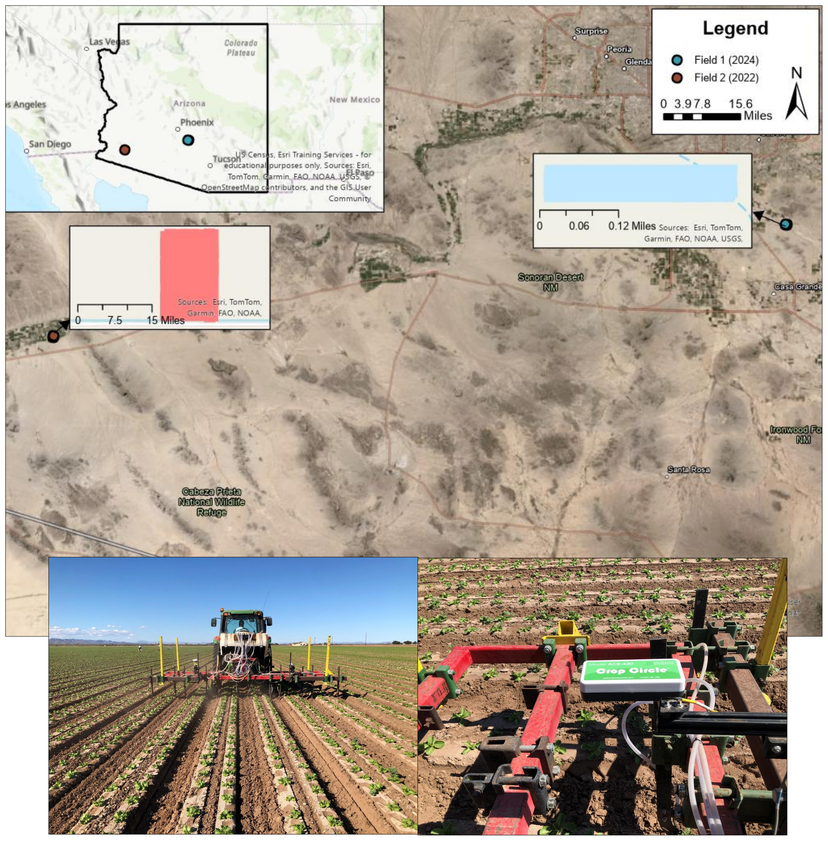


Fig. 1: Study area map and data collected using Crop Circle ACS-340, shown for field 1

Methods

Field 1 sensor data was separated into three MZD models and compared:

1. A soil-only model using on-the-go soil sensor variables, including electrical conductivity (ECa-dp), moisture, temperature.
2. A soil–terrain model that incorporates topographic variables such as elevation, slope, and curvature.
3. A soil–terrain–yield model that integrates crop performance data to capture productivity patterns.

Spatial Autocorrelation

Rasters for each soil variable were generated using a Simple Kriging model with a normal transformation.

Feature Extraction and Unsupervised Machine Learning

Principal Component Analysis (PCA) was applied for dimensionality reduction, followed by three cluster methods:

1. K-Means Clustering (KME)
2. Spatially Constrained Multivariate Clustering (SCM)
3. Iso Cluster Unsupervised Classification (ISO)

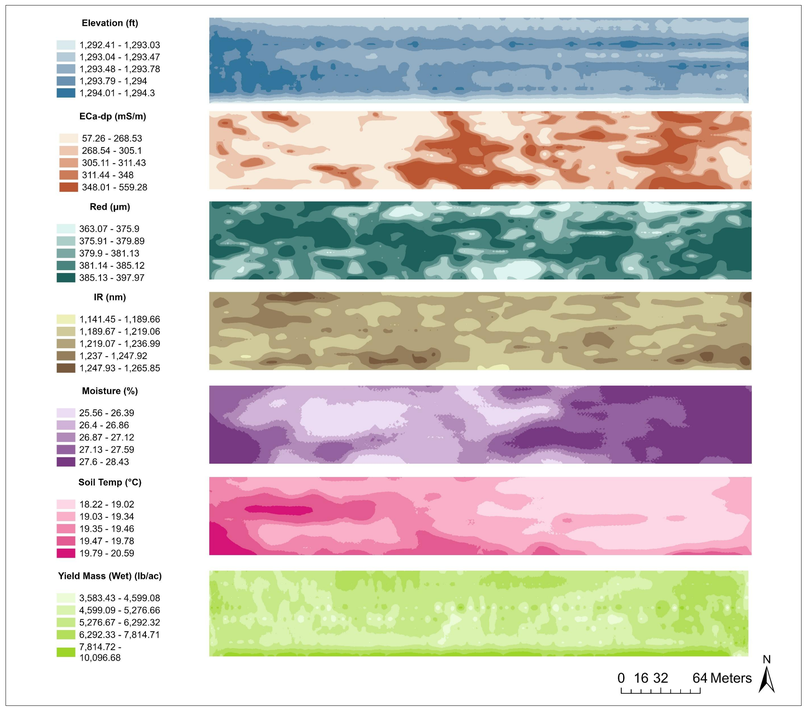


Fig. 2: Spatial maps of soil properties rendered for field 1

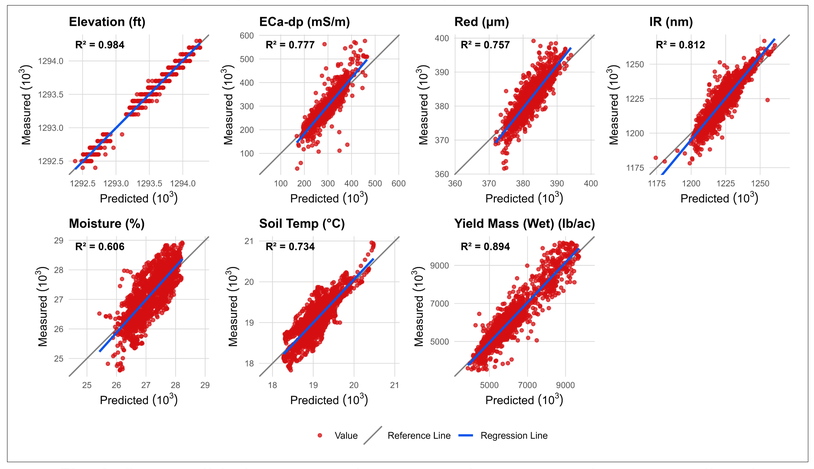


Fig. 3: Cross validation scatterplots comparing measured versus predicted soil properties

Results

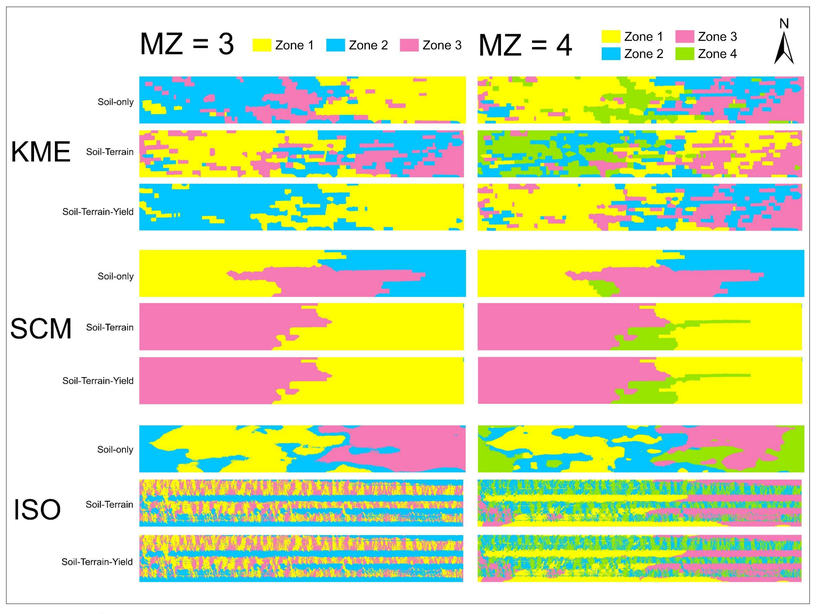


Fig. 4: MZ maps generated using various combinations of variable groupings and clustering algorithms, shown for 3 and 4 MZs

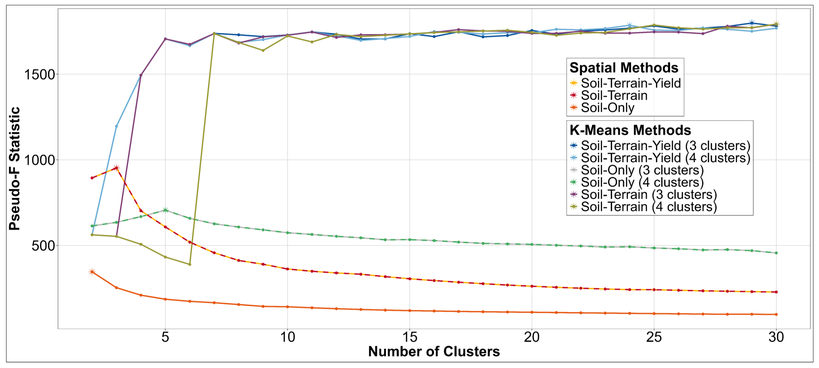


Fig. 5: Pseudo-F Statistic Chart comparing optimum number of clusters for various clustering algorithms

Note: Spatial Soil-Terrain and Spatial Soil-Terrain-Yield have identical graphs.
K-Means Soil-Only (3 Clusters) and K-Means Soil-Only (4 Clusters) have identical graphs.

Discussion

- SCM was the best classifier. Figure 4 demonstrates contiguous zones, whereas KME and ISO have significant cluster leakage.
- SCM applied on the Soil-Terrain and spatial Soil-Terrain-Yield MZD models have a small, negligible cluster on the top right of the field.

Future Work

- Perform the same pipeline on Field 2 with three MZD models: soil-only, soil-spectral, and soil-spectral-nitrates.
- Development of Ag-GPT, a local reproducible geospatial workflow for developing MZs from soil and yield point data.
- Peaks at k = 30 suggested for Soil-Terrain (3 clusters) and Soil-Terrain (4 clusters) models requires further investigation.

Acknowledgements

This research was funded by the National Science Foundation (NSF) Research Experience for Undergraduates (REU) grant OAC- 2447779 and the WiSys Sparks Grant T260013. The computational resources of this study were provided by the Blugold Center for High-Performance Computing under NSF grant CNS-1920220.

We thank the Office of Research and Sponsored Programs for supporting this research, and Learning & Technology Services for printing this poster.