Classification of Spatiotemporal Field Data for Biodiversity Estimation

John Salako[1, 2], Francesco Lodato[1], Bruno Basso[1]

(1) Earth and Environmental Sciences Department, Michigan State University

(2) Computational Mathematics, Science, and Engineering Department, Michigan State University

AGU24 Washington, D.C. | 9–13 December 2024

1. Summary

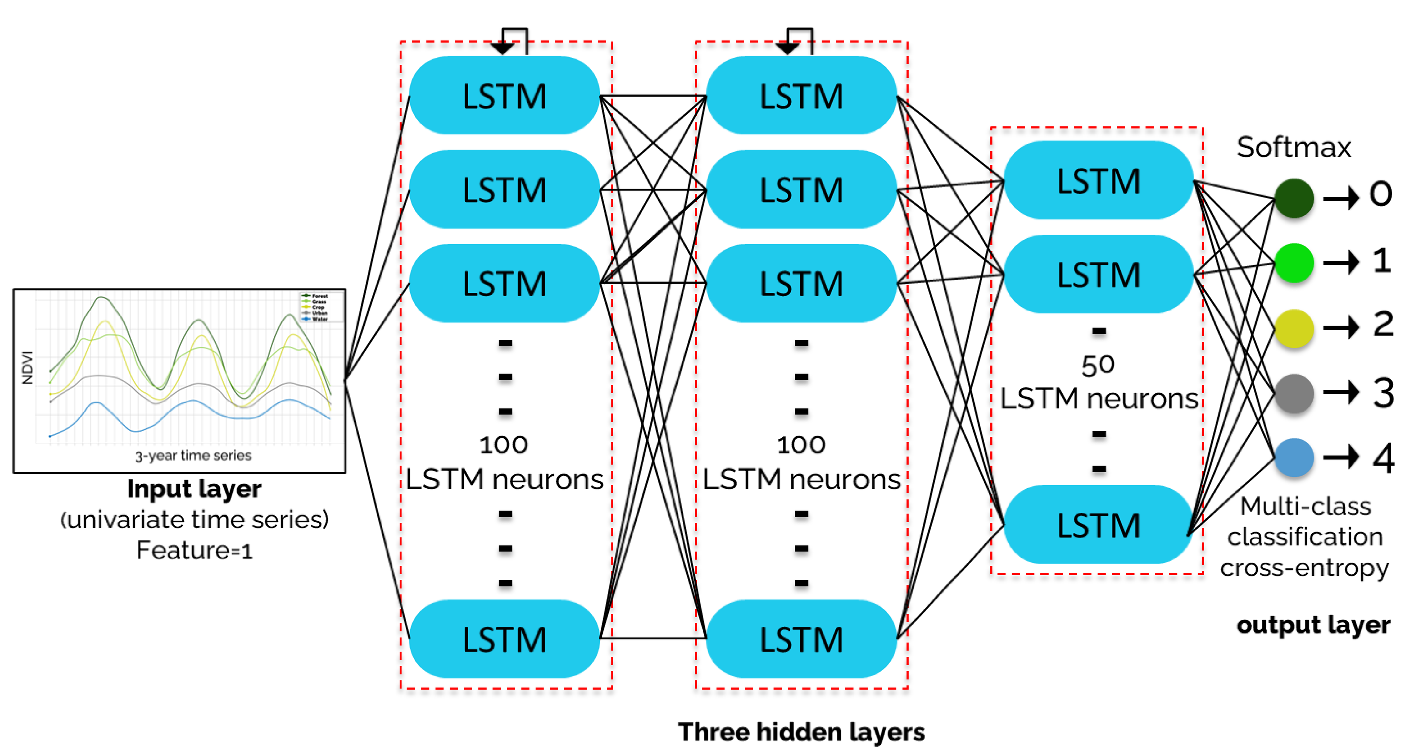
- Biodiversity estimation is critical for monitoring changes in land use and ecosystems.
- Our study combined remote sensing and analytical frameworks for Biodiversity monitoring and estimations
- We focused on identifying and quantifying buffer zones near agricultural fields in US Midwestern states (Illinois, Indiana, Iowa, Michigan, Minnesota, Missouri, Ohio, and Wisconsin)
- Crop, water, forest, grass, and urban categories were defined.
- Over 112,000 spatial data points were collected for model training. Aggregated monthly Sentinel-2 optical band timeseries data spanning three years were used.
- Vegetative indices, such as NDVI, were indicative for identifying unique signatures for each category.
- Long-Short-Term Memory (LSTM) networks achieved >95% classification accuracy.

2. Objective

- Use geo-referenced data points and Sentinel-2 satellite imagery to train machine-learning models to classify land cover types..
- Identify and estimate buffer zone areas at the edges of Corn and Soybean fields in the US Midwest

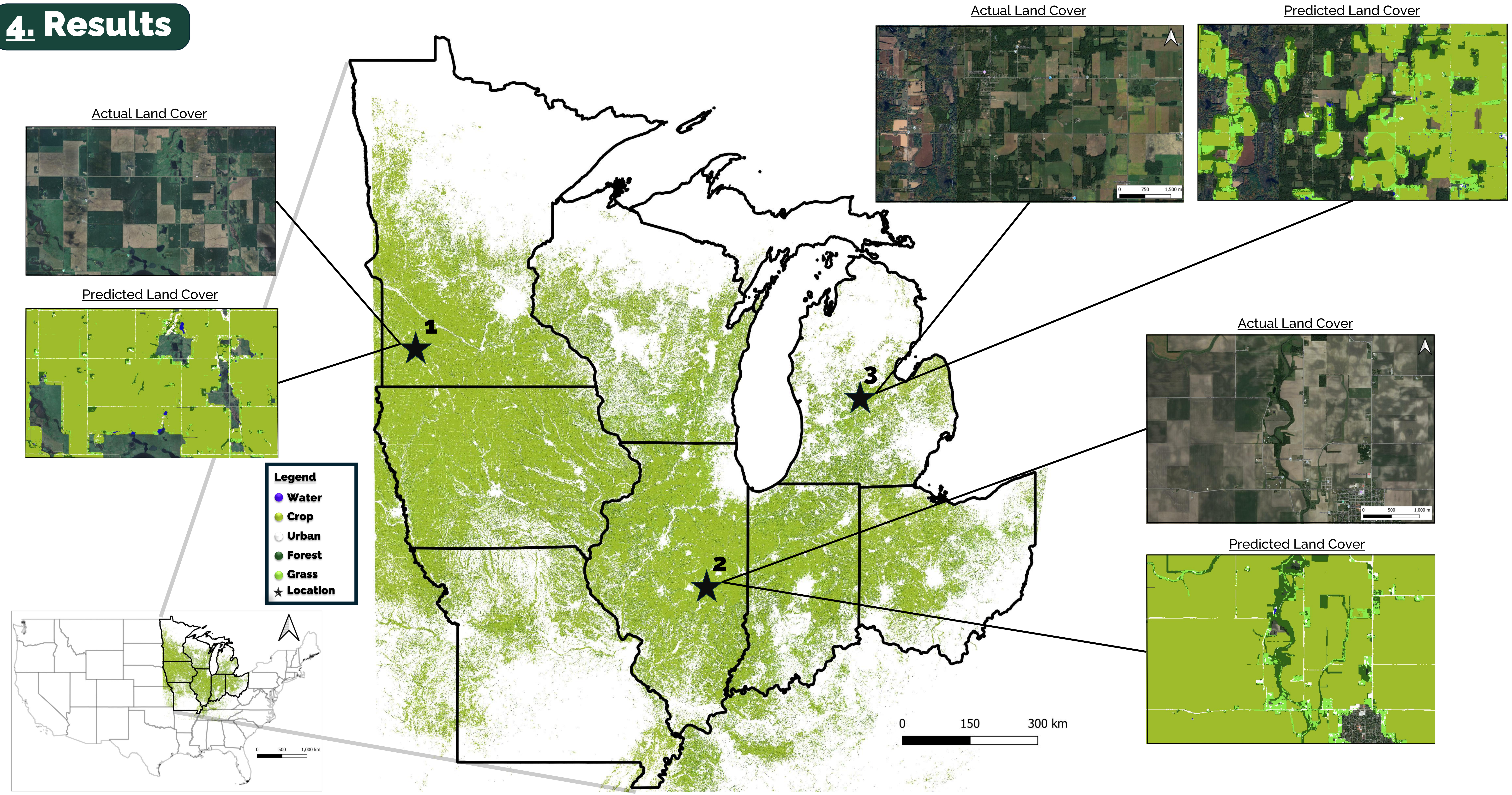
3. Methodology

- Corn and Soybean fields were extracted using the crop data layer (CDL) and downloaded as masks.
- A 100-meter buffer was added to the processed masks to capture biodiversity zones at the field edges.
- Five classes selected (112,350 georeferenced points): Water (6,029), Urban (17,200), Crop (33,804), Forest (42,699), Grass (12,618).
- Optical band time series data (01/21 – 12/23) from Sentinel-2 were downloaded and processed.
- LSTM models were used to train the NDVI data.
- NDVI provided the most accurate results.



LSTM neural network architecture

4. Results



Map of the US

SPATIOTEMPORAL CLASSIFICATION OF FIELD DATA, HIGHLIGHTING MILLIONS OF FIELDS ACROSS US MIDWEST

Three locations (above) were selected to demonstrate our model performance. The results highlight the capability of the LSTM model to capture individual land use classes across all states, allowing us to calculate and estimate buffer zones around the field edges.

State (Region)	Water (acres)	Crop (acres)	Urban (acres)	Forest (Buffer Zones) (acres)	Grass (acres)
Minnesota (1)	29.7	8968.9	323.3	246.38	298.1
Illinois (2)	0.96	3706.9	93.1	254.1	201.4
Michigan (3)	31.6	4911.3	54.6	1227.2	1009.9

Table showing the area of the different land cover of the selected regions

5. Conclusion

- We highlight the advancement of remote sensing for biodiversity monitoring, by demonstrating how integrating AI models with big data processing can efficiently analyze large datasets, uncover patterns, and make predictions.
- Timeseries NDVI data effectively captured the seasonality of agricultural fields, distinguishing Grass and Crop classes.
- LSTM models achieved over 95% classification accuracy, proving effective for land use monitoring and buffer zone estimation.
- The integration of remote sensing data and LSTM-based models offers a powerful framework for biodiversity estimation and land use analysis.