Learning county from pixels: Corn yield prediction with attention-weighted multiple instance learning

Xiaoyu Wang[1], Yuchi Ma[2], Qunying Huang[3], Zhengwei Yang[4], and Zhou Zhang[1*]

[1]Department of Biological Systems Engineering, University of Wisconsin-Madison

[2]Department of Earth System Science and Center on Food Security and the Environment, Stanford University

[3]Department of Geography, University of Wisconsin-Madison

[4]Research and Development Division, National Agricultural Statistics Service, United States Department of Agriculture

Biological Systems Engineering
UNIVERSITY OF WISCONSIN–MADISON

Introduction

- County-level corn yield prediction in the U.S. holds significant importance due to its central role in the country's agriculture and economy.
- Many previous studies treat a county as a unified entity, aggregating all pixels within the county to a single value (e.g. mean value).
- This study was designed to (1) examine each county at the pixel level and apply multiple instance learning to leverage detailed information within a county. (2) solve the "mixed pixel" problem by employing attention mechanism to automatically assign weights to different pixels, which can mitigate the influence of mixed pixels.

Materials and Methods

Data Acquisition

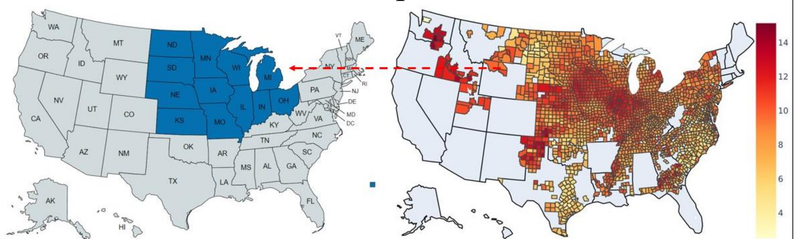


Figure 1. Study area

- **Study area:** Twelve corn belt states in the United States, including North Dakota, South Dakota, Minnesota, Wisconsin, Iowa, Illinois, Indiana, Ohio, Missouri, Kansas, Nebraska and Michigan.
- **Satellite data:** Three Vegetation indices including Green Chlorophyll Index (GCI), Enhanced Vegetation Index (EVI), and Normalized Difference Water Index (NDWI).
- **Weather data:** Daily mean air temperature (Tmean), maximum air temperature (Tmax), minimum air temperature (Tmin), maximum Vapor Pressure Deficit (VPDmax), minimum Vapor Pressure Deficit (VPDmin), and total precipitation (PPT).
- **Soil data:** Available Water Holding Capacity (AWC), Soil Organic Matter (SOM) and Cation Exchange Capacity (CEC).
- **Other data:** 5-year historical average yield and years.

"Mixed pixel" problem

(a) One MODIS pixel (b) One MODIS pixel masked by CDL

Figure 2. Illustration of "mixed pixel" problem

- MODIS dataset: satellite imagery with 500m resolution.
- CDL mask: Crop mask with 30m resolution.
- Mixed pixel: One MODIS pixel contains both crop field and other land types, which may introduce some noise into predictive model.

Methodology and metrics

- **Traditional method:**
  Treat each county as an instance for analyzing.
  Multiple instance learning (MIL):
  Treat each county as a bag and pixels in the county as many instances in a bag.
- **Our**
  Attention-MIL
- **Baseline**
  Instance-MIL
  Random forest
  Linear regression
  Ridge regression

Evaluation metrics:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{n}}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)_2}{\sum_{i=1}^{n}(y_i - \bar{y})^2}$$

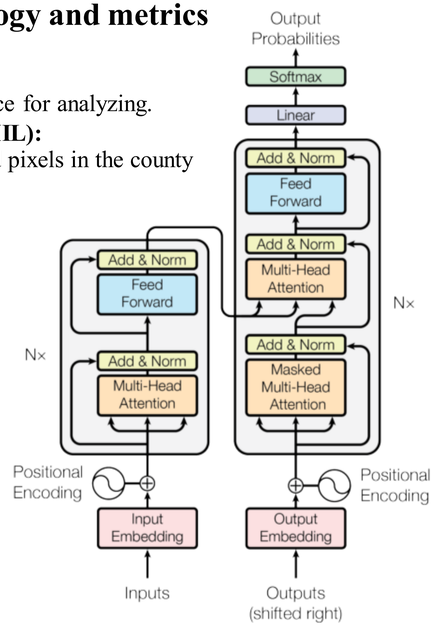


Figure 3. Attention module

Pipeline: multiple instance learning for county imagery

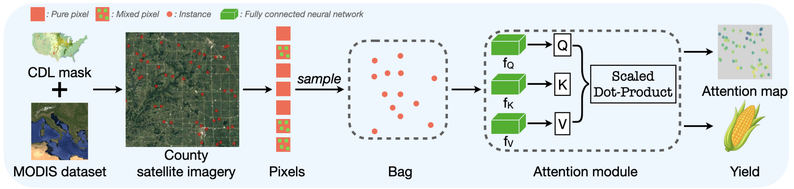


Figure 4. Pipeline of our method

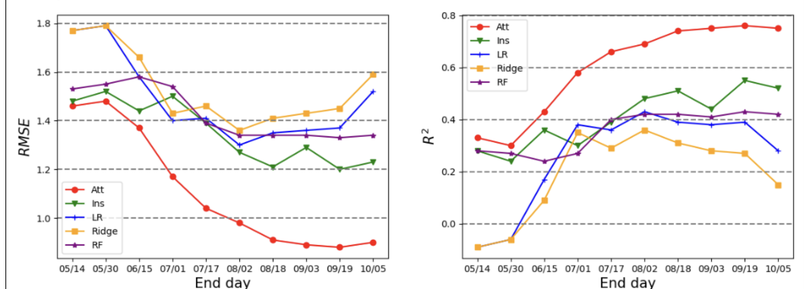
Results and Discussion

$RMSE$ Year \ Method	Att	Ins	LR	Ridge	RF
2018	1.04	1.42	1.49	1.49	1.41
2019	0.86	1.18	1.46	1.45	1.60
2020	0.87	1.35	1.53	1.81	1.41
2021	0.86	1.16	1.78	2.04	1.08
2022	0.88	1.07	1.35	1.18	1.19

R^2 Year \ Method	Att	Ins	LR	Ridge	RF
2018	0.73	0.50	0.45	0.45	0.50
2019	0.74	0.51	0.26	0.27	0.10
2020	0.67	0.20	-0.02	-0.42	0.13
2021	0.77	0.60	0.05	-0.24	0.64
2022	0.85	0.73	0.65	0.73	0.73

Figure 5. Model evaluation results in 2018-2022. We show their RMSE and R^2 in each year.

- The results show that our approach outperforms all other methods across all years and metrics.



(a) $RMSE$ (b) R^2

Figure 6. RMSE and R^2 in in-season county-level corn yield prediction in 2022.

Conclusion and Future work

- MIL is employed to leverage the pixel-level remote sensing observations, resolve the conflict between computational resources and information integrity, and address the lack of finer-grained yield records for pixel-level data processing.
- To tackle the mixed pixel problem caused by inconsistent resolutions among feature datasets and crop mask, an attention mechanism is incorporated to assign weights to pixels, thereby enhancing prediction accuracy.