Evaluating Climate Change Effects on Agricultural Yield Using Deep Learning

Raven Mott[1], Calvin Kamara[2], Firdous Kausar[3]
[1]Department of Computer Science ,College of Engineering ,Virginia State University, Petersburg, VA,
[2]Department of Computer Science , Bowie State University, Bowie, MD,
[3]Department of Computer Science and Data Science, SACS, Meharry Medical College, Nashville, TN

INTRODUCTION & MOTIVATION

- Crop yields swing with temperature, precipitation, humidity, and extreme events—traditional models miss these nonlinear, multi-variable effects.
- Escalating climate volatility makes reliable yield forecasts essential for global food security.
- We apply deep learning and ensemble ML (Random Forest, XGBoost, Gradient Boosting, LSTM) to a custom dataset combining USDA county-level yields (2017-2022) with WRF-HRRR climate simulations.
- Study covers four key U.S. crops—corn, cotton, soybeans, winter wheat—at county-level resolution nationwide.
- Inputs include raw weather variables plus engineered stress indicators such as hot-day count and drought duration.

OBJECTIVES

- Merge USDA county-level yields (2017-2022) with WRF-HRRR high-resolution weatherfields.
- Build monthly indicators—heat-day streaks, rain days, drought runs—to capture extremeevents.
- Test tree-based ensembles (RF, XGBoost, GBM) against a sequence model (LSTM) forcounty-level yield.
- Pinpoint the climate variables that most influence predictions for each crop and region.
- Release a reproducible pipeline that growers, analysts, and policy-makers can adapt tofuture climate-change scenarios.

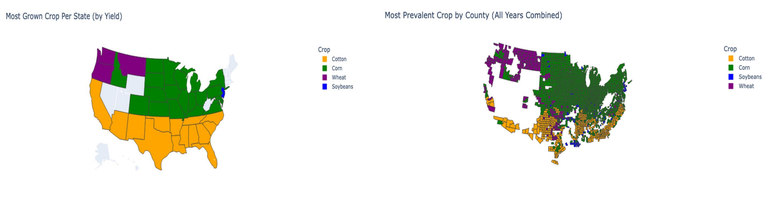
METHODOLOGY

Dataset Description:

- USDA Crop Yield Data (2017–2022): Corn, cotton, soybeans, winter wheat
- WRF-HRRR Climate Simulations: Temperature, rainfall, wind, humidity, radiation

Background: U.S. Crop Distribution (2017–2022)

- We analyze USDA yield data across all U.S. counties and states. Corn, cotton, wheat, and soybeans are the most prevalent crops. The maps below show the dominant crop by yield for each region.



Preprocessing:

- Cleaned and renamed yield columns
- Combined yearly files; calculated monthly climate indicators (heat days, rain, drought)

Feature Engineering:

- Encoded state/county/month
- Imputed missing data
- Standardized features

Integration:

- Merged yield and climate data by FIPS/month
- Produced monthly panel dataset

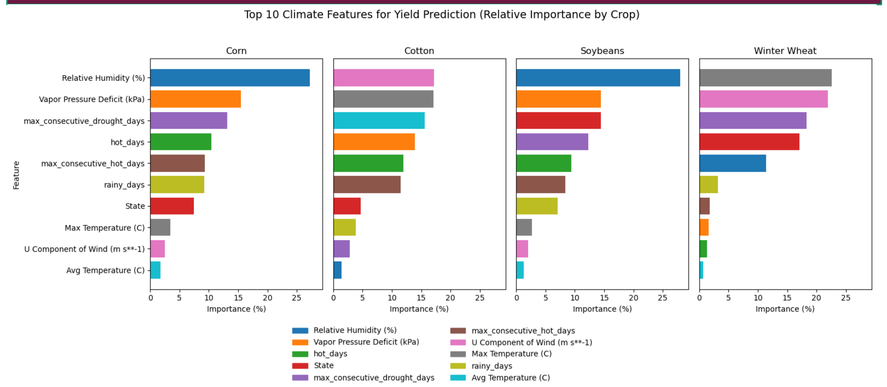
Modeling:

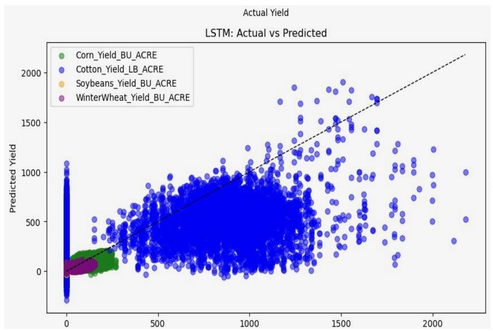
- Random Forest, XGBoost, GBM, LSTM
- Predicted crop yield per county; LSTM used for sequence learning

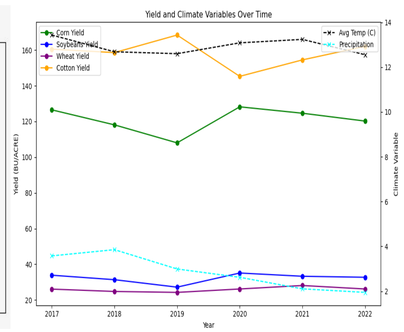
Evaluation:

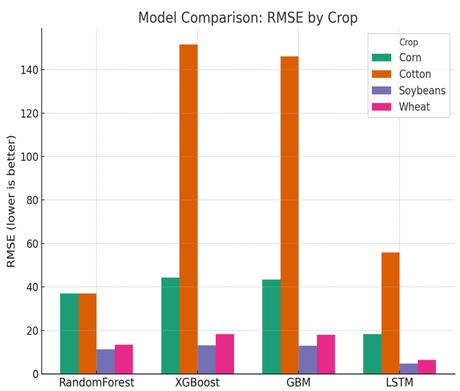
- Metrics: RMSE, MAE, R^2
- Cross-validation & train/test splits

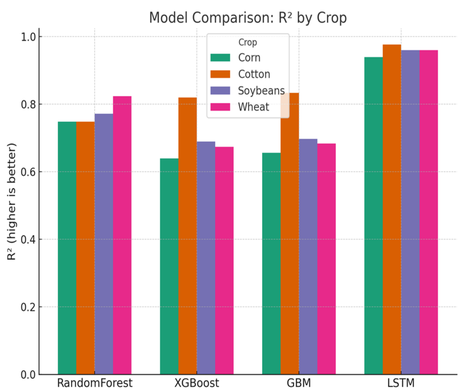
RESULTS AND OUTCOMES











Method	Corn RMSE ↓ / R^2 ↑ / Corr ↑	Cotton RMSE ↓ / R^2 ↑ / Corr ↑	Soybeans RMSE ↓ / R^2 ↑ / Corr ↑	Winter Wheat RMSE ↓ / R^2 ↑ / Corr ↑
RandomForest	37.03 / 0.748 / 0.869	37.03 / 0.748 / 0.869	11.25 / 0.772 / 0.881	13.41 / 0.824 / 0.911
XGBoost	44.38 / 0.639 / 0.804	151.50 / 0.820 / 0.908	13.13 / 0.689 / 0.832	18.26 / 0.674 / 0.827
GBM	43.34 / 0.656 / 0.816	146.06 / 0.833 / 0.915	12.95 / 0.697 / 0.838	18.01 / 0.683 / 0.834
LSTM	**18.29 / 0.939 / 0.970**	**55.83 / 0.976 / 0.988**	**4.73 / 0.960 / 0.980**	**6.41 / 0.960 / 0.980**

DISCUSSION

- **LSTM leads by a wide margin.**
  Sequence learning cuts RMSE 40-60 % relative to ensembles and lifts R2R^2R2 above 0.93 for all four crops. The gain is largest for cotton, whose yield responds to multi-week heat spells captured only by the temporal model.
- **Tree models aren't bad—just static.**
  RF, XGB, and GBM track corn and wheat reasonably well but miss yield swings tied to successive hot or wet months, highlighting the value of temporal context.
- **Feature importance matches agronomy.**
  Relative humidity and vapor-pressure deficit dominate soybean yield, while long drought streaks and max temperature govern cotton—consistent with field-trial literature.
- **Some crops remain tough.**
  Cotton's RMSE stays high (> 55 lb / acre) because ginning out-turn, pests, and irrigation practices aren't in the dataset. Future work should fold in management variables and remote-sensing vegetation indices.

CONCLUSION

- We built the **first county-level U.S. yield predictor** that merges USDA yields with sub-daily WRF-HRRR weather and engineered stress metrics.
- **LSTM outperforms** Random Forest, XGBoost, and GBM across all crops, proving that temporal dynamics matter for yield under climate variability.
- Key climate drivers differ by crop—insight that can steer targeted adaptation (e.g., drought-tolerant cotton in the Southeast, humidity management for soybeans).

ACKNOWLEDGEMENTS

This research was made possible through the support of the **MS-CC Undergraduate Summer Research Internship Program**, in partnership with **Fisk University** and **Meharry Medical College**. We would like to thank our mentos and research coordinators for their guidance throughout the project. Special appreciation goes to the developers and curators of the **USDA Crop Dataset** and **WRF-HRRR climate simulations dataset**, which formed the foundation of this study.

REFERENCES

1. Hu, T., Zhang, X., Khanal, S., Wilson, R., Leng, G., Toman, E. M., ... & Zhao, K. (2024). Climate change impacts on crop yields: A review of empirical findings, statistical crop models, and machine learning methods. *Environmental Modelling & Software*, *179*, 106119.
2. CropNet/CropNet dataset (2017–2022)—comprising Sentinel-2 imagery, WRF-HRRR computed data, and USDA crop yields—hosted on Hugging Face.
3. USDA (United States Department of Agriculture). County-level crop yield and production data for corn, cotton, soybeans, and winter wheat (2017–2022).
4. WRF-HRRR (Weather Research & Forecasting-based High-Resolution Rapid Refresh) climate model outputs, daily & monthly meteorological variables (2016–2022).
5. Iqbal, N., Shahzad, M. U., Sherif, E. S. M., Tariq, M. U., Rashid, J., Le, T. V., & Ghani, A. (2024). Analysis of wheat-yield prediction using machine learning models under climate change scenarios. *Sustainability*, *16*(16), 6976.

CONTACT

[1]Mottravenbusiness@gmail.com
[3]firdous.kausar@mmc.edu