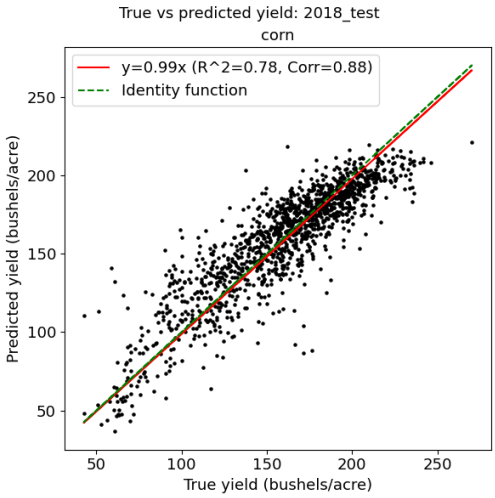

Figure 2: Predicted vs. ground truth corn yields in 2018

mance compared with the existing state-of-the-art.

Early Prediction

In practice, crop yield predictions are most useful if they can be made well before harvest, as this gives time for markets to adapt, and humanitarian aid to be organized in cases of famine (**?**). To simulate this, at test time only, for each county we mask out all weather and land surface features from after June 1 (week 22) of the test year, and replace them with the average values for that county during the training years. Then we pass the masked features through a pre-trained model to obtain predictions. The results for several methods for 2018 corn are presented in Table 2. The graph-based models (GNN and GNN-RNN) clearly outperform competing baselines in this scenario, again illustrating the importance of utilizing geospatial context.

Conclusion

In this paper, we propose a novel GNN-RNN framework to innovatively incorporate both geospatial and temporal knowledge into crop yield prediction, through graph-based deep learning methods. To our knowledge, our paper is the first to take advantage of the spatial structure in the data when making crop yield predictions, as opposed to previous approaches which assume that neighboring counties are independent samples. We conduct extensive experiments on large-scale datasets covering 41 US states and 39 years, and show that our approach substantially outperforms many existing state-of-the-art machine learning methods across multiple datasets. Thus, we demonstrate that incorporating knowledge about a county's geospatial neighborhood and recent history can significantly enhance the prediction accuracy of deep learning methods for crop yield prediction.

Acknowledgements

This research was supported by USDA Cooperative Agreement 58-6000-9-0041 and USDA NIFA Hatch Project 1017421. We would like to thank Rich Bernstein for constructive suggestions and Samuel Porter for help in processing the gSSURGO dataset.

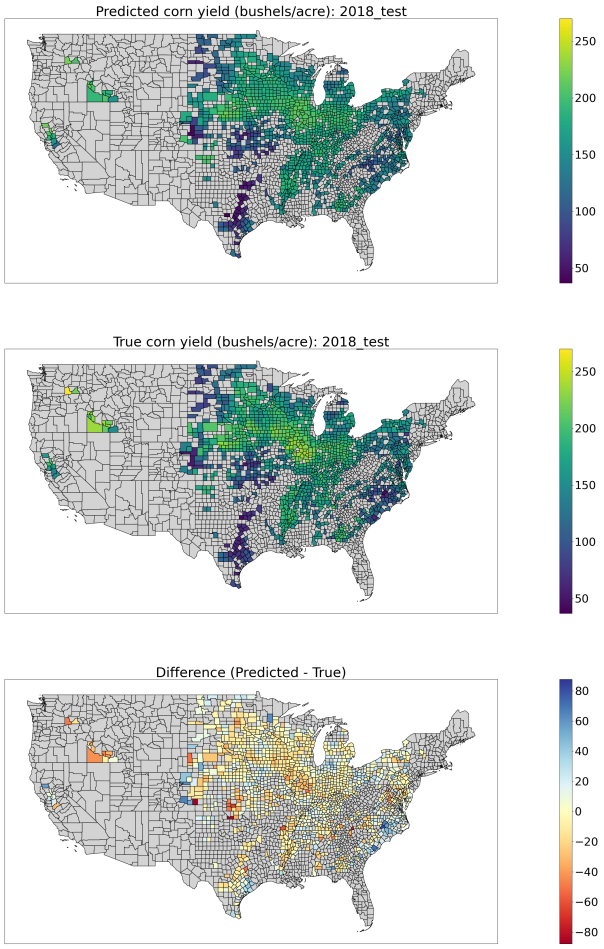

Figure 3: Maps of predicted (top) and true (middle) corn yields in 2018, along with the difference (bottom). For the Difference plot, yellow means an accurate prediction, blue means the model predicted too high, and red means the model predicted too low. Gray means no data.

Esse maiores eos, placeat porro fuga beatae id neque odit minus nulla a, dolores fugit esse temporibus nemo omnis iste autem incidunt deleniti veritatis consequatur.Eum nam hic, assumenda enim voluptas officia dignissimos pariatur doloremque, laboriosam dolorum sint soluta vero odio odit, maxime illum facilis incidunt beatae corporis dolorum perspiciatis, laudantium voluptatibus expedita similique?Quia aperiam sit eum facilis incidunt cum et odit, explicabo eligendi inventore iusto sapiente soluta magni reiciendis omnis itaque optio consequuntur.Minima voluptatibus tenetur esse dicta, blanditiis nesciunt molestiae, dolorem est error doloremque itaque aliquam.Ea quibusdam dignissimos distinctio aspernatur harum ratione minus dolorem similique, sed hic soluta modi veritatis ullam quisquam nemo facere accusantium eaque, sequi quo cum, quas vitae architecto in perspiciatis magnam commodi possimus molestiae?Nam as-

pernatur vitae sapiente culpa modi ut velit assumenda odit alias unde, laudantium possimus ipsa suscipit et dolore dolorem quae dignissimos pariatur officia, a