Utilizing machine learning to understand the future of corn prevented planting under climate change

Sophia Horigan[1], Haynes Stephens[2], Samantha Lapp[2], Raphael Rossellini[3], Eva Burgos[2], Dr. Liz Moyer[2]

[1]Department of Ecology and Evolution, University of Chicago; [2]Department of Geophysical Sciences, University of Chicago; [3]Department of Statistics, University of Chicago

Prevented planting is when farmers are unable to plant crops due to unfavorable conditions, resulting in empty fields, lost yields, and insurance payouts. Prevented planting has been steadily **increasing** over time, particularly for corn, the top crop in the United States.

Most prevented planting is due to **excess soil moisture**. But there are many facets of weather that make up soil moisture, including rainfall, temperature, humidity, and soil characteristics. It is therefore difficult to pinpoint exactly what weather patterns cause prevented planting.

This complexity makes it difficult to how prevented planting will change under the future weather that will result from **climate change**.

We used a random-forest model to:
1) tease apart the main weather patterns that drive the prevented planting of corn, and
2) predict the future of prevented planting under future weather.

A **random-forest model** is a supervised machine learning algorithm that averages the output of many decision tree to estimate the relationship between **features** and a **response**. In this case, our features are weather and soil conditions, and the response is the fraction of acres that are prevented. There are two main steps to use this model: **training** and **predicting**.

Model Training

We trained the model on county-level historical prevented planting and contemporary weather and soil conditions (fig 1). Although the dynamics underlying the model are a black box, an examination of the ranked feature importance indicates the relative importance of each weather/soil condition in driving prevented planting (fig 2).

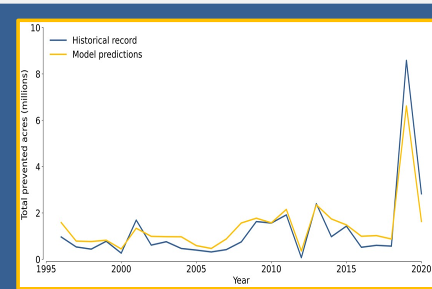

Figure 1: Model fit to historical prevented planting.

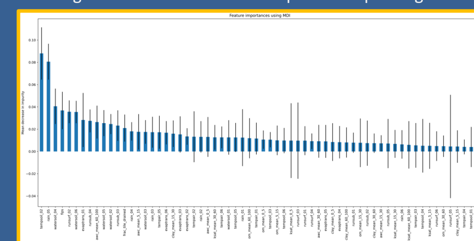

Figure 2: Ranked feature importance.

Key Takeaways

1. Overall, the model captures historical trends well.
2. The model fails to predict the full extent of the 2019 peak, indicating that it may underpredict future extremes.
3. February soil temperature and May rainfall appear to most strongly drive prevented planting.

Model Predictions

Once we have trained the model, we can use it to predict the future of prevented planting. To do this, we input future weather data from climate projections into the model and simulate the fraction of prevented acres per county.

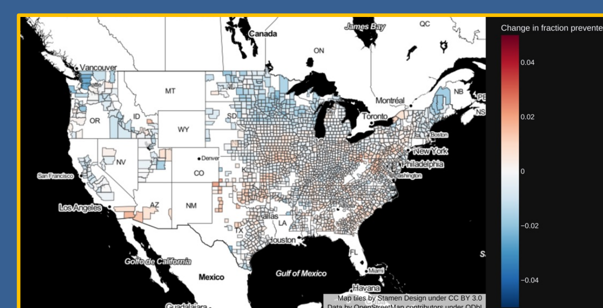

Figure 3: Projections of changes in county-level fraction of prevented acres from years 2020-2030 to years 2090-2100.

Key Takeaways

1. Generally, prevented planting will decrease in the north and increase in the south.
2. There is significant variation between climate models, increasing uncertainty.

Acknowledgements: Dr. Emily Padston, Jeff McMahon, Jacob Drucker, Ander Feistritzer, Giorgio Sarro

This material is based on work supported by the National Science Foundation under Grant No. 1735359.