U.S. soybean, corn yields could be increased through use of machine learning

September 29 2021

Research supported by a plant pathologist in Penn State's College of Agricultural Sciences suggests that machine-learning algorithms could increase corn and soybean yields. Credit: Pexels

Research guided by a plant pathologist in Penn State's College of Agricultural Sciences suggests that machine-learning algorithms that are programmed to recognize changing weather patterns could show producers and agricultural managers how to increase soybean and corn yields in the United States.

The approach could prove valuable in addressing climate change realities that have presented challenges in growing enough food for a rising global population, noted Paul Esker, associate professor of epidemiology and field crop pathology.

"Soybean and corn are among the most valuable crops in terms of food supply and economic output in the U.S. agricultural sector," said Esker, who pointed to U.S. Department of Agriculture statistics that place corn as the most widely produced crop in the U.S., with soybean following close behind.

Not only are these crops vital to food security in the U.S. and beyond, but their combined total value to the nation's economy is more than $100 billion. While Esker acknowledges that is an impressive figure, he points out that many scientists predict that that by 2050, the world must feed 9 billion people, so current outputs must increase.

"The demand for food will be 60% greater than it is today, so it is imperative that we advance agricultural research to develop more efficient—and sustainable—methods to increase production, especially when confronting climate change," he said.

A sustainable solution to this challenge is to increase crop yield without massive cropland area expansion. This, he maintains, can be achieved by identifying and adopting best management practices. However, that will require a more detailed understanding of how crop yield is influenced by climate change and growing-season weather variability, among other

factors.

Growers have access to a vast amount of crop-yield and management information that has emerged from ongoing agricultural experiments conducted across the U.S. Even with that knowledge, prediction is challenging because various factors interact with each other, noted project lead Spyridon Mourtzinis, of Agstat Consulting, Athens, Greece

"Such datasets have, to date, remained disconnected from each other and are difficult to combine, standardize and properly analyze," Mourtzinis said. "For example, variability in soil type can interact with weather conditions and mitigate or aggravate climate-related impacts on crop yields."

To overcome these challenges, the team explored the idea of leveraging the power of machine learning algorithms. A machine learning algorithm, Esker explained, is the method by which the computer learns from the data. The goal is to predict output values from given input data.

To that end, the team pored over crop yield and management data from variety trials conducted in 28 states between 2016 and 2018 for corn and between 2014 and 2018 for soybeans. Information was arranged based on soil type and management practices including irrigation, tillage method, seeding rate, row spacing and cultivar maturity.

Additionally, weather data were retrieved for each year for each set of coordinates. The team examined correlations among the weather variables and determined seven weather variables for corn and eight for soybean for specific regions.

For each crop, the researchers developed algorithms, or data sets, which they applied to various simulations and tested over two growing seasons in a randomly chosen field in southcentral Wisconsin. Boxplots, a

standardized way of displaying the distribution of data, were used to visually evaluate the results.

For that specific field and cropping system, use of the team's algorithms showed that corn with a May 1 sowing date realized a 6% increase in yield compared to corn planted on June 1. By creating scenarios with 256 background cropping system choices, the resultant algorithm-derived yield estimate difference for the same sowing date was smaller but still positive.

In the case of soybean, a May 1 sowing resulted in a 14% increase in yield compared to a June 1 sowing in the single-background cropping system. The result was consistent when yield differences due to sowing date were averaged across 128 background cropping system choices.

Mourtzinis and Esker acknowledged that there are limitations in the results, primarily due to the lack of information on rates of seed treatments, products applied to leaves and different management practices reported across different states.

Still, their results, published in *Scientific Reports*, suggest that the developed algorithms have the potential to lead the agriculture industry to substantial yield increases. "Our approach can accelerate agricultural research, identify sustainable practices and help overcome future food demands," Mourtzinis said.

The researchers also emphasize that machine learning algorithms should not be considered as a substitute of replicated trials. To the contrary, they pointed out, field trials performed by universities are needed as a source of unbiased data, which can be used to train even more comprehensive algorithms.

Also contributing to the research were James Specht, Department of

Agronomy and Horticulture, University of Nebraska-Lincoln, and Shawn Conley, Department of Agronomy, University of Wisconsin-Madison.

More information: Spyridon Mourtzinis et al, Advancing agricultural research using machine learning algorithms, *Scientific Reports* (2021). DOI: 10.1038/s41598-021-97380-7