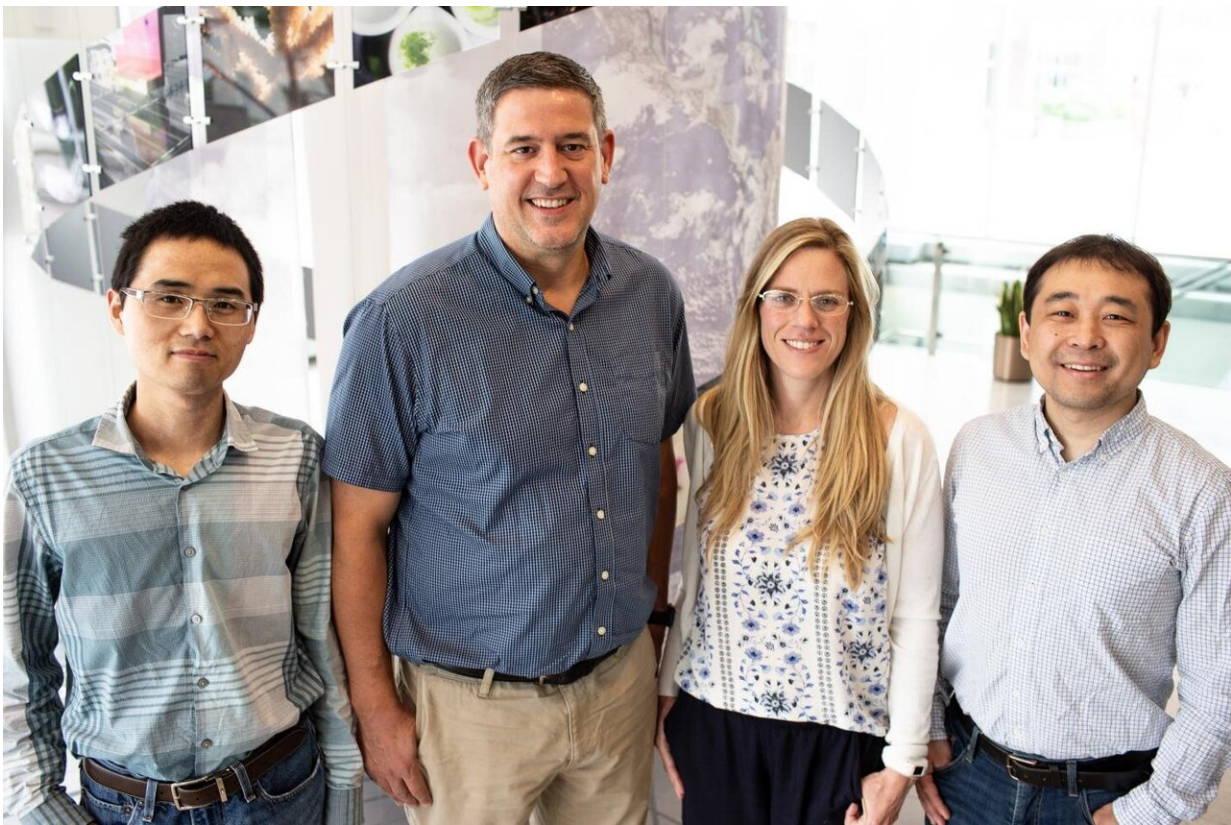


Scientists stack algorithms to improve predictions of yield-boosting crop traits

June 3 2019



From left to right: Postdoctoral Researcher Peng Fu, USDA-ARS Scientist Carl Bernacchi, Postdoctoral Researcher Katherine Meacham-Hensold, and Assistant Professor Kaiyu Guan. Credit: Claire Benjamin/RIPE Project

Hyperspectral data comprises the full light spectrum; this dataset of

continuous spectral information has many applications from understanding the health of the Great Barrier Reef to picking out more productive crop cultivars. To help researchers better predict high-yielding crop traits, a team from the University of Illinois have stacked together six high-powered, machine learning algorithms that are used to interpret hyperspectral data—and they demonstrated that this technique improved the predictive power of a recent study by up to 15 percent, compared to using just one algorithm.

"We are empowering scientists from many fields, who are not necessarily experts in [computational analysis](#), to translate their enormous datasets into beneficial results," said first author Peng Fu, a postdoctoral researcher at Illinois, who led this work for a [research project](#) called Realizing Increased Photosynthetic Efficiency (RIPE). "Now scientists do not need to scratch their heads to figure out which [machine learning](#) algorithms to use; they can apply six or more algorithms—for the price of one—to make more accurate predictions."

RIPE, which is led by Illinois, is engineering crops to be more productive by improving photosynthesis, the natural process all plants use to convert sunlight into energy and yields. RIPE is supported by the Bill & Melinda Gates Foundation, the U.S. Foundation for Food and Agriculture Research (FFAR), and the U.K. Government's Department for International Development (DFID).

In a recent study, published in *Remote Sensing of Environment*, the team introduced spectral analysis as a means to quickly identify photosynthetic improvements that could increase yields. In this new study, published in *Frontiers in Plant Science*, the team improved their previous predictions of photosynthetic capacity by as much as 15 percent using machine learning, where computers automatically applied these six algorithms to their [dataset](#) without human help.

"I've loved seeing what's possible when you can use [computational power](#) to exploit the data for all its worth," said co-author Katherine Meacham-Hensold, a RIPE postdoctoral researcher at Illinois, who led the previous study in *Remote Sensing of Environment*. "It's exciting to see what a data analyst like Peng can do with my data. Now other non-data-analyst scientists can test several powerful algorithms to figure out which one will help them leverage their data to the fullest extent."

However, more studies are needed to prove the relevance of this stacked [algorithm](#) technique to the plant science community and other fields of study.

"By applying the expertise of data analysts to address the needs of plant physiologists like myself, we ended up refining a technique that is relevant to other hyperspectral datasets," said co-author Carl Bernacchi, a RIPE research leader and scientist with the U.S. Department of Agriculture, Agricultural Research Service, who is based at Illinois' Carl R. Woese Institute for Genomic Biology. "The next step is to test more stacked machine learning algorithms on datasets from many more crop species and explore the utility of this technique to estimate other parameters, such as abiotic stresses from drought or disease."

"As scientists, we should try to use our domain knowledge to explain advanced performance from machine learning methods," said co-author Kaiyu Guan, an assistant professor in Illinois' College of Agriculture, Consumer, and Environmental Sciences (ACES). "Combining computational methods and domain disciplines allows us to possibly unravel what causes the measurable differences in hyperspectral datasets—which is an unsolved mystery in our work and worth future exploration."

More information: Katherine Meacham-Hensold et al. High-throughput field phenotyping using hyperspectral reflectance and partial

least squares regression (PLSR) reveals genetic modifications to photosynthetic capacity, *Remote Sensing of Environment* (2019). [DOI: 10.1016/j.rse.2019.04.029](https://doi.org/10.1016/j.rse.2019.04.029)

Peng Fu et al, Hyperspectral Leaf Reflectance as Proxy for Photosynthetic Capacities: An Ensemble Approach Based on Multiple Machine Learning Algorithms, *Frontiers in Plant Science* (2019). [DOI: 10.3389/fpls.2019.00730](https://doi.org/10.3389/fpls.2019.00730)

Provided by Carl R. Woese Institute for Genomic Biology, University of Illinois at Urbana-Champaign

Citation: Scientists stack algorithms to improve predictions of yield-boosting crop traits (2019, June 3) retrieved 24 April 2024 from <https://phys.org/news/2019-06-scientists-stack-algorithms-yield-boosting-crop.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.