

Project Proposal: Predicting US Corn Yields in the Midwest United States

Amy Steward (asteward@berkeley.edu), Michael White (msteckwhite@berkeley.edu), Yoko Morishita (yoko_morishita@berkeley.edu), Chad Adelman (chad_adelman@berkeley.edu)

Motivation:

Agricultural yield quantities have benefited or stymied humanity's progress since the invention of plant domestication. Despite the knowledge farmers have collected throughout the millennia, the yield from one year to the next has always been mercurial. This is particularly true in today's changing climate, with more frequent extreme temperature and precipitation events. Methods such as pest management and the application of weed killers and fertilizers can improve yields per acre of crop land, and help mitigate the uncertainty of agricultural output.

In this project, we aim to use machine learning to model the relationship between both human and natural elements and the yield of corn fields in the midwestern United States. We will focus on predicting corn yield in bushels per acre based on weather characteristics and human factors like irrigation and fertilization. Additional elements we could consider include soil moisture or nutrients, and seedling growth. By forecasting farm output and, by extension, the profitability of a growing season, farmers can better allocate resources and plan for financial burdens such as upgrading farm equipment. This also informs policymakers about overall production trends, which is crucial for setting agricultural policy, and managing food security initiatives.

Data:

Our main source of data comes from the National Agricultural Statistics Services or NASS. NASS has been around for 162 years, continuously taking annual surveys of United States crops and livestock grown. We intend to gather data on a yearly basis at the state level, focusing on an outcome variable of bushels per acre of corn. Informative data sets include moisture percentage, irrigated and non-irrigated corn yields and moisture content along with overall bushel per acre yields and some information on herbicide usage.

<https://quickstats.nass.usda.gov/>.

The dataset from National Oceanic and Atmospheric Administration (NOAA)'s Climate at a Glance provides year-to-date temperature and standardized precipitation data for the US national and state levels since 1895, enabling analysis of long-term trends and variability: https://www.ncei.noaa.gov/access/monitoring/climate-at-a-glance/national/time-series/110/pcp/yt_d/0/1895-2025.

Related work:

Recent advances in corn yield predictions show a growing trend in using machine learning technology to optimize productivity and resources. The studies below are focused on predicting yield of corn and other crops using features such as satellite images, meteorological data, and soil characteristics. These studies show the increasing influence that data driven models can have on forecasting future yield. Our project will aim to build on these studies by employing similar machine learning methods with more present data to see if our conclusion is aligned with past studies.

[An Artificial Intelligence Approach to Prediction of Corn Yields under Extreme Weather Conditions Using Satellite and Meteorological Data](#)

[Crop Yield Prediction Using Machine Learning](#)

[Projecting Future Changes in Potato Yield Using Machine Learning Techniques: A Case Study for Prince Edward Island, Canada](#)