ECON 484 Project Proposal

Noah Andersen, Lincoln Bay, Eliza Campbell, Yuki Sato

**Goal:**

The United States Department of Agriculture’s National Agricultural Statistics Service (NASS) performs an “Agricultural Census” to see the state of agricultural production and health in the United States economy (USDA, 2023). It provides critical information that cannot easily be found anywhere else to inform farmers, investors, and policymakers about the effects of weather, climate, and institutional decisions. However, this census is only collected once every five years, and an accompanying survey performed every year frequently fails to collect enough data to publish results for all counties. In fact, if not enough farmers in one county fail to respond, the NASS frequently withholds data from other counties in the state to prevent third parties from calculating the unresponsive county’s yield themselves. It’s not because there isn’t agricultural production there, either; which counties are not included changes year-to-year (Johanns & Thessen, 2020).

We will use satellite imagery taken from farms geolocated using the NASS-provided CroplandCROS tool months before harvest to generate a dataset filling in these gaps in non-census years for counties in which surveys produced insufficient responses.

**Data:**

We plan to obtain our data from two main sources: the United States Geological Survey (USGS) and the National Agricultural Statistics Service, published by the U.S. Department of Agriculture.

The National Agricultural Statistics Services publishes a “Census of Agriculture” every five years, dating back to 1840. This census contains data such as Land, Market Value of Agricultural Products Sold, Income From Farm-Related Sources, Farm Production Expenses, and more. This data is provided on the county level. The NASS also includes the locations of various types of crops (by latitude and longitude), which we can then tie back to the satellite imagery we will use.

The United States Geological Survey contains many datasets and topographical maps, as well as the satellite imagery that we plan to use in our analysis. This data is published several times a year.

After narrowing down the scope of our research and establishing a specific location (county) that we plan to analyze, we will likely incorporate more data into our project to increase the accuracy and reliability of the dataset we create.

NASS Census of Agriculture: <https://agcensus.library.cornell.edu/>

USGS satellite imagery & maps: <https://www.usgs.gov/educational-resources/data-tools-maps-and-satellite-imagery>

**Methods:**

The methods that we will use can be broken down into three steps:

1. **Crop the Satellite Imagery to our Area of Interest:**

The satellite imagery that we will be using will be standardized to a specific time frame before harvest season. With this criteria in mind, we will then overlay geospatially referenced satellite imagery (Geotiffs) onto geospatial open source farm location data. This will provide us with geospatial boundaries which we can then use to crop the satellite image to our area of interest.

1. **Generate a Crop Health Spectral Index:**

We then take the clipped satellite imagery of the agricultural fields we are interested in and leverage the NIR (Near Infra-red Reflectance) multispectral band coupled with the Red visible spectrum band. From these two bands we can generate the NDVI (Normalized Difference Vegetation Index) which isolates the reflectance of chlorophyll found in healthy vegetation. This index will then allow us to determine crop health based on a normalized scale.

1. **Predict Crop Yield for Withheld Counties to Complete our Dataset:**

Once we have the vegetation health feature we will apply it to our dataset of other features impacting crop yield. We will then fit a linear regression model to this data using crop yield as our dependent variable Y, crop health as our independent variable X, and using other crop related variables as controls such as weather, year, etc. Then with our fitted linear regression we can fill in the gaps between years when the NASS releases crop yield data.

References

Johanns, A. M., & Thessen, G. (2020, March). *Collection of county yield data, how does USDA NASS do it?*Www.extension.iastate.edu. https://www.extension.iastate.edu/agdm/articles/johanns/JohMar20.html

*USDA - National Agricultural Statistics Service - About NASS - Agency Overview*. (2023, December 29). Www.nass.usda.gov. https://www.nass.usda.gov/About\_NASS/index.php