

INTRODUCTION

- Beck's Hybrids develops and sells seed across the US
- Constantly developing new seed hybrids to perform better
- Part of development involves test plots for new variants
- Environmental factors can incur data and monetary losses
- Develop model for predicting least likely to be discarded plots
- Model based on public geospatial, soil, weather data

METHODOLOGY



Geospatial Data

- Automated data collection and processing through TNM API
- Difficulties:
 - LIDAR data collection not standardized across states
 - US states do not have complete coverage of LIDAR data
 - API is not up to date and sometimes returns incorrect links

Soil Data

- Web Soil Survey and "FedData" for data collection, visualized with R Shiny and Leaflet applications
- Difficulties:
 - Managing large file sizes and file structures
 - Inconsistencies from data sources, outdated sources
 - Very slow to load data

Weather Data

- Collection of data through weather stations
- Analyzed data via graphing and descriptive statistics
- Difficulties:
 - Inconsistent public data
 - Weather station location bias
 - Predicting weather difficult in itself
 - Many variables to consider (rain, wind, disasters, etc.)

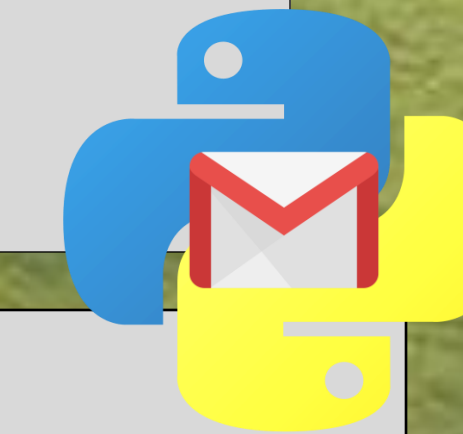
FUTURE GOALS

Machine Learning Goals

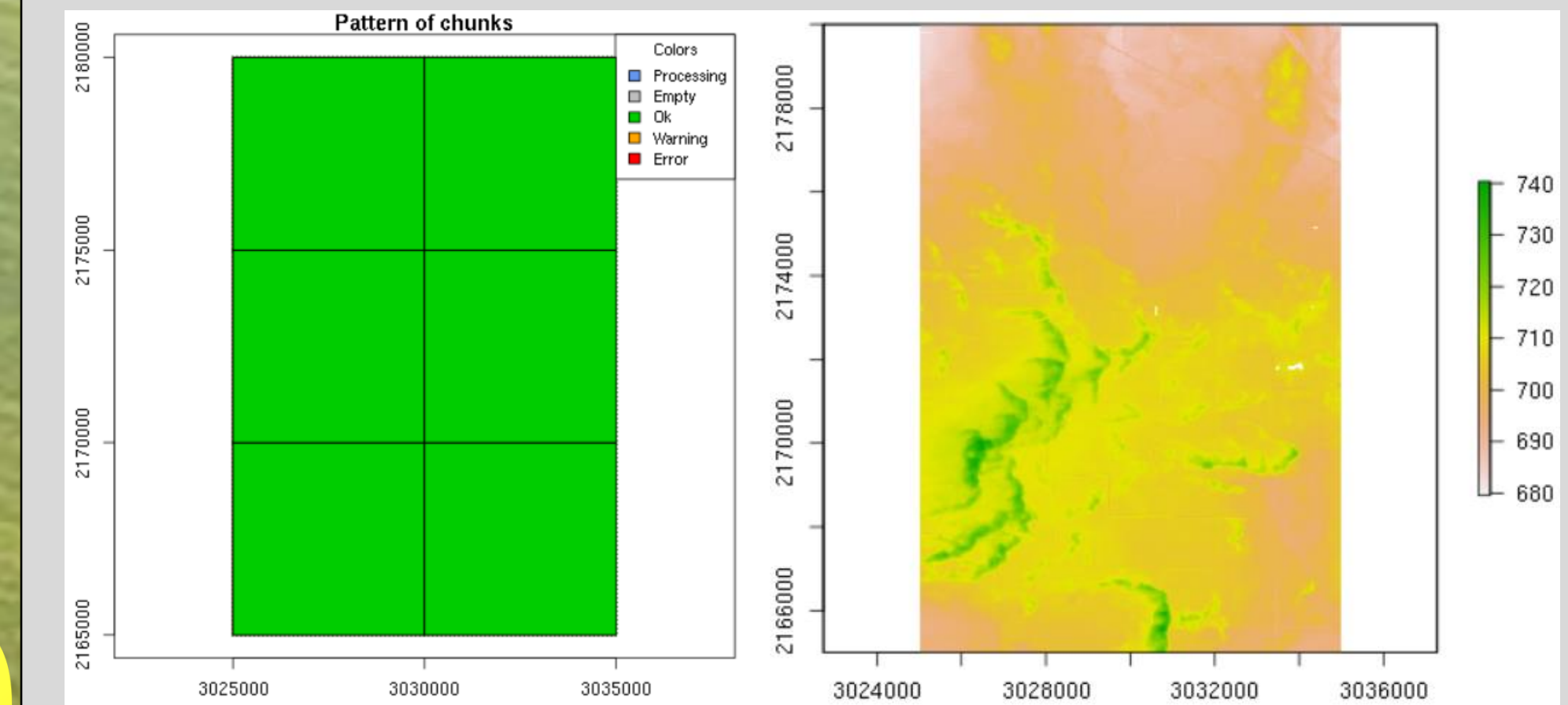
- Classify successful plots
- Merge geospatial, soil, weather data for training
- Early-stage Naive Bayes model achieves 90% accuracy on classification
- Develop Logistic Regression Model for predicting percentage of success

App Development Goal

- Build Shiny app akin to soil data extraction app
- Deploy R Shiny app as a desktop app to mitigate internet usage
- Integrate Model Estimators for faster prediction
- Incorporate automatic email messaging
- Automatically generate land quality and success estimation report



LIDAR Rastering



Soil Mapping in Shiny

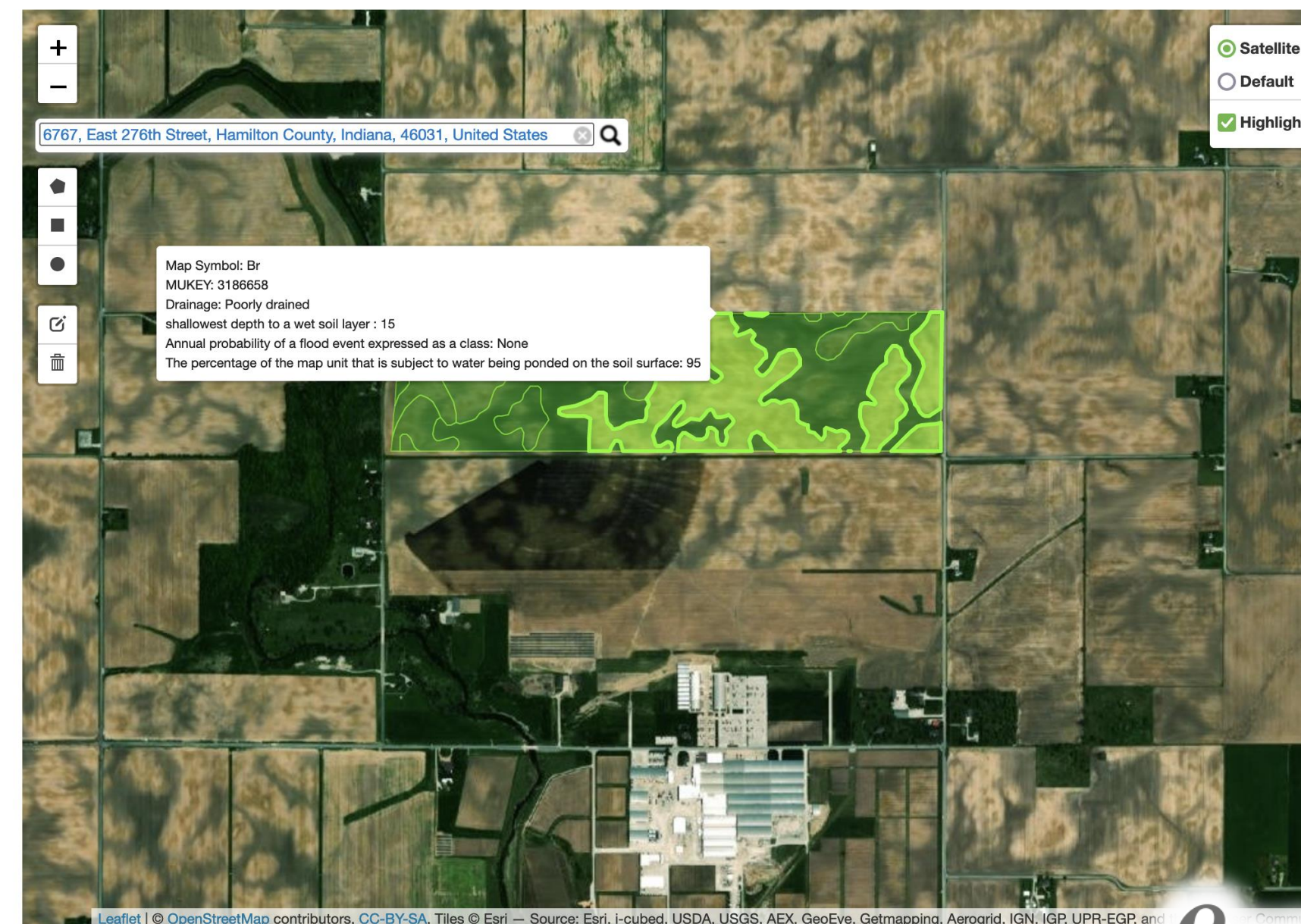
Area of Interest Soil Mapping

SSURGO Data Download

Instructions

1. **Search using OSM Geocoder:** The OSM Geocoder Search box allows for you to look up locations. Try searching: 6767, East 276th Street, Hamilton County, Indiana, 46031, United States. This is Beck's Hybrids HQ.
2. **Draw Polygon:** Click on the polygon shapes to draw a field. Try to not to draw anything bigger than a county because the data will take a minute to load. Make sure to draw 1 polygon.
3. **Download:** Click on the 'Download Map Data' button. A message should appear in the bottom right corner signifying data loading. This should take about a minute.
4. **Explore Field Data:** Hover above the field polygon shapes to see the field's soil characteristics

Area of Interest Selected



CONCLUSION

- Developed an understanding of digital agriculture limitations
- Learned about interoperability of data in agriculture
- Familiarized with limiting characteristic of production in agronomy
- Acquired knowledge of data extraction automation through APIs
- Grasped a better understanding of data science in business
- Learned about Shiny application development
- Gained experience with geospatial data methods
- Researched classification machine learning models

ACKNOWLEDGEMENTS

The Beck's Hybrids Plot Loss Tool Data Mine Team would like to thank our team at Beck's Hybrids: Will Hirschfeld, Steve Becker, Brandi Payne, Brad Fruth, Allyson Talyor, and Brooke Gajownik. Thank you for an opportunity to learn and challenge us with real world problems that require the implementation of data science. We would also like to thank our Data Mine mentors, David Glass, Kevin Amstutz, and Margarette Betz, who have helped us along this journey as well.

REFERENCES

- Bocinsky, Kyle. "Functions to Automate Downloading Geospatial Data Available from Several Federated Data Sources." *GitHub*, Oct. 2021, <https://github.com/ropensci/FedData>.
- Engel, Claudia A. "Using Spatial Data With R." *Chapter 2 Spatial Data Manipulation in R*, 2020, <https://cengel.github.io/R-spatial/spatialops.html>.
- NACSE, PRISM Climate Group. "Time Series Values for Individual Locations." *Prism Climate Group at Oregon State University*, USDA Risk Management Agency, 2022, <https://prism.oregonstate.edu/explorer/>.
- NRCS, USDA. "Web Soil Survey." *Web Soil Survey*, United States Department of Agriculture, 2019, <https://websoilsurvey.sc.egov.usda.gov/App/WebSoilSurvey.aspx>.
- University of Illinois. "Illinois Height Modernization (ILHMP): LIDAR DATA." *Illinois GeoSpatial Data Clearinghouse*, Prairie Research Institute, 2015, <https://clearinghouse.isgs.illinois.edu/data/elevation/illinois-height-modernization-ilhmp>
- USGS. "The National Map." *The National Map | U.S. Geological Survey*, United States Department of the Interior, 2022, <https://www.usgs.gov/programs/national-geospatial-program/national-map>.

