Using Observed Yield Maps to Optimize Variable Fertilizer Application

Research Paper

Presented to the

Center for Science and Technology

Charles Herbert Flowers High School

In Partial Fulfillment

of the Requirements for

Research Practicum

Marquis Skinner

May 6, 2022

CHARLES HERBERT FLOWERS HIGH SCHOOL

Using Observed Yield Maps to Optimize Variable Fertilizer Application

By:

Marquis Skinner

Has been approved on

May 6, 2022

APPROVED:

SAMANTHA COTTON, Program Coordinator

CARLISE BETHEL, Biological Sciences Committee Member

ALEXANDER DULCE, Engineering Committee Member

ERIN LUKOMSKA-SCHLAUCH, Biological Sciences Committee Member

KAREN SHELTON, Internship Coordinator, Committee Member

SHANICE WHITE, Computer Science Committee Member

ACCEPTED AND SIGNED:

Karen Shelton
Internship Coordinator
Approving Committee Member

Samantha Cotton
Science and Technology Program Coordinator
Approving Committee Member

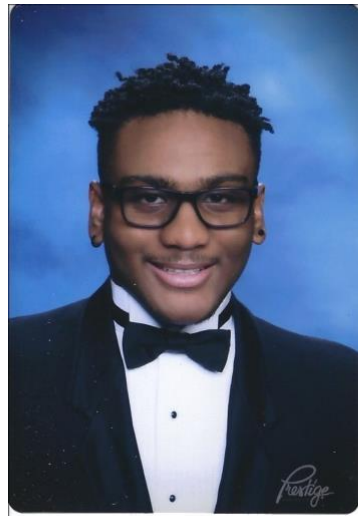
Abstract

Precision agriculture is a crop management strategy that uses data collected from an agricultural field to manage inputs such as fertilizer. For this research, a Python program was developed and used to optimize fertilizer application to maximize profit and minimize environmental impact. The measures of center were found for the results of the optimization. In addition, a sensitivity analysis was performed in order to examine how the cost of nitrogen fertilizer and corn value affected the optimal nitrogen application rate. Lastly, another analysis was performed to determine the correlation between precipitation and yield. Based on the measures of center (mean and median), this application of precision agriculture is effective. The average profit of the entire field increased from $2010 / ha to $2780 / ha. This is about a 38% increase after optimization. The median profit of the entire field increased from $2140 / ha to $2780 / ha. This is about a 30% increase after optimization. However, based on the sensitivity analysis, this application is not effective with no more than a ∓5% difference in optimal nitrogen application rate. It is believed that the difference in the yield potential between Iowa and Maryland may be the reason for such a large mean and median increase. Based on the findings of the precipitation yield analysis, precipitation and yield do not have significant correlation. With continued improvements to this research, this method of precision agriculture can be an effective method for improving current agricultural systems.

Acknowledgments

I would like my mentor Dr. Glenn Moglen, a Research Hydrologist at the Hydrology & Remote Sensing Laboratory at USDA-ARS for his support this entire year. I would also like to thank the Science and Technology Program, my internship coordinator Ms. Karen Shelton, and the Science and Technology Coordinator Mrs. Samantha Cotton. Their support has allowed me to thrive in this amazing opportunity. Lastly, I would like to express my appreciation to the Beltsville Agricultural Research Center for providing this opportunity.

Biographical Outline



Name: Marquis Skinner

Date of Birth: June 22, 2004

Place of Birth: Brooklyn, NY

City of Residence: Suitland, Maryland

Intended College: If you have more than one, pick one

Intended Major: Computer Science

Academic Achievements:

- Science & Technology Program (2018-2021)
- Principal's Honor Roll (2018-2021)
- Academic Honor Roll (2018-2021)
- National Honor Society (2019-2021)

Extracurricular Activities:

- Jaguar Players (2018-2020)
- Rainbow Club (Secretary) (2020-2022)
- National Chinese Honor Society (2020-2022)
- Math Honor Society (2020-2022)

Work/Volunteer Experience:

- USDA ARS (2021-2022)
- Tutor (2021)

Advice to Class of 2023:

- Take advantage of the resources and opportunities you are provided. Everyone wants to help you to succeed.

Table of Contents

List of Tables

List of Figures

Chapter One

The Problem and Its Setting

Agriculture is one of the most important industries in the world. Society relies on the agriculture industry to feed the humans that it relies on. As society moves toward a more sustainable future, agriculture has been one field in need of green innovation. The major reason being that agriculture has a significant impact on the environment. The application of fertilizer, for example, can have negative effects of the environment. "….nitrate contamination of ground and surface water" (Ferguson, Hergert, Schepers, & Crawford, 2015) due to leaching can affect surrounding ecosystems as excess nutrients are added to the biosphere. Leaching is the runoff of nutrient-rich fertilizer due to precipitation into nearby water bodies. Leaching is caused by excess fertilizer that is not taken up by crops. This excess fertilizer stays in the soil or on other surfaces where it can potentially be washed away. In order to mitigate this issue, one method was developed. That method is called precision agriculture.

Precision agriculture is a crop management strategy that uses data collected from an agricultural field in order to manage inputs. The effects of sufficient input management are maximized outputs, minimized cost, and limited negative environmental consequences. Managing inputs to improve outputs can also be referred to as optimization. Optimization is the act of managing a system in order to maximize or minimize a clearly defined goal.

One major factor that precision agriculture needs to account for in order to effectively apply fertilizer is spatial variability (Nielson, Wendroth, & Pierce, 1998). Spatial variability refers to differences in some quantity across a space. In this context, spatial variability refers to the differences across an agricultural field. One factor that varies across an agricultural field is crop response to fertilizer. "Corn yield response to N has been found to vary spatially within a field." (Kablan, et al., 2017). Yield is the measurement of the amount of harvested product from an agricultural field. Since crops do not respond uniformly to fertilizer application, variable distribution of fertilizer is necessary to address the needs of crops across a field in order to maximize yield and limit excess fertilizer application.

The backbone of precision agriculture is the relationship between the input which is fertilizer and the output which is crop yield. When this relationship between fertilizer and yield is graphed (with fertilizer being the independent variable and yield being the dependent variable), it can be seen that the relationship is not linear. The graph flattens as the x-values increase. This is called the yield rate curve. As more fertilizer is added, the less return there is in yield. In this relationship, there comes a point on the graph where "….the fertilizer rate at which crop yield increase is not large enough to pay for additional N application…." (Puntel, et al., 2016). This is called the Economic Optimum Nitrogen Rate (EONR). This point can be used to manage fertilizer application based on predicted yield.

The use of precision agriculture as a crop management system is relatively new. Practical applications of this system began thanks to advances in data collection and processing technology that are needed to make precision agriculture viable. "Implementation of spatially variable operations essentially requires four systems: sensing and mapping technologies, a location system, [Geographical Information Systems] and decision support software and precise and automatic application equipment." (Stafford & Bolam, 1998). There are multiple ways to collect agricultural data. One method that dominated in the '90s was soil sampling. "data can….be gathered at present by intensive manual sampling and subsequent laboratory analyses." (Stafford & Bolam, 1998). Currently, drones that use multispectral cameras and tractors that can measure yield while harvesting are the two most popular methods. To process and display these data, computers are used to turn the yield into useable values. These values are used to create yield maps. A yield map shows the spatial distribution of harvested product intensity.

There are numerous programs that can be used to simulate this fertilizer to yield relationship. APSIM is one such program. APSIM stands for Agricultural Productions Systems sIMulator. According to the authors, APSIM is “an open-source advanced simulator of agricultural systems that combines several process-based models in a modular design” (Puntel, et al., 2016). Open source means that the code can be seen and copied by anyone. According to the APSIM website, the APSIM model is internationally recognized for its capabilities (APSIM, n.d.).

In order to verify the results of models such as APSIM, statistical methods have to be employed. Statistics is the science of collecting and analyzing data. Since precision agriculture involves using data, statistics is a necessary component. “Relations between any number of soil, plant and seasonal weather variables proceeding across a farmer’s field can be identified with state-space analysis with a quantitative accounting of both measurement and model errors.” (Nielson, Wendroth, & Pierce, 1998). To find the model errors stated in this quote, it is necessary to used statistical methods such as ‘root mean squared error’. “To evaluate APSIM model goodness of fit, we used…. statistical methods. For the statistical evaluation, we computed the root mean square error (RMSE)….” (Puntel, et al., 2016).

Statement of the Problem

The application of fertilizer in farming can have negative effects on the environment. Excess fertilizer can leach into surrounding bodies of water. The excess nutrients from the fertilizer can result in algae blooms. These blooms can consume all oxygen in the water body, leading to anoxic conditions. These conditions can stress plant and animal life in the water body.

The relationship between yield and fertilizer application is not linear. Instead, the relationship can be modelled with a quadratic-plateau equation that approaches horizontal for large fertilizer application rate. This means the graph of this relationship eventually flattens or saturates. In terms of fertilizer and yield, this means that as more fertilizer is added, less additional crop is produced. This problem affects farmers economically as they are getting less return on their investment if they continue applying fertilizer at high rates.

The farmer trying to maximize profit and/or minimize negative environmental effects needs to consider crop value, fertilizer costs, and the spatial variability in crop yield potential across the farm field to develop a precise strategy for how to vary fertilizer application rates within the farm field.

Yield on an agricultural field in not uniform/even. It varies due to differences in soil type, topography, residual nutrients, and soil moisture across the field.

How can computer science and data science be used to address these problems?

Design Goal

For this research project, observed crop yield maps from the Beltsville Agricultural Research Center (BARC) was integrated with information on crop prices, fertilizer prices, and yield-fertilizer rate curves to maximize farmer profit and minimize negative environmental impacts. With this research, practical tools (computer algorithms) will be developed that can inform farmer decision-making with the goal of optimizing economic returns from a planned seasonal crop planting.

Variables and Specifications

Independent

- Nitrogen Fertilizer
- Cost of Fertilizer
- Price of corn grain

Dependent

- Yield
- Profit

Controlled

- Location – Beltsville Agricultural Research Center (BARC)
- Type of crop – Corn
- Harvesting and Measuring Tools
- Type of Fertilizer
- Amount of Fertilizer

Software specifications

- Programming language – Python
- Uses Excel to display produced spatial maps
- File size – TBD

Assumptions

The amount of fertilizer used was assumed to be the same amount yearly.

Limitations

Due to time constraints, research that covers all variables that affect corn grain yield was not accounted for. In addition, field testing on the effectiveness of the program was not accomplished.

Statistical Analysis

The data was analyzed using two statistical analyses. The first analysis was a sensitivity analysis to determine the how much corn and fertilizer price affected the optimal nitrogen application rate. The second analysis was a correlation analysis to determine how precipitation affected yield. Both of these analyses were quantitative studies.

Summary

The overarching topic of this research project is precision agriculture. Over the last 30 years, research has conducted in this field in order to make precision agriculture a viable solution for minimizing the environmental impact of agriculture. The goal of this research project is to use observed yield maps from the Beltsville Agricultural Research Center (BARC) in addition to information on crop prices, fertilizer prices, and yield-fertilizer rate curves to maximize farmer profit and minimize negative environmental impacts. This project will use precision agriculture techniques and implement them into a python program that can display prescribed fertilizer application, projected yield, and projected profit maps in Excel. This research project was completed over one year. Due to this time constraint, field testing was not used. All data produced by this project came from computer- based models.

Definitions of Terms and Abbreviations

1. **APSIM** - APSIM stands for Agricultural Productions Systems sIMulator. APSIM is "an open-source advanced simulator of agricultural systems that combines several process-based models in a modular design"
2. **Economic Optimum Nitrogen Rate (EONR)** - "….the fertilizer rate at which crop yield increase is not large enough to pay for additional N application…."(Puntel, et al., 2016)
3. **GIS (Geographic Information System) -** used to capture and analyze spatial and geographic data
4. **Leaching -** the runoff of nutrient-rich fertilizer due to precipitation into waterbodies ranging from groundwater to streams and lakes.
5. **Optimization -** the act of managing a system in order to maximize or minimize a clearly defined goal.
6. **Precision Agriculture -** a crop management strategy that uses data collected from an agricultural field in order to manage inputs.
7. **Spatial Variability -** refers to the differences in some quantity or characteristic across a space.
8. **Yield -** the measurement of the amount of harvested product from an agricultural field.
9. **Yield Curve -** The graphical relationship between fertilizer and yield.
10. **Yield Map** - A yield map shows the spatial distribution of harvested product intensity.

Chapter Two

The Review of Related Literature

The purpose of this research is to integrate observed crop yield maps from the Beltsville Agricultural Research Center (BARC) with information on crop prices, fertilizer prices, and yield-fertilizer rate curves in order to maximize farmer profit and minimize negative environmental impacts. Before starting this research, related literature needed to be reviewed. Only articles that were within the last five years or were important to the field were chosen. 10 articles have been reviewed.

From these 10 articles, three major methods of research were discovered: physical experimentation, modeling, and surveying. These three methods will be compared to gain more insight into each method.

Physical Experiment

What Is a Physical Experiment?

A physical experiment is a type of experiment that collects real-world data after performing some physical procedure. In terms of precision agriculture, this is an experiment that is performed on and affects real crops/soil. This is one of the most common types of experiment used in the field of precision agriculture.

Timeframe

The time frame for physical experiments varies. They typically range from two years (Stafford & Bolam, Near-ground and aerial radiometry imaging for assessing spatial variability in crop condition, 1998) to over a decade (Ferguson, Hergert, Schepers, & Crawford, 2015). This depends on the criteria of the experiment.

Tools

Tools are needed to collect data from crops to perform these experiments. The most common method of data collection, remote sensing, employs the use of technologies such as radiometers to collect data. (Stafford & Bolam, Near-ground and aerial radiometry imaging for assessing spatial variability in crop condition, 1998). Other tools such as crop harvesters (Baum, Archontoulis, & Licht, 2019) and soil samplers (Kablan, et al., 2017) are also used.

Data Collected

The data that is collected from these experiments can vary depending on the criteria of the project. Yield is the most common data point collected in precision agriculture. Nitrogen presence in soil (Ferguson, Hergert, Schepers, & Crawford, 2015) and crop condition (Stafford & Bolam, Near-ground and aerial radiometry imaging for assessing spatial variability in crop condition, 1998) are other data points that may be collected. In a single experiment, multiple data points can be collected and analyzed. One example of this is from a study titled 'Emerging concepts for solving the enigma of precision farming research'. "[Figure of] Spatial process of wheat grain yield, base saturation, and available water storage capacity…." (Nielson, Wendroth, & Pierce, Emerging concepts for solving the enigma of precision farming research, 1998). Here, the researchers for this study created a figure displaying multiple data points from their experiment.

Model

What is a Model?

A model or simulator is a computer program that is used to mimic real-world physical processes. They employ the use of statistics and other data science principles. The results from these models are not physically measured. They are instead merely predictions. In the case of precision agriculture, models are used to predict various agricultural factors, particularly yield.

Timeframe

The time frame needed to complete an experiment based on a model is usually shorter than that of a physical experiment. Models typically use data previously gathered (Puntel, et al., Modeling long-term corn yield response to nitrogen rate and crop rotation, 2016). This means the researchers do not need to wait to gather crop or soil data. This significantly shortens the time needed to complete the experiment. In addition, the advancement of computers has significantly shortened the time needed to perform mathematical calculations over large datasets.

Types of Models

There are many types of models used within the field of precision agriculture. One such mathematical model is “A multiple linear regression analysis….” (Morton, Buchleiter, & Heermann, 1998). More machine learning-based models are described in the article titled ‘County-level soybean yield prediction using deep CNN-LSTM model’. “Based on remote sensing data, great progress has been made in this field by using machine learning, especially the Deep Learning (DL) method, including Convolutional Neural Network (CNN) or Long Short-Term Memory (LSTM).” (Sun, Di, Sun, Shen, & Lai, 2019). Additionally, “Support vector machine (SVM), Gaussian process regression (GPR), and random forest (RF)….” (Han, et al., 2020) are types of models that can be used in precision agriculture.

Data Collected

The data that is collected from these experiments can vary depending on the criteria of the project. Yield is the most common data point that is modeled. Soil nitrogen and crop condition can also be modeled.

Reliability

Models are usually less reliable than physical experiments because the results from models are predictions. "Of possible concern is the reliance of the results on those relatively few points beyond the boundaries of the irrigation system that received less water than other parts of the field….Since no systematic field measurements of soil moisture were taken to verify the model estimates, the daily soil moisture value calculated by OPUS are simply unverified estimations." (Morton, Buchleiter, & Heermann, 1998)

Survey

What is a Survey?

A survey is a study performed on a population, usually people. The questions that are inquired depend on the criteria of the survey. In precision agriculture, this method is the least used as it does not directly result in any improvements in the field.

Data Collected

Answers from the population are collected. These answers are usually organized into categories that depend on the criteria of the survey. For example, from the study titled 'Precision agriculture technology adoption: a qualitative study of small-scale commercial "family farms" located in the North China Plain', the researchers divided their results into five categories. "Five central themes emerged from the data 'socio-political landscape', 'farming culture', 'agricultural challenges', 'adoption intentions (barriers/facilitators)' and 'practical support mechanisms'." (Kendall, et al., 2022).

Reliability

The reliability of the results of surveys is one constraint every survey considers. The size of the population surveyed, the diversity of the population, and the trustworthiness of the answers affect the overall reliability of the results.

Comparison

Similarities and Differences

- The main similarity between physical experiments and models is the data they produce. Both experiments and models produce similar points of data.
- While experiments produce measurable data, models produce predictions that can only be verified through physical means.

Strengths and Weaknesses

- Physical experiments produce more reliable results as compared to models.
- Model-based projects take significantly more time to complete than physical experiments

Summary

After considering the strengths, weaknesses, similarities, and differences between a physical experiment, surveys, and modeling, a model-based approach was chosen for this project. The time constraint of one year was the largest factor in this decision. Since physical experiments take a significant amount of time to complete, this choice was eliminated. Surveys do not fulfill the criteria of this project. Therefore, this choice was also eliminated. A model-based project will fit the time constraint and fulfill the criteria of this project.

Chapter Three

Method

Fertilizer application has environmental, agricultural, and economic consequences/impacts. Based on research in this field, observed yield maps can be used to optimize fertilizer application to limit its consequences. A python program was created to calculate the optimal fertilizer application based on observed yield maps from previous years.

The analysis of the data was quantitative. This research was performed in collaboration with Beltsville Agricultural Research Center. Data from their corn field was used in this analysis.

Method

The initial proposed design this research created was a program that used observed yield maps from previous years to prescribe fertilizer application in order to maximize yield and minimize negative environmental consequences. The output was an excel sheet representing a map of prescribed fertilizer application for the BARC field. In order to ensure the results of the design were correct, by-hand calculations were performed and peer-reviewed. If the by-hand and program results were the same, the results were assumed to be correct. If the results were different, the program was reviewed with the assumption that the by-hand calculations were correct.

After performing the calculations, it was discovered that the results were not explicitly useful. The optimized yield was not explicit in how beneficial this optimization was for the farmer. After conducting further research, it was determined that displaying the output as profit in addition to yield provided a better representation of the results. Another problem arose when testing how to implement the cost of corn and nitrogen fertilizer into the equation. Using real-time data solved this problem.

Instrumentation/Specifications.

1. HP Laptop 15-bs0xx
    a. Intel Core i3-7100U
    b. 8GB RAM
    c. Windows 10
2. Texas Instruments TI-84 PLUS CE Graphing Calculator
3. PyCharm Community Edition 2021.2.2 (Integrated Development Environment
4. Microsoft Excel

Procedure

1. Conduct initial research into site/location that is being used, tools used to collect the data, how the data is processed/displayed, how fertilizer-yield response is modelled, and how to find EONR
2. Begin testing.
    a. Obtain an observed yield map from a recent year.
    b. Write a python program that can access the observed yield map data and use the mathematical tools acquired from the initial research to process the data.
    c. To ensure the program results are accurate, by-hand calculations are needed. These results should be compared to the results of the program.
3. Complete Research Paper

a. Transfer Results of the program to the research paper. Explain the results and implications

***Program Design Outlines*.**

The program finds the file in which the observed yield map is stored in. The data from the map is stored in the program. Mathematical methods as shown in Figure 3.1 were integrated into the program and were used to process the inputted data. The program then stores the results as recommendation maps into excel files.

Figure 3.1

Fundamental Equations

$$f(x) = a + bx - cx^2$$

$$g(x) = mx$$

$$Profit = a + bx - mx - cx^2$$

$$\frac{dP}{dx} = (b-m) - 2cx = 0$$

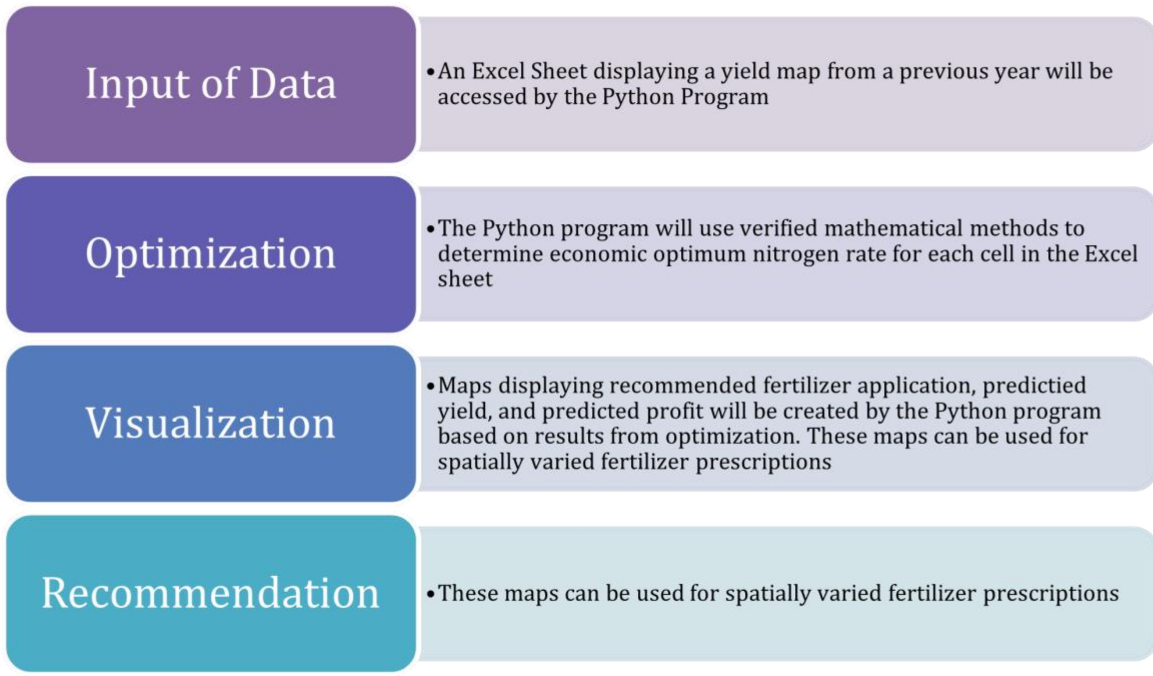
- ➔ The algebraic representation of yield-fertilizer response (a,b,and c are coefficients)
- ➔ Algebraic representation of cost of fertilizer (m is the rate/slope)
- ➔ Algebraic representation of profit for farmer (subtract Equation 2 from Equation 1)
- ➔ To find EONR, set the derivative of the profit equation to zero and solve for x

Solutions Selection

The final design was a Python Program thanks to the simple but powerful nature of Python and its libraries. The process is detailed in Figure 3.2.

Figure 3.2

Program Process



Design Analysis and Evaluation

Yield and fertilizer application data from a field at BARC was provided. The yield data were in cents/bushel. The fertilizer application rate was assumed to be uniform across the field. Fertilizer application was in $/ton. After the calculations were performed, the profit was in $/ha and the fertilizer application was in kg/ha.

A sensitivity test was performed in order to analyze how the price of corn and fertilizer affected the optimal nitrogen application rate calculated by the program. The ranges for the sensitivity test were determined by using historical corn and nitrogen fertilizer prices.

Figure 3.3

Graph of Historical Nitrogen Prices

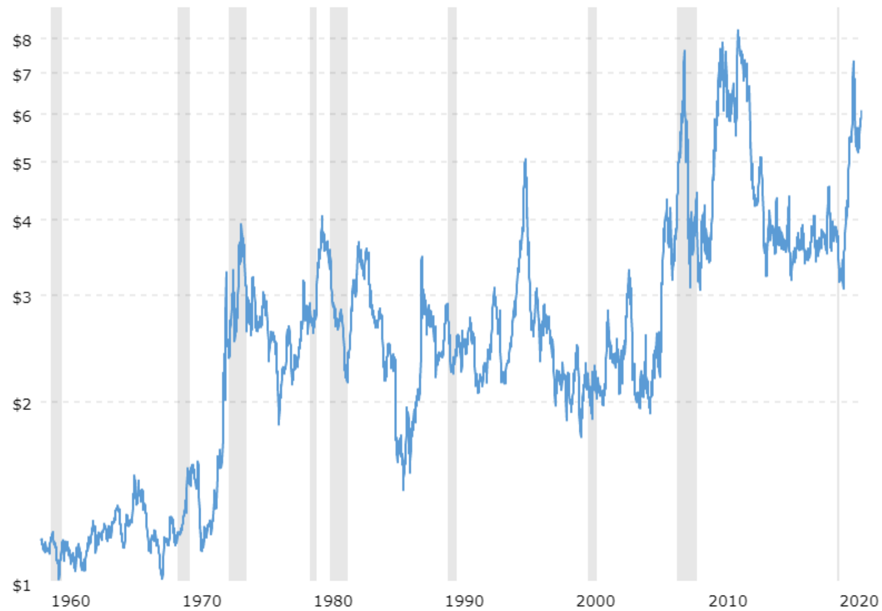
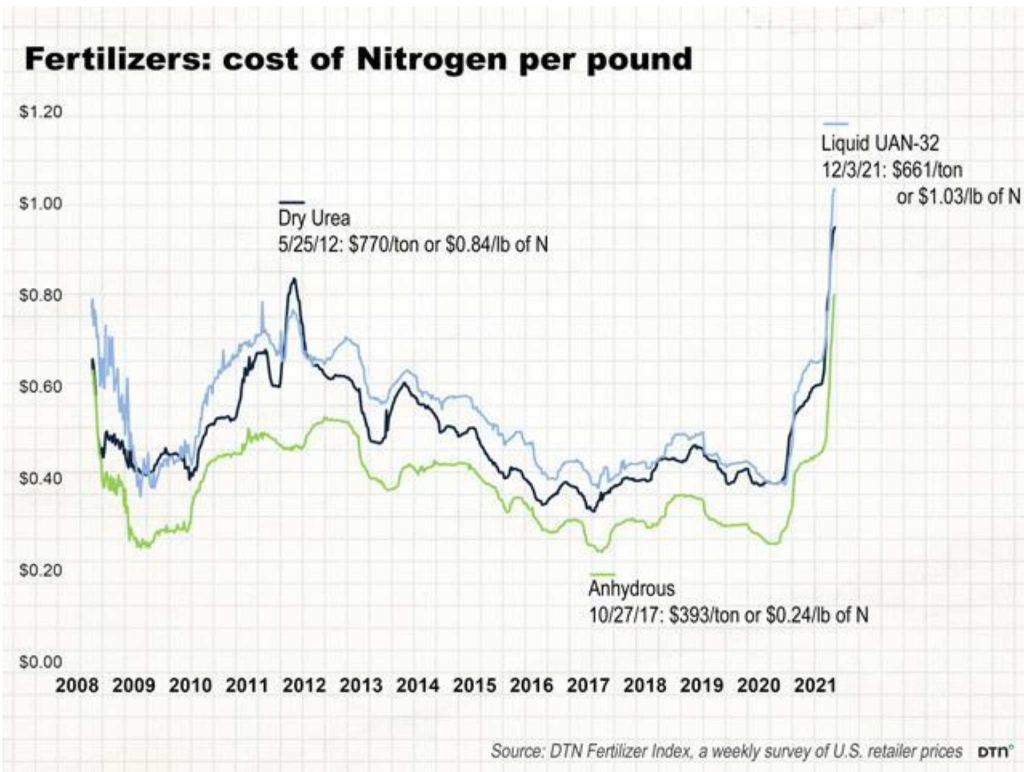


Figure 3.4

Graph of Historical Nitrogen Prices



Summary

Summarily, a computer program was created that could use observed yield maps from a corn field to optimize yield and maximize profit. The results from the optimization were in the form of Excel sheets representing maps. The location used for this testing was Beltsville Agricultural Research Center. All testing was done via modelling. No plants were used in this research. The results from this research are described in the next chapter.

Chapter Four

The Findings

This chapter presents the results gathered from this experiment. The interpretation of the results will be provided in the next chapter.

Presentation of the Data

Table 4.1

Measures of Center for Observed and Optimized Yield Data

Measure of Center	Observed	Optimized
Mean	$2010 / ha	$2780 / ha
Median	$2140 / ha	$2780 / ha

Note. This table demonstrates how the optimization of fertilizer affects the profit for the farmer.

Table 4.2

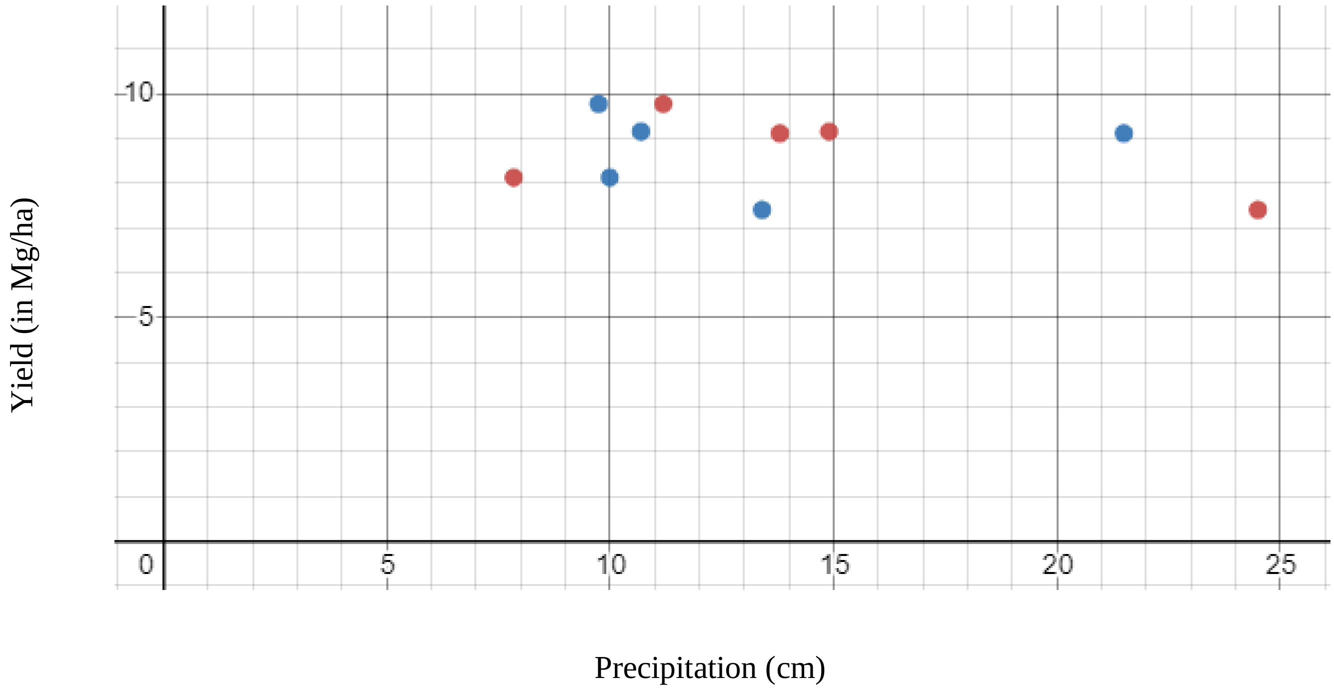
Sensitivity Analysis: Optimal Nitrogen Application Rate

Corn Price	Fertilizer (in $/kg)				
(in $/Mg)	0.2	0.4	0.6	0.8	1
118	211	206	201	196	191
157	212	209	205	201	197
236	214	211	209	206	204
315	214	212	210	209	207

Note. This sensitivity analysis demonstrates how the optimal nitrogen application rate is affected by the price of corn and nitrogen. Results are measured in kg N/ha.

Figure 4.1

Figure of Precipitation-Yield Relationship



Note. The data in this figure represents the Precipitation-Yield Relationship for five different years. The x-axis represents

*Red represents July. Blue represents August.

Design Analysis

One major modification in the program was the change from using the Xlsxwriter library (McNamara, 2021) to the Openpyxl library (Gazoni & Clark, 2021) for Excel file manipulation. Openpyxl provided additional flexibility in file types and read/write functionality. Another major modification was the change to a more object-oriented format. The creation of functions allowed other files to access functionality reducing redundancy.

Summary

As shown by Table 4.1, the average profit of the entire field increased from $2010 / ha to $2780 / ha. This is about a 38% increase after optimization. The median profit of the entire field increased from $2140 / ha to $2780 / ha. This is about a 30% increase after optimization. According to Table 4.2, the sensitivity analysis suggests that corn value and nitrogen fertilizer cost impact optimal fertilizer application by no more than $\mp$5%.

According to Figure 4.1, there appears to be little correlation between precipitation and yield. The small sample size was considered and will be further discussed in the next chapter.

Chapter Five

Conclusion

In summary, this project used observed crop yield maps from the Beltsville Agricultural Research Center (BARC) and integrated them with information on crop prices, fertilizer prices, and yield-fertilizer rate curves to maximize farmer profit and minimize negative environmental impacts. This project hopes to be able to improve the application of precision agriculture. Computer algorithms were developed and are provided in the appendix. These computer algorithms processed all of the data and produced fertilizer application rate maps. Statistical analysis was performed on the results

Summary of the Findings

The measures of center indicated that the optimization of fertilizer application increased profit by about 30-40%. The sensitivity analysis indicated that the price of corn and nitrogen fertilizer affects the optimal nitrogen rate by no more than 10%. Lastly, the correlation coefficient R was found for the precipitation-yield relationship. It was discovered that precipitation does not have a significant effect on yield.

Conclusions and Discussion

Based on the measures of center (mean and median) as shown in Table 4.1, this application of precision agriculture was effective. However, based on the sensitivity analysis shown in Table 4.2, this application was not effective. It was believed that the difference in the yield potential between Iowa and Maryland may be the reason for such a large mean and median increase. For future research, it was recommended to derive an equation that more accurately represents the yield-fertilizer relationship in Maryland. Based on the findings of the precipitation yield analysis shown in Figure 4.1, precipitation and yield do not have a significant correlation.

This may have been a result of the small sample size used. For future research, it is recommended that a larger sample size is used for better accuracy.

Recommendations

To improve this study, it is recommended to develop a yield-fertilizer rate curve that better represents the yield response in Maryland. It is also recommended to use a larger sample size to determine the correlation more accurately between yield and precipitation.

One major difficulty in this research was the creation of the yield-fertilizer rate curve. In the future, additional research into how to create yield-rate curves should be a priority.

In continuation of this research, more optimized and scalable computer algorithms should be developed. Also, to ensure accessibility for farmers, a more user-friendly design should be considered.

Future Implications

This research was important as it has the potential to modernize the way food is produced. As the environment is considered more in the advancement of societal processes, the results from this research may be considered a starting point for discovering a solution to address the environmental impacts of agriculture.

References:

APSIM. (n.d.). *APSIM: The Leading Software Framework for Agricultural Systems Modelling and Simulation*. From apsim: https://www.apsim.info/

Clark, C., & Gazoni, E. (2021). Openpyxl Python Library (Version 3.0.9) [Python Library]. https://foss.heptapod.net/openpyxl/openpyxl

Ferguson, R. B., Hergert, G. W., Schepers, J. S., & Crawford, C. A. (2015). Site-specific nitrogen management of irragated corn. *Proceedings of the Fourth International Conference on Precision Agriculuture*, (pp. 733-743).

Kablan, L. A., Chabot, V., Mailloux, A., Bouchard, M.-E., Fontaine, D., & Bruulsema, T. (2017). Variability in Corn Yield Response to Nitrogen Fertilizer in Eastern Canada. *Agronomy Journal*, 2231-2242.

Marcin, M. (n.d.). *History of Agriculture*. From crestcapital: https://www.crestcapital.com/tax/history_of_agriculture#:~:text=The%20history%20of%20agriculture%20begins%20in%20the%20Fertile%20Crescent.&text=Humans%20invented%20agriculture%20between%207%2C000,barley%2C%20chickpeas%2C%20and%20flax.

McNamara, J. (2021). XlsxWriter Python Library (Version 3.0.1) [Python Library]. https://github.com/jmcnamara/XlsxWriter

Nielson, D. R., Wendroth, O., & Pierce, F. J. (1998). Emerging Concepts for Solving the Enigma of Precision Farming Research. *Proceedings of the Fourth International Conference on Precisions Agriculture*, (pp. 303-318).

Puntel, L. A., Sawyer, J. E., Barker, D. W., Dietzel, R., Poffenbarger, H., Castellano, M. J., . . . Archontoulis, S. V. (2016). Modeling Long-Term Corn Yield Response to Nitrogen Rate and Crop Rotation. *Frontiers in Plant Science*.

Stafford, J. V., & Bolam, H. C. (1998). Near-Ground and Aerial Radiometry Imaging for Assessing Spatial Variability in Crop Condition. *Proceedings of the Fourth International Conference on Precision Agriculture*, (pp. 291-302).

Appendix

```
import openpyxl
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import statistics
import os

matrix = pd.read_excel('APU_5_2017_soybeans_clip (1).xlsx', header=5,
engine='openpyxl')
matrix.to_excel('Linear Optimization - APU_5_2017_soybeans_clip (1).xlsx')  #
Intakes and stores Excel file values into program

if os.path.exists("Adjusted.xlsx"):
    os.remove("Adjusted.xlsx")  # Removes file if it exists to prevent
errors.

list1 = matrix.stack()  # Flattens Dataframe into Series
list1_0 = []
std = round(list1.std(), 4)
list2 = list1.sort_values()  # Sorts the values in ascending order
list2_0 = []
list3 = []
median = 0.0
mean = 0.0
global mean1

workbook = openpyxl.Workbook()
workbook.save("AdjustedTemp.xlsx")
wb = openpyxl.load_workbook("AdjustedTemp.xlsx")
wrksht1 = wb.create_sheet("Optimized Yield", 0)
wrksht2 = wb.create_sheet("Optimized Fertilizer")
wrksht3 = wb.create_sheet("Optimized Fertilzer Quadratic")
cols = len(matrix.columns)

for item in list1:  # Creates a list without the NO DATA cells
    if item > 0:
        list1_0.append(item)

for item in list2:  # Creates numerically sorted list without the NO DATA
cells
    if item > 0:
        list2_0.append(item)
mean1 = statistics.mean(list1_0)
median = statistics.median(list2_0)
print(mean1)
print("Median"+str(median))

def linoptimize():
    print(list1.to_string())
    print("\n\n\n\n\n\n\n")
    print(list2.to_string())
    print("\n\n\n\n\n\n\n")

    #  Linear Optimization
```

```
    listapp = []  # List of optimized yield
    listfert = []  # List of fertilizer prescription
    list3 = np.array_split(list2, 2)  # Splits the list of numbers into two
lists
    sum = 0.0
    for i in list3:
        for j in i:
            sum = sum + j

    sum2_0 = 0
    sum2_0_1 = 0
    sum2_1 = 0
    sum2_1_1 = 0
    for item in list1:
        if median > item > 0:  # if the cell is not a NO DATA cell and it is
less than the median
            y1 = item

            x = 1
            m1 = (y1 - 50) / x
            # print(m1)
            if m1 < 0:
                y1_0 = 50
            else:
                y1_0 = 50 + (m1 * 1)  # Fertilizer application not adjusted
            y1 = 50 + (m1 * 0)  # Fertilizer application adjusted
            print(y1)
            listapp.append(y1)  # Creates list of optimized yield
            listfert.append(0)  # Creates list of prescribed fertilizer
treatment
            sum2_0 = sum2_0 + y1  # Sums all of the results (after process)
            sum2_0_1 = sum2_0_1 + y1_0  # Sums all of the results (before
process)
        elif item >= median and item > 0:  # if the cell is not a NO DATA
cell and it is greater than the median
            y2 = item
            x = 1
            m2 = (y2 - 50) / x
            # print(m2)
            if m2 < 0:  # if a value is less than 0, the yield will be 50
                y2_0 = 50
            else:
                y2_0 = 50 + (m2 * 1)  # Not adjusted fertilizer application
            y2 = 50 + (m2 * 2)  # Adjusted fertilizer application
            print(y2)
            listapp.append(y2)
            listfert.append(2)
            sum2_1 = sum2_1 + y2  # Sums all of the results (after process)
            sum2_1_1 = sum2_1_1 + y2_0  # Sums all of the results (before
process)
        else:  # if the cell is a NO DATA cell then it is just appended with
no optimization
            listapp.append(item)
            listfert.append(item)

    adjsum = 0.0
    for i in listapp:
```

```
        adjsum = adjsum + i

    mean = round(adjsum / len(listapp))

    print("Original sum:" + str(sum))
    print("Adjusted sum: " + str(sum2_0_1 + sum2_1_1))
    print("Sum after process: " + str(adjsum))
    print("Mean after process: " + str(mean))

    print(listapp)

    list3 = list2.to_numpy()

    k = 0
    for item in list3:  # Replaces all values less than 0 with 0
        if item < 0:
            list3[k] = 0
        k += 1

    print("Standard Deviation: " + str(np.std(list3)))
    print("Median: " + str(median))

    print(len(matrix.columns))
    print(len(matrix))


    i = 1
    j = 1

    for item in listapp:
        if j == cols+1:
            i = i + 1
            j = 1
        wrksht1.cell(row=i, column=j).value = item
        j = j + 1

    i = 1
    j = 1
    for item in listfert:
        if j == cols+1:
            i = i + 1
            j = 1
        wrksht2.cell(row=i, column=j).value = item
        j = j + 1

    # ----------------------------------------------------------------------------
-----------------------------------------------
    # Quadratic-plus-plateau
    # ----------------------------------------------------------------------------
-----------------------------------------------

xx = .55
yy = 208.5
# These are values determined by hand calculations with historical data
listquad2 = []

def quadoptimize(x,y):
```

```
    listquad = []
    listitem = []
    listprofit = []
    i = 1
    j = 1
    for item in list1:
        if item >0:
            m = (item*530)/(100*.405)
            listitem.append(m)
            print(m)
    print("------------------------------")
    for item in list1:
        if item>0:
            s = item / mean1  # Scaling Factor
            #  print(s)
            a = 4.95
            b = 0.073
            c = 0.0001689
            a = a * s * y          # ----------------------------------------
---------------------------
            b = (b * s * y) - x  # Each coefficient is multiplied by the
price conversion and scaling factor
            c = c * s * y         # ----------------------------------------
--------------------------
            deriv = b / (2 * c)
            profit = 834 + 14.7*(deriv)-0.0354*(deriv*deriv)
            # print(str(item) + " " + str(s) + " " + str(deriv))
            listquad.append(deriv)
            listquad2.append(deriv)
            listprofit.append(profit)
        else:
            listquad.append(item)
            listquad2.append(item)

    print("Adjusted Mean: "+str(statistics.mean(listprofit)))
    print("Original Yield Mean: "+str(statistics.mean(listitem)))
    print("Adjusted Median: "+str(statistics.median(listprofit)))
    print("Original Yield Median: "+str(statistics.median(listitem)))
    return listquad

def save():
    i = 1
    j = 1
    for item in listquad2:
        if j == cols+1:
            i = i + 1
            j = 1
        wrksht3.cell(row=i, column=j).value = item
        print(str(i) + " " + str(j) + " " + str(item))
        j = j + 1
    wb.save("Adjusted.xlsx")

    if os.path.exists("AdjustedTemp.xlsx"):
        os.remove("AdjustedTemp.xlsx")
    else:
        print("The file has already been removed.")
```

```
def histogram():
    print(len(matrix))
    print(len(matrix.columns))
    # Frequency Distribution of data
    kwargs = dict(alpha=0.5, bins=100)
    plt.hist(list3, **kwargs, color='g')  # list3 is an array of list2
    plt.title("Yield Distribution")
    plt.xlabel('Yield')
    plt.ylabel('Frequency')
    plt.axvline(mean, color='k', linestyle='dashed', linewidth=1)  # this
line shows mean on the graph
    plt.show()  # this creates the histogram

    #  EONR is calculated by the yield increase multiplied by the price of
corn. That product is then subtracted by the
    #   cost of N

linoptimize()
quadoptimize(xx, yy)
save()
```