The Impact of Genetically Engineered Upland Cotton on Land Utilization

Abstract

This paper examines the environmental impact of genetically modified upland cotton production in the United States. Linear regression models were employed to analyze the disparities in land usage between genetically engineered cotton and non-engineered cotton cultivation. The results revealed a significant difference in land usage, indicating the efficiency of genetically engineered cotton cultivation. This study aims to contribute to the growing discourse on sustainability within the agricultural domain.

Contents

Introduction

Cotton is one of the largest agricultural industries worldwide, commonly used in fabric production and a wide variety of other applications. The United States is one of the world's leading exporters of cotton, dominating around a third of the global cotton market. However, cotton cultivation faces significant challenges as its prone to a multitude of pests and disease. To address these issues, genetically engineered cotton was commercially introduced in the United States in 1995. Over the years, genetically modified cotton has been enhanced to increase yield capacity as well.

This paper aims to investigate the long term impact of genetically modified cotton on land usage and overall cotton cultivation trends within the United States, given that it has been over two decades since its introduction. Existing research mainly investigates genetically modified crops as a monolith, but this paper places its focus on Upland cotton.

Linear models were constructed to compare land usage discrepancies between GMO cotton and non modified cotton. The findings indicate that genetically engineered cotton generates higher yield and requires considerably less land for cultivation. Amidst growing sustainability concerns rising population and income, analysis of cotton production and land utilization are becoming more significant.

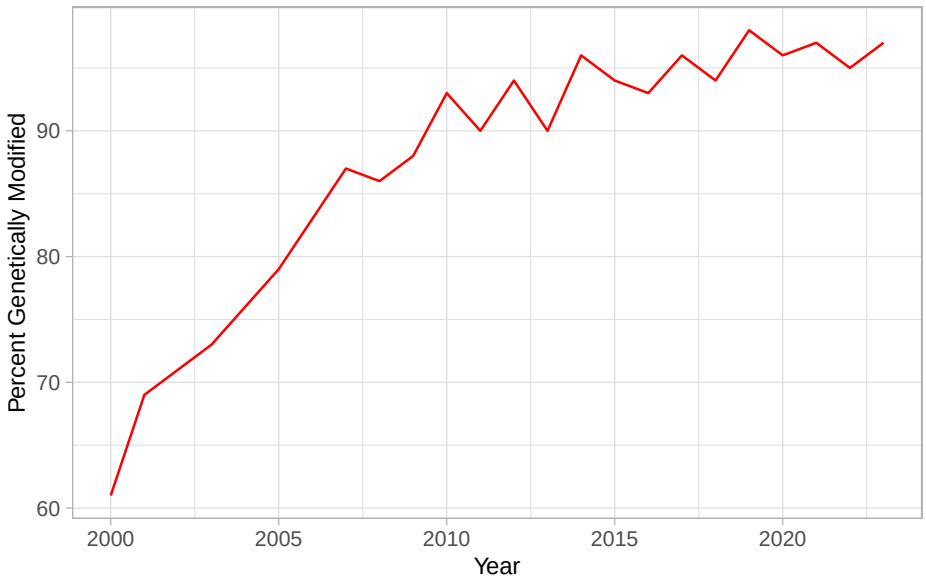
This paper is divided into 4 main sections: data, models, results, and the discussion. The data section examines the data used to construct the models presented in the models section. The results section uses the constructed models to make inferences and predictions. Finally, the last section discusses any other important details pertaining to the topic at hand.

Data

To analyze the effects of genetic engineering on upland cotton cultivation, raw data was obtained from the United States Department of Agriculture. These datasets were cleaned and transformed into a clean dataset using R and the tidyverse package. The resulting dataset consists of the variables "Year", "Percent GM", "Total Acres", "Acres Harvested", "Percent Abandoned", and "Yield".

The variable "Year" encompasses observations spanning from 2000 till 2023. "Percent GM" denotes the percentage of cotton planted that was genetically modified. "Total Acres" represents the total land area in acres dedicated to farming upland cotton in the United States, while "Acres Harvested" is the number of acres that were actually harvested. "Percent Abandoned" is the percentage of cotton that was not harvested and was abandoned. The variable "Yield" measures cotton harvested in pounds per acre.

This dataset was generated over other potential datsets since the original source has high credibility, and the cleaned variables "Year", "Percent GM", "Total Acres", "Acres Harvested", "Percent Abandoned", and "Yield" help investigate trends in upland cotton cultivation that have occurred in recent American history.



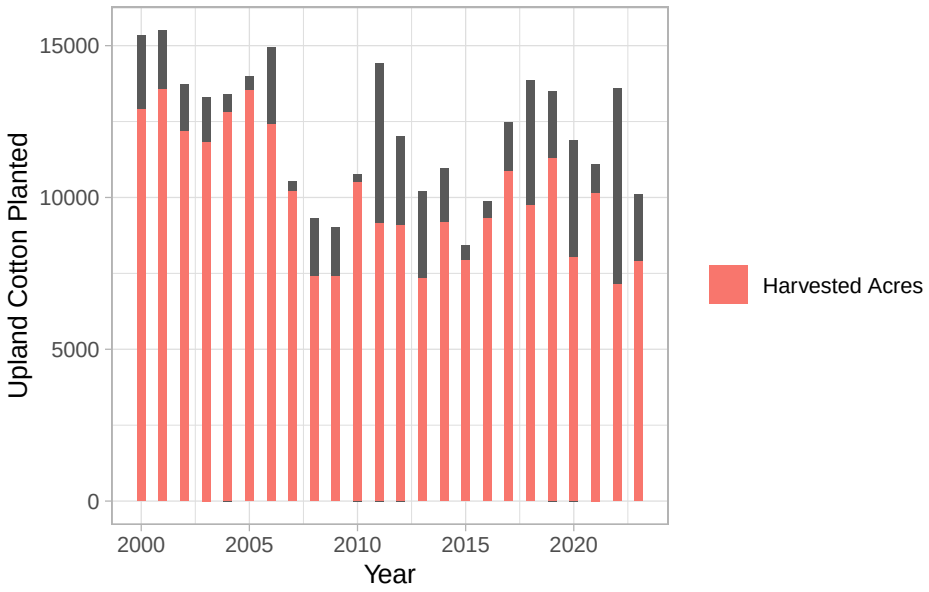
The percent of genetically modified Upland cotton planted in the United States increased drastically from 2000 to 2023. Starting with a minimum value of 61% in 2000, the percentage increased till it reached its peak in 2019 at 98%. The mean percentage of genetically modified upland cotton during this time period was 87%.

Measuring the percentage of genetically modified cotton is a difficult task as cross breeding can occur during cultivation. This type of measurement error is difficult to control as cultivating crops in complete isolation takes a lot of resources. This variable was generated using reports of farms tracked by the USDA.

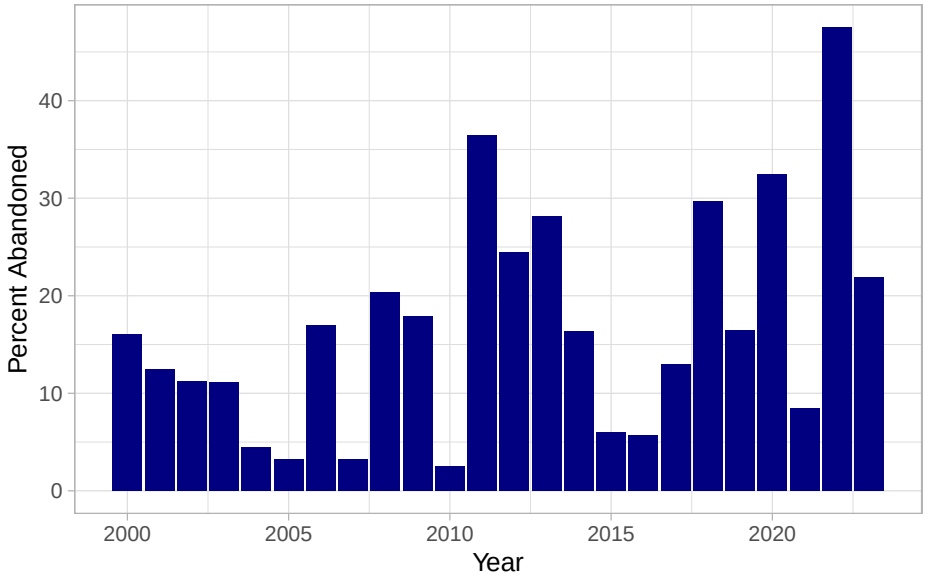
Acreage dedicated to Upland cotton in the United States varied drastically from 2000 to 2023, but notably decreased overtime. It reached a minimum value of 8,422 acres in 2015 and a peak of 15,499 acres in 2001. The mean acreage dedicated to planting upland cotton in the United States was 12,174 acres.

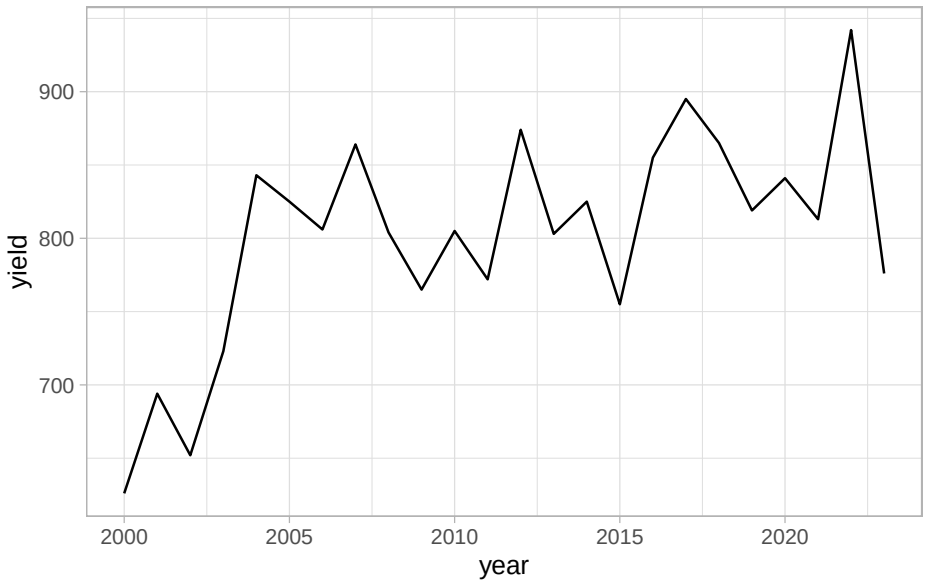
The total acreage of Upland cotton harvested follows a similar trend. The number of harvested acres has the overarching trend of decreasing over time. Harvested acres had a peak value of 13,560 in 2001 and a minimum value of 7,132 in 2022. From 2000 to 2023 the mean value of harvested acres was 10,072.

Measuring acreage is easier and more accurate than measuring the percent of genetically modified cotton. Active farm lands and plantations are tracked by the USDA and since most farming in the United States is commercial this is a fairly accurate measurement.



The percentage of abandoned cotton has a trend of increasing over time. In the year 2010 only 2.5% of the planted Upland cotton was abandoned compared to 2022 when around 48% of Upland cotton planted was abandoned. The mean percentage of cotton abandoned was 16%.





Upland cotton yield in the United States has been steadily increasing over time from 2000 to 2023. Total yield reached a maximum amount in 2022 with 942 pounds per acre compared to the minimum yield of 626 pounds per acre in 2000. The mean yield during this time period was 801 pounds per acre.

Models

Many linear regression models were employed to study the long term impacts of genetically engineering Upland cotton. First, 2 separate models were created with bootstrapping to infer the impact GMO cotton had on abandoned crop and yield.

Additional 2 multiple linear regression models were created to infer the discrepancies in land usage between GMO cotton and non GMO cotton.

Due to the limited data available bootstrapping with the boot package was used to enhance the accuracy of the coefficients estimated. All models were built using the tidymodels package.

The first model examines the impact of GMO cotton on the percentage of crop abandoned.

$$Y = 0.33X - 11.88$$

The second model examines the impact of GMO cotton on yield.

$$Y = 5.07X + 359.24$$

The third model uses GMO cotton and total acres of cotton planted as predictors for total yield, where GMO cotton is X_1 and total acres is X_2.

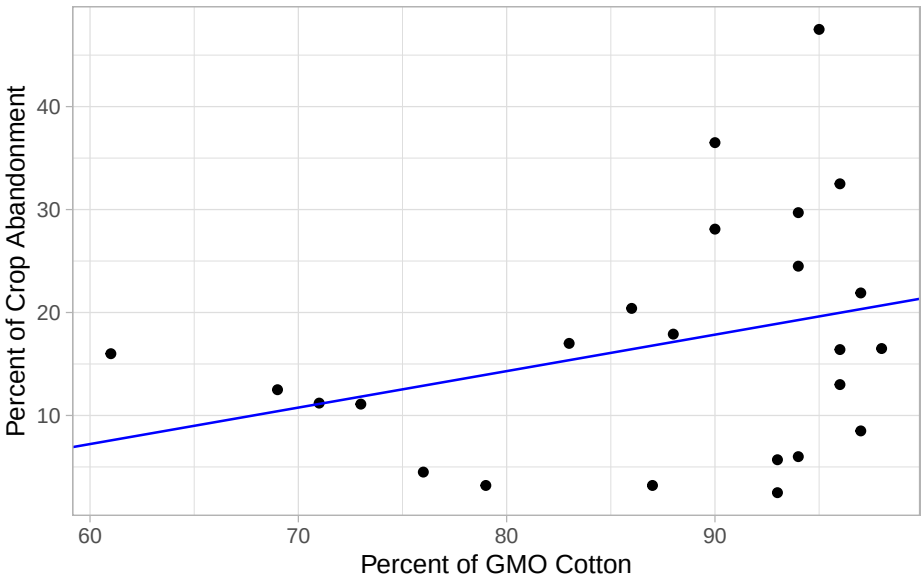
$$Y = 5.98X_1 + 0.01X_2 + 178.70$$

The last model uses GMO cotton and yield as predictors of total land used for cotton cultivation, where X_1 is GMO cotton and X_2 is yield.

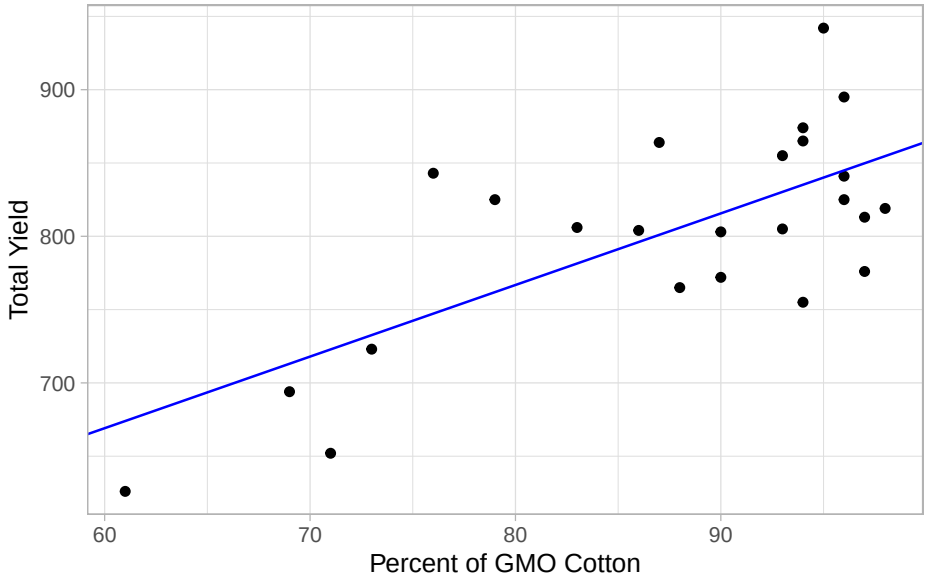
$$Y = -159.52X_1 + 9.86X_2 + 18199.24$$

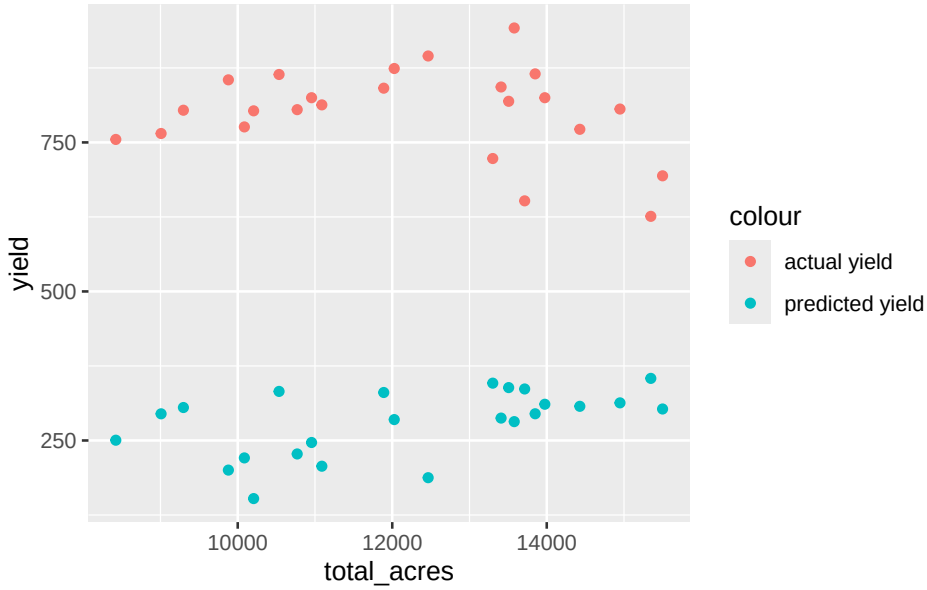
Results

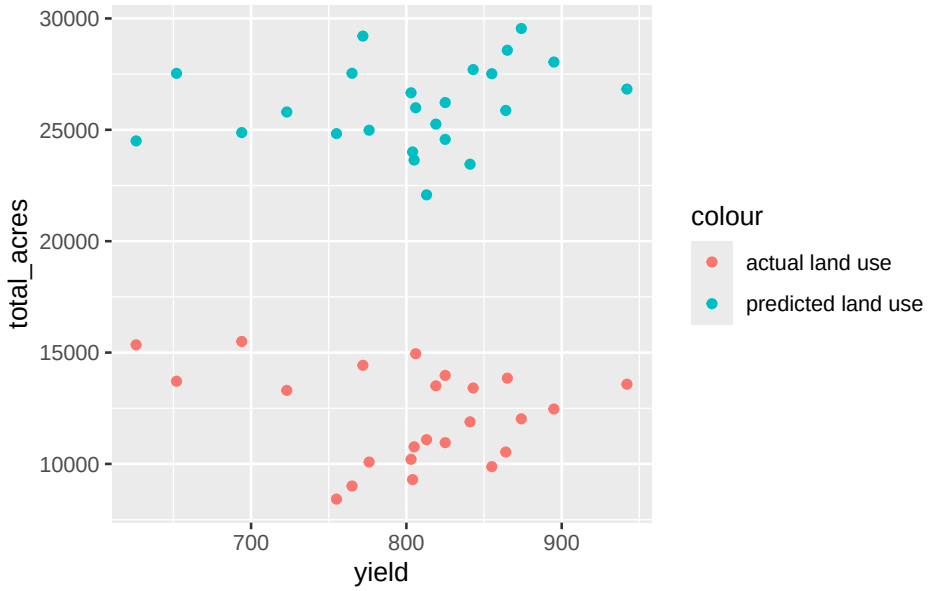
The first model $Y = 0.33X - 11.88$ depicts a positive relationship between GMO cotton and percent of crop abandonment.



The second model $Y = 5.07X + 359.24$ depicts a positive relationship between genetically engineered Upland cotton and total yield.


Total Yield
900
800
700
60
70
80
90
Percent of GMO Cotton


yield
750
500
250
10000
12000
14000
total_acres
colour
actual yield
predicted yield



Discussion

References