Improvements in Upland Cotton Production Through Genetic Engineering*

Tara Chakkithara

17 April 2024

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 and yield between genetically engineered cotton and non-engineered cotton cultivation. The results revealed a significant difference in land usage and yield, indicating the efficiency of genetically engineered cotton cultivation. This study aims to contribute to the growing discourse on sustainability within the agricultural domain.

Contents

1 Introduction

Making up around 90% of all cotton consumed, Upland cotton is one of the largest agricultural industries worldwide. It is commonly used in fabric produc-

*Analysis data and code available at https://github.com/pixel-echo/gmo tion but also has various other uses in kitchen projects and other applications (North Carolina State University, n.d.).

The United States is one of the world's leading exporters of cotton, dominating around a third of the global cotton market (United States Department of Agriculture 2022). However, cotton cultivation faces many challenges such as pests, disease, wind, and rain. For high quality cotton, it is essential the crop is picked as soon as it blooms, as prolonged exposure leaves it vulnerable to the challenges metioned (Weigmann 2024).

To address these issues, genetically engineered cotton was commercially introduced in the United States in 1995. There are three main types of genetically engineered cotton: insect-resistant (Bt), herbicide-tolerant (HT), and "stacked varieties that combine Ht and Bt traits" (Dodson 2022).

This paper aims to investigate the impact of genetically modified cotton on land usage and yield 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 using R (R Core Team 2023) and the "tidymodels" R package (Kuhn and Wickham 2020), to compare land usage and yield 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. Among growing sustainability concerns, along with rising population and income levels, 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 explains the results of the models sections. Finally, the last section discusses any other important details pertaining to the topic at hand.

2 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 (R Core Team 2023) and the tidyverse package (Wickham et al. 2019). 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.

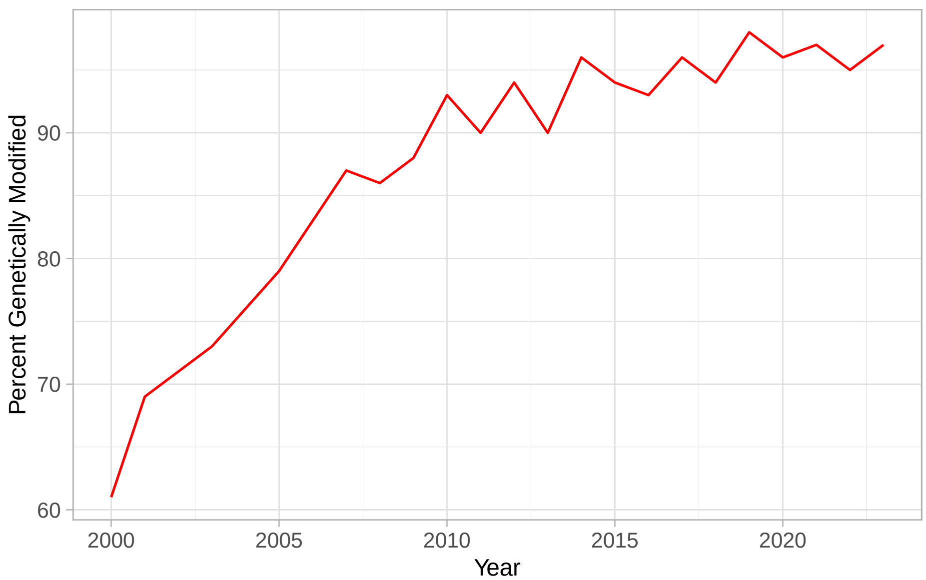

Figure 1: Percent of genetically modified Upland cotton from 2000 to 2023

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% as seen in Figure 1. 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.

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 as seen in Figure 2. From 2000 to 2023 the mean value of harvested acres was 10,072. Compared to other measurements measuring acreage is fairly accurate.

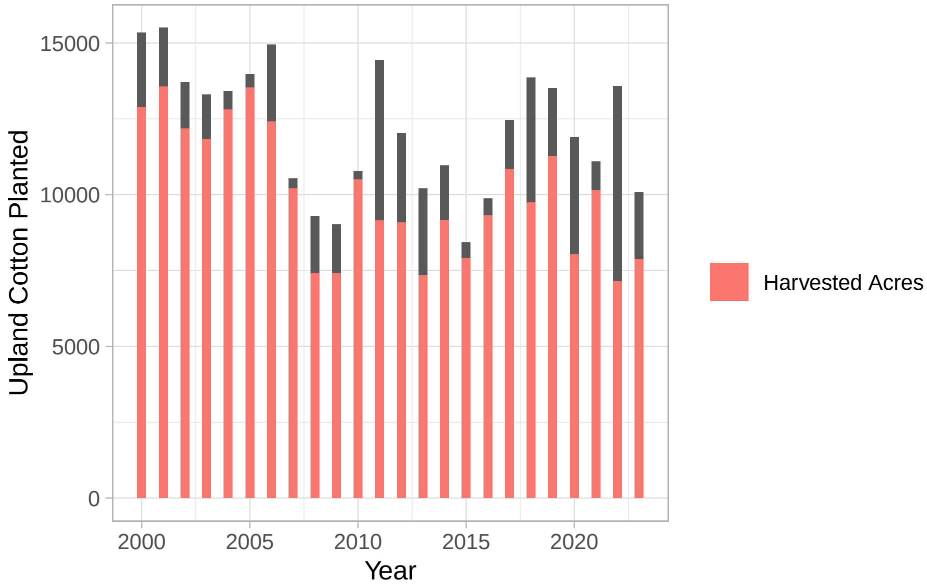

Figure 2: Total acres of cotton planted and harvested each year from 2000 to 2023

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% as seen in Figure 3.

As shown in Figure 4, 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.

3 Models

To study the impacts of genetic engineering on yield and land utilization, two simple linear models were constructed using the "tidymodels" package (Kuhn and Wickham 2020).

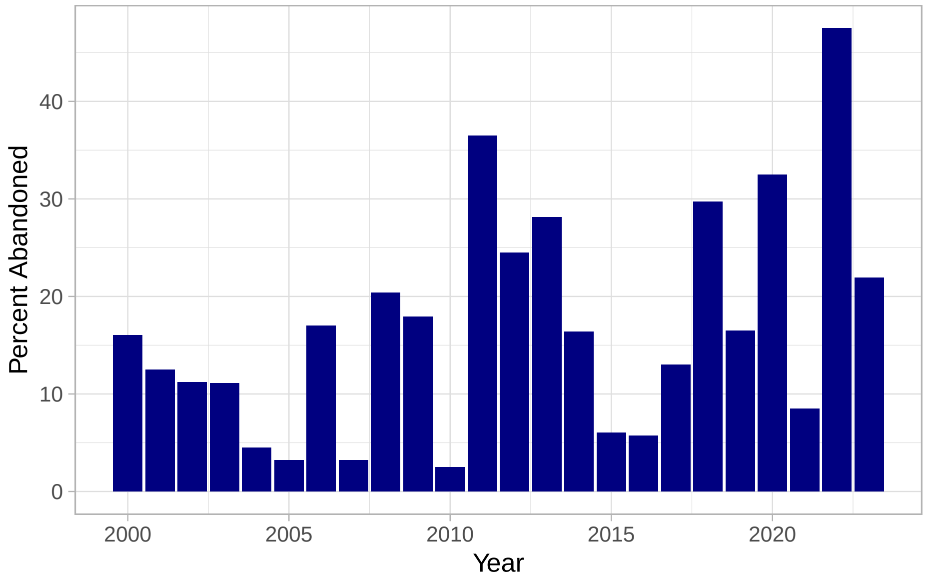

Figure 3: Total cotton abandoned each year from 2000 to 2023

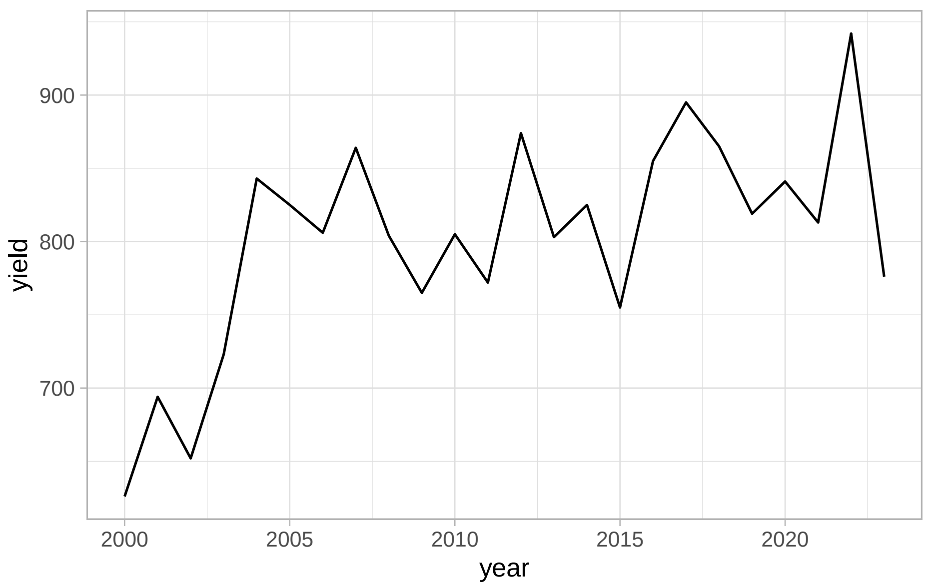

Figure 4: Upland cotton yield trends from 2020 to 2023

3.1 Yield

The first model constructed investigates the influence of genetically engineered cotton on the yield generated. Thus, in this model GMO cotton acts at the independent variable X and percentage of abandoned cotton represents the dependent variable Y_1. β_1 and β_0 represent the coefficients of the model.

$$Y_1 = \beta_1 X + \beta_0$$

3.2 Land Usage

The second model looks into the impact of genetically engineered cotton on land usage. In this model, GMO cotton is designated as the independent variable X and the total acres of cotton planted acts as the dependent variable Y_2. In this model, β_2 and β_3 represent the coefficients.

$$Y_2 = \beta_3 X + \beta_2$$

4 Results

Using the data, the coefficients for the first model were estimated to be $\beta_0 = 359.24$ (intercept) and $\beta_1 = 5.07$, with corresponding p-values 7.63×10^{-4} and 7.74×10^{-5} respectively. Since β_1 has a p-value less than 0.05, it is statistically significant, thus there is enough evidence that genetically engineered cotton impacts yield positively. This is also reflected in Figure 5.

For the second model, the coefficients β_2 (intercept) and β_3 were estimated to be 21741.90 and -109.55 respectively. β_3 had a p-value of 0.0061, making it statistically significant. Thus, there is enough evidence to conclude that GMO cotton impacts the total acres of cotton planted negatively. This model is shown in Figure 6.

5 Discussion

Based on the findings, we can conclude that genetically modified upland cotton is a more sustainable alternative compared to Non-GMO cotton, as it requires less land and produces more yield.

Based on Figure 3, we also know that despite land usage decreasing over time, crop waste has been increasing. This gives us a new focus in terms of sustainability. How can farm land be used more efficiently to reduce crop waste?

While these findings are exciting, this study has a couple weaknesses. First, all types of GMOs were considered during data analysis, without distinction. Secondly, technology has improved a lot from 2000 to 2023, other technological factors besides genetically modified seeds may be affecting yield and land usage.

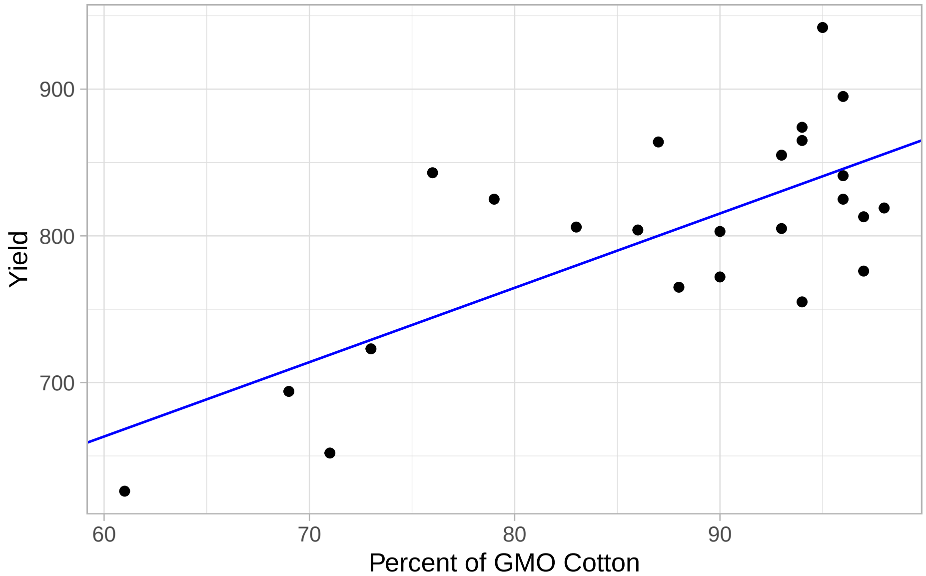

Figure 5: Linear regression model of the effect of genetically engineered cotton on total yield.

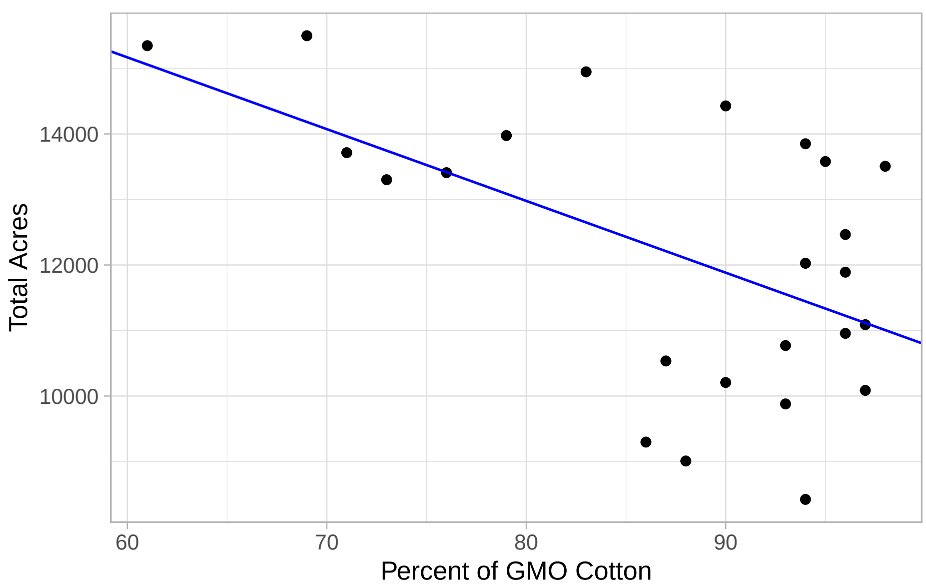

Figure 6: Linear regression model of the effect of GMO cotton on total land usage.

Further research is needed to explore these other factors, as they may also be crucial to sustainability.

References

Dodson, Laura. 2022. "Use of Genetically Engineered Cotton Has Shifted Toward Stacked Seed Traits." 2022. https://www.ers.usda.gov/amber-waves/2020/december/use-of-genetically-engineered-cotton-has-shifted-toward-stacked-seed-traits/.

Kuhn, Max, and Hadley Wickham. 2020. *Tidymodels: A Collection of Packages for Modeling and Machine Learning Using Tidyverse Principles.* https://www.tidymodels.org.

North Carolina State University. n.d. "Gossypium Hirsutum (Mexican Cotton, Upland Cotton)." https://plants.ces.ncsu.edu/plants/gossypium-hirsutum/.

R Core Team. 2023. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

United States Department of Agriculture. 2022. "Cotton and Wool." 2022. https://www.ers.usda.gov/topics/crops/cotton-and-wool/.

Weigmann, Hans-Deitrich H. 2024. "Cotton Fibre and Plant." 2024. https://www.britannica.com/topic/cotton-fibre-and-plant.

Wickham, Hadley, Mara Averick, Jennifer Bryan, Winston Chang, Lucy D'Agostino McGowan, Romain François, Garrett Grolemund, et al. 2019. "Welcome to the tidyverse." *Journal of Open Source Software* 4 (43): 1686. https://doi.org/10.21105/joss.01686.