Crop Abandonment: Effects of Weather, Irrigation, and Prices

by

Gan Jin

A Thesis Submitted to the Faculty of the

DEPARTMENT OF AGRICULTURAL AND RESOURCE ECONOMICS

In Partial Fulfillment of the Requirements

For the Degree of

MASTER OF SCIENCE

In the Graduate College

The University of Arizona

2013

STATEMENT BY AUTHOR

This thesis has been submitted in partial fulfillment of requirements for an advanced degree at the University of Arizona.

SIGNED:______________

Gan Jin .

APPROVAL BY THESIS DIRECTOR

This Thesis has been approved on the dates shown below

________________________ ______________

Professor George Frisvold Date
Agricultural and Resource Economics

Acknowledgement

First and foremost, I would like to thank my thesis advisor, Dr. George Frisvold, for his guidance and support. He has been a great mentor and has helped me both professionally and personally during my study at the Department of Agricultural and Resource Economics, the University of Arizona. I greatly appreciate the opportunity he has given me to be a Graduate Research Assistant and many doors that have opened while under his direction. I would also like to thank my committee members, Drs. Gary Thompson and Russell Tronstad. I am grateful for their support with my econometric modeling, their patience, and time. I would also like to thank Dr. Satheesh Aradhyula for his time, unconditional support, and patience and willingness to take time out of his busy schedules to assist me with any questions I had. Finally, I would like to extend my deepest gratitude to my AREC colleagues, the AREC department as a whole, and my friends and family. It has been my great pleasure and honor to be a member of AREC, I will always remember the support from the faculty members, the staffs, and my colleagues.

Table of Contents

LIST OF FIGURES

LIST OF TABLES

Abstract

Cotton is an important U.S. agricultural commodity, generating about 200,000 jobs among the various sectors from farm to textile mill and accounting for more than $25 billion in products and services annually. A double hurdle model is estimated to assess the effects of weather extremes, irrigation, crop variety choice, and changes in cotton prices on rates of county-level cotton acreage abandonment. Acres abandoned are those planted with cotton, but not subsequently harvested. The first step of the double hurdle model estimates factors affecting the probability that a county will have at least some acres abandoned. County level data were available for eight states and three crop types: upland irrigated cotton, upland non-irrigated cotton and irrigated Pima cotton. Abandonment rates were highest among upland non-irrigated acreage. Seasonal temperature and precipitation variables were significant predictors of abandonment behavior. A sub-sample of the data from Census of Agriculture years contained a variable for the number of cotton farms in a county. The probability that a county had some, positive amount of acres abandoned increased with the number of farms. The rate of abandonment for counties with abandoned acres, however, declined with the number of cotton farms. This results provides some justification of a more flexible double hurdle specification over a tobit specification to crop abandonment.

Chapter 1

Introduction

The United States cotton industry generates about 200,000 jobs among the various sectors from farm to textile mill and accounts for more than $25 billion in products and services annually. The predominant type of cotton grown in the U.S. is American Upland (*Gossypium hirsutum*), accounting for about 97% of the annual U.S. cotton crop. The other popular type is American Pima or extra-long staple (*Gossypium barbadense*). The markets for Pima cotton are mainly high-value products, for its long staple length of 1 ½ inches or longer, comparing to 1 to 1 ¼ inches for Upland cotton (Economic Research Service, United States Department of Agriculture, 2013). Like all other crops, cotton production suffers from abandonment due to various reasons. Figure 1.1 (Economic Research Service, United States Department of Agriculture, 2013) shows total acres planted and harvested from 1985 to 2009, across all cotton types.

Numerous studies have examined how climate and/or economic attributes can cause crop failure. Existing models vary drastically in complexity, largely in proportion to the dimensionality of the chosen environment specification (Starr and Kostrow 1978). For example, Brown (Brown 1959) used only linear functions of total precipitations for two periods, September-October and May-June, to explain variations of winter wheat yield in Utah. In contrast, Baier (1973) explained variations of spring wheat yield on selected plots in the Canadian Prairie Provinces. He used nonlinear functions of daily measures, throughout the growing season, of any three of minimum and maximum air temperature, soil moisture within the rooting zone of crops, the ratio of actual to potential evapotranspiration, and total incoming radiation from sky and sun as explanatory variables.

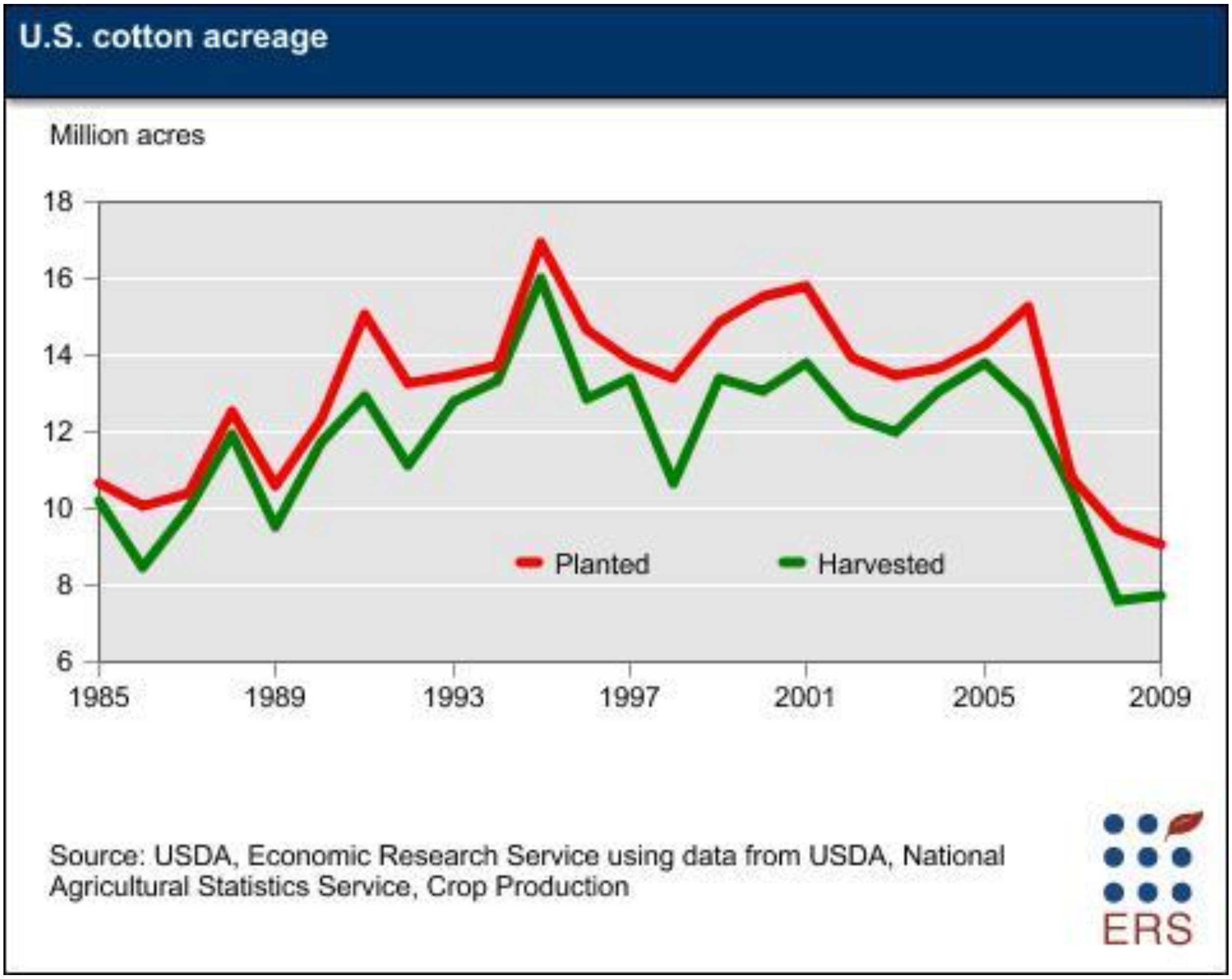


Figure 1.1. Total acres planted and harvested from 1985 to 2009 in United States, across all cotton types

Cotton production in this research is divided into 3 categories: non-irrigated Upland cotton, irrigated Upland cotton, and Pima cotton. The study timeframe is from 1990 through 2011. All the Pima cotton fields included in this research are irrigated. Figure 1.2 shows the shares of acres planted and abandoned by categories. It is obvious that the non-irrigated Upland cotton is the most likely to be abandoned, whereas Pima cotton is the least likely. The acreage abandonment of each category from 1990 through 2011, as shown in Figure 1.3, confirms the hypothesis.

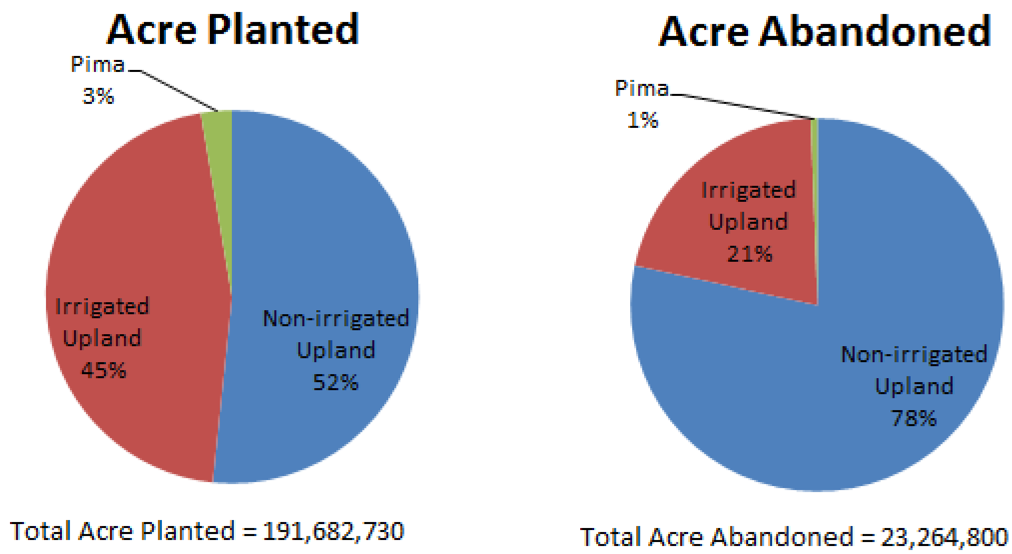


Figure 1.2. Share of acres planted and abandoned by category from 1990 through 2011

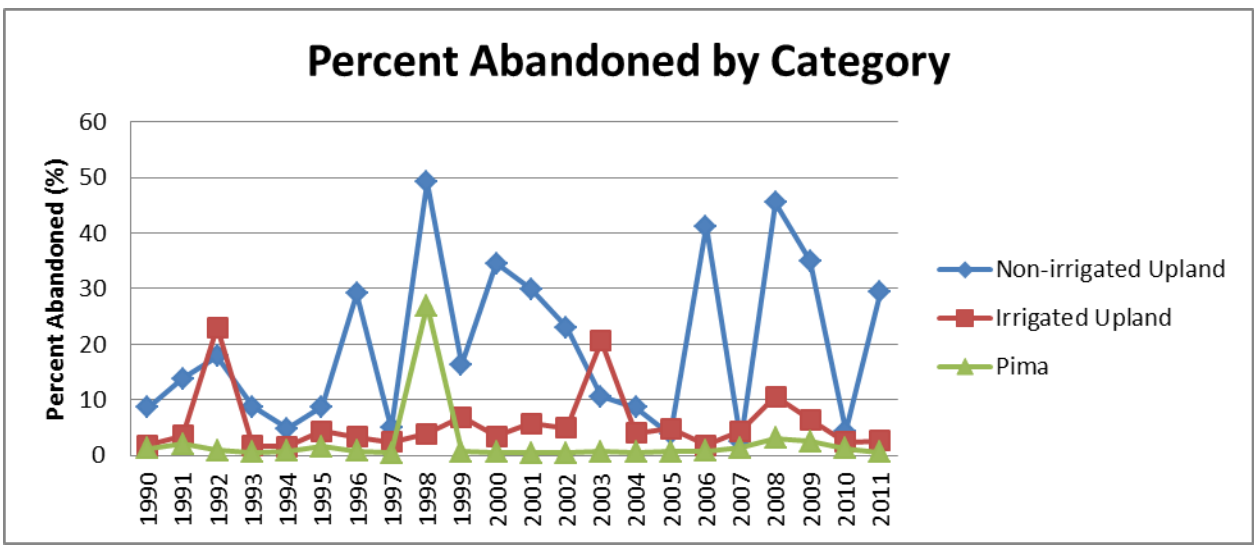


Figure 1.3. Percent abandoned of by category from 1990 through 2011

In this research, I have evaluated county-level cotton acreage abandonment of Arizona, Arkansas, California, Louisiana, Mississippi, New Mexico, Oklahoma, and Texas from 1990 through 2011. Those are the only states having separated irrigated/non-irrigated cotton records from 1990 through 2011 from QuickStats, USDA. A full description of variables included in this research is given in Chapter 5. For this analysis, I have used SAS 9.2 and Microsoft Excel 2010 for data management and Stata 12.1 for regression modeling and post-estimation predictions.

Chapter 2

Literature Review

2.1 Multivariate Analyses on Crop Abandonment

A few studies have been done estimating crop acreage abandonment. Michaels (Michaels 1983) used a time series Crop Reporting District (CRD) level model to estimate both the changes in large-area stress-tolerant crop yield and changes in abandonment, winter wheat in this research, from 1932 through 1975. The factors he included are changes in precipitation, temperature, and crop price in the western Great Plains. Crop yield is described as the total yield (measured in metric tons) over total hectares planted times 10. The formula for abandonment calculated as ((Acre Planted) – (Acre Harvested)) / (Acre Planted).

Michaels concludes that the abandonment is more likely to be associated with climate changes rather than economic factors. The overall model explained 77% of the total variance in abandonment, with 36.5% of the variation explained by the weather variables.

Michaels (Michaels 1985) extended the previous work described in Michaels (Michaels 1983). This paper extended the initial analysis to the spring wheat regions of North Dakota and South Dakota, and Minnesota. He also examined the effects of major weather factors with yield, rather than abandonment. The climate data consist of monthly mean temperature and total precipitation for May through August, as well as March to April. A principle component analysis was adopted to analyze the spatial correlation of abandonment. He then estimated abandonment on weather conditions, price, and yield. Again, he concluded that the abandonment is more likely due to climate conditions, rather than economic reasons.

Mendelsohn (2007) evaluated the crop failure contributed by climate, soil condition, and location. He used Agricultural Census data gathered in 1978, 1982, 1987, 1992, and 1997. The dependent variable was expressed as (failed cropland divided by all cropland), but in the denominator, he included idled land, and pastureland. He then regressed the crop failure rate on soil and climate variables. His conclusions are temperature and precipitation are likely to be significant, but the signs and significances varied by month. In addition, flatter terrains and soils with high water capacity reduced crop failure rates. Location was another important factor causing variations in crop failure rate. He commented that the results provided some insight into how global warming might influence crop failure rates.

2.2 Crop Abandonment, Insurance and Moral Hazard

Chen (2005) examined whether the insurance participation decision encouraged producers to abandon their crops. Data of individual insured units for her research were obtained from unpublished Risk Management Agency Corporate Database, USDA. The objective was to construct an intra-seasonal dynamic optimization model that incorporated crop producer's acreage abandonment decision with and without purchasing crop insurance. She concluded that insured farmers were more likely to abandon more crops to maximize the expected profits, when non-insured producers might take the risks to continue growing.

Chapter 3

Description of Data

3.1 Cotton Acres Planted/Harvested Data

This research covers 22 years (1990 ~ 2011) of cotton planted and harvested acreage records in 315 counties from 8 states in the Cotton Belt region of the United States. The 8 states included in this research are (in alphabetical order): Arizona (AZ), Arkansas (AR), California (CA), Louisianan (LA), Mississippi (MS), New Mexico (NM), Oklahoma (OK), and Texas (TX). Only those eight states have separated irrigated/non-irrigated cotton records available. The data include three types of cotton production: non-irrigated Upland cotton, irrigated Upland cotton, and Pima cotton. The annual cotton planted and harvested acre records are collected at the county-level. I have obtained the county-level cotton planted/harvested records using QuickStats tool, managed by National Agricultural Statistics Service, United States Department of Agriculture (National Agricultural Statistics Service, 2013). The options in QuickStats are:

- Select Commodity:

  Program: survey > Sector: CROPS > Group: FIELD CROPS > Commodity: COTTON > Category: AREA HARVESTED and AREA PLANTED > Data Item: COTTON, PIMA, IRRIGATED – ACRES HARVESTED and COTTON, PIMA, IRRIGATED – ACRES PLANTED and COTTON, UPLAND, IRRIGATED – ACRES HARVESTED and COTTON, UPLAND, IRRIGATED – ACRES PLANTED and COTTON, UPLAND, NON-

IRRIGATED – ACRES HARVESTED and COTTON, UPLAND, NON-IRRIGATED – ACRES PLANTED > Domain: TOTAL

- Select Location:

Geographic Level: COUNTY > State: ARKANSAS and CALIFORNIA and LOUISIANA and MISSISSIPPI and NEW MEXICO and OKLAHOMA and TEXAS > Year: 1990 through 2011 > Period Type: ANNUAL

The records of cotton production in Arizona are not listed in the data option. Given the fact that all the cotton fields growing in Arizona were irrigated, I have chosen the option "COTTON, PIMA, IRRIGATED – ACRES HARVESTED and COTTON, PIMA, IRRIGATED – ACRES PLANTED and COTTON, UPLAND – ACRES HARVESTED and COTTON, UPLAND – ACRES PLANTED" under "Data Item" tab and assigned all the planted/harvested acre records as irrigated Upland cotton records. I have not used the "Data Item" option "COTTON, UPLAND – ACRES HARVESTED and COTTON, UPLAND – ACRES PLANTED" for all the states is due to the fact that I could not fully separate the irrigated Upland cotton records from non-irrigated Upland cotton records.

I then have aggregated the records into one spreadsheet. All the observations with the county name "Other (COMBINED)" were omitted. There are only 5 records of non-irrigated Pima cotton, and I have deleted those records as well. All Pima cotton included in the dataset is irrigated. Some county records do not meet USDA's publishing standard. In addition, for this reason, the following records were omitted as well:.

- Irrigated Upland cotton from Otero County, NM in 1995

- Non-irrigated Upland cotton from Starr County, TX in 2006
- Pima cotton from Culberson County, TX in 1990
- Pima cotton from Ward County, TX in 1991

With the aggregated cotton production spreadsheet, I have created the cotton planted/harvested record dataset. A county can have up to three data entries in a specific year, covering non-irrigated Upland cotton records, irrigated Upland cotton records, and Pima cotton records with the help of two binary variables, d_irri and d_pima. d_irri takes the value of 0 and d_pima takes the value of 0 to control for non-irrigated upland cotton records, d_irri takes the value of 1 and d_pima takes the value of 0 to control for irrigated Upland cotton records, d_irri takes the value of 1 and d_pima takes the value of 1 to control for Pima cotton records.

In the final step, I calculated the acreage abandonment (labeled as pc_aban) using Equation 3.1.

Acreage abandonment = (planted acreage-harvested acreage) / (planted acreage) * 100 (3.1)

If pc_aban value is missing for a county in a certain year, it means this county did not have the type of cotton production indicated by d_irri and d_pima, and this observation is deleted. The cleaned cotton acres planted/harvested dataset has 6,501 records, organized in an unbalanced panel fashion. The numbers with respective percentage of pc_aban in each abandonment percentile is shown in Figure 3.1. The descriptive statistics of pc_aban is given in Table 3.1. All categories have a maximum abandonment of 100% except for Pima cotton. However, mean abandonment of all the categories never exceeds 15%. The category having the highest mean abandonment is non-irrigated Upland cotton, followed by irrigated Upland cotton. The mean abandonment of irrigated Upland cotton is 4.3%, dropped 10% from the non-irrigated Upland cotton. The Pima cotton category has the lowest mean abandonment, which is 3.2%.

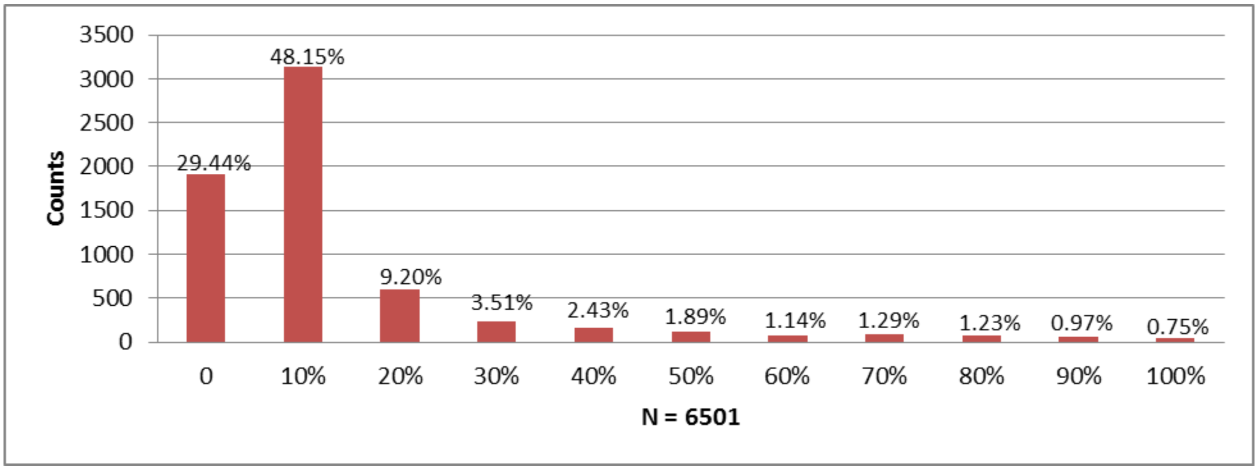


Figure 3.1. Numbers with respective percentage of percentage abandoned in each percentile

Data item	N	Maximum	Minimum	Mean	5%	95%
Full dataset	6,501	1	0	9.43	0	53.85
Non-irrigated Upland	3,384	1	0	14.26	0	69.65
Irrigated Upland	2,816	1	0	4.3	0	20
Pima	301	99.39	0	3.17	0	12.5

Table 3.1. Descriptive statistics of percentage abandoned

3.2 Weather Data

The weather records are provided by National Climatic Data Center (NCDC), National Environmental Satellite, Data, and Information Service (NESDIS), National Oceanic and Atmospheric Administration (NOAA), United States Department of Commerce (National Climatic Data Center, National Oceanic and Atmospheric Administration, 2013). The records are monthly, climate-division-level data collected from 1990 through 2010. The items recorded are described in Table 3.2. I have only used values of SP09 observed in September, PCP, and TMP in this research.

Item	Unit	Description
PCP	In.	Precipitation Index
TMP	°F	Temperature Index
PDSI	Unitless	Palmer Drought Severity Index
PHDI	Unitless	Palmer Hydrological Drought Index
ZNDX	Unitless	Palmer Z-Index
PMDI	Unitless	Modified Palmer Drought Severity Index
CDD	Unitless	Cooling Degree Days
HDD	Unitless	Heating Degree Days
SP01	Unitless	Standard Precipitation Index Over 1 Month
SP02	Unitless	Standard Precipitation Index Over 2 Months
SP03	Unitless	Standard Precipitation Index Over 3 Months
SP06	Unitless	Standard Precipitation Index Over 6 Months
SP09	Unitless	Standard Precipitation Index Over 9 Months
SP12	Unitless	Standard Precipitation Index Over 12 Months
SP24	Unitless	Standard Precipitation Index Over 24 Months

Table 3.2. Weather record items with respective units and descriptions

I then have manually recorded the counties within each climate division, based on the climate division map provided by Climate Prediction Center (CPC), NOAA, (Climate Prediction Center, National Oceanic and Atmospheric Administration, 2013). The climate division borders in California and New Mexico do not match the county borders, so I have picked the climate division covering the most area of a county in those two states as the designated climate division.

The final weather dataset includes county-level seasonal weather data and modifications on sp_09 records. Season 1 is the planning season, from January through March; season 2 is the planting season, from April through June; season 3 is the growing season, from July through

September; season 4 is the harvesting season, from October through December. In order to create the seasonal PCP and TMP records, I have summed monthly PCP values and averaged TMP values on a 3-month basis. The sp_09 index measures the average precipitation of the current nine months comparing to the long-term norm. I have used the sp_09 index measured in September for this analysis. For the states included in this analysis, September is either the end of the growing season or the beginning of the harvesting season, and by this time farmers have either abandoned some (even all) of their cotton fields, or made their decisions about abandoning their crops or not. Therefore, the precipitation of September can give a hint on the rainfall received of the crop season, in addition to the direct measurement of precipitations. I suspect sp_09 having a non-linear effect on crop abandonment, so I have derived 2 different variables, dry_sp09 and wet_sp09, regarding to sp_09 index values to count for the non-linear effect. I first have created 2 binary variables, d_dry and d_wet. d_dry equals to 1 if sp_09 index value is negative, otherwise 0. d_wet equals to 1 if sp_09 index value is positive, otherwise 0. dry_sp09 equals to the product of d_dry and the absolute value of sp_09 index. wet_sp09 equals to the product of d_wet and the absolute value of sp_09 index. At last, I have matched the weather variables into the county – climate division spreadsheet I created in the previous step.

3.3 Cotton Price Data

The cotton price data are monthly observed national level data, provided by National Cotton Council of America (National Cotton Council of America, 2013). For both non-irrigated and irrigated Upland cotton prices, I have used NYCE (New York Cotton Exchange) Near December Contract Price, starting from 1990. For Pima cotton prices, I used ELS Spot Prices, starting from

1990. I then have taken the difference by subtracting May prices from September prices, and deflated to 2005 dollars.

3.4 Full Dataset

I have merged the weather data and cotton price data into the cleaned cotton acres planted/harvested dataset by using SQL procedure with "left join" option. The criteria were county names, state, and year. A full description of full dataset will be given in the next chapter.

The county-level data used in this research have omitted the observations under the county name "Other (COMBINED)". Those observations are the combined cotton records from the counties that do not meet the USDA's standard for data publishing on a county-level within the same agricultural district. This procedure reduces the state total coverage. Appendix A lists the annual state-level planted acre coverage from the full dataset, and annual state-level planted acre coverage directly from NASS, USDA for non-irrigated Upland cotton, irrigated Upland cotton, and Pima cotton. In most cases, the differences in state-level acre coverage are within 5%. New Mexico State stopped recording irrigated and non-irrigated Upland cotton records separately from 2008, so the New Mexico county-level records from 2008 through 2011 are Upland cotton planted/harvested records, assuming all the cotton fields in New Mexico were irrigated after 2007. The irrigated upland cotton record differences in Arkansas stay below 2% across two reports until 2000, and then stay above 10% afterwards. The differences exceed 40% after 2008. The irrigated Upland coverage is better than non-irrigated Upland in Arkansas. The differences stay below 10% until 2008, and then exceed 10% afterwards. The difference goes up to 29.46% in 2011. The data coverage in Arizona, California, New Mexico, and Texas is good, except for Texas in 2011. The county-level report only covers 3.11% for irrigated Upland cotton and 9.46%

for non-irrigated Upland cotton. The data coverage in Louisiana is good until 2007. However, the coverage is not so good for Mississippi and Oklahoma in general.

3.5 Census Data

The agriculture census data are observed on 5-year basis, provided by USDA (United States Department of Agriculture, 2013). The advantage of using agriculture census data is that the census data have numbers of farms growing cotton. However, the drawback is that the census data are recorded every 5 years, so in the census subset only the observations from 1992, 1997, 2002, and 2007 are used, reducing the number of observations. There are minor variations of obtaining census data from different states. Here I will demonstrate the procedure of obtaining Arizona census data from 2007 as an example. After going to the website, I have picked 2007 under the “Census Publications” menu, clicked on “2007 Census of Agriculture” link under “Publications” section, clicked on “Full 2007 Census Report” link under “2007 Census Results” section, clicked on “All Counties by State by Table” link under “State and County Reports” section, selected “Arizona” from the link below the map, then clicked on the “Text” link and downloaded the full text file. Under Table 26, I have selected the table with the title “COTTON, ALL (BALES)”. At last, I have selected the numbers under “Farms” column in “Harvested” section in 2007 for all the recorded counties and imported the columns into an Excel spreadsheet. I have repeated the procedure 32 times and collected the county-level farm counts for all 8 states in 1992, 1997, 2002, and 2007. I have merged the census farm count dataset and full dataset by using SQL procedure with “left join” option. The criteria were county names, state, and year. The number of observations in the census subset is reduced to 1,204.

Chapter 4

Empirical Model and Methodology

4.1 Conceptual Framework

The objective of this research is to capture cotton field acre abandonment, giving the weather conditions and economic incentives, cotton type, and adoption of irrigation. Data from 8 states in South/Southwest U. S. from 1990 through 2011 are selected for this research. In this analysis, I have focused more on the weather impact on crop abandonment.

A tobit model or an alternative hurdle model can both predict the positive outcomes required for this research. However, an important limitation of the standard tobit model is that a single mechanism determines the choice between $y = 0$ versus $y > 0$ and the amount of y given $y > 0$ (Wooldridge 2002). In addition, tobit model assumes the distribution of error terms is normal distribution. However, the acreage abandonment is following a gamma distribution, instead of normal distribution. Meanwhile, the second tier of hurdle model recommends taking the natural log of the dependent variable, which will fix of the gamma distribution of abandonment in the full dataset. Therefore, I have chosen hurdle model over tobit model for this analysis.

4.2 Hurdle Model

Described by Wooldridge (Wooldridge 2002), the hurdle model is also known as two-tiered model. The first tier is predicting the probability of having a positive outcome, and the second tier is predicting the amount of the positive outcome.

For tier 1 estimation, I have chosen a probit model. Probit model is a special case of binary response model. The binary response model takes the general form:

$$P(y = 1|x) = P(y^{*} > 0|x) = P(e > -x\boldsymbol{\gamma}|x) = 1 - G(-x\boldsymbol{\gamma}) = G(x\boldsymbol{\gamma}) \qquad (4.1)$$

Where $\mathbf{y}^{*} = x\boldsymbol{\beta} + \mathbf{e}$, $\quad y = 1\ [y^{*} > 0]$, assuming e is symmetric above zero.

The probit model is a special case of Equation 4.1 with:

$$G(z) \equiv \Phi(z) \equiv \int_{-\infty}^{z} \varphi(v)dv \qquad (4.2)$$

Where φ(z) is the univariate standard normal density

$$\varphi(z) = (2\pi)^{-1/2} \exp(-z^{2}/2) \qquad (4.3)$$

$$\text{So } \hat{P}(y = 1|x) = \Phi(x\,\hat{\gamma}) \equiv \int_{-\infty}^{\mathbf{x}\,\hat{\gamma}} \varphi(v)dv \qquad (4.4)$$

Where $\hat{\gamma}$ is the estimated parameter vector of the probit model.

For the second tier, Cragg (1971) suggested to use OLS specification from the regression log(y) on x, where y are all positive. This estimation is simple because it is assumed that log(y) follows a classic linear model, given that all values of y are positive. The distribution of abandonment is a gamma distribution. By taking the natural log of the abandonment the distribution is assumed to follows a normal distribution. The expected value of lognormal OLS model is:

$$E(\hat{y}|x, y > 0) = \exp(x\hat{\boldsymbol{\beta}} + \hat{\sigma}^{2}/2) \qquad (4.5)$$

Where $\hat{\boldsymbol{\beta}}$ is the parameter estimate vector of the OLS model, $\hat{\sigma}^{2}$ is the model variance.

So the expected value of the hurdle model can be expressed as:

$$E(\hat{y}|x) = \hat{P}(y = 1|x) * E(\hat{y}|x, y > 0) = \Phi(x\hat{\gamma}) \exp(x\hat{\beta} + \hat{\sigma}^2/2) \tag{4.6}$$

Due to the restriction on having only the positive y values in the second tier OLS model, I cannot predict abandoned percentage for the counties with 0% abandonment in the dataset.

Chapter 5

Variables and Summary Statistics

5.1 Description of Variables

The dependent variables for the two-step hurdle model are: d_aban for probit model and ln_pc_aban for OLS model. d_aban is a binary variable, taking value of 1 when a positive abandonment is observed, otherwise 0. ln_pc_aban is the natural log of actual percent abandoned (pc_aban) calculated only when a positive percent abandoned is observed. Of all 6,501 records in the full dataset, 4,587 records show positive abandoned, so the OLS model is only applied to those records. Of all 1,204 records in the census subset, 811 records show positive percent abandoned, again the OLS model is only applied to those records.

In general, more rainfall and higher temperature are welcome by cotton farmers, but excess moisture and/or extreme temperature conditions can cause abandonment. Therefore, I expect the signs of the first order of the seasonal precipitation and temperature variables to be negative, and the signs of the second order seasonal precipitation and temperature variables to be positive. However, pcp_s4 and tmp_s3 are exceptions. Season 4 is the harvesting season, and more rainfall can have negative impact on the harvesting process. Season 3 is the growing season, and the higher temperature is healthier for cotton growth. Therefore, I expect the signs of the first order and second order of pcp_s4 and tmp_s3 to be positive and negative, respectively. The 9-month standard precipitation index (sp_09) measured in September is an indicator of the overall rainfall condition of the crop year, and the two variables I have derived from sp_09 can measure the impact of the higher and lower than normal rainfall on the same scale from the same direction, the higher value the index is, the wetter/dryer the condition is. Less rainfall can increase the abandonment, so I expect the signs of dry_sp09 are positive in both models. Excess moisture can contribute to the cotton abandonment, even though the farmers welcome as much rainfall as possible, so wet_sp09 has counter effects in the hurdle model. I expect the signs of wet_sp09 in the probit model and OLS model to be negative and positive respectively. If the price increases from May to September, the farmers have a stronger incentive to keep the cotton fields, so I expect the signs of price_may_sep_diff_def to be negative in both the probit model and OLS model. Like wet_sp09, farm count variable (farm) from census subset can have counter effects in probit model and OLS model. With a larger collection of farms in a county, the possibility to observe cotton abandonment is higher, but the actual abandonment may decrease. Therefore, I suspect the signs of farm in probit model and OLS model to be positive and negative. A full description of the independent variables included in this analysis and their expected signs are given in Table 5.1.

Figure 5.1 illustrates the nation-wide mean precipitation and temperature in each season from 1990 through 2011. Figure 5.2 illustrates the difference in price from 1990 through 2011. Table 5.2 shows the minimum, maximum, mean, 5 percentile, and 95 percentile values of each variable. The national mean precipitation values in each season show the precipitation varies among years, and there is no correlation between seasonal precipitations. However, there is a general trend of decreasing in precipitation between 1990 through 2011. On the other hand the seasonal temperature variables behave quite normally. There is a trend of average temperature increasing universally from 2009.

Variable Name	Description (Unit)	Expected Sign (probit/OLS)
pcp_s1	Summed precipitation, January - March (In.)	Negative/Negative
pcp_s1_sqr	Squared term of summed precipitation, January - March (Sqr. In.)	Positive/Positive
pcp_s2	Summed precipitation, April - June (In.)	Negative/Negative
pcp_s2_sqr	Squared term of summed precipitation, April - June (Sqr. In.)	Positive/Positive
pcp_s3	Summed precipitation, July - September (In.)	Negative/Negative
pcp_s3_sqr	Squared term of summed precipitation, July - September (Sqr. In.)	Positive/Positive
pcp_s4	Summed precipitation, October - December (In.)	Negative/Negative
pcp_s4_sqr	Squared term of summed precipitation, October - December (Sqr. In.)	Positive/Positive
tmp_s1	Averaged temperature, January - March (°F)	Negative/Negative
tmp_s1_sqr	Squared term of averaged temperature, January - March (Sqr. °F)	Positive/Positive
tmp_s2	Averaged temperature, April - June (°F)	Negative/Negative
tmp_s2_sqr	Squared term of averaged temperature, April - June (Sqr. °F)	Positive/Positive
tmp_s3	Averaged temperature, July - September (°F)	Negative/Negative
tmp_s3_sqr	Squared term of averaged temperature, July - September (Sqr. °F)	Positive/Positive
tmp_s4	Averaged temperature, October - December (°F)	Negative/Negative
tmp_s4_sqr	Squared term of averaged temperature, October - December (Sqr. °F)	Positive/Positive
dry_sp09	Absolute value of sp_09 index if sp_09 < 0 (Unitless)	Positive/Positive
wet_sp09	Absolute value of sp_09 index if sp_09 > 0 (Unitless)	Positive/Negative
d_irri	Binary if irrigation is adopted (Unitless)	Negative/Negative
d_pima	Binary if Pima cotton is grown (Unitless)	Negative/Negative
price_may_sep_diff_def	Predicted price difference between September and May, deflated into 2005 dollars (cents/pound)	Negative/Negative
farm	Numbers of farms growing cotton in a county (Unitless)	Positive/Negative

Table 5.1. Descriptions of the variables

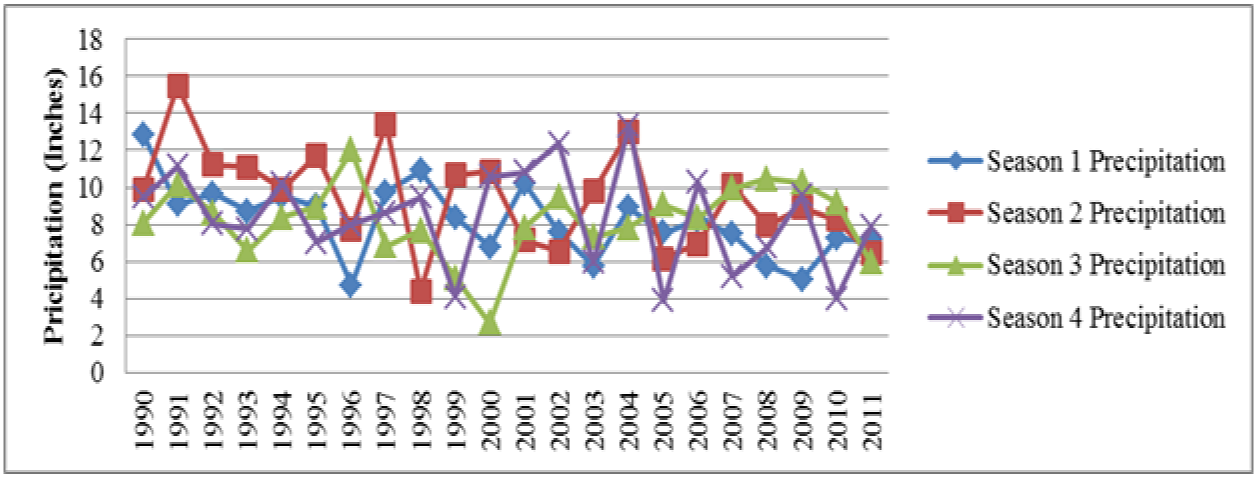


Figure 5.1a. Mean precipitation from 1990 through 2011

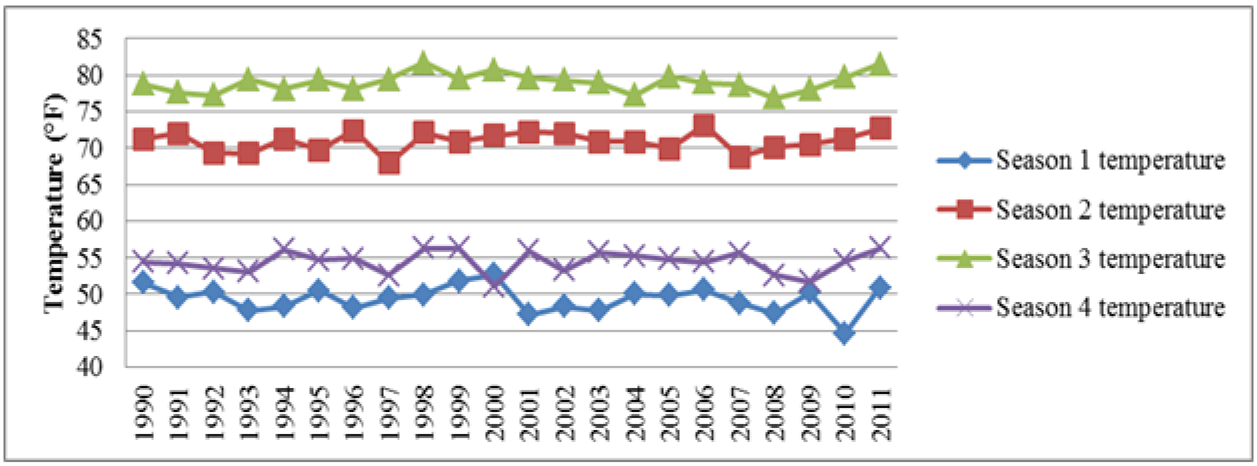


Figure 5.1b. Mean temperature from 1990 through 2011

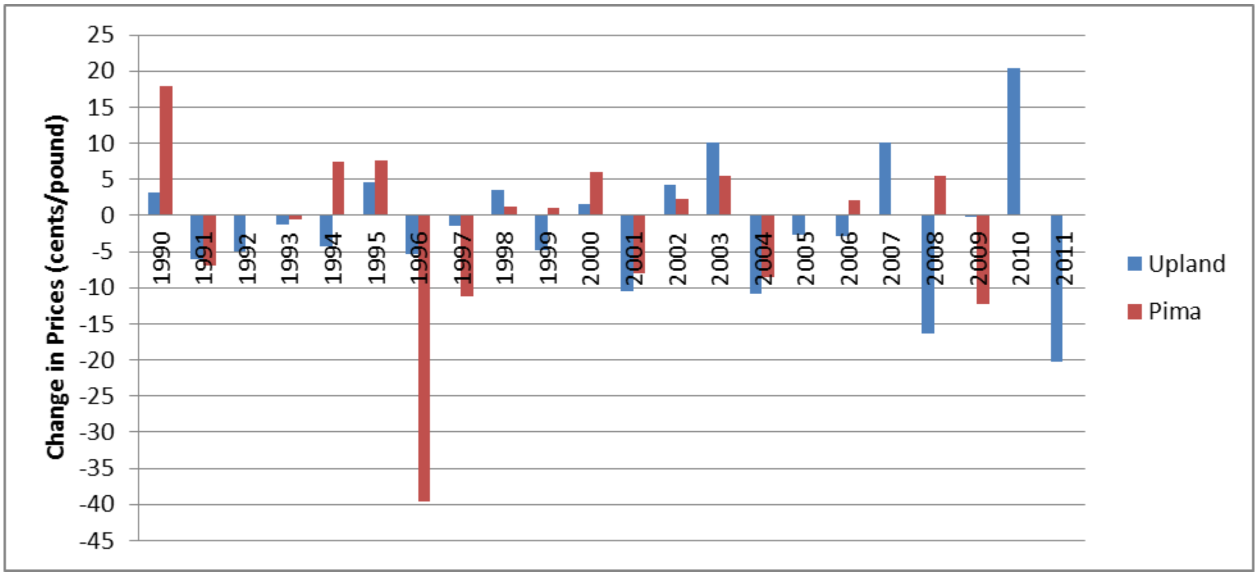


Figure 5.2. Change in price between May and September for Upland cotton and Pima cotton from 1990 through 2011

Variable	Minimum	Maximum	Mean	5 Percentile	95 Percentile
d_irri	0	1	0.48	0	1
d_pima	0	1	0.05	0	0
pcp_s1	0.06	39.23	8.51	1.02	20.59
pcp_s2	0	35.82	9.75	0.91	19.12
pcp_s3	0	24.43	8.21	1.01	15.66
pcp_s4	0	29.74	8.58	1.36	20.16
tmp_s1	37.37	70.03	49.47	42.13	59.47
tmp_s2	54.77	82.80	70.87	63.90	77.53
tmp_s3	68.97	92.50	78.95	73.63	84.50
tmp_s4	42.30	72.10	54.46	47.13	63.80
dry_sp09	0	3.09	0.27	0	1.15
wet_sp09	0	2.53	0.46	0	1.68
price_may_sep _diff_def	-39.66	20.47	-1.07	-10.90	10.19

Table 5.2. Variable descriptive statistics

Table 5.3 gives the means and standard deviations of all the variables, as well as the means and standard deviations of percent abandoned between irrigated Upland and non-irrigated Upland. All the mean values of precipitation and temperature variables are higher for non-irrigated Upland, and Z tests reject the hypotheses that the mean values are the same. A Z test was carried out between the mean abandonment of irrigated Upland and non-irrigated Upland. The test result also rejects the hypothesis. However, Z tests fail to reject the hypotheses that the mean values of Sp09 index in wetter conditions, and the price difference between May and September are the same.

	Irrigated Upland Mean	Irrigated Upland Standard Deviation	Non-Irrigated Upland Mean	Non-Irrigated Upland Standard Deviation
pcp_s1	7.95	6.62	9.24	6.31
pcp_s2	8.77	5.83	11.21	5.40
pcp_s3	7.66	4.24	9.08	4.03
pcp_s4	7.82	6.23	9.61	6.04
tmp_s1	48.81	5.15	49.89	5.33
tmp_s2	70.36	4.21	71.40	3.50
tmp_s3	78.58	3.68	79.27	2.82
tmp_s4	53.80	4.85	54.94	4.93
dry_sp09	0.29	0.47	0.25	0.41
wet_sp09	0.45	0.58	0.47	0.59
price_may_sep_diff_def	-1.08	7.43	-1.06	6.80
percent abandoned of all observations	4.30	10.70	14.26	21.79
percent abandoned for observations having positive abandonment	8.03	13.56	16.35	22.58

Table 5.3. Mean and standard deviation of each variable and abandonment

Chapter 6

Results and Implications

6.1 Regression Results

As discussed in Chapter 5, I have chosen the hurdle model over the tobit model for estimating both the weather and economic impacts on cotton acreage abandonment. The hurdle model consists of two steps. The first step is the probit model, predicting the possibility of a county abandoning any of the planted cotton fields under weather and economic conditions in a given year. The second step is the log-linear ordinary least square (OLS) model, predicting the percentage of planted cotton field abandoned under weather and economic conditions in a given year, given a positive abandonment is observed. The hurdle model is applied both on the full dataset and census subset. A brief discussion of the parameter estimates of the full dataset and census subset are given in the next two sections respectively. Due to the nature of the census records, the variable representing numbers of farms (farm) is only used in the census subset. The adjusted R^2 reported is the adjusted count R^2, calculated with the following equation:

$$\text{Adjusted Count } R^2 = \frac{(\text{number of correct prediction}) - (\text{number of most common outcome})}{(\text{number of observations}) - (\text{number of most common outcome})} \quad (6.1)$$

6.1.1 Parameter Estimates of Full Dataset

The parameter estimates with respective z-values or t-values, p-values, and 95% confidence intervals of the probit model and OLS model for the full dataset are recorded in Table 6.1 and Table 6.2 respectively.

The probit model predicts the likelihood of abandoning any cotton fields in a given year, when the OLS model predicts the percentage of acres is abandoned in the same year, under the precondition that a positive abandonment is observed. Therefore, some variables may have opposite effects in different models, for example wet_sp09. In general, cotton growth needs a large amount of water and warm environment, but excess moisture and/or extreme temperature may increase the likelihood of abandoning the planted crops. Therefore, the expected signs of the precipitation and temperature variables are negative for the first order and positive for the second order.

For the probit model of the full dataset, the adjusted count R^2 is 0.26, and it correctly predicts 5,091 binary outcomes out of all 6,501 observations. The prediction accuracy is 78.31%. All the parameter estimates have expected signs, except for tmp_s2 and tmp_s2_sqr. Most of the variables are significant at 5% level of confidence. Variable pcp_s2_sqr has a p-value of 0.35. Season 2 is from April through June, and it is the planting season for all the study states. Enough precipitation can create healthy growing conditions for cotton, so rainfall has a more of a linear effect, rather than a quadratic effect on the possibility to abandon any cotton field. Variable pcp_s4 and pcp_s4_sqr have p-values of 0.16 and 0.83 respectively. This may be because season 4 is the harvesting season and rainfall affects the planting season (season 2) and growing season (season 3) more than harvesting season. The temperature variables in season 1 and season 2 are all insignificant at 5% confidence level. This may be caused by the fact that season 1 and season 2 are the initial stage of cotton growing circle and temperature is not an important factor yet. Interestingly, the p-value of the cotton price difference between May and September (price_may_sep_diff_def) is 0.06, which is not significant at 5% level of significance.

Variable	Parameter Estimate	Standard Error	z	P>|z|	[95% Confidence Interval]	
d_irri	-1.35	0.042	-32.05	0	-1.43	-1.26
d_pima	-0.57	0.086	-6.62	0	-0.74	-0.40
pcp_s1	-0.12	0.013	-9.22	0	-0.14	-0.093
pcp_s1_sqr	0.0029	0.00045	6.48	0	0.0021	0.0038
pcp_s2	-0.03	0.011	-2.45	0.014	-0.05	-0.0056
pcp_s2_sqr	0.00036	0.00038	0.94	0.35	-0.00039	0.0011
pcp_s3	-0.12	0.017	-7.04	0	-0.15	-0.08
pcp_s3_sqr	0.0037	0.00077	4.76	0	0.0022	0.0052
pcp_s4	0.020	0.014	1.42	0.16	-0.0075	0.047
pcp_s4_sqr	-0.00011	0.00051	-0.21	0.83	-0.0011	0.00088
tmp_s1	-0.10	0.076	-1.35	0.18	-0.25	0.046
tmp_s1_sqr	0.00072	0.00077	0.93	0.35	-0.00079	0.0022
tmp_s2	0.15	0.15	1	0.32	-0.15	0.45
tmp_s2_sqr	-0.00083	0.0011	-0.76	0.45	-0.0030	0.0013
tmp_s3	0.69	0.20	3.52	0	0.30	1.08
tmp_s3_sqr	-0.0044	0.0012	-3.64	0	-0.0068	-0.0020
tmp_s4	-0.40	0.10	-3.93	0	-0.61	-0.20
tmp_s4_sqr	0.0035	0.00093	3.74	0	0.0017	0.0053
dry_sp09	-0.29	0.053	-5.59	0	-0.40	-0.19
wet_sp09	0.085	0.041	2.06	0.039	0.0042	0.17
price_may_sep_diff_def	-0.0048	0.0026	-1.85	0.064	-0.0099	0.00028
intercept	-15.87	7.29	-2.18	0.029	-30.16	-1.58
adjust count R^2	0.26					
sample size	6501					

Table 6.1. Parameter estimates with respecting z-values, p-values, and 95% confidence intervals of the probit model from the full dataset

For the OLS model, the adjusted R^2 is 0.35 and the model variance (σ^2) is 1.44. The predicted percentages of acre abandoned of all the observations have a correlation of 0.53 with the actual percentages of acre abandoned. Precipitation in season 2 turns out to be an interesting factor regarding percentage abandonment. Both pcp_s2 and pcp_s2_sqr have the opposite signs from my expectation, and pcp_s2_sqr is significant at 5% level of significance when pcp_s2 is not. This is the only case when the quadratic form of a variable is more important than the linear form of that variable in the full dataset. I will discuss more of this result in the marginal effect section.

Both the precipitation variables in season 4 are of the predicted signs but insignificant, maybe because rainfall matters more in the planting season and growing season but not as much in the harvesting season, but less rainfall is better for the harvesting process. Neither the temperature variables in season 1 are of the predicted signs and significant. Higher temperature may lead to a pest outbreak, and temperature does not have a strong impact on field abandonment because season 1 is only the planning season. dry_sp09 variable turns out to be insignificant in this model as well.

Variable	Parameter Estimate	Standard Error	t	P>|t|	[95% Confidence Interval]	
d_irri	-1.25	0.043	-29.31	0	-1.34	-1.17
d_pima	-0.27	0.12	-2.21	0.027	-0.50	-0.030
pcp_s1	-0.23	0.013	-17.56	0	-0.26	-0.21
pcp_s1_sqr	0.0060	0.00050	12.13	0	0.0051	0.0070
pcp_s2	0.015	0.012	1.28	0.20	-0.0080	0.038
pcp_s2_sqr	-0.00093	0.00039	-2.41	0.016	-0.0017	-0.00017
pcp_s3	-0.11	0.017	-6.37	0	-0.14	-0.074
pcp_s3_sqr	0.0034	0.00077	4.42	0	0.0019	0.0049
pcp_s4	0.025	0.013	1.91	0.056	-0.00061	0.051
pcp_s4_sqr	-0.00048	0.00049	-0.97	0.33	-0.0015	0.00049
tmp_s1	0.078	0.069	1.13	0.26	-0.058	0.21
tmp_s1_sqr	-0.0010	0.00070	-1.49	0.14	-0.0024	0.00033
tmp_s2	-0.61	0.17	-3.68	0	-0.93	-0.28
tmp_s2_sqr	0.0049	0.0012	4.09	0	0.0025	0.0072
tmp_s3	1.44	0.21	6.83	0	1.03	1.85
tmp_s3_sqr	-0.0091	0.0013	-6.95	0	-0.012	-0.0066
tmp_s4	-0.70	0.091	-7.7	0	-0.88	-0.52
tmp_s4_sqr	0.0057	0.00083	6.92	0	0.0041	0.0074
dry_sp09	0.090	0.053	1.69	0.091	-0.014	0.20
wet_sp09	0.22	0.041	5.35	0	0.14	0.30
price_may_sep_diff_def	-0.014	0.0027	-5.14	0	-0.019	-0.0085
intercept	-15.15	8.29	-1.83	0.068	-31.41	1.10
adjusted R^2	0.35					
sample size	6501					

Table 6.2. Parameter estimates with respecting t-values, p-values, and 95% confidence intervals of the OLS model from the full dataset

6.1.2 Parameter Estimates of Census Subset

The parameter estimates with respecting z-values or t-values, p-values, and 95% confidence intervals of the probit model and OLS model for the census subset are recorded in Table 6.3 and Table 6.4 respectively. With the census subset, I construct the model with an additional variable recording numbers of farms growing cotton in a certain county in a given year (farm). Census reports are only recorded every five years, so in the census subset I have observations only in 1992, 1997, 2002, and 2007, and the temporal variance of the data is smaller when comparing to the full dataset.

From the probit model, the adjusted count R^2 is 0.34, and it correctly predicts 943 binary outcomes out of all 1,204 observations. The predicting accuracy is 78.32%. The signs of the variable coefficient estimates are presumed to be the same as the signs of the same as the variable coefficients from the full dataset. However this is not the case. Both of the precipitation variables in season 2 are having the opposite signs as they were in the probit model from the full dataset, and the p-values are both 0. This indicates that the precipitation is important in season 2 for the census subset. I will discuss more in the marginal effect section. In addition, both the season 3 temperature variables are of the opposite signs, but they are insignificant at 5% level of confidence. Both the wet_sp09 variable and price_may_sep_diff_def variable have positive signs, which are the opposite of the expected signs as well, when they both give farmers incentives to keep their crops. As I have discussed before, more rainfall is welcome by the farmers, and if the cotton price increases during the crop year farmers are less likely to abandon their cotton fields. What's more, price_may_sep_diff_def is significant when both dry_sp09 and wet_sp09 is not.

Variable	Parameter Estimates	Standard Error	z	P>|z|	[95% Confidence Interval]	
d_irri	-1.44	0.10	-13.8	0	-1.65	-1.24
d_pima	-0.85	0.23	-3.74	0	-1.30	-0.40
pcp_s1	-0.29	0.052	-5.47	0	-0.39	-0.18
pcp_s1_sqr	0.010	0.0021	4.92	0	0.0062	0.014
pcp_s2	0.27	0.047	5.78	0	0.18	0.36
pcp_s2_sqr	-0.013	0.0021	-6.32	0	-0.017	-0.0090
pcp_s3	-0.21	0.050	-4.23	0	-0.31	-0.11
pcp_s3_sqr	0.0041	0.0021	1.93	0.054	-7.1E-05	0.0084
pcp_s4	0.072	0.044	1.65	0.099	-0.013	0.16
pcp_s4_sqr	-0.00019	0.0015	-0.13	0.90	-0.0031	0.0027
tmp_s1	-0.85	0.39	-2.16	0.031	-1.61	-0.078
tmp_s1_sqr	0.0083	0.0040	2.06	0.039	0.00042	0.016
tmp_s2	2.40	0.54	4.4	0	1.33	3.46
tmp_s2_sqr	-0.016	0.0039	-4.22	0	-0.024	-0.0088
tmp_s3	-0.87	0.66	-1.33	0.18	-2.16	0.41
tmp_s3_sqr	0.0052	0.0041	1.28	0.2	-0.0028	0.013
tmp_s4	-0.18	0.37	-0.49	0.63	-0.91	0.55
tmp_s4_sqr	0.0014	0.0035	0.39	0.70	-0.0055	0.0082
dry_sp09	0.11	0.15	0.71	0.48	-0.19	0.40
wet_sp09	0.22	0.14	1.56	0.12	-0.057	0.50
price_may_sep_diff_def	0.034	0.016	2.09	0.037	0.0021	0.066
farm	0.0059	0.00076	7.72	0	0.0044	0.0074
intercept	-21.67	18.95	-1.14	0.25	-58.81	15.47
adjusted count R^2	0.34					
sample size	1204					

Table 6.3. Parameter estimates with respecting z-values, p-values, and 95% confidence intervals of the probit model from the census subset

In the OLS model, the adjusted R^2 is 0.47 and the model variance (σ^2) is 1.11. The predicted percentages of acre abandoned have a correlation of 0.68 with the actual percentages acre abandoned. Precipitation variables in season 4 have the opposite signs in this model as they have in the OLS model for the full dataset. More moisture is harmful for the harvesting process, so the percent of cotton field abandoned should be monotonically increasing at a decreasing rate with the increase of rainfall in season 4. However, the parameter estimates of season 4 precipitation from this model show the opposite effect. Again, like pcp_s2 and pcp_s2_sqr in OLS model from the full dataset, pcp_s4_sqr turns out to be a lot more significant in this model than pcp_s4, even though neither of them is significant at 5% level. In addition, tmp_s1, tmp_s1_sqr, dry_sp09, and price_may_sep_diff_dep are having the opposite signs and they are not significant at 5% level of confidence.

Variabe	Parameter Estimate	Standard Error	t	P>|t|	[95% Confidence Interval]	
d_irri	-0.53	0.095	-5.58	0	-0.72	-0.34
d_pima	-0.25	0.25	-0.98	0.33	-0.75	0.25
pcp_s1	-0.074	0.051	-1.45	0.15	-0.17	0.026
pcp_s1_sqr	0.000059	0.0021	0.03	0.98	-0.0040	0.0041
pcp_s2	0.21	0.045	4.6	0	0.12	0.30
pcp_s2_sqr	-0.0071	0.0020	-3.54	0	-0.011	-0.0032
pcp_s3	-0.21	0.045	-4.73	0	-0.30	-0.12
pcp_s3_sqr	0.0037	0.0019	1.93	0.054	-6.9E-05	0.0074
pcp_s4	-0.021	0.037	-0.58	0.56	-0.094	0.051
pcp_s4_sqr	0.0024	0.0012	2	0.046	4.03E-05	0.0047
tmp_s1	-0.37	0.31	-1.2	0.23	-0.97	0.23
tmp_s1_sqr	0.0045	0.0031	1.43	0.15	-0.0017	0.011
tmp_s2	-0.28	0.47	-0.6	0.55	-1.21	0.64
tmp_s2_sqr	0.0043	0.0034	1.26	0.21	-0.0024	0.011
tmp_s3	0.26	0.62	0.42	0.67	-0.95	1.48
tmp_s3_sqr	-0.0034	0.0038	-0.88	0.38	-0.011	0.0041
tmp_s4	-0.52	0.33	-1.59	0.11	-1.16	0.12
tmp_s4_sqr	0.0029	0.0031	0.94	0.35	-0.0032	0.0089
dry_sp09	-0.19	0.13	-1.47	0.14	-0.44	0.063
wet_sp09	0.56	0.15	3.83	0	0.27	0.84
price_may_sep_diff_def	0.014	0.016	0.87	0.38	-0.018	0.045
farm	-0.0043	0.00049	-8.76	0	-0.0053	-0.0033
intercept	28.44	18.30	1.55	0.12	-7.47	64.36
adjusted R^2	0.47					
sample size	1204					

Table 6.4. Parameter estimates with respecting t-values, p-values, and 95% confidence intervals of the OLS model from the census subset

6.2 Marginal Effects

Marginal effect measures how the change in independent variable affects the change in expected value of dependent variable. The expected value of the dependent variable for the hurdle model is the product of the expected values of the first tier probit model and the second tier OLS model. The conventional method to calculate the marginal effect does not give a direct result on how either the predicted possibility of abandoning the cotton field, or the predicted acreage abandonment changes simultaneously when a dependent variable changes. Therefore, I have taken an alternative method to calculate and plot the marginal effect to capture the changes in expected values for both tiers of the hurdle model separately. The marginal effect for the hurdle model as a whole is presented in the next section.

In order to present the changes in probabilities or in predicted acreage abandonment regarding to the changes in independent variables, I have used the procedure described what follows:

1. Set non-irrigated Upland cotton as the baseline (set d_irri and d_pima at 0);
2. Calculate the mean, minimum, and maximum values of the first order independent variables;
3. Evenly sample 9 points between minimum and mean, and between mean and maximum values;
4. Square the minimum, mean, maximum, and 9 sampling point values to get the respective second order independent variables if necessary;
5. Calculate the predicted probabilities for probit model, or predicted acreage abandonment for OLS model, at the minimum, mean, maximum and sampling points of one independent variable, setting other independent variables constant at the mean values;

6. Plot the predicted probabilities for probit model, or predicted acreage abandonment for OLS model on the y-axis and the minimum, mean, maximum and sampling point values of the independent variable on the x-axis, then create the marginal effect curves for non-irrigated Upland cotton;
7. Switch d_irri to 1 and use it as a curve shifter, then plot the marginal effect curves for irrigated Upland cotton following step 5 on the same plain;
8. Switch d_pima to 1 and use it as an additional curve shifter, then plot the marginal effect curves for Pima cotton following the procedure described in step 5 on the same plain.

I have repeated the procedure 10 times and created the marginal effect curves of all 10 weather variables (pcp_s1, pcp_s2, pcp_s3, pcp_s4, tmp_s1, tmp_s2, tmp_s3, tmp_s4, dry_sp09, wet_sp09) for non-irrigated Upland cotton, irrigated Upland cotton, and Pima cotton. Here I use pcp_s2 and pcp_s3 marginal effect curves as examples. All the marginal effect curves for the probit model and OLS model are displayed in Appendix B and Appendix C respectively. It is worth noting that the tmp_s2 marginal effect curves for all cotton types and tmp_s4 marginal effect curve for non-irrigated Upland cotton exceed 100% for OLS model from census subset, so I manually set them to 100%.

The drawback of the method calculating the marginal effect is that it only calculates partial marginal effect. In this model, the dry and wet variables also measure the precipitation, and they are derived from precipitations. So when reporting the marginal effect of seasonal precipitation, the effect of dries and wet variables is ignored.

Season 2 is April through June. For most of the states May is the planting month, and all the states have done planting by the end of June, so season 2 is the planting season in the crop year. Precipitation in the planting season is welcomed by the farmers, so I expect the sign of the first order precipitation variable to be negative. However, excess moisture is not healthy for cotton seeds to sprout and farmers may have concerns about it, so I expect the sign of the second order precipitation variable to be positive. In general, with the increase of rainfall in season 2, the predicted possibility of abandoning cotton field and predicted acreage abandoned decreases at a decreasing rate. In addition, it is more likely to observe the precipitation variable reaching the optimal point on the predicted percentage acre abandoned curves, because rainfall is indeed harmful for growing conditions.

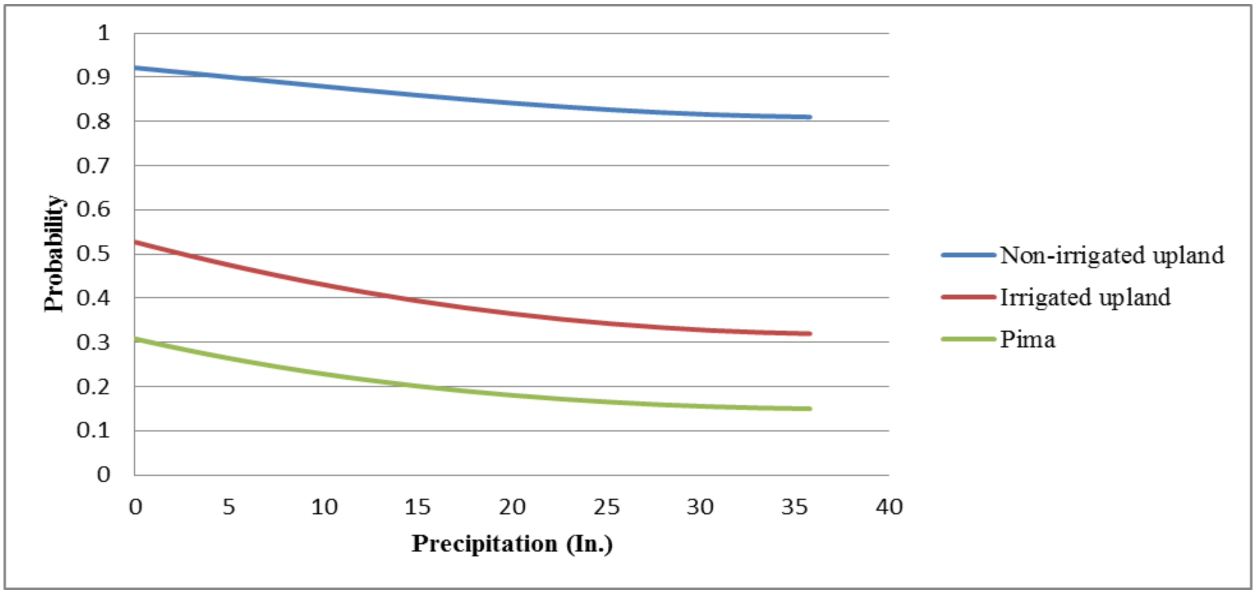


Figure 6.1.a. Marginal effect curves of season 2 precipitation of the probit model from full dataset

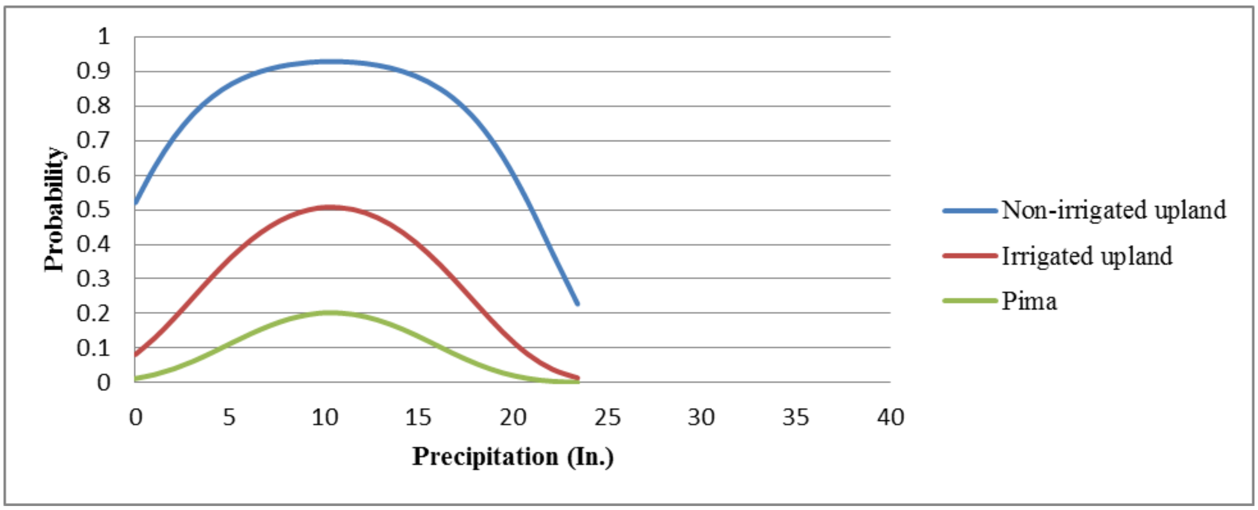


Figure 6.1.b. Marginal effect curves of season 2 precipitation of the probit model from census subset

Figure 6.1 shows marginal effect curves of pcp_s2 regarding to the possibilities to abandon cotton fields from full dataset and census subset. The marginal effect curves from the full dataset are exactly what I have expected. The sign of the first order precipitation variable is negative and the sign of the second order variable is positive. The abandon possibilities monotonically decrease at a decreasing speed. d_irri and d_pima act as curve shifter and drop the possibilities by a significant margin. The possibilities are high at the minimum point of precipitation. Probit model predicts a 92.14% chance for the farmers growing non-irrigated Upland cotton to abandon some crop. The curves do not reach the optimal point when measured at the maximum value of the season 2 precipitation. On the other hand the marginal effect curves from the census subset are not what I have expected. The expected abandonment possibilities increase until the maximum point then decrease, with the increase in precipitation received in season 2. Because of the limitation of the census subset, the range of precipitation in season 2 observed is a lot smaller than the range of season 2 precipitation in the full dataset.

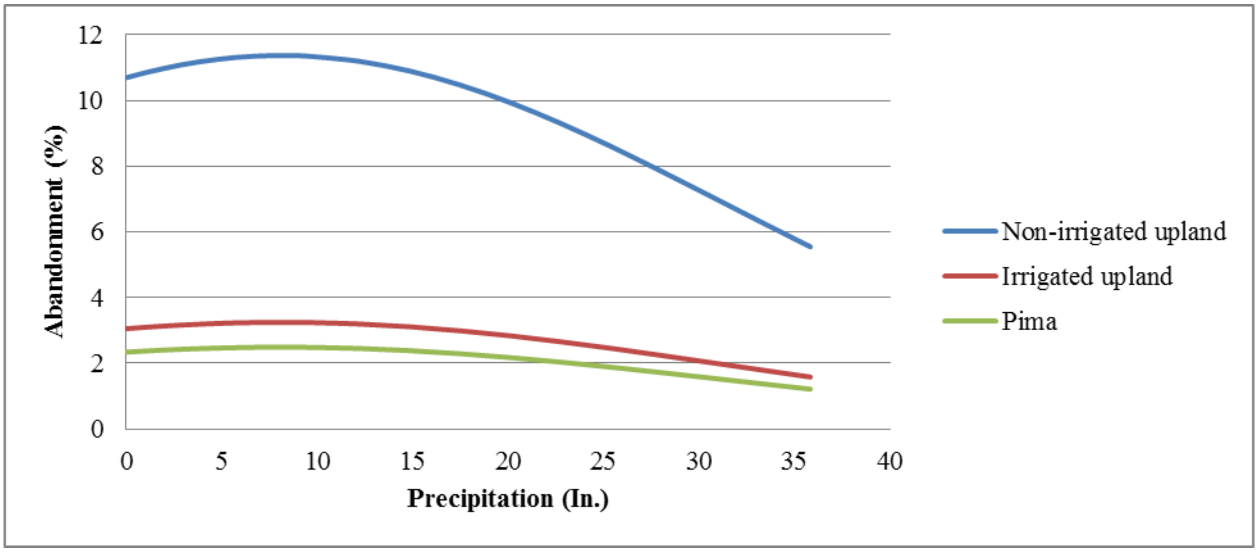


Figure 6.2.a. Marginal effect curves of season 2 precipitation of the OLS model from full dataset

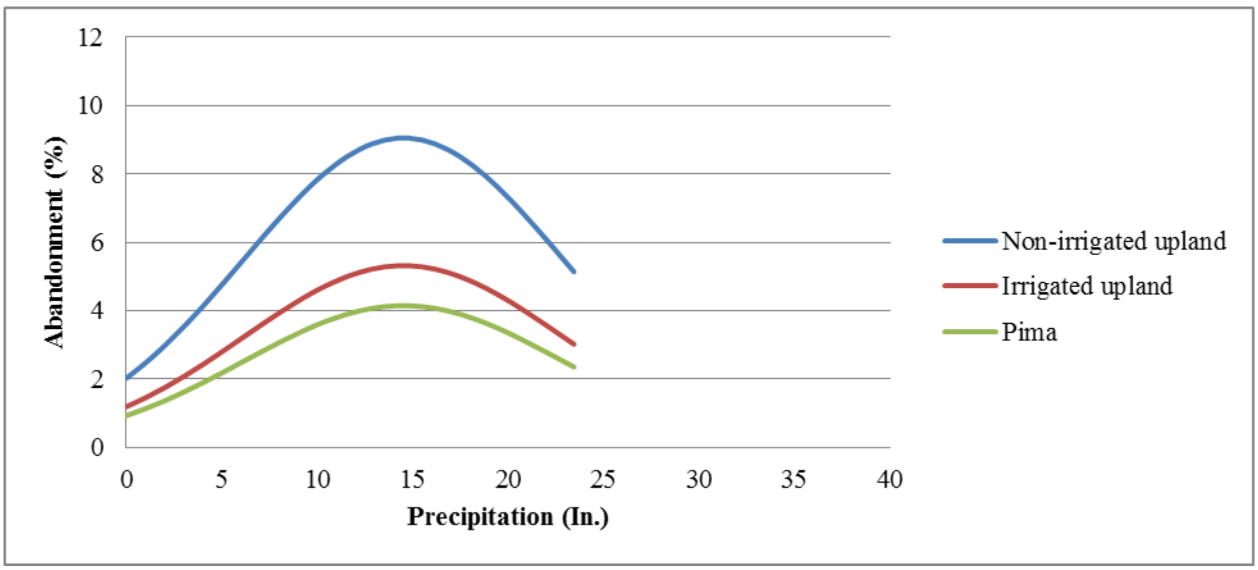


Figure 6.2.b. Marginal effect curves of season 2 precipitation of the OLS model from census subset

Figure 6.2 shows marginal effect curves of pcp_s2 regarding to the percentage acres abandoned from full dataset and census subset. The signs of the season 2 precipitation coefficients, both the first order and second order, are the opposite from my expectation, in the OLS model for both the full dataset and census subset. As I explained earlier, the predicted abandoned percentage should decrease at a decreasing speed as the precipitation in season 2 increases, and it is possible to observe the predicted possibility reaching the minimum point. Afterwards the possibility will increase at an increasing rate because excess moisture is harmful for cotton growth. However, according to the marginal curves of both the OLS models, the predicted possibilities are increasing at a decreasing rate then decrease at an increasing rate. The marginal effect curves of the full dataset behave a lot closer to my expectation. Even though they are increasing at a decreasing rate, but they reach the maximum point fairly fast then start decreasing at an increasing rate. The increasing rate is low, due to the fact that the parameter estimate of the second order precipitation is -0.00093.

If precipitation in season 2 gives confusing signals in predicting the probabilities of abandoning cotton field or percentage of the acres abandoned, precipitation in season 3, which is the growing season (from July through September), captures the predicted probabilities and percentages fairly well. In season 3, precipitation is again a welcoming factor. However, flood caused by excess moisture can lead to cotton field abandonment. So the expected signs of the parameter estimates of the first order precipitation and second order precipitation are negative and positive respectively, just like the precipitation variables in season 2. However, variable pcp_s3_sqr of the OLS model from the census subset turns out to be insignificant at 5% level with a p-value of 0.054. This might be caused by the smaller variance of the data in the census subset not capturing the full effect of precipitation in season 3.

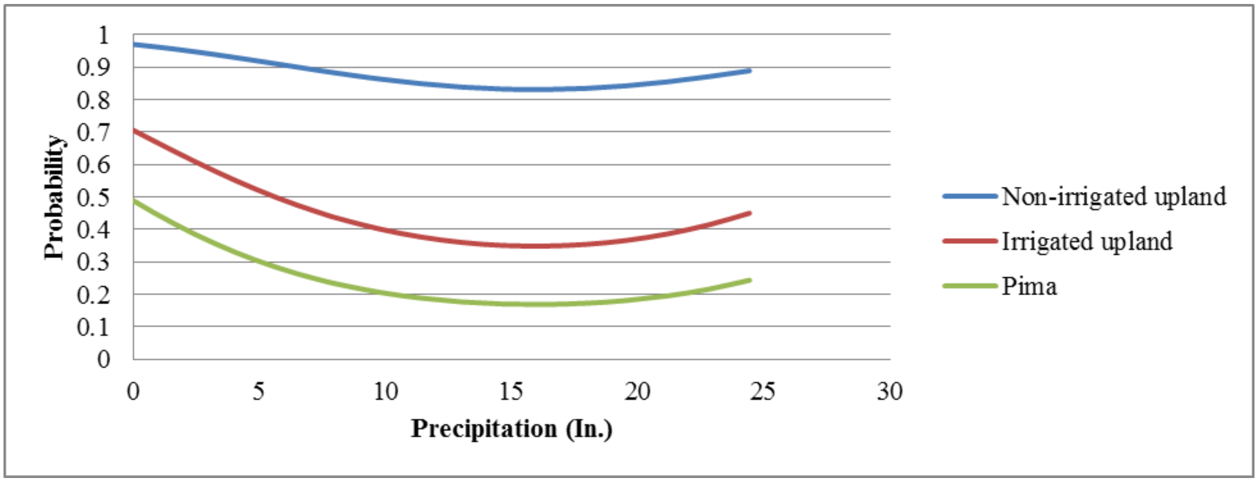


Figure 6.3.a. Marginal effect curves of season 3 precipitation of the probit model from full dataset

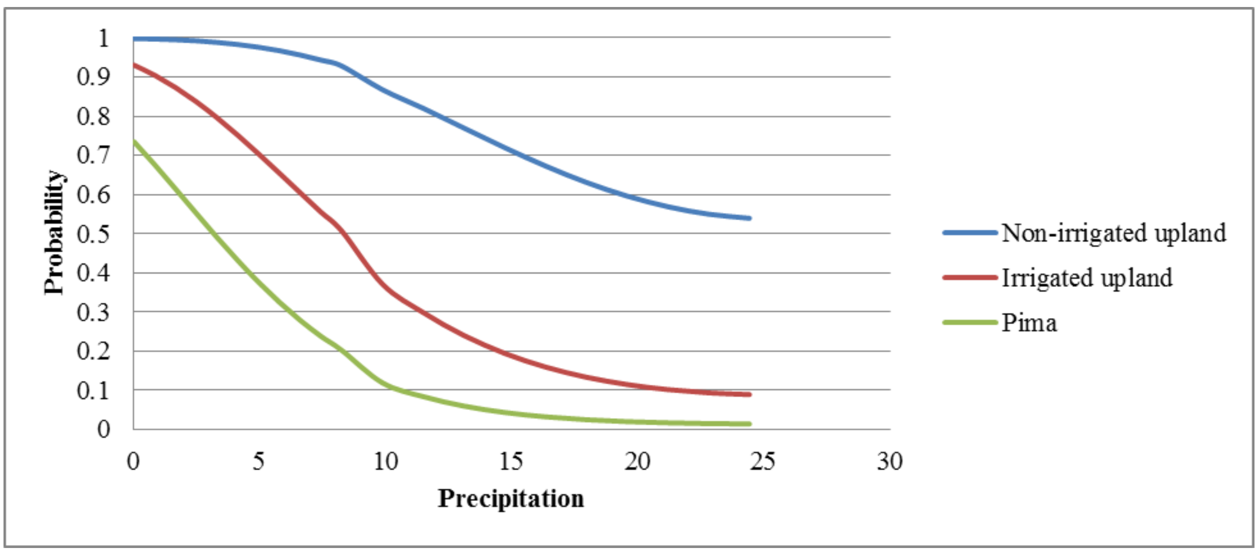


Figure 6.3.b. Marginal effect curves of season 3 precipitation of the probit model from census subset

Figure 6.3 shows marginal effect curves of pcp_s3 regarding to the possibilities to abandon cotton fields from full dataset and census subset. This time the marginal effect curves captured the increase in abandoning possibilities when precipitation in season3 increases. The marginal effect curves of the census subset are not as smooth as the other curves. It is because of the uneven sampling distance between minimum and mean values, and mean and maximum values. In addition, the curves do not capture the increase in possibilities due to excess moisture.

Figure 6.4 shows marginal effect curves of pcp_s3 regarding to the percentage acres abandoned from full dataset and census subset. Again, the marginal effect curves have captured the increase in predicted acreage abandonment. Interestingly, the percentage abandoned is predicted a lot higher in the census subset (35.10%) than in the full dataset (21.51%) for non-irrigated Upland cotton initially. Again, this might be caused by the smaller variation in the census subset.

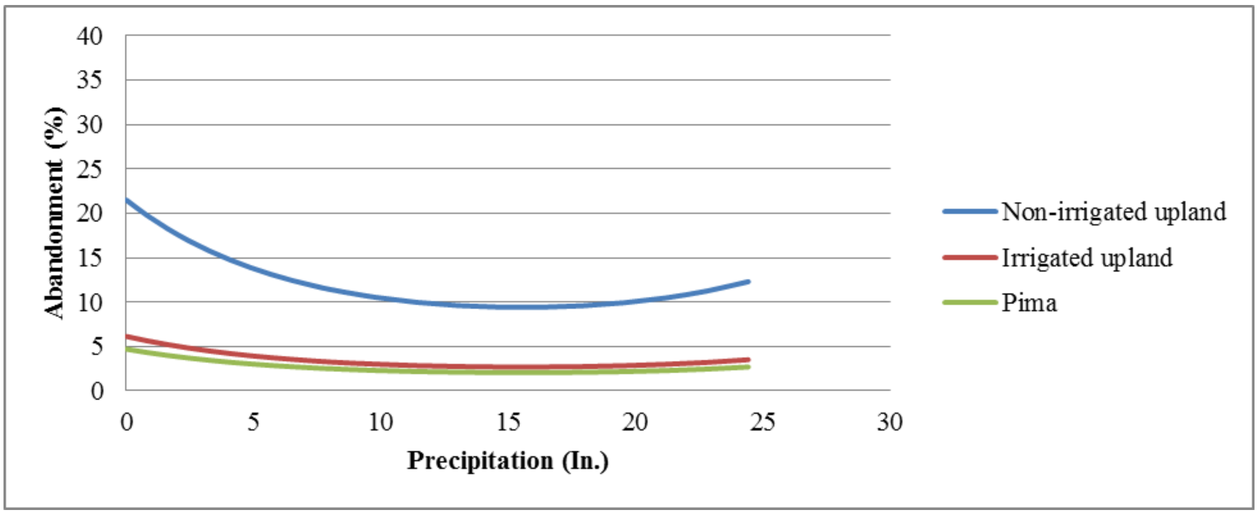


Figure 6.4.a. Marginal effect curves of season 3 precipitation of the OLS model from full dataset

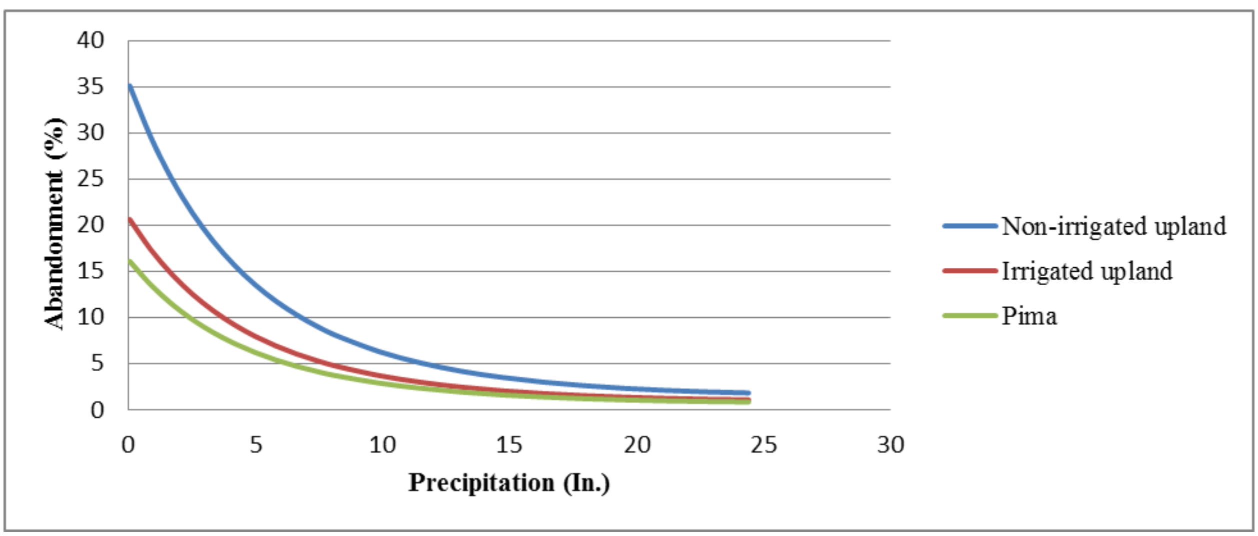


Figure 6.4.b. Marginal effect curves of season 3 precipitation of the OLS model from census subset

6.3 Expected Values of the Hurdle Model

I have used Equation 4.6 to calculate the expected values for full dataset.

In the last section, I have evaluated the marginal effects of all the weather variables regarding the probit model and OLS model respectively, but it would be interesting to evaluate the marginal effects of E(y|x) regarding all the weather variables. For this part, I have only used the results from the full dataset, due to the poor performance of the hurdle model from the census subset. Again, I have repeated the procedure 10 times and created the marginal effect curves of all 10 weather variables (pcp_s1, pcp_s2, pcp_s3, pcp_s4, tmp_s1, tmp_s2, tmp_s3, tmp_s4, dry_sp09, wet_sp09) for non-irrigated Upland cotton, irrigated Upland cotton, and Pima cotton, and I use pcp_s2 and pcp_s3 marginal effect curves as examples. All the marginal effect curves are listed in Appendix D.

The expected value marginal effect curves of pcp_s2 behave a lot better than taking the marginal effect curves for predicted possibilities and percentages separately. The curves are flat around the maximum point and decreases at an increasing rate afterwards. However, the marginal effect curves of pcp_s3 is classic. They decrease at a decreasing rate until the minimum point then start increasing at an increasing rate, due to the damage caused by excess moisture. The expected value for non-irrigated Upland cotton is more inelastic to rainfall in season 3, comparing to irrigated Upland and Pima cotton.

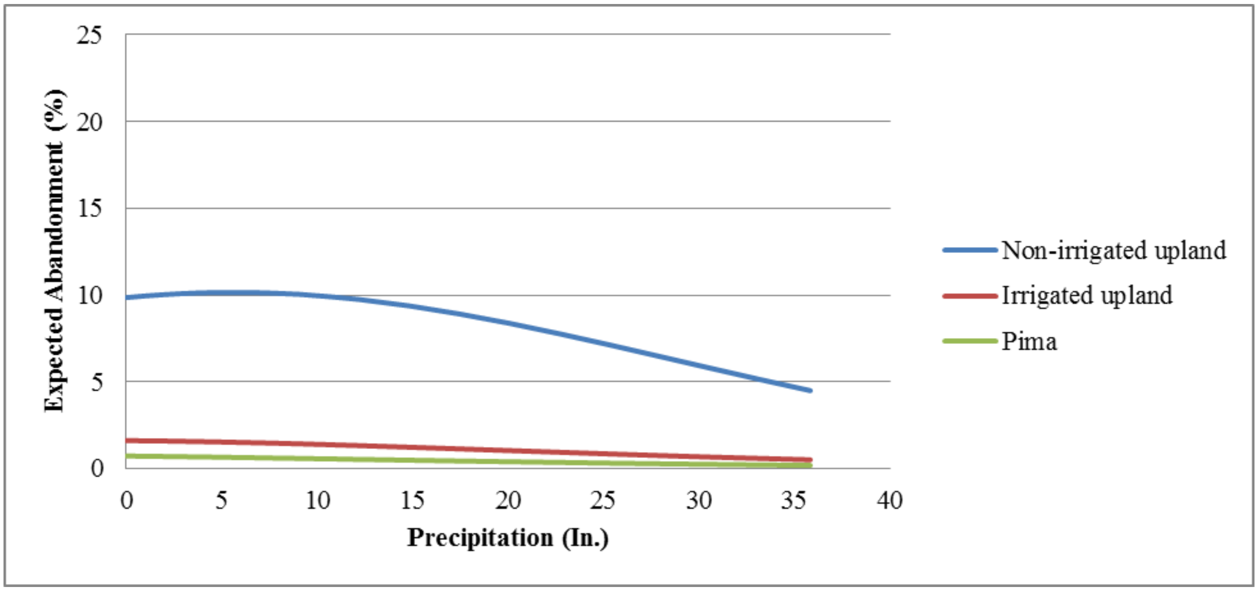


Figure 6.5.a. Marginal effect curves of season 2 precipitation for expected values

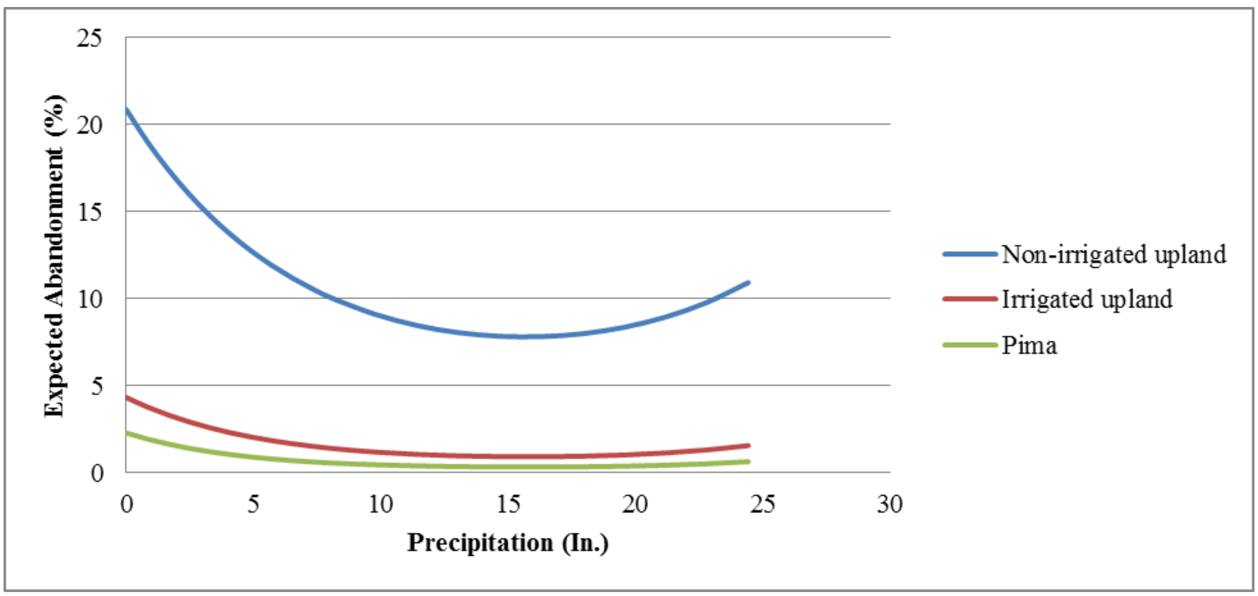


Figure 6.5.b. Marginal effect curves of season 3 precipitation for expected values

6.4 An Alternative Approach to Calculate Expected Values

In Equation 4.4, $\Phi(x\gamma)$ calculates the possibilities of abandoning any cotton fields, but it defeats the purpose of having the binary outcomes whether there will be cotton field abandoned or not. Therefore, I have proposed an alternative method to capture the possibility:

$$\hat{p} = \begin{cases} 0 & \text{if } \Phi(x\hat{\gamma}) < 0.5 \\ & . \\ 1 & \text{if } \Phi(x\hat{\gamma}) \geq 0.5 \end{cases} \tag{6.1}$$

Now 6.1 can be expressed as:

$$E(\hat{y}|x)_{alt} = \begin{cases} 0 & \text{if } \Phi(x\hat{\gamma}) < 0.5 \\ & . \\ \exp(x\hat{\beta} + \hat{\sigma}^2/2) & \text{if } \Phi(x\hat{\gamma}) \geq 0.5 \end{cases} \tag{6.2}$$

Using Equation 6.2, I have calculated the alternative expected percentage abandonment. The column correlations among actual percent abandoned (y), conventional percent abandoned ($\hat{y}$), and alternative percent abandoned ($\hat{y}_{alt}$) is recorded in Table 6.5.

	y	$E(\hat{y}\|x)$	$E(\hat{y}\|x)_{alt}$
y	1		
$E(\hat{y}\|x)$	0.5262	1	
$E(\hat{y}\|x)_{alt}$	0.5264	0.9958	1

Table 6.5. Correlations among actual percent abandoned (y), conventional percent abandoned ($\hat{y}$), and alternative percent abandoned ($\hat{y}_{alt}$)

From Table 6.5, it is clear that the alternative method to calculate the expected percent abandoned is more correlated to the true abandoned percentage, but only by a small proportion.

6.5 Special Case Study: Texas in 2011

Texas suffered an ongoing severe drought since 2011. According to a fact sheet released from the State of Texas, 2011 was "the driest year Texas has been since modern record keeping began in 1895". Nearly 67% of Texas was still in an "extreme" or "exceptional" drought (by U.S. Department of Agriculture standard) as of January 3, 2012 (Combs 2012). Table 6.6 shows the mean values of each weather variable across climate divisions in 2011 in Texas and means values across divisions over 22 years in Texas. Figure 6.6 shows the drought development from May, 2011 through September, 2011, covering the cotton production season in Texas. The images are from National Weather Service Weather Forecast Office (National Weather Service Weather Forecast Office, National Oceanic and Atmospheric Administration, 2012).

Variable	2011	22 Year Mean
pcp_s1	3.56	5.02
pcp_s2	3.40	8.45
pcp_s3	3.04	7.77
pcp_s4	6.55	6.03
tmp_s1	57.70	50.05
tmp_s2	78.87	71.50
tmp_s3	85.38	79.19
tmp_s4	57.70	50.05
dry_sp09	2.94	0.28
wet_sp09	0	0.48

Table 6.6. 2011 weather variable mean values and 22-year weather variable mean values in Texas

Comparing the 2011 state-level weather variable mean to the state-level weather variable mean over 22 years, 2011 is significantly drier, and not surprisingly, hotter. I would expect the percent abandoned to be high in the 22-year abandonment spectrum.

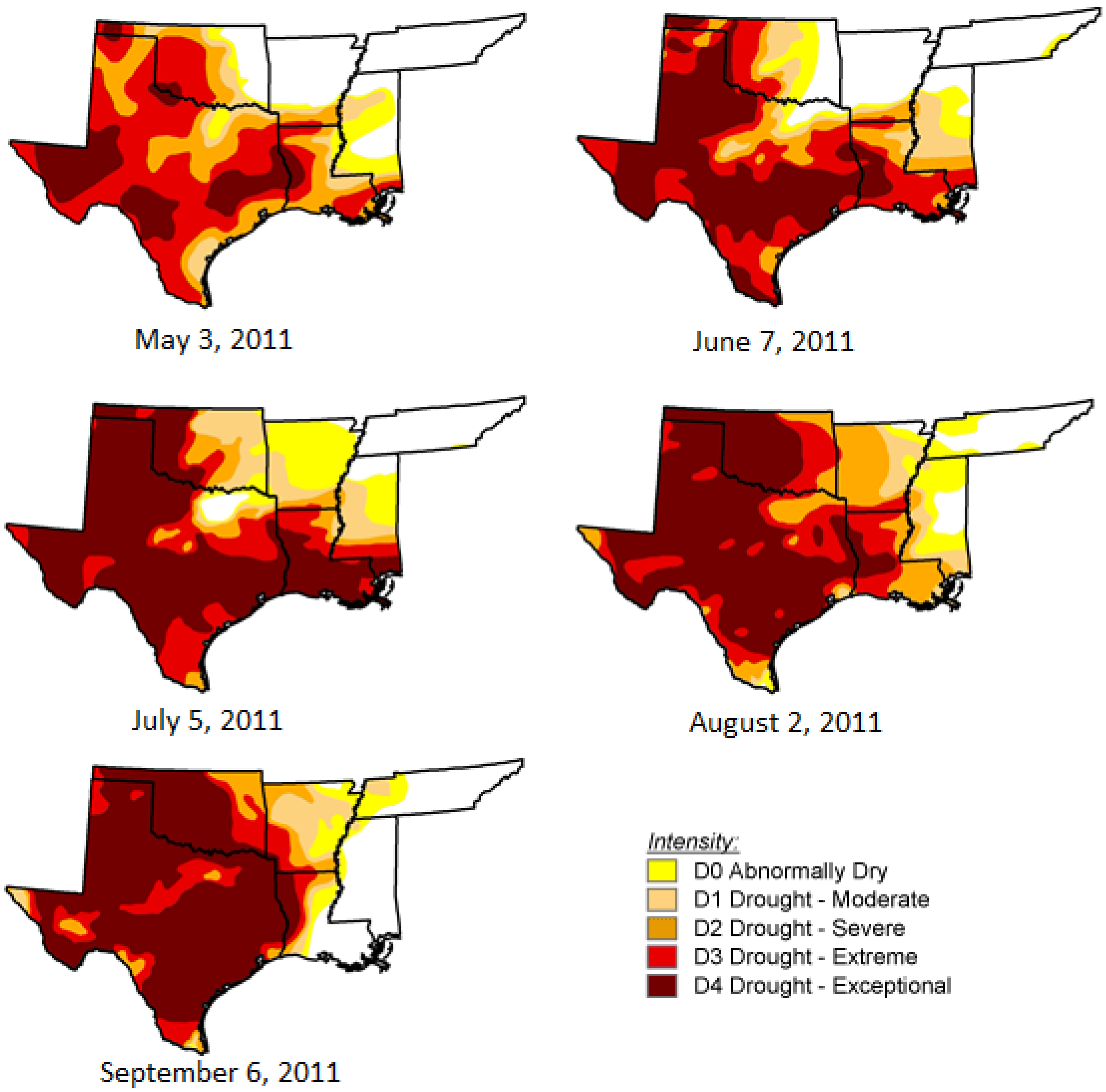


Figure 6.6. Drought intensity development from May through September in Southern Plain

The drought caused severe damages to agriculture production in Texas. In 2011, there were 1,819,200 acres of cotton planted in all eight states, with 569,200 acres planted alone in Texas. Of all the states, there were 214,800 acres abandoned, accounting for 11.81% of all the acres planted. In Texas alone, 188,000 acres were abandoned, covering 87.52% of all the acres abandoned in 2011. There were 479,900 acres of non-irrigated cotton fields planted and 183,700 acres abandoned; the percentage abandoned is 38.28%. Meanwhile, there were 89,300 acres of irrigated cotton planted and only 4,300 acres abandoned; the percentage abandoned is only 4.82%. However, as shown in Figure 6.7, the abandoned percentage signal is not very clear. The Pima cotton abandoned percentage in 2011 was relatively high, but it was insignificant comparing to the dramatic 67.39% in 1998. The non-irrigated Upland cotton abandoned percentage in 2011 increased significantly from 2010, but it was not the highest over 22 years. The irrigated Upland cotton abandoned percentage stayed low in the 22-year percent abandoned spectrum.

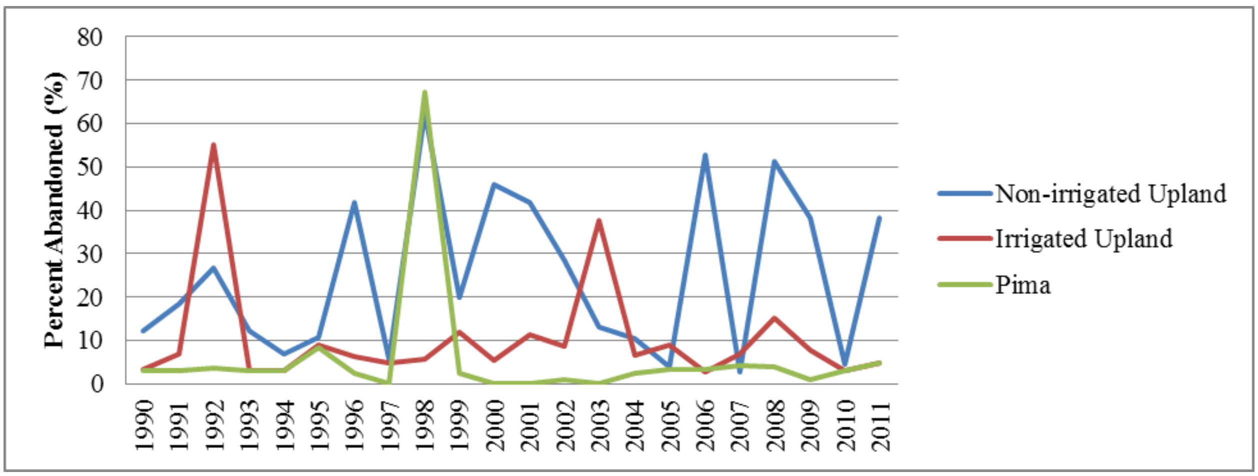


Figure 6.7. Percent abandoned curves sorted by commodity over 22 years in Texas

Of all the counties included in the full dataset, only 10 counties had grown non-irrigated Upland cotton in 2011. Six of them planted both non-irrigated and irrigated Upland cotton, the other 4 only planted non-irrigated Upland cotton, and none of them grew Pima cotton. Table 6.7 shows the predicted acre abandoned versus actual acre abandoned of non-irrigated Upland and irrigated Upland for those counties.

Non-irrigated	Actual Abandonment	Conventional Expected Value	Alternative Expected Value
BEE	17.07	41.32	47.58
BURLESON	37.50	41.32	47.58
CAMERON	21.20	89.44	94.38
FORT BEND	19.91	21.12	27.77
GUADALUPE	12.90	41.32	47.58
JIM WELLS	55.02	70.49	78.94
NUECES	12.88	41.32	47.58
TOM GREEN	96.17	90.21	97.12
TRAVIS	52.50	41.32	47.58
WHARTON	19.76	21.12	27.77

Table 6.7.a. Predicted acre abandoned versus actual acre abandoned for non-irrigated Upland cotton

Irrigated	Actual Abandonment	Conventional Expected Value	Alternative Expected Value
BEE	0	N/A	N/A
BURLESON	0	N/A	N/A
CAMERON	0.61	16.42	26.96
FORT BEND	40	2.07	0
TOM GREEN	12.41	15.20	27.75
WHARTON	0	N/A	N/A

Table 6.7.b. Predicted acre abandoned versus actual acre abandoned for irrigated Upland cotton

Because there was no crop abandoned in Bee County, Burleson County, and Wharton County, I could not get the prediction from OLS model for those counties, because of the restriction of the OLS model. In general, the model over predicts the percent abandoned, for both non-irrigated Upland cotton and irrigated Upland cotton.

Chapter 7

Conclusions and Future Work

This research attempts to capture the economic impact, and more importantly, weather impact on crop abandonment, considering irrigation practice. So far, most of the models evaluating the weather and/or economic impact on agriculture production used crop yield as dependent variable. Some models used acreage abandonment as dependent variable, but they ignored the reduction of crop abandonment due to irrigation. In this research, I have also used a binary variable for Pima cotton production as an additional curve shifter. So far, few researches have conducted empirical analysis of the effects of weather and economic variables on crop abandonment, considering irrigation.

In this research, I have used a 2-tier hurdle model estimating crop abandonment on a county-level dataset from 1990 through 2011. The first tier is to predict the possibility of abandonment, and the second tier is to predict the percentage of planted cotton fields abandoned for the counties observed positive abandonment. The model gives very promising results. Rainfall has a positive impact on reducing cotton field abandonment, but only to a certain point. Excess moisture can cause an increase in abandonment. Higher temperature creates a healthier condition for cotton growth, but abnormal high temperature, which is often correlated with reduction in rainfall, can increase abandonment as well. An increase in cotton price during the growing season gives farmers an incentive to keep the crop. The adoption of irrigation practice decreases the cotton abandonment. When measured at the mean weather conditions, the predicted probability to observe abandonment for non-irrigated upland cotton is 88.12% for full dataset and 92.82% for census subset. The probability reduces to 43.41% for full dataset and 50.72% for census subset if irrigation is adopted. The probability of abandonment for Pima cotton is 23.12% and 20.24% for full dataset and census subset respectively. In addition, the predicted abandonment for non-irrigated upland cotton is 11.37% for full dataset and 7.94% for census subset. The abandonment reduces to 3.24% for full dataset and 4.66% for census subset if irrigation is adopted. The abandonment for Pima cotton is 2.49% and 3.64% for full dataset and census subset respectively. Non-irrigated upland cotton is the least elastic to the changes in weather, and it has the highest range of both the predicted probability and abandonment. Irrigated upland cotton has a smaller range of the predicted probability and abandonment, when the Pima cotton has the smallest.

The model predicts fairly well for the full dataset, but not as well for the census subset. The census subset has some limitations. The census data are collected every 5 years, so the cotton planted/harvested record in the census subset only includes the observations in 1992, 1997, 2002, and 2007. The reduction in sample size significant reduces the variation in the dataset. However, the census data has the records of numbers of farms growing irrigated or non-irrigated cotton, which turns out to be an important variable. Interestingly, farm count variable has a positive coefficient in the probit model, and a negative coefficient in the OLS model. This suggests that it is more likely to observe abandonment in counties with more farms growing cotton, but the predicted abandonment decreases with the increase of farm count. Possible explanations might be higher diversification in counties having more farms growing cotton, or the possibility of those counties located in the marginal cotton production areas.

The results of the Texas special case study suggest the model's prediction is not as accurate at the extreme conditions. Texas suffered a severe yearlong drought in 2011, and the model tends to over-predict the abandonment given the abnormal weather behaviors.

Future work remains to be done in terms of adjusting the data generating method. For example, creating categorical variables or using other methods to control for spatial correlation. Meanwhile, other factors can be considered for the model, such as initial cost of planting the crop, the adoption of genetic modified seeds, or government programs, insurance policies and strategic behaviors. In addition, the model prediction of using county-level weather data, rather than climate division level weather data, can be an interesting topic. At last, it would be of great use to adopt this model in order to predict the damage of extreme climate events on crop field. What's more, full marginal effect analysis can be done on both the direct weather variables and the dry/wet indirect weather variables.

APPENDIX

Appendix A

State-Level Planted Acres of All Sates, 1990 – 2011

1. Planted Acres: Irrigated Upland summed from Full Dataset

Year	AR	AZ	CA	LA	MS	NM	OK	TX
1990	331,000	350,000	1,099,500	253,600	373,700	61,900	84,200	1,906,100
1991	416,700	360,000	979,500	268,900	358,000	63,400	81,000	2,244,000
1992	413,200	325,000	999,500	284,700	353,500	49,750	69,000	1,698,700
1993	414,000	316,000	1,049,500	259,600	413,000	51,100	68,000	1,959,800
1994	467,000	312,300	1,098,400	259,050	359,300	49,150	67,900	2,025,000
1995	567,000	364,200	1,168,100	350,000	389,700	56,450	74,600	2,332,800
1996	493,000	315,000	998,200	270,000	315,400	54,800	69,100	2,191,800
1997	619,000	325,000	878,800	155,000	294,900	59,400	74,300	2,032,600
1998	590,500	249,000	649,300	155,000	304,700	60,100	63,600	2,054,900
1999	511,000	269,300	609,500	155,000	401,400	79,000	77,500	2,236,900
2000	547,000	280,000	774,900	150,000	492,100	68,500	88,400	2,436,500
2001	528,000	295,000	629,700	289,900	716,800	68,000	69,700	2,213,000
2002	486,000	215,000	479,900	180,000	493,600	54,000	63,800	2,139,000
2003	519,700	214,400	549,700	176,000	446,800	53,000	74,300	2,187,900
2004	528,200	239,300	559,800	157,900	451,300	68,000	84,600	2,222,400
2005	621,100	227,700	429,800	228,600	518,800	56,000	92,300	2,212,300
2006	832,400	188,000	285,000	245,000	597,800	50,000	94,500	2,179,600
2007	402,600	164,000	193,500	71,100	296,700	34,800	55,100	1,744,700
2008	250,600	128,900	116,600	38,400	109,400	30,800	48,000	1,415,200
2009	271,700	137,800	53,500	25,000	90,600	24,000	-	1,623,200
2010	219,600	183,700	108,000	-	136,900	41,200	-	1,952,600
2011	158,700	233,800	172,200	-	193,100	44,000	-	76,900

2. Planted Acres: Irrigated Upland from QuickStats, USDA

Year	AR	AZ	CA	LA	MS	NM	OK	TX
1990	331,000	350,000	1,100,000	254,000	373,700	61,900	85,000	1,915,000
1991	416,700	360,000	980,000	269,200	430,000	63,400	82,000	2,255,000
1992	413,200	325,000	1,000,000	285,000	385,000	49,750	70,000	1,710,000
1993	415,000	316,000	1,050,000	260,000	440,000	51,200	69,500	1,970,000
1994	470,000	313,000	1,100,000	260,000	370,000	49,150	70,000	2,035,000
1995	570,000	365,000	1,170,000	350,000	400,000	56,600	76,000	2,350,000
1996	499,000	315,000	1,000,000	270,000	320,000	55,200	70,000	2,210,000
1997	626,500	325,000	880,000	155,000	309,500	59,900	75,000	2,050,000
1998	595,500	250,000	650,000	155,000	324,600	60,100	65,000	2,070,000
1999	639,000	270,000	610,000	155,000	431,000	79,000	80,000	2,250,000
2000	658,000	280,000	775,000	150,000	507,600	68,500	90,000	2,452,000
2001	689,000	295,000	630,000	290,000	731,000	68,000	72,000	2,238,000
2002	693,000	215,000	480,000	180,000	506,000	54,000	70,000	2,150,000
2003	732,000	215,000	550,000	176,000	457,000	53,000	82,000	2,250,000
2004	722,000	240,000	560,000	160,000	497,500	68,000	94,000	2,265,000
2005	864,000	230,000	430,000	230,000	567,600	56,000	100,000	2,300,000
2006	972,000	190,000	285,000	245,000	613,000	50,000	105,000	2,365,000
2007	723,000	170,000	195,000	101,000	320,000	43,000	80,000	1,840,000
2008	521,000	135,000	120,000	82,200	149,000	-	85,000	1,712,000
2009	459,000	145,000	71,000	54,000	123,700	29,000	90,000	1,756,000
2010	476,500	195,000	124,000	-	171,500	-	-	2,051,000
2011	606,000	250,000	182,000	-	277,000	-	-	2,476,000

3. Planted Acres: Non-irrigated Upland summed from Full Dataset

Year	AR	AZ	CA	LA	MS	NM	OK	TX
1990	437,900	-	-	553,000	854,400	7,100	293,000	3,573,500
1991	580,800	-	-	600,000	802,030	5,600	307,900	4,032,400
1992	583,800	-	-	599,300	958,600	5,250	195,300	3,776,900
1993	573,000	-	-	624,000	875,200	2,300	155,000	3,567,600
1994	502,000	-	-	635,050	889,900	5,850	141,500	3,401,100
1995	595,000	-	-	733,000	1,036,300	4,400	137,100	4,034,800
1996	491,000	-	-	617,600	550,200	3,800	155,650	3,476,100
1997	345,500	-	-	499,700	360,500	10,100	54,300	3,435,700
1998	316,500	-	-	378,900	317,900	6,200	41,900	3,567,400
1999	309,000	-	-	455,100	438,700	5,000	76,500	3,885,900
2000	282,000	-	-	555,100	521,700	3,500	108,700	3,417,600
2001	358,000	-	-	573,000	640,300	-	129,700	3,458,900
2002	250,500	-	-	337,900	434,200	-	80,200	3,150,900
2003	238,300	-	-	346,100	476,400	-	61,200	2,999,600
2004	181,900	-	-	325,900	374,800	-	87,300	3,108,700
2005	178,800	-	-	366,100	328,300	-	117,200	3,152,200
2006	192,900	-	-	372,800	519,900	-	167,500	3,324,700
2007	123,600	-	-	180,200	172,300	-	16,400	2,652,500
2008	83,100	-	-	86,800	116,700	-	8,000	2,163,700
2009	54,300	-	-	102,600	114,600	-	-	2,493,100
2010	48,900	-	-	-	147,700	-	-	2,703,500
2011	52,200	-	-	-	115,600	-	-	479,900

4. Planted Acres: Non-irrigated Upland from QuickStats, USDA

Year	AR	AZ	CA	LA	MS	NM	OK	TX
1990	439,000	-	-	556,000	856,300	7,100	295,000	3,585,000
1991	583,300	-	-	605,800	815,000	5,600	358,000	4,045,000
1992	586,800	-	-	605,000	965,000	5,250	300,000	3,790,000
1993	575,000	-	-	630,000	890,000	2,300	300,500	3,580,000
1994	510,000	-	-	640,000	910,000	5,850	290,000	3,415,000
1995	600,000	-	-	735,000	1,060,000	4,400	304,000	4,050,000
1996	501,000	-	-	620,000	800,000	3,800	220,000	3,490,000
1997	353,500	-	-	500,000	675,500	10,100	125,000	3,450,000
1998	324,500	-	-	380,000	625,400	6,200	95,000	3,580,000
1999	331,000	-	-	460,000	769,000	5,000	160,000	3,900,000
2000	302,000	-	-	560,000	792,400	3,500	190,000	3,948,000
2001	391,000	-	-	580,000	889,000	-	198,000	3,762,000
2002	267,000	-	-	340,000	664,000	-	130,000	3,450,000
2003	248,000	-	-	349,000	653,000	-	98,000	3,350,000
2004	188,000	-	-	340,000	612,500	-	126,000	3,585,000
2005	186,000	-	-	380,000	642,400	-	155,000	3,650,000
2006	198,000	-	-	390,000	617,000	-	215,000	4,035,000
2007	137,000	-	-	234,000	340,000	-	95,000	3,060,000
2008	99,000	-	-	217,800	216,000	-	85,000	3,288,000
2009	61,000	-	-	176,000	181,300	2,100	115,000	3,244,000
2010	68,500	-	-	-	248,500	-	-	3,499,000
2011	74,000	-	-	-	353,000	-	-	5,074,000

5. Planted Acres: Pima summed from Full Dataset

Year	AR	AZ	CA	LA	MS	NM	OK	TX
1990	-	124,700	25,600	-	-	19,300	-	56,300
1991	-	105,700	63,900	-	-	19,600	-	55,900
1992	-	102,700	109,900	-	-	13,000	-	35,000
1993	-	57,000	91,000	-	-	10,900	-	28,700
1994	-	47,400	80,700	-	-	10,800	-	26,300
1995	-	47,700	114,700	-	-	14,800	-	33,900
1996	-	42,000	164,700	-	-	13,450	-	36,300
1997	-	21,500	184,600	-	-	10,700	-	31,000
1998	-	14,100	199,900	-	-	7,300	-	97,200
1999	-	8,400	239,900	-	-	-	-	31,500
2000	-	4,200	144,700	-	-	3,200	-	15,000
2001	-	5,800	239,900	-	-	5,200	-	16,500
2002	-	6,300	209,800	-	-	6,600	-	18,500
2003	-	-	150,000	-	-	5,100	-	18,800
2004	-	1,200	214,400	-	-	7,700	-	20,500
2005	-	2,800	229,100	-	-	9,500	-	23,900
2006	-	4,800	274,600	-	-	11,150	-	30,100
2007	-	-	259,900	-	-	3,900	-	24,200
2008	-	-	155,000	-	-	2,600	-	15,600
2009	-	-	110,700	-	-	-	-	18,000
2010	-	-	180,400	-	-	-	-	17,000
2011	-	6,400	274,000	-	-	-	-	12,400

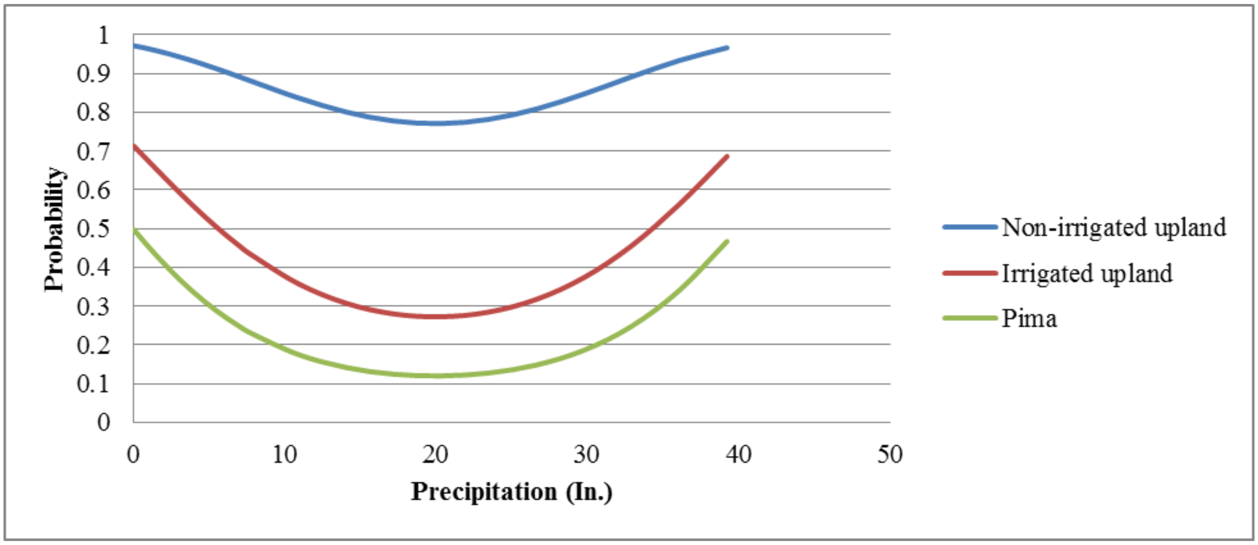
6. Planted Acres: Pima from QuickStats, USDA

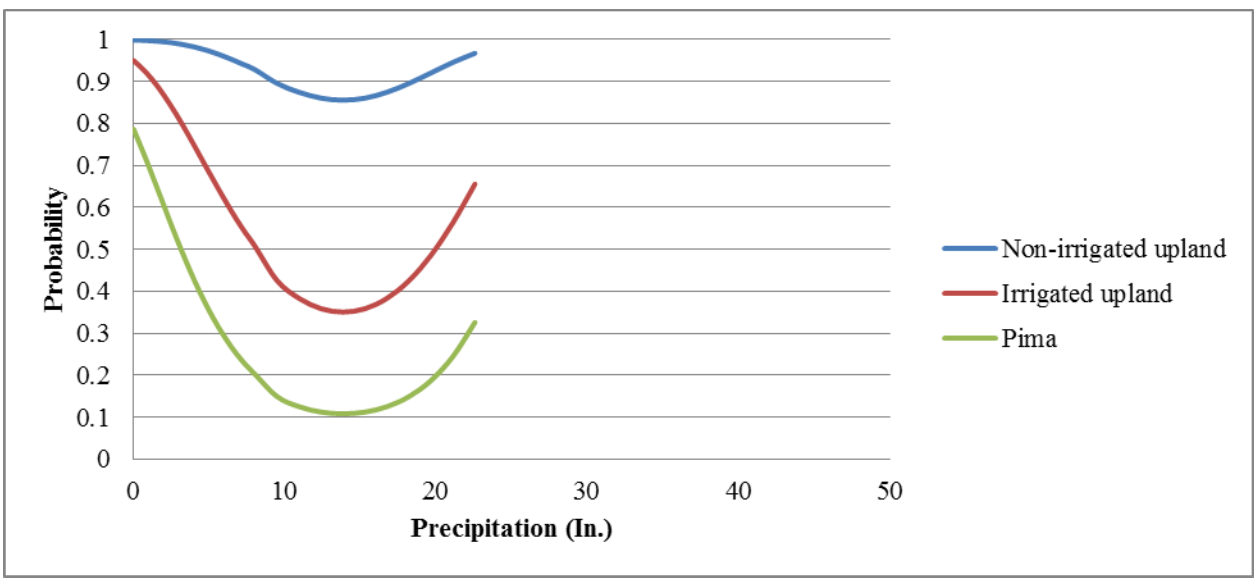
Year	AR	AZ	CA	LA	MS	NM	OK	TX
1990	-	125,000	25,700	-	1,300	19,300	-	60,000
1991	-	106,000	64,000	-	800	19,600	-	60,000
1992	-	103,000	110,000	-	400	13,000	-	37,000
1993	-	57,000	91,000	-	-	11,000	-	31,000
1994	-	48,000	81,000	-	-	11,000	-	28,500
1995	-	48,600	115,000	-	-	15,000	-	36,000
1996	-	42,000	165,000	-	-	14,000	-	37,000
1997	-	22,000	185,000	-	-	11,000	-	32,000
1998	-	15,900	200,000	-	-	7,300	-	105,000
1999	-	9,000	240,000	-	-	7,500	-	33,000
2000	-	5,000	145,000	-	-	4,200	-	16,000
2001	-	7,800	240,000	-	-	5,200	-	17,000
2002	-	8,300	210,000	-	-	7,100	-	18,500
2003	-	2,500	150,000	-	-	6,100	-	20,000
2004	-	3,000	215,000	-	-	10,600	-	21,000
2005	-	4,100	230,000	-	-	11,500	-	24,800
2006	-	7,000	275,000	-	-	13,000	-	31,000
2007	-	2,500	260,000	-	-	4,700	-	25,000
2008	-	800	155,000	-	-	2,600	-	15,600
2009	-	1,600	119,000	-	-	2,800	-	18,000
2010	-	2,500	182,000	-	-	2,700	-	17,000
2011	-	10,000	274,000	-	-	3,400	-	20,000

Appendix B

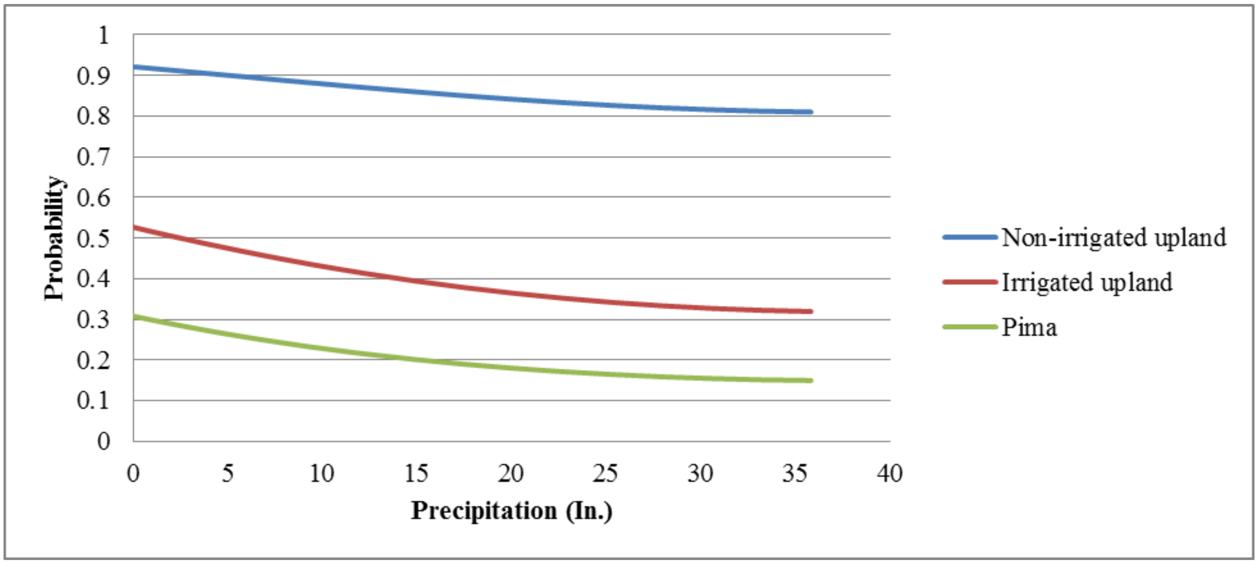
Marginal Effect Curves for Probit Model from Full Dataset and Census Subset

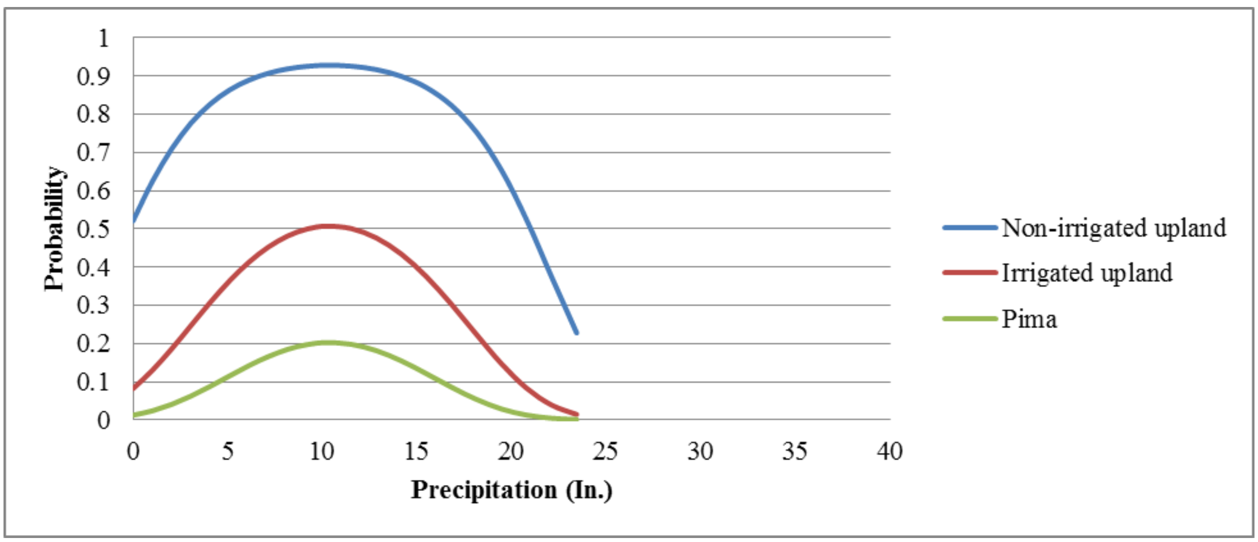
1. Marginal effect curves of season 1 precipitation for the probit model from full dataset (top) and census subset (bottom)



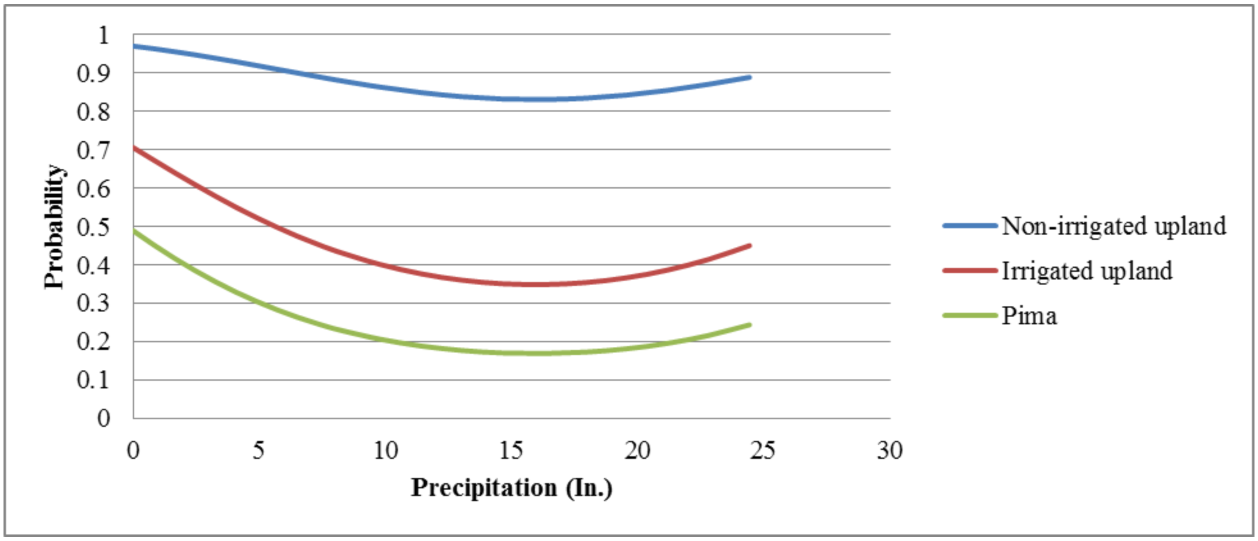


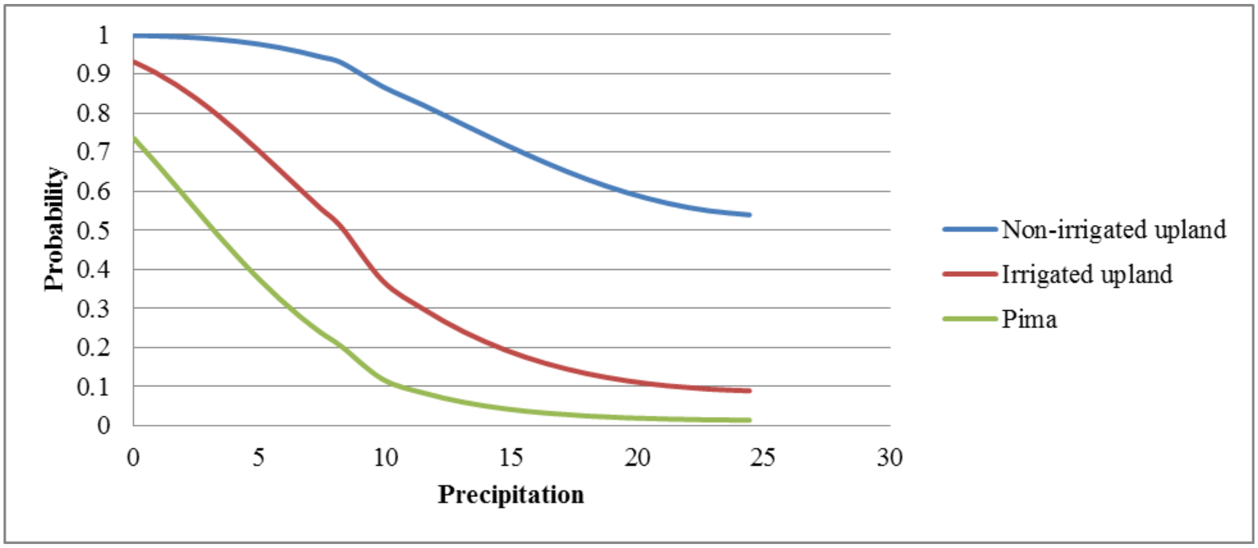
2. Marginal effect curves of season 2 precipitation for the probit model from full dataset (top) and census subset (bottom)



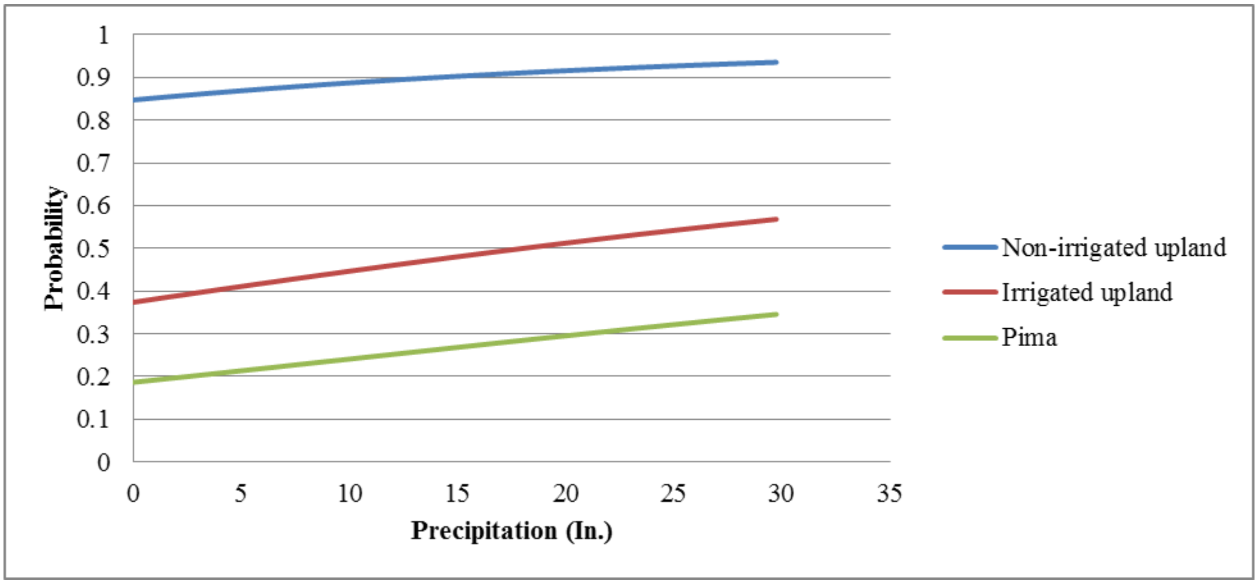


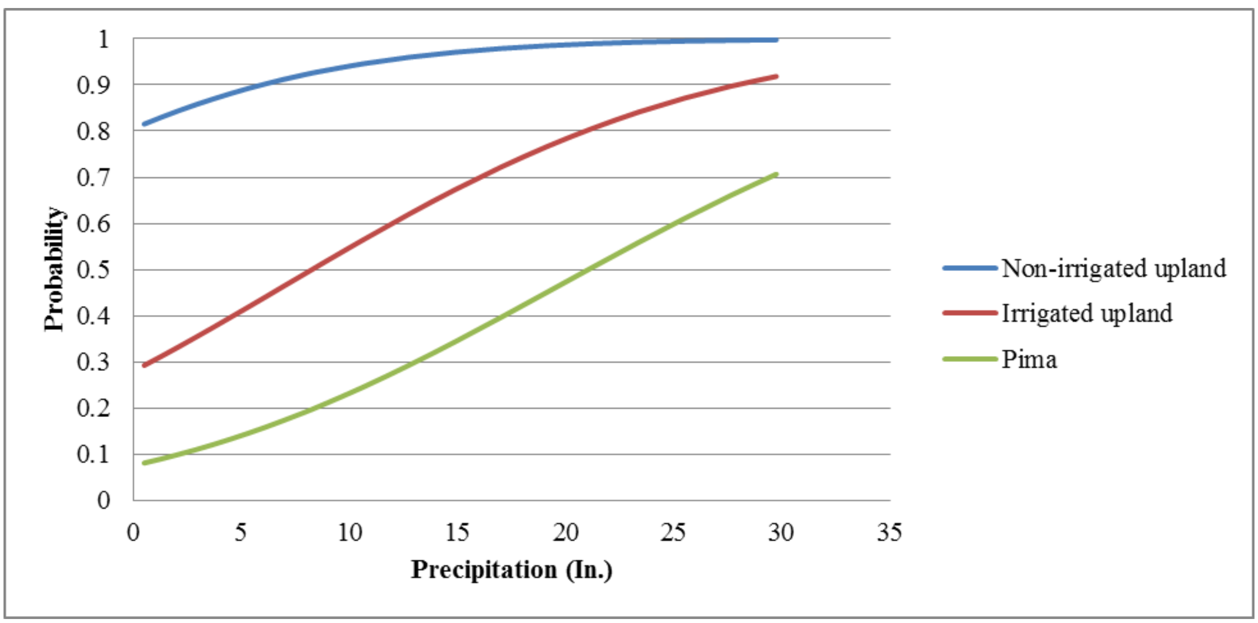
3. Marginal effect curves of season 3 precipitation for the probit model from full dataset (top) and census subset (bottom)



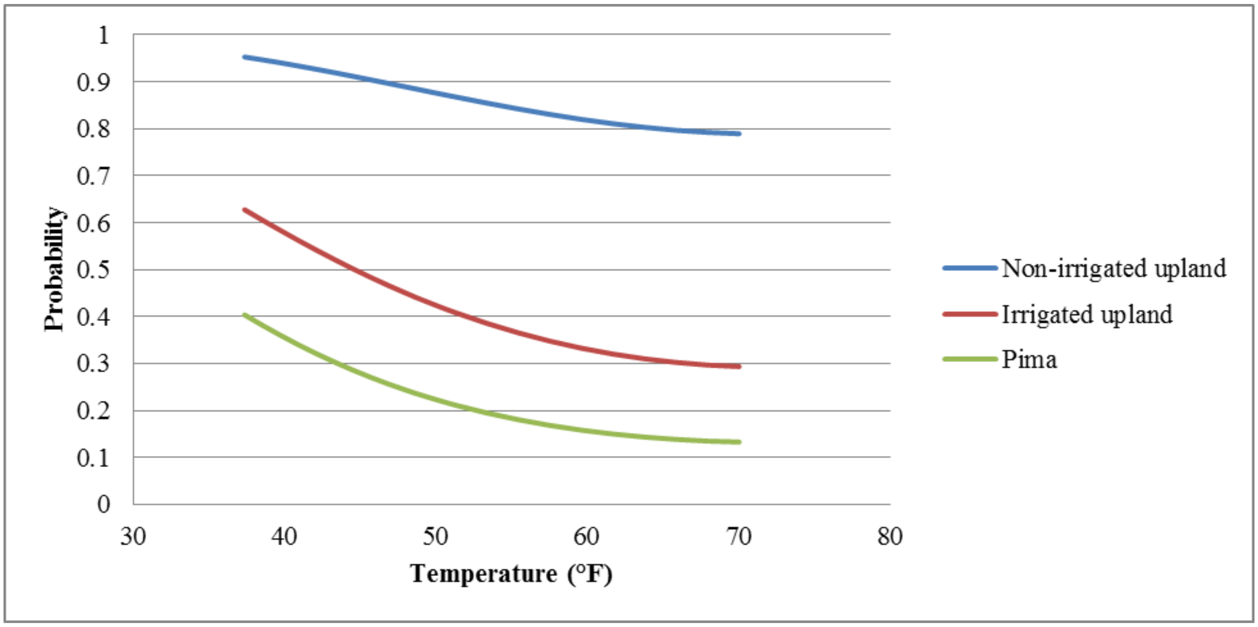


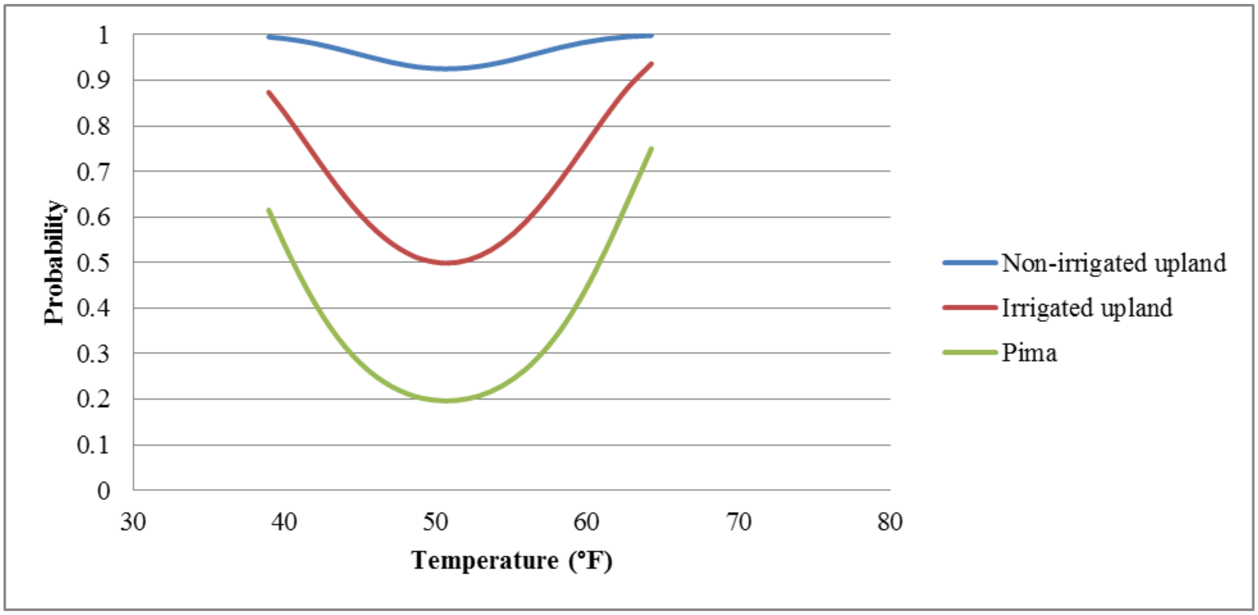
4. Marginal effect curves of season 4 precipitation for the probit model from full dataset (top) and census subset (bottom)



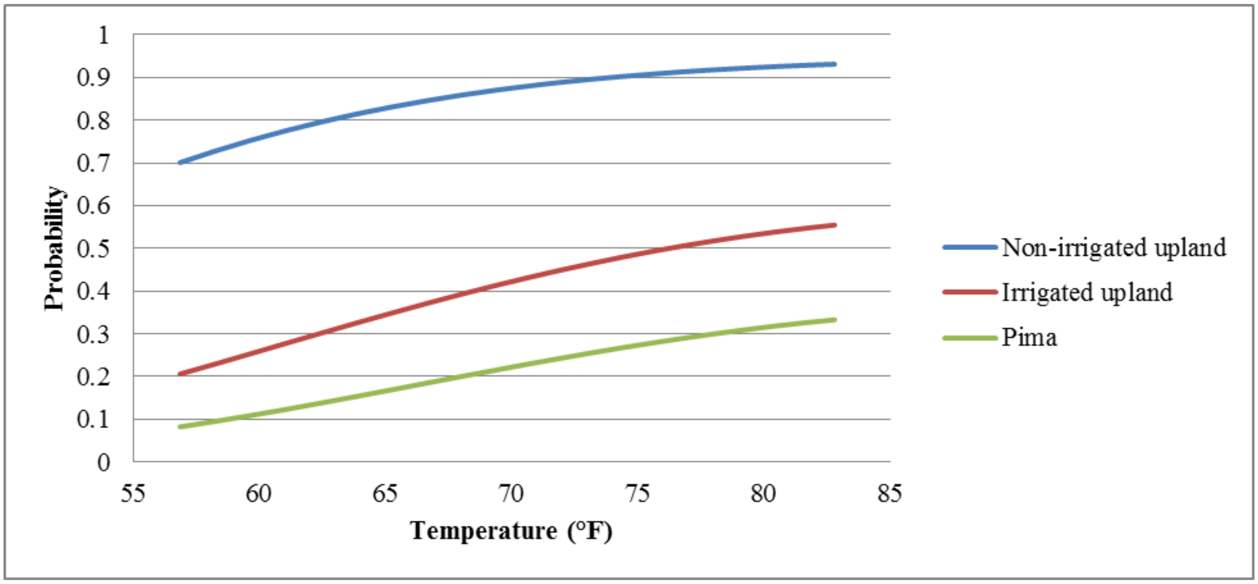


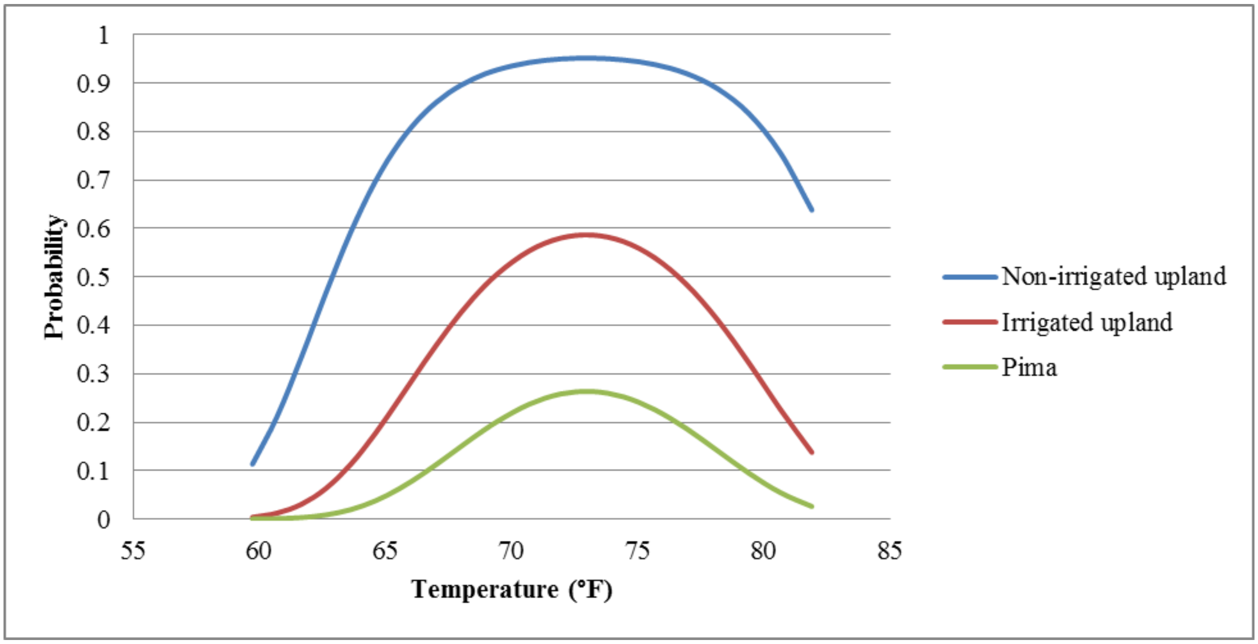
5. Marginal effect curves of season 1 temperature for the probit model from full dataset (top) and census subset (bottom)



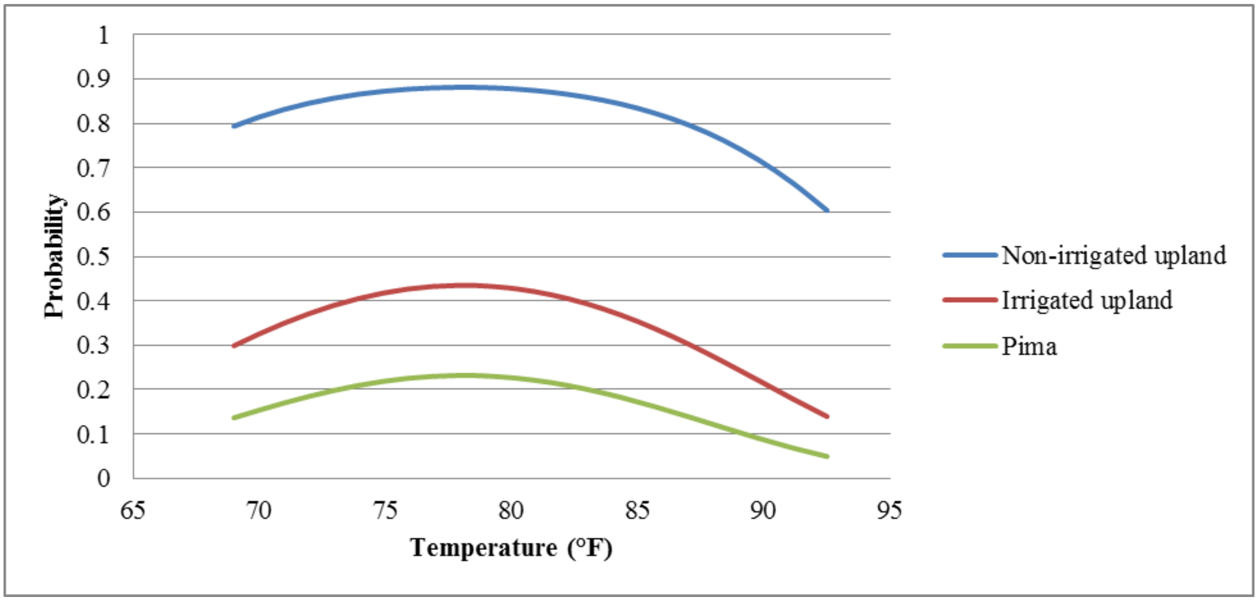


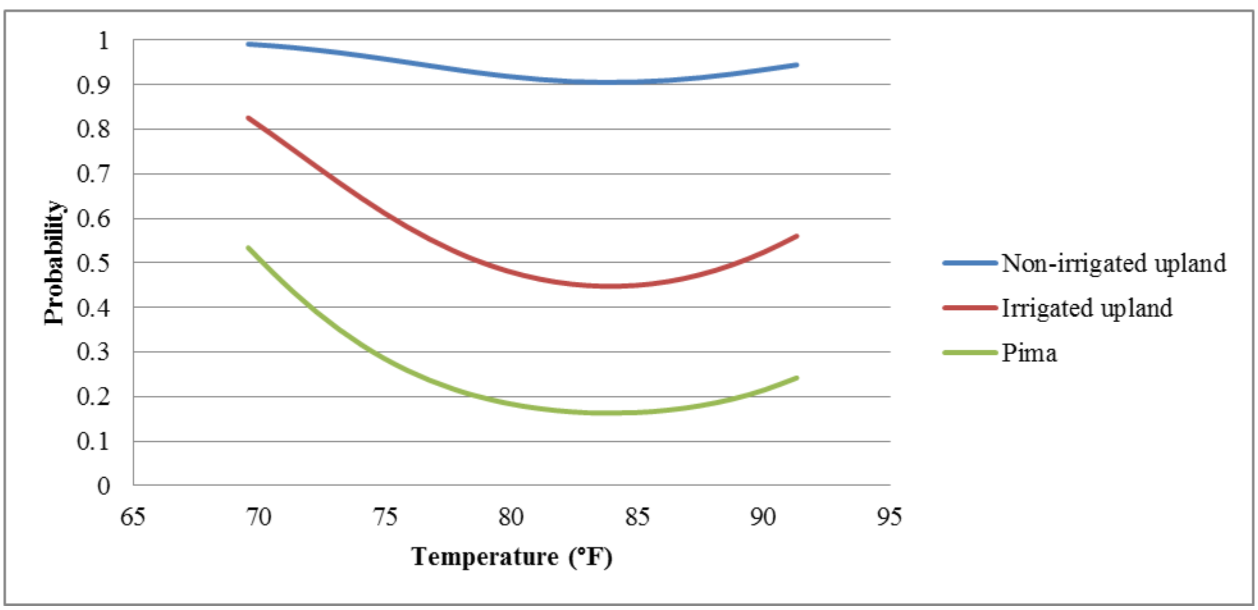
6. Marginal effect curves of season 2 temperature for the probit model from full dataset (top) and census subset (bottom)



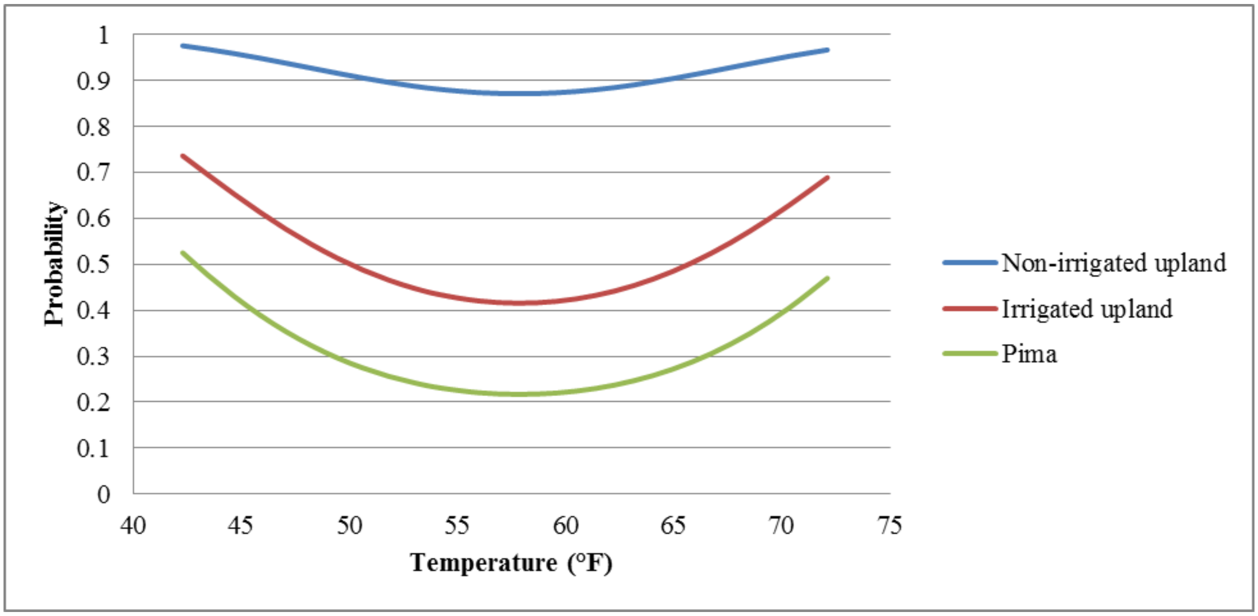


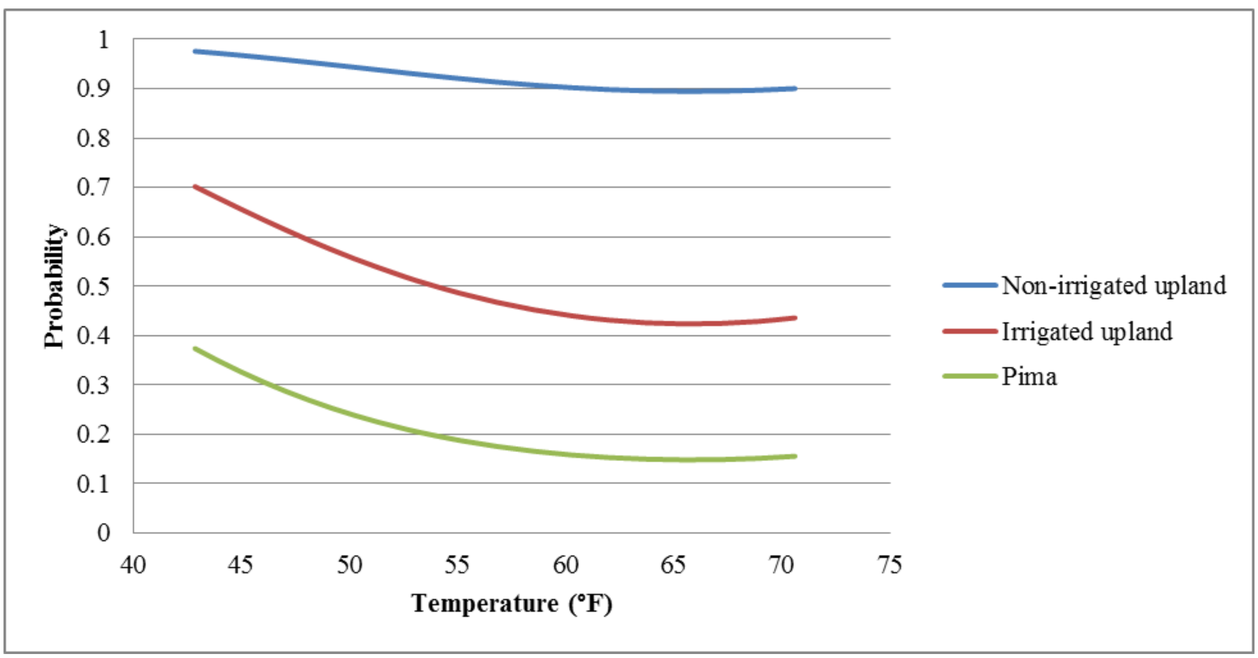
7. Marginal effect curves of season 3 temperature for the probit model from full dataset (top) and census subset (bottom)



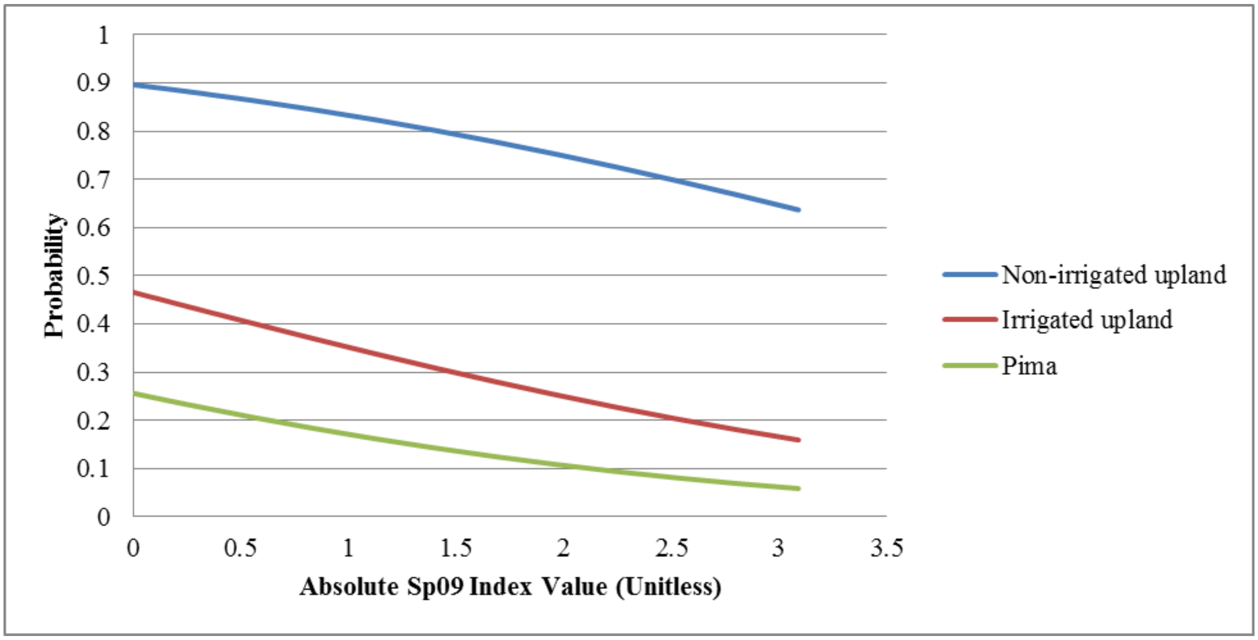


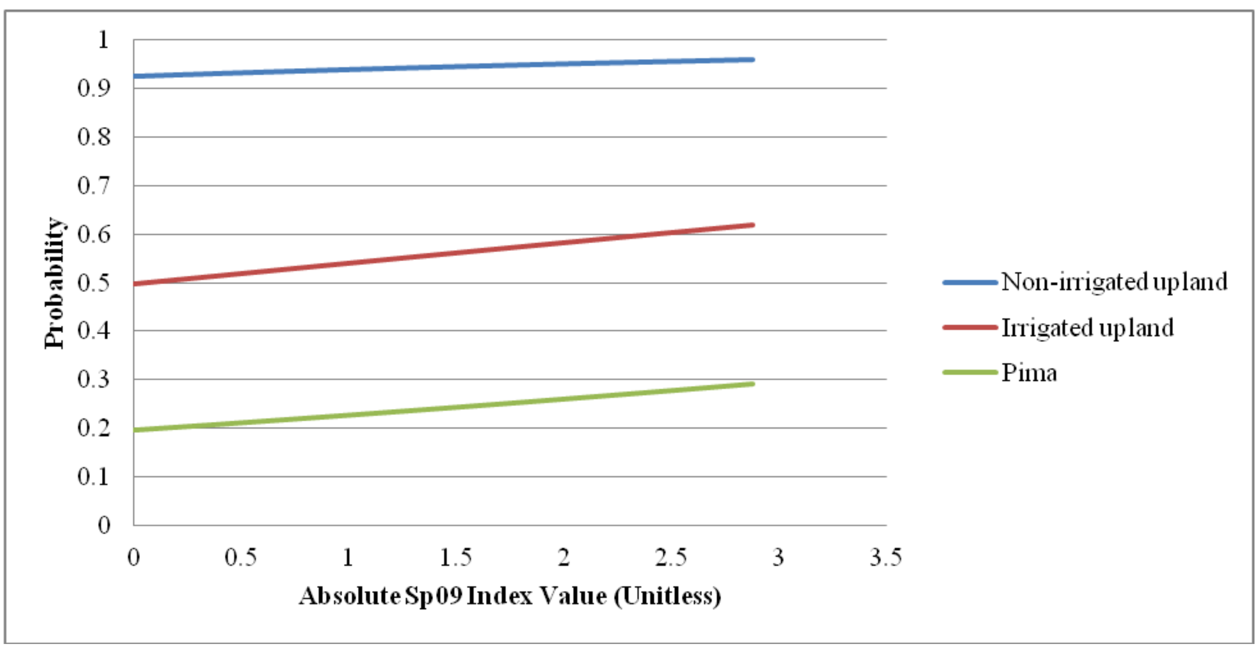
8. Marginal effect curves of season 4 temperature for the probit model from full dataset (top) and census subset (bottom)



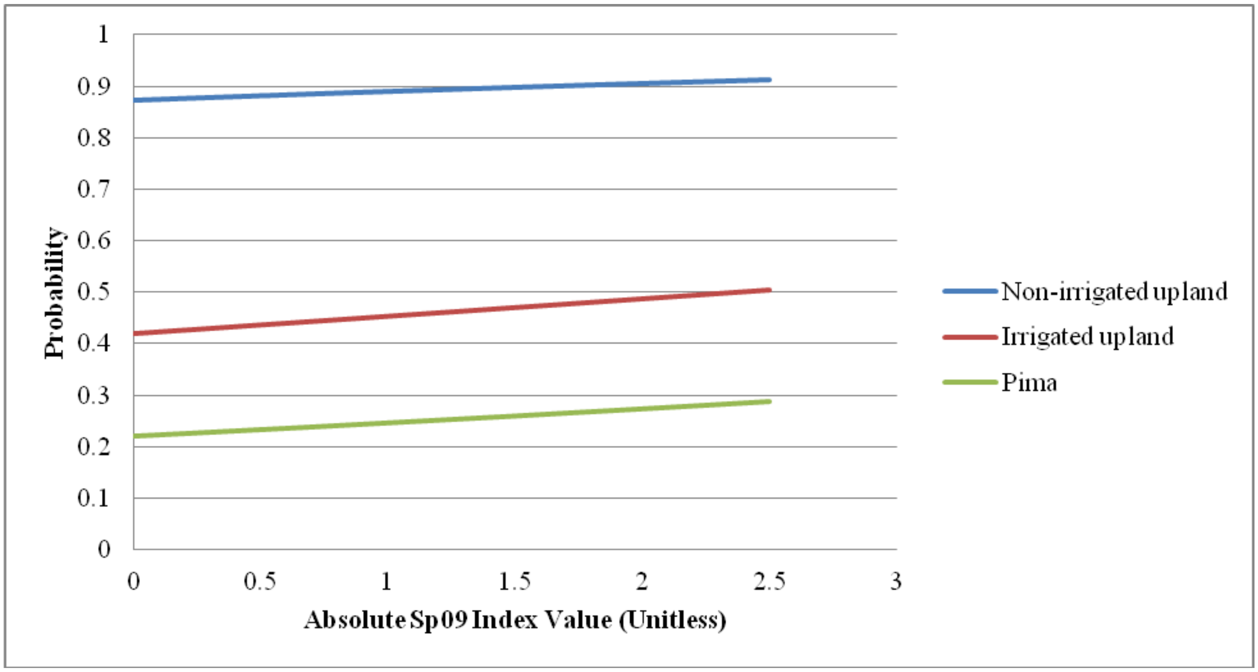


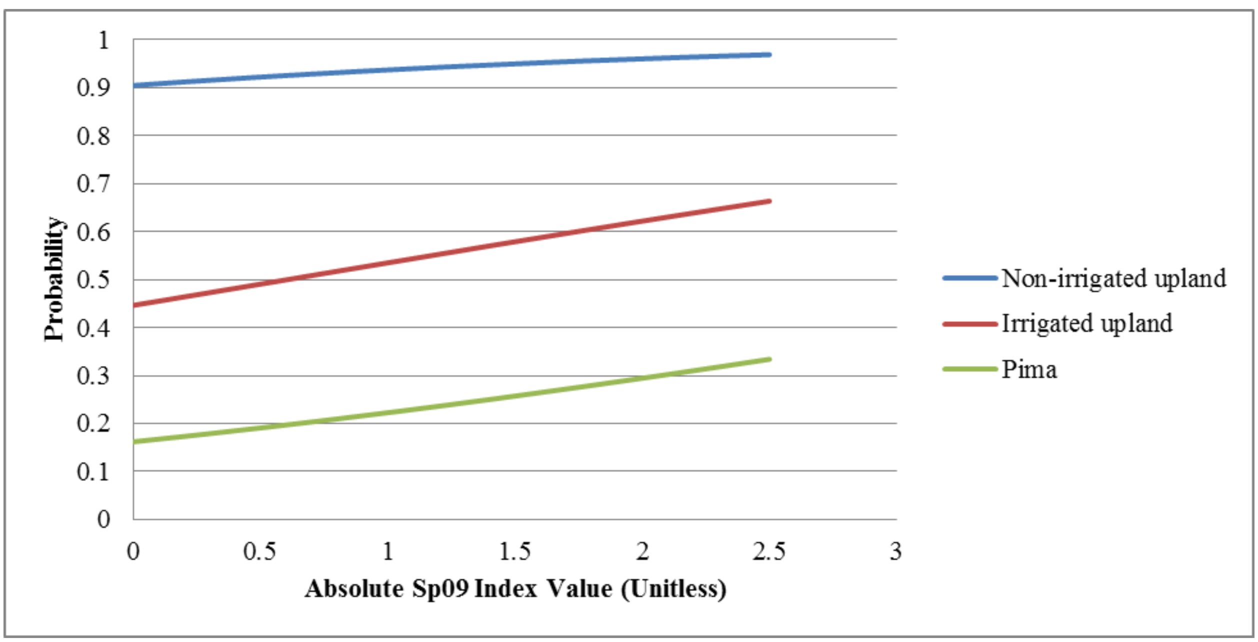
9. Marginal effect curves of 9-month standard precipitation index value under dry conditions for the probit model from full dataset (top) and census subset (bottom)





10. Marginal effect curves of 9-month standard precipitation index value under wet conditions for the probit model from full dataset (top) and census subset (bottom)

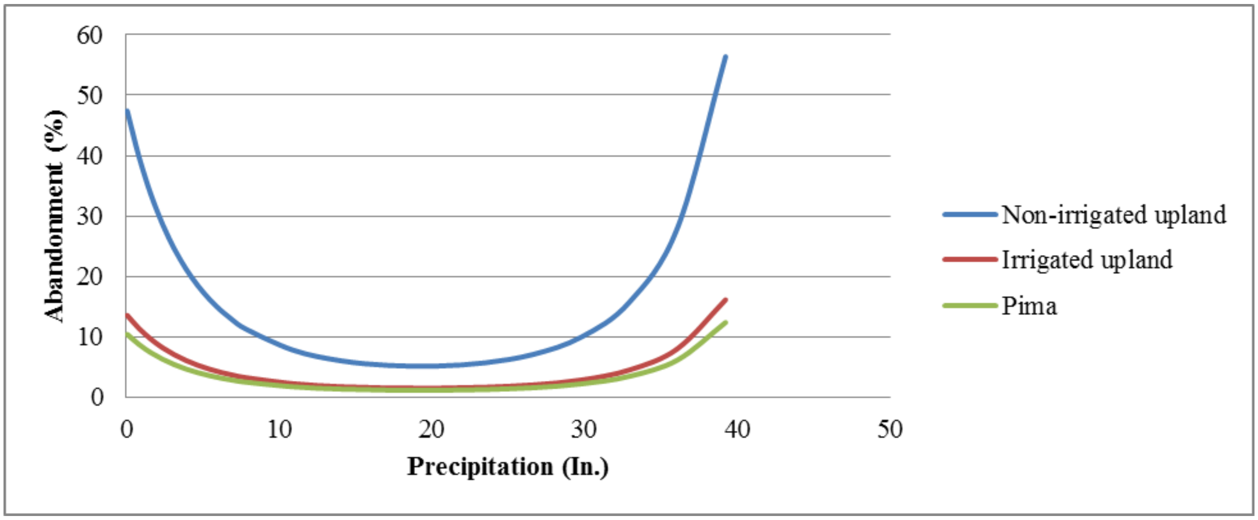


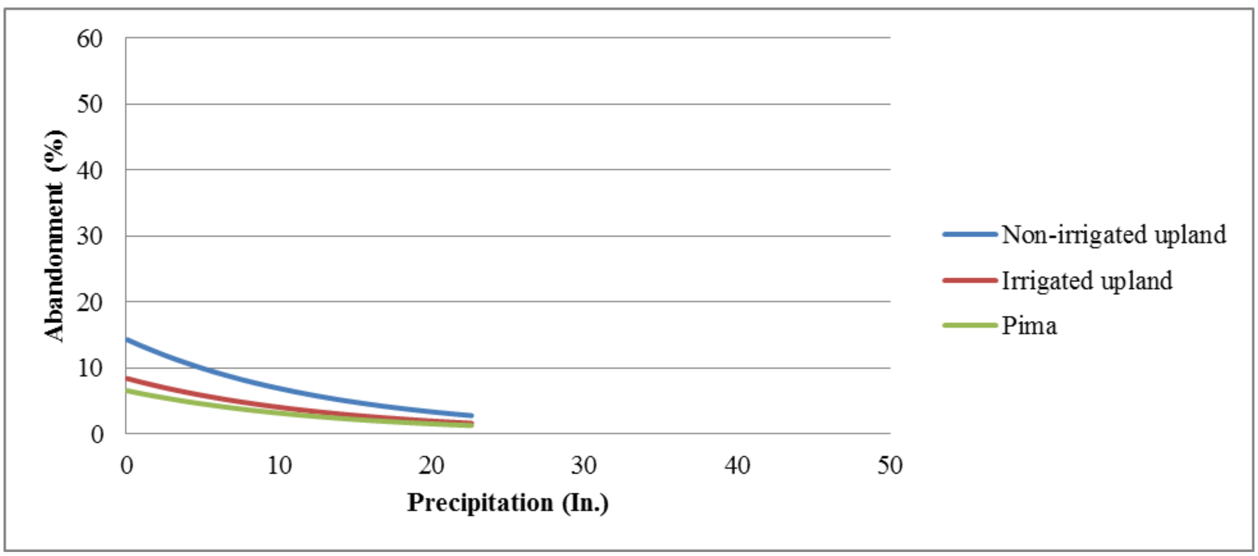


Appendix C

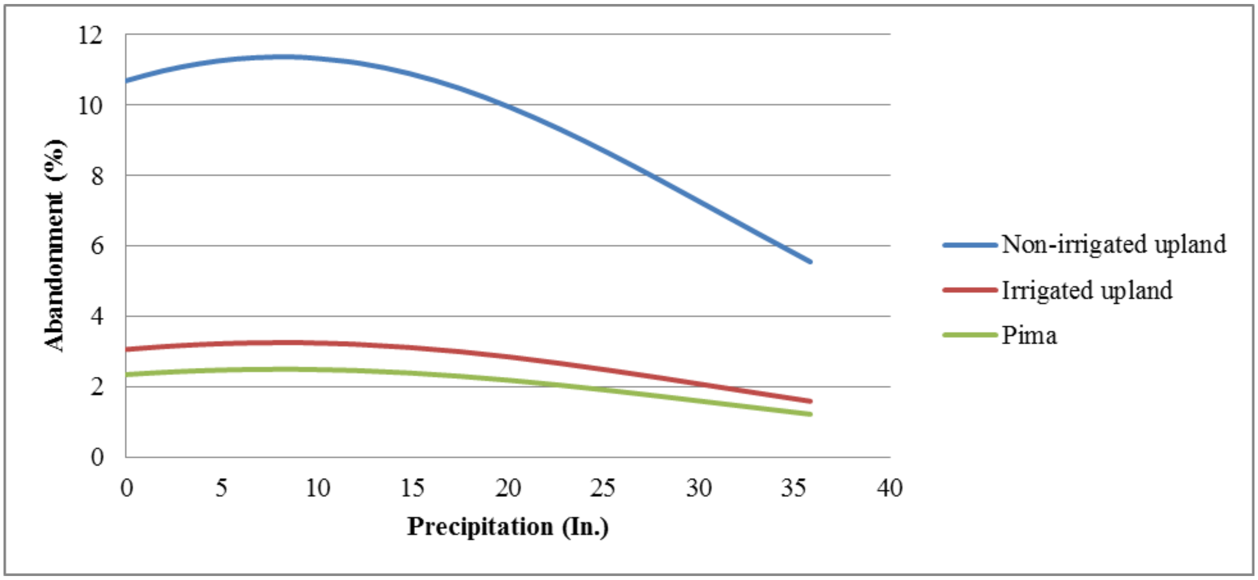
Marginal Effect Curves for OLS Model from Full Dataset and Census Subset

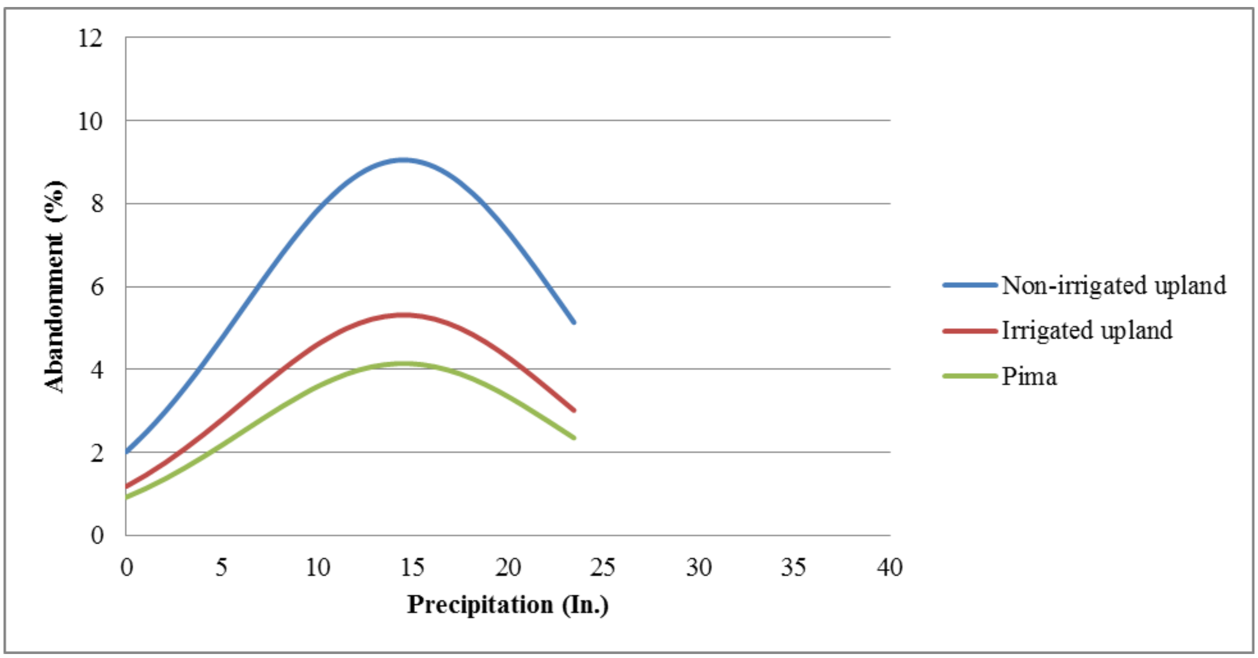
1. Marginal effect curves of season 1 precipitation for the OLS model from full dataset (top) and census subset (bottom)



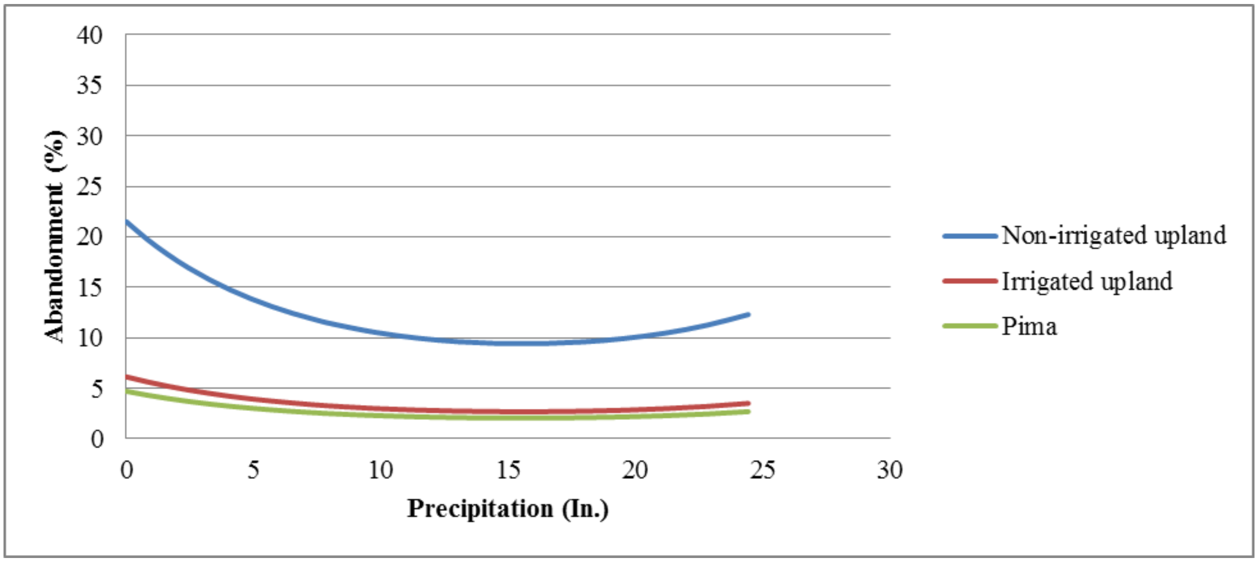


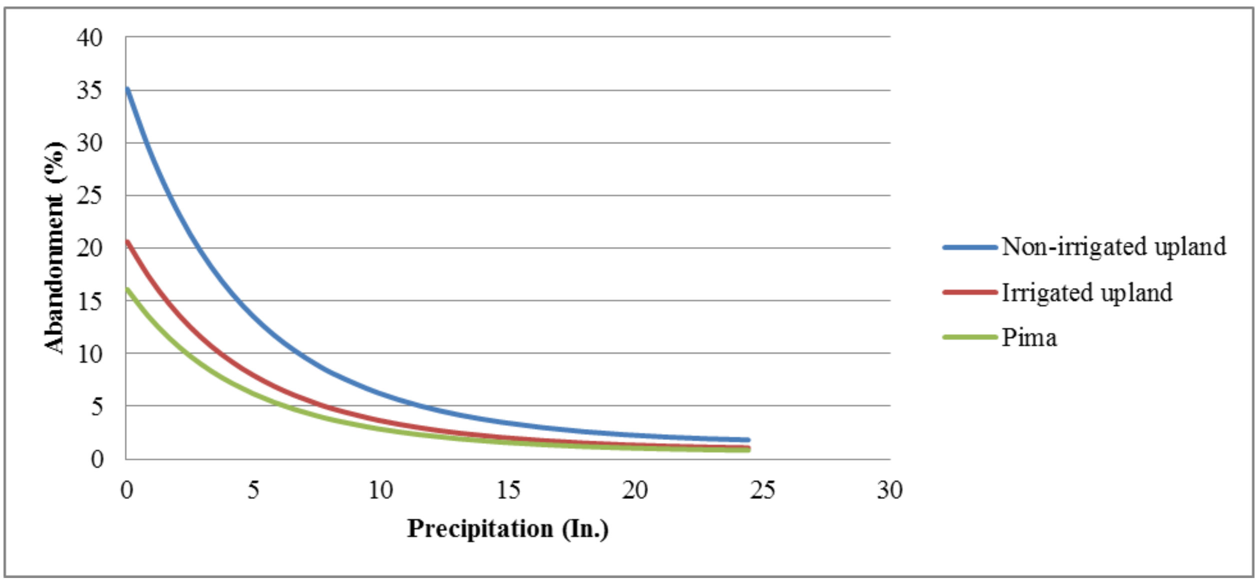
2. Marginal effect curves of season 2 precipitation for the OLS model from full dataset (top) and census subset (bottom)



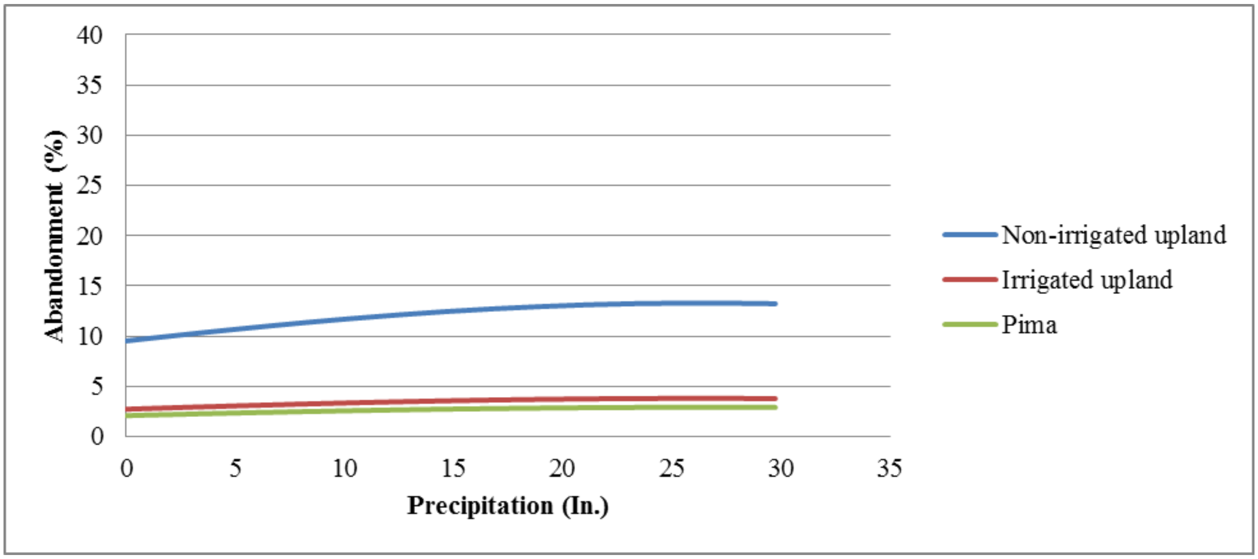


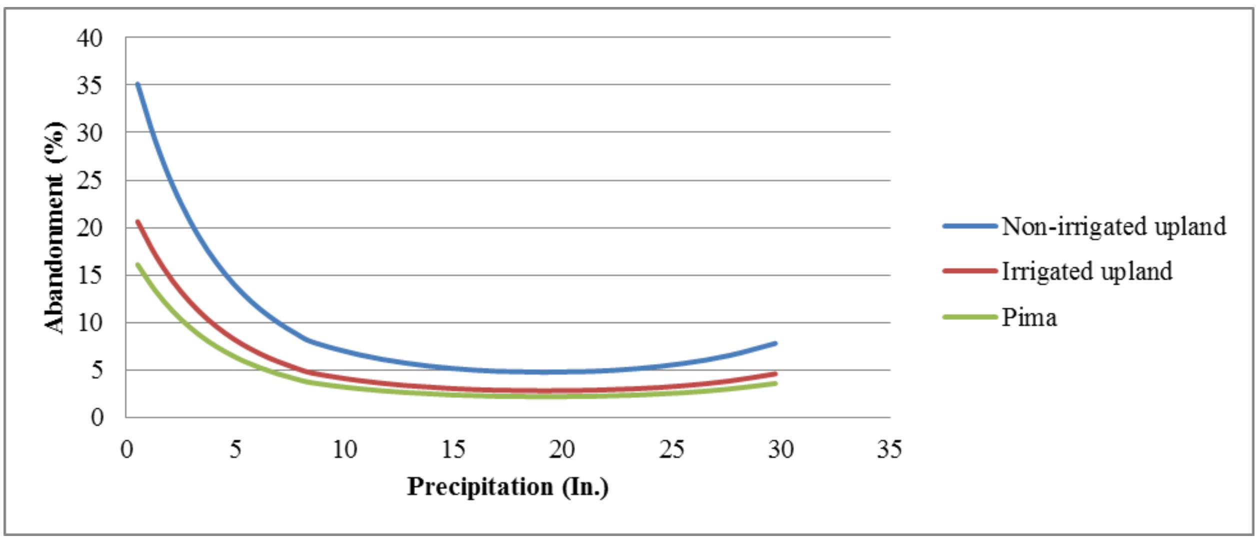
3. Marginal effect curves of season 3 precipitation for the OLS model from full dataset (top) and census subset (bottom)



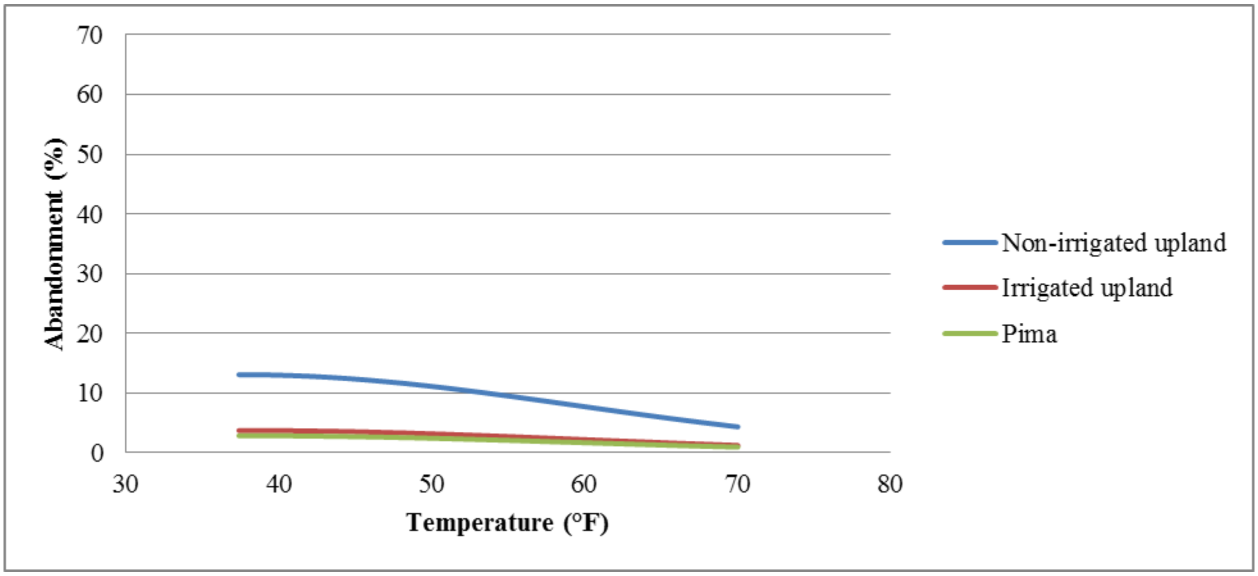


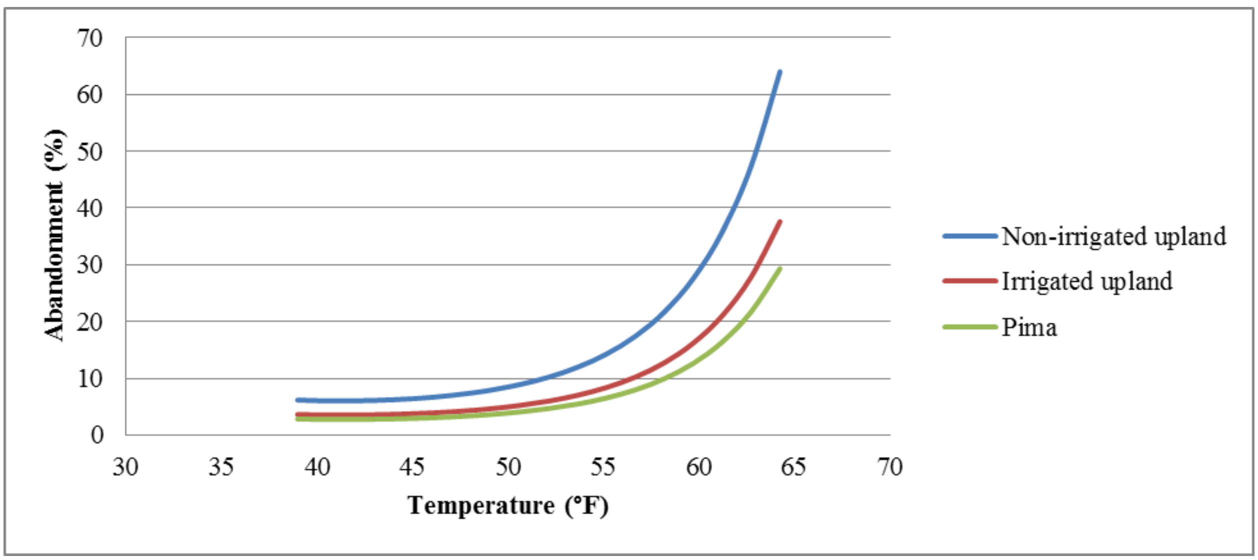
4. Marginal effect curves of season 4 precipitation for the OLS model from full dataset (top) and census subset (bottom)



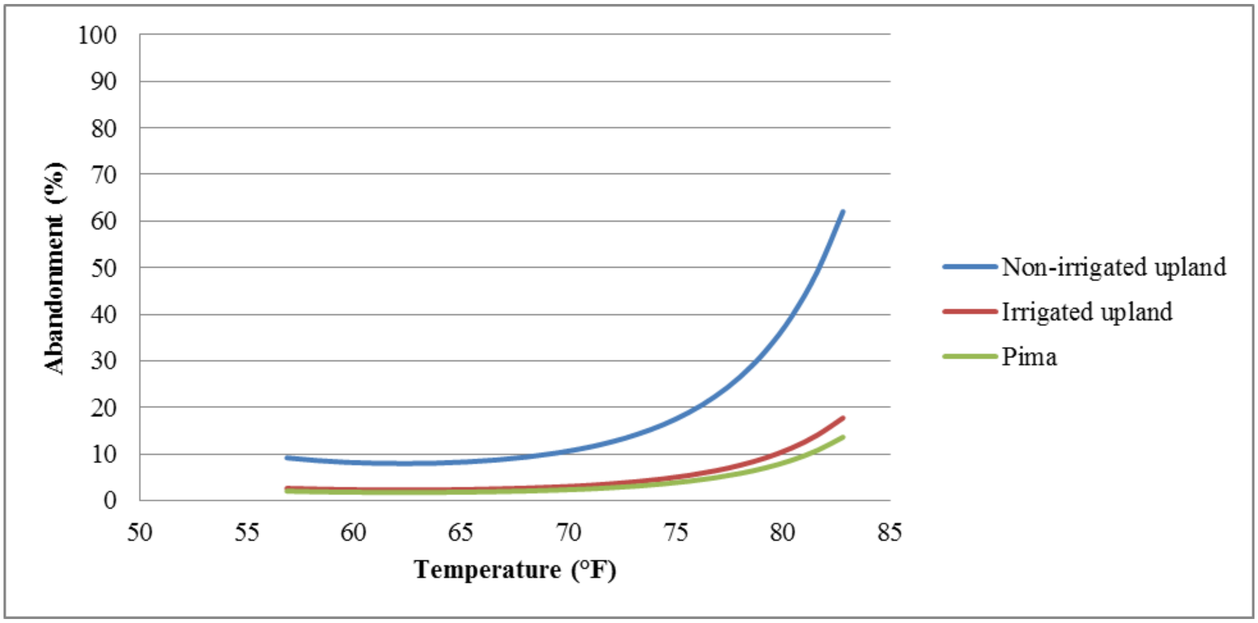


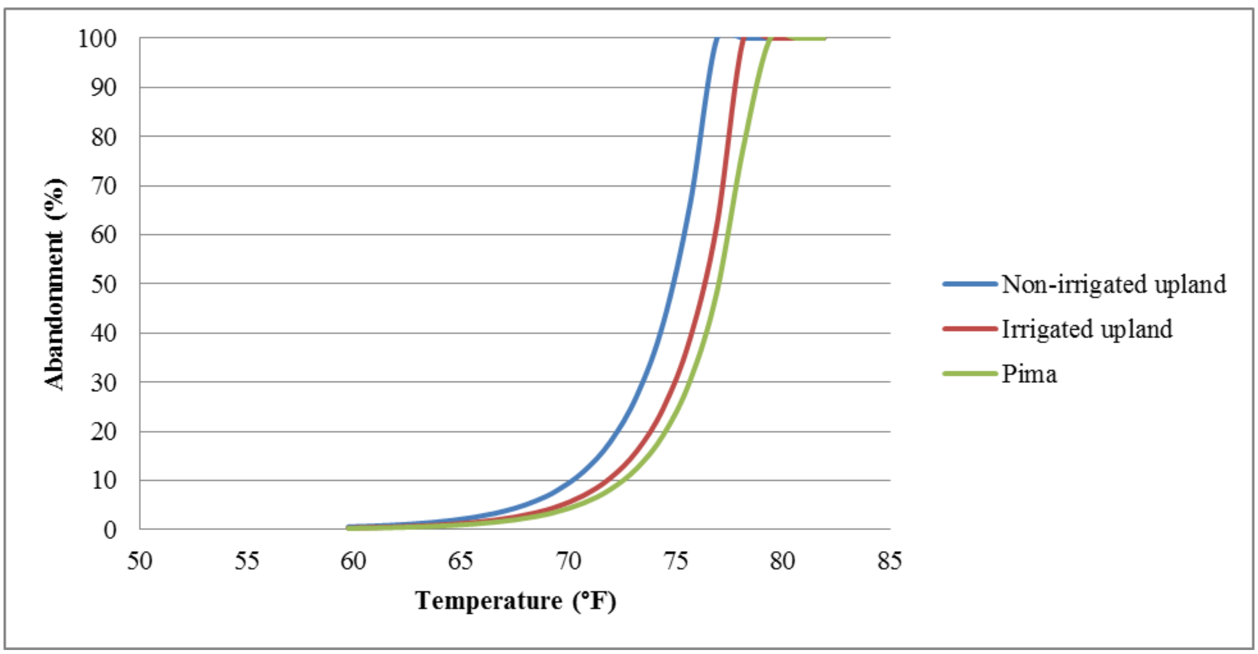
5. Marginal effect curves of season 1 temperature for the OLS model from full dataset (top) and census subset (bottom)



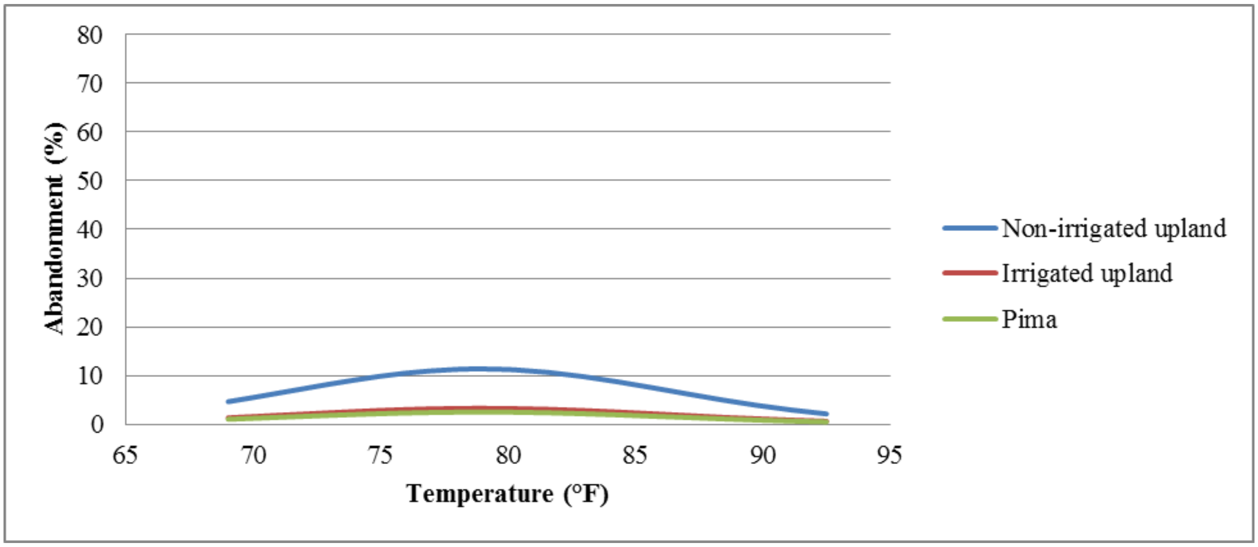


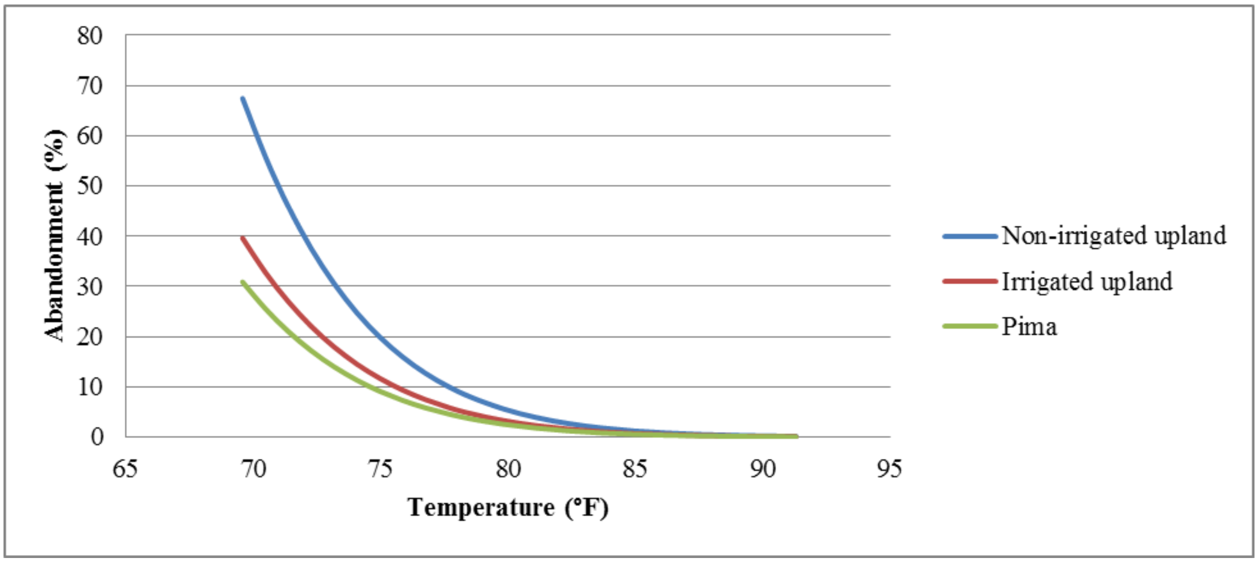
6. Marginal effect curves of season 2 temperature for the OLS model from full dataset (top) and census subset (bottom)



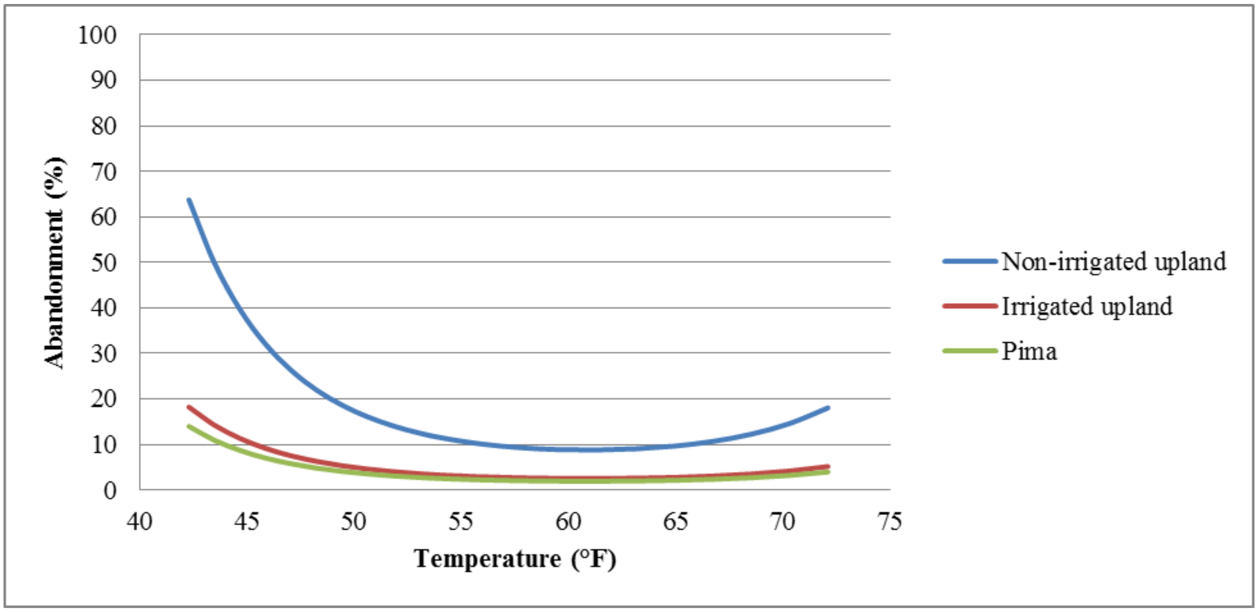


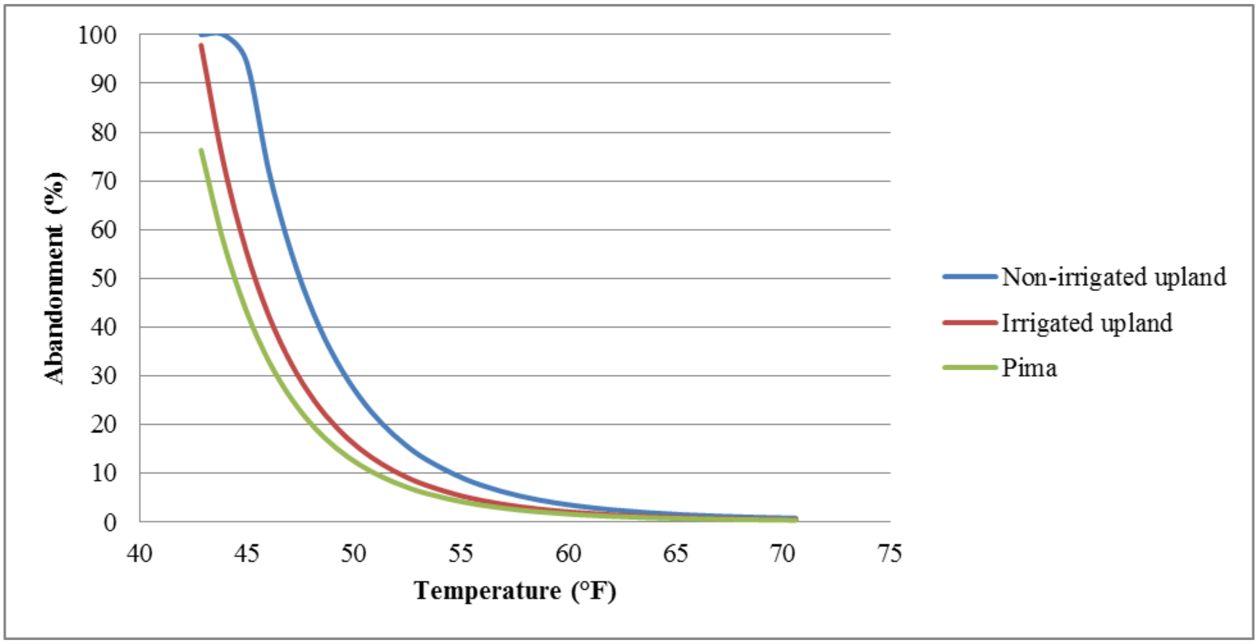
7. Marginal effect curves of season 3 temperature for the OLS model from full dataset (top) and census subset (bottom)



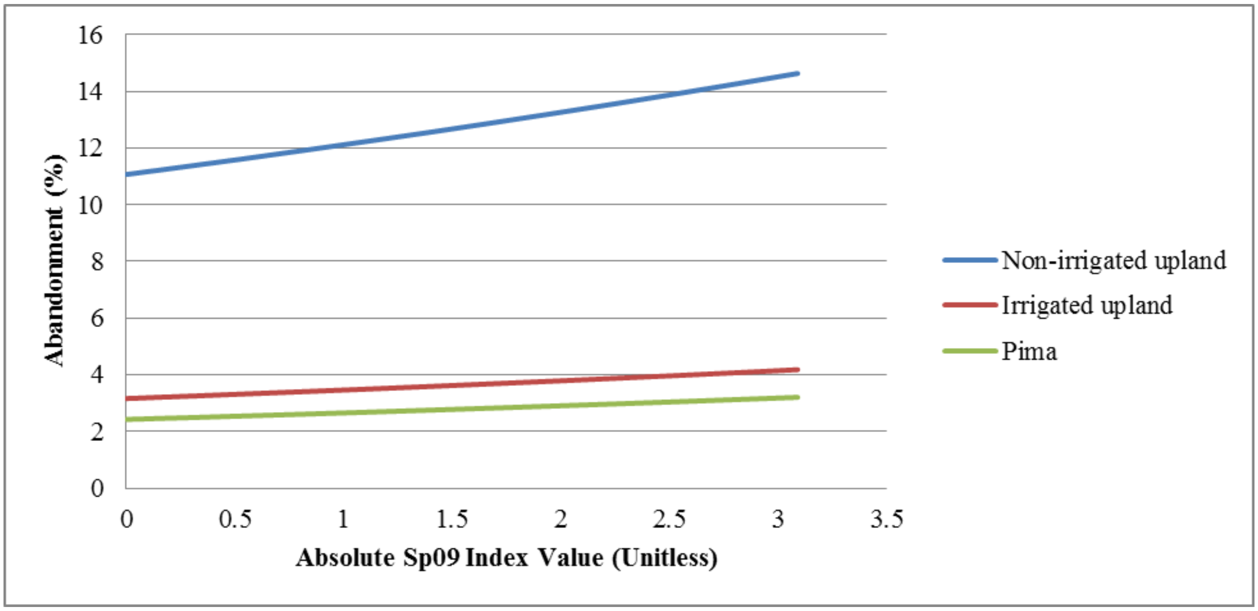


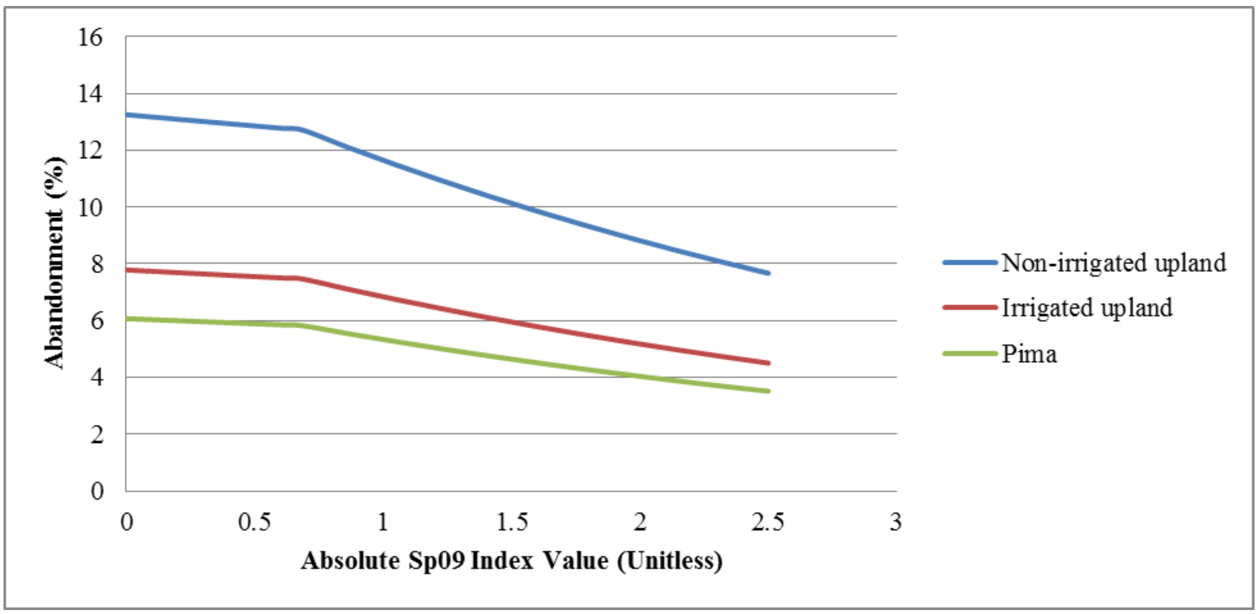
8. Marginal effect curves of season 4 temperature for the OLS model from full dataset (top) and census subset (bottom)



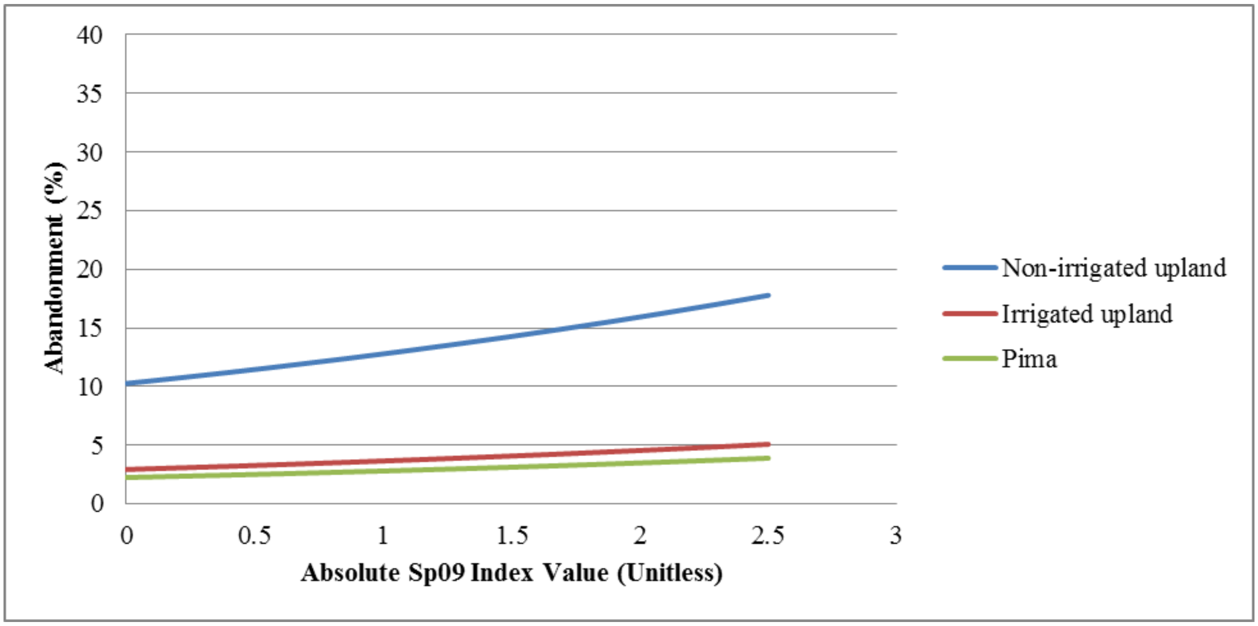


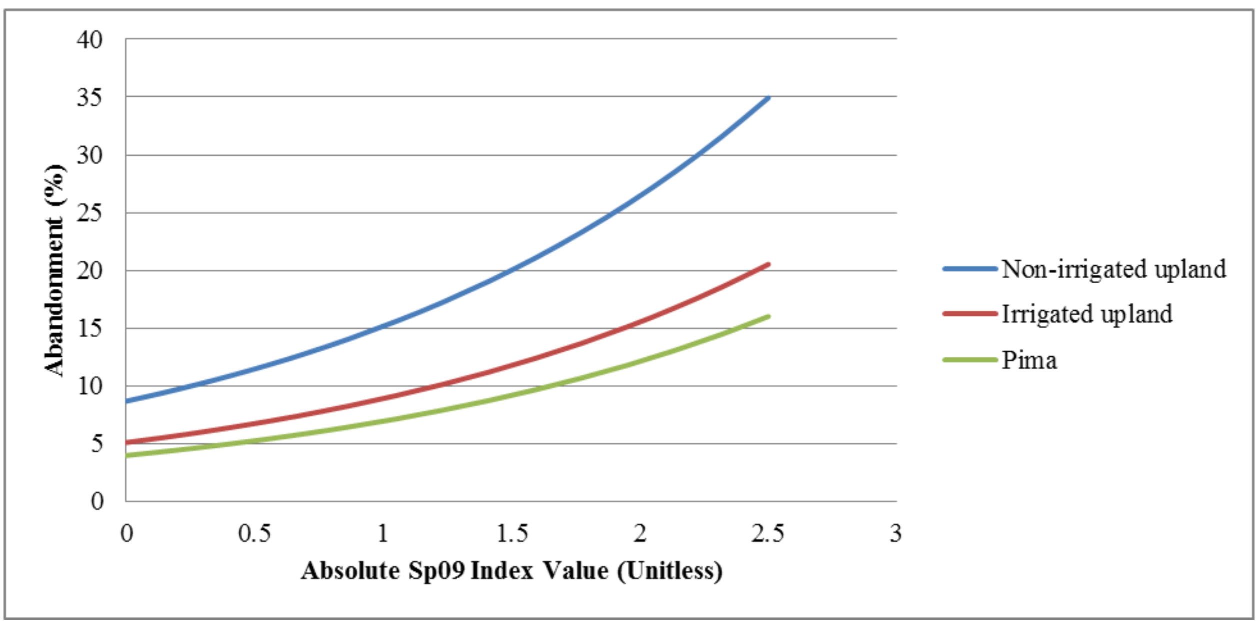
9. Marginal effect curves of 9-month standard precipitation index value under dry conditions for the OLS model from full dataset (top) and census subset (bottom)





10. Marginal effect curves of 9-month standard precipitation index value under wet conditions for the OLS model from full dataset (top) and census subset (bottom)

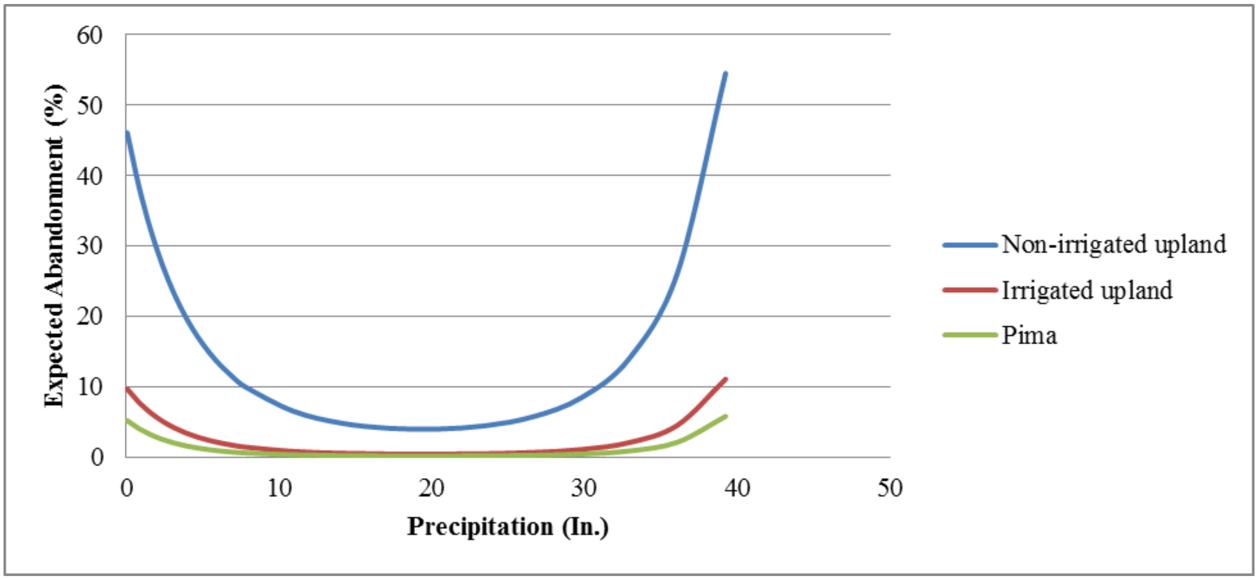




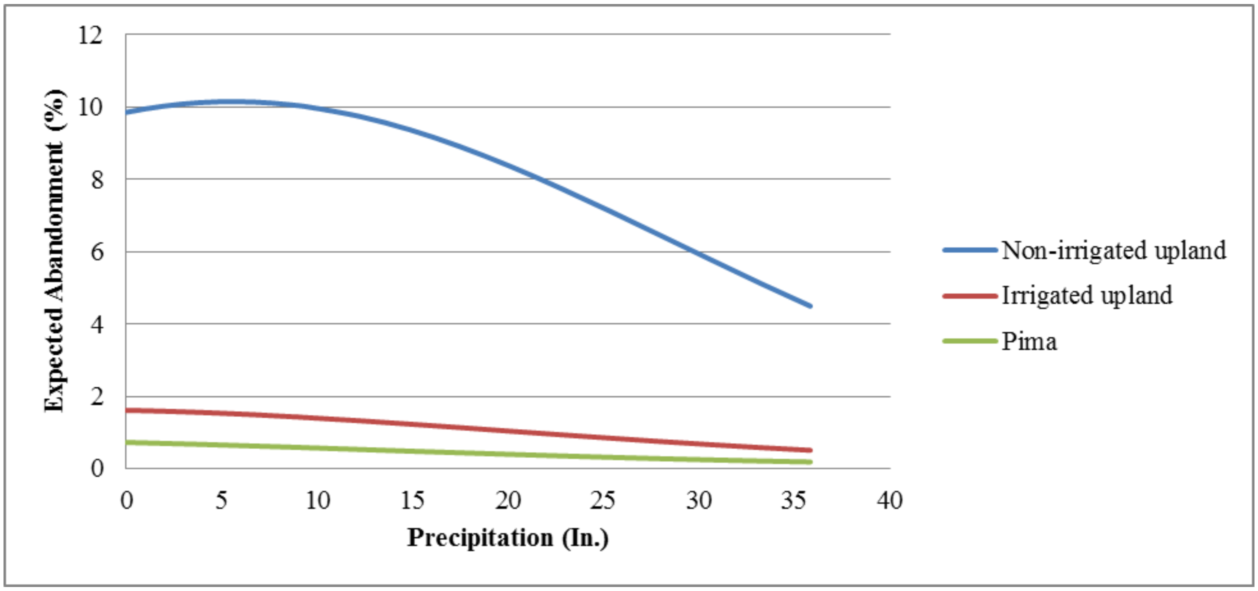
Appendix D

Marginal Effect Curves for Expected Values from Full Dataset

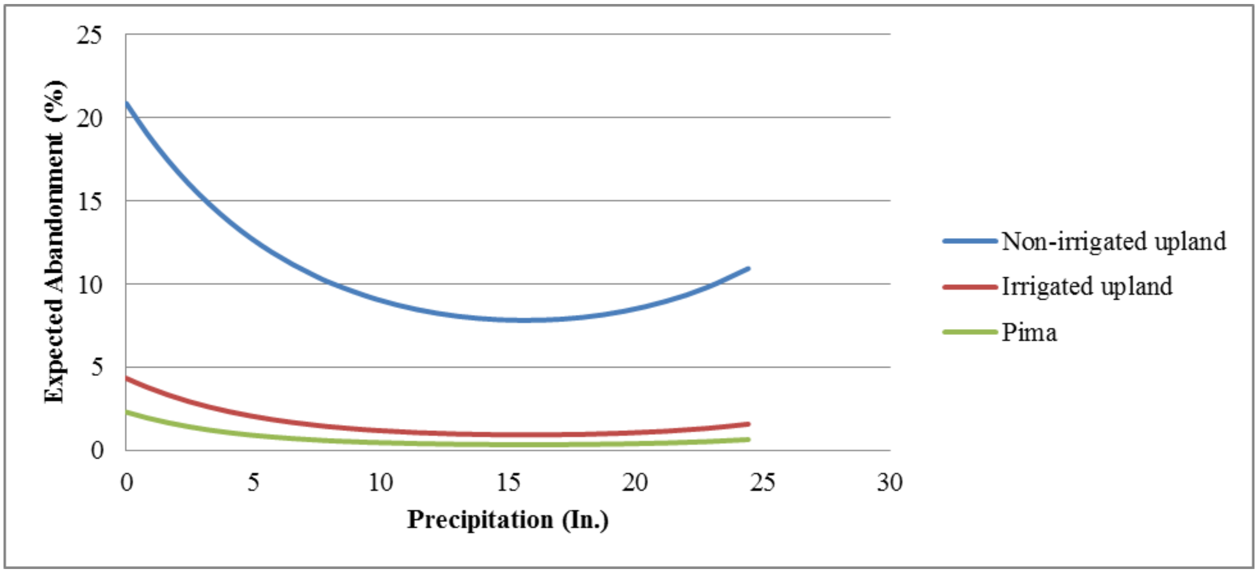
1. Marginal effect curves of season 1 precipitation for expected values from full dataset



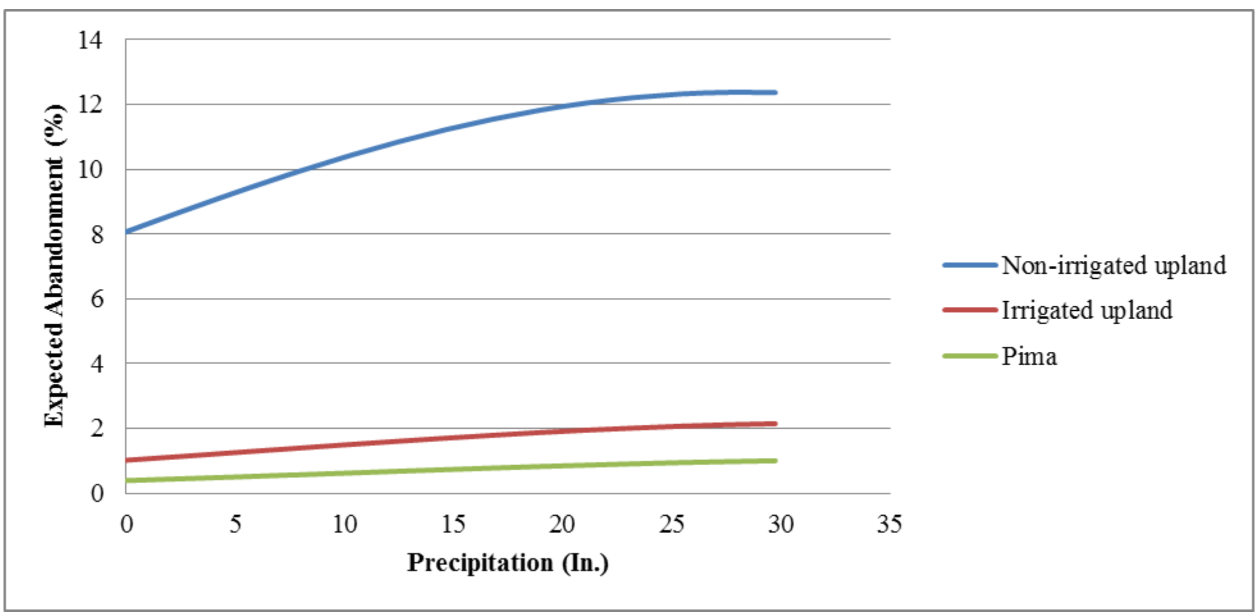
2. Marginal effect curves of season 2 precipitation for expected values from full dataset



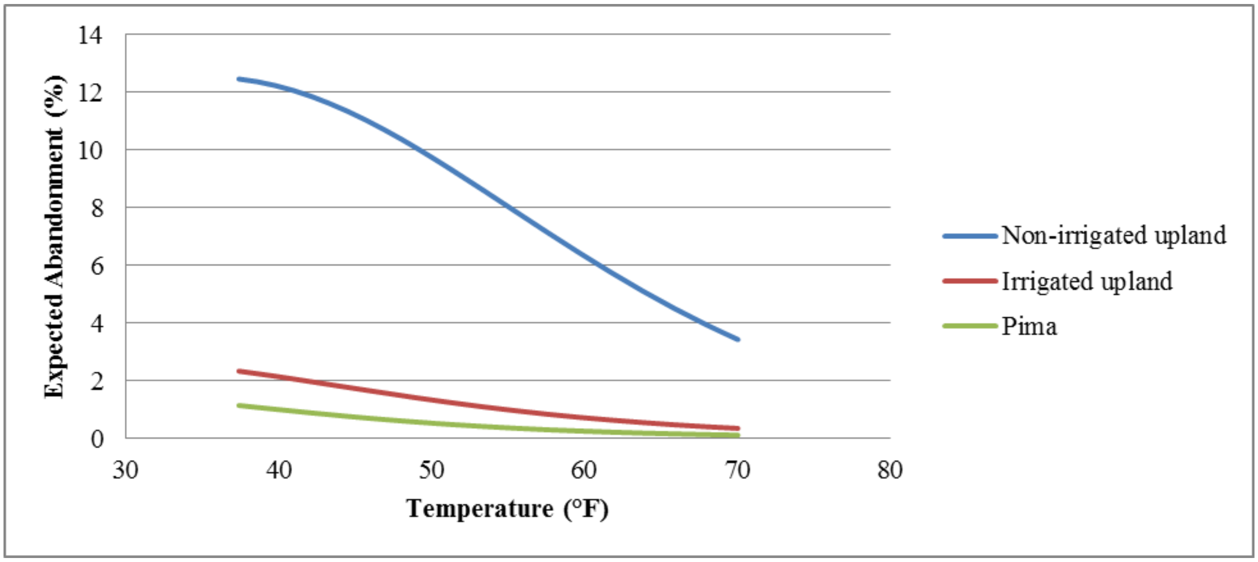
3. Marginal effect curves of season 3 precipitation for expected values from full dataset



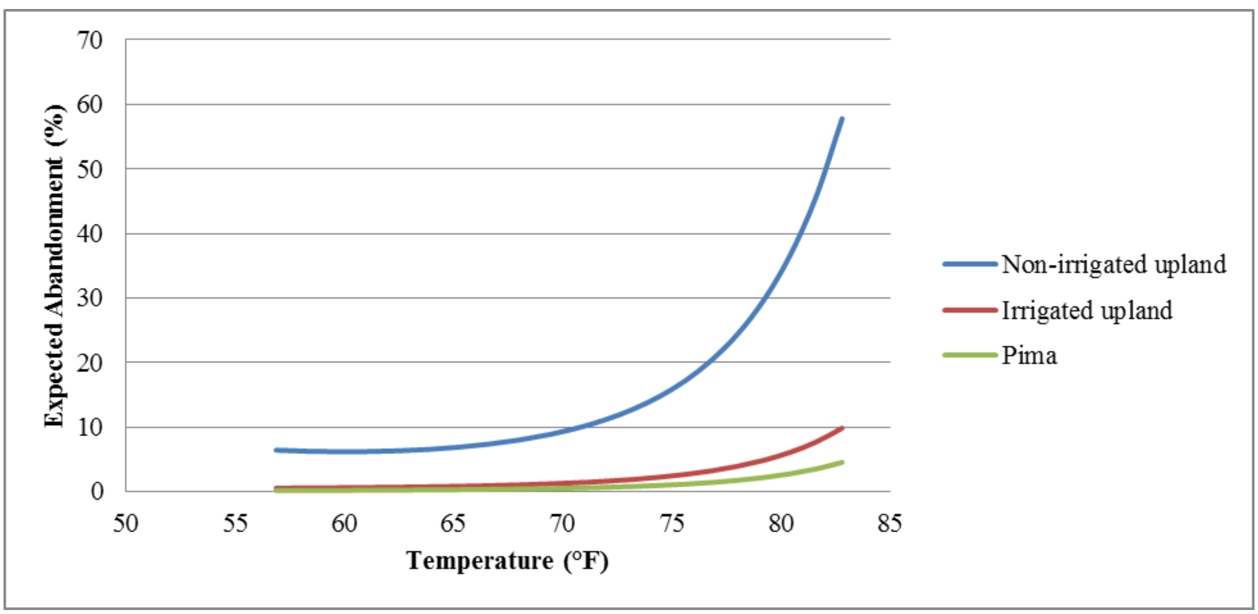
4. Marginal effect curves of season 4 precipitation for expected values from full dataset



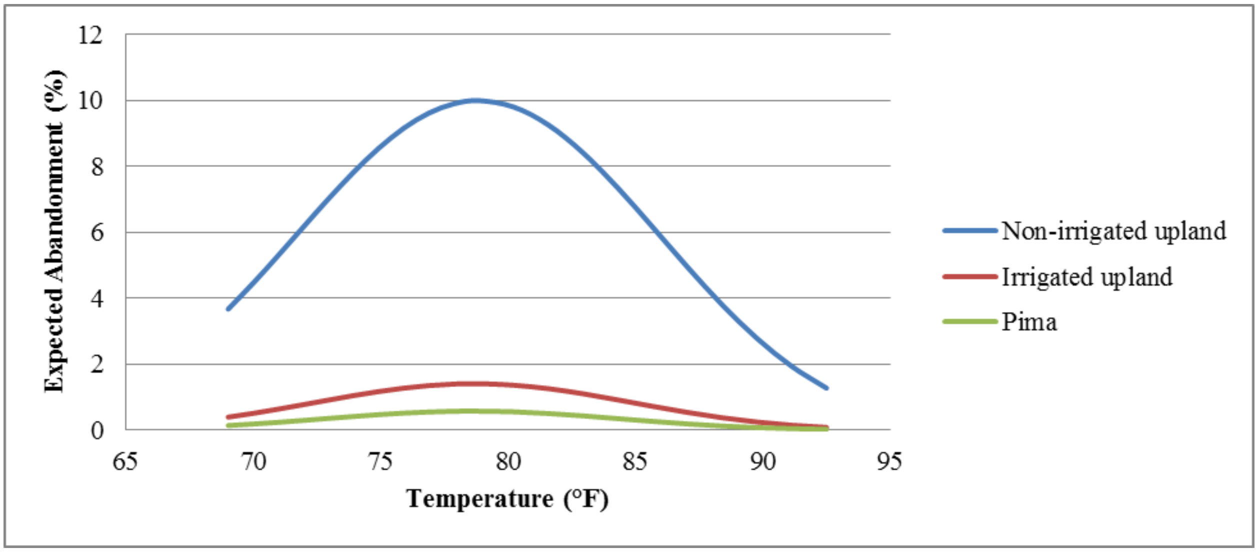
5. Marginal effect curves of season 1 temperature for expected values from full dataset



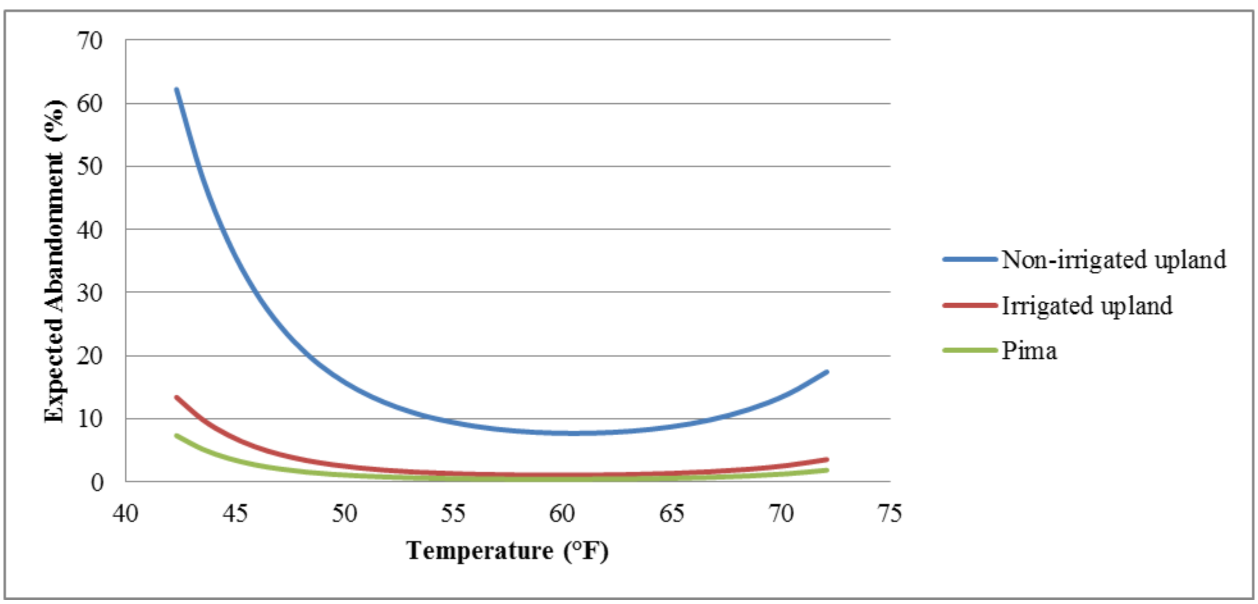
6. Marginal effect curves of season 2 temperature for expected values from full dataset



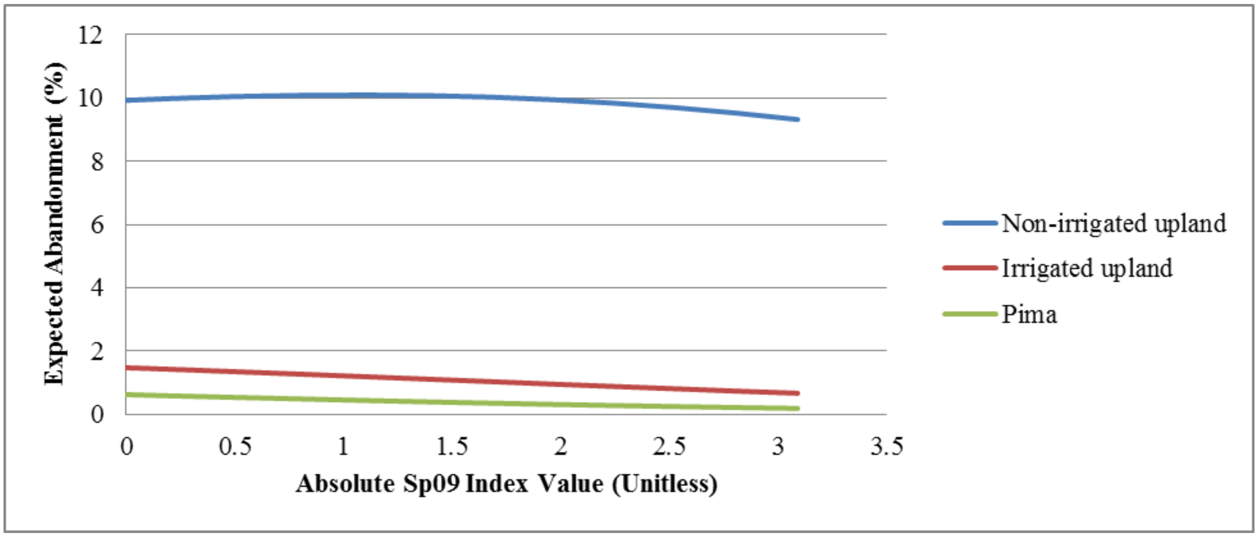
7. Marginal effect curves of season 3 temperature for expected values from full dataset



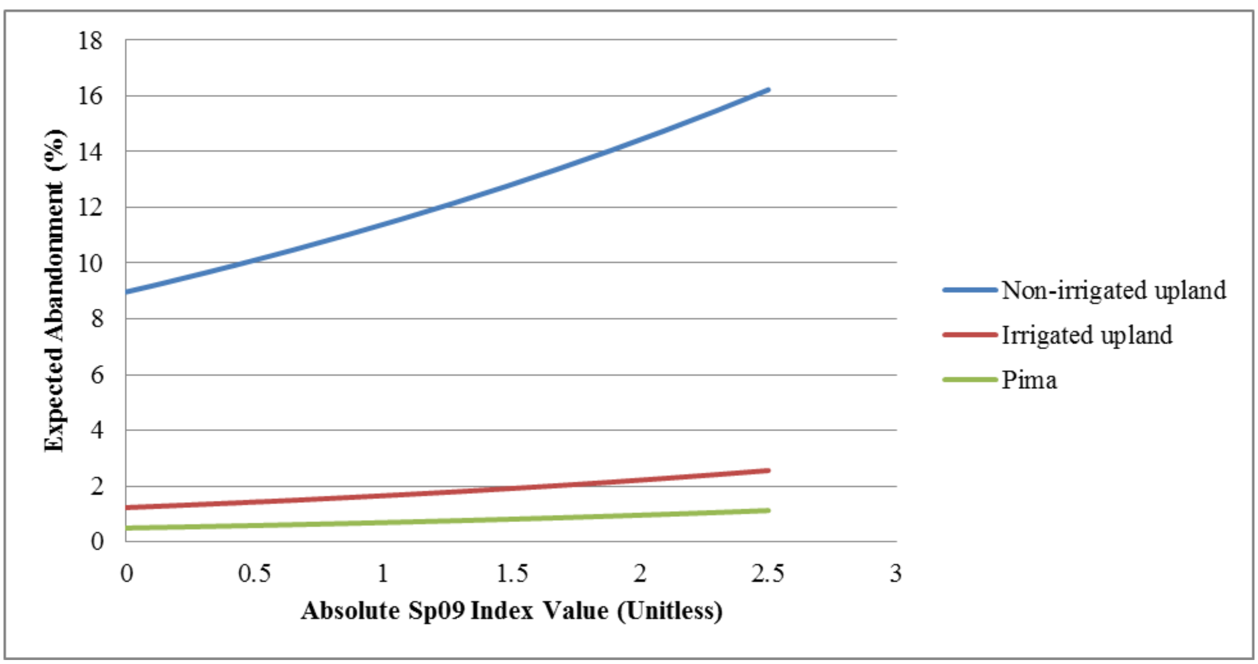
8. Marginal effect curves of season 4 temperature for expected values from full dataset



9. Marginal effect curves of 9-month standard precipitation index value under dry conditions for expected values from full dataset



10. Marginal effect curves of 9-month standard precipitation index value under wet conditions for expected values from full dataset



REFERENCES

Baier, W. (1973). "Crop-Weather Analysis Model: Review and Model Development." J. Appl. Meteor. 12(6): 937-947.

Brown, M. J. (1959). "The Relation of Weather Factors to the Yield of Winter Wheat in Box Elder County, Utah." Mon. Wea. Rev. Monthly Weather Review 87(3): 97-99.

Climate Prediction Center, National Oceanic and Atmospheric Administration. (2013). Webpage. (http://www.cpc.ncep.noaa.gov/products/analysis_monitoring/regional_monitoring/CLIM_DIVS/states_counties_climate-divisions.shtml).

Chen, S. L. (2005). "Acreage Abandonment, Moral Hazard and Crop Insurance." American Agricultural Economics Association.

Combs, S. (2012). "The impact of the 2011 drought and beyond". (http://www.window.state.tx.us/specialrpt/drought/pdf/96-1704-Drought.pdf)

Cragg, J. (1971), ''Some Statistical Models for Limited Dependent Variables with Applications to the Demand for Durable Goods,'' Econometrica 39, 829–844.

Economic Research Service, United States Department of Agriculture, (2013). Webpage. (http://www.agcensus.usda.gov/Publications/index.php).

Mendelsohn, R. (2007). "What causes crop failure?" Climatic Change 81(1): 61-70.

Michaels, P. J. (1983). "Price, weather, and 'acreage abandonment' in western great plains wheat culture". J. Climate Appl. Meteor. 22: 1296–1303.

Michaels, P. J. (1985). "Economic and climatic factors in 'acreage abandonment' over marginal cropland". J. Climatic Change 7(2): 185-202.

National Agricultural Statistics Service, United States Department of Agriculture. (2013). Webpage. (http://quickstats.nass.usda.gov/).

National Climatic Data Center, National Oceanic and Atmospheric Administration. (2013). Webpage. (http://www7.ncdc.noaa.gov/CDO/CDODivisionalSelect.jsp#).

National Cotton Council of America. (2013). Webpage. (http://www.cotton.org/econ/prices/monthly.cfm).

National Weather Service Weather Forecast Office, National Oceanic and Atmospheric Administration. (2012). Webpage. (http://www.srh.noaa.gov/tsa/?n=weather-event_2011drought)

Starr, T. B. and P. I. Kostrow (1978). "The Response of Spring Wheat Yield to Anomalous Climate Sequences in the United States." J. Appl. Meteor 17(8): 1101-1115.

United States Department of Agriculture. (2013). Webpage. (http://www.ers.usda.gov/topics/crops/cotton-wool/background.aspx#.Uav9ZUDqm8A).

Wooldridge, J. M. (2002). “Econometric Analysis of Cross Section and Panel Data”. Cambridge, Mass: MIT Press.