Group I: Informal Sensitivity - Almond Yields

Paloma Cartwright, Julia Parish, Quin Smith

2022/04/19

Informal Sensitivity - Almond Yields

This environmental model and sensitivity analysis was completed as an assignment for the course, Environmental Data Science 230 | Environmental Science & Management: Modeling Environmental Systems. The goal of this assignment was to develop a profit model for the California almond yield, conduct a simple informal sensitivity analysis, and graph analysis results. The source data and model design is based on research published in the paper, Impacts of future climate change on California perennial crop yields: Model projections with climate and crop uncertainties (Lobell 2006).

Load Functions

```
source(here("R/almond_model.R"))
source(here("R/compute_NPV.R"))
source(here("R/compute_profit_fromanmly.R"))
source(here("R/almond_model_test.R"))
```

Read in Climate Data

```
clim <- read.table(here("data/clim.txt"), header = TRUE) %>%
  clean_names()
```

Create Data Frame of Almond Prices

This data was based off of the USDA Almond Forecast, which can be found here (USDA/NASS 2019). For the years 1989 - 1994, the almond price was not provided we used the average price from 1995 - 2010 as the value for the missing years.

```
#years 1989-2010
year <- c(1989, 1990, 1991, 1992, 1993, 1994, 1995,
          1996, 1997, 1998, 1999, 2000, 2001, 2002,
          2003, 2004, 2005, 2006, 2007, 2008, 2009, 2010)

#$ per pound * 2000 = $ per ton
prices <- c(2.02, 2.02, 2.02, 2.02, 2.02, 2.02, 2.48,
            2.08, 1.56, 1.41, 0.86, 0.97, 0.91, 1.11,
            1.57, 2.21, 2.81, 2.06, 1.75, 1.45, 1.65, 1.79) * 2000

price_per_year <- data.frame(year, prices)
```

Calculate Yield Anomaly

Subset the data

```r
#subset the clim data for monthly averages
clim_avg <- clim %>%
  group_by(year, month) %>%
  summarize(mean_temp = mean(tmin_c),
            tot_precip = sum(precip)) %>%
  filter(month %in% c(1, 2))

# create a variable for avg. February temp. data
Tn2 <- clim_avg %>%
  group_by(year) %>%
  mutate(Tn2 = case_when(month == 2 ~ mean_temp)) %>%
  select(year, Tn2) %>%
  drop_na()

# create a variable for avg. January precipitation data
P1 <- clim_avg %>%
  group_by(year) %>%
  mutate(P1 = case_when(month == 1 ~ tot_precip)) %>%
  select(year, P1) %>%
  drop_na

# create a variable combined temp and precip data and added price
func_vars <- cbind(Tn2, P1 = P1$P1, price = price_per_year$prices)
```

Run Almond Yield Model

```r
#run almond_model() on subsetted clim data
almond_yield_anmly <- func_vars %>%
  mutate(yield_anmly = almond_model(Tn2 = Tn2, P1 = P1))
```

Assignment 03 Responses

1. Develop a profit model for almond yield

```r
#run compute_profit_fromanmly() over the almond data
profit <- compute_profit_fromanmly(anmly = almond_yield_anmly$yield_anmly,
                                   year = almond_yield_anmly$year,
                                   price = price_per_year$prices)
```

We first developed a profit model function to calculate the profit generated from almond yield anomaly (i.e. almond yield profit anomaly). We used historical California almond price data from the USDA as the basis for price. Price data was missing from 1989-1994. For these missing values, we used the average price between 1995-2010 ($2.02/lb `* 2000 lbs/ton`). Profit was measured as the net present value with a 12% discount rate, taken from the examples in class. A positive yield anomaly produced an increase in profit (positive $ value), while a negative yield anomaly produced a reduction in profit (negative $ value). The results of this analysis are shown in the figure below.

Profit model visualization

```r
profit_plot <- ggplot(data = profit,
                     aes(x = year, y = netpre)) +
  geom_line(color = "#E1BE6A") +
  geom_point(color = "#40B0A6") +
  scale_y_continuous(labels=scales::dollar_format()) +
  geom_text(aes(x = 1991, y = 400000, label = "$221,945"), stat = "unique",
          size = 3, color = "slategrey") +
  geom_text(aes(x = 1995, y = 5000000, label = "$4,824,702"), stat = "unique",
          size = 3, color = "slategrey") +
  geom_text(aes(x = 1997, y = 600000, label = "$415,452"), stat = "unique",
          size = 3, color = "slategrey") +
  geom_text(aes(x = 2005, y = 800000, label = "$601,725"), stat = "unique",
          size = 3, color = "slategrey") +
  geom_text(aes(x = 2008, y = 400000, label = "$194,039"), stat = "unique",
          size = 3, color = "slategrey") +
  labs(title = "California Almond Yield Profit Anomaly",
      subtitle = "1989 - 2010",
      caption = "Data source: Lobell 2006,USDA/NASS 2019",
      x = "Year",
      y = "Profit in current $") +
  theme(plot.title = element_text(face = "bold", size = 14),
        axis.title.x = element_text(face = "bold", size = 11),
        axis.title.y = element_text(face = "bold", size = 11),
        panel.background = element_rect(fill = "white"),
        panel.border = element_blank(),
        axis.line = element_line(colour = "black"),
        panel.grid.major = element_line(size = 0.25,
                                        linetype = 'solid', colour = "grey88"),
        panel.grid.minor = element_blank())

profit_plot
```

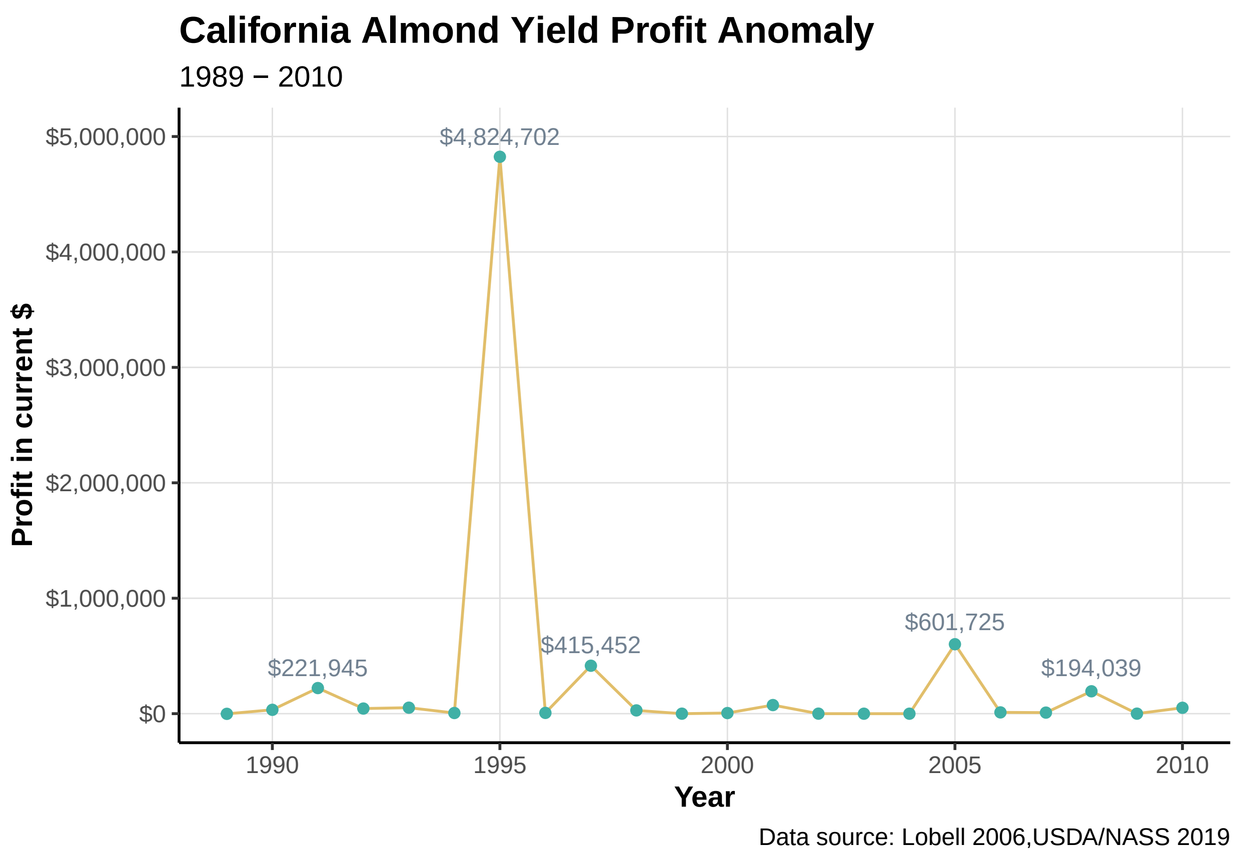

Figure 1: Annual profit generated from California almond yield anomaly

2. Conduct a simple informal sensitivity analysis of total almond yield profit using at least 2 parameters

```r
nsamples <- 100

# test the sensitivity of model to each param (temp, precip)
#Tn2 sensitivity
#uniform random sample
Tn2_sample <- runif(n = nsamples, min = min(func_vars$Tn2), max = max(func_vars$Tn2))

#run almond_model_test() over the variables, using sample of Tn2
results_Tn2 <- Tn2_sample %>%
  map(~almond_model_test(df = func_vars, P1 = func_vars$P1, price = func_vars$price, Tn2 = .x))

#create temporary data structure from purrr results
tmpTn2 <- map_df(results_Tn2,`[`, c("profit"))

#create dataframe from temporary results
Tn2_df <- data.frame(year = tmpTn2$profit$year,
                     netpre= tmpTn2$profit$netpre)

#create temporary data strucuture from purrr results for mean
tmpTn2mn <- map_df(results_Tn2,`[`, c("mean_netpre"))

#create dataframe from temporary results
Tn2mn_df <- data.frame(Tn2 = Tn2_sample, mean_netpre = tmpTn2mn$mean_netpre)

#P1 sensitivity
#uniform random sample
P1_sample <- runif(n = nsamples, min = min(func_vars$P1), max = max(func_vars$P1))

#run almond_model_test() over the variables, using sample of P1
results_P1 <- P1_sample %>%
  map(~almond_model_test(df = func_vars, Tn2 = func_vars$Tn2, price = func_vars$price, P1 = .x))

#create temporary data structure from purrr results
tmpP1 <- map_df(results_P1, `[`, c("profit"))

#create dataframe from temporary results
P1_df <- data.frame(year = tmpP1$profit$year,
                    netpre= tmpP1$profit$netpre)

#create temporary data strucuture from purrr results for mean
tmpP1mn <- map_df(results_P1,`[`, c("mean_netpre"))

#create dataframe from temporary results
P1mn_df <- data.frame(P1 = P1_sample, mean_netpre = tmpP1mn$mean_netpre)

#price sensitivity
#uniform random sample
price_sample <- runif(n = nsamples, min = min(func_vars$price), max = max(func_vars$price))

#run almond_model_test() over the variables, using sample of price
results_price <- price_sample %>%
```

```r
  map(~almond_model_test(df = func_vars, P1 = func_vars$P1, Tn2 = func_vars$Tn2, price = .x))

#create temporary data structure from purrr results
tmpprice <- map_df(results_price, `[`, c("profit"))

#create dataframe from temporary results
price_df <- data.frame(year = tmpprice$profit$year,
                     netpre= tmpprice$profit$netpre)

#create temp data structure from purrr results for mean
tmppricemn <- map_df(results_price,`[`, c("mean_netpre"))

#create dataframe from temporary results
pricemn_df <- data.frame(price = price_sample, mean_netpre = tmppricemn$mean_netpre)
```

To better understand the influence of the different variables in the Lobell (2006) almond yield anomaly model on profit, we conducted an informal sensitivity analysis using the purrr package in R. First, we constructed 100 samples of January total precipitation, February monthly average low temperatures, and almond yield price using random uniform distributions. We used actual minimum and maximum values from the data (and averages for price between 1989 and 1994) for each variable to constrain the uniform distribution range. We then used the map() function in purrr to run all the samples in each distribution over the data, keeping all other variables constant.

3. Create a single graph of the results

We first produced the distribution of annual profit anomalies as a function of uniform variation in the three different variables, shown in the boxplot graphs below.

Boxplots

```r
# function of Feb temperature box plot
Tn2_plot <- ggplot(Tn2_df,
       aes(x = as.factor(year), y = netpre, group=year)) +
  geom_boxplot(color = "gray61", fill = "grey61") +
  labs(subtitle = "A. Function of Minimum Temperature (°C)",
       x = " ",
       y = "Profit in current $") +
  scale_x_discrete(breaks=seq(1989, 2010, 2)) +
  scale_y_continuous(labels=scales::dollar_format()) +
  theme_minimal() +
  theme(axis.text.x = element_text(angle = 45, hjust=1))

# function of Jan precip box plot
P1_plot <- ggplot(P1_df,
       aes(x = as.factor(year), y = netpre, group=year)) +
  geom_boxplot(color = "gray61", fill = "lightskyblue1") +
  labs(subtitle = "B. Function of Precipitation (mm)",
       x = " ", y = " ") +
  scale_x_discrete(breaks=seq(1989, 2010, 2)) +
  scale_y_continuous(labels=scales::dollar_format()) +
  theme_minimal() +
  theme(axis.text.x = element_text(angle = 45, hjust=1))

# function of price box plot
```

```r
price_plot <- ggplot(price_df,
       aes(x = as.factor(year), y = netpre, group=year)) +
  geom_boxplot(color = "gray61", fill = "navajowhite") +
  labs(subtitle = "C. Function of Price",
       x = " ", y = "Profit in current $") +
  scale_y_continuous(labels=scales::dollar_format()) +
  theme_minimal() +
  theme(axis.text.x = element_text(angle = 45, hjust=1))

patchwork <- (Tn2_plot + P1_plot) / price_plot

patchwork + plot_annotation(
  title = "Annual Distribution of Almond Profit Anomaly",
  caption = "Data Sources: Lobell 2006, USDA/NASS 2019")
```

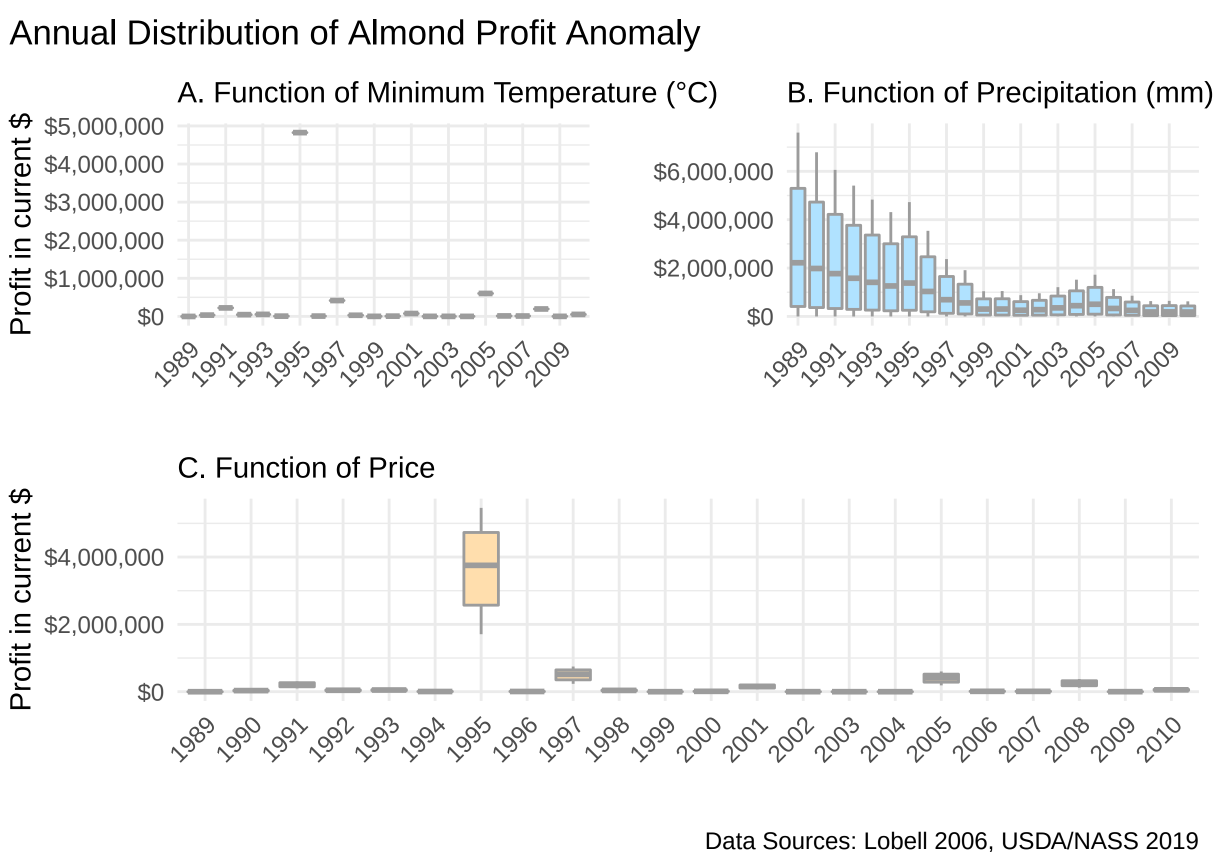

Figure 2: Annual Distribution of Almond Profit Anomaly

For variable `precipitation`, we see a downward trend in variation over the years.

For variable `February average minimum temperatures`, we see much less variation over the dataset.

For variable `price`, we see relatively small variations in profit for most years, in comparison to a wide variation in the year 1995. The fact that 1995 was an abnormally rainy year points to something about the interaction and influence of precipitation and price on the model that leads to high variation at high precipitation levels.

Scatter Plots

```r
meanTn2_plot <- ggplot(Tn2mn_df, aes(x = Tn2, y = mean_netpre)) +
  geom_point(color = "gray61", size = 1) +
  scale_y_continuous(labels=scales::dollar_format()) +
  labs(subtitle = "A. Function of Minimum Temperature",
       x = "Average Feburary Minimum Temperature (°C)",
       y = "Profit in current $") +
  theme_minimal() +
  theme(axis.title.x = element_text(size = 10))

meanP1_plot <- ggplot(P1mn_df, aes(x = P1, y = mean_netpre)) +
  geom_point(color = "lightskyblue1", size = 1) +
  scale_y_continuous(labels=scales::dollar_format()) +
  labs(subtitle = "B. Function of Precipitation",
       x = "Total January Precipitation (mm)", y = " ") +
  theme_minimal() +
  theme(axis.title.x = element_text(size = 10))

price_plot2 <- ggplot(price_df,
       aes(x = as.factor(year), y = netpre, group = year)) +
  geom_point(shape = 21, color = "gray61", fill = "navajowhite", size = 1.2) +
  scale_x_discrete(breaks=seq(1989, 2010, 2)) +
  scale_y_continuous(labels=scales::dollar_format()) +
  labs(subtitle = "C. Function of Price",
       x = "Year", y = "Profit in current $") +
  theme_minimal() +
  theme(axis.text.x = element_text(angle = 45, hjust=1, size = 10))

meanprice_plot <- ggplot(pricemn_df, aes(x = price, y = mean_netpre)) +
  geom_point(color = "grey40", size = 1.2) +
  scale_x_continuous(labels=scales::dollar_format()) +
  scale_y_continuous(labels=scales::dollar_format()) +
  labs(subtitle = "D. Function of Mean Price",
       x = "Price per ton",
       y = " ") +
  theme_minimal() +
  theme(axis.title.x = element_text(size = 10))

patchwork <- (meanTn2_plot + meanP1_plot) / (price_plot2 + meanprice_plot)

profit_plot2 <- patchwork + plot_annotation(
  title = "Mean Annual Profit from Almond Profit Anomaly",
  caption = "Data Sources: Lobell 2006, USDA/NASS 2019")

profit_plot2
```

Because the boxplot visualizations did not answer all of our questions about the variables' influence on the model, we plotted the average annual profit anomaly against the uniform distribution of each variable individually. From these visualizations, we found a positive linear relationship of almond price to the average annual profit anomaly, a negative, near-linear relationship of average February temperature lows, and a positive, exponential relationship of January precipitation. This provides an explanation as to why there was much more variation of profit anomaly at a higher precipitation level in 1995.

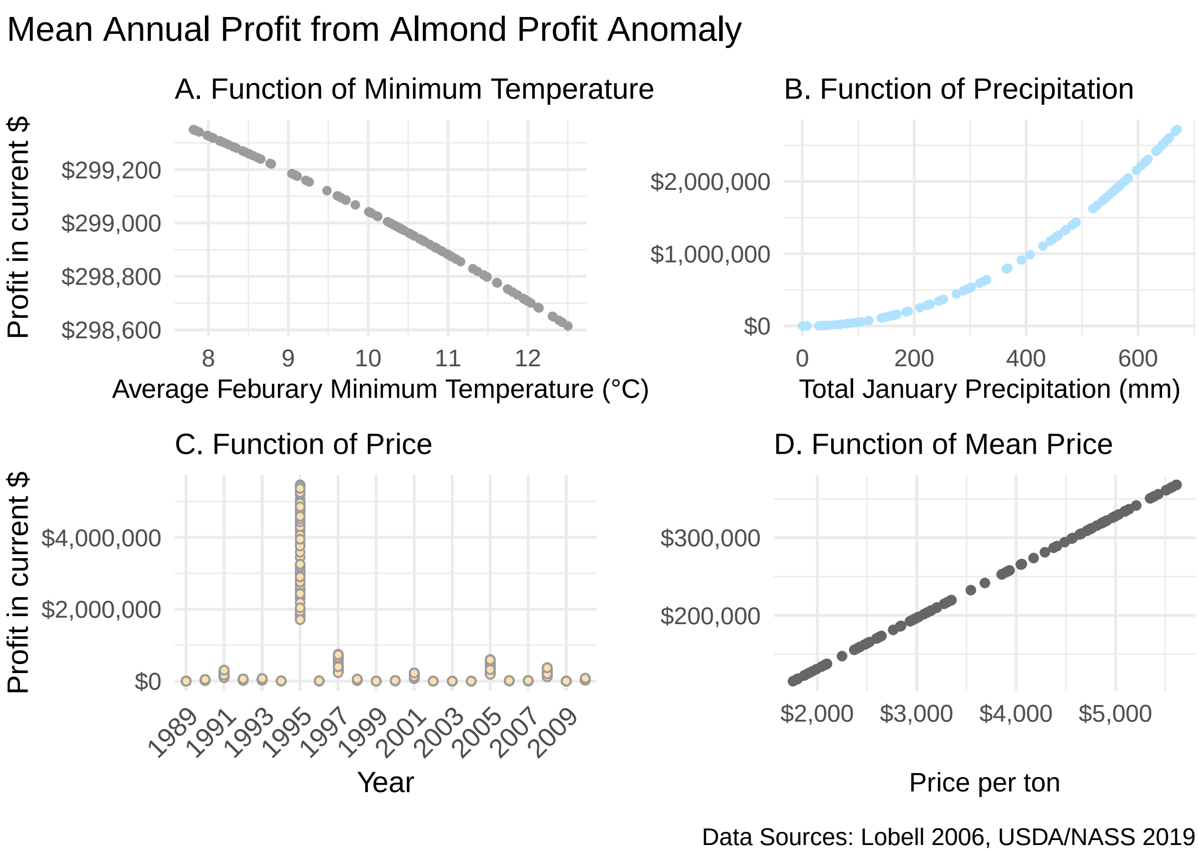

Figure 3: Mean Annual Profit from Almond Profit Anomaly

4. Output the graph as a stand alone image

```
ggsave(file = "meanannualprofit.png", plot = profit_plot2,
  scale = 1, width = 6, height = 6,
  units = "in", dpi = 300)
```

References

Lobell, D., Field, C., Nicholas, K., & Bonfils, C. (2006). Impacts of future climate change on California perennial crop yields: Model projections with climate and crop uncertainties. Agricultural and Forest Meteorology, 141, 208–218. https://doi.org/10.1016/j.agrformet.2006.10.006

Zhang, Z., Jin, Y., Chen, B., & Brown, P. (2019). California Almond Yield Prediction at the Orchard Level With a Machine Learning Approach. Frontiers in Plant Science, 10, 809. https://doi.org/10.3389/fpls.2019.00809

USDA/NASS, Pacific Regional Office. (2019). 2019 California Almond Forecast. USDA National Agricultural Statistics Service. www.nass.usda.gov/ca

Repository

Group I Github Repository https://github.com/juliaparish/EDS230_EnviroModeling_AlmondModel