

Assignment 3

Anna Calle and Jamie Miller

4/22/2020

Adjust your almond model to output ONLY the mean almond yield anomaly IF the users sets parameter (e.g mean_only = TRUE))

Perform a sensitivity analysis of how mean anomaly varies ALL of the parameters used in the yield model

Assume parameters are normally distributed with standard deviation of 20% mean value

Rank the parameters in term of their sensitivity

Graph uncertainty in mean yield anomaly across all parameter uncertainty (boxplot and cumulative distribution of the output).

Repeat using the LHS and Sobol methods

Repeat using twice as many parameter sets as you did in your first sensitivity analysis - and look at how this changes the sensitivity results

Submit R markdown and short write up describing what you learned from the sensitivity analysis. Please submit your markdown as an .html or PDF.

```
# read in the input data
clim <- read.table("clim.txt", sep=" ", header=T)

# source function
source("almond_yield.R")

almond_yield(mean_only = TRUE)
```

```
## [1] 181.4453
```

LHS

```
# names of our parameters: a = "Tmincoeff1", b = "Tmincoeff2", c = "Precipcoeff1", d = "Precipcoeff2", e = "Intercept"
factors = c("a", "b", "c", "d", "e")

# type of distributions they arise from
q = c("qnorm", "qnorm", "qnorm", "qnorm", "qnorm")

# parameters mean
a = -0.015
b = -0.0046
c = -0.07
d = 0.0043
e = 0.28

# parameters for those distributions
q.arg = list(list(mean=a, sd=0.2),
             list(mean=b, sd=0.2),
             list(mean=c, sd=0.2),
             list(mean=d, sd=0.2),
             list(mean=e, sd=0.2))

nsets=200
sens_YA = LHS(NULL,factors,nsets,q,q.arg)

# save parameter values
sens.pars = get.data(sens_YA)
head(sens.pars)
```

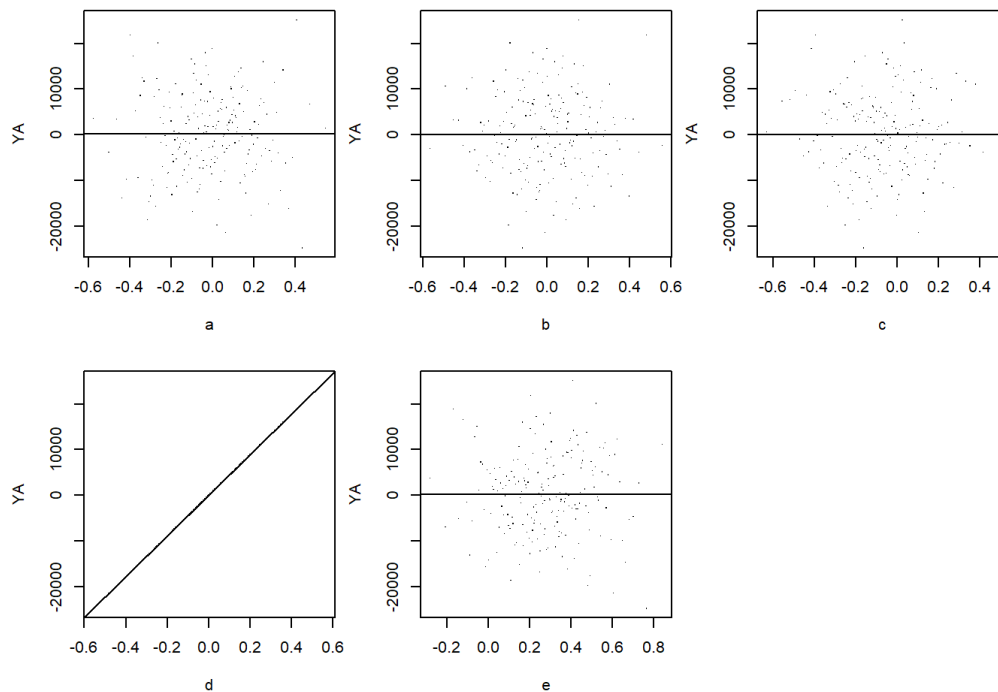
```
##           a           b           c           d           e
## 1 0.11832866 0.047364680 0.3137752 0.01307601 0.4465449
## 2 0.02922374 -0.209754004 -0.1484158 0.06671066 0.6360929
## 3 0.13442106 0.009197392 0.2214843 -0.23836792 0.4395554
## 4 0.35405162 0.084765393 -0.3544181 -0.13857349 0.1204446
## 5 0.36877525 0.158425267 -0.1245618 -0.36475162 0.3834447
## 6 0.12787349 0.339876778 -0.4039185 -0.02970258 0.3503569
```

```
# run model
sens_results= mapply(FUN=almond_yield, a=sens.pars$a, b=sens.pars$b, c=sens.pars$c, d=sens.pars$d, e=sens.pars$e, MoreArgs=1
ist(clim_data = clim, mean_only=TRUE))
head(sens_results)
```

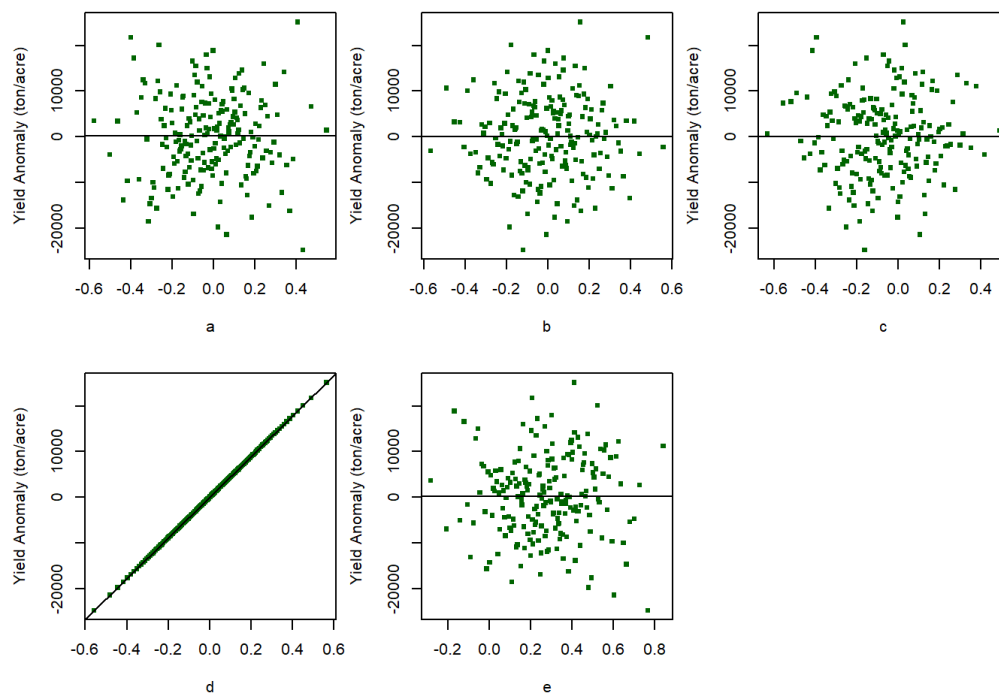
```
## [1] 628.7365 2924.0866 -10557.6349 -6191.0253 -16200.9631 -1337.5602
```

```
# use unlist to get a matrix
sens_res = matrix((unlist(sens_results)), ncol=1, byrow=TRUE)

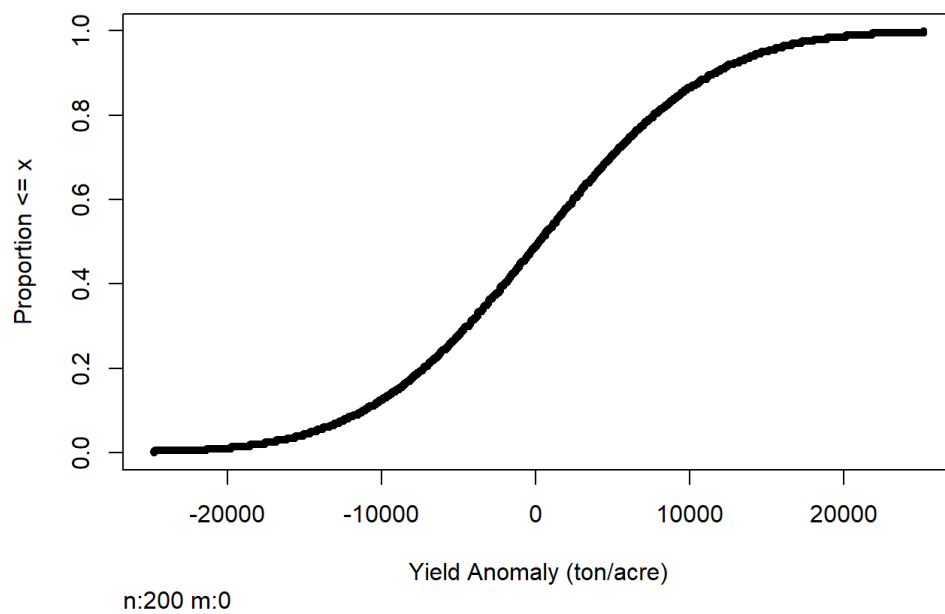
sens_YA = pse::tell(sens_YA, t(sens_res), res.names="YA")
plotscatter(sens_YA)
```



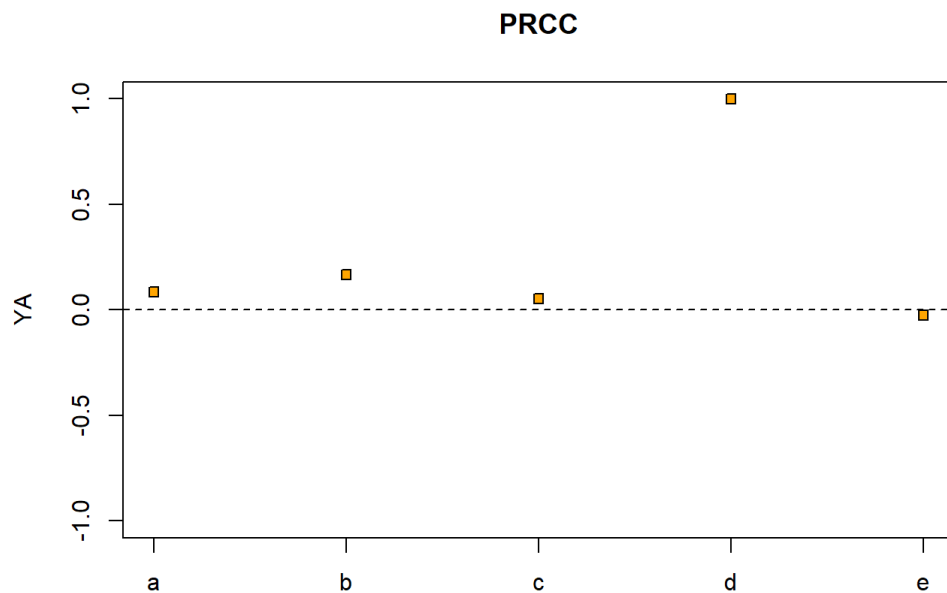
```
# note it can take standard R plotting parameters
plotscatter(sens_YA, col="darkgreen", cex=5, ylab="Yield Anomaly (ton/acre)")
```



```
# whats is the range of results
plotecdf(sens_YA, col="red", lwd=5, xlab="Yield Anomaly (ton/acre)")
```



```
# we can also plot partial correlation coefficients
plotprcc(sens_YA)
```



```
# rank parameters
sens_YA_prcc <- sens_YA$prcc[[1]]$PRCC %>%
  mutate(param = c("a", "b", "c", "d", "e")) %>%
  arrange(-(abs(original)))

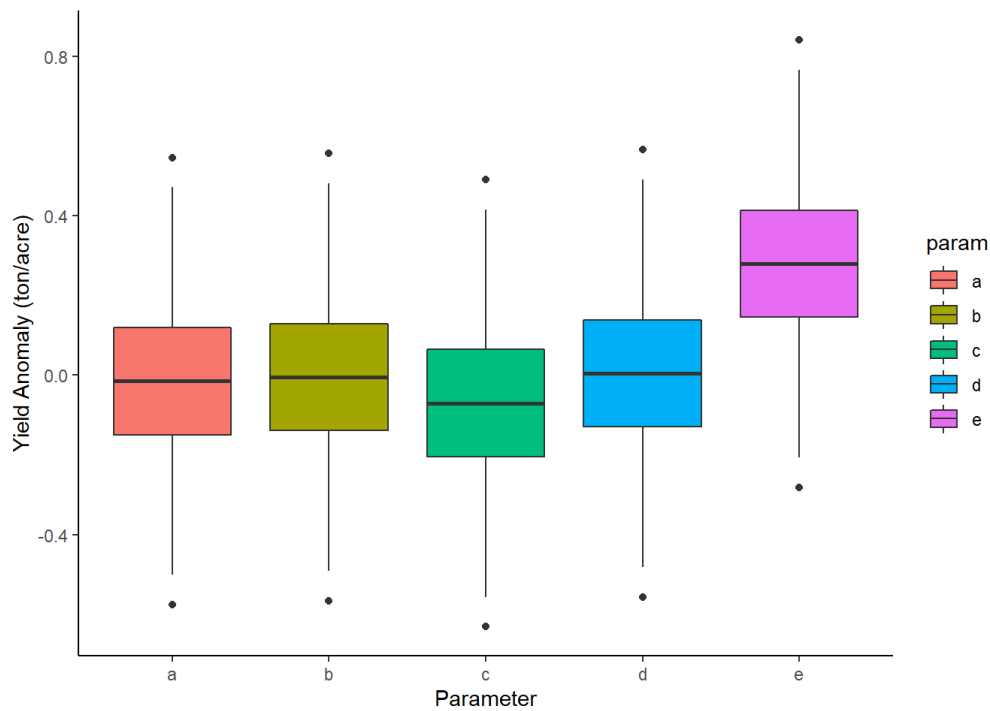
colnames(sens_YA_prcc) <- c("PRCC", "Parameter")
sens_YA_prcc
```

```
##      PRCC Parameter
## 1  0.99999855      d
## 2  0.16597958      b
## 3  0.08438616      a
## 4  0.05145218      c
## 5 -0.02661056      e
```

```
# boxplot
sens_YA_df <- sens_YA$data %>%
  gather(param, YA) %>%
  mutate(method = "LHS")

lhs_boxplot <- ggplot(sens_YA_df, aes(param, YA, fill=param)) +
  geom_boxplot() +
  labs(y="Yield Anomaly (ton/acre)", x = "Parameter") +
  theme_classic()

lhs_boxplot
```



Sobol

```
# names of our parameters
# number of parameters

np=200
a = rnorm(mean=a, sd=0.2, n=np)
b = rnorm(mean=b, sd=0.2, n=np)
c = rnorm(mean=c, sd=0.2, n=np)
d = rnorm(mean=d, sd=0.2, n=np)
e = rnorm(mean=e, sd=0.2, n=np)

# generate two examples of random number from parameter distributions

X1 = cbind.data.frame(a, b, c, d, e)

# repeat sampling
a = rnorm(mean=a, sd=0.2, n=np)
b = rnorm(mean=b, sd=0.2, n=np)
c = rnorm(mean=c, sd=0.2, n=np)
d = rnorm(mean=d, sd=0.2, n=np)
e = rnorm(mean=e, sd=0.2, n=np)

X2 = cbind.data.frame(a, b, c, d, e)

sens_YA_sobel = sobol2007(model = NULL, X1, X2, nboot = 100)

# run model for all parameter sets
res = mapply(FUN=almond_yield, a=sens_YA_sobel$X$a,
             b=sens_YA_sobel$X$b,
             c=sens_YA_sobel$X$c,
             d=sens_YA_sobel$X$d,
             e=sens_YA_sobel$X$e,
             MoreArgs=list(clim_data = clim, mean_only=TRUE))

sens_YA_sobel = sensitivity::tell(sens_YA_sobel,res, res.names="YA")

# first-order indices (main effect without co-variance)
sens_YA_sobel$S
```

```
##          original          bias  std. error    min. c.i.    max. c.i.
## a -9.728685e-06 -2.094582e-07 1.865329e-05 -4.875833e-05 2.345151e-05
## b  1.180416e-04  2.400955e-05 2.469925e-04 -3.618497e-04 6.514463e-04
## c -6.607578e-05  2.668393e-06 2.863918e-04 -7.009011e-04 5.234950e-04
## d  7.643375e-01  9.043822e-03 1.314557e-01  4.546345e-01 9.504441e-01
## e  2.752221e-06  1.970130e-07 1.812013e-06 -9.559774e-07 5.982450e-06
```

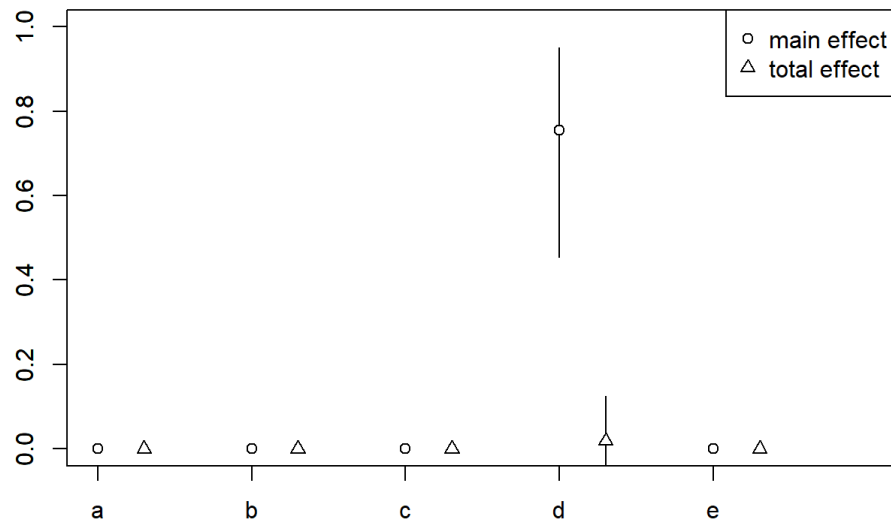
```
# total sensitivity index -note that this partitions the output variance - so values sum to 1
sens_YA_sobel$T
```

```
##          original          bias  std. error    min. c.i.    max. c.i.
## a  1.672587e-05  5.615473e-07 1.565104e-05 -1.697945e-05 4.694850e-05
## b -1.493057e-04 -2.179021e-05 1.520538e-04 -5.185836e-04 1.360672e-04
## c  2.285977e-04  1.879836e-05 2.084297e-04 -2.116282e-04 5.707380e-04
## d  1.764738e-02 -8.905817e-04 6.151947e-02 -1.172960e-01 1.255255e-01
## e -2.053647e-06  6.587798e-08 1.316474e-06 -4.900971e-06 2.705311e-07
```

```
# The difference between the main effect and total effect can tell us something about how the parameter influences results
# so in the main effect we include interactions with other parameters
print(sens_YA_sobel)
```

```
##
## Call:
## sobol2007(model = NULL, X1 = X1, X2 = X2, nboot = 100)
##
## Model runs: 1400
##
## First order indices:
##          original          bias  std. error    min. c.i.    max. c.i.
## a -9.728685e-06 -2.094582e-07 1.865329e-05 -4.875833e-05 2.345151e-05
## b  1.180416e-04  2.400955e-05 2.469925e-04 -3.618497e-04 6.514463e-04
## c -6.607578e-05  2.668393e-06 2.863918e-04 -7.009011e-04 5.234950e-04
## d  7.643375e-01  9.043822e-03 1.314557e-01  4.546345e-01 9.504441e-01
## e  2.752221e-06  1.970130e-07 1.812013e-06 -9.559774e-07 5.982450e-06
##
## Total indices:
##          original          bias  std. error    min. c.i.    max. c.i.
## a  1.672587e-05  5.615473e-07 1.565104e-05 -1.697945e-05 4.694850e-05
## b -1.493057e-04 -2.179021e-05 1.520538e-04 -5.185836e-04 1.360672e-04
## c  2.285977e-04  1.879836e-05 2.084297e-04 -2.116282e-04 5.707380e-04
## d  1.764738e-02 -8.905817e-04 6.151947e-02 -1.172960e-01 1.255255e-01
## e -2.053647e-06  6.587798e-08 1.316474e-06 -4.900971e-06 2.705311e-07
```

```
plot(sens_YA_sobel)
```



```
# compare with LHS and PRCC
sens_YA$prcc
```

```
## [[1]]
##
## Call:
## pcc.default(X = L, y = r, rank = T, nboot = nboot)
##
## Partial Rank Correlation Coefficients (PRCC):
##      original
## a  0.08438616
## b  0.16597958
## c  0.05145218
## d  0.99999855
## e -0.02661056
```

```
sens_YA_sobel$S
```

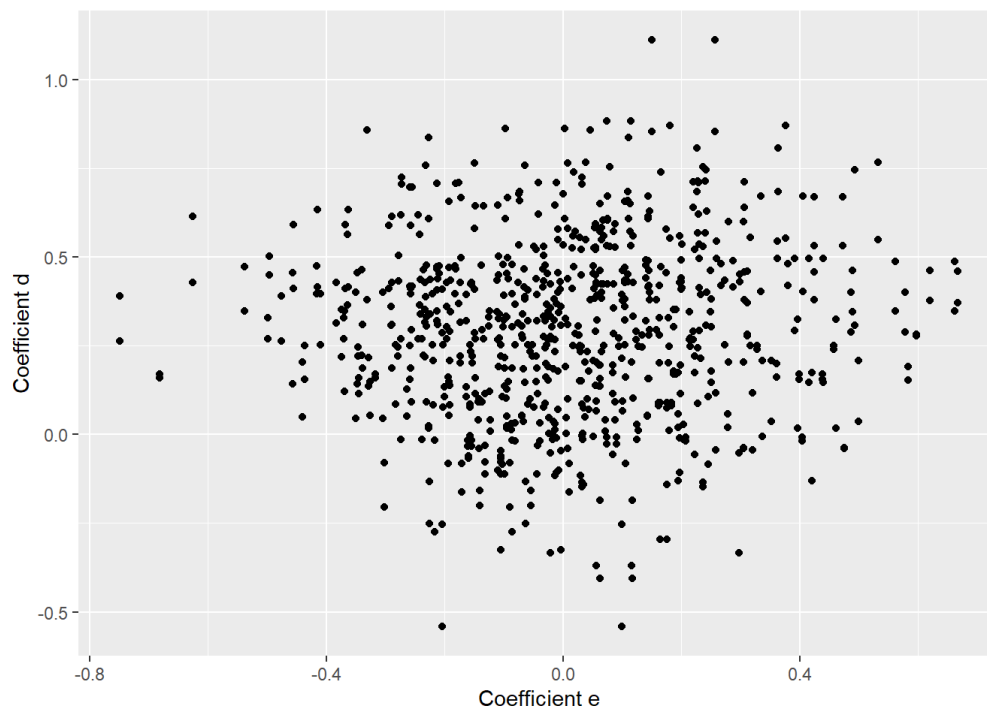
```
##      original      bias std. error  min. c.i.  max. c.i.
## a -9.728685e-06 -2.094582e-07 1.865329e-05 -4.875833e-05 2.345151e-05
## b  1.180416e-04  2.400955e-05 2.469925e-04 -3.618497e-04 6.514463e-04
## c -6.607578e-05  2.668393e-06 2.863918e-04 -7.009011e-04 5.234950e-04
## d  7.643375e-01  9.043822e-03 1.314557e-01  4.546345e-01 9.504441e-01
## e  2.752221e-06  1.970130e-07 1.812013e-06 -9.559774e-07 5.982450e-06
```

```
sens_YA_sobel$T
```

```
##      original      bias std. error  min. c.i.  max. c.i.
## a  1.672587e-05  5.615473e-07 1.565104e-05 -1.697945e-05 4.694850e-05
## b -1.493057e-04 -2.179021e-05 1.520538e-04 -5.185836e-04 1.360672e-04
## c  2.285977e-04  1.879836e-05 2.084297e-04 -2.116282e-04 5.707380e-04
## d  1.764738e-02 -8.905817e-04 6.151947e-02 -1.172960e-01 1.255255e-01
## e -2.053647e-06  6.587798e-08 1.316474e-06 -4.900971e-06 2.705311e-07
```

```
# make a data frame for plotting
both = cbind.data.frame(sens_YA_sobel$X, gs=sens_YA_sobel$y)

# Look at response of conductance to the two most important variables
ggplot(both, aes(d, e))+geom_point()+labs(y="Coefficient d", x="Coefficient e")
```



```
# rank S
rank_S <- sens_YA_sobel$S %>%
  select(original) %>%
  mutate(param = c("a", "b", "c", "d", "e")) %>%
  arrange(-(abs(original)))

colnames(rank_S) <- c("PRCC", "Parameter")
rank_S
```

```
##          PRCC Parameter
## 1  7.643375e-01      d
## 2  1.180416e-04      b
## 3 -6.607578e-05      c
## 4 -9.728685e-06      a
## 5  2.752221e-06      e
```

```
# rank T
rank_T <- sens_YA_sobel$T %>%
  select(original) %>%
  mutate(param = c("a", "b", "c", "d", "e")) %>%
  arrange(-(abs(original)))

colnames(rank_T) <- c("PRCC", "Parameter")
rank_T
```

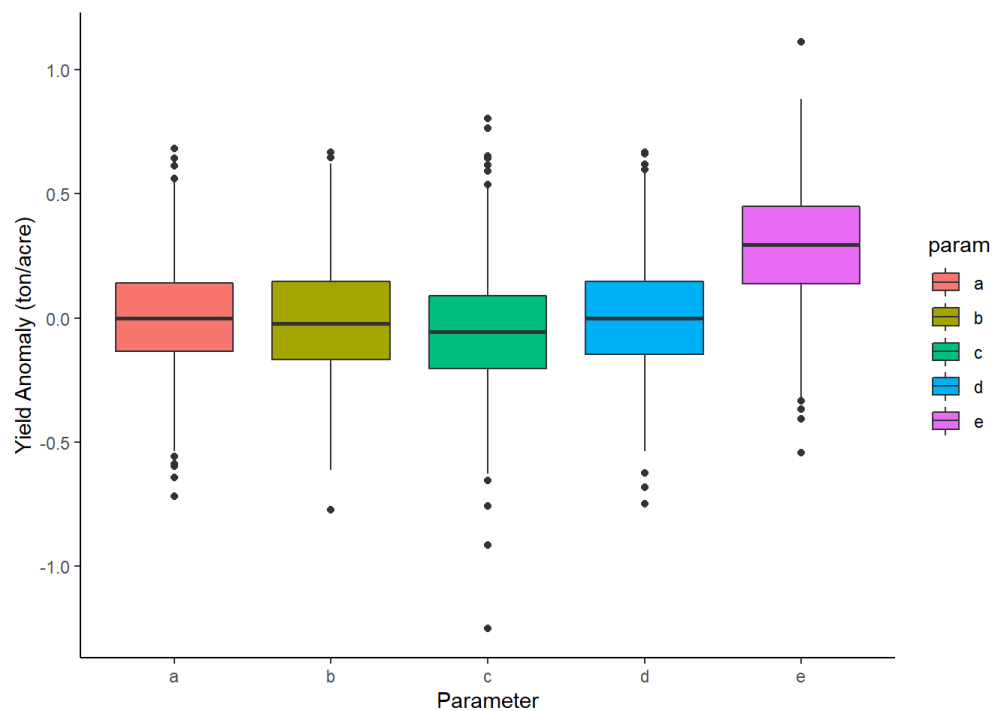
```
##          PRCC Parameter
## 1  1.764738e-02      d
## 2  2.285977e-04      c
## 3 -1.493057e-04      b
## 4  1.672587e-05      a
## 5 -2.053647e-06      e
```



```
# boxplot, comparing uncertainty of estimates
sens_YA_sobel_df <- sens_YA_sobel$X %>%
  gather(param, YA) %>%
  mutate(method = "Sobol")

sobel_boxplot <- ggplot(sens_YA_sobel_df, aes(param, YA, fill=param)) +
  geom_boxplot() +
  labs(y="Yield Anomaly (ton/acre)", x = "Parameter") +
  theme_classic()

sobel_boxplot
```

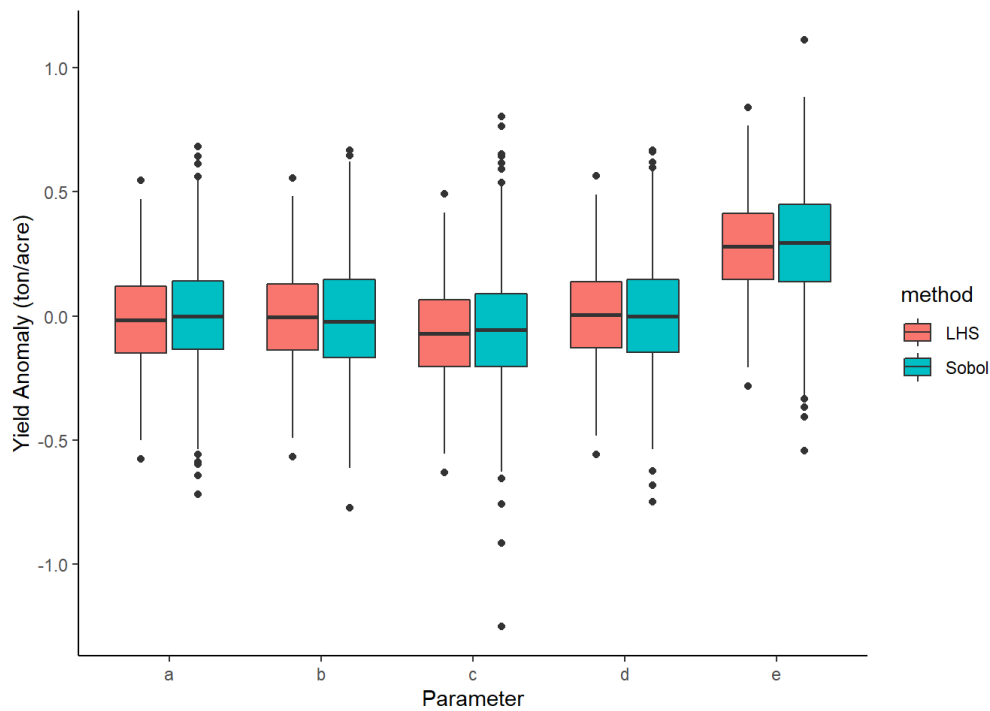


Comparing boxplots

```
# merging results for both methods
sens_both_df <- rbind(sens_YA_df, sens_YA_sobel_df)

# boxplot, comparing uncertainty of estimates
both_boxplot <- ggplot(sens_both_df, aes(x=param, y=YA, fill=method)) +
  geom_boxplot() +
  labs(y="Yield Anomaly (ton/acre)", x = "Parameter") +
  theme_classic()

both_boxplot
```



There was no significant difference in model sensitivities between the LHS and Sobol methods. Almond yield was the most sensitive to changes in variable D (Precipcoeff2). The partial rank regression coefficients (PRCC) were highest for parameter D, and lowest for parameter E. This effect was nearly identical between the LHS and Sobol methods. Parameter D is the squared value of the amount of January precipitation in the current harvest year. It makes sense that almond yield would be the most sensitive to this parameter because winter rainfall can have a significant impact on plant productivity in the current year. Prior year precipitation will impact individual plant growth from the prior season so may give a plant more branches on which to grow fruit. However, current season precipitation dictates how much energy that plant can dedicate to fruit production right now. Minimum temperatures can have significant impacts on plant productivity if there is a hard freeze, but most these event are not common and most farmers have management strategies to mitigate sub-optimal temperatures.