Sensitivity Analysis - Latin Hypercube Sampling

Group H - Paloma Cartwright, Juliet Cohen, Julia Parish

2022-04-26

Sensitivity Analysis - Latin Hypercube Sampling

This environmental model 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 code a function to compute atmosphere conductance and to conduct a formal sensitivity analysis using the Latin Hypercube Sampling (LHS) random sampling method. This assignment focuses on developing skills to create a atmospheric conductance model function, utilize the Latin Hypercube Sampling to generate near random sample of parameter values, and then plot the atmospheric conductance values.

Load Libraries

```
library(here)
library(SciViews) # ln() function
library(pse)
library(purrr)
library(ppcor)
library(kableExtra)
```

1. Code a function to compute atmospheric conductance

```
source(here("R/atmcon.R"))
```

2. Run `atmcon` model and provide a single estimate of atmospheric conductance for this forest.

```
ac_forest <- atmcon(vm = 250, h = 1000)
# ac_forest

ac_forest_rounded <- round(ac_forest[[1]], 2)

print(paste0("The atmospheric conductance for this vegetation is ",
             ac_forest_rounded, " centimeters per second."))
```

```
## [1] "The atmospheric conductance for this vegetation is 15.44 centimeters per second."
```

3. Conduct a sensitivity analysis

Consider the sensitivity of estimates to uncertainty in the following parameters and inputs:

- *vm*
- *h*
- *kd*
- *k0*

Windspeeds `vm` are normally distributed with a `mean` of `250 cm/sec` with a `standard deviation` of `30 cm/sec`.

For vegetation height, `h`, assume that height is somewhere between `9.5` and `10.5 m` (but any value in that range is equally likely).

The typical values of `kd` is 0.7 and `k0` is 0.1. Assume that they are normally distributed with standard deviation of 1% of their default values.

3.A. Use LHS to generate parameter values for the 4 parameters

```
# define 4 parameters to test
factors = c("vm", "h", "kd", "k")

vm_mean = 250 # units = centimeters / second
vm_sd = 30

# convert units of meters to centimeters
h_min = 9.5 * 100
h_max = 10.5 * 100

# static values
kd_value = 0.7
kd_sd = 0.01

k0_value = 0.1
k0_sd = 0.01

#define sample size
nsets = 100

# set distributions for defined parameters
# qnorm=windspeed, qunif=height, qnorm=k, qnorm=k0
q = c("qnorm", "qunif", "qnorm", "qnorm")

q.arg = list(list(mean = vm_mean, sd = vm_sd), #windspeed
             list(min = h_min, max = h_max), #height
             list(mean = kd_value, sd = kd_sd), #kd
             list(mean = k0_value, sd = k0_sd)) #k0
```

```r
# run LHS and generate samples
sens_ac = LHS(NULL, factors, nsets, q, q.arg) # NULL indicates there is no model

#sens_ac
#summary(sens_ac)
#sens_ac$data

sens_pars <- get.data(sens_ac)

sens_pars_head <- head(sens_pars)

sens_pars_table <- kable(sens_pars_head,
                         caption = "Subsample of Sample Parameter Values") %>%
  kable_styling(latex_options = "HOLD_position")

sens_pars_table
```

Table 1: Subsample of Sample Parameter Values

vm	h	kd	k
276.8942	981.5	0.6921081	0.0840181
228.3256	969.5	0.7056805	0.1026631
210.6826	1045.5	0.7062801	0.0742417
291.1661	980.5	0.6978530	0.0937199
293.1859	951.5	0.6840181	0.1045376
242.7872	958.5	0.6902589	0.0782991

3.B. Run the atmospheric conductance model, `atmcon`, for LHS derived parameters and return aerodynamic conductances

See Table 2.

```r
source(here("R/atmcon.R"))

ac_lhs <- pmap(sens_pars, atmcon)

atmospheric_conductances <- ac_lhs %>%
  map_dfr(`[`, "ac")

atmcon_head <- head(atmospheric_conductances)

atmcon_head_table <- kable(atmcon_head,
                           caption = "Subsample of Atmospheric Conductances") %>%
  kable_styling(latex_options = "HOLD_position")

atmcon_head_table
```

Table 2: Subsample of Atmospheric Conductances

ac
13.573586
14.553164
9.569449
16.379574
17.961076
10.872383

3.C. Plot conductance estimates in a way that accounts for parameter uncertainty

```
plot1 <- ggplot(data = atmospheric_conductances, aes(y = ac)) +
  geom_boxplot(fill='seashell1', color="grey58") +
  geom_text(aes(x = 0.01, y = 25, label = "24.77"), stat = "unique",
            size = 3, color = "slategrey") +
  labs(y = "Atmospheric Conductance (cm/s)",
       title = "Atmospheric Conductance - Forest Parameters",
       subtitle = "Variation in Estimated Conductance Outputs",
       caption = "Data: Values modeled based on parameters estimated using LHS.") +
  theme_minimal() +
  theme(axis.title.x = element_blank(),
        axis.text.x = element_blank(),
        axis.ticks.x = element_blank())
plot1
```

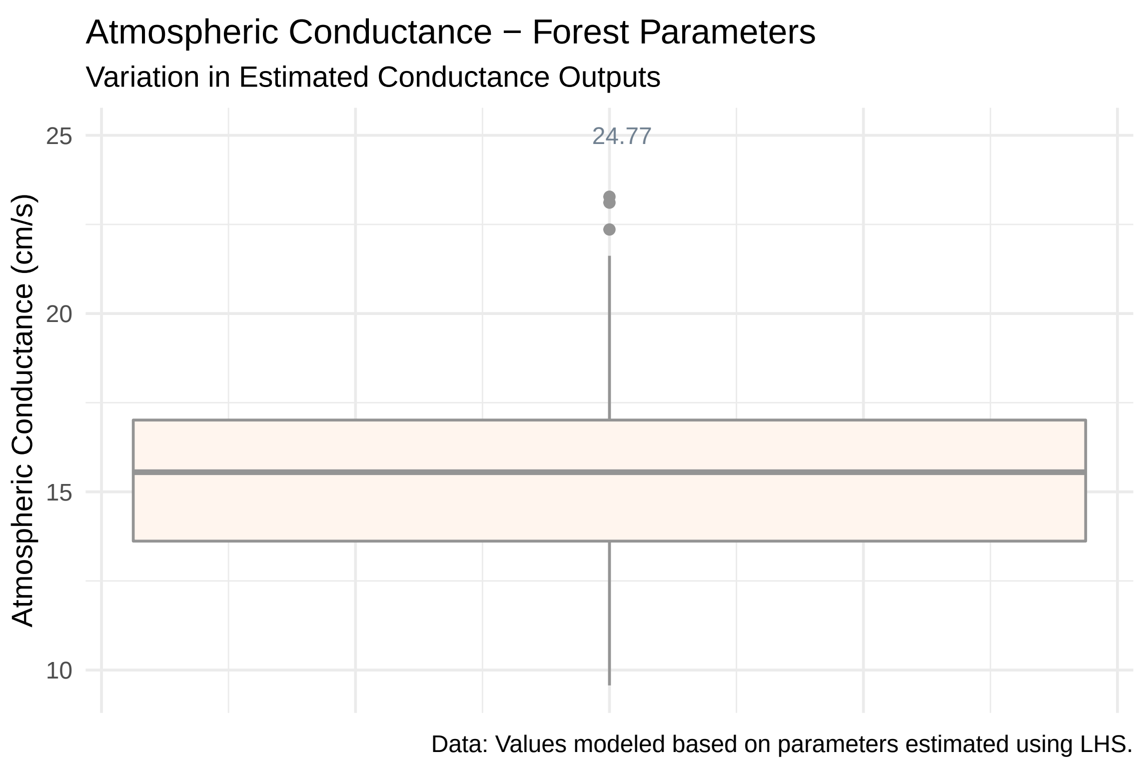

Figure 1: This boxplot shows distribution of atmospheric conducances with given estimated forest parameters

3.D. Plot conductance estimates against each of your parameters

```
# link LHS object (sens_conductance) to outputs
sens_conductance <- pse::tell(sens_ac,
                              t(atmospheric_conductances),
                              res.names = "Atmospheric Conductance")

# plot conductance vs params
plot2 <- pse::plotscatter(sens_conductance,
                          pch = 21, bg= "slateblue", col = "grey58", cex = 0.9)
```

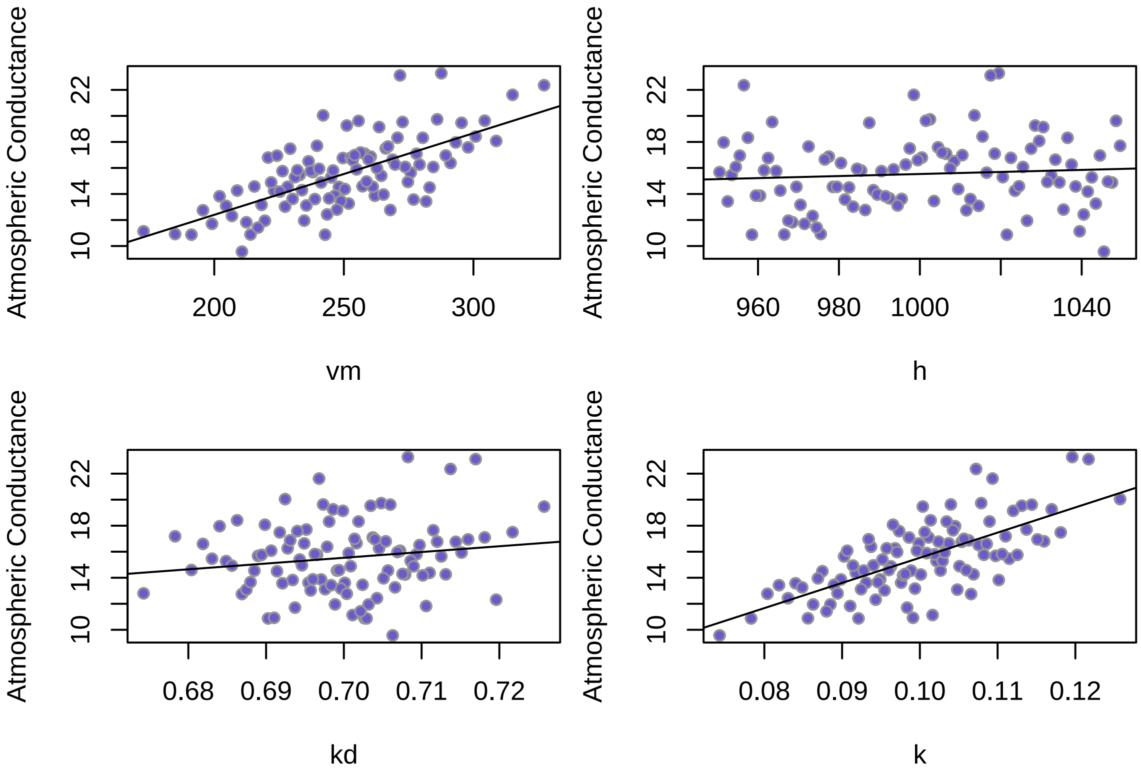

Figure 2: Analyzing atompheric conductance parameter sensitivity. Parameters: vm = windspeed, h = vegetation height, kd = constant, k = constant.

3.E. Estimate the Partial Rank Correlation Coefficients (PRCC)

See Table 3.

```
sens_con_df <- data.frame(sens_conductance$data)

# calculate pairwise correlations for each pair of variables
atmcon_prcc <- pcor(sens_con_df) # list

# convert from list to a dataframe, clean names, & subset columns
atmcon_prcc_df <- data.frame(atmcon_prcc) %>%
  janitor::clean_names() %>%
  dplyr::select(estimate_vm:p_value_k)

sens_con_table <- kable(atmcon_prcc_df,
                        caption = "Pairwise PRCC for each pair of variables given the other variables."
  kable_styling(latex_options = "HOLD_position")

sens_con_table
```

Table 3: Pairwise PRCC for each pair of variables given the other variables.

	estimate_vm	estimate_h	estimate_kd	estimate_k	p_value_vm	p_value_h	p_value_kd	p_value_k
vm	1.0000000	0.0000686	-3.40e-06	0.0002567	0.0000000	0.9994654	0.9999738	0.9979982
h	0.0000686	1.0000000	4.99e-05	-0.0002167	0.9994654	0.0000000	0.9996107	0.9983101
kd	-0.0000034	0.0000499	1.00e+00	0.0000238	0.9999738	0.9996107	0.0000000	0.9998142
k	0.0002567	-0.0002167	2.38e-05	1.0000000	0.9979982	0.9983101	0.9998142	0.0000000

```
plot3 <- pse::plotprcc(sens_conductance, col = "slateblue")
```

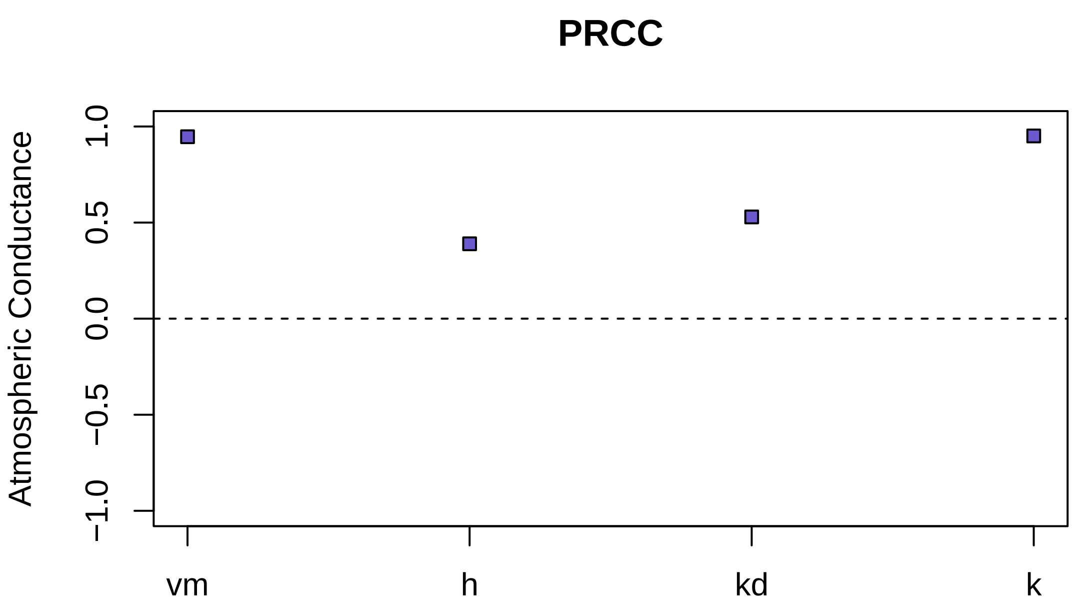

Figure 3: The partial rank correlation coefficient (PRCC) of the basic reproduction in model (atmcon) with respect to other model parameters. For each parameter, the absolute value of its PRCC represents the sensitivity of the parameter - the larger the value is, the more sensitive the parameter is to the corresponding parameter. Parameters: vm = windspeed, h = vegetation height, kd = constant, k = constant.

3.F. Discussion

What do the results tell about how aerodynamic conductance? The results from the scatter plot tells us that `wind speed` (`vm`) and `k` have the greatest impacts on aerodynamic conductance, as shown by the points being closer together and generally following the linear trend line. Parameters `kd` and `h` do not have a strong correlation with aerodynamic conductance, as shown by the points being more broadly scattered and not generally following a trend. Additionally, the partial rank correlation coefficient tells us that `wind speed` and `k` (`k0`) have the greatest impacts on aerodynamic conductance because those values are closest to 1. The other variables, `h` and `kd` have values closer to 0, which indicate they have little impact on aerodynamic conductance.

What does it suggest about what you should focus on if you want to reduce uncertainty in aerodynmaic conductance estimates? In order to reduce uncertainty in aerodynamic conductance estimates, the results discussed above suggest that we should vary `wind speed` (`vm`) and `k` rather than `h` or `kd`.

Does this tell you anything about the sensitivity of plant water use to climate change? These results imply that plant water use will *increase* with climate change, because as climate change becomes more extreme, weather events will become more extreme, such as those that are associated with high wind speeds like hurricanes and rain storms. Therefore, with more wind, the scatter plot shows higher aerodynamic conductance, so plants will require more water to survive moving forward with climate change.