

# Correlations

*Mahalia Clark*

*12/5/2019*

## Correlations

Let's look at the correlations between each variable and PC across the different time lags.

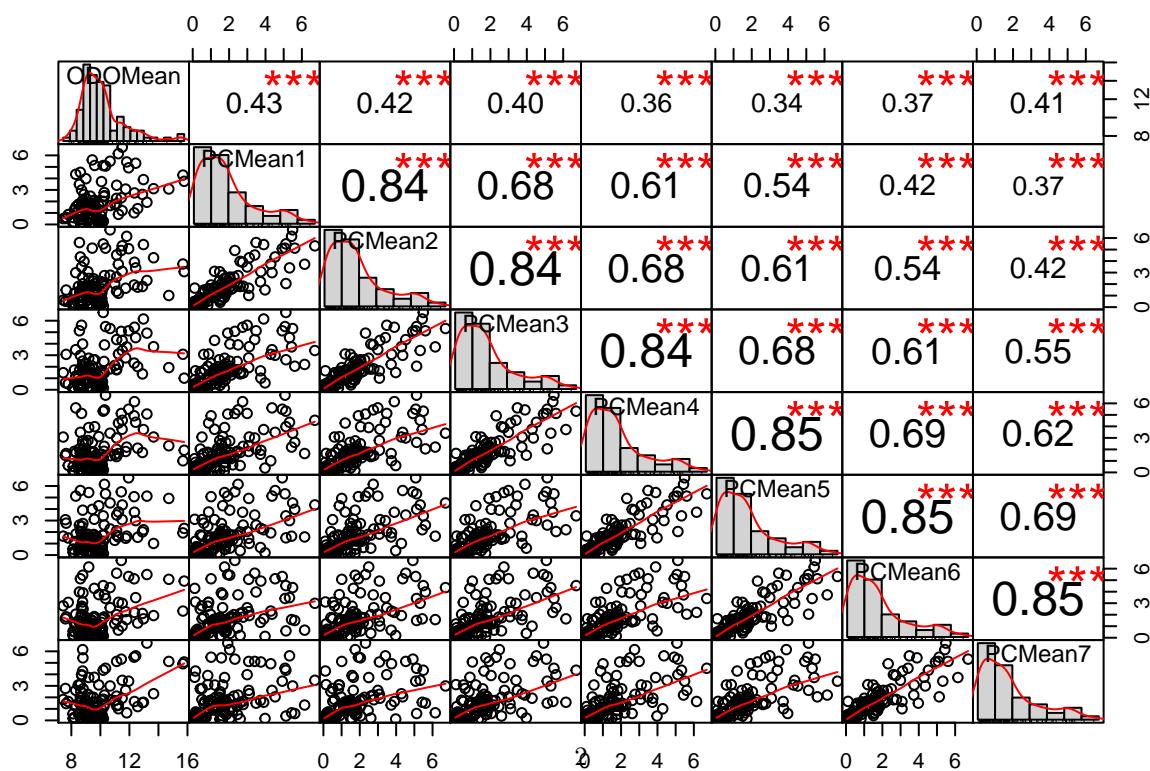
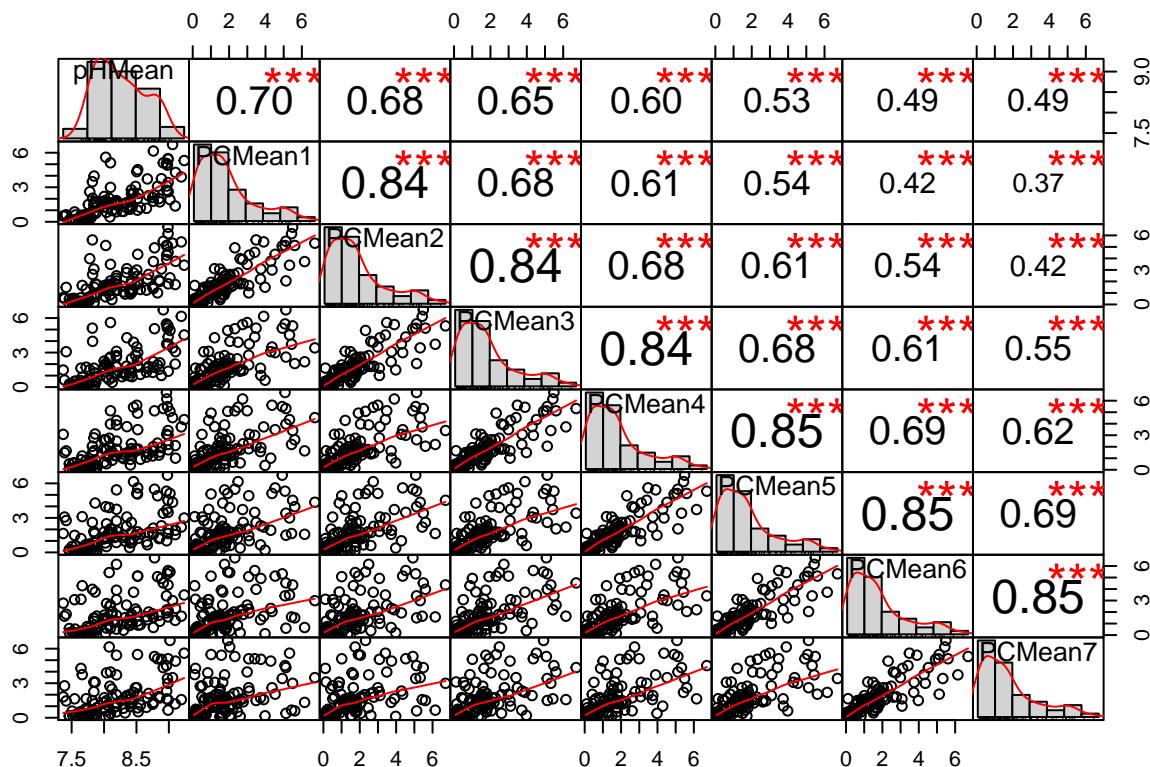
First we will look at potentially confounding variables: pH, ODO, and Chl. These are the ones we think are probably driven by photosynthesis, so by phytoplankton and cyanobacteria. Then we will look at those variables we used to run Feature Importance.

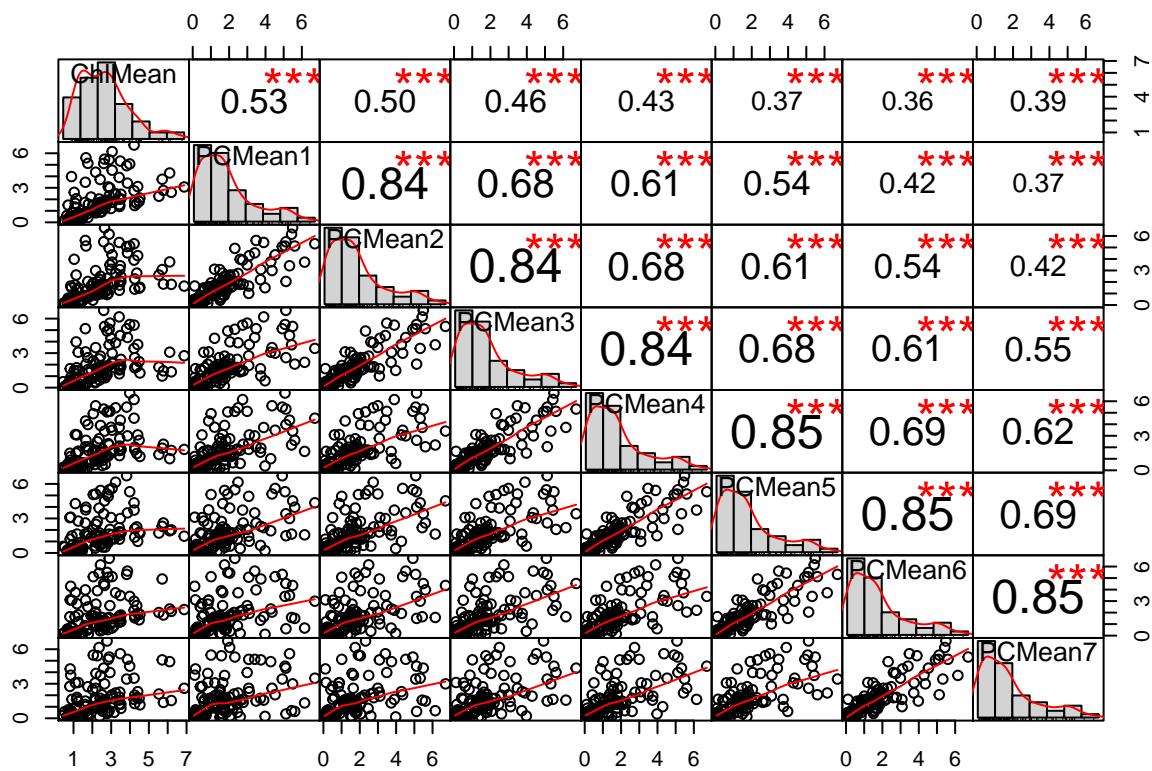
I only know how to make correlation matrices, so we'll do that for now (one for each variable with all 7 time steps of PC). By scanning across the top line of the matrix, we can see how the correlation of PC with that variable changes with increasing time lag.

Obviously, if we only care about the top line, I should probably just figure out how to calculate correlation coefficients and spit them out into a data table, and then I could have a row for each variable and columns for PC1-7 (and beyond!), (or vice versa, swap the rows and columns?) and then we could even graph how the correlation coefficient changes over time (increased lag). But I didn't think of that until I'd already created the rest of this document, so that can be a future/never project.

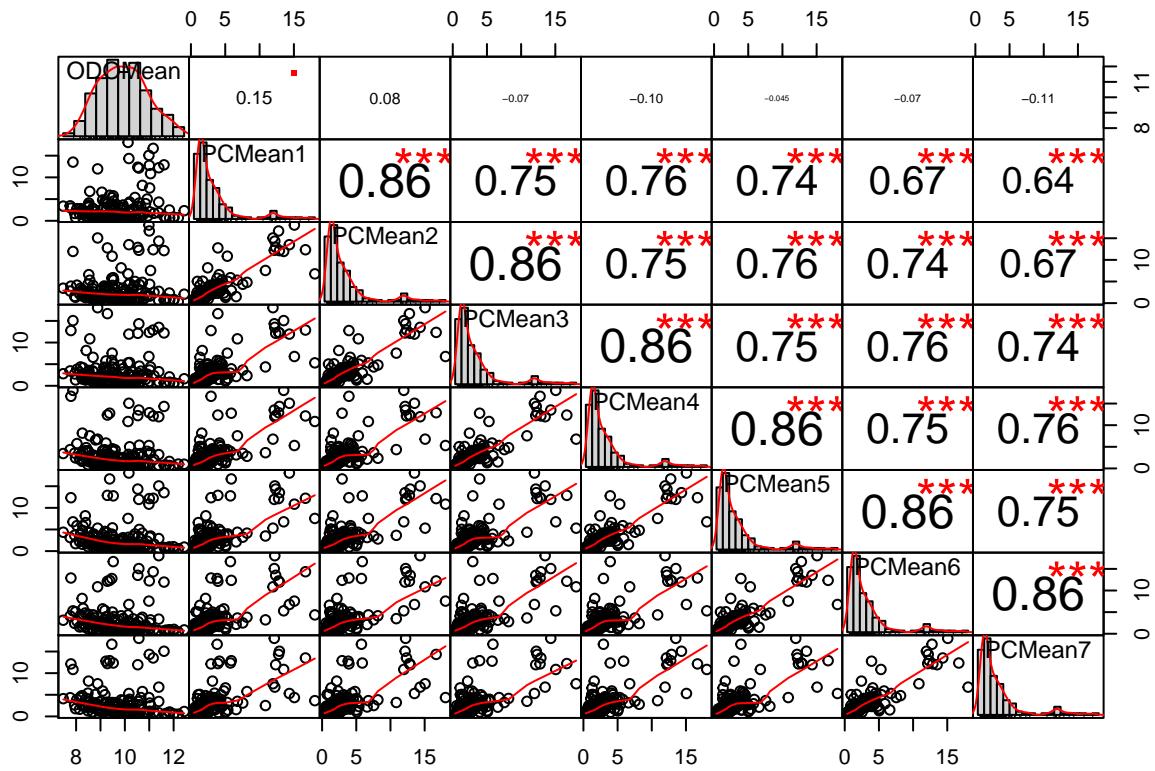
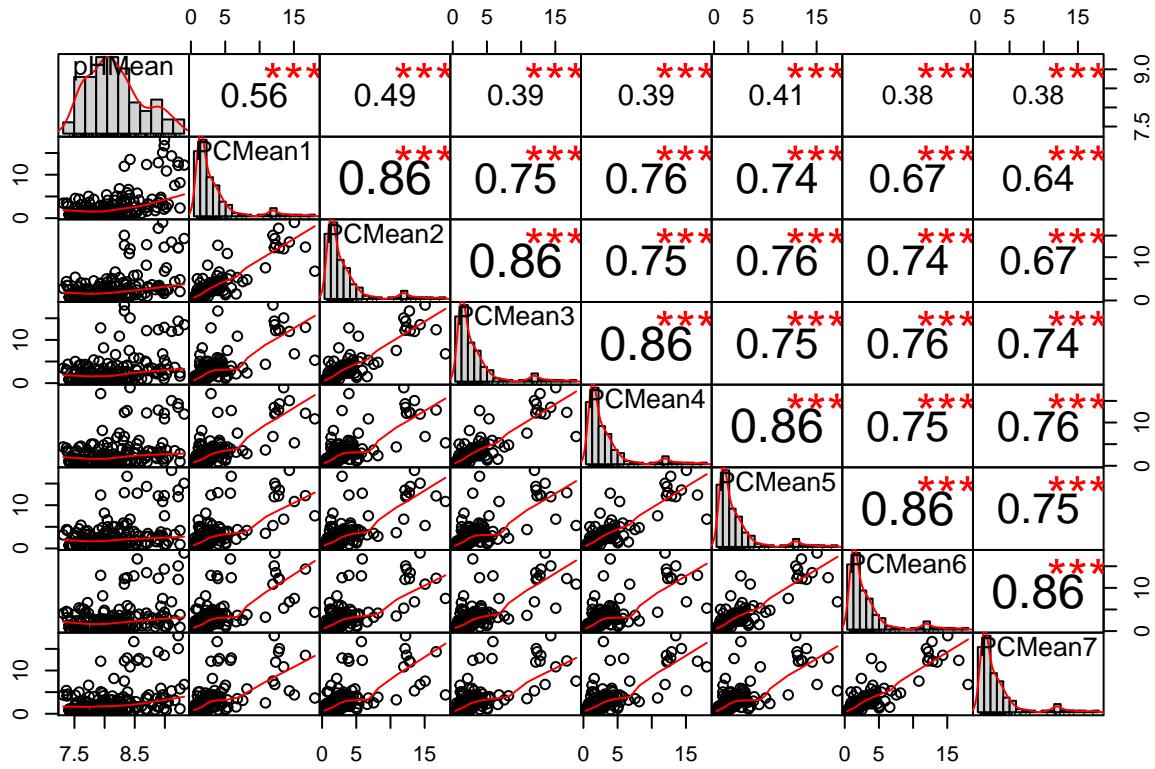
## Confounding Variables

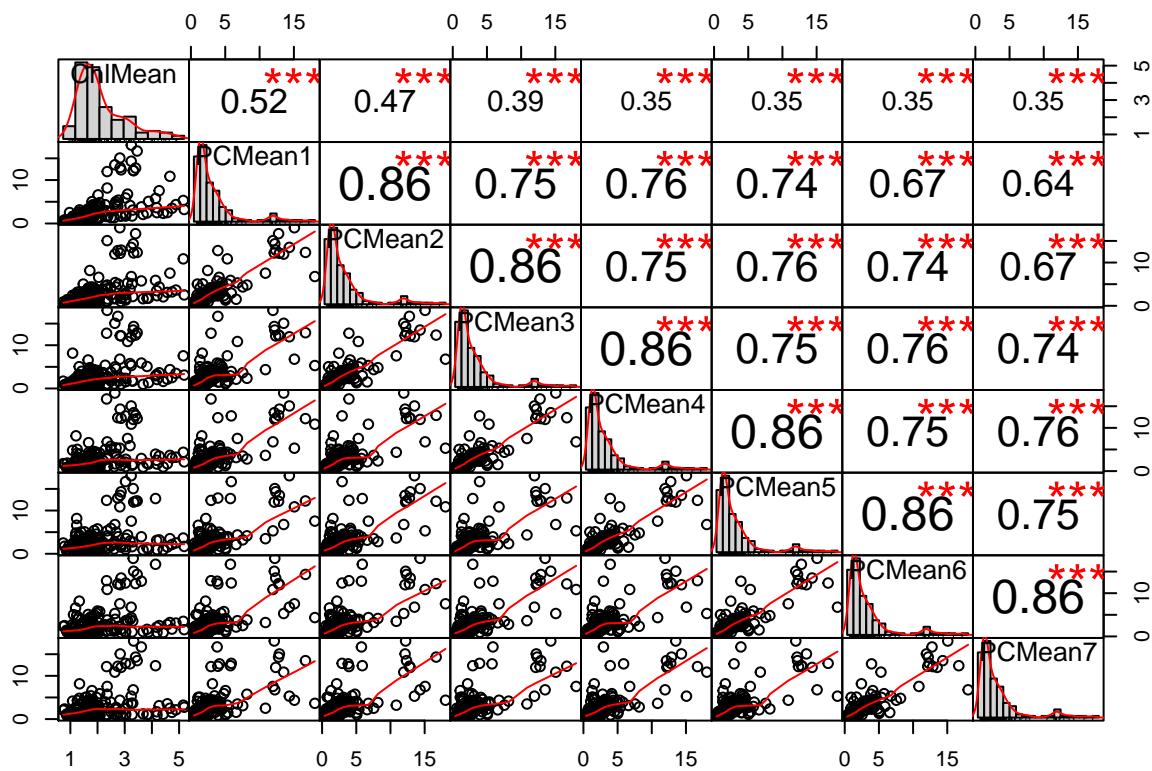
MB17



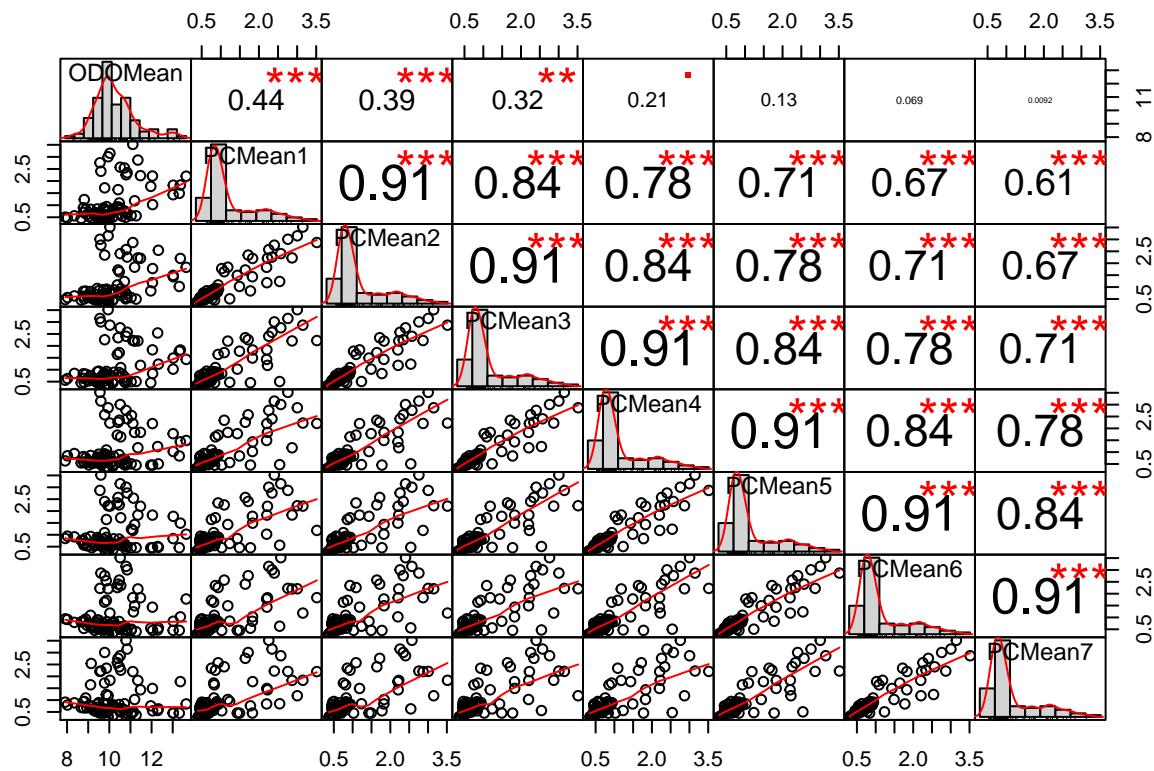
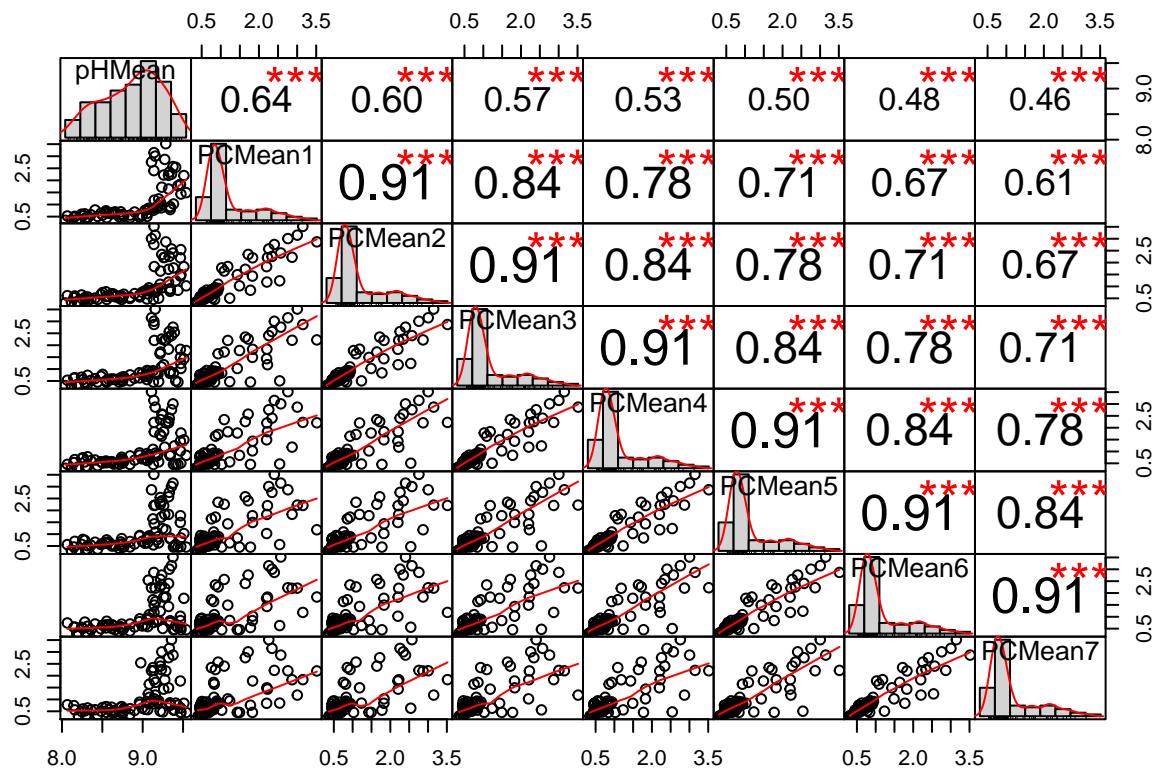


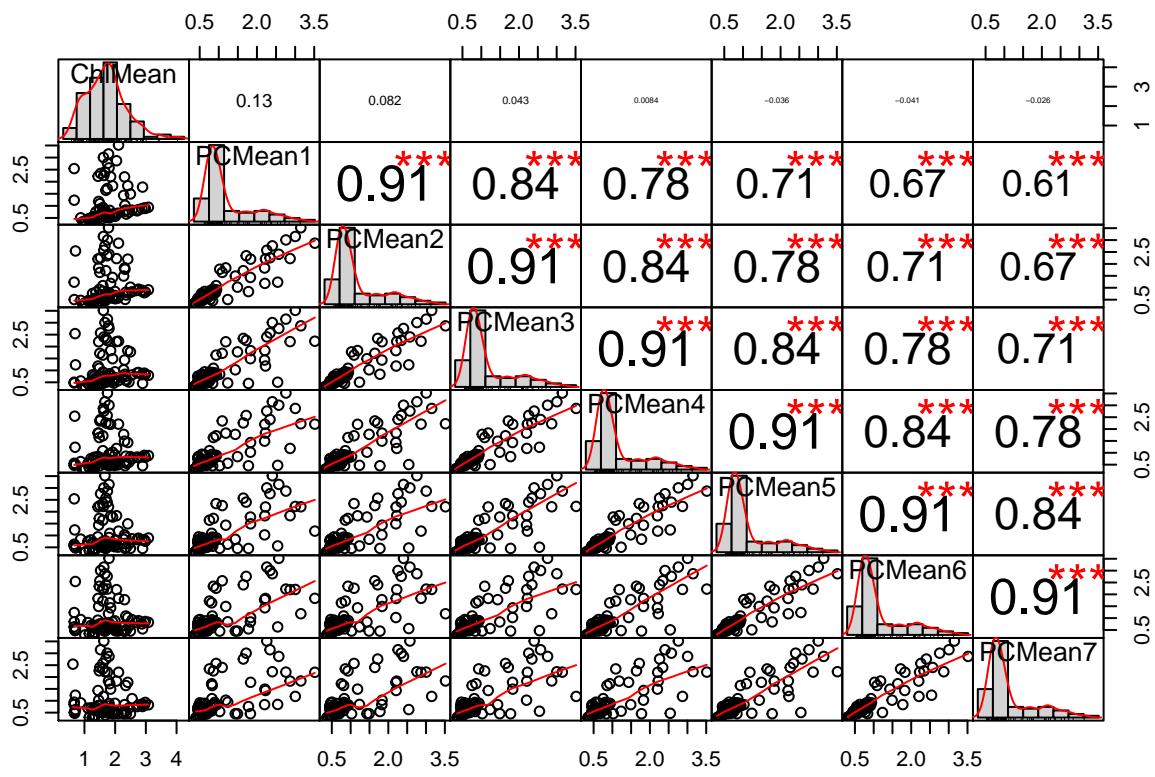
## MB18

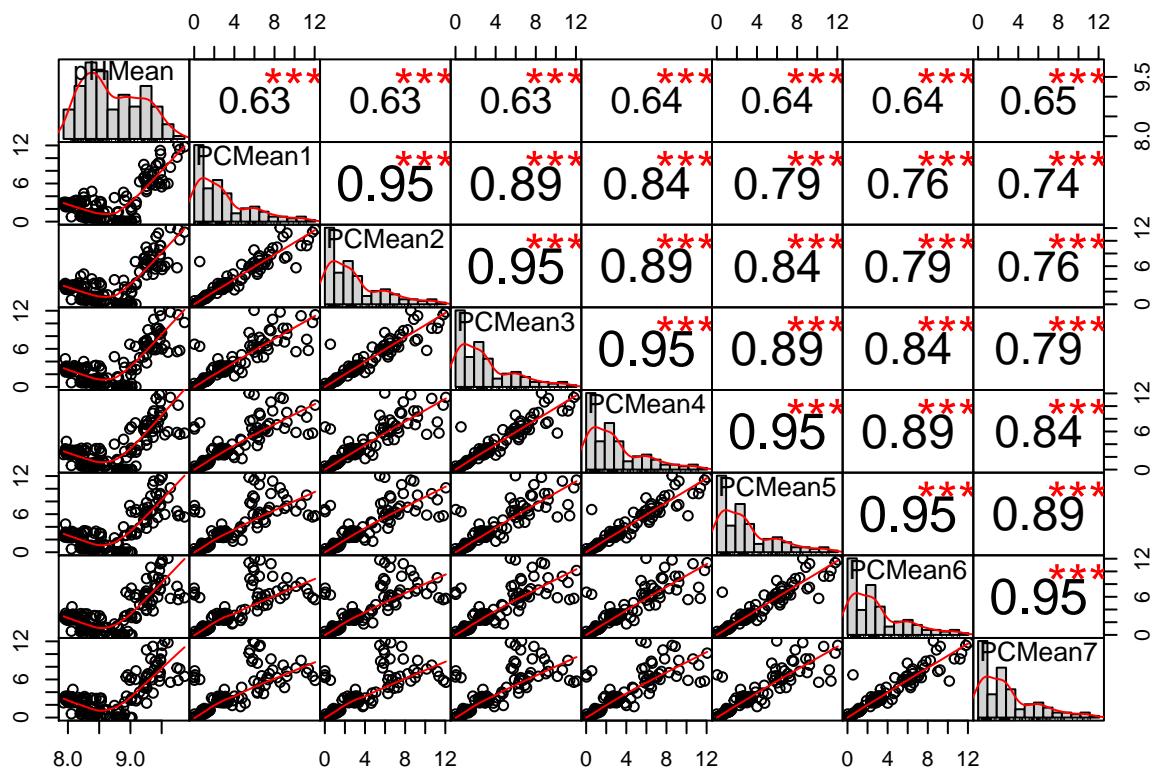




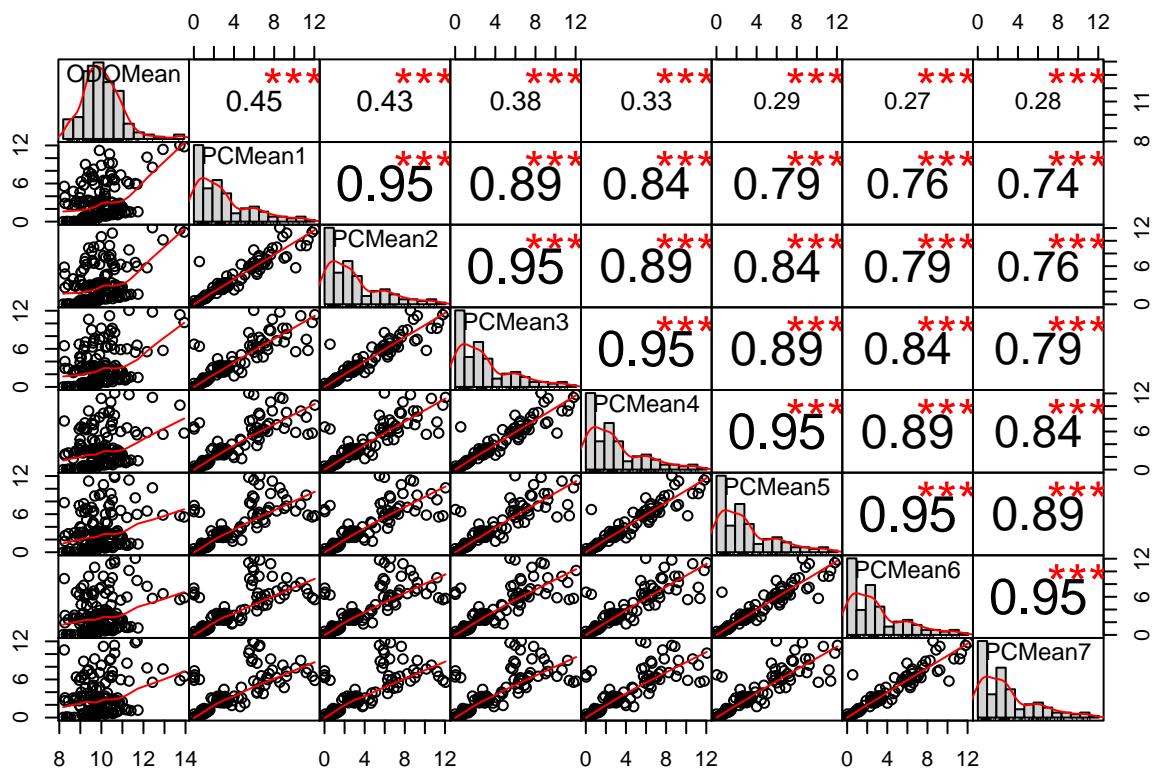
## SA17

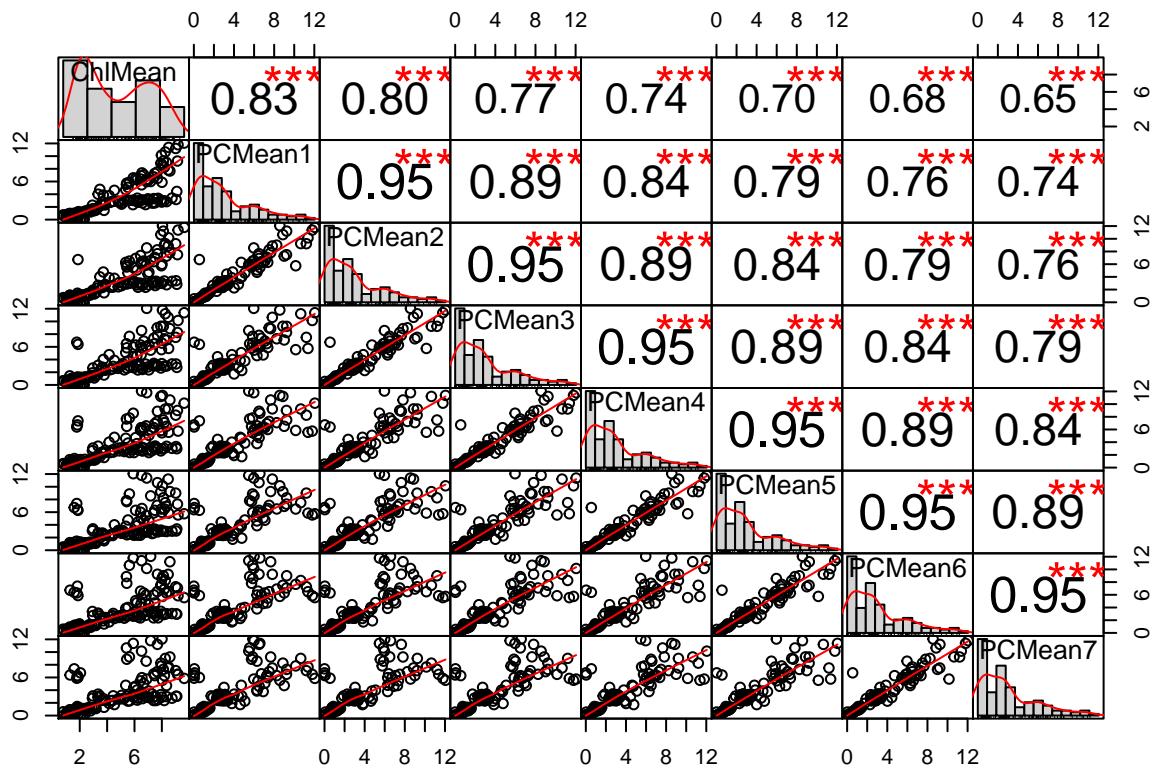






## SA18



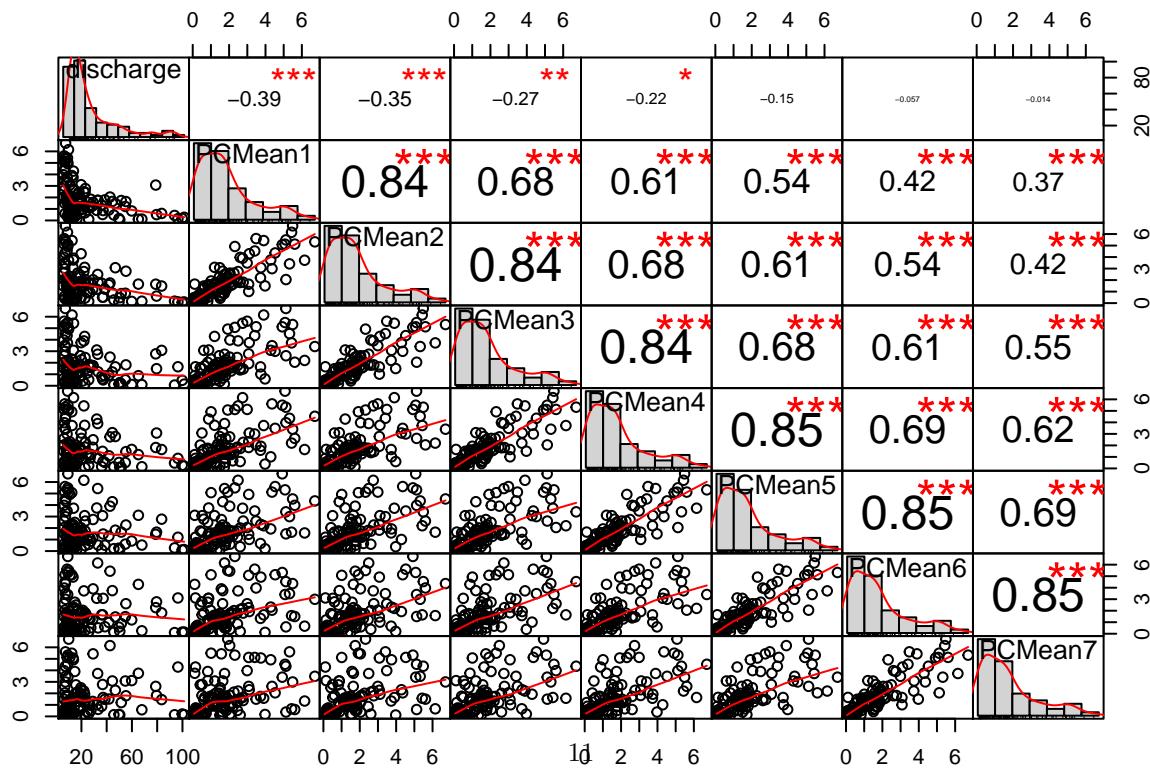
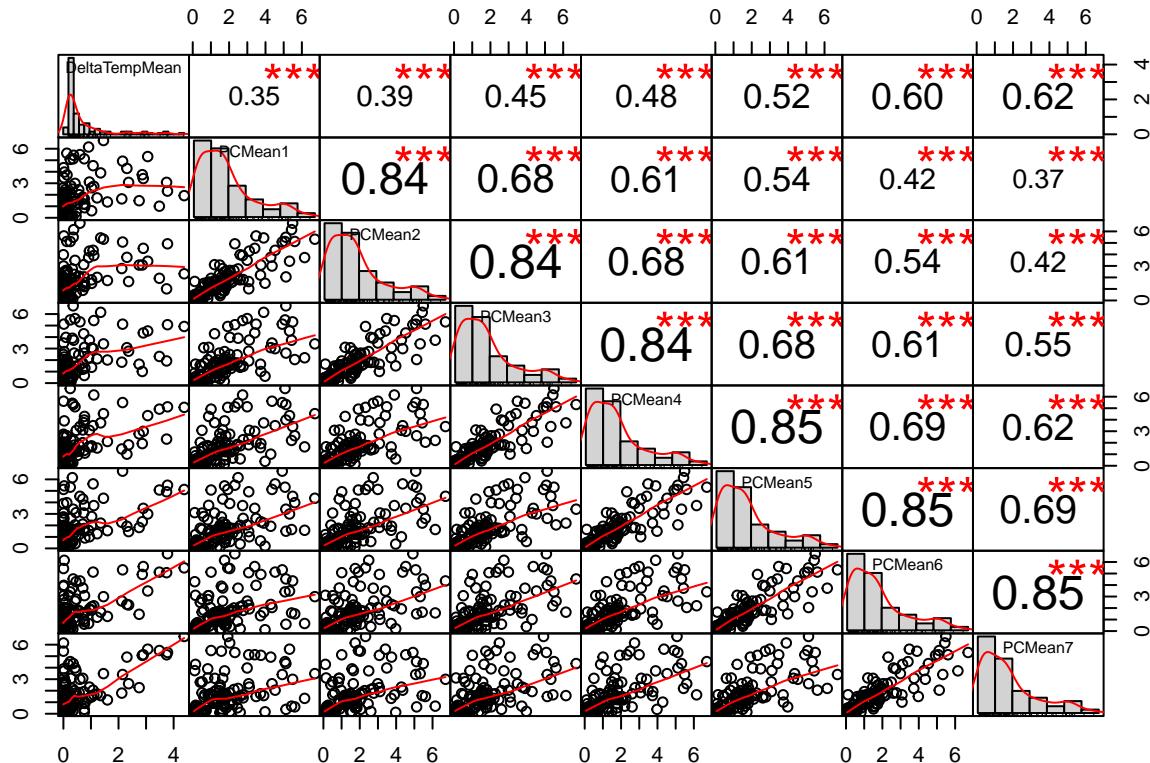


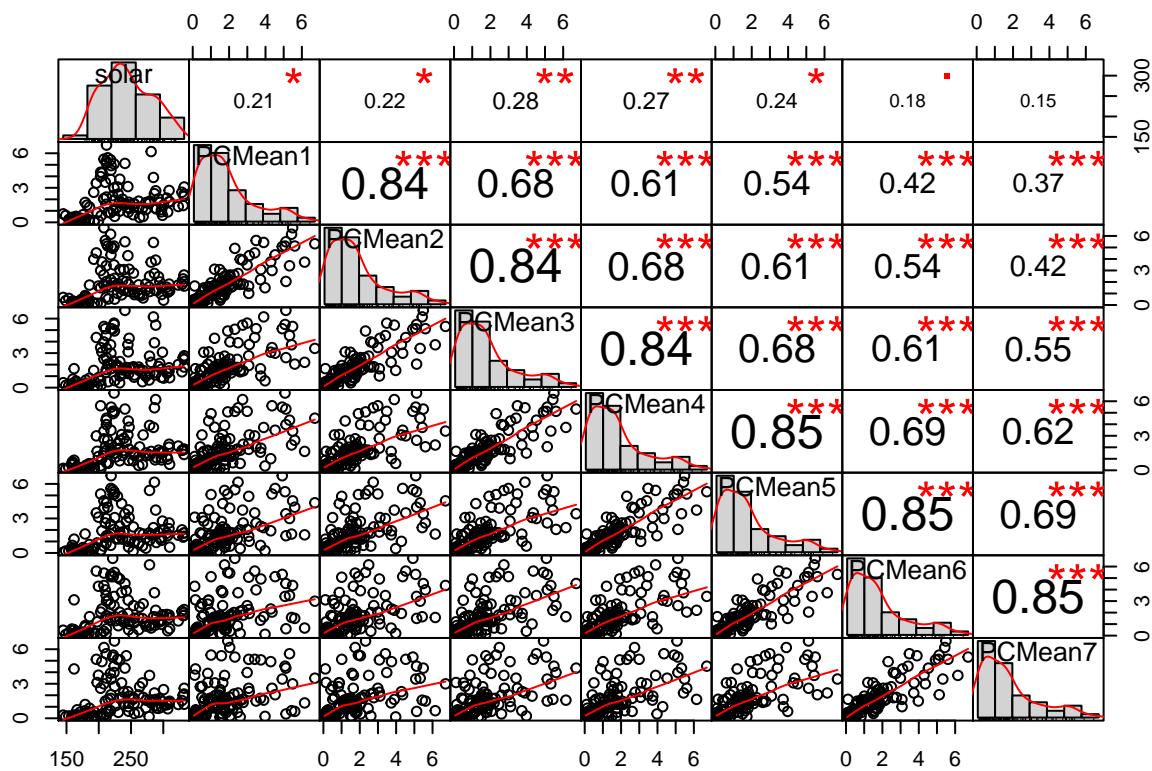
If we scan along the top line of each matrix, we see how the correlation of a given variable with PC changes with increasing time lag.

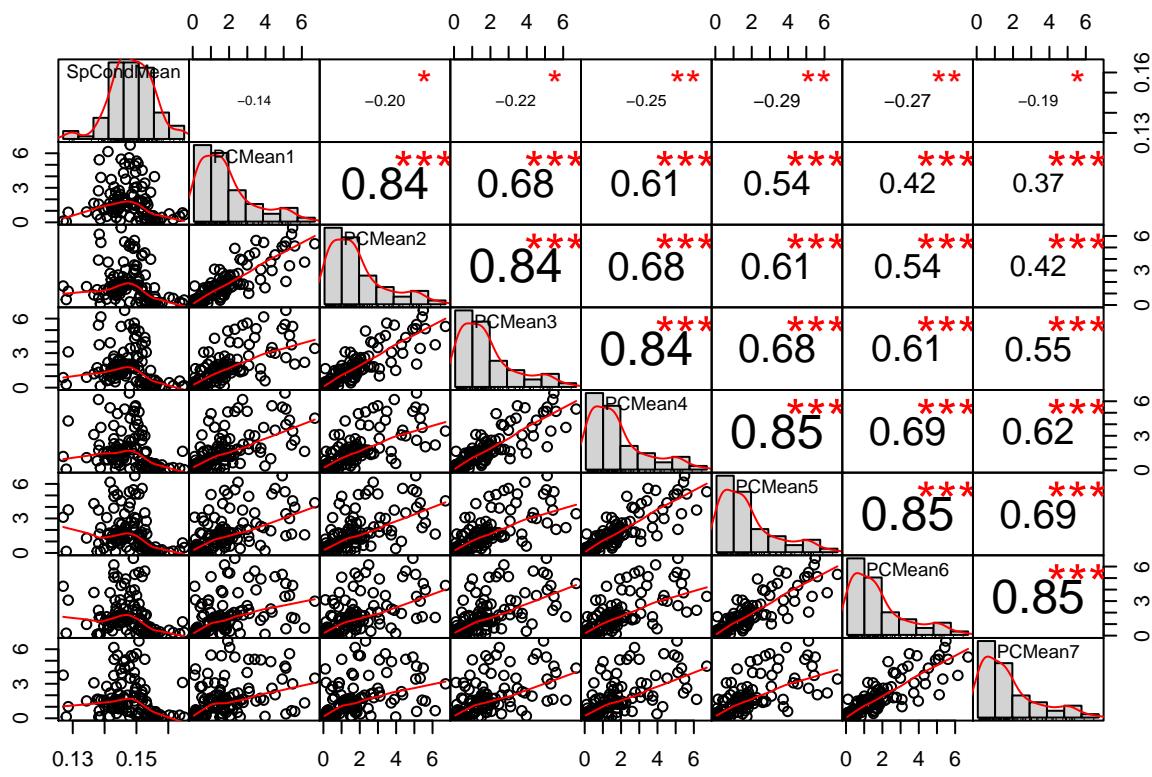
Most of these correlations become weaker and weaker with increasing time lag, although a couple remain steady (ODO in MB17, pH in SA18). Also interesting that ODO and PC are not at all correlated in MB2018.

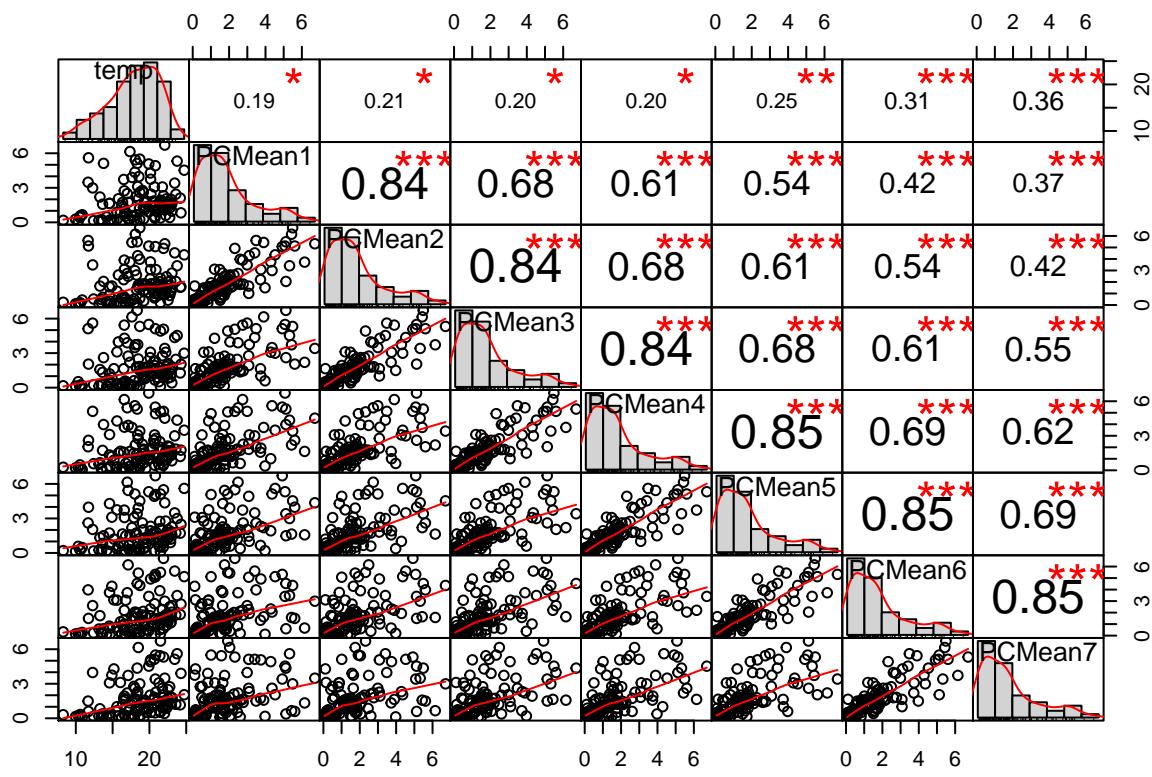
## Feature Importance variables

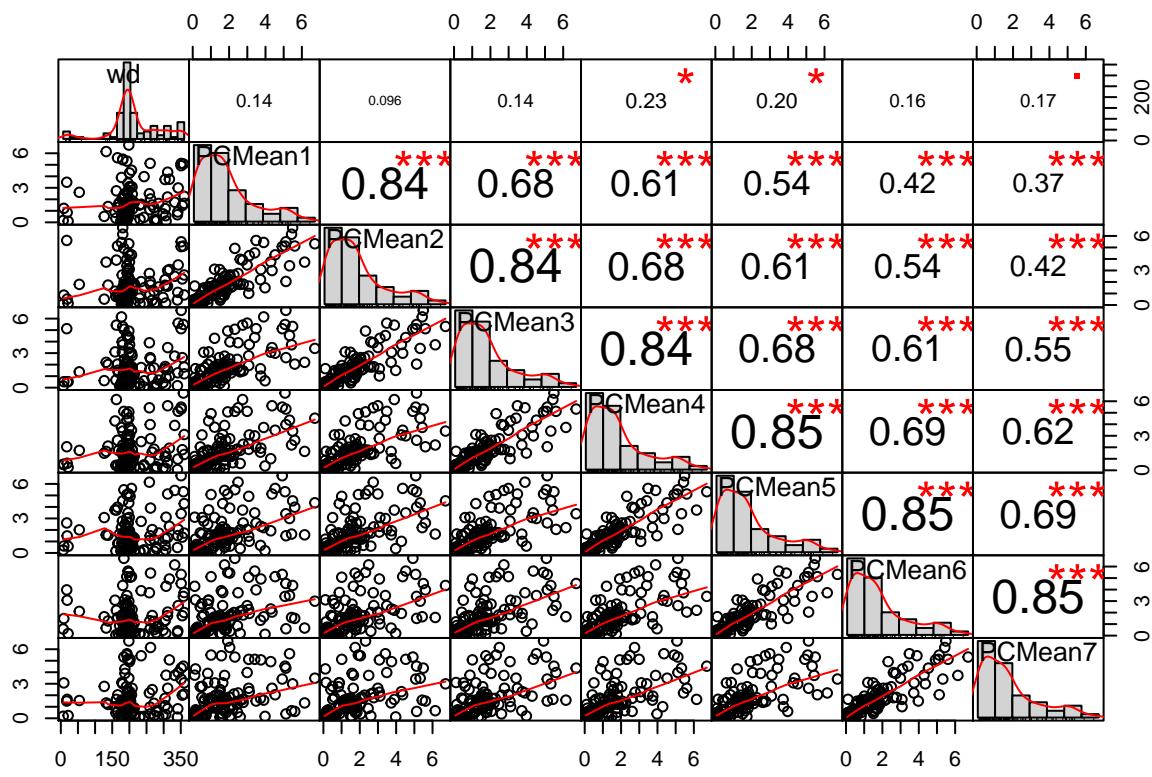
MB17

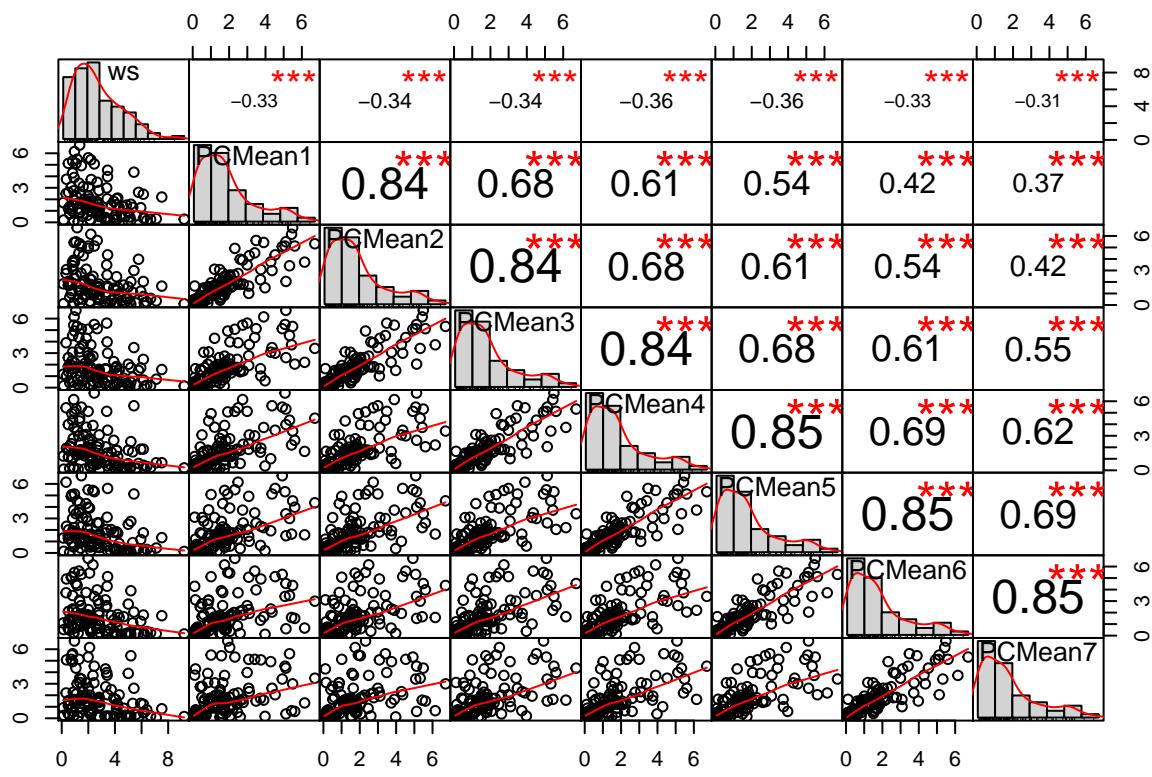




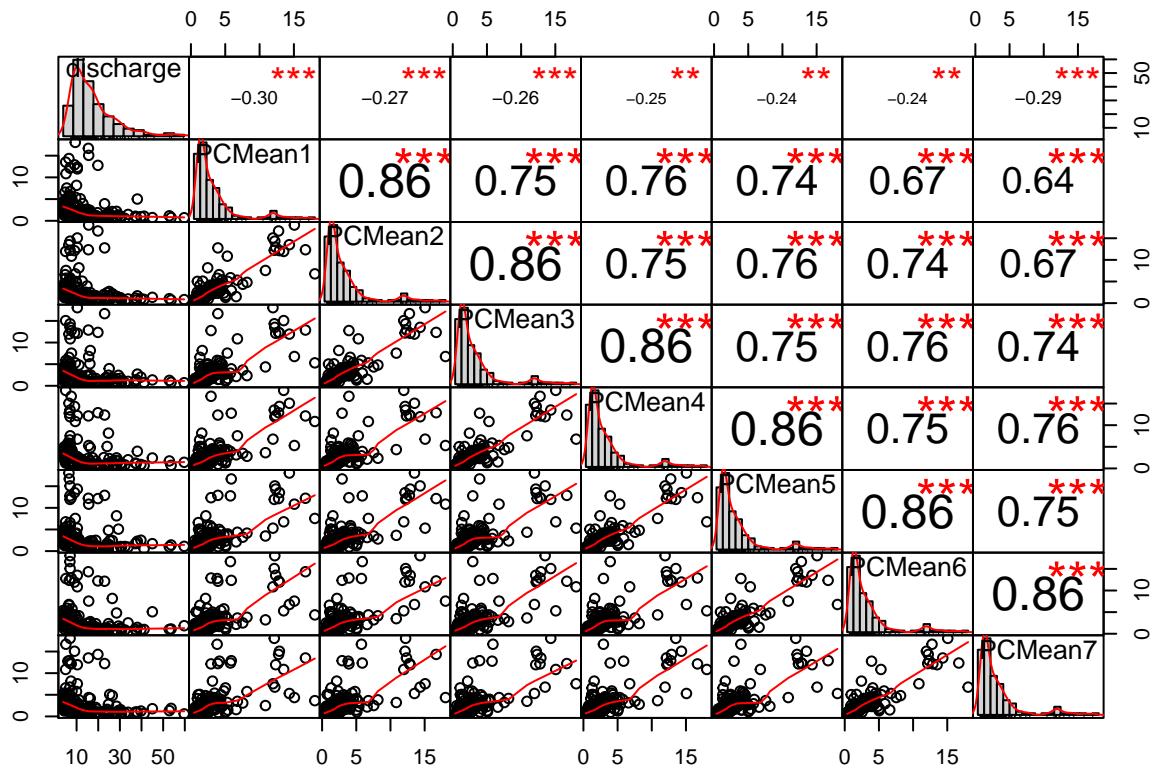
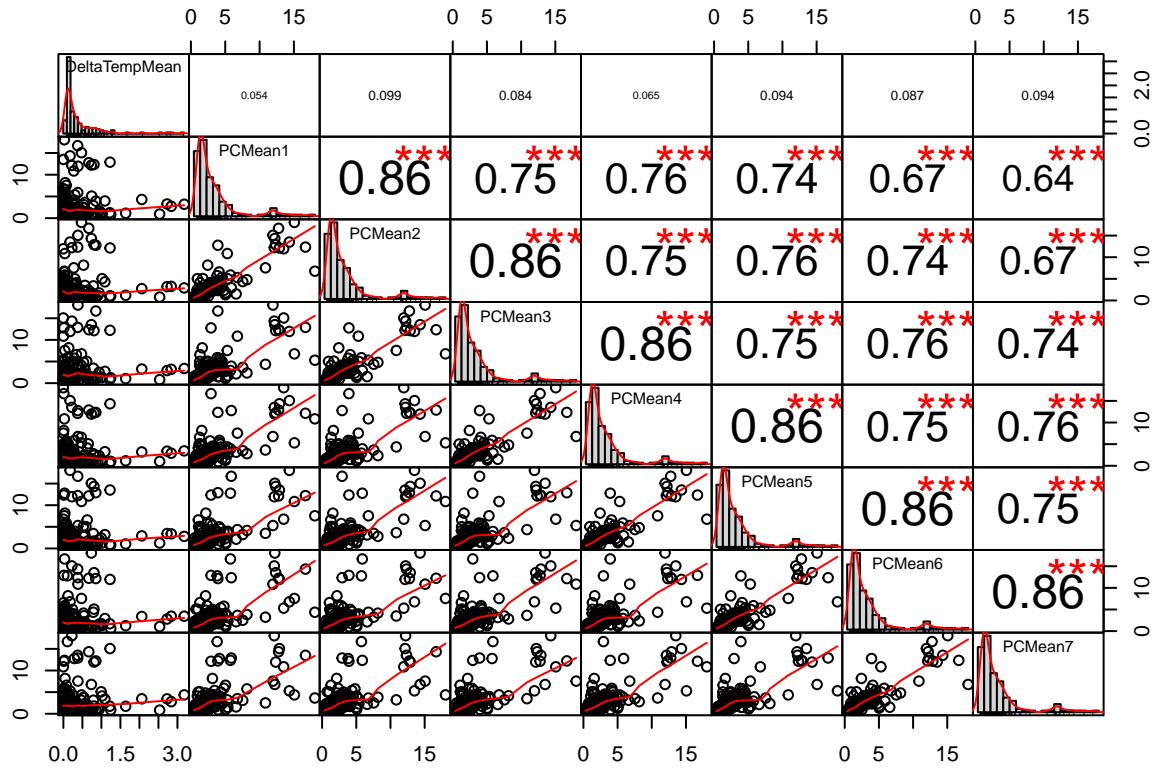


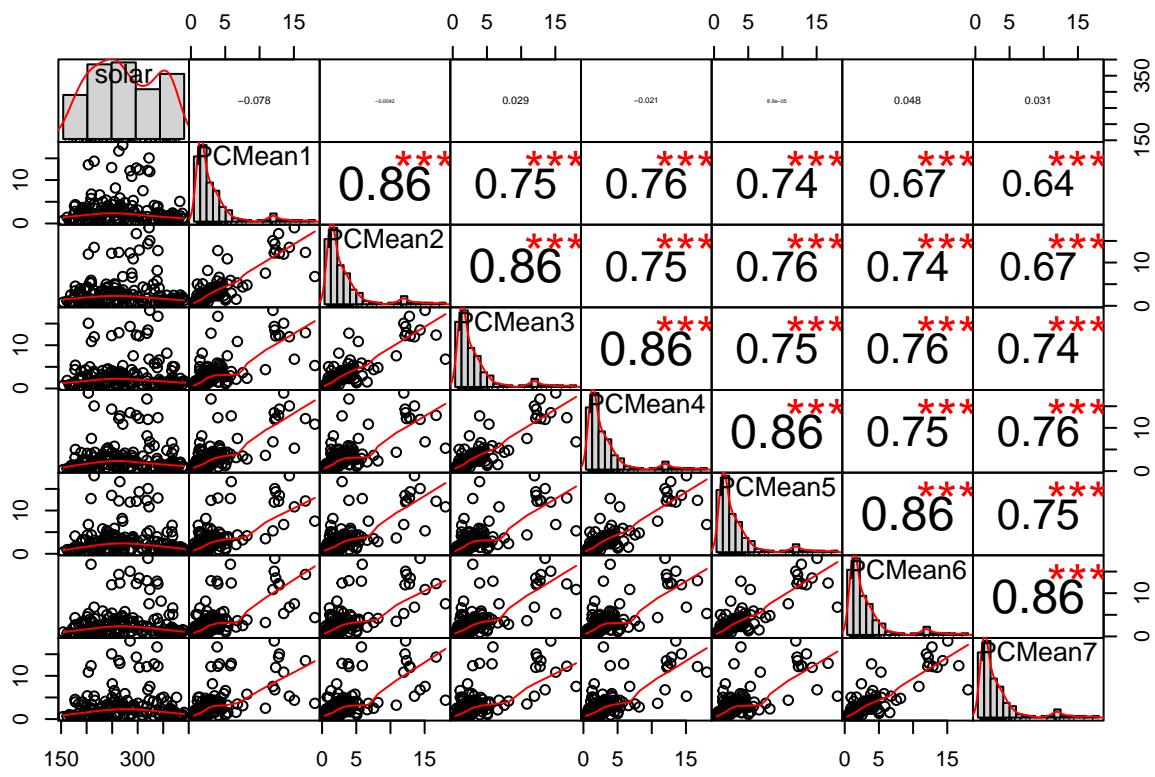


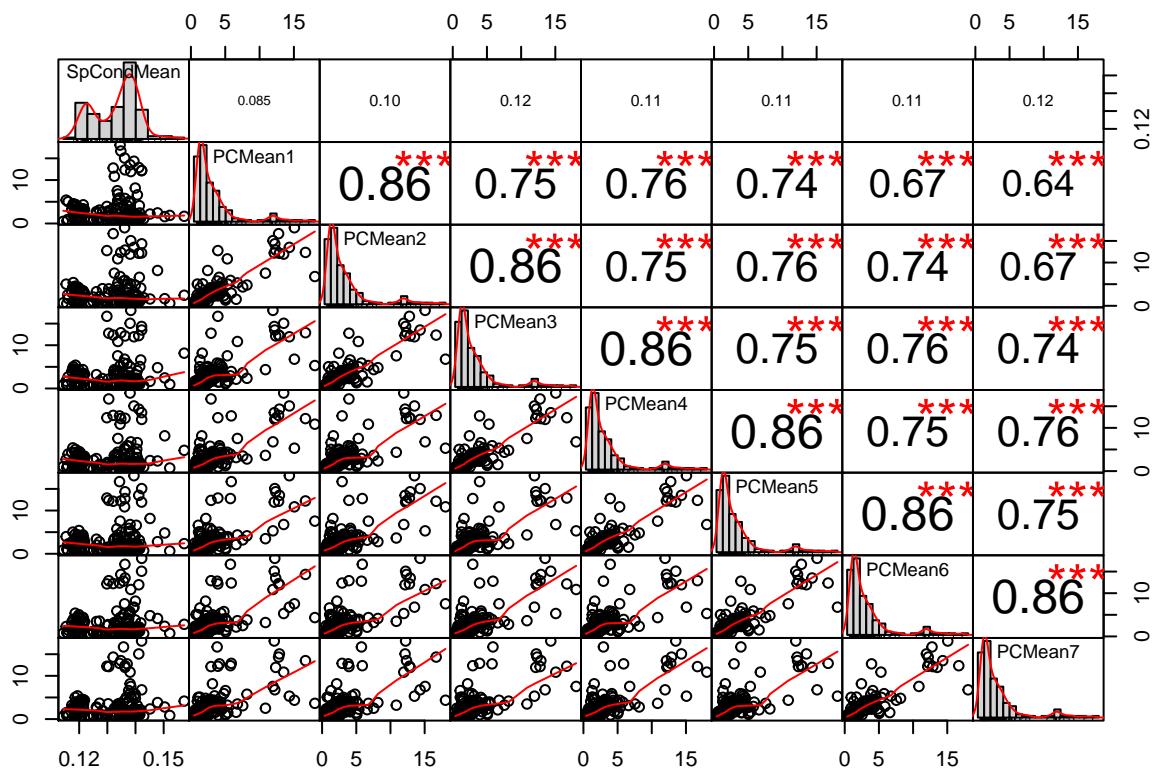


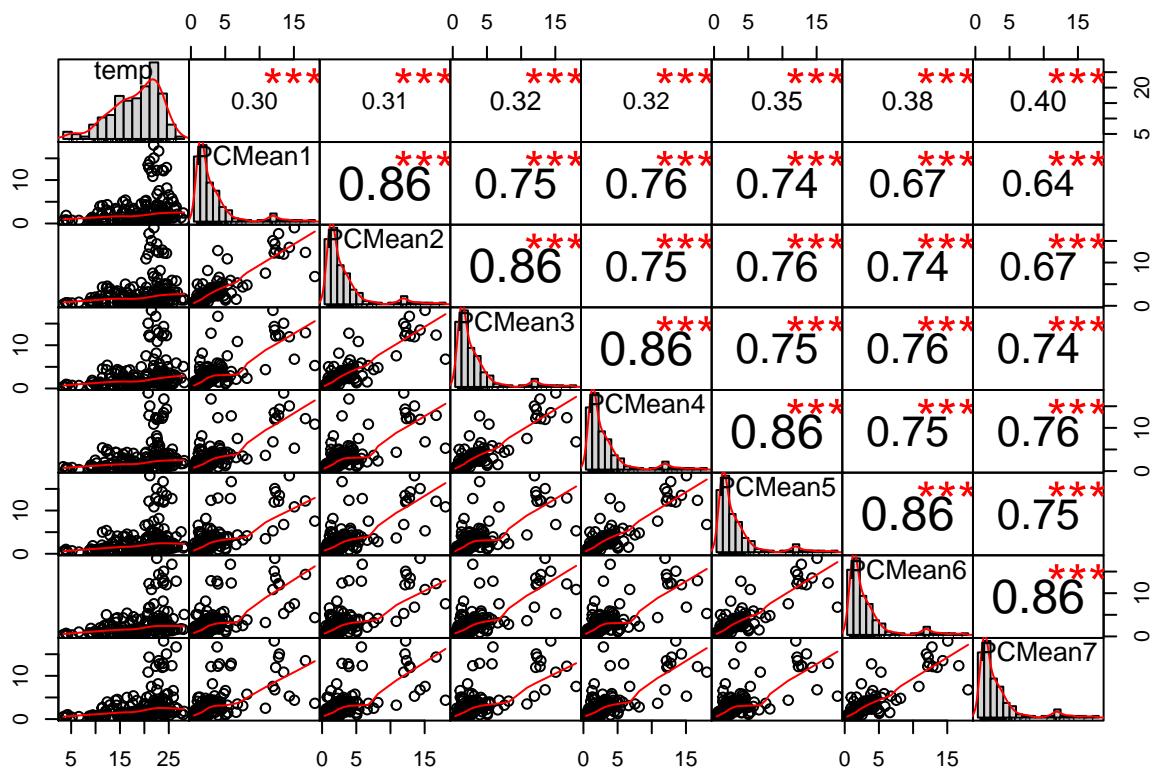


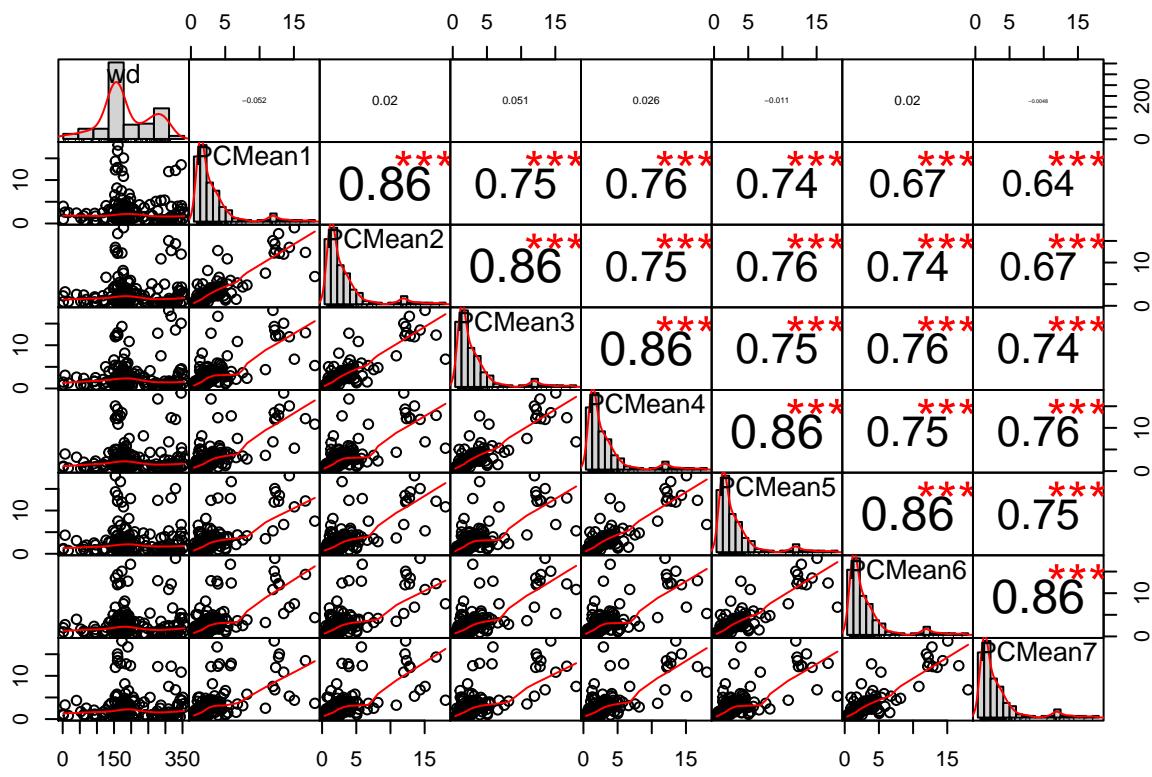
## MB18

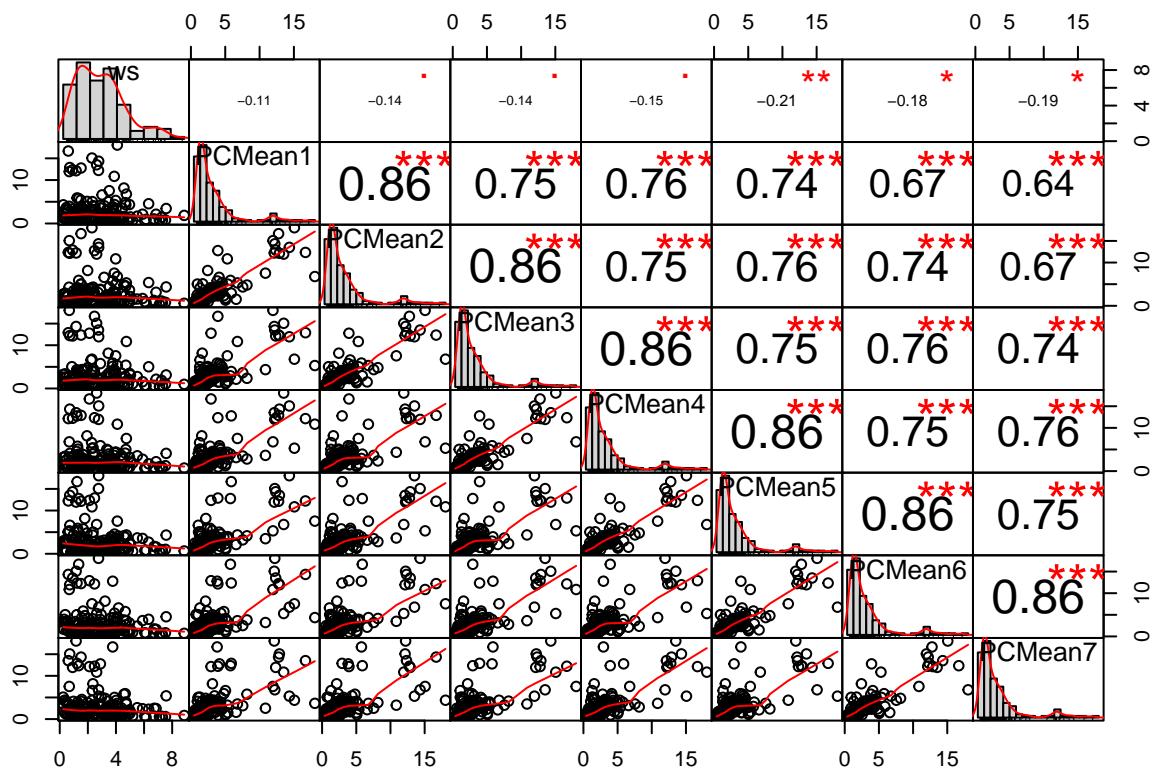


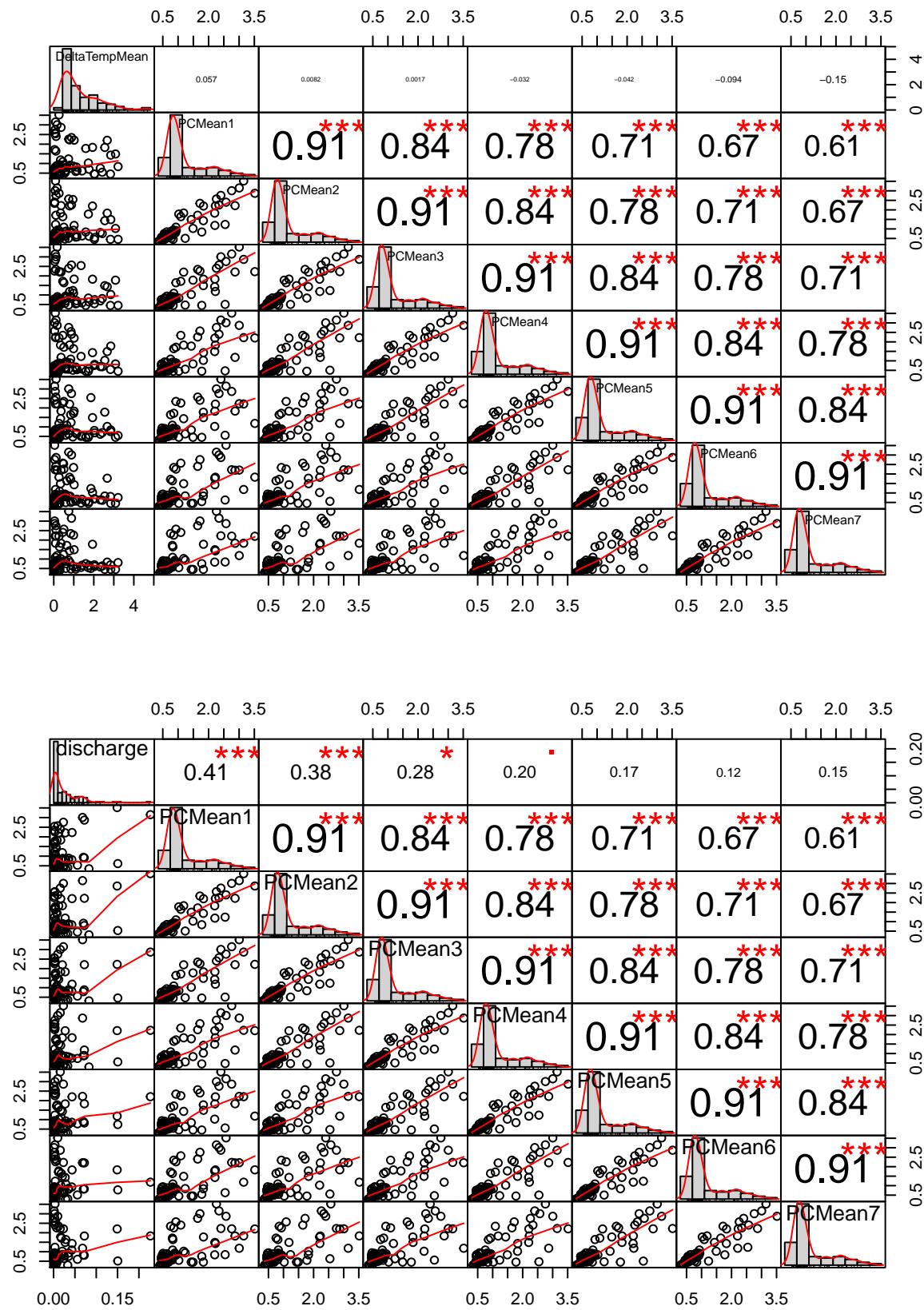


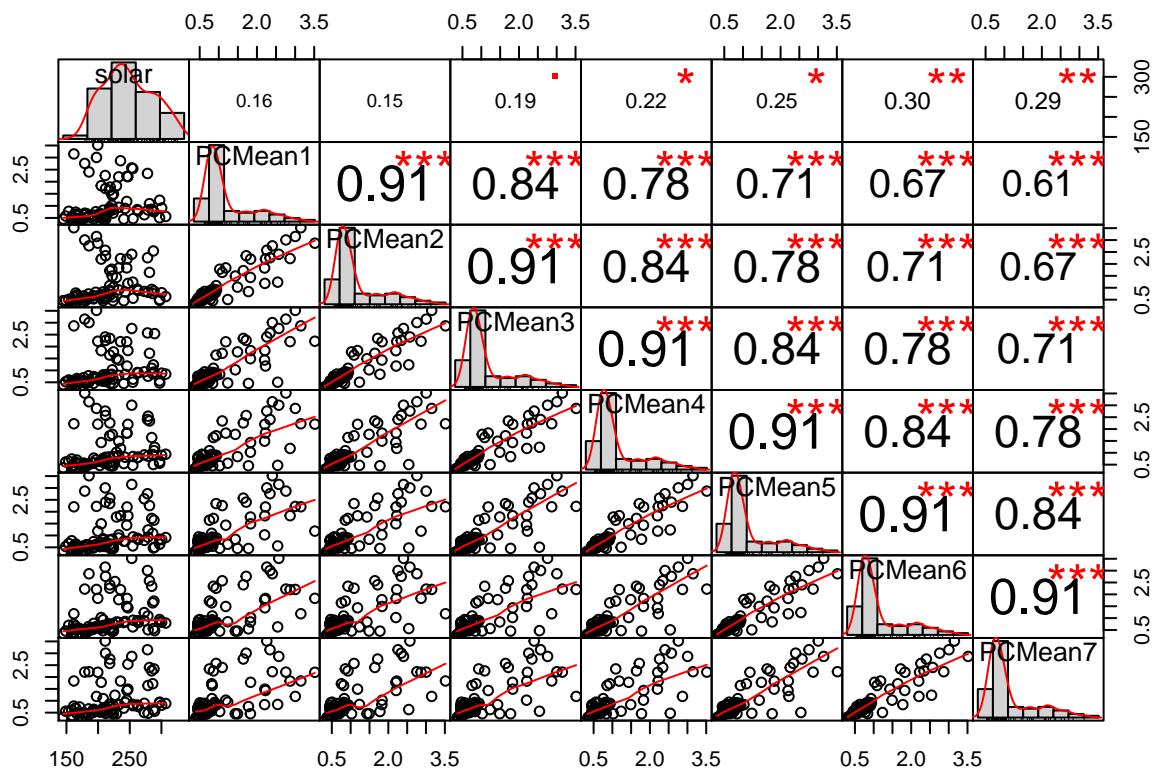


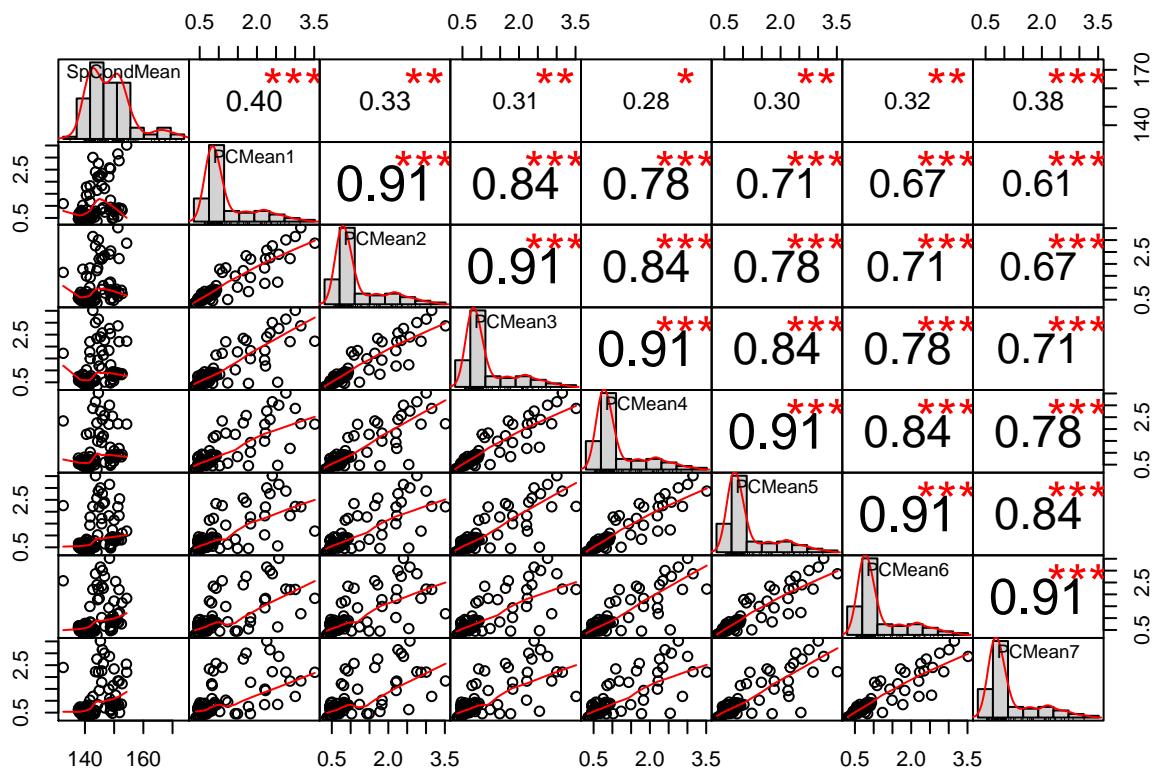


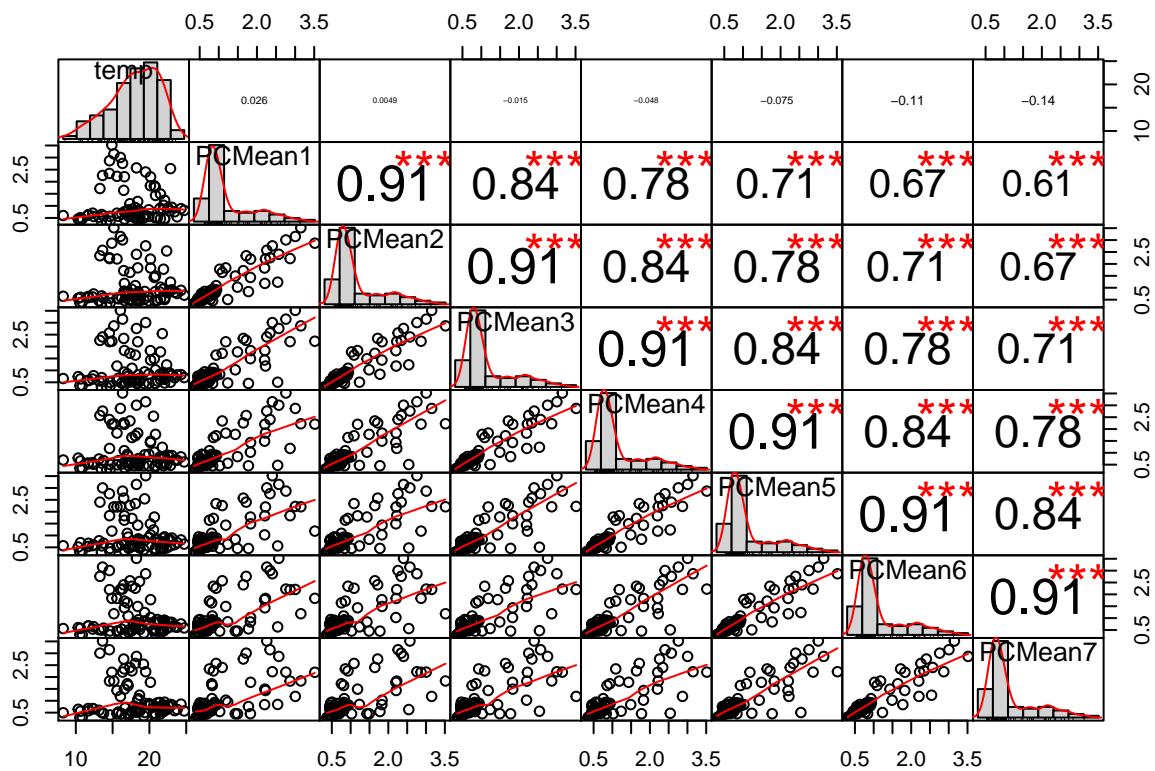


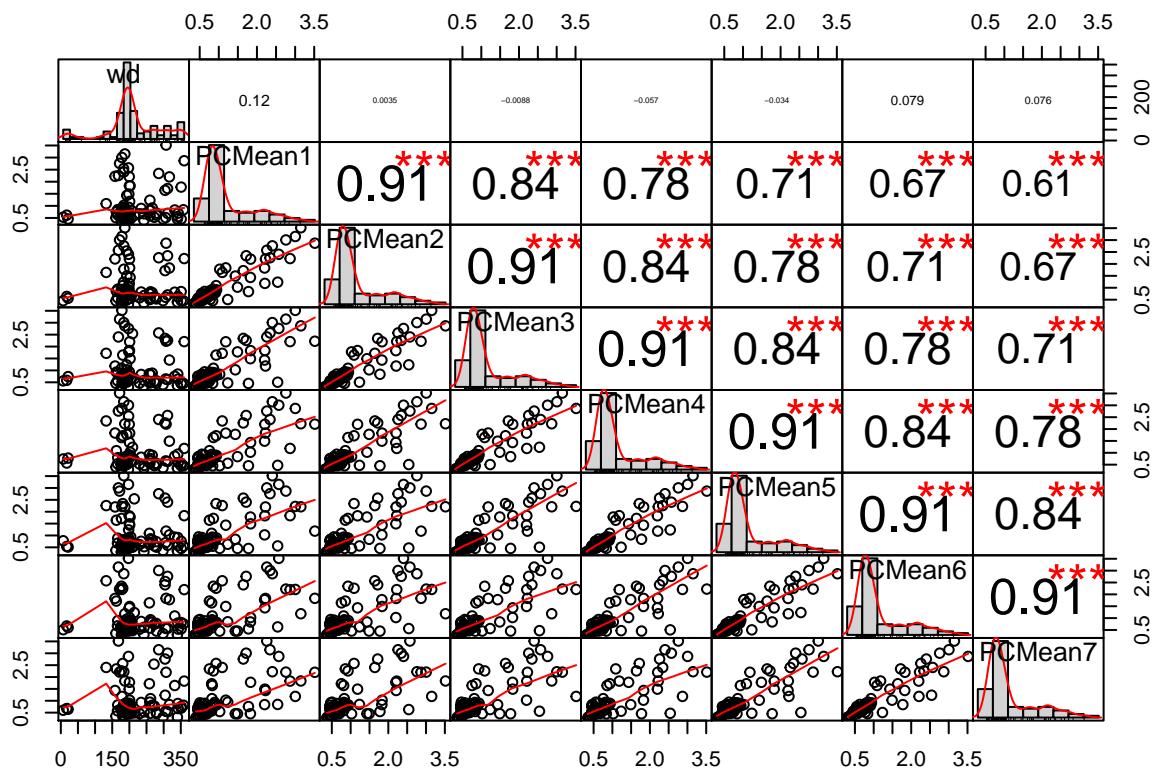


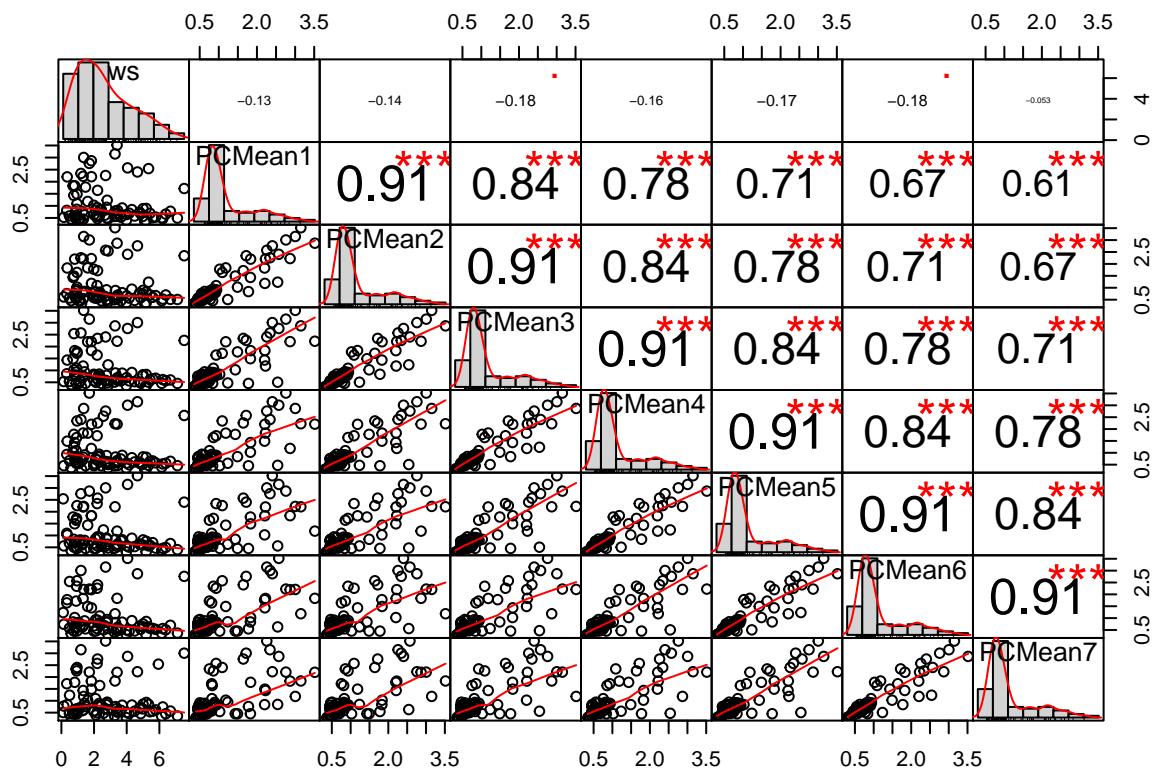




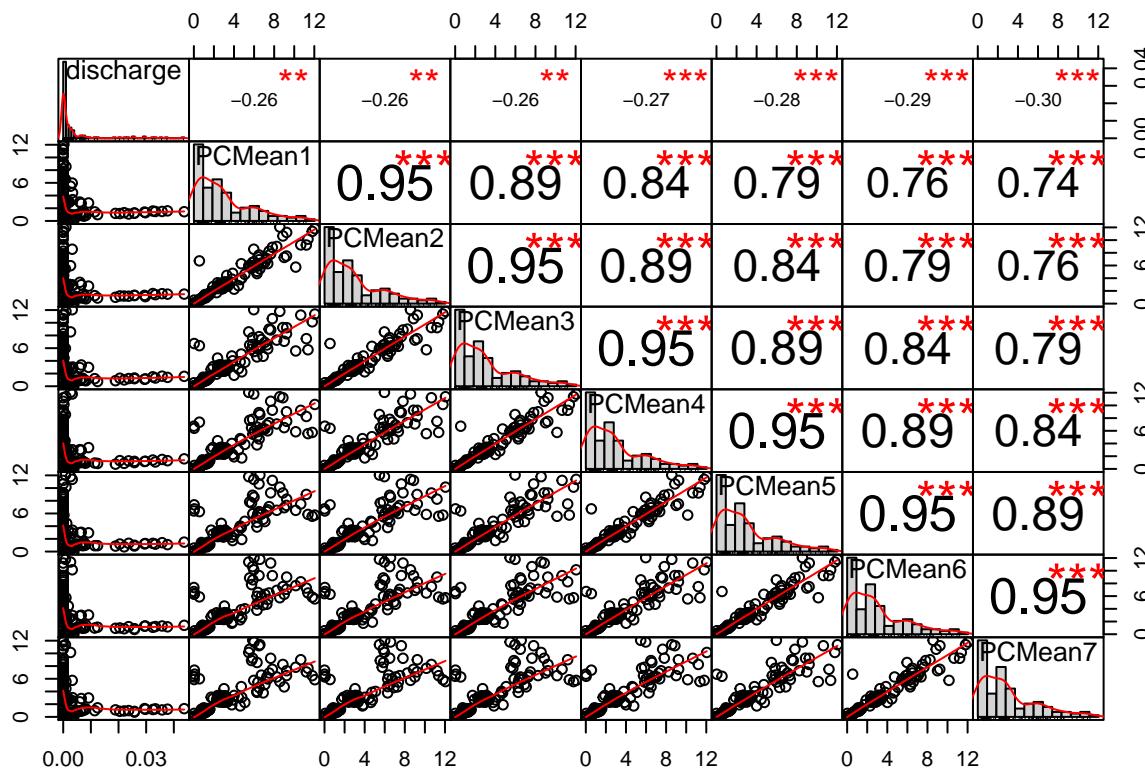
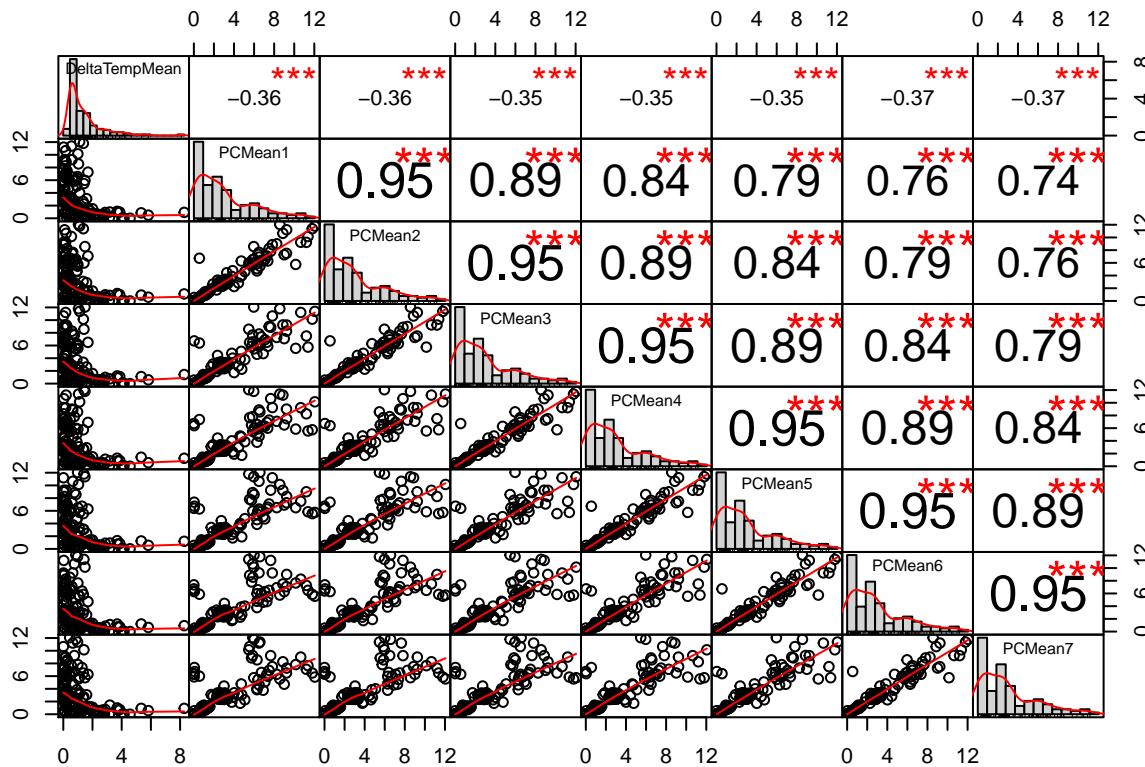


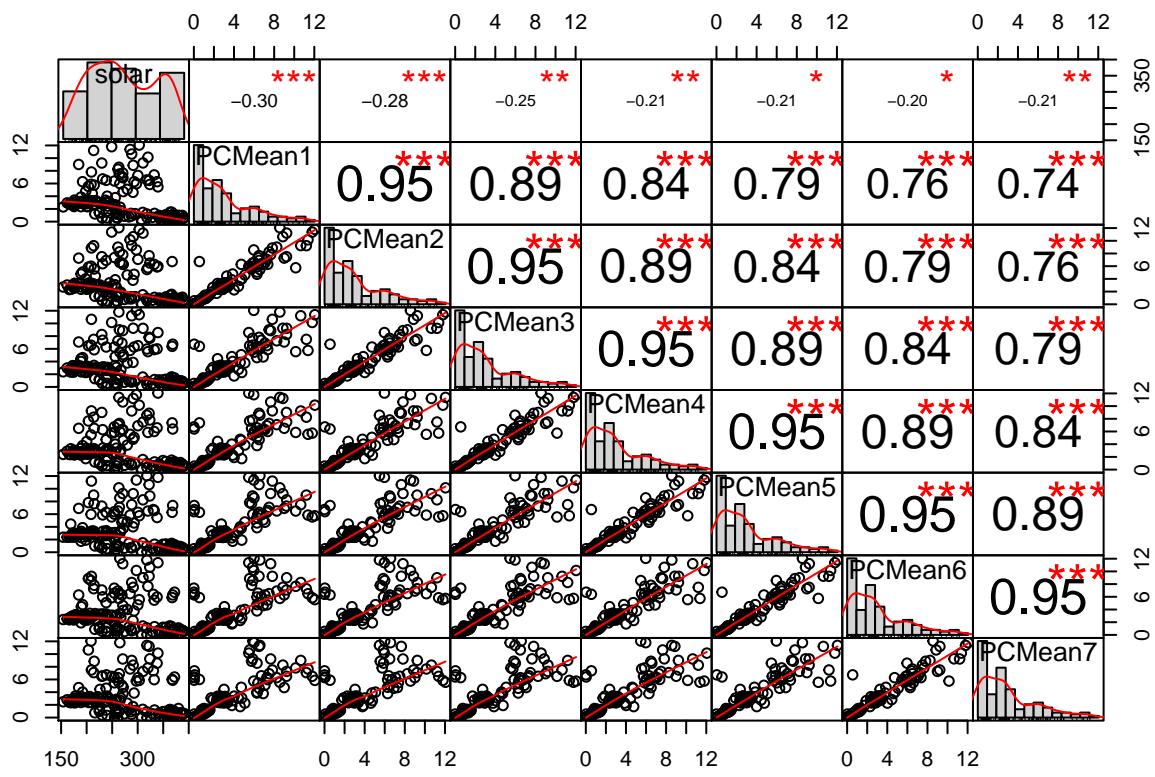


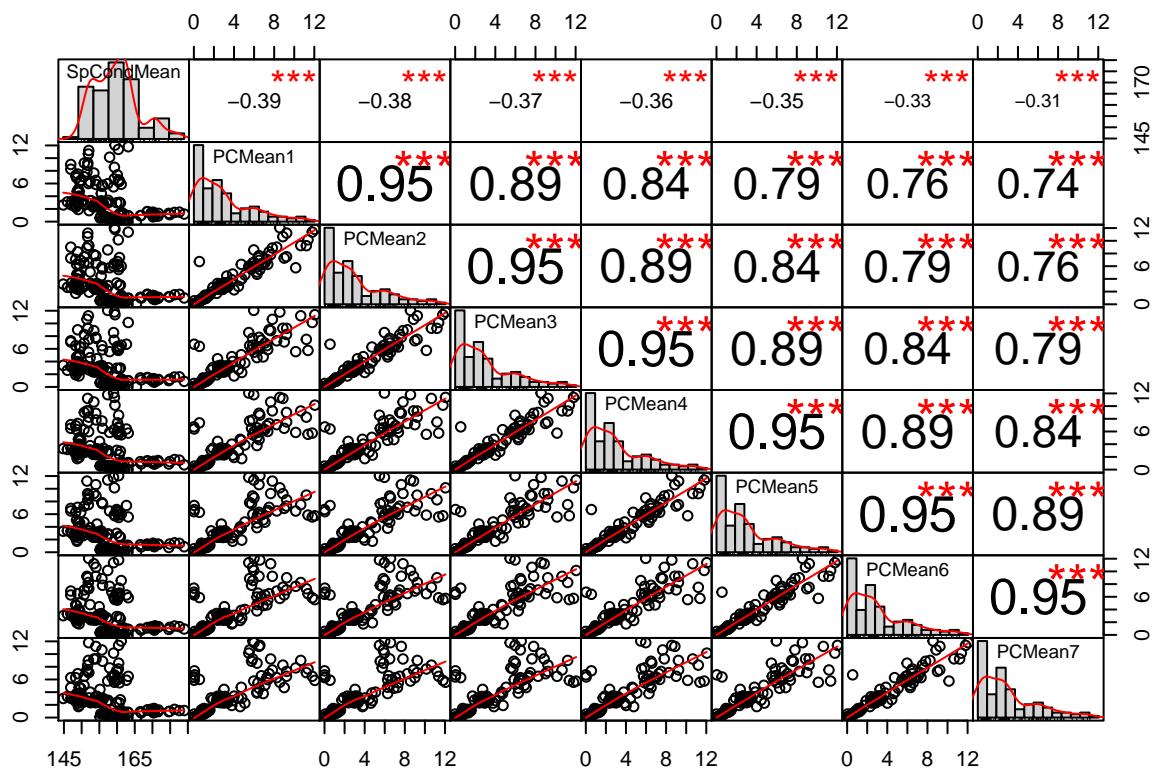


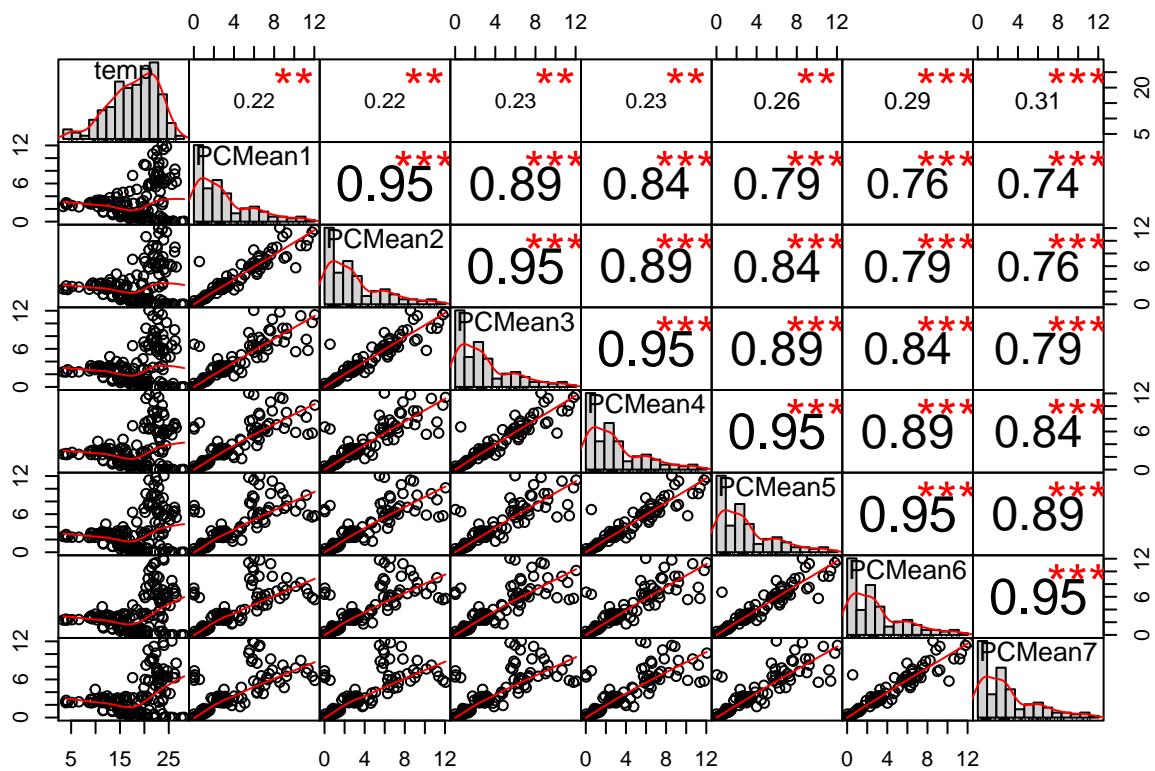


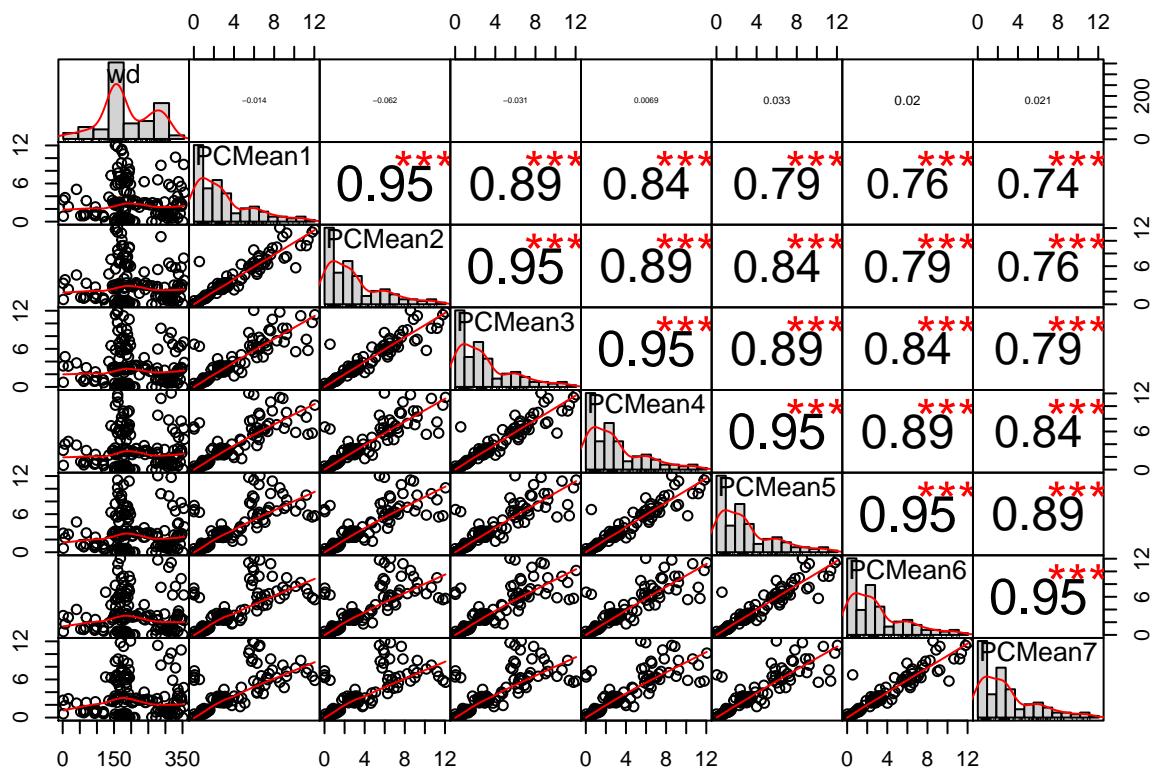
SA18

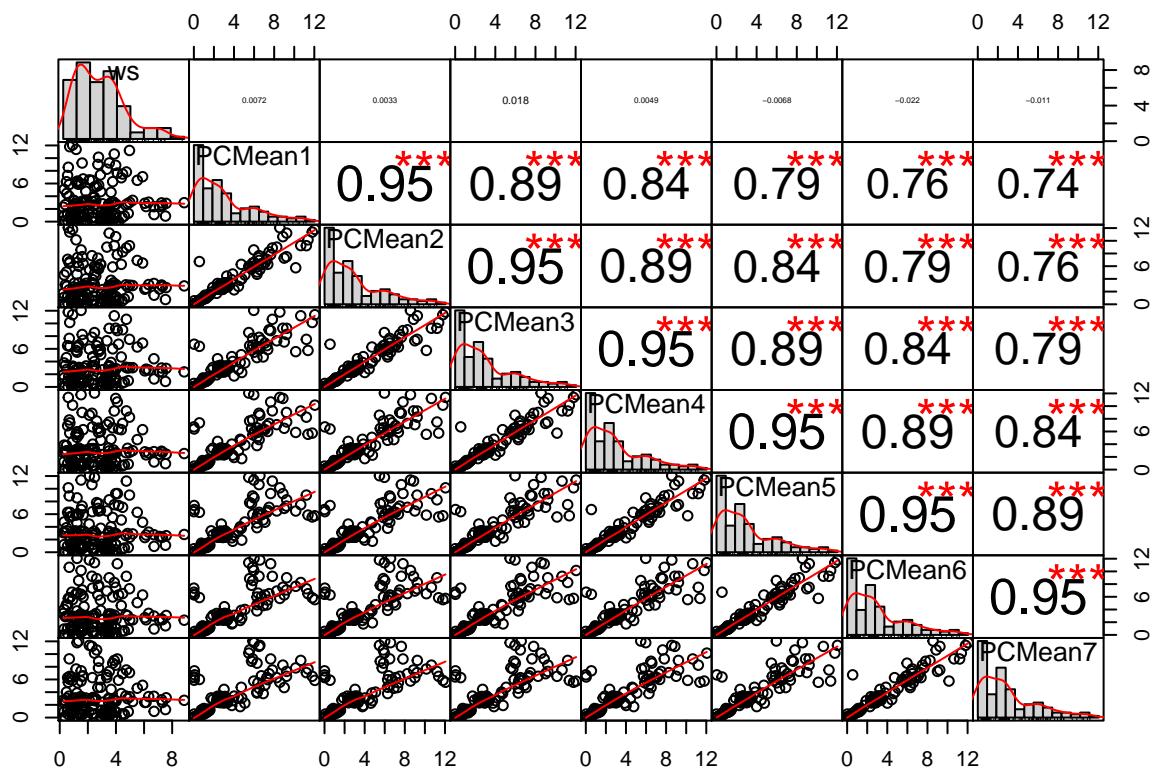






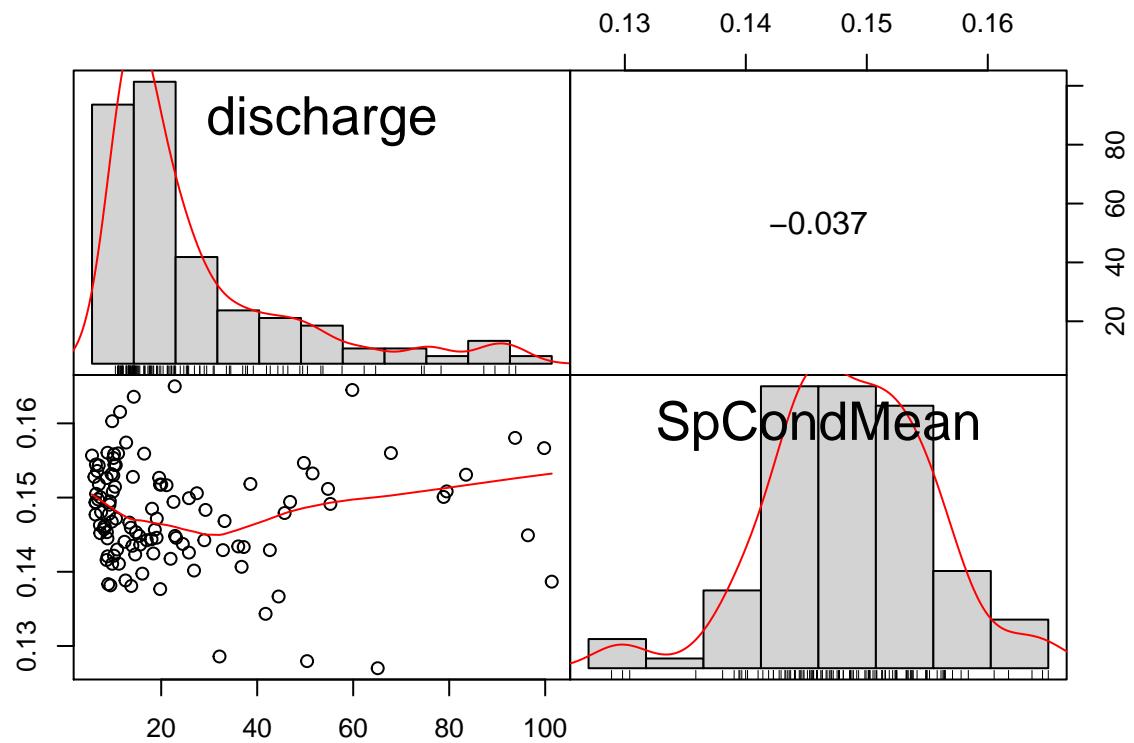






These sometimes show the same trends as feature importance, but not always.

Incidentally, discharge and salinity are NOT correlated in MB, but ARE correlated in St. Al-



bans.

