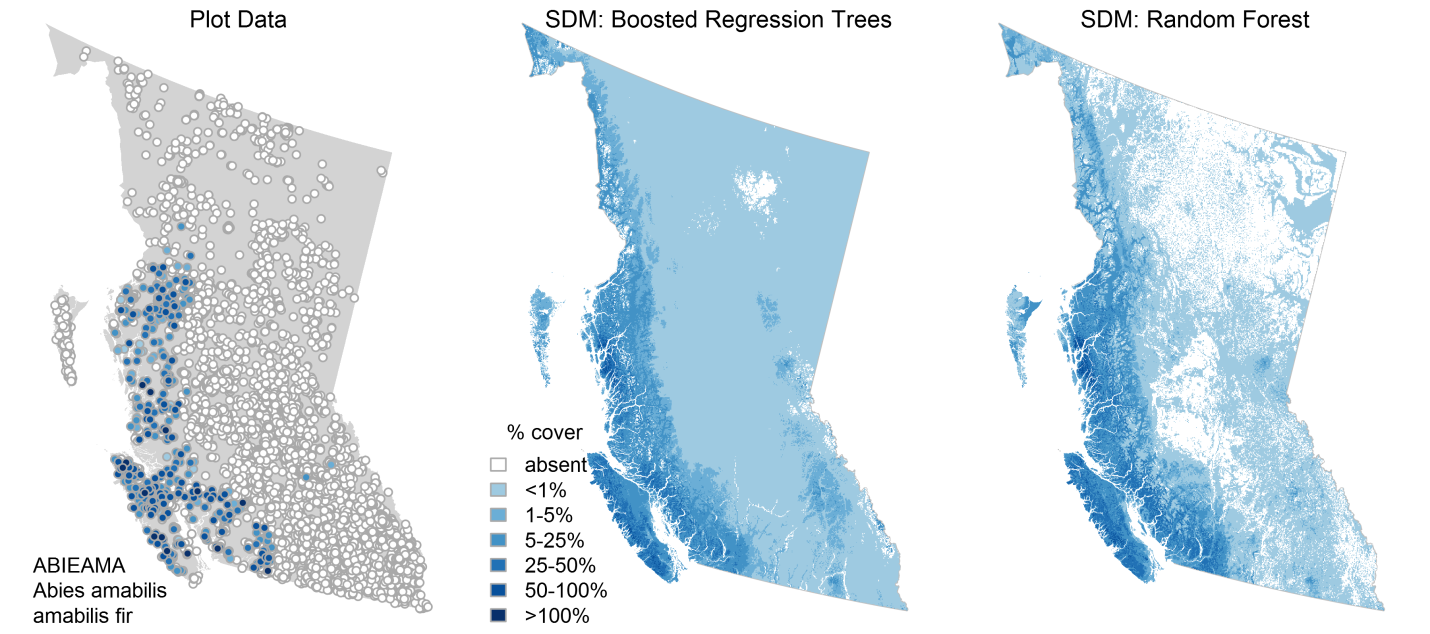
## Species abundance modeling with the BEC database

September 20th, 2016

BECveg\_SDMtrials\_ABIEAMA.R

I narrowed the BEC database down to a subset of plots and species, documented in the script “BECveg\_InputData\_Sept2016.R”. next step is to test out species abundance modeling methods. abundance modeling is quite uncommon, but has clear advantages (Howard et al. 2014), especially for the purpose of modeling BGC units based on vegetation, where abundance rather than presence is a critical input. the two obvious approaches are random forest and boosted regression trees (De’ath 2007, Elith et al. 2008, Young et al. 2012).

I first tried modeling the raw abundance values, and found that both algorithms badly overpredicted into the trace cover class:



I then used cover class as the response variable, and both methods gave an acceptable species abundance model (below). (one other methodological difference here is that I used the family=”poisson” instead of “Gaussian” for the BRT model). I’m not sure why there is this difference between % cover and cover class, but it is dramatic. It seems that both algorithms are sensitive to the scaling of the response variable.

I did a similar analysis for ABIELAS and PINUPON. Overall, the results look really good. note the absence of PINUPON from the bunchgrass zone: where there are enough plots, the climatic thresholds are well modeled. Comparing the two algorithms: random forest seems to be a bit more likely to predict presence in novel climates (e.g. Vancouver, Naikoon for ABIEAMA), though both models do predict into novel climates (e.g. Northwest BC).

Both methods appear to be roughly equivalent, but RF is a lot faster. For this reason, I’ll proceed with RF for the moment.

