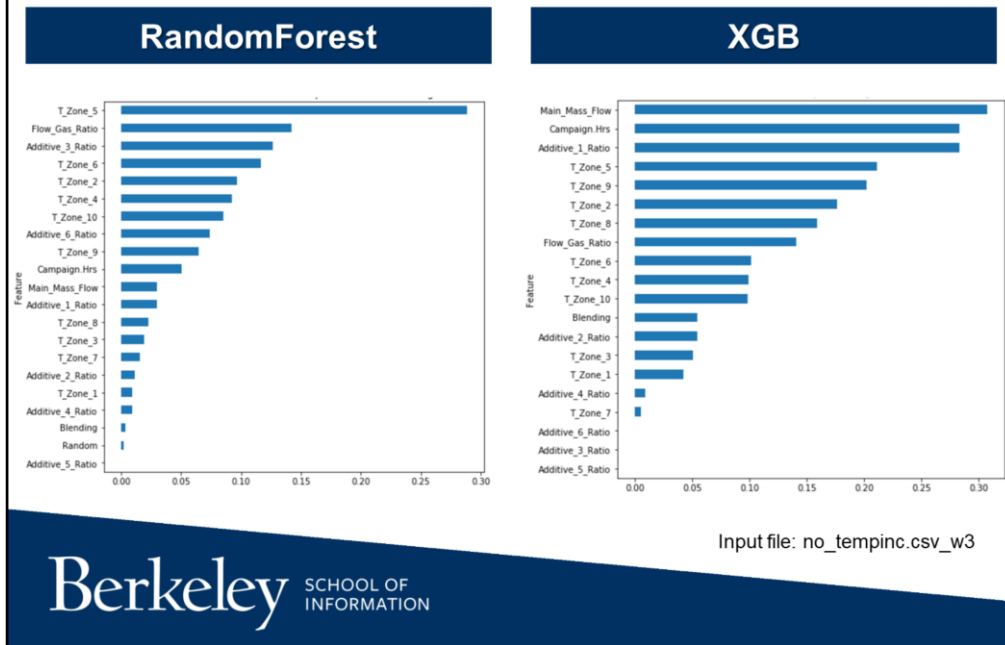


## Model permutation importance comparison



It is immediately evident based on the top five important variables identified in each model that T\_Zone\_2 and T\_Zone\_6 have significant impact on Quality; T\_Zone\_5 was ranked sixth by the XGB model so it also is likely a contributor to the outcome of the process. This also stresses the importance of energy balance in the process.

With regards to the material balance, the models can't seem to agree which of Main\_Mass\_Flow and Flow\_Gas\_Ratio are important. This is likely a gap in our design, as we only provided two variables relating to material balance.

In terms of the chemistry aspects of the process, the models start to diverge and tell conflicting stories. The additive ratio ranks do not agree, and the worse example of this is Additive\_3\_ratio which is ranked in the top 5 in RandomForest, but the lowest in the XGB model. At this point one would rely on the domain knowledge to determine if one model is making a critical error over the other. For example, is the additive just inert gas? In that case, we can confidently say the RandomForest model is erroneous as we do not expect any impact to Quality from inert components. Without intimate knowledge of the process, we cannot definitively determine why the models are diverging here.