PROGRESS REPORT

PROJECT TITLE: Agricultural production and the hydrologic cycle in the Upper Midwest
PROJECT NUMBER: MN CORN RES & PROMO COUNCIL/4101-15SP (no cost extension)
REPORTING PERIOD: Oct 01 to Dec 31, 2016
PRINCIPAL INVESTIGATOR: Tim Griffis
ORGANIZATION: Dept. Soil, Water, and Climate, University of Minnesota
PHONE NUMBER: 612-625-3117
EMAIL: timgriffis@umn.edu

1) PROJECT ACTIVITIES COMPLETED DURING THE REPORTING PERIOD. *(Describe project progress specific to goals, objectives, and deliverables identified in the project workplan.)*

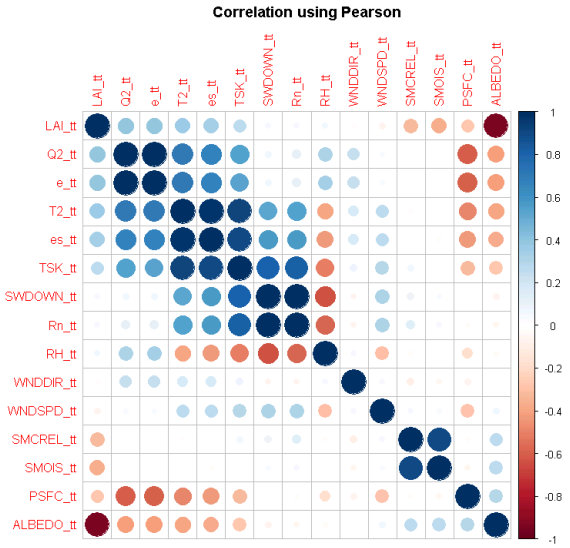
Over the past 3 months the majority of our work has focused on assessing the quality of ET forecasts for the 2016 growing season and trying to understand how to improve the forecasts. All of these forecasts were posted at our ETool website: http://www.biometeorology.umn.edu/research/etool

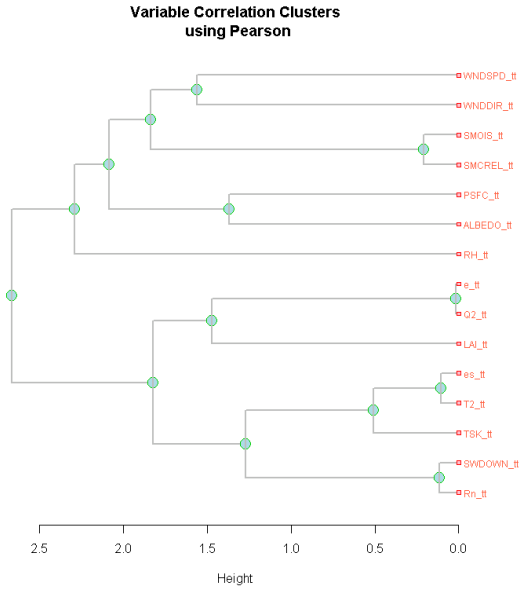
We are continuing our work with University of Minnesota U-Spatial team to improve the accessibility of our forecasts and to increase their functionality (i.e. making easily obtained estimates of precip, ET, water deficit, etc.). Based on the progress with the U-Spatial team it appears that our web interface will be ready for the 2017 growing season forecasts. We are excited about this development.

2) IDENTIFY ANY SIGNIFICANT FINDINGS AND RESULTS OF THE PROJECT TO DATE.

Machine learning is a widely used methodology in "big data" analyses. We are beginning to use these techniques to help understand the quality of our ET forecast and to seek ways to improve those forecasts. Given the high spatial and temporal resolution of the model forecast, machine learning techniques provide a powerful way to explore key relations among the driving variables.

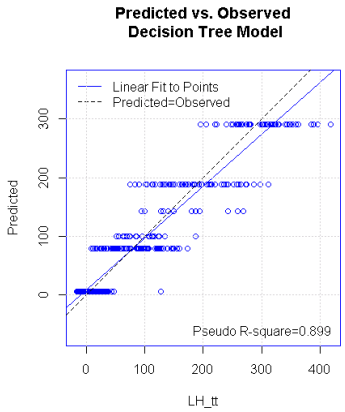
Here, we used the package 'rattle' in R to conduct our machine learning study. As an example, we selected hourly data from May 3 to October 11, 2015 at the grid point of our tall tower observations. In big data analyses, more data is better, no matter whether the variations are correlated. Therefore, we chose fifteen variables (albedo, MODIS LAI, air pressure, specific humidity, relative humidity, net radiation, soil moisture, relative soil moisture, downward shortwave radiation, air temperature, surface skin temperature, surface wind speed, wind direction, air vapor pressure and saturated vapor pressure) from as the machine learning inputs.

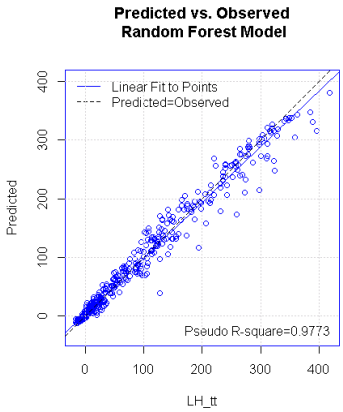


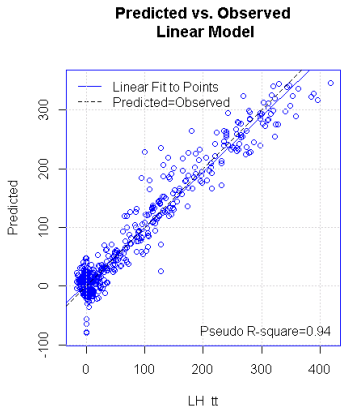


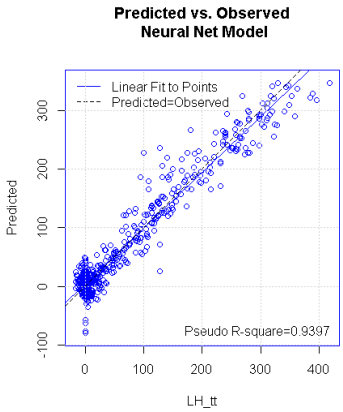
The correlation map above shows that LAI is highly negatively correlated to surface albedo. Soil moisture and relative soil moisture have weak or no correlation to the atmospheric variables.

The technique that we used provides four types of machine learning models: Decision Tree Model, Random Forest Model, Linear Model and Neural Net Model. In the machine learning process, data are separated into two parts, one part is used to train the model, and the other part is used to validate and test the model.

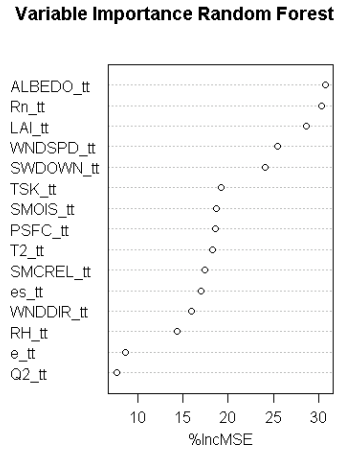








The validation results showed that Random Forest Model has the highest pseudo R-square (0.97).



Using the Random Forest Model, a larger IncMSE suggests a variable is more important. Albedo, net radiation and LAI ranked as the top three most important variables. As expected, this indicates that radiation is the dominant control on ET. The correlation between albedo and LAI indicates that an accurate representation of LAI is critical to the ET forecast. We have been testing and trying to improve the representation of LAI in the model. Soil moisture also ranked high among the fifteen variables. As we have suspected, the Random Forest Model also suggests that the soil moisture is important and represents a critical boundary conditions for the regional

weather model. This is an issue we are currently exploring. In our next steps we will make use of COSMOS soil water content estimates as input to the forecast model.

Ke Xiao's PhD thesis will have two chapters that describe these model applications/developments.

3) CHALLENGES ENCOUNTERED. *(Describe any challenges that you encountered related to project progress specific to goals, objectives, and deliverables identified in the project workplan.)*

4) FINANCIAL INFORMATION. *(Describe any budget challenges and provide specific reasons for deviations from the projected project spending.)*

5) EDUCATION AND OUTREACH ACTIVITIES. *(Describe any conferences, workshops, field days, etc attended and/or publications developed to disseminate project results.)*
PhD student Ke Xiao has developed new modeling and analytical skills as a result of this research opportunity.