Improving Soil Organic Carbon Predictions by Incorporating Long-term Soil Moisture and Soil Temperature Observations

Destiny Kerr and Tyson Ochsner

Department of Plant and Soil Sciences, Oklahoma State University

Introduction:

- Soil organic carbon (SOC) is the Earth's largest terrestrial carbon sink.
- Understanding the relationship between climate and SOC is crucial for modeling SOC and climate change.
- Current climate change models inflate uncertainties by using indirect covariates, such as precipitation and air temperature, to estimate SOC.
- Well known relationships between SOC and subsurface climate are not being utilized in SOC predictions.
- Using subsurface climate variables, such as soil moisture and soil temperature, could improve soil carbon estimates.

Methodology:

- Data for this project came from two sources; the Oklahoma Mesonet climate data and soil cores taken from each Mesonet site.
- Soil data are collected at 5, 25, and 60 cm depths.
- Daily climate data from the beginning of data collection to the soil sampling date were used to calculate the long term average of climate variables.
- Only sites with at least 10 years of daily data prior to sampling were included.
- SOC was calculated from the soil cores by the difference of soil inorganic carbon and soil total carbon.
- The software program, MATLAB, was used organize and analyze this data set.
- Correlation coefficients was used to identify and exclude candidate variables that were highly correlated with each other (r > 0.7).
- Least absolute shrinkage and selection operator regression analysis was used to select a subset of the candidate variables which allow accurate SOC predictions.
- Ten fold cross validation was used to quantify model accuracy.
- LASSO was used to produce 100 different models for the data, and the simplest model that produced a cross validation mean square error (MSE) within one standard error of the minimum MSE was selected.

Results:

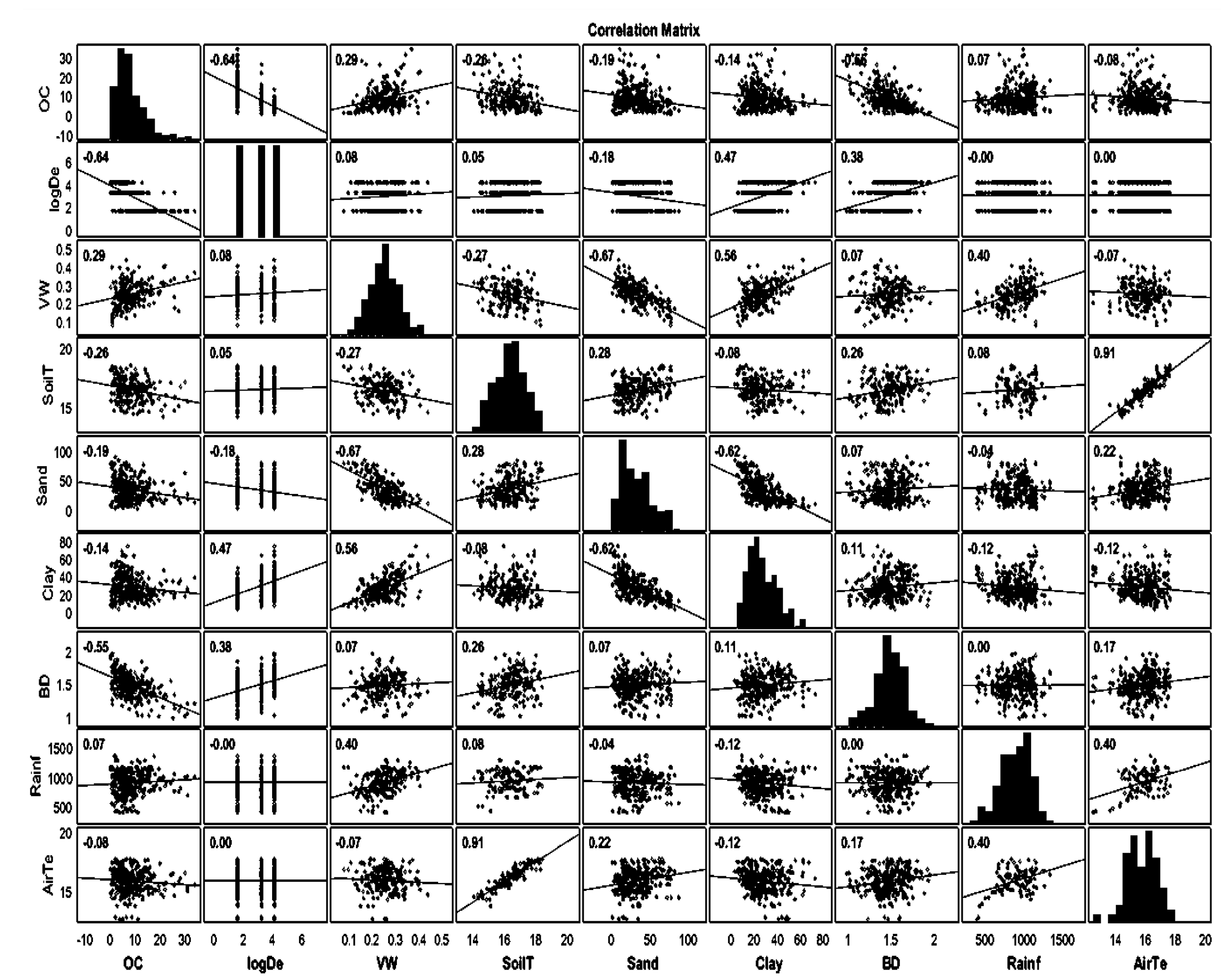

Figure 1. Correlation coefficient matrix of SOC and the eight candidate predictor variable inputs used in LASSO and the r value.

Table1 . List of the predictor variable inputs used in the LASSO regression , non-standardized and standardized coefficient of each predictor variable from the LASSO output.

Predictor Variable	Coefficient	Standardized Coefficient
ln Depth [ln(cm)]	-2.8031	-0.47
Water Content (cm^3/cm^3)	13.9402	0.15
Soil Temperature (°C)	-0.2637	-0.04
Sand Content (%)	-0.0165	-0.05
Clay Content (%)	0	0
Bulk Density (g/cm^3)	-5.5493	-0.16
Rainfall (mm/year)	0	0
Air Temperature (°C)	0	0

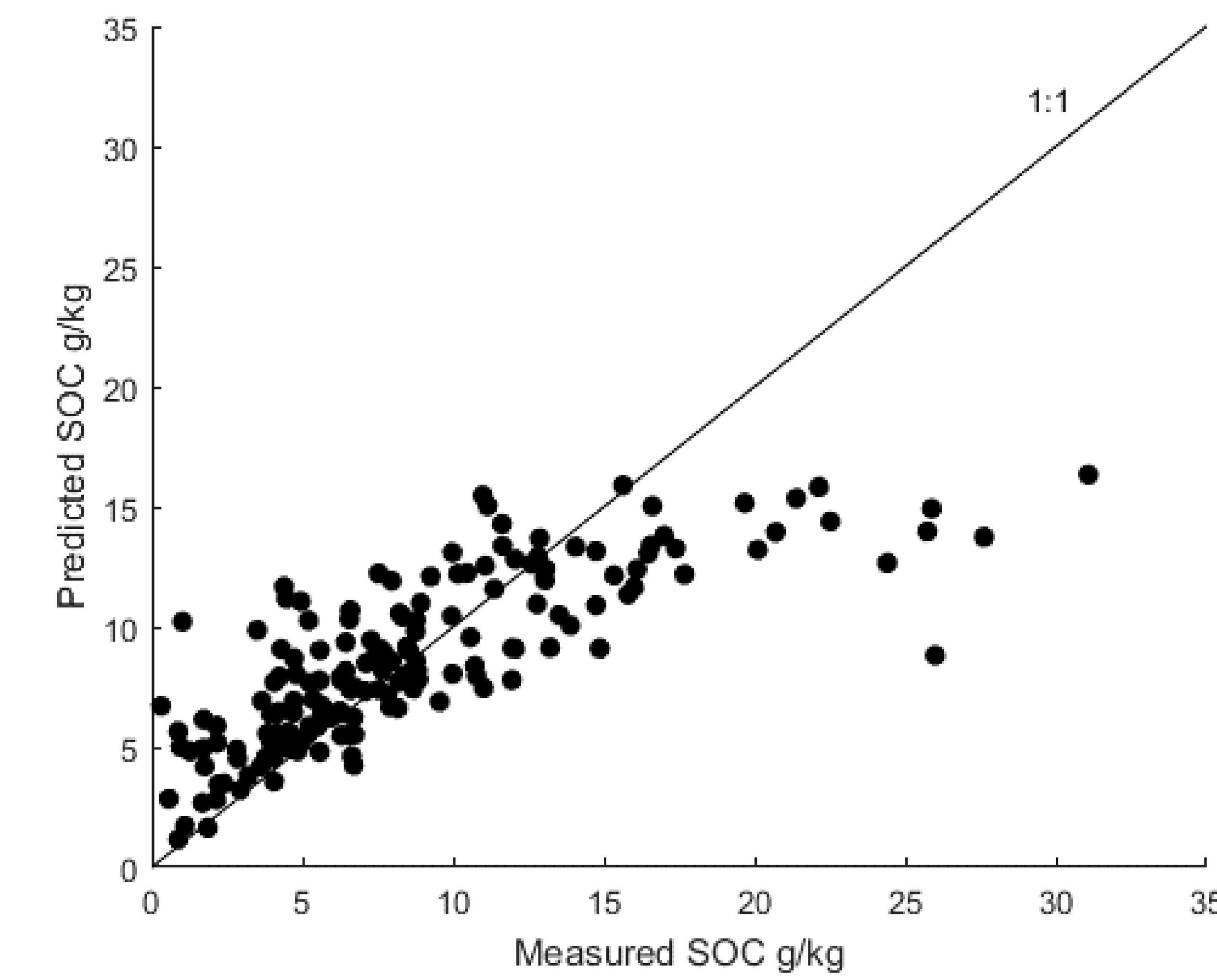

Figure 2. Graph of the measured versus predicted SOC., plotted against a 1:1 ratio line.

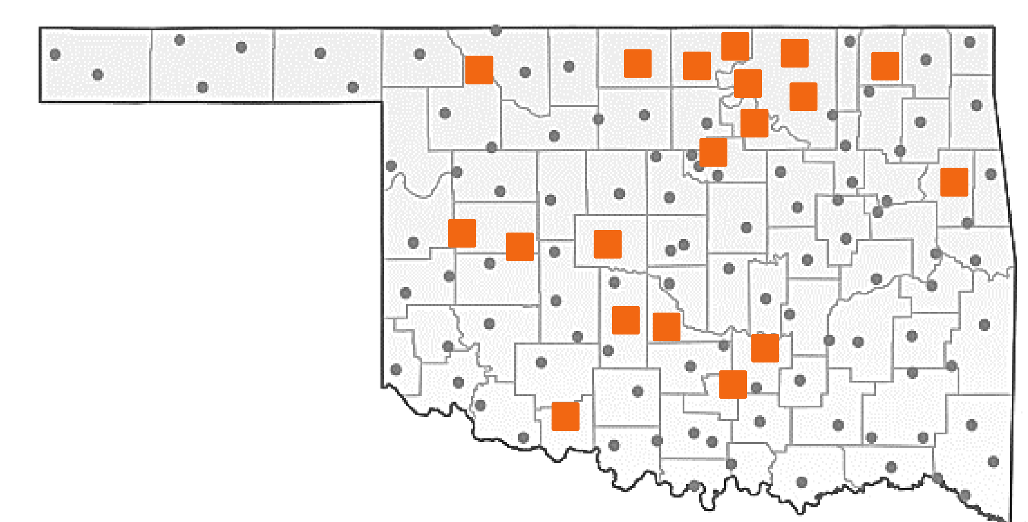

Figure 3. Map of Oklahoma Mesonet sites in gray. Sites with high SOC (>18 g/kg) are marked in orange.

Conclusions:

- LASSO pushed the coefficients of the clay content, rainfall, and air temperature variables to zero, meaning those variables were insignificant in this model.
- The standardized coefficients show depth, bulk density, and water content have the greatest influence when predicting SOC in this model.
- The model under predicts SOC > ~ 18 g kg^{-1} (Fig. 2) Figure 3 marks the site location of each of those high SOC measurements. The high SOC values may be associated with the type of vegetation (e.g. Tallgrass Prairie) or landscape position (e.g. flood plains).

Future work:

- Improve model to accurately predict higher SOC values.
- Incorporate additional predictor variables, such as vegetation and landscape position.