Evaluating Water Holding Capacity Across Spatial Scales with Neural Networks

Elissa Levine and Daniel Kimes

NASA-Goddard Space Flight Center
Code 923/Biospheric Sciences
Greenbelt, MD 20771
e-mail: elissa@ltpmail.gsfc.nasa.gov

Detailed soils characterization data is often stored in complex and large data sets that are difficult to aggregate and interpret. Aggregating or "scaling" this information from the pedon level to coarser generalizations requires subjective decisions by researchers that are not always consistent from one study to another. With the proper analysis tools, practical information about soil properties can be derived, and the relationship between models at different spatial resolutions can be defined. Neural networks have been shown to be a valuable tool for classifying or predicting soils information, and identifying complex nonlinear relationships in large data sets.

Neural networks "learn" the mathematical function of the relationship between a set of inputs and the desired output presented to them. Inputs are applied and weighted based on their "strength", and then are summed and passed through a transfer function to produce an output at each node within the network structure. Input and output variables can thus be related without any knowledge or assumptions about the underlying mathematical representation. They attempt to find the best nonlinear function, based on the network's complexity, without the constraint of linearity or pre-specified non-linearity used in regression and other traditional analyses.

Neural networks also can combine quantitative, descriptive, and ranked data in making predictions or classifications. This is especially useful when soil description, classification, and quantitative data are all available. In addition, when used with a genetic algorithm, neural networks are able to identify, from a large number of possible inputs, only those input variables that behave synergistically to produce the best network performance. In this way, better models can be derived by identifying those parameters most important for explaining the system.

In this study, neural networks were used to predict available water holding capacity across taxonomic classes. Water holding capacity is a critical measurement for many applications from field plot through global scales including understanding vegetation response to drought, responses to climate change, effects of land use on water dynamics, balancing global energy budgets, etc. The soils chosen for this study were part of the pedon database of the USDA, Natural Resources Conservation Service (NRCS). Data sets for each individual soil order were created, as well as a combined data set including all of the orders. Horizons from each profile in the data set were regarded as individual points in the data. A subset of quantitative, descriptive, and taxonomic parameters were chosen from the large number available from the data base.