Geostatistical spatial interpolation of soil water retention curve using PTFs based on particle size distribution

*Wataru Shiga[1], Yuji Kohgo[1], Hirotaka Saito[1]

1. Graduate School of Agriculture, Tokyo University of Agriculture and Technology

Predicting soil water retention curves (SWRC) or their model parameters at any unsampled locations using a geostatistical spatial interpolation technique requires a number of high quality retention data. Obtaining SWRC is, however, generally tedious, time consuming, and sometimes expensive. Therefore, pedotransfer functions (PTF), which allow one to predict soil hydraulic properties, such as SWRC, from easily measured soil properties, have been developed. In this study a geostatistical spatial interpolation technique was coupled with the PTF to predict water retention curves at given unsampled locations from available particle size distribution (PSD) data. Two PTFs are used in this study, one is the Arya and Paris (AP) model which predicts water retention curves from PSD and dry bulk densities, the other is based on the *k* -nearest neighbor (k-NN) algorithm which is a nonparametric method used in data mining. There are two approaches considered: (1) First, SWRC are predicted from PSD at given observed locations using one of the PTFs. SWRC are then predicted at given unsampled locations through the geostatistical spatial interpolation technique. This approach is referred to as the PTF-first and Interpolation-later approach or the PI approach. (2) First, PSD and the bulk densities are predicted at given unsampled locations using the geostatistical spatial interpolation technique. Then, SWRC are predicted at the unsampled locations by the PTFs from the interpolated PSD and bulk densities. This approach is then referred to as the Interpolation-first and PTF-later approach or the IP approach. Current study compares the performance of these two approaches to predict SWRC at any given unsampled location. The data used in this study were obtained from the Las Cruces trench site database, which contains water retention data for 447 soil samples. The dataset was then split into two sets, prediction and validation sets. This allows for the computation of prediction errors (root mean squared error or RMSE and mean absolute error or MAE). The results used AP model show that the performances of the PI and IP approaches were comparable. On the other hand, the results based on k-NN show that the PI arrroache outperformed the IP approach.

