Dear all, with a slight delay, here are the results of the using the 3D homosoil knn interpolation with idw. First, I have to say that I’m very satisfied with the results. The approach was tested on two case studies of mapping Organic Carbon, Bor and Edgeroi and in both case studies, homosil 3D knn has shown an improvement in prediction accuracy, about 5% in both case studies. comparison to the soil 3D knn.

3D soil interpolation implies using the observations approximately the same depth that belongs to the spatially neighboring profiles.

In this text, I will present what I did regarding the 3D soil interpolation by using homosoil knn with idw. Homosoil 3D knn do the interpolation of the 3D soil data observed at the soil horizons at the particular depth.

Methodology involves the following steps:

Step 1: Data preparing

Data preparation involves loading data, preparing data as soil profile collection, overlaying with spatial covariates grids, and preparing tidy data frame of soil data with the following structure: Each row represent one observation with the following structure: Profile ID, top, bottom, target variable observation, surface covariates, spatial coordinates (x, y). Such data frame is the input for functions.

Step 2: Splitting data in folds (outer and inner) suitable for nested cross validation. Nested cross-validation implies two times/levels data partitioning. First, the partitioning in the outer folds by using the procedure given in (Pejovic et. al 2018), and second, partitioning at the second level where each fold was additionally split by using the same procedure. In this way it was achieved that if we take one fold for testing, the remaining data was also split for applying cross-validation procedure for tuning parameters.

For this purpose, the R function “partitioning” was created.

Step 3: 3D soil interpolation by using “homosoil 3D knn” approach. Homosoil 3D knn (with idw) do the prediction at particular location in the 3D soil space (x, y, depth) by doing following steps:

1. Searching for the profiles which have observations within the prediction depth interval (prediction depth ± depth threshold).
2. From such subset of profiles, k-nearest profiles are selected based on the distance in the predictor space. For that purpose, the Gower distance was used.
3. Then, if at the k-nearest profiles, several soil horizons lie within the prediction depth interval, a weighted mean of the observations was calculated by using the portion of the horizon height that lies within the prediction depth interval as a weight. W = dh/depth\_th.
4. Then, the final prediction is calculated by using Inverse Distance Weighting approach which use the Gower distance instead of Euclidian distance.

Implementation.

I made R package soil3D\_knn. It consists of four functions. The first one is partitioning. The second one is soil3D\_knn. It performs the prediction based on the parameters which can be tuned. The tuning of the parameters is available via function tune\_soil3D\_knn. And finally, ncv\_soil3D\_knn performs nested cross-validation.