Geophysical Research Abstracts
Vol. 18, EGU2016-16089, 2016
EGU General Assembly 2016

3D-Digital soil property mapping by geoadditive models

Andreas Papritz
ETH Zurich, Institute of Biogeochemistry and Pollutant Dynamics, Department of Environmental Systems Science D-USYS, Zurich, Switzerland (papritz@env.ethz.ch)

In many digital soil mapping (DSM) applications, soil properties must be predicted not only for a single but for multiple soil depth intervals. In the GlobalSoilMap project, as an example, predictions are computed for the 0–5 cm, 5–15 cm, 15–30 cm, 30–60 cm, 60–100 cm, 100–200 cm depth intervals (Arrouays *et al.*, 2014). Legacy soil data are often used for DSM. It is common for such datasets that soil properties were measured for soil horizons or for layers at varying soil depth and with non-constant thickness (support).

This poses problems for DSM: One strategy is to harmonize the soil data to common depth prior to the analyses (e.g. Bishop *et al.*, 1999) and conduct the statistical analyses for each depth interval independently. The disadvantage of this approach is that the predictions for different depths are computed independently from each other so that the predicted depth profiles may be unrealistic. Furthermore, the error induced by the harmonization to common depth is ignored in this approach (Orton *et al.* 2016).

A better strategy is therefore to process all soil data jointly without prior harmonization by a 3D-analysis that takes soil depth and geographical position explicitly into account. Usually, the non-constant support of the data is then ignored, but Orton *et al.* (2016) presented recently a geostatistical approach that accounts for non-constant support of soil data and relies on restricted maximum likelihood estimation (REML) of a linear geostatistical model with a separable, heteroscedastic, zonal anisotropic auto-covariance function and area-to-point kriging (Kyriakidis, 2004.) Although this model is theoretically coherent and elegant, estimating its many parameters by REML and selecting covariates for the spatial mean function is a formidable task.

A simpler approach might be to use geoadditive models (Kammann and Wand, 2003; Wand, 2003) for 3D-analyses of soil data. geoAM extend the scope of the linear model with spatially correlated errors to account for nonlinear effects of covariates by fitting componentwise smooth, nonlinear functions to the covariates (additive terms). REML estimation of model parameters and computing best linear unbiased predictions (BLUP) builds in the geoAM framework on the fact that both geostatistical and additive models can be parametrized as linear mixed models Wand, 2003. For 3D-DSM analysis of soil data, it is natural to model depth profiles of soil properties by additive terms of soil depth. Including interactions between these additive terms and covariates of the spatial mean function allows to model spatially varying depth profiles. Furthermore, with suitable choice of the basis functions of the additive term (e.g. polynomial regression splines), non-constant support of the soil data can be taken into account. Finally, boosting (Bühlmann and Hothorn, 2007) can be used for selecting covariates for the spatial mean function. The presentation will detail the geoAM approach and present an example of geoAM for 3D-analysis of legacy soil data.

Arrouays, D., McBratney, A. B., Minasny, B., Hempel, J. W., Heuvelink, G. B. M., MacMillan, R. A., Hartemink, A. E., Lagacherie, P., and McKenzie, N. J. (2014). The GlobalSoilMap project specifications. In *GlobalSoilMap Basis of the global spatial soil information system*, pages 9–12. CRC Press.

Bishop, T., McBratney, A., and Laslett, G. (1999). Modelling soil attribute depth functions with equal-area quadratic smoothing splines. *Geoderma*, **91**(1–2), 27–45.

Bühlmann, P. and Hothorn, T. (2007). Boosting algorithms: Regularization, prediction and model fitting. *Statistical Science*, **22**(4), 477–505.

Kammann, E. E. and Wand, M. P. (2003). Geoadditive models. *Journal of the Royal Statistical Society. Series C: Applied Statistics*, **52**(1), 1–18.

Kyriakidis, P. (2004). A geostatistical framework for area-to-point spatial interpolation. *Geographical Analysis*, **36**(3), 259–289.

Orton, T., Pringle, M., and Bishop, T. (2016). A one-step approach for modelling and mapping soil properties based on profile data sampled over varying depth intervals. *Geoderma*, **262**, 174–186.

Wand, M. P. (2003). Smoothing and mixed models. *Computational Statistics*, **18**(2), 223–249.