



What is the best spatial soil inference system for mapping Available water capacity? A test in Languedoc-Roussillon (France)

Quentin Styc and Philippe Lagacherie

LISAH, Univ Montpellier, INRA, IRD, Montpellier SupAgro, Montpellier, France

Available Water Capacity (AWC) - i.e. the maximum quantity of water available for plant growth - is one of the soil property that is the most requested by users, including modellers. Yet, there is still few Digital Soil Mapping research that focus on this soil property. Since AWC is a composite soil property that involves several primary soil properties and several depth intervals, their mapping requires to select the most appropriate soil spatial inference system (Lagacherie & McBratney, 2007) between i) computing AWC at each observed site first and then build an unique DSM model for AWC or ii) building first DSM models for each of the involved soil properties and depth intervals then combining the maps for obtaining an AWC map or iii) selecting any of the possible soil spatial inference systems that may be envisaged between the two previous extreme options.

In this study we considered six possible soil spatial inference systems (SSIS) for mapping AWC over the Languedoc-Roussillon. These systems used Quantile Random Forests (Meinshauzen, 2006) and Survival Random Forests (Ishwaran et al., 2008) as DSM models, relief, lithology, climate and land use spatial data as soil covariates and 2024 legacy measured soil profiles as inputs for calibrating the DSM models. Each Soil Spatial Inference Systems outputs were validated by applying 20 times a 10-fold cross validation.

Performances of the six possible SSIS ranged from $R^2 = 0.04$, RMSE = 6.28 cm to $R^2 = 0.72$ RMSE = 3.35 cm, the best performance being obtained by mapping first a weighted mean of the primary soil properties involved in the calculation of AWC. Those results are discussed considering two antagonist consequences when mapping pooled soil properties: i) it smooths the variability of each soil properties and thus makes the mapping easier, and ii) it does not take into account the differences in the landscape properties that drive the variability of each soil property, which makes the mapping less efficient.

References

Ishwaran, H., Kogalur, U.B., Blackstone, E.H., Lauer, M.S., 2008. Random survival forests. *Ann. Appl. Stat.* 2, 841–860.

Lagacherie, P., McBratney, A.B., 2007. Spatial Soil Information Systems and Spatial Soil Inference Systems: Perspectives for Digital Soil Mapping, in: Lagacherie, P., McBratney, A.B., Voltz, M. (Eds.), *Digital Soil Mapping: An Introductory Perspective*. Elsevier, pp. 3–22. Elsevier, Amsterdam.

Meinshausen, N., 2006. Quantile Regression Forests. *J. Mach. Learn. Res.* 7, 983–999.