

Some methods for the quantification of prediction uncertainties for digital soil mapping

Soil Security Laboratory

2018

Soil scientists are quite aware of the current issues concerning the natural environment because our expertise is intimately aligned with their understanding and alleviation. We know that sustainable soil management alleviates soil degradation, improves soil quality and will ultimately ensure food security. Critical to better soil management is information detailing the soil resource, its processes and its variation across landscapes. Consequently, under the broad umbrella of “environmental monitoring”, there has been a growing need to acquire quantitative soil information (McBratney et al., 2003; Grimm and Behrens, 2010). The concerns of soil-related issues in reference to environmental management were raised by McBratney (1992) when stating that it is our duty as soil scientists, to ensure that the information we provide to the users of soil information is both accurate and precise, or at least of known accuracy and precision.

However, a difficulty we face is that soil can vary, seemingly erratically in the context of space and time (Webster, 2000). Thus the conundrum in model-based predictions of soil phenomena is that models are not “error free”. The unpredictability of soil variation combined with simplistic representations of complex soil processes inevitably leads to errors in model outputs.

We do not know the true character and processes of soils and our models are merely abstractions of these real processes. We know this; or in other words, in the absence of such confidence, we know we are uncertain about the true properties and processes that characterize soils (Brown and Heuvelink, 2005). The key is therefore to determine to what extent our uncertainties are propagated through a model of which effect the final predictions of a real-world process.

In modeling exercises, uncertainty of the model output is the summation of the three main sources generally described as: model structure uncertainty, model parameter uncertainty and model input uncertainty (Minasny and McBratney, 2002; Brown and Heuvelink, 2005). A detailed analysis of the contribution of each of the different sources of uncertainty is generally recommended. In this book chapter we will cover few approaches to estimate the uncertainty of model outputs. Essentially what this means is that given a defined level of confidence, model predictions from digital soil mapping will be co-associated with the requisite prediction interval or range. The approaches for quantifying the prediction uncertainties are:

-
- Universal kriging prediction variance.
 - bootstrapping
 - Empirical uncertainty quantification through data partitioning and cross validation.
 - Empirical uncertainty quantification through fuzzy clustering and cross validation

The data that will be used in this chapter is a small data set of subsoil pH that has been collected since 2001 to present from the Lower Hunter Valley in New South Wales, Australia. The soil data covers an area of approximately 220km². Validation of the quantification of uncertainty will be performed using a subset of these data. The mapping of the uncertainties will be conducted for a small region of the study area. The data for this section can be retrieved from the `ithir` package. The soil data is called `HV_subsoilpH` while the grids of environmental covariates is called `hunterCovariates_sub`.

References

- Brown, J. D. and G. B. M. Heuvelink
2005. *Encyclopaedia of Hydrological Sciences*, chapter Assessing uncertainty propagation through physically based models of soil water flow solute transport. John Wiley and Sons, Chichester.
- Grimm, R. and T. Behrens
2010. Uncertainty analysis of sample locations within digital soil mapping approaches. *Geoderma*, 155:154–163.
- McBratney, A. B.
1992. On variation, uncertainty and informatics in environmental soil management. *Australian Journal of Soil Research*, 30:913–935.
- McBratney, A. B., M. L. Mendonca Santos, and B. Minasny
2003. On digital soil mapping. *Geoderma*, 117:3–52.
- Minasny, B. and A. B. McBratney
2002. Uncertainty analysis for pedotransfer functions. *European Journal of Soil Science*, 53:417–429.
- Webster, R.
2000. Is soil variation random? *Geoderma*, 97:149–163.