



About the information content of spatial data – Improving the performance of boosted regression trees by using the differential evolution algorithm for parameter tuning

Anika Gebauer (1) and Mareike Ließ (2)

(1) Helmholtz Centre for Environmental Research - UFZ, Halle (Saale), Germany (anika.gebauer@ufz.de), (2) Helmholtz Centre for Environmental Research - UFZ, Halle (Saale), Germany (mareike.liess@ufz.de)

In certain soils data regionalisation remains a challenge. The organic carbon stored in soils plays a key role in reducing the atmospheric greenhouse gas concentration. Dry forest ecosystems in Ecuador are endangered to become a source of carbon emissions, because of deforestation. Nevertheless, these ecosystems' carbon stocks are particularly underexplored. To investigate the soil organic carbon stocks of a 16 km² dry forest in south-western Ecuador a limited number of soil profiles were sampled and spatial predictors were generated from a digital elevation model and spectral images. To exhaustively analyse the relationship between sampled soil properties and predictors, the differential evolution algorithm was used to optimise the parameter values of a boosted regression trees model. The objective function was optimised for both, RMSE and R², in order to obtain reasonable prediction results. However, the models, built on optimised model parameter values, still explained only up to 40% of the target variables' variance. Random effects, for example induced by animals or tree falls, were assumed to distort the relationship between sampled soil properties and the predictor variables.