



Evaluating different machine learning approaches for the interpolation of ambient air temperature at Mt. Kilimajaro, Tanzania

Tim Appelhans (1), Ephraim Mwangomo (1,4), Douglas Hardy (2), Andreas Hemp (3), and Thomas Naus (1)
(1) Philipps Universität Marburg, Geographie, Umweltinformatik, Marburg, Germany, (2) Climate System Research Centre and Department of Geosciences, University of Massachusetts Amherst, USA, (3) Plant physiology, Bayreuth University, Germany, (4) Tanzania National Parks, Arusha, Tanzania

Spatially high resolution climate information is required for a variety of applications in but not limited to functional biodiversity research. In order to scale the generally plot-based research findings to a landscape level, spatial interpolation methods of meteorological variables are required. Based on a network of 60 observation plots across the southern slopes of Mt. Kilimanjaro, the skill of 14 machine learning algorithms in predicting spatial temperature patterns is tested and evaluated against the heavily utilized kriging approach. Based on a leave-many-out testing design, regression trees generally perform better than linear and non-linear regression models. The best individual performance has been observed by the Cubist model followed by stochastic gradient boosting, random forest and model averaged neural networks which except for the latter are all regression tree-based algorithms. While these machine learning algorithms perform better than kriging in this quantitative evaluation, the overall visual interpretation of the resulting air temperature maps is ambiguous. Here, a combined Cubist and residual kriging approach might be the best solution.