ARTICLE IN PRESS

Spatial Statistics ▮ (▮▮▮▮) ▮▮▮–▮▮▮

Contents lists available at ScienceDirect

Spatial Statistics

journal homepage: www.elsevier.com/locate/spasta

Evaluating machine learning approaches for the interpolation of monthly air temperature at Mt. Kilimanjaro, Tanzania

Tim Appelhans[a,*], Ephraim Mwangomo[a,b], Douglas R. Hardy[c], Andreas Hemp[d], Thomas Nauss[a]

[a] *Environmental Informatics, Philipps University of Marburg, Germany*

[b] *Tanzania National Parks, Arusha, Tanzania*

[c] *Climate System Research Centre and Department of Geosciences, University of Massachusetts Amherst, USA*

[d] *Plant physiology, Bayreuth University, Germany*

ARTICLE INFO

Article history:
Received 31 October 2014
Accepted 28 May 2015
Available online xxxx

Keywords:
Spatial interpolation
Machine learning
Air temperature
Kriging
Cubist
Cross-validation

ABSTRACT

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 temperature 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 10-fold cross-validation testing design, regression trees generally perform better than linear and non-linear regression models. The best individual performance has been observed by the stochastic gradient boosting model followed by Cubist, 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 a quantitative evaluation, the overall visual interpretation of the resulting air temperature maps is ambiguous. Here, a combined Cubist and residual kriging approach can be considered the best solution.

* Corresponding author.
E-mail address: tim.appelhans@gmail.com (T. Appelhans).

http://dx.doi.org/10.1016/j.spasta.2015.05.008

1. Introduction

For biodiversity and ecosystem research, climate conditions are a major explanatory variable (e.g. Sala, 2000; Hawkins et al., 2003; Currie et al., 2004) and a common demand of biodiversity researchers is to get plot-scale information on the weather and climate conditions. The problem is complicated by the fact that the world wide weather station network is at its minimum or virtually absent in the regions of the world's biodiversity hotspots (see Myers et al., 2000) and that regional climate model simulations or satellite observations are generally to coarse (or less frequent) to fully meet the demands of the biodiversity community regarding plot-scale observations. So basically, the question to be addressed is how we can provide accurate weather status information for biodiversity and ecosystem research which often demands site specific information for certain intensively investigated research plots and area-wide information on a landscape scale.

The common solution for this problem is the installation of individual stations by the respective researchers working in the area. However, due to funding and man power restrictions, the stations can generally not be installed on each of the biodiversity observation plots (e.g. Fries et al., 2009). But even if this would be possible, area-wide weather and climate information is mandatory if the individual plot-based findings are to be transferred to the landscape scale. This necessitates the application of either spatial interpolation or downscaling techniques which potentially provide spatially high resolution weather datasets based on individual station observations or medium to high-resolution climate model simulations and/or meteorological satellite observations. Even though the latest CORDEX climate model runs for e.g. the African domain (Panitz et al., 2014) or the brand new microclim dataset (Kearney et al., 2014) have a resolution of 0.22° and 0.17° respectively, the grid edge length at the equator of roughly 24.5 km and accordingly 15 km is still to coarse to cover the local to landscape scale patterns relevant for many research approaches (e.g. the Mt. Kilimanjaro region is represented in the microclim dataset by 12 grid cells). Similar limitations hold true for the application of satellite observations which drop beyond about a 1 km by 1 km resolution as soon as the temporal resolution is daily or better. For the Kilimanjaro region as for any other mountain system the situation is further complicated because of the modification of meso-scale weather by micro-topographic site conditions (Loeffler et al., 2006). Hence, spatial interpolation or prediction techniques are still required to derive the high resolution meteorological fields demanded by functional biodiversity research.

An amplitude of studies exists on the utilization of different spatial interpolation methods for meteorological parameters across different regions. Until recently and neglecting quite specialized approaches like PRISM (Daly, 2006; Daly et al., 2007) or DAYMET (Thornton et al., 1997), these methods could be divided into simpler methods like distance weighting (e.g. Lennon and Turner, 1995; Willmott and Matsuura, 1995; Nalder and Wein, 1998) or polynomial interpolations (e.g. Tabios and Salas, 1985; Ashraf et al., 1997; Goodale et al., 1998; Xia et al., 1999; Ninyerola et al., 2000) and the more advanced geostatistical interpolation techniques of kriging and splining. The various forms of kriging (Krige, 1951) use linear weighting combinations at the known data points to predict the parameter of interest at points where no measurements are available (e.g. Holdaway, 1996; Ashraf et al., 1997; Diodato, 2005). Compared to kriging where the statistical model has to be subjectively selected, splining, i.e. fitting splines to the known data points is less dependent on the underlying statistical model but on the other hand it requires regularly spaced input data (Hulme et al., 1995; Hutchinson, 1995; Price et al., 2000; Xia et al., 2001).

Several evaluation studies of such geostatistical approaches have been carried out by various authors indicating that kriging produces generally more accurate results than other interpolation techniques (e.g. Ashraf et al., 1997; Goovaerts, 2000; Apaydin et al., 2004; Ustrnul and Czekierda, 2005; Chen et al., 2007; Hofstra et al., 2008). Even if kriging does not perform best with respect to typical validation indices (e.g. mean square error), the resulting interpolation fields might be more plausible than the ones from other techniques (e.g. Collins and Bolstad, 1996). For applications in regions with highly complex topography, different ancillary data sources like digital elevation models or land-use classifications have been used (e.g. Jarvis and Stuart, 2001; Hasenauer, 2003; Stahl et al., 2006; Baltas, 2007; Daly et al., 2007; Guler et al., 2007; Di Luzio et al., 2008) which are especially important for temperature interpolation techniques (e.g. Vicente-Serrano et al., 2003) or the derivation of sheltering

Please cite this article in press as: Appelhans, T., et al., Evaluating machine learning approaches for the interpolation of monthly air temperature at Mt. Kilimanjaro, Tanzania. Spatial Statistics (2015), http://dx.doi.org/10.1016/j.spasta.2015.05.008