Purposive sampling across different data sources for a robust forest risk assessment by species distribution modelling

Authors: Tobias Mette[1*];

[1] Bavarian State Institute for Forestry , LWF, Freising, Germany

*corresponding author: tobias.mette@lwf.bayern.de;

Abstract: Climate tolerance is of fundamental importance when assessing the risk affiliated with tree species cultivation decisions in climate change. A common approach is to use species distribution models (SDMs) to estimate the climatic niche of a tree species. The optimum data sources for SDMs are area-representative harmonic inventories over the entire distribution area of a species. But even large scale inventories like the US forest inventory or the pan-European forest condition survey do not necessarily cover the entire areal of a species in question. Also, the survey grid is mostly too coarse to capture the distribution margins of a species adequately, all the more of rare species. Hence, the modeller typically faces the challenge to (1) combine different data sources in such a way that (2) data-related distortions are minimized and (3) the target risk variable derived from the species distribution is robust regardless of a species' overall abundance. One key option to handle these problems is to adjust one's sampling design to the purpose of the SDM.

Since our purpose is a robust tree species cultivation risk we compare the effect of different sampling methods on species distribution models of six European tree species. Input data are national and pan-European forest inventories in combination with species distribution maps. The sampling methods range from flat to stratified sampling in the geographic and parameter data realm and different degrees of absence discrimination. As for the SDM we chose high-penalty generalized additive models using temperature and precipitation as input variables. The results across the different sampling methods prove more stable than expected especially for the common tree species. Yet, a stratified sampling in the climate parameter space with moderate absence discrimination to be most robust against heterogeneous data background and species abundance.

The interpretation of species distribution in terms of cultivation risk is an ongoing debate. Understanding the effect of sampling can account for methodology-related differences between SDMs and help focussing the discussion for instance on trait variances along the climatic gradient.