



Distribution models for 31 vegetation types in Norway

Peter Horvath (1,2), Rune Halvorsen (1), Frode Stordal (2), Hui Tang (2), and Anders Bryn (1)

(1) Natural History Museum, University of Oslo, Oslo, Norway, (2) Department of Geosciences, University of Oslo, Oslo, Norway

Climate models now incorporate complex vegetation modules to describe processes and interactions between land and atmosphere. However, the representation of boreal and arctic ecosystems in climate models is still burdened with considerable uncertainty and there is a clear need for more precise vegetation data for parameterisation and validation of climate models. These vegetation data have to have wall-to-wall coverage of the model area in order to be applicable for climate modelling. At present, only coarse vegetation maps based on remote sensing data have such coverage. In this study, we explore an alternative pathway to area-covering vegetation data, using distribution modelling (DM) to predict the current distributions of vegetation types in Norway from environmental information.

The vegetation data set used for parameterisation of the models consists of presence and absence of 31 different vegetation types (VT), derived from an extensive area frame survey of land cover and outfield land resources carried out by the Norwegian Institute of Bioeconomy Research (NIBIO). We use logistic regression, a type of Generalized Linear Models (GLM), to construct the models which are, in turn used to predict the probability of presence of each VT in every 100x100-m grid cell in Norway. A forward selection procedure with an F-ratio test was applied to 117 predictor variables representing different categories (climatic, snow, topography, geology and land cover).

All models were evaluated by calculation of the performance statistic AUC (Area Under the Receiver Operator Characteristic curve), using an independently collected vegetation dataset. AUC values for the 31 models suggest that VTs can be predicted with quite high degree of certainty. A total of 21 models had $AUC > 0.8$. The highest AUC value, $AUC = 0.989$, was obtained for is VT 5ab (broadleaf deciduous forest), whereas the lowest AUC value, $AUC=0.679$, was obtained for VT 11b (pastures). We found that AUC values of our models decreased significantly with the prevalence of VTs in the test data set, with average distance between test observations, and with distance to coast. Heath and mountain VTs are significantly better predicted than forest and wetland types. The reasons for these relationships are discussed.

Our study demonstrates that distribution modelling is a promising tool for spatial prediction of aggregated groups of species such as vegetation types, provided relevant predictor variables are available. We propose that predictions from such models will improve parametrization of vegetation modules of climate models.