MoGLI:
Potential Distribution Maps for Swiss Woody Species

a summary of methods and results

Rafael O. Wüest, Ariel Bergamini, Kurt Bollmann, Andri Baltensweiler

Correspondence to: Andri Baltensweiler (andri.baltensweiler@wsl.ch)

We used Swiss National Forest Inventory (NFI) data to model the potential distribution of the most common woody species for the forested area of Switzerland and provide potential distribution maps that fulfill minimal quality criteria with regard to predicting performance. Distribution maps were produced based on the principles of species distribution modeling, where the known occurrence of species from the NFI is statistically related to environmental predictors. The potential distribution was then be predicted for the entire forested area of Switzerland using the derived statistical relationship. Apart from traditional predictors such as climate or topography, MoGLI developed new predictors that describe forest structure using LiDAR (light detection and ranging; a method that scans the earth surface using laser-technology), and also used soil property maps to improve the accuracy of the spatial projections.

The resulting maps can be viewed in a simple web-GIS application available at: https://www.lfi.ch/produkte/mogli/mogli-en.php

Methods

The exact procedure is divided into three steps: occurrence and environmental data processing, statistical modelling, and spatial prediction.

Occurrence and Environmental Data

On the one hand, we used occurrence data of woody species from the Swiss National Forest Inventory (NFI). On the other hand, the project used information on environmental conditions available over the entire forested area of Switzerland. In addition to information on climate, topography and soil pH, satellite and LiDAR data were used for modelling. We used NDVI (normalized difference vegetation index) satellite data from the LANDSAT mission because they provide information on forest productivity. LiDAR data have been used to derive information on the vertical forest structure and can serve as a proxy to describe light conditions in forest stands. It has been shown that such structural features can especially improve the model quality for (shrub) species that require light such as the common juniper (*Juniperus communis* L. subsp. *communis*) and the blackthorn (*Prunus spinosa* L.), see Wüest et al. (2020) for more details.

Modeling Procedure

Five different statistical models (GLMs, GAMs, MaxEnt, random Forests, artificial neural networks) were used to explain the presence (or absence) of species at the NFI sampling locations. The quality of the models was checked by cross-validating for each species how well the models predict the occurrence (or absence) for the respective species. In cross-validation, the NFI data set was randomly divided into 70% training and 30% test data. All five models were then generated with the training data, and the predictions of these models were compared with the actual observed occurrences/absences in the test data. The quality of the predictions was assessed using the True Skills Statistics (*TSS*). This procedure was repeated a hundred times, and an average *TSS* value was calculated for each species. *TSS* values can vary between zero (worst possible model) and one (perfect model). Maps for species with insufficient model quality (*TSS_{CV}* < 0.5) have not been published to ensure that only reliable species maps are available for download.

Creating Maps

In the final step, the statistical models were used to predict the occurrence of the woody species for the entire Swiss forest area. The MoGLI project provides three products. The first product, the so-called ensemble map, is the average predicted probability of occurrence of a species. For this, the predicted values from the five statistical models were averaged, resulting in values between zero and one (the ensemble probability maps are provided in percent ranging from 0% to 100%). The second product provides information on how much the models differ in their predictions. This was calculated as the standard deviation across the five models, with large values indicating high discrepancies between the models (the standard deviation maps are provided as standard deviations multiplied by 100). The third product, so-called consensus maps, provides information about both the probability of occurrence and the uncertainty associated with the forecast. This was achieved by providing three classes, whereby two classes are assigned low uncertainty and indicate either the occurrence or absence of a species, while the third class does not provide any information on the potential current occurrence due to too much uncertainty. The consensus maps contain values ranging from one to three that indicate:
1 := occurrence unlikely
2 := occurrence uncertain
3 := occurrence likely

Results

Prediction Quality

Table 1 gives Details on the number of presences ($N_{presences}$) that were observed in the NFI. It also indicates how many climate (N_{clim}), terrain attribute (N_{terat}), linear feature (N_{lin}), and remotely sensed predictor variables (N_{rs}) were used. N_{total} declares the total number of predictor variables that the models used. TSS_{CV} indicates the average TSS_{CV} value in cross validation (where $TSS_{CV} = 0$ would indicate the worst possible model and $TSS_{CV} = 1$ would indicate a perfect model).

Table 1 *Modeled species for which the model quality was above the given threshold (TSScv > 0.5). For all of these species potential distribution maps are available (https://www.envidat.ch/dataset/mogli-sdm). The table further includes the number of presences of each species in the NFI data ($N_{presences}$), the total number of variables used for modeling the species (N_{total}) the number of variables used per category (climate: N_{clim}; terrain attributes: N_{terat}; linear distance features: N_{lin}; remote sensing: N_{rs}), and the average TSS of all cross-validation runs across all models (TSS_{CV}).*

Species	$N_{presences}$	N_{clim}	N_{terat}	N_{lin}	N_{rs}	N_{total}	TSS_{CV}
Abies alba Mill.	2307	20	33	6	9	68	0.524
Acer campestre L.	275	13	5	7	2	27	0.642
Acer opalus Mill.	73	5	1	0	1	7	0.787
Acer platanoides L.	236	12	3	5	3	23	0.508
Alnus glutinosa (L.) Gaertn.	85	5	3	0	0	8	0.600
Alnus incana (L.) Moench	275	7	17	0	3	27	0.528
Alnus viridis (Chaix) DC.	325	12	11	5	4	32	0.677
Amelanchier ovalis Medik.	33	2	0	0	1	3	0.639
Berberis vulgaris L.	175	12	2	0	3	17	0.622

Carpinus betulus L.	282	12	10	5	1	28	0.720
Castanea sativa Mill.	211	13	3	0	5	21	0.872
Clematis vitalba L.	230	12	5	4	2	23	0.563
Cornus sanguinea L.	389	14	15	6	3	38	0.636
Cotoneaster tomentosus Lindl.	53	3	0	0	2	5	0.603
Crataegus monogyna Jacq.	336	13	9	6	5	33	0.555
Crataegus laevigata (Poir.) DC.	207	13	3	4	0	20	0.557
Cytisus scoparius (L.) Link	57	5	0	0	0	5	0.887
Daphne laureola L.	61	4	2	0	0	6	0.704
Euonymus europaeus L.	245	9	8	7	0	24	0.657
Fagus sylvatica L.	3142	20	34	7	10	71	0.669
Fraxinus excelsior L.	2006	17	36	8	9	70	0.542
Hedera helix L.	907	15	35	6	10	66	0.622
Hippocrepis emerus (L.) Lassen	32	2	1	0	0	3	0.606
Ilex aquifolium L.	359	13	11	5	6	35	0.548
Juglans regia L.	189	11	3	4	0	18	0.563
Juniperus communis L. subsp. *communis*	151	8	3	0	4	15	0.615
Juniperus communis subsp. *alpina* Celak.	130	5	1	3	4	13	0.746
Laburnum alpinum (Mill.) Bercht. & J. Presl	36	2	1	0	0	3	0.551
Laburnum anagyroides Medik.	41	4	0	0	0	4	0.643
Larix decidua Mill.	1019	19	35	8	10	72	0.605
Ligustrum vulgare L.	294	12	7	7	3	29	0.669
Lonicera caerulea L.	62	5	0	1	0	6	0.504
Lonicera xylosteum L.	1179	15	33	7	10	65	0.509
Ostrya carpinifolia Scop.	32	3	0	0	0	3	0.853
Picea abies (L.) H. Karst.	4175	18	36	8	10	72	0.528
Pinus cembra L.	186	12	1	3	2	18	0.874
Pinus mugo subsp. *uncinata* (DC.) Domin	92	5	0	1	3	9	0.696
Pinus mugo Turra subsp. *mugo*	50	3	0	0	2	5	0.790
Pinus sylvestris L.	490	13	21	6	9	49	0.529
Prunus avium L.	610	14	31	6	10	61	0.510
Prunus padus L.	157	8	6	1	0	15	0.612
Prunus spinosa L.	153	9	2	4	0	15	0.551
Pseudotsuga menziesii (Mirb.) Franco	84	3	5	0	0	8	0.638
Quercus petraea Liebl.	295	15	5	5	4	29	0.581
Quercus pubescens Willd.	69	6	0	0	0	6	0.783
Quercus robur L.	344	10	18	6	0	34	0.607
Rhododendron ferrugineum L.	369	10	14	6	6	36	0.782
Ribes alpinum L.	83	6	2	0	0	8	0.579
Robinia pseudoacacia L.	36	3	0	0	0	3	0.681
Sorbus chamaemespilus (L.) Crantz	36	3	0	0	0	3	0.614
Taxus baccata L.	96	5	3	0	1	9	0.529
Tilia cordata Mill.	237	14	2	2	5	23	0.525

Ulmus glabra Huds.	476	16	18	6	7	47	0.526
Vaccinium myrtillus L.	369	14	9	5	8	36	0.575
Viburnum lantana L.	378	11	13	6	7	37	0.532
Viburnum opulus L.	240	10	8	6	0	24	0.531

References

Wüest, R. O., Bergamini, A., Bollmann, K., & Baltensweiler, A. (2020). LiDAR data as a proxy for light availability improve distribution modelling of woody species. *Forest Ecology and Management*, *456*, 117644. https://doi.org/10.1016/j.foreco.2019.117644