Assessing the predictive power of canopy height as a new predictor of plant distribution between grasslands and forests in the Swiss Western Alps

Plancherel C., Department of Ecology and Evolution, University of Lausanne.

1st step project, directed by Antoine Guisan, supervised by Daniel Scherrer

Master in Behaviour, Evolution and Conservation, University of Lausanne, 2016

Céline Plancherel

celine.plancherel@unil.ch

Abstract

Species distribution models (SDMs) are widely used in ecology and biogeography. Indeed, species distributions will shift under global warming and landuse change. Studying these movements is crucial to conservation efforts. Despite their wide use, improvements are still possible in SDMs by adding or refining relevant predictors. Here, we added canopy height as a predictor in topo-climatic models for plant species to see if it could improve their predictive power in terms of distinction between plants from grassland, plants from forest and plants found in both habitat. By analysis of four modelling techniques and four thresholding methods, we show that canopy height (plots' maximum) is an important predictor for the distribution of individual species. For the majority of species, canopy height was almost as important as seasonal information (degree-days), improving the predictive performance significantly. However, it does not allow better predictions for community composition, i.e. no better separation between different habitat specific sets of species.

Keywords: Species distribution model (SDMs), new predictor, continuous landscape, mountains, plants, community assembly, variable importance

Introduction

Understanding plant species distributions and above all predicting them are major subjects in ecology nowadays. Indeed, one can easily imagine the potential of such techniques in a context of global warming and conservation issues (Guisan & Thuiller 2005; Guisan & Theurillat 2000; Maiorano *et al.* 2011; Descombes *et al.* 2016). One important tool to this end, are species distribution models (SDMs). SDMs are more and more used and studied in ecology and benefits from regular improvements since it was created in the 70s (Nix *et al. 1977,* in Guisan & Thuiller 2005; Guisan & Thuiller 2005). However, there are still some important limitations and lacks that are currently under discussion such as the choice and improvement of predictors used in models, or their biological significance and meaning (Guisan & Thuiller 2005; Pottier *et al.* 2013; Pradervand *et al.* 2014*;* Mod *et al.* 2016).

When constructing SMDs, different steps are necessary. We have first the conceptualization, followed by the data preparation, the model fitting, its evaluation, then making spatial prediction and finally the assessment of model applicability (Guisan & Thuiller 2005). SDMs are based on the niche concept (i.e. that the realized niche is constant across space and time). Also, there are three principal kinds of influences that can be taken into account during conceptualization: limiting factors, disturbances and resources (Guisan & Thuiller 2005). The idea is to improve models by adding or correcting factors in these three categories. For instance, Dubuis *et al.* (2013) tried to improve “prediction of plant species and community composition by adding edaphic to topo-climatic variables”. In other words, they added new predictors that could be considered as limiting factors or resources depending on the point of view, and assessed their improvement on models. In a similar way, Pottier *et al.* (2013) evaluated the predictive power of SDMs along elevation gradient. Without really changing the predictors, they refined the model to evaluate it along a gradient.

Here, the idea is kind of mixing these two concepts by adding a new predictor to assess models’ predictive power along a gradient. Indeed, one important gradient not taken into account so far is the idea of continuous landscape between grassland and forest. For the moment, forest is often considered as a discrete variable (Mathys 2007). In concrete terms, models applied in forest and grassland ecosystem contain a (subjective) “binary” mask to strictly differentiate grassland and forest. However, this procedure raises problems, as some plants present in both habitat are often overpredicted or underpredicted in such model (as we merge dataset from forest and grassland). Indeed, we lose much information about this continuous variable when we

separate it in two categories, especially in intermediate cases, i.e. ecotones (Mathys 2007). Therefore, a more continuous variable is needed in SDMs to improve the representation of the forest to grassland gradient. We know that tree canopy cover, tree canopy height and stand width are three variables that vary along this gradient (Mathys 2007). Therefore, at least one of them should be a potential candidate for our purpose. Of the three, canopy height is most easily available for large spatial extents by remote sensing (LiDAR) and therefore seems to be the most logical choice to differentiate grassland from forest communities. Indeed, “the taller tree canopy height classes are better represented by forest than non-forest samples [and] the frequency of non-forest plots declined almost linearly with increasing tree canopy height” in a study based in the Jura mountains (Switzerland) (Mathys 2015).

Here, we evaluated the importance of the variable “canopy height” at different levels (maximum, median and minimum for each plot) to explain the distribution of 333 species that we can find in forest, grassland or both. The methodology we used is very similar to the one from Dubuis *et al.* (2013) study as the global idea is the same (assessing the predictive power of a new predictor in a model). Thus, we (1) made PCAs without any modelling results to see if canopy height was suitable to allow a better separation of forest and grassland species, (2) we used canopy height in single species SDMs to see if it improves the predictive power of individual species and (3) we created S-SDMs (stacked-species distribution models) to see if we can improve community predictions.

Materiel & methods

Study area

Our study area is a part of about 700 km^2 of the Western Swiss Alps (Canton de Vaud, Switzerland, 46°10′–46°30′ N; 6°50′–7°10′ E; fig. 1). It has an elevation from 375m asl. in Montreux to 3210m asl. on the top of the Diablerets massif and a temperate climate. Vegetation is known to be highly influenced by human activities and we often find pastures or meadows in deforested areas (Randin *et al.* 2009; Dubuis *et al.* 2013). We also find typical species from calcareous Alps along the altitudinal gradient (Randin *et al.* 2006).

Species data

Data on the presence-absence of species were provided from different older studies (for example, Scherrer *et al.* (in review) for forest species; D’Amen *et al.* 2015 for

grassland species). We have a total of 3989 sampling points (fig. 1). 3076 plots come from forested areas (in the broad sense; Hartmann *et al.* 2009), and they have a plot area of 314 m^2 (r=10 m). 913 plots come from open grassland areas (Randin *et al.* 2009) and they have a plot area of 64 m^2 (8x8 m). Also, the two subsets have different sampling strategies (grid based forest, random stratified for grasslands) (for details on sampling see Hartmann *et al.* 2009; Randin *et al.* 2009). Here, we pool these two datasets.

In total 1072 different species were found in the study plots. However, only the 333 species with at least 60 presence records were used in the model. Of the 333 species, 131 were only found in forest plots, 40 were only found in grassland plots, and 162 species were found in both datasets.

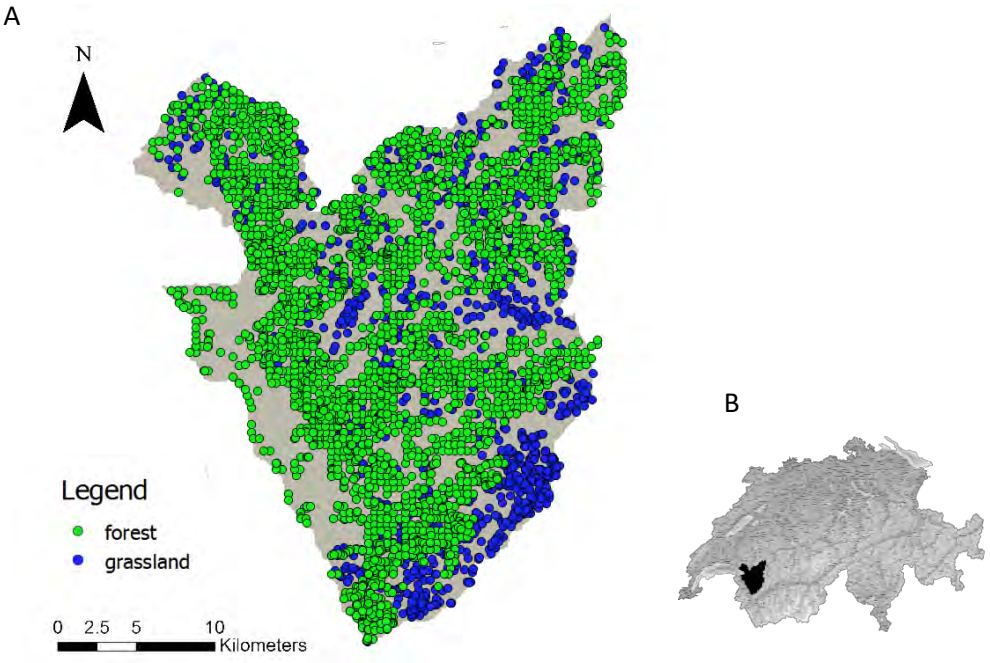


Figure 1 : Map of the study area with sampling plots. It is situated in the Western Alps in the Canton the Vaud. B represents its localization in Switzerland (Image source: *By Tschubby - Own work, CC BY-SA 3.0, https://commons.wikimedia.org/w/index.php?curid=14879468).* In A, green points represent our 3076 forest samples, blue points represent our 913 grassland samples. Created with QGIS 2.18.1.

Climate and topographic predictors

We had climatic and topographic information for all plots, i.e. 3989 plots, provided from a former database (Zimmermann & Kienast, 1999). We used three climatic predictors: degree-days above 3°C (seasonal information, ddeg300), side water balance (available water for plants, swb) and annual global solar radiation (srady). We also used two topographic predictors based on a high resolution digital elevation model

(25x25m, dem25). These were the slope and the topographic position. We therefore had five topo-climatic predictors (TC) providing relevant information about plant habitat. These five topo-climatic predictors have proved their efficiency in previous SDMs in the same study area (Engler *et al.* 2009; Randin *et al.,*2009a; Pellisier *et al.* 2010b, all in Dubuis *et al.* 2013).

To construct our models, we used these five topo-climatic variables and a sixth one representing either minimum, median, maximum canopy height or a random variable (which is used as a control, because we compare models with the same numbers of variables).

Canopy height was calculated based on high resolution LIDAR (light detection and ranging) data provided by the Swiss institute for topography (Swisstopo). The LIDAR remote sensing method was used to collect these data. It is based on the use of a laser to measure heights of vegetation or heights in general. Concretely, we measure the time the wave takes to return to the transmitter. Also, the short wavelengths used in this method allow great precision (Hanner 2010). By measuring canopy height, we also assess a proxy for another ecological information: light available on ground for other organisms (Lefsky *et al.* 2002).

Then, this high resolution (<1m) data was aggregated to 25x25m maps to create minimum (CH_min), median (CH_med) and maximum (CH_max) canopy maps (fig. 2/B/C/D) (Broennimann personal communication, September 2016).

Preliminary analysis by PCA

The first thing we did with this big dataset was visualizing the data. For this purpose, we loaded them in R 3.3.2 (R core team 2016) and made simple correlation plots between all canopy height variables (CH_min, CH_med, CH_max) and our topo-climatic predictors (TC) to see if there was any potential relation. We also performed correlation on forest species only, in order to obtain better results (i.e., preventing a zero inflated distribution of canopy height values). For some variables, we then used a loess function to better fit the data.

Then, we made a new step by computing some PCAs (on R) with topo-climatic variables and canopy height variables or randomraster (our random variable). The idea was to see if adding canopy height as a variable better separate forest species from grassland species. If this is the case, canopy height is a good candidate for modelling.

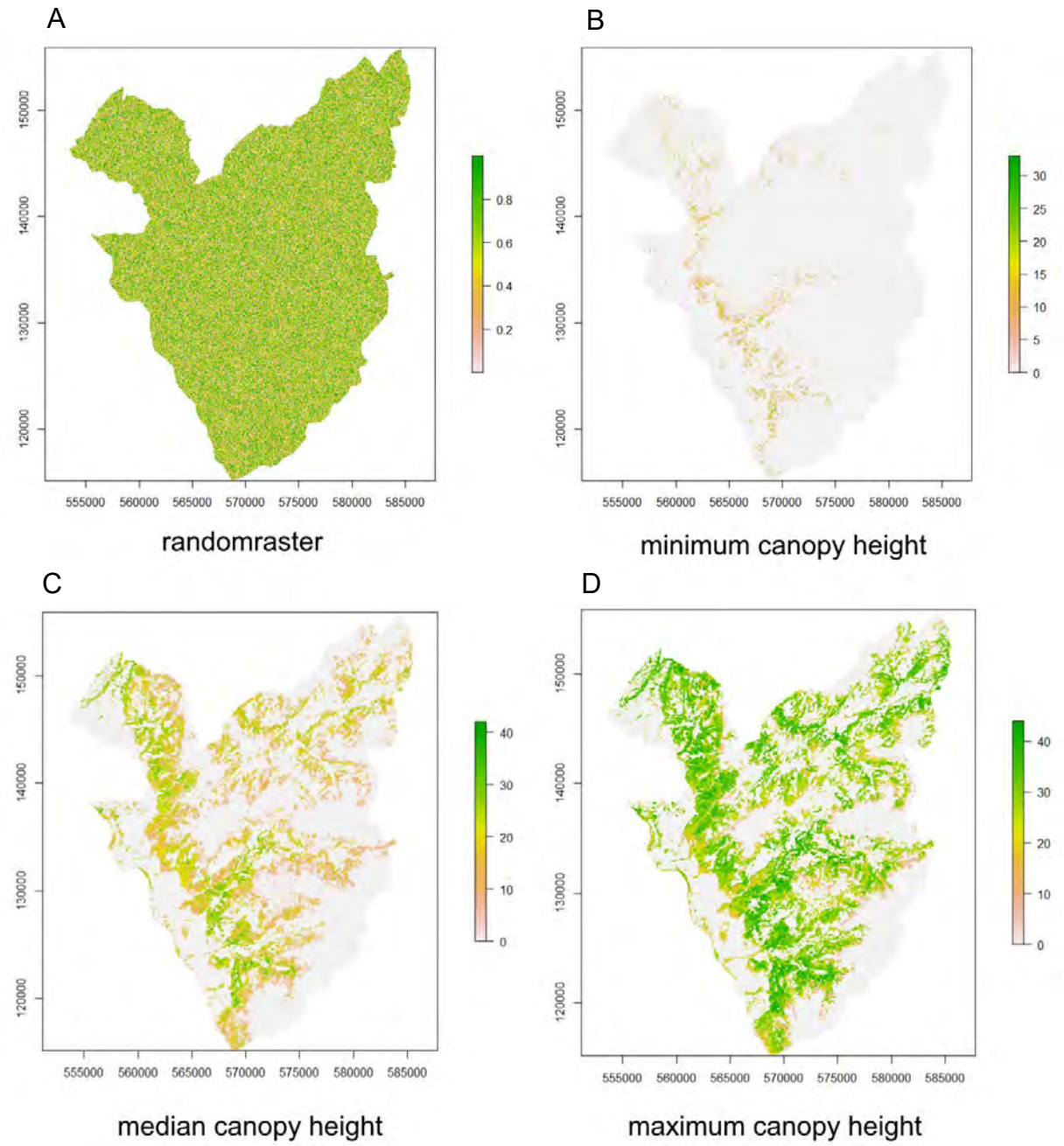


Figure 2 : Maps of randomraster (A) and canopy height layers (minimum (B), median (C), maximum (D)) at 25 m cell size. All maps are georeferenced. Gradation from white to green represent the gradation from lower value to higher value. For the randomraster, the values are uniform distributed random numbers from 0 to 1. For the canopy layers, values are in meters and varies from 0 to 44.

Species distribution modelling

We calibrated four sets of models for each of our 333 species: one using only the topo-climatic predictors with a randomraster (TC + randomraster), and three other using the topo-climatic predictors with one canopy height predictors (TC + CH_min, TC + CH_med, TC + CH_max).

We used four different statistical techniques, as they provide good results in predictions of species distribution (Elith *et al.* 2006), and as their uncertainty is limited by ensemble modelling (Araujo & New 2007). Two of them were based on regression methods: generalized linear models (GLM) and generalized additive models (GAM) (see Guisan *et al.* 2002); and the two others were tree based models: generalized boosted models (GBM, see Elith *et al.* 2008) and random forests (RF, see Prasad *et al.* 2006).

Models were run in R software with "biomod2" library (Thuiller 2016), following this global procedure: loading and formatting the data form biomod2, building "individual model" for each species, building "ensemble models", and making model projections (Georges & Thuiller 2013).

Model evaluation

In order to evaluate the predictive power of our models, we used four evaluation techniques. They are all based on true/false negative/positive (TN, FN, TP, FP) form confusion matrix (Tab. 1). True skill statistic (TSS) corresponds to the $sensitivity + specificity - 1$ (Dubuis *et al.* 2013; Allouche *et al.* 2006). It varies from -1 to 1 with random corresponding to 0, and we considered that a model has good evaluations values from 0.4. AUC, which means area under curve (on a plot with 1 – specificity in x-axis and sensitivity on y-axis). Random is at 0.5 and 1 means the model is perfect (i.e. it perfectly fits the data) (Dubuis *et al.* 2013; Fielding & Bell 1997). Cohens kappa is based directly on TP, FP, etc. (not only specificity and sensitivity) (KAPPA, Cohens 1960) and therefore avoid the tendency of overprediction for some cases with sensitivity and specificity. Like TSS, it varies from -1 to 1, with random at 0. Finally accuracy corresponds to:

$$\frac{TP + TN}{TP + FP + FN + TN}$$

So, it is the proportion of correct predictions (Accuracy, Allouche *et al.* 2006). These are percentages.

The relative importance of each predictor (variable importance) was determined by a repeated random permutation test (see Thuiller *et al.* 2009 for details).

		Observations	
		1	0
Projections	1	True positive (TP)	False positive (FP)
	0	False negative (FN)	True negative (TN)

sensitivity = TP / TP + FN

specificity = TN / TN + FP

Table 1: Confusion matrix with different possibilities in model projections and sensitivity/specificity formula. If we observe a species (1), we can have the right projection (1) and it is a true positive, or a predicted absence (0) and it's a false negative. When a species is absent in observations (0), we can have a predicted presence (1) and it's a false positive, or a predicted absence (0), and it's a true negative.

Single Species model analysis

The single species models (obtained by SDMs) permitted to obtain lots of data, including evaluation statistics (for all evaluation techniques: TSS, KAPPA, AUC and Accuracy), variable importance and projections.

We first analyzed model techniques (GLM, GBM, GAM, RF) by comparing their values for our four evaluation statistics with our four sets of predictors on R software. To do this, we proceed by pairwise comparisons (more precisely pairwise Wilcoxon test, because our data were not independent) between some data selected in dataset. We also visualized all results with boxplots.

Then, we analyzed data in order to determine which set of predictors provide better results (in terms of evaluation statistics values). Again we selected data and did Wilcoxon pairwise tests and boxplots.

Finally, we analyzed importance of the canopy height as a predictor with variable importance data. We looked at mean and standard deviation for all predictor sets.

Community composition predictions

With projections from SDMs, we could transform probabilities in binary response in order to have presences and absences data. This was made using four different thresholding methods (as the method chosen might make a difference to our conclusions, we tested different ones). First, the MaxTSS, which corresponds to choosing the threshold that maximizes TSS values. This recent technique presents the advantage of not being influenced by the variations in prevalence between absences and presences (Allouche *et al.* 2006). Then, MaxKappa, which is the same, but with Cohens Kappa values, and it allows the assessment of improvement over chance prediction (Cohen 1968; Huntley *et al.* 1995). Observed prevalence threshold is a

model-building data-only-based approach (Liu *et al.* 2005). It corresponds to choosing the threshold that makes predicted prevalence equal to observed prevalence in the calibration data. Finally, we used the average probability, which choose the mean of probabilities as threshold.

Then, we stacked our presence/absence maps for all species (S-SDMs) and compared species richness (also compared with species richness based on the sum of probability) and community composition (with the Sørensen index, which "estimates the similarity between the predicted and the observed communities" (Dubuis *et al.* 2013)) with Wilcoxon pairwise t-tests and boxplots.

Results

Preliminary analysis by PCA

Simple linear regressions on our correlation plots didn't inform us much about anything. Then, by fitting loess functions on correlation between maximum canopy height and several variables (degree-days, elevation and temperature), we perceived some interesting relations, closely dependent from each other. For instance, we see that maximum canopy height begins with increasing with higher elevation, then it decreases with even higher altitude (fig. 3).

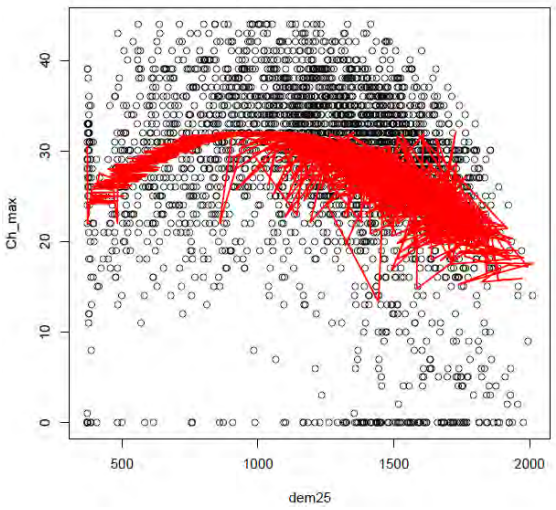


Figure 3 : Loess function on correlation plot between elevation (dem25) and maximum canopy height. We see a higher maximum canopy height with the increase of elevation. Then, in very high elevation, the canopy height decreases.

With PCAs (fig.4), we see that topo-climatic variables with randomraster (fig. 4/A) does not allow to separate forest and grassland species at all. The major axis of the PCA is the temperature (ddeg30), and it makes sense as we have a strong elevation gradient in our dataset and forest species don't grow high in alpine areas.

245

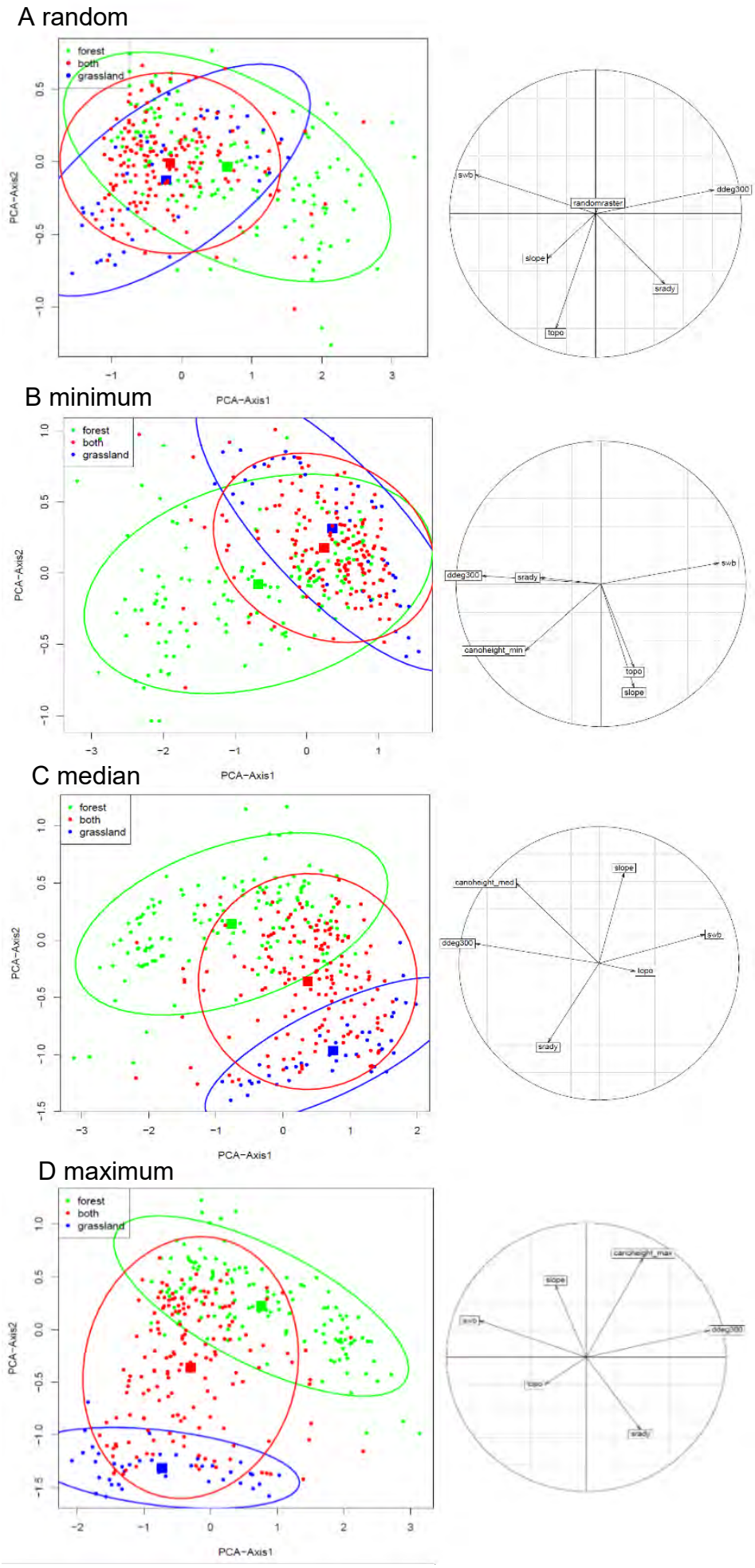


Figure 4: PCA with topo-climatic variables and randomraster (A), minimum (B), median (C) and maximum (D) canopy height on the distribution of species between habitats. Green points are species only found in forest, blue points are species only found in grassland, and red points are species found in both habitat. On the right, we see the axes of each PCA. We see that groups are better separated with median and maximum canopy height (C, D).

With minimum canopy height (fig. 4/B) we obtain very similar results. On the other hand, PCAs with medium and maximum canopy height (fig. 4/C and 4/D) are providing good groups separation. Temperature is still the major axis of the PCAs, but canopy height are now second axis.

Single Species models analysis

By comparing the evaluations statistics of all model techniques (GLM, GAM, GBM, RF) for each set of predictors (topo-climatic (TC) with randomraster, CH_min, CH_med and CH_max), we obtained that almost all model techniques provide different evaluation statistics (see all different value in the example on fig.5). Yet, GLM and GAM models are considered as significantly not different with accuracy evaluation techniques, for all set of predictors (*p-value* are comprised between *0.78* and *0.84*). The most important thing here is the observation of results with RF models. All comparisons allow to conclude to the same thing: RF are always higher than other models (see an example in fig. 5).

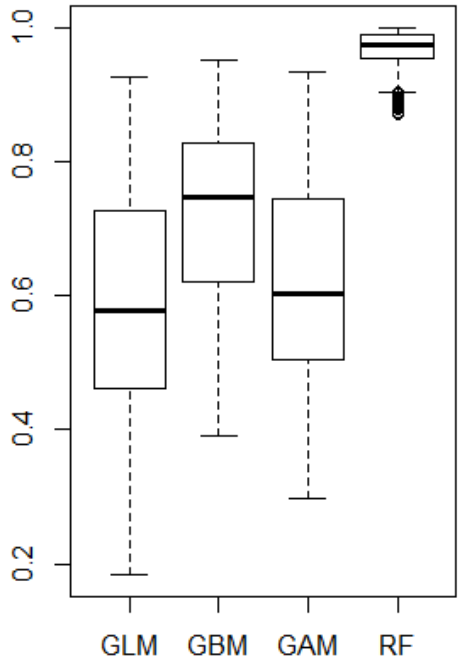
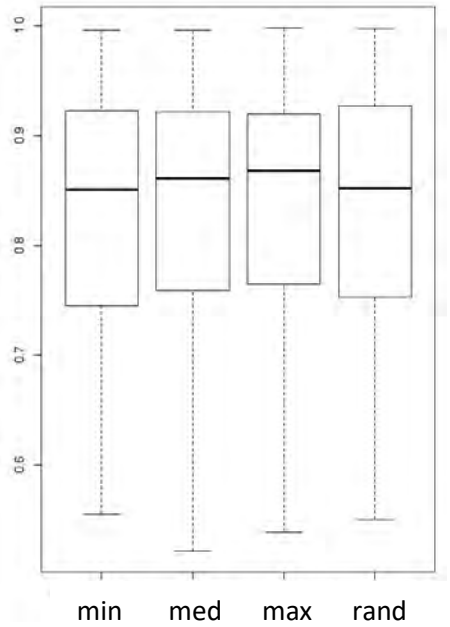


Figure 5 : Boxplot of TSS results with maximum canopy height. We see that TSS results for GLM and GAM are close (but still significantly different according to the results of Wilcoxon pairwise t-test, with *p-value = 0.035*). For all others pairwise comparisons, we have *p-values <2e-16.* We also note the big difference of result with RF.

To determine which is the best set of predictors (TC + CH_min/CH_med/CH_max/random), we compared their evaluation test results. First, we used results with the ensemble model of all model techniques (GLM, GBM, GAM, RF). We therefore have one value (average of all model techniques, one for each set of predictors) for each evaluation test (KAPPA, TSS, AUC, Accuracy) (see the example on fig.6).




```
        Pairwise comparisons using Wilcoxon rank sum test

data:  c(All_TSS[1, ], All_TSS[2, ], All_TSS[3, ], All_TSS[4, ]) and
c(rep("canopymin", 333), rep("canopymed", 333), rep("canopymax",  c(A
ll_TSS[1, ], All_TSS[2, ], All_TSS[3, ], All_TSS[4, ]) and     333),
rep("random", 333))

          canopymax canopymed canopymin
canopymed 1         -         -
canopymin 1         1         -
random    1         1         1

P value adjustment method: holm
```

Figure 6: Boxplot of TSS results for each set of predictors, with the ensemble model of all model techniques (on the left), and its Wilcoxon pairwise t-test (made on R) that compare all values (on the right). We see that TSS results are not significantly different (all *p-value = 1*), for all sets.

With this, we obtained a global impression that indicates if canopy height is a good predictor or not. We obtained that there is no difference between the different sets of predictors (*p-value* in Wilcoxon pairwise t-test are always bigger than *0.17,* and almost everytime equal to *1*).

Then, we made the same tests with all models techniques separately. For GLM, GBM and GAM models, the median and maximum canopy height were significantly different to minimum canopy height and random for all KAPPA, TSS and AUC tests. Maximum and medium canopy height are not significantly different from each other, as minimum canopy height and random. With boxplots (fig.8), we know that maximum and median canopy height provides significantly higher results than median canopy height and random. In figure 7, we have an example of the comparison results between sets of predictors with GLM model and KAPPA evaluation test.

```
        Pairwise comparisons using t tests with pooled SD

data:  c(GLM_KAPPA[1, ], GLM_KAPPA[2, ], GLM_KAPPA[3, ], GLM_KAP
PA[4,  and c(rep("canopymin", 333), rep("canopymed", 333), rep("
canopymax",      ]) and     333), rep("random", 333))

          canopymax canopymed canopymin
canopymed 0.067     -         -
canopymin 1.4e-11   2.9e-06   -
random    3.1e-13   2.0e-07   0.575

P value adjustment method: holm
```

Figure 7: Results for Wilcoxon pairwise t-test for GLM model and KAPPA evaluation test on R. We see that maximum canopy height is not significantly different from median canopy height (*p-value = 0.067)* and minimum canopy height is not significantly different from random (*p-value = 0.575*). All other pairs are significantly different (with *p-value ≤ 2.9e-06*). Maximum and median canopy height provides significantly different results than minimum canopy height and random.

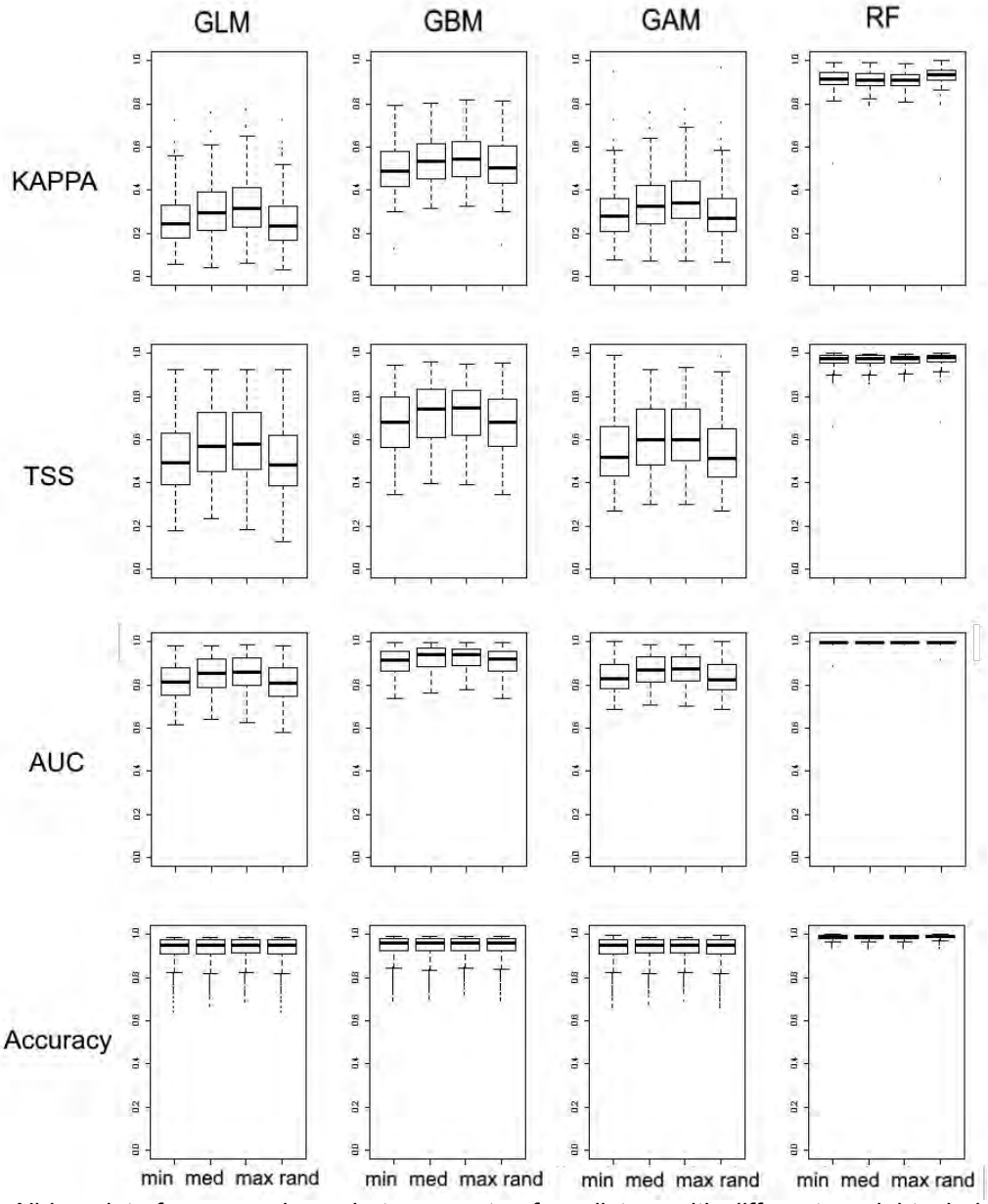


Figure 8: All boxplots for comparisons between sets of predictors with different model techniques and evaluation tests. We see that RF and Accuracy always give very high results. We also see global proximity between minimum canopy height and random, and between medium and maximum canopy height (especially with GLM, GBM, GAM and KAPPA, TSS and AUC). In these case, results for maximum and median canopy height are significantly higher than minimum canopy height and random,

For GLM, GMB, GAM models with Accuracy evaluation test, we always obtain no difference between predictor sets (all *p-value = 1*).

For RF models, we have slightly different and less clear results. We have, for TSS and AUC: all canopy heights significantly different from randomraster. For KAPPA, all canopy heights are different from random, but they are not all equivalent amongst themselves. For Accuracy, the only significant difference is between maximum canopy height and randomraster. So, we found no regular pattern with RF models.

Globally (fig. 8), when we have significate differences (i.e. in the majority of our results), it always includes maximum canopy height (at least) versus randomraster. Boxplots allow us to see that maximum canopy height is higher than randomraster. Besides, maximum canopy height and medium canopy height are almost every time not significantly different from each other.

The analysis of variable importance indicates that degree-days is always (for all sets of predictors and all model techniques) the most important variable (with values ranging from *0.40 to 0.63*). But for sets of predictors with maximum and median canopy height, the second most important variable is canopy height (with values from *0.27* to *0.41*).

Community composition analysis

The comparison of species richness deviation with different thresholding methods doesn't give any decisive result. Either we obtain that all set of predictors gives the same richness deviation, or we have two sets of predictors that aren't significantly different, and all the others are. For instance, with maxTSS, we obtain that all values are significantly different, except for sets with maximum canopy height and minimum canopy height (*p-value = 0.79*). For the average probability, again all values are significantly different, except for sets minimum canopy height and random (*p-value = 0.41*).

For the composition comparisons, we obtain even more scattered results. MaxTSS and Observation prevalence both give that no similarity values are significantly different (*p-values from 0.14 to 1*). MaxKAPPA indicates that none are significantly different, except for minimum canopy height and maximum canopy height (*p-value = 0.024)*, and Average probability indicates that all values are significantly different, except for minimum canopy height, which differ from random (*p-value = 0.58).*

Discussion

First, our preliminary analysis allow us to see that canopy height data seems logical and therefore correct. Indeed, canopy height increasing, then decreasing with elevation makes sense. But more importantly, PCAs allow us to conclude that canopy height seems to be a good candidate for modelling, as it better separates species groups. We also see that maximum and median canopy height both give good results. Indeed, degree-days stays the major axes in all PCA, but maximum and median canopy height are the second most important axes in each graphs, and allows good groups separation.

Then, with this idea, we use canopy height variable in models, with the first question being: does it improve single species models? We first saw that RF models largely differ in results than other model techniques (providing always very high evaluation values). The hypothesis is that RF models largely overfit the data. Then, knowing that, we can better analyze the results we obtained with sets of predictors' comparisons. We can first expect that results provided by RF models will not be much contrasted. And it is the case, as minimum, median and maximum canopy height are often considered as not significantly different. Still, we can already see better results for all canopy height predictors, with regard to the one with random values. Then, with other model techniques (GLM, GBM, GAM), we see that results obtained with Accuracy differ from the others (nothing differ significantly). Nevertheless, accuracy is known to be very controversial, due to the "accuracy paradox" (Abma 2009). Therefore, much important and reliable results are the ones obtained with GLM, GAM and GBM as model techniques, and with KAPPA, TSS and AUC as evaluation test. All these lead us to the same results: predictors sets including CH_max and CH_med better predict where a species will be (in space) than predictors sets with CH_min or random. Finally, analyzing variable importances allows us to conclude that degree-days is the most important variable in each case, and that canopy height is the second most important variable with CH_max and CH-med (for CH_min and random, second most important predictor are available water (swb) and solar radiation (srady)). From this whole part, we conclude that maximum (and median) canopy height is a good predictor in general. It globally improves the models and is an important variable. Note that these results are congruent with the ones from PCAs.

Then, the last step was stacking the models and analyzing the community predictions. Indeed, as canopy height allows good predictions for single species, we can hope for good results for community. Unfortunately, we can't conclude any

improvement about that by including canopy height in the models. As the results were really uneven, we can't draw any conclusions, but global impression would even be that including canopy height does not change anything.

We can conclude that canopy height gave promising results with Single Species model (and PCAs), and that it improves models for predict where a single species would be. But at this point, we cannot conclude that it can be used to predict community, i.e. separate grassland and forest vegetation.

Also, other limitations exist with the canopy height variable. For example, canopy height does not inform us about age of forest, and this could potentially lead to wrong prediction of species.

In a study about SDMs, they made the hypothesis that "the most accurate S-SDMs predictions are predominantly associated with strong environmental filtering in climatic harsh" (Pottier *et al.* 2013). They found that this was confirmed using canopy height and considering the assemblage specificity. With sensitivity, this was not true anymore. Besides, in another study, we learn that tree canopy cover is better than canopy height to estimate forest area (Mathys 2007). If we look at our results in this context, it could make sense. Maybe, if we add another variable in our model: canopy cover, in addition to canopy height, we maybe could obtain better result in specificity, and sensitivity (as canopy cover allows theoretically to separate forest from other area). Then, community predictions could maybe be improved.

Acknowledgments

I want to thank Daniel Scherrer, who supervised this work by providing very useful advice, help and quality teaching. I also thank all providers of data, in particular Olivier Broennmann, who adapted canopy height data, and Antoine Guisan, chief of the Spatial Ecology Group in Unil. I also thank all reviewers, and again particularly Daniel Scherrer for his wise comments and corrections.

References

Abma BJM (2009) Evaluation of requirements management tools with support for traceability-based change impact analysis. Master's thesis, University of Twente, Enschede.

Allouche O, Tsoar A, Kadmon R (2006) Assessing the accuracy of species distribution models: prevalence, kappa and the true skill statistic (TSS). *Journal of applied ecology*, **43(6)**, 1223-1232.

Araújo MB, New M (2007) Ensemble forecasting of species distributions. *Trends in ecology & evolution*, **22(1),** 42-47.

Broennimann O, Randin C, Zimmermann NE, Guisan A (2003) Swiss Eco-Climatic GIS data. [URL: https://www.unil.ch/ecospat/files/live/sites/ecospat/files/shared/CHpreds2.pdf, consulted on October 2016].

Cohen J (1960) A Coefficient of Agreement for Nominal Scales. *Educational and Psychological Measurement*, **20**, 37-46.

Cohen J (1968) Weighted kappa: nominal scale agreement with provision for scaled disagreement or partial credit. *American Psycological Association*, **70**, 213-220.

D'Amen M, Dubuis A, Fernandes RF, Pottier J, Pellissier L, Guisan A (2015). Using species richness and functional traits predictions to constrain assemblage predictions from stacked species distribution models. *Journal of Biogeography,* **42(7),** 1255-1266.

Descombes P, Petitpierre B, Morard E, Berthoud M, Guisan A, Vittoz P (2016) Monitoring and distribution modelling of invasive species along riverine habitats at very high resolution. *Biological Invasions*, **18(12)**, 3665-3679.

Elith J, Graham CH, Anderson RP, Dudik M, Ferrier S, Guisan A, Hijmans RJ, Huettmann F, Leathwick JR, Lehmann A, Li J, Lohmann LG, Loiselle BA, Manion G, Moritz C, Nakamura M, Nakazawa Y, Overton JM, Peterson AT, Phillips SJ, Richardson K, Scachetti-Pereira R, Schapire RE, Soberon J, Williams S, Wisz MS, Zimmermann NE (2006) Novelmethods improve prediction of species' distributions from occurrence data. *Ecography* **29.** 129–151.

Elith J, Leathwick JR, Hastie T (2008) A working guide to boosted regression trees. *Journal of Animal Ecology*, **77**, 802-813.

Fielding AH, Bell JF (1997) A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environmental Conservation*, **24**, 38-49.

411 Guisan A, Edwards TC, Hastie T (2002) Generalized linear and generalized additive
412 models in studies of species distributions: setting the scene. *Ecological*
413 *Modelling*, **157**, 89-100.
414 Guisan A, Theurillat JP (2000) Equilibrium modeling of alpine plant distribution: how
415 far can we go?. *Phytocoenologia*, **30(3/4),** 353-384.
416 Guisan A, Thuiller W (2005) Predicting species distribution: offering more than simple
417 habitat models. *Ecology letters*, **8(9)**, 993-1009.
418 Georges D, Thuiller W (2013) Multi-species distribution modeling with biomod2.
419 Hanner J (2010) Forest canopy height: why do we care?. *Culturing Science - biology*
420 *as relevant to us earthly beings*. [URL:
421 https://culturingscience.wordpress.com/2010/07/28/canopy-height/, consulted
422 on October 2016]
423 Hirzel A, Guisan A (2002). Which is the optimal sampling strategy for habitat
424 suitability modelling. *Ecological modelling*, **157(2),** 331-341.
425 Huntley B, Berry PM, Cramer W, McDonald AP (1995) Modelling present and
426 potential future ranges of some European higher plants using climate response
427 surfaces. *Journal of Biogeography*, **22**, 967-1001.
428 Lefsky MA, Cohen WB, Harding DJ, Parker GG, Acker SA, Gower ST (2002). Lidar
429 remote sensing of above-ground biomass in three biomes. *Global ecology and*
430 *biogeography*, **11(5),** 393-399.
431 Liu CR, Berry PM, Dawson TP, Pearson RG (2005) Selecting thresholds of occurrence
432 in the prediction of species distributions. *Ecography*, **28**, 385-393.
433 Maiorano L, Falcucci A, Zimmermann NE, Psomas A, Pottier J, Baisero D, Rondinini
434 C, Guisan A, Boitani L (2011) The future of terrestrial mammals in the
435 Mediterranean basin under climate change. *Philosophical Transactions of the*
436 *Royal Society of London B: Biological Sciences*, **366(1578),** 2681-2692.
437 Mathys L (2007) A discrete forest in a continuous landscape: investigating interactions
438 between natural resource management and sustainable development
439 Mod HK, Scherrer D, Luoto M, Guisan A (2016) What we use is not what we know:
440 environmental predictors in plant distribution models. *Journal of Vegetation*
441 *Science*, **27(6)**, 1308-1322.
442 Pottier J, Dubuis A, Pellissier L, Maiorano L, Rossier L, Randin CF, Vittoz P, Guisan,
443 A (2013) The accuracy of plant assemblage prediction from species distribution
444 models varies along environmental gradients. *Global Ecology and Biogeography*,
445 **22(1s)**, 52-63.

446 Pradervand JN, Dubuis A, Pellissier L, Guisan A, Randin C (2014) Very high resolution
447 environmental predictors in species distribution models Moving beyond
448 topography?. *Progress in Physical Geography*, **38(1)**, 79-96.
449 Prasad AM, Iverson LR, Liaw A (2006) Newer classification and regression tree
450 techniques: Bagging and random forests for ecological prediction. *Ecosystems*,
451 **9**, 181-199.
452 QGIS Development Team (2016) QGIS Geographic Information System. Open Source
453 Geospatial Foundation Project. http://www.qgis.org/
454 R Core Team (2016) R: A Language and Environment for Statistical Computing.
455 Vienna, Austria, R Foundation for Statistical Computing.
456 Randin CF, Dirnböck T, Dullinger S, Zimmermann NE, Zappa M, Guisan A (2006) Are
457 niche-based species distribution models transferable in space? *Journal of*
458 *Biogeography*, **33(10)**, 1689-1703.
459 Randin, C.F., Engler, R., Normand, S., Zappa, M., Zimmermann, N.E., Pearman, P.B.,
460 Vittoz, P., Thuiller, W. & Guisan, A. (2009) Climate change and plant distribution:
461 local models predict high-elevation persistence. *Global Change Biology*, **15**,
462 1557-1569.
463 Randin CF, Jaccard H, Vittoz P, Yoccoz NG, Guisan A (2009) Land use improves
464 spatial predictions of mountain plant abundance but not presence-absence.
465 *Journal of Vegetation Science*, **20(6)**, 996-1008.
466 Scherrer D, Massy S, Meier S, Vittoz P, Guisan A (in review). Assessing and predicting
467 shifts in mountain forest composition across 25 years of climate change. *Diversity*
468 *and Distribution.*
469 Thuiller W, Georges D, Engler R, Breiner F (2016). Package “biomod2”- Ensemble
470 Platform for Species Distibution Modeling version 3.3-7.
471 Thuiller W, Lafourcade B, Engler R, Araújo MP (2009). BIOMOD - A platform for
472 ensemble forecasting of species distributions. *Ecography.* **32**. 369-373.
473 Zimmermann NE, Kienast F (1999) Predictive mapping of alpine grasslands in
474 Switzerland: Species versus community approach. *Journal of Vegetation Science*
475 **10(4).** 469-482.
476