Electronic Supplementary Material S2

Evaluation of tree species models

Content

Response behaviour

In the field of ecology, the validation of the results by existing ecological knowledge is important, especially when using especially black-box methods used for estimates and predictions. Whenever the predictions and estimates of the model coincide with the ecological characteristics of the species, it confirms the judgment of the rationality of the model, and if it contradicts these, it reinforces the judgment that the results of the model are artefacts. In our study, the functional relationship of the abiotic environmental predictor variables on the probabilities of current and future distribution were analysed using partial dependence plots to detect in which direction the change of each predictor affects the value of the estimated probability of occurrence and how strong this effect is. In case the direction and relative magnitude of the effect coincide with our knowledge about the given species, the model is likely to give a realistic mapping of real relationships.

In the case of the continent-wide models, for some species the functional relationship between the predictors and occurrence probability was strongly predestined by the map pattern of the training data. When the occurrence-absence points used for the training areas show blocked structures, the dominance of the predictor variable corresponding to this geographical division in the estimates clearly interacts. This is typical for the continent-wide models' seasonal maximum temperatures, or the annual heat-moisture index and summer heat-moisture indices, or the day-degree values. For north- and south-framed occurrence patterns, there are usually two limiting factors that the models use: temperature is the limiting factor from the north, and the colder it becomes, the less likely the occurrence probability. From the south, humidity is limiting, so the greater the tendency to drought, the lower the occurrence probability is. Since in most cases the occurrence blocks are sharply and directly delimited on the map, the models also give sharp and well-defined boundaries in predictor values. A comparison of continent-wide with the regional models shows that the more diffuse the geographical distribution of occurrence-absence pairs, the smaller the value of explained variance.

Alnus glutinosa

The individual evaluation of the predictors for *A. glutinosa* in the continent-wide model showed that the most influential variable is the autumn maximum temperature. This is followed by the SHM index and the annual precipitation. The maximum autumn temperature rising approx. from 5 to 11°C has a very significant positive effect on the likelihood of the species occurring. The SHM value has the opposite negative effect on the probability of occurrence approx. from 0 to 180. The annual rainfall between 100 to 1000 mm also has a positive effect on the probability of occurrence. Outside these ranges, the predictors have no interpretable effect on the value of the target variable. A trend-like relationship between the increase in the value of continentality and the probability of occurrence could not be detected.

In the regional model, predictors of *A. glutinosa* water demand outweigh the impact of large-scale climatic variables. Thus, the vertical distance measured from the surface of the nearest water bodies and the AHM index indicating drought sensitivity are the most important variables, followed by two important degree days influencing the length of the growing season (DDbelow0 and DDabove5). The probability of occurrence is radically reduced by the increase in the vertical distance from a water body. Every meter counts up to about 100 m. It does not have a greater effect on the occurrence beyond 100 meters. Degree-days below 0°C not greater than 200 has a positive effect on the probability of occurrence. The need for more balanced climatic conditions is indicated by the occurrence optimum between 2300 and 2800 for degree-days above 5°C or when the AHM index is not outside the 10-30 range.

Fraxinus angustifolia

The species *F. angustifolia* requires a warmer, more balanced climate (continental model) and the connection to water (regional model). The temperature limit of the occurrence is close to the autumn maximum temperature of 12 °C. Tmax_at up to 16 °C has a strong effect of increasing the probability of occurrence. As a typical limiting variable has not been revealed by continental analyses, it is likely that our model according to temperature rise will project almost unlimited expansion for this tree species on a continental scale.

At the same time, in the regional models, besides the variables with the same effect as in the case of *A. glutinosa*, the horizontal measured distance to the nearest water body has already become a major predictor variable.

Fraxinus excelsior

The distribution pattern of training samples for *F. excelsior* showed a limited distribution similar to the alder, constrained from the north and south. We found several limiting climatic parameters here, which gave a chance to gain more realistic models. Besides the highly dominant expansion effect of the autumn maximum temperature, the SHM index and the degree of continentality proved to be significant limiting factors. This phenomenon fits into our picture of the species, which indicates a preference for more humid, rather cooler climate types. Thus, in continental models, with a southern decline, a northern spread is expected.

At the regional level, vertical and horizontal distances from rivers, the AHM index, and the number of degree-days below 18°C were the most significant predictors. According to the models, the higher locations are more ideal for ash than was found to be ideal for alder, fitting the role of the tree species in floodplain hardwood groves. Among the climate variables, at the regional level, the number of degree-days below 18°C has a positive effect on the probability of spread between values of 2500 and 3200, while the AHM index proved to be the most restrictive predictor. Overall, AHM is the most influential predictor for this tree species.

Populus nigra

The tree species *P. nigra* in the continent-wide model, the summer maximum temperature had the largest effect on the occurrence probability. In an ascending range from 20 to 26°C the summer temperature maximum stimulates a northward spreading. This experience roots in part the lowland and riverside distribution of *P. nigra*, and in part the bipolar distribution of training samples. The second most important variable was the number of degree-days below 0°C, which progressively decreased the probability from 0 to 450. In addition to the above, summer precipitation slightly increases the probability of occurrence up to 400 mm.

In the regional model, the pattern of training samples predicted that the horizontal and vertical distances of rivers strongly predestine the occurrence. In addition to the variables related to water bodies, three more significant variables were identified in the regional occurrence model of *P. nigra*: the Annual heat:moisture index, the values of degree-days above 0°C and 18°C. Of the degree-days, the former has a more negative effect and the latter has a more positive effect on the probability of occurrence. Interestingly, the increase in the Annual Heat-Moisture Index also has a positive effect on the output of the model, which indicates that drought is not an influencing factor accordingly.

Because the model did not really find a variable that would negatively affect the distribution of the tree species as climate change progressed, the projections mainly indicate area gains for *P. nigra*.

Quercus robur

For the *Q. robur*, the training samples of the continental models show a characteristic Southern- and Northern European absence and a strong presence along the east-west axis of the continent. Accordingly, the most important drivers of the continental models were the autumn maximum temperature, the summer precipitation amount, the degree of continentality, and the annual precipitation amount. As a decrease in these precipitation amounts has a negative effect on distribution, continental models predict a northward shift. No driver for the lowland nature of the tree species has been shown, so these predictors will not limit the potential for new dispersal in mountainous areas. Therefore, for future studies, it would be worthwhile to include a parameter limiting the vertical spread among the predictors, for example in the form of the current vertical distribution limit or taking into account soil types.

In the regional models, in addition to the distances from water bodies, the number of degree-days above 5°C and the extreme minimum temperature were appearing to be the strongest climatic drivers. The role of the AHM index can be said to be significant based on the importance values. Regarding the dependence plot, the interpretation of the effects is more difficult, as only the horizontal and vertical distance from the water bodies showed a clear trend in the effects of the variables: their increase counteracts the occurrence probability. The variable degree-days above 5°C has a positive effect on occurrence between values 2500 and 3000, while the AHM index has a more positive effect on the probability of occurrence between 10 and 25, and a more negative effect between 25 and 40. It is probably due to these effects that on the maps created by the model, the decline in probabilities around the rivers is more significant, and while in more distant areas where distance no longer takes significant effects the increase in degree-days and the increase in the AHM index values become dominant. However, due to the unlikely prediction of the regional model, estimates of *Q. robur* should be treated with caution.

Ulmus laevis

The *U. laevis*, showed a massive occurrence block in Central- and Eastern Europe and an absence block, which surrounds that from the North, West, and South, which can be indicated from the training data consistency. The order of importance of the predictor variables was also in line with this pattern. The most important climatic driver became the average temperature of the coldest month, which has the most positive effect on the probability of appearance in the range of -6 to -1°C, while it has a strong negative effect on the occurrence in the range of -1 to + 2°C. The AHM index has a strong positive effect on the probability of occurrence in the range between 20 and 25. However, higher AHM index values have a small negative effect on occurrence probability. An increase in winter precipitation is unfavourable, while an increase in the maximum temperature in spring (between 10 and 18°C), and an increase in the amount of summer precipitation (between 100 and 250 mm) has a beneficial effect on the value of the probabilities of spread. These drivers predict a very definite shift due to climate change towards Sweden, Finland, and Russia, while in the Central European region there is an opportunity to gain additional area only in the Carpathian region.

In the regional models, the temperature dependence between the most important predictor variables is represented by the degree-days below 0°C, and above 18°C, as well as the frost-free period. Days below 0°C decrease, while days above 18°C degrees increase the likelihood of occurrence. The wellknown dependence on water for *U. laevis* elm is best represented by the distance from rivers in the model, with which the incidence decreases drastically.

Ulmus minor

The main drivers for *U. minor* are the summer maximum temperature rise, the annual-, and summer rainfall sums. Dependency plots show a strong limitation below 20°C of maximum summer temperature and an increase up to 24°C with maximum inclination at 22°C. The positive influence of the frost-free period starts at 180 days with peak at around 320 days. Continentality values increase the probability of occurrence from around 13 up to 20 and decrease again at approximately 25. The model resulted in realistic continent-wide maps.

In the regional model, an increase in the number of frosty days is a clear limiting effect, complemented by a decrease in annual precipitation. Degree-days below 0°C decrease the occurrence probability until about 200, increasing degree-days above 5°C have a positive influence, and annual total precipitation abruptly increases the probability at around 800mm until 2000mm.