Project II: Tree species prediction

Names: Sarah Brüning, Tobias Koch, Marvin Müsgen, Eike Schott, Frauke von den Driesch
Umweltinformatik | Philipps-Universität Marburg

Outset and Goal

During the previous work step a comprehensive dataset consisting of spectral and structural properties was created. Together with the generated tree crown segmentation it provides the basis for predicting tree species only from remotely sensed observations. The goal of this project is to create a tree species map for Marburg Open Forest.

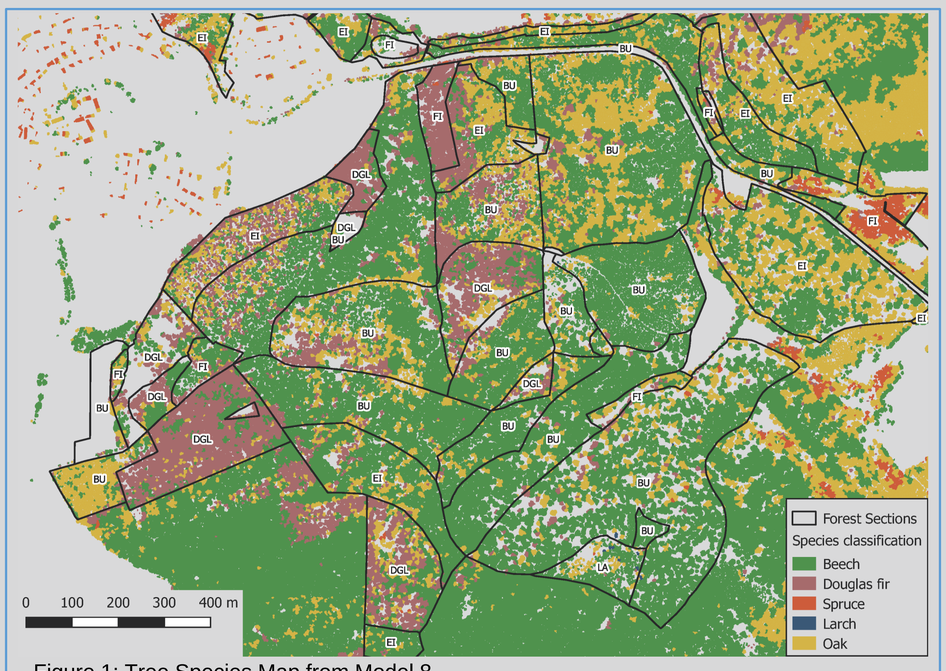

Figure 1: Tree Species Map from Model 8

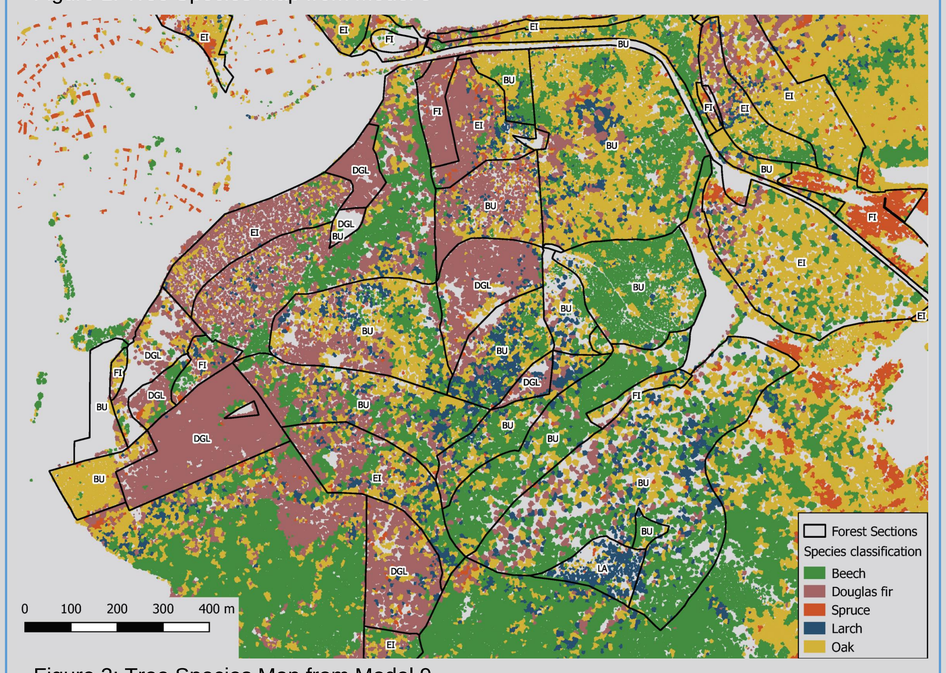

Figure 2: Tree Species Map from Model 9

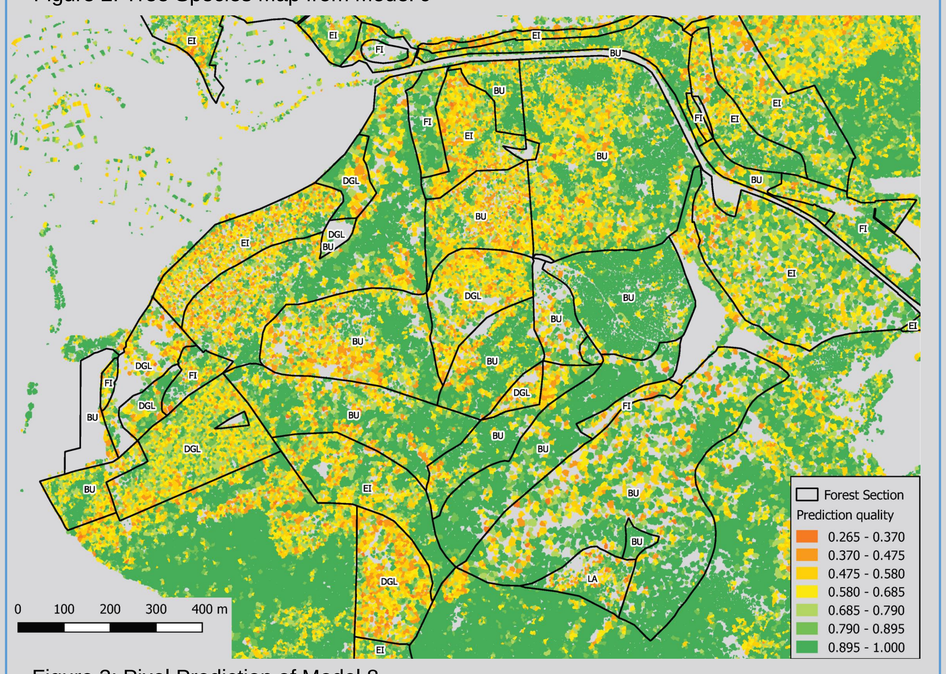

Figure 3: Pixel Prediction of Model 8

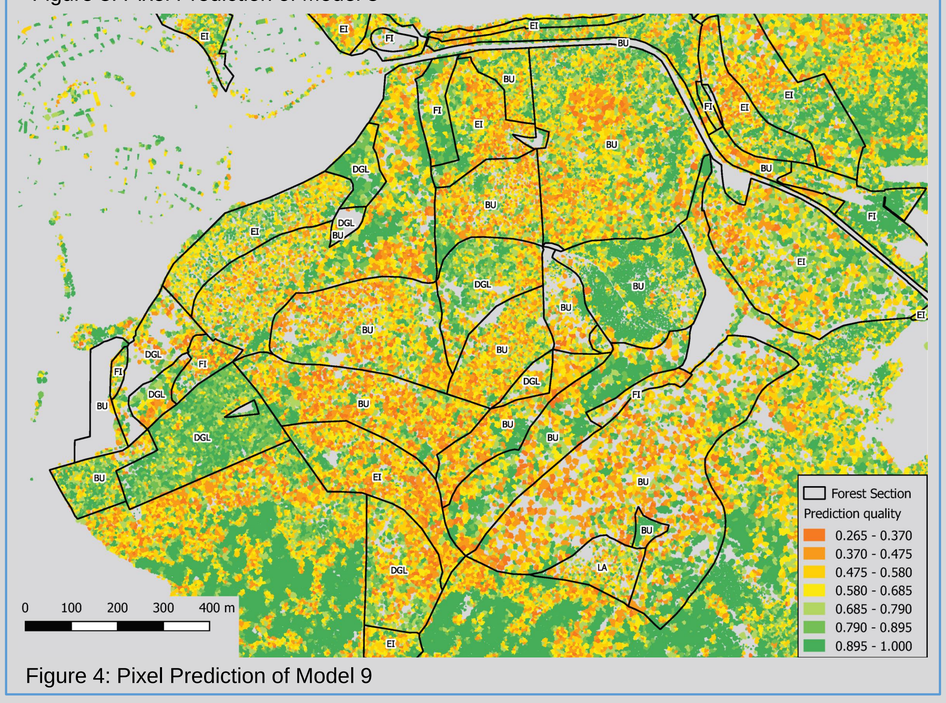

Figure 4: Pixel Prediction of Model 9

Model Creation

The final tree classification models are based on the results of the first project. The data basis is built up by the selected RGB indices and filters. Additionally the first and second principal components[3] calculated from the indices in combination with the selected filters are fed into the model training. The Models 8 and 9 (Fig. 1 and 2) are trained with the same training areas which were selected on the basis of the provided aerial image. The tree crown segmentation from the first project is not used for area selection due to the difference between the canopy height model and the aerial image. The designated training areas of Model 8 are weighted by tree species ratios (beech, douglas fir, spruce, larch, oak). For Model 9 equally large areas are dedicated for the different tree species so 1000 pixels per tree species were passed to the training. Then, a random sized sample of polygons is drawn for both models. Overall, seven forward feature selections were performed under usage of the "Caret" package[1]. The RGB indices, the colour channels red, green and blue as well as the first and second principal components[3] and these in combination with filters were passed to the forward feature selection. Results show the selection of only the first and second principal components and the principal com-ponents in combination with filters. Due to time related reasons the first and second principal components as well as their combination with filters are supplied to models 8 and 9 without running new forward feature selections. Finally, it is predicted on the remaining 80 % of pixels.

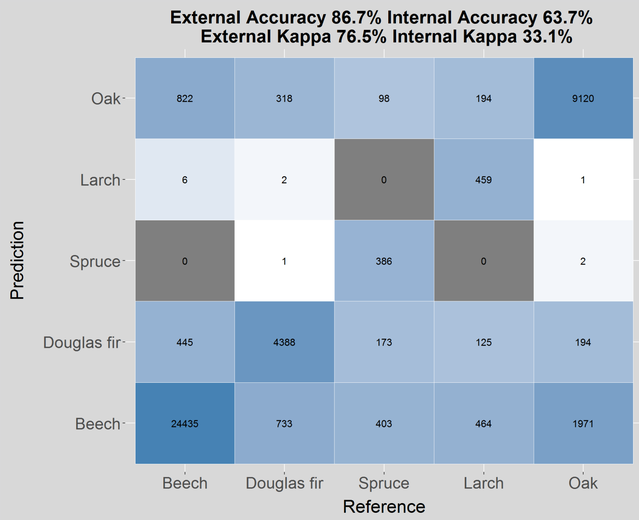

Figure 5: Pixel Prediction of Model 8

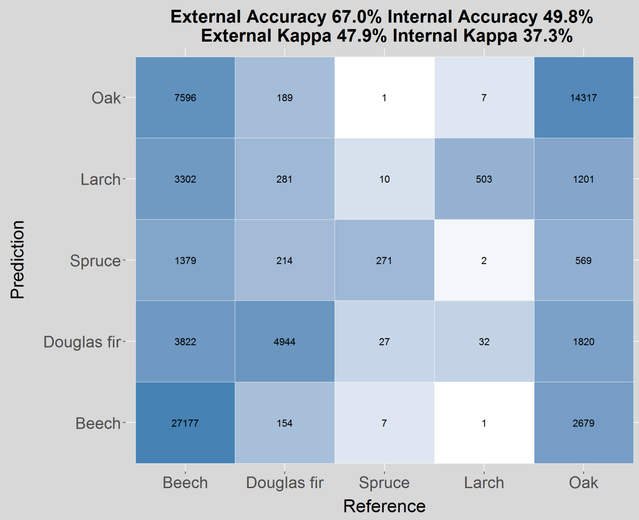

Figure 6: Pixel Prediction of Model 9

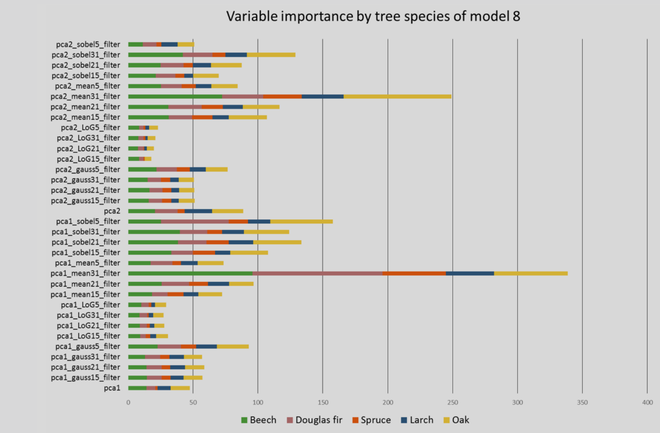

Figure 7: Variable importance by tree species of model 8

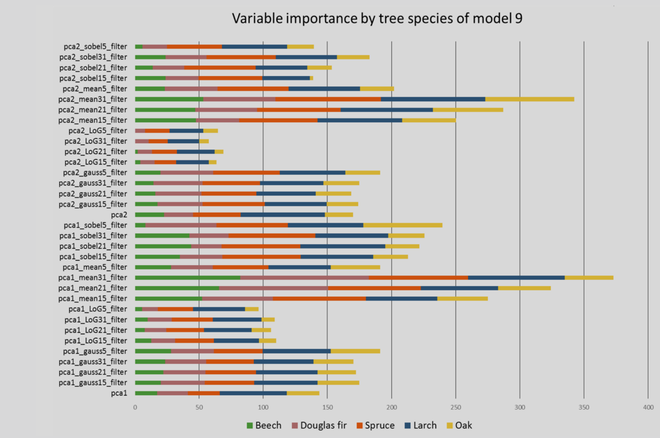

Figure 8: Variable importance by tree species of model 9

Discussion

The model performances are evaluated using Cohen's Kappa[2] which describes the mathematical extent of randomness and ranges between zero and one with high values indicating low extent of randomness. The statistical results of models 8 and 9 are listed in Figure 5 and 6. A polygon is classified by the species with the maximum amount of pixels within the polygon. The accuracy of classification of single polygons is shown in Figure 3 and 4. For that the number of pixels of the classified species is divided by the total number of pixels in the polygon. This measure describes the quality of the separately carried out classifications per tree. It is apparent that model 9 is worse than model 8 in predicting beeches which take up the majority in Marburg Open Forest. In contrast, the prediction quality for douglas firs in model 9 is better than in model 8. Larches, oaks and spruces are classified better as well in model 9.

Altogether it becomes apparent that both models have their own advantages. Model 8 performs better for the area of the beech dominated Marburg Open Forest than model 9. This can be traced to the species weighted training areas with a majority of beeches in model 8. It is up to discussion if model 9 will perform better in a forest with an unknown forest structure and species ratios due to the species balanced training areas for model 9.

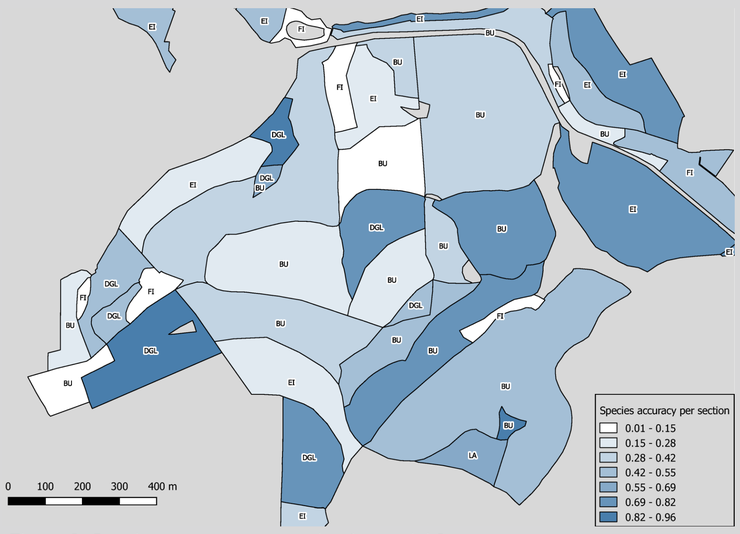

Figure 9: Species accuracy per section for model 9

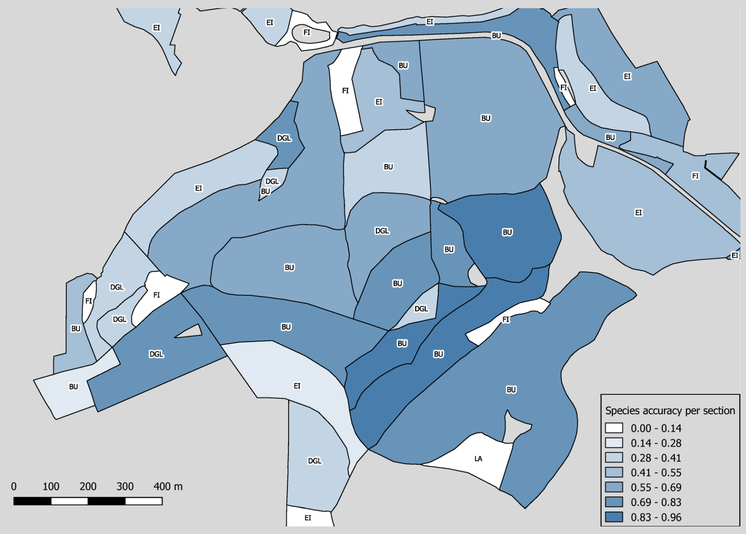

Figure 10: Species accuracy per section for model 8

Literature

1. J. Cohen (1960): A coefficient of agreement for nominal scales. In: Educational and Psychological Measurement, 20.
2. M. Kuhn (2011): The caret Package. <https://cran.r-project.org/web/packages/caret/index.html> last acces 13.02.2019 .
3. Jolliffe, I. (2002): Principal Components Analysis. Berlin/Heidelberg: Springer.