

Creating a comprehensive dataset to identify vegetation structures and predict species in a heterogenous Central European temperate forest

INTRODUCTION |

Monitoring the changes in land use and land cover is one of the greatest challenges facing nature conservancy today. The primary mission of the state-sponsored project “Natur 4.0” is to develop a new and state-of-the-art modularized environmental monitoring program. In our project, airborne data are used to calculate vegetation indices, which will form the basis for tree species prediction by machine learning algorithms. We calculate several spatial and Sobel filters to provide the algorithm with sufficient training data. To avoid unnecessary information, however, we compute a principal component analysis (PCA), so that the filters can be used properly.

STUDY AREA |

The area is a typical Central European temperate forest, which belongs to Philipps University of Marburg. Located west of Marburg, the forest is part of the Gladenbacher Hügelland in the Hessian Mittelgebirge. The mixed forest has clearly structured departments (see Figure 1) containing coniferous and deciduous forest and is dominated by *Fagus sylvatica* (European beech).

Table 1: List of vegetation indices calculated using RGB aerial images

Plot labeling	Index shortcut	Index name	Index Formula (R=Red, G=Green, B=Blue)
Vegind.1	VVI	Visible Vegetation Index	$(1-(R-30)/(R+30))*(1-(G-50)/(G+50))*(1-(B-1)/(B+1))$
Vegind.2	VARI	Visible Athmospheric Resistance Index	$(G-R)/(G+R-B)$
Vegind.3	NDTI	Normalized Difference Turbidity Index	$(R-G)/(R+G)$
Vegind.4	RI	Redness Index	$R**2/(B*G**3)$
Vegind.5	CI	Soil Color index	$(R-G)/(R+G)$
Vegind.6	BI	Brightness Index	$\text{sqrt}((R**2+G**2+B*2)/3)$
Vegind.7	SI	Spectral Slope Saturation index	$(R-B)/(R+B)$
Vegind.8	HI	Primary Color Hue Index	$(2*R-G-B)/(G-B)$
Vegind.9	TGI	Triangular Greenness Index	$-0.5[190(R670-R550)-120(R670-R480)]$
Vegind.10	GLI	Green Leaf index	$(2*G-R-B)/(2*G+R+B)$
Vegind.11	NGRDI	Normalized Green Red Difference Index	$(G-R)/(G+R)$

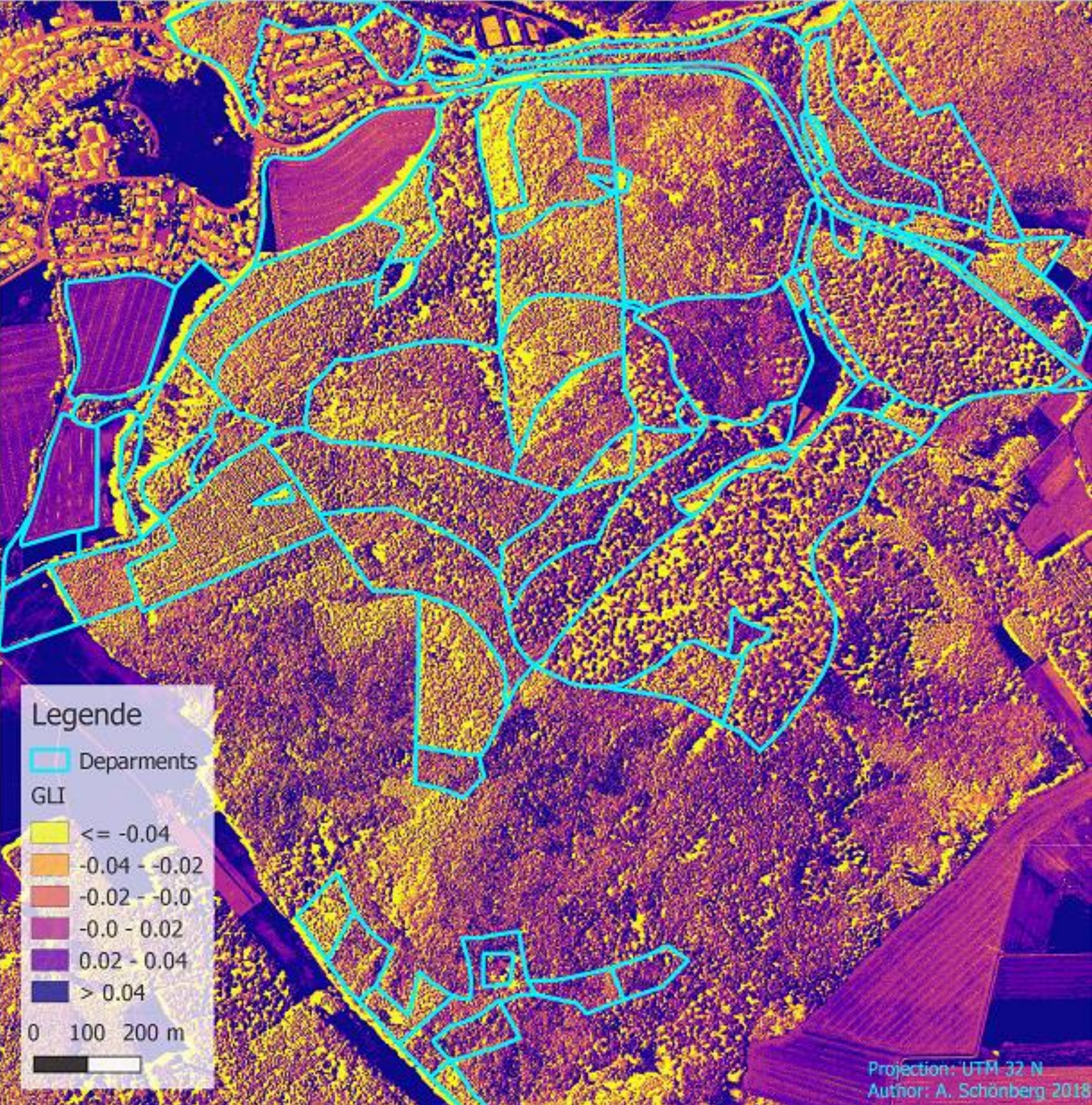
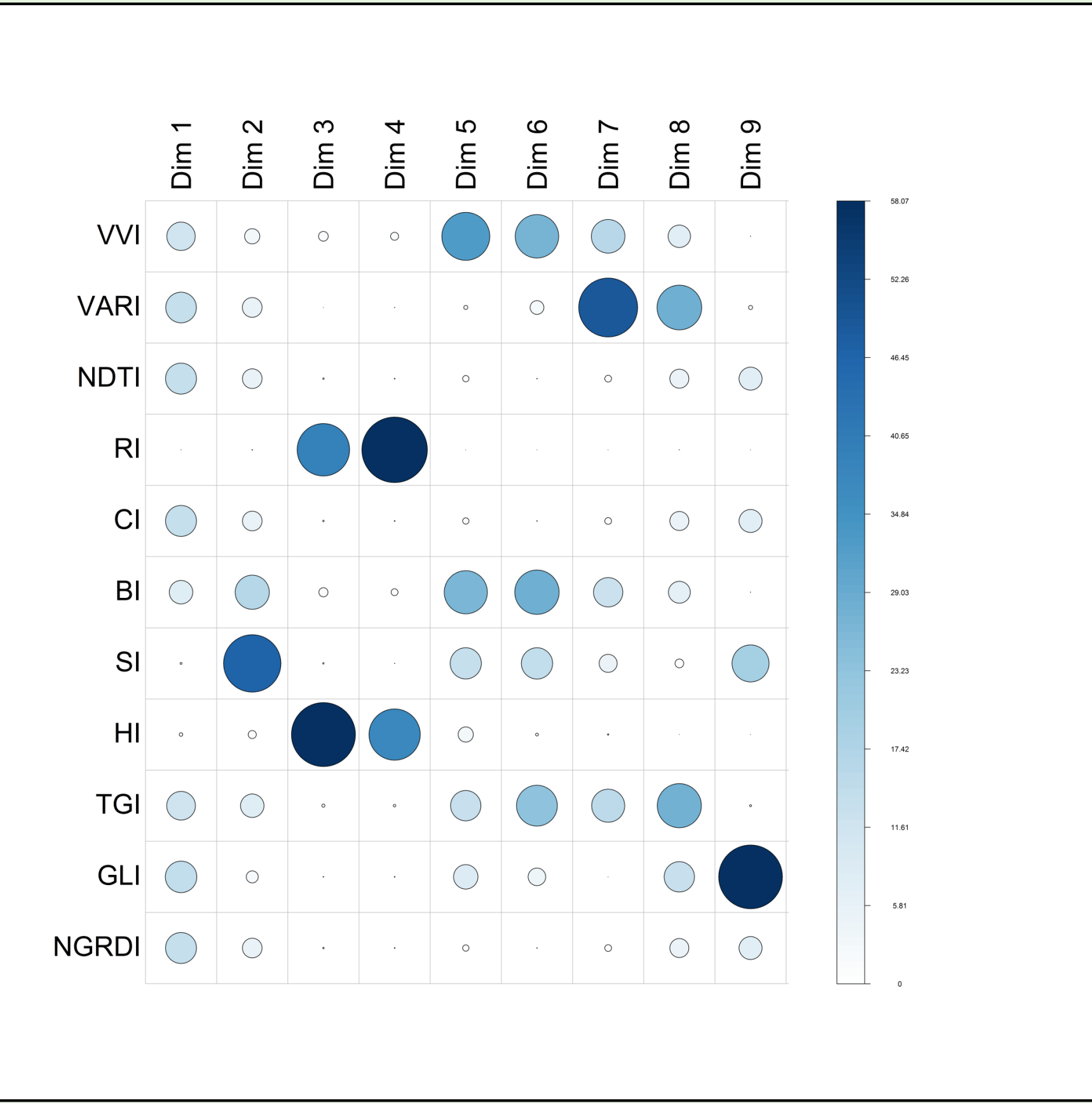
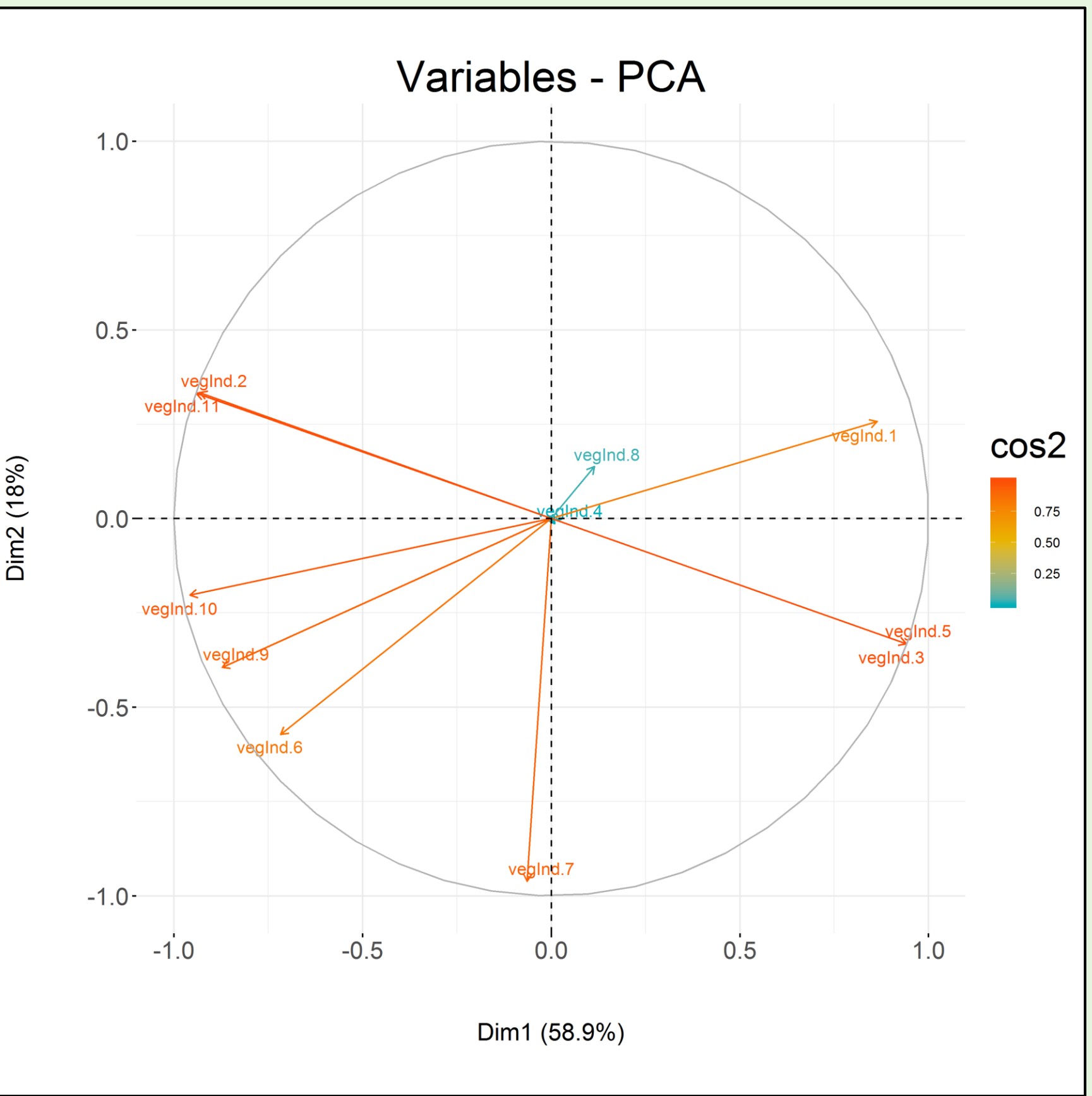
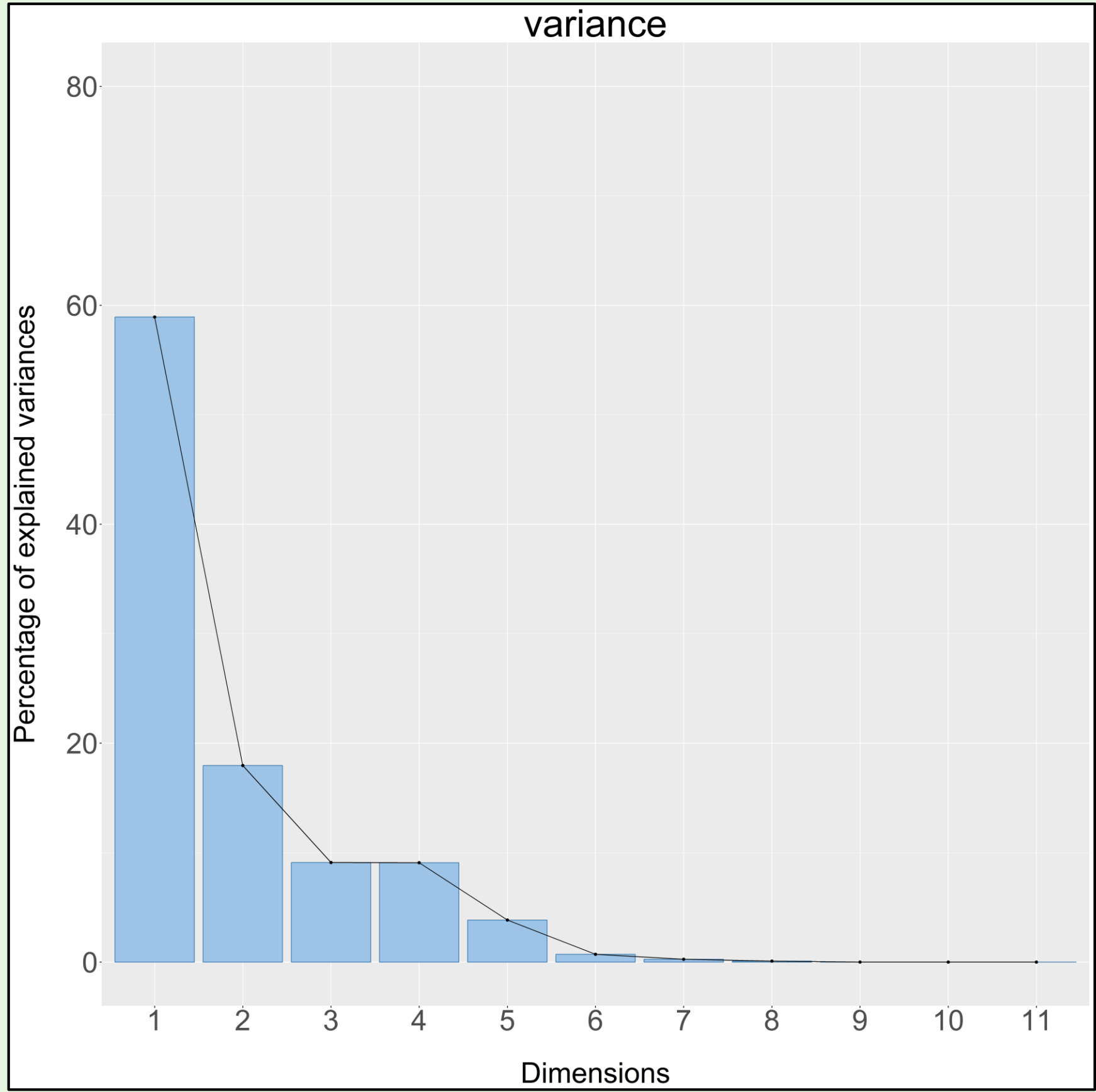


Figure 1: Study area with departments

METHODS AND WORKFLOW |

- Assimilate and process aerial RGB images
 - Use GIS software and related functionalites in R to create mosaic of RGB images provided by the Hessisches Landesamt für Naturschutz, Umwelt und Geologie (2016)
 - RGB images have many advantages: inexpensive to produce, contain enough information to run analyses and are easy to reproduce
- Compute various vegetation indices
 - e.g. the Green leaf index (GLI) and Triangular greenness index (TGI) (Hunt, 2013)
 - Such indices can determine plant health or nitrogen requirements and provide additional information about the environment
- Calculate principal components
 - Addresses statistical collinearity
 - Indices often include similar information. PCA eliminates redundancy, while summarizing the most significant information (Demsar, 2013)
- Apply filters to PCA
 - Spatial (mean, min, max, sd) and Sobel (full, horizontal, vertical) filters with moving windows of 5, 7 and 9 pixels
 - Filters smooth the information provided to later applications, which improve their performance in identifying features (see Figure 1)

In further analysis steps, this processed data will serve as predictor variables for training data in machine learning algorithms.



Figures 2, 3 & 4: Statistical indicators of the principal component analysis (PCA) on the visible vegetation indices.

References:
[1] DEMSAR, U., HARRIS, P., BRUNSDON, C., FOTHERINGHAM, A. & S. MCLOONE (2013): Principal Component Analysis on Spatial Data: An Overview. In: Annals of the Association of American Geographers 103(1). 106-128.
[2] HUNT, E., DORAISWAMY, P., MCMURTREY, J., DAUGHTRY, C., PERRY, E. & B. AKHMEDOV (2013): A visible band index for remote sensing leaf chlorophyll content at the canopy scale. In: International Journal of Applied Earth Observation and Geoinformation 21. 103-112.
[3] Hessisches Landesamt für Naturschutz, Umwelt und Geologie (2016).