# Supplementary Materials: Linking Land Surface Phenology and Vegetation-Plot Databases to Model Terrestrial Plant α-Diversity of the Okavango Basin

Rasmus Revermann, Manfred Finckh, Marion Stellmes, Ben Strohbach, David Frantz and Jens Oldeland

Table S1. Validation results for the two model types boosted regression trees (BRT) and random forest (RF) on the three subsets of the predictor variables (a) remote sensing and topography 'RS TOPO' (b) only climate data derived from CRU and ARC2 'CLIMATE CRU/ARC2', (c) all data 'ALL2' ('RS TOPO and 'CLIMATE CRU/ARC2'). The following performance measures were calculated: explained variance (expl. var. [%]), Pearson's correlation coefficient (r_P) between observed and predicted values, coefficient of determination (R^2), the root mean square error (RMSE, in species per 10^3 m^2) and the RMSE normalized by the mean, the relative root mean square error (rRMSE in per cent).The results for training and testing data are displayed (training 80% of the data and testing 20%).

		Expl. Var.	Correlation (r_P)		R^2		RMSE		rRMSE	
		Train	Train	Test	Train	Test	Train	Test	Train	Test
BRT	RS TOPO	54	0.78	0.69	0.61	0.48	10.1	11	30.2	31.9
	CLIMATE CRU/ARC2	64	0.85	0.78	0.73	0.61	8.5	9.8	24.3	27.0
	ALL2	67	0.86	0.76	0.74	0.57	8.3	10.3	23.8	28.5
RF	RS TOPO	43	0.94	0.70	0.89	0.49	5.9	10.9	17.2	30.4
	CLIMATE CRU/ARC2	50	0.93	0.79	0.87	0.62	5.9	9.6	17.0	26.4
	ALL2	54	0.95	0.78	0.90	0.61	5.5	9.7	15.8	26.8
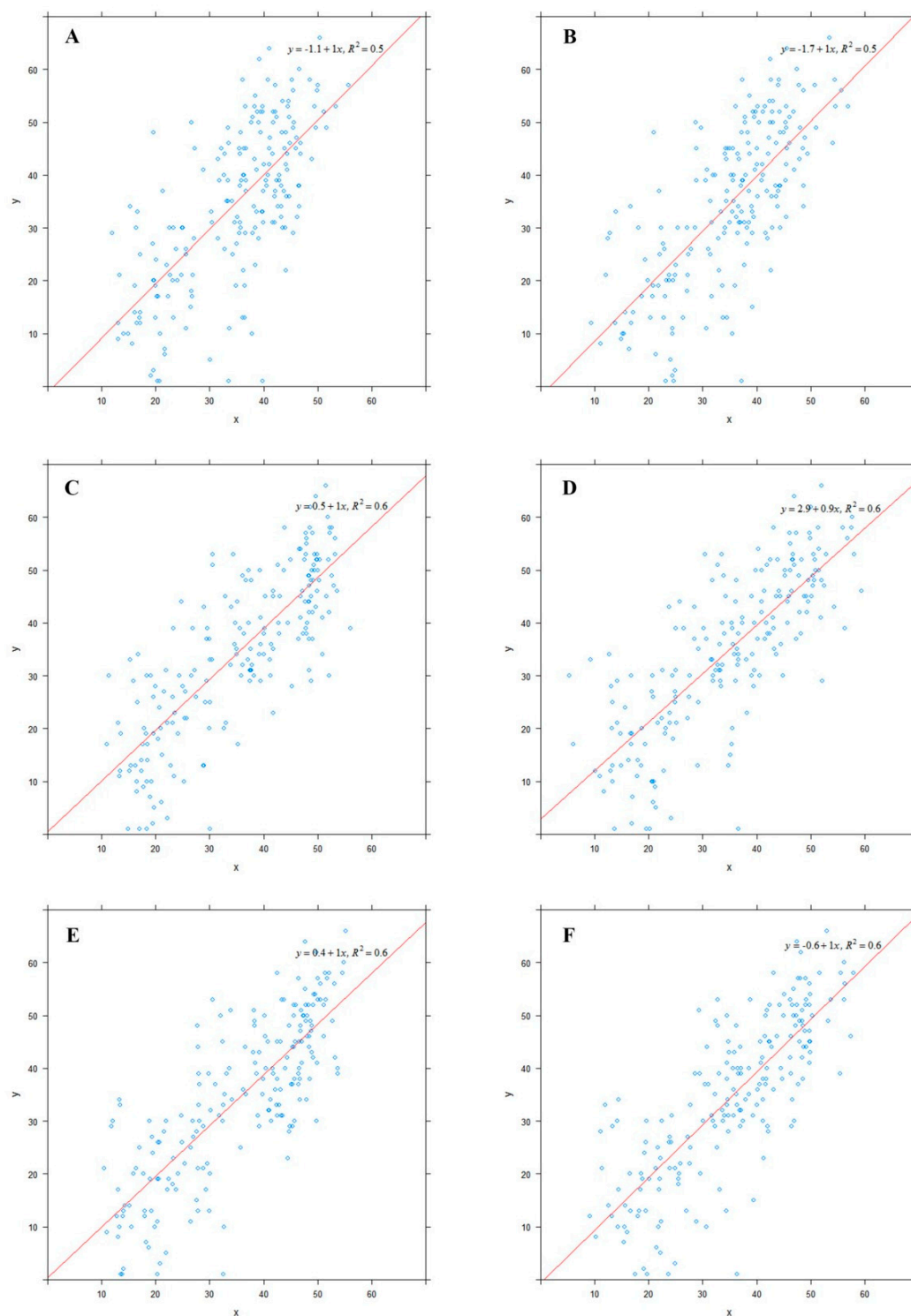
Figure S1. Observed values of alpha diversity plotted against predicted values on training data for (**A**) BRT on data set 'RS TOPO'; (**B**) RF on data set 'RS TOPO'; (**C**) BRT on data set 'CLIMATE'; (**D**) RF on data set 'CLIMATE'; (**E**) BRT on data set 'ALL'; (**F**) RF on data set 'ALL'.
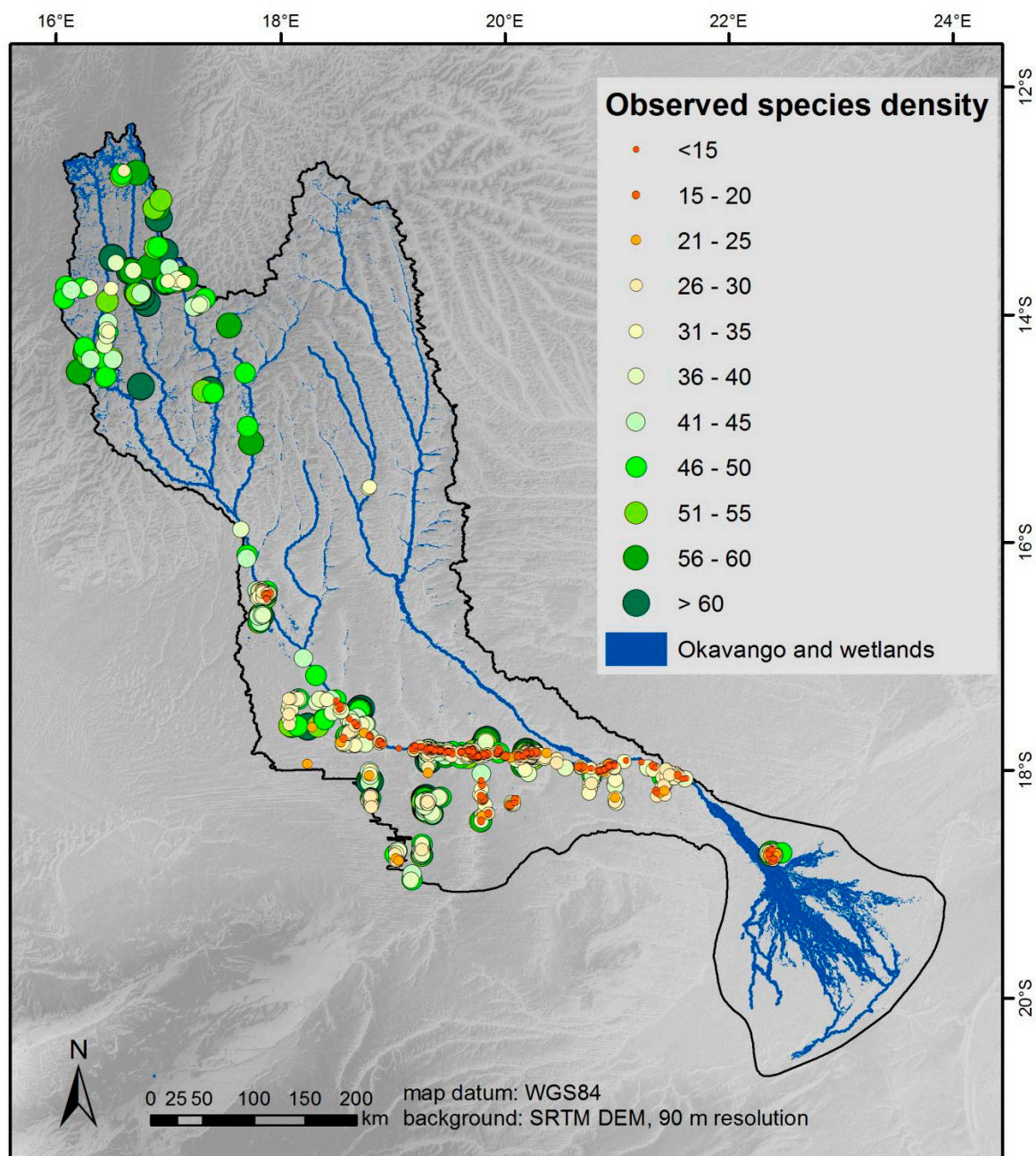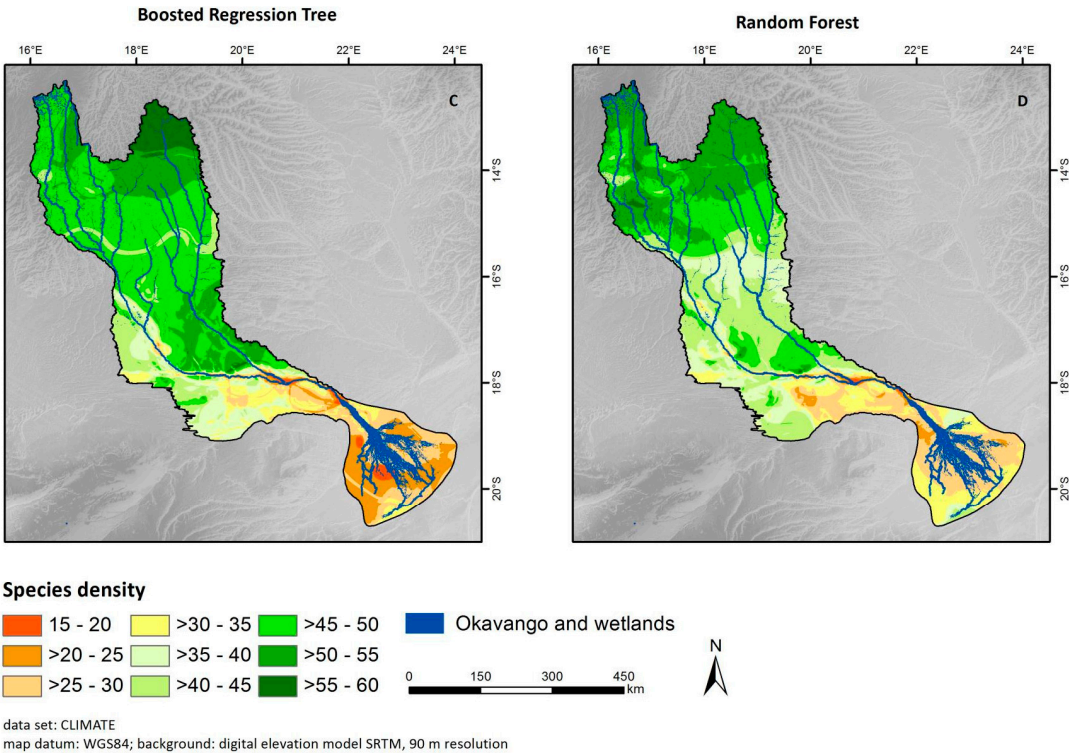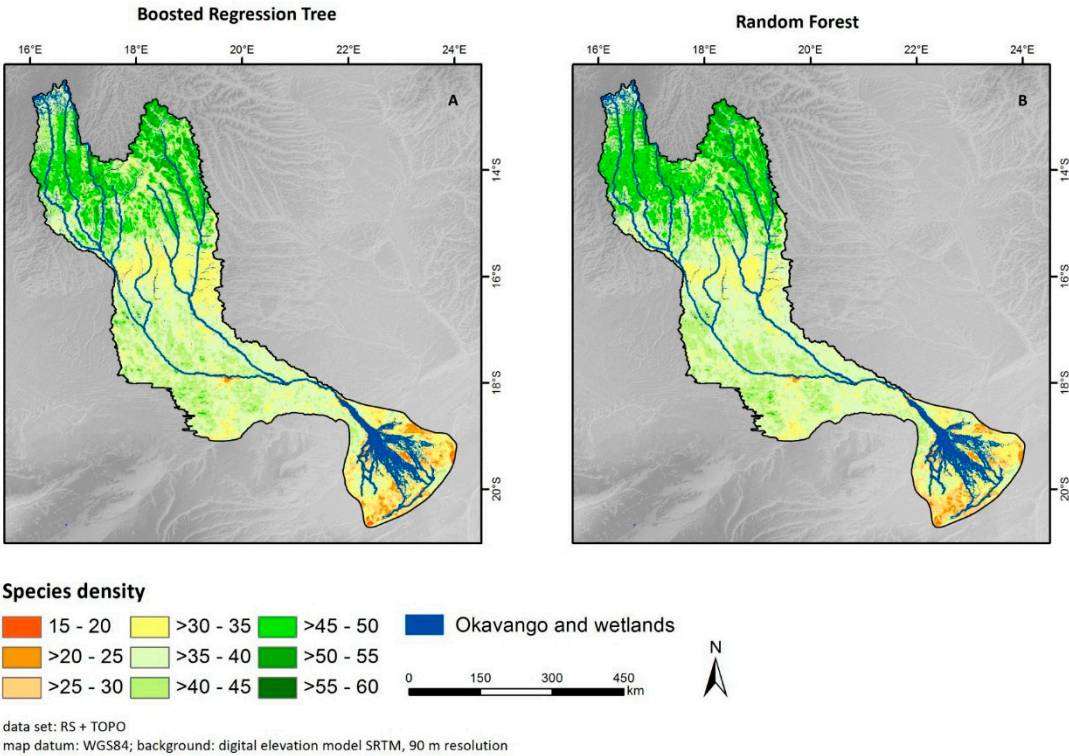
Figure S2. Observed plant alpha diversity (species density per 10^3 m^2). Data is based on 999 vegetation plots sized 20 × 50 m.
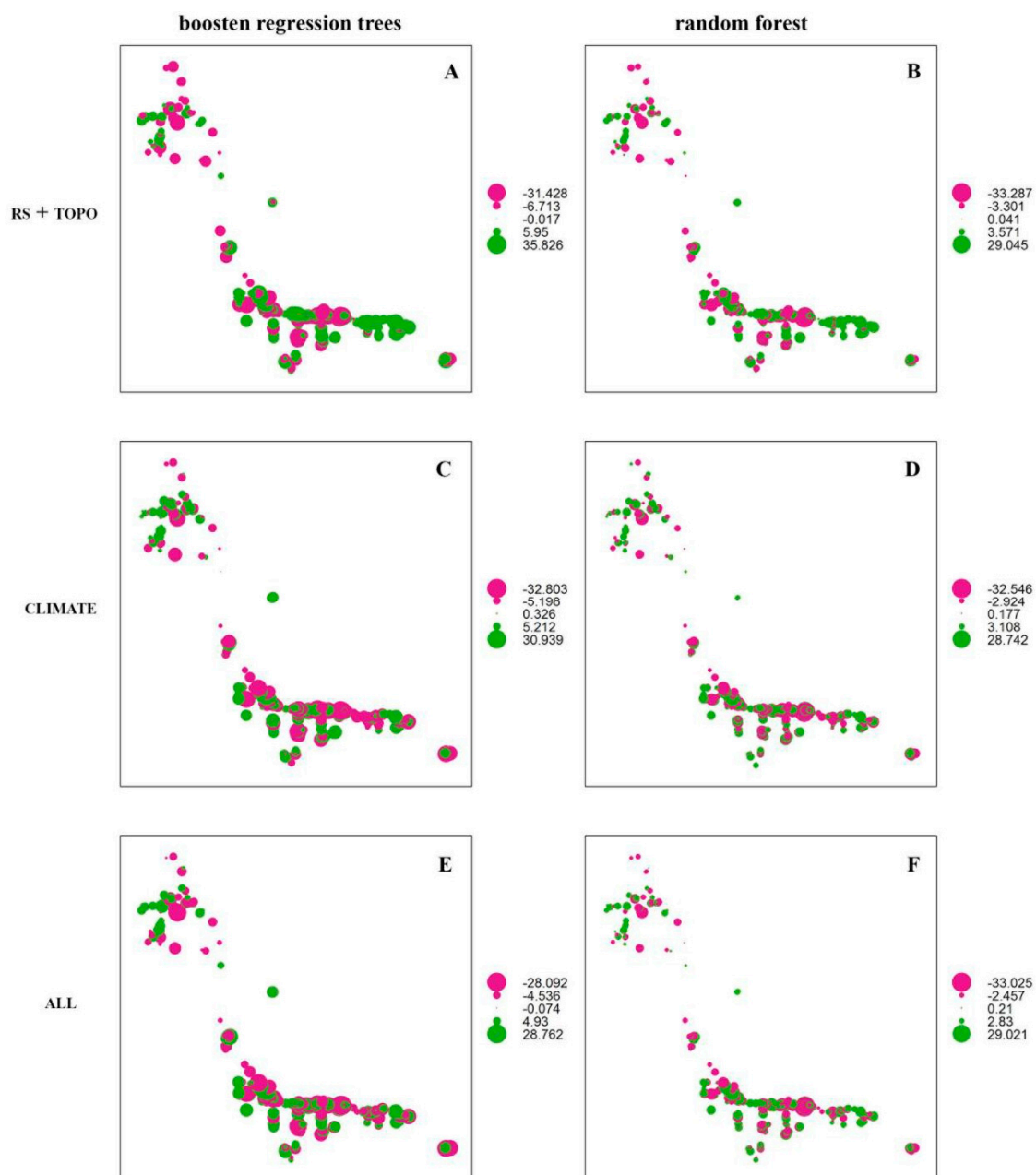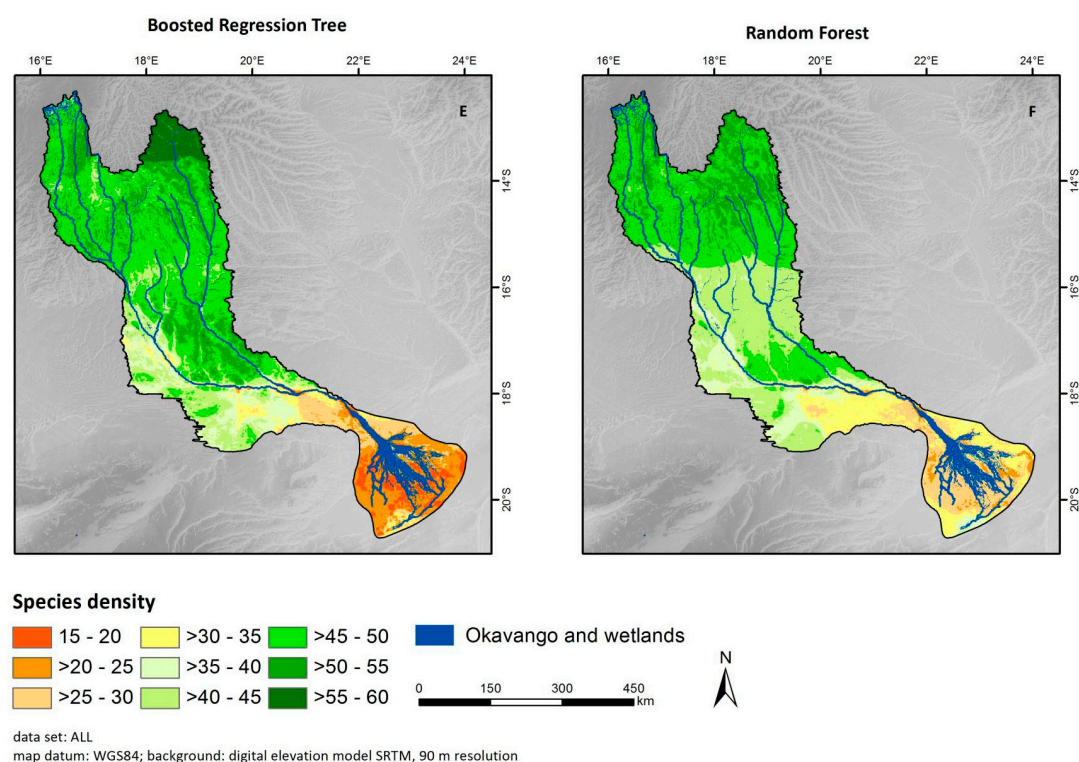
Figure S3. Model residual for the two model types: boosted regression trees (**A**,**C**,**E**) and random forest (**B**,**D**,**E**) on the three datasets: 'RS TOPO' (**A**, **B**); 'CLIMATE' (**C**, **D**); 'ALL'; (**G**, **H**). Furthermore, we calculated variograms to check for spatial autocorrelation but no sever spatial auto correlation was detected.
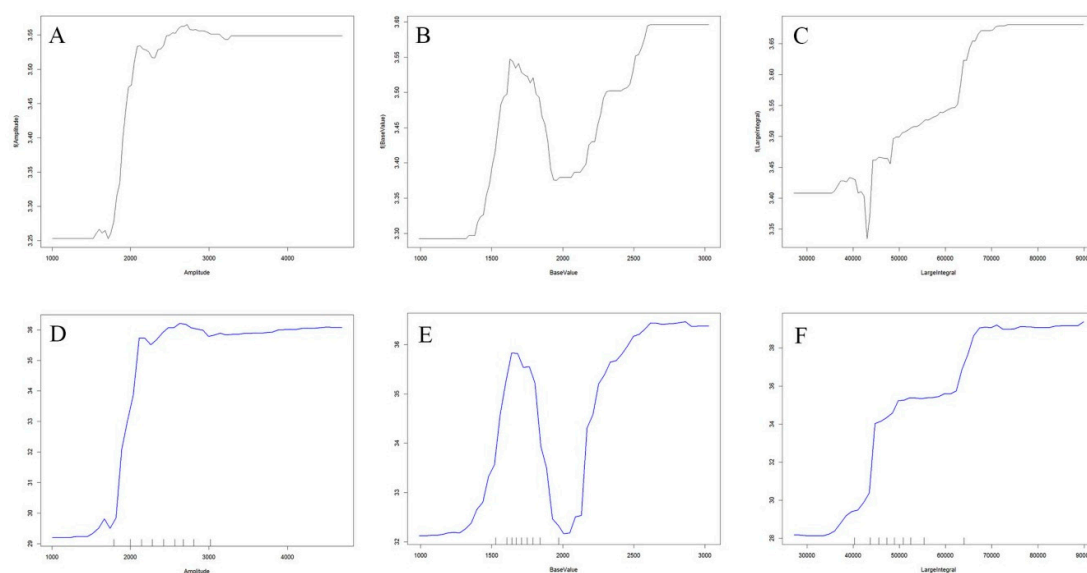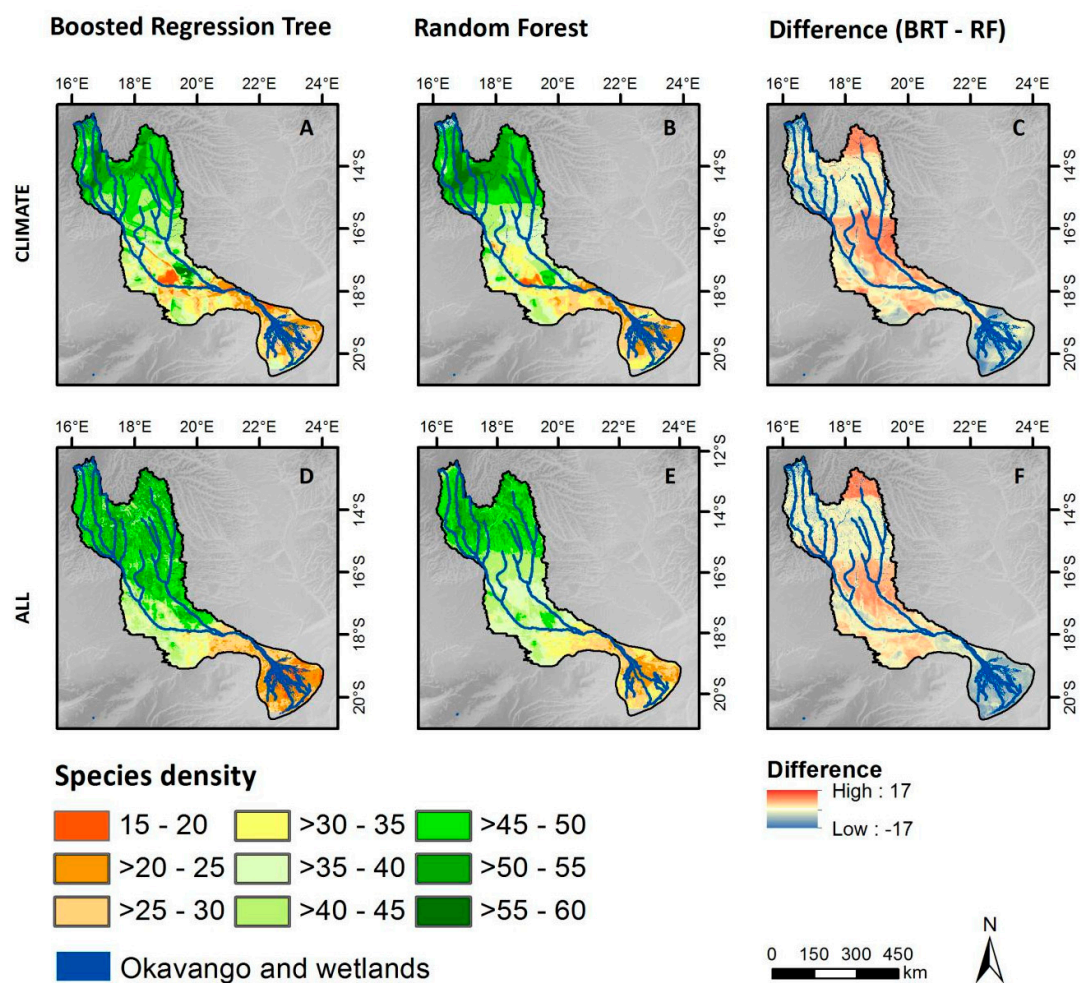
FigureS4. *Cont.*

Figure S4. Plant alpha diversity (species density per 10^3 m²) predicted by the two model types: BRT (**A**,**C**,**E**) and random forest (**B**,**D**,**F**) on the three data (sub-)sets: 'RS TOPO' (**A**,**B**); 'CLIMATE' (**D**,**E**); 'ALL' (**E**,**F**).

Figure S5. Partial dependence plots of the LSP metrics 'Amplitude' (**A**,**D**), 'BaseValue' (**B**,**E**) 'LargeIntegral' (**C**,**F**) for the two model types BRT (**A–C**) and RF (**D–F**).

Figure S6. Plant alpha diversity (species density per 10^3 m^2) predicted by the two model types: BRT (**A**,**D**) and random forest (**B**,**E**) on the second climate data set CRU/ARC2 (**A**,**B**); and on the entire data set comprising the second climate data set CRU/ARC2 and remote sensing data (**D**,**E**) and the difference between the two model algorithms (**C** and **F**). For a map on observed species density see Figure S1.