Supplement of Earth Syst. Sci. Data, 18, 2667–2687, 2026
https://doi.org/10.5194/essd-18-2667-2026-supplement
© Author(s) 2026. CC BY 4.0 License.

Open Access Earth System Science Data

Supplement of

A dataset on the structural diversity of European forests

Marco Girardello et al.

Correspondence to: Marco Girardello (marco.girardello@gmail.com) and Gonzalo Oton (gonzalo.oton@ec.europa.eu)

The copyright of individual parts of the supplement might differ from the article licence.

Table S1. Criteria used for filtering GEDI data

Description	Justification
Dropped the shots with quality_flag = 0	Guarantees good waveform quality
Observations limited to nighttime (solar elevation ≤ 0 degrees)	Mitigates influence of background noise from reflected solar radiation
Dropped the shots with degrade_flag = 0	Avoids potential geolocation errors under suboptimal conditions
Dropped the shots with footprint slope > 20 degrees	Prevents inaccuracies in vegetation height on steep terrain. Slope is determined by the Shuttle Radar Topography Mission (SRTM) data (see GEE code).
Dropped the shots with urban_proportion > 5%	Focuses on non-urbanized forest settings
Dropped the shots with landsat_water_persistence = 0	Eliminates areas with consistent water coverage
Dropped the shots with landsat_tree_cover < 10%	Confirms significant forest canopy presence
Dropped the shots with rh98 > 100 m	Maintains focus on typical forest canopy heights, avoiding outliers
Dropped the shots with rh50 ≤ 0 m	Confirms presence of vertical structure in vegetation
Dropped the shots with total canopy cover ≤ 0.1	Indicates significant presence of canopy materials
Dropped the shots with leaf-off_flag ≠ 0	Ensures data represents vegetation during growing season, avoiding underestimation of structural complexity in deciduous trees
Dropped the shots with num_detectedmodes = 0	A signal without modes is indicative of pure noise, lacking information on forest vertical structure
Dropped the shots with surface_flag = 0	Ensures presence of ground returns in waveform, essential for accurate vertical structure determination
Dropped the shots with sensitivity < 0.95	Provides high confidence in proper representation of forest's vertical structure

Table S2. List of the remote sensing predictors used for Random Forest Modelling. $\beta =$ SM stands for spatial mean. $\beta =$ ASM, ENT, DISS stands for the texture metrics Angular Second Moment (ASM), entropy (ENT), and dissimilarity (DISS) index, respectively, see Section 1 in the main text for more details. S1, S2, and AP2 stands for Sentinel-1, Sentinel-2 and AlOS-PALSAR-2, respectively.

Predictor	**Description**	β	**Composite**	**Sensor**
$\phi^{\beta}_{S1VVgs\mu}, \phi^{\beta}_{S1VHgs\mu}$	Growing season VV and VH mean	SM, ASM, ENT, DISS	Six-month	S1
$\phi^{\beta}_{S1VVgs\sigma}, \phi^{\beta}_{S1VHgs\sigma}$	Growing season VV and VH standard deviation	SM	Six-month	
$\phi^{\beta}_{S1VVpre\mu}, \phi^{\beta}_{S1VHpre\mu}$	Pre-peak growing season mean	SM	Two-month	
$\phi^{\beta}_{S1VVact\mu}, \phi^{\beta}_{S1VHact\mu}$	Peak growing season mean	SM	Two-month	
$\phi^{\beta}_{S1VVpost\mu}, \phi^{\beta}_{S1VHpost\mu}$	Post-peak growing season mean	SM	Two-month	
ϕ^{β}_{CO}	Coherence 12 days summer	SM	Summer	
$\phi^{\beta}_{AP2HH}, \phi^{\beta}_{AP2HV}$	Annual HH and HV	SM, ASM, ENT, DISS	Annual	AP2
ϕ^{β}_{NDVI}	Normalized difference vegetation index	SM, ASM, ENT, DISS	Six-month	S2
$\phi^{\beta}_{NDVI\sigma}$	Normalized difference vegetation index standard deviation	SM	Six-month	
ϕ^{β}_{NDWI}	Normalized difference water index	SM, ASM, ENT, DISS	Six-month	
ϕ^{β}_{NDRE}	Normalized difference red edge index	SM, ASM, ENT, DISS	Six-month	
ϕ^{β}_{MSAVI}	Modified soil adjusted vegetation index	SM, ASM, ENT, DISS	Six-month	
ϕ^{β}_{GNDVI}	Green normalized vegetation index	SM, ASM, ENT, DISS	Six-month	

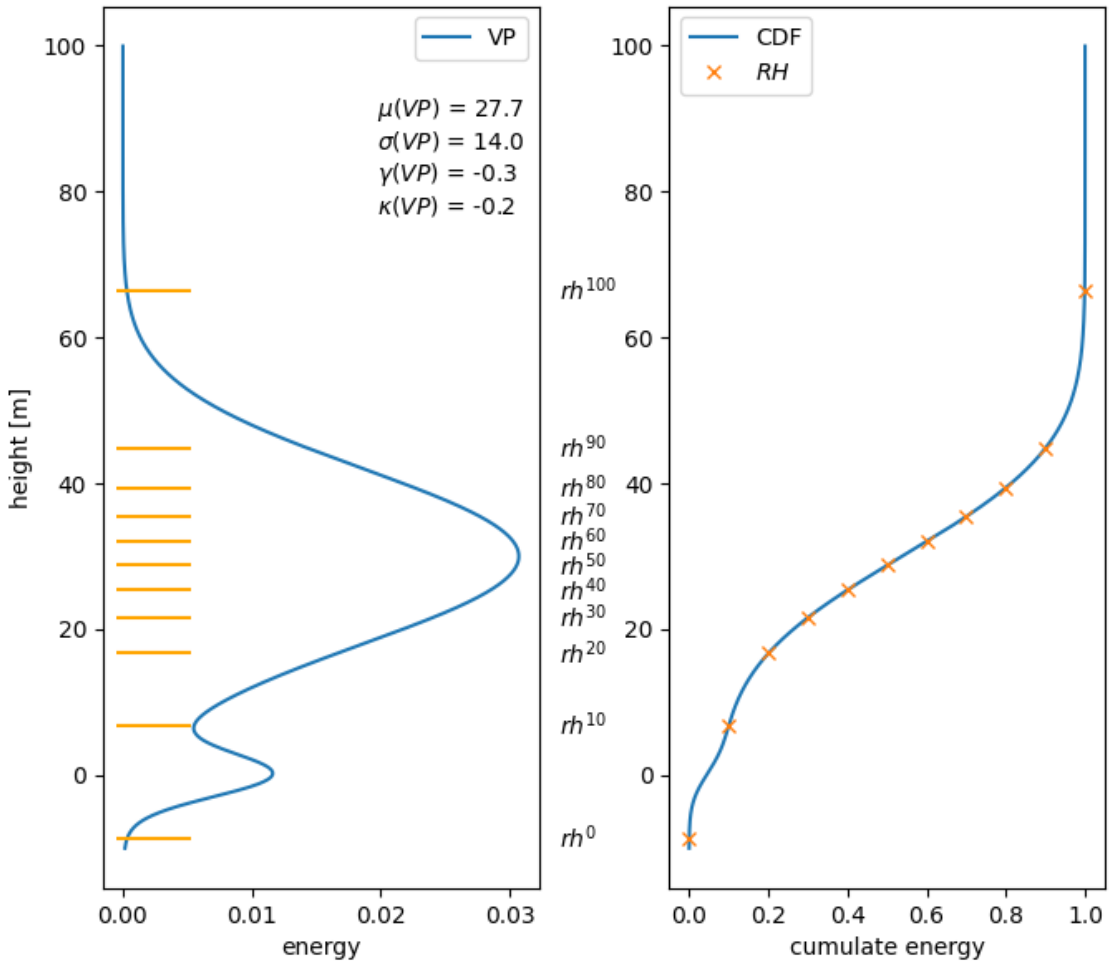

Figure S1. Vegetation vertical profile (denoted as VP) as detected by GEDI waveform (left panel). The numbers within the first panel indicate the four moments (for notation see Appendix A1 main text). The second panel shows the corresponding cumulative distribution function (CDF) of the relative heights (RH) (right panel).

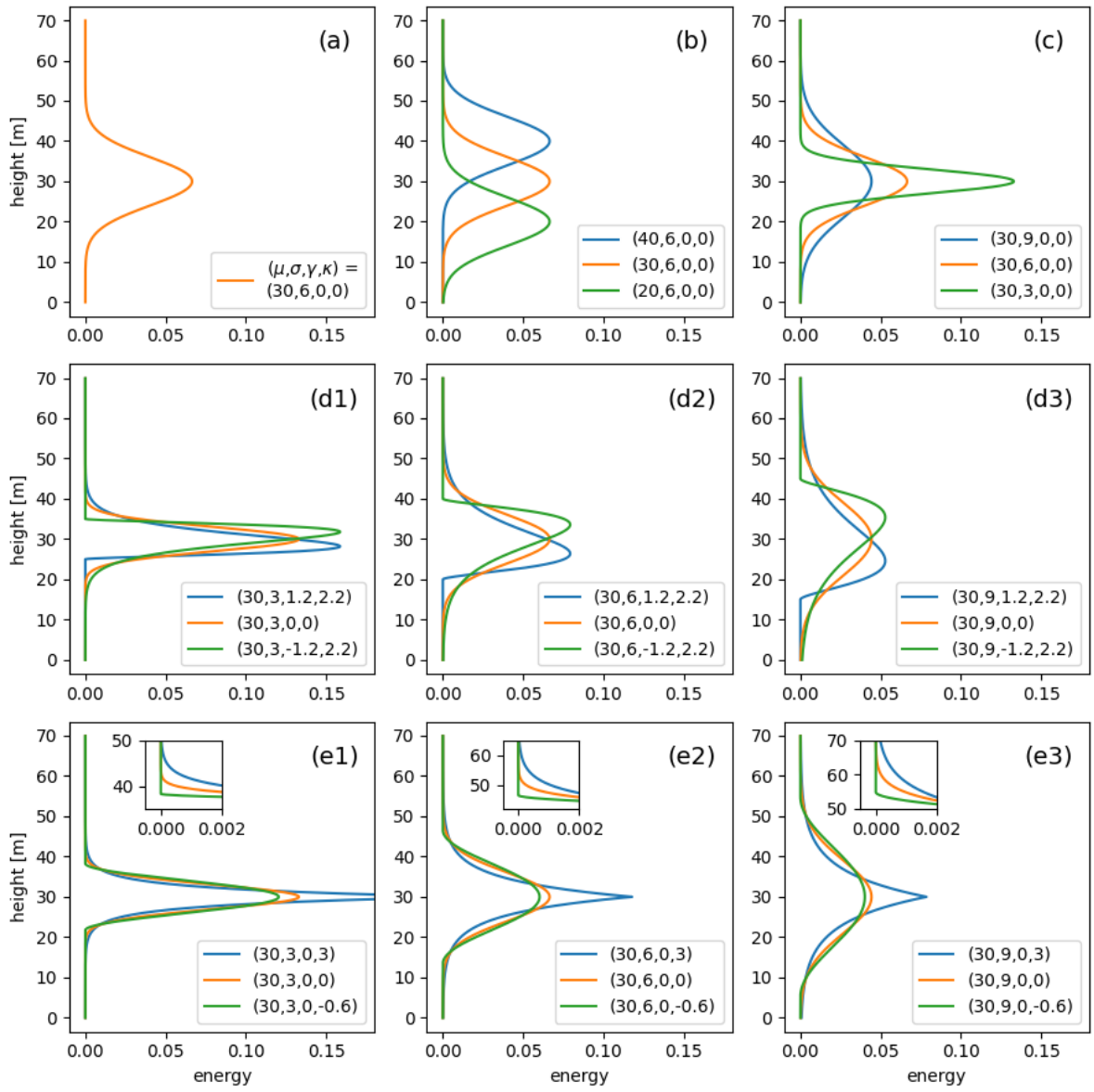

Figure S2. Examples of different unimodal vertical profiles and their relative moments μ (mean), σ (standard deviation), γ (skewness), and κ (excess kurtosis). All panels display the four moments for each profile shown; panels differ in which parameter is varied to demonstrate its effect on vertical profile shape. Panel a shows a normal distribution (γ = 0, κ = 0) as reference. Panels b-c demonstrate variation in μ and σ, which determine the coefficient of variation. Panels d demonstrate variation in γ (skewness). Panels e demonstrate variation in κ (kurtosis). Panel (a): $\mu = 30$, $\sigma = 6$, $\gamma = 0$, and $\kappa = 0$ (i.e. normal distribution). Panel (b): variation of μ (20, 30, 40 m) with fixed σ (6 m), fixed γ (0), and fixed κ (0). Panel (c): variation of σ (3, 6, 9 m) with fixed μ (30), fixed γ (0), and fixed κ (0). Panels d: variation of γ (-1.2, 0, 1.2) with fixed μ (30 m), fixed κ (2.2), and σ equal 3 m (panel d1), 6 m (panel d2), and 9 m (panel d3). Panels e: variation of κ (-0.6, 0, 3) with fixed μ (30 m), fixed γ (0), and σ equal 3 m (panel e1), 6 m (panel e2), and 9 m (panel e3).

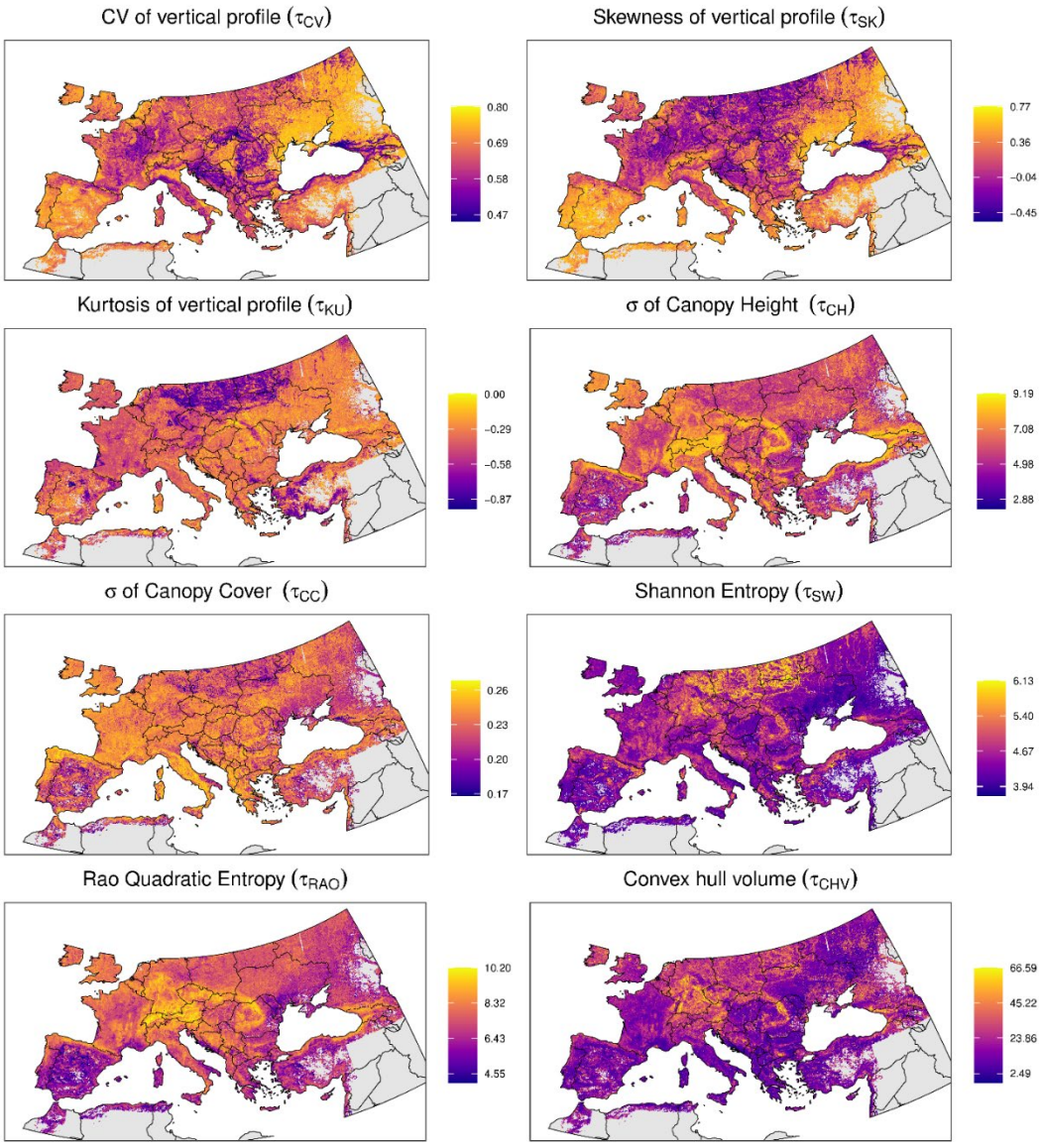

Figure S3 Mapped structural diversity at a 5 km resolution, derived from the Random Forest modelling. Each panel illustrates the geographic distribution of a specific metric (see methods for metric details). The colour palette transitions from purple to yellow, denote an increasing gradient of structural diversity, with warmer colours signifying higher values.

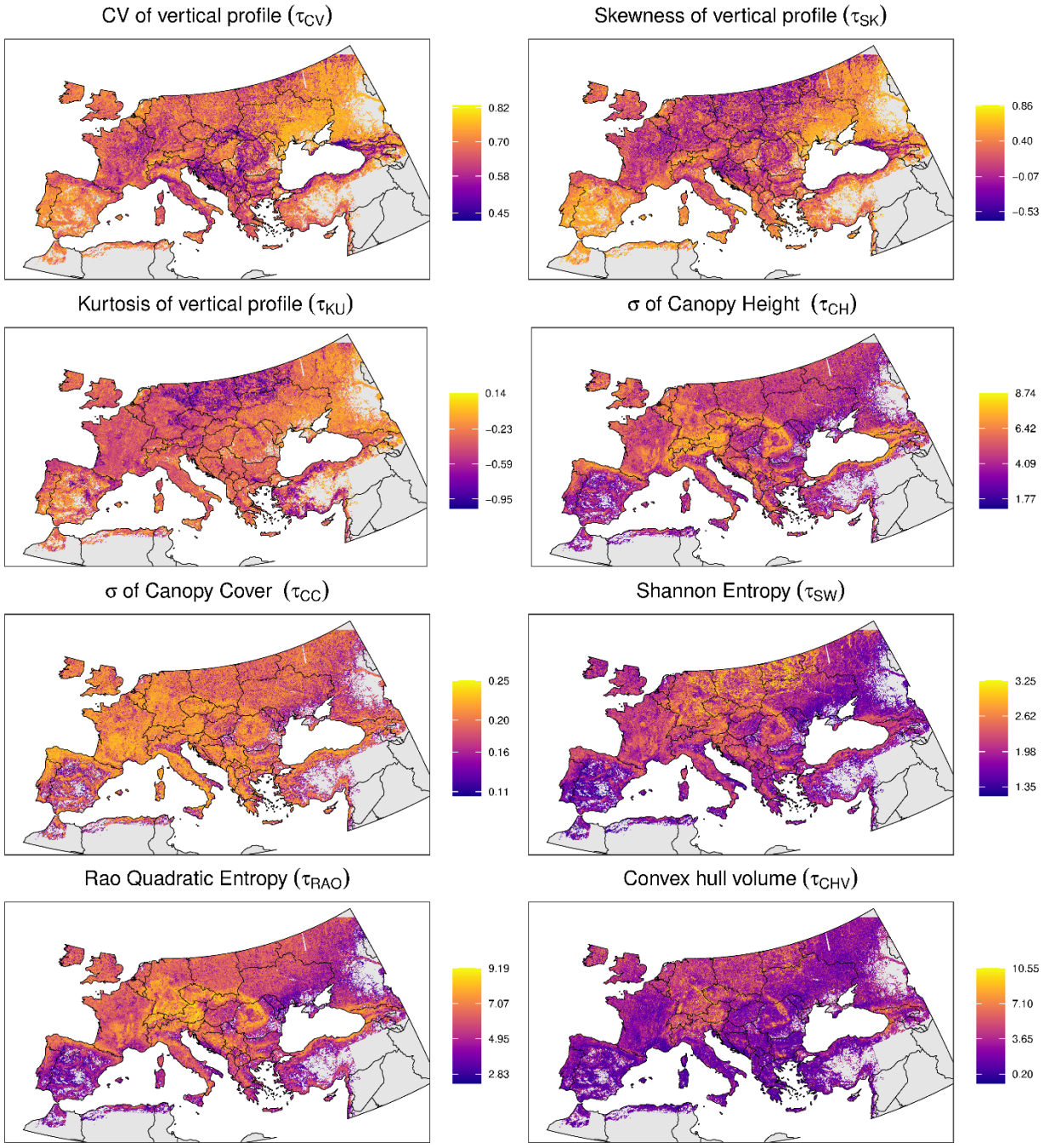

Figure S4 Mapped structural diversity at a 1 km resolution, derived from the Random Forest modelling. Each panel illustrates the geographic distribution of a specific metric (see methods for metric details). The colour palette transitions from purple to yellow, denote an increasing gradient of structural diversity, with warmer colours signifying higher values.

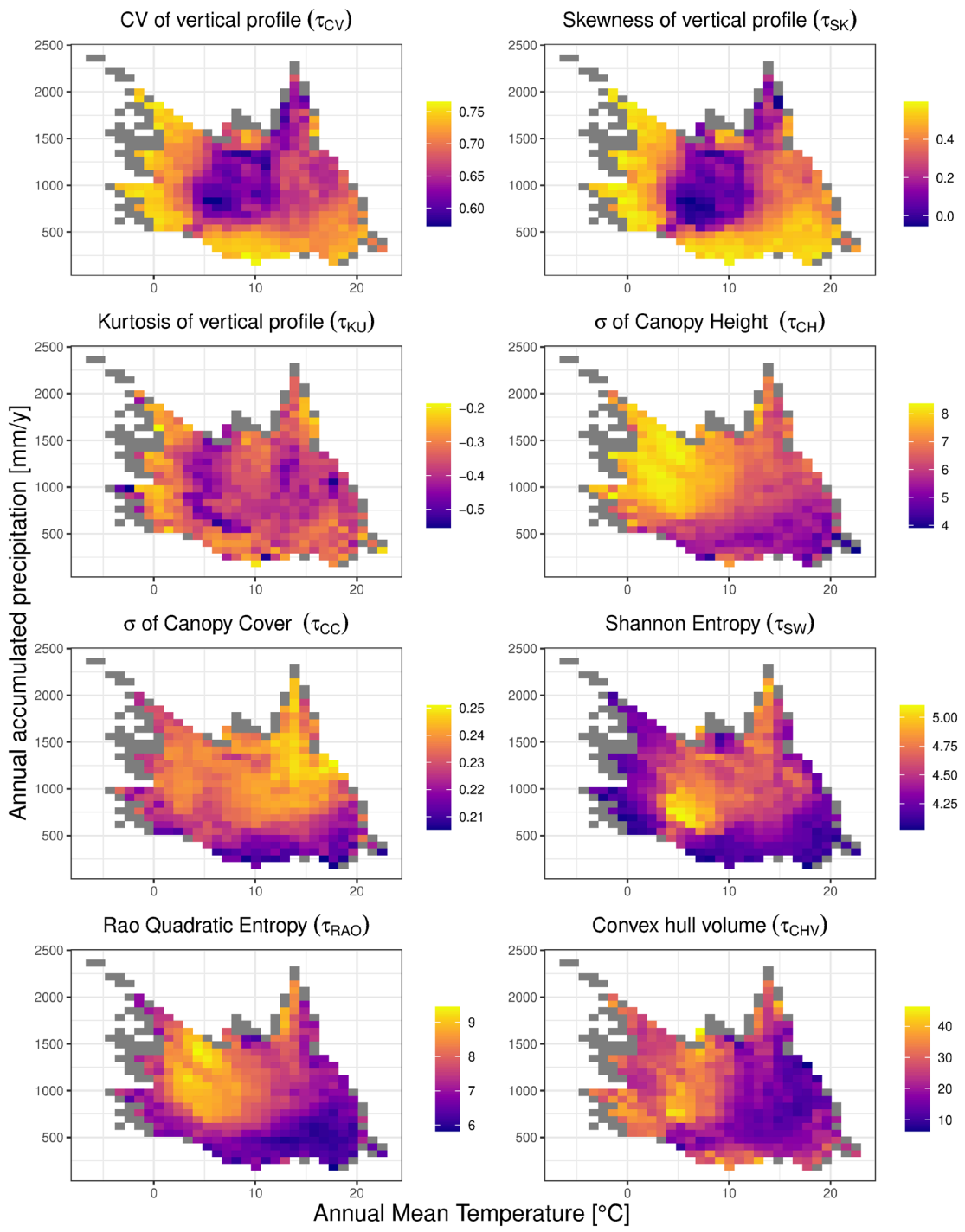

Fig. S5 Structural diversity variables mapped in climate space. The results refer to the dataset at 5 km resolution.

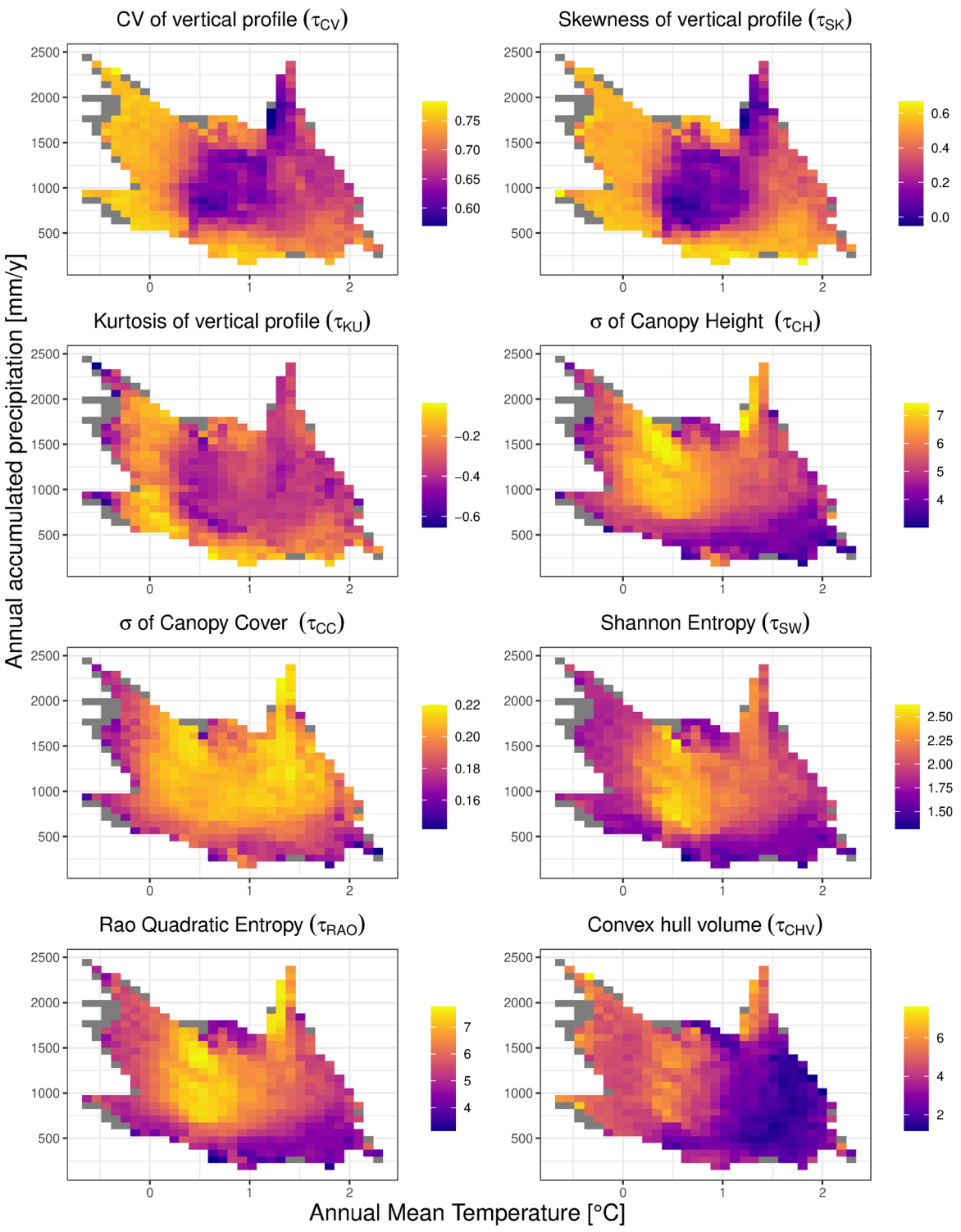

Fig. S6 Structural diversity variables mapped in climate space. The results refer to the dataset at 1 km resolution.

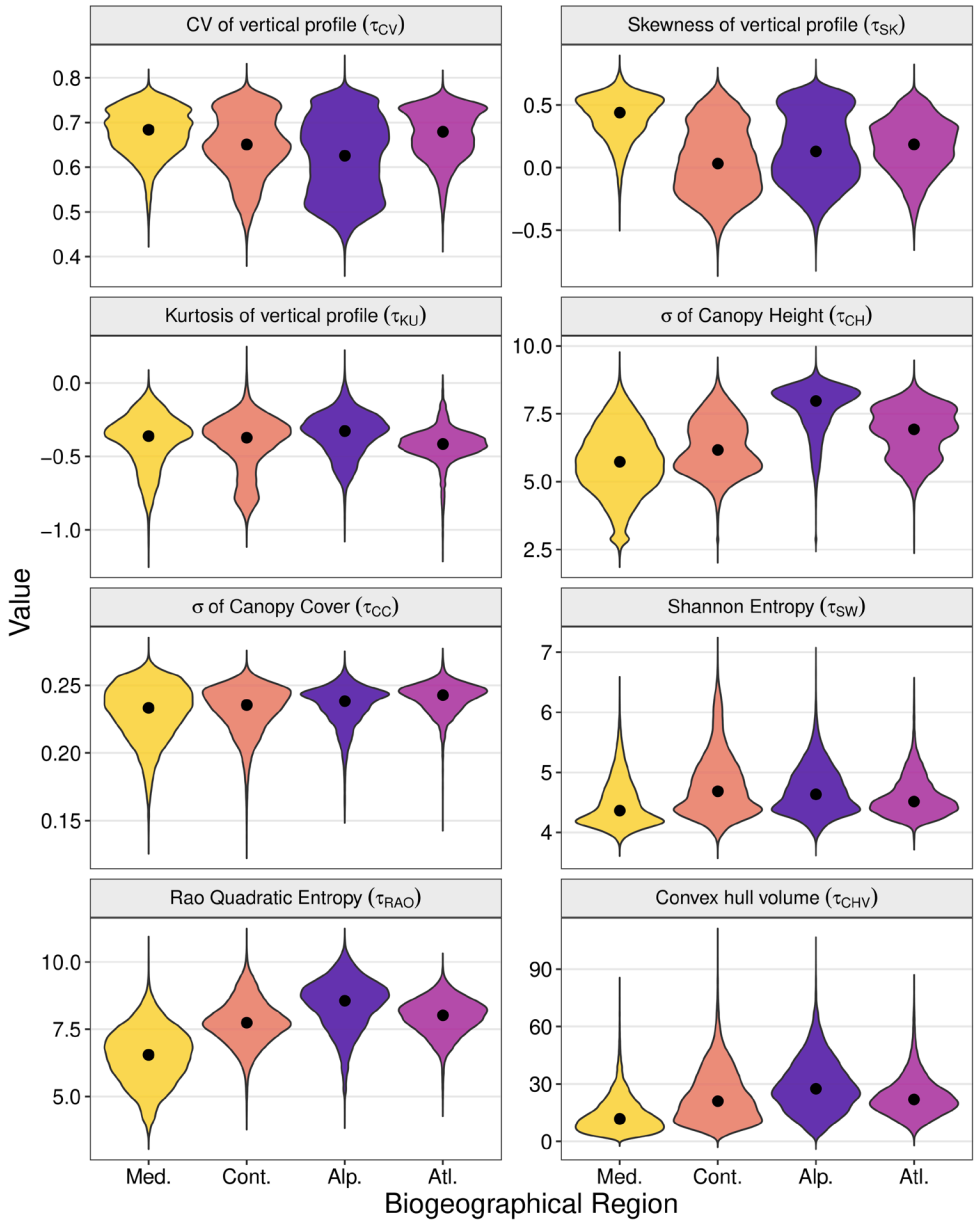

Fig. S7 Patterns of variability in forest structural diversity metrics across European biogeographic regions. Violin plots depict the probability density distribution of eight metrics at 5 km resolution. Width represents the kernel density estimate, with black points indicating median values. Colors denote the five biogeographic regions analysed. Med. = Mediterranean region, Cont. = Continental region, Alp. = Alpine region, Atl. = Atlantic region.

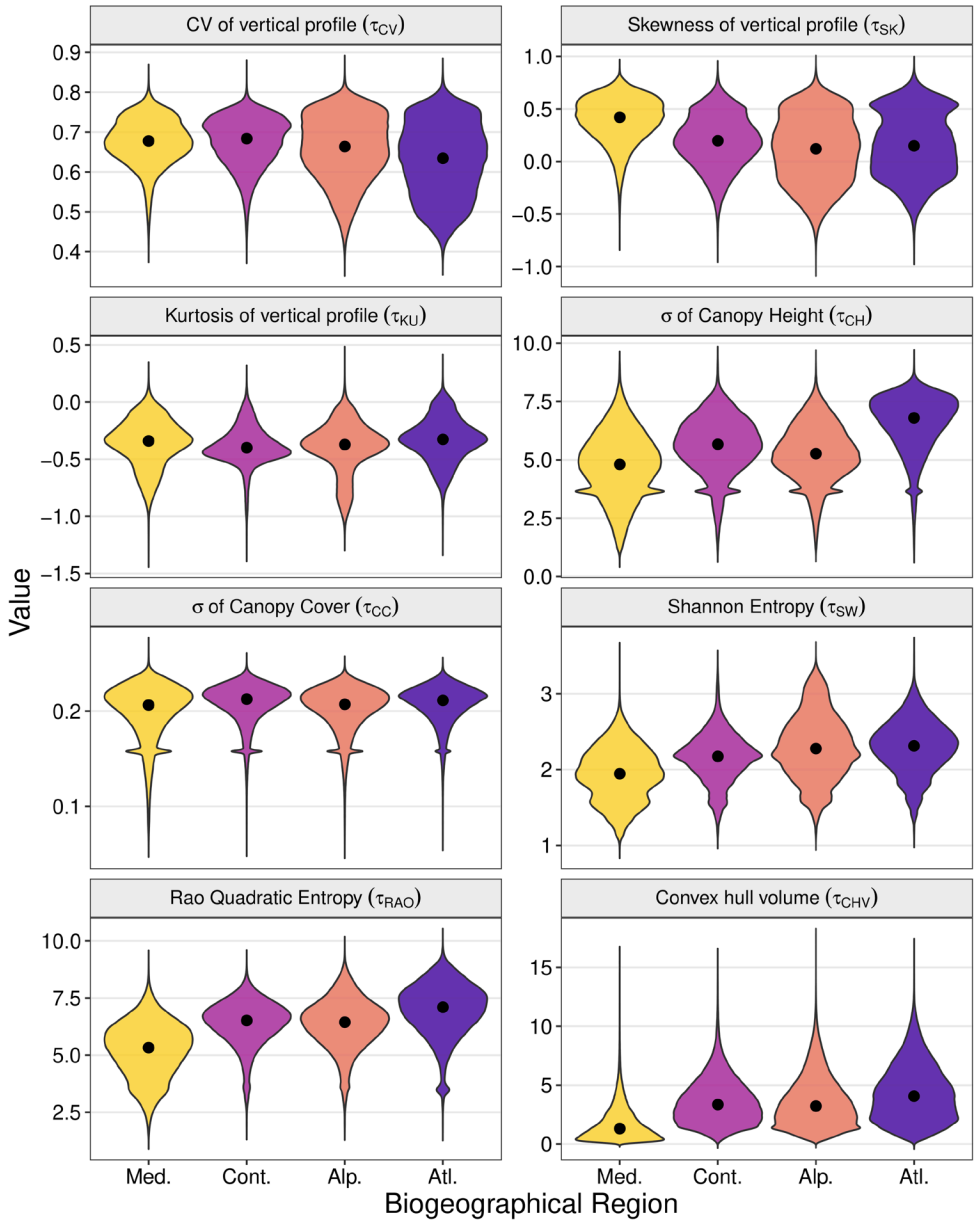

Fig. S8 Patterns of variability in forest structural diversity metrics across European biogeographic regions. Violin plots depict the probability density distribution of eight metrics at 1 km resolution. Width represents the kernel density estimate, with black points indicating median values. Colors denote the five biogeographic regions analysed. Med. = Mediterranean region, Cont. = Continental region, Alp. = Alpine region, Atl. = Atlantic region.

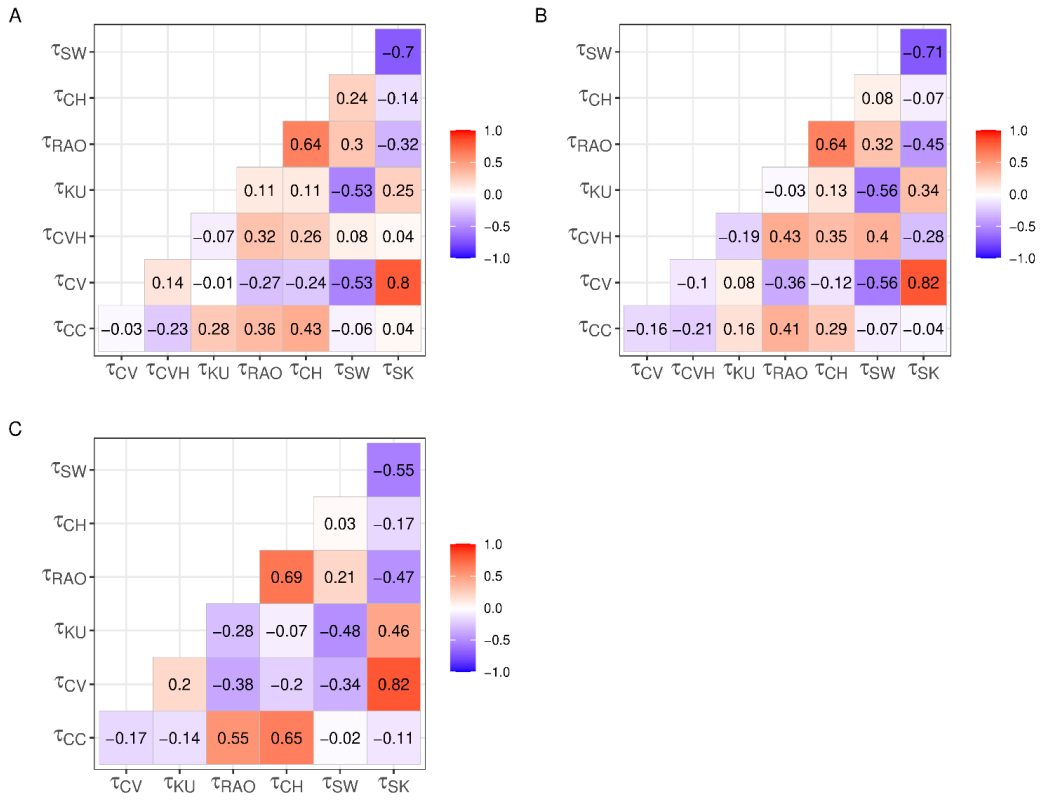

Figure. S9 Pairwise correlations for predicted structural diversity indices. Panel A shows the correlations among indices predicted at a 10 km x 10km resolution, Panel B at a 5km x 5km resolution, and Panel C a 1km x 1km. Each heatmap visualises the interrelationships of structural diversity indices derived from the Random Forest model, highlighting variations across resolutions.

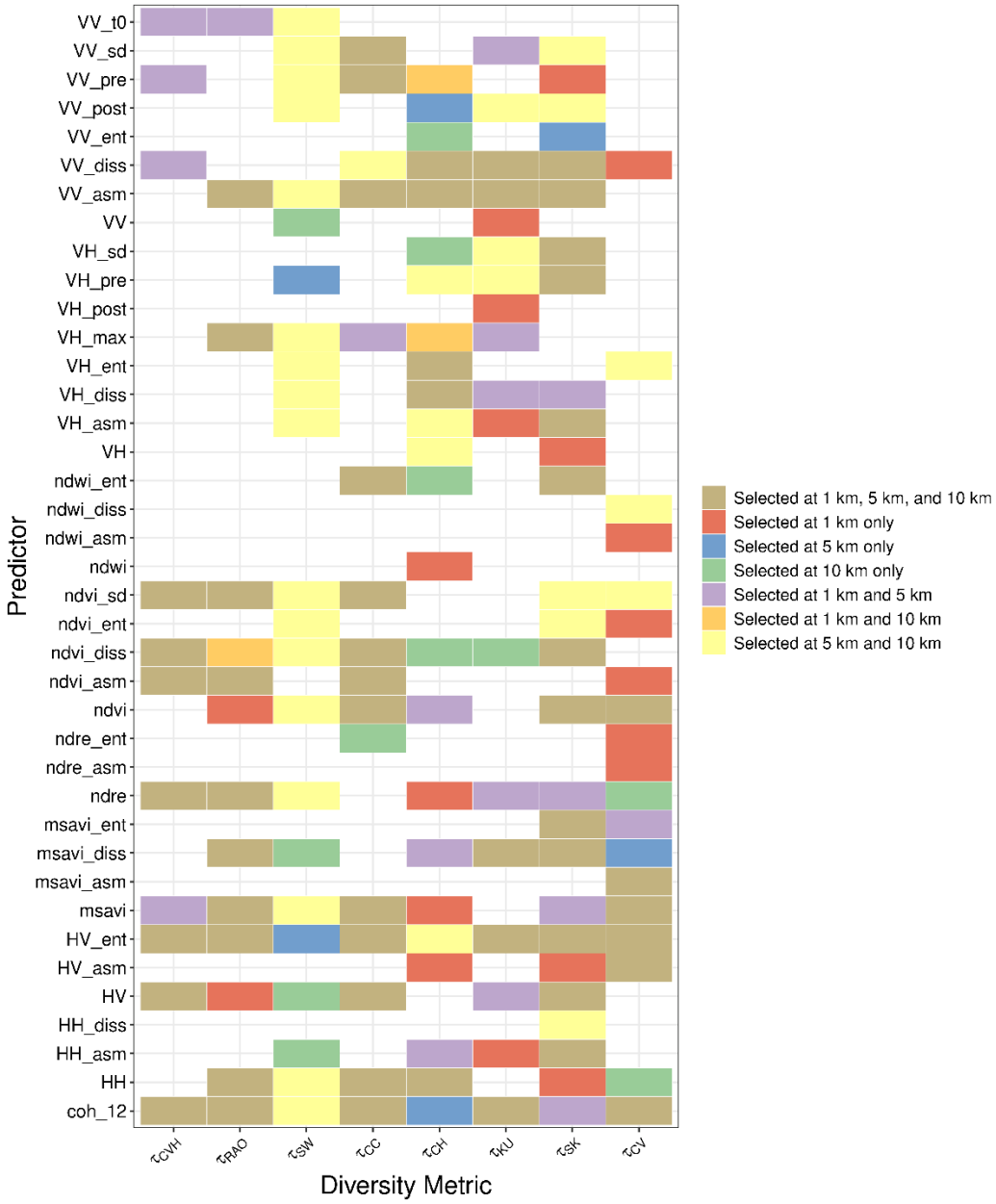

Figure S10 Summary of model selection results using 47 predictors as input variables for models fitted to 8 diversity metrics at different resolutions (1 km, 5 km, 10 km). The y-axis shows the predictors, and the x-axis shows the diversity metrics. Colours indicate whether a specific predictor was selected at one or more resolutions.

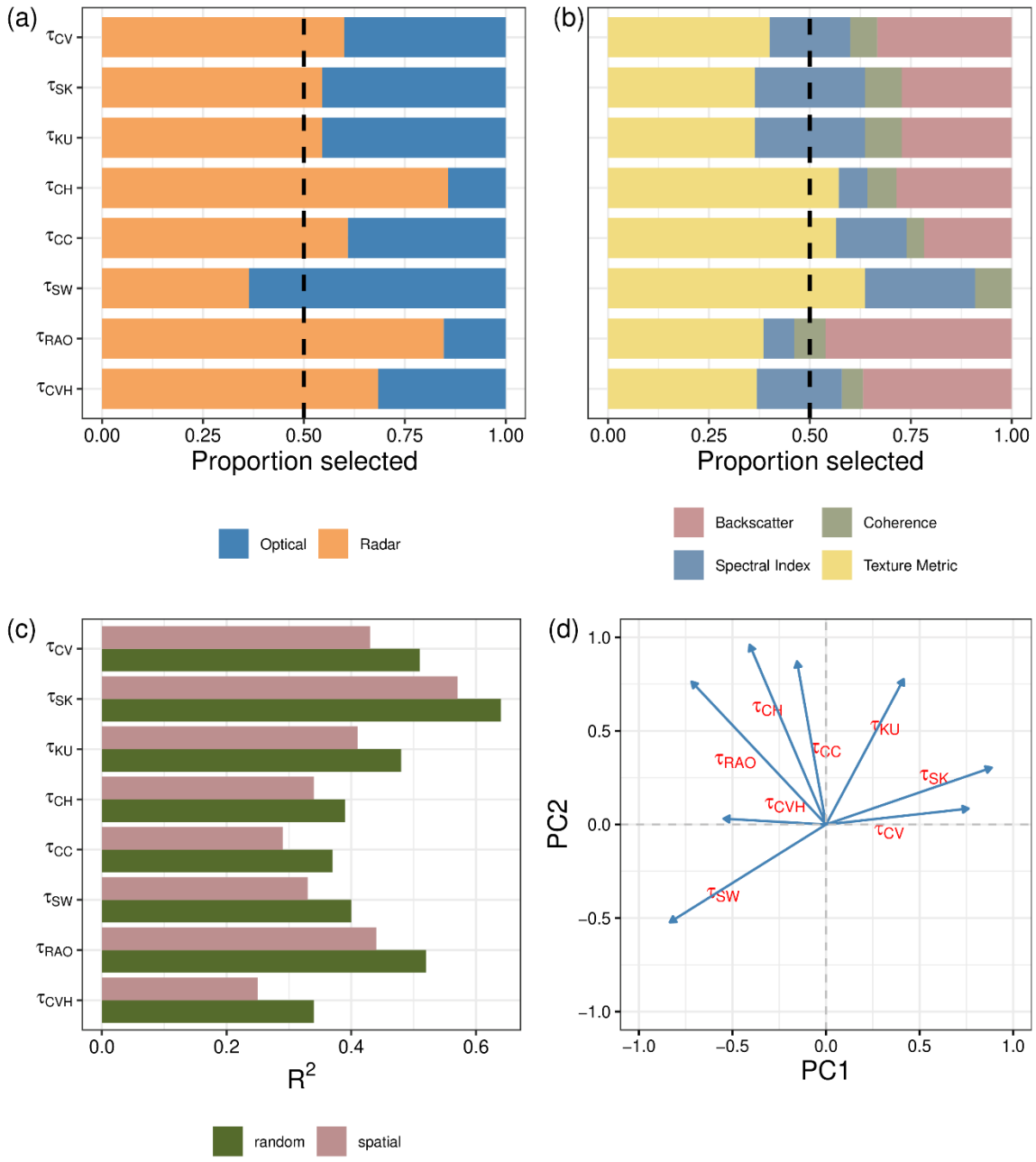

Figure S11. Results from the random forest modelling exercise at 5 km resolution. Panels display the variable selection frequencies (A and B) and model performance, as indicated by the R^2 values derived from two types of validation methods (C). Panel D shows the results of the Principal Component Analysis (PCA) conducted on the predicted metrics at this resolution.

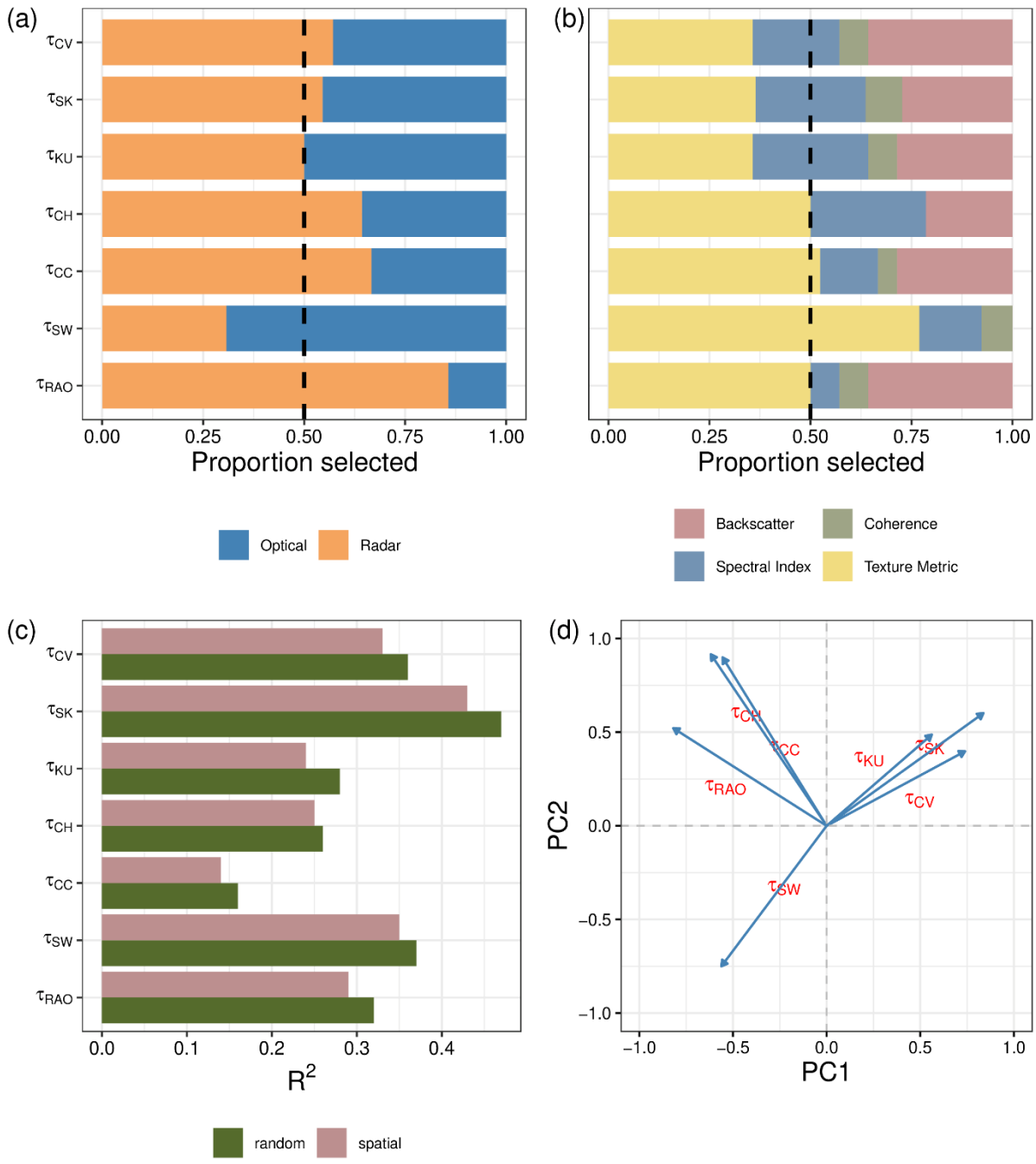

Figure S12 Results from the random forest modelling exercise at 1 km resolution. Panels display the variable selection frequencies (A and B) and model performance, as indicated by the R^2 values derived from two types of validation methods (C). Panel D shows the results of the Principal Component Analysis (PCA) conducted on the predicted metrics at this resolution.

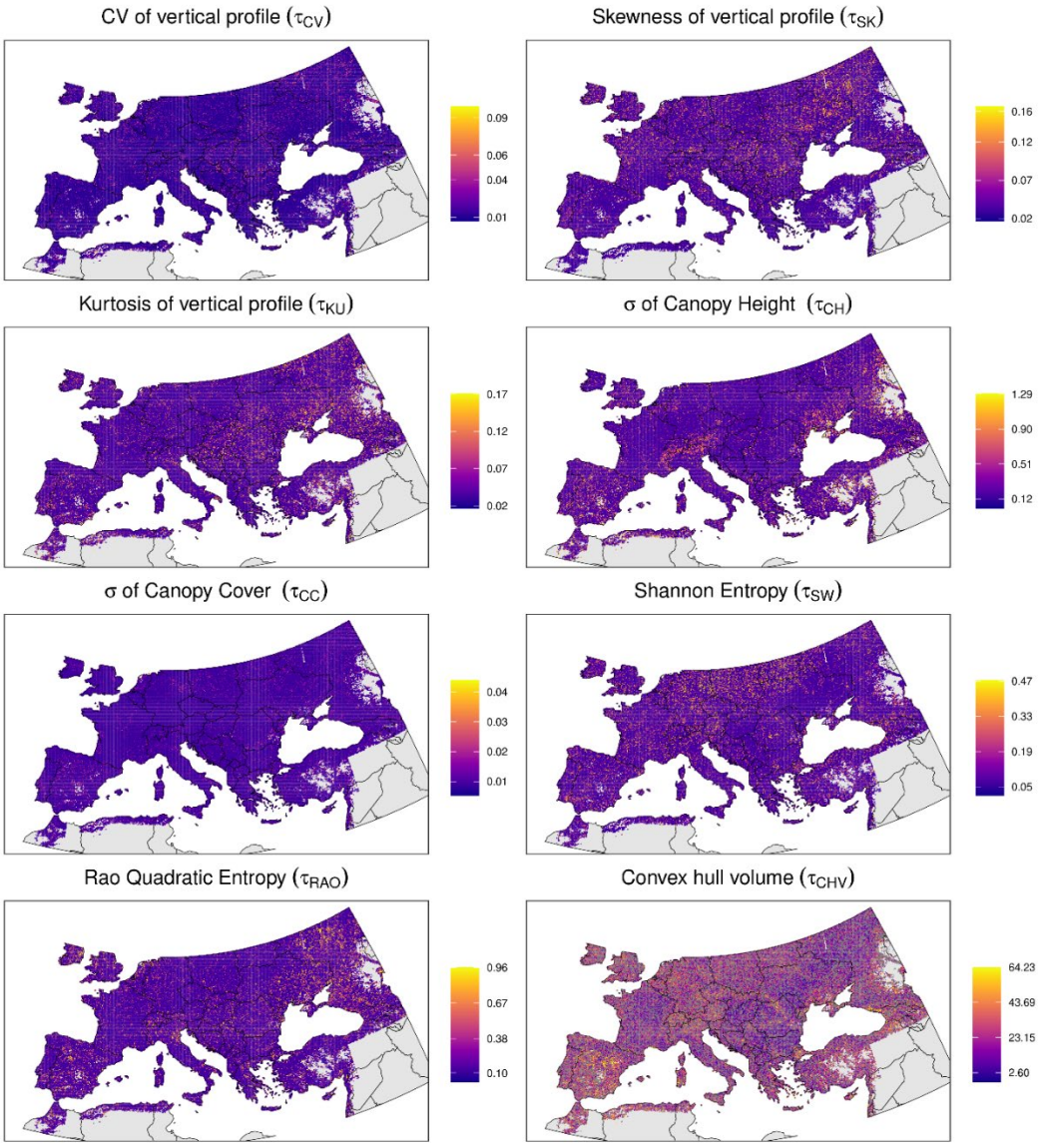

Figure S13. Standard deviation of the predictions for models trained at 10 km resolution using. The colour palette transitions from purple to yellow, denoting increasing standard error, with warmer colours signifying higher values.

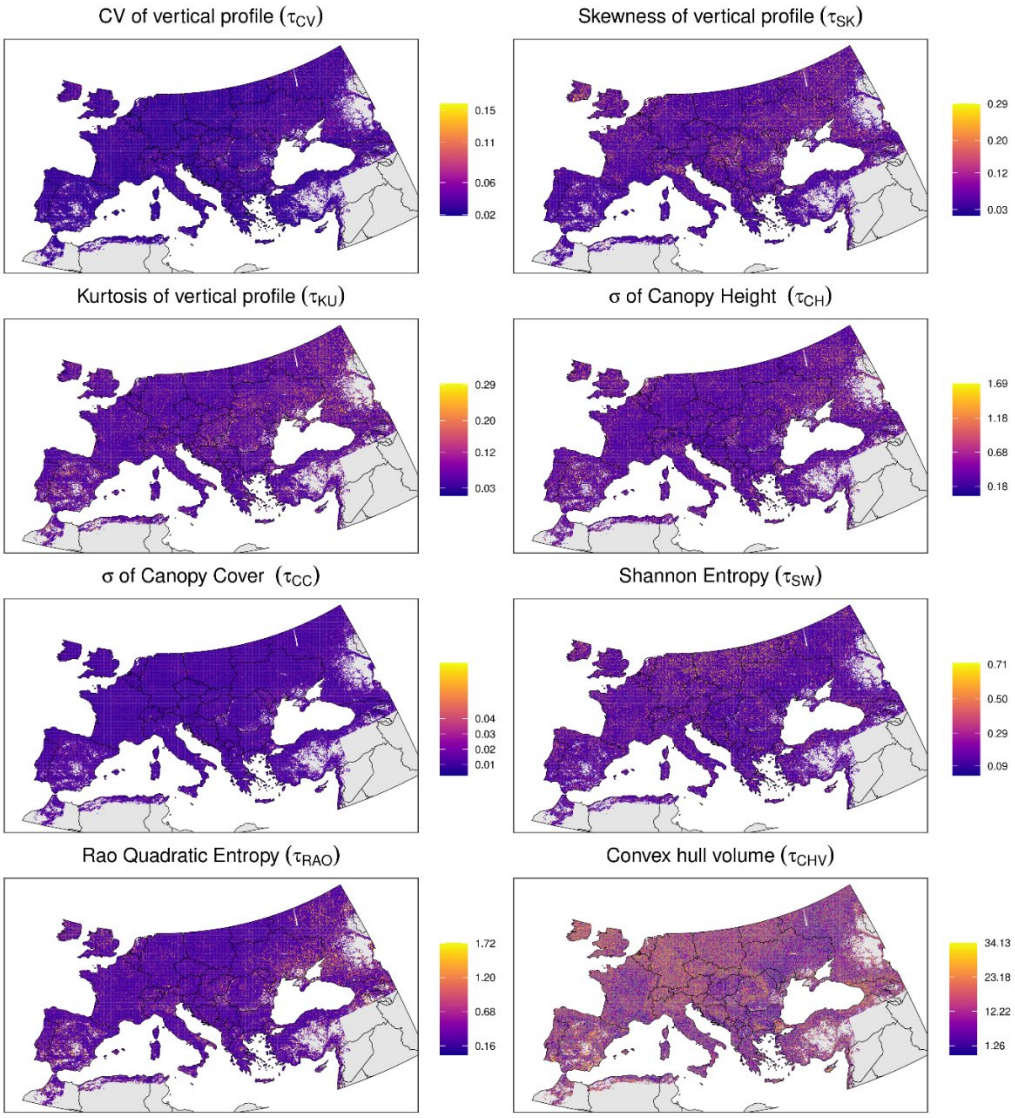

Figure S14 Standard deviation of the predictions for models trained at 5 km resolution using Random Forest modelling. The colour palette transitions from purple to yellow, denoting increasing standard error, with warmer colours signifying higher values.

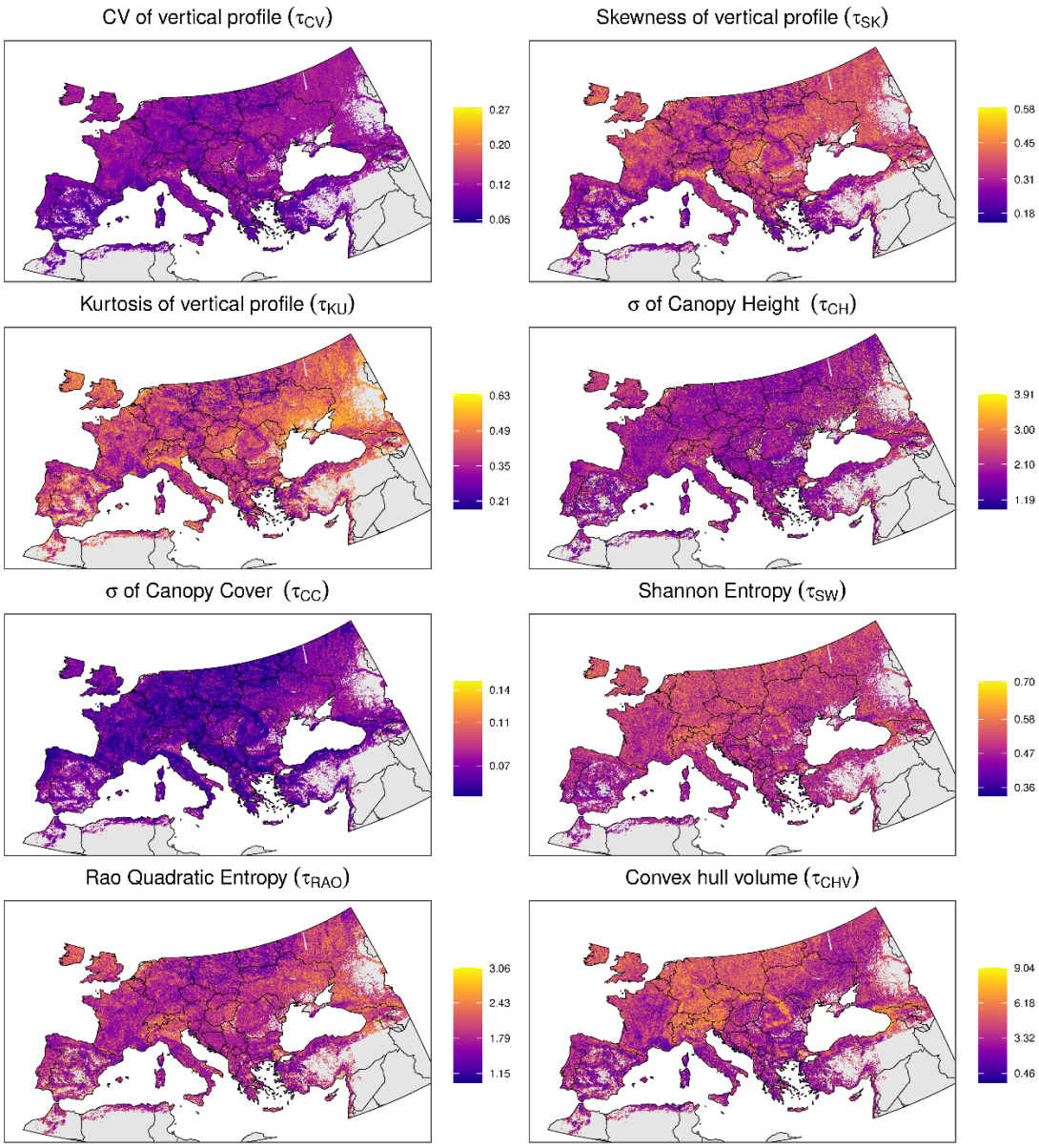

Figure S15 Standard deviation of the predictions for models trained at 1 km resolution using Random Forest modelling. The colour palette transitions from purple to yellow, denoting increasing standard error, with warmer colours signifying higher values.