

Supplementary Materials for

Prevailing European light regimes are mismatched with conditions for high plant diversity

Wanben Wu et al.

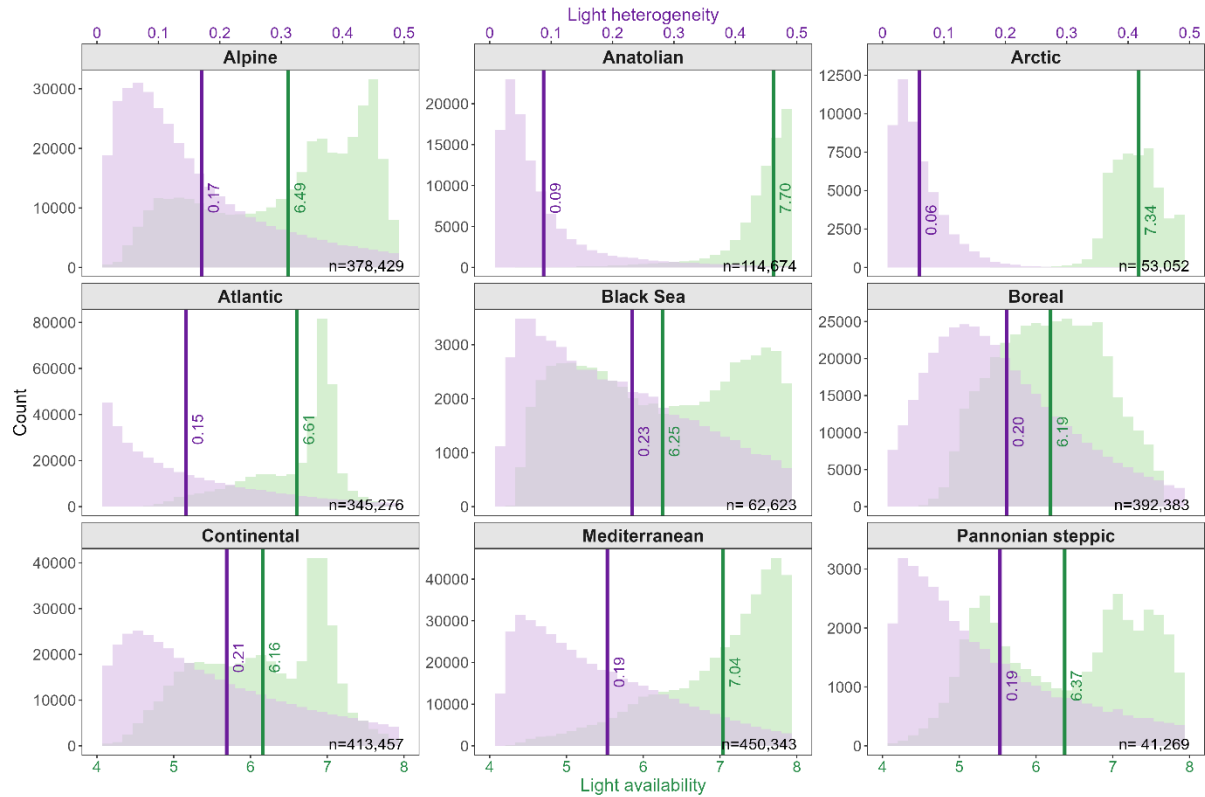
*Corresponding author. Email: wanben.wu@bio.au.dk

This PDF file includes:

Figs. S1 to S6

Tables S1 to S2

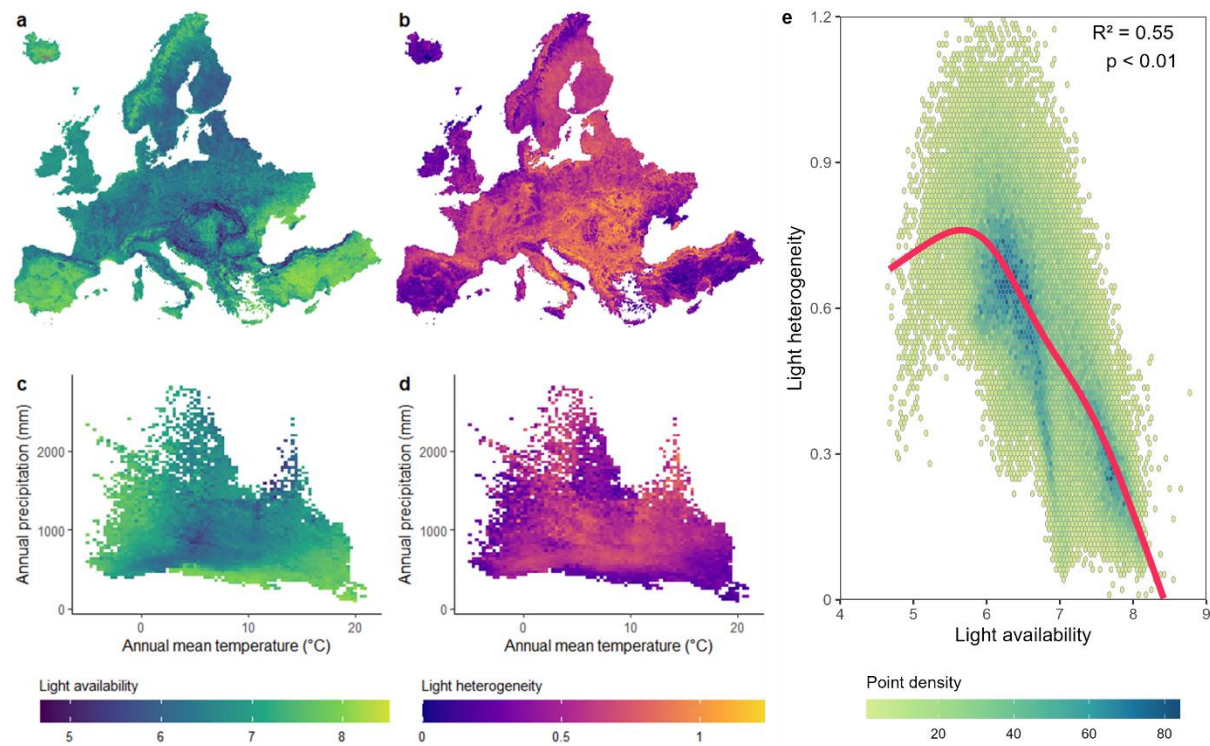
Fig. S1.



Light regime patterns across European biogeographic regions.

Distributions of light availability (green) and light heterogeneity (purple) across nine biogeographic regions, with vertical lines marking regional means. Light-regime metrics were derived from 10 m light-availability maps; heterogeneity represents the standard deviation within a 50 m radius moving window. The y-axis (Count) represents the number of 10 m pixels falling within each bin. Only natural and semi-natural areas were included. Sample sizes (n) denote the number of 10 m pixels randomly selected within each region based on the European biogeographic map⁷³.

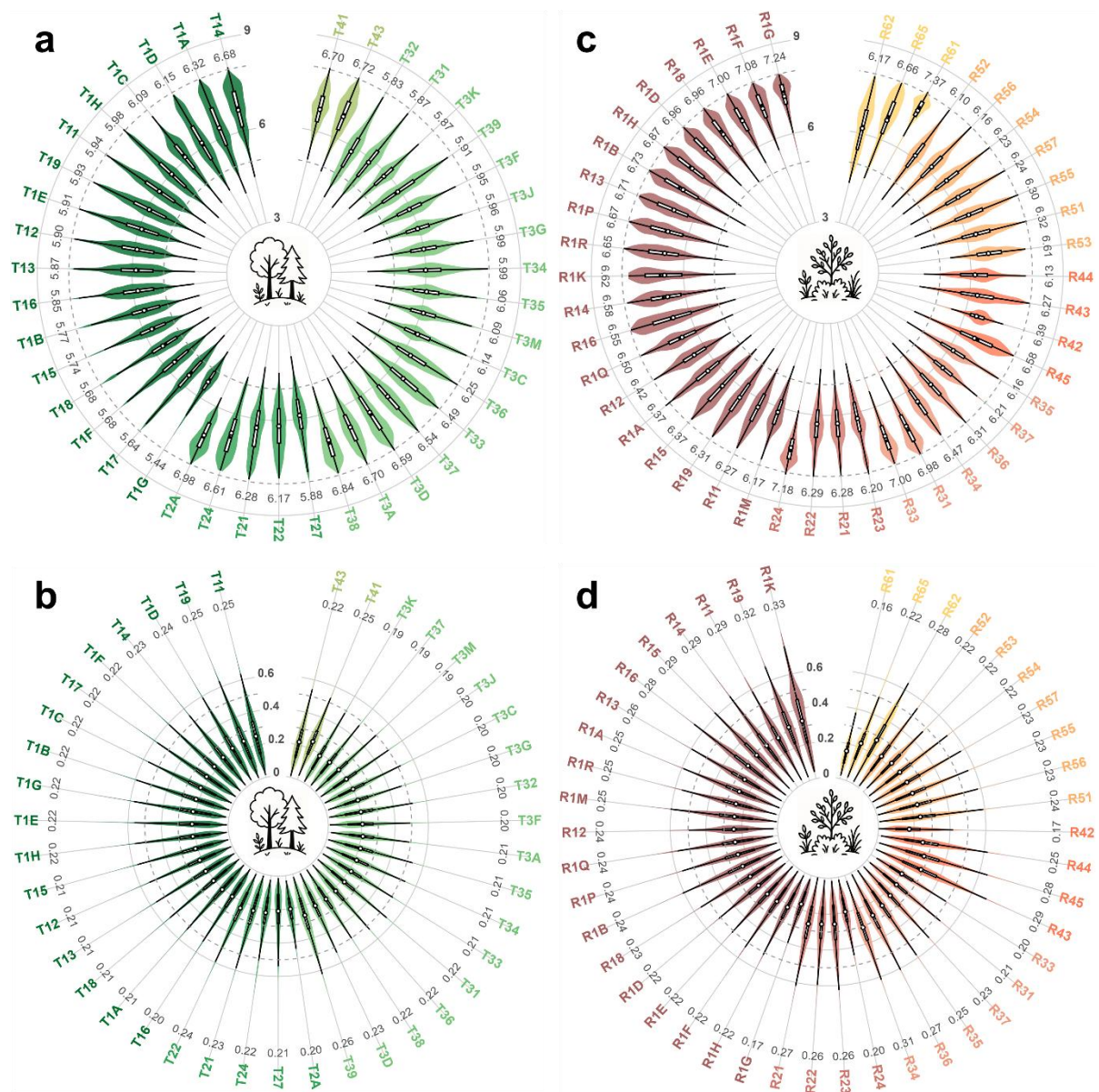
Fig. S2.



Spatial patterns and landscape-scale relationships of light regimes across Europe.

a, Spatial distribution of mean light availability aggregated to 10 km grid cells. **b**, Spatial distribution of light heterogeneity, computed as the standard deviation of 10 m light availability within each 10 km grid cell. **c–d**, Mean light availability (**c**) and heterogeneity (**d**) across climatological gradients of annual mean temperature and precipitation. Relationship between mean light availability and light heterogeneity across European landscapes. Hexagonal binning shows the density of observations in the two-dimensional light-regime space, with darker colours indicating higher point density. The solid line represents the fitted Generalized Additive Model (GAM) describing the smooth relationship between mean light availability and light heterogeneity. Shaded ribbons indicate the 95% confidence interval of the GAM smooth, reflecting uncertainty in the estimated mean relationship. Model fit statistics (R^2 and p value) are shown in the upper-right corner of the panel. Both metrics were derived from the 10 m light-availability map after masking non-natural land-cover types using the ESA WorldCover 2020 dataset²⁶ (excluding cropland [40], built-up areas [50], bare/sparse vegetation [60], snow and ice [70], and permanent water bodies [80]).

Fig. S3.

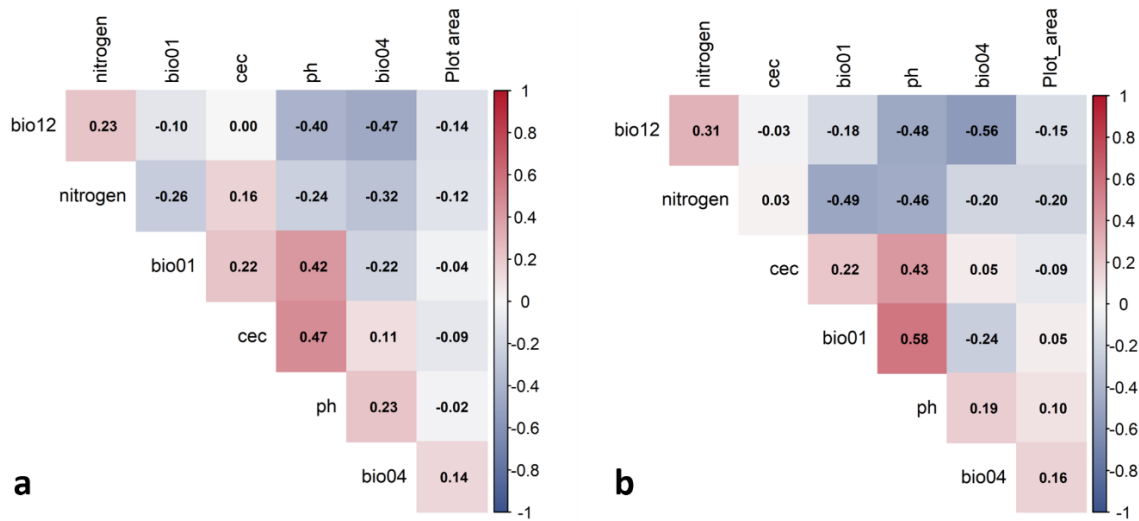


Light regime patterns across forest (T) and grassland (R) habitat types (EUNIS level 3). Panels a–b show, respectively, light availability and light heterogeneity across forest habitat types, while c–d present the corresponding patterns for grassland habitats. White points indicate habitat-level averages. Light-regime variables were derived from 10 m resolution light-availability maps; heterogeneity was quantified as the standard deviation within a 50 m radius moving window. Analyses were restricted to natural and semi-natural land-cover classes. Random samples were drawn within each habitat type using the European EUNIS habitat map. Forest habitats (T) include **T1 – Broadleaved deciduous forests**, comprising T11 – Temperate *Salix* and *Populus* riparian forest, T12 – *Alnus glutinosa*–*Alnus incana* forest on riparian and mineral soils, T13 – Temperate hardwood riparian forest, T14 – Mediterranean and Macaronesian riparian forest, T15 – Broadleaved swamp forest on non-acid peat, T16 – Broadleaved mire forest on acid peat, T17 – *Fagus* forest on

non-acid soils, T18 – *Fagus* forest on acid soils, T19 – Temperate and submediterranean thermophilous deciduous forest, T1A – Mediterranean thermophilous deciduous forest, T1B – Acidophilous *Quercus* forest, T1C – Temperate and boreal mountain *Betula* and *Populus tremula* forest on mineral soils, T1D – Southern European mountain *Betula* and *Populus tremula* forest on mineral soils, T1E – *Carpinus* and *Quercus* mesic deciduous forest, T1F – Ravine forest, T1G – *Alnus cordata* forest, and T1H – Broadleaved deciduous plantation of non site-native trees. Forest habitats further include **T2 – Broadleaved evergreen forests**, consisting of T21 – Mediterranean evergreen *Quercus* forest, T22 – Mainland laurophyllous forest, T24 – *Olea europaea*–*Ceratonia siliqua* forest, T27 – *Ilex aquifolium* forest, and T2A – Broadleaved evergreen plantation of site-native trees; **T3 – Coniferous forests**, including T31 – Temperate mountain *Picea* forest, T32 – Temperate mountain *Abies* forest, T33 – Mediterranean mountain *Abies* forest, T34 – Temperate subalpine *Larix*, *Pinus cembra* and *Pinus uncinata* forest, T35 – Temperate continental *Pinus sylvestris* forest, T36 – Temperate and submediterranean montane *Pinus sylvestris*–*Pinus nigra* forest, T37 – Mediterranean montane *Pinus sylvestris*–*Pinus nigra* forest, T38 – Mediterranean montane *Cedrus* forest, T39 – Mediterranean and Balkan subalpine *Pinus heldreichii*–*Pinus peuce* forest, T3A – Mediterranean lowland to submontane *Pinus* forest, T3C – *Taxus baccata* forest, T3D – Mediterranean *Cupressaceae* forest, T3F – Dark taiga, T3G – *Pinus sylvestris* light taiga, T3J – *Pinus* and *Larix* mire forest, and T3K – *Picea* mire forest; and **T4 – Lines of trees, small anthropogenic forests, recently felled forest, early-stage forest and coppice**, including T41 – Early-stage natural and semi-natural forest and regrowth and T43 – Recently felled area. Grassland and open habitats (R) include **R1 – Dry grasslands**, comprising R11 – Pannonian and Pontic sandy steppe, R12 – Cryptogam- and annual-dominated vegetation on siliceous rock outcrops, R13 – Cryptogam- and annual-dominated vegetation on calcareous and ultramafic rock outcrops, R14 – Perennial rocky grassland of the Italian Peninsula, R15 – Continental dry rocky steppic grassland and dwarf scrub on chalk outcrops, R16 – Perennial rocky grassland of Central and South-Eastern Europe, R18 – Perennial rocky calcareous grassland of subatlantic-submediterranean Europe, R19 – Dry steppic submediterranean pasture of the Amphi-Adriatic region, R1A – Semi-dry perennial calcareous grassland (meadow steppe), R1B – Continental dry grassland (true steppe), R1D – Mediterranean closely grazed dry grassland, R1E – Mediterranean tall perennial dry grassland, R1F – Mediterranean annual-rich dry grassland, R1G – Iberian oromediterranean siliceous dry grassland, R1H – Iberian oromediterranean basiphilous dry grassland, R1K – Balkan and Anatolian oromediterranean dry grassland, R1M – Lowland to montane, dry to mesic grassland usually dominated by *Nardus stricta*, R1P – Oceanic to subcontinental inland sand grassland on dry acid and neutral soils, R1Q – Inland sand drift and dune with siliceous grassland, and R1R – Mediterranean to Atlantic open, dry, acid and neutral grassland. Additional open habitats include **R2 – Mesic grasslands**, consisting of R21 – Mesic permanent pasture of lowlands and mountains, R22 – Low and medium altitude hay meadow, R23 – Mountain hay meadow, and R24 – Iberian summer pasture (vallicar); **R3 – Seasonally wet and wet grasslands**, including R31 – Mediterranean tall humid inland grassland, R33 – Mediterranean short moist grassland of mountains, R34 – Submediterranean moist meadow, R35 – Moist or wet mesotrophic to eutrophic hay meadow, R36 – Moist or wet mesotrophic to eutrophic pasture, and R37 – Temperate and boreal moist or wet oligotrophic grassland; **R4 – Alpine and subalpine grasslands**, including R42 – Boreal and Arctic acidophilous alpine grassland, R43 – Temperate acidophilous alpine grassland, R44 – Arctic-alpine calcareous grassland, and R45 –

Alpine and subalpine calcareous grassland of the Balkans and Apennines; **R5 – Woodland fringes and clearings and tall forb stands**, including R51 – Thermophilous forest fringe of base-rich soils, R52 – Forest fringe of acidic nutrient-poor soils, R53 – Macaronesian thermophilous forest fringe, R54 – *Pteridium aquilinum* vegetation, R55 – Lowland moist or wet tall-herb and fern fringe, R56 – Montane to subalpine moist or wet tall-herb and fern fringe, and R57 – Herbaceous forest clearing vegetation; and **R6 – Inland salt steppes**, including R61 – Mediterranean inland salt steppe, R62 – Continental inland salt steppe, and R65 – Continental subsaline alluvial pasture and meadow.

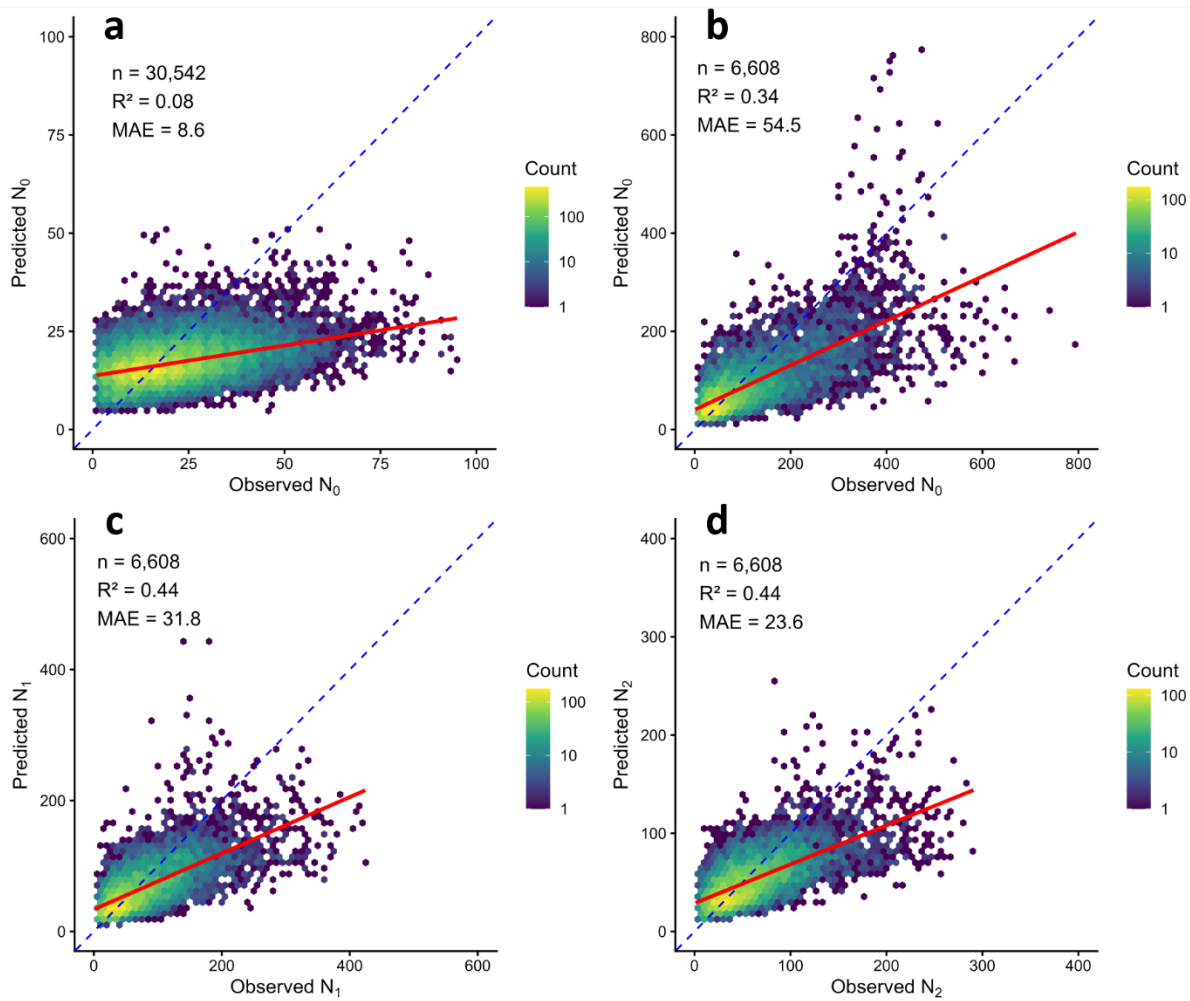
Fig. S4.



Correlation matrix of predictor variables included in the plant diversity models.

Pearson correlation coefficients among environmental predictors used in the final plant diversity models at the local scale (a) and landscape scale (b). Numbers within cells indicate pairwise correlation coefficients (r), with red and blue colours representing positive and negative correlations, respectively. Overall, correlations among predictors were moderate ($|r| < 0.6$), indicating low risk of severe multicollinearity. Variable abbreviations are as follows: bio01, annual mean temperature; bio04, temperature seasonality; bio12, annual precipitation; CEC, soil cation exchange capacity; nitrogen, soil nitrogen content; pH, soil pH; and Plot area, vegetation plot area.

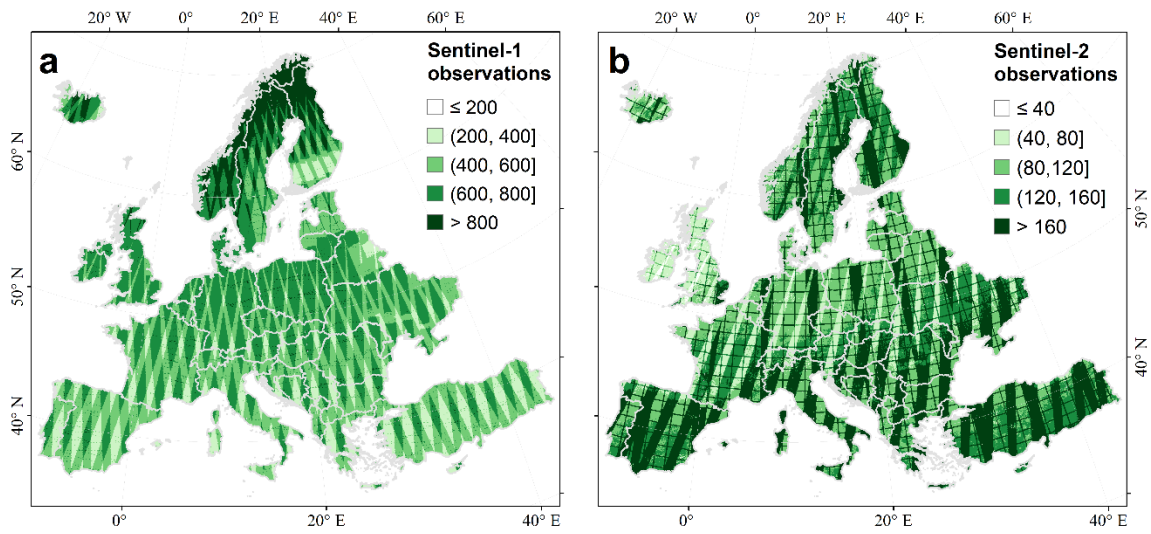
Fig. S5.



Cross-validated observed versus predicted plant diversity across spatial scales.

Hexbin plots show the relationship between observed and predicted diversity values from five-fold spatial cross-validation. Colours indicate the number of observations per hexagonal bin on a logarithmic scale. The blue dashed line represents the 1:1 relationship, and the red solid line shows the fitted linear regression. Panel **a** shows local-scale species richness (N_0), whereas panels **b–d** show landscape-scale diversity metrics, including species richness (N_0 ; **b**), the exponential of Shannon entropy (N_1 ; **c**), and the inverse Simpson index (N_2 ; **d**). Model performance statistics (sample size, coefficient of determination, and mean absolute error) are reported in each panel.

Fig. S6.



Spatial coverage of satellite data used. a, Number of Sentinel-1 observations acquired between 2019 and 2021 and used as inputs in the light-availability modelling framework. **b,** Number of clear-sky Sentinel-2 observations from 2019 to 2021 contributing spectral information for light-availability estimation.

Table S1.**Earth observation predictors used to reconstruct continental-scale light availability across Europe.**

Predictor variables included radar backscatter metrics from Sentinel-1 and PALSAR, spectral bands and vegetation indices derived from Sentinel-2 imagery, topographic variables, and geographic coordinates. Sentinel-1 and Sentinel-2 metrics were derived from observations acquired between 2019 and 2021. Abbreviations: VV and VH, vertical transmit–vertical receive and vertical transmit–horizontal receive polarization; HH and HV, horizontal transmit–horizontal receive and horizontal transmit–vertical receive polarization; NDVI, Normalized Difference Vegetation Index; EVI, Enhanced Vegetation Index; kNDVI, kernel Normalized Difference Vegetation Index; LSWI, Land Surface Water Index; MNDWI, Modified Normalized Difference Water Index; GCVI, Green Chlorophyll Vegetation Index; NMDI, Normalized Multi-band Drought Index; CSI, Combinational Shadow Index; BI, Brightness Index; BSI, Bare Soil Index; SATVI, Soil-Adjusted Total Vegetation Index; TPI, Topographic Position Index. Resolution refers to the native spatial resolution of each dataset.

Code	Products	Variables	Acquisition time	Resolution	Data Source
1-3	Sentinel-1	VV; VH; VVH	2019-2021	10 m	COPERNICUS/S1_G RD
4-5	PALSAR	HH; HV	2019-2021	25 m	JAXA/ALOS/PALSAR/ YEARLY/SAR
6-27	Sentinel-2	Aerosols; Blue; Green; Red; NIR; Red Edge 1–4; SWIR 1; SWIR 2; NDVI; EVI; kNDVI; LSWI; MNDWI; GCVI; NMDI; CSI; BI; BSI; SATVI	2019-2021	10/20/60 m	COPERNICUS/S2_S R_HARMONIZED
28-33	Topography	Elevation; Slope; Aspect; TPI _{150m} ; TPI _{300m} ; TPI _{1500m}	/	30 m	projects/sat-io/open- datasets/GLO-30
33-35	Location	Latitude; Longitude	/	10 m	/

Table S2.**Moran's I statistics for residual spatial autocorrelation in the final interaction models.**

Residual spatial autocorrelation was evaluated using Moran's I calculated from Pearson residuals of the final interaction models for local-scale species richness (N_0) and landscape-scale diversity metrics (N_0 , N_1 and N_2). Spatial neighbourhoods were defined using an 8-nearest-neighbour graph with row-standardized spatial weights. Moran's I values close to zero indicate that most spatial dependence was accounted for by the fitted spatial random field.

Scale	Diversity metric	Moran's I	P-value
Local	N_0	-0.04	1
	N_0	0.009	0.052
Landscape	N_1	0.0079	0.084
	N_2	0.0095	0.0504