Ellenberg-type indicator values for European vascular plant species

Lubomír Tichý, Irena Axmanová, Jürgen Dengler, Riccardo Guarino, Florian Jansen, Gabriele Midolo, Michael Nobis, Koenraad van Meerbeek, Svetlana Aćić, Fabio Attorre, et al.

Ellenberg-type indicator values for European vascular plant species

Lubomír Tichý[1] (0000-0001-8400-7741), Irena Axmanová[1] (0000-0001-9440-7976), Jürgen Dengler[2,3,4] (0000-0003-3221-660X), Riccardo Guarino[5] (0000-0003-0106-9416), Florian Jansen[6] (0000-0002-0331-5185), Gabriele Midolo[1] (0000-0003-1316-2546), Michael P. Nobis[7] (0000-0003-3285-1590), Koenraad Van Meerbeek[8,9] (0000-0002-9260-3815), Svetlana Aćić[10] (0000-0001-6553-3797), Fabio Attorre[11] (0000-0002-7744-2195), Erwin Bergmeier[12] (0000-0002-6118-4611), Idoia Biurrun[13] (0000-0002-1454-0433), Gianmaria Bonari[14] (0000-0002-5574-6067), Helge Bruelheide[15,4] (0000-0003-3135-0356), Juan Antonio Campos[13] (0000-0003-4770-0461), Andraž Čarni[16,17] (0000-0002-8909-4298), Alessandro Chiarucci (0000-0003-1160-235X)[18], Mirjana Ćuk[19] (0000-0002-8261-414X), Renata Ćušterevska[20] (0000-0002-3849-6983), Yakiv Didukh[21] (0000-0002-5661-3944), Daniel Dítě[22] (0000-0001-5251-9910), Zuzana Dítě[22] (0000-0002-2895-9024), Tetiana Dziuba[21] (0000-0001-8621-0890), Giuliano Fanelli[11] (0000-0002-3143-1212), Eduardo Fernández-Pascual[28] (0000-0002-4743-9577), Emmanuel Garbolino[23] (0000-0002-4954-6069), Rosario G. Gavilán (0000-0002-1022-445x)[24], Jean-Claude Gégout[25] (0000-0002-5760-9920), Ulrich Graf[7], Behlül Güler[26] (0000-0003-2638-4340), Michal Hájek[1] (0000-0002-5201-2682), Stephan M. Hennekens[27] (0000-0003-1221-0323), Ute Jandt[15,4] (0000-0002-3177-3669), Anni Jašková[1] (0000-0002-3510-1093), Borja Jiménez-Alfaro (0000-0001-6601-9597)[28], Philippe Julve[29], Stephan Kambach[15] (0000-0003-3585-5837), Dirk Nikolaus Karger[7] (0000-0001-7770-6229), Gerhard Karrer[30] (0000-0001-5172-2319), Ali Kavgacı[31] (0000-0002-4549-3668), Ilona Knollová[1] (0000-0003-4074-789X), Anna Kuzemko[1,21] (0000-0002-9425-2756), Filip Küzmič[16] (0000-0002-3894-7115), Flavia Landucci[1] (0000-0002-6848-0384), Attila Lengyel[32] (0000-0002-1712-6748), Jonathan Lenoir[33] (0000-0003-0638-9582), Corrado Marcenò[34] (0000-0003-4361-5200), Jesper Erenskjold Moeslund[35] (0000-0001-8591-7149), Pavel Novák[1] (0000-0002-3758-5757), Aaron Pérez-Haase[36] (0000-0002-5974-7374), Tomáš Peterka[1] (0000-0001-5488-8365), Remigiusz Pielech[37,38] (0000-0001-8879-3305), Alessandro Pignatti[11], Valerijus

Rašomavičius[39] (0000-0003-1314-4356), Solvita Rūsiņa[40] (0000-0002-9580-4110), Arne Saatkamp[41,42] (0000-0001-5638-0143), Urban Šilc[16] (0000-0002-3052-699X), Željko Škvorc[43] (0000-0002-2848-1454), Jean-Paul Theurillat[44,45] (0000-0002-1843-5809), Thomas Wohlgemuth[7] (0000-0002-4623-0894) & Milan Chytrý[1] (0000-0002-8122-3075)

[1]*Department of Botany and Zoology, Faculty of Science, Masaryk University, Brno, Czech Republic*

[2]*Vegetation Ecology Research Group, Institute of Natural Resource Sciences (IUNR), Zurich University of Applied Sciences (ZHAW), Wädenswil, Switzerland*

[3]*Plant Ecology, Bayreuth Center of Ecology and Environmental Research (BayCEER), University of Bayreuth, Bayreuth, Germany*

[4]*German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig, Leipzig, Germany*

[5]*Department of Biological Chemical and Pharmaceutical Sciences and Technologies (STEBICEF), University of Palermo, Palermo, Italy*

[6]*Faculty of Agricultural and Environmental Sciences, University of Rostock, Rostock, Germany*

[7]*Swiss Federal Research Institute WSL, Birmensdorf, Switzerland*

[8]*Department of Earth and Environmental Sciences, KU Leuven, Leuven, Belgium*

[9]*KU Leuven Plant Institute, KU Leuven, Leuven, Belgium*

[10]*Department of Botany, Faculty of Agriculture, University of Belgrade, Beograd, Serbia*

[11]*Department of Environmental Biology, Sapienza University of Rome, Roma, Italy*

[12]*Vegetation Ecology & Plant Diversity, Albrecht von Haller Institute of Plant Sciences, University of Göttingen, Göttingen, Germany*

[13]*Department of Plant Biology and Ecology, University of the Basque Country UPV/EHU, Bilbao, Spain*

[14]*Free University of Bozen-Bolzano, Bolzano, Italy*

[15]*Institute of Biology/Geobotany and Botanical Garden, Martin Luther University Halle-Wittenberg, Halle (Saale), Germany*

[16]*Research Centre of the Slovenian Academy of Sciences and Arts, Jovan Hadži Institute of Biology, Ljubljana, Slovenia*

[17] *School for Viticulture and Enology, University of Nova Gorica, Nova Gorica, Slovenia*

[18]*BIOME Lab, Department of Biological, Geological & Environmental Sciences, Alma Mater Studiorum - University of Bologna, Bologna, Italy*

[19]*Department of Biology and Ecology, Faculty of Science, University of Novi Sad, Novi Sad, Serbia*

[20]*Faculty of Natural Sciences and Mathematics, Ss. Cyril and Methodius University, Skopje, North Macedonia*

[21]*M.G. Kholodny Institute of Botany, National Academy of Sciences of Ukraine, Kyiv, Ukraine*

[22]*Plant Science and Biodiversity Center, Slovak Academy of Sciences, Bratislava, Slovakia*

[23]*Climpact Data Science, Nova Sophia - Regus Nova, Sophia Antipolis Cedex, France*

[24]*Botany Unit, Department of Pharmacology, Pharmacognosy and Botany, Complutense University, Madrid, Spain*

[25]*Université de Lorraine, AgroParisTech, INRAE, UMR Silva, Nancy, France*

[26]*Biology Education, Dokuz Eylül University, Buca, Izmir, Turkey*

[27]*Wageningen Environmental Research, Wageningen, the Netherlands*

[28]*IMIB Biodiversity Research Institute, University of Oviedo, Mieres, Spain*

[29]*Faculté de Gestion, Economie et Sciences, Lille Catholic University, Lille, France*

[30]*Department of Integrative Biology and Biodiversity Research, University of Natural Resources and Life Sciences Vienna, Vienna, Austria*

[31]*Burdur Food Agriculture and Livestock Vocational School, Burdur Mehmet Akif Ersoy University, Burdur, Türkiye*

[32]*Centre for Ecological Research, Institute of Ecology and Botany, Vácrátót, Hungary*

[33]*UMR CNRS 7058 "Ecologie et Dynamique des Systèmes Anthropisés" (EDYSAN), Université de Picardie Jules Verne, Amiens, France*

[34]*Department of Chemistry, Biology and Biotechnology, University of Perugia, Perugia, Italy*

[35]*Department of Ecoscience, Section for Biodiversity, Aarhus University, Aarhus, Denmark*

[36]*Department of Evolutionary Biology, Ecology and Environmental Sciences, University of Barcelona, Barcelona, Spain*

[37]*Department of Forest Biodiversity, Faculty of Forestry, University of Agriculture in Kraków, Kraków, Poland*

[38]*Foundation for Biodiversity Research, Wrocław, Poland*

[39]*Institute of Botany, Nature Research Centre, Vilnius, Lithuania*

[40]*Faculty of Geography and Earth Sciences, University of Latvia, Riga, Latvia*

[41]*Conservatoire Botanique National Méditerranéen, Hyères, France*

[42]*Aix Marseille Université, Université Avignon, CNRS, IRD, UMR IMBE, Marseille, France*

[43]*University of Zagreb, Faculty of Forestry and Wood Technology, Zagreb, Croatia*

[44]*Fondation J.-M. Aubert, Champex-Lac, Switzerland*

[45]*Department of Plant Sciences, University of Geneva, Chambésy, Switzerland*

Correspondence

Lubomír Tichý, Department of Botany and Zoology, Faculty of Science, Masaryk University, Kotlářská 2, 611 37 Brno, Czech Republic. E-mail: tichy@sci.muni.cz

Funding information

This research was funded through the 2019-2020 BiodivERsA joint call for research proposals, under the BiodivClim ERA-Net COFUND program, and with the funding organizations Technology Agency of the Czech Republic (SS70010002), the Swiss National Science Foundation SNF (project: FeedBaCks, 193907), and the German Research Foundation (DFG BR 1698/21-1, DFG HI 1538/16-1). E.F.P. was supported by the Jardín Botánico Atlántico (SV-20-GIJON-JBA), A.Č., U.Š. and F.K. werefunded by Slovenian Research Agency (ARRS P1-0236), I.B. and J.A.C. were funded by the Basque Government (IT1487-22), and SR was funded by the LIFE Integrated Project LIFE19 IPE/LV/000010.

Running title

Ellenberg-type indicator values for European plants

Abstract

Aims: Ellenberg-type indicator values are expert-based rankings of plant species according to their ecological optima on main environmental gradients. Here we extend the indicator-value system proposed by Heinz Ellenberg and co-authors for Central Europe by incorporating compatible systems developed for other European regions and creating a harmonized dataset of indicator values applicable at the European scale.

Methods: We collected European datasets of indicator values for vascular plants and selected 13 datasets that used the nine-, ten- or twelve-step scales defined by Ellenberg for light, temperature, moisture, reaction, nutrients and salinity. We compared these values with the original Ellenberg values and used those that showed consistent trends in regression slope and coefficient of determination. We calculated the average value for each combination of species and indicator value from these datasets. based on species co-occurrences in European vegetation plots, we also calculated new values for species that were not assigned an indicator value.

Results: We provide a new dataset of Ellenberg-type indicator values for 8,908 European vascular plant species (8,168 for light, 7,400 for temperature, 8,030 for moisture, 7,282 for reaction, 7,193 for nutrients and 7,507 for salinity), of which 398 species have been newly assigned to at least one indicator value.

Conclusions: The newly introduced indicator values are compatible with the original Ellenberg values. They can be used for large-scale studies of the European flora and vegetation or for gap-filling in regional datasets. The

European values and the original and taxonomically harmonized regional datasets of Ellenberg-type indicator values are available in Supplementary Information and the Zenodo repository.

Introduction

Bioindication of abiotic site conditions from environmental relationships of plant species has a long tradition (Cajander, 1926; Iversen, 1936). Seminal work was done by the German vegetation ecologist Heinz Ellenberg, who published a comprehensive dataset of indicator values for plant species (Ellenberg, 1974). These values were based on field observations and partly also measurements, mainly from Germany. Ellenberg defined indicator values for seven abiotic environmental variables: light, temperature, continentality, moisture, soil reaction, nutrient (nitrogen) content, and salinity. While the first three variables relate mainly to above-ground conditions, the last four describe substrate conditions (soil or water). Ellenberg originally defined indicator values for nitrogen content, but later studies suggested that they rather reflect general soil fertility, such as the combined availability of both nitrogen and phosphorus (Boller-Elmer, 1977; Briemle, 1986; Hill & Carey, 1997). Therefore, Ellenberg's original nitrogen values are nowadays more often called nutrient values (Ellenberg et al. 1992).

Ellenberg indicator values were defined on ordinal scales that characterize the relative position of the centroid of a species' realized one-dimensional niche related to the respective environmental variable. A low value corresponds to the position of the species optimum towards the lower end of the environmental gradient and, respectively, towards the higher end of the gradient for a high value. For example, low values of the light value are assigned to shade-tolerant species, whereas high values are assigned to species that occur in full light.

Ellenberg's system was inspired in part by the ideas of Cajander (1926), who used associations of plant species to evaluate forest types and productivity, and Iversen (1936), who arranged plants into response groups to environmental variables relevant to plant growth. However, Ellenberg (1948, 1950, 1952) was the first to use numerical codes instead of verbally defined levels of environmental gradients. Ellenberg (1948) also proposed using these codes to calculate community means based on simple species presence, or weighted values based on abundance (i.e., percentage cover in the plot). Subsequently, other authors (e.g., Zólyomi et al., 1967; Zlatník et al., 1970) adopted Ellenberg's concept of bioindication by creating regional systems of indicator values for other parts of Europe. Not only vascular plants but later also bryophytes and lichens were characterized by indicator values following the same system (Ellenberg et al., 1992).

Repeatedly updated and refined, Ellenberg indicator values (Ellenberg et al., 1992, 2001; Ellenberg & Leuschner, 2010) are a widely used tool for rapid estimation of environmental conditions without direct measurements (Diekmann, 2003; Holtland et al., 2010). In the Web of Science database, 907 articles with the keywords (including words used in abstracts) 'Ellenberg' AND 'Indicator' were registered between 1 January 1974 and 30 June 2022, indicating their importance to plant ecologists. Several studies found a good agreement between community means (weighted or non-weighted) calculated from Ellenberg indicator values and values of environmental variables measured *in situ* (Ellenberg et al., 1992;

Herzberger & Karrer, 1992; Hill & Carey, 1997; Ertsen et al., 1998; Schaffers & Sýkora, 2000; Wamelink et al., 2002; Diekmann, 2003; Chytrý et al., 2009; Sicuriello et al., 2014). Some authors also discussed the consistency of indicator values between different geographical areas (Diekmann & Lawesson, 1999; Gégout & Krizova, 2003; Godefroid & Dana, 2007; Wasof et al., 2013). Because Ellenberg's original dataset focused on plants occurring in the western part of Central Europe, other authors proposed indicator values for other European regions. These datasets included many species that were missing from Ellenberg's original dataset and often contained different values for the same species, reflecting shifted optima of their realized niches between regions (e.g. Landolt, 1977; Tsyganov, 1983; Jurko, 1990; Karrer, 1992; Borhidi, 1995; Mayor López, 1996; Böhling et al., 2002; Zarzycki et al., 2002; Hill et al., 2004; Pignatti, 2005; Landolt et al., 2010; Didukh, 2011; Chytrý et al., 2018; Domina et al., 2018; Guarino & La Rosa, 2019; Jiménez-Alfaro et al., 2021). Specialized datasets of indicator values for species limited to a specific habitat type but covering large areas were also created (e.g. Hájek et al., 2020 – mires; Dítě et al., 2022 – saline habitats).

The increasing number of synthetic and macroecological studies on European vegetation, catalyzed by the launch of the European database of vegetation plots (European Vegetation Archive, EVA; Chytrý et al., 2016), require a coherent system of species-level indicator values. Although regional systems of indicator values have been widely used for a long time, no consensual system of indicator values for European plants has been developed so far. Therefore, we have compiled a harmonized dataset of vascular plant indicator values for light, temperature, moisture, soil (or water) reaction (related to base saturation), nutrients (site productivity), and salinity suitable for a large part of Europe, using the same numerical scales as defined by Ellenberg. In this article, we describe the content of the new dataset and the methods used to compile it.

Methods

We compiled a database of thirteen published European datasets of indicator values for vascular plant species defined on the same nine-degree scale (or ten-degree scale for salinity and twelve-degree scale for moisture) as the original Ellenberg indicator values (Ellenberg et al., 1992, 2001). We refer to these datasets as *Ellenberg-type* indicator values. Datasets with scales containing a lower number of degrees, i.e., with a coarser resolution, were not included. If the scale had a higher number of degrees than nine (or ten for salinity or twelve for moisture), we accepted it, provided that: (1) the additional degrees represented an extension of the environmental gradient, while the other degrees retained the same meaning as in the original Ellenberg dataset (e.g. extending the nine-degree temperature scale originally defined for Central Europe to twelve degrees to reflect Mediterranean conditions; Pignatti, 2005) or (2) the additional degrees represented intermediate values on the nine- or twelve-degree scale (e.g. the 17-degree temperature scale and the 23-degree moisture scale in Didukh, 2011). We considered only datasets based entirely or largely on expert knowledge, and excluded those based on values re-calculated from vegetation plots without expert-based assessment of values for individual species (e.g. Lawesson et al., 2003 for the Faroe Islands).

The thirteen indicator-value datasets that met the above conditions included: Great Britain (Hill et al., 2000); the Cantabrian Mountains in Spain (Jiménez-Alfaro et al., 2021); France (Julve, 2015); Switzerland and the Alps (Landolt et al., 2010; temperature values only, as the other values use coarser scales than Ellenberg); Germany (Ellenberg et al., 2001, taken from Ellenberg & Leuschner, 2010); Czech Republic (Chytrý et al., 2018); Austria (Karrer, 1992); Hungary (Borhidi, 1995); Ukraine (Didukh, 2011; only the light, temperature and moisture values, as the others cannot be matched to the Ellenberg scales); Italy (Guarino & La Rosa, 2019, a corrected version prepared by R. Guarino for this study); South Aegean region of Greece (Böhling et al., 2002); European mires (Hájek et al., 2020); and saline habitats in Central Europe (Dítě et al., 2022). The scales of these thirteen datasets had twelve degrees for moisture and some of them also for temperature, ten degrees for salinity and nine degrees for the other values.

Therefore, we integrated the datasets using twelve-degree scales for temperature and moisture, a ten-degree scale for salinity and nine-degree scales for light, reaction and nutrients.

We did not include the indicator values for continentality because they are based on species geographical ranges. Continentality values may have an ambiguous meaning at the local scale since they may correlate with different factors, including seasonal differences in temperature and precipitation, diurnal differences in temperature, annual minimum temperatures and drought. Moreover, Berg et al. (2017) identified methodological weaknesses in the original Ellenberg approach to continentality values, proposed an improved protocol for their compilation, and defined new formally-verified values.

We unified the taxonomy and nomenclature of all vascular plant taxa across the thirteen datasets according to the Euro+Med PlantBase (http://europlusmed.org). We merged subspecies, varieties and forms at the species level and removed hybrids and rare alien species (mostly casual neophytes; Richardson et al., 2000). We also merged as 'aggregates' those taxonomically related species that are difficult to identify and, therefore, are often misidentified or not identified at all, such as species of the *Achillea millefolium* group in the *Achillea millefolium* aggr. The aggregates used were those defined in the Euro+Med PlantBase (Euro+Med, 2021) and the EUNIS-ESy expert system for EUNIS Habitat Classification (Chytrý et al., 2020). For infraspecific taxa within the same species or species within the same aggregate, we used their arithmetic mean as the indicator value for the species or aggregate to equally weight the indicator values of species in all datasets where the species occurs. In addition, we also calculated the median, minimum, and maximum. Some databases provided indicator values for both individual species and aggregates of species. Although some of these aggregates are not regularly used in vegetation science and do not fit the concept of Euro+Med and EUNIS, we kept them on the list to avoid losing information.

The new system of indicator values was prepared by calculating the arithmetic mean for each combination of species and environmental variable across all compatible regional datasets in which an indicator value was defined for the target species. As a first step, we tested whether the indicator values of each of the twelve datasets (other than the original Ellenberg dataset) were compatible with the Ellenberg values. We conducted two comparisons. For the first one, we tested a direct pairwise relationship between the original Ellenberg values (independent variable) for individual species (Ellenberg & Leuschner, 2010) and values for the same species in a different dataset (dependent variable; species-based regression). For the second comparison, we used vegetation plots from the EVA database (Chytrý et al., 2016) to calculate the unweighted means of the original Ellenberg values (independent variable) and indicator values from the other 12 datasets (dependent variable; plot-based regression). A total of 1,790,582 vegetation plots covering a wide range of vegetation types from across Europe were available for this approach. The Russian Federation, Georgia, Armenia, and Azerbaijan were not included due to their peripheral biogeographical location, lack of indicator-value datasets compatible with Ellenberg scales, and low density of plots in the EVA database. Species nomenclature was unified in the same way as in the indicator-value databases (see above). We selected only vegetation plots that contained at least five species with indicator values, both from the original Ellenberg dataset and from other indicator-value datasets, resulting in 622,402 plots for light indicator values, 413,832 for temperature, 615,301 for moisture, 490,617 for reaction, 575,406 for nutrients and 673,141 for salinity.

Based on the regression analyses described above, we selected datasets that showed consistent trends in both the direct species-based and indirect plot-based regressions against the original Ellenberg indicator values. In order to compare these trends, we selected two regression characteristics: (a) the coefficient of determination (R^2), which shows the amount of variation in the dependent variable that is explained by the regression. However, the same R^2 can be obtained with vastly different slopes. Therefore, we also adopted (b) the second criterion of the slope, which mainly indicates differences at the ends (extremes) of the indicator value range. Based on the empirical assessment of the regression results, we selected only indicator values for which the regression slope was within the range from 0.5 to 1.2 and R^2 was higher than 0.5. The only exception was the salinity dataset for Central Europe (Dítě et al., 2022), which, in contrast to Ellenberg salinity values, did not include any non-halophytic species.

When different indicator values occurred in different datasets for the same species and the same environmental variable, we calculated the mean of these values. If the difference between the minimum and maximum values across all original taxa that were merged into the same species or aggregate was more than three indicator value units across all datasets, and the range crossed the central value (i.e. a value of 5 for the 9-degree scales and a value of 6.5 for the 12-degree scales), we reported no indicator value. The condition of crossing the central degree filtered out generalist species occurring under intermediate conditions, while preserving values for species occurring under more extreme conditions. All indicator values resulting from either the averaging or median calculation that had more than one decimal place were rounded to one decimal place.

To assign indicator values to species for which indicator values were not available in any of the datasets but which occurred in at least 50 EVA vegetation plots, we used the method described by Chytrý et al. (2018). First, for each of these target species, we searched for the set of other species that had the most similar occurrence pattern across EVA plots. We measured the degree of co-occurrence of species pairs using the *phi* coefficient of association (Sokal & Rohlf, 1995). For each species with no indicator value, we listed all species with an indicator value that had a similar occurrence pattern (interspecific association of *phi* > 0.1). If there were at least five such species, we calculated the mean (rounded to one decimal place) of their indicator values and assigned it as the indicator value for the target species with no indicator value. If more than 20 species met these conditions, we considered only the 20 species with the highest *phi* value. If there were fewer than five such species, no new indicator values were calculated.

Mean indicator values always have a narrower range than the original scale of indicator values (see Hill et al., 2000), which reduces the compatibility between the newly calculated and original indicator values. To standardize the range of indicator values for species with newly-calculated values, we first calculated indicator values for species that occurred in at least one dataset of indicator values and for which we knew the original indicator values in the regional datasets. For a set of these species, we calculated a linear regression between the values estimated from species co-occurrence (independent variable) and average indicator values from the regional datasets (dependent variable). Then we used the formula of the regression line to adjust indicator values for species with values estimated only from species co-occurrence, i.e., those for which indicator values were not previously available.

Any subjective adjustment of indicator values was avoided. However, indicator values for obligatory epiphytic hemiparasites germinating on trees (*Arceuthobium, Loranthus* and *Viscum*) were not included in the final list in the case of nutrients, reaction and salinity.

We tested the validity of the harmonized indicator values using an example of indicator values for temperature by regressing them on an independent source of gridded temperature data. We calculated unweighted community-mean temperature indicator values across species in each EVA plot that contained at least five species (413,832 plots) and related them to modeled mean summer temperatures from the Chelsa database (Karger et al., 2017; bio10 – daily mean air temperatures of the warmest quarter for the period of 1981–2010). Data processing and analyses were performed using the programs JUICE v. 7.1 (Tichý, 2002) and R v. 4.0.3 (R Core Team, 2022).

Results

Of the 12 Ellenberg-type indicator-value datasets (i.e., excluding the original Ellenberg dataset), 11 were found to be at least partially compatible with the original Ellenberg dataset (Table 1, Appendix S1) after being tested with species-based regression and plot-based regression (Appendix S2). Outlier datasets that did not meet our compatibility conditions were excluded from further analyses. Indicator values for the Cantabrian Mountains were excluded entirely. For the Southern Aegean dataset, we retained the indicator values for moisture and salinity, but excluded the other values for lack of compatibility. For the Ukrainian dataset, we retained the indicator values for light and moisture, but excluded temperature (thermal climate or thermoregime).

Table 1. Regional datasets of Ellenberg-type indicator values used as a potential source for the European dataset. Numbers are given where indicator values are present in the source dataset and were considered for the calculation. The numbers are, in turn, counts of species or aggregates (after nomenclature standardization) with indicator values. 'NA' (not accepted) – the indicator value exists, and the authors declared that it follows the Ellenberg concept, but it did not meet our compatibility criteria and was excluded from further analyses. 'NC' (not considered) – indicator value exists, but its concept or scale differs from Ellenberg indicator values. '–' – indicator value does not exist in the source dataset. Information on the percentage distribution of indicator values across three indicator value ranges within each dataset is provided in Appendix S1.

Source		**Light**	**Temperature**	**Moisture**	**Reaction**	**Nutrients**	**Salinity**
Germany	Ellenberg & Leuschner (2010)	**2478**	**2191**	**2407**	**3778**	**2315**	**2495**
Austria	Karrer (1992)	**1006**	**724**	**938**	**1198**	**855**	**1000**
Cantabrian Range	Jiménez-Alfaro et al. (2021)	NA	NA	NA	NA	NA	–
Czech Republic	Chytrý et al. (2018)	**2191**	**2194**	**2194**	**2192**	**2192**	**2194**
European mires	Hájek et al. (2020)	–	–	**1479**	–	–	–
France	Julve (2015)	**3815**	**3763**	**3750**	**3758**	**3764**	**3792**
Great Britain	Hill et al. (2000)	**1684**	–	**1684**	**1684**	**1684**	**1684**
South Aegean	Böhling et al. (2002)	NA	NA	**1831**	NA	NA	**1922**
Hungary	Borhidi (1995)	**2028**	**2028**	**2028**	**2026**	**2028**	**2028**
Italy	Guarino & La Rosa (2019)	**5136**	**4985**	**5092**	**4869**	**5049**	**5121**
Saline habitats	Dítě et al. (2022)	–	–	–	–	–	**335**
Switzerland	Landolt et al. (2010)	NC	**4380**	NC	NC	NC	NC
Ukraine	Didukh (2011)	**2877**	NA	**2895**	NC	NC	NC
FINAL		**8168**	**7400**	**8030**	**7282**	**7193**	**7507**

The final dataset contained 8,908 European vascular plant species with at least one indicator value Indicator values defined for all six environmental variables were defined for 5,398 species. At least one indicator value was newly assigned for 398 species that were not listed in any regional dataset. Correlation matrix, histograms, and the relative frequency of indicator values for species or mean indicator values for vegetation plots with the relationship between each combination of the environmental variables was shown in Fig. 2 and 3.

The set of 1,790,582 vegetation plots from the EVA database contained 11,161 species of vascular plants after standardizing the nomenclature. Of these, 7,918 (70.9%) had at least one indicator value derived from at least one of the 12 retained datasets or estimated from species co-occurrences. The new indicator values were defined mainly for frequent species. Therefore, at least one indicator value was available for 99.7% of all species occurrences in the EVA vegetation plots.

Linear regressions between plot mean values calculated from the new dataset of European values for temperature and the mean summer temperature from the Chelsa dataset showed a stronger relationship (R^2 = 0.49) than regressions calculated from each regional dataset taken individually (Appendix S3). Community means for temperature values showed negligible differences in slope and coefficient of determination when calculated with or without the the species for which the value is derived from the EVA-based estimations.

Discussion

We created an extensive dataset of indicator values for six main environmental variables that affect plant distribution and community composition under natural conditions. This dataset covers a large part of Europe and is suitable for European studies of flora and vegetation. Although it does not include all the European species, it contains most of the widespread and common species, and represents the broadest harmonized source permitting sound comparisons. Our indicator values were created by mathematically integrating data from the original Ellenberg values and 11 compatible datasets for other regions in Europe. In addition, we estimated indicator values for species for which no values had been published, based on species co-occurrences in vegetation plots from the EVA database.

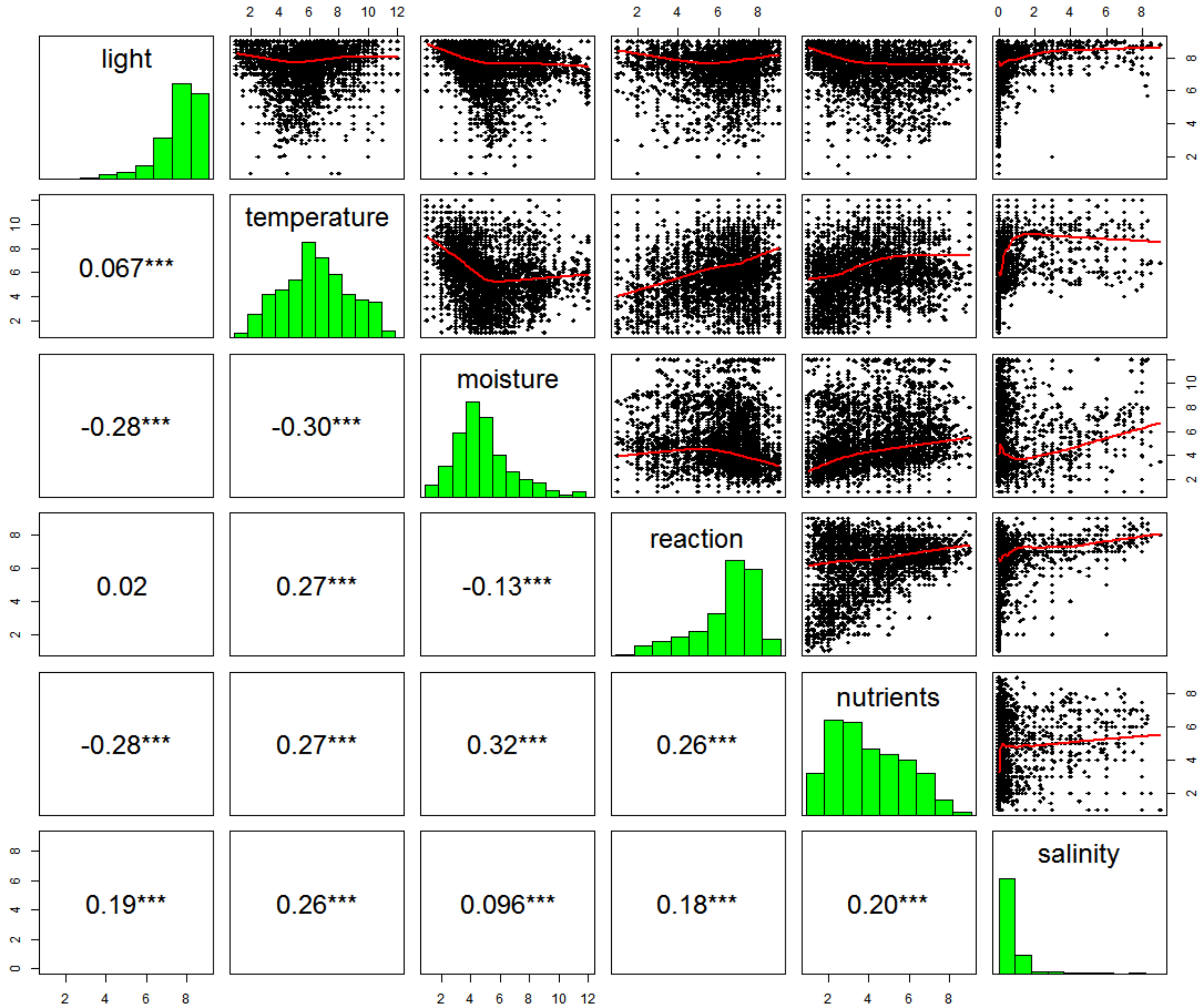


Fig. 1: Correlation matrix of Ellenberg-type indicator values for Europe. Histograms on the diagonal show the relative frequency of species for a particular value along the environmental gradient, boxes below the diagonal show Pearson correlation coefficients with their significance, and scatter plots above the diagonal show the distribution of species in a pairwise comparison between two corresponding indicators (each black dot represents one species)..

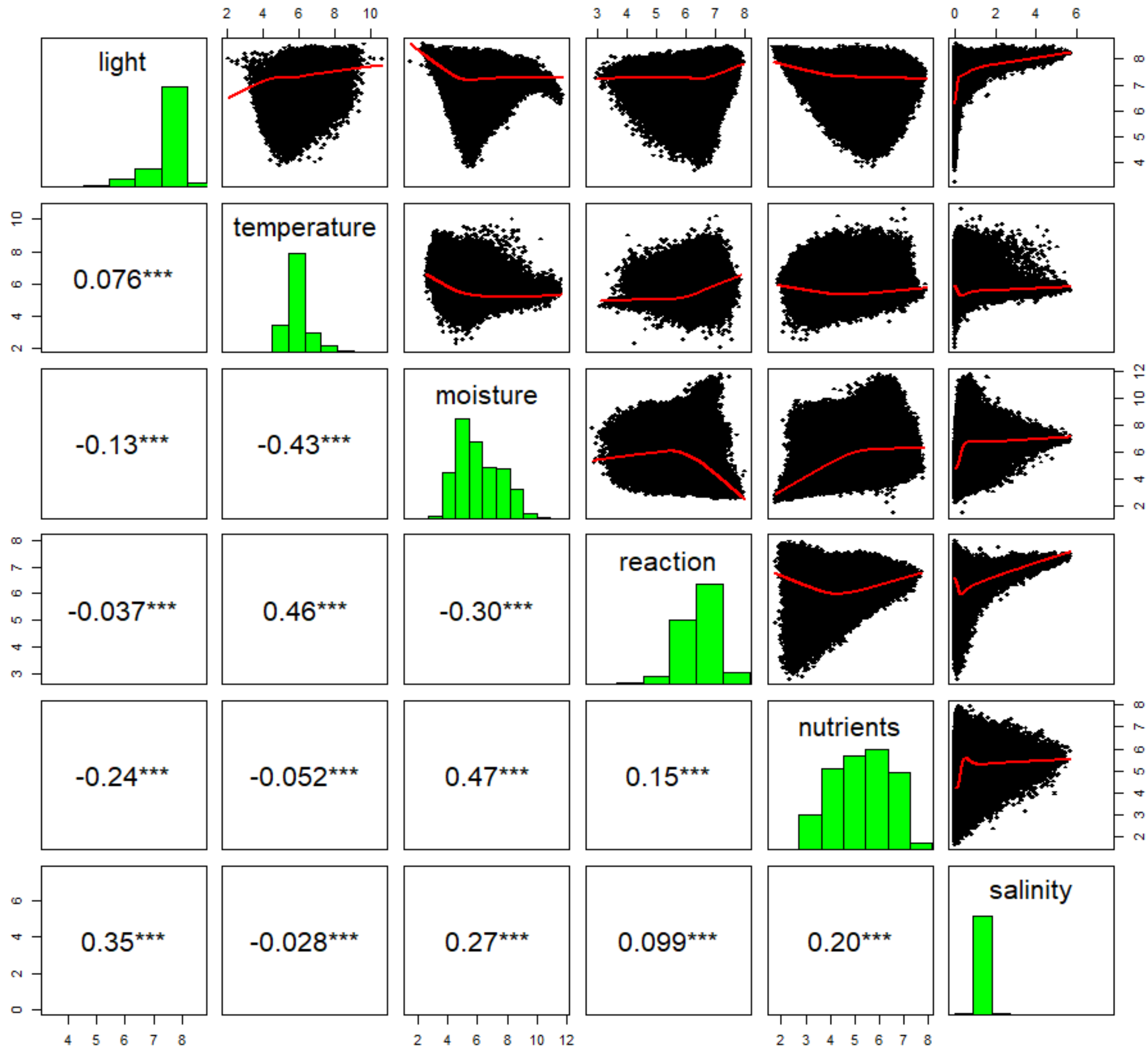


Fig. 2: Correlation matrix of community means of Ellenberg-type indicator values for Europe calculated for EVA vegetation plots. Histograms on the diagonal show the relative frequency of plots for a particular value along the environmental gradient, boxes below the diagonal show Pearson correlation coefficients with their significance, and scatter plots above the diagonal how the distribution of vegetation plots in a pairwise comparison between two corresponding indicators (each black dot represents one vegetation plot).

Alternative approaches to calculating Ellenberg-type indicator values from vegetation plots were proposed by ter Braak & Gremmen (1987) and Hill et al. (2000). They calculated indicator values by reciprocal averaging of community means of species indicator values from vegetation plots. ter Braak & Gremmen (1987) also proposed the maximum likelihood method. However, both methods utilized community means as a source for species' indicator estimation or correction. Our experience from a previous study (Chytrý et al., 2018) shows that the calculation of indicator values for new species from community means can be negatively affected by the fact that a few widespread and common generalist species are found in many plots and represent a relatively high proportion of the total number of species in individual plots. For example, only 477 out of 11,164 vascular plant species in the selection from the EVA database used for this study occur in more than 1% of plots. There are many vegetation plots in which these widespread species are the only species with an indicator value. In the case of temperature, for instance, this concerns 10.4% of all plots. As a result, some specialized species with missing indicator values may receive inappropriate values if only the average values for generalist species are used. Therefore, we suggest using only the values for the most specialized and most similarly distributed species for calculating new indicator values based on vegetation plots. The advantage of the method used in this work, as proposed by Chytrý et al. (2018), is that it does not average all species in plots, but assigns missing indicator values based on averaging the values for a limited number of species with the most similar patterns of co-occurrence. Although this method calculates indicator values only for species that frequently co-occur with other species that already have indicator values, the calculated values are more reliable.

Ellenberg (1974) and other authors defined indicator values on ordinal scales, which has sometimes been criticized (Dierschke, 1994). Ellenberg et al. (2001) argued that at least part of their scales have equidistant segmentation of the interval scale, which allows calculating community means. ter Braak & Barendregt (1986) showed that community means calculated from indicator values best estimate environmental conditions when each indicator value is the centroid of the symmetric (normally distributed) species response curve to the given environmental variable. Other authors (Pignatti et al., 2001; Marcenò & Guarino, 2015; Wildi, 2016) have also shown that in large datasets, Ellenberg indicator values can be evaluated with parametric tests because they tend to be normally distributed. Because many recent studies have also estimated environmental conditions using community means (e.g. Ahl et al., 2021; Baumann et al., 2021; Dwyer et al., 2021; Jaroszewicz et al., 2021), we considered all scales of published indicator values to be interval scales. Differences among published sources were smoothed by calculating means with decimal precision. The new dataset of indicator values retains the range of the original Ellenberg scales of nine, ten or twelve degrees, so it is compatible with other datasets defined on the same scales.

As our indicator-value dataset is prepared for broad-scale analyses, it uses a relatively coarse taxonomic resolution at the level of species or, in some cases, species aggregates. However, different subspecies of the same species or different narrowly-defined species within an aggregate may differ substantially in their ecological requirements for some environmental variables (e.g. Landolt et al., 2010). Therefore, for some species or aggregates in our dataset, no indicator value was given for some environmental variables. As a result, only 4,946 (44.3%) of the vascular plant species occurring in the EVA vegetation plots had an indicator value for all six environmental variables. Another reason for the relatively low number of such species was that we used only six datasets that contained indicator values for less than six environmental variables compatible with the Ellenberg scales (Hill et al., 2000; Böhling et al., 2002; Landolt et al., 2010; Didukh, 2011; Hájek et al., 2020; Dítě et al., 2022).

The original Ellenberg values had been estimated primarily by expert knowledge. Cornwell & Grubb (2003) demonstrated that Ellenberg species values for different environmental conditionsare often not independent. They found a significant rank correlation for the relationship between nutrients and moisture ($r_s = 0.362$, $p = 0.001$), which is also found in our harmonized dataset (Fig. 1). Similar trends of the relationship between environmental factors can be seen in Fig. 2, where we compared unweighted community means calculated for vegetation plots of the EVA database. The reason of partial intercorrelations between indicators for individual species is not so obvious as for community means, in which the problem is much more evident because the indication of ecological factors is related to exact site conditions. However, comparing Fig. 1 and Fig. 2, the interpretation of inter-correlations is not trivial.

Independent verification of the validity of our dataset of indicator values in relation to accurately measured local environmental variables is difficult because there are no standardized measurements of local conditions at the European scale at the sites where the vegetation was sampled. The only exception is temperature, which has both local and macroscale components considered in the indicator values. Therefore, the community mean indicator values can be compared with interpolated data from temperature measurements at climate stations. Such data represent macroclimate, but Ellenberg (1974) also derived temperature indicator values from species occurrence in altitudinal belts in Germany and the Alps. There was a strong relationship between mean summer temperatures from the Chelsa database (Karger et al., 2017) and community mean temperature indicator values for vegetation plots from the EVA database. However, we did not account for differences in local conditions, such as slope, aspect and shading from trees, shrubs and adjacent topographic features, which can affect local temperatures but are not available for all vegetation plots. Community means calculated from directly assigned indicator values and those calculated using species co-occurrences showed negligible differences in R^2 values (Appendix S3), largely due to the robustness of calculating community means, as also shown in Ewald (2003). Species with indicator values calculated based on species co-occurrences represented only about 3% of the species in the EVA database, and these were mainly rare species.

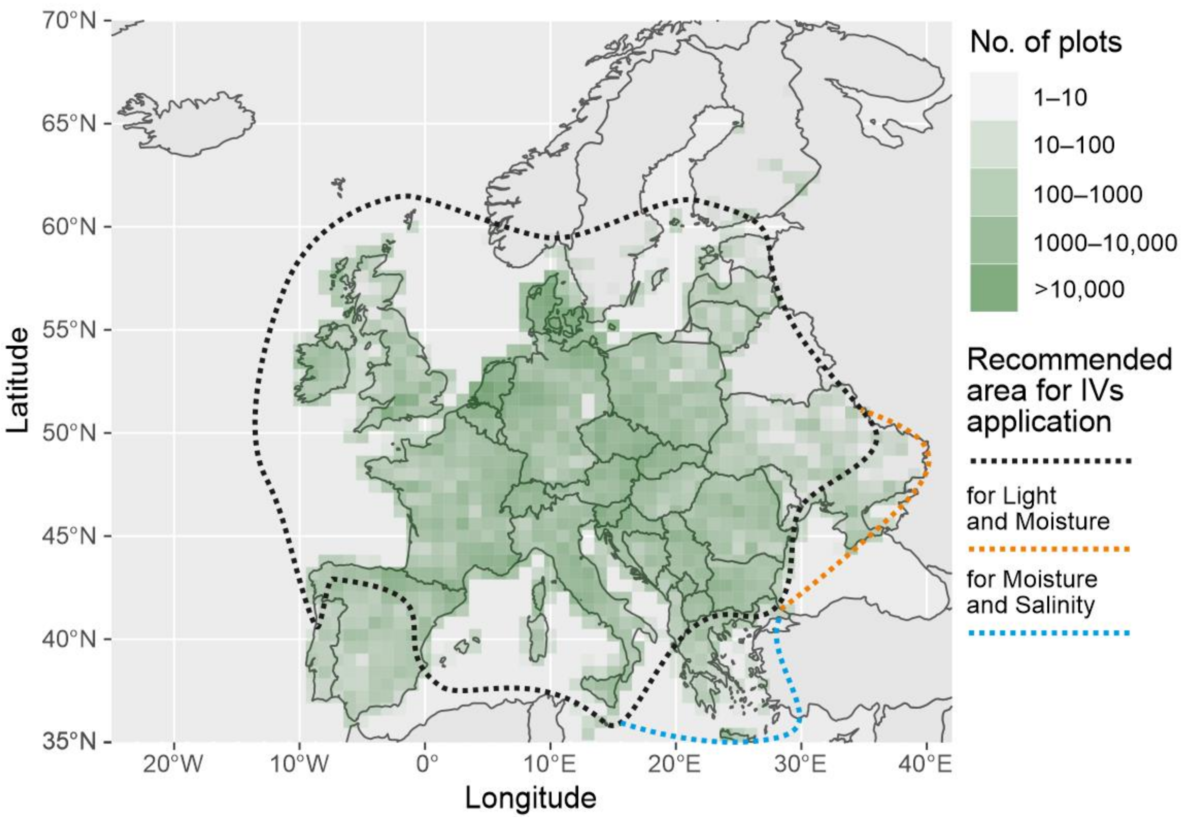


Fig. 3: Europe divided into a grid of 0.6° for latitude and 1° for longitude. Shades of green represent the density of 413,705 georeferenced vegetation plots from the EVA database that contain at least five species with indicator value for each environmental variable: light, temperature, moisture, reaction, nutrients and salinity. The black dotted line defines the approximate area, for which we recommend using the dataset of indicator values for all environmental variables. The orange dotted line indicates an additional area where light and moisture values can be safely used, and the blue-dotted line is an additional area where moisture and salinity values can be safely used.

The 12 regional datasets of species indicator values integrated into our unified dataset cover most of central and western Europe. However, their reliability decreases with distance from their area of origin (Herzberger & Karrer, 1992; Englisch & Karrer, 2001; Coudun & Gégout, 2005; Godefroid & Dana, 2007), as some species may change their realized niche or be represented by genotypes adapted to different fundamental niches (ecotypic adaptation; Hájková et al. 2008). For example, the niche width of some European species increases northward, making Ellenberg indicator values less applicable in northern Europe (Diekmann, 1995; Hedwall et al., 2019). In contrast, some species shift and narrow their niche toward the edges of their distribution range (Papuga et al., 2018) relative to their center of distribution (Englisch & Karrer, 2001). This is consistent with our comparisons of regional datasets, which showed the largest deviations from the original Ellenberg values for datasets from regions that are geographically and climatically farthest away from Germany, e.g. the Cantabrian Mountains in Spain (Jiménez-Alfaro et al., 2021) and the South Aegean region of Greece (Böhling et al., 2002). It is also likely that local endemics, e.g., in the Cantabrian Mountains and the Aegean region, outcompete species with broader geographic ranges from a part of the full realized niches of the latter, resulting in a shift of their environmental requirements and the narrowing of the realized niche. Therefore, we did not (or only partially) consider these datasets from distant areas. As a result, we consider the new dataset of indicator values to be mainly representative of Central and Western Europe, and southern and eastern adjacent biogeographical areas (Fig. 3). For the Mediterranean region, especially for the Iberian Peninsula, Greece, Turkey and probably also southeastern Ukraine, new systems of ecological indicator values need to be developed, based on local observations, expert knowledge and careful comparisons with indicator values already established in other parts of Europe.

Although the primary motivation for our work was to create a dataset of Ellenberg-type indicator values that can be used for broad-scale international studies of macroecological patterns of the European flora and vegetation, this dataset can also be used in local studies. Its advantage is that it retains the traditional Ellenberg scales. Thus, if a local study uses a regional system of Ellenberg-type indicator values from a nearby region, our harmonized European dataset can be used to add values for species that are missing from the regional system but occur in the study area. It is likely that most regional systems of indicator values provide more accurate estimates of site conditions in their region than the European dataset, which is based on averaging indicator values from different regions. For example, species that behave as generalists on the European scale and thus were not assigned an indicator value in the European dataset may have narrower niches and be good indicators in particular regions. Therefore, it is reasonable to continue to use regional systems of indicator values for local studies in regions where such systems exist. Nevertheless, if local studies from different regions use the European system of indicator values, their results can be directly compared. The next step would be to test the explanatory power of the new indicator values to predict measured climate and soil variables in plots.

Acknowledgements

We thank Cajo ter Braak for helpful comments on the manuscript, Jan Divíšek for the first version of the climate data used for testing, and database custodians and all researchers who collected the vegetation-plot data stored in the EVA database.

Author contributions

LT and MC conceived the research idea; IA standardized the nomenclature and prepared the data; RG revised the Italian indicator values; LT proposed analyses and performed all calculations; LT, MC and IA wrote the text; GM helped visualize the appendices; all authors commented on the manuscript.

Data availability statement

The vegetation-plot data used in this study are stored in the European Vegetation Archive database (EVA; http://euroveg.org/eva-database) under project number 142, product (a). Tables of original indicator values for each region and harmonized indicator values for Europe can be downloaded from the Zenodo repository (https://doi.org/10.5281/zenodo.6984813), where also the future updates will be available.

Supporting information of the paper

Appendix S1. Percentages of indicator values in regional datasets selected as a potential source for a harmonized European dataset of indicator values.

Appendix S2. Evaluation of 12 regional systems of ecological indicator values based on their relationship to Ellenberg indicator values.

Appendix S3. Comparison of mean Ellenberg-type indicator values for temperature calculated for vegetation plots and mean summer temperature for plot locations obtained from climatic datasets.

References

Ahl, L., Aas, G., Walentowski, H., Höltken, A.M. & Feulner, M. (2021) Niche differentiation between *Malus sylvestris* and its hybrid with *Malus domestica* indicated by plant community, soil and light. *Journal of Vegetation Science*, 32, e13078. https://doi.org/10.1111/jvs.13078

Baumann, M., Dittrich, S., Körner, M. & von Oheimb, G. (2021) Temporal changes in the ground vegetation in spruce forests in the Erzgebirge (Ore Mountains) — bryophytes are better indicators of the impact of liming and of sulphur and nitrogen deposition than the herb layer. *Applied Vegetation Science*, 24, e12598. https://doi.org/10.1111/avsc.12598

Berg, C., Welk, E. & Jäger, E.J. (2017) Revising Ellenberg's indicator values for continentality based on global vascular plant species distribution. *Applied Vegetation Science*, 20, 482–493. https://doi.org/10.1111/avsc.12306

Böhling, N., Greuter, W. & Raus, T. (2002) Zeigerwerte der Gefäßpflanzen der Südägäis (Griechenland) (Indicator values of vascular plants in the South Aegean (Greece)). *Braun-Blanquetia*, 32, 1–108.

Boller-Elmer, K.C. (1977) Stickstoff-Düngungseinflüsse von Intensiv-Grünland auf Streu- und Moorwiesen. *Veröffentlichungen des Geobotanischen Institutes der Eidg. Tech. Hochschule, Stiftung Rübel, in Zürich*, 63, 1-103. https://doi.org/10.3929/ethz-a-000123290

Borhidi, A. (1995) Social behaviour types, the naturalness and relative ecological indicator values of the higher plants in the Hungarian flora. *Acta Botanica Hungarica*, 39, 97–181.

Briemle, G. (1986) Vergleich der Stickstoff-Mineralisation mit der N-Zahl Ellenberg's am Beispiel einer Streuwiese im Alpenvorland – Erste Erfahrungen mit zweijaihrigen Nmin- Untersuchungen (Comparison of nitrogen mineralization with Ellenberg's N-value using the example of a litter meadow in the foothills of the Alps – first experiences with two-year Nmin investigations). *Natur und Landschaft*, 61, 423–427.

Cajander, A.K. (1926) The theory of forest types. *Acta Forestalia Fennica*, 29, 1–108. https://doi.org/10.14214/aff.7193.

Chytrý, M., Hejcman, M., Hennekens, S.M. & Schellberg, J. (2009) Changes in vegetation types and Ellenberg indicator values after 65 years of fertilizer application in the Rengen Grassland Experiment, Germany. *Applied Vegetation Science*, 12, 167–176. https://doi.org/10.1111/j.1654-109X.2009.01011.x

Chytrý, M., Hennekens, S.M., Jiménez-Alfaro, B., Knollová, I., Dengler, J., Jansen, F. et al. (2016) European Vegetation Archive (EVA): an integrated database of European vegetation plots. *Applied Vegetation Science*, 19, 173–180. https://doi.org/10.1111/avsc.12191

Chytrý, M., Tichý, L., Dřevojan, P., Sádlo, J. & Zelený, D. (2018) Ellenberg-type indicator values for the Czech flora. *Preslia*, 90, 83–103. https://doi.org/10.23855/preslia.2018.083

Chytrý, M., Tichý, L., Hennekens, S.M., Knollová, I., Janssen, J.A.M., Rodwell, J.S. et al. (2020) EUNIS Habitat Classification: Expert system, characteristic species combinations and distribution maps of European habitats. *Applied Vegetation Science*, 23, 648–675. https://doi.org/10.1111/avsc.12519

Cornwell, W.K. & Grubb, P.J. (2003) Regional and local patterns in plant species richness with respect to resource availability. *Oikos*, 100, 417–428. https://doi.org/10.1034/j.1600-0706.2003.11697.x

Coudun, C. & Gégout, J.-C. (2005) Ecological behaviour of herbaceous forest species along a pH gradient: a comparison between oceanic and semicontinental regions in northern France. *Global Ecology and Biogeography*, 14, 263–270. https://doi.org/10.1111/j.1466-822X.2005.00144.x

Didukh, Ya.P. (2011) *The ecological scales for the species of Ukrainian flora and their use in synphytoindication.* Kyiv: Phytosociocentre.

Diekmann, M. (1995) Use and improvement of Ellenberg's indicator values in deciduous forests of the Boreo-nemoral zone in Sweden. *Ecography*, 18, 178–189. https://doi.org/10.1111/j.1600-0587.1995.tb00339.x

Diekmann, M. (2003) Species indicator values as an important tool in applied plant ecology – a review. *Basic and Applied Ecology*, 4, 493–506. https://doi.org/10.1078/1439-1791-00185

Diekmann, M. & Lawesson, J.E. (1999) Shifts in ecological behaviour of herbaceous forest species along a transect from northern central to North Europe. *Folia Geobotanica*, 34, 127–141. https://doi.org/10.1007/BF02803080

Dierschke, H. (1994) *Pflanzensoziologie: Grundlagen und Methoden (Plant sociology: principles and methods).* Stuttgart: Ulmer. 683 pp.

Dítě, D., Šuvada, R., Tóth, T. & Dítě, Z. (2022) Inventory of halophytes in central Europe. *Preslia,* (reviewed).

Domina, G., Galasso, G., Bartolucci, F.. & Guarino, R. (2018) Ellenberg Indicator Values for the vascular flora alien to Italy. *Flora Mediterranea*, 28, 53–61. https://doi.org/*10.7320/FlMedit28.053*

Dwyer, C., Millett, J., Pakeman, R.J. & Jones, L. (2021) Environmental modifiers of the relationship between water table depth and Ellenberg's indicator of soil moisture. *Ecological Indicators*, 132, article 108320. https://doi.org/10.1016/j.ecolind.2021.108320

Ellenberg, H. (1948) Unkrautgesellschaften als Mass fuer den Säeuregrad, die Verdichtung und andere Eigenschaften des Ackerbodens. *Berichten der Landtechnik,* 4, 130–146.

Ellenberg, H. (1950) *Landwirtschaftliche Pflanzensoziologie. I. Unkrautgemeinschaften als Zeiger für Klima und Boden (Agricultural Plant Sociology. I. Weed communities as indicators of climate and soil)*. Stuttgart: Ulmer. 141 pp.

Ellenberg, H. (1952) *Landwirtschaftliche Pflanzensoziologie. II. Wiesen und Weiden und ihre standörtliche Bewertung (Agricultural Plant Sociology. II. Meadows and pastures and their site assessment)*. Stuttgart: Ulmer. 143 pp.

Ellenberg, H. (1974) Zeigerwerte der Gefäßpflanzen Mitteleuropas (Indicator values of vascular plants in Central Europe). *Scripta Geobotanica*, 9, 1–97.

Ellenberg, H. & Leuschner, C. (2010) Zeigerwerte der Pflanzen Mitteleuropas (Indicator values of vascular plants in Central Europe). In: Ellenberg H. & Leuschner, C., *Vegetation Mitteleuropas mit den Alpen (Vegetation of Central Europe including the Alps)*. 6th ed. Stuttgart: Ulmer. 1334 pp. https://doi.org/10.36198/9783825281045.

Ellenberg, H., Weber, H.E., Düll, R., Wirth, V., Werner, W. & Pauliβen, D. (1992) Zeigerwerte von Pflanzen in Mitteleuropa. (Indicator values of plants in Central Europe) 2nd ed. *Scripta Geobotanica*, 18, 1–258.

Ellenberg, H., Weber, H.E., Düll, R., Wirth, V. & Werner, W. (2001) Zeigerwerte von Pflanzen in Mitteleuropa (Indicator values of plants in Central Europe). 3rd ed. *Scripta Geobotanica*, 18, 1–262.

Englisch, T. & Karrer, G. (2001) Zeigerwertsysteme in der Vegetationsanalyse – Anwendbarkeit, Nutzen und Probleme in Österreich (Indicator value systems in vegetation analysis - applicability, utility and problems in Austria). *Berichte der Reinhold-Tüxen-Gesellschaft*, 13, 83–102.

Ertsen, A.C.D., Alkemade, J.R.M. & Wassen, M.J. (1998) Calibrating Ellenberg indicator values for moisture, acidity, nutrient availability and salinity in the Netherlands. *Plant Ecology*, 135, 113–124.

Euro+Med (2021) *Euro+Med PlantBase – the information resource for Euro-Mediterranean plant diversity*. Available at http://ww2.bgbm.org/EuroPlusMed/ [accessed 2021]

Ewald, J. (2003) The sensitivity of Ellenberg indicator values to the completeness of vegetation relevés. *Basic and Applied Ecology*, 4, 507–513. https://doi.org/10.1078/1439-1791-00155

Gégout, J.C. & Krizova, E. (2003) Comparison of indicator values of forest understory plant species in western Carpathians (Slovakia) and Vosges Mountains (France). *Forest Ecology and Management*, 182, 1–11. https://doi.org/10.1016/S0378-1127(03)00068-9

Godefroid, S. & Dana, E.D. (2007) Can Ellenberg's indicator values for Mediterranean plants be used outside their region of definition? *Journal of Biogeography*, 34, 62–68. https://doi.org/10.1111/j.1365-2699.2006.01582.x

Guarino, R. & La Rosa, M. (2019) Digital Italian Flora (Italian). In: Pignatti, S., Guarino, R. & La Rosa, M. (Eds), *Flora d'Italia,* 2nd edition. Bologna: Edagricole, Edizioni Agricole di New Business Media.

Hájek, M., Dítě, D., Horsáková, V., Mikulášková, E., Peterka, T., Navrátilová, J. et al. (2020) Towards the pan-European bioindication system: Assessing and testing updated hydrological indicator values for vascular plants and bryophytes in mires. *Ecological Indicators*, 116, article 106527. http://dx.doi.org/10.1016/j.ecolind.2020.106527

Hájková, P., Hájek, M., Apostolova, I., Zelený, D. & Dítě, D. (2008) Shifts in the ecological behaviour of plant species between two distant regions: evidence from the base richness gradient in mires. *Journal of Biogeography*, 35, 282–294. https://doi.org/10.1111/j.1365-2699.2007.01793.x

Hedwall, P.-O., Brunet, J. & Diekmann, M. (2019) With Ellenberg indicator values towards the north: Does the indicative power decrease with distance from Central Europe? *Journal of Biogeography*, 46, 1041–1053. https://doi.org/10.1111/jbi.13565

Herzberger, E. & Karrer, G. (1992) Test der internen Konsistenz ökologischer Zeigerwerte am Beispiel der Vegetationsaufnahmen der Österreichischen Waldboden-Zustandsinventur (Test of the internal consistency of ecological indicator values using the example of vegetation plots of the Austrian forest soil status inventory). *FBVA-Berichte*, 70, 93–102.

Hill, M.O. & Carey, P.D. (1997) Prediction of yield in the Rothamsted Park Grass Experiment by Ellenberg indicator values. *Journal of Vegetation Science*, 8, 579– 586. https://doi.org/10.2307/3237210

Hill, M.O., Roy, D.B., Mountford, J.O. & Bunce, R.G.H. (2000) Extending Ellenberg's indicator values to a new area: an algorithmic approach. *Journal of Applied Ecology*, 37, 3–15. https://doi.org/10.1046/j.1365-2664.2000.00466.x

Hill, M.O., Preston, C.D. & Roy, D.B. (2004) *PLANTATT – Attributes of British and Irish plants: status, size, life history, geography and habitats*. Huntingdon: Centre for Ecology & Hydrology.

Holtland, W.J., ter Braak, C.J.F. & Schouten, M.G.C. (2010) Iteratio: calculating environmental indicator values for species and relevés. *Applied Vegetation Science*, 13, 369–377. https://doi.org/10.1111/j.1654-109X.2009.01069.x

Iversen, J. (1936) Biologische Pflanzentypen als Hilfsmittel in der Vegetationsforschung. Ein Beitrag zur ökologischen Charakterisierung und Anordnung der Pflanzengesellschaften (Biological plant types as tools in vegetation research. A contribution to the ecological characterization and arrangement of plant communities). *Meddelelser fra Skalling Laboratoriet Kobenhavn*, 4, 1–224.

Jaroszewicz, B., Borysowicz, J. & Cholewińska, O. (2021) Forest floor plant diversity drives the use of mature spruce forests by European bison. *Ecology and Evolution*, 11, 636–647. https://doi.org/10.1002/ece3.7094

Jiménez-Alfaro, B., Carlón, L., Fernández-Pascual, E., Acedo, C., Alfaro-Saiz, E., Alonso Redondo, R. et al. (2021) Checklist of the vascular plants of the Cantabrian mountains. *Mediterranean Botany*, 42, e74570. https://doi.org/10.5209/mbot.74570

Julve, P. (2015) *Baseflor. Index botanique, écologique et chorologique de la flore de France (Baseflor. Botanical, ecological and chorological index of the flora of France).* Available at http://philippe.julve.pagesperso-orange.fr/catminat.htm [accessed 2022].

Jurko, A. (1990) *Ekologické a socioekonomické hodnotenie vegetácie [Ecological and socioeconomic assessment of vegetation (Slovak)]*. Bratislava: Príroda.

Karger, D.N., Conrad, O., Böhner, J., Kawohl, T., Kreft, H., Soria-Auza, R.W. et al. (2017) Climatologies at high resolution for the earth's land surface areas. *Scientific Data*, 4, 170122. https://doi.org/10.1038/sdata.2017.122

Karrer, G. (1992) Österreichische Waldboden-Zustandsinventur. Teil VII: Vegetationsökologische Analysen (Austrian forest soil status inventory. Part VII: Vegetation ecology analyses). *Mitteilungen Forstliche Bundesversuchsanstalt Wien*, 168, 193–242.

Landolt, E. (1977) Ökologische Zeigerwerte zur Schweizer Flora (Ecological indicator values for the Swiss flora). *Veröffentlichungen des Geobotanischen Institutes der Eidg. Tech. Hochschule, Stiftung Rübel, in Zürich*, 64, 1–208.

Landolt, E., Bäumler, B., Erhardt, A., Hegg, O., Klötzli, F., Lämmler, W. et al. (2010) *Flora indicativa – Ökologische Zeigerwerte und biologische Kennzeichen zur Flora der Schweiz und der Alpen (Flora* *indicativa - Ecological indicator values and biological features for the flora of Switzerland and the Alps)*. Bern: Haupt Verlag.

Lawesson, J.E., Fosaa, A.M. & Olsen, E. (2003) Calibration of Ellenberg indicator values for the Faroe Islands. *Applied Vegetation Science*, 6, 53–62. https://doi.org/10.1111/j.1654-109X.2003.tb00564.x

Marcenò, C. & Guarino, R. (2015) A test on Ellenberg indicator values in the Mediterranean evergreen woods (*Quercetea ilicis*). *Rendiconti Lincei. Scienze Fisiche e Naturali*, 26, 345–356. https://doi.org/10.1007/s12210-015-0448-8

Mayor López, M. (1996) *Indicadores ecológicos y grupos socioecológicos en el Principado de Asturias (Ecological indicators and socio-ecological groups in the Principality of Asturias).* Oviedo: Universidad de Oviedo.

Papuga, G., Gauthier, P., Pons, V., Farris, E. & Thompson, J.D. (2018) Ecological niche differentiation in peripheral populations: a comparative analysis of eleven Mediterranean plant species. *Ecography*, 41, 1650–1664. https://doi.org/10.1111/ecog.03331

Pignatti, S., Bianco, P.M., Fanelli, G., Guarino, R., Petersen, L. & Tescarollo, P. (2001) Reliability and effectiveness of Ellenberg's indices in checking flora and vegetation changes induced by climatic variations. In: Walter, G.R., Burga, C.A. & Edwards, P.J. (Eds), *Fingerprints of climate changes: adapted behaviour and shifting species ranges*, New York: Springer, pp. 281–304. https://doi.org/10.1002/joc.871

Pignatti, S. (2005) Valori di bioindicazione delle piante vascolari della flora d'Italia (Bioindicator values of vascular plants of the flora of Italy). *Braun-Blanquetia*, 39, 1–97.

R Core Team (2022) *R: A Language and Environment for Statistical Computing.* Vienna: R Foundation for Statistical Computing. URL: https://www.R-project.org/

Richardson, D.M., Pyšek, P., Rejmánek, M., Barbour, M.G., Panetta, F.D. & West, C.J. (2000) Naturalization and invasion of alien plants: concepts and definitions. *Diversity and Distributions*, 6, 93–107. https://doi.org/10.1046/j.1472-4642.2000.00083.x

Schaffers, A.P. & Sýkora, K.V. (2000) Reliability of Ellenberg indicator values for moisture, nitrogen and soil reaction: a comparison with field measurements. *Journal of Vegetation Science*, 11, 225–244. https://doi.org/10.2307/3236802

Sicuriello, F., De Nicola, C., Dowgiallo, G. & Testi, A. (2014) Assessing the habitat conservation status by soil parameters and plant ecoindicators. *iForest*, 7, 170–177. https://doi.org/10.3832/ifor0963-007

Sokal, R.R. & Rohlf, F.J. (1995) *Biometry*, 3rd edition. New York: Freeman.

ter Braak, C.J.F. (1987) *Unimodal models to relate species to environment.* Agricultural mathematics Group, Wageningen. http://edepot.wur.nl/201452.

ter Braak, C.J.F. & Barendregt, L.G. (1986) Weighted averaging of species indicator values: its efficiency in environmental calibration. *Mathematical Biosciences*, 78, 57–72. https://doi.org/10.1016/0025-5564(86)90031-3

ter Braak, C.J.F. & Gremmen, N.J.M. (1987) Ecological amplitudes of plant species and the internal consistency of Ellenberg's indicator values for moisture. *Vegetatio*, 69, 79–87.

Tichý, L. (2002) JUICE, software for vegetation classification. *Journal of Vegetation Science*, 13, 451–453. https://doi.org/10.1111/j.1654-1103.2002.tb02069.x

Tsyganov, D.N. (1983) *Phytoindication of ecological regimes in the mixed coniferous-broad-leaved forest subzone* (Russian). Moskva: Nauka.

Wamelink, G.W.W., Joosten, V., van Dobben, H.F. & Berendse, F. (2002) Validity of Ellenberg indicator values judged from physico-chemical field measurements. *Journal of Vegetation Science*, 13, 269–278. https://doi.org/10.1111/j.1654-1103.2002.tb02047.x

Wasof, S., Lenoir, J., Gallet-Moron, E., Jamoneau, A., Brunet, J., Cousins, S.A.O. et al. (2013) Ecological niche shifts of understorey plants along a latitudinal gradient of temperate forests in northwestern Europe. *Global Ecology and Biogeography*, 22, 1130–1140. https://doi.org/10.1111/geb.12073

Wildi, O. (2016) Why mean indicator values are not biased. *Journal of Vegetation Science*, 27, 40–49. https://doi.org/10.1111/jvs.12336

Zarzycki, K., Trzcińska-Tacik, H., Różański, W., Szeląg, Z., Wołek, J. & Korzeniak, U. (2002) Ecological indicator values of vascular plants of Poland. Kraków: W. Szafer Institute of Botany, Polish Academy of Sciences.

Zlatník, A., Križo, M., Svrček, M. & Manica, M. (1970) Lesnická botanika speciální (Forestry botany). Praha: Státní zemědělské nakladatelství.

Zólyomi, B., Baráth, Z., Fekete, G., Jakucs, P., Kárpáti, I., Kovács, M. et al. (1967) Einreihung von 1400 Arten der ungarischen Flora in ökologische Gruppen nach TWR-Zahlen (Classification of 1400 species of the Hungarian flora in ecological groups following the TWR numbers). *Fragmenta Botanica Musei Historico-naturalis Hungarici*, 4, 101–142.

Supporting information of the paper

Tichý, L. et al. (2022) Ellenberg-type indicator values for European vascular plant species. *Journal of Vegetation Science.*

Appendix S1. Percentages of indicator values in regional datasets selected as a potential source for a harmonized European dataset of indicator values.

Percentages are given where indicator values are present in the source dataset and were included in the calculation. Three categories (cat1, cat2, cat3) express percentages of species (or aggregates, after nomenclature standardization) in three indicator value ranges – 1.0–3.0; 3.1–6.9; 7.0–9.0 (7.0–12.0 for temperature and moisture). ‘NA’ (not accepted) – the indicator value exists, and the authors stated that it follows the Ellenberg concept, but it did not meet our compatibility rules and was excluded from further analyses. ‘NC’ (not considered) – the indicator value exists, but its concept or scale differs from the Ellenberg indicator values. ‘–‘ – the indicator value does not exist in the source dataset.

Source		**Light** % 1–3 | 3–7 | 7–9	**Temperature** % 1–3 | 3–7 | 7–12	**Moisture** % 1–3 | 3–7 | 7–12	**Reaction** % 1–3 | 3–7 | 7–9	**Nutrients** % 1–3 | 3–7 | 7–9	**Salinity** % 0–3 | 3–7 | 7–9
Germany	Ellenberg & Leuschner (2010)	**3| 21 |76**	**19| 57 |24**	**18| 54 |28**	**7| 57 |36**	**48| 34 |18**	**98| 1 |1**
Austria	Karrer (1992)	**2| 24 |74**	**24| 57 |17**	**13| 57 |30**	**6| 57 |37**	**46| 32 |22**	**97| 2 |1**
Cantabrian Range	Jiménez-Alfaro et al. (2021)	NA	NA	NA	NA	NA	–
Czech Republic	Chytrý et al. (2018)	**3| 27 |70**	**5| 68 |17**	**19| 55 |36**	**7| 36 |57**	**33| 47 |20**	**98| 1 |1**
European mires	Hájek et al. (2020)	–	–	**5| 53 |42**	–	–	–
France	Julve (2015)	**1| 17 | 82**	**20| 47 |33**	**9| 66 |25**	**9| 37 |54**	**31| 50 |19**	**90| 6 |4**
Great Britain	Hill et al. (2000)	**1| 21 |78**	–	**10| 58 |32**	**7| 41 |52**	**38| 47 |15**	**97| 3 |0**
South Aegean	Böhling et al. (2002)	NA	NA	**31| 54 |15**	NA	NA	**96| 3 |1**
Hungary	Borhidi (1995)	**3| 24 |73**	**2| 58 |40**	**28| 34 |36**	**2| 35 |63**	**41| 42 |17**	**98| 1 |1**
Italy	Guarino & La Rosa (2019)	**2| 18 |80**	**14| 38 |48**	**42| 42 |16**	**12| 44 |44**	**52| 37 |11**	**98| 1 |1**
Saline habitats	Dítě et al. (2022)	–	–	–	–	–	**44| 55 |1**
Switzerland	Landolt et al. (2010)	NC	**18| 34 |48**	NC	NC	NC	NC
Ukraine	Didukh (2011)	**1| 17 |82**	NA	**1| 77 |22**	NC	NC	NC
FINAL		**1| 15 |84**	**12| 42 |46**	**27| 59 |14**	**9| 40 |51**	**49| 43 |8**	**96| 3 |1**

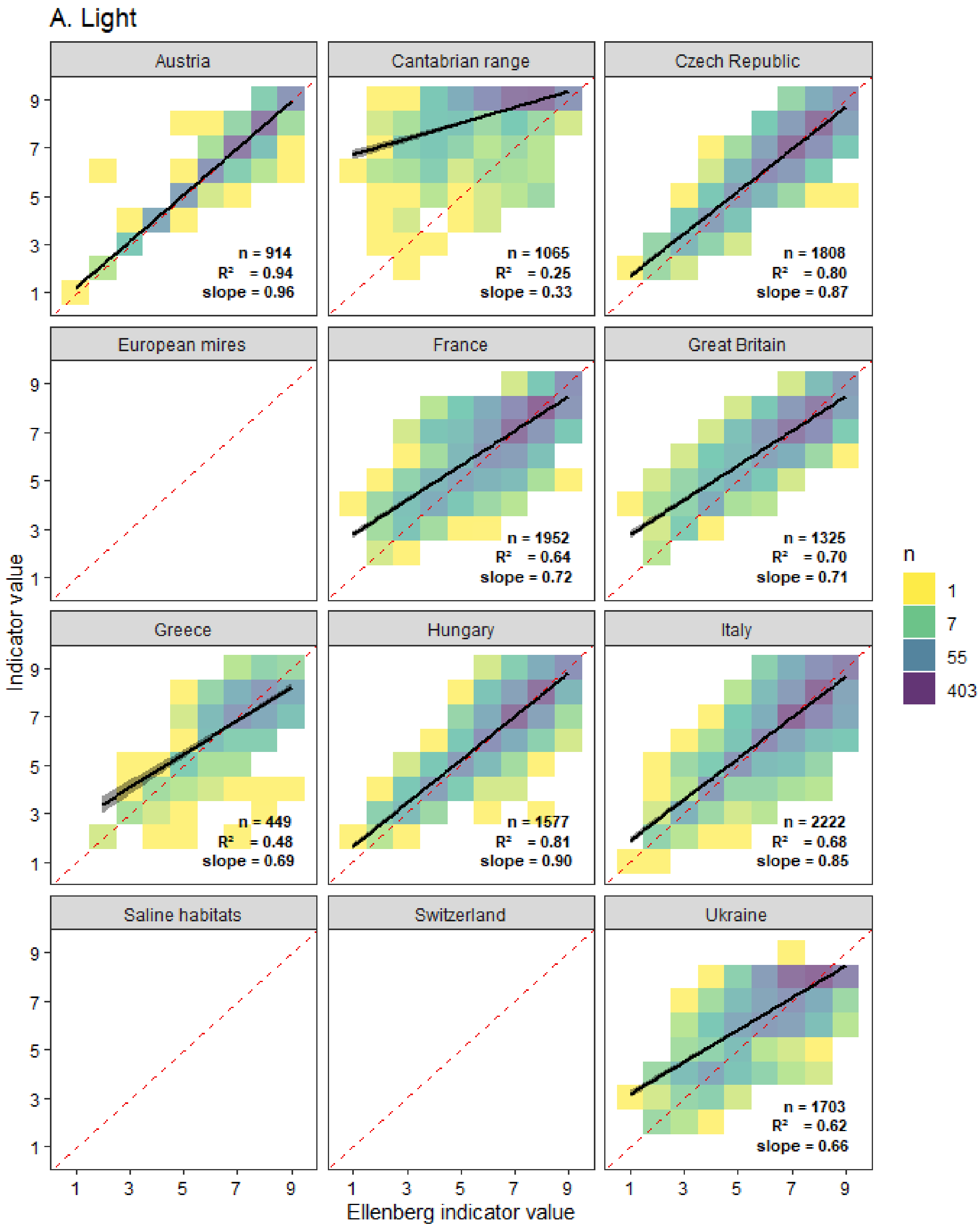
Supporting information of the paper

Tichý, L. et al. (2022) Ellenberg-type indicator values for European vascular plant species. *Journal of Vegetation Science.*

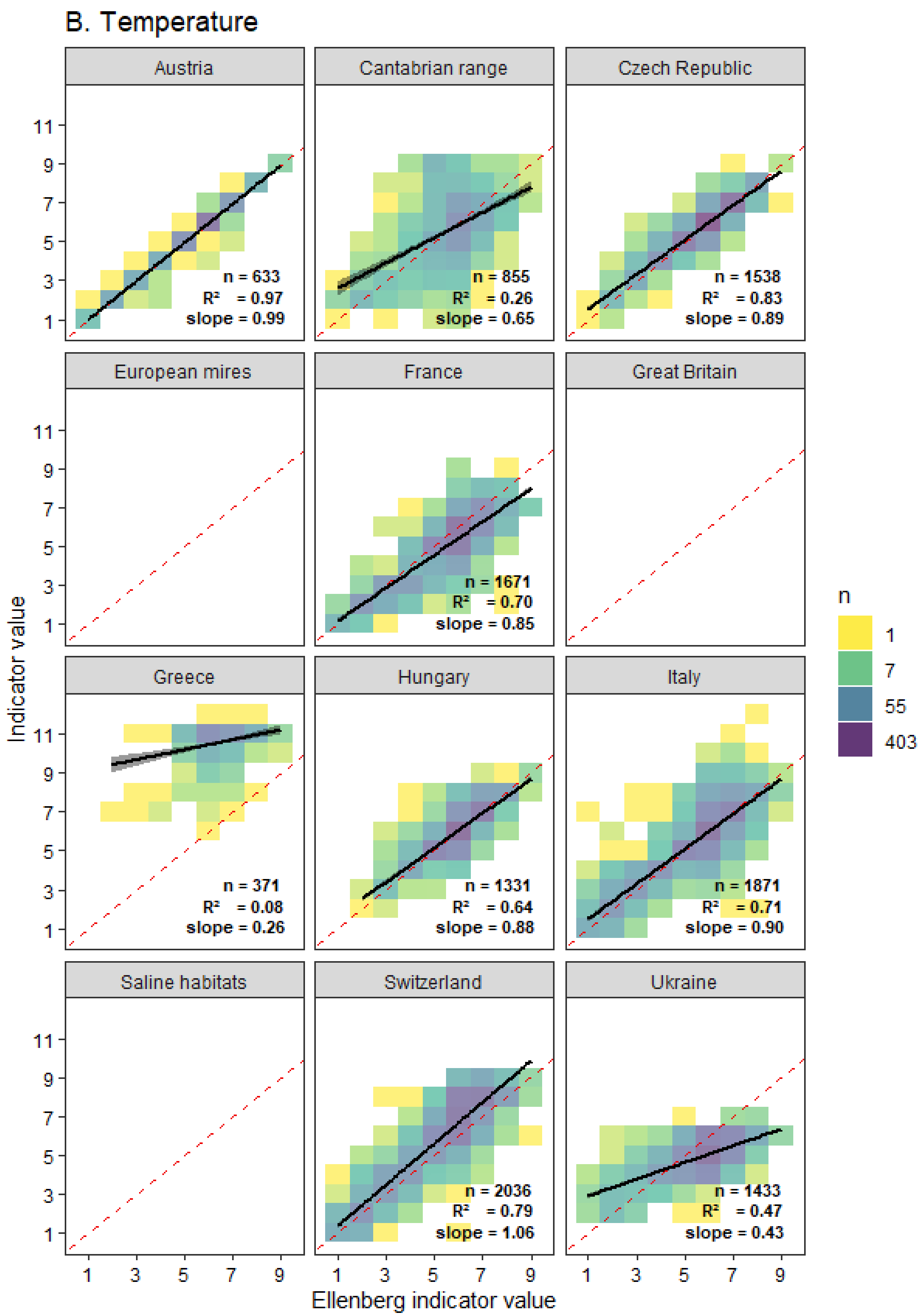
Appendix S2. Evaluation of 12 regional systems of ecological indicator values based on their relationship to Ellenberg indicator values

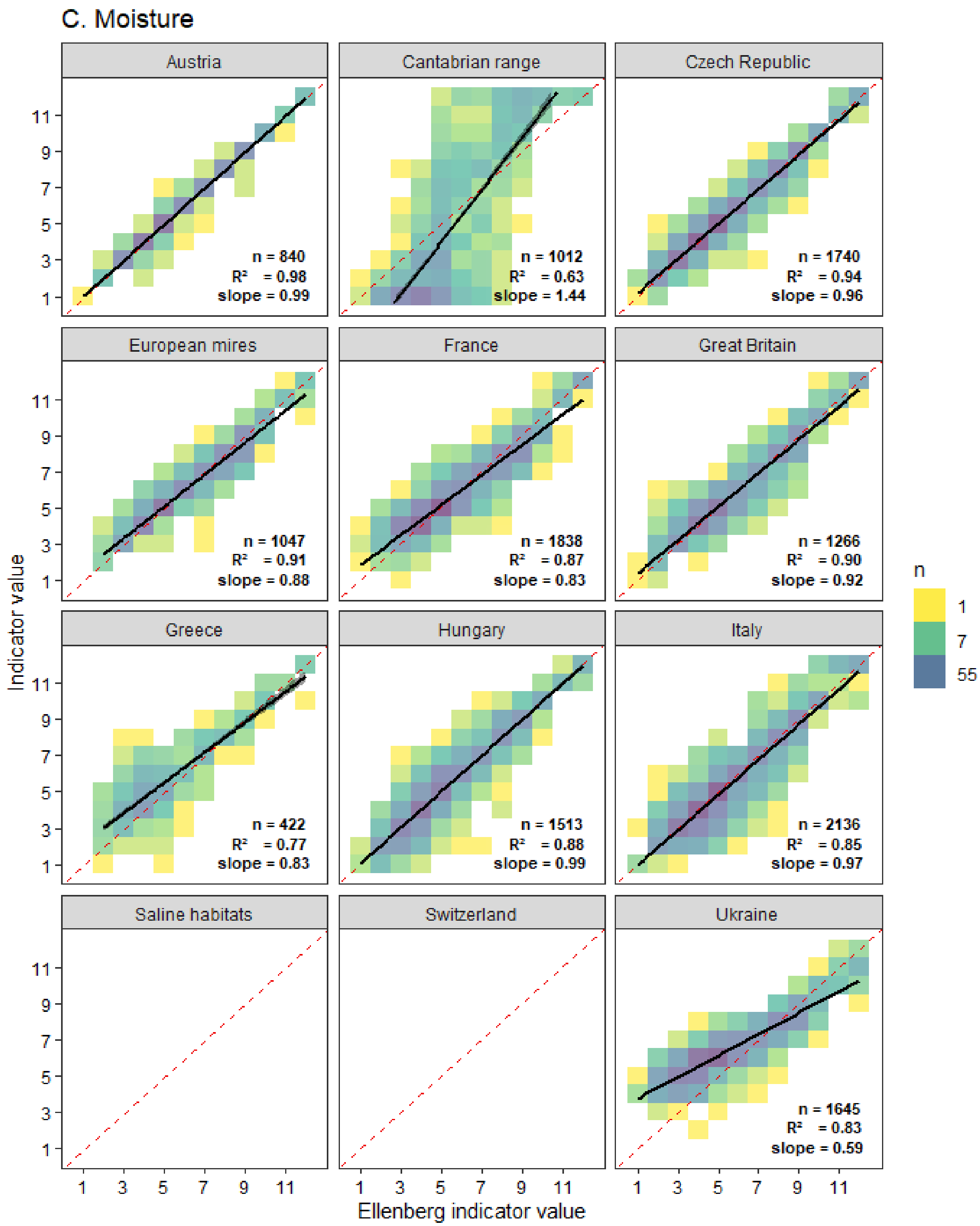
A. Direct comparisons based on species indicator values

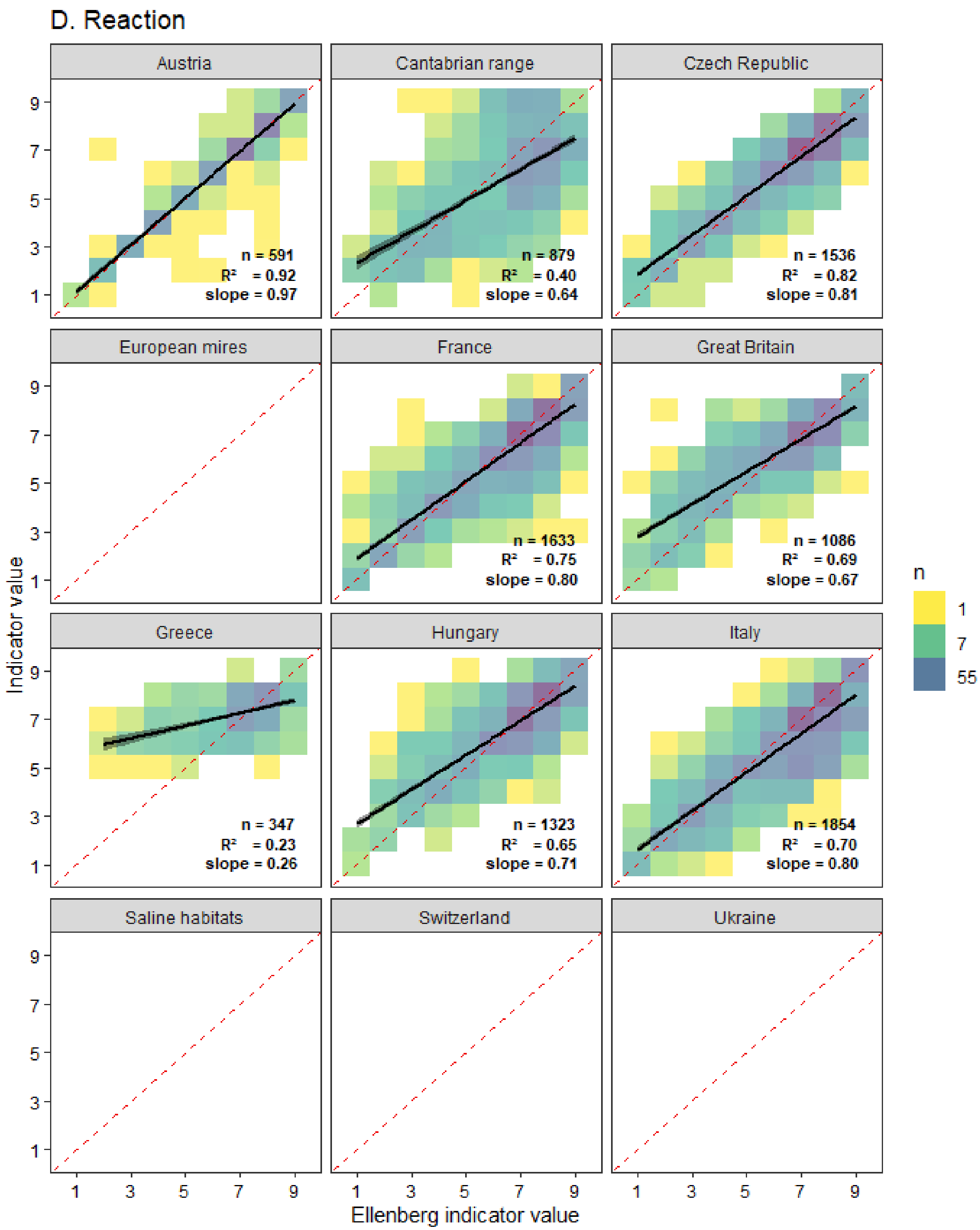
Fig. 1A–F: Comparison of regional datasets of indicator values for Austria, Cantabrian Mountains, Czech Republic, European mires, France, Great Britain, Greece, Hungary, Italy, Saline habitats, Switzerland and the Alps, and Ukraine with Ellenberg indicator values (Germany; Ellenberg & Leuschner, 2010). Most values are scaled from 1 to 9, but the range is broader (from 1 to 12) for temperature (Greece and Italy) and moisture (all datasets).


A. Light
Austria
Cantabrian range
Czech Republic
European mires
France
Great Britain
Greece
Hungary
Italy
Saline habitats
Switzerland
Ukraine
n = 914
R² = 0.94
slope = 0.96
n = 1065
R² = 0.25
slope = 0.33
n = 1808
R² = 0.80
slope = 0.87
n = 1952
R² = 0.64
slope = 0.72
n = 1325
R² = 0.70
slope = 0.71
n = 449
R² = 0.48
slope = 0.69
n = 1577
R² = 0.81
slope = 0.90
n = 2222
R² = 0.68
slope = 0.85
n = 1703
R² = 0.62
slope = 0.66
Indicator value
Ellenberg indicator value
n
1
7
55
403

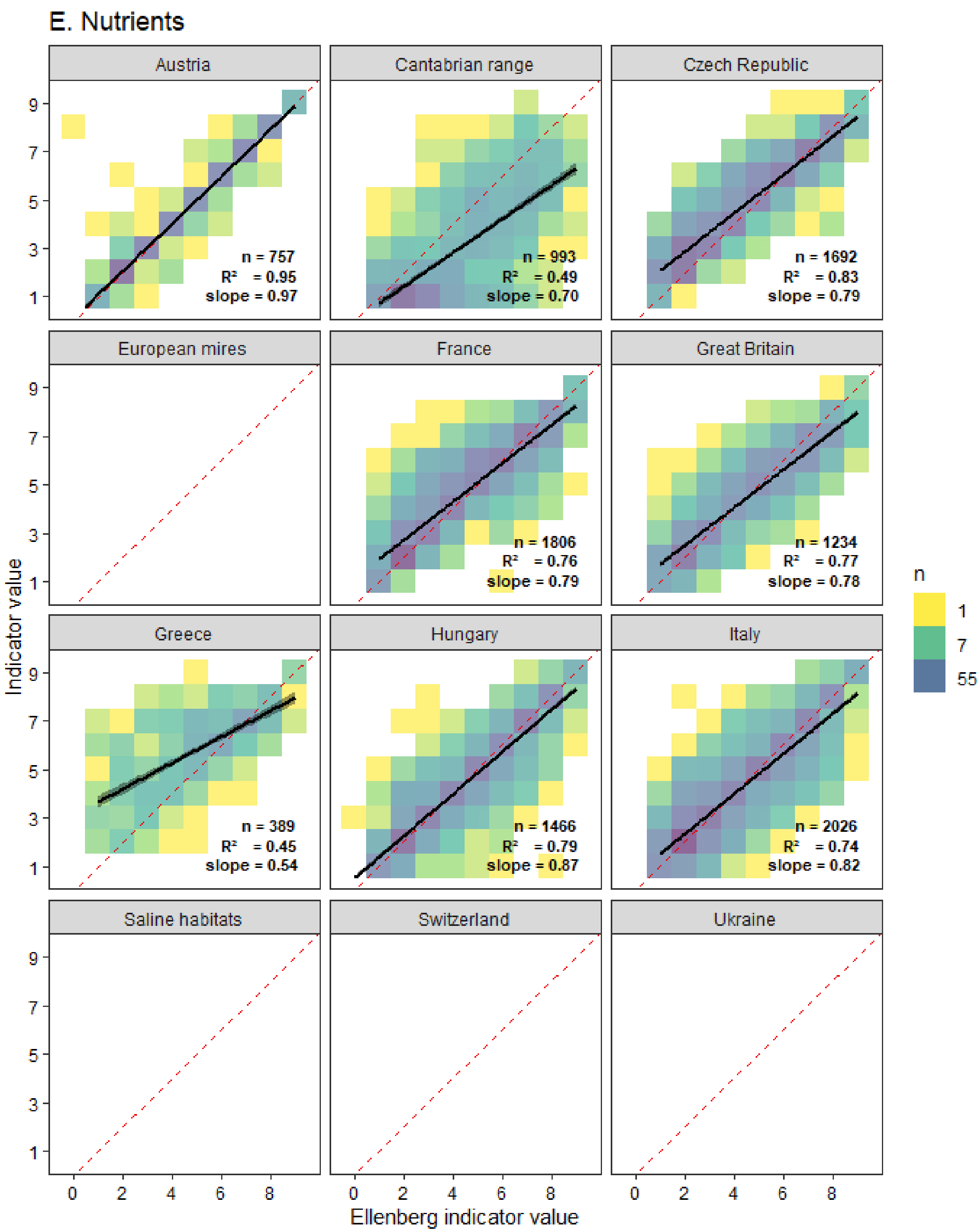
737

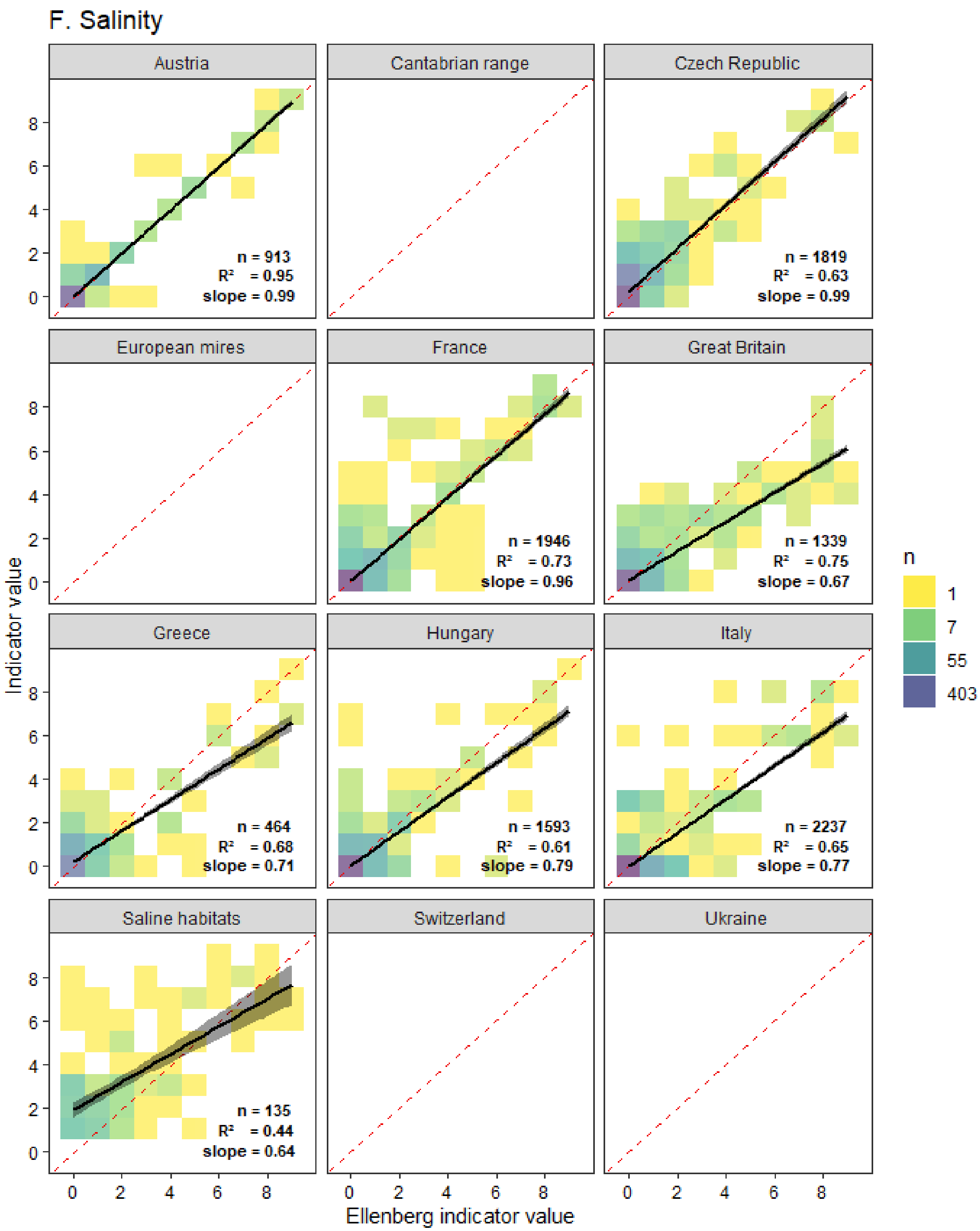

B. Temperature
Austria
n = 633
R² = 0.97
slope = 0.99
Cantabrian range
n = 855
R² = 0.26
slope = 0.65
Czech Republic
n = 1538
R² = 0.83
slope = 0.89
European mires
France
n = 1671
R² = 0.70
slope = 0.85
Great Britain
Greece
n = 371
R² = 0.08
slope = 0.26
Hungary
n = 1331
R² = 0.64
slope = 0.88
Italy
n = 1871
R² = 0.71
slope = 0.90
Saline habitats
Switzerland
n = 2036
R² = 0.79
slope = 1.06
Ukraine
n = 1433
R² = 0.47
slope = 0.43
Indicator value
Ellenberg indicator value
n
1
7
55
403


C. Moisture
Austria
Cantabrian range
Czech Republic
European mires
France
Great Britain
Greece
Hungary
Italy
Saline habitats
Switzerland
Ukraine
n = 840
R² = 0.98
slope = 0.99
n = 1012
R² = 0.63
slope = 1.44
n = 1740
R² = 0.94
slope = 0.96
n = 1047
R² = 0.91
slope = 0.88
n = 1838
R² = 0.87
slope = 0.83
n = 1266
R² = 0.90
slope = 0.92
n = 422
R² = 0.77
slope = 0.83
n = 1513
R² = 0.88
slope = 0.99
n = 2136
R² = 0.85
slope = 0.97
n = 1645
R² = 0.83
slope = 0.59
Indicator value
Ellenberg indicator value
n
1
7
55


D. Reaction
Austria
Cantabrian range
Czech Republic
European mires
France
Great Britain
Greece
Hungary
Italy
Saline habitats
Switzerland
Ukraine
n = 591
R² = 0.92
slope = 0.97
n = 879
R² = 0.40
slope = 0.64
n = 1536
R² = 0.82
slope = 0.81
n = 1633
R² = 0.75
slope = 0.80
n = 1086
R² = 0.69
slope = 0.67
n = 347
R² = 0.23
slope = 0.26
n = 1323
R² = 0.65
slope = 0.71
n = 1854
R² = 0.70
slope = 0.80
Indicator value
Ellenberg indicator value
n
1
7
55

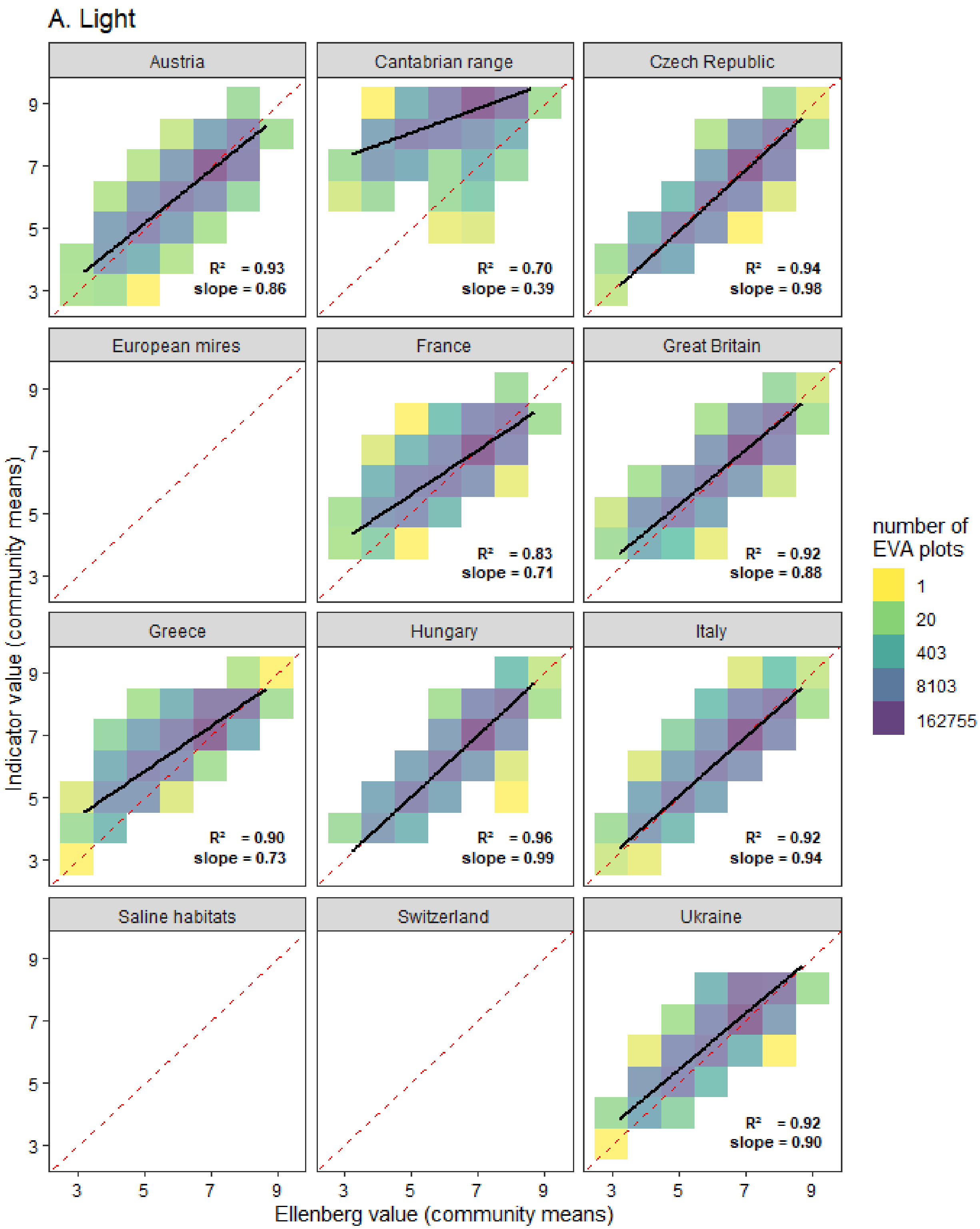
740


E. Nutrients
Austria
n = 757
R² = 0.95
slope = 0.97
Cantabrian range
n = 993
R² = 0.49
slope = 0.70
Czech Republic
n = 1692
R² = 0.83
slope = 0.79
European mires
France
n = 1806
R² = 0.76
slope = 0.79
Great Britain
n = 1234
R² = 0.77
slope = 0.78
Greece
n = 389
R² = 0.45
slope = 0.54
Hungary
n = 1466
R² = 0.79
slope = 0.87
Italy
n = 2026
R² = 0.74
slope = 0.82
Saline habitats
Switzerland
Ukraine
Indicator value
Ellenberg indicator value
n
1
7
55

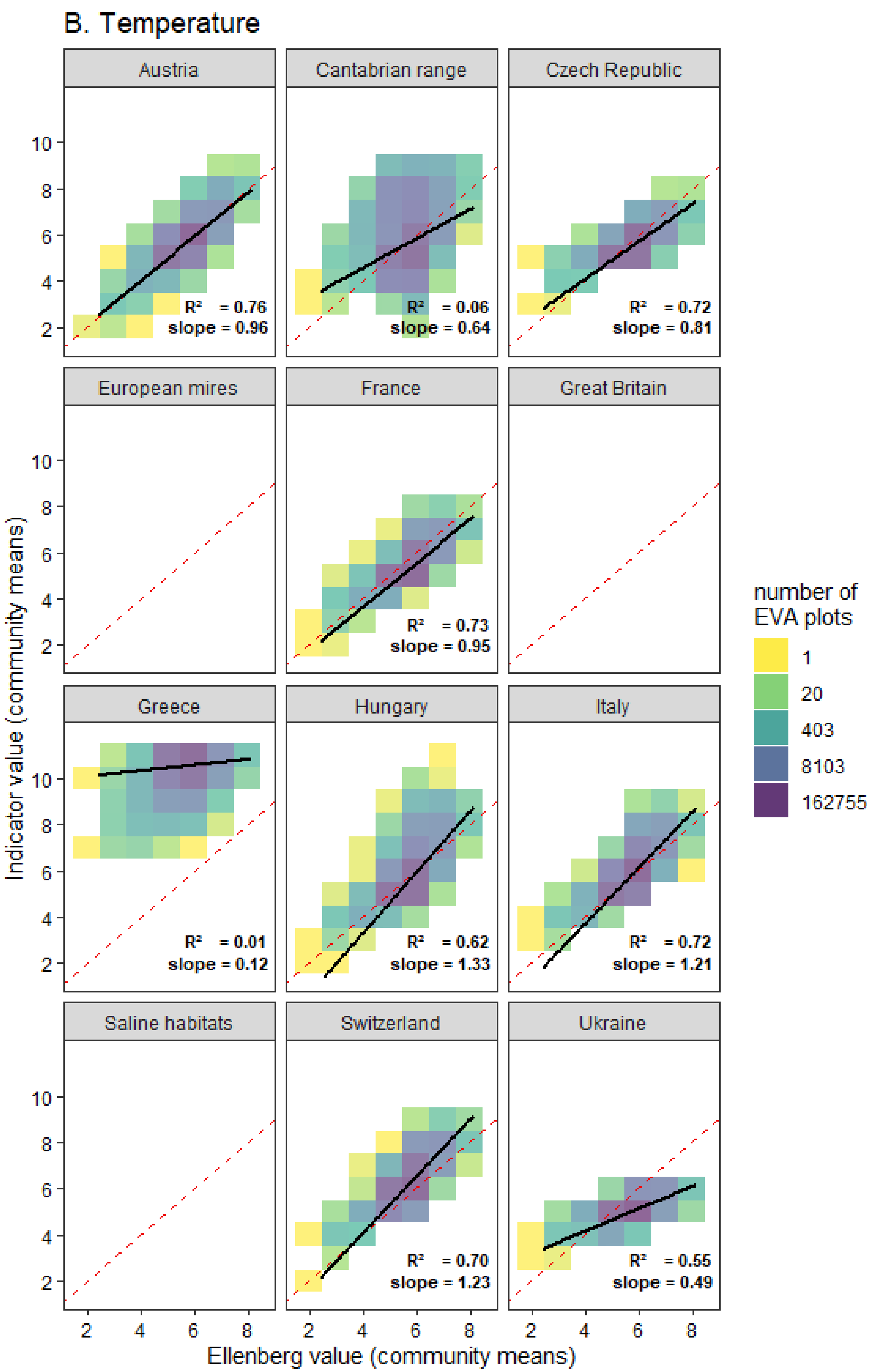

F. Salinity
Austria
n = 913
R² = 0.95
slope = 0.99
Cantabrian range
Czech Republic
n = 1819
R² = 0.63
slope = 0.99
European mires
France
n = 1946
R² = 0.73
slope = 0.96
Great Britain
n = 1339
R² = 0.75
slope = 0.67
Greece
n = 464
R² = 0.68
slope = 0.71
Hungary
n = 1593
R² = 0.61
slope = 0.79
Italy
n = 2237
R² = 0.65
slope = 0.77
Saline habitats
n = 135
R² = 0.44
slope = 0.64
Switzerland
Ukraine
Indicator value
Ellenberg indicator value
n
1
7
55
403

B. Indirect comparisons based on mean indicator values for vegetation plots

Fig. 2A–F: The relationships between the unweighted community means for light, temperature, moisture, reaction, nutrients and salinity calculated for individual vegetation plots using the different regional systems of indicator values and the original Ellenberg indicator values. The dataset includes 622,402 plots for light, 413,832 plots for temperature, 615,301 plots for moisture, 490,617 plots for reaction, 575,406 plots for nutrients and 673,141 plots for salinity from the EVA database; calculations were performed only for plots containing at least five species with the respective indicator value.

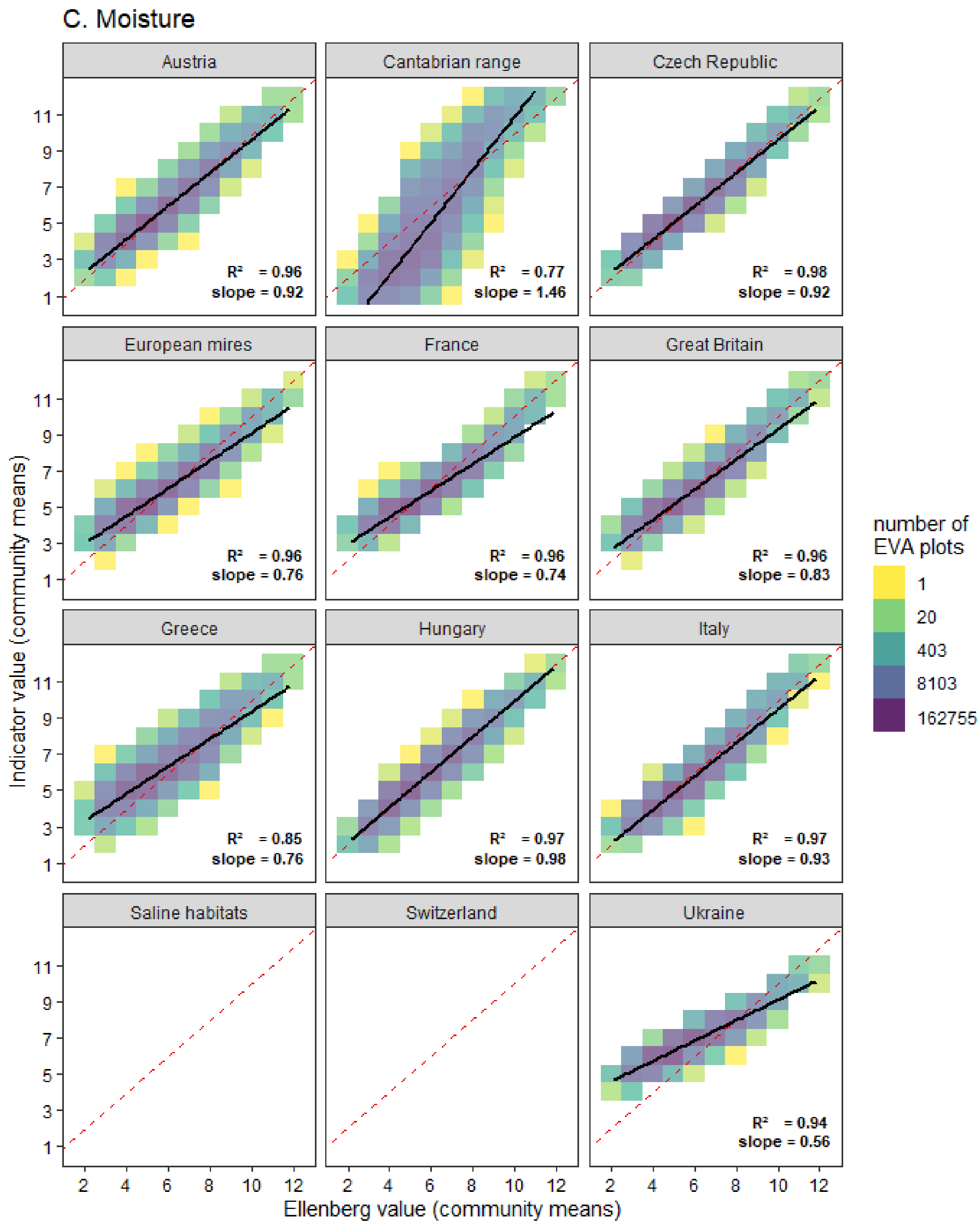

A. Light
Austria
Cantabrian range
Czech Republic
European mires
France
Great Britain
Greece
Hungary
Italy
Saline habitats
Switzerland
Ukraine
R² = 0.93
slope = 0.86
R² = 0.70
slope = 0.39
R² = 0.94
slope = 0.98
R² = 0.83
slope = 0.71
R² = 0.92
slope = 0.88
R² = 0.90
slope = 0.73
R² = 0.96
slope = 0.99
R² = 0.92
slope = 0.94
R² = 0.92
slope = 0.90
Indicator value (community means)
Ellenberg value (community means)
number of EVA plots
1
20
403
8103
162755

750

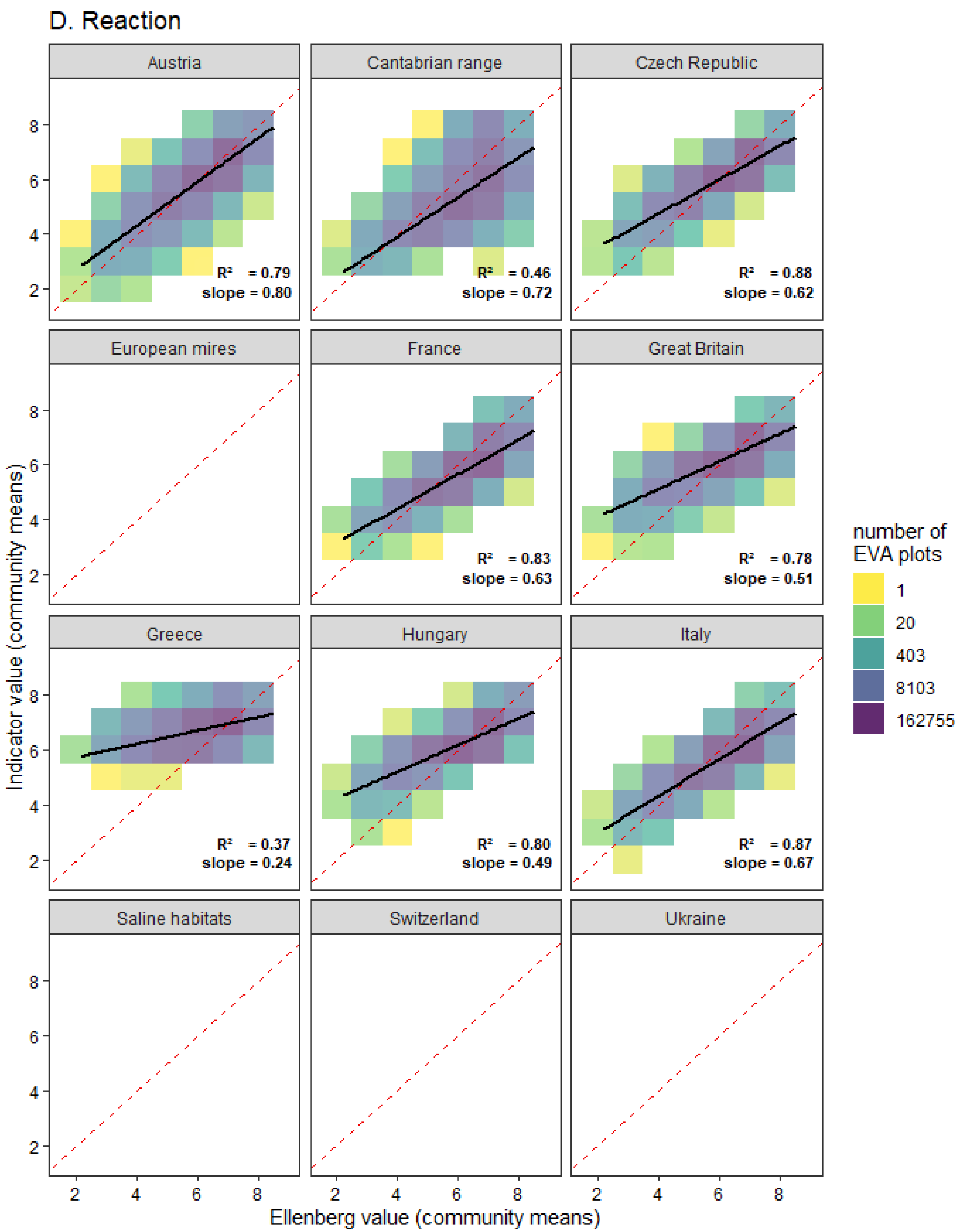


B. Temperature
Austria
Cantabrian range
Czech Republic
European mires
France
Great Britain
Greece
Hungary
Italy
Saline habitats
Switzerland
Ukraine
R² = 0.76
slope = 0.96
R² = 0.06
slope = 0.64
R² = 0.72
slope = 0.81
R² = 0.73
slope = 0.95
R² = 0.01
slope = 0.12
R² = 0.62
slope = 1.33
R² = 0.72
slope = 1.21
R² = 0.70
slope = 1.23
R² = 0.55
slope = 0.49
Indicator value (community means)
Ellenberg value (community means)
number of
EVA plots
1
20
403
8103
162755

751


C. Moisture
Austria
Cantabrian range
Czech Republic
European mires
France
Great Britain
Greece
Hungary
Italy
Saline habitats
Switzerland
Ukraine
R² = 0.96
slope = 0.92
R² = 0.77
slope = 1.46
R² = 0.98
slope = 0.92
R² = 0.96
slope = 0.76
R² = 0.96
slope = 0.74
R² = 0.96
slope = 0.83
R² = 0.85
slope = 0.76
R² = 0.97
slope = 0.98
R² = 0.97
slope = 0.93
R² = 0.94
slope = 0.56
Indicator value (community means)
Ellenberg value (community means)
number of EVA plots
1
20
403
8103
162755

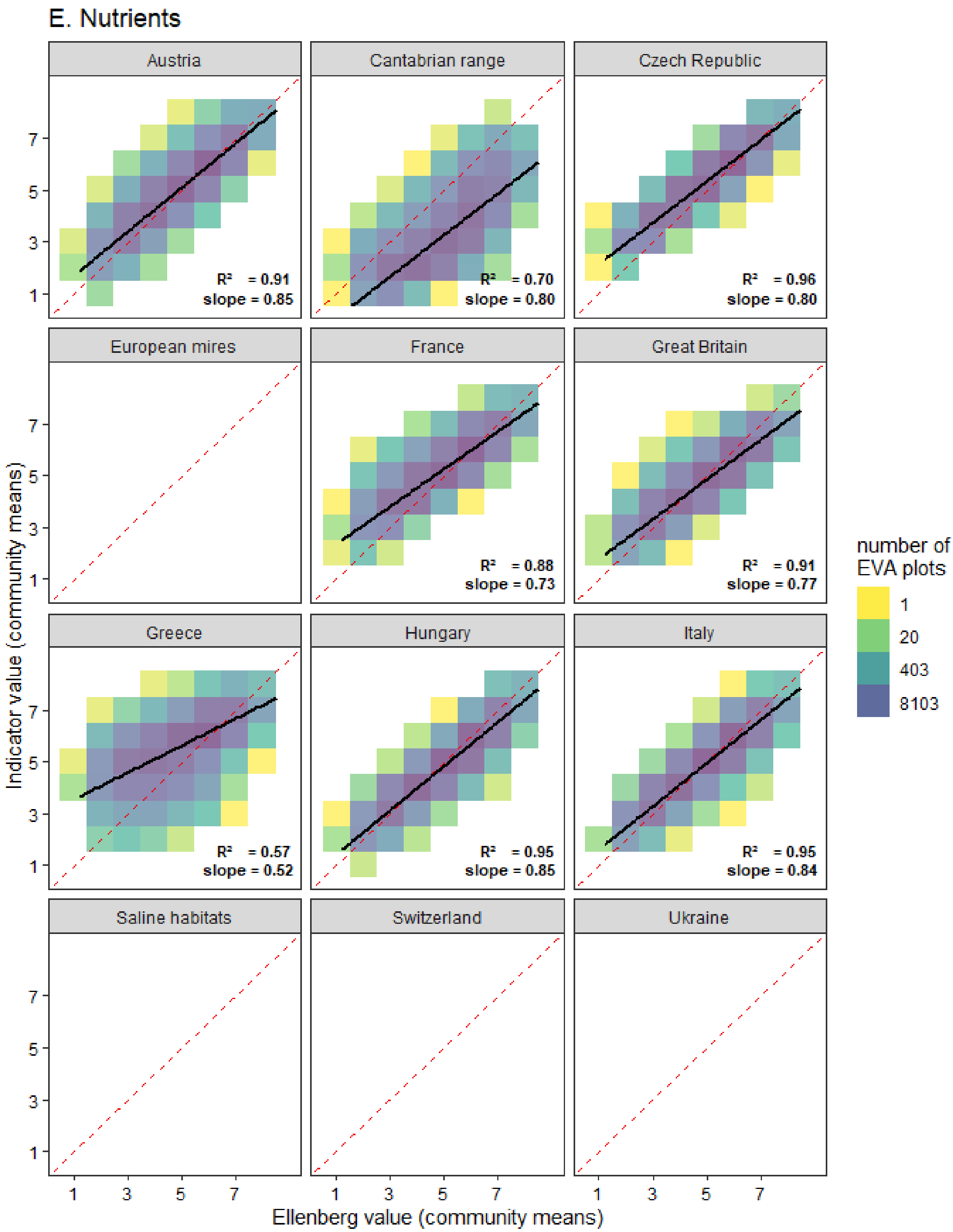
752



D. Reaction
Austria
Cantabrian range
Czech Republic
European mires
France
Great Britain
Greece
Hungary
Italy
Saline habitats
Switzerland
Ukraine
R² = 0.79
slope = 0.80
R² = 0.46
slope = 0.72
R² = 0.88
slope = 0.62
R² = 0.83
slope = 0.63
R² = 0.78
slope = 0.51
R² = 0.37
slope = 0.24
R² = 0.80
slope = 0.49
R² = 0.87
slope = 0.67
Indicator value (community means)
Ellenberg value (community means)
number of EVA plots
1
20
403
8103
162755

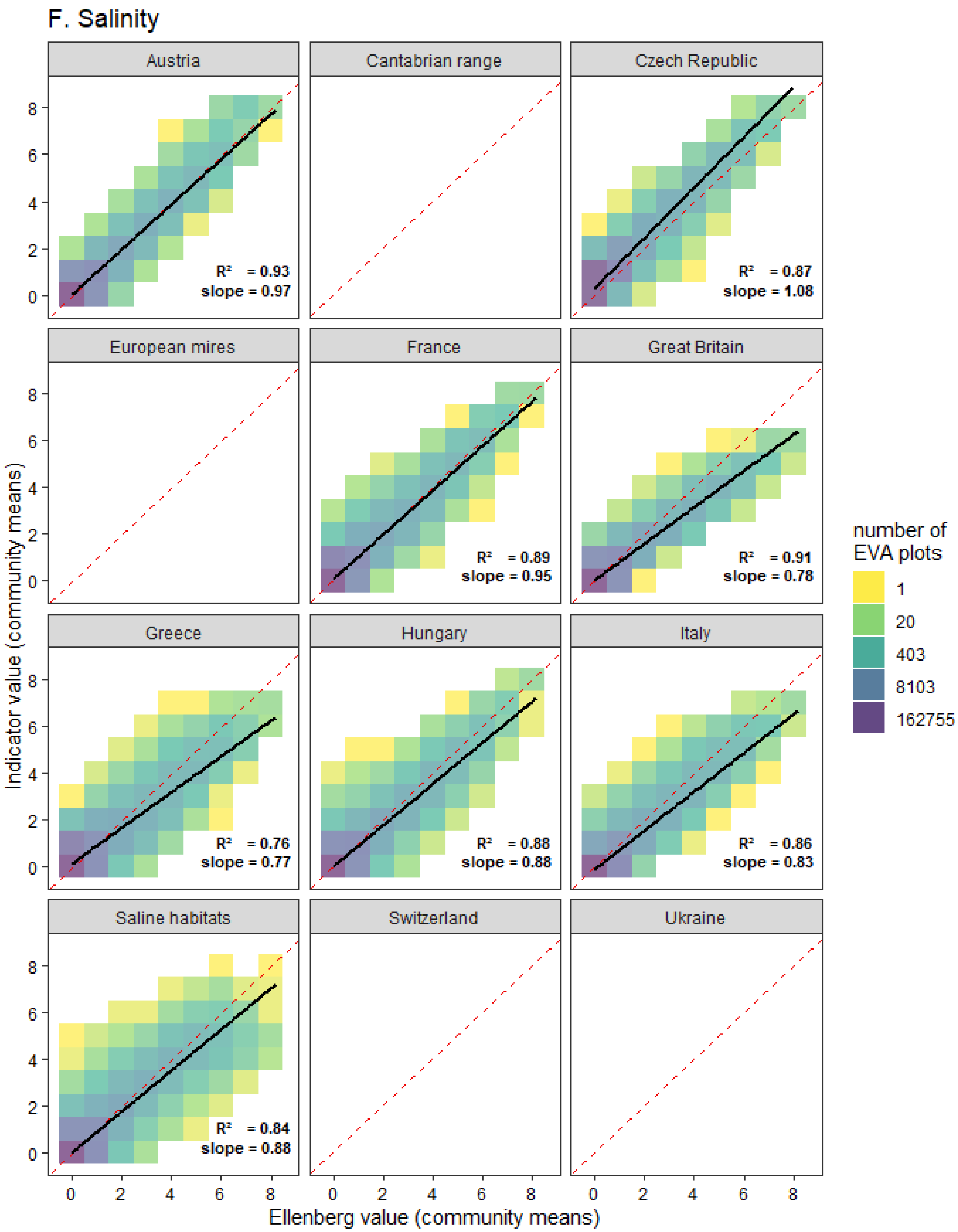
753

754



E. Nutrients
Austria
R² = 0.91
slope = 0.85
Cantabrian range
R² = 0.70
slope = 0.80
Czech Republic
R² = 0.96
slope = 0.80
European mires
France
R² = 0.88
slope = 0.73
Great Britain
R² = 0.91
slope = 0.77
Greece
R² = 0.57
slope = 0.52
Hungary
R² = 0.95
slope = 0.85
Italy
R² = 0.95
slope = 0.84
Saline habitats
Switzerland
Ukraine
Indicator value (community means)
Ellenberg value (community means)
number of EVA plots
1
20
403
8103

755



F. Salinity
Austria
Cantabrian range
Czech Republic
European mires
France
Great Britain
Greece
Hungary
Italy
Saline habitats
Switzerland
Ukraine
R² = 0.93
slope = 0.97
R² = 0.87
slope = 1.08
R² = 0.89
slope = 0.95
R² = 0.91
slope = 0.78
R² = 0.76
slope = 0.77
R² = 0.88
slope = 0.88
R² = 0.86
slope = 0.83
R² = 0.84
slope = 0.88
Indicator value (community means)
Ellenberg value (community means)
number of
EVA plots
1
20
403
8103
162755

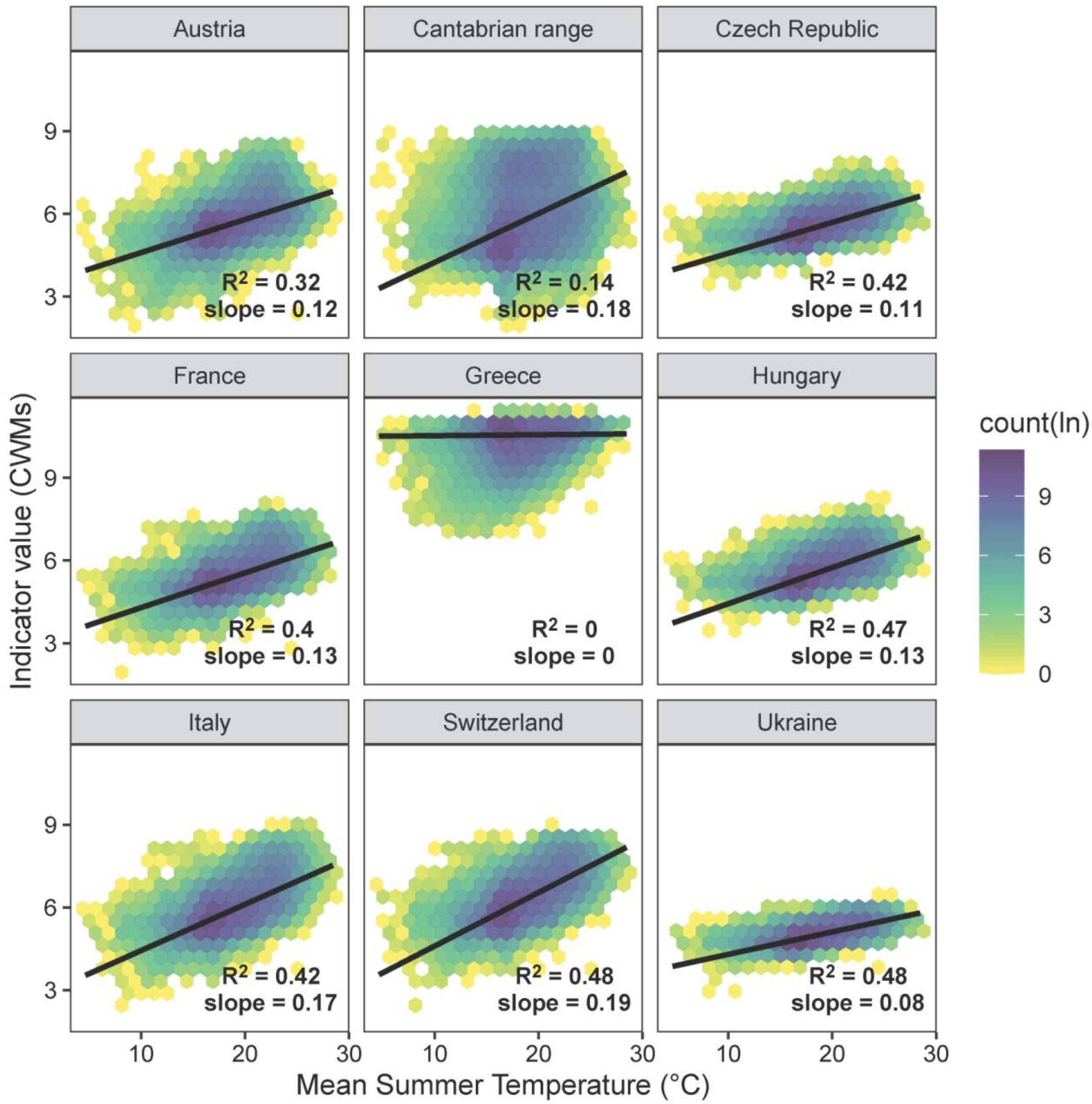
Supporting information of the paper

Tichý, L. et al. (2022) Ellenberg-type indicator values for European vascular plant species. *Journal of Vegetation Science.*

Appendix S3. Comparison of mean Ellenberg-type indicator values for temperature calculated for vegetation plots and mean summer temperature for plot locations obtained from climatic datasets.

A. Source datasets.

Fig. 1: Comparisons are based on 364,104 georeferenced plots from the EVA database (subset of the dataset selected for comparison of unweighted community means for temperature in Appendix S2B) that contain at least five species with the Temperature indicator value.



B. Harmonized Ellenberg-type indicator values for Europe.

Fig. 2: The dataset included 364,104 georeferenced plots and 8,146 species from the EVA database similarly as in Appendix S2. (IV) Unweighted community means for temperature were calculated using 5,196 species with temperature indicator value in at least one regional dataset. (IV.full) Unweighted community means for temperature were calculated using 5,553 species with temperature indicator value

defined in at least one regional dataset or estimated using a similar distribution for species for which indicator value was not defined in any regional dataset.