Supplementary Materials for

European plants lagging behind climate change pay a climatic debt in the North, but are favored in the South

François Duchenne, Gabrielle Martin & Emmanuelle Porcher

Correspondence to: francois.duchenne@mnhn.fr

This PDF file includes:

Other Supplementary Materials for this manuscript include the following:

Supplementary Methods:

DOIs

For computational reasons we split the extractions from the GBIF into 10 parts, accessible through the 10 followings DOI:

GBIF.org (01 May 2020) GBIF Occurrence Download https://doi.org/10.15468/dl.z8kcad

GBIF.org (01 May 2020) GBIF Occurrence Download https://doi.org/10.15468/dl.fyakne

GBIF.org (01 May 2020) GBIF Occurrence Download https://doi.org/10.15468/dl.7u7dgx

GBIF.org (01 May 2020) GBIF Occurrence Download https://doi.org/10.15468/dl.ske4qr

GBIF.org (01 May 2020) GBIF Occurrence Download https://doi.org/10.15468/dl.bhfw37

GBIF.org (01 May 2020) GBIF Occurrence Download https://doi.org/10.15468/dl.zy2nhr

GBIF.org (01 May 2020) GBIF Occurrence Download https://doi.org/10.15468/dl.67dgze

GBIF.org (30 April 2020) GBIF Occurrence Download https://doi.org/10.15468/dl.3b8bc8

GBIF.org (30 April 2020) GBIF Occurrence Download https://doi.org/10.15468/dl.p5awfb

GBIF.org (30 April 2020) GBIF Occurrence Download https://doi.org/10.15468/dl.kjv523

Plant species selection

We removed crop and ornamental species from this list, using an FAO reference list for crop species (http://www.fao.org/fileadmin/templates/ess/documents/world_census_of_agriculture/appendix4_r7.pdf) complemented by visual inspection of the list. We also considered invasive species separately; they were identified using the European Union list for invasive species (https://ec.europa.eu/environment/nature/invasivealien/index_en.htm), complemented by visual inspection. By definition, invasive species have increased in abundance/occupancy in the

recent past; this allowed us to check that occupancy trends as calculated below were able to detect these increases (Fig. S3).

To homogenize taxonomy among these lists and our dataset, we used the GBIF backbone taxonomy, which is designed to avoid synonymy issues among datasets, all along the paper.

Potential drivers of occupancy trends

For each species, we calculated historical climatic indices for the six bioclimatic variables studied here. To do so, we averaged SCIs over 1951-1980 (i.e. stopping before the recent sharp temperature increase, Fig. S1a) for each species, weighting each year by the number of records of the given species.

For nutrient (reflecting mainly nitrophily) and moisture preferences we used the Ellenberg Indicator Values (EIV) from France (Julve 1998), United-Kingdom (Fitter & Peat 1994), Italy (Pignatti *et al.* 2005), Czech Republic (Chytrý *et al.* 2018) and Germany (Ellenberg *et al.* 1992). All EIVs are on the same scale, but they are a relative measure of species preference, depending on the species assemblages used. As these species assemblages depend on the source, a given value can reflect distinct nutrient/moisture preferences among sources. Using species shared across EIV sources, we corrected biases among data sources, using the following formula:

$$EIV_{x_s}{}' = EIV_{x_s} + \frac{\sum_{j=1}^{n} EIV_{Baseflor_j} - EIV_{x_j}}{n} \quad (5)$$

where EIV_{x_s} is the Ellenberg indicator value of species *s* in source *x* and *n* the number of species shared between the source *x* and Baseflor, which was used as a reference because it contains information for the largest number of species. Finally, for species present in several data sources we used the average corrected EIVs, over all sources. The whole procedure was performed independently for moisture and nutrient preferences.

Pollen vector (i.e. insects, wind or self-fertilization) was retrieved from the Baseflor (Julve 1998), Ecoflora (Fitter & Peat 1994) and BiolFlor (Kühn *et al.* 2004) databases. Many species are associated with several types of pollen vectors, both within and among databases. We encapsulated this variability into a single variable: pollinator dependency, the percentage of times "insects" appear as a pollen vector for a given species, across all databases. Pollinator dependency ranges from 0, for species that are never associated with insect pollination in the trait databases and that should be therefore independent of pollinators for their reproduction, to 100, for species that are only associated with insect pollination, and that should be strictly dependent on pollinators for their reproduction.

The lifespan of each species was extracted using the R package *TR8* (Gionata 2015) from BiolFlor and LEDA (Kleyer *et al.* 2008) databases and coded following Martin et al.'s (2019) categories but with three levels only: strict annual plants, intermediate plants (biennial, annual/perennial, etc.) and strict perennial plants.

Habitat affinity was calculated following the same principles as for SCIs, but averaged over the whole time period (1951-2014). We used the EUNIS habitat classification (Davies *et al.* 2004) at the first level, but merging all aquatic, wetland and coastal habitats together (Table S2). As for SCIs and occupancy calculation we used a 10km×10km grid cell. In general, the finest spatial resolution is best to calculate species habitat affinity or detect changes in occupancy or species climatic indices. Hence, we chose to limit spatial aggregation as much as possible, but we were constrained first by computation times and second, more importantly, by the spatial resolution of GBIF data: we discarded a spatial aggregation at 1km² because GBIF datasets are often defined at 5 or 10km². For each 10km×10km grid cell, we calculated the percentage of area covered by each habitat. Then, for each species and each habitat, we calculated the weighted mean of the habitat coverage over the range of each plant species, weighting the contribution of each cell (10×10km²) by the ratio of the number of records of this

species on the number for records for all plant species. For each species, we therefore obtained 7 habitat affinity indices, each ranging from 0 to 1 and summing to 1 across habitats. They correspond to the fraction of a given habitat in the species distribution.

Species traits, SCI and occupancy trends are available in Table S1.

Phylogenetic signal in SCI and occupancy trends

Estimating phylogenetic signal in species response informs us on plausible evolutionary constraints on mechanisms underlying SCI and occupancy trends. To analyze the phylogenetic structure of SCI and occupancy trends, we used the *Daphne* phylogeny of European flora (Durka & Michalski 2012). Of the 4,120 species we analyzed, 1,335 were not included in the phylogeny, thus we excluded them for phylogenetic analysis, and focused on the 2,785 remaining species. We assessed the phylogenetic structure using Pagel's λ, implemented in the *phylosignal* R package (Keck *et al.* 2016), and tested its significance by randomizing the tips of the phylogeny 1,000 times, for both SCI and occupancy trends.

Evidencing the role of the recent climatic change in observed effects

To confirm that the costs of the shifts in experienced climatic conditions occur only after the acceleration of climate change (1980-2000, Fig. S5), we performed the same set of analyses on the earliest data, taking the first 40 years (1951-1990), i.e. including 1980-1990 to retain enough records and years to estimate SCI and occupancy trends. We calculated SCI trends and occupancy trends between 1951 and 1990. For SCI trends, we used only records from 1951 to 1990. For occupancy trends, we used all records and the same model as in equation (2) but with a broken-line model for the year effect, as implemented in the *segmented* R package (Muggeo 2008), with a breakpoint in 1990. Such method enables estimation of random site effects and effect of the species list length on the entire dataset while modelling a trend for 1951-1990. Results are shown in Fig. S6.

References

Chytrý, M., Tichý, L., Dřevojan, P., Sádlo, J. & Zelený, D. (2018). Ellenberg-type indicator values for the Czech flora. *Preslia*, 90, 83–103.
Davies, C.E., Moss, D. & Hill, M.O. (2004). EUNIS habitat classification revised 2004, 310.
Durka, W. & Michalski, S.G. (2012). Daphne: a dated phylogeny of a large European flora for phylogenetically informed ecological analyses. *Ecology*, 93, 2297–2297.
Ellenberg, H., Weber, H.E., Düll, R., Wirth, V., Werner, W., Paulißen, D., *et al.* (1992). Zeigerwerte von pflanzen in Mitteleuropa.
Fitter, A.H. & Peat, H.J. (1994). The Ecological Flora Database. *J. Ecol.*, 82, 415–425.
Gionata, B. (2015). TR8: an R package for easily retrieving plant species traits. *Methods Ecol. Evol.*, 6, 347–350.
Julve, P. (1998). Baseflor. Index botanique, écologique et chorologique de la flore de France. *Inst. Cathol. Lille Lille*.
Keck, F., Rimet, F., Bouchez, A. & Franc, A. (2016). phylosignal: an R package to measure, test, and explore the phylogenetic signal. *Ecol. Evol.*, 6, 2774–2780.
Kleyer, M., Bekker, R.M., Knevel, I.C., Bakker, J.P., Thompson, K., Sonnenschein, M., *et al.* (2008). The LEDA Traitbase: a database of life-history traits of the Northwest European flora. *J. Ecol.*, 96, 1266–1274.
Kühn, I., Durka, W. & Klotz, S. (2004). BiolFlor — a new plant-trait database as a tool for plant invasion ecology. *Divers. Distrib.*, 10, 363–365.
Muggeo, V.M.R. (2008). segmented: an R Package to Fit Regression Models with Broken-Line Relationships. *R News*, 8, 20–25.
Pignatti, S., Menegoni, P. & Pietrosanti, S. (2005). Indicazione attraverso le piante vascolari. Valori di indicazione secondo Ellenberg (Zeigerwerte) per le specie della Flora d'Italia. *Braun-Blanquetia*, 3, 91–97.

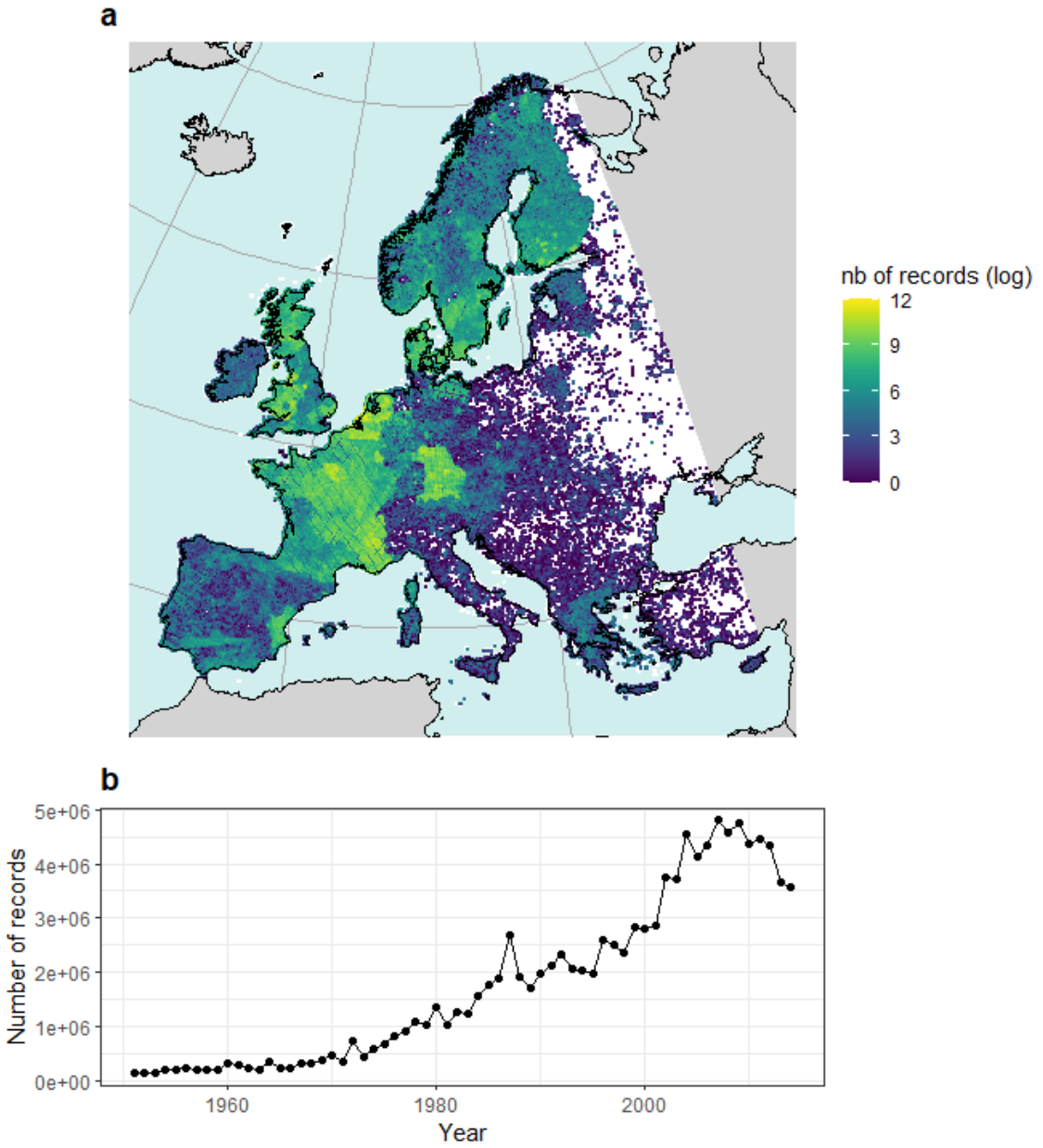


Figure S1: ***Spatial and temporal distribution of the records used.*** *(a) Spatial distribution of the records used, grouped by ~100km² grid cells. White grid cells correspond to cells with no data and grey cells are outside of the study area. Hexagonal patterns in France are due to the main data source from France (CBN dataset: https://www.gbif.org/fr/dataset/75956ee6-1a2b-4fa3-b3e8-ccda64ce6c2d). This dataset was aggregated spatially using WSG84 coordinates, leading to grid cells with heterogeneous area on a map projection. Here we re-projected the entire dataset using grid cells with the same area across Europe. The superimposition of two grids that are not orthogonal, because of distinct projections, led to such periodic patterns. (b) Number of annual records through time, showing a continuous geometric increase over years in the number of data points.*

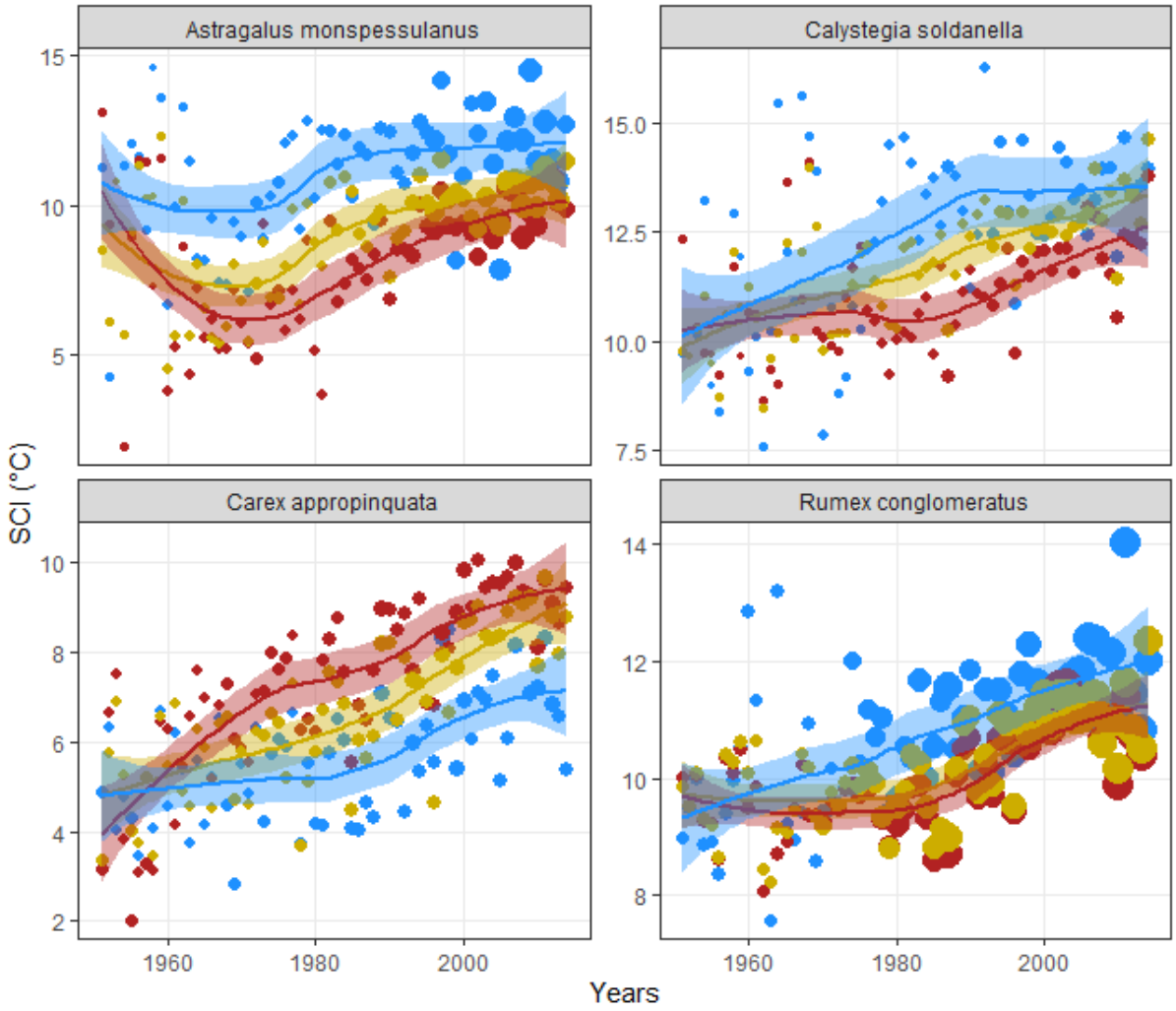


Figure S2: *Examples of interannual variability in SCIs, using annual mean temperature. SCIs represented by blue points are those used in the paper. They are calculated with a mean weighted by the ratio* $\frac{N_{ijk}}{\sum_{s=1}^{nsp} N_{sjk}}$ *where* N_{ijk} *is the number of record of species i, grid cell k and year j.* $\sum_{s=1}^{nsp} N_{sjk}$ *is the total number of records over all plant species for a given year and grid cell. SCIs represented by yellow points are calculated with a mean weighted by* N_{ijk}*, which accounts for sampling pressure (represented by* $\sum_{s=1}^{nsp} N_{sjk}$*). The red curve is the mean weighted by* $\begin{matrix} 1 \; if \; N_{ijk} > 0 \\ 0 \; if \; N_{ijk} = 0 \end{matrix} \times \sum_{s=1}^{nsp} N_{sjk}$*, which represents variation in temperature but also temporal bias in the sampling pressure over the geographic range of species i. Circle size is proportional to the number of grid cells included in the (weighted) mean. Curves are the results of locally estimated scatterplot smoothing regressions implemented in ggplot2.*

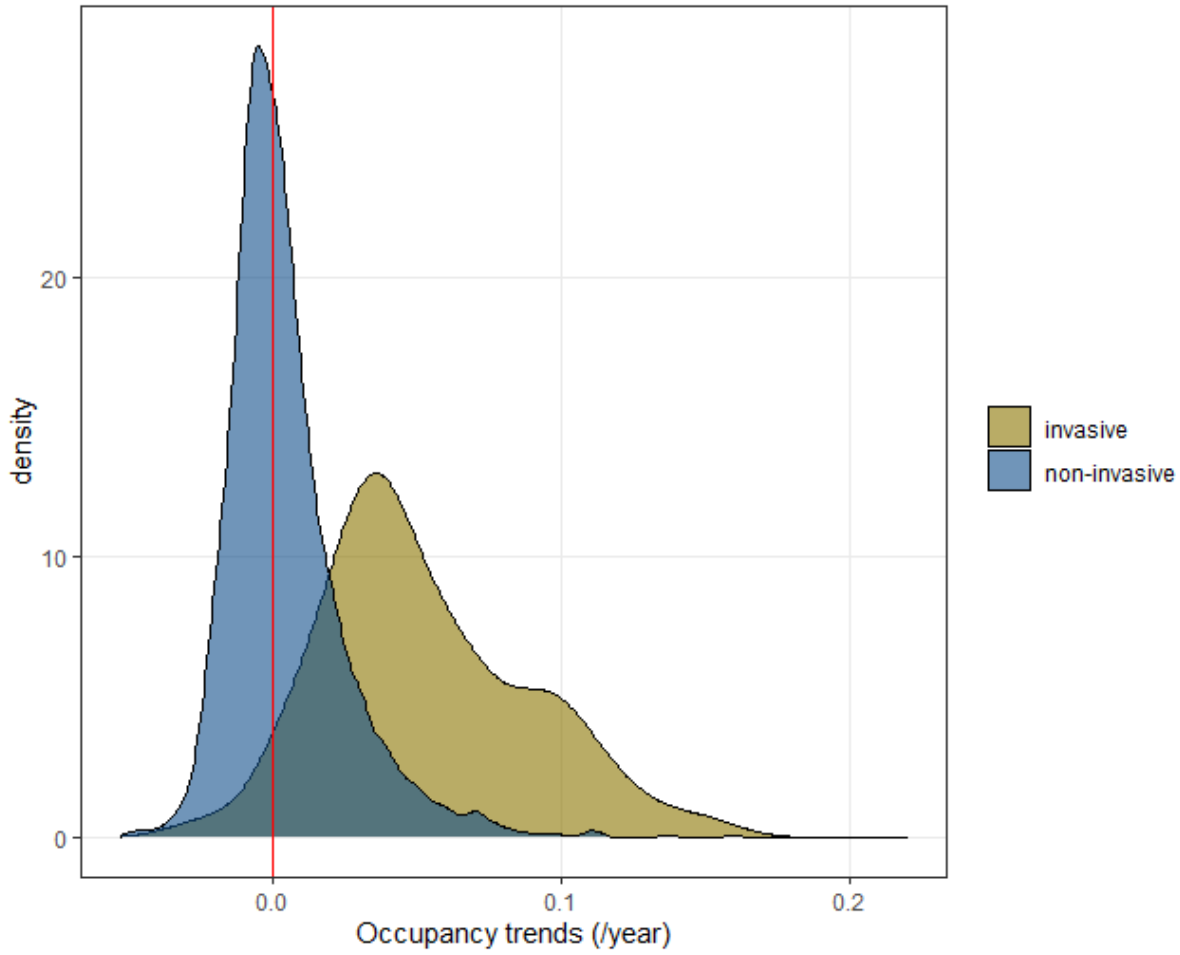
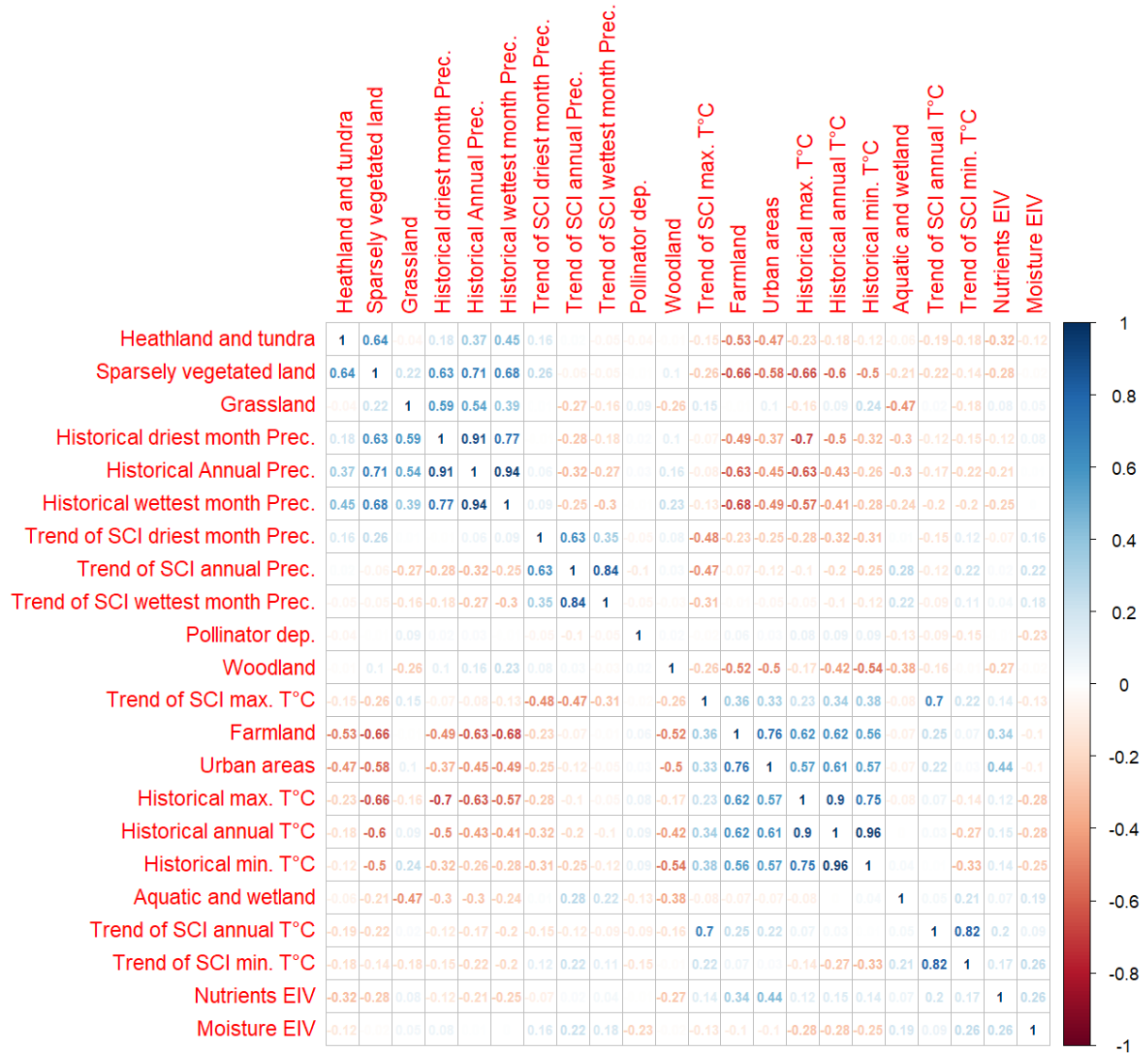


Figure S3: ***Occupancy trends of non-invasive vs. invasive species****. Density distribution of occupancy trends for the 4,120 native or naturalized ("non-invasive") species and for the 58 species that were identified as "invasive" in the species selection. The red vertical line indicates zero.*



174 ***Figure S4:* Correlation matrix among variables potentially added to phylogenetic regression**
175 ***and linear mixed-effect models.*** *The red to blue color ramp represents the sign and strength*
176 *of the correlation. Variables are ordered in the matrix so that highly correlated variables are*
177 *clustered.*

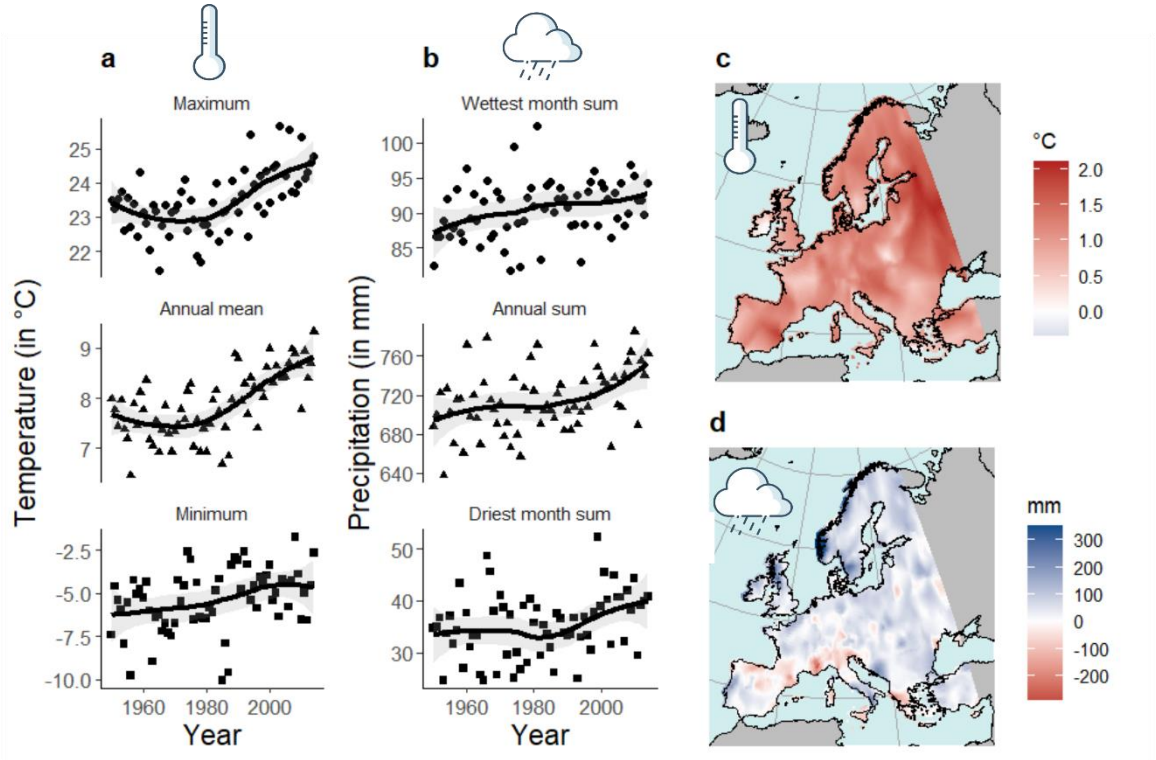


Figure S5: ***Change in bioclimatic variables across time and space****. (a) and (b) show temporal variations in temperature and precipitation respectively, averaged over the study area. Circles depict maximum temperature or precipitation of the wettest month (bio5 & 13), triangles annual mean temperature or annual precipitation (bio1 & 12), and squares minimum temperature or precipitation of the driest month (bio6 & 14). The black lines correspond to LOESS (Locally Estimated Scatterplot Smoothing) curves obtained from the ggplot2 library in R. (c) and (d) illustrate the spatial variation in temporal changes of annual mean temperature and precipitation respectively. Temporal changes were measured here as the difference between the 1950-1960 average and the 2004-2014 average. In (d) the upper bound of the color scale is truncated to 350mm instead of 575mm to preserve readability.*

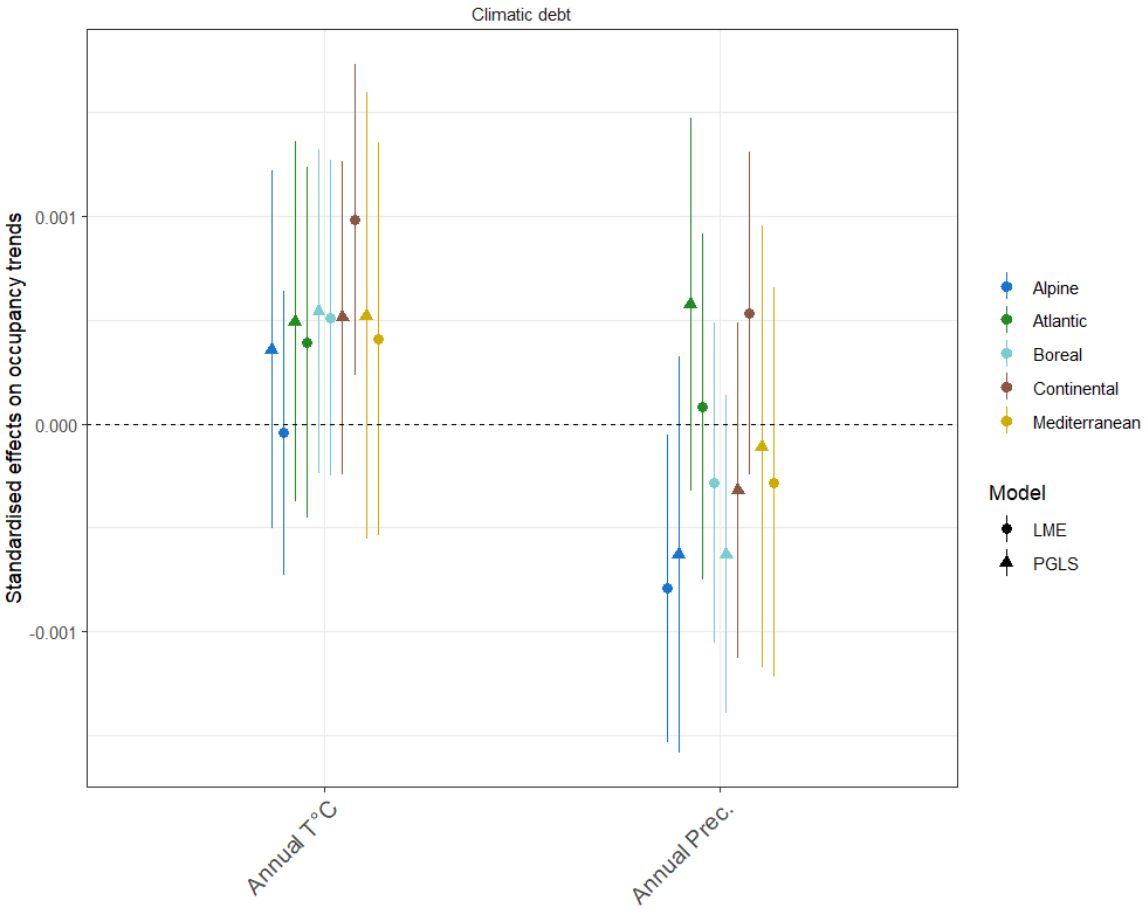


Figure S6: ***Correlations between occupancy trends and SCI trends from 1951 through 1990.*** *Estimates (±$CI_{95\%}$) from phylogenetic regressions (PGLS) and linear mixed-effect models (LME) explaining occupancy trends with temporal trends in SCIs and other species traits.*

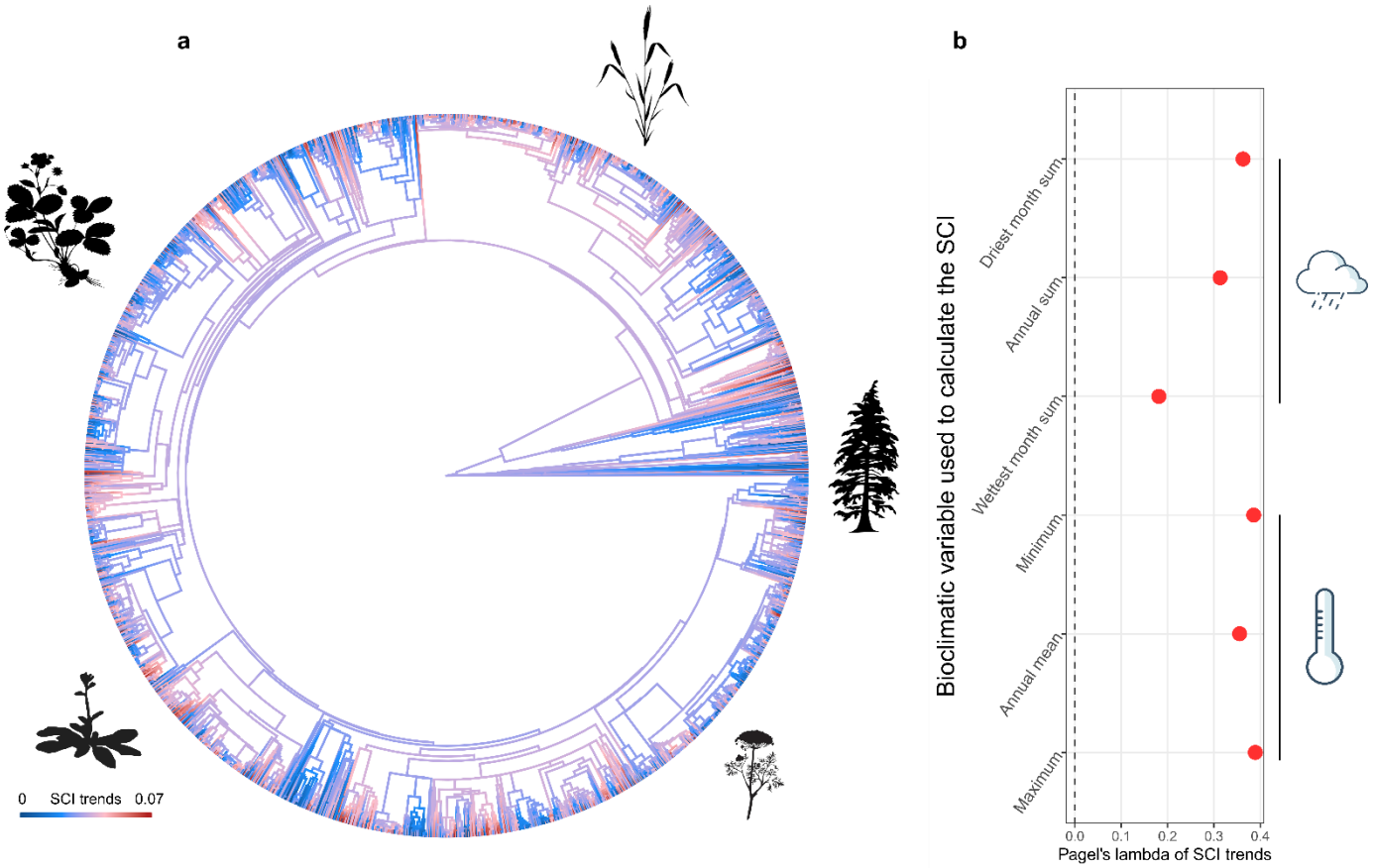
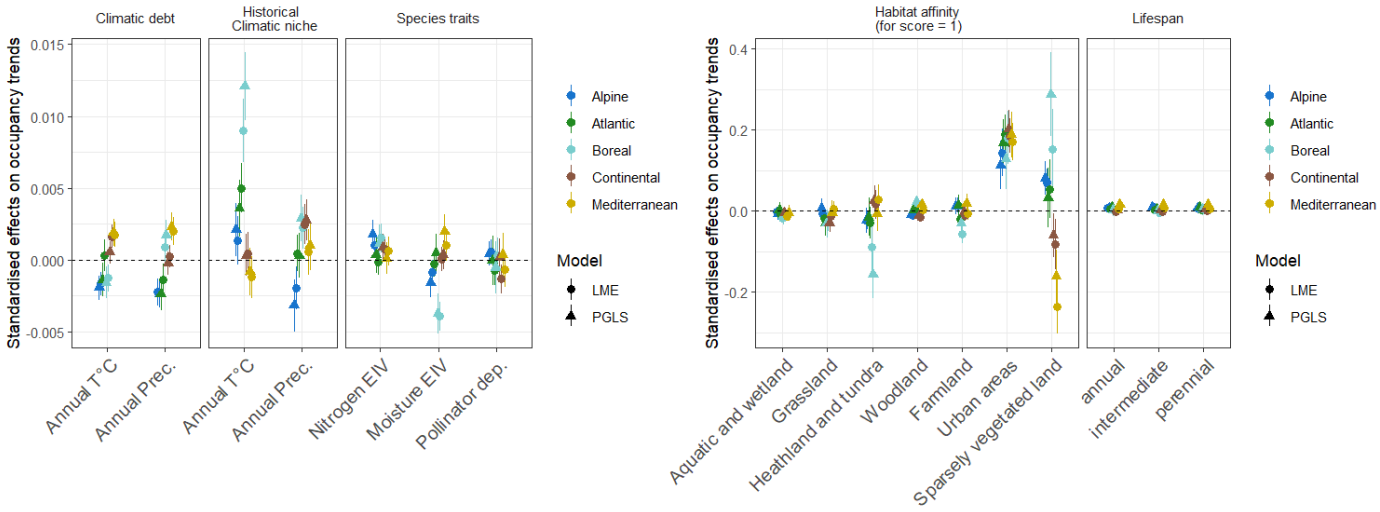


Figure S7: ***Phylogenetic signal in the linear trends in species climatic indices (SCIs) for the 2,785 species included in the phylogeny***. *(a) Phylogenetic signal in the SCI trends related to the annual mean temperature bioclimatic variable (bio1). The color scale is bounded between the 5th and 95th quantile to preserve readability. (b) Pagel's λ for SCI trends related to the six bioclimatic variables. Zero (dashed black line) indicates an absence of phylogenetic signal. Red circles correspond to a significant phylogenetic signal (p-value < 0.05, calculated from 1000 randomizations).*



***<u>Figure S8:</u>* Correlates of occupancy trends for each biogeographic region.** *The three left panels represent the estimates (±$CI_{95\%}$) from phylogenetic regression (PGLS) and linear mixed-effects model (LME), explaining occupancy trends with temporal trends in species climatic indices (SCIs) and other species traits. The two right panels show predicted averaged occupancy trends (±$CI_{95\%}$) for each habitat, considering a theoretical perfect affinity to each habitat (score = 1 & lifespan = annual), and for lifespan categories, predicted at the average of all other variables.*

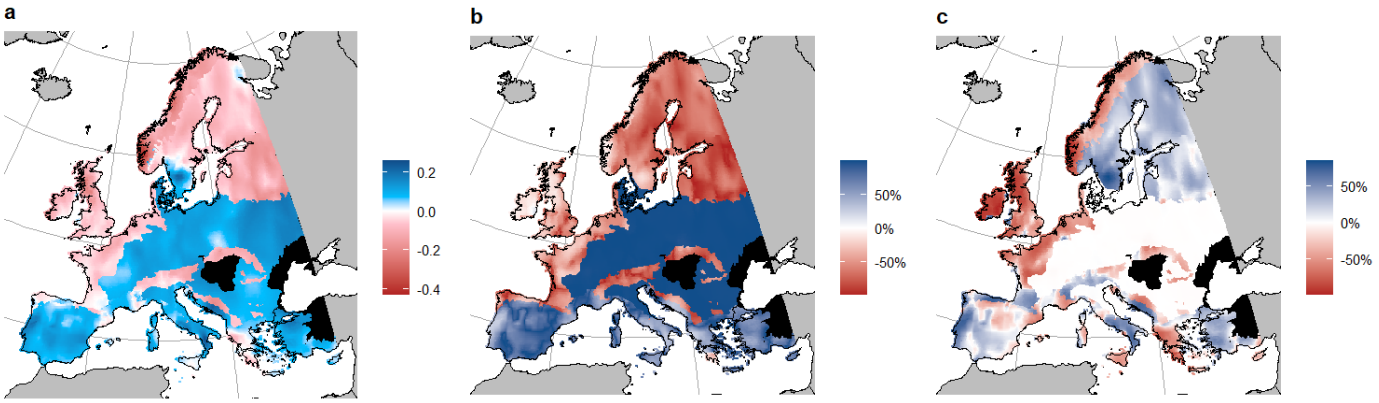


Figure S9: ***Climatic debt/bonus in Europe and its climatic drivers, considering non-significant predictors.*** *Same figure as Figure 4 of the paper but also including the effects of trends in species climatic indices (SCIs) that are not significant. (a) Climatic debt/bonus averaged over all species over the last 65 years. The gradient from white to red indicates a climatic debt (cost of climate change in terms of species occupancy), while the gradient from white to blue indicates a climatic bonus (benefits of climate change in terms of species occupancy); white represents no cost on average for plants. Relative contribution of (b) temperature and (c) precipitation SCI trends to the climatic debt, in percentage. Black regions are biogeographic regions with too few data. The maps were generated using predictions averaged over the linear mixed-effects model (LME) and the phylogenetic regression (PGLS).*

Table S1 (separate file):
SCI, occupancy trends and species traits for the 4,120 species studied + the 58 invasive species present in the initial species list.

Table S2: EUNIS habitat classification, and the grouped habitat classification used in the study.

EUNIS categories	Our categories
A - Marine habitats	Aquatic and wetland
B - Coastal habitats	Aquatic and wetland
C - Inland surface waters	Aquatic and wetland
D - Mires, bogs and fens	Aquatic and wetland
E - Grasslands and land dominated by forbs, mosses or lichens	Grassland
F - Heathland, scrub and tundra	Heathland and tundra
G - Woodland, forest and other wooded land	Woodland
H - Inland unvegetated or sparsely vegetated habitats	Sparsely vegetated land
I - Arable land and market gardens	Farmland
J - constructed, industrial and other artificial habitats	Urban areas

Table S3: Mean and associated $CI_{95\%}$ for SCI trends over the 4,120 studied species.

Variable	upper CI95%	Mean	lower CI95%
Annual mean Temperature (bio1)	0.03290744	0.032226788	0.0315461347
Maximum Temperature (bio5)	0.04568614	0.044922632	0.0441591201
Minimum Temperature (bio6)	0.03366092	0.032643692	0.0316264601
Annual Precipitation (bio12)	0.27474310	0.202730377	0.1307176590
Wettest month Precipitation (bio13)	0.01677690	0.008501243	0.0002255894
Driest month Precipitation (bio14)	0.12556774	0.118927077	0.1122864159