

ECOGRAPHY

Research

Global distribution and bioclimatic characterization of alpine biomes

Riccardo Testolin, Fabio Attorre and Borja Jiménez-Alfaro

R. Testolin (<https://orcid.org/0000-0002-8916-7231>) ✉ (riccardo.testolin@uniroma1.it) and F. Attorre, Dept of Environmental Biology, Sapienza Univ. of Rome, Rome, RM, Italy. – B. Jiménez-Alfaro, Research Unit of Biodiversity (CSIC/UO/PA), Univ. of Oviedo, Mieres, Principado de Asturias, Spain.

Ecography

43: 1–10, 2020

doi: 10.1111/ecog.05012

Subject Editor: Wolf Eiserhardt

Editor-in-Chief: Miguel Araújo

Accepted 3 February 2020



Although there is a general consensus on the distribution and ecological features of terrestrial biomes, the allocation of alpine ecosystems in the global biogeographic system is still unclear. Here, we delineate a global map of alpine areas above the treeline by modelling regional treeline elevation at 30 m resolution, using global forest cover data and quantile regression. We then used global datasets to 1) assess the climatic characteristics of alpine ecosystems using principal component analysis, 2) define bioclimatic groups by an optimized cluster analysis and 3) evaluate patterns of primary productivity based on the normalized difference vegetation index. As defined here, alpine biomes cover 3.56 Mkm² or 2.64% of land outside Antarctica. Despite temperature differences across latitude, these ecosystems converge below a sharp threshold of 5.9°C and towards the colder end of the global climatic space. Below that temperature threshold, alpine ecosystems are influenced by a latitudinal gradient of mean annual temperature and they are climatically differentiated by seasonality and continentality. This gradient delineates a climatic envelope of global alpine biomes around temperate, boreal and tundra biomes as defined in Whittaker's scheme. Although alpine biomes are similarly dominated by poorly vegetated areas, world ecoregions show strong differences in the productivity of their alpine belt irrespectively of major climate zones. These results suggest that vegetation structure and function of alpine ecosystems are driven by regional and local contingencies in addition to macroclimatic factors.

Keywords: alpine biomes, bioclimatic characterization, global treeline, Google Earth Engine, NDVI, primary productivity

Introduction

The knowledge of the extent and climatic characteristics of terrestrial biomes, further linked to their functional aspects (e.g. productivity), is key to understanding ecological and biogeographical phenomena (Mucina 2019). Among terrestrial environments, alpine ecosystems (i.e. high-elevation habitats above the climatic treeline) are the only biogeographic unit represented across all continents and latitudes (Körner 2003); they are characterized by a varied history of climatic changes and strong microhabitat differentiation (Körner 2003); they also contain global biodiversity hotspots – e.g. the



www.ecography.org

© 2020 The Authors. Ecography published by John Wiley & Sons Ltd on behalf of Nordic Society Oikos
This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

tropical Andes (Myers et al. 2000) – and support about 10 000 plant species as a whole, many of which are endemics (Körner 2003). Alpine ecosystems supply fresh water to more than half of the human population (Pomeroy 2015) and may stock up to 1% of the global terrestrial carbon pool (Körner 1995); they are also home to the habitats most threatened by land use (Nagy and Grabherr 2009) and anthropogenic climate change (Hughes 2000).

Despite the relevance of alpine ecosystems to global biodiversity, their biogeographic delineation in the biome system is still unclear. The definition of world biomes has been traditionally based on vegetation physiognomy and macroclimate for grouping areas with similar dominant ecosystems (Mucina 2019). Alpine ecosystems, which are characterized by the absence of trees in response to low temperatures, form a continuum of shrubland and grassland habitats occurring above the climatic treeline across latitudes. First attempts of characterizing alpine ecosystems were based on their relationships with pre-defined biomes. For example, the temperature-based system of Holdridge (1947) linked the alpine and nival altitudinal belts to the frigid and polar zones, encompassing all that is widely known as arctic and alpine tundra. In his influential characterization of world biomes, Whittaker (1975) acknowledged the differences between arctic and alpine biome types by identifying two temperature-related ecoclines (i.e. latitudinal and altitudinal) involved in the transition from forest to treeless tundra. Similarly, Walter's classification of terrestrial ecosystems (Walter and Box 1976) separated latitudinal zonobiomes from alpine orobiomes. In a different biogeographical context, Olson et al. (2001) identified the 'Montane Grasslands and Shrublands' habitat type encompassing tropical and subtropical mountain ranges. Yet, this classification is not based on vegetation patterns and left out temperate and boreal alpine regions of the northern hemisphere, that were either included in forest biomes or arctic tundra. Similarly, Faber-Langendoen et al. (2016) defined 'Tropical Montane and High Montane Grasslands and Shrublands' in a recent classification of world vegetation formations, separating these areas from the non-tropical alpine tundra.

Despite large divergences in the interpretation of alpine ecosystems, an empirical characterization of the ecological or functional properties of the global alpine belt and its relationships with terrestrial biomes is missing. Most studies are limited to continental (Körner et al. 2003) or regional scales (Noroozi and Körner 2018) while current estimates of the global alpine area are either based on average regional treeline elevations and expert evaluation (Körner 2003) or coarse resolution delineation of altitudinal belts (Körner et al. 2011). This knowledge gap has so far hindered any comparative analysis of alpine ecosystems in relation to their biogeographical patterns, despite their well-known similarities in dominant vegetation (Körner 2003). Recent developments in publicly accessible cloud computing platforms like Google Earth Engine (Gorelick et al. 2017) and the increasing availability of large-scale datasets are

facilitating the exploration of natural patterns at the global scale (Hansen et al. 2013, Bastin et al. 2017, 2019), allowing us to answer long standing questions about the biogeography of alpine ecosystems.

Here, we developed a method for estimating the global extent of alpine areas based on empirical data and a high-resolution map of their distribution. We then characterized alpine ecosystems to assess their bioclimatic and productivity patterns through global climatic variables and the normalized difference vegetation index (NDVI) from global to regional scales. Our first aim was to re-evaluate prevailing questions about the distribution of alpine ecosystems, such as 1) what is the spatial distribution and extent of alpine ecosystems worldwide? 2) How are alpine ecosystems related to major lowland biomes as for their bioclimatic and productivity patterns? By producing the first empirical dataset of global alpine biomes, our second aim was to advance their comparative ecology and to provide a spatial tool that will assist future mountain research across disciplines.

Material and methods

Study area

Our study focuses on global mountain regions with an alpine zone, i.e. a vegetation belt above the climatic treeline (Körner et al. 2011). Arctic or subarctic regions dominated by treeless vegetation (Baffin Island, Greenland, Iceland, Svalbard and Novaya Zemlya) were excluded because the arctic tundra, although analogous to the alpine zone, is not defined by elevational gradients (Quinn 2008) and represents a different zonobiome (Walter and Box 1976). We used a global inventory of mountain areas based on topographic ruggedness (Körner et al. 2017) and a raster of climatic belts (Körner et al. 2011) to obtain a preliminary GIS layer of mountain polygons that contained the 'upper montane' belt and at least one pixel of the 'alpine' belt. The workflow is illustrated in Supplementary material Appendix 1 Fig. A1. This step excluded mountain polygons where the alpine belt is absent or scarcely represented, which was necessary to optimize our workflow. Nevertheless, we verified that our study area encompasses the majority of known alpine areas in all continents. Our working dataset included 345 mountain regions covering nearly 11 Mkm² representing mid- to high-altitude mountain areas worldwide, thus ranging from mountain forests to the highest unforested summits.

Identification of alpine areas

We used the Google Earth Engine computing platform (Gorelick et al. 2017) to select alpine areas within the mountain polygons. First, we deleted forest areas as they were mapped at the global scale using satellite images from the year 2015 at 30 m spatial resolution (Hansen et al. 2013) (Supplementary material Appendix 1 Fig. A1). All pixels with forest cover >

0% were removed. As the resulting mask contained many scattered unforested pixels, we applied a low pass filter to reduce high frequency information by performing a linear convolution using a 11×11 pixels moving window. We then upsampled the image to 50m spatial resolution, to reduce its size while keeping most of the detail, and considerably speed up the following operations. The resulting raster, representing unforested mountain areas, was vectorized. At this spatial resolution, we were not able to isolate any unforested area in Mont Cameroun (Cameroon), Virunga Mountains (Democratic Republic of the Congo, Rwanda and Uganda), Hidaka-sanmyaku (Japan) and Kaimanawa Mountains (New Zealand) (Supplementary material Appendix 1 Fig. A2), despite being included in the initial dataset. Thus, the following operations have been carried out on a set of 341 mountain regions. We also removed unforested area polygons smaller than 5 km^2 to avoid the inclusion of many scattered patches of alpine and sub-alpine habitats in large mountain ranges that would have considerably increased computation time, with the only exception of Mount Meru (Tanzania) and Ruwenzori (Democratic Republic of the Congo, Uganda), whose unforested areas were kept despite all being smaller than 5 km^2 .

At this point the dataset consisted of unforested areas within and above the upper montane belt that may include, besides the alpine zone delimited by the climatic treeline, other mountain areas where the original forest was suppressed either by anthropogenic disturbance, local environmental or topographic conditions (Holtmeier 2009). To retain alpine areas only, we modeled regional treeline elevation with linear quantile regression using R software (R Core Team) and the ‘quantreg’ package (Koenker 2018), based on equally spaced points sampled every 5km along the unforested polygon boundaries using QGIS 2.18 (QGIS Development Team 2016) (Supplementary material Appendix 1 Fig. A1). For each point we extracted elevation and northness (cosine of aspect) from the SRTM-3 global digital elevation model (DEM) (Farr et al. 2007, NASA and JPL 2013). In mountains located above 60° latitude, thus not covered by the SRTM-3 DEM, elevation and northness values were derived from the ASTER-2 global DEM (NASA/METI/AIST/Japan Space Systems and U.S./Japan ASTER Science Team 2009). For each mountain range we modeled the 99th percentile of the distribution of forest border elevation values measured at the point locations, controlling for northness. We opted for the 99th percentile by analogy with the concept of treeline, defined as the line that connects the highest patches of forest – in our case forest pixels – within a series of slopes of similar exposure (Körner 2003). In mountain ranges spanning more than 5 degrees in latitude, latitude was also included in the model. We chose 5 degrees as a reasonable interval at which latitudinal changes of treeline elevation likely override possible disturbance-induced treeline shifts. To avoid singularities and ensure model convergence, we added random noise constrained between -0.5 and 0.5 m to elevation values.

Our methodology relies on the assumption that, within each mountain region, the remaining traces of the climatic treeline can be used to model its elevation across the whole region. However, some arid mountain ranges naturally lack a treeline because low water availability prevents the establishment of trees regardless of temperature (Körner 2012). In our dataset, this was the case for some ranges in the driest parts of South America (21 ranges) and central Asia (34 ranges). To consistently identify the potential treeline elevation in these regions and ensure the continuity of alpine areas extent across adjacent mountains, we applied the treeline elevation quantile regression models derived from the closest neighboring mountain range, controlling for local northness and latitude. As an example, we used the model for Himalaya to estimate the treeline elevation of the surrounding treeless mountains by applying it to each DEM, accounting for the difference in latitude. Similarly, for the arid central Andean mountains, treeline elevation was estimated by applying a hybrid model fitted using the points sampled along the unforested areas polygons of the closest surrounding mountain ranges (Cordillera Oriental Peru Bolivia Chile and the Cordillera Frontal) thus assuming a linear decrease in treeline elevation between the two. A complete list of the treeless mountain ranges for which such procedure was applied, together with the corresponding neighboring ranges whose treeline model was used, is reported in Supplementary material Appendix 1 Table A1.

Finally, we extracted the area above the modeled treeline elevation within each mountain range, obtaining an estimate of the global extent of alpine ecosystems (Supplementary material Appendix 1 Fig. A1). Given the high spatial resolution of the obtained alpine layer, we again filtered out polygons smaller than 5 km^2 to reduce graininess and streamline further operations. To visualize the global patterns of the estimated treeline elevations, we calculated their mean for each mountain range by sampling equally spaced points every 20km along the alpine polygons’ boundaries and extracting their elevation in Google Earth Engine using the same DEMs described above. We applied a weighted loess fit with $\text{span}=0.4$ to the mean values of treeline elevation along latitudes to describe the global pattern and to allow visual comparison with the treeline model compiled by Körner (Körner 2003) using worldwide field observations. To account for the uneven latitudinal distribution of mountain ranges (i.e. nonequal variance of treeline elevation along latitude), a greater weight was assigned to the mountains at under-represented latitudes. All distances were calculated in equidistant cylindrical Plate Carrée projection, while areas were calculated in equal area pseudo-cylindrical Eckert IV projection.

Bioclimatic characterization

To outline the climatic space occupied by the mapped alpine ecosystems, we overlapped the values of mean annual temperature and annual precipitation with the widely recognized

Whittaker's biomes classification (Whittaker 1975), using the global climatic dataset CHELSA (Karger et al. 2017) at 30 arc-sec spatial resolution. At each pixel location, we also extracted elevation values from SRTM-3 and ASTER-2 global digital elevation models upscaled to 30 arc-sec resolution. To evaluate the climatic differences among different alpine ecosystems, we ran a principal component analysis (PCA) on 19 centered and scaled bioclimatic variables (Karger et al. 2017). The variables with the greatest factor loadings were used to interpret the PCA axes. To describe the climatic variation among global alpine ecosystems and delineate alpine regions of similar climatic conditions, we performed a cluster analysis based on the first four PCA axes, which captured almost 90% of the variance. We used Euclidean distances on PCA axes to overcome the strong multicollinearity of some of the original environmental variables and exploit their orthogonal properties (Weigelt et al. 2013). Prior to clustering, PCA axes were multiplied by the square root of their eigenvalues to weight their influence on the classification outcome according to their importance.

We employed the clustering large applications (CLARA) algorithm (Kaufman and Rousseeuw 1990), an extension of the k-medoids method for large datasets, using the clara function in the R package 'cluster' (Maechler et al. 2018). This method considers a subset of the data with fixed size and applies the k-medoids algorithm to generate an optimal set of medoids for the sample. The quality of the resulting medoids is measured by the average dissimilarity between every object in the entire data set and the medoid of its cluster. The sampling and clustering process are then repeated for a fixed number of subsets of the entire dataset and the final clustering results correspond to the set of medoids with the lowest average dissimilarity. This method requires a specified number of clusters (k) in advance. For this, we explored a limited number of clusters, from 2 to 10, to facilitate presentation and interpretation of results. Given the size of our dataset (almost 6 million records), we adopted a heuristic approach for the choice of the best k . We ran the clara function on 100 random subsets of 1000 cells for each k and replicated the process 100 times. Each time, the best k was based on the highest average silhouette width. Finally, among the runs with $k = \text{best } k$, we chose the one with the greatest average silhouette width.

To further assess the reliability of the k-medoids-based clustering, we performed a hierarchical clustering using the first four weighted PC axes of a subset of 30 000 records. To make sure that the sample captured most of the climatic variability of alpine ecosystems, we stratified it equally among three latitudinal belts in each hemisphere: tropical ($0-23.5^\circ$), temperate ($23.5-50^\circ$) and subpolar ($> 50^\circ$). The clustering was performed using the hclust function of the 'fastcluster' package (Müllner 2013), with the Ward2 clustering method. To assess the best number of k using this method, we repeated the process for 100 subsets and recorded the average silhouette width when cutting the tree from 2 to 10 clusters. Then, we cut the dendrogram in order to get $k = \text{best } k$. Finally, we

compared the two clustering outcomes using PC biplots and an alluvial plot.

Primary productivity

We estimated the primary productivity of alpine ecosystems based on the normalized difference vegetation index (NDVI), an indirect measure of vegetation cover and biomass related to the physical properties of plants (Cihlar et al. 1991, Pettorelli et al. 2005). Despite its limitations, the NDVI has been widely used as a proxy for ecosystem properties including global grassland productivity (Gao et al. 2016) and above ground biomass in the alpine belt (Liu et al. 2017). To minimize the problems related to single-date remote sensing studies of vegetation (Holben 1986), we calculated the maximum NDVI value at each pixel of 30×30 m using Landsat 8 Annual Greenest-Pixel images (from 2013 to 2018) in Google Earth Engine. The resulting NDVI values reflect the maximum productivity of each pixel during the growing season in recent times, independently of within- or between-year climatic variation. However, the length of the growing season could change at different latitudes and so would the total annual productivity, which remains undetectable using this methodology. Nevertheless, this allows us to interpret the relative proportion of ecosystems ranging from the smallest (rocky habitats) to the greatest (shrubby habitats) peak productivity across the study regions. The final composite was upscaled to 30 arc-sec resolution. We then removed artifact NDVI values that were negative or > 1 . To investigate the relationships between productivity, climate and dominant vegetation types, we compared the distributions of NDVI values among clusters and ecoregions (Olson et al. 2001) using kernel density plots. We also fitted a generalized additive model to NDVI data using the bam function in the R package 'mgcv' (Wood 2011), assuming a Gamma distributed conditional response, with smooth terms for the first four PC axes described above, as a proxy of the alpine climatic space. To reduce the effect of spatial autocorrelation, we sampled 100 000 random points and fit the model on this subset, including also a smooth term for geographic coordinates. To assess the influence of macroclimate in vegetated areas only, we also ran a model on a subset of 100 000 points with NDVI values > 0.1 .

Results

Despite the presence of a few outliers, treeline elevation shows an increasing trend toward the tropics and decreases again close to the equator, almost symmetrically in both hemispheres (Fig. 1b). Based on regional treeline models, we isolated alpine ecosystems above the climatic treeline worldwide (Fig. 1a) and estimated their extent to 3.56 Mkm^2 , corresponding to 2.64% of total land area outside Antarctica. Asia hosts almost three fourths of the global alpine area with 2.59 Mkm^2 , followed by South America (15%; 0.55 Mkm^2),

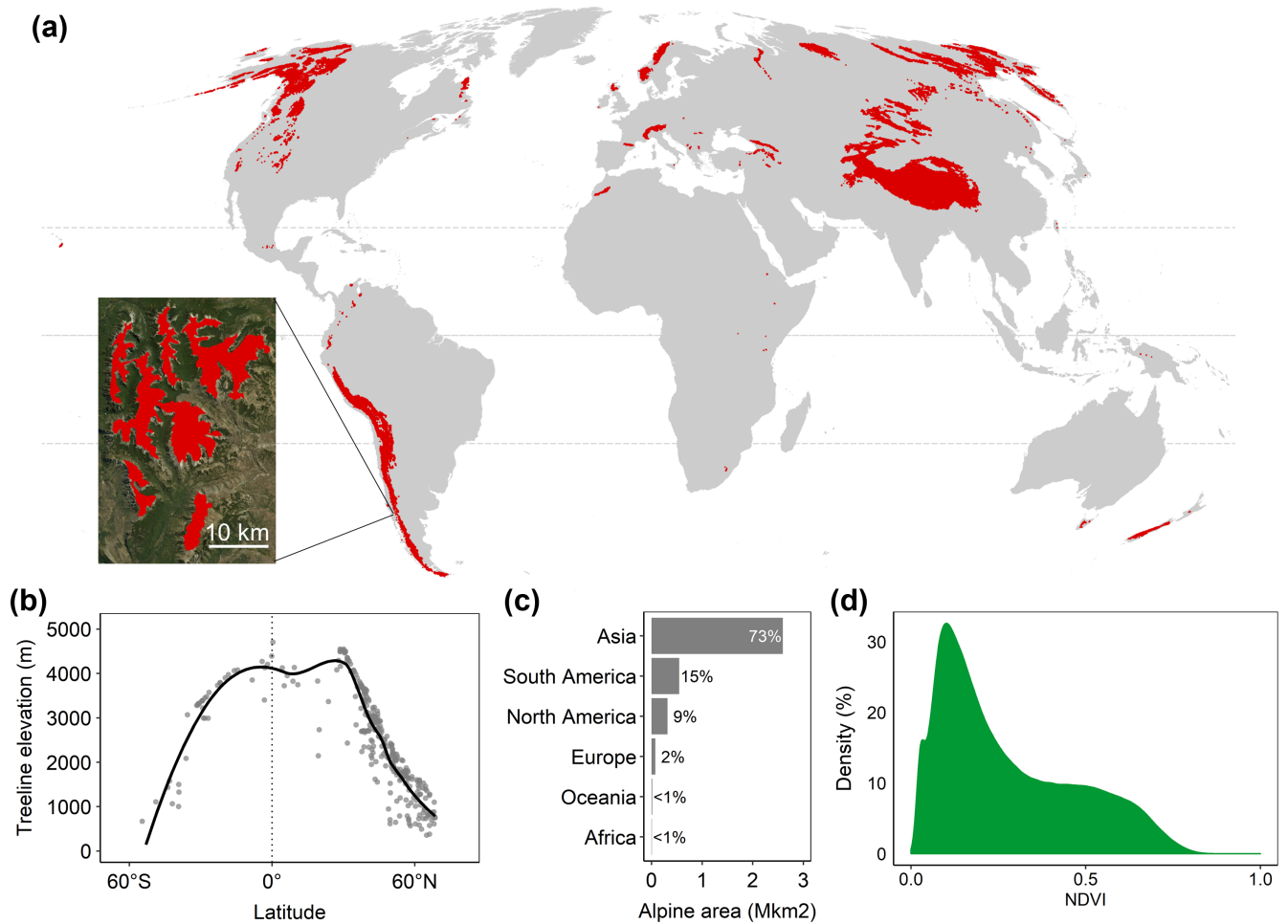


Figure 1. Extent and distribution of global alpine areas. (a) Spatial distribution of alpine areas based on a 30 m spatial resolution map. Dashed lines represent the equator and the tropics. (b) Distribution of treeline elevation values along latitude for 326 mountain ranges worldwide. Each dot represents the mean elevation and latitude for a mountain range. The black line represents a weighted loess fit. (c) Alpine areas extent and share for each continent. (d) Density plot of NDVI values above the treeline.

North America (9%; 0.32Mkm²) and Europe (2%; 0.08Mkm²), while Oceania and Africa together contribute to only 1% of the global alpine area (Fig. 1c). The distribution of maximum NDVI values above the treeline peaks at 0.1, reaching a median value at 0.2 and with a decreasing frequency of higher values (Fig. 1d).

The mapped alpine ecosystems are grouped toward the colder end of the global climatic space (Fig. 2), with 99% of the grid cells situated below a mean annual temperature of 5.9°C and tropical alpine ecosystems lying on this threshold. The first two axes of the PCA of the 19 bioclimatic variables (Fig. 3) explained 66% of the global variation of the alpine climate and correspond to differences in seasonality and continentality of global alpine ecosystems. Clustering the whole dataset of alpine regions using the CLARA algorithm with the first four weighted PCA axes highlighted the presence of four groups (best $k=4$ in 95% of iterations; max average silhouette width=0.40) (Supplementary material Appendix 1 Fig. A3a), that were interpreted as 1) oceanic, 2) hemiboreal, 3) continental and 4) subtropical alpine

ecosystems (Fig. 3, 4a). The hierarchical clustering based on the 30 000 records subset highlighted only two groups (best $k=2$ in 100% of iterations; mean average silhouette width=0.42) (Supplementary material Appendix 1 Fig. A3b), with the first encompassing most of the continental, hemiboreal and subtropical groups, while the second taking up most of the oceanic cluster (Supplementary material Appendix 1 Fig. A4a–c).

The four alpine clusters obtained by CLARA have similar NDVI values distribution (Fig. 4a) and they are comparable to the pattern observed at the global level (Fig. 1d). Nevertheless, this concordance disappears at the ecoregion scale, where the distribution of NDVI values, hence primary productivity, varies remarkably even within the same climatic cluster (Fig. 4b). The generalized additive model of NDVI using the PC climatic axes values as predictors explained 58% of the deviance, with highly significant parametric coefficients and smooth terms ($p<0.001$). Similarly, the model fit to a subset of pixels representing vegetated areas ($NDVI>0.1$) explained 53% of the deviance.

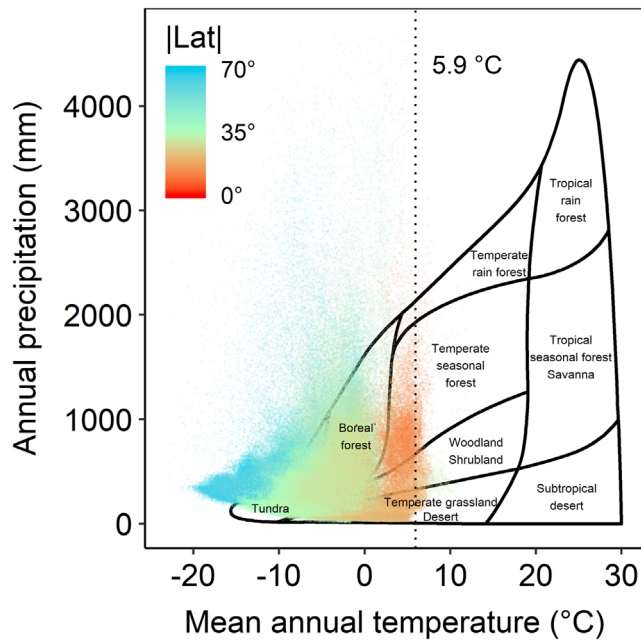


Figure 2. Values of temperatures and precipitation in alpine areas overlaid to Whittaker's biome plot. The dotted line represents the 99th percentile of the distribution of temperature values. Points are colored according to their distance from the equator ($|Lat|$). To improve readability, the figure is based on a random subset of 500 000 30 arc-second cells in alpine areas.

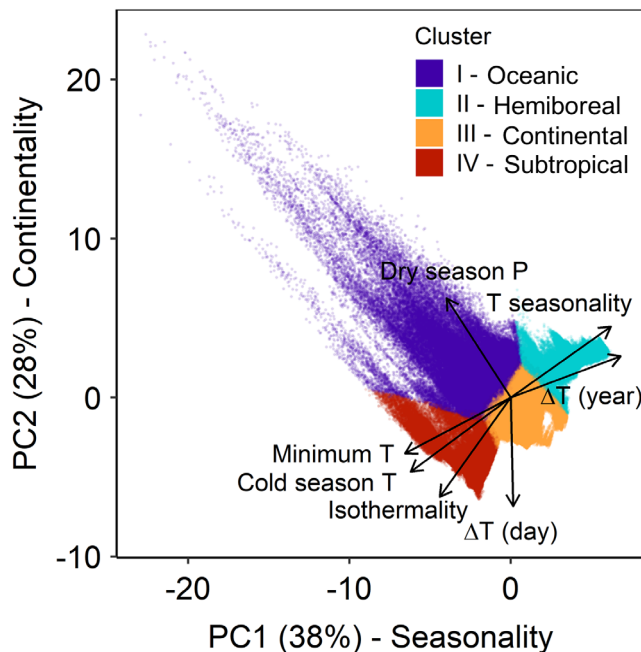


Figure 3. Biplot of principal component analysis of 19 climatic variables in alpine areas. The arrows indicate the loadings of selected climatic variables (correlation > 0.3 with one of the two axes). Points are colored according to climatic clusters. Variables names' abbreviations and symbols (P: precipitation, T: temperature, Δ : difference). To improve readability, the figures are based on a random subset of 500 000 30 arc-second cells in alpine areas.

Discussion

Spatial distribution and extent of alpine ecosystems

This study provides, for the first time, a global map of alpine ecosystems at 30 m spatial resolution (Fig. 1a) obtained through the analysis of global data sources on land cover and remote sensing (Supplementary material Appendix 1 Fig. A1). Our estimate of global alpine areas (3.56 Mkm^2) is very close to the 3.55 Mkm^2 reported by a previous study based on topography and broad climatic models (Körner et al. 2011). However, the two figures are not entirely comparable, since those authors included arctic mountain regions and excluded large parts of alpine plateaus (e.g. the Tibetan Plateau) which nonetheless may host alpine vegetation as defined here. In contrast, we based our approach on the presence of treeless vegetation thriving above the treeline, thus focusing on vegetation patterns rather than topography. For this reason (i.e. lying above the potential modeled treeline elevation, regardless of local terrain ruggedness) the flat Tibetan Plateau contributes to the total global alpine area in our map, while large portions of the Andean Altiplano were not included. Likewise, Arctic mountain regions and the rest of the Arctic tundra were excluded because they are located above the latitudinal treeline independently from elevational gradients.

We note that our estimation of treeline elevation is based on empirical forest cover data that consider as trees any vegetation taller than 5 m (Hansen et al. 2013). Although our map may include high-mountain forests with low ($< 5 \text{ m}$) trees, the trends of the NDVI suggest that this effect is not relevant, or at least such forests have low cover and are mostly located in disrupted subalpine zones. In many mountain regions of the world, the treeline has been lowered by thousands of years of human activity (Holtmeier 2009) but some remnants of the climatic treeline usually survive on the least accessible slopes, even in very exploited regions like the European Alps (Holtmeier 2009). Through the analysis of a high-resolution map, we assume that our quantile regression was mainly based on the few remaining forest patches at the climatic treeline. Indeed, the resulting pattern of global treeline elevation closely resembles previous observations derived by field measurements (Körner 2003), showing well known patterns like the higher elevation of southern hemisphere treelines, when compared to the northern at the same distance from the equator (Cieraad et al. 2014, Karger et al. 2019). It also shows the general decreasing trend in treeline elevation close to the equator already reported by Körner (2003), despite some unexpectedly high afroalpine treelines (e.g. Ruwenzori: $4706 \pm 49 \text{ m}$; Mount Kenia: $4390 \pm 6 \text{ m}$). However, these values could have arisen from a misinterpretation of the local *Dendrosenecio* woodlands vegetation in the original forest cover map. Indeed, afroalpine vegetation is characterized by the presence of these giant rosettes forming open groves above the treeline (Shugart 2005) that might have accounted for tree cover in Hansen et al. (2013). Furthermore, some treelines were higher than expected also at mid latitudes, especially in large, longitudinally stretching ranges like the

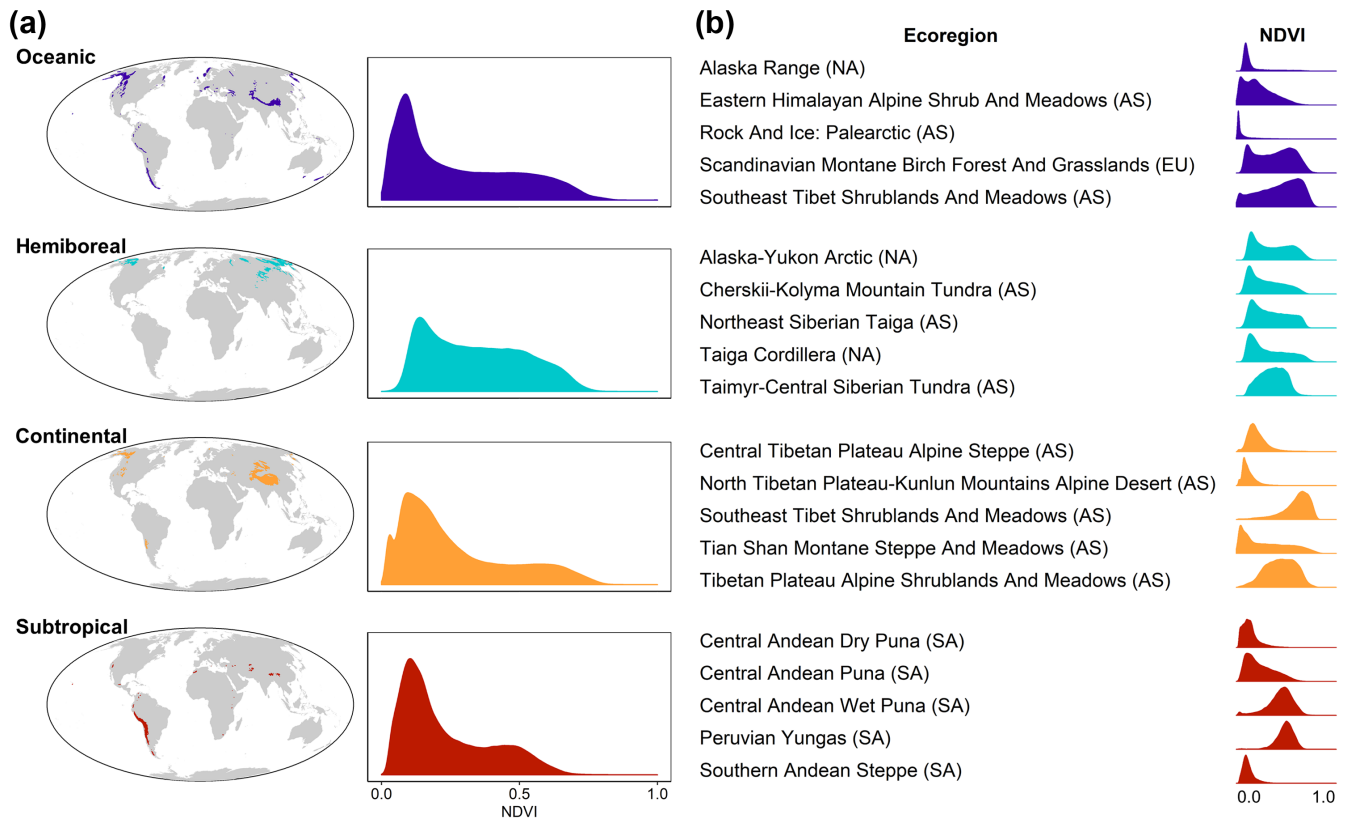


Figure 4. Distribution of climatic groups and NDVI values. (a) Global distribution of the four alpine climatic clusters and density plots of their NDVI values estimated using multitemporal Landsat eight Greenest Pixel imagery in the four alpine clusters. Y-axes are scaled to the same values. (b) Distribution of NDVI values of the alpine belt for the five largest ecoregions in each cluster. Continents are coded and reported in brackets (AS: Asia; EU: Europe; NA: North America; SA: South America). Y-axes of NDVI density plots are not scaled.

European Alps (2360 ± 19 m). This is probably due to the mass elevation effect that, combined with lower wind speeds, leads to higher treeline elevations approaching the center of large mountain ranges (Holtmeier 2009). As we did not account for these factors in our analyses, in some large mountain regions treelines may be skewed toward the upper values of their potential range, providing a rather conservative estimate of the alpine areas' extent. Likewise, the initial exclusion of some mountain regions that reportedly host an alpine belt, e.g. Iberian Peninsula mountains (Barrio et al. 2013) and Alborz mountains (Noroozi and Körner 2018), as well as the removal of smaller patches of alpine areas during the dataset cleaning process (see Material and methods; Supplementary material Appendix 1 Fig. A2), were carried out for the sake of conservativeness. Despite the acknowledged contribution of small, isolated and endemics-rich patches of lower alpine habitats and isolated regions to the overall alpine biodiversity (Körner 2003), their removal likely had negligible effects on the estimation of the global alpine area.

Climatic and productivity patterns of global alpine biomes

Since our definition of alpine ecosystems is based on land cover data rather than a-priori assumptions about climate-treeline

relationships, it allowed us to perform a climatic characterization without risk of circularity (Peters 1976). Plotting the climatic envelope of the mapped alpine ecosystems in the classic representation of Whittaker's world biomes, we found a mean annual temperature threshold of 5.9°C adjacent to tropical alpine regions (Fig. 2). Although this temperature corresponds to global climatic models at 1 km resolution, it is in line with previous studies based on field measurements that found mean annual temperatures at the treeline in tropical mountains between 5°C and 6°C (Holtmeier 2009). Below this temperature threshold, the bioclimatic space of alpine ecosystems is driven by a major latitudinal temperature gradient that mainly overlaps with the position of tundra and boreal biomes in the Whittaker's scheme (Whittaker 1975). As mountain ranges approach the equator, the alpine belt from tropical and subtropical biomes decouples from the climatic space occupied by the corresponding lowland zones (Fig. 2, Supplementary material Appendix 1 Fig. A5). This reflects the outstanding compression of life zones (Körner 2003) that is found in tropical mountains where, within a few thousand meters difference in elevation, the diversity of habitats spans from the lowland rainforest to the glaciated mountain tops (Körner and Spehn 2019). In contrast, alpine biomes at the highest latitudes are centered on the environmental space of tundra, hence climatically close to

their reference biome. They are also located at lower elevations (Supplementary material Appendix 1 Fig. A5), making an exclusively climate-based distinction between arctic tundra and alpine tundra particularly challenging.

We also found that the climatic variation within global alpine biomes is mainly linked to seasonality and continentality, and less to temperature (Fig. 3). This finding agrees with predictions on the primary role of seasonality and humidity gradients in defining alpine regions (Whittaker 1975, Nagy and Grabherr 2009), which had not been confirmed yet at the global scale due to the lack of global data. Since we had a rather complete sample of global alpine areas, we chose the CLARA algorithm to highlight the presence of main clusters with the whole dataset, in accordance with previous studies that used semi-quantitative approaches for choosing the number of clusters (Metzger et al. 2013, Weigelt et al. 2013). The optimal classification in four clusters provided a meaningful biogeographic interpretation, in comparison with the two groups suggested by hierarchical clustering with a stratified subset. Although different clustering approaches usually lead to contrasting results, especially when applied to environmental data (Weigelt et al. 2013), the results provided in the two classifications were still comparable. Oceanic alpine regions were clearly differentiated in both cases; they are distributed across all continents and latitudes, encompassing mountain ranges characterized by an oceanic influence in terms of higher precipitation and relative temperature stability (Fig. 3, 4a). Oceanic regions include the whole alpine belt of Europe and Oceania and large parts of North American ranges and the Andes, together with the Himalayas, at the interface between the seasonal humid, tropical climate of the Indian subcontinent and the cold mountain desert of the Tibetan Plateau. The other three groups defined by the CLARA algorithm were aggregated in the hierarchical clustering with the most continental subset of the oceanic group (Supplementary material Appendix 1 Fig. A4a–c), but still suggesting a clear differentiation in the climatic space. Continental alpine regions are subject to much lower precipitation rates and greater daily temperature variability, including most of central Asian mountains and the driest portion of Rocky Mountains and central Andes, which are isolated from the influence of the Pacific Ocean (Fig. 3, 4a). Interestingly, these oceanic and continental regions often occur in close vicinity within the same mountain range, sometimes even the same ecoregion (e.g. Southeast Tibet Shrubland and Meadows, Fig. 4b). This happens because topography affects macroclimatic patterns, with the most exposed slopes forming a barrier to humid air streams, hence causing rainfalls on the one side and much drier conditions on the opposite. In contrast, hemiboreal alpine regions occur mainly at boreal latitudes of the northern hemisphere and have lower annual and seasonal temperature minima (Fig. 3, 4a). They comprise most of the Siberian mountains and the northernmost ranges of North America, while subtropical alpine regions are mainly represented in the Andes and other tropical or subtropical regions that exhibit higher temperature minima and a much more stable climate

throughout the year, despite marked diurnal variations (Fig. 3, 4a). The latitudinal overlap of oceanic alpine regions with the others is in part an inherent consequence of the clustering approach. Indeed, the portion of oceanic alpine regions at higher latitudes is characterized by relatively greater seasonality than the one located closer to the equator. As a matter of fact, the oceanic group occupies a rather wide section of the global environmental space of alpine ecosystems (Fig. 3). However, in a global perspective, oceanic regions as a whole are separated by hemiboreal and subtropical ones, forming a coherent, independent group. Furthermore, the spatial distribution and climatic characteristics of the oceanic alpine cluster are consistent with the oceanic group of the Köppen–Geiger climate classification (Köppen 1936). Oceanic climate is indeed characterized by relatively stable temperatures, the absence of a dry season and covers both coastland and inland areas of all continents, including mountain areas at subtropical latitudes like African mountains and parts of the Himalaya.

We also used global remote-sensing information to characterize alpine biomes using NDVI as a surrogate of photosynthetic activity and vegetation productivity (Whittaker 1975, Mucina 2019). In line with previous estimates (Bradley et al. 2017), the peak of the distribution of NDVI values in alpine ecosystems (Fig. 1d) indicates dominance of bare or scarcely vegetated areas, while the lower frequencies of higher values represent the most productive vegetation found in these areas (i.e. grasslands and dwarf shrubs). Despite differences in temperature seasonality and amount of precipitation, alpine ecosystems grouped by climatic similarity show analogous patterns of NDVI variation, which in turn reflect a global system characterized by a large portion of poorly vegetated and low-productive areas. However, when looking at the NDVI values distribution among WWF ecoregions (Olson et al. 2001) within the same climatic group (i.e. comparing alpine ecosystems with similar climate across the globe, Fig. 4b), visible differences can be observed in all groups. These results suggest that, although macroclimate is able to explain 58% of the overall variation in NDVI, this is not the only factor shaping vegetation structure and function in the alpine belt, which may also differ strongly among regions. The growth of alpine plant species and the dominance of specific life forms also depend on regional and local factors like fine scale topography, disturbance and biogeographic history (Körner 2003, Jiménez-Alfaro et al. 2014), as additional factors explaining the heterogeneity of regional biodiversity across mountain regions. More studies are therefore needed to characterize the functional properties of alpine biomes, by combining remote-sensing indices with data collected on the ground from different vegetation types across regions.

Conclusions

This study provides a fine-scale estimate of the worldwide extent of alpine biomes and their bioclimatic characterization. Rather than relying on temperature thresholds, our study provides an empirical view on this decades-old issue,

using big data sources and a consistent definition of the study system. Although this study is based on a well-accepted definition of alpine ecosystems, we note that there could be different views on the interpretation of alpine versus arctic tundra, the inclusion or exclusion of rugged areas within the alpine zones, or the definition of world biomes under different frameworks. Our approach provides a conservative estimate of the extent of alpine areas, but our methodological framework had little effect on their bioclimatic characterization. Indeed, our workflow can be easily applied from local to global scales and adjusted according to specific aims and conceptual assumptions.

By considering the assumptions of our approach, this study also provides the first spatial and bioclimatic characterization of alpine biomes using a consistent, data-driven methodology. In general terms, we highlight that global alpine biomes occupy a relatively well-defined and continuous climatic space, which is geographically and climatically independent from other biomes regardless of latitude. Alpine biomes are mainly driven by seasonality and continentality gradients, but major groups defined over this variation may be heterogeneous in the structure and function of dominant vegetation, reflecting regional differences and the coexistence of multiple plant life-forms. Our findings are likely to be consistent under other assumptions on alpine biomes, given that we have analyzed most alpine regions in the world, but less so for those approaches including arctic tundra into the same methodological framework, because this will add a new source of climatic variability. For the assessment of alpine biomes as defined here (i.e. high-elevation regions above the climatic treeline), our results may help in the evaluation of these relevant ecosystems at global and regional scales. The associated data sources of this study also provide useful tools for biodiversity assessment, ecological modeling, habitat monitoring or the analysis of climate change adaptation of different biota.

Data availability statement

The data produced in this study and used for the analyses, including the 1) Google Earth Engine code used for extracting unforested areas, 2) R code used for modelling treeline elevation and extracting alpine areas, 3) table with means and standard deviations of treeline elevations for world mountain ranges, 4) shapefiles of alpine areas at 30 m resolution and 5) shapefiles of alpine areas at 1 km resolution divided by clusters and ecoregions are available from Figshare Digital Repository: <<https://dx.doi.org/10.6084/m9.figshare.11710002>> (Testolin et al. 2020).

Acknowledgements – The authors would like to thank F. M. Sabatini (German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig; Martin-Luther Univ. Halle-Wittenberg, Inst. of Biology, Geobotanic Lab) and A. Farcomeni (Univ. of Rome Tor Vergata) for their useful advice. They would also like to thank NASA and Google for making global data and computation tools freely available. The ASTER GDEM is a product of NASA and METI. SRTM is a product of NASA.

Funding – BJ-A thanks the Marie Curie Clarín-COFUND program of the Principality of Asturias-EU (ACB17-26) and the regional grant IDI/2018/000151 for research groups.

Conflicts of interest – The authors declare no conflict of interest.

References

- Barrio, I. C. et al. 2013. Alpine ecology in the Iberian Peninsula: what do we know, and what do we need to learn? – *Mt. Res. Dev.* 33: 437–442.
- Bastin, J.-F. et al. 2017. The extent of forest in dryland biomes. – *Science* 358: 635–638.
- Bastin, J.-F. et al. 2019. The global tree restoration potential. – *Science* 79: 76–79.
- Bradley, Z. C. et al. 2017. Observed long-term greening of alpine vegetation – a case study in the French Alps. – *Environ. Res. Lett.* 12: 114006.
- Cieraad, E. et al. 2014. Southern Hemisphere temperate tree lines are not climatically depressed. – *J. Biogeogr.* 41: 1456–1466.
- Cihlar, J. et al. 1991. Relation between the normalized difference vegetation index and ecological variables. – *Remote Sens. Environ.* 35: 279–298.
- Faber-Langendoen, D. et al. 2016. Classification and description of world formation types. – *Gen. Tech. Rep. RMRS-GTR-346*, U.S. Dept of Agriculture, Forest Service, Rocky Mountain Research Station, Fort Collins, CO.
- Farr, T. G. et al. 2007. The shuttle radar topography mission. – *Rev. Geophys.* 45: RG2004.
- Gao, Q. et al. 2016. Climatic change controls productivity variation in global grasslands. – *Sci. Rep.* 6: 1–10.
- Gorelick, N. et al. 2017. Google Earth Engine: planetary-scale geospatial analysis for everyone. – *Remote Sens. Environ.* 202: 18–27.
- Hansen, M. C. et al. 2013. High-resolution global maps of 21st-century forest cover change. – *Science* 342: 850–853.
- Holben, B. N. 1986. Characteristics of maximum-value composite images from temporal AVHRR data. – *Int. J. Remote Sens.* 7: 1417–1434.
- Holdridge, L. R. 1947. Determination of world plant formations from simple climatic data. – *Science* 105: 367–368.
- Holtmeier, F.-K. 2009. Mountain timberlines. Ecology, patchiness and dynamics. – Springer.
- Hughes, L. 2000. Biological consequences of global warming: is the signal already apparent? – *Trends Ecol. Evol.* 15: 56–61.
- Jiménez-Alfaro, B. et al. 2014. Biogeographic deconstruction of alpine plant communities along altitudinal and topographic gradients. – *J. Veg. Sci.* 25: 160–171.
- Karger, D. N. et al. 2017. Climatologies at high resolution for the earth's land surface areas. – *Sci. Data* 4: 170122.
- Karger, D. N. et al. 2019. Why tree lines are lower on islands – climatic and biogeographic effects hold the answer. – *Global Ecol. Biogeogr.* 28: 839–850.
- Kaufman, L. and Rousseeuw, P. J. 1990. Finding groups in data: an introduction to cluster analysis. – Wiley.
- Koenker, R. 2018. Quantreg: quantile regression. – R package ver. 5.36, <<https://cran.r-project.org/package=quantreg>>.
- Köppen, W. 1936. Das geographische system der klimate. – In: Köppen, W. and Geiger, G. (eds), *Handbuch der Klimatologie*. Borntraeger, pp. 1–44.
- Körner, C. 1995. Alpine plant diversity: a global survey and functional interpretations. – In: Chapin III, F. S. and Körner, C.

- (eds), Arctic and alpine biodiversity: patterns, causes and ecosystem consequences. Springer, pp. 45–62.
- Körner, C. 2003. Alpine plant life. Functional plant ecology of high mountain ecosystems. – Springer.
- Körner, C. 2012. Alpine treelines: functional ecology of the global high elevation tree limits. – Springer.
- Körner, C. and Spehn, E. M. 2019. Mountain biodiversity: a global assessment. – Routledge.
- Körner, C. et al. 2003. A bioclimatic characterisation of Europe's alpine areas. – In: Nagy, L. et al. (eds), Alpine biodiversity in Europe. Springer, pp. 13–28.
- Körner, C. et al. 2011. A definition of mountains and their bioclimatic belts for global comparisons of biodiversity data. – *Alp. Bot.* 121: 73–78.
- Körner, C. et al. 2017. A global inventory of mountains for biogeographical applications. – *Alp. Bot.* 127: 1–15.
- Liu, S. et al. 2017. Spatiotemporal dynamics of grassland above-ground biomass on the Qinghai-Tibet Plateau based on validated MODIS NDVI. – *Sci. Rep.* 7: 1–10.
- Maechler, M. et al. 2018. cluster: cluster analysis basics and extensions. – R package ver. 2.0.7-1.
- Metzger, M. J. et al. 2013. A high-resolution bioclimate map of the world: a unifying framework for global biodiversity research and monitoring. – *Global Ecol. Biogeogr.* 22: 630–638.
- Mucina, L. 2019. Biome: evolution of a crucial ecological and biogeographical concept. – *New Phytol.* 222: 97–114.
- Müllner, D. 2013. fastcluster: fast hierarchical, agglomerative clustering routines for R and python. – *J. Stat. Softw.* 59: 1–18.
- Myers, N. et al. 2000. Biodiversity hotspots for conservation priorities. – *Nature* 403: 853–858.
- Nagy, L. and Grabherr, G. 2009. The biology of alpine habitats. – Oxford Univ. Press.
- NASA and JPL 2013. NASA Shuttle Radar Topography Mission Global 1 arc second number [SRTMGL1N_003]. – NASA EOSDIS Land Processes DAAC, accessed 2019-01-10.
- NASA/METI/AIST/Japan Spacesystems and U.S./Japan ASTER Science Team. 2009. ASTER Global Digital Elevation Model [ASTGTM v002]. – NASA EOSDIS Land Processes DAAC, accessed 2019-01-10.
- Noroozi, J. and Körner, C. 2018. A bioclimatic characterization of high elevation habitats in the Alborz mountains of Iran. – *Alp. Bot.* 128: 1–11.
- Olson, D. M. et al. 2001. Terrestrial ecoregions of the world: a new map of life on earth. – *Bioscience* 51: 933.
- Peters, R. H. 1976. Tautology in evolution and ecology. – *Am. Nat.* 1104: 1–12.
- Pettorelli, N. et al. 2005. Using the satellite-derived NDVI to assess ecological responses to environmental change. – *Trends Ecol. Evol.* 20: 503–510.
- Pomeroy, J. 2015. Research network to track alpine water. – *Nature* 521: 32.
- QGIS Development Team 2016. Qgis geographic information system. – Open Source Geospatial Foundation Project, <www.qgis.org/>.
- Quinn, J. A. 2008. Arctic and alpine biomes. – Greenwood.
- Shugart, H. H. 2005. Remote sensing detection of high elevation vegetation change. – In: Huber, U. M. et al. (eds), Global change and mountain regions. Advances in global change research, vol. 23. Springer, pp. 457–465.
- Testolin, R. et al. 2020. Data from: Global distribution and bioclimatic characterization of alpine biomes. – Figshare Digital Repository, <<https://dx.doi.org/10.6084/m9.figshare.11710002>>.
- Walter, H. and Box, E. 1976. Global classification of natural terrestrial ecosystems. – *Vegetatio* 32: 75–81.
- Weigelt, P. et al. 2013. Bioclimatic and physical characterization of the world's islands. – *Proc. Natl Acad. Sci. USA* 110: 15307–15312.
- Whittaker, R. H. 1975. Communities and ecosystems. – Macmillan Publishing.
- Wood, S. 2011. Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. – *J. R. Stat. Soc. B* 73: 3–36.

Supplementary material (available online as Appendix ecog-05012 at <www.ecography.org/appendix/ecog-05012>. Appendix 1.