Check for updates

Bioregionalization of Albania: Mismatch between the flora and the climate suggests that our models of Southern European bioregions are in need of a revision

Luca Malatesta · Michele De Sanctis · Emile Ammann · Fabio Attorre · Francesca Buffi · Vito Emanuele Cambria · Carlo Fratarcangeli · Petrit Hoda · Ermelinda Mahmutaj · Marjol Meco · Alfred Mullaj · Lulëzim Shuka · Giuliano Fanelli

Abstract We analysed the floristic subdivisions of Albania by hierarchical clustering of all the vascular plant species of Albania over a grid of 25 km cells, adapting data from the *Vascular Species Distribution Atlas in Albania* of (Barina Z (ed.) (2017) Distribution atlas of vascular plants in Albania. Hungarian Natural History Museum, Budapest). We identified the principal axes of variation of the flora and the bioclimatic variables that explain it. We also analysed the bioclimate of Albania by means of a self-organizing map (SOM) of the main climatic variables and the application of the global bioclimatic model of Rivas-Martínez. We compared the analysis based on the flora with that based on the bioclimate. The results divided Albania into eight floristic units corresponding to the main vegetation types (*Carpinus orientalis* woodlands, *Quercus cerris* woodlands, *Quercus petraea* and *Fagus sylvatica* woodlands, *Quercus coccifera* woodlands, *Quercus coccifera* and *Quercus ithaburensis* woodlands, *Pinus heldreichii* woodlands, alpine vegetation of North Albania, and alpine vegetation of eastern Albania). Two main trends of

L. Malatesta · M. De Sanctis (✉) · E. Ammann · F. Attorre · F. Buffi · V. E. Cambria · C. Fratarcangeli · G. Fanelli
Department of Environmental Biology, University Sapienza of Rome, P.le Aldo Moro 5, 00185 Rome, Italy
e-mail: michele.desanctis@uniroma1.it

L. Malatesta
e-mail: luca.malatesta@uniroma1.it

E. Ammann
e-mail: ammann.1662487@studenti.uniroma1.it

F. Attorre
e-mail: fabio.attorre@uniroma1.it

F. Buffi
e-mail: francesca.buffi@uniroma1.it

V. E. Cambria
e-mail: vitoemanuele.cambria@uniroma1.it

C. Fratarcangeli
e-mail: carlo.fratarcangeli@uniroma1.it

G. Fanelli
e-mail: giuliano.fanelli@gmail.com

P. Hoda · E. Mahmutaj
Research Center of Flora and Fauna, Faculty of Natural Sciences, University of Tirana, Tirana, Albania
e-mail: hodapetrit@yahoo.com

E. Mahmutaj
e-mail: ermelinda.mahmutaj@fshn.edu.al

M. Meco · A. Mullaj · L. Shuka
Department of Biology, Faculty of Natural Sciences, University of Tirana, Tirana, Albania
e-mail: mariol.meco@yahoo.com

A. Mullaj
e-mail: alfredmullaj@gmail.com

L. Shuka
e-mail: lulezim.shuka@fshn.edu.al

variation can be recognized: four elevational belts, strongly correlated with average July temperature, and a north–south subdivision, weakly correlated with annual rainfall and thermal excursion. The bioclimatic analysis (SOM and Rivas-Martínez model) resolves and adds detail to the axis of the elevational belts but cannot discern the north–south subdivision apart from the high mountains. We obtained a division into eight floristic units ordered according to two gradients: four elevational belts, each in turn subdivided into a north-central and a southern province. These two gradients seem to have a broader meaning in Southern Europe, particularly the north–south subdivision at about 30° latitude. The partial mismatch between the bioclimate and the flora indicates that bioclimatic models of Europe are in need of a revision and suggests that the history, in addition to the present-day climate (for instance, the vegetation of the last glacial maximum), can play a role in shaping the biogeography of Southern Europe.

Introduction

Albania is a small territory that has remained isolated for a long time. Only in recent years has it commenced opening up to the rest of the world, thanks also to the procedure for joining the European Union, a labourious process that is already profoundly changing the face of the country.

The long isolation has meant that botanical and phytogeographical knowledge has lagged far behind the rest of Europe. For years, much of what we knew about the Albanian flora and vegetation was due to the tireless and adventurous activity of two great botanists, Antonio Baldacci from Italy and Friedrich Markgraf from Germany, who undertook their studies back at the beginning of the past century (Baldacci 1896; Markgraf 1932; Lack and Barina 2020). Today, the situation is rapidly changing in the country, as a new generation of local botanists increasingly emerges in the international arena.

Markgraf (1932) is still the cornerstone of Albania's phytogeographical acquaintance and floristic regionalization. In this work, the authors proposed a subdivision

into elevational belts (or 'phytoclimatic zones') and phytogeographical sectors based on floristic data and personal observations. The various existing European phytogeographical syntheses, such as that of the European Union (EEA 2000), basically follow the schematization of Markgraf. They place Albania on the border between the Eurasian and Mediterranean regions, emphasizing how the peculiarities of this Balkan country derive from this border position. At a distance of ninety years from the initial publication, an update of Markgraf's synthesis takes on significant value. Today's increased amount of information available, such as the atlas of Vangjeli (2015) and that of Barina (2017), makes this exercise more accessible. Barina's atlas is particularly useful because it reports the distribution of the Albanian plant species in great detail. This atlas has enabled the attempt to regionalize the Albanian flora presented in this paper.

Modern phytoclimatic models can offer further methodological advancements to this aim. For example, the inspirational work of Rivas-Martínez (Rivas-Martínez et al. 2011) has never been applied to Albania so far, except for a small area (Dajti) near Tirana (Topi and Mankolli 2012).

We approached the problem of the bioregionalization of Albania from three different perspectives: (1) a floristic analysis based on the distribution of species in a 25-km grid; (2) a climate analysis based on the main bioclimatic variables spatialized over the same grid as the floristic analysis; and (3) the application of the Rivas-Martínez model to Albania.

The different approaches are complementary and facilitate a better understanding of the patterns of the phytogeography of Albania.

Material and Methods

Study Area

With a total area of 28,750 km^2 and an average elevation of 708 m a.s.l., Albania is one of the most mountainous countries in Europe (Zickel et al. 1994). Mountains and hills constitute natural barriers that limit access to much of the inland. The remainder is occupied by the coastal plains of the Adriatic and Ionian Seas and a complex and often inhospitable system of floodplains that lie intricately between imposing mountain ranges.

The sedimentary cover of carbonate deposits is abundant on a Palaeozoic basement of metamorphic and magmatic rocks (Fig. 1). Dolomites and limestones of different formations are interspersed with Miocene deposits of flysch and clays derived from turbidites brought to the surface during the Alpine orogeny (Barina 2017).

The presence of ophiolites and ultrabasic rocks is frequent in many mountain groups distributed in the north–south direction in continental Albania (Bortolotti et al. 1996). These formations are of particular floristic interest due to numerous endemism of plant species adapted to the heavy metals abundant in these substrates (De Sanctis et al. 2018). Pliocene and recently
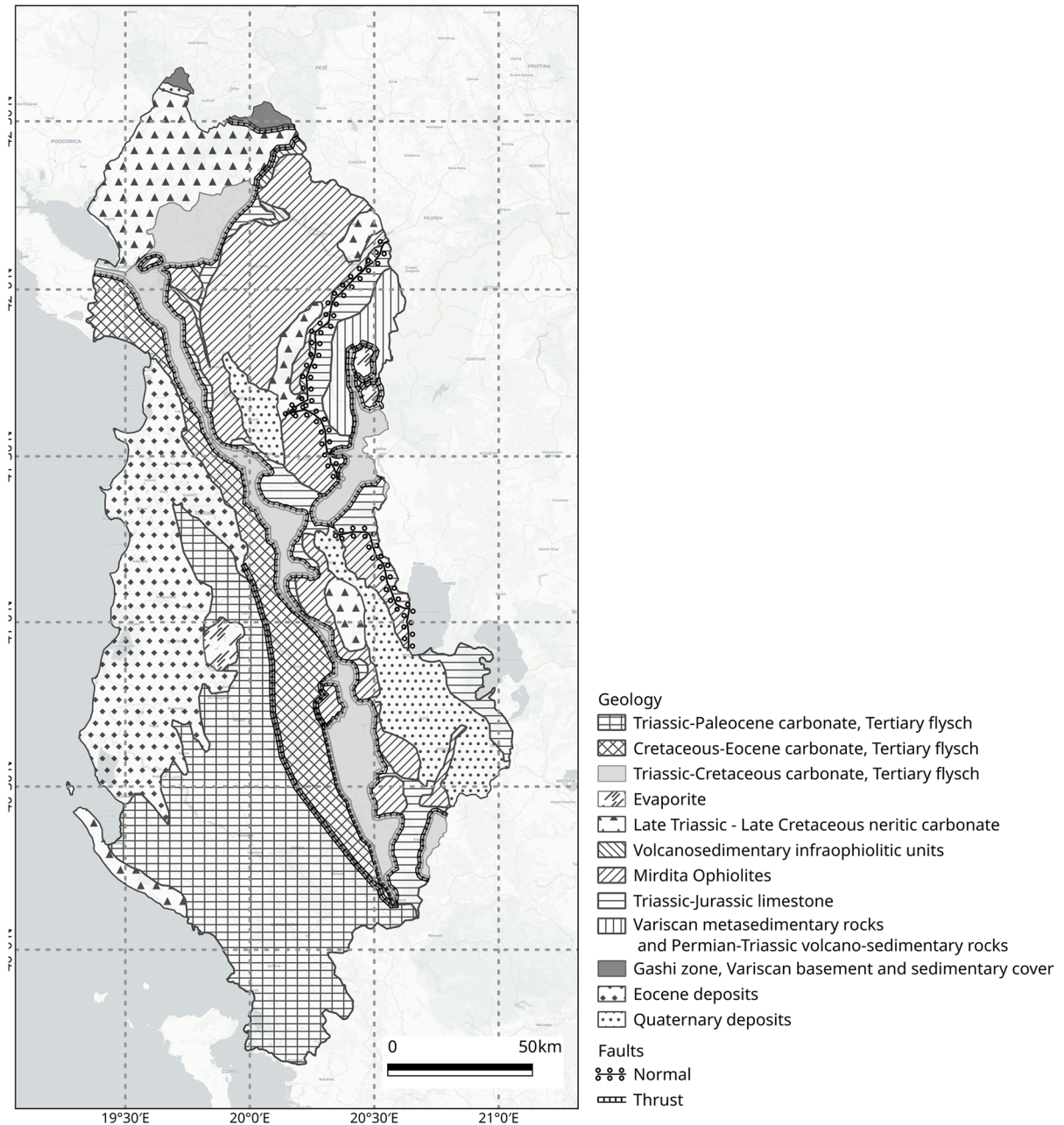

Geology

⊞ Triassic-Paleocene carbonate, Tertiary flysch
⊠ Cretaceous-Eocene carbonate, Tertiary flysch
▨ Triassic-Cretaceous carbonate, Tertiary flysch
▨ Evaporite
⌐ Late Triassic - Late Cretaceous neritic carbonate
▨ Volcanosedimentary infraophiolitic units
▨ Mirdita Ophiolites
⊟ Triassic-Jurassic limestone
▥ Variscan metasedimentary rocks
     and Permian-Triassic volcano-sedimentary rocks
▨ Gashi zone, Variscan basement and sedimentary cover
⊡ Eocene deposits
⋯ Quaternary deposits

Faults

⚬⚬⚬ Normal
⊔⊔⊔ Thrust

Fig. 1 Simplified geological map of Albania (after Tremblay et al. 2015)

formed deposits are confined to coastal plains and numerous alluvial valleys throughout the territory (Barina 2017; Meco and Aliaj 2000).

The flora of Albania is extremely rich considering its modest area, including about 4,000 autochthonous vascular plant species distributed in 165 families and 910 genera (Vangjeli et al. 2000a, b). Bearing in mind that Albania is much less explored than the rest of Europe and that many new species have been discovered in recent years, a significant increase in the number of species is expected in the next decade.

Dataset

This work is based on the distribution atlas by Barina (2017). The taxonomy and nomenclature of vascular plants follow the Euro+Med PlantBase (2006). The database offers detailed maps of the distribution of 3,168 taxa and includes: (i) 50,000 floristic reports collected between 2004 and 2016 by the authors with GPS coordinates; (ii) 16,500 herbarium samples deposited at the Hungarian Museum of Natural History in Budapest.; (iii) 21,000 herbarium data entries extrapolated from European collections; and (iv) 29,000 published data entries.

We considered only autochthonous species and subspecies. Data from taxonomically unstable groups and unconfirmed reports from the same localities in recent times were discarded.

A square grid of 25 km × 25 km sides for a total of 67 cells was superimposed on the species distribution maps (on glossy paper). In the Albanian context, this resolution effectively intercepts a large part of the floristic variability of the territory while maintaining the right level of detail to define the primary phytogeographical limits. A higher resolution would not have corresponded with the detail of the original data. It would also have resulted in many empty cells in which species are present but have not been detected. The presence / absence species data for each cell were used to create a plot vs species matrix (67 plots vs 3,140 species and subspecies).

Data Analysis

Cluster Analysis

To identify a possible division into groups of floristically similar cells, hierarchical classification was performed using the 'cluster' package (version 1.15.2)

in R version 3.6.1 (R Core team 2019). The clustering technique used was hierarchical aggregation. We adopted the Jaccard distance, which measures dissimilarity between sample sets, and Ward's minimum variance method (Murtagh and Legendre 2014) to calculate the distances (or dissimilarities) between groups and obtain a dendrogram of classification.

We chose the cutting level of the dendrogram based on average silhouette width, a popular cluster validation index for estimating the number of clusters. It measures the degree to which an object belongs to its cluster, based on the average dissimilarity between it and all those in its cluster, compared to the same measure determined for the next closest cluster.

Environmental Variables

As potential environmental drivers, we tested the nineteen bioclimatic variables included on the WorldClim website (www.worldclim.org/bioclim – Fick and Hijmans 2017). We used version 2.1 of the dataset, released in January 2020. We also considered some derived bioclimatic variables: average annual temperature / maximum temperature of the warmest month (a derived index, linked with temperature range: 'bio7 / bio5'), maximum temperature of warmest month / average annual rainfall (Emberger index approximation: 'bio5 / bio12', Emberger 1945), water resource (annual precipitation/potential evapo-transpiration: 'wr') and the Mediterranean index (summer dryness bio9 / bio18: 'BioMed'). Because the resolution of the WorldClim model is 1 km^2, we considered the mean value of each environmental variable inside our 25 × 25-km grid as the value for each cell (Hijmans et al. 2005).

To check for multicollinearity in the complete set of covariates, a correlation analysis was performed, and highly correlated variables were eliminated ($R^2 > 0.75$). The final set of variables included was: bio5, bio7, bio12, bio5 / bio12, wr.

Gradient Analysis

For the gradient analysis, we adopted NMDS (nonmetric multidimensional scaling), using the 'meta-MDS' function of the 'vegan' package (version 2.5.6) in R. We adopted the Jaccard dissimilarity index for sorting. To analyse the relationships between the

distribution pattern of plant communities and environmental variables, we superimposed environmental vectors on the NMDS sorting using the 'envfit' function.

Bioclimatic Classification by SOM (Self-Organizing Map)

The bioclimatic classification was obtained by means of a dimensionality reduction based on neural networks, a method named self organizing map (SOM). The SOM (Chon et al. 1996; Kohonen 2001) is an unsupervised neural network built in a predefined two-dimensional space (the user defines the number of nodes desired): in our case a grid of 5×5 hexagonal nodes (25 total nodes, every node having 6 neighbours). After the SOM grid is constructed, each of the 67 cells subject to the analysis (the 25×25-km cells into which the country was divided), with its associated bioclimatic parameters, is presented to the neural network and assigned to a node based on an 'activation function' that uses bioclimatic variables as training. Following the assignment, the parameters of the selected node (importance of each of the bioclimatic variables within it) and of the nearby ones are updated, and the map gradually organizes itself autonomously (hence the name). The final result preserves the topology of the elements (in our case 25×25-km grid cells): Similar cells will be found in the same node or in adjacent nodes, and as the distance between the nodes increases, the dissimilarity between them will also increase. Self-organizing maps are often used in ecology and environmental research (Kim et al. 2023; Dhakal et al. 2023; Tedesco and Fettweis 2020) due to their effectiveness in reducing the complexity of datasets while preserving the topology of elements. After the dimensionality reduction obtained with SOM, agglomerative hierarchical clustering was performed using Ward's minimum variance method and Manhattan distance. The SOM was generated using the 'kohonen' package (version 3.0.8) in R (www.R-project.org).

Initially, we tried a classification of bioclimatic variables with standard cluster analysis (not shown among the results), but its results were utterly uncorrelated with species-based clustering (being almost exclusively driven by aridity and mediterraneity) and did not distinguish the elevational belts identified by the species-based clustering (differences in elevation distribution tested non-significant). On the other hand, SOM analysis on species data also results in a weak clustering (average silhouette width: < 0.3). Therefore, we decided to adopt different methods for the two datasets analysed, using hierarchical clustering for species presence-absence and resorting to a more advanced machine learning approach for a bioclimatic classification.

Bioclimatic Classification According to Rivas-Martínez

The Rivas-Martínez bioclimatic model was developed for the Iberian Peninsula and subsequently extended to cover all vegetation worldwide (Rivas-Martínez et al. 2011). The model is based on the calculation of a series of bioclimatic indices that characterize thermicity (Tp), thermal continentality (Ic, Icext) and humidity (Ios, Ios2, Ios3). The combination of these indices produces a subdivision of the territory that in the Iberian Peninsula corresponds excellently with phytogeographic limits on a floristic basis.

For the bioclimatic characterization, according to Rivas-Martínez, monthly data of rainfall, mean annual temperature, mean maximum temperature and mean minimum temperature for ten Albanian stations were used. The data were collected from various sources, including FAO and Weatherbase (www.fao.org; www.weatherbase.com).

First, we drew thermo-hydrograms (Fig. S1 in the electronic supplementary material) according to Walter and Lieth (1967) using the 'climatol' package (version 3.1.2) of the R software. Rivas-Martínez indices were then calculated for all ten stations. Using the R language and environment for statistical computing, we interpolated and modelled the values of the indices over the whole country by means of regression–kriging, a spatial prediction technique commonly used in climatology (Moral 2010; Sekulić et al. 2020), based on a linear regression of the response variable (Rivas-Martínez indices) on a set of predictors, followed by the kriging of the residuals. In this study, latitude, longitude and elevation were used as predictors for the linear regression, based on a generalized linear model. These analyses were performed using the R packages 'gstat' (version 2.0-5), 'mgcv' (version 1.8-28) and 'raster' (version 2.8-19) in R the environment. In this way, we produced a map of thermotypes and a map of bioclimates for comparison with floristic regionalization.

Chorotypes

To characterize in a phytogeographical sense the groups obtained from the analysis of the floristic clusters, the percentage of chorotypes for each cell was calculated. To this end, a chorotype was assigned to each species in common with Italy, based on Pignatti (1982), and for species not present in Italy based on the Flora of Albania (Flora e Shqipërisë – Vangjeli et al. 2000a, b). For a limited number of species not indicated in any of the two sources, the chorotype was evaluated based on the maps of distribution by countries provided by the EuroMed PlantBase (ww2. bgbm.org/EuroPlusMed).

The chorotypes used were: European and Circumboreal, Sub-tropical, Orophytic, Atlantic, Mediterranean (steno-Mediterranean + eu-Mediterranean), Balkan, Pontic, Cosmopolite, Arctic–Alpine, and Endemic (Pignatti 1982). Besides those, we added a 'Balkan' chorotype, which also includes species distributed in Greece and the Balkans.

Results

Cluster Analysis

At the first level, the dendrogram divides the 67 cells into two main clusters corresponding to the Mediterranean coastal belt and the temperate mountainous inland belt. At the second level of the dendrogram, four clusters are obtained (Fig. 2a): The coastal belt is divided between southern and northern central Albania, and the inland belt is divided between a hilly and montane belt on one side and an alpine and boreal belt on the other side (Fig. 2b).

At the level of eight clusters (Fig. 3a), Albania is divided the north-south direction (Fig. 3b) and the south-west / north-east direction. In the south-west / north-east direction, we observe a coastal belt with a Mediterranean character, then a hilly belt with a sub-Mediterranean character, followed by a mountainous belt, and finally an alpine belt. These are further divided into a central-northern and a southern Mediterranean belt that separate approximately at the latitude of Vlora. This does not apply to the sub-Mediterranean belt, which is only present in the north.
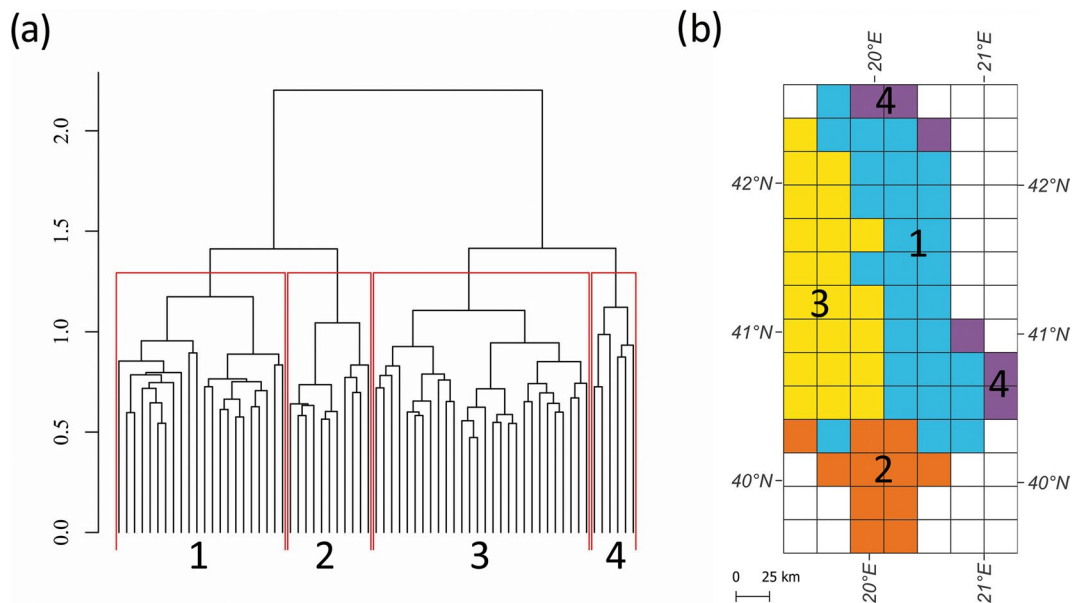

Fig. 2 A – Dendrogram produced by cluster analysis using the Jaccard distance and Ward's minimum variance method cut at the level of four clusters. **b** – Distribution of the four clusters identified in Albania. The four groups in the figure describe the coastal region (clusters 2 and 3), the mountain region (cluster 1) and the alpine region (cluster 4)
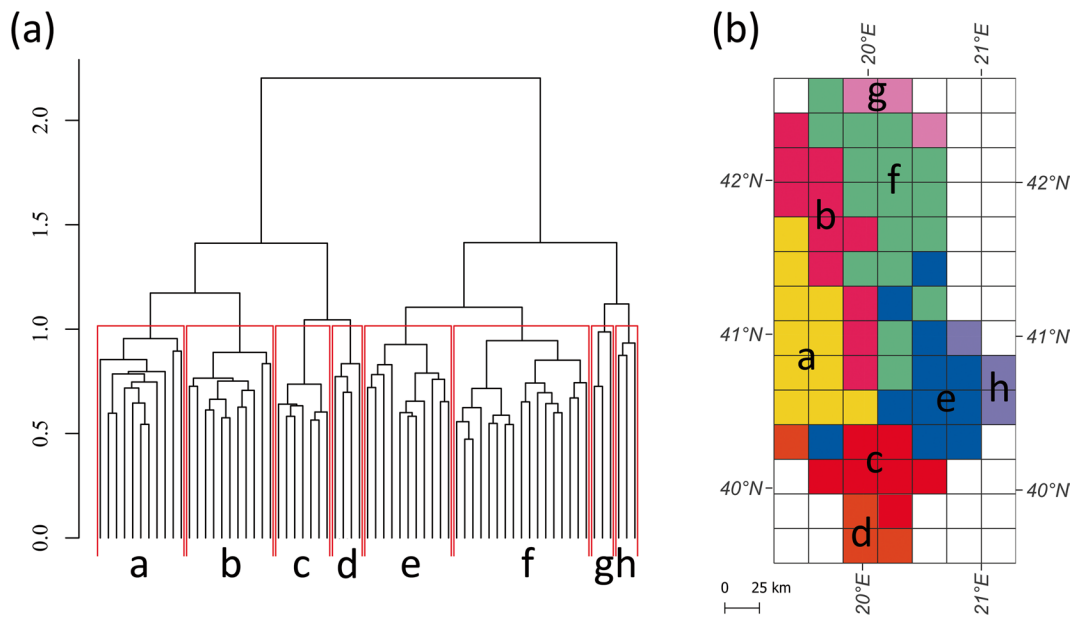
Fig. 3 **A** – Dendrogram produced by cluster analysis using the Jaccard distance and Ward's minimum variance method cut at the level of eight clusters. **b** – Distribution of the eight clusters identified in Albania

Gradient Analysis

Two gradients are evident on the NMDS plane (Fig. 4). The first axis expresses a gradient from low coastal areas (positive values) to higher situated inland areas (negative values) and corresponds broadly with elevational belts. The second separates the areas in central and northern Albania (positive values) from those in the southern Albania (negative values).

The bioclimatic variables bio5 and bio12 are strongly correlated with the two axes of the NMDS (Fig. 4; Table 1). Variable bio5 (maximum summer temperature) explains the gradient from the inland towards the coast whereas bio12 (annual rainfall) explains the distinction between Albania's central north and its south. Variable bio7 (temperature excursion) is also strongly correlated with the second axis in a similar direction and in the direction opposite to it.

Bioclimatic Classification According to Rivas-Martínez

The bioclimatic indices Tp, Ios2 and Icext were spatialized across the territory of Albania (Fig. 5a, b).

A series of thermotypes were identified by applying the Rivas-Martínez indices interpolated over the Albanian territory (Fig. 5a). They subdivide the various phytogeographical contexts as follows: (1) The coastal and lowland belt is assigned to the meso-Mediterranean, both in the south and the north of Albania; (2) a narrow band of hills in northern Albania is assigned to the colline belt; (3) mountains are divided into the sub-mountain and the high mountain belt; and (4) the highest peaks are divided between the subalpine belt and the boreal belt.

There is a good correspondence with the floristic classification, even if the Mediterranean aspects are included in the meso-Mediterranean. Furthermore, the Albanian situation leads to a different interpretation of the thermotypes than that of Rivas-Martínez, which is essentially based on the western Mediterranean (Iberian Peninsula). The macro-bioclimates correspond substantially with the thermotypes (Fig. 5a, b). Still, a clear differentiation is observed between the continental and the oceanic Mediterranean. The former is distributed in a belt between Shkoder and Tirana, the latter in the lowland belt between Durres and Saranda. Ombrotypes vary little between sub-humid and low-humid areas. As mentioned in the Introduction, Albania has a rather rainy

Fig. 4 NMDS result showing how clusters are spatially separated according to different gradients
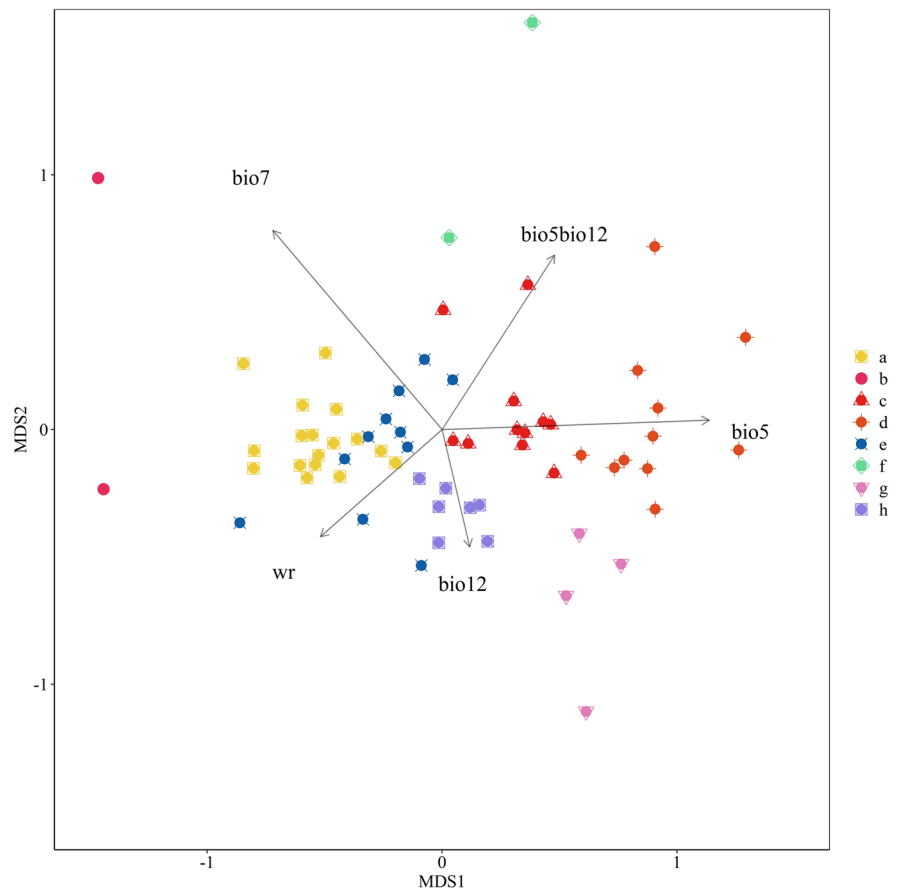

Table 1 Scores of environmental vectors in the NMDS plot. Each score defines the position of the corresponding variable's vector on the axis of the plot shown in Fig. 4 (NMDS1 is the *x* axis, NMDS2 is the *y* axis). The correlation of the variables with each axis is reported in the third column using the R^2 metric. Significance codes: '***' – ≤ 0.001, '*' – ≤ 0.01 but > 0.05.

	NMDS1	NMDS2	R^2	$Pr(> r)$	Sig. codes
bio5	0.99948	0.03227	0.5756	0.001	***
bio7	−0.67744	0.73557	0.5022	0.001	***
bio12	0.24768	−0.96884	0.1002	0.029	*
bio5 / bio12	0.57304	0.81953	0.3101	0.001	***
wr	−0.77666	−0.62991	0.1976	0.001	***

climate for a Mediterranean country. The north-south differentiation found in the dendrogram does not appear in the Rivas-Martínez classification.

Bioclimatic Classification by Means of Self-Organizing Maps (SOM)

The SOM was divided into eight clusters (Fig. 6), which correspond with the elevational belts (Fig. 7):

Cluster 1 corresponds to the coastal plain of Central Albania, cluster 3 to the first reliefs near the coast, cluster 4 to the coastal hills, cluster 6 to the mountain belt, cluster 7 to the high areas of Northern Albania, cluster 5 to the dry area of the great lakes Ohrid and Prespa, cluster 2 distinguishes the northern coastal strip of Albania on the border with Montenegro characterized by high rainfall, and cluster 8 corresponds to the Albanian Alps on the border with Montenegro.
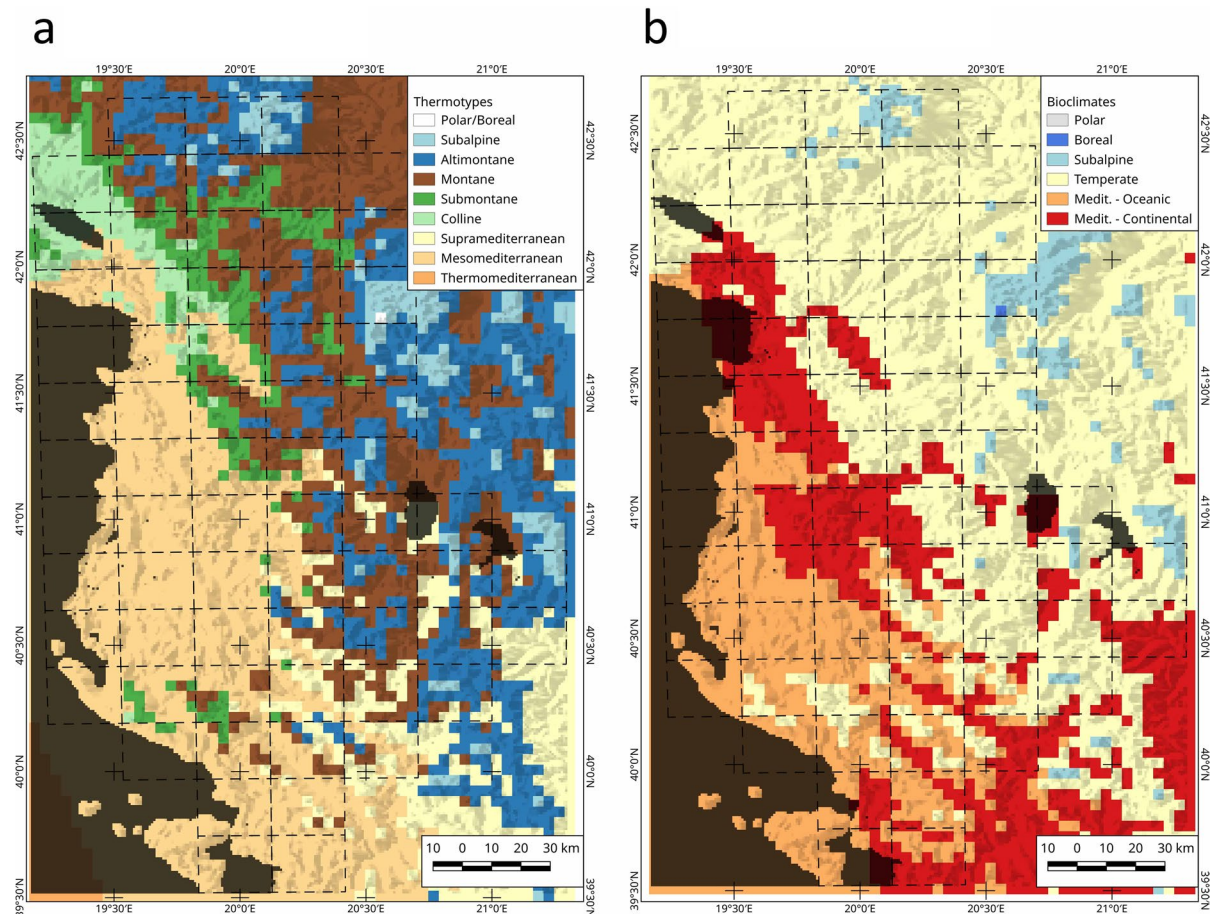
Fig. 5 **A** – Maps of Rivas-Martínez thermotypes and **b** – bioclimates according to Rivas-Martínez indices

The learning curve of the SOM (Fig. S2 in the electronic supplementary material) shows a satisfactory progression of classification accuracy, with mean sample distances to the closest node reaching a plateau at $d \approx 0.03$.

Overall, the SOM returns a classification somewhat similar to Rivas-Martínez's. As far as climate variables are concerned (Figs. 8, 9, Fig. S3 in the electronic supplementary material), Wr and bio5 / bio12 have a right-to-left and a left-to-right trend in the map grid, respectively, that explains much of the classification outcomes. Variable bio12 (annual rainfall) has a top left to bottom right pattern that correlates well with clusters 2, 3 and 4. Variable bio5 is higher in the top left and bottom right corners whereas bio7 (annual temperature range) goes from the bottom right corner towards the top left corner and corresponds to clusters 5, 6 and 7, respectively.

The percentage of the various chorotypes corresponds well to the clusters of the floristic analysis (Fig. S4 in the electronic supplementary material): (1) European and circumboreal species are more frequent in the inland belt than in the coastal belt and correspond with the first subdivision of the dendrogram. (2) Sub-tropicals exhibit an opposite trend to boreal species. They are widespread in the coastal plain but lacking in southern Albania, which is partially differentiated from this chorotype. (3) Southeastern European chorotypes are concentrated in cluster g (Albanian Alps), which is distinctly alpine. (4) Atlantic species do not show a particular distribution pattern. They have low percentages in all squares. (5) Mediterranean species (steno-Mediterranean + eu-Mediterranean) are widespread in the coastal and Mediterranean areas but decline in the inland belt in correspondence with the first subdivision of the

Fig. 6 Distribution of the
eight clusters obtained by
the SOM method
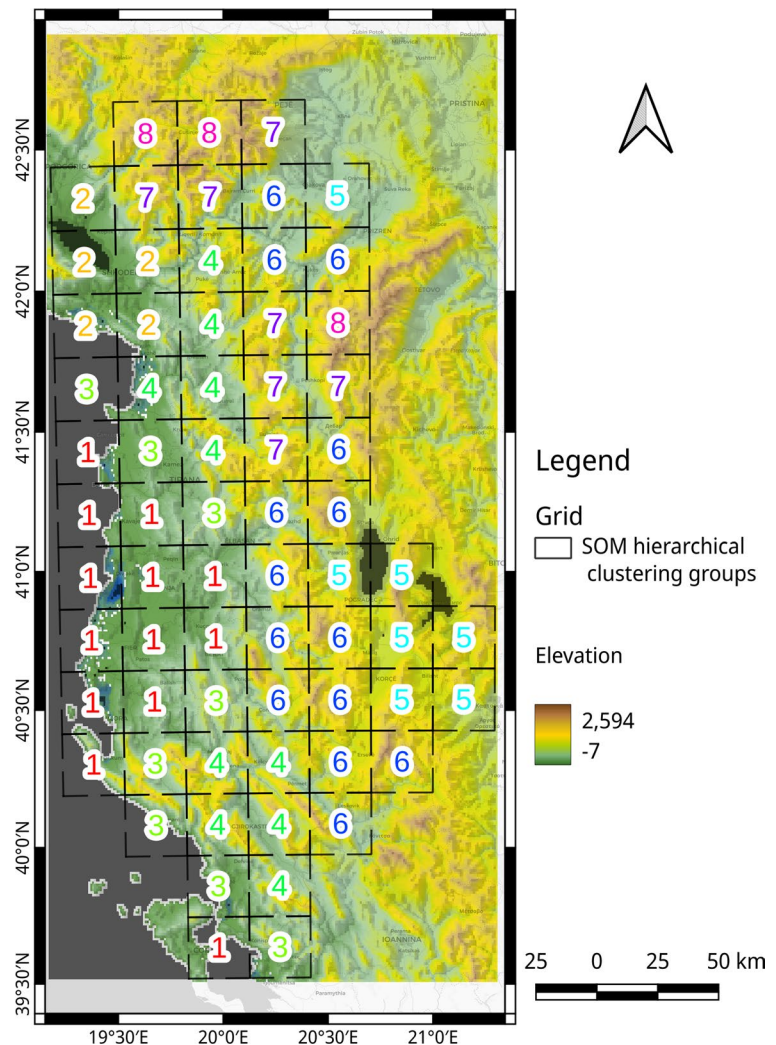
dendrogram. (6) Balkan species have a bimodal inter-exposure distribution. They are prevalent in southern Albania but increase again in the north in correspondence with the extensive ophiolite outcrops. (7) Pontic elements are mainly concentrated in a part of cluster h, which has the most continental character in Albania. (8) Cosmopolitan species are widespread in the central-northern coastal area, which is the most anthropized. (9) The Arctic–Alpine species are concentrated within cluster g (Albanian Alps); the alpine character of the Albanian Alps has long been known. (10) The richest part of endemic species is in the mountainous part of south Albania.

To better characterize the rich Albanian Balkan component, we made a distinction between the South Balkan elements (i.e. species distributed in Greece, Bulgaria and Macedonia) and the wider Balkan elements.

The South Balkans also present an interesting pattern. They are clearly more frequent in southern Albania within clusters c, d and e of the floristic analysis, although they are locally frequent in some individual cells in northern Albania.

Discussion

Markgraf (1927, 1932) identified four elevational belts and two sectors. The four elevational belts are the belt of the maquis and shibljak (corresponding to what is currently called meso-Mediterranean and thermo-Mediterranean), the belt of dry woods
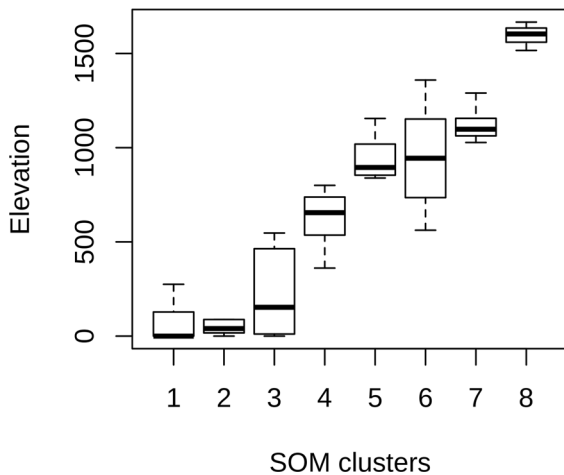
Fig. 7 Range of elevations in the different clusters retrieved from the SOM, represented through boxplots

(supra-Mediterranean), the belt of cloud forest (montane) and the alpine belt (subalpine according to current nomenclature). Markgraf regarded the first and second belts as belonging to the Mediterranean region and the montane and subalpine belts to the Central European region, a subdivision followed by all subsequent authors (e.g. EEA 2000; Olson et al. 2001). Markgraf also divided Albania into a northern and a southern sector spanned by the Mali i Thatë – Guri i Topit-Vlora line (the 'grosse Querlinie', henceforward GQ).

Our results agree strikingly with the subdivision of Markgraf. We identified four elevational belts, which are in turn divided according to the position north or south of a line corresponding precisely with the transverse line of Markgraf (Figs. 2B, 3b). The eight clusters that we have retrieved correspond with the dominant vegetation types of Albania: cluster a – *Carpinus orientalis* woodlands, cluster b – *Quercus cerris* woodlands, cluster c – *Quercus coccifera–Quercus ithaburensis* woodlands, cluster d – *Quercus coccifera* woodlands, cluster e – *Pinus heldreichii* woodlands, cluster f – *Quercus petraea–Fagus sylvatica* woodlands, cluster g – northern alpine vegetation, and cluster h – alpine vegetation of eastern (continental) Albania (see e.g. Mitrushi 1955, 1966). Due to the scale of this analysis, these vegetation types can be found elsewhere in Albania. For instance, the Mediterranean element of the Albanian flora extends in some parts deep into continental areas, but such

details cannot be captured with quadrats 25 km in size. Cluster g (Albanian Alps) is differentiated chorologically by a high percentage of Arctic–Alpine and South-European orophytic clusters that identify the Alpine character recognized by all authors. This differentiation is less pronounced in the dendrogram. It is difficult to understand whether this result depends on the generally punctiform character of many Albanian endemics (e.g. *Astragalus autranii*, *Campanula aureliana*, *Oxytropis tomoriensis*, *Tulipa albanica*) or whether it indicates an individuality of Albania with respect to the rest of the Balkans, as some endemics distributed throughout the Albanian territory would suggest, such as, for example, *Festucopsis serpentinii* and *Forsythia europaea (*Bogdanović et al. 2015; Tan et al. 2015; Shuka et al. 2017, 2020). A comparison on a Balkan scale only can solve this problem.

The Rivas-Martínez bioclimatic classification is widely used in the phytogeographical classification of Europe (Blasi 1996; Mucina et al. 2016). The bioclimatic classification of Albania corresponds in part with the subdivisions identified on a floristic basis. In particular, the vegetation belts are consistent with the floristic classification and are even more detailed. For example, it separates the submontane and altimontane, a subdivision reported in the study of beech woodlands by Fanelli et al. (2020), where lower situated beech woodlands are clearly distinguished from ones growing at higher elevations. At high elevations, Rivas-Martínez distinguishes a subalpine and a boreal belt that do not appear at the rather coarse spatial resolution of our grid.

On the other hand, the Rivas-Martínez regionalization is not effective in differentiating the south from the north of Albania. The clear distinction between clusters c and b and between clusters g and h, which belong to the same elevational belt but occur on different sides of the GQ, disappears once the regionalization of Rivas-Martínez is applied. In other words, the bioclimatic model of Rivas-Martínez is extremely effective for the inner Central European part of Albania, allowing fine distinctions that also correspond with phytosociological differentiations. At the same time, however, it fails to capture the peculiarities of the central-eastern Mediterranean characteristic of the coastal parts of Albania.

The analysis based on the SOM is similar to that of Rivas-Martínez. The clusters identified by the SOM method (Fig. 8) are related to elevation and thus to
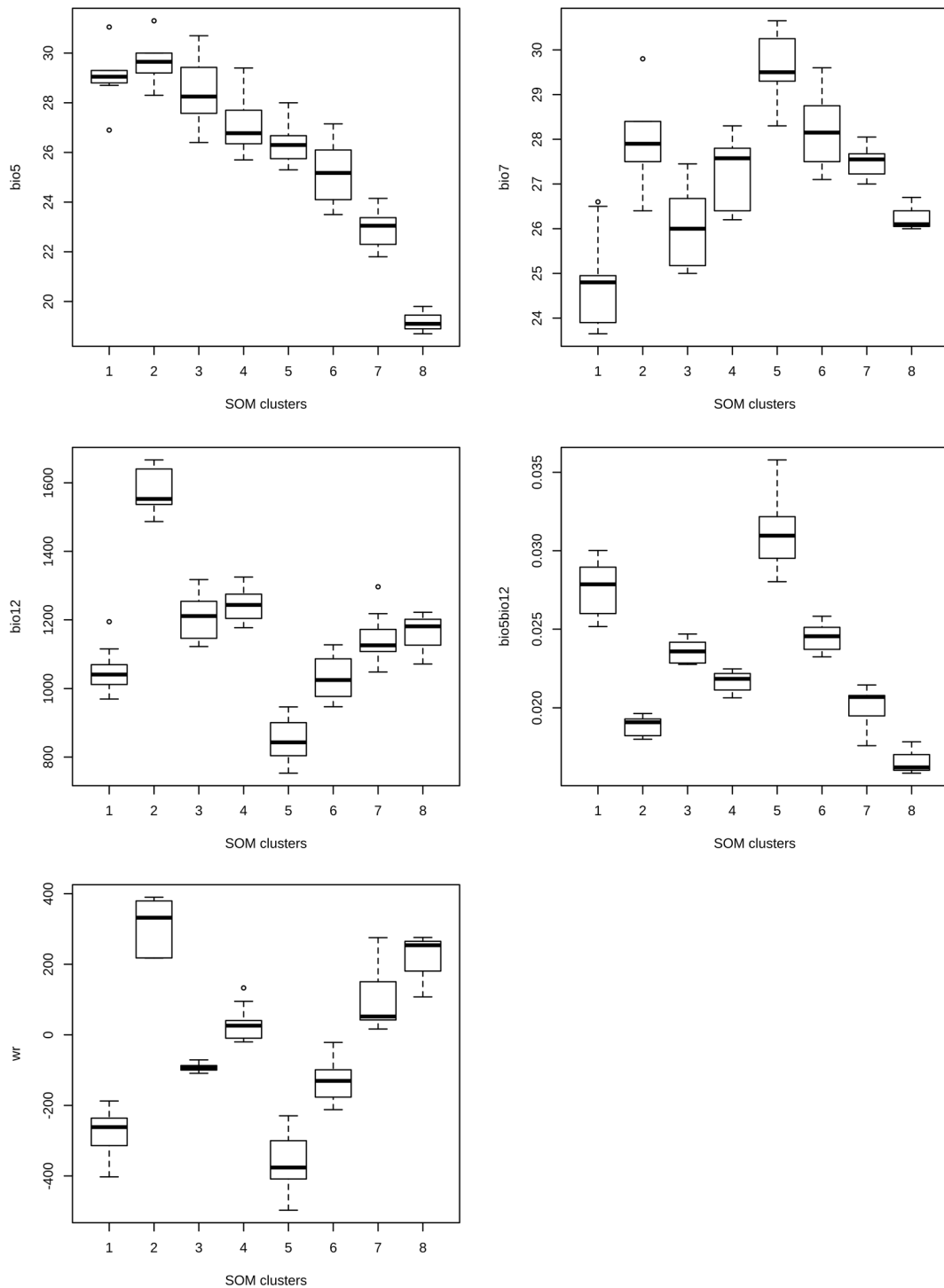
Fig. 8 Distribution of the various clusters retrieved from the SOM analysis according to environmental variables
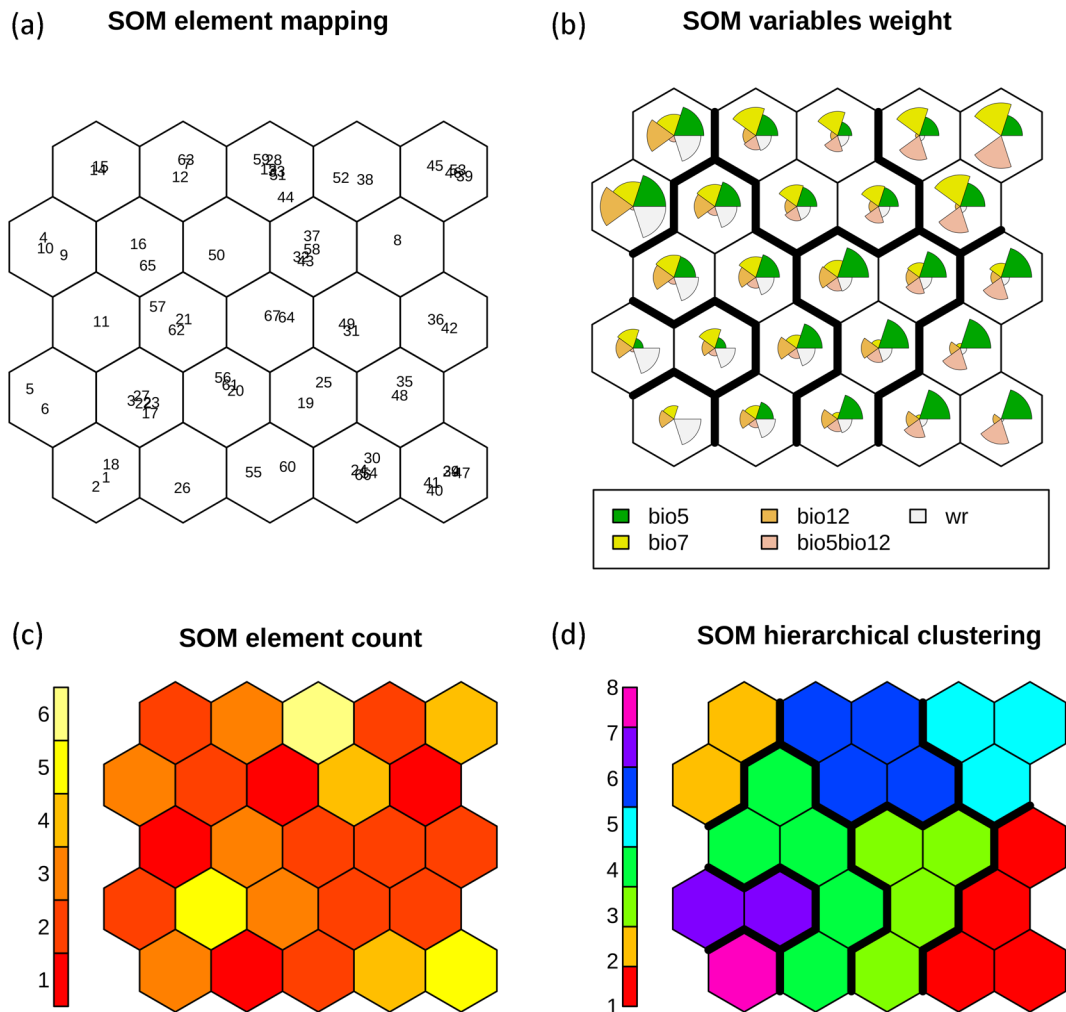
Fig. 9 **A**: Mapping of the 25x25 km cells in the SOM nodes; **b**: Fan diagram representing the magnitude of each predictor in the SOM nodes; **c**: Element count in the SOM nodes; d: Outcome of the agglomerative hierarchical clustering on SOM nodes

elevational belts (Fig. 7). On the other side, SOM cannot identify the north–south distinction across the GQ (Fig. 10).

The GQ, identified by Markgraf and confirmed by our quantitative analysis, is particularly evident in Albania. However, its meaning is not limited to Albania but extends far beyond this country's limits (Fig. 10). Kosanin (1918) identified Northern Macedonia as a southern area with a Mediterranean character and a northern region with a purely Central European character; the boundary between the two areas is the continuation of Markgraf's GQ. In Bulgaria, Tzonev et al. (2013) applied the bioclimatic categories of Rivas-Martínez, identifying a clear division between

the southern part and the rest of Bulgaria. Here, too, the dividing line between the two domains continues Markgraf's line. In Italy, where the distinction between the Adriatic and Tyrrhenian sides tends to be emphasized, Attorre et al. (2014) highlights a clear difference between Calabria, Sicily, Sardinia, Salento and the rest of Italy from a phytogeographical point of view. Again, the geographical boundary is a continuation of the GQ. In the Iberian Peninsula, great importance has traditionally been attached to the boundary between the Eurasian, which is essentially limited to the Cantabrian Cordillera, the northern Pyrenees and the Mediterranean. However, Rivas-Martínez clearly distinguishes the thermo-Mediterranean vegetation

Fig. 10 Extension of the line separating southern from northern / central Albanian areas (the grosse Querlinie, GQ) from other Mediterranean countries (Macedonia: Bulgaria: Tzonev et al. 2013; Italy: Attorre et al. 2014; Spain: Rivas-Martínez et al. 2011)

of Andalusia from the less thermophilic vegetation of Extremadura and Castile. This distinction corresponds to a latitude approximately similar to that of Markgraf.

Looking at other continents, it is striking that in California, only the area south of the 32nd parallel is considered truly Mediterranean (Castri and Mooney 1973). The sectors to the north are not considered Mediterranean at all. This latitude corresponds exactly with the line we followed from Albania and found both east and westwards.

We have repeatedly underlined how Albania is subdivided according to elevational plans, separating the lower coastal areas from the higher inland areas, and secondly in a north–south direction separating the central-northern sectors from the southern ones. When we investigate the bioclimatic factors that explain this subdivision, we notice that the first

one, in an elevational sense, is perfectly explained by several climatic factors. In particular, bio5 is strongly correlated with the coast–inland gradient (Fig. 4), which also corresponds largely with the Rivas-Martínez bioclimatic model and the SOM of climatic variables.

Instead, the north–south differentiation that we have seen to have an importance far beyond the small Albanian territory is difficult to explain in purely bioclimatic terms. On the NMDS diagram (Fig. 4), this differentiation appears to be related to bio7 and bio12, but both the Rivas-Martínez classification and the SOM based on bioclimatic variables fail to distinguish central-northern Albania from southern Albania.

This discrepancy can be explained in two ways: (1) It is possible that the generally used climatic indices are not the most suitable to explain the north–south

Albanian differentiation; it is pretty clear that the central-northern territories are drier and more continental than the southern ones (as evidenced by the NMDS biplot), but it is difficult to summarize this differentiation quantitatively. Studying the thermo-hydrometric diagrams (Fig. S1), the stations north of the GQ are characterized by high rainfall at the end of the year and low rainfall at the beginning of the year, while those below the GQ have high rainfall at both the beginning and the end of the year. This pattern (the distribution of rainfall outside the dry season) deserves further investigation, perhaps even at a larger scale than the Albanian one; it is a pattern that has not been considered in any of the various indices in the literature to characterize the bioclimate. It is also possible that the dominant air masses are different in northern-central and southern Albania (Isidorov 1955).

(1) It is still possible that the north-south differentiation is not completely based on the current climate but that it has an origin in past events, evolutionary history and the climate of ancient geological epochs. The fact that the southern Balkan component (i.e. essentially an endemic element, albeit more broadly understood than the concept of endemism in the strict sense) correlates very well with the separation between north and south-central Albania (Fig. S2) might suggest that the areas south of the GQ experienced less climatic variability in the transition from the Tertiary to the Holocene. Suggestive of this is the fact that Italian regions south of the continuation of the GQ (i.e. Sicily, Sardinia, Calabria and Salento) are characterized by a particularly high rate of endemism.

Whether the north–south distinction is due to current climate or evolutionary history requires phylogenetic and phylogeographic studies, and it should be borne in mind that the two hypotheses are not mutually exclusive but could interact with each other.

In this study, the standard approach in phytogeography was followed to (a) identify the floristic and vegetation subdivisions and (b) find the correspondence of this subdivision with one or more environmental factors. Some authors believe there is an immediate correspondence between environmental variables and vegetation. They follow the opposite approach, meaning that they start from a subdivision of the territory based on environmental variables and only then look for correspondences of this subdivision with floristic features of the territory. This

approach fails in Albania because it may work very well for the elevational gradient but not with the north–south gradient, which is evident and difficult to explain in terms of simple bioclimatic indices. This is particularly evident with the SOM classification.

As regards the classification approaches, our results highlight that the preliminary dimensionality reduction performed using SOM better combines the several environmental variables, preserving the similarity among the grid cells and providing a more consistent result compared to standard cluster analysis. On the other hand, SOM analysis of species data does not add any further relevant information, probably because of the nature of the data (SOM does not effectively handle presence / absence data). This seems to suggest that using different methods is in fact the most effective approach in our case, as the two datasets considered are completely different in terms of the properties observed and the structure of the information.

Conclusion

The foundations of Albanian phytogeography were laid by the oft-quoted Friedrich Markgraf in his seminal work *Pflanzengeographie von Mittelalbanien* (Markgraf 1932). Albania is divided in a southwest–north-east direction into elevational planes and a north–south direction into a transitional Mediterranean area in the north and a narrow Mediterranean region in the south.

This work reformulates Markgraf's intuitions in a more modern phytogeographical and bioclimatic framework, considering the acquisitions of the second half of the twentieth century, and clarifies the significance of Albania in the broader Mediterranean context. On the one hand, there is a need to adapt or perhaps reformulate Rivas-Martínez's biogeographical scheme in the territories of Southeastern Europe, where the Spanish author's schemes do not always capture the fundamental phytogeographical limits. On the other, the importance of what Markgraf called GQ emerges, which separates the sub-Mediterranean / Mediterranean transitional territories from the Mediterranean proper, also referred to as the eu-Mediterranean. This distinction is of little value in the western Mediterranean, influenced substantially by the Azores anticyclone,

while it is of extreme importance in the central-eastern Mediterranean between the Mediterranean proper and Central Europe proper. The great confusion around this problem stems precisely from not recognizing that the Mediterranean can be defined in two senses, which in Albania correspond to the differentiation between the coast and the inland and in a second direction between the north and south. This orthogonality of the two boundaries makes their interpretation particularly easy in this country. In other Mediterranean countries, however, the two different definitions of the Mediterranean partially overlap, so they are difficult to discern.

In addition to representing a crucial phytogeographical junction for understanding the Mediterranean and the Balkans, Albania is confirmed as a territory of very high biodiversity, as it includes both purely Central European and purely Mediterranean aspects. In his first geographic survey, Markgraf defined Albania as a country on the border of the Mediterranean, which therefore receives both the floristic contributions of Europe and those of the Mesogeum. Elsewhere, this situation is perhaps found only in Liguria, another territory with very high biodiversity and endemism that maintains some phytogeographic links with Albania.

Declarations

References

Attorre F, Francesconi F, De Sanctis M, Alfò M, Martella F, Valenti R, Vitale M (2014) Classifying and mapping potential distribution of forest types using a finite mixture model. *Folia Geobot* 49: 313–335

Baldacci A (1896) Rivista critica della collezione botanica fatta nel 1894 in Albania [Critical revision of the botanical collection made in 1894 in Albania]. *Bull Herb Boissier* 4:609–653

Barina Z (ed.) (2017) *Distribution atlas of vascular plants in Albania*. Hungarian Natural History Museum, Budapest

Blasi C (1996) Il fitoclima d'Italia [Phytoclimate of Italy]. *Giorn Bot Ital* 130:166–176

Bogdanović S, Rešetnik I, Brullo S, Shuka L (2015) *Campanula aureliana* (Campanulaceae), a new species from Albania. *Pl Syst Evol* 301:1555–1567

Bortolotti V, Kodra A, Marroni M, Mustafa F, Pandolfi L, Principi G, Saccani E (1996) Geology and petrology of ophiolitic sequences in the Mirdita region (northern Albania). *Ofioliti* 2:3–20

Castri F, Mooney H (1973) Mediterranean type ecosystem. Springer-Verlag, Berlin

Chon T-S, Park YS, Moon KH, Cha EY (1996) Patternizing communities by using an artificial neural network. *Ecol Modelling* 90:69–78

De Sanctis M, Fanelli G, Gjeta E, Mullaj A, Attorre F (2018) The forest communities of Shebenik-Jabllanicë National Park (Central Albania). *Phytocoenologia* 48:51–76

Dhakal T., Kim T. S., Lee D. H., Jang G. S. (2023) Examining global warming factors using self-organizing map and Granger causality network: a case from South Korea. *Ecol Processes* 12:9

EEA (2000) Digital map of European ecological regions. Available at https://www.eea.europa.eu/data-and-maps/data/digital-map-of-european-ecological-regions[Accessed 12 March 2022]

Euro+Med (2006) Euro+Med PlantBase – the information resource for Euro-Mediterranean plant diversity. Available at http://ww2.bgbm.org/EuroPlusMed [Accessed 20 March 2022]

Emberger L (1945) Climate biogeographic classification. *Collection of Botanical Geological and Zoological Laboratories Works. Volume 7*. Montpellier Faculty of Sciences, pp 3-43

Fanelli G, Hoda P, Mersinllari M, Mahmutaj E, Attorre F, Farcomeni A et al. (2020) Phytosociological overview of the Fagus and Corylus forests in Albania. *Veg Classif Surv* 1:175–189

Fick SE, Hijmans RJ (2017) WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. *Int J Climatol* 37:4302–4315

Hijmans RJ, Cameron SE, Parra JL, Jones PG, Jarvis A (2005) Very high resolution interpolated climate surfaces for global land areas. *Int J Climatol* 25:1965–1978

Isidorov GC (1955) *Klima e Shqiperisë* [Climate of Albania]. Instituti i Shkencave, Tiranë

Kim TS, Dhakal T, Kim SH, Lee JH, Kim SJ, Jang GS (2023) Examining village characteristics for forest management using self-and geographic self-organizing maps: a case from the Baekdudaegan mountain range network in Korea. *Ecol Indicators* 148:110070

Kohonen T (2001) Learning vector quantization. In Kohonen T (ed.) *Self-organizing maps*. Springer-Verlag, Berlin, pp 245–261

Lack HW, Barina Z (2020) The early botanical exploration of Albania (1839–1945). *Willdenowia* 50:519–558

Markgraf F (1927) An den Grenzen des Mittelmeergebiets Pflanzengeographie von Mittelalbanien [On the borders of the Mediterranean area Phytogeography of Central Albania]. *Repert Spec Nov Regni Veg Beih* 45:1–217

Markgraf F (1932) Pflanzengeographie von Albanien [Phytogeographya of Albania]. *Biblioth Bot* 26:1–132

Meco S, Aliaj S (2000) *Geology of Albania*. Gebruder Borntraeger, Berlin

Mitrushi I (1955) *Drurët dhe Shkurret e Shqiperisë* [Trees and Shrubs of Albania]. Instituti i Shkencave, Tiranë

Mitrushi I (1966) *Dendroflora e Shqiperise* [Dendroflora of Albania]. Fakulteti i Shkencave Te Natyres Universiteti Shteteror, Tirane

Moral FJ (2010) Comparison of different geostatistical approaches to map climate variables: application to precipitation. *Int J Climatol* 30:620-631

Mucina L, Bültmann H, Dierßen K, Theurillat JP, Raus T, Čarni A et al. (2016) Vegetation of Europe: hierarchical floristic classification system of vascular plant, bryophyte, lichen, and algal communities. *Appl Veg Sci* 19:3–264

Murtagh F, Legendre P (2014) Ward's hierarchical agglomerative clustering method: Which algorithms implement Ward's criterion? *J Classific* 31:274–295

Olson DM, Dinerstein E, Wikramanayake ED, Burgess ND, Powell GV, Underwood EC, Kassem KR (2001) Terrestrial ecoregions of the world: a new map of life on earth: A new global map of terrestrial ecoregions provides an innovative tool for conserving biodiversity. *BioScience* 51:933–938

Pignatti S (1982) *Flora d'Italia* [Flora of Italy]. Edagricole, Bologna

R Core Team (2019). R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Available at https://www.R-project.org.

Rivas-Martínez S, Rivas-Sáenz S, Penas-Merino A (2011) Worldwide bioclimatic classification system. *Global Geobot* 1:1–638

Sekulić A, Kilibarda M, Protić D, Tadić MP, Bajat B (2020) Spatio-temporal regression kriging model of mean daily temperature for Croatia. *Theor Appl Climatol* 140:101–114

Shuka L, Tan K, Vold G (2017) Report 56: *Oxytropis tomoriensis*. In Vladimirov V, Aybeke M, Tan K (ed.) *New floristic record in the Balkans: 32*. *Phytol Balcan* 23:131–133

Shuka D, Tan K, Hallaçi B, Shuka L (2020) Additions to the flora of North Albania. *Phytol Balcan* 26:507–512

Tan K, Shuka L, Gjeta E, Vold G (2015) The rediscovery of *Astragalus autranii* (Fabaceae) on Mt Tomor, south central Albania. *Phytotaxa* 234:83–89

Tedesco M, Fettweis X (2020) Unprecedented atmospheric conditions (1948–2019) drive the 2019 exceptional melting season over the Greenland ice sheet. *Cryosphere* 14:1209–1223

Topi I, Mankolli H (2012) The climate indicators and vegetation in the central area of Albania. *BALWOIS 2012 Conference*. Ohrid, Republic of Macedonia, 27 May, 2 June 2012, paper 145

Tremblay A, Meshi A, Deschamps T, Goulet F, Goulet N (2015) The Vardar zone as a suture for the Mirdita ophiolites, Albania: constraints from the structural analysis of the Korabi-Pelagonia zone. *Tectonics* 34:352–375

Tzonev R, Pavlova D, Sánchez-Mata D, de la Fuente V (2013) Contribution to the knowledge of Bulgarian serpentine grasslands and their relationships with Balkan serpentine syntaxa. *Pl Biosyst* 147:955–969

Vangjeli J, Ruci B, Mullaj A, Xhulaj M (2000a) Flora e vegetazione in Albania. In Marchiori S, De Castro F, Myrta A (eds) *La cooperazione Italo-Albanese per La valorizzazione della biodiversità*. CIHEAM 2000a, Bari, pp 51–66

Vangjeli J, Ruci B, Mullaj A, Paparisto K, Qosja Xh (2000b) *Flora e Shqipërisë* [Flora of Albania], Volume IV. Akademia e Shkencave e Republikës së Shqipërisë, Tiranë

Vangjeli J (2015) *Atlas i flores e Shqiperise, Volume I: Pteridophyta-Apiaceae*. Akademia e Shkencave e Shqiperise, Tiranë

Walter H, Lieth H (1967) *Klimadiagramamm-Weltatlos* [Atlas of climate diagrams of the world]. Gustav Fischer, Jena

Zickel RE, Iwaskiw WR, Keefe EK (1994) Albania: a country study Library of Congress Federal Research Division. Available at https://lccnlocgov/93042885