



RESEARCH PAPER

Journal of
Biogeography

WILEY

Network- and distance-based methods in bioregionalization processes at regional scale: An application to the terrestrial mammals of Iran

Gholam Hosein Yusefi^{1,2,3} | Kamran Safi^{4,5} | José Carlos Brito^{1,2}

¹CIBIO/InBIO, Centro de Investigação em Biodiversidade e Recursos Genéticos da Universidade do Porto, Vairão, Portugal

²Departamento de Biologia da Faculdade de Ciências da, Universidade do Porto, Porto, Portugal

³Mohitban Society, Tehran, Iran

⁴Department of Migration, Max Planck Institute of Animal Behavior, Radolfzell, Germany

⁵Department of Biology, University of Konstanz, Konstanz, Germany

Correspondence

Gholam Hosein Yusefi, CIBIO/InBIO, Centro de Investigação em Biodiversidade e Recursos Genéticos da Universidade do Porto, R. Padre Armando Quintas, 4485-661 Vairão, Portugal.
Email: gh.yusefi@cibio.up.pt

Funding information

Norte Portugal Regional Operational Programme (NORTE2020), Grant/Award Number: AGRIGEN-NORTE-01-0145-FEDER-000007; Portuguese Foundation of Science and Technology (FCT), Grant/Award Number: IF/00459/2013 and PD/BD/52605/2014

Handling Editor: Şerban Procheş

Abstract

Aim: In recent years, novel approaches have been proposed to improve current bioregionalization methods, but these have not been thoroughly compared. We assessed the applicability of the recently developed network-based clustering method (Infomap algorithm) in bioregionalization analysis at regional spatial scales and compared the results with commonly used distance-based methods (hierarchical clustering algorithm). We also identified climate regions by using a model-based cluster analysis (Gaussian algorithm). Finally, we quantified the representation of climate regions and bioregions in current protected areas (PAs).

Location: Iran.

Taxa: Terrestrial mammals.

Methods: To define bioregions we used the Infomap algorithm and distance-based clustering methods based on species distribution data (over 14,000 occurrence records for 188 species). The Infomap algorithm was applied using the interactive web application "INFOMAP BIOREGIONS" and the distance-based clustering was based on unweighted pair-group method using arithmetic averages (UPGMA). To identify climate regions we used principal components analysis and a model-based cluster analysis both based on 15 climatic variables as well as a terrain ruggedness index.

Results: The Infomap algorithm detected nine biogeographical units: seven bioregions and two transition zones. The distance-based method suggested five bioregions. The identified bioregions differed between methods with some consistent spatial patterns across methods. Temperature and precipitation explained 85.8% of the environmental variation. Eight climate regions were identified. In general, climate variation and bioregional patterns are currently poorly represented in PAs (<25% coverage).

Main conclusions: The network-based method allowed identifying bioregions at regional scale and was apparently more sensitive than the conventional distance-based method. The detection of transition zones by the Infomap algorithm was an advantage, and stressed the fact that the distribution of Iranian mammalian fauna is complex, especially in the south-eastern part where contact areas between several bioregions are found. The identified bioregions (especially the distance-based bioregions) and climate regions tended to match well with previous bioregionalization



studies and the global terrestrial ecoregions. When thoroughly compared and understood, bioregions and climate regions provide a framework for regional biodiversity conservation planning.

KEYWORDS

beta diversity, biogeographical regionalization, climate regions, hierarchical clustering algorithm, Infomap Bioregions, map equation, Palaearctic, Saharo-Arabian, transition zone

1 | INTRODUCTION

Knowledge about regional patterns in biodiversity distribution forms the basis for numerous lines of theoretical research, and provides valuable information for conservation planning (Richardson & Whittaker, 2010; Whittaker et al., 2005). Bioregionalization reveals how different sets of species are spatially grouped (Morrone, 2018), by displaying the relationships among regions grouped into geographically meaningful clusters, i.e. biogeographical regions or bioregions (Escalante, 2017). Bioregions are important units for research in biogeography, ecology and evolution (Holt et al., 2013; Kreft & Jetz, 2010), and for biodiversity conservation (Dinerstein et al., 2017; Riddle, Ladle, Lourie, & Whittaker, 2011). Regional patterns of species distributions in turn correlate with climate. As such, identification of climate regions, i.e. areas with geographically distinct climatic patterns, can help defining priority areas for protection of climatic refugia or evolutionary processes (Brito et al., 2016; Michalak, Lawler, Roberts, & Carroll, 2018). Bioregions and climate regions can serve as conservation units and provide a framework for regional biodiversity conservation planning. Climate regions are especially important where a lack of precise data on species distribution may limit species-based conservation prioritization (Brito et al., 2016).

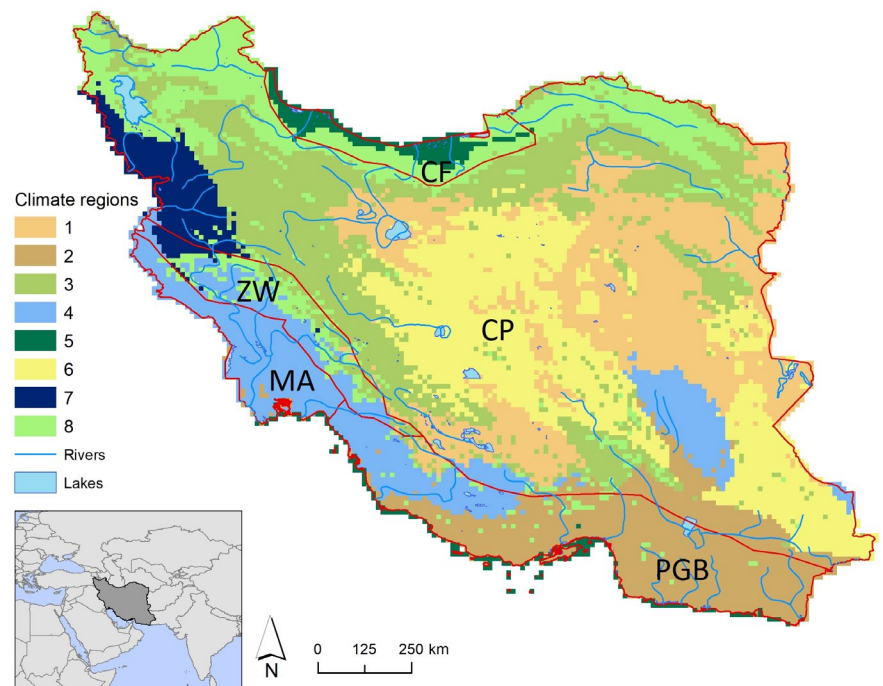
Bioregionalization exercises based on species occurrence data often rely on species turnover and distance-based clustering methods (e.g. González-Orozco, Laffan, Knerr, & Miller, 2013; He, Kreft, Gao, Wang, & Jiang, 2017; Kreft & Jetz, 2010). Recently, a network-based method based on a community-detection algorithm known as Infomap has been put forward for identifying bioregions (Vilhena & Antonelli, 2015). This method clusters both species and grid cells together, as opposed to clustering grid cells with distances inversely proportional to the number of shared species (as in distance-based methods) (Vilhena & Antonelli, 2015). The advantages and better performance of network- over distance-based methods are widely discussed in the literature (e.g. Bloomfield, Knerr, & Encinas-Viso, 2018; Ferrari, 2018; Sidor et al., 2013; Vilhena & Antonelli, 2015). However, network-based methods have been mostly used to identify bioregions at global and continental scale (see Bernardo-Madrid et al., 2019; Bloomfield et al., 2018; Droissart et al., 2018; Edler, Guedes, Zizka, Rosvall, & Antonelli, 2017; Hazzi, Moreno, Ortiz-Movliav, & Palacio, 2018; Vilhena & Antonelli, 2015) and their application potential at regional and

local spatial scales remains unknown. More importantly, comparisons in bioregionalization processes between network- and distance-based methods are available at continental (e.g. Bloomfield et al., 2018) but not regional scales. This is particularly the case for biogeographic crossroads or biogeographical transitional areas (i.e. the boundaries between biogeographical regions or realms, where biotic elements overlap and contain a mixture of different geographical components), which seem to have long been neglected in biogeographical studies (Ferro & Morrone, 2014). Transitional areas harbour high levels of both species and phylogenetic diversity (Ferro & Morrone, 2014; Morrone, 2009), and are of crucial importance for evolutionary and ecological research, as well as for conservation.

Iran is a prime example of a biogeographical transitional area located in the intersection of three main biogeographical realms, Palaearctic, Saharo-Arabian and Oriental (Ficetola, Mazel, & Thuiller, 2017; Holt et al., 2013). At a regional scale, it is a proverbial bridge between the western Mediterranean and Arabian Peninsula, Central Asia and India. Iran displays high physiographic complexity, encompassing seven terrestrial biomes and 19 ecoregions (Table S1 in Appendix S1; Dinerstein et al., 2017), and a wide variety of ecosystems and habitat types, from temperate forests to extreme deserts (Firouz, 2005; Zehzad, Kiabi, & Madjnoonian, 2002; Zohary, 1973). From the biodiversity viewpoint, Iran is a rich region within western Asia (Firouz, 2005), partially classified as global biodiversity hotspot (Mittermeier, Turner, Larsen, Brooks, & Gascon, 2011). After the seminal zoogeographical division of Iran by Blanford (1876), there were multiple bioregionalization exercises based on plants (Frey & Probst, 1986; Zohary, 1963, 1973), birds (Zarudny, 1911), and reptiles (Anderson, 1968). Still, the bioregionalization of Iran is underexplored due to limitations in the availability of high-resolution species distribution data. The few bioregionalization studies available (e.g. Blanford, 1876) were expert-based and restricted to qualitative descriptions of species diversity and distribution (Figure 1). As such, a quantitative assessment of regional biogeographical units is necessary to uncover the representative array of major distinct ecosystems, vital for conservation planning. Likewise, quantifying representation levels of distinct biogeographical units and climate regions in the current network of Protected Areas (PAs) (covering c. 10% of Iran; DoE-GIS, 2016) is essential to flag potential representation gaps (Brito et al., 2016; Dinerstein et al., 2017).

Here, we conducted a regional-scale analysis to identify biogeographical units of Iran. To achieve this, we used a large mammal

FIGURE 1 The study area and identified climate regions (eight units): 1—S Khorasan-Sistan; 2—Makran-Hormuzgan-S Fars; 3—S Alborz-E Zagros-Kerman Mts; 4—Persian Gulf Shores-Khuzestan-Lut; 5—Hyrcanian; 6—Central Basin-N Baluchestan; 7—Kordestan; and 8—N Alborz-Caucasus-Kopet Dag Mts. The red lines indicate the five zoogeographical regions identified by Blanford (1876): Central Plateau (CP), Caspian Forest (or Hyrcanian) (CF), Zagros Woodland (ZW), Mesopotamian Area (MA) and the Persian Gulf and Baluchestan shores (PGB). The map is based on a 0.1° grid cell [Colour figure can be viewed at wileyonlinelibrary.com]



assemblage of recent database containing over 14,000 mammalian species occurrences (Yusefi, Faizolahi, Darvish, Safi, & Brito, 2019), and network theory and clustering-based approaches. We compared the results of these two methods with regards to the number of identified bioregions and transitions zones, and their spatial distribution. We also quantified climate regions, and evaluated the levels of protection of climate regions and bioregions.

2 | MATERIALS AND METHODS

2.1 | Study area

Iran (1,648,195 km²) is located in south-west Asia and displays two large mountain ranges, extending northwest-northeast (Alborz) and northwest to south-central (Zagros), that separate the humid Caspian Sea coast and Anatolian/Mesopotamian region from the arid Central Basin (Figure 1; Figure S1 in Appendix S1). The Central Basin covers most of the country and displays high physiographic complexity (Firouz, 2005).

2.2 | Species distribution data

Distribution data for terrestrial mammals were retained from a database of 14,910 unique and georeferenced historical and current occurrence records representing 188 native species, seven orders, and 34 families (Yusefi et al., 2019). This database is the most comprehensive dataset on Iran's mammals and was built from an extensive review of published and unpublished observations (details in Yusefi et al., 2019). To avoid biases resulting from insufficient or uneven sampling, we limited analyses to grid cells with more than

five species or with $\geq 50\%$ of their area within the country borders (Rueda, Rodríguez, & Hawkins, 2013).

2.3 | Identification of bioregions

2.3.1 | Network-based method

The network-based method (Rosvall & Bergstrom, 2008) is implemented in the application "INFOMAP BIOREGIONS" (Edler et al., 2017; <http://bioregions.mapequation.org>). INFOMAP BIOREGIONS bins the species records (georeferenced point occurrences or range maps) into quadratic grid cells and generates a bipartite network of species and the grid cells. The network is clustered with Infomap clustering algorithm for bipartite networks where the geographical grid cells are merged into different clusters, which become the resulting bioregions (Edler et al., 2017). Infomap in turn works with respect to an information-theoretic map equation which attempts to simplify the network structure by using an efficient random walk process, and to minimize the length of code of a modular description of pathways on a network by using minimum description length principle (Rosvall & Bergstrom, 2008; Rosvall, Axelsson, & Bergstrom, 2009; for further details see Bernardo-Madrid et al., 2019). INFOMAP BIOREGIONS relies on an adaptive spatial resolution to reflect differences on data density, where sampling units vary in size according to record density (observations available) (Edler et al., 2017). Thus, we set the maximum cell size to 1° (which is convenient for the size of the study area; M. Rosvall, pers. comm.), and the minimum cell size was chosen between 1/8° and 1°, setting the limits for the adaptive resolution algorithm.



Maximum and minimum cell sizes refer to the largest and smallest size of the grid cells used. The minimum cell capacity, i.e. the minimum number of observations per cell, was set to 5, and we used maximum cell capacities (i.e. the number of records in a grid cell) from 10 to 300. Under the adaptive resolution approach, the algorithm subdivides grid cells with more records than the maximum cell capacity until it reaches the minimum cell size provided by the user (for additional details see Edler et al., 2017). The analysis was run using simulations of different combinations of minimum cell sizes and maximum cell capacities to ensure that the results were qualitatively unaffected by the values used, and to find the optimal number of clusters (M. Rosvall pers. comm.). Also, we used the recommend number of trials = 1, and number of cluster cost = 1.0. Given that the network-based method also allows to identify transition zones between bioregions (Bloomfield et al., 2018; Droissart et al., 2018; Vilhena & Antonelli, 2015), the biogeographical units retrieved may represent either distinct bioregions or transition zones. Here, we defined bioregions as clusters containing more than 30 species and more than 100 records, while clusters with fewer species or records were considered as transition zones (Droissart et al., 2018; M. Rosvall pers. comm.). INFOMAP BIOREGIONS also identifies the most common and the most indicative species in each grid cell and bioregion (Edler et al., 2017). The former corresponds to the species with larger number of occurrences and the latter are defined as the ratio between species frequency in one bioregion against all bioregions.

2.3.2 | Distance-based method

Simpson's dissimilarity index (β_{sim}) was used to quantify pairwise similarities among grid cells. The β_{sim} index calculates as: $\beta \sim 1 - \frac{a}{\min(b,c)+a}$, where a is the number of species common to two grid cells and b and c are the numbers of unique species (not shared). It ranges from 0.0 to 1.0, with higher values indicating higher dissimilarity (Koleff, Gaston, & Lennon, 2003). We chose Simpson's index because it reduces effects of species richness imbalances between samples (Kreft & Jetz, 2010). Clustering was conducted based on the unweighted pair-group method using arithmetic averages (UPGMA), as this algorithm is widely proven to produce robust models in agglomerative hierarchical clustering (Kreft & Jetz, 2010). Then, the optimal number of clusters was determined by using the Kelley-Gardner-Sutcliffe penalty function for a hierarchical cluster tree (Kelley, Gardner, & Sutcliffe, 1996). Data preparation and analyses were conducted in R 3.4.4 (R Development Core Team, 2018) (Table S1 in Appendix S2). Analyses were performed using species distributions at a spatial resolution of 1° grid cells (to be comparable with the network-based method). The model transfers the grid cells into a presence-absence data matrix. To assess scale effects on bioregion detection, additional analyses (results not shown) were done at smaller grid-cell resolutions (25 km × 25 km and 50 km × 50 km). Additionally, we quantified the level of spatial

overlap between bioregions identified by both methods using ArcView GIS.

2.4 | Climatic variables and identification of climate regions

Climatic variation was identified using principal component analysis (PCA) based on 16 environmental variables. The three first PCA components were the basis for a model-based cluster analysis, and model selection statistics were used to identify the best clustering solution. The level of spatial overlap of identified climate regions and bioregions was quantified (detailed explanations on these methods are provided in Appendix S3).

2.5 | Gap analysis

The level of representation of the identified climate regions and bioregions in areas under protection of Department of Environment of Iran was quantified using ArcView GIS. These areas are classified into three categories that include: National Parks ($n = 31$), Wildlife Refuges ($n = 49$), and PAs ($n = 166$) (Darvishsefat, 2006; DoE-GIS, 2016). Bioregions were downscaled from 1° to 0.1° (~10 km × 10 km) specifically for these analyses.

3 | RESULTS

3.1 | Bioregions

The simulations based on network-based method showed important effects of the chosen minimum cell size and the maximum cell capacity on the estimated number of clusters (Figure S1 in Appendix S4). The simulations suggested that the number of clusters converged when minimum cell sizes and maximum cell capacities used were maximized. In this case, the number of clusters became more consistent when we increased the maximum cell capacity from 10 to 300, and the minimum cell size from 1/8° to 1° (Figure S1 and Figure S2 in Appendix S4). A final list of seven bioregions and two transition zones was identified as the most likely number of clusters (Table 1; Figure 2). Of these, the large bioregions 1 and 2 largely matched the Alborz-Zagros mountains and Central Basin lowlands, respectively. The two transition zones were located in boundary areas, one where the Iranian Plateau and Arabian Peninsula meet, and the other between the two largest identified bioregions (1 and 2). The most indicative species and the common species of each bioregion were identified (Table S1 in Appendix S4). For instance, the most indicative species of bioregion 2 were desert-adapted species: a rodent (*Dipus sagitta*), six carnivores (including *Acinonyx jubatus*, *Felis margarita*, *Caracal caracal*), and three herbivores (*Equus hemionus*, *Gazella bennettii* and *Ovis vignei*).

The distance-based method suggested five clusters as the most likely number of bioregions and one large bioregion covered the entire Central Basin and Alborz and Zagros mountains, except in Caucasian region (Table 1; Figure 2).

**TABLE 1** Climate regions, bioregions and transitions zones identified by network-based and distance-based methods, with percentage of Iran occupied by each region (% area) and the percentage covered by protected areas (% protected)

	Name	% area	% protected
Network-based method			
Bioregions			
1	Alborz-Zagros-Kopet Dag Mts.	36.2	9.5
2	Central Basin	27.7	18.8
3	Baluchestan-Khorasan	22.6	4.3
4	Persian Gulf Shores-Khuzestan	6.2	4.0
5	Makran Mts.	2.8	6.5
6	Turkmen Plain	0.6	0.0
7	Makran lowlands	2.8	13.2
Transition zones			
8	Abarkooh-Shahreza ridges	0.6	0.0
9	Arvand-Shadegan lowlands	0.6	50.0
Distance-based method			
Bioregions			
1	N Zagros-Caucasus	11.9	10.5
2	Hyrcanian	2.5	8.5
3	Central Basin-Alborz-Zagros-Kopet Dag Mts.	67.5	13.0
4	Persian Gulf Shores-Khuzestan	8.1	7.2
5	Makran	10.0	5.5
Climate regions			
1	S Khorasan-Sistan	18.7	16.2
2	Makran-Hormuzgan-S Fars	11.9	6.9
3	S Alborz-E Zagros-Kerman Mts	20.6	8.6
4	Persian Gulf Shores-Khuzestan-Lut	11.3	5.7
5	Hyrcanian	2.7	8.9
6	Central Basin-N Baluchestan	17.8	8.2
7	Kordestan	3.4	5.0
8	N Alborz-Caucasus-Kopet Dag Mts.	13.7	15.0

3.2 | Climate regions

Using spatially explicit climate data with the PCA and model-based clustering techniques, we identified eight climate regions (Table 1; Figure 1), and the range of variation of climatic factors tended to be distinct across climate regions (Figure S3 in Appendix S3). Low temperatures were observed in climate regions located in high mountainous areas, such as climate regions 3, 7 and 8, while climate regions 5 and 7 exhibited the highest precipitation levels (Figure S3 in Appendix S3; see Appendix S3 for more details).

3.3 | Spatial overlap between bioregions and climate regions

Network- and distance-based methods identified bioregions with similar spatial patterns (Figure S1 in Appendix S5). For instance, bioregions 1 and 3 in the distance-based method were located in

the bioregion 1 of the network-based method. Bioregions 4 and 5 in each method also tended to match spatially.

Bioregions detected by both methods generally matched climate regions (Figure S2 in Appendix S5). For instance, bioregions 5 and 6 identified by the network-based method completely overlapped with climate regions 2 and 8, respectively. Bioregion 1 mostly overlapped with climate regions 3 and 8, and bioregion 2 broadly overlapped with the range of climate regions 1 and 6. A similar overlap pattern was observed among distance-based bioregions. For instance, the range of bioregions 2 and 4 largely matched with climate regions 5 and 4, respectively (Figure S2 in Appendix S5).

3.4 | Gap analysis

Only bioregions 2 and 7 identified by the network-based method, had more than 10% of their area represented in PAs, while bioregion 6 received no protection at all (Table 1; Figure 3). According to the distance-based method, only bioregions 1 and 3 have slightly more than 10% of their area represented in PAs. With the exception of the climate regions 1 and 8, <10% of each climate region is represented in the current PAs (Table 1).

4 | DISCUSSION

The recently developed network-based method was efficient in delimiting biodiversity patterns in a broad transitional area at a regional scale, as it was found previously at larger spatial scales. We address the advantages and shortcomings of this novel approach, which may provide insights into the utility of INFOMAP BIOREGIONS in biogeographical analyses. The bioregions and climate regions identified here represent the first comprehensive climatic and bioregional classification of Iran and contribute significantly to the understanding of Iran's biogeography.

4.1 | Bioregions

The network-based method used here identified two main bioregions within Iran and their border spatially fits with the limits between the Palaearctic and Saharo-Arabian zoogeographic realms identified by Holt et al. (2013). The two main bioregions correspond to the Alborz-Zagros mountains and Central Basin lowlands, and belong mainly to the temperate broadleaf, mixed, and coniferous forests, and deserts and xeric shrublands biomes, respectively. This finding bears similarities with the zoogeographic map of Kreft and Jetz (2010) derived from global distributions of mammals at family level resulting from UPGMA hierarchical clustering of grid cell assemblages based on β_{sim} dissimilarity matrices. However, the Saharo-Arabian region is grouped with tropical Africa by Kreft and Jetz (2010), while it has been considered to be part of the Palaearctic in most bioregionalizations, and also the recognition of Saharo-Arabian region as a distinct realm is often criticized (Kreft & Jetz, 2013; Rueda et al., 2013).

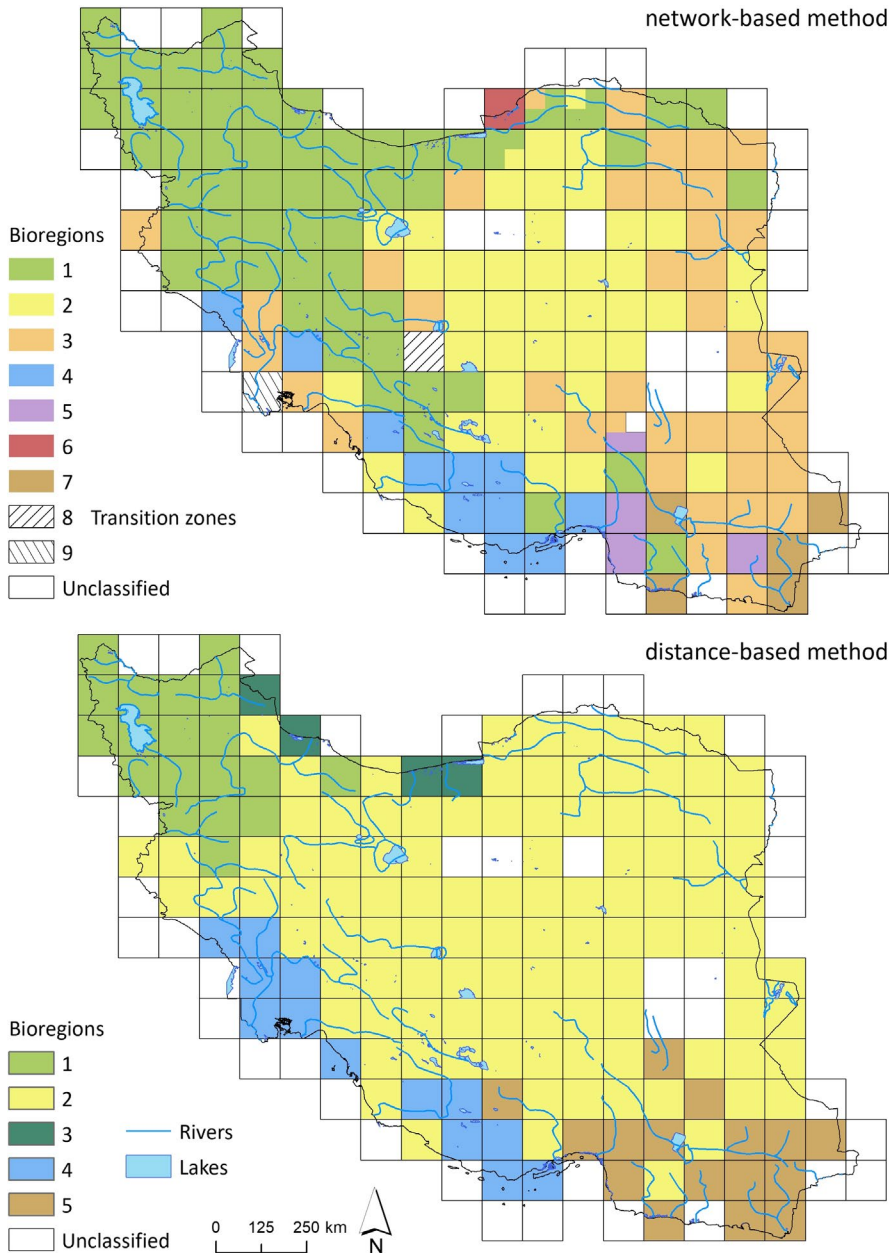


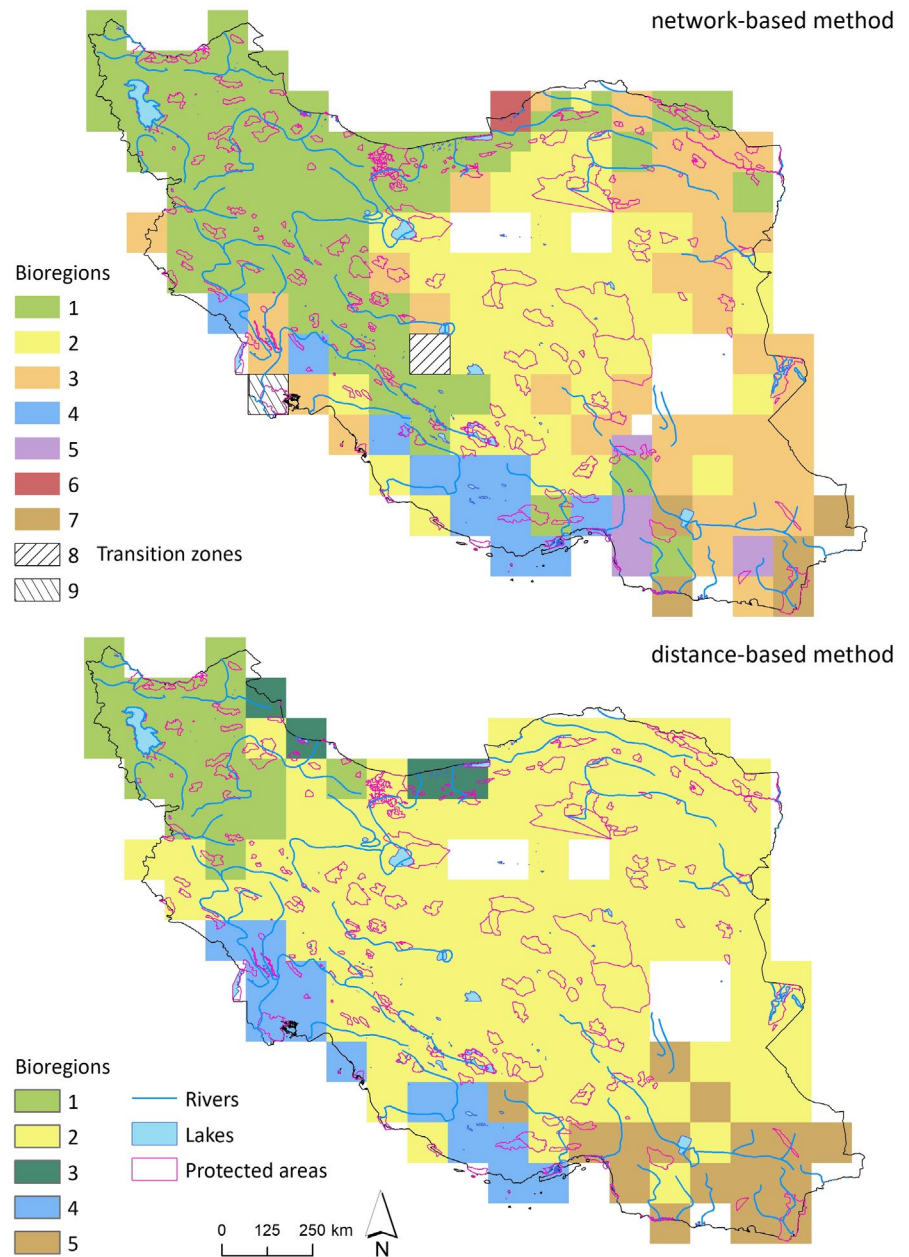
FIGURE 2 Bioregions and transition zones identified by network-based (top: 1—Alborz-Zagros-Kopet Dag Mts.; 2—Central Basin; 3—Baluchestan-Khorasan; 4—Persian Gulf Shores-Khuzestan; 5—Makran Mts.; 6—Turkmen Plain; 7—Makran lowlands; 8—Abarkooh-Shahreza ridges; and 9—Arvand-Shadegan lowlands) and by distance-based (bottom: 1—N Zagros-Caucasus; 2—Central Basin-Alborz-Zagros-Kopet Dag Mts.; 3—Hyrcanian; 4—Persian Gulf Shores-Khuzestan; and 5—Makran) methods for Iran. The maps are based on a 1° grid cells [Colour figure can be viewed at wileyonlinelibrary.com]

The network-based method supported the complexity of south-eastern Iran, where three distinct bioregions intersect. This region is significantly influenced by faunal elements from the Oriental realm and interchange occurs between Saharo-Arabian/Palaearctic and Oriental faunas, as it is the case with reptiles in neighbouring Indus Valley (Ficetola, Falaschi, Bonardi, Padoa-Schioppa, & Sindaco, 2018). Furthermore, the network-based method identified two transition zones (which appear as distinct biogeographical units) in the south-west, indicating that the distribution of Iranian fauna is very complex in this area too. Our results support the idea that Iran is a biogeographical transitional area between the temperate Palaearctic, desertic Saharo-Arabian, and tropical Oriental faunas (Ficetola et al., 2017; Holt et al., 2013). The complexity of southwest Asia in general and Iran in particular has been historically shown by studies based on different animal

and plant taxa (Blanford, 1876; Zohary, 1963, 1973). A similarly complex pattern, for instance, can be seen in the zoogeographical regions derived from global distributions of mammal families (Kreft & Jetz, 2010) or from reptile diversity in the Western Palaearctic (Ficetola et al., 2018).

In general, the results of the distance-based method are in line with the bioregions proposed by the network-based method, with some differences (Figure S1 in Appendix S5). The former method detected one large bioregion (Central Basin-Alborz-Zagros-Kopet Dag Mts.) which comprised two bioregions detected by the network-based method, the central lowland deserts and the high mountains of Alborz and Zagros. Although there are some similarities in the species assemblages between these two areas, there are also marked differences regarding habitat types. Other differences in the bioregions detected by the two methods emerged in south-eastern

FIGURE 3 Spatial overlap between protected areas and the identified bioregions and transition zones by network-based (top: 1—Alborz-Zagros-Kopet Dag Mts.; 2—Central Basin; 3—Baluchestan-Khorasan; 4—Persian Gulf Shores-Khuzestan; 5—Makran Mts.; 6—Turkmen Plain; 7—Makran lowlands; 8—Abarkooh-Shahreza ridges; and 9—Arvand-Shadegan lowlands) and by distance-based (bottom: 1—N Zagros-Caucasus; 2- Central Basin-Alborz-Zagros-Kopet Dag Mts.; 3- Hyrcanian; 4—Persian Gulf Shores-Khuzestan; and 5—Makran) methods. The maps are based on a 1° grid cells [Colour figure can be viewed at wileyonlinelibrary.com]



Iran, where only one bioregion (Makran) was identified by the distance-based method but two bioregions, for lowlands and highlands, were retrieved by the network-based method. In addition, the network-based method detected the lowland desert of Turkmen Plain as a separate bioregion, which although small but harbours high level of diversity and faunal differences. In contrast, the distance-based method was able to detect the Hyrcanian and Caucasian regions as distinct bioregions. The former is covered by dense broad-leaved forests and has been recognized as a separate biogeographical unit within Iran by other studies (Anderson, 1968; Frey & Probst, 1986; Zarudny, 1911; Zohary, 1963, 1973).

The bioregions revealed by the network- and distance-based methods tended to match expert-based biogeographical units. For instance, Blanford (1876) identified five regions, the Central Plateau, the Caspian Forest (Hyrcanian), the Mesopotamian area, Zagros

woodlands and the Persian Gulf and Baluchestan shores (Figure 1), which were also identified as bioregions, especially by the distance-based method (Figure 2). However, the Caucasian region was not identified as part of the Central Basin in the distance-based method. Surprisingly, the distinction of Caucasus was predicted by Blanford (1876) as he stated: "Beside the above, the province of Adarbajian in North-western Persia, ...might perhaps be distinguished". On the contrary, the network-based results were somehow different from Blanford's proposed regions, especially regarding the Alborz and Zagros mountains, that were here identified as distinct bioregion but treated as part of Central Plateau by Blanford (1876). Blanford's zoogeographical regions are largely consistent with the climate regions found here, while the vast Central Plateau was here subdivided into four regions (N Alborz-Caucasus-Kopet Dag Mts., S Alborz-E Zagros-Kerman Mts, S Khorasan-Sistan, and Central Basin-N Baluchestan).



Other studies have found different numbers of biogeographical units: three in plants (Hyrcanian, Irano-Turanian and Nubo-Sindian; Frey & Probst, 1986; Zohary, 1963, 1973), nine in birds (Zarudny, 1911), and 13 in reptiles (Anderson, 1968), which only partly correspond to the bioregions we identified.

4.2 | The network-based method (INFOMAP BIOREGIONS)—pros and cons

The inherent advantages of the network-based method, which reflects the benefits of using community detection algorithms, have been discussed in detail elsewhere (e.g. Bernardo-Madrid et al., 2019; Leroy et al., 2019). Indeed, the community-detection algorithm [Infomap] presents several advantages that were instrumental in our choice: (a) it is able to group taxa within bioregions or assign each species to a specific bioregion (Carstensen et al., 2012); (b) biogeographic structuring is based on species occurrence data directly, instead of similarity measurements as in the distance-based methods (Sidor et al., 2013; Vilhena & Antonelli, 2015); (c) it includes a metric (codelength) that clusters nodes and identifies the best or optimal solutions (partitions on the geographical space) at the same time, by minimizing codelength, and avoids the subjective criterion of deciding the number of bioregions (Bernardo-Madrid et al., 2019); (d) it is considered more appropriate because it can deal with incomplete species distribution data (Edler et al., 2017); and (e) it is less impacted by sampling biases (Vilhena & Antonelli, 2015). Besides, the interactive and user-friendly web application, INFOMAP BIOREGIONS offers several advantages over distance-based methods, including: (a) it was much easier and faster to apply, while the distance-based method required time-consuming data preparation and analytical procedures; (b) it was able to identify transition zones at a regional scale; and (c) it was able to recover the complexity of the country in general, and in south-eastern Iran in particular. The two latter advantages are common to other community detection algorithms.

Despite these advantages, INFOMAP BIOREGIONS also exhibit some drawbacks. First, it lacks a statistical criterion to decide which are the optimal minimum and maximum sizes of grid cells, as well as the maximum number of registers within them. In fact, the determination of the number of clusters is somehow subjective and depends on the chosen parameter values. The number and spatial distribution of the bioregions identified by INFOMAP BIOREGIONS differed considerably under different combinations of chosen parameter values (see Figures S1 and S2 in Appendix S4). Here, we used a simulation approach to test the influence of two parameters, i.e. minimum cell size and maximum cell capacity, on the number and distribution of the retrieved clusters. After closely examining different combinations of parameter values, we decided to use a threshold value (when the estimated clusters appeared to be similar) that would represent a safe margin of error (Figure S1 in Appendix S4). We acknowledge that this threshold was chosen subjectively, but it seems that at the moment there is no feasible alternative to this simulation approach (D. Edler pers. comm.). Nevertheless, obtaining different delineations by using alternative grid cell sizes can also affect

other approaches (Kreft & Jetz, 2010). Moreover, the method also shows some limitations in identifying the most indicative species, as INFOMAP BIOREGIONS only represents species with the highest relative abundance (Edler et al., 2017). This indicator metric is based on the proportion of the records of a species in a given region (or proportion of the distribution of a given species inside a bioregion); however, it could also depend on the proportion of the bioregion occupied by a given species (Bernardo-Madrid et al., 2019), also known as fidelity and affinity, respectively (Dufrêne & Legendre, 1997). Thus, the indicator value of INFOMAP BIOREGIONS may be biased for species with small distributions. It is also important to mention that INFOMAP BIOREGIONS uses species distribution data either as point occurrence or range polygons, but the potential effects of using different types of species distribution data remain unstudied. Given that range polygons may include extensive areas from which a particular species is absent, they are likely to overestimate the true geographic ranges of taxa (Gaston & Fuller, 2009), and hence the retrieved bioregions may differ from point-based analyses due to the differences in the number of grid cells with presence data between these two data-types. This is somehow evident in the bioregion map produced by Edler et al. (2017) for world mammals using species occurrences, which is clearly less structured than the bioregion map of world amphibians generated with Infomap Bioregions, using the IUCN species range maps. According to the literature, high resolution species occurrence data are generally more precise than range polygons and may provide different results (Graham & Hijmans, 2006; Kreft & Jetz, 2010). The shortcomings of species turnover and UPGMA clustering methods were discussed previously (e.g. Bradshaw, Colville, & Linder, 2015; Ferrari, 2018; Vilhena & Antonelli, 2015), regardless of being a successful approach (Kreft & Jetz, 2010).

4.3 | Comparison between bioregions, climate regions and ecoregions

We used the concept of “bioregion” in a similar way to the “ecoregion” concept as used by Dinerstein et al. (2017). According to the literature (see Vilhena & Antonelli, 2015 for a discussion of terminology), both of these biogeographical units, i.e. bioregions and ecoregions, include distinct assemblages of species and both are important units for conservation and management. Overall, the bioregions and climate regions identified here showed rather strong consistency with the 13 main terrestrial ecoregions recognized in Iran (Dinerstein et al., 2017), with a reasonable spatial agreement (Figures S3 and S4 in Appendix S5): (a) the climate regions S Khorasan-Sistan and Central Basin-N Baluchestan broadly overlap with the range of the ecoregion Central Persian desert basin; (b) the climate region Makran-Hormozgan-S Fars mostly overlaps with the range of the ecoregion South Iran Nubo-Sindian desert and semi-desert; (c) the climate Hyrcanian region is almost completely located within the ecoregion Caspian Hyrcanian mixed forests; (d) the climate regions S Alborz-E Zagros-Kerman Mts largely overlap the range of the ecoregion Zagros Mountains forest steppe. The concordance between climate



regions, ecoregions and derived bioregions suggests that patterns of biodiversity distribution may be related to climate variation or may reflect climatic patterns, especially in the case of local-scale studies (González-Orozco et al., 2013).

The identified bioregions were broadly consistent with the terrestrial ecoregions (Figure S3 in Appendix S5). For instance: (a) Alborz-Zagros-Kopet Dag Mts. bioregion mostly located in the ecoregion Zagros Mountains forest steppe; (b) Central Basin and Baluchestan-Khorasan bioregions cover the ecoregion Central Persian desert basins; and (c) Turkmen Plain bioregion encompassed to a large extent the ecoregion Caspian lowland desert ecoregion. In contrast, the bioregions identified here differed markedly from global zoogeographical analyses of mammal distributions, either using clustering methods based on species turnover (Kreft & Jetz, 2010) or network-based methods (Edler et al., 2017), which is probably related to the much finer spatial scale we used here.

4.4 | Protection level of climate regions and bioregions

In general, the climate regions we identified are poorly represented in the current network of PAs of Iran, especially the regions Kordestan and Persian Gulf Shores-Khuzestan-Lut, and the same is true for the bioregions (Table 1). Only the Central Basin bioregion is well-represented, mostly because several PAs have been specifically designated in the region in the last years for the conservation of the Critically Endangered Asiatic cheetah (Khalatbari, Yusefi, Martínez-Freiría, Jowkar, & Brito, 2018). In contrast, the bioregions Turkmen Plain, Baluchestan-Khorasan and Persian Gulf shores-Khuzestan have the lowest protection levels. The Turkmen Plain is of special concern given its small size. In general, many parts of Iran are under-protected, despite their importance for local/regional wildlife conservation (Ghoddousi et al., 2016). The recent regional red list assessment of Iranian land mammals suggested that several species are of particular concern due to their precarious conservation status (Yusefi et al., 2019).

We conclude that the user-friendly web application “INFOMAP BIOREGIONS” is well-suited for the identification of biogeographical units at regional scales (even outperform conventional method), but with some limitations to the models. As the method is relatively new in biogeographical studies, additional research is needed and practitioners should remain vigilant about a final verdict on this approach. Incorporating phylogenetic information (e.g. Holt et al., 2013) may provide further insights into the bioregionalization schemes (Bloomfield et al., 2018; but see Hattab et al., 2015). It is important to note that the results presented here do not represent a comprehensive taxon-wide analysis of Iran's biogeography, but rather a comparison of results generated from different methodological approaches. Although the bioregions and climate regions identified here are purely biogeographical, they may facilitate future attempts to prioritize areas for protection in a diverse country, where a lack of data hampers biodiversity conservation planning. We found that mammal distribution follows a

complex pattern which is reflected in the high number of biogeographical units at small spatial scales, each of which is composed of a unique combination of species and therefore potential conservation units. The understanding of Iran's biogeographical patterns is challenging given the country's physiographic complexity and high beta diversity. We acknowledge that, when available, the use of additional information from other animal and plant taxa could improve the model outputs and allow an even better understanding of the biogeography of the country.

ACKNOWLEDGEMENTS

We thank M. Rosvall and D. Edler for discussions on network-based clustering, F. Santarém, P. Pereira, A.E. Sultán, A. Scharf and P. Tarroso for helping with data analyses, and K. Faizolahi, T. Ghadirian, M. Soofi, A. Qashqaei, H. Abolghasemi, Iranian DoE and provincial offices, Conservation of Asiatic Cheetah Project CACP, Mohitban Society, Persian Wildlife Heritage Foundation PWHF, and several local NGOs for sharing unpublished species records. We also thank R. Bernardo-Madrid and an anonymous reviewer for providing comments that greatly improved our manuscript. Research financially supported by Science and Technology Foundation (FCT; grant PD/BD/52605/2014; BIODIV Doctoral Programme), FCT contract (IF/00459/2013), AGRIGEN-NORTE-01-0145-FEDER-000007, and by Norte Portugal Regional Operational Programme (NORTE2020), under the PORTUGAL 2020 Partnership Agreement, through the European Regional Development Fund (ERDF).

DATA AVAILABILITY STATEMENT

All relevant data are within the paper and its Supporting Information files.

ORCID

Gholam Hosein Yusefi  <https://orcid.org/0000-0002-3258-8561>

REFERENCES

- Anderson, S. C. (1968). Zoogeographic analysis of the lizard fauna. In W. B. Fisher (Ed.), *The land of Iran. The Cambridge history of Iran*, Vol. 1 (pp. 305–371). Cambridge, UK: Cambridge University Press.
- Bernardo-Madrid, R., Calatayud, J., González-Suarez, M., Rosvall, M., Lucas, F., Rueda, M., ... Revilla, E. (2019). Human activity is altering the world's zoogeographical regions. *Ecology Letters*, <https://doi.org/10.1111/ele.13321>
- Blanford, W. T. (1876). *Eastern Persia: An account of the journeys of the Persian boundary commission, 1870–71–72, The Zoology and Geology* Vol. 2. London: Macmillan and Co.
- Bloomfield, N. J., Knerr, N., & Encinas-Viso, F. (2018). A comparison of network and clustering methods to detect biogeographical regions. *Ecography*, 41, 1–10. <https://doi.org/10.1111/ecog.02596>
- Bradshaw, P. L., Colville, J. F., & Linder, H. P. (2015). Optimising regionalisation techniques: Identifying centres of endemism in the extraordinarily endemic-rich Cape Floristic Region. *PLoS ONE*, 10, e0132538. <https://doi.org/10.1371/journal.pone.0132538>



- Brito, J. C., Tarroso, P., Vale, C. G., Martínez-Freiria, F., Boratyński, Z., Campos, J. C., ... Carvalho, S. B. (2016). Conservation biogeography of the Sahara-Sahel: Additional protected areas are needed to secure unique biodiversity. *Diversity and Distributions*, 22, 371–384. <https://doi.org/10.1111/ddi.12416>
- Carstensen, D. W., Dalsgaard, B., Svenning, J.-C., Rahbek, C., Fjelds, J., Sutherland, Q. J., & Olesen, J. M. (2012). Biogeographical modules and island roles: A comparison of Wallacea and the West Indies. *Journal of Biogeography*, 39, 739–749. <https://doi.org/10.1111/j.1365-2699.2011.02628.x>
- Darvishsefat, A. A. (2006). *Atlas of protected areas of Iran* (p. 157). Tehran, Iran: University of Tehran Press.
- Dinerstein, E., Olson, D., Joshi, A., Vynne, C., Burgess, N. D., Wikramanayake, E., ... Saleem, M. (2017). An ecoregion-based approach to protecting half the terrestrial realm. *BioScience*, 67, 534–545. <https://doi.org/10.1093/biosci/bix014>
- DoE-GIS (2016) *Areas under protection by the Department of Environment (DoE) of Iran*. Department of the Environment of Iran, GIS & Remote Sensing Section, GIS. November 2016.
- Droissart, V., Dauby, G., Hardy, O. J., Deblauwe, V., Harris, D. J., Janssens, S., ... Couvreur, T. L. P. (2018). Beyond trees: Biogeographical regionalization of tropical Africa. *Journal of Biogeography*, 45, 1153–1167. <https://doi.org/10.1111/jbi.13190>
- Dufrêne, M., & Legendre, P. (1997). Species assemblages and indicator species: The need for a flexible asymmetrical approach. *Ecological Monographs*, 67, 345–366. <https://doi.org/10.2307/2963459>
- Edler, D., Guedes, T., Zizka, A., Rosvall, M., & Antonelli, A. (2017). Infomap bioregions: Interactive mapping of biogeographical regions from species distributions. *Systematic Biology*, 66, 197–204.
- Escalante, T. (2017). A natural regionalization of the world based on primary biogeographic homology of terrestrial mammals. *Biological Journal of the Linnean Society*, 120, 349–362.
- Ferrari, A. (2018). Biogeographical units matter. *Australian Systematic Botany*, 30, 391–402. <https://doi.org/10.1071/SB16054>
- Ferro, I., & Morrone, J. J. (2014). Biogeographical transition zones: A search for conceptual synthesis. *Biological Journal of the Linnean Society*, 113, 1–12. <https://doi.org/10.1111/bij.12333>
- Ficetola, G. F., Falaschi, M., Bonardi, A., Padoa-Schioppa, E., & Sindaco, R. (2018). Biogeographical structure and endemism pattern in reptiles of the Western Palearctic. *Progress in Physical Geography*, 42, 220–236. <https://doi.org/10.1177/0309133318765084>
- Ficetola, G. F., Mazel, F., & Thuiller, W. (2017). Global determinants of zoogeographical boundaries. *Nature Ecology and Evolution*, 1, 89. <https://doi.org/10.1038/s41559-017-0089>
- Firouz, E. (2005). *The complete fauna of Iran* (p. 322). London, UK: I. B. Tauris & Co Ltd.
- Frey, W., & Probst, W. (1986). A synopsis of the vegetation of Iran. In H. Kurschner (Ed.), *Contributions to the Vegetation of Southwest Asia* (pp. 9–24). Wiesbaden, Germany: L. Reichert.
- Gaston, K. J., & Fuller, R. A. (2009). The sizes of species' geographic ranges. *Journal of Applied Ecology*, 46, 1–9. <https://doi.org/10.1111/j.1365-2664.2008.01596.x>
- Ghoddousi, A., Khaleghi Hamidi, A., Soofi, M., Khorozyan, I., Kiabi, B. H., & Waltert, M. (2016). Effects of ranger stations on predator and prey distribution and abundance in an Iranian steppe landscape. *Animal Conservation*, 19, 273–280. <https://doi.org/10.1111/acv.12240>
- González-Orozco, C. E., Laffan, S., Knerr, N., & Miller, J. T. (2013). A biogeographical regionalization of Australian *Acacia* species. *Journal of Biogeography*, 40, 2156–2166.
- Graham, C. H., & Hijmans, R. J. (2006). A comparison of methods for mapping species ranges and species richness. *Global Ecology and Biogeography*, 15, 578–587. <https://doi.org/10.1111/j.1466-8238.2006.00257.x>
- Hattab, T., Albouy, C., Ben Rais Lasram, F., Le Loc'h, F., Guilhaumon, F., & Leprieur, F. (2015). A biogeographical regionalization of coastal Mediterranean fishes. *Journal of Biogeography*, 42, 1336–1348. <https://doi.org/10.1111/jbi.12505>
- Hazzi, N. A., Moreno, J. S., Ortiz-Movliav, C., & Palacio, R. D. (2018). Biogeographic regions and events of isolation and diversification of the endemic biota of the tropical Andes. *Proceedings of the National Academy of Sciences*, 115, 7985–7990. <https://doi.org/10.1073/pnas.1803908115>
- He, J., Kreft, H., Gao, E., Wang, Z., & Jiang, H. (2017). Patterns and drivers of zoogeographical regions of terrestrial vertebrates in China. *Journal of Biogeography*, 44, 1172–1184. <https://doi.org/10.1111/jbi.12892>
- Holt, B. G., Lessard, J.-P., Borregaard, M. K., Fritz, S. A., Araújo, M. B., Dimitrov, D., ... Rahbek, C. (2013). An update of Wallace's zoogeographic regions of the world. *Science*, 339, 74–78. <https://doi.org/10.1126/science.1228282>
- Kelley, L. A., Gardner, S. P., & Sutcliffe, M. J. (1996). An automated approach for clustering an ensemble of NMR-derived protein structures into conformationally-related subfamilies. *Protein Engineering*, 9, 1063–1065. <https://doi.org/10.1093/protein/9.11.1063>
- Khalatbari, L., Yusefi, G. H., Martínez-Freiria, F., Jowkar, H., & Brito, J. C. (2018). Range contraction of the Asiatic cheetah during last century is related to prey availability and climate change. *Hystrix, Italian Journal of Mammalogy*, 29, 145–151.
- Koleff, P., Gaston, K. J., & Lennon, J. J. (2003). Measuring beta diversity for presence-absence data. *Journal of Animal Ecology*, 72, 367–382. <https://doi.org/10.1046/j.1365-2656.2003.00710.x>
- Kreft, H., & Jetz, W. (2010). A framework for delineating biogeographical regions based on species distributions. *Journal of Biogeography*, 37, 2029–2053. <https://doi.org/10.1111/j.1365-2699.2010.02375.x>
- Kreft, H., & Jetz, W. (2013). Comment on 'An update of Wallace's zoogeographic regions of the world'. *Science*, 341, 343. <https://doi.org/10.1126/science.1237471>
- Leroy, B., Dias, M. S., Giraud, E., Huguency, B., Jézéquel, C., Leprieur, F., ... Tedesco, P. A. (2019). Global biogeographical regions of freshwater fishes. Retrieved from <https://www.biorxiv.org/content/biorxiv/early/2019/03/25/319566.full.pdf>
- Michalak, J. L., Lawler, J. J., Roberts, D. R., & Carroll, C. (2018). Distribution and protection of climatic refugia in North America. *Conservation Biology*, 32, 1414–1425. <https://doi.org/10.1111/cobi.13130>
- Mittermeier, R. A., Turner, W. R., Larsen, F. W., Brooks, T. M., & Gascon, C. (2011). Global biodiversity conservation: The critical role of hotspots. In F. E. Zachos & J. C. Habel (Eds.), *Biodiversity hotspots: Distribution and protection of priority conservation areas* (pp. 3–22). Berlin: Springer-Verlag.
- Morrone, J. J. (2009). *Evolutionary Biogeography: An integrative approach with case studies* (p. 301). New York: Columbia University Press.
- Morrone, J. J. (2018). The spectre of biogeographical regionalization. *Journal of Biogeography*, 45, 282–288. <https://doi.org/10.1111/jbi.13135>
- R Development Core Team (2018) R: a language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from <http://www.R-project.org/>
- Richardson, D. M., & Whittaker, R. J. (2010). Conservation biogeography – foundations, concepts and challenges. *Diversity and Distributions*, 16, 313–320. <https://doi.org/10.1111/j.1472-4642.2010.00660.x>
- Riddle, B. R., Ladle, R. J., Lourie, S. A., & Whittaker, R. J. (2011). Basic biogeography: Estimating biodiversity and mapping nature. In R. J. Ladle & R. J. Whittaker (Eds.), *Conservation biogeography* (pp. 47–92). Oxford, UK: Blackwell Publishing.
- Rosvall, M., Axelsson, D., & Bergstrom, C. T. (2009). The map equation. *The European Physical Journal Special Topics*, 178, 13–23. <https://doi.org/10.1140/epjst/e2010-01179-1>
- Rosvall, M., & Bergstrom, C. T. (2008). Maps of information flow reveal community structure in complex networks. *Proceedings of the*



- National Academy of Sciences of the United States of America, 105, 1118–1123.
- Rueda, M., Rodríguez, M. A., & Hawkins, B. A. (2013). Identifying global zoogeographical regions: Lessons from Wallace. *Journal of Biogeography*, 40, 2215–2225. <https://doi.org/10.1111/jbi.12214>
- Sidor, C. A., Vilhena, D. A., Angielczyk, K. D., Huttenlocker, A. K., Nesbitt, S. J., Peacock, B. R., ... Tsuji, L. A. (2013). Provincialization of terrestrial faunas following the end-Permian mass extinction. *Proceedings of the National Academy of Sciences of the United States of America*, 110, 8129–8133. <https://doi.org/10.1073/pnas.1302323110>
- Vilhena, D. A., & Antonelli, A. (2015). A network approach for identifying and delimiting biogeographical regions. *Nature Communications*, 6, 6848. <https://doi.org/10.1038/ncomms7848>
- Whittaker, R. J., Araújo, M. B., Paul, J., Ladle, R. J., Watson, J. E. M., & Willis, K. J. (2005). Conservation biogeography: Assessment and prospect. *Diversity and Distributions*, 11, 3–23. <https://doi.org/10.1111/j.1366-9516.2005.00143.x>
- Yusefi, G. H., Faizolah, K., Darvish, J., Safi, K., & Brito, J. C. (2019). The species diversity, distribution and conservation status of the terrestrial mammals of Iran. *Journal of Mammalogy*, 100, 55–71. <https://doi.org/10.1093/jmammal/gyz002>
- Zarudny, N. A. (1911). Verzeichnis der Vögel Persiens. *Journal Für Ornithologie*, 59, 185–241. <https://doi.org/10.1007/BF02091053>
- Zehad, B., Kiabi, B. H., & Madjnoonian, H. (2002). The natural areas and landscape of Iran: An overview. *Zoology in the Middle East*, 26, 7–10. <https://doi.org/10.1080/09397140.2002.10637915>
- Zohary, M. (1963). On the geobotanical structure of Iran. *The Bulletin of the Research Council of Israel*, 11D supplementary, 1–113.
- Zohary, M. (1973). *Geobotanical foundations of the Middle East* (Volumes 1 & 2) (p. 756). Stuttgart, Germany: Gustav Fischer Verlag.

BIOSKETCH

This study is part of G.H.Y.'s PhD project (University of Porto, CIBIO-InBIO) focused on diversity and distribution of Iranian mammals. His main research interests include biogeography, mammalian taxonomy, ecology and conservation. J.C.B. leads BIODESERTS research group at CIBIO-InBIO, focused in assessing biodiversity patterns and evolutionary processes in deserts and arid regions.

Authors' contributions: G.H.Y. and J.C.B. conceived the ideas; G.H.Y. collected the data; G.H.Y., J.C.B. and K.S. analysed the data, G.H.Y. wrote the paper and J.C.B. and K.S. contributed to the writing.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

How to cite this article: Yusefi GH, Safi K, Brito JC. Network- and distance-based methods in bioregionalization processes at regional scale: An application to the terrestrial mammals of Iran. *J Biogeogr.* 2019;46:2433–2443. <https://doi.org/10.1111/jbi.13694>