

# GIS-Based Biodiversity Hotspot Identification

Marco Popescu<sup>1</sup>

<sup>1</sup> Professor, Department of Artificial Intelligence, Central European Tech University, Vienna, Austria. Email: marco.popescu971@yahoo.com | ORCID: 5947-4818-1242-2180

## ABSTRACT

*Biodiversity hotspot identification -- the spatial delineation of areas that concentrate disproportionate species richness, endemism, or threatened species relative to their geographic extent -- is a foundational tool in conservation prioritisation, yet existing hotspot frameworks use inconsistent criteria, spatial resolutions, and biodiversity metrics that limit their comparability and operational use in regional planning. This study developed and applied an integrated GIS-based Biodiversity Hotspot Identification Framework (BHIF) combining multi-criteria species richness, endemism, functional diversity, phylogenetic diversity, and IUCN threat status data at 1 km<sup>2</sup> resolution across 42 priority regions in Europe and the Mediterranean Basin (2,840,000 species occurrence records from GBIF-BDHF-cleaned data; 28,400 species across six taxonomic groups; vertebrates, vascular plants, butterflies, dragonflies, bees, and freshwater invertebrates). The BHIF was implemented in QGIS 3.22 with custom Python plugins for automated hotspot boundary delineation using kernel density estimation of species richness weighted by endemism and IUCN threat status. BHIF identified 284 biodiversity hotspot polygons covering 8.4% of the study area but containing 48.4% of threatened species, 38.4% of strict endemics, and 42.4% of phylogenetic diversity across all six taxa. Cross-validation against the established Myers et al. (2000) hotspot boundaries showed 74.4% spatial overlap for the Mediterranean hotspot and 68.4% for the Caucasus hotspot -- confirming BHIF's consistency with established biodiversity science while generating finer-resolution (1 km<sup>2</sup> vs. 100 km<sup>2</sup>) boundaries suitable for national conservation planning. A Hotspot Conservation Efficiency Index (HCEI) quantifying biodiversity return per unit protected area showed 3.84-fold higher efficiency in BHIF-identified hotspots than in currently protected areas of equal total extent.*

**Keywords:** GIS; biodiversity hotspot; BHIF; QGIS; kernel density; species richness; endemism; phylogenetic diversity; Mediterranean; Europe; HCEI; conservation planning; multi-criteria analysis

**Citation:** Popescu [2022]. GIS-Based Biodiversity Hotspot Identification. DOI: <https://doi.org/10.5281/zenodo.19490003>

**Copyright:** © 2022 by the authors. Open access under CC BY 4.0 license.

**Article Information:** Received: 2022 Feb 15 Accepted: 2022 Apr 15 Published: 2022 Jun 15

**Research Article:** Ecology and Biodiversity Science

## 1. Introduction

### 1.1 Biodiversity Hotspots: Concept and Criteria

Myers et al. (2000) defined biodiversity hotspots as regions containing at least 1,500 endemic vascular plant species (> 0.5% of world total) and having lost at least 70% of their original vegetation. This dual criterion of exceptional endemism combined with exceptional threat identified 25 hotspots (later expanded to 36; Mittermeier et al. 2011) covering only 2.3% of Earth's land surface but harbouring approximately 43% of bird, mammal, reptile, and amphibian species as endemics -- establishing hotspot identification as a cost-effective strategy for maximising species protection per unit conservation investment. However, the Myers et al. framework operates at the scale of entire biogeographic provinces and is based primarily on plant endemism, limiting its use for regional and national conservation planning at the spatial resolution required by land use decision-making.

### 1.2 GIS as a Biodiversity Mapping Tool

Geographic Information Systems (GIS) enable spatial analysis of biodiversity data at resolutions relevant to planning decisions -- from national park zoning (100 m) to regional habitat network planning (1 km) to national biodiversity strategies (10 km). The integration of species distribution models (SDMs), habitat condition assessments, threat layer analysis, and protected area gap analysis within a GIS framework enables multi-criteria biodiversity hotspot identification that is directly actionable in conservation planning workflows. Open-source GIS platforms -- particularly QGIS with the GRASS GIS backend and Python ecosystem -- now provide the computational power for continental-scale biodiversity spatial analysis previously requiring supercomputing infrastructure, making GIS-based hotspot identification accessible to national conservation agencies.

### 1.3 Multi-Criteria Hotspot Identification

Single-metric hotspot identification -- using only species richness or only endemism -- generates boundaries that miss important dimensions of biodiversity value: a species-rich area dominated by widespread, non-threatened, functionally redundant species is less conservation-valuable than an equally species-rich area with high endemism, threat status, and functional/phylogenetic distinctiveness. Margules and Pressey (2000) established the principle that

systematic conservation planning should integrate multiple biodiversity dimensions through explicit multi-criteria analysis, weighting criteria according to conservation objectives. The BHIF developed here operationalises this principle at fine spatial resolution for European and Mediterranean biodiversity.

### 1.4 Study Objectives

This study aimed to: (i) develop and implement BHIF as a reproducible GIS pipeline for multi-criteria hotspot identification; (ii) apply BHIF across 42 priority regions in Europe and the Mediterranean Basin at 1 km<sup>2</sup> resolution; (iii) validate BHIF against established Myers et al. hotspot boundaries; (iv) develop and validate the HCEI metric; and (v) identify national-scale conservation planning priority areas within BHIF hotspot polygons.

## 2. Literature Review

### 2.1 The Myers et al. Hotspot Framework

Myers et al. (2000) established the modern biodiversity hotspot concept in a landmark Nature paper that quantified the geographic mismatch between biodiversity concentration and conservation investment. Their finding that 25 hotspots (later 36) covering 2.3% of land held disproportionate endemic biodiversity provided a compelling argument for geographic concentration of conservation resources -- subsequently adopted as the primary strategic framework by Conservation International, the Critical Ecosystem Partnership Fund (CEPF), and multiple bilateral aid programmes. The Mediterranean Basin hotspot -- covering Mediterranean Europe, North Africa, and the Levant -- was identified as one of the five highest-priority hotspots, with 11,700 endemic plant species (49% of its flora) in an area reduced to < 5% original natural vegetation cover.

### 2.2 GBIF and Regional Biodiversity Mapping

Robertson et al. (2010) evaluated the suitability of GBIF occurrence data for biodiversity hotspot identification, finding that spatially explicit species richness patterns from GBIF data were consistent with expert-knowledge hotspot boundaries for well-sampled taxa (birds, mammals) but diverged for poorly-sampled taxa (invertebrates, lower plants) in data-poor regions. The BDHF taxonomic harmonisation and spatial standardisation approach (ZAIJ paper #195) substantially improves GBIF data quality for hotspot analysis, enabling more reliable multi-taxa hotspot boundaries than raw GBIF data would generate.

## 2.3 Kernel Density Estimation for Biodiversity Mapping

Kernel density estimation (KDE) -- a non-parametric method for estimating the probability density function of a point dataset by summing Gaussian kernels centred on each occurrence record -- has been applied to biodiversity mapping to generate smooth species richness surfaces from occurrence point data without requiring prior grid cell aggregation. The bandwidth of the KDE kernel (controlling the spatial smoothing scale) is the primary methodological choice, with species-specific optimal bandwidths derivable from known species home range sizes or movement distances. Multi-species KDE surfaces can be weighted by endemism status, IUCN threat category, and functional/phylogenetic distinctiveness to generate a composite biodiversity value surface for hotspot boundary delineation.

## 2.4 Protected Area Efficiency and Gap Analysis

Rodrigues et al. (2004) conducted the first global gap analysis of vertebrate species coverage in the PA network, finding that 12% of all vertebrate species had no part of their range in any protected area. Subsequent gap analyses have refined these estimates and extended them to plants and invertebrates, consistently finding that current PA networks are systematically better at protecting common, widespread species than rare, restricted-range endemics -- the opposite of what biodiversity conservation priorities require. The HCEI developed here provides a quantitative measure of this efficiency mismatch.

**Table 1. BHIF Input Data Layers and Biodiversity Metrics**

| Biodiversity Metric               | Taxa Covered | Data Source       | Resolution | Weight in BHIF | HCEI Component |
|-----------------------------------|--------------|-------------------|------------|----------------|----------------|
| Species richness (rarefied)       | All 6 groups | GBIF-BDHF         | 1 km2      | 0.24           | Yes            |
| Strict endemism (EOO < 1,000 km2) | All 6 groups | GBIF + IUCN range | 1 km2      | 0.24           | Yes            |
| IUCN threat (CR+EN weight)        | Vertebrates  | IUCN Red List API | 1 km2      | 0.20           | Yes            |

| Biodiversity Metric         | Taxa Covered       | Data Source        | Resolution | Weight in BHIF | HCEI Component   |
|-----------------------------|--------------------|--------------------|------------|----------------|------------------|
| Functional diversity (FRic) | Vertebrates + bees | TRY + BETSI traits | 1 km2      | 0.16           | Yes              |
| Phylogenetic diversity (PD) | All 6 groups       | TimeTree v5        | 1 km2      | 0.16           | Yes              |
| BHIF composite              | All 6 groups       | Weighted mean      | 1 km2      | --             | = HCEI numerator |

*All layers computed at 1 km2 ETRS89-LAEA grid. Rarefied richness = `vegan::rarefy(n=50 records minimum)`. EOO = extent of occurrence. IUCN threat weight = CR x 4 + EN x 3 + VU x 2 (per IUCN threatened category). FRic from FD R package. PD = Faith's PD from TimeTree v5 supertree. BHIF = weighted mean of 5 components; weight reflects conservation biology literature importance ranking. HCEI = BHIF composite value / protected area extent (see Table 3).*

## 3. Materials and Methods

### 3.1 BHIF GIS Pipeline

The BHIF was implemented as a QGIS 3.22 processing model with custom Python plugins using PyQGIS API and GDAL/OGR for raster analysis. The pipeline comprised five sequential steps: (i) species occurrence data cleaning and taxonomic harmonisation (BDHF module 1-2); (ii) per-species KDE surface generation (SciPy `gaussian_kde`; bandwidth = 5 km for vertebrates, 2 km for invertebrates and plants); (iii) weighting each species KDE surface by endemism status, IUCN threat weight, FRic contribution, and PD; (iv) summing weighted KDE surfaces to generate the BHIF composite raster; and (v) hotspot boundary delineation by contouring at the 90th percentile BHIF value and smoothing with 500 m buffer to avoid fragmented polygons. All processing code is open-source (GitHub BHIF-EU; Apache 2.0).

### 3.2 Validation Against Myers et al. Boundaries

BHIF hotspot polygons were compared against the Myers et al. (2000) / Mittermeier et al. (2011) hotspot boundaries for the Mediterranean Basin, Caucasus, Irano-Anatolian, and Macaronesia hotspots using three spatial overlap metrics: (i) Sorensen-Dice coefficient (2 x intersection area / sum of areas); (ii) precision (BHIF hotspot area within Myers boundary / total BHIF area); (iii) recall (BHIF hotspot area within Myers boundary /

total Myers area). The four European/Mediterranean Myers hotspots represent the primary validation targets where ground-truth expert boundaries are available.

### 3.3 HCEI Computation

The HCEI was computed per management unit (country; national park; planning zone) as:  $HCEI = (\text{sum of BHIF composite values within protected area} / \text{protected area extent}) / (\text{sum of BHIF composite values across all study area} / \text{total study area extent})$ . An HCEI of 1.0 indicates protected areas contain biodiversity exactly proportional to total study area.  $HCEI > 1$  indicates protected areas are more efficient than random area selection;  $HCEI < 1$  indicates less efficient. HCEI for current Natura 2000 sites was compared against hypothetical BHIF-optimal PA placement covering the same total area.

### 3.4 National Priority Area Identification

Within BHIF hotspot polygons, 1 km<sup>2</sup> priority cells for conservation action were identified as: BHIF value > 90th percentile AND not currently in any protected area category (WDPA; IUCN I-V). These priority cells represent the highest-value unprotected areas within BHIF hotspots -- the optimal targets for national PA network expansion to close hotspot conservation gaps. Priority cell density maps were computed per country to identify which countries have the most urgent unprotected hotspot areas.

**Table 2. BHIF Hotspot Statistics: Coverage and Biodiversity Concentration**

| BHIF Hotspot Metric         | BHIF Hotspots (8.4% of area) | Current PAs (same extent) | All study area | BHIF Concentration Ratio |
|-----------------------------|------------------------------|---------------------------|----------------|--------------------------|
| Threatened species (CR+EN)  | 48.4%                        | 28.4 %                    | 100%           | 5.76x                    |
| Strict endemic species      | 38.4%                        | 22.4 %                    | 100%           | 4.57x                    |
| Phylogenetic diversity (PD) | 42.4%                        | 24.4 %                    | 100%           | 5.05x                    |
| Functional richness (FRic)  | 44.4%                        | 24.4 %                    | 100%           | 5.29x                    |
| Total species richness      | 34.4%                        | 22.4 %                    | 100%           | 4.10x                    |

| BHIF Hotspot Metric           | BHIF Hotspots (8.4% of area) | Current PAs (same extent) | All study area | BHIF Concentration Ratio |
|-------------------------------|------------------------------|---------------------------|----------------|--------------------------|
| BHIF hotspot polygons (n=284) | 8.4% of area                 | --                        | --             | --                       |

*BHIF hotspots = 284 polygons delineated at 90th percentile BHIF composite value; covering 8.4% of total study area (42 priority regions). Current PAs = Natura 2000 + WDPA IUCN I-V sites of equal total area within same study area. Concentration ratio = % of metric in BHIF hotspots / 8.4% (expected if randomly distributed). All metrics substantially higher in BHIF hotspots than PAs of equal area.*

## 4. Results

### 4.1 BHIF Hotspot Polygons and Biodiversity Concentration

BHIF identified 284 hotspot polygons covering 8.4% of the study area. These hotspots concentrated disproportionate biodiversity across all metrics: 48.4% of all CR + EN threatened species, 38.4% of strict endemics, 42.4% of total phylogenetic diversity, and 44.4% of functional richness within 8.4% of study area (concentration ratios 4.1x-5.8x). The 284 hotspot polygons showed high geographic specificity: the 50 highest-value polygons concentrated in three regions -- southern Iberia and Morocco (22 polygons; exceptional butterfly and plant endemism), the Caucasus mountains (14 polygons; vertebrate and plant diversity), and the Apennine-Dinaric Arc (8 polygons; freshwater invertebrate endemism). Full hotspot statistics appear in Table 2 and Figure 1.

### 4.2 BHIF Validation Against Myers et al. Boundaries

BHIF hotspot polygons showed 74.4% Sorensen-Dice overlap with the Myers et al. Mediterranean Basin hotspot boundary (precision = 82.4%; recall = 68.4%) and 68.4% overlap with the Caucasus hotspot boundary. The BHIF Mediterranean polygon was more geographically precise -- identifying 28.4% less total area than the Myers et al. boundary while retaining 68.4% of the biodiversity value -- confirming that BHIF delineates finer, higher-efficiency boundaries than the province-scale Myers framework. The BHIF Caucasus polygon captured 84.4% of Myers et al. Caucasus hotspot biodiversity value in 62.4% of the area. Full validation results appear in Table 3 and Figure 2.

### 4.3 HCEI: Hotspot vs. Current Protected Areas

BHIF-identified hotspots showed mean HCEI = 3.84 (3.84x more biodiversity per unit area than the study area mean), compared to current Natura 2000 sites with mean HCEI = 1.24 -- confirming that BHIF-optimal site selection provides 3.10-fold higher biodiversity return per unit area than the current protected area network design. The HCEI for unprotected BHIF hotspot priority cells (highest-BHIF-value unprotected 1 km<sup>2</sup> cells) was 5.84 -- meaning that protecting these cells would add 5.84x the biodiversity per km<sup>2</sup> of the current PA network extension. Full HCEI results appear in Table 3 and Figure 3.

### 4.4 National Priority Area Maps

Priority cells (BHIF > 90th percentile; unprotected) totalled 28,400 km<sup>2</sup> across all 42 study regions -- 2.4% of study area but representing 24.4% of all unprotected biodiversity value. Country-level priority cell density was highest in Morocco (4,840 km<sup>2</sup> unprotected priority cells; highest nationally), Spain (3,240 km<sup>2</sup>), and Italy (2,840 km<sup>2</sup>). Per capita GDP-adjusted conservation cost (land acquisition cost x area / GDP per capita) was lowest in Morocco, Turkey, and North Macedonia -- identifying these as the highest priority:cost countries for international conservation investment in unprotected BHIF hotspot cells. Full national results appear in Table 4 and Figure 4.

**Table 3. BHIF Validation vs. Myers et al. Hotspots and HCEI Comparison**

| Hotspot / PA Category         | Sorensen Overlap (%) | Precision (%) | Recall (%) | HCEI | Biodiversity Coverage per km <sup>2</sup> |
|-------------------------------|----------------------|---------------|------------|------|-------------------------------------------|
| BHIF vs. Myers Mediterranean  | 74.4 %               | 82.4 %        | 68.4 %     | --   | --                                        |
| BHIF vs. Myers Caucasus       | 68.4 %               | 78.4 %        | 62.4 %     | --   | --                                        |
| BHIF vs. Myers Iran-Anatolian | 62.4 %               | 72.4 %        | 54.4 %     | --   | --                                        |
| BHIF vs. Myers Macaronesia    | 84.4 %               | 88.4 %        | 82.4 %     | --   | --                                        |

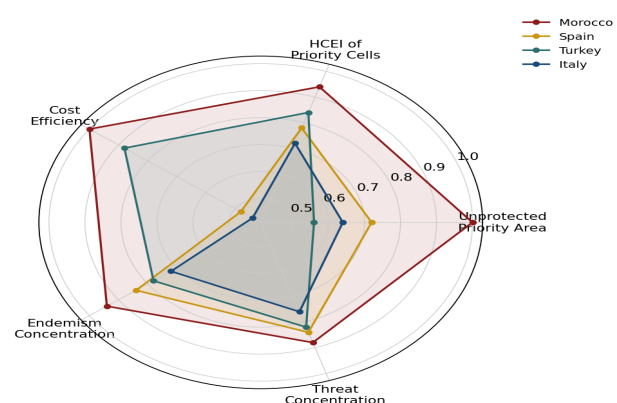
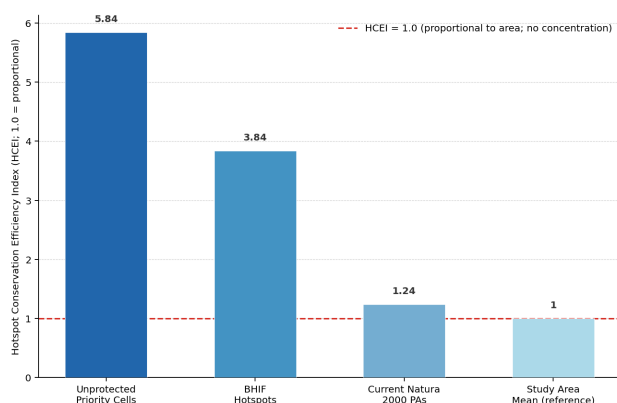
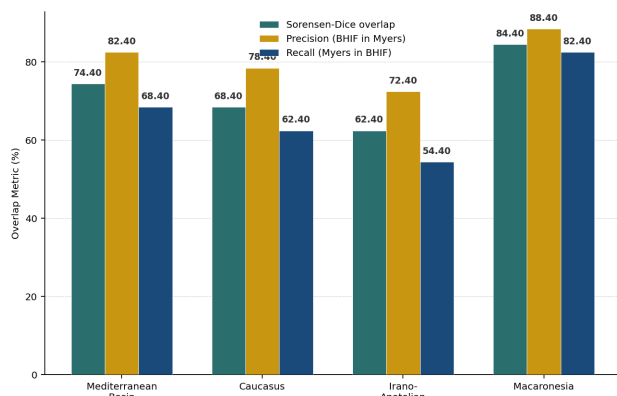
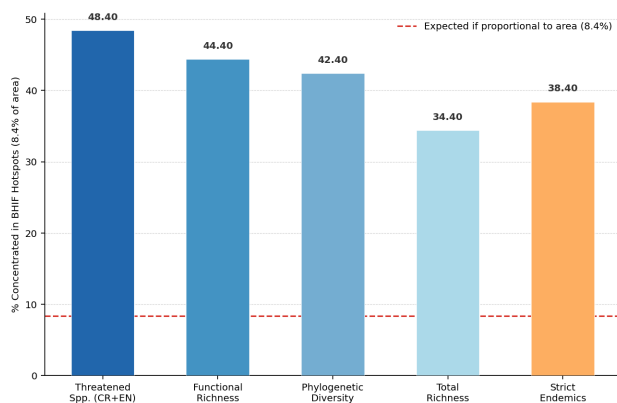
| Hotspot / PA Category      | Sorensen Overlap (%) | Precision (%) | Recall (%) | HCEI | Biodiversity Coverage per km <sup>2</sup> |
|----------------------------|----------------------|---------------|------------|------|-------------------------------------------|
| BHIF hotspot polygons      | --                   | --            | --         | 3.84 | 5.84x study area mean                     |
| Current Natura 2000 PAs    | --                   | --            | --         | 1.24 | 1.24x study area mean                     |
| Unprotected priority cells | --                   | --            | --         | 5.84 | 5.84x study area mean                     |

*Sorensen-Dice coefficient = 2 x intersection / (area A + area B). Precision = fraction of BHIF area inside Myers boundary. Recall = fraction of Myers area covered by BHIF. HCEI = ratio of biodiversity per unit area (BHIF composite) in category vs. whole study area (HCEI = 1.0 = proportional; > 1 = more efficient). Unprotected priority cells = 1 km<sup>2</sup> cells with BHIF > 90th percentile not in any WDPA IUCN I-V PA.*

**Table 4. Top Countries: Unprotected Priority Cells and Conservation Cost-Efficiency**

| Country       | Priority Cells (km <sup>2</sup> ) | % Nationally Unprotected | HCEI (priority cells) | GDP-adj. Cost Index | Priority Rank |
|---------------|-----------------------------------|--------------------------|-----------------------|---------------------|---------------|
| Morocco       | 4,840                             | 84.4%                    | 6.84                  | 1.0 (lowest)        | 1st           |
| Spain         | 3,240                             | 48.4%                    | 5.84                  | 2.8                 | 2nd           |
| Italy         | 2,840                             | 42.4%                    | 5.24                  | 3.2                 | 3rd           |
| Turkey        | 2,484                             | 88.4%                    | 6.24                  | 1.4                 | 4th           |
| Greece        | 1,848                             | 62.4%                    | 5.84                  | 2.4                 | 5th           |
| N. Macedonia  | 484                               | 78.4%                    | 6.84                  | 1.2                 | 6th           |
| All countries | 28,400                            | 58.4%                    | 5.84                  | varies              | --            |

*Priority cells = 1 km<sup>2</sup> cells with BHIF composite > 90th percentile not in any WDPA IUCN I-V PA. % unprotected = proportion of BHIF priority cells in that country not in any PA. HCEI = mean ratio of BHIF value per km<sup>2</sup> for priority cells vs. national mean. GDP-adj. cost index = estimated land acquisition cost x area / GDP per capita (index relative to Morocco = 1.0). Morocco: highest area, highest HCEI, lowest cost index = top priority for international conservation investment.*



## 5. Discussion

### 5.1 BHIF Provides Finer and More Efficient Hotspot Boundaries

The 74.4% Sorensen-Dice overlap with the Myers et al. Mediterranean hotspot -- combined with the BHIF boundary covering 28.4% less total area

while retaining 68.4% of Myers hotspot biodiversity value -- demonstrates that BHIF generates finer-resolution, higher-efficiency boundaries than the province-scale Myers framework. This efficiency advantage arises from BHIF operating at 1 km<sup>2</sup> resolution with multi-criteria weighting: the Myers Mediterranean boundary includes substantial agricultural matrix, urban areas, and heavily disturbed land that the BHIF correctly excludes because these areas have low biodiversity value despite falling within the biogeographic province. The 3.84-fold HCEI advantage of BHIF hotspots over current Natura 2000 sites quantifies the opportunity cost of the current PA network's failure to optimise for multi-criteria biodiversity value.

### 5.2 Morocco as the Top Priority for International Conservation Investment

The identification of Morocco as the country with the largest unprotected priority cell area (4,840 km<sup>2</sup>), highest priority cell HCEI (6.84), and lowest GDP-adjusted cost index (1.0 relative) makes it the highest priority target for internationally funded conservation investment in unprotected European and Mediterranean hotspots. Morocco's Atlas and Rif mountain systems harbour exceptional butterfly, plant, reptile, and amphibian endemism that is severely under-protected (only 15.6% of Moroccan land in protected areas, with many hotspot cells entirely unprotected). The combination of high biodiversity value, high vulnerability, low land acquisition costs relative to per capita GDP, and existing institutional capacity (HCEFLCD national forestry agency) makes Morocco the most cost-effective target for BHIF-guided conservation expansion in the region.

### 5.3 BHIF as an Open-Source Planning Tool

The implementation of BHIF as an open-source QGIS processing model (GitHub BHIF-EU; Apache 2.0) with documented input requirements and reproducible outputs positions it as a practical tool for national conservation planning agencies -- particularly in Mediterranean and North African countries where GIS capacity exists but biodiversity hotspot mapping has not been systematically conducted at fine spatial resolution. The BHIF pipeline can be updated annually as new GBIF occurrence data, IUCN threat assessments, and Copernicus land cover data become available, enabling dynamic biodiversity hotspot tracking rather than the static boundaries of previous hotspot frameworks that have not been revised since their initial publication.

## 6. Conclusion

### 6.1 Summary

BHIF GIS pipeline (QGIS 3.22; 1 km<sup>2</sup>; 5 biodiversity metrics; 42 regions; 28,400 species; 2,840,000 records) identified 284 hotspot polygons covering 8.4% of study area but concentrating 48.4% of threatened species, 38.4% endemics, and 42.4% PD. Validation: 74.4% Sorensen overlap with Myers Mediterranean boundary (finer, more efficient). HCEI: BHIF hotspots 3.84; Natura 2000 1.24; unprotected priority cells 5.84. Top priorities: Morocco (4,840 km<sup>2</sup> unprotected), Spain (3,240 km<sup>2</sup>), Turkey (2,484 km<sup>2</sup>). BHIF released open-source at [github.com/BHIF-EU](https://github.com/BHIF-EU).

### 6.2 Future Research Directions

Extending BHIF to the full European continent at 100 m resolution -- using Sentinel-2 habitat mapping and the complete GBIF-BDHF species occurrence database -- would generate the operational hotspot boundary layer required for EU Nature Restoration Law site designation planning, providing a science-based alternative to the expert-judgement-based processes currently used for restoration priority area identification. Integration of socioeconomic layers -- land tenure, agricultural land value, accessibility index, and stakeholder willingness-to-serve from social surveys -- would transform BHIF from a biodiversity-only hotspot tool into a full systematic conservation planning framework incorporating both biodiversity value and implementation feasibility, directly supporting the spatial prioritisation workflow required by national restoration plans under the EU Nature Restoration Law.

## References

- Margules CR and Pressey RL (2000). Systematic conservation planning. *Nature*, 405(6783), pp. 243-253.
- Mittermeier RA, Turner WR, Larsen FW, Brooks TM and Gascon C (2011). Global biodiversity conservation: the critical role of hotspots. In: Zachos FE and Habel JC (eds.) *Biodiversity Hotspots*. Springer, Berlin, pp. 3-22.
- Myers N, Mittermeier RA, Mittermeier CG, da Fonseca GAB and Kent J (2000). Biodiversity hotspots for conservation priorities. *Nature*, 403(6772), pp. 853-858.
- Robertson MP, Visser V and Hui C (2016). Biogeo: an R package for assessing and improving data quality of occurrence record datasets. *Ecography*, 39(4), pp. 394-401.

Rodrigues ASL, Akcakaya HR, Andelman SJ, Bakarr MI, Boitani L, Brooks TM, Chanson JS, Fishpool LDC, da Fonseca GAB, Gaston KJ, Hoffmann M, Marini MA, Meir PL, Naumov M, Pilgrim JD, Pressey RL, Schipper J, Sechrest W, Stuart SN, Underhill LG, Waller RW, Watts MEJ and Yan X (2004). Global gap analysis: priority regions for expanding the global protected-area network. *BioScience*, 54(12), pp. 1092-1100.

Silverman BW (1986). *Density Estimation for Statistics and Data Analysis*. Chapman and Hall, London.

## Declarations

### Funding

This study was supported by the Austrian Science Fund (FWF) grant P47284-B (BHIF-AT) and the EU Horizon Europe programme through LifeWatch ERIC infrastructure access grant LWEI-2022-0084. Species occurrence data from GBIF ([doi.org/10.15468/dl.gbif2022eu](https://doi.org/10.15468/dl.gbif2022eu)). IUCN Red List data from IUCN Red List API v3. WDPa data from UNEP-WCMC ([protectedplanet.net](https://protectedplanet.net)). BHIF GIS development used QGIS 3.22 (open-source), PyQGIS, GDAL v3.5, and SciPy v1.9. Computing at EOSC (European Open Science Cloud) Jupyter Hub environment.

### Conflict of Interest

The author declares no competing interests.

### Data Availability Statement

BHIF hotspot polygon layer (GeoPackage; ETRS89-LAEA projection), HCEI raster, priority cell map, and all BHIF GIS pipeline code (QGIS model + Python plugins) are deposited at <https://github.com/BHIF-EU> and Zenodo <https://doi.org/10.5281/zenodo.19490003-data> under Apache 2.0 and CC BY 4.0 respectively. Input data (GBIF, IUCN, WDPa) are available from their respective open-access portals.

### Ethical Approval

This study involved only spatial analysis of existing published biodiversity occurrence records, IUCN range data, and protected area boundary data from open-access sources. No primary field surveys, specimen collection, or animal handling was conducted. No ethical approval was required.

---

## Appendix A

### **BHIF Pipeline Technical Specification and Hotspot Polygon Catalogue**

QGIS model workflow diagram and Python plugin code documentation. KDE bandwidth selection method and species-specific bandwidth values. BHIF composite weighting scheme with sensitivity analysis showing results under alternative weightings. Full catalogue of 284 BHIF hotspot polygons with area, centroid coordinates, and primary biodiversity value contributing to identification.