



Challenges and limitations of modelling historical spatial data on nature: 19th century flora and fauna of Peloponnese, Greece

Eleni Gkadolou^a (corresponding author), Katerina Koutsovoulou^a, Georgios Toliás^{b,c} and Vassilis Detsis^a

elenigadolou@gmail.com, kkoutsovoulou@gmail.com, gtoliás@eie.gr, detsis@hua.gr

^a Department of Economics and Sustainable Development, Harokopio University, Athens, Greece

^b Institute of Historical Research, National Hellenic Research Foundation, Greece

^c École pratique des hautes études – Paris Sciences et Lettres, France

Abstract. The aim of the study is the consolidation of a methodology suitable for organizing, utilizing and visualizing information regarding species distribution provided as text in historical sources. The work of the French Scientific Expedition in 1829 in Peloponnese, Greece, was used as a case study. We propose a system organized in three geographical levels: for information referring to a certain locality the form of a grid is appropriate, otherwise polygons depicting historical administrative areas or the whole region of Peloponnese should be preferred. There are three important caveats to avoid. First, species presence referring to an administrative area or region does not equal with presence in every locality and should not be transferred to the level of grid mapping, respectively reference of presence in the region should not be transferred to the administrative units' level. Second, historical sources refer to species names that often are no longer valid; this kind of data must be referred to currently valid species names. Third, absence of reference of species presence should not be misinterpreted as absence of the species.

Keywords: historical flora, historical fauna, database, map, spatial

1 Introduction

Current ecological patterns can sometimes be attributed to processes that took place in the distant past, a pattern that is especially important in areas with a long history of human presence and human-induced nature alterations (e.g., Gamboa-Badilla et al., 2020; Saatkamp et al., 2020). The major assumption, explicitly or implicitly, underlying the task of studying past regimes is that we can document a valid reconstruction of former configurations and transformations of nature (Meyer and Crumley, 2012). This exercise is of limited value unless the information can be convincingly related to a

particular place. Progress in the exploration of how geographical change has occurred over time has been slow due to a) the complexity of the uniform cartographic representation of data derived from various heterogeneous sources and b) the insufficient means available to deal with the unavoidable uncertainties concerning the information as such and the geographical location of reference.

The use of layers within a GIS provides a basis to deal with the inherent spatial and temporal complexities providing means to overcome the difficulties (Gregory and Healey, 2007). Despite the progress that has been made in various fronts, a prototyping of spatial historical information that can support its optimal utilization and allow synergies among applications is still wanted.

The goal of the present study is to develop a spatial database for mapping, analysing and visualizing data about nature concerning time scales that span decades to centuries. This infrastructure would allow the spatial linking between historical and current data about species, habitats and land cover types and support a reconstruction and interpretation of their historical spatial distribution and its dynamics.

We use as reference material species lists of Peloponnese, Greece, which were published by the “French Scientific Expedition” that took place in 1829. In the following the methodological process is presented while keeping the biology-related details to the absolute minimum.

2 Materials and methods

2.1 Sources

Naturalist missions in 18th -19th century have published their results in large corpuses including taxonomies, sketches, narrative descriptions and geographic

references of their observations. In Greece, one of the most important sources of historical information regarding nature is the work of the French Scientific Expedition in 1829 in Peloponnese. The full results of the Expedition's survey were published in 8 volumes (three of them concerned flora and fauna) along with a map and an atlas. The map was georeferenced and spatial data were extracted and stored in a geo-database in terms of the previous project entitled "The historical landscape at the end of the Greek Revolution: The French Scientific Expedition of Moreas, 1829" (Gkadolou, 2019), available at: <https://moree1829.gr>. This database (called "Moreas database" onwards) includes historical place names of 1828 and their location (in form of a gazetteer) and it was used as a reference in order to geo-locate the flora and fauna data of this research. The species data used in this study were drawn from two of the volumes: Saint-Hilaire et al. (1832) and Bory de Saint-Vincent et al. (1832).

2.2 Species lists

The first step towards species identification was to perform a cataloguing exercise to link the species mentioned in the Expedition's lists to currently valid species names. Reliable sources and databases were then selected for each species group followed by communication with experts, where required. In order to assess taxonomic uncertainties, species information and drawings presented in the Expedition's reports were used, when possible.

In total 1,558 plant and lichen species were recorded by the Expedition. The present study focuses on the 1,376 Vascular Plant species (Angiosperms, Gymnosperms, Pteridophytes) recorded for Peloponnese and excludes species found elsewhere or lacking location information. Also, some species mentioned in the lists could not be attributed to a valid species name. Concerning vertebrate animal species, the Expedition's reports similarly document 21 mammal species including 9 domesticated ones, 58 bird species, of which 55 wild species and 3 domestic and a total of 31 amphibian and reptile species for the Peloponnese.

2.3 Locating species presence

The Expedition's lists provide spatial descriptions for the locations of the species presence. According to the level of detail, these descriptions were classified into three categories: a) regional scale i.e., "Le Péloponèse", b) administrative scale, based on the administrative units of 1828 (e.g., La Laconie), and c) local scale on which

attestations are based on a specific place (e.g., L'île Sapiencie - Sapiencie island) or broader area (e.g., La côte maritime entre Coron et Modon - the seacoast between Koroni and Methoni). For geo-locating the latter, the "Moreas database" was used (see 2.1 Sources).

Some of the local scale attestations are quite vague. This vagueness comes either directly from the original narration (e.g., Les environs de Modon et sites analogues, aux lieux herbeux - The surroundings of Methoni and similar sites, in grasslands) or indirectly from the "today" interpretation of the linguistic meaning of that period's attestation (e.g., La région inférieure - The lower region). For those cases, several assumptions had to be taken based mostly on the geomorphological characteristics of the area.

A 10 km x 10 km grid was constructed (Fig. 1). After geo-locating the data by matching the descriptions with the historical place names from the "Moreas database", each location was assigned to the cells it falls within. The size of the cells was decided by considering the descriptions retrieved from the text that could be approximately quantified. Thus, the scale used was defined by the available description and not by the attributes of the different species. In the case of the administrative or the local scale a species can be referred to more than one unit depending on its distribution (Fig. 1).

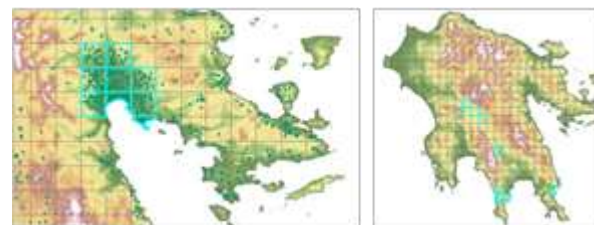


Figure 1: The location "Plaine d'Argos"-Argos plain represented by nine cells (left image) and the different locations where *Ballota acetabulosa* (L.) Benth is present (right image).

In order to store and manage the spatial and descriptive information of the species, a database was developed according to the conceptual schema illustrated in Fig. 2 and the main entities (classes) are:

- "Place" for representing the geographic location of the species as this is recorded in the texts characterized by the level of detail of the description
- "SpatialUnit" that refers to the grid cells
- "InformationResource" (the historical texts)
- "Time", i.e., year of recording

- “Flora” for representing the flora species
- “Fauna”, further classified to “Mammal”, “Bird” and “Reptile-Amphibian”

The attributes of “Flora” and “Fauna” include information on the taxonomy of the species (both at the time of publication and the currently valid status), the location description in the texts and data concerning the modern status of the species.

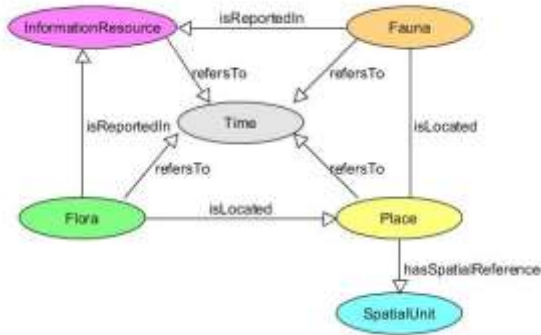


Figure 2: Diagram of the main classes of the database.

2.4 Software and Data Availability Sub-Section

The project is still ongoing, so the database is not finalized yet.

3 Results

Plant species identification and their assignment to current taxonomic nomenclature was achieved for 1,219 out of 1,322 species (92,2%). However, 236 species (17,9%) were excluded because evidence suggested that either these species were erroneously identified, or the taxonomy of their groups was so radically altered that linking them to a currently valid name is impossible.

This process left 983 species that have been geo-located (Figs. 1, 3 and Tab. 1), the majority of which have been mapped in the local scale. It is noted that a species might be located in more than one area of different levels of detail according to the description in the texts (e.g., a species is present in Kalamata city and in the administrative unit of Argolide).



Figure 3: The distribution of Eurasian jay (*Garrulus glandarius*) in Arkadie and Argolide administrative units of 1828.

Nineteen out of 21 mammal species (90,5%), all 58 birds (100%) and 27 out of 29 reptiles/amphibians recorded from the Peloponnese (93%) were identified and assigned to current taxonomic nomenclature. Four mammals, 2 reptiles and one amphibian were excluded from the database for the same reasons explained for plants above. All registered occurrences have been located on the map according to the level of detail of the description of the geographic location (Tab. 1). Only few animal species were recorded at the local level (Fig. 4).



Figure 4: The distribution of the six species of birds mapped at the local level.

It must be noted that in some cases different species (records) in the Expedition lists may be assigned to a single one based on current taxonomy. Therefore, in total 953 plant species (983 records), 17 mammals (17 records), 58 birds (58 records), 4 amphibians (5 records) and 17 reptiles (21 records) are included in the database.

Table 1: The number of species recorded according to the level of detail of the description of the geographic location.

Level of detail	Vascular plants	Mammals	Birds	Reptiles- Amphibians
Regional	778	13	49	15
Administrative unit	154	-	3	-
Local	159	4	6	11

4 Discussion

Until the late 20th century most species collections or citations of species presence were in form of a text relating the point of collection or observation to an easily recognizable landmark that was expected to be found in a map. This kind of data are usually being pooled as species presence in wider regions so that more detailed geographical information is lost. Efforts to retrieve this information are only slowly beginning to emerge. In terms of this research, we dealt with the following issues regarding historical geographic information.

The availability of contemporary cartographic material is extremely important (Buldrini et al., 2019). The historical place names from the “Moreas database” forms a valuable dataset of the toponyms used during the first half of the 19th century for Greece and are matched to the current ones. Since then, many place names changes have occurred and thus, in lack of an official historical gazetteer, spatialization of historical sources (extraction of geographical information) is difficult or implemented in a not formalized way. This gazetteer is now being implemented as Linked Open Data and part of the World Historical Gazetteer and linked to other gazetteers in the framework of a larger effort for creating an “ecosystem of past places”. As a result, the gazetteer will be then easily integrated in Name Entity Recognition tools (such as Recogito Annotation platform) and will permit the spatial annotation of historical texts based on automated matching to the gazetteer records. The reason for not using Name Entity Recognition tools in this phase was that the available tools cannot handle efficiently

descriptions that do not refer to a toponym but to a location relative to an identifiable toponym (e.g., “around A”, “in the high altitudes of B”, “between A and B”) (McDonough et al., 2019). Even if in our case the matching of textual attestations to the historical place names was not implemented automatically, yet it is of great value that flora and fauna data of an official resource of 19th century were correlated to this gazetteer. As a result, flora and fauna data can be also re-used as Linked Open Data. Furthermore, moving a step forward, we proceeded from point locations of species data (that is typical of relevant applications) to polygon or raster ones as an effort to implement a closer to reality geo-location of species and to preserve the actual geographical descriptions from the historical texts.

With respect to the actual distribution of the listed species there are two issues that must be dealt with. The first is the well-known bias in favour of presence. The produced maps should be treated as presence maps only and absence should not be inferred solely from lack of reference. The second issue is the bias against spatial accuracy in the case of common species (McClenachan et al., 2015), contrary to species found in few places, which are explicitly mentioned and can be mapped in greater detail. We suggest in similar cases the use of polygons that correspond to the higher level of description. These data should not be transferred to finer resolutions (in our case the raster grid), since this would be wrongly taken to imply that a species was present in each single cell.

The visualization of the spatial data is also an important issue regarding the interpretation of information and the creation of digital maps. Different criteria for visualizing the geospatial data must be applied in order to explore and interpret data and infer new knowledge as well as to indicate the levels of accuracy of the historical information. These criteria should aim at identifying spatial patterns, relationships between entities and statistical comparisons. Taxonomic ambiguities present a great challenge. Many species has changed name in the meanwhile, while whole groups of species have been reconsidered, new species or genus described, and formerly separate groups are now considered synonyms. In many cases a valid name can be related to the name in the dataset, in other cases, and provided that one or more specimens were deposited in a museum, the examination of the original material is the only way to be certain. Without going into detail, clarifying the object of mapping is one of the more complicated and tedious tasks. Mapping the species by

the names given to them by the collectors or surveyors very often does not produce any useful information.

A more effective optical character recognition taking into account species names, so as to avoid mistaken transcriptions, as well as interlinking of species taxonomic databases could speed up the procedure of species synonyms identification and assignment to current taxonomic nomenclature. However, interpretation of species synonyms will require human intervention for the foreseeable future, even if the relevant databases were interlinked, since organism groups are continuously being split, merged or renamed and an effective ontology capable of replacing free text descriptions has yet to be developed (Deans et al. 2012). Even in the same collection organisms referred to with a single name may belong to different currently accepted species, as well as the opposite, if a particular organism group or groups is not being constantly revised.

5 Acknowledgments

This research is co-financed by Greece and the European Union (European Social Fund- ESF) through the Operational Programme «Human Resources Development, Education and Lifelong Learning 2014-2020» in the context of the project “Degradation and Conservation of Nature: A Digital Atlas of the Peloponnese, 19th-21st century” (MIS 5049018).

6 References

Bory de Saint-Vincent, J.-B.-G.-M., Fauché, J.-B., Chaubard, A., *Expédition scientifique de Morée. Section des sciences physiques. Tome III. 2e partie. Botanique*, Paris 1832.

Buldrini, F., Muzzi, E. and Pezzi, G.: Past landscape features may be captured by historical floristic data. A case study in northern Italy, *Bot. Lett.*, 166(1), 93–103, <https://doi.org/10.1080/23818107.2018.1563567>, 2019.

Deans, A.R., Yoder, M.J., Balhoff, J.P.: Time to change how we describe biodiversity, *Trends in Ecology & Evolution*, 27(2), 78-84, <https://doi.org/10.1016/j.tree.2011.11.007>, 2012.

Gamboa-Badilla, N., Segura, A., Bagaria, G., Basnou, C. and Pino, J.: Contrasting time-scale effects of land-use legacy on species richness, diversity and composition in Mediterranean scrubland communities,

Landsc. Ecol., 35(12), 2745–2757, <https://doi.org/10.1007/s10980-020-01087-5>, 2020.

Gkadolou, E.: Spatial Organization and Semantic Modelling of Historical Data: The Case of the French Scientific Mission, 1828-29, *Int. J. Comput. Methods Herit. Sci.*, 3(2), 43–57, <https://doi.org/10.4018/IJCMHS.2019070103>, 2019.

Gregory, I. N. and Healey, R. G.: Historical GIS: structuring, mapping and analysing geographies of the past, *Prog. Hum. Geogr.*, 31(5), 638–653, <https://doi.org/10.1177/0309132507081495>, 2007.

McClenachan, L., Cooper, A. B., McKenzie, M. G. and Drew, J. A.: The Importance of Surprising Results and Best Practices in Historical Ecology, *BioScience*, 65(9), 932–939, <https://doi.org/10.1093/biosci/biv100>, 2015.

Meyer, W. J. and Crumley, C. L.: Historical Ecology: Using What Works to Cross the Divide, in *Atlantic Europe in the First Millennium BC*, edited by T. Moore and X.-L. Armada, pp. 109–134, Oxford University Press, <https://doi.org/10.1093/acprof:osobl/9780199567959.03.0003>, 2012.

Saatkamp, A., Henry, F. and Dutoit, T.: Romans Shape Today’s Vegetation and Soils: Two Millennia of Land-Use Legacy Dynamics in Mediterranean Grasslands, *Ecosystems*, <https://doi.org/10.1007/s10021-020-00581-w>, 2020.

Saint-Hilaire, I.G., Saint-Hilaire, E.G., Deshayes, G.P., *Expédition scientifique de Morée. Section des sciences physiques. Tome III. 1re partie. Zoologie. 1re section. Animaux vertébrés, mammifères et oiseaux, reptiles et poissons, mollusques et polypiers*, Paris 1832.

Wehrend, S. and Lewis, C.: A problem-oriented classification of visualization techniques, in *Proceedings of the First IEEE Conference on Visualization: Visualization '90*, pp. 139–143, <https://doi.org/10.1109/VISUAL.1990.146375>, 1990.