KARST WATERS INSTITUTE

LABORATOIRE SOUTERRAIN

Special Publication 6

Mapping Subterranean Biodiversity
Cartographie de la biodiversité souterraine

Proceedings of an international workshop held
March 18 through 20, 2001
Laboratoire Souterrain du CNRS, Moulis, Ariege, France

Edited by
David C. Culver, Louis Deharveng, Janine Gibert, and Ira D. Sasowsky

Mapping subterranean biodiversity: structure of the database and mapping software CKMAP and report of status for Italy

Fabio Stoch

Museum of Natural History
Verona, Italy

Address for correspondence:
viale XXV Aprile 24,
I-34015 Muggia (Trieste)
Italy

Italy has some of the best species-based distribution data of any country in the world (Stoch 2000) and a complete checklist of its fauna is available (Minelli *et al.*1993-95). Based on the checklist and the amendments introduced by taxonomists contributing to the database, the Italian fauna includes nearly 57,500 species and 3750 subspecies, approximately 10% of them being endemics. This fauna is probably one of the richest in Europe (Minelli 1996). The distributional database of Italian fauna includes nearly 5000 species (mainly invertebrates) and over 250,000 distribution records. Considering that 27% of the Italian territory is karstic, and more than 33,000 caves are included in the Italian survey (Stoch 2001), it is reasonable to suppose that the number of obligate cave-dwelling species is high as well. This assumption is reinforced by the results given by Culver *et al.* (1999), who demostrated a close relationship between number of species and number of caves per county in the United States where nearly 45,000 caves are known. Recent papers and books (Cobolli-Sbordoni *et al.*1994, Stoch 2001) confirm the high diversity and the antiquity of troglobitic and stygobitic fauna of Italy, where some of the richest hotspots of subterranean biodiversity are located (Culver and Sket 2000).

Unfortunately, knowledge of the distribution of subterranean biodiversity in Italy remains poor in comparison to epigean taxa. Considering that most of obligate cave-dwelling species in Italy are narrow endemics (Cobolli-Sbordoni et al., 1994; Stoch, 1995, 2001), a complete inventory of subterranean species distribution is needed for at least two reasons: a) endemic species have been a popular focus for studies of biodiversity and conservation (Williams et al., 2000) and a significant aid in the identification of priority areas for conservation (Deharveng, 2000 and contributors in the same volume); b) endemism is one of the criteria adopted by the Habitat Directive of European Community for selecting key-species for "Nature 2000" network (European Commission, 1992). The purpose of this short communication is to illustrate the methods adopted in Italy for implementing the database, the structure of the mapping program CKMAP (Stoch, 2000a) and the first results on the distribution of Italian obligate cave-fauna extracted from the database.

The digital database project of Italian fauna is financially supported by the Ministry of Environment and is being assembled by the Museum of Natural History in Verona and the Unione Zoologica Italiana. The data were compiled by well-experienced taxonomists (over 130 contributors) from published atlases and literature; additional records were obtained from museum and private collections, as well as from unpublished data sets given by the specialists (Stoch 2000). The ongoing project will include at the end of 2001 over 10,000 species of the Italian fauna; up to now (spring 2001) it includes 5337 invertebrate taxa (species and subspecies) and nearly 250,000 georeferenced distributional records. Within this data set, 899 species and their subspecies (903 total taxa) are subterranean, comprising 16.9% of the total taxa.. It is divided among subterranean habitats as follows: 317 (35.1% of the total number of subterranean species and subspecies) are exclusive of soil litter or MSS (Milieu Souterrain Superficiel); 265 (29.3%) are stybobionts; and 321 (35.6%) are terrestrial troglobionts and eutroglophiles up to now known only from cave habitats. A list of the taxa for which distributional data are available is given in Table 1. They include gastropods, oligochaeta, crustaceans, pseudoscorpions, centipedes, and insects (mainly coleopterans and orthopterans). The number of endemic species is incredibly high (685 subterranean species and subspecies - 75.85% - are narrow endemics, and 637 - 70.5% - are endemic to Italy).

The selection of the taxa to be included in the database followed precise guidelines (Pearson 1995, Stoch 2000). Criteria meeting the requirements of a conservation-oriented database include the degree of taxonomic knowledge, habitat specialization, and potential bioindicator value. The basic requirement was the inclusion of all the species and subspecies present in selected families (or higher taxa) to avoid any subjective choice.

The database was implemented on PC computers using the Windows®-based software MS Access®. This is a relational database-package useful for files not larger than 2 Gb; for larger data sets (*e.g.* millions of data entries) more sophisticated software is required. This is probably not the case even for a complex country like Italy, and it is not needed for subterranean biodiversity databases where data are usually scarce even in larger countries. For example, Culver *et al.* (1999) used less than 3000 locations in their survey of USA cave fauna. In any case, with millions of data entries, the user has two choices: to change database software package, or to divide the large database into smaller units (e.g. on a regional or taxonomic basis).

Specialists contributed their files in MS Access® tables, MS Excel® sheets, or even ASCII text delimited files; with modern database the data formats can be easily interchanged; these formats were adopted by Mac users as well.

The database structure is very simple. This is essential when exploring and querying thousands of data. Furthermore, a simple structure allows an easy link with other databases (at a regional level or at an European level) and metadata structures. For this reason, the number of "core" tables is limited to three (four in the new version): 1) taxon table (divided into a taxon table and a species table in the new version); b) distributional data table; c) references table (the last one, very simple, does not need to be considered further). The relational data structure is illustrated in fig. 1. Species and distributional data are joined using the species code; species names are repeated in the distribution table as well, were synonyms may be reported for literature records.

The taxon table includes taxon (subspecies, species, and genus) code, name and authorship, and several explanatory fields (dealing with standardized chorological groups, habitat types, feeding habits, sizes, etc.); higher taxa codes and names are added to this table (older version) or included in a relational taxon table (joined using a family-code in CKMAP 2000 version). The separation of the taxon table into two joined tables allows an easier programming of data exploration and more facilities to link other databases based on a common "higher taxon" structure (considering the high number of Italian endemics at the species level). The most important part of the table is the code, because it has a hierarchical structure (three digits for phylum code, subsequent 3 digits for class code, and so on for order, family, and - in the species table - genus, species and subspecies codes; every triplet is separated by a point; sometimes additional digits are added to avoid changes in the whole coding structure while adding new taxa). For example, the stygiobiont isopod family Sphaeromatidae has the code 029.010.008.051, derived from: phylum Arthropoda (029); class Malacostraca (029.010); order Isopoda (029.010.008). The sphaeromatid isopod *Monolistra racovitzai racovitzai* Strouhal, 1928 is joined to its family using the previous code and has a further code (030.197.0.005.0.001): Italian checklist reference (030 – this part is the number of the volume where the species is listed, and may be omitted) followed by genus *Monolistra* (197.0), species *M. racovitzai* (197.0.005.0), subspecies *M. r. racovitzai* (197.0.005.0.001). This code was appended to the checklist (full explanation of the usage of one digit extensions in Minelli *et al.* 2000). Intermediate taxa (as subclasses, subfamilies, etc.) are not included for the moment in the database. Hierarchy allows to select all the species or subspecies belonging to higher taxonomic units, as well as to reproduce the correct taxonomic sequence using a tree-like structure, currently available in Windows® programming languages.

The distributional data table includes: species code, species name (or synonym), locality entries (region, province, generic locality, detailed locality), data source (literature citation, linked with the references table; museum or private collection; original data), coordinates, and a field dealing with the accuracy of the coordinates. Localities are reported in two fields: the first one (generic locality) includes only data obtained from the index of a common road atlas of Italy (edited by the Touring Club Italiano – TCI – scale 1:250000), *e.g.*, a well known place name close to (or identical with) the locality reported in the original citation (literature, collection label); the second one (detailed locality) reports fine-scale data and habitat data (e.g. cave name), when available. The TCI place name is used for two purposes: a) easy location of the spot on a widely distributed commercial atlas; b) the possibility of assigning coordinates based on this place name (the full list is reported in a separate database table) in the case we have no more detailed geographical coordinates. For a cave-oriented database, we can join this table to a cave registry, generally available in most European countries.

Georeferencing of data is a complex problem; the database uses a UTM grid (10x 0 or 1x1 km grid cells, names following MGRS convention, datum European 1950), which is used also for mapping species distribution using the program CKMAP or commercial GIS software. Routines to convert between geographical coordinates and metric coordinates in different datums (ED 50, WGS 84 and the national datum Rome 40) were implemented in the database. Grid size (10 km or 1 km) gives the precision of the coordinates reported in the database. Precision is not synonymous with accuracy: the accuracy of georeferencing (included in a separate field) may be defined as good, poor, etc. depending on the quality of the data reported in the literature or in the labels accompanying the specimens.

The software programmed for exploring and mapping the data set is called CKMAP (Stoch, 2000a). It is not a new GIS package, but a simple mapping program for the exploratory analysis of data, with powerful statistical functions as in WORLDMAP (Williams, 1996). It has the following advantages over other software packages (see Fig. 2): a) hierarchical structure of the species tree, which allows an easy exploration of the species and taxon tables using a Windows Explorer @ – like structure; b) immediate mapping of the distribution of every subspecies, species, or higher rank taxon selected on the tree, using UTM 10 x 10 grid cells; c) the possibility of overlaying layers representing regions, provinces, DTM, reserve networks and hydrography (geological and climatic maps will be implemented); d) facilities to export distributional maps to commercial GIS software packages like ArcView@ and MapInfo@; e) interactive maps, easily zoomed, which allows the listing of all records included in the selected UTM grid cell; and f) database tables easily accessible and editable in spreadsheet format; g) biodiversity, rarity and endemism statistics and other calculations easily performed using simple button clicks or menus. CKMAP is intended as a user-friendly interactive software package which allows the exploration and mapping of hundred thousands data and complex base maps in fractions of second on the new Pentium@ based-machines.

Two outputs from the database and CKMAP software are presented as examples of the importance of mapping subterranean biodiversity and a synthesis of the distributional data of subterranean species and subspecies in Italy updated to spring 2001.

In the first example, Figure 3 displays the geographical variation of the number of obligate subterranean endemic species and subspecies among UTM 10 x 10 km grid cells in Italy. All of the taxa endemic to Italy (637 species and subspecies) were included in the analysis. Considering that several important taxa (like carabid beetles) will be added in a near future, the figure is given as a provisional report of status of subterranean fauna mapping in Italy, and no effort is made to explain the observed patterns. However, the map reported in Fig. 3 deserves some comments: a) regional trends of endemics richness are evident, and mark the extent of karstic areas; b) a higher number of endemics is displayed in the Prealps of northern Italy and, to a lesser extent, in Liguria, northern Appennines and Sardinia island, while richness decreases moving southwards. This pattern follows the overall trend of species richness in Italy delineated by Stoch (2000), and clearly shows a "peninsular effect" (Massa 1982) reflecting several historical and ecological factors (Stoch 1995), mainly size and age of karstic massifs. Unfortunately, heterogeneous sampling effort may influence the general trend, and careful statistical data analysis is needed to confirm this pattern.

The second example shows the overlay of the plots of subterranean endemics richness and European nature reserves network "Nature 2000" in an area located in the most important karstic massif of Italy, the Prealps. Figure 4 shows that the natural reserves network fails to represent subterranean endemic species. This failure is due to two main factors: a) the criteria for choices of natural reserves are usually oriented towards protection of "charismatic" surface species assemblages (mainly vertebrates – especially birds – and plants [Araújo 1999]); b) although caves are included in the list of habitats of communitary interest (appendix 1 of the Habitat Directive), and endemism is a key concept for nature protection in Europe (article 1), no troglobiont or stygobiont species are included in the list of appendix 2 (species of communitary interest: European Commission, 1992). For this reason, apart from the presence of bats or troglobitic amphibians, no "key species" are available to protect subterranean ecosystems in western Europe. Mapping the distribution of cave-dwelling species is thus an essential approach to propose an integration of the species lists of Habitat Directive (hence of the current reserve network) taking in account the protection of subterranean biodiversity.

Acknowledgements

This "state of the art" was presented to the international workshop "Mapping Subterranean Endemism" in Moulis (18-20 March 2001). I am grateful to: dr. Sandro La Posta (Italian Ministry of Environment) for the support given to the work, for providing access to the Italian reserve network maps and for the permission to use the data; dr. Giuseppe Messana (Florence), president of the SIBS (Societé International de Biospéléologie), for the interest and encouragement in presenting this paper; and to Prof. Dave Culver (Washington) for his kind help and revision of the English text. The data of subterranean species were entered into the database by the following specialists: M. Bodon, F. Giusti, G. Manganelli, S. Cianfanelli (Gastropoda); B. Sambugar (Oligochaeta); G. Gardini (Pseudoscorpionida); R. Argano, F. Stoch (isopods); S. Ruffo (Amphipoda, Thermosbaenacea, Bathynellacea); V. Cottarelli, R. Berera, D. Galassi, F. Stoch (Harpacticoida); M. Zapparoli (Chilopoda); P. Fontana, M. La Greca (Orthoptera); F. Penati, P. Vienna, S. Zoia, L. Latella, R. Poggi, G. Sabella, A. Zanetti, G. Osella (Coleoptera); all the intellectual rights on individual data belong to the specialists. Financial support is provided by the Ministry of Environment through a grant to the Museum of Natural History in Verona.

References

Araújo, M.B. 1999. Distribution patterns of biodiversity and the design of a representative reserve network in Portugal. *Diversity Distribution* 5:151-163.

Cobolli-Sbordoni, M., M. Paoletti, V. Sbordoni, A. Vigna Taglianti. 1994. Italie, pp. 737-751. In C. Juberthie and V. Decu [eds.], *Encyclopaedia Biospeologica, Vol. 1.* Moulis, France.

Culver D.C., H.H. Hobbs III, M.C. Christman, and L.L. Master. 1999. Distribution map of caves and cave animals in the United States. *J. Cave Karst Studies* 61:139-140.

Culver, D.C., and B. Sket. 2000. Hotspots of subterranean biodiversity in caves and wells. *J. Cave Karst Studies* 62:11-17.

Deharveng, L. 2000. The Workshop Endemism: introduction. *Belg. J. Entomol.* 2:5-8.

European Commission. 1992. Council Directive 92/43/EEC of 21 May 1992 on the conservation of natural habitats and of wild fauna and flora. O.J. L206, 22.07.1992.

Massa, B. 1982. Il gradiente faunistico nella penisola italiana e nelle isole. *Atti Società Italiana Scienze Naturali, Museo Civico di Storia naturale di Milano* 123: 353-374.

Minelli, A. 1996. La checklist delle specie della fauna italiana. Un bilancio del progetto. *Bollettino Museo Civico Storia naturale Verona* 20:249-261.

Minelli, A., Ruffo S. & La Posta S. [eds.], 1993-1995. – *Checklist delle specie della fauna italiana.* Fascicoli 1-110. Edizioni Calderini, Bologna.

Minelli, A., F. Stoch and S. Zoia. 1999. Aggiornamenti all checklist delle specie dell fauna Italia. *I. Contributo. Boll. Soc. Entomol. Ital.* 131-269-278.

Pearson, D.L. 1995. Selecting indicator taxa for the quantitative assessment of biodiversity, pp. 75-79. In D.L. Hawksworth [ed.], *Biodiversity measurement and estimation.* The Royal Society, Cambridge University Press, Cambridge.

Stoch, F. 1995. The ecological and historical determinants of crustacean diversity in groundwaters, or: why are there so many species? *Mémoires Biospéologie* 22: 139-160.

Stoch, F. 2000. How many endemic species? Species richness assessment and conservation priorities in Italy. *Belg. J. Entomol.* 2:125-133.

Stoch, F. 2000. *CKMAP 2000: software e help on-line, versione 3.0.* Ministero dell'Ambiente, Roma & Museo Civico di Storia Naturale, Verona.

Stoch F. [ed.]. 2001. *Grotte e fenomeno carsico.* Quaderni Habitat, Ministero dell'Ambiente, Roma & Museo Friulano Storia Naturale, Udine: 158 pp.

Williams P.H. 1996. *WORLDMAP 4 WINDOWS: software and help document 4.1.* Distributed privately, London. Available at http://www.nhm.ac.uk/science/projects/worldmap.

Williams P., C. Humphries, M. Araújo, R. Lampinen, W. Hagemeijer, J. P. Gasc, and T. Mitchell-Jones. 2000. Endemism and important areas for representing European biodiversity: a preliminary exploration of atlas data for plants and terrestrial vertebrates. *Belg. J. Entomol.* 2:21-46.

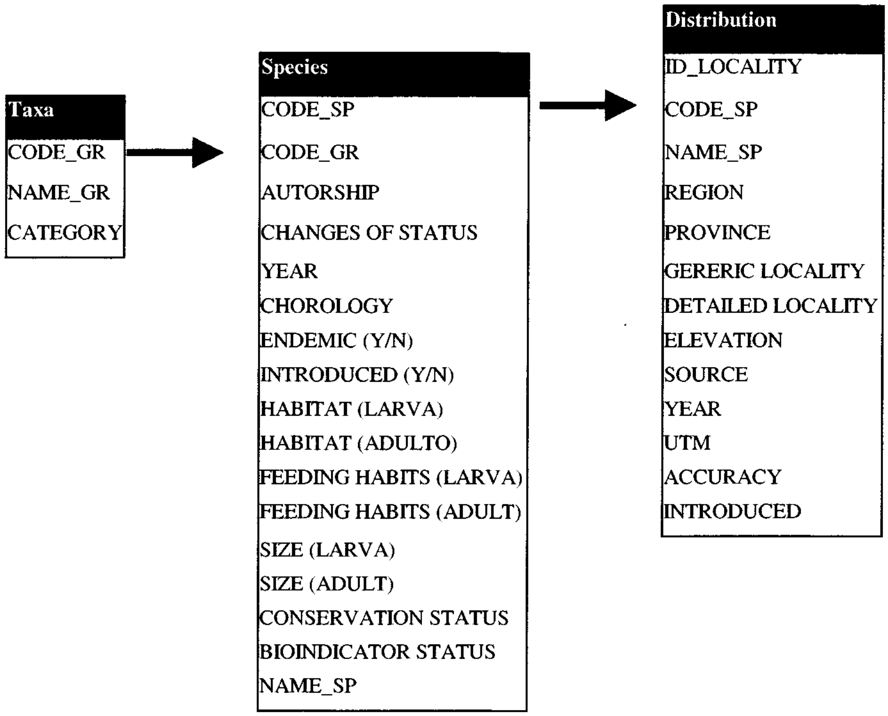

Fig. 1. Structure of the main tables of the database and their relationships.

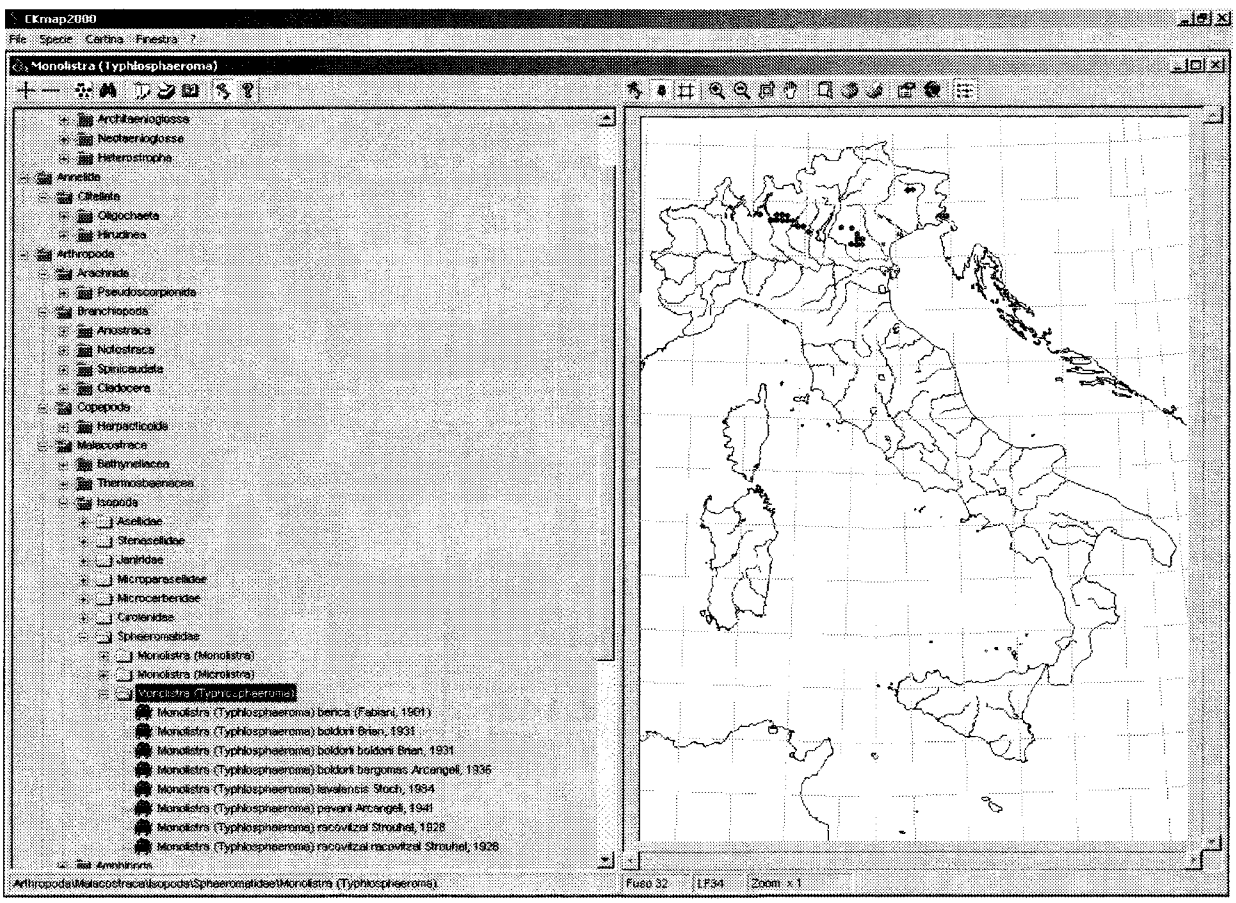

Fig. 2. Main screen of the CKMAP software

Table 1. List of the taxa including the subterranean species and subspecies entered in the Italian database up to spring 2001; numbers in brackets are comprehensive of endogeans and species of MSS.

		N sp+ssp	N records	Total sp.
Gastropoda	Hydrobiidae	49	1709	**49**
Oligochaeta	Lumbriculidae	5	15	
	Tubificidae	8	20	**13**
Pseudoscorpionida	Chthoniidae	30 (39)	202	
	Neobisiidae	41 (43)	425	
	Syarinidae	4	18	
	Larcidae	1	1	
	Chernetidae	3	21	**90**
Harpacticoida	Ectinosomatidae	3	3	
	Ameiridae	23	86	
	Tetragonicipitidae	1	1	
	Canthocamptidae	40	206	
	Cylindropsyllidae	11	32	
	Parastenocarididae	32	100	
	Latiremidae	1	3	**111**
Bathynellacea	Bathynellidae	7	7	**7**
Thermosbaenacea	Halosbaenidae	1	1	
	Monodellidae	3	7	**4**
Isopoda	Asellidae	24	105	
	Stenasellidae	3	15	
	Janiridae	1	1	
	Microparasellidae	7	18	
	Microcerberidae	1	1	
	Cirolanidae	2	2	
	Sphaeromatidae	12	102	**50**
Amphipoda	Bogidiellidae	8	15	
	Gammaridae	6	28	
	Hadziidae	4	9	
	Salentinellidae	2	48	
	Ingolfiellidae	1	1	
	Metaingolfiellidae	1	1	
	Pseudoniphargidae	5	8	
	Metacrangonyctidae	1	1	
	Atyidae	2	30	
	Palaemonidae	1	7	**31**
Chilopoda	Lithobiidae	14	122	
	Cryptopidae	1	2	
	Himantariidae	1	2	
	Geophilidae	1	1	**17**
Orthoptera	Rhaphidophoridae	18	423	
	Gryllidae	1 (6)	1 (205)	
	Gryllotalpidae	(8)	(189)	
	Acrididae	(3)	(445)	**35**
Coleoptera	Histeridae	8	35	
	Cholevidae	113 (268)	679 (4585)	
	Pselaphidae	81 (160)	2626 (3092)	
	Staphylinidae	3	5	
	Curculionidae	1 (57)	2 (207)	**496**

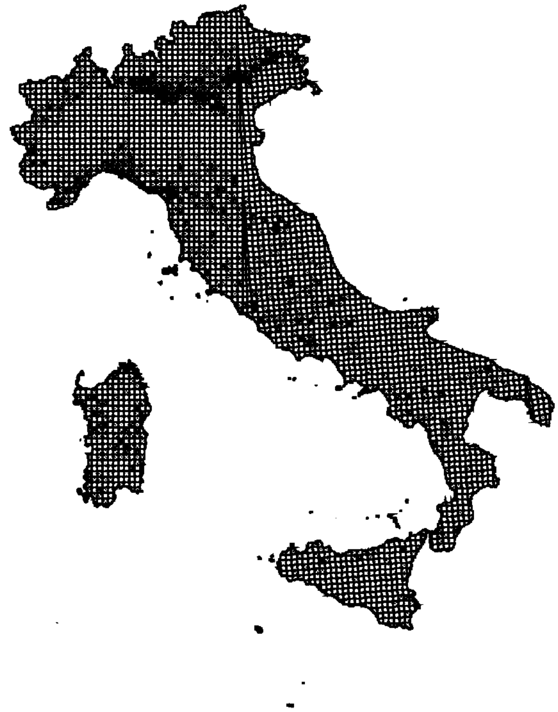

Fig. 3. Geographical variation of the number of obligate subterranean endemic species and subspecies (637 taxa) among UTM 10 x 10 km grid cells in Italy

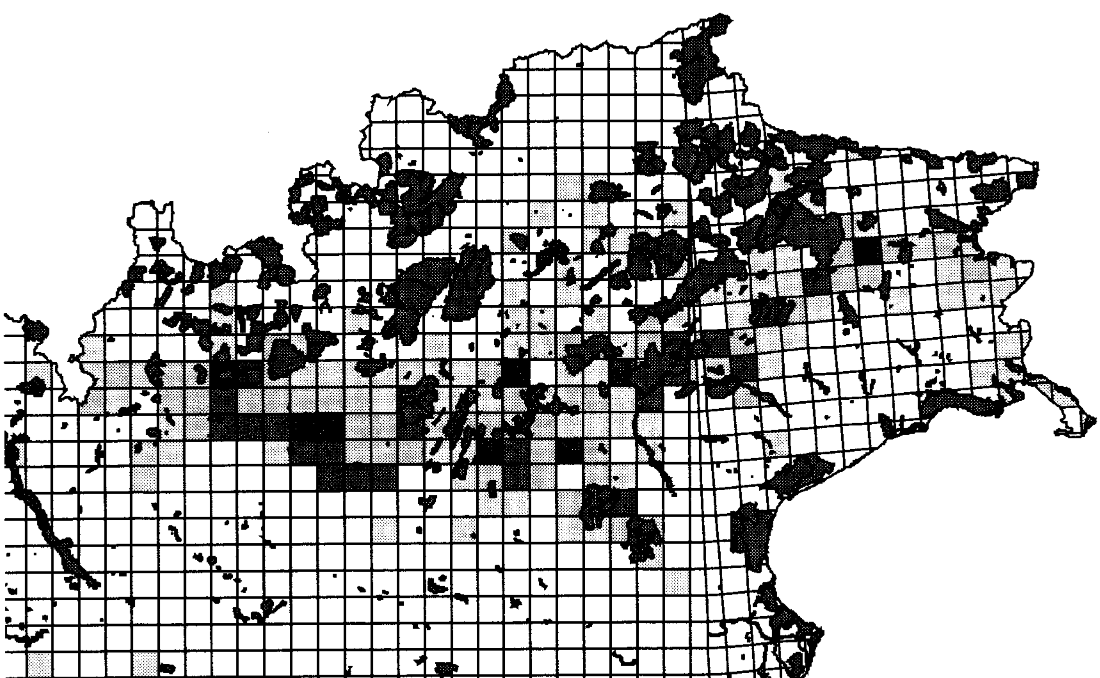

Fig. 4. Overlay of the plots of subterranean endemics richness and European nature reserves network "Nature 2000" on a UTM 10 x 10 km grid in Italy