Global Ecology and Biogeography

WILEY

Global Ecology and Biogeography
A Journal of Macroecology

DATA ARTICLE OPEN ACCESS

# EGCop: An Expert-Curated Occurrence Dataset of European Groundwater-Dwelling Copepods (Crustacea: Copepoda)

Francesco Cerasoli[1] | Barbara Fiasca[1] | Mattia Di Cicco[1] | Emma Galmarini[1] | Ilaria Vaccarelli[1,2] | Stefano Mammola[3,4,5] | Florian Malard[6] | Fabio Stoch[7] | Diana M. P. Galassi[1]

[1]Department of Life, Health, and Environmental Sciences, University of L'Aquila, L'Aquila, Italy | [2]University Institute of Higher Studies in Pavia, Pavia, Italy | [3]Molecular Ecology Group (MEG), water Research Institute (IRSA), National Research Council (CNR), Italy | [4]Laboratory for Integrative Biodiversity Research (LIBRe), Finnish Museum of Natural History (LUOMUS), University of Helsinki, Helsinki, Finland | [5]NBFC, National Biodiversity Future Center, Palermo, Italy | [6]University Claude Bernard Lyon 1, CNRS, ENTPE, Villeurbanne, France | [7]Department of Organismal Biology, Université Libre de Bruxelles (ULB), Brussels, Belgium

**Correspondence:** Diana M. P. Galassi (dianamariapaola.galassi@univaq.it)

**Received:** 20 August 2024 | **Revised:** 13 November 2024 | **Accepted:** 18 December 2024

**Handling Editor:** Fabien Leprieur

**Funding:** This research was funded by Biodiversa+, the European Biodiversity Partnership under the 2021–2022 BiodivProtect joint call for research proposals, co–funded by the European Commission (GA N° 101052342) and with the funding organisations Ministry of Universities and Research (Italy), Agencia Estatal de Investigación–Fundación Biodiversidad (Spain), Fundo Regional para a Ciência e Tecnologia (Portugal), Suomen Akatemia–Ministry of the Environment (Finland), Belgian Science Policy Office (Belgium), Agence Nationale de la Recherche (France), Deutsche Forschungsgemeinschaft e.V. (Germany), Schweizerischer Nationalfonds (Grant N° 31BD30_209583, Switzerland), Fonds zur Förderung der Wissenschaftlichen Forschung (Austria), Ministry of Higher Education, Science and Innovation (Slovenia), and the Executive Agency for Higher Education, Research, Development and Innovation Funding (Romania). S.M. acknowledges support by NBFC, funded by the Italian Ministry of University and Research, P.N.R.R., Missione 4, Componente 2, "Dalla ricerca all'impresa", Investimento 1.4, Project CN00000033. FC is granted by the PNR 2021–2027 (Italian Ministry of University and Research, DM 737/2021).

**Keywords:** biodiversity data | Copepoda | Europe | groundwater | protected areas | subterranean ecosystems

## ABSTRACT

**Motivation:** Subterranean biodiversity is increasingly threatened by multiple intertwined anthropogenic impacts, including habitat loss, pollution, overexploitation of resources, biological invasions and climate change. Worryingly, subterranean biodiversity is still poorly represented in conservation agendas, also due to persisting gaps in our knowledge of the organisms thriving in the often-secluded and difficult-to-access subterranean ecosystems. This is even more apparent for small-sized (body size < 1 mm) groundwater-dwelling metazoans, among which copepods (Crustacea: Copepoda) represent the dominant group in terms of both species richness and biomass.

We present a dataset including 6986 occurrence records of 588 species/subspecies of European obligate groundwater-dwelling copepods. We curated all records to make their taxonomy consistent with the current systematics of Copepoda, while assessing uncertainty in the geographic coordinates by coupling in-depth web and literature searches with GIS analyses. We suggest the data provided can be used to explore a range of eco-evolutionary questions—from the drivers of the distribution of groundwater fauna to the assembly of groundwater communities—as well as to prompt the conservation of groundwater biodiversity and more.

**Main Types of Variables Contained:** Occurrence records of groundwater-dwelling copepods, with details about specimen taxonomy, source of the record, occurrence locality and habitat type.

Diana Maria Paola Galassi and Fabio Stoch should be considered joint senior authors.

**Spatial Location and Grain:** Geographical Europe (including western Russian Federation), along with Turkey and Georgia. Occurrence records were assigned projected geographic coordinates (EPSG:3035) at 100 m resolution but with varying spatial uncertainty.
**Time Period and Grain:** 1907–2017.
**Major Taxa and Level of Measurement:** Crustacea: Copepoda. Most records have species-level identification, while some of them are identified at the subspecies level.
**Software Format:** Comma-separated values file (.csv) and Excel file (.xlsx), with UTF-8 encoding and meta-data provided following the Darwin Core standard.

## 1 | Introduction

Spatially accurate occurrence data are pivotal in ecology and conservation, allowing us to explore the drivers of species distribution, disentangle the factors structuring biotic communities, assess the anthropogenic impacts on biodiversity and ultimately define conservation targets (Rondinini et al. 2006; Ducatez, Tingley, and Shine 2014; Naimi et al. 2014; Console et al. 2020).

In the digital era, data availability has steeply increased thanks to international research networks developing publicly accessible online repositories, such as the Global Biodiversity Information Facility (GBIF) and the World Register of Marine Species (WoRMS), which contain taxonomic and distributional information for thousands of taxa across multiple ecosystem types. Yet, data shortfalls related to taxonomy, distribution patterns, phylogenetic relationships and other biodiversity facets persist in most biological groups (Hortal et al. 2015). Furthermore, broad digitised datasets often convey notable uncertainties in terms of taxonomic identification of specimens, spatial uncertainty of georeferenced records and data completeness with respect to the overall distribution of the target species, particularly for citizen science data (Soroye et al. 2022; Backstrom et al. 2024). Thus, the construction and sharing of biodiversity datasets undergoing thorough validation by expert taxonomists and ecologists should be encouraged to provide reliable data for (macro)ecological, biogeographical and conservation-oriented research (Hampton et al. 2013; Hobern et al. 2019; García-Girón, Bini, and Heino 2023; Santi et al. 2024). This is even more compelling for ecosystems whose biodiversity is still poorly known owing to access and sampling challenges, such as most subterranean ones (Zagmajster et al. 2010; Ficetola, Canedoli, and Stoch 2019; Mammola et al. 2021).

Terrestrial (e.g., caves, lava tubes and mines) and aquatic (e.g., lakes, rivers, pools and interstitial habitats) subterranean environments host a wide array of taxa specialised to a life in darkness. Given the multiple challenges of accessing these systems, distributional, functional and phylogenetic data for subterranean organisms have only started accumulating in recent decades (Zagmajster et al. 2014; Lunghi et al. 2018; Borko et al. 2021; Mammola et al. 2022). This is especially true for groundwater and groundwater-dependent ecosystems (Marmonier et al. 2023; Saccò et al. 2024), which show a general lack of publicly available and expert-curated datasets describing their biodiversity.

To partially fill this knowledge gap, we present here a novel dataset comprising occurrence records for all obligate groundwater-dwelling species (i.e., those accomplishing their life cycle exclusively in groundwater) of copepods (Crustacea: Copepoda), hereafter termed 'groundwater copepods', currently known from Europe, except for those exclusively occurring in anchialine environments (i.e., enclosed water bodies with an underground connection to the sea).

Copepods are crustaceans inhabiting all aquatic habitats on Earth, from the deep sea to moist soil litter (Macêdo et al. 2024). In groundwater, copepods are among the most representative animal groups, in terms of both abundance and species richness (Galassi, Huys, and Reid 2009), and show a high frequency of narrow endemics, with high species turnover even at regional-to-local scale (Galassi, Stoch, and Brancelj 2013; Cerasoli et al. 2023; Galmarini et al. 2023). Additionally, several groundwater copepods are remnants of ancient phylogenetic lineages which disappeared from surface waters after colonising subterranean environments (Galassi, Huys, and Reid 2009).

A partial version of the dataset we present was part of the 'European groundwater crustacean dataset' (EGCD), used to analyse the patterns of range size and beta diversity of all European groundwater crustaceans (Zagmajster et al. 2014). That version, which is not publicly available, was also used to highlight diversity hotspots for European groundwater harpacticoid copepods and to delimit the corresponding conservation-relevant areas (Iannella et al. 2020a, 2021). Over the last few years, we put much effort into improving data quality to provide (macro)ecologists, conservationists and biogeographers with an updated, reliable and openly accessible occurrence dataset for groundwater copepods of Europe. Specifically, we made necessary nomenclatural updates owing to taxonomic revisions (e.g., Galassi et al. 2017, 2019; Iepure, Bădăluţă, and Moldovan 2021), added records for additional species and reduced the spatial uncertainty of the georeferenced records whenever possible.

## 2 | Methods

The dataset, named 'EGCop' (standing for 'European Groundwater-dwelling Copepods'), contains 6986 records of groundwater copepods, spanning 36 European countries (also considering the Russian Federation, Turkey and Georgia). The spatial extents covered by the data are 28.398° W—43.939° E and 28.133° N—79.135° N. The names of the variables included in the dataset and the corresponding subjects are reported in Table 1.

**TABLE 1** | Variables, and corresponding subjects, included as columns in EGCop. For those variables included in the Darwin Core List of Terms, the name of the relevant term is reported as variable's subject, along with a link to the Darwin Core webpage where the term is explained.

| Variable name | Variable subject |
|---|---|
| ID | Numeric identifier of each record |
| Order | "dwc:order" (http://rs.tdwg.org/dwc/terms/order) |
| Family | "dwc:family" (http://rs.tdwg.org/dwc/terms/family) |
| Subfamily | "dwc:subfamily" (http://rs.tdwg.org/dwc/terms/subfamily) |
| Genus | "dwc:genus" (http://rs.tdwg.org/dwc/terms/genus) |
| Subgenus | "dwc:subgenus" (http://rs.tdwg.org/dwc/terms/subgenus) |
| SpecificEpithet | "dwc:specificEpithet" (http://rs.tdwg.org/dwc/terms/specificEpithet) |
| InfraspecificEpithet | "dwc:infraspecificEpithet" (http://rs.tdwg.org/dwc/terms/infraspecificEpithet) |
| ScientificName | "dwc:scientificName" (http://rs.tdwg.org/dwc/terms/scientificName) |
| ScientificNameAuthorship | "dwc:scientificNameAuthorship"(http://rs.tdwg.org/dwc/terms/scientificNameAuthorship) |
| AcceptedNameUsage | "dwc:acceptedNameUsage" (http://rs.tdwg.org/dwc/terms/acceptedNameUsage) |
| NamePublishedInYear | "dwc:namePublishedInYear" (http://rs.tdwg.org/dwc/terms/namePublishedInYear) |
| Country | "dwc:country" (http://rs.tdwg.org/dwc/terms/country) |
| Locality | "dwc:locality" (http://rs.tdwg.org/dwc/terms/locality) |
| LocationRemarks | "dwc:locationRemarks" (http://rs.tdwg.org/dwc/terms/locationRemarks) |
| Habitat | "dwc:habitat" (http://rs.tdwg.org/dwc/terms/habitat) |
| DecimalLongitude | "dwc:decimalLongitude" (http://rs.tdwg.org/dwc/terms/decimalLongitude) |
| DecimalLatitude | "dwc:decimalLatitude" (http://rs.tdwg.org/dwc/terms/decimalLatitude) |
| GeodeticDatum | "dwc:geodeticDatum" (http://rs.tdwg.org/dwc/terms/geodeticDatum) |
| Easting_EPSG3035 | Easting coordinates in ETRS89-extended/LAEA Europe (EPSG: 3035) projected coordinate system |
| Northing_EPSG3035 | Northing coordinates in ETRS89-extended/LAEA Europe (EPSG: 3035) projected coordinate system |
| Coord.Uncertainty | Spatial uncertainty in the provided coordinates (spanning from 0.1 km to "Region") |
| Coord.Validator | Person who performed the assessment of the spatial uncertainty in the provided coordinates |
| AssociatedReferences | "dwc:associatedReferences" (http://rs.tdwg.org/dwc/terms/associatedReferences) |
| Source | Literature item, specimens' collection, or research project from which the single records were obtained |
| ResearchGroup | Researchers involved in the project during which the single records were collected |

### 2.1 | Taxonomic Validation

We carried out taxonomic validation of all the records using specialised literature (Borutzky 1952; Dussart 1967, 1969; Boxshall and Halsey 2004; Wells 2007) and online resources such as the Virtual Copepod Library (https://www.marinespecies.org/copepoda/aphia.php?p=sources), critically analysed in the light of the deep expertise in copepod taxonomy and biogeography available in the research group (e.g., Galassi and De Laurentiis 2004; Stoch 2006; Galassi, Huys, and Reid 2009; Galassi et al. 2017, 2019; Stoch and Galassi 2010). We updated generic allocations of species that have changed following recent taxonomic revisions, in accordance with the International Code of Zoological Nomenclature (ICZN 1999). To make the dataset as comprehensive as possible, we retained also records of specimens identified only at the genus level (e.g., *Parastenocaris* sp.), as well as those of new species still lacking formal description (e.g., *Nitocrella* sp. J1).

### 2.2 | Spatial Validation

We further curated the dataset in R version 4.2.2 (R Core Team 2022), using various packages devoted to data handling and visualisation.

We validated the geographic coordinates (WGS84 datum, EPSG: 4326) of all the records as follows:

1. Browsing web resources (e.g., cave cadastres, technical reports by public water operators and websites of local authorities) to gather additional information for each sampling locality (reported in the 'Locality' column), and later retrieving the corresponding geographic coordinates using Google Earth Pro software.
2. Georeferencing the sampling localities displayed in maps reported within the consulted literature, using the QGIS software (version 3.22.8).

We re-projected all coordinates to the ETRS89-extended/LAEA Europe (EPSG: 3035) reference system, subsequently assigning each record to one of the following categories of spatial uncertainty: 0.1, 0.2, 0.5, 1, 2, 3, 5, 10, 20 km; Comm.Centr (standing for 'commune centroid'); Catchment; Region. To assign a record to an uncertainty category of $x$ kilometres, a point location was identified through the coordinates of that record. Then, we verified that the corresponding sampling locality (e.g., a karst spring, a well, the portion of a river flowing across a certain commune) was entirely encompassed by a circular buffer, around that point, having a radius equal or lower than $x$ kilometres. Worth to note, the coordinates (and related uncertainties) of the specimens sampled from caves refer to the cave entrance.

### 2.3 | Data Characterisation

We further analysed the curated dataset by assessing coverage of the occurrence records by protected areas (PAs) and groundwater habitat types, considering only the records with spatial uncertainty ≤ 5 km.

For PAs, we used spatial vector data of the following protection networks: Natura 2000 (N2k) and Emerald Network (Em.Net), relying on the datahub of the European Environment Agency (EEA, https://www.eea.europa.eu/en/analysis); Nationally designated PAs (Nat.designated), extracted from the World Database on Protected Areas (WDPA), updated to May 2024 (https://www.protectedplanet.net/en/thematic-areas/wdpa?tab=WDPA).

For groundwater habitat types, we used the map by Cornu, Eme, and Malard (2013), aggregating the corresponding features as follows: all aquifers in unconsolidated sediments were labelled as 'Uncons.Sediments', irrespective of their permeability; aquifers in consolidated rocks with high permeability were labelled as 'Cons.Rocks_High.Perm', while those in consolidated rocks with medium-to-low permeability were labelled as 'Cons.Rocks_MedLow.Perm'; practically non-aquiferous rocks were termed 'NonAq.Rocks', while freshwater bodies, lacustrine waters and glaciers were aggregated as 'Freshwaters & Glaciers'.

## 3 | Results

Most records corresponded to species of the orders Cyclopoida ($n$ = 3664; 184 species/subspecies) and Harpacticoida ($n$ = 3288; 395 species/subspecies). Calanoida were represented by 32 records (7 species), while Gelyelloida were only represented by unique records of the two known species (*Gelyella droguei* Rouch & Lescher-Moutoué, 1977 from France, and *Gelyella monardi* Moeschler & Rouch, 1988 from Switzerland).

Occurrences were spread across most of Europe for both Cyclopoida and Harpacticoida. However, Harpacticoida were reported in Fennoscandia more than Cyclopoida, while the latter were more represented than Harpacticoida in Great Britain and Eastern Europe. Records of Calanoida were mostly concentrated in the Balkans, with a single record in France and another one in the Crimean Peninsula (Figure 1).

We retrieved most records from literature published from the late 1950s onwards, with a peak in the early 2000s (Figure 2a).

After the spatial validation, most records showed low uncertainty in coordinates (Figure 2b): records having spatial uncertainty ≤ 1 km amounted to 4675 while less than 300 records still showed high uncertainty (i.e., 20 km, 'Catchment' or 'Region').

About 60% of the records ($n$ = 4200) fell within at least one of the three categories of protected areas considered, primarily the N2k network (Figure 2c). As for groundwater habitat types, 'Cons.Rocks_High.Perm' was the most represented for Cyclopoida, while aquifers in 'Cons.Rocks_MedLow.Perm' encompassed most of Harpacticoida records, followed by unconsolidated sediments (Figure 2d). However, as the spatial uncertainty of some records amounted to a few kilometres, these shares are approximations.

## 4 | Discussion

EGCop includes about 7000 expert-curated occurrence records of European groundwater copepods. Thanks to the performed spatial validation, most records show low spatial uncertainty (i.e., ≤ 1 km). This will facilitate spatially explicit (macro)ecological and biogeographical analyses and may also favour future re-sampling of specimens, allowing researchers to perform molecular analyses and build phylogenies for measuring phylogenetic diversity, a facet still poorly investigated for groundwater copepods.

The number of records in EGCop equals about one-third of the crustacean occurrences ($n$ = 21,700) forming EGCD, the most comprehensive dataset of European groundwater crustaceans assembled so far (Zagmajster et al. 2014). Further, the addition of new records and the performed taxonomic validation led to 588 species/subspecies comprised in EGCop compared to 547 species/subspecies of copepods represented in EGCD, out of a total of 1570 species/subspecies of European groundwater crustaceans (Zagmajster et al. 2014). These shares highlight the notable contribution of copepods to the biodiversity of groundwater, where they also represent a key node within the truncated subterranean food webs (Gibert and Deharveng 2002; Galassi, Huys, and Reid 2009) and provide fundamental ecosystem services such as natural attenuation of microbial biofilms and removal of contaminants (Griebler and Avramov 2015).

The gathered occurrences are more rarefied at higher latitudes, especially in areas completely covered by ice sheets

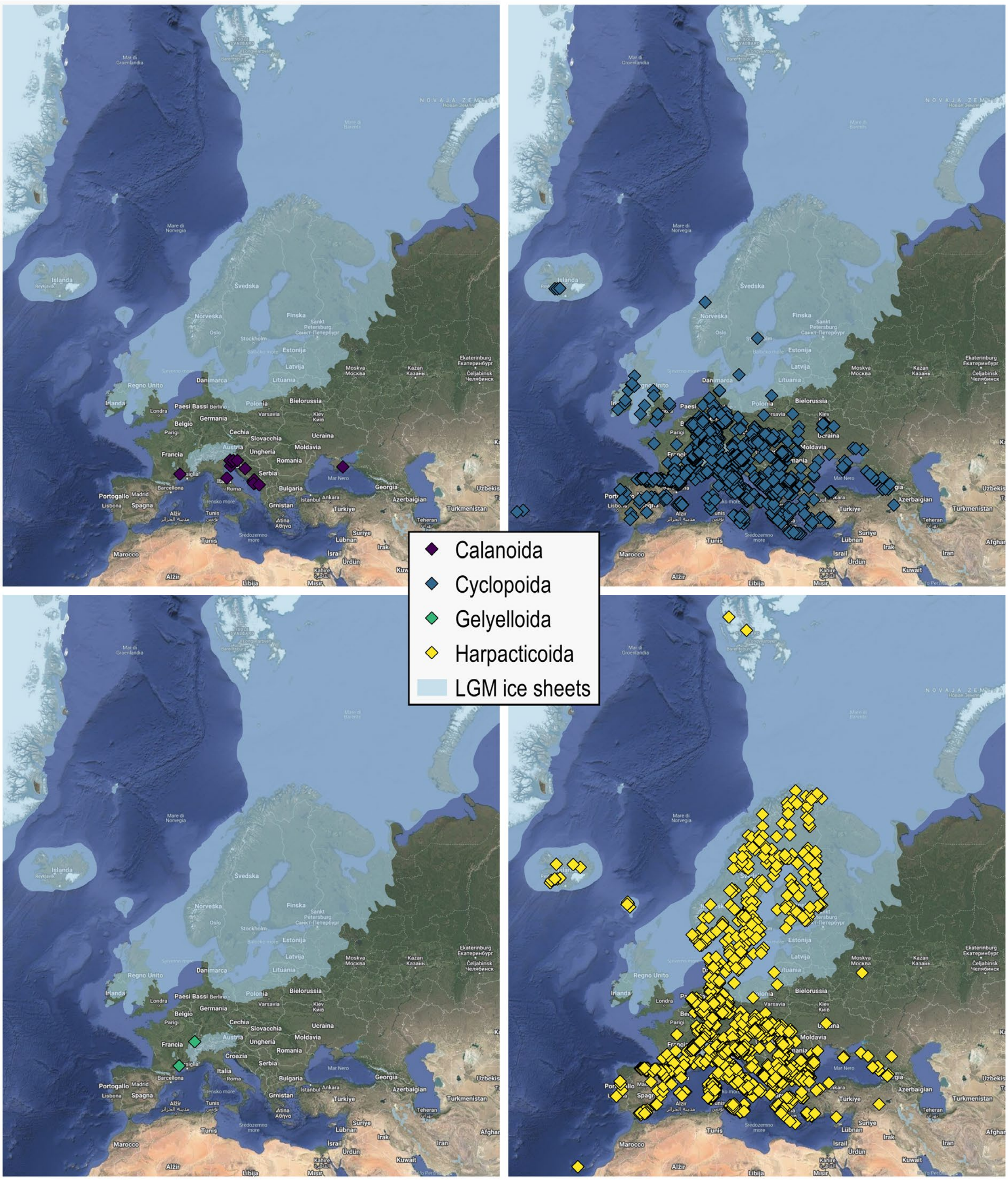


**FIGURE 1** | Spatial arrangement of the gathered occurrence records, grouped by copepod order. Records with large spatial uncertainty in coordinates (i.e., only available at the 'Region' or 'Catchment' scale) are not reported in the maps. Extent of ice sheets during the Last Glacial Maximum (LGM, ~21,000 y.a.), derived from Becker et al. (2015), is reported in the maps in light cyan; ice sheets covering some areas of southern Europe (e.g., French Central Massif, Italian Apennines and Pyrenees) during LGM are not clearly discernible at the scale of visualisation of the maps. The basemap is the 'Google Hybrid' layer available within the QGIS 'QuickMapServices' plugin.

during the last glacial maximum (Figure 1), aligning with previous evidence about the strong influence of Quaternary climatic oscillations on the distribution of subterranean species in Europe (Holdhaus 1932; Stoch and Galassi 2010; Růžička, Smilauer, and Mlejnek 2013; Mammola, Isaia, and Arnedo 2015; Mammola, Schönhofer, and Isaia 2019; Iannella et al. 2020a; Knüsel, Alther, and Altermatt 2024). Further, the high proportion of records in areas characterised by consolidated rocks with high permeability (26.4%) confirms the critical importance of karst regions for invertebrate groundwater biodiversity (Iannella et al. 2020a, 2020b; Borko et al. 2021). On the other hand, areas with a prevalence of unconsolidated sediments showed a notable richness of groundwater copepods as well, hosting 23.8% of the analysed records. However,

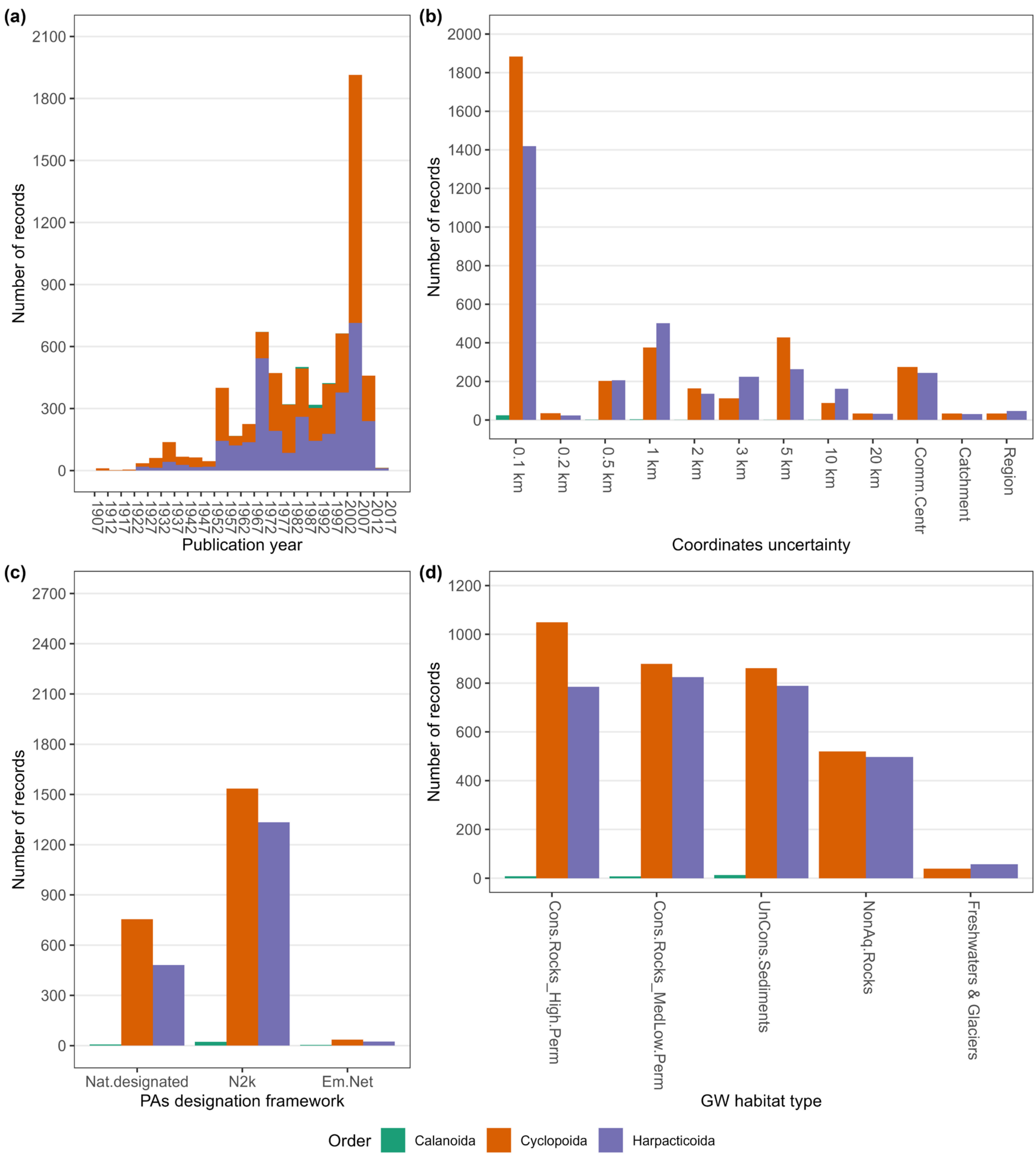


**FIGURE 2** | Distribution of occurrence records, grouped by order (Gelyelloida were excluded from spatial analyses due to the low number of occurrences), according to the following: (a) publication year of the reference book, research article, or technical report reporting the record; (b) spatial uncertainty in the provided projected geographic coordinates (EPSG: 3035); (c) category of surface protected area(s) each record falls within, (d) groundwater habitat type each record falls within. In (c) and (d), records with spatial uncertainty in coordinates higher than 5 km were not included. GW = groundwater; PA = protected area.

biogeographic patterns emerging from our data, such as the preponderance of harpacticoids in Fennoscandia, should be considered with caution. Indeed, while being the largest dataset about European groundwater copepods to date, we do not pretend EGCop to be exhaustive.

Notably, more than 4000 records fall within surface-designated protected areas (PAs), particularly those belonging to the Natura 2000 (N2k) network. Considering that N2k was established by the European Union's Habitats Directive in 1992 and implemented by Member States some years later, and that several records derive from research published in the early 2000s, one may ask whether sampling efforts targeting European groundwater fauna have been biased towards PAs. Answering this question, however, would require a deeper assessment of recent literature about groundwater fauna, analysing how sampling campaigns have been spatially distributed across Europe in the last three decades. Importantly, the presence of several records within PAs should not automatically suggest that groundwater fauna is exhaustively protected in Europe. Indeed, conservation measures adopted within surface-designated PAs do not usually consider the three-dimensional nature of subterranean ecosystems, thus not being necessarily effective in preserving the latter along their vertical dimension (Mammola et al. 2024). In this context, EGCop will be part of a broader effort—the 'DarCo' project—to collate high-quality data about European subterranean fauna, map the main direct (e.g., pollution, groundwater over-abstraction) and indirect (e.g., global warming, biological invasions) threats to subterranean biodiversity (Vaccarelli et al. 2023; Mammola et al. 2024), and propose evidence-based conservation measures (Mammola, Meierhofer, et al. 2022).

By publicly sharing this dataset we hope to stimulate further research on groundwater copepods, shedding light on their biogeography and ecology as well as including them in conservation efforts aiming to halt biodiversity loss.

### 4.1 | Code Availability Statement

The R code used to handle and analyse the presented data is provided as Supporting Information.

**Acknowledgements**

Francesco Cerasoli thanks Camilla Diamanti for her huge support in browsing the web to look for detailed information needed to estimate the spatial uncertainty for hundreds of occurrence records. Open access publishing facilitated by Universita degli Studi dell'Aquila, as part of the Wiley - CRUI-CARE agreement.

**Conflicts of Interest**

The authors declare no conflicts of interest.

**Data Availability Statement**

The presented dataset and the related metadata are accessible within the DRYAD repository at the following link https://doi.org/10.5061/dryad.9w0vt4bqj.

**References**

Backstrom, L. J., C. T. Callaghan, N. P. Leseberg, C. Sanderson, R. A. Fuller, and J. E. M. Watson. 2024. "Assessing Adequacy of Citizen Science Datasets for Biodiversity Monitoring." *Ecology and Evolution* 14, no. 2: e10857. https://doi.org/10.1002/ece3.10857.

Becker, D., J. Verheul, M. Zickel, and C. Willmes. 2015. "LGM Paleoenvironment of Europe - Map." *CRC806-Database*. https://doi.org/10.5880/SFB806.15.

Borko, Š., P. Trontelj, O. Seehausen, A. Moškrič, and C. Fišer. 2021. "A Subterranean Adaptive Radiation of Amphipods in Europe." *Nature Communications* 12, no. 1: 3688. https://doi.org/10.1038/s41467-021-24023-w.

Borutzky, V. E. 1952. *Fauna of U.S.S.R. Crustacea 3(4). Freshwater Harpacticoida*. Moscow-Leningrad: Academy of Sciences of the U.S.S.R.

Boxshall, G., and S. H. Halsey. 2004. *An Introduction to Copepod Diversity*. London, UK: The Ray Society.

Cerasoli, F., B. Fiasca, T. Di Lorenzo, et al. 2023. "Assessing Spatial and Temporal Changes in Diversity of Copepod Crustaceans: A Key Step for Biodiversity Conservation in Groundwater-Fed Springs." *Frontiers in Environmental Science* 11: 1051295. https://doi.org/10.3389/fenvs.2023.1051295.

Console, G., M. Iannella, F. Cerasoli, P. D'Alessandro, and M. Biondi. 2020. "A European Perspective of the Conservation Status of the Threatened Meadow Viper *Vipera ursinii* (Bonaparte, 1835) (Reptilia, Viperidae)." *Wildlife Biology* 2020, no. 2: 1–12. https://doi.org/10.2981/wlb.00604.

Cornu, J. F., D. Eme, and F. Malard. 2013. "The Distribution of Groundwater Habitats in Europe." *Hydrogeology Journal* 21, no. 5: 949–960. https://doi.org/10.1007/s10040-013-0984-1.

Ducatez, S., R. Tingley, and R. Shine. 2014. "Using Species Co-Occurrence Patterns to Quantify Relative Habitat Breadth in Terrestrial Vertebrates." *Ecosphere* 5, no. 12: 11–12. https://doi.org/10.1890/ES14-00332.1.

Dussart, B. 1967. *Les Copepodes des Eaux Continentales. Tome I: Calanoides et Harpacticoides*. Paris: N. Boubee & Cie.

Dussart, B. 1969. *Les Copépodes des Eaux Continentales d'Europe Occidentale. Tome II. Cyclopoïdes et Biologie*. Paris: N. Boubee & Cie.

Ficetola, G. F., C. Canedoli, and F. Stoch. 2019. "The Racovitzan Impediment and the Hidden Biodiversity of Unexplored Environments." *Conservation Biology* 33, no. 1: 214–216. https://doi.org/10.1111/cobi.13179.

Galassi, D. M. P., and P. De Laurentiis. 2004. "Towards a Revision of the Genus *Parastenocaris* Kessler, 1913: Establishment of *Simplicaris* Gen. Nov. From Groundwaters in Central Italy and Review of the *P. brevipes*-Group (Copepoda, Harpacticoida, Parastenocarididae)." *Zoological Journal of the Linnean Society* 140, no. 3: 417–436. https://doi.org/10.1111/j.1096-3642.2003.00107.x.

Galassi, D. M. P., B. Fiasca, T. Di Lorenzo, A. Montanari, S. Porfirio, and S. Fattorini. 2017. "Groundwater Biodiversity in a Chemoautotrophic Cave Ecosystem: How Geochemistry Regulates Microcrustacean Community Structure." *Aquatic Ecology* 51: 75–90. https://doi.org/10.1007/s10452-016-9599-7.

Galassi, D. M. P., F. Fiers, M.-J. Dole-Olivier, and B. Fiasca. 2019. "Discovery of a New Species of the Genus *Stygepactophanes* From a Groundwater-Fed Spring in Southern France (Crustacea, Copepoda, Harpacticoida, Canthocamptidae)." *ZooKeys* 812: 69–91. https://doi.org/10.3897/zookeys.812.29764.

Galassi, D. M. P., R. Huys, and J. W. Reid. 2009. "Diversity, Ecology and Evolution of Groundwater Copepods." *Freshwater Biology* 54, no. 4: 691–708.

Galassi, D. M. P., F. Stoch, and A. Brancelj. 2013. "Dissecting Copepod Diversity at Different Spatial Scales in Southern European Groundwater." *Journal of Natural History* 47, no. 5–12: 821–840. https://doi.org/10.1080/00222933.2012.738834.

Galmarini, E., I. Vaccarelli, B. Fiasca, et al. 2023. "Regional Climate Contributes More Than Geographic Distance to beta Diversity of Copepods (Crustacea Copepoda) Between Caves of Italy." *Scientific Reports* 13, no. 1: 21243. https://doi.org/10.1038/s41598-023-48440-7.

García-Girón, J., L. M. Bini, and J. Heino. 2023. "Shortfalls in Our Understanding of the Causes and Consequences of Functional and Phylogenetic Variation of Freshwater Communities Across Continents." *Biological Conservation* 282: 110082. https://doi.org/10.1016/j.biocon.2023.110082.

Gibert, J., and L. Deharveng. 2002. "Subterranean Ecosystems: A Truncated Functional Biodiversity." *Bioscience* 52, no. 6: 473–481. https://doi.org/10.1641/0006-3568(2002)052[0473:SEATFB]2.0.CO;2.

Griebler, C., and M. Avramov. 2015. "Groundwater Ecosystem Services: A Review." *Freshwater Science* 34, no. 1: 355–367. https://doi.org/10.1086/679903.

Hampton, S. E., C. A. Strasser, J. J. Tewksbury, et al. 2013. "Big Data and the Future of Ecology." *Frontiers in Ecology and the Environment* 11, no. 3: 156–162. https://doi.org/10.1890/120103.

Hobern, D., B. Baptiste, K. Copas, et al. 2019. "Connecting Data and Expertise: A New Alliance for Biodiversity Knowledge." *Biodiversity Data Journal* 7: e33679. https://doi.org/10.3897/BDJ.7.e33679.

Holdhaus, K. 1932. "Die europäische Höhlenfauna in ihren Beziehungen zur Eiszeit." *Zoogeographica* 1: 1–32.

Hortal, J., F. de Bello, J. A. F. Diniz-Filho, T. M. Lewinsohn, J. M. Lobo, and R. J. Ladle. 2015. "Seven Shortfalls That Beset Large-Scale Knowledge of Biodiversity." *Annual Review of Ecology, Evolution, and Systematics* 46, no. 1: 523–549. https://doi.org/10.1146/annurev-ecolsys-112414-054400.

Iannella, M., B. Fiasca, T. Di Lorenzo, M. Biondi, M. Di Cicco, and D. M. P. Galassi. 2020a. "Jumping Into the Grids: Mapping Biodiversity Hotspots in Groundwater Habitat Types Across Europe." *Ecography* 43, no. 12: 1825–1841. https://doi.org/10.1111/ecog.05323.

Iannella, M., B. Fiasca, T. Di Lorenzo, M. Biondi, M. Di Cicco, and D. M. Galassi. 2020b. "Spatial Distribution of Stygobitic Crustacean Harpacticoids at the Boundaries of Groundwater Habitat Types in Europe." *Scientific Reports* 10, no. 1: 19043. https://doi.org/10.1038/s41598-020-76018-0.

Iannella, M., B. Fiasca, T. Di Lorenzo, et al. 2021. "Getting the 'Most out of the Hotspot' for Practical Conservation of Groundwater Biodiversity." *Global Ecology and Conservation* 31: e01844. https://doi.org/10.1016/j.gecco.2021.e01844.

ICZN. 1999. *International Code of Zoological Nomenclature*. Fourth ed, 306. London, UK: International Trust for Zoological Nomenclature.

Iepure, S., C. A. Bădăluţă, and O. T. Moldovan. 2021. "An Annotated Checklist of Groundwater Cyclopoida and Harpacticoida (Crustacea, Copepoda) From Romania With Notes on Their Distribution and Ecology." *Subterranean Biology* 41: 87–108. https://doi.org/10.3897/subtbiol.41.72542.

Knüsel, M., R. Alther, and F. Altermatt. 2024. "Pronounced Changes of Subterranean Biodiversity Patterns Along a Late Pleistocene Glaciation Gradient." *Ecography* e07321. https://doi.org/10.1111/ecog.07321.

Lunghi, E., F. Cianferoni, F. Ceccolini, et al. 2018. "Field-Recorded Data on the Diet of Six Species of European Hydromantes Cave Salamanders." *Scientific Data* 5, no. 1: 180083. https://doi.org/10.1038/sdata.2018.83.

Macêdo, R. L., M. Toutain, J. Reid, et al. 2024. "Substantial Unrealised Global Biodiversity of Continental Microcrustaceans." *Journal of Plankton Research* 46, no. 3: 338–347. https://doi.org/10.1093/plankt/fbae020.

Mammola, S., F. Altermatt, R. Alther, et al. 2024. "Perspectives and Pitfalls in Preserving Subterranean Biodiversity Through Protected Areas." *NPJ Biodiversity* 3, no. 1: 2. https://doi.org/10.1038/s44185-023-00035-1.

Mammola, S., M. Isaia, and M. A. Arnedo. 2015. "Alpine Endemic Spiders Shed Light on the Origin and Evolution of Subterranean Species." *PeerJ* 3: e1384. https://doi.org/10.7717/peerj.1384.

Mammola, S., E. Lunghi, H. Bilandžija, et al. 2021. "Collecting Eco-Evolutionary Data in the Dark: Impediments to Subterranean Research and How to Overcome Them." *Ecology and Evolution* 11, no. 11: 5911–5926. https://doi.org/10.1002/ece3.7556.

Mammola, S., M. B. Meierhofer, P. A. Borges, et al. 2022. "Towards Evidence-Based Conservation of Subterranean Ecosystems." *Biological Reviews* 97, no. 4: 1476–1510. https://doi.org/10.1111/brv.12851.

Mammola, S., M. Pavlek, B. A. Huber, et al. 2022. "A Trait Database and Updated Checklist for European Subterranean Spiders." *Scientific Data* 9, no. 1: 236. https://doi.org/10.1038/s41597-022-01316-3.

Mammola, S., A. L. Schönhofer, and M. Isaia. 2019. "Tracking the Ice: Subterranean Harvestmen Distribution Matches Ancient Glacier Margins." *Journal of Zoological Systematics and Evolutionary Research* 57, no. 3: 548–554. https://doi.org/10.1111/jzs.12264.

Marmonier, P., D. M. P. Galassi, K. Korbel, M. Close, T. Datry, and C. Karwautz. 2023. "Chapter 5 - Groundwater Biodiversity and Constraints to Biological Distribution." In *Groundwater Ecology and Evolution*, edited by F. Malard, C. Griebler, and S. Rétaux, Second ed., 113–140. London, UK: Academic Press. https://doi.org/10.1016/B978-0-12-819119-4.00003-2.

Naimi, B., N. A. Hamm, T. A. Groen, A. K. Skidmore, and A. G. Toxopeus. 2014. "Where Is Positional Uncertainty a Problem for Species Distribution Modelling?" *Ecography* 37, no. 2: 191–203. https://doi.org/10.1111/j.1600-0587.2013.00205.x.

R Core Team. 2022. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

Rondinini, C., K. A. Wilson, L. Boitani, H. Grantham, and H. P. Possingham. 2006. "Tradeoffs of Different Types of Species Occurrence Data for Use in Systematic Conservation Planning." *Ecology Letters* 9, no. 10: 1136–1145. https://doi.org/10.1111/j.1461-0248.2006.00970.x.

Růžička, V., P. Smilauer, and R. Mlejnek. 2013. "Colonization of Subterranean Habitats by Spiders in Central Europe." *International Journal of Speleology* 42: 133–140. https://doi.org/10.5038/1827-806x.42.2.5.

Saccò, M., S. Mammola, F. Altermatt, et al. 2024. "Groundwater Is a Hidden Global Keystone Ecosystem." *Global Change Biology* 30, no. 1: e17066. https://doi.org/10.1111/gcb.17066.

Santi, F., R. Testolin, P. Zannini, et al. 2024. "MEDIS—A Comprehensive Spatial Database on Mediterranean Islands for Biogeographical and Evolutionary Research." *Global Ecology and Biogeography* 33, no. 8: e13855. https://doi.org/10.1111/geb.13855.

Soroye, P., B. P. M. Edwards, R. T. Buxton, et al. 2022. "The Risks and Rewards of Community Science for Threatened Species Monitoring." *Conservation Science and Practice* 4, no. 9: e12788. https://doi.org/10.1111/csp2.12788.

Stoch, F. 2006. "Crustacea Copepoda Cyclopoida." *Checklist and Distribution of the Italian Fauna* 10: 93–95.

Stoch, F., and D. M. P. Galassi. 2010. "Stygobiotic Crustacean Species Richness: A Question of Numbers, a Matter of Scale." In *Fifty Years After the "Homage to Santa Rosalia": Old and New Paradigms on Biodiversity in Aquatic Ecosystems*, edited by L. Naselli-Flores and G. Rossetti, 217–234. Netherlands: Springer. https://doi.org/10.1007/978-90-481-9908-2_16.

Vaccarelli, I., R. Colado, S. Pallares, et al. 2023. "A Global meta-Analysis Reveals Multilevel and Context-Dependent Effects of Climate Change on Subterranean Ecosystems." *One Earth* 6, no. 11: 1510–1522. https://doi.org/10.1016/j.oneear.2023.09.001.

Wells, J. 2007. "An Annotated Checklist and Keys to the Species of Copepoda Harpacticoida (Crustacea)." *Zootaxa* 1568, no. 1: 1–872. https://doi.org/10.11646/zootaxa.1568.1.1.

Zagmajster, M., D. C. Culver, M. C. Christman, and B. Sket. 2010. "Evaluating the Sampling bias in Pattern of Subterranean Species Richness: Combining Approaches." *Biodiversity and Conservation* 19: 3035–3048. https://doi.org/10.1007/s10531-010-9873-2.

Zagmajster, M., D. Eme, C. Fišer, et al. 2014. "Geographic Variation in Range Size and beta Diversity of Groundwater Crustaceans: Insights From Habitats With Low Thermal Seasonality." *Global Ecology and Biogeography* 23, no. 10: 1135–1145. https://doi.org/10.1111/geb.12200.

## Supporting Information

Additional supporting information can be found online in the Supporting Information section.