



OPEN

DATA DESCRIPTOR

The Danube Fish Database: documenting species distributions across a major European river basin

Yusdiel Torres-Cambas *et al.*[#]

The Danube River Basin (DRB) harbors the highest documented fish species richness of any European river, yet native populations face increasing threats from physical infrastructures that impede longitudinal and lateral connectivity, unsustainable fisheries, the introduction of non-native species, and climate change. Spanning across 19 countries, the DRB presents conservation challenges that demand coordinated, transboundary data sharing. The present database compiles and standardizes fish occurrence datasets that have been previously unavailable, fragmented and often restricted by federal agencies, research institutes, and conservation organizations, integrating also data from sources such as the Global Biodiversity Information Facility, the Joint Danube Surveys, the European Fish Index, and national monitoring programmes. It contains 133,131 occurrence records across 114 fish species, representing 30 families and 17 orders, with a temporal range from 1856 to 2024, organized into 39 columns. By supporting fish community conservation, invasive alien species monitoring, and climate impact assessments, this database provides a vital resource for developing evidence-based management strategies in the DRB.

Background & Summary

The Danube River Basin (DRB) hosts the highest reported fish species richness among European rivers, with more than 100 species, though the exact number varies depending on the source, ranging between 102 and 124 species. It supports around 20% of Europe's freshwater fish fauna, including nearly 30 endemic species—some restricted to single lagoons or rivers—while unfortunately also harboring a similarly high number of introduced alien species^{1–3}. However, extensive industrialization and river regulation since the 19th century have placed significant pressure on native fish communities in the DRB⁴. Anthropogenic impacts, including physical infrastructures that impede longitudinal and lateral connectivity, unsustainable fisheries, the introduction of non-native species, and climate change have collectively contributed to the decline of native fish populations^{5–10}. This trend reflects both regional and global patterns, as freshwater fish are the second most threatened animal group in Europe, surpassed only by molluscs, and are among the most endangered vertebrates worldwide^{11,12}. Within the DRB, this vulnerability is evident, with approximately 25 native fish species classified as globally threatened¹³.

Beyond ecological concerns, the DRB presents unique governance challenges. As the most international river basin in the world, spanning across 19 countries, it requires coordinated transnational efforts to manage freshwater biodiversity effectively¹⁴. Differences in administrative frameworks, monitoring standards, and data accessibility create obstacles for conservation planning and large-scale ecological assessments¹⁰. Consequently, there is a pressing need for harmonized data-sharing initiatives to support evidence-based conservation strategies.

To address the challenges of varying and scattered data sources in this heterogeneous river basin, this database¹⁵ is an effort to provide fish occurrence data in the DRB open access, complementing repositories, such as the Global Biodiversity Information Facility (GBIF, <https://www.gbif.org/>), the Joint Danube Surveys (JDS 1–4, <https://www.danubesurvey.org/jds4/>), and the European Fish Index¹⁶ (EFI+). The database¹⁵ was assembled through collaborations with federal agencies, universities, research institutes, NGOs, national park administrations, private sector stakeholders, and independent researchers. Data sources include research and monitoring projects as well as data collected for EU-mandated status assessments. While some of these datasets were previously available (e.g. for Austria^{17,18} and Hungary^{19–22}), they often remained difficult to access

[#]A full list of authors and their affiliations appears at the end of the paper.

Dataset name	Method	Contact persons
Sallai-Vital	Electrofishing and fyke nets. Additional records, particularly for <i>Hypophthalmichthys</i> spp., were obtained from rod-and-reel anglers through social media and subsequently validated by ZV.	Márton Sallai, Zoltán Vitál.
WWF_BG	Sampling was conducted in the Danube River using drifting trammel nets 100–150 m in length and 2 m in height, with a main net mesh size of 18–20 mm, along transects ranging from 800 to 2,500 m, mainly between May and September during both daytime and nighttime hours.	Stoyan Dobrev Mihov.
VUVH	Occurrence records were collected by electrofishing in the Morava, Danube, Little Danube, Váh, Hron, and Ipel rivers during surveys conducted between 2015 and 2019.	Maros Kubala.
UL_FGG	Occurrence records were collected in the Sava River between 2010 and 2023 using multiple sampling methods, including electrofishing, multimesh nets, traps, angling, and observational surveys.	Klaudija Lebar, Maša Čarf, Simon Rusjan.
UB-IMSI	Occurrence records were collected between 1965 and 2024 from the Danube, Tisza, Sava, Velika Morava, Tamiš, Timok, Mlava, Porečka, Pek, Karaš, Nera, Drina, Studva, and other associated rivers and channels using a range of approaches, including electrofishing, stagnant and drift net fishing, and fisheries data; in projects We Pass 1 and 2, fish were captured by net fishing and tagged for telemetry.	Dušan Nikolić, Gorčin Cvijanović, Marija Smederevac-Lalić, Milica Jaćimović.
RBI-DANUBEPARKS	Occurrence records were collected between 2003 and 2024 from the Sava River and its backwaters, multiple habitats within Kopački Rit Nature Park, and several Croatian rivers and water bodies (Drava, Kupa, Kupčina, Orljava, Sunja, Una, Lomnica, Odra, Lipnica, Mrtva Odra Meander, and Lonja) using electrofishing, combined electrofishing and net sampling, and recreational sampling.	Damir Valić, Sandi Orlic, Matej Marusic, Marina Piria.
INCDDD	Occurrence records were collected in 2018 and 2019 from the Danube River, including the Saint George, Chilia, and Tulcea branches, Sulina branch and the Lakes in the Delta, as well as the Somova–Parcheș Complex, using net fishing.	Irina Cernisencu, Iulian Nichersu.

Table 1. Overview of individual datasets present in the database¹⁵ that contain primary, unpublished data, including the methods used to collect occurrence records and the corresponding contact persons. Dataset names correspond to the values in the datasetName column of the database.

due to restricted user permissions or language barriers. By consolidating and standardizing these records, this collection enhances the availability and usability of fish biodiversity data across the DRB, fostering informed decision-making for freshwater ecosystem conservation.

The database¹⁵ provides essential data to address conservation challenges for fish in Europe and the DRB. It can be used to identify suitable habitats for native species in projects aiming at prioritizing habitat reconnection, taking into account estimated dam removal costs and fish pass construction costs, while also minimizing the spread of invasive species²³. Additionally, it supports monitoring species responses to climate change, aiding in the prediction of vulnerability and the development of adaptive management strategies²⁴. The database¹⁵ is also valuable for tracking non-native species, enabling early detection of range expansions and providing crucial data for risk assessments and control measures to mitigate their spread^{25–28}. Furthermore, fish occurrence data can be integrated with spatially explicit threat data, such as land use intensity, pollution, connectivity barriers, invasive species and climate change^{29–32} to prioritize management actions and contribute to the implementation of the European Union Nature Restoration Law aimed at restoring free-flowing rivers¹⁰. This approach will help determine which actions to implement and where to implement them to effectively achieve biodiversity recovery targets³³.

Methods

We compiled the occurrence data for fish species within the DRB from a variety of sources. Contributions included data from Austria^{17,18}, Bulgaria, Croatia, Germany^{34–38}, Hungary (¹⁹ BioAqua Pro Ltd., Duna-Ipoly National Park Directorate²², Duna-Drava National Park Directorate, Fertő-Hanság National Park Directorate^{20,21}, WWF Hungary^{20–22}), Romania, Serbia, Slovakia, and Slovenia (given the partners in the DANUBE4all project, <https://www.danube4allproject.eu/>), as well as from GBIF³⁹, the Fourth Joint Danube Survey (JDS4), and the EFI+ database. Further information on datasets containing primary, unpublished data, including sampling methods and contact persons, is provided in Table 1.

In total, 506,290 entries were collected and subsequently subjected to quality checks and cleaning procedures. Following the data collation, formatting, and quality control, we wrapped all the R-code we employed into a new `danubeoccurR` R-package⁴⁰, to facilitate and streamline the entire process for reproducibility and possible re-use in other regions and datasets. We also incorporated additional R-packages, `hydrographR`⁴¹ and `specleanr`⁴² into the workflow to aid in data manipulation and geospatial analysis (Fig. 1).

For the retrieval of fish species occurrences from GBIF within the DRB, we fetched records based on a preliminary, unpublished checklist compiled by national fish experts from the Fifth Joint Danube Survey (JDS5), which is available as a dataset named `species_checklist` in the `danubeoccurR` R-package⁴⁰. This procedure can be applied using the function `download_gbif_records()` from the `danubeoccurR` package. We then performed several cleaning and standardization steps to ensure the data’s suitability for modelling. For GBIF data, we removed duplicates and filtered records based on coordinate precision (i.e. number of decimal places in the latitude and longitude values), uncertainty (i.e. radius in meters around the given coordinates within which the actual location is expected to be found), and geographic buffers by applying the `clean_gbif()` function. This process excluded records with low precision (e.g. less than 3 decimal places) or high uncertainty (above 1000 m) and removed those located near country centroids (within 1000 m), capital city centroids (within 1000 m), and zoo/herbarium centroids (within 1000 m). For other datasets, coordinates were standardized from degrees, minutes, and seconds (DMS) to decimal degrees using the `dms_to_decimal()` function, while date information was also standardized using the `extract_date()` function to ensure uniformity across the dataset.

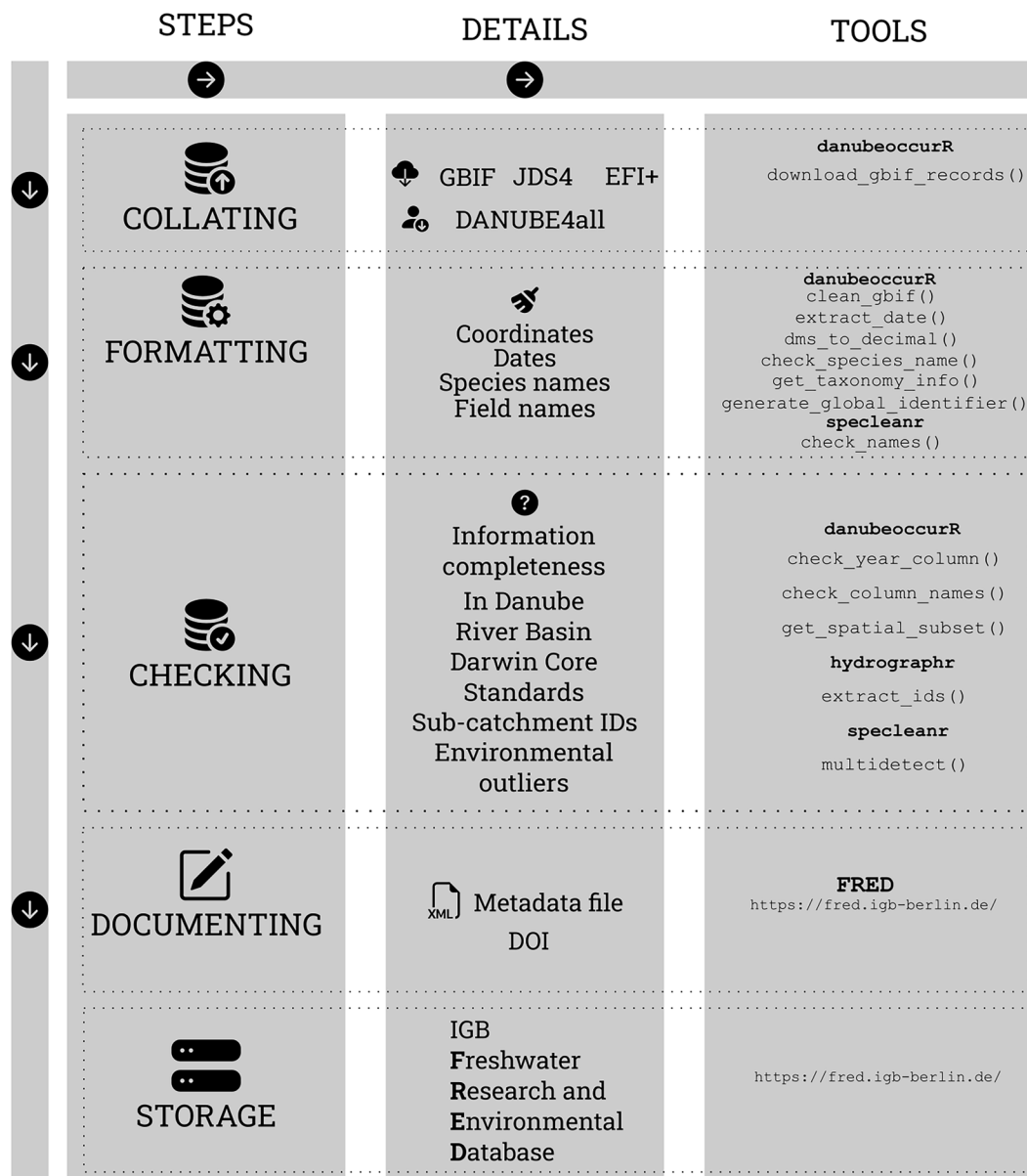


Fig. 1 Workflow to compile a database of fish species occurrences from the DRB. The process involves gathering information from various sources, standardizing it to a common format, performing quality checks, and documenting the data using dedicated R packages. The finalized data is stored on an open-source platform for broader accessibility and future use.

A critical aspect of our data preparation was ensuring taxonomic accuracy across the various data sources. This was achieved by using the `check_names()` function from `specleanr`⁴². The function identifies synonyms based on FishBase⁴³, corrects spelling errors, and suggests names based on a 90% similarity threshold to ensure the most accurate matches. If automated correction was not feasible, as it was the case for 82 occurrences, we performed manual updates using the `check_species_name()` function in `danubeoccurR`⁴⁰. This function provides an interactive interface in RStudio⁴⁴ for users to manually edit species records. Finally, the resulting FishBase species list was compared against the species list in the FreshVets v1.0 database⁴⁵, which is based on the Eschmeyer Catalog of Fishes (CoF)⁴⁶. FishBase species that were not present in FreshVets v1.0 (seven species) were double-checked and updated directly against the Eschmeyer Catalog of Fishes using the most up-to-date version of the CoF⁴⁶.

In addition to taxonomic verification, we applied data enrichment steps to complete taxonomic information for each occurrence record. Taxonomic information such as genus, family, and order was populated using the `get_taxonomy_info()` function. Persistent identifiers (UUIDs) were also assigned to each record using the `generate_global_identifier()` function to align with Darwin Core and other standards. Additionally, to enable further spatial analysis, sub-catchment IDs were extracted using the `extract_ids()` function from `hydrographr`⁴¹, using the sub-catchment layers from the Hydrography90m global dataset⁴⁷.

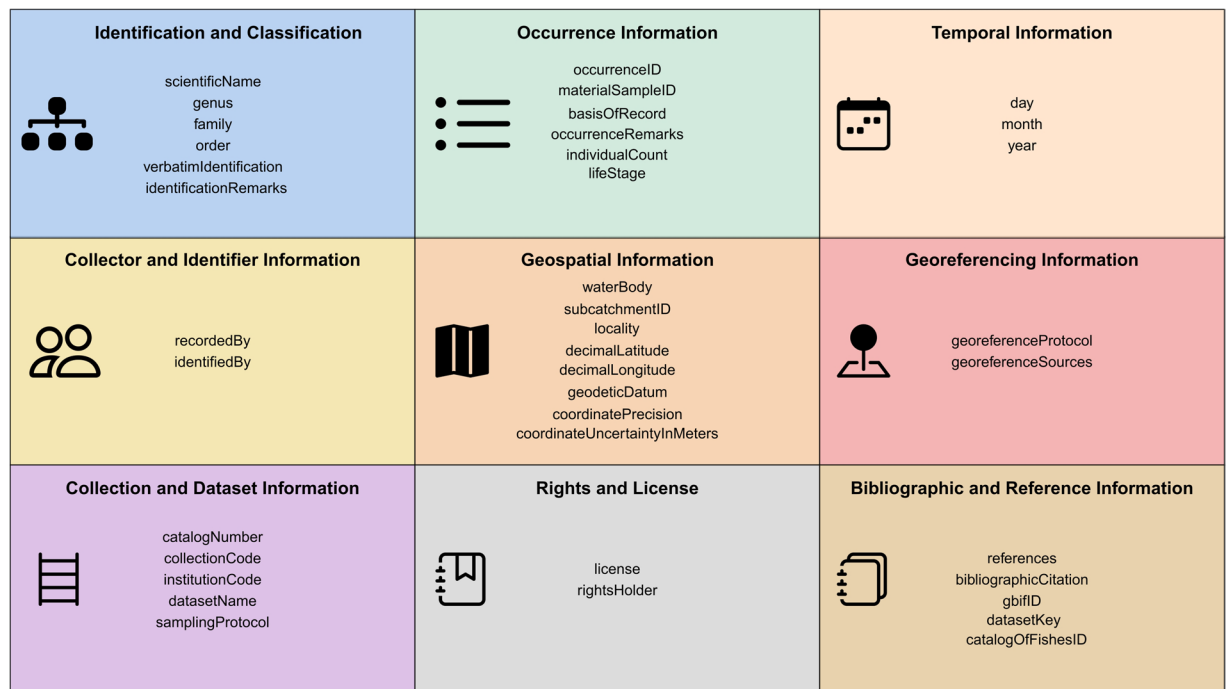


Fig. 2 Structure of a database of fish species occurrences from the Danube River Basin. The database fields are organized into nine categories, each represented by a box in the figure. Field names follow Darwin Core Standards.

Following these procedures, we performed additional data validation steps. We verified column names using the `check_column_name()` function to ensure conformity with Darwin Core terms. The “year” column was checked for missing values and validated to be numeric and within an acceptable range using the `check_year_column()` function. Coordinate validation was performed using the `check_coordinates()` function to confirm that latitude and longitude values were in the correct decimal degree format (WGS84). To ensure that only records within the DRB were retained, we applied the `get_spatial_subset()` function, which performs a spatial filtering operation to keep only those points falling within a specified polygon. In this case, the polygon represents the DRB and was obtained from the Hydrography90m global dataset⁴⁷. Additionally, potential environmental outliers in the occurrence data were detected and flagged using the `multidetector()` function from `specleanr`^{42,48}. Identifying environmental outliers is a critical step in preparing data for species distribution models (SDMs), as these outliers can skew model predictions by introducing patterns that do not represent broader ecological trends⁴⁸.

Data Records

Repository and Access. The Danube Fish Database¹⁵ is openly available in the Freshwater Research and Environmental Database (FRED), which is the central data repository of the Leibniz Institute of Freshwater Ecology and Inland Fisheries (IGB), at <https://doi.org/10.18728/igb-fred-1078.7> (Fig. 1). The database¹⁵ is structured to facilitate accessibility and reproducibility, with records provided in standardized formats to ensure compatibility with biodiversity and ecological data frameworks.

Structure and File Formats. The database¹⁵ consists of two files, each corresponding to a distinct aspect of fish occurrence records across the DRB. These files are available at <https://doi.org/10.18728/igb-fred-1078.7> and are organized as follows:

- `occurrence_records.csv` - Primary occurrence data. The database of fish species occurrences from the DRB contains 39 fields grouped into nine categories presented in Fig. 2.
- `metadata.pdf` - A metadata file describing column definitions and data sources.

Data Overview

The database¹⁵ contains 133,131 occurrence records, representing 114 bony fish species across 30 families and 17 orders. Temporal coverage ranges from 1856 to 2024. The median number of records per species is 371, with the most recorded species, the bleak (*Alburnus alburnus*), having 13,582 occurrences. In contrast, the least recorded species *Acipenser naccarii*, *Ictalurus punctatus*, *Ponticola syrman* and *Salmo marmoratus*, each have only one occurrence. Eighty-two percent of the species have ten or more occurrence records. A summary checklist of species and their occurrence counts is provided in Supplementary Table 1. The spatial distribution of records is

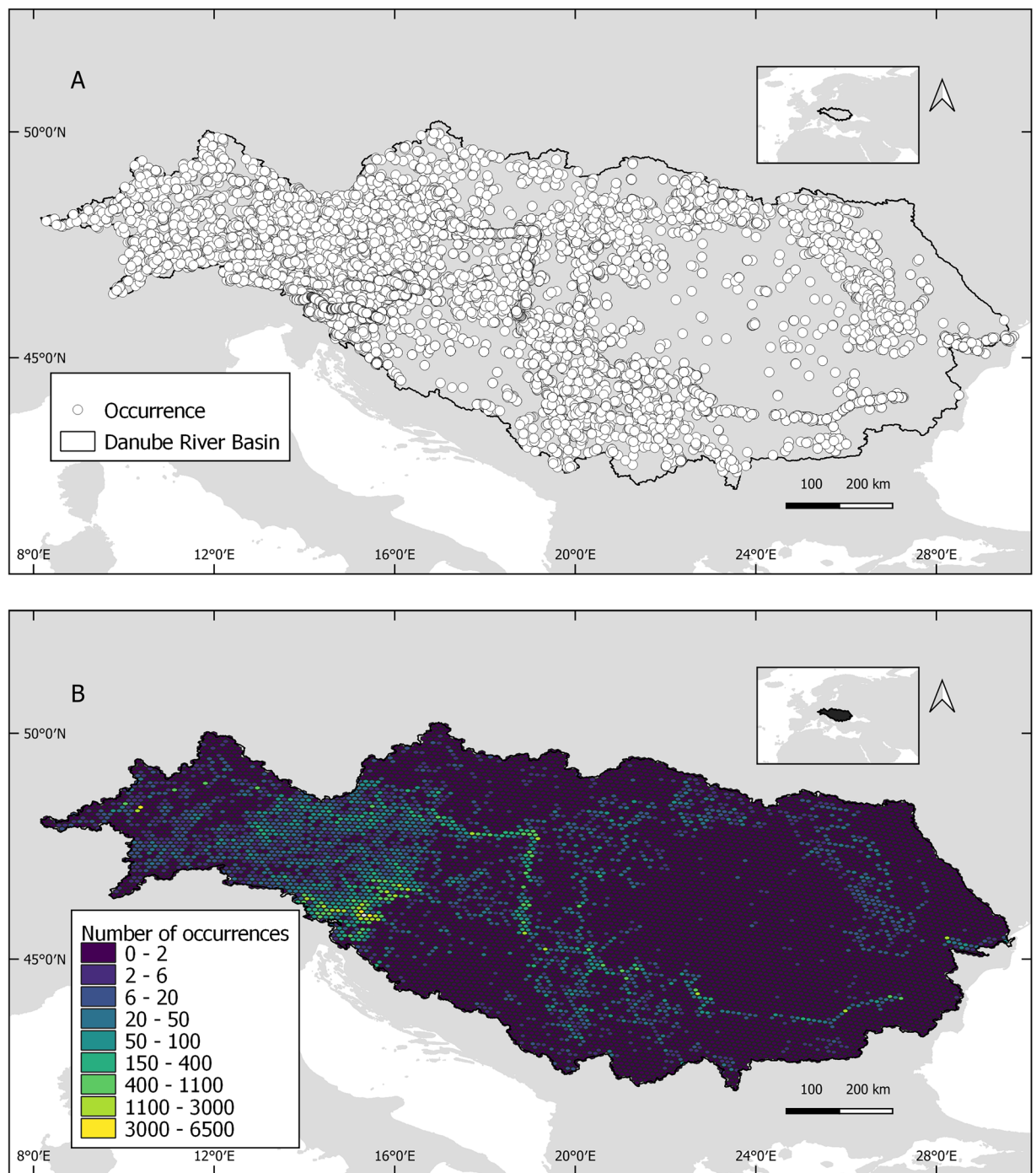


Fig. 3 Fish occurrence records across the Danube River Basin are shown for 114 species. Figure 3A illustrates individual occurrence points, while Figure 3B shows the number of occurrences per 10 km² grid cell. The database was compiled from online databases and contributions from DANUBE4all project partners in Austria, Bulgaria, Croatia, Germany, Hungary, Romania, Serbia, Slovakia, and Slovenia, with additional data sourced from GBIF, the Fourth Joint Danube Survey (JDS4), and the EFI+ database. Basin boundaries follow the hydrography90m dataset⁴⁷.

shown in Fig. 3, while a higher-resolution visualization is available on the IGB-GeoNode (https://geo.igb-berlin.de/layers/geonode:danube4all_fish_occurrence_records), the geospatial information platform of the Leibniz Institute of Freshwater Ecology and Inland Fisheries.

Technical Validation

To ensure the technical quality of the fish species occurrence database for the DRB, several validation steps were undertaken:

- **Taxonomic Validation:** FishBase⁴³ and the CoF⁴⁶ were consulted to ensure alignment with the most up-to-date fish taxonomy, following the semi-automated procedures described in the Methods section. Two species, *Salvelinus umbla* and *Vimba elongata*, are treated as valid taxa in FishBase, but as synonyms of *S. alpinus* and *V. vimba*, respectively, in the CoF. We adopted the CoF treatment as the current valid taxonomic status. For records originally classified as *Salvelinus umbla* and *Vimba elongata*, a note was added to the identificationRemarks and verbatimIdentification fields in the database¹⁵.
- **Spatial Distribution Validation:** The spatial distribution of species occurrences was assessed to ensure that the recorded locations were geographically plausible and consistent with known habitats for each species. To achieve this, the occurrences were first compared against environmental maps to identify and flag potential environmental outliers (see Methods), and subsequently cross-checked by the respective data provider.
- **Data Completeness and Consistency:** The database was examined for missing values, duplicates, and inconsistencies in key fields such as coordinates, dates, and species names. Gaps in data were identified, and missing or inconsistent records were flagged for review and potential correction.

Usage Notes

To facilitate data reuse, we provide key recommendations for accessing, processing, and interpreting the database.

- **Data Access and Format:** The database is available in a standard tabular format (CSV) using Darwin Core-compliant terminology to ensure compatibility with biodiversity databases. Users should refer to the metadata file for a detailed description of the column names. For convenience, a custom function named `split_and_save_csv()` is provided in `danubeooccurR`⁴⁰ to split the occurrence database into independent datasets.
- **Geospatial Considerations:** Users should be aware that some records may have variable spatial precision, particularly historical occurrences. The function `snap_points_on_map()` in `danubeooccurR`⁴⁰ allows users to manually adjust occurrence points for greater precision. Additionally, for applications that require spatial alignment between occurrences and a stream network, we recommend using the snapping functionalities available in the online platform GeoFRESH⁴⁹ for small datasets (e.g. fewer than 100 occurrences), or functions in the R-package `hydrographr`⁴¹ for custom workflows and larger datasets.
- **Taxonomic Standardization:** Users are advised to cross-check species names against current taxonomic databases because taxonomic revisions can occur after the database's publication. To facilitate future taxonomic updates, we also included a column (`catalogOfFishesID`) containing the Eschmeyer Catalog of Fishes species identification number (`spid`) for each taxon⁴⁶. Because `spid` is the identifier used by the Catalog to retrieve species records directly, it provides a stable reference for following taxonomic changes and updating species names in subsequent versions of the database.
- **Data Quality and Potential Limitations:** While efforts were made to standardize and clean the data, users should consider potential sources of bias, including sampling effort variations, taxonomic misidentifications, or incomplete historical records. Some records have been flagged as environmental outliers based on inconsistencies between species occurrence and expected environmental conditions. These flagged records should be reviewed carefully and may require further investigation or validation before inclusion in analyses.

Spatial bias, resulting from sampling or observation bias, can negatively impact certain applications of occurrence data, such as SDMs⁵⁰. To mitigate sampling bias, spatial thinning can be applied. This technique removes the fewest records necessary to significantly reduce bias while retaining the maximum amount of useful information⁵¹. Alternatively, in SDMs, it can also be addressed by incorporating covariates that account for accessibility and sampling effort during model calibration⁵².

Records of *Acipenser naccarii*, should be interpreted with caution. For instance, two specimens of *Acipenser naccarii* recorded in 1995 were sourced from GBIF, originally from the Naturkunde Museum Stuttgart (<https://id.smns-bw.org/smns/collection/html/775282/1237352/779394>), found in Teichanlage Wöllershof, a pond complex located in Wöllershof, in the Oberpfalz (Upper Palatinate) region of Bavaria, Germany. Since this area is known for its network of fish ponds, it is likely that the record comes from aquaculture.

- **Further species information:** The Danube Fish Database represents a valuable basis for various research fields. Users can flexibly join the dataset to additional information on the biology, autoecology, traits, origin, and conservation status of fish species or communities. A noteworthy example is the freshwater ecology.org database⁵³ from where information on autecological characteristics, traits, and geographic origin can be extracted via the `fwtraits` R-package⁵⁴. Species-level information, including morphological or physiological characteristics can be obtained from FishBase and conservation status from the IUCN Red List via APIs (e.g. (<https://api.iucnredlist.org/>)). Genetic data can be accessed via databases such as BOLD (<https://www.boldsystems.org/>) or NCBI (<https://www.ncbi.nlm.nih.gov/>).

Data availability

The Danube Fish Database is openly available in the Freshwater Research and Environmental Database (FRED), which is the central data repository of the Leibniz Institute of Freshwater Ecology and Inland Fisheries (IGB), at <https://doi.org/10.18728/igb-fred-1078.7>.

Code availability

All functions and example procedures for data sourcing, cleaning, formatting, and quality control of fish species occurrences in the DRB are available through the R package `danubeoccurR`⁴⁰, which can be installed from GitHub as follows:

```
# Install remotes package if not already installed
install.packages("remotes")

# Install danubeoccurR from GitHub
remotes::install_github("ytorres-cambas/danubeoccurR")
```

The package includes vignettes that demonstrate key procedures, including:

- **Data Sourcing and Cleaning:** Retrieve occurrence records from GBIF and perform data cleaning. See the vignette General workflow to collate, format, data enrichment and check, sections 2–4 (<https://ytorres-cambas.github.io/danubeoccurR/articles/download-and-clean-gbif.html>).
- **Taxonomic Standardization:** Verify species names and supplement missing taxonomic information such as genus, family, and order. See the vignette General workflow to collate, format, data enrichment and check, sections 6 and 7 (<https://ytorres-cambas.github.io/danubeoccurR/articles/download-and-clean-gbif.html>).
- **Adding Additional Information:** Incorporate geospatial data such as administrative names, subcatchment identification numbers and other information such as persistent identifiers for occurrences. See the vignette General workflow to collate, format, data enrichment and check, section 8–10 (<https://ytorres-cambas.github.io/danubeoccurR/articles/download-and-clean-gbif.html>).
- **Quality Control:** Ensure data completeness and proper formatting, verifying that each record includes at least a species name, date, and coordinates. Also, check column names and coordinate and date format. See the vignette General workflow to collate, format, data enrichment and check (<https://ytorres-cambas.github.io/danubeoccurR/articles/download-and-clean-gbif.html>).
- **Snapping Points:** Manually adjust occurrence points for improved spatial accuracy. See the vignette Manually snap species occurrence points (<https://ytorres-cambas.github.io/danubeoccurR/articles/manually-snap-points.html>).

The package is implemented in R (version 4.4.3) and relies on dependencies such as `sfc`^{55,56}, `tidyverse`⁵⁷, `rgbif`⁵⁸, `spleanr`⁴² and `CoordinateCleaner`⁵⁹, which are automatically installed during package installation. Users can access documentation, report issues, or contribute to development through the GitHub repository <https://github.com/ytorres-cambas/danubeoccurR>.

Additionally, an example of the procedure used for detecting environmental outliers is available on the `spleanr`⁴² package's website. For more details, see the vignette Environmental outlier detection in fish species records (<https://anthonybasooma.github.io/spleanr/articles/eOutlier.html>).

Received: 12 November 2025; Accepted: 13 April 2026;

Published online: 27 May 2026

References

1. Kováč, V. Current Status of Fish Communities in the Danube. In Liska, I. (ed.) *The Danube River Basin*, vol. 39 of *The Handbook of Environmental Chemistry*, https://doi.org/10.1007/698_2015_377 (Springer, Berlin, Heidelberg, 2015).
2. Kottelat, M. & Freyhof, J. *Handbook of European Freshwater Fishes* (Publications Kottelat, Cornol, Switzerland, 2007).
3. Tedesco, P. A. *et al.* A global database on freshwater fish species occurrence in drainage basins. *Sci. Data* **4**, 1–6, <https://doi.org/10.1038/sdata.2017.141> (2017).
4. Keckeis, H. & Schiemer, F. Understanding conservation issues of the Danube River. In Fuiman, L. A. & Werner, L. A. (eds.) *Fishery Science. The unique contributions of early life stages*, 272–288 (Blackwell Publishing, 2002).
5. Friedrichs-Manthey, M. *et al.* Three hundred years of past and future changes for native fish species in the upper Danube River Basin—Historical flow alterations versus future climate change. *Diversity and Distributions* **30**, e13808, <https://doi.org/10.1111/ddi.13808> (2024).
6. Lenhardt, M. *et al.* Human Impacts on Fish Fauna in the Danube River in Serbia: Current Status and Ecological Implications. In Bănăduc, D., Curtean-Bănăduc, A., Pedrotti, F., Cianfaglione, K. & Akeroyd, J. R. (eds.) *Human Impact on Danube Watershed Biodiversity in the XXI Century*, 257–279, https://doi.org/10.1007/978-3-030-37242-2_13 (Springer International Publishing, 2020).
7. Bănăduc, D., Rey, S., Trichkova, T., Lenhardt, M. & Curtean-Bănăduc, A. The Lower Danube River-Danube Delta-North West Black Sea: A pivotal area of major interest for the past, present and future of its fish fauna — A short review. *Science of The Total Environment* **545–546**, 137–151, <https://doi.org/10.1016/j.scitotenv.2015.12.058> (2016).
8. Mueller, M., Bierschenk, A. M., Bierschenk, B. M., Pander, J. & Geist, J. Effects of multiple stressors on the distribution of fish communities in 203 headwater streams of Rhine, Elbe and Danube. *Science of The Total Environment* **703**, 134523, <https://doi.org/10.1016/j.scitotenv.2019.134523> (2020).
9. Bierschenk, A. M., Mueller, M., Pander, J. & Geist, J. Impact of catchment land use on fish community composition in the headwater areas of Elbe, Danube and Main. *Science of The Total Environment* **652**, 66–74, <https://doi.org/10.1016/j.scitotenv.2018.10.218> (2019).
10. Stoffers, T. *et al.* Reviving europe's rivers: Seven challenges in the implementation of the nature restoration law to restore free-flowing rivers. *WIREs Water* **11**, e1717, <https://doi.org/10.1002/wat2.1717> (2024).
11. Costa, M. J., Duarte, G., Segurado, P. & Branco, P. Major threats to European freshwater fish species. *Science of The Total Environment* **797**, 149105, <https://doi.org/10.1016/j.scitotenv.2021.149105> (2021).
12. Darwall, W. R. T. & Freyhof, J. Lost fishes, who is counting? The extent of the threat to freshwater fish biodiversity. In Closs, G. P., Krkosek, M. & Olden, J. D. (eds.) *Conservation of Freshwater Fishes*, Conservation Biology, 1–36 (Cambridge University Press, 2015).
13. Sommerwerk, N. *et al.* The danube river basin. In Tockner, K., Zarfl, C. & Robinson, C. T. (eds.) *Rivers of Europe*, 59–112, <https://doi.org/10.1016/B978-0-12-369449-2.00003-5> (second edition edn, Elsevier, 2009).
14. Habersack, H. *et al.* Challenges of river basin management: Current status of, and prospects for, the River Danube from a river engineering perspective. *Science of The Total Environment* **543**, 828–845, <https://doi.org/10.1016/j.scitotenv.2015.10.123> (2016).
15. Torres-Cambas, Y. *et al.* Danube fish occurrence database. Dataset, <https://doi.org/10.18728/igb-fred-1078.7> (2026).

16. Schinegger, R., Pletterbauer, F., Melcher, A. & Schmutz, S. Metadata describing the european fish index plus (efi+) database. *Freshwater Metadata Journal* **17**, 1–12, <https://doi.org/10.15504/fmj.2016.17> (2016).
17. GZÜV. Gewässerzustandsüberwachungsverordnung (GZÜV). Verordnung über die Überwachung des Zustands der Gewässer. BGBl. II Nr. 479/2006, i.d.g.F. (2006).
18. Bundesministerium für Land- und Forstwirtschaft, Regionen und Wasserwirtschaft (BMLRW) & Umweltbundesamt GmbH. WISA Datenabfrage. <https://wasser.umweltbundesamt.at/wisa-datenabfrage/#/> (2025).
19. Czeglédi, I. Database of fish occurrence records from the Danube River Basin (Hungary). Dataset, <https://doi.org/10.5158/ARP/2DHQF1> ARP, Version V1 (2025).
20. Takács, P. Halközösségek felmérése az nbmr országos programja keretében (kerka, lajta, mura, répcé, pinka, mosoni-duna, zala). Technical Report, Commissioned by Fertő-Hanság National Park Directorate (2017).
21. Sallai, Z. Közösségi jelentőségű halfajok felmérése a Fertő-Hanság Nemzeti Park Igazgatóság területén. Technical Report, Nimfea Természetvédelmi Egyesület, Szarvas (2008).
22. Duna-Ipoly National Park Directorate (DINPI). Species occurrence database of the Duna-Ipoly National Park Directorate (DINPI). <https://doi.org/10.18426/obm.23mj4em6bca0> (2017).
23. Cooper, A. R. *et al.* Prioritizing native migratory fish passage restoration while limiting the spread of invasive species: A case study in the Upper Mississippi River. *Science of The Total Environment* **791**, 148317, <https://doi.org/10.1016/j.scitotenv.2021.148317> (2021).
24. Radinger, J. *et al.* The future distribution of river fish: The complex interplay of climate and land use changes, species dispersal and movement barriers. *Global Change Biology* **23**, 4970–4986, <https://doi.org/10.1111/GCB.13760> (2017).
25. Bailey, S. A. *et al.* Trends in the detection of aquatic non-indigenous species across global marine, estuarine and freshwater ecosystems: A 50-year perspective. *Diversity and Distributions* **26**, 1780–1797, <https://doi.org/10.1111/DDI.13167> (2020).
26. Jarić, I., Cvijanović, G., Hegediš, A. & Lenhardt, M. Assessing the range of newly established invasive species in rivers using probabilistic methods. *Hydrobiologia* **680**, 171–178, <https://doi.org/10.1007/S10750-011-0914-Y/FIGURES/3> (2012).
27. Panov, V. E. *et al.* Assessing the risks of aquatic species invasions via European inland waterways: from concepts to environmental indicators. *Integrated Environmental Assessment and Management* **5**, 110–126, https://doi.org/10.1897/IEAM_2008-034.1 (2009).
28. Haubrock, P. J., Kurtul, I. & Kouba, A. Tracking aquatic non-native macroinvertebrate species in Germany using long-term data. *Environmental Sciences Europe* **36**, 1–12, <https://doi.org/10.1186/S12302-024-00986-7/FIGURES/5> (2024).
29. Belletti, B. *et al.* More than one million barriers fragment Europe's rivers. *Nature* **588**, 436–441, <https://doi.org/10.1038/s41586-020-3005-2> (2020).
30. Karger, D. N. *et al.* CHELSA-W5E5: daily 1 km meteorological forcing data for climate impact studies. *Earth System Science Data* **15**, 2445–2464, <https://doi.org/10.5194/ESSD-15-2445-2023> (2023).
31. Brun, P., Zimmermann, N. E., Hari, C., Pellissier, L. & Karger, D. N. Global climate-related predictors at kilometer resolution for the past and future. *Earth System Science Data* **14**, 5573–5603, <https://doi.org/10.5194/ESSD-14-5573-2022> (2022).
32. International Commission for the Protection of the Danube River (ICPDR). DanubeGIS: The Danube River Basin Geographic Information System. Accessed: 2025-03-13 (2025).
33. Salgado-Rojas, J., Hermoso, V. & Álvarez Miranda, E. prioriactions: Multi-action management planning in R. *Methods in Ecology and Evolution* <https://doi.org/10.1111/2041-210X.14220> (2023).
34. Pander, J., Nagel, C., Ingermann, H. & Geist, J. Water level induced changes of habitat quality determine fish community composition in restored and modified riverbanks of a large alpine river. *International Review of Hydrobiology* **107**, 46–59, <https://doi.org/10.1002/iroh.202002079> (2022).
35. Pander, J. & Geist, J. Seasonal and spatial bank habitat use by fish in highly altered rivers – a comparison of four different restoration measures. *Ecology of Freshwater Fish* **19**, 127–138, <https://doi.org/10.1111/j.1600-0633.2009.00397.x> (2010).
36. Pander, J. & Geist, J. Can fish habitat restoration for rheophilic species in highly modified rivers be sustainable in the long run? *Ecological Engineering* **88**, 28–38, <https://doi.org/10.1016/j.ecoleng.2015.12.006> (2016).
37. Pander, J., Nagel, C. & Geist, J. Integration of constructed floodplain ponds into nature-like fish passes supports fish diversity in a heavily modified water body. *Water* **13**, 1018, <https://doi.org/10.3390/w13081018> (2021).
38. Pander, J. & Geist, J. The contribution of different restored habitats to fish diversity and population development in a highly modified river: A case study from the river ginz. *Water* **10**, 1202, <https://doi.org/10.3390/w10091202> (2018).
39. GBIF.Org User. GBIF Occurrence Download, <https://doi.org/10.15468/dl.yug2ad> (2024).
40. Torres-Cambas, Y. *danubeoccurR: Tools for downloading, cleaning, and managing fish occurrence records for the Danube River Basin*. R package version 0.1.0 (2025).
41. Schürz, M. *et al.* hydrographr: An R package for scalable hydrographic data processing. *Methods in Ecology and Evolution* **00**, 1–11, <https://doi.org/10.1111/2041-210X.14226> (2023).
42. Basooma, A. *specleanr: Detecting environmental outliers in biogeographical modeling workflows* R package version 0.1.0 (2025).
43. Froese, R. & Pauly, D. Fishbase. <https://www.fishbase.org> (2025). World Wide Web electronic publication., version (04/2025).
44. Posit Software, PBC. RStudio: Integrated Development Environment for R Version 2024.12.1.563 (2024).
45. Miqueleiz, I. *et al.* Ecoregional distributions of the world's freshwater vertebrate species. *Scientific Data* **12**, 1286, <https://doi.org/10.1038/s41597-025-05622-4> (2025).
46. Fricke, R., Eschmeyer, W. N. & Van der Laan, R. Eschmeyer's catalog of fishes: Genera, species, references. Electronic version (2026).
47. Amatulli, G. *et al.* Hydrography90m: a new high-resolution global hydrographic dataset. *Earth System Science Data* **14**, 4525–4550, <https://doi.org/10.5194/essd-14-4525-2022> (2022).
48. Basooma, A. *et al.* specleanr: an r package for automated flagging of environmental outliers in ecological data for modeling workflows. *Ecography* <https://doi.org/10.1002/ecog.08221> (2025).
49. Domisch, S. *et al.* GeoFRESH - an online platform for freshwater geospatial data processing. *International Journal of Digital Earth* **17**, 2391033, <https://doi.org/10.1080/17538947.2024.2391033> (2024).
50. Warton, D. I., Renner, I. W. & Ramp, D. Model-based control of observer bias for the analysis of presence-only data in ecology. *PLOS ONE* **8**, e79168, <https://doi.org/10.1371/JOURNAL.PONE.0079168> (2013).
51. Aiello-Lammens, M. E., Boria, R. A., Radosavljevic, A., Vilela, B. & Anderson, R. P. spThin: An R package for spatial thinning of species occurrence records for use in ecological niche models. *Ecography* **38**, 541–545, <https://doi.org/10.1111/ECOG.01132> (2015).
52. Chauvier, Y. *et al.* Novel methods to correct for observer and sampling bias in presence-only species distribution models. *Global Ecology and Biogeography* <https://doi.org/10.1111/geb.13383> (2021).
53. Schmidt-Kloiber, A. & Hering, D. An online tool that unifies, standardises and codifies more than 20,000 european freshwater organisms and their ecological preferences. *Ecological Indicators* **53**, 271–282 (2015).
54. Basooma, A. *et al.* Introducing fwtraits – an r package for obtaining freshwater biodiversity trait information. *Basic and Applied Ecology* **89**, 81–91, <https://doi.org/10.1016/j.baae.2025.10.010> (2025).
55. Pebesma, E. Simple features for R: standardized support for spatial vector data. *The R Journal* **10**, 439–446, <https://doi.org/10.32614/RJ-2018-009> (2018).
56. Pebesma, E. & Bivand, R. *Spatial Data Science: With Applications in R* (Chapman and Hall/CRC, Boca Raton, 2023).
57. Wickham, H. *et al.* Welcome to the tidyverse. *Journal of Open Source Software* **4**, 1686, <https://doi.org/10.21105/joss.01686> (2019).
58. Chamberlain, S. *et al.* *rgbif: Interface to the Global Biodiversity Information Facility API* (2025). (second edition edn, Elsevier, 2009).
59. Zizka, A. *et al.* CoordinateCleaner: standardized cleaning of occurrence records from biological collection databases. *Methods in Ecology and Evolution* **–7**, <https://doi.org/10.1111/2041-210X.13152> R package version 3.0.1 (2019).

Acknowledgements

We acknowledge the anonymous reviewer's suggestion to follow the Eschmeyer Catalog of Fishes (CoF) taxonomy and to include the corresponding CoF identifier in the database. Funding was provided by the European Union's Horizon Europe innovation action programme through the projects DANUBE4all (grant agreement no. 101093985), DANUBELifelines (grant agreement no. 101213836), and AquaINFRA (grant agreement no. 101094434). The financial support by the Austrian Federal Ministry for Digital and Economic Affairs, the National Foundation for Research, Technology and Development, and the Christian Doppler Research Association is gratefully acknowledged. AB acknowledges support by HR21 Doctoral School, BOKU University Vienna. MK acknowledges support and funding by Ministry of Environment of the Slovak republic under "Monitorovanie a hodnotenie stavu vôd" (grant agreement no. 24110110001), "Monitorovanie a hodnotenie stavu vôd - II. etapa" (grant agreement no. 24110110158) and "Monitorovanie a hodnotenie stavu vôd - III. etapa" (grant agreement no. 310011A366). SCJ acknowledges funding for the Leibniz Competition project "Freshwater Megafauna Futures". The authors also acknowledge funding from the German Research Foundation (NFDI4Earth, DFG project no. 460036893, <https://www.nfdi4earth.de>).

Author contributions

Conceptualization: Y.T.-C., M.Ts., M.S.-L., S.D. Data collation: A.A., A.B., A.M., A.Sa., A.Se., A.Sp., B.B., B.K., B.K.-M., B.P., B.T., C.N., D.N., D.V., G.C., G.T., I.Ce., I.Cz., I.N., J.G., J.P., K.L., L.P., M.B., M.Č., M.J., M.K., M.M., M.S., M.S.-L., M.Te., M.to., M.Ts., P.M., P.S., P.T., R.S., S.D.-M., S.O., S.R., TE, T.H., T.S., T.T., V.F., Y.T.-C., Z.F., Z.M., Z.Sa., Z.Sz., Z.V. Data collation coordination: M.S.-L. and M.Ts. Coding, data formatting, cleaning and visualization: Y.T.-C., A.B., V.B. Manuscript writing: Y.T.-C., M.Ts., S.D., S.J. and M.S.-L. wrote the first draft, with contributions from all authors. Technical validation: all authors reviewed the final database.

Funding

Open Access funding enabled and organized by Projekt DEAL.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41597-026-07249-5>.

Correspondence and requests for materials should be addressed to Y.T.-C.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2026

Yusdiel Torres-Cambas^{1,2}✉, András Ambrus³, Miklós Bán⁴, Bálint Bánó⁵, Anthony Basooma⁶, Vanessa Bremerich¹, Florian Borgwardt^{6,7}, Maša Čarf⁸, Irina Cernisencu⁹, Gorčin Cvijanović¹⁰, István Czeglédi⁵, Sami Domisch¹, Tibor Erős⁵, Zoltán Fehér¹¹, Vivien Füstös^{12,13}, Juergen Geist¹⁴, Thomas Hein^{6,7}, Milica Jaćimović¹⁰, Sonja C. Jähnig^{1,15}, Béla Kiss¹⁶, Maroš Kubala¹⁷, Klaudija Lebar¹⁸, Borislava Kostadinova Margaritova¹⁹, Matej Marušić²⁰, Paul Meulenbroek^{6,7}, Stoyan Dobrev Mihov¹⁹, Attila Mozsár⁵, Zoltán Müller¹⁶, Christoffer Nagel¹⁴, Iulian Nichersu⁹, Dušan Nikolić¹⁰, Sandi Orlić^{21,22}, Joachim Pander¹⁴, Polona Pengal²³, Marina Piria^{24,25}, László Polyák¹⁶, Bálint Preiszner⁵, Simon Rusjan¹⁷, Márton Sallai²⁶, Zoltán Sallai²⁷, Péter Sály²⁸, Andrea Samu¹¹, Brigitte Sasano²⁹, Astrid Schmidt-Kloiber⁶, András Sevcsik³⁰, Marija Smederevac-Lalić¹⁰, András Specziár⁵, Twan Stoffers^{1,31}, Zoltán Szalóky²⁸, Renáta Szita³², Gábor Takács³², Péter Takács⁵, Maxim Teichert²⁹, Milcho Todorov³³, Balázs Tóth³⁰, Teodora Trichkova³³, Damir Valić²¹, Zoltán Vitál²⁶ & Martin Tschikof^{6,7}

¹Leibniz Institute of Freshwater Ecology and Inland Fisheries, Department of Community and Ecosystem Ecology, Müggelseedamm 310, D-12489, Berlin, Germany. ²Universidad de Oriente, Department of Biology, Patricio Lumumba sn, Santiago de Cuba, Cuba. ³Independent Researcher, Jurisich u. 16., H-9495, Kópháza, Hungary. ⁴HUN-REN DE Behavioural Ecology Research Group, University of Debrecen, Egyetem tér 1, H-4010, Debrecen, Hungary. ⁵HUN-REN Balaton Limnological Research Institute, Tihany, H-8237, Hungary. ⁶BOKU University, Institute of Hydrobiology and Aquatic Ecosystem Management, Gregor-Mendel-Strasse 33, 1180, Vienna, Austria. ⁷Christian Doppler Laboratory for Meta Ecosystem Dynamics in Riverine Landscapes, BOKU University, Vienna, Institute of Hydrobiology and Aquatic Ecosystem Management, 1180, Vienna, Austria. ⁸Fisheries Research Institute of Slovenia, Spodnje Gameljne 61a, 1211, Ljubljana - Šmartno, Slovenia. ⁹Danube Delta National Institute for Research and Development, 165 Babadag street, Tulcea, 820112, Romania. ¹⁰University of Belgrade - Institute for Multidisciplinary Research, National Institute of the Republic of Serbia, Kneza Višeslava 1, 11030, Belgrade, Serbia. ¹¹WWF Hungary, Álmos vezér útja 69/A, 1141, Budapest, Hungary. ¹²Department of Hydraulic and Water Resources Engineering, Faculty of Civil Engineering, Budapest University of Technology and Economics, H-1111, Budapest, Hungary. ¹³HUN-REN-BME Water Management Research Group, Budapest, H-1111, Hungary. ¹⁴Aquatic Systems Biology Unit, Technical University of Munich (Technische Universität München), Mühlenweg 22, 85354, Freising-Weihenstephan, Germany. ¹⁵Geography Department, Humboldt-Universität zu Berlin, 10099, Berlin, Germany. ¹⁶BioAqua Pro Ltd., Debrecen, H-4032, Hungary. ¹⁷Water Research Institute, Nábr. arm. gen. L. Svobodu 5 (7), 812 49, Bratislava, Slovakia. ¹⁸University of Ljubljana, Faculty of Civil and Geodetic Engineering, Jamova 2, 1000, Ljubljana, Slovenia. ¹⁹WWF Bulgaria, 147 Knyaz Boris I Str., Floor 1, Sofia, 1000, Bulgaria. ²⁰DANUBE PARKS, Danube River Network of Protected Areas, Orth an der Donau, Austria. ²¹Ruđer Bošković Institute, Bijenička cesta 54, 10000, Zagreb, Croatia. ²²University of Montenegro, Cetinjski put 2, Podgorica, Montenegro. ²³Institute for Ichthyological and Ecological Research, Business Unit Ljubljana, Staretova ulica 1, 1233, Dob, Slovenia. ²⁴University of Zagreb Faculty of Agriculture, Zagreb, Croatia. ²⁵Department of Ecology and Vertebrate Zoology, Faculty of Biology and Environmental Protection, University of Łódź, Łódź, Poland. ²⁶Research Center for Fisheries and Aquaculture, Hungarian University of Agriculture and Life Sciences, Gödöllo, Hungary. ²⁷Vaskos Csabak Bt., Békésszentandrás, Hungary. ²⁸HUN-REN Institute of Aquatic Ecology, Centre for Ecological Research, Budapest, H-1113, Hungary. ²⁹Federal Agency for Water Management, Institute for Aquatic Ecology and Fisheries Management, Scharfling 18, 5310, Mondsee, Austria. ³⁰Duna-Ipoly National Park Directorate, Budapest, H-1121, Hungary. ³¹Aquaculture biology and Fisheries ecology group, Wageningen University & Research, Wageningen, The Netherlands. ³²Fertő-Hanság National Park Directorate, Sarród, H-9435, Hungary. ³³Institute of Biodiversity and Ecosystem Research, Bulgarian Academy of Sciences, 1 Tsar Osvoboditel Blvd., 1000, Sofia, Bulgaria. ✉e-mail: yusdiel.torres@gmail.com