Ecological Informatics 85 (2025) 102915

Contents lists available at ScienceDirect

Ecological Informatics

journal homepage: www.elsevier.com/locate/ecolinf

# ReLTER: An R interface for environmental observation in long term ecological research

Alessandro Oggioni [a,*], Micha Silver [b], Paolo Tagliolato [a], Arnon Karnieli [c]

[a] CNR - Institute for Electromagnetic Sensing of the Environment, (IREA), Via Alfonso Corti 12, I-20133, Milano, Italy

[b] Ben Gurion University, The Albert Katz International School for Desert Studies, The Jacob Blaustein Institutes for Desert Research, Sede Boker Campus, 8499000, Israel

[c] Ben Gurion University, Remote Sensing Laboratory, French Associates Institute for Agriculture and Biotechnology of Drylands, The Jacob Blaustein Institutes for Desert Research, Sede Boker Campus, 8499000, Israel

ARTICLE INFO

Dataset link: https://github.com/ropensci/ReLTER/

*Keywords:*
eLTER
Ecology
Taxonomy
Earth observation
Harmonisation

ABSTRACT

Ecological datasets are inherently complex, heterogeneous, and often span different time series. The links between spatial and temporal data and ecological/environmental science are framed in the discipline of ecoinformatics. The rapid growth of ecoinformatics has led to a demand for software collections specifically designed for environmental data analysis. This article introduces a new R package, `ReLTER`, simplifying user interaction with web data and the Dynamic Ecological Information Management System-Site and Dataset Registry (DEIMS-SDR). This package was motivated by users within the eLTER community and the emerging eLTER Research Infrastructure Management System (eLTER-RI), as well as other stakeholders to facilitate interaction with data and tools provided by eLTER-RI. Furthermore, DEIMS-SDR is currently used as a site catalogue by European environmental research projects, and has the potential to develop as the a reference for the EU ecological site registry.

The `ReLTER` package, written in the `R` language, is based on several auxiliary packages that enrich the data collected within the Long-Term Ecological Research (eLTER) network with data from other repositories. More than forty functions have been implemented that allow the retrieval of species occurrence data from third-party online networks, such as GBIF, iNaturalist, and OBIS, as well as earth observation images from various internet archives, and to harmonize the downloaded data (in terms of structure, semantics, taxonomy, etc.) for combining and linking them in subsequent analyses. This paper describes the functionality of the `ReLTER` package and its potential within the global ecological community.

## 1. Introduction

Since the 1990s, computers have become a mainstay of academic research and business, and today digital storage has replaced analogue storage (Hilbert and López, 2011). Datasets, especially in the field of ecology, are inherently complex, heterogeneous, and often span short or long time series, and require new techniques for acquisition, analysis, sharing, searching, transfer, visualisation, and more. The evolution of data management infrastructures has been slow but steady (Granell et al., 2009) marking the first steps towards interoperability at various levels (Heiler, 1995; Kissling et al., 2015; RDA-CODATA Legal Interoperability Interest Group, 2016; Hardisty et al., 2019), and the addition of principles related to findability, accessibility and reusability (Wilkinson et al., 2016; Mons et al., 2017), finally are moving towards a FAIR 2.0 with full semantic interoperability (Vogt et al., 2024). Activities of this kind are therefore intended for, and especially beneficial to European research infrastructures (Kolar et al., 2024), which leverage solutions from computer science and eco-informatics to share their data with a FAIR approach and effectively enable a connection between spatial and temporal data with ecological and environmental sciences (Recknagel, 2011).

One of the largest ecosphere observing networks is the International Long-Term Ecological Research Network (ILTER) (Mirtl et al., 2018; Muelbert et al., 2019; Vihervaara et al., 2013). Founded in 1993, ILTER currently consists of about 759 long-term observation facilities worldwide (as of September 2024 - https://deims.org/search/sites?f[0]=network_ri:ILTER), covering a broad range of ecosystems and aiming to understand global environmental change. LTER-Europe (Long-Term Ecosystem Research in Europe) is the community and networking

* Corresponding author.
*E-mail addresses:* alessandro.oggioni@cnr.it (A. Oggioni), silverm@post.bgu.ac.il (M. Silver), paolo.tagliolatoacquavivadar@cnr.it (P. Tagliolato), karnieli@bgu.ac.il (A. Karnieli).

https://doi.org/10.1016/j.ecoinf.2024.102915
Received 10 April 2024; Received in revised form 21 November 2024; Accepted 22 November 2024
Available online 10 December 2024

**Table 1**
A selective list of R packages designed to access data from research infrastructures, general-purpose data repositories, or large global organisations. The websites referenced in this table were last accessed on 18.09.2024.

| Name of R package | Purpose | Reference | Web page of package |
|---|---|---|---|
| `pangaear` | For data retrieval interface for the World Data Center PANGAEA repository (https://www.pangaea.de/) | Chamberlain et al. (2024) | https://docs.ropensci.org/pangaear/ |
| `rdryad` | To interface with the Dryad data sharing organisation (https://datadryad.org) | Ram et al. (2024) | https://docs.ropensci.org/rdryad/ |
| `rinat` | R wrapper for accessing to the iNaturalist (https://www.inaturalist.org/) observations trough API | Barve and Hart (2024) | https://docs.ropensci.org/rinat/ |
| `rgbif` | R package which gives you access to GBIF (https://www.gbif.org) mediated data | Chamberlain and Boettiger (2017) | https://docs.ropensci.org/rgbif/ |
| `MODIStsp` | R package devoted to automatizing the creation of time series of raster images derived from MODIS Land Products data | Busetto and Ranghetti (2016) | https://docs.ropensci.org/MODIStsp/ |
| `zen4R` | An interface to Zenodo (http://zenodo.org/) REST API data repository | Blondel and Barde (2024) | https://github.com/eblondel/zen4R/wiki |
| `CopernicusMarine` | For downloading and using the marine data from Copernicus | de Vries (2024) | https://pepijn-devries.github.io/CopernicusMarine/ |
| `argoFloats` | A package for downloading and analyzing collections of oceanographic Euro-Argo ERIC (https://www.euro-argo.eu) float datasets | Vitale et al. (2020) | https://argocanada.github.io/argoFloats/ |
| `RFlux` | Allows processing and cleaning eddy covariance flux measurements provided by ICOS ERIC (https://www.icos-cp.eu) | Vitale et al. (2020) | https://github.com/icos-etc/RFlux?tab=readme-ov-file |

component of the European LTER, and as of September 2024 (https://deims.org/search/sites?f[0]=network_ri:LTEREurope) includes more than 632 LTER long-term observation facilities, referring to these both as research sites and socio-ecological platforms (Holzer et al., 2019; Berthet et al., 2022). In 2018, LTER-Europe became part of the ESFRI (European Strategy Forum for Research Infrastructures) roadmap as eLTER Research Infrastructure (RI). The eLTER-RI focuses on long-term, site-based research and monitoring in the fields of ecosystem, biodiversity, critical zone, and social ecology. In this sense, eLTER RI has to deal with a vast amount of diverse, long-term data through ecoinformatics processes, including the collection, processing, analysis, synthesis, and transfer of data (Acri et al., 2020; Oggioni et al., 2023; Dirnböck et al., 2024; Lippl et al., 2024).

In ecology and environmental science, integration is fundamental (Lindenmayer and Likens, 2010; Schildhauer, 2018; Kühl et al., 2020; Wicquart et al., 2022) to identify patterns of complex physical, chemical and biological phenomena. Most scientific data infrastructures, among them the European Research Infrastructure Consortium (ERIC), provide programmatic access to their data catalogue, enabling data retrieval as part of a scientific software pipeline. In parallel, software packages that exchange information with specific services ensure the inclusion of these scientific infrastructures in the realm of reproducible open science. Such trends are becoming increasingly generalized with initiatives such as `ROpenSci` or `CRAN` and many packages have been developed to fulfill these purposes ( Table 1).

The `R` package described here is the result of requirements and needs gathered within the community of ecologists involved in eLTER (Peterseil et al., 2023a,b; Wurm et al., 2024), to address the scope and needs in long-term ecological research facilities. The result is important not only for the community of ecologists in the eLTER network but also for the global ecological community. Indeed, as in other packages, this package incorporates a 'one stop shop' approach in which complex and fragmented functions are consolidated and simplified for the benefit of ecology researchers who might not be familiar with the complex world of R-packages.

This article emphasizes how the `ReLTER` package enables programmatic access to data from existing archives and research infrastructures, combining and harmonizing them for new and reproducible analyses and merging data collected from the observing facilities registered in the Dynamic Ecological Information Management System-Site and Dataset Registry (DEIMS-SDR) (https://deims.org accessed on 18.09.2024) (Wohner et al., 2019). The innovative and relevant aspects of the package for the global ecological community can be summarized in these three points:

- the capability of the package to integrate and harmonize data from third-party repositories at various levels (e.g., format, units of measurement, time, semantic level, and taxonomic);
- the use of DEIMS-SDR - the hub developed by eLTER that allows for the registration and discovery of long-term research facilities on ecosystems worldwide — as a geographic information base for querying third-party repositories;
- the assumption that DEIMS-SDR will not only be used by ecological researchers within the eLTER network to independently record long-term observation facilities, but will also become the standard tool for archiving information on these resources. Already now, DEIMS-SDR stores information from observation facilities exclusively belonging to more than 40 international networks, European research projects and other European research infrastructures.

## 2. The ReLTER package

Researchers in ecology or environmental sciences often face difficulties using data analysis tools because they are not specifically developed to meet the community's requirements (Recknagel, 2011; Boyd and Foody, 2011). The design approach for the `ReLTER` package aims to address these challenges by providing a set of functions that ensure transparent access to global tabular and spatial datasets, simplifying operations, and effectively achieving harmonisation of the results obtained. These functions can be applied to specific research projects through a few computing platforms: scripts within the `R` language (R Core Team, 2021), using a docker container from the Rocker project (Boettiger and Eddelbuettel, 2017), or as computing notebooks in cloud platforms.

The various functions in this package offer straightforward access to databases of DEIMS-SDR site or platform metadata, global biodiversity and taxonomy, and global remote sensing-derived spatial data sets. This access is facilitated by drawing on, consolidating, and simplifying existing packages and functions in the extensive R computing framework. This framework includes the following R packages for spatial data: `sf` (Pebesma, 2018), `terra` (Hijmans, 2023), and `MODIStsp` (Busetto and Ranghetti, 2016), packages to acquire biodiversity and taxonomy data: `taxize` (Chamberlain and Szocs, 2013), and `worrms` (Chamberlain and Vanhoorne., 2023), and also `jqr` (FitzJohn et al., 2022) for querying JSON formatted metadata. Thus, ecological researchers can view `ReLTER` as both a starting point for data analysis in R and a 'one-stop shop' for acquiring global tabular and spatial datasets. Rather than serving as a portal, `ReLTER` acts as a connector, offering a series of functions to access and harmonize data shared by the eLTER network and other communities, research infrastructures, and repositories.

Currently, the `ReLTER` package contains some forty functions, created for facilitating the work of researchers from the eLTER and ILTER communities, and ecologists in general, in analyzing the metadata about observation facilities or resource types registered in DEIMS-SDR, but also and especially in enriching the knowledge about the eLTER long-term observation facilities (a 'facility' is defined as an in-situ long-term observation site or socio-ecological platform, delimited in space - e.g., https://deims.org/f30007c4-8a6e-4f11-ab87-569db54638fe - accessed on 18.09.2024) in the network with data from other repositories (Fig. 1). The goal of DEIMS-SDR is to serve as long-term observation facilities across all biomes, ensuring that data linked to each site is well-documented and accessible to the public, especially to scientists. Certain functions within `ReLTER` are designed for package configuration, while others generate plots, return tables, spatial vectors, or raster objects, and some facilitate data transformation into the eLTER data template (Peterseil and Geiger, 2022). The design goal is to allow the merging, harmonization, and standardization of information collected by DEIMS-SDR and other data repositories. In particular, the development of this package has focused on data retrieval, combination, and harmonization. Information or data from the different repositories, once collected by the `ReLTER` functions, are harmonized in at least five aspects: format, semantics, taxonomy, unit of measurement, and temporal domain. Refer to Fig. 1 for an overview of the interaction between DEIMS-SDR, other data repositories, and `ReLTER` functions.

In terms of format, all the functions that output tabular results (e.g., `get_site_info()`, `get_activity_info()`, `get_site_speciesOccurrences()`, `get_sos_obs()`) produce a `tibble`. In R environments, `tibble` handles two-dimensional arrays in which each column contains values of one variable, and each row includes one set of variables. According to the `tibble` package authors (Kirill and Hadley, 2023) (https://tibble.tidyverse.org - accessed on 18.09.2024) a `tibble` extends the base `data.frame` but with some simplifying assumptions: no change to variable names or types and no partial matching. Wickham and Grolemund (2017) describe a `tibble` as "data frames, but with tweaks to some older behavior to make life a little easier".

Given the strong geographic character of data from DEIMS-SDR and ecological data from other global repositories in general, the harmonization of formats is not limited to alphanumeric data. By adopting the spatial data format provided by the `sf` package (Pebesma, 2018) (https://r-spatial.github.io/sf/ - accessed on 18.09.2024), which allows adding a geometry column of any of the Open Geospatial Consortium (OGC) simple feature types, harmonized spatial data is ensured. Thus, the returned geospatial features from relevant functions can then undergo geographical operations (GEOS - https://libgeos.org - accessed on 18.09.2024), coordinate reference system transformation (PROJ - https://proj.org/en/9.3/ - accessed on 18.09.2024), and storage into spatial database archives such as PostGIS http://postgis.net (accessed on 18.09.2024) as part of a research task.

Semantic harmonization is achieved by associating a variable name within a `tibble` to a Uniform Resource Identifier (URI) of a Linked Data (Berners-Lee, 2006) entity, e.g., to a Simple Knowledge Organisation System (SKOS) (World Wide Web Consortium, 2009) term from a controlled vocabulary (Listing 5 - lines 34–40). The use of controlled vocabularies and unambiguous references (identical URIs are used for identical terms throughout the package) to terms in a SKOS vocabulary ensures semantic interoperability within this package. It allows for harmonization with other data sources as well. For example, variables named "Time" and "Air_Temperature" are always referred to http://www.opengis.net/def/property/OGC/0/PhenomenonTime (accessed on 18.09.2024) and http://vocabs.lter-europe.net/EnvThes/22035 (accessed on 18.09.2024). This association between variables and a SKOS term is implemented in `ReLTER` by attaching attributes to variables. Within R, the use of attributes allows the tagging of additional information to an R object (Adler, 2012). In `ReLTER`, objects are tagged with a list of URIs as an attribute named "uris", where the order of the links corresponds to the order of the variables in the table, achieving *de facto* semantic labeling. In the current version (v2.1), the following vocabularies were used for tagging: EnvThes eLTER Vocabulary (http://vocabs.lter-europe.net/EnvThes - accessed on 18.09.2024), NERC Vocabulary Server (NVS - https://vocab.nerc.ac.uk - accessed on 18.09.2024), Darwin Core standard (http://rs.tdwg.org/dwc/terms - accessed on 18.09.2024), and Open Geospatial Consortium (OGC - https://www.ogc.org - accessed on 18.09.2024).

A special type of semantic harmonization concerns taxonomic harmonization, which is also implemented in `ReLTER` through specific functions (e.g., `taxon_id_pesi()`, `taxon_id_worms()`). These functions harmonize datasets containing lists of species, also through interaction with the user, who decides the best name among those suggested by the taxonomic backbones.

Two specific R packages assist in harmonizing units of measurement (UOM) and temporal variables: `units` (Pebesma et al., 2016) (https://r-quantities.github.io/units/ - accessed on 18.09.2024) and `lubridate` (Grolemund and Wickham, 2011) (https://lubridate.tidyverse.org - accessed on 18.09.2024). These packages simplify the assignment, conversion, and derivation of units, including the exact definitions of time zones, thus fostering and facilitating data aggregation in eLTER RI and with other infrastructures.

## 3. ReLTER packages use cases

In this section, four use cases are presented as examples, demonstrating the main functions and potential of the `ReLTER` package. For a complete description of all functions, refer to the package documentation: (https://docs.ropensci.org/ReLTER/ - accessed on 18.09.2024) or vignettes (https://ropensci.r-universe.dev/ReLTER/ReLTER.pdf - accessed on 18.09.2024) of the package.

### 3.1. Discovering DEIMS-SDR resources

In the first use case, users seek to discover all datasets related to a specific site ecosystem (e.g., Alpine and Lakes) monitored within a network (e.g., LTER-Italy) in order to reach these datasets for future analysis. To facilitate this type of discovery, dedicated functions were developed to interact with the DEIMS-SDR API specification (https://deims.org/api - accessed on 18.09.2024) and to retrieve the necessary information. For each resource type (e.g., sites or long-term observation facilities, datasets, sensors, activities, and networks), functions were created to read the content of a resource and return a `tibble` with spatial information and semantic enrichment where appropriate (see Section 2 above): `get_site_info()`, `get_dataset_info()`, `get_sensor_info()`, `get_activities_info()`, and `get_network_*()`. Each of these functions has a mandatory parameter, the ID of the resource. This corresponds to the unique, resolvable, and persistent code assigned by eLTER-RI to every resource uploaded to DEIMS-SDR, and referred to as DEIMS.ID.

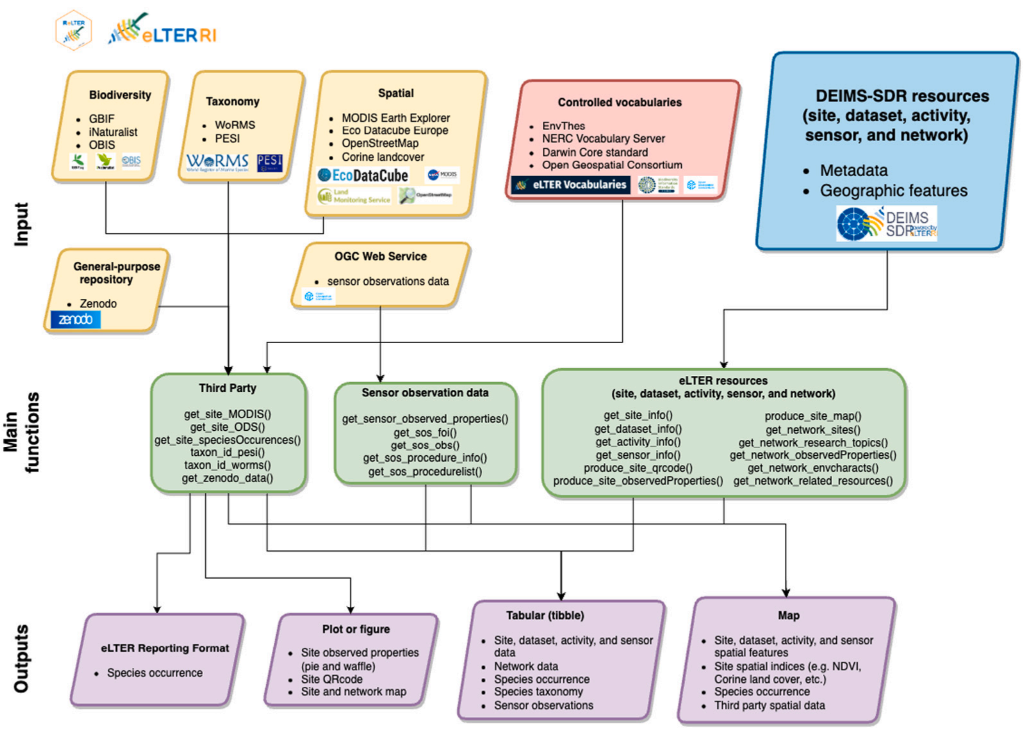


**Fig. 1.** Flow of inputs from third-party data sources through the ReLTER functions to produce various outputs.

```
1  listEnvCharacts <- ReLTER::get_network_sites(
2    networkDEIMSID =
3      "https://deims.org/network/7fef6b73-e5cb-4cd2-b438-
           ed32eb1504b3"
4  )
5
6  selected_sites <- listEnvCharacts$uri %>%
7    purrr::map_dfr(.f = function(x) {
8      x %>%
9        ReLTER::get_site_info(category = "EnvCharacts") %>%
10       dplyr::select(title, uri, geoCoord, ecosystemType) %>%
11       tidyr::unnest()
12   }) %>%
13   dplyr::filter(label == c("Lakes", "Alpine"))
14
15 # A tibble:
16 title  uri  geoCoord  label
17 <chr> <chr> <chr>    <chr>
18 1 Torgnon grassland [...]     https://deims.org/[...] POINT
        (7.57 45.[...] Alpine
19 2 Appennino centrale [...]    https://deims.org/[...] POINT
        (13.36 42.[...]  Alpine
20 3 Istituto Scientifi[...]     https://deims.org/[...] POINT
        (7.86 45.[...] Lakes
21 [...]
22
23 related_resources <- selected_sites$uri %>%
24   purrr::map_dfr(.f = function(x) {
25     x %>%
26       ReLTER::get_site_info(category = "RelateRes") %>%
27       dplyr::select(title, uri, geoCoord, relatedResources) %>%
28       tidyr::unnest()
29   }) %>%
30   dplyr::filter(!is.na(relatedResourcesTitle))
31
32 resource_info <- ReLTER::get_dataset_info(
33   datasetid = related_resources$uri1[43],
34   show_map = TRUE
35   ) %>%
36   dplyr::select(
37     title, abstract, doi,
38     legal.rights, method.instrumentation,
39     method.methodDescription, boundaries
40   )
```

**Listing 1** Example code demonstrating the combination of the functions `get_network_sites()`, `get_site_info()`, and `get_dataset_info()` to identify dataset(s) shared by alpine lake long-term observation facilities within the Italian eLTER-RI network.

A user who wants to know what data are available through the DEIMS-SDR on the Alpine Lake long-term observation facilities of the Italian LTER network needs only the unique identifier (DEIMS.ID) of the network as the parameter to the `get_network_sites()` function (Listing 1 - lines 1–4). This function returns a spatial point vector object, including title and DEIMS.ID (URI), coordinates, and label, of all the DEIMS-SDR long-term observation facilities belonging to a national eLTER Network.

By iterating through the URIs (DEIMS.ID of the site) from the above output table (`listEnvCharacts`) within the `get_site_info()` function, it is possible to filter and select only those long-term observation facilities that have 'Lakes' and 'Alpine' as equivalent terms in the `ecosystemType` under the category environmental characteristics (Listing 1 - lines 6–13).

The same `get_site_info()` function, this time using the URIs of the reduced list (`selected_sites`, Listing 1 - lines 23–30) of the Alpine lake long-term observation facilities only, with a specific request for the site metadata related to the `RelatedRes` category, obtains a list of all the datasets (Listing 1 - lines 23–30) distributed through DEIMS-SDR.

Finally, with the use of the `get_dataset_info()` function (Listing 1 - lines 32–38), using any DEIMS.ID (`uri1`) of a dataset present in the `tibble` `related_resources`, it is possible to obtain a new `tibble` containing all the metadata of that dataset. Significant elements of metadata that can be found are title, abstract, DOI if provided by the user, etc. (Fig. 2). Finally, thanks to the `show_map` parameter, available in all the functions presented in this section, it is also possible to obtain a map of the geographical feature, in this case, a point where the dataset was collected or which the dataset represents (Fig. 3).

| title | abstract | doi | legal.rights | method.instrumentation | method.methodDescription |
|---|---|---|---|---|---|
| <chr> | <chr> | <chr> | <chr> | <chr> | <chr> |
| A long-term (1986-2010) phytoplankton [...] | Lake Candia This georeferenced dataset describes a 25-year (1986-2010) monitoring studies of phytoplankton [...] | https://doi.org/10.15468/8k32hw | NA | NA | Study extent Phytoplankton integrated samples in the whole euphotic zone were gathered approximately monthly, at the station of lake maximum depth (7.7 m). [...] |

**Fig. 2.** A section of the `tibble` is obtained from the function `get_dataset_info()` (result of listing 1 - lines 32–40). The fields contain the information provided by DEIMS-SDR about a specific related resource (dataset).

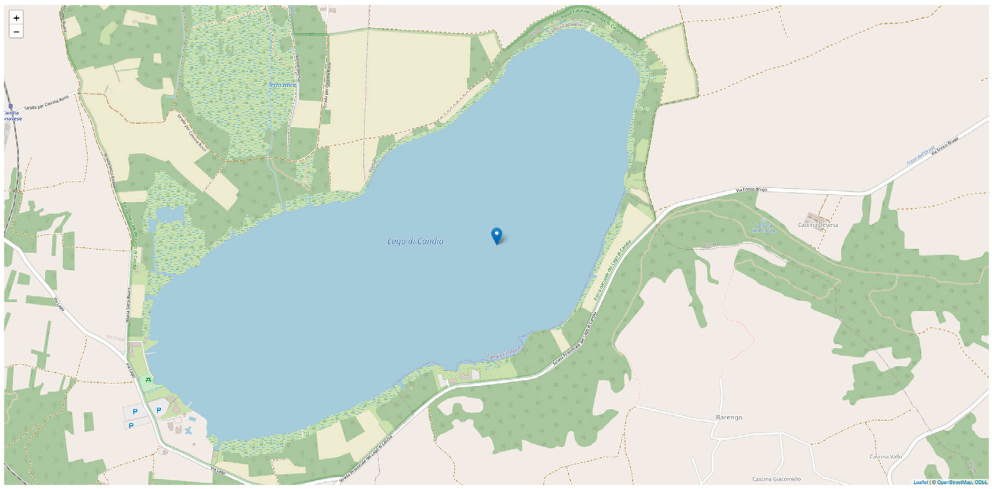


**Fig. 3.** The map represents the feature of where the selected dataset (Listing 1) was collected.

### 3.2. Dataset provided by external repositories

In the second use case, users who have registered their site or platform into DEIMS-SDR might need to determine whether relevant data are available in third-party repositories.

The ability to merge data collected in DEIMS-SDR site or platform with data from external, third-party data repositories expands the scope of knowledge for studying the DEIMS-SDR site or platform itself. The functions described below implement this merging of data:

- using geoprocessing operations such as clipping (Wiki.GIS.com, 2012), to cut a piece of a large geo-referenced dataset using one or more geographical features from another dataset (e.g., `get_site_speciesOccurrence()` and `get_site_MODIS()`, `get_site_ODS()` described in the next section (see sub-Section 3.3);
- requesting specific data from generic repositories (e.g., `get_zenodo_data()`) or web services (e.g., for retrieving INSPIRE data from Sensor Observation Services `get_sos_obs()`) (INSPIRE MIG sub-group MIWP-7a, 2016);
- interfacing with reliable sources to enrich a dataset with certified taxonomic information (e.g. author of the scientific name and taxon id — Listing 5, lines 26–32), effectively enriching the original dataset (e.g., `taxon_id_pesi()` or `taxon_id_worms()`).

The `get_site_speciesOccurrence()` function acquires species occurrence data detected in a specific DEIMS-SDR site or platform from the Global Biodiversity Information Facility (GBIF - https://www.gbif.org - accessed on 18.09.2024), iNaturalist (https://www.inaturalist.org/ - accessed on 18.09.2024) and the Ocean Biodiversity Information System (OBIS - https://obis.org/ - accessed on 18.09.2024), using R packages: `rgbif` (Chamberlain and Boettiger, 2017) and `spocc` (Owens et al., 2023).

In the example given in the Listing 2, the Gulf of Venice DEIMS-SDR site or platform (DEIMS.ID https://deims.org/758087d7-231f-4f07-bd7e-6922e0c283fd - accessed on 18.09.2024) is used to subset occurrences provided by iNaturalist and OBIS data portals. The results obtained are shown in (Fig. 4) containing all species occurrences and a series of `tibbles`, one for each of the repositories browsed, with the information provided by the repositories themselves.

```
occ_GoV <- ReLTER::get_site_speciesOccurrences(
  deimsid =
    "https://deims.org/758087d7-231f-4f07-bd7e-6922e0c283fd",
  list_DS = c("inat", "obis"),
  show_map = TRUE,
  limit = 50
)
```

**Listing 2** Example code of the function `get_site_speciesOccurence()` to request the first 50 species occurrences (`limit = 50`), present in the iNaturalist and OBIS repositories (`list_DS`) and included within the boundaries of the eLTER Gulf of Venice site (`deimsid`).

Zenodo (European Organization For Nuclear Research and OpenAIRE, 2013) is a general-purpose open repository developed under the European OpenAIRE program and operated by CERN. It allows researchers to deposit research papers, data sets, research software, reports, and any other research digital products. The `get_zenodo_data()` retrieves data stored in Zenodo and allows it to be merged with others (Listing 3). This function first downloads the data set and then creates a tibble of the same data set.

```
record <- ReLTER::get_zenodo_data(
  doi = "10.5281/zenodo.7041152"
)
```

**Listing 3** Code for download the dataset deposits in Zenodo.

The `get_sos_obs()` function obtains the observations shared by Sensor Observation Service (SOS) (Listing 4 and Fig. 5). This feature was developed since many observations made at the eLTER network long-term observation facilities are sensor-based; therefore, the network has adopted the standard web service SOS as the standard for

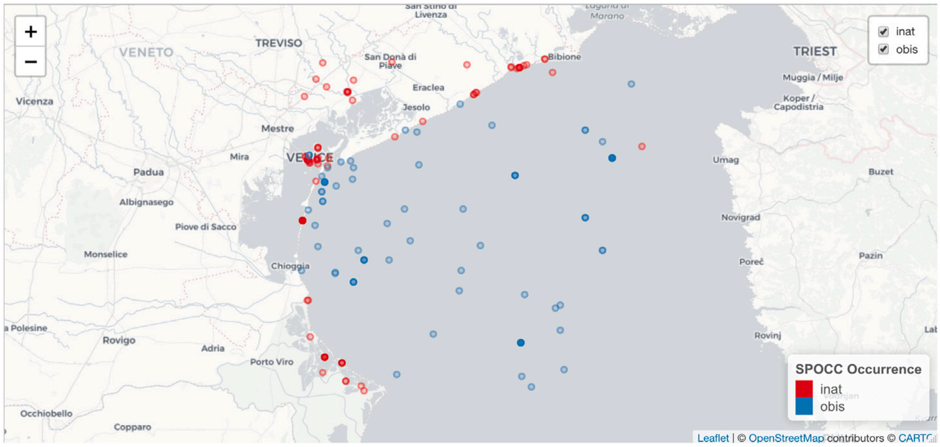


**Fig. 4.** Map representation of the distribution of the first 50 occurrences (`limit = 50`, Listing 2), within the boundary of the Gulf of Venice site, for each repository consulted. In red for occurrences relating to iNaturalist and in blue those relating to OBIS. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

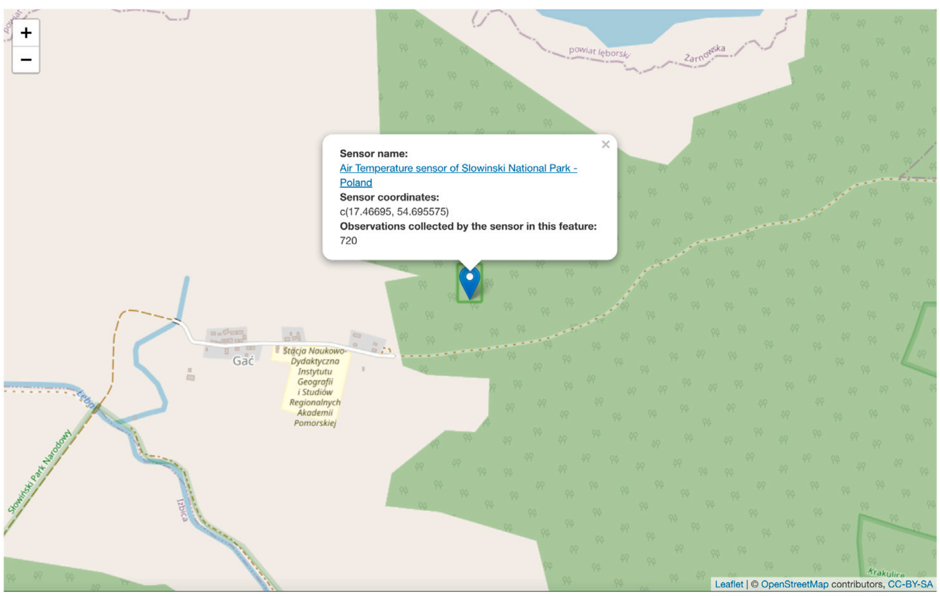


**Fig. 5.** The map shows the location where the sensor made observations.

distributing these types of observations. `get_sos_obs()` returns a map with the position of the sensor and a tibble with values harmonized in terms of format, a unit of measurement, time, and semantics.

```
airTemp <- ReLTER::get_sos_obs(
  sosURL = "http://getit.lteritalia.it/observations/service",
  procedure = "http://www.get-it.it/sensors/getit.lteritalia.it
        /procedure/noOwnerDeclared/noModelDeclared/
        noSerialNumberDeclared/SI001469-SPNAirTemp", #720
        Slowinski National Park
  show_map = TRUE
)
```

**Listing 4** An example of requesting data to the SOS web service, specifying the service endpoint (sosURL) and unique sensor identifier (procedure).

To query for species taxa, the `taxon_id_worms()` function allows for taxonomic validation of the submitted dataset through the World Register of Marine Species — WoRMS (Ahyong et al., 2023) backbone service. The function returns a `tibble` object that includes all columns of the input table (`input`, Listing 5 - lines 1–16), plus new columns about accepted scientific name, scientific name authorship, systematic categories (such as kingdom, phylum, class, order, family, and genus), and synonyms (Listing 5 - lines 17–24). The function may perform validation automatically, choosing among the taxa included in WoRMS partial name matching or a series of similar records for interactive choice by the user. This option is activated through the refine attribute (Listing 5 - line 15). Similarly, the function `taxon_id_pesi()` uses the taxonomic backbone service developed by PESI (Pan-European Species-directories Infrastructure, http://www.eu-nomen.eu/pesi/ accessed on 18.09.2024) (de Jong et al., 2015). In both cases, the columns of the output table are associated with Darwin Core vocabulary terms (Listing 5 - lines 25–31). These functions also use all aspects of harmonization.

```
1 phytoplankton <- tibble::tibble(
2   ID = c(1, 2, 3, 4, 5, 6, 7),
3   species = c(
4     "Asterionella formosa", "Chrysococcus sp.",
5     "Cryptomonas rostrata", "Dinobryon divergens",
```

```
    "Mallomonas akrokomos", "Melosira varians",
    "Cryptomonas rostrata"
  )
)
> head(phytoplankton, 4)
# A tibble: 4 x 2
     ID species
  <dbl> <chr>
1     1 Asterionella formosa
2     2 Chrysococcus sp.
3     3 Cryptomonas rostrata
4     4 Dinobryon divergens

taxon_tbl <- ReLTER::taxon_id_worms(
  input = phytoplankton,
  taxaColumn = 2,
  verbose = TRUE,
  refine = TRUE
)
> head(taxon_tbl)
# A tibble: 6 x 7
     ID originalNameUsage     scientificName
           scientificNameAuthorship taxonID order          family
  <dbl> <chr>                 <chr>
                          <int> <chr>          <chr>
1     1 Asterionella formosa Asterionella formosa Hassall, 1850
                   148954 Fragilariales Fragilariaceae
2     2 Chrysococcus sp.      NA                   NA
                              NA NA                NA
3     3 Cryptomonas rostrata Cryptomonas rostrata Skuja, 1948
                    572024 NA              NA
4     4 Dinobryon divergens   Dinobryon divergens   O.E.Imhof,
      1887            157248 Chromulinales Dinobryaceae
5     5 Mallomonas akrokomos Mallomonas akrokomos Ruttner, 1913
                   249722 Synurales      Mallomonadaceae
6     6 Melosira varians      Melosira varians      C.Agardh,
      1827             149043 Melosirales    Melosiraceae

> attributes(taxon_tbl)$uri
[1] ""
[2] "http://rs.tdwg.org/dwc/terms/originalNameUsage"
[3] "http://rs.tdwg.org/dwc/terms/scientificName"
[4] "http://rs.tdwg.org/dwc/terms/scientificNameAuthorship"
[5] "http://rs.tdwg.org/dwc/terms/taxonID"
[6] "http://rs.tdwg.org/dwc/terms/order"
[7] "http://rs.tdwg.org/dwc/terms/family"
```

**Listing 5** An example of taxon enrichment and Semantic harmonisation using `taxon_id_worms()` function.

### 3.3. Remote sensing data functions

The third use case, similar to the previous (see sub- Section 3.2), exclusively involves data from third-party repositories but related to remote sensing. In this case, users want to obtain spatial data for their site or platform, such as vegetation indices like the Normalized Difference Vegetation Index (NDVI) or data related to land cover.

Recognizing the importance of Earth Observation (EO) data in environmental monitoring programs, ReLTER includes functions for accessing remote sensing data. ReLTER developers implemented functions to acquire EO data, allowing transparent access to remote sensing imagery, along with the tabular data archives of biodiversity, taxonomy, and eLTER monitoring long-term observation facilities.

Remote sensing data are exposed through two ReLTER functions: `get_site_MODIS()` and `get_site_ODS()`.

The first adds the ability to acquire either of two MODIS products (https://modis.gsfc.nasa.gov/data/ - accessed on 18.09.2024): vegetation indices or land surface temperature and crop to an eLTER study area. Users specify a date range, and the function acquires all images during the requested date interval. Then, average pixel values over the study area of each time step are computed, and a time series plot is displayed.

The second EO function utilizes the high-resolution remote sensing data available from https://ecodatacube.eu (accessed on 18.09.2024). This service offers many EO products covering Europe, all homogenized to 30 m resolution and projected to the Lambert Azimuthal Equal-Area (LAEA) coordinate system used widely in Europe. The `get_site_ODS()` function exposes the following datasets: three landcover layers, OpenStreetMap (OSM) buildings, seasonal NDVI, as well as NDVI trend, and three forest layers. What is more, similar to `get_site_MODIS()`, the downloaded layers are cropped to the boundaries of a given DEIMS-SDR site or platform.

The code snippet below (Listing 6) demonstrates finding and downloading a boundary polygon for a specific DEIMS-SDR site or platform (Braila Islands, https://deims.org/d4854af8-9d9f-42a2-af96-f1ed9cb25712- accessed on 18.09.2024), then acquiring two RS datasets, cropped to that polygon, from Eco Datacube Europe (see Fig. 6).

```
# ReLTER functions to get DEIMS-SDR site or platform polygon
# and raster data from Braila Island LTER
library(tmap)
library(terra)
braila <- get_ilter_generalinfo(
  country_name = "Romania", site_name = "Braila")
braila_id = braila$uri
braila_polygon = get_site_info(braila_id, "Boundaries")

braila_landcover = get_site_ODS(braila_id, dataset = "landcover
     ")
names(braila_landcover) <- "Landcover"
braila_ndvi_summer = get_site_ODS(braila_id, "ndvi_summer")
names(braila_ndvi_summer) <- "NDVI summer"

# Plot NDVI and Landcover maps
tmap_mode("plot")
ndvi_map <- tm_shape(braila_polygon) +
  tm_borders(col = "purple") +
  tm_shape(braila_ndvi_summer) +
  tm_raster(alpha=0.7, palette = "RdYlGn", midpoint = NA) +
  tm_layout(legend.position = c("left", "top"))

landcover_map <- tm_shape(braila_polygon) +
  tm_borders(col = "purple") +
  tm_shape(braila_landcover) +
  tm_raster(alpha=0.7, palette = "Dark2") +
  tm_layout(legend.position = c("left", "top"))
```

**Listing 6** Example using RS data from Eco Datacube Europe

Both functions return R raster objects, handled by the `terra` package (Hijmans, 2023). Thus, users can perform further analyses on these images, such as correlating with *in situ* data or comparing between several long-term observation facilities.

### 3.4. Production of site or network figure

The fourth use case demonstrates how the functions implemented in the ReLTER package can be used to visualize the distribution of long-term observation facilities within the eLTER network, long-term observation facilities registered in DEIMS-SDR, or specific characteristics recorded in DEIMS-SDR related to long-term observation facilities (e.g., observed property). In order to support and facilitate the production of maps, charts, or barcodes of the information obtained from DEIMS-SDR, specific functions are included in the ReLTER package.

The `produce_site_map()` takes a boundary polygon for a specific DEIMS-SDR site or platform and produces a figure with the site boundaries overlaid on an OpenStreetMap (OpenStreetMap contributors, 2024), and with compass rose, map scale, country viewport, site title, and DEIMS.ID (Fig. 7), ready for use in publications or reports requiring a geographical representation of the described site.

A comparable approach is implemented to create a map of all the DEIMS-SDR long-term observation facility members of a country network. The function `produce_network_points_map()`, starting from the DEIMS.ID of the network, generates a figure (see Fig. 8) displaying the locations of all the long-term observation facilities in the country network, as well as the country borders. The aim is once again to provide a pre-made figure.

The function `produce_site_qrcode()` has been created to translate any DEIMS.ID into a QR code. This barcode could be used

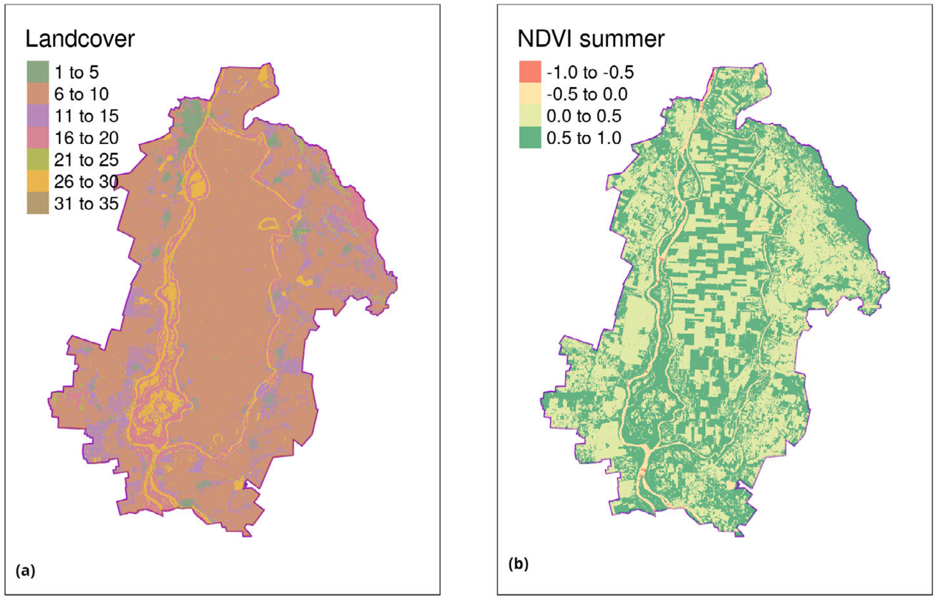


**Fig. 6.** Two remote sensing maps at the Braila eLTER site in Romania. (a) Landcover (Corine) classification, and (b) NDVI during summer months. Both are acquired from the Eco Datacube Europe (https://ecodatacube.eu - accessed on 18.09.2024) website.

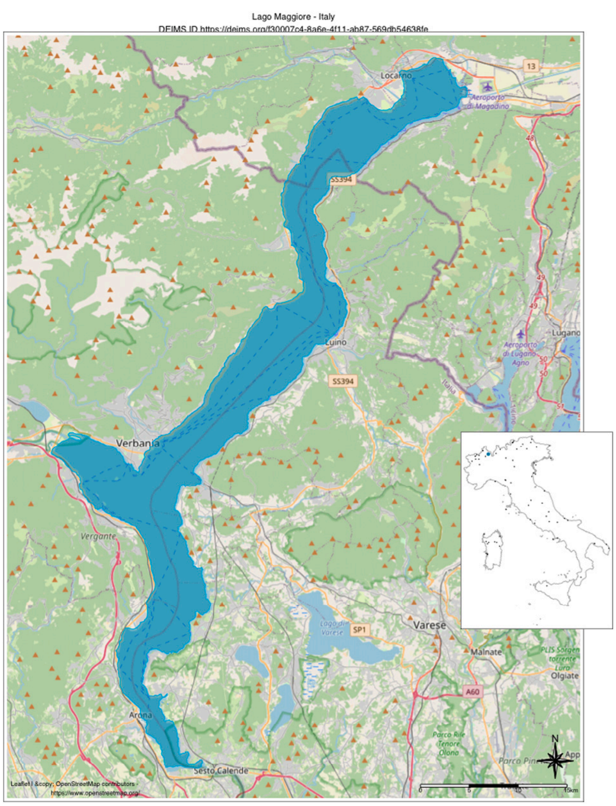


**Fig. 7.** Lake Maggiore (DEIMS.ID https://deims.org/f30007c4-8a6e-4f11-ab87-569db54638fe - accessed on 18.09.2024) map with country viewport to show where the site is located within the country.

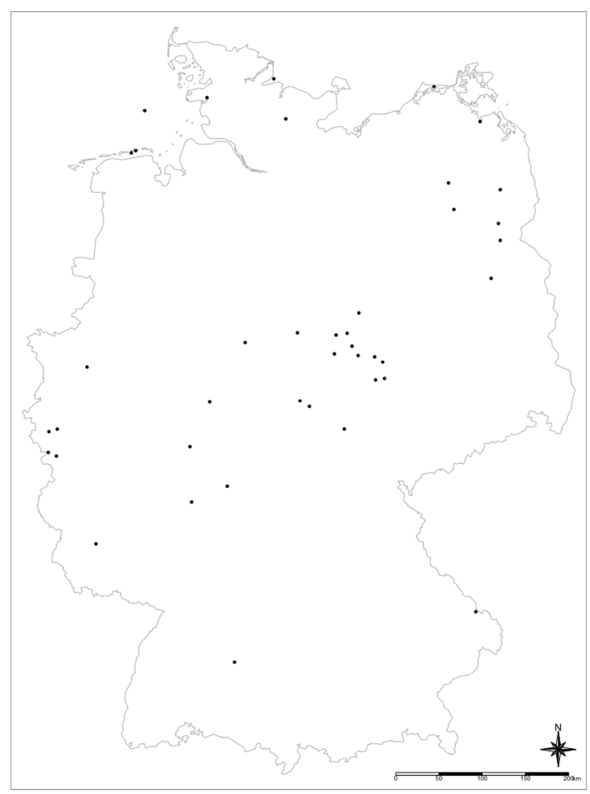

**Fig. 8.** Geographic distribution of the German eLTER site network (DEIMS.ID https://deims.org/networks/e904354a-f3a0-40ce-a9b5-61741f66c824 - accessed on 18.09.2024) is shown on the map.

anywhere to provide information about any resource stored in the DEIMS SDR repository simply by scanning this two-dimensional matrix.

Finally, ReLTER contains a pair of functions (`produce_site_observedProperties_pie()` and `produce_site_observedProperties_waffle()`) that produce a pie and a waffle representation of the observed properties at a site. The observed properties refer to the chemical, physical, and biological variables measured at the research site, and they are grouped into groups (e.g., agricultural, atmospheric, biological). The pie chart (Fig. 9A) shows the percentage distribution of the observed properties in the different groups, while the waffle chart (Fig. 9B) shows the number of observed properties per

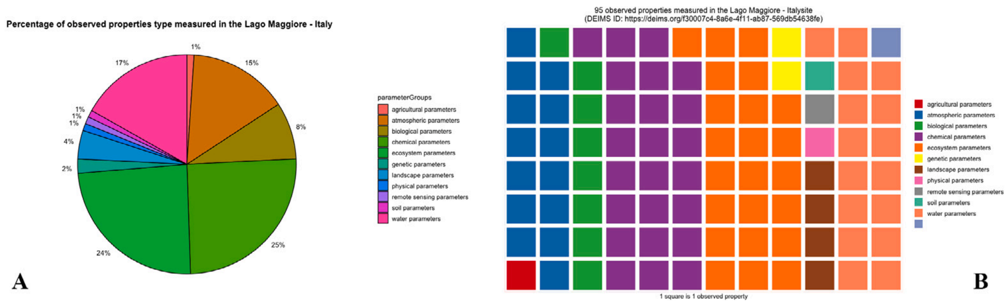


**Fig. 9.** Pie and waffle charts of observed properties measured in Lake Maggiore (DEIMS.ID - https://deims.org/f30007c4-8a6e-4f11-ab87-569db54638fe - accessed on 18.09.2024).

group, both provide immediate visualisation of the research conducted at the site and can be used as such in publications, reports or for compare long-term observation facilities.

## 4. Discussion and conclusions

In recent decades, ecological science has played a significant role in studying the dynamics of Earth biosystems. This recognized status has strengthened in light of accelerating environmental changes due to long-term climatic shifts and anthropogenic activities. Demanding prompt solutions, ecological research strives to understand life processes influencing organism abundance, distribution, and diversity. More than in other fields, data integration is ubiquitous and essential in ecology (Schildhauer, 2018; Kühl et al., 2020), effectively enhancing the understanding of these changes (Carpenter et al., 2009), but certain challenges, such as standardisation and reproducibility, still require further effort (Huber et al., 2021). Disciplines such as ecoinformatics make possible improved understanding through dedicated and open tools that encourage the reproducibility of analyses.

In this regard, the ILTER network and the European research infrastructure eLTER-RI are aimed at conducting advanced ecological and socio-ecological research to understand environmental phenomena in the context of global change. Scientists create, collect, manage, maintain, analyze, and document continuous long-term ecological data accumulated into diverse datasets, including parameters about different spheres: biological, hydrological, geological, ecological, and sociological. Complementary to this, various remote sensing systems and other research infrastructures or networks acquire data to provide an efficient means of exploring the physical, chemical, and biological properties to support the sustainable development of human civilisation and the global environment by monitoring and assessing natural and man-made environmental conditions and changes.

Resting on the stable backbone of the open-source `R` software ecosystem, the `ReLTER` package advances standardisation and enables researchers to conduct open and reproducible data analyses (Pebesma et al., 2012). By adopting the approach and licenses of open source software, the scientific research pipelines using this package and its specific functions are transparent to future scrutiny, in line with the European Commission's open source software strategy (European Commission, 2020). The data resources accessible by `ReLTER` are also open. Therefore, research procedures using the tools they contain can be shared, examined, reproduced, and modified by others.

The DEIMS-SDR registry offers the eLTER-RI research infrastructure and other continent-wide projects a data depository for long term, observations, and metadata. By interfacing with this registry, the `ReLTER` package allows users to: i. perform discovery of resources archived in DEIMS-SDR (see sub- Section 3.1); ii. access third-party repositories to enhance ongoing research (see sub-Sections 3.2 and 3.3); iii. create visualizations or charts based on the information available in DEIMS-SDR; and iv. harmonize the obtained results in terms of format, semantics, taxonomy, units of measurement, and time (see Section 3). The functions in the `ReLTER` package, especially those that interact with DEIMS-SDR, have been implemented to catch certain exceptions or incorrect data found in the eLTER repository. Users of the package can contribute to its improvement by reporting errors through the issue reporting page (https://github.com/ropensci/ReLTER/issues - accessed on 07.11.2024).

The current version of `ReLTER` (v2.0) suffers from dependency on many other packages: there are currently almost 40 required during installation. The authors are aware that this number of dependencies could affect future functionality, requiring constant updating to adapt to API changes or deprecation of external packages in future R releases. In the version currently under development, a meticulous effort has already begun to consolidate the code by removing some dependencies and integrating functionality directly in `ReLTER` to achieve a more efficient and feature-rich package, resilient to external changes.

Additional improvements are also under consideration. Notably: i. lazy loading, such that external packages will be installed and loaded only on execution of a function that requires that package; ii. integration of data from other research infrastructure repositories (e.g., LifeWatch, ICOS, DISSCO, ICOS, TERENO, PANGAEA).

`ReLTER` has filled the standardization gaps for the ecological research community using DEIMS-SDR as the main repository for sharing information about long-term observation facilities, networks, and activities. Specialised functions are implemented to allow uniformity of procedures, to make data analysis reproducible, and to reduce effort and to facilitate interdisciplinary data reuse. Based on these characteristics, `ReLTER` has currently been integrated into the Virtual Research Environment (VRE), which is in the final stages of development for the ITINERIS project (Italian Integrated Environmental Research Infrastructures System funded by EU by National Recovery and Resilience Plan-Next Generation EU). Given the involvement of some of the authors in this activity and others being planned and the release of the package as an open project on GitHub and ROpenSci, ongoing maintenance and integration of new functions in the package are assured.

## 5. Availability

The `ReLTER` package is available, as a open-source software (GNU General Public License) on the GitHub repository at https://github.com/ropensci/ReLTER/.

## CRediT authorship contribution statement

**Alessandro Oggioni:** Writing – review & editing, Writing – original draft, Software, Funding acquisition, Conceptualization, Supervision. **Micha Silver:** Writing – review & editing, Writing – original draft, Software, Conceptualization. **Paolo Tagliolato:** Writing – review & editing, Writing – original draft, Software, Conceptualization. **Arnon Karnieli:** Writing – review & editing, Writing – original draft.

## Acknowledgments

This work has been partially funded from the European Union's Horizon 2020 research and innovation program under the H2020 eLTER-Plus Project (https://elter-ri.eu/elter-plus - accessed on 18.09.2024), grant agreement No 871128 (DOI - https://doi.org/10.3030/871128).

## Data availability

The ReLTER package is available, as a open-source software (GNU General Public License) on the GitHub repository at https://github.com/ropensci/ReLTER/.

## References

Acri, F., Bastianini, M., Bernardi Aubry, F., Camatti, E., Boldrin, A., Bergami, C., Cassin, D., De Lazzari, A., Finotto, S., Minelli, A., Oggioni, A., Pansera, M., Sarretta, A., Socal, G., Pugnetti, A., 2020. [A long-term (1965–2015) ecological marine database from the LTER-Italy northern adriatic sea site: plankton and oceanographic observations]. Earth Syst. Sci. Data 12 (1), 215–230. http://dx.doi.org/10.5194/essd-12-215-2020, URL: https://essd.copernicus.org/articles/12/215/2020/.

Adler, J., 2012. R in a Nutshell, second ed. O'Reilly, Tokyo, Includes index.

Ahyong, S., Boyko, C., Bailly, N., Bernot, J., 2023. World Register of Marine Species (WoRMS). WoRMS Editorial Board, URL: https://www.marinespecies.org. (Last accessed 11 March 2024).

Barve, V., Hart, E., 2024. rinat: Access 'iNaturalist' data through APIs. R package version 0.1.9 https://docs.ropensci.org/rinat/ (website) https://github.com/ropensci/rinat (devel).

Berners-Lee, T., 2006. Linked data - design issues. W3C, http://www.w3.org/DesignIssues/LinkedData.html.

Berthet, E.T., Bretagnolle, V., Gaba, S., 2022. Place-based social-ecological research is crucial for designing collective management of ecosystem services. Ecosyst. Serv. 55, 101426. http://dx.doi.org/10.1016/j.ecoser.2022.101426, URL: https://www.sciencedirect.com/science/article/pii/S2212041622000225.

Blondel, E., Barde, J., 2024. zen4R: R Interface to Zenodo REST API. Zenodo, http://dx.doi.org/10.5281/zenodo.11500665.

Boettiger, C., Eddelbuettel, D., 2017. The R Journal: An introduction to Rocker: Docker containers for R. R J. 9, 527–536. http://dx.doi.org/10.32614/RJ-2017-065.

Boyd, D., Foody, G., 2011. An overview of recent remote sensing and GIS based research in ecological informatics. Ecol. Inform. 6 (1), 25–36. http://dx.doi.org/10.1016/j.ecoinf.2010.07.007, URL: https://www.sciencedirect.com/science/article/pii/S1574954110000828, Special Issue: 5th Anniversary.

Busetto, L., Ranghetti, L., 2016. MODIStsp: an R package for preprocessing of MODIS Land Products time series. Comput. Geosci. 97, 40–48. http://dx.doi.org/10.1016/j.cageo.2016.08.020, URL: https://github.com/ropensci/MODIStsp.

Carpenter, S.R., Armbrust, E.V., Arzberger, P.W., Chapin, III, F.S., Elser, J.J., Hackett, E.J., Ives, A.R., Kareiva, P.M., Leibold, M.A., Lundberg, P., Mangel, M., Merchant, N., Murdoch, W.W., Palmer, M.A., Peters, D.P.C., Pickett, S.T.A., Smith, K.K., Wall, D.H., Zimmerman, A.S., 2009. Accelerate synthesis in ecology and environmental sciences. BioScience 59 (8), 699–701. http://dx.doi.org/10.1525/bio.2009.59.8.11, arXiv:https://academic.oup.com/bioscience/article-pdf/59/8/699/711931/59-8-699.pdf.

Chamberlain, S., Boettiger, C., 2017. R Python, and Ruby Clients for GBIF Species Occurrence Data. PeerJ, Inc., http://dx.doi.org/10.7287/peerj.preprints.3304v1.

Chamberlain, S., Szocs, E., 2013. Taxize - taxonomic search and retrieval in R. F1000Research URL: https://f1000research.com/articles/2-191/v2.

Chamberlain, S., Vanhoorne., B., 2023. worrms: World register of marine species (WoRMS) client. URL: https://CRAN.R-project.org/package=worrms, R package version 0.4.3.

Chamberlain, S., Woo, K., MacDonald, A., Zimmerman, N., Simpson, G., 2024. pangaear: Client for the 'Pangaea' Database. URL: https://github.com/ropensci/pangaear (devel), R package version 1.1.0, https://docs.ropensci.org/pangaear/ (documentation).

de Jong, Y., Kouwenberg, J., Boumans, L., Hussey, C., 2015. PESI - a taxonomic backbone for Europe. Biodivers. Data J. 3, e5848. http://dx.doi.org/10.3897/BDJ.3.e5848.

de Vries, P., 2024. CopernicusMarine: Search download and handle data from copernicus marine service information. URL: https://github.com/pepijn-devries/CopernicusMarine, R package version 0.2.3.002, https://pepijn-devries.github.io/CopernicusMarine/.

Dirnböck, T., Bahn, M., Diaz-Pines, E., Djukic, I., Englisch, M., Gartner, K., Gollobich, G., Hofbauer, A., Ingrisch, J., Kitzler, B., Knaebel, K., Kobler, J., Maier, A., Wohner, C., Offenthaler, I., Peterseil, J., Pröll, G., Venier, S., Zechmeister, S., Zolles, A., Glatzel, S., 2024. High-resolution Carbon cycling data from 2019 to 2021 measured at six Austrian LTER sites. Earth Syst. Sci. Data Discuss. 2024, 1–23. http://dx.doi.org/10.5194/essd-2024-110, URL: https://essd.copernicus.org/preprints/essd-2024-110/.

European Commission, 2020. Open source software strategy 2020 – 2023. URL: https://commission.europa.eu/document/download/97e59978-42c0-4b4a-9406-8f1a86837530_en.

European Organization For Nuclear Research, OpenAIRE, 2013. Zenodo. CERN, http://dx.doi.org/10.25495/7GXK-RD71, URL: https://www.zenodo.org/.

FitzJohn, R., Ooms, J., Chamberlain, S., Stefan Milton Bache, 2022. jqr: Client for 'jq', a 'JSON' Processor. URL: https://CRAN.R-project.org/package=jqr, R package version 1.2.3.

Granell, C., Gould, M., Manso, M.A., Bernabe, M.A., 2009. Spatial data infrastructures. In: Karimi, H.A. (Ed.), Handbook of Research on Geoinformatics. IGI Global, Hershey, PA, USA, pp. 36–41. http://dx.doi.org/10.4018/978-1-59140-995-3.ch005, URL: https://services.igi-global.com/resolvedoi/resolve.aspx?doi=10.4018/978-1-59140-995-3.ch005.

Grolemund, G., Wickham, H., 2011. Dates and times made easy with lubridate. J. Stat. Softw. 40 (3), 1–25, URL: https://www.jstatsoft.org/v40/i03/.

Hardisty, A.R., Michener, W.K., Agosti, D., Alonso García, E., Bastin, L., Belbin, L., Bowser, A., Buttigieg, P.L., Canhos, D.A., Egloff, W., De Giovanni, R., Figueira, R., Groom, Q., Guralnick, R.P., Hobern, D., Hugo, W., Koureas, D., Ji, L., Los, W., Manuel, J., Manset, D., Poelen, J., Saarenmaa, H., Schigel, D., Uhlir, P.F., Kissling, W.D., 2019. The Bari Manifesto: An interoperability framework for essential biodiversity variables. Ecol. Inform. 49, 22–31. http://dx.doi.org/10.1016/j.ecoinf.2018.11.003, URL: https://www.sciencedirect.com/science/article/pii/S1574954118301961.

Heiler, S., 1995. Semantic interoperability. ACM Comput. Surv. 27 (2), 271–273. http://dx.doi.org/10.1145/210376.210392.

Hijmans, R.J., 2023. terra: Spatial Data Analysis. URL: https://CRAN.R-project.org/package=terra, R package version 1.7-3.

Hilbert, M., López, P., 2011. The world's technological capacity to store, communicate, and compute information. Science 332 (6025), 60–65. http://dx.doi.org/10.1126/science.1200970, arXiv:https://www.science.org/doi/pdf/10.1126/science.1200970, URL: https://www.science.org/doi/abs/10.1126/science.1200970.

Holzer, J., Adamescu, C., Cazacu, C., Díaz-Delgado, R., Dick, J., Méndez, P., Santamaría, L., Orenstein, D., 2019. Evaluating transdisciplinary science to open research-implementation spaces in European social-ecological systems. Biol. Cons. 238, 108228. http://dx.doi.org/10.1016/j.biocon.2019.108228, URL: https://www.sciencedirect.com/science/article/pii/S0006320719305397.

Huber, R., D'Onofrio, C., Devaraju, A., Klump, J., Loescher, H.W., Kindermann, S., Guru, S., Grant, M., Morris, B., Wyborn, L., Evans, B., Goldfarb, D., Genazzio, M.A., Ren, X., Magagna, B., Thiemann, H., Stocker, M., 2021. Integrating data and analysis technologies within leading environmental research infrastructures: Challenges and approaches. Ecol. Inform. 61, 101245. http://dx.doi.org/10.1016/j.ecoinf.2021.101245, URL: https://www.sciencedirect.com/science/article/pii/S1574954121000364.

INSPIRE MIG sub-group MIWP-7a, 2016. Technical Guidance for Implementing Download Services Using the OGC Sensor Observation Service and ISO 19143 Filter Encoding. Technical Report, EU INSPIRE Maintenance and Implementation Group (MIG).

Kirill, M., Hadley, W., 2023. tibble: Simple data frames. https://tibble.tidyverse.org/, https://github.com/tidyverse/tibble.

Kissling, W.D., Hardisty, A., García, E.A., Santamaria, M., De Leo, F., Pesole, G., Freyhof, J., Manset, D., Wissel, S., Konijn, J., Los, W., 2015. Towards global interoperability for supporting biodiversity research on essential biodiversity variables (EBVs). Biodiversity 16 (2–3), 99–107. http://dx.doi.org/10.1080/14888386.2015.1068709.

Kolar, J., Brečko, B.N., Horvat, M., Kralj, A., Sterle, P., Šarler, B., Pipan, T., Rozman, D., Čeh, M., Malnar, B., 2024. ESFRI Landscape Analysis 2024. ESFRI, Brussels, URL: https://landscape2024.esfri.eu/media/coqdoq0q/20240604_la2024.pdf.

Kühl, H.S., Bowler, D.E., Bösch, L., Bruelheide, H., Dauber, J., Eichenberg, D., Eisenhauer, N., Fernández, N., Guerra, C.A., Henle, K., Herbinger, I., Isaac, N.J., Jansen, F., König-Ries, B., Kühn, I., Nilsen, E.B., Pe'er, G., Richter, A., Schulte, R., Settele, J., van Dam, N.M., Voigt, M., Wägele, W.J., Wirth, C., Bonn, A., 2020. Effective biodiversity monitoring needs a culture of integration. One Earth 3 (4), 462–474. http://dx.doi.org/10.1016/j.oneear.2020.09.010, URL: https://www.sciencedirect.com/science/article/pii/S2590332220304796.

Lindenmayer, D.B., Likens, G.E., 2010. The science and application of ecological monitoring. Biol. Cons. 143 (6), 1317–1328. http://dx.doi.org/10.1016/j.biocon.2010.02.013, URL: https://www.sciencedirect.com/science/article/pii/S0006320710000522.

Lippl, F., Maringer, A., Kurka, M., Abermann, J., Schöner, W., Hirschmugl, M., 2024. A benchmark data set for long-term monitoring in the eLTER site Gesäuse–Johnsbachtal. Earth Syst. Sci. Data Discuss. 2024, 1–17. http://dx.doi.org/10.5194/essd-2024-12, URL: https://essd.copernicus.org/preprints/essd-2024-12/.

Mirtl, M., Borer, E., Djukic, I., Forsius, M., 2018. Genesis, goals and achievements of Long-Term Ecological Research at the global scale: A critical review of ILTER and future directions. Sci. Total Environ. 626, 1439–1462. http://dx.doi.org/10.1016/j.scitotenv.2017.12.001, URL: https://www.sciencedirect.com/science/article/pii/S0048969717334204.

Mons, B., Neylon, C., Velterop, J., Dumontier, M., da Silva Santos, L.O.B., Wilkinson, M.D., 2017. Cloudy, increasingly FAIR; revisiting the FAIR Data guiding principles for the European Open Science Cloud. Inf. Serv. Use 37 (1), 49–56. http://dx.doi.org/10.3233/ISU-170824, Publisher: IOS Press.

Muelbert, J.H., Nidzieko, N.J., Acosta, A.T.R., Beaulieu, S.E., 2019. ILTER – The international long-term ecological research network as a platform for Global Coastal and ocean observation. Front. Mari. Sci. 6, http://dx.doi.org/10.3389/fmars.2019.00527, URL: https://www.frontiersin.org/articles/10.3389/fmars.2019.00527.

Oggioni, A., Ruggiu, D., Morabito, G., Pugnetti, A., Sparber, K., Cozza, R., Panzani, P., Ruffoni, T., Austoni, M., 2023. A long-term (1986–2010) phytoplankton dataset from the LTER-Italy site Lake Candia. J. Limnol. 82 (s1), http://dx.doi.org/10.4081/jlimnol.2023.2122, URL: https://www.jlimnol.it/jlimnol/article/view/2122.

OpenStreetMap contributors, 2024. Planet dump retrieved from https://planet.osm.org, https://www.openstreetmap.org. (Last accessed 11 March 2024).

Owens, H., Barve, V., Chamberlain, S., 2023. spocc: Interface to species occurrence data sources. https://github.com/ropensci/spocc (devel), https://docs.ropensci.org/spocc/ (user manual).

Pebesma, E., 2018. Simple features for R: Standardized support for spatial vector data. R J. 10 (1), 439–446. http://dx.doi.org/10.32614/RJ-2018-009.

Pebesma, E., Mailund, T., Hiebert, J., 2016. Measurement units in R. R J. 8 (2), 486–494. http://dx.doi.org/10.32614/RJ-2016-061.

Pebesma, E., Nüst, D., Bivand, R., 2012. The R software environment in reproducible geoscientific research. EOS Trans. Am. Geophys. Union 93 (16), 163. http://dx.doi.org/10.1029/2012EO160003, URL: https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2012EO160003.

Peterseil, J., Djukic, I., Oggioni, A., Zilioli, M., Tagliolato, P., Bolton, W., Wohner, C., Marangi, C., Poppe, C., Provencale, A., Alice, B., Parisi, A., Pons, M.-N., 2023a. eLTER PLUS Deliverable D4.1 Workflow for Retrieval and Harmonisation of Legacy Data. Zenodo, http://dx.doi.org/10.5281/zenodo.8279757.

Peterseil, J., Parland-von Essen, J., Lukkarinen, A., Pursula, A., Oggioni, A., Zilioli, M., Watkins, J., Rennie, S., Brown, M., Magagna, B., Goldfarb, D., Dirnböck, T., Haase, P., Baatz, R., Dick, J., Orenstein, D., 2023b. eLTER PLUS Deliverable D10.2 Report on Data and ICT needs from Research Challenges (RCs) to be used in Virtual Access (VA). Zenodo, http://dx.doi.org/10.5281/zenodo.8279774.

Peterseil, J., Geiger, S., 2022. Field specification for data reporting. http://dx.doi.org/10.5281/zenodo.6373410.

R Core Team, 2021. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria, URL: https://www.R-project.org/.

Ram, K., Chamberlain, S., Boettiger, C., 2024. rdryad: Access for dryad web services. URL: https://docs.ropensci.org/rdryad, R package version 1.0.1.95, https://github.com/ropensci/rdryad.

RDA-CODATA Legal Interoperability Interest Group, 2016. Legal Interoperability of Research Data: Principles and Implementation Guidelines. Zenodo, http://dx.doi.org/10.5281/zenodo.162241.

Recknagel, F., 2011. Ecological informatics: A discipline in the making. Ecol. Inform. 6 (1), 1–3. http://dx.doi.org/10.1016/j.ecoinf.2010.12.002, URL: https://linkinghub.elsevier.com/retrieve/pii/S1574954110001135.

Schildhauer, M., 2018. Data integration: Principles and practice. In: Recknagel, F., Michener, W.K. (Eds.), Ecological Informatics: Data Management and Knowledge Discovery. Springer International Publishing, Cham, pp. 129–157. http://dx.doi.org/10.1007/978-3-319-59928-1_8.

Vihervaara, P., D'Amato, D., Forsius, M., Angelstam, P., 2013. Using long-term ecosystem service and biodiversity data to study the impacts and adaptation options in response to climate change: insights from the global ILTER sites network. Curr. Opin. Environ. Sustain. 5 (1), 53–66. http://dx.doi.org/10.1016/j.cosust.2012.11.002, URL: https://linkinghub.elsevier.com/retrieve/pii/S187734351200187X.

Vitale, D., Fratini, G., Bilancia, M., Nicolini, G., Sabbatini, S., Papale, D., 2020. A robust data cleaning procedure for eddy covariance flux measurements. Biogeosciences 17 (6), 1367–1391. http://dx.doi.org/10.5194/bg-17-1367-2020, URL: https://bg.copernicus.org/articles/17/1367/2020/.

Vogt, L., Strömert, P., Matentzoglu, N., Karam, N., Konrad, M., Prinz, M., Baum, R., 2024. FAIR 2.0: Extending the FAIR guiding principles to address semantic interoperability. arXiv:2405.03345.

Wickham, H., Grolemund, G., 2017. R for Data Science: Import, Tidy, Transform, Visualize, and Model Data, first ed. O'Reilly Media, Inc.

Wicquart, J., Gudka, M., Obura, D., Logan, M., Staub, F., Souter, D., Planes, S., 2022. A workflow to integrate ecological monitoring data from different sources. Ecol. Inform. 68, 101543. http://dx.doi.org/10.1016/j.ecoinf.2021.101543, URL: https://www.sciencedirect.com/science/article/pii/S1574954121003344.

Wiki.GIS.com, 2012. Clip. http://wiki.gis.com/wiki/index.php/Clip. (Last accessed 11 March 2024).

Wilkinson, M.D., Dumontier, M., Aalbersberg, I.J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.-W., da Silva Santos, L.B., Bourne, P.E., Bouwman, J., Brookes, A.J., Clark, T., Crosas, M., Dillo, I., Dumon, O., Edmunds, S., Evelo, C.T., Finkers, R., Gonzalez-Beltran, A., Gray, A.J., Groth, P., Goble, C., Grethe, J.S., Heringa, J., 't Hoen, P.A., Hooft, R., Kuhn, T., Kok, R., Kok, J., Lusher, S.J., Martone, M.E., Mons, A., Packer, A.L., Persson, B., Rocca-Serra, P., Roos, M., van Schaik, R., Sansone, S.-A., Schultes, E., Sengstag, T., Slater, T., Strawn, G., Swertz, M.A., Thompson, M., van der Lei, J., van Mulligen, E., Velterop, J., Waagmeester, A., Wittenburg, P., Wolstencroft, K., Zhao, J., Mons, B., 2016. The FAIR Guiding Principles for scientific data management and stewardship. Sci. Data 3 (1), 160018. http://dx.doi.org/10.1038/sdata.2016.18.

Wohner, C., Peterseil, J., Poursanidis, D., Kliment, T., Wilson, M., Mirtl, M., Chrysoulakis, N., 2019. DEIMS-SDR – A web portal to document research sites and their associated data. Ecol. Inform. 51, 15–24. http://dx.doi.org/10.1016/j.ecoinf.2019.01.005, URL: https://www.sciencedirect.com/science/article/pii/S1574954118302528.

World Wide Web Consortium, 2009. SKOS Simple Knowledge Organization System Reference. URL: https://www.w3.org/TR/skos-reference/. (Last accessed 11 March 2024).

Wurm, F., Langlet, A., Vadrot, A., 2024. Surveying biodiversity data needs among stakeholders. URL: https://zenodo.org/communities/marco-bolo.