

Italian Fluvial Sediment Transport Database: A Comprehensive Hub for Archiving and Analyzing Hydrological Data

Marco Luppichini¹, Francesco Comiti², Andrea Andreoli³, Barbara Aldighieri⁴, Paolo Billi⁵, Irene Maria Bollati⁶, Andrea Brenna⁶, Diletta Chirici⁷, Paolo Ciavola⁵, Silvia Cilli⁵, Marco Donnini⁹, Simona Francalanci¹⁰, Lorenzo Innocenti¹⁰, Laura Longoni¹¹, Luca Mao¹², Ivan Marchesini⁹, Ilenia Murgia⁸, Giacomo Pellegrini¹², Daniele Penna^{8,9,13}, Lorenzo Picco², Riccardo Rainato¹⁴, Leonardo Schippa¹⁵, Luca Solari¹⁰, Gianluca Tronti⁷, Bruno Testa⁴, Matteo Verdone⁸, Monica Bini^{1,16,17}

1. Department of Earth Sciences, University of Pisa, Via S. Maria 52, 56126 Pisa, Italy
2. University of Padua, Italy
3. Free University of Bozen-Bolzano, Italy
4. Consiglio Nazionale delle Ricerche – Istituto di Geoscienze (CNR-GEO), Via Mario Bianco 9, 20131 Milano, Italy
5. Department of Physics and Earth Sciences, Università degli Studi di Ferrara, Via Saragat 1, 44122 Ferrara, Italy
6. Consorzio di Bonifica della Romagna, Via R. Lambruschini, 195, 47521 Cesena (FC)
7. Earth Science Department “A. Desio”, Università degli Studi di Milano, Italy
8. Department of Agricultural, Food, Environmental and Forestry Sciences and Technologies, University of Florence, Via San Bonaventura 13, 50145 Firenze, Italy
9. Consiglio Nazionale delle Ricerche – Istituto di Geoscienze (CNR-GEO), Via della Madonna Alta 126, 06128 Perugia, Italy
10. Department of Civil and Environmental Engineering, University of Florence, Via S. Marta 3, 50139 Firenze, Italy
11. Department of Civil and Environmental Engineering, Politecnico di Milano, Piazza Leonardo da Vinci 32, 20133 Milano (MI), Italy;
12. Department of Geography, Climate, Catchments, Coasts and Communities Research Group, University of Lincoln, Brayford Pool, Lincoln LN67TS, UK
13. Forest Engineering Resources and Management Department, Oregon State University, Corvallis, OR, USA
14. Regione Veneto, Italy

15. Department of Engineering, University of Ferrara, Via Saragat 1, 44122 Ferrara, Italy

16. CIRSEC – Centro Interdipartimentale di Ricerca per lo Studio degli Effetti del
Cambiamento Climatico, University of Pisa, Italy

17. Istituto Nazionale di Geofisica e Vulcanologia (INGV), Rome, Italy

Corresponding author: Marco Luppichini, marco.luppichini@dst.unipi.it

Abstract

Fluvial sediment transport plays a key role in geomorphological and hydrological processes, influencing river morphology, watershed and coastal sediment balances, and the environmental response to climatic and anthropogenic changes. In Italy, the availability of homogeneous and long-term data is strongly limited due to the discontinuity of monitoring activities and the fragmentation of existing sources. This study presents the development of a relational database and web application specifically designed for the collection, storage, and consultation of sediment transport data, aimed at integrating and enhancing heterogeneous datasets from both historical and contemporary monitoring networks: the Italian Fluvial Sediment Transport Database (ITASED). The database architecture, developed in PostgreSQL with the PostGIS extension, enables a direct link between observational data and their associated geographical, instrumental, and methodological metadata, ensuring traceability, interoperability, and the possibility to perform multi-temporal analyses. The web interface, built using open-source technologies, allows controlled data entry and interactive data exploration through maps, tables, and export functions. Analysis of the archived datasets highlights strong instrumental, temporal, and spatial heterogeneity: some physical-chemical parameters (e.g., pH and electrical conductivity) show standardized measurement protocols, whereas sediment transport variables exhibit high methodological variability. Time series range from sub-daily to annual observations, with denser coverage in northern Italian basins (e.g., Po, Adige, Piave, Tagliamento) since 1924. Despite these inconsistencies, integration within a unified relational framework enhances the value of a largely underused data heritage. The experience gained through this project enabled the identification of both the limitations and the potential of Italy's sediment monitoring system, providing operational guidance for methodological standardization, metadata improvement, and data harmonization at the national scale. The ITASED represents a concrete step toward establishing a shared sedimentological archive that supports scientific research, environmental management, and sustainable river basin planning.

Keywords: Sediment transport database, Fluvial monitoring, Data standardization, Hydro-sediment dynamics, Italian river basins

1 Introduction

Sediment transport in rivers plays a key role in geomorphological and hydrological processes, shaping the morphological evolution of fluvial environments, influencing the sediment balances at the catchment scale and in coastal areas (Fortesa et al., 2021; Vercruysse et al., 2017; Zarfl & Dunn, 2022). Sediments carried by rivers contribute to the genesis of floodplains, deltas, and sediment replenishment in coastal systems, processes that have significant ecological and economic implications (Leatherman et al., 2000; Luppichini & Bini, 2025; K. Zhang et al., 2004). Their dynamics result from a complex interplay of multiple watershed local characteristics, including climatic conditions, geomorphological features, geological substrate, vegetation cover, and anthropogenic activities (Cox et al., 2021; Wei et al., 2023; J. Zhang et al., 2024). This interdependence among the aforementioned factors is central to the study of fluvial systems and represents a critical research field for understanding the mechanisms driving these phenomena and for developing targeted strategies for sustainable land management (Asselman, 1995; Gurnell et al., 2002; Hooke, 2006; Naik & Jay, 2011).

The extent and nature of these interactions vary considerably across basins. For example, climatic fluctuations, including changes in precipitation patterns and temperature, along with the intensification of natural disturbances, directly affect sediment mobilization and transport by modifying both sediment availability and the transport capacity of river flows (Jadhav et al., 2024). Furthermore, anthropogenic activities such as afforestation-deforestation dynamics, urban expansion, infrastructure development, in-channel sediment mining and water resource management have profoundly modified sediment transport regimes. These interventions can alter erosional processes in mountainous catchments, disrupt natural sediment fluxes through the construction of dams, weirs and check-dams, and modify sediment transport capacity via water abstraction or river channelization and training (Asselman, 1995; Favaro & Lamoureux, 2015; Naik & Jay, 2011; Rodríguez-Blanco et al., 2016; Surian & Rinaldi, 2004; Thodsen et al., 2008).

Fare clic o toccare qui per immettere il testo. Despite the importance of these phenomena, a comprehensive understanding of the interactions between climatic variability, anthropogenic pressures, and sediment dynamics is often hampered by the lack of systematic monitoring data and by the fragmented accessibility of existing records. This limitation is particularly critical because fluvial sediment transport is an extremely complex process, featuring high spatial and temporal variability, and it is composed of two fractions (suspended load and bedload), which are controlled by partly distinct mechanisms. Therefore, such fractions may respond differently to natural and human-driven pressures. Suspended sediment transport is generally more directly linked to the mobilization of fine material from hillslopes (featuring different land uses), floodplains and channel banks, and is strongly influenced by precipitation intensity, runoff generation and vegetation cover (Battista et al., 2020; Fryirs, 2013; Vercruysse et al., 2017). In comparison to the suspended fraction, bedload transport, is more strongly dependent on the availability

106 within the catchment of mobile coarser fractions, as well as on their connectivity within and to
107 the channel network. Additional important factors for bedload fluxes are channel morphology,
108 grain-size distribution, transport thresholds, and the occurrence and legacy of flood events.
109 (Surian & Rinaldi, 2004)

110 For this reason, the use of undifferentiated or poorly documented sediment datasets can limit
111 the interpretation of hydro-sedimentary processes and may lead to misleading estimates of
112 total sediment transport. Properly separated datasets for suspended load, bedload, and,
113 where available, total load are therefore needed to distinguish the processes involved, evaluate
114 the role of natural and anthropogenic drivers, and assess sediment budgets at the catchment
115 scale. This requirement is particularly relevant for long-term analyses, where changes in
116 monitoring methods, sampling frequency, instrumentation, and metadata completeness can
117 strongly affect data comparability. This condition, with some degree of variability, is widely
118 observed at the global scale (Walling & Fang, 2003). In particular, the scarcity of long,
119 continuous, and consistent time series represents a significant obstacle to quantifying the
120 specific contribution of climatic and anthropogenic factors to the different components of
121 sediment transport and to accurately predicting how future environmental changes may alter
122 riverine and coastal sediment balances (Luppichini et al., 2024; Syvitski et al., 2005)

123 In Italy, this issue is particularly pronounced. From the 1930s to the 1990s, the Italian
124 Hydrographic and Ocean Service (*Servizio Idrografico e Mareografico Italiano*, SIMI) played a
125 crucial role in systematically collecting suspended sediment transport data from major rivers
126 across the country. However, with the dissolution of the service in the 1990s, the responsibility
127 for monitoring was transferred to regional authorities, which faced significant challenges in
128 maintaining continuity (Billi & Spalevic, 2022). Nonetheless, monitoring activities persisted in
129 some basins and research observatories, although often with limited resources and reduced
130 operational capacity. Examples include long-term or repeated monitoring experiences in the
131 Adige basin, in the Piave basin and Rio Cordon instrumented catchment, and in other northern
132 Italian basins such as the Po and Tagliamento, where sediment transport and related
133 hydrological variables continued to be investigated through regional monitoring activities and
134 research-based initiatives (Bonfrisco et al., 2025; Pellegrini et al., 2023a; Rainato et al., 2017,
135 2025a). In central Italy, additional datasets from rivers such as the Arno, Ombrone, and Versilia
136 further document the persistence of local monitoring and research efforts, although generally
137 with more fragmented spatial and temporal coverage (Francalanci et al., 2013a, 2013b). This
138 has led to a highly fragmented data landscape: on one hand, there are paper-based historical
139 records that are difficult to access and are not fully exploited; on the other hand, more recent
140 data have been collected sporadically by public agencies or research groups, often without
141 national-level coordination. The lack of a centralized platform designed to collect and
142 harmonize these datasets has hindered their use for long-term analyses and to support water
143 and land management policies.

144 To address this critical issue, the present work addresses two main objectives. The first is to
145 create and strengthen a network of experts who, both in Italy and abroad, have been collecting

146 data on sediment transport for several decades. Establishing this network aims to foster
147 knowledge exchange and best practices, encouraging greater collaboration among institutions,
148 universities, and land management authorities. The second objective is to develop a
149 centralized database designed to store both historical and future data, integrating them into a
150 single, accessible, and well-structured repository. Although the database is centred on
151 sediment transport, it was designed with a broader hydro-sedimentary perspective and
152 therefore also includes associated hydrological and physicochemical variables, such as flow
153 rate, flow height, temperature, pH, electrical conductivity, and major ions. These variables
154 provide essential context for interpreting sediment dynamics and for analysing the
155 relationships between river flow conditions, water properties, and sediment transport
156 processes. As part of this process, a web-based call for sediment transport data was launched
157 (in 2025) to collect existing sediment transport and related hydrological datasets from public
158 institutions, research groups, river basin authorities, and monitoring agencies. In this study, the
159 term “data call” refers to a targeted request for available datasets and metadata, especially for
160 data not already accessible through open repositories or standardised national platforms. This
161 step was also intended to consolidate the expert network mentioned above and to identify
162 dispersed or poorly accessible datasets suitable for harmonisation within the proposed
163 database.

164 Comparable harmonization efforts have long been implemented in hydrology through
165 international initiatives such as the Global Runoff Data Centre (GRDC) (Färber et al., 2025) and
166 the Global Streamflow Indices and Metadata archive (GSIM) (Do et al., 2018). These databases
167 integrate streamflow records from multiple national agencies and research institutions,
168 applying standardized metadata structures and quality-control procedures to ensure
169 consistency and comparability across regions. In contrast, sediment transport data are still
170 largely fragmented, often dispersed among individual research projects or local monitoring
171 programs. The development of the present database represents a first step toward a similar
172 level of coordination in the field of fluvial sediment monitoring, promoting interoperability and
173 long-term data accessibility.

174 The adopted IT infrastructure, comprising the database architecture, server environment, and
175 web-based interface, combines a solid, well-organised backend with a flexible, intuitive
176 frontend. This approach ensures not only data security and integrity but also broader
177 dissemination of information, making it accessible to both technical experts and users with
178 limited expertise. The availability of a centralized platform marks a significant step forward in
179 addressing the current gaps in sediment transport monitoring, providing a valuable tool for land
180 management and scientific research. In the long term, this initiative could facilitate greater
181 integration of sediment data at the European level, contributing to a more comprehensive
182 understanding of sediment dynamics in response to climate change and anthropogenic
183 pressures.

184 2 Material and Methods

185 The conceptual and logical design of the database and its graphical interface was carried out
186 in close collaboration with the research groups and technical partners involved to understand
187 their specific needs for storing and accessing sediment transport data. This participatory
188 process supported the development of a pragmatic and shared standardization strategy, aimed
189 at organizing heterogeneous datasets within a common relational structure while preserving
190 the traceability of original data sources, measurement methods, instruments, units, and
191 access conditions. The rationale behind this strategy was not to define a universal standard for
192 sediment transport data, but to provide a flexible and reproducible framework suitable for the
193 available datasets and adaptable to future contributions from different institutions or
194 monitoring traditions. The physical model of the database was implemented as PostgreSQL
195 database, while data entry, consultation, and updating were managed through a custom-
196 developed web application designed to provide a structured and user-friendly tool for data
197 management and integrated analysis.

198 Before being uploaded to the database, the datasets underwent a data-curation workflow
199 aimed at harmonising heterogeneous sources while preserving their original traceability.
200 Candidate datasets were selected according to their relevance to sediment transport or
201 associated hydrological and physicochemical variables, the availability of temporal
202 information, and the possibility of linking the observations to a river, monitoring station, and
203 watershed. Parameter names were harmonised using the internal database vocabulary, so that
204 equivalent terms from different sources could be stored under a common parameter definition.
205 Measurement units were checked against the units associated with each parameter in the
206 database; when necessary, values were converted to the adopted standard unit, while the
207 original source and methodological information were retained to preserve traceability. Missing
208 metadata, such as instrument type, measurement method, sampling frequency, or access
209 conditions, were not reconstructed when unavailable, but were explicitly recorded as missing
210 or unknown. Potential duplicated observations were checked using combinations of station,
211 parameter, date/time, value, and source information before upload. This workflow allowed
212 heterogeneous historical and recent datasets to be integrated into a common relational
213 structure while documenting the main limitations inherited from the original sources.

214 2.1 PostgreSQL Database

215 The database was developed in PostgreSQL and installed on a server machine accessible via
216 the HTTPS protocol (<https://lca.dst.unipi.it/>). PostgreSQL is an open-source relational
217 database management system (RDBMS), known for its adherence to SQL standards and
218 advanced features. It is designed to handle workloads ranging from small individual databases
219 to large-scale management systems, typical of complex applications. PostgreSQL has been
220 extended with the PostGIS library, which adds support for geospatial data, making it an
221 advanced tool for storing, managing, and analyzing spatial and geographic data. It complies

222 with the standards of the Open Geospatial Consortium (OGC), including the Simple Features
 223 for SQL and SQL/MM Spatial (Strobl, 2008).

224 The relational schema is shown in Figure 1. The information has been divided into 15 tables that
 225 allow for the storage of highly varied data.

226 The "watershed" table represents the highest level of the geographic system and describes the
 227 watersheds, which represent the fundamental territorial unit for monitoring sediment
 228 transport. This table includes spatial information (*geom*), enabling integration with GIS
 229 systems, and descriptive attributes for the 272 main watersheds of Italy
 230 ([http://sdi.isprambiente.it/geoserver/hy/bacini_principali/wfs?service=wfs&version=2.0.0&re](http://sdi.isprambiente.it/geoserver/hy/bacini_principali/wfs?service=wfs&version=2.0.0&request=GetCapabilities)
 231 [quest=GetCapabilities](http://sdi.isprambiente.it/geoserver/hy/bacini_principali/wfs?service=wfs&version=2.0.0&request=GetCapabilities)). Each watershed is linked to the corresponding river (in the "rivers"
 232 table). The "rivers" table represents the main watercourses that flow through the watersheds.
 233 Each river is associated with a watershed via the field *id_watershed*. This relationship allows
 234 rivers to be organized by their respective watersheds. The monitored rivers are described using
 235 a unique identifier (*id*) and a textual description (*description*), useful for identifying the river. The
 236 rivers serve as reference points for data acquisitions.

237 The "acquisition" table is the main repository for the data obtained from river monitoring. Each
 238 acquisition is geolocated using coordinates (*coordx*, *coordy* using the WGS84 coordinate
 239 reference system). The table also includes the monitored river (*id_river*), the measured
 240 parameter (*id_parameter*), the instrument used (*id_instrument*), the responsible user (*id_user*),
 241 additional information such as data availability (*availability*), specific reference (*reference*), and
 242 the monitoring section code (*section_code*). The acquisitions are also linked to a detail of the
 243 measurement type through *id_meas_type*.

244 The monitored parameters are defined in the "parameter" table. Each parameter describes a
 245 measurable element (e.g., pH, flow rate, velocity, suspended sediment concentration, etc.) and
 246 includes a textual description (*description*), a link to a general category through the field
 247 *id_categorize*, associated with the "macro_categories" table, which groups the parameters into
 248 broader categories (e.g., sediment concentration, river flow, chemical analysis, grain size, etc.).
 249 Parameters can be measured using different units of measurement and instruments, as
 250 described in the "parameter_has_unit" and "parameter_has_instruments" tables. The
 251 "instruments" table lists the tools used to collect field measurements. The
 252 "parameter_has_instruments" relationship establishes which instruments are used to
 253 measure each parameter. Each parameter is measured in a specific unit (e.g., °C, mg/L),
 254 defined in the "meas_unit" table. This n:n association is managed through the
 255 "parameter_has_unit" table.

256 Finally, the collected data are stored in the "data" table, which archives the actual numerical
 257 values recorded during acquisitions. Each record in the table includes the measured value
 258 (*value*), the collection date (*date*), and a reference to the specific acquisition via *id_acquisition*.

259 This approach allows for the direct connection of the measured data to the geographical
260 context, the monitored parameter, the instrument used, and the monitoring conditions.

261 Each data entry is linked to a responsible user (in the "users" table), which contains information
262 about the system users, including personal details (*name, surname, email*) and technical
263 information (*affiliation, associated schema, last login date*). It is related to the "user_type"
264 table, which categorizes users based on their role or authorization. Users can be active or
265 inactive (in the "activations" table) and can have different roles (editor or viewer) with varying
266 permissions on the database through the web application management system.

267 The database and the web application are hosted on dedicated server machines provided by
268 the University of Pisa and specifically configured for this project. This institutional hosting
269 solution allows the platform to operate without direct maintenance costs for individual
270 researchers and provides a more stable framework than project- or researcher-dependent
271 storage solutions.

272 User roles are assigned by the database administrators after registration. By default, newly
273 registered users are assigned a viewer role, allowing data consultation while preventing
274 uncontrolled modifications or imports. When a user intends to contribute new datasets, an
275 exchange with the database administrators is carried out during the account activation or role-
276 upgrading phase, in order to evaluate the user's affiliation, expertise, type of contribution, and
277 consistency of the proposed data with the database structure. Users recognized as data
278 contributors can then be assigned an editor role. This approach is intended to preserve data
279 integrity while keeping the platform open to future collaborations with institutions, research
280 groups, and monitoring agencies.

281 At the present stage, quality control is mainly implemented as a first-level validation based on
282 data traceability, controlled vocabularies, mandatory metadata fields, and consistency checks
283 before records are written to the database. For historical datasets, the quality of the archived
284 records necessarily depends on the completeness and reliability of the original sources and
285 associated metadata. Therefore, the database does not remove all uncertainties related to past
286 monitoring activities, but makes them explicit and traceable through the relational structure.
287 For newly measured data, the same controlled workflow represents a first level of quality
288 assurance, while the development of more detailed and documented quality-control
289 guidelines is identified as a priority for future updates of the platform.

290

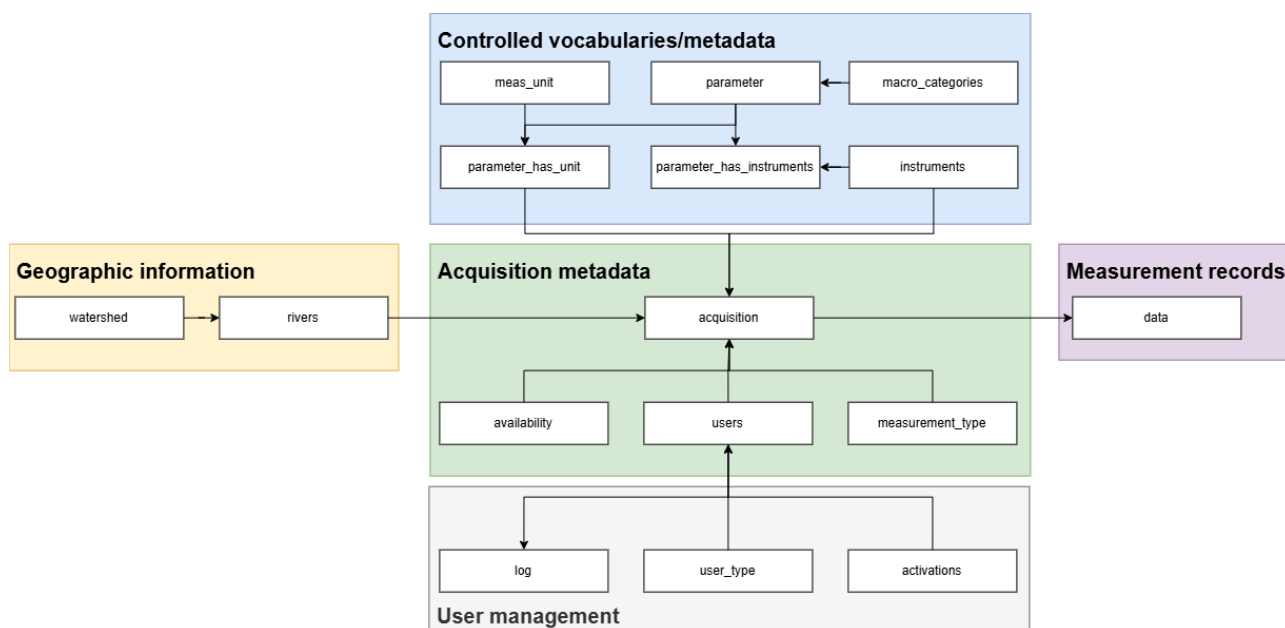


Figure 1 Relational schema of the database PostgreSQL

2.2 Web application

The web application (Figure 2; <https://lca.dst.unipi.it/SedimentTransport>) was developed to facilitate and control the database population and to provide tools for visualizing the data contained within it. It was built using open-source JavaScript and PHP libraries, primarily based on tools such as Bootstrap and OpenLayer. Access to the web application is granted after completing a registration form, during which the user selects their login credentials. Once the user is created, they must be approved by the database administrators, who will assign a user role (editor or viewer).

The application features a two-section layout (Figure 2): on the left, there is a menu with various options (Home, Import Data, View Data), while the main screen is displayed in the center. The homepage provides summary information such as the number of records in the database (“Data”), the number of measurement points (“Stations”), the number of rivers with at least one data entry (“Rivers with available data”), the number of active users in the last six months (“Active Users”), and the number of user affiliations (“Affiliations”). It should be noted that the web platform is continuously updated; therefore, the numbers displayed in the interface may differ from those reported in the Results, which refer to the frozen database version used for the analyses presented in this manuscript.

The homepage, which appears as shown Figure 2, also displays a map with the measurement points available in the database. Additionally, a table is shown summarizing the data in the database. The maps in the web application are interactive: when clicking on the layers, specific information about the measurement site is displayed.

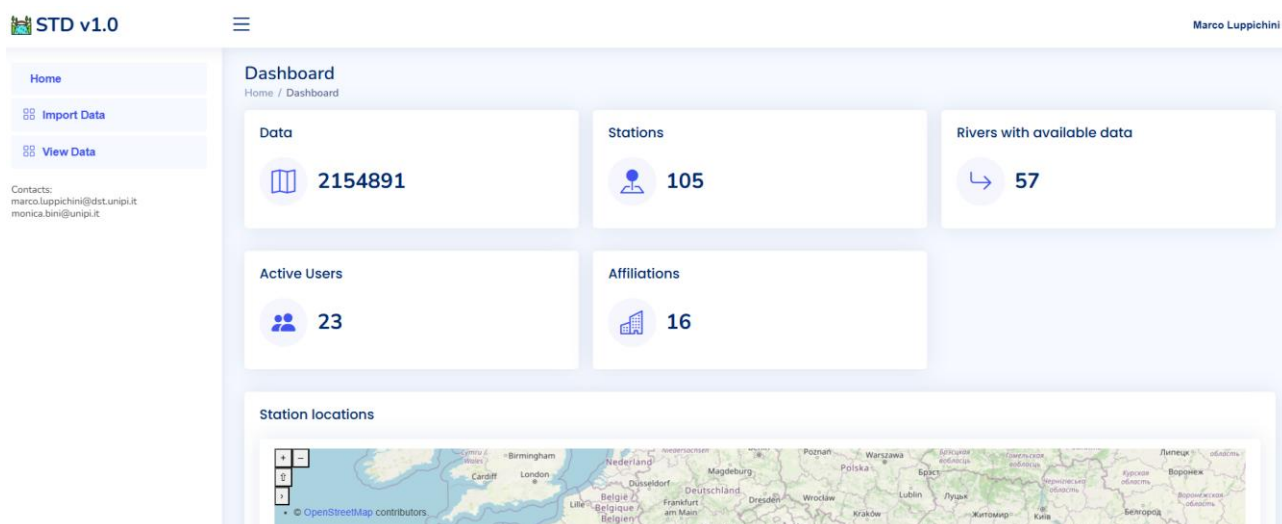


Figure 2 Home page of the web application (accessed on 16 February 2026). The web platform is continuously updated; therefore, the numbers shown in the interface may differ from those reported in the Results, which refer to the frozen database version used for the analyses presented in this manuscript.

The data entry process is carried out through the "Import Data" page. This feature is visible and accessible only to users with an editor role and consists of an initial form for general information to complete the acquisition table in the database (such as the river, measurement point, measured parameter, etc.). The dropdown menus are interconnected, and the available options change based on previous selections, guiding the user in selecting the correct entries.

The second part of the form consists of a tab with two options: "Insert Single Data" or "Insert Multi Data" (Figure 3). The "Insert Single Data" (Figure 3) allows the user to enter one record at a time by selecting the date and entering the corresponding value. The "Insert Multi Data" tab allows multiple values to be entered by uploading a Microsoft Excel file containing a specific template (the Excel file template is provided in the Supplementary Materials). The template consists of only two columns: "date" and "value". A data validation procedure is performed on the Excel file before writing data to the database. Any errors or successful operations are displayed to the user via appropriate alerts. The upload validation includes checks on the structure of the input file, the presence and format of the required fields, the readability of dates, the numeric type of measured values, and the consistency between the uploaded data and the acquisition metadata previously selected by the editor, including parameter, unit, river, monitoring station, and geographic coordinates. The procedure also prevents the upload of records with missing dates or missing values and flags potential duplicated observations based on the combination of acquisition, date, and value. These checks are intended to ensure database consistency and prevent structural errors during data entry.

For data consultation and download, the "View Data" page can be used, accessible from the left-hand menu of the application (Figure 4). The page contains a map showing the locations of the measurement stations. By clicking on the points, users can access information about the historical time series associated with the selected station.

343 The page also includes a table that lists all the information associated with each data record.
344 Through the provided form, users can query the data by applying filters, and the requested data
345 will be displayed in a table that can be further filtered.

346 Data in the table can be searched using the "Search" function or by applying column filters in
347 the last row of the table. Data can be exported in various formats using the "Copy", "Excel",
348 "CSV", "PDF", or "Print" options.

349 Data imported under a specific availability limit (on request) are only partially visible to users
350 who are not the data owner. However, users can request the full data by contacting the data
351 owner.

STD - beta version

Marco Luppichini

Home

Import Data

View Data

Data entry

Data / Import Data

Import

Activity: Bedload Transport

Category: Chemical Analysis

X coordinate [EPSG: 4326]:

Y coordinate [EPSG: 4326]:

Parameter: Major ions - Ca

Measurement Unit: mg/L

Instrument: Laboratory procedure

River: Adige River [ADIGE]

Section Code:

Location:

Availability: Open Access

Reference:

Note:

Insert Single Data Insert Multi Data

Date: gg/mm/aaaa -- Value: Insert Data

352
353 Figure 3 Import data page of the web application (accessed on 24 October 2025)

354

View Data

Main Watershed*: SERCHIO

Parameter*: Flow rate

Start Date: gg/mm/aaaa

End Date: gg/mm/aaaa

Availability*: On Request, Open Access

Search Data

Copy Excel CSV PDF Print

25 entries per page

Search:

TYPE	RIVER	INSTRUMENT	PARAMETER	UNIT	LOCATION	SECTION CODE	WATERSHED	DATE	VALUE	AVAILABILITY	USER
►River parameters	Serchio	Unknown	Flow rate	m3/s	Vecchiano		SERCHIO	2021-12-15 00:00:00	41.926	Open Access	Marco Luppichini
►River parameters	Serchio	Unknown	Flow rate	m3/s	Vecchiano		SERCHIO	2005-07-12 00:00:00	28.313	Open Access	Marco Luppichini
►River	Serchio	Unknown	Flow rate	m3/s	Vecchiano		SERCHIO	2009-11-28	56.077	Open Access	Marco

355
356 Figure 4 View data page of the web application (accessed on 24 October 2025)

3 Results

Before describing the content of the data records, we provide a spatial overview of the monitoring network included in the database. Figure 5 shows the location of the monitoring stations together with the corresponding river network and watershed boundaries. This representation allows readers to evaluate the geographical distribution of the available observations and to identify the main monitored river basins across the Italian territory.

The monitoring stations are not uniformly distributed across the country, reflecting both the availability of regional monitoring programmes and the heterogeneous development of sediment transport observations in Italy. A higher density of stations is evident in northern and central Italy, particularly in Alpine and Apennine river systems, where several monitored basins are represented. The map also highlights the connection between individual monitoring points and their upstream drainage areas, providing useful spatial context for the interpretation of the database.

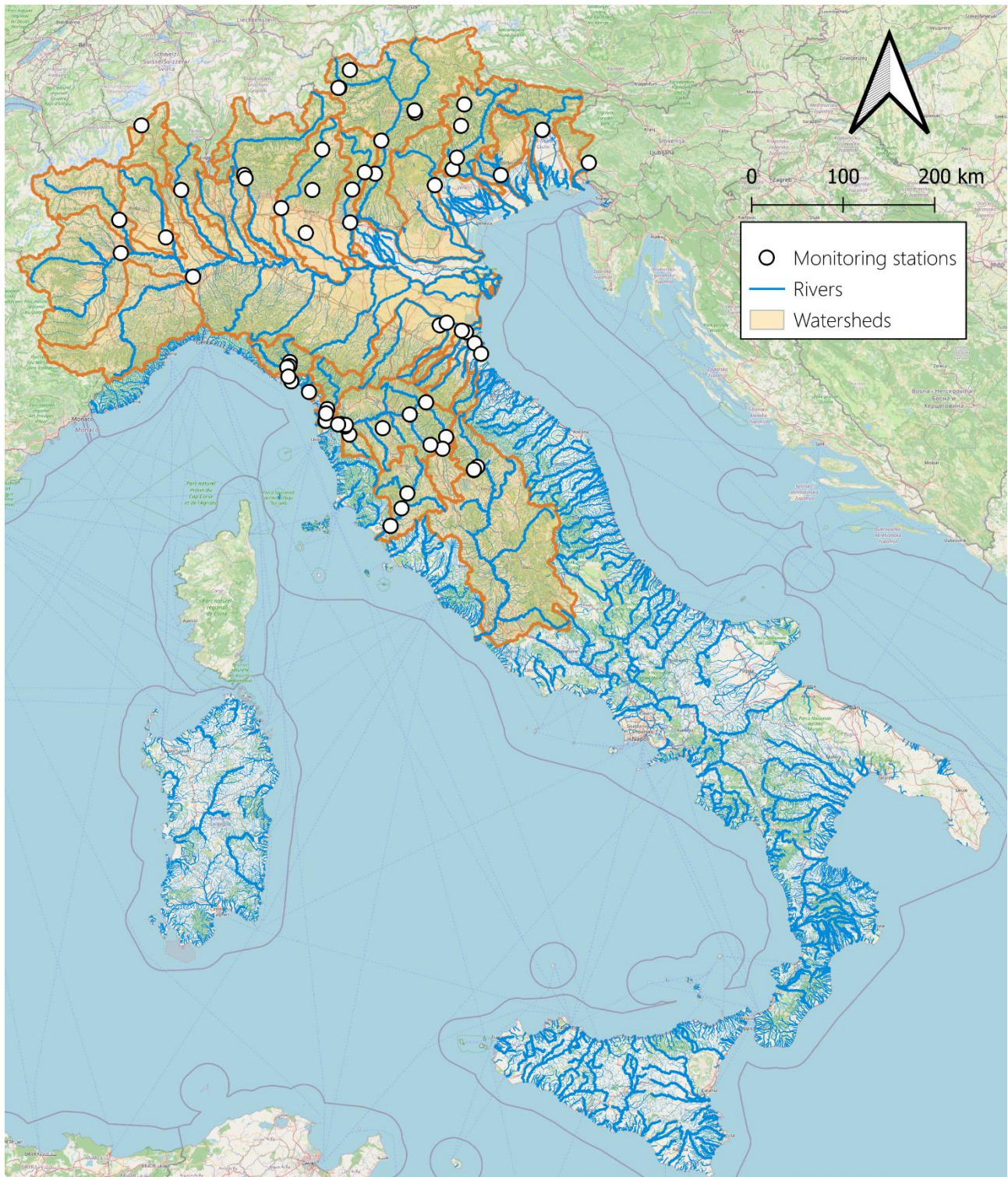


Figure 5 Spatial distribution of the monitoring stations included in the database. White circles indicate monitoring stations, blue lines represent the river network, and orange outlines delimit the corresponding watersheds. The map provides an overview of the monitored river basins and their distribution across the Italian territory.

3.1 Data Records

The data collected within the database amount to 1,988,116 records and are characterized by a high degree of spatial, temporal, and instrumental heterogeneity. The database currently

includes information from 29 river basins and 39 rivers, distributed across 76 measurement sites. A total of 386 time series are archived. The dataset covers 21 different parameters, acquired using 19 different types of instruments, highlighting the diversity of monitoring approaches and measurement techniques adopted across different hydrological contexts.

To provide a concise overview of the database contents before the detailed analysis of instruments, sampling frequency, temporal coverage, and spatial distribution, *Table 1* summarises the parameters currently populated in the database. For each parameter and measurement unit, the table reports the number of available records, the number of time series, the number of rivers and measurement sites represented, and the overall temporal coverage. *Table 1* highlights the heterogeneous structure of the database, both in terms of data volume and temporal extent. Some variables, such as flow height, flow rate, temperature, turbidity, and electrical conductivity, are represented by large numbers of records, whereas several sediment-related parameters, including annual bedload discharge, annual suspended sediment load, bedload sediment load, and grain-size descriptors, are represented by smaller datasets. This summary provides a compact reference for interpreting the following figures, which analyse instrumentation, sampling frequency, temporal coverage, and spatial distribution in greater detail.

Figure 6 shows the percentage distribution of the instruments used to acquire each parameter in the database, expressed as the proportion of total records available for that parameter. The analysis highlights that, in some cases, the instrumentation employed appears to be extremely homogeneous, while in others a considerable heterogeneity emerges, related to the nature of the parameter and the different measurement methodologies available.

Parameters closely linked to basic physico-chemical measurements, such as electrical conductivity and pH, show a clear predominance of a specific instrument (the *conductivity meter* and the *pH meter*, respectively), with percentages close to 100%. A similar pattern can be observed for parameters derived from annual aggregations, such as Annual Bedload Discharge (ABD), which are exclusively associated with specific calculation methods and measuring instruments (e.g., *geophones*).

Conversely, parameters related to instantaneous or monthly sediment transport and hydraulic flow characteristics show greater variability in the instruments used. In particular, the Suspended Sediment Concentration (SSC) was acquired using four different types of instruments (*bottles*, *DH59*, and *acoustic doppler current profilers*). This heterogeneity reflects the coexistence, within the various monitoring networks, of both direct and indirect approaches to measuring suspended sediment concentration, as well as operational differences between monitoring sites and acquisition periods.

A similar situation can be observed for turbidity, which is measured using optical instruments (e.g., *nephelometers* and *turbidimeters*) but with varying percentages, suggesting differences in both the temporal resolution of the observations and the technological availability of the monitoring sites. Hydraulic parameters such as flow height and flow rate also display a more

417 fragmented distribution, with an alternative use of *radar sensors*, *pressure transducers*, and
 418 *float gauges*, consistent with the diversity of environmental conditions and the differing
 419 operational choices among monitoring stations.

420 Overall, there is a clear distinction between parameters that benefit from standardized
 421 measurement procedures and dedicated instruments and more complex parameters for which
 422 the variety of methodological approaches and operational tools constitutes an intrinsic feature
 423 of the acquisition process. This differentiation underscores the importance of a database
 424 structure capable of accommodating heterogeneous information while retaining details of the
 425 instrumental and methodological context associated with each dataset, thereby ensuring
 426 traceability and comparability among data derived from different sources.

427 It is also noteworthy that seven out of the sixteen analyzed parameters show significant
 428 percentages of data associated with an instrument classified as “*unknown*”. Data associated
 429 with an “unknown” instrument should therefore be interpreted with caution. This classification
 430 does not necessarily indicate low-quality or unreliable measurements, as many of these
 431 records may derive from historical sources or previously processed datasets considered valid
 432 by the original data providers. However, the absence of information on the measurement
 433 instrument limits the possibility of assessing methodological comparability, instrumental
 434 uncertainty, and potential biases among datasets. Consequently, these records can still be
 435 useful for exploratory analyses, historical reconstructions, and broad assessments of spatial
 436 and temporal data availability, but they are less suitable for detailed comparisons among
 437 measurement techniques, uncertainty analyses, or model calibration procedures requiring
 438 fully documented metadata. In this sense, the database does not eliminate the limitations
 439 inherited from the original data sources; rather, it makes them explicit and traceable, providing
 440 a basis for future metadata improvement and data-quality refinement.

441 *Table 1 Summary of the parameters currently populated in the Italian Fluvial Sediment Transport Database. For each parameter*
 442 *and measurement unit, the table reports the number of records, number of time series, number of rivers, number of*
 443 *measurement sites, and temporal coverage. The table provides a compact overview of the database contents and*
 444 *complements the detailed analyses of instruments, sampling frequency, temporal coverage, and spatial distribution shown in*
 445 *the following figures.*

Parameter	Measurement Unit	No. of Records	No. of time series	No. of Rivers	No. of measurement sites	Start Date	End Date
Annual Bedload Discharge (ABD)	t/yr	6	1	1	1	2014-01-01	2020-01-01
Annual Suspended Sediment Load (ASSL)	t/yr	6	1	1	1	2014-01-01	2020-01-01
Bedload discharge	t/day	56	9	6	6	2003-11-03	2019-12-03
Bedload Sediment Load (BSL)	kg/s	159	7	7	7	1987-01-01	2012-11-05
Electrical potential	mV	34	23	15	22	2011-02-09	2011-12-10
Electric conductivity	S/cm	1	1	1	1	2011-04-23	2011-04-23
Electric conductivity	μS/cm	154686	33	19	27	2011-02-09	2025-03-21

Flow height	m	680983	20	16	16	2003-11-03	2025-08-07
Flow rate	kg/s	1	1	1	1	2023-03-14	2023-03-14
Flow rate	m3/s	145562	51	32	43	1924-01-01	2023-10-18
Grain size - D50	mm	38	4	3	3	2005-04-12	2019-12-03
Grain size - D90	mm	36	4	3	3	2005-04-12	2019-12-03
Major Ions	mg/L	384	27	16	23	2011-02-09	2023-03-31
Monthly Suspended Sediment Load (MSSL)	10^3 tonn	2719	8	5	8	1936-01-01	1995-12-01
Monthly Suspended Sediment Load (MSSL)	Max kg/m3	2770	8	5	8	1931-01-01	2011-02-09
Monthly Suspended Sediment Load (MSSL)	Max kg/s	2709	8	5	8	1936-01-01	1995-12-01
Monthly Suspended Sediment Load (MSSL)	Mean kg/m3	2718	8	5	8	1936-01-01	1995-12-01
Monthly Suspended Sediment Load (MSSL)	Mean kg/s	2718	8	5	8	1936-01-01	1995-12-01
Monthly Suspended Sediment Load (MSSL)	Min kg/m3	2768	8	5	8	1931-01-01	1995-12-01
Monthly Suspended Sediment Load (MSSL)	Min kg/s	2709	8	5	8	1936-01-01	1995-12-01
Monthly Suspended Sediment Load (MSSL)	t/km2	2718	8	5	8	1936-01-01	1995-12-01
pH	NaN	62	28	16	23	2011-02-09	2023-03-31
Reanalysis SSL	kg/s	8860	10	2	2	1924-01-01	2024-10-01
Suspended Sediment Concentration (SSC)	g	70911	2	1	1	2020-01-09	2023-09-20
Suspended Sediment Concentration (SSC)	g/L	94452	25	9	20	2003-11-03	2025-04-16
Suspended Sediment Load (SSL)	kg/s	46	5	5	5	1998-12-22	2012-11-05
Suspended Sediment Load (SSL)	t/yr	94288	2	1	1	2020-01-09	2023-09-20
Temperature	°C	323950	35	20	29	2011-02-09	2025-08-07
Total Sediment Load (TSL)	kg/s	46	5	5	5	1998-12-22	2012-11-05
Turbidity	NTU	557906	10	6	7	1987-01-01	2025-08-07
Width of bedload zone (Wbl)	m	52	7	4	4	2003-11-03	2019-12-03
Width of suspended load zone (Wsb)	m	58	9	6	6	2003-11-03	2019-12-03

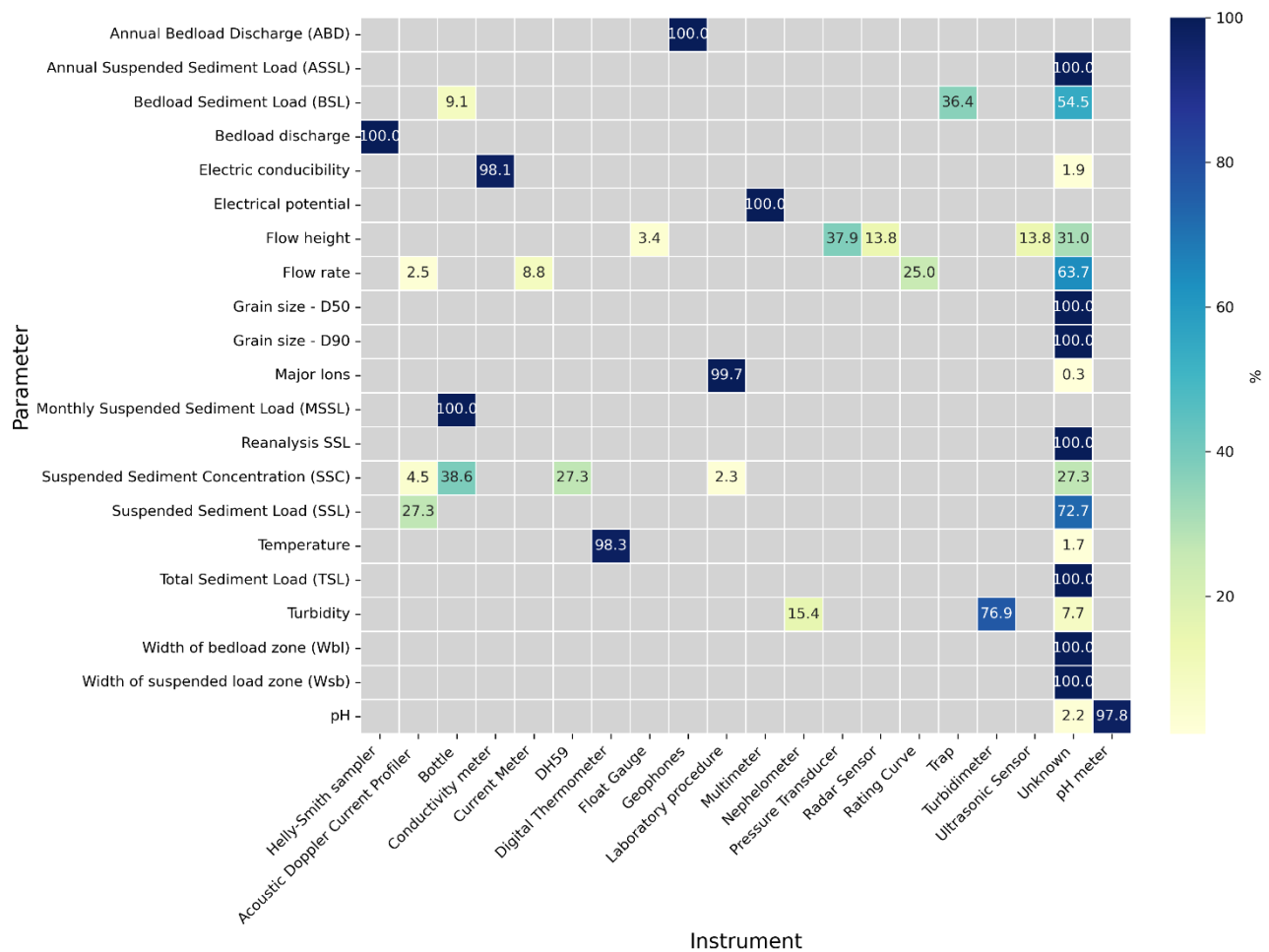


Figure 6 Frequency distribution of instruments used for the acquisition of each parameter stored in the database. Grey cells indicate the absence of values (0%). The figure highlights both homogeneous cases, where a single instrument is dominant, and heterogeneous situations, particularly for sediment transport and hydraulic parameters, where multiple measurement techniques coexist.

Figure 7 summarizes the typical sampling frequencies of the datasets stored in the database. For each parameter, the modal temporal resolution (i.e., the most frequently observed sampling interval) was identified and compared with the full range of observed frequencies, spanning from minimum to maximum values. The cumulative distribution further quantifies the proportion of records associated with each temporal resolution class.

The distribution highlights marked heterogeneity in the temporal resolution of the data, spanning sub-hourly to annual frequencies. Some parameters, particularly those of a physical or hydraulic nature (e.g., flow height), are acquired at high frequencies (2 minutes or 1 hour), consistent with continuous or semi-continuous monitoring activities. Other datasets, especially those related to sediment transport or derived parameters, show daily or monthly frequencies, typical of manual field measurements or periodic sampling associated with laboratory procedures.

The cumulative curve reveals five main acquisition regimes, corresponding to frequencies of 10 minutes, 15 minutes, 1 hour, 1 day, and 1 month, which represent the prevailing operational scales within the monitoring networks considered. The most pronounced steps in the curve

467 occur at these classes, indicating that a substantial portion of the total data is concentrated in
468 these temporal intervals.

469 It is also evident that there is a wide variability in the effective acquisition frequency. This
470 behavior reflects the presence of irregular or discontinuous time series, where the nominal
471 frequency does not always correspond to actual data continuity (for example, hourly series
472 characterized by interruptions or inactive periods). This variability underscores the complexity
473 of harmonizing heterogeneous datasets and highlights the importance of a database structure
474 capable of handling both the nominal sampling frequency and its actual temporal regularity.

475 Overall, the figure illustrates how the collected data integrate multi-temporal information,
476 combining continuous acquisitions with periodic or event-based sampling. This temporal
477 diversity enables a more comprehensive representation of sediment and hydraulic dynamics
478 across different spatial and temporal scales.

479

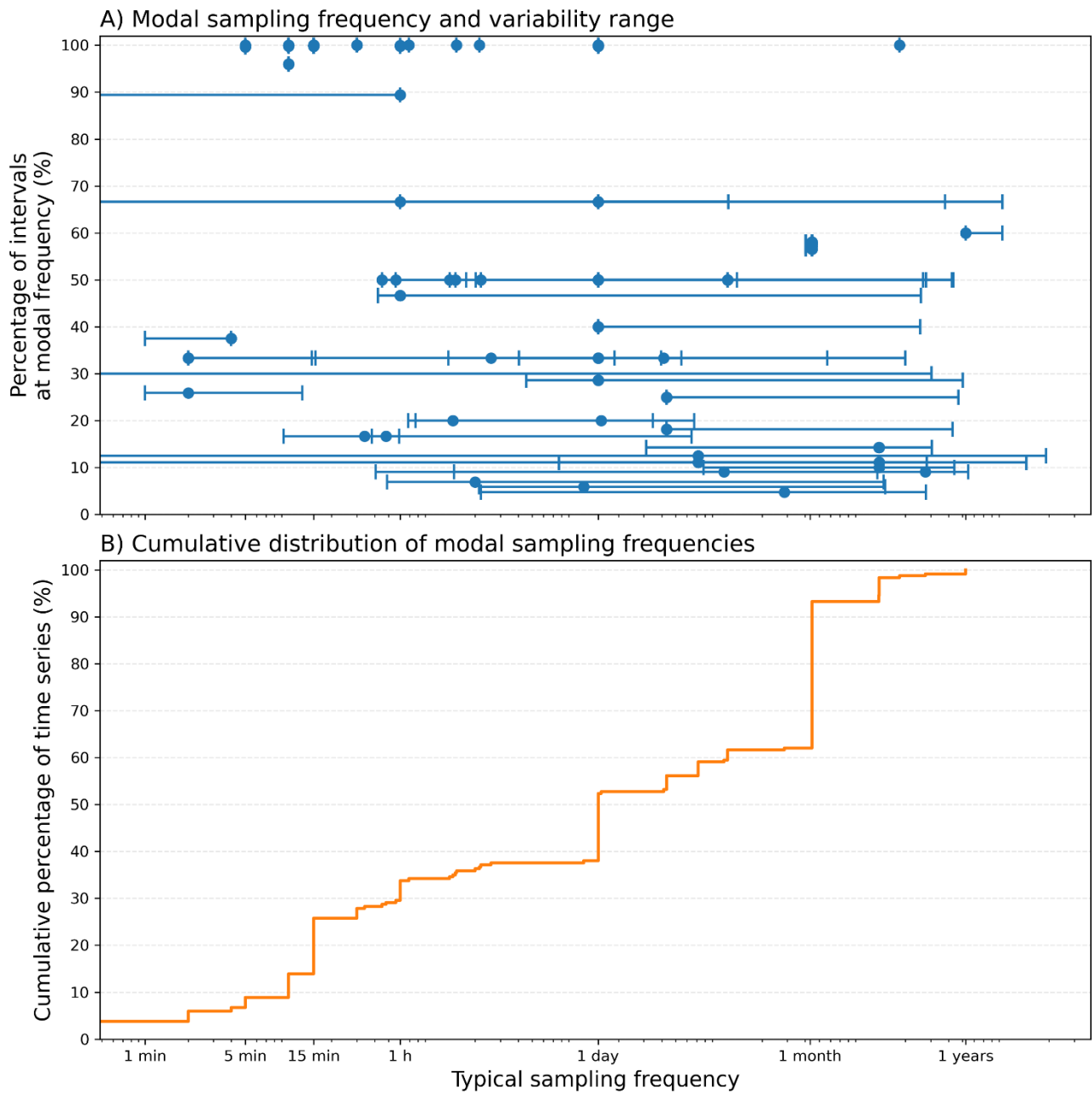


Figure 7 Sampling-frequency characteristics of the time series included in the database. **(A)** Modal sampling frequency of each time series, expressed as the percentage of intervals corresponding to the most frequent temporal resolution; horizontal bars indicate the variability range of observed sampling intervals. **(B)** Cumulative distribution of modal sampling frequencies, showing the proportion of time series acquired at temporal resolutions equal to or finer than each frequency class. The figure highlights the coexistence of high-frequency, daily, monthly, and lower-frequency records within the database.

Figure 8 represents the annual temporal coverage of the data available for the parameters included in the database. For each parameter, the number of active time series was calculated for each year, where an active time series is defined as a series with at least one available record in that year. The total number of time series available for each parameter is reported in square brackets to contextualise the extent of the dataset.. It should be noted that the bars represent annual data availability and do not imply homogeneous monitoring frequency within or among years.

493 In this sense, Figure 8 should be interpreted together with the monitoring-frequency analysis
494 reported in Figure 7, which specifically describes the typical temporal resolution of the records
495 included in the database.

496 The analysis reveals that the temporal extent of the datasets varies widely across parameters,
497 reflecting the diversity of monitoring networks and measurement techniques adopted over
498 time. Some parameters exhibit very long time series, often beginning in the 1940s or 1950s and
499 continuing up to the present day. This is particularly evident for flow rate, which shows the
500 largest number of series and the longest temporal span, making it one of the most consistently
501 monitored variables in terms of temporal coverage. However, the frequency of observation may
502 vary substantially across stations, parameters, and periods. Older records, in particular, may
503 be less frequent, less regular, or less standardised than more recent datasets, depending on
504 the original monitoring programme and data source. Other hydrological parameters, such as
505 flow height and Monthly Suspended Sediment Load (MSSL), also display extended observation
506 periods and numerous records, partly due to the long-standing tradition of monitoring
507 discharge and sediment dynamics in some of the major Italian river basins (e.g., Arno and
508 Ombrone rivers).

509 Conversely, parameters of a more specific nature or associated with newer measurement
510 technologies—such as turbidity, electrical conductivity, pH, and major ions—show shorter and
511 more discontinuous time series, often limited to the past two decades. This reflects both the
512 recent introduction of dedicated automatic sensors and the historical heterogeneity in the
513 availability of physicochemical data.

514 Parameters related to sediment transport also exhibit significant variability. While annual
515 datasets such as ABD and Annual Suspended Sediment Load (ASSL) are generally shorter and
516 concentrated within specific periods, others—such as Bedload Sediment Load (BSL) and
517 SSC—show more fragmented temporal coverage, alternating between phases of intensive
518 monitoring and intervals without data acquisition. In several high-mountain catchments,
519 suspended sediment monitoring is often seasonal, as sensors are removed during winter
520 months (typically from November to May) due to harsh environmental conditions. This pattern
521 reflects not only logistical constraints but also the strong morphoclimatic contrasts that
522 characterise Italy, where monitoring strategies are necessarily adapted to markedly different
523 environmental settings.

524 Overall, the figure highlights the heterogeneous temporal coverage among the different types
525 of data: some measurements, such as flow rate or flow height, provide long temporal records,
526 while others, more specialised, offer detailed but discontinuous insights into fluvial dynamics
527 and sediment transport processes. This heterogeneity, and particularly the lack of
528 simultaneous measurements among some parameters and monitoring stations, may limit
529 integrated analyses of hydrological, physicochemical, and sediment transport processes.
530 Therefore, Figure 8 should be interpreted as an aggregated descriptive overview of temporal
531 data availability rather than as a direct measure of uniform monitoring effort or sampling

frequency through time. Differences in temporal coverage and monitoring frequency are an inherent characteristic of the compiled database and are further documented by the monitoring-frequency analysis shown in Figure 7. These aspects should therefore be considered when using the dataset for comparative, temporal, integrated, or trend-based analyses.



Figure 8 Annual temporal coverage of the parameters included in the database. Each panel refers to one parameter, and the bars indicate the number of active time series for each year, defined as time series with at least one available record in that year.

542 Figure 9 summarizes the spatial distribution of these data across the analyzed river basins. The
543 four panels respectively show: the number of rivers for which observations are available (top
544 left), the total number data records (top right), the number of monitored parameters (bottom
545 left), and the total number of measurement points in each basin (bottom right). The database
546 currently includes data from 29 river basins out of the 272 that cover the entire Italian
547 peninsula. At present, the dataset is primarily populated with basins from northern and central
548 Italy, while southern regions and the major islands remain underrepresented.

549 The analysis reveals a notable heterogeneity in spatial coverage and data density. Some basins,
550 particularly those in northern Italy, such as the Po, Adige, and Tagliamento, show a higher
551 concentration of information, both in terms of the number of measurement points and the
552 range of parameters observed. In these cases, monitoring networks are more complex and
553 include several rivers and sub-basins, reflecting a stronger historical and institutional focus on
554 sediment transport and fluvial dynamics monitoring.

555 Conversely, the basins of central Italy generally have fewer observation points and monitored
556 parameters. In these areas, monitoring activities are more sporadic or limited to short periods,
557 often associated with specific research projects or local management initiatives.

558 The panel showing the total number of data also highlights that the amount of available
559 information depends not only on the number of monitoring sites but also on the temporal
560 frequency of observations. Basins with fewer stations but high-frequency monitoring can,
561 indeed, provide data volumes comparable to those of larger networks with lower-frequency
562 time series.

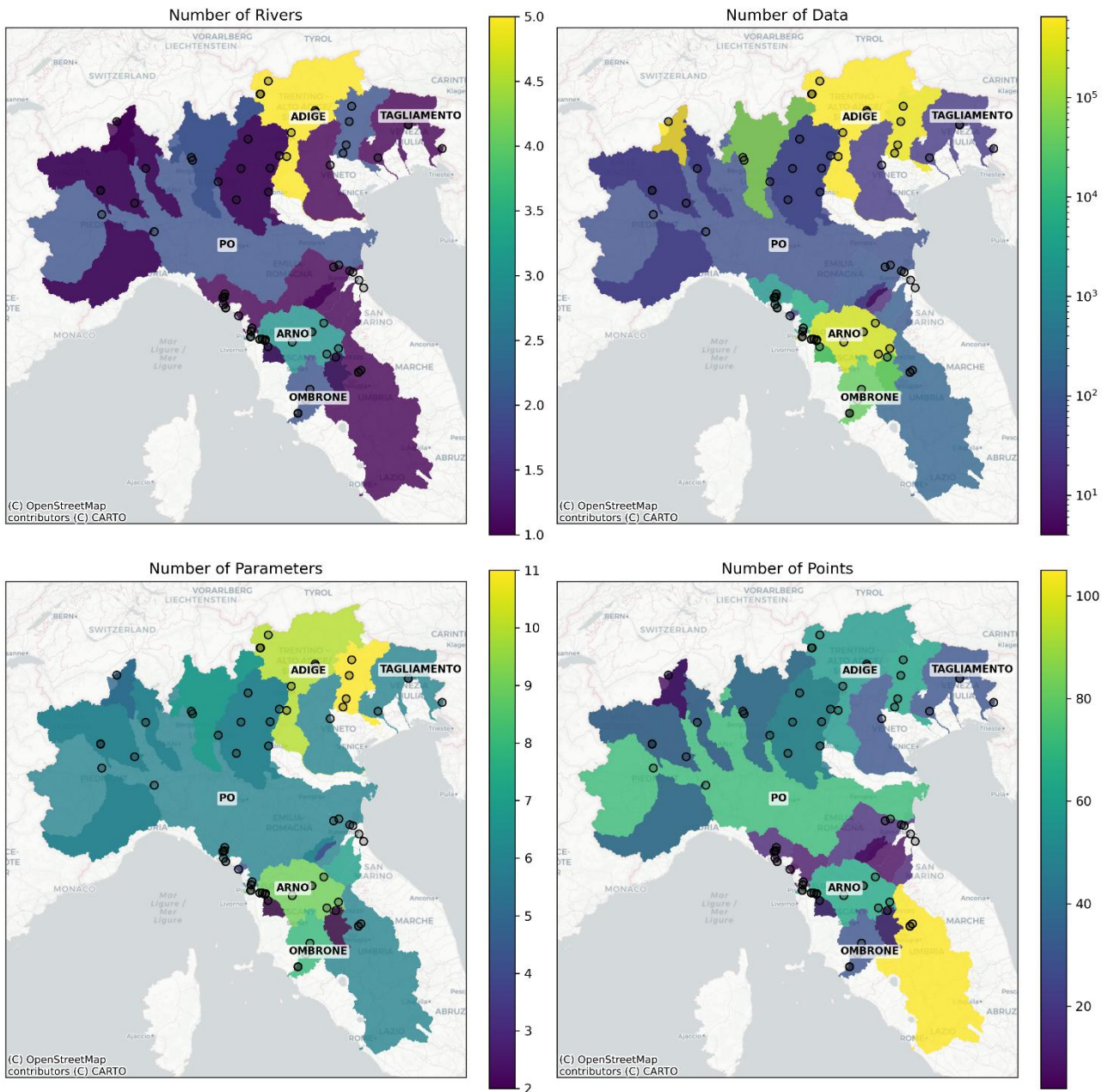


Figure 9 Distribution of sediment transport data by river basin. The panels show (clockwise from top left): the number of rivers analyzed, the total number of recorded data, the number of measurement points, and the number of monitored parameters for each basin. The figure highlights the strong spatial variability in data coverage and density, with higher concentrations in northern Italy compared to central and southern basins. The points indicate the location of the measurement sites.

3.2 Sediment load parameters: basin-scale analysis

The analysis presented in this section focused exclusively on data related to parameters directly connected to sediment transport, including both suspended and bedload components. Specifically, the following parameters were considered: *suspended sediment concentration deviation (SSCD)*, *suspended sediment load (SSL)*, *suspended sediment concentration (SSC)*, *monthly suspended sediment load (MSSL)*, *annual suspended sediment*

575 load (ASSL), annual bedload discharge (ABD), bedload sediment load (BSL), total sediment
576 load (TSL), total suspended solids (TSS), turbidity, grain size – D50 and grain size – D90.

577 *Figure 10* shows the actual distribution of sediment transport data, distinguishing between
578 suspended and bedload components. In the top-left panel, river basins highlighted in red
579 correspond to those for which the database contains at least one sediment transport
580 parameter (*basins with sediment data*), while the remaining basins, outlined in black, are
581 included in the database but associated only with other types of parameters (hydrological,
582 physico-chemical, etc.).

583 The other panels provide a detailed view of the number of available data for each specific
584 parameter within every river basin, allowing the identification of the most monitored areas. The
585 SSL and SSC are the most commonly represented parameters, with extensive data in the river
586 basins of northern Italy (particularly the Piave) and central Italy. The MSSL, on the other hand,
587 shows a more localized distribution, with significant values in a few basins, mainly in Tuscany.

588 Annual parameters such as ASSL and ABD show data concentrated in a single northern basin,
589 reflecting the more complex and resource-demanding nature of these measurements. The BSL
590 shows a similar pattern, with higher values in the Arno basin (central Italy), where direct
591 measurements of bedload transport are available.

592 The TSL is present only in a limited number of basins, generally derived from the combination
593 of suspended and bedload monitoring data. Finally, turbidity emerges as one of the most
594 widespread parameters in the central and northern basins, due to the greater availability of
595 optical sensors and its direct correlation with suspended sediment concentration.

596

597

598

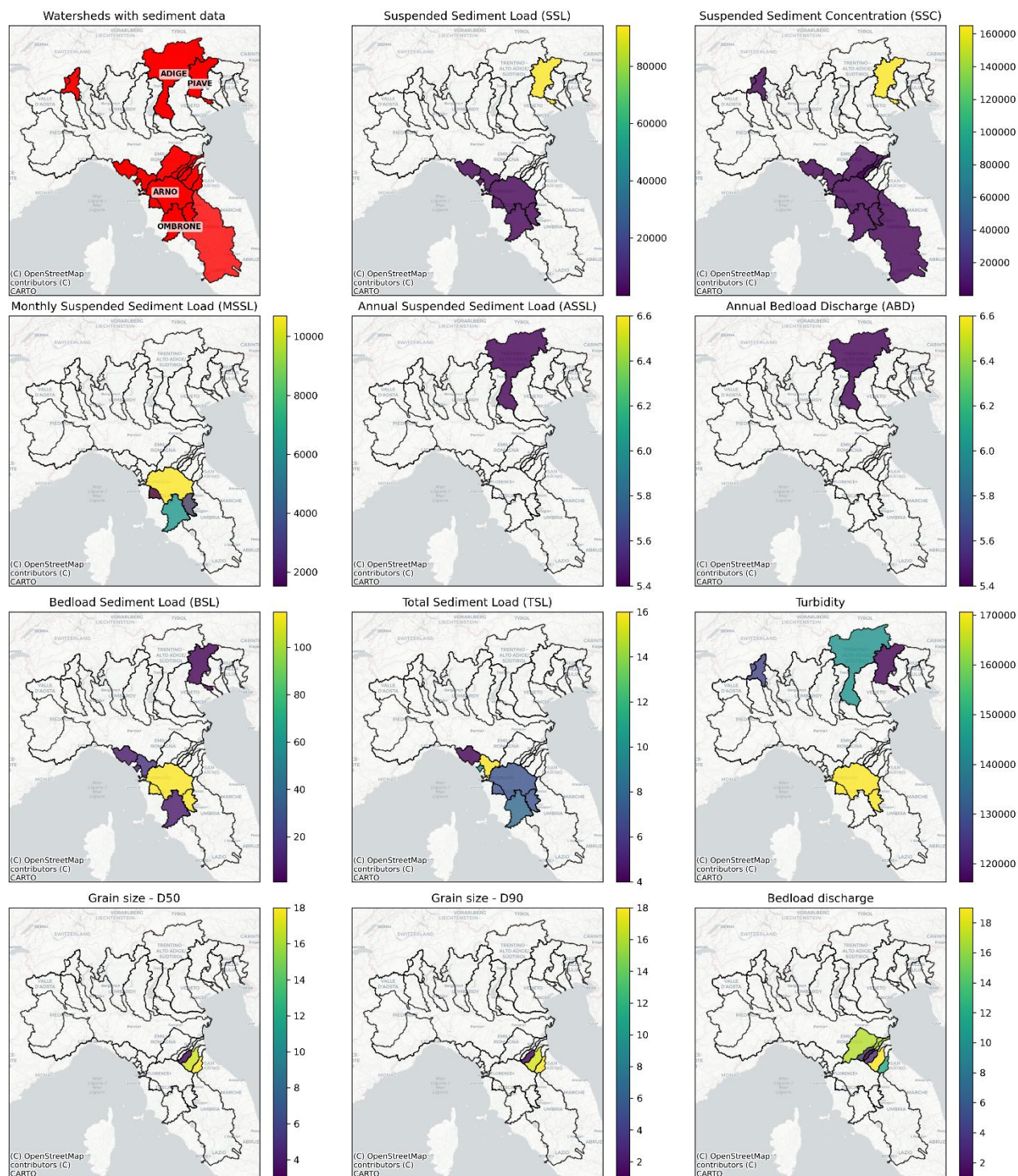


Figure 10 Spatial distribution of sediment transport data for each parameter considered. In the top-left panel, red basins indicate watersheds containing at least one sediment transport parameter, while the remaining basins, outlined in black, are included in the database but associated only with other types of parameters. The other panels show, for each sediment-related parameter, the number of available records within each basin. Colour scales are panel-specific and should not be directly compared across parameters; in each panel, colours represent the number of available records for the corresponding parameter.

4 Discussion and Conclusions

The aim of this study was to develop a relational database and a web interface dedicated to the storage, consultation, analysis, and export of fluvial sediment transport data. This initiative responds to the growing need in geosciences and environmental management for information systems capable of integrating, standardizing, and making accessible heterogeneous datasets that are historically fragmented and dispersed (Bailo et al., 2023; Parsons & Fox, 2013; Vitolo et al., 2015). This approach is consistent with European directives and policies on environmental and public-sector data sharing, particularly the INSPIRE Directive 2007/2/EC and the Open Data Directive (EU) 2019/1024, as well as with international open data practices. These frameworks emphasise that interoperability, accessibility, reusability, and metadata completeness are fundamental elements for the effective use and enhancement of hydrological and sedimentological data (Cho & Cromptoets, 2019; Jacob et al., 2025; Matthews et al., 2023). The relevance of such an integrated information system also extends to the implementation of the Water Framework Directive (2000/60/EC). The Common Implementation Strategy guidance on integrated sediment management explicitly recognises sediment quantity, continuity, and quality as important components of river basin management and stresses the need for appropriate information to support their assessment within River Basin Management Plans (European Commission, 2022). In this context, a harmonised and accessible sediment transport database can provide a valuable knowledge base for assessing sediment dynamics, identifying monitoring gaps, and supporting integrated sediment management at the river-basin scale.

In this context, the comparison with established hydrological infrastructures such as GRDC (Färber et al., 2025) and GSIM (Do et al., 2018) highlights both the relevance and the current gap in sediment data integration. While streamflow monitoring has progressively evolved toward coordinated international frameworks supported by standardized metadata protocols and long-term institutional cooperation, sediment transport data have not yet reached a comparable level of structural organization. This disparity is not only technical but also institutional, reflecting differences in monitoring continuity, data ownership, and harmonization practices. The database developed in this study contributes to reducing this gap at the national scale by providing a structured environment for data standardization and controlled access. Although it does not yet represent a fully institutionalized infrastructure comparable to large international repositories, it establishes the foundational architecture upon which broader coordination and future integration efforts can be built.

The results, showing the data distribution in the time interval 1924-2025, revealed marked variability in the instrumentation used to measure and monitor the different parameters. Some parameters, such as electrical conductivity or pH, are characterized by nearly uniform instrumentation, reflecting their standardized measurement methods. Conversely, more complex parameters such as Suspended Sediment Concentration (SSC) or Turbidity, show a wide diversification of instruments, indicating the coexistence of direct and indirect

645 measurement approaches. This heterogeneity, while limiting the direct comparability of data,
646 also constitutes an informative resource, as it reflects the diversity of local characteristics and
647 operational contexts in which the monitoring activities were carried out (Fratianni & Acquavota,
648 2017; Fredi & Lupia Palmieri, 2017; Wang et al., 2022).

649 However, the presence of data without information on the measuring instrument ("unknown
650 instrument") highlights the importance of traceability and metadata completeness, a topic
651 already emphasized in the sediment monitoring literature (Walling, 2006).

652 The analysis of sampling frequencies and the temporal extent of time series revealed
653 considerable temporal heterogeneity in the datasets. Sampling frequencies range from sub-
654 hourly to annual measurements, while the duration of the series varies from a few years to more
655 than seven decades. This discontinuity limits the possibility of conducting consistent and
656 comparable temporal analyses and represents one of the most common challenges in
657 managing long-term environmental data (Ellingsen et al., 2017; Nejad et al., 2024; Syvitski et
658 al., 2005). It is therefore necessary to distinguish between nominal and effective acquisition
659 frequency, document the actual regularity and completeness of the series, and develop
660 automatic quality control procedures. In this context, the proposed database supports
661 preliminary data-quality assessment by allowing users to rapidly visualise temporal coverage,
662 acquisition frequency, and data gaps. These diagnostics help identify discontinuities, irregular
663 sampling intervals, and periods requiring further verification, thereby providing a basis for
664 documenting the effective completeness of each time series and guiding subsequent quality-
665 control procedures.

666 From a spatial perspective, the basin-scale analyses showed a high concentration of data in
667 Alpine river basins, particularly the Adige (Bonfrisco et al., 2025) and Piave (Pellegrini et al.,
668 2023b; Rainato et al., 2017, 2025b). Central Italian basins are less represented, although data
669 are available for some smaller river basins, such as the Versilia River (Francalanci et al., 2013b).
670 Southern Italian basins are almost entirely absent from the database. (Bonfrisco et al.,
671 2025)(Pellegrini et al., 2023b; Rainato et al., 2017, 2025b)(Francalanci et al., 2013b)This
672 distribution reflects the historical development of monitoring networks, primarily established
673 in areas with higher anthropogenic pressure and hydrogeological risk, but it also highlights the
674 need for a more balanced territorial coverage. Parameters such as SSC and Turbidity are
675 widespread and well-documented, whereas others, such as BSL or TSL, show limited coverage,
676 being measured only at a few locations. This imbalance highlights the technical and
677 operational complexity of bedload measurements and also suggests the need for strategies
678 that combine direct and indirect approaches to broaden the available dataset. It should be
679 noted, however, that the apparent absence of data in southern Italy may reflect both a lower
680 historical monitoring intensity and a limited response to the present data call. Therefore, the
681 current spatial pattern should not be interpreted as a complete lack of sediment monitoring
682 activities in these regions, but rather as an indication of the present state of data collection
683 within the framework of this database. Further engagement with regional institutions may help
684 improve territorial representativeness in future updates. In addition, the database is currently

685 dominated by datasets derived from research initiatives, while data routinely collected by
686 regional and national environmental agencies are still only partially represented. This likely
687 results in a more fragmented picture than the actual potential availability of sediment
688 monitoring data at the national scale, further underscoring the need for stronger institutional
689 collaboration in future database expansions.

690 The experience gained in building the database made it possible to identify both the limitations
691 and the potential of the data collection and integration process. Among the main limitations
692 are the difficulty of harmonising pre-existing formats, the lack of uniform metadata, the need
693 for manual data validation procedures, and the dependence on voluntary contributions from
694 individual researchers and institutions conducting monitoring activities. This latter aspect may
695 affect both the completeness and representativeness of the dataset and also raises broader
696 issues related to data responsibility, long-term maintenance, and institutional accountability.

697 To address these limitations, future development of the database will focus on several practical
698 steps. First, a minimum metadata template will be defined for new contributions, including
699 information on measurement methods, instruments, units, sampling frequency, temporal
700 coverage, uncertainty, and monitoring conditions. Second, semi-automatic validation
701 procedures will be implemented to identify missing metadata, inconsistent units, duplicated
702 records, anomalous values, and temporal gaps before data are integrated into the database.
703 Third, the platform will be progressively used as an operational reference for data producers,
704 helping them identify data gaps and harmonise future monitoring activities before new field
705 campaigns are planned.

706 However, technical solutions alone are not sufficient. The long-term effectiveness of the
707 database will also depend on clearer governance mechanisms, stable institutional
708 involvement, and agreed procedures for data submission, validation, updating, and citation.
709 Bureaucratic constraints, fragmented responsibilities among agencies, and the absence of
710 formal obligations to share sediment transport data may remain major obstacles to the
711 creation of a fully comprehensive national archive. For this reason, one of the future priorities
712 will be to strengthen collaboration with public authorities, river basin districts, regional
713 agencies, and research institutions, with the aim of moving from voluntary data contributions
714 toward more structured and documented data-sharing workflows.

715 Nevertheless, the relational structure and the developed platform demonstrate the
716 effectiveness of a scalable system capable of integrating heterogeneous information and
717 providing a coherent overview of hydro-sedimentary observations at the national scale. While
718 historical datasets may inevitably retain some limitations inherited from their original sources,
719 future data collected within or in relation to this framework can achieve higher levels of
720 metadata completeness, traceability, quality control, and reusability.

721 The results of this study suggest several methodological guidelines for the future management
722 of sediment transport data. First, it is essential to promote the standardization of
723 instrumentation and measurement protocols, ensuring the traceability of procedures and the

comparability of results. Second, the systematic recording and preservation of metadata, including metadata related to effective acquisition frequency should become a consolidated practice. Third, spatial coverage of monitoring networks must be rebalanced by strengthening underrepresented areas and promoting the integration of local data into a coordinated national system. Finally, adopting interoperable and open-source data architectures is key to ensuring the scalability and long-term usability of the system.

In conclusion, this work demonstrates that, despite the fragmentation and gaps in the available data, it is possible to build a coherent and functional database capable of enhancing the existing information assets. The gained experience highlights the need to develop dedicated information infrastructures for an integrated management of sediment transport data, supporting both scientific research and environmental planning. This study, therefore, represents a concrete step toward the creation of a national reference system for the monitoring and analysis of fluvial sediment transport. The developed database provides a robust and effective framework for data organization, harmonization, and accessibility, and it is designed as an open and updatable tool, allowing future uploads and continuous integration of new datasets. Its long-term effectiveness, however, will largely depend on the progressive expansion and completion of the dataset through the inclusion of high-quality data, i.e., datasets accompanied by consistent metadata, clear methodological documentation, and adequate temporal and spatial coverage.

5 Author Contributions

Conceptualization: Luppichini M., Bini M., Comiti F.;

Methodology: Luppichini M., Bini M., Comiti F.;

Software: Luppichini M.;

Data curation: Luppichini M., Donnini M., Andreoli A., Innocenti L., Picco L., Chirici D., Brenna A., Billi P.; Longoni L.; Ciavola P.;

Formal analysis: Luppichini M., Bini M.;

Investigation: Luppichini M.;

Visualization: Luppichini M.;

Writing – original draft: Luppichini M., Bini M.;

Writing – review & editing: All authors;

Supervision: Comiti F., Bini M.;

Project administration: Bini M., Comiti F.;

Funding acquisition: Bini M.

757 6 Competing Interests

758 The authors declare that they have no conflict of interest.

759 7 Acknowledgements

760 This work is the result of the activities carried out within the working group “Geomorphology
761 and sustainable development of alluvial plains under anthropogenic pressure and climate
762 change (GEOALL)” of the Italian Association of Physical Geography and Geomorphology
763 (AIGeo).

764 8 Funding

765 This research was funded by the Autorità di Bacino Distrettuale dell’Appennino
766 Settentrionale—Misure di prevenzione tese a supportare ed ottimizzare la pianificazione di
767 gestione, la programmazione e realizzazione degli interventi di cui al PGRA (F54J16000020001).

768 9 Data availability

769 The Italian Fluvial Sediment Transport Database, described in this paper, is publicly available
770 at Zenodo: <https://doi.org/10.5281/zenodo.18799025> (Luppichini et al., 2026). The archived
771 version (v1.0) represents a frozen and citable release of the PostgreSQL relational database and
772 includes the complete dataset and database structure necessary to ensure reproducibility and
773 long-term accessibility.

774 The database can also be explored through a dedicated web application
775 (<https://lca.dst.unipi.it/SedimentTransport/>), which provides dynamic access to the data
776 following a brief registration procedure. While the web interface facilitates data consultation
777 and querying, the Zenodo archive constitutes the reference version for citation and
778 reproducibility purposes.

779

780 10References

- 781 Asselman, N. E. M. (1995). The impact of climate change on suspended sediment transport in the
782 river Rhine. In S. Zwerver, R. S. A. R. van Rompaey, M. T. J. Kok, & M. M. Berk (Eds.), *Studies in*
783 *Environmental Science* (Vol. 65, pp. 937–942). Elsevier. [https://doi.org/10.1016/S0166-](https://doi.org/10.1016/S0166-1116(06)80119-6)
784 [1116\(06\)80119-6](https://doi.org/10.1016/S0166-1116(06)80119-6)
- 785 Bailo, D., Paciello, R., Michálek, J., Cocco, M., Freda, C., Jeffery, K., & Atakan, K. (2023). The EPOS
786 multi-disciplinary Data Portal for integrated access to solid Earth science datasets. *Scientific*
787 *Data*, 10. <https://doi.org/10.1038/s41597-023-02697-9>

788 Billi, P., & Spalevic, V. (2022). Suspended sediment yield in Italian rivers. *CATENA*, 212, 106119.
789 <https://doi.org/10.1016/j.catena.2022.106119>

790 Bonfrisco, M., Coviello, V., Vignoli, G., Miyata, S., Piantini, M., Engel, M., Nadalet, R., Dinale, R., &
791 Comiti, F. (2025). Temporal variability of bedload vs suspended sediment load and respective
792 hysteresis cycles in a glacier-fed Alpine river. *Journal of Hydrology*, 660, 133215.
793 <https://doi.org/10.1016/j.jhydrol.2025.133215>

794 Cho, G., & Cromptvoets, J. (2019). The INSPIRE directive: some observations on the legal framework
795 and implementation. *Survey Review*, 51, 310–317.
796 <https://doi.org/10.1080/00396265.2018.1454686>

797 Cox, J., Huismans, Y., Knaake, S., Leuven, J., Vellinga, N., Vegt, M., Hoitink, A., & Kleinhans, M. (2021).
798 Anthropogenic Effects on the Contemporary Sediment Budget of the Lower Rhine-Meuse Delta
799 Channel Network. *Earth's Future*, 9. <https://doi.org/10.1029/2020ef001869>

800 Do, H. X., Gudmundsson, L., Leonard, M., & Westra, S. (2018). The Global Streamflow Indices and
801 Metadata Archive (GSIM) – Part 1: The production of a daily streamflow archive and metadata.
802 *Earth System Science Data*, 10(2), 765–785. <https://doi.org/10.5194/essd-10-765-2018>

803 Ellingsen, K., Yoccoz, N., Tveraa, T., Hewitt, J., & Thrush, S. (2017). Long-term environmental
804 monitoring for assessment of change: measurement inconsistencies over time and potential
805 solutions. *Environmental Monitoring and Assessment*, 189, 1–16.
806 <https://doi.org/10.1007/s10661-017-6317-4>

807 Färber, C., Plessow, H., Mischel, S. A., Kratzert, F., Addor, N., Shalev, G., & Looser, U. (2025). GRDC-
808 Caravan: extending Caravan with data from the Global Runoff Data Centre. *Earth System Science*
809 *Data*, 17(9), 4613–4625. <https://doi.org/10.5194/essd-17-4613-2025>

810 Favaro, E. A., & Lamoureux, S. F. (2015). Downstream patterns of suspended sediment transport in
811 a High Arctic river influenced by permafrost disturbance and recent climate change.
812 *Geomorphology*, 246, 359–369. <https://doi.org/10.1016/j.geomorph.2015.06.038>

813 Fortesa, J., Ricci, G., García-Comendador, J., Gentile, F., Estrany, J., Sauquet, E., Datry, T., & De
814 Girolamo, A. (2021). Analysing hydrological and sediment transport regime in two
815 Mediterranean intermittent rivers. *Catena*, 196, 104865.
816 <https://doi.org/10.1016/j.catena.2020.104865>

817 Francalanci, S., Paris, E., & Solari, L. (2013). A combined field sampling-modeling approach for
818 computing sediment transport during flash floods in a gravel-bed stream. *Water Resources*
819 *Research*, 49(10), 6642–6655. <https://doi.org/10.1002/wrcr.20544>

820 Fratianni, S., & Acquafotta, F. (2017). The Climate of Italy. In M. Soldati & M. Marchetti (Eds.),
821 *Landscapes and Landforms of Italy* (pp. 29–38). Springer International Publishing.
822 https://doi.org/10.1007/978-3-319-26194-2_4

- 823 Fredi, P., & Lupia Palmieri, E. (2017). Morphological Regions of Italy. In M. Soldati & M. Marchetti
824 (Eds.), *Landscapes and Landforms of Italy* (pp. 39–74). Springer International Publishing.
825 https://doi.org/10.1007/978-3-319-26194-2_5
- 826 Gurnell, A. M., Piegay, H., Swanson, F. J., & Gregory, S. V. (2002). Large wood and fluvial processes.
827 *Freshwater Biology*, 47(4), 601–619. <https://doi.org/10.1046/j.1365-2427.2002.00916.x>
- 828 Hooke, J. M. (2006). Human impacts on fluvial systems in the Mediterranean region.
829 *Geomorphology*, 79(3), 311–335. <https://doi.org/10.1016/j.geomorph.2006.06.036>
- 830 Jacob, D., Ehrenmann, F., David, R., Tran, J., Mirande-Ney, C., & Chaumeil, P. (2025). An ecosystem
831 for producing and sharing metadata within the web of FAIR Data. *GigaScience*, 14.
832 <https://doi.org/10.1093/gigascience/giae111>
- 833 Jadhav, H., Kandekar, A., & Das, S. (2024). Climatic and Anthropogenic Influences on Long-Term
834 Discharge and Sediment Load Changes in the Second-Largest Peninsular Indian Catchment.
835 *Water*. <https://doi.org/10.3390/w16243648>
- 836 Leatherman, S. P., Zhang, K., & Douglas, B. C. (2000). Sea level rise shown to drive coastal erosion.
837 *Eos, Transactions American Geophysical Union*, 81(6), 55–57.
838 <https://doi.org/10.1029/00EO00034>
- 839 Luppichini, M., & Bini, M. (2025). 40-Year shoreline evolution in Italy: Critical challenges in river delta
840 regions. *Estuarine, Coastal and Shelf Science*, 315, 109166.
841 <https://doi.org/10.1016/j.ecss.2025.109166>
- 842 Luppichini, M., Comiti, F., Andreoli, A., Aldighieri, B., Billi, P., Bollati, I. M., Brenna, A., Chirici, D.,
843 Ciavola, P., Cilli, S., Donnini, M., Francalanci, S., Innocenti, L., Longoni, L., Mao, L., Marchesini,
844 I., Murgia, I., Pellegrini, G., Penna, D., ... Bini, M. (2026). *Italian Fluvial Sediment Transport*
845 *Database (version 1.0)*. Zenodo. <https://doi.org/10.5281/zenodo.18799025>
- 846 Luppichini, M., Lazzarotti, M., & Bini, M. (2024). Climate change as main driver of centennial decline
847 in river sediment transport across the Mediterranean region. *Journal of Hydrology*, 636,
848 131266. <https://doi.org/10.1016/j.jhydrol.2024.131266>
- 849 Matthews, F., Verstraeten, G., Borrelli, P., Vanmaercke, M., Poesen, J., Steegen, A., Dégre, A.,
850 Rodríguez, B., Biëlders, C., Franke, C., Alary, C., Zúmr, D., Patault, E., Nadal-Romero, E., Smolska,
851 E., Licciardello, F., Swerts, G., Thodsen, H., Casalí, J., ... Panagos, P. (2023). EUSEDcollab: a
852 network of data from European catchments to monitor net soil erosion by water. *Scientific Data*,
853 10. <https://doi.org/10.1038/s41597-023-02393-8>
- 854 Naik, P. K., & Jay, D. A. (2011). Distinguishing human and climate influences on the Columbia River:
855 Changes in mean flow and sediment transport. *Journal of Hydrology*, 404(3), 259–277.
856 <https://doi.org/10.1016/j.jhydrol.2011.04.035>
- 857 Nejad, A. S., Alaiz-Rodríguez, R., McCarthy, G., Kelleher, B., Grey, A., & Parnell, A. (2024). SERT: A
858 transformer based model for multivariate temporal sensor data with missing values for

859 environmental monitoring. *Comput. Geosci.*, 188, 105601.
860 <https://doi.org/10.1016/j.cageo.2024.105601>

861 Parsons, M. A., & Fox, P. A. (2013). Is Data Publication the Right Metaphor? *Data Science Journal*.
862 <https://doi.org/10.2481/dsj.WDS-042>

863 Pellegrini, G., Mao, L., Rainato, R., & Picco, L. (2023). Surprising suspended sediment dynamics of an
864 alpine basin affected by a large infrequent disturbance. *Journal of Hydrology*, 617, 128933.
865 <https://doi.org/10.1016/j.jhydrol.2022.128933>

866 Rainato, R., Mao, L., García-Rama, A., Picco, L., Cesca, M., Vianello, A., Preciso, E., Scussel, G. R., &
867 Lenzi, M. A. (2017). Three decades of monitoring in the Rio Cordon instrumented basin:
868 Sediment budget and temporal trend of sediment yield. *Geomorphology*, 291, 45–56.
869 <https://doi.org/10.1016/j.geomorph.2016.03.012>

870 Rainato, R., Pellegrini, G., Lenzi, M. A., & Picco, L. (2025). Interaction between climatic conditions,
871 water- and sediment-fluxes in an alpine basin: Long-term monthly and seasonal analysis.
872 *Geomorphology*, 472, 109590. <https://doi.org/10.1016/j.geomorph.2024.109590>

873 Rodríguez-Blanco, M. L., Arias, R., Taboada-Castro, M. M., Nunes, J. P., Keizer, J. J., & Taboada-Castro,
874 M. T. (2016). Potential impact of climate change on suspended sediment yield in NW Spain: A
875 case study on the corbeira catchment. *Water (Switzerland)*, 8(10).
876 <https://doi.org/10.3390/w8100444>

877 Strobl, C. (2008). PostGIS. In S. Shekhar & H. Xiong (Eds.), *Encyclopedia of GIS* (pp. 891–898). Springer
878 US. https://doi.org/10.1007/978-0-387-35973-1_1012

879 Surian, N., & Rinaldi, M. (2004). Channel adjustments in response to human alteration of sediment
880 fluxes: examples from Italian rivers. *IAHS Publication*, 288, 276–282.

881 Syvitski, J. P. M., Vörösmarty, C. J., Kettner, A. J., & Green, P. (2005). Impact of Humans on the Flux of
882 Terrestrial Sediment to the Global Coastal Ocean. *Science*, 308(5720), 376–380.
883 <https://doi.org/10.1126/science.1109454>

884 Thodsen, H., Hasholt, B., & Kjærsgaard, J. H. (2008). The influence of climate change on suspended
885 sediment transport in Danish rivers. *Hydrological Processes*, 22(6), 764–774.
886 <https://doi.org/10.1002/hyp.6652>

887 Vercruyssen, K., Grabowski, R., & Rickson, R. (2017). Suspended sediment transport dynamics in
888 rivers: Multi-scale drivers of temporal variation. *Earth-Science Reviews*, 166, 38–52.
889 <https://doi.org/10.1016/j.earscirev.2016.12.016>

890 Vitolo, C., El-Khatib, Y., Reusser, D., Macleod, C., & Buytaert, W. (2015). Web technologies for
891 environmental Big Data. *Environ. Model. Softw.*, 63, 185–198.
892 <https://doi.org/10.1016/j.envsoft.2014.10.007>

893 Walling, D. E. (2006). Human impact on land–ocean sediment transfer by the world’s rivers.
894 *Geomorphology*, 79(3), 192–216. <https://doi.org/10.1016/j.geomorph.2006.06.019>

895 Walling, D. E., & Fang, D. (2003). Recent trends in the suspended sediment loads of the world's rivers.
896 *Global and Planetary Change*, 39(1), 111–126. [https://doi.org/10.1016/S0921-8181\(03\)00020-](https://doi.org/10.1016/S0921-8181(03)00020-1)
897 1

898 Wang, J., Haining, R., Zhang, T., Xu, C., Hu, M., Yin, Q., Li, L., Zhou, C., Li, G., & Chen, H. (2022).
899 Statistical Modeling of Spatially Stratified Heterogeneous Data. *Annals of the American*
900 *Association of Geographers*, 114, 499–519. <https://doi.org/10.1080/24694452.2023.2289982>

901 Wei, D., Liu, S., Wu, Y., Feng, S., Gao, H., Qin, C., Ren, D., Tang, W., & Zhang, Y. (2023). Impacts of
902 human activities and climate change on water and sediment evolution in four large subtropical
903 river basins in China. *Ecological Indicators*. <https://doi.org/10.1016/j.ecolind.2023.110958>

904 Zarfl, C., & Dunn, F. (2022). The delicate balance of river sediments. *Science*, 376, 1385–1386.
905 <https://doi.org/10.1126/science.abq6986>

906 Zhang, J., Wang, J., Zhao, N., Shi, J., & Wang, Y. (2024). Analysis of Changes in Runoff and Sediment
907 Load and Their Attribution in the Kuye River Basin of the Middle Yellow River Based on the Slope
908 Change Ratio of Cumulative Quantity Method. *Water*. <https://doi.org/10.3390/w16070944>

909 Zhang, K., Douglas, B. C., & Leatherman, S. P. (2004). *GLOBAL WARMING AND COASTAL EROSION*.
910 www.ipcc.ch

911