HYDROLOGICAL SCIENCES JOURNAL
https://doi.org/10.1080/02626667.2026.2742382

SPECIAL ISSUE: HISTORY OF HYDROLOGY

OPEN ACCESS Check for updates

A step towards the systematic reconstruction of the Italian pluviometric and thermometric heritage: the ANNHYDRO project

Barbara Lastoria[a], Francesca Archi[a], Matteo Balistrocchi[b], Daniela Biondi[c], Anna Botto[a], Giovanni Braca[a], Paolo Colosio[b], Francesca Conocchiella[c], Francesco Cruscomagno[c], Silvano Fortunato Dal Sasso[d], Mauro Fiorentino[e], Vito Iacobellis[f], Audrey Maria Noemi Martellotta[f], Federica Mesto[f,g], Leonardo Mita[d], Cosimo Peruzzi[a], Teresa Pizzolla[d], Roberto Ranzi[b], Vincenzo Totaro[f], Robertino Tropeano[a], Piergiorgio Vasta[c] and Martina Bussettini[a]

[a]BIO-ACAS, Istituto Superiore per la Protezione e la Ricerca Ambientale (ISPRA), Rome, Italy; [b]Dipartimento di Ingegneria Civile, Architettura, Territorio, Ambiente e di Matematica (DICATAM), Università di Brescia, Brescia, Italy; [c]Dipartimento di Ingegneria Informatica, Modellistica, Elettronica e Sistemistica (DIMES), Università della Calabria, Rende, Italy; [d]Dipartimento per l'Innovazione Umanistica, Scientifica e Sociale (DIUSS), Università della Basilicata, Matera, Italy; [e]Dipartimento di Ingegneria (DIING), Università della Basilicata, Potenza, Italy; [f]Dipartimento di Ingegneria Civile, Ambientale, del Territorio, Edile e di Chimica (DICATECh), Politecnico di Bari, Bari, Italy; [g]Dipartimento di Scienze del Suolo, della Pianta e degli Alimenti (DISSPA), Università degli Studi di Bari, Bari, Italy

ABSTRACT

The reconstruction of historical hydro-meteorological records is crucial for assessing long-term climate variability and for supporting adaptation policies. This paper introduces a recent and comprehensive effort to rescue the Italian climate heritage contained in the hydrological yearbooks printed in the 1920–1990s. Conducted within the framework of the ANNHYDRO project, this study describes the activities adopted for recovering daily data of rainfall and maximum-minimum temperature. Over 1 300 000 references to rainfall data and 110 000 references to temperature data were inventoried, assembled into gauge-specific series and georeferenced to build a comprehensive database. The historical evolution of the hydrological yearbooks is outlined, highlighting how historical monitoring and reporting practices have shaped the spatiotemporal availability of observations and contributed to the Italian hydrological heritage. The critical role of source traceability and metadata consistency, ensured by the one-to-one linkage between original publications and their digital counterpart, is discussed along with potential applications of the database.

ARTICLE HISTORY
Received 23 February 2026
Accepted 4 September 2026

EDITOR
S. Archfield

GUEST EDITOR
O. Batelaan

KEYWORDS
historical climate data rescue; digitization of climate data; Italy

1 Introduction

The awareness of the vital importance of the atmospheric system to life on Earth has been finding evidence since ancient times (Koutsoyiannis and Mamassis 2021). A considerable number of such traces have been preserved to the present day, providing us with valuable information about past climate and extreme events (Brázdil *et al.* 2005, WMO 2007).

Among climate variables, rainfall and air temperature measured at the Earth's surface can be considered those with the longest tradition of monitoring and are essential for improving our long-term comprehension of the climate system and producing advanced statistical analyses (Camuffo 2002, Marani and Zanetti 2015, Balistrocchi and Grossi 2020, Ranzi *et al.* 2021, 2025, Montanari *et al.* 2023, Volpi *et al.* 2024, Camuffo *et al.* 2025) needed for hydrological design (Grimaldi *et al.* 2011). Since the first gauges were installed, a shift towards a more quantitative approach to monitoring started, allowing the collection of long and reliable series of observed climate variables. To date, the availability of climate data has made possible the development and the advancement of modelling applications, as well as the improvement of our knowledge of Earth dynamics (Bojinski *et al.* 2014, Zeng *et al.* 2019).

Technological advances – such as those based on the use of remote sensing techniques – improve climate investigation in both its spatial and temporal evolutions, integrating intrinsic limitations of point measurements (Karl *et al.* 1989, Bacchi and Kottegoda 1995, New *et al.* 2001, Gourley *et al.* 2010, Wang *et al.* 2023). Here lies the foundation for the increasing number of gridded datasets that have been developed over the last 20 years, and that have become almost essential in facilitating the implementation of large-scale studies based on climate data, such as climate model evaluations and hydrological applications (Deidda *et al.* 2013, Izzaddin *et al.* 2024). Despite their spatial and temporal coverage and their ease of access and manipulation, gridded products are often affected by uncertainties that should be properly managed to improve their application (Newman *et al.* 2019). For example, it is well known how the accuracy of such products based on ground observations can be strongly influenced by the number of gauges, thus highlighting the importance of developing high-quality regional datasets of measured variables, especially – but not only – in terms of stations density, metadata availability, and compliance to international standard resolutions (Böhm *et al.* 2001, Auer *et al.* 2005, Merino *et al.* 2021). Several

CONTACT Vincenzo Totaro vincenzo.totaro@poliba.it Dipartimento di Ingegneria Civile, Ambientale, del Territorio, Edile e di Chimica (DICATECh), Politecnico di Bari, Via Edoardo Orabona, 4, Bari 70125, Italy

© 2026 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

authors have highlighted how observed data are very useful for integrating and testing gridded products, providing key information about their reliability and uncertainty (Huffman *et al.* 2009, Hersbach *et al.* 2020, Tepetidis *et al.* under review).

Despite the considerable importance that data availability assumes in modern hydrology and climatology, to date the majority of data remains unexplored and not digitized, especially before 1960 (Brunet *et al.* 2014, Engström *et al.* 2023, Noone *et al.* 2024). Data rescue assumes a key role in preserving this information from deterioration and providing it in digital format (Downs and Chen 2017, WMO 2024). To this aim, several projects have been initiated globally, such as Rainfall Rescue and MEDARE (Hawkins *et al.* 2023, Noone *et al.* 2024). With respect to Europe, we mention HISTALP in Austria, the MeteoSwiss historical digitization programmes in Switzerland, the UK Met Office MIDAS/ACRE and Rainfall Rescue projects, and the AEMET climate-archive recovery initiatives in Spain.

However, most efforts for data retrieval of gauged climate have been confined to monthly and annual scales. The interest in refining such temporal scales to daily resolution has increased over the last two decades, but advances have been counterbalanced by difficulties related to the accessibility and practical retrieval of recorded data (Menne *et al.* 2012, Serrano-Notivoli and Tejedor 2021). At the same time, it should be noted that assembling long series (especially rainfall and temperature) requires digitalizing historical records and obtaining sufficient metadata to ensure their quality (Brunetti *et al.* 2004).

In this framework, Italy has been well recognized to have a significant history of meteorological data collection, which dates back to the 17th century, continuing over the subsequent two centuries with regular observations carried out in isolated observatories (Brunetti *et al.* 2004, 2006b). In the decades following the unification of the country in the 1860s, several efforts were undertaken both to preserve the national scientific heritage and to establish an organized gauge network, designed to uniformly cover the entire territory and to provide regularly standardized observations, which were disseminated in printed volumes (Eredia 1908, 1919, Brunetti *et al.* 2006b). A turning point can be traced back to 1917 with the establishment of the National Hydrographic Service (Servizio Idrografico Italiano, in Italian; hereafter SII), a national authority aimed at monitoring and publishing hydrological climate data at the national scale by exercising a consolidated organization of network gauges and local offices. Such data (among which stand out rainfall and air temperature) were printed in paper form, mainly in the consolidated publications often known as hydrological yearbooks (or hydrographic bulletins). These publications continued until the end of the 20th century, when the responsibility for maintaining them moved from national coordination to regional agencies.

In the face of this considerable data heritage, the complete inventory, transcription in digital format and free availability of rainfall and temperature data at daily scale recorded from historical networks is still a challenge for Italy, as it is for many other countries (Brunet and Jones 2011, Libertino *et al.* 2018). With respect to Italy, the online publication of daily rainfall and temperature data in ready-to-use format has been carried out only in some regions, totally or at least in some parts (e.g. Calabria, Umbria, Emilia-Romagna, Marche, Toscana, Piemonte and the autonomous provinces of Trento and Bolzano). Moreover, such data (rainfall, in particular) were used in several studies with data collection and management left to targeted initiatives (e.g. Testa *et al.* 2006, Ciccarelli *et al.* 2008, Fatichi and Caporali 2009, Moccia *et al.* 2021). At the same time, several studies addressed the reliability of gridded products with different benchmarks of ground-measured variables across Italy, highlighting how the uncertain and sparse spatial and temporal coverage of ground-based air temperature and rainfall measurements can influence the performance of such products (Sarigil *et al.* 2024, Moccia *et al.* 2025).

This overview makes quite evident the lack of a comprehensive and systematic review and analysis of all the data reported in the Italian hydrological yearbooks and bulletins, as well as the digitization, validation and analysis of the temporal and spatial continuity of this heritage. Several activities were put in place for retrieving this heritage of data at the national scale and orientated to different types of variables, such as the SIREN project (Mazzoglio *et al.* 2025a) for daily discharge data recorded from SII, the ReData project (Ceppi *et al.* 2026) focused on the meteorological data collected by the Italian Royal Central Meteorological Office from 1879 to 1940, and the Cli-DaRe@School project, concentrated on monthly data of rainfall before 1950 and temperature between 1921 and 1950 collected in publications edited by SII (Manara *et al.* 2025a, 2025b).

Against this background, the need for complete recollection, digitalization and open accessibility of the Italian hydrological heritage in terms of available measurements of rainfall and temperature at daily scale emerges clearly. The ANNHYDRO project complements these efforts by establishing the first comprehensive initiative aimed at providing the systematic digitized collection of daily rainfall and maximum–minimum temperature series from the hydrological yearbooks at a national scale, offering daily temporal resolution and supporting metadata for benchmarking and extending.

Funded by the Italian National Recovery and Resilience Plan through the RETURN Extended Partnership (Mission 4, Component 2, Investment 1.3; CUP J33C22002840002), ANNHYDRO was conceived with the aim of building a national open-access dataset of daily rainfall and temperature data collected from SII, with a partnership led by the Italian Institute for Environmental Protection and Research (Istituto Superiore per la Protezione e la Ricerca Ambientale, in Italian; hereafter ISPRA) and supported by the Polytechnic University of Bari, the University of Basilicata, the University of Brescia and the University of Calabria. The proposed methodology to be implemented in ANNHYDRO was orientated towards the rational collection of all information contained in the hydrological yearbooks and bulletins from 1917 to the late 1990s and was initially conducted in 2024–2025.

In this paper we describe the activities put in place for retrieving daily data, with a particular focus on rainfall and maximum and minimum temperature collected between 1951 and the late 1990s, when the hydrological yearbooks were published in the consolidated structure conceived after the end of World War II. More specifically, beyond presenting the resulting dataset, this paper addresses methodological

challenges of broad interest to the data rescue community, highlighting three main contributions:

(1) The assessment of the consistency and evolutionary dynamics of a historical monitoring network at the national scale, showing how past management policies influence the current availability of climate data;
(2) The importance of establishing a unique and stable metadata inventory, addressing the critical issues of station relocations and nomenclature changes over time;
(3) The need for a structured approach to quality control and cross-validation between different paper and digital sources.

The paper is organized as follows: in Section 2 the history of the Italian monitoring service and the methodological structure of the ANNHYDRO project are illustrated; in Section 3 the main results of our work are described, and some potential future applications are discussed. Finally, in the Conclusions section we highlight the general added value of the project and propose future steps to implement a definitive and rational collection of the Italian hydrological heritage.

2 Background, project structure, and methodological approach

2.1 Background

The crucial point in the collection and management of Italian hydrological data can be dated back to the Lieutenancy Decree of 17 June 1917, no. 1055, which enacted the SII (Manara *et al.* 2025b). This agency was entrusted with the commitment – among other duties – to monitor rainfall, temperature, and other hydrological variables (e.g. wind speed and direction). The original mandate of SII was twofold: on the one hand, to assess water resource availability; on the other, to provide reliable data on extreme events for weather-related hazard estimation (De Marchi 1931). To this aim, the Italian territory was initially partitioned into local hydrographic compartments (Compartimenti Idrografici, in Italian) according to hydrographic criteria, whose number and borders changed several times over the years. Hydrological data at a daily scale were published from 1917 to 1926, in books known as hydrographic bulletins (Bollettini Idrografici, in Italian), and from 1927 in volumes called hydrological yearbooks (Annali Idrologici, in Italian). The data format moved from monthly tables (Fig. 1(a)) to an annual table format (Fig. 1(b)). This latter publication was generally issued in two volumes, listed as Part 1 and Part 2: thermometric and pluviometric data were included in Part 1, while hydrometric data were reported in Part 2.

The format of the hydrological yearbooks changed many times, partly due to the events of World War II, only reaching in 1951 the consolidated structure that remained unchanged thereafter. In addition, SII published other special volumes, as the complete list of all thermo-pluviometric stations operational in 1970 (Ministero Dei Lavori Pubblici and Idrografico 1976; hereafter Pubblicazione 27) and the characteristics of Italian rivers (Ministero Dei Lavori Pubblici and Idrografico 1980), that are still considered pivotal moments in the hydrological history of Italy.

This framework persisted until the enactment of Law no. 183 of 18 May 1989, which unified the SII with the Mareographic Service (Servizio Mareografico, in italian), thereby giving rise to the Italian Hydrographic and Mareographic Service (Servizio Idrografico e Mareografico Nazionale, in Italian; hereafter SIMN). SIMN was structured with 14 local compartments and a Central Direction (see Fig. 2) and was closed in 2002 when, with the publication of the decree of the Italian President of the Council of Ministers (DPCM) of July 24, the commitments of the SIMN compartments were transferred to the local administrations (Regions), establishing the end of standard

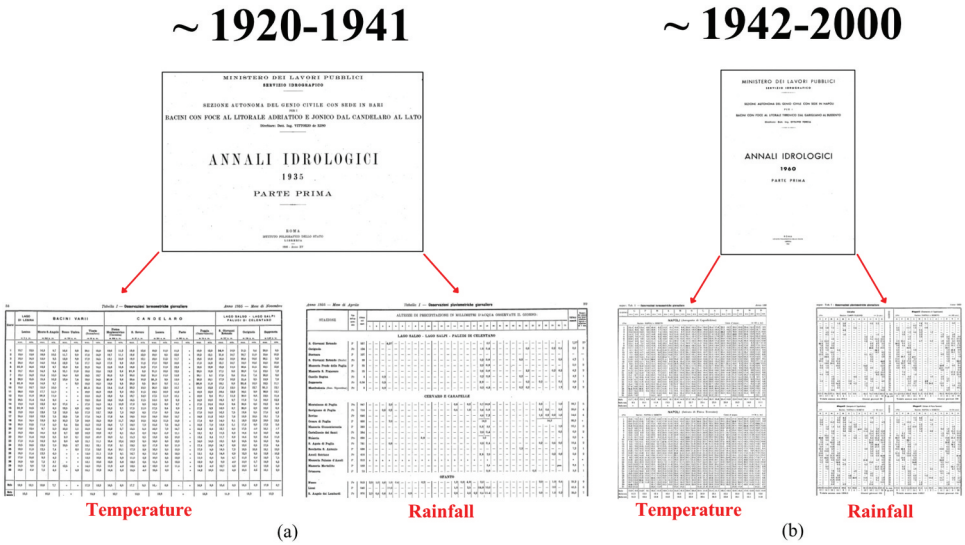


Figure 1. Typical format of Italian hydrological yearbooks for daily temperature and rainfall: (a) years ~1920–1941, where annual data were split into monthly tables; (b) years ~1942–2000: where annual data were reported in a single table for each station.

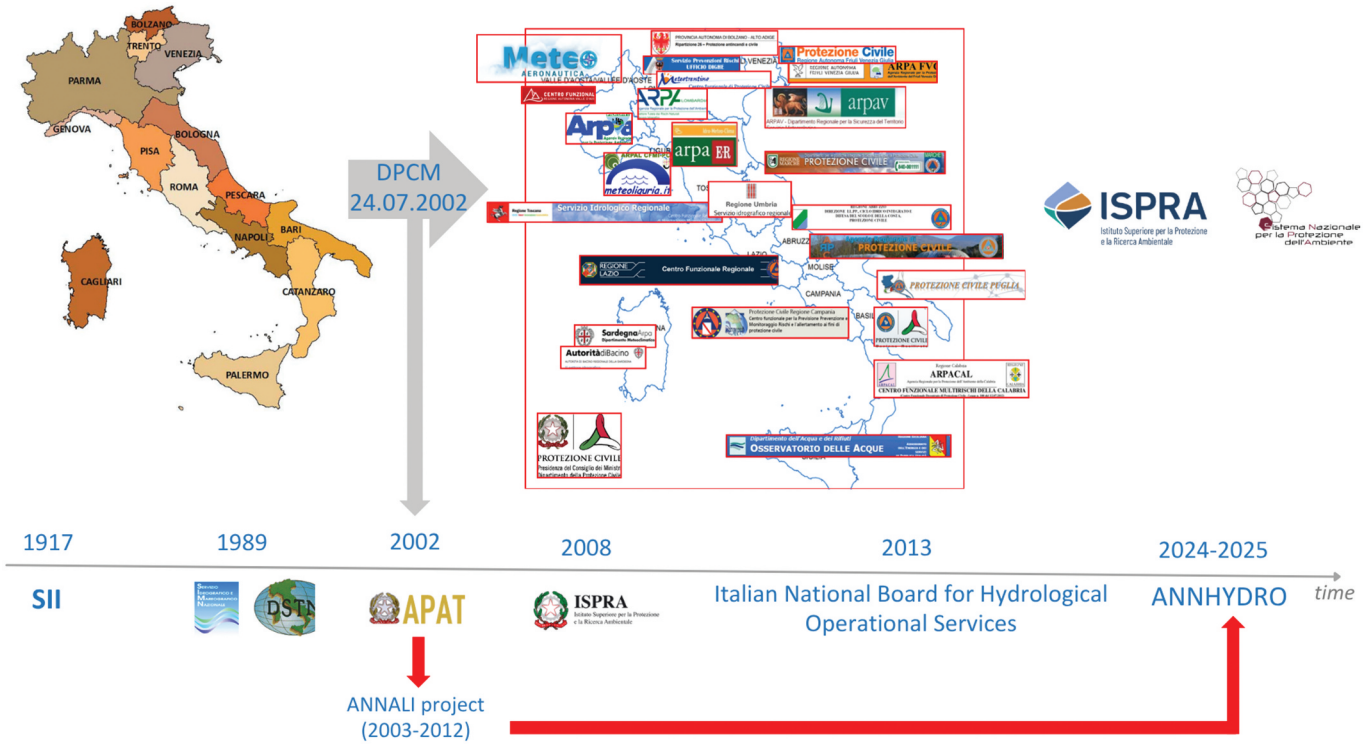


Figure 2. Evolution of hydro-thermo-pluviometric monitoring authorities in Italy.

publication at the consolidated compartment scale of hydrological yearbooks. The Central Direction, instead, was first replaced by the Agency for Environmental Protection and Technical Services (Agenzia per la Protezione dell'Ambiente e per i Servizi Tecnici, in Italian; hereafter APAT) and, since 2008, by ISPRA (Fig. 2).

The first attempt to digitize the entire historical hydrological information contained in hydrological yearbooks and hydrographic bulletins was carried out first by APAT and then by ISPRA from 2003 to 2012 within an initiative named the "ANNALI project" (Progetto Annali, in Italian), conceived to preserve the memory of the extremely relevant heritage of SII and SIMN and to make it available as an open-source dataset (Lastoria *et al.* 2025). During this project, hydrological yearbooks and hydrographic bulletins published in paper format were acquired as digital images, creating an archive accessible on the APAT website since 2005, when the images were indexed and organized into a browsing interface (http://www.bio.isprambiente.it/annalipdf/). This interface allows the user to query the archive by selecting compartments, year, and volume (Part 1 or Part 2), and easily moving between the published pages.

From these images textual information was extracted with optical character recognition (OCR) systems and organized in a structured query language (SQL) relational database where records were linked to the corresponding position in the hydrological yearbooks. With the aim of supporting the reader with a tangible example of the yearbooks' structure, we present two typical pages containing rainfall and maximum and minimum temperatures at daily scale in the hydrological yearbooks after 1951:

(1) For rainfall data (Fig. 3), each page hosts data for up to four stations, identified by name, related catchment, gauge type, and elevation (in m a.s.l.). For each day of the year, the 24 h antecedent cumulative precipitation observed at 9.00am is printed.
(2) For temperature data (Fig. 4), each page hosts data for up to two stations, identified by name, related catchment (sometimes supported with the river name), gauge type, and elevation (in m a.s.l.). For each day of the year, the antecedent 24 h minimum and maximum temperature values observed at 9.00am are printed.

During the acquisition process, each page was named with an alphanumeric string (file name) containing a coded reference to the compartment, year, part and page of the yearbook where the image was published. Such codes make it possible to uniquely link the SQL records to the corresponding page of the yearbook. In the SQL database each station is also identified in data tables by a composite primary key, which is the combination of three numerical codes associated with the compartment, the basin, and the station. The primary composite key gives access to a "parameter table" where information related to each station is reported: file name, station and basin name, hydrological variable (e.g. rainfall, temperature), temporal aggregation (e.g. daily, monthly). However, the same station can have many key codes, because basin and station names reported in the hydrological yearbook changed over the years, due to typographical errors, different operators working in different years, and changes in municipality nomenclature.

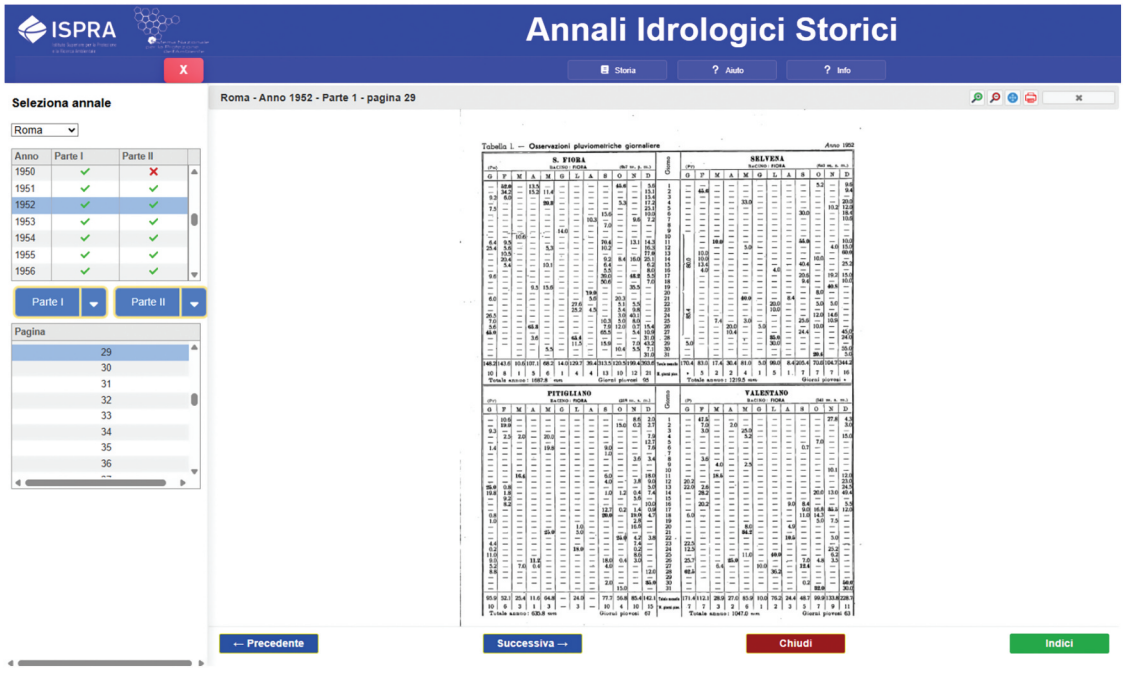


Figure 3. Typical format for reporting daily rainfall since 1951 (source: ISPRA website: http://www.bio.isprambiente.it/annalipdf/; Compartment: Roma, Year: 1952, Part I, page: 29 - Last access: 12/10/2026).

In hydrological yearbooks, graphical symbols or specific text formats sometimes appear together with numerical values of records, indicating, for example, rain cumulated over multiple days, uncertain or unmeasurable values, snow mixed with rain, etc. In order to maintain this information in the SQL database, a two-character alphanumeric code was associated with numeric values of rainfall and temperature.

For the random verification of this database, the textual information of a considerable number of hydrological yearbook pages was manually digitized as Microsoft Excel files linked to hydrological yearbooks by an alphanumeric string containing a coded reference to the compartment, year, part and page. In this database, the above-mentioned graphical symbols were often attached as characters to the numerical value of rainfall and temperature.

In this way, two independent databases were made available: the SQL database and the MS Excel-based one, with the latter to be used only for verification purposes of the SQL. However, the inhomogeneities and inconsistencies existing in both datasets made it very complex to analyse their temporal and data integrity and their quality, precluding the return of the information in a ready-to-use format without a preventative data re-organization and validation.

Moving forward from this point, the project was aimed at: (i) producing a database with all the time series of rainfall and temperature at daily scale contained in the hydrological yearbooks, (ii) implementing a reliable procedure for first-level control of data reliability, and (iii) publishing the database on an open-access web infrastructure.

In this paper we illustrate all the procedures related to the first two points.

2.2 Methodological framework

The reconstruction of the daily time series of rainfall and temperature published in the hydrological yearbooks was carried out by implementing the following activities:

(i) Development of the metadata inventory featuring the main characteristics of every station whose daily data are included in the hydrological yearbooks year by year.
(ii) Aggregation of single-year time series into multi-year series for every station, building a new inventory with stable metadata.
(iii) Extraction of all the time series from the SQL database (and retrieval of those missing), matching with the metadata inventory defined at point (ii), and construction of a preliminary version of the ANNHYDRO database (hereafter dbv0).
(iv) Construction of a twin database by extracting the time series from MS Excel files (hereafter dbv1), with quality verification at a monthly scale. As in point (iii), missing series are manually retrieved.
(v) Comparison of dbv0 with data extracted from dbv1 and finalization of the ANNHYDRO database.

Figure 5 provides a synthetic overview of the data-processing pipeline.

In the following subsections a detailed description of these activities is provided to outline a clear view of the project contents. Collection and analysis of metadata were implemented according to the national guidelines published by ISPRA (Braca *et al.* 2013).

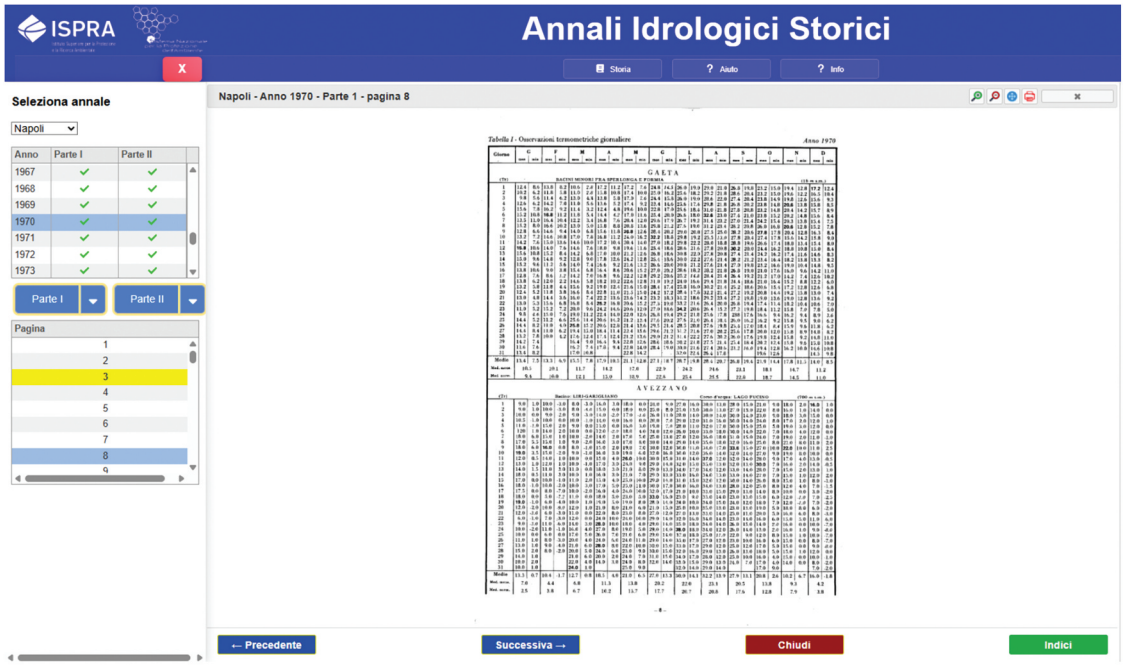


Figure 4. Typical format for reporting daily maximum and minimum temperature since 1951 (source: ISPRA website: http://www.bio.isprambiente.it/annalipdf/, Compartment: Napoli, Year: 1970, Part I, page: 8 - Last access: 12/10/2026).

2.2.1 *Daily series and metadata inventory*

To ensure the complete recovery of all available data, an inventory of all the daily series reported year by year in the hydrological yearbooks was carried out. For this purpose, a complete list of all the stations with daily data was manually compiled by collecting, in a standardized format (see the template in Table 1), the metadata associated with each station. These metadata were taken from the information printed along with the daily data, as illustrated in Figs. 3 and 4, which show examples of the original pages. In particular, the metadata record explicitly includes the precise position of the original daily series in the hydrological yearbooks. The final version of the inventory supplies all data reported in Table 1. The exact match between the metadata inventory records and the printed daily series allows straightforward access to the original data and metadata, thus providing a remarkable advantage in the evaluation of digitized data quality and reliability.

Within the proposed methodological framework, to ensure the consistency and traceability of the rescued dataset, the metadata inventory plays a crucial twofold role. First, it makes possible the unambiguous inventory of all the records included in the yearbooks/bulletins. Second, the univocal alphanumeric string, obtained by merging the compartment, catchment, year and order fields, is built to access both databases.

Furthermore, to tackle possible changes in station metadata within the same year (only for yearbooks published before 1951) or among years, the aggregation of all records belonging to an individual station was carried out by defining a unique station name and identifying all the corresponding alternative names, herein referred to as *aliases*. To precisely carry out this operation, accurate and stable metadata about stations are essential (Beltrano *et al.* 2012). Before the onset of World War II, the hydrological yearbooks provided detailed metadata and station-specific updates, documenting new installations, suppressions and changes in the registration officer. Examples are illustrated in Figs. 6 and 7.

After 1951, the metadata in Fig. 6 were only sporadically reported, and the information in Fig. 7 no longer included the person responsible for the station management and the registration officer's name. In addition, catchment information was limited to the main catchment, whose nomenclature also changed over time and/or was refined into specific sub-catchments.

The assembly of yearly records into a single time series with a unique station name was therefore preceded by the identification of all aliases associated with each station. This was achieved by comparing station names, elevations, and – when elevations were uncertain – the starting year of the observation record and/or the height of the gauge orifice above ground. To enhance the available information from other sources, comparable analyses conducted in previous projects – such as Valutazione delle Piene in Italia (VAPI) (Ferrari and Versace 1994) – were also used to support this procedure, along with metadata from specific expertise, and interpretation of publication practices. Consequently, the resulting metadata inventory inevitably retains a certain degree of subjectivity, partly due to uncertainties and inconsistencies present in the original yearbooks/bulletins.

To georeference the stations associated with each time series, the official geographical coordinates reported in Pubblicazione 27 were used as the primary source. If this source failed to provide these data, alternative sources were examined, such as those listed below (in order of reliability):

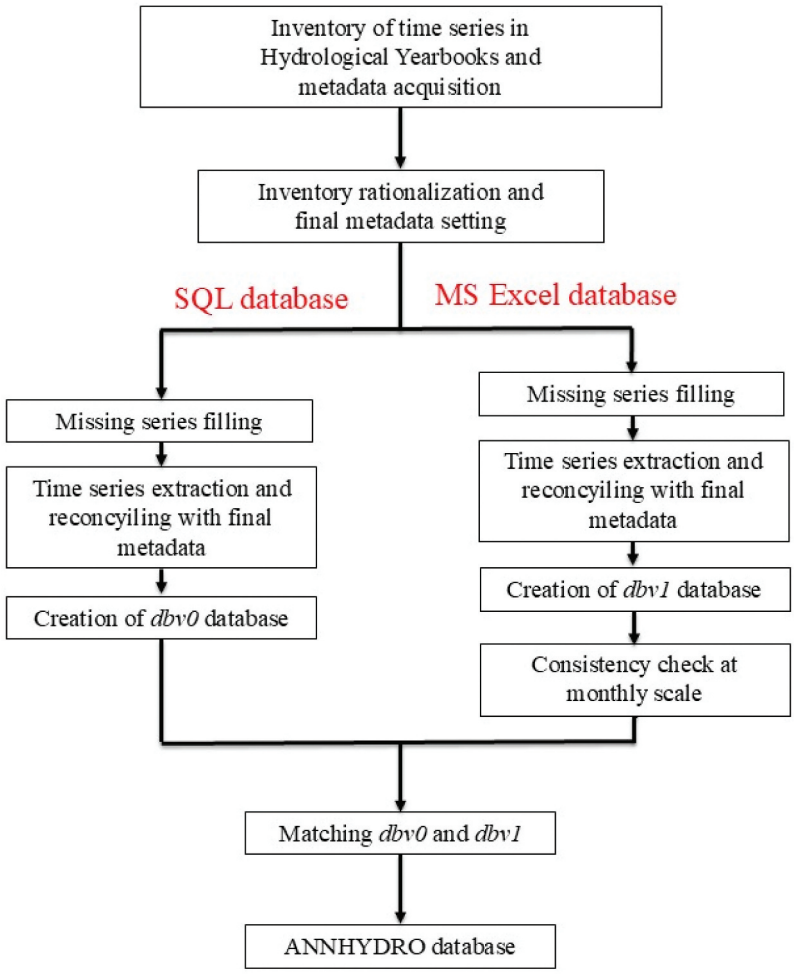


Figure 5. Workflow of the ANNHYDRO data-processing pipeline.

Table 1. Template of the information collected for the daily station inventory: field names and descriptions.

Field name	Field description
Compartment	Code of the compartment that the station belongs to
Year	Year of the hydrological yearbook
Page	Page number in the hydrological yearbook
Order/ Position	Order or position of the station in the page (integer value)
Catchment	Catchment name reported for the station
River	Name of the relevant river (only for temperature records)
Elevation	Station elevation (m a.s.l.)
Gauge type	Gauge type (selected from an enumeration list)
Station name	Station name (exactly as reported in the hydrological yearbook)

- Official metadata provided by regional/provincial agencies responsible for environmental monitoring;
- Toponym coordinates provided by the Army Geographical Institute dataset (available from the geoportal of Ministry of Environment and Energy Security, http://www.pcn.minambiente.it/viewer/);
- Toponym coordinates estimated by the use of geographical information systems.

The final result of this process is an improved and internally consistent metadata inventory, featuring unique station names, elevations, catchment names and coordinates for all the stations reported in the original yearbooks/bulletins.

Finally, the cross-check between final inventories and SQL and MS Excel databases allowed the identification of the missing files, and/or the missing data (or records) of the "ANNALI Project", i.e. those files/data for which a manual recovery from the hydrological yearbook was required and, then, carried out.

2.2.2 *Time series extraction and database population*

Once all series are available and the final inventory established, each inventory record is linked to SQL and MS Excel databases by means of its alphanumeric codes. This procedure enables the extraction of all the time series from both SQL and MS Excel databases, resulting in two independent sources, dbv0 and dbv1, respectively.

It is worth mentioning that in the ANNALI project dbv1 was conceived with the purpose to serve as a verification tool for dbv0, without any check on its reliability. In order to improve the overall quality of dbv1, a further step for quality assessment of data by using some synthetic measurements such as the cumulated monthly or annual rainfall and the mean monthly temperatures was introduced (see Section 3.2). In cases where substantial discrepancies were found, a manual search for the underlying error was undertaken.

NOTIZIE SOMMARIE

SUL LAVORO SVOLTO DALLA SEZIONE E SULLA SITUAZIONE DEL SERVIZIO AL TERMINE DELL'ANNO

Durante l'anno 1933 è stata intensificata la sistemazione della rete esistente al 31 dicembre 1932, impiantando nuove stazioni, trasformandone e sopprimendone altre, che non rispondevano più agli scopi per i quali erano state impiantate.

RETE TERMOMETRICA. — È entrata in funzione la nuova stazione di Lucugnano, ed è stata trasformata in termometrica la stazione termografica di Barletta.

Per effetto delle suddette variazioni la consistenza della rete termometrica, alla fine dell'anno 1933, risultava di 46 stazioni, con una densità approssimativa di una stazione ogni 435 *kmq*.

PROSPETTO I — **Stazioni di nuovo impianto o comunque trasformate.**

STAZIONE	Tipo	Data dell'impianto o della trasformazione	STAZIONE	Tipo	Data dell'impianto o della trasformazione
		1933			1933
Lucugnano	Tm	1 Gennaio	Gallipoli	Pr	17 Novem.
Masseria Piccinna	P	15 Luglio	Maglie	Pr	19 Novem.
Francavilla Fontana	P	18 Luglio	Otranto	Pr	23 Novem.
Torre S. Giovanni	P	25 Luglio	S. Maria di Leuca (Semaf.)	Pr	12 Dicem.
Spinazzola	Pr	23 Ottobre	Brindisi (Semaforo)	Pr	16 Dicem.
Montemilone	Pr	23 Ottobre	Barletta	Tm	31 Dicem.

RETE PLUVIOMETRICA. — Sono entrate in funzione durante l'anno 1933 le stazioni pluviometriche di Masseria Piccinna, Francavilla Fontana e Torre S. Giovanni; ed è da segnalare la trasformazione da pluviometriche in pluviografiche delle stazioni di Spinazzola, Montemilone, Gallipoli, Maglie, Otranto, S. Maria di Leuca (Semaforo) e Brindisi (Semaforo).

Sono state soppresse, per altro, la pluviografica di Foggia (Campo Sperimentale Agrario) e la pluviometrica di Alimini (Bonifica), per l'assoluta impossibilità di ottenerne il regolare funzionamento.

In dipendenza di dette variazioni la rete delle stazioni risultava, al termine dell'anno, di 150, delle quali 50 dotate di apparecchio registratore: la densità media era, quindi, di una stazione ogni 134 *kmq*.

RETE IDROMETRICA. — Non ha subito, nel corso dell'anno, alcuna variazione rispetto all'anno precedente.

MISURE DI PORTATA. — Le misure di portata eseguite durante l'anno ascesero a 192, delle quali 133 effettuate alle stazioni permanenti di misura e 59 in altre località per indagini varie.

I bacini imbriferi per i quali è stato possibile lo studio dei *bilanci idrologici* annuali sono in numero di 6, cioè di due in più rispetto all'anno decorso (v. Sezione E — Portate e Bilanci Idrologici).

PROSPETTO II — **Stazioni soppresse.**

STAZIONE	Tipo	Data della soppressione o della trasformazione	STAZIONE	Tipo	Data della soppressione o della trasformazione
		1933			1933
Foggia (Campo Sper. Agr.)	Pr	19 Gennaio	Maglie *	P	19 Novem.
Foggia (Campo Sper. Agr.)	Fr	19 Gennaio	Masseria Coviello	F	22 Novem.
Veglie	F	14 Maggio	Otranto *	P	23 Novem.
Alimini (Bonifica)	P	15 Luglio	Masseria Salmenta	F	6 Dicem.
Spinazzola *	P	23 Ottobre	S. Maria di Leuca (Sem.) *	P	12 Dicem.
Montemilone *	P	23 Ottobre	Brindisi (Semaforo) *	P	16 Dicem.
Gallipoli *	P	17 Novem.	Barletta	Tr	31 Dicem.
Salve	F	18 Novem.			

* La stazione, in realtà, non è stata soppressa — chè anzi continua a funzionare — ma essendo stata trasformata in pluviografica (v. Prospetto I) la stessa va segnalata soltanto provvista di apparecchio registratore.

RETE FREATIMETRICA. — Per la impossibilità di ottenerne il regolare funzionamento sono state soppresse le stazioni freatimetriche di Veglie, Salve, Masseria Coviello e Masseria Salmenta e la freatimetrografica di Foggia (Campo Sperimentale Agrario).

Per effetto di tali soppressioni il numero delle stazioni freatimetriche si è ridotto a 102, rispetto a 107 del 1932.

OSSERVAZIONI TORBIOMETRICHE. — È stato continuato il prelevamento delle torbide

Figure 6. Example of a page from the Italian hydrological yearbooks containing information about network management (Compartment of Bari, year 1933, Part 2, page 3).

TAB. I. — *Elenco e caratteristiche delle stazioni pluviometriche.*

STAZIONE	BACINO PRINCIPALE	Bacini secondari di. 1° ordine	2° ordine	3° ordine	Tipo dell'apparecchio	Coordinate geografiche Longit.	Latit.	Quota sul mare (metri)	Altezza della bocca dell'apparecchio sul suolo (metri)	Anno dell'inizio delle osservazioni	Ente da cui dipende la stazione	Cognome e nome dell'osservatore	OSSERVAZIONI
	(*Segue*)												
Viesti	VARII	—	—	—	P	3° 44′	41° 53′	50	2,00	1921	R. Marina	Capo semaforista.	Funzionò anche dal 1884 al 1900
Monte S. Angelo	VALLE DI MATTINATA	—	—	—	Pn	3° 30′	41° 42′	843	15,00	1920	S. I.	Tancredi Giovanni.	
Torremaggiore	CANDELARO	Candelaro	—	—	Pr	2° 50′	41° 41′	169	18,00	1917	A. P.	Borrelli Michele.	Funzionò come pluviometro nel 1909 e 1910.
S. Marco in Lamis	id.	Candelaro	—	—	Pr	3° 11′	41° 43′	560	15,00	1917	id.	Leone Michele.	
Alberona	id.	Salsola	Salsola	—	Pr	2° 40′	41° 26′	700	1,50	1917	id.	De Mattei Leonardo.	
Lucera	id.	Candelaro	Candelaro	—	P	2° 53′	41° 30′	251	1,70	1911	Privato	Curato Giandomenico.	Funzionò anche dal 1881 al 1886
Faeto	id.	Celone	Celone	—	Pr	2° 43′	41° 19′	905	14,00	1917	A. P.	Pavia Annita.	Funzionò come pluviometro negli anni 1884, 1885 e 1886.
Troia	id.	id.	id.	—	P	2° 51′	41° 22′	439	10,80	1907	U. C. M.	Frisoli Gerardo.	
Foggia	id.	id.	Celone	—	Pr	3° 6′	41° 28′	87	37,00	1873	id.	Nigri Vincenzo.	
Cerignola	LAGO SALSO, ECC.	—	—	—	P	3° 27′	41° 16′	124	11,00	1921	S. I.	Caputi Luigi.	Funzionò anche negli anni 1882-1884; 1887-1888 e 1913-1915.
Manfredonia (Bonifica Sipontina)	id.	—	—	—	P	3° 25′	41° 35′	2	9,90	1900	G. C. Foggia	Tartaglia Guido.	
Monteleone di Puglia	CERVARO	Cervaro	—	—	Pn	2° 48′	41° 10′	847	1,30	1920	S. I.	Volpe Felice.	

Figure 7. Example of a page from the Italian hydrological yearbooks containing information about pluviometric station characteristics (Compartment of Pescara, year 1921, Part 2, page 12). Metadata contained in columns are (from left to right): station name (col. 1), catchment and sub-catchment names (main catchment col. 2 and sub-catchment order cols. 3–5), gauge type (col. 6), station coordinates (cols. 7–8), station elevation a.s.l. (col. 9), height of gauge mouth above ground (col. 10), monitoring start year (col. 11), body responsible for station management (col. 12), recording officer's initials and/or surname (col. 13), and details about previous recording periods (col. 14). For several years, coordinates were removed and catchment information reduced.

2.2.3 *Building of the database*

As a final step, the dbv0 had to be cross-checked with dbv1, in order to verify that the data associated with each station were correct and to minimize errors. All the inconsistencies were to be subsequently examined against the original hydrological yearbooks and corrected accordingly. The final version of the ANNHYDRO database was then finalized.

3 Results and discussion

3.1 *Inventory and metadata analysis*

The final census of the hydrological yearbook records derived from the metadata inventory resulted in approximately 1 300 000 raingauge records and 110 000 temperature gauge records. The temporal distribution of the station records from

1951 – each corresponding to an annual daily series – is illustrated in Fig. 8, where compartments are aggregated by macro-areas. It must be noted that the standardized formats of daily data publication – illustrated in Figs. 3 and 4 – were adopted by the hydrological yearbooks in 1951 for all compartments except for the Venezia Compartment, which began using this format in 1955, and for the Trento and Bolzano Compartments, which adopted it in 1975 following their separation from the Venice Compartment.

Figure 8 provides an informative overview of the evolution of the Italian monitoring network. A first consideration is that the pluviometric network is substantially more consistent than the thermometric one. In Fig. 8 there are two separate shaded areas, which are related to those years for which publication of hydrological yearbooks was not completed. The area on the left extended until 1954 and was motivated by the policy of Venezia Compartment (see above). The area on the right, in contrast, reflects the availability of printed hydrological yearbooks at the time of Progetto Annali digitalization. It is interesting to note that the absolute number of operational stations reflects not only the size of each compartment but also the specific morphological and climatic characteristics of the corresponding catchments. These considerations are further supported by Fig. 9, which illustrates the evolution of the spatial distribution of all retrieved gauges. Both Fig. 8 and Fig. 9 indicate that the rain and the temperature gauge networks followed contrasting trends in the number of available daily series: decreasing for the raingauges and increasing for the temperature gauges.

The metadata inventory analysis revealed that uncertainty involves primarily the fields with the station name, the elevation and the gauge type. As an illustrative example, we report some results cumulated over the compartments of Bolzano, Parma and Trento. On average, 44.5% of stations underwent at least one name change over the years, underscoring a substantial challenge for the construction of homogeneous time series and, consequently, for the practical reliability of the resulting dataset (Hawkins *et al.* 2023). Additionally, in certain years, the same station name was assigned to two distinct stations. More than 50% of the stations are affected by an elevation change. In most cases, these changes do not appear to result from actual station relocations but rather from typographical errors. An emblematic case of uncertain metadata is shown in Table 2, where the changes in station name, elevation, and gauge type for the raingauge of Napoli (Istituto di Fisica Terrestre) are evident. The raingauge is located in the city of Napoli and its position – to our knowledge – never changed over time. As can be seen, the gauge elevation above sea level changes, whereas the gauge height above the ground remains constant. This last characteristic is confirmed by the raingauge position, always placed on the top of the same building. However, a different elevation of the site is reported for these three consecutive years (Longobardi and Villani 2010). Lastly, the type of raingauge appears to have changed between 1970 and 1971 from a recording raingauge (Pr) to a simple raingauge (P).

Another example is provided in Table 3 for the Parma Compartment, where selected metadata records for the Sonico raingauge are reported. In this case as well, both the monitor starting year and the gauge height above the ground remain constant. Nevertheless, the station elevation exhibits considerable and alternating variations of approximately 300 m. The change in elevation after 1982 is consistent with the elevation reported for the associated thermometric station, which is equipped with a maximum–minimum thermometer and is placed at 641 m a.s.l. for the three years when data were published (1988, 1990 and 1991). Therefore, in this study, variations in the station elevation were not assumed to be a key criterion in the identification of breakpoints in series homogeneity. According to the published information, the type of gauge changed after 1958, when a snow-raingauge (Pn) was replaced by a simple raingauge (P). Moreover, when identifying aliases associated with an individual series that exhibited changes in the station name, the monitor starting year was adopted as the primary guiding criterion to ensure series homogeneity.

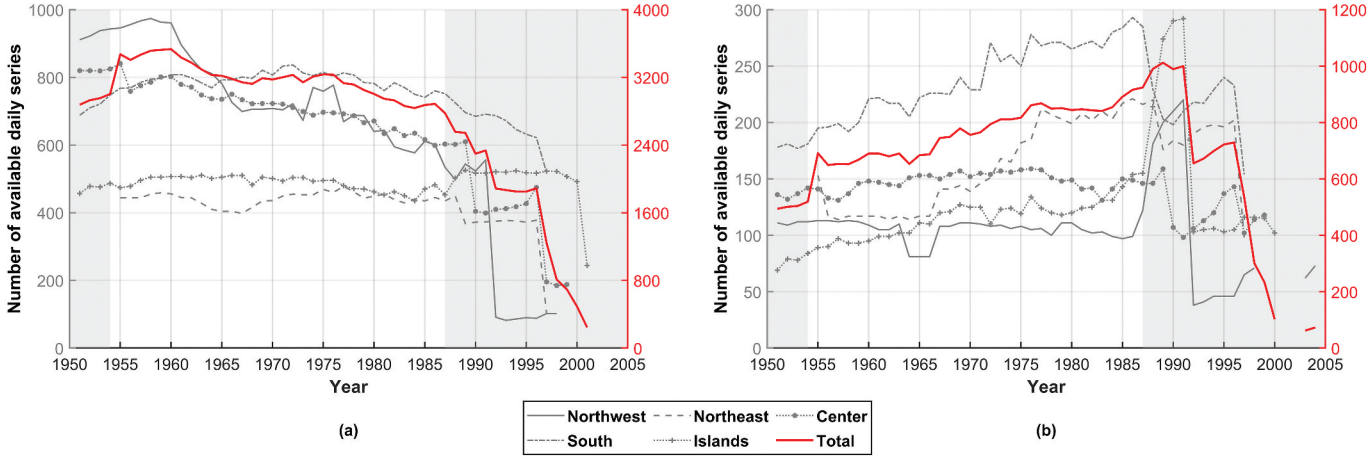


Figure 8. Yearly distribution of gauge stations with available time series of daily (a) rainfall and (b) maximum–minimum temperatures, covering the period from 1951 to the early 2000s, shown for the entire Italian territory and disaggregated by macro-areas: Northwest (Parma and Genova), Northeast (Venezia, Bolzano, and Trento), Center (Pisa, Bologna, and Roma), South (Napoli, Pescara, Bari, and Catanzaro), and Islands (Cagliari and Palermo). The right y-axis (red) refers exclusively to the thick red line representing the total number of stations. Grey-shaded areas indicate the periods when not all compartments were operational.

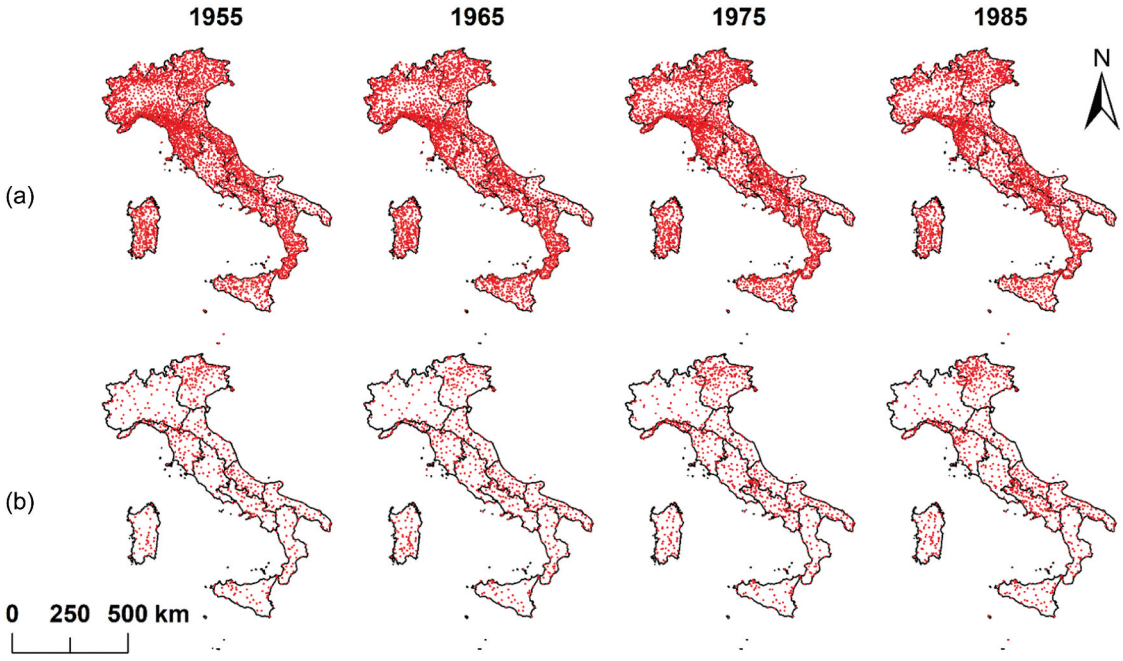


Figure 9. Spatial distribution of (a) precipitation and (b) temperature gauges in the ANNHYDRO database for different decades.

Table 2. Example of uncertain reporting of metadata in the Compartment of Napoli.

Year	Station name	Gauge type	Elevation (m a.s.l.)	Gauge height above ground (m)	Monitoring start year
1969	Napoli (Ist. Fisica Terrestre)	Pr	54	30.4	1908
1970	Napoli (Ist. Fisica Terrestre)	Pr	10	30.4	1908
1971	Napoli Ist. Fis. Terr.	P	40	30.4	1908

Table 3. Example of uncertain reporting of metadata in the Compartment of Parma.

Year	Station name	Gauge type	Elevation (m a.s.l.)	Gauge height above ground (m)	Monitoring start year
1951–1956	Sonico	Pn	1090	1.70	1928
1957–1958	Sonico	Pn	1080	1.70	1928
1960–1981 1985–1986	Sonico	P	1090	1.70	1928
1982–1984 1987–1991	Sonico	P	641	1.70	1928
1988 1990–1991	Sonico	Tm	641	2.00	1928

Assessing the reliability of such assumptions will require further analyses, particularly those aimed at the digitalization of SII/SIMN paper archives and the recovery of metadata information (Auer *et al.* 2001, Brönnimann *et al.* 2019). Careful management of this material is necessary for identifying and addressing potential sources of inhomogeneity in the recovered data, which may hinder the assessment of natural climate variability (Auer *et al.* 2005). Such studies are also necessary to explain potential inconsistencies in publication policies and data restitution among the local compartments of the SII (Auer *et al.* 2007). For example, drawing on insights from centuries-old time series, it will be important to investigate whether changes in stations – particularly gauge-orifice elevation – correspond to actual modifications in location and monitoring conditions (Böhm *et al.* 2001, Brunetti *et al.* 2004). These considerations are also applicable to the current state of the monitoring system, in order to prevent future issues in data availability arising from the fragmentation of the climate monitoring system among different local public institutions (Brunetti *et al.* 2006a).

Once this phase is completed, the yearly series will undergo a reconciling operation with the assigned aliases and the final time series will be correctly finalized.

3.2 Consistency check and final database generation

After dbv0 and dbv1 are organized in their final version, dbv1 underwent a quality assessment. This step was necessary to remove errors introduced during data digitalization, basically due to poor print quality or typographical mistakes. Owing to the different nature and aggregation of rainfall and temperature data, this operation was carried out according to these criteria:

Rainfall: comparison between the cumulative monthly rainfall obtained by summing the daily values and that printed in the original hydrological yearbooks;

Temperature: comparison between the monthly means of maximum and minimum temperatures and those printed in the original hydrological yearbooks.
In both cases, verification was performed through manual comparison of the two values and correction of notable errors contained in dbv1. However, we deem it important to highlight that a considerable number of errors were identified in the hydrological yearbooks too, most notably:

- The sum of the daily rainfall values does not match the corresponding monthly total (the same applies to the temperature mean), raising doubts regarding which value should be considered the true one.
- The number of days in some months does not correspond to the expected count (e.g. November with 31 values).
- Monthly statistics are affected by suspected outliers in daily values.
- The same daily data are assigned to different stations.

In such cases, a conservative approach that relied only on those daily values for the expected number of days in a month was adopted. Anomalous values, instead, were marked with a specific quality code. Proper investigation of such situations could be tackled only by analysing the paper archives of the SIMN. Although this operation requires a considerable amount of expertise, it will provide the chance to recover (at least) continuous records, also leveraging previous experiences in Italy (Deidda *et al.* 2007). Finally, dbv0 and dbv1 will be compared, detected inconsistencies will be fixed and the final version of the ANNHYDRO database will be produced.

Data will be made available through HIS Central portal, which is currently experiencing structural updates to enhance international accessibility. This repository will allow researchers worldwide to download complete daily rainfall and temperature time series with associated metadata and quality documentation.

Future work could also be addressed with respect to validation by means of a complete verification of dbv1, a comparison with similar datasets already digitalized and available online for users, in order to cross-validate all available information, and to provide a unique and consistent database of climate data at daily scale. Although manual procedures are recognized to be time consuming, we deem that human supervision is unavoidable for ensuring the effectiveness of the expected outputs (Ashcroft *et al.* 2018).

Although the ANNHYDRO database represents a significant step towards the systematic reconstruction of Italy's hydro-meteorological heritage, different sources of uncertainty can be acknowledged. First, metadata inconsistencies across years and compartments (e.g. changes in station names, elevations and aliases) introduce uncertainty in the station-matching procedure and may result in potential relocations. Second, georeferencing accuracy varies according to the availability and reliability of coordinate sources. Third, the original hydrological yearbooks contain occasional transcription or typographical errors, discrepancies between daily values and monthly aggregates, and incomplete months. The explicit retrieval, analysis and discussion of such issues do not compromise the value of the project, but, instead, can be considered an added value.

3.3 Potential applications

The availability of a complete national-scale rainfall and temperature database at daily scale is suitable for a high number of applications. Such databases have become almost essential in facilitating the implementation of large-scale studies based on climate data, such as climate model evaluation and hydrological applications, including the assessment of biases in major gridded products (e.g. E-OBS, ERA5-Land, WorldClim) as a function of historical observational density, as well as applications in statistical downscaling and homogenization of climate time series. Among these applications, the analysis of extreme events particularly benefits from the spatial completeness of the database, providing a more complete understanding of the physical events that occurred at the national scale, regardless of sub-national administrative limits (such as the clustering in compartments). Figure 10 shows an example of the monthly distribution of annual maxima of daily rainfall data for two compartments: Parma and Rome. In addition, setting high thresholds for rainfall data can highlight significant extreme events that occurred in the whole country, allowing researchers and stakeholders to gain spatially distributed information about these events at the national scale (Fig. 11). The ANNHYDRO database can also be used to retrieve the history of some historical floods that occurred in Italy. In Fig. 12, we list a sequence of some relevant events detected by applying the 200-mm threshold for daily rainfall.

3.4 Implications and lessons learned

The added value of this project, compared to other studies on the recovery of past climate data, lies also in its design aiming at preserving the full information and preservation on the sources and their information contents. Also, this project does not end with the mere digitalization of the daily data, but also provides a systematic and critical elaboration of the single time series aimed at organizing and providing homogeneous outputs. However, much work remains to be done. Despite this, this project has provided the basis for a number of considerations, both with respect to the Italian case and, more generally, regarding data rescue and collection activities. Here we list and discuss some of the most important points.

(1) *Published data as the tip of an iceberg*: printed hydro-meteorological data are often only a synthesis of wider monitoring activities/programmes. Existing archives contain much more information about measurements and network management than the synthesis published in printed volumes (Chimani *et al.* 2022, Ross and Packman 2025). For example, digitizing strips containing continuous measurements of rainfall or discharges can constitute a precious resource for improving the

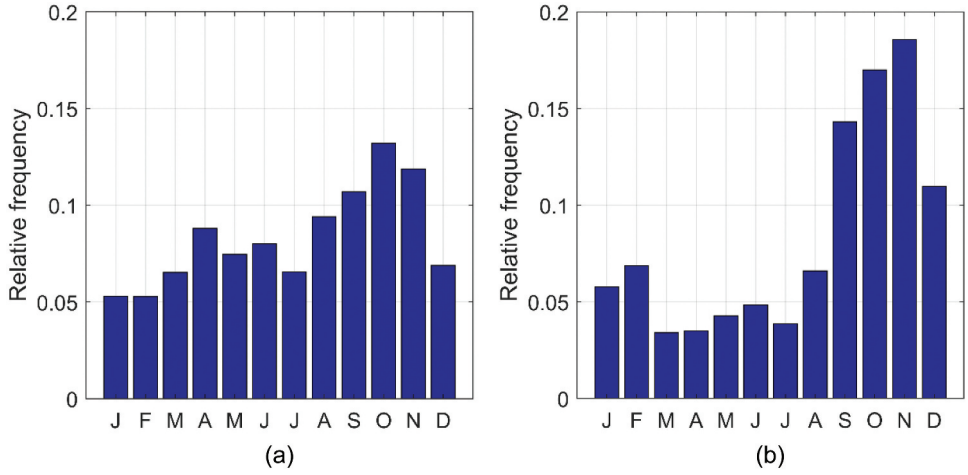


Figure 10. Monthly distribution of annual maxima of daily rainfall for the (a) Parma and (b) Roma compartments.

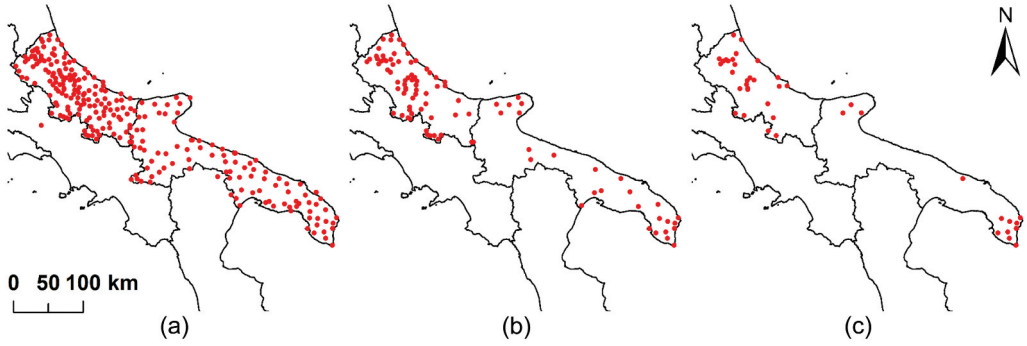
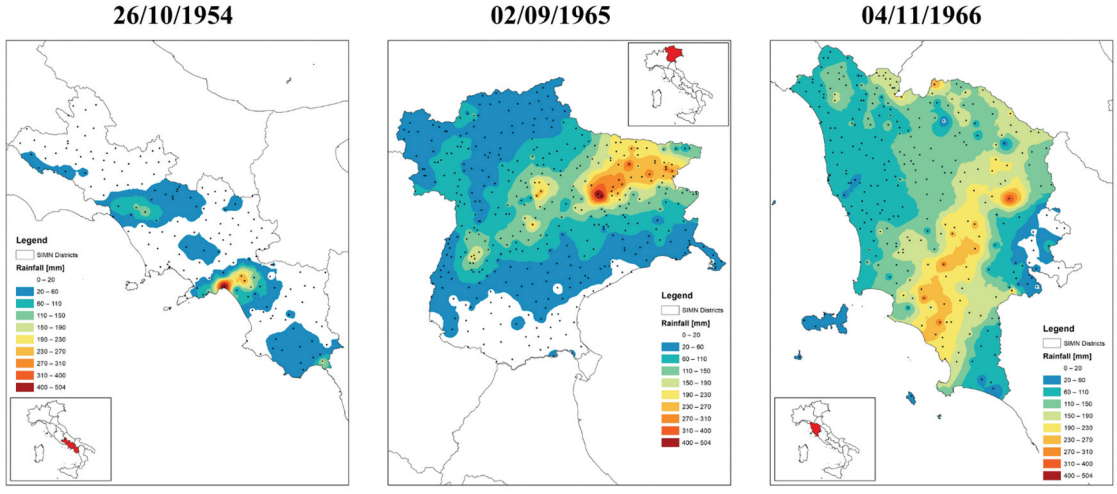


Figure 11. Spatial distribution of extreme daily rainfall exceeding the (a) 100 mm, (b) 150 mm, and (c) 200 mm threshold for the compartments of Bari and Pescara.



Highest daily rainfall records	Highest daily rainfall records	Highest daily rainfall records
Salerno: 504.0 mm	Diga Cellina: 500 mm	Badia Agnano: 338.7 mm
Gauro: 274.0 mm	Barcis: 500 mm	Grosseto: 291.8 mm
Giffoni Vallepiana: 247.2 mm	Villasantina: 347.3 mm	Lupo (Giuncarico): 290.0 mm

Figure 12. List of some notable exceptional rainfall events, their spatial distribution and the three highest recorded values for each event. All these events are associated with catastrophic floods, such as the well-known Florence flood of 1966.

knowledge of weather dynamics (Deidda *et al.* 2007, Wilkinson *et al.* 2019). Concerns about the deterioration of materials and storage conditions in libraries should highlight the need for a concerted effort aimed at collecting and preserving volumes containing historical climate information (Verticchio *et al.* 2021, 2022). This consideration also applies to the need to recover the full patrimony of former SII/SIMN, but also to the

considerable patrimony of information still unexplored (Brugnara *et al.* 2015).

(2) *Remove the (temporal) boundaries*. Data rescue activities are generally characterized by a relevant amount of work at the expense of uncertain results in terms of both data quality and speed (Brönnimann *et al.* 2006, Wilkinson *et al.* 2019, Manara *et al.* 2025b). Also, in the case where data rescue is just a step in a wider project activity, time and financial constraints have been demonstrated to impact the final result (Downs and Chen 2017, Luterbacher *et al.* 2024). In this sense, assigning temporal boundaries before the beginning of a systematic data rescue initiative can prove counterproductive, as a broader horizon may be needed which can overcome temporal and financial constraints. This can be also helpful for rationalizing potential funding sources. Partners of the ANNHYDRO project propose to continue the activities in the following years, with a specific focus on data validation and homogenization, trying to exploit rational and fruitful approaches carried out in Italy (Ceppi *et al.* 2026, Manara *et al.* 2025a, 2025b, Mazzoglio *et al.* 2025a). Integrating these ongoing individual efforts, which are often carried out through the commendable contribution of unfunded research, is crucial to avoid the multiplication of information and to build a truly collective and coherent national framework. Extending the data rescue to other data (e.g. sub-hourly rainfall extremes, historical hydrometric data) is an accessible perspective, as well as creating an inventory of volumes containing potentially useful data to be rescued and prioritizing future activities (Ashcroft *et al.* 2018, Wilkinson *et al.* 2019, Pappert *et al.* 2021).

(3) *Comparison of a centralized hydro-meteorological service with a distributed/federal one*. This point can be regarded in two different directions, the first being related to the past experience of SII/SIMN: a decentralized structure with a central office ensured a regular publication of climate data in a consolidated structure, even if (as shown in Section 3) differences in local compartment policies can be detected. The second direction is related to the near future: a fragmentated structure, although it has the advantage of monitoring a lower portion of territory, at the same time provides inhomogeneous data acquisition, storage and diffusion at national level, in its own way making digital data rescue necessary to implement large-scale studies at national and international levels (Braca *et al.* 2021, Avino *et al.* 2024, Cammalleri *et al.* 2024, Mazzoglio *et al.* 2025b). The experience of SII/SIMN needs to be projected in the future, demonstrating the need for a consolidated structure at the national level for the systematic collection and sharing of climatic data.

4 Conclusions

This paper highlighted the fundamental importance of *data rescue* for the valorization of historical hydrological *knowledge*. The methodological framework, along with the main phases and the implications of the retrieved database, presents what can be considered the first Italian national-scale product arising from a systematic analysis of daily data of rainfall and temperature printed in the hydrological yearbooks. Among the main results of the project, the rationalization of the pluviometric and thermometric heritage of the former SII/SIMN at the national scale stands out, along with the recognition of the substantial informational value still embedded in the original sources. These outcomes highlight the need for further and more detailed investigations aimed at fully exploiting this heritage, while acknowledging its relevance as a national scientific asset.

Despite the significant goals achieved, the outcomes of the ANNHYDRO project must be regarded as a foundation for further future developments rather than a finish line. Open questions emerged during the project, highlighting the need for common strategies and data management for current climate and hydrological monitoring, aimed at preserving the history of monitoring network management. In this sense, we claim the relevance of a more cautionary analysis and use of data archives, against the tendency to superficially obtain and analyse bigger datasets, whilst neglecting errors that may derive from inhomogeneous collecting methodologies (Morbidelli *et al.* 2021, Pelosi *et al.* 2022).

Building on the background developed within ANNHYDRO, the project provides a concrete basis for a renewed national interest in climate data collection and analysis, not only orientated towards the investigation of extremes but also addressing a wider spectrum of hydrological issues. The comparison with the outputs of past and ongoing activities, the harmonization of different datasets available online, the establishment of national commissions, and the digitalization of paper charts represent relevant avenues for future research and coordinated initiatives. The outcomes of this work highlight that the rationalization and valorization of historical climate data cannot be effectively addressed through isolated initiatives, but require coordinated and methodologically consistent efforts. Concerted efforts among stakeholders represent a necessary step towards defining a systematic approach for the rationalization and public dissemination of Italy's climate data heritage, providing a solid basis for future hydrological and climatological research.

Acknowledgements

The authors thank Giulia Acconciaioco, Luca De Carlo, Clemente Antonio Festa, Michele Gimma, Santo Grieco, Fausta Lorusso, Alessandro Mazzeo, Tiziana Piemontese and Francesco Rotunno for their support in creating the hydrological yearbook inventory and in the validation of the database.

Author contributions

CRediT: **Barbara Lastoria:** Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Supervision, Validation, Visualization, Writing – original draft; **Francesca Archi:** Data curation, Formal analysis, Investigation, Validation; **Matteo Balistrocchi:** Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Validation, Visualization, Writing – original draft; **Daniela Biondi:** Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Supervision, Validation, Visualization, Writing –

original draft; **Anna Botto:** Data curation, Formal analysis, Investigation, Validation; **Giovanni Braca:** Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Supervision, Validation, Visualization, Writing – original draft; **Paolo Colosio:** Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Validation, Visualization, Writing – original draft; **Francesca Conocchiella:** Data curation, Investigation, Validation, Visualization; **Francesco Cruscomagno:** Data curation, Formal analysis, Validation, Visualization; **Silvano Fortunato Dal Sasso:** Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Supervision, Validation, Visualization, Writing – original draft; **Mauro Fiorentino:** Conceptualization, Funding acquisition, Methodology, Project administration, Supervision, Writing – review & editing; **Vito Iacobellis:** Conceptualization, Funding acquisition, Methodology, Project administration, Visualization, Writing – review & editing; **Audrey Maria Noemi Martellotta:** Data curation, Formal analysis, Investigation, Validation, Visualization; **Federica Mesto:** Data curation, Formal analysis, Investigation, Validation; **Leonardo Mita:** Data curation, Investigation, Methodology, Validation; **Cosimo Peruzzi:** Data curation, Formal analysis, Methodology, Validation; **Teresa Pizzolla:** Data curation, Formal analysis, Investigation, Validation, Visualization, Writing – original draft; **Roberto Ranzi:** Conceptualization, Funding acquisition, Investigation, Methodology, Project administration, Supervision, Validation, Writing – original draft, Writing – review & editing; **Vincenzo Totaro:** Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Validation, Visualization, Writing – original draft; **Robertino Tropeano:** Data curation, Formal analysis, Investigation, Validation; **Piergiorgio Vasta:** Data curation, Investigation, Validation, Visualization; **Martina Bussettini:** Conceptualization, Data curation, Formal analysis, Funding acquisition, Methodology, Project administration, Supervision, Visualization, Writing – review & editing.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

The ANNHYDRO project was funded by the Italian National Recovery and Resilience Plan (NRRP) through the RETURN Extended Partnership [Mission 4, Component C2, Investment 1.3; CUP J33C22002840002].

ORCID

Barbara Lastoria http://orcid.org/0009-0006-3139-856X
Matteo Balistrocchi http://orcid.org/0000-0001-8302-1103
Daniela Biondi http://orcid.org/0000-0002-3570-9605
Giovanni Braca http://orcid.org/0000-0003-4687-0426
Paolo Colosio http://orcid.org/0000-0002-6426-780X
Silvano Fortunato Dal Sasso http://orcid.org/0000-0003-1376-7764
Mauro Fiorentino http://orcid.org/0000-0003-4250-9465
Vito Iacobellis http://orcid.org/0000-0002-8691-8271
Leonardo Mita http://orcid.org/0000-0003-1232-8225
Cosimo Peruzzi http://orcid.org/0000-0002-1418-9575
Teresa Pizzolla http://orcid.org/0000-0003-3665-5511
Roberto Ranzi http://orcid.org/0000-0002-7408-9891
Vincenzo Totaro http://orcid.org/0000-0003-2160-421X
Robertino Tropeano http://orcid.org/0009-0005-3190-9780
Martina Bussettini http://orcid.org/0000-0002-1346-749X

References

Ashcroft, L., *et al.*, 2018. A rescued dataset of sub-daily meteorological observations for Europe and the southern Mediterranean region, 1877–2012. *Earth System Science Data*, 10 (3), 1613–1635. doi:10.5194/essd-10-1613-2018

Auer, I., *et al.*, 2005. A new instrumental precipitation dataset for the greater alpine region for the period 1800–2002. *International Journal of Climatology*, 25, 139–166. doi:10.1002/joc.1135

Auer, I., *et al.*, 2007. HISTALP - Historical instrumental climatological surface time series of the Greater Alpine Region. *International Journal of Climatology*, 27 (1), 17–46. doi:10.1002/joc.1377

Auer, I., Böhm, R., and Schöner, W., 2001. Austrian long-term climate 1767–2000 (ALOCLIM): multiple instrumental climate time series from Central Europe. *Osterreichische Beiträge zu Meteorologie und Geophysik*, 25, 147.

Avino, A., *et al.*, 2024. Are rainfall extremes increasing in southern Italy? *Journal of Hydrology*, 631 (September 2023), 130684. doi:10.1016/j.jhydrol.2024.130684

Bacchi, B. and Kottegoda, N.T., 1995. Identification and calibration of spatial correlation patterns of rainfall. *Journal of Hydrology*, 165 (1–4), 311–348. doi:10.1016/0022-1694(94)02590-8

Balistrocchi, M. and Grossi, G., 2020. Predicting the impact of climate change on urban drainage systems in northwestern Italy by a copula-based approach. *Journal of Hydrology: Regional Studies*, 28, 100670. doi:10.1016/j.ejrh.2020.100670

Beltrano, M.C., Esposito, S., and Iafrate, L., 2012. The archive and library of the former Italian Central Office for Meteorology and Climatology. *Advances in Science and Research*, 8 (1), 59–65. doi:10.5194/asr-8-59-2012

Böhm, R., *et al.*, 2001. Regional temperature variability in the European Alps: 1760–1998 from homogenized instrumental time series. *International Journal of Climatology*, 21 (14), 1779–1801. doi:10.1002/joc.689

Bojinski, S., *et al.*, 2014. The concept of essential climate variables in support of climate research, applications, and policy. *Bulletin of the American Meteorological Society*, 95 (9), 1431–1443. doi:10.1175/BAMS-D-13-00047.1

Braca, G., *et al.*, 2013. Linee guida per l'analisi e l'elaborazione statistica di base delle serie storiche di dati idrologici (Guidelines for the analysis and basic statistical processing of hydrological time series). *ISPRA Manuals and Guidelines*, 84 (2013), 166 (in Italian).

Braca, G., *et al.*, 2021. Il Bilancio Idrologico Gis BAsed a scala Nazionale su Griglia regolare – BIGBANG: metodologia e stime. Rapporto sulla disponibilità naturale della risorsa idrica. *Rapporti*, n. 339 (2021), 113.

Brázdil, R., *et al.*, 2005. Historical climatology in Europe - The state of the art. *Climatic Change*, 70, 363–430. doi:10.1007/s10584-005-5924-1

Brönnimann, S., *et al.*, 2006. A guide for digitising manuscript climate data. *Climate of the Past*, 2 (2), 137–144. doi:10.5194/cp-2-137-2006

Brönnimann, S., *et al.*, 2019. Unlocking pre-1850 instrumental meteorological records a global inventory. *Bulletin of the American Meteorological Society*, 100, ES389–ES413. doi:10.1175/BAMS-D-19-0040.1

Brugnara, Y., *et al.*, 2015. A collection of sub-daily pressure and temperature observations for the early instrumental period with a focus on the 'year without a summer' 1816. *Climate of the Past*, 11 (8), 1027–1047. doi:10.5194/cp-11-1027-2015

Brunet, M., *et al.*, 2014. Data sources for rescuing the rich heritage of Mediterranean historical surface climate data. *Geoscience Data Journal*, 1 (1), 61–73. doi:10.1002/gdj3.4

Brunet, M. and Jones, P., 2011. Data rescue initiatives: bringing historical climate data into the 21st century. *Climate Research*, 47 (1), 29–40. doi:10.3354/cr00960

Brunetti, M., *et al.*, 2004. Changes in daily precipitation frequency and distribution in Italy over the last 120 years. *Journal of Geophysical Research: Atmospheres*, 109 (D5), 1–16. doi:10.1029/2003jd004296

Brunetti, M., *et al.*, 2006a. Precipitation variability and changes in the greater Alpine region over the 1800–2003 period. *Journal of Geophysical Research: Atmospheres*, 111. doi:10.1029/2005JD006674

Brunetti, M., *et al.*, 2006b. Temperature and precipitation variability in Italy in the last two centuries from homogenised instrumental time series. *International Journal of Climatology*, 26 (3), 345–381. doi:10.1002/joc.1251

Cammalleri, C., *et al.*, 2024. Testing trends in gridded rainfall datasets at relevant hydrological scales: a comparative study with regional ground

observations in Southern Italy. *Journal of Hydrology: Regional Studies*, 55 (August), 101950. doi:10.1016/j.ejrh.2024.101950

Camuffo, D., 2002. History of the long series of daily air temperature in Padova (1725–1998). *Climatic Change*, 53 (1–3), 7–75. doi:10.1023/A:1014958506923

Camuffo, D., Becherini, F., and Della Valle, A., 2025. Recovery of the long series of precipitation in Pisa, Italy: trend, anomaly and extreme events. *Climate*, 13 (4), 1–38. doi:10.3390/cli13040073

Ceppi, A., *et al*., 2026. Digitizing historical daily weather bulletins through citizen scientists: the ReData project. *PLOS Climate*, 5 (3), 1–24. doi:10.1371/journal.pclm.0000865

Chimani, B., *et al*., 2022. Data rescue in selected countries in connection with the EUMETNET DARE activity. *Geoscience Data Journal*, 9 (1), 187–200. doi:10.1002/gdj3.128

Ciccarelli, N., *et al*., 2008. Climate variability in north-western Italy during the second half of the 20th century. *Global and Planetary Change*, 63 (2–3), 185–195. doi:10.1016/j.gloplacha.2008.03.006

De Marchi, G., 1931. Compiti ed attività del Servizio. *In*: *Il Servizio Idrografico Italiano*. Rome, Italy: Ministero dei Lavori Pubblici, Tipografia del Senato, 7–66 (in Italian).

Deidda, R., *et al*., 2007. An automatic system for rainfall signal recognition from tipping bucket gage strip charts. *Journal of Hydrology*, 333 (2–4), 400–412. doi:10.1016/j.jhydrol.2006.09.011

Deidda, R., *et al*., 2013. Regional climate models' performance in representing precipitation and temperature over selected Mediterranean areas. *Hydrology and Earth System Sciences*, 17 (12), 5041–5059. doi:10.5194/hess-17-5041-2013

Downs, R.R. and Chen, R.S., 2017. Curation of scientific data at risk of loss: data rescue and dissemination. Curating Res. data Vol. 1 Pract. Strateg. your Digit. Repos. 263–278.

Engström, J.E., *et al*., 2023. Data rescue of historical wind observations in Sweden since the 1920s. *Earth System Science Data*, 15 (6), 2259–2277. doi:10.5194/essd-15-2259-2023

Eredia, F., 1908. Le precipitazioni atmosferiche in Italia dal 1880 al 1905. *Annali dell'Ufficio Centrale di Meteorologia e Geodinamica*, Serie II, vol XXV. Rome, Italy: Civelli, 313 (in Italian).

Eredia, F., 1919. *Osservazioni pluviometriche raccolte a tutto l'anno 1915 dal R. Ufficio Centrale di Meteorologia e Geodinamica*. Ministero dei Lavori Pubblici (Servizio Idrografico Centrale), Rome, Italy: Tipografia Nazionale Bertero (in Italian).

Fatichi, S. and Caporali, E., 2009. A comprehensive analysis of changes in precipitation regime in Tuscany. *International Journal of Climatology*, 29 (13), 1883. doi:10.1002/joc.1921

Ferrari, E. and Versace, P., eds., 1994. La Valutazione delle Piene in Italia. *Rapporto Nazionale di Sintesi*. Rome, Italy: Consiglio Nazionale delle Ricerche (CNR), Gruppo Nazionale per la Difesa dalle Catastrofi Idrogeologiche (GNDCI), 484 (in Italian).

Gourley, J.J., *et al*., 2010. Intercomparison of rainfall estimates from radar, satellite, gauge, and combinations for a season of record rainfall. *Journal of Applied Meteorology and Climatology*, 49 (3), 437–452. doi:10.1175/2009JAMC2302.1

Grimaldi, S., *et al*., 2011. 2.18 – statistical Hydrology. *In*: P. Wilderer, ed. *Treatise on water science*. Oxford: Elsevier, 479–517. doi:10.1016/B978-0-444-53199-5.00046-4

Hawkins, E., *et al*., 2023. Millions of historical monthly rainfall observations taken in the UK and Ireland rescued by citizen scientists. *Geoscience Data Journal*, 10 (2), 246–261. doi:10.1002/gdj3.157

Hersbach, H., *et al*., 2020. The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, 146 (730), 1999–2049. doi:10.1002/qj.3803

Huffman, G.J., *et al*., 2009. The TRMM multi-satellite precipitation analysis (TMPA). *In*: M. Gebremichael and F. Hossain, eds. *Satellite rainfall applications for surface hydrology*. Dordrecht: Springer, Netherlands, 3–22.

Izzaddin, A., *et al*., 2024. A new diagram for performance evaluation of complex models. *Stochastic Environmental Research and Risk Assessment*, 38 (6), 2261–2281. doi:10.1007/s00477-024-02678-3

Karl, T.R., *et al*., 1989. The recent climate record: what it can and cannot tell us. *Reviews of Geophysics*, 27 (3), 405–430. doi:10.1029/RG027i003p00405

Koutsoyiannis, D. and Mamassis, N., 2021. From mythology to science: the development of scientific hydrological concepts in Greek antiquity and its relevance to modern hydrology. *Hydrology and Earth System Sciences*, 25 (5), 2419–2444. doi:10.5194/hess-25-2419-2021

Lastoria, B., *et al*., 2025. Un passo avanti verso la ricostruzione dell'informazione termopluviometrica storica italiana degli annali idrologici: il progetto ANNHYDRO. *In*: V. Iacobellis, V. Totaro, V and E. Toth, eds. *Atti del Le Giornate dell'Idrologia della Società Idrologica Italiana 2025*, 156–157. doi:10.5281/zenodo.17298436

Libertino, A., Ganora, D., and Claps, P., 2018. Technical note: space-time analysis of rainfall extremes in Italy: clues from a reconciled dataset. *Hydrology and Earth System Sciences*, 22 (5), 2705–2715. doi:10.5194/hess-22-2705-2018

Longobardi, A. and Villani, P., 2010. Trend analysis of annual and seasonal rainfall time series in the Mediterranean area. *International Journal of Climatology*, 30 (10), 1538–1546. doi:10.1002/joc.2001

Luterbacher, J., *et al*., 2024. The importance and scientific value of long weather and climate records; Examples of historical marine data efforts across the globe. *Climate*, 12 (3), 39. doi:10.3390/cli12030039

Manara, V., *et al*., 2025a. Engaging high school students in rescuing and digitizing data from historical observations in Italy: the citizen science project Cli-DaRe@School. *Bulletin of the American Meteorological Society*, 106 (3), E509–E524. doi:10.1175/BAMS-D-24-0078.1

Manara, V., *et al*., 2025b. A new dataset of Italian precipitation records for the period 1921–1950 from the Cli-DaRe@School citizen science project. *Bulletin of Atmospheric Science and Technology*, 6 (1), 25. doi:10.1007/s42865-025-00111-3

Marani, M. and Zanetti, S., 2015. Long-term oscillations in rainfall extremes in a 268 year daily time series. *Water Resources Research*, 51 (1), 639–647. doi:10.1002/2014WR015885

Mazzoglio, P., *et al*., 2025a. Who is saving our streamflow data? Exploring volunteer profiles and their engagement in the SIREN data rescue project. *PLOS ONE*, 20 (10), 1–19. doi:10.1371/journal.pone.0333091

Mazzoglio, P., *et al*., 2025b. Mapping the uneven temporal changes in ordinary and extraordinary rainfall extremes in Italy. *Journal of Hydrology: Regional Studies*, 58, 102287. doi:10.1016/j.ejrh.2025.102287

Menne, M.J., *et al*., 2012. An overview of the global historical climatology network-daily database. *Journal of Atmospheric and Oceanic Technology*, 29 (7), 897–910. doi:10.1175/JTECH-D-11-00103.1

Merino, A., *et al*., 2021. Evaluation of gridded rain-gauge-based precipitation datasets: impact of station density, spatial resolution, altitude gradient and climate. *International Journal of Climatology*, 41 (5), 3027–3043. doi:10.1002/joc.7003

Ministero Dei Lavori Pubblici, and Servizio Idrografico, 1976. Elenco delle stazioni termopluviometriche del Servizio Idrografico Italiano (situazione al 1970). *Pubblicazione n. 27 del Servizio*. Rome, Italy: Istituto Poligrafico dello Stato – Libreria, 266 (in Italian).

Ministero Dei Lavori Pubblici and Servizio Idrografico, 1980. Dati caratteristici dei corsid'acqua italiani. *Pubblicazione n. 17 del Servizio*, 5th ed. Rome, Italy: Istituto Poligrafico dello Stato (in Italian).

Moccia, B., *et al*., 2021. Probability distributions of daily rainfall extremes in Lazio and Sicily, Italy, and design rainfall inferences. *Journal of Hydrology: Regional Studies*, 33, 100771. doi:10.1016/j.ejrh.2020.100771

Moccia, B., *et al*., 2025. What is our pick? Assessment of satellite and reanalysis precipitation datasets over Italy. *Journal of Hydrology: Regional Studies*, 60, 102487. doi:10.1016/j.ejrh.2025.102487

Montanari, A., *et al*., 2023. Why the 2022 Po River drought is the worst in the past two centuries. *Science Advances*, 9 (32), eadg8304. doi:10.1126/sciadv.adg8304

Morbidelli, R., *et al*., 2021. Effect of time-resolution of rainfall data on trend estimation for annual maximum depths with a duration of 24 hours. *Water (Switzerland)*, 13, 8–10. doi:10.3390/w13223264

New, M., *et al*., 2001. Precipitation measurements and trends in the 20th century. *International Journal of Climatology*, 21 (15), 1889–1922. doi:10.1002/joc.680

Newman, A.J., *et al*., 2019. Methodological intercomparisons of station-based gridded meteorological products: utility, limitations,

and paths forward. *Journal of Hydrometeorology*, 20 (3), 531–547. doi:10.1175/JHM-D-18-0114.1

Noone, S., *et al*., 2024. Investigating the potential for students to contribute to climate data rescue: introducing the Climate Data Rescue Africa project (CliDaR-Africa). *Geoscience Data Journal*, 11 (4), 758–774. doi:10.1002/gdj3.248

Pappert, D., *et al*., 2021. Unlocking weather observations from the Societas Meteorologica Palatina (1781–1792). *Climate of the Past*, 17 (6), 2361–2379. doi:10.5194/cp-17-2361-2021

Pelosi, A., *et al*., 2022. Regional assessment of sub-hourly annual rainfall maxima. *Water (Switzerland)*, 14, 1–14. doi:10.3390/w14071179

Ranzi, R., *et al*., 2021. A multi-century meteo-hydrological analysis for the Adda river basin (Central Alps). Part II: daily runoff (1845–2016) at different scales. *International Journal of Climatology*, 41 (1), 181–199. doi:10.1002/joc.6678

Ranzi, R., *et al*., 2025. A multi-century meteo-hydrological analysis in the Italian Alps: daily streamflow (1862–2022) at different time scales. *Hydrological Processes*, 39. doi:10.1002/hyp.70305

Ross, C. and Packman, S., 2025. Rainfall data rescue activities at the UK met office national meteorological archive - current activities, use in climate monitoring, and next steps. *Proceedings of Science*. 490. 10.22323/1.490.0358

Sarigil, G., Neri, M., and Toth, E., 2024. Evaluation of national and international gridded meteorological products for rainfall-runoff modelling in Northern Italy. *Journal of Hydrology: Regional Studies*. 56. 10.1016/j.ejrh.2024.102031

Serrano-Notivoli, R. and Tejedor, E., 2021. From rain to data: a review of the creation of monthly and daily station-based gridded precipitation datasets. *WIREs Water*, 8 (6). doi:10.1002/wat2.1555

Tepetidis, N., *et al*., (2026). *From deterministic to stochastic: a hybrid machine learning framework for reliable satellite precipitation merging over Greece*. Submitted manuscript. Available at SSRN 6574189. Available from: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=6574189 [Accessed 12 Oct 2026].

Testa, O., *et al*., 2006. A study of rainfall in the Roman area in the years 1951–2000. *Nuovo Cim. della Soc. Ital. di Fis. C*, 29, 191–213. doi:10.1393/ncc/i2005-10176-3

Verticchio, E., *et al*., 2021. Climate-induced risk for the preservation of paper collections: comparative study among three historic libraries in Italy. *Building and Environment*, 206, 108394. doi:10.1016/j.buildenv.2021.108394

Verticchio, E., *et al*., 2022. Conservation risks for paper collections induced by the microclimate in the repository of the Alessandrina Library in Rome (Italy). *Heritage Science*, 10 (1), 1–15. Springer International Publishing. doi:10.1186/s40494-022-00714-5

Volpi, E., Mancini, C.P., and Fiori, A., 2024. What can we learn from long hydrological time-series? The case of rainfall data at Collegio Romano, Rome, Italy. *Journal of Hydrology X*, 23, 100176. doi:10.1016/j.hydroa.2024.100176

Wang, X., *et al*., 2023. Traditional and novel methods of rainfall observation and measurement: a review. *Journal of Hydrometeorology*, 24 (12), 2153–2176. doi:10.1175/JHM-D-22-0122.1

Wilkinson, C., *et al*., 2019. Best practice guidelines for climate data rescue v1, of the copernicus climate change service data rescue service (March), 49.

WMO, 2007. Guidelines on climate data management. World Meterological Organization, Geneva, WCDMP-no. 60, WMO-TD no. 1376, 67 pp.

WMO. (2024) *Guidelines on good practices for data rescue*. Availalble from: https://library.wmo.int/idurl/4/55395 [Accessed 12 Oct 2026].

Zeng, Y., *et al*., 2019. Towards a traceable climate service: assessment of quality and usability of essential climate variables. *Remote Sensing*, 11 (10), 1186. doi:10.3390/rs11101186