POLITECNICO DI TORINO
Repository ISTITUZIONALE

Measuring floods by their scars: toward a data-driven flood severity scale

Original
Measuring floods by their scars: toward a data-driven flood severity scale / Giacalone, A., D'Andrilli, A., Mazzoglio, P., Claps, P.. - In: NATURAL HAZARDS. - ISSN 1573-0840. - ELETTRONICO. - 122:(2026). [10.1007/s11069-026-08383-4]

Availability:
This version is available at: 11583/3015330 since: 2026-09-10T07:19:25Z

Publisher:
Springer

Published
DOI:10.1007/s11069-026-08383-4

Terms of use:

This article is made available under terms and conditions as specified in the corresponding bibliographic description in the repository

Publisher copyright

(Article begins on next page)

16 November 2026

Natural Hazards (2026) 122:632
https://doi.org/10.1007/s11069-026-08383-4

ORIGINAL PAPER

Measuring floods by their scars: toward a data-driven flood severity scale

Alessandro Giacalone[1] · **Anna D'Andrilli**[1] · **Paola Mazzoglio**[1] · **Pierluigi Claps**[1]

Received: 12 February 2026 / Accepted: 8 August 2026
© The Author(s) 2026

Abstract
To classify and compare extreme events, it is useful to rely on metrics that capture their impacts. In the natural hazard field, widely accepted impact-based classification systems already exist, such as the Fujita scale for tornadoes, the TORRO scale for hail, and the Mercalli scale for earthquakes. While magnitude-based measures like the Richter scale have become standard for seismic events, phenomena such as tornadoes and hailstorms are still most effectively characterized through the impact rather than through direct physical measurements. For flood events, assessing and comparing their "magnitude" is often problematic, as peak discharge data are frequently unavailable, incomplete, or difficult to relate to a meaningful river section. Moreover, even with available precipitation data, it is known that similar rainfall amounts can produce very different effects in distinct environments. For these reasons, this paper suggests the introduction of a flood severity (FLOSEV) scale. The FLOSEV scale is devised using the historical information from the AVI (Aree Vulnerate Italiane) project, an extensive collection of flood and landslide events in Italy spanning centuries but with a unique systematic inventory from the early years of 1900. Each AVI record contains detailed information on affected municipalities, number of fatalities, and damage to properties, here used to suggest an approximate log-scale classification of flood events based on five severity levels, ranging from minor disruptions to widespread destruction. The proposed criteria have been applied to recent major events and appear to attribute them to reasonable classes, which supports the robustness and practical relevance of the scale.

Keywords Floods · Classification · Damage · Impact · Database · Italy

✉ Paola Mazzoglio
paola.mazzoglio@polito.it

[1] Department of Environment, Land and Infrastructure Engineering, Politecnico di Torino, Torino, Italy

Published online: 09 September 2026

1 Introduction

1.1 Challenges in flood documentation and event severity assessment

Floods are among the most frequent and devastating natural disasters worldwide (UNDRR and CRED 2020). They can result in significant loss of life, widespread destruction of infrastructure and property, and substantial damage to the environment and ecosystems (Douben 2006; Cunado and Ferreira 2014; Mazzoglio et al. 2019; Merz et al. 2021). Given their profound societal, economic, and environmental impacts, floods have long been a subject of human concern. The inherent human interest in documenting floods reflects both the immediacy of their impact and the sustained need to learn from them (Bonini 1666; Bersani and Bencivenga 2001; Masoero et al. 2013; Grazzini et al. 2020) also to convey the related experience in increased resilience of communities (Mazzoglio et al. 2021). By preserving knowledge of such events, societies aim to protect lives, safeguard resources, and ensure a more sustainable coexistence with natural hazards.

Unfortunately, most of the historical information about floods of the last century is of poor quality or, sometimes, almost inaccessible. This is largely because, until relatively recent times, no systematic efforts were made to collect and preserve such data (Brázdil et al. 2006; Claps et al. 2024; Mazzoglio et al. 2025). Yet, historical flood records are invaluable for comparing past events with modern ones, analyzing the spatio-temporal evolution of flooding phenomena, and assessing whether floods have become more severe over time, particularly in the context of the ongoing global warming, even if performing such comparisons is inherently challenging (Barredo 2007; Bertola et al. 2020, 2021). One major issue in these assessments lies in the different amounts and quality of information available for historical versus modern events, as well as between bigger and smaller events (Berz 2000). While recent floods are well-documented thanks to advances in technology, monitoring systems, and reporting mechanisms, historical records often rely on fragmented and uncertain documentation despite, in recent years, efforts have been dedicated to reconstructing historical information (see, e.g., Barredo 2007; Gaume et al. 2009; Payrastre et al. 2011; Archer and Fowler 2021; Papagiannaki et al. 2022). Recent developments in natural language processing (NLP) and large language models (LLMs) have begun to address some of these limitations, offering new ways to extract structured impact information from unstructured sources such as newspaper articles and social media (see e.g., Lai et al. 2022; Sodoge et al. 2023, 2024; de Brito et al. 2025; Li et al. 2026). Although these approaches are proving to substantially reduce the manual effort required to build impact databases, their application is currently limited to online data related to the last few decades.

When evaluating the severity of a flood event, a multitude of factors must be carefully considered, including: hydrological data, (water levels, flow rates, and rainfall intensity), impact on human life (number of fatalities, injuries, and displaced people), structural damage (destruction of infrastructure, homes, and assets), geographic extent, economic impact and environmental consequences (effects on ecosystems, land use, and water quality). Information on all these factors is not always available, particularly in the case of hydrological observations, whose accessibility, standardization, and comparability may remain limited or uneven across datasets and regions (Addor et al. 2019). The complexity of the information needed underscores the importance of devoting as much effort as possible to comprehensively classify historical flood data. In light of these gaps, a valuable tool for comparing

flooding events, even across different periods, could be a numerical indicator, simple enough to calculate through impact assessment, that can quantify the magnitude of an event.

For hydrological variables like precipitation and discharge, it is relatively straightforward to construct numerical indicators that measure event severity. Examples include the relative severity (also known as growth factor or extreme rain multiplier, see Laver et al. 2025), which compares the magnitude of an event X_{oss} to a baseline (usually the mean value, $\overline{X}$). More used is the return period, which estimates how often an event of such magnitude is statistically expected to occur (Fernandez and Salas 1999a, 1999b; Volpi 2019; Claps et al. 2022) but with highly questionable robustness (see e.g. Koutsoyiannis 2007). For comparison purposes, in fact, the use of the relative severity has the advantage that average values used as baseline can be reliable even for short records and are the most robust property of a sample. Furthermore, the concept of a return period is not always easily understood by the general population (Bell and Tobin 2007).

However, hydrological indicators are not always well correlated to the actual flood impact. Areas where hydrological events are frequent and characterized by high absolute values of rainfall and discharge tend to develop resilience over time, investing in structural protection measures and adapting land use practices, and therefore experience less damage even during extreme events (see e.g. the two record-breaking extreme rainfall events occurred in 2011 and 2021 in Liguria region, Italy; Mazzoglio 2022). Conversely, moderate hydrological events can cause disproportionate damage in areas with low adaptive capacity, poor infrastructure, or high exposure of population and assets. This decoupling between event intensity and societal impact means that hydrological metrics alone are insufficient to characterize the severity of a flood event from a disaster risk perspective. On the other hand, developing indicators based on the impact of hydrometeorological events requires integrating diverse and often incomplete types of data, such as human casualties, infrastructure damage and economic losses. The heterogeneity of these data sources, their varying levels of completeness across regions and time periods, and the difficulty of standardizing them represent significant practical challenges. Nevertheless, once a systematic way to gather and harmonize these indicators is established, impact-based metrics offer a crucial advantage: they allow for the comparison of events in situations where traditional severity metrics (e.g., return period or growth factor) are difficult or impossible to calculate. This can be due to limited data availability or to the lack of hydrological information, caused by the absence of monitoring instruments or to a difficulty even to determine a reference river section, as in the case of inundations caused by the collapse of infrastructures. In such contexts, an impact-based scale not only complements hydrological assessments but may represent the only viable approach to classification and comparison. By applying consistent classification criteria across events spanning more than a century (including periods and regions where instrumental data are absent or unreliable) it becomes possible to identify areas of persistent vulnerability, track long-term changes in flood risk exposure, and evaluate the effectiveness of mitigation measures over extended time horizons. This temporal depth and methodological consistency constitute the primary added value of the scale over existing classification frameworks. In this regard, the experience accumulated with impact-based scales for other natural hazards, like the Mercalli scale for earthquakes, the Fujita scale for tornadoes and the TORRO scale for hailstorms, illustrates concretely what such frameworks make possible. These scales share three key advantages: they enable intuitive communication of event severity also to non-specialist audiences, including the general public, media,

and civil protection authorities; they provide a common basis for comparing events across different locations and time periods, regardless of the availability of instrumental measurements; and they support operational decision-making by linking severity levels to response protocols and resource allocation. All three advantages are directly transferable to the flood context. In the following subsection, a resume and a comparison of these three important scales for natural hazards are presented. The analysis of these examples will drive us in a procedure for the selection of a suitable severity scale for floods.

1.2 Damage scales in use for natural hazards

The Mercalli scale is one of the most known systems to measure an earthquake's intensity, as it quantifies the severity according to the human perception and the observable effects of the event on people, objects and structures (Blong 2003; Tertulliani 2019). It was introduced by Giuseppe Mercalli in 1902 and it significantly differs from magnitude scales such as Richter, that estimate the energy released from an earthquake at the hypocenter. The first scale proposed by Mercalli was an attempt to improve the De Rossi-Forel scale, a ten-level effect-based scale created in 1881, which Mercalli did not find nuanced enough for an accurate description of events, especially of the more intense ones. Over the subsequent decades, the Mercalli scale evolved further. Recognizing the need for finer distinctions after major earthquakes, seismologists like Cancani and Sieberg expanded it to 12 levels, leading to the Mercalli-Cancani-Sieberg (MCS) scale. Later adaptations, most notably the Modified Mercalli Intensity (MMI) scale developed by Wood and Neumann in the US, adjusted the descriptions to better reflect contemporary building practices. Today, the MMI scale is widely used internationally, particularly in North America, while Europe often utilizes the European Macroseismic Scale (EMS-98; Grünthal 1998).

The Fujita scale is a system developed in 1971 by Tetsuya Fujita to classify the intensity of a tornado based on the observable damage to structures, where damages are used to estimate the relative wind speed that caused it (Fujita 1971). The primary reason why it was created in this way lies in the inherent difficulty of accurately measuring wind speeds and the need to quickly classify, within a certain range of precision, a large number of events. It is divided into 6 levels, from F0 to F5, each of which is coupled with a description of the damage. This system was almost immediately adopted by the meteorological and engineering community, which quickly noticed the limitations of the scale, i.e. possible inconsistencies in the rating of a tornado, the overrating of the wind speed associated with some descriptors of damage, and the variability caused by the different resistance of structures. Thus, on the basis of the Fujita scale, a new and improved measuring system was created: the Enhanced Fujita scale. This new version enables a greater precision in the classification of an event, introducing 28 Damage Indicators (DI), each coupled with a range of Degrees of Damage (DoD). For every DoD, estimated upper and lower bounds for wind speed were assessed through a process of expert elicitation, which involved a group of experts making their best estimate of the expected, upper and lower bound wind speeds to cause each DoD and calculating the mean and standard deviation of the values. Like the original one, this new scale is also divided into 6 levels, from EF0 to EF5, each of which is assigned an estimated wind speed range.

Finally, with reference to hailstorms, we can refer to the TORRO scale (Webb et al. 1986). It was developed in 1986 by the TORRO (Tornado and Storm Research Organiza-

tion) and it classifies hailstorms into eleven different levels, from H0 to H10, based on both the hail diameter as well as the observable damage (https://www.torro.org.uk/research/hail/hscale, accessed on 26/05/2026). The scale classifies the size of the hailstones in 10 size codes, ranging from size code 1 (pea-sized, 5 mm in diameter), to size code 10 (melon-sized, over 100 mm in diameter), to which typical damage impacts are associated. The value of the TORRO scale lies in the nuance it offers, allowing the distinction between hailstorms of similar size based on the observable damage.

The review of these three scales yields several insights that could directly inform the design of a similar classification for floods. First, all three frameworks share a common feature: they classify events based on observable consequences rather than on physical measurements that require instrumentation. Second, the evolution from the original Fujita scale to the Enhanced Fujita scale illustrates a recurring challenge in impact-based classification: the need to balance simplicity with discriminatory power. The introduction of multiple Damage Indicators and Degrees of Damage in the EF scale increased precision but also data requirements. Third, the number of levels across these scales ranges from 6 (Fujita) to 12 (Mercalli), with Blong (2003) suggesting that 5 to 13 levels represent a reasonable range used in damage intensity scales.

1.3 Elements to devise a flood severity scale

Some attempts to formulate impact-based severity classifications for flood events have been proposed in the past, although their application has often been limited to specific flood types or operational contexts. For instance, a Flash Flood Severity Index (FFSI) has been introduced to categorize flash flood events based on observed impacts on buildings, infrastructure, and transportation networks, with severity levels ranging from minor to catastrophic (Schroeder et al. 2016). This type of classification has proven useful for risk communication. However, by construction, it focuses on flash floods only and does not explicitly account for the spatial extent of events or for detailed information on affected populations, which are often less relevant for localized, short-duration phenomena. Diakakis et al. (2020) proposed a 10-level flash flood impact severity scale based on impacts on the built environment, on the man-made mobile objects, on the natural environment, and on the population. While this classification is highly detailed and capable of capturing fine distinctions among events, its systematic application, especially to historical datasets, is challenging. The scale relies on high-resolution impact information, including the local occurrence of fatalities (e.g., within the same building, campsite, vehicle, convoy of vehicles, or specific public spaces). Such detailed spatial information is rarely available. As a result, despite its conceptual robustness, this approach is difficult to apply consistently when working with long-term or heterogeneous flood archives.

Building upon these previous efforts, but aiming for a broader applicability, this paper develops and proposes a flood severity (FLOSEV) scale designed to classify flood events of different types and spatial scales. The proposed scale is conceptually similar to established impact-based systems such as the Mercalli, Fujita, and TORRO scales used for other natural hazards, but it is specifically tailored to the characteristics of flood events and the type of information commonly available in historical records. To this end, we use the AVI (*Aree Vulnerate Italiane*—Areas Affected by Landslides or Floods in Italy) database as the primary source of information. To our knowledge, AVI represents the most comprehensive

European dataset of its kind, providing systematic, event-level impact records over a period exceeding one century. This unique combination of detail and temporal depth makes it the most suitable foundation available for developing and calibrating an impact-based severity scale. After retrieving and standardizing all available records, we have processed the data to resolve inconsistencies, merge entries referring to the same event, and extract the impact variables relevant to our analysis. Based on these fields, we have developed a set of criteria to classify each event into five severity levels. The resulting scale is then tested on more than a century of events to assess its coherence, distribution, and applicability. It should be emphasised, however, that AVI serves here just as a comprehensive information source used for calibration: the scale is not intended as an Italy-specific classification, but as a general framework applicable to flood events in any geographical context, when analogous impact data are available, with the possibility of geographic fine-tuning of the thresholds, if needed. In any case, we demonstrate the application using two recent, well-documented floods to evaluate if the FLOSEV scale correctly rates their relevance in contemporary and well-documented contexts. With this work, we aim to demonstrate that an impact-based approach can complement traditional physical metrics, provide a quantitative framework for classifying floods, and enable historical and comparative analyses of flood severity where instrumental data are lacking. Our approach shifts the focus from the physical magnitude of the event to its tangible consequences on society and infrastructure.

The article is structured as follows. The dataset used in this work is described in Sect. 2; the methodology is summarized in Sect. 3, while the application and calibration of the scale are reported in Sect. 4. The results are discussed in Sect. 5, which also includes the application of the scale to two recent flood events. Finally, Sect. 6 provides the conclusions.

2 Data

Italy represents an ideal case for developing and testing a flood classification scale. The country is particularly prone to flooding and hydrogeological instabilities: overall, 94% of Italian municipalities are at risk of landslides, floods and/or coastal erosion, with an estimated 1.3 million residents exposed to landslide risk and 6.8 million exposed to flood risk (Trigila et al. 2021). Secondly, Italy benefits from the availability of different catalogues reporting flooding events: while some cover the entire national territory (Guzzetti et al. 1994, 2005; Guzzetti and Tonelli 2004), others include information only for limited areas (Govi and Turitto 1994; Monticelli 1998; Bencivenga et al. 1999; Vennari et al. 2016) or refer to the whole world, including Italy (Delforge et al. 2025).

In this work, we selected the AVI (*Aree Vulnerate Italiane*—Areas Affected by Landslides or Floods in Italy; Guzzetti et al. 1994; Guzzetti et al. 2005) catalogue as the main source for its high level of detail in the information reported for historical geo-hydrological events. The AVI project was commissioned in 1989 by the Department of Civil Protection to the National Group for Prevention of Hydro-geologic Catastrophes (GNDCI) of the Italian National Research Council (CNR—*Consiglio Nazionale delle Ricerche*). It aimed to build an inventory of information on areas historically affected by landslides and floods in Italy, and to gather data about these natural disasters (Guzzetti et al. 1994). The primary objective of this project was the mapping of geo-hydrologic risk for the entire nation based on histori-

cal events. The work was led by a Central Coordination Unit responsible for the direction of 14 regional teams and 2 CNR support and analysis groups.

The first phase of this research (1990 to 1992) consisted in the collection of information from interviews, newspaper articles and technical reports. This information was organized in four different types of records: single articles or sources of information (S0), source forms, i.e. collection of data relative to the same event coming from the same source (S1), and event forms, collecting all the available information for a certain event (S2). Additionally, for selected events, a more detailed form (S3) was created to comprehensively describe the most notable events of the catalogue with more sources and geographical information, while S4 records were created by simplifying the S3 versions.

From 1992 to 2002, the catalogue was consistently updated to include new information, so that the current catalogue contains events from the year 1030 to the year 2002. The vast majority of the events are relatively recent: just 798 records report an event that happened earlier than 1900, while the remaining 7586 records report events from 1900 onwards. This reflects the difficulty of retrieving reliable information about historical events. From 1950 onward, a clear increase in the number of records is visible, as a result of the fact that many newspapers introduced regional and local chronicles, which made information on less impactful events and on floods affecting smaller or more remote locations more readily available to the compilers.

Since 1999, the AVI catalogue, originally in physical format (paper documents), was digitized (Guzzetti and Tonelli 2004) and inserted in the SICI (*Sistema Informativo sulle Catastrofi Idrogeologiche*—Information System on Hydrological and Geomorphological Catastrophes), along with other catalogues, to form a single information system on hydrological and geomorphological historical events. In the SICI website (https://sici.irpi.cnr.it/, accessed on 26/05/2026), 8503 AVI records are available, each reporting a flooding event. These records constitute an invaluable resource since they are the only publicly available homogeneous source of information about floods in Italy.

Each AVI record contains information about one flooding occurrence. Some problems relative to the classification of events arise due to the way AVI records were compiled: a relative level of inhomogeneity can be seen in their compilation standards, especially when deciding whether multiple records should be created for the same event, or all the information should be inserted in a single record. To exemplify this complexity, we can take the infamous 1966 Firenze flood. This is the most notable instance of record separation in the database: 186 records in the catalogue report an event that happened on the 4th November 1966, 141 of which refer to an event that hit Tuscany. Most of these records contain information related to a single city. This compilation method is different from what was done for most other events, which would have been reported in fewer records. The spike in the events per year graph (Fig. 1, light blue bars) relative to 1966 can be explained by this unusual way of compiling records for this event.

The aforementioned example is not the only instance of event grouping or separation in the catalogue: there are several ways in which events happening during the same day in the same province may have been counted in more than one record. Sometimes the same event is submitted by different operating units, thus creating a doubling of the information. Other times, the same event is submitted by the same operating unit but by different researchers. Another possibility is that different operating units adopted different standards in the categorization of events into one or more records. All these inconsistencies are considered

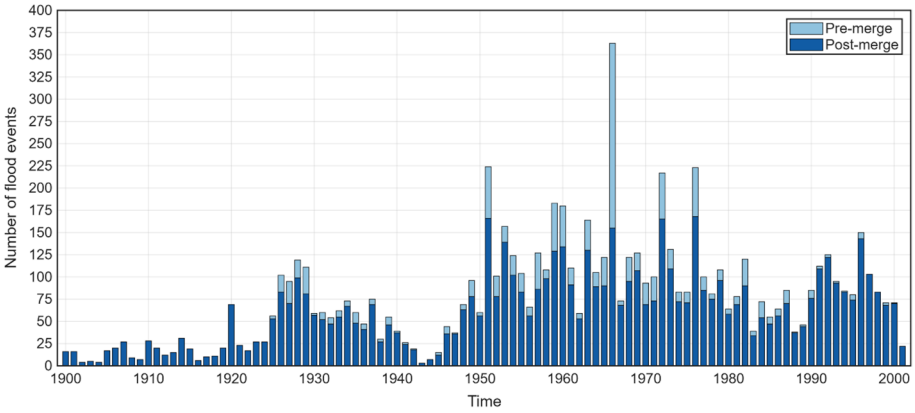


Fig. 1 Time evolution of the number of flood events in Italy (contained in the AVI catalogue). In light blue, the time evolution of the original dataset. In dark blue, the time evolution of the post-processed catalogue (information on the merging procedure is reported later, in Sect. 3)

when analysing the database. To avoid confusion, hereafter the term "record" is used to designate individual AVI entries, including multiple entries referring to the same occurrence, whereas "event" denotes the aggregated occurrences obtained through the unification of related records.

Each AVI record contains text-only information reported in a structured format to grant uniformity among all the events. In most records, some fields lack information: for events with a very localized impact or for older historical occurrences, the available sources often did not provide enough detail to populate all sections. More in detail, the fields of each record are:

- *Administrative information* the region, the province, the municipality and the ISTAT (*Istituto Nazionale di Statistica*) code (a unique identifier for each Italian municipality) affected by the event. No geographical coordinates are available.
- *General information about the event* this section includes the date of the flooding (typically in the day, month and year format). When precise information about the date is not available, the indication of the day or even the month may be absent (this is frequent in the case of older events). Other fields in this section are physiographic environment (i.e., information about the morphology of the territory, like plains, hills, mountains, etc.), duration of the events (usually reported with the number of days), affected area, affected basins, the watercourses involved in the event, and additional notes.
- *Cartographic information* this section includes the reference to the technical cartography of the area (*Tavoletta IGM*, in Italian). The triggering causes are also reported in this field.
- *Hydrological information* if available, hydrological data of the event are reported in this field.
- *Information about damage* this section presents a description of the damage of the event and is the richest in information since it contains several subsections focused on damage to agriculture, livestock assets, properties/assets, people, and land use, a monetary assessment, indirect effects, additional notes on damages and measures taken, and emer-

gency information. Within the sub-field related to damages to properties/assets, a distinction is made among different types of assets: these include public buildings (e.g., schools, hospitals, administrative offices), civil structures (e.g., residential areas, rural settlements), industrial facilities, public infrastructure (e.g., roads, railways, bridges), network infrastructures (e.g., water supply, sewage, power lines, telecommunications), and cultural or artistic heritage sites (e.g., churches, museums, historical buildings). For each asset type, the severity of the damage is specified and categorized into different levels, typically "mild," "serious," or "total." This classification allows for a detailed assessment of the impacts across various sectors, highlighting which categories of assets were most affected during each event. The most important component of the damage section is damage to people. If the events had any, this section reports the number of fatalities, missing persons, injured, temporarily and permanently displaced persons and indicates whether this number is confirmed or estimated.

- *Available documentation* this section refers to the general documentation of the event, including the references, the newspapers or other documents examined, the record type (S4, S3, S1, S0) and the operational unit responsible for the redaction of the record.

A more detailed description of the structure of the database is reported in Table 1.

3 Methods

3.1 Data preparation

The records available on the official AVI website are distributed across separate web pages, each identified by a unique record ID. This structure makes it difficult to perform direct queries or extract information about multiple events. To overcome this limitation, we downloaded the records with a Python script and built a local database. For this purpose, we employed the *requests* module, a widely used Python library that simplifies making HTTP requests to interact with web resources with minimal effort, handle responses, and manage tasks like authentication, headers, and cookies. We first extracted a list of years in which flood events were registered, obtained from the webpage summarizing the number of events per year (http://wwwdb.gndci.cnr.it/php2/avi/piene_riassunto_anno.php?lingua=it, accessed on 26/05/2026). Every page lists events of a specific year and follows the same standard URL structure, differing only by the year parameter (e.g., http://wwwdb.gndci.cnr.it/php2/avi/piene_anno.php?anno=2001&lingua=it, accessed on 26/05/2026). This allowed us to iterate through the list of years and systematically access all event pages. In each of these pages, the record IDs of events of that year can be found, so we were able to create a list of all the record IDs of events in the database. Subsequently, we developed a script that, given a record ID, uses the *requests* module to fetch the raw text content from the webpage associated with that event ID, cleans it up and extracts the data.

In this offline version of the database, we computed the affected people as the sum of number of missing persons, number of injured, number of temporarily displaced people, number of permanently displaced people reported in AVI. We then compiled a csv file containing all the events, with the information retrieved from each record and the information relative to the affected population. Each row represents a single event and includes all the

associated attributes. Of the 8,503 records that we retrieved from the website, we processed and stored in a local database 8,384 records, excluding those lacking information on the year of occurrence. The database includes 4,482 unique cities, with 21,506 city entries, as some cities appear multiple times due to recurring flood events.

A post-processing was carried out to standardize the dataset, addressing several inconsistencies identified in the original AVI records. As previously mentioned, during the compilation of the database, especially in the case of major flooding that impacted multiple cities or wide areas, the authors sometimes created multiple records for a single event. The fragmentation of this information can artificially inflate the number of floodings included, and our processing aimed to mitigate this issue by harmonizing and consolidating the information where appropriate. Thus, we compiled a second database, in which we merged the data referring to what we identified as the same event (see Fig. 1, dark blue bar, for the temporal evolution of the number of events). The criterion for merging two records was the following: if two records refer to events that happened in the same date and the provinces mentioned by the two records have some overlap (i.e., at least one province appeared in both records) we merged the records into a single event. We consider this approach reasonable, as it aligns with the compilation logic of the original database, where multiple municipalities within the same province were often reported as affected by the same flood. A second issue identified in the database concerned events reporting "zero" affected municipalities; these entries were manually corrected. A third error involved the number of fatalities: several events contained inconsistencies, and all entries reporting more than 10 fatalities were manually checked and corrected. As fourth editing step, we removed the 1985 Val di Stava disaster since it should be considered as an industrial disaster, as it was derived from the collapse of two mine-tailing dams. As a result of this merging process, the total number of entries decreased from 8,384 to 7,059. In this work, however, we opted to exclude events before the year 1900, since we considered the relative information less reliable due to a less systematic approach to the data collection; by disregarding the events recorded before 1900, the final count reaches 6292 events. A graphical representation of the entire procedure is depicted in Fig. 2.

In Fig. 3 we show, respectively, the spatial distribution of records (Fig. 3a) and events (Fig. 3b) in the database: after the merging procedure, some regions show a marked reduction in the number of occurrences compared to the original record count. This drop reflects

Table 1 Structure of each record of the AVI dataset

Field	Sub-field	Assets
Administrative information	Region Province Municipality ISTAT code	
General information about the event	Date Physiographic environment Duration of the event Area Notes	
Cartographic information	Tavoletta IGM Triggering causes	
Hydrological information		

Table 1 (continued)

Field	Sub-field	Assets
Damages	To agriculture	
	To livestock assets	
	To people	Number of fatalities
		Missing persons
		Injured persons
		Displaced persons
	To properties/assets	Civil Buildings – Scattered houses
		Civil Buildings – Urban areas
		Civil Buildings – Rural settlements
		Cultural Heritage Buildings – Historic and architectural assets
		Cultural Heritage Buildings – Churches
		Cultural Heritage Buildings – Museums
		Cultural Heritage Buildings – Works of art
		Industrial Buildings – General storage facilities
		Industrial Buildings – Livestock facilities
		Industrial Buildings – General manufacturing facilities
		Industrial Buildings – Extractive
		Industrial Buildings – Chemical
		Industrial Buildings – Electricity production
		Public Buildings – Other public services
		Public Buildings – Libraries
		Public Buildings – Hospitals
		Public Buildings – Schools
		Public Buildings – Public administration offices
		Network Infrastructure – Aqueduct
		Network Infrastructure – Power line
		Network Infrastructure – Sewer network
		Network Infrastructure – Gas pipeline
		Network Infrastructure
		– Telecommunications line
		Network Infrastructure – Oil pipeline
		Network Infrastructure – Penstock
		Transport Infrastructure – Highway
		Transport Infrastructure – Railway
		Transport Infrastructure – Municipal road
		Transport Infrastructure – Provincial road
		Transport Infrastructure – State road
		Transport Infrastructure – Cableway
		Public Interest Structures – Airports
		Public Interest Structures – Other
		Public Interest Structures – River regulation works
		Public Interest Structures – Bridges and viaducts
		Public Interest Structures – Dams
		Public Interest Structures – Port, River ports
		Public Interest Structures – Incinerators
		Public Interest Structures – Water treatment plants
		Public Interest Structures – Landfills
	To land use	
	Monetary assessment	
	Indirect effects	
	Notes on damages and measurements taken	
	Emergency information	
Available documentation	References	
	Newspapers	
	Record type	
	Responsible for the redaction of the record	

differences in how the AVI records were compiled in those areas. Once these records are consolidated and merged, the resulting number of unique events becomes more consistent with that of neighbouring regions.

3.2 Rationale for definition of severity classes

Having all the above categories of information in mind, we then made some assumptions to devise the creation of five classes to assess the severity of an event. We aimed to devise a scale with a quasi-logarithmic selection capacity, with each level containing approximately an order of magnitude more events than the previous, so that a clear distinction could be noted between events of different levels. This choice was motivated by the right-skewed frequency distribution of flood impacts in historical records: truly catastrophic events are exponentially rarer than moderate ones, and a logarithmic spacing of levels naturally mirrors this empirical regularity. In addition, the classification criteria combine quantitative variables (such as the number of casualties and affected people, which vary over several orders of magnitude across events) with qualitative indicators of damage extent, which by their nature introduce discrete jumps rather than continuous gradations. A quasi-logarithmic structure accommodates both types of information. With this in mind, we chose 5 classes in total because a higher number of classes seemed to us likely to produce more uncertainty in classification, with unclear boundaries between levels, as also pointed out by Blong (2003). A lower number of classes, on the other hand, may have caused events that are vastly different in gravity to be classified in the same level. Indeed, severity/damage scales usually have 5 to 13 levels (Blong 2003).

In defining the meaning of each level, we relied on well-established concepts in disaster studies and in similar impact-based classification frameworks proposed in the literature for floods, according to which the severity of an event is primarily determined by the number of people affected and by the extent of damage to infrastructure and the geographical area involved (Gad-el-Hak 2008; Caldera and Wirasinghe 2022). For example, Diakakis et al. (2020) developed a detailed flash flood impact severity scale that, in addition to human consequences, explicitly incorporates impacts on buildings and infrastructure. Similarly, the District Flood Severity Index (DFSI; Saharia et al. 2025), developed for India, integrates the number of flood events in a district, their duration, affected area, district population, and human consequences such as fatalities and injuries to characterize flood severity at the district scale. We thus acknowledge that elements like number of fatalities, affected areas and damage extent are widely recognized as the key factors that distinguish small-scale events from large-scale disasters and were therefore explicitly considered in the design of the FLOSEV scale. However, translating these concepts into operational criteria requires accounting for a fundamental asymmetry in the availability of impact data across the historical record. Quantitative information on human impacts (number of fatalities, injured, and displaced persons) is available for the vast majority of events, as it was systematically reported even in early newspaper articles and administrative records. By contrast, quantitative and detailed economic damage assessments are available only for recent events and only for the most severe ones, typically those for which detailed post-event studies were conducted. For the majority of historical events, the only information available on damage to assets and infrastructure is mostly qualitative. In the AVI database, for instance, damage is not reported with quantitative measures (e.g., number of destroyed buildings, kilometers of damaged

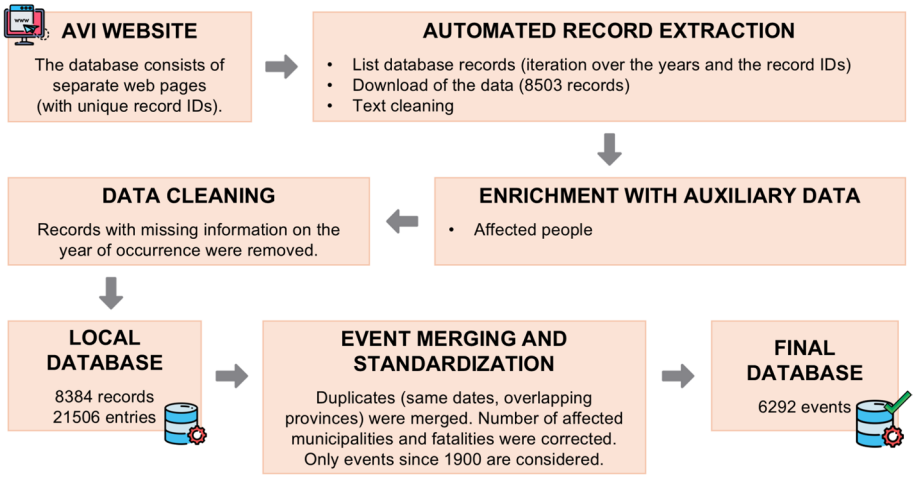


Fig. 2 Methodology used for data preparation

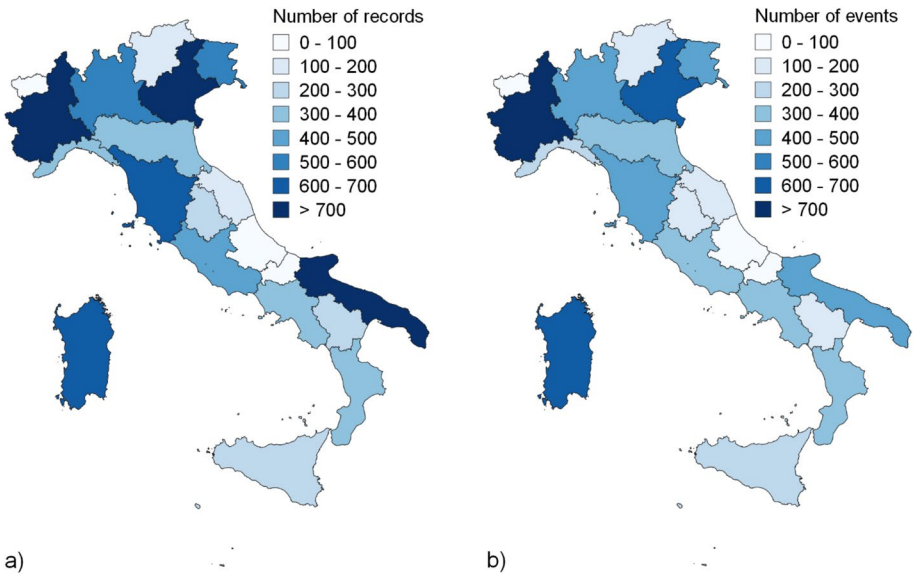


Fig. 3 Number of records (**a**) and events (**b**) in the AVI database for each Italian region

roads, etc.), but rather in an almost binary/qualitative form: for each asset class, the database simply records whether damage occurred and, if so, its qualitative severity (mild, serious or total). Rather than treating this asymmetry as a limitation, the scale is explicitly designed to accommodate it, combining quantitative human impact variables with qualitative damage indicators within a unified, rule-based classification scheme.

Accordingly, in the FLOSEV scale, each level is grounded exclusively in impact-based criteria without any direct reference to the statistical frequency of events. The expected

number of events per level, discussed in the following section, serves only as a posteriori validation that the scale produces a reasonable and interpretable distribution of events, and as an intuitive reference for the reader, but plays no role in the definition of the levels themselves. On the basis of this reasoning, we selected 5 levels as:

- *Level 5* - Floods as extraordinary catastrophes of enormous scale, which cause an immense amount of damage to people and assets. Given the exceptional nature of these events, the threshold values that characterized this level are selected considering not only the Italian data but taking also into account the major flood events that occurred after 1900 in Europe.
- *Level 4* - Floods as extremely severe events, that caused extensive destruction to assets. These are characterised by a very large number of fatalities and long-lasting impacts on the population and the environment. Level-4 floods are meant to be extraordinarily rare and destructive. They should have been observed just a few times in a century all over the country.
- *Level 3* - Floods as severe events of national resonance but which do not reach the scale of the higher levels. They are characterised by a medium to large number of fatalities and a high level of damage to assets. They should be observed some tens of times in a century.
- *Level 2* - Floods as events of medium severity that do not reach the level of damage of the other levels but are still notable. They are characterised by modest damage to people or by severe damage to assets. They should have been recorded few hundred times in a century.
- *Level 1* - The rest of the events, which do not present a high level of damage to people or assets and range from inconveniences to mild damage to assets or people.

To define the numerical thresholds separating each severity level, we adopted a percentile-rank based approach applied to the full dataset. For each event, we computed the percentile rank of 3 key variables: number of fatalities, number of affected municipalities, and number of total-damage indicators. Each variable was ranked independently on a scale from 0 to 1. A composite score was then obtained for each event as the median of the three percentile ranks, providing a balanced measure of overall severity that weights human impact, spatial extent, and asset destruction equally without privileging any single dimension. The thresholds defining the boundaries between levels were derived by identifying, for each level and for each variable separately, the minimum observed value at which the percentile rank reaches the cutoff corresponding to that level. The resulting classification assigns events to Levels 2, 3 and 4 if they meet at least two out of three criteria (fatalities $\geq$ X, number of affected municipalities $\geq$ Y, total-damage indicators $\geq$ Z). Level 5 is assigned to events with at least 150 fatalities, regardless of the other indicators, reflecting the exceptional and unambiguous nature of catastrophic loss of life at that scale. The 150 fatalities threshold has been defined with reference not only to the most impactful events recorded in Italy, but also to major European flood events (like the 2021 Ahr Valley flood in Germany and the 2024 Spanish flood "DANA"), so that any future event exceeding this threshold can be consistently classified as Level 5 without requiring retrospective comparison with the dataset. Events not meeting the criteria for Level 2 are assigned to Level 1.

Regarding the 2-out-of-3 classification rule, this choice reflects a deliberate trade-off between robustness and flexibility. A stricter AND rule (requiring all three thresholds to be exceeded simultaneously) would produce a more conservative classification but could systematically underclassify events for which one variable is poorly documented or genuinely atypical. A more permissive OR rule (requiring only one threshold to be exceeded) would risk over-classifying events that are extreme in one dimension but unremarkable in others; for instance, a flood affecting many small municipalities but causing no fatalities and minimal asset damage. The 2-out-of-3 rule occupies a middle ground: it requires convergent evidence from at least two independent dimensions of severity before assigning an event to a given level, reducing the influence of outliers in any single variable while remaining applicable even when one of the three variables is missing or uninformative.

4 Application and calibration of the FLOSEV scale

The methodology described in Sect. 3 was tested on the 6292 events that happened in the 103-year period 1900–2002. The original classification that resulted from the application of the methodology, with the corresponding threshold values and the number of events falling in each level is reported in Table 2. The distribution of events across the 5 levels follows the quasi-logarithmic pattern sought: Level 1 accounts for the vast majority of events (5548), while each successive level contains roughly an order of magnitude fewer events, down to 2 events at Level 5. This confirms that the percentile-rank calibration procedure successfully produced a scale where severity levels are well-separated and internally coherent.

To avoid thresholds that might appear overly precise, tailored to a specific dataset and therefore arbitrary we rounded all values to the nearest 5. This rounding introduces a small degree of approximation but makes the scale more interpretable. As a result, the final thresholds (reported in Table 3) are slightly more generic than those derived directly from the percentile-rank procedure, but remain fully consistent with the underlying distributional logic and do not alter the classification.

Overall, also the thresholds exhibit a broadly consistent quasi-logarithmic progression across levels, though with varying multipliers depending on the variable: fatalities increase by a factor of 5 between Level 2 and Level 3, by a factor of 7 between Level 3 and Level 4 and by a factor of 4 between Level 4 and Level 5; affected municipalities increase by a factor of 3 at both transitions; and total damage indicators increase by a factor of 5 between Level 2 and Level 3, and by a factor of 2 between Level 3 and Level 4, reflecting the fact that the number of total-damage indicators is inherently bounded by the number of asset categories recorded in the AVI framework, which limits the discriminatory range of this variable at higher severity levels. For Level 4, we additionally tested the effect of rounding the total-damage indicator threshold upward to 15 instead of 10. This adjustment caused the number of Level-4 events to drop from 9 to 7, as two events previously classified at Level 4 were reclassified to Level 3. We considered this outcome unsatisfactory, as it led to an underestimation of the severity of those events. We therefore retained the threshold of 10 total-damage indicators for Level 4.

As can be seen in Table 3, the distribution of events across the five levels follows a clear hierarchical pattern. The number of events in each class is roughly an order of magnitude greater than that of the subsequent, higher-severity level. This outcome aligns well

with our initial goal of constructing a logarithmic-type scale, where the frequency of events decreases systematically as severity increases. Such a relationship supports the internal consistency of the proposed classification and mirrors the typical behavior of natural hazard distributions, where extreme events are rarer but disproportionately more destructive. This framework enables a systematic evaluation of flood impacts and provides a consistent basis for comparison across time and space.

5 Discussion

By applying our final classification criteria to the database of flood events that occurred in Italy over the 103-year period from 1900 to 2002, we were able to re-examine individual events to evaluate if the classification actually reflects event severity. The two Level-5 events in our database are the 1963 Vajont flood (Genevois and Ghirotti 2005) and the 1954 Salerno flood (Esposito et al. 2004). This Vajont flood was caused by a landslide that led to a catastrophic wave, yet leaving untouched the Vajont Dam. The resulting flood destroyed the city of Longarone, just downstream, causing the death of 1917 people (the highest death toll for any event in the twentieth century in Italy). For this reason, we consider this event deserving of the highest level on the scale, even though the flood was primarily induced by a landslide that fell into the lake area. This event also reported 21 indicators of total damage (a number only surpassed by two other events in our dataset, which nonetheless fall below the Level-5 fatality threshold): this means that this flood caused total destruction to a great quantity of assets such as civil buildings, public and industrial buildings, monuments and infrastructure. Regarding the second one, the extraordinary rainfall recorded at Salerno in the 25/10/1954 (Esposito et al. 2004) produced a flood that hit the city and five other close cities along the coast. It is among the most destructive events in the entire database. It caused a total of 318 fatalities and it registered 25 damage indicators, of which 12 total, 12 severe and 1 mild.

In the Level-4 list of 13 events, one can recognize other widely known inundations, like the infamous one of Florence of the 4th November 1966, which, despite its fame, did not reach the huge level of mortality and destruction of the Level-5 events. It is worth mentioning that, with 13 events in a century, Level-4 inundations can be considered as almost the "flood of the decade" in Italy, i.e., those that are hard to forget. Using the AVI dataset, we

Table 2 Original thresholds attributed to each level of the FLOSEV scale

Level	Fatalities	Affected municipalities	Total damages	Selection rule	Number of events
1	–	–	–	All the events not included in Level 2	5548
2	1	5	1	At least 2 out of 3 between: fatalities ≥ 1, affected municipalities ≥ 5, total damages ≥ 1	660
3	4	16	6	At least 2 out of 3 between: fatalities ≥ 4, affected municipalities ≥ 16, total damages ≥ 6	73
4	33	46	14	At least 2 out of 3 between: fatalities ≥ 33, affected municipalities ≥ 46, total damages ≥ 14	9
5	150	–	–	Fatalities ≥ 150	2

Table 3 Final thresholds attributed to each level of the FLOSEV scale

Level	Fatalities	Affected municipalities	Total damages	Selection rule	Number of events
1	–	–	–	All the events not included in Level 2	5548
2	1	5	1	At least 2 out of 3 between: fatalities ≥ 1, affected municipalities ≥ 5, total damages ≥ 1	659
3	5	15	5	At least 2 out of 3 between: fatalities ≥ 5, affected municipalities ≥ 15, total damages ≥ 5	70
4	35	45	10	At least 2 out of 3 between: fatalities ≥ 35, affected municipalities ≥ 45, total damages ≥ 10	13
5	150	–	–	Fatalities ≥ 150	2

were also able to investigate the temporal evolution of flood severity in Italy throughout the twentieth century. As shown in Fig. 4, the average severity level remains relatively stable over time, showing no significant long-term variations.

A weak but positive relationship emerges between the number of events and their average severity (Fig. 5). This suggests that periods with more recorded floods tend to display slightly higher severity levels, although the association remains modest. However, this pattern may also be influenced by differences in data availability, reporting intensity, and spatial coverage rather than by a genuine physical link between flood frequency and severity. Regions with a more detailed historical record or a denser population may have more events documented, including minor floods, which inflate the number of records without necessarily affecting the average level. Conversely, in areas where only major floods were systematically reported, the number of events may be smaller but associated with higher mean severity.

Considering records after the year 1900, we can see that a higher number of records was collected in the northern area of Italy (Fig. 3). This could reflect the greater abundance of newspaper data available for the compilation of the records. However, there does not seem to be a correlation between the average level of the events per region (Fig. 6a) and the number of records collected for a region (Fig. 6b). This is not surprising since the majority of the records collected are low-level events.

To further test the effectiveness of the FLOSEV scale we applied the proposed classification to two recent flood events occurred in 2020 and in 2022: the idea is to assess whether the scale remains meaningful if used to classify events in the current, possibly richer, information context. We would also check how well the scale reflects present-day flood impacts, considering the significant progress made in both structural and non-structural flood risk management measures over recent decades.

In a first step, we analyze the flood that affected both northern and southern Piedmont, particularly the Cuneo valleys, on 2nd-3rd October 2020, caused by extratropical storm Alex (ARPA Piemonte 2020; Società Italiana di Geologia Ambientale 2021). This event represents a notable example of an extreme hydro-meteorological phenomenon in the Alpine context, and is also supported by detailed data, allowing an accurate classification according to the FLOSEV scale. The event was notable for its rapid development and the strong severity of the rainfall, which quickly exceeded the absorption and runoff capacity of the

involved catchments. The Limone Pancani (CN) meteorological station recorded 589 mm of total precipitation, of which 538 mm fell in just 12 h, a value more than 3.5 times greater than the highest rainfall ever recorded in that area (ARPA Piemonte 2025).

With over 350 affected municipalities (Regione Piemonte 2020), the event would rank as the second most widespread in the whole Italian database if included. The high number is partly due to the greater availability of data on affected assets and modern information tools compared to historical events. In total, 3 fatalities and approximately 770 temporarily displaced persons were recorded across an area of about 25,000 km^2 (Regione Piemonte 2020). This flood event is characterized by the exceptional damage produced, which exerted significant impacts on the territory and economic activities. The damages predominantly involved infrastructure and built assets, compromising the operational functionality of transport networks towards France, critical services, and industrial facilities. For a systematic damage reporting, the full AVI classification framework can be applied to categorize the various types of damage (Table 4).

Based on the proposed classification, the flood event of 2nd–3rd October, was assigned Level 2. Despite the destructive extent of asset damage and the exceptional rainfall recorded, the event does not fall into higher levels because the number of fatalities was limited and the total destruction did not affect civil buildings. This highlights the progress made over time in flood risk management and prevention, compared to historical events in the database that were assigned higher levels.

As a second case study, we selected the flood that struck the Marche Region between 15 and 16 September 2022, an event of exceptional intensity that mainly affected the provinces of Ancona and Pesaro-Urbino, impacting an area of approximately 5,000 km^2 (Morelli et al. 2023). The event was triggered by a stationary V-shaped self-regenerating thunderstorm, which produced exceptionally high rainfall concentrated over a very short period. At the Cantiano rain gauge station, a total of 419 mm of rainfall in about nine hours was recorded, with peak intensities of 90 mm/h, corresponding to more than 30% of the area's annual average precipitation (Confuorto et al. 2025). These values represent an extremely rare occurrence for the Italian territory and rank among the highest ever recorded in Italy for similar durations, highlighting the exceptional nature of the event (Tartaglione 2025). The social relevance of this flood, together with the availability of detailed pluviometric and hydro-

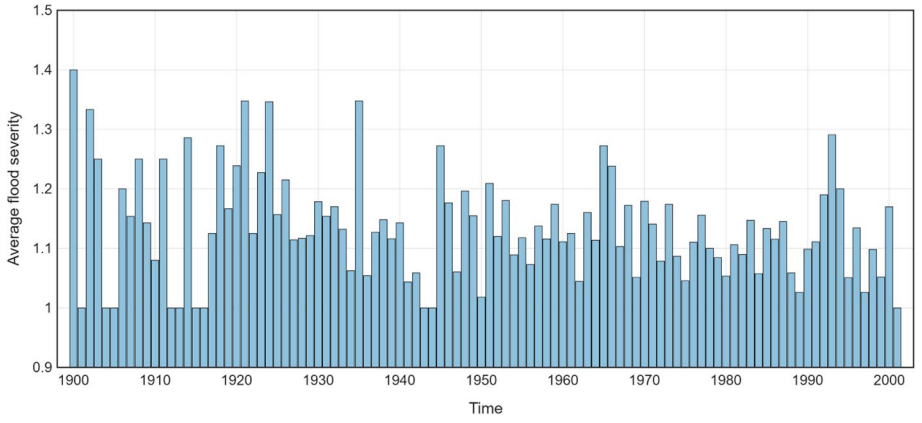


Fig. 4 Temporal evolution of the average severity level for each year

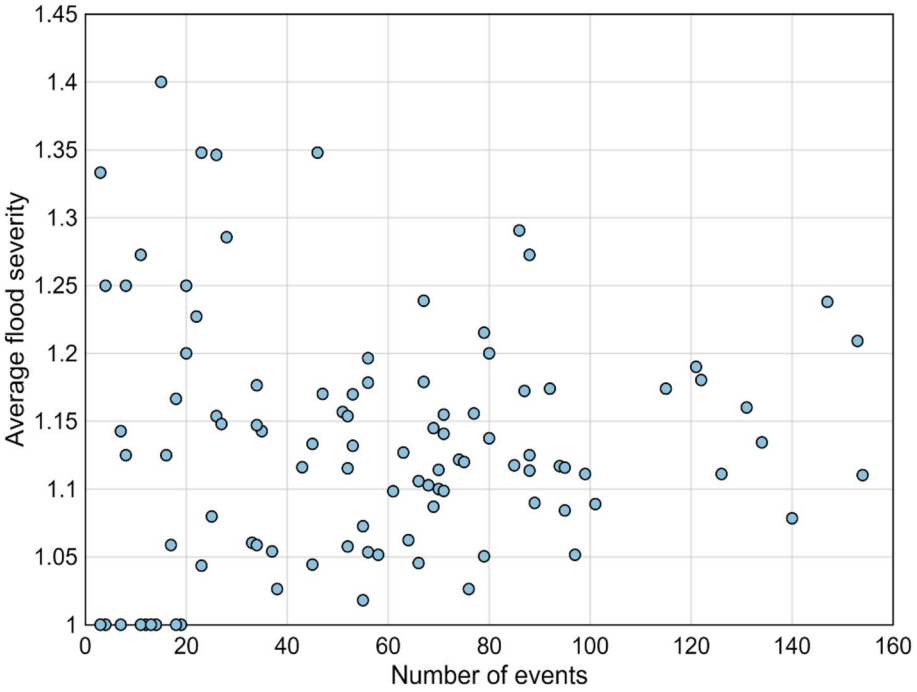


Fig. 5 Average level of event severity with respect to the number of events

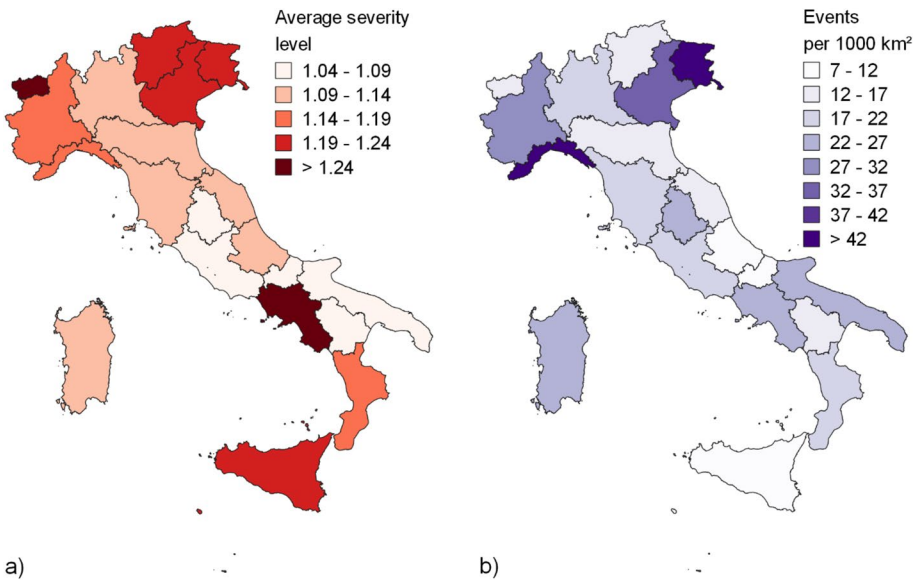


Fig. 6 Average severity level for each region (**a**). Number of events per region per 1000 km^2 of surface area (**b**)

logical data, allows for an accurate classification according to the FLOSEV scale while providing a significant example of an extreme hydro-meteorological event in the Apennine context.

The direct impacts on the population caused by the flood in the Marche region were very significant, resulting in 13 fatalities and over 1,200 displaced persons (Regione Marche—Direzione Protezione Civile e Sicurezza del Territorio—Centro Funzionale Regionale 2022). The event caused widespread and substantial damage to buildings and infrastructure (Rrokaj et al. 2025), classified in Table 5 according to the AVI categories, with impacts in tens of municipalities (La Notizia Giornale 2022; POLARIS 2022). The exceptional rainfall event also triggered 1,687 landslides in Marche and Umbria, with direct impacts on the urban fabric and transportation network, particularly in the Misa River basin (Santangelo et al. 2023).

Based on the adopted classification criteria, the Marche flood of September 2022 was assigned a Level 3 severity. The event caused the destruction of bridges, the complete submersion of the ground floors of numerous urban buildings, and damage to more than one hundred structures due to landslides, as well as a number of fatalities sufficient to fully meet the requirements for this severity level. In essence, as Level-3 type of event, this Marche flooding resulted in the list of the highest 100 floods of the last centuries.

In addition to applying our classification framework, we also tested alternative impact-based approaches for comparison purposes. In particular, we considered the Flash Flood Impact Severity Scale (FFISS) proposed by Diakakis et al. (2020). The FFISS distinguishes impacts across four categories: built environment, man-made mobile objects, natural environment, and population. Each category is organized into ten increasing severity levels, and the framework was designed not only for classification but also to support the spatial mapping of flash-flood impacts. Its application, however, relies on a highly detailed evidence base, including the typology of affected elements in the built environment, the severity and type of damage, the involvement of mobile objects (plus their typology and weight), the effects on the natural environment, and the consequences for the population (Table 6). Information of this kind is rarely available in a homogeneous and complete form, even for recent, well-documented events such as the 2020 Piedmont and 2022 Marche floods. For the Piedmont event, e.g., the flood caused severe disruption to roads, bridges, railways, service networks, and industrial facilities over a very large area. Yet, a full FFISS assessment would require more specific evidence on individual affected elements, including their location (information available), the type and severity of damage (info not available), possible loss of functionality (info not available), the involvement of vehicles and other mobile objects (info partially available in selected locations), and the effects on the natural environment (info partially available in selected locations). A similar limitation emerges for the Marche event.

By contrast, in both case studies, the available information is adequate for the application of our FLOSEV classification, which relies on the number of fatalities and the extent of material damage. Such information is often easier to retrieve from technical reports or newspaper archives, making this method particularly suitable for long-term historical analyses, as shown by the many projects focused on compiling flood damage data from newspaper-based sources (Zêzere et al. 2014; Gil-Guirado et al. 2019; Diakakis et al. 2020; Papagiannaki et al. 2013; Naulet et al. 2005; Paprotny et al. 2024).

Table 4 Damage classification of the Storm Alex case study

Field	Sub-field	Asset	Level
To agriculture	–	–	Mild
To livestock assets	–	–	Mild
To properties/assets	Civil Buildings	Scattered houses	Serious
		Urban areas	Serious
		Rural settlements	Serious
	Public Buildings	Other public services	Serious
	Network Infrastructure	Aqueduct	Serious
		Power line	Serious
		Sewer network	Mild
		Gas pipeline	Serious
		Telecom-munications line	Serious
	Transport Infrastructure	Railway	Serious
		Municipal road	Total
		Provincial road	Serious
		State road	Total
	Public Interest Structures	Other	Serious
		River regula-tion works	Total
		Bridges and viaducts	Serious
To land use	–	–	Mild

For these reasons, detailed classifications like the one suggested by Diakakis et al. (2020) are not directly applicable to datasets like AVI. More broadly, the two approaches serve different purposes. The framework by Diakakis et al. (2020) is primarily designed for the detailed assessment of recent events, where high-resolution post-event data are available. In contrast, the objective of our work is to develop a classification system that can be consistently applied to both historical and contemporary events, relying on a level of information that is commonly available across long-term and heterogeneous datasets. This difference in scope explains the distinct methodological choices adopted in the two approaches.

6 Conclusions

In this paper, we propose a systematic impact-based classification of the severity of flood events, tested on flood events that affected Italy throughout the twentieth century using the AVI database. Unlike traditional hydrological metrics, the proposed flood severity (FLOSEV) scale does not rely on discharge, water level, or rainfall data. Instead, it uses information on the impact of an event (fatalities, number of affected municipalities, and the extent of damage to assets) as the main factors driving the attribution of a severity level. The

Table 5 Damage classification of the Marche region flood

Field	Sub-field	Asset	Level
To agriculture	–	–	Serious
To livestock assets	–	–	Mild
To properties/assets	Civil Buildings	Scattered houses	Serious
		Urban areas	Total
		Rural settlements	Serious
	Cultural Heritage Buildings	Churches	Mild
	Industrial Buildings	General storage facilities	Serious
		Livestock facilities	Serious
		General manufacturing facilities	Serious
	Public Buildings	Other public services	Serious
		Public administration offices	Serious
	Network Infrastructure	Aqueduct	Serious
		Power line	Serious
		Sewer network	Total
	Transport Infrastructure	Highway	Serious
		Railway	Serious
		Municipal road	Total
		Provincial road	Total
		State road	Serious
	Public Interest Structures	River regulation works	Total
		Bridges and viaducts	Total
To land use	–	–	Mild

resulting scale distinguishes 5 levels of severity, producing a ranking of events with a quasi-logarithmic distribution, where each level contains approximately an order of magnitude more events than the one above. The classification relies on a percentile-rank calibration and a 2-out-of-3 threshold rule, which requires convergent evidence from at least two independent dimensions of severity before assigning an event to a given level. Because it does not depend on instrumental hydrological data, which are not systematically available until relatively recent years, the scale is suitable for classifying events across different historical periods. Therefore, as long as information such as the number of fatalities and the extent of property damage is available, events can be assessed and compared.

The approach is not without limitations. Collecting and verifying impact information can be complex and time-consuming, and it may introduce biases toward larger cities or events of greater social and economic resonance. Historical newspapers, for instance, tended to

Table 6 Data requirements for the application of FFISS and FLOSEV

Impact category	FFISS	FLOSEV
Built environment	Type of affected buildings (residential, industrial, public, …) or infrastructure Flood depth (basement, ground floor, upper floors) Affected part of the asset Nature of damage (structural vs non-structural) Degree of structural damage (inundation, cracks, subsidence, deformation, partial collapse, total collapse) Involvement of strategic infrastructure (roads, bridges, railways, levees, utilities) and impact on them	Aggregated information on damage to buildings and infrastructure Overall assessment of damage severity
Man-made mobile objects	Type/weight of mobile object get transported Number and spatial location of affected objects Distinction between isolated and widespread involvement Human impacts associated with mobile objects	Not required for classification
Natural environment	Damage to vegetation or forests (type of vegetation uprooted or destroyed) Presence of inundated agricultural plots Type of impact on animals Soil erosion and sediment deposition Presence of trash spread in the watercourses Landslides triggered Presence and entity of contamination events	Not required for classification
Human population	Number of trapped, injured, and deceased people Spatial location of cases Specific circumstances of involvement Changes in everyday activities Evacuation cases Entity of the injuries Distinction between isolated cases and grouped cases type. For grouped cases, number of fatalities occurred locally, in the same building, campsite, vehicle, etc.…	Number of fatalities and affected people in aggregated form

focus their coverage on areas closer to their headquarters or on events of greater social or economic relevance, potentially leading to an underrepresentation of concentrated or more remote flood events.

Although calibrated on Italian data, the FLOSEV scale is designed as a general framework rather than an Italy-specific classification, and can in principle be transferred to other geographical contexts. Potential datasets for such assessment would include those provided by insurance companies, such as NatCatSERVICE (owned by Munich Re), as well as global catastrophe databases like EM-DAT. In this possible follow-up, the most significant issue concerns differences in data structure and field definitions between the AVI database and other datasets. While correspondences can likely be established for many variables, some fields may not have direct equivalents and would require specific adaptation to make the scale operational in a new context. The most problematic field is expected to be the one related to damage indicators, which reflects a specific way of recording and interpreting damage in the AVI dataset. This variable could potentially be replaced by a more general indicator of damage severity, although the appropriate formulation and implementation of such a substitution remains to be defined. Despite the problems associated with the above-mentioned biases, this approach allows a viable assessment of basic information about fatalities and damages and can be readily applied to whatever context other than Italy, as

it is based only on the degree of destruction and the extent of the affected area, in terms of involved municipalities.

In summary, the flood severity scale provides a simple and flexible framework, based on an extensive database, for comparing flood impacts across space and time. Its quite viable computation allows it to be applied even when data quality varies, while its reliance on observable impacts makes it suitable for long-term analyses. Although further work is needed to explore its applicability in different geographical and data contexts, the scale provides a practical starting point for assessing flood event severity even when impacts are available in a descriptive way.

Author contributions All authors contributed to the study conception and design. Data collection and analysis were performed by Alessandro Giacalone and Anna D'Andrilli. Visualizations were created by Paola Mazzoglio and Anna D'Andrilli. The first draft of the manuscript was written by Paola Mazzoglio and Alessandro Giacalone and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Funding Open access funding provided by Politecnico di Torino within the CRUI-CARE Agreement. No funding was received for conducting this study.

Declarations

Conflict of interest The authors have no relevant financial or non-financial interests to disclose.

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

References

Addor N, Hong XD, Alvarez-Garreton C, Coxon G, Fowler K, Mendoza PA (2019) Large-sample hydrology: recent progress, guidelines for new datasets and grand challenges. Hydrol Sci J 65(5):712–725. https://doi.org/10.1080/02626667.2019.1683182

Archer D, Fowler H (2021) A historical flash flood chronology for Britain. J Flood Risk Manag 14(3):e12721. https://doi.org/10.1111/jfr3.12721

ARPA Piemonte (2020) Rapporto evento 02 ottobre 2020. ARPA Piemonte. https://www.arpa.piemonte.it/sites/default/files/media/2024-07/rapporto_evento_02_ottobre_2020_.pdf.

ARPA Piemonte (2025) Banca dati storica. ARPA Piemonte. https://www.arpa.piemonte.it/rischi_naturali/snippets_arpa_graphs/precipitazioni_intense/?statid=PIE-004110-901-2006-10-05.

Barredo JI (2007) Major flood disasters in Europe: 1950–2005. Nat Hazards 42:125–148. https://doi.org/10.1007/s11069-006-9065-2

Bell HM, Tobin GA (2007) Efficient and effective? The 100-year flood in the communication and perception of flood risk. Environ Hazard 7(4):302–311. https://doi.org/10.1016/j.envhaz.2007.08.004

Bencivenga M, Di Loreto E, Liperi L (1999) Piene storiche del Tevere a Roma. L'acqua 3:17–24

Bersani P, Bencivenga M (2001) Le piene del Tevere a Roma dal V secolo a.C. all'anno 2000. Tevere 22:4–9

Bertola M, Viglione A, Lun D, Hall J, Blöschl G (2020) Flood trends in Europe: are changes in small and big floods different? Hydrol Earth Syst Sci 24:1805–1822. https://doi.org/10.5194/hess-24-1805-2020

Bertola M, Viglione A, Vorogushyn S, Lun D, Merz B, Blöschl G (2021) Do small and large floods have the same drivers of change? A regional attribution analysis in Europe. Hydrol Earth Syst Sci 25:1347–1364. https://doi.org/10.5194/hess-25-1347-2021

Berz G (2000) Flood disasters: lessons from the past–worries for the future. Proc Inst Civ Eng Water Marit Eng 142:3–8. https://doi.org/10.1680/wame.2000.142.1.3

Blong R (2003) A review of damage intensity scales. Nat Hazards 29:57–76. https://doi.org/10.1023/A:1022960414329

Bonini FM (1666) Del Tevere incatenato overo l'arte di frenar le correnti. Francesco Moneta, Roma

Brázdil R, Kundzewicz ZW, Benito G (2006) Historical hydrology for studying flood risk in Europe. Hydrol Sci J 51(5):739–764. https://doi.org/10.1623/hysj.51.5.739

Caldera HJ, Wirasinghe SC (2022) A universal severity classification for natural disasters. Nat Hazards 111:1533–1573. https://doi.org/10.1007/s11069-021-05106-9

Claps P, Ganora D, Mazzoglio P (2022) Rainfall regionalization techniques. In: Morbidelli R (ed) Rainfall. Elsevier, Amsterdam, pp 327–350. https://doi.org/10.1016/B978-0-12-822544-8.00013-5

Claps P, Evangelista G, Ganora D, Mazzoglio P, Monforte I (2024) FOCA: a new quality-controlled database of floods and catchment descriptors in Italy. Earth Syst Sci Data 16:1503–1522. https://doi.org/10.5194/essd-16-1503-2024

Confuorto P, Franceschini R, Scarpitta L, Casagli N, Morelli S, Raspini F, Tofani V, Moretti S (2025) Event-based landslide inventory through very high-resolution optical images and field surveys. Geoenviron Disasters 12:23. https://doi.org/10.1186/s40677-025-00328-6

Cunado J, Ferreira S (2014) The macroeconomic impacts of natural disasters: the case of floods. Land Econ 90(1):149–168. https://doi.org/10.3368/le.90.1.149

de Brito MM, Sodoge J, Kreibich H, Kuhlicke C (2025) Comprehensive assessment of flood socioeconomic impacts through text-mining. Water Resour Res 61:e2024WR037813. https://doi.org/10.1029/2024WR037813

Delforge D, Wathelet V, Below R, Sofia CL, Tonnelier M, van Loenhout JA, Speybroeck N (2025) EM-DAT: the emergency events database. Int J Disaster Risk Reduct 124:105509. https://doi.org/10.1016/j.ijdrr.2025.105509

Diakakis M, Deligiannakis G, Antoniadis Z, Melaki M, Katsetsiadou NK, Andreadakis E, Spyrou NI, Gogou M (2020) Proposal of a flash flood impact severity scale for the classification and mapping of flash flood impacts. J Hydrol 590:125452. https://doi.org/10.1016/j.jhydrol.2020.125452

Douben K-J (2006) Characteristics of river floods and flooding: a global overview, 1985–2003. Irrig Drain 55:S9–S21. https://doi.org/10.1002/ird.239

Esposito E, Porfido S, Violante C (2004) Il nubifragio dell'ottobre 1954 a Vietri sul Mare-Costa di Amalfi Salerno. CNR GNDCI, Report 2870:1–381

Fernández B, Salas JD (1999a) Return period and risk of hydrologic events. I: mathematical formulation. J Hydrol Eng 4(4):297–307

Fernández B, Salas JD (1999b) Return period and risk of hydrologic events. II: Applications. J Hydrol Eng 4(4):308–316

Fujita T (1971) Proposed characterization of tornadoes and hurricanes by area and intensity. SMRP Research Paper no. 91. Department of Geophysical Sciences, University of Chicago, Chicago.

Gad-el-Hak M (2008) Introduction. In: Gad-el-Hak M (ed) Large-scale disasters. Cambridge University Press, New York, pp 1–4. https://doi.org/10.1017/CBO9780511535963.003

Gaume E, Bain V, Bernardara P, Newinger O, Barbuc M, Bateman A et al (2009) A compilation of data on European flash floods. J Hydrol 367(1–2):70–78. https://doi.org/10.1016/j.jhydrol.2008.12.028

Genevois R, Ghirotti M (2005) The 1963 Vaiont landslide. G Geol Appl 1:41–52

Gil-Guirado S, Pérez-Morales A, Lopez-Martinez F (2019) SMC-Flood database: a high-resolution press database on flood cases for the Spanish Mediterranean coast (1960–2015). Nat Hazards Earth Syst Sci 19(9):1955–1971. https://doi.org/10.5194/nhess-19-1955-2019

Govi M, Turitto O (1994) Ricerche bibliografiche per un catalogo sulle inondazioni, piene torrentizie e frane in Valtellina e Valchiavenna. Associazione Mineraria Subalpina, Quaderni di Studi e Documentazione, Quaderno 16.

Grazzini F, Fragkoulidis G, Pavan V, Antolini G (2020) The 1994 Piedmont flood: an archetype of extreme precipitation events in Northern Italy. Bull Atmos Sci Technol 1:283–295. https://doi.org/10.1007/s42865-020-00018-1

Grünthal G (1998) European macroseismic scale 1998 (EMS-98).

Guzzetti F, Tonelli G (2004) Information system on hydrological and geomorphological catastrophes in Italy (SICI): a tool for managing landslide and flood hazards. Nat Hazards Earth Syst Sci 4:213–232. https://doi.org/10.5194/nhess-4-213-2004

Guzzetti F, Cardinali M, Reichenbach P (1994) The AVI project: a bibliographical and archive inventory of landslides and floods in Italy. Environ Manage 18:623–633. https://doi.org/10.1007/BF02400865

Guzzetti F, Stark CP, Salvati P (2005) Evaluation of flood and landslide risk to the population of Italy. Environ Manag 36:15–36. https://doi.org/10.1007/S00267-003-0257-1

Koutsoyiannis D (2007) A critical review of probability of extreme rainfall: principles and models. In: Ashley R, Garvin S, Pasche E, Vassilopoulos A, Zevenbergen C (eds), Advances in urban flood management. Taylor & Francis, London, chap 7. https://doi.org/10.1201/9780203945988.

La Notizia Giornale (2022). Alluvione Marche: la situazione nei comuni colpiti. https://www.lanotiziagiornale.it/alluvione-marche-la-situazione-nei-comuni-colpiti/.

Lai K, Porter JR, Amodeo M, Miller D, Marston M, Armal S (2022) A natural language processing approach to understanding context in the extraction and geocoding of historical floods, storms, and adaptation measures. Inf Process Manag 59(1):102735. https://doi.org/10.1016/j.ipm.2021.102735

Lavers DA, Villarini G, Cloke HL, Simmons A, Roberts N, Lombardi A, Burgess SN, Pappenberger F (2025) How bad is the rain? Applying the extreme rain multiplier globally and for climate monitoring activities. Meteorol Appl 32(2):e70031. https://doi.org/10.1002/met.70031

Li N, Thiery W, Zahra S, de Brito MM, Worou K, Kurfalı M, Lampe S, Muñoz P, Flynn C, Trigoso C, Nivre J, Zscheischler J, Messori G (2026) Wikimpacts 1.0: a new global climate impact database based on automated information extraction from Wikipedia. Nat Hazards Earth Syst Sci 26:2609–2636. https://doi.org/10.5194/nhess-26-2609-2026

Masoero A, Claps P, Asselman NEM, Mosselman E, Di Baldassarre G (2013) Reconstruction and analysis of the Po River inundation of 1951. Hydrol Process 27:1341–1348. https://doi.org/10.1002/hyp.9558

Mazzoglio P, Macchia S, Gallo E, Winter J, Claps P (2021) Disaster tales as communication tool for increasing risk resilience. Int J Disaster Risk Sci 12:341–354. https://doi.org/10.1007/s13753-021-00341-x

Mazzoglio P, Bertola M, Listo T, Princivalle L, Lombardo L, Viglione A, Claps P, Mattozzi A (2025) Who is saving our streamflow data? Exploring volunteer profiles and their engagement in the SIREN data rescue project. PLoS ONE 20(10):e0333091. https://doi.org/10.1371/journal.pone.0333091

Mazzoglio P (2022) Geographically-based approaches to the statistical analysis of rainfall extremes. PhD thesis, Politecnico di Torino, Italy.

Mazzoglio P, Laio F, Sandu C, Boccardo P (2019) Assessment of an Extreme Rainfall Detection System for Flood Prediction over Queensland (Australia). Proceedings 18:1. https://doi.org/10.3390/ECRS-3-06187.

Merz B, Blöschl G, Vorogushyn S et al (2021) Causes, impacts and patterns of disastrous river floods. Nat Rev Earth Environ 2:592–609. https://doi.org/10.1038/s43017-021-00195-3

Monticelli P (1998) Ricostruzione storica degli eventi alluvionali nelle Valli di Lanzo tra il 1400 ed il 1900. Regione Piemonte, Direzione Servizi Tecnici di Prevenzione, Quaderno 12.

Morelli S, Bonì R, Guidi E, De Donatis M, Pappafico G, Francioni M (2023) L'alluvione delle Marche del 15 settembre 2022: cause e conseguenze. La Dinamica fluviale. La conoscenza del fiume per la pianificazione e la salvaguardia del territorio, 24:136–147. https://hdl.handle.net/11576/2723812.

Naulet R, Lang M, Ouarda TB, Coeur D, Bobee B, Recking A, Moussay D (2005) Flood frequency analysis on the Ardeche river using French documentary sources from the last two centuries. J Hydrol 313:58–78. https://doi.org/10.1016/j.jhydrol.2005.02.011

Papagiannaki K, Lagouvardos K, Kotroni V (2013) A database of high-impact weather events in Greece: a descriptive impact analysis for the period 2001–2011. Nat Hazards Earth Syst Sci 13(3):727–736. https://doi.org/10.5194/nhess-13-727-2013

Papagiannaki K, Petrucci O, Diakakis M et al (2022) Developing a large-scale dataset of flood fatalities for territories in the Euro-Mediterranean region. FFEM-DB Sci Data 9:166. https://doi.org/10.1038/s41597-022-01273-x

Paprotny D, Terefenko P, Śledziowski J (2024) HANZE v2.1: an improved database of flood impacts in Europe from 1870 to 2020. Earth Syst Sci Data 16:5145–5170. https://doi.org/10.5194/essd-16-5145-2024

Payrastre O, Gaume E, Andrieu H (2011) Usefulness of historical information for flood frequency analyses: developments based on a case study. Water Resour Res. https://doi.org/10.1029/2010WR009812

Regione Piemonte (2020) Evento alluvionale del 2–3 ottobre 2020: Relazione a supporto della richiesta di dichiarazione dello stato di emergenza. https://www.regione.piemonte.it/web/media/27777/download.

POLARIS (2022). Alluvione nelle Marche. https://polaris.irpi.cnr.it/event/alluvione-nelle-marche/

Regione Marche - Direzione Protezione Civile e Sicurezza del Territorio - Centro Funzionale Regionale (2022) RAPPORTO DI EVENTO preliminare: Maltempo 15, 16 e 17 settembre 2022 https://www.regione.marche.it/portals/0/Protezione_Civile/Manuali%20e%20Studi/Rapporto_Evento_preliminare_20220915.pdf

Rrokaj S, Arrighi C, Ballocci M, Bertoli G, da Porto F, De Lucia C, Di Bacco M, Di Fluri P, Domeneghetti A, Donà M, Gallazzi A, Gennaro A, Hammouti M, Lelli G, Mozzon S, Petruccelli N, Saler E, Scorzini AR, Sterlacchini S, Treglia G, Voltolina D, Zazzeri M, Molinari D (2025) Survey data of damaged residential buildings and business premises from the 2022 record-breaking flood in the Marche region, Italy. Earth Syst Sci Data 17:6487–6496. https://doi.org/10.5194/essd-17-6487-2025

Saharia M, Jain SK, Prakash V, Malik H, Sreejith OP, Joshi D (2025) A district-level flood severity index for flood management in India. Nat Hazards 121:17815–17831. https://doi.org/10.1007/s11069-025-07493-9

Santangelo M, Althuwaynee O, Alvioli M, Ardizzone F, Bianchi C, Bornaetxea T, Brunetti MT, Bucci F, Cardinali M, Donnini M, Esposito G, Gariano SL, Grita S, Marchesini I, Melillo M, Peruccacci S, Salvati P, Yazdani M, Fiorucci F (2023) Inventory of landslides triggered by an extreme rainfall event in Marche-Umbria, Italy, on 15 September 2022. Sci Data 10:427. https://doi.org/10.1038/s41597-023-02336-3

Schroeder AJ, Gourley JJ, Hardy J, Henderson J, Parhi P, Rahmani V, Reed KA, Schumacher RS, Smith BK, Taraldsen MJ (2016) The development of a flash flood severity index. J Hydrol 541:523–532. https://doi.org/10.1016/j.jhydrol.2016.04.005

Società Italiana di Geologia Ambientale (2021) L'evento alluvionale del 2–3 ottobre 2020 in Piemonte. SIGEA. https://www.sigeaweb.it/documenti/gda-supplemento-4-2021.pdf.

Sodoge J, Kuhlicke C, de Brito MM (2023) Automatized spatio-temporal detection of drought impacts from newspaper articles using natural language processing and machine learning. Weather Clim Extrem 41:100574. https://doi.org/10.1016/j.wace.2023.100574

Sodoge J, Kuhlicke C, Mahecha MD, de Brito MM (2024) Text mining uncovers the unique dynamics of socio-economic impacts of the 2018–2022 multi-year drought in Germany. Nat Hazards Earth Syst Sci 24:1757–1777. https://doi.org/10.5194/nhess-24-1757-2024

Tartaglione N (2025) Unprecedented flooding in the Marche Region (Italy): analyzing the 15 September 2022 event and its unique meteorological conditions. Meteorology (Basel) 4(1):3. https://doi.org/10.3390/meteorology4010003

Tertulliani A (2019) Storia delle scale macrosismiche. Quaderni di Geofisica 150:1–60. https://doi.org/10.13127/qdg/150

Trigila A, Iadanza C, Lastoria B, Bussettini M, Barbano A (2021) Dissesto idrogeologico in Italia: pericolosità e indicatori di rischio - Edizione 2021. ISPRA, Rapporti 356/2021.

UNDRR and CRED (2020) The human cost of disasters: an overview of the last 20 years (2000–2019). UNDRR, Geneva. https://www.undrr.org/media/48008/download?startDownload=20251030.

Vennari C, Parise M, Santangelo N, Santo A (2016) A database on flash flood events in Campania, southern Italy, with an evaluation of their spatial and temporal distribution. Nat Hazards Earth Syst Sci 6(12):2485–2500. https://doi.org/10.5194/nhess-16-2485-2016

Volpi E (2019) On return period and probability of failure in hydrology. Wires Water 6:e1340. https://doi.org/10.1002/wat2.1340

Webb JCD, Elsom DM, Meaden GT (1986) The TORRO hailstorm intensity scale. J Meteorol 11(114):337–339

Zêzere JL, Pereira S, Tavares AO, Bateira C, Trigo RM, Quaresma I, Santos PP, Santos M, Verde J (2014) Disaster: a GIS database on hydro-geomorphologic disasters in Portugal. Nat Hazards 72(2):503–532. https://doi.org/10.1007/s11069-013-1018-y

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