Journal of Environmental Management 322 (2022) 116051

Contents lists available at ScienceDirect

Journal of Environmental Management

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

Research article

AReGeoDatHa: Apulian Regional GeoDatabase for geo-hydrological Hazards

Carmela Vennari[a,*], Paola Salvati[b], Cinzia Bianchi[b], Domenico Casarano[c], Mario Parise[d], Alessia Basso[c], Ivan Marchesini[b]

[a] Consiglio Nazionale delle Ricerche (CNR), Istituto di Ricerca per la Protezione Idrogeologica, Rende, CS, Italy
[b] Consiglio Nazionale delle Ricerche (CNR), Istituto di Ricerca per la Protezione Idrogeologica, Perugia, Italy
[c] Consiglio Nazionale delle Ricerche (CNR), Istituto di Ricerca per la Protezione Idrogeologica, Bari, Italy
[d] Earth and Environmental Sciences Department, University Aldo Moro, Bari, Italy

ARTICLE INFO

Keywords:
Geodatabase
Geo-hydrological hazard
Geo-hydrological damage
Source of information quality index

ABSTRACT

In many territories, geo-hydrological disasters represent a problem of considerable importance, given that they can cause substantial property damage, above all, life losses. In Italy, the geo-hydrological risk is widespread depending on the geomorphological settings, climate conditions and anthropic pressure. Accurate knowledge regarding the spatial and temporal extensions of past events, as well as the associated damages, allows the elaboration of hazard and risk maps. The resulting data are strongly dependent on the accuracy and efficiency of the sources used for data collection. In the framework of a long-term regional project, we compiled a geospatial database of geo-hydrological hazards and damage that occurred in the Apulian region of southern Italy over the last few decades. We used a validated database structure and improved it for practical and scientific purposes, thereby making it compliant with the needs of the stakeholders (i.e., regional civil protection agency) as it considers the regional geo-hydrological peculiarity and, at the same time, it follows the FLOOD and INSPIRE European directives.

Based on the research and analysis of different information sources, we propose a method for quantitatively evaluating the availability, efficiency, and accuracy of each utilized source (i.e., source of information total quality index—SIQI) in an objective and reproducible manner. The SIQI index of the sources used is analyzed and discussed.

We discuss the collected data on geo-hydrological processes and/or damage between 2008 and 2019, analyze their temporal and geographical distribution, and make comparisons with other national catalogues and daily geo-hydrological warning alerts.

1. Introduction

Understanding where geo-hydrological risk and related impacts will be greatest is largely based on rigorous analysis of detailed data on past events (Glade et al., 2001; Salvati et al., 2009; Taylor et al., 2015; Napolitano et al., 2018; Avand et al., 2021).

The knowledge of the spatial and temporal distribution of geo-hydrological phenomena, the quantification of the associated damages and the characterization of their triggers can facilitate (i) the assessment of the impacts of possible future weather/climate or seismic triggers, (ii) the identification of areas exposed to relevant impacts caused by natural phenomena. Understanding past events also provides ideas for activating specific damage mitigation actions in relation to possible future scenarios. It also favors the development, use and validation of early warning systems. For establishing a magnitude–frequency relationship regarding hazardous events, it is generally necessary to collect historical data and perform statistical analysis (van Westen et al., 2008; van Westen and Greiving, 2017). Available landslide/flood inventories can be used for analyzing possible changes in the frequency and intensity of events and, consequently, possible variations in their triggering factors (Mertz et al., 2014; Stoffel et al., 2014; Gariano et al., 2015, 2021), even in light of the impending climate changes. In this regard, collecting, organizing, and managing data about geo-hydrological phenomena provides indispensable knowledge

---

* Corresponding author.
E-mail address: carmela.vennari@irpi.cnr.it (C. Vennari).

https://doi.org/10.1016/j.jenvman.2022.116051
Received 31 January 2022; Received in revised form 14 July 2022; Accepted 18 August 2022

regarding vulnerability, hazard, and risk assessments (Guzzetti et al., 2005, 2012; van Westen et al., 2006; Van Den Eeckhaut and Hervás, 2012; Blahut et al., 2012; Mirus et al., 2020; Rohan et al., 2021).

Although it is widely recognized that collecting information on past and recent geo-hydrological disasters is an exhausting and time-intensive activity (Blahut et al., 2012), it is still the starting point for building a reliable database where data on processes, damage, and costs can be used for producing rigorous statistical analyses for successive hazard and risk evaluation. As it is quite rare to obtain all the data useful to reconstruct the complete history of a harmful event, it is crucial to investigate the following questions: (i) What are the most effective resources and storage systems for optimizing the activity, and (ii) what is the quality of the resulting databases? In this context, the need for "systematically evaluating, recording, sharing, and publicly accounting for disaster losses" was acknowledged as a major priority of the United Nations Hyogo framework for action on disaster risk reduction (UNISDR, 2015), while requesting the development and improvement of relevant databases. In recent years, many agencies, organizations, and research institutes have started working on disaster data collection and management, following the EU directive on open data and the reuse of public sector information (i.e., directive (EU) 2019/1024), which encourages member states to make available for reuse as much information as possible. Consequently, the data collection activity is highly interesting to the scientific community, but it is predominantly applied in public organizations regarding civil protection issues and amplified by the expected increases in frequency and intensity of the triggering factors. The intensification of hydro-meteorological hazards related to climate change (IPCC, 2014), combined with the increase of the exposed elements, is expected to rapidly increase the impact of disasters (van Westen et al., 2011; Tanoue et al., 2016; Ellena et al., 2020), as well as the risks of economic activities in terms of damages.

In this paper, we describe the regional database of geo-hydrological events as a relevant result of a project involving the CNR IRPI and the Apulia regional civil protection agency. Research, selection, and storage activities were devoted to collecting information on major events, low-intensity phenomena, damages, and related costs spent by the public and private sectors. Contextually, we present novel criteria designed for ranking the availability, efficiency, and accuracy of each source of information used for compiling the database. We believe that the results of the work, the proposed criterion, and the procedure described can be used for optimizing future data collection campaigns in Italy and elsewhere. To the best of our knowledge, and because of the nature and quality of the information collected, the database described in this paper represents an ideal extension at a regional scale of the large national AVI project, concluded in 2001 (Guzzetti et al., 1994; Guzzetti and Tonelli, 2003) for both landslide and flood hazards. Results allowed us to compare the temporal and spatial distributions of phenomena that occurred in the Apulia region between 2008 and 2019 with those registered in the AVI database between 1959 and 1998.

The paper is structured as follows: In section 2, we provide a background of earlier works regarding published databases and collection of information on geo-hydrological hazards, damage, and related costs. In section 3, we describe the procedure adopted for constructing our database. In section 4, we describe the study area and apply the procedure to the case study. In section 5, we analyze the collected data. In section 6, we conclude by discussing the results and implications.

2. Previous works

A number of databases with information on natural hazards, including geo-hydrological hazards, were compiled from global to local scales and used for research, insurance, and economic purposes. A relatively large number of authors have attempted to organize and publish databases on geo-hydrological hazards, the majority of which were hazard- rather than impact-based. A brief list of database examples conducted at the national or regional scales are given here, representing those databases organized for geo-hydrological hazards, having newspapers, online news, and/or other organized and published databases as sources of information.

At present, few global multi-hazard databases are available, e.g., EM-DAT (http://www.emdat.be), NatCat (Munich Re, 2011), and Sigma (Swiss Re, 2017). The first published database of historical information on landslides and floods at the national scale, based on newspaper screening and analysis, was built during the AVI project (Aree Vulnerate Italiane; in English: "damaged urban areas") (Guzzetti et al., 1994; Guzzetti and Tonelli, 2004) and contains more than 9300 landslides and flood phenomena that occurred in Italy during 1919–2001. More recently, an extensive catalogue of flooding phenomena in Greece was compiled by Diakakis et al. (2012), based on a large number of reports and documents collected from several sources, demonstrating that urban environments exhibit higher flood recurrence rates than mountainous and rural areas. A set of newspapers were selected by Zêzere et al. (2014) and were systematically surveyed for collecting data on disasters and constructing a consistent and validated hydrogeomorphology database for Portugal. A variety of information sources, such as scientific publications, field data, and agency and press archives, provided basic information on the occurrence dates and impacts of damaging events for building a landslide database for Germany (Damm and Klose, 2015). To supplement existing records of the national landslide databases in Great Britain (i.e., by the British geological survey (BGS)), Taylor et al. (2015) updated the records on landslide events and their impacts by searching an electronic archive of regional newspapers and implementing a method for obtaining information from regional newspaper archives. A national database on damage caused by landslides, floods, and debris flows was built by Andres and Badoux (2019) for Switzerland; they collected information on damage from 1972 to 2017 for identifying potential connections to climate change. Regarding Italy, Calvello and Pecoraro (2018) realized "FraneITalia", i.e., a catalogue of 8931 landslides, which occurred in Italy from 2010 to 2017, using exclusively online news outlets. With regard to floods, Isacco et al. (2018) prepared "Floodbook", i.e., an online platform aiming at becoming a database for flood-damaged areas in Italy, while filling the knowledge gap regarding flood events occurring in small and ungauged basins. Sinkhole events, both of natural and anthropogenic origin, have been collected by Parise and Vennari (2013, 2017) by searching different sources of information for building a chronological catalogue of sinkholes in Italy. In Italy, to investigate the dependence of mortality on gender and age, both during fatal landslide and flood events, Salvati et al. (2018) organized data on 2063 fatalities that occurred in the country over the 50-year, 1965–2014 period, using multiple sources of information, including newspapers.

Valenzuela et al. (2017), based mainly on chronicles, created the spatiotemporal database "BAPA" containing 2063 landslides that occurred in northwestern Spain from 1980 to 2015, as well as information on related damages and costs. They also created a website and an app for increasing the number of landslides collected and minimizing the bias caused by the use of newspaper information. For cities on the Spanish Mediterranean coast, Gil-Guirado et al. (2019) used the digital archives of the main newspapers and built a database at municipal scale and with daily temporal resolution containing 3008 flood cases from 1960 to 2015, and collected information on damage, intensity, and severity of the events. Graff et al. (2019), using information from several geographical databases, proposed an approach for characterizing the elements at risk for the Normandy region of France.

To address the lack of available precise damage data for calibrating and validating flood-risk assessment models in southern France, Saint-Martin et al. (2018) published DamaGIS, which aimed to collect and assess flood-related damage data at the local scale using all types of information. Examples of site-specific dataset recorded events for limited areas became available by Paliaga et al. (2019), who published an inventory of geo-hydrological phenomena in the Genova municipality in northwestern Italy regarding landslides and floods, their interactions with built-up areas, and cultural heritage sites, thereby allowing a first risk analysis. Garcia-Urquia and Axelsson (2014) compiled a database for rainfall-induced landslides occurring in Tegucigalpa, Honduras, based on news reported by two local newspapers.

Although the outlined literature represents a partial list of existing catalogues, the analysis highlights the following: (i) The heterogeneity of the inventory structures, (ii) their frequent focus on a single type of hazard, (iii) the shortage of inventories including information on triggers, other related phenomena, connected damage, and costs (De Groeve et al., 2014), and (iv) the lack of structured information on source-evaluation criteria, based on efficiency, quality, and availability indices.

3. Methodology

In this section, we describe the procedure devised for registering information regarding historical geo-hydrological phenomena (i.e., landslides, floods, and sinkholes), as well as related triggers, damages, mitigation works, and costs. The steps of the procedure are illustrated in Fig. 1.

3.1. Agreement on definitions

The first step of the procedure (Fig. 1A) was theoretical and concerned the full sharing, as well as understanding, of some definitions and choices. First of all, it was defined the most effective IT solution for facilitating the data management and publication. Then the type of hazards to be recorded were selected taking into consideration the stakeholder's requests and the characteristics and extension of the study area. It was also essential to set the level of detail (the data granularity) with which hazards and impacts can or should be described. A quality index was conceptualized, designed and applied to evaluate the quality of the information sources (bibliography) used to compile the inventory.

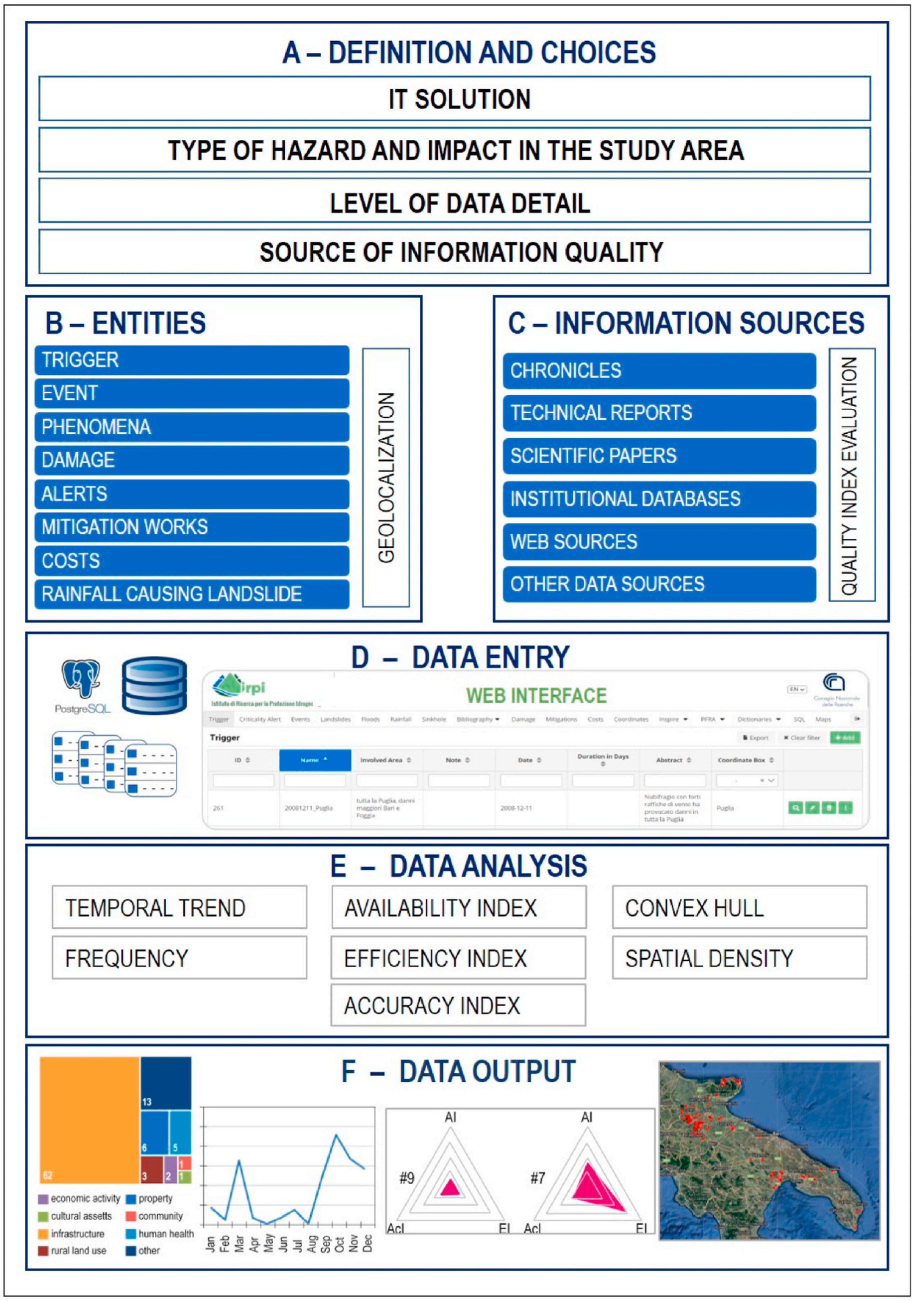


Fig. 1. Outline of the procedure designed for building AReGeoDatHa.

3.2. Entities

The second step (Fig. 1B) involved the identification of different entities to be used for describing the geo-hydrological phenomena and related features. The starting point of this activity was the work of Napolitano et al. (2018), in which groups of entities are identified and considered fundamental to characterize landslides, floods and their related effects. These entities include: Trigger (i.e., the major meteorological, seismic or anthropic causal factor of the geo-hydrological phenomena), Event (i.e., the spatio-temporal ensemble of single phenomena caused by the Trigger), Phenomenon (e.g., debris flow, flash flood), Damages (e.g., people, roads, buildings), Mitigation (i.e., safety interventions, riverbank restorations), Cost (i.e., cumulative or partial monetary value related to the damage or a mitigation), Geolocalization (of Trigger, Event, Phenomenon, Damage, Mitigation, by means of points, lines, polygons). More details about the different entities are provided in section 3.4.

For the purposes of this work, we have included three further entities: Sinkholes (a geo-hydrological hazard that frequently occurs in Apulia region), Criticality Alerts (i.e., warning levels issued by authorities in charge the day before the occurrence of the geo-hydrological event) and Rainfall Causing Landslide (i.e., duration and cumulated precipitation responsible for triggering a landslide). The two latter entities are closely related to each other and useful for civil protection issues. Criticality alerts are important to understand if and to what extent the warnings issued by civil protection authority are then confirmed by the occurrence of geo-hydrological phenomena. Rainfall Causing Landslide is relevant for the definition of rainfall thresholds predicting the occurrence of slope movements (e.g., Brunetti et al., 2010) and useful for early warning systems implementation. This latter entity therefore has the purpose of identifying accurate and very specific information on individual landslides. Although it contains information on the most frequent cause of the landslide occurrence (rain), it does not fall within the more general definition of the Trigger. The Trigger entity is in fact used to represent and describe the most general meteorological factor, which affects wide territories and cannot be traced back to the single accumulation and duration of rainfall recorded on a single sensor located near a single landslide.

3.3. Sources of information

The third step involved the selection of the most appropriate information sources regarding geo-hydrological data (Fig. 1C). For our purposes, we selected newspaper articles, technical reports, and scientific papers, in addition to existing institutional databases and other web sources. Reliable, accurate, and complete information on geo-hydrological phenomena is seldom obtained from a single information source. In most cases, inter-comparisons of multiple sources improve the information and allow the exclusion of less reliable sources. Inter-comparisons are made based on the time and/or location of known entities. The time and/or location of phenomena and damage are useful for quickly looking into other sources for additional and more detailed information; if more sources are available, additional details can be added. In the days immediately after the most serious damaging events, newspaper articles and chronicles report rough economic estimates of the damage. Only after the official inspections made available by the competent authorities, is it possible to collect certified data on the costs and mitigation works financed or authorized. Regarding rainfall data, considering their implication in the estimation of rainfall thresholds for landslide triggering, it is necessary to consider only data provided by, or derived from, designated agencies (in our case provided by National Department of Civil Protection). For cost, and mitigation, inter-comparisons are made with scientific sources, technical reports or dedicated papers. Overall, four main categories of information were used: (i) newspaper articles, (ii) technical reports and scientific papers, (iii) institutional databases, and (iv) web sources. Multiple sources were allowed for checking the truthfulness and goodness of the information by implementing an inter-comparison activity for validating the information content and maintaining the quality standard of the collected data. The procedure used for the selection and collection of the information is described step by step in Appendix A.

3.4. Data-entry interface and database structure

The fourth step (Fig. 1D) involved selecting the database structure and the related data-entry interface to be used for registering and storing information. In this regard, LAND-deFeND (Napolitano et al., 2018), a freely available and validated database structure (http://geomorphology.irpi.cnr.it/tools/land-defend-database-structure), has been modified and implemented with additional features and a web interface for data-entry and consultation. The objective was twofold: (i) to be as compliant as possible with the European directives (INSPIRE, FLOOD and their national implementations) and (ii) to facilitate data entry by operators. LAND-deFeND is a relational database with a hierarchical structure based on the definition of the following four groups of entities: (i) Natural (NE), (ii) Human (HE), (iii) Geospatial (GE), and (iv) Bibliographical (BE). The NE group includes the physical processes such as the Trigger (blue in Fig. 2A), which may encompass one or multiple Events (green in Fig. 2A), which in turn represent one or multiple Phenomena (yellow in Fig. 2A).

This hierarchical structure, which fits perfectly with the entities defined in section 3.2 and which consider the Trigger entity as an essential element (followed by the Event and Phenomenon in cascade), is in accordance with the Italian implementations of the INSPIRE Directive. Question marks in the figure highlight that the database structure is designed to allow the registration of entities even when information on some intermediate entity is missing. The HE group represents the interaction of the physical processes with the anthropic sphere and includes Damages (red triangles in Fig. 2A), Mitigation Works (brown triangles in Fig. 2A) and Costs (purple and light blue ellipses in Fig. 2A, depending, respectively, if a single cost is cumulated, i.e. related to multiple damage/mitigation, or single).

In Fig. 2B we show another way to represent the conceptual schema of the database, made to include the GE (grey boxes) which can be associated to the NE (yellow boxes) and HE (red boxes). In the same figure we can see that, in the database, the BE can be related to (i.e., allow the storage of the information source on) NE, as the Trigger or the Phenomena and/or HE (as the Damage and the Mitigations). Compared to the original, the updated version of the database structure includes additional entities such as Sinkholes, Rainfall causing landslides and Criticality alert. Sinkholes were added because of the characteristics of the Apulian territory, where this type of geological hazard, which can have both natural and anthropogenic origin, is highly frequent (Parise, 2010b, 2015; Fiore and Parise, 2013).

Geo-hydrological hazards (i.e., landslides, floods, and sinkholes) occur in response to a single Trigger, such as an intense rainfall event, a prolonged rainfall period, a rapid snowmelt event, a human action or an earthquake. Multiple damaging processes (i.e., phenomena) that occur in response to a single trigger can cause a cumulative socio-economic impact. Fig. 2B shows the possible relations among the entities inside a single trigger (T_2) that correspond to a typical geo-hydrological disaster.

The data-entry operational sequence depends on the hierarchical relationships existing through the NE, HE, GE and BE entities (Fig. 2A). The Trigger is at the top; thus, it is mandatory to insert at least one Trigger to register any other entities. The Trigger must necessarily be characterized by a geospatial component indicating the area affected by meteorological, earthquake or human induced events. In this regard, the web interface allows users to record geometries by drawing multi-polygons, lines, or points on top of a base layer (i.e., *OpenStreetMap)*, thereby allowing, in the case of the multi-regional/national database, the analysis of the spatiotemporal evolution of the meteorological

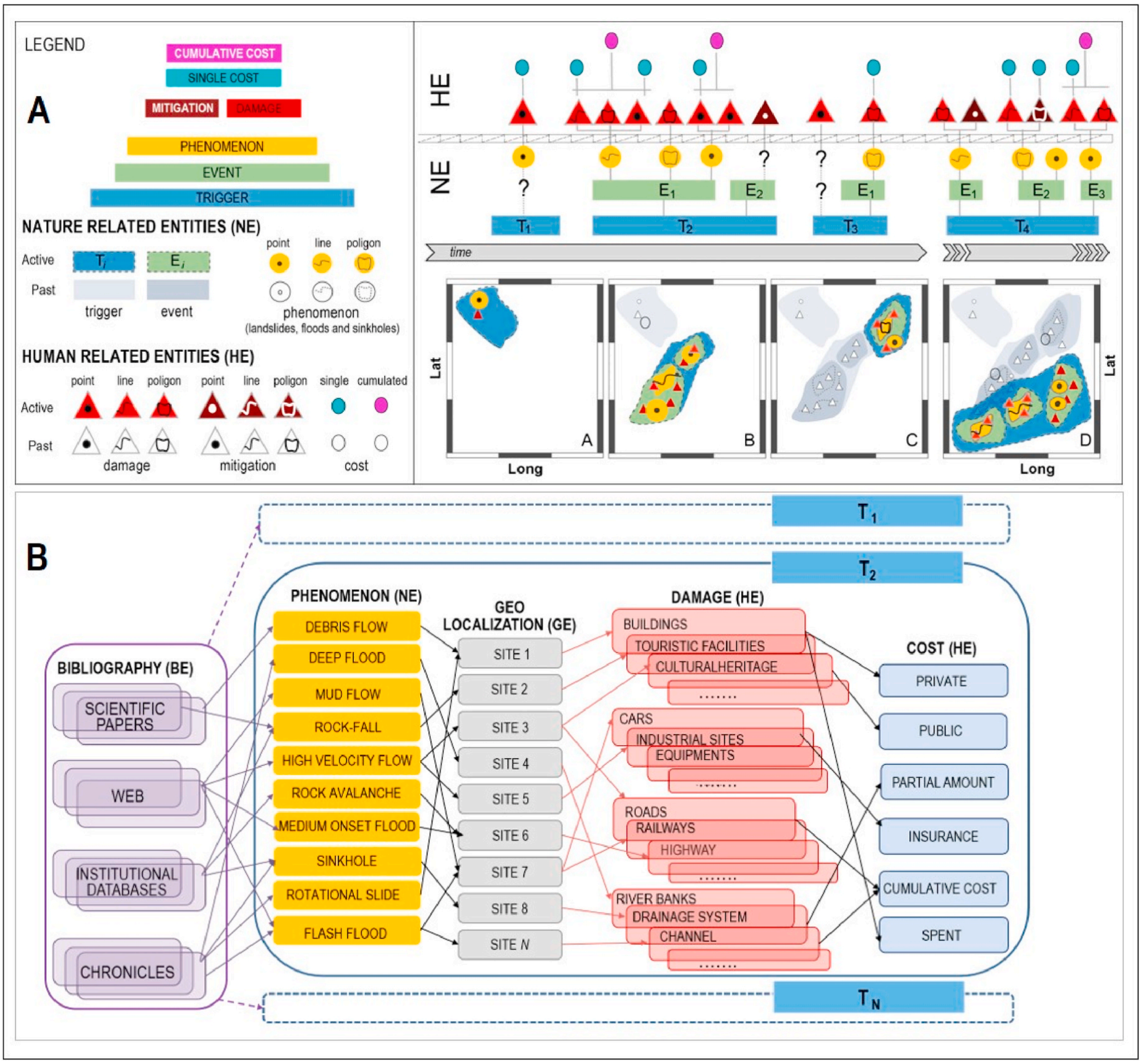


Fig. 2. Conceptual representations (A and B) of the LAND-deFeND database structure. A): temporal (above the grey time line) and spatial-temporal (below the grey time line) relations between the Nature-related Entities (NE) and the Human-related Entities (HE) (modified after Napolitano et al., 2018). B): schematic relationships between the four groups of entities: in the example the Bibliographic entities (BE) can be related to one (T_2) or more (T_N) Trigger (NE), Geolocalization (Geospatial entities GE) can be related to one or more phenomenon (NE) or damage (HE).

perturbation, which produces different types of phenomena and damage based on its persistency, rainfall intensity, and the geomorphological assets of the involved territories.

Events are defined as spatio/temporal ensembles of Phenomena; therefore, more than one Event can belong to the same Trigger, even if they differ regarding the location and start/end time. Geometry entities (i.e., coordinates and shapes) can be used to geolocate and characterize Trigger, Event, geo-hydrological phenomena (Landside, Flood, Sinkhole), Damage, and Mitigation works. Cost entities can be registered in the database to better describe damage and mitigation works they are related to. A singe Cost can be related to one or multiple Damage/Mitigation, depending on the granularity of the available information. The BE (e.g., newspapers, books, scientific papers, etc.) can be registered in the database, together with a pdf or an image file. All the entities mentioned in the present section were translated into PostgreSQL tables, when the conceptual model was implemented in a physical one. Please refer to Napolitano et al. (2018) for details about the conceptual, logical and physical modeling of the database structure.

3.5. *Data analysis and views*

Data elaboration constitutes the fifth step of the procedure. The web interface includes predefined views that we implemented, allowing for statistical analysis and producing maps compliant with the INSPIRE and FLOOD directives (Fig. 1E, F). Additional views were implemented for generating elaborate plots and other geospatial analyses able to quantify the impacts of geo-hydrological events, classifying the related damage, and quantifying the restoration and mitigation costs.

A specific tool of the web interface allows an expert user to manage and query the stored data while using the SQL language. Predefined views in accordance with the INSPIRE directive and the preliminary flood-risk assessments of the FLOOD directive are available at the web interface. The data can be exported easily in different formats.

3.5.1. *Source of information quality index - SIQI*

The collection of data on geo-hydrological phenomena represents the basic activity for all the subsequent data analyses; therefore, the collected information must be as reliable and accurate as possible for maximizing the number of details included. For this purpose, we paid particular attention to the evaluation of different sources of information.

To quantitatively measure the characteristics of each source of information we devised a criterion for ranking the availability, efficiency, and accuracy of each information source used during the data collection activity that, cumulatively, provides a total quality index, i.e., the source of information total quality index—SIQI (Fig. 1E). Table 1 lists the three indices with increasing degree of complexity, i.e., the availability, efficiency, and accuracy indices (AI, EI, and AcI, respectively). Each index is generated by a combination of three numeric indicators obtained through different evaluation grids.

AI is generated by means of three double choice (i.e., 0 or 1) indicators: the x indicator is dedicated to the cost options that could be free (1) or paid for (0); the y indicator describes the type of support that could be digital (1) or paper (0); and the z indicator is for evaluating the usability that could be public (1) or private (0). To generate the final index score, we assigned to each indicator a number using the following equation (Eq. (1)) that is both positional and numeric:

$$AI = x^*2^0 + y^*2^1 + z^*2^2 \qquad \text{Eq.1}$$

EI is generated by means of three double choice (i.e., 0 or 1) indicators: the α indicator is dedicated to the terminology that could be proper (1) or general (0), the β indicator is dedicated to the media file that could be reliable (1) or stock images (0), and the γ indicator is dedicated to the informative content that could be good (1) or poor (0). To evaluate the γ indicator in the most objective way, for each news piece or document in bibliography, a form was completed containing six different questions about the presence (1) or absence (0) of information regarding the place, date, triggering factor, type of processes, damages, and cost estimate. When the sum of the single score was <3, the γ indicator was evaluated as poor (i.e., 0); if it was ≥3, the γ indicator was good (i.e., 1). To generate the final score, we assigned a number to each indicator using Eq. (2):

$$EI = \alpha^*2^0 + \beta^*2^1 + \gamma^*2^2 \qquad \text{Eq.2}$$

Table 1
List of the availability, efficiency, and accuracy indices (AI, EI, and AcI, respectively). For each index the relative equation is reported together with the indicators, items, options, and possible score used for its calculation.

Index	Indicators	Items	Options		Score
Availability -AI Eq. 1	x—Cost	for free/paid for			1/0
	y—Support	digital/paper			1/0
	z—Usability	public/private			1/0
Efficiency - EI Eq. 2	α—Terminology	proper/general			1/0
	β—Media file	reliable/stock			1/0
	γ—Informative content	place provided: yes/no	1/0	∑ ≥3;	1;
		date provided: yes/no	1/0	∑ <3	0
		triggering factor provided: yes/no	1/0		
		type of process provided: yes/no	1/0		
		type of damage provided: yes/no	1/0		
		cost estimate provided: yes/no	1/0		
Accuracy - AcI Eq. 3	a—Temporal	≤24 h	high		2
		24 ÷ 168 h	medium		1
		>168 h	low		0
	b—Spatial	≤1 km	high		2
		1 ÷ 10 km	medium		1
		>10 km	low		0
	c—Dimensional details	2 or 3 D size estimate	high		2
		1D size estimate	medium		1
		Qualitative or no dimensional data	low		0

AcI is constituted of three triple choice (i.e., 0, 1, 2) indicators, and it assesses the degree of accuracy of the information (i.e., data) reported by the different sources. To quantitatively assess the three indicators, we referred to the accuracy intervals defined in the dictionary tables of the LAND-deFeND structure, which are specifically dedicated to measuring the temporal and spatial accuracy (see Table 1) of each record in the database. The tables list specific intervals expressed in hours (i.e., temporal) or km (i.e., spatial), which were used for assigning a numeric score to each accuracy indicator (see Table 1). The a accuracy indicator measures the temporal accuracy, while the b accuracy indicator quantifies the spatial accuracy, both of which are classified as high (2), medium (1), or low (0). The c accuracy indicator evaluates the presence or absence of dimensional details (i.e., length, width, depth, area, and volume) describing the physical processes (i.e., phenomena), and it is classified as follows: high (i.e., 2) when 2D or 3D size estimates (i.e., areal or volume) are provided, medium (i.e., 1) when 1D linear dimensional data are provided (e.g., length), or low (i.e., 0) when no information or only qualitative estimations (e.g., *a large landslide*) are provided. To generate the final score, we assigned a number to each indicator using the following equation:

$$AcI = \alpha^*2^1 + b^*3^1 + c^*4^1 \qquad \text{Eq.3}$$

Finally, we define the SIQI according to the following equation, as the sum of the normalized score indices ()’:

$$SIQI = \left[(AI_{\#})' + (EI_{\#})' + (AcI_{\#})' \right] / 3 \qquad \text{Eq.4}$$

4. Application of the procedure

The procedure described in section 3 allowed the realization of the Apulian Regional GeoDatabase for geo-hydrological Hazards (AreGeoDatHa).

4.1. Apulia Region study area

Apulia is the southeastern part of the Italian Peninsula, extending in the Adriatic and Ionian Seas toward the Balkans and Greece, and is an elongated peninsula with over 860 km long coastline (Fig. 3). Geologically, it acted as a foreland during the building up of the Apenninic Chain in the Miocene, and consists of several thousands of meters thick carbonate bedrock formed in the Tertiary that remained essentially undeformed (Patacca et al., 1990; Patacca and Scandone, 2007). It was subsequently covered by Quaternary clastic carbonates and was characterized since the Lower Pleistocene by a general uplifting until it reached its present configuration (Doglioni et al., 1994). Topographically, it exhibits its maximum elevations (i.e., slightly over 1000 m a.s.l.) in its northern sectors, namely in Daunia (NW) and in the Gargano promontory (NE).

Apulia is one of the main karst regions in the Mediterranean area and is characterized throughout most of its extent by soluble rocks (i.e., mostly carbonates, with subordinate evaporites). However, there are differences among the main sectors of Apulia, with Daunia (i.e., the NW sector, which is the transition to the southern Apennine Chain) strongly differing from the rest of the region. Daunia terrains mainly consist of prevailing fine-grained material, forming the typical flysch sequences in the area. Regarding landslide types, due to the mostly clayey presence and the low to medium relief (i.e., gentle slopes in the 7–15° range), most of the phenomena belong to the flow-type categories, even though they often start as rotational or translational slides (Zumpano et al., 2020; Spalluto et al., 2021). The poor geotechnical characteristics of the clay-rich flysch units (Cotecchia et al., 2015) are the origin of most of the observed phenomena and reactivation (Parise et al., 2012; Refice et al., 2019; Wasowski and Pisano, 2020; Diprizio et al., 2021).

Fig. 3. Map of the study area (Apulia region).

Sinkhole occurrence with various typologies and triggers has been documented in Apulia for many years (Delle Rose and Parise, 2002; Bruno et al., 2008; Del Prete et al., 2010; Fidelibus et al., 2011; Festa et al., 2012; Margiotta et al., 2012, 2016, 2021). Sinkholes have repeatedly caused serious damages to infrastructures, buildings, and human life, with more than 650 reported evacuees and injuries, as well as one fatality (Fiore and Parise, 2013; Parise and Vennari, 2013, 2017). Given the widespread presence of underground voids of both natural and anthropogenic origin, sinkholes are among the most serious geological hazards in Apulia, with most of the dangerous situations being related to anthropogenic underground cavities. These cavities have been excavated since the ancient times because of distinct albeit complementary needs: urban development demanding building materials, or the need for continuous development of agricultural practices at the surface (Parise, 2010a; Negri et al., 2015).

With time, the loss of memory (which is very often the reason for geohazards; see Calcaterra and Parise, 2001), combined with the urban expansion above the areas originally intended for underground quarrying, resulted in the development of sinkholes, typically occurring catastrophically and without any warning signs at the surface.

Regarding the precipitation patterns in the Apulia region, the average annual rainfall over the entire region is approximately 640 mm. Except for the stations at higher altitudes, i.e., located in Gargano and reporting annual rainfall between 800 and 1200 mm, annual rainfall ranges between approximately 450 mm at part of the Tavoliere delle Puglie plain and the Ionian coast near Taranto, and 800 mm in the inland areas of Murgia and Salento, with record values being approximately 700 and 800 mm, respectively (Casarano et al., 2019). More than 66% of the average annual rainfall is observed in autumn to winter; in summer rainfall exceeds 100 mm only in the rainiest areas of Gargano. In the two decades 1980s and 1990s recurrent drought periods were observed (Polemio and Casarano, 2008: Lionello et al., 2014; Doglioni and Simeone, 2019).

4.2. Advancement in the database structure

The particular abundance of sinkhole phenomena occurring in the Apulia region was the basis for modifying LAND-deFeND database structure and implementing the possibility for recording this phenomenon. The choice required not only deploying a new database entity, but also registering a number of different dictionaries (Table B 1; Appendix B) that are used for characterizing single sinkholes. The *sinkholes* table includes 23 fields, some of which are common with other types of phenomena (i.e., landslides and floods), while others are related through foreign keys to six new dictionary tables that have been added for including the nature (Parise and Vennari, 2013; Parise, 2015), origin (Gutierrez et al., 2014; Parise, 2019), classification, triggering factor, and type of the new entity. In addition, we implemented the table *criticality alerts*, containing information about the alert warning system published daily by the regional civil protection agency. The table records contained the type of criticality (i.e., hydrological, geo-hydrological, storm), level (i.e., no significant phenomena expected, ordinary, moderate, high), date (i.e., day, hour), name of the authority in charge, validity (i.e., time) of the forecast, and territory (i.e., alert zone). Many of these fields are secondary to dictionary tables.

To characterize the rainfall event responsible for landslides triggering a new entity, a Rainfall causing landslide was introduced in the physical model through the creation of a table (i.e., *rainfall_event)* for registering the duration and cumulative rainfall data measured at the time of the triggering of one or more landslides. We used data derived from the "IRPI-DPC rainfall-events catalogue" (Vennari et al., 2013). Raingauge data of the closest rain gauge to the landslide in which the rain was measured (*cumulated_rainfall_mm, rainfall_duration_hour*) were calculated using the tool proposed by Melillo et al. (2018) for defining objective and reproducible rainfall thresholds (Brunetti et al., 2010; Peruccacci et al., 2012). In the table *landslide*, a foreign key defines the relationship with the *rainfall_event* table (which represents the Rainfall causing landslides entity*)*.

The data-entry web interface (step 4, section 3.4) was used for recording information from 2008 to 2019 regarding 54 geo-hydrological disasters (i.e., Triggers), and 133 data on landslides, 222 on floods, and 38 on sinkholes, resulting in 663 records dedicated to damages in Apulia.

Data analysis (section 3.5; Fig. 1E) allowed the statistical evaluation of the data, as described in the following section.

5. Results

Results of this work are various and concern (i) a description of the data recorded in the database, (ii) the evaluation of the geographical persistency and temporal frequency of geo-hydrological processes and their related damage and direct cost, (iii) the analysis of the data in relation to earlier published catalogues available for the Apulia region (AVI archive) and to the daily geo-hydrological criticality alerts bulletins issued by the responsible authorities, and (iv) the analysis of the Source of Information Quality Index (SIQI) values.

5.1. Recorded data

Data on geo-hydrological events generating one or multiple phenomena and/or damage during 2008–2019, were entered in AReGeoDatHa and are now available for downloading at https://doi.org/10.5281/zenodo.5898539 including the metadata and the layer description.

The actual version of the database contains i) 488 natural entities, ii) 864 human entities, iii) 772 geospatial entities, and iv) 156 information-source entities. Fig. 4 shows an overview of the geographical distribution of the sites for which precise or approximate information on the geospatial entities (i.e., latitude and longitude) could be collected in the Apulian database, related to natural (in green) and human (in red) entities. Visual inspection of Fig. 4 reveals that, during the study period, sites affected by geo-hydrological phenomena were concentrated in the Gargano and Daunia areas (i.e., northwest), especially regarding landslide phenomena, and along the Adriatic coast and in proximity to the Gulf of Taranto (i.e., southeast), regarding mostly flooding. In the white boxes, we highlight the coexistence of different geographical entities (i.e., points, lines, and polygons) referred to natural or human entities.

Overall, 54 Triggers and 41 Events have been surveyed, most of which have slope and fluvial dynamics. There are triggers with no events and triggers with more than one event, as far as the database structure allows to do. The triggers with the highest number of events, which differ in spatial extension, occurred during October 2009, involving the Bari, Foggia, and Brindisi Provinces (i.e., ~12.500 km^2), and in November and December 2013, involving the Foggia, Bari, and Taranto Provinces (i.e., ~13.500 km^2). The trigger with more events that differed for temporal extension occurred in September 2014 in the Gargano area (i.e., northwest). This distinction was made on the basis of rainfall analysis, through which four main rainfall events were identified (Martinotti et al., 2017).

Fig. 5 summarizes the data grouped by phenomenon type (left: landslides; middle: floods; and right: sinkholes). The landslides recorded in AReGeoDatHa are 133 and are classified as follows: 81 generic, 13 rock falls, 12 earthflows, 10 mud flows, 10 debris flows, 4 slides, 2 flow slides, 1 rock slide. Most of the landslides affected the Foggia Province (i.e., north), while San Marco in Lamis is the most impacted municipality (dark red in Fig. 5a). Regarding the temporal accuracy of landslide phenomena, for 44 and 26% of the cases, the weeks and days of occurrence are known, while for 15% of the cases, the temporal occurrence uncertainty is in the order of hours. To record geospatial entities of landslides, we used *coordinates* as points or polygons. The spatial accuracy of geo-referenced landslide phenomena is very high, i.e., less than 100 m for 80% of the elements.

The trigger that caused the highest number of landslides (i.e., 40) was in March–April 2016 (Fig. 5d); it was an extreme weather event mainly focused on the Foggia Province, and in particular the Daunia area (i.e., northwest). A state of emergency was declared, several inspections were conducted by the regional civil protection agency, and, consequently, precise data regarding the location of phenomena and damaged elements were available. The highest monthly occurrence rate of landslides occurred in March and the second highest in September (Fig. 5e). Rainfall-triggered landslides with associated rainfall parameters from the IRPI-CNR rainfall-event catalogue were recorded in the database for 22 distinct landslide phenomena. The most damaged elements are the road network, railways, and network infrastructure, and are grouped together in the infrastructure category (in orange in Fig. 5m), as regulated by the European FLOOD directive. Very few landslides damaged cultural assets or economic activities. These results are in line with those documented in Daunia, where damages due to landslides and land-use changes play a significant role in the depopulation of the area (Wasowski et al., 2010; Pisano et al., 2017a, b).

In AReGeoDatHa, 222 flood phenomena were recorded during 2008–2019. Their distribution at the municipal scale is represented in Fig. 5b, where Bovino and Castelluccio Valmaggiore (in the Foggia Province) are the most affected villages during the study period. The trigger that caused the largest number of floods (i.e., 50) affected the Foggia, Bari, and Taranto Provinces during November–December 2013

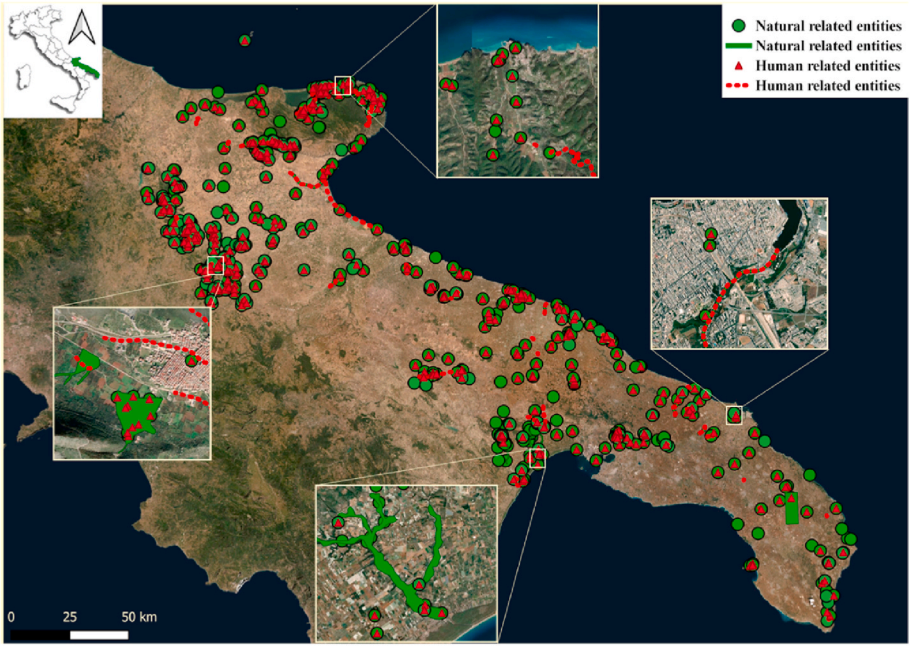


Fig. 4. Natural (green) and human (red) entity coordinates recorded in AReGeoDatHa. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

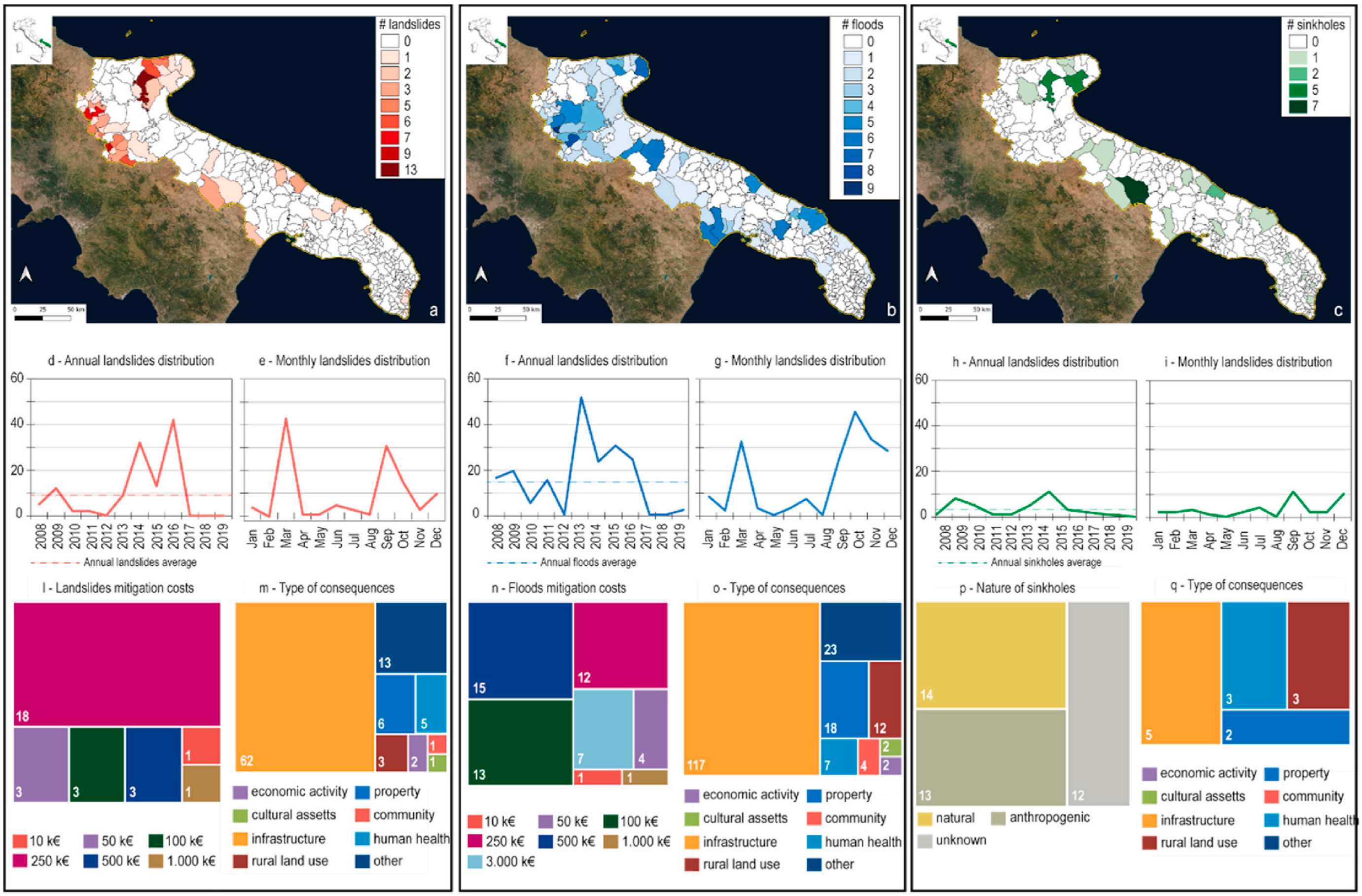


Fig. 5. Main features recorded in AReGeoDatHa for landslides (left), floods (middle), and sinkholes (right): (a, b, c) municipal distributions, (d, f, h) annual distribution, (e, g, i) monthly distribution, (l, n) mitigation costs due to landslide and flood phenomena, (p) nature of sinkholes, (m, o, q) type of consequence.

(Fig. 5f). Most of the surveyed floods have daily temporal accuracies, while 14% of the floods have temporal uncertainties in the order of a few hours. The monthly distribution of the recorded floods revealed a major peak in September and a secondary peak in March (Fig. 5g).

The spatial accuracy of the flood *coordinates* (points, lines, or polygons) was very high: less than 100 m for 80% of the records surveyed. Regarding the source of flooding (as required by directives 2007/2/EC), 56% of the floods are fluvial and 33% are pluvial. Floods are the phenomena that cause most of the recorded damage in the Apulian catalogue. To record damage information, the FLOOD Directive classification was used (Directive 2007/2/EC), which considers 11 categories and 111 subcategories of damaged objects. We implemented the classification with new damaged objects not included in the directive, considering that AreGeoDatHa includes also landslides and sinkholes. The purpose of the FLOOD Directive is to establish a framework for the assessment and management of flood risks, aiming at the reduction of adverse consequences for human health, the environment, cultural heritage, and economic activities. For this reason, for any type of phenomena, we also analyzed the "type of consequence" that the damage produced: human health, community, waterbody status, protected areas, environment, cultural assets, property, economic activity, infrastructure, rural land use, pollution sources, and other.

Communication infrastructure (orange in Fig. 5o) as longitudinal defense works (i.e., embankments, bank walls, paintbrushes), public and private buildings, and the agricultural sector (i.e., rural land use) were repeatedly affected (Fig. 5o). Damage to people, and in particular, the number of deaths recorded in the database, is mainly due to floods.

Over the study period, there are 38 sinkholes recorded in the Apulian geodatabase; the number of processes at the municipal scale is shown in Fig. 5c. Most of the sinkhole phenomena occurred in 2014; the worst month was September (Fig. 5h, i). These sinkholes were triggered by the meteorological event of September 2014 (Parise et al., 2018). The high percentage compared to the total sample is linked to the field inspections by the IRPI researcher team after the event. Most of the sinkholes are connected to cavities of natural origin (35.9%), while 33.3% of them are related to anthropogenic cavities (Fig. 5p). The remaining percentage refers to sinkholes for which the nature of the underground cavities is unknown. The temporal accuracy of the recorded sinkhole phenomena was weekly. The spatial geolocation accuracy of the sinkholes is very high for 75% of the phenomena, i.e., less than 100 m. Referring to damage caused by sinkholes, Fig. 5q shows that infrastructure and rural land use are the most damaged categories.

The costs related to mitigation works, financed and/or conducted, are mainly derived from the interventions financed after state of emergency statements (from 2004 to 2017), provided by the Apulian civil protection department. Regarding post-landslide interventions, costs exceeded 4,000,000 €, with interventions ranging between 3500 and 1,700,000 € (many of the costs are cumulated) (Fig. 5l). Road networks, safety measures, hydraulic works, construction of gabionades, and restoration work are among the most frequent interventions. Costs reserved for flood mitigation work were significantly higher, amounting to 14.000000 €. These are interventions with costs between 6.000 and 2.250.000 € (Fig. 5n): restoration of hydraulic functionality (i.e., embankment routes, damages to the hydraulic network, restoration of the river's viability), road network, water supply, and sewerage works for which the cumulative costs are provided.

5.2. Alert bulletins and other catalogues

The analysis of the geo-hydrological alerts daily issued by the civil protection authorities can be very helpful for understanding the presence of missed alerts. Warning levels were compared with the occurrence of geo-hydrological phenomena. Of the 41 damaging events collected for a total of 97 days, from 2008 to 2019, the corresponding alerts were the following:

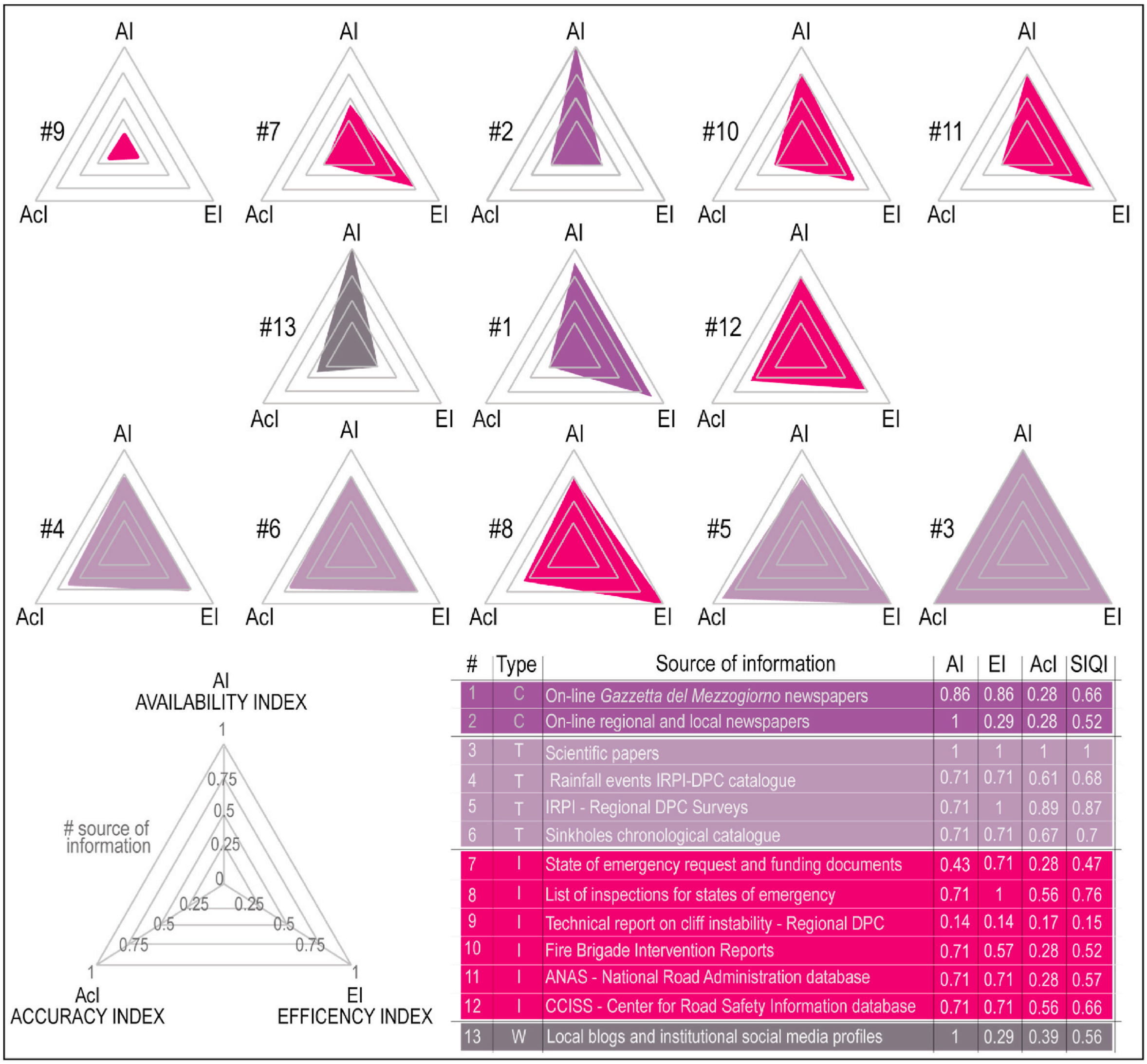
i 27.8% absence of significant phenomena had been forecasted;
ii 24.7% ordinary criticality for hydraulic and/or geo-hydrological risk;
iii 38.2% moderate criticality for hydraulic and/or geo-hydrological risk;
iv 9.3% high criticality for hydraulic and/or geo-hydrological risk.

We observed that more than a quarter of the events that caused damage were not predicted; the percentage rises to more than half if we also consider the cases in which ordinary conditions were predicted (but which actually include the possibility of damage to persons and property due to localized phenomena).

Currently AreGeoDatHa covers a period of 12 years (i.e., 2008–2019), during which the landslide, flood, and sinkhole frequencies were 9.75, 15.4, and 3.9 events/year, respectively; these values are larger than those previously reported (i.e., AVI project; Guzzetti et al., 1994; Guzzetti and Tonelli, 2003). In the 40-year 1959–1998 period, the AVI database for the Apulia region accounted for 165 landslides (i.e., 4.12 events/year), 371 floods (i.e., 9.3 events/year), and 28 sinkholes (i.e., 0.12 events/year). However, this comparison must be carefully considered because (i) intensive urbanization and land-use changes have made the territory more vulnerable to geo-hydrological phenomena, and (ii) in recent years, there are relatively more sources of information available. Nevertheless, since the source of information that provided most of the information in this study is the same used by AVI project (i.e., the newspaper "La Gazzetta del Mezzogiorno"), we cannot exclude that the increased number of events causing damages may also be, at least partly, related to the increased frequency of extreme weather phenomena (Diodato et al., 2019).

5.3. SIQI index values

When preparing catalogues of geo-hydrological events/phenomena, data collection is conducted using diverse information sources, as their combination can improve the quality of the recorded data. The SIQI



#	Type	Source of information	AI	EI	AcI	SIQI
1	C	On-line *Gazzetta del Mezzogiorno* newspapers	0.86	0.86	0.28	0.66
2	C	On-line regional and local newspapers	1	0.29	0.28	0.52
3	T	Scientific papers	1	1	1	1
4	T	Rainfall events IRPI-DPC catalogue	0.71	0.71	0.61	0.68
5	T	IRPI - Regional DPC Surveys	0.71	1	0.89	0.87
6	T	Sinkholes chronological catalogue	0.71	0.71	0.67	0.7
7	I	State of emergency request and funding documents	0.43	0.71	0.28	0.47
8	I	List of inspections for states of emergency	0.71	1	0.56	0.76
9	I	Technical report on cliff instability - Regional DPC	0.14	0.14	0.17	0.15
10	I	Fire Brigade Intervention Reports	0.71	0.57	0.28	0.52
11	I	ANAS - National Road Administration database	0.71	0.71	0.28	0.57
12	I	CCISS - Center for Road Safety Information database	0.71	0.71	0.56	0.66
13	W	Local blogs and institutional social media profiles	1	0.29	0.39	0.56

Fig. 6. Triangular charts comparing the SIQI, AI, EI, and AcI rescaled to a common range [0–1], with 0 and 1 assigned to the least and best performing values, respectively. Triangle areas indicate the SIQI, with larger (smaller) triangles reflecting higher (lower) quality. C: newspaper articles and chronicles; T: technical reports and scientific papers; I: institutional databases; W: web sources.

quantitatively measures the characteristics of each source of information, thereby allowing the quantitative ranking of the availability, efficiency, and accuracy of each piece of information, and cumulatively giving a final quality index.

In the following, we first briefly describe the main characteristics of the four types of information sources used in this work (i.e., newspapers articles, technical reports and scientific papers, institutional databases, and web sources), and we subsequently show the SIQI values calculated based on these characteristics. Appendix A reports a more detailed description of the information sources used.

Newspaper articles (#1–2 in Fig. 6) provide complete temporal coverage and information on most of the recorded data; however, they mainly focus on events with larger impacts on community life and public interests. No standard criterion emerged in their descriptions, while news terminology was not properly used. The accuracy of spatial and temporal information can be very variable, and it is usually correlated with the severity of the event. A validation process for this type of source was conducted for maintaining the quality standard of the collected data. Despite these limitations, newspapers articles constitute a uniform basis for most recorded events.

Technical reports and scientific papers (#3–4–5–6 in Fig. 6) provide thorough descriptions of single processes (e.g., a single landslide, flood mechanism, etc.) and very detailed technical information, but are limited to specific case studies. Data from scientific sources are characterized by high detail and rigorous terminology; however, the investigated time coverage, study area, and type of phenomena are limited to the specific aims of each study. Technical reports are mainly available for a few, very extensive, damage disasters, and scientific papers are available for specific study purposes. Depending on the specific aim of a single research project, scientific sources (i.e., papers and reports) can include information about restoration and mitigation works and their related costs. The scientific community can also make available specific databases compiled using newspapers articles and chronicles or other institutional databases. These latter sources are fundamental in acquiring data on damage compensation and cost assessments.

Institutional databases (#7–8–9–10–11–12 in Fig. 6) are more accurate because of the standardization of data type, terminology, and validation process; conversely, the information they provide is limited to their operational and/or institutional purposes. Databases may aim at specific aspects, cartography, phenomena typology, triggering mechanisms, hazards, vulnerability, and risk analyses. The collected information is not directly available even to insiders, and, in general, it is necessary to ask for data availability. Usually, this source of information is limited by the transition from analog to digital formats, thereby exhibiting problems related to software maintenance or compatibility, and even in the preservation of hardware support. Their operational reorganizations produce data gaps; consequently, gaps in temporal series are not rare. Analysis of paper archives is time intensive; therefore, analysis is often limited to digital data.

Web sources (#13 in Fig. 6) are represented by online newspapers, blogs, and social media; they can provide detailed information including photos and videos, and are useful for better characterizing the natural, human, and geospatial entities related to geo-hydrological hazards. In addition, online sources provide detailed information regarding the time occurrence because they are continually updated. Conversely, the reliability of these sources of information is often verified. Often, the same news are published on more than one website or blog, and despite the abundance of online information, no additional details are given. We usually refer to institutional profiles for news published on social media, or research other sources for inter-comparison, which increases the accuracy and detail of the data, and it is necessary for validating the information from the web sources.

The calculated SIQI values are shown in Fig. 6, where for each source of information, a triangular chart is provided, together with a table containing the corresponding values. The areas of the triangles measure the quality of the sources of information, with larger (smaller) triangles reflecting higher (lower) quality. The shapes of the triangles reveal the contribution of the three indices to the overall quality of the information. The scientific papers (i.e., #3) was the type of source with the highest scores for all three indices; conversely, the worst type of source was the technical reports on coastal cliff instability (i.e., #9) produced by the regional civil protection agency, which did not provide useful information. Although this latter source is one of the few certified and institutionally available sources, the systematic reading of all documents did not produce useful data. This is not the case for the other two certified institutional sources (IRPI - regional DPC surveys #5, list of inspections relating to states of emergency #8) that exhibit a very high SIQI score and both turned out to be very reliable sources. Inspection of Fig. 6 reveals that high AcI values are rare, and only one (i.e., #3) reaches the maximum score; the statistics of this index (i.e., *mean* $\mu = 0.48$; *mode* $Mo = 0.28$) are the lowest among those of all three indices. Conversely, the sources with a low or very low AI value (i.e., #9–7) are rare; the mean value of this index ($\mu = 0.73$) is the highest among those of the three indices. Lastly, EI exhibits a more gradual distribution (i.e., *mean* $\mu = 0.67$; *mode* $Mo = 0.71$).

6. Discussion and conclusion

Geodatabases can be built using different types of information sources, ranging from newspapers articles and chronicles to scientific reports, and they can be used for different purposes, such as for performing hazard and risk analysis, calibrating risk assessment models, and designing civil defense management plans. The goodness of results mostly depends on the quality and reliability of the data. In this work, particular attention was paid to the quality of information on geo-hydrological phenomena and their effects. When dealing with the collection, organization, and storage of geo-hydrological data, two of the many problems are certainly the most relevant: (i) finding available, reliable, and accurate sources of information, and (ii) making use of a flexible database structure that allows the collection of as many data as possible. To address the first issue, we conceptualized and designed a criterion for evaluating and classifying the different sources of information used for the data search. This need has grown over the years due to the considerable increase in the number and type of documentary sources because of the spread of the web. To the best of our knowledge, this is the first time that a quantitative analysis has been applied to the information sources with the aim of finding different performance indices, the combination of which gives a quality index (SIQI). Beyond the ranking obtained and analyzed (section 5.3) regarding the different sources used for compiling AReGeoDatHa, the designed criterion is in itself an important result: it allows the assessment of three different indices, which are useful for quantifying the quality of the data source. The approach designed and tested for the Apulian catalogue can be easily applied to other information sources in future research. The designed indices become a means for objectively and quantitatively evaluating the sources used for compiling the catalogue, and if included as metadata, they could reinforce the trust in the collected data. The indices could add value to similar databases that use local or national newspapers or other non-institutional sources that do not recognize the quality and reliability of their data.

Quality indices are widely used in many fields, e.g. journalism. Globally recognized trust indicators (https://thetrustproject.org/about/) have been created for amplifying the journalism's commitment to transparency, accuracy, inclusion, and fairness. Similarly, the indices presented in this paper can easily assess whether the catalogue information content comes from credible sources.

Using different sources of information and covering long periods, the details on the collected information may not be complete or uniform. For this purpose, a flexible database structure, specially designed for recording and managing data on geo-hydrological processes, should be used. The LAND-deFeND structure was adopted for the Apulian database. The original structure (Napolitano et al., 2018) was implemented and adapted to the specific stakeholder requests, thereby confirming its broad applicability. The new version includes the following new entities: (i) natural and anthropogenic sinkholes, (ii) rainfall causing landslides, and (iii) dedicated tables for recording the warning alerts issued by the regional/national civil protection agency. Sinkhole occurrence in Apulia is essentially due to the karst nature of most of the region, combined with the extremely frequent artificial cavities of various typologies (Galeazzi, 2013; Parise and Vennari, 2013). Sinkholes activated by natural causes typically occur because of storms or, less frequently, earthquakes (i.e., triggers). Because of their importance in Apulia, sinkholes have been integrated into the conceptual model as a natural entity, belonging to the phenomena in the logical model of the database.

To increase the information content of the Apulian database, specifically with regard to the triggering factors of landslides, an additional entity, i.e., the Rainfall causing landslide, has been added. At regional and global scales, empirical rainfall thresholds are among the most used tools for predicting rainfall-induced landslides (Rossi et al., 2012). Storing information on rainfall events responsible for landslide triggering is very useful for civil protection actions. The introduction of the *rainfall_event table* (which implement the Rainfall causing landslides entity) allows relating rainfall data with data on landslide processes and the consequences they cause. The new relations allow for deeper and detailed data analyses and broaden the application possibilities in the scientific field, as well as in the operational field of civil protection forecasting systems. Finally, in order to verify and validate the alerts issued by the regional and national civil protection authorities regarding the expected geo-hydrological criticality for a given territory, a specific table (*criticality_alert*) has been included in LAND-deFeND. Including these new tables was essential because the efficiency of the forecasting system and effectiveness of phenomena prediction can only be assessed *a posteriori* through the analysis of the information collected. The first results show that, for the time period studied, in at least a quarter of the cases, the warnings were not able to predict the events that caused damage.

For most landslides, floods, and sinkholes in the catalogue, information on the damage and cost is available. The cost table stores information on the economic value of the damage, construction works, remediation, and mitigation measures. Typically, information on the cost caused by a trigger, event, or phenomenon is often generic (Munich Re, 2011); for example, it encompasses costs for damage remediation and risk mitigation. To consider this possibility, the damage and mitigation tables include a foreign key for linking them to the cost table, so that information on multiple damages and mitigation works can be related to a single, cumulated monetary value in the cost table. In addition to the monetary value, fields on the cost table allow separating economic costs that are estimated or officially allocated but not yet spent or spent. This distinction is important when determining the actual cost of a damaging event or trigger. In the Apulian database, the majority of the costs recorded are officially allocated, since there is available information regarding events for which a state of emergency has been declared; therefore, precise data have been given regarding the placement of the damaged element and related costs. This is because in Italy, damages caused by geo-hydrological hazards are fixed primarily using public resources, while often following specific legislation. In our experience, for more severe events, which in terms of civil protection have led to higher emergency levels, more and more detailed information is available.

The AReGeoDatHa database can be meaningfully compared with the AVI archive data (http://sici.irpi.cnr.it/). Comparative analysis of the frequencies of the two available catalogues reveals that the frequency values calculated for the 12-year 2008–2019 period are higher than those calculated for the 40-year 1959–1998 period of the AVI database. Since the present work and the AVI project share the same primary source of information (i.e., the newspaper “La Gazzetta del Mezzogiorno”), it cannot be excluded that the increased number of damaging events may be at least partly related to the increased frequency of extreme weather phenomena. In addition, in recent decades, the intensification of human activities and land-use changes have led to the spread of new urban areas and buildings, infrastructure, and industrial and rural activities. Meanwhile, the development and dissemination of the web and social media have led to increased availability of information, even for small events (Stoffel et al., 2014) such as single landslides interrupting local roads or isolated rainfall events that produce local flooding. These kinds of events were certainly under-reported in earlier catalogues, even though they were the most frequent and of particular interest regarding civil protection matters.

This work confirms that using different information sources, analyzed and classified with a quality index, can increase the amount of data and quality of extracted information. The designed criteria could easily be applied to new databases, even in different research sectors; the quantification of the quality of collected data is their added value.

The presented case study confirms the flexibility and efficiency of the used database structure, as it facilitates recording all the available information regarding geo-hydrological hazards in a single structure. It has the advantage of collecting in a single database information on different phenomena, related damages, costs, and mitigation works. It facilitates the addition of new data thanks to a simple user-friendly web interface, the analysis of recorded data, and the management of the temporal and spatial uncertainty of geo-hydrological events. The regional database could be useful for land use and civil protection planning, while it may elucidate areas more susceptible to geo-hydrological hazards during an event, thanks to the separate geo-location of phenomena and damage. It could be helpful to select sites where monitoring processes and/or prevention works need to be adopted, based on data concerning past events, incurred costs, and mitigation works. From this perspective, the main users of the database will be local administrators and civil protection agencies. The database structure is compliant with European directives; therefore, its potential application at the national or European scale is guaranteed. Since the EU Member States are required to maintain and update their own national catalogue of past flood events, (as required by article 4.2b of the Preliminary Flood Risk Assessment) AReGeoDatHa allows to download the data for updating the national catalogue on flood events.

Credit author statement

Carmela Vennari: Conceptualization, Methodology, Investigation, Software, Formal analysis, Data Curation, Validation, Writing - Original Draft, Writing - Review & Editing, Visualization. **Paola Salvati:** Conceptualization, Methodology, Formal analysis, Validation, Writing - Original Draft, Writing - Review & Editing, Visualization. **Cinzia Bianchi:** Data Curation, Validation, Visualization. **Domenico Casarano:** Investigation, Data Curation, Validation, Writing - Original Draft. **Mario Parise**: Validation, Writing - Original Draft. **Alessia Basso:** Data Curation. **Ivan Marchesini:** Supervision, Conceptualization, Methodology, Software, Formal analysis, Validation, Writing - Original Draft, Writing - Review & Editing.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Carmela Vennari reports financial support was provided by Regional Civil Protection of Apulian Region.

Data availability

Data download: https://doi.org/10.5281/zenodo.5898539.

Acknowledgements

The research is supported by the Civil Protection of the Apulia region, in the framework of the project 'Integrated assessment of geo-hydrological instability phenomena in the Apulia region, interpretative models and definition of rainfall thresholds for landslide triggering' funded by the P.O.R. Puglia 2014–2020, Asse V - Azione 5.1. (Project identification number: B82F16003840006).

Appendix A

Procedure used for the collection of the information

The database of geo-hydrological events developed for the Apulian region considers the phenomena that have occurred from 1992 to 2019. Four main categories of information were used: (i) newspapers articles, (ii) technical reports and scientific papers, (iii) institutional databases, and (iv) web sources.

For the collection of the information (section 3.3; Fig. 1C), we proceeded by date using a temporal approach starting from the oldest (i.e., from 1992) to the more recent (i.e., from 2019) available information sources, and from the large digital archive of the local newspaper "La Gazzetta del Mezzogiorno" (#1 in Fig. 6). Such an approach allows the trigger to be circumscribed in an occurrence time interval expressed in number of days. We scrutinized the online archive on which keyword searches were carried out and about 900 articles were selected with useful information on geo-hydrological phenomena as well as on damages, and restoration and mitigation works related to them.

Subsequently, the recognition of any type of processes (i.e., phenomena) and human-related entities is made by using additional local or regional newspapers (#2 in Fig. 6) available during the days of the trigger occurrence, thereby allowing the data amount and details to be enriched, and new phenomena or damage to be recognized. To validate the information from the newspapers articles, an inter-comparison activity allows us to obtain validated data from different information sources including blogs and institutional social media profiles (#13 in Fig. 6).

More detailed and accurate information result in greater numbers of elements allowing a complete re-enactment of geo-hydrological events occurring in the study area. Normally, the availability of multiple sources depends largely on the severity and extent of the damage occurring during an event, as well as the timespan covered by the catalogue.

There were multiple collected institutional sources. The regional civil protection agency provided technical documents with reliable information: technical and emergency reports (#7 in Fig. 6), lists of the funded interventions following states of emergency (#8 in Fig. 6), and a list of coastal cliff collapses (#9 in Fig. 6). These documents contain information on both phenomena and damage and restoration costs, even though the format of these sources is very heterogeneous, ranging from paper or PDF files to spreadsheets. We acquired also the intervention sheets of the provincial fire brigade command (#10 in Fig. 6), containing detailed useful information on their interventions following damages caused by geo-hydrological phenomena. To obtain information on the events that affected the road network, a list of road interruptions caused by geo-hydrological events was requested and acquired by the national roads authority (#11 in Fig. 6) and road safety information coordination center (#12 in Fig. 6). In addition, we analyzed all scientific publications (#3 in Fig. 6) and databases from earlier projects (#4, #6 in Fig. 6), along with studies and technical reports (#5 in Fig. 6) from the center of competence for civil protection (CC) CNR-IRPI, whenever they were involved in the states of emergency.

Appendix B

Table B 1
input tables, dictionary tables and views present in LanDeFeND.

INPUT	DICTIONARY	VIEW
Coordinates	Administrative Data	Inspire-Flood Observed Event
Trigger	Alert Zone	Inspire-Landslide Observed Event
Events	Bibliographic Source	Inspire-Sinkhole Observed Event
Criticality Alert	Bibliographic Topic	Inspire-Observed Event
Landslides	Cave Type	Inspire-Landslide Exposed Event
Floods	Civili Protection Authority	Inspire-Flood Exposed Event
Rainfall Event	Cost	Inspire- Observed Event
Sinkholes	Criticality Level	PFRA
Damage	Critically Type	PFRA-Damage
Mitigations	Event	PFRA-Events
Costs	Expose Element Category	PFRA-Phenomena
Bibliography	Flood	
Biblio Association	Flood Mechanism	
	Funding	
	Geographical Accuracy	
	Landslide	
	Level Total Damage	
	Object	
	Object Type	
	Predisposing Factor	
	Progress	
	Sinkhole Nature	
	Sinkhole Origin	
	Source of Flooding	
	Temporal Accuracy	
	Triggering Factor	
	Triggering Reliability	
	Type Category	
	Type of Image	
	Unit of Management	

Web sites

https://www.emdat.be/, accessed 15 September 2021.
https://www.eea.europa.eu/data-and-maps/data/copernicus-land-monitoring-service-eu-dem, accessed 21 December 2018.
https://thetrustproject.org/about/, accessed 17 September 2021.
http://sici.irpi.cnr.it/, accessed 20 October 2021.

References

Andres, N., Badoux, A., 2019. The Swiss flood and landslide damage database: normalisation and trends. J. Flood Risk Manag. 12, e12510 https://doi.org/10.1111/jfr3.12510.

Avand, M., Moradi, H.R., Ramazanzadeh Lasboyee, M., 2021. Spatial prediction of future flood risk: an Approach to the effects of climate change. Geosciences. https://doi.org/10.3390/geosciences11010025.

Blahut, J., Poretti, I., De Amicis, M., Sterlacchini, S., 2012. Database of geo-hydrological disasters for civil protection purposes. Nat. Hazards 60 (3), 1065–1083. https://doi.org/10.1007/s11069-011-9893-6.

Brunetti, M.T., Peruccacci, S., Rossi, M., Luciani, S., Valigi, D., Guzzetti, F., 2010. Rainfall thresholds for the possible occurrence of landslides in Italy. Nat. Hazards Earth Syst. Sci. 10, 447–458. https://doi.org/10.5194/nhess-10-447-2010.

Bruno, E., Calcaterra, D., Parise, M., 2008. Development and morphometry of sinkholes in coastal plains of Apulia, Southern Italy. Preliminary sinkhole susceptibility assessment. Eng. Geol. 99, 198–209. https://doi.org/10.1016/j.enggeo.2007.11.017.

Calcaterra, D., Parise, M., 2001. The contribution of historical information in the assessment of the landslide hazard. In: Glade, T., Albini, P., Frances, F. (Eds.), The Use of Historical Data in Natural Hazard Assessments. Advances in Natural and Technological Hazards Research, vol. 17. Springer, Dordrecht. https://doi.org/10.1007/978-94-017-3490-5_15.

Calvello, M., Pecoraro, G., 2018. FraneItalia: a catalog of recent Italian landslides. Geoenviron. Disasters 5 (1), 1–16. https://doi.org/10.1186/s40677-018-0105-5.

Casarano, D., Dragone, V., Polemio, M., 2019. Groundwater resources at salinisation risk: effects of climate and exploitation changes in the case of Apulian coastal aquifers (south-eastern Italy). Acque Sotterranee - Ital. J. Groundwater 53–57. https://doi.org/10.7343/as-2019-374, 2019AS28- 374.

Cotecchia, F., Vitone, C., Santaloia, F., Pedone, G., Bottiglieri, O., 2015. Slope instability processes in intensely fissured clays: case histories in the southern Apennines. Landslides 12, 877–893. https://doi.org/10.1007/s10346-014-0516-7.

Damm, B., Klose, M., 2015. The landslide database for Germany: closing the gap at national level. Geomorphology 249, 82–93. https://doi.org/10.1016/j.geomorph.2015.03.021.

De Groeve, T., Poljansek, K., Ehrlich, D., Corbane, C., 2014. European Commission - JRC Scientific and Policy Reports. Current Status and Best Practices for Disaster Loss Data Recording in EU Member States – A Comprehensive Overview of Current Practice in the EU Member States. https://doi.org/10.2788/18330, 2014JRC92290.

Del Prete, S., Iovine, G., Parise, M., Santo, A., 2010. Origin and distribution of different types of sinkholes in the plain areas of Southern Italy. Geodin. Acta 23 (1/3), 113–127. https://doi.org/10.3166/ga.23.113-127.

Delle Rose, M., Parise, M., 2002. Karst subsidence in south-central Apulia Italy. Int. J. Speleol. 31 (1/4), 181–199. https://doi.org/10.5038/1827-806X.31.1.11.

Diakakis, M., Mavroulis, S., Deligiannakis, G., 2012. Flood in Greece, a statistical and spatial approach. Nat. Hazards 62, 485–500. https://doi.org/10.1007/s11069-012-0090-z.

Diodato, N., Ljungqvist, F.C., Bellocchi, G., 2019. A millennium-long reconstruction of damaging hydrological events across Italy. Sci. Rep. 9 (1), 1–10. https://doi.org/10.1038/s41598-019-46207-7.

Diprizio, G., Andriani, G.F., Loiotine, L., Liso, I.S., Parise, M., 2021. An integrated approach to elaborate 3-D geological and geotechnical models: a case study from the Daunia Sub-Apennine (Apulia, southern Italy). Ital. J. Eng. Geol. Environ., sp. issue 1, 67–78. https://doi.org/10.4408/IJEGE.2021-01.S-06.

Doglioni, A., Simeone, V., 2019. Effects of climatic changes on groundwater availability in a semi-arid Mediterranean region. In: Shakoor, A., Cato, K. (Eds.), IAEG/AEG Annual Meeting Proceedings, San Francisco, California, vol. 4. Springer, Cham. https://doi.org/10.1007/978-3-319-93133-3_14, 2018.

Doglioni, C., Mongelli, F., Pieri, P., 1994. The Puglia uplift (SE Italy): an anomaly in the foreland of the Apenninic subduction due to buckling of a thick continental lithosphere. Tectonics 13 (5), 1309–1321. https://doi.org/10.1029/94TC01501.

Ellena, M., Ricciardi, G., Barbato, G., Buffa, A., Villani, V., Mercogliano, P., 2020. Past and future hydrogeological risk assessment under climate change conditions over urban settlements and infrastructure systems: the case of a sub-regional area of Piedmont, Italy. Nat. Hazards 102, 275–305. https://doi.org/10.1007/s11069-020-03925-w.

Festa, V., Fiore, A., Parise, M., Siniscalchi, A., 2012. Sinkhole evolution in the Apulian karst of southern Italy: a case study, with some considerations on sinkhole hazards. J. Cave Karst Stud. 74 (2), 137–147. https://doi.org/10.4311/2011JCKS0211.

Fidelibus, M.D., Gutierrez, F., Spilotro, G., 2011. Human-induced hydrogeological changes and sinkholes in the coastal gypsum karst of Lesina Marina area (Foggia Province, Italy). Eng. Geol. 118, 1–19. https://doi.org/10.1016/j.enggeo.2010.12.003.

Fiore, A., Parise, M., 2013. Cronologia degli eventi di sprofondamento in Puglia, con particolare riferimento alle interazioni con l'ambiente antropizzato. Mem. Descr.lla Carta Geol. Italia 93, 239–252.

Galeazzi, C., 2013. The typological tree of artificial cavities: a contribution by the Commission of the Italian Speleological Society. Opera Ipogea 1, 9–18.

Garcia-Urquia, E., Axelsson, K., 2014. The use of press data in the development of a database for rainfall-induced landslides in Tegucigalpa, Honduras, 1980–2005. Nat. Hazards 73 (2), 237–258. https://doi.org/10.1007/s11069-014-1043.

Gariano, S.L., Brunetti, M.T., Iovine, G., Melillo, M., Peruccaci, S., Terranova, O., Vennari, C., Guzzetti, F., 2015. Calibration and validation of rainfall thresholds for shallow landslide forecasting in Sicily, southern Italy. Geomorphology 228, 653–665. https://doi.org/10.1016/j.geomorph.2014.10.019.

Gariano, S.L., Supplizi, G.V., Ardizzone, F., Salvati, P., Bianchi, C., Morbidelli, R., Saltalippi, C., 2021. Long-term analysis of rainfall-induced landslides in Umbria, central Italy. Nat. Hazards 106 (3), 2207–2225. https://doi.org/10.1007/s11069-021-04539-6.

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, 1955–1971. https://doi.org/10.5194/nhess-19-1955-2019.

Glade, T., Albini, P., Francés, F. (Eds.), 2001. The Use of Historical Data in Natural Hazard Assessments, vol. 17. Springer Science & Business Media. https://doi.org/10.1007/978-94-017-3490-5.

Graff, K., Lissak, C., Thiery, Y., Maquaire, O., Costa, S., Laignel, B., 2019. Analysis and quantification of potential consequences in multirisk coastal context at different spatial scales (Normandy, France). Nat. Hazards 99 (2), 637–664. https://doi.org/10.1007/s11069-019-03763-5.

Gutierrez, F., Parise, M., De Waele, J., Jourde, H., 2014. A review on natural and human-induced geohazards and impacts in karst. Earth Sci. Rev. 138, 61–88. https://doi.org/10.1016/j.earscirev.2014.08.002.

Guzzetti, F., Tonelli, G., 2003. A database of historical information on landslides and floods in Italy. In: EGS-AGU-EUG Joint Assembly (abstract id. 3069). https://ui.adsabs.harvard.edu/abs/2003EAEJA.....3069G/abstract.

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. Manag. 18 (4), 623–633. https://doi.org/10.1007/BF02400865.

Guzzetti, F., Salvati, P., Stark, C.P., 2005. Evaluation of risk to the population posed by natural hazards in Italy. In: Hung, O., Fell, R., Couture, R., Eberhardt, E. (Eds.), Landslide Risk Management. A.A. Balkema Publishers, pp. 381–389.

Guzzetti, F., Mondini, A.C., Cardinali, M., Fiorucci, F., Santangelo, M., Chang, K.T., 2012. Landslide inventory maps: new tools for an old problem. Earth Sci. Rev. 112 (1–2), 42–66. https://doi.org/10.1016/j.earscirev.2012.02.001.

IPCC, Clarke, L., Dahe, Q., Dasgupta, P., 2014. Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment. Report of the Intergovernmental Panel on Climate.

Isacco, S., Claps, P., Grasso, S., Ferrari, E., Guercio, M.B., Musumeci, R.E., Scarcella, G.E., Versace, P., Laio, F., 2018. Floodbook: a social platform for flood hydrology. EPiC Ser. Eng. 3, 941–949. https://scholar.googleusercontent.com/scholar?q=cache:nHzKz8QeImcJ:scholar.google.com/+Floodbook:+a+social+platform+for+flood+hydrology&hl=it&as_sdt=0,5.

Lionello, P., Abrantes, F., Gacic, M., Planton, S., Trigo, R., Ulbrich, U., 2014. The climate of the Mediterranean region: research progress and climate change impacts. Reg. Environ. Change 14, 1679–1684. https://doi.org/10.1007/s10113-014-0666-0.

Margiotta, S., Negri, S., Parise, M., Valloni, R., 2012. Mapping the susceptibility to sinkholes in coastal areas, based on stratigraphy, geomorphology and geophysics. Nat. Hazards 62 (2), 657–676. https://doi.org/10.1007/s11069-012-0100-1.

Margiotta, S., Negri, S., Parise, M., Quarta, T.A.M., 2016. Karst geosites at risk of collapse: the sinkholes at Nociglia (Apulia, SE Italy). Environ. Earth Sci. 75 (1), 1–10. https://doi.org/10.1007/s12665-015-4848-y.

Margiotta, S., Marini, G., Fay, S., D'Onghia, F.M., Liso, I.S., Parise, M., Pinna, M., 2021. Hydro-stratigraphic conditions and human activity leading to development of a sinkhole cluster in a Mediterranean water ecosystem. Hydrology 8 (3), 111. https://doi.org/10.3390/hydrology8030111.

Martinotti, M.E., Pisano, L., Marchesini, I., Rossi, M., Peruccacci, S., Brunetti, M.T., Melillo, M., Amoruso, G., Loiacono, P., Vennari, C., Vessia, G., Trabace, M., Parise, M., Guzzetti, F., 2017. Landslides, floods and sinkholes in a karst environment: the 1–6 September 2014 Gargano event, southern Italy. Nat. Hazards Earth Syst. Sci. 17 (3), 467–480. https://doi.org/10.5194/nhess-17-467-2017.

Melillo, M., Brunetti, M.T., Peruccacci, S., Gariano, S.L., Roccati, A., Guzzetti, F., 2018. A tool for the automatic calculation of rainfall thresholds for landslide occurrence.

Environ. Model. Software 105, 230–243. https://doi.org/10.1016/j.envsoft.2018.03.024.

Mertz, M., Inthorn, J., Renz, G., Rothenberger, L.G., Salloch, S., Schildmann, J., Wöhlke, S., Schicktanz, S., 2014. Research across the disciplines: a road map for quality criteria in empirical ethics research. BMC Med. Ethics 15, 17. https://doi.org/10.1186/1472-6939-15-17.

Mirus, B.B., Jones, E.S., Baum, R.L., Godt, J.W., Slaughter, S., Crawford, M.M., Lancaster, J., Stanley, T., Kirschbaum, D.B., Burns, W.J., Schmitt, R.G., Lindsey, K. O., McCoy, K.M., 2020. Landslides across the USA: occurrence, susceptibility, and data limitations. Landslides 17, 2271–2285. https://doi.org/10.1007/s10346-020-01424-4.

Munich, Re, 2011. NatCat SERVICE Loss Database for Natural Catastrophes Worldwide.

Napolitano, E., Marchesini, I., Salvati, P., Donnini, M., Bianchi, C., Guzzetti, F., 2018. LAND-deFeND–An innovative database structure for landslides and floods and their consequences. J. Environ. Manag. 207, 203–218. https://doi.org/10.1016/j.jenvman.2017.11.022.

Negri, S., Margiotta, S., Quarta, T.A.M., Castiello, G., Fedi, M., Florio, G., 2015. Integrated analysis of geological and geophysical data for the detection of man-made underground caves in an area in Southern Italy. J. Cave Karst Stud. 77 (1), 52–62. https://doi.org/10.4311/2014ES0107.

Paliaga, G., Faccini, F., Luino, F., Turconi, L., 2019. A spatial multicriteria prioritizing approach for geo-hydrological risk mitigation planning in small and densely urbanized Mediterranean basins. Nat. Hazards Earth Syst. Sci. 19 (1), 53–69. https://doi.org/10.5194/nhess-19-53-2019.

Parise, M., 2010a. The impacts of quarrying in the Apulian karst. In: Carrasco, F., La Moreaux, J.W., Duran Valsero, J.J., Andreo, B. (Eds.), Advances in Research in Karst Media. Springer, pp. 441–447. https://doi.org/10.1007/978-3-642-12486-0_68.

Parise, M., 2010b. Hazards in karst. In: Bonacci, O. (Ed.), Proceedings International Interdisciplinary Scientific Conference "Sustainability of the Karst Environment. Dinaric Karst and Other Karst Regions", vol. 2. IHP-UNESCO, Series on Groundwater, Plitvice Lakes (Croatia), pp. 155–162, 23-26 September 2009. https://scholar.googleusercontent.com/scholar?q=cache:XdHrVdvzBr8J:scholar.google.com/+Hazards+in+karst&hl=it&as_sdt=0,5&scioq=The+Typological+Tree+of+Artificial+Cavities:+A+Contribution+by+the+Commission+of+the+Italian+Speleological+.

Parise, M., 2015. A procedure for evaluating the susceptibility to natural and anthropogenic sinkholes. Georisk 9 (4), 272–285. https://doi.org/10.1080/17499518.2015.1045002.

Parise, M., 2019. Sinkholes. In: White, W.B., Culver, D.C., Pipan, T. (Eds.), Encyclopedia of Caves, third ed. Academic Press, Elsevier. https://doi.org/10.1016/B978-0-12-814124-3.00110-2.

Parise, M., Vennari, C., 2013. A chronological catalogue of sinkholes in Italy: the first step toward a real evaluation of the sinkhole hazard. In: Land, L., Doctor, D.H., Stephenson, B. (Eds.), Proceedings of the 13th Multidisciplinary Conference on Sinkholes and the Engineering and Environmental Impacts of Karst. National Cave and Karst Research Institute, Carlsbad, NM, pp. 383–392. https://doi.org/10.5038/9780979542275.1149.

Parise, M., Vennari, C., 2017. Distribution and features of natural and anthropogenic sinkholes in Apulia. In: Renard, P., Bertrand, C. (Eds.), EuroKarst 2016, Neuchatel. Advances in the Hydrogeology of Karst and Carbonate Reservoirs. https://doi.org/10.1007/978-3-319-45465-8_3.

Parise, M., Federico, A., Palladino, G., 2012. Historical evolution of multi-source mudslides. In: Eberhardt, E., Froese, C., Turner, A., Lerouil, S. (Eds.), Landslides and Engineered Slopes. Protecting Society through Improved Understanding. Proceedings 11th Int. Symp. Landslides, pp. 401–407. Banff (Canada). https://scholar.googleusercontent.com/scholar?q=cache:UTCbKbwkMi4J:scholar.google.com/+Historical+evolution+of+multi-source+mudslides&hl=it&as_sdt=0,5&scioq=The+Typological+Tree+of+Artificial+Cavities:+A+Contribution+by+the+Commission+of+the+Italian+Speleological+Society.

Parise, M., Pisano, L., Vennari, C., 2018. Sinkhole clusters after heavy rainstorms. J. Cave Karst Stud. 80 (1), 28–38. https://doi.org/10.4311/2017ES0105.

Patacca, E., Scandone, P., 2007. Geology of the southern Apennines. Boll. Soc. Geol. Ital. Vol. Spec. 7, 75–119.

Patacca, E., Sartori, R., Scandone, P., 1990. Tyrrhenian basin and Apenninic arcs: kinematic relations since late Tortonian times. Mem. Soc. Geol. Ital. 45, 425–451.

Peruccacci, S., Brunetti, M.T., Luciani, S., Vennari, C., Guzzetti, F., 2012. Lithological and seasonal control on rainfall thresholds for the possible initiation of landslides in central Italy. Geomorphology 139, 79–90. https://doi.org/10.1016/j.geomorph.2011.10.005.

Pisano, L., Zumpano, V., Malek, Z., Micu, M., Rosskopf, C.M., Parise, M., 2017a. Multi-temporal landslide susceptibility maps and future scenarios for expected land cover changes (Southern Apennines, Italy). In: Mikos, M., Vilimek, V., Yin, Y., Sassa, K. (Eds.), Advancing Culture of Living with Landslides. WLF 2017. Springer, Cham. https://doi.org/10.1007/978-3-319-53483-1_45.

Pisano, L., Zumpano, V., Malek, Z., Rosskopf, C.M., Parise, M., 2017b. Variations in the susceptibility to landslides, as a consequence of land cover changes: a look to the past, and another towards the future. Sci. Total Environ. 601–602, 1147–1159. https://doi.org/10.1016/j.scitotenv.2017.05.231.

Polemio, M., Casarano, D., 2008. Climate change, drought and groundwater availability in southern Italy. In: Dragoni, W., Sukhija, B.S. (Eds.), Climate Change and Groundwater, vol. 288. Geological Society London Special Publications, pp. 39–51. https://doi.org/10.1144/SP288.4.

Refice, A., Spalluto, L., Bovenga, F., Fiore, A., Miccoli, M.N., Muzzicato, P., Nitti, D.O., Nutricato, R., Pasquariello, G., 2019. Integration of persistent scatterer interferometry and ground data for landslide monitoring: the Pianello landslide (Bovino, Southern Italy). Landslides 16, 447–468. https://doi.org/10.1007/s10346-018-01124-0.

Rohan, T.J., Wondolowski, N., Shelef, E., 2021. Landslide susceptibility analysis based on citizen reports. Earth Surf. Process. Landforms 46 (4), 791–803. https://doi.org/10.1002/esp.5064.

Rossi, M., Peruccacci, S., Brunetti, M.T., Marchesini, I., Luciani, S., Ardizzone, F., Balducci, V., Bianchi, C., Cardinali, M., Fiorucci, F., Mondini, A.C., Reichenbach, P., Salvati, P., Santangelo, M., Bartolini, D., Gariano, S.L., Palladino, M., Vessia, G., Viero, A., Antronico, L., Borselli, L., Deganutti, A.M., Iovine, G., Luino, F., Parise, M., Polemio, M., Guzzetti, F., Tonelli, G., 2012. SANF: national warning system for rainfall-induced landslides in Italy. In: Eberhardt, E., Froese, C., Turner, A.K., Leroueil, S. (Eds.), Landslides and Engineered Slopes. Protecting Society through Improved Understanding. Proceedings 11th Int. Symp. Landslides, vol. 2, pp. 1895–1899. https://doi.org/10.13140/2.1.4857.9527. Banff (Canada), 3-8 June 2012.

Saint-Martin, C., Javelle, P., Vinet, F., 2018. DamaGIS: a multisource geodatabase for collection of flood-related damage data. Earth Syst. Sci. Data 10, 1019–1029. https://doi.org/10.5194/essd-10-1019-2018.

Salvati, P., Balducci, V., Bianchi, C., Guzzetti, F., Tonelli, G., 2009. A WebGIS for the dissemination of information on historical landslides and floods in Umbria, Italy. GeoInformatica 13, 305–322. https://doi.org/10.1007/s10707-008-0072-1.

Salvati, P., Petrucci, O., Rossi, M., Bianchi, C., Pasqua, A.A., Guzzetti, F., 2018. Gender, age and circumstances analysis of flood and landslide fatalities in Italy. Sci. Total Environ. 610–611, 867–879. https://doi.org/10.1016/j.scitotenv.2017.08.064.

Spalluto, L., Fiore, A., Miccoli, M.N., Parise, M., 2021. Activity maps of multi-source mudslides from the Daunia Apennines (Apulia, southern Italy). Nat. Hazards 106, 277–301. https://doi.org/10.1007/s11069-020-04461-3.

Stoffel, M., Tiranti, D., Huggel, C., 2014. Climate change impacts on mass movements – case studies from the European Alp. Sci. Total Environ. 493, 1255–1266. https://doi.org/10.1016/j.scitotenv.2014.02.1022.102.

Swiss, Re, 2017. Sigma Natural Catastrophes and Man-Made Disasters.

Tanoue, M., Hirabayashi, Y., Ikeuchi, H., 2016. Global-scale river flood vulnerability in the last 50 years. Sci. Rep. 6, 1–9. https://doi.org/10.1038/srep36021.

Taylor, F.E., Malamud, B.D., Katy Freeborough, K., Demeritt, D., 2015. Enriching Great Britain's national landslide database by searching newspaper archives. Geomorphology 249, 52–68. https://doi.org/10.1016/j.geomorph.2015.05.019, 2015.

UNISDR, 2015. Hyogo Framework for Action 2005-2015, Building the Resilience of Nations and Communities to Disasters.

Valenzuela, P., Domínguez-Cuesta, M.J., García, M.A.M., Jiménez-Sánchez, M., 2017. A spatio-temporal landslide inventory for the NW of Spain: BAPA database. Geomorphology 293, 11–23. https://doi.org/10.1016/j.geomorph.2017.05.010.

Van Den Eeckhaut, M., Hervás, J., 2012. State of the art of national landslide databases in Europe and their potential for assessing landslide susceptibility, hazard and risk. Geomorphology 139, 545–558. https://doi.org/10.1016/j.geomorph.2011.12.006.

van Westen, C.J., Greiving, S., 2017. Multi-hazard risk assessment and decision making. In: Environmental Hazards Methodologies for Risk Assessment and Management. IWA Publishing, London, pp. 31–39. https://doi.org/10.2166/9781780407135_0031.

van Westen, C.J., Van Asch, T.W., Soeters, R., 2006. Landslide hazard and risk zonation—why is it still so difficult? Bull. Eng. Geol. Environ. 65, 167–184. https://doi.org/10.1007/s10064-005-0023-0.

van Westen, C.J., Castellanos, E., Kuriakose, S.L., 2008. Spatial data for landslide susceptibility, hazard, and vulnerability assessment: an overview. Eng. Geol. 102 (3–4), 112–131. https://doi.org/10.1016/j.enggeo.2008.03.010.

van Westen, C.J., Alkema, D., Damen, M.C.J., Kerle, N., Kingma, N.C., 2011. Multi-hazard Risk Assessment. Distance Education Course; Guide Book. United Nation University – ITC School on Disaster Geoinformation Management (UNU-ITC DGIM).

Vennari, C., Vessia, G., Parise, M., Rossi, M., Loiacono, P., Amoruso, G., Trabace, M., Giandonato, P., 2013. Slope movements in Daunia (Apulia): collecting historical events for the definition of rainfall thresholds. Rend. Online Soc. Geol. Ital. XX (2013), 322–324. ISSN 2035-8008.

Wasowski, J., Pisano, L., 2020. Long-term InSAR, borehole inclinometer, and rainfall records provide insight into the mechanism and activity patterns of an extremely slow urbanized landslide. Landslides 17, 445–457. https://doi.org/10.1007/s10346-019-01276-7.

Wasowski, J., Lamanna, C., Casarano, D., 2010. Influence of land-use change and precipitation patterns on landslide activity in the Daunia Apennines, Italy. Q. J. Eng. Geol. Hydrogeol. 43 (4), 387–401. https://doi.org/10.1144/1470-9236/08-101.

Zêzere, J.L., Pereira, S., Tavares, A.O., Bateira, C., Trigo, R.M., Quaresma, I., Santos, P.P., 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.

Zumpano, V., Ardizzone, F., Bucci, F., Cardinali, M., Fiorucci, F., Parise, M., Pisano, L., Reichenbach, P., Santaloia, F., Santangelo, M., Wasowski, J., Lollino, P., 2020. The relation of spatio-temporal distribution of landslides to urban development (a case study from the Apulia region, Southern Italy). J. Maps 17 (4), 133–140. https://doi.org/10.1080/17445647.2020.1746417.