Journal of Environmental Management 207 (2018) 203–218

Contents lists available at ScienceDirect

Journal of Environmental Management

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

Research article

LAND-deFeND – An innovative database structure for landslides and floods and their consequences

Elisabetta Napolitano, Ivan Marchesini*, Paola Salvati, Marco Donnini, Cinzia Bianchi, Fausto Guzzetti

CNR IRPI, Via Madonna Alta 126, I-06128, Perugia, Italy

ARTICLE INFO

Article history:
Received 25 July 2017
Received in revised form
20 October 2017
Accepted 7 November 2017

Keywords:
Database
Geo-spatial data
Landslide
Flood
Damage
Cost

ABSTRACT

Information on historical landslides and floods – collectively called "geo-hydrological hazards – is key to understand the complex dynamics of the events, to estimate the temporal and spatial frequency of damaging events, and to quantify their impact. A number of databases on geo-hydrological hazards and their consequences have been developed worldwide at different geographical and temporal scales. Of the few available database structures that can handle information on both landslides and floods some are outdated and others were not designed to store, organize, and manage information on single phenomena or on the type and monetary value of the damages and the remediation actions. Here, we present the LANDslides and Floods National Database (LAND-deFeND), a new database structure able to store, organize, and manage in a single digital structure spatial information collected from various sources with different accuracy. In designing LAND-deFeND, we defined four groups of entities, namely: nature-related, human-related, geospatial-related, and information-source-related entities that collectively can describe fully the geo-hydrological hazards and their consequences. In LAND-deFeND, the main entities are the nature-related entities, encompassing: (i) the "phenomenon", a single landslide or local inundation, (ii) the "event", which represent the ensemble of the inundations and/or landslides occurred in a conventional geographical area in a limited period, and (iii) the "trigger", which is the meteo-climatic or seismic cause (trigger) of the geo-hydrological hazards. LAND-deFeND maintains the relations between the nature-related entities and the human-related entities even where the information is missing partially. The physical model of the LAND-deFeND contains 32 tables, including nine input tables, 21 dictionary tables, and two association tables, and ten views, including specific views that make the database structure compliant with the EC INSPIRE and the Floods Directives. The LAND-deFeND database structure is open, and freely available from http://geomorphology.irpi.cnr.it/tools.

1. Introduction

Information on historical landslides and floods – hereafter collectively referred to as "geo-hydrological hazards" – is important to understand the complexities and dynamics of past events, and proves useful to construct and validate landslide and flood prediction models and to design appropriate mitigation measures. Databases and digital catalogues on geo-hydrological hazards store, organize, and manage information on the physical characteristics, the geographical location, and the temporal occurrence of past landslide and flood events (Hervás, 2013). A number of databases and digital catalogues with information on natural hazards – including geo-hydrological hazards – were compiled and used for research, insurance and economic purposes (e.g., Guzzetti and Tonelli, 2004; Munich Re, 2011; Menoni et al., 2016; Swiss Re, 2017). Modern databases on natural hazards and their consequences exploit geographical information system (GIS) technology to locate geographically the historical events, and to store geographic information on the events. The information stored in the databases is then made available through dedicated Web-GIS, or using Web Services and Spatial Data Infrastructures (SDI).

Only a few databases exist at the global scale (Table 1), including the U.S. Geological Survey Advanced National Seismic System (ANSS) Composite Catalog of earthquakes (https://earthquake.usgs.gov/data/comcat/), the Global Active Archive of Large Flood Events, compiled by the Dartmouth Flood Observatory of the University of

* Corresponding author.
E-mail address: ivan.marchesini@irpi.cnr.it (I. Marchesini).

https://doi.org/10.1016/j.jenvman.2017.11.022

Table 1
Catalogues available at different scales and for different natural hazards. CF, capable faults. EQ, earthquakes. FL, floods. LS, landslides. MS, mass disasters. NC, natural catastrophes. NH, natural hazards. References: [1] Munich Re, 2011; [2] Swiss Re, 2017; [3] Boschi et al., 1997; [4] Guidoboni et al., 2007; [5] Locati et al., 2014; [6] Llasat et al., 2013; [7] Michetti et al., 2000; [8] Basili et al., 2008; [9] Rovida et al., 2016; [10] Martino et al., 2014; [11] APAT & Trigila, 2007; [12] Guzzetti et al., 1994; [13] Guzzetti and Tonelli, 2004; [14] Petrucci and Versace, 2004.

	Hazard	Name	Coverage	Institution	Web page	Ref	ID
Global	EQ	COMprehensive earthquake CATalog [ComCat]	since 1898	USGS - ANSS	earthquake.usgs.gov/data/comcat/	–	1
	FL	Global active archive of large flood events	since 1985	University of Colorado - DFO	floodobservatory.colorado.edu/Archives/index.html	–	2
	MS	Emergency events Database [EM-Dat]	since 1900	CRED	www.emdat.be/	–	3
	NC	NatCat SERVICE	since 1980	Munich Re	natcatservice.munichre.com/	[1]	4
	NC	Sigma data	since 1970	Swiss Re	www.sigma-explorer.com/	[2]	5
Europe	EQ	Catalogue of strong earthquakes in Italy and Mediterranean area [CFTI4Med]	461 BC-1997 760 BC-1500	INGV & SGA	storing.ingv.it/cfti4med/	[3, 4]	6
	EQ	European Archive of Historical EArthquake Data [AHEAD]	1000–1899	INGV & AHEAD partners	www.emidius.eu/AHEAD/main/	[5]	7
	FO	FLOODHYMEX	1981–2010	HyMeX	mistrals.sedoo.fr/HyMeX/	[6]	8
Italy	CF	ITalyHAzard from CApable faults [ITHACA]	–	ISPRA	sgi.isprambiente.it/GMV2/index.html	[7]	9
	EQ	Database of Individual Seismogenic Sources [DISS]	–	INGV	diss.rm.ingv.it/diss/index.php/54-database-access	[8]	10
	EQ	Parametric catalogue of Italian earthquakes [CPTI15]	1000–2015	INGV	emidius.mi.ingv.it/CPTI15-DBMI15/index_en.htm	[9]	11
	EQ	Italian catalogue of earthquake-induced ground failures [CEDIT]	1000–2016	Università La Sapienza, CERI	www.ceri.uniroma1.it/index_cedit.html	[10]	12
	LS	Inventario dei Fenomeni Franosi in Italia [IFFI]	1918–2014	ISPRA	www.progettoiffi.isprambiente.it/cartanetiffi/cartografia.asp	[11]	13
	LS-FL	Aree Vulnerate Italiane [AVI]	1918–2001	CNR GNDCI	sici.irpi.cnr.it/avi.htm	[12]	14
	LS-FL	Sistema Informativo sulle Catastrofi Idrogeologiche [SICI]	18th-20thC.	CNR GNDCI	sici.irpi.cnr.it/	[13]	15
Region in Italy	LS-FL	Catasto dissesti Regionale	20th-21st C.	Regione Valle d'Aosta	http://catastodissesti.partout.it/#	–	16
	LS-FL	Aree storicamente inondate e fenomeni di dissesto idrogeologico [ASICal]	119–2004	Università della Calabria, Camilab	www.camilab.unical.it/web/camilab/prodotti-products	[14]	17
	LS-FL	Geologia e dissesto	1918–2005	ARPA Piemonte	www.arpa.piemonte.gov.it/approfondimenti/temi-ambientali/geologia-e-dissesto	–	18
	LS	Cartografia del dissesto della Regione Emilia-Romagna	Middle age - 2013	Regione Emilia-Romagna	applicazioni.regione.emilia-romagna.it/cartografia_sgss/user/viewer.jsp?service=dissesto	–	19
	NH	Archivio storico online degli eventi calamitosi della Provincia Autonoma di Trento [ARCA]	339–2005	Provincia Autonoma di Trento	www.protezionecivile.tn.it/territorio/Banchedati/	–	20

Colorado (http://floodobservatory.colorado.edu/), the NatCat SERVICE (Munich Re, 2011), the EM-Dat database (http://www.emdat.be), and the Sigma database (Swiss Re, 2017). The latter three databases contain loss data related to natural hazards, and were designed for insurance and reinsurance purposes (Kron et al., 2012; Wirtz et al., 2014). The global databases contain information primarily on highly destructive events that typically have impacted large areas. For this reason, the global databases are known to be incomplete, and they lack systematic information on low to medium intensity events, and on local events. The incompleteness of the global databases hampers their use for quantitative risk assessment studies (Van Den Eeckhaut and Hervàs, 2012).

For Italy, multiple databases and catalogues on natural hazards are available, including catalogues of earthquakes (e.g., Guidoboni et al., 2007), floods (e.g., European Commission, Joint Research Centre, 2014) and landslides (e.g., APAT & Trigila, 2007), with the scale and the geographical coverage of the catalogues in the range from regional to national (Table 1). Analysis of the existing databases and catalogues on geo-hydrological hazards in Italy reveals some important limitations. A common problem is that the existing databases do not separate the geographical location of the natural events (i.e., a landslide, a flood) from the location of their consequences (e.g., a landslide or flood fatality, a damaged or inundated building, a failed bridge) and of the mitigation measures, if any. This hampers the possibility to execute reliable *ex-post* damage and recovery analyses (Donnini et al., 2017). Another problem is that, having being designed well before the publication of hazard-related European Commission (EC) Directives, the databases do not consider or adopt the implementation rules required by relevant EC Directives concerning natural hazard catalogues and digital archives. Lastly, the inherent heterogeneity of the different Italian national and regional databases (Table 1) prevent their effective merging in a single database structure (Guzzetti and Tonelli, 2004). This limits the full exploitation and widespread use of the information in the databases and catalogues.

In an attempt to overcome these limitations, we designed LAND-deFeND – an acronym for LANDslides and Floods National Database – a new database structure capable of storing and organizing non-homogeneous information on historical and recent landslide and flood events. LAND-deFeND complies with the EC Floods Directive on flood risk assessment and management (2007/60/EC) that mandates that the description of relevant past flood events has to be recorded using a standardized structure (European Commission – DG Environment, 2013), and with the EC INfrastructure for Spatial InfoRmation in Europe (INSPIRE) Directive that defines standards to ensure the compatibility of different spatial data infrastructures, fostering data availability, quality, and accessibility in Europe (2007/2/EC).

In this context, the scope of LAND-deFeND is to manage in a single standardized structure spatial information on landslide and flood events collected at different geographical scales (from the national to the local scale), and by different institutions and organizations (e.g., national and local administrations, research centres).

In the paper, we first describe the conceptual, logical, and physical models laying at the base of the design and the implementation of the new LAND-deFeND database structure (Section 2). Next, we discuss the relationship between LAND-deFeND and the two relevant cited EC directives concerning geospatial information and flood hazard and risk (Section 3). Next, we compare LAND-deFeND to other existing database structures for geo-hydrological hazards (Section 4), and demonstrate the use of LAND-deFeND showing two case studies (Section 5). This is followed by a discussion of the relevant features of the new LAND-deFeND database structure (Section 6). We then conclude summarizing the main findings of this research effort (Section 7).

2. LAND-deFeND design and implementation

To design the LAND-deFeND database structure, we adopted a data modeling approach that included the design of a conceptual and logical data model, and its implementation in a physical data model (Codd, 1970). The conceptual data model defines the main entities and their basic properties, and it specifies the requirements for the database structure. The logical data model defines the relationships among the entities, reduces redundancy, and improves data integrity (i.e., the normalization rules, Codd, 1972; Date, 1999). Lastly, the physical data model consists in the actual implementation of the database structure, which is obtained exploiting a Relational DataBase Management System (RDBMS) software that defines data types, database functions, access constraints, and converts entities into tables, relationships into primary and foreign keys, and attributes into columns.

2.1. Conceptual model

The LAND-deFeND conceptual model identifies four groups of entities that collectively describe all the relevant characteristics of the geo-hydrological events and their socio-economic consequences and environmental impacts. We named the four groups of entities based on their main characteristics as Nature-related Entities (NE), Human-related Entities (HE), GeoSpatial-related Entities (GSE), and Information Source-related Entities (ISE).

The core of the database structure consists of the Nature-related Entities (NE), which comprise:

- The "phenomenon" i.e., the single geo-hydrological hazard. This includes a single landslide or a local flooding.
- The "event" i.e., an ensemble of floods and/or landslides occurred in a given geographical area (e.g., a catchment, a municipality, a region, the territory managed by a river basin authority) in a period, in the range from hours to weeks. We note that the LAND-deFeND "event" entity differs from the "event" defined by the EC Floods Directive (2007/60/EC; European Commission - DG Environment, 2013), which uses the "event" to describe both a generic flood event at the scale of the hydrological basin (e.g., a flood of the Po River), and a small inundation due to a levee break at the scale of a small river segment.
- The "trigger" i.e., the meteorological or seismic trigger that causes (i.e., triggers) the considered geo-hydrological hazards (i.e., the single or multiple landslides or floods).

Of the remaining entities, the Human-related Entities (HE) represent the public or private properties, goods and services that were damaged by the geo-hydrological hazards, and the related restoration and risk mitigation costs. The Geospatial-related Entities (GSE) represent the geographical location of the phenomena, of the damage, and of the remedial and mitigation works. Finally, the Information Source-related Entities (ISE) represent the sources of the information stored in the database, including bibliographic, chronicle, media, Internet, and other sources.

Fig. 1 illustrates the relations among the three NE ("phenomenon", "event", "trigger"), their possible geographical representation – as points, lines, or polygons, depending on the geometry and the accuracy of the information – and the relations between the NE and the HE. The latter, shown by red triangles, represent the consequences resulting from the interference of a natural event with human activities and interests. In the upper part of Fig. 1 we describe the temporal relations among the three NE main entities i.e., the “phenomenon”, “event”, and “trigger”, and the HE. In the lower part of Fig. 1, we exemplify the geographical relations between the two groups of entities. Three main relationships arise from the interaction of the NE main entities and the HE, namely:

- The “trigger” – “event”; where the “trigger” (blue in Fig. 1) is the entity that has the largest temporal (from one to several days) and geographical (from local to regional in scale) ranges, and comprises single or multiple “events” (green in Fig. 1). As an example, a meteorological front driven by a low-pressure (the “trigger”) moving across different geographical regions can generate inundations and landslides at different sites (“events”, each including one or more phenomena), in a period ranging from a few to several days.
- The “event” – “phenomenon”; where the “event” includes single or multiple phenomena (floods, landslides). The temporal extent of an “event” encompasses the date of occurrence of all the single related phenomena, and may not correspond or be enclosed in the period of the corresponding “trigger”. This is because a single phenomenon (e.g., a landslide) can occur (or can continue) after the end of the original triggering factor. This is the case of an “event” that includes a deep-seated landslide that occurs hours to days after the end of the triggering rainfall (Petley and Allison, 1997). In Fig. 1, the case is represented by event E_3 of trigger T_4.
- The “phenomenon” – Human-related Entities (HE); where a “phenomenon” (a single landslide or inundation) may be linked to one or more HE that was (were) affected and damaged by the phenomenon. This is the case of a small urban inundation causing damage at multiple sites. The “phenomenon” and the HE have attributes describing their geographical coordinates. As an example, a polygon describing the location and the shape of a landslide (a phenomenon) encompasses a point describing the location of a single building (a HE) damaged by the landslide.

The three relations discussed above do not encompass all the possible relations that may exist among the various considered entities. To cope with this problem, we designed the LAND-deFeND structure to be flexible, and capable of exploiting relations between the NE and the HE even where the information is partially missing. This is exemplified in Fig. 1 where the relations between (i) a trigger (T_1) and a phenomenon (in this case the event is unknown), (ii) the HE and the event (E_2) of the trigger (T_2) (here the phenomenon is unknown), and (iii) the trigger (T_3) and the HE (here in one case the event and the phenomena are unknown). We consider relevant the ability of LAND-deFeND to link entities even when the information is missing partially. This allows to store, organize, and manage information even where it is partial, contributing to reduce the incompleteness typical of catalogues and databases on natural hazards and their related damage (Guzzetti and Tonelli, 2004; Van Den Eeckhaut and Hervàs, 2012).

Different “triggers”, “events” or “phenomena” occurring in different periods in the same area may affect regions that were previously affected by past “triggers”, “events” or “phenomena”. The condition is shown by the grey areas and the white symbols in maps B, C and D in Fig. 1. In the LAND-deFeND conceptual model, the location of the “phenomenon” and the HE is crucial, as it allows to analyse the recurrence of the geo-hydrological events (i.e., their return period, or average time frequency), and of the related damage. We note that this is a requirement of the EC Floods Directive (2007/60/EC).

2.2. Logical model

The logical model is the design level at which the entities become tables, the attributes become table fields, and the relationships among the tables are established through primary and foreign keys (Codd, 1972; Date, 1999). Fig. 2 illustrates the logical model for the LAND-deFeND database structure. To avoid misinterpretations, we use double inverted commas when referring to the conceptual entities (e.g., “trigger”, “event”, “damage”) and first capital letters when referring to the corresponding tables (e.g., Trigger, Event, Damage). At this level of design of the database structure, we describe the main fields of the four groups of tables. The first two tables are derived from the NE and the HE, and are named Nature-related Tables (NT) and Human-related Tables (HT). Next, the GeoSpatial-related Tables (GST) and the Information Source-related Tables (IST) were introduced to consider the geospatial and the information source attributes of the entities. In the following, we describe the most relevant table fields, and the relations among the four mentioned table groups.

The “phenomenon” entity is transformed and split in the two tables, Landslide and Flood, which are used to describe slope and fluvial processes, respectively. The Landslide and Floods tables contain fields to store temporal and geo-spatial information, including (i) the date/time of occurrence and the duration of the event, and the corresponding uncertainties, (ii) the geographical location and shape of each landslide or flood (position and geometry). The geometrical fields store the geo-spatial information as a point, line or polygon, depending on the accuracy and/or the scale of the information. In the Event table the main fields store (i) the temporal information, namely the starting date and the duration of the event, in days, and (ii) the primary (main) type of the geo-hydrological hazard (e.g., fluvial inundation, slope failure, both). The Trigger table stores information on the type of trigger i.e., meteorological or seismic, and the fields include the starting date and the duration of the trigger (e.g., an intense or prolonged rainfall period, an earthquake), and free-text fields for ancillary information on the trigger (e.g., a brief summary of the trigger).

The Human-related Tables (HT) include fields describing the damaged public or private properties, goods and services, the related economic costs, the human impact (deaths, missing persons, injured people) and the mitigation actions designed to protect a given area, building, structure or infrastructure. In this context, the Damage table stores information on the effects of a specific geo-hydrological hazard on buildings, structures and infrastructures, and on the population, and can be linked to one or more Nature-related Tables, depending on the accuracy of the information (Fig. 1). As for the single phenomenon (a landslide, a flood), the damage information is mapped geographically using points, lines or polygons, depending on the scale and/or on the accuracy of the information.

Information on construction works and remedial or mitigation actions taken to protect an area from a geo-hydrological hazard (e.g., a new retaining wall, a drainage system, a new levee) are stored in the Mitigation table, which has the same relations to the Cost table as the Damage table (Fig. 2). A single work, remediation or mitigation measure is mapped geographically as a point, line or a polygon, depending on the scale and the accuracy of the information.

The Cost table stores information on the economic value of the damage, of the construction works, and of the remediation and the mitigation measures. In addition to the monetary value, fields on the Cost table allow separating economic costs that are (i) estimated, (ii) officially allocated but not spent yet, or (iii) spent. The distinction is important when attempting to determine the actual cost of a damaging event or trigger (Donnini et al., 2017). In real cases, the information on the cost caused by a trigger, event or phenomenon is often generic e.g., it encompasses costs for damage remediation and for risk mitigation. To consider this possibility, the

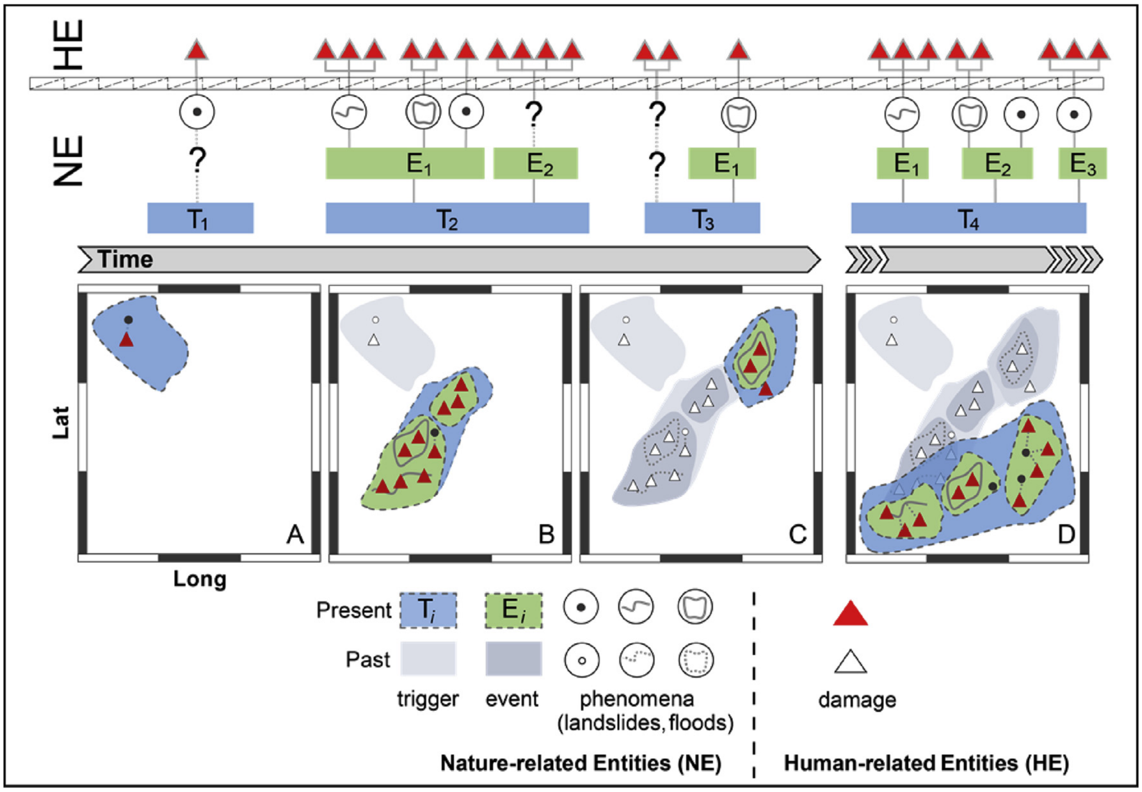

Fig. 1. Conceptual model schema adopted by the proposed LAND-DeFeND database structure. The upper and the lower parts of the Figure exemplify the temporal (above the time line) and the spatial-temporal (below the time line) relations between the Nature-related Entities (NE) and the Human-related Entities (HE). Below the time line, NE and HE are represented with coloured or grey scale symbols depending on the time of occurrence. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

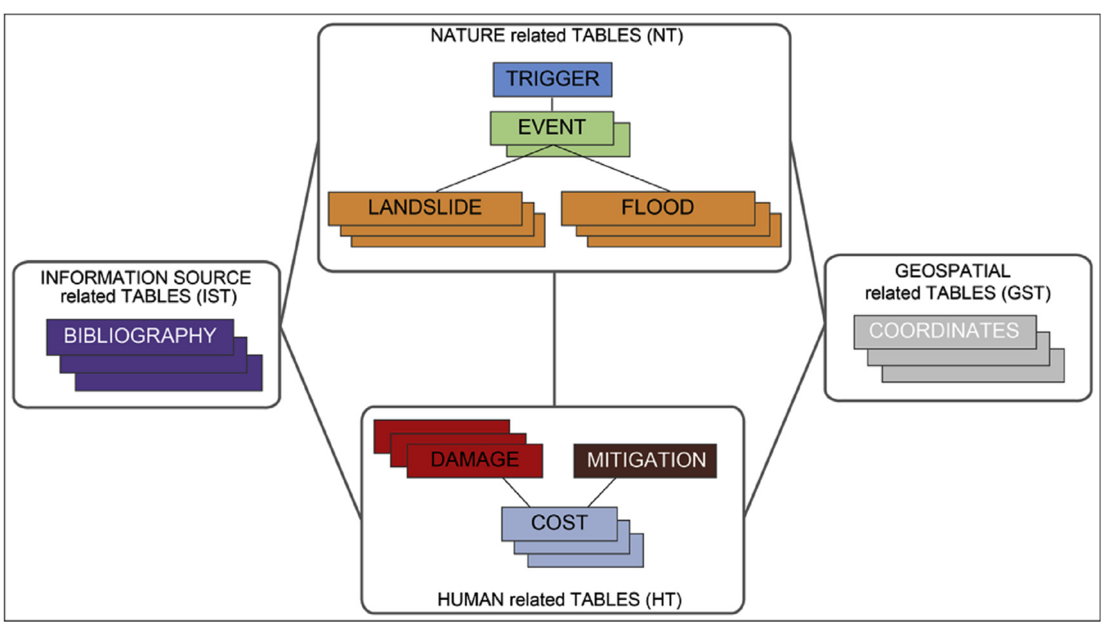

Fig. 2. Graphical representation of the logical model of the LAND-DeFeND database structure. The main relations between groups of tables are shown. See text for explanation.

Damage and Mitigation tables include a foreign key to link them to the Cost table, so that information on multiple damage and mitigation works can be related to a single, cumulated monetary value in the Cost table.

The GeoSpatial-related Tables (GST) include tables used to store the geographical location, the geometry (shape), and the positional accuracy of the single geo-hydrological hazards, damage, and mitigation works and actions. The GST are related to both the NT and the HT. We note that the location of an area that was affected by a given geo-hazard, where damage occurred or where mitigation works were executed, is stored using a single geographical entity (a point, a line, a polygon). This is also the case of future geo-hydrological phenomena in the same area, for which it will not be necessary − unless needed − to define a new geographical entity. The approach reduces unnecessary redundancy, contributes to enforcing referential integrity, and avoids problems related to the representation of the same geographical feature with multiple geometric entities.

Lastly, the Information Source-related Tables (IST) includes tables used to classify the different sources for the information, considering all the possible types of information sources.

2.3. Physical model

For the physical implementation of the LAND-deFeND database structure we adopted PostgreSQL, version 9.1, an open source RDBMS, with the PostGIS extension, version 2.1. Tables, fields, and relationships (designed in the logical model (Section 2.2) and described in detail in Appendix A) were translated into PostgreSQL (i) physical tables, (ii) fields, and (iii) primary and foreign keys (shown with the *id* and *fk* suffixes, respectively). Field data types were defined, and checks and constraints were adopted to guarantee the necessary semantic integrity.

Physical modeling required the creation of 32 tables (Fig. 3) and ten views (Fig. 5). The nine tables in Fig. 3 marked by black outlines (i.e., Trigger, Event, Landslide, Flood, Damage, Mitigation, Cost, Coordinates, Bibliography) are input tables. The other tables include 21 dictionary tables (grey background) used to define the fields of the tables and to protect the reference integrity through primary−foreign key relationships, and two association tables (grey outline) used to define the many−to−many relationships.

Inspection of Fig. 3 reveals that the core of the LAND-deFeND database structure is the TEP table that maintains a unique link among the three main Nature-related Tables (NT) i.e., the Trigger, Event and Phenomenon tables. We used this association table to maintain the hierarchical relationship shown in Fig. 1, linking "phenomenon" to "event" and/or to "trigger" − a mandatory condition to properly record the information. The NT table is also used when a damage, cost, or mitigation work has to be assigned to a phenomenon (i.e., a single landslide, a single flood), to an event or trigger. The TEP table was created to automatically fill the records related to the unique combination of "trigger" (T), "trigger" − "event" (T−E), "trigger" − "event" − "phenomenon" (T−E−P), each identified by foreign keys in three different records (Fig. 3). The information on two (or more) geo-hydrological hazards (landslides and/or floods) related to the same event are stored in the TEP table with two (or more) single records, each representing the combination of the identifiers of the "trigger", of the "event", and of the "phenomenon" (i.e. the landslide or flood). As an example, in Fig. 1, the trigger (T_4) is linked to a single event (E_2), which is linked to two phenomena (P_1 and P_2). In the database, the two corresponding records in the TEP table are: (i) T_4−E_2−P_1 and (ii) T_4−E_2−P_2.

To facilitate the data entry in the LAND-deFeND database structure, we developed (i) a specific procedure − described in Appendix B − that exploits QGIS® software (QGIS Development Team, 2017) as a client (Fig. 4), and (ii) a dedicated data entry web interface. Development of the two data-entry tools was facilitated by the fact that LAND-deFeND is independent from the applications used for entering and viewing the data.

We note here that the LAND-deFeND database structure is not designed to store the geometry (i.e., the geographical extent) of the individual triggers or events. This was decided because the geometry of the triggers or of the events is very often uncertain, and difficult − or even impossible − to decide accurately (Guzzetti et al., 2005). However, since information on the location and the extent of a trigger or an event is important, we designed two dedicated database views (trigger_convexhull and event_convexhull, see Fig. 5, Appendix A) to provide a representation of the geographical extent of the triggers and of the events. In the two views, all the geometries (point, line, polygon) associated to a single trigger or event are grouped, and encompassed by a minimum convex polygon, which is computed if the trigger or the event includes three or more geometries (i.e., point, line, polygon). As shown in Fig. 5, and described in Section 3, the LAND-deFeND database structure includes eight views to show (and export) data in the format mandated by the EC INSPIRE (2007/2/EC) and Floods (2007/60/EC) Directives.

3. LAND-deFeND and EC directives

LAND-deFeND complies to the EC INfrastructure for Spatial InfoRmation in Europe (INSPIRE) Directive for spatial data compatibility (2007/2/EC), and to the EC Floods Directive for flood risk assessment and management (2007/60/EC). The INSPIRE Directive (2007/2/EC) defines standards to ensure the compatibility of different spatial data infrastructures, with the aim of improving the availability, quality, and accessibility of data throughout Europe (INSPIRE Thematic Working Group Natural Risk Zones, 2013). Article 4.2b of the Floods Directive requires the description of past significant flood events using a common structure to store the data (2007/60/EC), and prescribes that the individual EU Member States maintain and update their own national catalogue of past flood events.

To be compliant to the EC Floods Directive reporting schemas (European Commission - DG Environment, 2013), we included in the tables of the LAND-deFeND database structure mandatory and optional fields defined by article 4.2b of the Preliminary Flood Risk Assessment. In dedicated dictionary tables, we included all the values required to compile the mandatory and the optional fields of the EC Floods Directive reporting schemas. Exploiting the information stored in the mandatory and the optional fields, LAND-deFeND can provide the same output prescribed by the EC Floods Directive, through dedicated database views (Fig. 5 and Appendix A). We adopted the same approach (Appendix A) to ensure compatibility with the 'Observed Event' and the 'Exposed Element' feature types defined by the INSPIRE data specification on Natural Risk Zones (INSPIRE Thematic Working Group Natural Risk Zones, 2013).

Fig. 5 shows the eight views designed for compliance with the two mentioned EC Directives i.e., two views for the Floods Directive (FD_FloodEvent and FD_TypeofPotentialConsequence) and six views for the INSPIRE directive (INSPIRE_ExposedElement, INSPIRE_F_ExposedElement, INSPIRE_L_ExposedElement, INSPIRE_ObservedEvent, INSPIRE_F_ObservedEvent, and INSPIRE_L_ObservedEvent). Since the INSPIRE data specification on Natural Risk Zones (INSPIRE Thematic Working Group Natural Risk Zones, 2013) separates landslides and floods within the observed event category, we prepared two different views for landslides and floods. In Fig. 5, the red headers identify views related to damage information corresponding to the 'Exposed Elements' of the INSPIRE Directive (2007/2/EC), and to the 'Potential Consequence' of the EC Floods Directive (2007/60/EC).

4. Comparison with published databases on geo-hydrological hazards

In Fig. 6, we compare LAND-deFeND with the 13 databases and catalogues dealing with geo-hydrological hazards listed in Table 1, and with the database structures prescribed by the EC Floods Directive (European Commission - DG Environment, 2013), and the

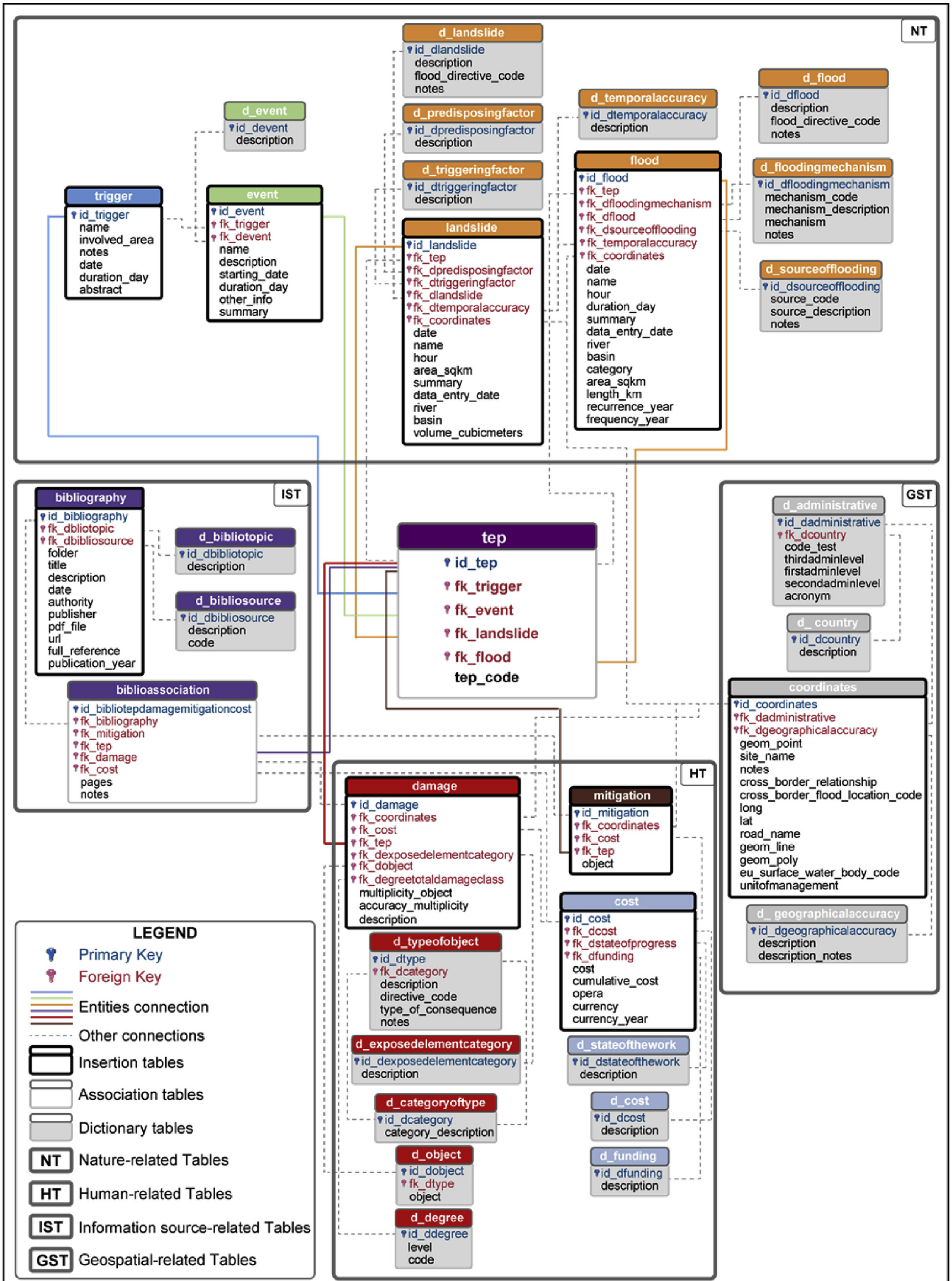

Fig. 3. Physical structure of the LAND-DeFeND database structure. See text for explanation.

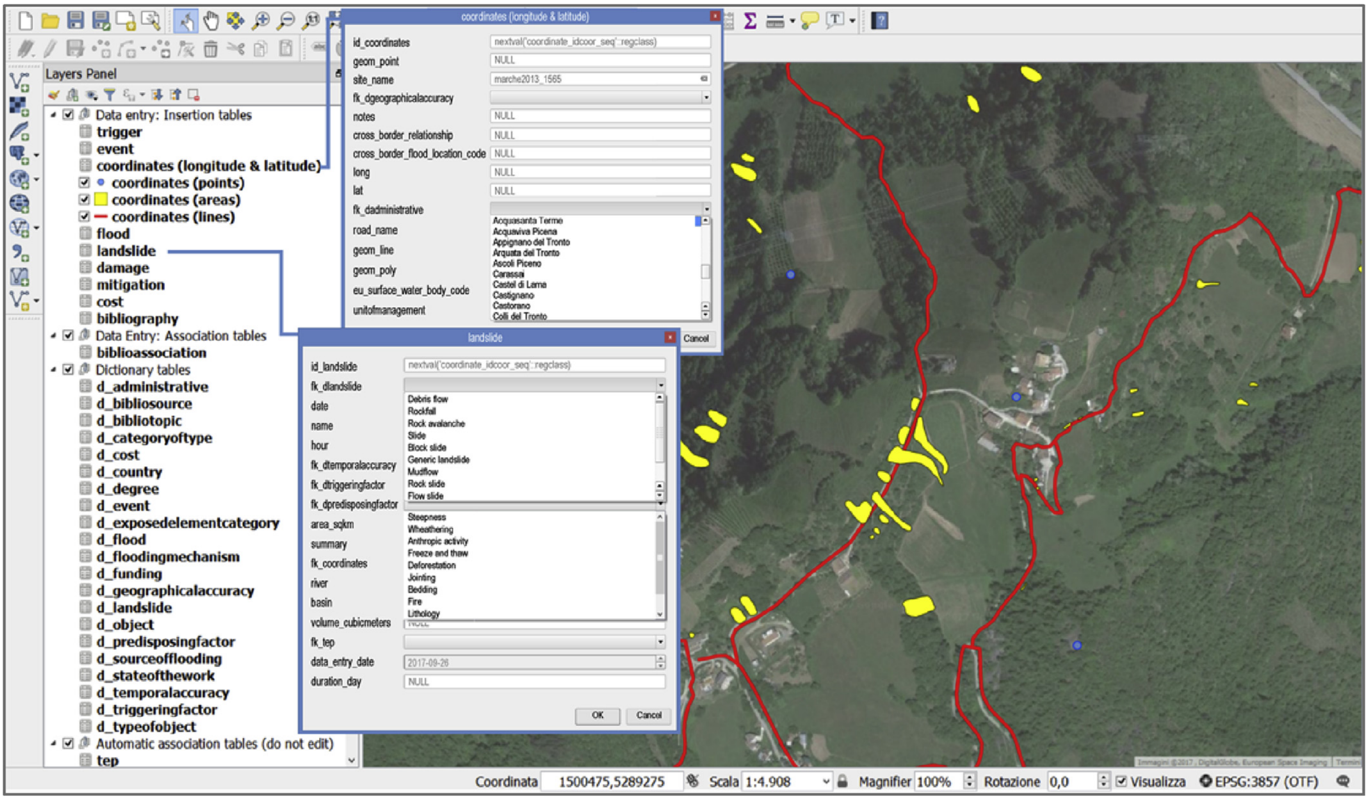

Fig. 4. LAND-deFeND QGIS® interface, used for data entry and for the editing of the geographical information.

INSPIRE data specification on Natural Risk Zones (INSPIRE Thematic Working Group Natural Risk Zones, 2013). For the comparison, we consider ten key characteristics (columns in Fig. 6) of the LAND-deFeND database structure that we consider important to store, organize, and manage information on geo-hydrological hazards and their consequences.

Inspection of Fig. 6 reveals that six (of 13) of the considered databases and schemas treat only a single geo-hydrological hazard, including four databases for floods (DFO, Sigma data, FLOOD-HYMEX, Flood Directive) and two databases for landslides (IFFI, Cartografia del Dissesto della Regione Emilia-Romagna). For four databases (IFFI, AVI, Flood Directive, and the INSPIRE data specification in Natural Risk Zones) the implementation schema is publicly available, facilitating their implementation and interoperability. Nine databases consider information on damage (including fatalities, injured people, and homeless), monetary cost and mitigation actions. None of the databases used the hierarchical subdivision of geo-hydrological hazards in different entities ("trigger", "event", "phenomenon") adopted by LAND-deFeND, and none provided a physical model for immediate implementation of the database structure. We consider the later a severe limitation that hampers the adoption of a specific database or schema. Compared to other databases structures, LAND-deFeND has the ability to record and manage geographical information using different types of geometries (i.e., points, lines, polygons). This helps locating the different phenomena and the related damage effectively e.g., a levee break along a river (using a point), a damaged road (represented by a line), and a flooded area (outlined by a polygon). Only four of the considered databases and schemas can handle different types of geometries (IFFI, Catasto Dissesti Regione Valle d'Aosta, Geologia e Dissesto ARPA Piemonte, Cartografia del Dissesto della Regione Emilia-Romagna). Finally, LAND-deFeND is compliant with the prescriptions mandated by the INSPIRE (2007/2/EC) and the Floods (2007/60/EC) Directives. Inspection of Fig. 6 shows that the other considered databases are not compliant with the two EC Directives.

5. Case studies

When designing and implementing a new database structure, a goal is to ensure that the data entry meets the data modeling rules and the design specifications required by the adopted formal data modeling techniques (Simsion and Witt, 2005). We checked the new LAND-deFeND database structure by populating it with about 1000 records, collectively listing 83 triggers, 101 events, 218 floods, 506 landslides, and 920 damages — 706 of which with associated cost data — occurred in Italy in the 16-year period between 2000 and 2015. The data vary in their level of detail, and cover a wide range of cases; from regional events that have impacted large areas and have produced severe and extensive damage, to site specific events that have caused little or no damage.

Here, we describe two examples of the use of the LAND-deFeND database. The two examples represent two major rainfall triggers that have caused widespread landslides and floods, human impact, and severe and widespread economic damage. The first example shows the effects of an intense rainfall event that hit the Messina Province, NE Sicily, southern Italy, on 1 October 2009 (Ardizzone et al., 2012). The second example shows the effects of a prolonged rainfall period from November to December 2013 in the Marche Region, central Italy (Donnini et al., 2017).

On 1 October 2009, a Mediterranean cyclonic vortex originating from the Balearic Islands generated an intense storm cell that dumped intense rainfall along the Ionian Coast of Sicily, SW of the city of Messina, with the cumulated rainfall locally exceeding 220 mm in seven hours. The intense rainfall caused flash floods and widespread — mostly shallow — landslides and debris flows that

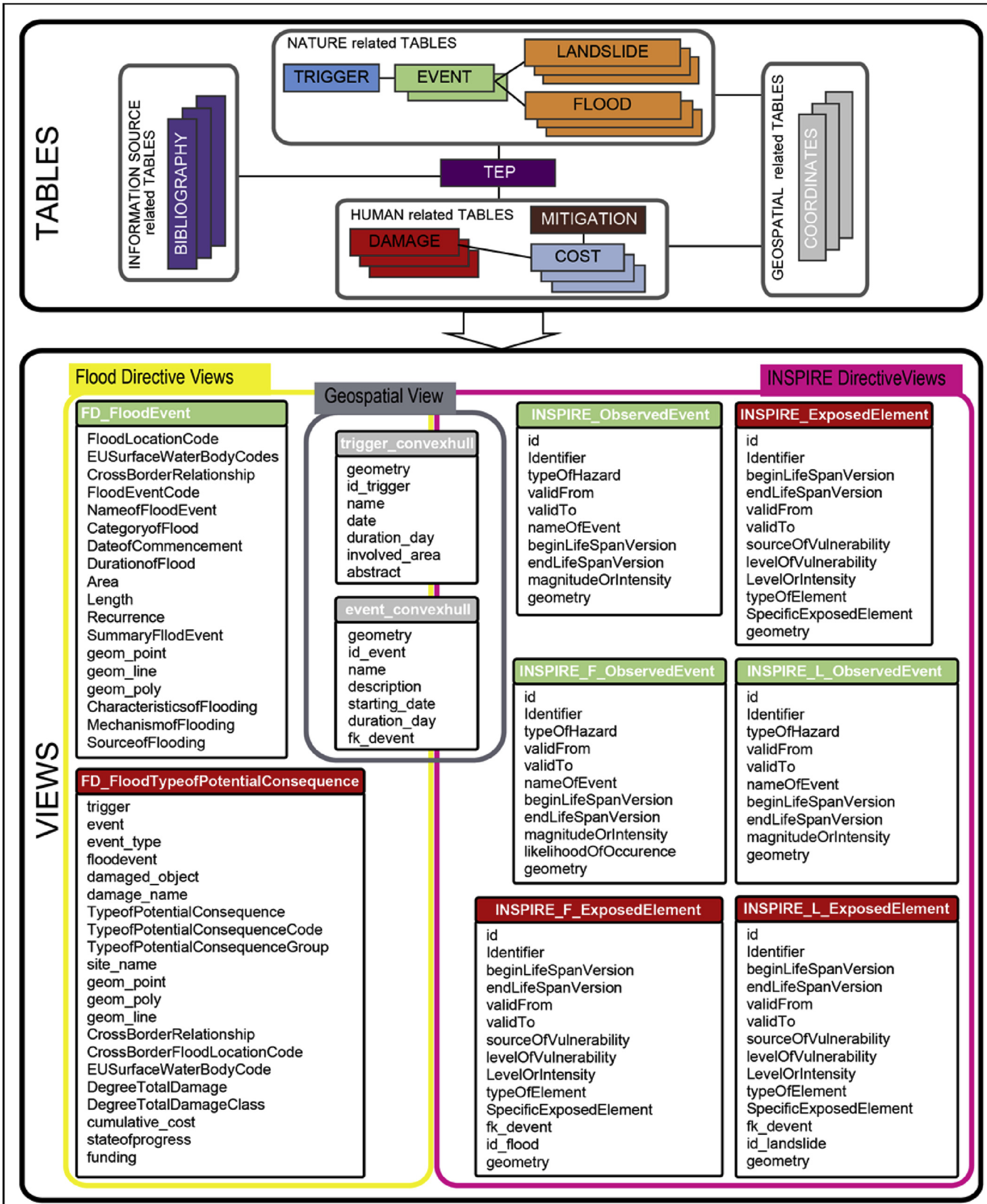

Fig. 5. Graphical representation of the 10 views for the data output of the LAND-deFeND database structure. To be compliant with the Event, Coordinates, and Damage table colours, the headers views maintain the same colours. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

	ID (Table 1)	Landslides	Floods	Database schema published	Physical model available	Hazard hierarchy	Data on damage	Data on costs	Different type of geometry	INSPIRE compliant	Flood Directive compliant
Global active archive of large flood events - DFO	2		•				•	•			
EM-Dat	3	•	•				•	•			
NatCat SERVICE	4	•	•				•	•			
Sigma data	5		•				•	•			
FLOODHYMEX	8		•				•	•			
IFFI	13	•		•			•		•		
AVI	14	•	•	•			•	•			
SICI	15	•	•				•	•			
Catasto Dissesti Regionale Valle d'Aosta	16	•	•				•		•		
ASICal	17	•	•				•				
Geologia e Dissesto ARPA Piemonte	18	•	•				•	•	•		
Cartografia del dissesto della Regione Emilia-Romagna	19	•					•		•		
ARCA	20	•	•				•	•			
Flood Directive			•	•							•
INSPIRE data specification on Natural Risk Zones		•	•	•						•	•
LAND-deFeND		•	•	•	•	•	•	•	•	•	•

Fig. 6. Main characteristics of the LAND-deFeND database structure, compared to the structures adopted by the databases dealing with geo-hydrological hazards listed in Table 1, and with the structures and schemas prescribed by the EC Floods Directive and the INSPIRE data specification on Natural Risk Zones. See text for explanation.

affected public and private buildings and roads, in urban and rural areas (Ardizzone et al., 2012). Landslides and floods caused 31 deaths, seven missing persons, and numerous injured people. After the event, a total of € 193.6-million were allocated by the national and the regional governments for the necessary recovery and risk mitigation actions (Donnini et al., 2017).

The Messina event is a prototype example of a single, geographically limited meteorological trigger that produced abundant floods and landslides, that in turn caused multiple types of damages at different sites in a relatively short period of time (a few hours). For the test case, we stored in the LAND-deFeND database information on 140 individual phenomena caused by the single intense rainfall trigger, including 91 landslides and 49 floods, and on the location and type of 100 damaged elements, including private homes, public buildings, trenches, roads and railways, aqueducts, and warehouses. We obtained the information from different sources, including (i) official documents and reports produced by local, regional, and the national governments that provided most of the damage and cost data, and (ii) newspapers and scientific papers, that provided most of the technical information, including the flood and landslide initiation mechanisms, and the location and volume of the landslides and the debris flows. For each damaged element, we stored in the database the total amount of funding – allocated or spent – for the specific restoration and risk reduction works, and we related each damage – and the related cost – to the specific phenomenon (the landslide, the flash flood) that had caused the damage.

The information stored in the LAND-deFeND database allowed for different analyses. In Fig. 7 we show a map that portrays the spatial distribution of the costs of the restoration and risk reduction measures. The damage and the cost data shown in the map can be aggregated in various ways, providing quantitative figures for (i) the total number and the total length of the damaged roads, and (ii) the total cost of the restoration and risk reduction measures. The damage and cost information can be segmented on (i) the type of roads (main or secondary road), (ii) the type of the damaging phenomenon (e.g., debris flow, shallow slide, flash flood), or on (iii) the stage of the cost (e.g., allocated, spent). These analyses provide a wealth of information to understand the impact of the high intensity rainfall event on the different types of vulnerable elements present in the affected area (Donnini et al., 2017).

The second example shows that LAND-deFeND can store, organize, and manage information on geo-hydrological events triggered in a long period – from November to December 2013 – in a relatively large geographical area.

Between November and December 2013, a series of prolonged rainfall events caused floods and flash floods and triggered numerous landslides in the hilly and mountainous terrain of the Marche region, central Italy. Some of the landslides caused damage, particularly to the road network. Several secondary roads were interrupted, numerous small rural settlements were isolated, and hundreds of people were evacuated. We executed a reconnaissance survey in the municipalities of Acquasanta Terme and Roccafluvione where landslides were particularly abundant. We mapped each landslide in the field, and we transferred the landslide information in a GIS, in vector format. Overall, we identified and mapped 1593 landslides.

From regional and local government reports, we obtained information on the type and extent of the damage caused by landslides to the road network. Overall, we obtained cost data for 110 damaged roads, for a total of € 11.8-million (Donnini et al., 2017). In addition, we searched newspapers and other chronicles sources searching for information on the type and the extent of damage caused by landslide and floods. Exploiting the spatial join functions available in PostGIS, (i) we identified the landslides that had intersected the road network, and (ii) we linked each landslide

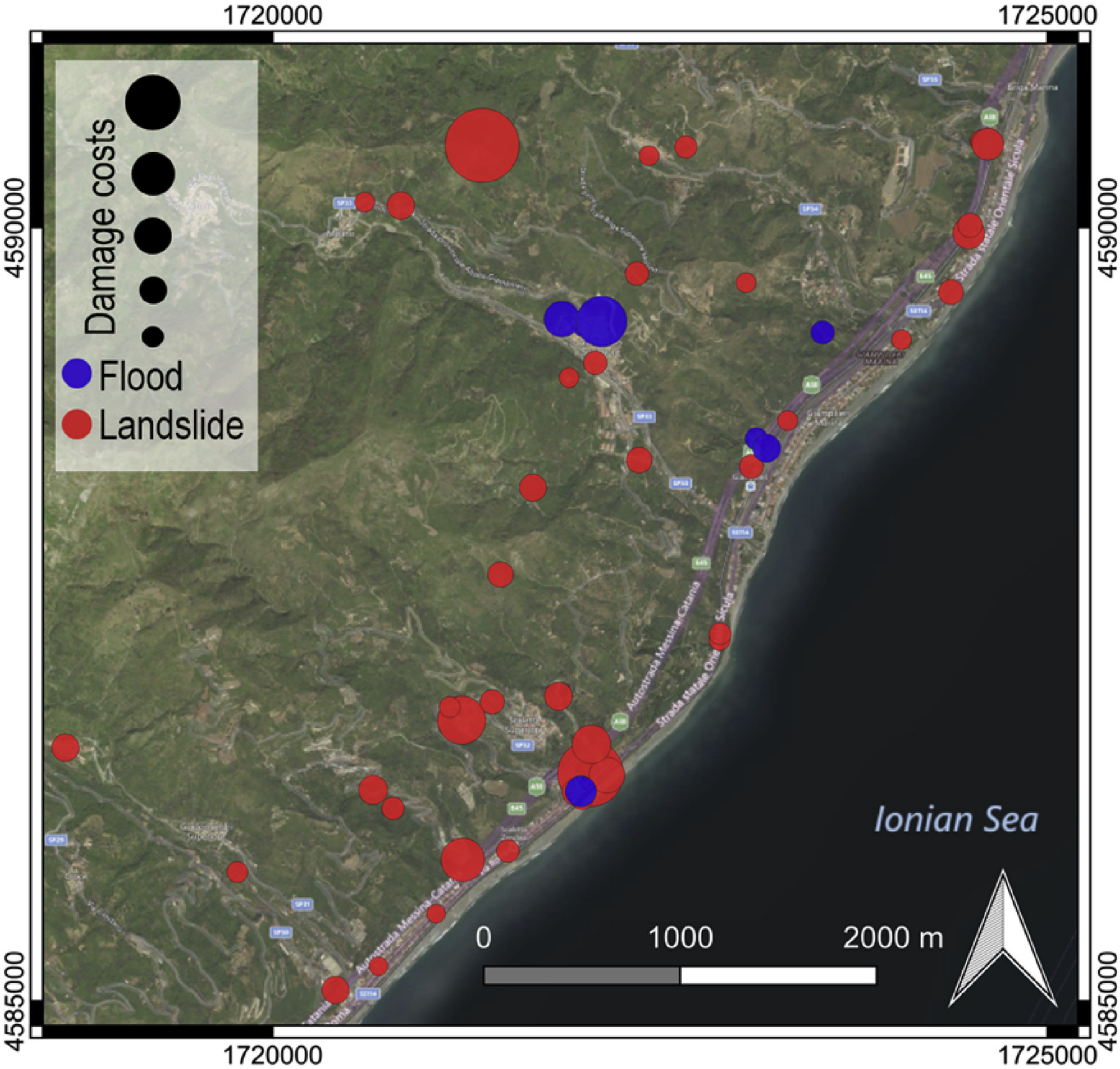

Fig. 7. Messina test case, NE Sicily, southern Italy. The map shows the sites where floods (blue dot) and landslides (red dot) have caused damage. Size of dots is proportional to the amount euro (the cost) spent or allocated for restoration and risk reduction works. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

along the road network to the information on the costs sustained for repairing the damaged roads. In Fig. 8 we show a map where landslides and damaged roads are classified – and coloured – based on (i) the average restoration cost per damaging landslide, and (ii) the number of damaging landslides per kilometres. This information is useful to execute landslide impact studies along road networks.

The Messina and the Marche case studies illustrate how complex information on geo-hydrological hazards and their consequences can be stored, organized and managed in the LAND-deFeND database structure, and then used to perform different economic, impact, vulnerability and risk analyses.

6. Discussion

Most commonly, geo-hydrological hazards (i.e., landslides and floods) occur in response to a single trigger e.g., an intense rainfall event, a prolonged rainfall period, a rapid snowmelt event, an earthquake. Multiple phenomena occurring in response to a single trigger cause a cumulative socio-economic impact, which is often difficult to quantify and to attribute to each single damaging phenomenon (e.g., a single landslide, a group of landslides, a single inundation). The problem is particularly severe when dealing with historical events, and their related information. On the other hand, it is incorrect – and not useful – to describe the natural and the socio-economic consequences of geo-hydrological hazards without identifying their triggering factors, the multiple ground effects, and the direct and indirect consequences to single individuals, communities, and the environment.

LAND-deFeND represents an effort to concentrate and manage in a single digital database structure all the issues that can arise when storing, organizing, managing and analysing information on geo-hydrological hazards obtained from different sources, covering different periods, and with different levels of accuracy. We consider

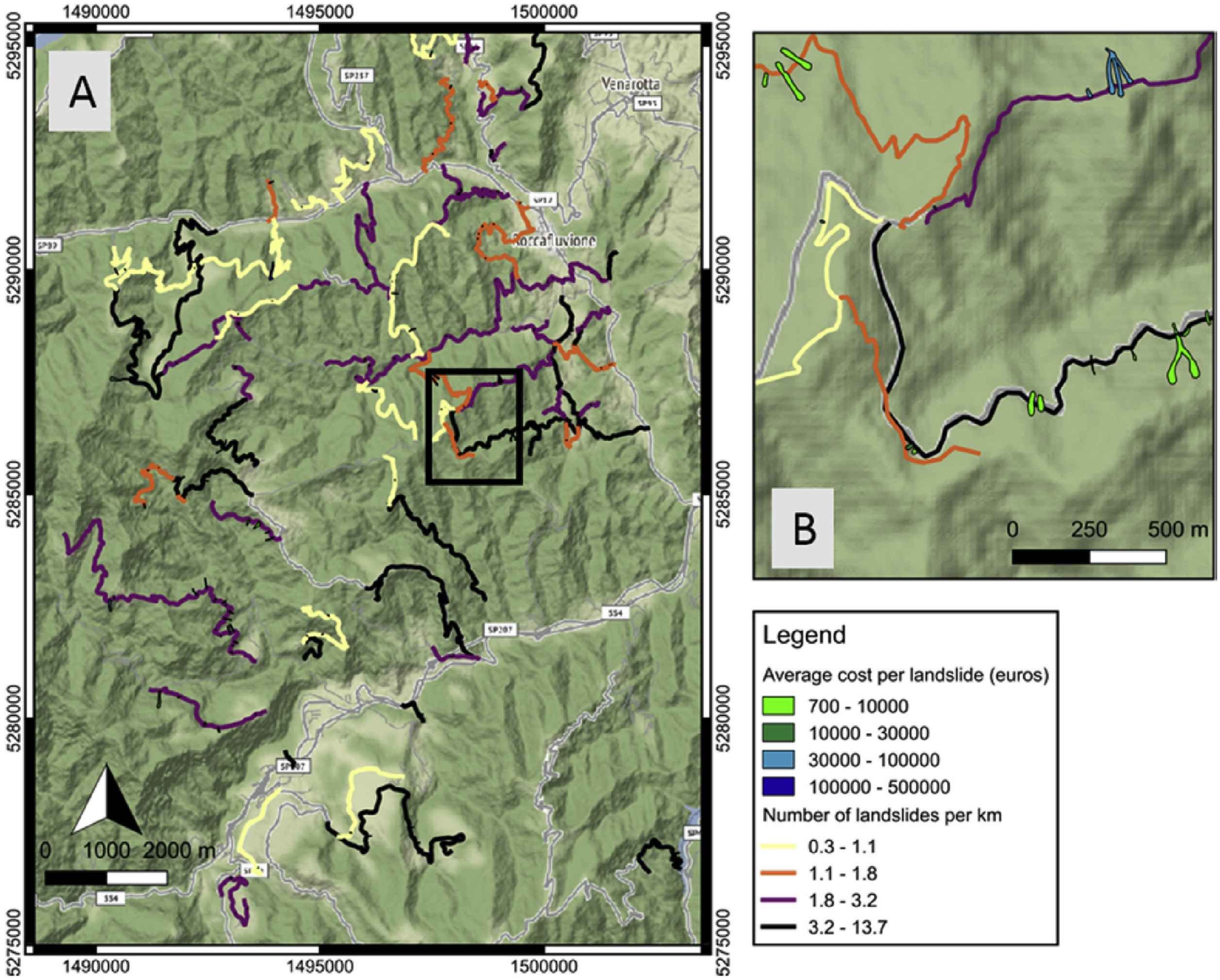

Fig. 8. Map A shows landslides and roads classified, respectively, on the basis of the average restoration cost and on the number of landslides occurred per kilometre. Map B shows the enlargement of the framed box in map A.

the new database structure a step towards an improved and more effective way of storing, organizing and managing information on geo-hydrological hazards and their consequences, where a comprehensive view of the hazards is necessary, in terms of magnitude (e.g., number of incurred phenomena, number of fatalities, people involved, number of affected buildings, structures, infrastructure), and of the corresponding impacts, including economic impact. The possibility to link information on the intensity of a phenomenon (e.g., the return period of a flood, the area, volume or velocity for a landslide) to the extent and magnitude of the damage (e.g., the type and number of damaged objects, the incurred costs) proves useful to obtain vulnerability curves; which represent key information for *ex-ante* impact, vulnerability and risk analyses. Thanks to the multiple relations in LAND-deFeND structure (Fig. 3) it is possible to adjust numerous and heterogeneous sources of information, concerning all the entities and to address the process of data collection for the future data. The Marche case study (Section 5) proves how LAND-deFeND structure facilitates the harmonization of datasets collected for different purposes.

A relevant innovation of LAND-deFeND consists in the distinction of the three main Nature-related Entities (NE) i.e., the “trigger”, the “event”, and the “phenomenon”, and the three Human-related Entities (HE) i.e., the “damage”, the “mitigation”, and the “cost”. In the database structure, a single phenomenon, event, damage or mitigation action is linked to a single meteorological or seismic trigger, as a separate but inherently linked entity. We consider the introduction of the three NE a valuable asset, useful to overcome the limitations inherent to existing database structures (Fig. 5). The EC Floods Directive states that “flood means the temporary covering by water of land not normally covered by water. This shall include floods from rivers, mountain torrents, Mediterranean ephemeral water courses, and floods from the sea in coastal areas” (2007/60/EC). Based on this definition, in the EC Floods Directive reporting schema (European Commission - DG Environment, 2013) data are aggregated at the event scale, which can be a major regional flood (e.g., where multiple mechanisms of flooding occurred at different sites, and landslides may have occurred) or a small inundation (e.g., a debris blockage of a bridge crossing a minor tributary). In the LAND-deFeND database structure, use of the “event” and “phenomenon” entities allows to store, organize and manage information on both regional and local inundations, and on their consequences.

The LAND-deFeND database structure can also store monetary information representing cost data. This allows for economic evaluations at different geographical (national, regional, municipal) and temporal (from days to decades) scales, and proves useful for the estimation of the total cost caused by a single entity i.e., a single “phenomenon”, “event”, or “trigger”, as shown in the two case studies presented in Section 5. Availability of information at different geographical and temporal scales also proves useful for analyses of the frequency of the geo-hydrological hazards and their consequences (Crovelli, 2000; Guzzetti et al., 2006).

An important aspect of the LAND-deFeND database structure is the possibility of geo-referencing a single damage (e.g., a landslide hitting a road, a levee break along a river, a collapsed bridge), and the corresponding monetary cost, independently from knowing (or not knowing) the phenomenon that has caused the damage. It is also possible to record information on the restoration cost of a single damaged object (e.g., a road, a bridge, a house) without knowing necessarily the exact location and geometry of the "phenomenon" (i.e., a landslide, a flood) that has caused the damage. This is obtained by linking the "damage" to the "event" entities. We consider this an advantage over existing database structures for geo-hydrological hazards, including the AVI project (Guzzetti et al., 1994) and the IFFI project (APAT & Trigila, 2007) databases, in which the geo-spatial information (geographic position and geometry) is related to the damaged objects, and to the phenomenon (the landslide), respectively.

Among the few available database structures capable of storing, organizing, and managing information on both landslides and floods, some are outdated (e.g., the Italian SICI - http://sici.irpi.cnr.it/), and others are designed to store information on high impact catastrophic events (e.g., the EM-Dat - http://www.emdat.be). LAND-deFeND can handle low (e.g., few fatalities) and very low (e.g., no fatalities, little damage) magnitude geo-hydrological hazards, which in many regions are the most frequent (Salvati et al., 2010, 2015). In Italy, every year an average of 100 damaging landslides are reported by government agencies (e.g., Trigila and Iadanza, 2015). However, the real number of landslides in Italy every year is larger – possibly more than one order of magnitude larger – as demonstrated by the frequent heavy or prolonged rainfall events that trigger hundreds or thousands of landslides (Guzzetti and Cardinali, 1989; Cardinali et al., 2006; Ardizzone et al., 2012; Donnini et al., 2017). The mismatch between the recorded and the occurred damaging landslides can result in an underestimation of the real impact of landslides. To understand the physical and socio-economic impact of geo-hydrological hazards, the low and very low magnitude events are important, and should not be overlooked.

In designing the LAND-deFeND database structure, we decided not to include fields that, albeit potentially relevant, would probably remain empty for most of the records, due to information scarcity. As highlighted by Trigila et al. (2010), a large number of the fields representing specific physical characteristics of landslides (e.g., depth of the sliding surface, distribution and style of activity) in the IFFI project database (APAT & Trigila, 2007) have remained empty. Similarly, in the AVI project database (Guzzetti et al., 1994; Guzzetti and Tonelli, 2004) some fields also remained mostly empty, with geotechnical information available for 1.7% of the total number of the recorded landslides, landslide length available for 9%, landslide width for 10%, and landslide volume for only 3% of all the landslides. To limit the problem, we designed LAND-deFeND to limit – as much as possible – the number of the fields and tables. This also facilitated the development and maintenance of the data-entry interfaces, and simplified the data-entry processes.

7. Conclusions

To facilitate the organization and management of information on landslides and floods, we designed and implemented the database structure LANDslides and Floods National Database – LAND-deFeND. The structure allows for storing in a single digital database physical, geographical and socio-economic data on geo-hydrological hazards and their consequences with different levels of detail; from extremely detailed information on single or multiple phenomena at a single site or in a small area, to aggregated information on damage caused by many phenomena caused by a trigger in a large region in a period of several days.

The core of the LAND-deFeND database is in three Nature-related entities i.e., the "phenomenon", the "event", and the "trigger", which collectively allow for a hierarchical organization of the information capable of handling effectively multiple conditions and cases; from a single, local phenomenon (e.g., a single landslide affecting a house), to regional events caused by vast meteorological or seismic triggers that produced vast and widespread damage (e.g., many landslides and floods in a large river basin, and the related consequences at multiple sites).

LAND-deFeND is the result of a conceptual effort aimed to build the more appropriate data structure for the management of geo-hydrological hazards at the national and regional scale. We expect that the LAND-deFeND database structure will be useful to different government agencies, land management and planning authorities, civil protection services, and to research scientists. Use of the new database structure will facilitate quantitative geo-hydrological risk analyses, the design of mitigation measures, and the adoption of effective adaptation strategies to mitigate the consequences of geo-hydrological hazards.

Acknowledgements

We thank Alessandro Sarretta for advices and comments to the text and Paola Giostrella for collecting data. Study financed partially by the Italian PRIN Project – Previsione spazio-temporale di fenomeni franosi ad alto impatto nel quadro dei cambiamenti climatici, and by a grant of the Italian National Department for Civil Protection (DPC). Elisabetta Napolitano was supported by a PRIN grant, Marco Donnini by a Fondazione Assicurazioni Generali grant, and Cinzia Bianchi by a DPC grant.

Appendix A

Here, we describe the tables and relations defined by the physical model of LAND-deFeND (Section 2.3, Fig. 3), we explain the content of the main data input, association and dictionary tables, and we describe the database views (Fig. 5). To facilitate the description of Fig. 3, and to avoid misinterpretations, we use (i) italic (e.g., *id_trigger, id_event, involved_area*) to refer to table fields (columns), and (ii) first capital letters for input tables (e.g., Trigger, Event, Damage) and association tables (e.g., Biblioassociation, Tep). The *fk_* prefix indicates a foreign key (e.g., *fk_dlandslide, fk_coordinates, fk_tep*), shown in magenta in Fig. 3. The d_ prefix refers to dictionary tables listing predefined values (e.g., d_event, d_landslide, d_municipality), shown with grey backgrounds in Fig. 3.

Nature-related tables (NT)

The "Trigger", "Event", and "Phenomenon" (Landslide, Flood) tables have unique codes (*id*) in the corresponding input tables (i.e., *id_trigger*, *id_event*, *id_landslide*, *id_flood*). In the "Trigger" table, the triggers are described in terms of their name (*name*), time of occurrence (*date*), duration in days (*duration_day*), and affected area (*involved_area*). A text field (*abstract*) can be used to store additional general information. In the "Event" table, the event is defined by its name (*name*) and described in a summary field (*description*). The temporal information is stored in the fields *duration_day* and *starting_date*, and ancillary information is stored in *other_info* and *summary*. To classify the type of event, we use the dictionary table d_event, which includes three values, 'slope dynamics', 'fluvial dynamics', and 'slope and fluvial dynamics'.

The "Landslide" and "Flood" tables have fields describing the name (*name*), the date (*date*), the hour (*hour*) of occurrence of the event, and the extent of the affected area, in square kilometres (*area_sqkm*). Other fields that the EC Floods Directive (2007/60/EC) ascribes to the Event table, in LAND-deFeND are placed in the "Landslide" and the "Flood" tables, which have common foreign key fields. The *fk_coordinates* field defines the spatial location of a single flood or landslide. The *fk_dtemporalaccuracy* field defines the confidence on the date and *hour* fields (hour, part of the day, day, and week). The *fk_tep* field connects the single Landslide or Flood to the Tep.

The "Flood" table has fields to specify the name of the river (*river*) and the river basin (*basin*), and fields for the spatial and temporal characteristics of the phenomena (i.e., *frequency_year, recurrence_year, length_km, category*). The field required by the Floods Directive to describe events (*category*) is filled automatically with the value 'past', according to article 4.2b of the directive (2007/60/EC). Using the dictionary table d_sourceofflooding, one discriminates in *fk_dsourceofflooding* the process (e.g., pluvial, fluvial, groundwater rise) causing the flood. Other specific information required by the EC Floods Directive can be stored, including (i) the flooding mechanism (e.g., overflow, blockage, defence or infrastructure failure – in *fk_dfloodingmechanism*) and (i) the flood type (e.g., flash flood, snow melt flood, slow onset flood, deep flood – in *fk_dflood*). Notwithstanding the EC Floods Directive advise to link the flooding mechanisms to the Event, we chose to specify them in the "Flood" table (*fk_dfloodingmechanism*). This is because the multiple and complex mechanisms of flooding that can occur during a major flood event are better defined in space and time if specified in the "Flood" table rather than in the more generic "Event" table.

The EC Floods Directive does not consider slope processes (i.e., landslides) and associates landslide processes (e.g., debris flow) to a flood event (2007/60/EC). Adopting the widely accepted landslide classifications of Cruden and Varnes (1996) and Hungr et al. (2014), we attributed debris flows to the landslide phenomenon, in the "Landslide" table. Other fields in the "Landslide" table include (i) the type of process (e.g., debris flow, slide, soil slip - in *fk_dlandslide*), (ii) the predisposing factor (e.g., terrain slope, jointing, bedding - in *fk_dpredisposingfactor*), and (iii) the triggering factor (e.g., snowmelt, water table fluctuation, heavy rainfall - in *fk_dtriggeringfactor*).

Human related tables (HT)

The input Human-related Tables are the "Damage", "Mitigation" and "Cost" tables.

In the "Damage" table, *fk_dobject* specifies the object damaged or destroyed by floods or landslides, using a dictionary table (d_object) that lists more than 100 types of objects (e.g., school, farm, highway, greenhouse) classified in *fk_dtype* according to the EC Floods Directive categories (e.g., B10 - social, B11 - human health, B30 - cultural heritage, B32 - landscape). In d_typeofobject, the field *fk_dcategory* specifies the four EC Floods Directive macro-categories (i.e., human health, environment, cultural heritage, and economic activity) listed in d_categoryoftype. In the Damage table, the field *fk_dexposedelemencategory* (linked to d_exposedelementcategory) classifies the damaged object adopting the EC INSPIRE Directive classification. The database structure allows for inserting other types of damaged objects, available for the next data entry. In the "Damage" table, it is also possible to specify the intensity of the damage (*fk_degreetotaldamageclass*) using qualitative classes (i.e., low, medium, high) specified in d_level. The number of the damaged or destroyed objects, and the accuracy of the information, are listed in *multeplicity_objects* and *accuracy_multiplicity*. The *fk_coordinates* defines the geographical location of the damage, and the *fk_tep* connects the single damage to the Tep table.

The "Mitigation" table contains a brief description of the type of work (*object*), and it is linked to the "Coordinates" and "Cost" tables using the *fk_coordinates,* and *fk_cost* fields. As for the "Damage" table, *fk_tep* connects the single damage to the Tep table.

In the "Cost" table, the monetary values are given as estimated, allocated, or spent (*fk_dcost*), the state of the work (in progress or executed in *fk_dstateofthework*), and the funding body (local authority, central government, European Commission, private, insurance - *fk_dfunding*). *Cost* is used for entering the amount, *currency,* and *currency_year* to better specify the cost information, and *opera* is used to provide a short description of the works.

In the database structure, a one-to-many relationship links the "Cost" table to the "Damage" and the "Mitigation" tables, using *fk_costs*. Hence, single cost information can be connected to one or more damage, or to one or more mitigation works.

Geospatial-related tables (GST)

The "Coordinates" table stores information on the geographical location and shape of the "Flood", "Landslide", "Mitigation", "Damage" and "Cost" tables. The geometries are in three fields (*geom_point*, *geom_line*, *geom_polygon*) with different data types (multipoint, multiline, multipolygon), to represent both the location and shape. The spatial accuracy of the geometries (from very rough to very accurate) is stored in *fk_dgeographicalaccuracy,* linked to the dictionary table d_geographical accuracy. Latitude and longitude data are stored in *lat* and *long* and an automatic procedure fills the *geom_point* field with corresponding multipoint geometries. Site names (*site_name*) and road names (*road_name*) can be given, where available.

The management unit (in Italy, the River Basin Authority or Districts) is provided in *unit_of_management*. In *fk_dadministrative* different local authorities are given, as in d_administrative, a dictionary valid for Italy (with *firstadminlevel* for the Region, *secondadminlevel* for the Province, and *thirdadminlevel* for the Municipality), that can be customized for other States (see Appendix B).

The information required by the EC Floods Directive on the flood location, namely 'FloodLocationCode', 'EUSurfaceWaterBodyCodes', 'CrossBorderRelationship', and 'CrossBorderFloodLocationCode', are named as: *id_geography*, *eu_surface_water_body_codes*, *cross_border_relationship*, and *cross_border_flood_location_code*.

Information source-related tables (IST)

The "Bibliography" table is linked to NT and HT, and contains reference information to the data sources. The title, the document description and date are given in *title*, *description*, and *date*. The publisher or publishing authority are given in *authority* and *publisher*. The folder in the archive where the source is stored is given in *folder*. The PDF file of the document and the Uniform Resource Locator (url) are given in *pdf_file* and *url*. The *full_reference* field records an extended reference of the document (e.g., the full reference of a scientific paper).

The "IST" table contains two dictionary tables linked to the foreign keys of Bibliography: d_bibliotopic, linked to *fk_dbibliotopic* that specifies the type of phenomenon/a referred by the bibliographic source (i.e., flood, landslide, both), and d_bibliosource, linked to *fk_dbibliosource* that specifies the type of source (e.g., scientific publication, newspaper, web-site). The Biblio association table establishes many – to – many relationships among the

Bibliography, HT and NT tables. In this association table, *fk_tep* links Bibliography to NT, and *fk_damage*, *fk_mitigation*, and *fk_cost* link Bibliography to HT.

Floods directive views

To be compliant with the article 4.2b of the EC Floods Directive (2007/60/EC), we designed database views that assemble all the LAND-deFeND fields needed to describe a flood event as required by the EC Floods Directive, and we use the same field names specified in the EC Floods Directive user guide. Two separated views were designed, the PFRA_FloodEvent (Preliminary Flood Risk Assessment Flood Event) view and the PFRA_FloodTypeofPotentialConsequence (Preliminary Flood Risk Assessment Flood Consequences) view. Since the data are referred to past events, the consequences (damage) that the Floods Directive called 'potential' in LAND-deFeND are named 'observed'.

INSPIRE views

We designed database views to encompass all the LAND-deFeND fields needed to describe the two features type spatial objects ObservedEvent and ExposedElement defined by the EC INSPIRE Data Specification on Natural Risk Zones (INSPIRE Thematic Working Group Natural Risk Zones, 2013).

Geospatial views

The LAND-deFeND database structure includes two database views named Trigger_convexhull and Event_convexhull. In these views, all the fields of the "Trigger" and the "Event" tables are shown together with a geospatial field generated using PostGIS functions. The geospatial field contains the delineation of the minimum convex polygons that include the location of the damage pertaining to each trigger or event. Once loaded in a GIS (e.g., using a direct PostGIS connection or exporting them to a common geospatial vector format), the database views allow visualizing the boundaries of the polygons, representing triggers or events, which include all the damages related to them.

Appendix B

The LAND-deFeND database structure is available as a binary database dump file of the PostgreSQL® DBMS at the following link: http://geomorphology.irpi.cnr.it/tools. A PostgreSQL® server must be running on the computer where the database structure is restored. PostgreSQL® server is open source, and is available for different operative systems (https://www.postgresql.org). PostGIS®, the spatial database extender for PostgreSQL®, is also required and can be installed during or after the installation of the PostgreSQL® server. Depending on the operative system, different methods can be used to instal PostGIS®. In GNU/Linux platforms, we recommend using the specific package management tools (e.g., apt, yum, etc.). Using Windows®, PostGIS® can be installed together with PostgreSQL® using the StackBuilder, a package manager that can be used to download and install additional PostgreSQL® tools and drivers. The StackBuilder is included into the PostgreSQL® installer. We tested the current version of LAND/deFeND with PostgreSQL® 9.1.4 and PostgreSQL® 9.6.5 (current stable release), and with PostgGIS® 2.0.0 and PostGIS® 2.3.3 (current stable release).

In the following, we assume that the pgAdmin® client, an open source software for the management of PostgreSQL® databases (https://www.pgadmin.org/), is available to the user. pgAdmin® can be installed using specific package management tools (when using GNU/Linux®) and is included in the PostgreSQL® installer (in Windows®). We tested the following procedure using PGadmin3 and PGadmin4 (current stable release) releases.

When all the software is installed, the procedure for the restoration of the LAND-deFeND database structure can be executed from the pgAdmin® graphical user interface, where the user can double click on the 'Server' node icon, and select the 'Databases' node. From the top menu bar, the user can then create a new database (menu 'Edit' and then 'Create').

In the 'New database' dialog window, the user shall enter a name for the new database. The user can conveniently choose the name 'landdefend' to exploit immediately the QGIS® interface described at the end of this Appendix). Once the new database has been created, a left-click on the node allows to choose the 'Restore' item, and to access to the corresponding window where the user can browse, select, and restore the dump file of the database structure. At the end of the procedure, the message 'Process returned exit code 0' is evidence that the database structure was imported correctly (warnings or minor errors may occur).

For restoring the database structure in a GNU/Linux environment using the command line, the following instructions can be executed on the server running PostgreSQL®:

#become the 'postgres' user

sudo su postgres

#create an empty database

createdb landdefend

#restore the database

pg_restore LAND-deFeND_structure_20170926_02.backup -d landdefend –no-owner –role=postgres

#come back to the original user

exit

In addition, the d_administrative table should be compiled with the administrative information for the country where the database structure is used. The following steps need to be executed:

- The user has to prepare a CSV (comma separated values) file with the following columns (in brackets the data type): id (integer), third administrative level (character varying, maximum 50 characters), first administrative level (character varying, maximum 50 characters), second administrative level (character varying, maximum 50 characters), acronym (character varying, maximum 2 characters), country (character varying, maximum 2 characters). An example of the content of a line of the CSV file for Italy is: 12058091, Roma, Lazio, Roma, RM, IT.
- Using pgAdmin, the user shall navigate through the database structure, and then right clicks on the name of the Table named d_administrative. In the contextual menu, the item 'import' allows choosing the csv file, and to import the data into the table.

To facilitate the adoption and use of the LAND-deFeND database structure, we prepared a QGIS® (QGIS Development Team, 2017) project for data entry and visualization of the tables and maps (Fig. 4). The QGIS® project is provided together with the database dump file. It assumes that the new PostgreSQL® database is named -landdefend-, and is running on the 'localhost' database server available through the standard 5432 port. To use directly the QGIS® project, QGIS® must be installed on the same computer running the PostgreSQL®server. The QGIS® project has been tested using the QGIS® releases 2.14 and 2.18 (current stable release).

References

APAT, Trigila, A., 2007. Rapporto sulle frane in Italia: il Progetto IFFI: metodologia, risultati e rapporti regionali. APAT.

Ardizzone, F., Basile, G., Cardinali, M., Casagli, N., Del Conte, S., Del Ventisette, C., Fiorucci, F., Garfagnoli, F., Gifli, G., Guzzetti, F., Iovine, G., Mondini, A.C., Moretti, S., Panebianco, M., Raspini, F., Reichenbach, P., Rossi, M., Tanteri, L., Terranova, O., 2012. Landslide inventory map for the briga and the giampilieri catchments, NE sicily, Italy. J. Maps 8 (2), 176–180. https://doi.org/10.1080/17445647.2012.694271.

Basili, R., Valensise, G., Vannoli, P., Burrato, P., Fracassi, U., Mariano, S., Monica, M., Tiberti, M.M., Boschi, E., 2008. The Database of Individual Seismogenic Sources (DISS), version 3: summarizing 20 years of research on Italy's earthquake geology. Tectonophysics 453, 20–43. https://doi.org/10.1016/j.tecto.2007.04.014.

Boschi, E., Guidoboni, E., Ferrari, G., Valensise, G., Gasperini, P., 1997. Catalogo dei forti terremoti in Italia dal 461 a.C. al 1990, vol. 2. ING-SGA, Bologna, p. 644.

Cardinali, M., Galli, M., Guzzetti, F., Ardizzone, F., Reichenbach, P., Bartoccini, P., 2006. Rainfall induced landslides in December 2004 in south-western Umbria, central Italy: types, extent, damage and risk assessment. Nat. Hazards Earth Syst. Sci. 6 (2), 237–260. https://doi.org/10.5194/nhess-6-237-2006.

Codd, E.F., 1970. A relational model of data for large shared data banks. Communications of the ACM Archive 13 (6), 377–387. https://doi.org/10.1145/362384.362685.

Codd, E.F., 1972. Further Normalization of the data base relational model. In: Rustin, R. (Ed.), Database Systems. Prentice Hall, Englewood Cliffs, New Jersey, pp. 33–64.

Crovelli, R.A., 2000. Probability Models for Estimation of Number and Costs of Landslides. United States Geological Survey Open File Report 00–249.

Cruden, D.M., Varnes, D.J., 1996. Landslide types and processes. In: Turner, A.R., Schuster, R.L. (Eds.), Landslides: Investigation and Mitigation, Sp. Rep. 247. Transportation Research Board, National Research Council. National Academy Press, Washington, DC, pp. 36–72.

Date, C.J., 1999. An Introduction to Database Systems, eighth ed. Addison-Wesley Longman, ISBN 0-321-19784-4.

Donnini, M., Napolitano, E., Salvati, P., Ardizzone, F., Bucci, F., Fiorucci, F., Santangelo, M., Cardinali, M., Guzzetti, F., 2017. Impact of event landslides on road networks: a statistical analysis of two Italian case studies. Landslides. https://doi.org/10.1007/s10346-017-0829-4.

European Commission, Joint Research Centre, 2014. Floods in Emilia Romagna and Liguria, Italy (2014-01-22). European Commission, Joint Research Centre [Dataset]. http://data.europa.eu/89h/134ea47f-e50c-4996-b069-31f405def59f. (Accessed July 2017).

Guidoboni, E., Ferrari, G., Mariotti, D., Comastri, A., Tarabusi, G., Valensise, G., 2007. CFTI4Med, Catalogue of Strong Earthquakes in Italy (461 B.C.-1997) and Mediterranean Area (760 B.C.-1500). INGV-SGA. http://storing.ingv.it/cfti4med/. (Accessed July 2017).

Guzzetti, F., Cardinali, M., 1989. Carta Inventario dei Fenomeni Franosi della Regione dell'Umbria ed aree limitrofe. CNR, Gruppo Nazionale per la Difesa dalle Catastrofi Idrogeologiche Publication, p. 204.

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., 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., Stark, C.P., Salvati, P., 2005. Evaluation of flood and landslide risk to the population of Italy. Environ. Manag. 36, 15–36. https://doi.org/10.1007/s00267-003-0257-1.

Guzzetti, F., Galli, M., Reichrnbach, P., Ardizzone, F., Cardinali, M., 2006. Landslide hazard assessment in the Collazzone area, Umbria, Central Italy. Nat. Hazards Earth Syst. Sci. 6, 115–131. SRef-ID: 1684-9981/nhess/2006-6-115.

Hervás, J., 2013. Landslide inventories in Europe and policy recommendations for their interoperability and harmonisation, ISBN 978-92-79-27994-2. https://doi.org/10.2788/75587. A JRC contribution to the EU-FP7 SafeLand project.

Hungr, O., Leroueil, S., Picarelli, L., 2014. The Varnes classification of landslide types, an update. Landslides 11, 167–194. https://doi.org/10.1007/s10346-013-0436-y.

Kron, W., Steuer, M., Low, P., Wirtz, A., 2012. How to deal properly with a natural catastrophe database analysis of flood losses. Nat. Hazards Earth Syst. Sci. 12, 535–550. https://doi.org/10.5194/nhess-12-535-2012.

Locati, M., Rovida, A., Albini, P., Stucchi, M., 2014. The AHEAD portal: a gateway to European historical earthquake data. Seismol Res. Lett. 85 (3), 727–734. https://doi.org/10.1785/0220130113.

Llasat, M.C., Llasat-Botija, M., Petrucci, O., Pasqua, A.A., Rosselló, J., Vinet, F., Boissier, L., 2013. Towards a database on societal impact of Mediterranean floods within the framework of the HYMEX project Nat. Hazards Earth Syst. Sci. 13, 1337–1350. https://doi.org/10.5194/nhess-13-1337-2013. www.nat-hazards-earth-syst-sci.net/13/1337/2013/. (Accessed July 2017).

Martino, S., Prestinizi, A., Romeo, R.W., 2014. Earthquake-induced ground failures in Italy from a reviewed Database. Nat. Hazards Earth Syst. Sci. 14, 799–814. https://doi.org/10.5194/nhess-14-799-2014.

Menoni, S., Molinari, D., Ballio, F., Minucci, G., Mejri, O., Atun, F., Berni, N., Pandolfo, C., 2016. Flood damage: a model for consistent, complete and multipurpose scenarios. Nat. Hazards Earth Syst. Sci. 16, 2783–2797. https://doi.org/10.5194/nhess-16-2783-2016.

Michetti, A.M., Serve, L., Vittori, E., 2000. ITHACA Italy Hazard from Capable Faults: a Database of Active Faults of the Italian Onshore Territory. CD-ROM and Explication Notes. published by ANPA, Roma.

Munich Re, 2011. NatCat SERVICE Loss Database for Natural Catastrophes Worldwide. http://www.munichre.com/en/reinsurance/business/non-life/georisks/natcatservice/default.aspx. (Accessed 30 May 2017).

Petley, D.N., Allison, R.J., 1997. The mechanics of deep-seated landslides. Earth Surf. Process. Landforms 22, 747–758. https://doi.org/10.1002/(SICI)1096-9837(199708)228<747::AID-ESP767>3.0.CO;2-#.

Petrucci, O., Versace, P., 2004. ASICal: a database of landslides and floods occurred in Calabria (Italy). In: Proc. 1st Italian-Russian Workshop: New Trends in Hydrology CNR-GNDCI N. 2823, Bios, Cosenza, Italy. Database available at: http://www.camilab.unical.it. (Accessed July 2017).

QGIS Development Team, 2017. QGIS Geographic Information System. Open Source Geospatial Foundation Project. http://www.qgis.org/. (Accessed July 2017).

Rovida, A., Locati, M., Camassi, R., Lolli, B., Gasperini, P. (Eds.), 2016. CPTI15, the 2015 Version of the Parametric Catalogue of Italian Earthquakes. Istituto Nazionale di Geofisica e Vulcanologia, Milano, Bologna. http://doi.org/10.6092/INGV.IT-CPTI15.

Salvati, P., Bianchi, C., Rossi, M., Guzzetti, F., 2010. Societal landslide and flood risk in Italy. Nat. Hazards Earth Syst. Sci. 10, 465–483. https://doi.org/10.5194/nhess-10-465-2010.

Salvati, P., Bianchi, C., Fiorucci, F., Marchesini, I., Rossi, M., Guzzetti, F., 2015. Flood risk to the population and its temporal variation in Italy. In: Moramarco, Barbetta, S., Brocca, L. (Eds.), Advanced in Watershed Hydrology. Water Resources Publications, LLC, Chelsea, Michigan, US chapter 22, 433–456. ISBN-13: 978-188720185-8, ISBN-10: 188720185-8.

Simsion, G.C., Witt, G.C., 2005. Data Modeling Essentials, third ed. Morgan Kauffman Publishers, ISBN 0-12-644551-6.

Swiss Re, 2017. Sigma Natural Catastrophes and Man-made Disasters. http://institute.swissre.com/research/overview/sigma/. (Accessed July 2017).

Trigila, A., Iadanza, C., Spizzichino, D., 2010. Quality assessment of the Italian Landslide Inventory using GIS processing. Landslides 7, 455–470. https://doi.org/10.1007/s10346-010-0213-0.

Trigila, A., Iadanza, C., 2015. Rapporto di sintesi sul dissesto idrogeologico in Italia 2014. RT/SUO-IST 014/2015. ISPRA. http://www.isprambiente.gov.it/it/temi/suolo-e-territorio/dissesto-idrogeologico/sintesi_dissesto_idrogeologico_ispra_2014.pdf.

Van Den Eeckhaut, M., Hervàs, J., 2012. State of the art of national landslide databases in Europe and their potential for assessing landslide susceptibility, hazard and risk. Geomorphology 139–140, 545–558. https://doi.org/10.1016/j.geomorph.2011.12.006.

Wirtz, A., Kron, W., Low, P., Steuer, M., 2014. The need for data: natural disasters and the challenges of database management. Nat. Hazards 70, 135–157. https://doi.org/10.1007/s11069-012-0312-4.

European directives

2007/60/EC. Directive 2007/60/EC of the European Parliament and of the Council of 23 October 2007 on the Assessment and Management of Flood risks.

2007-2-EC. Directive 2007/2/EC of the European Parliament and of the Council of 14 March 2007 establishing an Infrastructure for Spatial Information in the European Community (INSPIRE). Off. J. European Union L108, 14 (http://inspire.jrc.ec.europa.eu/).

INSPIRE Thematic Working Group Natural Risk Zones, 2013. Data Specification on Natural Risk Zones. D2.8.III.12, INSPIRE Data Specification on Natural Risk Zones – Technical Guidelines. INSPIRE Directive 2007/2/EC. http://inspire.ec.europa.eu/id/document/tg/nz.

European Commission - DG Environment, 2013. Technical Support in Relation to the Implementation of the Floods Directive (2007/60/EC). A User Guide to the Floods Reporting Schemas. http://cdr.eionet.europa.eu/help/Floods/Floods_603_2016/resources/User%20Guide%20to%20the%20Floods%20schema%20v6.0.pdf.