



A predictive model of societal landslide risk in Italy

Mauro Rossi, Fausto Guzzetti*, Paola Salvati, Marco Donnini, Elisabetta Napolitano, Cinzia Bianchi

CNR IRPI, via della Madonna Alta 126, I 06128 Perugia, Italy



ARTICLE INFO

Keywords:

Landslide
Fatalities
Societal risk
Sparse data
Point model
Italy

ABSTRACT

We propose a novel approach to evaluate the spatial and the temporal distribution of societal landslide risk from historical, sparse, point information on fatal landslides and their direct human consequences. We test the approach using a record of 5571 fatalities caused by 1017 landslides at 958 sites across Italy, in the 155-year period 1861–2015. Adopting a Zipf distribution, we model societal landslide risk for the whole of Italy, and for seven physiographic and 20 administrative subdivisions of Italy. Results confirm that the Zipf distribution is adequate to describe the frequency (and the probability) of fatal landslides, and show that societal landslide risk varies in Italy depending on the largest magnitude landslide F , the number of fatal events E , and the scaling exponent of the Zipf distribution s , which controls the relative proportion of low vs. large magnitude landslides. To model societal landslide risk, we then test different grid spacings, g and circular kernel sizes, r finally adopting $g = 10$ km and $r = 55$ km. Using such geometrical constraints, we prepare maps of the variables F , E and s , revealing the complexity of landslide risk in Italy, which cannot be described properly with a single metric. For each grid cell, we assign the $\{F, E, s\}$ variables to the red, green and blue bands of a composite image to obtain a single view of landslide risk to the population of Italy. Next, we prepare risk scenarios for landslides of increasing magnitudes, which we validate checking the anticipated return period of the fatal events against information on 130 fatal landslides between 1000 and 1860, and eleven fatal landslides between January 2016 and August 2018. Despite incompleteness in the old part of the record for the low magnitude landslides, and the short length and limited number of events in the recent period 2016–2018, the anticipated return periods are in good agreement with the occurrence of fatal landslides in both validation periods. Despite the known difficulty in modelling sparse datasets, the approach provided a coherent and realistic representation of societal landslide risk in Italy. Our results give new insight on the spatial and temporal variations of societal landslide risk in Italy. We expect this to contribute to improve existing zonings of landslide risk in Italy; to foster the efficacy of national and regional landslide early warning systems; and to design and implement better landslide communication, mitigation and adaptation strategies. Our approach is general and not constrained to the information on fatal landslides available for Italy. We therefore expect the approach to be used to model societal landslide risk in other geographical areas for which adequate information is available, and to model the fatal consequences of other hazards.

1. Introduction

Landslides are a widespread hazard that in many areas of the world cause significant societal damage (Badoux et al., 2016; Dowling and Santi, 2013; Froude and Petley, 2018; Grahn and Jaldell, 2017; Guzzetti, 2000; Guzzetti et al., 2005b; Li et al., 2016; Lin and Wang, 2018; Pereira et al., 2015; Petley, 2012; Salvati et al., 2016, 2013, 2010). Due to their large natural variability, landslides and their damaging consequences remain difficult to predict (Guzzetti et al., 2012). In many areas, this limits the ability to mitigate landslide risk and to

reduce their damaging consequences, and particularly the direct consequences on the population, including deaths, missing and injured people.

Italy is one of the few countries in the world for which there is a long and accurate catalogue of landslides with human consequences (Salvati et al., 2018; Van Den Eeckhaut and Hervás, 2012). The catalogue has been updated repeatedly, and has been used to define the landslide risk to the population (Guzzetti, 2000; Salvati et al., 2016, 2013, 2012, 2010). Here, we use this unique record of historical, sparse, point fatal landslide data to prepare a spatially distributed model of

* Corresponding author.

E-mail address: Fausto.Guzzetti@irpi.cnr.it (F. Guzzetti).

<https://doi.org/10.1016/j.earscirev.2019.04.021>

Received 18 October 2018; Received in revised form 13 April 2019; Accepted 13 April 2019

Available online 20 May 2019

0012-8252/ © 2019 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

social landslide risk in Italy. The scope of the model is to provide a complete and consistent picture of the social landslide risk in Italy, and to allow the construction of social landslide risk scenarios in Italy.

The paper is organized as follows. After the introduction of the nomenclature used in the paper (Section 2), we present a record of historical fatal landslides in Italy (Section 3). Next, we define societal landslide risk and we present estimates of societal landslide risk for Italy, and for different physiographical domains and administrative subdivisions of Italy (Section 4). Next, we introduce a spatially distributed model of societal landslide risk, which we evaluate with independent information, we propose scenarios and we study the temporal variation of societal landslide risk in Italy (Section 5). This is followed by a discussion of the model outcomes and their limitations, and of the proposed method (Section 6). We conclude summarizing the main results obtained (Section 7).

2. Nomenclature

We use the term “landslide” to encompass all types of mass movements including e.g., rock falls and topples, debris flows, soil slips, earthflows, rockslides, rock avalanches, shallow and deep-seated slides, and complex and compound slope failures (Cruden and Varnes, 1996; Hungr et al., 2014). A “landslide event” consists of one or more landslides caused by the same trigger (e.g., a rainstorm, a prolonged rainfall period). “Landslide fatalities” are individuals who lost their lives due to, or as a consequence of a landslide, and who would be – or would have been – alive without the landslide event (Salvati et al., 2018). “Fatalities” are the sum of the deaths and the missing persons, and “casualties” are the sum of the fatalities and the injured people (Guzzetti et al., 2005a). “Consequences” is a synonym for “fatalities”, and “fatal event” a synonym for “fatal landslide”. The “magnitude” of a fatal landslide event measures the landslide consequences, and it is given by the number of the landslide fatalities. “Societal” (or collective) landslide risk is the risk posed by a landslide on society as a whole (Australian Geomechanics Society, Sub-Committee on Landslide Risk Management, 2000; Fell et al., 2005; Fell and Harford, 1997; Guzzetti, 2000; Guzzetti et al., 2005a; Salvati et al., 2010). Unless specified otherwise, we use “landslide risk” and “societal risk” as synonyms of “societal landslide risk”.

3. Record of fatal landslides in Italy

Using different sources of information, including archives, chronicles, newspapers, scientific journals, books, event records, damage and technical reports, web sites, blogs and other sources, Guzzetti (2000), Guzzetti et al. (2005b) and Salvati et al. (2016, 2013, 2010, 2003) have compiled and updated repeatedly a catalogue of historical landslides with direct human consequences to the population of Italy, including deaths, missing persons, injured people, homeless and evacuees, from 68 BC to August 2018. The 2083.7-year-long record lists 1178 fatal landslides that have caused 14,923 fatalities (including 14,887 deaths and 36 missing persons) at 1079 sites. In addition, the record lists information on 230,233 homeless and evacuees caused by 2206 landslides in the same period. For 54 fatal landslides in the catalogue the exact number of the fatalities is unknown, and only qualitative information is reported (e.g., many, some, hundreds). For 16 landslides the number of the fatalities is known but the exact or approximate location of the landslide is unknown. The catalogue further lists 450 landslides that have resulted in unknown numbers of homeless and evacuees.

A typical fatal landslide has caused deaths at a single location, but exceptions exist and five landslides in the catalogue have caused fatalities at multiple sites. The Vajont rockslide of 9 October 1963 killed 1917, of whom 1709 at Longarone and 208 at Erto and Casso. The Val Pola rock avalanche of 28 July 1987 killed 22 at Aquilone and seven along the road from Aquilone to Sant'Antonio Morignone. For these and

other similar cases, we used the total number of fatalities caused by the landslides (i.e., 1917 for the Vajont rockslide and 29 for the Val Pola rock avalanche). Seventy landslide events have triggered multiple fatal landslides at different sites. As an example, the “Sarno” debris flows event of 5 May 1998 (Capparelli and Versace, 2014; Cascini et al., 2011) killed 135 persons at Episcopio (caused by four landslides), eleven at Quindici, six at Bracigliano (caused by two landslides), five at Siano, and two at Sarno. For these cases, since the fatal landslides were distinct, we attributed to each landslide the corresponding number of deaths and missing persons. Further details on the sources of information and on the methods used and the problems encountered in the compilation of the catalogue of historical fatal landslides in Italy are given in Guzzetti et al. (2005b) and Salvati et al. (2016, 2013, 2010).

As for other records of historical hazards and their consequences (Albini et al., 2014; Glade et al., 2004; Guzzetti, 2000; Ibsen and Brunsden, 1996; Kirschbaum Bach et al., 2009; Li et al., 2016; Mudelsee et al., 2003; Rossi et al., 2010; Simkin et al., 2001; Stucchi et al., 2013; Van Den Eckhaut and Hervás, 2012), the completeness and accuracy of the information in the Italian catalogue of landslides with human consequences varies with time, and they improve in the recent part of the record. In this work, to model societal landslide risk in Italy we use the part of the record between 1861 and 2015 (Fig. 1), considering only landslides for which the location of the fatalities is known (Fig. 2, Table 1). In this 155-year period (t_0) the record lists information on 5761 persons who were killed (5725) or went missing (36) as a result of 1017 fatal landslides at 958 different sites (Fig. 2a). This is an average of 6.6 fatal landslides per year, and an average of 37.2 landslide fatalities per year. For twelve of the 155 years in the record (7.7%) no fatal landslides were recorded, of which 10 years before 1900 and 2 years (1904, 1944) after 1900. In the t_0 period, the largest fatal landslide was the Vajont rockslide that killed 1917 (<http://www.vajont.net>, Guzzetti et al., 2005b), the second largest was the Stava mudflow that killed 268 on 19 July 1985, and the third largest was the Cetara mud

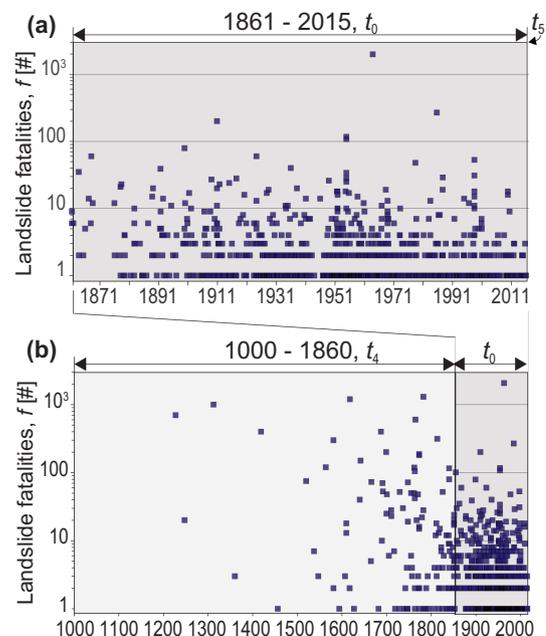


Fig. 1. Historical fatal landslides in Italy. (a) 1026 fatal landslides in the 155-year, t_0 period 1861–2015 that have resulted in 5876 fatalities, including 1017 fatal landslides that have caused 5761 fatalities for which the location of the fatalities is known. (b) 1163 fatal landslides in the 1016-year period 1000–2015 that have resulted in 14,881 fatalities, of which 137 fatal landslides in the 861-year, t_4 period 1000–1860 that have caused 9005 fatalities, including 8000 fatalities caused by 130 landslides for which the geographical location of the fatalities is known. See Table 1 for length and statistics of the fatal landslides in the different periods.

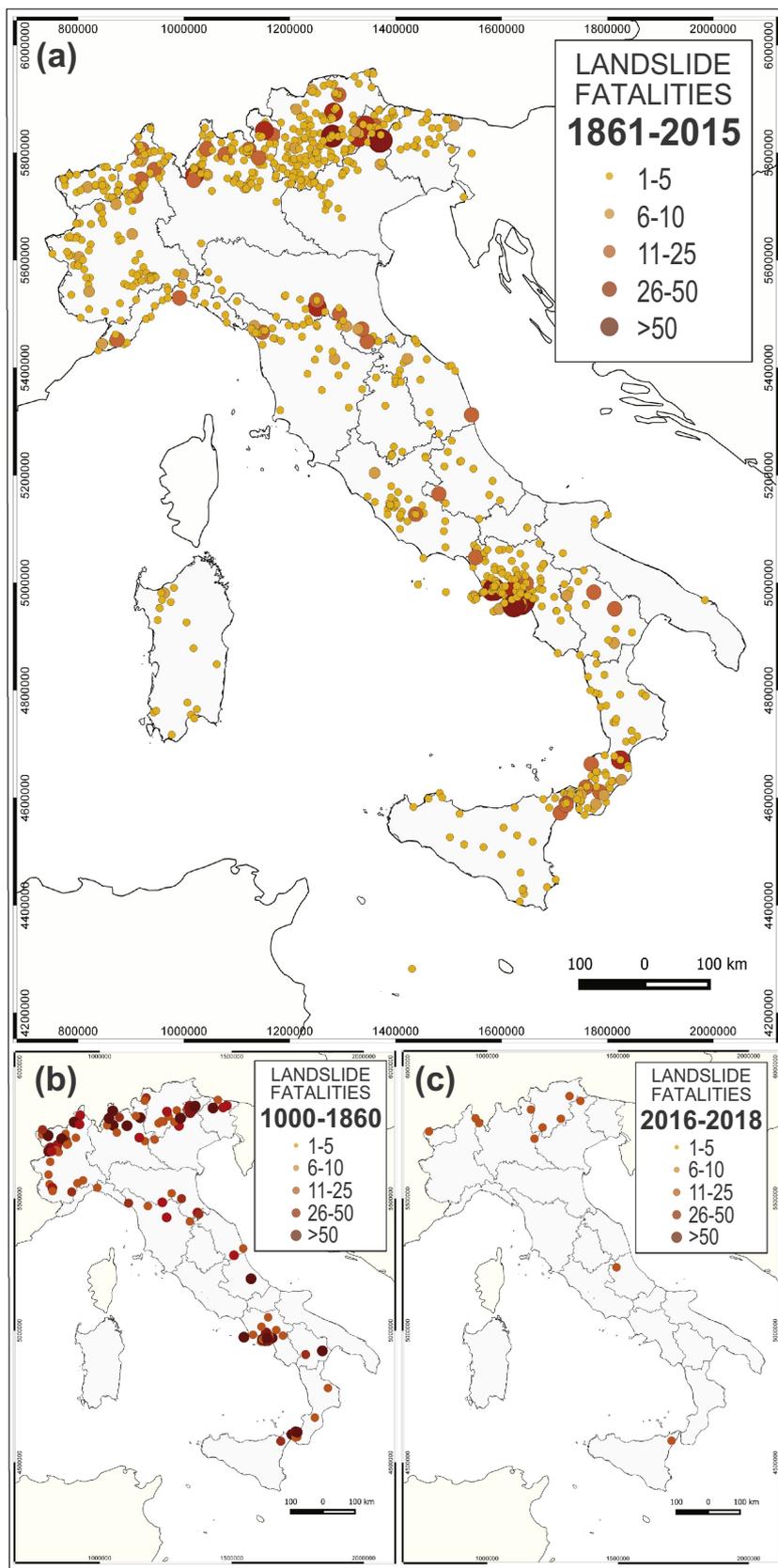


Fig. 2. Maps showing location of fatal landslides in Italy for which the location of the fatalities is known. The magnitude of the fatal landslides, measured by the number of the fatalities, is shown in five classes using dots of increasing size and shades of yellow and red colours. (a) $E = 1017$ fatal landslides ($1 \leq f \leq 1917$) at 958 sites in the 155-year, t_0 period 1861–2015. (b) $E = 130$ fatal landslides ($1 \leq f \leq 1300$) at 119 sites in the 861-year, t_4 period 1000–1860. (c) $E = 11$ fatal landslides ($1 \leq f \leq 2$) at eleven sites in the 2.7-year, t_5 period January 2016 – August 2018. See Table 1 for statistics of fatal landslides in the periods. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

and debris flows that killed 200 on 24 October 1910 (Guzzetti et al., 2005b). In the t_0 period, the vast majority of the fatal landslides (1013, 99.6%) and of the landslide fatalities (5753, 99.9%) were meteorologically induced, chiefly rainfall induced, with only a very few fatal

landslides (4, 0.4%) and only eight landslide fatalities (0.1%) caused by four earthquakes.

To evaluate the model of societal landslide risk in Italy, we further use the portions of the catalogue between 1000 and 1860 (t_4), and

Table 1

Statistics of fatal landslides in Italy, for different periods. Only landslides for which the location of the fatalities is known are considered.

Period	ID	Length [yr]	F [#]	F_{tot} [# %]	F_{ya} [#]	E [#]	E_{ya} [#]	Coverage [km ² %]
1861–2015	t_0	155	1917	5761	37.9	1017	6.6	235,900 78%
1966–2015	t_1	50	268	1289 22.6%	25.7	399 39.4%	8.0	166,000 55%
1916–1965	t_2	50	1917	3541 62.0%	70.8	488 48.2%	9.7	180,000 60%
1866–1915	t_3	50	200	882 15.4%	17.6	126 12.4%	2.5	84,600 28%
1866–2015	t_{1-3}	150	1917					
1000–1860	t_4	861	1300	8000	10.5	130	0.2	n.a.
1/2016–8/2018	t_5	2.7	2	13	4.8	11	7.1	n.a.

F , largest number of fatalities caused by a single landslide. F_{tot} , total number of fatalities. F_{ya} , yearly average number of fatalities. E , number of fatal landslides. E_{ya} , yearly average number of fatal landslides. Coverage, total (km²) and percentage (%) of the Italian territory covered by societal landslide risk models.

between January 2016 and August 2018 (t_5). The older portion of the historical record (t_4), albeit certainly incomplete for the lower magnitude events, particularly in the earlier and the intermediate parts of the record (Guzzetti et al., 2005b), lists information on 9005 fatalities caused by 137 fatal landslides, including 8000 fatalities caused by 130 landslides for which the location of the fatalities is known at 119 sites (Fig. 2b). The youngest portion of the record (t_5) lists information on 13 fatalities caused by 11 landslides at 11 different sites (Fig. 2c).

4. Societal landslide risk in Italy

Societal risk is represented and quantified constructing “frequency-consequences” plots that show the frequency (or the probability) of the fatal events against the size of the consequences, measured by the number of the fatalities (or the casualties) (Australian Geomechanics Society, Sub-Committee on Landslide Risk Management, 2000; Fell and Harford, 1997; Friedman, 2015; Guzzetti, 2000; Guzzetti et al., 2005b, 2005a). Empirical, non-cumulative distributions of population and fatality data exhibit a distinct power law, scale-invariant behaviour (Bohorquez et al., 2009; Salvati et al., 2010; Zipf, 1949), and are typically modelled using the Pareto, Zeta or Zipf distributions (Bohorquez et al., 2009; Clauset et al., 2009; Newman, 2005; Reed, 2001; White et al., 2008). The Zipf distribution, widely used in linguistics, ecology, economics, geography and other natural and social sciences (Clauset et al., 2009; Newman, 2005; Reed, 2001; Zipf, 1949), is a discrete distribution defined for a population of finite size that prescribes a power-law probability for the size of a random event that takes an integer value of at least one ($n \in \mathbb{N}^+$, $n \geq 1$) (Newman, 2005). Since the number of landslide fatalities is discrete and finite, Guzzetti et al. (2005b) have argued that the Zipf distribution is a good descriptor of the probability of fatal landslides of a given magnitude.

For a Zipf distribution, the probability mass function (PMF), equivalent to the probability density function (PDF) for discrete data, is given by:

$$PMF(f; s, F) = \frac{1}{f^s \sum_{f=1}^F \frac{1}{f^s}} \quad (1)$$

where $f \in \{1, 2, \dots, F\}$ is the number of the fatalities caused by a landslide i.e., the magnitude of the fatal event, F is the largest number of fatalities caused by a single fatal landslide in the empirical record, and $s \in \mathbb{R}^+$ is the scaling exponent of the Zipf distribution model that measures the proportion of small versus large magnitude fatal events in the record.

To determine the PMF from the empirical data in the historical record shown in Fig. 1a adopting a Zipf distribution model, we used a maximum likelihood estimation (MLE) approach to estimate the value of the s parameter (White et al., 2008), and a bootstrapping re-sampling technique (Davison and Hinkley, 1997; Efron, 1979) to determine its variability (uncertainty) σ_s . Results are summarized in Table 2 and portrayed in Fig. 3 that shows the estimated PMF of fatal landslides (a) for the whole of Italy, (c) for seven physiographical subdivisions, and

Table 2

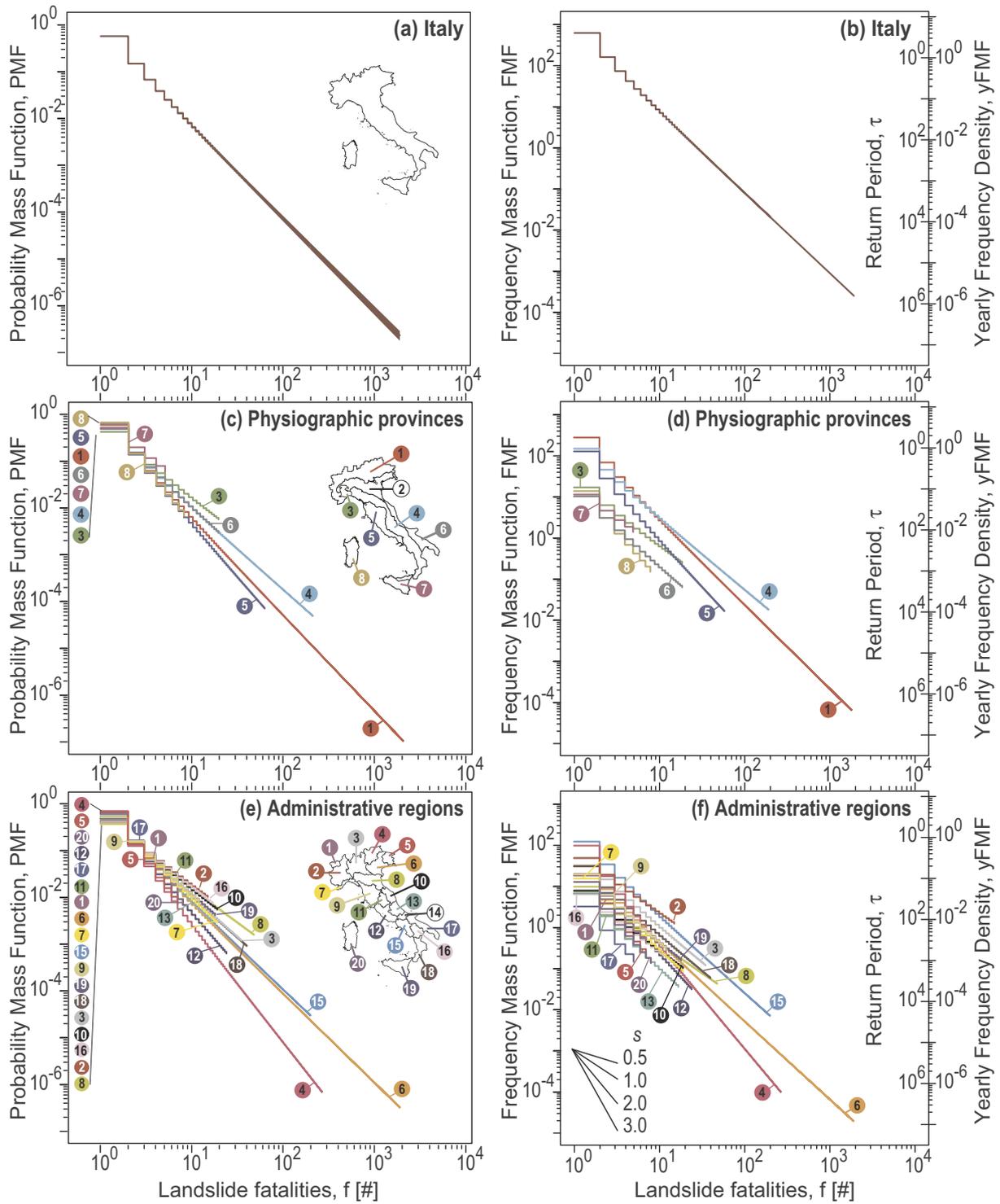
Statistics of fatal landslides and related modelled societal landslide risk in the 155-year 1861–2015, t_0 period, in Italy, in seven physiographical subdivisions (topographic divisions of Guzzetti and Reichenbach (1994), and 20 administrative subdivisions (the Italian administrative Regions). Short names are used in Fig. 3.

Area		F	F_{tot}	E	s	σ_s
Italy	ITA	1917	5761	1026	1.952	0.032
Physiographical subdivisions						
Alpine Mountain System	ALPS	1917	3354	444	2.040	0.053
North Italian Plain	POPL	1	1	1	n.a.	n.a.
Alps-Apennines Trans. Zone	ALAP	19	117	37	1.451	0.183
Apennines Mountain System	APEN	200	1716	291	1.719	0.052
Tyrrhenian Borderland	TYRR	60	439	186	2.208	0.104
Adriatic Borderland	ADRI	19	62	18	1.746	0.295
Sicily	SICI	5	40	21	1.315	0.383
Sardinia	SARD	8	32	19	2.194	0.433
Administrative subdivisions						
Valle d'Aosta	VAO	7	53	30	1.952	0.327
Piedmont	PIE	15	395	121	1.364	0.106
Lombardy	LOM	35	348	97	1.691	0.109
Trentino–Alto Adige	TAA	268	521	138	2.475	0.142
Friuli–Venezia Giulia	FVG	7	43	28	2.352	0.395
Veneto	VEN	1917	2089	58	1.897	0.127
Liguria	LIG	19	87	29	1.771	0.235
Emilia–Romagna	EMR	48	175	26	1.410	0.173
Tuscany	TUS	12	90	37	1.684	0.226
Marche	MAR	19	66	18	1.483	0.266
Umbria	UMB	6	22	12	1.914	0.541
Lazio	LAZ	24	104	48	2.178	0.218
Abruzzo	ABR	17	34	12	1.854	0.390
Molise	MOL	2	6	3	n.a.	n.a.
Campania	CAM	200	1385	231	1.846	0.065
Basilicata	BAS	17	62	15	1.426	0.295
Puglia	PUG	5	9	5	1.929	0.906
Calabria	CAL	40	238	62	1.703	0.135
Sicily	SIC	18	117	37	1.648	0.200
Sardinia	SAR	8	32	19	2.194	0.433

F , largest number of fatalities caused by a single landslide. F_{tot} , total number of fatalities. E , number of fatal landslides. s , scaling exponent of the estimated Zipf distribution model. σ_s , standard deviation of the scaling exponent of the Zipf distribution model.

(e) for 20 regional administrative subdivisions of Italy. Similarly, the three plots on the right side of Fig. 3 show the corresponding estimated Frequency Mass Functions (FMF, left y-axes), the related yearly Frequency Mass Functions (yFMF, outside right y-axes), and the expected return period, τ of the fatal landslides (inside right y-axes). To obtain the FMFs, we multiplied the corresponding PMFs by the number of the fatal landslides, E in the record. To obtain the yFMF, we normalized the FMF by the length of the historical record, $T = 155$ years (t_0 in Table 1). The return period τ is $1/\text{yFMF}$.

For our physiographical analysis, we used the topographic subdivision of Italy proposed by Guzzetti and Reichenbach (1994) (map in Fig. 3c), who classified the Italian landscape into eight provinces using



Physiographic provinces

- 1 Alpine Mountain System, ALPS 2 Po-Plain, POPL 3 Alpine-Appennines Transition Zone, ALAP 4 Apennines Mountain System, APEN
- 5 Tyrrhenian Lowland, TYRR 6 Adriatic Lowland, ADRI 7 Sicily, SICI 8 Sardinia, SARD

Administrative regions

- 1 Valle d'Aosta, VAO 2 Piedmont, PIE 3 Lombardy, LOM 4 Trentino-Alto Adige, TAA 5 Friuli-Venezia Giulia, FVG
- 6 Veneto, VEN 7 Liguria, LIG 8 Emilia-Romagna, EMR 9 Tuscany, TUS 10 Marche, MAR
- 11 Umbria, UMB 12 Lazio, LAZ 13 Abruzzo, ABR 14 Molise, MOL 15 Campania, CAM
- 16 Basilicata, BAS 17 Puglia, PUG 18 Calabria, CAL 19 Sicily, SICI 20 Sardinia, SAR

(caption on next page)

Fig. 3. Societal landslide risk models for the 155-year, t_0 period 1861–2015, (a, b) in Italy, (c, d) in seven physiographical provinces of Italy (Guzzetti and Reichenbach, 1994), and (e, f) in the 20 Italian administrative Regions. Plots to the left show Probability Mass Function (PMF) Zipf distribution models. Plots to the right show Frequency Mass Function (FMF) Zipf distribution models, with the annual FMF (yFMF, outside right y-axes) for a period $T = 155$ -year, and the corresponding estimated return period (τ , inside right y-axes). In plots c and e, columns of coloured dots show the order of PMF Zipf distribution models, from highest to lowest, for $f = 1$.

a semi-quantitative, stepwise approach that combined a cluster analysis of four derivatives of terrain elevation, with the visual interpretation of morphometric, geological and structural maps. For the geographical analysis, we used the administrative subdivision of Italy in 20 Regions (map in Fig. 3e).

Fig. 3a shows the PMF of the Zipf distribution model obtained for the $E = 1026$ fatal landslides in Italy from 1861 to 2015 (Fig. 1a), in the range of fatalities from $f = 1$ (593 events) to $F = 1917$ (the Vajont landslide). Fig. 3b shows the corresponding FMF, the yFMF for a period $T = 155$ years, and the estimated return period, τ . The scaling exponent of the Zipf model is $s = 2.21$, with an associated uncertainty $\sigma_s = 0.03$ (Table 2).

Considering the physiographical subdivisions, inspection of Fig. 3c reveals that for very small magnitude landslides ($f = 1$), the PMF is lowest in the Alps–Apennines transition zone (3) followed by the Apennines mountain system (4), and it is highest in Sardinia (8) followed by the Tyrrhenian borderland (5). For large magnitude landslides ($f \geq 25$), the PMF is largest in the Apennines mountain system (4) and lowest in the Tyrrhenian borderland (5) followed by the Alps mountain system (1) (Table 2). Inspection of Fig. 3d reveals that for very small magnitude landslides ($f = 1$) the FMF is largest in the Alps (1) followed by the Apennines (4), and lowest in the Adriatic borderland (6) followed by Sicily (7) and Sardinia (8), whereas for large magnitude landslides ($f \geq 25$) the FMF is largest in the Apennines (4), followed by the Alps (1), and is lowest in the Alps–Apennines transition zone (3) and in the Tyrrhenian borderland (5). The range of the PMF and the FMF Zipf models is largest in the Alps ($F = 1917$, due to the Vajont rockslide) and it is smallest in Sicily ($F = 5$, $E = 21$) and in Sardinia ($F = 8$, $E = 19$) (Table 2).

Considering instead the administrative subdivisions, Fig. 3e shows that for very low magnitude landslides ($f = 1$) the PMF is largest in Trentino–Alto Adige (4) followed by Sicily (19), and is lowest in Emilia–Romagna (8) followed by Piemonte (2); whereas for large ($f \geq 25$) and very large ($f \geq 50$) magnitude landslides the PMF is largest in Emilia–Romagna (8) followed by Lombardy (3) and Calabria (18), and is lowest in Trentino–Alto Adige (4) with intermediate values in Campania (15) and Veneto (6) (Table 2). Examining the FMFs for the 20 Italian regions (Fig. 3f), one finds that for very low magnitude landslides the frequency of fatal landslides is largest in Campania (15) followed by Trentino–Alto Adige (4), and is smallest in Puglia (17), whereas for very large magnitude landslides the FMF is largest in Campania (15) followed by Lombardy (3), and is smallest in Trentino–Alto Adige (4). The magnitude range of the PMF and the FMF Zipf models is largest in Veneto (6, $F = 1917$, due to the Vajont rockslide) and is smallest in Puglia (17, $F = 5$, $E = 5$) (Table 2).

In all the plots shown in Fig. 3, the position and range of the Zipf curves reveal differences in the societal landslide risk levels in the different geographical subdivisions (Guzzetti et al., 2005b; Salvati et al., 2016, 2010), complicating the evaluation of societal landslide risk. This is exemplified in Fig. 4 that shows the PMF (a) and the FMF (b) of two Zipf models of societal landslide risk for two hypothetical geographical areas i.e., the “red”, R and the “blue”, B areas. In the figure, the curve for the red area R covers a larger range than the curve for the blue area ($F_R > F_B$), has a larger number of fatal landslides ($E_R > E_B$), and it is less steep than the blue curve ($s_B > s_R$). It is difficult to decide which area has the largest societal risk. The larger fatality range in the red area ($F_R > F_B$) suggests that in the red area one expects landslides of a larger magnitude than in the blue area. The larger number of fatal landslides in the red area, and the corresponding

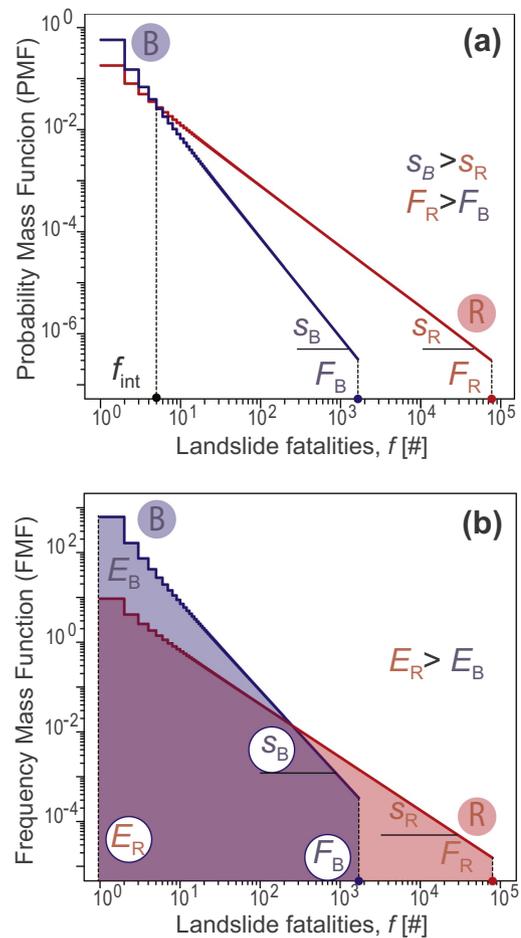


Fig. 4. Interpretation of probability/frequency – consequences plots. For two hypothetical study areas, R (red) and B (blue) areas, plot (a) shows Probability Mass Function (PMF), and plot (b) shows Frequency Mass Function (FMF) of Zipf distribution models. F_R , F_B , number of fatalities (landslide magnitude) for the most catastrophic landslide. s_R , s_B , scaling exponent of the Zipf distribution models. E_R , E_B , number of fatal landslides. See text for explanation. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

smaller number of fatal landslides in the blue area ($E_R > E_B$), indicate that fatal landslides are more frequent in the red than in the blue area. Considering the scaling exponent of the Zipf models ($s_B > s_R$) one expects a larger relative proportion of large magnitude fatal landslides in the red area than in the blue area. These observations suggest that societal landslide risk is larger in the red than in the blue area. However, considering the PMFs (Fig. 4a), we note that for small-magnitude landslides ($f \leq f_{int}$ with $f_{int} = 5$) the probability of experiencing a fatal event is larger in the blue than in the red area, and that for larger magnitude landslides ($f > f_{int}$) the probability of fatal events is larger in the red than in the blue area.

We emphasize that the position and the range of the Zipf models in Fig. 3 depend on three factors, namely, (i) the magnitude range of the fatal events, defined by the size of the event with the largest number of fatalities i.e., the largest magnitude landslide, F in the record, (ii) the number of fatal landslides, E in the record, in each geographical

subdivision, and (iii) the steepness of the curves, controlled by the model scaling parameter s . In the next section, we use these three variables $\{F, E, s\}$ to construct a model to describe the spatial and the temporal distribution of societal landslide risk in Italy, based on the available historical record of fatal landslides (Figs. 1a and 2a).

5. A model of societal landslide risk in Italy

We assume that the largest magnitude fatal landslides F , the number of fatal landslides E , and the scaling parameter of the Zipf distribution model s , together represent a good measure of societal landslide risk in Italy.

5.1. Model construction

Based on this assumption, we first partitioned the entire Italian territory (301,340 km²) in a regular square grid of size g , in km. Second, for each grid cell ($g \times g$ in size, in km²), we selected from the record of historical fatal landslides (Figs. 1a and 2a) all the events within a circular kernel of radius r , in km. Third, we used the same MLE approach used in Section 4 to determine the s_k parameter that controls the Zipf model, and its variability σ_s , for the given geographical sub-set of historical fatal landslides in the range from 1 to F_k , the largest number of fatalities caused by a landslide in the selected geographical sub-set. Lastly, we attributed to each grid cell the three model variables $\{F_k, E_k, s_k\}$, where E_k measures the number of fatal landslides in the selected sub-set, and the subscript “k” indicates values computed for the geographical sub-set of the historical record inside the kernel of size r .

To test the sensitivity of the geographical analysis to the size of the sampling grid cell and of the radius of the circular kernel, we repeated the operation for sampling grids of three sizes i.e., $g = 10, 25$ and 50 km (corresponding to grid cells of $g \times g = 100, 625$ and 2500 km²), and for circular kernels of seven sizes i.e., $r = 10, 25, 40, 55, 70, 85$ and 100 km.

Inspection of Fig. 5, which summarizes the results, reveals that the 2-sided Kolmogorov-Smirnov D statistics decreases with increasing kernel size, and it is not affected significantly by the size of the sampling grid (Fig. 5f), with the associated p -values constantly > 0.95 (Fig. 5e). We take these statistical evidences as an indication that the

Zipf distribution is well suited to model the empirical landslide fatality data in Italy (Guzzetti et al., 2005b). We further note that the p -values increase with the size of the kernel (Fig. 5e), and we attribute the result to the increasing number of empirical data found in larger kernels.

Further inspection of Fig. 5 shows that increasing the size of the kernel, the average number of fatal landslides in each grid cell increases nearly independently of the size of the sampling grid (Fig. 5a). This was expected, as a larger kernel covers a larger area where more fatal landslides are found, independently of the size of the sampling grid. Increasing the size of the kernel also increases the average scaling exponent of the Zipf distribution model (Fig. 5b), indicating that over large (small) areas the proportion of fatal landslides with a large (small) number of fatalities is smaller (larger) than in small (large) areas. This also was expected, as large and very-large magnitude events are rare in the t_0 portion of the record (10 landslides ≥ 50 fatalities (1.0%) and five landslides ≥ 100 fatalities (0.5%)). We further note that above a minimum kernel size ($r > 25$ km), the variability of the scaling parameter s , measured by its standard deviation σ_s , decreases with increasing kernel size, indicating a less uncertain estimation of the Zipf distribution models (Fig. 5c). This is confirmed by the mean value of significance of s parameter (referred as s, p -value in the figure which is obtained by a Z test to verify the parameter difference from zero) with increasing kernel size up to $r = 40$ km, and it decreases less rapidly for kernel sizes $r \geq 55$ km (Fig. 5d). We attribute the result to the larger number of empirical data found in larger kernels and used to estimate the Zipf distribution models.

In conclusion, we chose $g = 10$ km and $r = 55$ km as the “optimal” parameters used to construct our spatially distributed predictive model of societal landslide risk in Italy. The selection provides a high spatial resolution of the prediction (grid cells of 100 km²) without losing model performance. Further, using a sampling kernel area of about 9503 km² ($r = 55$ km), in each grid cell the single predictive models were determined with an average of 30 fatal landslides. This guarantees that the scaling exponent of the Zipf distribution model is robust, and its variability remains limited (on average $\sigma_s < 0.5$).

5.2. Model outcomes

The results of our modelling effort are illustrated in Fig. 6 where we

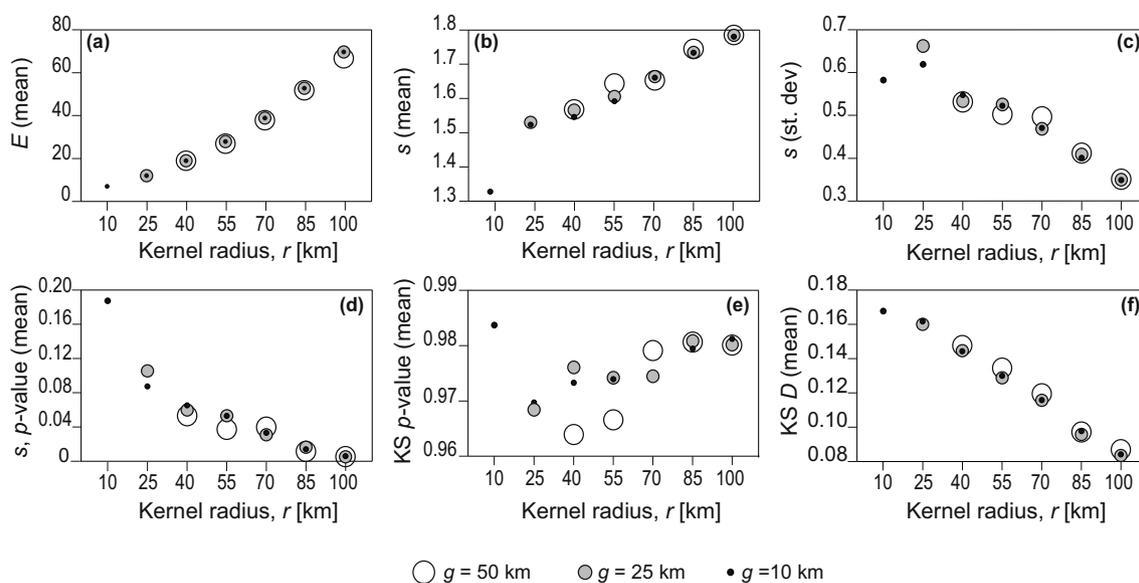


Fig. 5. Model sensitivity analysis to the sampling grid cell g , in km, and the kernel radius r , in km. Three grid sizes were tested ($g = 10, 25, 50$ km), shown by circles of increasing size. Seven kernel sizes were tested ($r = 10, 25, 40, 55, 70, 85, 100$ km), corresponding to kernel areas of about 314, 1963, 5027, 9503, 15,394, 22,698, and 31,416 km². E , number of fatal landslides. s , scaling exponent of the Zipf distribution models. KS, 2-sided-Kolmogorov-Smirnov test. D , 2-sided Kolmogorov-Smirnov statistics that measures the difference between the empirical (observed) and the Zipf-modelled data. See text for explanation.

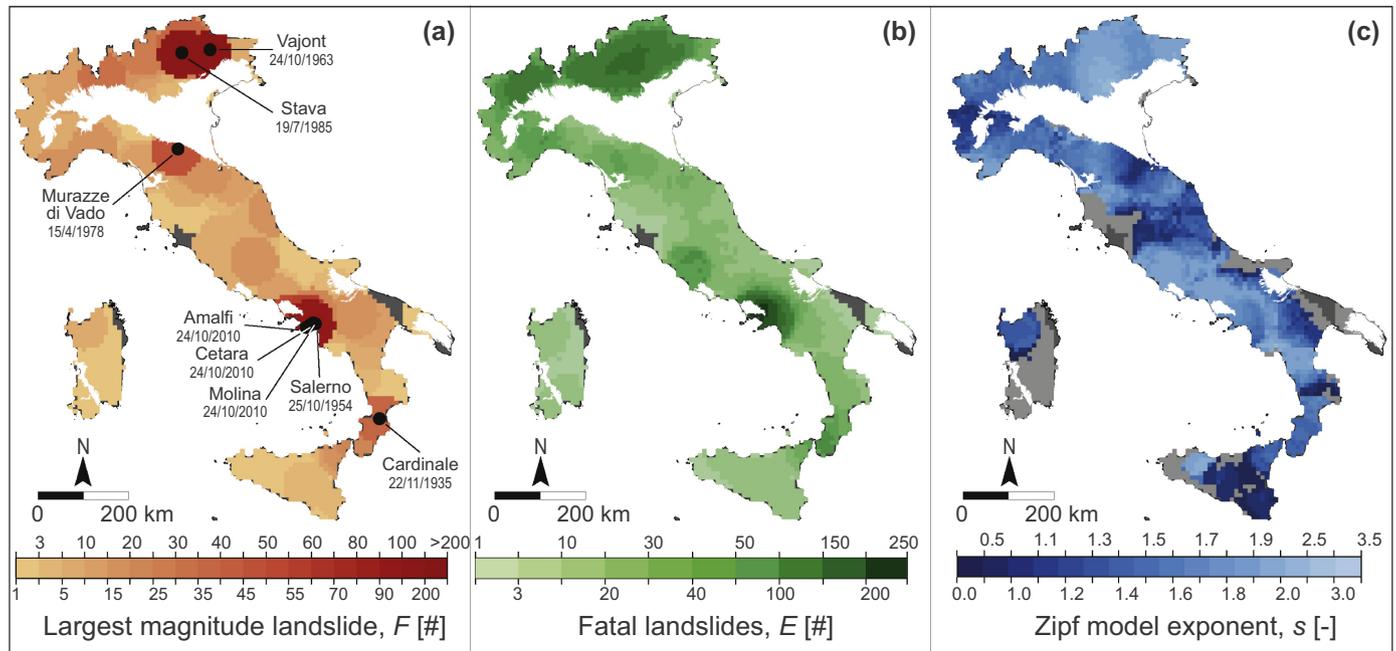


Fig. 6. Spatially distributed model of societal landslide risk in Italy. Maps show the geographical distribution of (a) the largest magnitude landslide i.e., the landslide with the largest number of fatalities, F (shades of red), (b) the total number of fatal landslides, E (shades of green), and (c) the scaling exponent of the Zipf model, s (shades of blue). In (a), (b) and (c) white areas are plains where landslides are not expected and landslide fatalities were not reported in the catalogue, and dark grey shows areas for which landslide fatality data were not available. Plains were outlined in a GIS considering (i) local terrain gradient from a coarse-scale DEM, (ii) Recent fluvial, alluvial or river deposits from small-scale geological maps, (iii) the size of the individual areas, and (iv) expert-based local refinements. In (c), light grey shows areas where the model scaling exponent, s was not calculated by the maximum likelihood estimation (MLE) procedure, mostly due to the lack of sufficient fatality data. Map (a) shows locations and dates of the eight largest magnitude landslides. All data shown/computed for the 155-year, t_0 period 1861–2015 (Table 1). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

show the geographical distributions of (a) the maximum number of landslide fatalities, F_k (i.e., the largest magnitude landslide), (b) the number of fatal landslides, E_k and (c) the Zipf distribution model scaling exponent, s_k that controls the relative proportion of low (few fatalities) vs. large (many fatalities) magnitude landslides. Where the modelled s_k parameter is small (large) the model predicts a relatively larger (smaller) number of large magnitude fatal landslides with many fatalities compared to the small magnitude events with only a few fatalities.

In the three maps, white areas are large plains (17.4%) where landslides are not expected, and landslide fatality data were not listed in the catalogue (Fig. 2). The dark grey colour shows grid cells for which no fatal landslides were found in the model kernel (4.0%), and the light grey colour (in map c) shows grid cells for which the scarcity of fatality data in the model kernel did not allow the MLE procedure to calculate the scaling exponent of the Zipf distribution model (17.6%). Overall, the model was computed for 235,900 km², 78.4% of the Italian territory (Table 1). The part of the Italian territory where the model was not computed (22.6%) matches reasonably well the percentage of territory considered as “plain” (23.2%) by the Italian National Institute of Statistics (2014) that for its classification used criteria based chiefly on elevation. We attribute the minor difference to the model interpolation procedure, which depends on the model grid size ($g = 10$ km) and the size of the kernel ($r = 55$ km), and to the fact that some fatal landslides (e.g., some of the high mobility “Sarno” debris flows) caused fatalities in areas classified as “plains” by the Istituto Nazionale di Statistica (2014).

In the historical record, very large magnitude landslides ($F_k \geq 50$) were recorded in the NE Alps ($F_k = 1917$ for the Vajont rockslide, and $F_k = 288$ for the Stava mudflow), and in Campania, southern Italy ($F_k = 200$ for the Cetera debris flow of 24 October 1910, $F_k = 117$ for the Vietri and $F_k = 108$ for the Salerno debris flows of 25 October 1954). Large magnitude landslides ($25 \geq F_k \geq 50$) were reported in the Emilia–Romagna ($F_k = 48$, the Murazze di Vado landslide of 15 April

1978), Calabria ($F_k = 40$, the Cardinale landslide of 22 November 1935) and Lombardy ($F_k = 35$, the Lemma (Faggeto Lario) landslide of 17 October 1863) regions (Fig. 6a).

The geographical distribution of the number of fatal landslides, E_k reveals a different picture (Fig. 6b). Very large numbers of fatal landslides are found in Campania ($E_{\max} = 210$, $E_{\text{avr}} = 89.5$) and in most of the Alps (126, 58.0). Large numbers of fatal events ($30 \geq E_k \geq 100$) are also reported in the Alps–Apennines transition zone (41, 29.9) (Fig. 6b), in the Rome metropolitan area and its rural surroundings (45, 38.1), in the Apuane Alps (36, 24.8), in the southern part of Calabria and in the Peloritani range, NE Sicily (47, 30.4). Conversely, low numbers of fatal landslides ($E_k \leq 10$) are found along the Tyrrhenian coast of Tuscany (10, 2.9), in parts of the Adriatic borderland, and in most of the Puglia (7, 2.5), Sicily (10, 5.6) and Sardinia (10, 5.2) regions.

A yet different picture of societal landslide risk in Italy is given by Fig. 6c, in which the darker (lighter) shades of blue show larger (smaller) estimated scaling exponents of the Zipf distribution models. Where the blue colour is lighter, the model curves are steeper ($s_k \geq 1.8$) and the corresponding area is characterized by a larger proportion of small magnitude landslides with only a few fatalities, and a proportionally smaller number of large magnitude landslides with many fatalities. Conversely, where the blue colour is darker, the model curves are less steep (gentler, $s_k \leq 1.2$), and the area is characterized by a larger proportion of large magnitude landslides with many fatalities and a smaller proportion of small magnitude landslides with only few fatalities. Fig. 6c shows that the model curves are very steep ($s_k \geq 2.0$) in the eastern and the central Alps, inside and around Rome, and around Palermo, in Sicily. The Zipf model curves are gentler ($s_k \leq 1.3$) in the western Alps, and particularly in Piedmont, in large sectors of the northern, central and southern Apennines, and in Sicily and Sardinia. Regardless of the number of the fatal landslides in the historical record, where the model curves are gentler (steeper), the model predicts a larger (smaller) proportion of large magnitude landslides compared to the areas where the model curves are steeper (gentler).

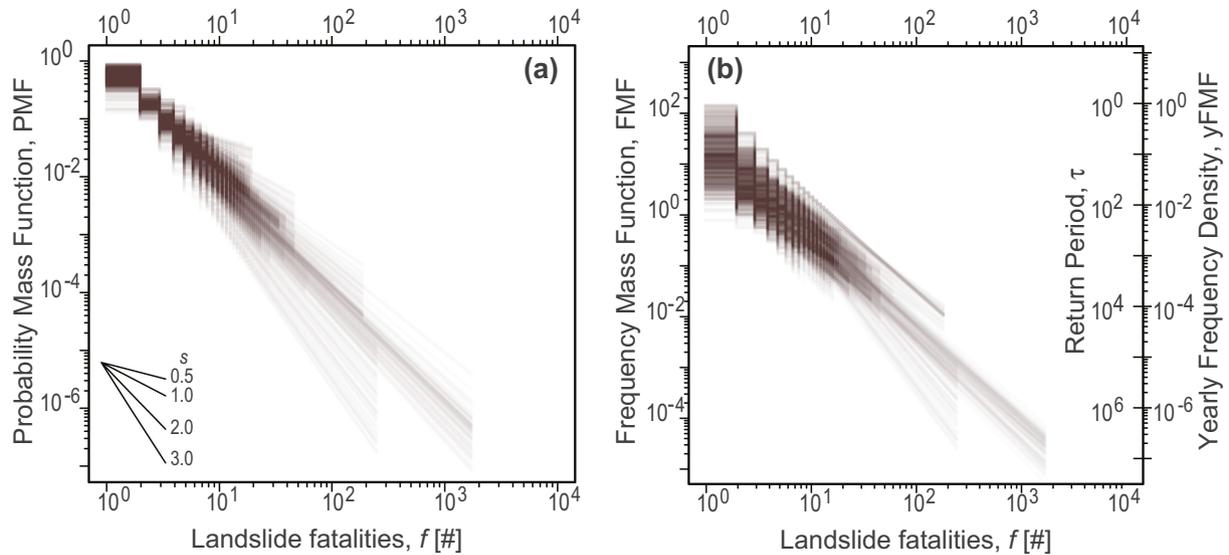


Fig. 7. Zipf's distribution models for 2320 grid cells in Italy. Models grid spacing $g = 10$ km and circular kernel radius $r = 55$ km. (a) Probability Mass Function (PMF, left axis). (b) Frequency Mass Function (FMF), with yearly FMF (yFMF, outside right y-axes) for $T = 155$ years, and return period (τ , inside right y-axes).

Fig. 7 shows the ensemble of 2320 Zipf's distribution models of societal landslide risk computed for all the grid cells in the Italian territory, with a grid spacing $g = 10$ km and a circular kernel of radius $r = 55$ km. The two plots show the ranges of (a) the Probability Mass Function (PMF) and (b) the Frequency Mass Function (FMF, left-y-axis) together with the yearly Frequency Mass Function (yFMF, outside right y-axes) and the return period (τ , inside right y-axes) of fatal landslides of different magnitudes. In the two plots, darker and lighter colours show larger and smaller numbers of model curves defined by similar model values $\{F_k, E_k, s_k\}$, and they provide a visual estimate of the most and least common modelling conditions that control societal landslide risk in Italy.

The three maps in Fig. 6 provide different, largely independent and complementary pictures of societal landslide risk in Italy. To combine the information shown in the three maps, we prepared Fig. 8 that shows a three-band, RGB (red, green, blue) false colour composite obtained assigning the F_k values (Fig. 6a) to the red band, the E_k values (Fig. 6b) to the green band, and the s_k values (Fig. 6c) to the blue band. The resulting map shows a linear combination of the $\{F_k, E_k, s_k\}$ variables that control societal landslide risk in Italy. We note that the appearance of the map depends on the histogram stretching used for displaying the three variables and on the software used to prepare the map (QGIS version 3.2.2, <https://www.qgis.org/>). Use of a different histogram stretching scheme, and of a different software may result in a somewhat different visual result.

Inspection of Fig. 8 reveals that societal landslide risk is high (light blue colours) or very high (pink colours) in the NE Alps, in the central Alps and in parts of the western Alps, and in Campania. In the Alps, high and very-high risk is due to the very large landslide magnitude (Fig. 6a), the large number of fatal landslides, and the larger relative proportion of medium-to-low magnitude landslides, measured by high-to-medium exponents of the Zipf models. Similarly, in Campania, high risk is due to the very large number and the large magnitude (Fig. 6a) of the fatal landslides, with intermediate values of the model exponent. Societal risk is intermediate (violet to light violet) in the NW Alps, in Liguria and in parts of the Alps–Apennines transition zone, in parts of the Apennines and of the Tyrrhenian borderland, and in Calabria and NE Sicily. In the western Alps, the medium risk is due to the large number of fatal landslides and the large proportion of medium to large magnitude landslides. In Rome and its rural surroundings, and in parts of the Tyrrhenian borderland, societal risk depends on the large number of fatal landslides, the steep Zipf model curves, and the small maximum landslide magnitude. In Calabria and NE Sicily, societal risk depends on the medium-large number of fatal landslides, the average steepness of

the Zipf model, and an intermediate value of the maximum landslide magnitude. Finally, societal landslide risk is low in central Tuscany, in Basilicata and Puglia, and in large parts of Sicily and Sardinia. In these areas, the low risk level depends on the low number of fatal landslides and small maximum landslide magnitude, and a larger proportion of low magnitude compared to medium and large magnitude landslides, measured by the low values of the Zipf model exponent.

5.3. Landslide risk scenarios

The model outcomes shown in Figs. 6 and 8 allow for designing scenarios of societal landslide risk in Italy. In Fig. 9, we show a set of 24 maps giving different and complementary information on the modelled societal landslide risk, for landslides of four magnitudes i.e., $f = 1, 5, 10$ and 25 fatalities. From left to right, in each row the maps show the geographical distribution of the PMF (Eqs. (1) and (B1)), the corresponding FMF (Eq. (B3)), the Complementary Cumulative Distribution Function (CCDF, Eq. (B2)), the Complementary Cumulative Frequency Distribution Function (CCFDF, Eq. (B4)), the yearly Complementary Cumulative Distribution Function (yCCFDF, Eq. (B5)) calculated considering a period $T = 155$ years, and the expected return period for the fatal landslides, $\tau_{\text{CCFDF}} = 1/\text{yCCFDF}$ (Eq. (B6)). As for the maps shown in Figs. 6 and 8, white areas show plains where landslides are not expected; in the dark grey areas landslide fatality data were not available; and in the light grey areas fatality data were too scarce to calculate the Zipf distribution models.

Visual examination of the maps (Fig. 9) allows for the following considerations. The PMF and the FMF of the fatal landslides are largest for the very low magnitude landslides ($f = 1$) and they reduce rapidly with the increase of the landslide magnitude. This was expected, as in the historical record the proportion of large magnitude landslides with many or very many fatalities is significantly smaller than the proportion of low magnitude landslides with one or a few fatalities. For very low magnitude landslides the $\text{PMF} \geq 0.30$ in most of Italy, and $\text{PMF} \geq 0.50$ in most of NE Italy, in large parts of the Apennines range and the Tyrrhenian borderland, and in NW Sicily.

The FMF is very large ($\text{FMF} > 80$) in the coastal area of Campania, in southern Italy, and subordinately ($\text{FMF} > 50$) in the NE Italian Alps. Most of the Alps, large parts of the Alps–Apennines transition zone, limited parts of the NW Apennines, the area encompassing Rome and its surroundings, the southern part of Calabria and an area in NE Sicily have $\text{FMF} > 8$. We take these as evidences of the fact that single landslide fatalities can be expected in most of the mountain areas and in

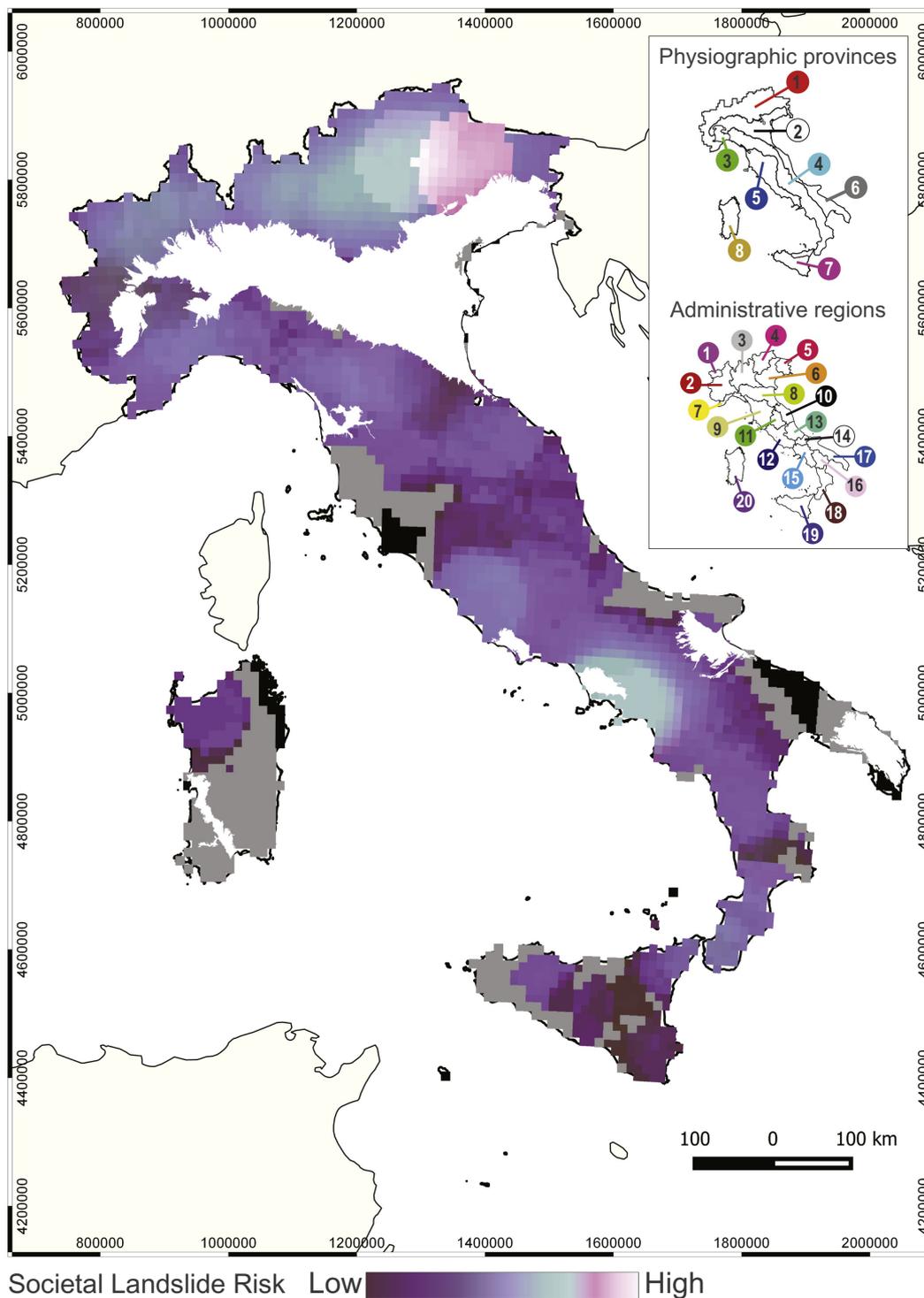


Fig. 8. False colour composite showing geographical distribution of societal landslide risk in Italy, for the 155-year, t_0 period 1861–2015 (Table 1). Map obtained rescaling from 1 to 256 the values of the F_k (Fig. 6a), the E_k (Fig. 6b) and the s_k (Fig. 6c) model variables, and assigning the three parameter maps to the red, green and blue bands. White areas are plains where landslides are not expected and landslide fatalities were not reported in the catalogue; dark grey shows areas for which landslide fatality data were not available; light grey shows areas where the model scaling exponent s was not calculated by the maximum likelihood estimation (MLE) procedure. See text for explanation. Legend: Physiographical subdivisions (Guzzetti and Reichenbach, 1994); 1, ALPS, Alpine Mountain System; 2, POPL, North Italian Plain; 3, ALAP, Alps–Apennines Transition Zone; 4, APEN, Apennines Mountain System; 5, TYRR, Tyrrhenian Borderland; 6, ADRI, Adriatic Borderland; 7, SICI, Sicily; 8, SARD, Sardinia. Administrative subdivisions: 1, VAO, Valle d’Aosta; 2, PIE, Piedmont; 3, LOM, Lombardy; 4, TAA, Trentino–Alto Adige; 5, FVG, Friuli–Venezia Giulia; 6, VEN, Veneto; 7, LIG, Liguria; 8, EMR, Emilia–Romagna; 9, TUS, Tuscany; 10, MAR, Marche; 11, UMB, Umbria; 12, LAZ, Lazio; 13, ABR, Abruzzo; 14, MOL, Molise; 15, CAM, Campania; 16, BAS, Basilicata; 17, PUG, Puglia; 18, CAL, Calabria; 19, SIC, Sicily; 20, SAR, Sardinia. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

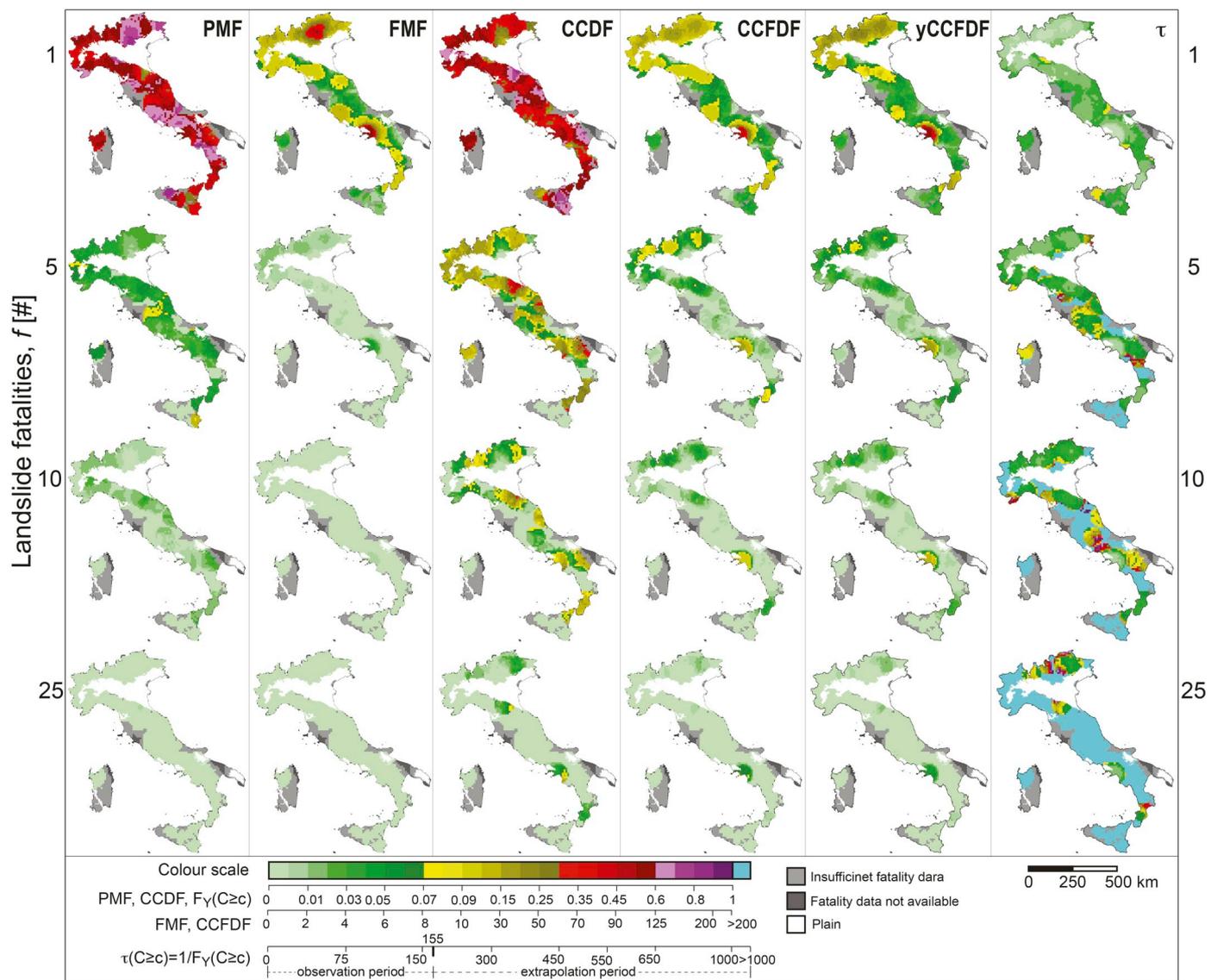


Fig. 9. Societal landslide risk scenarios for Italy. From left to right, maps show the Probability Mass Function (PMF, Eq. (B1)), the Frequency Mass Function (FMF, Eq. (B3)), the Complementary Cumulative Distribution Function (CCDF, Eq. (B2)), the Complementary Cumulative Frequency Distribution Function (CCFDF, Eq. (B4)), the yearly Complementary Cumulative Distribution Function (yCCFDF, Eq. (B5)), and the projected return period (τ_{CCFDF} , Eq. (B6)). See text for explanation. White areas are plains; dark grey shows areas for which landslide fatality data were not available; light grey shows areas where the model scaling exponent s was not calculated due to the lack of sufficient fatality data. See caption of Fig. 6 for source of plains geographical information. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of the article.)

large parts of the hills of Italy. For medium landslide magnitudes ($f = 10$), the picture is somewhat different with most of Italy exhibiting $\text{PMF} \leq 0.03$, and $\text{FMF} \leq 1$. The pattern is even more evident for the large magnitude landslides ($f = 25$), for which all of Italy has $\text{PMF} \leq 0.01$ and $\text{FMF} \leq 1.0$. This is evidence that very large magnitude fatal landslides are rare in Italy, but they can be expected in significant parts of the mountains and the hills of Italy.

The maps showing the geographical distribution of the Complementary Cumulative Distribution Function (CCDF), also known as the “survival” or “risk” function, provide a more diverse picture of societal landslide risk in Italy. For very low magnitude landslides, most of the Italian territory, and particularly the hills and the mountains, have $\text{CCDF} \geq 0.30$, indicating that the probability of experiencing $f \geq 1$ landslide fatalities is large almost everywhere in Italy. For large ($f \geq 10$) and very large ($f \geq 25$) magnitude landslides, the probability is large ($\text{CCDF} \geq 0.1$) or very large ($\text{CCDF} \geq 0.2$) in SE Emilia–Romagna, and in places in Campania, Basilicata and southern Calabria (Fig. 9).

The yCCFDF was obtained dividing the CCFDF by the length of the

observation period, $T = 155$ years (t_0 in Table 1), and therefore it shows a scaled version of the CCFDF. For very low magnitude landslides the annual frequency is large ($\text{yCCFDF} \geq 0.3$) in Campania, and is $\text{yCCFDF} \geq 0.02$ in small parts of the Alps. For medium magnitude landslides ($f = 10$), the annual frequency is large ($\text{yCCFDF} \geq 0.08$) in Campania, and is $\text{yCCFDF} \geq 0.25$ in large parts of the Alps, in parts of the Emilia–Romagna region, and in southern Calabria. For large magnitude landslides ($f = 25$) the annual frequency is large ($\text{yCCDF} \geq 0.04$) in Campania, and is $\text{yCCFDF} \geq 0.02$ in limited parts of the NE Alps and of the Emilia–Romagna region. Similarly, the return period – the reciprocal of yCCFDF i.e., $\tau_{\text{CCFDF}} = 1/\text{yCCFDF}$ – also shows a scaled version of the CCDF and the yCCFDF. For the lowest magnitude landslides ($f = 1$) the return period is short ($\tau < 30$ years) in most of the Alps, in the Alps–Apennines transition zone, in large parts of the Apennines range, of the Tyrrhenian and the Adriatic borderlands, and in NE Sicily. For large magnitude landslides ($f \geq 25$), most of the hills and mountains of Italy exhibit a very large return period ($\tau_{\text{CCFDF}} > 1000$ years, light blue colour), whereas parts of the central and the eastern Alps, of the Emilia–Romagna, Campania and Calabria regions

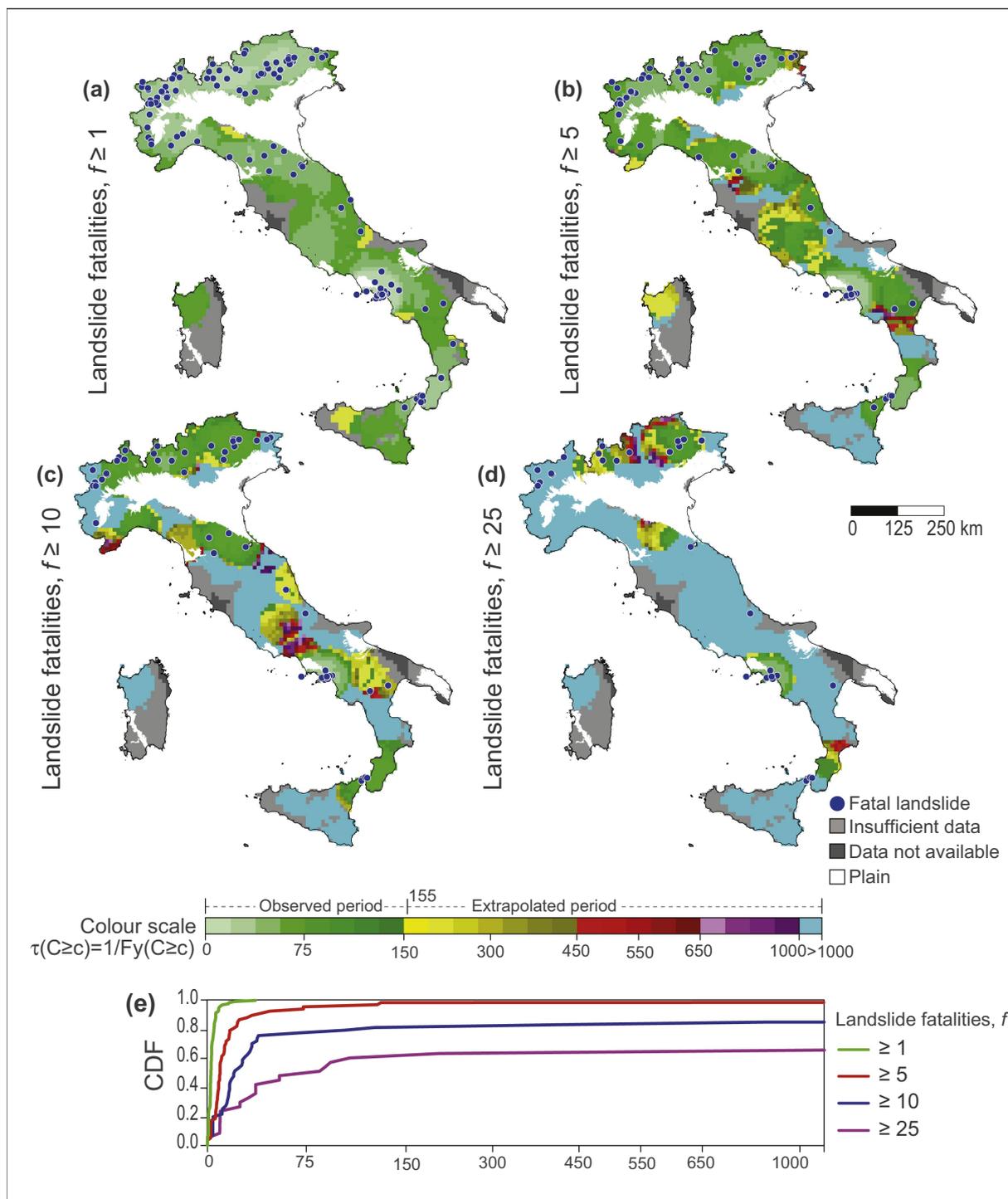


Fig. 10. Validation of societal landslide risk scenarios for Italy. Maps (a–d) show the expected return periods, τ_{yCCDF} for landslides with $f \geq 1$, $f \geq 5$, $f \geq 10$ and $f \geq 25$ fatalities. Blue dots show location of fatal landslides with $f \geq 1$, $f \geq 5$, $f \geq 10$ and $f \geq 25$ fatalities in the 861-year, t_4 period 1000–1860 (Figs. 1b and 2b, Table 1). (e) Plot shows empirical Cumulative Distribution Functions (CDF) for fatal landslides of four magnitudes ($f \geq 1$, $f \geq 5$, $f \geq 10$, $f \geq 25$ fatalities). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

have $\tau_{yCCDF} \geq 150$ years. In these areas societal landslide risk should be considered high or very high.

5.4. Validation of the landslide risk scenarios

We validated the societal landslide risk scenarios shown in Fig. 9 using information on 130 fatal landslides in the magnitude range $1 \leq f \leq 1300$ occurred at 119 sites (Fig. 2a) in the 861 years, t_4 period

1000–1860 (Fig. 1b, Table 1). With this independent information, we checked the anticipated (modelled) return period, τ_{yCCDF} for fatal landslides of magnitude $f \geq 1$, $f \geq 5$, $f \geq 10$ and $f \geq 25$ expected fatalities (Fig. 9). Fig. 10 shows four maps that portray the geographical distribution of the expected return periods for the four considered landslide magnitude scenarios, together with the location of the fatal landslides (blue dots) with $f \geq 1$ ($E = 137$ fatal landslides), $f \geq 5$ ($E = 74$), $f \geq 10$ ($E = 58$) and $f \geq 25$ ($E = 37$) fatalities occurred in the

t_4 validation period 1000–1860.

Visual analysis of Fig. 10a reveals that, regardless of their magnitude (i.e., for $f \geq 1$), the majority of the fatal landslides have occurred where the return period was anticipated to be small, $\tau_{yCCDF} \leq 30$ years. This is confirmed by the empirical Cumulative Density Function (eCDF) for $f \geq 1$ (Fig. 10e) that increases very rapidly with the return period, reaching the maximum value for $\tau_{yCCDF} = 78$ years. The CDF shows that 50% of the landslides with $f \geq 1$ have occurred in areas where $\tau_{yCCDF} \leq 6$ years, and 90% of the landslides with $f \geq 1$ have occurred where $\tau_{yCCDF} \leq 14$ years. Similarly, 50% (90%) of the landslides with $f \geq 5$ have occurred where $\tau_{yCCDF} \leq 21$ (87) years. The figures indicate that in the t_4 validation period the 67 (48.9%) reported very small and small magnitude landslides ($f \leq 5$) have occurred at a lower occurrence frequency i.e., with a higher return period than what was anticipated by the scenario (Fig. 9). We explain this result with the known incompleteness of the landslide record, particularly for the old and very old periods and for landslides with one or a few fatalities (Guzzetti et al., 2005b). For the large and very large magnitude landslides ($f \geq 25$), examination of Fig. 10d reveals that 25 landslides (67.8%) occurred where $\tau_{yCCDF} \leq 600$ years, and 12 landslides where $\tau_{yCCDF} \geq 1000$ years i.e., where the return period was anticipated to be very low. The evidence is confirmed by the plot of the empirical CDF for $f \geq 25$ (Fig. 10e) that increases slowly until about 60%, and reaches 90% for $\tau_{yCCDF} > 1000$ years. We take the evidence as an indication of the inherent difficulty of predicting accurately the temporal recurrence of very large magnitude fatal landslides in Italy.

We performed a second independent validation of the expected return period of fatal landslides in Italy (Fig. 9) using eleven fatal landslides occurred in the 2.7-year period from January 2016 to August 2018 (Fig. 2c, t_5 in Table 1). These recent landslides are in the magnitude range $1 \leq f \leq 2$ and have all occurred where the anticipated return period for $f \geq 1$ (Fig. 9) was $\tau_{yCCDF} \leq 27$ years i.e., in the areas where fatal landslides were expected with a low return period, and a correspondingly large temporal frequency of the fatal events. We acknowledge that the very recent sample is small ($E = 11$), it covers a short period (2.7 years) and a limited range of landslide magnitudes ($1 \leq f \leq 2$), but we consider the result an additional evidence of the ability of the model to predict societal landslide risk in Italy.

5.5. Temporal variation of societal landslide risk

To assess the temporal variation of societal landslide risk in Italy, we segmented the 150-year period 1866–2015 in three 50-year sub-periods i.e., t_1 , 1966–2015, t_2 , 1916–1965 and t_3 , 1866–1915 (Table 1), which collectively cover a very large part (96.8%) of the t_0 period 1861–2015. For each sub-period, we repeated the analysis performed before on the t_0 period – described in Section 5.1 – using the same “optimal” pair of geometric model parameters ($g = 10$ km and $r = 55$ km), and we then compared the geographical distributions of the model variables $\{F_k, E_k, s_k\}$ obtained for the three sub-periods. Inspection of the results, summarized in Fig. 11, reveals a general similarity of the societal landslide risk models obtained for the recent, t_1 (1966–2015) and the intermediate, t_2 (1916–1965) sub-periods, which both differ notably from the model obtained for the old, t_3 (1866–1915) sub-period. We maintain that the differences depend chiefly on the different completeness of the landslide record for the three sub-periods (Guzzetti et al., 2005b), which also affected the proportion of the Italian territory for which the risk models could be prepared (Table 1).

In the old, t_3 (1866–1915) sub-period the historical record lists $E = 126$ fatal landslides that have caused $F_{tot} = 882$ fatalities. In the 50-year sub-period, this corresponds to a yearly average of $E_{ya} = 2.5$ fatal landslides and $F_{ya} = 17.6$ landslide fatalities. These average values are lower than the corresponding figures for the intermediate, t_2 (1916–1965, $E_{ya} = 9.7$, $F_{ya} = 70.8$) and the recent, t_1 (1966–2015, $E_{ya} = 8.0$, $F_{ya} = 25.7$) sub-periods. The incompleteness of the catalogue for the old part of the record is also reflected in the smaller area covered

by the model for the t_3 sub-period, which covers only 28% of the Italian territory, compared to 60% of the model for the intermediate t_2 sub-period, and 55% of the model for the t_1 sub-period (Table 1). In the areas where model comparisons were possible, an analysis of the geographical distribution of the models of societal landslide risk for the old sub-period, t_3 reveals (i) a smaller number of fatal landslides ($E_k = 126$), (ii) a smaller largest number of fatalities caused by a single landslide ($F_k = 200$), and (iii) a lower scaling exponent s_k of the Zipf distribution models, indicating a larger relative proportion of large and very large magnitude landslides (Fig. 11). We consider the later a further evidence of the incompleteness of the old part of the record, t_3 for the small and very small magnitude fatal events (Guzzetti et al., 2005b).

An analysis of the changes between the geographical distributions of the $\{F_k, E_k, s_k\}$ variables for the models prepared for the intermediate, t_2 and the recent, t_1 sub-periods (Fig. 11) revealed that the differences between the largest number of fatalities caused by a single landslide in the two sub-periods were negligible ($-5 < F_k \leq 5$, 48.9%) or minor ($-15 < F_k \leq -5$ and $5 < F_k \leq 15$, 30.4%), with local exceptions in NE Italy due to the Vajont landslide (October 1963, in the t_2 sub-period) and the Stava mudflow (July 1985, in the t_1 period), in Campania, due to the Vietri and Salerno debris flows (October 1954, in the t_2 period) and the “Sarno” debris flows event (May 1998, in the t_1 period), in Emilia-Romagna, due to the Murazze di Vado landslide (April 1978, in the t_1 period), and in Calabria, due to the Cardinale landslide (November 1935, in the t_2 period) (Fig. 6a). We note that the geographical extent of the differences is controlled by the size of the model kernel, $r = 55$ km (Fig. 11).

The differences between the total number of fatal landslides in the two periods were mostly negligible ($-5 < E_k \leq 5$, 59.4%), with minor exceptions in Campania, Veneto and Trentino–Alto Adige (NE Italy), and in Rome and its surroundings (Fig. 11). We note that the distribution of the differences of E_k is skewed towards positive values, indicating a larger number of fatal landslides in the intermediate, t_2 sub-period, compared to the recent, t_1 sub-period. Exceptions are present in NE Piedmont and in NE Trentino–Alto Adige, in western Liguria and in NE Tuscany. In these areas, the fatal landslides were more numerous in the recent, t_1 sub-period.

The analysis of the differences in the geographical distribution of the Zipf model scaling exponent provides a more diversified picture, with a significant part of the territory characterized by negligible ($-0.25 < s_k \leq 0.25$, 35.9%) or minor ($-0.75 < s_k \leq -0.25$ and $0.25 < s_k \leq 0.75$, 41.0%) variations of the modelled scaling exponents (Fig. 11). We note that the distribution of the differences between the s_k in the two sub-periods is skewed towards negative values, particularly for the larger differences found in SW Piedmont and western Liguria, in NE Trentino–Alto Adige, and in parts of the northern and central Apennines range. This is the result of a relatively fewer number of large magnitude landslides compared to the small magnitude landslides in the recent t_1 sub-period, in these areas; and an indication of a reduction of the landslide risk to the population in these areas.

6. Discussion

6.1. Model assumptions and limitations

To construct our model of societal landslide risk in Italy (Figs. 6 and 8) from the historical record of fatal landslides (Figs. 1 and 2) we made a number of assumptions that may affect our risk assessment.

First, we assumed that the largest magnitude fatal landslides, the number of fatal landslides, and the scaling exponent of the Zipf distribution model, together are a good measure of landslide risk to the population, and that their geographical and temporal variations identify and measure changes in societal landslide risk in Italy. We maintain that both assumptions are reasonable. The three maps in Fig. 6 show different and complementary information on the spatial distribution of societal landslide risk in Italy; however, none of the three individual variables tells the full story, and only their combination provides a

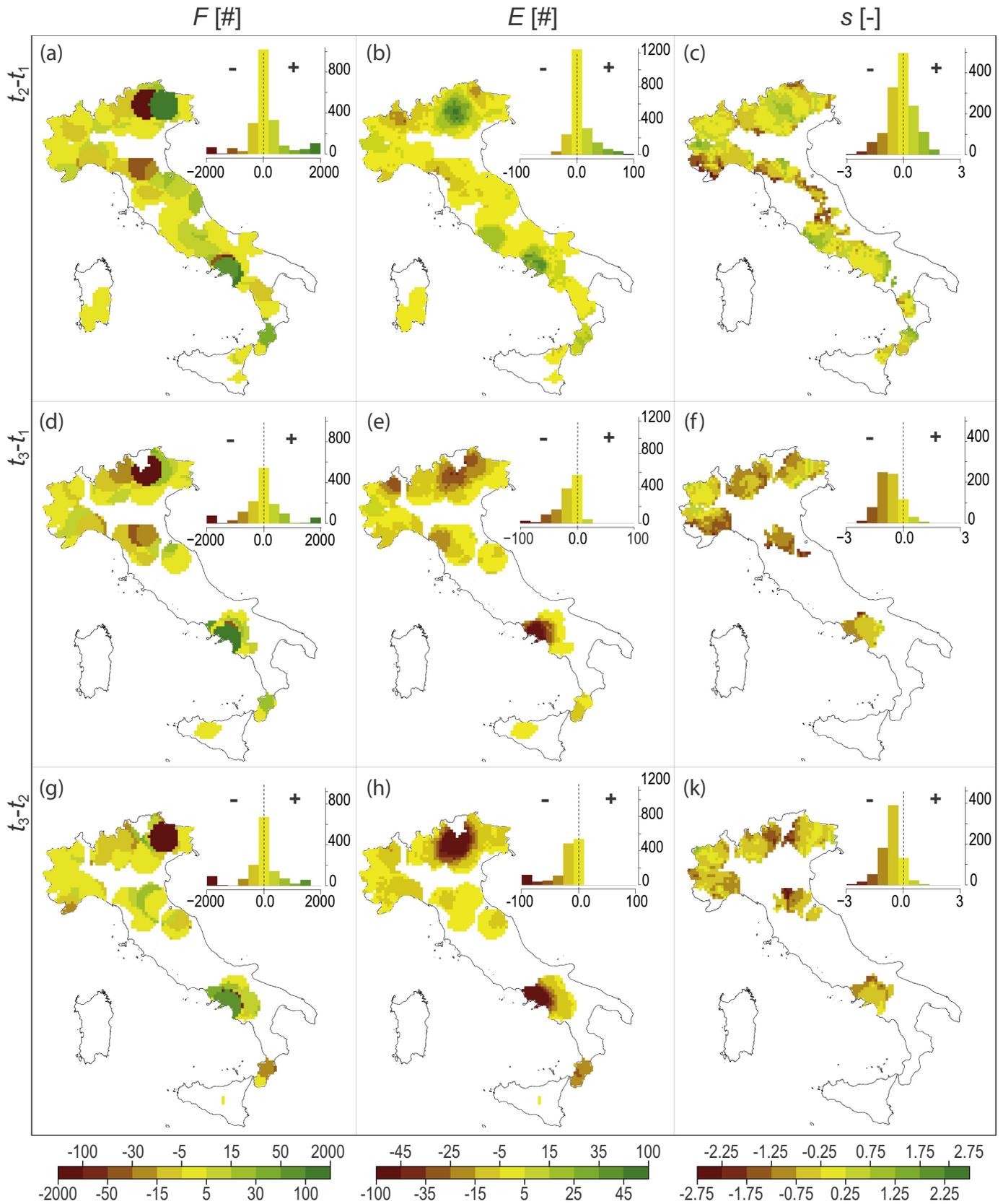


Fig. 11. Temporal variation of societal landslide risk in Italy. Maps (a–k) show the geographical distribution of the differences between the Zipf distribution model variables $\{F_k, E_k, s_k\}$ computed for three 50-year sub-periods, t_1 (1966–2015), t_2 (1916–1965), and t_3 (1866–1915). Bar charts show the differences for the entire modelled areas, which are different for the different sub-periods. See Table 1 for statistics of fatal landslides in the three sub-periods, and text for explanation. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

realistic evaluation of societal landslide risk. Similarly, our temporal analysis (Fig. 11) revealed variations in the geographical distribution of landslide risk. Some of the variations depend on the incompleteness of the old part of the record, but other variations outline real changes in landslide risk to the Italian population in the considered period. Based on the results of the temporal analysis, one could argue that the estimation of societal landslide risk in Italy would be better pursued using the fatality data in the recent, t_1 and the intermediate, t_2 sub-periods, which are the most complete. However, the selection of a shorter period (1916–2015) would have resulted in fewer fatal landslides ($E = 887$) and, consequently, a set of more uncertain Zipf models (i.e., models with larger σ_s) and a smaller area covered by the models (Table 1). We therefore consider the selection of the t_0 period 1866–2015 appropriate.

Second, we assumed that the Zipf distribution (Clauset et al., 2009; Guzzetti et al., 2005b; Newman, 2005; Zipf, 1949) was adequate to represent the frequency and the probability of different magnitude fatal landslides in Italy. Inspection of the Zipf distribution models obtained for the whole of Italy, and for the physiographical and the administrative subdivisions (Fig. 3), confirms that the Zipf distribution is an adequate model for the magnitude of the fatal landslides in Italy. The reduced uncertainty associated to the model estimates (σ_s , Table 2), which depends on the number of the fatal landslides, and the fact that the scaling exponent s does not change significantly if the size of the modelling kernel and the grid size are changed within reasonable ranges (Fig. 5), further confirms that the Zipf distribution is a robust descriptor of the magnitude of the fatal landslides in Italy.

Third, we assumed that the physical conditions that control landslide hazard – and hence societal risk – have not changed in the examined period. We note that the geological conditions (e.g., lithology, structure, seismicity) have not changed significantly in the 155-year period. Morphological modifications have occurred locally, but widespread changes have not occurred. Meteorological and climate conditions have changed, with a general reduction in the number of wet days balanced by an increase in the intensity of the rainfall events (Brunetti et al., 2006, 2004, 2001); however, the changes were probably not large enough to influence significantly landslide risk and its geographical and temporal distributions (Gariano and Guzzetti, 2016). Land use and land cover, which also contribute to controlling landslide hazard (Sidle, 2006), have changed in about half of the Italian territory, with the changes accelerated in the last decades (Falcucci et al., 2006); however, we do not have indications that these changes have influenced widely the geographical or the temporal distribution of societal landslide risk in the considered period.

Lastly, we assumed that the anthropic factors that condition societal landslide risk (e.g., the population distribution and density) have not changed in Italy in the 155-year considered period. This is a strong assumption; because the population of Italy has almost tripled from 1861 (22.2 millions) to 2015 (60.7 millions) (<http://demo.istat.it/pop2015/index.html>). However, we note that the growth of the population was largest in the plains, where landslides do not occur and our model is not applicable (Figs. 6 and 8). From the 1920s, and progressively more in the second half of the 20th century, there was a generalized migration from the mountain areas – that locally suffered a net loss of inhabitants – and partly from the rural areas to the urban areas, which are located chiefly in the plains (Guzzetti et al., 2005b; fig. 14.5 in Salvati et al., 2016), where we did not experience fatal landslides in the past according to our record (Fig. 2), and in the lowland hills (Guzzetti et al., 2005b; Salvati et al., 2016). We also note that the increase in the size of the population is matched by a significant increase in the extent of the built-up areas and the length of the road and railway networks along which many fatal landslides have occurred. Lastly, we note that in many mountain and hilly areas, tourism changes the size of the population (i.e., the number of vulnerable elements) seasonally, weekly, and even daily. We conclude that it is difficult to determine the contrasting effects of the uneven spatial and temporal variations of the population on landslide risk in Italy. We further conclude that determining societal landslide risk based on the statistical analysis of past

fatal landslides remains difficult.

6.2. Modelling a sparse dataset

The record of historical fatal landslides used in this study has temporal and geographical dimensions, with both dimensions characterized by sparse point measures; in time for the day of the fatal landslides (Fig. 1) and in space for the point location of the fatal events (Fig. 2). In the 56,612-day period (155 years) between 1 January 1861 and 31 December 2015 (t_0 , Table 1), for the vast majority of the days in the record (55,741, 98.5%) no fatal landslides were reported, and only 871 days (1.5%) have one or more landslides listed in the record. This corresponds to 1 day with one or more fatal landslides every 65 days; a very sparse record. Similarly, considering the 3009, 10 km × 10 km grid cells used for modelling, only 509 cells (16.9%) have one or more fatal landslides reported in the cell, and the remaining 2500 cells (83.1%) have no fatal landslides reported. This corresponds to one cell with fatal landslides every six cells; a rather sparse spatial information.

Working with sparse point datasets is problematic (Chao, 1989; Shepperd and Cartwright, 2001; Witt and Malamud, 2013), and even more so if the datasets are sparse in multiple dimensions and they cover large areas and long periods, like our record (Figs. 1 and 2). Our relatively simple modelling approach proved effective in handling the multivariate (F_k, E_k, s_k), multi-dimensional (space, time), sparse catalogue of fatal landslides in Italy. We expect that the same approach can be used to model similar catalogues of fatal landslides in other geographical areas (Badoux et al., 2016; Dowling and Santi, 2013; Grahn and Jaldell, 2017; Li et al., 2016; Lin and Wang, 2018; Pereira et al., 2015; Petley, 2012), or even globally (Froude and Petley, 2018; Petley, 2012). We also expect the approach to be able to model the fatal point consequences of other hazards (Salvati et al., 2018), including floods and flash floods (Guzzetti et al., 2005b; Pereira et al., 2015; Salvati et al., 2012, 2010; Špitalar et al., 2014), snow avalanches (Boyd et al., 2009), meteorological hazards (Badoux et al., 2016; Borden and Cutter, 2008; Myung and Jang, 2011; Rappaport, 2000) and earthquakes (Albini et al., 2014; Allen et al., 2010; Doocy et al., 2013; Li et al., 2014; Spence et al., 2011; Stucchi et al., 2013). When modelling the consequences of hazards different from landslides, investigators should check that the Zipf distribution is an adequate descriptor of the fatal consequences of the studied hazard. They should also consider if the three variables used in this study, $\{F_k, E_k, s_k\}$ are adequate to evaluate the risk posed to the population by the investigated hazard. We stress that the modelling approach is not limited to the type and number of variables used in this study, nor to the Zipf distribution. The model variables and the Zipf distribution can be changed, to tailor the modelling to the specific hazard conditions. We further recommend that investigators perform a thorough analysis of the sensitivity of their models to the geometric characteristics of the data and the model, testing different grid spacing and kernel sizes. This is because the “optimal” model spacing and grid size may vary depending on the type of hazard, the completeness and sparseness of the record, and the extent of the study area.

6.3. Considerations on societal landslide risk in Italy

Societal landslide risk in Italy cannot be described by a single metric (Fig. 6). We expect the same to be the case in other geographical areas, and globally (Froude and Petley, 2018; Petley, 2012). Our analysis showed that landslide risk in Italy varies geographically and temporally. Geographically, landslide risk to the population is very high in the NE Alps and in the coastal area of Campania, southern Italy, and is high in large parts of the central Alps, in Liguria, in parts of the northern Apennines range, and in southern Calabria and NE Sicily (Fig. 8). Conversely, landslide risk to the population is low in parts of central Italy, in Sicily and in Sardinia. The analysis has further revealed that the probability of experiencing fatal landslides is significant almost everywhere in Italy, and that large magnitude fatal events, albeit rare, can occur in large

parts of the mountains and the hills of Italy. The result is a consequence of the high propensity of the Italian territory to landslide risk to the population. Societal landslide risk depends on landslide hazard and on the vulnerability, distribution and abundance of the population (Australian Geomechanics Society, Sub-Committee on Landslide Risk Management, 2000; Fell and Harford, 1997; Glade et al., 2004; Varnes and International Association of Engineering Geology, 1984). However, the degree of the weightings of the two components vary, locally. In the NE Alps landslide risk is very high primarily because the hazard is high, and in Campania because the population density is high. Similarly, in Rome risk is high because of the large population density.

Our analysis outlined temporal variations in societal landslide risk, with a generalized higher risk (i.e., more fatal events, more destructive events) in the t_2 , sub-period 1916–1965 than in the t_1 , sub-period 1966–2015 for most of the country, with local exceptions in Campania, Veneto, Trentino–Alto Adige and Lazio where higher risk levels were experienced in the more recent t_1 , sub-period (Fig. 11). The observed temporal variations may depend on changes in landslide hazards, in the distribution and abundance of the vulnerable elements, or both. Examination of the temporal trend of the fatal landslides (Fig. 1) reveals that the frequency of very low magnitude landslides (with one or two fatalities) has remained unchanged between 1861 and 2015, whereas the magnitude of the most catastrophic landslides has decreased. Gariano and Guzzetti (2016) have argued that this is the result of a combination of natural (hazard) and societal (vulnerability) causes, and that the reduced trend in the magnitude of the most destructive landslides is due largely to improved monitoring and warning systems, and due to the increased availability of information on landslides and their consequences. However, in an examination of the perception of the Italian population to landslide (and flood) risk, Salvati et al. (2014) have concluded that in most of the country the perception of the threat posed by landslides (and floods) does not match the long-term risk posed by hydrological and geological hazards.

In Fig. 9 we show scenarios of societal landslide risk in Italy, for landslides of four different magnitudes. We note here that calculation of other scenarios for landslides of different magnitudes is possible, and may provide additional information for a more refined zonation of landslide risk to the population of Italy. The validation of the proposed scenarios performed using fatal landslides in the old (t_4 , 1000–1860) and the very recent (t_5 , 2016–2018) periods measured the consistency of our model (Fig. 11). In particular, the eleven recent landslides with fatal consequences in the t_5 period confirmed the ability of the model to predict the expected average return period of fatal landslides in Italy. This opens to the possibility of using the model scenarios to predict the future consequences of fatal landslides in Italy. The rationale behind the projections will be that “the past is a proxy for the future” i.e., that the physical conditions that control landslide hazard and the societal factors that control the size and distribution of the population and their vulnerability to landslides, will not change significantly in the period of the projection. Examinations of the sign and magnitude of the changes is beyond the scope of this work. Here, we only note that Gariano and Guzzetti (2016) have argued that, due to global warming, the frequency and intensity of severe rainfall events, a primary trigger of rapid-moving landslides that cause many fatalities in Italy (Guzzetti et al., 2005b), is expected to increase. Hence, the landslide risk posed by rainfall-induced landslides to the population of Italy is expected to increase.

We note that in NE Italy the model is conditioned by the two largest magnitude landslides in the historical record i.e., the Vajont rockslide that killed 1917, and the Stava mudflow that killed 268 (Fig. 6a). These two highly catastrophic landslides were clearly related to the influence of human interventions and engineering works in close proximity; the presence of an artificial dam and reservoir (Vajont) and of poorly constructed mining-waste embankments (Stava). This has conditioned the model of societal landslide risk in NE Italy (Figs. 6 and 8). However, we argue that several other cases exist in the catalogue of landslides conditioned by human intervention and engineering works. One could argue that with the present state-of-the-art knowledge on landslides, modern monitoring technologies

and improved engineering techniques and methods should now anticipate or prevent similar catastrophic landslides could be anticipated or prevented, and the direct consequences to the population averted or reduced significantly. We argue that this is difficult to quantify, and to consider in a predictive model of societal landslide risk. We further argue that “black swans” (Makridakis and Taleb, 2009; Taleb, 2007) i.e., landslides that cannot be anticipated with the present understanding of landslide phenomena, are possible and should not be ignored when ascertaining landslide risk to the population. Our model does not consider the “physics of the phenomena” (e.g., the geological conditions that may be more (or less) prone to the initiation of landslides, or the meteorological or seismic triggers of the landslides). This may prove an advantage when attempting to predict rare and unexpected fatal events based on the historical record. We stress that the exact position and magnitude of a fatal landslide depend on local conditions and the dynamics of the event (Salvati et al., 2018), which are not considered by our model. We therefore caution that the model outcomes (Figs. 6 and 8) cannot be used to determine landslide risk to individuals at any specific location. For the purpose, more detailed investigations and analyses are necessary (Reichenbach et al., 2004).

Lastly, we acknowledge that our model does not consider all the possible uncertainties that may affect societal landslide risk in Italy. With this respect, a counterfactual analysis (Balke and Pearl, 1994; Pearl, 2000) may provide supplementary information for uncertainty assessment and risk management (Woo, 2018).

7. Conclusions

We proposed an original approach to evaluate the spatial and the temporal distribution of societal landslide risk from historical, sparse, point information on fatal landslides and their consequences. We tested the approach in Italy, a country for which a long and accurate record of historical landslides with fatal consequences is available (Guzzetti et al., 2005b; Salvati et al., 2016, 2013, 2010, 2003). To model societal landslide risk in Italy, we used the portion of the record covering the 155-year period from 1861 to 2015 (Fig. 1), considering only the landslides for which the location of the fatalities was known (Fig. 2). To validate the societal risk model with independent information, we used the portions of the record from 1000 to 1860, and from January 2016 to August 2018 (Fig. 2, Table 1). Despite some known incompleteness in the old part of the record (1000–1860), and the short length of the recent period 2016–2018, the validation confirmed the ability of the approach, and of the resulting societal risk model to anticipate the frequency of fatal landslides of varying magnitudes in Italy.

To model societal landslide risk, for the whole of Italy and for seven physiographic (Guzzetti and Reichenbach, 1994) and 20 administrative subdivisions of Italy, we adopted the Zipf distribution, broadly used in different fields of the natural and the social sciences (Clauset et al., 2009; Newman, 2005; Zipf, 1949). Results confirmed that the Zipf distribution is an adequate and robust descriptor of the magnitude of the fatal landslides in Italy (Fig. 3). We anticipate the distribution to be an adequate model of landslide fatalities elsewhere, and we encourage investigators to adopt it to describe the frequency and the probability of fatal landslides. We foresee that this will facilitate the comparison of societal landslide risk levels in different areas.

The ensemble of the Zipf model curves obtained for Italy and for the subdivisions of Italy (Fig. 3) revealed differences in the geographical distribution of landslide risk. We found that in any given area the differences depended on (i) the number of fatalities caused by the most catastrophic landslide i.e., the largest magnitude landslide, (ii) the total number of fatal landslides, regardless of their magnitude, and (iii) the proportion of low, medium and large magnitude landslides, controlled by the exponent of the Zipf distribution. Maps of the three variables were locally different (Fig. 6) revealing the complexity of landslide risk in Italy. We conclude that societal landslide risk in Italy – and most probably elsewhere in the world – cannot be described by a single metric; and that only a combination of the variables can provide a

reliable representation of societal landslide risk levels (Fig. 8).

We found landslide risk to be particularly high in the NE Alps and in the coastal area of Campania, southern Italy, and high in large parts of the central Alps, in Liguria, in parts of the northern Apennines, and in southern Calabria and NE Sicily (Fig. 8). Instead, we found risk to be low or very low in large parts of central Italy, in most of Puglia, Sicily and Sardinia. Our analysis revealed generalized higher risk levels between 1916 and 1965 than in the more recent period 1966–2015. We conclude that, despite the significant increase in the size of the population, societal landslide risk has decreased in Italy in the recent period. Exceptions exist in Campania, Veneto, Trentino–Alto Adige and Lazio that have experienced higher risk levels in the recent period 1966–2015 (Fig. 11). We further conclude that in these areas landslide risk to the population has increased over the years.

To the best of our knowledge, this is the first work that attempts to anticipate the return period of fatal landslides, in Italy and elsewhere. Our analysis revealed that for very low magnitude landslides (with one or two fatalities), the return period is short (< 30 years) in most of the Alps, in the Alps–Apennines transition zone, in large parts of the Apennines range, of the Tyrrhenian and the Adriatic borderlands, and in NE Sicily. We therefore determine that landslide risk is high in these areas. For large magnitude landslides (with ≥ 25 fatalities), the return period is very long (> 1000 years) in most of the hills and the mountains, with significant exceptions in the central and the eastern Alps, in Campania and in Calabria, that exhibit return periods ≥ 150 years. We suggest that in these areas the risk of experiencing a catastrophic landslide should be considered high (Fig. 9).

Speculations on the causes of the geographical and the temporal variations of landslide risk to the population of Italy are beyond the scope of this work. However, we note that the frequency of the very low magnitude landslides (with one or two fatalities) has remained unchanged almost everywhere, whereas the magnitude of the most catastrophic landslides has decreased over time (Fig. 1). This may be the

result of a combination of natural and societal causes (Gariano and Guzzetti, 2016). The later includes the production of landslide risk zoning and the enforcement of landslide mitigation strategies, the availability of monitoring and warning systems, and better information on landslides and their consequences.

Comparison of our quantitative societal landslide risk assessments (Figs. 3 and 7) against risk acceptability criteria is beyond the scope of this work. However, we stress that public authorities (e.g., civil protection) and private businesses (e.g., insurance and re-insurance companies) involved or interested in risk assessment and management may use the results of this study to enhance their risk management and mitigation strategies.

We conclude stressing that our approach to model societal risk is general, and it can be used to ascertain the societal risk posed by other single location point hazards, provided sufficient information on the time and place of occurrence of the fatal consequences is available. We anticipate that the application of the approach to model other hazards will allow the comparison of societal risk levels posed by different hazards, will contribute to ascertain risk levels where multiple hazards coexist, will facilitate the evaluation of societal risk levels against risk acceptance criteria, and it will contribute to risk management.

Acknowledgments

Our work was partially supported by grants awarded by the Italian National Department for Civil Protection (DPC), the Italian Ministero dell'Istruzione, dell'Università e della Ricerca, Progetti di Rilevante Interesse Nazionale (MIUR-PRIN) 2010–2011, and the Fondazione Assicurazioni Generali (FAG). CB was supported by a grant of the DPC, EN was supported by a grant of MIUR-PRIN, and MD was supported by a grant of FAG. We thank Cees van Westen and another anonymous referee. Their constructive comments have contributed to improve the quality and clearness of the work.

Appendix A. Variables and acronyms used in the text

Variable	Explanation
f	Number of fatalities caused by a landslide [#]
g	Size of model square grid cell [km]
r	Radius of model circular kernel [km]
s	Zipf distribution parameter (scaling exponent) [–]
s_k	Zipf distribution parameter (scaling exponent) in the model kernel [–]
D	2-sided Kolmogorov-Smirnov statistics
E	Number of fatal landslides in the historical record [#]
E_k	Number of fatal landslides inside the modelling kernel [#]
E_{avr}	Average number of fatal landslides in the model kernel [#]
E_{max}	Maximum number of fatal landslides in the model kernel [#]
E_{ya}	Yearly average of fatal landslides in a period [#]
F	Largest number of fatalities caused by a landslide in the record [#]
F_k	Largest number of fatalities caused by a landslide in the model kernel [#]
F_{tot}	Total number of fatalities in a period [#]
F_{ya}	Yearly average number of fatalities in a period [#]
T, t	Length of the historical record, period, or sub-period [yr]
α_s	Variability of the Zipf distribution parameter, s [–]
τ	Return period of fatal landslide [yr]

Acronym	Explanation
CCDF	Complementary Cumulative Distribution Function
CCFDF	Complementary Cumulative Frequency Distribution Function
CDF	Cumulative Distribution Function
FMF	Frequency Mass Function
MLE	Maximum Likelihood Estimation
PMF	Probability Mass Function
yCCFDF	Yearly Complementary Cumulative Frequency Distribution Function
yFDF	Yearly Frequency Mass Function
τ_{yCCFDF}	Return period, for Complementary Cumulative Frequency Distribution Function

Appendix B. Equations used in the work

The Probability Mass Function (PMF) of the Zipf distribution model adopted to evaluate societal landslide risk in Italy is

$$PMF = \frac{1}{f^s \sum_{f=1}^F \frac{1}{f^s}} \quad (B1)$$

where, $f \in \{1, 2, \dots, F\}$ is the number of the landslide fatalities that measures the magnitude of the fatal landslide, F is the largest number of fatalities caused by a single fatal landslide in the record, and $s \in \mathbb{R}^+$ is the scaling exponent that controls the steepness of the Zipf distribution. The PMF gives information on the expected probability of observing a fatal landslide of a given magnitude, f .

The Complementary Cumulative Distribution Function (CCDF), also known as the “survival” function or the “risk curve”, is

$$CCDF = 1 - CDF \quad (B2)$$

where CDF is the Cumulative Distribution Function i.e., the integral of the PMF, $CDF = \int_0^1 PMF$. A monotonically decreasing function, the CCDF gives the exceedance probability of observing more than a given number of fatalities, $z > f$.

The Frequency Mass Function (FMF) is

$$FMF = PMF \times E = \int_0^1 \frac{1}{f^s \sum_{f=1}^F \frac{1}{f^s}} \times E, \quad (B3)$$

where E is the total number of fatal landslides in the record. For a given landslide magnitude f , the FMF gives the expected number (frequency) of fatal landslides of the same magnitude.

The Complementary Cumulative Frequency Distribution Function (CCFDF) is

$$CCFDF = 1 - CFDF, \quad (B4)$$

where CFDF is the Cumulative Frequency Distribution Function, the integral of the FMF, $CFDF = \int_0^1 FMF$. A monotonically decreasing function, for a given landslide magnitude f , the CCFDF gives the exceedance probability of observing more than a given number of landslide fatalities, $z \geq f$ information on the frequency of fatal landslides of equal or larger magnitude, $z \geq f$.

The yearly Complementary Cumulative Frequency Distribution Function (yCCFDF) is

$$yCCFDF = \frac{CCFDF}{T}, \quad (B5)$$

where T is the length of the historical record, in years. A monotonically decreasing function, the yCCFDF gives the annual frequency of fatal landslides of magnitude $f \geq h$.

The return period of a fatal landslide is

$$\tau_{yCCFDF} = \frac{1}{yCCFDF}, \quad (B6)$$

that gives the average recurrence interval between fatal landslides of a given magnitude, f .

References

- Albini, P., Musson, R.M.W., Rovida, A., Locati, M., Gomez Capera, A.A., Viganò, D., 2014. The global earthquake history. *Earthquake Spectra* 30, 607–624. <https://doi.org/10.1193/122013EQS297>.
- Allen, C.D., Macalady, A.K., Chenchouni, H., Bachelet, D., McDowell, N., Vennetier, M., Kitzeberger, T., Rigling, A., Breshers, D.D., Hogg, E.H.T., Gonzalez, P., Fensham, R., Zhang, Z., Castro, J., Demidova, N., Lim, J.-H., Allard, G., Running, S.W., Semerci, A., Cobb, N., 2010. A global overview of drought and heat-induced tree mortality reveals emerging climate change risks for forests. *For. Ecol. Manag.* 259, 660–684. <https://doi.org/10.1016/j.foreco.2009.09.001>.
- Australian Geomechanics Society, Sub-Committee on Landslide Risk Management, 2000. *Landslide risk management concepts and guidelines*. Aust. Geomech. 49–92.
- Badoux, A., Andres, N., Techel, F., Hegg, C., 2016. Natural hazard fatalities in Switzerland from 1946 to 2015. *Nat. Hazards Earth Syst. Sci.* 16, 2747–2768. <https://doi.org/10.5194/nhess-16-2747-2016>.
- Balke, A., Pearl, J., 1994. Counterfactual probabilities: computational methods, bounds and applications. In: *Uncertainty Proceedings 1994*. Elsevier, pp. 46–54. <https://doi.org/10.1016/B978-1-55860-332-5.50011-0>.
- Bohorquez, J.C., Gourley, S., Dixon, A.R., Spagat, M., Johnson, N.F., 2009. Common ecology quantifies human insurgency. *Nature* 462, 911–914. <https://doi.org/10.1038/nature08631>.
- Borden, K.A., Cutter, S.L., 2008. Spatial patterns of natural hazards mortality in the United States. *Int. J. Health Geogr.* 7, 64. <https://doi.org/10.1186/1476-072X-7-64>.
- Boyd, J., Haegeli, P., Abu-Laban, R.B., Shuster, M., Butt, J.C., 2009. Patterns of death among avalanche fatalities: a 21-year review. *CMAJ* 180, 507–512. <https://doi.org/10.1503/cmaj.081327>.
- Brunetti, M., Colacino, M., Maugeri, M., Nanni, T., 2001. Trends in the daily intensity of precipitation in Italy from 1951 to 1996. *Int. J. Climatol.* 21, 299–316. <https://doi.org/10.1002/joc.613>.
- Brunetti, M., Maugeri, M., Monti, F., Nanni, T., 2004. Changes in daily precipitation frequency and distribution in Italy over the last 120 years. *J. Geophys. Res.* 109, D05102. <https://doi.org/10.1029/2003JD004296>.
- Brunetti, M., Maugeri, M., Monti, F., Nanni, T., 2006. Temperature and precipitation variability in Italy in the last two centuries from homogenised instrumental time series. *Int. J. Climatol.* 26, 345–381. <https://doi.org/10.1002/joc.1251>.
- Capparelli, G., Versace, P., 2014. Analysis of landslide triggering conditions in the Sarno area using a physically based model. *Hydrol. Earth Syst. Sci.* 18, 3225–3237. <https://doi.org/10.5194/hess-18-3225-2014>.
- Cascini, L., Cuomo, S., Della Sala, M., 2011. Spatial and temporal occurrence of rainfall-induced shallow landslides of flow type: a case of Sarno-Quindici, Italy. *Geomorphology* 126, 148–158. <https://doi.org/10.1016/j.geomorph.2010.10.038>.
- Chao, A., 1989. Estimating population size for sparse data in capture-recapture experiments. *Biometrics* 45, 427. <https://doi.org/10.2307/2531487>.
- Clauset, A., Shalizi, C.R., Newman, M.E.J., 2009. Power-law distributions in empirical data. *SIAM Rev.* 51, 661–703. <https://doi.org/10.1137/070710111>.
- Cruden, D.M., Varnes, D.J., 1996. *Landslide Types and Processes*. Transportation Research Board, National Academy of Sciences, Washington, D.C.
- Davison, A.C., Hinkley, D.V., 1997. *Bootstrap Methods and their Application*, Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press <https://doi.org/10.1017/CBO9780511802843>.
- Doocy, S., Daniels, A., Packer, C., Dick, A., Kirsch, T.D., 2013. The human impact of earthquakes: a historical review of events 1980–2009 and systematic literature review. *PLoS Currents*. <https://doi.org/10.1371/currents.dis.67bd14fe457f1db0b5433a8ee20fb833>.
- Dowling, C.A., Santi, P.M., 2013. Debris flows and their toll on human life: a global analysis of debris-flow fatalities from 1950 to 2011. *Nat. Hazards*. <https://doi.org/10.1007/s11069-013-0907-4>.
- Efron, B., 1979. Bootstrap methods: another look at the jackknife. *Ann. Stat.* 7, 1–26. https://doi.org/10.1007/978-1-4612-4380-9_41.
- Falucci, A., Maiorano, L., Boitani, L., 2006. Changes in land-use/land-cover patterns in Italy and their implications for biodiversity conservation. *Landsch. Ecol.* 22, 617–631. <https://doi.org/10.1007/s10980-006-9056-4>.
- Fell, R., Harford, D., 1997. *Landslide risk management*. In: Cruden, D.M., Fell, R. (Eds.), *Landslide Risk Assessment*. Balkema, pp. 51–109.
- Fell, R., Ho, K.K.S., Lacasse, S., Leroy, E., 2005. *State of the Art Paper 1 a Framework for Landslide Risk Assessment and Management*. pp. 23.
- Friedman, J.A., 2015. Using power laws to estimate conflict size. *J. Confl. Resolut.* 59, 1216–1241. <https://doi.org/10.1177/0022002714530430>.

- Froude, M.J., Petley, D.N., 2018. Global fatal landslide occurrence from 2004 to 2016. *Nat. Hazards Earth Syst. Sci.* 18, 2161–2181. <https://doi.org/10.5194/nhess-18-2161-2018>.
- Gariano, S.L., Guzzetti, F., 2016. Landslides in a changing climate. *Earth Sci. Rev.* 162, 227–252. <https://doi.org/10.1016/j.earscirev.2016.08.011>.
- Glade, T., Anderson, M.G., Crozier, M.J., 2004. *Landslide Hazard and Risk*. John Wiley & Sons.
- Grahn, T., Jaldell, H., 2017. Assessment of data availability for the development of landslide fatality curves. *Landslides* 14, 1113–1126. <https://doi.org/10.1007/s10346-016-0775-6>.
- Guzzetti, F., 2000. Landslide fatalities and the evaluation of landslide risk in Italy. *Eng. Geol.* 58, 89–107. [https://doi.org/10.1016/S0013-7952\(00\)00047-8](https://doi.org/10.1016/S0013-7952(00)00047-8).
- Guzzetti, F., Reichenbach, P., 1994. Towards a definition of topographic divisions for Italy. *Geomorphology* 11, 57–74. [https://doi.org/10.1016/0169-555x\(94\)90042-6](https://doi.org/10.1016/0169-555x(94)90042-6).
- Guzzetti, F., Salvati, P., Stark, C.P., 2005a. Evaluation of risk the population posed by natural hazards in Italy. In: Hungr, O., Fell, R., Couture, R., Eberhardt, E. (Eds.), *Landslide Risk Management*. Taylor & Francis, pp. 381–389.
- Guzzetti, F., Stark, C.P., Salvati, P., 2005b. 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., 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, 42–66. <https://doi.org/10.1016/j.earscirev.2012.02.001>.
- 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>.
- Ibsen, M.L., Brunsden, D., 1996. The nature, use and problems of historical archives for the temporal occurrence of landslides, with specific reference to the south coast of Britain, Ventnor, Isle of Wight. *Geomorphology* 15, 241–258. [https://doi.org/10.1016/0169-555x\(95\)00073-e](https://doi.org/10.1016/0169-555x(95)00073-e).
- Statistica, Istituto Nazionale di, 2014. *Annuario Statistico Italiano, 2014th ed.* Istituto Nazionale di Statistica.
- Kirschbaum Bach, D., Adler, R.F., Hong, Y., Hill, S., Lerner-Lam, A., 2009. A global landslide catalog for hazard applications: method, results, and limitations. *Nat. Hazards* 52, 561–575. <https://doi.org/10.1007/s11069-009-9401-4>.
- Li, Y., Chen, T., Weng, W., Yuan, H., 2014. A Chinese earthquake database for casualty modelling. In: *Proceedings of the 11th International ISCRAM Conference – University Park, Pennsylvania, USA, May 2014*. Presented at the 11th International ISCRAM Conference, pp. 5.
- Li, W., Liu, C., Hong, Y., Zhang, X., Wan, Z., Saharia, M., Sun, W., Yao, D., Chen, W., Chen, S., Yang, X., Yue, Y., 2016. A public Cloud-based China's Landslide Inventory Database (CsLID): development, zone, and spatiotemporal analysis for significant historical events, 1949–2011. *J. Mt. Sci.* 13, 1275–1285. <https://doi.org/10.1007/s11629-015-3659-7>.
- Lin, Q., Wang, Y., 2018. Spatial and temporal analysis of a fatal landslide inventory in China from 1950 to 2016. *Landslides*. <https://doi.org/10.1007/s10346-018-1037-6>.
- Makridakis, S., Taleb, N.N., 2009. Living in a world of low levels of predictability. *Int. J. Forecast.* 25, 840–844. <https://doi.org/10.1016/j.ijforecast.2009.05.008>.
- Mudelsee, M., Borngen, M., Tetzlaff, G., Grunewald, U., 2003. No Upward Trends in the Occurrence of Extreme Floods in Central Europe. vol. 425. pp. 4.
- Myung, H.-N., Jang, J.-Y., 2011. Causes of death and demographic characteristics of victims of meteorological disasters in Korea from 1990 to 2008. *Environ. Health* 10, 82. <https://doi.org/10.1186/1476-069X-10-82>.
- Newman, M.E.J., 2005. Power laws, Pareto distributions and Zipf's law. *Contemp. Phys.* 46, 323–351. <https://doi.org/10.1080/00107510500052444>.
- Pearl, J., 2000. The logic of counterfactuals in causal inference. *J. Am. Stat. Assoc.* 95, 428–435.
- Pereira, S., Zêzere, J.L., Quaresma, I., Santos, P.P., Santos, M., 2015. Mortality patterns of hydro-geomorphologic disasters. *Risk Anal.* <https://doi.org/10.1111/risa.12516>.
- Petley, D.N., 2012. Global patterns of loss of life from landslides. *Geology* 40, 927–930. <https://doi.org/10.1130/G33217.1>.
- Rappaport, E.N., 2000. Loss of life in the United States associated with recent Atlantic tropical cyclones. *Bull. Am. Meteorol. Soc.* 81, 2065–2073.
- Reed, W.J., 2001. The Pareto, Zipf and other power laws. *Econ. Lett.* 74, 15–19. [https://doi.org/10.1016/S0165-1765\(01\)00524-9](https://doi.org/10.1016/S0165-1765(01)00524-9).
- Reichenbach, P., Galli, M., Cardinali, M., Guzzetti, F., Ardzzone, F., 2004. Geomorphological mapping to assess landslide risk: concepts, methods and applications in the Umbria region of Central Italy. In: Glade, T., Anderson, M.G., Crozier, M.J. (Eds.), *Landslide Hazard and Risk*. John Wiley & Sons, Ltd, pp. 429–468.
- Rossi, M., Witt, A., Guzzetti, F., Malamud, B.D., Peruccacci, S., 2010. Analysis of historical landslide time series in the Emilia-Romagna region, northern Italy. *Earth Surf. Process. Landf.* 35, 1123–1137. <https://doi.org/10.1002/esp.1858>.
- Salvati, P., Guzzetti, F., Reichenbach, P., Cardinali, M., Stark, C.P., 2003. *Map of Landslides and Floods with Human Consequences in Italy*.
- 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., Rossi, M., Guzzetti, F., 2012. Flood risk in Italy. In: Kundzewicz, Z.W. (Ed.), *Changes in Flood Risk in Europe*. IAHS, pp. 277–292.
- Salvati, P., Marchesini, I., Balducci, V., Bianchi, C., Guzzetti, F., 2013. A new digital catalogue of harmful landslides and floods in Italy. In: Margottini, C., Canuti, P., Sassa, K. (Eds.), *Landslide Science and Practice*. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 409–414. https://doi.org/10.1007/978-3-642-31310-3_56.
- Salvati, P., Bianchi, C., Fiorucci, F., Giostrella, P., Marchesini, I., Guzzetti, F., 2014. Perception of flood and landslide risk in Italy: a preliminary analysis. *Nat. Hazards Earth Syst. Sci.* 14, 2589–2603. <https://doi.org/10.5194/nhess-14-2589-2014>.
- Salvati, P., Rossi, M., Bianchi, C., Guzzetti, F., 2016. Landslide risk to the population of Italy and its geographical and temporal variations. In: Chavez, M., Ghil, M., Urrutia-Fucugauchi, J. (Eds.), *Extreme Events: Observations, Modeling, and Economics*, Geophysical Monograph Series. John Wiley & Sons, Inc, Hoboken, NJ, pp. 177–194. <https://doi.org/10.1002/9781119157052.ch14>.
- 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>.
- Shepperd, M., Cartwright, M., 2001. Predicting with sparse data. *IEEE Trans. Softw. Eng.* 27, 12.
- Sidle, R.C., 2006. Hazard Assessment and Prediction Methods, in: *Landslides: Processes, Prediction, and Land Use*. pp. 1–23. <https://doi.org/10.1029/18W006>.
- Simkin, T., Siebert, L., Blong, R., 2001. Volcano fatalities—lessons from the historical record. *Science* 291 (5502), 255. <https://doi.org/10.1126/science.291.5502.255>.
- Spence, R., So, E., Scawthorn, C. (Eds.), 2011. *Human Casualties in Earthquakes*. Springer, Netherlands, Dordrecht. <https://doi.org/10.1007/978-90-481-9455-1>.
- Špitalar, M., Gourley, J.J., Lutoff, C., Kirstetter, P.-E., Brilly, M., Carr, N., 2014. Analysis of flash flood parameters and human impacts in the US from 2006 to 2012. *J. Hydrol.* 519, 863–870. <https://doi.org/10.1016/j.jhydrol.2014.07.004>.
- Stucchi, M., Rovida, A., Gomez Capera, A.A., Alexandre, P., Camelbeek, T., Demircioglu, M.B., Gasperini, P., Kouskou, V., Musson, R.M.W., Radulian, M., Sesetyan, K., Vilanova, S., Baumont, D., Bungum, H., Fäh, D., Lenhardt, W., Makropoulos, K., Martínez Solares, J.M., Scotti, O., Živčić, M., Albini, P., Batlo, J., Papaioannou, C., Tatevossian, R., Locati, M., Meletti, C., Viganò, D., Giardini, D., 2013. The SHARE European Earthquake Catalogue (SHEEC) 1000–1899. *J. Seismol.* 17, 523–544. <https://doi.org/10.1007/s10950-012-9335-2>.
- Taleb, N.N., 2007. Black swans and the domains of statistics. *Am. Stat.* 61, 198–200. <https://doi.org/10.1198/000313007X219996>.
- 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* 139–140, 545–558. <https://doi.org/10.1016/j.geomorph.2011.12.006>.
- Varnes, D.J., *International Association of Engineering Geology, 1984. Landslide Hazard Zonation—A Review of Principles and Practice*. UNESCO, Paris.
- White, E.P., Enquist, B.J., Green, J.L., 2008. On estimating the exponent of power-law frequency distributions. *Ecology* 89, 905–912. <https://doi.org/10.1890/07-1288.1>.
- Witt, A., Malamud, B.D., 2013. Quantification of long-range persistence in geophysical time series: conventional and benchmark-based improvement techniques. *Surv. Geophys.* 34, 541–651. <https://doi.org/10.1007/s10712-012-9217-8>.
- Woo, G., 2018. Counterfactual disaster risk analysis. *Variance* 10 (2), 279–291.
- Zipf, G.K., 1949. *Human Behavior and the Principle of Least Effort: An Introduction to Human Ecology*. Addison-Wesley Press, Cambridge, MA. <https://doi.org/10.2307/2572028>.