Comparison between two methodologies for assessing historical earthquake parameters and their impact on seismicity rates In the Western Alps

Ludmila Provost, Andrea Antonucci, Andrea Rovida, Oona Scotti

COMPARISON BETWEEN TWO METHODOLOGIES FOR ASSESSING HISTORICAL EARTHQUAKE PARAMETERS AND THEIR IMPACT ON SEISMICITY RATES IN THE WESTERN ALPS.

Ludmila Provost, Institut de Radioprotection et Sûreté Nucléaire (IRSN), PSE-ENV, SCAN, BERSSIN, Fontenay-aux-Roses, 92262, France

Andrea Antonucci, Istituto Nazionale di Geofisica e Vulcanologia, Milano, 20133, Italy

Andrea Rovida, Istituto Nazionale di Geofisica e Vulcanologia, Milano, 20133, Italy

Oona Scotti, Institut de Radioprotection et Sûreté Nucléaire (IRSN), PSE-ENV, SCAN, BERSSIN, Fontenay-aux-Roses, 92262, France

Corresponding author: Ludmila Provost, ludmila.provost@irsn.fr

ACKNOWLEDGMENTS

The authors would like to thank Francesco Visini for introducing to the Albarello et al (2001) algorithm used to compute completeness times. Part of the work presented in this paper was done during the lock-down due to the pandemic situation, the authors would like to thank their respective lock-down companions for their patience during our weekly online meetings. The authors would like to thank the two anonymous reviewers for their useful suggestions to improve the article.

ABSTRACT

We investigate the differences in seismicity rate estimates from two historical earthquake catalogues obtained with two methodologies (Boxer and QUake-MD) calibrated on a common dataset of macroseismic intensities and calibration events. The two methodologies were then applied to a test data set of historical earthquakes covering the France, Italy and Switzerland Alpine region. Differences between the resulting magnitude estimates and instrumental magnitudes show a standard deviation of 0.4 for both methodologies, with a mean residual of 0.01 for Boxer and -0.04 for Quake-MD. A systematic difference in magnitude estimates between the two methodologies that correlates with the depth estimated by Quake-MD has been observed. This is attributed to the difference in the treatment of the depth parameter between Boxer and QUake-MD. Nevertheless, differences in magnitude estimates between the two methodologies show a mean residual of 0.006 and a standard deviation of 0.35 resulting in seismicity rates that are not significantly different considering the associated uncertainties. Such results made us believe that the European community could gain in the reduction of epistemic uncertainties associated with the estimate of historical earthquake parameters by agreeing on a common macroseismic and

calibration dataset across borders. These efforts should be strongly encouraged. On the other hand, we show that even in the ideal conditions of this benchmark (same calibration events and same macroseismic intensity dataset), methodological differences can lead to systematic differences in magnitude estimates. It is therefore paramount to explore different methodologies for a more realistic quantification of the epistemic uncertainties in estimates of maximum magnitudes and seismic activity rates.

KEYWORDS

Historical earthquake, parametric earthquake catalogue, seismic activity rate, Western Alps

DECLARATIONS

This work was funded by IRSN and INGV. The authors do not have any conflict of interest. Data used in this paper are available in the Electronical supplements, and/or from the websites of AHEAD (https://www.emidius.eu/AHEAD/) and CPTI15 (https://emidius.mi.ingv.it/CPTI15-DBMI15/). The two codes used for deriving earthquake parameters from intensity data are available from: https://emidius.mi.ingv.it/boxer/; https://github.com/ludmilaprvst/QUake-MD. The third code Matlab_PSHA, used to compute completeness times and annual rates, will be published soon.

Figures were made with Matlplotlib (J. D. Hunter, "Matplotlib: A 2D Graphics Environment", Computing in Science & Engineering, vol. 9, no. 3, pp. 90-95, 2007.) and maps with QGIS.

INTRODUCTION

Researches in many fields, such as seismology, earthquake geology, tectonics, and engineering seismology require the knowledge of the seismicity of the target area to be as long as possible. To this purpose, earthquake parameters determined from intensity data are used to extend the coverage of instrumental earthquake catalogues back to historical times. In Europe, the long history, the type and amount of written documents, and the extensive historical seismological research conducted in the last 50 years translated into one of the world's longest and most detailed record of effects of past earthquakes. This knowledge is represented by several earthquake catalogues (for examples see the AHEAD database; Locati *et al.*, 2014; Rovida and Locati, 2015 and databases at different geographical scales, as well as a wealth of studies on single earthquakes, periods, or areas published over the years.

However, the recent analysis in Rovida *et al.* (2020a) highlighted a high fragmentation of repositories of data and a lack of homogeneity in the knowledge of European pre-instrumental seismicity. In particular, the authors show that the geographical distribution of earthquakes with published sets of intensity data is not uniform across Europe, and in some areas, especially in the north and east, historical earthquakes are known only through parametric catalogs, independently of the earthquake size and period of occurrence. Such a fragmentation results from the national perspective adopted in the last decades for the compilation of earthquake databases, with national earthquake catalogues relying on their own input intensity data and their own methodology to determine earthquake locations and magnitudes.

A step forward in the harmonization of the knowledge of European past seismicity is represented by the European Archive of Historical Earthquake Data AHEAD (Locati *et al.*, 2014; Rovida and Locati, 2015). It collects the multiplicity of earthquake data of different types and creates relations among them with the aim of making the knowledge of each European earthquake in the period 1000-1899 promptly available. Providing the conclusive macroseismic intensity distribution for every European pre-instrumental earthquake would represent an idealistic ambition, and AHEAD reveals the discrepancies existing in the intensity distributions of earthquakes at country borders, and eases their comparison and the identification of the most representative of the knowledge of each earthquake. Data provided in AHEAD were fully exploited for the compilation of the SHARE European Earthquake Catalogue (SHEEC) 1000-1899 (Stucchi *et al.*, 2013) the input catalogue for the European Seismic Hazard Model ESHM13 (Woessner *et al.*, 2015), and for its update called EPICA – European PreInstrumental earthquake Catalogue (Rovida and Antonucci, 2021).

Once a unique base of input data, as harmonized as possible, is built, the compilation of a uniform European earthquake catalogue encounters the problem of assessing robust parameters from it. Several methods for the determination of earthquake parameters from macroseismic information have been proposed since the early ages of modern seismology (see Cecic *et al.*, 1996; and the introduction to Gasperini et al., 2010), from the shape and size of hand-drawn isoseismals, to the analysis and inversion of the spatial distribution of intensity data (e.g. Bakun and Wentworth, 1997; Gasperini et al., 1999; 2010; Musson and Jimenéz, 2008 Provost and Scotti, 2020). However, Bakun *et al.*, (2011) showed some limitations in the most recent and widely used of these methods, each with its own pros and cons, that make the selection of the most reliable solution difficult and not straightforward (Stucchi *et al.*, 2013).

In Europe, one of the most evident examples of the variety of data and methods is at the shared borders of Switzerland, France and Italy, where complete and internally consistent macroseismic intensity databases generated the three most advanced historical earthquake catalogues in Europe, respectively ECOS-09 (Fäh *et al.*, 2011), F-CAT17 (Manchuel *et al.*, 2018), and CPTI15 (Rovida *et al.*, 2020b). However, being based on different macroseismic databases, and different procedures to assess earthquake parameters they usually provide inconsistent solutions for the very same earthquakes. The consequence of these discrepancies affects all the cross-border elaborations based on these historical catalogues. Among such elaborations, the definition of seismic activity rates, one of the key components of seismic hazard assessment, is particularly influenced by the reliability of the magnitude estimates, as demonstrated in several papers (e.g. Rong *et al.*, 2011; Musson, 2012; Mucciarelli, 2014; Beauval *et al.*, 2020). However, none of these works fully explored the effect of the different approaches used for assessing magnitude from macroseismic data.

This work aims at first exploring the differences in the macroseismic magnitudes obtained with two alternative methods, Boxer (Gasperini *et al.*, 1999, 2010) and QUake-MD (Provost and Scotti, 2020) to the intensity distributions of pre-instrumental earthquakes at the border between France and Italy (Western Alps region, Figure 1). Then, the annual seismic rates calculated from the two different sets of magnitudes are compared in order to quantify the effect of the alternative magnitude determinations.

1. SELECTED METHODOLOGIES

The methodologies selected in this study are Boxer (Gasperini *et al.*, 1999, 2010) and QUake-MD (Provost and Scotti, 2020). The two methodologies have a downloadable associated software (https://emidius.mi.ingv.it/boxer/; https://github.com/ludmilaprvst/QUake-MD, respectively) making them good candidates for a comparison study in terms of reproducibility.

Boxer incorporates different strategies for jointly determining epicenter location, epicentral intensity I_0, and magnitude M from intensity data. The "Method 0" (Gasperini *et al.*, 2010) assumes the epicenter as the barycenter of the points with the maximum observed intensities and calculates it as the trimmed (between the 20^{th} and 80^{th} percentiles) mean of their coordinates. The epicentral intensity I_0 is assumed equal to the maximum observed value Imax, or to Imax – 1 if less than three Imax values are present.

The earthquake magnitude is computed as the weighted average of the values independently obtained from each *i*-th intensity class through the equation by (Sibol *et al.*, 1987):

$$M(I) = a(I) + b(I)I_0 + c(I)\left[\log(\pi D_{epi})\right]^2 \qquad \text{Equation 1}$$

where M is magnitude, D_{epi} is the average distance of points with observed intensity I to the calculated epicenter, I_0 is the epicentral intensity, and a(I), b(I), and c(I) are empirical coefficients. The weight is inversely proportional to the number of intensity data and to the square of the standard deviation of the regression for the corresponding intensity class. The explicit computation of depth is excluded from Boxer, and, according to Sibol et al. (1987) and Galanopulos, 1961, the effect of the source depth on the magnitude determination is taken into account by means of the epicentral intensity term in Equation 1.

QUake-MD, acronym for Quantifying Uncertainties for earthquakes' Magnitude and Depth, quantifies uncertainties in magnitude/depth estimates for earthquakes known only by their macroseismic distribution by taking into account the quality of intensity data and the epistemic uncertainties of intensity prediction equations (IPE). Intensity data quality, when available, is used to weight the intensity data points (IDP) in the application of the intensity prediction equations for the magnitude/depth inversion. Intensity data quality is also used to compute uncertainties of the inverted depth and magnitude. IPE epistemic uncertainties are taken into account by the use of different IPEs calibrated for the target region (see Provost and Scotti 2020 for an example). Uncertainties associated to the inverted depth and magnitude combined with the use of different IPEs can be used to build a probability density function of the plausible depth, magnitude and epicentral intensity associated to the considered earthquake. In this exercise, we used the barycenter of the probability density function to compare with other depth and magnitude solutions, i.e. Boxer’s outputs or instrumental solutions. QUake-MD IPEs have the following mathematical formulation for describing the attenuation of intensity with epicentral distance:

$$I = C_1 + C_2 M + \beta \log\left(\sqrt{{D_{epi}}^2 + H^2}\right) + \gamma \sqrt{{D_{epi}}^2 + H^2} \qquad \text{Equation 2}$$

where H is the hypocentral depth, C1 and C2 are the magnitude coefficients, β the geometrical attenuation coefficient and γ the intrinsic attenuation coefficient. In QUake-MD, epicentral location is an input. For the benchmark exercise, Boxer’s epicentral locations were taken as input.

2. CALIBRATION OF METHODOLOGIES ON THE SAME DATASET

2.1 DEFINITION OF A COMMON DATASET

The empirical coefficients of the relations among intensity, magnitude, and depth used in both Boxer and QUake-MD (Equations 1 and 2) need to be calibrated inverting known data of well-assessed recent earthquakes. The reliability of the calibration of both QUake-MD and Boxer depends on the characteristics of the selected calibration dataset. In general, a dataset for calibration purposes must first include earthquakes with reliable instrumental assessments of the magnitude, each with a consistent and plentiful set of intensity data. Such earthquakes must belong to the same tectonic context in order to reflect the same (or similar) propagation and attenuation features of the seismic waves. The dataset should consist of a minimum of 20-30 earthquakes in order to sample the widest possible range of both site intensity values and magnitudes (and epicentral intensities in the case of Boxer). For calibrating the equation used in QUake-MD, a reliable estimate of the focal depth, with associated uncertainty is also necessary.

The above criteria are not always easily satisfied, especially in areas of low or moderate seismicity (Bakun and Scotti, 2006; Gomez-Capera *et al.*, 2015) such as the Western Alps.

To match all the requirements, we built a calibration dataset of 28 earthquakes (Table 1) selected from the macroseismic databases of France (SisFrance/BRGM-EDF-IRSN 2016; Jomard *et al.*, 2021), Switzerland (ECOS-09; Fäh *et al.*, 2011), and Italy (DBMI15; Locati *et al.*, 2019). In case of multiple datasets for the same earthquake, we selected that with highest number of IDPs and with the most complete spatial coverage. To consider as many earthquakes as possible, including large ones, we took into account also earthquakes in the eastern Alps, i.e. outside the study area but within the same seismotectonic context, and those with early although reliable instrumental magnitude measurements from the 1950s and 1960s. Lacking any harmonized instrumental catalogue for the studied border area, magnitudes and depths are selected from different sources available in the literature, as shown in table 1, and include both native moment magnitudes derived from moment tensor solutions, and proxy Mws derived from magnitude estimates in other scales. The magnitude of the selected earthquakes ranges between 3.3 and 6.5 (figure 1, table 1).

In the QUake-MD calibration process a-priori depth constraints are explored between a minimum and maximum value given for each event (Hmin and Hmax in table 1) based on either literature or on the statistical analysis in Visini et al (2021). A default value of 5 km is associated to events based on LDG estimates. A minimum lower depth limit is fixed at 1 km and a minimum value of 2 km depth uncertainty is assumed. For the application of QUake-MD, depth inversion limits are fixed between 1 and 21 km.

#	Date	Area	NObs	Ix	Ref	Lat	Lon	I0	Mw	Mw Std	Magnitude reference	Conv.	H	Hmin	Hmax	Depth reference
1	1936-10-18	Alpago Cansiglio	250	9	DBMI15	46,089	12,380	9	5,84	0,14	CPTI15	Yes	13	10	16	(Sandron *et al.*, 2014)
2	1946-01-25	Ayent	243	8	ECOS-09	46,350	7,400	8	5,80	0,20	(Bernardi *et al.*, 2005)	Yes	10	1	21	Visini et al., 2021
3	1946-05-30	Ayent	96	7	ECOS-09	46,300	7,420	7	5,50	0,20	(Bernardi *et al.*, 2005)	Yes	10	1	21	Visini et al., 2021
4	1954-05-19	Mayens	80	6	ECOS-09	46,280	7,310	6	5,30	0,43	(Bernardi *et al.*, 2005)	Yes	10	5	15	ECOS-09
5	1959-04-05	Alpes Provencales	207	7-8	SisFrance	44,533	6,817	7-8	5,52	0,13	CPTI15	Yes	10	1	21	Visini et al., 2021
6	1960-03-23	Brig	307	7	ECOS-09	46,370	8,020	8	5,00	0,20	(Bernardi *et al.*, 2005)	Yes	5	1	10	ECOS-09
7	1962-04-25	Vercors	506	7-8	SisFrance	45,033	5,567	7-8	4,98	0,50	(Gomez-Capera *et al.*, 2015)	Yes	12	7	17	LDG
8	1963-04-25	Vercors	156	7	SisFrance	44,933	5,667	7	4,7	0,19	(Cara *et al.*, 2015)	No	13	8	18	LDG
9	1964-03-14	Alpnach	362	7	ECOS-09	46,870	8,320	7	5,30	0,30	(Bernardi *et al.*, 2005)	Yes	5	1	10	ECOS-09
10	1971-09-29	Vorstegstock	295	6	ECOS-09	46,900	9,010	6	4,90	0,15	(Bernardi *et al.*, 2005)	No	10	1	21	Visini et al., 2021
11	1976-05-06	Friuli	749	9-10	DBMI15	46,241	13,119	9-10	6,45	0,10	CPTI15	No	5,7	3	8	(Slejko *et al.*, 1999)
12	1976-12-13	Garda Settentrionale	128	7	ECOS-09	45,894	10,799	7	4,59	0,23	CPTI15	Yes	5,9	1	15	ISC Bulletin
13	1984-04-17	Vercors	90	5-6	SisFrance	44,983	5,167	5-6	4,00	0,19	(Cara *et al.*, 2015)	No	2	1	7	LDG
14	1988-02-01	Friuli	273	6	DBMI15	46,348	13,076	6	4,82	0,17	CPTI15	Yes	5,1	3	7	OGS Bulletin
15	1989-09-13	Prealpi Vicentine	779	6-7	DBMI15	45,882	11,264	6-7	4,85	0,10	CPTI15	No	9	7	11	OGS Bulletin
16	1990-02-11	Torinese	201	6	DBMI15	44,918	7,558	6	4,56	0,15	CPTI15	Yes	24	20	28	CSI1.1
17	1991-11-20	Vaz	374	6	ECOS-09	46,730	9,530	6	4,60	0,30	(Bernardi *et al.*, 2005)	No	6	3	9	ECOS-09

18	1994-12-14	Genevois	522	6	SisFrance	45,917	6,367	6	4,28	0,07	(Bernardi *et al.*, 2005; Braunmiller *et al.*, 2005)	No	7	4	10	GRN
19	1996-02-27	Dolomiti Friulane	150	5-6	DBMI15	46,309	12,577	5-6	4,31	0,10	CPTI15	Yes	9,5	7	12	OGS Bulletin
20	1996-04-13	Dolomiti Friulane	164	5-6	DBMI15	46,312	12,559	5-6	4,45	0,12	CPTI15	Yes	11,5	9	14	OGS Bulletin
21	1996-07-15	Avant-Pays Savoyard	782	7	SisFrance	45,917	6,083	7	4,62	0,09	(Braunmiller *et al.*, 2005)	No	3	1	4	GRN
22	1997-10-31	Prealpes De Digne	90	6	SisFrance	44,233	6,467	6	4,27	0,18	RCMT	No	6	3	9	GRN
23	1999-01-11	Drac	98	5-6	SisFrance	45,067	5,733	5-6	3,65	0,19	(Cara *et al.*, 2015)	No	0	1	3	GRN
24	1999-02-14	Fribourg	123	5	ECOS-09	46,780	7,210	5	3,93	0,22	(Bernardi *et al.*, 2005; Braunmiller *et al.*, 2005)	No	2	1	5	ECOS-09
25	1999-09-13	Bas-Plateaux Dauphinois	89	4-5	SisFrance	45,450	5,417	4-5	3,30	0,15	SED-TDMT	No	4	1	7	GRN
26	2001-07-17	Val Venosta	657	6	DBMI15	46,697	11,074	6	4,78	0,07	CPTI15	No	1	1	12	CSI1.1
27	2004-11-24	Garda Occidentale	176	7-8	DBMI15	45,685	10,521	7-8	4,99	0,07	CPTI15	No	5,4	3	8	INGV Bulletin
28	2011-07-25	Torinese	105	5-6	DBMI15	45,016	7,365	5-6	4,67	0,07	CPTI15	No	10	1	21	Visini *et al.*, 2021

Table 1: List of the earthquake and associated parameters of the calibration dataset. Nobs is the number of intensity data points, Ix is the maximum intensity. The column Ref contains the macroseismic database for the intensity data points. The column Conv. indicates if the magnitude in column Mw is a converted magnitude (Yes) or a native one (No). H is the depth of the earthquakes and Hmin and Hmax are the constraints used in QUake-MD calibration. The column Depth reference gives the reference associated to column H. CPTI15 stands for Catalogo Parametrico dei Terremoti Italiani 2015 (Rovida et al., 2020b) , ISC for International Seismological Centre (Storchak *et al.*, 2013)**, LDG for Laboratoire de Geophysique** (Duverger *et al.*, 2021)**, OGS for Istituto Nazionale di Oceanografia e di Geofisica Sperimental (**http://www.crs.inogs.it/bollettino/RSFVG/**), GRN for l'Observatoire des Sciences de l'Univers de Grenoble (https://www.osug.fr/), CSI1.1 for Castello et al., 2006, INGV for Istituto Nazionale di Geofisica e Vulcanologia (http://terremoti.ingv.it/)**

,

ECOS-09 for Earthquake Catalogue of Switzerland (Fäh et al., 2011) and DBMI15 for Database Macrosismico Italiano (Locati *et al.*, 2019)**.**

In the definition of the calibration dataset we discarded intensities not expressed as numerical values (e.g. "felt" data) and those with intensity lower than 2, while uncertain intensities (e.g. 5-6) are considered as independent classes and assumed as half degrees (e.g. 5.5). Although expressed in different macroseismic scales, for our purposes they can be considered as equivalent, because conversions between scales introduce much higher uncertainties than the differences between the application of the two scales (e.g. Musson *et al.*, 2010; Allen *et al.*, 2012).

QUake-MD is sensitive to the completeness of the data belonging to each intensity class, which we evaluated comparing the population of each class with the adjacent ones. Using this method, we defined an intensity of completeness for each earthquake (available in Online Resource 1): intensity classes with values smaller than this intensity of completeness are considered as not complete and then not used in QUake-MD calibration. For Boxer we instead used all the data in each class because the number of "complete" IDPs in some intensity classes is not sufficient for calibrating that intensity class, and the calibrations resulting from the complete and the whole datasets are similar. The intensity datasets compiled are made of 8052 data for the whole dataset and 5667 for the "complete" one as shown in figure 2.

2.2 CALIBRATION PROCESS

For QUake-MD, we divided the calibration process into two steps, and we assumed γ equal to zero in Equation 2. The first step calibrates the attenuation coefficient β and the second step estimates the magnitude coefficients C1 and C2. The attenuation coefficient β is calibrated with the (Kövesligethy, 1907) mathematical formula:

$$I = I_0 + \beta \log\left(\sqrt{{D_{epi}}^2 + H^2}\right) \qquad \text{Equation 3.}$$

In order to represent the epistemic uncertainty linked to the choice of the dataset, the attenuation is also calibrated with subsets of the calibration dataset as well as with the entire calibration dataset. The subsets were selected according the number of IDP (more than 20 IDP, 100 IDP or 200 IDP), the date of the earthquake (all the earthquakes or only those after 1980), the number of intensity classes (at least 2, 3 or 4 classes), the difference between the epicentral intensity and the intensity of completeness (difference greater than 1, 2 or 3) or the distance of completeness (without this criteria or intensity distributions complete for the first 25 km, i.e. without abrupt changes in data density with distance). Intensity and distance of completeness are defined by expert opinion, based on the change in the slope of the intensity decay with distance and are provided in the Online Resource 1. The calibration with the whole dataset and its subsets results in a distribution of attenuation coefficient values. In the second step, the magnitude coefficients are calibrated with Equation 2. No calibration subsets were built for this step. The magnitude coefficients are calibrated for each attenuation coefficient value obtained in the first step. Three weighting schemes are used in this step: a uniform weighting scheme, a magnitude standard deviation weighting scheme and a magnitude class weighting scheme (see Online Resource 2 for more details about the three weighting schemes). Depth is sequentially inverted in the two steps, within the depth limits defined in Table 1, with the depth value given for each earthquake +/- the associated uncertainties (see Online Resource 2 for more details about depth limits and uncertainties).

The Boxer code incorporates a sub-routine to calibrate the coefficients of Equation 1. As described in Appendix 1 to Gasperini *et al.*, 2010, the average epicentral distances, trimmed between the 20th and 80th percentiles, of the points in a given intensity class for a given earthquake are calculated first. Classes with less than 4 data and those with intensity greater than the epicentral intensity, calculated by Boxer itself, are excluded from the procedure. The coefficients are then determined by fitting the equation separately for each intensity class and for each earthquake through the minimum weighted sum of the squares. The weights are: i) directly proportional to the number of data points in each intensity class, normalized with the total number of earthquakes in the calibration set that contain intensity of the considered class; and ii) inversely proportional to the square of the standard deviation of the input instrumental magnitude of the corresponding earthquake.

The described procedure applied to the calibration dataset resulted in the determination of the coefficient for the intensity classes from 2 to 7, including intermediate degrees (e.g. 5-6).

Both Boxer's and QUake-MD calibrated coefficients are available in the Online Resource3 and Online Resource 4 respectively.

The reliability of the obtained calibrations of both methods for magnitude estimates is illustrated by the magnitude residual analysis shown in figure 3. The differences between instrumental and macroseismic magnitude are within ±0.5 Mw units, with an average of 0.03 for Boxer and 0.05 for QUake-MD. Instrumental Mw smaller than 4 tend to be overestimated by Boxer, although there are only 3earthquakes of this magnitude. Interestingly, these 3 earthquakes (#23, #24 and #25 in Table 1) are all superficial with depth between 0 km and 4 km. Outputs of Boxer and QUake-MD are available in Online Resource 5.

As QUake-MD estimates depth along with magnitude, the reliability of QUake-MD calibration is checked for both magnitude and depth in the next section.

2.3 TESTING THE CALIBRATED MODELS

2.3.1 TEST DATASET

To further test the performance of the obtained calibrations of both QUake-MD and Boxer we applied both methods to an independent test dataset. The dataset was built starting from the same sources of data (i.e. the SisFrance, ECOS-09 and DBMI15 macroseismic databases) and with the same criteria as the calibration dataset, although with a longer time coverage and a majority of converted Mw instead of true ones (Online resource 6). As a whole the test dataset is made of 102 earthquakes in the Alps covering the period 1905 to 2014, with Mw values ranging between 2.7 and 6.1 and 9396 IDPs with maximum intensity from 3 to 9 (figure 4). The test dataset is available in Online resource 6.

2.3.2 COMPARISON OF THE MAGNITUDE ESTIMATES FOR THE TEST DATASET

We compared first the residual between instrumental magnitude and macroseismic magnitudes (figures 5a and 5b). Most magnitude residuals for both methodologies are in the +/- 0.5 range. Standard deviation of the magnitude residuals for both methodologies is equal to 0.4. Boxer's mean magnitude residual is equal to -0.01 and QUake-MD's mean magnitude residual is equal to -0.04. We observe a systematic overestimation of Boxer's and QUake-MD's magnitudes for instrumental magnitudes smaller than 4.0, which exceeds 0.5 for magnitudes smaller than 3.5. Magnitudes greater than 5.0 seem underestimated, although the small number of earthquakes with this magnitude and especially greater than 5.5, does not allow us to conclude on a systematic underestimation.

Events with residuals exceeding 0.5 are related to four small magnitude events Mw< 3.3 for which predictions are extrapolated beyond magnitude used for the calibration or to a specific earthquake for which the instrumental estimate may be overestimated (1920 with magnitude 5.43 Mw). Indeed the instrumental magnitudes reported before the deployment of the World Wide Standardized Seismic Network - WWSSN (i.e., 1960-1964) can be overestimated (Vannucci *et al.*, 2021).

Both methodologies present the same bias with magnitude which seems to follow a linear correlation with instrumental magnitudes, i.e. an overestimation of small magnitudes and an underestimation of high magnitudes, with a pivot magnitude equal to 4.5. To confirm this linear correlation we computed the Pearson correlation coefficient for both QUake-MD's and Boxer's magnitude residuals. Pearson correlation coefficient in absolute value can vary from 0 to 1 depending on the strength of the linear relationship between two datasets. An absolute value of the correlation coefficient smaller than 0.5 indicates a weak linear correlation between the two datasets. The Pearson coefficient value for the magnitude of the whole test dataset is equal to 0.6 and 0.8 for QUake-MD and Boxer respectively (violet dashed line in figure 5c and 5d), thus a linear correlation exists between instrumental magnitude and residual magnitude for both methodologies. We then computed the Pearson coefficient for different sliding magnitude window of width equal to 1, from magnitude 2.5 to magnitude 5.5 (see figures 5c and 5d). QUake-MD's residuals present a weak linear correlation with the instrumental magnitude for the windows between 3.5 and 5.5, with Pearson coefficient values close to 0.2 for the windows between 4.0 and 5.5 (0.6 for Boxer for the same magnitude window), whereas Boxer's residuals present a weak linear correlation for the 4.5-5.5 magnitude window.

As a conclusion, both methodologies have the same bias with instrumental magnitude: small magnitudes are overestimated and high magnitudes seem underestimated with a pivot point around Mw 4.5. The linear correlation is weaker around the pivot magnitude equal to 4.5, especially for QUake-MD. The bias does not seem to have a link with the conversion of the instrumental magnitudes to Mw: the residual magnitudes for both methodologies with native Mw follow the same pattern as the converted ones.

The comparison of QUake-MD's and Boxer's magnitudes shows (figure 6) that the differences of most magnitudes are smaller than 0.5, with a mean residual between QUake-MD's and Boxer's magnitudes equal to 0.3. Interestingly, there is a systematic difference of magnitude between the two methods due to the inclusion of depth in the QUake-MD inversion scheme which leads to greater magnitudes when the associated depth is deeper than 8 km and lower magnitudes when the associated depth is shallower than 8 km.

This 8 km value is close to the inverted average depth of 7.77 km computed by Pasolini *et al.* (2008) for Italy based on the same mathematical formulation as Equation 2 used in QUake-MD in this study.

2.3.3 COMPARISON OF THE DEPTH ESTIMATES FOR THE TEST DATASET FOR QUAKE-MD

Figure 7 shows the result of the calibration in terms of depth for QUake-MD for the calibration dataset (figure 7a) and the test dataset (figure 7b). On a total number of 28 earthquakes of the calibration dataset, 19 earthquakes, i.e.

67 %, and 66 earthquakes on a total number of 87 earthquakes of the test dataset, i.e 75%, are accurately estimated within an error of +/-5 km. Combined with the good results in magnitude estimates presented in the paper, we consider that the QUake-MD's calibration is satisfactory for the exercise presented in the paper.

The largest depth residuals can be explained by a combination of a poorly known instrumental depth (associated uncertainty up to 11 km) and of a poor macroseismic field quality (lack of data). In those cases, it is quite difficult to know what the actual depth of the earthquake is. A detailed analysis of the two calibration earthquakes presenting the largest residual is available in Online Resource 2.

3. APPLICATION OF THE TWO METHODOLOGIES: IMPACT ON ANNUAL SEISMICITY RATES

3.1 PRESENTATION OF THE DATASET WITHIN THE STUDY AREA

In this section we explore the impact of the methodological differences in the estimate of macroseimic magnitudes on the estimate of earthquake rates.

Focusing on the study area, in the Western Alps (see figure 8), we built two catalogues that share the same instrumental parameters for the period 1981-2015 and same input intensity data for earthquakes in the period 1000-1980, but are processed with Boxer and QUake-MD separately. The earthquakes within the study area are mapped in the figure 8. The catalogues with the earthquakes in the period 1000-1980 processed with Boxer and QUake-MD are called hereafter Boxer's catalogue and QUake-MD's catalogue, respectively.

Macroseismic data were selected among those provided in AHEAD for the period 1000-1899, and for the period 1900-1980 in its Italian counterpart - the Italian Archive of Historical Earthquake Data ASMI (https://doi.org/10.13127/asmi; Rovida *et al.*, 2017), that covers the study area. For each earthquake we selected the dataset with the largest number of data and derived from the most complete historical investigation (see Online Resource 7). For the instrumental part we used the instrumental catalogue at the basis of CPTI15, which consists of instrumental locations from Italian instrumental catalogues selected according to a temporal priority scheme and Mw following the procedures in Lolli *et al.*, 2020.

The selected events consist of 232 earthquakes, 189 of which are in the period 1000-1980. The magnitude of the latter was determined with both QUake-MD and Boxer and we kept in the catalogue 65 events after applying the following criteria to ensure the reliability of the results: Boxer magnitudes calculated with all IDPs (not from epicentral intensity) and only earthquakes with more than 3 IDP. Following these criteria, the older event in the output catalogues occurred in 1564. The catalogue, with both Boxer and QUake-MD estimates, is available in Online Resource 7.

3.2 MAGNITUDE COMPARISON

Boxer’s and QUake-MD's magnitude estimates are in good agreement (see figure 9), with a mean residual equal to 0.006 and a standard deviation equal to 0.35. Except for three events, differences between the two methodologies in magnitude estimates are less than 0.5 units. The same difference between the two methods due to the inclusion of depth in QUake-MD inversion scheme as in figure 6 is observed: QUake-MD magnitudes greater than Boxer’s are associated to depth estimates greater than 8 km, and smaller QUake-MD magnitudes are associated to depth shallower than 8 km. Thus, depending on whether the region of study is dominated by shallow or deep events, the use of one method or another may lead to important differences in the estimate of the observed maximum magnitude.

3.3 ANNUAL SEISMIC ACTIVITY RATES COMPARISON

In order to compute comparable annual seismic activity rates from the Boxer’s and QUake-MD catalogues, we applied the same declustering and completeness estimate methods to both catalogues. The space-time windowing declustering technique of Knopoff and Gardner, 1972 and Gardner and Knopoff, 1974, with window’s length modified according to Burkhard and Grünthal (2009) as implemented in the last version of the ZMap software (Wiemer, 2001) , identifies the same clusters for both catalogues. However, the main event of each cluster was not necessarily the same for both catalogues. Nevertheless, the magnitudes associated to the main events for both catalogue in this study area were close with a mean difference of 0.1. These differences did not affect the following steps of earthquake rates estimates, because the main events of each cluster were classified in the same magnitude class.

In order to apply the same methodology to both catalogues, we used the algorithm of Albarello *et al.*, 2001. The method tests differences in occurrence rates of events in different time intervals, using a binomially distributed random variable. The completeness probability is then calculated using a conditional probability. We chose a minimum magnitude of completeness at Mw 3.5 and considered a magnitude bin width of 0.5.

Mean and quantiles of the obtained distribution of completeness times for each catalogue (figure 10) differ mainly for magnitudes greater than 4.5 due to the reduced number of earthquakes in this magnitude range and their difference in higher magnitude estimates.

Finally, we estimated earthquake rates applying the same completeness times to both catalogues corresponding to the most recent completeness time of both distributions for each bin. In this way, each catalog should be theoretically complete. We used a minimum magnitude of 3.5 and a magnitude bin width of 0.25.

The incremental annual rates for Boxer's and QUake-MD's catalogues are presented figure 11a. Not surprisingly, the annual rates of the two first magnitude bins are identical: according to the completeness time (equal to 1976) only instrumental earthquakes are in these bins. The two following magnitude bins, in which both instrumental and historical earthquakes are present, similar annual rates are estimated for both catalogues. The annual rates for magnitudes higher than 4.5 are quite different.

The small number of earthquakes in the highest bin leads to annual seismic activity rates more sensitive to magnitude differences between the two catalogues due to the different treatment of depth in the two methodologies. However, given the small number of events in these bins, the observed differences in annual seismic activity rates should not be considered statistically significant as shown by the error bars computed with the Weichert method shown in Figure 11a.

DISCUSSIONS AND CONCLUSIONS

The purpose of this paper is to perform a benchmarking exercise to quantify the differences in the estimates of historical earthquake magnitudes determined with different methodologies, namely Boxer and the QUake-MD, calibrated with the same sets of macroseismic and instrumental data. We then compare the annual seismic activity rates resulting from the two parametric earthquake catalogues obtained from the two methodologies applied to the same intensity distributions of historical earthquakes.

One of the main differences between the two methodologies is the treatment of the hypocentral depth of the earthquakes, and the determination of the epicentral location. In Boxer, the depth is implicitly taken into account through the epicentral intensity value in the estimate of magnitude from the macroseismic field. Magnitude estimates are then less influenced by eventually biased depth estimates. However, in case of sources with depth significantly different from the mean of the calibration dataset, the use of the epicentral intensity as proxy for depth may affect the accuracy of the magnitude estimate. Estimating depth is paramount especially for events which have depth that maybe significantly different from the mean depth of the study area. In QUake-MD both depth and magnitude are deduced from the macroseismic field. However, estimates of depth can be biased by poor quality macroseismic fields. Since magnitude and depth values are correlated, magnitude can be also biased. In those cases, fixing depth or reducing the acceptable depth limits in QUake-MD is strongly encouraged (Provost and Scotti, 2020). In this case, if magnitude estimates are on average less influenced by potentially biased depth estimates, the magnitude of events much deeper than the average depth of the calibration events are most likely underestimated.

We show that the two methodologies offer equivalent results in terms of magnitude estimates when calibrated with the same dataset, although the they are quite different in their approach to estimate magnitude. Systematic differences between the two methods (higher magnitudes for QUake-MD estimates compared to Boxer's for events deeper than 8 km) are shown to be due to difference in the way depth is considered. As a consequence, according to the results presented here, earthquake rates based on the catalogues compiled with the two methodologies are similar, as long as earthquakes in any given magnitude bin present a range of depths balanced around 8 km, as for example for M< 4.5. On the other hand, for areas with either prevailing very shallow or deep seismic activity, a bias could exist between the magnitude estimates of the two methodologies, and thus on the corresponding annual seismic activity rates. Finally, although differences in annual seismic activity rates observed in the study area between the two methodologies for M>4.5 are statistically poorly constrained, conclusions may differ in regions with higher seismic activity, in particular when considering only magnitudes Mw > 4.5 to compute frequency-magnitude distributions.

Estimate of the parameters of historical earthquakes are associated with large uncertainties depending on the method chosen to estimate those parameters, the calibration database for the chosen method or the completeness of the intensity distributions, that sum to the uncertainties deriving from the assessment of macroseismic intensity from documental information. Such uncertainties should be thoroughly considered throughout a seismic hazard assessment, whereas they are usually disregarded. We believe that sharing the same data for the calibration earthquakes, i.e. instrumental magnitude, depth, and location and macroseismic data among neighboring countries, as well as the same macroseismic data for the historical earthquakes, would help reducing the differences between magnitude estimates across borders. However, building on the experience of this study and depending on the purpose of a seismic hazard assessment study, it may be worth exploring different methodologies to convert macroseismic data into earthquake parameters to ensure that the range of plausible earthquake rates are sampled. Indeed, the uncertainties in earthquake parameters derived from one methodology only could be insufficient to represent the actual epistemic uncertainty of such estimates.

FIGURES CAPTIONS

Fig. 1 : Epicentral location, magnitude, and number of intensity data points (IDP) of each earthquake in the calibration dataset. Numbers indicate the earthquakes as in Table 1. Each earthquake is represented by a circle. The circle size represents the magnitude of the earthquake: the bigger the magnitude earthquake, the bigger the size. Color of the circle represent the number of IDP (column NObs in table 1) associated to each earthquake: the darker the color, the smaller the number of IDP.

Fig. 2 : Calibration dataset: frequencies of the intensity values for the whole (light gray on the figure) and the complete (dark gray on the figure) dataset (NN = non-numerical values). Complete dataset means all data within complete intensity classes, i.e. intensities greater than their associated intensity of completeness. The whole dataset includes all available IDPs, excluding intensity levels smaller than 2.

Fig. 3 : Residuals between instrumental and macroseismic magnitude estimates for QUake-MD (a) and Boxer (b) as a function of instrumental magnitude for the calibration earthquakes used in this study. The horizontal dashed lines represent the +/- 0.5 values of magnitude residuals. A dark blue color (black points in non-color version) is used to identify earthquakes with native Mw and a yellow color (light gray in non-color version) is used to identify earthquakes with Mw magnitudes coming from conversion of another type of magnitude.

Fig. 4 : Test dataset: epicentral location, magnitude, and number of intensity data (IDP) for each earthquake in the test dataset. Each earthquake is represented by a circle. The circle size represents the magnitude of the earthquake: the bigger the magnitude earthquake, the bigger the size. Color of the circle represent the number of IDP (column NObs in Table 1) associated to each earthquake: the darker the color, the smaller the number of IDP.

Fig. 5 : Comparison between instrumental and macroseismic magnitude residual for QUake-MD (c) and Boxer (d) as a function of instrumental magnitude for the test data set. A dark blue color (black points in non-color version) is used to identify earthquakes with native Mw and a yellow color (light gray in non-color version) is used to identify earthquakes with Mw magnitudes coming from conversion of another type of magnitude. Top figures: Pearson coefficient for the entire dataset (dashed violet lines) and for sliding magnitude windows of one unit width (gray lines) for QUake-MD (a) and Boxer (b). The subplots (a) and (b) a horizontal dotted line is added at Pearson coefficient value of 0.5.

Fig. 6 : Comparison of QUake-MD and Boxer magnitudes for the 102 earthquakes of the test dataset. A 1:1 black line is added, as well as dashed lines representing the +/- 0.5 magnitude difference around the 1:1 line. The colors represent the values of QUake-MD depth. White color is associated to depth equal to 8 km. Earthquakes with depths deeper than 8 km are mostly above the 1:1 line and earthquakes with depths shallower than 8 km are mostly under the 1:1 line.

Fig. 7: Residual depth as a function of magnitude and depth residuals for the calibration dataset (left figure) and the test dataset (right figure). Residuals are computed by subtracting instrumental values and QUake-MD's values. In both figures, the solid black lines represent the 0 residual, the horizontal dashed lines represent the +/- 5 km depth residual and the vertical dashed lines represent the +/- 0.5 unit magnitude residual.

Fig. 8 : Earthquakes in the study area (black polygon). Blue (the darker gray in non-color version) points represent the post-1980 earthquakes and the red (the lighter gray in non-color version) points represent

the historical earthquakes, i.e. earthquakes that occurred before 1980, including 1980. The insert map shows the location of the study area at a larger scale.

Fig. 9 : Comparison of Boxer and QUake-MD magnitudes for the study area A 1:1 black line is added, as well as dashed lines representing the +/- 0.5 magnitude difference around the 1:1 line. The colors represent the values of QUake-MD depth. White color is associated to depth equal to 8 km. Earthquakes with depths deeper than 8 km are mostly above the 1:1 line and earthquakes with depths shallower than 8 km are mostly under the 1:1 line.

Fig. 10 : Completeness time distribution obtained by the application of the Albarello 2001 algorithm for the Boxer and QUake-MD catalogues, with the magnitude-time distribution for the Boxer's catalogue (a) and QUake-MD's catalogue (b). Filled lines represent the mean of the completeness time distribution and the dashed lines the quartiles. A blue (darker gray in non-color version) color is used for Boxer and red (lighter gray) color is used for QUake-MD.

Fig. 7 :(a) Incremental annual rates for Boxer and QUake-MD catalogues. Vertical lines represent the annual seismicity rates uncertainty computed as described in Weichert (1980). A blue (darker gray in non-color version) color and a square symbol are used for Boxer and red (lighter gray) color and a circle symbol are used for QUake-MD. The QUake-MD symbols are slightly shifted to the right for better readability of the figure. (b) Number of earthquakes per magnitude bin used to compute the annual seismicity rates for the Boxer's catalogue (blue bars) and QUake-MD's catalogue (red bars).

ONLINE RESOURCE CAPTIONS

ESM_1.xlsx : Additional information to the calibration list presented in table 1 of the main text

ESM_2.pdf : additional information about QUake-MD calibration procedure and macroseismic fields of two calibration earthquake.

ESM_3.xlsx : output file of Boxer calibration with Boxer's model coefficients

ESM_4.xlsx : QUake-MD's coefficients after calibration

ESM_5.xlsx : Results of the application of Boxer and QUake-MD on the calibration dataset

ESM_6.xlsx : List of the test dataset

ESM_7.xlsx : Boxer's and QUake-MD catalogues for the study area

REFERENCES

Albarello, D., R. Camassi, and A. Rebez (2001). Detection of space and time heterogeneity in the completeness of a seismic catalog by a statistical approach: An Application to the Italian Area, *Bull. Seismol. Soc. Am.* **91**, no. 6, 1694–1703, doi: 10.1785/0120000058.

Allen, T. I., D. J. Wald, and C. B. Worden (2012). Intensity attenuation for active crustal regions, *J. Seismol.* **16**, no. 3, 409–433, doi: 10.1007/s10950-012-9278-7.

Bakun, W. H., A. G. Capera, and M. Stucchi (2011). Epistemic Uncertainty in the Location and Magnitude of Earthquakes in Italy from Macroseismic Data, *Bull. Seismol. Soc. Am.* **101**, no. 6, 2712–2725, doi: 10.1785/0120110118.

Bakun, W. H., and O. Scotti (2006). Regional intensity attenuation models for France and the estimation of magnitude and location of historical earthquakes, *Geophys. J. Int.* **164**, no. 3, 596–610, doi: 10.1111/j.1365-246X.2005.02808.x.

Bakun, W. H., and C. M. Wentworth (1997). Estimating Earthquake Location and Magnitude from Seismic Intensity Data, *Bull. Seismol. Soc. Am.* **87**, no. 6, 1502–1521.

Beauval, C., P.-Y. Bard, and L. Danciu (2020). The influence of source- and ground-motion model choices on probabilistic seismic hazard levels at 6 sites in France, *Bull. Earthq. Eng.* **18**, no. 10, 4551–4580, doi: 10.1007/s10518-020-00879-z.

Bernardi, F., J. Braunmiller, and D. Giardini (2005). Seismic Moment from Regional Surface-Wave Amplitudes: Applications to Digital and Analog Seismograms, *Bull. Seismol. Soc. Am.* **95**, no. 2, 408–418, doi: 10.1785/0120040048.

Braunmiller, J., sed M. W. Group, N. Deichmann, sed M. W. Group, D. Giardini, sed M. W. Group, S. Wiemer, and sed M. W. Group (2005). Homogeneous Moment-Magnitude Calibration in Switzerland, *Bull. Seismol. Soc. Am.* **95**, no. 1, 58–74, doi: 10.1785/0120030245.

Burkhard, M., and G. Grünthal (2009). Seismic source zone characterization for the seismic hazard assessment project PEGASOS by the Expert Group 2 (EG1b), 1, *Swiss J. Geosci.* **102**, no. 1, 149–188, doi: 10.1007/s00015-009-1307-3.

Cara, M., Y. Cansi, A. Schlupp, P. Arroucau, N. Béthoux, E. Beucler, S. Bruno, M. Calvet, S. Chevrot, A. Deboissy, *et al.* (2015). SI-Hex: a new catalogue of instrumental seismicity for metropolitan France, *Bull. Soc. Geol. Fr.* **186**, no. 1, 3–19, doi: 10.2113/gssgfbull.186.1.3.

Castello, B., G. Selvaggi, C. Chiarabba, and A. Amato (2006). CSI Catalogo della sismicità italiana 1981-2002, versione 1.1, *Ist. Naz. Geofis. E Vulcanol. INGV*, doi: 10.13127/CSI.1.1.

Cecic, I., R. M. W. Musson, and M. Stucchi (1996). Do seismologists agree upon epicentre determination from macroseismic data? A survey of ESC Working Group ’ Macroseismology’, 5, *Ann. Geophys.* **39**, no. 5, doi: 10.4401/ag-4031.

Duverger, C., G. Mazet-Roux, L. Bollinger, A. G. Trilla, A. Vallage, B. Hernandez, and Y. Cansi (2021). A decade of seismicity in metropolitan France (2010–2019): the CEA/LDG methodologies and observationsUne décennie de sismicité en France métropolitaine (2010–2019) : les méthodes et observations du CEA/LDG, *Bull. Société Géologique Fr.* **192**, no. 1, doi: 10.1051/bsgf/2021014.

Fäh, D., D. Giardini, P. Kästli, N. Deichmann, M. Gisler, G. Schwarz-Zanettti, S. Alvarez-Rubio, S. Sellami, B. Edwards, B. Allmann, *et al.* (2011). ECOS-09, Earthquake Catalogue of Switzerland - Release 2011, SED/ECOS/R/001/20110417, Swiss Seismological Service, ETH Zürich.

Galanopulos, A. G. (1961). On magnitude determination by using macroseismic data, *Ann. Geophys.* **14**, 225–253.

Gardner, J. K., and L. Knopoff (1974). Is the sequence of earthquakes in Southern California, with aftershocks removed, Poissonian?, *Bull. Seismol. Soc. Am.* **64**, no. 5, 1363–1367.

Gasperini, P., F. Bernardini, G. Valensise, and E. Boschi (1999). Defining seismogenic sources from historical earthquake felt reports, *Bull. Seismol. Soc. Am.* **89**, no. 1, 94–110.

Gasperini, P., G. Vannucci, D. Tripone, and E. Boschi (2010). The Location and Sizing of Historical Earthquakes Using the Attenuation of Macroseismic Intensity with Distance, *Bull. Seismol. Soc. Am.* **100**, no. 5A, 2035–2066, doi: 10.1785/0120090330.

Gomez-Capera, A. A., A. Rovida, P. Gasperini, M. Stucchi, and D. Viganò (2015). The determination of earthquake location and magnitude from macroseismic data in Europe, *Bull. Earthq. Eng.* **13**, no. 5, 1249–1280, doi: 10.1007/s10518-014-9672-3.

Jomard, H., O. Scotti, S. Auclair, P. Dominique, K. Manchuel, and D. Sicilia (2021). The SISFRANCE database of historical seismicity. State of the art and perspectives, *Comptes Rendus Géoscience* **353**, no. S1, 1–24, doi: 10.5802/crgeos.91.

Knopoff, L., and J. K. Gardner (1972). Higher Seismic Activity During Local Night on the Raw Worldwide Earthquake Catalogue, *Geophys. J. Int.* **28**, no. 3, 311–313, doi: 10.1111/j.1365-246X.1972.tb06133.x.

Kövesligethy, R. (1907). Seismischer Stärkegrad und Intensität der Beben, *Gerlands Beitr. Zur Geophys.* **8**, 21–103.

Locati, M., R. D. Camassi, A. N. Rovida, E. Ercolani, F. M. A. Bernardini, V. Castelli, C. H. Caracciolo, A. Tertulliani, A. Rossi, R. Azzaro, *et al.* (2019). Italian Macroseismic Database DBMI15, version 2.0, report, doi: https://doi.org/10.13127/DBMI/DBMI15.2.

Lolli, B., G. Paolo, E. Biondini, and G. Vannucci (2020). The relationship between ML and Mw for small earthquakes (ML < 2-4) in Italy, **22**, 18821.

Meletti, C., W. Marzocchi, V. D'Amico, G. Lanzano, L. Luzi, F. Martinelli, B. Pace, A. Rovida, M. Taroni, and F. Visini (2021). The new Italian seismic hazard model (MPS19), *Ann. Geophys.* **64**, no. 1, doi: https://doi.org/10.4401/ag-8579.

Mucciarelli, M. (2014). Uncertainty in PSHA related to the parametrization of historical intensity data, *Nat. Hazards Earth Syst. Sci.* **14**, no. 10, 2761–2765, doi: https://doi.org/10.5194/nhess-14-2761-2014.

Musson, R. (2012). The Effect of Magnitude Uncertainty on Earthquake Activity Rates, *Bull. Seismol. Soc. Am.* **102**, 2771–2775, doi: 10.1785/0120110224.

Musson, R. M. W., G. Grünthal, and M. Stucchi (2010). The comparison of macroseismic intensity scales, *J. Seismol.* **14**, no. 2, 413–428, doi: 10.1007/s10950-009-9172-0.

Musson, R. M. W., and M. J. Jimenéz (2008). Macroseismic estimation of earthquake parameters, NA4 deliverable, NEREIS Project.

Pasolini, C., D. Albarello, P. Gasperini, V. D'Amico, and B. Lolli (2008). The Attenuation of Seismic Intensity in Italy, Part II: Modeling and Validation, *Bull. Seismol. Soc. Am.* **98**, no. 2, 692–708, doi: 10.1785/0120070021.

Provost, L., and O. Scotti (2020). QUake-MD: Open-Source Code to Quantify Uncertainties in Magnitude–Depth Estimates of Earthquakes from Macroseismic Intensities, *Seismol. Res. Lett.* **91**, no. 5, 2520–2530, doi: 10.1785/0220200064.

Rong, Y., M. Mahdyiar, B. Shen-Tu, and K. Shabestari (2011). Magnitude problems in historical earthquake catalogues and their impact on seismic hazard assessment, *Geophys. J. Int.* **187**, no. 3, 1687–1698, doi: 10.1111/j.1365-246X.2011.05226.x.

Rovida, A., and M. Locati (2015). Archive of Historical Earthquake Data for the European-Mediterranean Area, in *Perspectives on European Earthquake Engineering and Seismology: Volume 2* A. Ansal(Editor), Springer International Publishing, Cham, Geotechnical, Geological and Earthquake Engineering, 359–369, doi: 10.1007/978-3-319-16964-4_14.

Rovida, A., M. Locati, R. Camassi, A. Antonucci, F. M. A. Bernardini, C. H. Caracciolo, and L. Maffezzoni (2017). L'Archivio Storico Macrosismico Italiano (ASMI).

Rovida, A., M. Locati, R. Camassi, B. Lolli, and P. Gasperini (2020). The Italian earthquake catalogue CPTI15, *Bull. Earthq. Eng.* **18**, no. 7, 2953–2984, doi: 10.1007/s10518-020-00818-y.

Sandron, D., G. Renner, A. Rebez, and D. Slejko (2014). Early instrumental seismicity recorded in the eastern Alps, *Boll. Geofis. Teor. Ed Appl.* **55**, no. 4, 755–788, doi: 10.4430/bgta0118.

Scotti, O., D. Baumont, G. Quenet, and A. Levret (2004). The French macroseismic database SISFRANCE: objectives, results and perspectives, *Ann. Geophys.* **47**, no. 2/3, 11.

Sibol, M. S., G. A. Bollinger, and J. B. Birch (1987). Estimation of magnitudes in central and eastern North America using intensity and felt area, *Bull. Seismol. Soc. Am.* **77**, no. 5, 1635–1654.

Slejko, D., R. Camassi, I. Cecic, D. Herak, M. Herak, S. Kociu, V. Kouskouna, J. Lapajne, K. Makropoulos, C. Meletti, *et al.* (1999). Seismic hazard assessment for Adria, *Ann. Geofis.* **42**, no. 6.

Storchak, D. A., D. D. Giacomo, I. Bondár, E. R. Engdahl, J. Harris, W. H. K. Lee, A. Villaseñor, and P. Bormann (2013). Public Release of the ISC–GEM Global Instrumental Earthquake Catalogue (1900–2009), *Seismol. Res. Lett.* **84**, no. 5, 810–815, doi: 10.1785/0220130034.

Vannucci, G., B. Lolli, and P. Gasperini (2021). Inhomogeneity of Macroseismic Intensities in Italy and Consequences for Macroseismic Magnitude Estimation, *Seismol. Res. Lett.* **92**, no. 4, 2234–2244, doi: 10.1785/0220200273.

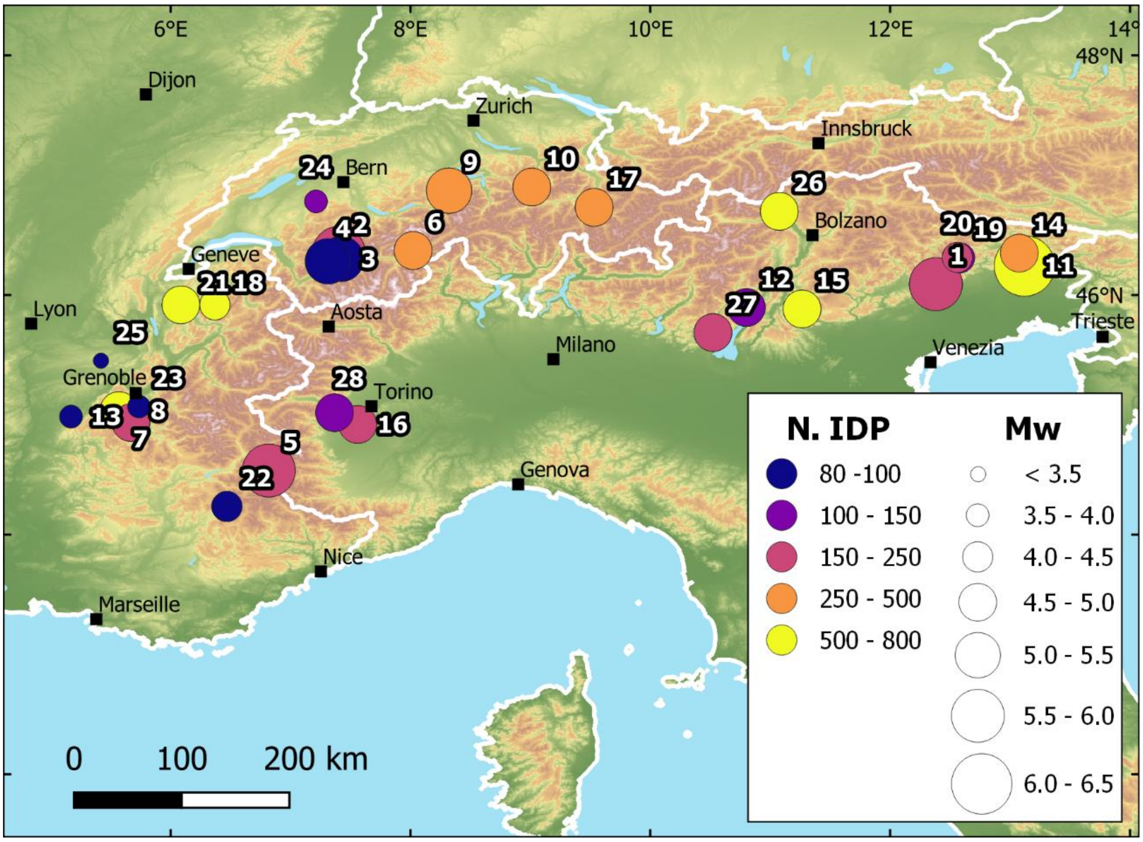
Visini, F., B. Pace, C. Meletti, W. Marzocchi, A. Akinci, R. Azzaro, S. Barani, G. Barberi, G. Barreca, R. Basili, *et al.* (2021). Earthquake Rupture Forecasts for the MPS19 Seismic Hazard Model of Italy, 2, *Ann. Geophys.* **64**, no. 2, SE220–SE220, doi: 10.4401/ag-8608.

Weichert, D. (1980). Estimation of the earthquake recurrence parameters for unequal observation periods for different magnitudes, *Bull. Seismol. Soc. Am.* **70**, no. 4, 1337–1346.

Wiemer, S. (2001). A Software Package to Analyze Seismicity: ZMAP | Seismological Research Letters | GeoScienceWorld, *Seismol. Res. Lett.* **72**, no. 3, 373–382, doi: https://doi.org/10.1785/gssrl.72.3.373.

FIGURES

593



594 **Fig. 1**

595

596

597

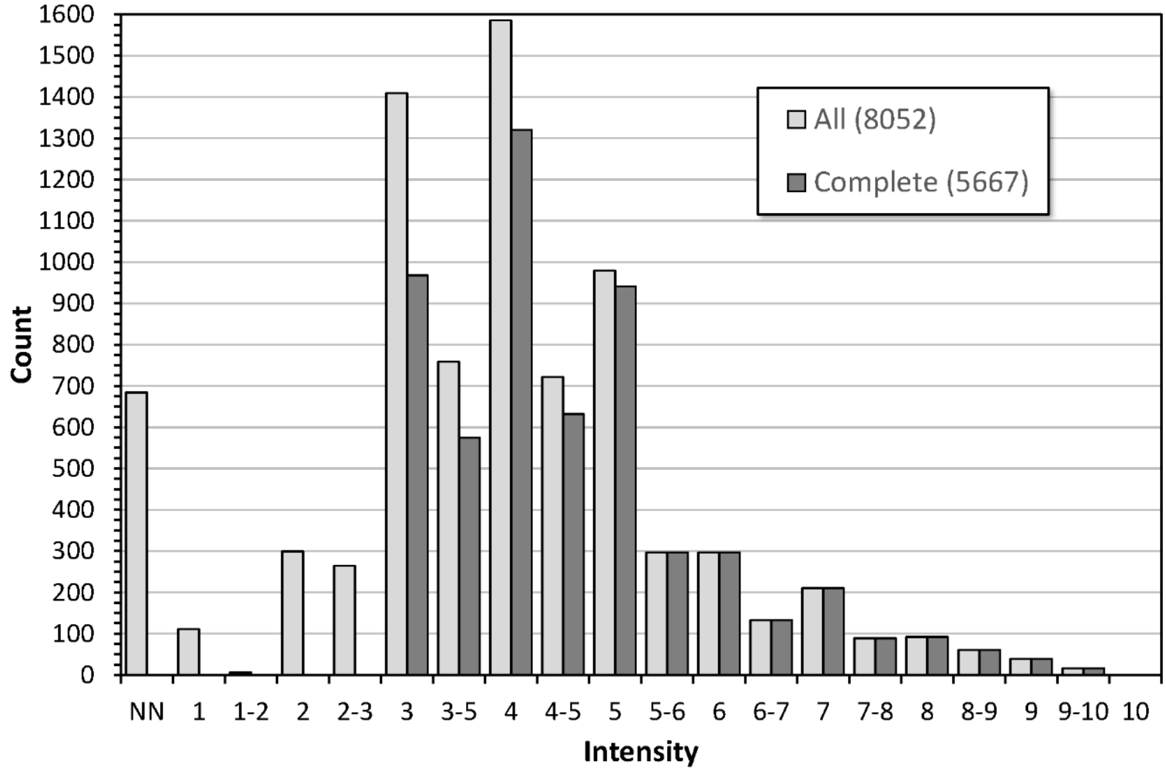

All (8052)
Complete (5667)
Count
1600
1500
1400
1300
1200
1100
1000
900
800
700
600
500
400
300
200
100
0
NN 1 1-2 2 2-3 3 3-5 4 4-5 5 5-6 6 6-7 7 7-8 8 8-9 9 9-10 10
Intensity

Fig. 2

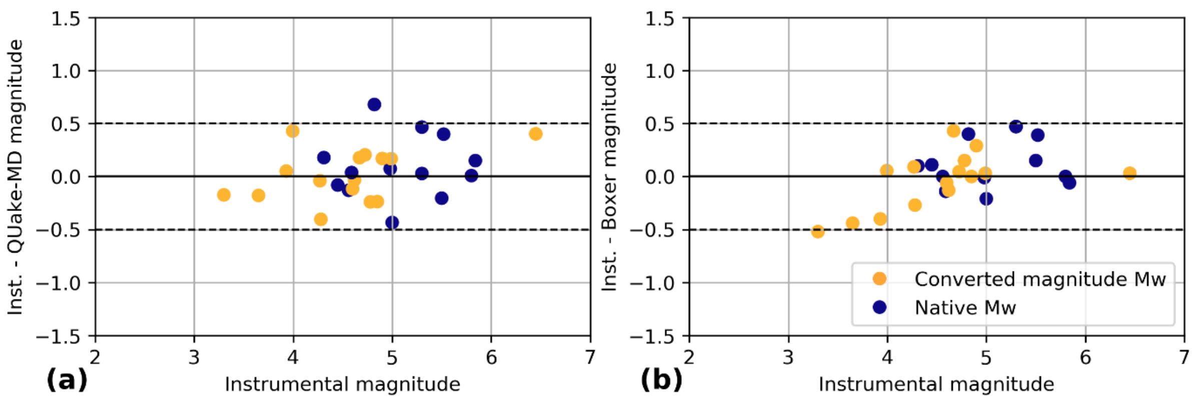

Inst. - QUake-MD magnitude
Instrumental magnitude
(a)
Inst. - Boxer magnitude
Instrumental magnitude
(b)
Converted magnitude Mw
Native Mw
1.5
1.0
0.5
0.0
−0.5
−1.0
−1.5
2 3 4 5 6 7

Fig. 3

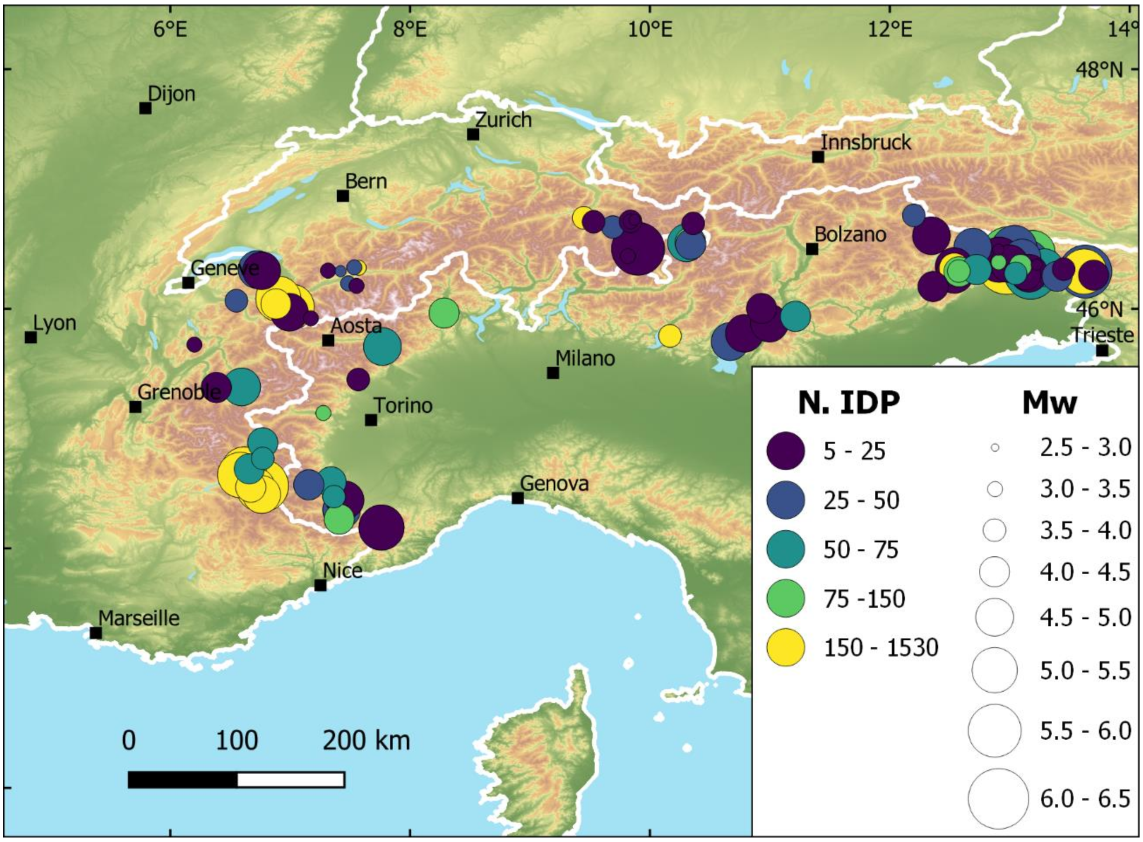

6°E
8°E
10°E
12°E
14°
48°N
46°N
Dijon
Zurich
Innsbruck
Bern
Bolzano
Geneve
Lyon
Aosta
Milano
Trieste
Grenoble
Torino
Genova
Nice
Marseille
N. IDP
5 - 25
25 - 50
50 - 75
75 -150
150 - 1530
Mw
2.5 - 3.0
3.0 - 3.5
3.5 - 4.0
4.0 - 4.5
4.5 - 5.0
5.0 - 5.5
5.5 - 6.0
6.0 - 6.5
0
100
200 km

Fig. 4

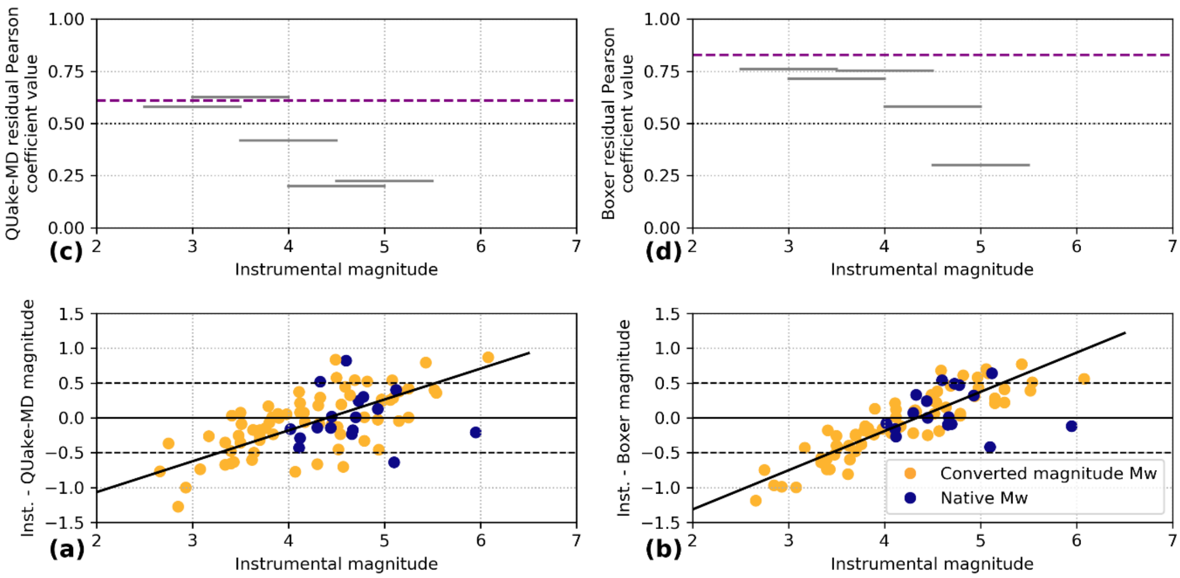

QUake-MD residual Pearson coefficient value
Boxer residual Pearson coefficient value
Instrumental magnitude
(c)
(d)
Inst. - QUake-MD magnitude
Inst. - Boxer magnitude
(a)
(b)
Converted magnitude Mw
Native Mw

Fig. 5

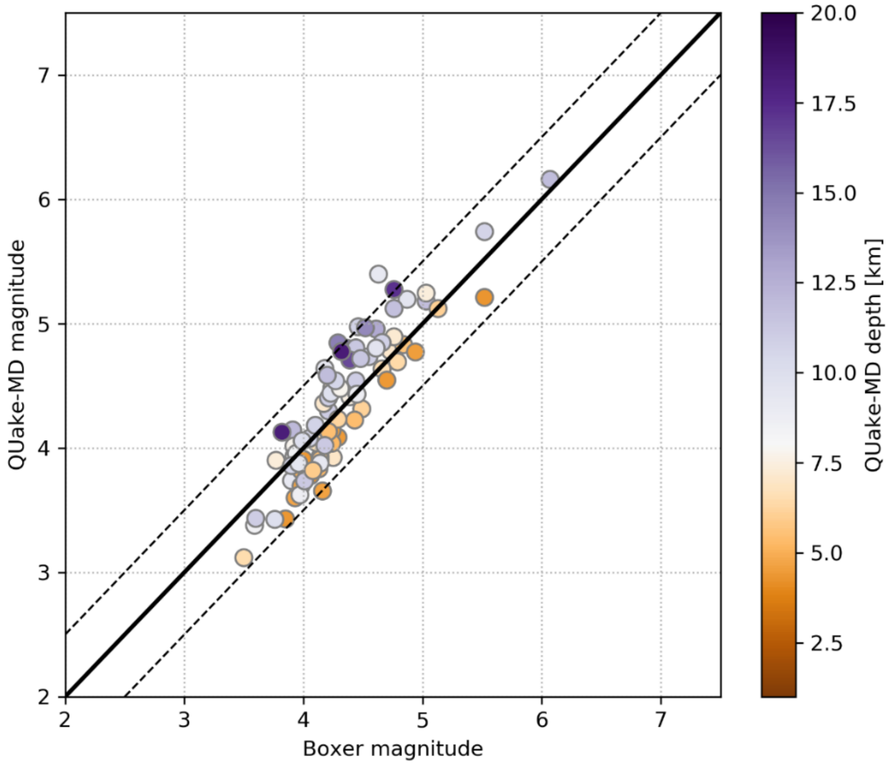

QUake-MD magnitude
Boxer magnitude
QUake-MD depth [km]
20.0
17.5
15.0
12.5
10.0
7.5
5.0
2.5

Fig. 6

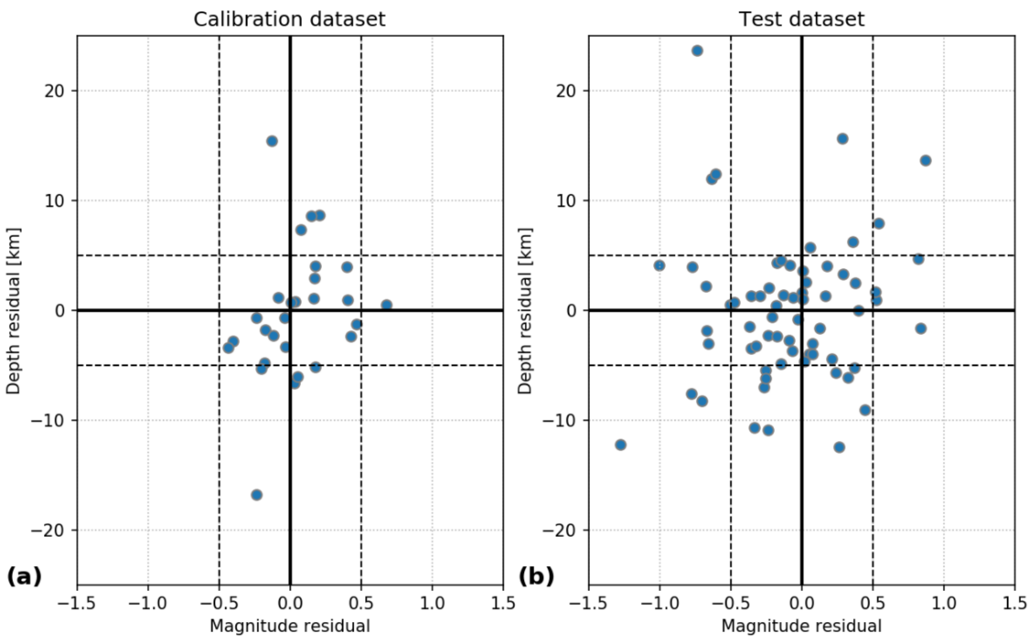

Calibration dataset
Test dataset
Depth residual [km]
Magnitude residual
(a)
(b)

Fig. 7

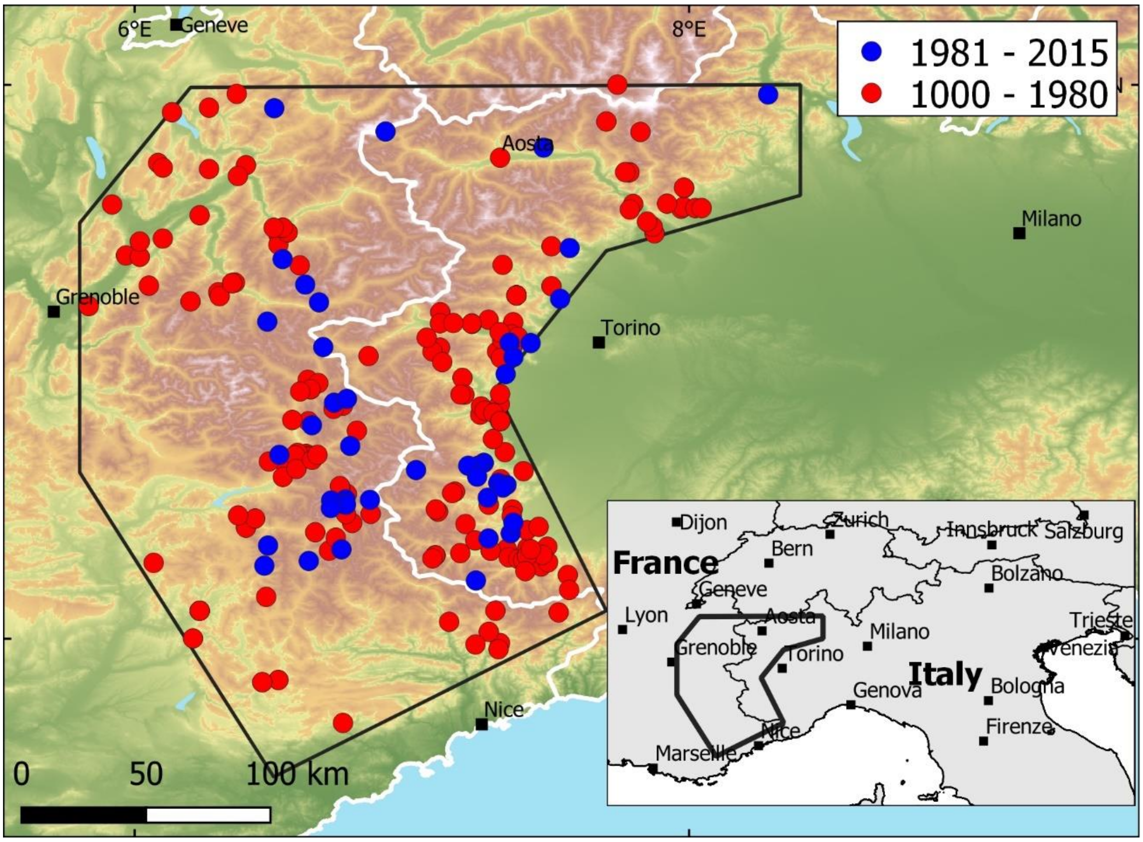

6°E
Geneve
8°E
1981 - 2015
1000 - 1980
Aosta
Milano
Grenoble
Torino
Nice
0
50
100 km
Dijon
Zurich
Innsbruck
Salzburg
Bern
France
Bolzano
Geneve
Lyon
Aosta
Milano
Trieste
Grenoble
Torino
Venezia
Italy
Genova
Bologna
Firenze
Marseille
Nice

Fig. 8

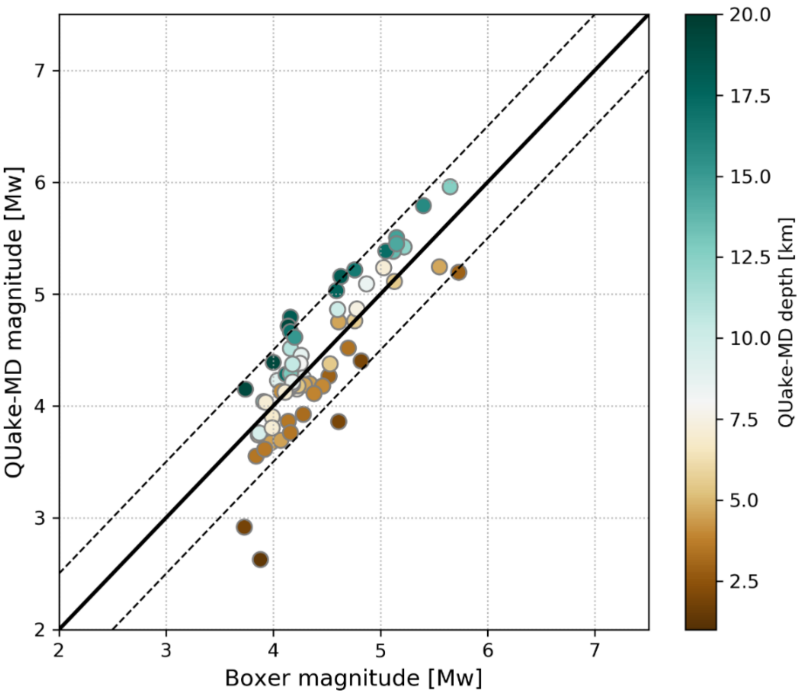
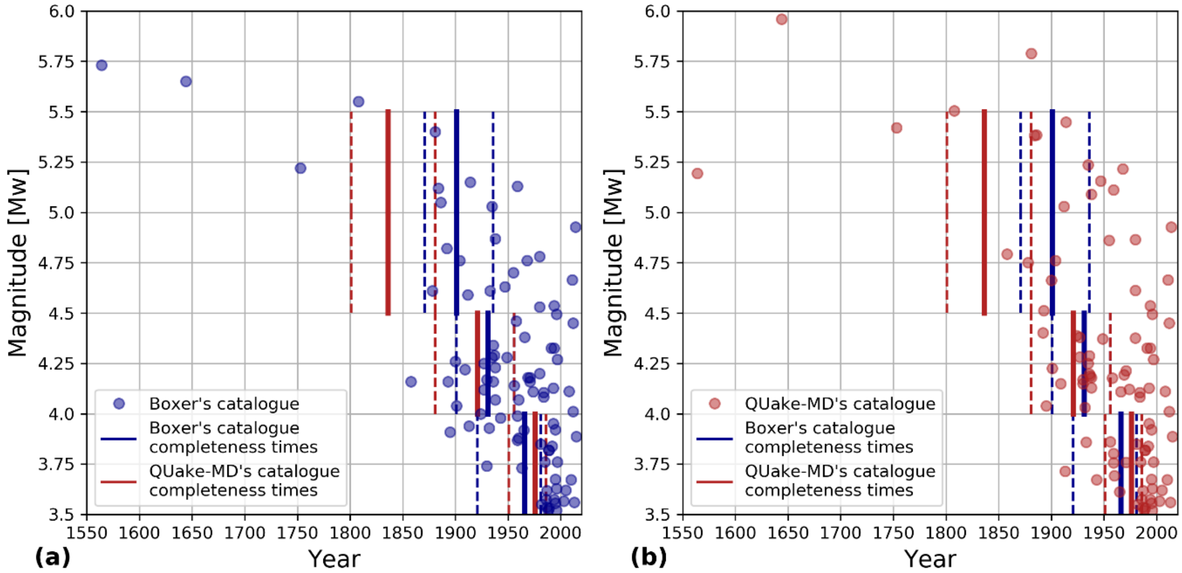

QUake-MD magnitude [Mw]
Boxer magnitude [Mw]
QUake-MD depth [km]
2
3
4
5
6
7
20.0
17.5
15.0
12.5
10.0
7.5
5.0
2.5

Fig. 9



Magnitude [Mw]
Year
Boxer's catalogue
Boxer's catalogue completeness times
QUake-MD's catalogue completeness times
(a)
QUake-MD's catalogue
Boxer's catalogue completeness times
QUake-MD's catalogue completeness times
(b)

Fig. 10

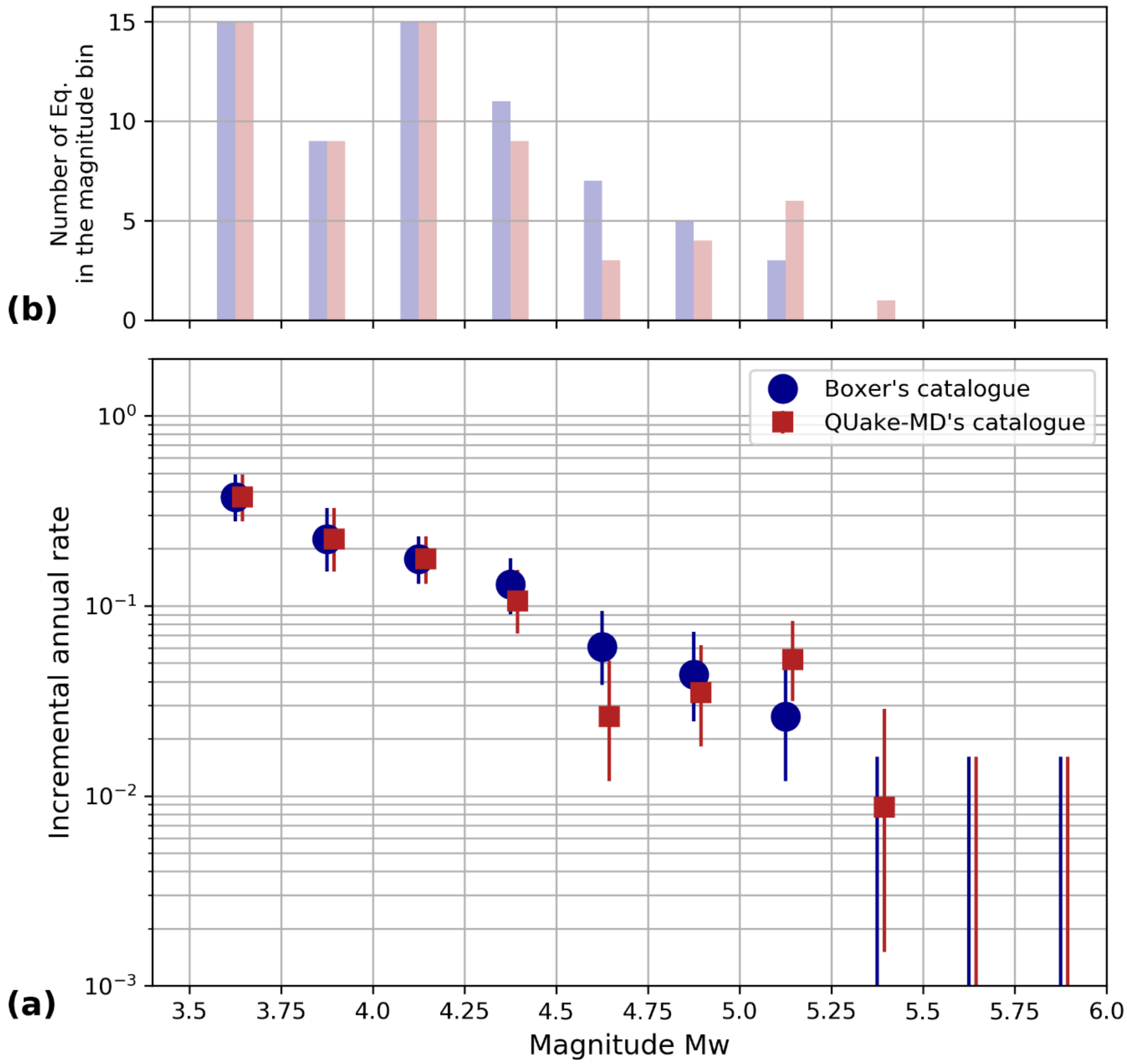

(b)
Number of Eq.
in the magnitude bin
15
10
5
0
(a)
Incremental annual rate
10^{0}
10^{-1}
10^{-2}
10^{-3}
Boxer's catalogue
QUake-MD's catalogue
3.5
3.75
4.0
4.25
4.5
4.75
5.0
5.25
5.5
5.75
6.0
Magnitude Mw

Fig. 11