



Global occurrence and impact of small-to-medium magnitude earthquakes: a statistical analysis

Cecilia I. Nieves¹ · Julian J. Bommer² · Helen Crowley³ · Jan van Elk⁴

Received: 26 March 2019 / Accepted: 21 August 2019 / Published online: 26 September 2019
© The Author(s) 2019

Abstract

Despite their much smaller individual contribution to the global counts of casualties and damage than their larger counterparts, earthquakes with moment magnitudes M_w in the range 4.0–5.5 may dominate seismic hazard and risk in areas of low overall seismicity, a statement that is particularly true for regions where anthropogenically-induced earthquakes are predominant. With the risk posed by these earthquakes causing increasing alarm in certain areas of the globe, it is of interest to determine what proportion of earthquakes in this magnitude range that occur sufficiently close to population or the built environment do actually result in damage and/or casualties. For this purpose, a global catalogue of potentially damaging events—that is, earthquakes deemed as potentially capable of causing damage or casualties based on a series of pre-defined criteria—has been generated and contrasted against a database of reportedly damaging small-to-medium earthquakes compiled in parallel to this work. This paper discusses the criteria and methodology followed to define such a set of potentially damaging events, from the issues inherent to earthquake catalogue compilation to the definition of criteria to establish how much potential exposure is sufficient to consider each earthquake a threat. The resulting statistics show that, on average, around 2% of all potentially-damaging shocks were actually reported as damaging, though the proportion varies significantly in time as a consequence of the impact of accessibility to data on damage and seismicity in general. Inspection of the years believed to be more complete suggests that a value of around 4–5% might be a more realistic figure.

Keywords Damaging earthquakes · Earthquake impact · Seismic risk · Earthquake recurrence · Earthquake catalogue compilation

1 Introduction

The focus of earthquake engineering has been traditionally set on protecting the built environment and its inhabitants from the potential consequences of severe dynamic loading imposed by moderate and large earthquakes. However, interest in smaller magnitude seismic events has risen in recent years, likely due to a combination of factors including the

✉ Julian J. Bommer
j.bommer@imperial.ac.uk

Extended author information available on the last page of the article

occurrence of some unusually destructive events with magnitudes below what is generally considered a damaging earthquake, the continuously-evolving development of seismic risk assessment methodologies focused on the existing building stock rather than new constructions, and the heightened recognition of the potential risk posed by induced seismicity.

The consequences that have been reported for earthquakes with moment magnitudes M_w in the range 4.0–5.5 have been investigated and gathered in the form of a database whose scope, development and resulting characteristics are explored in separate publications (Nievas et al. 2019b, c). To further address the risk posed by these small-to-medium magnitude earthquakes, the objective of the present work is to quantify the proportion of upper-crustal earthquakes with magnitudes in this range and occurring in close proximity to exposed assets or population, that have resulted in reported damage and/or casualties. This statement implies three conditions of the earthquakes analysed herein: (i) they occur in the upper crust, (ii) their moment magnitude is equal to or larger than 4.0 and equal to or smaller than 5.5, and (iii) they occur sufficiently close to human settlements to pose a threat to the population and/or the built environment.

The condition of working with upper-crustal earthquakes stems from two reasons. The first is to aim at the potential applications of the results to the context of seismic risk associated with anthropogenically-induced earthquakes, which, by nature, occur within the regions of the crust accessible to human activities. The second is that deeper earthquakes in this magnitude range are too unlikely to pose a relevant threat. The term “upper-crustal” is used herein in contrast with deeper subduction earthquakes and earthquakes related to relatively specific conditions, like those found in Bucamaranga (Colombia-Venezuela), Vrancea (Romania), or the Hindu Kush (Afghanistan-Pakistan).

The upper magnitude limit of M_w 5.5 was selected based on enduring widespread recognition by the earthquake engineering community that earthquakes of greater magnitude can be destructive (e.g., Bommer and Crowley 2017). The lower limit of M_w 4.0 was selected on the dual basis of it being highly unlikely that smaller events could cause appreciable damage, and the potential difficulties associated with selecting a value below the completeness threshold of the source earthquake catalogues.

The last requisite is the proximity to exposed assets or individuals and is fundamental to the objective of this work, as only earthquakes that occur sufficiently close to human settlements can pose a threat. The most obvious earthquakes that need to be filtered out in this regard are those occurring in the oceans and largely uninhabited areas of the planet, the challenge residing largely on the determination of what is *sufficiently close* and what is *too far away*, as will be discussed further herein.

The study has been carried out on earthquakes occurring worldwide during the 15 years spanning from 2001 through 2015. The paper starts with a comprehensive description of the methodology, which is then followed by two sections providing details on the different steps of the process adopted to identify the set of potentially damaging events, as well as sensitivity analyses carried out on preliminary versions of the catalogue and database of damaging earthquakes, whose outcome informed many of the decisions made in the final work presented herein. The fifth section then presents the resulting catalogue and statistics on the observed proportion of damaging earthquakes. The paper finishes with a detailed discussion of the results obtained and the conclusions reached.

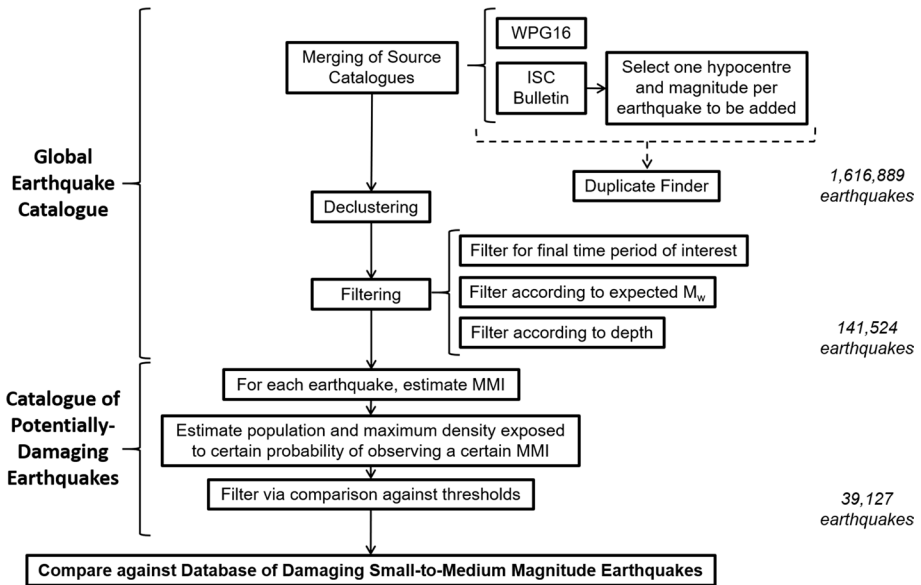


Fig. 1 Outline of the methodology. Numbers in italic on the right indicate the number of earthquakes at each stage

2 Outline of the methodology

The methodology followed is composed of three key stages, as shown in Fig. 1. The first is the compilation of a homogeneous-magnitude global catalogue of upper-crustal earthquakes with moment magnitudes M_w in the range 4.0–5.5. This was done by means of merging an updated version of the homogeneous-magnitude catalogue of Weatherill et al. (2016), referred to as WPG16* hereafter (version 3c, provided by the authors), with events from the Bulletin of the International Seismological Centre (ISC) that were not contained therein.

The global earthquake catalogue produced herein was first built for the period 1st January 1998–31st August 2018, longer than the 15-year-long period of interest so as to be able to adequately subject it to a declustering process, and then filtered to the final period 1st January 2001–31st December 2015. The latter period was selected for a series of reasons:

- At the time of starting this analysis, the Reviewed ISC Bulletin (i.e., the Bulletin that has been processed by specifically-designed algorithms and the ISC staff¹) was available up to 31st January 2016.
- The magnitude of completeness of the WPG16* catalogue stays relatively stable at around 4.0 from the late 1990s.
- From around the 2000s there is a significant increase in the availability of data regarding damage and/or casualties caused by small-to-medium magnitude earth-

¹ Throughout this paper, “ISC Bulletin” refers to the Bulletin that contains data from all contributing agencies/networks and “Reviewed ISC Bulletin” refers to the subset of the later that has been re-processed by the ISC staff.

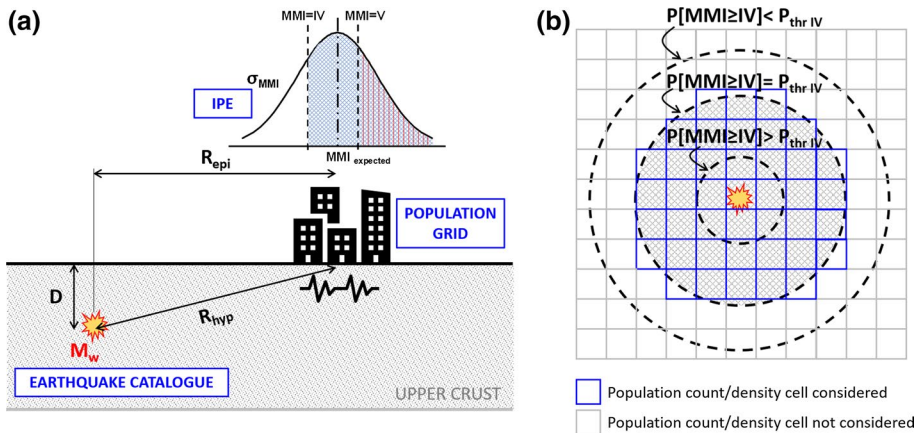


Fig. 2 Schematic representation of the process used to determine whether an earthquake occurred *sufficiently close* to population and the built environment so as to pose a threat or not: **a** side view and **b** plan view. The plan view exemplifies the criterion used to include cells from the population grid, using only the probability of observing $\text{MMI} \geq \text{IV}$ and its associated probability threshold ($P_{\text{thr IV}}$)

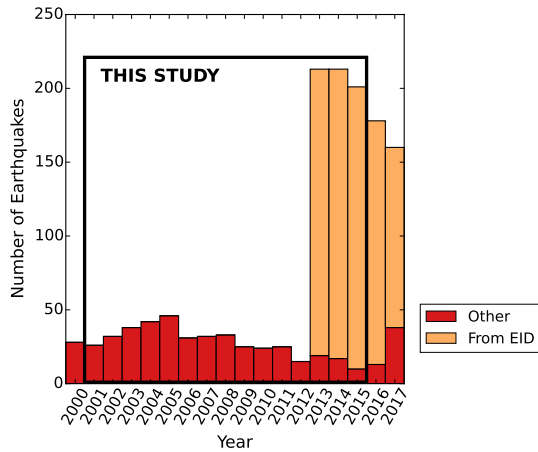
quakes thanks to the increasing popularity and penetration of the Internet and the contribution of the Earthquake Impact Database (EID) from the year 2013 onward (Nievas et al. 2019b, c).

- It is long enough to be of statistical significance and, at the same time, short enough not to pose an excessive computational demand.

Similarly, for what concerns magnitudes, the global earthquake catalogue was first built including earthquakes of all magnitudes but then filtered for the range of interest, defined as $3.95 \leq M_w < 5.55$ to account for the fact that magnitudes are usually reported with one decimal place. A magnitude-dependent maximum depth criterion was then used to keep only the upper-crustal events.

The second main stage of the process was the identification of potentially-damaging earthquakes within the filtered global earthquake catalogue. The purpose of this stage was to eliminate earthquakes happening in very sparsely populated areas, oceans or deserts that clearly pose no threat or minimal threat to human settlements. *Sufficient* proximity to inhabited areas was defined by means of comparing the number of people and the maximum density of people exposed to a certain probability of having observed Modified Mercalli Intensities (MMI) of IV and V against the number of people and density of people frequently used to define an urban settlement, as shown schematically in Fig. 2. Complete distributions of values of MMI were calculated at the geometric centre of each cell of the grids of population and density of Gridded Population of the World GPW v4.0 (CIESIN 2016), using intensity prediction equations (IPEs) and accounting for uncertainty in the value of M_w and hypocentral depth, as well as in the intensity prediction model itself. The number of people in the cells for which the probability of observing $\text{MMI} \geq \text{IV}$ and $\text{MMI} \geq \text{V}$ exceeded a series of thresholds (that were dependent on the magnitude and defined to be consistent with the aforementioned maximum depth criterion adopted) was added up and the maximum population density of those same cells was noted. The earthquake was considered to pose a threat if the population count

Fig. 3 Distribution of dates of the earthquakes that make up the Database of Damaging Small-to-Medium Magnitude Earthquakes (Nievas et al. 2019b, c) for the years 2000–2017 (M_w 4.0–5.5). The rectangle encloses the period 1st January 2001–31st December 2015, which was considered for the present study and encompasses 996 earthquakes



was equal to or larger than 2500 or if the maximum density exceeded 300 people/km², and thus became part of the catalogue of potentially damaging earthquakes.

The third and final stage consisted in identifying the *actually* damaging earthquakes within this last catalogue, which was done through the Database of Damaging Small-to-Medium Magnitude Earthquakes (Nievas et al. 2019b, c). As explained in the cited publications, the database gathers earthquakes in the magnitude range of interest that have reportedly caused damage and/or casualties, noting that the level of reliability of the original sources is undetermined and that the consequences range from slight cracks in non-structural components all the way to collapses, and from injuries due to people escaping in panic to deaths attributed to structural failures. In this sense, the concept of *actually* damaging implies the existence of reports of damage and/or casualties, but not necessarily of verified scientific measurements or observations, or of a certain degree of damage or injury. A point worth noting is that databases of earthquake damage are subject to similar completeness problems to any earthquake catalogue in which availability and accessibility of data play a fundamental role. This is illustrated in Fig. 3, which shows the number of earthquakes contained in the database for each year in the time period of interest for the present work (enclosed by the black rectangle). As can be observed and has been discussed in Nievas et al. (2019b, c), the number of damaging earthquakes identified by the EID is significantly larger than those present in other sources. If the rate of around 190 damaging earthquakes per year observed for 2013–2017 were true for all years before 2013 as well, it would imply that less than 20% of the damaging earthquakes of the 2000s may have been captured in the database (Nievas et al. 2019b, c). Moreover, it is not possible to ascertain whether the EID has, in fact, identified every earthquake causing damage and/or casualties, and it is thus conceivable that the number of damaging events could be larger in reality.

3 Global earthquake catalogue

3.1 Merging of source catalogues

The compilation of earthquake catalogues entails a series of challenges (e.g., Mueller 2019; Weatherill et al. 2016), from which the present work is not exempt. The most

relevant herein are the need to select one origin (i.e., hypocentral location and origin time) and one value of magnitude per earthquake, knowing that different seismological agencies and/or studies often yield different results, and the need to homogenise magnitudes calculated using different scales. The relevance and complexity of these tasks led the Global Earthquake Model (GEM) Foundation to develop a set of tools to aid in the generation of earthquake catalogues and resulted in the global WPG16* catalogue (Weatherill et al. 2016). It was thus decided to take the WPG16* catalogue as a starting point for the present work, with the aim of taking advantage of an existing homogeneous-magnitude catalogue generated by means of transparent and reproducible criteria. The sources of the WPG16* catalogue were the ISC-GEM catalogue (v4.0) (Storchak et al. 2015), the Global Centroid Moment Tensor catalogue (GCMT; Dziewonski et al. 1981; Ekström et al. 2012), the Pacheco and Sykes (1992) catalogue, the Engdahl et al. (1998) Bulletin (ISC-EHB), and the ISC Reviewed Bulletin. The latter was used to retrieve origin and magnitude estimates by the National Earthquake Information Center (NEIC) of the United States Geological Survey (USGS), GCMT (those not found in the original catalogue), Japan's National Research Institute for Earth Science and Disaster Prevention (NIED), and the ISC itself. Moment magnitude values from GCMT and ISC-GEM were retrieved and assigned to earthquakes as reported, while moment magnitudes by NEIC and NIED, as well as body-wave (m_b) and surface-wave (M_s) magnitudes by NEIC and the ISC, were converted into GCMT-equivalent values of moment magnitude by means of empirical conversion models derived from the data available from the catalogue itself (see Weatherill et al. 2016 for further details).

This focus of the WPG16* catalogue on a constrained number of source catalogues/agencies and magnitude scales results in many earthquakes being excluded from it. For this reason, earthquakes reported in the ISC Bulletin that were not already present in the WPG16* catalogue (by not being part of the Reviewed ISC Bulletin) were incorporated to generate the present global earthquake catalogue. The ISC Bulletin was selected as the natural complement of the WPG16* catalogue both because it was one of its fundamental sources and because it gathers and reports origins and magnitudes provided by an extensive number of contributing agencies from around the world. Using the toolkit published alongside the paper of Weatherill et al. (2016), the ISC Bulletin was queried (on 21st September 2018) to retrieve all earthquakes dated between 1st January 1998 and 31st August 2018 with magnitude estimates (in any scale) of at least 2.5 with no restriction on the author, the only condition being that they were not flagged (within the event comments of the ISC Bulletin, methodology suggested by Weatherill et al. 2016) as “explosion”, “chemical” or “nuclear”. The resulting set of earthquakes was then compared against the WPG16* catalogue in terms of event IDs and origin IDs (WPG16* keeps the event IDs of the ISC Bulletin when the origin is retrieved from there), distance between reported hypocentres and difference in origin time. Tolerances for the latter were set to 100 km and 60 s, as locations and origin times estimated by different agencies can easily vary by several tens of km and several tens of seconds, especially because those of all agencies reported in the ISC Bulletin are used for this comparison. Smaller tolerances could easily result in earthquakes from the WPG16* catalogue not being matched properly with their ISC Bulletin counterparts. A more refined search for potentially duplicated entries resulting from these choices was conducted at a later stage, as will be described below. As a general rule, when an earthquake retrieved from the ISC Bulletin was already present in the WPG16* catalogue, the entry from WPG16* was kept. However, if two or more entries from the ISC Bulletin seemed to match one of the WPG16* catalogue, it was assumed that the events had been updated in the ISC Bulletin, and so the entries from the WPG16* catalogue were discarded and those

from the ISC Bulletin were kept. Updates, merges and eliminations of events occur in the ISC Bulletin, firstly, when events from the ISC Bulletin are manually processed to generate the ISC Reviewed Bulletin, and, secondly, any time the ISC identifies an issue with a reported event. As the WPG16* catalogue was compiled with a version of the ISC Bulletin older than the one compared against for this work, such cases were found to exist.

After having identified the events that needed to be incorporated into the catalogue from the ISC Bulletin, one origin and one magnitude estimate needed to be selected per earthquake. The process started with a search for a magnitude estimate in a relevant scale by a relevant agency. Magnitude scales considered were moment magnitude (M_w), surface-wave magnitude (M_s), body-wave magnitude (m_b), local magnitude (M_L), and duration magnitude (M_d)—this order of preference selected based on the observed variability in their scaling with M_w —which were converted into moment magnitude as discussed in Sect. 3.2. Within each scale, estimates from main agencies were preferred over estimates from local/regional agencies, in view of the former offering reliable worldwide coverage and the reliability of the latter being variable for different countries (establishing country-specific rules to prioritise main agencies in some countries and local agencies in some others would have added considerable complexity to the problem). Moreover, this choice ensures consistency with the WPG16* catalogue, which was compiled mostly from global agencies/sources (except for NIED, included to improve coverage in Japan). Agencies/authors classified as “main” herein were the ISC-GEM and ISC-EHB catalogues, the ISC, NEIC/USGS, the GCMT, the European-Mediterranean Seismological Centre (EMSC), the International Data Centre (IDC) of the Comprehensive Nuclear-Test-Ban Treaty Organization (CTBTO), the Geophysical Survey of the Russian Academy of Sciences, the China Earthquake Networks Center, the Experimental (GSETT3) International Data Center and the IASPEI Working Group on Reference Events. All other agencies providing estimates at the country-level were classified as “local” and were ranked according to their relevance within each country. Whenever there was an obvious national organization, it was selected as the most relevant for the country, unless it was already classified as a main agency, which would already take precedence over local ones. For agencies that were not particularly known at the international level and for which it was, thus, not possible to determine their relevance, the information provided within the ISC website regarding their activity and level of contribution was used as a supplementary criterion. Well-established and well-instrumented local networks were ranked next, followed by national and seismic laboratories, as well as state or provincial level agencies. Universities and temporary experiments were considered last. Details on the ranking adopted for all agencies can be found in the report by Nievas et al. (2017).

While the hierarchies defined for main agencies could be directly applied to any earthquake in the world, local agencies had to be selected and sorted as a function of the epicentral location. Average coordinates calculated considering all origin estimates available for a particular earthquake were compared against bounding boxes for countries (Nearby UK 2017). Countries whose bounding boxes contained the average epicentral coordinates were classified as primary countries. Neighbouring countries of relevance were then defined based on the intersection of the bounding boxes of the primary countries with those of other countries. Two categories of neighbouring countries were defined by dilating the bounding boxes by 70 or 500 km. The 70-km dilation intended to acknowledge that an earthquake occurring close to the border between two countries might be at risk of not being considered within all relevant countries as a consequence of the shape of the country profile. The 500-km dilation was defined with the purpose of extending the list of potential agencies, in case no estimation from closer agencies was found first, and was prompted by

the observation of cases for which no location or magnitude could be selected due to the epicentre falling far enough from all relevant bounding boxes (e.g., within the Mediterranean sea). The agencies from primary countries and 70-km neighbours were subsequently ranked going first by ranking and then by country. For example, if country A had agencies A1, A2 and A3, country B had agencies B1 and B2, and country C had agencies C1, C2 and C3, the final hierarchy was A1, B1, C1, A2, B2, C2, A3, C3. Then, the same logic was applied to the agencies from the 500-km neighbours, which were ranked after the former. Within each subgroup of primary countries, 70-km neighbours and 500-km neighbours, countries were ranked in alphabetical order.

A subsequent visual inspection of both the WPG16* catalogue and the catalogue resulting from merging the latter with additional events retrieved from the ISC Bulletin—*merged catalogue*, hereafter—suggested the possible presence of duplicate events, that is, separate entries that are likely to refer to the same earthquake. Such cases can exist due to two main reasons. Firstly, because they may exist in the ISC Bulletin itself as a result of retrieving events from the still-unreviewed Bulletin (the process of grouping information received from all the contributing agencies into events is automatic), the impossibility of manually verifying every single event (the ISC has criteria for the selection of events that are verified), and, most importantly, the variability that exists in estimates of location and origin time from different agencies, of which the previous two are consequences. The second reason for the existence of potentially duplicate events in the merged catalogue is the inherent difficulty in unequivocally identifying the same earthquake across different sources during the merging process, both when the WPG16* catalogue was compared against the ISC Bulletin for the present work and when the WPG16* itself was compiled, due mainly to the aforementioned variability in the location and origin time estimates. An algorithm was thus developed to identify and resolve cases of potentially duplicated events, based on a combination of criteria including the overlapping or not of the origin times when considering the variability stemming from the existence of different estimates, the distance between the hypocentres, the difference in moment magnitude, the listing of the same or different agencies for each event in the ISC Bulletin, and the existence or not of a solution marked as *#PRIME* in the ISC Bulletin for one or both events. Interested readers can refer to the report by Nievas et al. (2017) for details on this algorithm.

By the end of the merging process, the merged catalogue contained 1,616,889 earthquakes for the period 1st January 1998–31st August 2018, of which 376,900 (23.3%) were retrieved from the WPG16* catalogue and 1,239,989 (76.7%) were added from the ISC Bulletin. Their M_w -depth distribution is depicted in Fig. 4.

3.2 Magnitude homogenisation

With moment magnitude (M_w) as the target homogenous magnitude to represent earthquakes in the catalogue, empirical conversion models were used to convert M_s , m_b , M_L and M_d into M_w . For the case of M_s and m_b , average models resulting from the combination of those derived by Scordilis (2006), the Generalised Orthogonal Regression (GOR) models of Di Giacomo et al. (2015) and those of Weatherill et al. (2016) were used. These were selected due to them being fit using robust orthogonal regression techniques on large homogeneously-processed global datasets (on the order of several tens of thousands of events). The uncertainty of the average models was taken as the maximum uncertainty of all involved models, resulting in 0.317 for m_b and ranging between 0.17 and 0.20 in the case of M_s , for which the model is piece-wise linear. The exponential models of Di

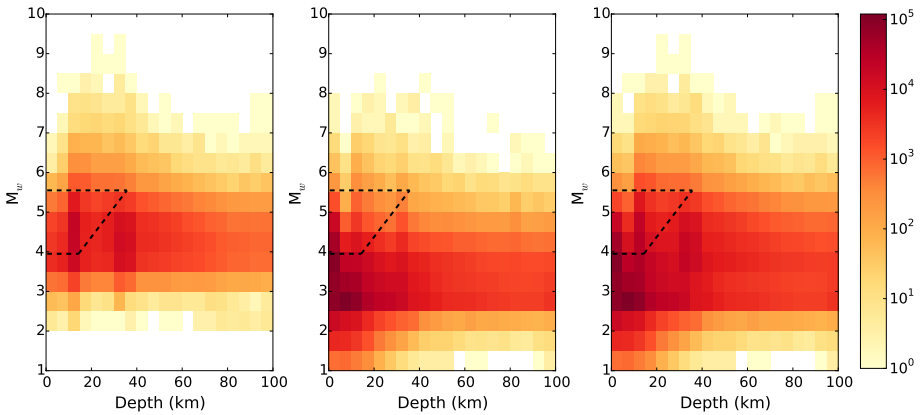


Fig. 4 Moment magnitude-depth distribution of the events retrieved from the WPG16* catalogue (left), the ISC Bulletin (centre) and the whole of the merged catalogue (right). Time period: 1st January 1998–31st August 2018. Number of earthquakes per cell according to colour scale; white cells indicate zero earthquakes. Horizontal axis truncated to 100 km but values go up to 832 km. Black dashed lines enclose the magnitude-depth combinations considered in the present work, discussed further in Sect. 3.3

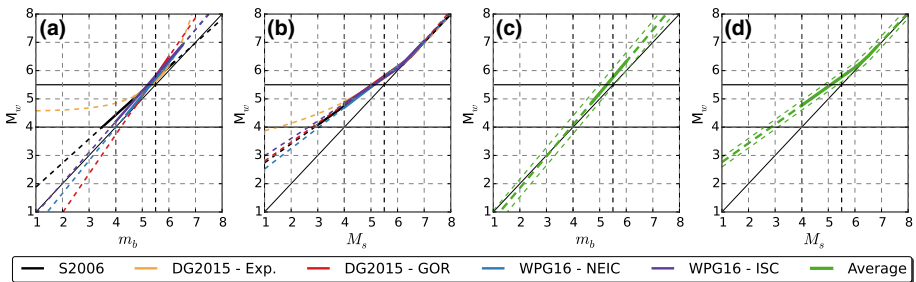


Fig. 5 Empirical conversion models: **a** from m_b to M_w , different models, **b** from M_s to M_w , different models, **c** from m_b to M_w , average model, **d** from M_s to M_w , average model. S2006: Scordilis (2006). DG2015: Di Giacomo et al. (2015) (*Exp.* exponential, *GOR* Generalised Orthogonal Regression). WPG16: Weatherill et al. (2016) (derived separately for m_b and M_s from NEIC and ISC). Thick continuous lines indicate model within its derivation or application range. Thick dashed lines indicate extrapolation of the model beyond the latter. Thin dashed lines indicate \pm one standard deviation in (c, d). Magnitude range 4.0–5.5 indicated

Giacomo et al. (2015) were excluded as they saturate at low values of M_s and m_b , tending to M_w 2.863 and M_w 4.555 when M_s and m_b tend to minus infinity, respectively. Figure 5 shows the original and average models.

From their detailed assessment of models available to convert from local magnitude M_L into moment magnitude M_w , Dost et al. (2018) concluded that the average of the studies conform with $M_w \approx M_L$ for values of M_L above 2.0–3.0. In view of this, and the fact that the relation between M_w and M_L varies regionally (e.g., Scordilis 2006), rendering the selection of more refined models to be applied worldwide trivial, $M_w = M_L$ equivalence was adopted herein. A model uncertainty of 0.25 magnitude units was used for these cases, as this is a reasonable value when contrasted against existing models to convert from M_L to M_w (e.g., Goertz-Allman et al. 2011; Grünthal et al. 2009; Papazachos et al. 1997). Though less reliable than M_L , duration magnitude M_d (based on the duration of the coda of the seismogram) is commonly calculated by local agencies as a proxy for M_L . The poor scaling

with other magnitude scales and the high level of noise that can be observed in the data (Weatherill *pers. comm.* 2016) suggest that a simple assumption of $M_d = M_L = M_w$ is the most reasonable, as more refined analysis do not guarantee any real improvement over this hypothesis. A larger standard deviation of 0.3 was adopted.

The final uncertainty of M_w (σ_{M_w}) stemming from the measurement error reported in the ISC Bulletin (σ_{measure}) and the uncertainty of the conversion model f used ($\sigma_{\text{conversion}}$) was calculated as per Eq. (1) via standard error propagation. Values of 0.1 and 0.3 were adopted as measurement errors for moment magnitude and all other magnitude scales, respectively, when measurement errors were not reported or reported as null in the ISC Bulletin and the WPG16* catalogue.

$$\sigma_{M_w} = \sqrt{\sigma_{\text{conversion}}^2 + \left(\frac{\partial f}{\partial M_{\text{original}}} \right)^2 \cdot \sigma_{\text{measure}}^2} \quad (1)$$

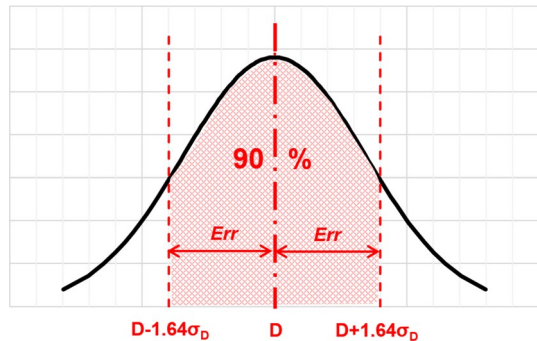
It is noted that the process just described was applied to the events added from the ISC Bulletin and not to those taken directly from the WPG16* catalogue, as the latter already provides moment magnitude and uncertainty for each earthquake it contains (except for M_w values retrieved from the ISC-GEM and the GCMT catalogue, to which a measurement error of 0.1 was assigned in the present work), the latter calculated as per Eq. (1) as well.

While full probability distributions of magnitude, represented by a normal distribution, were used for the calculation of estimated values of intensity, expected (median) M_w values were used for determining whether earthquakes were included or not in the catalogue, for declustering purposes, and for determining maximum depths as per the criteria described in the following section.

3.3 Focal depths

Hypocentral depths are relevant for this work for two main reasons. Firstly, because they define whether an earthquake can be considered upper-crustal or not and, secondly, because they have a significant influence over the resulting ground shaking and seismic intensities (i.e., the potential for the earthquake to be damaging). Estimates of depth were retrieved both from the WPG16* catalogue and the ISC Bulletin, in the case of the latter according to the hierarchies of agencies by means of which origins were selected. Based on the description of the IASPEI Seismic Format (ISF) available on the ISC website and on the work of Bondár and Storchak (2011), the depth errors (*Err*) reported in the WPG16* catalogue and the ISC Bulletin were interpreted as being the distance from the median reported depth (D) to either side of the 90%-confidence interval resulting from the vertical projection of the four-dimensional location error ellipsoids, as schematically depicted in Fig. 6. As such, the standard deviation of a normal distribution centred around the reported depth that can be used to describe the uncertainty in the depth estimate is calculated as $Err/1.64$. Strictly speaking, this distribution should be truncated at a depth of 0 km, but the documentation from the ISC suggests that the reported error is calculated in an unbounded space. In the present work, a truncated normal distribution was used when calculating the probability of the depth being equal to or smaller than a certain limit, but the standard deviation of the unbounded distribution (σ_D) was used for the propagation of uncertainties, as will be explained in Sect. 4.1, to avoid inconsistency with the original estimate of the ISC. These same assumptions were applied to estimates authored by contributing agencies.

Fig. 6 Schematic representation of the relationship between the depth error reported in the ISC Bulletin (Err) and the standard deviation of a normal distribution describing the uncertainty in the depth estimate, as interpreted for this work



However, depth can only be resolved when the available records contain sufficient information, and it is not uncommon for the depth to be fixed to pre-defined values in order to locate an earthquake. The ISC, for example, only attempts a free-depth solution when at least one of four pre-defined criteria are met, and otherwise fixes the depth to a default value (Bondár and Storchak 2011), although in the past a free-depth solution was always initially sought (Adams et al. 1982). The way in which fixed depths are selected has changed in time as well, alternatives including the use of a depth determined by means of the depth-phase stack technique, or that reported by an external agency, or a conventional value (typically 0, 10 or 33 km), or a location-specific value derived from statistical data of past events of the ISC, as is currently done (Bondár and Storchak 2011). The USGS currently adopts 10 km as a fixed depth, although it used to adopt 33 km in the past (USGS, 2018). Within both the WPG16* catalogue and the ISC Bulletin, a depth solution was considered fixed if either flagged as fixed or lacking a reported depth error or having been assigned a 0-km depth error.

This distinction between free- and fixed-depth solutions was relevant for the determination of whether an earthquake was upper-crustal or not. To begin with, judgement-based magnitude-dependent maximum depths were defined, as shown in the plot on the left of Fig. 7, and were applied taking as a reference the expected value of M_w . Fixed-depth solutions were directly compared against these limits, earthquakes becoming part of the catalogue filtered by depth when the former were equal to or smaller than the latter. In the case of free-depth solutions, considerations had to be made to account for the existence of disproportionately large reported depth errors, that is, for pairs of depth-uncertainty that yield very large values of the coefficient of variation (e.g., a 200 km error reported for a depth of 5 km). After analysing the probabilities of an earthquake with depth D and associated error σ_D being equal to or smaller than a certain depth limit D_{lim} , it was decided to consider an earthquake passing the maximum depth criterion if the probability of D being equal to or smaller than D_{lim} was at least 50%. As can be observed in the plot on the right of Fig. 7, this condition can be satisfied by either earthquakes with small values of D (in comparison with D_{lim}), even if their associated uncertainty is large, or earthquakes with values of D close to D_{lim} with small uncertainties. In order to be consistent with this analysis, earthquakes with fixed-depth solutions were assigned the maximum standard error that caused the probability of D being smaller than or equal to D_{lim} (for the corresponding expected magnitude) to be at least of 50%. In other words, they were assigned the ratio of σ_D to D that corresponded to the 0.5-contour of Fig. 7 for their ratio D/D_{lim} .

It is noted that while full probability distributions of depth, represented by a truncated normal distribution, were used for the calculation of estimated values of intensity, expected

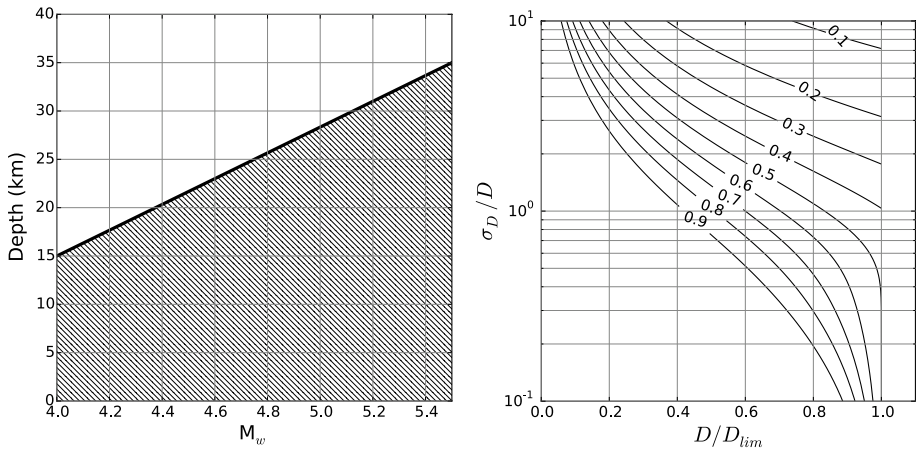


Fig. 7 Magnitude-dependent maximum depth limits (left) and probability of the depth D being equal to or smaller than the maximum-depth limit D_{lim} as a function of the ratio of reported depth to depth limit (D/D_{lim}) and depth error to reported depth (σ_D/D), for a truncated normal distribution with 0-km depth upper bound (right)

values of depth as reported in the ISC Bulletin and the WPG16* catalogue were used for declustering purposes.

3.4 Declustering

Understanding the extent to which small-to-medium magnitude earthquakes can result in consequences for the population or the built environment should, ideally, stem from analysing events in which all the reported consequences are due to one specific and perfectly identifiable event. However, earthquakes often occur as parts of sequences and, as explained in Nievas et al. (2019b, c), determining the amount of damage or number of casualties attributable to any particular earthquake of a sequence or the extent to which previous shocks had an influence or not on the damage observed after a later shock is extremely complex. This does not depend just on the magnitude of the events and their distribution in space and time, but also on the speed with which damaged buildings are retrofitted or replaced, the quality and efficacy with which this is done, and the extent of previously-unrepaired damage. Depending largely on the socio-economic setting, all of these are variables that are difficult to incorporate in an analysis like the one presented herein. However, the simplified assumption that main shocks—or main shocks together with foreshocks—could represent earthquakes for which previous damage conditions did not exist can be of use for assessing the impact of earthquakes without the influence of other events. The accuracy of such an assumption is clearly debatable but is adopted herein due to its practicality for a global study.

Identification of foreshocks, main shocks and aftershocks, the process commonly known as *declustering*, was carried out herein using the algorithm of Gardner and Knopoff (1974), as implemented in the OpenQuake Hazard Modeller's Toolkit (Pagani et al. 2014; Weatherill 2014), modified to take into consideration hypocentral depth. Gardner and Knopoff (1974) time and distance windows were used, and the same time span was considered for

both foreshocks and aftershocks. As with any declustering algorithm, the resulting classification should be interpreted with care, as algorithms of this kind work under assumptions of temporal and spatial clustering and are not informed by data on geophysical features such as active faults.

4 Catalogue of potentially damaging events

After having declustered the merged catalogue and filtering it for the time period and magnitude range of interest, as well as according to the maximum depth criterion, 141,524 earthquakes were retained from the original 1,616,889 (8.7%). Of the resulting 141,524 events, 51,969 (36.7%) were classified as main shocks, 27,192 (19.2%) as foreshocks, and 62,363 (44.1%) as aftershocks. The next stage then deployed this set of 141,524 earthquakes and aimed at determining which of them could be considered as potentially damaging, based on the number and density of people likely to have been exposed to intensities equal to or larger than MMI IV and V, as schematically shown in Fig. 2.

4.1 Prediction of Modified Mercalli Intensities

Complete distributions of Modified Mercalli Intensities were calculated for the surroundings of each earthquake by means of the IPEs of Atkinson and Wald (2007), which comprise two sets of coefficients, one for the Central and Eastern United States (CEUS) and another for California. These two represent, respectively, a stable continental region (SCR) with very low attenuation (e.g., Bakun and McGarr 2002), and an archetypal active crustal region (ACR). The low attenuation characteristic of SCRs translates into larger macroseismic intensities being observed at larger distances than would occur in an ACR. Within the whole range of SCRs, the attenuation level is closely related to how cratonic or not the region is, as cratonic areas are typified by old, hard rocks that are better at transmitting seismic waves than younger ones. As Chen et al. (2018) point out, determining with absolute accuracy what is a cratonic environment and what is not can be challenging at a global scale. Acknowledging this, Chen et al. (2018) have proposed a fuzzy approach to this classification that ultimately led to what they call the craton index (CI) and which they mapped for the whole world at 0.5° intervals. The CI is a quantitative measure of how cratonic ($CI=1$) or non-cratonic ($CI=0$) a site is, and is thus useful for this work for it allows to distinguish simultaneously between extremely cratonic SCRs like the CEUS and less-cratonic SCRs like north-western Europe, or even ACRs, as the latter are, by definition, non-cratonic. Moreover, it allows one to do so with a certain degree of smoothness in the transitions from one to the other. This last point is very important, as a rigid fixed boundary between SCRs and ACRs, or between more cratonic and less cratonic SCRs, could lead to severe inconsistencies based on moving very short distances to one side or another of the boundary.

Based on these considerations, the approach adopted consisted in using the craton index to weigh the values of MMI calculated using the Atkinson and Wald (2007) IPE for ACRs (MMI_{ACR}) and for SCRs (MMI_{SCR}), as per Eq. (2). This weighting scheme assumes that MMI_{ACR} and MMI_{SCR} represent the two possible ends of the spectrum of macroseismic intensities to be expected. It is noted that, while this approach may appear to establish an equivalence between ACRs and non-cratonic SCRs, this assumption is only applied herein in terms of attenuation properties. It is possible that non-cratonic SCRs still have

differences in attenuation when confronted with ACRs, but the approximation is deemed sufficient for the purposes of this work. The value of CI used was that corresponding to the epicentral coordinates of each earthquake.

$$MMI_{final} = CI \cdot MMI_{SCR} + (1 - CI) \cdot MMI_{ACR} \quad (2)$$

Each of the two models used to calculate MMI_{SCR} and MMI_{ACR} has an associated standard deviation, stemming from the variability of the data around the model itself ($\sigma_{IPE_{SCR}}$ and $\sigma_{IPE_{ACR}}$, for the IPE using the coefficients for SCRs and ACRs, respectively) and the uncertainties in the values of depth and magnitude [σ_D and σ_{M_w} , respectively, the latter as per Eq. (1)]. The total standard deviation was thus calculated as per Eq. (3), in which the derivative terms refer to that of the IPE with respect to depth ($\partial IPE / \partial D$) and magnitude ($\partial IPE / \partial M_w$). As Atkinson and Wald (2007) do not provide a model for the standard deviation and only mention an average standard deviation of the residuals of 0.4 MMI units, which appears as small in comparison with those of other models (e.g., Allen et al. 2012; Văcăreanu et al. 2015), the model for the standard deviation of Allen et al. (2012) was used instead. The latter yields values that vary between 0.8 and 1.2 MMI units for the range of depths of interest in this work.

$$\sigma_{MMI_{ACR}} = \sqrt{\sigma_{IPE_{ACR}}^2 + \sigma_D^2 \cdot \left(\frac{\partial IPE_{ACR}}{\partial D} \right)^2 + \sigma_{M_w}^2 \cdot \left(\frac{\partial IPE_{ACR}}{\partial M_w} \right)^2} \quad (3)$$

The final uncertainty was then calculated as per Eq. (4), which is analogous to Eq. (2). This linear combination of uncertainties was preferred over its theoretically-accurate calculation as the linear combination of two random variables because the latter presented an undesired non-physical behaviour. It is not uncommon that $\sigma_{MMI_{ACR}}$ and $\sigma_{MMI_{SCR}}$ are numerically similar. In the extreme case in which they were the same, the theoretically-correct final uncertainty would result in the uncertainty in the prediction of MMI being smaller when the site is neither fully cratonic nor fully non-cratonic and at its largest when the site lies on one of the two extremes, a behaviour which clearly holds no physical meaning.

$$\sigma_{MMI_{final}} = CI \cdot \sigma_{MMI_{SCR}} + (1 - CI) \cdot \sigma_{MMI_{ACR}} \quad (4)$$

The uncertainty in moment magnitude, hypocentral depth and that associated with the intensity prediction model are not the only sources of uncertainty in the problem of estimating earthquake intensity. Horizontal errors in epicentral location, for example, were not incorporated in the present work due to the intrinsic difficulty that results from the need to define a distribution of epicentral distances based on the error ellipse of the epicentral coordinates, which would be a function of the relative location of the site with respect to the ellipse.

The influence of site effects and soil conditions has not been considered in the present work for two reasons. Firstly, because the IPEs of Atkinson and Wald (2007) do not include terms that represent site effects. Secondly, because doing so requires the use of a world model of site conditions, the kind of which only exists in terms of proxy parameters (e.g., the slope-based proxy V_{s30} world model of Wald and Allen 2007), and thus results in the incorporation of further uncertainties. Moreover, the purpose of this work is not to

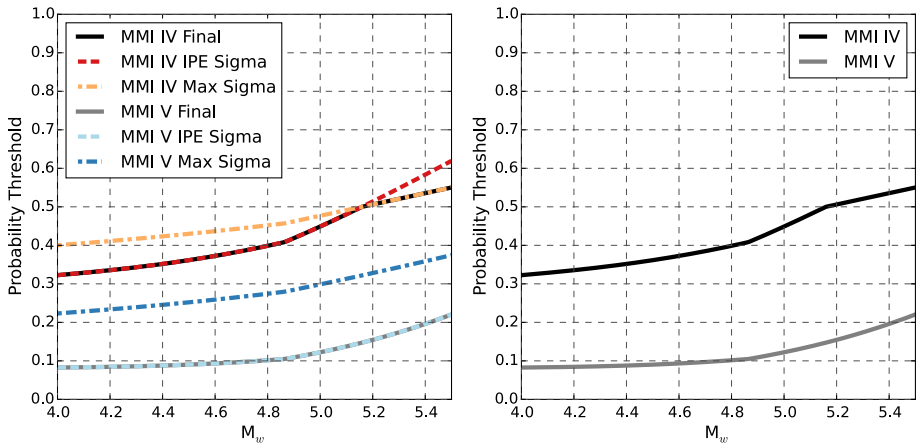


Fig. 8 Magnitude-dependent probability thresholds for $MMI \geq IV$ and $MMI \geq V$: calculation by components (left) and finally-adopted values (right). “IPE Sigma” and “Max Sigma” refer to the case of only considering the uncertainty of the intensity prediction equation versus considering a maximum feasible uncertainty, respectively

determine in detail the intensity potentially generated by each earthquake of the catalogue at each site but to be able to classify earthquakes as potentially damaging or not, based on their proximity to population and assets. Similarly, the influence of style-of-faulting on resulting ground motions—and, consequently, intensities—has not been included in this work due to it neither being considered within the IPEs of Atkinson and Wald (2007) nor focal mechanisms being available for all earthquakes in the catalogue.

4.2 Estimation of exposed population and exposure criteria

The previous step results in the intensity predicted to have been caused by a specific earthquake at a site being defined in terms of a distribution with median and standard deviation calculated by means of Eqs. (2) and (4). The sites selected for these calculations were the geometrical centroids of the cells of the grid of population and density of Gridded Population of the World GPW v4.0 (CIESIN, 2016). Each cell was counted as being exposed to potentially damaging shaking if the probabilities of observing MMI values of IV and V or larger exceeded the pre-defined thresholds defined in the plot on the right of Fig. 8, whose derivation will be explained below. For example, for an earthquake with M_w 4.0 a cell needed to have a probability of observing $MMI \geq IV$ of at least 32% and a probability of observing $MMI \geq V$ of at least 8% in order to be counted, while the thresholds climb up to 55% and 22% for an earthquake with M_w 5.5. The population count of grid cells complying with both probability thresholds was added up and their population density was noted, the largest of all being finally selected. The earthquake was then considered to have been potentially damaging if the population exposed to $MMI \geq IV$ and $MMI \geq V$ with probabilities of at least those defined in Fig. 8 was at least 2500 people and/or the maximum population density of all compliant grid cells was equal to or larger than 300 people/km².

The probability thresholds shown in Fig. 8 were defined on the premise that an earthquake occurring at the maximum depth (for its magnitude) indicated by the plot on the left of Fig. 7 needed to be able to cause these probabilities at least in the cell that contained its epicentre. Figure 9 illustrates the probabilities of observing $MMI \geq IV$ and $MMI \geq V$

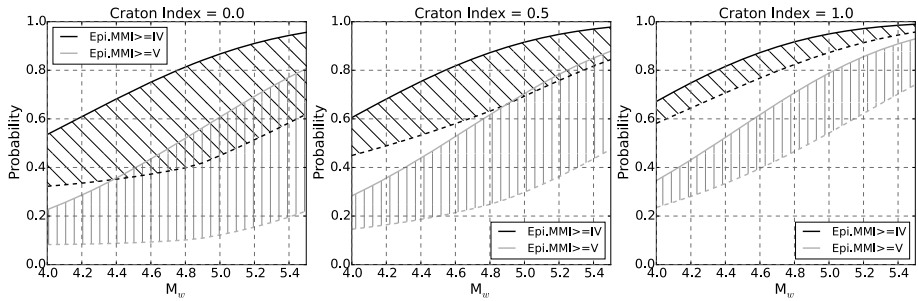


Fig. 9 Range of probabilities of observing MMI values of IV (black) or V (grey) and above at the epicentre for the whole range of hypocentral depths considered herein, calculated as per Sect. 4.1 considering only the uncertainty of the IPE (i.e., no uncertainty in M_w or depth). Continuous and dashed lines correspond to 0-km depth and maximum depth limits (defined by the plot on the left of Fig. 7), respectively

right above the epicentre for the whole range of depths considered at each magnitude (0-km depth in continuous lines, maximum depth limits in dashed lines), with the only uncertainty being the variability of the IPE ($\sigma_{IPE\ SCR}$ and/or $\sigma_{IPE\ ACR}$). Selecting probability thresholds larger than those defined in the plot on the right of Fig. 8 would be equivalent to establishing more restrictive depth limits than those shown in Fig. 7. For example, a 50% probability of observing $MMI \geq V$ at the epicentre due to earthquakes with hypocentral depths equal to their corresponding maximum depth limits would be impossible to achieve for all magnitudes (M_w 4.0–5.5) for values of the craton index equal to or smaller than 0.5. While the plots in Fig. 9 only consider the uncertainty in the IPE, the complete uncertainty also includes the uncertainties in depth and magnitude, σ_D and σ_{M_w} , as per Eq. (3). The probabilities of observing a particular value of MMI are clearly affected by an increase in the final uncertainty. In order to overcome this limitation, the probability thresholds were defined as the minimum of that resulting from considering the standard deviation of the IPE alone and that resulting from incorporating maximum feasible uncertainties for depth and magnitude. As this analysis involves the maximum depths considered (D_{lim}), the uncertainty in depth cannot be very large, as the 0.5-probability contour line becomes asymptotic at a D/D_{lim} value of 1.0 in the plot on the right of Fig. 7, and thus a value of 0.001% of D was adopted. For the case of magnitude, a final uncertainty [resulting from combining measurement errors with the uncertainty of conversion models as per Eq. (1)] of 1.657 was used, as this is the largest value found in WPG16* for the time period of interest and is larger than the largest found in the merged catalogue as well. The way in which the final thresholds are calculated is depicted in the plot on the left of Fig. 8, in which “IPE Sigma” and “Max Sigma” refer to the two alternative levels of uncertainty considered. As can be observed, for $MMI \geq V$ the final threshold is that calculated only with the standard deviation of the IPE, as the expected MMI does not reach V for any value of M_w , while for $MMI \geq IV$, the final thresholds are those calculated from a combination of both levels of uncertainty. The thresholds were calculated using a CI of zero, as this leads to the lowest probabilities (see Fig. 9) and guarantees consistency with the maximum depth limits in all environments.

The 2500 people and/or 300 people/km² thresholds were established by investigating the definitions of urbanisation used by different sources, all of which highlight that defining what an urban settlement is and making the distinction between urban and rural populations is neither trivial nor objective. According to the definitions of UNICEF (2012), at

least 2000 people define an urban settlement, though the number varies greatly around the world and can range between 200 and 50,000. For the 2010 census, the United States Census Bureau defined an urbanised area as that having 50,000 people or more and a density of at least 1000 people per square mile (386 people/km²), and an urban cluster as that having between 2500 and 50,000 people. In the European Union, urban areas are defined by first identifying grid cells of 1 km² in which the population density is equal to or larger than 300 people/km², grouping adjoining cells that satisfy this criterion and verifying that the resulting group of cells adds up to, at least, 5000 people (Eurostat; Dijkstra and Poelman 2018). Recognising the relevance of the size of the grid used to calculate the density (highlighted by many of the sources), the maximum density of at least 300 people/km² exposed to the probability thresholds of MMI defined in Fig. 8 was adopted, as Eurostat specifies the grid cell size and it is consistent with that of GPW v4.0 at the Equator, though the 30 arc-second resolution of GPW implies that the area of the cells decreases when moving towards the Poles. The minimum cumulative population count of 2500 was selected due to it being the threshold to define an urban cluster according to the United States Census Bureau, and being only slightly higher than that used by UNICEF (2012). The threshold of 5000 people set by the European Union was deemed too high in comparison with the alternatives found and likely to be unrepresentative of the conditions in many other parts of the world.

It is noted that this approach implicitly considers that an earthquake whose depth is the maximum contemplated for its magnitude (according to the plot on the left of Fig. 7) is potentially damaging if the population count or density right at the epicentre complies with the 2500 people and/or 300 people/km² thresholds. This is equivalent to imagining a circle of radius tending to zero in the plot on the right of Fig. 2. If the earthquake is shallower, the radius of the circle is greater and more population cells fall within it. In this sense, the choice of the specific MMI levels to consider is not of major influence, as the key lies in the probability thresholds being defined as those right at the epicentre of an earthquake with maximum depth occurring in a non-cratonic environment.

4.3 Sensitivity analyses

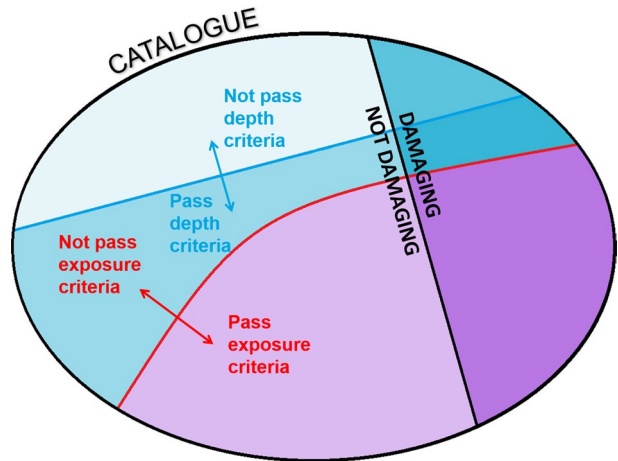
This section explores the influence of three (sets of) variables on the final proportion of damaging to total earthquakes: (1) the thresholds of population count and density, (2) the parameters used within the declustering algorithm, and (3) the depth criterion. The sensitivity analyses were conducted over a preliminary version of the catalogue and the Database of Damaging Small-to-Medium Magnitude Earthquakes (Nievas et al. 2019b, c) as well, which did not contain the earthquakes incorporated directly from the EID. This preliminary version differed from the final one in many ways, including (but not limited to) the study of a slightly different time period (1st July 1999–30th June 2014), the use of a step function to define the maximum depth limits, the consideration of only M_w , M_s and M_L magnitude scales without conversion models, the use of a different IPE, no craton index and no uncertainties in the calculation of MMI, and the use of a probability of at least 50% of observing $MMI \geq IV$ as the condition to count a grid cell or not. Details on the preliminary version and on the sensitivity analyses can be found in Nievas et al. (2017, 2019a). The sensitivity analyses have been used to inform many of the decisions made in the final work presented herein.

The extent to which each variable affects the final results was measured in terms of the variability of the proportion of damaging to total number of earthquakes for a series of cases, reported in the four right-most columns of the results tables, of which Table 1 below

Table 1 Generic table for the comparison of alternatives of the sensitivity analyses

Case	Catalogue				Damaging				Damaging/Catalogue			
	Pass depth		Pass exposure		Pass depth		Pass exposure		Pass depth		Pass exposure	
	All	MS	All	MS	All	MS	All	MS	All	MS	All	MS
X	A	B	C	D	E	F	G	H	E/A	F/B	G/C	H/D

Fig. 10 Schematic representation of filtering levels used in Table 1



is a generic representation. Columns labelled “MS” indicate that only main shocks are considered, while those labelled “All” indicate that all earthquakes are considered, irrespective of their classification as fore-, main- or aftershocks. The meaning of all other subdivisions of the catalogue used in Table 1 are illustrated in Fig. 10. The four right-most columns of Table 1 present the proportion of damaging to total earthquakes in each case, as indicated by the ratios E/A, F/B, etc., which make reference to table positions A through H.

4.3.1 Thresholds of population count and density

As discussed in Sect. 4.2, defining the population counts and densities that make the distinction between an urban settlement and more scarcely-populated rural areas is neither trivial nor objective. For this reason, a series of alternative values were considered herein, as shown in Table 2, all under the same condition of expected $MMI \geq IV$ for the grid cells to be counted. Criterion labelled C1 is the benchmark (“BMK” herein and in everything that follows) and was used for all other sensitivity analyses. Criteria C1L and C1H stem from increasing/decreasing the thresholds of C1 by 10%, while criteria C2 through C6 explore some of the other thresholds discussed in Sect. 4.2, combined with drastic reductions in either of the two. Finally, C7 represents the most extreme case possible, that in which just one person is sufficient to deem the earthquake potentially damaging. Results are shown in Table 3.

Leaving the extreme C7 aside for the time being, the largest increase in the number of earthquakes that satisfy the exposure criterion with respect to the benchmark occurs for C6

Table 2 Alternatives considered in the sensitivity analyses on the population count and density thresholds

Criteria	Population count (ppl)	Max. density (ppl/km ²)
BMK = C1	2500	300
C1L	2250	270
C1H	2750	330
C2	1000	300
C3	2500	150
C4	2500	400
C5	5000	300
C6	1000	150
C7	1	1

and represents 5.6% of the total and 5.3% of the main shocks, followed by C2, which represents 5.2% of the total and 4.7% of the main shocks. The largest decrease occurs for C5 and represents 5.4% of the total and 4.7% of the main shocks. All these suggest that it is the population count threshold that has the largest impact in the results. However, the overall change in the number of earthquakes that pass the exposure criterion is not significant for criteria C1 through C6, even when taking very low values of both the population count and maximum density thresholds (e.g., C6). The proportion of damaging to total number of earthquakes varies between 0.75 and 0.84% for all kinds of shocks and between 1.47 and 1.62% for the main shocks with these criteria. The population and density thresholds of the benchmark were finally adopted for this work, as they have been both rationally selected and have proven to produce results that are close to the average of the upper and lower bounds of the ranges resulting from criteria C1 through C6.

Results for C7 are, as would be expected, the most extreme deviation from the benchmark. Interestingly, though, consideration of just a minimum of 1 person exposed to estimated MMI values of IV or larger only increases the percentage of earthquakes that pass both the maximum depth and exposure criteria with respect to those that pass the maximum depth one from 35.2 to 41.4% (36.4% to 42.5% when only main shocks are considered). In other words, only an additional ~6% of the earthquakes that pass the maximum depth criterion are expected to have had an exposure to $\text{MMI} \geq \text{IV}$ of at least 1 person. Due to this increase in the number of earthquakes in the catalogue, the proportion of damaging to total earthquakes decreases to 0.68% for all kinds of shocks and 1.33% for main shocks. While these numbers are more meaningfully different from the benchmark than all other criteria listed in Table 3, they show that even extreme criteria do not cause unreasonable variations in the final results.

4.3.2 Declustering

As depicted in Table 4, the algorithm of Gardner and Knopoff (1974) requires three parameters: the kind of distance window, the kind of time window and the foreshock parameter (fs). Three distance and time window models are available in the OpenQuake Modeller's Toolkit (Pagani et al. 2014, Weatherill 2014): those proposed by Gardner and Knopoff (1974), those of Grünthal (as reported in van Stiphout et al. 2012) and those of Uhrhammer (1986), all of which are magnitude-dependent. The implementation in OpenQuake requires the same kind of window (i.e., by the same author) to be applied both in time and

Table 3 Results for the sensitivity analysis over the population count and density thresholds

Case	Catalogue				Damaging				Damaging/Catalogue			
	Pass depth		Pass exposure		Pass depth		Pass exposure		Pass depth		Pass exposure	
	All	MS	All	MS	All	MS	All	MS	All (%)	MS (%)	All (%)	MS (%)
BMK	101,248	32,842	35,654	11,968	287	189	282	185	0.28	0.58	0.79	1.55
C1L	101,248	32,842	35,981	12,052	287	189	282	185	0.28	0.58	0.78	1.54
C1H	101,248	32,842	35,377	11,881	287	189	282	185	0.28	0.58	0.80	1.56
C2	101,248	32,842	37,503	12,534	287	189	282	185	0.28	0.58	0.75	1.48
C3	101,248	32,842	35,921	12,075	287	189	282	185	0.28	0.58	0.79	1.53
C4	101,248	32,842	35,557	11,925	287	189	282	185	0.28	0.58	0.79	1.55
C5	101,248	32,842	33,733	11,403	287	189	282	185	0.28	0.58	0.84	1.62
C6	101,248	32,842	37,658	12,599	287	189	282	185	0.28	0.58	0.75	1.47
C7	101,248	32,842	41,918	13,963	287	189	283	186	0.28	0.58	0.68	1.33

The numbers of damaging events for BMK are slightly different from those in Tables 5 and 6 due to the presence of potentially duplicated events and the change in the order in which the depth and exposure filtering was carried out for the sake of decreasing the computational demand that repeating calculations incurs

space, reducing the three input parameters to effectively two. The foreshock parameter (f_s) is used to define whether foreshocks are sought ($f_s > 0$) or not ($f_s = 0$). A value between zero and unity indicates the fraction of the aftershocks time window that is used to identify foreshocks, a value of unity meaning that the same time is applied for both. Values of 0.5 and 1.0 were considered herein.

The results shown in Table 5 reveal that the Uhrhammer (1986) windows produce the largest number of events identified as main shocks, followed by the Gardner and Knopoff (1974) windows and, finally, the Grünthal windows (van Stiphout et al. 2012). As the variation is larger for the whole of the catalogue than for the damaging subset, the proportion of damaging to total events is the lowest when the Uhrhammer (1986) windows are used and the highest for the Grünthal windows (van Stiphout et al. 2012). The foreshock parameter (f_s) does not appear to have a significant influence on the results from Gardner and Knopoff (1974), the largest impact occurring in combination with the Grünthal windows. As can be observed, even if the final proportions of damaging to total earthquakes are influenced by the choices made during the declustering process, it is noted that all numbers remain within a reasonable order of magnitude: between 0.43 and 0.74% when filtering only according to depth and between 1.23 and 1.93% when filtering according to exposure as well.

In view of the above, Gardner and Knopoff (1974) time and space windows and a foreshock parameter of 1.0 were selected to be applied to the final version of this work, as results from this combination lie reasonably within the middle of the sensitivity study, and these time and space windows were those originally proposed alongside the algorithm. It cannot be over-emphasised, however, that there is no right answer in terms of declustering algorithms. Being based only on empirically-derived temporal and spatial windows and ignorant of fault geometry and other physical constraints, one can only hope for a sufficiently reasonable classification of earthquakes into main shocks, foreshocks and aftershocks.

Table 4 Alternatives considered in the sensitivity analyses on the parameters used for declustering

Case	Algorithm	Distance window	Time window	fs
BMK	Gardner and Knopoff (1974)	Gardner and Knopoff (1974)		1.0
2	Gardner and Knopoff (1974)	Gardner and Knopoff (1974)		0.5
3	Gardner and Knopoff (1974)	Uhrhammer (1986)		1.0
4	Gardner and Knopoff (1974)	Uhrhammer (1986)		0.5
5	Gardner and Knopoff (1974)	Grünthal (van Stiphout et al. 2012)		1.0
6	Gardner and Knopoff (1974)	Grünthal (van Stiphout et al. 2012)		0.5

Table 5 Results for the sensitivity analysis over the declustering algorithm and parameters

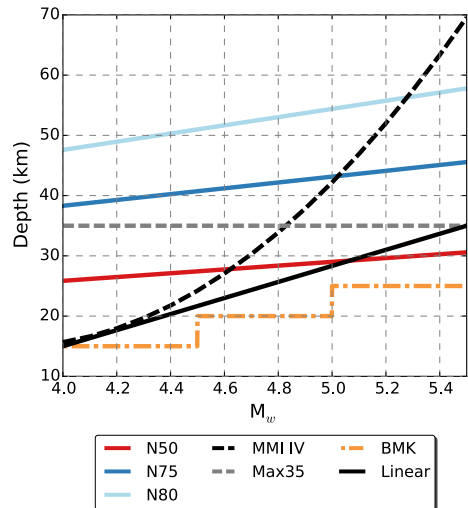
Case	Catalogue				Damaging				Damaging/Catalogue			
	Pass depth		Pass exposure		Pass depth		Pass exposure		Pass depth		Pass exposure	
	All	MS	All	MS	All	MS	All	MS	All	MS	All	MS
									(%)	(%)	(%)	(%)
BMK	101,248	32,842	35,654	11,968	285	188	280	184	0.28	0.57	0.79	1.54
2	101,248	35,354	35,654	12,735	285	197	280	193	0.28	0.56	0.79	1.52
3	101,248	48,845	35,654	16,799	285	213	280	209	0.28	0.44	0.79	1.24
4	101,248	50,384	35,654	17,270	285	216	280	212	0.28	0.43	0.79	1.23
5	101,248	22,845	35,654	8564	285	168	280	165	0.28	0.74	0.79	1.93
6	101,248	25,399	35,654	9435	285	179	280	175	0.28	0.70	0.79	1.85

4.3.3 Maximum depth criteria

Sensitivity regarding the maximum depth was conducted using seven alternative maximum depth criteria, as plotted in Fig. 11. The curve labelled BMK is the step function that was used for the compilation of the preliminary catalogue used for all other sensitivity analyses. N50, N75 and N80 stem from fitting a linear regression to the 50th, 75th and 80th percentiles of depth distributions of the earthquakes contained in the Significant Earthquake Database of the National Centers for Environmental Information of the National Oceanic and Atmospheric Administration (NOAA) of the United States (NOAA database hereafter), while Max35 represents a magnitude-independent maximum of 35 km. Finally, MMI IV represents the maximum depths that allow for the epicentre to have an MMI of at least IV, calculated using the IPE of Allen et al. (2012). Details on the reasons for selecting these alternatives can be found in the report by Nieves et al. (2019a).

As opposed to the two sensitivity analyses already presented, for which a probability of at least 50% of observing $\text{MMI} \geq \text{IV}$ was the condition to count a population grid cell or not, probability thresholds of observing $\text{MMI} \geq \text{IV}$ were defined herein as a function of magnitude, so as to guarantee that an earthquake occurring at the maximum depth considered for its magnitude was able to cause these probabilities at least in the cell that contained its epicentre. This is equivalent to what was done to define the final probability thresholds depicted in Fig. 8 and serves the purpose of not capping the maximum depths alternatives

Fig. 11 Alternatives considered in the sensitivity analyses on maximum depth criteria



by over imposing a more restrictive exposure criterion. Results labelled BMK* in Table 6 correspond to those obtained using the 50%-probability criterion.

The results presented in Table 6 show that while the number of earthquakes included in the database increases when the maximum depth criterion is more relaxed (i.e., when greater depths are considered), the number of non-damaging events increases faster than the number of damaging ones, causing the proportion of damaging to total to drop from 0.28% (all) and 0.57% (only main shocks) to 0.17% and 0.41% in the most extreme cases. This tendency is maintained when adding the exposure criterion. As N75, N80 and Max35 allow for greater depths to be considered for smaller magnitudes (in contrast, for example, with MMI IV, which does the opposite), these three criteria result in a larger number of smaller-magnitude earthquakes with lower damaging capacity being added to the world catalogue that ultimately yields smaller proportions of damaging earthquakes.

The Linear criterion was selected to be applied to the final version of this work due to it yielding conservative results (the largest proportion of damaging earthquakes) and being consistent with the definition of crustal earthquakes, as inferred from both existent mapped depths to Moho (e.g., Artemieva and Thybo, 2013) and reported seismogenic thicknesses of shallow crustal earthquake ruptures (e.g., Stafford, 2014).

5 Statistics of damaging earthquakes in the M_w 4.0–5.5 range

5.1 Resulting catalogue

The resulting catalogue of potentially damaging events—that is, earthquakes deemed as potentially capable of causing damage or casualties based on the criteria defined in the previous sections—consists of 39,127 (27.6%) of the 141,524 earthquakes that comprise the global earthquake catalogue. Their geographical distribution is depicted in Fig. 12, in which epicentres marked in light grey are those of the global catalogue that do not meet the exposure criteria and are thus not deemed potentially damaging. As depicted in the

Table 6 Results for the sensitivity analysis over the maximum depth criterion

Case	Catalogue				Damaging				Damaging/Catalogue			
	Pass depth		Pass exposure		Pass depth		Pass exposure		Pass depth		Pass exposure	
	All	MS	All	MS	All	MS	All	MS	All (%)	MS (%)	All (%)	MS (%)
BMK*	101,248	32,842	35,654	11,968	285	188	280	184	0.28	0.57	0.79	1.54
BMK	101,248	32,842	28,954	9932	285	188	253	169	0.28	0.57	0.87	1.70
MMI IV	141,176	42,049	43,758	14,417	355	240	349	236	0.25	0.57	0.80	1.64
N50	125,335	37,832	40,133	13,234	314	207	301	196	0.25	0.55	0.75	1.48
N75	198,957	56,704	64,753	20,800	359	243	353	238	0.18	0.43	0.55	1.14
N80	218,894	61,440	80,927	25,030	374	253	371	251	0.17	0.41	0.46	1.00
Max35	186,699	53,989	51,889	16,849	342	231	328	219	0.18	0.43	0.63	1.30
Linear	113,316	35,529	34,459	11,667	316	209	303	198	0.28	0.59	0.88	1.70

plot on the left of Fig. 13, the number of potentially damaging earthquakes varies per year, possibly reflecting the effects of the incompleteness of the catalogue. The tendency for the number of earthquakes to increase in time is likely due to both the gain in detectability of smaller-magnitude earthquakes in the last decade or so, as a consequence of the improvement in seismological networks worldwide, and the increase in the number of agencies that contribute to the ISC Bulletin as time goes by. As will be discussed in Sect. 5.3, a similar effect is observed in the Database of Damaging Small-to-Medium Magnitude Earthquakes. The moment magnitude distribution of the catalogue of potentially damaging events shown in the plot on the right of Fig. 13 follows, in general terms, the Gutenberg–Richter relationship (Gutenberg and Richter 1944).

5.2 Reportedly damaging earthquakes not in the resulting catalogue

Of the 996 earthquakes of the Database of Damaging Small-to-Medium Magnitude Earthquakes (DDSMME) that occurred in the period 1st January 2001–31st December 2015, only 740 can be found within the 39,127 events that make up the catalogue of potentially damaging earthquakes. Of the remaining 256, 30 can be found within the 141,520 events that make up the global earthquake catalogue but do not pass the exposure criteria, 213 can be found in the initial (unfiltered) merged catalogue but do not comply with the maximum depth limits, six lie outside the defined magnitude range, and seven are not found at all, not even in the initial merged catalogue.

The seven reportedly damaging earthquakes that cannot be found in the merged catalogue include two cases mentioned only in local catalogues and/or the news (events M188 and M129 of DDSMME), one case mentioned in the USGS website and the NOAA database but not the ISC Bulletin (A119 of DDSMME), one case that can be found in the ISC Bulletin but only has one magnitude estimate in terms of M_L , which was not considered in either WPG16* or the present work (E13-373 of DDSMME), and three cases that exist in the ISC Bulletin but do not make it to the merged catalogue due to them being (likely

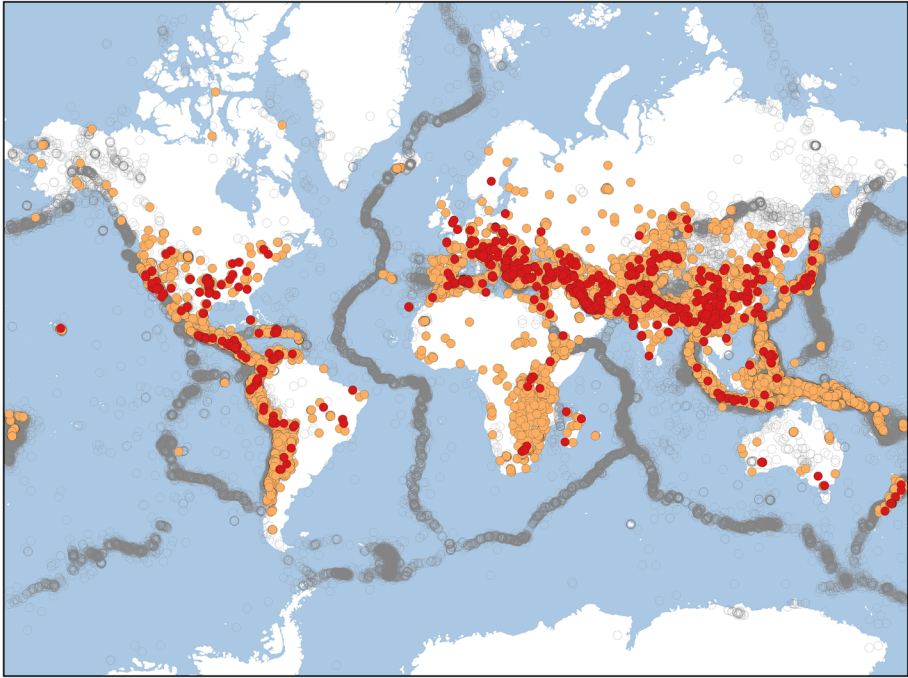


Fig. 12 Global earthquake catalogue (light grey; 141,524 events), catalogue of potentially damaging earthquakes (orange; 39,127 events) and damaging earthquakes within the latter (red; 740 earthquakes)

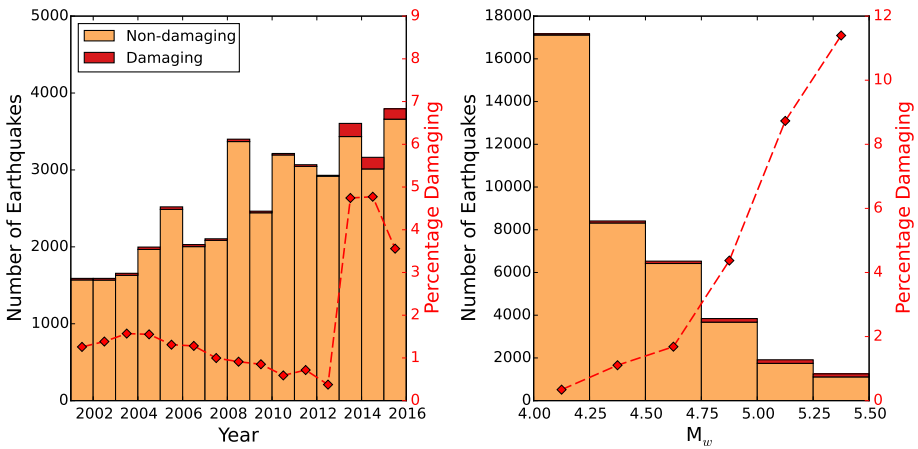


Fig. 13 Distribution of potentially damaging earthquakes in time (left) and by moment magnitude (right). Damaging and non-damaging events indicated in red and orange, respectively. Red rhombuses indicate the percentage of damaging events per bin

mis-)classified as explosions (E13-184, E14-353 and E15-274). A closer look at these last three revealed that there seem to exist cases in which one agency labels the event as a suspected explosion while others label it as having an anthropogenic origin other than an explosion. For example, six events located in Groningen, the Netherlands, an area known to be subject to seismicity induced by gas extraction activities, were identified as having been labelled as suspected explosions by the University of Upssala (UPP) and as simply anthropogenic by the ISC. A small investigation on the matter (Nievas et al. 2019a) led to the conclusion that this potential kind of misclassification is not likely to have a large impact on the results, notwithstanding the fact that classification of events into tectonic or induced earthquakes and non-earthquake phenomena is extremely complex and filtering based on the use of keywords is subject to issues like the one identified herein.

Of the six cases of reportedly damaging earthquakes that appear to lie outside the defined magnitude range, three correspond to earthquakes retrieved from the WPG16* catalogue to which values of M_w directly calculated by agencies not considered in the latter were assigned in the DDSMME (events A81, C99 and E13-257), one corresponds to simply selecting magnitude estimates from different agencies in DDSMME and the merged catalogue (E14-262), and the last two are reported to have M_w 5.55 within the WPG16* catalogue, which does not pass the criterion $M_w < 5.55$, while having M_w 5.549 in DDSMME (A48 and E14-052). The first four events illustrate the impact of the uncertainty that stems from different agencies obtaining different estimates of magnitude, while the last two illustrate the arbitrary nature of numerical precision within computer calculations.

The 213 reportedly damaging earthquakes that do not comply with the maximum depth limits are worthy of a more detailed investigation, as they represent 21.4% of the total number of earthquakes in the DDSMME in the period 1st January 2001–31st December 2015. Their distribution in time is not uniform, as 139 of them correspond to the period 2013–2015. As shown in Fig. 14a, b, a number of these earthquakes could comply with the criteria if the depth values reported in DDSMME were the ones assigned to them in the global earthquake catalogue herein. In order to assess the extent to which the choice of a different depth estimate (and its associated uncertainty) could have caused these earthquakes to be part of the catalogue, all depth estimates from different authors reported in the ISC Bulletin were gathered for each of them and the percentage of those that comply with the maximum depth limits was calculated.² Results plotted in Fig. 14c show that around 27% of these 213 earthquakes have depth estimates that pass the criterion less than 5% of the times, while around 65% do so less than 50% of the times, and suggest that exclusion of these events from the filtered catalogue was not simply due to an unlucky selection of depth estimates from those reported in the ISC Bulletin. While it would have been possible to assign the depth values reported in DDSMME to these earthquakes and, in this way, guarantee the consideration of the cases that fall within the trapezium of considered depths in Fig. 14a, this was not done herein to avoid inconsistencies with the criteria used for the compilation of the catalogue. Just like for these earthquakes, different depth estimates could have been assigned to all other earthquakes in the catalogue with more than one available estimate. Manually changing these values but not others would introduce a bias toward incorporating more damaging than non-damaging earthquakes.

² There are cases in the ISC Bulletin of depths reported by an agency B citing agency A as the source, that is, cases in which agency B adopts the estimates of agency A and reports them. Strictly speaking, these should be counted only once, as they are not independent estimates but a repetition. However, identifying such cases is difficult and was thus not done.

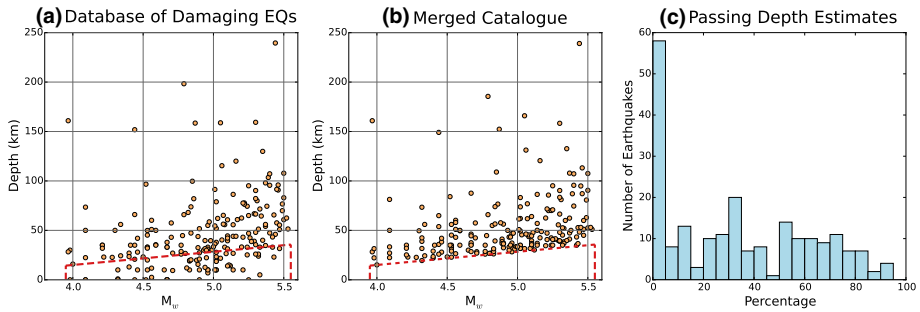


Fig. 14 Characterisation of the 213 reportedly damaging earthquakes that do not comply with the maximum depth limits: **a** depth as reported in DDSMME versus M_w as in the merged catalogue, **b** depth as in the merged catalogue versus M_w as in the merged catalogue, **c** percentage of depth estimates of each earthquake that comply with the maximum depth limits. Red dashed lines in (a, b) indicate the magnitude–depth range considered herein. Both (a, b) do not show one point with M_w 5.25 and depth 617 km

The existence of many supposedly deep yet reportedly damaging earthquakes can be due to (1) poor hypocentral depth resolution due to lack of the necessary good quality waveforms (i.e., the depths having actually been shallower than the reported estimates), (2) erroneous assignation of an earthquake to an entry of the EID (i.e., the selected earthquake is not the one that caused the damage/casualties; see Nievas et al. 2019b, c), (3) the earthquakes having caused unexpectedly large ground motions, (4) erroneous reporting of damage or casualties for an earthquake that did not cause either, or (5) reported consequences being indirectly related to the earthquake. An example of the latter is a M_w 5.44 event that occurred on 4th April 2013 at 06.31 UTC in Afghanistan at a depth of 239 km (E13-078 in DDSMME), for which the EID reports damage level 1—which translates into non-structural damage according to their conventions—and three injuries due to people falling off the stairs. The EID provides a link to the Earthquake-Report (2013) website, which states that the injuries were due to the panic that the shaking caused in Peshawar, and that no damage occurred. An extreme depth value of 617 km, not shown in Fig. 14a, b, corresponds to a case identified by the EID as a M_w 5.25 ($m_b = 5.1$) event occurring on 27th November 2015 at 00.52 UTC in Acre, Brazil (E15-323 in DDSMME), the source for which was a local news website reporting on cracks on a wall of a church observed after an “earthquake of magnitude 5.1 that occurred in the dawn of Friday (27)” (Globo, 2015). While the article appears to intend to make reference to the aforementioned earthquake, there seems to be some confusion with the local time, as this earthquake did not occur during the local dawn but in the evening of the previous day. This apparent confusion and the fact that the photos published on the website show minor cracks suggests that there is a possibility that these were caused by a M_w 6.7 earthquake that occurred less than 24 h earlier, reported at a similar depth, which is deemed normal for this region of the world. These two examples illustrate the fact that the events contained in the Database of Damaging Small-to-Medium Magnitude Earthquakes are associated with different degrees of both reported consequences and reliability.

Of the 30 earthquakes that comply with the maximum depth limits but not with the exposure criteria, ten would not pass the latter even if thresholds of 1 person and 1 person/km² (criterion C7 in Table 2) were used instead of 2500 people and 300 people/km². Of these ten, only one is reportedly associated with a casualty (a death attributed to a M_w 5.33 earthquake in Vietnam that occurred on 8th November 2005—M169 in DDSMME—and

which is classified as a shaking death by PAGER-CAT, Allen et al. 2009, of which no further details have been found), three and two are reported to have caused non-structural and limited damage by the EID, respectively, another three are reported to have caused “few, slight”, “minor”, “some, minor” damage, and a final one is assigned a 50–100 damaged-building range by NOAA. The other 20 of the 30 earthquakes that do not comply with the exposure criteria are associated with a variety of consequences, with only two having reportedly caused damage to over 30,000 buildings and injuries to 28–100 people each, while the remaining 18 are associated with no casualties and much lower numbers and mild qualitative descriptions of damaged buildings.

Apart from the influence of not being able to count 256 of the 996 earthquakes (25.7%) that make up the Database of Damaging Small-to-Medium Magnitude Earthquakes on the statistics on the proportion of damaging earthquakes presented in Sect. 5.3, it is of interest to understand the overall number of casualties and damaged buildings that these earthquakes represent. To this end, Figs. 15 and 16 depict the number of injuries, total deaths, damaged and destroyed buildings attributed to individual earthquakes (isolated markers) and their sum per year (markers connected by lines), both for the 740 earthquakes that make up the catalogue of potentially damaging events (coloured markers, thick continuous lines), for the remaining 256 earthquakes (grey void markers) and the total 996 earthquakes (thin dashed lines). With the exception of some peaks in the number of deaths, particularly in the years 2002 and 2006 and highly influenced by 50 landslide-caused deaths associated with two earthquakes, the total numbers of the whole of the Database of Damaging Small-to-Medium Magnitude Earthquakes appear to stay relatively well-represented by the 740 earthquakes that can be used for this statistical study. Overall (and, again, with the exception of the deaths), the proportion of casualties and damaged/destroyed buildings these 256 earthquakes account for is smaller than the 25.7% they represent of the total 996 events (ranging between 5 and 19%), though this does not imply their consequences are negligible (compare the continuous and dashed lines in Figs. 15 and 16). It should be noted as well that Fig. 16 contains only the earthquakes for which numerical values were available, though there exist many cases of damage being reported only in qualitative terms, particularly in the years 2013–2014, and for which comparisons are much more difficult (for details regarding the processing of different kinds of data please refer to Nievas et al. 2019c).

5.3 Proportion of damaging earthquakes

Table 7 summarises the results obtained for the whole catalogue of potentially damaging earthquakes. In the 15 years it spans, 1.9% of the earthquakes considered to have posed a threat under the conditions defined for this study reportedly caused damage and/or casualties. The proportion rises to 3.3% when considering only events classified as main shocks, going down slightly to 2.7% if main shocks and foreshocks are considered together. The reduction in the proportion observed for aftershocks may be due to three reasons: (1) the difficulties associated with recording damage and consequences stemming from any particular aftershock of a relatively destructive main shock (the consequences tend to be reported for the sequence as a whole), (2) the likelihood that earthquakes in the range M_w 4.0–5.5 that are classified as aftershocks occur in areas of high seismicity, capable of generating large-magnitude earthquakes that would have aftershocks of these magnitudes, where buildings might be better designed to withstand the shaking and societies are more used to frequent low-amplitude ground motions, so that minor damage does not get reported,

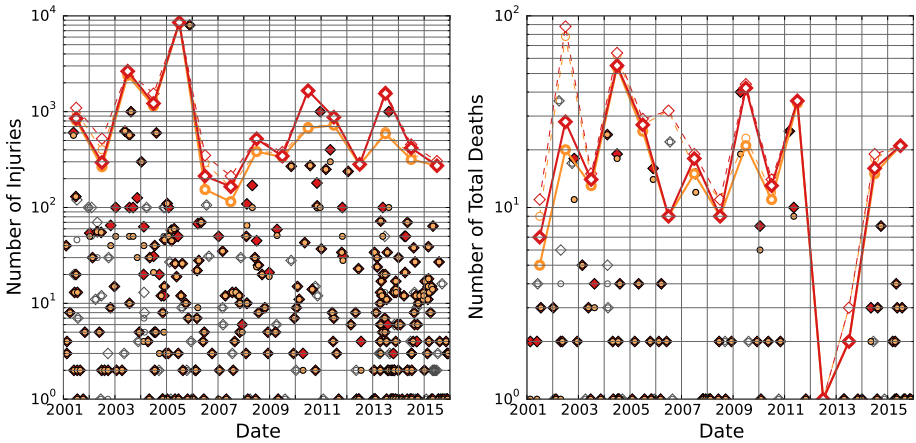


Fig. 15 Number of injuries (left) and total deaths (excluding heart attacks, right) reported in the Database of Damaging Small-to-Medium Magnitude Earthquakes. Isolated circles and rhombuses indicate values per each individual earthquake, in colour those that are found within the catalogue of potentially damaging events, in grey those that are not. Circles and rhombuses connected with lines indicate the total per year, in thick continuous lines those found within the catalogue, in thin dashed lines the total. Orange circles and red rhombuses refer to lower and upper bounds of values reported in the original sources, respectively. Zero values not shown

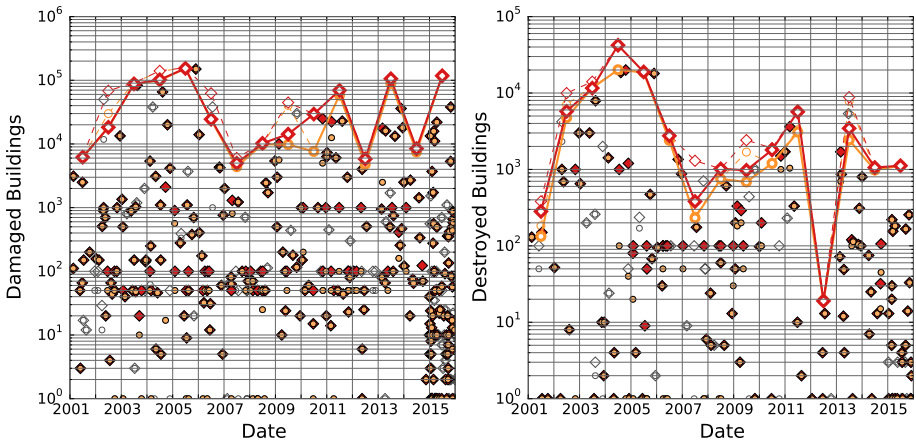


Fig. 16 Number of damaged (left) and destroyed (right) buildings reported in the Database of Damaging Small-to-Medium Magnitude Earthquakes (only cases in which numerical values, i.e., not qualitative descriptions, are available). Colours and lines as per caption of Fig. 15. Zero values not shown

and (3) the known tendency of aftershocks to have lower stress drops than main shocks and thus result in lower ground motions, when they occur in the same part of the fault that was ruptured by the main shock (e.g., Abrahamson et al. 2014, Wooddell and Abrahamson 2014). Interestingly, this means that including the aftershocks in the statistics does not lead to an increase in the resulting proportions of damaging earthquakes—as could have been expected based on the influence of incremental damage, as discussed in Sect. 3.4—but the

Table 7 Proportion of damaging events within the catalogue of potentially damaging earthquakes

	Total	Damaging	Proportion (%)
All shocks	39,127	740	1.9
Main shocks and foreshocks	21,475	584	2.7
Main shocks	15,123	498	3.3
Foreshocks	6352	86	1.4
Aftershocks	17,652	156	0.9

opposite. As shown in the plot on the right of Fig. 13, the proportion of damaging earthquakes increases significantly with magnitude, as would be expected.

However, as noted earlier, the earthquakes that make up the catalogue of potentially damaging events are not uniformly distributed in time, and neither are those of the Database of Damaging Small-to-Medium Magnitude Earthquakes, as a consequence of the incompleteness of both. As shown on the plot on the left of Fig. 13, the proportions of damaging earthquakes increase significantly in the last 3 years of the 15-year period under study, due to the influence of the EID in the number of damaging earthquakes identified per year (see Fig. 3). This tendency can be observed in main shocks as well as in foreshocks and aftershocks, as shown in Fig. 17. On average, 4.3% of the potentially damaging earthquakes (6.2% of main shocks and foreshocks) that occurred in the period 2013–2015 were identified as damaging, while this value reaches only 1.0% for the period 2001–2012. Individually, 2013, 2014 and 2015 are reported to have seen 4.7%, 4.8% and 3.6% of damaging cases within the catalogue of potentially damaging earthquakes (7.7%, 6.3% and 4.9% considering only main shocks and foreshocks). It is thus possible that a more realistic annual percentage of damaging earthquakes may be around 4.3%—or even larger—and not the average 1.9% obtained for the whole period of 2001–2015 (or 6.2% instead of 2.7% considering only main shocks and foreshocks). The potential implications of this on the total number of casualties or damaged buildings observed per year is difficult to assess. For example, of the total 43 upper-bound deaths reported for 2013–2015, 48.8% correspond to events retrieved from the EID. However, it would not be reasonable to assume that the 88 upper-bound deaths reported for the year 2002 alone (Fig. 15) represent only the equivalent 51.2% not retrieved from the EID to which another 84 deaths assumed to have occurred but not to have been recorded properly would need to be added. The large variability observed for the total consequences per year does not allow to make such inferences. The situation is even more complicated for the number of damaged and destroyed buildings, as so many earthquakes are associated only with qualitative descriptions. It is interesting to note, however, that the overall numbers of casualties and damaged buildings depicted in Figs. 15 and 16 do not peak for the years 2013–2015 as do the numbers of earthquakes (though many earthquakes retrieved from the EID are associated with descriptions of damaged buildings instead of numbers). Moreover, it is unlikely that earthquakes with extensive consequences occurred before 2013 had passed unnoticed, and it is known that EID has the capacity to report on earthquakes associated with even the slightest damage (Nievas et al. 2019b, c). All this suggests that the unreported earthquakes from 2001 to 2012 are not likely to have had an extreme impact on the overall annual consequences, despite their significance in terms of proportion of damaging events.

It is noted that the effect of the inclusion of data found in the EID and not elsewhere cannot be regarded as a statistical artefact of changing the data sources but rather as an

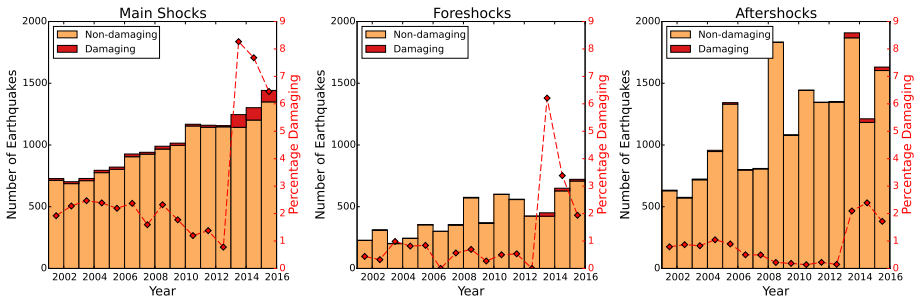


Fig. 17 Distribution in time of the catalogue of potentially damaging events. Damaging and non-damaging events indicated in red and orange, respectively. Red rhombuses indicate the percentage of damaging events per year

expression of the inherent problem of accessibility to post-earthquake assessment data in the case of small-to-medium magnitude earthquakes. It cannot, thus, be overemphasised that the apparent increase in the proportion of damaging earthquakes in the last 3 years of the present study is unlikely to be due to a real increase (i.e., earthquakes occurring after 2013 having actually been more damaging than before 2013) but a result of the improvements in detectability and accessibility in this time period (i.e., the proportion is likely to have been relatively more stable but only more detectable from 2013 onward); in other words, numbers for years prior to the EID are likely underestimated. Consequently, we regard the values obtained for 2013–2015 as potentially closer to reality (despite all their limitations) than those averaged over the 15 years of the study.

The potential impact of modifying the criteria used to define the catalogue of potentially damaging earthquakes for the present study can be illustrated by taking a look at the relation between the percentage of damaging events and the total number of potentially damaging earthquakes obtained with the different population thresholds used for the sensitivity study in Sect. 4.3.1 and enumerated in Table 2. As can be observed in Fig. 18, the percentage of damaging earthquakes decreases when the size of the catalogue of potentially damaging events increases, reaching minimum values when the most flexible criterion C7 (1 person, 1 person/km²) is used. The same effect could be expected if deciding to consider, for example, larger maximum depths (i.e., deeper hypocentres) that would allow for at least some of the 213 earthquakes that do not comply with the limits defined in Fig. 7 to be included in the final ratio. It is of great interest to note as well that the order of magnitude of the percentage of damaging events stays within a reasonable range.

If the whole 256 reportedly damaging earthquakes that do not form part of the catalogue were added to the statistics without adding any non-damaging event as well, the percentage of damaging would rise from 4.3% to only 5.8% for the period with the highest proportion, 2013–2015, and from 4.8 to 6.6% for 2014 alone, the year with the largest increase in percentage.

6 Discussion and conclusions

In a context of increasing interest of the earthquake engineering community in small-to-medium magnitude earthquakes, this paper has presented a study conducted to determine the proportion of upper-crustal events in the range M_w 4.0–5.5 that occur sufficiently close

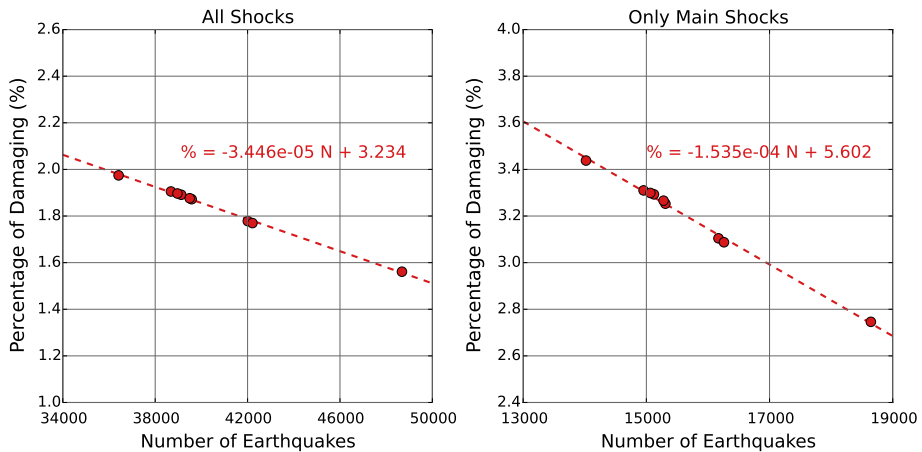


Fig. 18 Percentage of damaging events versus size of the catalogue of potentially damaging earthquakes obtained with different population thresholds (Table 2), considering all shocks (left) and only main shocks (right). Linear regressions shown in red

to inhabited areas worldwide that actually cause damage and/or casualties. The compilation of this global catalogue of potentially damaging earthquakes has required addressing a series of challenges, many of which are common to the building of earthquake catalogues in general, while many other are far more specific, such as the definition of what *sufficiently close to inhabited areas* means.

Results show that, on average, around 2% of all shocks were identified as damaging. However, the proportion varies significantly in time (most likely due to incompleteness of the catalogue and the Database of Damaging Small-to-Medium Magnitude Earthquakes) and when only main shocks and foreshocks are considered. Interestingly, the proportion rises to 2.7% in the latter case—which can be thought of as an approximation to the case of damage observations independent from previous earthquakes in a sequence—while that corresponding to aftershocks is smaller (0.9%). This is possibly due to a combination of the fact that, in many cases, consequences due to aftershocks are recorded and reported jointly with the main shock, the likelihood of many aftershocks of magnitude M_w 4.0–5.5 occurring in areas of high seismicity—which may tend to ignore minor consequences in contrast with more devastating events they often experience—and the known tendency of aftershocks to have lower stress drops than main shocks when they rupture the same segment of the fault as the latter.

The variation of the proportion of damaging earthquakes across the 15 years studied herein is thought to be related mostly to the impact of the access to data regarding damage and seismicity in general. The detectability of small-magnitude earthquakes and the worldwide network coverage have increased during the time period under study, as has the number of agencies that contribute with their data to the ISC Bulletin. In terms of reported consequences, the initiation of the EID in the year 2013 marks a significant jump in the number of earthquakes reported as damaging or being associated with casualties. For the 3 years comprised in the period 2013–2015, the proportion of damaging earthquakes rises to 4.3% and 6.2% for all shocks and main shocks plus foreshocks, respectively. As it is still possible that earthquakes that caused damage and/or casualties be missing from the EID, these proportions could be even higher. Nevertheless, it is fundamental to recall that

the extent of the consequences of the earthquakes considered as damaging is very variable and can range from very light damage and/or injuries caused by reacting in panic, all the way through to structural collapses and death. As a consequence, these 4.3% or 6.2%—or potentially larger values—cannot and should not be interpreted as implying a homogeneous set of earthquakes with serious life-threatening consequences but as having a mixture of a variety of potential outcomes. For an insight regarding the kind of consequences observed, the reader is invited to refer to Nievas et al. (2019b, c).

While 21.4% of the total 996 identified damaging earthquakes did not comply with the maximum depth limits set to define upper-crustal events, 3.0% did not pass the filtering according to population exposure criteria, 0.6% ended up outside of the magnitude range of interest, and a further 0.7% could not be found in the ISC Bulletin at all, the impact of these on the overall statistics is not expected to be large. It would be unrealistic to expect that a change in the criteria used to compile the catalogue of potentially damaging earthquakes would lead to the inclusion of only these damaging events and not a proportional or even larger amount of non-damaging ones, as it was observed that the proportion of damaging earthquakes decreased with an increasing size of the catalogue. Moreover, even if only the 256 reportedly damaging events that do not form part of the catalogue were added, the percentage of damaging would rise from 4.3% to only 5.8% for the period with the highest proportion, 2013–2015, and from 4.8 to 6.6% for 2014 alone, the year with the largest increase in percentage. A series of sensitivity analyses have shown as well that results are relatively stable with respect to the maximum depth limits, the parameters used for declustering, and different population and density thresholds.

While the incompleteness of the catalogue of potentially damaging earthquakes could be resulting in larger proportions of damaging earthquakes (think of the very small 0.7% of damaging earthquakes that could not be found in the ISC Bulletin at all), the under-reporting of damage and consequences is doing the opposite. In this sense, the complexities entailed by the identification of small-to-medium magnitude earthquakes that have caused damage and/or casualties suggest that the proportions of damaging earthquakes determined in this work may be lower bounds. This is not, however, the only source of uncertainty in the present analysis. Exploration of the decisions regarding the maximum depths and population/density thresholds suggests that any loosening of the criteria, that is, any modification that leads to the catalogue of potentially damaging earthquakes to be larger, would likely result in smaller proportions of damaging earthquakes, while the opposite would be true if more restrictive conditions were to be imposed. However, the set of criteria adopted is judged to be appropriate and reasonable for the purposes of this study, particularly when considering the relatively constrained variations in the results observed in the sensitivity analyses. Further investigation is nevertheless warranted to refine the study by, for example, exploring statistics within smaller sub-ranges of magnitude, considering all depth and magnitude estimates for each earthquake, and examining the specific characteristics of the most significant examples of damaging small-magnitude events.

Acknowledgements The authors would like to thank Dr Dmitry Storchak, Dr Domenico Di Giacomo, and James Harris, from the International Seismological Centre, and Dr Istvan Bondár, for kindly and promptly answering our questions regarding the ISC location procedures and the ISC Bulletin. Our gratitude goes as well to Dr Peter Stafford for sharing with us his view on the characterisation of uncertainty in hypocentral depth and to Dr Graeme Weatherill for fruitful discussions on catalogue compilation and regionalisation, as well as his assistance in the use of the GEM Catalogue Toolkit https://github.com/GEMScienceTools/catalogue_toolkit. We are also very thankful to the two anonymous reviewers whose comments helped improve the quality of the manuscript. QGIS Geographic Information System (<https://qgis.org>) used to generate the map in Fig. 12. This work was funded by Nederlandse Aardolie Maatschappij B.V. (NAM).

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References


- Abrahamson NA, Silva WJ, Kamai R (2014) Summary of the ASK14 ground motion relation for Active Crustal Regions. *Earthq Spectra* 30(3):1025–1055
- Adams RD, Hughes AA, McGregor DM (1982) Analysis procedures at the International Seismological Centre. *Phys Earth Planet Inter* 30:85–93
- Allen TI, Marano KD, Earle PS, Wald DJ (2009) PAGER-CAT: a composite earthquake catalog for calibrating global fatality models. *Seismol Res Lett* 80(1):1325–1345
- Allen TI, Wald DJ, Worden CB (2012) Intensity attenuation for active crustal regions. *J Seismol* 16:409–433
- Artemieva IM, Thybo H (2013) EUNaseis: a seismic model for Moho and crustal structure in Europe, Greenland, and the North Atlantic region. *Tectonophysics* 609:97–153
- Atkinson GM, Wald DJ (2007) “Did You Feel It?” intensity data: a surprisingly good measure of earthquake ground motion. *Seismol Res Lett* 78(3):362–368
- Bakun WH, McGarr A (2002) Differences in attenuation among the stable continental regions. *Geophys Res Lett* 29(23):4
- Bommer JJ, Crowley H (2017) The purpose and definition of the minimum magnitude limit in PSHA calculations. *Seismol Res Lett* 88(4):1097–1106
- Bondár I, Storchak D (2011) Improved location procedures at the International Seismological Centre. *Geophys J Int* 186:1220–1244
- Center for International Earth Science Information Network-CIESIN-Columbia University (2016) Gridded Population of the World, Version 4 (GPWv4). NASA Socioeconomic Data and Applications Center, Palisades. <http://dx.doi.org/10.7927/H4NP22DQ>
- Chen Y-S, Weatherill G, Pagani M, Cotton F (2018) A transparent and data-driven global tectonic regionalisation model for seismic hazard assessment. *Geophys J Int* 213(2):1263–1280
- Di Giacomo D, Bondár I, Storchak DA, Engdahl ER, Bormann P, Harris J (2015) ISC-GEM: Global Instrumental Earthquake Catalogue (1900–2009), III. Re-computed M_S and m_b , proxy M_W , final magnitude composition and completeness assessment. *Phys Earth Planet Inter* 239:33–47
- Dijkstra L, Poelman H (2018) Regional typologies overview. Eurostat—Statistics Explained platform. https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Regional_typologies_overview. Accessed 7 Jan 2019
- Dost B, Edwards B, Bommer JJ (2018) The relationship between M and ML : a review and application to induced seismicity in the Groningen Gas Field, the Netherlands. *Seismol Res Lett* 89(3):1062–1074
- Dziewonski AM, Chou T-A, Woodhouse JH (1981) Determination of earthquake source parameters from waveform data for studies of global and regional seismicity. *J Geophys Res* 86:2825–2852
- Earthquake Impact Database (EID). https://docs.google.com/spreadsheets/d/1oveZ4ZOLdJFnKu2aZfMg4aV1AR_OgHfXrM-65bJw4w/edit#gid=0. Accessed 5 Jan 2018 (year 2017), 22 Nov 2017 (years 2013–2016). Additional information: <https://earthquake-report.com/2018/02/10/earthquake-impact-database-2018/>. Accessed 20 Nov 2018
- Earthquake-Report (2013) Strong but very deep earthquake in Afghanistan shakes the entire region (many countries). <https://earthquake-report.com/2013/04/04/strong-earthquake-hindu-kush-region-afghanistan-on-april-4-2013/>. Accessed 15 Feb 2019
- Ekström G, Nettles M, Dziewonski AM (2012) The global CMT project 2004–2010: Centroid moment tensors for 13,017 earthquakes. *Phys Earth Planet Inter* 200–201:1–9
- Engdahl ER, van der Hilst R, Buland R (1998) Global teleseismic earthquake relocation with improved travel times and procedures for depth determination. *Bull Seismol Soc Am* 88(3):722–743
- Gardner JK, Knopoff L (1974) Is the sequence of earthquakes in Southern California, with aftershocks removed, Poissonian? *Bull Seismol Soc Am* 64(5):1363–1367
- Global Earthquake Model (GEM). <https://www.globalquakemodel.org/>. Accessed 18 Feb 2019
- Globo (2015) Após terremoto no AC, igreja em Tarauacá sofre rachaduras (Church in Tarauacá suffers cracks after earthquake in AC, in Portuguese). <http://g1.globo.com/ac/acre/noticia/2015/11/apos-terremoto-no-ac-igreja-em-tarauaca-sofre-rachaduras.html>. Accessed 15 Feb 2019

- Goertz-Allman BP, Edwards B, Bethmann F, Deichmann N, Clinton J, Fäh D, Giardini D (2011) A new empirical magnitude scaling relation for Switzerland. *Bull Seismol Soc Am* 101(6):3088–3095
- Grünthal G, Wahlström R, Stromeyer D (2009) The unified catalogue of earthquakes in central, northern, and northwestern Europe (CENEC)—updated and expanded to the last millennium. *J Seismol* 13:517–541
- Gutenberg B, Richter C (1944) Frequency of earthquakes in California. *Bull Seismol Soc Am* 34(4):185–188
- International Seismological Centre (ISC) IASPEI Seismic Format (ISF). <http://www.isc.ac.uk/standards/isf/>. Accessed 12 Feb 2019
- Mueller CS (2019) Earthquake catalogs for the USGS National Seismic Hazard Maps. *Seismol Res Lett* 90(1):251–261
- National Geophysical Data Center/World Data Service (NGDC/WDS) National Geophysical Data Center, National Oceanic and Atmospheric Administration (NOAA). Significant Earthquake Database. <https://www.ngdc.noaa.gov/hazard/earthqk.shtml>. <https://doi.org/10.7289/v5td9v7k>. Accessed 5 Sep 2018
- Nearby UK (2017) Country bounding boxes: <http://www.nearby.org.uk/downloads.html>. Accessed 17 Aug 2017
- Nievas CI, Bommer JJ, Crowley H (2017) Global occurrence and impact of small-to-medium magnitude earthquakes: a statistical analysis. Report for the Nederlandse Aardolie Maatschappij BV (NAM), The Netherlands. <https://nam-onderzoeksrapporten.data-app.nl/reports/download/groningen/en/5e983926-879b-489a-9199-6ed63640427e>
- Nievas CI, Bommer JJ, Crowley H (2019a) Global occurrence and impact of small-to-medium magnitude earthquakes: a statistical analysis—Version 2. Report for the Nederlandse Aardolie Maatschappij BV (NAM), The Netherlands. <https://nam-onderzoeksrapporten.data-app.nl/reports/download/groningen/en/cedfd158-c5f9-4b63-8a48-cb367479a3b2>
- Nievas CI, Bommer JJ, Crowley H, van Elk J, Ntinalexis M, Sangirardi M (2019b) A database of damaging small-to-medium magnitude earthquakes. *J Seismol* (under review)
- Nievas CI, Ntinalexis M, Kazantzidou-Firtinidou D, Borozan J, Sangirardi M, Crowley H, Bommer JJ (2019c) A database of damaging earthquakes of moment magnitude from 4.0 to 5.5 - version 2. Report for the Nederlandse Aardolie Maatschappij BV (NAM), The Netherlands. <https://nam-onderzoeksrapporten.data-app.nl/reports/download/groningen/en/bb24d606-1c5d-40dc-a9af-968ed182c400>
- Pacheco J, Sykes LR (1992) Seismic moment catalogue of large shallow earthquakes, 1900 to 1989. *Bull Seismol Soc Am* 82(3):1306–1349
- Pagani M, Monelli D, Weatherill G, Danciu L, Crowley H, Silva V, Henshaw P, Butler L, Nastasi M, Panzeri L, Simonato M, Viganò D (2014) OpenQuake engine: an open hazard (and risk) software for the Global Earthquake Model. *Seismol Res Lett* 85:692–702
- Papazachos BC, Kiratzi AA, Karacostas BG (1997) Towards a homogeneous moment-magnitude determination for earthquakes in Greece and the surrounding area. *Bull Seismol Soc Am* 87(2):474–483
- Scordilis EM (2006) Empirical global relations converting M_S and m_b to moment magnitude. *J Seismol* 10:225–236
- Stafford PJ (2014) Source-scaling relationships for the simulation of rupture geometry within Probabilistic Seismic-Hazard Analysis. *Bull Seismol Soc Am* 104(4):1620–1635
- Storchak D, Di Giacomo D, Engdahl E, Harris J, Bondár I, Lee W, Bormann P, Villaseñor A (2015) The ISC-GEM Global Instrumental Earthquake Catalogue (1900–2009): introduction. *Phys Earth Planet Inter* 239:48–63
- Uhrhammer R (1986) Characteristics of Northern and Central California Seismicity. *Earthq Notes* 57(1):21
- United Nations International Children's Emergency Fund (UNICEF) (2012) The state of the world's children 2012: Children in an urban world. Report published by Brodbeck Press, Inc. https://www.unicef.org/aids/files/SOWC_2012-Main_Report_EN_21Dec2011.pdf. Accessed 7 Jan 2019
- United States Census Bureau. 2010 Census urban and rural classification and urban area criteria. <https://www.census.gov/geo/reference/ua/urban-rural-2010.html>. Accessed 7 Jan 2019
- United States Geological Survey (USGS) Earthquake Catalog. <https://earthquake.usgs.gov/earthquakes/search/>. Accessed 5 Sep 2018
- United States Geological Survey (USGS) Why do so many earthquakes occur at a depth of 10 km? Short online note. <https://www.usgs.gov/faqs/why-do-so-many-earthquakes-occur-a-depth-10km>. Accessed 19 Oct 2018
- Văcăreanu R, Iancovici M, Neagu C, Pavel F (2015) Macro seismic intensity prediction equations for Vrancea intermediate-depth seismic source. *Nat Hazards* 79(3):2005–2031
- van Stiphout T, Zhuang J, Marsan D (2012) Theme V—models and techniques for analysing seismicity. Technical report. Community Online Resource for Statistical Seismicity Analysis. <http://www.corssa.org>

- Wald DJ, Allen TI (2007) Topographic slope as a proxy for seismic site conditions and amplification. *Bull Seismol Soc Am* 97:1379–1395
- Weatherill GA (2014) OpenQuake hazard modeller's toolkit - user guide. Global Earthquake Model (GEM), Technical Report
- Weatherill GA, Pagani M, Garcia J (2016) Exploring earthquake databases for the creation of magnitude-homogeneous catalogues: tools for application on a regional and global scale. *Geophys J Int* 206:1652–1676. WPG16* World Catalogue version of 21 July 2017, pers. comm
- Wooddell KE, Abrahamson NA (2014) Classification of main shocks and aftershocks in the NGA-West2 Database. *Earthq Spectra* 30(3):1257–1267

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

Affiliations

Cecilia I. Nieves¹  · Julian J. Bommer²  · Helen Crowley³  · Jan van Elk⁴

Cecilia I. Nieves
cecilia.nieves@gfz-potsdam.de

Helen Crowley
helen.crowley78@gmail.com

Jan van Elk
Jan.VanElk@shell.com

- ¹ Seismic Hazard and Risk Dynamics, GFZ German Research Centre for Geosciences, Potsdam, Germany
- ² Civil and Environmental Engineering, Skempton Building, Imperial College London, South Kensington Campus, London SW7 2AZ, UK
- ³ Seismic Risk Consultant, Pavia, Italy
- ⁴ Nederlandse Aardolie Maatschappij B.V. (NAM), Schepersmaat 2, 9405 TA Assen, The Netherlands