NAM

Global Occurrence and Impact of Small-to-Medium Magnitude Earthquakes: A Statistical Analysis (Part 2)

Cecilia I. Nievas, Julian J. Bommer & Helen Crowley

Datum March 2019

Editors Jan van Elk & Dirk Doornhof

General Introduction

Earthquakes are complex phenomena, which can have a large impact on the people living in the vicinity of the epicentre. This can include damage to buildings ranging from cracks to collapse, damage to infra-structure and psychological effects on the community. Case studies of historical earthquakes can therefore contribute to the understanding of the diversity of the effects earthquakes can have on the built environment, the natural environment, the local economy, the community and individual people. Case studies of earthquakes are therefore an element of the assurance of the risk assessment (Ref. 1).

Many case histories of earthquakes are available, but these focus primarily on larger earthquakes, like the 1906 earthquake in San Francisco and the 1960 earthquake in Chile. These tectonic earthquakes are considerably larger than the earthquakes expected to contribute to the hazard of induced earthquakes in Groningen.

Several studies trying to learn from smaller earthquake in the range from 4 to 5.5 have been initiated. Because case histories for earthquakes in the magnitude range relevant for Groningen are more difficult to find, a compendium of earthquakes was compiled (Ref. 2 and 5). Additionally, an overview of all (potentially) human-induced earthquakes (Ref. 3) was prepared. This was followed by a statistical evaluation of earthquakes (Ref. 4), focusing on earthquakes that occur in the upper crust in proximity to population centres and with a magnitude between M4 and M5.5. These earthquakes are most interesting for the human-induced earthquakes in Groningen.

This report presents an extension of the statistical evaluation of small-to-medium magnitude earthquakes.

References

1. Groningen Meet- en Regelprotocol, NAM, May 2017
2. A Database of Damaging Earthquakes of Moment Magnitude from 4.0 to 5.5, Cecilia Inés Nievas, Helen Crowley, Michail Ntinalexis and Julian J Bommer, June 2016
3. Human-induced Earthquakes, Gillian R. Foulger, Miles Wilson, Jon Gluyas and Richard Davies, June 2016
4. Global Occurrence and Impact of Small-to-Medium Magnitude Earthquakes: A Statistical Analysis Part 1, Cecilia I. Nievas, Julian J. Bommer & Helen Crowley, November 2017
5. A Database of Damaging Earthquakes of Moment Magnitude from 4.0 to 5.5, Cecilia I Nievas, Michail Ntinalexis, Danai Kazantzidou-Firtinidou, Jovana Borozan, Marialuigia Sangirardi, Helen Crowley & Julian J Bommer, March 2019

NAM

Title	**Global Occurrence and Impact of Small-to-Medium Magnitude Earthquakes: A Statistical Analysis (Part 2)**	**Date**	March 2019
		Initiator	NAM
Author(s)	Cecilia I. Nievas, Julian J. Bommer & Helen Crowley	**Editors**	Jan van Elk Dirk Doornhof
Organisation	Team of Academic Experts	**Organisation**	NAM
Place in the Study and Data Acquisition Plan	Study Theme: Hazard and Risk Assessment Comment: Earthquakes are complex phenomena, which can have a large impact on the people living in the vicinity of the epicentre. This can include damage to buildings ranging from cracks to collapse, damage to infra-structure and psychological effects on the community. Case studies of historical earthquakes can therefore contribute to the understanding of the diversity of the effects earthquakes can have on the built environment, the natural environment, the local economy, the community and individual people. Case studies of earthquakes are therefore an element of the assurance of the risk assessment. Many case histories of earthquakes are available, but these focus primarily on larger earthquakes, like the 1906 earthquake in San Francisco and the 1960 earthquake in Chile. These tectonic earthquakes are considerably larger than the earthquakes expected to contribute to the hazard of induced earthquakes in Groningen. Several studies trying to learn from smaller earthquake in the range from 4 to 5.5 have been initiated. Because case histories for earthquakes in the magnitude range relevant for Groningen are more difficult to find, a compendium of earthquakes was compiled. Additionally, an overview of all (potentially) human-induced earthquakes was prepared. This was followed by a statistical evaluation of earthquakes, focusing on earthquakes that occur in the upper crust in proximity to population centres and with a magnitude between M4 and M5.5. These earthquakes are most interesting for the human-induced earthquakes in Groningen. This report presents an extension of the statistical evaluation of small-to-medium magnitude earthquakes.		

Directly linked research	(1) Hazard and Risk Assessment (2) Meet – en Regelprotocol
Used data	Open Literature.
Associated organisation	Team of Academic Experts
Assurance	Report is based on compilation of academic papers and open literature.

Global Occurrence and Impact of Small-to-Medium Magnitude Earthquakes: A Statistical Analysis

Cecilia I. Nievas, Julian J. Bommer & Helen Crowley

Version 2

March 2019

Table of Contents

1. INTRODUCTION

While earthquake engineering has traditionally focused on moderate- and large-magnitude earthquakes capable of causing severe damage and resulting in large numbers of casualties, the interest in events of smaller magnitudes has increased significantly in recent years. The reasons for this are many and include the occurrence of relatively destructive small-magnitude earthquakes (*e.g.*, 2017 Ischia, Italy) as well as the increasing attention put on the assessment of existing buildings, many of which were not designed to withstand earthquake loading or have been left to deteriorate with time. A fundamental reason of this increased interest that is particularly relevant to the present work is, however, the potential risk posed by induced seismicity (*e.g.*, Taylor *et al.*, 2018).

Studies have shown that events with moment magnitude (**M**) in the range 4.0-5.5 dominate the seismic hazard and risk estimates due to induced earthquakes in the Groningen field and, potentially, in many other areas of the world in which anthropogenic earthquakes pose a larger threat than tectonic seismicity (Bourne *et al.*, 2015; van Elk *et al.*, 2019). The objective of the present work is, thus, to quantify the proportion of upper crustal earthquakes with magnitudes in the range **M**4.0-5.5—occurring in close proximity to exposed assets or population—that cause damage and/or casualties. This very statement contains three conditions of the earthquakes analysed herein: (i) they occur in the upper crust, (ii) their moment magnitude is equal to or larger than **M**4.0 and equal to or smaller than **M**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 the context of this work being the seismic risk associated with anthropogenically-induced earthquakes. The categorisation is made here in contrast with deeper subduction earthquakes and earthquakes related to relatively peculiar conditions, like those found in Bucamaranga (Colombia-Venezuela), Vrancea (Romania), or the Hindu Kush (Afghanistan-Pakistan). None of the deeper events are comparable with anthropogenically-induced earthquakes, which, by nature, occur within the regions of the crust accessible to human activities.

The last requisite is the proximity of the earthquakes to exposed assets or individuals and is fundamental to the objective of this work, as only events that occur sufficiently close to human settlements to pose a threat to them are of interest herein. The most obvious earthquakes that need to be left out in this regard are those occurring in largely uninhabited areas of the planet, such as oceans and deserts. But there are many other regions of the world with such scarce population that small earthquakes could not pose a significant threat anyway. The key here becomes the definition of a reasonable criterion to determine what is *sufficiently close* and what is *too far away*, concepts that are clearly linked to the earthquake magnitude and the depth at which it occurs. An extremely costly answer to this question could be given by means of a complete earthquake-by-earthquake risk scenario analysis in which the hazard, exposure and vulnerability are all explicitly modelled, and an estimate of losses and casualties is made. As this is clearly not reasonable when dealing with hundreds of thousands of events located all around the globe, this issue was tackled by means of

Intensity Prediction Equations (IPEs), that is, models that yield an expected level of macroseismic intensity at a certain distance from a specific earthquake, in combination with a world population grid to represent exposure and a set of thresholds of *sufficient number of people* subject to a *sufficient level of seismic intensity*, as schematically shown in Figure 1.1. In this way, population was used as a proxy for the built environment.

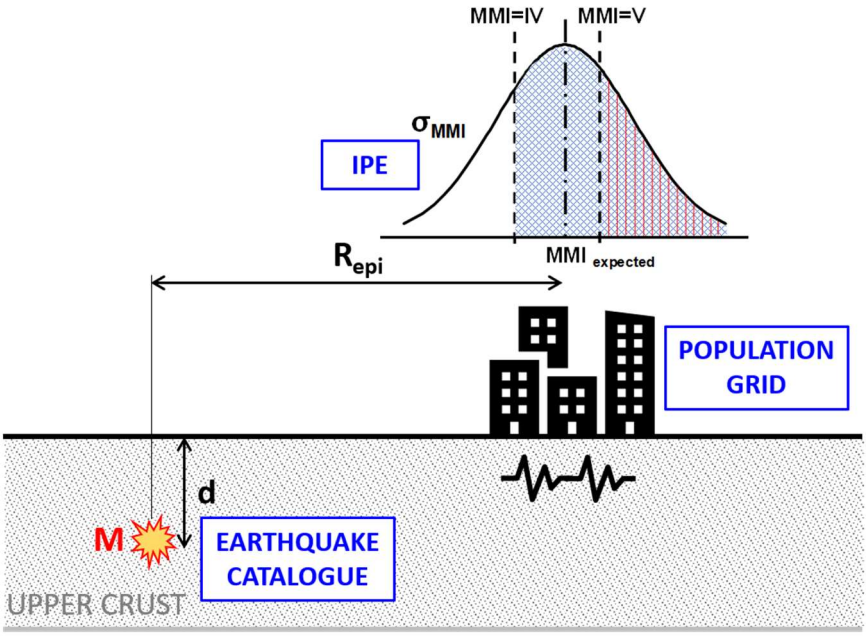

Figure 1.1. Schematic representation of the process used to determine whether an earthquake occurred *sufficiently close* to population and the built environment or not.

The identification of which of these earthquakes caused damage and/or injuries was carried out separately from the present work and resulted in a Database of Damaging Small-to-Medium Magnitude Earthquakes whose description can be found in Nievas *et al.* (2019). An exploratory version of the present statistical study, referred to as *Version 1* in the whole of this report, was presented in Nievas *et al.* (2017). The present version, referred to as *Version 2*, builds upon Version 1, addressing the points identified as possible improvements in the report for Version 1. The main areas in which modifications have been introduced are the explicit consideration of uncertainty in depth, magnitude and intensity prediction models, the use of empirical models to convert from several magnitude scales into moment magnitude, the modification of the maximum depth criteria, the introduction of probabilistic thresholds for the counting or not of population cells for the determination of expected degree of exposure, and the flagging of potential anthropogenically-induced earthquakes. These and other changes introduced in Version 2 are discussed in Chapter 2, alongside the outline of the methodology used. Chapter 3 presents a series of sensitivity analyses carried out over Version 1 that led to decisions made for the generation of Version 2. Chapter 4 goes deep into the details of the changes introduced in Version 2 and the calculation process, while the resulting world catalogue and the statistical analysis are presented in Chapter 5. Conclusions and discussions can finally be found in Chapter 6.

2. GENERAL OVERVIEW

2.1 Outline of the Methodology

The objective of the present work is to quantify the proportion of upper crustal earthquakes with magnitudes in the range **M**4.0-5.5 occurring in close proximity to exposed assets or population all around the world that cause damage and/or casualties. As discussed in the Introduction, this statement imposes three conditions that refer to (i) the earthquakes occurring in the upper crust, (ii) their magnitudes being in the range **M**4.0-5.5, and (iii) their location being sufficiently close to human settlements to pose a threat. The core of the work is, thus, the definition of the criteria needed to compile a catalogue of earthquakes that comply with these conditions, which we refer to as *catalogue of potentially damaging earthquakes*.

Figure 2.1 schematically shows the outline of the process followed for the compilation. The catalogue was built for a 15-year time window in the near past, a period long enough to be of statistical significance and, at the same time, short enough to not pose an excessive computational demand, using the magnitude-homogeneous catalogue of Weatherill *et al.* (2016) (WPG16 hereafter) and the Bulletin of the International Seismological Centre (ISC, see Web References). However, the process started with a longer period with additional 3 years considered before and 2.5-3 years after, so as to be able to identify aftershocks and foreshocks by means of a declustering algorithm. Each earthquake, incorporated either from the WPG16 catalogue or the ISC Bulletin, was characterised by one estimate of hypocentral location and one estimate of magnitude. Different estimates found in the ISC Bulletin were sorted according to hierarchies of agencies defined as explained in the report for Version 1. An algorithm designed to identify potential duplicate entries of the same earthquake was run over the resulting merged catalogue, as in Version 1 as well. Declustering was then carried out using the algorithm of Gardner & Knopoff (1974), as implemented in the OpenQuake Hazard Modeller's Toolkit (Weatherill, 2014), modified to take into consideration hypocentral depth. As it is extremely difficult to distinguish the damage caused by each of the events in a sequence, the resulting flagging of foreshocks and aftershocks was used to be able to generate separate statistics considering only main shocks or considering all events.

After declustering, the earthquakes were filtered for the final time period of interest (the 15 years mentioned earlier), and according to their moment magnitude and depth, considering only earthquakes in the range **M**4.0-5.5 occurring in the shallow crust, defined by means of magnitude-dependent maximum depth criteria. The set of events that make up this group is referred to herein as *global catalogue of upper-crustal earthquakes*. Earthquakes suspected to have been caused or triggered by anthropogenic activities were flagged.

The (estimated) population exposed to (estimated) Modified Mercalli Intensities (MMI) above a certain threshold was used to eliminate earthquakes happening in extremely underpopulated areas, oceans or deserts that clearly pose no threat or minimal threat to human settlements. For each earthquake, the procedure was as follows. Using intensity prediction equations (IPE), the MMI estimated to have been generated by the earthquake

were calculated using a spatial grid around the epicentre. Cells whose MMI values (or probability of observing a value, depending on the version) were above a certain threshold were identified. Within these cells, population counts were added and the maximum density across them all was determined. The earthquake was kept within the catalogue if the maximum density or the population count were above 300 people/km^2 and 2,500 people, respectively, these thresholds being selected based on definitions of urbanisation by different sources, as explained in the report for Version 1 and in Section 3.4.1 herein. Earthquakes kept after this filtering were the ones making up the final catalogue of crustal small-to-medium magnitude earthquakes to have occurred near urbanised areas, or *catalogue of potentially damaging earthquakes*. The final step then consisted in identifying within this catalogue the earthquakes that make up the Database of Damaging Small-to-Medium Magnitude Earthquakes (Nievas *et al.*, 2019), and determining the proportion of the former that the latter represents.

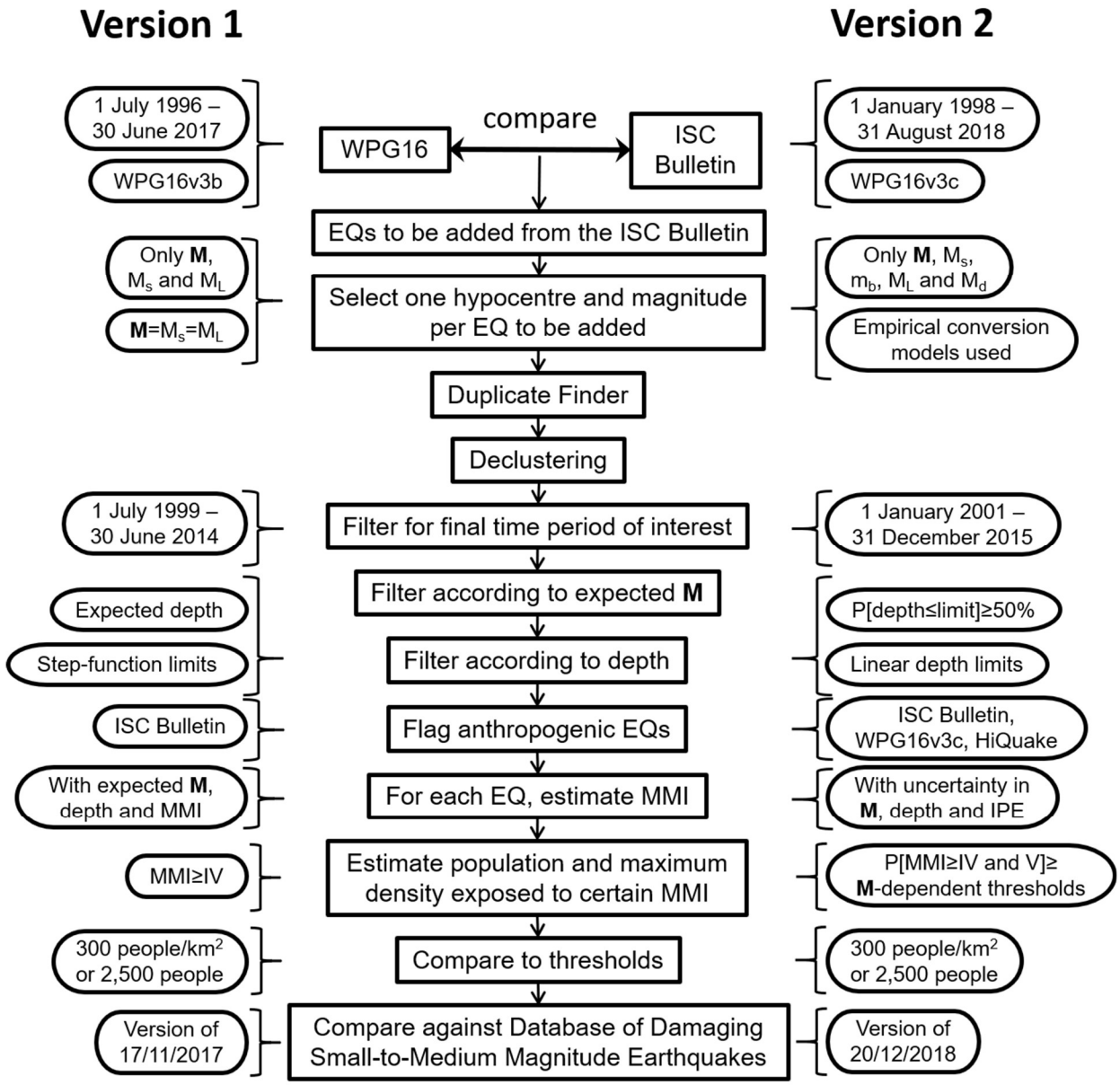

Figure 2.1. Outline of the methodology and main differences between Version 1 (left) and Version 2 (right).

2.2 Changes Introduced in Version 2

The left and right ends of Figure 2.1 highlight the main differences between Versions 1 and 2 of this work. The sources used to build Version 2 were still the catalogue of Weatherill *et al.* (2016) and the ISC Bulletin, but the time window was shifted one year and a half so as to take advantage of the processing of the ISC Bulletin that took place between the two versions. This was done so as to cover more of the recent earthquake history, in recognition of the much larger amount of data on damage and consequences of seismic events available for the period 2013-2017 thanks to the Earthquake Impact Database (EID). The 15-year time window defined for Version 2 starts, thus, on 1st January 2001 and finishes on 31st December 2015, as the ISC Reviewed Bulletin covers until 31st January 2016 at the time of gathering the data (September 2018). As for the case of Version 1, an additional 3-year period was considered before the starting date for the declustering of the catalogue, though a shorter 2.5-year period that brings the final date to almost the present time was used after the end date. This reduction is not expected to have a significant impact on the results, as aftershocks that may occur in the future do not affect the declustering process, and main shocks to occur whose foreshocks have already occurred would only lead to a potential misclassification of the latter as main shocks. However, as it is generally accepted that the foreshock time window may be shorter than the aftershock time window, and as, for the sake of damage and casualties, a foreshock of a main shock still to come is equivalent to a main shock in itself (the consequences of the main shock have not yet occurred and are, thus, neither reported yet nor mixed with those of the foreshock), this half a year difference is not likely to significantly affect this study. The declustering of the merged catalogue was carried out in the same way as for Version 1, using the algorithm of Gardner & Knopoff (1974).

Version 2 uses version v3c of the catalogue of Weatherill *et al.* (2016), referred to as WPG16v3c hereafter, instead of version v3b. Both versions of WPG16, *i.e.* v3b and v3c, cover the same time period and take the same original sources. The main difference between the two is that v3c uses the magnitude conversion models reported in the paper of Weatherill *et al.* (2016) while v3b used different ones that led to relatively small values of m_b to be converted into comparatively large values of moment magnitude, as will be discussed in Section 4.2.1.

Moment magnitude **M**, surface-wave magnitudes M_s, body-wave magnitudes m_b, local magnitudes M_L, and duration magnitudes M_d were considered for the earthquakes retrieved from the ISC Bulletin in Version 2, all of the latter transformed into moment magnitude by means of empirical conversion models. Details on this topic are discussed in Section 4.4.1.

The maximum depth criterion defined in Version 1 for the selection of upper crustal events by means of a step function was transformed into a linear relation between 15 km at **M**4.0 and 35 km at **M**5.5, based on the considerations explained later in Section 3.2. In Version 2, the reported depth of earthquakes with fixed-depth solutions was directly compared against these maximum depth limits, while earthquakes with free-depth solutions were subject to a probabilistic approach by which depths with a probability equal to or larger than

50% of lying below the limit were kept in the catalogue. This approach was adopted based on an analysis of the impact of uncertainties in depth on the uncertainties in predicted macroseismic intensities (see Section 4.5.2).

A couple of modifications were introduced in the calculation of the estimated Modified Mercalli Intensities (MMI) as well. Firstly, the IPE of Allen *et al.* (2012) was replaced by that of Atkinson & Wald (2007), which provides different sets of coefficients for Active Crustal and Stable Continental Regions (ACRs and SCRs, respectively). The two were used to represent these two extreme conditions and were weighted by means of a craton index (Chen *et al.*, 2018), as will be explained in Section 4.6.2. Secondly, the uncertainty in magnitude and depth values was explicitly incorporated to the calculation of the probability of observing a certain level of macroseismic intensity at a particular location, as explained in Section 4.6.2 as well. Thirdly, while in Version 1 population cells were counted when the estimated mean MMI was equal to or greater than IV, the criterion for counting cells in Version 2 was that the probability of observing MMI≥IV and MMI≥V was greater than a pre-defined threshold which is lower than the 50% implicit in using only the mean MMI in Version 1. Moreover, these thresholds were defined to be consistent with the maximum depth criteria.

While uncertainty in depth and magnitude was taken into account in Version 2 for the computation of estimated values of MMI, expected values were used both to select events within the magnitude range of interest and to decluster the catalogue, as any other choice would have led to unnecessary complications. The use of expected values for the inclusion or not of events based on their magnitude effectively implies that earthquakes with at least a 50% probability of lying in the range **M**4.0-5.5 are included. Selecting a higher probability threshold would lead to the range being effectively narrower and the very contradictory and confusing situation of finding seismic events reported by renown agencies with magnitude 4.0 or 5.5 being excluded from the present analysis. Acknowledging that magnitude estimates are often reported with up to one decimal place, the limits effectively used to define whether an earthquake fell in the magnitude range of interest or not were 3.95≤**M**<5.55. In Version 1, a stricter 4.0≤**M**≤5.5 had been used instead.

Flagging of induced events was carried out both by means of keywords from the ISC Bulletin identified using the toolkit published alongside the paper of Weatherill *et al.* (2016), as for Version 1 (though adding the keyword “rockburst”), and making use of the Human-Induced Earthquake Database (HiQuake) (Foulger *et al.*, 2016; Foulger *et al.*, 2018; Wilson *et al.*, 2017; Induced Earthquakes, 2017) and the flagging of WPG16v3c, as will be explained in Section 4.7.

The Database of Damaging Small-to-Medium Magnitude Earthquakes (Nievas *et al.*, 2019) has significantly increased in volume between Versions 1 and 2 of this work, particularly due to the addition of earthquakes listed in the Earthquake Impact Database (EID). The characteristics of the earthquakes that make up this database and fall into the time period of interest for the present work are described in Section 4.3

3. SENSITIVITY ANALYSES OF VERSION 1

3.1 General

This chapter explores the influence of three (sets of) variables on the final proportion of damaging to total earthquakes obtained in Version 1 of this work: (i) the depth criterion, (ii) the declustering algorithm and parameters therein used, and (iii) the population thresholds.

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 3.1 below 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. "Catalogue" makes reference to the database of crustal earthquakes with magnitude **M**4.0-5.5 for the period 1st July 1999 – 30th June 2014, while "Damaging" indicates those events that are found within the Database of Damaging Small-to-Medium Magnitude Earthquakes as well. "Depth" indicates that the depth criteria have been applied to filter the database (*i.e.*, magnitude-dependent maximum depth limits as per Table 3.2 below), while "Exposure" indicates that depth and exposure criteria (*i.e.*, earthquakes with at least 2,500 people or a maximum density of at least 300 people/km^2 exposed to MMI≥IV) have been applied to carry out the filtering. Figure 3.1 depicts a schematic representation of these filtering levels. The four right-most columns 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, B, C… in Table 3.1.

Apart from results corresponding to each alternative considered, each results table contains the values obtained in Version 1 as well. Table 3.1 indicates where these values can be found in the report for Version 1 (Nievas *et al.*, 2017).

Table 3.1. Generic table for the comparison of alternatives of the sensitivity analysis.

Criteria	Catalogue Depth		Catalogue Exposure		Damaging Depth		Damaging Exposure		Depth		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*
Original v1	101,248	32,842	35,654	11,968	287	N/A	282	185	0.28%	N/A	0.79%	1.55%
Version 1	Table 2.6		Table 2.7		Table 4.3 & S. 4.1.2	N/A	Table 4.3		N/A		Table 4.5	

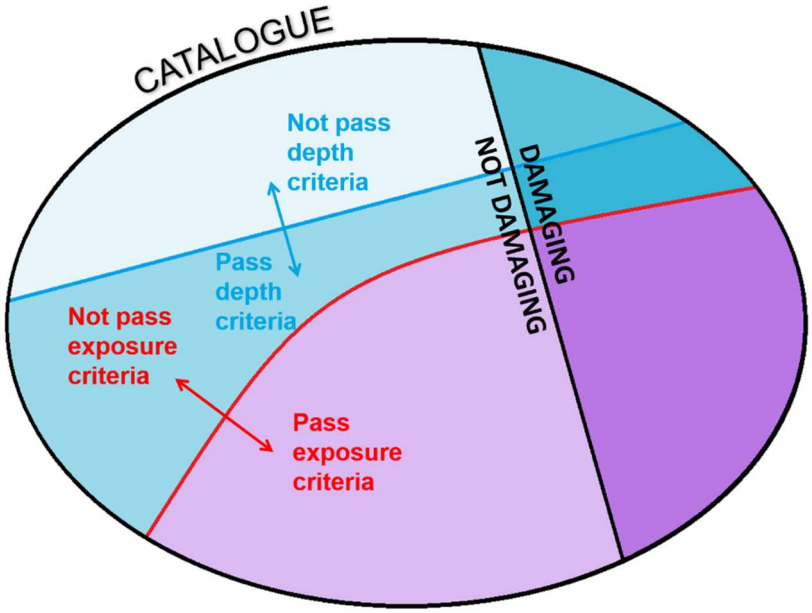

Figure 3.1. Schematic representation of filtering levels used in Table 3.1 and all results tables.

3.2 Maximum Depth Criteria

3.2.1 Maximum Depth Criterion in Version 1

The maximum depth criterion adopted in Version 1 was defined by the step function in Table 3.2.

Table 3.2. Adopted depth criterion.

Magnitude Range	Maximum Depth
4.0 ≤ **M** < 4.5	15 km
4.5 ≤ **M** < 5.0	20 km
5.0 ≤ **M** ≤ 5.5	25 km

3.2.2 Constraints and Considerations

Due to the characteristics of induced seismicity, only upper crustal events are being considered in the elaboration of the database. However, the definition of a threshold for a hypocentral depth to be considered an upper crustal event or not is not trivial and is heavily influenced by the inherent uncertainty associated with the determination of hypocentral depths, particularly for small magnitude events occurring in areas with low density of seismograph networks.

The main natural physical constraint is imposed by the depth to Moho, below which it is not possible to sustain a shear rupture due to the medium becoming partially ductile. A look at the map with the model of the depth to Moho elaborated by Artemieva & Thybo (2013) for Europe, Greenland and the North Atlantic region suggests that depths of up to 35 km might

be reasonable below land (*i.e.*, in non-oceanic environments) in most seismic areas, though smaller areas around the Alps and the Balkans have a deeper Moho. This value is in agreement with the seismogenic thicknesses (which are mostly shallower than the depth to Moho) plotted by Stafford (2014) using a dataset consisting of sub-set of events from the NGA-West 2 project (Ancheta *et al*., 2013) and a series of additional strike-slip events considered by Hanks & Bakun (2002, 2008).

A fundamental issue to bear in mind when working with hypocentral depths is that of the difficulties associated with constraining this quantity during the location process. Depth can only be resolved when the available records contain sufficient information. For example, the ISC only attempts to compute 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 & Storchak, 2011). In the past, however, the location algorithm of the ISC did always seek a free-depth solution before moving onto a fixed one if the latter did not converge, as described by Adams *et al.* (1982). A corollary of this is that it may be possible to find earthquakes in the ISC Bulletin for which a free-depth solution is reported but for which this value is not of great accuracy. The way in which fixed depths are selected has evolved in time as well. Before 2009, the depth was first fixed to the depth determined by means of the depth-phase stack technique, then to a depth reported by an external agency if the former did not work, and finally to a conventional value of 0 or 33 km (Adams *et al.*, 1982). A presentation by Bolton *et al.* (2007) suggests that 0 was used for explosions, 10 for mid-ocean ridge events, and 33 km in all other cases. The choice of conventional depths eventually extended to be 0, 10, 33, 50, 100, 150, 200, 250, *etc*. From around 2002, while keeping the choice of values from the ones enumerated above, the practice of selecting always the same value for the same region became more established (Storchak, 2018, *personal comm.*). Since 2009 until the present, depths are fixed with the following order of preference: a location-specific value derived from statistical data of past events of the ISC, the median of the reported depths if the initial depth is larger than 100 km, a conventional depth assigned to the corresponding Flinn-Engdahl geographic region (Young *et al.*, 1996). As a consequence, it is not really possible, for example, to know whether an earthquake that has been assigned a standard value such as 33 km was shallow or not. This is not an issue of the ISC procedures, but an inherent difficulty that is faced whenever the available waveforms are not sufficient to accurately constrain hypocentral depth. While the present description focuses on the practices of the ISC, other agencies have their own analogous procedures. The United States Geological Survey currently adopts 10 km as a fixed depth, though a value of 33 km was used in the past (USGS, 2018).

An important consideration to be made when setting the maximum depth criteria is that depth has a direct influence on the macroseismic intensities observed and it is thus a significant parameter for Intensity Prediction Equations (IPE). In Version 1 it was decided to use the population count and the maximum population density enclosed by isoseismals of Modified Mercalli Intensity (MMI) of IV or higher to determine if an earthquake was sufficiently close to inhabited areas to pose a potential threat (see further discussion in Section 4.6.4). Consequently, it is of interest to take into consideration the maximum depths that are expected to generate a median MMI of at least IV right at the epicentre. As shown in Figure 3.2, these depths vary between around 15 km for **M**4.0 and 70 km for **M**5.5, and

the relationship between maximum depth and moment magnitude can be represented by a parabola. The difference between this relation and that established by Table 3.2 implies that many earthquakes can be discarded according to Table 3.2 and still cause MMI values of IV.

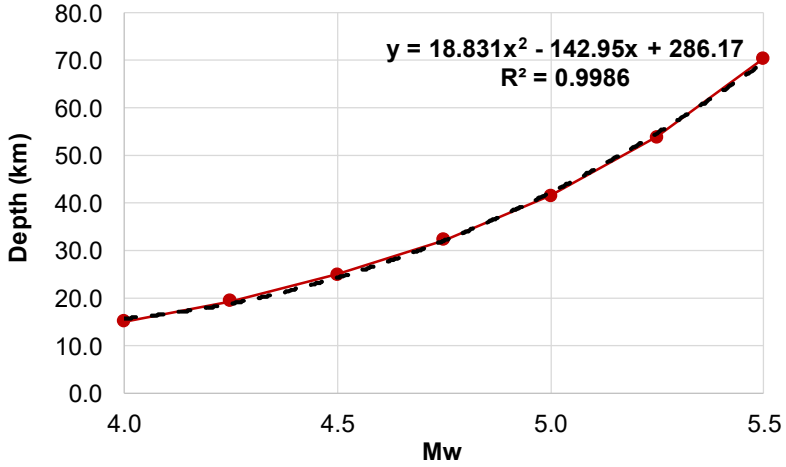

Figure 3.2. MMI IV maximum depth criterion. Red markers indicate values calculated directly from the IPE of Allen et al. (2012). Dashed black lines indicate approximate parabola.

A final factor to consider in this sensitivity analysis is the observed depths of earthquakes reported to have been damaging. For this purpose, it is of interest to take a look at the magnitude-depth distribution of earthquakes that make up the Significant Earthquake Database of the National Oceanic and Atmospheric Admnistration (NOAA) of the United States (NOAA database hereafter), plotted for the magnitude range of interest in Figure 3.3. Being part of the NOAA database, these earthquakes are reported to have caused damage and/or casualties. It is interesting to note that depths around 26-30 km, which represent the immediate upper-bound of the maximum depths defined in Table 3.2, represent the 50th percentile of the cases. While it is clear that the number of earthquakes in the lower magnitude range is not enough to make general statistical inferences (note the scarcity of data in Figure 3.3 for the lower magnitudes), in the middle-to-upper range this implies that around half of the earthquakes that are reported to have had some kind of consequence in the NOAA database would get discarded with the criteria set in Table 3.2. Looking at the data, the limits set by the 75th and 80th percentiles (*i.e.* values of depth not exceeded by 75 and 80% of the earthquakes, respectively) appear of interest for further exploration, because they include a larger number of earthquakes while still being relatively reasonable numerically (*i.e.*, they are not extremely unreliable as a value of, say, 200 km would be). These observations do not imply that extremely deep medium-magnitude earthquakes are actually capable of causing damage (as the depth estimates could potentially be inaccurate) but are rather an example of how the uncertainty associated with depth determination may affect an analysis of the kind we are carrying out herein.

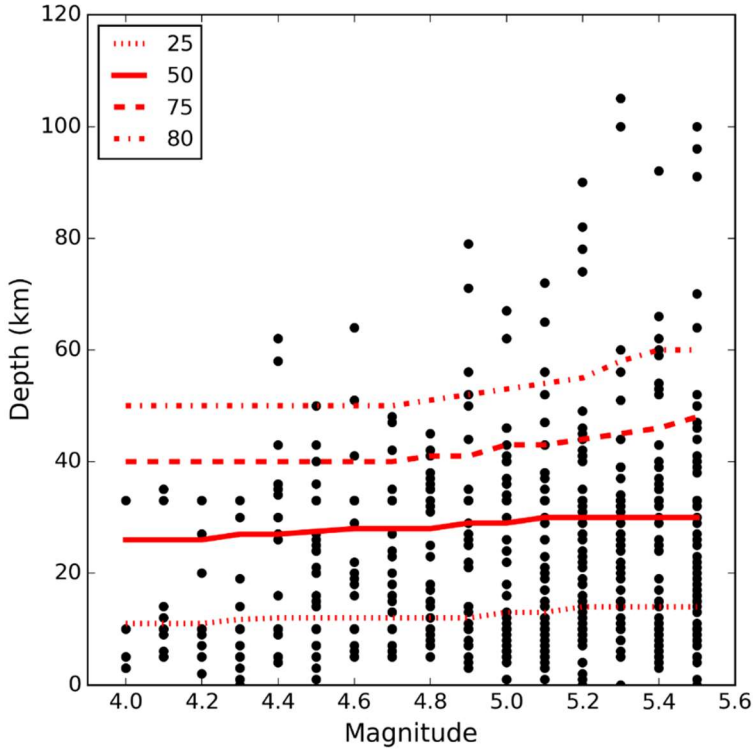

Figure 3.3. Magnitude-depth distribution of earthquakes in the NOAA database with preferred (**M**>M_s>m_b>M_L>M_{fa}>M_{unk}) magnitude in the range 4.0-5.5. Red lines indicate depth percentiles as per legend.

3.2.3 Alternative Maximum Depth Criteria

Apart from the criteria set by Table 3.2, hereafter labelled "Original", the following alternative criteria were used for the sensitivity analysis, based on the discussion above:

- **MMI IV:** Maximum depth that causes the expected MMI at the epicentre to be IV according to the IPE (median) of Allen *et al.* (2012). This criterion yields, approximately, depth limits of 15, 25, 40 and 70 km at magnitudes 4.0, 4.5, 5.0 and 5.5. For the sake of simplicity, the criterion was approximated by the parabola shown in Figure 3.2.

- **NOAA 50, NOAA 75 and NOAA 80:** Different percentiles of the depth distribution of earthquakes in the range **M**4.0-5.5 reported to have caused damage and/or casualties in the NOAA database (linear regression used):
    - 50^{th} percentile, linearly varying between 25.9 and 30.6 km, *i.e.*, values not exceeded by 50% of the earthquakes with **M**4.0-5.5 in the NOAA database;
    - 75^{th} percentile, linearly varying between 38.3 and 45.6 km, *i.e.*, values not exceeded by 75% of the earthquakes with **M**4.0-5.5 in the NOAA database;
    - 80^{th} percentile, linearly varying between 47.6 and 57.8 km, *i.e.*, values not exceeded by 80% of the earthquakes with **M**4.0-5.5 in the NOAA database.

- **MAX 35km:** Representative of the depth to Moho (seismogenic depth) in central and western Europe and containing the standard fixed depth of 33 km used in the past.

- **Linear:** a linear variation from 15 km at **M**4.0 to 35 km at **M**5.5.

- **Original:** the criterion used in Version 1, defined in Table 3.2.

Figure 3.4 shows a summary of the seven alternative maximum depth criteria used for the sensitivity analysis. It is noted that these alternative criteria were defined for the purpose of carrying out the present sensitivity analysis and their listing here does not imply that they are all equally reasonable or worth of consideration. The goal of including extreme cases such as NOAA's 80th percentile was to observe the behaviour of the results when pushing decisions to the edge.

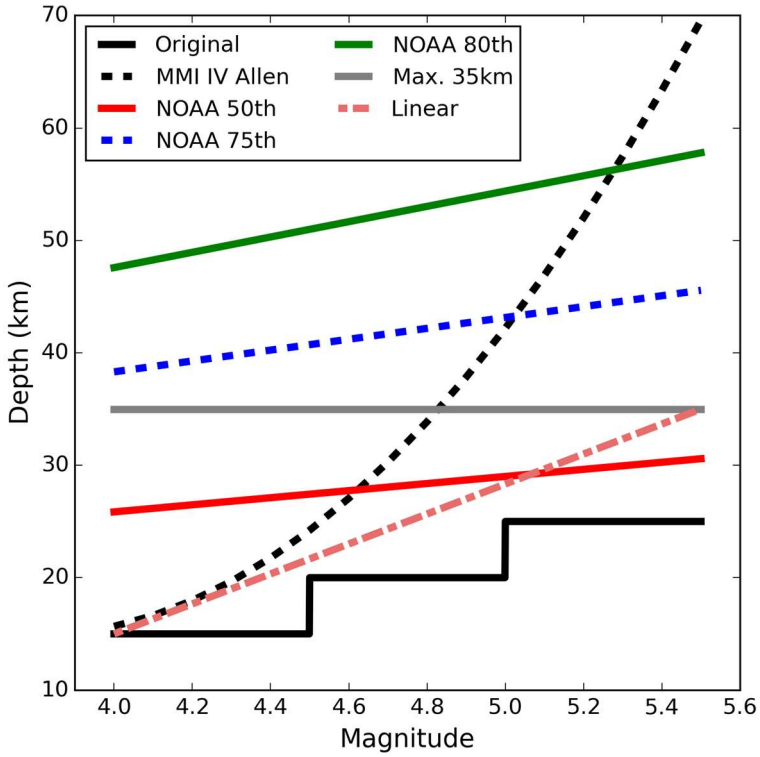

Figure 3.4. The six alternative maximum depth criteria used for the sensitivity analysis.

3.2.4 Results

Table 3.3 presents the results obtained using each of the alternative maximum depth criteria described above (please refer to Section 3.1 and Table 3.1 for guidance on the interpretation of Table 3.3). The row labelled "Original" corresponds to the results obtained when applying the depth criterion defined in Table 3.2. The numbers should be the same as those in the row labelled "Original v1", which come directly from Version 1, but vary slightly in the columns corresponding to the damaging events. This is due to the presence of potentially duplicate events and the change in the order in which the filtering was carried out in Version 1 and for the present sensitivity analysis for the sake of simplicity. As a consequence, the row labelled "Original" is taken as the benchmark, representative of the original work. It is noted that, while the number of earthquakes included in the database increases when the

maximum depth criterion is more relaxed (*i.e.*, when deeper 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 case. However, when adding the exposure criterion, the proportions of damaging to total remain much more stable (0.77-0.80% and 1.54-1.64%).

Table 3.3. Results for the sensitivity analysis over the maximum depth criterion. Catalogue vs. Damaging refers to the complete catalogue and the subset of it reported to have caused damage. All vs. MS refers to all shocks and only main shocks, respectively. Depth refers to earthquakes being filtered according to the maximum depth criteria. Exposure refers to earthquakes filtered both by depth and exposure criteria. The last four columns depict the proportion of damaging to total earthquakes.

Criteria	Catalogue Depth		Catalogue Exposure		Damaging Depth		Damaging Exposure		Depth		Exposure	
	All	MS	All	MS	All	MS	All	MS	All	MS	All	MS
Original v1	101,248	32,842	35,654	11,968	287	N/A	282	185	0.28%	N/A	0.79%	1.55%
Original	101,248	32,842	35,654	11,968	285	188	280	184	0.28%	0.57%	0.79%	1.54%
MMI IV	141,176	42,049	43,815	14,429	355	240	349	235	0.25%	0.57%	0.80%	1.63%
NOAA 50	125,335	37,832	40,279	13,253	314	207	309	203	0.25%	0.55%	0.77%	1.53%
NOAA 75	198,957	56,704	43,577	14,343	359	243	342	231	0.18%	0.43%	0.78%	1.61%
NOAA 80	218,894	61,440	43,825	14,432	374	253	349	236	0.17%	0.41%	0.80%	1.64%
MAX 35km	186,699	53,989	43,008	14,145	342	231	330	221	0.18%	0.43%	0.77%	1.56%
Linear	113,316	35,529	38,979	12,903	316	209	311	205	0.28%	0.59%	0.80%	1.59%

Before the application of the exposure criterion (third and fourth columns from right to left), two groups of relatively comparable proportions of damaging to total earthquakes can be identified: the 0.25-0.28% range resulting from the Original, MMI IV, NOAA50 and Linear criteria, and the 0.17-0.18% range that the NOAA75, NOAA80 and MAX 35km criteria produce. A first observation of interest is that this occurs even if the individual numbers of earthquakes in the database do change. For example, the MMI IV criterion yields a total number of depth-filtered earthquakes 39.4% larger than the Original criterion (141,176 vs. 101,248), but the proportion of damaging to all decreases only from 0.28% to 0.25%. Causing the same decrease in the proportion, the NOAA50 criterion features only a 23.8% increase in the total number of earthquakes (125,335 vs. 101,248). A second observation of interest is that the MMI IV and Linear criteria are more restrictive in the low-magnitude range and much more flexible in the medium-to-larger magnitude range, and still produce percentages comparable to those from the Original and NOAA50 criteria.

The significant decrease in the proportion of damaging to total earthquakes when only depth filtering is applied observed for the NOAA75, NOAA80 and MAX 35km criteria clearly show the consequence of incorporating a larger number of smaller-magnitude earthquakes to the world catalogue, earthquakes with a clearly low capacity of causing damage or casualties. As the two right-most columns of Table 3.3 show, the exposure criterion then takes care of these cases, as its application is equivalent to imposing the MMI IV criterion. The focus in this observation is put on smaller magnitude earthquakes because of the large difference in the depth-filtered results obtained with these three criteria against those of the MMI IV

criterion, which, as shown in Figure 3.4, is comparably lose in the larger-magnitude range, or even more so, but much more restrictive for the lower magnitudes.

3.2.5 Conclusions

Given the relative stability of the results once the exposure criterion is applied due to the latter acting like an implicit maximum-depth limit (see Section 4.6.4 for further discussion on this matter), defining which criterion to finally adopt is more of a conceptual decision than a practical one if the exposure criterion is held independent from the depth criterion. In this scenario, adopting criteria such as NOAA 50, NOAA 75, NOAA 80 or MAX 35km could be tempting under the pretext of not excluding potential earthquakes of interest. However, they lack a physical meaning (especially the two former) in terms of why those limits would represent upper-crustal earthquakes capable of posing a threat to the built environment. The MMI IV criterion, on the other hand, would be consistent with the exposure criterion used in Version 1, but would likely allow to include too many subduction earthquakes that are not of interest for the present study. While the fact that many earthquakes known to have caused damage or casualties may be discarded on the basis of a seemingly restrictive maximum depth limit is acknowledged herein, being able to include some likely-misplaced hypocentres does not justify the unrealistic increase in the overall number of earthquakes considered to pose a threat to the population and the built environment. Acknowledging that a 25-km limit might be too restrictive for **M**5.5 in light of the discussion of Section 3.2.2 and that a linearly-varying function would probably be more sensible than a step one, a new maximum depth criterion varying linearly from 15 km at **M**4.0 to 35 km at **M**5.5 was adopted for Version 2. As discussed earlier, this is consistent with mapped depths to Moho and reported seismogenic thicknesses of shallow crustal earthquakes. Further details can be found in Section 4.5.2.

As the exposure criterion was modified for Version 2 as well, this sensitivity analysis is repeated in Section 4.6.5.

3.3 Declustering

3.3.1 General Considerations

Declustering is the process by which Poissonian events are separated from non-Poissonian, *i.e.,* dependent, events. While in Probabilistic Seismic Hazard Assessment (PSHA) this is done to be able to calculate earthquake rates from independent events, the purpose herein is to avoid focusing on events that fall within our magnitude range of interest but are aftershocks of bigger events. There are many declustering algorithms available in the literature, of which two were considered: that of Gardner & Knopoff (1974) and the Afteran algorithm of Musson (1999). Both have been applied from their implementation in the OpenQuake Hazard Modeller's Toolkit (Weatherill, 2014), modified to take into consideration hypocentral depth.

The algorithm of Gardner & Knopoff (1974) requires three parameters: the kind of distance window, the kind of time window and the foreshock parameter. Three distance and time window models are available in OpenQuake: those proposed by Gardner & Knopoff (1974), those of Grünthal (as reported in van Stiphout *et al.*, 2012) and those of Uhrhammer (1986). Figure 3.5 shows the definition of these models, all of which are magnitude-dependent. The implementation in OpenQuake requires for the same kind of window (*i.e.*, by the same author) to be applied both in time and space, reducing the three input parameters to effectively two.

The foreshock parameter (fs) is used to define whether foreshocks are sought (fs>0) or not (fs=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.

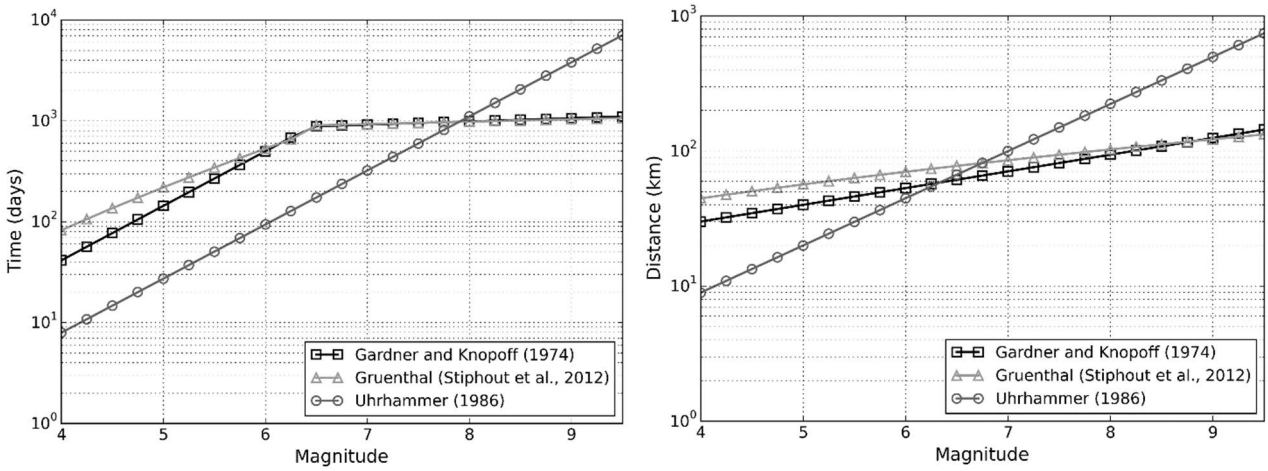

Figure 3.5. Time (left) and distance (right) windows according to the three models used.

The Afteran algorithm (Musson, 1999) uses a moving time window (tw) instead of a fixed one. While the algorithm of Gardner & Knopoff (1974) defines the time window based on the magnitude of what is considered a potential main shock and only looks for earthquakes within that time window, the Afteran algorithm first applies the window starting to count from the potential main shock and then applies it again starting from the last identified aftershock, the process being repeated until no additional aftershocks have been identified within the last added time window. Little guidance exists with respect to the length of the time window to use. Musson (1999) applied a time window of 100 days, based on the inspection of known clusters, though he admits that there was still some subjectivity in that choice. By default, the algorithm applies the same time windows and philosophy to search for foreshocks. The distance window can be any of the three described above (see Figure 3.5, right).

3.3.2 Alternative Declustering Options

Table 3.4 shows the 15 possibilities considered for this sensitivity analysis, the first row corresponding to what was done in Version 1. Two possible foreshock parameters were considered within the Gardner & Knopoff (1974) algorithm, namely 100% and 50% of the aftershock time window. Three moving time windows were used with the Afteran algorithm:

the original 100 days proposed by Musson (1999), 200 days, representing the situation in stable continental regions, and 20 days, representing a minimum conceptually admissible.

Table 3.4. Alternative declustering setups considered in this sensitivity analysis.

Algorithm	Distance Window	Time Window	Foreshocks Code
Gardner & Knopoff (1974)	Gardner & Knopoff (1974)	Gardner & Knopoff (1974)	1.0
			0.5
	Uhrhammer (1986)	Uhrhammer (1986)	1.0
			0.5
	Grünthal (van Stiphout *et al* ., 2012)	Grünthal (van Stiphout *et al* ., 2012)	1.0
			0.5
Afteran (Musson, 1999)	Gardner & Knopoff (1974)	20.0	-
		100.0	-
		200.0	-
	Uhrhammer (1986)	20.0	-
		100.0	-
		200.0	-
	Grünthal (van Stiphout *et al* ., 2012)	20.0	-
		100.0	-
		200.0	-

3.3.3 Results

Table 3.5 summarises the results obtained. Please refer to Section 3.1 and Table 3.1 for guidance on the interpretation of Table 3.5 To aid visualization, Figures 3.6 and 3.7 depict the proportion of damaging to total main shocks obtained with each alternative declustering option. As the declustering process affects only the number of earthquakes identified as main shocks but not the number of earthquakes in the catalogue as a whole, columns of Table 3.5 labelled “All” (*i.e.*, including main, fore- and aftershocks) do not vary.

In general terms, the Uhrhammer (1986) windows produce the largest number of events identified as main shocks, followed by the Gardner & 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 when the Grünthal windows (van Stiphout *et al.*, 2012) are used instead. However, the choice of the time window in the Afteran algorithm influences this tendency, which is much more marked when the Gardner & Knopoff (1974) algorithm is used instead. It is worth noting at this point that the foreshock parameter (fs) does not appear to have a significant influence on the results from Gardner & Knopoff (1974).

Within the results stemming from the use of the Afteran algorithm, those obtained with a time window of 20 days yield the largest number of main shocks. Due to the larger stability of the number of main shocks within the damaging subset, this is then reflected in some of the lowest proportions of damaging to total being obtained.

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 0.38 and 0.74% when filtering only according to depth and between 1.12 and 1.93% when filtering according to exposure as well.

Table 3.5. Results for the sensitivity analysis over the declustering decisions. Catalogue vs. Damaging refers to the complete catalogue and the subset of it reported to have caused damage. All vs. MS refers to all shocks and only main shocks, respectively. Depth refers to earthquakes being filtered according to the maximum depth criteria. Exposure refers to earthquakes filtered both by depth and exposure criteria. The last four columns depict the proportion of damaging to total earthquakes. GK= Gardner & Knopoff (1974). Uhr= Uhrhammer (1986). Grü: Grünthal (van Stiphout *et al.*, 2012).

Criteria	Catalogue Depth		Catalogue Exposure		Damaging Depth		Damaging Exposure		Depth		Exposure	
	All	MS	All	MS	All	MS	All	MS	All	MS	All	MS
GK - GK - fs=1.0 (Original)	101,248	32,842	35,654	11,968	285	188	280	184	0.28%	0.57%	0.79%	1.54%
GK - GK - fs=0.5	101,248	35,354	35,654	12,735	285	197	280	193	0.28%	0.56%	0.79%	1.52%
GK - Uhr - fs=1.0	101,248	48,845	35,654	16,799	285	213	280	209	0.28%	0.44%	0.79%	1.24%
GK - Uhr - fs=0.5	101,248	50,384	35,654	17,270	285	216	280	212	0.28%	0.43%	0.79%	1.23%
GK - Grü - fs=1.0	101,248	22,845	35,654	8,564	285	168	280	165	0.28%	0.74%	0.79%	1.93%
GK - Grü - fs=0.5	101,248	25,399	35,654	9,435	285	179	280	175	0.28%	0.70%	0.79%	1.85%
Afteran - GK - tw=20.0	101,248	59,159	35,654	19,925	285	233	280	229	0.28%	0.39%	0.79%	1.15%
Afteran - GK - tw=100.0	101,248	41,332	35,654	13,945	285	189	280	186	0.28%	0.46%	0.79%	1.33%
Afteran - GK - tw=200.0	101,248	32,731	35,654	11,128	285	167	280	164	0.28%	0.51%	0.79%	1.47%
Afteran - Uhr - tw=20.0	101,248	59,131	35,654	19,692	285	225	280	220	0.28%	0.38%	0.79%	1.12%
Afteran - Uhr - tw=100.0	101,248	45,286	35,654	15,119	285	186	280	182	0.28%	0.41%	0.79%	1.20%
Afteran - Uhr - tw=200.0	101,248	39,049	35,654	13,166	285	171	280	167	0.28%	0.44%	0.79%	1.27%
Afteran - Grü - tw=20.0	101,248	53,600	35,654	18,036	285	228	280	224	0.28%	0.43%	0.79%	1.24%
Afteran - Grü - tw=100.0	101,248	33,369	35,654	11,041	285	170	280	166	0.28%	0.51%	0.79%	1.50%
Afteran - Grü - tw=200.0	101,248	24,970	35,654	8,353	285	148	280	144	0.28%	0.59%	0.79%	1.72%

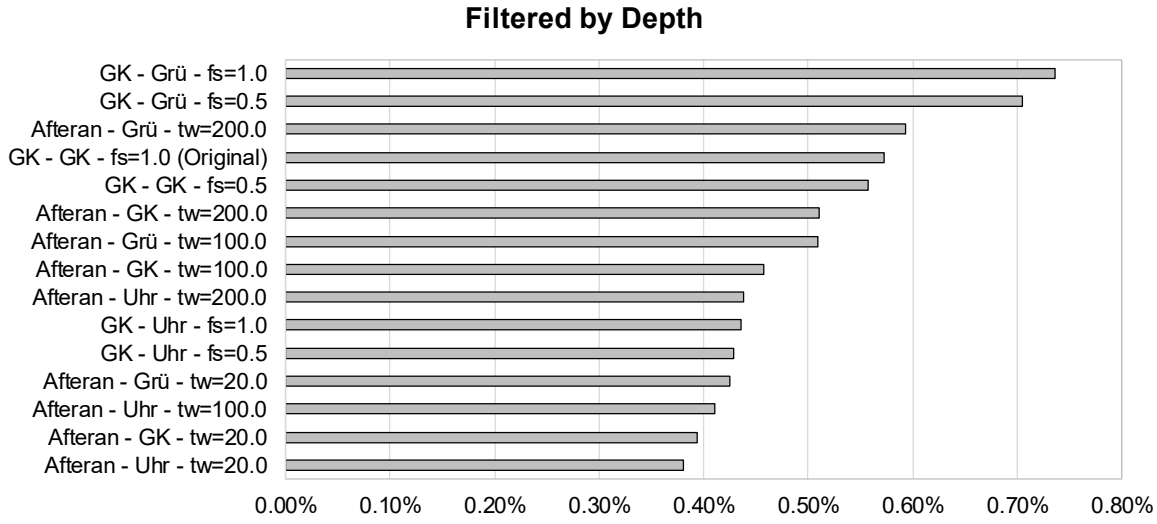

Figure 3.6. Proportion of damaging to total number of main shocks filtered according to maximum depth criteria.

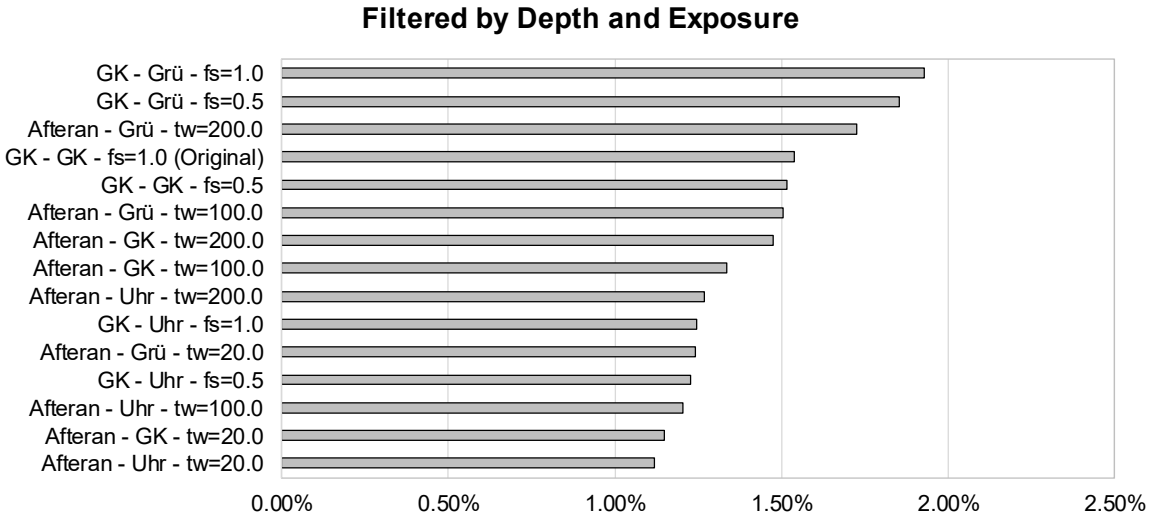

Figure 3.7. Proportion of damaging to total number of main shocks filtered according to depth and exposure criteria.

3.3.4 Conclusions

In view of the results presented above, it was decided to keep on using the declustering algorithm of Gardner & Knopoff (1974) with Gardner & Knopoff (1974) time and space windows and a foreshock parameter of 1.0 for Version 2, as had been done in Version 1. The reasons for this are the following:

- Results from this particular choice lie reasonably within the middle of the sensitivity study (78th percentile).
- The algorithm of Gardner & Knopoff (1974) is a well-established and broadly used declustering technique, much more frequently applied than the Afteran algorithm (Musson 1999).
- Within the alternative time-space windows and foreshock parameters considered, this choice lies strictly in the middle.

It cannot be over-emphasised that there is no right answer in terms of declustering algorithms. Being based only on empirically-derived temporal and spatial windows and ignorant on fault geometry and other physical constraints, one can only hope for a sufficiently reasonable classification of earthquakes into main shocks, foreshocks and aftershocks.

3.4 Exposure Thresholds

3.4.1 General Considerations

As briefly described in the Introduction (Chapter 1) and Section 2.1, whether an earthquake posed a threat to the population and the built environment or not was gauged in Version 1 in terms of the population count and maximum population density exposed to estimated Modified Mercalli Intensities (MMI) of IV or larger. Intensity Prediction Equations (IPEs) were

used to determine the latter, and Gridded Population of the World (GPW) v4.0 (CIESIN, 2016) allowed to determine the former. The thresholds to consider the exposure to have been large *enough* to pose a real threat were established by investigating the definitions of urbanisation used by different sources. As described in the report for Version 1 (Section 2.2 therein), defining what an urban settlement is and making the distinction between urban and rural populations is neither trivial nor objective. The following criteria had been discussed back then:

- UNICEF (2012): The minimum population to define an urban settlement is around 2,000 people, though this number varies greatly around the world and can range between 200 and 50,000.
- United States Census Bureau, 2010 census: Urbanised area defined as that having 50,000 people or more and a density of at least 1,000 people per square mile (386 people/km^2), and urban cluster defined as that having between 2,500 and 50,000 people.
- European Union (Eurostat; Dijkstra & Poelman, 2018): Urban areas defined by first identifying grid cells of 1km^2 in which the population density is equal to or larger than 300 people/km^2, grouping adjoining cells that satisfy this criterion and verifying that the resulting group of cells adds up to, at least, 5,000 people.

Recognising the relevance of the size of the grid used to calculate the density (highlighted by many of the sources), a maximum density of at least 300 people/km^2 exposed to MMI values of at least IV was adopted for Version 1, as Eurostat specifies the grid cell size and it coincides at the Equator with that of GPW v4.0, used in both Versions 1 and 2 of the present work (GPW v4.0 uses a 30-arcsec resolution, which implies that the area of the cells decreases when moving towards the Poles). A minimum cumulative population count of 2,500 was selected as well, as it corresponded to the threshold to define an urban cluster according to the United States Census Bureau, and is only slightly higher than that mentioned by UNICEF (2012). The threshold of 5,000 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 places of the world.

3.4.2 Alternative Thresholds

The following thresholds (summarised in Table 3.6) were considered in the present sensitivity analysis:

- **Criterion 01:** The one applied in Version 1. Population equal to or larger than 2,500 people or maximum density equal to or larger than 300 people/km^2 exposed to estimated MMI equal to or larger than IV.
- **Criterion 01L:** 10% lower thresholds than Criterion 01 (*i.e.*, population equal to or larger than 2,250 people or maximum density equal to or larger than 270 people/km^2 exposed to estimated MMI equal to or larger than IV).

- **Criterion 01H:** 10% higher thresholds than Criterion 01 (*i.e.*, population equal to or larger than 2,750 people or maximum density equal to or larger than 330 people/km^2 exposed to estimated MMI equal to or larger than IV).
- **Criterion 02:** Drastic arbitrary reduction in the population count threshold, keeping the maximum density threshold. Population equal to or larger than 1,000 people or maximum density equal to or larger than 300 people/km^2 exposed to estimated MMI equal to or larger than IV
- **Criterion 03:** Drastic arbitrary reduction in the maximum density threshold, keeping the population count threshold. Population equal to or larger than 2,500 people or maximum density equal to or larger than 150 people/km^2 exposed to estimated MMI equal to or larger than IV
- **Criterion 04:** Using the maximum density of the United States Census Bureau (1,000 people/mile^2 ≈ 400 people/km^2). Population equal to or larger than 2,500 people or maximum density equal to or larger than 400 people/km^2 exposed to estimated MMI equal to or larger than IV
- **Criterion 05:** Using the population count threshold defined by the European Union. Population equal to or larger than 5,000 people or maximum density equal to or larger than 300 people/km^2 exposed to estimated MMI equal to or larger than IV
- **Criterion 06:** Drastic arbitrary reduction in both thresholds. Population equal to or larger than 1,000 people or maximum density equal to or larger than 150 people/km^2 exposed to estimated MMI equal to or larger than IV
- **Criterion 07:** Extreme case of considering a threshold of 1 person or 1 person/km^2 exposed to estimated MMI equal to or larger than IV as sufficient to consider the earthquake a threat.

Table 3.6. Summary of the alternative thresholds of cumulative population and maximum density considered herein.

Criteria	Cumulative Population	Maximum Density (ppl/km²)
Crit. 01	2,500	300
Crit. 01L	**2,250**	**270**
Crit. 01H	**2,750**	**330**
Crit. 02	**1,000**	300
Crit. 03	2,500	**150**
Crit. 04	2,500	**400**
Crit. 05	**5,000**	300
Crit. 06	**1,000**	**150**
Crit. 07	**1**	**1**

3.4.3 Results

Table 3.7 summarises the results obtained with the different population thresholds defined above. Please refer to Section 3.1 and Table 3.1 for guidance on the interpretation of Table 3.7. As the difference between different alternatives lies only in the exposure criterion, all columns related to the application of the maximum depth criterion only remain the same. Leaving the extreme Criterion 07 aside for the time being, the largest increase in the number of earthquakes that satisfy the exposure criterion with respect to Version 1 (Crit. 01) occurs for Criterion 06 and represents 5.6% of the total and 5.3% of the main shocks, followed by Criterion 02, which represents a 5.2% of the total and a 4.7% of the main shocks. The largest decrease occurs for Criterion 05 and represents a 5.4% of the total and 4.7% of the main shocks. The similarity in the results for Criterion 02 and Criterion 06 suggest that it is the population count threshold that is having the largest impact in the results. Criterion 06 is effectively a combination of Criteria 02 and 03 with a logical OR condition.

Table 3.7. Results for the sensitivity analysis over the thresholds of population exposure. Catalogue vs. Damaging refers to the complete catalogue and the subset of it reported to have caused damage. All vs. MS refers to all shocks and only main shocks, respectively. Depth refers to earthquakes being filtered according to the maximum depth criteria. Exposure refers to earthquakes filtered both by depth and exposure criteria. The last four columns depict the proportion of damaging to total earthquakes.

Criteria	Catalogue Depth		Catalogue Exposure		Damaging Depth		Damaging Exposure		Depth		Exposure	
	All	MS	All	MS	All	MS	All	MS	All	MS	All	MS
Original v1	101,248	32,842	35,654	11,968	287	N/A	282	185	0.28%	N/A	0.79%	1.55%
Crit. 01	101,248	32,842	35,654	11,968	287	189	282	185	0.28%	0.58%	0.79%	1.55%
Crit. 01L	101,248	32,842	35,981	12,052	287	189	282	185	0.28%	0.58%	0.78%	1.54%
Crit. 01H	101,248	32,842	35,377	11,881	287	189	282	185	0.28%	0.58%	0.80%	1.56%
Crit. 02	101,248	32,842	37,503	12,534	287	189	282	185	0.28%	0.58%	0.75%	1.48%
Crit. 03	101,248	32,842	35,921	12,075	287	189	282	185	0.28%	0.58%	0.79%	1.53%
Crit. 04	101,248	32,842	35,557	11,925	287	189	282	185	0.28%	0.58%	0.79%	1.55%
Crit. 05	101,248	32,842	33,733	11,403	287	189	282	185	0.28%	0.58%	0.84%	1.62%
Crit. 06	101,248	32,842	37,658	12,599	287	189	282	185	0.28%	0.58%	0.75%	1.47%
Crit. 07	101,248	32,842	41,918	13,963	287	189	283	186	0.28%	0.58%	0.68%	1.33%

As can be observed, the number of damaging earthquakes within those that pass the exposure criterion remain the same in all cases except for Crit. 07. The five earthquakes that pass the depth criterion but not the exposure one are those listed in Section 4.1.2 of the report for Version 1, four of them being main shocks and one being classified as a foreshock. Of these five, only the **M**5.0 2010 Porangahau (New Zealand) one is expected to have caused MMI≥IV in areas with some population, the total count being 329 people and the maximum density 128 people/km^2. It is this case that passes exposure criterion 07.

What is most worth of notice is the fact that the change in the number of earthquakes that pass the exposure criterion is not significant for criteria 01 through 06, even when taking very low values of both the population count and maximum density thresholds (*i.e.*, Crit. 06). 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. Results for the criterion used in Version 1 (Crit. 01) are close to the average of the upper and lower bounds of these ranges.

Results for Criterion 07 are, as would be expected, the most extreme deviation from the criterion used in Version 1. Interestingly, though, consideration of only 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 only 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 MMI≥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 Criterion 01 than all other criteria listed in Table 3.6, they show that even extreme criteria do not cause unreasonable variations in the final results.

3.4.4 Conclusions

In view of the stability of the results obtained in this sensitivity analysis, the population count and maximum density thresholds are kept at 2,500 people and 300 people/km^2 as in Version 1. The consideration of Criterion 07 in this analysis has shown that decreasing the thresholds may result in a much larger increase in the total number of earthquakes than in the number of damaging ones.

4. DETAILS ON THE METHODOLOGY AND CHANGES INTRODUCED IN VERSION 2

4.1 General

This chapter, together with the report for Version 1 (Nievas *et al.*, 2017), explain in detail the methodology used to compile the catalogue of potentially damaging upper crustal earthquakes, whose outline has been presented in Section 2.1.

The following aspects have not changed overall from what was described for Version 1:

- the general scheme of compiling a catalogue for a period longer than that of interest and including all magnitudes and depths, which is then declustered and filtered for the magnitude range **M**4.0-5.5, upper-crustal depths and exposure criteria based on predicted population count and density exposed to a certain seismic intensity, determined by means of intensity prediction equations;
- comparison and merging of the WPG16 catalogue and the ISC Bulletin;
- selection of one set of hypocentral coordinates and magnitude estimate per earthquake based on a pre-defined hierarchy of agencies;
- resolution of issues regarding, for example, the existence of magnitude estimates with two authors, repetition of magnitude-scale-author combinations, *etc.*;
- the algorithm used to identify potential duplicate events within the merged catalogue;
- flagging of potentially induced events by searching for keywords in the comments to the ISC Bulletin, as suggested by Weatherill *et al.* (2016);
- population count and density thresholds used.

Aspects that have changed and are thus expanded upon in this chapter are:

- the 15-year period of interest has been shifted from 01/07/1999-30/06/2014 to 01/01/2001-31/12/2015;
- version 3c of the WPG16 catalogue used instead of version 3b;
- consideration of more magnitude scales and use of empirical magnitude conversion models;
- the maximum depth criteria used, which now takes into account uncertainty and differentiates between fixed- and free-depth solutions;
- explicit consideration of uncertainty in magnitude, depth and intensity prediction equations to define uncertainty in intensity;
- definition of magnitude-dependent probability thresholds of observing intensities of IV and V;
- the use of an updated version of the Database of Damaging Small-to-Medium Earthquakes.

Moreover, in Version 2 we have explored the use of the Human-Induced Earthquake Database (HiQuake) (Foulger *et al.*, 2018) for flagging potentially induced events, complementing the flagging done based on keyword-searching within the ISC Bulletin.

4.2 Use of Version 3c of the WPG16 Catalogue

4.2.1 Changes in WPG16 Itself

Version 2 uses version v3c of the catalogue of Weatherill *et al.* (2016), referred to as WPG16v3c hereafter, instead of version v3b. Both versions of WPG16, *i.e.* v3b and v3c, cover the same time period and take the same sources. The modifications between the latter and those described in the paper (Weatherill *et al.*, 2016) are the following (Weatherill, 2017, *pers. comm.*):

- integration of the ISC-GEM v.4.0 catalogue (Storchak *et al.*, 2015; Di Giacomo *et al.*, 2018),
- extension of the Global Centroid Moment Tensor Project (GCMT; Dziewonski *et al.*, 1981; Ekström *et al.*, 2012) catalogue up to the end of 2016,
- extension of the ISC Reviewed Bulletin to June 2014, and
- inclusion of estimations of the National Earthquake Information Center (NEIC) of the United States Geological Survey (USGS), as reported in the ISC Bulletin, up to the end of 2016.

Apart from minor changes on the GEM Earthquake Catalogue Toolkit (Weatherill *et al.*, 2016) used to generate it (Weatherill, 2018, *pers. comm.*), the main difference between version v3b and v3c is that v3c uses the magnitude conversion models reported in the paper of Weatherill *et al.* (2016) (Equations 4.9 through 4.14 in Section 4.4) while v3b used different ones, as depicted in the plots of Figures 4.1 and 4.2. For the case of converting from M_s, both v3b and v3c used two-segment piecewise linear models, but with different regression coefficients (and a different crossover point for the case of values from the ISC). In the case of m_b, v3c uses linear models while v3b applied exponential models instead. The latter tend to saturate for small values of m_b, causing the smallest of m_b to translate into **M**>4.56 (ISC) or **M**>4.11 (NEIC/USGS), as shown in Figure 4.2. WPGv3c is thus deemed more realistic than WPGv3b. The potential consequence of this change could be a reduction in the number of earthquakes from Version 1 to Version 2. Results (Chapter 5) show the exact opposite, most likely due to the latter effect being counterbalanced by the inclusion of m_b and M_d in Version 2, which is discussed in Section 4.4.

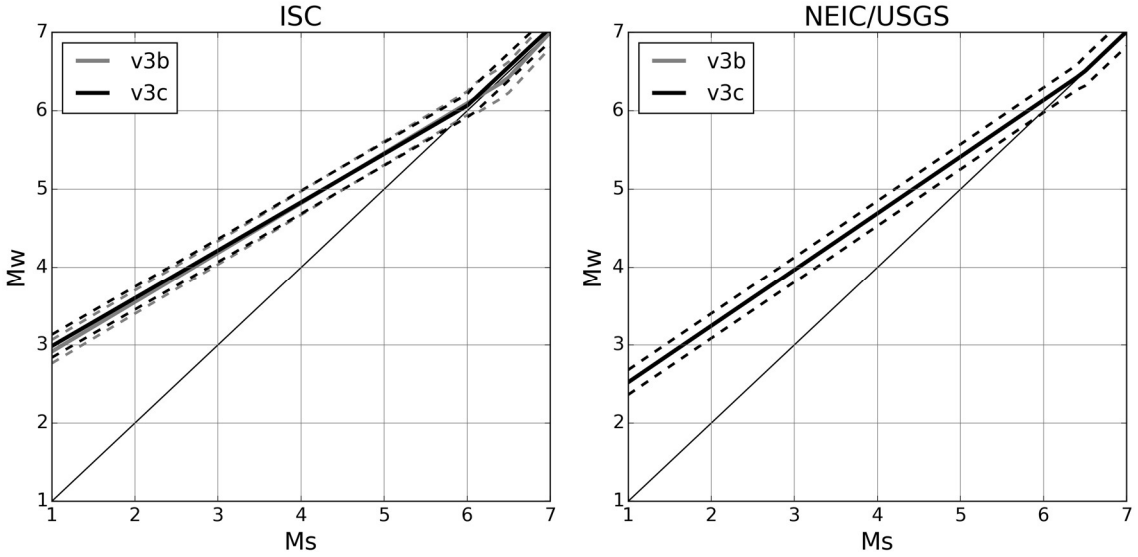

Figure 4.1. Relation between M_s (from the ISC, left, and NEIC/USGS, right) and **M** (GCMT) according to the models fitted by Weatherill et al. (2016).

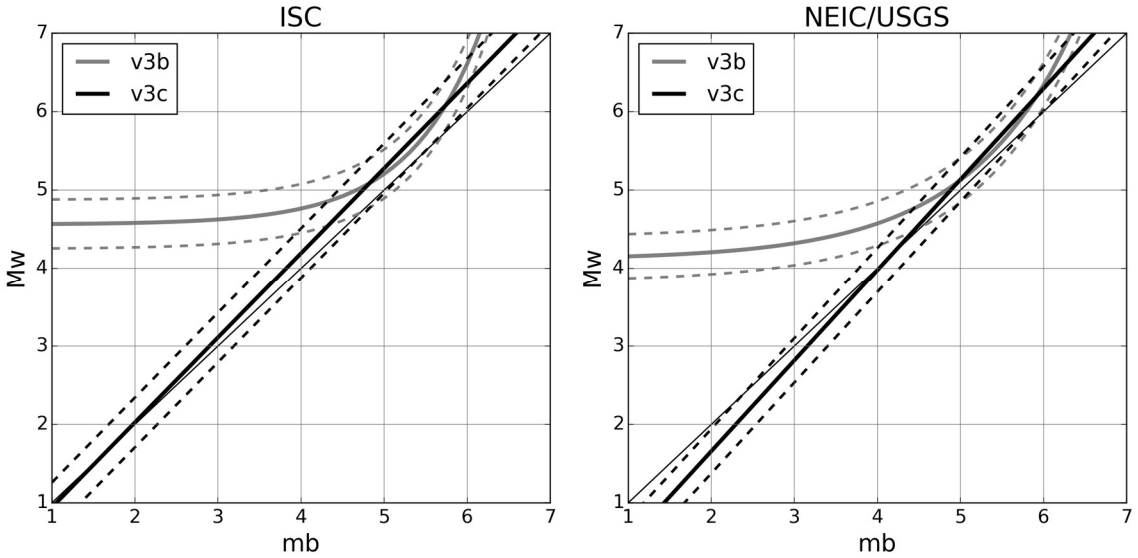

Figure 4.2. Relation between m_b (from the ISC, left, and NEIC/USGS, right) and **M** (GCMT) according to the models fitted by Weatherill et al. (2016).

Figure 4.3 shows the distribution of earthquake magnitudes in time for the WPG16v3b and WPG16v3c catalogues. As can be observed, the magnitude of completeness stays relatively stable at around **M**4.0 from the late 1990s onward in WPG16v3c. The significant change in the density between the two plots for magnitudes smaller than **M**4.0 is due to the change in the conversion models.

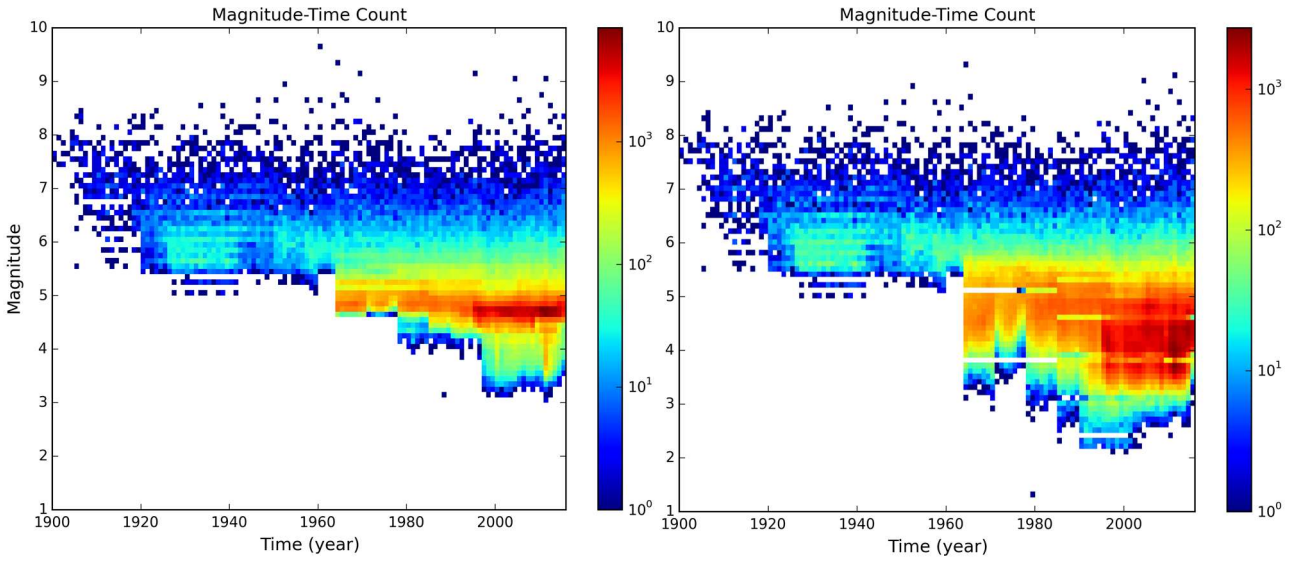

Figure 4.3. Density of earthquake magnitudes (colour scales) with time for the WPG16v3b (left) and WPG16v3c (right) catalogues.

4.2.2 Derived Changes

As opposed to v3b, WPG16v3c differentiates between GCMT estimates retrieved from the ISC Bulletin and those retrieved directly from the GCMT catalogue by labelling the latter GCMTo ("o" for "original"). This allowed to refine the code used to compare the events from the ISC Bulletin against those from the WPG16v3c catalogue, as the event IDs of those cases retrieved from the GCMTo could not be directly compared against the event IDs of the ISC Bulletin but those labelled as GCMT could. This increased the accuracy of the comparison.

Detailed observation of the WPG16v3c catalogue revealed the existence of several pairs of potential duplicates related to the comparison between the ISC Bulletin and the GCMT original catalogue carried out during its compilation. As these represented a very specific case of duplication of events, it was treated separately from the duplicate finder procedure applied over the merged catalogue (see Section 2.7.7 of the report for Version 1). The WPG16v3c was thus first purged of these GCMTo-related duplicate events before using it for comparison against the ISC Bulletin in the present work. The complete WPG16v3c catalogue contains 631,601 entries for the period 1 January 1900 – 31 December 2016, which are reduced to 631,386 when the 215 entries without depth information are withdrawn (see Appendix I of Version 1). Within these 631,386 entries, 511 were fully retrieved from the original GCMT catalogue and were thus assigned "GCMTo" as the agency. The word "fully" is being used herein to indicate that both the origin and magnitude estimates were those of the GCMT catalogue, as opposed to cases in which the magnitude was retrieved from the GCMT catalogue but the origin was not. Of these 511 events, 66 had potential duplicates, defined as events occurring within 60 seconds and 100 km of the GCMTo-identified ones, and with a maximum difference in magnitude of 0.35. The first two limits–60 seconds and 100 km–were selected to match those used for the comparison of the ISC Bulletin against WPG16v3c, which were as well used for the compilation of WPG16v3c itself, while the maximum difference in magnitude was the same as that used in the duplicate

finder (Section 2.7.7 of Version 1). Of these 66 pairs of potential duplicate events, 60 corresponded to cases in which the non-GCMTo event had a direct estimate of moment magnitude as well, while the remaining six had been assigned a proxy **M**. The aforementioned 60 events were thus purged, while the remaining six were left to be dealt with by means of the duplicate finder, as the comparison between direct and proxy values of **M** entails larger uncertainties. After this purge, 631,326 events are left, of which 379,348 correspond to the period 1 January 1998 – 31 December 2016 (end of the WPG16 catalogue), and 317,653 correspond to the period 1 January 2001 – 31 December 2015, that is, the final time window of interest in the present work.

A further small additional advantage of the WPG16v3c catalogue over WPG16v3b is that the former contains a field that indicates when an event has been flagged as induced using the keywords “geothermal", "reservoir", "mining", “anthropogenic" and “rockburst” that can be found in the comments of the ISC Bulletin for each earthquake entry.

4.3 Updated Version of the Database of Damaging Small-to-Medium Magnitude Earthquakes

The Database of Damaging Small-to-Medium Magnitude Earthquakes (Nievas *et al.*, 2019) has significantly increased in volume between Versions 1 and 2 of this work, particularly due to the addition of earthquakes listed in the Earthquake Impact Database (EID). The present database consists of 1,960 earthquakes with magnitudes in the range **M**4.0-5.5 that occurred from the year 1900 through 2017 for which reports of damage and/or casualties have been found.

While details on the compilation and the whole of the database can be found in Nievas *et al.* (2019), this section provides a brief overview of the 996 earthquakes that correspond to the period of interest for the present study (1st January 2001 – 31st December 2015), whose magnitude distribution is depicted in Figure 4.4. Of these, 90 and 70 earthquakes are reported to have caused damage to infrastructure and landslides, respectively (magnitude distribution in Figure 4.5). Only four out of the 996 earthquakes reportedly caused liquefaction, all of them with magnitudes above **M**5.0.

Figure 4.6 depicts the distribution of these 996 reportedly damaging earthquakes in time. As can be observed, the number of identified earthquakes increases sharply in 2013, the first year for which the Earthquake Impact Database (EID) was compiled. As discussed in Nievas *et al.* (2019), the EID differs significantly from other data sources in the fact that information on damage and/or casualties is collected online in almost-real time, thanks to a network of collaborators, allowing for a larger number of events to be registered. The impact of the EID on the Database of Damaging Small-to-Medium Magnitude Earthquakes illustrates the role that data accessibility plays in the perception of the impact of non-catastrophic hazards. The extent to which it dominates the database from 2013 onward serves as an indication of the completeness of the Database of Damaging Small-to-Medium Magnitude Earthquakes, as it is unlikely that the increase observed in Figure 4.6 be due to a real increase in the number

of damaging earthquakes. It is noted that the label “From EID” is used herein to refer to those entries of the EID that were processed and incorporated to the Database of Damaging Small-to-Medium Magnitude Earthquakes by means of the semi-automatic process described in Nievas *et al.* (2019), and does not imply that those labelled “Other” cannot be found in the EID as well.

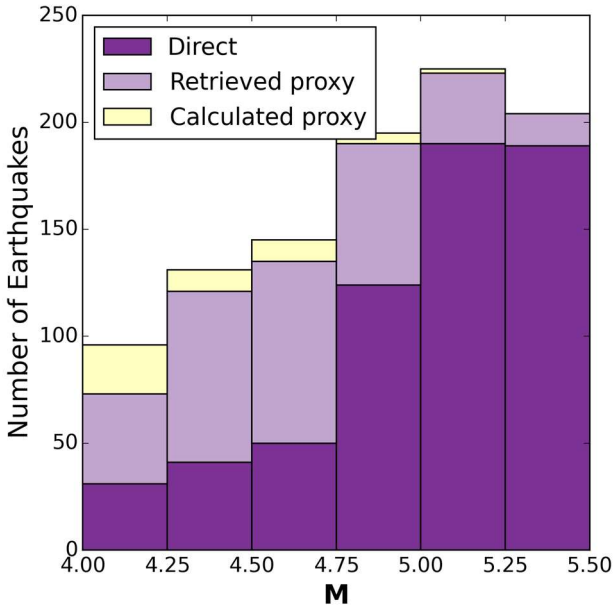

Figure 4.4. Distribution of moment magnitudes for the 996 earthquakes of the Database of Damaging Small-to-Medium Magnitude Earthquakes in the period of interest (1st January 2001 – 31st December 2015). Direct = **M** directly calculated from inversion (retrieved from seismological agencies, scientific literature, etc.). Retrieved proxy = proxy **M** retrieved from existent catalogues. Calculated proxy = proxy **M** calculated from other magnitude scales by Nievas et al. (2019).

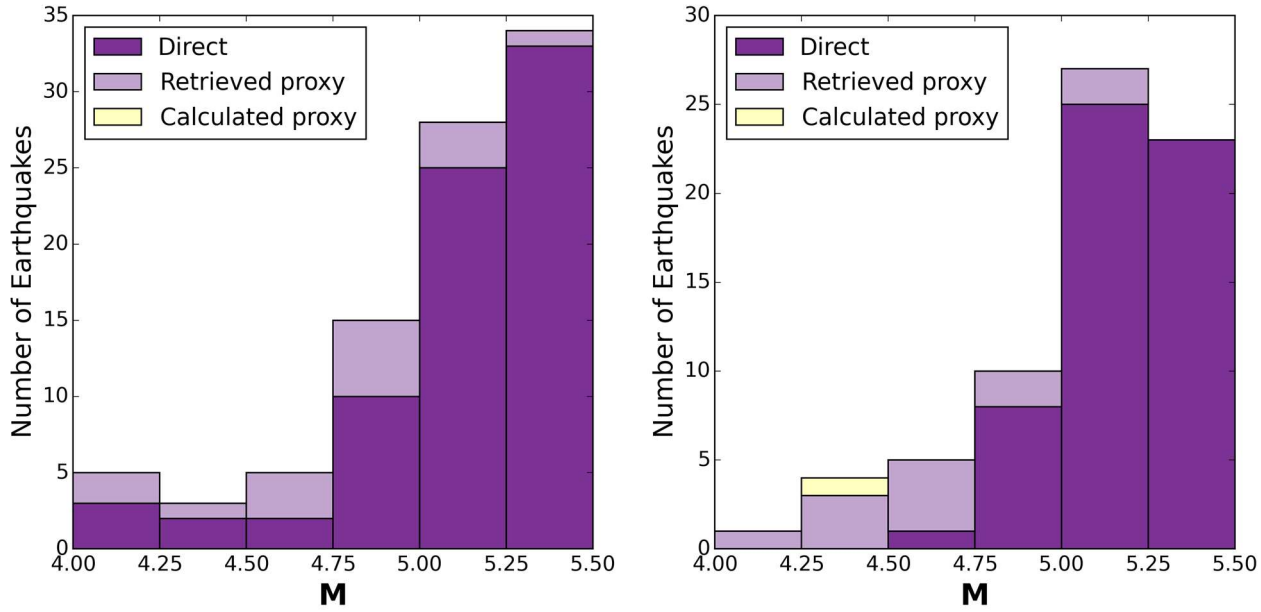

Figure 4.5. Distribution of magnitudes of the 90 earthquakes reported to have affected infrastructure (left) and the 70 reported to have caused landslides (right).

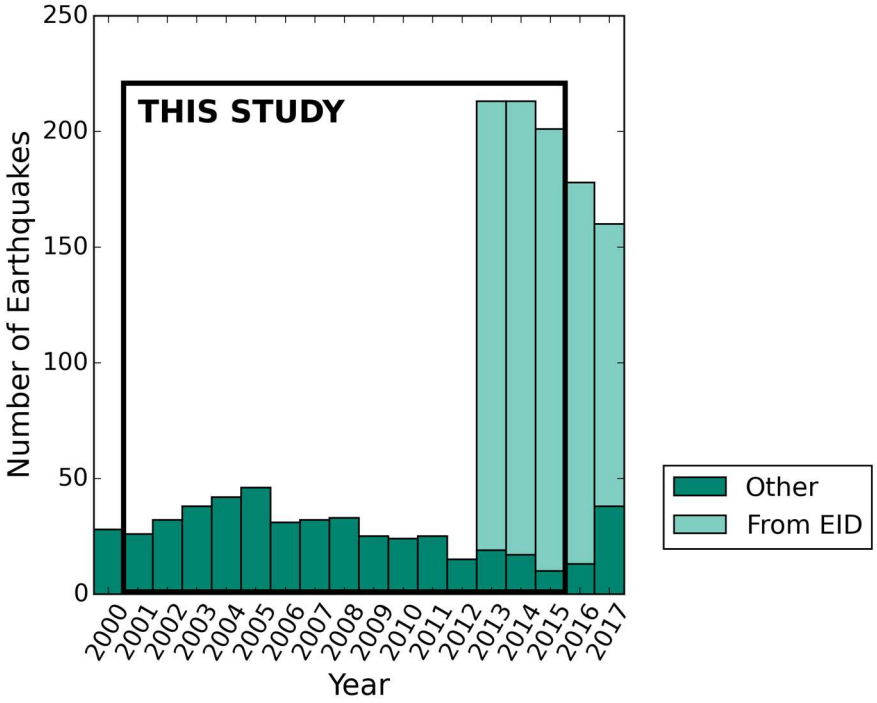

Figure 4.6. Distribution of dates of the earthquakes that make up the Database of Damaging Small-to-Medium Magnitude Earthquakes (Nievas *et al.*, 2019) for the years 2000-2017, with the rectangle enclosing the period of interest for the present study. “From EID” and “Other” refer to entries processed in a semi-automatic fashion from the Earthquake Impact Database and all other entries, respectively.

The drop in the number of earthquakes within the **M**5.25-5.50 bin with respect to the **M**5.00-5.25 one is likely due to a combination of, firstly, a natural tendency for the number of earthquakes to decrease as the Gutenberg-Richter law appears to take over the tendency of larger-magnitude earthquakes to have consequences in terms of damage and/or casualties (the same is observed for the whole of the Database of Damaging Small-to-Medium Magnitude Earthquakes), and, secondly, the greater proportional contribution of earthquakes reported in the EID for the 2001-2015 period than for the 1900-2017 period, with their magnitude distribution being relatively more uniform for all magnitude bins than that of all other cases.

As summarised in Table 4.1, these 996 earthquakes comprise a total of 362-407 deaths (excluding 18 heart attacks, as discussed in Nievas *et al.*, 2019) and 17,821-21,068 injuries. Numbers of casualties and damaged buildings are given in ranges (*i.e.*, lower and upper bounds) because sometimes different values are reported by different sources and also because, in some cases, the actual numbers are not known and only estimated ranges are available. Of the 362-407 deaths, 101-106 are due to landslides, and these are associated with 18 of the 70 earthquakes reported to have caused landslides, rockslides, mudslides and/or snow slides. As depicted in Figure 4.7, landslides are named as the cause for over one quarter of the total number of deaths, representing the second most-common cause after structural failures. While 23.6-26.0% of the deaths have been found to be directly attributed to structural failures in the literature and diverse consulted sources, a further 21.3-25.1% have been marked by the authors as “probably structural failures” based on the

reports on damage and the extent of the physical consequences observed, while an additional 15.0-16.0% have been found to be indicated as due to structural failures in the PAGER-CAT database (Allen *et al.*, 2009), based on similar criteria. These proportions change slightly when considering instances instead of number of deaths, as shown in Figure 4.8, understanding instances as the number of times a cause is named (*i.e.*, if all deaths associated to an earthquake are attributed to one cause, that is one instance of the cause). Within the 94 earthquakes associated with deaths (see Table 4.2), only 11-12 name more than one cause of death.

Table 4.1. Summary of casualties observed for the whole database during the time period of interest (1st January 2001 – 31st December 2015), excluding heart attacks attributed to earthquakes.

Casualties	Without EID		From EID		All	
	Lower	Upper	Lower	Upper	Lower	Upper
Total Deaths	341	386	21	21	362	407
Injuries	16,896	19,190	925	1,878	17,821	21,068

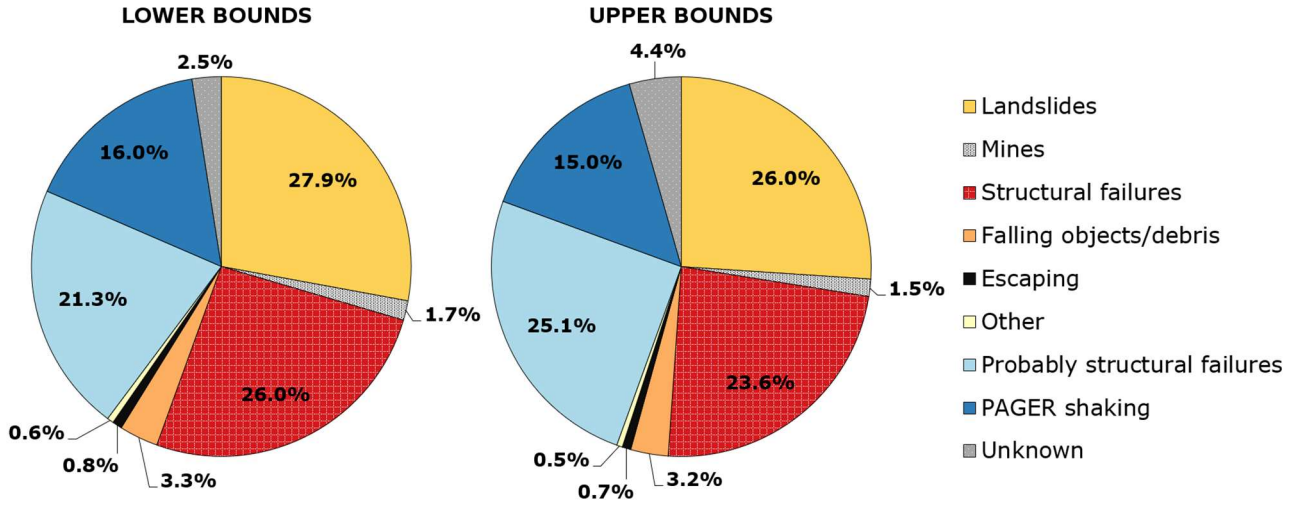

Figure 4.7. Causes of death for all total deaths (362, lower bounds, left; 407, upper bounds, right) for the earthquakes of the Database of Damaging Small-to-Medium Magnitude Earthquakes in the period of interest (1st January 2001 – 31st December 2015), excluding deaths attributed to heart attacks, in terms of proportions of the number of deaths.

Despite the total numbers of deaths and injuries presented in Table 4.1, it is noted that the most common situation is for earthquakes in the database to have caused no casualties at all, as shown in Table 4.2. Earthquakes causing neither deaths nor injuries account for 65.0% of the total 996, while those causing some injuries but no deaths amount to 25.6%. Deaths *and* injuries are associated to just 7.5% of all these earthquakes, while having some deaths but no injuries is the most uncommon outcome and represents only 1.9% of the cases. Apart from the deaths shown in Tables 4.1 and 4.2, there are twelve earthquakes for which the only deaths are attributed to heart attacks.

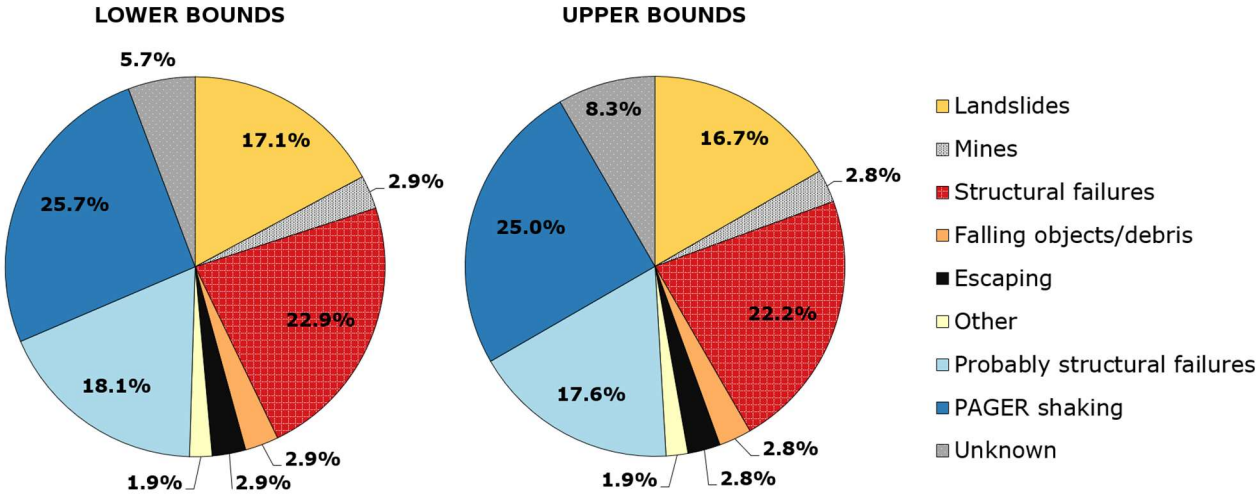

Figure 4.8. Causes of death for all total deaths for the earthquakes of the Database of Damaging Small-to-Medium Magnitude Earthquakes in the period of interest (1st January 2001 – 31st December 2015), excluding deaths attributed to heart attacks, in terms of proportions of number of instances (105 instances, lower bounds, left; 108 instances, upper bounds, right).

Table 4.2. Classification of earthquakes according to the existence or not of associated casualties for the whole database during the time period of interest (1st January 2001 – 31st December 2015), excluding heart attacks attributed to earthquakes.

Casualties	Number of Earthquakes						Percentage (%) of Earthquakes					
	W/o EID		From EID		All		W/o EID		From EID		All	
	Low	Up	Low	Up	Low	Up	Low	Up	Low	Up	Low	Up
No deaths & no injuries	181	181	466	466	647	647	43.6	43.6	80.2	80.2	65.0	65.0
No deaths & some injuries	154	154	101	101	255	255	37.1	37.1	17.4	17.4	25.6	25.6
Some deaths & no injuries	12	12	7	7	19	19	2.9	2.9	1.2	1.2	1.9	1.9
Some deaths & some injuries	68	68	7	7	75	75	16.4	16.4	1.2	1.2	7.5	7.5

Figures 4.9 and 4.10 depict the distribution of these injuries and deaths in time and according to magnitude. The tendency for the number of injuries and deaths to increase with magnitude is clear in both cases, as it is clear from Figure 4.11 that the number of earthquakes associated with either injuries or deaths increases with magnitude too. The drop observed for the **M**5.25-5.50 bin reflects the magnitude distribution of the 996 earthquakes in the period of interest (see Figure 4.4 and corresponding discussion). This drop does not occur when considering the whole of the database (1900-2017), though the relative increase from the **M**5.00-5.25 to the **M**5.25-5.50 bin is smaller than from **M**4.75-5.00 to **M**5.00-5.25 in that case. For a discussion regarding some cases of earthquakes with seemingly unusual number of injuries/deaths for their magnitude, please refer to Nievas *et al.* (2019).

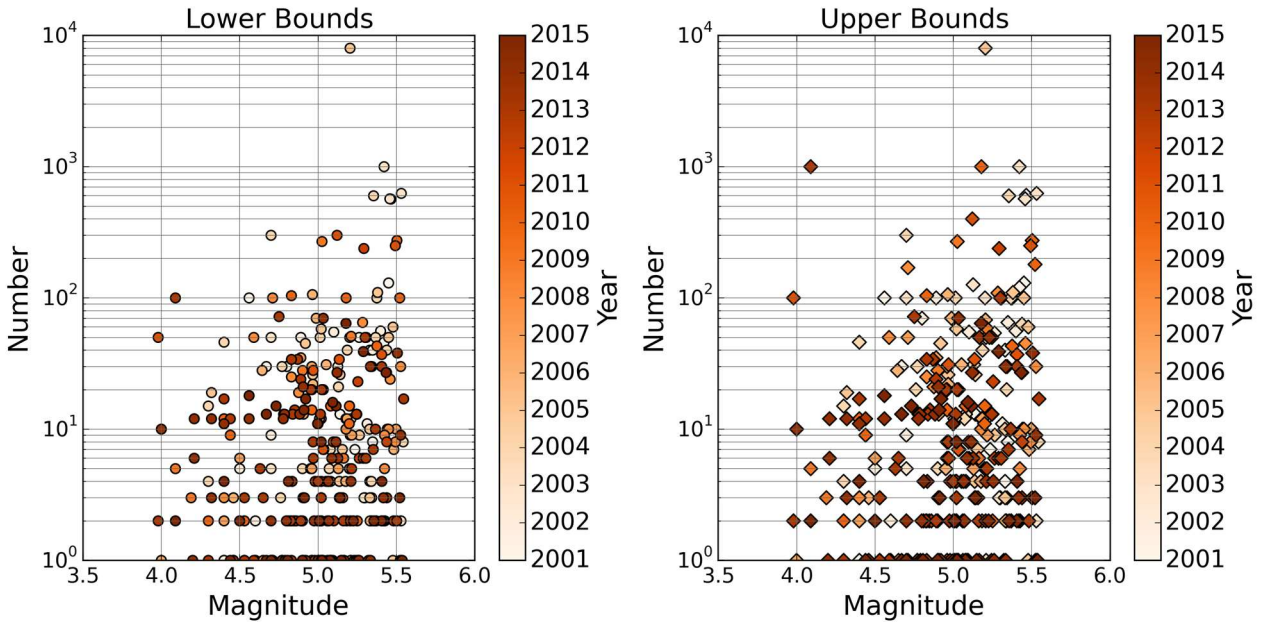

Figure 4.9. Number of injuries as a function of moment magnitude for the earthquakes of the Database of Damaging Small-to-Medium Magnitude Earthquakes in the period of interest (1st January 2001 – 31st December 2015). Colour scale indicates year of occurrence. Zero values not shown.

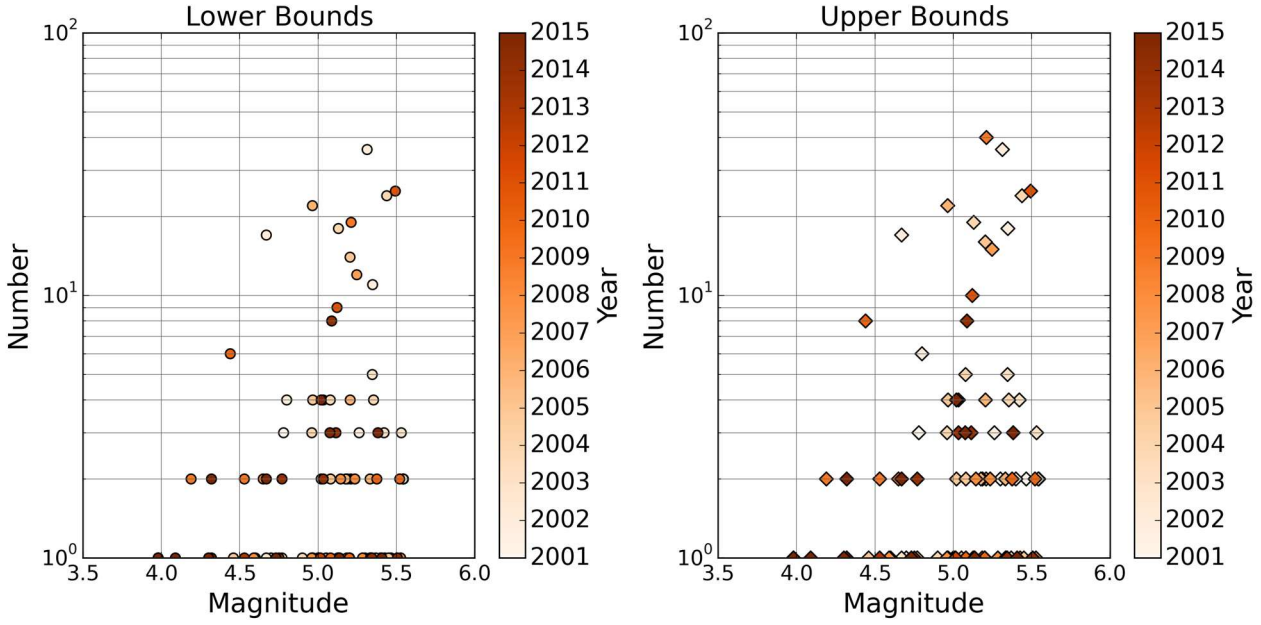

Figure 4.10. Number of deaths as a function of moment magnitude for the earthquakes of the Database of Damaging Small-to-Medium Magnitude Earthquakes in the period of interest (1st January 2001 – 31st December 2015). Colour scale indicates year of occurrence. Zero values not shown.

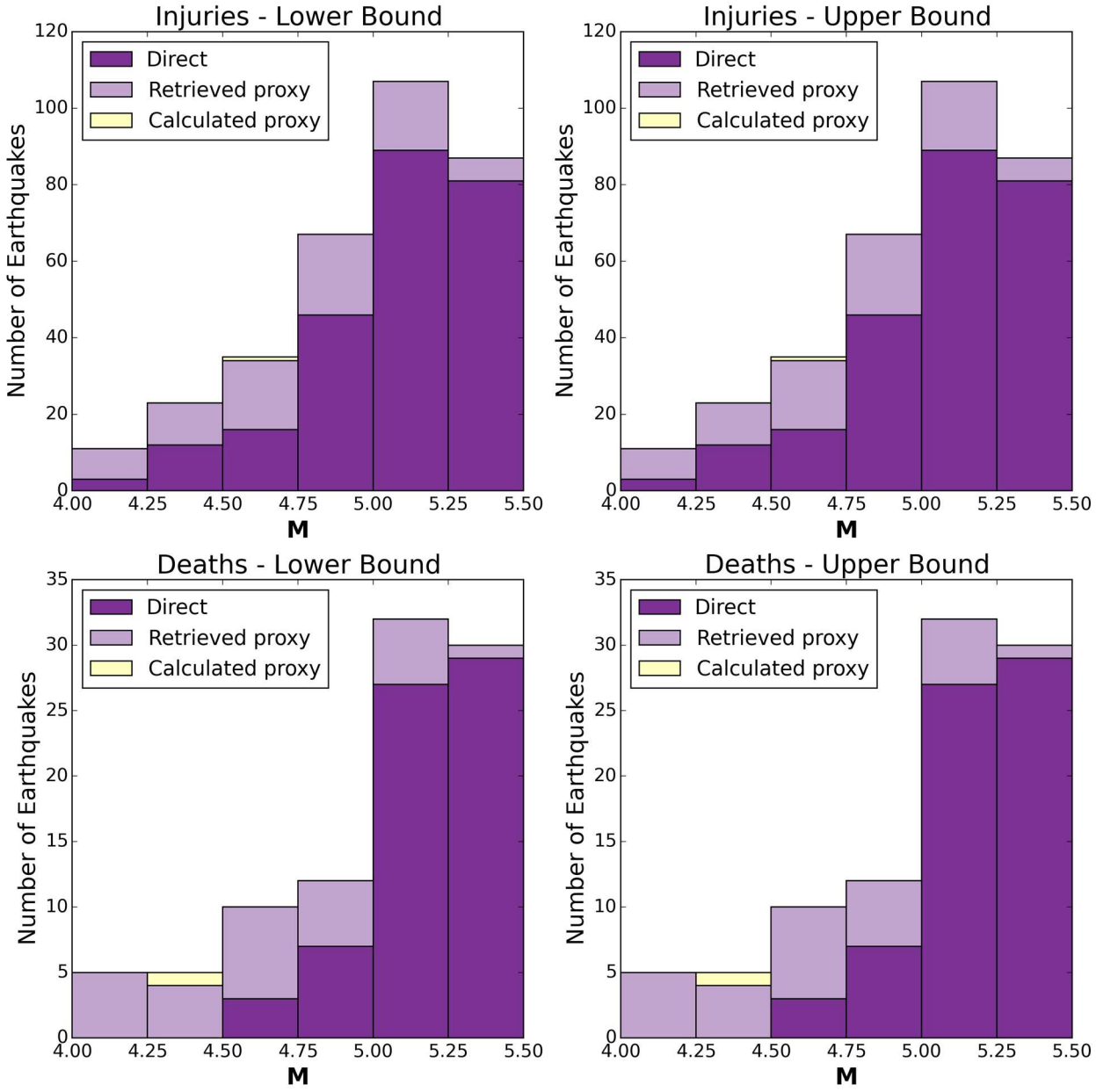

Figure 4.11. Magnitude distribution of the earthquakes of the Database of Damaging Small-to-Medium Magnitude Earthquakes in the period of interest (1st January 2001 – 31st December 2015) with associated injuries (top) and total deaths excluding heart attacks (bottom), considering lower- (left) and upper-bound (right) figures. Direct, Retrieved proxy, Calculated proxy as per description of Figure 4.4.

As explained in Nievas *et al.* (2019), the fields in the Database of Damaging Small-to-Medium Earthquakes concerning the number of damaged and destroyed buildings do not always contain specific numbers but many times hold verbal descriptions and ranges. Figure 4.12 shows the proportion of each kind of data for the 996 earthquakes of interest for the present study: “Number” (a specific number), “EID estimation” (a number specified by the EID to have been estimated and not necessarily observed), “NOAA range” (a range of values estimated by NOAA based on verbal descriptions of damage according to the criteria specified in their website), “Other range” (a range of numbers stemming from different sources reporting different values, “Text” (a verbal description such as “limited” or “some, severe”), “Some of” (only in the case of destroyed buildings, it implies that “some of” the

buildings reported as damaged were actually destroyed), and “Not available”. According to the experience gained during the compilation of the Database of Damaging Small-to-Medium Earthquakes, the authors believe that a non-available number of destroyed buildings when a number of damaged buildings is available usually indicates that none of the damage instances observed corresponded to destruction. However, this may not always be the case, and there are certainly cases in which there is no data for either damaged or destroyed buildings but a monetary value of loss is reported, clearly implying that damage did happen but the numbers of affected buildings is currently unknown.

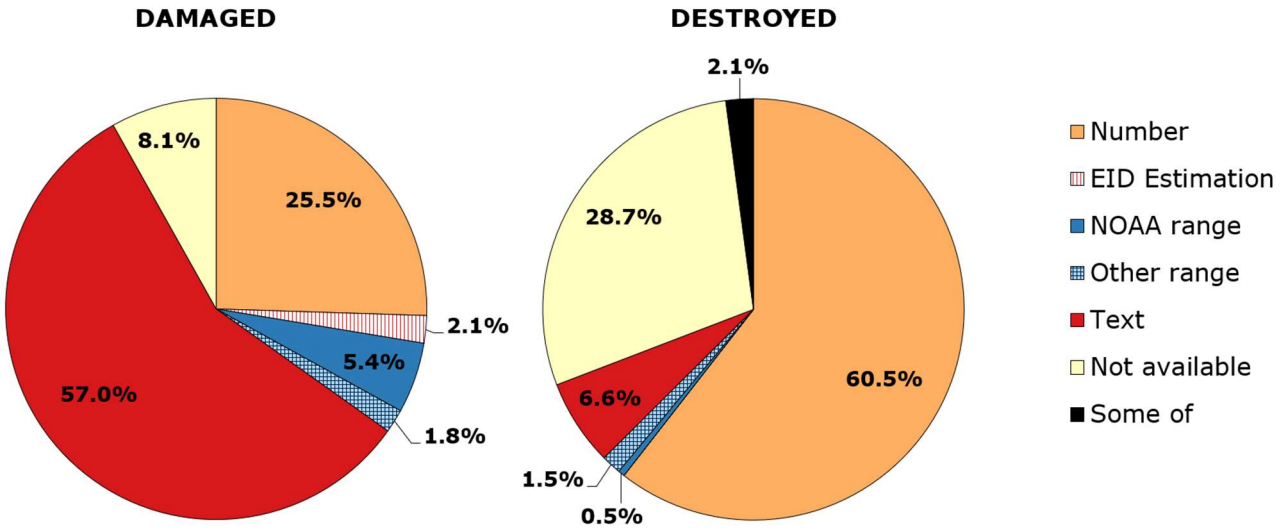

Figure 4.12. Kind of information available on damaged (left) and destroyed (right) buildings for the 996 earthquakes of the Database of Damaging Small-to-Medium Magnitude Earthquakes in the period of interest (1st January 2001 – 31st December 2015).

Figures 4.13 and 4.14 depict the verbal descriptions available for 586 and 66 earthquakes with damaged and destroyed buildings, respectively. Of greater interest are, of course, Figures 4.15 and 4.16, which show the number of damaged and destroyed buildings as a function of magnitude and time of occurrence for the cases in which a number or range of numbers are available (“Number”, “EID estimation”, “NOAA range”, “Other range”, and “Some of” cases). The tendency for the numbers of damaged and destroyed buildings to increase with magnitude, similar to that observed for casualties, is clear here as well. For more information regarding some cases of earthquakes with seemingly unusual number of damaged/destroyed buildings for their magnitude, please refer to Nievas *et al.* (2019).

The 347 and 644 earthquakes for which data on damaged and destroyed buildings (respectively) was available as “Number”, “EID estimation”, “NOAA range”, “Other range”, and “Some of”, are reportedly associated with 796,579-932,191 damaged and 77,829-112,186 destroyed buildings.

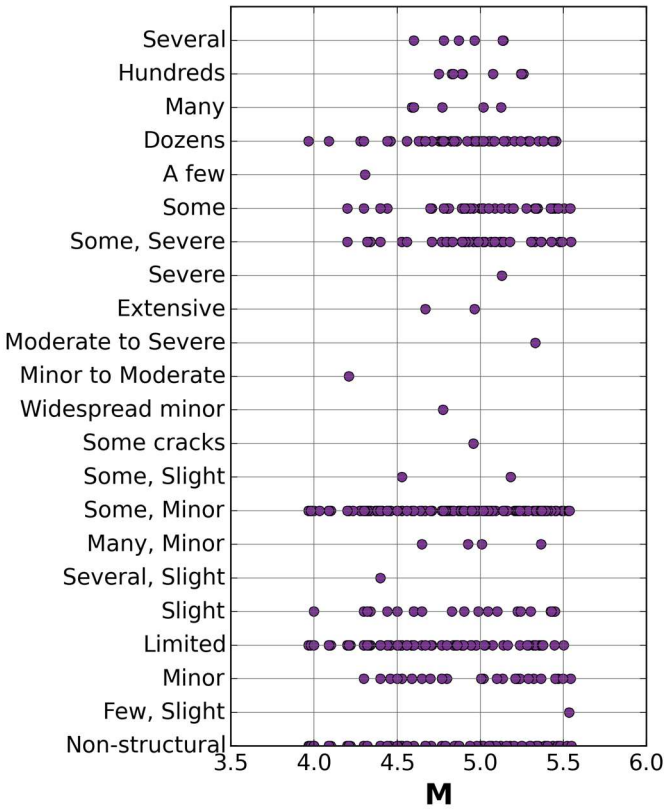

Figure 4.13. Descriptions of damage of damaged buildings as a function of magnitude for the 586 out of the 996 earthquakes for which only verbal descriptions were found.

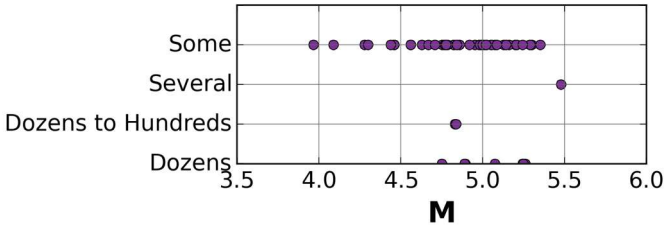

Figure 4.14. Descriptions of damage of destroyed buildings as a function of magnitude for the 66 out of the 996 earthquakes for which only verbal descriptions were found.

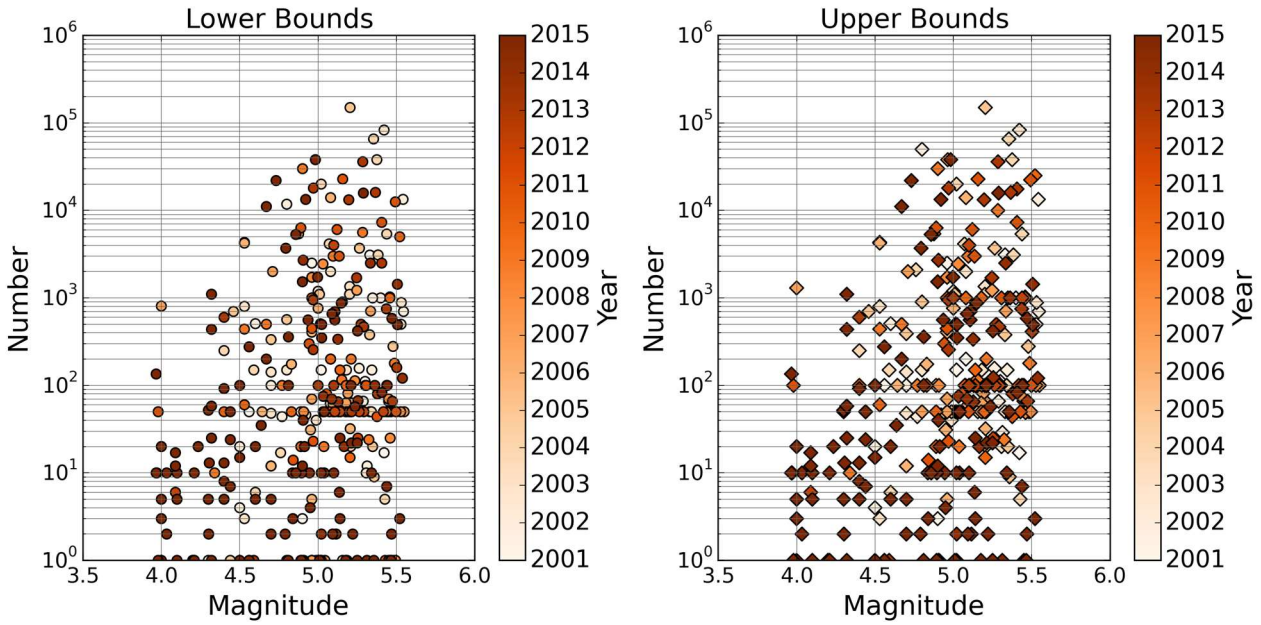

Figure 4.15. Number of damaged buildings as a function of moment magnitude for the earthquakes of the Database of Damaging Small-to-Medium Magnitude Earthquakes in the period of interest (1st January 2001 – 31st December 2015) for which numbers of damaged buildings are available. Colour scale indicates year of occurrence. Zero values not shown.

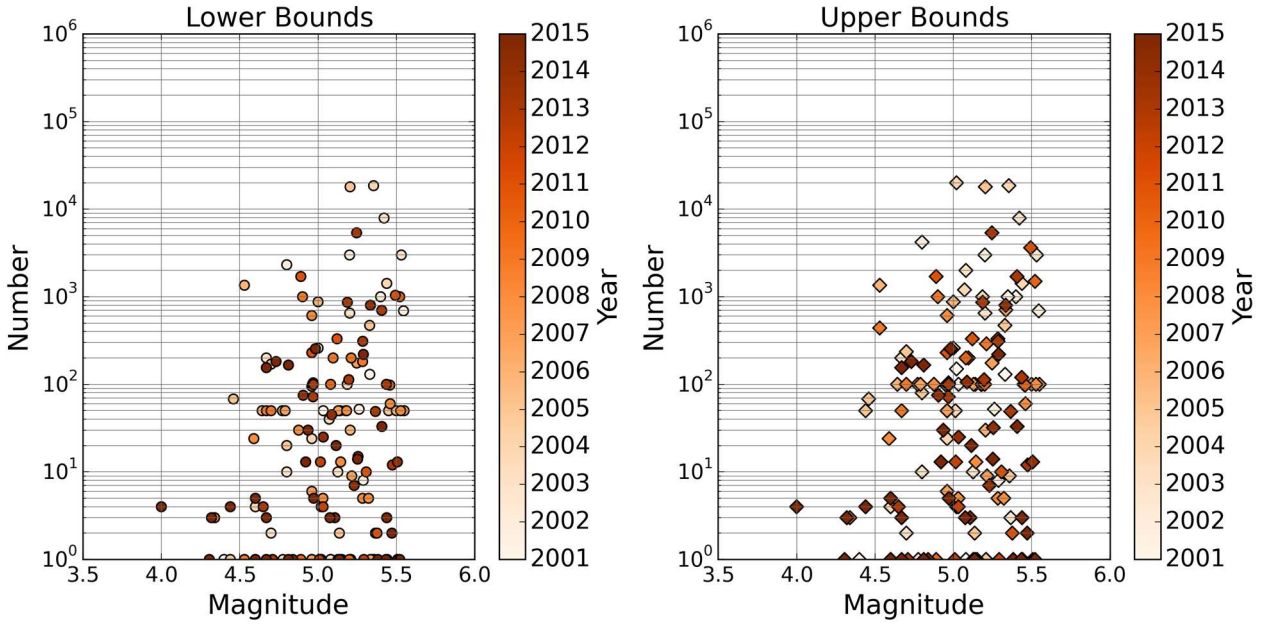

Figure 4.16. Number of destroyed buildings as a function of moment magnitude for the earthquakes of the Database of Damaging Small-to-Medium Magnitude Earthquakes in the period of interest (1st January 2001 – 31st December 2015) for which numbers of destroyed buildings are available. Colour scale indicates year of occurrence. Zero values not shown.

4.4 Treatment of Magnitude

Section 2.5 of the report for Version 1 discusses the relation between moment magnitude **M** and other magnitude scales, such as surface-wave magnitude M_s, body-wave magnitude m_b, duration magnitude M_d, and Richter local magnitude M_L. In the compilation of Version 1 of the world database only moment magnitude **M**, surface-wave magnitude M_s and local magnitude M_L were considered when retrieving information from the ISC Bulletin, and a relation of direct equivalence (**M**=M_s=M_L) was assumed among the three for the sake of keeping things simple and accessible during the first exploratory version of this work. A refinement of this criterion was made for Version 2, both to incorporate other two magnitude scales of common use and to convert all scales into moment magnitude **M**. Moreover, while in Version 1 the magnitude limits were set to 4.00≤**M**≤5.50, the criterion 3.95≤**M**<5.55 was applied in Version 2 to account for the fact that magnitudes are often reported to just one decimal place (*i.e.*, 3.95 would be rounded up to 4.0 and values approaching 5.55 would be rounded down to 5.5 in such reports).

4.4.1 Magnitude Scales and Conversion Models

The decision to consider only **M**, M_s and M_L resulted in events for which only magnitude values in other scales were available being discarded in Version 1. Among these, there were 20,598 instances of duration magnitude M_d and 24,975 instances of body-wave magnitude m_b in the range 4.0-5.5, each instance being a magnitude estimate and not an individual event (*i.e.*, one event can have, for example, one instance of M_d and another instance of m_b). Under the conservative assumption that each instance corresponds to a separate event, these numbers represent 1.85% and 2.26% of the total of earthquakes in Version 1 of the catalogue, respectively. Both duration magnitude M_d and body-wave magnitude m_b were considered herein for Version 2.

Aiming at the homogenisation of the different magnitude scales considered into moment magnitude **M**, a look was taken at three valuable sets of empirical conversion models available in the literature: those of Scordilis (2006), Di Giacomo *et al.* (2015) and Weatherill *et al.* (2016). The models of Scordilis (2006) were derived using moment magnitude **M** from the Global Centroid Moment Tensor Project (GCMT) and the United States Geological Survey / National Earthquake Information Center (USGS/NEIC), all the data points being treated as one set, for shallow earthquakes (h≤70 km). The models are the following:

- **M** from M_s:

  3.0≤M_s≤6.1: $$\boldsymbol{M} = 0.67 \cdot M_s + 2.07,\ \sigma = 0.17 \quad (4.1)$$

  6.2≤M_s≤8.2: $$\boldsymbol{M} = 0.99 \cdot M_s + 0.08,\ \sigma = 0.20 \quad (4.2)$$

- **M** from m_b:

  3.5≤m_b≤6.2: $$\boldsymbol{M} = 0.85 \cdot m_b + 1.03,\ \sigma = 0.29 \quad (4.3)$$

Di Giacomo *et al*. (2015) derived two kinds of models for each magnitude scale, one using exponential non-linear regression and the other using linear Generalised Orthogonal Regression (GOR), all obtained using only moment magnitude **M** from the GCMT and m_b and M_s as recomputed for the ISC-GEM catalogue (Storchak *et al.*, 2015):

- **M** from M_s, exponential model:

$$\boldsymbol{M} = e^{(-0.222 \quad .233 \cdot M_S)} + 2.863 \tag{4.4}$$

- **M** from M_s, GOR model:

$M_s \leq 6.47$: $$\boldsymbol{M} = 0.67 \cdot M_s + 2.13 \tag{4.5}$$

$M_s > 6.47$: $$\boldsymbol{M} = 1.10 \cdot M_s - 0.67 \tag{4.6}$$

- **M** from m_b, exponential model:

$$\boldsymbol{M} = e^{(-4.664 + 0.859 \cdot m_b)} + 4.555 \tag{4.7}$$

- **M** from m_b, GOR model:

$$\boldsymbol{M} = 1.38 \cdot m_b - 1.79 \tag{4.8}$$

Weatherill *et al.* (2016) derived different models for m_b and M_s calculated by the International Seismological Centre (ISC) or the USGS/NEIC, in all cases in relation with moment magnitude **M** from the GCMT. The resulting models are the following:

- **M** from M_s determined by the ISC:

$M_s \leq 6.00$: $$\boldsymbol{M} = 0.616 \cdot M_s^{ISC} + 2.369,\ \sigma = 0.147 \tag{4.9}$$

$M_s > 6.00$: $$\boldsymbol{M} = 0.994 \cdot M_s^{ISC} + 0.1,\ \sigma = 0.174 \tag{4.10}$$

- **M** from M_s determined by the USGS/NEIC:

$M_s \leq 6.47$: $$\boldsymbol{M} = 0.723 \cdot M_s^{NEIC} + 1.798,\ \sigma = 0.159 \tag{4.11}$$

$M_s > 6.47$: $$\boldsymbol{M} = 1.005 \cdot M_s^{NEIC} - 0.026,\ \sigma = 0.187 \tag{4.12}$$

- **M** from m_b determined by the ISC:

$$\boldsymbol{M} = 1.084 \cdot m_b^{ISC} - 0.142,\ \sigma = 0.317 \tag{4.13}$$

- **M** from m_b determined by the USGS/NEIC:

$$\boldsymbol{M} = 1.159 \cdot m_b^{NEIC} - 0.659, \sigma = 0.283 \quad (4.14)$$

Figure 4.17 shows the behaviour of these models in a broad magnitude range. Solid colour lines depict the range for which the model was derived, while dashed colour lines indicate extrapolation of the models to lower or higher magnitudes. The range of applicability is explicitly stated only by Scordilis (2006) and was instead inferred from the data plots of Di Giacomo *et al*. (2015) and Weatherill *et al.* (2016). All models were derived with very scarce data below the magnitude 4.0 level. The consequences of this are apparent in the case of the exponential models of Di Giacomo *et al*. (2015), being particularly extreme in the case of m_b, for which the model can never reach values of **M** below 4.555. This means that even for m_b tending to minus infinity, **M** still is 4.555, which is clearly non-physical. In the case of M_s, the model tends to **M**2.863 when M_s tends to minus infinity. Due to this extreme behaviour below the magnitude range for which they were derived, the exponential models of Di Giacomo *et al*. (2015) were not considered any further for use within this work.

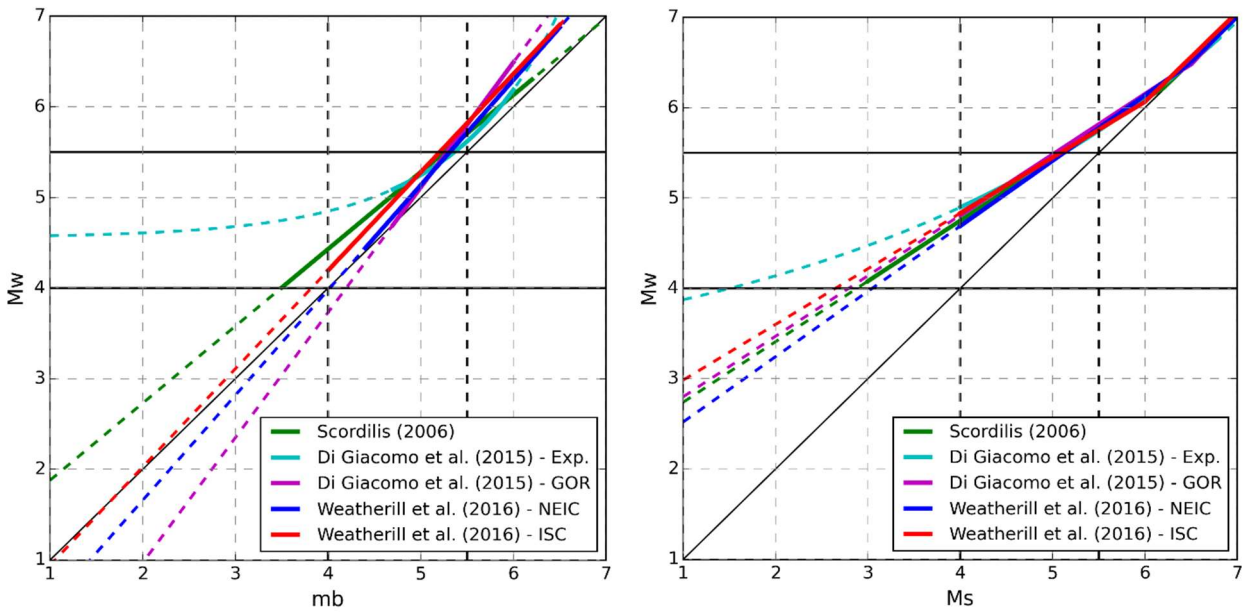

Figure 4.17. Relation between m_b and **M** (left) and M_s and **M** (right) according to the different models considered herein.

An average of the four models was finally used for conversion from m_b and M_s into **M**. The resulting average models are depicted in Figure 4.18. In this case, the applicability range has been taken as the most restrictive bounds of the four individual models. The standard deviation of the average model was taken as the maximum of those of the individual models. In the case of the model for m_b, this corresponds to 0.317, stemming from the model of Weatherill *et al*. (2016), and for values in the range 0.17-0.20 (stemming from the model of Scordilis, 2006) for the model for M_s, which is piece-wise linear. As Di Giacomo *et al*. (2015) do not report the standard deviations of their models, only those of the other models were considered.

Strictly speaking, the standard deviation of the average models should be calculated as that of the linear combination of a series of normal distributions, as per Equation 4.15. However, this yields a much smaller value than that of any of the individual models. While this might be correct from the point of view of theoretical rigour, it lacks meaning in the present context, as taking an average model results from the uncertainty of not being able to determine which model applies best to data from various agencies from all over the world, if any, which cannot be reflected in a smaller standard deviation. Moreover, these average models were needed the most for earthquakes lying far below the values of m_b and M_s for which they were derived, and thus applied in a range for which not much data is available. A standard deviation smaller than that of any of the individual models would certainly not reflect this.

$$\sigma_{av} = \sqrt{\sum_{i=1}^{k} \left(\frac{1}{n}\right)^2 \cdot \sigma_i^2} \tag{4.15}$$

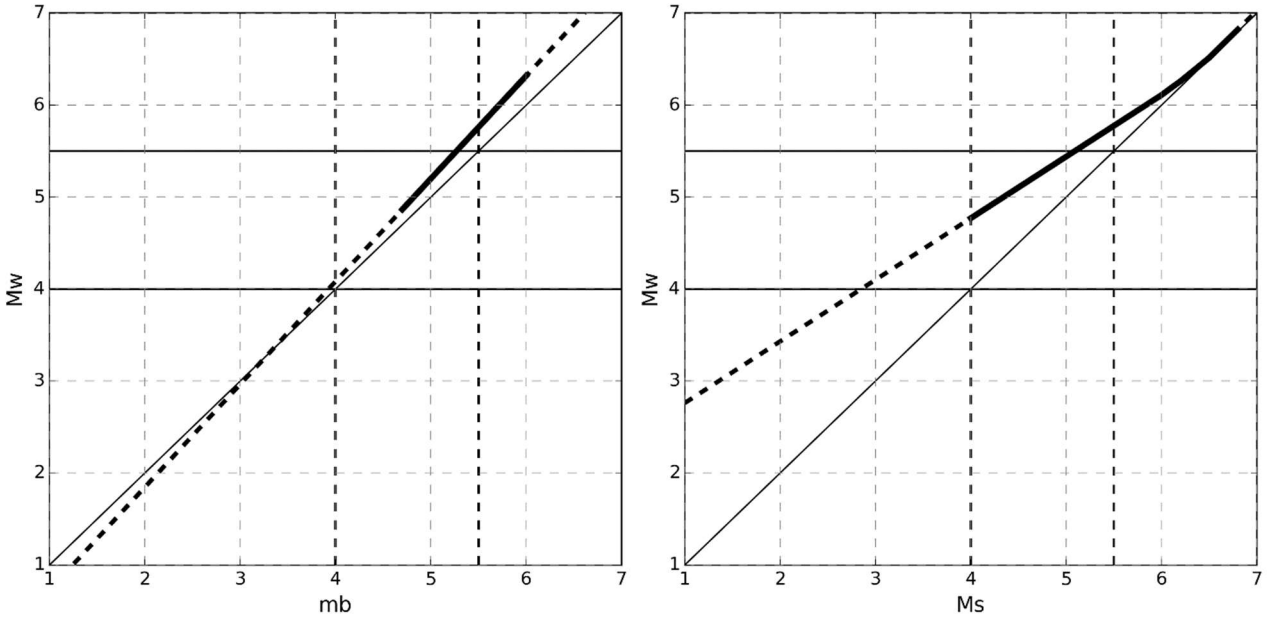

Figure 4.18. Average relation between m_b and **M** (left) and M_s and **M** (right) according to the four different models finally considered.

From their detailed assessment of models available to convert from local magnitude M_L into moment magnitude **M**, Dost *et al.* (2018) concluded that the average of the studies conform with **M** ≈ M_L for values of M_L above 2.0-3.0, while **M** lies systematically above M_L for M_L<2.0. Given that this threshold is significantly smaller than the magnitudes of interest for the present study, and taking into account the fact that the relation between **M** and M_L varies across different regions (*e.g.*, Scordilis, 2006), the **M**=M_L equivalence was adopted herein, as had been done already for Version 1. A standard deviation of 0.25 was applied to these cases, as this appears as a reasonable value when contrasted against existing models to convert from M_L to **M** (*e.g.*, Goertz-Allman *et al.*, 2011; Grünthal *et al.*, 2009; Papazachos *et al.*, 1997), and in relation with those reported above for models based on M_s and m_b.

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. As already

discussed in the report for Version 1, M_d appears to underestimate M_L at large magnitudes and overestimate it at low magnitudes, at least for values of the two calculated by agencies like the Italian National Institute of Geophysics and Volanology (INGV) and the National Observatory of Athens (NOA, Greece). However, the tendency does not seem to be so pronounced for those of the Kandili Observatory and Earthquake Research Institute (KOERI, Turkey) or the Disaster and Emergency Management of the Presidency (AFAD, Turkey) (Weatherill 2017, *pers. comm.*). The poor scaling with other magnitude scales and the high level of noise that can be observed in the data suggest that a simple assumption of M_d=M_L=**M** is the most reasonable, as more refined analysis do not guarantee any real improvement over this assumption. A standard deviation of 0.3 was adopted.

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 WPG16v3c catalogue, as the latter already provides moment magnitude and standard deviation for each earthquake it contains. The rationale for not reprocessing all of the events from the WPG16v3c catalogue was that, given their different nature with respect to those added herein, the additional computational effort would not yield an improved homogenisation of the set. Weatherill *et al.* (2016) only converted values of M_s and m_b calculated either by the ISC or the USGS/NEIC, and no other, for example. In each case, they applied conversion equations calibrated for each of the two authoring agencies, as described earlier. As earthquakes already in WPG16v3c are not added from the ISC Bulletin (unless an update is needed, see Section 2.7.4 of the report for Version 1; Nievas *et al.*, 2017), any estimates of M_s and m_b that were converted using the average models depicted in Figure 4.18 were calculated by a different agency, whose scaling with respect to moment magnitude could be slightly different from those of the ISC or the USGS/NEIC. In other words, the use of one conversion model for all agencies implies overseeing the potential differences across agencies. Given that the volume of earthquakes in the final catalogue of potentially damaging earthquakes that is taken directly from WPG16v3c is around 50%, it was deemed better to keep the agency-specific conversions for such a large number of events.

4.4.2 Expected Magnitude and Associated Uncertainty

As uncertainty in magnitude is explicitly taken into consideration in Version 2, each event ends up being described not only by a moment magnitude estimate, obtained either directly from an agency or by means of conversion models, but also by a standard deviation around that estimate. Whenever the value of **M** is taken directly from that reported by an agency in the ISC Bulletin, the standard deviation is the measurement error read directly from the Bulletin as well. However, if a conversion model *f* is applied to obtain **M**, then standard error propagation must be carried out to obtain the final standard deviation from the combination of both the measurement error $\sigma_{measure}$ and the uncertainty of the conversion model $\sigma_{conversion}$ as per Equation 4.16, where $M_{original}$ is the original magnitude scale (*e.g.*, Weatherill *et al.*, 2016).

$$\sigma_{\boldsymbol{M}} = \sqrt{\sigma^2_{conversion} + \left(\frac{\partial f}{\partial M_{original}}\right)^2 \cdot \sigma^2_{measure}} \tag{4.16}$$

In order to understand the usual order of magnitude of the measurement error, $\sigma_{measure}$, all measurement errors reported in the ISC Bulletin for the period 1st January 2009 – 30th November 2015 were scrutinised. The mean and median resulted in 0.168 and 0.100, respectively. Of the 1,887,188 values analysed, 99.1% of the values were equal to or smaller than 0.5, 94.9% of the values were equal to or smaller than 0.3, 1,061 cases were equal to or larger than 1.0, while 220 had values larger than 9.0. From the documentation of the ISC Bulletin it is not possible to know what such values mean, though the relatively large number of values equal to 9.9 suggests this might be a code used by some agencies to indicate that the value was not available. Considering only measurement errors below 1.0, the mean and median were 0.166 and 0.100, respectively, and 95.0% of the values were equal to or smaller than 0.3. In view of these values, and considering the fact that the lack of availability of the measurement error contributes to an increased uncertainty, a value of 0.3 was adopted when the measurement error was not reported or was reported as null in the ISC Bulletin. In the case of moment magnitudes, a lower value of 0.1 was adopted instead, as this reflects a typical measurement error found in the Global Moment Tensor Database (*e.g.*, Weatherill et al., 2016). Magnitude estimates with reported measurement error larger than 1.0 were not accepted as valid estimates and were rejected at the time of verifying the hierarchy of agencies and magnitude scales. Assignment of default values of measurement errors was carried out as well for the earthquakes taken directly from the WPG16v3c catalogue that had direct values of moment magnitude with no associated error.

4.4.3 Summary of the Way in Which Magnitude Was Treated

The most important points to bear in mind regarding the treatment of magnitude are the following:

- One value of expected magnitude and associated measurement error was selected for each earthquake, according to the hierarchy of agencies (authors) described in the report for Version 1.
- Moment magnitudes of earthquakes retrieved from the WPG16v3c catalogue were taken directly as reported in it. Default values of measurement errors were assigned when moment magnitudes had been directly available and WPG16v3c reported no error.
- For earthquakes added from the ISC Bulletin, only magnitudes in terms of moment magnitude **M**, surface-wave magnitude M_s, body-wave magnitude m_b, local magnitude M_L, and duration magnitude M_d were considered.
- Surface-wave magnitudes M_s and body-wave magnitudes m_b were converted into moment magnitudes using an average model stemming from four models available in the literature.
- Local magnitudes M_L and duration magnitudes M_d were converted into moment magnitudes assuming $\mathbf{M}=M_L=M_d$.
- Final uncertainties were calculated by means of combining both measurement errors and the standard deviations of conversion models.
- Expected moment magnitudes were used for determining whether the earthquakes were included or not in the catalogue.

- Earthquakes included in the catalogue comply with the criterion 3.95≤**M**<5.55.
- Expected (reported) magnitudes were used for declustering purposes.
- Expected (reported) magnitudes were used to define maximum depths as per Section 4.5.2 below.
- Full probability distributions of magnitude, represented by a normal distribution, were used for the calculation of expected values of Modified Mercalli Intensities.

4.5 Treatment of Depth

4.5.1 Retrieving Expected Depth and Associated Measurement Error

Similar to the case of magnitude, uncertainty in depth is explicitly taken into consideration in Version 2. Each earthquake in the catalogue was characterised by its expected depth and its associated standard deviation. For this purpose, the depth error reported in the ISC Bulletin (and in WPG16v3c) was interpreted as being the distance from the median reported depth to either side of the 90%-confidence interval resulting from the vertical projection of the four-dimensional location error ellipsoids, as depicted in Figure 4.19 (interpretation based on the description of the IASPEI Seismic Format available on the ISC website and on the work of Bondár & Storchak, 2011). 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 related to *Err* by means of Equation 4.17. In a strict physical sense, this distribution should be truncated at a depth of 0 km. However, as the documentation from the ISC suggests that the reported standard error is calculated in an unbounded space, 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 was used for the propagation of uncertainties, as will be explained in Section 4.6.2, to avoid inconsistency with the original estimate of the ISC. Figure 4.20 shows examples of how the truncated normal distribution compares to its parent unbounded one. The further away the distribution moves from the 0-km truncation limit, the more the two are alike.

$$\sigma_D = Err/1.64 \tag{4.17}$$

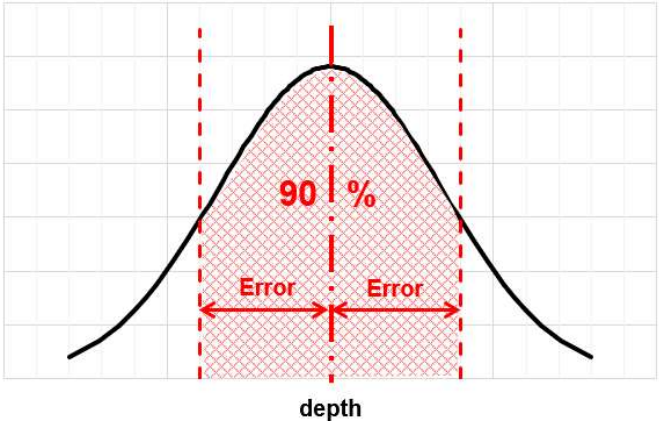

Figure 4.19. Interpretation of the depth error reported in the ISC Bulletin (marked as “Error”).

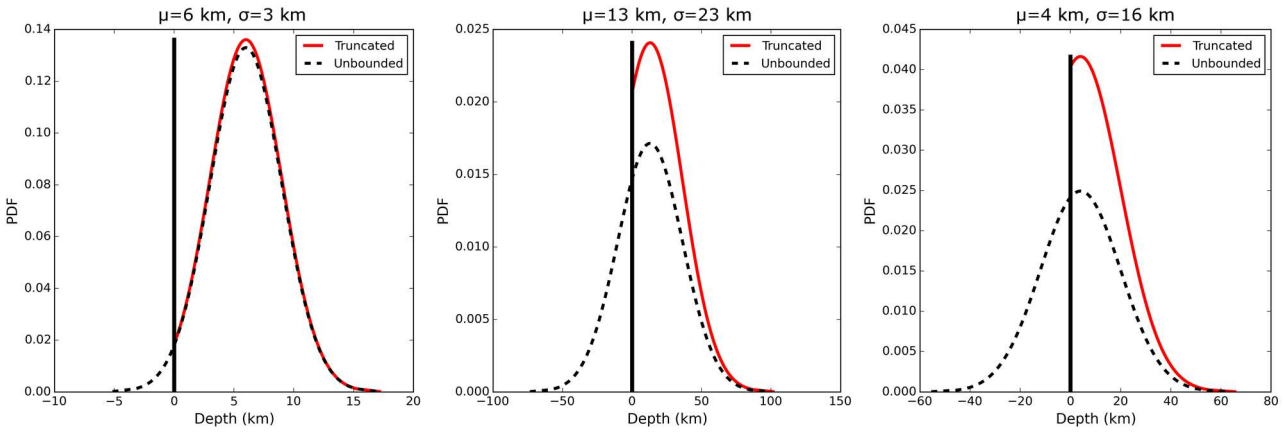

Figure 4.20. Examples of truncated vs. unbounded normal distributions for different values of location (μ) and scale (σ) of the unbounded distribution.

However, and as already explained in Section 3.2.2, it is not always possible to determine the hypocentral depth during the inversion process, and the procedures applied to calculate a free-depth solution and/or select a fix depth have changed in time. At the time of writing, the ISC is undergoing the process of re-building the whole Bulletin using the locator described by Bóndar & Storchak (2011) and Bóndar, I. (2015). At this point, four periods can be distinguished:

- 1964-1979: has been already re-built using the same procedures as those applied from 2009 till the present time.
- 1980-2005: old locator, Jeffreys-Bullen velocity model. Default depths could be 0, 10, 33, 50, 100, 150, 200, 250, etc. Depths reported by local agencies could be used as fixed depths well. Free-depth solution always attempted at first.
- 2006-2008: old locator, ak135 velocity model. Same default depths as during 1980-2005, but 33 km was changed to 35 km. Free-depth solution always attempted at first.
- 2009-present time: new locator. Free-depth solution only attempted if one of four criteria is met. Fixed default depths in order of preference are: a location-specific value derived from statistical data of past events of the ISC, the median of the reported depths if the initial depth is larger than 100 km, a conventional depth assigned to the corresponding Flinn-Engdahl geographic region.

It is noted that “present time” in the last bullet point makes reference to the last month of the Reviewed Bulletin, which is constantly updated.

While the ISC uses flags such as “f” or “d”, which indicate depths fixed to default regional values or calculated by means of depth-phase stacking, respectively, the most significant indication that a depth value has been fixed is the lack of a reported standard error. As a consequence, the lack of a standard error was used to identify fix-depth solutions both from the ISC Bulletin and the WPG16v3c catalogue. When this occurred, the event was assigned an error calculated as explained in the upcoming Section 4.5.2.

These same assumptions, which are based on the practices of the ISC for generating the Reviewed Bulletin, were also applied to their contributing agencies, even though it is possible that their practices be different. For example, it is possible that some agencies may not report a depth error even when they carried out a depth-free location search. However, it is impossible to manually review documentation that might (or not) exist for each agency, and there is no one document that gathers it all.

4.5.2 Maximum Depth Criteria

Two main changes were introduced in Version 2 regarding the maximum hypocentral depths to be considered as corresponding to upper-crustal events. The first one involves the maximum depths themselves, with the step function defined by Table 3.2 and depicted in Figure 4.21 replaced by a linear variation from 15 km at **M**4.0 to 35 km at **M**5.5, as shown in Figure 4.21 as well. The constraints and considerations taken into account to adopt this criterion were discussed in Section 3.2.

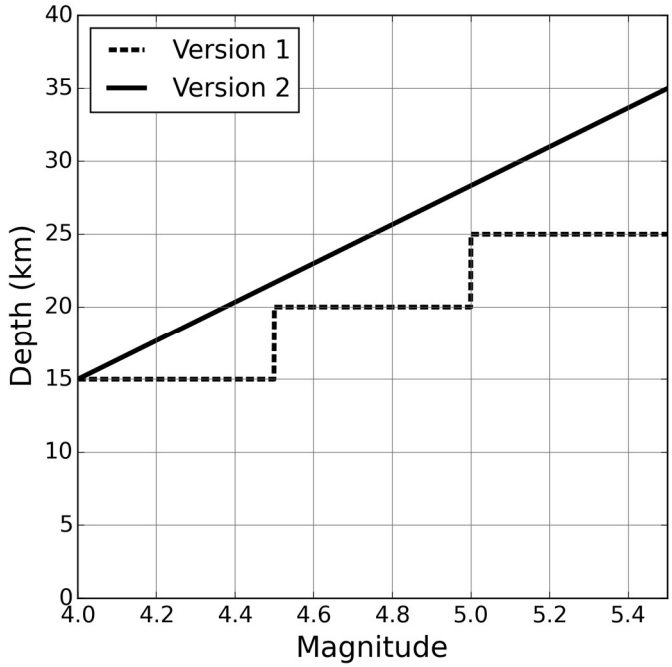

Figure 4.21. Maximum depth criteria adopted for Versions 1 and 2.

The second change is associated with the incorporation of the uncertainty in depth to the compilation process, the details of which are explained in Section 4.6.2. For the case of earthquakes with fixed-depth solutions, their expected (reported) depth was compared against the maximum depth limits corresponding to the expected magnitude[1]. If smaller, they were kept, if larger, they were discarded. Earthquakes with free-depth solutions, on the other hand, were kept if the probability of their depth being equal to or smaller than the

[1] This was done once one origin and magnitude estimate was selected for each earthquake, as described in Section 2.6 of the report for Version 1 (Nievas *et al.*, 2017), and after applying the corresponding magnitude conversion models, as discussed in Section 4.4 above.

maximum depth limit (corresponding to their expected magnitude) was equal to or larger than 50% and discarded otherwise.

This last criterion was established to deal with earthquakes for which unreasonably large errors are reported (*e.g.*, error of 200 km for an expected value of 4 km). Unreasonably large errors imply that the depth was, in fact, not constrained, and are reflected further down the process in unreasonably large standard deviations of Modified Mercalli Intensities (MMI). As shown in Figure 4.22, the probability of their depth being equal to or smaller than the maximum depth limit (colour scale) correlates well with the final uncertainty in MMI (except for the cases in which a 0-km depth with an extremely large error is reported, which does not inflate the standard deviation because of the 0-km depth but which clearly correspond to unconstrained cases; these are the blue dots with error/expected ratios of several hundreds or thousands, which were re-assigned a value of 0.1 km in order to be included in the analysis). The criterion was finally defined by plotting probability contours for different ratios of error/expected and expected/limit depth values, as shown in Figure 4.23. These probabilities are independent of the actual values of depth, error, and maximum-depth limit. An expected/limit ratio of 1.0 implies that the reported depth is exactly equal to the maximum depth limit. In this case, the probability can never reach 50%, but tends to it for very small values of the measurement error. When moving to the left from this point in the horizontal axis, the admissible error/expected ratios can get larger because the mean moves away from the limit towards a more comfortable area. This plot shows why other metrics such as the coefficient of variation alone (*i.e.* the vertical axis of Figure 4.23) are not sufficient.

In order to be consistent with this last analysis, earthquakes with fixed-depth solutions were assigned the maximum standard error that caused the probability of the depth being smaller than or equal to the maximum depth (for the corresponding expected magnitude) to be at least of 50%. In other words, they were assigned the ratio of error to expected depth that corresponded to the 0.5-contour of Figure 4.23 for their expected to maximum depth ratio. Numerically this was carried out by defining this contour for a relevant number of values of the ratio of expected to maximum depth and interpolating for in-between values. Depths fixed to 0 km were assigned an arbitrary value of 0.1 km, to make the calculation numerically feasible. The smallest ratio of expected to maximum depth is, thus, 0.1/35.0=0.00286, larger than the 0.002 minimum value for which the 0.5-contour was calculated. A minimum standard error to expected depth ratio of 0.00001 was assigned for a ratio of expected to maximum depth of 1.0; this translates into a standard error of 0.00015 for an earthquake with a fixed depth of 15 km and **M**4.0 and of 0.00035 for an earthquake with a fixed depth of 35 km and **M**5.5.

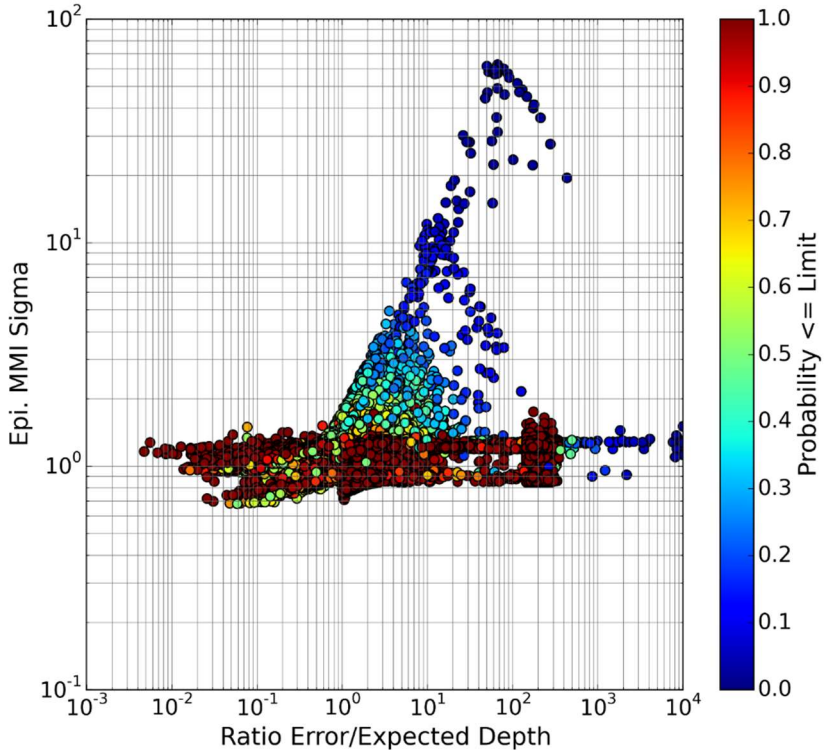

Figure 4.22. Standard deviation of MMI at the epicentre[2] versus ratio of depth error to reported expected depth. Colour scale indicates probability of depth being equal to or smaller than the maximum depth limit for the corresponding magnitude. This plot includes fixed-depth-solution earthquakes, which were assigned an error equal to the limit and a probability of 1.0 when their reported depth complies with the maximum depth limit. Preliminary (*i.e.*, non-final) results.

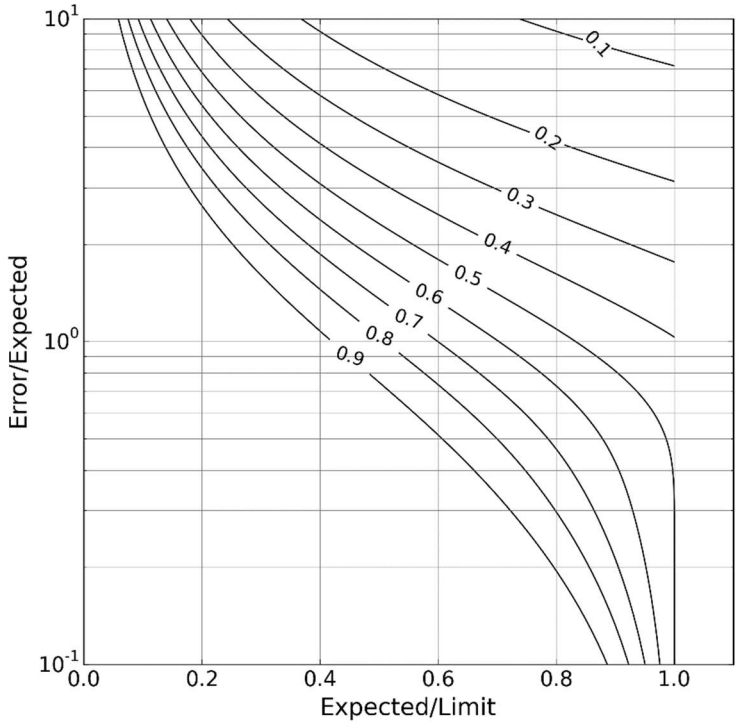

Figure 4.23. Probability of the depth being equal to or smaller than the maximum-depth limit as a function of the ratio of expected depth to depth limit (horizontal axis) and depth error to expected depth (vertical axis), for a truncated normal distribution (0-km depth lower bound).

[2] Calculated as per Section 4.6.2.

As a consequence of these criteria, earthquakes that do not pass the depth criteria are those with:

- fixed-depth solutions with expected depth larger than the limit, irrespective of sigma;
- free-depth solutions with expected depth larger than the limit, irrespective of sigma;
- free-depth solutions with depth smaller than the limit but sigma such that the probability of the depth being smaller than the limit is less than 50% (*i.e.*, depth solutions with “large” sigma).

The latter implies, for example, that a **M**4.7 earthquake with a fixed-depth solution of 10 km passes the criterion while an earthquake with a free-depth solution of 10 km and sigma of 90 km does not. While this could seem like a contradiction, given that a fixed-depth solution does not result from an actual calculation of the hypocentral depth, the depths adopted for fixed-depth solutions tend to be based on past observations of earthquakes in the epicentral area (at least since 2009 for those calculated by the ISC) or general knowledge on usual depth distributions, and can thus be deemed as a potentially superior estimate than a poorly-constrained free-depth solution.

Of the 327,271 events that make up the merged catalogue filtered by magnitude, 141,520 (43.2%) pass the depth criteria and 185,751 (56.8%) don’t. Of these 185,751 events that do not, 100,582 (54.1%) correspond to free-depth solutions. Of these, 1,972 events would pass the maximum depth limit if the uncertainty was not taken into account, that is, 1,972 have expected depths that comply with the maximum depths defined in Figure 4.21 but less than 50% probability of their depths being smaller than those limits when considering their uncertainty as well. These events are shown on the map of Figure 4.24. As they only represent 1.4% (=1,972/141,520) of the set that does pass the depth criteria and many of them are located in uninhabited areas of the planet and would not pass the exposure criterion, the influence of including these events or not is minor.

4.5.3 Summary of the Way in Which Depth Was Treated

The most important points to bear in mind regarding the treatment of depth are the following:

- One value of expected depth and associated measurement error was selected for each earthquake, according to the hierarchy of agencies (authors) described in the report for Version 1.
- Inclusion or not of the earthquake in the catalogue was determined based on a magnitude-dependent maximum-depth criterion, as illustrated by the solid black line in Figure 4.21. Expected magnitudes (without uncertainty) were used for this purpose.
- Earthquakes with free-depth solutions were included if the probability of their depth being equal to or smaller than that defined in Figure 4.21 was at least of 50%.
- Earthquakes with fixed-depth solutions were included if their reported fix depth was equal to or smaller than that defined in Figure 4.21. The maximum standard deviation that causes the probability of the depth being equal to or smaller than that defined in Figure 4.21 to be at least of 50% was assigned to these earthquakes.

- Expected (reported) depths were used for declustering purposes.
- Full probability distributions of depth, represented by a truncated normal distribution, were used for the calculation of expected values of Modified Mercalli Intensities.

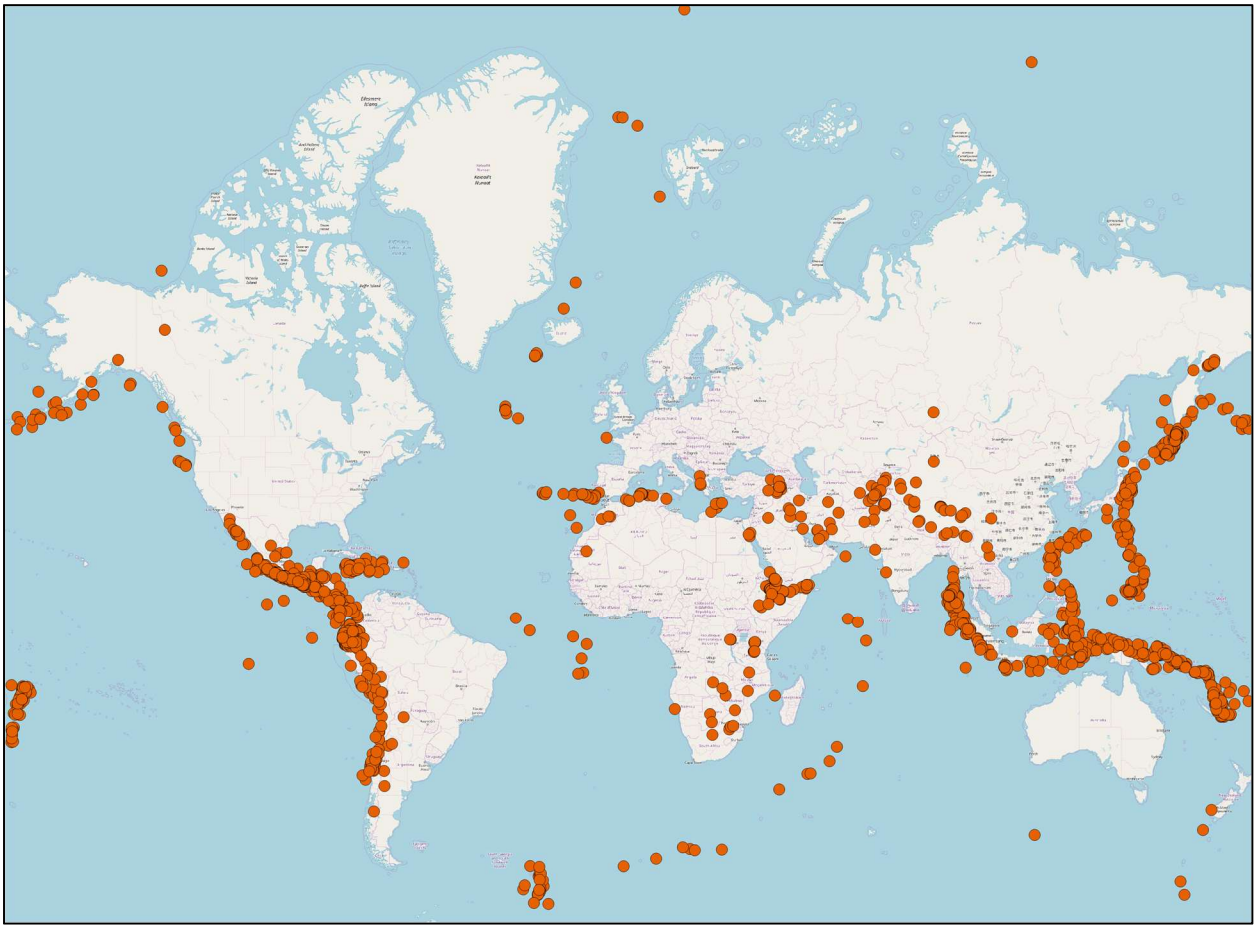

Figure 4.24. Earthquakes with free-depth solutions whose expected depths would comply with the limits defined in Figure 4.21 but have less than 50% probability of complying with these limits if the uncertainty in depth is considered.

4.6 Proximity to Inhabited Areas

4.6.1 Overview

As briefly described in the Introduction (Chapter 1) and Section 2.1, the proximity of the earthquakes to exposed assets or individuals was evaluated in terms of the number and density of population estimated to have been exposed to a certain level of seismic intensity. The former was obtained from Gridded Population of the World (GPW) v4.0 (CIESIN, 2016), while the latter was calculated by means of Intensity Prediction Equations (IPEs). The same approach had been adopted in Version 1, but a series of modifications have been introduced herein:

- change of IPEs used;
- use of a craton index to take into account different degrees of attenuation in different parts of the world;

- explicit consideration of uncertainty in depth, magnitude and the IPE; and
- change in the decision of whether to count a cell or not, derived from the consideration of uncertainty.

4.6.2 Intensity Prediction Equations and the Craton Index

In Version 1 it was discussed that five earthquakes identified as having caused damage and/or casualties did not comply with the exposure criterion. These were:

- 5th August 2005 18:07:12 UTC, **M**5.0, Ho Chi Minh, Vietnam (C97): There is not much information about the consequences of this earthquake, except for the ISC Bulletin reporting a statement by the USGS saying that minor damage occurred at Ho Chi Minh city, and NOAA assigning an economic loss of less than 1 million USD. The epicentre was located over 90 km away from the cost, and the hypocentral depth is reported as 10 km. With these parameters, the IPE of Allen *et al.* (2012) yields an epicentral intensity of VI, and a distance to the MMI III isoseismal of 88 km. As a consequence, the estimation of population exposure was zero in Version 1.

- 8th November 2005 07:54:37 UTC, **M**5.3, Ho Chi Minh, Vietnam (M169): There is not much information about the consequences of this earthquake either, except for the ISC Bulletin, the USGS catalogue and Ngo *et al.* (2008) reporting one death with no specified cause. Phuong & Truyen (2014) describe this earthquake as the main shock of a series of which the event of 5th August 2005 was a foreshock, but do not make any reference to damage or casualties caused by the series. With epicentral coordinates very close to the previous event, and slightly larger magnitude and shallower depth, the IPE of Allen *et al.* (2012) yields an epicentral intensity of VII, and 125 km to the MMI III isoseismal, which does not seem to be enough for any population to fall within predicted MMI values of IV and larger.

- 20th April 2010 00:17:10 UTC, **M**4.5, Kalgoorlie-Boulder, Australia (M55): Information about the damage caused by this earthquake is more detailed than for the previous two cases. Edwards *et al.* (2010) describe damage observed in around 60 buildings and a few minor injuries. The earthquake is believed to be connected to mining activities in the area. It is clear that this earthquake is of relevance to the database, but it gets filtered out based on low population exposure. The epicentral MMI predicted by the model of Allen *et al.* (2012) is slightly above V, and the MMI IV isoseismal (of Version 1) passes right by the side of the towns of Kalgoorlie and Boulder, where the damage was observed.

- 6th September 2010 22:48:34 UTC, **M**5.0, Porangahau, New Zealand (C51): Information regarding the damage caused by this earthquake appears to be contradictory. It has an estimate of economic losses in NOAA, and the USGS catalogue indicates the occurrence of slight damage. However, online newspapers suggest that there were no reports of damage (Otago Daily Times, 2010; Radio New Zealand, 2010). It is noted, though, that it is possible that damage reports may not

have made the news, but may have been collected by the scientific community. Unlike the previous cases, the MMI IV isoseismal predicted by Allen *et al.* (2012) encloses population, but not enough to satisfy the exposure criterion.

- 26th November 2012 05:33:49 UTC, **M**5.1, Ruoqiang, China (C68): According to NOAA, this earthquake caused damage to around 50 to 100 buildings, amounting to monetary losses of less than 1 million USD. Information in online news sites does not seem to be abundant (at least in English), but the George Herald (2012) mentions homes having been damaged. It is clear that the area is severely underpopulated.

Inconsistencies of this kind can be caused by many reasons, including:

- lack of accuracy in depth reports (*i.e.*, believing the hypocentre was deeper than it was),
- lack of accuracy in epicentral coordinates (*i.e.*, the event was actually closer to population centres than was modelled),
- lack of accuracy in magnitude estimations (*i.e.*, the event was actually bigger than modelled),
- the arbitrary nature of the exposure criterion (*e.g.*, a cumulative population count of 2,490 people and a maximum density of 290 people/km^2 do not pass the thresholds of 2,500 people and 300 people/km^2 set in Version 1 but are almost the same in practice),
- the differences in attenuation properties of different parts of the crust, which are not taken into consideration in the Intensity Prediction Equation (IPE) of Allen *et al.* (2012).

Prompted by the case of the **M**4.5 2010 Kalgoorlie-Boulder (Australia) earthquake, for which damage was reported but which did not pass the exposure criterion in Version 1, the focus herein falls on the last point. South-western Australia is a well-known Stable Continental Region (SCR), whose attenuation behaviour is significantly different from those of Active Crustal Regions (ACR) for which the IPE of Allen *et al.* (2012) was derived. Stable Continental Regions present less attenuation than the latter, which translates into larger macroseismic intensities being observed at larger distances than would occur in an ACR. Consequently, the IPE of Allen *et al.* (2012) is likely to underestimate the macroseismic intensities generated at long epicentral distances by earthquakes in SCRs, and an IPE developed for SCRs would probably yield more realistic values. In this context, the IPE of Atkinson & Wald (2007) derived from data from the Central and Eastern United States (CEUS), one of the most prominent SCRs in the world, appears as an attractive alternative. However, it is known not only that attenuation does vary across different SCRs, but also that the CEUS presents one of the lowest attenuations of all SCRs (Bakun & McGarr, 2002). Within the whole range of SCRs, the attenuation level is closely related to how cratonic or not the site is, as cratonic regions are typified by old, hard rocks that are better at transmitting seismic waves than younger regions. 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, said authors have proposed a fuzzy approach to this

classification that ultimately led to what they call the craton index and is shown in Figure 4.25, mapped for the whole world at 0.5º intervals. The craton index (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.

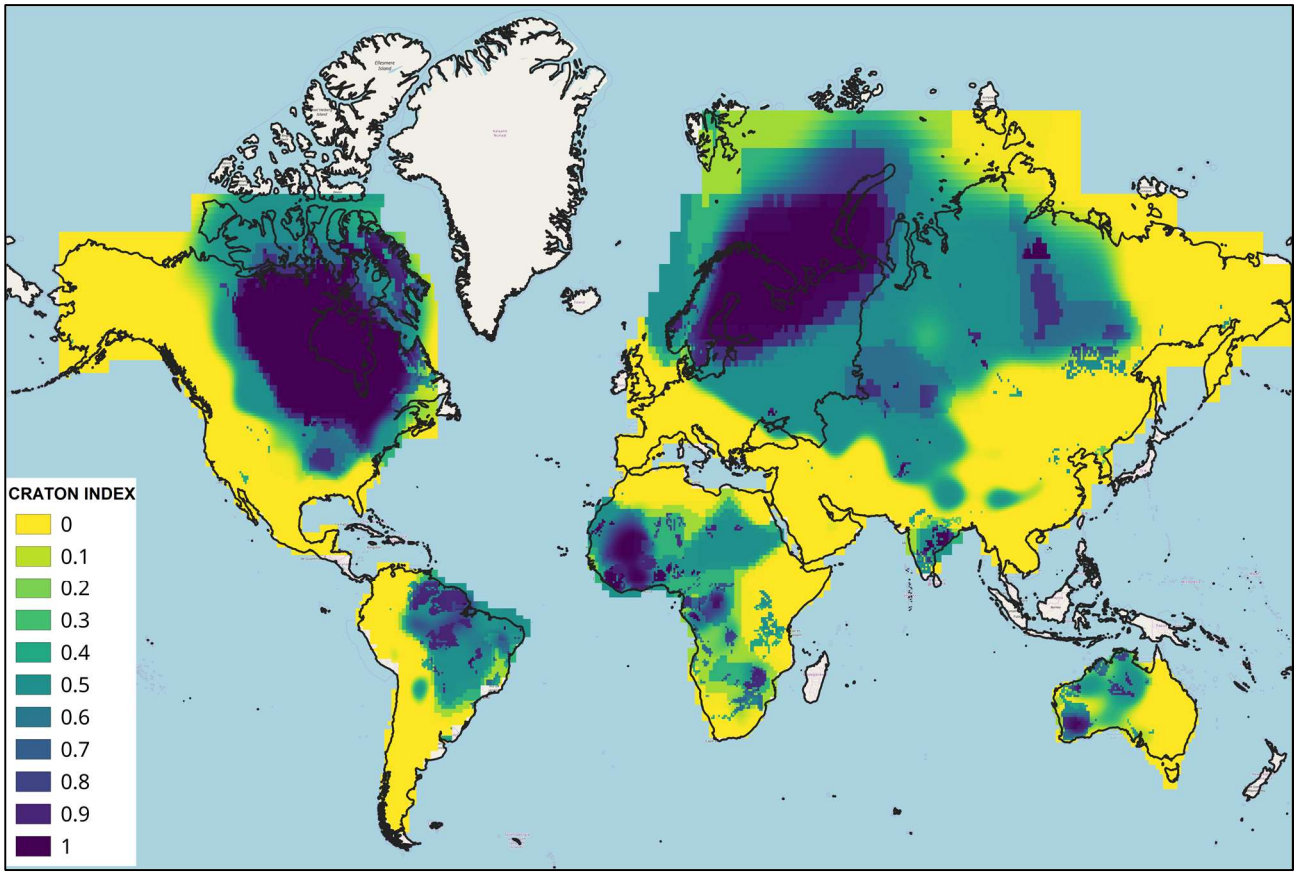

Figure 4.25. The craton index according to Chen *et al.* (2018): 0= non cratonic, 1=cratonic. Data courtesy of Dr. G. Weatherill.

Based on these considerations, the approach adopted in Version 2 for the determination of the population exposure consists in using the craton index (CI) to weigh the values of Modified Mercalli Intensity (MMI) calculated using an IPE for ACRs (MMI_{ACR}) and an IPE for SCRs (MMI_{SCR}), as shown in Equation 4.18. This weighting scheme assumes that MMI_{ACR} and MMI_{SCR} represent the two possible ends of the spectrum of macroseismic intensities to be expected, and that the larger the value of CI, the closest the value of MMI generated by an earthquake occurring at a particular location is to MMI_{SCR}, while the more CI tends to zero, the more the MMI tends to MMI_{ACR}.

$$MMI_{final} = CI \cdot MMI_{SCR} + (1 - CI) \cdot MMI_{ACR} \quad (4.18)$$

It is noted that, while the aforementioned approach appears 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.

At the time of deciding which IPEs to use for the calculation of MMI_{SCR} and MMI_{ACR}, it became clear that the combination of that of Atkinson & Wald (2007) for the CEUS and that of Allen *et al.* (2012) would not be the optimum, as the two models have different mathematical forms that lead to them crossing each other for magnitudes, depths and epicentral distances of interest for this work. This means that, for certain combinations of the input parameters, the IPE of Allen *et al.* (2012) predicts much higher values of MMI at short distances than that of Atkinson & Wald (2007) for the CEUS, but this situation reverts at longer distances. It is noted that both models are being applied outside of the ranges for which they were derived when using very short epicentral distances and very shallow depths, so this crossing is not necessarily an indication of an inherent flaw of any of the models but a consequence of them being extrapolated to other ranges of interest. Moreover, the model of Allen *et al.* (2012) was derived for earthquakes with moment magnitude larger than 5.0, which is higher than our smallest magnitude of interest (**M**4.0). For these reasons, it was decided that Version 2 would use the two models of Atkinson & Wald (2007) derived for the CEUS and California to calculate MMI_{SCR} and MMI_{ACR}, respectively.

The models of Atkinson & Wald (2007) were derived using earthquakes with moment magnitudes and distances to the fault in the ranges [**M**2.0-7.8, 6-1,000 km] and [**M**2.4-7.8 and 1-500 km] for the CEUS and California, respectively. The mathematical shape is the same for both, with different coefficients. In the plots of Figure 4.26, different shades of grey correspond to different values of the craton index. A craton index of zero (lightest colour) corresponds to the IPE of Atkinson & Wald (2007) for California, while a craton index of one (black) corresponds to the IPE of Atkinson & Wald (2007) with CEUS parameters. These plots show that no strange behaviour stems from combining the two models weighed by the craton index. Nevertheless, it is noted, once more, that the models of Atkinson & Wald (2007) are well constrained only for distances to the fault larger than 10 km and is thus being applied beyond its range of derivation herein.

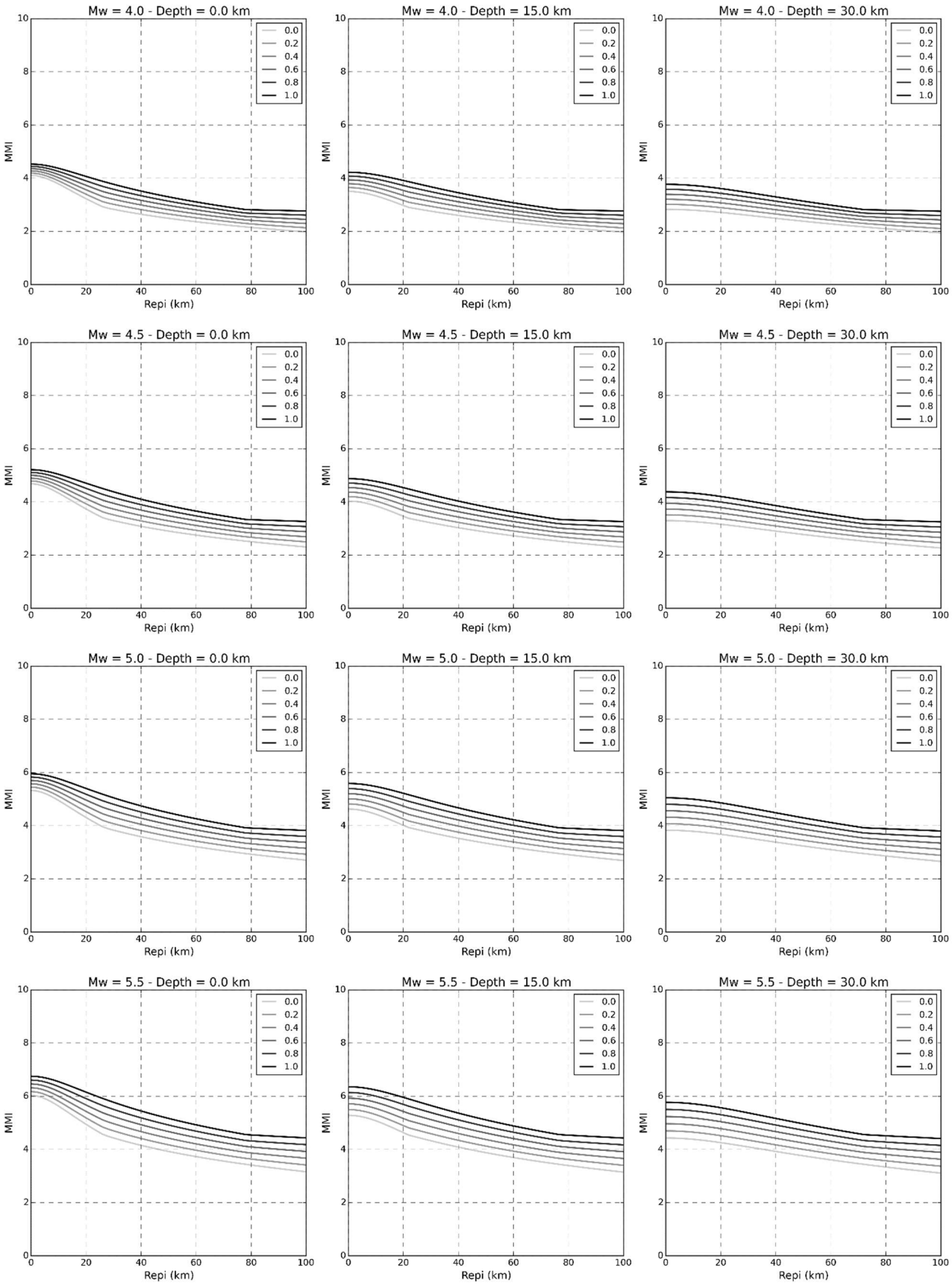

Figure 4.26. Variation of the final MMI estimation according to Equation 4.18 for different values of the craton index (shades of grey: light grey= non cratonic, black=cratonic), magnitude and depth.

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 ($\sigma_{IPE\ SCR(ACR)}$) and the uncertainties in the values of depth and magnitude (σ_h and $\sigma_{\mathbf{M}}$, respectively, the latter as per Equation 4.16). The total standard deviation was thus calculated as per Equation 4.19, in which the derivative terms refer to the derivative of the IPE with respect to depth (∂IPE/∂h) and magnitude (∂IPE/$\partial\mathbf{M}$).

$$\sigma_{\substack{MMI\ SCR \\ ACR}} = \sqrt{\sigma^2_{\substack{IPE\ SCR \\ ACR}} + \sigma^2_h \cdot \left(\frac{\partial IPE_{\substack{SCR \\ ACR}}}{\partial h}\right)^2 + \sigma^2_{\boldsymbol{M}} \cdot \left(\frac{\partial IPE_{\substack{SCR \\ ACR}}}{\partial \boldsymbol{M}}\right)^2} \tag{4.19}$$

As Atkinson & 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. Figure 4.27 shows the resulting values for three different depths and a range of epicentral distances. The standard deviation is only dependent on the hypocentral depth in this model. As can be observed, values vary between 0.8 and 1.2 MMI units for the range of depths of interest in this work.

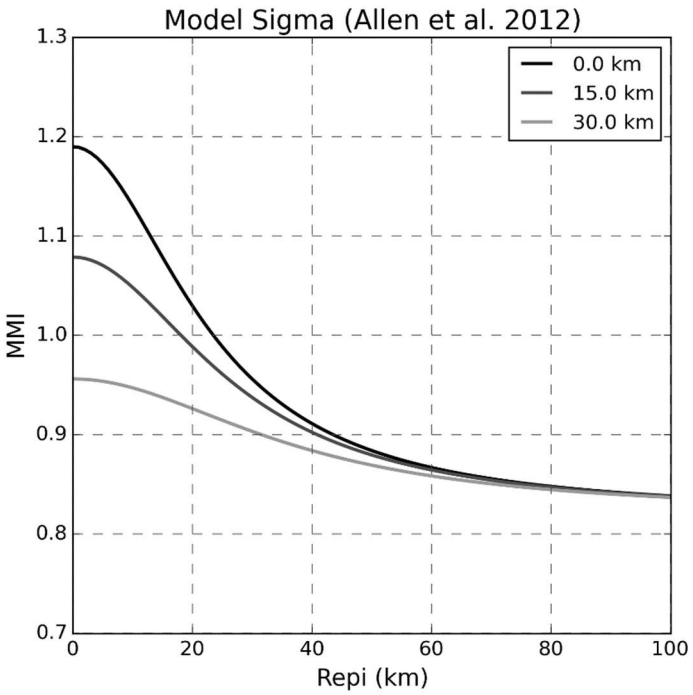

Figure 4.27. Standard deviation of the IPE of Allen *et al*. (2012) based on hypocentral distance, for different depths (shades of grey) and epicentral distances.

Strictly speaking, $\sigma_{IPE\ SCR}$ and $\sigma_{IPE\ ACR}$ should be combined using the same CI and (1-CI) weighting factors applied in Equation 4.18 for the expected values, as per Equation 4.20.

$$\sigma_{MMI\ final} = \sqrt{CI^2 \cdot \sigma^2_{MMI\ SCR} + (1 - CI)^2 \cdot \sigma^2_{MMI\ ACR}} \tag{4.20}$$

However, a closer look to this equation reveals its undesired non-physical behaviour. As $\sigma_{MMI\ ACR}$ and $\sigma_{MMI\ SCR}$ greatly depend on the standard deviation of the IPE, which is the same for both, and not so much on the derivatives of the IPE with respect to magnitude and depth, where the different coefficients for California and the CEUS intervene, they can be very close in value. In the extreme case in which they are the same (*i.e.*, $\sigma_{MMI\ ACR} = \sigma_{MMI\ SCR}$) the final standard deviation becomes a function of the craton index as per Equation 4.21. The factor multiplying $\sigma_{MMI\ SCR}$ behaves as shown in Figure 4.28, with a minimum for CI=0.5 and maxima for both CI=0 and CI=1.

$$\sigma_{MMI\ final} = \sigma_{MMI\ SCR} \cdot \sqrt{CI^2 + (1 - CI)^2} = \sigma_{MMI\ SCR} \cdot \sqrt{2 \cdot CI^2 - 2 \cdot CI + 1} \tag{4.21}$$

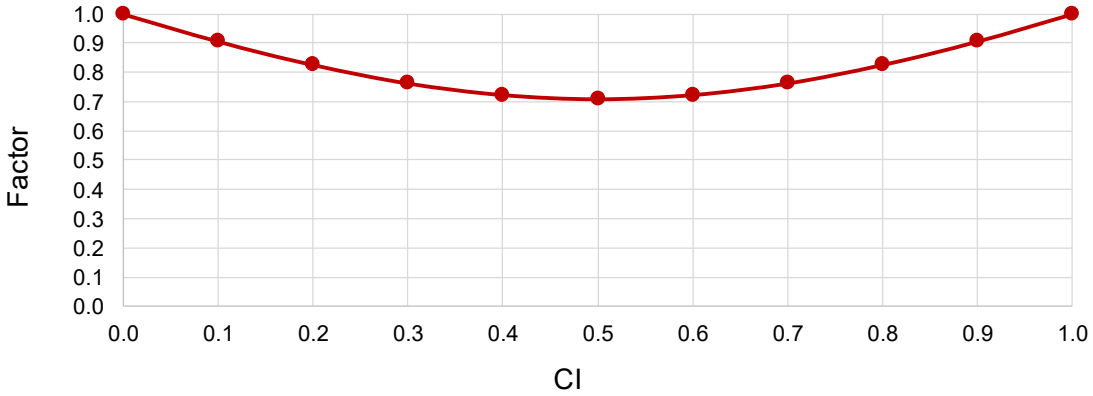

Figure 4.28. Factor multiplying the standard deviation of MMI against the craton index.

As a consequence, the standard deviation has a minimum at CI=0.5 and is maximum for both CI=0 and CI=1 or, in other words, the uncertainty in the prediction of MMI is 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. This does not hold real physical meaning and is, thus, a statistical artefact that results in the probability of observing certain values of MMI under certain conditions behaving strangely with respect to the craton index, as illustrated by the plots on the top of Figure 4.29. The black lines depict the expected MMI plus/minus one standard deviation, according to the scale on the left, while the red and blue continuous lines depict the probability of observing MMI≥IV and MMI≥V, respectively. The blue line on the top left plot first decreases for increasing values of CI, reaches a minimum, and then increases, when the expected physical behaviour would be a steady increase.

In view of this, the final standard deviation was taken as a weighted average of the two, $\sigma_{MMI\ ACR}$ and $\sigma_{MMI\ SCR}$, as per Equation 4.22. The plots on the bottom of Figure 4.29 show how this equation yields the desired behaviour.

$$\sigma_{MMI\ final} = CI \cdot \sigma_{MMI\ SCR} + (1 - CI) \cdot \sigma_{MMI\ ACR} \tag{4.22}$$

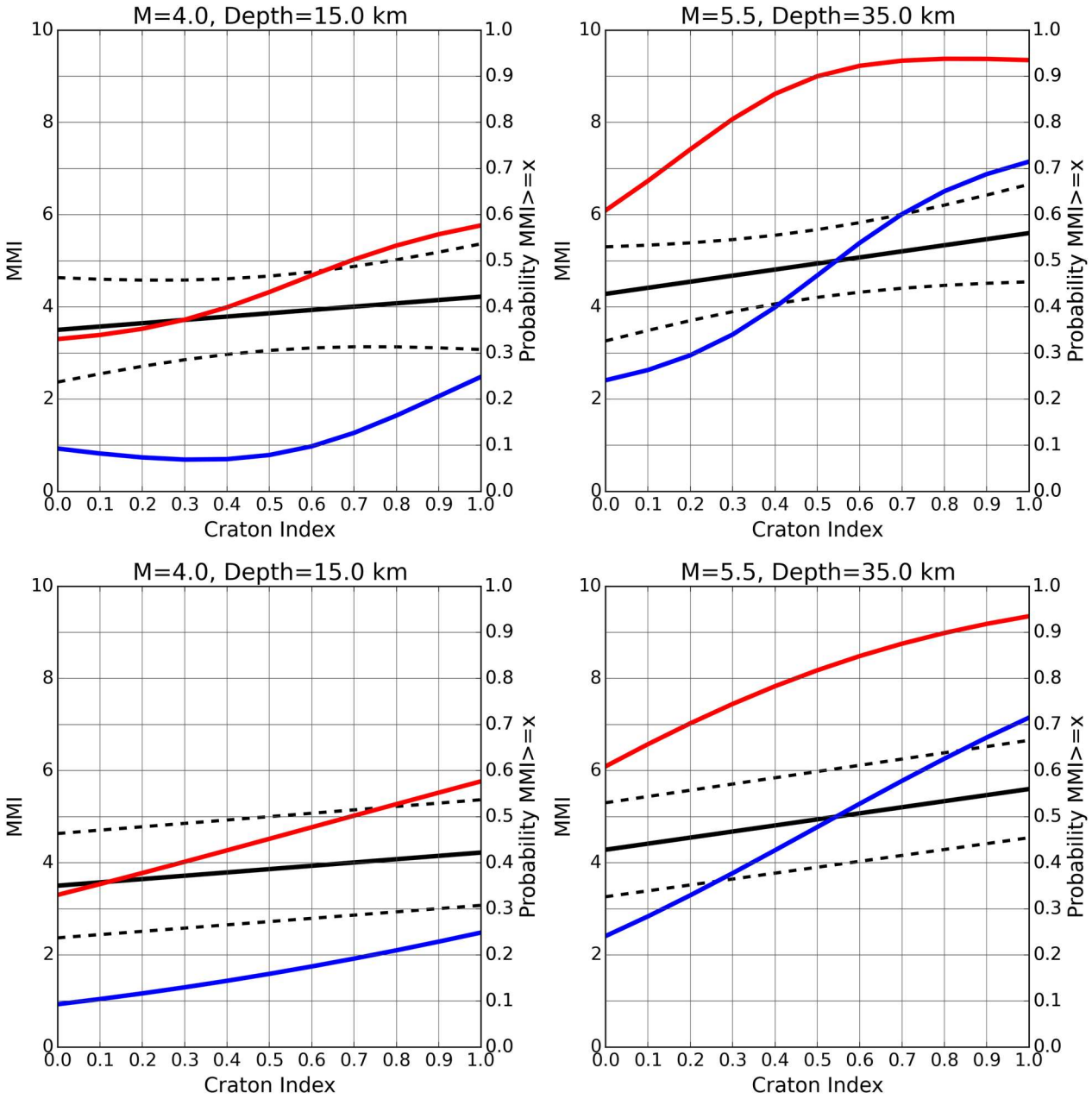

Figure 4.29. Expected MMI ± one standard deviation (black lines, left-hand scale) and probabilities of observing MMI≥IV (red) and MMI≥V (blue), according to right-hand scale, using Equations 4.20 (top) and 4.22 (bottom) for the final standard deviation. Uncertainty in moment magnitude = 0.3. Uncertainty in depth= 20% of depth value. Epicentral distance = 0 km.

It is of interest to note that the use of the craton index of Chen *et al.* (2018) results in oceanic crust being treated as an active crustal region, which is debatable. Being of a completely different nature, oceanic crust cannot be classified into cratonic or non-cratonic. While there appears to be evidence to support the idea that attenuation properties of the oceanic crust are closer to those of a stable continental region than an active crustal one (*e.g.*, Vilanova *et al.*, 2012), the similarity or dissimilarity with the latter has not been yet explicitly quantified. For this reason, and noting that oceanic crust tends to start at several tens if not hundreds of kilometres off shore, the default craton index value of 0 was used herein for earthquakes nucleated within the oceanic crust.

It is noted that 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. Directly linked to this is the fact that real earthquake ruptures are finite and the consideration of epicentral distance instead of other forms of distance that can better represent the actual rupture reduces the latter to a point source. The Atkinson & Wald (2007) IPEs are intended to be used with the distance from the fault, though the authors explicitly mention that the latter can be considered equivalent to hypocentral distance for small to moderate earthquakes. As finite-rupture geometries are only calculated for earthquakes deemed of relevance for the scientific community, it would be impossible to use any metric other than hypocentral distance herein.

The influence of site effects and soil conditions has not been considered herein either for two reasons. Firstly, because the IPEs of Atkinson & Wald (2007) do not include site-specific terms. 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 Vs30 world model of Wald & Allen, 2007), and thus results in the incorporation of further uncertainties. Moreover, the purpose of this work is not to determine in detail the intensity potentially generated by each earthquake of the catalogue at each site but to be able to classify earthquakes in 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 & Wald (2007) nor focal mechanisms being available for all earthquakes in the catalogue.

4.6.3 Estimation of Exposed Population Accounting for Uncertainty

The procedure described in Section 4.6.2 was used to estimate the expected Modified Mercalli Intensities (MMI) and their associated uncertainties across a spatial grid of site locations for each earthquake. Expected (median) MMI values were calculated as per Equation 4.18 using expected values of depth and magnitude, while the uncertainties in depth and magnitude contributed to the overall final uncertainty as per Equations 4.19 and 4.22. Each cell of the spatial grid was thus associated with a distribution of MMI values from which the probability of observing intensities greater than IV or V can be calculated, as shown schematically in Figure 4.30. The spatial grid was defined to match the grids of population count and density of Gridded Population of the World GPW v4.0 (CIESIN, 2016).

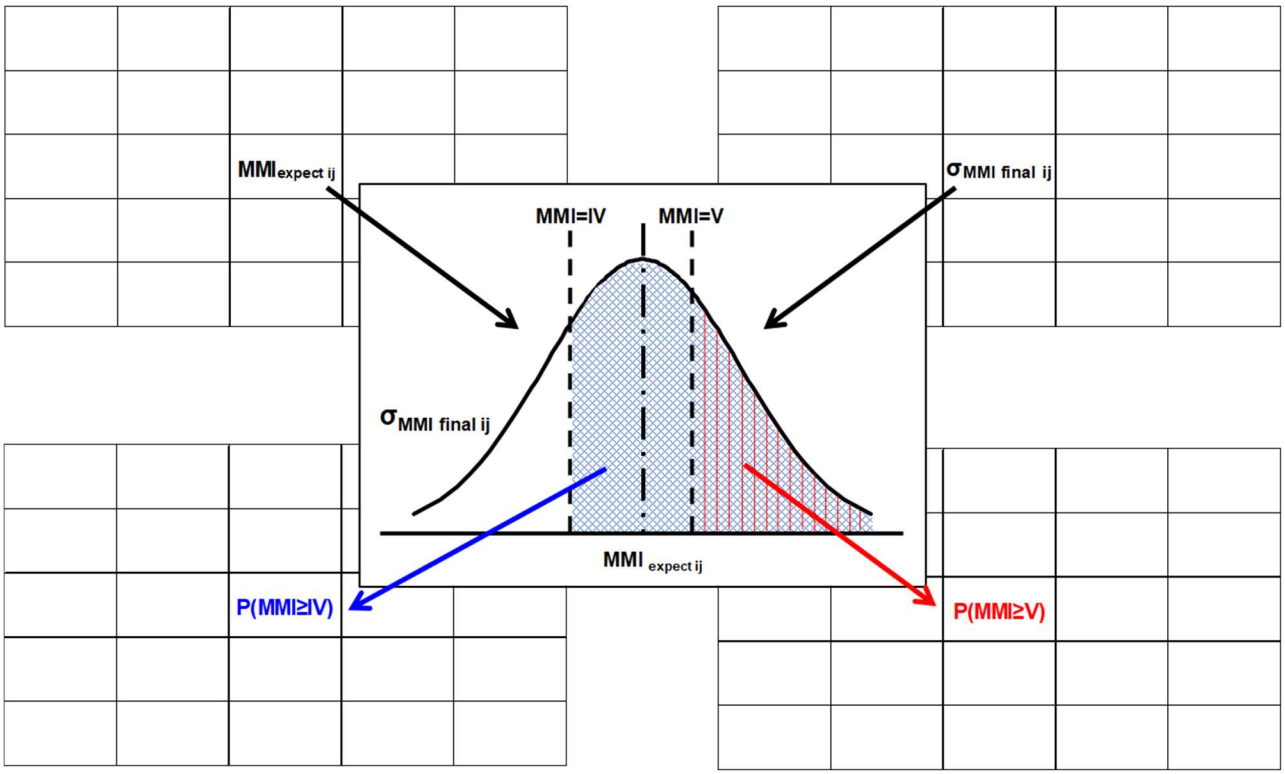

Figure 4.30. Schematic representation of the process of determining the probability of observing MMI values of IV (blue) or V (red) and higher for each longitude-latitude cell.

Each cell was counted as being exposed to potentially damaging shaking if the probabilities of observing MMI values of IV and V exceeded the pre-defined thresholds defined in Figure 4.31, whose calculation is explained in Section 4.6.4 below. For example, for an earthquake with **M**5.5 a cell needed to have a probability of observing MMI≥IV of at least 55% and a probability of observing MMI≥V of at least 22% in order to be counted. For those cells that complied with both probability thresholds, their population count was added up and the largest of their densities was selected. The earthquake was then included in the catalogue of potentially damaging earthquakes if the population exposed to MMI≥IV *and* MMI≥V with probabilities of at least those defined in Figure 4.31 was at least 2,500 people *and/or* the maximum population density of all compliant grid cells was equal to or larger than 300 people/km^2.

The selected population and density thresholds, 2,500 people and 300 people/km^2, are the same as in Version 1, and a discussion on their selection can be found in the corresponding report (Nievas *et al.*, 2017) as well as Section 3.4.1 of the present report. The difference with Version 2 lies in how the cells are selected to be considered or not. As Version 1 did not take into account uncertainties, the condition was that the cell had a probability of 50% of observing MMI≥IV.

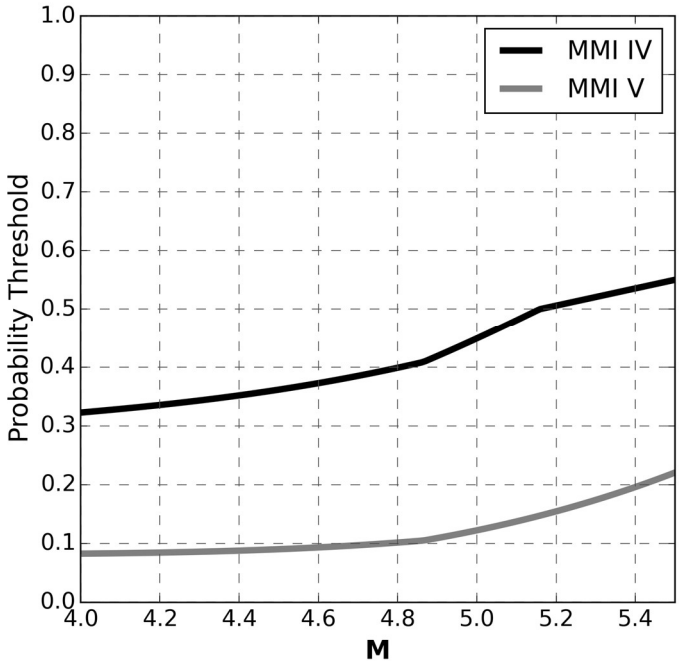

Figure 4.31. Magnitude-dependent probability thresholds for MMI≥IV and MMI≥V.

4.6.4 Probability Thresholds in the Passing Criterion

Two fundamental aspects of the earthquakes considered for the present work are, as explained in Chapter 1, that they are upper-crustal events and that they have occurred sufficiently close to human settlements to pose a threat to the population and/or the built environment. While, on the one hand, these conditions refer to different constraints, both are deeply interrelated by their direct connection with hypocentral depth. This implies that a constraint on one is inevitably reflected on the other. This becomes evident when plotting the probability of observing epicentral macroseismic intensities above a certain threshold as a function of moment magnitude and hypocentral depth, as shown in Figure 4.32. These plots show, for example, that a **M**4.0 earthquake occurring at a depth of 10 km has a 35% (Allen *et al.*, 2012), 15% (Atkinson & Wald, 2007, California) or 30% (Atkinson & Wald, 2007, CEUS) probability of causing a Modified Mercalli Intensity (MMI) of V or above at the epicentre[3]. This means that if the criterion to consider that this earthquake poses a threat was the number of people with a 50% chance or more of being exposed to shaking of MMI V or above the answer would be, straight away, that this earthquake presents no threat. Setting this criterion would effectively translate into assigning a maximum depth of around 7.3 km to earthquakes with magnitude **M**4.0, according to the model of Allen *et al.* (2012), or not considering earthquakes of this magnitude at all, according to the model for California of Atkinson & Wald (2007), irrespective of the maximum depth used to define an upper-crustal earthquake.

[3] None of the IPEs should be applied at the epicentre for hypocentral depths smaller than 6-10 km. However, the results shown herein are not affected by this limitation.

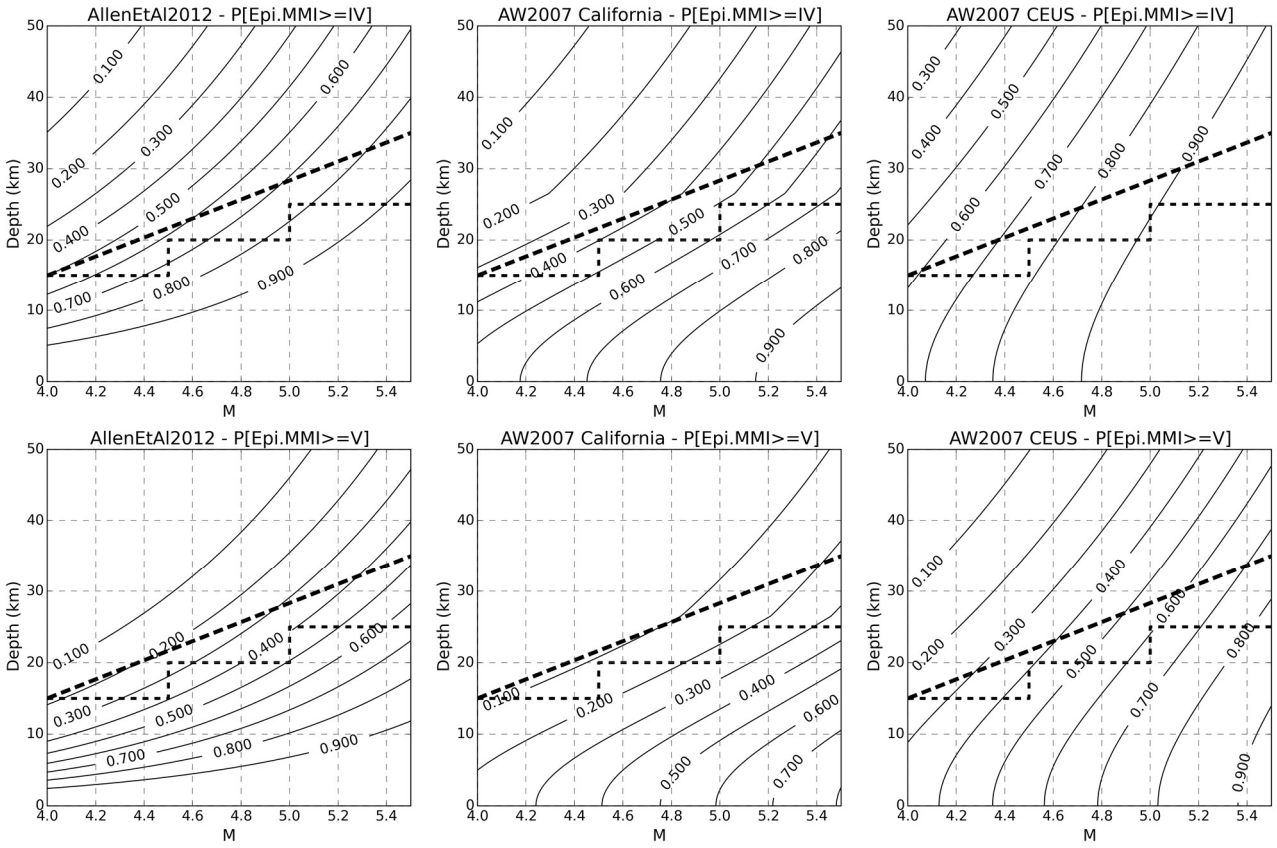

Figure 4.32. Probability of observing Modified Mercalli Intensities of IV (top) or V (bottom) and above at the epicentre as a function of moment magnitude and hypocentral depths, according to three IPEs: Allen *et al.* (2012; AllenEtAl2012) and Atkinson & Wald (2007; AW2007) for California and the Central and Eastern United States (CEUS) (standard deviation model of Allen *et al.*, 2012, in all cases). Thinner and thicker dashed lines indicate the maximum depth criteria used in Versions 1 and 2, respectively. All IPEs applied beyond their derivation range.

Along the same line of thought, Figure 4.33 shows the 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 with the maximum depth criterion used in Version 1, using the IPE of Allen *et al*. (2012). The largest probabilities correspond to 0-km depths, while dashed lines indicate the probabilities for the maximum depths. A criterion of at least 50% probability of observing MMI of V or more would imply generating a horizontal cut at 0.5 and discarding all the possible depths below it. This was one of the reasons for using MMI values of IV as a threshold in Version 1, and not MMI of V, despite the former being characterised, by definition, by the rattling of dishes, windows and doors (but without any of these breaking or cracking), the creaking of walls and frames, and the swinging of hanging objects, and the latter being defined as involving some instances of damage such as broken dishes and/or cracked windows (Wood & Neumann, 1931). Figure 4.34 is analogous to Figure 4.33 but using the IPE of Atkinson & Wald (2007) and the maximum depth criterion used in Version 2.

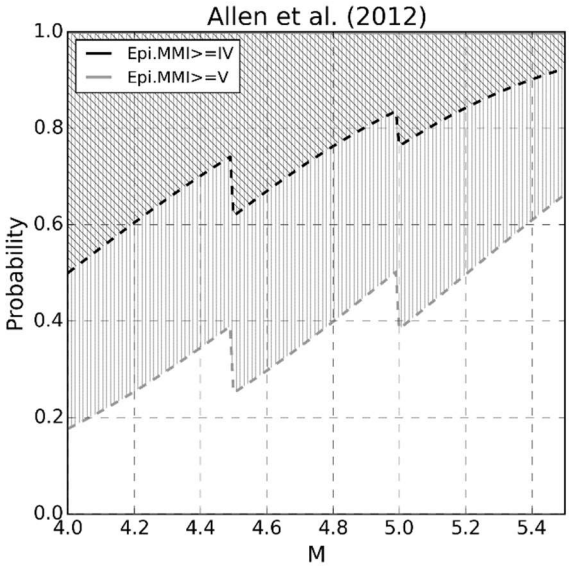

Figure 4.33. Range of probabilities of observing MMI values of IV (black) or V (grey) and above at the epicentre for the whole range of hypocentral depths considered with the maximum depth criterion used in Version 1, using the IPE of Allen *et al.* (2012). Continuous lines (close to a probability of 1.0 for all magnitudes) and dashed lines correspond to 0-km depth and maximum considered depth, respectively.

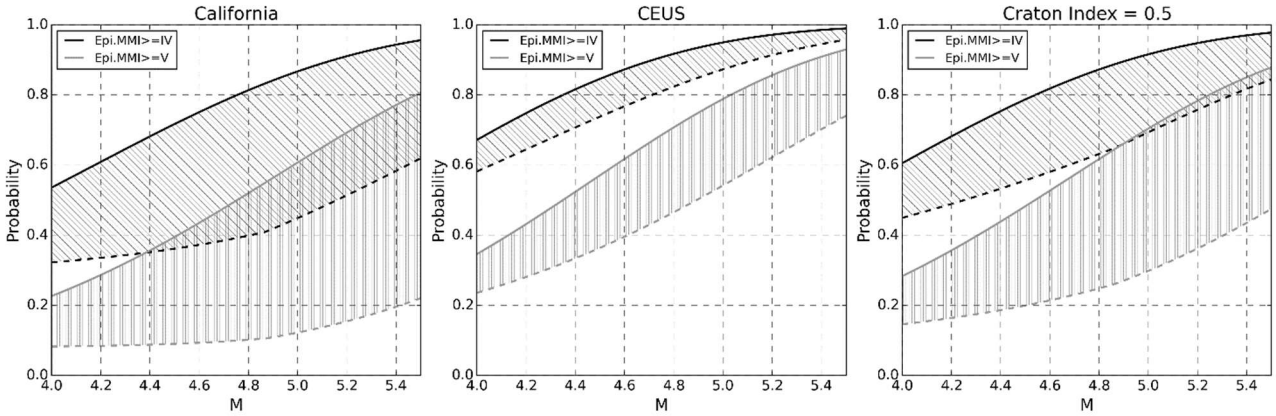

Figure 4.34. Range of probabilities of observing MMI values of IV (black) or V (grey) and above at the epicentre for the whole range of hypocentral depths considered with the maximum depth criterion used in Version 1, using the IPE of Atkinson & Wald (2007) with coefficients derived for California (left), the CEUS (centre) and results weighted by a craton index of 0.5 (right, equivalent to average of resulting MMI distributions from California and CEUS), and the standard deviation model of Allen *et al.* (2012). Only the uncertainty of the IPE considered. Continuous and dashed lines correspond to 0-km depth and maximum considered depth, respectively.

Two alternative approaches emerge from this situation: either to fix a maximum depth criterion and be consistent with it when defining the intensity/exposure criteria (*i.e.*, not introducing a more restrictive limit on maximum depth disguised within the latter), or to leave the maximum depth to be controlled by the intensity/exposure criteria, only introducing a maximum cap beyond which it would be impossible to consider the earthquake as upper-crustal. Figure 4.35 illustrates what the maximum depths corresponding to different probabilities of observing MMI≥IV or V at least right above the hypocentre (with two different levels of uncertainty in the magnitude estimation) would look like. The uncertainty in

magnitude ($\sigma_{\mathbf{M}}$) has no influence when considering a 50% probability of observing MMI IV or larger due to a 50% probability depending only on the median and not the standard deviation of the distribution. As can be observed, maximum depths defined in this way for a non-cratonic setting (California, craton index = 0, plot on the left of Figure 4.35) appear as relatively reasonable, particularly when ignoring the uncertainty in magnitude and for a 50% probability of observing MMI IV or a 20% probability of observing MMI V or larger. However, values get unrealistically large for purely-cratonic environments (CEUS, craton index = 1), and even for a craton index of 0.5. This behaviour of the IPE would thus result in the need to introduce caps to the values so defined. In contrast with the alternative about to be presented, this need for caps appears as far less consistent.

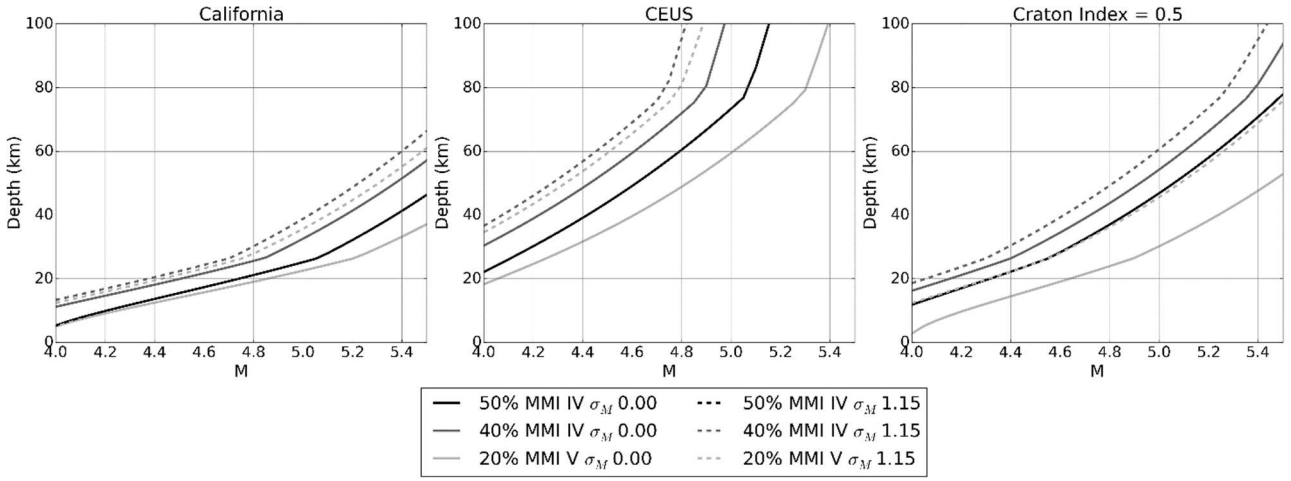

Figure 4.35. Maximum depths defined from the probabilities of observing different MMI levels calculated using the IPE of Atkinson & Wald (2007) with coefficients derived for California (left), the CEUS (centre) and results weighted by a craton index of 0.5 (right, equivalent to average of resulting MMI distributions from California and CEUS). Null error in depth. Uncertainty in magnitude ($\sigma_{\mathbf{M}}$) as per legend.

Having adopted a maximum depth criterion as described in Section 4.5.2, the adoption of intensity/exposure criteria that do not introduce further constraints on it implies the need to guarantee that an earthquake that occurred at the maximum depth considered for its magnitude passed the intensity threshold criteria at least at the epicentre, so as to be able to consider, at least, population and assets located right above it. The way in which this can be achieved is by setting the probability threshold for a particular MMI value to match the probability of observing that value or larger with the deepest hypocentre considered for its particular magnitude, in a non-cratonic environment, which is the most restrictive. In other words, the probability thresholds to be used for the intensity/exposure criteria would be the dashed lines in Figure 4.34. Taking MMI equal to or larger than V as the intensity of interest, for example, this approach would imply that a probability of at least 8.26% of observing MMI equal to or larger than V at a site is necessary for the site to be counted when subject to an earthquake of magnitude **M**4.0 occurring in an active crustal region (craton index = 0, equivalent to using the California coefficients, left-side of Figure 4.34). However, a further consideration needs to be made. The probability of observing a certain MMI value depends not only on the expected MMI but also on its associated uncertainty, which is a function of the standard deviation of the IPE, the measurement error of the depth and the combined

measurement error and standard deviation of the magnitude conversion models. The plots of Figure 4.34 consider only the first of these three. As a consequence, for any magnitude-depth combination for which the expected MMI is smaller than IV, the probability of observing an MMI equal to or larger than IV when considering all sources of uncertainty is larger than when considering only the standard deviation of the IPE. The final standard deviation is always larger than the latter, as all the terms in the equation of the propagation of error are squared, and this results in larger probabilities of observing MMI values larger than the median. The counterpart of this is that it also results in smaller probabilities of observing MMI values equal to or larger than a threshold smaller than the median for that combination of magnitude-depth.

In order to overcome this limitation, the probability thresholds can thus be 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. In the case of depth, a measurement error of 0.001% of the maximum depth was considered, as this was the value adopted for ratios of expected to maximum depth tending to unity, which is the case here as the maximum depth for each magnitude is being used. A value needs to be adopted, as it is clear from Figure 4.23 that the 0.5-probability contour line becomes asymptotic with the expected to maximum depth ratio of 1.0. For the case of magnitude, a final uncertainty (resulting from combining measurement errors with the uncertainty of conversion models as per Equation 4.16) of 1.657 was used, as this is the largest value found in WPG16v3c for the time period of interest and is larger than the largest found in the merged catalogue as well (see Appendix I for details). The criteria resulting from this approach are depicted in Figure 4.36 for MMI IV, V and VI. As can be observed, for MMI≥V and MMI≥VI the final thresholds are always those calculated only with the standard deviation of the IPE, as the expected MMI does not reach V or VI in any of these cases. For MMI IV, the final thresholds are those calculated from a combination of both uncertainties.

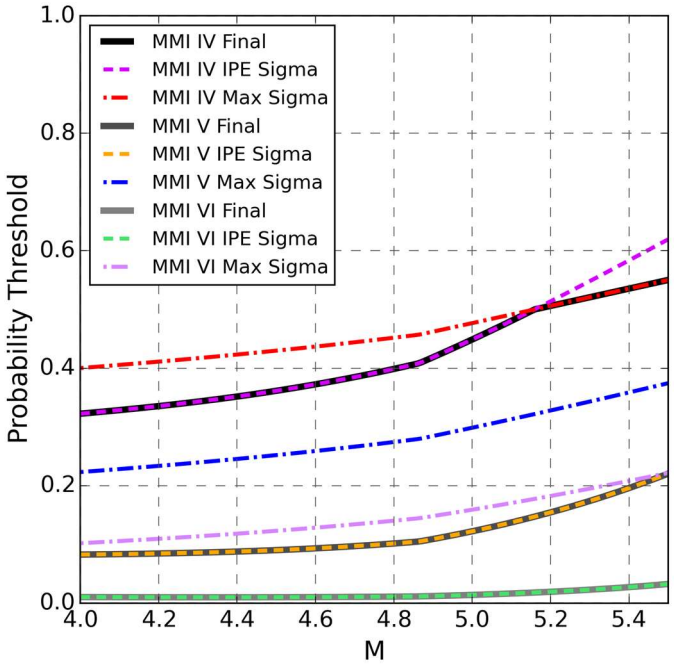

Figure 4.36. Probability thresholds for different MMI levels and uncertainty cases. Final adopted thresholds shown in continuous black and grey lines.

As, in Version 2, macroseismic intensity was determined considering a craton index used to weight the MMI values obtained using the California and CEUS coefficients of the IPE of Atkinson & Wald (2007), as explained in Section 4.6.2, the probability threshold could thus become both magnitude- and location-dependent. However, this would lead to a contradiction in the resulting behaviour of the exposure criterion in cratonic vs. non-cratonic environments, as discussed in Appendix I. For this reason, a craton index of 0 was used to calculate the probability thresholds shown in Figure 4.36 and adopted hereafter. Figure 4.37 depicts the maximum epicentral distances that satisfy the condition that the probability of MMI being equal to or larger than IV, V or VI (left, centre and right plots) be larger than the thresholds defined in Figure 4.36 for a hypocentre located at a depth of 0.1 km with an error of 0.1 km as well. These distances are the largest radii from the epicentre within which population and population densities may be counted to determine whether the earthquake satisfies the exposure criterion or not.

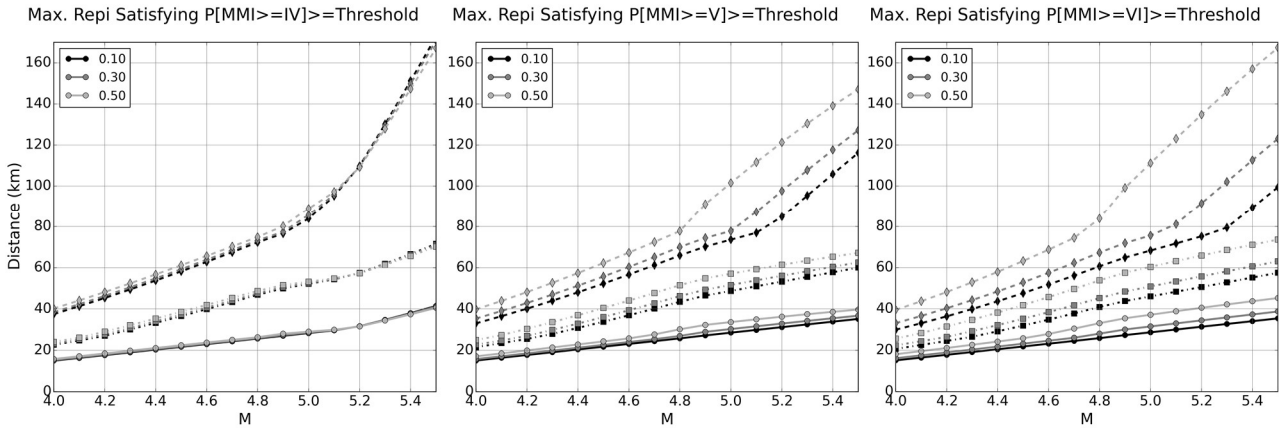

Figure 4.37. Maximum epicentral distances satisfying the condition that the probability of MMI being equal to or larger than IV (left), V (centre) and VI (right) is larger than the thresholds defined in Figure 4.36, for different uncertainties in magnitude (colour scale). Continuous, dotted and dashed lines indicate craton index = 0 (California coefficients), craton index = 0.5, and craton index = 1 (CEUS coefficients), respectively. Thresholds independent from craton index (i.e. determined using craton index = 0).

This approach is deemed more consistent than that of not applying maximum depth criteria other than that resulting implicitly from the consideration of a certain probability of exceeding a certain estimated MMI level (Figure 4.35), as it allows for an initial decoupling of the problem to apply reasonable depth limits according to the condition of the events being upper crustal and capable of causing sufficient shaking, and then move on to estimating the intensity they may have caused at different locations, instead of applying caps where the effective maximum depths exceed reasonable upper crustal values, which would result in different criteria being applied to different earthquakes depending on their moment magnitudes. As a consequence, this criterion has been adopted as described in Section 4.6.3. The thresholds depicted in Figure 4.31 stem from Figure 4.36 and the considerations herein described.

It is relevant to highlight that this approach implicitly assumes that an earthquake that occurs at a depth equal to the maximum accepted for its magnitude (Figure 4.21) is potentially

damaging if the grid cell right above its hypocentre complies with the 2,500 people and/or the 300 people/km^2 population count and density thresholds. Due to this assumption, the choice of the specific MMI levels to consider (IV and V herein) is not of major influence, as the key lies in the definition of the probability limits as those that are calculated right at the epicentre of an earthquake that occurs at the maximum considered depth.

4.6.5 Sensitivity to Maximum Depth Criteria Accounting for Changes in Exposure Criteria Introduced in Version 2

As explained in Section 3.2.5, the sensitivity analysis carried out on the maximum depth criterion in Section 3.2 still applied the exposure criterion defined in Version 1 (50% probability of observing MMI values of IV as per the IPE of Allen *et al.*, 2012), which acted as an implicit maximum-depth limit due to it not being linked to the exposure criteria. As this has been changed in Version 2 as per the considerations of Section 4.6.4, the sensitivity analysis is repeated herein, with probability thresholds of observing MMI values of IV defined as a function of the maximum depth assigned to each magnitude, with the purpose of guaranteeing that an earthquake occurring at such a depth is able to cause these probabilities at least in the cell right above the hypocentre. The catalogue of Version 1 and the IPE of Allen *et al.* (2012) have been maintained as in Section 3.2.

The results obtained are shown in Table 4.3. As can be observed, the exposure criterion no longer acts as a homogeniser of the results as was the case in Table 3.3 (two right-most columns), and the proportion of damaging earthquakes now reflects the same tendency of the number of non-damaging events increasing faster than the number of damaging ones after the exposure criterion has been applied. This tendency is illustrated in Figure 4.38. These results suggest that the adoption of the linear criterion leads to conservative results, *i.e.*, larger proportions of damaging earthquakes than the other criteria being compared.

Table 4.3. Results for the sensitivity analysis over the maximum depth criterion. Catalogue vs. Damaging refers to the complete catalogue and the subset of it reported to have caused damage. All vs. MS refers to all shocks and only main shocks, respectively. Depth refers to earthquakes being filtered according to the maximum depth criteria. Exposure refers to earthquakes filtered both by depth and exposure criteria. The last four columns depict the proportion of damaging to total earthquakes. Exposure criterion dependent on maximum depth criterion.

Criteria	Catalogue Depth		Catalogue Exposure		Damaging Depth		Damaging Exposure		Depth		Exposure	
	All	MS	All	MS	All	MS	All	MS	All	MS	All	MS
Original v1	101,248	32,842	35,654	11,968	287	N/A	282	185	0.28%	N/A	0.79%	1.55%
Original	101,248	32,842	28,954	9,932	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%
NOAA 50	125,335	37,832	40,133	13,234	314	207	301	196	0.25%	0.55%	0.75%	1.48%
NOAA 75	198,957	56,704	64,753	20,800	359	243	353	238	0.18%	0.43%	0.55%	1.14%
NOAA 80	218,894	61,440	80,927	25,030	374	253	371	251	0.17%	0.41%	0.46%	1.00%
MAX 35km	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%

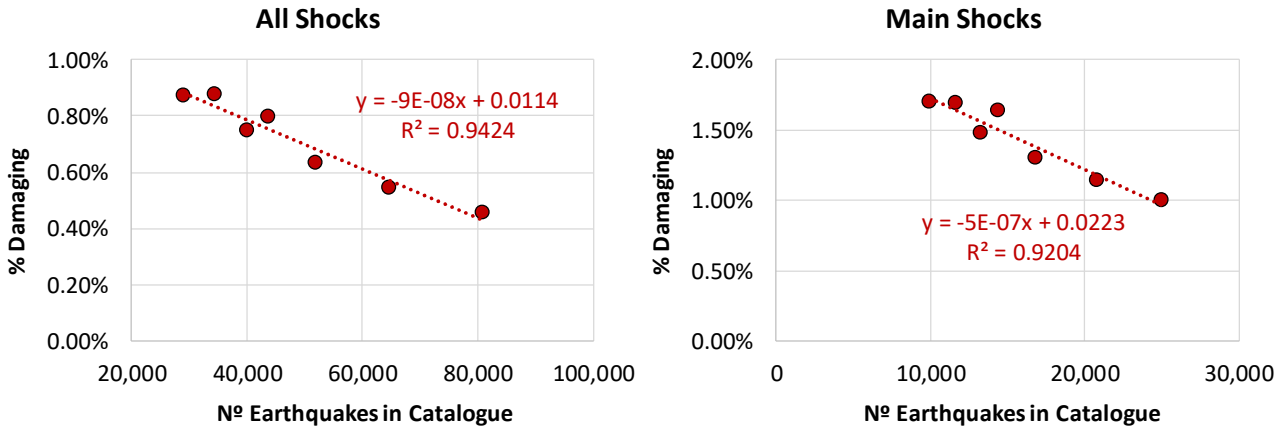

Figure 4.38. Proportion of damaging to total number of earthquakes in the catalogue of potentially damaging events (Version 1) for different maximum-depth criteria: all shocks (left) and only main shocks (right).

4.7 Flagging of Induced Events

As in Version 1, a first flagging of induced events was carried out using the technique suggested by Weatherill *et al.* (2016) of searching for certain keywords within the event comments of the ISC Bulletin. The toolkit published alongside the paper of Weatherill *et al.* (2016) was thus used with the keywords “geothermal”, “reservoir”, “mining”, “anthropogenic”, and “rockburst” (this last one added in Version 2)[4]. Two additional verifications were introduced in Version 2. First, version 3c of WPG16 includes a field for anthropogenic events. This field was used alongside the code developed in Version 1 and earthquakes were flagged as induced if any of the two indicated so. Lack of agreement was only encountered in three cases in which the ISC Bulletin had clearly been updated with respect to the version used to generate WPG16v3c, and four cases of split/duplicate events that exist in WPG16v3c but can no longer be found in the Bulletin. The use of this field of WPG16v3c allowed to eliminate the “not classified” category used in Version 1 for the earthquakes that could not be verified back then.

The second verification introduced in Version 2 is the use of the Human-Induced Earthquake Database (HiQuake) (Foulger *et al.*, 2016; Foulger *et al.*, 2018; Wilson *et al.*, 2017; Induced Earthquakes, 2017) as an additional source to indicate potentially induced events, by means of the procedure described in what follows. The version of HiQuake used herein was downloaded from the corresponding website in November 2017.

[4] “Induced” was not included as a keyword because it picked up too many cases of earthquakes in which the description was "XXX was induced by slip from YYY" (Weatherill, 2019, pers. comm.).

4.7.1 The Human-Induced Earthquake Database (HiQuake)

The Human-Induced Earthquake Database is not a list of individual earthquakes identified as being induced, but a list of projects or precursory events that have each generated series of earthquakes. Among many other properties, each entry is described in terms of the latitude and longitude of the project/precursory event, the category of causative activity to which it belongs (*e.g.*, mining, geothermal, *etc.*), the time at which the project started/ended and at which the seismicity was first and last monitored, the maximum magnitude observed, *etc.*, though not all entries contain all fields of information. The process followed to transform this database into a tool to flag potentially induced events is described in the following section.

The database, as well as the report about it of Foulger *et al.* (2016), group observed induced seismicity according to the human activities believed to have caused it. These are:

- Impoundment of water reservoirs: Seismicity may be induced by the change in local stresses due to the addition of the mass of water.

- Construction of tall buildings: Similarly, seismicity may be triggered by the change in local stresses due to the weight of the buildings. According to Foulger *et al.* (2016), the case of the Taipei 101 building in Taiwan is the only one for which a report of induced seismicity of this kind exists.

- Coastal engineering: Similarly, seismicity may be triggered by the change in local stresses due to the weight of shingle accumulated along harbours and coastlines. The only seismic event for which this potential triggering cause has been suggested is the 2007 Folkestone earthquake in the United Kingdom, thought the supporting evidence is not strong, as will be discussed below.

- Quarrying: On the other extreme, seismicity may be induced by the change in local stresses due to large masses being removed from the Earth's surface.

- Groundwater extraction: Seismicity may be triggered by the change in local stresses and pore pressures due to the extraction of water from the subsurface.

- Mining and solution mining: Seismicity may be induced by the strong perturbations in local stresses that deep mine excavations may cause, as initial stresses drop to atmospheric pressure from relatively large values. In the case of solution mining, a lixiviant liquid is injected via boreholes, circulated through the rock to dissolve the target mineral, and finally extracted by means of a well. Seismicity may be thus triggered by the changed in local stresses and pore pressures caused by the injection and extraction process.

- Tunnel excavation: Similarly, seismicity may be induced by analogous perturbations in local stresses caused by the excavation of tunnels for diverse purposes (roads, railways, power stations, fluid transport, *etc.*).

- Hydrocarbons: Seismicity may be induced by the changes in pore pressures and underground stresses generated by the removal of fluids from oil/gas reservoirs. As many times activities related to the extraction of hydrocarbons take place alongside other activities such as waste-water disposal or the injection of fluids to enhance oil recovery, it is often difficult to determine the exact cause of the induced seismicity.

- Geothermal production: Seismicity may be induced by the relief of stress caused by the cooling and contraction of the geothermal heat source by means of rock fracturing.

- Military waste: Seismicity may be induced by the change in local stresses caused by the injection of military contaminated wastewater into the subsurface.

- Wastewater disposal: Seismicity may be induced by the change in local stresses caused by the injection of waste fluids into the subsurface. It is common in these cases for the relatively cold fluids to cause thermal fracturing when put in touch with much hotter geological formations deep below the surface.

- Water injected for enhanced oil recovery: Once natural reservoir energy decreases to a point in which the natural differential pressure is not enough to drive the hydrocarbons toward the well and the surface, enhanced oil recovery techniques, usually consisting on the injection of an external fluid (liquid or gas) to maintain the reservoir pressure, are applied (Schlumberger 2018a, 2018b). Seismicity may then be induced by the change in local stresses derived from these injections.

- Enhanced Geothermal Systems (EGS): While traditional geothermal production makes use of the natural presence of water in relatively high-permeability rocks, in enhanced geothermal systems fluid is injected at high pressure into a well with the purpose of fracturing the rock by means of hydraulic and thermal processes, and thus increasing permeability and creating a heat exchanger. Seismicity is a natural consequence of an EGS, as rock fracturing is a fundamentally-intrinsic part of the process.

- Geothermal reinjection: Seismicity may be induced by the change in local stresses derived from re-injecting water into exploited geothermal fields with the purpose of maintaining pressure.

- Shale-gas hydrofracturing (“fracking”): As the name describes, shale-gas hydrofracturing consists on fracturing gas-bearing shale formations with the purpose of increasing permeability and, ultimately, extracting the gas. As with EGS, seismicity

is a natural consequence of fracking, as rock fracturing is a fundamentally-intrinsic part of the process too.

- Mine flooding: Seismicity may be induced by the natural groundwater recharge that often occurs when mines are abandoned, and the subsequent change in local stresses and pore pressures in the area.

- Research: Seismicity may be induced voluntarily within the context of scientific projects aimed at studying the mechanics by which induced earthquakes occur. Research projects of this kind include cases of injection, waste disposal, secondary recovery, and mine flooding.

- Natural gas storage: Seismicity may be induced by the changes in underground stresses generated by the storage of natural gas in depleted hydrocarbon reservoirs, aquifers and salt cavern formations, carried out with the purpose of stabilizing gas supply.

- CO2 for oil recovery: See “water injected for enhanced oil recovery” above.

- Carbon Capture and Storage (CCS): See “natural gas storage” above.

- Nuclear explosions: Seismicity may be induced by energy released by nuclear explosions (the explosions themselves are not considered in HiQuake or herein).

4.7.2 Parsing HiQuake

An initial purge of the database was carried out to simplify later processing. This purge consisted on the elimination of 134 entries with maximum observed magnitudes in the range 2.2-2.5 (not including the latter), as they are too small to be relevant for our magnitude range of interest, and the elimination of further 25 entries with reported seismicity/monitoring end date earlier than 1996. Further eliminations were carried out during the parsing process, as discussed below. The following minor edits were needed to facilitate processing as well:

- Whenever the project start/end date were giving two dates specifying details of the activity, the earlier date was kept. For example, for an entry saying “1998 (extraction), 2013 (storage)”, 1998 was kept.
- Whenever the project or seismicity start dates were specifying a year and a month (as a text), the first day of the month was adopted. For example, an entry saying “2005 (September)” was translated into 01/09/2005.
- Whenever the project or seismicity end dates were specifying a year and a month (as a text), the last day of the month was adopted. For example, an entry saying “2006 (December)” was translated into 31/12/2006.
- Following the same reasoning, whenever only a year was specified, the first or last day of the year (depending on whether the field was start or end date, respectively) was adopted.

- Following the same reasoning, whenever only a decade was specified, the first or last day of the decade (depending on whether the field was start or end date, respectively) was adopted.
- Whenever something like “1990s (mid)” was specified, it was translated into 01/01/1995.
- Whenever a range was specified in the fields of maximum distance of earthquakes to project, or distance of maximum observed magnitude to project, the largest distance was adopted.
- Whenever the research needed to carry out the parsing of HiQuake led to finding a start/end date of project/seismicity that was not specified in HiQuake, it was added.

An entry associated with the 2015 **M**7.8 Gorkha earthquake in the Indo-Gangetic plains (Nepal) was removed from HiQuake before parsing the database as well. Inclusion in HiQuake of this entry was due to a paper by Kundu *et al.* (2015) in which the potential influence of groundwater changes on the 2015 **M**7.8 earthquake was investigated and which concluded that earthquakes that occur beneath the Himalayan arc are influenced by human-driven groundwater extraction. However, it is clear that the area is an active margin capable of large events, some of which have already occurred in the past, and for which anthropogenic processes are unlikely to be a main driver or a factor controlling the size of the rupture. Having left this entry within HiQuake would have led to an incredibly large number of earthquakes being classified as potentially induced in an area of the world which is clearly capable of intense tectonic activity.

The parsing process started by assigning four parameters to each of the entries in the database: start date, end date, radius and offset. The start and end date designate a time window during which the activity is assumed to have potentially induced seismicity in the area. The area was defined, by means of the radius and offset, as a circle centred on the coordinates reported in HiQuake, which were assumed to make reference to the geometric centres of the activities and/or the observed seismicity. This spatial influence was then translated into a series of grids with a 0.1º step (around 11 km at the Equator) covering the whole world, one per earthquake causative activity. Given the nature of the grids, the areas originally defined lost their original perfectly-circular shape. A cell was considered to be included within the area if the distance between its midpoint (defined as the average of the latitudes and longitudes of the edges) and the centre of the circle was equal to or smaller than the radius of influence.

The repetition of this process for each anthropogenic activity and each entry falling within it resulted in one grid per anthropogenic activity and a final grid for the whole of the database considered together. A particular 0.1ºx0.1º cell of these grids may either be empty, indicating that no induced seismic activity was inferred to have taken place there from the data contained in HiQuake, or contain a series of dates and quality flags. The dates (start and end) indicate the time window within which each entry may have generated seismicity in the cell. The flags were used to indicate the overall quality of the data, as the spatial and temporal influence of each entry in HiQuake is not always defined and a series of assumptions were needed, given that case-by-case determination of a maximum spatial and

temporal window of potential influence of a particular activity on the seismicity of the area is outside the scope of the present work. These assumptions were dependent on the kind of anthropogenic activity, as will be explained below.

Whenever the surface areas of the project (*i.e.*, reservoir, gas field, etc) were reported in HiQuake, they were used to calculate the radius of influence, assuming a circular shape. Only 42 of the 728 original entries of the database contain this information. Similarly, the maximum distances between earthquakes and projects reported in HiQuake (73 cases out of 728) were used as additional radial offsets beyond the aforementioned radius, instead of assuming them to be measured form the reported longitude-latitude coordinates, for two reasons. Firstly, it was observed that many of these values were quite small in relation to the expected size of the activity (*e.g.*, 100 m reported for the 16.6-km long West Qinling Rail Tunnels, China). Secondly, the uncertainty in earthquake location needed to be accounted for as well and taking, for example, a radius of only 100 m around the reported longitude-latitude might make it impossible for any earthquakes to be effectively associated with these entries. A minimum offset of 5 km was imposed in all cases (*i.e.,* the value reported in HiQuake as the maximum distance between earthquakes and project was adopted only if larger than 5 km). Distances from projects to maximum-magnitude earthquakes associated to them reported in HiQuake were used as maximum distances between earthquakes and projects when the latter were not available. Whenever these parameters were not available in Hi-Quake, activity-dependent pre-defined values (described below) were used. Quality flags of 'A' (high) or 'C' (low) were assigned to the radii and offset values depending on whether they could be calculated/retrieved from information available in HiQuake or if the pre-defined values were adopted.

Definition of the start and end dates of influence of each entry was fed by a combination of parameters from HiQuake: the dates of the start and end of the project, the dates of the start and end of the seismic activity and/or its monitoring, and the year of the maximum observed magnitude. For all causative activities, the order of preference of the start date was the following:

- the date of start of the seismicity/monitoring, if available, or
- the date of start of the project, if available, or
- a date older than all events in the catalogue, effectively assuming the causative activity has always existed.

Qualities A, B and C were assigned to each of these cases, ranking from best quality data ("A") to worse quality data ("C").

The end date was assigned using a similar hierarchical approach, but in an activity-dependent manner, as will be explained below. In all cases, the date selected was compared against the (31st of December of the) year in which the maximum magnitude earthquake was observed, plus a magnitude-dependent time-offset equal to the Gardner & Knopoff (1974) time window commonly used for declustering of earthquake catalogues (rounded up to the next round number of years). The maximum magnitude reported in HiQuake was used

as input for the Gardner & Knopoff (1974) model, treating all scales as moment magnitude for the sake of simplicity, with a value of 1 year used whenever this information was not available. Radii were also defined as a function of the causative activity. The anthropogenic activities are listed below in the same order as they appear in the report about HiQuake of Foulger *et al.* (2016):

- Impoundment of water reservoirs: Information on the size of large water reservoirs in the word was sought from different sources. The first position in the ranking presented by Avakyan & Ovchinnikova (1971) appears to present an unusual relation between surface area and depth when contrasted against other cases in the list. For this reason, and due to the fact that the area seems extremely large to be representative of most reservoirs, this case is discarded. The remaining cases suggest a radius of around 40 km from the epicentral coordinates reported in the database could be a reasonable approximation (surface areas reported in the source converted to radii assuming a circular shape). An additional 5-km offset was defined to be added to either the 40-km radius or the radius resulting from assuming a circular area when taking the values directly from HiQuake. These adopted values are consistent with the statement by Foulger *et al.* (2016) saying that earthquakes may be induced by water impoundment "throughout relatively large regions". The end date was defined as (in order of preference) the date of end of the seismicity/monitoring (quality "A"), or the date of the end of the project plus 20 years (quality "B"), or no date limit was assigned (quality "C"). The 20-year offset was selected as an envelope of the examples of seismicity induced by the impoundment of water reservoirs reported by Foulger *et al.* (2016).

- Construction of tall buildings: There is only one case of induced seismicity attributed to this cause, and it is that of the Taipei 101 building in Taiwan. Lin (2005), the source reported in HiQuake, studied the variation in seismicity in the Taipei basin around the time of construction of the building and concluded that it is possible that the increment in seismic activity that occurred may have been induced by the increase in vertical loading. From a figure in the aforementioned paper it is possible to infer that the study area was a rectangle approximately centred in the building, with sides of around 15.6 and 16.7 km in the E-W and N-S directions, respectively. From these values, a radius of influence of 8 km and an additional 5-km offset were adopted. Also from a figure by Lin (2005), start and end of seismicity dates were defined as 1st July 1997 and 1st July 2005, respectively.

- Coastal engineering: Similar to the case of erection of tall buildings, there is also only one case listed as induced by coastal land gain: the 2007 Folkestone earthquake. While the evidence for listing this as a case of induced seismicity is not strong (only an abstract for the American Geophysical Union 2007 Fall Meeting; Klose, 2007), it was kept for this analysis as a one-by-one study of the earthquakes in HiQuake is not possible, and its removal would be inconsistent with the overall treatment of the database. Based on the coordinates reported in HiQuake and the most distant epicentral coordinates reported by Nievas *et al.* (2019) for this earthquake, an

influence radius of 25 km with a 5-km offset was adopted. As the additional mass that can allegedly induce seismicity is still present in the area, no end date was considered.

- Quarrying: The only case mentioned by Foulger *et al.* (2016) as being caused by quarrying appears in HiQuake listed under the more general Mining term. As it would not be possible to know if other mining cases in HiQuake correspond to quarrying as well without analysing each case, and being quarrying an activity very similar in essence to the more general *mining*, this category was not considered separately herein (see Mining below).

- Groundwater extraction: Of the five cases listed in HiQuake, two were not considered for the flagging of induced events in this work: the case of the Indo-Gangetic plains and the case of the San Joaquin Valley, in California. The reason is simple: as the information provided in the database is not enough to constrain in time and space the influence of groundwater extraction on the seismicity of these areas, and both the Indo-Gangetic plains and the San Joaquin Valley cover very large surfaces, inclusion of these cases would lead to a very large number of earthquakes being erroneously classified as induced. The intent of this is not to deny the potential influence of groundwater extraction on the seismicity of these places, but to acknowledge that both the Himalayan Thrust and California are, by nature, seismically active, and classifying so many of their earthquakes as induced is likely to be a large overestimation. For the remaining cases, a radius of 15 km and an offset of 5 km was used, based on details on the Guadalentin River Basin (Spain) and the Parana Basin (Brazil) obtained from González & Fernández (2015) and Assumpção *et al*. (2010). The end date was defined as (in order of preference) the date of end of the seismicity/monitoring (quality “A”), or the date of the end of the project (quality “B”), or no date limit was assigned (quality “C”).

- Mining and solution mining: Based on the size of the North Antelope Rochelle mine (Wyoming, USA; measured on Google Maps), the largest coal mine in the world, a radius of 20 km and an offset of 5 km were adopted. Given the nature of the anthropogenic activity and the potential capacity of abandoned mines to induce seismicity, the date of the end of the project was not used to assign the end date. The end date was, thus, defined as (in order of preference) the date of end of the seismicity/monitoring (quality “A”), or no date limit was assigned (quality “C”).

- Tunnel excavation: Based on the length of the West Qinling Rail Tunnels (China, 16.6 km) (Robbins, 2018), the Erlangshan Tunnel (Sichuan-Tibet Highway, 8.6 km) (Wikipedia), the Shimizu Tunnel (Japan, 9.7 km) (Wikipedia), the Kan-Etsu Tunnel (Japan, 11 km) (Structurae, 2018), and the Simplon Tunnel (Switzerland/Italy, 19.8 km) (Wikipedia), all of which are listed in HiQuake, a radius of 10 km and an offset of 5 km were adopted. Looking at construction times of these same tunnels, a time offset of 15 years from the date of start of the project or the monitoring of the seismicity was adopted when more specific dates were not available (several dates were added to

HiQuake while gathering the information presented herein). The end date was defined as (in order of preference) the date of end of the seismicity/monitoring (quality "A"), or the date of end of the project (quality "B"), or the date of start of the seismicity/monitoring plus 15 years (quality "B"), or the date of start of the project plus 15 years (quality "B"), or no date limit was assigned (quality "C").

- Hydrocarbons: Radii of 35 and 39 km were adopted for gas and oil fields, respectively, based on median values obtained from PetroData, a global dataset of oil and gas fields compiled by Lujala *et al.* (2007) and available online (see Table 4.4). PetroData Shapefiles were processed to obtain the areas of the oil and gas fields around the world, from which radii assuming circular shapes were calculated. When the exact resource (oil or gas) was not specified, or when both oil and gas were mentioned, a radius of 47 km (corresponding to fields labelled as "oil and gas" in PetroData) was used. An offset of 5 km was used in all cases. A time offset of 15 years from the date of end of the project was adopted, based on the model by van Elk *et al.* (2018) for the development of the seismicity in the Groningen field, the Netherlands, after a cessation of gas production. The end date was defined as (in order of preference) the date of end of the seismicity/monitoring (quality "A"), or the date of end of the project plus 15 years (quality "B"), or no date limit was assigned (quality "C").

Table 4.4. Median radii (km) of gas, oil and oil and gas fields obtained from PetroData (Lujala *et al.*, 2007).

Onshore	Gas	29.96
	Oil	30.08
	Oil and Gas	49.30
Offshore	Gas	29.90
	Oil	30.01
	Oil and Gas	40.34
All	Gas + Oil and Gas	35.16
	Oil + Oil and Gas	38.40
	Oil and Gas	47.21

- Geothermal production: Based on a map of the Enhanced Geothermal System (EGS) of the Cooper Basin (Australia) by Chen (2010), a radius of 20 km was adopted. The same 5-km offset as in all other cases was used as well. The end date was defined as (in order of preference) the date of end of the seismicity/monitoring (quality "A"), or the date of end of the project plus 1 year (quality "B"), or no date limit was assigned (quality "C"). The 1-year offset was adopted as the longest delay between geothermal activity and seismicity reported by Foulger *et al.* (2016) for geothermal production and EGS.

- Military waste: HiQuake contains only one entry related to the injection of military waste into the subsurface, which occurred in the 1960s. As seismicity the project finished in 1966 and seismicity had essentially ceased by the 1980s, significantly

before the start of our time window of interest, no further attention was given to this category.

- Wastewater disposal: Based on distances between injection wells and seismicity observed in the past, a radius of 40 km was adopted. This value is relatively conservative as smaller distances appear to be significantly more common (*e.g.*, 10-15 km for the Painesville, Guy-Greenbrier and Prague earthquakes, according to McGarr, 2014). However, Goebel *et al.* (2017) have reported the triggering of earthquakes 40 km away (or more) in Oklahoma. As in all other cases, a 5-km offset was adopted as well. The end date was defined as (in order of preference) the date of end of the seismicity/monitoring (quality "A"), or the date of end of the project plus 10 years (quality "B"), or no date limit was assigned (quality "C"). The 10 years were selected based on a 1-decade delay of seismicity from the date of start of injection reported by Foulger *et al.* (2016) and a plot by Dieterich *et al.* (2015) that shows rates of around 1 and 0.1 earthquakes/year, 10 and 15 years after shut-in, respectively.

- Water injected for enhanced oil recovery: As Enhanced Oil Recovery (EOR) techniques are applied to aid in the extraction of hydrocarbons, the same criteria described above under "Hydrocarbons" were applied.

- Enhanced Geothermal Systems (EGS): The same criteria as described above for "Geothermal production" were applied.

- Geothermal reinjection: The same criteria as described above for "Geothermal production" were applied.

- Shale-gas hydrofracturing ("fracking"): As HiQuake reports a 0.6-to-7-km maximum distances of earthquakes to project for 11 fracking entries, and maximum distances reported by others are within this limit (*e.g.,* 0.4 km to the best recorded seismic event of 2011 in Blackpool, United Kingdom, Clarke *et al.*, 2014), a 7-km radius was adopted, together with the standard 5-km offset. The end date was defined as (in order of preference) the date of end of the seismicity/monitoring (quality "A"), or the date of end of the project plus 1 year (quality "B"), or no date limit was assigned (quality "C"). The 1-year time offset from the time of the end of the project was selected based on reports of time delays between injections and seismicity of days to several months (Foulger *et al.*, 2016).

- Mine flooding: The same criteria as described above for "Mining and solution mining" were applied.

- Research: Based on reports of seismicity taking place within a few metres from the injection point (Guglielmi *et al.*, 2015; Foulger *et al.*, 2016) a 1-km radius with the standard 5-km offset was adopted. Some larger values (2-4 km) have been observed for other projects (*e.g.*, Matsushiro, as reported by Foulger *et al.*, 2016, or Rangely, Colorado, as reported in HiQuake). However, it was felt that increasing the radius of

influence based only on some cases was overconservative under this category. The end date was defined as (in order of preference) the date of end of the seismicity/monitoring (quality “A”), or the date of end of the project (quality “B”), or no date limit was assigned (quality “C”). No time offset was adopted due to the nature of the activity.

- Natural gas storage: From Figure 1 of Cesca *et al.* (2014), which suggests seismicity up to a 25-km radius, and Figure 17 of Juez-Larré *et al.* (2016), which suggests a maximum length of the Norg reservoir (the largest of the three mentioned in the report by Foulger *et al.*, 2016) of around 20 km, a 25-km radius was adopted, together with the usual 5-km offset. The end date was defined as (in order of preference) the date of end of the seismicity/monitoring (quality “A”), or no date limit was assigned (quality “C”). The three entries of HiQuake that fall under this category correspond to the Castor project, reported by Foulger *et al.* (2016) to have been still causing seismicity in 2016 even if the project appears to have been discontinued, and to the Norg and Gripskerk reservoirs in the Netherlands, for which no dates are specified (discarding two entries with observed maximum magnitudes of 0.7 and 1.5). As it is understood that seismicity is still ongoing in all three, the end of the project was not used to define the end date of entries corresponding to this category.

- CO2 for oil recovery: The same criteria as described above for “Hydrocarbons” were applied.

- Carbon Capture and Storage (CCS): The only two entries associated to this category were eliminated during the initial purge, as the maximum reported observed magnitudes were 1.26 and 1.7.

- Nuclear explosions: All 20 cases of seismicity induced by nuclear explosions reported in HiQuake have a date for the start of the “project”, but no other dates (except for two cases that include year of observation of the maximum magnitude earthquake). Based on all the cases described by Foulger *et al.* (2016), it is clear that seismicity that results from nuclear explosions starts occurring very soon after the explosion itself and seems to last from days to weeks. As the 20 cases correspond to nuclear explosions of the 1960s and 1970s and the period of interest for the present study starts in 2001, all these entries were removed from HiQuake for the purpose of the present work.

4.7.3 Using HiQuake to Flag Potentially Induced Earthquakes

Once the information contained in HiQuake was parsed, the resulting grids were used to flag potentially induced earthquakes in the merged catalogue. The epicentral coordinates of each event were used to access the grid and determine whether the earthquake fell in a cell marked as having potential induced seismicity or not. If positive, then the start and end dates of all activities contributing to the induced seismicity in the cell were compared against the time of occurrence of the earthquake under scrutiny. If the latter was contained in the former, the earthquake was marked as potentially induced.

Figures 4.39 through 4.42 show the cells marked as areas of potential induced seismicity resulting from the parsing procedure described above. The quality indicators shown in the colour scale refer to the best quality of start time, end time, radius and offset in the cell. This implies, for example, that if two different entries of HiQuake make reference to one cell, and one is assigned a start time quality of A and the other of B, quality A is shown in the plot, though the time windows need not overlap. As can be observed, the map on Figure 4.39, corresponding to start time, can be deemed as that of better overall quality from the four, while that on Figure 4.41, referring to radii, appears as the worst.

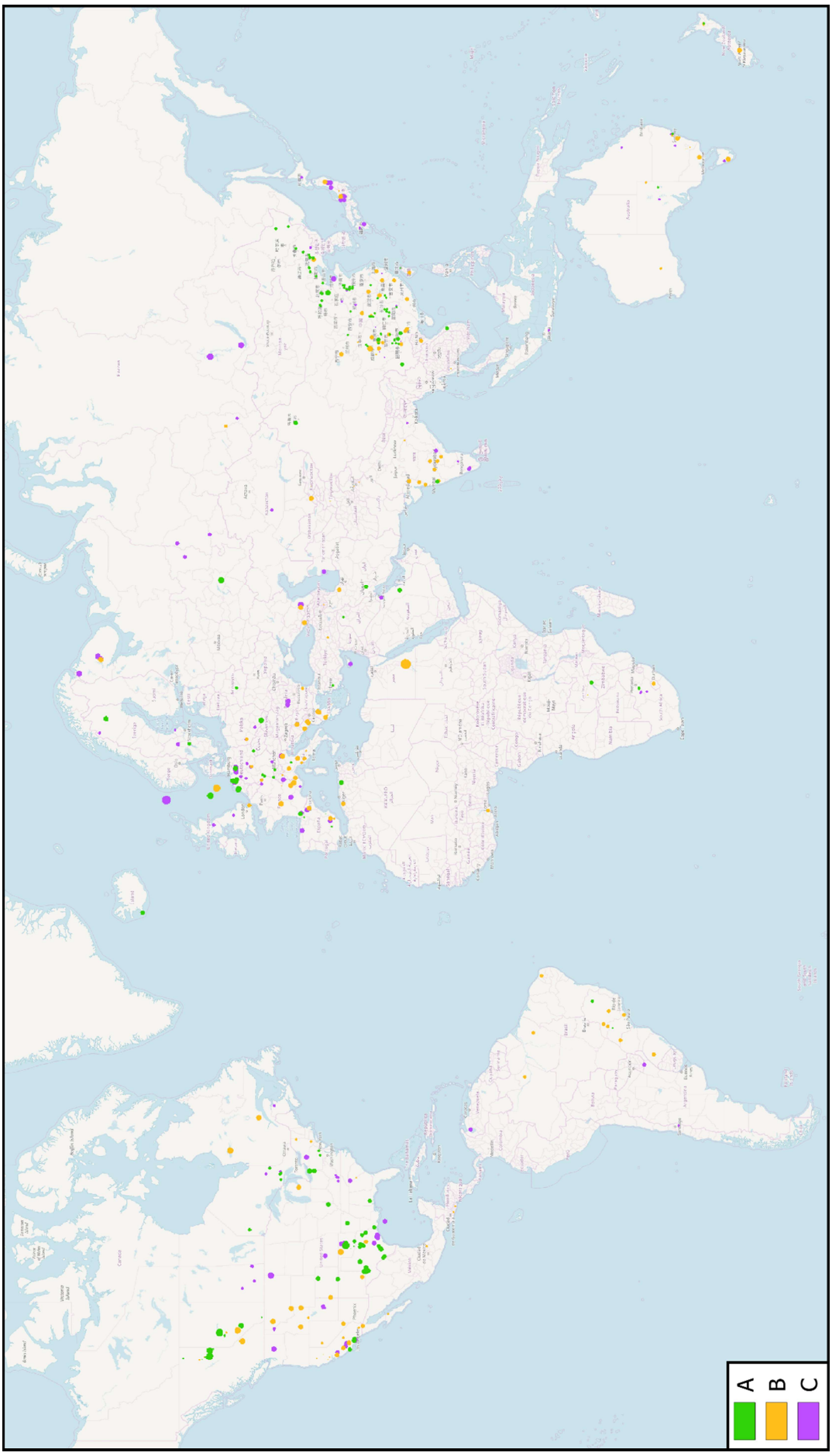

Figure 4.39. Best quality of start time in cell (A=best, C=worst).

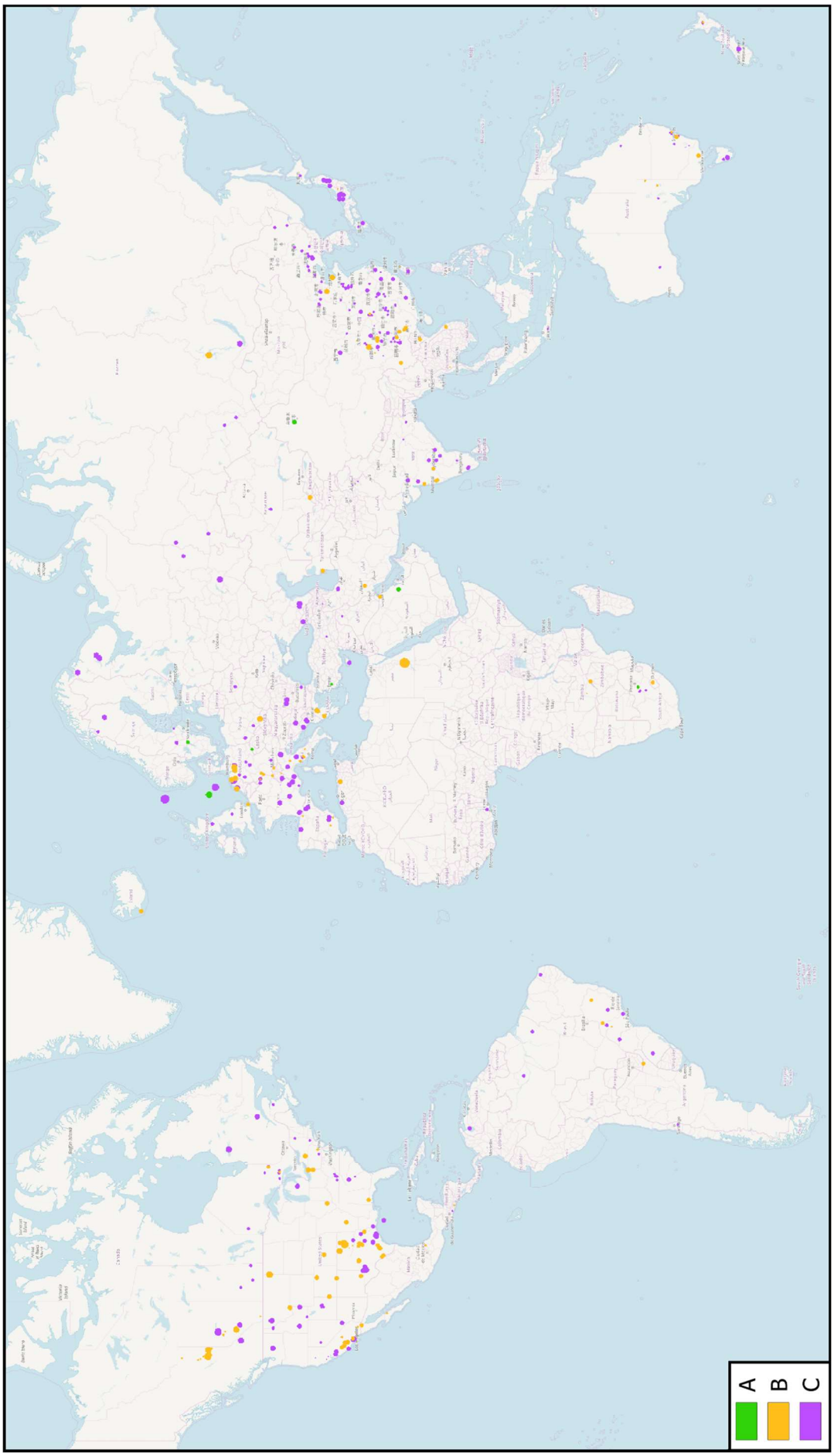

Figure 4.40. Best quality of end time in cell (A=best, C=worst).

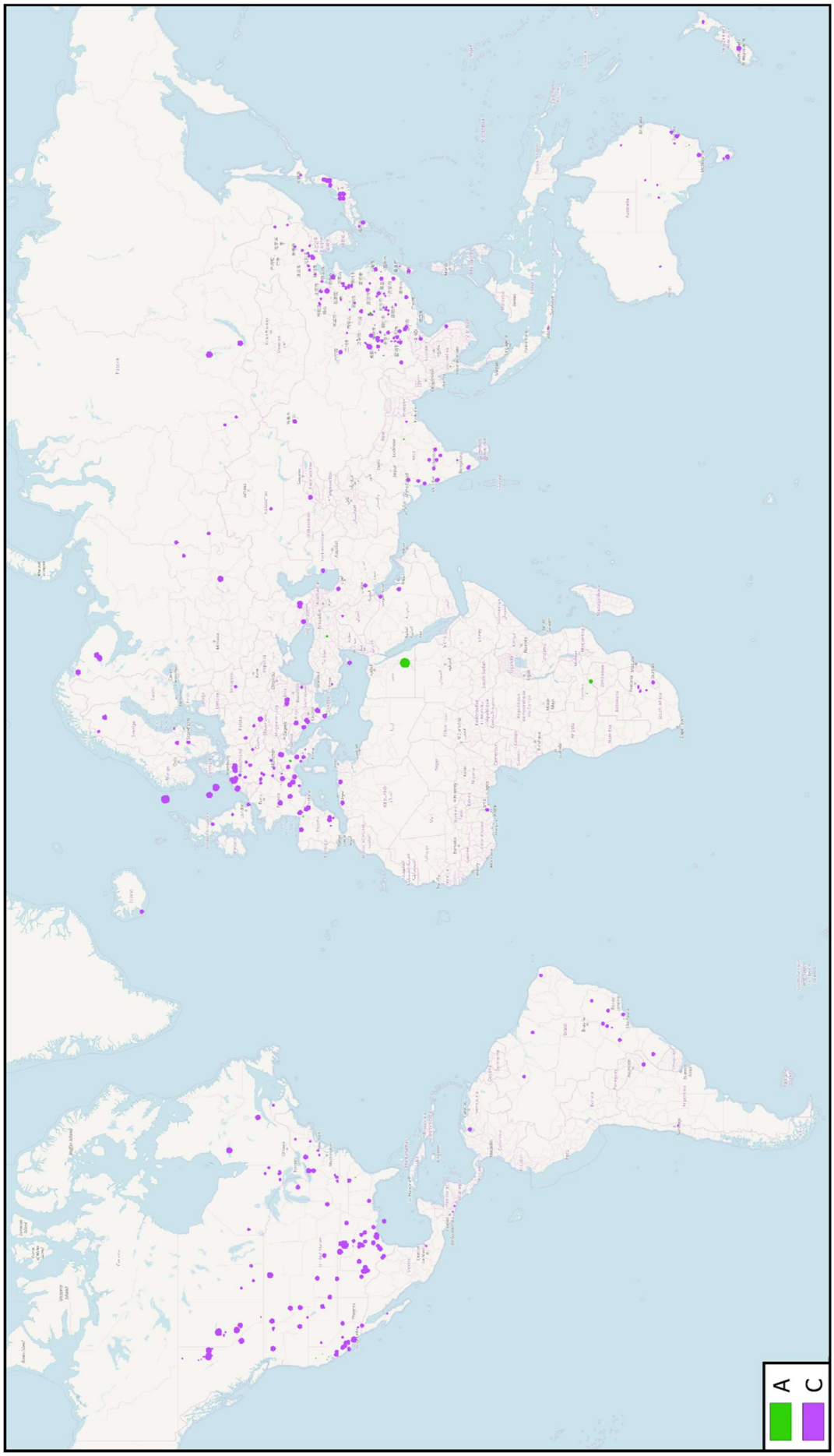

Figure 4.41. Best quality of radii in cell (A=best, C=worst).

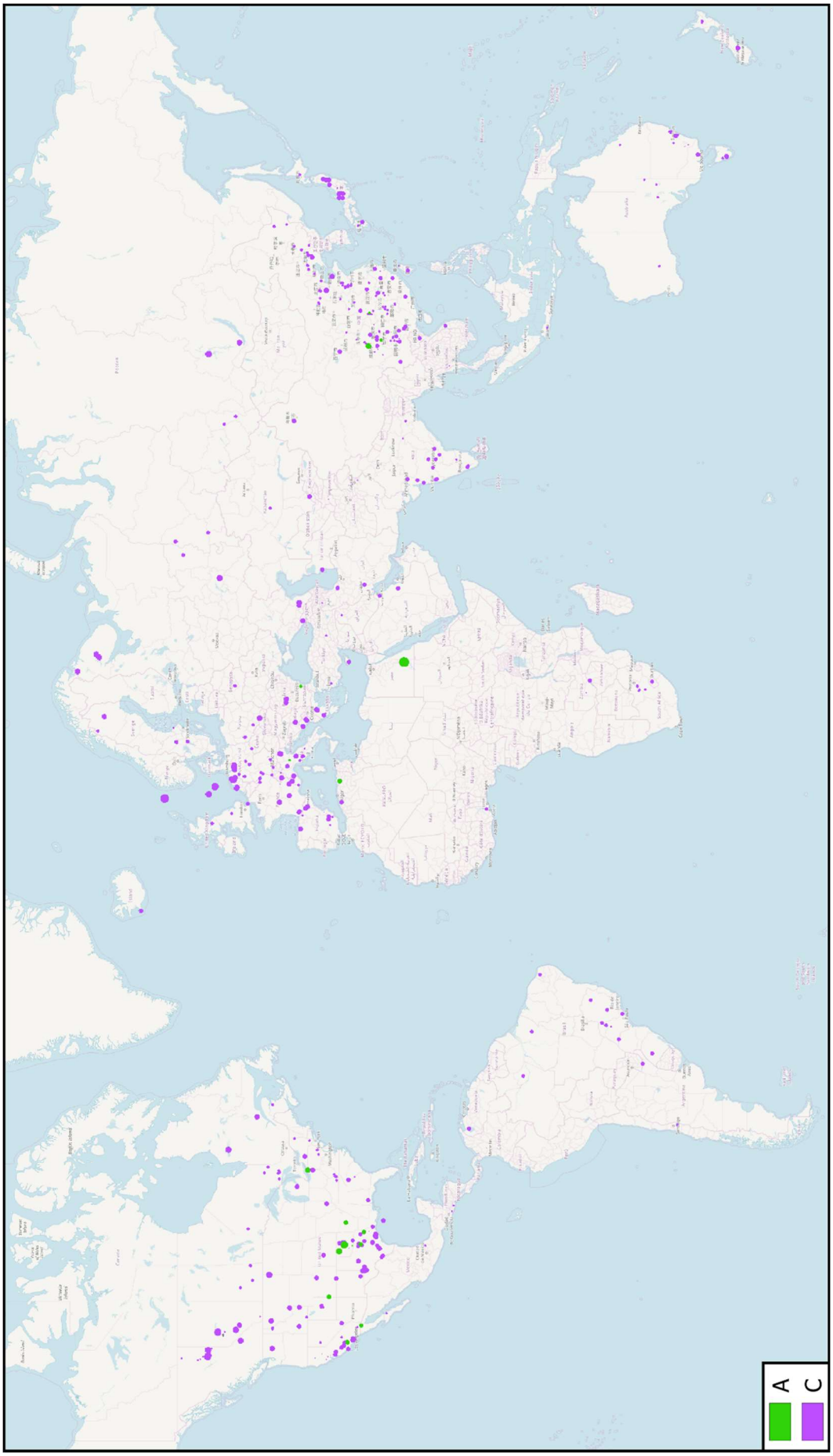

Figure 4.42. Best quality of offset in cell (A=best, C=worst).

4.8 Manual Modification of Events with Issues

While not being a technical change with respect to Version 1, the manual modification of outlier events was different in Version 2 due to the ISC Bulletin being updated in between. In Version 1, the magnitude of two events was manually modified due to the ISC Bulletin reporting a clearly wrong M_S of 9.8 for event ID 9514154, with other estimates being m_b=2.4 and M_L=4.0, and an M_S of 9.9 for event ID 610730892, with its other estimate being M_L=2.2. In the version of the Bulletin used for Version 2 of this work, the M_S of 9.8 has been eliminated for event ID 9514154, while event ID 610730892 has been removed from the Bulletin overall.

A manual inspection of the events in the merged catalogue with the largest values of moment magnitude led to the manual elimination of the earthquake with event ID 7453155, assigned a proxy **M**8.66 stemming from M_S=8.6 calculated by the International Data Centre (IDC) of the Comprehensive Nuclear-Test-Ban Treaty Organization (CTBTO), as the ISC Bulletin says of this event "*Amplitude readings of this event reported by the IDC are likely to be contaminated by the recordings of main Sumatera event at 00h58m*" and the IDC estimates are the only ones available for it. Declustering of a preliminary version of the merged catalogue had classified this event as the main shock of a cluster that includes earthquakes with moment magnitude values of up to 6.16.

5. RESULTS

5.1 General

This chapter presents the characteristics of the resulting catalogue of potentially damaging earthquakes and the in-between catalogues leading to it, as well as the statistical analysis stemming from its comparison against the Database of Damaging Small-to-Medium Magnitude Earthquakes (Nievas *et al.*, 2019). The procedure followed along this work is schematically shown in Figure 5.1.

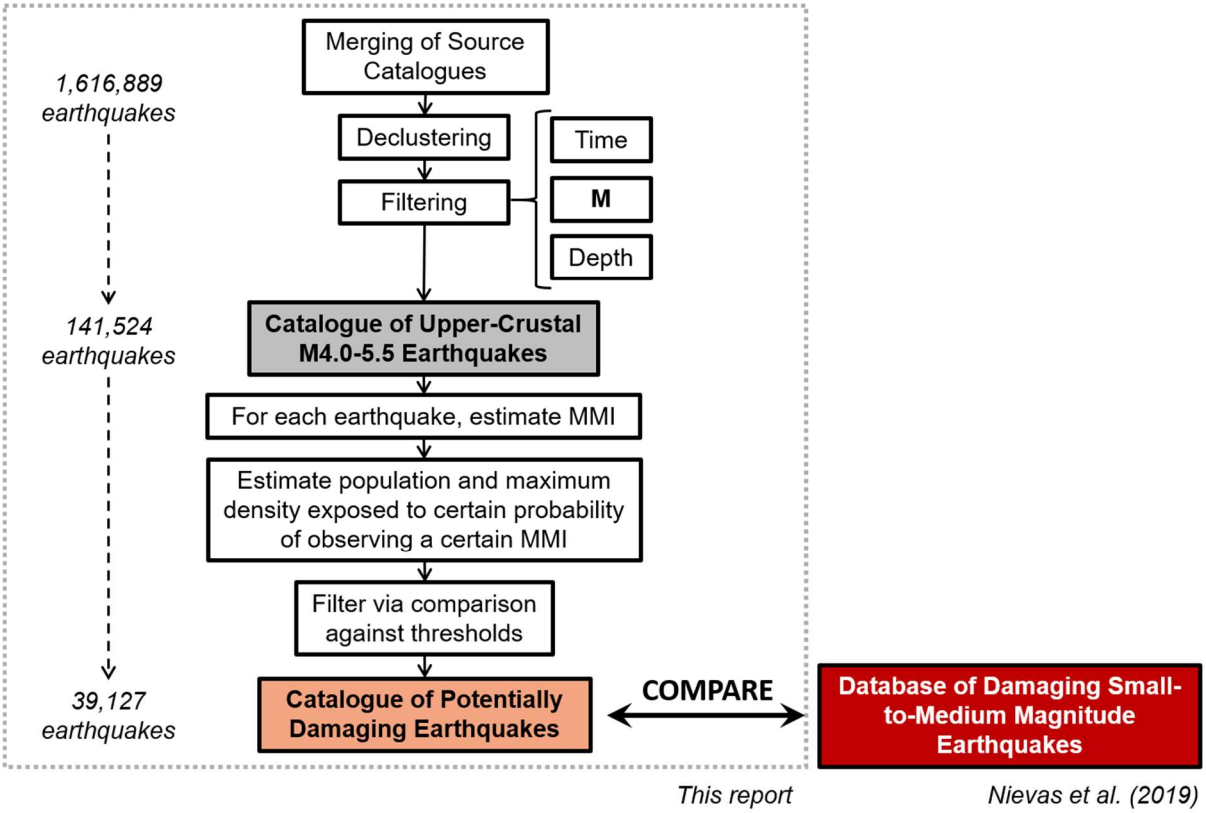

Figure 5.1. Schematic outline of the work.

5.2 Composition of the Resulting Catalogues

As shown in Table 5.1, the catalogue that resulted from merging the WPG16v3c catalogue with the events added from the ISC Bulletin contains 1,616,889 earthquakes for the period 1st January 1998 – 31st August 2018, whose distribution in time, by magnitude and in depth are depicted in Figures 5.2 through 5.4. Of these, 1,239,989 (76.7%) were added from the ISC Bulletin, while the remaining 376,900 (23.3%) were retrieved from the WPG16v3c catalogue. These proportions change significantly after filtering by magnitude and depth for the final time period of interest (1st January 2001 – 31st December 2015): of the 141,524 earthquakes that pass the filtering, 72,955 (51.5%) correspond to the set retrieved from the WPG16v3c catalogue and 68,569 (48.5%) were added from the ISC Bulletin. This change is due to the set of earthquakes added from the ISC Bulletin containing a very large proportion of smaller-magnitude events, which is consistent with the fact that Weatherill *et*

al. (2016) only considered a reduced number of well-established agencies and magnitude scales for the compilation of their catalogue, and many events of smaller magnitudes get reported only by local agencies and using magnitude scales such as M_d or M_L. The final catalogue of potentially damaging earthquakes, that is, the catalogue that results from applying the exposure criterion explained in Section 4.6, contains 39,127 earthquakes with similar proportions: 18,752 (47.9%) come from the WPG16v3c catalogue, while 20,375 (52.1%) were added from the ISC Bulletin.

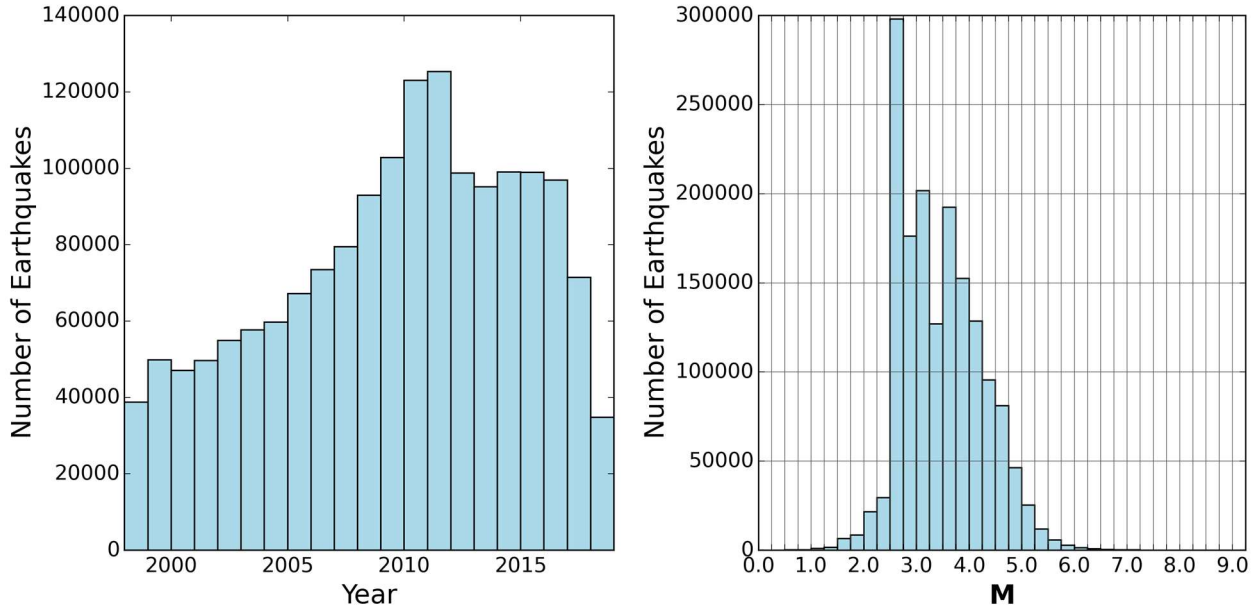

Figure 5.2. Distribution in time (left) and by magnitude (right) of the 1,616,889 earthquakes that make up the merged catalogue. Note that the bar for the year 2018 only covers 8 months.

Table 5.1. Composition of the resulting earthquake catalogues. Columns labelled “WPG16v3c” and “ISC Bulletin” indicate the number of earthquakes retrieved from the WPG16v3c catalogue and the ISC Bulletin, respectively.

Catalogue	Time Period	M range	Depth	Exposure	WPG16v3c	ISC Bulletin	Total
Merged	01/01/1998-31/08/2018	Any	Any	Any	376,900	1,239,989	1,616,889
Upper Crustal	01/01/2001-31/12/2015	3.95≤**M**<5.55	Section 4.5.2	Any	72,955	68,569	141,524
Potentially Damaging	01/01/2001-31/12/2015	3.95≤**M**<5.55	Section 4.5.2	Section 4.6	18,752	20,375	39,127

The proportions of earthquakes that were retrieved from the WPG16v3c catalogue and added from the ISC Bulletin is quite different from that of Version 1, for which the final catalogue of potentially damaging earthquakes contained 35,654 earthquakes for the period 1st July 1999 – 30th June 2014, of which 32,035 (89.8%) came from the WPG16v3c catalogue and the remaining 3,619 (10.2%) were added. Apart from potential variations in the temporal distribution of earthquakes and their detectability for the two 15-year periods considered in each version (see plot on the left of Figure 5.2), the difference is likely due to two main reasons. Firstly, the fact that earthquakes with magnitude estimates only in terms of M_d or M_L were included in Version 2 and discarded in Version 1. Earthquakes with estimates in only these scales were not considered by Weatherill *et al.* (2016) for the compilation of the WPG16v3c catalogue. Consequently, all of these have been added from the ISC Bulletin. Secondly, the use of exponential conversion equations for the compilation of the WPG16v3b catalogue used in Version 1 led to a large number of earthquakes being awarded magnitudes above **M**4.0 due to the saturation of such equations for low values of m_b and M_s. Many of these earthquakes have been assigned magnitudes smaller than **M**4.0 and, thus, outside of the range of interest in the WPG16v3c catalogue used for Version 2.

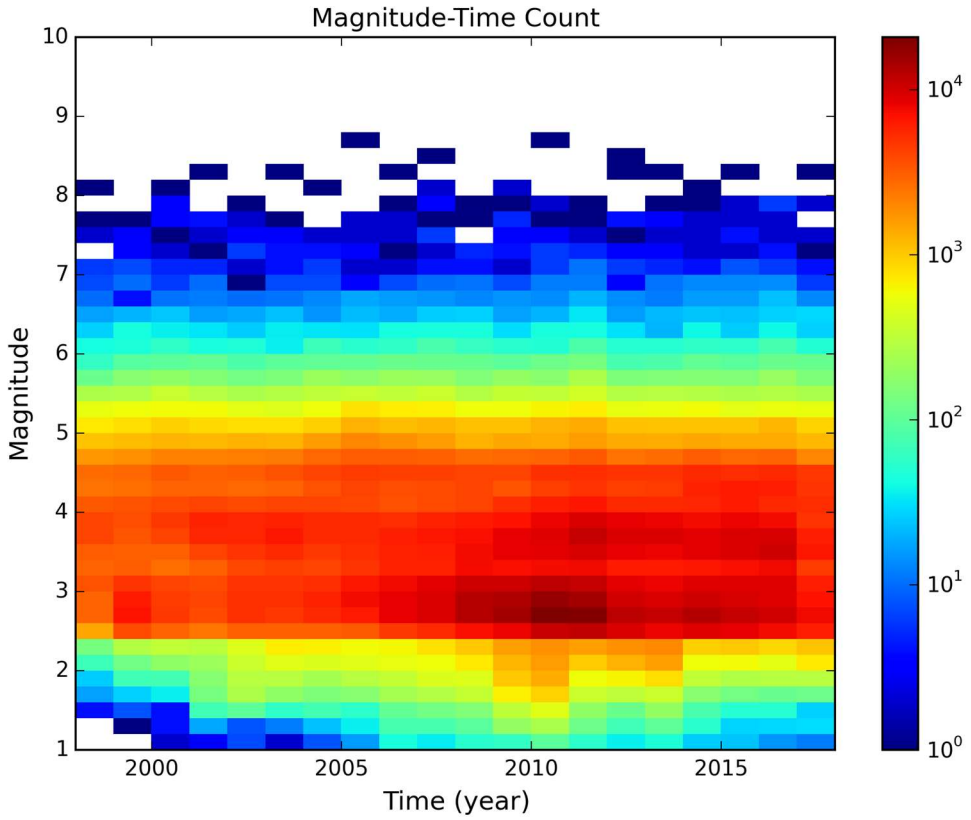

Figure 5.3. Density of earthquake magnitudes (colour scale) with time for the merged catalogue spanning from 1st January 1998 through 31st August 2018 (year 2018 not shown).

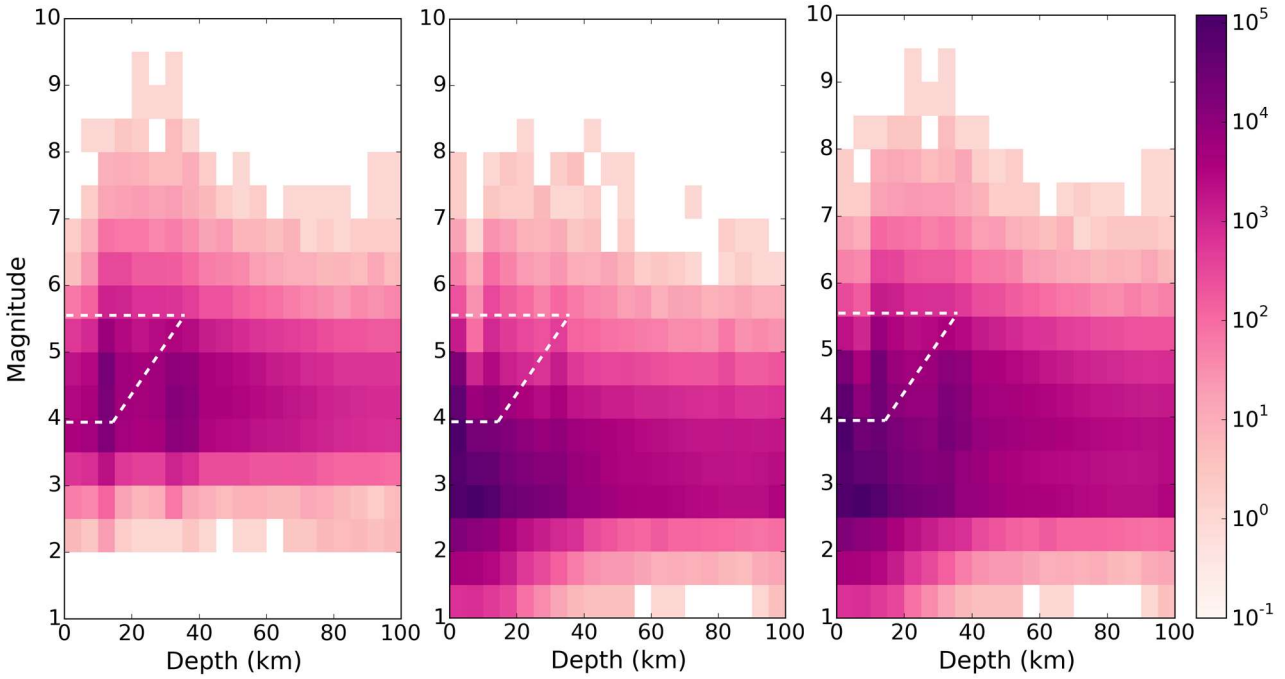

Figure 5.4. Moment magnitude-depth distribution of the earthquakes retrieved from the WPG16v3c catalogue (left), the ISC Bulletin (centre), and of the resulting merged catalogue (right) spanning from 1st January 1998 through 31st August 2018. Number of earthquakes per cell according to colour scale (interpret 10E-1 as zero, logarithmic colour scale used). The horizontal axis has been truncated to 100 km but values go up to 832 km. White dashed lines enclose the magnitude-depth combinations considered in the present work (notwithstanding uncertainty).

As explained in Section 2.1, an algorithm designed to identify potential duplicate entries of the same earthquake was run over the merged catalogue. Using relaxed temporal and spatial windows of 120 seconds and 130 km, respectively, 36,514 potential cases of duplicate events were identified and further analysed with the algorithm, details of which are described in Section 2.7.7 of the report for Version 1. Only in 264 cases the algorithm concluded that the two events were duplicates of the same earthquake and one of them was eliminated. It is noted that the 1,616,889 earthquakes mentioned herein as forming the merged catalogue are those left after the identification of these duplicates.

The map on Figure 5.5 depicts the 141,524 earthquakes that make up the catalogue of upper-crustal earthquakes, that is, those earthquakes from the merged catalogue that occurred in the period 1st January 2001 – 31st December 2015, with magnitudes in the range $4.0 \leq \mathbf{M} \leq 5.5$ (implemented as $3.95 \leq \mathbf{M} < 5.55$) and depths that satisfy the maximum depth criteria defined in Section 4.5.2. Epicentres depicted in grey are those that are later discarded due to not complying with the exposure criteria defined in Section 4.6, while those depicted in salmon and red are the 39,127 events that make up the final catalogue of potentially damaging earthquakes.

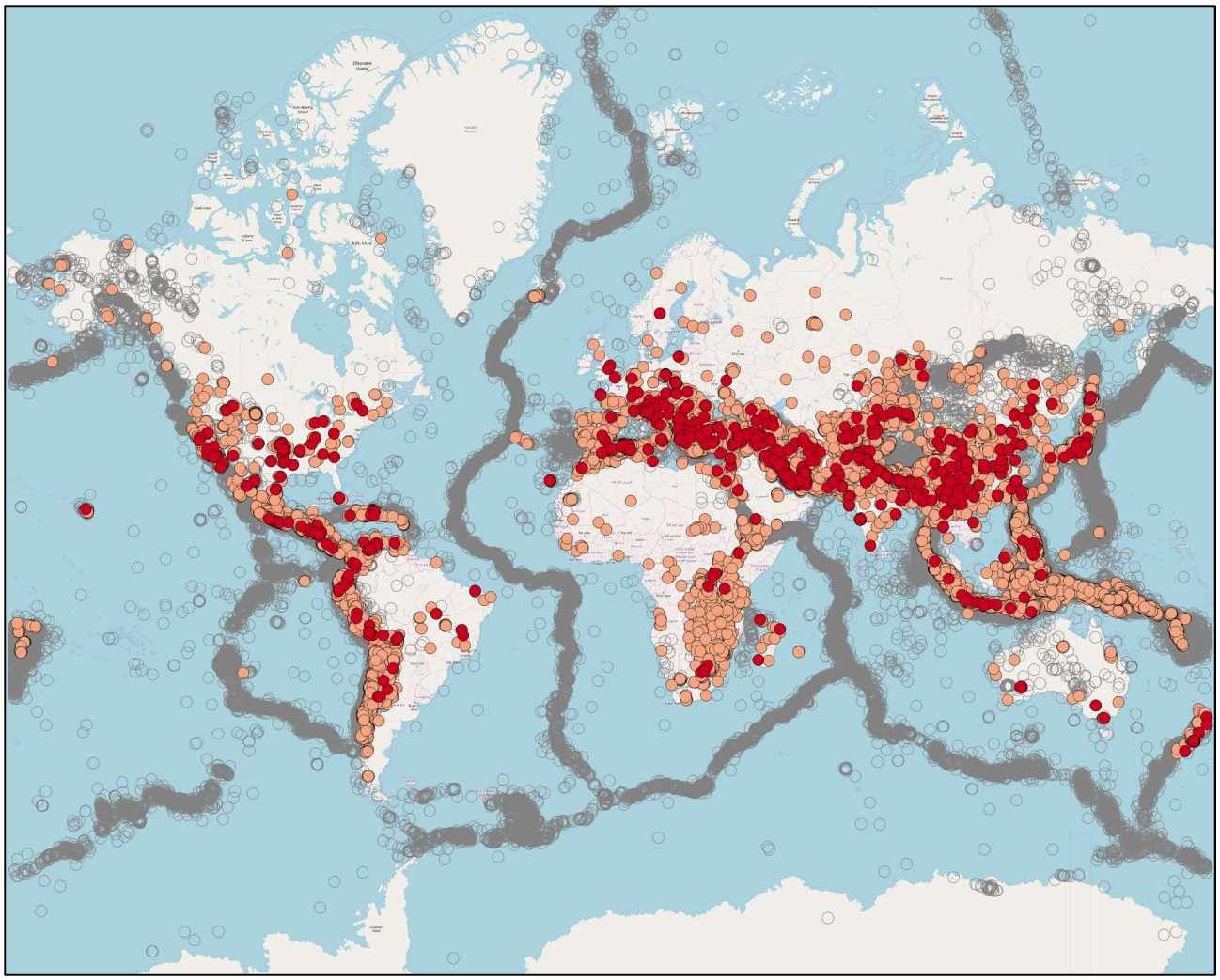

Figure 5.5. Catalogue of upper-crustal earthquakes with magnitudes in the range 3.95≤**M**<5.55 (light grey; 141,524 events), catalogue of potentially damaging earthquakes (salmon; 39,127 events) and damaging earthquakes within the latter (red; 740 earthquakes).

The plots in Figure 5.6 depict the distribution in time and by magnitude of the catalogue of potentially damaging earthquakes, classified in terms of main shocks (MS), aftershocks (AS) and foreshocks (FS). As can be observed, the number of earthquakes appears to increase significantly in time. This is most likely reflecting the improvements that have taken place in the seismological networks worldwide in the last decade or so, as well as potentially the increase in the number of agencies contributing to the ISC Bulletin, rather than indicating an actual increase in seismicity. The moment magnitude distribution of the catalogue appears to follow, in general terms, the Gutenberg-Richter relationship (Gutenberg & Richter, 1944).

As explained in Section 4.7, two methodologies were used to identify human-induced events within the catalogue of potentially damaging earthquakes. Table 5.2 compares the results obtained with each of the two, showing that the resulting number of induced earthquakes is significantly larger when HiQuake is used. This comes as no surprise when considering that a large proportion of the earthquakes classified as induced based on HiQuake correspond to low-quality spatial and temporal windows, as shown in Figure 5.7. As explained in Section 4.7.2, if the dates of start/end of seismicity/monitoring were reported in HiQuake, a quality flag of “A” was assigned, corresponding to the highest reliability, while an intermediate

quality “B” was assigned if the start and end dates could be inferred by the start/end date of the project in conjunction with assumptions regarding the temporal extent of seismicity associated to the specific anthropogenic activity, and a low quality (“C”) was assigned if no start or end dates could be specified, in which case the causative activity was assumed to have been always present. Similarly, quality flags of “A” (high) or “C” (low) were assigned to the radius and offset used to define the spatial influence of an activity, depending on whether they were calculated from information available in HiQuake or if pre-defined values depending on the kind of causative activity were adopted.

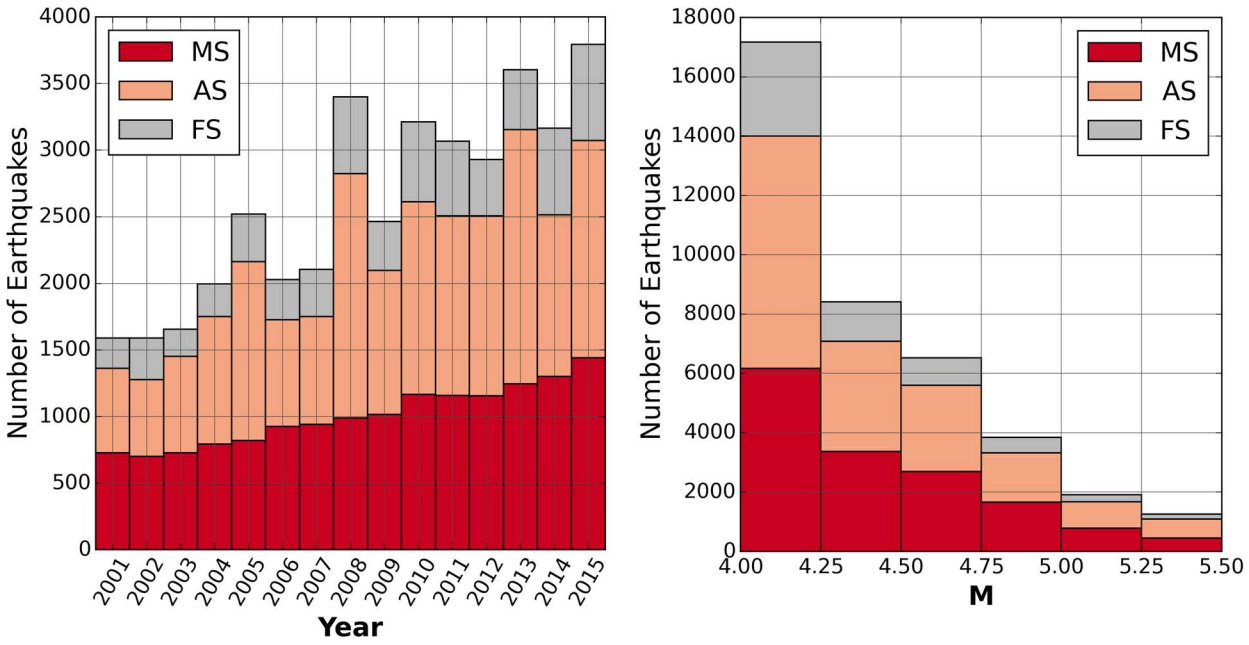

Figure 5.6. Distribution in time (left) and by magnitude (right) of the 39,127 earthquakes that make up the catalogue of potentially damaging **M**4.0-5.5 upper-crustal earthquakes. Colour scale indicates main shocks (MS), aftershocks (AS) and foreshocks (FS).

The pie plots in Figure 5.7 depict the best quality indicators for each earthquake classified as induced by using HiQuake. This means that if the earthquake was classified as induced due to two or more different entries of HiQuake, the highest quality of the two is shown. It is clear from the plots that most spatial areas of influence were defined by adopting generic values for different kinds of causative activities and not from actual observations. In this sense, the quality of the starting time appears to be much better. The second plot from the left indicates that around 60% of the entries of HiQuake were not assigned an end time are were therefore assumed to still be able to induce seismicity.

The plots of Figure 5.7 are very similar when only the 2,436 earthquakes that are classified as induced based on HiQuake but not on the ISC Bulletin/WPG16v3c are considered. However, when only considering the 203 cases flagged as induced with the two methodologies, the proportions of bad-quality indicators is much larger, with only one earthquake featuring A quality in any indicator (the start date in this case), around 90% featuring B quality for the start date, all earthquakes featuring C quality in indicators for radii and offsets, and most activities not having a defined finishing time of influence. The corollary

of this observation is that agreement between the two methodologies to flag induced earthquakes cannot be deemed a synonym of data quality within HiQuake.

Table 5.2. Classification of the catalogue of potentially damaging earthquakes into induced and not induced, according to the two methodologies used.

CASE		Based on HiQuake		Total	Total Percentage
		Induced	Not Induced		
Based on ISC Bulletin and WPG16v3c	Induced	203	138	341	0.9%
	Not Induced	2,436	36,350	38,786	99.1%
Total		2,639	36,488	39,127	100.0%
Total Percentage		6.7%	93.3%	100.0%	

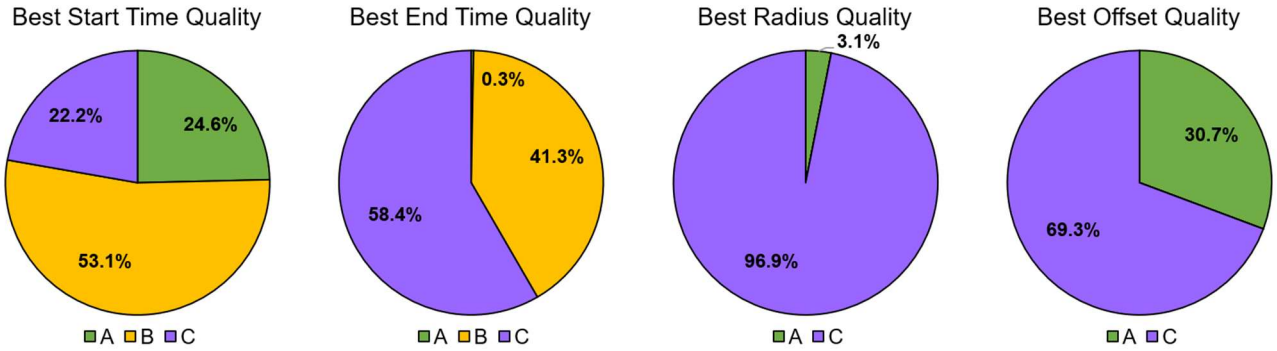

Figure 5.7. Best quality indicator for a particular earthquake, ranging from highest (A) to lowest (C) quality, for the 2,639 cases classified as induced with the methodology based on HiQuake.

The maps in Figures 5.8 and 5.9 show the earthquakes classified as induced based on the ISC Bulletin/WPG16v3c and HiQuake, respectively, while those in Figures 5.10 through 5.12 compare results obtained with the two methodologies: the 203 earthquakes classified as induced by both (Figure 5.10), the 138 earthquakes classified as induced based on the ISC Bulletin/WPG16v3 but not on HiQuake, and the 2,436 earthquakes classified as induced based on HiQuake but not on the ISC Bulletin/WPG16v3. The 192 earthquakes in Japan classified as induced based on HiQuake might call the reader's attention in Figures 5.9 and 5.12. These represent around 20% of all the potentially damaging earthquakes located in Japan, and are not flagged as induced based on the ISC Bulletin/WPG16v3c. It is possible that the HiQuake-based net may have been too wide in this case, especially when considering that the end time quality and the spatial quality of the related entries is C in all Japan cases (see Figures 4.40 through 4.42), and the start time qualities are split between B and C. Something similar may occur in the area around California (not including Mexicali, Mexico, close to the frontier with he USA), where 95 earthquakes (out of slightly over 300) are identified as induced based on HiQuake, only one of which is identified as induced based on the ISC Bulletin/WPG16v3c as well. However, the quality of the spatial and temporal windows seem in this case slightly better than those of Japan (see Figures 4.39 through 4.42).

Figure 5.8. The 341 earthquakes classified as induced using flags from the ISC Bulletin and the WPG16v3c catalogue.

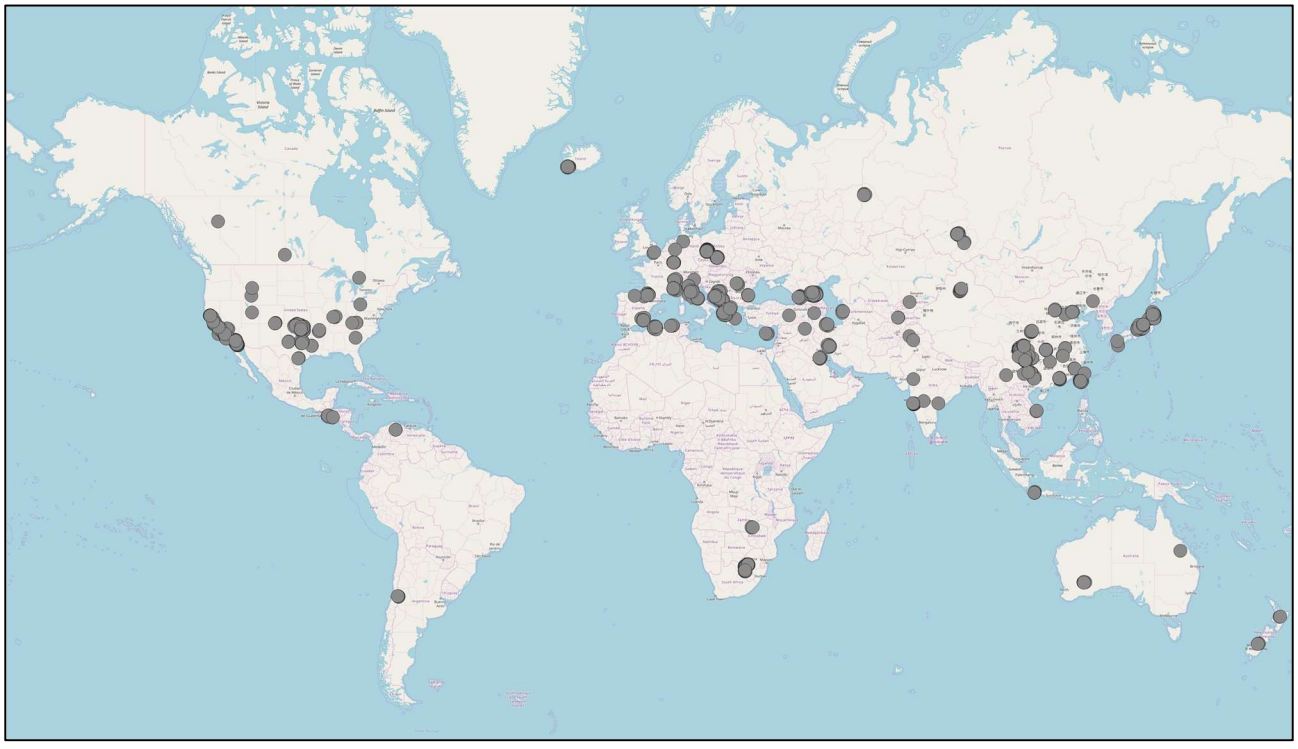

Figure 5.9. The 2,639 earthquakes classified as induced using HiQuake and the methodology described in Section 4.7.

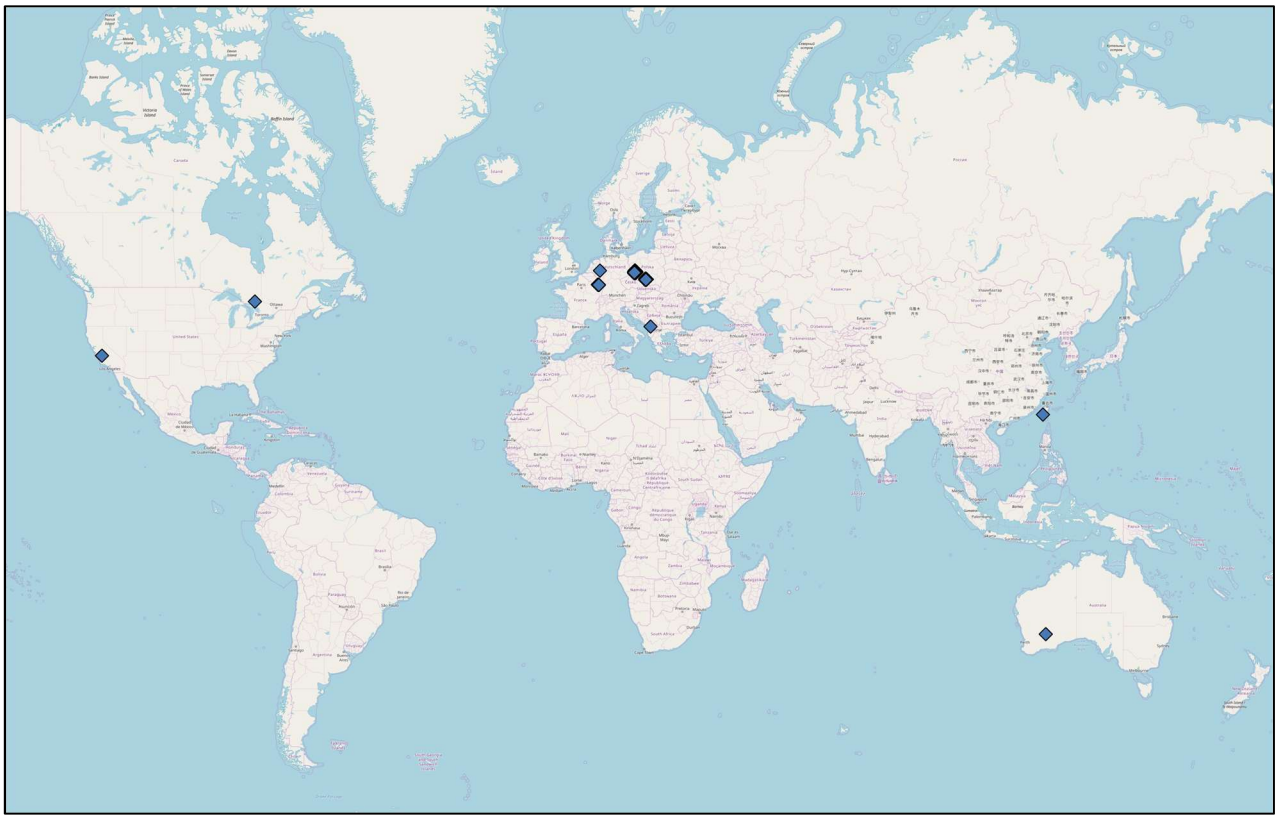

Figure 5.10. The 203 earthquakes classified as induced using both flags from the ISC Bulletin and the WPG16v3c catalogue as well as HiQuake and the methodology described in Section 4.7.

Figure 5.11. The 138 earthquakes classified as induced using flags from the ISC Bulletin and the WPG16v3c catalogue but as not induced when using HiQuake and the methodology described in Section 4.7.

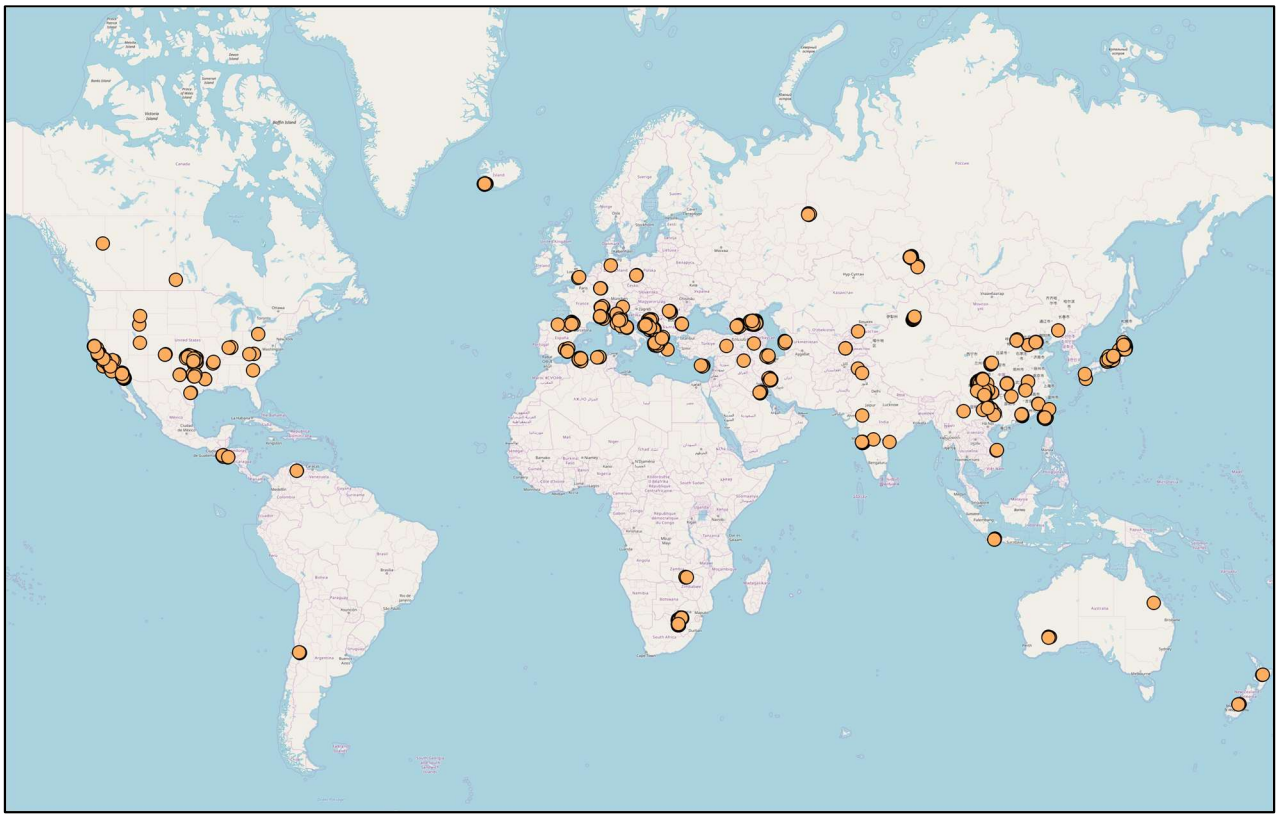

Figure 5.12. The 2,436 earthquakes classified as not induced using flags from the ISC Bulletin and the WPG16v3c catalogue but classified as induced when using HiQuake and the methodology described in Section 4.7.

While the aforementioned considerations suggest that the number of earthquakes identified as anthropogenically induced based on HiQuake might be an overestimation, the fact that no earthquake in Oklahoma, USA, has been identified as induced based on the ISC Bulletin/WPG16v3c implies the existence of an underestimation when flagging induced events based only on the latter. Based on HiQuake, 211 earthquakes in Oklahoma have been classified as induced. For a few of these the ISC Bulletin lists suggested bibliography that makes reference to induction by injection, but the bibliography from the ISC Bulletin is not being used to flag induced events, due to potential misinterpretation issues that could arise from it (*i.e.*, publication titles containing some of the keywords but not making reference to the earthquake being anthropogenically induced). However, the link between wastewater injection and seismic activity in Oklahoma has long been recognised (*e.g.*, Goebel *et al.*, 2017; Petersen *et al.*, 2016). It is of interest to note as well that further 122 potentially damaging earthquakes located within Oklahoma have not been flagged as induced based on HiQuake either. Comparison of these earthquakes against the map shown by Petersen *et al.* (2016) and Figures 4.39 through 4.42 suggests that either HiQuake[5] is missing entries associated with Oklahoma or the spatial influence of those existing in HiQuake is larger than modelled when parsing the database (Section 4.7.2).

[5] Version of November 2017.

5.3 Proportion of Damaging Earthquakes

As shown in Figure 5.1, the last stage of the study consisted in comparing the 39,127 events of the catalogue of potentially damaging earthquakes against the 996 ones contained in the Database of Damaging Small-to-Medium Magnitude Earthquakes. Only 740 of these 996 earthquakes were found in the catalogue (depicted in red in Figure 5.5). While the reasons for this to occur and the implications for the present study will be discussed and explained in detail in Section 5.4, the present section focuses on the 39,127 events from the catalogue and the subset of 740 damaging earthquakes within it. Table 5.3 and Figure 5.13 summarise the proportion that these represent of the total number of potentially damaging earthquakes. In Table 5.3, results are subclassified according to the kind of shock (all, mainshocks plus foreshocks, mainshocks, foreshocks, aftershocks) and time period. The variation of the results in time is apparent in the plot on the left of Figure 5.13, which clearly shows the increase in the proportion of damaging earthquakes obtained for the years 2013-2015. This is due to the increase in the number of damaging earthquakes identified from the year 2013 onward thanks to the contribution of data from the EID, as was shown earlier in Figure 4.6 (Section 4.3), and reflects the effect that data accessibility has on an analysis of this kind. The same is true for the overall number of potentially damaging earthquakes, which also increases in time, for the reasons discussed when presenting Figure 5.6 earlier. In other words, the increase in the number of earthquakes and proportion of damaging earthquakes in time depicted in Figures 5.13 and 5.14 are most likely reflecting the variability of the completeness of the sets and not an actual increase in the quantities.

Table 5.3. Number and proportion of damaging events within the catalogue of potentially damaging events for the 15 years that this study spans (2001-2015) and subdivided into 2001-2012 and 2013-2015. MS, FS and AS refer to main shocks, foreshocks and aftershocks, respectively.

Shocks	2001-2015			2001-2012			2013-2015		
	Total	Damaging	%	Total	Damaging	%	Total	Damaging	%
All	39,127	740	1.9%	28,563	283	1.0%	10,564	457	4.3%
MS + FS	21,475	584	2.7%	15,662	224	1.4%	5,813	360	6.2%
MS	15,123	498	3.3%	11,133	202	1.8%	3,990	296	7.4%
FS	6,352	86	1.4%	4,529	22	0.5%	1,823	64	3.5%
AS	17,652	156	0.9%	12,901	59	0.5%	4,751	97	2.0%

As shown in Table 5.3, while the proportion of damaging earthquakes is 1.9% for the complete 15-year period under study, it rises up to 4.3% for the period 2013-2015 and drops to just 1.0% in 2001-2012. In sight of the considerations regarding completeness of both the catalogue of potentially damaging earthquakes and the Database of Damaging Small-to-Medium Magnitude Earthquakes made earlier, it is thus likely that 4.3% (or even more) is a more realistic annual percentage of damaging earthquakes than 1.9%. As shown in Figure 5.14, this tendency can be observed in main shocks as well as in foreshocks and aftershocks.

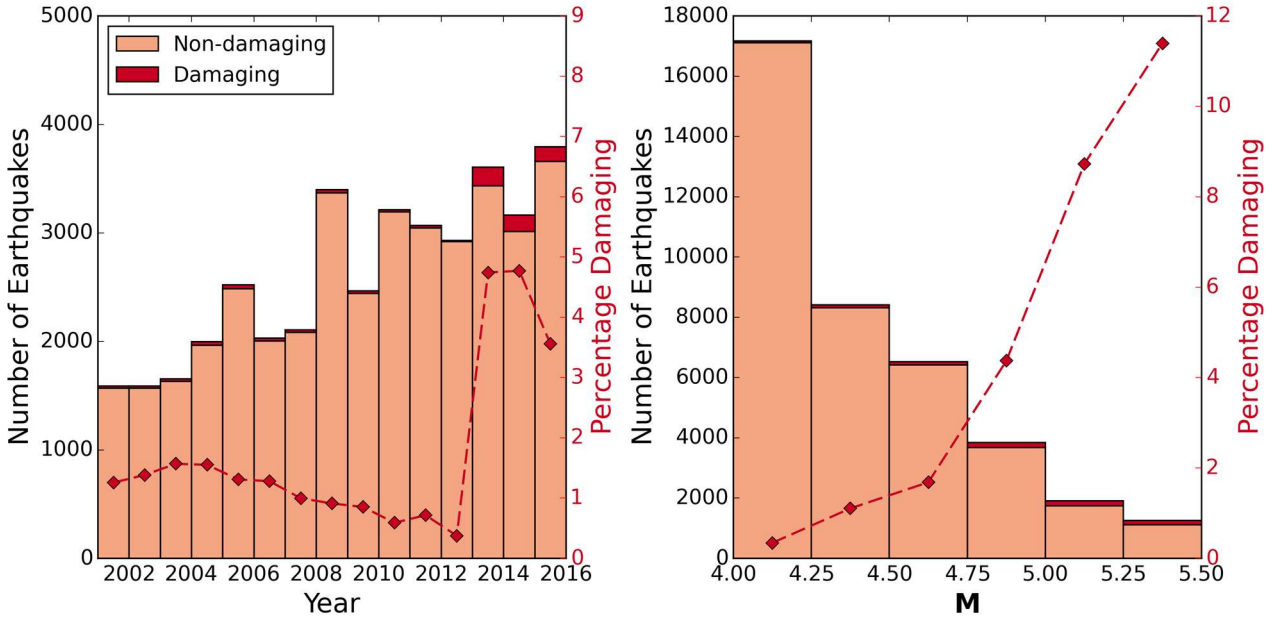

Figure 5.13. Number and proportion of damaging earthquakes in time (left) and by moment magnitude (right). Red rhombuses indicate the percentage of damaging events per year (left) and magnitude bin (right).

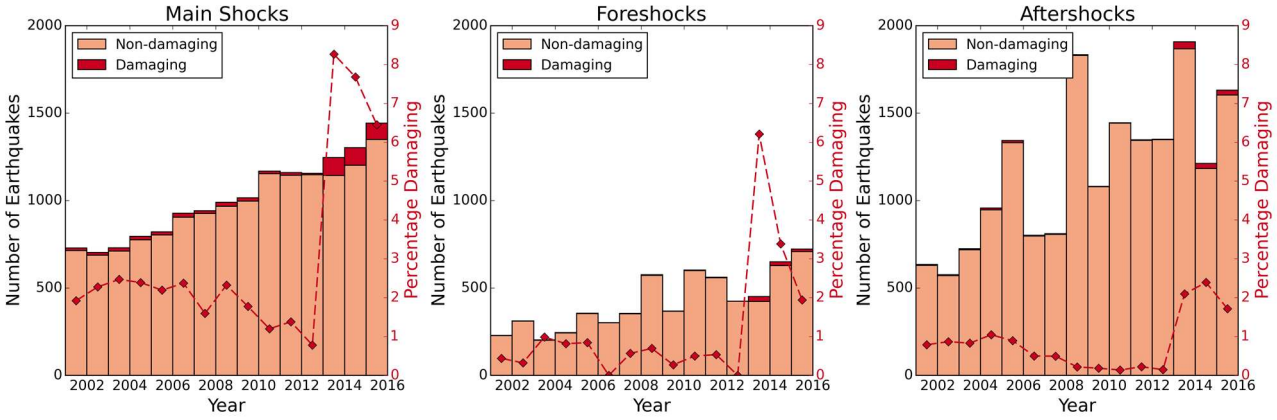

Figure 5.14. Number and proportion of damaging earthquakes in time for earthquakes classified as main shocks (left), foreshocks (centre) and aftershocks (right). Red rhombuses indicate the percentage of damaging events per year.

The potential implications of the effects of completeness observed herein on the total number of casualties or damaged buildings per year is difficult to assess. For example, around half (48.8%) of the total 43 upper-bound deaths reported for 2013-2015 correspond to events retrieved from the EID, as depicted in Figure 5.15. However, it is unlikely that the 88 upper-bound deaths reported for the year 2002 alone 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. Such inferences cannot be made, due mainly to (i) the large variability observed for the total number of deaths and injuries per year, (ii) the fact that the overall number of casualties do not peak for the years 2013-2015 as does the number of earthquakes, and (iii) the fact that it is unlikely that earthquakes with extensive consequences that occurred before 2013 had passed unnoticed. All this suggests that the unreported earthquakes from 2001-2012 probably did

not have an extreme impact in the overall annual casualties, despite their significance in terms of proportion of damaging events. The situation is more complicated for the number of damaged and destroyed buildings, as so many earthquakes are associated only with qualitative descriptions, particularly in the case of those retrieved from the EID (as shown in Figure 5.16; only for the year 2015 are there numbers of damaged/destroyed buildings available from the EID, 2013-2014 contain qualitative descriptions). While it is still unlikely in this case that earthquakes with extensive consequences that occurred before 2013 passed unnoticed, the fact that buildings classified as damaged can range from anything from having minor non-structural cracks to serious damage[6] does not allow to draw conclusions on the potential number of unreported damaged buildings. In other words, large numbers of buildings with minor damage may have passed unreported in the period before 2013.

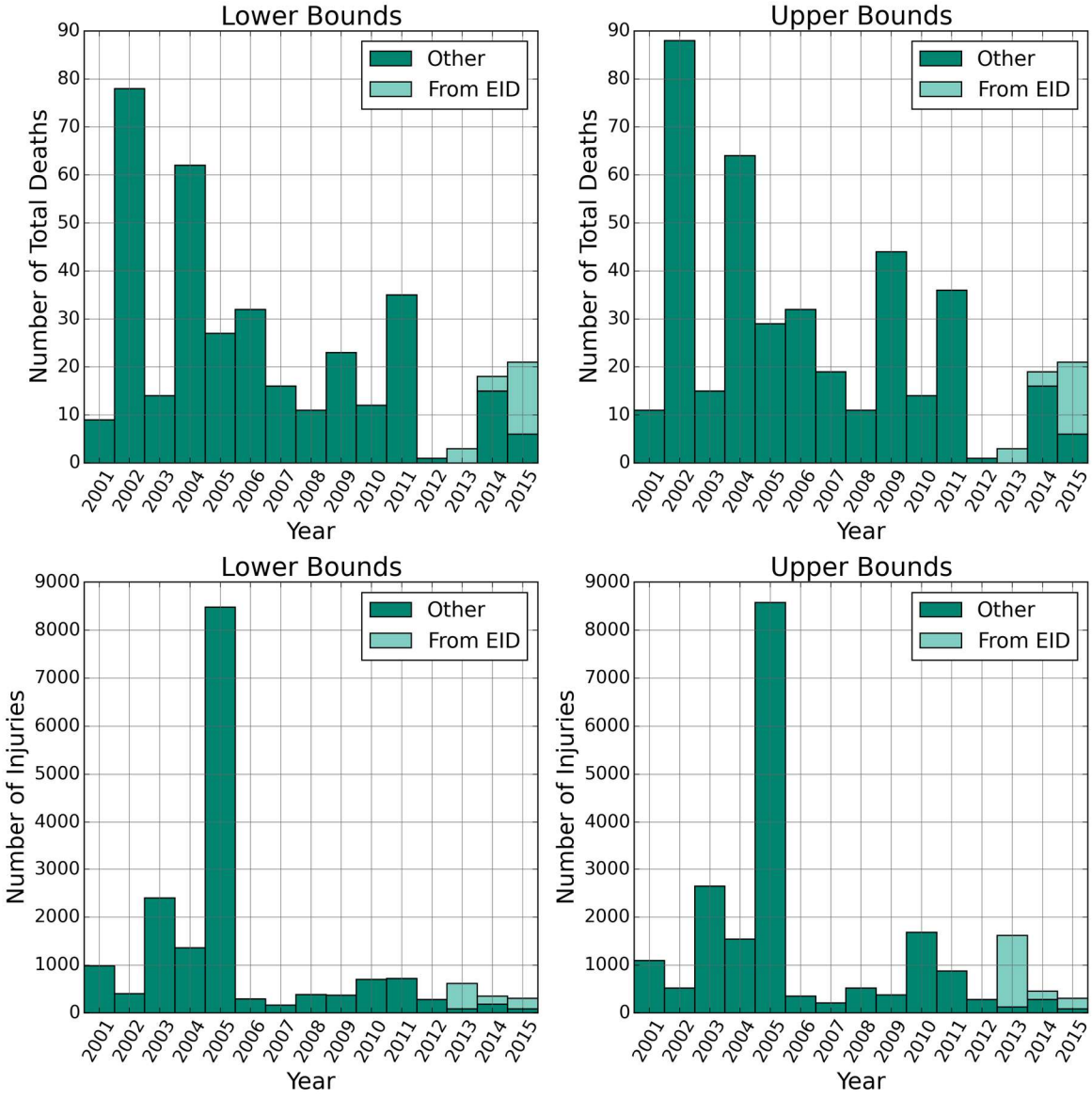

Figure 5.15. Number of total deaths (top) and injuries (bottom) for the 996 events that make up the Database of Damaging Small-to-Medium Magnitude Earthquakes classified according to whether they were retrieved from the EID or other sources. Lower (left) and upper (right) bounds shown.

[6] Seriously damaged buildings may still be counted as damaged and not destroyed either due to lack of details in the original sources or because the latter only specify that “some of” the damaged buildings were actually destroyed.

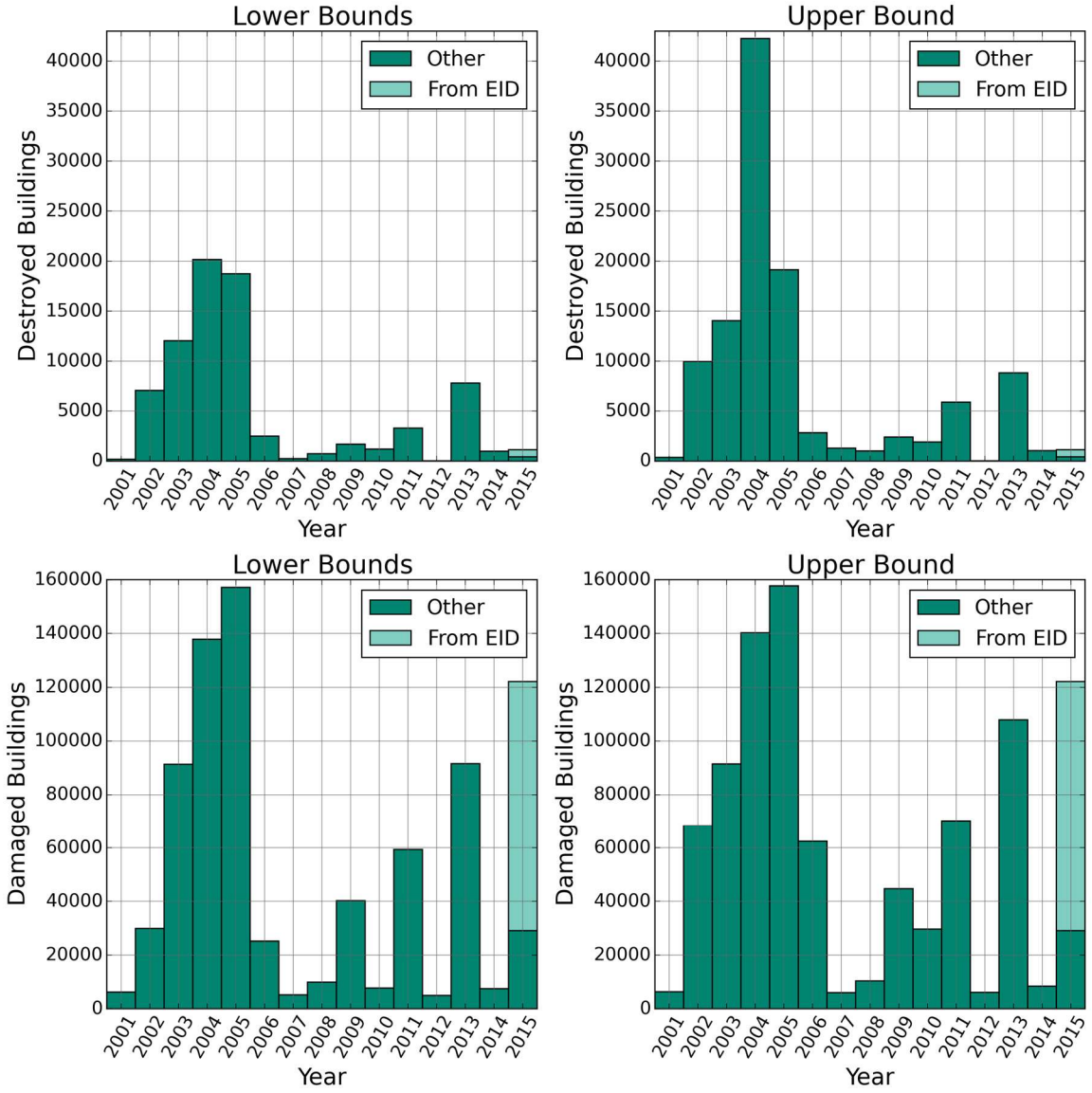

Figure 5.16. Number of damaged (bottom) and destroyed (top) buildings for the events that make up the Database of Damaging Small-to-Medium Magnitude Earthquakes for which numerical information is available, classified according to whether they were retrieved from the EID or other sources. Lower (left) and upper (right) bounds shown.

When considering only main shocks, the proportions rise from 1.9% to 3.3% for 2001-2015, from 1.0% to 1.8% for 2001-2012 and from 4.3% to 7.4% for 2013-2015. In contrast, they drop to 0.9% for 2001-2015, 0.5% for 2001-2012 and 2.0% for 2013-2015 in the case of aftershocks. This reduction observed for aftershocks might be associated with three reasons. Firstly, the fact that, in many cases, the damage and/or casualties caused by an aftershock get reported together with the consequences from the main shock, both for the sake of simplicity and because it is often impossible to distinguish one from the other, if they occur too closely-spaced in time. Secondly, it is likely that earthquakes in the range **M**4.0-5.5 that are classified as aftershocks occur in areas of high seismicity, where buildings might be better designed to withstand the shaking and societies are more used to frequent shaking, so that minor damage does not get reported. Finally, there is a known tendency for 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 & Abrahamson, 2014). It is interesting that the corollary of this appears to be 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, that is, the fact that previous shocks weaken structures that end up being more vulnerable to later shocks—but the opposite.

As depicted in the plot on the right of Figure 5.13, the proportion of damaging earthquakes increases significantly with magnitude, as would be expected both from the Gutenberg-Richter relationship (Gutenberg and Richter, 1944) (Figure 5.6) and the distribution of moment magnitude of the Database of Damaging Small-to-Medium Magnitude Earthquakes (Figure 4.4).

When discriminating earthquakes flagged as induced by either of the two methodologies described in Section 4.7, the overall proportion of damaging earthquakes (*i.e.*, for the whole 15-year period, all shocks) changes very little for the non-induced subset, as shown in Table 5.4, but rises to 2.3-2.9% for the induced subset. This might be due to the likely higher propensity to report earthquake damage in areas of the world where a connection between seismicity and anthropogenic activities is strongly suspected. In such places, seismic risk tends to be perceived as an imposed and, therefore, avoidable risk, while regions of the world where seismic hazard is dominated by tectonic activity tend to perceive risk as natural. The corollary of this observation could be either that the proportion of damaging earthquakes in the non-induced subset be too low or that the higher proportion of damaging earthquakes within the induced set be somewhat artificially inflated, or even a combination of the two. It is nevertheless noted that the number of induced earthquakes is significantly smaller than that of non-induced ones and, consequently, the significance of the sample could be questioned.

Table 5.4. Proportion of damaging events within the catalogue of potentially damaging earthquakes, classified in terms of kind induced/non-induced as per flagging in the ISC Bulletin / WPG16v3c and based on the HiQuake database.

Classification Strategy	INDUCED			NON-INDUCED		
	Total	Dam.	% Dam.	Total	Dam.	% Dam.
ISC Bulletin / WPG16v3c	341	8	2.3	38,786	732	1.9
ISC Bulletin / WPG16v3c / HiQuake	2,777	80	2.9	36,350	660	1.8

Within the results presented in Table 5.4 for the induced subset, it is interesting to note that the large increase in the number of earthquakes flagged as induced when using the methodology based on HiQuake (and/or combined with that based on the ISC Bulletin/WPG16v3c) does not translate into an increase of the same magnitude in the proportion of damaging events. As was shown in Table 5.2, the proportion of induced earthquakes based on the ISC Bulletin/WPG16v3c was 0.9% while that based on HiQuake was 6.7%, and that based on both methodologies is 7.1%. However, the resulting proportions of damaging earthquakes vary only between 2.3% and 2.9%. This is interesting

because there is a possibility that the flagging based on HiQuake result in several false-positive cases (*i.e.*, events flagged as induced when they were not), given the large number of HiQuake entries for which the spatial and temporal windows could not be properly constrained (see Figure 5.7), and this still seems to result in a relatively small change in the proportion of damaging events. It is nevertheless noted that flagging of induced earthquakes is very difficult, particularly at a global scale, and neither of the two methodologies used herein can be deemed perfect (note the Oklahoma earthquakes not being flagged with the ISC Bulletin/WPG16v3c, see Section 5.2).

In order to illustrate the potential impact of modifying the criteria used to define the catalogue of potentially damaging earthquakes for the present study, the different population counts and density thresholds used for the sensitivity analysis of Section 3.4 (Table 3.6) were applied to define alternative versions of the catalogue of potentially damaging earthquakes. Figure 5.17 shows the relation between the latter and the corresponding percentage of damaging events. As can be observed, the percentage of damaging earthquakes decreases when the size of the catalogue of potentially damaging events increases. Minimum values are reached when the most flexible criterion C7 (1 person, 1 person/km^2) is used. A similar 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 Section 4.5.2 to be included in the final ratio (see Section 5.4.3 for details on these earthquakes). It is noted, however, that the order of magnitude of the percentage of damaging events stays within a reasonable range.

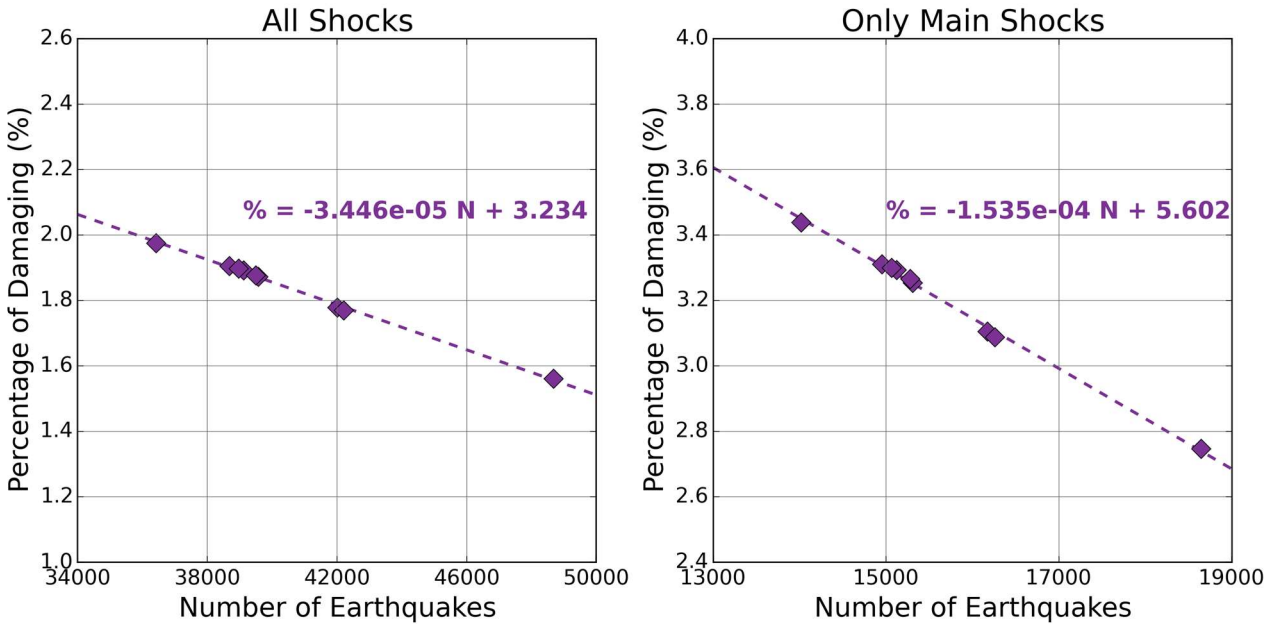

Figure 5.17 Percentage of damaging earthquakes against total number of potentially damaging earthquakes, all shocks (left) and only main shocks (right) considered. Linear regressions shown in purple.

5.4 Reportedly Damaging Earthquakes That Are Not Part of the Statistics

As mentioned in Section 4.3, the Database of Damaging Small-to-Medium Magnitude Earthquakes (Nievas *et al.*, 2019) contains 996 events in the period 1st January 2001 – 31st December 2015. Of these (see Figure 5.18):

- 740 (74.3%) can be found within the 39,127 earthquakes that make up the catalogue of potentially damaging earthquakes,
- 30 (3.0%) can be found within the 141,520 earthquakes that make up the global catalogue of upper-crustal earthquakes but do not pass the exposure criteria set in Section 4.6 and are thus not within the set of 39,127,
- 213 (21.4%) can be found in the initial unfiltered merged catalogue but do not comply with the maximum depth limits set in Section 4.5.2,
- 6 (0.6%) lie outside the magnitude range of interest,
- and 7 (0.7%) are not found in the initial unfiltered merged catalogue either.

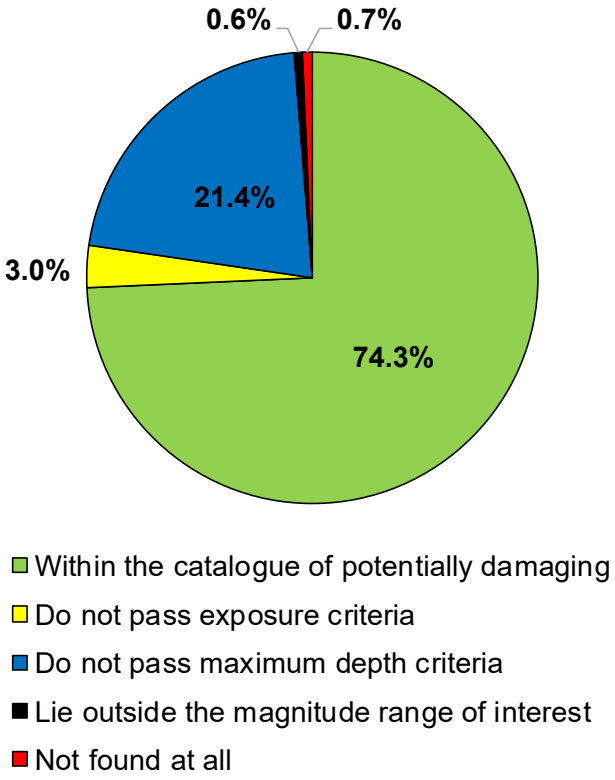

Figure 5.18. Situation within the global catalogues compiled herein of the 996 earthquakes that make up the Database of Damaging Small-to-Medium Magnitude Earthquakes (Nievas et al., 2019) for the period 1st January 2001 – 31st December 2015.

The four sets of damaging events not found in the catalogue of potentially damaging earthquakes, whose joint number of events is 256, are investigated in the subsections below.

5.4.1 Reportedly Damaging Earthquakes Not Found in the Merged Catalogue

The seven cases of reportedly damaging earthquakes that cannot be found in the merged catalogue at all are the following (codes in parentheses indicate ID within the Database of Damaging Small-to-Medium Magnitude Earthquakes):

- 1st May 2005 12:23:00 UTC, Chuschi, Peru (M188): For unknown reasons, this earthquake and its consequences are only mentioned by governmental reports and offices from Peru, and is, therefore, not part of the merged catalogue. While information on damage and casualties is not fully consistent and is, in many ways, subject to the interpretation of the choice of words, an analysis of the available sources suggests that at least around 2,000 people were affected by this earthquake, more than 300 buildings were damaged, and over 200 were destroyed (INDECI, 2005a, 2005b; Tavera *et al*., 2016).

- 13th March 2007 08:04:00 UTC, Manica, Mozambique (M129): This earthquake could not be found in either the USGS catalogue or the ISC Bulletin, and it is not mentioned either in EM-DAT or in NOAA. As a consequence, it is not part of the merged catalogue. It was, however, reported in the news that six school children were injured when rushing into their school building due to the fear to the tremors (Earthweek, 2007; IOL, 2007). There does not seem to have been any damage or any other casualties.

- 24th December 2006 22:43:26 UTC, Rajasthan, India (A119): This earthquake could not be found in the ISC Bulletin, but exists in the USGS catalogue[7], where slight damage is reported. More details on the damage observed are given in the Amateur Seismic Centre (ASC) website[8], which focuses on earthquakes that occur in the Indian Subcontinent. The NOAA database lists this earthquake but all its fields are blank.

- 13th July 2013 04:15:13 UTC, Poland (E13-184): The ISC Bulletin[9] has (at the time of writing) four statements regarding the origin of this earthquake: “suspected mining induced” (twice), “suspected explosion”, and “anthropogenic event”. The fact that the keyword “explosion” is contained in these statements led to this event being discarded from the merged catalogue. No scientific publications are listed by the ISC. The EID says “some buildings suffered (additional) cracks”.

- 18th December 2013 14:14:29 UTC, Java, Indonesia (E13-373): This earthquake can be found in the ISC Bulletin but with only one magnitude estimate in M_{Lv} scale, which was considered neither by Weatherill *et al*. (2016) in the compilation of WPG16v3c nor in the present work.

[7] https://earthquake.usgs.gov/earthquakes/eventpage/usp000f0xp/impact
[8] http://asc-india.org/lib/20061224-jaipur.htm
[9] http://www.isc.ac.uk/cgi-bin/web-db-v4?event_id=603231302&out_format=IMS1.0&request=COMPREHENSIVE

- 30th November 2014 16:28:12 UTC, Poland (E14-353): The ISC Bulletin[10] has (at the time of writing) three statements regarding the origin of this earthquake: “suspected mining induced”, “suspected explosion”, and “anthropogenic event”. The fact that the keyword “explosion” is contained in these statements led to this event being discarded from the merged catalogue. No scientific publications are listed by the ISC. The EID assigns damage level 2 to this earthquake, which translates into “some houses damaged, mainly minor damage”. The source cited therein[11] speaks of (an unspecified number of) cracks in the walls of residential buildings.

- 18th October 2015 21:44:53 UTC, Oblast Sverdlovsk, Russia (E15-274): At the time of writing (January 2019), the ISC Bulletin[12] does not have any comment related to an anthropogenic origin of this event. However, the version of the Bulletin used for this work (downloaded on 21st September 2018) had a comment saying “suspected mining explosion” and another saying “anthropogenic event”. Within the EID, this event is flagged as mining-related and associated with damage to two buildings, though the cited source is no longer available online.

The identification of three reportedly damaging earthquakes potentially misclassified as explosions led to the need to look at the issue in more detail and to try to estimate how many cases of the kind might exist. Before getting into the details it is useful to remember that rejection of events and classification of events as having an anthropogenic origin was carried out by means of GEM’s Catalogue Toolkit (Weatherill *et al.*, 2016), which searches for keywords of relevance within the events in the ISC Bulletin to make such decisions. Earthquakes that contained the keywords “explosion”, “chemical” and “nuclear” were rejected both in the compilation of the WPG16v3c catalogue and in the present work, while earthquakes that contained the keywords “geothermal”, “reservoir”, “mining”, “anthropogenic”, or “rockburst” were flagged as associated with an anthropogenic origin. These keywords were sought within the comments to the events that are reported to the ISC by the contributing agencies, right below the origin estimate of the corresponding agencies, which implies that one event might contain different comments from different agencies.

Based on what was observed for E13-184, E14-353 and E15-274, and similar cases found for Groningen, the Netherlands, which are most likely induced by gas extraction activities, the hypothesis to test was that there exist cases that are likely anthropogenic in origin but not explosions, misclassified as explosions by some agencies. The ISC Bulletin files for the period 1st January 1998 – 31st August 2018 were read again, filtering out only earthquakes containing the keywords “chemical” or “nuclear”, but not “explosion”, using the toolkit published alongside the paper of Weatherill *et al.* (2016). Then, events that had flags of both “explosion” and at least one of the keywords used to identify anthropogenic events, namely “geothermal”, “reservoir”, “mining”, “anthropogenic”, and “rockburst”, but neither “mining

[10] http://www.isc.ac.uk/cgi-bin/web-db-v4?event_id=606064137&out_format=IMS1.0&request=COMPREHENSIVE
[11] https://twojepajeczno.pl/wiadomosci/trzesienie-ziemi/
[12] http://www.isc.ac.uk/cgi-bin/web-db-v4?event_id=607921370&out_format=IMS1.0&request=COMPREHENSIVE

explosion" nor "blast", were identified. There were 1,111 events complying with these criteria, of which Figure 5.19 shows their moment magnitudes and distribution in time. As can be observed, only 62 of these fall within the range **M**4.0-5.5, and only 44 of these 62 fall within the final time period of interest (1st January 2001 – 31st December 2015), which implies that their contribution to the overall catalogue would have been small.

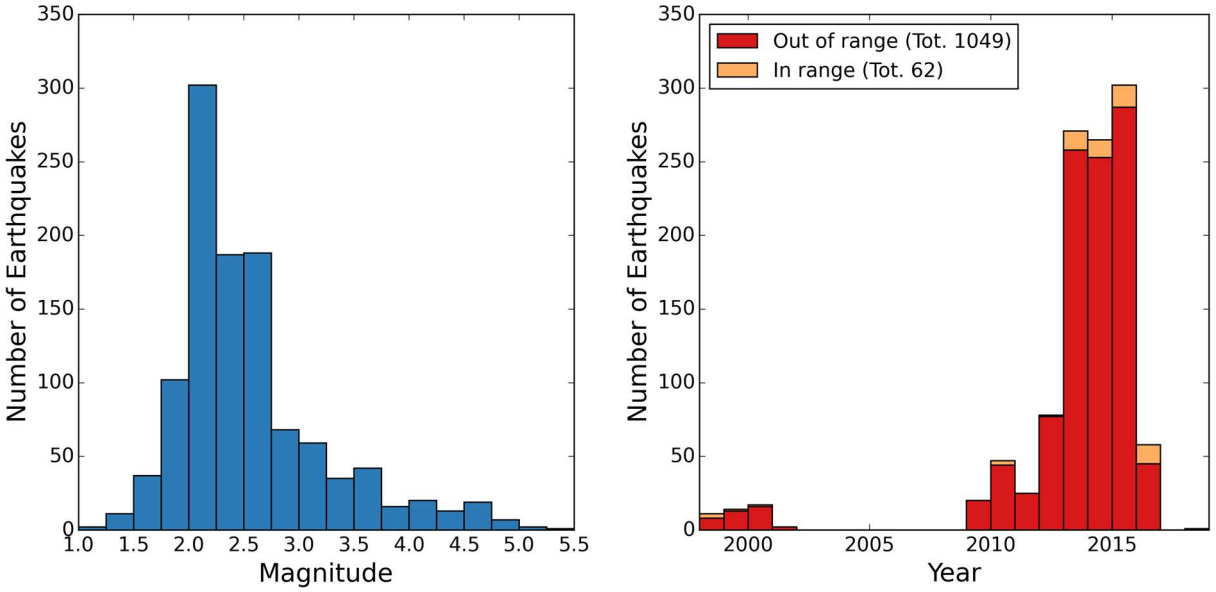

Figure 5.19. Moment magnitudes (left) and year of occurrence (right) of earthquakes of (potential) anthropogenic origin potentially misclassified as explosions. In the plot on the right, orange and red bars indicate moment magnitudes within and outside the range of interest (**M**4.0-5.5), respectively.

A closer look at the 62 earthquakes whose magnitude falls in the range **M**4.0-5.5 revealed that in 38 cases was the "suspected explosion" comment associated with the University of Upssala (UPP), while other agencies labelled the events as "anthropogenic" and "mining-induced". Five other cases had their "suspected explosion" comment associated with the Institute of Seismology of the University of Helsinki (HEL), with other agencies saying "mining-induced". In other eight cases both UPP and either the Geological Survey of Denmark and Greenland (DNK, seven cases) or the NORSAR foundation (NAO, one case) provide the label "suspected explosion", while other agencies say "suspected mining-induced" and/or "anthropogenic". The strong repetition of a particular agency labelling earthquakes as suspected explosions while other agencies point at anthropogenic origins of a different nature is, at least, worth of attention, as it may reflect an issue with the way an agency generates its flags. It is noted that the 1,111 earthquakes identified included six located in the area of Groningen, all of them with **M** below 4.0 and, thus, not within the subset of 62 discussed herein, which are known to be induced by gas extraction activities. Another easily recognisable cluster located in correspondence with an area known to be associated with induced seismicity is that around the Polkowice mine in Poland (*e.g.*, Gibowicz, 1998; HiQuake database; Orlecka-Sikora *et al.*, 2009).

The 62 (and 1,111) earthquakes do include, of course, some cases of verifiable explosions, such as the nuclear test in North Korea of 6th January 2016, and other cases which have been labelled as explosions by more than one agency. An example of these is a set of

events that occurred in Murmansk Oblast, in north-western Russia, over which some agencies only state “mining”, but which may be either induced by blasting operations or correspond to the blasting operations themselves (Adushkin, 2013). A fundamental issue to highlight is thus that distinguishing between natural earthquakes, induced earthquakes and non-earthquake phenomena is not trivial. With many detailed studies on specific earthquake sequences still having difficulty in being conclusive, criteria used by seismological agencies to classify events as they occur is likely even more subject to potential misclassifications. In this sense, the identification of these 1,111 events herein is simply a way of quantifying the extent to which events classified as explosions and as having some other anthropogenic origins exist, and does not intend to be conclusive in any way. We cannot know, without additional investigation, whether these events were explosions, as we cannot verify whether other events classified only as explosions were indeed such, or even whether events not classified in any particular way were explosions.

5.4.2 Reportedly Damaging Earthquakes Out of Magnitude Range

The six cases of reportedly damaging earthquakes that fall out of the magnitude range of interest for the present study are the following:

- 9th June 2004 11:41:16 UTC, Western Saudi Arabia (A81): In the Database of Damaging Small-to-Medium Magnitude Earthquakes, this event is represented by a **M**4.4, retrieved from the ISC Bulletin, reportedly calculated by the Geophysical Institute of Israel (GII). However, within the merged catalogue this event comes from WPG16v3c, where it has been assigned a **M**3.87, stemming from a conversion from the ISC-calculated m_b=3.7, as Weatherill et al. (2016) did not include GII among the considered agencies. According to the USGS, this earthquake caused minor damage in the epicentral area. NOAA assigns an economic loss of less than 1 million USD to this event but provides no further details.

- 18th June 2007 14:29:48 UTC, Qom, Iran (A48): In the Database of Damaging Small-to-Medium Magnitude Earthquakes, this event is represented by a **M**5.549, calculated from the seismic moment M_o of 2.36E24 dyn·cm reported in the GCMT catalogue. However, within the merged catalogue this event comes from WPG16v3c, where it has been assigned a **M**5.55 from GCMT as well. The problem is clearly that of floating precision as **M**5.549 satisfies the condition 3.95≤**M**<5.55 but **M**5.55 does not. According to the USGS, this earthquake caused slight damage to some buildings, while the NOAA estimates 50-100 damaged buildings, some of which can be expected to have been destroyed, and economic losses of less than 1 million USD.

- 20th September 2013 02:06:33 UTC, Austria (E13-257): In the Database of Damaging Small-to-Medium Magnitude Earthquakes, this event is represented by a **M**4.5, retrieved from the ISC Bulletin, reportedly calculated by MedNet Regional Centroid - Moment Tensors (MED RCMT). However, within the merged catalogue this event comes from WPG16v3c, where it has been assigned a **M**3.65, stemming from a conversion from the ISC-calculated m_b=3.5, as Weatherill et al. (2016) did not include

MED RCMT among the considered agencies. According to comments contained within the ISC Bulletin, this earthquake caused slight damage to buildings in the form of fissures and spalling of plaster, while the EID describes the damage as limited.

- 17th February 2014 05:56:01 UTC, Luzon, Philippines (E14-052): In the Database of Damaging Small-to-Medium Magnitude Earthquakes, this event is represented by a **M**5.549, calculated from the seismic moment M_o of 2.36E24 dyn·cm reported in the GCMT catalogue. However, within the merged catalogue this event comes from WPG16v3c, where it has been assigned a **M**5.55 from GCMT as well. The problem is clearly that of floating precision as **M**5.549 satisfies the condition 3.95≤**M**<5.55 but **M**5.55 does not. The EID reports non-structural damage for this earthquake.

- 28th August 2014 17:49:21 UTC, Lombardy, Italy (E14-262): In the Database of Damaging Small-to-Medium Magnitude Earthquakes, this event is represented by a **M**4.0, retrieved from the ISC Bulletin, reportedly calculated by MedNet Regional Centroid - Moment Tensors (MED RCMT). However, in the merged catalogue it is assigned a **M**3.6 retrieved from the ISC Bulletin, but calculated by the INGV. This is clearly related to the need to assign one value of magnitude to each earthquake, despite different agencies arriving to different results. The EID reports non-structural damage for this earthquake.

- 8th March 2015 20:47:27 UTC, Kosjeric, Serbia (C99): In the Database of Damaging Small-to-Medium Magnitude Earthquakes, this event is represented by a **M**4.5, retrieved from the ISC Bulletin, reportedly calculated by MedNet Regional Centroid - Moment Tensors (MED RCMT). However, within the merged catalogue this event comes from WPG16v3c, where it has been assigned a **M**3.86, stemming from a conversion from the NEIC-calculated m_b=3.9, as Weatherill *et al.* (2016) did not include MED RCMT among the considered agencies. Interestingly, the ISC Bulletin reports as well an estimated value of M_{wr} by NEIC of 4.4, which is significantly closer to the **M**4.5 of MED RCMT than to the m_b-derived proxy **M**3.86, but M_{wr} was not considered in WPG16v3c. NOAA reports 100 buildings damaged by this earthquake and estimates an economic loss of less than 1 million USD.

The cases of events A81, C99, E13-257 and E14-262 illustrate the impact of the variability of the estimates of magnitude carried out by different agencies. Events A48 and E14-052 illustrate instead the arbitrariness of float precisions within computer calculations.

5.4.3 Reportedly Damaging Earthquakes that Do Not Pass the Depth Criterion

Figure 5.20 depicts the epicentral locations and depths according to the merged catalogue of the 213 earthquakes that can be found in the initial unfiltered merged catalogue but do not comply with the maximum depth limits set in Section 4.5.2. As shown in Figure 5.21, their distribution in time is not uniform and a great number of them (139 out of 213 = 65.3%) occurred in 2013 through 2015.

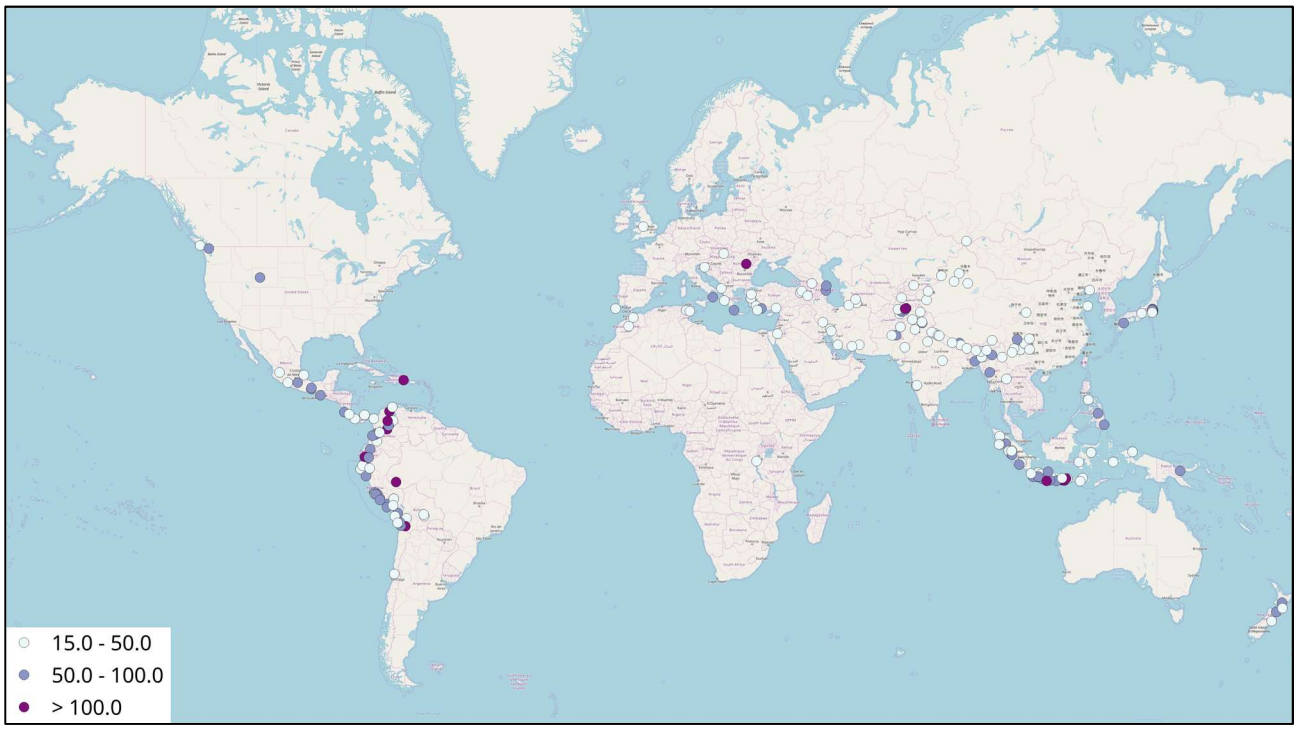

Figure 5.20. Location of the 213 reportedly damaging earthquakes that do not pass the maximum depth criteria defined in Section 4.5.2. Colour scale indicates hypocentral depth as in the global merged catalogue.

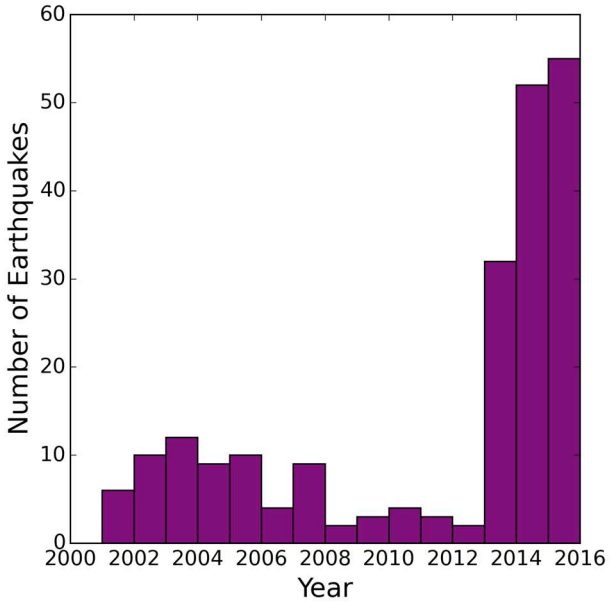

Figure 5.21. Year of occurrence of the 213 reportedly damaging earthquakes that do not pass the maximum depth criteria defined in Section 4.5.2.

The plots in Figure 5.22 depict the hypocentral depths as reported in the Database of Damaging Small-to-Medium Magnitude Earthquakes (plot on the left) and as reported in the merged catalogue (plot on the right) against moment magnitude from the merged catalogue for these 213 events. As can be observed, a number of these earthquakes would pass the maximum depth criteria if the values reported in the Database of Damaging Small-to-

Medium Magnitude Earthquakes were adopted. This observation raises the question of how much the choice of a particular depth estimate for each earthquake in the merged catalogue could have influenced the inclusion or not of events in the final catalogue of potentially damaging earthquakes. In order to assess this for these 213 events, all depth estimates from different authors reported in the ISC Bulletin were retrieved for each of them and the percentage of those that comply with the maximum depth limits was determined. It is noted that it can occur that the ISC Bulletin reports depths by an agency B citing agency A as the source (*i.e.*, the ISC Bulletin got the data from agency B but agency B adopted the value calculated by agency A). If the depth is then reported both by agencies A and B, then this value should be counted only once. However, identifying such cases is difficult and there is no guarantee either that every time an agency B adopts a value from another agency A it is reported as such within the Bulletin, and it was thus not done. Consequently, each reported value was counted as an independent estimation. Figure 5.23 shows the results obtained in terms of numbers of earthquakes for which each range of percentage of passing the depth criteria occurred. As can be observed, 58 earthquakes out of 213 (27%) have depth estimates that pass the depth criteria less than 5% of the times, while around 65% have done so less than 50% of the times. These values suggest that exclusion of these events from the filtered catalogue was not due to a simple unfortunate selection of depth estimates when assigning one location and magnitude from the ISC Bulletin to each earthquake.

Potential causes for the existence of many supposedly deep yet reportedly damaging small-to-medium magnitude earthquakes are:

1. poor resolution of the depth estimates due to lack of good quality waveforms resulting in estimated depths been larger than the real unknown value;
2. erroneous assignation of an earthquake to an entry of the EID during the compilation of the Database of Damaging Small-to-Medium Magnitude Earthquakes, as this was done by means of a semi-automatic process, as described in Nievas *et al.* (2019) (*i.e.*, the selected earthquake is not the one that caused the damage/casualties);
3. the earthquakes having caused unexpectedly large ground motions for their depths;
4. erroneous reporting of damage or casualties for an earthquake that did not cause either (in the original sources used to compile the Database of Damaging Small-to-Medium Magnitude Earthquakes);
5. reported consequences being only indirectly related with the earthquake.

An example of the last potential cause is ID E13-078 of the Database of Damaging Small-to-Medium Magnitude Earthquakes: a **M**5.44 event that occurred on 4th April 2013 at 06.31 UTC in Afghanistan at a depth of 239 km, for which the EID reports damage level 1, which translates (according to their damage scale) into non-structural damage, and three injuries due to people falling off the stairs. An article from Earthquake-Report[13] to which the EID links states that no damage occurred and the injuries were due to the panic that the shaking caused in Peshawar.

[13] https://earthquake-report.com/2013/04/04/strong-earthquake-hindu-kush-region-afghanistan-on-april-4-2013/

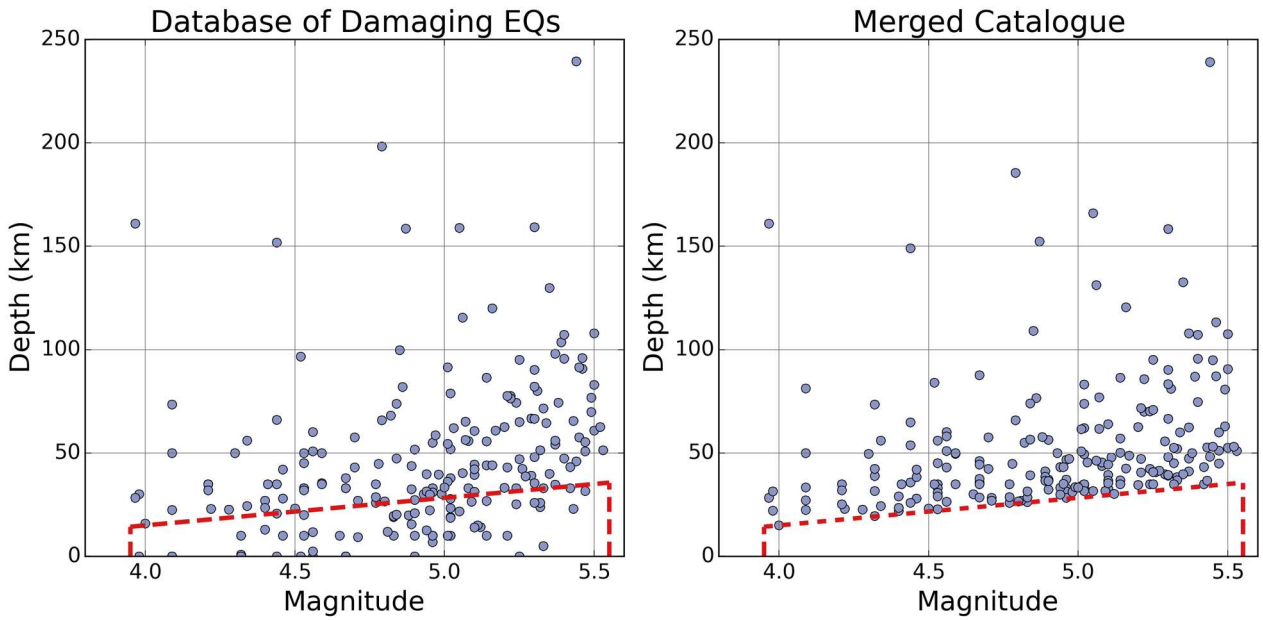

Figure 5.22. Depth as reported in the Database of Damaging Small-to-Medium Magnitude Earthquakes (left) and depth as reported in the merged catalogue (right) versus **M** from the merged catalogue for the 213 reportedly damaging earthquakes that do not pass the maximum depth criteria defined in Section 4.5.2. Red dashed lines indicate the magnitude-depth range considered herein. In both plots there is one point with magnitude 5.25 and depth 617 km that is not shown.

An example of point (4) above is as a **M**5.25 (m_b=5.1) event occurring at 617 km depth (not shown in Figure 5.22) on 27th November 2015 at 00.52 UTC in Acre, Brazil (ID E15-323 in the Database of Damaging Small-to-Medium Magnitude Earthquakes) contained in the EID, whose reported source was a local news website[14]. The latter referred to cracks on a wall of a church observed after an "earthquake of magnitude 5.1 that occurred in the dawn of Friday (27)". While the magnitude matches, the time does not. This earthquake did not occur during the local dawn but in the evening of the previous day. This suggests that there is a possibility that the minor cracks that can be observed in the photographs of the news website were caused by a **M**6.7 earthquake that occurred less than 24 hours earlier, reported at a similar depth, which is deemed normal for this region of the world.

Assigning to these 213 earthquakes the depth values reported in the Database of Damaging Small-to-Medium Magnitude Earthquake (or at least to those events for which doing so would have implied later inclusion in the filtered catalogue) could have been a possibility. However, this was not done so as to avoid inconsistencies with the criteria used for the compilation of the catalogue, as these 213 earthquakes are not the only ones that could have been assigned a different depth estimate along the compilation process. Manually changing the depths for these earthquakes but not for others would introduce a bias toward incorporating more damaging than non-damaging earthquakes in the catalogue.

[14] http://g1.globo.com/ac/acre/noticia/2015/11/apos-terremoto-no-ac-igreja-em-tarauaca-sofre-rachaduras.html

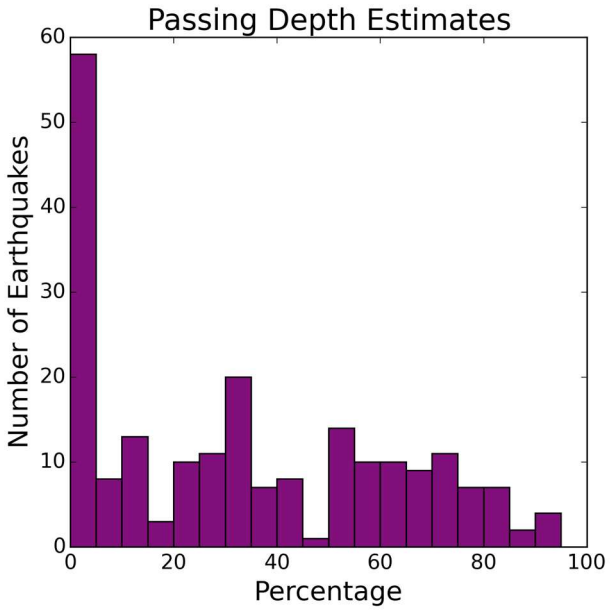

Figure 5.23. Percentage of depth estimates of each earthquake of the 213-set that comply with the maximum depth limits (consideration of uncertainty for free-depth solutions included).

5.4.4 Reportedly Damaging Earthquakes that Do Not Pass the Exposure Criteria

Given that there are 30 earthquakes that exist in the Database of Damaging Small-to-Medium Magnitude Earthquakes but do not pass the exposure criteria, it is of interest to repeat the sensitivity test carried out in Section 3.4 for the population and density thresholds, this time with the catalogue of Version 2. The same alternatives as before were used for this purpose (see definition in Table 3.6). Results obtained are shown in Table 5.5. Two things strike the eye. Firstly, that the number of earthquakes that are added when loosening the population and density thresholds to the maximum (Crit. 07, 1 person and 1 person/km^2) is only 20 out of the 30 that do not pass Crit. 01 (the final adopted criterion). Secondly, that the ratio of damaging to total number of potentially damaging earthquakes decreases as the number of earthquakes in the catalogue increases, as depicted in Figure 5.24.

Table 5.5. Number of potentially damaging and damaging earthquakes obtained applying the different population and density thresholds defined in Table 3.6.

Criterion	Potentially Damaging	Damaging	Ratio
Crit. 01	39,127	740	1.89%
Crit. 01L	39,579	741	1.87%
Crit.01H	38,682	737	1.91%
Crit. 02	42,009	747	1.78%
Crit. 03	39,494	741	1.88%
Crit. 04	38,963	739	1.90%
Crit. 05	36,416	719	1.97%
Crit. 06	42,210	747	1.77%
Crit. 07	48,683	760	1.56%

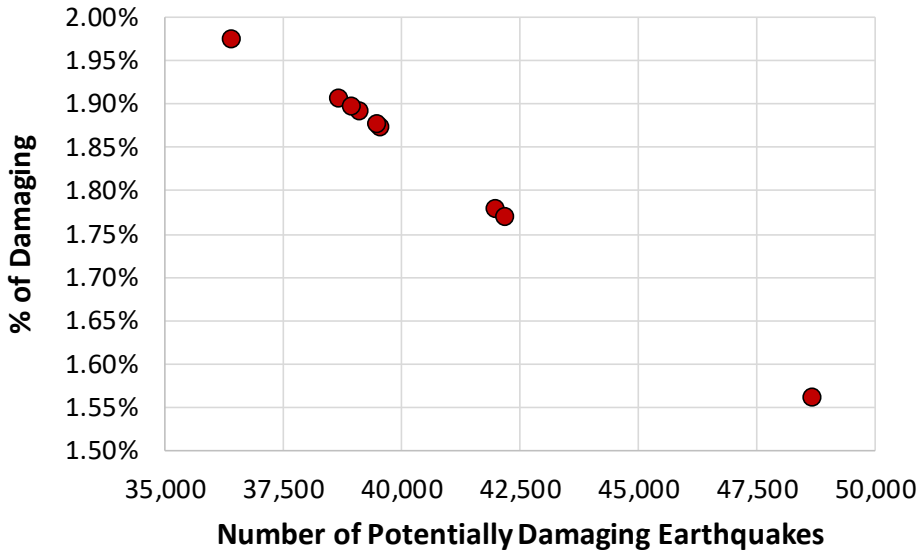

Figure 5.24. Percentage of damaging earthquakes vs. number of events in the catalogue of potentially damaging earthquakes for different population and density thresholds.

Of the 10 earthquakes that do not pass any of the criteria considered in Table 5.5, not even for Crit. 07:

- Three did not pass the exposure criterion in Version 1 either. Of these:
  - one—a **M**5.33 earthquake in Vietnam that occurred on 8th November 2005 (M169)—is reportedly associated with a casualty, which is classified as a shaking death by PAGER-CATA but for which no further information has been found;
  - one (C68) has been assigned a range of 50 to 100 damaged buildings by NOAA;
  - one (C97) is reported to have caused "minor" damage, but no information on the number of affected buildings has been found.
- Seven were not part of the version of the Database of Damaging Small-to-Medium Magnitude Earthquakes used for Version 1. Of these:
  - three (E13-205, E13-367, E14-212) are reported to have caused only non-structural damage by the EID;
  - two (E13-246, E14-175) are reported to have caused limited damage by the EID;
  - two (E14-114, C266) are reported to have caused "few, slight" and "some, minor" damage, but no information on the number of affected buildings has been found.

A variety of consequences appears to have been reported for the other 20 of the 30 earthquakes that do not comply with the exposure. While 18 of the 20 are not associated with any casualties and are reportedly associated only with mild qualitative descriptions of damaged buildings, two are reported to have caused damage to over 30,000 buildings and injuries to 28-100 people each (2004 **M**5.38 and 2009 **M**4.90 earthquakes in China, M230 and M64). Around half of them have been only assigned qualitative descriptions of damage.

Of the 30 earthquakes that do not pass the exposure criteria, only one is associated with a death (the 2005 Vietnam case mentioned above, M169), and only two are associated with injuries (the two Chinese earthquakes just mentioned).

Similarly to the case of the earthquakes that do not pass the depth criterion (Section 5.4.3), potential causes of the existence of these 30 earthquakes that do not comply with the exposure criteria are:

1. erroneous assignation of an earthquake to an entry of the EID during the compilation of the Database of Damaging Small-to-Medium Magnitude Earthquakes, as this was done by means of a semi-automatic process, as described in Nievas *et al.* (2019) (*i.e.*, the selected earthquake is not the one that caused the damage/casualties);
2. the earthquakes having caused unexpectedly large ground motions at distances for which this is usually not expected;
3. erroneous reporting of damage or casualties for an earthquake that did not cause either (in the original sources used to compile the Database of Damaging Small-to-Medium Magnitude Earthquakes);
4. reported consequences being only indirectly related with the earthquake.

5.4.4.1 Those of Version 1 That Prompted the Change in MMI Estimation

As explained in Section 4.6.2, five cases of earthquakes that exist in the Database of Damaging Small-to-Medium Magnitude Earthquakes but did not pass the exposure criterion had been identified in Version 1 of this work. Of these five, only one changed its status when using two IPEs combined by means of the Craton Index: the 20th April 2010 **M**4.5 Kalgoorlie-Boulder (Australia) earthquake. This outcome is completely reasonable, given that south-east Australia is quite cratonic in nature. Of the remaining four, the two Vietnamese and the one in New Zealand lay in areas with craton indices equal to zero, while the Chinese earthquake is associated with an index of 0.09, which was not large enough to make a difference.

It is interesting to note that the 2010 **M**5.0 Porangahau, New Zealand, earthquake passes the exposure criteria when Crit. 07 is used, that is, when the population and density thresholds are 1 person and 1 person/km^2, respectively. Even in Version 1 the MMI IV isoseismal predicted by Allen *et al.* (2012) enclosed some population, though it was not enough to pass the pre-defined thresholds.

5.4.4.2 Changes from Version 1 to Version 2

While in Version 1 only five earthquakes that were part of (Version 1 of) the Database of Damaging Small-to-Medium Magnitude Earthquakes were not part of the world catalogue due to not complying with the exposure criterion, the number rises up to 30 in Version 2. It is thus of relevance to understand what has changed in between.

These two groups of five (Version 1) and 30 (Version 2) earthquakes share four earthquakes in common (C68, C97, M169, C51), as one of the five from Version 1 now passes the exposure criteria (20th April 2010 **M**4.5 Kalgoorlie-Boulder, Australia, M55). Of the remaining 26 earthquakes of Version 2, 23 were not in Version 1 of the Database of Damaging Small-to-Medium Magnitude Earthquakes, and hence were not sought for in the world catalogue, while two (C76, M64) were in Version 1 of the Database of Damaging Small-to-Medium Magnitude Earthquakes but did not pass the depth criteria in Version 1 (they do pass the depth criteria in Version 2). The last remaining earthquake of this group, M230, used to pass the exposure criterion in Version 1 but does not pass in Version 2 due to the change in the IPE from that of Allen *et al.* (2012) to that of Atkinson & Wald (2007), which yields lower values of MMI than the former when the craton index is zero (*i.e.*, when coefficients for California are used).

5.4.5 Characterisation of Earthquake Consequences and Discussion

As 256 earthquakes out of 996 events that make up the Database of Damaging Small-to-Medium Earthquakes for the period 1st January 2001 – 31st December 2015 end up not being part of the statistics of damaging events due to them not forming part of the catalogue of potentially damaging earthquakes, it is of interest to investigate the kind and extent of consequences that these events represent.

Table 5.6 shows the number of total deaths and injuries associated with these 256 earthquakes, while Table 5.7 classifies the 256 earthquakes according to the kind of casualties associated with them. Lower and upper bounds may exist either because of a disagreement in the numbers reported by different sources or because a range was assigned in the NOAA database based on qualitative descriptions (only in the case of injuries for these earthquakes). When contrasting these values against those shown in Table 4.1 it is observed that the deaths and injuries associated with these earthquakes represent 27-29% and 5-6% of those of the complete database in the period of interest, respectively. It is noted that the 106-109 reported deaths correspond to only 20 events, while over 90% of the 256 earthquakes are not associated with deaths and 71% are not associated with casualties at all.

Table 5.6. Summary of casualties observed for the 256 earthquakes of the Database of Damaging Small-to-Medium Earthquakes that are not part of the catalogue of potentially damaging earthquakes, excluding heart attacks.

Casualties	Without EID		From EID		All	
	Lower	Upper	Lower	Upper	Lower	Upper
Total Deaths	102	105	4	4	106	109
Injuries	841	1,153	86	86	927	1,239

Further seven heart attack deaths are associated with six events of these 256. For five of these six earthquakes the only deaths are those attributed to heart attacks. In one of these six cases, the heart attack death is the only reported consequence.

Table 5.7. Classification of earthquakes according to the existence or not of associated casualties for the 256 earthquakes of the Database of Damaging Small-to-Medium Earthquakes that are not part of the catalogue of potentially damaging earthquakes, excluding heart attacks.

Kind of Casualties	Number	Percentage (%)
No deaths & no injuries	182	71.1
No deaths & some injuries	54	21.1
Some deaths & no injuries	6	2.3
Some deaths & some injuries	14	5.5

Figures 5.25 and 5.26 depict the causes of death for these 106-109 deaths in terms of proportion of the total number of deaths and proportion of instances, respectively. Instances is used herein to refer to number of times in which a cause was mentioned, with more than one cause (and, hence, instance) being possible for each earthquake. As can be observed, around half of the deaths were due to the occurrence of landslides, while 42% are likely to have been due to structural failures (only 11% confirmed, remaining 31% assumed based on the extent of damage, labels “Probably structural failures” and “PAGER shaking”). In terms of instances, the proportion associated with landslides is much smaller (17-18%, four earthquakes), while that associated with structural failures is larger (56-59%, 13 earthquakes). Of these 13 earthquakes, one does not pass the exposure criteria, while twelve do not comply with the depth criteria. Of the latter, only four have 50% or more depth estimates that comply with the maximum depth criterion, while the remaining eight have less (all depth estimates retrieved from the ISC Bulletin, as was done above to generate Figure 5.23).

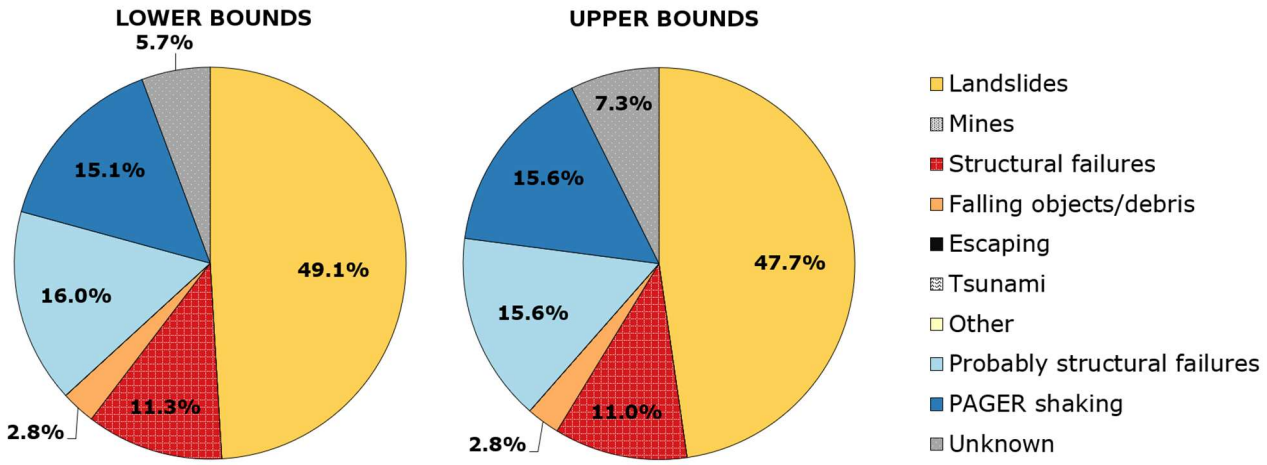

Figure 5.25. Causes of death for all total deaths excluding heart attacks (106-109 deaths), in terms of proportions of the number of deaths, associated with the 256 reportedly damaging earthquakes that do not form part of the catalogue of potentially damaging earthquakes.

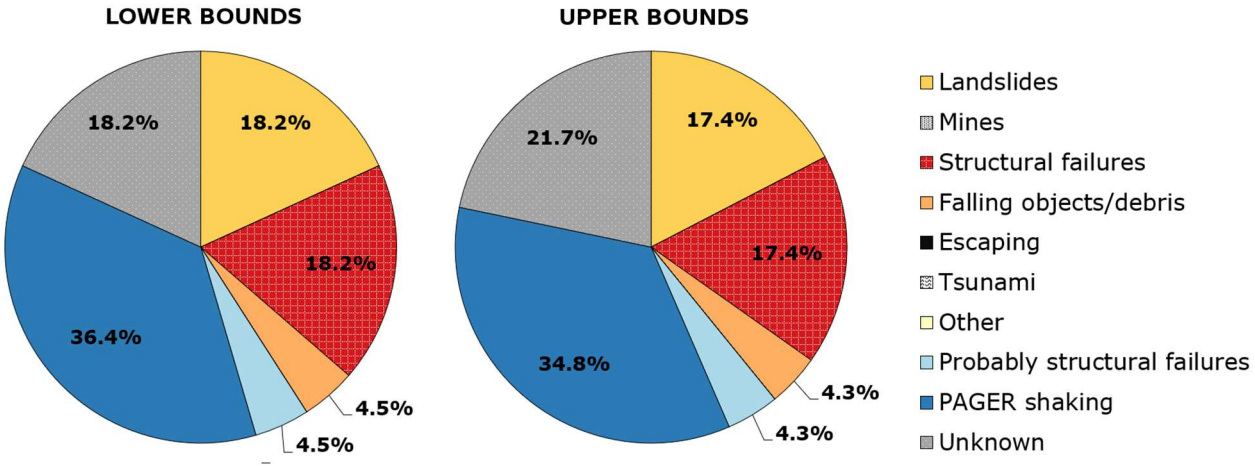

Figure 5.26. Causes of death for all total deaths excluding heart attacks, in terms of proportions of instances (*i.e.*, mentions of the cause) of deaths (22-23 instances in total, 20 earthquakes), associated with the 256 reportedly damaging earthquakes that do not form part of the catalogue of potentially damaging earthquakes.

Of the 256 earthquakes:

- 14 are reported to have caused landslides, 4 of which are associated with 52 deaths as mentioned above;
- 18 are reported to have affected infrastructure:
- none is reported to have caused liquefaction.

As explained in Section 4.3, the fields in the Database of Damaging Small-to-Medium Earthquakes concerning the number of damaged and destroyed buildings do not always contain specific numbers but many times hold verbal descriptions and ranges. Figure 5.27 shows the proportion of each kind of data that is available for these 256 earthquakes, while Figure 5.28 depicts the number of damaged and destroyed buildings for those cases in which a number or range is available and Figure 5.29 illustrates the verbal descriptions found regarding damaged buildings. For the case of destroyed buildings, only the verbal descriptor “Some” was observed for ten earthquakes.

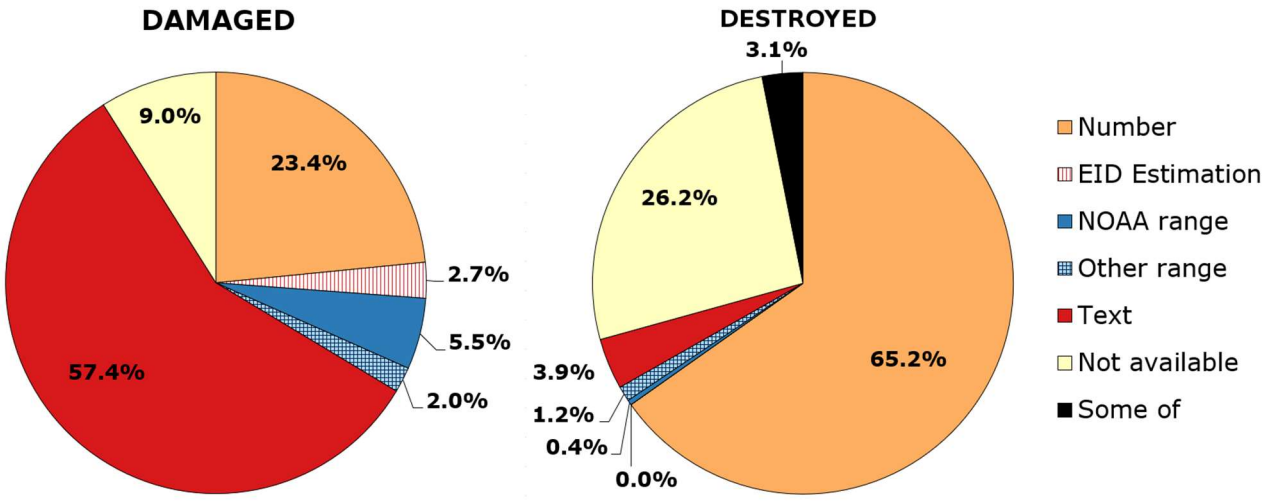

Figure 5.27. Kind of information available regarding damaged (left) and destroyed (right) buildings for the 256 reportedly damaging earthquakes that do not form part of the catalogue of potentially damaging earthquakes.

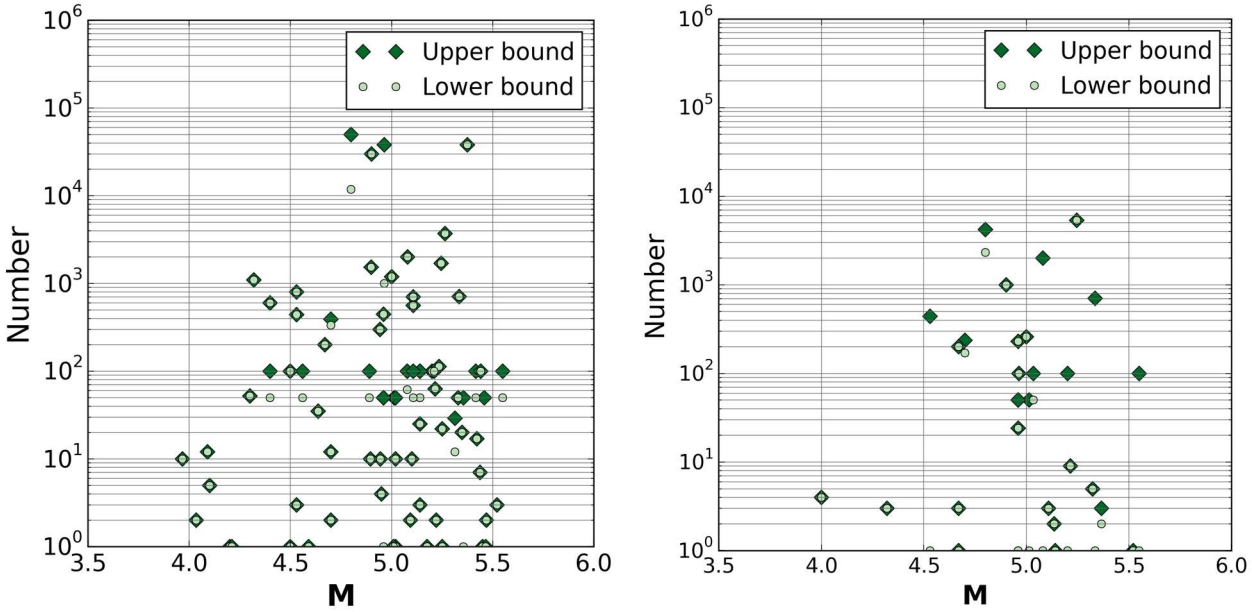

Figure 5.28. Number of damaged (left) and destroyed (right) buildings as a function of magnitude for the 256 reportedly damaging earthquakes that do not form part of the catalogue of potentially damaging earthquakes for which data was available as “Number”, “EID estimation”, “NOAA range”, “Other range”, and “Some of” (86 cases, left; 179 cases, right). Zero values not shown (9 cases, left; 150 cases, right[15]).

[15] These 9 and 150 cases are those in which zero buildings are specifically reported. They do not include cases in which no information is available.

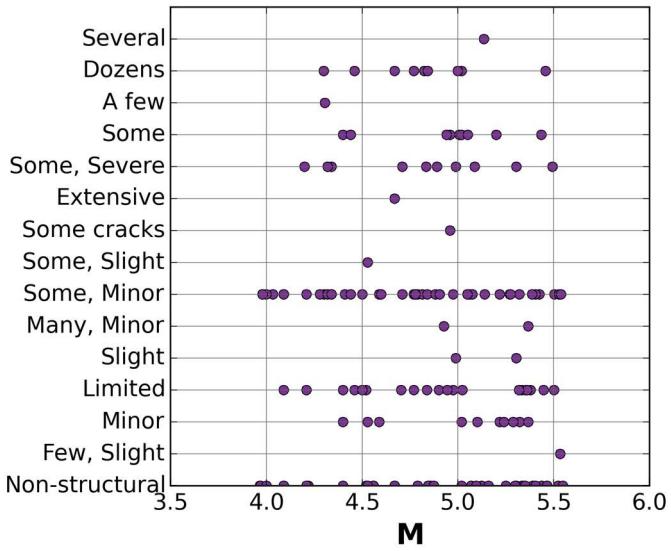

Figure 5.29. Descriptions of damage of damaged buildings as a function of magnitude for the whole of the 256 reportedly damaging earthquakes that do not form part of the catalogue of potentially damaging earthquakes for which only verbal descriptions were found (147 cases).

What is perhaps of greater interest is to compare the casualties and number of damaged/destroyed buildings for the 256 reportedly damaging earthquakes that are not part of the catalogue of potentially damaging earthquakes and the 740 that are and have been used for statistical purposes. Figures 5.30 and 5.31 do this by depicting in isolated circles and rhombuses the values corresponding to individual earthquakes, in colour those that are found within the catalogue and in grey those that are not, while markers connected with lines indicate the total values per year, in thick continuous lines for those found within the catalogue and in thin dashed ones for the total 996 earthquakes, with red and orange referring to upper and lower bounds, respectively. While the total number of deaths associated with the 256 earthquakes that are not part of the catalogue represent 27-29% of the deaths of the whole set of 996 earthquakes, it is noted from the plot on the right of Figure 5.30 that the discrepancies are mostly concentrated in the years 2002 and 2006, and just three earthquakes account for 75 of the 106-109 deaths that correspond to the 256 earthquakes, that is, 18-21% of the total number of deaths of the whole set of 996 earthquakes. Moreover, 50 of those 75 deaths were reportedly caused by landslides. It is clear from the plot on the left of Figure 5.30 that the 256 earthquakes account for only a small percentage of the total number of injuries.

For those earthquakes for which data on damaged/destroyed buildings was available as “Number”, “EID estimation”, “NOAA range”, “Other range”, and “Some of”, the total number of damaged buildings associated with the 256 earthquakes that are not part of the catalogue represent 12.4-18.7% of the total of the whole set of 996 earthquakes, while the number of destroyed buildings stands for 12.5-13.6%. However, the number of buildings damaged by the 256-building subset is greatly influenced by the numbers corresponding to just four earthquakes which account for 82.0-89.4% of the total of the subset (compare continuous and dashed lines in the plot on the left of Figure 5.31). Similarly, six earthquakes account for 77.5-89.5% of the total number of destroyed buildings of the 256-building subset, two of which correspond to cases in which all the information known is that “some of” the damaged

buildings were actually destroyed and, consequently, the reported number of damaged buildings is adopted as the upper-bound value of destroyed ones.

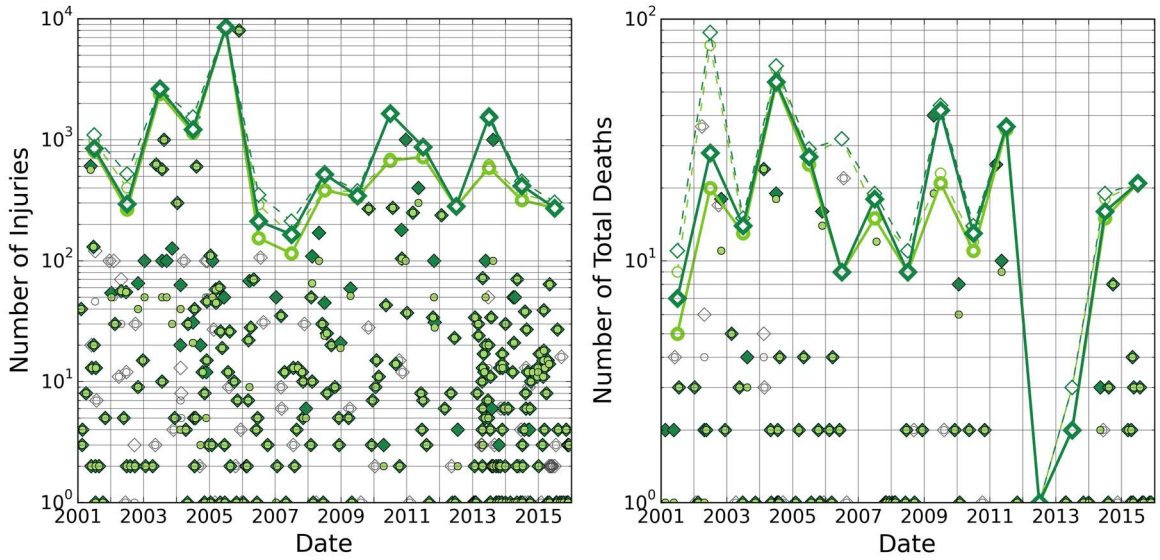

Figure 5.30. Number of injuries (left) and total deaths (excluding heart attacks, right) reported in the Database of Damaging Small-to-Medium Magnitude Earthquakes. Isolated circles (lower bounds) and rhombuses (upper bounds) indicate values per each individual earthquake, in colour those that are found 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 those of the whole set of 996 reportedly damaging earthquakes. Light and dark green refer to lower and upper bounds of values reported in the original sources, respectively. Zero values not shown.

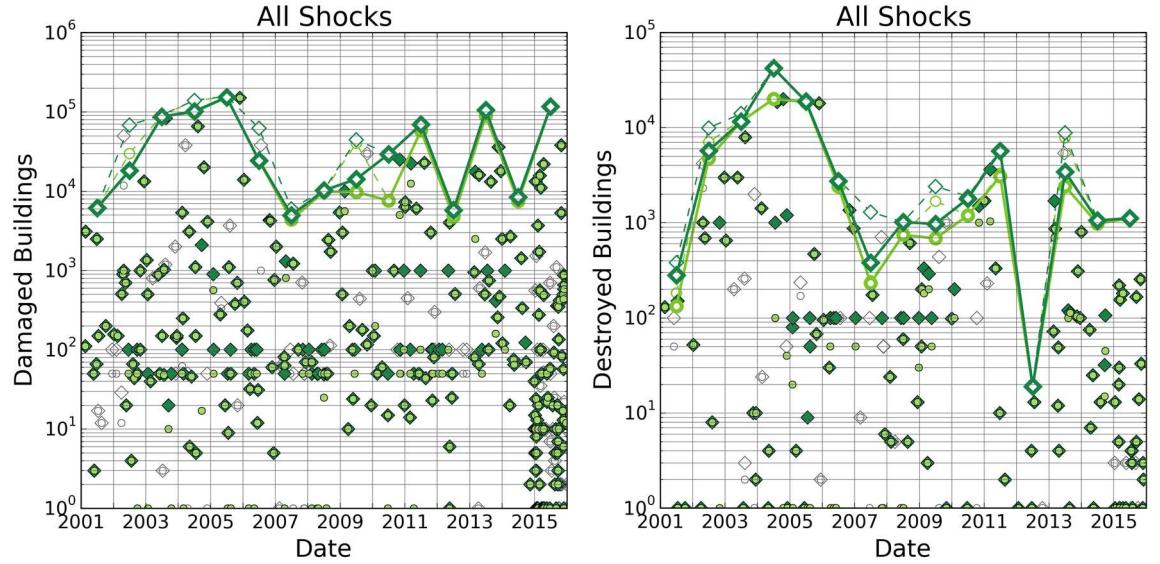

Figure 5.31. Number of damaged (left) and destroyed (right) buildings reported in the Database of Damaging Small-to-Medium Magnitude Earthquakes (only cases in which numerical values, i.e., not qualitative descriptions, are available). Colours and lines as per caption of Figure 5.30. Zero values not shown.

While Figure 5.30 covers all 256 and 996 earthquakes, it is important to bear in mind that cases for which only qualitative descriptions are available are not considered in Figure 5.31. Carrying out a comparison in terms of qualitative descriptions is much more difficult (confront Figure 5.29 and Figure 4.13).

While the numbers, plots and discussions above indicate that the consequences of the 256 reportedly damaging earthquakes that are not part of the catalogue are not negligible, it is of interest to note that the proportion of casualties and damaged/destroyed buildings they represent is overall smaller than the proportion of earthquakes. The exception would be the number of deaths, though it has been noted that almost half of those of the 256 earthquakes are attributed to landslides.

Apart from the possible causes of the existence of these reportedly damaging events that are not part of the catalogue that were already mentioned in the previous sub-sections, it is of relevance to highlight that the catalogue results from a series of analysis decisions whose grounds have been explained, but which could have been different. Both the maximum depth and exposure criteria were defined to be consistent with current knowledge on the hypocentral depth of upper-crustal earthquakes and on assumptions regarding their capability of causing damage. What should be highlighted is that setting very loose criteria that basically render *any* earthquake as potentially damaging results in the proportions of *actually* damaging earthquakes reducing drastically. The idea of setting up these criteria rises exactly from this need to constrain the dataset to a point in which the analysis becomes meaningful, the other extreme being a simple count of all earthquakes recorded and confrontation against number of earthquakes reported to have caused damage, even if the former occurred in highly-uninhabited areas.

If the 256 reportedly damaging earthquakes had been added to the catalogue of potentially damaging earthquakes, the proportions of damaging events would have increased as depicted in Figure 5.32. These values result from including only the 256 events and no other non-damaging event, and is thus quite pessimistic, as modifying the criteria used to compile the catalogue would have led to the inclusion of non-damaging events as well. As can be observed, the proportion of damaging earthquakes would have increased from 1.9% to 2.5% (dot-dashed lines), while that for 2013-2015 would have raised from 4.3% to 5.8% (dashed lines). The year with the largest percentual increase would have been 2014, which would have gone up from 4.8% to 6.6% (individual markers).

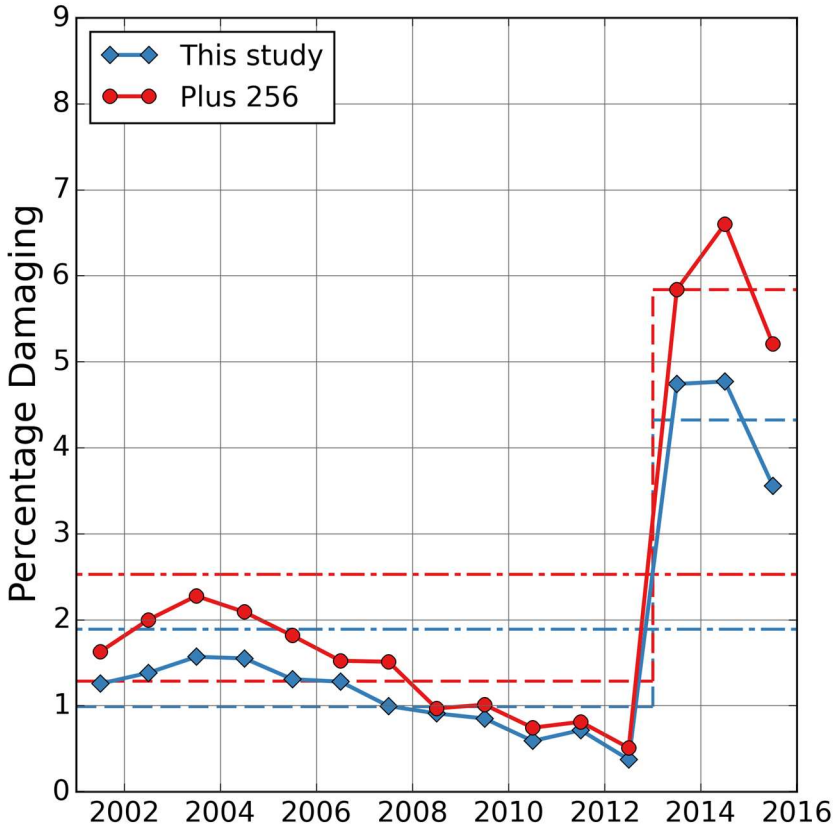

Figure 5.32. Proportion of damaging earthquakes as calculated in this study (blue) and that would result from adding the 256 damaging events that are not part of the catalogue of potentially damaging earthquakes (red). Individual markers correspond to annual values, dashed lines to the 2001-2012/2013-2015 values, and dot-dashed lines to the 15-year period values. All shocks.

6. CONCLUSIONS

With seismic hazard and risk in the Groningen field, the Netherlands, dominated by earthquakes with moment magnitude in the range 4.0-5.5, this work has investigated the proportion of upper crustal earthquakes in this magnitude range that occur *sufficiently close* to the built environment to cause damage and/or casualties. The two main challenges involved in this task were the definition of a *catalogue of potentially damaging (upper crustal) earthquakes* and the compilation of a Database of Damaging Small-to-Medium Magnitude Earthquakes, both at the global level. While the latter was carried out and reported on separately (Nievas *et al.*, 2019), the former has been the focus of the present work, an exploratory version of which (referred to as *Version 1* herein) has been presented in Nievas *et al.* (2017). The *catalogue of potentially damaging earthquakes* was compiled for the years 2001-2015 using the magnitude-homogeneous catalogue of Weatherill *et al.* (2016) and the Bulletin of the International Seismological Centre (ISC); it contains 39,127 events. A magnitude-dependent maximum depth criterion was applied in combination with intensity prediction equations and the population count/density grids of Gridded Population of the World (GPW) v4.0 (CIESIN, 2016) to provide a quantifiable definition of *sufficiently close*. The uncertainty in magnitude, depth and the intensity prediction model were explicitly taken into account for the calculation of the final uncertainty in the predicted intensity.

Both the catalogue of potentially damaging earthquakes and the Database of Damaging Small-to-Medium Magnitude Earthquakes are likely subject to completeness issues, which result in an increase in time of the number of both damaging and non-damaging events by the end of the 15-year period under study. In the case of the former, they probably stem from the increase in detectability of smaller-magnitude earthquakes that have taken place in the last decade or so as a consequence of the improvement of seismological networks worldwide, and the incorporation of more agencies to the contributors of the ISC Bulletin. In the case of the latter, the impact of accessibility to damage data from small-to-medium magnitude earthquakes is very visible in the abrupt change observed from the year 2013 onward. This was the first year for which the Earthquake Impact Database (EID), a very relevant source of the Database of Damaging Small-to-Medium Magnitude Earthquakes, was compiled in near-real time. While the incompleteness of the catalogue of potentially damaging earthquakes could be resulting in larger proportions of damaging earthquakes, the under-reporting of damage and consequences could be doing the opposite. All this is an expression of the inherent problem of insufficient recording and accessibility to data stemming both from technical limitations that are overcome in time and from changes in the perspectives of the earthquake engineering community regarding the relevance of smaller-magnitude events.

The results obtained show that 740 (1.9%) of all 39,127 shocks considered to have been potentially damaging during the 15-year period between 2001 and 2015 were identified as damaging. However, these percentage rises to 4.3% for the 3-year period of 2013-2015, due to the aforementioned increase in the identified number of damaging earthquakes from 2013 onward and drops to 1.0% for the period 2001-2012. It is thus likely that 4.3% is a more representative value than 1.9%. However, it is not possible to guarantee either that all

earthquakes causing damage and/or casualties have been captured in the EID, despite the much larger number of earthquakes it contains with respect to other databases. In any case, it is relevant to recall that, in the present study, all earthquakes identified to have caused damage and/or casualties are treated equally as damaging events, even if the extent of their consequences is quite varied and ranges from very light damage and/or injuries (caused, for example, by reacting in panic) all the way through to structural collapses and death. As a corollary, this 4.3% or potentially larger values cannot be regarded as a homogeneous set of earthquakes with serious consequences to life and the built environment but as having a mixture of a variety of potential outcomes.

When discriminating into main shocks plus foreshocks on one hand and aftershocks on the other, the proportions rise for the former and decrease for the latter. Of all aftershocks only 0.9% were identified as damaging, while the proportion of damaging earthquakes was found to be 2.7% of all main shocks plus foreshocks. We believe the reasons for this may be the fact that consequences due to aftershocks are often recorded and reported jointly with the main shock, the likelihood of many aftershocks of magnitude **M**4.0-5.5 occurring in areas of high seismicity, where minor consequences might be ignored when contrasted against more devastating events they often experience, and the known tendency of aftershocks to have lower stress drops than main shocks when the same segment of the fault is ruptured by both.

Not all the 996 events contained in the Database of Damaging Small-to-Medium Magnitude Earthquakes (Nievas *et al.*, 2019) for the 15-year period of interest were found in the catalogue of potentially damaging earthquakes: 213 (21.4%) did not comply with the maximum depth limits set to define upper-crustal events, 30 (3.0%) did not pass the filtering according to population exposure criteria, 6 (0.6%) ended up outside of the magnitude range of interest, and a further 7 (0.7%) could not be found in the ISC Bulletin at all. While details on the reasons for each of these subgroups to exist are discussed in the main body of the report, it is of relevance to highlight here those associated with the 213 events that did not comply with the maximum depth criteria. Potential causes for the existence of these supposedly deep yet reportedly damaging earthquakes might be: poor resolution of the depth estimates (*i.e.*, the earthquakes having been shallower than reported), erroneous matching of entries from the EID with earthquakes from the ISC Bulletin, as discussed in Nievas *et al.* (2019), the earthquakes having caused unexpectedly large ground motions for their depths, erroneous reporting in the original sources of damage or casualties for an earthquake that did not cause either, and reported consequences being only indirectly related with the earthquake. While many of these are beyond our control, one could think that a loosening of the maximum depth criteria could have solved, at least partially, this issue. However, it is noted that the use of a more relaxed maximum depth criteria would have likely led to the inclusion of many more non-damaging than damaging events and result in lower proportions of damaging earthquakes as was observed in the sensitivity analyses, which showed overall decreasing proportions of damaging events whenever the catalogue of potentially damaging earthquakes grew larger. Even if only the 256 reportedly damaging events that do not form part of the catalogue (but no non-damaging earthquakes) were added, the percentage of damaging would rise from 1.9% to 2.5% for the whole 15-year period, from 4.3% to 5.8% for 2013-2015, and from 4.8% to 6.6% for 2014 alone, the year

with the largest increase in percentage. These values together with a series of sensitivity analyses carried out using the exploratory Version 1 (and reported herein) have shown that results are relatively stable with respect to the maximum depth limits, the parameters used for declustering, and different population and density thresholds.

When discriminating earthquakes flagged as induced, the proportion of damaging earthquakes considering all the shocks in the 15-year period of interest changes very little for the non-induced subset, but rises to 2.3-2.9% for the induced one. This might be due to a higher likelihood of people reporting earthquake damage in regions where a connection between seismicity and anthropogenic activities is strongly suspected and, consequently, the seismic risk is perceived as imposed and avoidable. This might also imply that either the proportion of damaging earthquakes in the non-induced subset is too low, or that the higher proportion of damaging earthquakes within the induced set be somewhat artificially inflated, or even a combination of the two.

It is nevertheless noted that the number of induced earthquakes is highly uncertain, as it has been made clear that none of the two methodologies used herein to identify them can be deemed to be flawless. Apart from searching for anthropogenic-related keywords in the comments of the ISC Bulletin (Weatherill *et al.*, 2016), the Human-Induced Earthquake Database (HiQuake) (Foulger *et al.*, 2018), which contains cases found in the literature as potentially induced, was used to flag induced events that fell within pre-defined spatial and temporal windows. However, in many cases the data were not enough to constrain these windows properly, and it is thus possible that the number of flags obtained with this method be too large, or at least too large in certain areas. While only 341 (0.9%) earthquakes out of 39,127 were flagged as induced using the ISC Bulletin, 2,639 (6.7%) were flagged using HiQuake, with an overlap of only 203 flagged as induced with both and a resulting total of 2,777 (7.1%) flagged by either methodology. While the use of unconstrained time windows with HiQuake (when no better data was available) may have led to an overestimation of the number of induced earthquakes, a comparison of results obtained with each methodology revealed that most of the earthquakes occurring in Oklahoma, which are strongly believed to be induced, did not contain induced-related keywords in the ISC Bulletin and were thus not flagged as induced based on this method. At the same time, some of the Oklahoma earthquakes were flagged as induced based on HiQuake, but not all. It is thus noted that the proportions of damaging earthquakes presented herein in terms of induced/non-induced subsets are only indicative, and that the use of HiQuake for this purpose should be deemed as exploratory. The technique used herein could be improved by further investigating the spatial and temporal windows associated with each kind of causative man-made activity.

Further future refinements to the present study may include, for example, considering all depth and magnitude estimates for each earthquake instead of establishing a hierarchy of agencies/sources to select one value per earthquake, incorporating the uncertainty in epicentral location to the estimation of seismic intensity, exploring statistics within smaller sub-ranges of magnitude, and examining the specific characteristics of the most significant examples of damaging small-magnitude events.

7. 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 his assistance in the use of the GEM Catalogue Toolkit https://github.com/GEMScienceTools/catalogue_toolkit as well as fruitful discussions on catalogue compilation and regionalisation.

This work was funded by Nederlandse Aardolie Maatschappij B.V. (NAM) as part of the Study and Data Acquisition Program for induced seismicity in Groningen.

8. REFERENCES

8.1 Bibliography

Abrahamson, N.A., W.J. Silva & R. Kamai (2014). Summary of the ASK14 ground motion relation for Active Crustal Regions. *Earthquake Spectra* **30**(3),1025-1055.

Adams, R.D., A.A. Hughes & D.M. McGregor (1982). Analysis procedures at the International Seismological Centre. *Physics of the Earth and Planetary Interiors* **30**, 85-93.

Adushkin, V.V. (2013). Blasting-induced seismicity in the European part of Russia. *Physics of the Solid Earth* **49**(2), 258-277.

Afshari, K. & J.P. Stewart (2016). Physically parametrised prediction equations for significant duration in active crustal regions. *Earthquake Spectra* **32**(4), 2057-2081.

Allen, T.I., K.D. Marano, P.S. Earle & D.J. Wald (2009). PAGER-CAT: A composite earthquake catalog for calibrating global fatality models. *Seismological Research Letters* **80**(1), 1325–1345.

Allen, T.I., D.J. Wald & C.B. Worden (2012). Intensity attenuation for active crustal regions. *Journal of Seismology* **16**, 409-433.

Ancheta, T. D., R.B. Darragh, J.P. Stewart, E. Seyhan, W.J. Silva, B.S.J. Chiou, K.E. Wooddell, R.W. Graves, A.R. Kottke, D.M. Boore, T. Kishida & J.L. Donahue (2013). NGA-West2 Database. *Earthquake Spectra* **30**(3), 989–1005.

Artemieva, I.M. & H. Thybo (2013). EUNAseis: A seismic model for Moho and crustal structure in Europe, Greenland, and the North Atlantic region. *Tectonophysics* **609**, 97–153.

Assumpção, M., T.H. Yamabe, J.R. Barbosa, V. Hamza, A.E.V. Lopes, L. Balancin & M.B. Bianchi (2010). Seismic activity triggered by water wells in Paraná Basin, Brazil. *Water Resources Research* **46**(7).

Atkinson, G.M. & D.J. Wald (2007). “Did You Feel It?” intensity data: A surprisingly good measure of earthquake ground motion. *Seismological Research Letters* **78**(3), 362-368.

Avakyan, A.B. S.P. & Ovchinnikova (1971). Data on world reservoirs. *Hydrotechnical Construction* **5**(8), 773-777.

Bakun, W.H. & A. McGarr (2002). Differences in attenuation among the stable continental regions. *Geophysical Research Letters* **29**(23), 4pp.

Bolton, M., D. Storchak & J. Harris (2007). *Updating default depths in the ISC Bulletin*. Presentation of the International Seismological Centre (ISC). Available online at (last accessed 31st August 2018): ftp://ftp.isc.ac.uk/pub/iwref/docs/bolton.pdf.

Bommer, J.J., P.J. Stafford & J.E. Alarcón (2009). Empirical equations for the prediction of the significant, bracketed, and uniform duration of earthquake ground motion. *Bulletin of the Seismological Society of America* **99**(6), 3217-3233.

Bóndar, I. (2015). *ISCLOC: The location algorithm used by the ISC – Version 2.2.6*. Report of the International Seismological Centre. Berkshire, United Kingdom.

Bóndar, I. & D. Storchak (2011). Improved location procedures at the International Seismological Centre. *Geophysical Journal International* **186**, 1220-1244.

Bourne, S.J., S.J. Oates, J.J. Bommer, B. Dost, J. van Elk & D. Doornhof (2015). A Monte Carlo method for probabilistic seismic hazard assessment of induced seismicity due to conventional gas production. *Bulletin of the Seismological Society of America* **105**(3), 1721-1738.

Center for International Earth Science Information Network - CIESIN - Columbia University (2016). *Gridded Population of the World, Version 4 (GPWv4)*. Palisades, NY: NASA Socioeconomic Data and Applications Center. http://dx.doi.org/10.7927/H4NP22DQ.

Cesca, S., F. Grigoli, S. Heimann, A. González, E. Buforn, S. Maghsoudi, E. Blanch & T. Dahm (2014). The 2013 September–October seismic sequence offshore Spain: a case of seismicity triggered by gas injection? *Geophysical Journal International* **198**(2), 941–953.

Chen, D. (2010). Concepts of a basic EGS model for the Cooper Basin, Australia. *Proceedings of the 2010 World Geothermal Congress*, Bali, Indonesia, 25-29 April.

Chen, Y.-S., G. Weatherill, M. Pagani & F. Cotton (2018). A transparent and data-driven global tectonic regionalisation model for seismic hazard assessment. *Geophysical Journal International* **213**(2), 1263–1280.

Clarke, H., L. Eisner, P. Styles & P. Turner (2014). Felt seismicity associated with shale gas hydraulic fracturing: The first documented example in Europe. *Geophysical Research Letters* **41**, 8308–8314.

Dieterich, J.H., K.B. Richards-Dinger & K.A. Kroll (2015). Modeling injection-induced seismicity with the physics-based earthquake simulator RSQSim. *Seismological Research Letters* **86**(4), 1102-1109.

Di Giacomo, D., I. Bondár, D.A. Storchak, E.R. Engdahl, P. Bormann & J. Harris (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. *Physics of the Earth and Planetary Interiors* **239**, 33-47.

Di Giacomo, D., E.R. Engdahl & D.A. Storchak (2018). The ISC-GEM Earthquake Catalogue (1904–2014): status after the Extension Project. *Earth System Science Data* **10**, 1877-1899.

Dost, B., B. Edwards & J.J. Bommer (2018). The relationship between M and ML: A review and application to induced seismicity in the Groningen Gas Filed, the Netherlands. *Seismological Research Letters* **89**(3), 1062–1074.

Dziewonski, A.M., T.-A. Chou & J.H. Woodhouse (1981). Determination of earthquake source parameters from waveform data for studies of global and regional seismicity. *Journal of Geophysical Research* **86**, 2825-2852.

Ekström, G., M. Nettles & A.M. Dziewonski (2012). The global CMT project 2004-2010: Centroid moment tensors for 13,017 earthquakes. *Physics of the Earth and Planetary Interiors* **200-201**, 1-9.

Dijkstra, L. & H. Poelman (2018). *Regional typologies overview*. Eurostat - Statistics Explained platform. Available online at: https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Regional_typologies_overview. Last accessed 7th January 2019.

Foulger, G.R., M. Wilson, J. Gluyas & R. Davies (2016). *Human-induced earthquakes*. Report for the Nederlandse Aardolie Maatschappij BV (NAM), The Netherlands. Report available online at: https://www.nam.nl/feiten-en-cijfers/onderzoeksrapporten.html#iframe=L2VtYmVkL2NvbXBvbmVudC8/aWQ9b25kZXJ6b2Vrc3JhcHBvcnRlbg. Last accessed 28th November 2017.

Foulger, G.R., M.P. Wilson, J.G. Gluyas, B.R. Julian & R.J. Davies (2018). Global review of human-induced earthquakes. *Earth-Science Reviews* **178**, 438-514.

Gardner, J. K. & L. Knopoff (1974). Is the sequence of earthquakes in Southern California, with aftershocks removed, Poissonian? Bulletin of the Seismological Society of America **64**(5), 1363 – 1367.

George Herald (2012). *Strong earthquake strikes western China*. Available online at: https://www.georgeherald.com/news/News/International/40572/strong-earthquake-strikes-western-china-20170711. Last accessed: 4th February 2019.

Gibowicz, S.J. (1998). Partial stress drop and frictional overshoot mechanism of seismic events induced by mining. In Talebi, S. (ed.): *Seismicity caused by mines, fluid injections, reservoirs, and oil extraction* (pp. 5-20). Birkhäuser, Basel.

Goebel, T.H.W., M. Weingarten, X. Chen, J. Haffener & E.E. Brodsky (2017). The 2016 Mw5.1 Fairview, Oklahoma earthquakes: Evidence for long-range poroelastic triggering at >40 km from fluid disposal wells. *Earth and Planetary Science Letters* **472**, 50-61.

Goertz-Allman, B.P., B. Edwards, F. Bethmann, N. Deichmann, J. Clinton, D. Fäh & D. Giardini (2011). A new empirical magnitude scaling relation for Switzerland. *Bulletin of the Seismological Society of America* **101**(6), 3088–3095.

González, P.J. & J. Fernández (2011). Drought-driven transient aquifer compaction imaged using multitemporal satellite radar interferometry. *Geology* **39**(6), 551-554.

Grünthal, G., R. Wahlström & D. Stromeyer (2009). The unified catalogue of earthquakes in central, northern, and northwestern Europe (CENEC)—updated and expanded to the last millennium. *Journal of Seismology* **13**, 517–541.

Guglielmi, Y., F. Cappa, J.-P. Avouac, P. Henry & D. Elsworth (2015). Seismicity triggered by fluid injection-induced aseismic slip. *Science* **348**(6240), 1224-1226.

Gutenberg, B. & C. Richter (1944). Frequency of earthquakes in California. *Bulletin of the Seismological Society of America* **34**,185–188.

Hanks, T. C. & W. H. Bakun (2002). A bilinear source-scaling model for Mlog A observations of continental earthquakes. *Bulletin of the Seismological Society of America* **92**(5), 1841–1846.

Hanks, T. C. & W. H. Bakun (2008). M-log A observations for recent large earthquakes. *Bulletin of the Seismological Society of America* **98**(1), 490–494.

Juez-Larré, J., G. Remmelts, J.N. Breunese, S.F. van Gessel & O. Leeuwenburgh (2016). Using underground gas storage to replace the swing capacity of the giant natural gas field of Groningen in the Netherlands. A reservoir performance feasibility study. *Journal of Petroleum Science and Engineering* **145**, 34-53.

Klose, C.D. (2007). Coastal land loss and gain as potential earthquake trigger mechanism in SCR's. *AGU Fall Meeting Abstracts*. San Francisco, California, USA.

Kundu, B., N.K. Vissa & V.K. Gahalaut (2015). Influence of anthropogenic groundwater unloading in Indo-Gangetic plains on the 25 April 2015 Mw 7.8 Gorkha, Nepal earthquake. *Geophysical Research Letters* **42**, 10,607–10,613.

Lin, C.-H. (2005). Seismicity increase after the construction of the world's tallest building: An active blind fault beneath the Taipei 101. *Geophysical Research Letters* **32**(22), 4pp.

Lujala, P. J. Ketil Rod & N. Thieme (2007). Fighting over oil: Introducing a new dataset. *Conflict Management and Peace Science* **24**(3), 239-256. PetroData shapefiles downloaded on 27th September 2018 from: https://www.prio.org/Data/Geographical-and-Resource-Datasets/Petroleum-Dataset/Petroleum-Dataset-v-12/.

McGarr, A. (2014). Maximum magnitude earthquakes induced by fluid injection. *Journal of Geophysical Research: Solid Earth* **119**, 1008-1019.

Musson, R.M.W. (1999). Probabilistic seismic hazard maps for the North Balkan Region. *Annali di Geofisica* **42**(2), 1109–1124.

Nievas, C.I., J.J. Bommer & H. Crowley (2017). *Global occurrence and impact of small-to-medium magnitude earthquakes: a statistical analysis (Version 1)*. Report for the Nederlandse Aardolie Maatschappij BV (NAM), The Netherlands.

Nievas, C.I., M. Ntinalexis, D. Kazantzidou-Firtinidou, J. Borozan, M. Sangirardi, H. Crowley & J.J. Bommer (2019). *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.

Orlecka-Sikora, B., E.E. Papadimitriou & G. Kwiatek (2009). A study of the interaction among mining-induced seismic events in the Legnica-Głogów Copper District, Poland. *Acta Geophysica* **57**(2), 413-434.

Pagani M., D. Monelli, G. Weatherill, L. Danciu, H. Crowley, V. Silva, P. Henshaw, L. Butler, M. Nastasi, L. Panzeri, M. Simionato & D. Vigano (2014). OpenQuake engine: an open hazard (and risk) software for the Global Earthquake Model. *Seismological Research Letters* **85**, 692–702.

Papazachos, B.C., A.A. Kiratzi & B.G. Karacostas (1997). Towards a homogeneous moment-magnitude determination for earthquakes in Greece and the surrounding area. *Bulletin of the Seismological Society of America* **87**(2), 474–483.

Petersen, M.D., C.S. Mueller, M.P. Moschetti, S.M. Hoover, A.L. Llenos, W.L. Ellsworth, A.J. Michael, J.L. Rubinstein, A.F. McGarr & K.S. Rukstales (2016). Seismic-hazard forecast for 2016 including induced and natural earthquakes in the Central and Eastern United States. *Seismological Research Letters* **87**(6), 1327-1341.

Scordilis, E. M. (2006). Empirical global relations converting M_S and m_b to moment magnitude. *Journal of Seismology* **10**, 225–236.

Schlumberger (2018a). *Oilfield Glossary: primary recovery*. Available online at (last accessed 9th October 2018): https://www.glossary.oilfield.slb.com/en/Terms/p/primary_recovery.aspx.

Schlumberger (2018b). *Oilfield Glossary: secondary recovery*. Available online at (last accessed 9th October 2018): https://www.glossary.oilfield.slb.com/en/Terms/s/secondary_recovery.aspx.

Stafford, P.J. (2014). Source-scaling relationships for the simulation of rupture geometry within Probabilistic Seismic-Hazard Analysis. *Bulletin of the Seismological Society of America* **104**(4), 1620–1635.

Storchak, D., D. Di Giacomo, E. Engdahl, J. Harris, I. Bondár, W. Lee, P. Bormann & A. Villaseñor (2015). The ISC-GEM Global Instrumental Earthquake Catalogue (1900–2009): Introduction. *Physics of the Earth and Planetary Interiors* **239**, 48–63.

Taylor, O.-D.S., A.P. Lester, T.A. Lee III & M.H. McKenna (2018). Can repetitive small magnitude-induced seismic events actually cause damage? *Advances in Civil Engineering* 2018, 5 pp.

Uhrhammer, R. (1986). Characteristics of Northern and Central California Seismicity. *Earthquake 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 Brodock Press, Inc. Available online at: https://www.unicef.org/aids/files/SOWC_2012-Main_Report_EN_21Dec2011.pdf. Last accessed 7th January 2019.

Unites States Census Bureau. *2010 Census urban and rural classification and urban area criteria*. Available online at: https://www.census.gov/geo/reference/ua/urban-rural-2010.html. Last accessed 7th January 2019.

United States Geological Survey (USGS) (2018). *Why do so many earthquakes occur at a depth of 10km?* Short online note, available at: https://www.usgs.gov/faqs/why-do-so-many-earthquakes-occur-a-depth-10km. Last accessed 19th October 2018.

Văcăreanu, R., M. Iancovici, C. Neagu & F. Pavel (2015). Macroseismic intensity prediction equations for Vrancea intermediate-depth seismic source. *Natural Hazards* **79**(3), 2005-2031.

van Elk, J., S.J. Bourne, S.J. Oates, J.J. Bommer, R. Pinho & H. Crowley (2019). A probabilistic model to evaluate options for mitigating induced seismic risk. *Earthquake Spectra* **35**(2), 537-564.

van Elk, J., A. Mar-Or, L. Geurtsen, P. Valvatne, E. Kuperus & D. Doornhof (2018). *Seismic risk assessment for a selection of gas production scenarios for the Groningen field – Addendum to: Induced seismicity in Groningen assessment of hazard, building damage and risk (November 2017)*. Report for the Nederlandse Aardolie Maatschappij BV (NAM), The Netherlands.

van Stiphout, T., J. Zhuang & D. Marsan (2012). *Theme V - Models and techniques for analysing seismicity*. Technical report. Community Online Resource for Statistical Seismicity Analysis: http://www.corssa.org.

Vilanova, S.P., J.F.B.D. Fonseca & C. Oliveira (2012). Ground-motion models for Seismic-Hazard Assessment in Western Iberia: Constraints from instrumental data and intensity observations. *Bulletin of the Seismological Society of America* **102**(1), 169–184.

Weatherill, G. A. (2014). *OpenQuake Hazard Modeller's Toolkit - User Guide*. Global Earthquake Model (GEM), Technical Report.

Weatherill, G.A., M. Pagani & J. García (2016). Exploring earthquake databases for the creation of magnitude-homogeneous catalogues: tools for application on a regional and global scale. *Geophysical Journal International* **206**, 1652-1676. WPG16v1 World Catalogue (original) version of 2016. WPG16v2 World Catalogue version of 5th July 2017. WPG16v3b World Catalogue version of 21st July 2017. WPG16v3c World Catalogue version of 19th December 2018.

Wilson, M.P., G.R. Foulger, J.G. Gluyas, R.J. Davies & B.R. Julian (2017). HiQuake: The human-induced earthquake database. *Seismological Research Letters*. DOI: 10.1785/0220170112.

Wood, H.O. & F. Neumann (1931). Modified Mercalli Intensity of 1931. *Seismological Society of America Bulletin* **21**(4), 277-283.

Wooddell, K.E. & N.A. Abrahamson (2014). Classification of main shocks and aftershocks in the NGA-West2 Database. *Earthquake Spectra* **30**(3),1257–1267.

Young, J.B., B.W. Presgrave, H. Aichele, D.A. Wiens & E.A. Flinn (1996). The Flinn-Engdahl regionalization scheme: the 1995 revision. *Physics of the Earth and Planetary Interiors* **96**, 223-297.

8.2 Web References

Earthquake Impact Database (EID):
https://docs.google.com/spreadsheets/d/1oveZ42OLdJFnKu2aZfMg4aV1AR_OgHftxRm-65bJw4w/edit#gid=0. Last accessed: 5th January 2018 (year 2017), 22nd November 2017 (years 2013-2016). Additional information: https://earthquake-report.com/2018/02/10/earthquake-impact-database-2018/. Last accessed: 20th November 2018.
Facebook page: https://www.facebook.com/earthquakeimpactdatabase/. Last accessed: 27th March 2018.

EM-DAT – The Emergency Events Database (Université Catholique de Louvain, Brussels, Belgium; Cred. Prof. Dr. D. Guha-Sapir): http://www.emdat.be/. Last accessed: 15th November 2017.

Eurostat – Urban-rural typology (last accessed 13th September 2017):
http://ec.europa.eu/eurostat/statistics-explained/index.php/Urban-rural_typology.

Global Centroid Moment Tensor Project (GCMT): https://www.globalcmt.org/.

Induced Earthquakes (2017): http://inducedearthquakes.org/. Last accessed: 28th November 2017.

International Seismological Centre (ISC), Online Bulletin: http://www.isc.ac.uk/iscbulletin/search/. *Internatl. Seismol. Cent.*, Thatcham, United Kingdom, 2014. Downloaded on 21st September 2018.
List of agencies contributing with the ISC Bulletin (last accessed 17th August 2017):
http://www.isc.ac.uk/iscbulletin/agencies/.
Review procedure (last accessed 8th November 2017): http://www.isc.ac.uk/iscbulletin/review/.

National Geophysical Data Center / World Data Service (NGDC/WDS): Significant Earthquake Database. National Geophysical Data Center, National Oceanic and Atmospheric Admnistration (NOAA). DOI:10.7289/V5TD9V7K. https://www.ngdc.noaa.gov/hazard/earthqk.shtml. Last accessed: 15th November 2017.

Robbins (2018): http://www.therobbinscompany.com/projects/west-qinling-rail-tunnels/. Last accesed: 27th September 2018.

Structurae (2018): https://structurae.net/structures/kan-etsu-tunnel. Last accesed: 27th September 2018.

UNICEF – The State of the World's Children 2012 – Definitions (last accessed 13th September 2017):
https://www.unicef.org/sowc2012/pdfs/SOWC-2012-DEFINITIONS.pdf.

Unites States' Census Bureau - 2010 Census Urban and Rural Classification and Urban Area Criteria: https://www.census.gov/geo/reference/ua/urban-rural-2010.html. Last accessed: 13th September 2017.

Wikipedia:
Erlangshan Tunnel: https://en.wikipedia.org/wiki/Mount_Erlang_Tunnel. Last accesed: 27th September 2018.
Shimizu Tunnel: https://en.wikipedia.org/wiki/Shimizu_Tunnel. Last accesed: 27th September 2018.
Simplon Tunnel: https://en.wikipedia.org/wiki/Simplon_Tunnel. Last accesed: 27th September 2018.

8.3 Other Resources

Earthquake catalogues processed with the aid of:

- Tools available in OpenQuake (Pagani *et al.*, 2014).
- The OpenQuake Hazard Modeller's Toolkit (Weatherill, 2014).
- The toolkit published alongside the paper of Weatherill *et al.* (2016).

QGIS free and open source Geographic Information System: http://qgis.org/.

Shapefiles of countries' administrative boundaries from DIVA-GIS: http://www.diva-gis.org/gdata.

Shapefiles of continents' boundaries:
https://www.arcgis.com/home/item.html?id=5cf4f223c4a642eb9aa7ae1216a04372.

World maps data copyrighted by OpenStreetMap contributors. Available from https://www.openstreetmap.org.

APPENDIX I: SUPPORTING INFORMATION FOR THE DEFINITION OF THE PROBABILITY THRESHOLDS

I.1. Maximum Magnitude Uncertainty

When defining the probability thresholds in Section 4.6.4 it was necessary to consider extreme values of uncertainty in depth and magnitude. For the case of depth the choice was simple, as the probability thresholds are defined using the maximum depth accepted for each magnitude and this leads to uncertainties tending to zero in order to be able to satisfy the maximum depth criterion.

For the case of magnitude, uncertainties stem from the combination of measurement errors and the uncertainty of empirical conversion models, as per Equation 4.16. The following considerations were made in order to select a value:

- Values found in the merged catalogue before filtering according to the exposure criterion (after filtering according to the maximum depth criterion): The maximum is 1.234, though extreme values like this are rarely observed. The median is 0.263, the 95th percentile 0.462 and the 99th percentile 0.548. Figure I.1 shows the distribution of values observed.

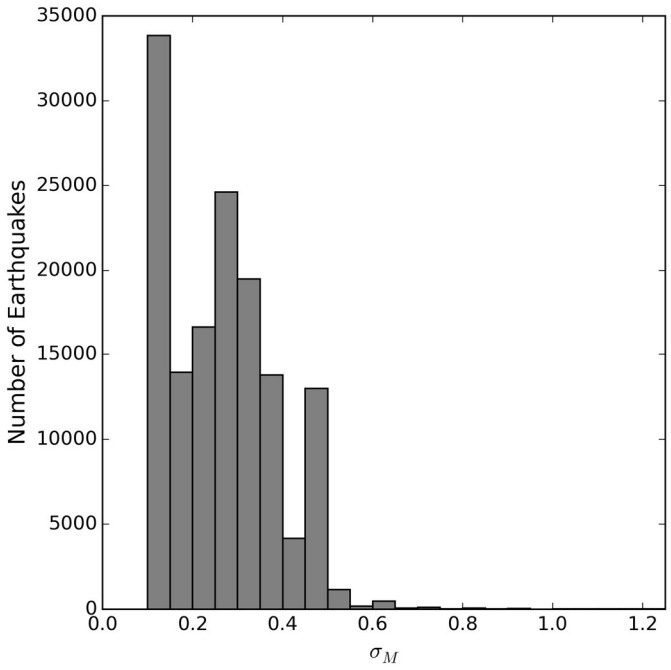

Figure I.1. Distribution of values of final uncertainty in magnitude (obtained from the combination of measurement errors and the uncertainty of conversion models) for the 141,524 earthquakes in the merged catalogue before the application of the exposure criterion.

- Values found in WPG16v3c for the period of interest (1 January 2001 – 31 December 2015): The three largest values are 1.657, 1.550 and 1.339. The former is not part of

the merged catalogue because the associated magnitude is **M**3.87, while the two latter are not included because of being too deep. The fourth largest value is 1.234, which becomes the largest in the merged catalogue as stated in the previous point.

- Maximum values that can occur when converting non-**M** values reported in the ISC Bulletin into **M** by means of empirical models: Having imposed a maximum measurement error of 1.0 (discarding values with larger errors), the largest possible final uncertainty is 1.162, which corresponds to a standard deviation of 0.317 of the m_b-to-**M** average conversion model, a model derivative of 1.118 (average slope), and a measurement error of 1.0.

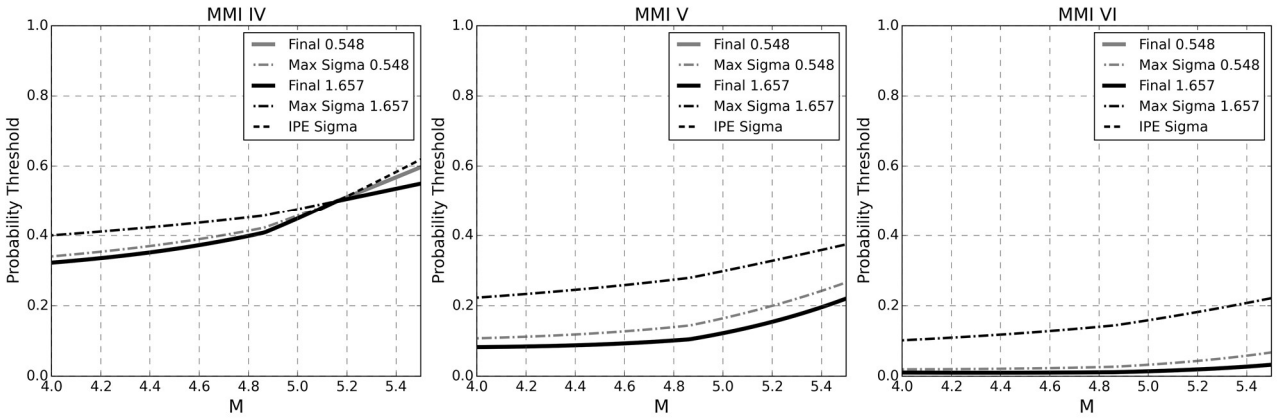

Figure I.2. Comparison of probability thresholds for MMI≥IV (left), MMI≥V (centre) and MMI≥VI (right) using a measurement error in depth of 0.001% of the maximum depth and two alternative values of final uncertainty in moment magnitude: 0.548 (grey) and 1.657 (black).

The adoption of high values such as 1.234–1.657 seems extreme when it is clear from Figure I.1 that they are not common. However, adoption of smaller values could result in discarding earthquakes that would otherwise be reasonably eligible. In order to understand the extent to which this decision can affect the probability thresholds adopted, a small comparison was carried out using the largest of the values found (*i.e.*, 1.657) and 0.548, the 99th percentile of the merged catalogue, the smallest value that would seem reasonable to adopt from the observation of Figure I.1. Dashed lines in Figure I.2 show the probability thresholds obtained accounting only for the uncertainty in the IPE, while dash-dotted lines show those obtained when including maximum uncertainties in depth and magnitude, and thick continuous lines indicate what would be the adopted as thresholds, that is, the minimum of the two. As can be observed, the probabilities arising from considering only the uncertainty of the IPE dominate the final thresholds for MMI≥V and MMI≥VI (the continuous black lines cover the continuous grey lines and dashed lines in the plot). In the case of MMI≥IV, uncertainty in magnitude becomes relevant only for magnitudes above around **M**5.1. In other words, the only difference between adopting a value as high as 1.657 and a more reasonable one of 0.548 is the difference between the continuous black and grey lines for **M**>5.1 in the plot on the left of Figure I.2 (MMI≥IV). Given the small overall difference and the risk of discarding earthquakes lying on the more potentially damaging side of the

spectrum (because of having larger magnitudes) that stems from using a smaller value, 1.657 was finally adopted.

I.2. Should the Thresholds Depend on the Craton Index?

MMI probability thresholds were defined in Section 4.6.4 under the premise that they should guarantee that an earthquake that occurred at the maximum depth considered for its magnitude satisfies them at least right above the epicentre. Having used the craton index to allow for intensity predictions to take into account the different attenuation properties displayed by cratonic and non-cratonic environments, whether a certain earthquake causes a certain MMI at the epicentre depends on the craton index. However, if the probability thresholds were to be dependent on the craton index, a non-desirable behaviour would arise, as will be explained in what follows.

Figure I.3 extends Figure 4.36 of Section 4.6.4, generated with a craton index of zero, to additional cases of craton indices of 1 (centre plot) and 0.5 (right-most plot). As can be observed, the thresholds, that is, the probabilities of observing MMI≥IV, or MMI≥V, or MMI≥VI at the epicentre with an earthquake of a certain magnitude located at the maximum possible depth according to the depth criterion of Section 4.5.2, increase for increasing values of the craton index. This stems from the smaller attenuation of cratonic environments, which is reflected in the behaviour of the IPE of Atkinson & Wald (2007), which is depicted in Figure I.4. Applying craton index-dependent probability thresholds would imply, for example, that an earthquake with magnitude **M**4.0 should have a probability of at least 32% of generating MMI≥IV at the epicentre in a non-cratonic environment and of at least 54% in a cratonic environment to have a chance of being considered in the catalogue (the final inclusion depends on the population exposed to probabilities of MMI≥IV larger than the threshold).

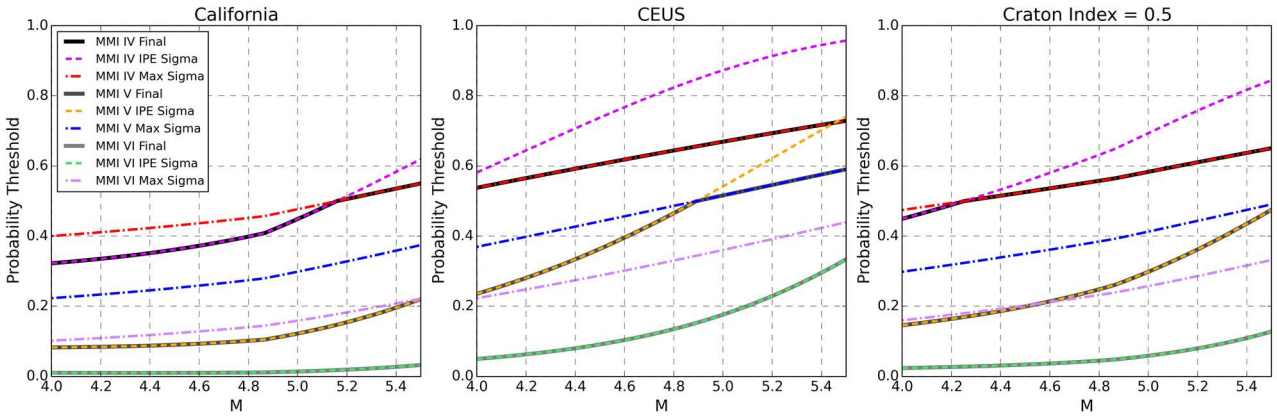

Figure I.3. Craton index-dependent probability thresholds for different MMI levels, uncertainty cases and three values of the craton index: 0 (left), 1 (centre) and 0.5 (right). What would constitute final adopted thresholds are shown in continuous black and grey lines. The plot on the left is the same as Figure 4.36 in Section 4.6.4.

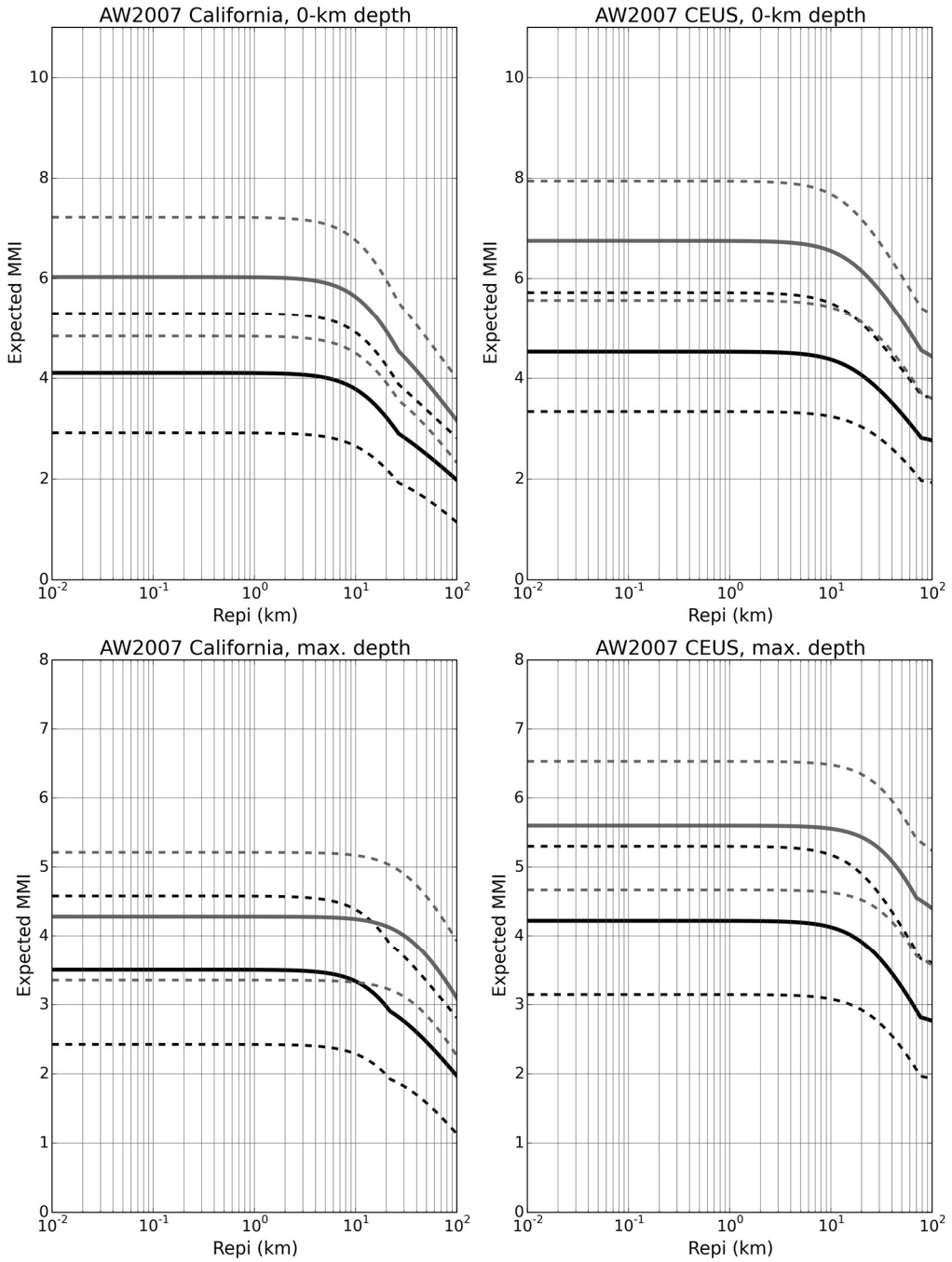

Figure I.4. Behaviour of the IPE of Atkinson & Wald (2007) with coefficients for California (left) and the CEUS (right), for a depth of 0 km (top) and equal to the maximum depth (bottom) corresponding to the two magnitudes considered: **M**4.0 (black lines) and **M**5.5 (grey lines). Models applied out of applicability range for hypocentral distances smaller than approximately 10 km.

The plots in Figure I.5 depict the maximum epicentral distances that satisfy the condition that the probability of MMI being equal to or larger than IV (top), V (centre) and VI (bottom) is larger than the thresholds defined in Figure I.3, for different uncertainties in magnitude and craton indices of 0 (left) and 1 (right). All these plots were created considering a null depth and a 0.1-km depth error, as this would yield the largest epicentral distances. These plots indicate, for example, that if the threshold for MMI V was used as a criterion, radii up to around 20 km would be considered for a **M**4.0 and of up to around 50 km would be

considered for a **M**5.0 in a non-cratonic setting (centre plot on the left of Figure I.5) when evaluating the population exposed.

As it may catch the reader's attention, it is noted that the crossing of results for different uncertainties in magnitude that occurs for MMI≥IV for a craton index of 0 (Figure I.5, top left) and for MMI≥V for a craton index of 1 (Figure I.5, centre right) is due to defining the probability thresholds as the minimum of that resulting from only the IPE uncertainty and that resulting from maximum feasible uncertainties, as explained in Section 4.6.4 and depicted in Figure I.3.

The main feature of the plots in Figure I.5 that should strike the eye is, however, the fact that the maximum epicentral distances are larger for a craton index = 1 than for a craton index = 0 only when MMI≥IV is considered, but not for the case of MMI≥V or MMI≥VI, the tendency being more extreme for the largest values of uncertainty in magnitude considered. In order to highlight this behaviour, some of the curves from Figure I.5 have been plotted again in Figure I.6, but superimposing those for Craton Indices of 0 (continuous lines) and 1 (dashed lines). The focus has been set on values of uncertainty in magnitude equal to or smaller than 0.5, which represent a large percentage of the merged catalogue. Given the lower attenuation of cratonic environments, one would expect the maximum epicentral distances to be always larger for a craton index of 1 than for a craton index of 0. In other words, one would expect the area that satisfies that the probability of MMI≥V to be larger in a cratonic environment than a non-cratonic environment. However, establishing craton index-dependent probability thresholds leads to this contradiction. For this reason, the probability thresholds corresponding to a craton index of 0 (California coefficients, left plot of Figure I.3) have been adopted independently from the location of the earthquake. Figure I.6 then becomes Figure 4.37 (Section 4.6.4), and maximum epicentral distances increase with increasing values of the craton index, as desired.

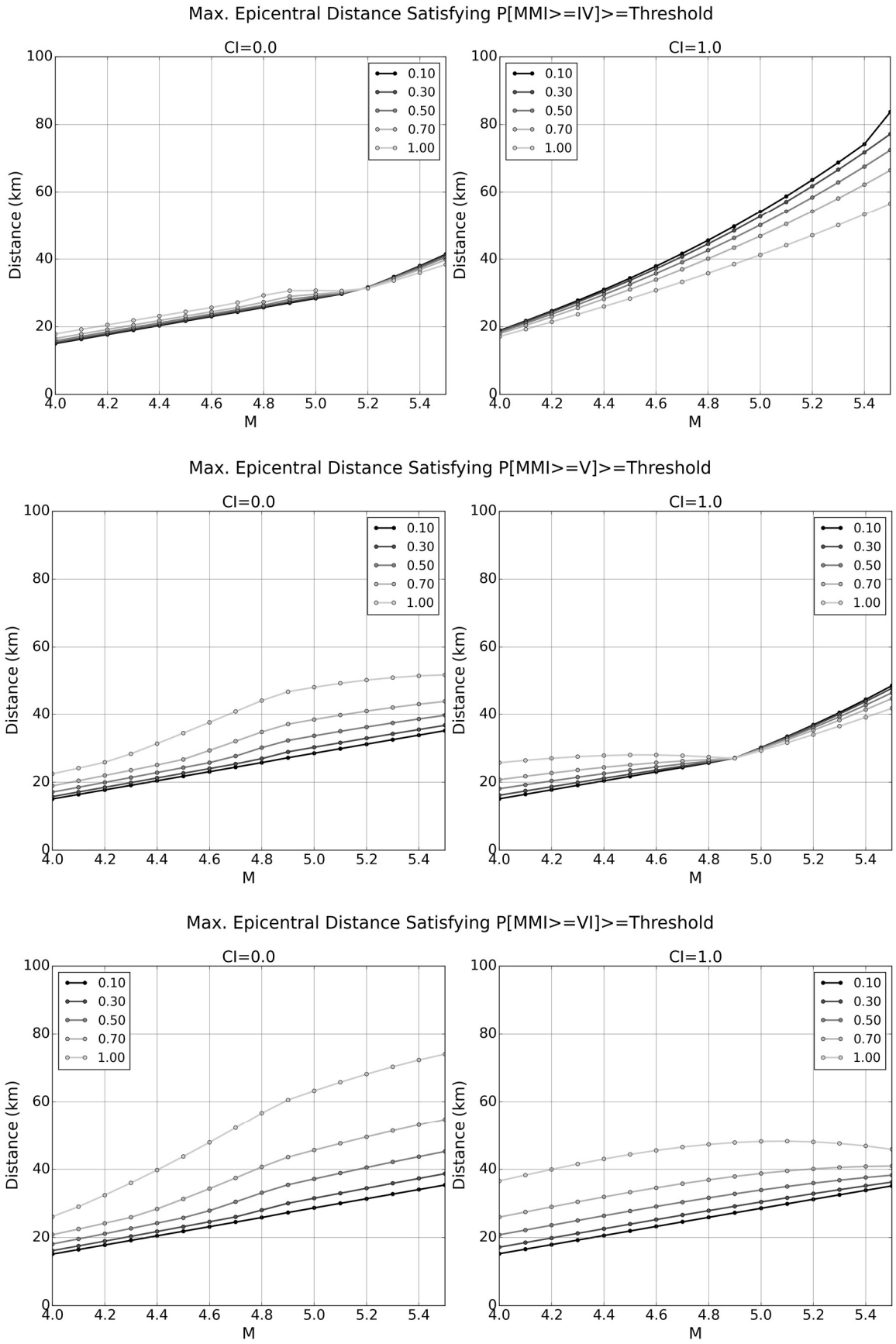

Figure I.5. Maximum epicentral distances satisfying the condition that the probability of MMI being equal to or larger than IV (top), V (centre) and VI (bottom) is larger than the thresholds defined in Figure 4.36, for different uncertainties in magnitude (colour scale). craton index = 0 (California coefficients, left) and craton index = 1 (CEUS coefficients, right).

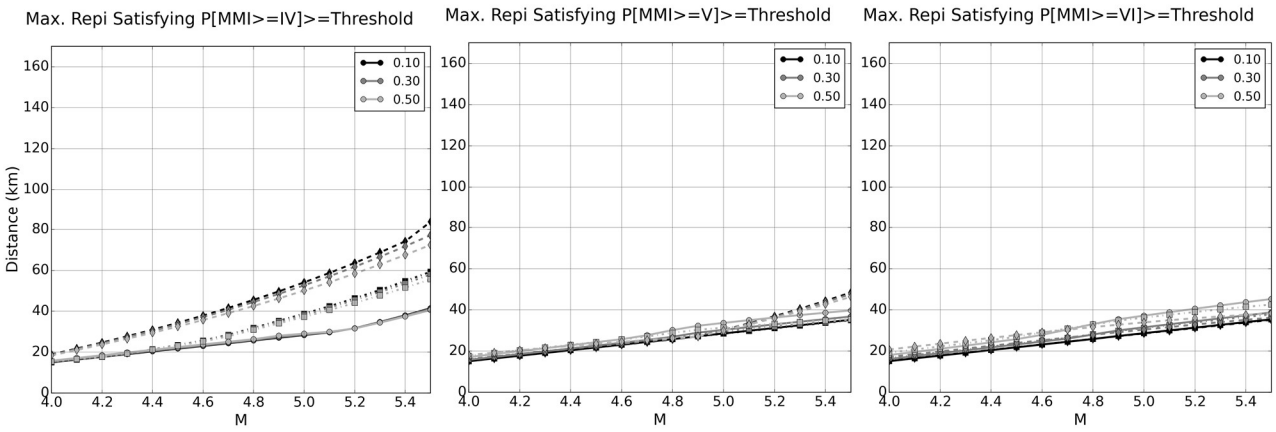

Figure I.6. Maximum epicentral distances satisfying the condition that the probability of MMI being equal to or larger than IV (left), V (centre) and VI (right) is larger than the thresholds defined in Figure 4.36, for different uncertainties in magnitude (colour scale). Continuous, dotted and dashed lines indicate craton index = 0 (California coefficients), craton index = 0.5, and craton index = 1 (CEUS coefficients), respectively. Thresholds dependent on craton index.

I.3. Maximum Epicentral Distances

As discussed in Section 4.6.4 and depicted in Figure 4.37, the exposure criterion translates into the definition of maximum epicentral distances within which the population count and population density is investigated. While Figures 4.37 and I.6 investigate the resulting values of maximum epicentral distances with reasonable uncertainties in magnitude (0.1-0.5, see Figure I.1), their behaviour in more extreme cases is investigated herein. Figures I.7 and I.8 depict the maximum epicentral distances resulting from considering uncertainties in magnitude $\sigma_{\mathbf{M}}$ of 0.263 (median of the merged catalogue[16]), 0.548 (99th percentile of the merged catalogue), 1.0, 1.234 (maximum of the merged catalogue), 1.657 (maximum of the WPG16v3c catalogue), and 2.0 (0-km hypocentral depth assumed in all cases). As can be observed, the behaviour of the relationship between the maximum epicentral distances and magnitude is reasonable for $\sigma_{\mathbf{M}}$ of 0.263 and 0.548, which cover 99% of the merged catalogue, but deviates from the expected trends for increasing values of $\sigma_{\mathbf{M}}$. While still reasonable overall for $\sigma_{\mathbf{M}}$ of 1.0 and 1.234, larger values of $\sigma_{\mathbf{M}}$ cause the maximum epicentral distances to reach a peak, decrease and increase again with increasing **M**. This would imply, for example, that an earthquake with **M**5.0 is treated as capable of producing damage at longer distances that an earthquake with **M**5.5, when considering the probability of exceeding MMI V (bottom row of Figures I.7 and I.8), which is clearly not true.

[16] In this sentence, "merged catalogue" makes reference to the merged catalogue filtered according to the maximum depth criteria and for the time period and magnitude range of interest.

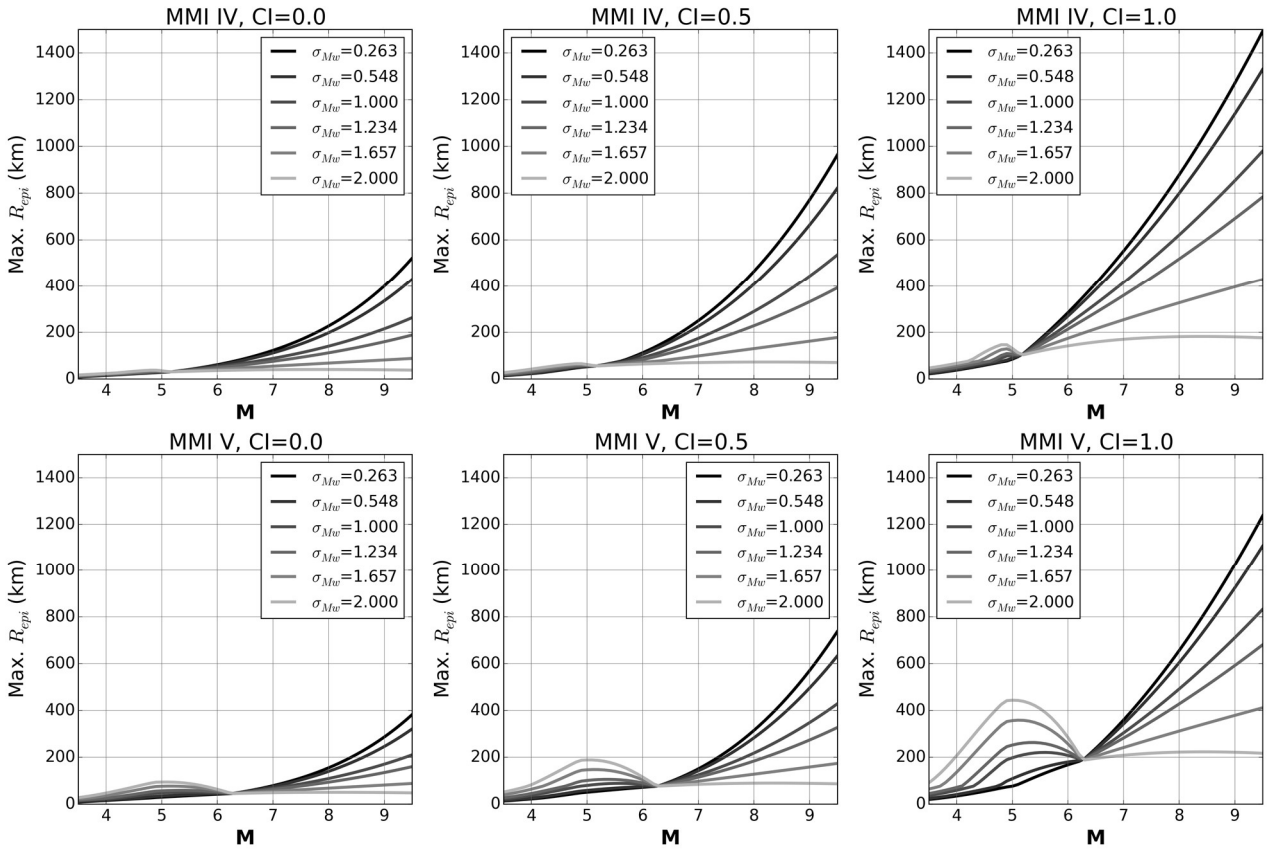

Figure I.7. Maximum epicentral distances satisfying the condition that the probability of MMI being equal to or larger than IV (top) and V (bottom) is larger than the thresholds defined in Figure 4.36, for different uncertainties in magnitude (colour scale) and values of the craton index.

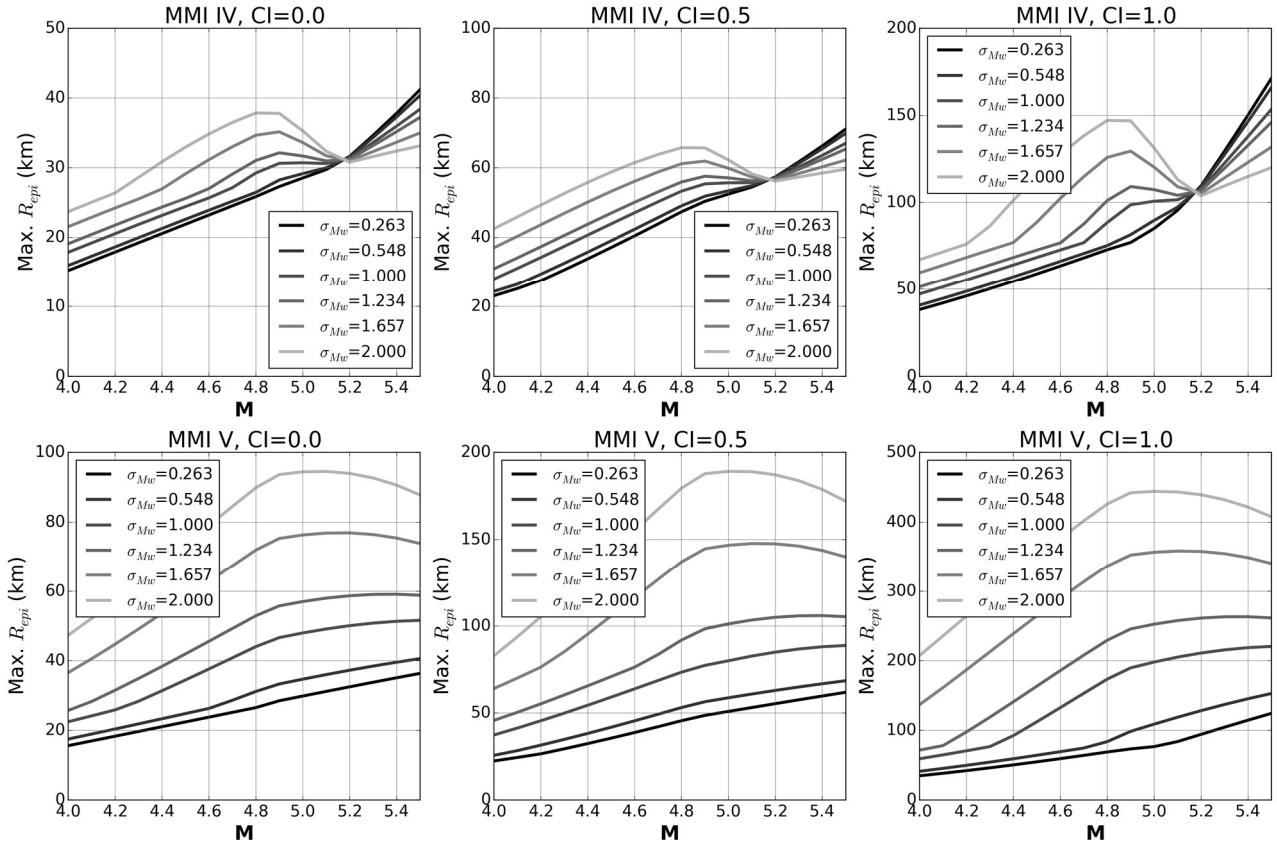

Figure I.8. Detail of Figure I.7 for the magnitude range **M**4.0-5.5.

This behaviour appears to emerge from a combination of the non-linearity and abrupt changes in the slopes of the probability thresholds defined in Figure 4.31, the mathematical shape of the derivative of the IPE of Atkinson & Wald (2007) with respect to magnitude, which intervenes in the calculation of the uncertainty in MMI as described by Equation 4.19, and the larger values of uncertainty of the IPE for smaller hypocentral distances (see Figure 4.27). However, it is clear that it is quite large uncertainties in magnitude ($\sigma_{\mathbf{M}}$) that trigger it. This last point is fundamental. Uncertainties in magnitude of 1.5 or 2.0 units imply that an earthquake with reported/assigned **M**4.0 could easily be a **M**2.5 or a **M**5.5 and the problem is evident: the damaging capacity of one is extremely different from that of the other, as the energy released by the **M**5.5 is over 31,000 times that released by the **M**2.5. Table I.1 illustrates the impact of the uncertainty in magnitude over the final uncertainty in MMI. As can be observed, an uncertainty in magnitude of 1 unit results in very large uncertainties in MMI of 1.5 to 2.0 units at 15 km from the epicentre for earthquakes with reported magnitudes in the range of interest (4.0-5.5). However, uncertainties in magnitude of this calibre would correspond to extreme outlier cases in which the earthquake is very poorly constrained, and it is thus not rare that the model cannot perform well under such circumstances.

Table I.1. Uncertainties in MMI calculated with the IPE of Atkinson & Wald (2007) for a 0-km hypocentral depth, two epicentral distances and three values of $\sigma_{\mathbf{M}}$.

CI	**M**	**$\sigma_{\mathbf{M}}$ = 0.263**		**$\sigma_{\mathbf{M}}$ = 0.548**		**$\sigma_{\mathbf{M}}$ = 1.000**	
	R_{epi}:	**15 km**	**50 km**	**15 km**	**50 km**	**15 km**	**50 km**
0.0	4.0	1.11	0.91	1.21	0.98	1.47	1.17
	5.5	1.14	0.93	1.32	1.09	1.75	1.45
1.0	4.0	1.13	0.93	1.28	1.06	1.65	1.39
	5.5	1.16	0.96	1.40	1.18	1.93	1.67

The impact of the non-linearity and abrupt changes in the slopes of the probability thresholds defined in Figure 4.31 is visible in the crossing points that appear in the plots in Figures I.7 and I.8 at around **M**5.2 for MMI≥IV and **M**6.2 for MMI≥V, the former being depicted in Figure 4.36. The potential influence of the mathematical shape of the derivative of the IPE of Atkinson & Wald (2007) with respect to magnitude can be inferred by assigning arbitrary epicentral distances to each magnitude, as was done in Figure I.9 (0-km hypocentral depth assumed). As can be observed, the resulting derivatives also appear to behave oddly with respect to magnitude. Analysing the three terms that conform Equation I.1, where c_2, c_3, c_7 and h are coefficients that vary for California and the CEUS, it is clear that the derivative increases linearly with magnitude, and that the final curvature is provided by the last term, which depends on hypocentral depth. The resulting values of the derivative become amplifying factors of the uncertainty in magnitude when they are inserted into Equation 4.19. It is this synergy that appears to cause the peak-decrease-increase behaviour observed in Figures I.7 and I.8 for large values of $\sigma_{\mathbf{M}}$.

$$\frac{\partial IPE}{\partial \boldsymbol{M}} = c_2 + 2c_3(\boldsymbol{M} - 6) + c_7 \log \sqrt{R_{hyp}^2 + h^2} \quad \text{(I.1)}$$

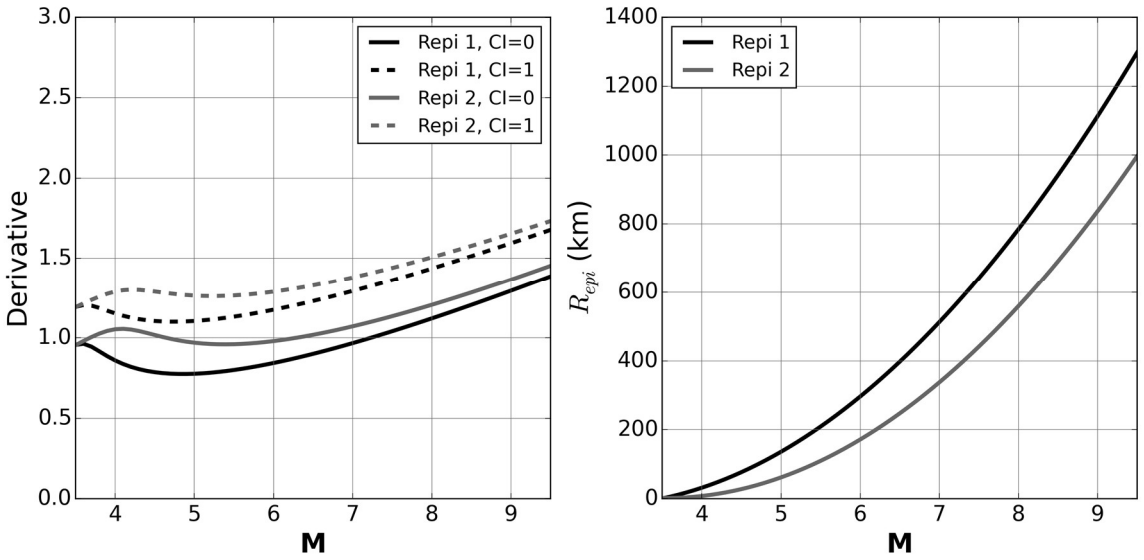

Figure I.9. Derivative of the IPE of Atkinson & Wald (2007) as a function of **M** (left) for magnitude-dependent epicentral distances defined by parabolas (right).

As the undesirable behaviour only occurs for unrealistically large values of uncertainty in magnitude and not for values found in 99% of the merged catalogue, it is concluded that this is not expected to have a significant impact on the results presented.