Earthquakes Parameters from Citizen Testimonies: A Retrospective Analysis of EMSC Database

Published Version:
Gianfranco Vannucci, Paolo Gasperini, Laura Gulia, Barbara Lolli (2023). Earthquakes Parameters from Citizen Testimonies: A Retrospective Analysis of EMSC Database. SEISMOLOGICAL RESEARCH LETTERS, 95(2A), 969-996 [10.1785/0220230245].

Availability:
This version is available at: https://hdl.handle.net/11585/960394 since: 2024-02-22

Published:
DOI: http://doi.org/10.1785/0220230245

Earthquakes parameters from citizen testimonies. A retrospective analysis of EMSC database

Gianfranco Vannucci[1*], Paolo Gasperini[2,1], Laura Gulia[2] and Barbara Lolli[1]

[1]Istituto Nazionale di Geofisica e Vulcanologia, Sezione di Bologna

[2]Dipartimento di Fisica e Astronomia, Università di Bologna

*** Corresponding author**

Declaration of Competing Interests:

The authors acknowledge there are no conflicts of interest recorded.

Abstract

We aim to compute macroseismic parameters (location and magnitude) using the BOXER code for the first time on the citizen testimonies, i.e., individual intensity data points (IDPs) at the global scale collected and made available by the LastQuake system of the European-Mediterranean Seismological Centre (EMSC).

IDPs available for different earthquakes are selected to eliminate those that are geographically inconsistent with most data, then they are clustered spatially based on various methods. For each cluster with at least 3 IDPs, a macroseismic data point (MDP), corresponding to an intensity value assessed for given localities as in classical macroseismic studies, is computed by various central tendency estimators (average, median, trimmed averages). Finally, macroseismic parameters are obtained by MDP distribution using two location methods of BOXER code. For each earthquake, we used raw and corrected intensities and 132 different combinations of grouping methods, estimators and BOXER methods.

We assigned a ranking to the combinations that best reproduce instrumental parameters and used such a ranking to select preferred combinations for each earthquake. We analysed retrospectively the reliability of the parameters as a function of time and space. The results are essentially identical using original and corrected intensities and show higher reliability for BOXER's method 1 than for method 0, they are dependent on the geographical area and generally improves over time and with the number of IDPs collected. These findings are useful for future real-time analyses and for evaluating the location and magnitude of earthquakes whenever a sufficient number of IDPs are available and with a distribution such that MDPs can be derived, and the BOXER method applied.

Introduction

The macroseismic intensity, i.e. the quantification of the severity of the ground motion, based on earthquake effects on humans, objects, natural environment and buildings, is a tool for studying pre-instrumental earthquakes used for seismic hazard assessment and seismic risk mitigation. The intensity assessed by macroseismic experts or other methods (e.g. Vannucci et al., 2015) through macroseismic scales (e.g. MCS -Sieberg, 1912, 1932-, EMS -Grünthal et al., 1998-) is quantified using a damage scenario at the scale of localities and their geographic distribution allows to assess reliable epicentre location and magnitude, using various software codes (e.g. Bakun and Wentworth 1997; Gasperini et al. 1999, 2010; Pettenati and Sirovich 2003; Musson and Jiménez 2008). Gasperini et al. (2010) have shown how macroseismic intensities make it possible to calculate location, magnitude and, in the most favourable cases (e.g., earthquakes with magnitude ≥ 5.7), also the orientation of the source, with an accuracy comparable to instrumental methods. Vannucci et al. (2019) also demonstrated that if the intensities are well distributed and quickly available after the occurrence of the earthquake, they can constrain well the macroseismic source and provide useful information to civil protection and stakeholders even before reliable instrumental data be available. Therefore, the macroseismic intensities do not only provide information on pre-instrumental earthquakes but also on contemporary ones by taking advantage of the geographic abundance of information coming from different localities that are much denser than the instrumental stations. Such data also provide a direct check of the theoretical models of energy propagation (like SHAKEMAP, see data and resource section) for local calibration of expected effects.

Presently, the development of specific software applications allows to collect and elaborate testimonies of the shaking felt by individual citizen. Indeed, since several years, community intensities are collected by different agencies e.g. “Did you feel it?” (DYFI, Wald et al., 1999, 2011, Dewey et al., 2000), of the U.S. Geological Survey (USGS), “Hai sentito il terremoto?” (HSIT, Tosi et al., 2015) of the Istituto Nazionale di Geofisica e Vulcanologia (INGV), New Zealand GeoNet questionnaires (GeoNet, Goded et al., 2018), LastQuake system (Bossu et al., 2015, 2018) of the European-Mediterranean Seismological Centre (EMSC). These data are collected at different spatial scales, and with different methodologies. In particular, individual data points (IDPs), i.e. macroseismic intensity according to the European Macroseismic Scale (EMS98, Grünthal, 1998) and assessed by each eyewitness citizen, are collected and made available by LastQuake system. The IDP database is based on a worldwide community of people, whose number increases over time. Our aim is to use this basic information to develop methods to compute the location and the magnitude of the earthquake.

Through the LastQuake system EMSC collected 1874376 IDPs (with intensity ≥ 2) of 51359 global earthquakes (with magnitude ranging between 0.4 and 8.4) from 2012 to February 2023 (Fig. 1). Such data are freely available at EMSC website (see data and resources section). The number of collected IDPs generally increased over time (Fig. S1 in supplementary material) as the popularity of the application increased and the users became more and more involved in such activity (Bossu et al., 2017). Each collected IDP provides latitude, longitude, raw (R) and corrected (i.e. revaluated) intensity (C). The raw intensity is assessed through the selection by each citizen/observer of thumbnails that best represent the observed seismic effects, i.e. by the correspondence between eyewitness observations and felt scenario representations, while the corrected intensity is computed, according to Bossu et al. (2017), to best reproduce DYFI intensities for a reference dataset of 17 earthquakes.

The number of collected intensities is a decreasing function of their value: the higher the intensity, the lower the number of reports, since higher intensities are generally limited to the areas close to the source (near field), while lower intensities occur at longer geographical distances (far field), with larger numbers of people and reports. This trend is generally valid for raw intensities except for the extreme intensities 2 and 12 (Fig. 1).

Based on the geometric spreading of seismic energy, the effects of the earthquake should "ideally" propagate in any direction from the epicentre. However, cities and citizens are not evenly distributed throughout the territory and the distribution of IDPs suffers sometimes from the lack of coverage in uninhabited areas. In general, the greater the earthquake magnitude the wider the area of effects and the higher the number of felt reports, but both the number and the distribution of IDPs are subject to a number of factors: geomorphological ones (presence of seas, lakes, mountains, deserts), demographic ones (variable population density, presence or absence of cities), technological ones (internet coverage) and political ones (free or equitable access to internet, e.g. Hough and Martin 2021).

The lack of IDPs in the epicentral area for earthquakes of strong magnitude and destructive effects could even be due just to the strength of such effects (e.g., destruction of buildings, infrastructures and casualties) that might prevent the people to pay attention to the reports so leaving empty the epicentral zone ("doughnut effect") (Bossu et al., 2018). The IDPs may be absent where restrictive policies on the use of smartphone applications are in force (e.g., in China, North Korea etc., see Fig. 1). Hence, in some regions of the world where earthquakes are known to occur but where there are only a few IDPs (Fig. 1), the distribution of IDPs can be uneven: IDPs are not well distributed around the epicentre so that the maximum azimuthal gap of IDPs with respect to the epicentre is larger than 180 degrees.

Another factor to consider is the presence of some IDPs that are inconsistent with the distribution of the most of other ones. These anomalous IDPs can be divided into two types: "intensity outliers" and "geographic outliers".

Intensity outliers are IDPs for which the assigned intensity values appear significantly inconsistent with respect to the other ones in the neighbour. They can be due to a) the wrong judgement by the citizen who has emphasised the effects for emotional reasons or misjudgement, in most cases overestimating the macroseismic intensity; b) misreporting of intensity with the selection of the last of the available thumbnails in LastQuake system, which might also explain the unreasonably high frequency observed for the degree 12 of the raw intensity (Fig. 1).

Geographic outliers are IDPs located in areas far from the instrumental epicentre. They could be due to various reasons: a) reports sent from a computer (not a smartphone) for which there is a wrong reporting of the geographical position due to the link to fixed network servers located up to tens of kilometres away from the observing site; b) use of Virtual Private Network (VPN) with geo-location up to thousands of kilometres away; c) persons reporting an earthquake and its intensity on behalf of others, so associating the information with the geo-referenced location of the reporting smartphone; d) bad association between felt report and event; e) shocks due to other causes (e.g. quarry blasts). Geographic outliers, if present, are generally a small fraction of the total number of IDPs. Their presence in most cases enlarges the area covered by testimonies. During periods of intense seismicity (seismic sequences), the IDPs can be erroneously attributed to another shock occurred almost simultaneously but located at long distance, generating intensity and geographical outliers. However, only very few earthquakes show a totally inconsistent association between instrumental epicentre and location of IDPs, i.e., IDPs are located too far from the epicentre and cannot represent its effects, making these earthquakes unreliable and unusable for further analyses. The number of these unreliable epicentre-IDPs associations decreased in the course of time, probably owing to the increasing consciousness of people submitting their reports and to a more careful use of the application by the users.

Crowdsourcing projects such as DYFI (Wald et al., 2011) and HSIT (Tosi et al., 2015) collect intensities by citizens based on written questionnaires and have already approached the problem of outliers by grouping single reports and derive intensities at geographical localities as commonly done in standard macroseismic surveys. Geographical outliers can be detected on the basis of empirical magnitude-distance relationships, evidencing intensities at anomalous distances from the epicentre, while possible intensity outliers can be filtered out from the felt scenario by imposing an intensity threshold (e.g. <11, as in Bossu et al. 2017). This can be justified by considering that the assessment of very heavy damage or destruction by citizens involved in them is unlikely because they usually do not pay attention to sending smartphone reports while they are in danger of life. Following this approach, EMSC always consider degrees 11 and 12 as outliers and provides corrected intensities by eliminating these values. However, the remaining, high, intensities (e.g. 8, 9, 10), which define very severe and general damage, may also be unreliable, thus representing anomalous intensity values anyway.

Both DYFI and HSIT join individual IDPs using ZIP codes and municipal territories to obtain the MDPs, but, his is not possible for LastQuake data because they are provided at a global scale where such geographical subdivisions are not available.

Method: from IDP to macroseismic parameters

The distribution of IDPs provides a reasonable indication at a glance of the area of the effects and of the possible epicentre location. To compute the earthquake parameters such as location and magnitude, we use the BOXER code (Gasperini et al., 1999, 2010), a software widely used for macroseismic analysis for present (e.g., Vannucci et al., 2019) and past earthquakes (e.g., Rovida et al., 2020). However, the use of single IDP is too sensitive to the presence of outliers and then it is preferable to use instead Macroseismic Data Points (MDPs), i.e., intensities assigned to clusters of IDPs. We adopt the term MDP in analogy to an intensity value assessed for given localities as in classical macroseismic studies, although it has a different origin. Therefore, the quantitative computation of macroseismic parameters follows two main steps (Fig. 2): first, the grouping of IDPs and the assessment of intensity on MDPs; second, the processing of MDPs to compute location and magnitude of the earthquake by BOXER.

Starting from the IDP distribution, we use an original code to constrain the area within which to select IDPs and outside which to eliminate geographic outliers. The IDPs are grouped using different grouping methods. If the number of IDPs in each cluster is larger than a given minimum threshold (e.g. 3, 5), the MDP intensities are computed by various statistical estimators of central tendency as for example the average, the median and the trimmed mean, so reducing the effects of intensity outliers. All MDP intensities are finally processed by the BOXER code to obtain macroseismic parameters and their uncertainties (Fig. 2).

In this retrospective analysis of the EMSC database we clustered IDPs to derive MDPs both using raw and corrected intensities.

We discarded all intensities 2 and 12 thus reducing the range of the raw intensities to the interval from 3 to 11. This because even for true intensity estimates made by macroseismic experts, intensity 2 corresponds to a so weak perception of ground shaking (felt by very few people in particularly receptive conditions indoors) that it might remain unobserved in most cases, and it is also difficult to be distinguished from degree 3 (felt by few people indoors). For example, Bakun and Wentworth (1997), for their location and sizing method, choose to aggregate the degree 2 with degree 3, while Tosi et al. (2015) considered degree 2 equivalent to "not-felt" (degree 1) for HSIT data. We therefore preferred to simply discard degree 2, even considering the lower reliability of our intensities based on citizen testimonies. On the other hand, true intensities 12 were really never observed.

a) Classification of EMSC events

To select IDPs useful for computations and statistical retrospective analyses, we must firstly eliminate possible geographical outliers. In this retrospective analysis, we use the known instrumental epicentre and magnitude to constrain the geographic area of IDP coverage by a Maximum Distance Prediction Equation (MDPE, Fig. 3), an empirical function, aimed at discarding only the furthest geographical outliers, that links the magnitude of an event with the maximum distance of IDPs with respect to epicentre:

$$MDPE = a + b * M + \exp{(c * M)} \quad (1)$$

where M is the magnitude and a=50, b=70, c=0.9 are fixed coefficients defined empirically by a trial-and-error procedure. The purpose is a quick preliminary selection of IDPs, deleting those located at distances longer than that predicted by the MDPE (Fig. 3) in order to significantly reduce the time required to assess the MDPs, considering the retrospective analysis of thousands of earthquakes of the EMSC dataset. The maxima and minima of the latitudes and longitudes of the IDPs within the MDPE radius defines a rectangular area for next elaborations (solid black lines in Fig. 3). For the sake of clarity, the application of the MDPE equation 1 and the filtering of geographical outliers is only done for this retrospective analysis, whereas for event by event future near real-time analyses the procedure requires no filter and knowledge of location and instrumental magnitude (see Appendix A for more details).

We classify the earthquakes considering if:

a) the epicentre is located inland or offshore;

b) the epicentre is located in or out the defined rectangular area;

c) there are geographic outliers. Consequently, we assign a two-character code: the first one indicating whether the epicentre is inland or offshore (L or S, respectively), while the second one is:

1: if the epicentre is inside the area, without outliers;

2: if the epicentre is inside the area, with outliers;

3: if the epicentre is outside the area, without outliers;

4: if the epicentre is outside the area, with outliers.

We provide in Figure 4 a scheme of the 8 main categories, identified by various codes (e.g., L1, L4, S3, …), some real examples of earthquakes classified following the previous scheme are plotted in Fig. S2 of the supplementary material.

b) From IDPs to MDPs: Data clustering

This procedure (see details in Appendix A) is structured in three steps (Fig. 5):

A) definition of spatial areas or clusters where grouping IDPs;

B) evaluation of the occurrence of IDPs in each spatial area or cluster;

C) assessment of MDPs.

Step A ("IDPs in/out" in Fig. 5): IDPs available for each earthquake can be clustered using different methods: within a given radius (RA), over a square grid (SQ), over a hexagonal grid (HE), within a radius and over a square grid (RS, i.e. RA+SQ), within a radius and over a hexagonal grid (RH, i.e. RA+HE) or by DBSCAN (DB) method (see Appendix for details of each of such methods). For the first 5 methods, fixed geometries are used to constrain the clustering areas, whereas for DB method, the shape and the size of clustering areas can vary with the distribution of the data. We use a partitioning approach in which each IDP is assigned to only one cluster and cannot be shared by more clusters as in "hierarchical clustering" method (e.g., Amorese et al., 2015).

Step B ("Occurrence" in Fig. 5): each cluster of IDPs collects intensities. The minimum number of IDPs to calculate a MDP intensity in an area/cluster could be taken in analogy with agencies that collect and provide "crowdsourced" intensities: 5, like HSIT (Tosi et al., 2015) and 3, like DYFI (Wald et al., 2011). Areas/clusters with a number of IDPs lower than the threshold are not evaluated and the MDPs are not assessed. After several tests with different thresholds, we decided to use 3 IDPs (as done by DYFI).

Step C ("MDPs" in Fig. 5): on the IDPs in each area/cluster, we apply various statistical estimators of central tendency to derive both location (geographical coordinates) and the final MDP intensity of each cluster. We use the average (mnsa), the median (mdna) and the trimmed mean with four different intervals of the distribution of the sorted intensity values: 10%-90% (mn10), 15%-85% (mn15), 20%-80% (mn20), 25%-75% (mn25). Note that trimmed means are computed only if the tails of the distributions have at least one IDP, otherwise the simple average is used. The use of central tendency estimators reduces the effects intensity outliers because these are averaged with other IDPs in the clustering area and do not influence the final MDP intensity assessment too much. The approach followed is more conservative compared to HSIT and DYFI by preserving the intensities assessed by citizens as much as possible.

MDPs available are therefore the results of the combination of grouping and central tendency methods using both raw (R) and corrected (C) intensities. Consequently, even the computed MDPs are hereinafter and analogously indicated as raw or corrected.

To calculate MDP, we used a minimum threshold of 3 IDPs, deleting geographical outliers and using the intensity in the range 3-11 degrees (3-10 for corrected intensities). Hence, the initial number of 51359 earthquakes in the EMSC dataset reduces to 22761 (Table S1 of the supplementary material). The selected earthquakes whose instrumental epicentre is located inland are about 2/3 of the total, covering a wide range of magnitudes. It is important to note that a threshold of 5 IDPs would immediately eliminate further 4291 earthquakes and that only 2603 earthquakes have more than 100 IDPs while 20159 earthquakes have IDPs ranging between 3 and 100 (panel B of Table S1).

c) From MDPs to macroseismic parameters

The BOXER code (Gasperini et al., 1999, 2010) calculates macroseismic parameters such as epicentre, magnitude and their uncertainties using available MDPs. Among the different computation methods available in the code, we use only the n. 0 and n. 1, hereinafter indicated as BOXER-0 and BOXER-1 (or Bx0 and Bx1), respectively. Method 0 computes the epicentre as the barycentre of the sites with most severe effects. Method 1 computes the centre of the entire intensity distribution by a minimisation of squared residuals of an attenuation function (Pasolini et al., 2008). BOXER-0 can locate even the earthquakes with only one MDP whereas BOXER-1 needs more than one MDP (we set a minimum of 5 MDPs) with the obvious consequence of reducing the total number of events for which macroseismic parameters can computed. However, the latter method allows in most favourable cases to assessing the epicentre also for earthquakes located offshore or in uninhabited areas. Macroseismic magnitude can also be estimated by different methods, depending on the number and the distribution of MDPs. The classical method described in Gasperini et al. (1999) uses both the epicentral intensity I0 and the average distances R_I of various classes of intensities I. However, as the I0 computed by questionnaire data is usually unreliable, we modified the original algorithm to only use the R_I. In any case, at least 4 MDPs are required (two intensity classes with two MDPs each) to compute a magnitude. The alternative methods described by Gasperini et al. (2010), based on a linear relation between I0 and M cannot be used in the present work for the poor reliability of I0, as well the new method described in Gasperini et al. (2010) because it was found to systematically underestimate the magnitudes.

Results and discussion

In Table 1 we show the distribution of the number of earthquakes as a function of the number of MDPs using both raw and corrected intensities. For ~7600 of the 22761 initial earthquakes, we do not even have a single MDP. Consequently, the earthquakes with at least one MDP for which we can provide the location are ~15000 (Table 1). Hence, only 2/3 of the events can be compared with instrumental locations to quantify the ability of BOXER code to provide reliable macroseismic parameters.

For each earthquake of the dataset with at least 3 IDPs (22761 earthquakes), we combined 11 different grouping methods and 6 different central tendency estimators to assess MDPs (Appendix and Table A1). Moreover, macroseismic locations and magnitudes are computed by 2 methods (BOXER-0 and BOXER-1). Hence, in total, we have 132 different alternative combinations of methods to test. The minimum threshold of 3 IDPs for locate an earthquake is a minimum but not sufficient condition because it is necessary that they all belong to the same clustering area to have a single MDP.

The comparison between macroseismic epicentres and magnitudes with instrumental data provides an estimate of the reliability of the computed parameters for different combinations both using raw and corrected intensities. Instrumental locations and magnitudes of each earthquake are taken from the EMSC webservice. In particular magnitudes are homogenized to Mw by using empirical formulas at the global scale of Lolli et al. (2014). For each earthquake it is possible to evaluate the combinations of methods which separately minimize the distance between macroseismic and instrumental epicentre and the difference between the macroseismic and instrumental magnitude, but they are generally different for different earthquakes.

However, we can establish a ranking of combinations by counting the number of earthquakes for which each combination best reproduces the instrumental parameters. To objectively compile such ranking, we consider datasets of earthquakes for which both the epicentre and the magnitude can be computed using all combinations. Such datasets include 1144 and 1082 earthquakes for raw and corrected intensities, respectively.

For each earthquake, we assign a score 3 to the combination of methods having, separately, the minimum epicentral distance and the minimum absolute magnitude difference, a score of 1 to all combinations with distances and differences within 5% of the minimum ones and no greater than 1 km and 0.2 m.u. and a score of 0 for all the other cases. We used such nonparametric approach (instead of, for example, the total root mean square error) because we are unsure that macroseismic locations and magnitudes are normally distributed, even considering the possible presence of intensity outliers in some earthquakes.

Such scores are reported in Table 2 for raw intensities (and in Table S2 of the supplementary material for corrected intensities). Neither for localization distances nor for differences in magnitude, there is a combination which clearly overperforms all the other ones and which we can choose as the "preferred" one to use prospectively.

The best performing combinations are different for epicentral location and magnitude and for raw and corrected intensities. For epicentral location from raw intensities (Tables 2 and 3), the first 43 combinations in the ranking use BOXER method 1 and the first 5 the grouping method DB2. For corrected intensities (Tables S2 and S3 in the supplementary material) the first 27 use BOXER method 1 and 4 of the first 5 use the grouping method DB2.

For magnitude estimation, the results are less coherent. Using the raw intensities (Tables 2 and 4), in the highest rankings we have an alternation of both BOXER methods and different grouping methods with a certain prevalence of BOXER-1 and grouping methods RA and DB. Using corrected intensities (Tables S2 and S4 in the supplemental material), the first 11 combinations use BOXER-1 while the preferred grouping methods varies from DB to RA and RH.

The better agreement of BOXER-1 with respect to BOXER-0, concerning the distance from the instrumental epicentre and the good performance of the grouping method DB2 can be immediately evidenced by plotting (Fig. 6) the values of Tables 2 and S2 for the raw and corrected intensities, respectively: the greater the distance of each combination from the centre of each Radar plot the higher the score obtained by the combination. We also observe a prevalence of the median as central tendency estimator that minimise the difference with the instrumental data for various grouping and BOXER combinations. About the difference in magnitude, the values are similar to each other, showing the lowest values for the DB2 grouping method but there is not a clear prevalence of one BOXER or central tendency estimator method with respect to the others.

In general, not all earthquakes can be located and sized by the best performing combination, hence, to determining the parameters for as many earthquakes as possible, even combinations other than the "top" ranking one must be used. To verify which combinations are mostly useful, we compute epicentres and magnitudes in our complete datasets of 22761 earthquakes, using the combinations with higher ranking that are able, separately, to compute such parameters.

In the bottom sections of Tables 3 and 4 for raw intensities, we report the numbers of earthquakes located (15103) and sized (5703) by each combination according to such procedure. Note that the total number of located earthquakes is about ~2/3 of the 22761 earthquakes, while magnitudes can only be estimated for ~1/5 of the earthquakes. This because, for the location, one MDP is sufficient, while for the magnitude, at least 4 MDPs are needed. Hence, it is not possible to locate 7658 and to size 17058 earthquakes. The results for corrected intensities are shown in Tables S3 and S4 of the supplementary material, with similar values for earthquakes located (15100) and sized (5625), and not-located (7661) and not-sized (17136).

For raw intensities (Table 2), combinations using BOXER-1 and BOXER-0 can locate ~1/3 and ~2/3 of the 15103 earthquakes, respectively. In detail, 3321 (22%) earthquakes can be located using the "top" scoring combination (DB2-20% trimmed average-BOXER-1), other 1816 (12%) earthquakes can be located by different combinations using BOXER-1. Overall, BOXER-1 locates at best 5137 earthquakes i.e. all the events that have number of MDPs ≥5. BOXER-0 locates 9966 earthquakes, 6409 (42.4%) of which by the combination "3500RH3-average", 1864 (12.3%), by the "2000RA-median" 1633 (10.8%) by the "DB2- 20% trimmed average". The latter three combinations correspond to the 44th, 66th, 69th positions in the ranking, respectively (Table 3). Overall, 17 combinations are used to locating 15103 earthquakes.

The situation is similar for corrected intensities (Table S2 of the supplementary material) where 3319 (22%) earthquakes can be located by the same top scoring combination for raw intensities, 1756 (11.6%) by other combinations using BOXER-1, 10025 (66.4%) by combinations using BOXER-0. Overall, BOXER-1 locates 5075 earthquakes of 5134 earthquakes with the number of MDPs ≥5 (Table 1). Also, for corrected intensities, 17 combinations locate 15100 earthquakes. Excluding the top scoring combination (DB2-20% trimmed average-BOXER-1), median and average are generally used for locating earthquakes (Table 3), in agreement with the highest-ranking values in Table 2 and Figure 6.

For raw intensities (Table 4), combinations using BOXER-1 and BOXER-0 assign the magnitude at best to 3767 and 1936 events, respectively (i.e. ~2/3 and ~1/3 of the total of 5703 earthquakes). This preference for BOXER-1 is even more pronounced with corrected intensities (Table S4 of supplementary material) with 4959 (88%) of the total of 5625 events, whereas combinations with BOXER-0 assess the magnitude at best for only 666 events (12%). Using raw intensities (Table 4), 3060 (53.7%) magnitudes can be determined by the top scoring combination (DB2-10% trimmed mean-BOXER-1), other 707 (12.4%) by combinations using BOXER-1, 1936 (34.4%) by combinations using BOXER-0. In all 71 combinations are used to compute the 5703 magnitudes. Using corrected intensities, 2984 (53%) magnitudes can be determined by the top scoring combination (DB2-mean-BOXER-1), other 1965 (35.1%) by combinations using BOXER-1, 666 (11.8%) by combinations using BOXER-0. In all 83 combinations are used to compute 5625 magnitudes. Note that all the grouping, central tendency and BOXER methods are necessary to compute epicentres of magnitudes for all earthquakes.

From a first analysis of the correspondence between macroseismic and instrumental parameters in Fig. 7 , it is quite evident a geographical heterogeneity: a fairly good agreement is observed in Europe and North America and some greater discrepancy in other areas of the World. For this reason, we will analyse the results not only at a global scale but also for the 5 macro-areas indicated in Fig. 7: Europe (EU), Asia and Oceania (AO), North America (US), South America (SA), Africa (AF). It is obvious to relate the agreement and disagreement between macroseismic and instrumental parameters with the number of IDPs available in the different areas. In fact, the larger number of IDPs in the EU and US with greater density and continuity (Fig. 1) corresponds to a higher average number of MDPs in the same areas for each analysed earthquake (Table 5 for raw and corrected intensities). Such larger number of MDPs per earthquake therefore manages to better constrain location and macroseismic magnitude, improving the agreement with the instrumental data at the global scale (see Fig. S3 and cTable S5 of the supplementary material).

Both at the global scale and for different macro-areas, we calculated the frequency histograms in various ranges of distances and magnitude differences (Fig. 8 with numerical values in Tables S6 and S8 of the supplementary material for raw intensities and Fig. S4, Tables S8 and S9 of the supplementary material for corrected ones). The lower the values, the better the fit of macroseismic to instrumental values. All earthquakes (a) have also been divided into categories or subsets, depending on whether they are located inland (L) or offshore (S), have the maximum gap between available MDPs and epicentre less than 180 degrees (g), and, for epicentral distance only, have at least 3 MDP(n). We do not consider the latter subdivision for magnitudes because the minimum number of MDPs for computing them is 4. As well, for a gap <180 degrees, 3 MDPs are required at least. This comparison between macroseismic and instrumental parameters is displayed in Fig. 8 both in terms of number of events and of percentage of the total number.

Both for the distance and for the difference in magnitude, at the global scale, the earthquakes located on land (L) are ~2/3 of the total (a) while ~1/3 are located offshore (S). The agreement is generally better for the former ones than for the latter ones and improves by a few percentage points by only considering earthquakes with at least 3 MDPs (n). The agreement further improves for earthquakes with maximum azimuthal gap lower than 180 degrees, which number, however, is about 1/4 of the total for the location and to about one half for the magnitude. For about 30% of earthquakes, the distance exceeds 50 km and for about 15% of them it exceeds 100 km. Only 40% of the earthquakes have magnitude differences less than 0.6 m.u. This indicates a certain difficulty of the macroseismic magnitudes in reproducing the instrumental ones.

Analysing the results by macro-areas, the correspondence between macroseismic and instrumental data shows significant variations: ~2/3 of the earthquakes are concentrated in Europe, while the other macro-areas have about 1200-2200 earthquakes with location and 400-700 with magnitude except for the African area having about 200 events (Fig. 8). Compared to the data at a global scale, a clearly better agreement between macroseismic and instrumental parameters is observed for the EU and the US areas and a worse agreement for AO, SA and AF (Fig. 8).

It is also clear that events in the sea (S) have worse agreement with the instrumental data than all the other datasets (a, L, n, g). Compared to the whole dataset (a), the trend of improvement of the agreement is evident for the subsets L, n and g. It follows that the number of MDPs (n), possibly well distributed around the epicentre (g), are factors that improve the quality of the final macroseismic data, making the calculated parameters more reliable. Increasing more and more the number of IDPs and then of MDPs is a goal and a mean to obtain realistic estimates of macroseismic parameters. The use of corrected intensities leads to results substantially similar to those calculated with raw intensities, with some slight improvements at the shortest distances and smaller magnitude difference (Fig. S5 of the supplementary material).

To show the evolution over time of the agreement between macroseismic and instrumental data, we subdivided the results by year, from 2012 to 2022 (excluding 2023 which has only two months of data). Fig. 9 shows the overall results of the distance and magnitude difference at a global scale, both in terms of number of earthquakes and of percentage. For each year, the earthquakes are divided into subsets (SaLng for distance and SaLg for magnitude difference) analogously to Fig. 8. It is possible to observe how the number of events whose macroseismic parameters are estimated increases over time, except for year 2022 in which it decreases. The agreement between macroseismic and instrumental parameters remains similar to each other even within the subsets of earthquakes.

Over time, the distances and the differences in magnitude decrease: in 2020-2022 for the subsets "L", "n" and "g", the percentage of earthquakes located within 10 km from the instrumental epicentre is about 20-30%, about 30-50% within 20 km, about 50-70% within 30 km and about 80% within 50 km. For the differences in magnitude, a slight percentage improvement over time is observed with about 25-40% of the earthquakes of the subsets "L" and "g" within about 0.3 m.u. and about 65% of events within 0.6 m.u. The trend of improvement over time is even more visible considering events beyond certain values (e.g. 100 km away and 1 degree of magnitude) which halves their percentages compared to the first few years. It should also be noted that some years like 2016 and 2017 have percentages in line or even better in terms of agreement than most recent years.

The temporal behaviour in the different macro-areas compared to the global scale (Fig. 9) shows different results both in terms of percentage and of the number of earthquakes. For Europe (Fig. 10), it can be observed that the number of earthquakes slightly decreases in 2018 and in 2022, but increases the percentage of earthquakes that have relatively shorter distances and smaller magnitude differences. Furthermore, over the years we can note a marked decrease in the percentages of earthquakes with distances longer than 100 km and magnitude differences greater than 1 degree (Fig. 10). In the other macro-areas (AO, US, SA, AF), we have about 1/3 of the total number of earthquakes analysed. For certain years and/or certain subsets of earthquakes, the small number of events available makes the statistics scarcely significant. The North America (US) area has similar or even slightly better agreement than that of Europe with the exception of years 2012 and 2013 when the statistics are insignificant due to the low number of events (Figs. S6 and S7 of the supplementary material for raw and corrected intensities, respectively). In the other macro-areas (Figs. S8 and S9 of the supplementary material) the percentage of well localized events and well assigned magnitudes also drops significantly due to the reduced number of MDPs per earthquake (Table 5). Using the corrected intensity gives similar results (Figs. S10-S15 of Supplementary Material). For the sake of clarity, we also provide in Appendix B an example of the entire procedure from IDPs to macroseismic parameters for the 2020/09/19 California earthquake (06:38 UTC, M=4.5).

Conclusions

We analysed the database of individual intensities provided by citizens (1874376 IDPs) collected and made available online by the EMSC for 51359 earthquakes. The database provides two intensity values: raw and corrected (i.e. eliminating intensities >10 and applying an empirical formula to the raw data, according to Bossu et al., 2017). On both the raw and corrected datasets we applied various methods for grouping the IDPs. We tested the combinations of 11 clustering methods and 6 central tendency estimators (mean, median, trimmed means with various trimming intervals) to derive a MDP intensity for each cluster with at least 3 IDPs. The MDPs thus available were processed with methods 0 and 1 of the BOXER code (Gasperini et al., 2010). Therefore, for each event there are 132 possible combinations of methods, for each type of intensity, which allow to compute epicentre and macroseismic magnitude. The threshold of at least 3 IDPs for deriving an MDP, significantly lowers the number of earthquakes for which macroseismic parameters can actually be calculated. Furthermore, at least 4 MDPs are required for the calculation of magnitude. Therefore, it is possible to compute an epicentre and a magnitude for ~15000 and ~5700 earthquakes, respectively.

The calculated macroseismic parameters can be compared with the instrumental ones to evaluate the reliability of the entire methodology. To identify the combination that minimizes the difference with the instrumental data, separately for distance and magnitude, we selected about a thousand earthquakes for which the parameters could be calculated for all the possible combinations. A score was assigned to each combination based on its ability to well reproduce the instrumental parameters of each earthquake. This systematic approach shows similar score values for several combinations of methods, especially concerning the difference in magnitude. Considering the distance alone, however, the better overall results are obtained by BOXER-1 compared to BOXER-0. The score assigned to the different combinations for all earthquakes defines a ranking that can be used to select the most preferable ones in a prospective view.

Since not all earthquakes can be located and sized by the best performing combination, other combinations must also be used to determine the parameters for as many earthquakes as possible. In particular most earthquakes can only be located using BOXER-0 because it requires less MDPs than BOXER-1 to be applied.

In addition to the complete dataset of available earthquakes (a), we also considered subsets of events with epicentre located on land (L), offshore (S), with number of MDPs≥3 (n) and with azimuthal gap between MDPs and instrumental epicentre < 180 degrees (g).

The analyses we brought, not only at a global scale but also for 5 macro-areas (Europe, Asia and Oceania, North America, South America, Africa), show substantially similar results between raw and corrected intensities. The distribution of available earthquakes shows a clear concentration in Europe with ~2/3 of the total data. In general, the fit between macroseismic and instrumental parameters shows an increasing trend from the dataset of earthquakes located offshore (S) up to the dataset of earthquakes with a gap (g) of less than 180 degrees, with intermediate results for other datasets (a, L, n). Moreover, compared to the global scale, some macro-areas (Europe, North America) have a better fit than others (Asia and Oceania, South America, Africa). We can argue that the larger numbers of MDPs per earthquake that we have in Europe and North America, has a role in improving the agreement with instrumental parameters. In the practice, future near real-time analyses will take advantage of knowing the macro-area where an event occurs to give a preliminary assessment of the likely reliability of calculated parameters.

Analysing the results as a function of time and macro-areas, we can observe increasing trends for subsets as well as for the complete dataset, except for certain areas or certain years for which the low number of events makes the statistics poorly significant. With the increase over time of the number of MDP available per earthquakes, an improvement of the fit between macroseismic and instrumental parameters is generally observed. In certain areas such as Europe and North America, 60-70% of the events are localized within about 30 km from the instrumental epicentre with a magnitude difference <0.6 m.u. and, above all, there is a strong reduction over time of extreme differences (more than 100 km of distance or >1 of magnitude). In other areas however the agreement is still not so good probably due to the still low number of MDPs. It is therefore desirable to continue to increase the number of IDPs and to overcome the economic and political barriers which today exclude large areas of the Earth from the possibility of providing such information.

Finally, the reporting of IDPs could also be influenced by the thumbnails representing the different scenarios associated with various degrees used in the LastQuake system. In particular, the types of houses and furniture depicted in them are more similar to European and North American environments than to those of other macro-areas and this makes it more difficult to apply the EMS98 scale to the damage scenario.

The processing performed by applying the BOXER code to the IDPs data in an original way is essential and preparatory for future applications in near real-time. When, for an event, EMSC starts collecting IDPs from citizens, an automatic procedure can be run. If the number of IDPs is enough to allow their grouping into MDPs it will be possible to assess location and magnitude with BOXER following a preferential ranking order. The greater the number of MDPs, the greater the reliability of the result. In particular, we believe that the threshold of 5 MDPs allowing the application of the BOXER-1 method is a discriminating element in order to give greater reliability to the results. Obviously, further comparative tests of the results at time intervals will have to be conducted, exploiting the delay time information with respect to the time T0 origin of the event, but all this will be the subject of further specific work and is beyond the scope of the present purposes.

Data and resources

Boxer code: freely available at: https://emidius.mi.ingv.it/boxer/

Cities500.txt database, available at https://www.geonames.org

DYFI, “Did you feel it?”, available at http://earthquake.usgs.gov/earthquakes/dyfi/

IDPs, individual intensities data points are downloaded by EMSC via webservices, (www.seismicportal.eu/testimonies-ws/). e.g.: http://www.seismicportal.eu/testimonies-ws/api/search?unids=20210629_0000012&includeTestimonies=true (last accessed January 2023).

EMSC – European Seismological Centre, https://emsc.csem.org/

GHSL database, available at https://ghsl.jrc.ec.europa.eu/index.php (last accessed January 2023)

GeoNames database, available at https://www.geonames.org (last accessed February 2022)

GeoNet New Zealand questionnaires, available at https://www.geonet.org.nz

HSIT, “Hai sentito il terremoto?”, available at http://www.haisentitoilterremoto.it/

SHAKEMAP -A Tool for Earthquake Response, available at https://earthquake.usgs.gov/data/shakemap/

Supplementary material for this article includes figures and tables that provide further information and details of the main text. Moreover, similar elaborations, plots and figures are given for the "corrected" intensities in as for the "raw" intensities in the main text.

Acknowledgements

This paper benefitted from funding provided by the H2020 EU project RISE n. 821115 and Istituto Nazionale di Geofisica e Vulcanologia-Sezione di Bologna. We thank Rémy Bossu and Matthieu Landès for data, discussion and suggestions in the preparation and drafting of the article.

References

Amorese, D., R. Bossu, and G. Mazet-Roux (2015). Automatic clustering of macroseismic intensity data points from internet questionnaires: Efficiency of the partitioning around medoids (PAM), Seismol. Res. Lett. 86, no. 4, 1171–1177.

Bakun, W. H., and C. M., Wentworth (1997). Estimating earthquake location and magnitude from seismic intensity data, Bull. Seismol. Soc. Am., 87/6, 1502-1521.

Bossu, R., M. Laurin, G. Mazet-Roux, F. Roussel, and R. Steed (2015). The importance of smartphones as public earthquake-information tools and tools for the rapid engagement with eyewitnesses: A case study of the 2015 Nepal earthquake sequence, Seismol. Res. Lett. 86, 6, 1587–1592

Bossu, R., F. Roussel, L. Fallou, M. Landès, R. Steed, G. Mazet-Roux, A. Dupont, L. Frobert, and L. Petersen (2018). LastQuake: From rapid information to global seismic risk reduction, Int. J. Disast. Risk Reduc. 28, 32-42.

Bossu, R., M. Landès, F. Roussel, R. Steed, G. Mazet-Roux, S. S. Martin, and S. Hough (2017). Thumbnail-based questionnaires for the rapid and efficient collection of macroseismic data from global earthquakes, Seismol. Res. Lett. 88, 1, 72-81.

Dewey, J., D. Wald, and L. Dengler (2000). Relating conventional USGS modified Mercalli intensities to intensities assigned with data collected via the Internet, Seismol. Res. Lett. 71, 264.

Ester M., H.P. Kriegel, J. Sander, and X. Xiaowei (1996). A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise. KDD'96: Proceedings of the Second International Conference on Knowledge Discovery and Data Mining, 226–231

Florczyk, A., C. Corbane, M. Schiavina, M. Pesaresi, L. Maffenini, M. Melchiorri, P. Politis, F. Sabo, S. Freire, D. Ehrlich, T. Kemper, P. Tommasi, D. Airaghi, and L. Zanchetta (2019). GHS Urban Centre Database 2015, multitemporal and multidimensional attributes, R2019A. European Commission, Joint Research Centre (JRC)PID: https://data.jrc.ec.europa.eu/dataset/53473144-b88c-44bc-b4a3-4583ed1f547e, https://ghsl.jrc.ec.europa.eu/ghs_stat_ucdb2015mt_r2019a.php

Gasperini, P., F. Bernardini, G. Valensise, and E. Boschi (1999). Defining seismogenic sources from historical earthquake felt reports. Bull. Seismol. Soc. Am. 89, 94-110.

Gasperini, P., G. Vannucci, D. Tripone, and E. Boschi (2010). The location and sizing of historical earthquakes using the attenuation of macroseismic intensity with distance. Bull. Seismol. Soc. Am. 100, 2035-2066, doi: /10.1785/0120090330.

Goded T., N. Horspool, S. Canessa, A. Lewis, K. Geraghty, A. Jeffrey, and M. Gerstenberger (2018), New macroseismic intensity assessment method for New Zealand web questionnaires, Seismol. Res. Lett., 89(2A), 640-652, doi.org/10.1785/0220170163.

Grünthal, G. (ed.) (1998). European macroseismic scale 1998, Conseil de l'Europe. Cahiers du Centre Européen de Géodynamique et de Séismologie,13, Luxembourg, 99 pp.

Hough, S. E., and S. S. Martin (2021), Which earthquake accounts matter?, Seismol. Res. Lett., 92 (2A), 1069-1084, doi.org/10.1785/0220200366.

International Seismological Centre (2022), On-line Bulletin,https://doi.org/10.31905/D808B830

Lolli B., P. Gasperini, and G. Vannucci (2014). Empirical conversion between teleseismic magnitudes (mb and Ms) and moment magnitude (Mw) at the Global, Euro-Mediterranean and Italian scale, Geophys. J. Int., 199, 805-828, doi: 10.1093/gji/ggu264

Musson, R.M.W., and M.J Jiménez (2008). Macroseismic Estimation of Earthquake Parameters. NA4 Deliverable D3, NERIES Project.

Pasolini, C., D. Albarello, P. Gasperini, V. D'Amico, and B. Lolli (2008). The attenuation of seismic intensity in Italy part II: modeling and validation. Bull. Seismol. Soc. Am. 98, 692-708. Doi.org/10.1785/0120070021.

Pettenati, F., and L. Sirovich (2003). Tests of source-parameter inversion of the U.S. Geological Survey intensities of the Whittier Narrows, 1987 earthquake. Bull. Seismol. Soc. Am., 93, 47-60.

Rovida, A., M. Locati, R. Camassi, B. Lolli, and P. Gasperini (2020). The Italian earthquake catalogue CPTI15, Bull. Earthq. Eng., 18, 2953-2984, doi: 10.1007/s10518-020-00818-y.

Sieberg, A. (1912): Über die makroseismische Bestimmung der Erdbebenstärke. Ein Beitrag zur seismologische Praxis, G.Gerlands Beiträge zur Geophysik, 11 (2-4), 227-239 (in German).

Sieberg, A. (1932): Erdebeben, in Handbuch der Geophysik, Vol. 4, (B. Gutenberg Ed.), 552-554 (in German).

Tosi, P., P. Sbarra, V. De Rubeis, and C. Ferrari (2015), Macroseismic intensity assessment method for web-questionnaires. Seism. Res. Lett, 86, 985-990, doi: 10.1785/022014022.

Vannucci G., D. Tripone, P. Gasperini, G. Ferrari, and B. Lolli (2015). Automated assessment of macroseismic intensity from written sources using the Fuzzy sets. Bulletin of Earthquake Engineering, 13, 2769-2803, doi:10.1007/s10518-015-9759-5

Vannucci G., P. Gasperini, B. Lolli, and L. Gulia (2019). Fast characterization of sources of recent Italian earthquakes from macroseismic intensities. Tectonophysics 750, 70-92, doi: 10.1016/j.tecto.2018.11.002

Wald, D. J., V. Quitoriano, L. Dengler, and J. W. Dewey (1999). Utilization of the Internet for rapid community intensity maps, Seismol. Res. Lett. 70, 87–102.

Wald, D.J., V. Quitoriano, B. Worden, M. Hopper, and J.W. Dewey (2011). USGS "Did You Feel It?" Internet-based macroseismic intensity maps. Annals Geophys., 54, 6; doi: 10.4401/ag-5354

Authors' addresses

Gianfranco Vannucci
Istituto Nazionale di Geofisica e Vulcanologia, Sezione di Bologna, Viale Berti Pichat 6/2, 40127 Bologna, Italy , email: gianfranco.vannucci@ingv.it

Paolo Gasperini
Dipartimento di Fisica e Astronomia, Università di Bologna, Viale Berti Pichat 6/2, 40127 Bologna, Italy, email: paolo.gasperini@unibo.it

Laura Gulia
Dipartimento di Fisica e Astronomia, Università di Bologna, Viale Berti Pichat 6/2, 40127 Bologna, Italy, email: laura.gulia@unibo.it

Barbara Lolli
Istituto Nazionale di Geofisica e Vulcanologia, Sezione di Bologna, Viale Berti Pichat 6/2, 40127 Bologna, Italy, email: barbara.lolli@ingv.it

Tables

Table 1 – number of earthquakes (n. Eqks) as a function of the number of MDPs. The last rows (in bold) show the cumulative number of earthquakes with MDPs numbers ≥1, ≥3, ≥5 .

n. MDPs	n. Eqks	
	Raw Int.	Corrected Int.
0	7658	7661
1	4940	4940
2	2550	2551
3	1473	1477
4	1003	998
5	742	742
6	516	515
7	424	426
8	369	367
9	279	280
10	256	255
11-15	763	762
16-20	430	430
21-30	439	440
31-50	380	378
51-75	207	207
76-100	104	104
101-150	101	101
151-200	46	47
201-500	66	65
501-1000	12	12
1001-2000	3	3
all	22758	22758
>=1	**15103**	**15100**
>=3	**7613**	**7609**
>=5	**5137**	**5134**

Table 2 – Scores (see text) obtained, using raw intensities, by the 132 combinations for both the distances between macroseismic and instrumental locations (Di, upper part of the table) and the differences between macroseismic and instrumental magnitudes (dM, lower part of the table). The comparison refers to the dataset of common earthquakes (n.eqks) for which the parameters can be calculated by all the combinations of methods. Grouping methods are indicated (see Appendix) as a function of population density (den), clustering method (MG) and grid/radius of the area (size), while central tendency estimators are indicated by acronyms: average (mean), median (mdna) and trimmed averages with 10%, 15%, 20% and 25% of tail trimming: (mn10, mn15, mn20, mn25, respectively). Bx0 and Bx1 indicate BOXER methods 0 and 1 respectively. Results for corrected intensities are reported in Table S2 of the supplementary material.

Dset: A n.eqks: 1144			**Raw intensities**										
		den	*2000*	*3500*	*5500*				*3500*	*3500*			
		MG	*RA*	*RA*	*RA*	*SQ*	*SQ*	*HE*	*RS*	*RH*	*DB*	*DB*	*DB*
		size				*1*	*2*	*2*	*3*	*3*	*0.5*	*1*	*2*
Di	**Bx0**	mean	56	61	54	45	45	43	51	86	36	46	51
		mdna	70	55	52	64	56	42	62	54	51	49	60
		mn10	42	46	45	36	36	38	50	52	35	43	40
		mn15	42	42	47	30	37	47	52	48	36	53	41
		mn20	54	62	43	59	34	52	53	59	39	40	67
		mn25	53	63	48	58	38	45	54	67	44	48	54
	Bx1	mean	103	88	126	98	87	100	93	127	92	96	129
		mdna	106	91	119	112	93	97	100	109	79	78	144
		mn10	97	92	114	78	99	91	75	121	87	85	116
		mn15	92	97	90	87	81	75	72	90	74	85	145
		mn20	84	94	74	78	67	86	68	107	78	81	146
		mn25	82	101	92	79	75	88	90	101	86	77	130
dM	**Bx0**	mean	324	318	324	318	273	291	285	231	312	267	246
		mdna	330	342	309	300	258	258	258	234	300	252	279
		mn10	297	282	315	288	282	306	243	279	285	297	303
		mn15	285	297	309	303	294	267	246	252	330	237	315
		mn20	303	342	318	300	297	225	240	243	300	282	240
		mn25	288	258	303	285	309	252	246	261	291	270	258
	Bx1	mean	324	285	318	273	240	249	294	264	255	246	315
		mdna	294	258	282	285	249	222	258	324	237	285	324
		mn10	306	318	366	255	273	249	264	291	222	279	384
		mn15	273	315	288	246	276	276	270	297	243	276	357
		mn20	315	312	249	270	285	246	258	330	270	285	279
		mn25	261	297	285	255	255	234	258	321	264	252	291

Table 3 - Upper part: ranking order of the 132 combinations of methods based on distance scores in Table 2 (upper part) for raw intensities. Lower part: numbers of events for which macroseismic parameters can be computed by each combination, following the order of the ranking. “nd” indicates the number of earthquakes which cannot be located or sized by any combination. Acronyms as in Table 2. Results for corrected intensities in Table S3 of the supplementary material.

Dset: A			**Raw intensities**										
		den	*2000*	*3500*	*5500*				*3500*	*3500*			
		MG	*RA*	*RA*	*RA*	*SQ*	*SQ*	*HE*	*RS*	*RH*	*DB*	*DB*	*DB*
		size				*1*	*2*	*2*	*3*	*3*	*0.5*	*1*	*2*
Di n eqks: 1144	**Bx0**	mean	81	75	87	110	109	113	97	44	129	105	96
		mdna	66	82	94	71	80	115	74	85	95	99	76
		mn10	118	106	107	126	127	123	98	92	130	112	121
		mn15	116	117	103	132	125	104	93	100	128	89	119
		mn20	83	73	114	78	131	91	90	77	122	120	69
		mn25	88	72	102	79	124	108	84	70	111	101	86
	Bx1	mean	16	40	7	22	42	20	29	6	32	26	5
		mdna	15	34	9	12	28	23	19	13	54	57	3
		mn10	24	33	11	58	21	35	61	8	43	47	10
		mn15	30	25	37	41	52	60	65	38	64	48	2
		mn20	49	27	63	56	68	46	67	14	55	51	1
		mn25	50	18	31	53	62	39	36	17	45	59	4
Di n eqks: 15103 nd: 7658	**Bx0**	mean	-	-	-	-	-	-	-	6409	-	-	-
		mdna	1864	-	-	-	3	-	20	-	-	-	-
		mn10	-	-	-	-	-	-	-	-	-	-	-
		mn15	-	-	-	-	-	-	-	-	-	-	-
		mn20	-	-	-	-	-	37	-	-	-	-	1633
		mn25	-	-	-	-	-	-	-	-	-	-	-
	Bx1	mean	-	-	49	-	-	127	-	1040	12	10	-
		mdna	116	-	-	232	-	-	148	-	-	-	-
		mn10	-	-	-	-	76	-	-	-	-	-	-
		mn15	-	-	-	-	-	-	-	-	-	-	-
		mn20	-	-	-	-	-	-	-	-	-	-	3321
		mn25	-	6	-	-	-	-	-	-	-	-	-

Table 4 - As in Table 3 for magnitude difference (dM) scores and raw intensities. Results for corrected intensities in Table S4 of the supplementary material.

Dset: A			Raw intensities										
		den	*2000*	*3500*	*5500*				*3500*	*3500*			
		MG	*RA*	*RA*	*RA*	*SQ*	*SQ*	*HE*	*RS*	*RH*	*DB*	*DB*	*DB*
		size				*1*	*2*	*2*	*3*	*3*	*0.5*	*1*	*2*
dM n eqks: 1144	**Bx0**	mean	10	15	11	17	77	49	56	129	25	86	116
		mdna	6	5	28	39	94	96	100	128	36	106	70
		mn10	41	66	23	54	69	31	120	72	57	40	33
		mn15	65	43	29	34	48	85	117	108	7	125	24
		mn20	32	4	18	37	42	130	123	121	38	68	122
		mn25	55	95	35	64	27	105	118	91	50	81	92
	Bx1	mean	12	60	16	80	124	109	47	87	101	114	21
		mdna	46	99	67	58	112	132	93	13	126	62	9
		mn10	30	19	2	102	78	110	89	51	131	71	1
		mn15	79	20	53	113	76	75	84	45	119	74	3
		mn20	22	26	111	83	59	115	98	8	82	61	73
		mn25	90	44	63	103	104	127	97	14	88	107	52
dM n eqks: 5703 nd: 17058	**Bx0**	mean	36	3	5	157	7	15	19	2	7	2	1
		mdna	180	88	1	8	7	27	13	14	13	1	28
		mn10	3	-	1	-	1	134	1	38	1	16	13
		mn15	2	1	2	19	29	11	9	3	504	-	79
		mn20	4	196	1	11	7	1	-	-	-	13	-
		mn25	-	-	-	-	158	2	2	39	-	-	1
	Bx1	mean	-	-	-	-	-	-	25	-	-	-	2
		mdna	-	-	-	-	-	-	5	58	-	1	29
		mn10	2	-	156	-	-	-	-	2	-	-	3060
		mn15	-	-	-	-	-	3	2	10	-	-	44
		mn20	3	-	-	-	-	-	-	359	-	2	-
		mn25	-	-	-	-	-	-	-	4	-	-	-

Table 5 - Average number of MDPs per earthquake (nMDPs/Eqk) at global scale and for macro-areas (as in Fig. 7) using raw and corrected intensities. Values for distance (Di) and difference of magnitude (dM) are shown.

n. MDPs/Eqk		**Global (W)**	**Europe (EU)**	**Asia, Oceania (AO)**	**North America (US)**	**South America (SA)**	**Africa (AF)**
Raw Int	Di	5.1	5.6	3.8	6.5	3.2	4.4
	dM	11.4	11.7	9.4	13.9	8.9	7.5
Corrected Int	Di	5.1	5.5	3.8	6.5	3.2	4.4
	dM	11.5	11.9	9.6	14.2	9.1	7.8

List of Figures Captions

Figure 1 - Top panel: In colours, numbers of IDPs on a regular grid with mesh of 1 degree both in latitude and longitude. In black, seismicity from the revised catalogue of International Seismological Centre (ISC, 2022), with M>3 in the time span 2013-2020. Bottom panels: frequency distribution over intensity bin of 0.5 degrees of IDPs of the EMSC database. Raw intensities and corrected ones (Bossu et al., 2017) in red and blue colours, respectively.

Figure 2 - procedure used from IDPs to assessment of macroseismic parameters.

Figure 3 - example of geographical (circled in red) and intensity outliers (with raw intensity=12, circled in blue), for the 2013/04/16 10:44 M=7.8 earthquake, number of IDPs: 408. The black dashed circle indicates the Maximum Distance Prediction Equation (MDPE) used to delete farthest IDPs and define the area (solid black line) of minimum and maximum latitude and longitude of selected IDPs. The black star indicates the instrumental epicentre.

Figure 4 - scheme of classification of the distribution of IDPs. The star indicates the instrumental epicentre, the circular dashed black line is a circle with MDPE radius, the rectangular black line delimits the area of location of usable IDPs, i.e. the minimum and maximum latitude and longitude of usable IDPs, without any geographic outliers (circled in red colour).

Figure 5 - Methods of clustering of IDPs into MDPs through three steps: column A: IDPs available are grouped (or not) following the various methods; column B: for each area of grouping the occurrence of a sufficient number of IDPs is assessed (numbers in green) or not (numbers in red); column C: IDPs are used to compute a combined intensity (MDPs), indicated with different colours

and symbols, for selected area/clusters (in white colours) and by using different central tendency estimators.

Figure 6 - Radar diagrams of the data represented in Tables 2 and S2 for Distance (Di, upper part) and difference of magnitude (dM, lower part) for BOXER-0 (Bx0) and BOXER-1 (Bx1). The light grey areas refer to BOXER-0 and the dark grey ones to BOXER-1. Coloured symbols (small circles) refer to central tendency estimators used to compute MDPs, plotted as a function of methods used to cluster raw IDPs (codified as in Table 2). The number of earthquakes (and then the agreement with instrumental data) increases from the centre of each circle outwards.

Figure 7 - Plot of distance (Di, lower panel, 15103 earthquakes) and magnitude difference of (dM, upper panel, 5703 earthquakes) between "preferred" macroseismic parameters and instrumental data for raw intensity. Five zones (EU=Europe, AO=Asia and Oceania, US=Nord America, SA= South America, AF=Africa) are shown.

Figure 8 - Statistical results of the comparison between macroseismic and instrumental parameters (represented in Fig. 7). Plots display numbers of event (N) and percentages (%) for magnitude differences (dM, upper panel) and distances (Di, lower panel). Columns refer to global scale (W) and different macro-areas (EU=Europe, AO=Asia and Oceania, US=Nord America, SA= South America, AF=Africa). The columns of each zone (see the legend in lowest left corner) indicate, from left to right, the earthquakes located offshore (S), all the earthquakes (a), earthquakes located inland (L), earthquakes with the number of MDPs ≥3 (n) and earthquakes with azimuthal gap < to 180 degrees (g). The area in grey highlights the macro-areas with respect to the global area (W). The scales of the numbers of earthquakes (N) are different for the global area (left) and the macro-areas (right).

Figure 9 - Same as in Figure 8, at the global scale and for different years.

Figure 10 - Same as in Figure 9 for Europe.

Figures

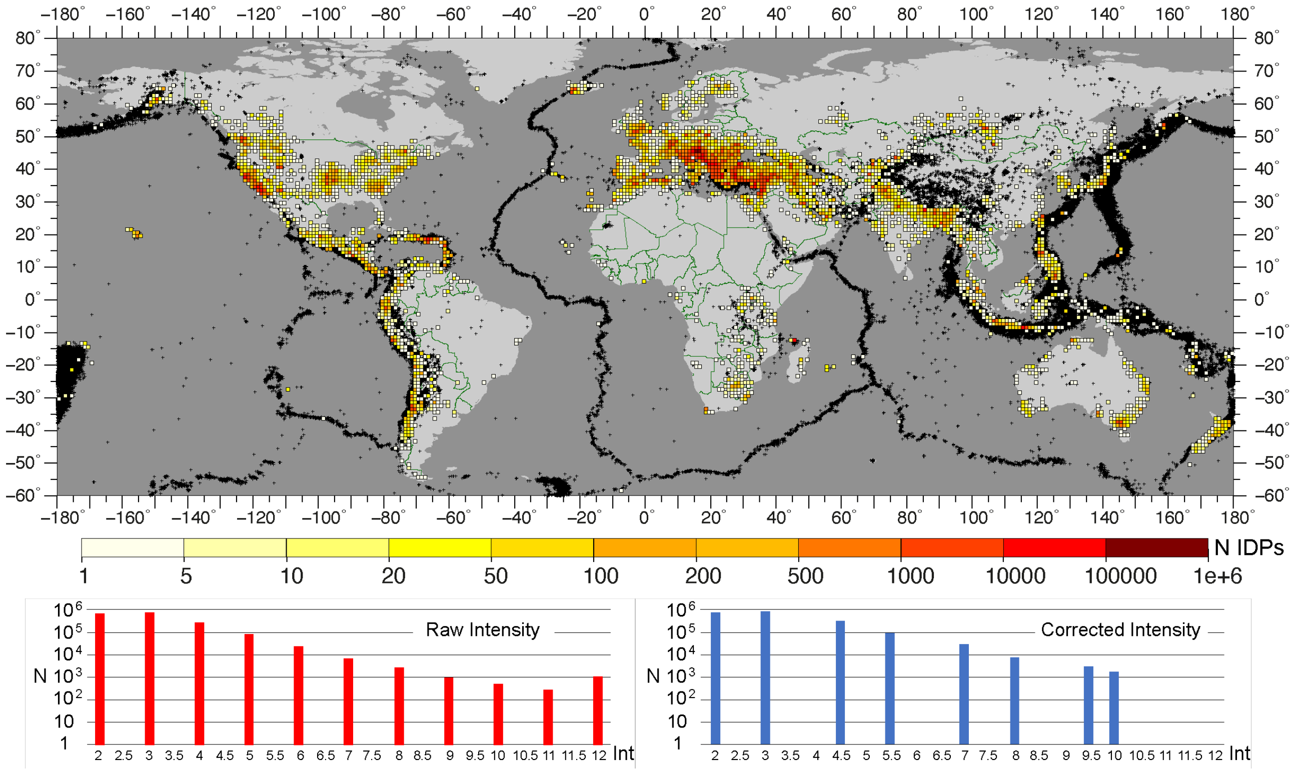


Figure 1 - Top panel: In colours, numbers of IDPs on a regular grid with mesh of 1 degree both in latitude and longitude. In black, seismicity from the revised catalogue of International Seismological Centre (ISC, 2022), with M>3 in the time span 2013-2020. Bottom panels: frequency distribution over intensity bin of 0.5 degrees of IDPs of the EMSC database. Raw intensities and corrected ones (Bossu et al., 2017) in red and blue colours, respectively.

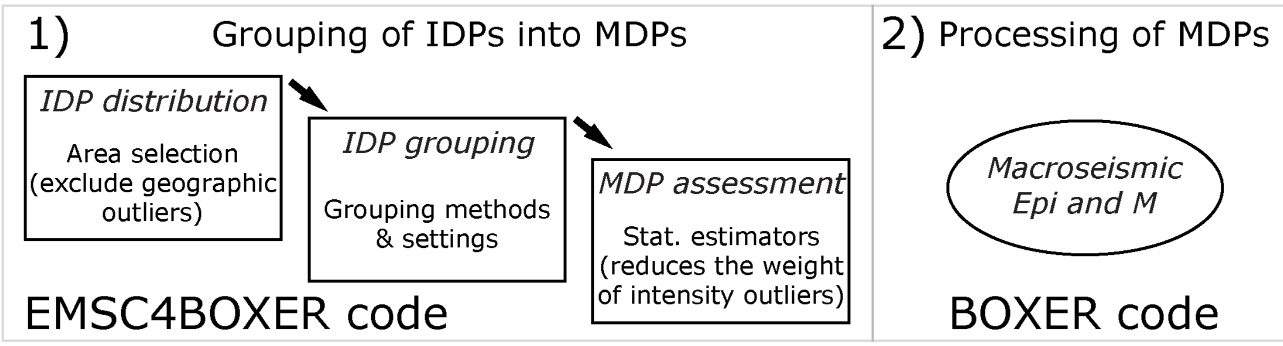


Figure 2 - procedure used from IDPs to assessment of macroseismic parameters.

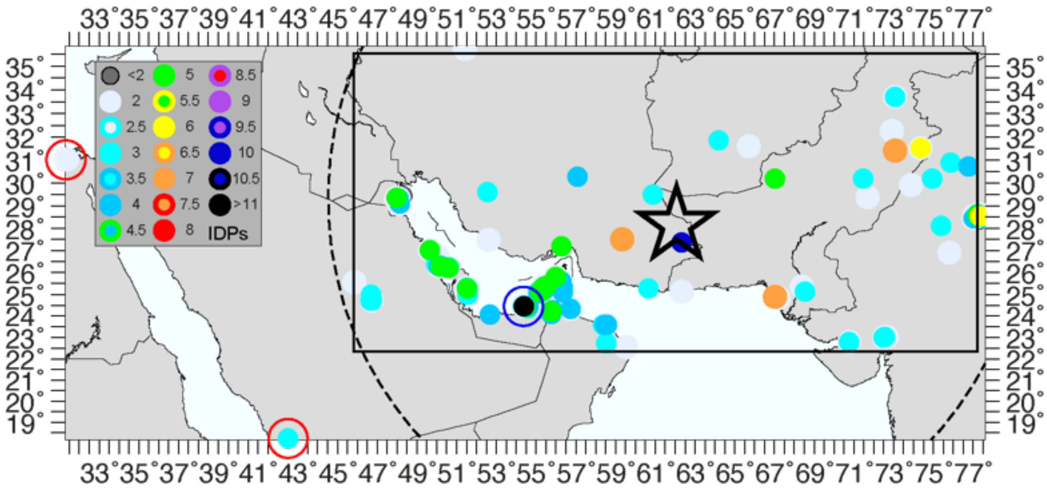


Figure 3 - example of geographical (circled in red) and intensity outliers (with raw intensity=12, circled in blue), for the 2013/04/16 10:44 M=7.8 earthquake, number of IDPs: 408. The black dashed circle indicates the Maximum Distance Prediction Equation (MDPE) used to delete farthest IDPs and define the area (solid black line) of minimum and maximum latitude and longitude of selected IDPs. The black star indicates the instrumental epicentre.

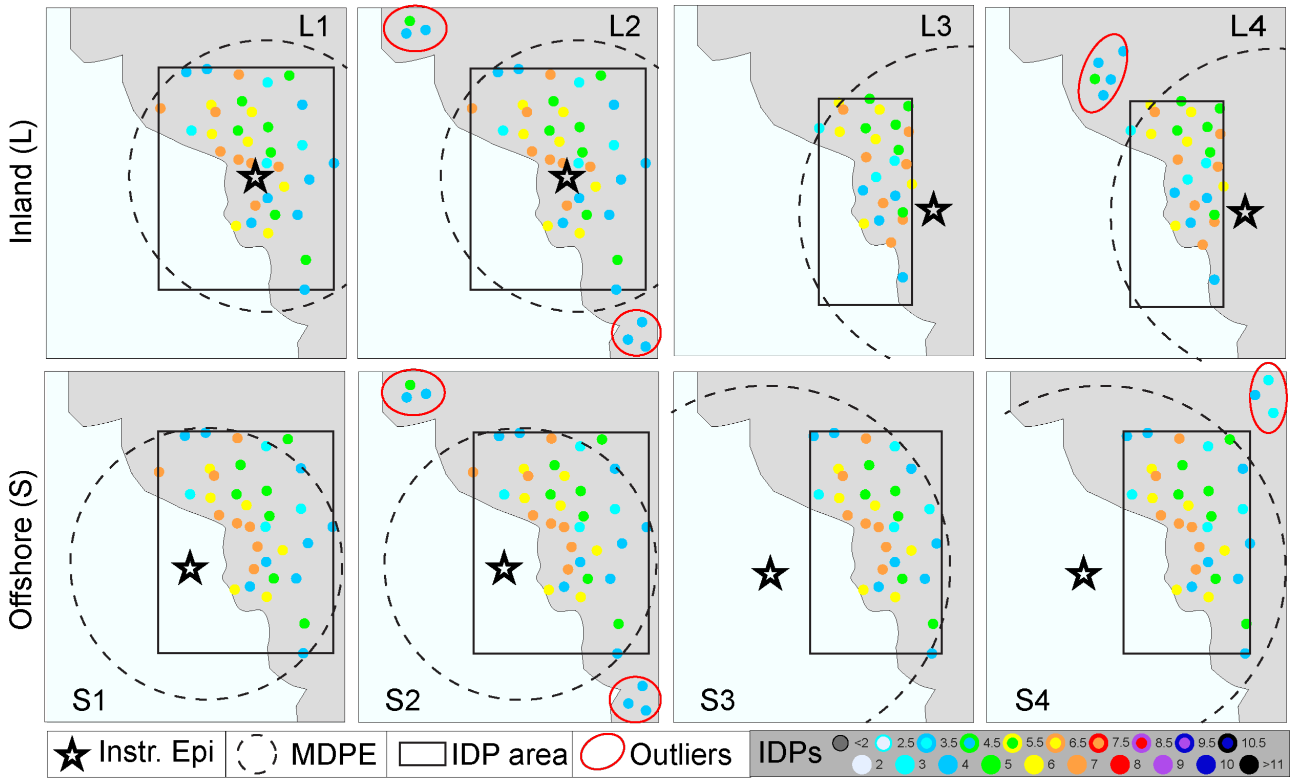


Figure 4 - scheme of classification of the distribution of IDPs. The star indicates the instrumental epicentre, the circular dashed black line is a circle with MDPE radius, the rectangular black line delimits the area of location of usable IDPs, i.e. the minimum and maximum latitude and longitude of usable IDPs, without any geographic outliers (circled in red colour).

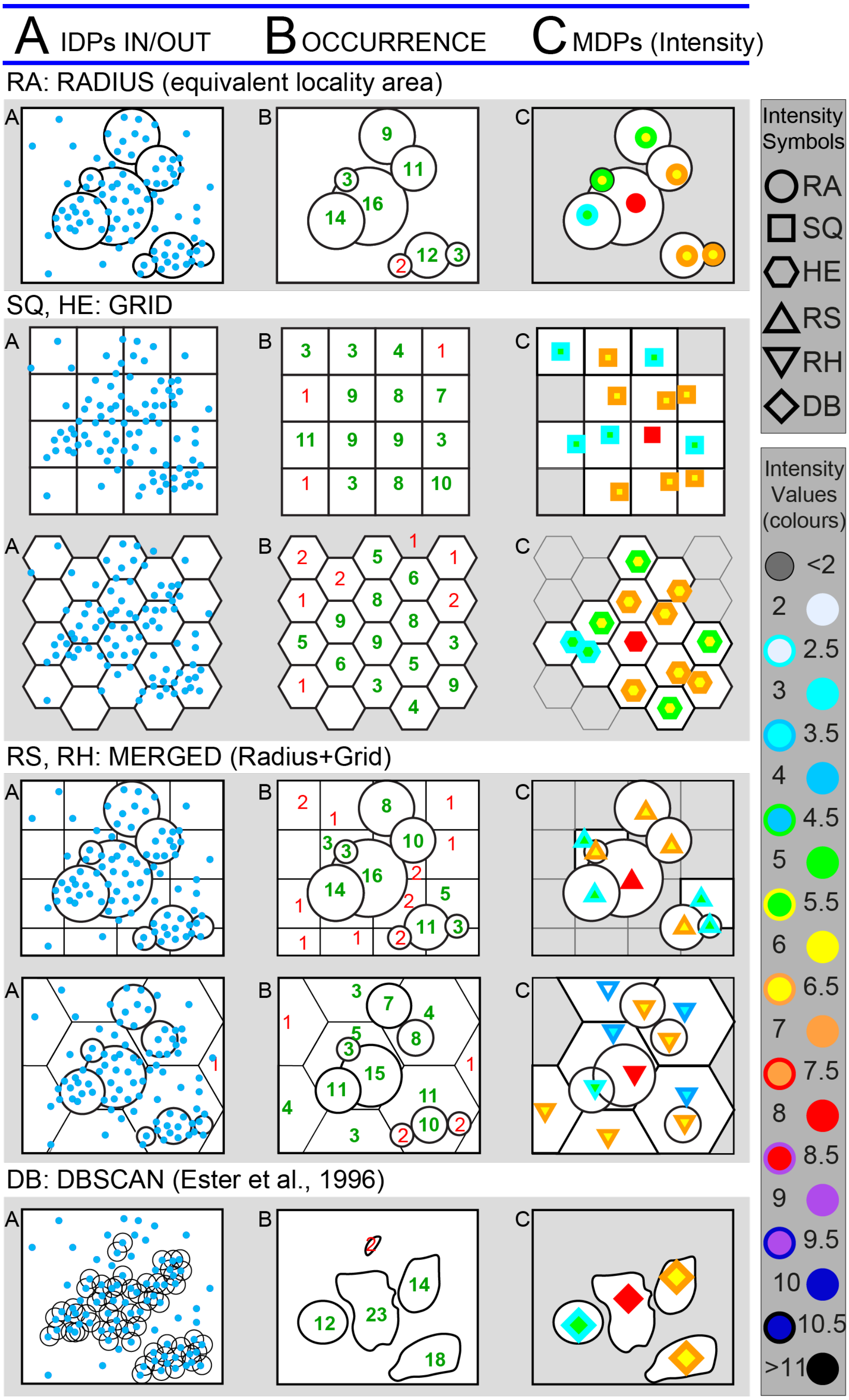


Figure 5 - Methods of clustering of IDPs into MDPs through three steps: column A: IDPs available are grouped (or not) following the various methods; column B: for each area of grouping the occurrence of a sufficient number of IDPs is assessed (numbers in green) or not (numbers in red); column C: IDPs are used to compute a combined intensity (MDPs), indicated with different colours

and symbols, for selected area/clusters (in white colours) and by using different central tendency estimators.

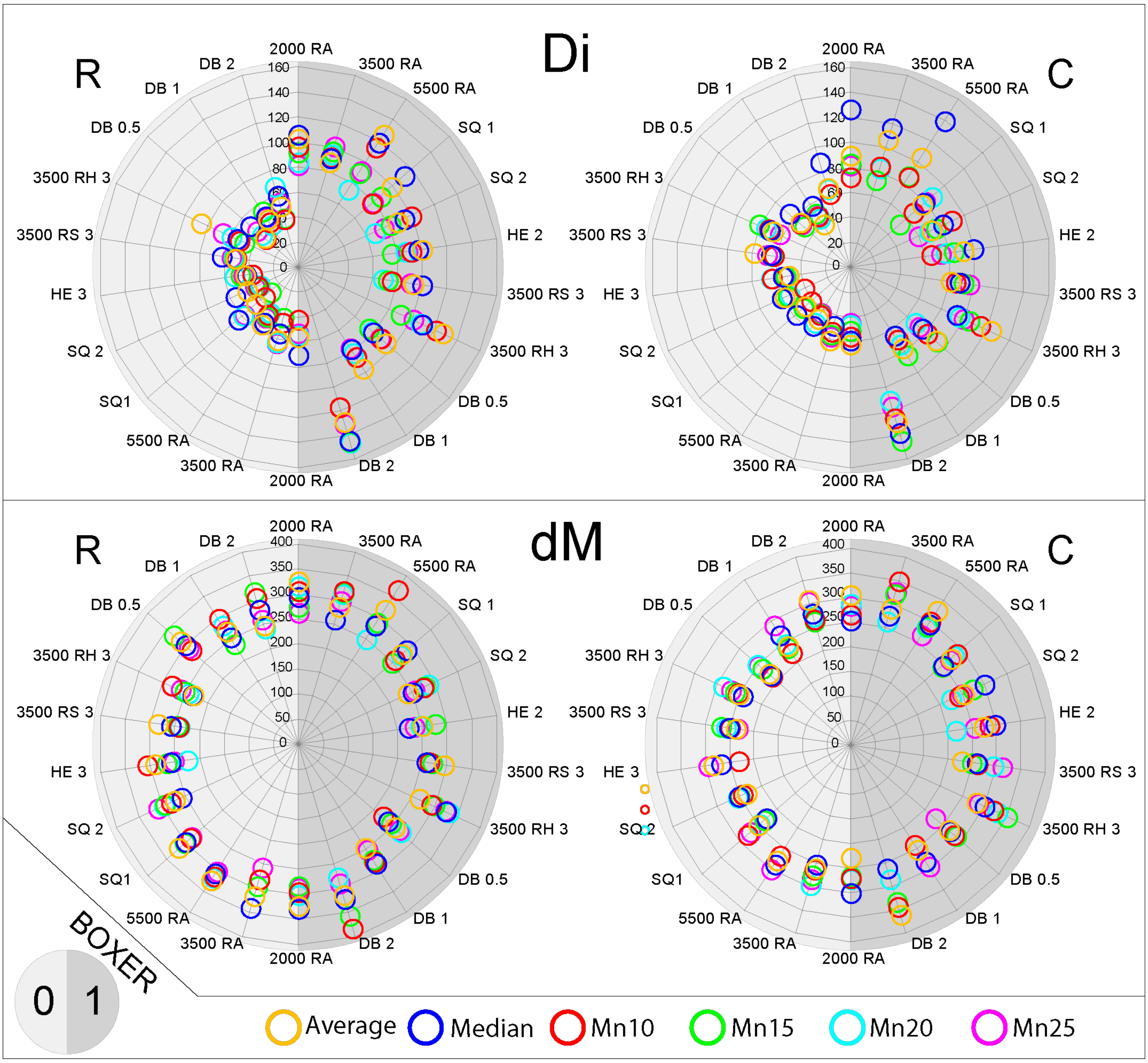


Figure 6 - Radar diagrams of the data represented in Tables 2 and S2 for Distance (Di, upper part) and difference of magnitude (dM, lower part) for BOXER-0 (Bx0) and BOXER-1 (Bx1). The light grey areas refer to BOXER-0 and the dark grey ones to BOXER-1. Coloured symbols (small circles) refer to central tendency estimators used to compute MDPs, plotted as a function of methods used to cluster raw IDPs (codified as in Table 2). The number of earthquakes (and then the agreement with instrumental data) increases from the centre of each circle outwards.

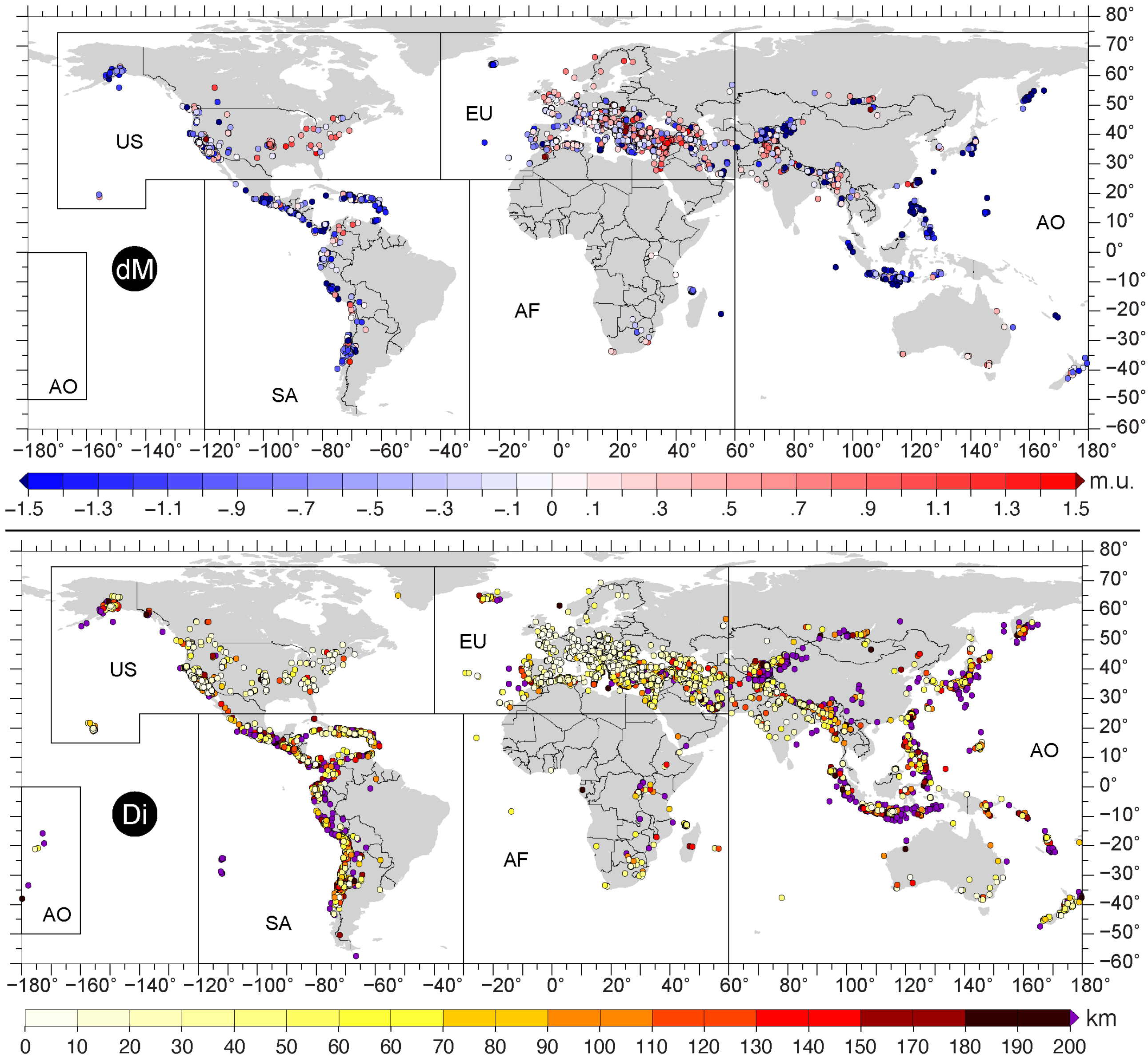


Figure 7 - Plot of distance (Di, lower panel, 15103 earthquakes) and magnitude difference of (dM, upper panel, 5703 earthquakes) between "preferred" macroseismic parameters and instrumental data for raw intensity. Five zones (EU=Europe, AO=Asia and Oceania, US=Nord America, SA= South America, AF=Africa) are shown.

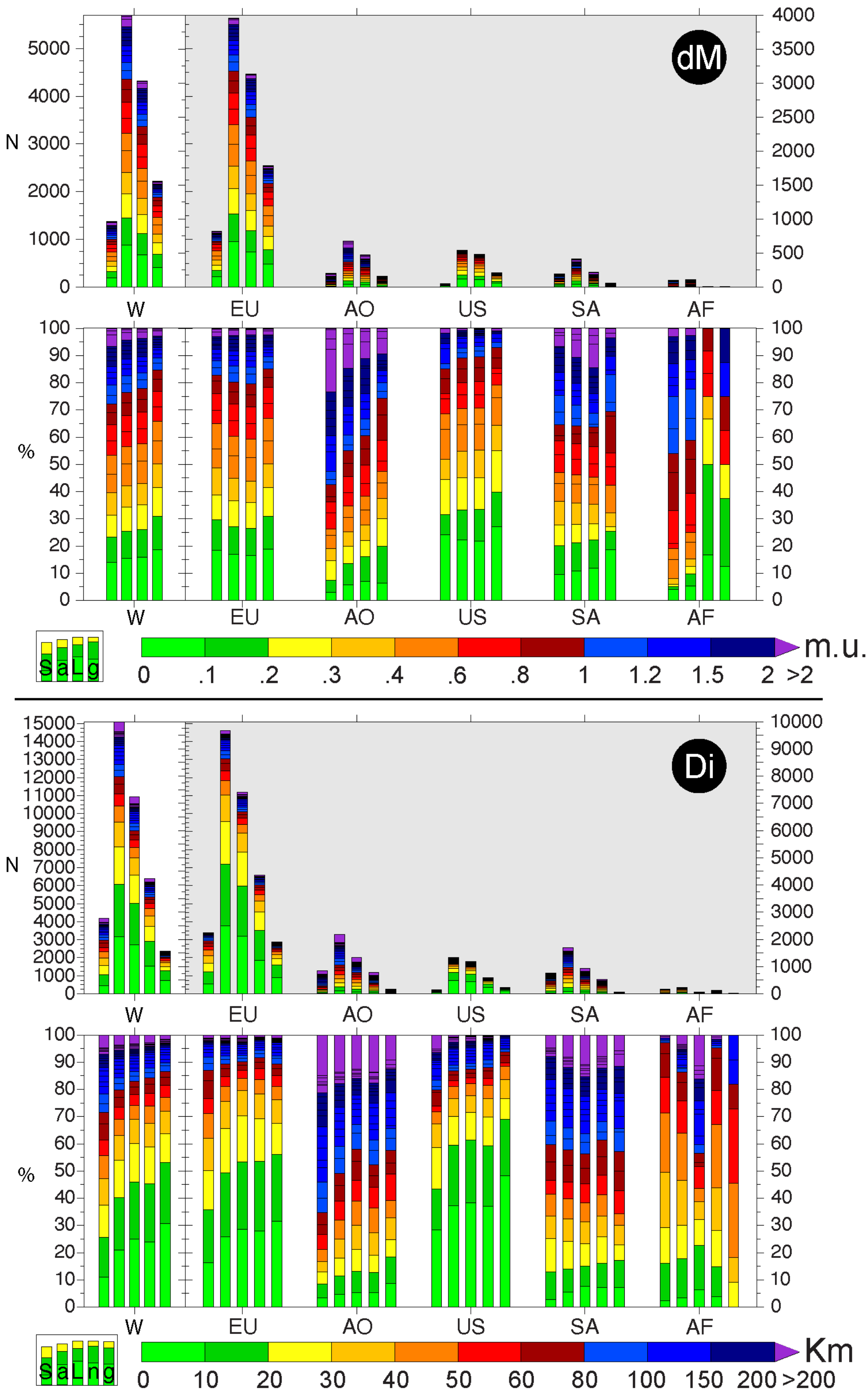


Figure 8 - Statistical results of the comparison between macroseismic and instrumental parameters (represented in Fig. 7). Plots display numbers of event (N) and percentages (%) for magnitude differences (dM, upper panel) and distances (Di, lower panel). Columns refer to global scale (W) and

different macro-areas (EU=Europe, AO=Asia and Oceania, US=Nord America, SA= South America, AF=Africa). The columns of each zone (see the legend in lowest left corner) indicate, from left to right, the earthquakes located offshore (S), all the earthquakes (a), earthquakes located inland (L), earthquakes with the number of MDPs ≥ 3 (n) and earthquakes with azimuthal gap $<$ to 180 degrees (g). The area in grey highlights the macro-areas with respect to the global area (W). The scales of the numbers of earthquakes (N) are different for the global area (left) and the macro-areas (right).

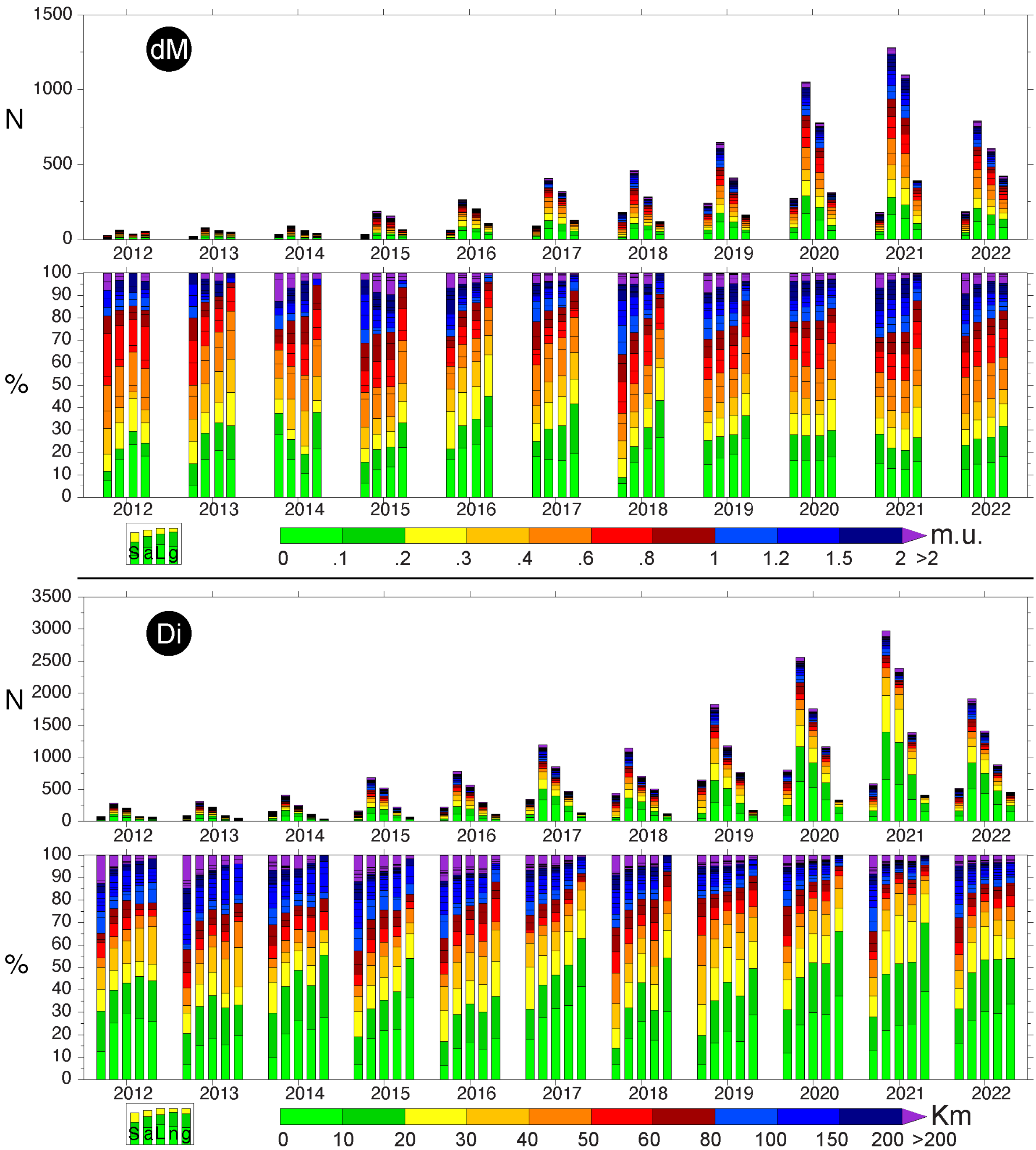


Figure 9 - Same as in Fig. 8, at the global scale and for different years.

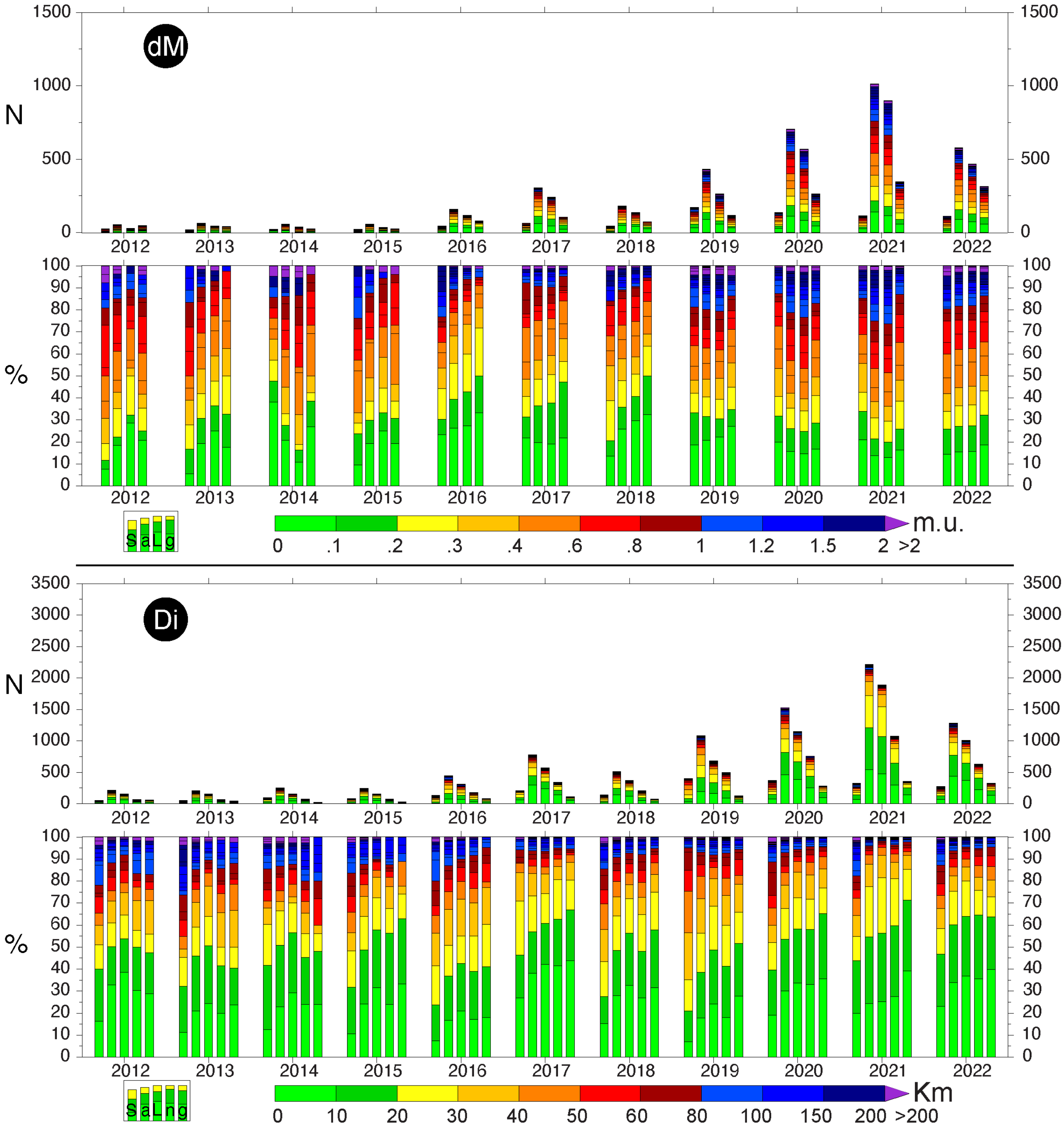


Figure 10 - Same as in Fig. 9 for Europe.

Appendix A - method for grouping IDPs and compute MDPs

The transformation from IDPs to MDPs implies a delimitation of the IDPs in a felt area and in some cases a selection of IDPs discarding the out-of-area data.

For retrospective statistical analyses or to derive relationships from available IDPs, the area is limited to the threshold distance defined by MDPE (eq. 1). The use of such a filter does not change or modify the number of MDPs that are actually calculated, as it only eliminates isolated IDPs (i.e. geographic outliers), but it does significantly reduce the calculation time required to create the subsequent geographic grids (for more than 15,000 earthquakes) on which to check cell by cell the relative occurrence of IDPs. IDPs available for each earthquake can be clustered or not in areas by 6 different methods (Fig. 5):

1. radius (RA)
2. square grid (SQ)
3. hexagonal grid (HE)
4. radius and square grid (RS, i.e. RA+SQ)
5. radius and hexagonal grid (RH, i.e. RA+HE)
6. DBSCAN method (DB)

In details:

1) RA method uses georeferenced localities (from a database) as cluster centres. Starting from the identification location, the radius constrains a representative surface of the location. IDPs can be in or out of the “city-equivalent” area. We use a database of global localities (i.e. the open source cities500.txt, see data and resource section) that also provide the number of inhabitants for each locality, however without population density information. Even if some databases as GHS Urban Centre Database (Florczyk et al., 2019, from GHSL (see data and resource section) collects open source information of the areas in km^2 only 13,000 cities in the World as collected. By setting a population density (e.g. 2000, 3500, 5500

inhabitants/km^2) it is possible derive a "city-equivalent" area, i.e. a spatial area roughly proportional to the number of inhabitants. The radius is computed as $\sqrt{(area/\pi)}$. IDPs located within the radius from the locality belong (IN) to the locality or not (OUT). Within each area, the clustering of IDPs starts with the localities with the smallest number of inhabitants and continues by grouping the remaining IDPs following the localities with increasing numbers of inhabitants.

2, 3) SQ and HE methods use a regular equal-areal grid with squared and hexagonal mesh, respectively. The centre of development of the grid is fixed to the average of coordinates of all IDPs inside the area.

4, 5) RS (RA+SQ) and RH (RA+HE) methods combine the method of clustering 1 with 2 and 3 respectively: first the RA method is applied, then remaining IDPs are grouped by the SQ or HE method. This approach overcomes in certain cases the simplification of equating the locality area to a circle based on a fixed population density and allows to retrieve information about IDPs outside of RAs but in a sufficient number so that to compute residual MDPs over grid.

6) DB (Density-Based Spatial Clustering of Applications with Noise-DBSCAN, Ester et al., 1996) is a method based on the grouping of IDPs located less than an arbitrary distance that successively can aggregate neighbour clusters and IDPs. If an IDP is close to another one (that is it is located at a distance, or "EPS" radius smaller than a given value) the two IDPs are grouped together in the same aggregation area. However, if one of aggregated IDPs is close to another IDP at a distance smaller than the EPS radius, then the latter IDP is joined together the aggregation area to which the former IDP belongs. This technique proceeds in a chain by joining IDPs to the cluster and is able to discover clusters of arbitrary shape. IDPs at a distance greater than the EPS radius from all the IDPs of the cluster are external or belong to other, distinct, aggregation areas.

For each grouping method (1-6) setting parameter (e.g. population density, grid side or EPS distance) constrain the areas for IDP grouping. We derived MDPs using various combinations of grouping methods and central tendency estimators and varying the reference settings. In particular (Table A1) we used:

- three population densities of 2000, 3500 and 5500 inhabitants/km^2 using the RA method;
- regular equidimensional grids with side of the mesh of 1 and 2 km for squared cells (SQ) and 2 km for hexagonal cells (HE);
- a population density of 3500 inhabitants/km^2 and side of the grid of 3 km both for RS and RH methods
- three eps radii (0.5, 1, 2 km) for DB methods

The methods for assessing MDPs are not equivalent to each-other in terms of computing time. In table A1 the last column gives a raw evaluation of the computational speed of the method (high, average, low speed and relative comparison with "+" and "-" symbols). Grouping methods based on SQ and HE grids require more computer time than RA or DB (Table A1) methods. The RS and RH methods are intermediate between the previous approaches. The construction of grids requires a complete coverage of the whole area and the smaller the size of the grid side, the longer the time to construct the grid and therefore to search for IDPs within each cell. The tessellation with hexagonal cells, due to a higher complexity, is more time-consuming than that with square cells. RS and RH methods use grids with sides slightly wider than SQ and HE ones, so they are faster than SQ and HE methods. In any case, the higher the level of detail one wants to achieve as spatial coverage, the more the time needed to perform computations. The DB methods (Fig. 5) is independent of external data, like locality databases, or of grid tessellation and is based only on the available information (location and intensity) of the IDPs.

We tested for the all the earthquakes of the EMSC dataset the combinations of different grouping methods and settings. To simplify the discussion of analyses and statistics we then selected some settings only indicated in Table A1. Combinations can be represented by combining acronyms:

"3500RH3-mean" uses both the radius (R) of "city-equivalent" area, based on a population density of 3500 inhabitants/km2 and hexagonal cells (H) of side 3 km as grouping method and the median to derive the MDP intensity, while "2000RA-mdna" uses a radius (R) with population density of 2000 inhabitants/km2 and the median.

Other methods of data clustering (e.g. based on polylines of the limit urban areas at different sites) are not available on a global scale with the same quality: some countries may have these data even for small locations while others do not. Note that determining whether or not an IDP falls within a polyline is a time-consuming calculation.

For future near real-time analyses the instrumental location and magnitude of events are unknown when IDPs are made available since the event time (T0). The IDPs collected at the subsequent time steps (T1, T2, Tn...) directly define the maximum and minimum latitude and longitude of the survey area because the analysis of only one event at a time does not create problems of excessive calculation time. All the IDPs (even the geographical outliers) will be tested to verify their occurrence in the grouping areas for the assessment of the MDPs. In any case, geographic outliers are generally isolated (i.e. below the expected threshold (3) of minimum IDP occurrence to assign an MDP) and do not contribute to the creation of MDPs.

Intensities: raw (R), corrected (C)				Speed test
Grouping methods	central tendency estimators (mean, mdna, mn10, mn15, mn20, mn25)			
	Population density (Den) (inhabitants/km^2)	Side of cells (gr) (in km)	EPS radius (eps) (in km)	
RA	2000, 3500,5500			Fast
SQ		1, 2		Slow+
HE		2		Slow-
RS	3500	3		Medium+
RH	3500	3		Medium-
DB			0.5, 1, 2	Fast

Table A1 - Summary of grouping methods and settings used to derive MDPs from IDPs. The speed test is a relative indication of the processing time of MDPs from slowest to fastest, with further intermediate levels (+, -). In total, 11 rouping Methods, 6 central tendency estimators, 2 type of Intensity (R and C) are used for comparative analyses on the EMSC earthquakes.

Appendix B – example of the procedure from IDPs to Macroseismic parameters

To better represent the procedure from the IDPS to the choice of preferred macroseismic parameters of an earthquake, we show as an example, the earthquake of 19 September 2020 6:38 UTC, Lat: 34.02, Lon: -118.08, Mw=4.5) referred to the raw intensities only.

For this event, EMSC provides 2192 IDPs (Fig. B.1). The grouping of IDPs into MDPs involves selecting IDPs by discarding geographical outliers (not present in the example, however) and grouping them into MDPs using clustering methods and central tendency estimators for a total of 66 possible MDP distributions (see also Appendix A and Table A.1 for details).

Fig. B.2 shows the MDPs obtained by applying 11 grouping methods and the median as the central tendency distribution. For each group of 66 combinations, the BOXER provides location and magnitude with methods 0 and 1, giving a total of 132 locations and magnitudes. Fig. B.3 (with numerical values in Table B.1) shows the epicentres and the differences in magnitude with respect to the instrumental values. Most of the macroseismic epicentres are very close to the true instrumental one (the maximum distance is about 19 km) generally with small differences in magnitude (the overall range is between -0.3 and 1.5 m.u.).

To choose a preferred location and magnitude, we applied the ranking order (Tables 3 and 4). For the example earthquake, macroseismic parameters are available for all 132 possible combinations (Table B.1), so the first ranked combination was chosen for both distance (DBSCAN with eps 2 km, trimmed mean 20 and BOXER-1, Table 3) and magnitude (DBSCAN with eps 2 km, trimmed mean 10 and BOXER-1, Table 4). The macroseismic preferred solution is located at latitude: 33.9829, longitude: -118.0443, with magnitude: 4.62 (Fig. B.3). The distance with respect to instrumental epicentre is 5.28 km and the difference of magnitude 0.1 m.u.

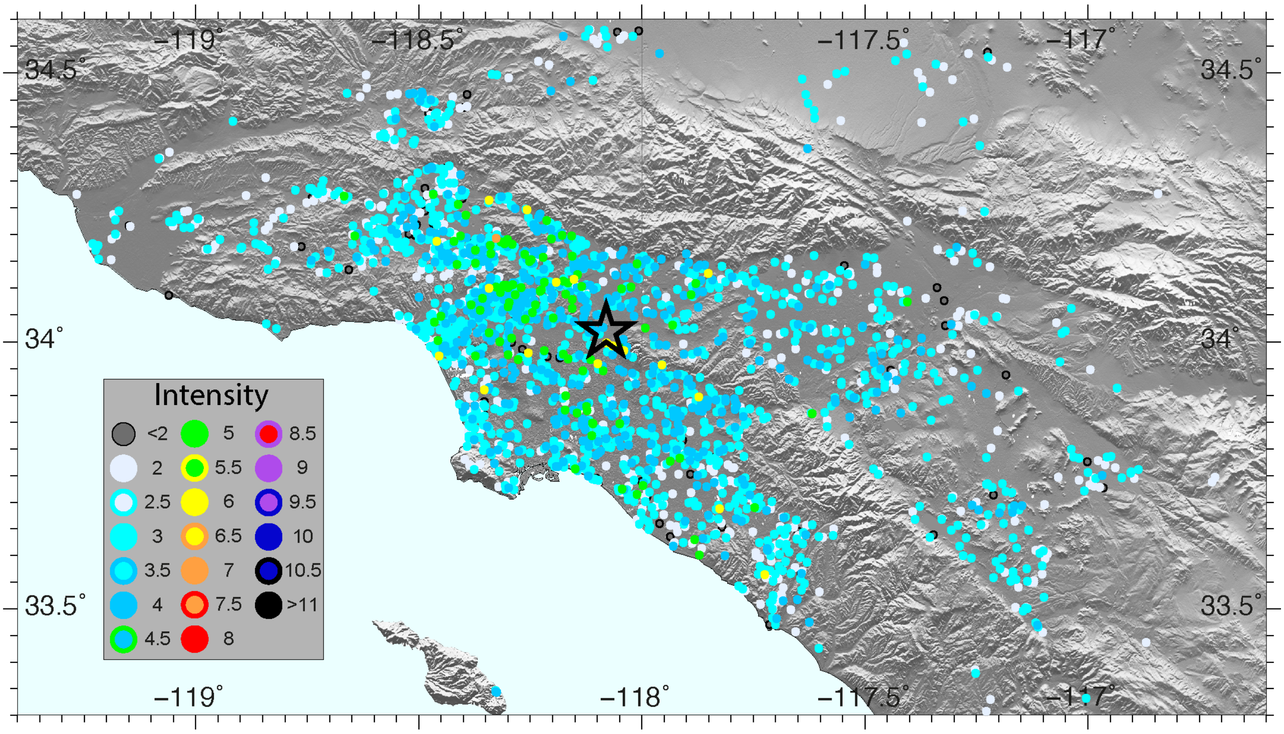
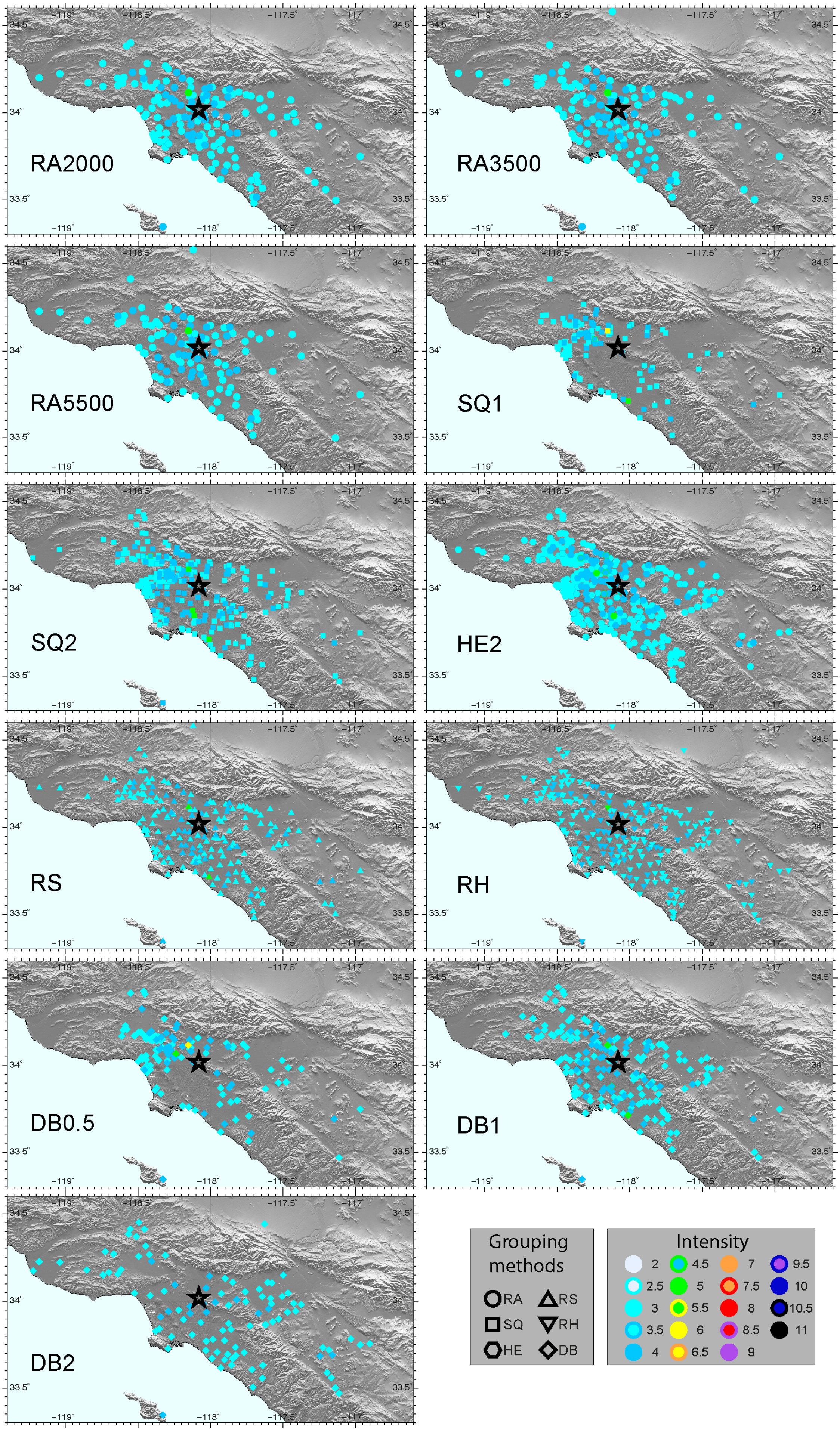
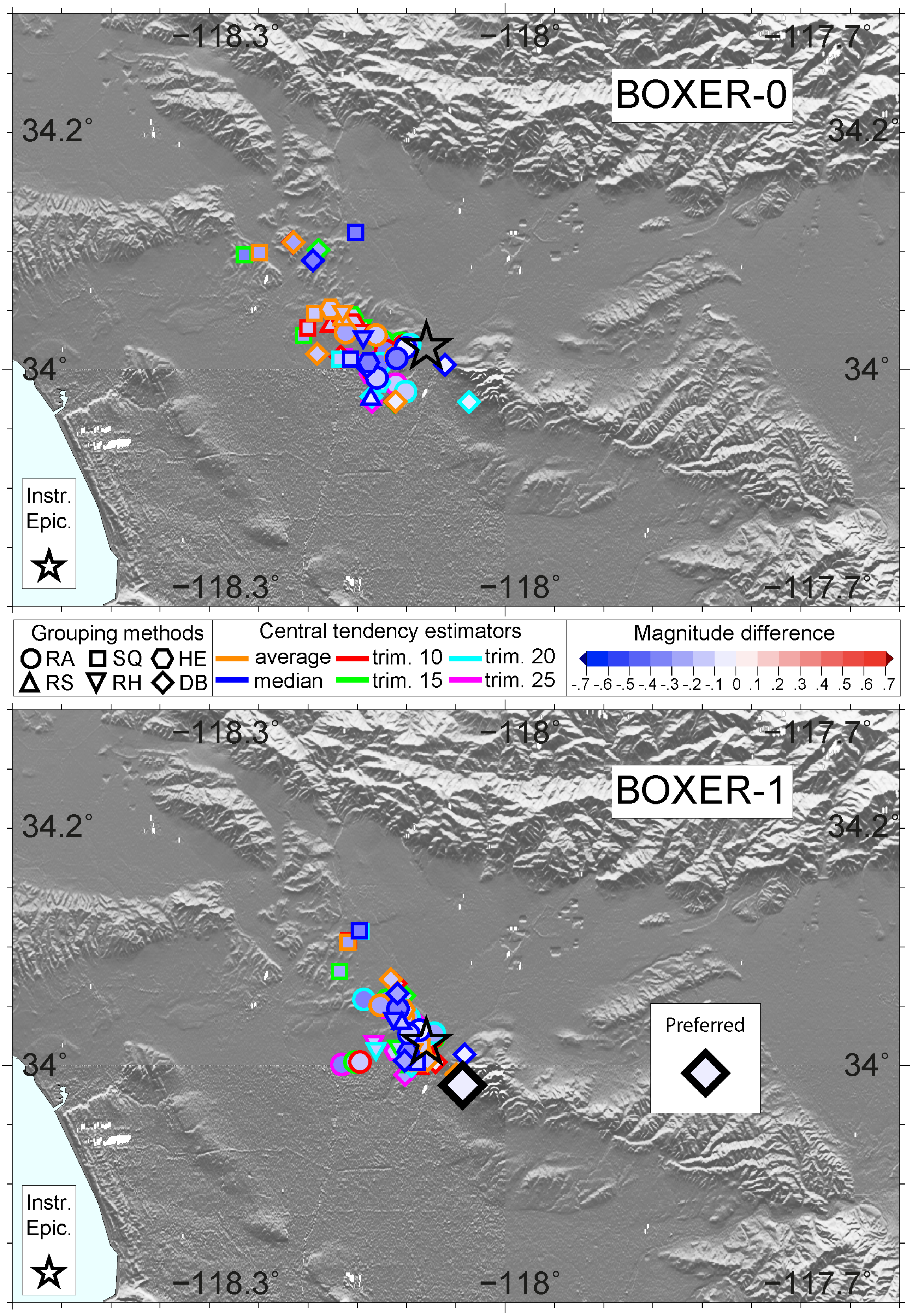


Figure B.1: Plot of 2692 IDPs (raw intensities) of the event of 2020/09/19. The star represents the instrumental epicentre.



RA2000
RA3500
RA5500
SQ1
SQ2
HE2
RS
RH
DB0.5
DB1
DB2
Grouping methods
RA
RS
SQ
RH
HE
DB
Intensity
2
2.5
3
3.5
4
4.5
5
5.5
6
6.5
7
7.5
8
8.5
9
9.5
10
10.5
11

Figure B.2: MDPs for 11 different grouping methods (RA, SQ, HE, RS, RH, DB and relative settings, see Appendix A and in Table A1) and the median as central tendency estimator. The star represents the instrumental epicentre.



BOXER-0
−118.3°
−118°
−117.7°
34.2°
34°
Instr.
Epic.
Grouping methods
RA SQ HE
RS RH DB
Central tendency estimators
average
median
trim. 10
trim. 15
trim. 20
trim. 25
Magnitude difference
-.7 -.6 -.5 -.4 -.3 -.2 -.1 0 .1 .2 .3 .4 .5 .6 .7
BOXER-1
Preferred

Figure B.3: Macroseismic parameters (location and difference of magnitude with respect to instrumental one) with BOXER-0 and BOXER-1 for a total of 132 different MDPs distributions. The preferred solution following the ranking order is indicated (numerical values in Table B.1).

BX mth	Lon (deg+km)	Lat (deg+km)	M	Di (km)	dM (m.u.)	n MDPs	Den	MGs	gr/eps (km)	CTEs
0	-118.1107±3.6	34.0106±2.4	4.49±.22	3.02	-0.01	171	2000	RA	-	mnsa
	-118.1012±2.7	34.0185±2.4	4.65±.24	1.96	0.15					mdna
	-118.1027±3.4	34.0200±2.6	4.49±.22	2.09	-0.01					mn10
	-118.1057±3.0	34.0224±2.2	4.5±.22	2.38	0					mn15
	-118.0968±3.5	34.0214±2.7	4.49±.19	1.55	-0.01					mn20
	-118.1075±3.4	34.0197±2.4	4.49±.19	2.54	-0.01					mn25
	-118.1308±3.8	34.0290±2.8	4.52±.23	4.79	0.02	146	3500			mnsa
	-118.1101±3.6	34.0097±2.8	4.32±.26	3	-0.18					mdna
	-118.1212±3.4	34.0162±4.4	4.39±.24	3.82	-0.11					mn10
	-118.1247±3.6	34.0241±3.7	4.4±.24	4.14	-0.1					mn15
	-118.1268±3.7	34.0051±3.5	4.38±.25	4.62	-0.12					mn20
	-118.1318±3.8	34.0057±2.7	4.37±.25	5.03	-0.13					mn25
	-118.162±3.9	34.0311±4.1	4.49±.24	7.65	-0.01	126	5500			mnsa
	-118.1302±3.2	33.9933±2.7	4.46±.50	5.5	-0.04					mdna
	-118.1334±3.6	33.9983±4.3	4.49±.22	5.48	-0.01					mn10
	-118.1328±3.6	33.9983±4.3	4.49±.22	5.44	-0.01					mn15
	-118.1011±3.4	33.9812±3.7	4.41±.28	4.73	-0.09					mn20
	-118.1108±3.9	33.9877±3.5	4.44±.36	4.58	-0.06					mn25
	-118.2493±2.9	34.0989±1.0	4.42±.19	17.9	-0.08	184	-	SQ	1	mnsa
	-118.1518±	34.1158±	4.3±.24	12.54	-0.2					mdna
	-118.2493±2.9	34.0989±1.0	4.43±.19	17.89	-0.07					mn10
	-118.2647±2.3	34.0972±1.1	4.38±.19	19.06	-0.12					mn15
	-118.1518±	34.1158±	4.34±.23	12.54	-0.16					mn20
	-118.1518±	34.1158±	4.33±.23	12.54	-0.17					mn25
	-118.1938±3.1	34.0475±1.6	4.58±.13	10.93	0.08	347			2	mnsa
	-118.1567±3.4	34.0089±1.5	4.57±.12	7.17	0.07					mdna
	-118.2±2.8	34.0354±1.5	4.57±.14	11.2	0.07					mn10
	-118.2043±2.8	34.0289±1.6	4.55±.14	11.5	0.05					mn15
	-118.1676±3.7	34.0091±1.6	4.59±.12	8.16	0.09					mn20
	-118.1631±3.7	34.0111±1.6	4.58±.12	7.73	0.08					mn25
	-118.1771±2.0	34.0511±1.5	4.58±.14	9.6	0.08	356	-	HE		mnsa
	-118.1386±3.1	34.0058±1.3	4.55±.12	5.63	0.05					mdna
	-118.1544±2.4	34.0388±1.6	4.58±.14	7.17	0.08					mn10
	-118.1554±2.7	34.0448±1.8	4.54±.13	7.47	0.04					mn15
	-118.1274±3.3	33.9897±1.5	4.54±.13	5.52	0.04					mn20
	-118.1371±3.0	33.9966±1.5	4.54±.13	5.87	0.04					mn25
	-118.1647±3.0	34.0423±1.8	4.54±.14	8.19	0.04	358	3500	RS	3	mnsa
	-118.1364±4.3	33.9753±8.0	4.58±.12	7.19	0.08					mdna
	-118.1776±2.7	34.0368±2.0	4.52±.14	9.18	0.02					mn10

	-118.1722±2.4	34.0366±2.1	4.5±.14	8.7	0					mn15
	-118.1364±4.3	33.9753±8.0	4.55±.13	7.19	0.05					mn20
	-118.1364±4.3	33.9753±8.0	4.55±.13	7.19	0.05					mn25
	-118.1652±3.2	34.0490±2.5	4.55±.17	8.49	0.05	268	3500	RH	3	mnsa
	-118.1441±3.4	34.0276±1.8	4.47±.19	5.97	-0.03					mdna
	-118.1475±3.9	34.0330±2.4	4.55±.17	6.38	0.05					mn10
	-118.1458±3.9	34.0370±2.7	4.56±.17	6.35	0.06					mn15
	-118.1423±5.3	34.0055±2.4	4.49±.19	5.96	-0.01					mn20
	-118.1446±5.0	34.0076±2.3	4.5±.19	6.12	0					mn25
	-118.2151±4.1	34.1075±3.9	4.48±.19	15.8	-0.02	179	-	DB	0.5	mnsa
	-118.1953±4.0	34.0920±2.6	4.34±.24	13.3	-0.16					mdna
	-118.2152±4.1	34.1075±3.9	4.49±.19	15.81	-0.01					mn10
	-118.1897±4.7	34.1010±4.4	4.49±.20	13.54	-0.01					mn15
	-118.1953±4.0	34.0920±2.6	4.39±.23	13.3	-0.11					mn20
	-118.1953±4.0	34.0920±2.6	4.39±.24	13.3	-0.11					mn25
	-118.191±2.8	34.0133±3.8	4.6±.17	10.26	0.1	241			1	mnsa
	-118.1405±4.6	34.0022±3.4	4.41±.21	5.91	-0.09					mdna
	-118.1668±3.6	34.0107±4.0	4.65±.17	8.07	0.15					mn10
	-118.1669±3.6	34.0107±4.0	4.63±.17	8.08	0.13					mn15
	-118.1355±5.7	33.9773±4.0	4.62±.15	6.98	0.12					mn20
	-118.1354±5.4	33.9726±3.8	4.62±.15	7.34	0.12					mn25
	-118.1115±6.5	33.9731±1.3	4.61±.32	5.96	0.11	128			2	mnsa
	-118.061±8.2	34.0041±11.8	4.65±.23	2.48	0.15					mdna
	-118.1119±6.5	33.9731±1.3	4.62±.32	5.99	0.12					mn10
	-118.112±6.4	33.9725±1.3	4.61±.31	6.05	0.11					mn15
	-118.0366±13.9	33.9725±20.3	4.61±.29	6.63	0.11					mn20
	-118.0368±13.9	33.9728±20.3	4.6±.29	6.59	0.1					mn25
1	-118.0934±2.3	34.0103±2.7	4.49±.23	1.64	-0.01	171	2000	RA	-	mnsa
	-118.0959±2.1	34.0307±2.6	4.66±.29	1.89	0.16					mdna
	-118.08±2.2	34.0264±4.1	4.5±.22	0.72	0					mn10
	-118.0786±2.3	34.0249±3.5	4.51±.23	0.56	0.01					mn15
	-118.0801±2.4	34.0314±3.2	4.5±.25	1.27	0					mn20
	-118.1013±2.3	34.0367±2.7	4.5±.26	2.7	0					mn25
	-118.1202±3.2	34.0522±3.4	4.55±.24	5.16	0.05	146	3500			mnsa
	-118.1267±3.1	34.0573±2.9	4.34±.26	5.98	-0.16					mdna
	-118.1176±3.2	34.0552±3.9	4.42±.25	5.23	-0.08					mn10
	-118.1344±4.1	34.0659±3.2	4.43±.26	7.15	-0.07					mn15
	-118.1441±4.3	34.0597±3.2	4.39±.26	7.38	-0.11					mn20
	-118.1683±3.7	33.9997±3.0	4.36±.26	8.45	-0.14					mn25
	-118.1394±3.8	34.0625±3.2	4.51±.30	7.23	0.01	126	5500			mnsa
	-118.1114±2.5	34.0323±3.4	4.47±.50	3.2	-0.03					mdna
	-118.1567±3.0	34.0008±2.8	4.49±.30	7.39	-0.01					mn10
	-118.1608±3.0	34.0008±2.7	4.49±.30	7.75	-0.01					mn15
	-118.1143±3.2	34.0493±3.4	4.4±.24	4.54	-0.1					mn20
	-118.1164±3.4	34.0479±3.7	4.39±.27	4.57	-0.11					mn25
	-118.1637±2.1	34.1082±2.9	4.55±.21	12.48	0.05	184	-	SQ	1	mnsa
	-118.1542±1.6	34.1147±1.5	4.3±.24	12.55	-0.2					mdna
	-118.1628±2.0	34.1088±2.6	4.55±.21	12.48	0.05					mn10
	-118.176±2.7	34.1026±6.6	4.5±.21	12.75	0					mn15
	-118.1533±1.7	34.1142±1.6	4.33±.23	12.46	-0.17					mn20
	-118.1534±1.7	34.1144±1.5	4.32±.23	12.49	-0.18					mn25

	-118.099±2.0	34.033±2.8	4.63±.15	2.27	0.13	347			2	mnsa
	-118.1007±2.1	34.0123±2.5	4.6±.13	2.09	0.1					mdna
	-118.102±2.1	34.0239±2.7	4.62±.15	2.07	0.12					mn10
	-118.1023±2.0	34.0278±2.8	4.61±.16	2.23	0.11					mn15
	-118.1121±2.3	34.0272±2.9	4.62±.13	3.06	0.12					mn20
	-118.1142±2.3	34.0316±2.9	4.6±.14	3.4	0.1					mn25
	-118.1206±2.1	34.052±2.4	4.6±.16	5.17	0.1	356	-	HE	2	mnsa
	-118.1109±2.1	34.0143±2.2	4.56±.16	2.92	0.06					mdna
	-118.0914±1.9	34.0193±2.5	4.57±.15	1.06	0.07					mn10
	-118.1056±2.0	34.0341±2.9	4.55±.15	2.84	0.05					mn15
	-118.1114±2.1	34.0096±2.0	4.54±.14	3.12	0.04					mn20
	-118.1269±2.2	34.0191±2.3	4.54±.14	4.33	0.04					mn25
	-118.1092±2.1	34.0056±1.8	4.54±.17	3.13	0.04	358	3500	RS	3	mnsa
	-118.1342±2.5	34.051±3.0	4.56±.13	6.07	0.06					mdna
	-118.1014±2.0	34.0014±1.7	4.53±.17	2.86	0.03					mn10
	-118.1186±2.4	34.0058±1.7	4.51±.18	3.89	0.01					mn15
	-118.1226±2.4	34.0526±2.7	4.54±.14	5.34	0.04					mn20
	-118.1348±2.3	34.0542±2.6	4.53±.14	6.32	0.03					mn25
	-118.1092±3.5	34.0141±3.0	4.53±.20	2.77	0.03	268	3500	RH	3	mnsa
	-118.1257±2.5	34.0375±2.1	4.47±.19	4.63	-0.03					mdna
	-118.1139±2.8	34.0213±3.9	4.54±.19	3.13	0.04					mn10
	-118.1209±3.3	34.0185±4.0	4.55±.20	3.78	0.05					mn15
	-118.1385±2.7	34.0154±2.5	4.5±.20	5.42	0					mn20
	-118.1407±2.8	34.0219±2.6	4.51±.19	5.6	0.01					mn25
	-118.154±3.0	34.1121±4.4	4.56±.23	12.3	0.06	179	**-**	**DB**	0.5	mnsa
	-118.1536±2.7	34.1112±3.2	4.38±.24	12.2	-0.12					mdna
	-118.151±2.8	34.1116±4.2	4.57±.22	12.1	0.07					mn10
	-118.151±2.9	34.1124±4.0	4.54±.22	12.18	0.04					mn15
	-118.1516±2.8	34.1098±3.5	4.43±.23	11.97	-0.07					mn20
	-118.1521±2.8	34.1103±3.4	4.43±.24	12.04	-0.07					mn25
	-118.0917±2.5	34.0065±3.2	4.61±.18	1.85	0.11	241			1	mnsa
	-118.1157±2.7	34.0076±3.0	4.41±.20	3.56	-0.09					mdna
	-118.0806±2.6	34.0043±2.8	4.66±.19	1.74	0.16					mn10
	-118.0917±2.5	34.0117±3.0	4.65±.20	1.42	0.15					mn15
	-118.1091±2.8	34.0011±2.8	4.62±.19	3.41	0.12					mn20
	-118.1164±2.9	33.9924±2.6	4.62±.20	4.55	0.12					mn25
	-118.0483±2.9	33.9885±2.9	4.59±.30	4.56	0.09	**128**			**2**	mnsa
	-118.0378±3.7	33.9972±3.4	4.65±.46	4.64	0.15					mdna
	-118.0481±2.9	33.9853±2.7	**4.62±.27**	4.86	**0.12**					**mn10**
	-118.0469±2.8	33.9844±2.6	4.6±.28	5	0.1					mn15
	-118.0443±2.7	**33.9829±2.5**	4.61±.27	**5.28**	0.11					**mn20**
	-118.0407±4.7	33.9799±4.3	4.6±.29	5.74	0.1					mn25

Table B1. Numerical values of the data in Fig. B.3 for BOXER method (Bx mth) 0 and 1, 11 grouping methods (MGs) and used settings (Den, gr/eps as in Table A.1) e 6 central tendency estimators (CTEs). Macroseismic latitudes, longitudes and magnitudes also report the uncertainties values computed by BOXER. Distance (Di) and difference of magnitude (dM) with respect to instrumental values are indicated. The preferred data (BOXER-1, DBSCAN with eps 2 km, trimmed mean 20 for location and mean 10 for magnitude) in bold characters.

Supplementary material of:

Earthquakes parameters from citizen testimonies. A retrospective analysis of EMSC database

Gianfranco Vannucci[1*], Paolo Gasperini[2,1], Laura Gulia[2] and Barbara Lolli[1]

[1]Istituto Nazionale di Geofisica e Vulcanologia, Sezione di Bologna

[2]Dipartimento di Fisica e Astronomia, Università di Bologna

*** Corresponding author**

Declaration of Competing Interests:

The authors acknowledge there are no conflicts of interest recorded.

Description of the Supplemental Material

This supplementary material contains figures and tables that provide further information and details of the main text. Moreover, similar elaborations, plots and figures are given for the “corrected” intensities in as for the “raw” intensities in the main text.

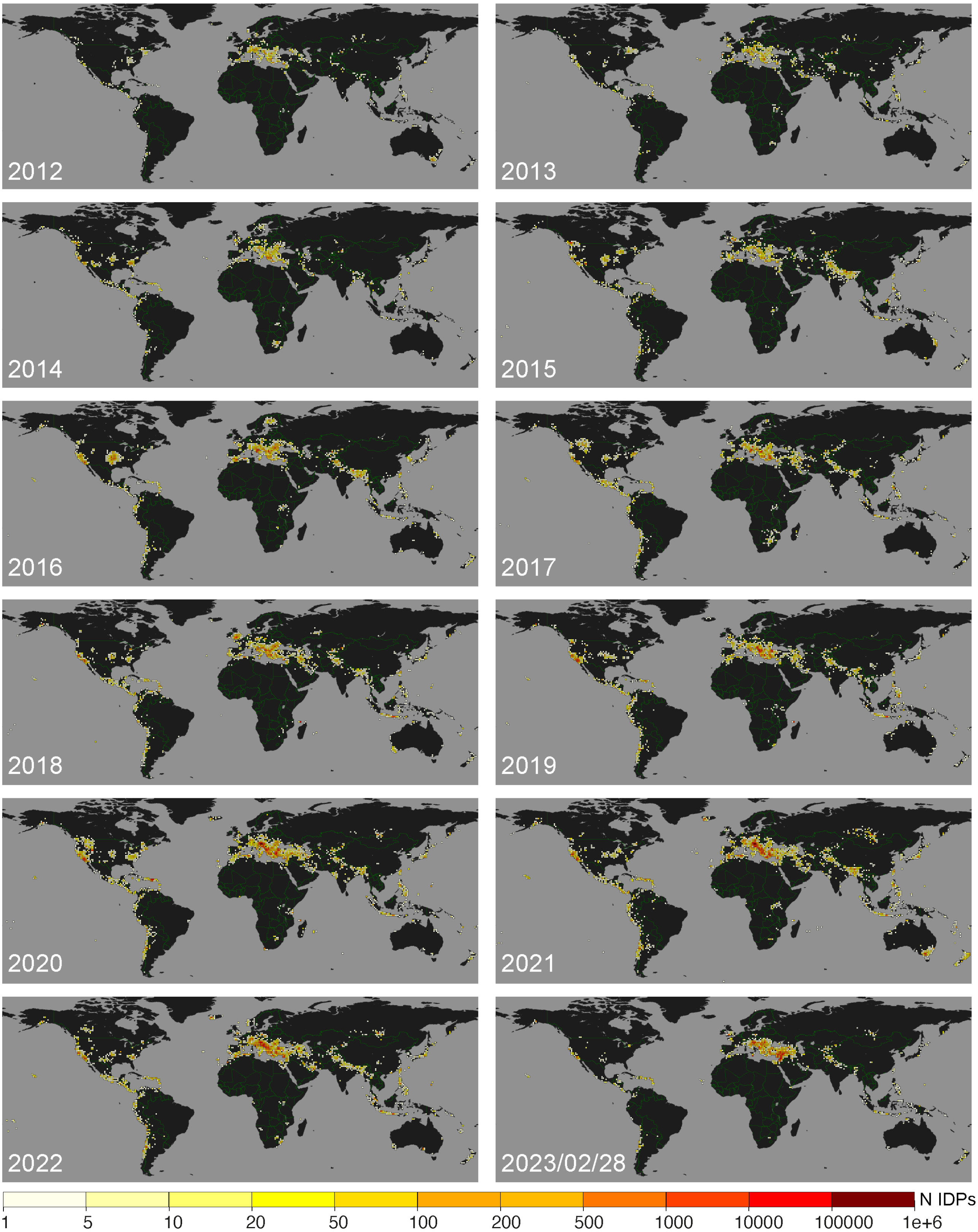


Figure S1 - IDPs occurrence per year over a grid of 1x1 degree both in latitude and longitude in function of the year

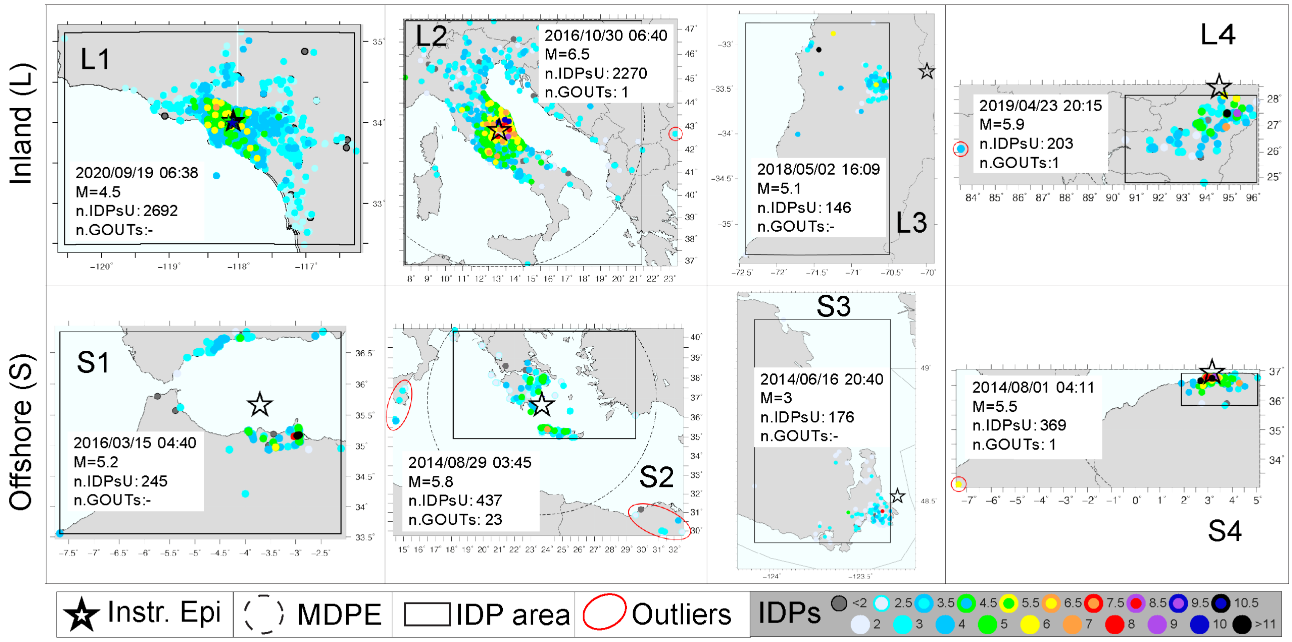


Figure S2 - examples of earthquakes classified following the scheme of Fig. 4. If no MDPE radius is shown all the IDPs are selected, without geographic outliers.

A)	N Eqks	Epicentre DISTANCE from Coast Line (km)				
		≤2	>2,≤5	>5,≤10	>10,≤50	>50
L1	9042	-	-	-	-	-
L2	1333	-	-	-	-	-
L3	5499	-	-	-	-	-
L4	166	-	-	-	-	-
TOT L	**16040**	**-**	**-**	**-**	**-**	**-**
S1	1972	128	289	343	948	264
S2	386	34	75	68	135	74
S3	4140	69	187	318	2100	1466
S4	223	6	16	26	83	92
TOT S	**6721**	**237**	**567**	**755**	**3266**	**1896**
L+S	**22761**	**237**	**567**	**755**	**3266**	**1896**

B)	N Eqks	n.IDPs Usable				GAP IDPs-Epicentre (degrees)				
		≥300	≥100, <300	≥5, <100	≥3, <5	≤90	>90, ≤120	>120, ≤180	>180, ≤360	360
L1	9042	519	901	6869	753	2110	1333	3240	2359	0
L2	1333	263	280	780	10	521	215	330	267	0
L3	5499	6	41	3444	2008	0	0	0	5447	52
L4	166	2	12	119	33	0	0	0	165	1
TOT L	**16040**	**790**	**1234**	**11212**	**2804**	**2631**	**1548**	**3570**	**8238**	**53**
S1	1972	80	186	1558	148	205	233	784	750	0
S2	386	79	86	214	7	80	48	118	140	0
S3	4140	18	81	2730	1311	0	0	1	4102	37
S4	223	13	36	153	21	0	0	1	222	0
TOT S	**6721**	**190**	**389**	**4655**	**1487**	**285**	**281**	**904**	**5214**	**37**
L+S	**22761**	**980**	**1623**	**15867**	**4291**	**2916**	**1829**	**4474**	**13452**	**90**

C)	N Eqks	MAGNITUDE range								
		0-1	1-2	2-3	3-4	4-5	5-6	6-7	7-8	8-9
L1	9042	20	677	1710	3484	2313	716	102	20	0
L2	1333	13	48	127	308	425	293	93	25	1
L3	5499	6	173	1289	2059	1391	532	44	5	0
L4	166	2	3	17	37	59	35	13	0	0
TOT L	**16040**	**41**	**901**	**3143**	**5888**	**4188**	**1576**	**252**	**50**	**1**
S1	1972	0	1	242	775	598	283	55	15	3
S2	386	0	0	21	92	102	94	61	13	3
S3	4140	0	0	451	1189	1452	845	181	21	1
S4	223	0	0	5	18	80	83	29	7	1
TOT S	**6721**	**0**	**1**	**719**	**2074**	**2232**	**1305**	**326**	**56**	**8**
L+S	**22761**	**41**	**902**	**3862**	**7962**	**6420**	**2881**	**578**	**106**	**9**

Table S1: Occurrence of the 22,761 EMSC earthquakes with number of IDPs ≥3, after deleting geographic outliers and using IDPs in the intensity range of 3-11 degrees. Occurrences are shown for 8 categories of classification: inland (L), offshore (S) and classification number 1-4, see Fig. 4. Occurrence in function of: minimum distance (for S-earthquakes only) of instrumental epicentre from the coast line (panel A), number of usable IDPs and the maximum gap between IDPs and epicentre (panel B), range of magnitude (panel C). The offshore/inland location of the instrumental epicentre is established through the high-resolution polylines of the Global Self-consistent Hierarchical High-resolution Geography (GSHHG, see data and resource section).

Dset: A n eqks: 1082		*den MG size*	**Corrected intensities**										
			2000 RA	*3500 RA*	*5500 RA*	*SQ 1*	*SQ 2*	*HE 2*	*3500 RS 3*	*3500 RH 3*	*DB 0.5*	*DB 1*	*DB 2*
Dist	**Bx0**	mean	62	61	49	48	58	50	78	69	54	40	66
		mdna	59	49	56	58	61	55	65	71	65	58	87
		mn10	56	54	46	42	41	64	62	72	53	48	61
		mn15	51	56	45	50	56	53	64	81	49	51	66
		mn20	46	61	53	49	54	65	64	75	49	46	65
		mn25	45	58	44	50	55	54	68	63	52	49	65
	Bx1	mean	89	106	104	76	71	91	81	123	90	77	129
		mdna	126	116	139	78	81	99	88	93	76	65	138
		mn10	72	84	85	66	89	65	85	114	81	69	126
		mn15	83	72	86	52	75	83	91	104	91	84	145
		mn20	72	82	85	86	69	77	91	96	68	74	111
		mn25	81	83	86	79	59	73	96	99	73	84	116
dM	**Bx0**	mean	228	261	279	249	234	285	237	252	222	240	309
		mdna	300	252	288	228	252	267	252	243	216	267	279
		mn10	270	264	267	279	243	231	249	255	213	225	267
		mn15	267	276	279	234	249	267	267	261	240	237	264
		mn20	273	297	288	246	255	231	258	288	252	246	273
		mn25	300	285	300	270	249	294	234	270	249	291	306
	Bx1	mean	306	291	327	270	252	273	228	282	264	252	360
		mdna	255	276	294	249	300	297	258	300	267	282	261
		mn10	267	348	300	285	243	285	261	318	279	240	342
		mn15	306	321	288	246	273	273	249	351	282	240	333
		mn20	288	264	285	279	225	216	294	327	267	252	285
		mn25	285	327	267	267	249	255	312	285	228	294	261

Table S2 - as in Table 2 for Corrected intensities.

Dset: A			Corrected intensities										
		den	*2000*	*3500*	*5500*				*3500*	*3500*			
		MG	*RA*	*RA*	*RA*	*SQ*	*SQ*	*HE*	*RS*	*RH*	*DB*	*DB*	*DB*
		size				*1*	*2*	*2*	*3*	*3*	*0.5*	*1*	*2*
Dist n eqks: 1082	**Bx0**	mean	85	88	117	123	94	113	48	66	102	132	72
		mdna	90	119	99	92	87	100	73	64	75	93	28
		mn10	96	103	126	130	131	82	84	62	108	122	89
		mn15	112	97	128	114	98	106	81	42	118	111	70
		mn20	124	86	107	121	104	79	80	55	116	125	74
		mn25	127	95	129	115	101	105	68	83	110	120	76
	Bx1	mean	25	12	14	53	63	23	43	7	24	51	4
		mdna	6	8	2	49	44	16	27	19	52	77	3
		mn10	59	35	34	71	26	78	33	10	46	65	5
		mn15	39	61	30	109	54	40	21	13	20	37	1
		mn20	60	41	32	29	67	50	22	17	69	56	11
		mn25	45	38	31	47	91	57	18	15	58	36	9
Dist n eqks: 15100 nd: 7661	**Bx0**	mean	980	-	-	-	-	-	33	-	-	-	-
		mdna	-	-	-	-	-	-	-	-	-	-	8509
		mn10	-	-	-	-	-	-	-	-	-	-	-
		mn15	-	-	-	-	-	-	-	428	-	-	-
		mn20	-	-	-	-	-	75	-	-	-	-	-
		mn25	-	-	-	-	-	-	-	-	-	-	-
	Bx1	mean	-	-	-	-	-	-	-	835	-	-	-
		mdna	235	6	142	-	-	300	-	-	-	-	-
		mn10	-	-	-	-	93	-	-	-	-	-	-
		mn15	-	-	-	-	-	-	-	-	56	-	3319
		mn20	-	-	-	-	-	-	-	-	-	-	-
		mn25	-	-	-	-	-	-	89	-	-	-	-

Table S3 - Same as Table 3 for corrected intensities.

Dset: A			Corrected intensities										
		den	*2000*	*3500*	*5500*				*3500*	*3500*			
		MG	*RA*	*RA*	*RA*	*SQ*	*SQ*	*HE*	*RS*	*RH*	*DB*	*DB*	*DB*
		size				*1*	*2*	*2*	*3*	*3*	*0.5*	*1*	*2*
dM n eqks: 1082	**Bx0**	mean	125	79	51	104	119	35	117	91	129	112	12
		mdna	20	96	30	126	93	67	97	111	130	72	46
		mn10	62	78	74	47	109	121	102	88	132	127	68
		mn15	69	53	48	118	98	64	65	82	114	116	76
		mn20	56	22	32	108	87	122	85	33	92	106	55
		mn25	17	36	19	61	103	25	120	60	99	29	14
	Bx1	mean	13	28	6	59	90	57	123	44	75	94	1
		mdna	86	52	24	101	21	23	84	16	73	45	81
		mn10	63	3	18	39	110	40	80	10	50	113	4
		mn15	15	9	31	107	58	54	105	2	43	115	5
		mn20	34	77	42	49	128	131	26	7	66	95	37
		mn25	38	8	70	71	100	89	11	41	124	27	83
dM n eqks: 5625 nd: 17136	**Bx0**	mean	2	-	2	1	2	44	3	12	-	1	69
		mdna	20	2	2	2	19	11	5	10	1	8	15
		mn10	5	1	-	10	2	1	2	9	-	2	5
		mn15	1	5	6	-	7	11	22	10	1	-	3
		mn20	-	7	-	-	26	1	10	49	10	-	1
		mn25	58	-	4	9	-	57	-	9	1	20	70
	Bx1	mean	51	-	3	1	-	-	-	-	-	-	2984
		mdna	-	-	2	-	186	86	1	25	-	1	-
		mn10	-	105	-	52	-	2	-	19	-	1	16
		mn15	13	1	-	-	3	1	-	1062	10	-	13
		mn20	-	-	-	7	-	-	2	62	2	-	-
		mn25	-	13	-	-	-	-	231	-	-	4	-

Table S4 - Same as Table 4 for corrected intensities.

N MDPs	Raw intensity			Corrected intensity		
	N eqks	Di (km)	dM (m.u.)	N eqks	Di (km)	dM (m.u.)
=1	6627	62.97	-	6587	63.81	-
1-3	2980	57.01	-	3188	53.44	-
3-5	1894	46.52	0.81	1865	44.90	0.82
5-9	1978	45.52	0.77	1922	45.97	0.76
9-19	1046	43.67	0.55	956	45.17	0.59
19-29	244	50.06	0.52	245	47.76	0.50
29-99	222	42.12	0.46	225	44.44	0.44
59-99	66	37.76	0.37	66	35.28	0.34
>99	46	30.03	0.39	46	30.48	0.34

Table S5: number of earthquakes (N eqks), average distance (Di) and average absolute difference of magnitude (dM) for increasing ranges of the number of MDPs (N MDPs).

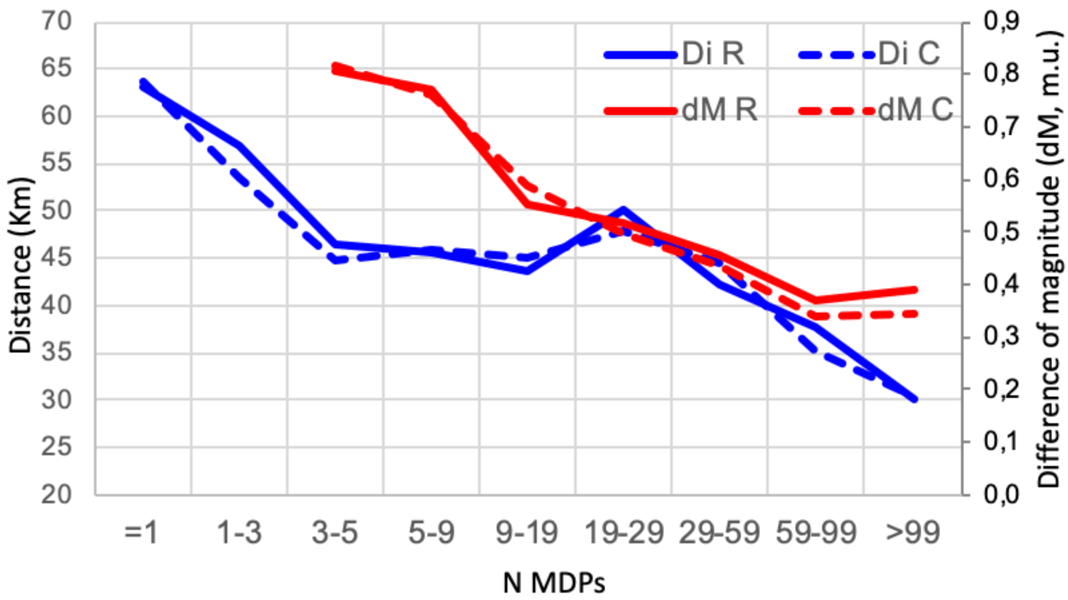


Figure S3: Plot of the values in Table S5: average distance (Di) and average absolute difference of magnitude (dM) for raw (R) and corrected (C) intensities as a function of the number of MDPs (N MDPs).

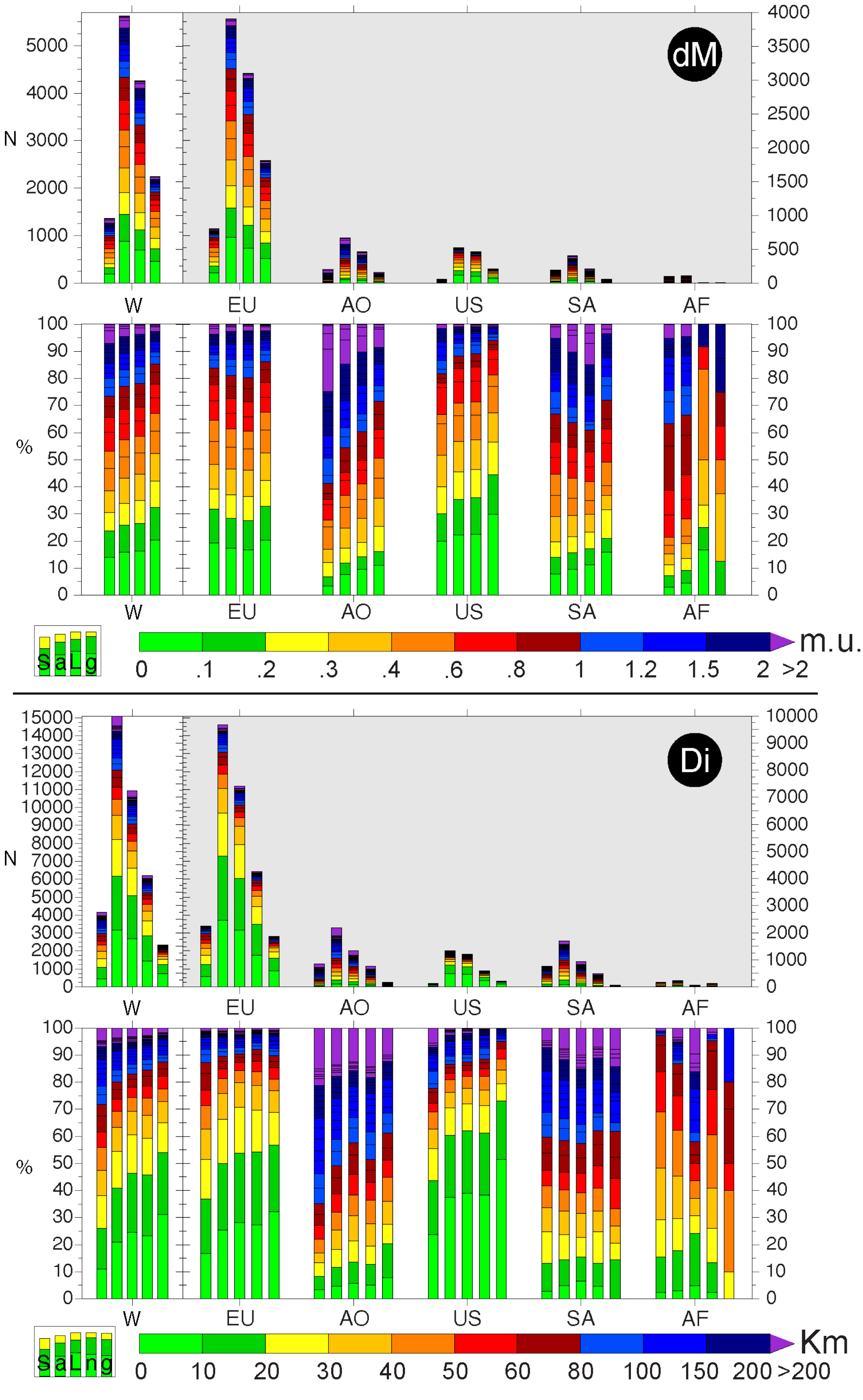


Figure S4 - as in Fig. 8 for corrected intensities.

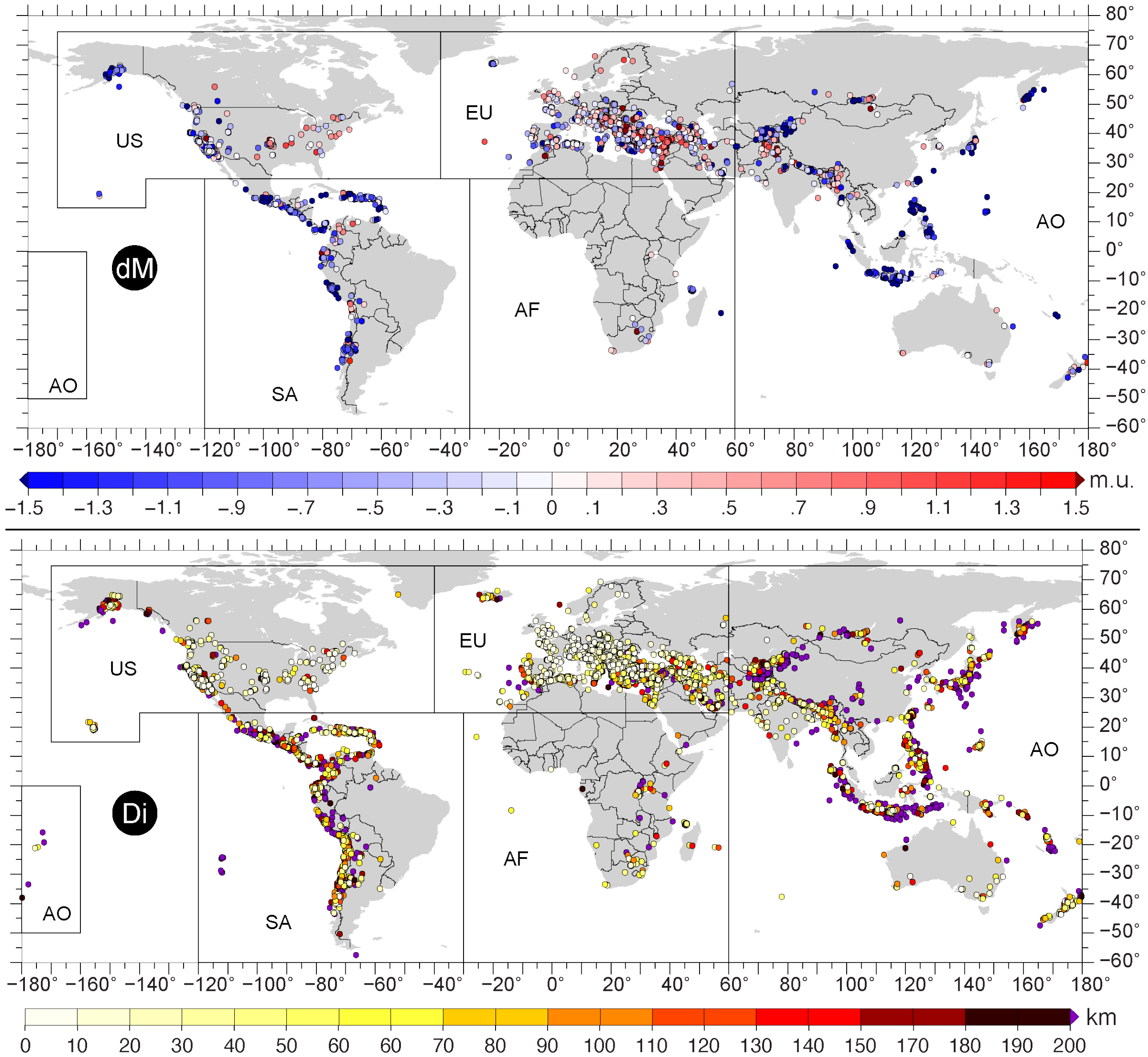


Figure S5 - as in Fig. 7 for corrected intensities.

Di Raw	Global (W)					Europe (EU)					Asia and Oceania (AO)					North America (US)					South America (SA)					Africa (AF)				
N	**S**	**a**	**L**	**n**	**g**	**S**	**a**	**L**	**n**	**g**	**S**	**a**	**L**	**n**	**g**	**S**	**a**	**L**	**n**	**g**	**S**	**a**	**L**	**n**	**g**	**S**	**a**	**L**	**n**	**g**
0-10	458	3180	2722	1527	729	365	2488	2123	1220	600	29	100	71	41	15	39	492	453	223	109	21	92	71	38	5	4	8	4	5	0
10-20	607	2904	2297	1369	537	441	2281	1840	1117	466	43	148	105	58	17	21	295	274	134	47	78	146	68	46	7	24	34	10	14	0
20-30	496	2056	1560	858	249	321	1570	1249	676	216	36	144	108	50	11	21	141	120	64	17	95	172	77	51	4	23	29	6	17	1
30-40	412	1372	960	560	200	268	959	691	396	164	33	150	117	64	14	12	85	73	40	16	64	139	75	40	5	35	39	4	20	1
40-50	349	915	566	404	121	206	530	324	233	93	36	155	119	71	11	6	60	54	29	11	62	128	66	41	3	38	41	3	30	3
50-60	237	665	428	300	106	125	357	232	171	77	47	150	103	60	17	3	27	24	16	3	39	103	64	37	6	23	28	5	16	3
60-70	249	549	300	210	69	144	277	133	110	50	31	123	92	33	7	7	28	21	16	6	52	105	53	36	5	15	16	1	15	1
70-80	179	405	226	148	49	86	183	97	73	33	37	96	59	30	8	1	19	18	8	3	48	98	50	32	5	7	9	2	5	0
80-90	132	354	222	127	52	58	149	91	60	37	44	122	78	42	8	2	14	12	8	2	27	66	39	16	5	1	3	2	1	0
90-100	150	320	170	109	25	55	138	83	52	17	50	95	45	30	6	3	15	12	8	1	41	71	30	19	1	1	1	0	0	0
100-110	110	270	160	104	31	36	101	65	48	23	44	81	37	29	2	2	16	14	10	2	28	70	42	16	3	0	2	2	1	1
110-120	72	214	142	76	37	23	88	65	35	25	23	57	34	17	5	4	19	15	8	2	22	45	23	15	4	0	5	5	1	1
120-130	92	214	122	71	22	28	88	60	30	13	38	67	29	23	6	2	16	14	8	2	23	41	18	10	1	1	2	1	0	0
130-140	76	198	122	71	28	18	78	60	27	21	37	60	23	24	5	1	14	13	6	2	20	46	26	14	0	0	0	0	0	0
140-150	59	152	93	44	16	5	45	40	17	9	29	49	20	13	5	1	13	12	6	1	24	43	19	8	1	0	2	2	0	0
150-160	55	135	80	48	16	6	34	28	11	5	34	58	24	22	8	1	9	8	3	0	14	31	17	12	3	0	3	3	0	0
160-170	44	108	64	39	9	9	32	23	13	6	17	37	20	10	1	0	4	4	3	1	18	34	16	12	1	0	1	1	1	0
170-180	39	87	48	24	9	2	14	12	6	4	21	35	14	9	2	0	8	8	1	1	16	30	14	8	2	0	0	0	0	0
180-190	38	82	44	17	4	6	18	12	5	2	20	35	15	8	2	2	6	4	1	0	10	22	12	3	0	0	1	1	0	0
190-200	26	66	40	22	6	5	21	16	6	3	13	23	10	9	2	1	7	6	2	0	7	15	8	5	1	0	0	0	0	0
>200	293	857	564	261	66	43	217	174	50	36	178	390	212	137	22	9	35	26	8	0	61	203	142	64	8	2	12	10	2	0
%	**S**	**a**	**L**	**n**	**g**	**S**	**a**	**L**	**n**	**g**	**S**	**a**	**L**	**n**	**g**	**S**	**a**	**L**	**n**	**g**	**S**	**a**	**L**	**n**	**g**	**S**	**a**	**L**	**n**	**g**
0-10	11	21	25	24	31	16	26	29	28	32	3	5	5	5	9	28	37	38	37	48	3	5	8	7	7	2	3	6	4	0
10-20	15	19	21	21	23	20	24	25	26	25	5	7	8	7	10	15	22	23	22	21	10	9	7	9	10	14	14	16	11	0
20-30	12	14	14	13	10	14	16	17	16	11	4	7	8	6	6	15	11	10	11	8	12	10	8	10	6	13	12	10	13	9
30-40	10	9	9	9	8	12	10	9	9	9	4	7	9	8	8	9	6	6	7	7	8	8	8	8	7	20	17	6	16	9
40-50	8	6	5	6	5	9	5	4	5	5	4	7	9	9	6	4	5	5	5	5	8	8	7	8	4	22	17	5	23	27
50-60	6	4	4	5	4	6	4	3	4	4	6	7	8	8	10	2	2	2	3	1	5	6	7	7	9	13	12	8	13	27
60-70	6	4	3	3	3	6	3	2	3	3	4	6	7	4	4	5	2	2	3	3	7	6	6	7	7	9	7	2	12	9
70-80	4	3	2	2	2	4	2	1	2	2	4	4	4	4	5	1	1	2	1	1	6	6	5	6	7	4	4	3	4	0
80-90	3	2	2	2	2	3	2	1	1	2	5	6	6	5	5	1	1	1	1	1	4	4	4	3	7	1	1	3	1	0
90-100	4	2	2	2	1	2	1	1	1	1	6	4	3	4	3	2	1	1	1	0	5	4	3	4	1	1	0	0	0	0
100-110	3	2	1	2	1	2	1	1	1	1	5	4	3	4	1	1	1	1	2	1	4	4	5	3	4	0	1	3	1	9
110-120	2	1	1	1	2	1	1	1	1	1	3	3	3	2	3	3	1	1	1	1	3	3	2	3	6	0	2	8	1	9
120-130	2	1	1	1	1	1	1	1	1	1	5	3	2	3	3	1	1	1	1	1	3	2	2	2	1	1	1	2	0	0
130-140	2	1	1	1	1	1	1	1	1	1	4	3	2	3	3	1	1	1	1	1	3	3	3	3	0	0	0	0	0	0
140-150	1	1	1	1	1	0	0	1	0	0	3	2	2	2	3	1	1	1	1	0	3	3	2	2	1	0	1	3	0	0
150-160	1	1	1	1	1	0	0	0	0	0	4	3	2	3	5	1	1	1	1	0	2	2	2	2	4	0	1	5	0	0
160-170	1	1	1	1	0	0	0	0	0	0	2	2	2	1	1	0	0	0	1	0	2	2	2	2	1	0	0	2	1	0
170-180	1	1	0	0	0	0	0	0	0	0	3	2	1	1	1	0	1	1	0	0	2	2	2	2	3	0	0	0	0	0
180-190	1	1	0	0	0	0	0	0	0	0	2	2	1	1	1	1	0	0	0	0	1	1	1	1	0	0	0	2	0	0
190-200	1	0	0	0	0	0	0	0	0	0	2	1	1	1	1	1	1	1	0	0	1	1	1	1	1	0	0	0	0	0
>200	7	6	5	4	3	2	2	2	1	2	21	18	16	18	13	7	3	2	1	0	8	12	15	12	11	1	5	16	2	0

Table S6 - numerical parameter of histograms in Fig. 8 for distance (Di) and raw intensities: global and zones as in Fig. 7.

dM Raw	Global (W)				Europe (EU)				Asia and Oceania (AO)				North America (US)				South America (SA)				Africa (AF)			
N	**S**	**a**	**L**	**g**	**S**	**a**	**L**	**g**	**S**	**a**	**L**	**g**	**S**	**a**	**L**	**g**	**S**	**a**	**L**	**g**	**S**	**a**	**L**	**g**
0-0.1	194	876	682	416	152	667	515	337	6	39	33	10	13	119	106	57	19	45	26	11	4	6	2	1
0.1-0.2	126	567	441	269	91	406	315	214	9	52	43	22	4	60	56	27	21	44	23	4	1	5	4	2
0.2-0.3	113	512	399	240	75	375	300	190	15	43	28	16	7	63	56	32	15	28	13	1	1	3	2	1
0.3-0.4	114	451	337	190	82	326	244	155	9	36	27	12	4	53	49	20	17	33	16	3	2	3	1	0
0.4-0.5	93	450	357	188	57	326	269	152	10	36	26	9	6	52	46	21	13	29	16	6	7	7	0	0
0.5-0.6	97	368	271	162	77	290	213	145	5	28	23	7	3	32	29	10	8	14	6	0	4	4	0	0
0.6-0.7	71	314	243	119	41	226	185	100	12	34	22	6	3	27	24	9	13	24	11	4	2	3	1	0
0.7-0.8	81	341	260	124	50	240	190	104	8	40	32	12	1	24	23	4	10	24	14	3	12	13	1	1
0.8-0.9	62	265	203	98	37	176	139	67	4	32	28	15	3	28	25	8	4	15	11	8	14	14	0	0
0.9-1	45	217	172	77	19	143	124	57	9	32	23	10	2	21	19	8	8	13	5	1	7	8	1	1
1-1.1	44	182	138	55	22	136	114	45	4	15	11	4	0	11	11	4	9	11	2	2	9	9	0	0
1.1-1.2	51	166	115	46	20	100	80	35	6	24	18	5	1	9	8	0	12	21	9	6	12	12	0	0
1.2-1.3	36	140	104	31	10	81	71	23	15	32	17	4	3	10	7	3	4	13	9	1	4	4	0	0
1.3-1.4	28	105	77	32	12	66	54	21	5	18	13	3	1	9	8	4	7	9	2	3	3	3	0	1
1.4-1.5	31	106	75	19	13	66	53	19	7	21	14	0	0	3	3	0	6	11	5	0	5	5	0	0
1.5-1.6	29	101	72	25	12	62	50	24	7	22	15	1	1	3	2	0	5	10	5	0	4	4	0	0
1.6-1.7	21	82	61	16	5	49	44	14	6	17	11	1	0	2	2	0	6	10	4	1	4	4	0	0
1.7-1.8	23	92	69	26	11	57	46	23	6	24	18	2	1	2	1	0	3	7	4	0	2	2	0	1
1.8-1.9	16	60	44	16	5	41	36	12	7	11	4	2	0	3	3	1	4	5	1	1	0	0	0	0
1.9-2	13	56	43	11	5	25	20	4	7	20	13	4	0	3	3	1	1	8	7	2	0	0	0	0
2-2.5	59	164	105	41	15	70	55	29	32	61	29	9	1	3	2	2	8	27	19	1	3	3	0	0
2.5-3	18	55	37	13	8	19	11	9	8	26	18	4	0	1	1	0	2	9	7	0	0	0	0	0
> 3	14	33	19	10	3	11	8	7	8	13	5	2	0	0	0	0	3	9	6	1	0	0	0	0
%	**S**	**a**	**L**	**g**	**S**	**a**	**L**	**g**	**S**	**a**	**L**	**g**	**S**	**a**	**L**	**g**	**S**	**a**	**L**	**g**	**S**	**a**	**L**	**g**
0-0.1	14	15	16	19	18	17	16	19	3	6	7	6	24	22	22	27	10	11	12	19	4	5	17	13
0.1-0.2	9	10	10	12	11	10	10	12	4	8	9	14	7	11	12	13	11	11	10	7	1	4	33	25
0.2-0.3	8	9	9	11	9	9	10	11	7	6	6	10	13	12	12	15	8	7	6	2	1	3	17	13
0.3-0.4	8	8	8	9	10	8	8	9	4	5	6	8	7	10	10	9	9	8	7	5	2	3	8	0
0.4-0.5	7	8	8	8	7	8	9	9	5	5	6	6	11	10	10	10	7	7	7	10	7	6	0	0
0.5-0.6	7	6	6	7	9	7	7	8	2	4	5	4	6	6	6	5	4	3	3	0	4	4	0	0
0.6-0.7	5	6	6	5	5	6	6	6	6	5	5	4	6	5	5	4	7	6	5	7	2	3	8	0
0.7-0.8	6	6	6	6	6	6	6	6	4	6	7	8	2	4	5	2	5	6	6	5	12	12	8	13
0.8-0.9	5	5	5	4	5	4	4	4	2	5	6	9	6	5	5	4	2	4	5	14	14	13	0	0
0.9-1	3	4	4	3	2	4	4	3	4	5	5	6	4	4	4	4	4	3	2	2	7	7	8	13
1-1.1	3	3	3	2	3	3	4	3	2	2	2	3	0	2	2	2	5	3	1	3	9	8	0	0
1.1-1.2	4	3	3	2	2	3	3	2	3	4	4	3	2	2	2	0	6	5	4	10	12	11	0	0
1.2-1.3	3	2	2	1	1	2	2	1	7	5	4	3	6	2	1	1	2	3	4	2	4	4	0	0
1.3-1.4	2	2	2	1	1	2	2	1	2	3	3	2	2	2	2	2	4	2	1	5	3	3	0	13
1.4-1.5	2	2	2	1	2	2	2	1	3	3	3	0	0	1	1	0	3	3	2	0	5	4	0	0
1.5-1.6	2	2	2	1	1	2	2	1	3	3	3	1	2	1	0	0	3	2	2	0	4	4	0	0
1.6-1.7	2	1	1	1	1	1	1	1	3	3	2	1	0	0	0	0	3	2	2	2	4	4	0	0
1.7-1.8	2	2	2	1	1	1	1	1	3	4	4	1	2	0	0	0	2	2	2	0	2	2	0	13
1.8-1.9	1	1	1	1	1	1	1	1	3	2	1	1	0	1	1	0	2	1	0	2	0	0	0	0
1.9-2	1	1	1	0	1	1	1	0	3	3	3	3	0	1	1	0	1	2	3	3	0	0	0	0
2-2.5	4	3	2	2	2	2	2	2	16	9	6	6	2	1	0	1	4	6	9	2	3	3	0	0
2.5-3	1	1	1	1	1	0	0	1	4	4	4	3	0	0	0	0	1	2	3	0	0	0	0	0
> 3	1	1	0	0	0	0	0	0	4	2	1	1	0	0	0	0	2	2	3	2	0	0	0	0

Table S7 - numerical parameter of histograms in Fig. 8 for difference of magnitude (dM) and raw intensities: global and zones as in Fig. 7.

Di Corr.	Global (W)					Europe (EU)					Asia and Oceania (AO)					North America (US)					South America (SA)					Africa (AF)				
N	**S**	**a**	**L**	**n**	**g**	**S**	**a**	**L**	**n**	**g**	**S**	**a**	**L**	**n**	**g**	**S**	**a**	**L**	**n**	**g**	**S**	**a**	**L**	**n**	**g**	**S**	**a**	**L**	**n**	**g**
0-10	458	3152	2694	1453	726	374	2462	2088	1164	600	28	103	75	39	14	30	497	467	224	112	22	83	61	23	0	4	7	3	3	0
10-20	623	3003	2380	1383	535	454	2359	1905	1134	457	42	148	106	59	22	25	300	275	135	47	79	161	82	41	9	23	35	12	14	0
20-30	501	2051	1550	844	258	331	1584	1253	662	226	41	146	105	51	13	15	135	120	59	14	90	158	68	56	4	24	28	4	16	1
30-40	390	1345	955	534	180	250	927	677	381	150	30	154	124	63	15	9	77	68	34	11	68	150	82	37	4	33	37	4	19	0
40-50	348	896	548	384	112	197	516	319	221	80	44	146	102	67	16	8	61	53	29	9	63	133	70	42	4	36	40	4	25	3
50-60	243	665	422	287	106	128	349	221	159	79	42	145	103	50	11	4	30	26	17	8	43	111	68	40	7	26	30	4	21	1
60-70	245	560	315	215	66	139	282	143	111	45	34	128	94	35	8	5	28	23	13	6	50	103	53	38	5	17	19	2	18	2
70-80	176	398	222	143	56	89	184	95	76	39	34	96	62	30	10	2	16	14	8	0	45	93	48	24	6	6	9	3	5	1
80-90	150	367	217	139	38	65	152	87	60	30	47	118	71	43	6	2	19	17	10	0	35	75	40	25	2	1	3	2	1	0
90-100	133	310	177	97	26	45	127	82	42	17	45	97	52	35	7	6	23	17	13	2	36	62	26	7	0	1	1	0	0	0
100-110	104	261	157	91	39	32	99	67	45	28	40	77	37	25	4	1	11	10	4	3	31	71	40	15	2	0	3	3	2	2
110-120	81	227	146	89	32	25	85	60	35	21	31	73	42	29	7	3	18	15	6	1	22	47	25	19	3	0	4	4	0	0
120-130	85	196	111	56	13	24	77	53	21	10	36	63	27	21	2	2	13	11	7	1	22	41	19	7	0	1	2	1	0	0
130-140	71	193	122	55	23	16	75	59	22	17	34	57	23	18	4	1	14	13	3	1	20	47	27	12	1	0	0	0	0	0
140-150	59	152	93	48	18	5	47	42	18	11	31	54	23	16	5	1	14	13	7	1	22	35	13	7	1	0	2	2	0	0
150-160	59	126	67	35	12	12	40	28	11	7	31	52	21	16	3	1	3	2	0	0	15	28	13	8	2	0	3	3	0	0
160-170	50	108	58	32	12	10	30	20	11	6	21	42	21	7	2	0	4	4	3	1	19	32	13	11	3	0	0	0	0	0
170-180	38	90	52	30	9	5	18	13	8	3	17	30	13	9	4	1	11	10	4	1	15	31	16	9	1	0	0	0	0	0
180-190	32	79	47	20	4	3	13	10	5	3	19	34	15	8	1	0	6	6	2	0	10	25	15	5	0	0	1	1	0	0
190-200	27	69	42	24	7	6	23	17	10	5	13	22	9	6	2	1	9	8	2	0	7	15	8	6	0	0	0	0	0	0
>200	285	852	567	255	63	39	218	179	52	32	178	390	212	140	22	9	34	25	6	0	57	198	141	54	9	2	12	10	3	0
%	**S**	**a**	**L**	**n**	**g**	**S**	**a**	**L**	**n**	**g**	**S**	**a**	**L**	**n**	**g**	**S**	**a**	**L**	**n**	**g**	**S**	**a**	**L**	**n**	**g**	**S**	**a**	**L**	**n**	**g**
0-10	11	21	25	23	31	17	25	28	27	32	3	5	6	5	8	24	38	39	38	51	3	5	7	5	0	2	3	5	2	0
10-20	15	20	22	22	23	20	24	26	27	24	5	7	8	8	12	20	23	23	23	22	10	9	9	8	14	13	15	19	11	0
20-30	12	14	14	14	11	15	16	17	16	12	5	7	8	7	7	12	10	10	10	6	12	9	7	12	6	14	12	6	13	10
30-40	9	9	9	9	8	11	10	9	9	8	4	7	9	8	8	7	6	6	6	5	9	9	9	8	6	19	16	6	15	0
40-50	8	6	5	6	5	9	5	4	5	4	5	7	8	9	9	6	5	4	5	4	8	8	8	9	6	21	17	6	20	30
50-60	6	4	4	5	5	6	4	3	4	4	5	7	8	7	6	3	2	2	3	4	6	7	7	8	11	15	13	6	17	10
60-70	6	4	3	3	3	6	3	2	3	2	4	6	7	5	4	4	2	2	2	3	6	6	6	8	8	10	8	3	14	20
70-80	4	3	2	2	2	4	2	1	2	2	4	4	5	4	6	2	1	1	1	0	6	5	5	5	10	3	4	5	4	10
80-90	4	2	2	2	2	3	2	1	1	2	6	5	5	6	3	2	1	1	2	0	5	4	4	5	3	1	1	3	1	0
90-100	3	2	2	2	1	2	1	1	1	1	5	4	4	5	4	5	2	1	2	1	5	4	3	1	0	1	0	0	0	0
100-110	3	2	1	1	2	1	1	1	1	2	5	4	3	3	2	1	1	1	1	1	4	4	4	3	3	0	1	5	2	20
110-120	2	2	1	1	1	1	1	1	1	1	4	3	3	4	4	2	1	1	1	0	3	3	3	4	5	0	2	6	0	0
120-130	2	1	1	1	1	1	1	1	0	1	4	3	2	3	1	2	1	1	1	0	3	2	2	1	0	1	1	2	0	0
130-140	2	1	1	1	1	1	1	1	1	1	4	3	2	2	2	1	1	1	1	0	3	3	3	2	2	0	0	0	0	0
140-150	1	1	1	1	1	0	0	1	0	1	4	2	2	2	3	1	1	1	1	0	3	2	1	1	2	0	1	3	0	0
150-160	1	1	1	1	1	1	0	0	0	0	4	2	2	2	2	1	0	0	0	0	2	2	1	2	3	0	1	5	0	0
160-170	1	1	1	1	1	0	0	0	0	0	3	2	2	1	1	0	0	0	1	0	2	2	1	2	5	0	0	0	0	0
170-180	1	1	0	0	0	0	0	0	0	0	2	1	1	1	2	1	1	1	1	0	2	2	2	2	2	0	0	0	0	0
180-190	1	1	0	0	0	0	0	0	0	0	2	2	1	1	1	0	0	1	0	0	1	1	2	1	0	0	0	2	0	0
190-200	1	0	0	0	0	0	0	0	0	0	2	1	1	1	1	1	1	1	0	0	1	1	1	1	0	0	0	0	0	0
>200	7	6	5	4	3	2	2	2	1	2	21	18	16	18	12	7	3	2	1	0	7	12	15	11	14	1	5	16	2	0

Table S8 - as in Table S6 for corrected intensities and histograms in Fig. S4.

dM Corr.	Global (W)				Europe (EU)				Asia and Oceania (AO)				North America (US)				South America (SA)				Africa (AF)			
N	**S**	**a**	**L**	**g**	**S**	**a**	**L**	**g**	**S**	**a**	**L**	**g**	**S**	**a**	**L**	**g**	**S**	**a**	**L**	**g**	**S**	**a**	**L**	**g**
0-0.1	192	887	695	459	155	675	520	368	7	51	44	18	12	117	105	64	15	39	24	9	3	5	2	0
0.1-0.2	131	563	432	269	102	434	332	226	7	29	22	8	6	70	64	31	12	25	13	3	4	5	1	1
0.2-0.3	92	453	361	219	60	334	274	172	11	36	25	15	6	54	48	26	11	24	13	6	4	5	1	0
0.3-0.4	112	524	412	230	73	376	303	185	10	51	41	17	7	59	52	23	18	32	14	3	4	6	2	2
0.4-0.5	111	444	333	185	69	315	246	150	17	50	33	12	7	50	43	21	15	26	11	2	3	3	0	0
0.5-0.6	88	356	268	149	63	263	200	122	5	31	26	12	2	25	23	9	15	30	15	5	3	7	4	1
0.6-0.7	95	333	238	128	57	224	167	100	12	36	24	12	6	41	35	12	12	24	12	4	8	8	0	0
0.7-0.8	74	295	221	115	49	215	166	98	4	21	17	5	1	26	25	8	11	23	12	3	9	10	1	1
0.8-0.9	55	249	194	98	24	176	152	82	4	28	24	9	1	11	10	4	10	18	8	2	16	16	0	1
0.9-1	52	225	173	69	25	151	126	54	8	33	25	8	1	14	13	3	10	19	9	4	8	8	0	0
1-1.1	43	197	154	58	24	139	115	46	6	26	20	6	1	16	15	3	4	8	4	3	8	8	0	0
1.1-1.2	46	151	105	40	15	100	85	33	13	24	11	3	2	8	6	2	12	15	3	2	4	4	0	0
1.2-1.3	27	116	89	27	8	66	58	20	3	20	17	4	1	7	6	1	9	17	8	2	6	6	0	0
1.3-1.4	25	121	96	37	12	81	69	28	5	23	18	6	2	3	1	1	3	11	8	2	3	3	0	0
1.4-1.5	25	104	79	20	10	63	53	15	9	23	14	3	1	8	7	2	2	7	5	0	3	3	0	0
1.5-1.6	26	90	64	17	5	52	47	14	8	18	10	2	1	5	4	0	9	11	2	0	3	4	1	1
1.6-1.7	28	78	50	20	11	44	33	14	11	21	10	4	0	1	1	0	5	11	6	1	1	1	0	1
1.7-1.8	17	78	61	21	10	49	39	18	3	18	15	2	1	3	2	0	3	8	5	1	0	0	0	0
1.8-1.9	18	62	44	18	3	30	27	14	9	18	9	1	1	5	4	2	2	6	4	1	3	3	0	0
1.9-2	12	46	34	10	4	17	13	7	3	16	13	1	0	1	1	0	5	12	7	2	0	0	0	0
2-2.5	62	162	100	40	16	68	52	27	32	60	28	11	1	4	3	1	8	25	17	1	5	5	0	0
2.5-3	19	53	34	9	7	17	10	7	12	27	15	1	0	1	1	1	0	8	8	0	0	0	0	0
> 3	15	38	23	12	6	17	11	9	7	12	5	2	0	0	0	0	2	9	7	1	0	0	0	0
%	**S**	**a**	**L**	**g**	**S**	**a**	**L**	**g**	**S**	**a**	**L**	**g**	**S**	**a**	**L**	**g**	**S**	**a**	**L**	**g**	**S**	**a**	**L**	**g**
0-0.1	14	16	16	20	19	17	17	20	3	8	9	11	20	22	22	30	8	10	11	16	3	5	17	0
0.1-0.2	10	10	10	12	13	11	11	12	3	4	5	5	10	13	14	14	6	6	6	5	4	5	8	13
0.2-0.3	7	8	8	10	7	9	9	10	5	5	5	9	10	10	10	12	6	6	6	11	4	5	8	0
0.3-0.4	8	9	10	10	9	10	10	10	5	8	9	10	12	11	11	11	9	8	7	5	4	5	17	25
0.4-0.5	8	8	8	8	9	8	8	8	8	7	7	7	12	9	9	10	8	6	5	4	3	3	0	0
0.5-0.6	6	6	6	7	8	7	6	7	2	5	6	7	3	5	5	4	8	7	7	9	3	6	33	13
0.6-0.7	7	6	6	6	7	6	5	6	6	5	5	7	10	8	7	6	6	6	6	7	8	7	0	0
0.7-0.8	5	5	5	5	6	6	5	5	2	3	4	3	2	5	5	4	6	6	6	5	9	9	8	13
0.8-0.9	4	4	5	4	3	5	5	5	2	4	5	6	2	2	2	2	5	4	4	4	16	15	0	13
0.9-1	4	4	4	3	3	4	4	3	4	5	5	5	2	3	3	1	5	5	4	7	8	7	0	0
1-1.1	3	4	4	3	3	4	4	3	3	4	4	4	2	3	3	1	2	2	2	5	8	7	0	0
1.1-1.2	3	3	2	2	2	3	3	2	6	4	2	2	3	2	1	1	6	4	1	4	4	4	0	0
1.2-1.3	2	2	2	1	1	2	2	1	1	3	4	2	2	1	1	0	5	4	4	4	6	5	0	0
1.3-1.4	2	2	2	2	1	2	2	2	2	3	4	4	3	1	0	0	2	3	4	4	3	3	0	0
1.4-1.5	2	2	2	1	1	2	2	1	4	3	3	2	2	2	1	1	1	2	2	0	3	3	0	0
1.5-1.6	2	2	2	1	1	1	2	1	4	3	2	1	2	1	1	0	5	3	1	0	3	4	8	13
1.6-1.7	2	1	1	1	1	1	1	1	5	3	2	2	0	0	0	0	3	3	3	2	1	1	0	13
1.7-1.8	1	1	1	1	1	1	1	1	1	3	3	1	2	1	0	0	2	2	2	2	0	0	0	0
1.8-1.9	1	1	1	1	0	1	1	1	4	3	2	1	2	1	1	1	1	1	2	2	3	3	0	0
1.9-2	1	1	1	0	1	0	0	0	1	2	3	1	0	0	0	0	3	3	3	4	0	0	0	0
2-2.5	5	3	2	2	2	2	2	1	16	9	6	7	2	1	1	0	4	6	8	2	5	5	0	0
2.5-3	1	1	1	0	1	0	0	0	6	4	3	1	0	0	0	0	0	2	4	0	0	0	0	0
> 3	1	1	1	1	1	0	0	1	3	2	1	1	0	0	0	0	1	2	3	2	0	0	0	0

Table S9 - as in Table S7 for corrected intensities and histograms in Fig. S4.

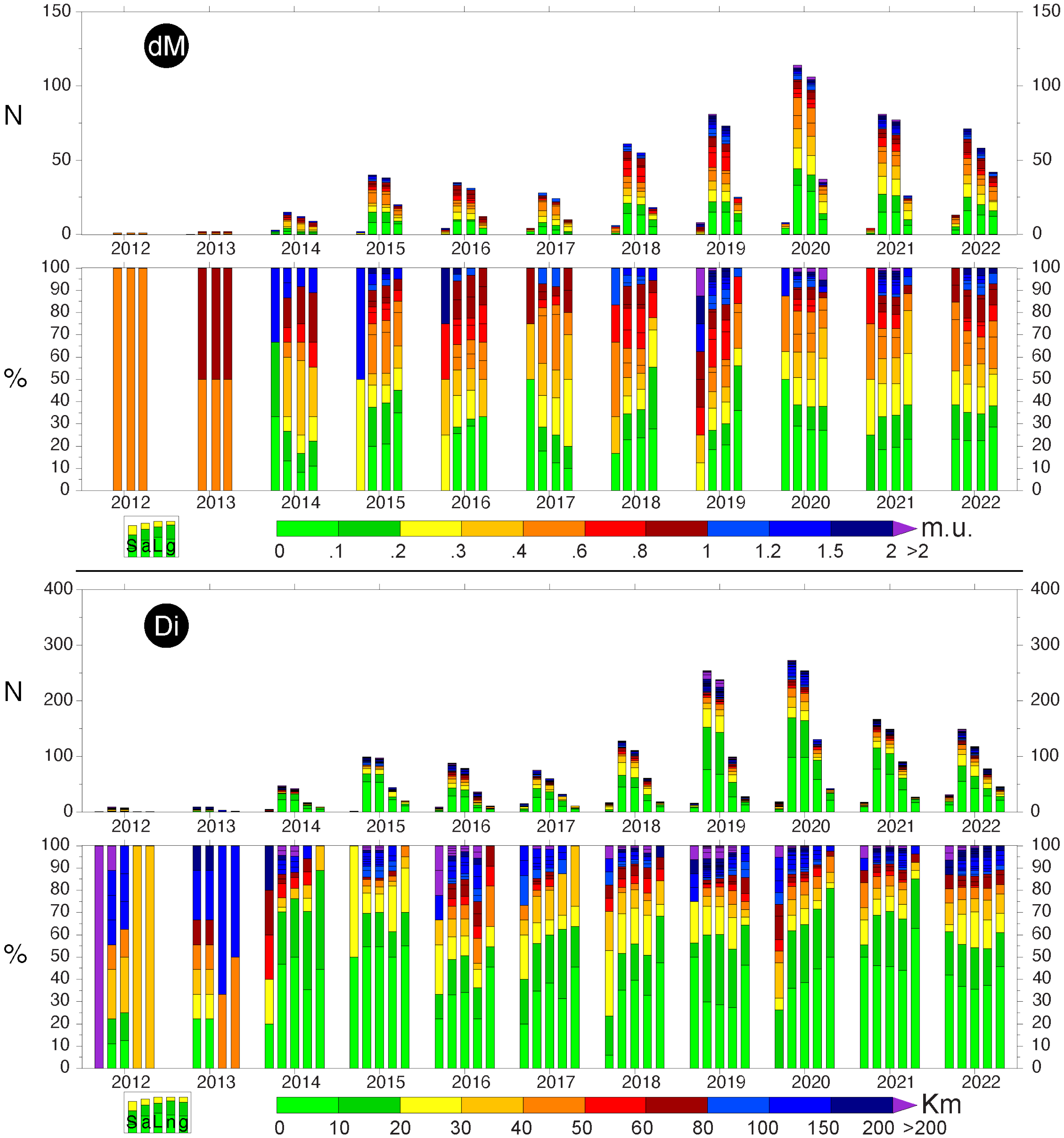


Figure S6 - as in Fig. 9 for US macro-area.

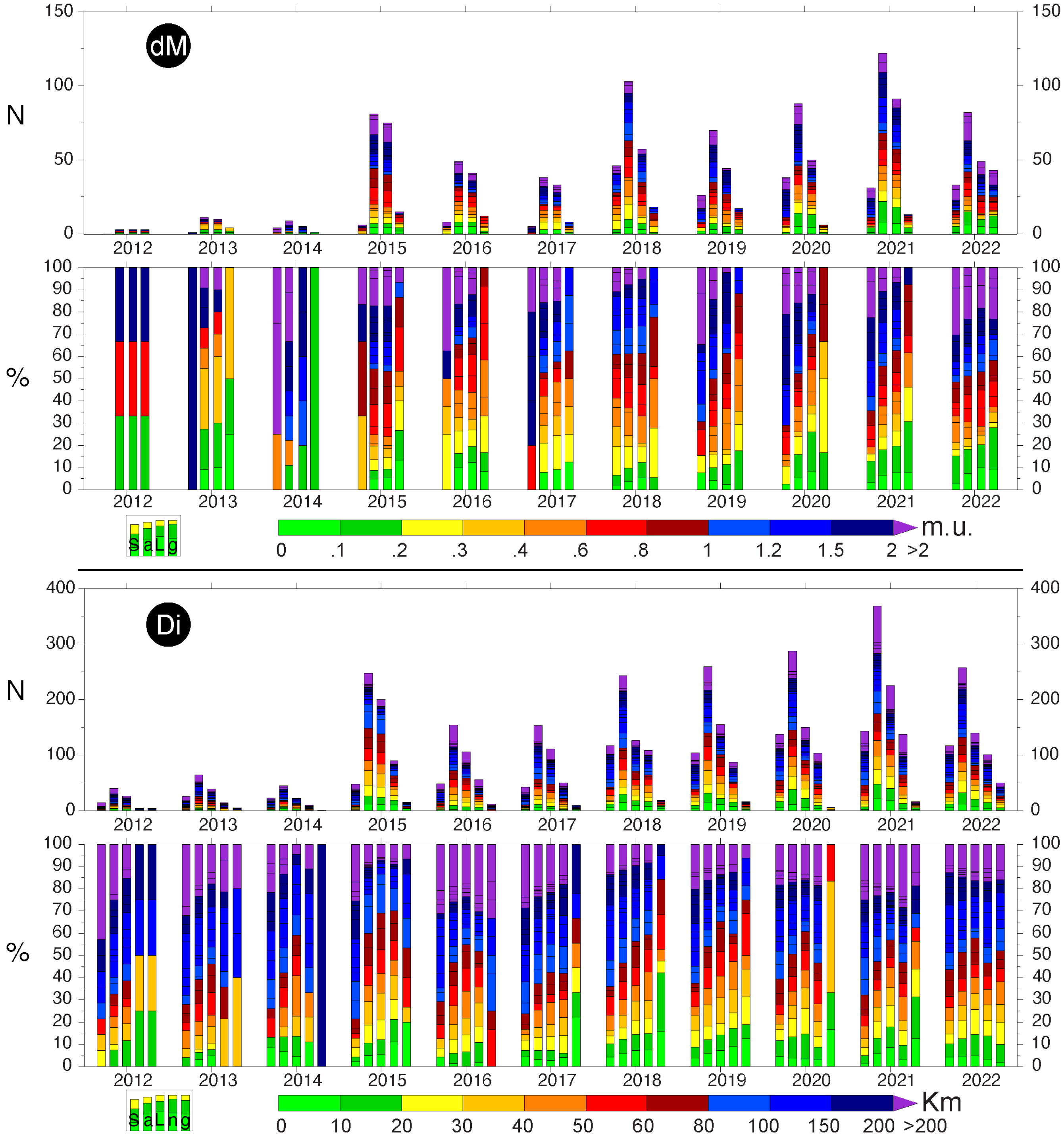


Figure S7 - as in Fig. 9 for AO macro-area.

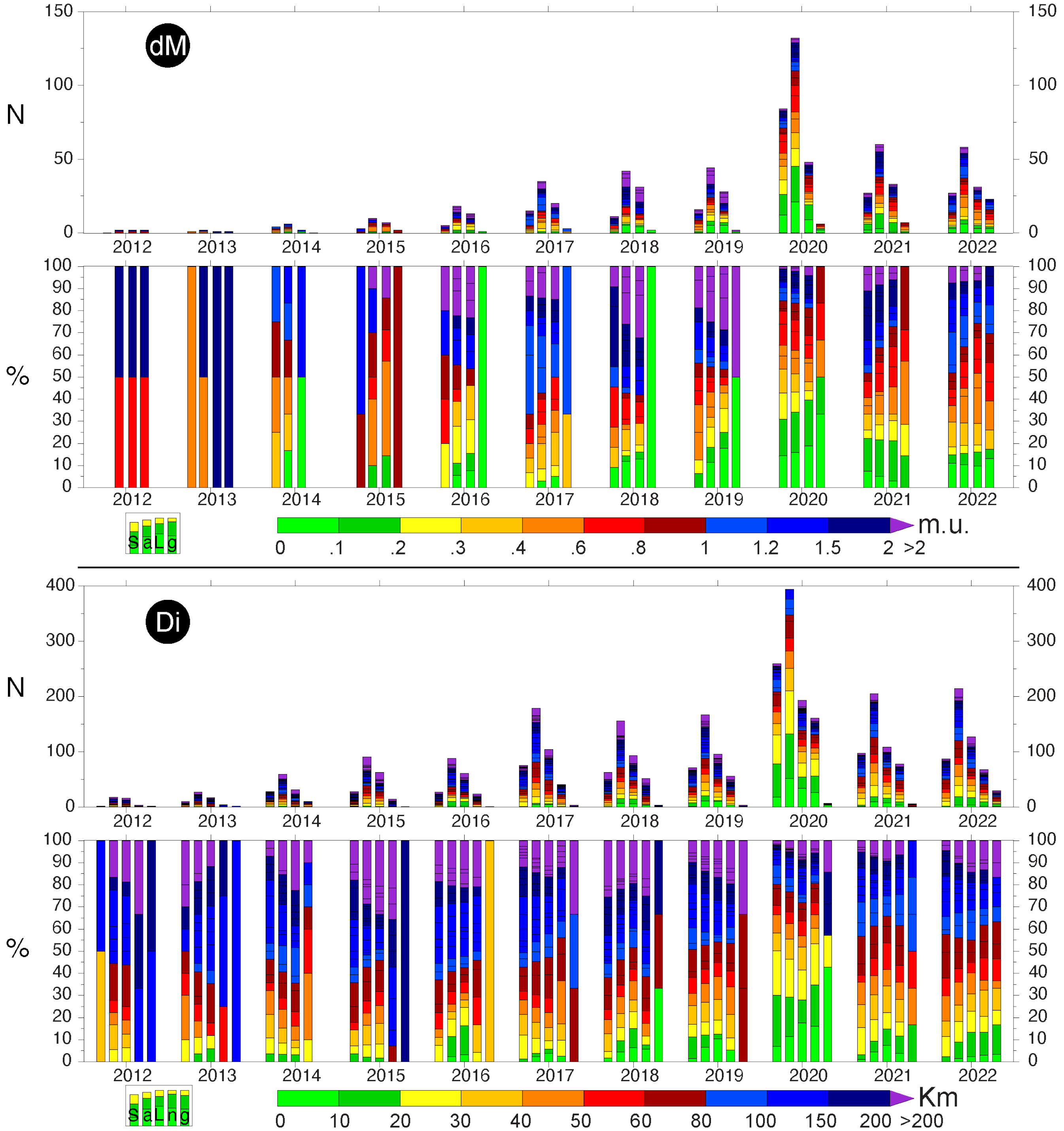


Figure S8 - as in Fig. 9 for SA macro-area.

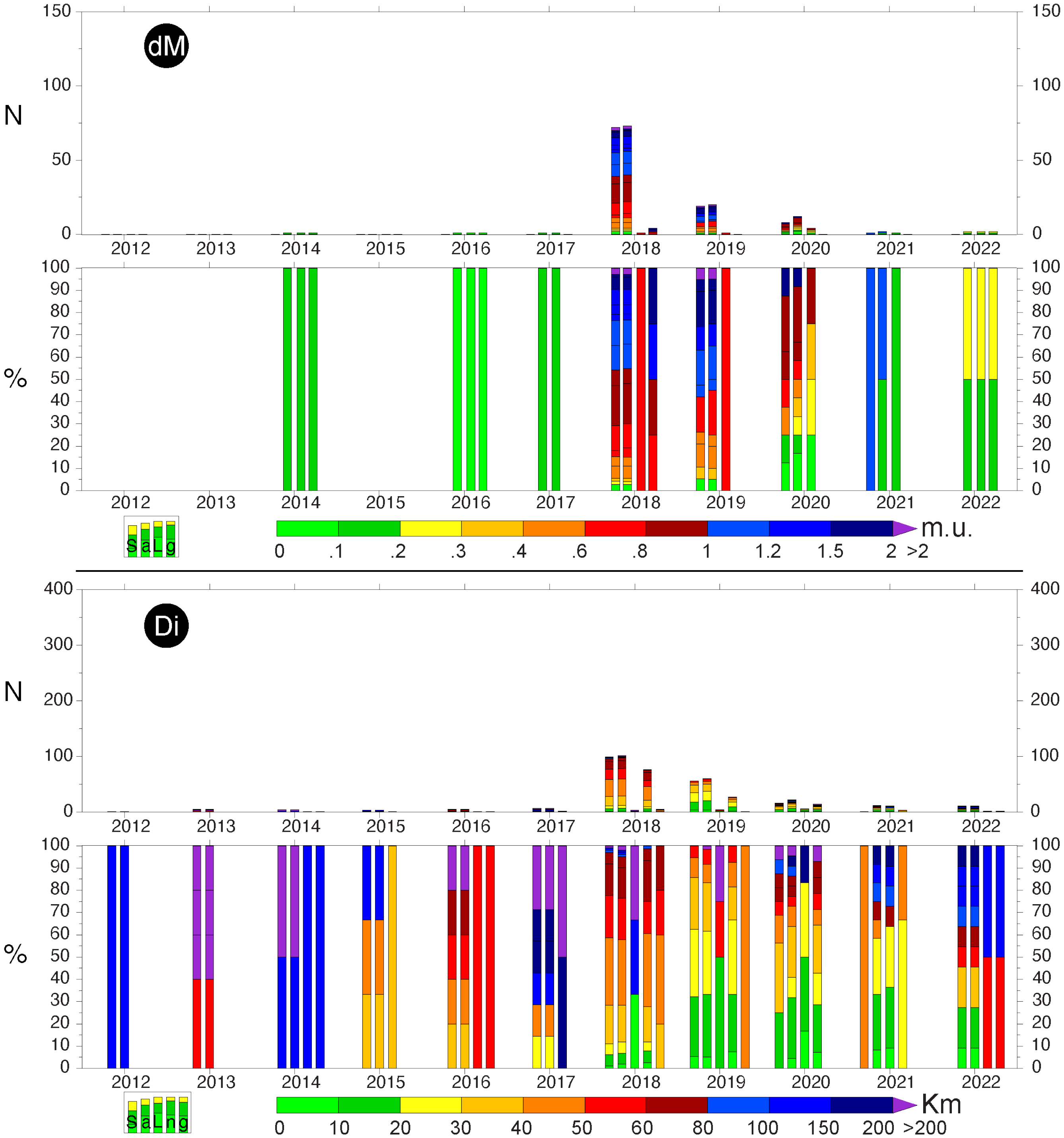


Figure S9 - as in Fig. 9 for AF macro-area.

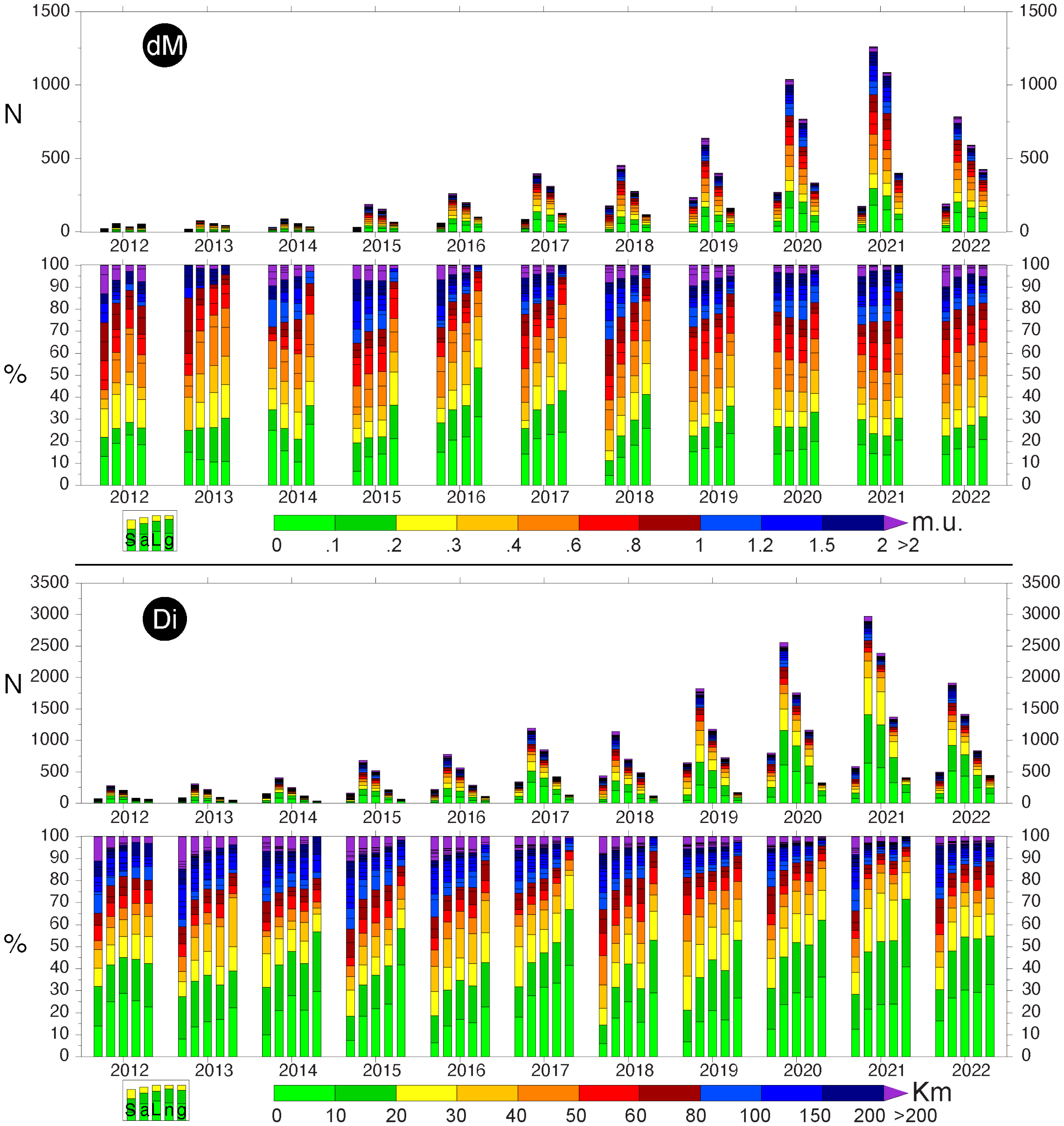


Figure S10 - as in Fig. 9 for corrected intensities.

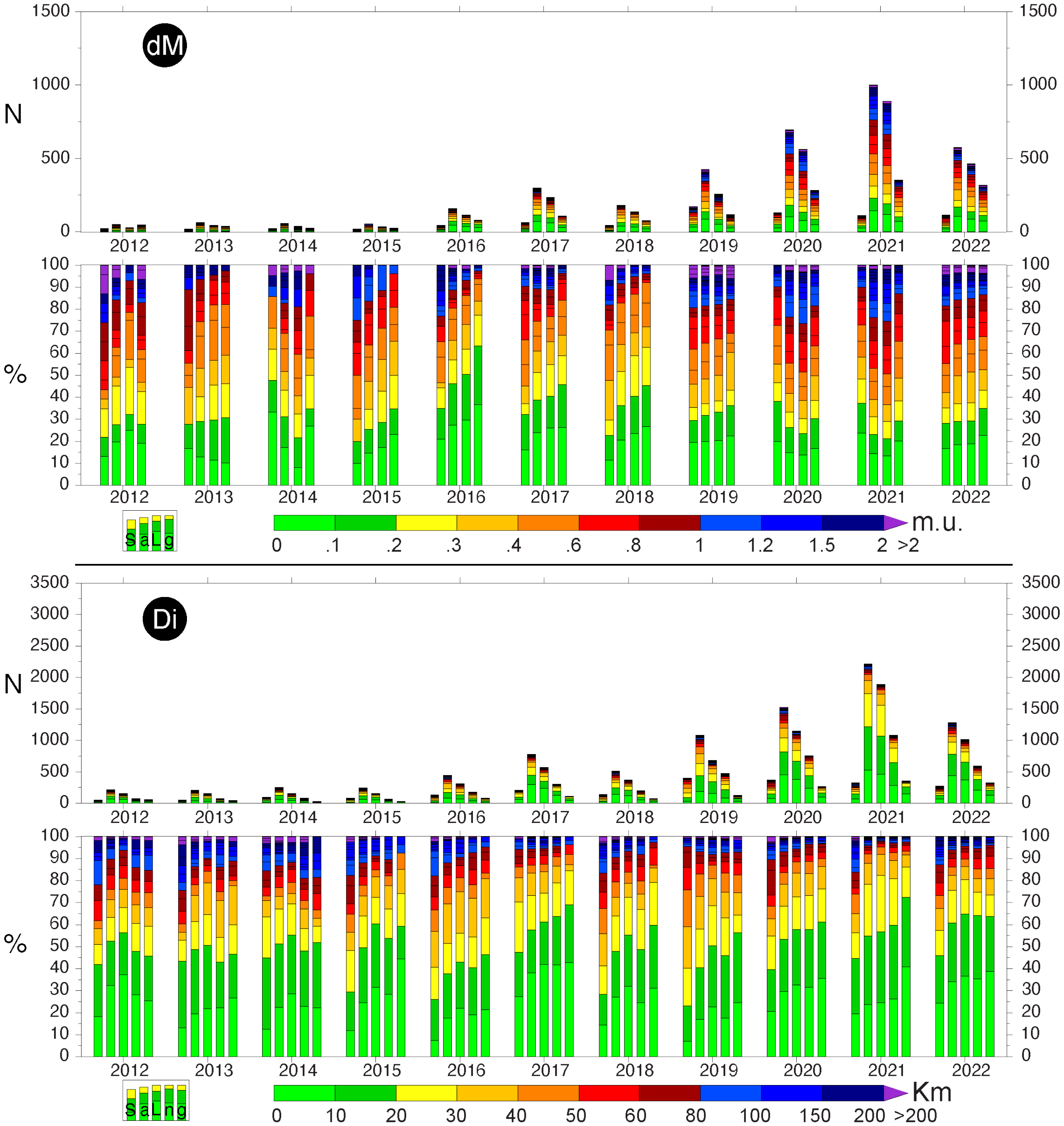


Figure S11 - as in Fig. 10 for corrected intensities.

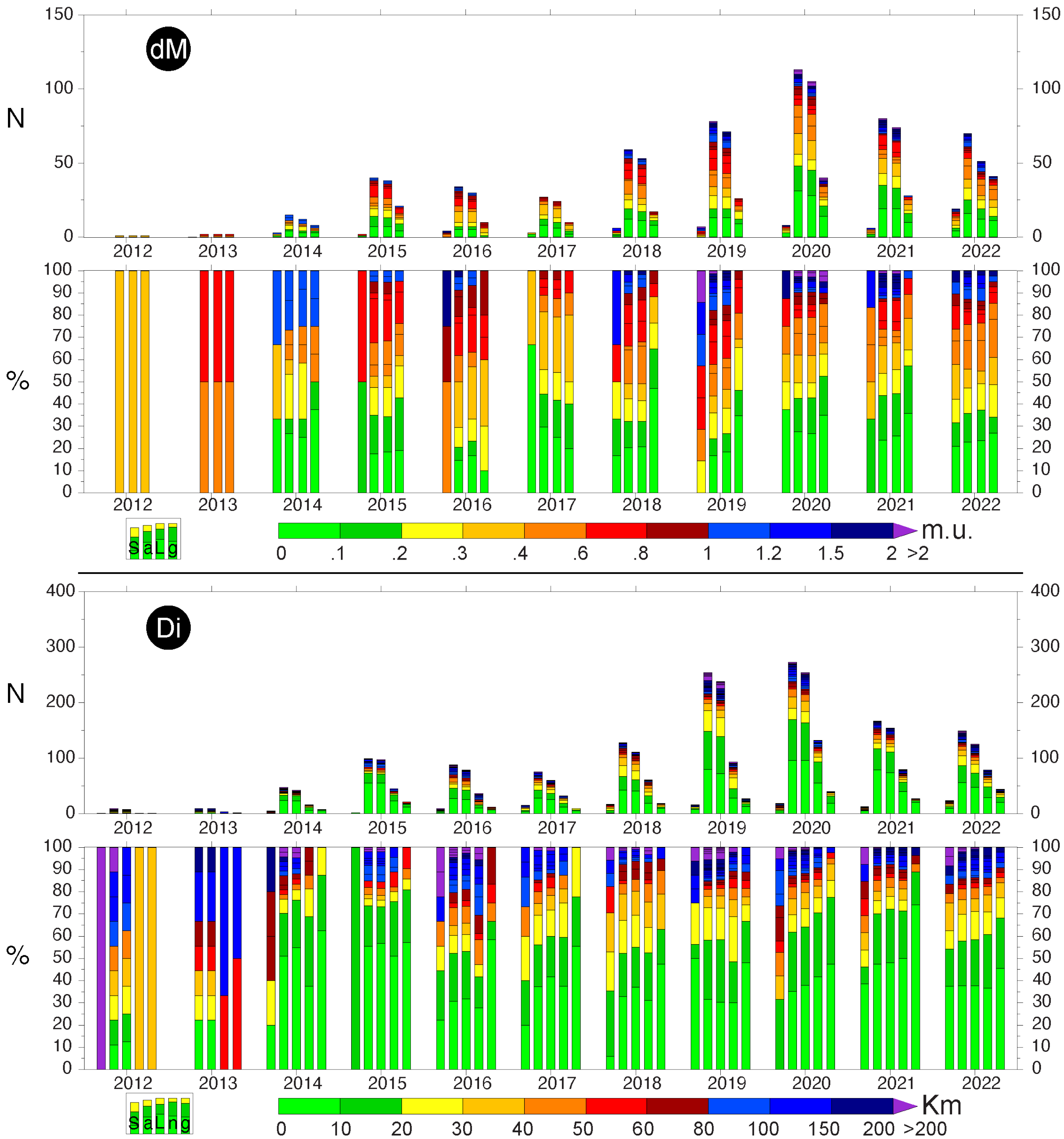


Figure S12 - as in Fig. S6 for corrected intensity.

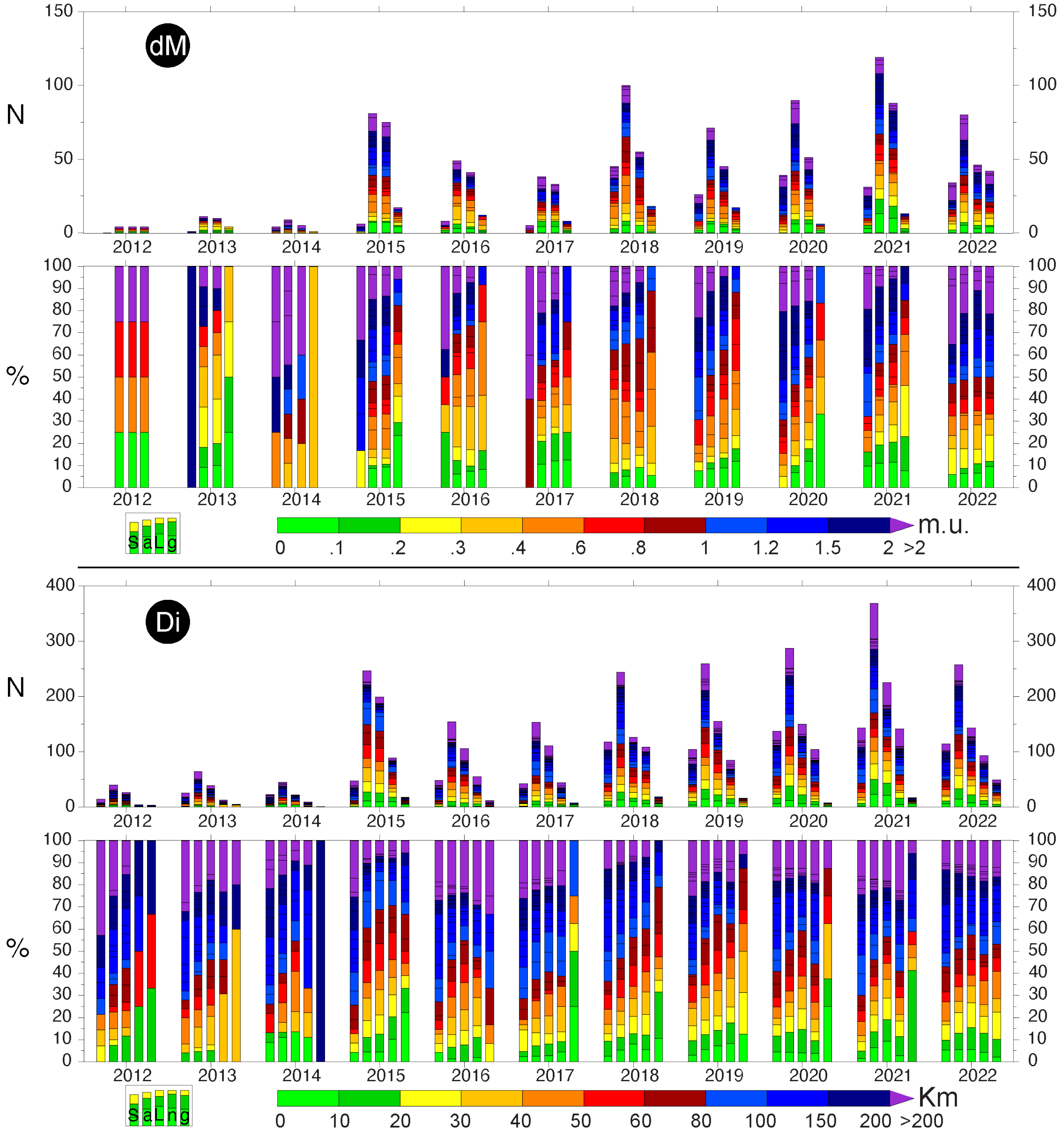


Figure S13 - as in Fig. S7 for corrected intensities.

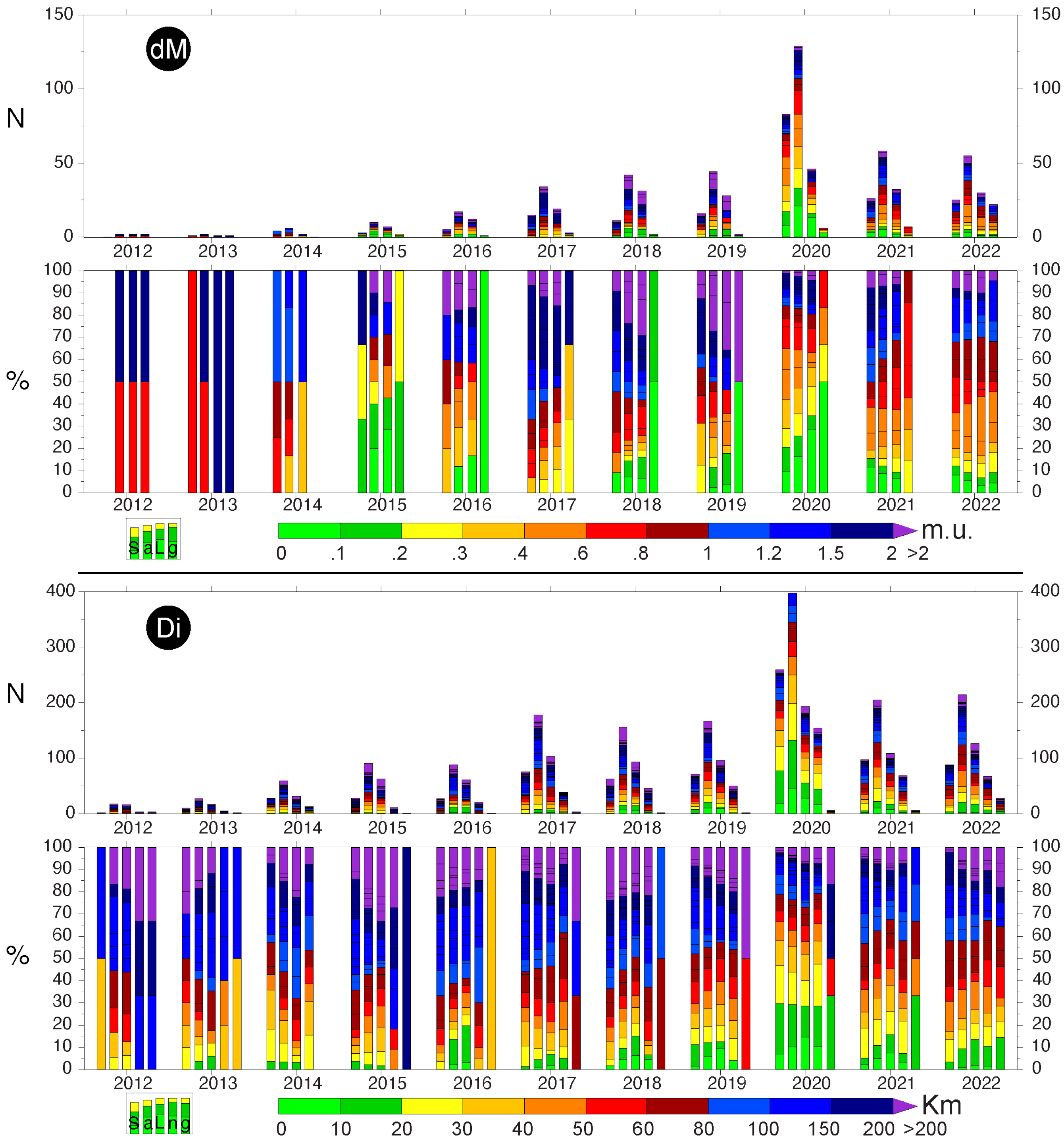


Figure S14 - as in Fig. S8 for corrected intensity.

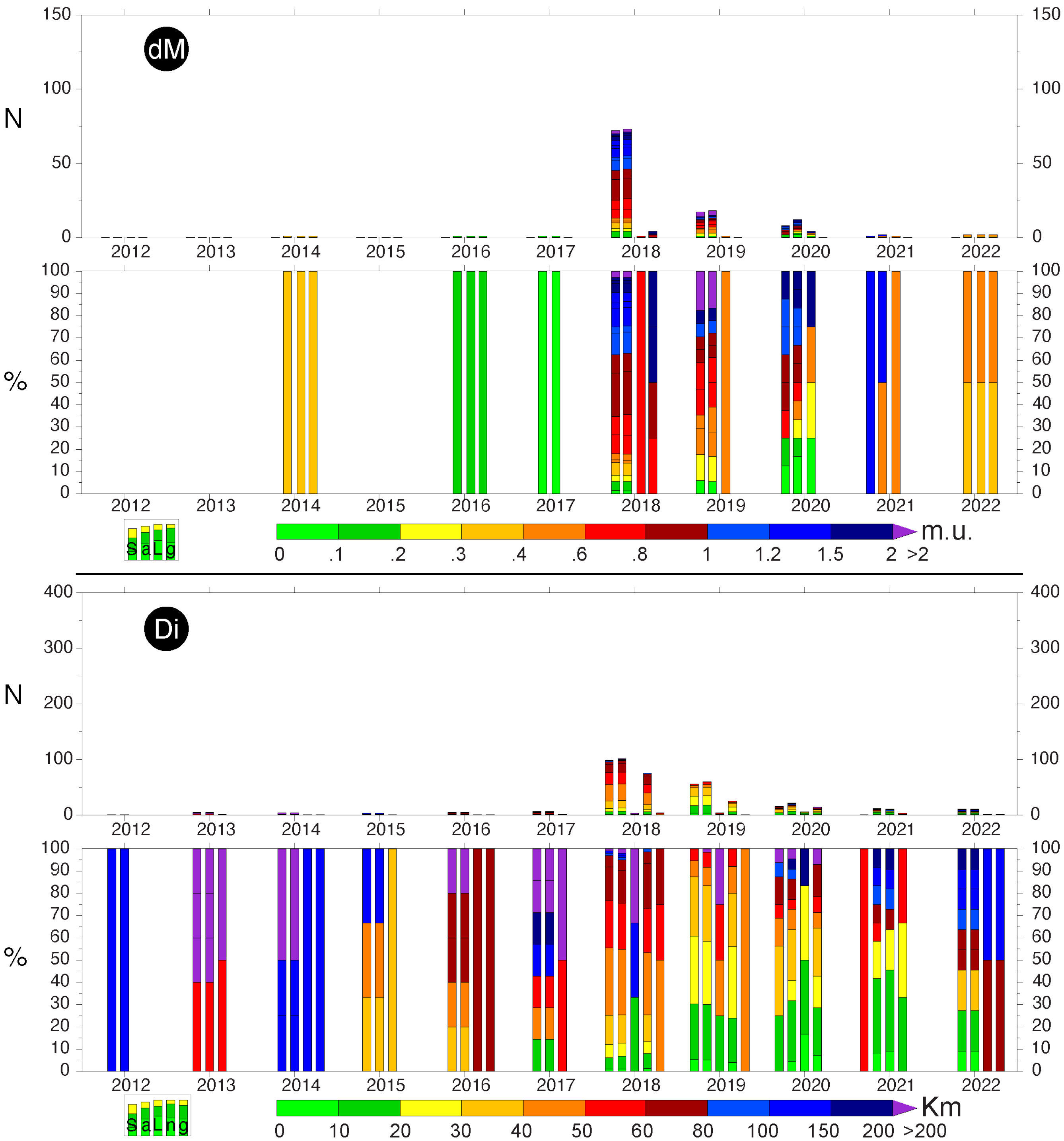


Figure S15 - as in Fig. S9 for corrected intensity.

Data and Resources

GSHHG "Global Self-consistent Hierarchical High-resolution Geography" database, available at www.soest.hawaii.edu/pwessel/gshhg/

EMSC ID, available at https://seismicportal.eu/eventdetails.html?unid= "EMSC ID").