A strategy to build a unified data set of moment magnitude estimates for low-to-moderate seismicity regions based on European–Mediterranean data: application to metropolitan France

Aurore Laurendeau, Christophe Clement, Oona Scotti

Geophysical Journal International

Geophys. J. Int. (2022) **230,** 1980–2002
Advance Access publication 2022 February 26
GJI Seismology

https://doi.org/10.1093/gji/ggac085

A strategy to build a unified data set of moment magnitude estimates for low-to-moderate seismicity regions based on European–Mediterranean data: application to metropolitan France

A. Laurendeau, C. Clément and O. Scotti

Institut de Radioprotection et Sûreté Nucléaire (IRSN), PSE-ENV, SCAN, BERSSIN, Fontenay-aux-Roses 92262, *France. E-mail: aurore.laurendeau@irsn.fr*

Accepted 2022 February 24. Received 2022 February 21; in original form 2021 April 29

SUMMARY

Seismic hazard assessment (SHA) requires earthquake events characterized in terms of the moment magnitude (M_w) scale both in the earthquake catalogues and in the ground motion databases (SHA data sets). However, published M_w estimates for the same event may actually differ due to the adoption of different computation methods, input data, or seismic networks. The main objective of this paper is to develop a novel strategy to assign a unique M_w value for each earthquake and unified to an M_w reference data set that can be applied to low-to-moderate seismicity regions ($M_w < 4.5$) such as Metropolitan France, the target region of our study. The strategy consists of first collecting all M_w estimated directly by spectral or temporal inversion in a large area of the European–Mediterranean zone, including regional and specific studies in our target region. Then, the sources of M_w estimates are ranked in five different categories, from global to specific studies. A GCMT (Global Centroid Moment Tensor project) standard M_w value is recomputed and a quality flag is assigned depending on the level of the information provided by the sources. M_w estimates are then compared at the European–Mediterranean scale with those provided by the CMT services, chosen as reference. M_w estimates are subsequently corrected for their systematic deviations by progressively including lower magnitude estimates as one moves from high ranked categories to lower ones at each step of the unification procedure. In this work, the collected data set contains 6752 M_w estimates of 4454 shallow events (depth < 40 km) that occurred between 1963 and 2019 and computed by 34 sources (bulletins, publications,. . . , etc). A unified M_w value could be attributed to 185 events ($2.36 \leq M_w \leq 5.35$) in France and neighbouring regions and 3351 events ($3.06 \leq M_w \leq 7.57$) in the European–Mediterranean region. Finally, we showed that this M_w data set is consistent with European SHA data sets but leads to slightly larger M_w values compared to the French SHA data sets. We also noted that some variability still subsist in M_w data set compared to the existing SHA data sets, especially for $M_w < 5.0$. This new unified M_w data set is a first step toward the building of future catalogues. It will contribute to greatly improve the coherence of M_w definitions in the next generation of SHA data sets for both metropolitan France and Europe.

Key words: Europe; Earthquake hazards; Earthquake source observations; Statistical seismology.

1. INTRODUCTION

For a coherent application of SHA, it is paramount to dispose of magnitudes (M) defined in the same scale and ideally estimated using the same approach in both the earthquake catalogues and in ground motion databases, the two fundamental data sets of any seismic hazard assessment (hereafter referred to as SHA data sets). Introduced by Kanamori (1977), the moment magnitude (M_w) is considered today by the community as the magnitude scale of reference for SHA data sets (e.g. Grünthal & Wahlström 2012; Laurendeau *et al.* 2013; Kotha *et al.* 2020). This magnitude has the advantage of not saturating for the largest earthquakes compared to other magnitude scales. In addition, M_w is directly linked to the source parameters (Aki 1966), especially to the seismic moment (M_0). Different methods depending on the target magnitude range have been implemented to estimate M_w based on waveform inversions in time or in frequency domains (referred to hereafter as the **'direct'** methods). At the global scale, M_w are routinely estimated for $M_w \geq 5.5$ events since 1976 (e.g. Dziewonski *et al.* 1981) and from

$M_w \geq 4.5$ since the end of the 1990s (e.g. Ekström *et al.* 1998). M_w can be estimated regionally down to M_w 3.5 (e.g. Dreger & Helmberger 1993) and locally to even lower magnitudes if the network allows it (e.g. Godano *et al.* 2013).

The different methods and standard procedures (e.g. IASPEI 2013) used to estimate M_w do not necessarily lead to the exact same M_w estimate for an earthquake. For example, Scognamiglio *et al.* (2016) reported for the 2012 May 20 Emilia earthquake M_w values ranging between 5.63 and 6.12. Several studies have already shown that the M_w estimated from a regional network are generally smaller than those using teleseismic recordings (Gasperini *et al.* 2012; Konstantinou 2015; Kishida *et al.* 2018). Differences are generally due to a combination of the use of a higher frequency signal, a larger azimuthal gap and the size of the rupture that can exceed the regional network size.

Different strategies are adopted in the literature to define a reference M_w data set with a **'unique'** M_w estimate for each event.

A first strategy is to prefer the M_w estimates given by one source over another following a priority scheme (e.g. Fäh *et al.* 2011; Grünthal & Wahlström 2012). This strategy has the advantage that the original M_w are kept as they are, but the heterogeneities between the M_w sources are also kept. The European–Mediterranean Earthquake Catalogue (EMEC, Grünthal & Wahlström 2012) follows this strategy by prioritizing M_w estimate sources in this way: (1) the specific published studies; (2) the Swiss Moment Tensor Solutions (Braunmiller *et al.* 2005); (3) the European–Mediterranean Regional Centroid Moment Tensor Solutions (Pondrelli 2002; Pondrelli *et al.* 2004, 2007, 2011) and (4) regional catalogues of European countries if they exist. Otherwise, the majority of the M_w was deduced from empirical magnitudes converted from magnitude scale conversion laws. In this case, the laws are developed from an M_w reference data set containing events prior to 1993 (Grünthal & Wahlström 2003). Both the Pan-European Engineering Strong Motion data set (ESM, Bindi *et al.* 2018; Lanzano *et al.* 2019) and the recently published GMM of Kotha *et al.* (2020) used the M_w estimated following the EMEC strategy.

A second strategy is to correct the systematic discrepancies in M_w estimates relative to a source chosen as reference to achieve **'unified'** M_w values and then average the estimates for each event (e.g. Gasperini *et al.* 2012; Konstantinou 2015; Kishida *et al.* 2018). In the Italian catalogue, Gasperini *et al.* (2012) chose the M_w estimates from the Global Centroid Moment Tensor project (GCMT, also known as Harvard M_w) as reference and applied offset corrections to the other M_w sources (up to +0.2 to the Italian regional M_w estimates). Then, the weighted average (Wa) M_w values are used directly in the Italian seismic catalogue (CPTI15, Rovida *et al.* 2016) and to develop magnitude scale conversion laws (Gasperini *et al.* 2013). This unification procedure was also followed in other regions, such as Greece (Konstantinou 2015) and Taiwan (Kishida *et al.* 2018), where discrepancies were found with regional M_w values up to +0.18 and +0.43, respectively.

Implementing catalogues following similar strategies would help improving the consistency of the SHA data sets both at regional (e.g. Woessner *et al.* 2015) and national scales (e.g. Drouet *et al.* 2020). Currently, the M_w unification strategy has only been applied to some moderate-to-high seismicity countries by collecting direct M_w estimates from a limited number of sources (global and regional approaches). In low-to-moderate seismicity regions there is a need to estimate magnitudes for $M_w < 4.5$ when performing probabilistic SHA (PSHA). Generally speaking, none or few direct M_w estimates exist for this magnitude scale range. Thus, M_w are deduced from other magnitude scales by applying magnitude scale conversion laws established from a reference M_w data set (referred to hereafter as the proxy M_w). SHA data sets are then built from these various heterogeneous magnitude definitions. Heterogeneities can impact both Gutenberg–Richter fitting relationships (Beauval *et al.* 2020; Drouet *et al.* 2020) as well as the standard deviation of GMMs (Kotha *et al.* 2016; Ktenidou *et al.* 2018).

Metropolitan France SHA data sets are particularly concerned by this issue and the most recent national efforts well illustrate the challenges. On the one hand, Cara *et al.* (2015) built the instrumental SI-Hex (acronym for 'Sismicité Instrumentale de l'Hexagone') catalogue covering the period 1962–2009. This catalogue includes only proxy M_w estimated from various conversion laws established for several magnitude scales and different periods using several inhomogeneous M_w reference data sets (Cara *et al.* 2017; Laurendeau *et al.* 2019). The majority of the magnitude scale conversion laws are based on the local magnitude (M_L) provided by the Laboratory for Detection and Geophysics (LDG). However, several studies (Cara *et al.* 2017; Drouet *et al.* 2020) have shown that SI-Hex M_w estimates are lower compared to other M_w estimates that suggest revising the SI-Hex catalogue strategy. On the other hand, the French Seismological and Geodetic network of the European Plate Observing System (Résif-Epos) ground-motion data set published by Traversa *et al.* (2020) containing 468 events recorded in France and surroundings between 1996 and 2016 is also concerned by this issue. In fact, this data set consists of heterogeneous M_w estimates either taken from the SI-Hex catalogue or from the M_L provided by the French national seismic network (RéNaSS) converted applying the magnitude scale conversion law of Grünthal *et al.* (2009), developed for EMEC.

In this paper, the main objective is to develop a new strategy to assign a unique and unified M_w value to seismic events that occurred in low-to-moderate seismic areas ($M_w < 4.5$) with a focus on metropolitan France. We first present the M_w estimation methods and the sources collected to compile a data set of direct and standard M_w estimates for the European–Mediterranean region. Then, we introduce the strategy allowing to obtain a data set with a unique and unified M_w value for each event. Finally, this new unified M_w data set is presented and the impact of the strategies to build M_w data sets is discussed by intercomparing our data set with existing SHA data sets in both metropolitan France and Europe.

2. COMPILING A DATA SET OF 'DIRECT' M_w ESTIMATES

The first step is to collect all 'direct' M_w estimated using an inversion in time or in frequency domains. Our target region is an area around the metropolitan France, called hereafter 'FRANCE' (6°W–10°E; 41°N–52°N, Fig. 1), corresponding to the SI-Hex extended zone used in Cara *et al.* (2017). We collect instrumental events available since 1962, corresponding to the beginning of the instrumental period for metropolitan France (Duverger *et al.* 2021). In addition, we collect M_w estimates for events that occurred in a larger area required for the unification of the M_w estimates discussed later in Section 3.2. This area covers a European–Mediterranean region (20°W–50°E; 25°N–60°N, Fig. 1), called

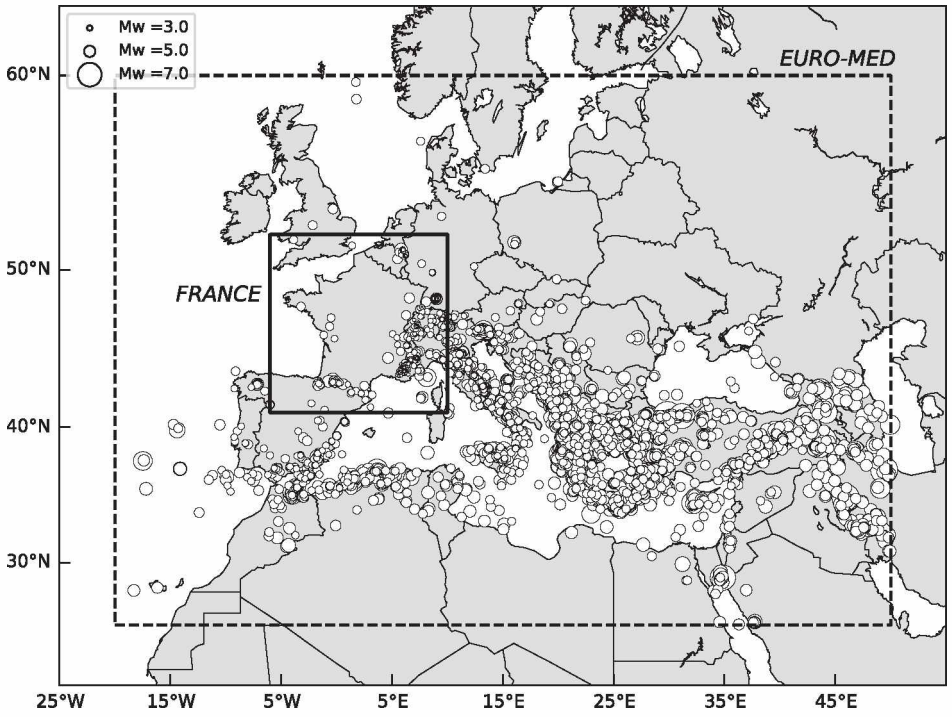

Figure 1. Location map of the 4454 earthquakes with M_w estimates (from 1963 to 2019) collected in this study (presented in Section 4.1) and of the two regions: (FRANCE), an extended area around metropolitan France [6°W–10°E; 41°N–52°N] and (EURO-MED) the European–Mediterranean area [20°W–50°E; 25°N–60°N].

hereafter 'EURO-MED'. This area was previously used in the Gasperini *et al.* (2012)'s study. We exclude events deeper than 40 km which are most likely related to subduction events.

In the following subsections, first we introduce the main specificities and nomenclature of the various methods used for estimating M_w of events collected in this study. Then, we list the sources collected and specify data selection criteria. We review the level of information provided by the different sources, recompute an M_w value following the GCMT standards and attribute quality flags depending on the level of information provided by the sources. The main characteristics of this data set are finally presented.

2.1 Direct M_w estimates: existing methods and associated nomenclatures

Below we briefly recall the difference between the M_w estimation methods. In order to distinguish them we follow the National Earthquake Information Center (NEIC) three-letter nomenclature.

Mwc identifies M_w estimates based on centroid inversion of a combination of very long-period body waves, mantle waves and surface waves (e.g. Dziewonski *et al.* 1981; Ekström *et al.* 1998), *Mwb* those based on body waves (e.g. Sipkin 1982, 1994) and *Mww* those based on ultra-long-period W-phase (e.g. Kanamori & Rivera 2008). These methods were originally developed to estimate $M_w \geq 5.5$ events based on teleseismic distant stations. Recent developments (Hayes *et al.* 2009; Ekström *et al.* 2012) allowed to consider lower magnitudes ($M_w \geq 4.5$). International agencies such as GCMT, the Regional CMT (RCMT), NEIC and the GEOFON program at the German Research Centre for Geosciences [GeoForschungsZentrums (GFZ)] in Potsdam publish M_w computed with this kind of methods.

Methods based on the full waveforms from regional distance stations are identified by *Mwr* and are usually applied to $Mwr \geq 3.5$. The classical tool used to estimate *Mwr* is the Time Domain Moment Tensor inversion (TDMT) scheme developed by Dreger (2003) which is based on inversion of the full waveforms mainly of intermediate and long-period surface waves. The solution is determined by fitting the synthetic seismograms to the observed data, with the variance reduction (VR) parameter measuring its quality. This method was initially implemented by the Berkeley Seismological Laboratory in 1992 for the earthquakes in northern California (Dreger & Helmberger 1993) and it is currently used by several regional observatories in Europe such as ETHZ (acronym for 'Eidgenössische Technische Hochschule Zürich') in Switzerland (Clinton *et al.* 2006; Baer *et al.* 2007), INGV (acronym for 'Istituto Nazionale di Geofisica e Vulcanologia') in Italy (Scognamiglio *et al.* 2009) and IGN (acronym for 'Instituto Geográfico Nacional') in Spain (Rueda & Mezcua 2005). GEOFON also estimates $Mwr < 4.5$ with this method (Saul *et al.* 2011). Based on the work of Clinton *et al.* (2006), a VR threshold larger than 60 per cent is largely used to qualify the solution as reliable. For INGV, Scognamiglio *et al.* (2009) qualify *Mwr* as reliable for VR $\geq$ 40 per cent when at least four stations are used. VR will depend on many factors, including the number of stations available, their azimuthal distribution, their proximity to the source, the accuracy of velocity model used, the quality of the recordings. In the literature, similar methods based on the inversion of the full waveforms recorded at regional distance stations are also developed. For example, the FMNEAR method (Delouis

2014) uses both the waveform inversion of near-source seismic recordings and a linear finite source model for moderate-to-large events (*Mwr* > 5.5–9.0) or a single point source for *Mwr* < 5.5.

Finally, M_w can be computed by measuring the Fourier spectral amplitude of the low frequency plateau of the displacement spectrum of *P* and *S* waves when a local network exists relatively close to the seismic source (e.g. Lancieri *et al.* 2012; Godano *et al.* 2013), hereafter called *Mws*. In this case, an average M_0 value and its standard deviation are estimated from the individual M_0 value deduced for each three component spectrum for each station, while a single M_0 value is determined using the time domain inversion method for the best solution associated to a focal mechanism. The M_w estimate is finally the average of the different measurements and it is important to have a good azimuthal distribution of the recordings to average the effect of the radiation pattern (e.g. Stork *et al.* 2014).

Finally, for sources that do not specify the approach used, the *Mwx* notation is used (e.g. *Mwx* GEOFON).

2.2 Compilation of M_w sources and associated metadata

The main characteristics and references of the M_w estimates given by the different sources used in this study are summarized in Tables 1 and 2.

It should be noted that some institutes publish their M_w estimates only online through webservices or on their institutional webpage. The webpages/webservices consulted in this study are listed in the Supporting Information S1. Depending on the institute, the level of information published can be heterogeneous and the M_w estimates and the associated metadata can be more or less easy to collect automatically. For example, NEIC collects and disseminates information from other institutes, such as *Mwc* GCMT, and provides the results of its own estimations. It is possible to retrieve the event characteristics via a request through the webservices in a quakeml format. In this format, data is grouped into information type blocks, including origins, magnitudes and focal mechanisms. In this study, we have used the preferred origin and the focal mechanism blocks including M_0, the moment tensor elements and focal mechanism axes, taking care to only extract metadata from a NEIC analysis. However, gathering all metadata from NEIC was challenging due to differences in the quakeml entries throughout the time period covered by the data, the presence of duplicate estimates and the difficulty to merge metadata from different blocks. The reader should thus be aware that in spite of the numerous checks performed mistakes may still persist.

Few institutes have attributed a quality flag to their Mw estimates and/or have provided the metadata required to be able to qualify an M_w estimate, such as in particular the azimuthal gap which is generally used to provide a quality to end-users (e.g. Pondrelli *et al.* 2006). In this study, we selected in each source the events for which M_w were considered reliable by their authors (see Table 1, last column).

SismoAzur (http://sismoazur.oca.eu) is the source publishing the most M_w estimates for French metropolitan events. Its data set is composed of M_w estimates collected from different origins selected in this order of priority: (i) from manually revised solutions communicated by B. Delouis, part of which is available on the BCSF website (Le Bureau Central Sismologique Français, http://www.franceseisme.fr/english.php) (ii) from the SismoAzur website (last version accessed in July 2020), which provided events that occurred since 2019 at the global scale and since 2016 at the regional scale, with automatic solutions associated with a quality A or B that depends on the quality of the focal mechanisms and (iii) from the SismoAzur website (last version accessed in April 2019), which provided M_w values for events since 2014 independently of the focal mechanism quality and that we consider reliable in terms of M_w estimate even if the focal mechanism is less reliable.

Fig. 2 displays the number of selected events with an M_w estimate as a function of M_w bins and over given periods from the global (called GLOBAL) to regional scales (EURO-MED, FRANCE). For more than 37 700 earthquakes for which *Mwc* GCMT is available at the global scale, only 11 moderate magnitude earthquakes are associated to an M_w estimate for the FRANCE zone. Similarly, other international institutes (GEOFON, NEIC) that compute M_w estimates for a more recent period provide very few events for the FRANCE area. The European RCMT catalogue contains around 1700 events in EURO-MED area, 35 of which are located in the FRANCE area. The regional approaches allow characterizing more events in the FRANCE area. Especially, neighbouring countries of France (Switzerland, Italy and Spain) automatically compute *Mwr* for $M_L \geq 3.5$ from their own regional networks which allow having M_w estimates for some events at the borders of these countries. Fig. 2 highlights the temporal and spatial completeness of the *Mwc* GCMT data set associated to RCMT on a global and European scale compared to other M_w sources, particularly the regional methods that have been developed after the deployment of the seismic networks.

The collected M_w estimates available in specific published studies (see Table 2) can be grouped in five categories. The first one corresponds to a collection of M_w estimates coming from specific studies with events occurred since the early 1900s compiled to build an earthquake catalogue or to derive magnitude scale conversion laws. The second group is made up of M_w estimates obtained as part of a regional study using a unique approach. The third category corresponds to M_w estimated in the context of the development of alternative M_w estimation methods. A fourth group includes two specific studies in which M_w were computed for low magnitude earthquake from two local networks: one located in the Sampeyre region in Italy, close to the border of France for which Godano *et al.* (2013) computed *Mws* for 730 earthquakes of $0.7 \leq M_w \leq 3.15$ occurring between 2010/10/13 and 2010/11/12 and another one located around the Middle Durance Fault, France, where IRSN (Nechtschein 2003; Volant *et al.* 2003) computed *Mws* ranging between 0.58 and 2.99 for 232 earthquakes that occurred between 1998 and 2007 (Volant *et al.* 2000; Cushing *et al.* 2008). These M_w estimates, currently only available on request from the authors, will be useful in the future to develop new magnitude scale conversion laws. Finally, the last category includes M_w estimated for a specific earthquake occurred in metropolitan France or close to the surrounding political border.

Table 1. Nomenclature, main characteristics and references of M_w methods routinely computed for seismic events for each source listed from global to regional scale approaches. The last column indicates if a specific selection is applied to the events of the original collection. When this is the case, the selection is based on quality criteria of the M_w estimate following the authors' recommendations.

Type	Source	Main references	Time period	Magnitude range	Method	Region	Selection for this study
Mwc	GCMT	Dziewonski *et al.* (1981); Ekström *et al.* (2012)	1976[1]–31/12/2019	*Mwc* ≥ 5.5 until 2003 and then *Mwc* ≥ 5.0	Centroid moment tensor inversion using mantle (125–350 s) and body waves (40–125s), and since 2004 intermediate period surface waves (50–150s) in addition	Global	All
Mwb	NEIC	Sipkin (1982, 1994)	02/03/1996–31/12/2019	5.5 ≤ *Mwb* ≤ 7.5	Centroid moment tensor inversion of long-period body waves (~20–200 s). Applicability limited for larger magnitude than 7.5 due to source complexity		All
Mww	NEIC	Kanamori & Rivera (2008); Hayes *et al.* (2009)	28/03/2005–31/12/2019	5.5 ≤ *Mww* (4.5 ≤ *Mww* if high quality records at a regional scale)	Centroid moment tensor inversion of the W-phase (~50–2000 s)		All
Mwr	NEIC	Dreger & Helmberger (1993); Dreger (2003)	05/10/2008–31/12/2019	3.5 ≤ *Mwr* ≤ 7.0 (the minimum magnitude depends on the density of stations available)	Moment tensor inversion of the whole seismogram at regional distances (~10–100 s). Applicability limited for larger magnitude than 7.0 due to source complexity		All
Mwx	GEOFON	GEOFON Data Centre (1993); Saul *et al.* (2011)	27/02/2008–31/12/2019	4.0 ≤ *Mwr*	Mw estimates from different methods implemented in SeisComp3 that examine different parts of the wave train: teleseismic body waves, mantle and surface waves, and long-period phase waves (W Phase)		All
Mwc	RCMT	Ekström *et al.* (1998); Pondrelli (2002); Pondrelli *et al.* (2011)	01/01/1997–31/12/2019	4.5 ≤ *Mwr* ≤ 5.5	Centroid moment tensor inversion using mainly intermediate-period surface waves but also mantle and long period surface waves	Europe	Only definitive solutions[2] of quality A or B
Mwc	Italian-CMT	Pondrelli & Salimbeni (2006); Pondrelli *et al.* (2006);	06/05/1976–06/12/2015	4.0 ≤ *Mwc*	Compilation for Italy of Mw from GCMT and RCMT with new Mw computed for the period 1976–1996	Italy	Solutions of quality A or B and only the events that occurred between 1976 and 1996
Mwr	SismoAzur	Delouis (2014)	07/04/2014–31/12/2019	2.8 ≤ *Mwr*	The FMNEAR method is based on both the waveform inversion of near source seismic records and on a linear finite-source model	Global but mainly the south of Europe	All
Mwr	IAG	Stich *et al.* (2003, 2006, 2010); Martín *et al.* (2015)	20/10/1986–19/08/2014	3.1 ≤ *Mwr*	Full waveform inversion of the size, the depth and the double-couple focal mechanisms from the time domain inversion methodology introduced by Dreger & Helmberger (1993)	Iberian-Maghreb	when the quality is given selection of a and b
Mwr	IGN-TDMT	Dreger & Helmberger (1993); Dreger (2003); Rueda & Mezcua (2005)	04/02/2002–31/12/2019	3.1 ≤ *Mwr*		Spain	if VR ≥ 60 per cent
Mwr	INGV-TDMT	Dreger & Helmberger (1993); Dreger (2003); Scognamiglio *et al.* (2006); Scognamiglio *et al.* (2009)	24/11/2004–31/12/2019	2.8 ≤ *Mwr*		Italy	Solutions of quality Aa, Ab, Ba
Mwr	SED-TDMT	Dreger & Helmberger (1993); Dreger (2003); Clinton *et al.* (2006); Baer *et al.* (2007)	13/09/1999–28/08/2014	2.8 ≤ *Mwr*		Switzerland	if VR ≥ 60 per cent

[1]An estimate of M_w is given by GCMT for some earlier events (between 1962 and 1976).[2]Reviewed solutions until 31/01/2019.

Table 2. List of individual publications used in this study classified according to their main objectives.

Type of study	Application	References of the study
1. Compilation of a large M_w data set from different authors/approaches in the literature	- The ISC-GEM (International Seismological Centre—Global Earthquake Model) catalogue Version 7.0	- Di Giacomo *et al.* (2015, 2018); Lee & Engdahl (2015)
	- M_w used as reference in the EMEC catalogue	- Scherbaum & Stoll (1983); Grünthal & Wahlström (2003)
2. Compilation of a regional M_w data set defined by using a unique approach	- M_w used as reference for conversion law of Ml for the ECOS catalogue	- Braunmiller *et al.* (2005)
	- M_w used as reference to convert M_s for the ECOS catalogue	- Bernardi (2005)
	- Catalogue for the Pyrenees	- Chevrot *et al.* (2011)
3. Analysis of a large data set of events for a methodological development	- Moment tensor determination at the European scale from regional recordings	- Braunmiller *et al.* (2002)
	- Estimation of M_w from the near-field spectra of strong-motion records	- Delouis *et al.* (2009)
4. Low M_w defined from specific local networks	- Swarm of Sampeyre, Italy	- Godano *et al.* (2013)
	- IRSN Durance (Volant *et al.* 2003; Cushing *et al.* 2008)	- Nechtschein (2003)
5. Specific studies around single events	Baer *et al.* (2007); Diehl *et al.* (2018); Perrot *et al.* (2005); Larroque *et al.* (2009); Larroque *et al.* (2016); Nechtschein & Lesueur (2011); Courboulex *et al.* (1998); Courboulex *et al.* (1999); Courboulex *et al.* (2001); Courboulex *et al.* (2013); Dufumier *et al.* (2000); Dufumier (2002)	

2.3 Recomputing GCMT standard M_w values and assigning quality flags

M_0 and M_w estimates are computed following expert choices that may differ from one institute to another. To have homogeneous M_w values, our objective was to recompute M_0 and M_w following the GCMT standard procedure. However, the metadata provided by the different sources collected in this study are heterogeneous (Supporting Information S2) ranging from international institutes furnishing several metadata, allowing to recompute M_0, to specific studies providing only an M_w value with one significant digit (SD). Thus we developed a strategy to assign a quality to our standard M_w estimates.

The GCMT standard (Method1) is based on the Hanks & Kanamori (1979) and IASPEI (2013) formula:

$$M_W = \frac{2}{3}(\log_{10}(M_0) - 9.1) \tag{1}$$

where M_0 is the moment tensor in N m, is computed from the principal axes of the moment tensor solution as the average of the modulus of the two largest eigenvalues (λ_1 and λ_2) of the moment tensor:

$$M_0 = \frac{1}{2}(|\lambda_1| + |\lambda_2|) \tag{2}$$

In the case of NEIC, M_0 is estimated by the Silver & Jordan (1982) formula based on the individual moment tensor elements (Method2). Differences up to 0.1 units of M_w between these two methods can be observed (Fig. 3, left-hand panel). Because the choice of the M_0 computation formula may be a first source of discrepancy, we also traced the origin of the discrepancies between the original published M_w values and the ones from the GCMT standard procedure. We also found discrepancies due to the accuracy of published M_0, that is the number of SD used to compute M_w. In Fig. 3 (right-hand panel), *Mwc* collected on the GCMT websites are compared to *Mwc* values obtained from M_0 which has been rounded from four to one SD. This comparison reveals that four SD are necessary to reproduce the *Mwc* value published by GCMT. With three and two SD, we can still retrieve the *Mwc* GCMT value for the majority of the events with a maximum deviation of 0.01. However, when only one SD is provided the error could be larger than 0.1 M_w unit. Finally, a third reason for M_w discrepancies, already pointed out by Kagan (2003), is the application of a slightly different formula to compute M_w from M_0 (such as eq. 1), mainly due to different round numbers if the 2/3 factor is already applied. For example, Gasperini *et al.* (2012) found M_w values 0.03 unit lower than the original ones in the ETHZ collection due to the use of a slightly different formula.

Thus, depending on the level of metadata available from each source, an information quality flag, called 'FLAG quality standard Mw', ranging from A to E was assigned to qualify it with respect to the standard procedure as follows:

A. GCMT standard procedure: the principal axes of the moment tensor are available with sufficient SD allowing to recompute M_0 (Method1) and then M_w following the IASPEI (2013)'s recommendations (eq. 1).

B. M_0 is given with at least two SD and other information, such as the individual moment tensor elements (Method2) allows to check the M_0 value or if M_0 is computed from the spectral method (*Mws*).

C. M_0 is given with at least two SD but no information is available to check M_0.

D. M_0 is not available, M_w is available with at least three SD and the equation used to compute M_w is known (M_w can thus be corrected for discrepancies by using eq. 1).

E. Only M_w is available without information about its computation or if M_w is available with less than three SD.

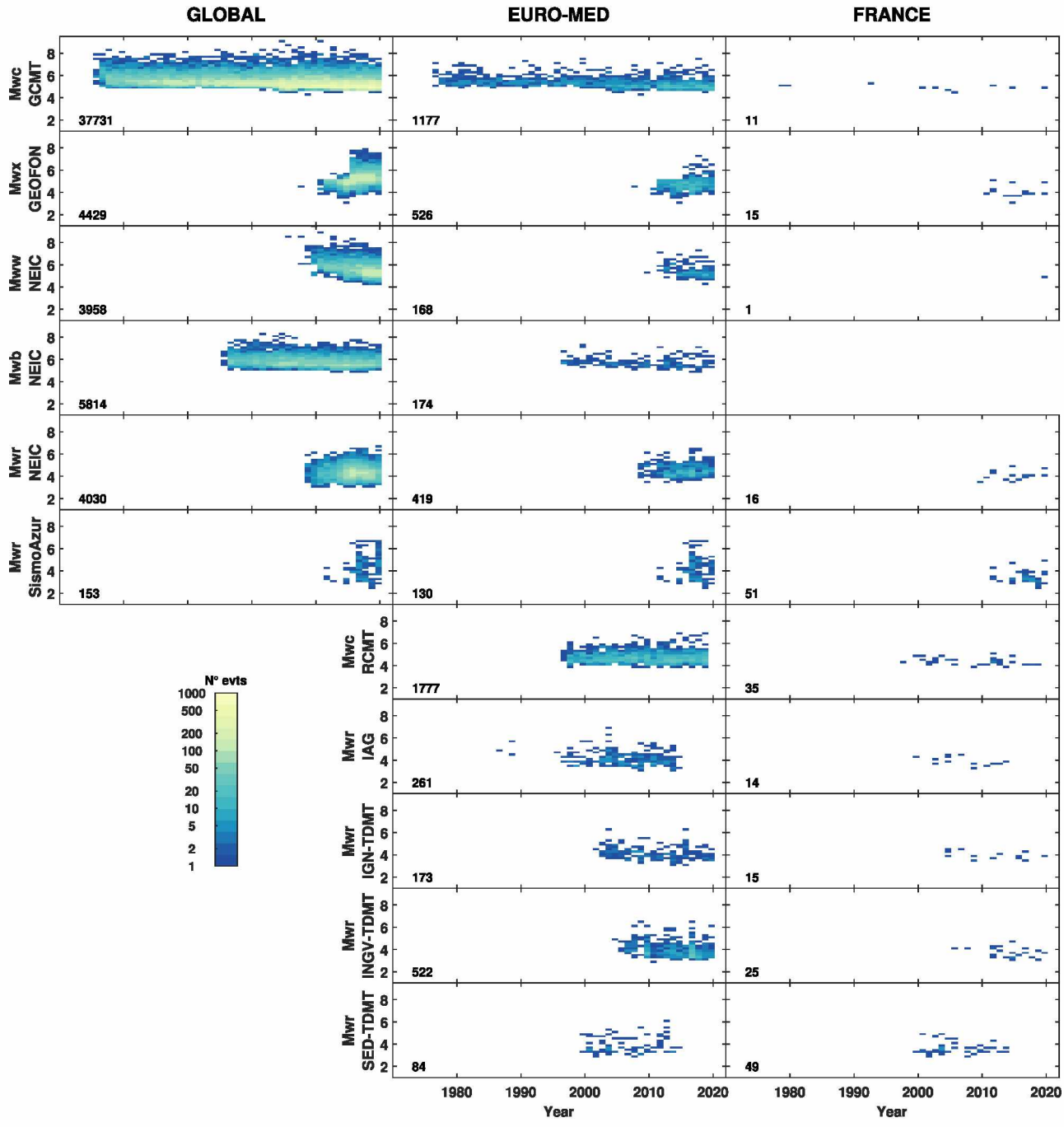

Figure 2. Temporal distribution of earthquakes in terms of M_w estimates provided by the institutes listed in Table 1 for the three geographic areas (GLOBAL, EURO-MED and FRANCE). The magnitude and time bins are of 0.2 and 1 yr, respectively. The total number of earthquakes is indicated at the bottom left and the colourbar indicates the number of events for each bin.

2.4 Merging the compiled M_w

We collected in the EURO-MED region all M_w values given by the different sources mentioned in Section 2.2, with the exception of (i) events for which the solutions were not considered reliable by the authors (see last column of Table 1) and (ii) events with significant differences between the M_0 values published and recomputed by us [difference of $\log_{10}(M_0) > 0.4$].

The different M_w sources were then merged. However, merging two data sets can be a challenging task due to possible large differences in the time origins and the epicentral locations between two solutions or on the contrary when these parameters are too close due to sequences of events. We developed an automatic merging procedure that allows identifying duplicate events between two collections mainly using the differences in time (Δt, s) and in distance (Δd, km). Details of the processing applied for each event is described in the Supporting Information S3.

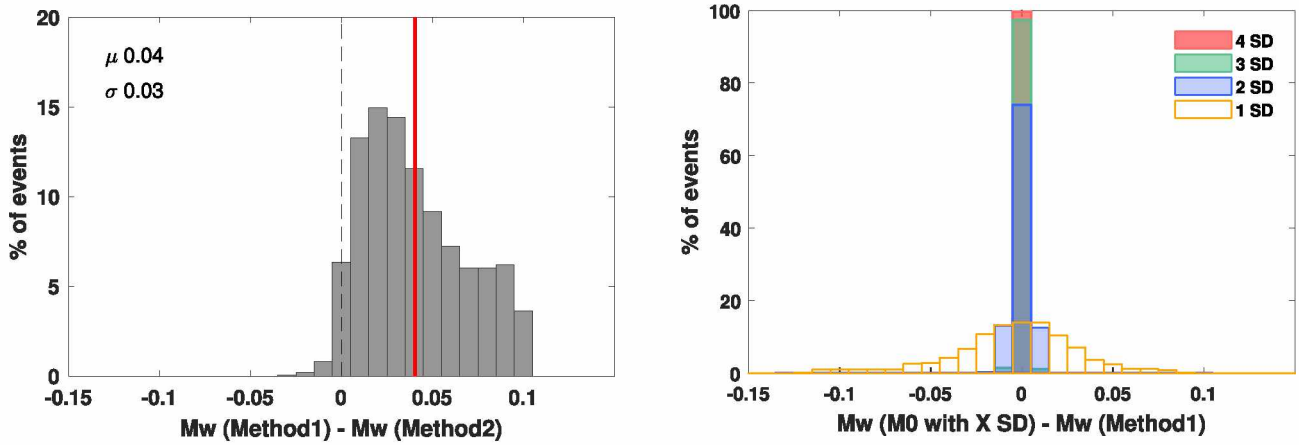

Figure 3. Comparison of M_w values from the GCMT collection defined with three significant digits (SD) and computed from different M_0 values. Left-hand panel: M_0 is computed from the principal axes (Method1) and from the moment tensor elements (Method2). Right-hand panel: M_0 is the value provided by GCMT and rounded to 4/3/2/1 SD (with X the number of digits) and M_0 is computed from Method1. For this example, the GCMT collection is used for earthquakes occurred at the global scale since 2003 for which M_0 is given with four SD.

Table 3. Number of 'direct' M_w collected and percentage of associated 'standard' M_w quality factors estimated in this study shown separately for earthquakes located in the FRANCE region and those located outside in the so-called EURO-MED_clip region (*cf.* Supporting Information S4, sheet 'Mw_sources_used').

Region	Number of M_w estimates	Number of events	Range of Mw	Number (*per cent*) of Mw estimates associated with the following standard M_w quality				
				A	B	C	D	E
FRANCE	1425	1288	0.12–6.30	237 (*17*)	988 (*69*)	90 (*6.3*)	54 (*3.8*)	56 (*3.9*)
EURO-MED_clip	5327	3166	2.82–7.62	3509 (*66*)	1636 (*31*)	62 (*1.2*)	70 (*1.3*)	50 (*0.9*)

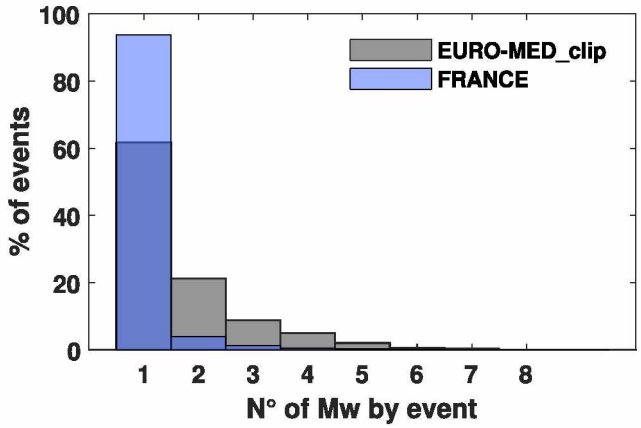

Figure 4. Histogram of the percentage of events as a function of the number of M_w sources available for each event for FRANCE and EURO-MED_clip regions.

The collection of 34 sources allowed identification of 6752 'direct' M_w estimates (sheet 'Mw_sources_used' in Supporting Information S4) concerning 4454 events that occurred between 1963 and 2019 in the EURO-MED region. Table 3 lists the number of events and the percentage of 'standard' M_w estimates for each quality level by distinguishing those within the FRANCE region from those outside in the so-called EURO-MED_clip region. M_w estimates are computed for a large majority of events applying the standard procedure (A: 66 per cent) or with a procedure close to the standard one (A + B: 97 per cent) for the EURO-MED_clip and for FRANCE region (A + B: 86 per cent). Many events of the EURO-MED_clip data set (62 per cent) are associated to an M_w value from only one source (Fig. 4); only four events are associated to a maximum of eight sources. In the FRANCE area, 94 per cent of the events are associated to only one source and three events are associated to a maximum of six sources (Table 4). Table 4 illustrates well the range of discrepancies in M_w estimates that can exists among sources (ranging from 0.14 to 0.42 unit of M_w). *Mwc* GCMT is the source which always estimates the highest M_w value in this case, but the M_w of these earthquakes are below the validity threshold value of this method. Conversely, the *Mwr* INGV-TDMT or an equivalent regional study such as the one of Chevrot *et al.* (2011) provide the lowest M_w values. In addition, M_w estimates based on equivalent methods can lead to significant differences: for example, the *Mwr* values range from 4.16 to 4.42 for the Vallorcine earthquake. Table 4 also presents

Table 4. Examples of M_w estimates collected in the case of four earthquakes that occurred in or close to metropolitan France. ΔM_w indicates the difference between the largest and the lowest ones.

M_w sources	M_w types	Event 1366 2004/09/18 12:52:18 Spanish Basque region	Event 1481 2005/09/08 11:27:17 Vallorcine	Event 3423 2014/04/07 19:27:00 Barcelonette	Event 4432 2019/11/11 10:52:46 Le Teil
GCMT	*Mwc*	4.68	4.58	4.95	4.91
RCMT	*Mwc*	4.52	4.51	4.93	
GEOFON	*Mwx*			4.86	4.87
NEIC	*Mww*				4.84
NEIC	*Mwr*			4.82	4.77
INGV-TDMT	*Mwr*		4.16	4.71	
IGN-TDMT	*Mwr*	4.45			
IAG	*Mwr*	4.49			
SismoAzur	*Mwr*			4.79	4.86
SED-TDMT	*Mwr*		4.42		
Chevrot *et al.* (2011)	*Mwr*	4.36			
Delouis *et al.* (2009)	*Mws*		4.47		
IRSN-RAP	*Mws*	4.50	4.31		
ΔM_w		0.32	0.42	0.24	0.14

the five M_w values currently collected for the recent 11 November 2019 Le Teil earthquake, for which the M_w estimates seem more stable, with discrepancies of 0.14 unit of M_w.

3. A STRATEGY TO UNIFY M_w ESTIMATES

In this section, a new strategy is implemented to assign a unique and unified M_w value to each earthquake event by adapting the unification procedures previously proposed in the literature (Gasperini *et al.* 2012; Konstantinou 2015; Kishida *et al.* 2018) in order to integrate lower M_w events, essential for low-to-moderate seismicity countries. We developed a hybrid strategy combining (i) a priority scheme of sources classified in five categories ranging from global to specific analyses and (ii) a unification of the M_w estimates with a reference, which is a combination of the M_w values given by the CMT services. In this strategy, the definition of the M_w reference data set progressively evolves by including lower M_w estimates as one moves to lower categories. In this way, the majority of sources can be unified. Unfortunately, the last category includes M_w estimates that cannot currently be unified. The hybrid strategy is summarized in Fig. 5 and more details will be given in the following paragraphs.

3.1 Ranking of the sources

Ranking (left part of Fig. 5) of the different sources of M_w (from global to specific published studies) depends on the method used to estimate M_w (see Section 2.1), the standard M_w quality of the published data set (see Section 2.3) and the number of earthquakes with an M_w estimate in each data set for the EURO-MED region (Fig. 2). Indeed, to estimate homogeneous M_w on a larger scale than our target area and over a longer period of time, we have favoured sources estimating M_w on large data sets.

The first category ('I. CMT services') constitutes our first reference M_w data set to which the other M_w sources are unified. The *Mwc* GCMT are considered as the reference by the seismological community due its large spatial, temporal and magnitude range coverage (Fig. 2). Several authors have chosen *Mwc* GCMT as the prime reference in their catalogues (e.g. CPTI15; ISC-GEM). Following the strategy of Pondrelli *et al.* (2006) for the Italian-CMT catalogue, we constituted this first category, referred to as $Mw_{unified\_I}$, as a combination of centroid moment tensor solutions selected according to the following priority scheme:

(1) *Mwc* GCMT $\geq$ 5.5;
(2) *Mwc* RCMT;
(3) *Mwc* Italian-CMT (for earthquakes occurring between 1976 and 1996 new *Mwc* were computed in a similar way to RCMT for Italy);
(4) *Mwc* GCMT $<$ 5.5 if no M_w estimate provided by sources of categories II and III (as the quality of the M_w estimate is lower in the low M_w range, we preferred to use M_w from lower categories when available).

The category II includes additional global sources that provide moment tensor solutions estimated in a uniform way at the global scale and with information allowing M_0 to be defined accurately (see Section 2.3). For this category II, the solutions however are provided with a smaller temporal window than I (Fig. 2). All regional services that systematically furnish *Mwr* since 2000–2010 are included in the category III. Then, the category IV includes publications that provide M_w estimates computed using the same method for a collection of events but

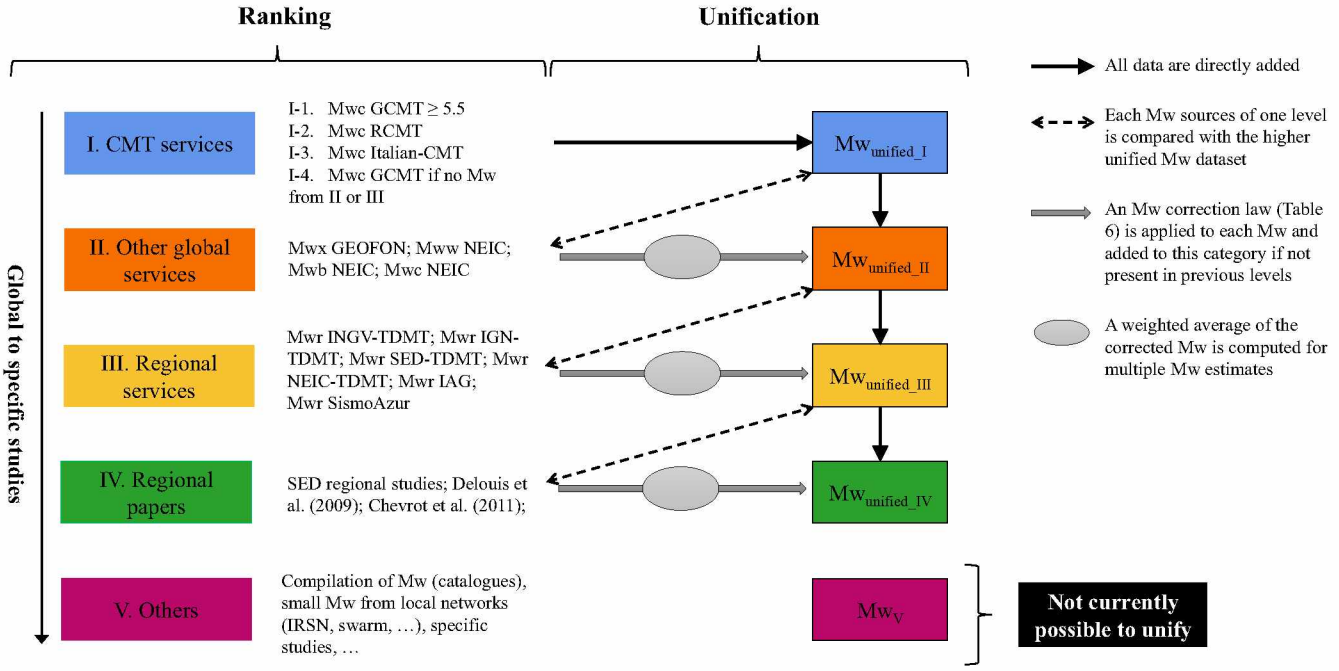

Figure 5. Scheme of the hybrid strategy used to build a reference data set of unique and unified M_w estimates. The colours chosen to differentiate the categories will be used in the following figures.

Table 5. Number and percentage of earthquakes with an M_w estimate for the five categories of sources for FRANCE and EURO-MED_clip.

Category	FRANCE		EURO-MED_clip	
	Number	Per cent	Number	Per cent
I	51	4	3004	56
II	16	1	838	16
III	170	12	1408	26
IV	70	5	77	1
V ($M_w \geq 2.5$)	80	6	-	-
V	1118	78	-	-
Total	1425	100	5327	100

not estimated with this method in a systematic way over time (see Table 2). In the case of Switzerland, we combined the M_w values available within different publications in a unique data set called hereafter 'SED regional studies' following a priority scheme similar to ECOS-09 (Fäh *et al.* 2011):

(1) Bernardi (2005).
(2) Baer *et al.* (2007).
(3) Braunmiller *et al.* (2005).
(4) Braunmiller *et al.* (2002).

Finally, all the other M_w estimates are integrated into the last category V.

For the II to IV categories, the different M_w sources in each category were considered as equivalent. For earthquakes associated to several M_w estimates, an average M_w value is thus computed. Indeed, it is difficult to give a preference to a source within a category; it would be better to individually analyse each event, but this task is too time-consuming when building a large data set.

The resulting classification for each earthquake in our data set is reported in the Supporting Information S4 (sheet 'Mw_sources_used', field 'Mw_ranking'). Table 5 indicates the number of M_w for each of the five categories of sources. For EURO-MED_clip area, 56 per cent of earthquakes were described by an M_w value provided by the category I, conversely for the FRANCE area, 12 per cent are from the category III including the regional approaches and 78 per cent from category V. It should be noted however that only 6 per cent of the category V earthquakes have magnitudes values $M_w \geq 2.5$.

3.2 Unification of the M_w values

The second step (right-hand part of Fig. 5) is to unify the M_w values to be consistent with a reference. Thus, the M_w values from one source (X) have to be compared with an M_w data set chosen as reference (Y) and then eventually corrected to be unified with this reference.

3.2.1 Choice of statistical tools to perform regressions

In the literature, different regression techniques are implemented to compare magnitudes considering the associated uncertainties (e.g. Stromeyer *et al.* 2004; Bethmann *et al.* 2011; Gasperini *et al.* 2012, 2013). Lolli & Gasperini (2012) found comparable results from these techniques. When monitoring services do not report an uncertainty associated with the Mw estimates, ordinary orthogonal regression (OOR) are often performed (e.g. Grünthal *et al.* 2009; Gasperini *et al.* 2012) under the assumption of equal variances for the two magnitudes [$\eta = \sigma^2(Y)/\sigma^2(X) = 1$]. Castellaro & Bormann (2007) and Gutdeutsch *et al.* (2011) recommended in such cases to compare OOR results with the ordinary linear least square regression (OLS) ($Y \rightarrow aX + b$) and the inverted OLS_inv ($X \rightarrow cY + d$) technique. This ensures that the $\eta = 1$ hypothesis can be considered reasonable (i.e. the OOR line is close to the angular midst formed by OLS and OLS_inv lines).

In this study, we used the OOR technique with the hypothesis of $\eta = 1$. The OLS techniques were used as a tool to quantify the variances. We also checked two hypothesis: (i) that the mean difference between two magnitudes estimates, $\overline{\Delta M}$ is equal to zero can be rejected and thus an offset correction is required and (ii) that the slope of the linear conversion law is equal to unity ($Y = X + b$) can be rejected and thus a scaling correction is needed. The two statistical tests were rejected if the p-value, computed as described in Konstantinou (2015), is lower or equal to 0.01, that is if the mean difference is significantly different from zero and the slope is significantly different from unity. Thus, if only the first test is rejected, only an offset correction is applied such as $Y = X + b$. If the second test is rejected (whatever the result of the first test), a complete scaling correction such as $Y = aX + b$ is applied.

3.2.2 Implementation of the unification procedure

The M_w value intercomparisons were carried out progressively following the proposed ranking by setting up for each category of sources a new data set, called $Mw_{unified\_X}$ (with X for the category number), which serve as a reference for the lower category (see Fig. 5). Table 6 presents the M_w correction laws obtained for all sources and Fig. 6 displays four intercomparison examples of M_w sources from category I to IV compared with their own reference. Figures showing the intercomparisons for all sources are included as Supporting Information S5. For each M_w intercomparison, the four conversion laws obtained from the different regression techniques discussed before are presented. We find that the $\eta = 1$ hypothesis is reasonable and that the variances associated to the M_w values are relatively low given the small angle observed between OLS and OLS_inv laws (explaining why the OLS and OLS_inv lines are overlying in Fig. 6).

For the category I, the $Mw_{unified\_I}$ data set is directly built following the priority scheme of the three CMT sources. Fig. 6a displays the comparison between *Mwc* GCMT and a combination of *Mwc* from RCMT (1997–2019) and Italian-CMT (1976–1996 only for Italy). The similarity of the OOR, OLS and OLS_inv laws and a low standard deviation (σ) of 0.06 attest to the compatibility of these *Mwc* values for $Mwc \geq 5.5$. For lower *Mwc* values, the dispersion is larger (σ of 0.08) and a scaling bias can be observed with increasingly higher *Mwc* GCMT values compared to *Mwc* (RCMT, Italian-CMT) with decreasing *Mwc* values. As already mentioned by Gasperini *et al.* (2012), GCMT tends to overestimate *Mwc* for $M_w < 5.5$, that is why we preferred M_w from other CMT services or lower categories (II, III) for *Mwc* GCMT < 5.5 when estimates are available.

For the categories II to IV, the M_w values of each source were compared with its M_w of reference, $Mw_{unified\_X}$, defined at the level of the higher category (see Table 6 and Fig. 5). According to the results of the statistical test, an offset or a scaling correction was defined and applied to each source (see Table 6). Then, a weighted average (Wa) is computed when more than one M_w estimate is available for one event inside a category, using the standard deviations of the laws as a weight (Table 6). We verified the agreement between the new unified M_w values with their reference by confirming that corrections are no longer necessary (see Table 7). Finally, the M_w estimates from a category were used when the corresponding events were not already associated with an M_w estimate from the higher levels. These different steps are represented schematically in Fig. 5.

For category II, a simple offset correction is required for the different sources (see Table 6). Using events occurring from 2011 to 2019, we defined an offset correction of + 0.1 for *Mwx* GEOFON compared to the reference $Mw_{unified\_I}$ (Fig. 6b). This result is stable regardless of the M_w range: similar tendencies are observed for earthquake of $M_w < 5.5$ (offset of + 0.1) and of $M_w \geq 5.5$ (offset of + 0.09).

Concerning the sources of category III, *Mwr* SED-TDMT is the source of M_w with the most dispersion (σ of 0.14) but not requiring offset correction. *Mwr* IGN- and INGV-TDMT required an offset correction of +0.1 and +0.24, respectively. For the other sources of this category III, the statistical tests revealed that a scaling correction is required. In fact, in the case of *Mwr* INGV-TDMT, the p-value from the scaling test (hypothesis of $a = 1$ in the law Y = aX + b) was close to the chosen threshold value of 0.01. The intercomparisons of *Mwr* SismoAzur with the reference $Mw_{unified\_II}$, presented in Fig. 6(c), shows that a scaling correction is required and the dispersion of the M_w values is low (σ of 0.09), indicating a clear trend. In the case of *Mwr* NEIC and IAG, a scaling correction was applied in spite of the p-value being close to the threshold of 0.01.

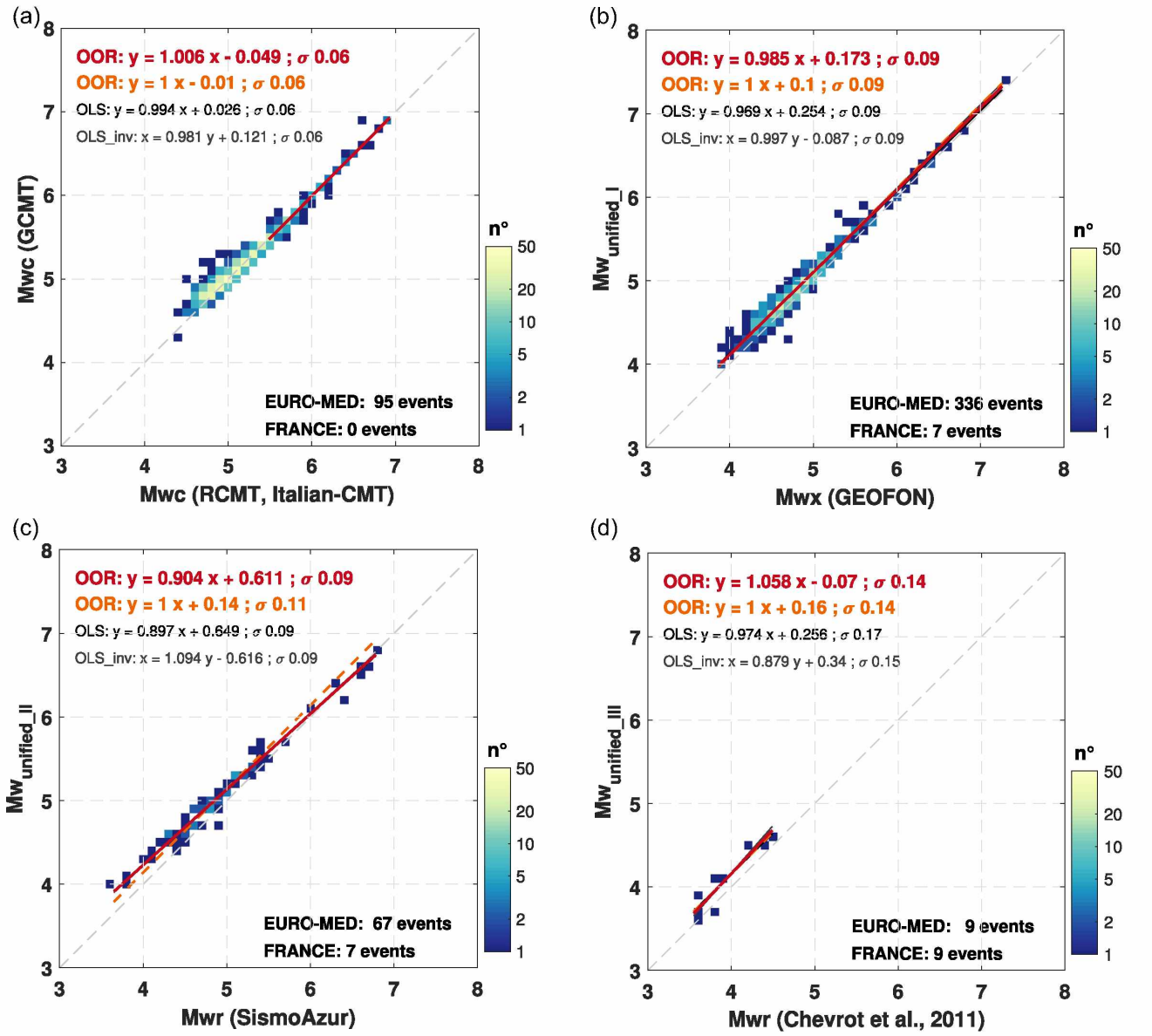

Figure 6. Selection of intercomparisons between the M_w estimates of the different sources with the corresponding data set of reference. The data are represented in terms of density of events indicated by the colourbar for bins of 0.1 unit of M_w. Different regression techniques are computed for events occurring in the EURO-MED zone: OOR is represented in case of a classical linear regression ($Y = aX + b$) (red line) and for a linear regression for which the slope is fixed to one ($a = 1$; $Y = X + b$) (orange line); OLS (dark grey line) and OLS_inv (grey line). The light grey dashed line corresponds to $Y = X$. The number of events included in EURO-MED and FRANCE zones are indicated at the right bottom of each subplot.

For category IV, we found an offset correction of −0.04 for the *Mwr* of the SED specific studies. The most important offset correction was applied to the M_w values from the specific study of Chevrot *et al.* (2011) for earthquacks of the Pyrenees region. We define a clear correction of 0.16 but only from nine events (Fig. 6d). Note that for this category IV, the restricted data sets available limits the robustness of the regression. The estimated corrections are still in the same range as those of other sources with larger data sets.

Finally, for category V, with the exception of 20 earthquakes out of 1118 (Table 5), already characterized by other sources of the previous categories (see Supporting Information S5), no comparisons could be performed and therefore these M_w values were not unified.

4. RESULTS AND DISCUSSION

4.1 A data set of direct, unique and unified M_w estimates

The resulting data set contains 4454 shallow seismic events (depth < 40 km) occurred from 1963 to 2019 in the European–Mediterranean region associated to a direct and unique M_w value (table given as Supporting Information S4, sheet 'Mw_unified'). For 3351 earthquakes, a unified M_w value based on a 'CMT services' reference is provided. For the remaining 1103 earthquakes, the M_w estimate of the original study

Table 6. Intercomparisons of M_w estimates, from the different sources used in this study and ranked by category, with the M_w values of reference ($Mw_{reference}$) defined at the level of the higher category. The corresponding number of data used for the comparisons coming from EURO-MED (FRANCE) is indicated. M_w correction laws were established with their corresponding standard deviations (σ intersources) for each source according to the p-value results. P-values in bold indicate that offset and/or scaling corrections are needed ($p < 0.01$).

$Mw_{reference}$	M_w	Number of data in EURO-MED (FRANCE)	p-Value ($\overline{\Delta M} = 0$)	p-Value (a = 1)	Mw correction laws	σ Inter sources
$Mw_{unified\_I}$	*Mwx* GEOFON	336 (7)	**<0.01**	0.15	$Mw_{reference} = Mw + 0.10$	0.09
	Mww NEIC	124 (0)	**<0.01**	0.27	$Mw_{reference} = Mw + 0.04$	0.09
	Mwb NEIC	162 (0)	**<0.01**	0.65	$Mw_{reference} = Mw + 0.09$	0.11
$Mw_{unified\_II}$	*Mwr* NEIC	299 (13)	**<0.01**	**<0.01**	$Mw_{reference} = 0.952.Mw + 0.342$	0.09
	Mwr SismoAzur	67 (7)	**<0.01**	**<0.01**	$Mw_{reference} = 0.904.Mw + 0.611$	0.09
	Mwr INGV-TDMT	219 (10)	**<0.01**	0.02	$Mw_{reference} = Mw + 0.24$	0.11
	Mwr IAG	69 (3)	0.03	**<0.01**	$Mw_{reference} = 0.911.Mw + 0.46$	0.12
	Mwr IGN-TDMT	52 (3)	**<0.01**	0.44	$Mw_{reference} = Mw + 0.1$	0.11
	Mwr SED-TDMT	27 (10)	0.18	0.07	$Mw_{reference} = Mw$	0.14
$Mw_{unified\_III}$	*Mwr* of SED specific studies	61 (13)	**<0.01**	0.61	$Mw_{reference} = Mw - 0.04$	0.11
	Mws of Delouis *et al.* (2009)	7 (4)	0.28	0.72	$Mw_{reference} = Mw$	0.12
	Mwr of Chevrot *et al.* (2011)	9 (9)	**0.01**	0.73	$Mw_{reference} = Mw + 0.16$	0.14

Table 7. Comparison between the unified M_w values at one level (after applying the M_w correction laws presented in Table 6) with the unified M_w values of the category defined at the higher level.

$Mw_{reference}$	M_w	Number of data in EURO-MED (FRANCE)	p-Value ($\overline{\Delta M} = 0$)	p-Value ($a = 1$)	σ inter sources
$Mw_{unified\_I}$	$Mw_{unified\_II}$	481 (7)	0.85	0.47	0.09
$Mw_{unified\_II}$	$Mw_{unified\_III}$	549 (30)	0.83	0.04	0.10
$Mw_{unified\_III}$	$Mw_{unified\_IV}$	73 (24)	0.45	0.30	0.11

is reported (all in category V). It should be noted that in this data set, we have reported the location given by the source providing the M_w estimates. However, for the building of an earthquake catalogue, a more regional location would be preferable.

This data set was compiled with a focus on metropolitan France, which means that not all data sources were collected at the European scale. For this reason, for easier distinction, the events were classified according to EURO-MED, FRANCE and an additional FRANCE_20 geographical region (see Fig. 7, bottom), corresponding to a zone extended up to 20 km beyond the borders of metropolitan France (similar to the definition used in the SI-Hex catalogue, Cara *et al.* 2015). Table 8 provides the number of events for each of the three geographic areas and for the different categories. Fig. 7 displays the geographic distributions of all earthquakes and Fig. 8 provides the temporal limits of each subset.

For the FRANCE region, unified M_w values are provided for 185 earthquakes (Table 8) with magnitude ranging from M_w 2.36 to 5.35. The majority are located in the southwest and southeast of France (Fig. 7, bottom panel), the most seismically active regions, and also in the neighboring countries for which the M_w computed from regional approaches are collected (such as Switzerland, Italy and Spain). On the contrary, only five earthquakes (including four with unified M_w estimates) are currently associated to a direct M_w value in the western part of France. As shown in Fig. 8, it was only from the late 1990s that our data set for the FRANCE_20 area began to include M_w estimates in a more systematic way, which is also concomitant with the development of the French accelerometric network (Péquegnat *et al.* 2008).

Concerning the EURO-MED_clip region, the methodology consists of collecting all M_w estimates from categories I and II but for categories III and IV, only sources concerning the FRANCE region were collected. This data set contains mainly earthquakes with magnitude ranging from M_w 3.06 to 7.57 located in Italy, Greece, Turkey and at the Iberian-Maghreb boundary (Fig. 7, top panel). Most earthquakes are characterized in category I (CMT services) since 1976 (Fig. 8, top panel). From 2010, some M_w estimates from category II are provided by other global networks, particularly GEOFON (Table S4). Concerning the category III, M_w estimates are mainly provided since 1995 by the Italian, Swiss and Spanish regional networks with also a contribution coming from *Mwr* NEIC and SismoAzur for Greek and Turkish events.

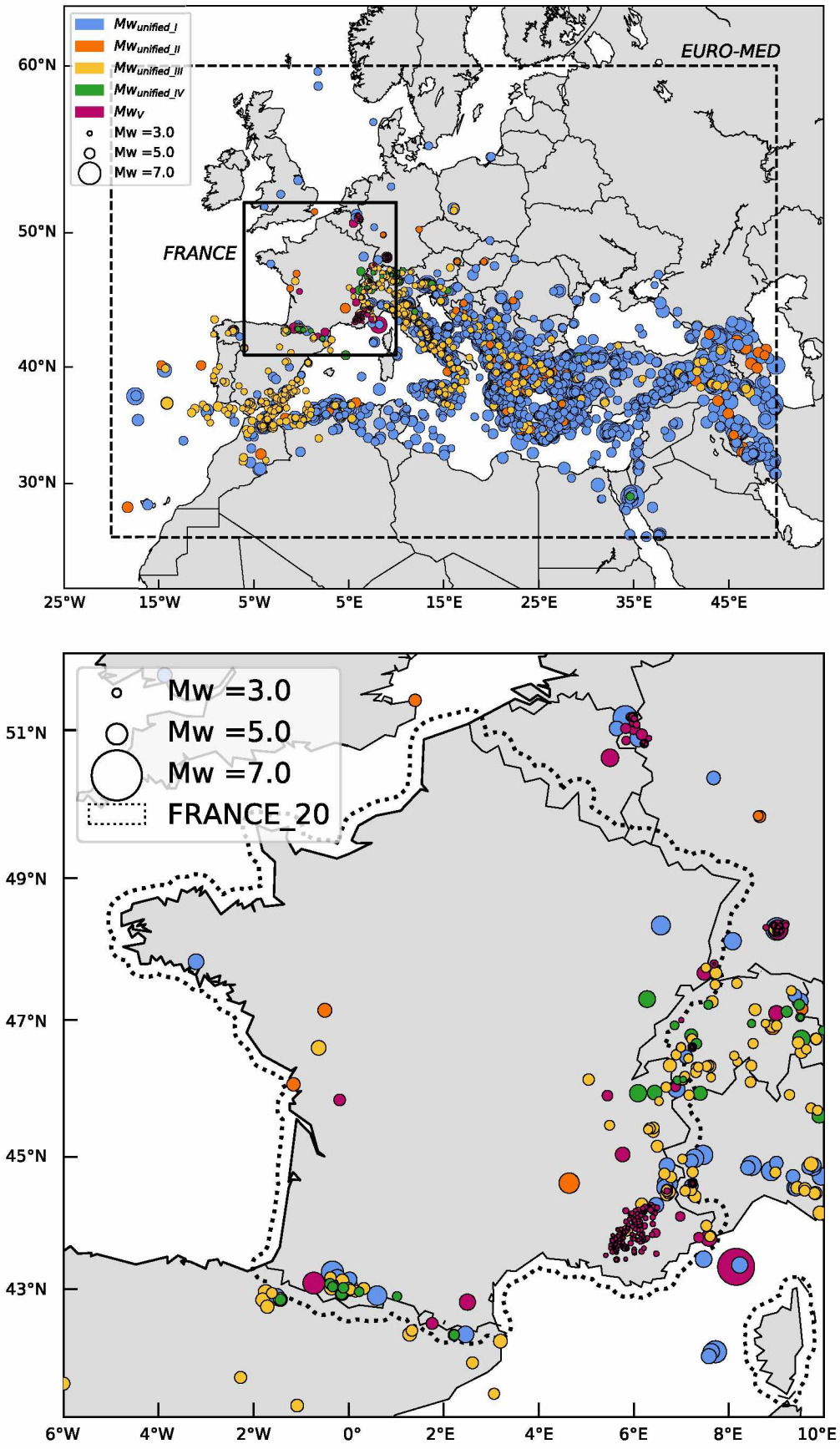

Figure 7. Location maps of the seismic events included in our resulting data set represented according to the categories of the M_w sources for EURO-MED, FRANCE (top panel) and FRANCE_20 (bottom panel). The colours indicate the category of the M_w origin (Mw_V values represented in purple are not unified).

Table 8. Number of earthquakes with an M_w estimate from the five categories of sources shown separately for earthquakes located in a zone extended up to 20 km beyond the borders of metropolitan France (FRANCE_20), located in the FRANCE_clip region and then in the EURO-MED_clip region.

Category	FRANCE_20	FRANCE_clip	EURO-MED_clip
I	14	29	2336
II	3	5	181
III	47	56	635
IV	18	13	14
V	247	856	0
V ($M_w \geq 2.5$)	16	49	0
Total	**329**	**959**	**3166**
Range of M_w	0.12–5.14	0.63–6.30	3.06–7.57
Number *(per cent)* of earthquakes with an unified M_w value	**82** (25)	**103** (11)	**3166** (100)
Range of unified M_w	2.36–5.14	2.36–5.35	3.06–7.57

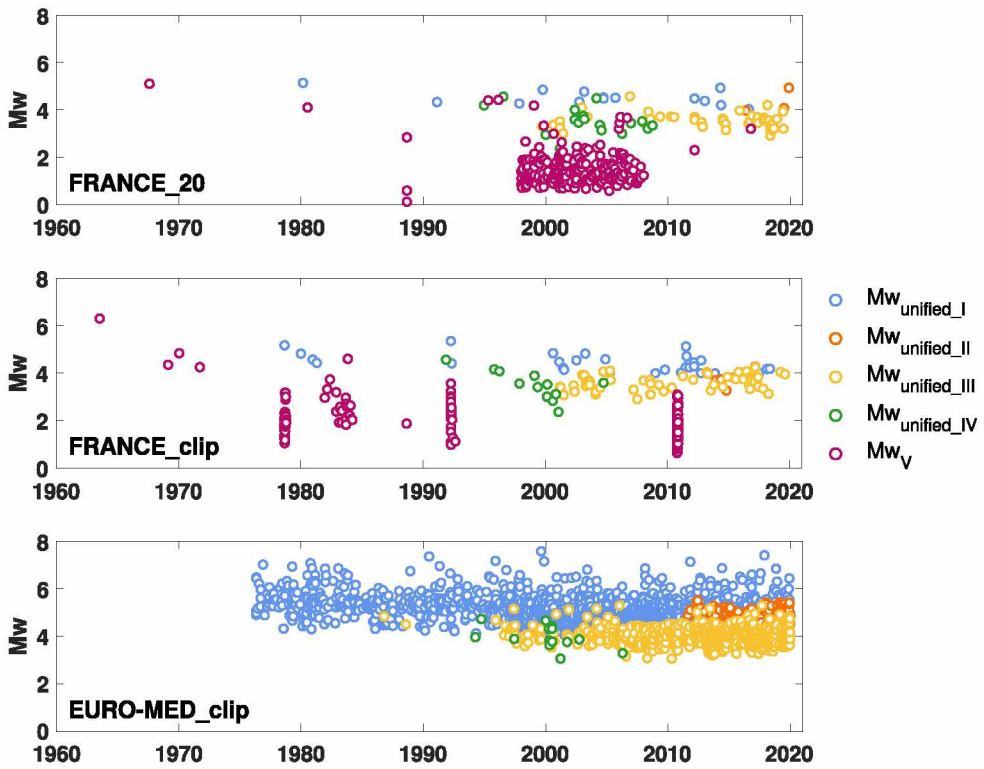

Figure 8. Temporal distribution of M_w estimates for earthquakes represented for the three geographic zones; the colours indicate the different categories of the M_w estimates.

Conversely, very few M_w values were collected for the northeastern part of Europe characterized by a low-to-moderate seismic activity, such as metropolitan France.

All data sources used here were last accessed on July 2020, therefore the latest *Mwr* IAG results, or the *Mwc* RCMT published after 31/01/2019, are not included. Consequently, the data set that we provided for $M_w > 5.0$ includes events up to 31/01/2019, for $3.5 < M_w < 5.0$, up to 19/08/2014 and only for the area on the border with Spain.

4.2 Comparison with existing SHA data sets

In the following, we performed a series of M_w intercomparisons with previously existing SHA data sets to discuss the impact of the strategy used to define M_w estimates. We also aim to underline where we believe future studies are needed to further improve SHA data sets for Europe, in general, and for metropolitan France, in particular.

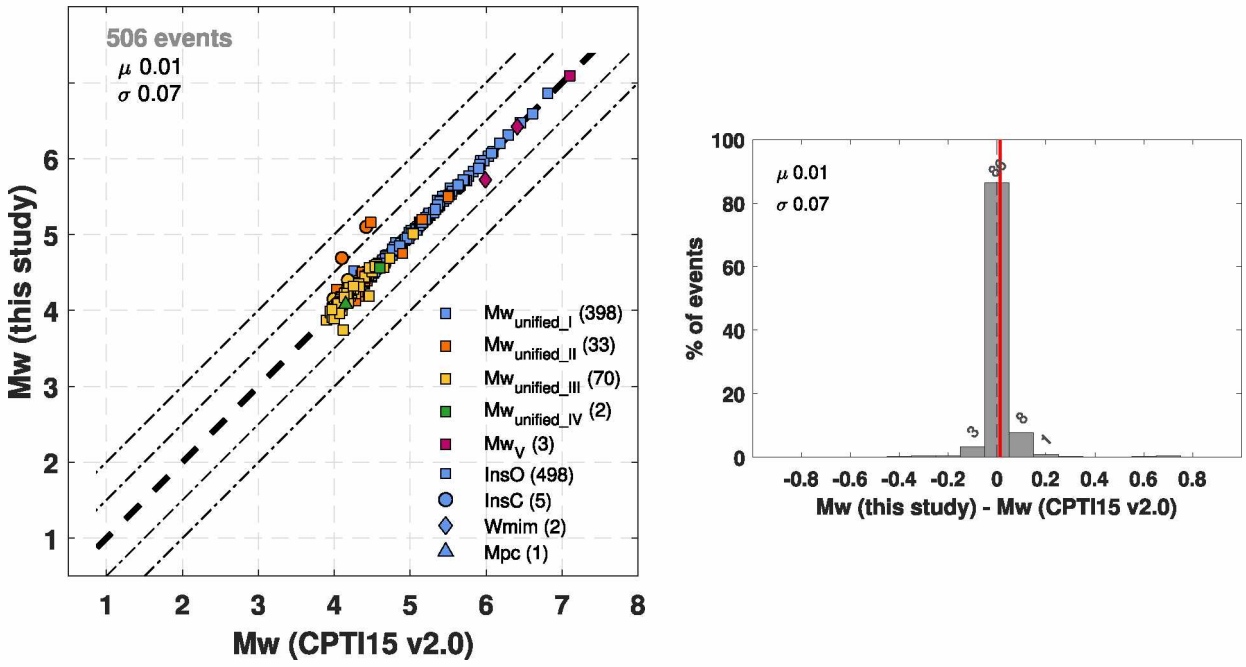

Figure 9. Intercomparison of the M_w estimates included in our data set, M_w (this study), and the Italian catalogue (CPTI15 v.2.0). M_w values (left-hand panel) and histograms of the differences between the two M_w values (right-hand panel). The average (μ) and the standard deviation (σ) of the differences (represented by a red line at right) are given. For CPTI15, the origin of the M_w estimate method is represented by various markers (InsO—M_w is from moment tensor solutions (direct); InsC—M_w proxy; Wmim—M_w is the mean of instrumental and macroseismic determination; Mpc—M_w is from parametric catalogue of macroseismic determinations). The number of data represented by a coloured marker is indicated in parentheses. The bold dotted line represents the case for which M_w (X) = M_w (Y) and thin dotted lines deviations spaced by 0.5 unit of M_w.

4.2.1 Comparing M_w estimates with the Italian earthquake catalogue

The average differences (μ) and the associated standard deviations (σ) of the comparison between M_w estimates from this study and from the Italian CPTI15 v2.0 catalogue (Rovida *et al.* 2019) are close to zero ($\mu = 0.01$ and $\sigma = 0.07$, Fig. 9) (even over the entire M_w range). We observe a slight increase in dispersion for the lowest magnitudes ($M_w \leq 4.6$) which concerns mainly earthquakes with M_w estimates from category III. This confirms that applying similar procedures will improve consistency between earthquake catalogues across borders.

The observed M_w value differences may be related to the strategy applied to assign a unique M_w value to an event. Indeed, in the Italian catalogue, a weighted average is computed with the available M_w estimates (maximum of five sources) while these same sources can be ranked into different categories in our procedure. Secondly, the M_w intercomparisons between sources, such as in Table 6, were carried out on slightly different data sets (in terms of input data selection criteria, such as the quality, the depth and the period covered) and slight differences are observed in the offset corrections (*Mwb* NEIC: +0.05 for Gasperini *et al.* (2012)/+0.09 for this study; *Mwr* INGV-TDMT: +0.213/+0.24; *Mwr* of SED specific studies: −0.05/−0.04).

It should also be pointed out that for few earthquakes, discrepancies between the M_w values can be larger than 0.5. Further inspection showed that: (i) for two earthquakes (13/03/2014 and 30/10/2016), M_w were estimated from another magnitude scale (flag InsC) in the case of the Italian catalogue and (ii) for the 30/10/2016 event (n°3885), CPTI15 v2.0 uses *Mwr* INGV-TDMT corrected at 4.48 and we use *Mwx* GEOFON corrected at 5.16, which gives a large discrepancy around 0.68 unit of M_w. In the new Homogenized Instrumental Seismic Catalogue (HORUS) of Italy (Lolli *et al.* 2020), this event is now associated to an M_w value of 4.79, corresponding to the weighted average of the corrected *Mwx* GEOFON and *Mwr* INGV-TDMT, reducing the discrepancy to 0.37 compared to our M_w estimate. This example illustrates well the importance of the choice of ranking of M_w sources.

4.2.2 Comparing M_w estimates with the French earthquake catalogue

Only 36 unified M_w estimates with $M_w > 3.0$ could be compared with the M_w of the SI-Hex catalogue (Cara *et al.* 2015, 2017) (Fig. 10). SI-Hex M_w estimates are on average lower ($\mu = +0.11$) and relatively dispersed ($\sigma = 0.18$) compared to our unified M_w. This average difference is also observed for the nine M_w estimates from the category I ($\mu = +0.12$) and is stronger for the $Mw_{unified\_III}$ ($\mu = +0.18$). In the SI-Hex catalogue, most earthquakes with $M_w > 3.0$ are described by an M_w value estimated from Mcoda, a magnitude defined from the inversion of the amplitude of the coda wave envelope (Denieul *et al.* 2015), then calibrated to M_w using as a reference 'direct' M_w values from different sources. In addition, when the extended version of the SI-Hex catalogue (obtained on request from authors) is used (see Supporting Information S6), the M_w values are even lower on average and very dispersed ($\mu = +0.36$ and $\sigma = 0.34$). This extended version of the SI-Hex catalogue is used in the recent French PSHA study of Drouet *et al.* (2020).

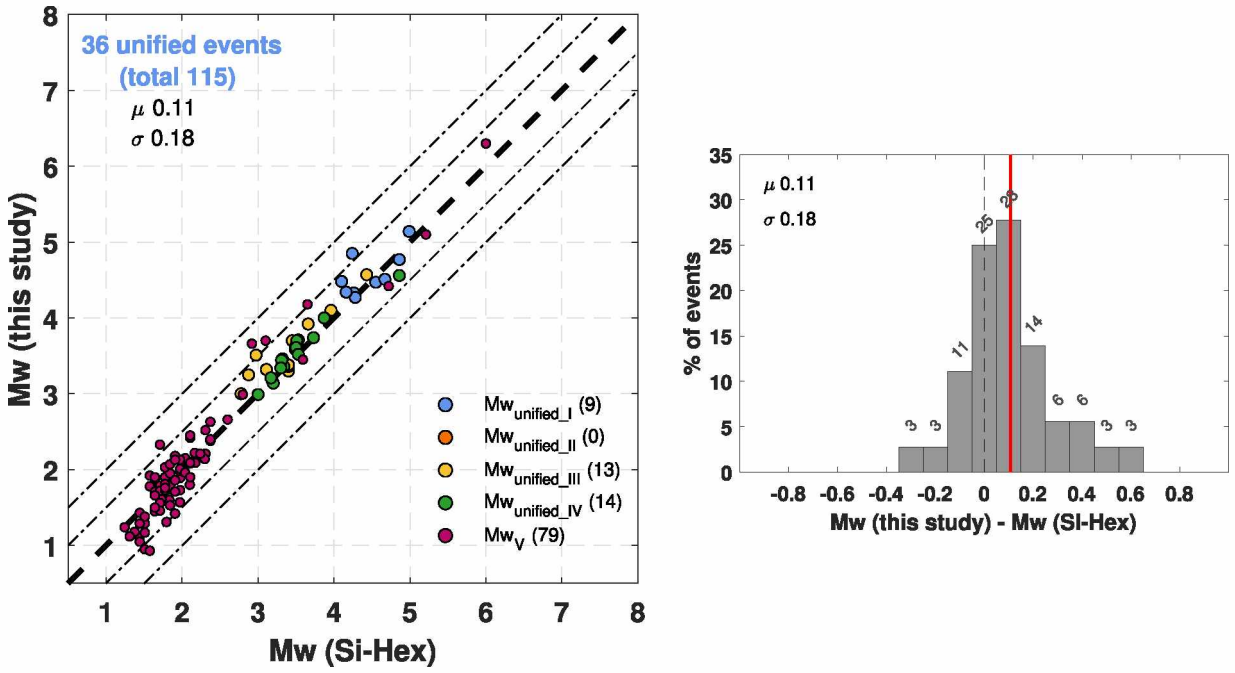

Figure 10. Intercomparisons of the M_w values for earthquakes included in our data set, M_w (this study), and the SI-Hex catalogue. M_w values (left-hand panel) and histograms show the differences between the two M_w values (right-hand panel). The average (μ) and the standard deviation (σ) of the differences (represented by a red line at right) are given only for the 36 earthquakes with a unified M_w estimate. The number of data represented by a coloured circle is indicated in parentheses. The bold dotted line represents the case for which M_w (X) = M_w (Y) and the other thin dotted lines deviations spaced by 0.5 unit of M_w.

For M_w SI-Hex < 3.0, our data set is only composed of M_w values from category V, that is to say not unified M_w values. These M_w values are mainly estimated from the local network around the Durance Fault (Nechtschein 2003). We observe differences, and probably a scaling bias, between these M_w values and the ones from SI-Hex, which have been estimated for a majority after applying a double magnitude scale correction law (Cara *et al.* 2017), from a regional network ML to ML_{LDG} and then to M_w, using a magnitude scale conversion law based on Godano *et al.* (2013)'s M_w data set. This observed scaling bias could be also due to the application of the OLS regression technique used to define the magnitude scale conversion law, as pointed out by Laurendeau *et al.* (2019).

4.2.3 Comparing unified M_w estimates with those used in strong ground motion data sets

We compared our unified M_w estimates with the ones currently available in the strong ground motion databases both for Europe and France (Fig. 11).

For Europe, we analysed the recently published ESM data set (Bindi *et al.* 2018; Lanzano *et al.* 2019) which contains 2179 events located in a larger region than ours (26°W–68.5°E; 23.5°N–72°N). In the ESM data set 61 per cent of the earthquakes are associated to an M_w estimate (defined following a similar strategy to EMEC, hereafter ESM-EMEC data set). We found a subset of 890 events in common with our data set (Fig. 11a). For metropolitan France, we analysed the recently published Résif-Epos ground motion data set (Traversa *et al.* 2020). In this data set, different strategies were used to define M_w estimates but all are proxy M_w. We found 89 out of 468 events have a unified M_w value (Fig. 11b).

For M_w > 5.0, only covered by ESM-EMEC, ESM-EMEC and our unified M_w estimates are similar for most events. Indeed, a majority of these M_w estimates are defined by sources similar to the ones used in this study, such as RCMT. For M_w < 5.0, M_w estimates can differ from our data set by up to ±0.4 and ±0.8 for ESM-EMEC and the French data sets, respectively. For ESM-EMEC, it seems that the largest discrepancies are when EMEC M_w are estimated from proxies. These direct M_w values were probably not available when estimating these EMEC M_w. Differences between direct and proxy M_w may also explain the larger differences observed in the case of the French data set. In addition, the M_w values of Traversa *et al.* (2020) data set have a slight tendency to be lower compared to the M_w estimates from our study (μ = +0.05), as already observed for the SI-Hex catalogue (Fig. 10).

4.2.4 Discussion about the strategies to build M_w data sets

These intercomparisons highlight that the strategy for defining M_w has a strong impact on M_w estimates provided in SHA data sets. Indeed, there is a good agreement of the M_w values when the strategies are similar (case of CPTI15 v2.0) and a significant dispersion when different

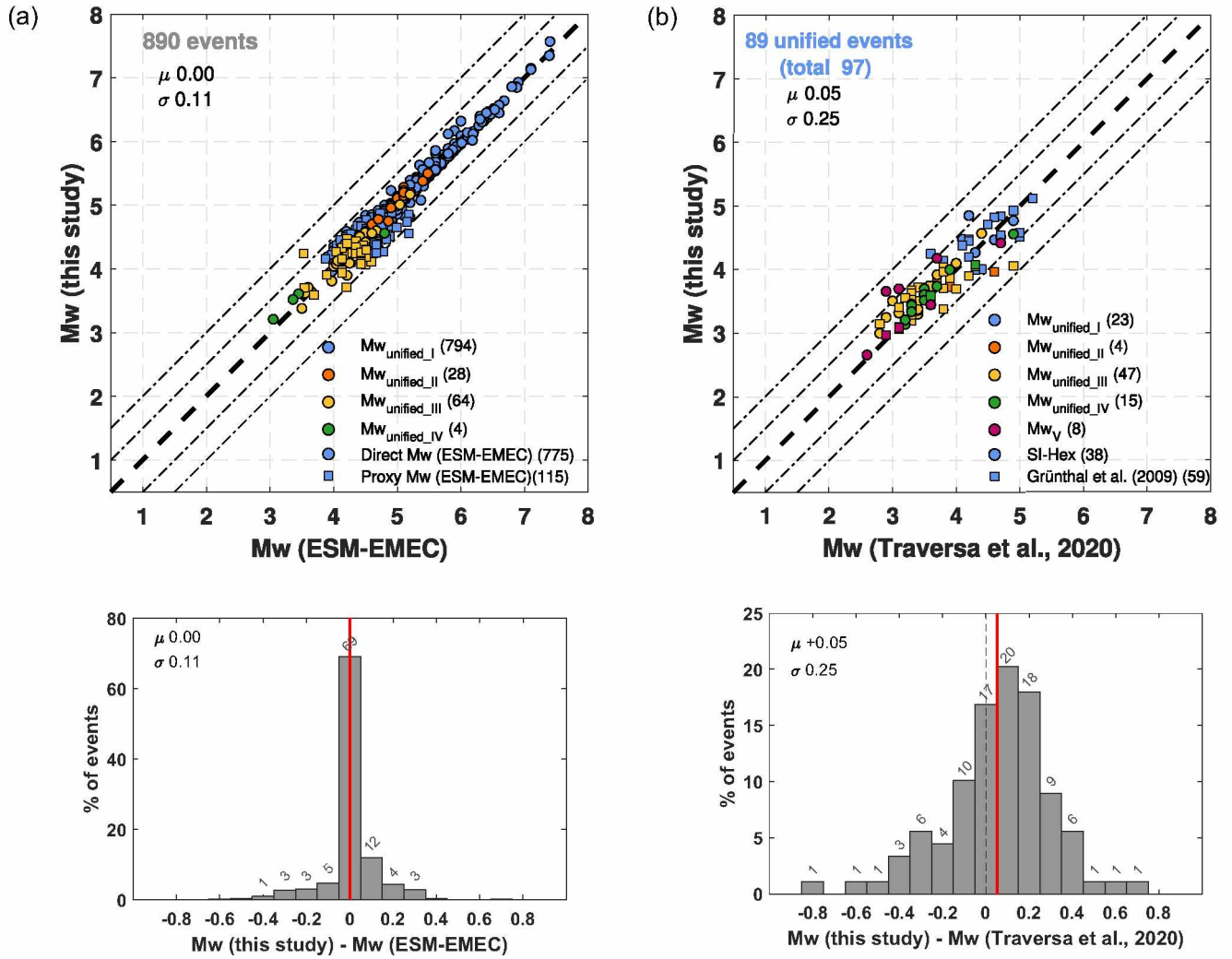

Figure 11. Intercomparisons of the M_w values for earthquakes included in our data set, Mw (this study), and common events found (a) in the ESM-EMEC data set (Bindi *et al.* 2018; Lanzano *et al.* 2019) and (b) in the Résif-Epos ground motion data set (Traversa *et al.* 2020) for which a marker specifies the original procedure used to define proxy M_w. M_w values (top panels) and histograms of the differences between the two M_w values with the red line representing the mean of the difference between the M_w values of the two data sets (bottom panels).

strategies are used (ESM-EMEC, SI-Hex, Traversa *et al.* 2020, data sets). This dispersion is particularly important when the comparison is between direct and proxy M_w estimates, and also for $M_w < 5.0$ events, for which M_w from different sources can be used.

In this study, we have chosen to apply M_w correction laws to the direct M_w estimates in order to unify them considering the same source chosen as a reference. However, the differences observed between data sets seem to be more related to the choice of the M_w source: the average difference between two data sets is small compared to the offset/scaling corrections defined in Table 6. For example, in the case of the ESM-EMEC data set, there is no observed difference on average, the M_w corrections that were applied are not reflected in the average difference (for categories II + III + IV).

In our strategy, the ranking of sources (Fig. 5) can therefore be the factor that has the most impact on the final M_w estimate. We have favoured the M_w of the CMT services, whereas for example in the case of ESM-EMEC, the specific studies have priority. Our choice was guided by the need to have homogeneous M_w at the largest spatio-temporal scales possible and compatible with the majority of the metadata in the GMMs However, M_w estimates from the regional methods are more sensitive to the density and spatial distribution of stations, to the velocity model used (e.g. Scognamiglio *et al.* 2016) which can lead to larger differences in the M_w estimates as observed for some specific earthquakes, such as the 2005 September 8 Vallorcine earthquake (Table 4), and to the systematic offsets as observed by various M_w correction laws (Table 6) for these regional methods. Metropolitan France may be particularly concerned by these limitations: indeed, the heterogeneous spatial distribution of stations is imposed by the presence of numerous coastlines and a lower amount of stations in the less seismically areas. Furthermore, the small quantity of M_w values estimated from specific studies would not allow the implementation of an M_w unification strategy with a unique source as reference.

The table called 'M_w_sources_used' of the Supporting Information S4 provides the M_w values from the different sources, and end-users are free to make their own choices. It is important to encourage the community to pursue the effort of homogenizing M_w estimates at large scales in order to better understand the origin of the inconsistencies and to try to avoid them.

4.3 Traceability, reproducibility and quality of the M_w collections

The establishment of this unified M_w data set from various and heterogeneous M_w collections highlights some difficulties related to its traceability, reproducibility and quality. We intend to stress these points here, as it would also help in the selection or not of an M_w estimate by a source in an automatic way.

4.3.1 Traceability of M_w results over time

(i) A first difficulty is that the original results used in a study are no longer accessible online. This is already the case for the SismoAzur collection used in this study but also for M_w collections used in the CPTI15 (Gasperini *et al.* 2012) and the HORUS (Lolli *et al.* 2020) catalogues for ETHZ and NEIC data sets.

(ii) A second difficulty arises from the evolution over time of the approaches used to compute or select M_w which are not always easily traceable for end-users. For example, in the case of ETHZ a series of authors computed M_w using different approaches. For SismoAzur, previously published results were withdrawn from online publication following different choices of focal mechanism quality criteria.

(iii) A third difficulty arises from updates of the M_w values after performing a new analysis without tracing the date of the last modification or of the previously published M_w values. For example, the 2010/08/31 event (n° 2120) for which our M_w value is different from the one in CPTI15 while the same source *Mwr* INGV-TDMT was used. Therefore, it seems important to update the original collections each time a new data set is compiled.

(iv) It should also be pointed out that evolutions in data formats, such as the entries of the quakeml files in the case of NEIC, require routinely checking and adapting the data extraction codes to the new formats.

4.3.2 Reproducibility of results to define a 'standard' M_w estimate

As discussed in Section 2.3, for many sources (mainly the oldest ones), the level of information associated to an M_w estimate does not allow the M_0 and M_w values to be reproduced either because they are not provided by their authors or they are not provided with sufficient accuracy (see Supporting Information S2). When M_w are mixed together in the same data set, it seems important to define them with the same standards. The different authors do not necessarily use the GCMT standards and thus it is necessary to recompute homogeneously the M_w values, which is not always possible. This lack of information can induce discrepancies (smaller than those between sources) but which remain notable (larger than 0.1 for some cases).

4.3.3 Quality of the M_w solutions

In this study, we focused on the mean M_w values to characterize the seismic events. However, the question of the uncertainty values associated to the M_w estimates is a paramount question for the development of magnitude scale conversion laws (Castellaro & Bormann 2007; Gutdeutsch *et al.* 2011), or to take into account uncertainties in a PSHA study (e.g. Drouet *et al.* 2020). In this part, we discuss only the difficulties for end-users to appreciate the quality of an M_w estimate.

In fact, the sources which estimate M_w do not necessarily report an associated error and it is not easy to define *a posteriori* a quality with the information provided by their authors. To give an indication about the quality of one moment tensor solution, some authors have added letters depending on different criteria (e.g. Pondrelli *et al.* 2006; Scognamiglio *et al.* 2009): the number of data analysed, the azimuthal coverage of the recordings, the quality of the fit between the synthesized and the observed data (VR, the centroid location, . . .), the moment tensor double-couple (DC) percentage. From one study to another, the criteria for defining the reliability of one solution are not necessarily the same. That is why in this study, we relied on the opinions of the authors to consider or not a moment tensor solution. However, for a large number of M_w collected in this study (∼40 per cent), information on the reliability of the solution was not provided, and thus in this case all solutions were considered reliable. Depending on the parametric studies and the methodologies used, uncertainty errors can range from 0.07 (e.g. Gasperini *et al.* 2012) to 0.8 (Stork *et al.* 2014). It would be interesting in the future to define international standards for quality criteria associated to each M_w estimate to be able to work more easily with different M_w estimates from different sources. This would allow assigning an error value to the M_w estimates consistent between the different sources.

5 CONCLUSIONS AND PERSPECTIVES

We propose a novel strategy to build unique and unified M_w reference data sets based on (i) intercomparisons of M_w estimates at the European–Mediterranean scale, (ii) ranking sources of M_w in five levels from global to specific studies and (iii) using a M_w value threshold that are useful for low-to-moderate seismicity regions such as metropolitan France, the target region in our study.

We collected 6752 direct M_w estimates from 34 sources (from webservices to specific published studies) for 4454 shallow seismic events (depth < 40 km) that occurred in the European–Mediterranean area (20°W–50°E; 25°N–60°N) between 1963 and 2019. We recomputed an M_w value following the GCMT standards when it was possible and attributed quality flags depending on the level of the information provided by the sources.

The proposed unification strategy allowed assigning to 185 events ($2.36 \leq M_w \leq 5.35$) direct, unique and unified M_w value using the CMT services as reference for the FRANCE region (including neighbouring countries). For the European–Mediterranean region, 3351 events ($3.06 \leq M_w \leq 7.57$) have a unified M_w value. This strategy could also be extended by collecting regional and specific M_w studies in other European countries in order to have a more complete unified M_w data set for Europe.

We observed comparable M_w values on average with the European SHA data sets (the CPTI15 v2.0 and ESM-EMEC data sets) and slightly larger M_w values compared to the French SHA data sets (the SI-Hex and Traversa *et al.* 2020, data sets). For all these data sets with which we compared our unified M_w values, we noted an increase of the discrepancies for $M_w < 5.0$, especially when different strategies were used. We attributed this variability to differences in the choice of M_w source, reflecting the wide variability of M_w estimates between sources for $M_w < 5.0$.

Despite all our efforts, the unified M_w estimates remains limited in number of events characterized, in their spatio-temporal distribution and in the magnitude range. To reduce the dependence on the need for M_w proxies in SHA data sets, future efforts will need to focus on estimating 'direct' M_w for a larger number of earthquakes and reducing differences between sources for $M_w < 5.0$ earthquakes, an M_w range that concerns metropolitan France in particular.

Our data set aims to be upgradeable over time, by integrating new M_w estimates and then applying again the hybrid strategy. For example, the recordings used to estimate the M_w of category V (not unified today), could be used again to estimate M_w from a unique method in order to unify them.

Our unified M_w data set will certainly already improve the M_w estimates of the SHA data sets (instrumental catalogues and ground motion data sets) not only in France but also in Europe: we have unified M_w estimates for 1/5 of the Traversa *et al.* (2020) data set and 2/5 of the ESM-EMEC data set. This unified M_w data set could be also used directly as metadata to describe the events included in the recently published database of earthquake focal mechanisms in metropolitan France and conterminous Western Europe, FMHEX20 (Mazzotti *et al.* 2021). Furthermore, the impact of such unification strategy on the interevent variability of GMMs will have to be assessed.

Finally, in this work, we shared our feedback on the traceability, reproducibility and quality of the M_w collections in the requested databases. Efforts should be pursued in this direction in order to define international standards, to facilitate the choice of the best M_w estimate for the end-users.

ACKNOWLEDGMENTS

This research is funded by IRSN (30001295–80). AL is the principal investigator for the research presented in the manuscript. CC initiated the work on earthquake catalogues and his experience benefited to this analysis. OS and CC followed and provided advice for the good progress of the work. OS helped synthesizing the ideas for the paper. AL and OS did the revision of the paper. The authors would like to thank the researchers who helped us with their M_w databases such as B. Delouis, M. Godano and others. We would like to thank B. Delouis and C. Satriano for the discussions about the methods. This work has also benefited from discussions within the framework of Résif, and especially the Seismicity Transverse Action and the French Accelerometric Network (RAP). We would also like to thank very much the two reviewers (Rashad Sawires and an anonymous reviewer) and the editor Victor Cruz-Atienza for their careful reading and comments which allowed the text to be restructured, simplified and the discussions to be refocused.

DATA AVAILABILITY

The data used in this paper are available in its online supplementary material.

REFERENCES

Aki, K., 1966. Generation and propagation of G waves from the Niigata earthquake of June 16, 1964. Part 2. Estimation of earthquake moment, released energy, and stress-strain drop from the G wave spectrum, *Bull. Earthq. Res. Inst.,* **44,** 73–88.

Baer, M. *et al.*, 2007. Earthquakes in Switzerland and surrounding regions during 2006, *Swiss J. Geosci.,* **100,** 517–528.

Beauval, C., Bard, P.-Y. & Danciu, L., 2020. The influence of source- and ground-motion model choices on probabilistic seismic hazard levels at 6 sites in France, *Bull. Earthq. Eng.,* **18,** 4551–4580.

Bernardi, F., 2005. Seismic moment from regional surface-wave amplitudes: applications to digital and analog seismograms, *Bull. seism. Soc. Am.,* **95,** 408–418, doi.org/10.1785/0120040048.

Bethmann, F., Deichmann, N. & Mai, P.M., 2011. Scaling relations of local magnitude versus moment magnitude for sequences of similar earthquakes in Switzerland, *Bull. seism. Soc. Am.,* **101,** 515–534.

Bindi, D., Kotha, S.-R., Weatherill, G., Lanzano, G., Luzi, L. & Cotton, F., 2018. The pan-European engineering strong motion (ESM) flatfile: consistency check via residual analysis, *Bull. Earthq. Eng.,* **17,** 583–602.

Braunmiller, J., Kradolfer, U., Baer, M. & Giardini, D., 2002. Regional moment tensor determination in the European–Mediterranean area—initial results, *Tectonophysics,* **356,** 5–22.

Braunmiller, J., Deichmann, N., Giardini, D. & Wiemer, S., & SED Magnitude Working Group, 2005. Homogeneous moment-magnitude calibration in Switzerland, *Bull. seism. Soc. Am.,* **95,** 58–74.

Cara, M. *et al.*, 2015. SI-Hex: a new catalogue of instrumental seismicity for metropolitan France, *Bull. Soc. Geol. Fr.,* **186,** 3–19.

Cara, M., Denieul, M., Sèbe, O., Delouis, B., Cansi, Y. & Schlupp, A., 2017. Magnitude Mw in metropolitan France, *J. Seismol.,* **21,** 551–565.

Castellaro, S. & Bormann, P., 2007. Performance of different regression procedures on the magnitude conversion problem, *Bull. seism. Soc. Am.,* **97,** 1167–1175.

Chevrot, S., Sylvander, M. & Delouis, B., 2011. A preliminary catalog of moment tensors for the Pyrenees, *Tectonophysics,* **510,** 239–251.

Clinton, J.F., Hauksson, E. & Solanki, K., 2006. An evaluation of the SCSN moment tensor solutions: robustness of the Mw magnitude scale, style of faulting, and automation of the method, *Bull. seism. Soc. Am.,* **96,** 1689–1705.

Courboulex, F. *et al.*, 1998. Source study and tectonic implications of the 1995 Ventimiglia (border of Italy and France) earthquake (ML = 4.7), *Tectonophysics,* **290,** 245–257.

Courboulex, F., Deichmann, N. & Gariel, J.-C., 1999. Rupture complexity of a moderate intraplate earthquake in the Alps: the 1996M5 Epagny-Annecy earthquake, *Geophys. J. Int.,* **139,** 152–160.

Courboulex, F., Duval, A.-M., Deschamps, A., Lomax, A. & Larroque, C., 2001. Les enseignements du petit séisme de Peille (Alpes-Maritimes, France), *Comptes Rendus Académie Sci. - Ser. IIA - Earth Planet. Sci.,* **333,** 105–112.

Courboulex, F., Dujardin, A., Vallée, M., Delouis, B., Sira, C., Deschamps, A., Honoré, L. & Thouvenot, F., 2013. High-frequency directivity effect for an Mw 4.1 earthquake, widely felt by the population in southeastern France. Short Note, *Bull. seism. Soc. Am.,* **103,** 3347–3353.

Cushing, E.M. *et al.*, 2008. A multidisciplinary study of a slow-slipping fault for seismic hazard assessment: the example of the Middle Durance Fault (SE France), *Geophys. J. Int.,* **172,** 1163–1178.

Delouis, B., 2014. FMNEAR: determination of focal mechanism and first estimate of rupture directivity using near-source records and a linear distribution of point sources, *Bull. seism. Soc. Am.,* **104,** 1479–1500.

Delouis, B., Charlety, J. & Vallee, M., 2009. A method for rapid determination of moment magnitude Mw for moderate to large earthquakes from the near-field spectra of strong-motion records (MWSYNTH), *Bull. seism. Soc. Am.,* **99,** 1827–1840.

Denieul, M., Sèbe, O., Cara, M. & Cansi, Y., 2015. M_w estimation from crustal coda waves recorded on analog seismograms, *Bull. seism. Soc. Am.,* **105,** 831–849.

Di Giacomo, D., Bondár, I., Storchak, D.A., Engdahl, E.R., Bormann, P. & Harris, J., 2015. ISC-GEM: global Instrumental Earthquake Catalogue (1900–2009), III. Re-computed MS and mb, proxy MW, final magnitude composition and completeness assessment, *Phys. Earth planet. Inter.,* **239,** 33–47.

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

Diehl, T. *et al.*, 2018. Earthquakes in Switzerland and surrounding regions during 2015 and 2016, *Swiss J. Geosci.,* **111,** 221–244.

Dreger, D.S., 2003. 85.11 TDMT_INV: time domain seismic moment tensor INVersion, *Int. Geophys.,* **81,** 1627, doi:10.1016/S0074-6142(03)80290-5.

Dreger, D.S. & Helmberger, D.V., 1993. Determination of source parameters at regional distances with three-component sparse network data, *J. geophys. Res.,* **98,** 8107–8125.

Drouet, S., Ameri, G., Dortz, K.L., Secanell, R. & Senfaute, G., 2020. A probabilistic seismic hazard map for the metropolitan France, *Bull. Earthq. Eng.,* **18,** 1865–1898.

Dufumier, H., 2002. Synthesis of magnitude and focal mechanism computations for the M ≥ 4.5 earthquakes in France for the period 1995–2000, *J. Seismol.,* **6**(2), 163–181.

Dufumier, H., Souriau, A., Sylvander, M., Judenherc, S. & Granet, M., 2000. Magnitude and focal mechanism computations for the Saint-Beat earthquake of 4 October 1999, *Comptes Rendus Acad. Sci. - Ser. IIa Sci. Terre Planetes,* **331,** 331–338.

Duverger, C., Mazet-Roux, G., Bollinger, L., Trilla, A.G., Vallage, A., Hernandez, B. & Cansi, Y., 2021. A decade of seismicity in metropolitan France (2010–2019): the CEA/LDG methodologies and observationsUne décennie de sismicité en France métropolitaine (2010–2019) : les méthodes et observations du CEA/LDG, *Bull. Société Géologique Fr.,* **192**(1),.

Dziewonski, A.M., Chou, T.-A. & Woodhouse, J.H., 1981. Determination of earthquake source parameters from waveform data for studies of global and regional seismicity, *J. geophys. Res.,* **86,** 2825–2852.

Ekström, G., Morelli, A., Boschi, E. & Dziewonski, A.M., 1998. Moment tensor analysis of the Central Italy Earthquake Sequence of September–October 1997, *Geophys. Res. Lett.,* **25,** 1971–1974.

Ekström, G., Nettles, M. & Dziewoński, A.M., 2012. The global CMT project 2004–2010: centroid-moment tensors for 13,017 earthquakes, *Phys. Earth planet. Inter.,* **200-201,** 1–9.

Fäh, D. *et al.*, 2011. ECOS-09 Earthquake Catalogue of Switzerland Release 2011, Report and Database. Public catalogue, 17.4. 2011 (No. SED/RISK/R/001/20110417), p. 42, Report SED/RISK.

Gasperini, P., Lolli, B., Vannucci, G. & Boschi, E., 2012. A comparison of moment magnitude estimates for the European—Mediterranean and Italian regions, *Geophys. J. Int.,* **190,** 1733–1745.

Gasperini, P., Lolli, B. & Vannucci, G., 2013. Empirical calibration of local magnitude data sets versus moment magnitude in Italy, *Bull. seism. Soc. Am.,* **103,** 2227–2246.

GEOFON Data Centre, 1993. GEOFON Seismic Network, approximately 80 active stations; greater than 440 MB/day, Deutsches GeoForschungsZentrum GFZ, doi:10.14470/TR560404.

Godano, M. *et al.*, 2013. The October–November 2010 earthquake swarm near Sampeyre (Piedmont region, Italy): a complex multicluster sequence, *Tectonophysics,* **608,** 97–111.

Grünthal, G. & Wahlström, R., 2003. An earthquake catalogue for central, northern and northwestern Europe based on Mw magnitudes (Scientific Technical Report No. STR 03/02), p. 142, *Deutsches GeoForschungsZentrum GFZ.* doi:10.2312/GFZ.b103-030104.

Grünthal, G. & Wahlström, R., 2012. The European-Mediterranean Earthquake Catalogue (EMEC) for the last millennium, *J. Seismol.,* **16,** 535–570.

Grünthal, G., Stromeyer, D. & Wahlström, R., 2009. Harmonization check of Mw within the central, northern, and northwestern European earthquake catalogue (CENEC), *J. Seismol.,* **13,** 613–632.

Gutdeutsch, R., Castellaro, S. & Kaiser, D., 2011. The magnitude conversion problem: further insights, *Bull. seism. Soc. Am.,* **101,** 379–384.

Hanks, T.C. & Kanamori, H., 1979. A moment magnitude scale, *J. geophys. Res.,* **84,** 2348.

Hayes, G.P., Rivera, L. & Kanamori, H., 2009. Source Inversion of the W-phase: real-time implementation and extension to low magnitudes, *Seismol. Res. Lett.,* **80,** 817–822.

IASPEI, 2013. Summary of Magnitude Working Group recommendations on standard procedures for determining earthquake magnitudes from digital data, Retrieved from http://isc-mirror.iris.washington.edu/docs/papers/download/Summary_WG_recommendations_20130327.pdf.

Kagan, Y.Y., 2003. Accuracy of modern global earthquake catalogs, *Phys. Earth planet. Inter.,* **135,** 173–209.

Kanamori, H., 1977. The energy release in great earthquakes, *J. geophys. Res.,* **82,** 2981–2987.

Kanamori, H. & Rivera, L., 2008. Source inversion of W phase: speeding up seismic tsunami warning, *Geophys. J. Int.,* **175,** 222–238.

Kishida, T., Darragh, R.B., Bozorgnia, Y., Kuo, C.-H. & Si, H., 2018. Homogeneous estimation of moment magnitude for small-to-moderate magnitude earthquakes located near the border between Japan and Taiwan, *Seismol. Res. Lett.,* **89,** 1093–1100.

Konstantinou, K.I., 2015. Moment magnitude estimates for earthquakes in the Greek Region: a comprehensive comparison, *Bull. seism. Soc. Am.,* **105,** 2555–2562.

Kotha, S.R., Bindi, D. & Cotton, F., 2016. Partially non-ergodic region specific GMPE for Europe and Middle-East, *Bull. Earthq. Eng.,* **14,** 1245–1263.

Kotha, S.R., Weatherill, G., Bindi, D. & Cotton, F., 2020. A regionally-adaptable ground-motion model for shallow crustal earthquakes in Europe, *Bull. Earthq. Eng.,* **18,** 4091–4125.

Ktenidou, O.-J., Roumelioti, Z., Abrahamson, N., Cotton, F., Pitilakis, K. & Hollender, F., 2018. Understanding single-station ground motion variability and uncertainty (sigma): lessons learnt from EUROSEISTEST, *Bull. Earthq. Eng.,* **16,** 2311–2336.

Lancieri, M., Madariaga, R. & Bonilla, F., 2012. Spectral scaling of the aftershocks of the Tocopilla 2007 earthquake in northern Chile: scaling of Tocopilla aftershocks, *Geophys. J. Int.,* **189,** 469–480.

Lanzano, G. *et al.*, 2019. The pan-European Engineering Strong Motion (ESM) flatfile: compilation criteria and data statistics, *Bull. Earthq. Eng.,* **17,** 561–582.

Larroque, C., Delouis, B., Godel, B. & Nocquet, J.-M., 2009. Active deformation at the southwestern Alps–Ligurian basin junction (France–Italy boundary): evidence for recent change from compression to extension in the Argentera massif, *Tectonophysics,* **467,** 22–34.

Larroque, C., Delouis, B., Sage, F., Régnier, M., Béthoux, N., Courboulex, F. & Deschamps, A., 2016. The sequence of moderate-size earthquakes at the junction of the Ligurian basin and the Corsica margin (western Mediterranean): the initiation of an active deformation zone revealed?,*Tectonophysics,* **676,** 135–147.

Laurendeau, A., Cotton, F., Ktenidou, O.-J., Bonilla, L.-F. & Hollender, F., 2013. Rock and stiff-soil site amplification: dependency on VS30 and Kappa (0), *Bull. seism. Soc. Am.,* **103,** 3131–3148.

Laurendeau, A, Clément, C. & Scotti, O., 2019. Catalogue de sismicité instrumental unifié en Mw pour la France métropolitaine: la problématique des conversions de magnitude, in *Presented at the 10ème Colloque National de l'AFPS,* Strasbourg, France, **2,** p. 66–77.

Lee, W.H.K. & Engdahl, E.R., 2015. Bibliographical search for reliable seismic moments of large earthquakes during 1900–1979 to compute MW in the ISC-GEM Global Instrumental Reference Earthquake Catalogue, *Phys. Earth planet. Inter.,* **239,** 25–32.

Lolli, B. & Gasperini, P., 2012. A comparison among general orthogonal regression methods applied to earthquake magnitude conversions: earthquake magnitude regression methods, *Geophys. J. Int.,* **190,** 1135–1151.

Lolli, B., Randazzo, D., Vannucci, G. & Gasperini, P., 2020. The Homogenized Instrumental Seismic Catalog (HORUS) of Italy from 1960 to Present, *Seismol. Res. Lett.,* **XX,** 1–15.

Martín, R., Stich, D., Morales, J. & Mancilla, F., 2015. Moment tensor solutions for the Iberian-Maghreb region during the IberArray deployment (2009–2013), *Tectonophysics,* **663,** 261–274.

Mazzotti, S. *et al.*, 2021. FMHex20: an earthquake focal mechanism database for seismotectonic analyses in metropolitan France and bordering regions, *Bull. Société Géologique Fr.,* **192**(1), 10.

Nechtschein, S., 2003. Calcul d'une magnitude de moment Mw pour les séismes enregistrés par le réseau de la Durance (No. Rapport DEI /SARG / 03-01), p. 31, IRSN.

Nechtschein, S. & Lesueur, C., 2011. Calcul de la magnitude de moment de neuf séismes français récents à partir de la base de données sismiques du RAP (No. DEI/SARG/2011-004), p. 22, IRSN.

Péquegnat, C., Guéguen, P., Hatzfeld, D. & Langlais, M., 2008. The French Accelerometric Network (RAP) and National Data Centre (RAP-NDC), *Seismol. Res. Lett.,* **79,** 79–89.

Perrot, J. *et al.*, 2005. Analysis of the Mw 4.3 Lorient earthquake sequence: a multidisciplinary approach to the geodynamics of the Armorican Massif, westernmost France, *Geophys. J. Int.,* **162,** 935–950.

Pondrelli, S., 2002. European-Mediterranean Regional Centroid-Moment Tensors Catalog (RCMT), Istituto Nazionale di Geofisica e Vulcanologia (INGV), doi:10.13127/rcmt/euromed.

Pondrelli, S. & Salimbeni, S., 2006. Italian CMT Dataset, Istituto Nazionale di Geofisica e Vulcanologia (INGV), doi:10.13127/rcmt/italy.

Pondrelli, S., Morelli, A. & Ekström, G., 2004. European-Mediterranean regional centroid-moment tensor catalog: solutions for years 2001 and 2002, *Phys. Earth planet. Inter.,* **145,** 127–147.

Pondrelli, S., Salimbeni, S., Ekström, G., Morelli, A., Gasperini, P. & Vannucci, G., 2006. The Italian CMT dataset from 1977 to the present, *Phys. Earth planet. Inter.,* **159,** 286–303.

Pondrelli, S., Salimbeni, S., Morelli, A., Ekström, G. & Boschi, E., 2007. European–Mediterranean Regional Centroid Moment Tensor catalog: solutions for years 2003 and 2004, *Phys. Earth planet. Inter.,* **164,** 90–112.

Pondrelli, S., Salimbeni, S., Morelli, A., Ekström, G., Postpischl, L., Vannucci, G. & Boschi, E., 2011. European–Mediterranean Regional Centroid Moment Tensor catalog: solutions for 2005–2008, *Phys. Earth planet. Inter.,* **185,** 74–81.

Rovida, A., Locati, M., Camassi, R., Lolli, B. & Gasperini, P., 2016. Catalogo Parametrico dei Terremoti Italiani, versione CPTI15, Istituto Nazionale di Geofisica e Vulcanologia (INGV), doi:10.6092/INGV.IT-CPTI15.

Rovida, A., Locati, M., Camassi, R., Lolli, B. & Gasperini, P., 2019. Catalogo Parametrico dei Terremoti Italiani (CPTI15), versione 2.0, Istituto Nazionale di Geofisica e Vulcanologia (INGV), doi:10.13127/CPTI/CPTI15.2.

Rueda, J. & Mezcua, J., 2005. Near-real-time seismic moment-tensor determination in Spain, *Seismol. Res. Lett.,* **76,** 455–465.

Saul, J., Becker, J. & Hanka, W., 2011. Global moment tensor computation at GFZ Potsdam, in *Presented at the AGU 2011 Fall Meeting,* San Francisco, CA, Retrieved from: http://gfzpublic.gfz-potsdam.de/pubman/faces/viewItemOverviewPage.jsp?itemId=escidoc:244630:1.

Scherbaum, F. & Stoll, D., 1983. Source parameters and scaling lawsd of the 1978 Swabian Jura aftershocks, *Bull. seism. Soc. Am.,* **73,** 1321–1343.

Scognamiglio, L., Tinti, E. & Michelini, A., 2009. Real-time determination of seismic moment tensor for the Italian region, *Bull. seism. Soc. Am.,* **99,** 2223–2242.

Scognamiglio, L., Tinti, E. & Quintiliani, M., 2006. Time Domain Moment Tensor (TDMT), Istituto Nazionale di Geofisica e Vulcanologia (INGV), doi:10.13127/TDMT.

Scognamiglio, L., Magnoni, F., Tinti, E. & Casarotti, E., 2016. Uncertainty estimations for moment tensor inversions: the issue of the 2012 May 20 Emilia earthquake, *Geophys. J. Int.,* **206,** 792–806.

Silver, P.G. & Jordan, T.H., 1982. Optimal estimation of scalar seismic moment, *Geophys. J. Int.,* **70,** 755–787.

Sipkin, S.A., 1982. Estimation of earthquake source parameters by the inversion of waveform data: synthetic waveforms, *Phys. Earth planet. Inter.,* **30,** 242–259.

Sipkin, S.A., 1994. Rapid determination of global moment-tensor solutions, *Geophys. Res. Lett.,* **21,** 1667–1670.

Stich, D., Ammon, C.J. & Morales, J., 2003. Moment tensor solutions for small and moderate earthquakes in the Ibero-Maghreb region: Ibero-Maghrebian moment tensor solutions, *J. geophys. Res.,* **108,** doi:10.1029/2002JB002057.

Stich, D., Serpelloni, E., Lis Mancilla, F. de & Morales, J., 2006. Kinematics of the Iberia–Maghreb plate contact from seismic moment tensors and GPS observations, *Tectonophysics,* **426,** 295–317.

Stich, D., Martín, R. & Morales, J., 2010. Moment tensor inversion for Iberia–Maghreb earthquakes 2005–2008, *Tectonophysics,* **483,** 390–398.

Stork, A.L., Verdon, J.P. & Kendall, J.-M., 2014. The robustness of seismic moment and magnitudes estimated using spectral analysis: the robustness of seismic moment and magnitudes estimates, *Geophys. Prospect.,* **62,** 862–878.

Stromeyer, D., Grünthal, G. & Wahlström, R., 2004. Chi-square regression for seismic strength parameter relations, and their uncertainties, with applications to an M_W based earthquake catalogue for central, northern and northwestern Europe, *J. Seismol.,* **8,** 143–153.

Traversa, P. *et al.*, 2020. RESIF RAP and RLBP dataset of earthquake ground motion in Mainland France, *Seismol. Res. Lett.,* **91,** 2409–2424.

Volant, P., Berge-Thierry, C., Dervin, P., Cushing, M., Mohammadioun, G. & Mathieu, F., 2000. The South Eastern Durance fault permanent network: preliminary results, *J. Seismol.,* **4,** 175–189.

Volant, P., Lomax, A., Nechtschein, S., Cushing, M., Ait-Ettajer, T. & Berge-Thierry, C., 2003. Localisation 3D et calcul de magnitude pour les événements du réseau Durance, in *Presented at the 6ème colloque National de l'Association Française de Génie Parasismique,* 1–3 July, Ecole Polytechnique, Palaiseau, France, Vol. **I,** pp. 21–32.

Woessner, J. *et al.*, 2015. The 2013 European Seismic Hazard Model: key components and results, *Bull. Earthq. Eng.,* **13,** 3553–3596.

SUPPORTING INFORMATION

Supplementary data are available at *GJI* online.

Supporting_information_S1_to_S6.doc

Please note: Oxford University Press is not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the paper.