Bull Earthquake Eng (2018) 16:2169–2193
https://doi.org/10.1007/s10518-017-0178-7

ORIGINAL RESEARCH PAPER

Exploration tree approach to estimate historical earthquakes Mw and depth, test cases from the French past seismicity

Paola Traversa[1] · David Baumont[2] · Kévin Manchuel[1] · Emmanuelle Nayman[1] · Christophe Durouchoux[1]

Received: 14 October 2016 / Accepted: 19 June 2017 / Published online: 14 July 2017

Abstract The estimation of the seismological parameters of historical earthquakes is a key step when performing seismic hazard assessment in moderate seismicity regions as France. We propose an original method to assess magnitude and depth of historical earthquakes using intensity data points. A flowchart based on an exploration tree (ET) approach allows to apply a consistent methodology to all the different configurations of the earthquake macroseismic field and to explore the inherent uncertainties. The method is applied to French test case historical earthquakes, using the SisFrance (BRGM, IRSN, EDF) macroseismic database and the intensity prediction equations (IPEs) calibrated in the companion paper (Baumont et al. Bull Earthq Eng, 2017). A weighted least square scheme allowing for the joint inversion of magnitude and depth is applied to earthquakes that exhibit a decay of intensity with distance. Two cases are distinguished: (1) a "Complete ET" is applied to earthquakes located within the metropolitan territory, while (2) a "Simplified ET" is applied to both, offshore and cross border events, lacking information at short distances but disposing of reliable data at large ones. Finally, a priori-depth-based magnitude computation is applied to ancient or poorly documented events, only described by single/sporadic intensity data or few macroseismic testimonies. Specific processing of "felt" testimonies allows exploiting this complementary information for poorly described earthquakes. Uncertainties associated to magnitude and depth estimates result from both, full propagation of uncertainties related to the original macroseismic information and the epistemic uncertainty related to the IPEs selection procedure.

Keywords Seismic hazard · Historical seismicity · Seismicity catalog · Macroseismic intensity

✉ Paola Traversa
paola.traversa@edf.fr

1 EDF-CEIDRE-TEGG, 905 avenue du camp de menthe, 13090 Aix-en-Provence, France

2 Geoter – FUGRO, 20 Zac Pujol, 13390 Auriol, France

1 Introduction

When performing seismic hazard assessment studies, it is essential to take into account the historical seismicity, especially in low-to-moderate seismic activity regions such as France. Indeed, the instrumental period, which started in 1962 with the deployment of the CEA-LDG seismic network, does not include major earthquakes that occurred in the pre-instrumental period, and that are decisive while performing seismic hazard assessment for French critical facilities. The location and magnitude of these earthquakes must be assessed based on macroseismic data, which is the only information available for historical earthquakes.

Over the past decades different methods have been proposed to estimate the depth and the magnitude of earthquakes from macroseismic data. A common way of estimating the depth is based on the work of Kövesligethy (1907) and Sponheuer (1960), which is at the origin of the intensity attenuation models derived during the last decades worldwide. These authors show that intensity, when a sufficiently large number of points are available, exhibits a regular decreasing pattern with distance, which can be accounted for using a simple energy radiation model involving a point source. Their model can be written as follows (Burton et al. 1985):

$$\Delta I = I_{epc} - I_j = k \log\left(R_j/h\right)^m + k\alpha \log(e)\left(R_j - h\right) \tag{1}$$

where I_{epc} is the epicentral intensity; I_j and $R_j = \sqrt{D_j^2 + h^2}$ are the intensity and mean hypocentral radius of the jth isoseismal (D_j being the epicentral distance and h the focal depth); k is the proportionality factor between intensity scale degrees and ground motion amplitude; m describes the geometric spreading of the wave amplitude (1 for body waves), α is the absorption coefficient, depending on wave frequency and soil conditions and e is Euler's constant. Different authors apply the Sponheuer's (1960) model to estimate focal depths of historical earthquakes, using either isoseismals (e.g. Ambraseys 1985; Levret et al. 1994; Musson 1996) or individual intensity observations (Scotti et al. 1999; Gasperini et al. 2010).

The methods proposed in the literature to estimate the magnitude of earthquakes using macroseismic data mainly refer to three approaches (Gasperini et al. 2010). The first dates back to the work of Richter (1935), considers that the epicentral intensity is representative of the strength of the seismic source and relates it to the instrumental magnitude using an empirical relationship. However, when disposing of a macroseismic field for the considered earthquake, this approach is restrictive. The second approach relates the magnitude to the felt area (e.g. Galanopulos 1961; Toppozada 1975; Sibol et al. 1987), to the area within all available isoseismal lines (Gasperini and Ferrari 2000), or to the isoseismal radii (e.g. Ambraseys 1985). Johnston (1996) and Bakun and Wentworth (1997) point out that magnitudes determined using isoseismal areas are robust estimates only when the amount and the spatial distribution of the macroseismic observations available (hereafter referred to as intensity data points—IDPs) are sufficient to determine isoseismal areas with precision. To overcome the practical difficulties that can arise in determining isoseismal areas from sparse intensity datasets, Bakun and Wentworth (1997) propose a third approach, in which a magnitude M_j is estimated for each IDP using an intensity attenuation model (hereafter called Intensity Prediction Equation—IPE). Then, the preferred earthquake magnitude is the mean of all M_j:

$$M_I = \frac{1}{N}\sum_{j=1}^{N} M_j \tag{2}$$

This approach has been applied worldwide for historical earthquake magnitude estimation, e.g. France (Bakun and Scotti 2006), Northern Rhine area (Hinzen and Oemisch 2001), Eastern North America (Bakun et al. 2003), Southern California (Bakun 2006), Ecuador (Beauval et al. 2010).

In this work, the estimation of the seismological parameters (M_W and depth) of historical earthquakes from available IDPs makes use of isoseismal-radii-based IPEs in which the isoseismal radii are defined according to different types of metrics. Detailed description of these metrics can be found in the companion paper by Baumont et al. (2017), who reappraise the strategy developed by Baumont and Scotti (2011) in Stucchi et al. (2013). The computation of the seismological parameters is then integrated in an exploration tree (ET) framework. Depending on the characteristics of the macroseismic field available for a given earthquake, different strategies, i.e. different alternative paths, are followed. This flow chart allows to consider all possible macroseismic description configurations in the same framework. When applied in a systematic manner, it allows to obtain consistent (homogeneous) magnitude and depth estimates over the whole historical period (e.g. 463–1965 in France, see the companion paper by Manchuel et al. 2017). Depending on the case, the adopted strategy implies: (1) a "Complete ET" magnitude and depth simultaneous inversion scheme, when dealing with well documented events that exhibit a decay of macroseismic intensity with distance; (2) a "Simplified ET" magnitude and depth simultaneous inversion scheme when dealing with either, offshore or cross border events, which lack information at short epicentral distances but dispose of reliable data at large ones; (3) a priori-depth-based magnitude computation when dealing with poorly described (mainly ancient) events, only known by single/sporadic intensity data or worse, only few "felt" testimonies. In all cases the uncertainties associated to magnitude and depth estimates are quantified. These uncertainties result from the variability in the amount and in the quality of the macroseismic information available for the considered earthquake, as well as from the standard deviation of the adopted IPE model. Then the ET approach allows performing realistic characterization of the epistemic uncertainty related to the IPE(s) selection process.

2 Data and data processing

Since the beginning of the nuclear program launched in France in the 70's, a vast program of investigations was set up to collect all the available information on past earthquakes. This program led to the creation of the SisFrance database by BRGM, EDF and IRSN (www.sisfrance.net). For each earthquake, the SisFrance database provides information on the epicentral location, the epicentral intensity as well as the estimates of the effects induced by the earthquake at various localities (IDPs), expressed in steps of half intensity values according to the MSK 1964 intensity scale. Each IDP is associated to a quality factor that is representative of the level of confidence associated with the numerical estimate (quality A: certain intensity, quality B: fairly certain intensity, quality C: uncertain intensity). SisFrance does not provide a quantitative evaluation of the uncertainty these quality factors are representative of. In the present work, while building the isoseismals, we assign a weight to each IDP as function of its quality factor, as follows:

$W_{QA} = 4$, $W_{QB} = 3$ and $W_{QC} = 2$ for A, B and C qualities, respectively. When used, a weight $W_{QD} = 1$ is assigned to quantified "felt" testimonies (see Sect. 2.2). Considering these weights as the inverse of the squared standard deviation associated to the intensity value, we obtain: $\sigma_{I,QA} = 0.5$ MSK, $\sigma_{I,QB} = 0.58$ MSK, $\sigma_{I,QC} = 0.71$ MSK and $\sigma_{I,QD} = 1$ MSK) A quality factor is also associated to the epicentral intensity value (quality A: certain epicentral intensity; quality B: fairly certain epicentral intensity; quality C: uncertain epicentral intensity, quality E: arbitrary epicentral intensity; quality K: fairly certain epicentral intensity, resulting from a calculation based on an attenuation law). The uncertainty values we assign to the epicentral intensities are the following: $\sigma_{Io,QA} = 0.250$ MSK, $\sigma_{Io,QB} = \sigma_{I0,QK} = 0.375$ MSK, $\sigma_{Io,QC} = 0.500$ MSK, $\sigma_{Io,QE} = 0.750$ MSK.

Finally, the SisFrance database provides quality factors associated to the epicentral location [quality A: exact location (a few km); quality B: fairly certain location (around 10 km); quality C: imprecise location (between 10 and 20 km); quality D: strongly assumed location (from a few km to 50 km); quality E: arbitrary location].

In this work, the 2014 edition of the SisFrance database has been used.

2.1 Treatment of "felt" testimonies

When the description of the effects produced by an earthquake at a given locality is not detailed enough in the original sources, it is not always possible to assess an intensity level (note that the term "felt" testimony refers to this kind of information). This holds in particular for ancient events, whose IDP distribution is often strongly incomplete and sometimes essentially composed of "felt" testimonies. Nevertheless, "felt" testimonies can be very valuable and should be exploited when more quantitative information is either, insufficient or lacking. However, one of the issues that are faced when dealing with "felt" testimonies is that they might be representative of several ground shaking levels and may vary with time. Indeed, for recent earthquakes, intensities as low as II MSK can be collected at large distances (i.e. during the 2003 February 22nd, Rambervillers earthquake), whereas for older events the felt area is likely to be related to larger intensities.

In order to shed some light on this issue, IDP distributions are built for various periods of time (Fig. 1). Half-intensity levels are assimilated to the lower intensity value (i.e. intensities II–III MSK are assimilated to the value of II MSK, intensities III–IV MSK to the value of III MSK and so on). The considered periods of time are defined by trial and error in order the resulting distributions are characterized by the sharpest drop at low intensity. Even though this approach carries some subjectivity, sensitivity analyses do not highlight any "cliff-edge" effect, i.e. the characteristics of the distributions evolve smoothly when modifying the limits of the selected time periods.

Figure 1 (top) shows that for the most recent period, the SisFrance database is statistically not complete below intensity III MSK. Indeed, a drop in the number of intensity data is observed below intensity III MSK, while one would expect the number of observations to increase from intensity III MSK to intensity II MSK. Similarly, during the period [1875–1980], the data distribution is characterized by a lack of intensities smaller than IV MSK. Finally, before 1875 the distribution is characterized by a lack of intensities smaller than V MSK.

Assuming that the "felt" testimonies correspond to intensities poorly represented in the SisFrance database, they are introduced into the analysis adopting the following strategy:

- For earthquake occurred starting from 1980, the intensity value associated with "felt" testimonies is randomly chosen between II and II–III MSK;

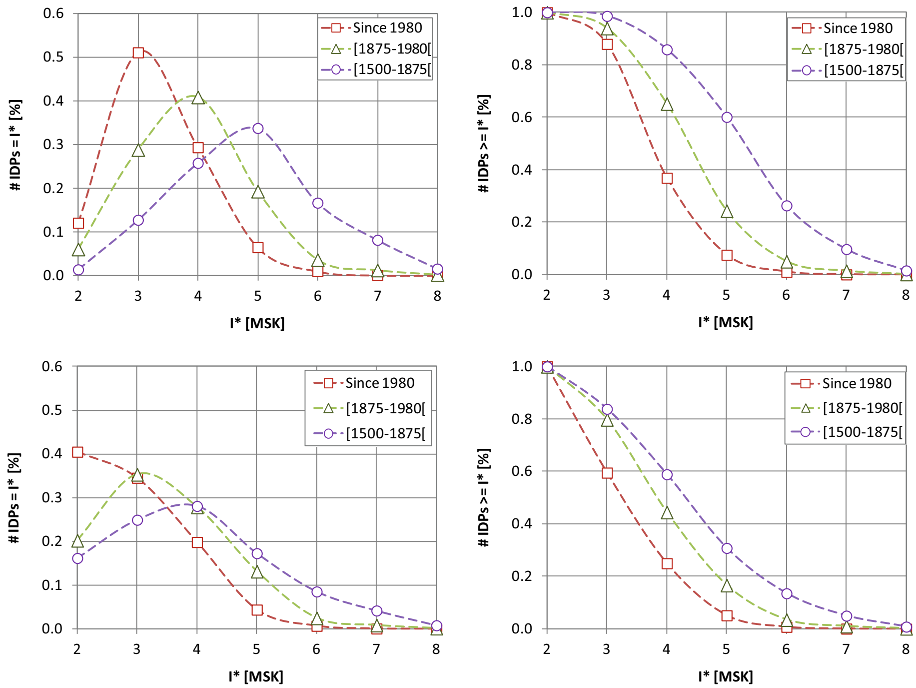

Fig. 1 *Top* Frequency (*left*) and cumulative (*right*) distributions of SisFrance IDPs for different time periods: *red-square line* earthquakes post-1980; *green-triangle line* earthquakes occurred within the period [1875–1980]; *violet-diamond line* earthquakes occurred before 1875. *Bottom* same as *top*, but considering the "felt" testimonies

- For earthquakes occurred during the period [1875–1980], the intensity value associated with "felt" testimonies is randomly chosen between III and III–IV MSK;
- For earthquakes occurred before 1875, the intensity value associated with "felt" testimonies is randomly chosen between IV and IV–V MSK.

As shown in Fig. 1 (bottom), taking into account "felt" testimonies allows to partially correct the effect of the lack of small intensity data and to improve the overall data completeness. In this work, as explained in the following section, this statistical correction is applied before estimating magnitude and depth of poorly documented earthquakes, allowing to enlarge the exploitable information.

2.2 Processing of SisFrance macroseismic intensity data

Prior to applying the seismological parameters estimation methodology, IDPs associated to the considered earthquake undergo an intensity binning procedure aimed at building the isoseismals. Either, for earthquakes disposing of less than 50 intensity observations, or earthquakes for which the proportion of felt testimonies is larger than 10% of the total number of observations, we assign an intensity value to the available "felt" testimonies depending on the epoch they derive from. Following the analysis described in Sect. 2.1, we apply:

- Earthquakes occurred starting from 1980: I_{felt} = II MSK;

- Earthquakes occurred during the period [1875–1980]: I_{felt} = III MSK;
- Earthquakes occurred before 1875: I_{felt} = IV MSK.

The IDPs “created” in this way are assigned a quality factor D.

Consistently to the procedure adopted in the companion paper (Baumont et al. 2017), six different intensity binning strategies are applied, aimed at obtaining six different representations of the available data. The following six metrics are defined:

1. R_{AVG}, corresponding to the weighted barycenter of the IDPs within a given class of intensity (i.e. for the considered intensity class, the isoseismal radius is taken as the weighted mean of the $\log_{10}$ IDP hypocentral distances, and the corresponding intensity is the weighted mean of IDP intensity values). The width of the intensity class is fixed to 1.0 MSK. In the mean computation, the weight assigned to each IDP is function of the SisFrance quality factor associated to it ($W_{QA} = 4$, $W_{QB} = 3$ and $W_{QC} = 2$ for A, B and C qualities, respectively. When used, a weight $W_{QD} = 1$ is assigned to quantified felt testimonies). The uncertainty associated to each R_{AVG} isoseismal radius is defined as the weighted standard deviation of the considered IDPs;
2. The R_{OBS} metric is identical to the R_{AVG} but using an intensity bin width equal to zero;
3. The R_{P50} metric is analogue to R_{OBS} metric, but the weighted 50th percentile (using an intensity bin width equal to zero) is considered. The uncertainties on the R_{P50} isoseismal radii are defined as the half of the difference between the percentiles 84th and 16th of the IDPs hypocentral distances (expressed in the $\log_{10}$ scale);
4. The R_{P84} metric is analogue to the R_{OBS} metric, but the weighted 84th percentile (using an intensity bin width equal to zero) is considered. The uncertainties on the RP_{84} isoseismal radii are defined as the half of the difference of the percentiles 98th and 50th of the hypocentral distances (expressed in the log scale);
5. The R_{F50} metric (to be intended as a kind of “felt radius”) is equivalent to the R_{P50} metric but, rather than describing the decay of the macroseismic intensity with respect to the distance, only the intensity bin being representative of far field reliable information is considered. To obtain this isoseismal, the *W* parameter is defined, as follows:

$$W_j = W_{n_j} \cdot W_{q_j} \cdot R_{P50_j} \tag{3}$$

   where j represents the jth isoseismal, W_{n_j} is the square root of the number of IDPs in the isoseismal j, W_{q_j} is the sum of the weights associated to IDPs of isoseismal j (defined as function of the quality index assigned by SisFrance to each IDP) and R_{P50_j} is the jth isoseismal radius. The isoseismal presenting the largest value of parameter W is selected as R_{F50}.
6. The R_{F84} metric is equivalent to the R_{F50} metric, but the 84th percentile is considered.

Finally, for all these metrics, the uncertainties associated to the isoseismals intensity values are deduced from the uncertainties on the hypocentral radii using the apparent slope β of the intensity decay with the log distance ($I = I_0 - \beta \log_{10} {}^{R}/_{H}$), measured on the isoseismal radii above completeness for each event.

Figure 2 illustrates the application of the six binning strategies for the September 1866 Brenne earthquake.

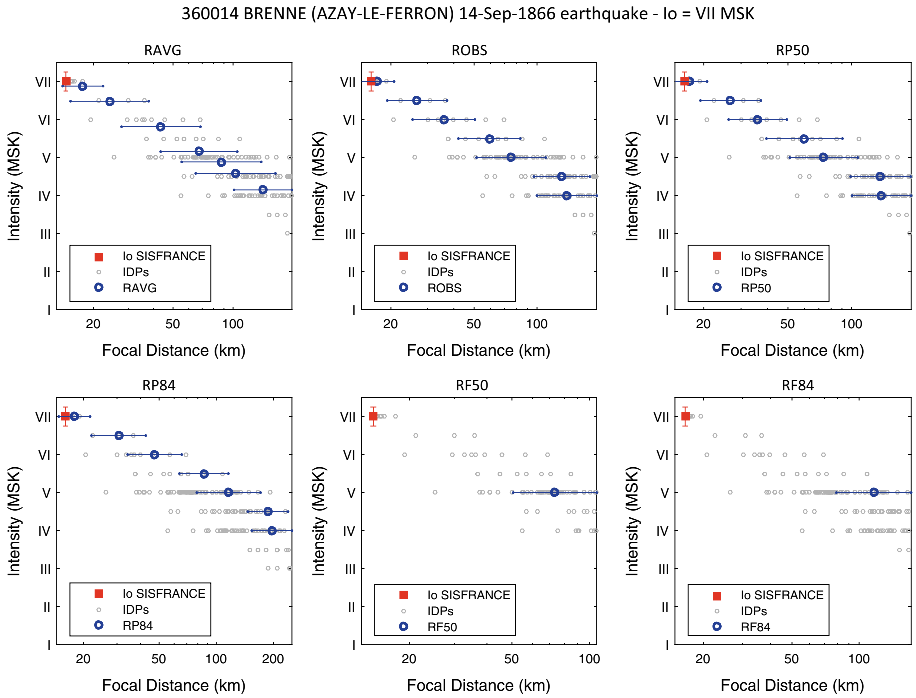

Fig. 2 Illustration of the six different binning strategies adopted to determine isoseismal radii from the observed data. *Light grey circles* IDPs; *blue circles* isoseismal radii for the corresponding metric; *blue horizontal lines* standard deviation associated to each isoseismal radius; *red square* epicentral intensity; *red vertical lines* standard deviation associated to the epicentral intensity value. Starting from the *upper left panel* and following clock-wise: R_{AVG}, R_{OBS}, R_{P50}, R_{F84}, R_{F50} and R_{P84} metrics

2.3 Intensity of completeness

Basing the estimate of the earthquake seismological parameters on the decay of macroseismic data with distance relies on the hypothesis that the macroseismic dataset is complete. If such hypothesis holds for recent events down to very weak intensities, it may fail for historical events, for which weakest intensity information either has not been passed down to these days or is particularly difficult to collect. If not taken into account, these cutoffs in the data collection may introduce a bias in the inversion scheme. The following procedure has therefore been set up to identify the intensity of completeness in the macroseismic dataset of a given event, value below which the intensity data has been discarded from the inversion procedure. The lack of completeness is detected using the R_{AVG} metric by identifying the inflection point in the cumulative number of observations per intensity bin, defined as the point where the second derivative of the cumulative number of observations as function of the intensity bin value changes sign. The intensity of completeness (IC) is therefore the intensity value corresponding to the inflection point. Figure 3a illustrates the macroseismic field for the September 1866 Brenne earthquake. After building the isoseismal radii according to the R_{AVG} binning strategy (Fig. 3b), the Intensity of Completeness (IC) is identified with the red dashed line on Fig. 3c.

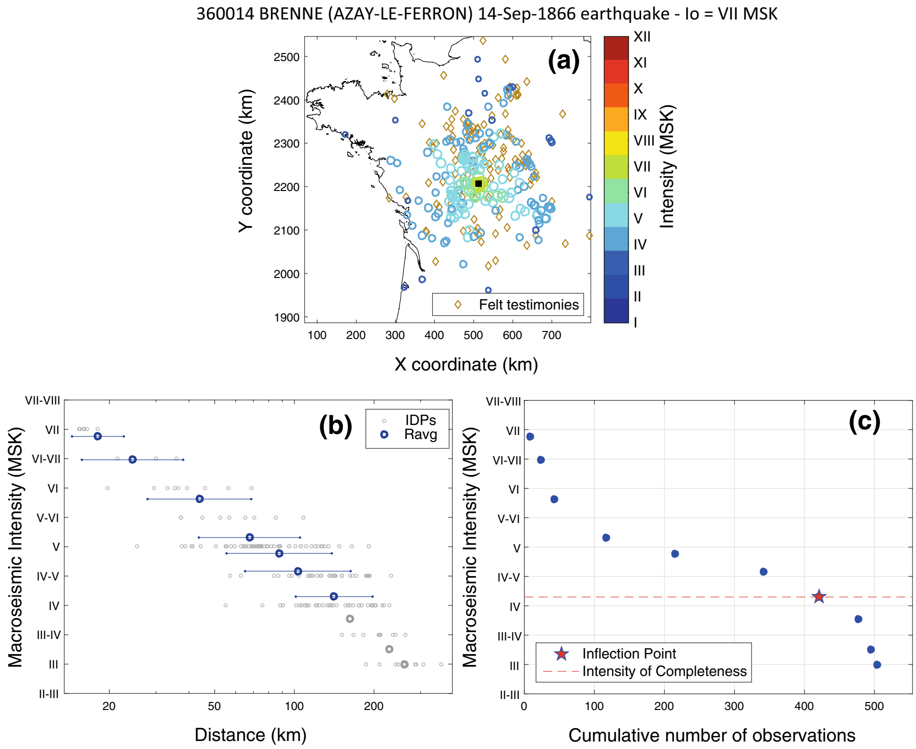

Fig. 3 The September 1866 Brenne earthquake (SisFrance reference number: 360014). **a** IDPs map, *black square* earthquake epicentral location (SisFrance); *color circles* IDPs (*circle colors* and *sizes* are related to the intensity value, following the *colorbar* on the *right side*; *gold diamonds* felt testimonies. **b** Decay of intensity with distance, *light grey circles* IDPs; *bold grey circles* isoseismal radii obtained using the R_{AVG} intensity binning strategy; *blue circles* isoseismal radii above completeness; *blue lines* standard deviation associated to each isoseismal radius. **c** Isoseismal radii as function of the cumulative number of observations for each intensity bin, *red star* computed inflection point (see text); *dashed line* intensity of completeness

3 Flowchart and exploration tree: ET-approach

When calibrating intensity prediction equations (IPEs), Baumont et al. (2017) show that the result is pretty sensitive to both, the chosen parametrization and the calibration dataset. As a consequence, choosing a single (best) IPE to be used in the estimate of the seismological parameters for historical earthquakes clearly appears not reasonable. On the other hand, the use of a set of models looks a rational way to overcome this limitation. The results of both, residual analysis and goodness of fit evaluation procedures performed by Baumont et al. (2017) on the whole set of IPEs they propose, guide us in the selection of the subset of IPEs that are the most appropriate to be used in the estimation of the seismological parameters of historical earthquakes in metropolitan France.

For detailed description of the earthquake datasets used to calibrate the IPEs, as well as their associated uncertainties and regional dependencies, the reader can refer to the Baumont et al. (2017) companion paper. Among the IPEs proposed by these authors we select the following “families” of models:

1. IPEs including both, geometrical spreading and intrinsic attenuation in their functional form;
2. IPEs defined for either, regional attenuation domains or an average attenuation at the national scale. In the latter case, the spatial variability of attenuation is not taken into account, while in the former case, the regional attenuation domains are defined on the basis of the Mayor et al. (2017) intrinsic attenuation maps at 1 Hz. In practice, within both, the IPE calibration phase, and the applications presented in this work, two main attenuation domains are considered: a low attenuation zone covering the Western part of France, and a high attenuation zone covering the rest of the French territory (see Fig. 2 in Baumont et al. 2017). These two alternative parametrizations are considered as different branches in the ET;
3. IPEs for which the epicentral intensity value was included either, as a constraint or simply as an additional datum during the calibration phase. Similarly to the previous item, these two alternative modelizations are considered as different branches in the ET;
4. IPEs calibrated using IDPs $\leq$ VII MCS intensity scale for the Italian calibration earthquake datasets.

Within each of these four alternatives, which constitute as many ET branches, all the models determined on the basis of different data selection choices (e.g. calibration datasets characterized by Completeness Distance (distance corresponding to the Intensity of Completeness value) Dc $\geq$ 30 km and 50 km and calibration datasets characterized by a number of complete intensity classes N_CLASS $\geq$ 3, 5 and 7) are retained, giving six additional ET branches. Finally, for each of these models, the formulations following the six metrics described above (i.e. R_{AVG}, R_{OBS}, R_{P50}, R_{P84}, R_{F50}, R_{F84}) are considered, either all of them or two of them (R_{F50}, R_{F84}), following the case.

In summary, a maximum of $N = 4 \times 6 \times 6 = 144$ different IPEs are retained in the calculations. The ET approach allows us to deal with such a number of alternative models and to characterize the epistemic uncertainty related to the model selection process (i.e. the variability in the magnitude and depth estimates resulting from the fact that we do not know the "true" model that can explain the data exactly).

As mentioned before, all SisFrance earthquake macroseismic field configurations are treated following this same philosophy. However, depending on the macroseismic information we dispose of for a given earthquake, two main cases can be identified, as shown by the flow chart illustrated in Fig. 4 and described in the following sections.

3.1 Case of several IDPs available

When several IDPs are available for a given earthquake, it is possible to build isoseismal radii according to the six metrics described above. In this case, magnitude and depth can be simultaneously computed through a maximum likelihood inversion scheme based on a weighted least squares (WLSQ) criterion (Tarantola 2005). The inversion procedure will be described in Sect. 4.1. Two subcases are further distinguished.

3.1.1 Earthquakes located within the French territory

In this case, we assume that the macroseismic data is sufficiently well distributed in both, the near and the far field. The so-called "Complete ET" (Fig. 4) is applied, i.e. the simultaneous inversion of magnitude and depth is performed for each selected IPE (i.e. 144

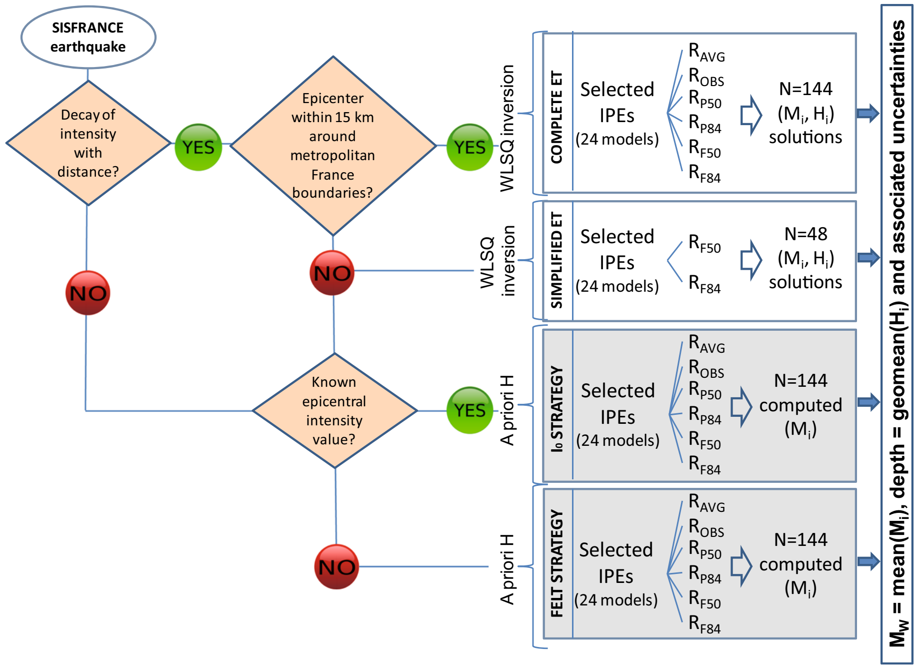

Fig. 4 Flowchart and exploration tree (ET) scheme adopted for the estimation of magnitude, depth and associated uncertainties of SisFrance earthquakes

models), resulting in 144 individual couples (M_i, H_i) allowing the different models to explain the decay of the intensity with distance, as well as the epicentral intensity value.

3.1.2 Earthquakes located either, offshore or outside the French boundaries

For most offshore or cross-border earthquakes, the macroseismic field provided by SisFrance is generally poorly described in the near field, with very few (if any) IDPs in the epicentral area. On the other hand, macroseismic observations are available at larger distances. Although the reason of this is obvious in the case of offshore earthquakes, for cross-border earthquakes this lack of data can be considered as inherent to the Sisfrance database, which is implemented in France and therefore mostly contains intensity observations collected within the French territory. Efforts of data homogenization and sharing at the European scale are ongoing (e.g. AHEAD platform, Stucchi et al. 2013; Locati et al. 2014) and should possibly be enhanced through EU founded project. In the future, these efforts will allow to dispose of homogeneous cross-border macroseismic data to be used in the estimation of the seismological parameters of earthquakes located close to the border, on either, one or the other side. In these cases we propose to adopt a "Simplified ET" (Fig. 4), in which all IPEs are used, but only the R_{F50} and R_{F84} metrics are retained. Indeed, since the near field information for these types of earthquakes is either, nonexistent or strongly incomplete, we believe that restricting the inversion scheme to metrics representative of the earthquake effects in the far field allows focusing the computation on unbiased isoseismals. Again, simultaneous inversion of individual magnitude and depth couples is performed, resulting in 48 individual couples (M_i, H_i) allowing the different models to explain the far-field macroseismic information, as well as the epicentral intensity value.

3.2 Case of single/sporadic (if any) intensity observations available

This is unfortunately the case for the majority of SisFrance earthquakes. Two subcases are distinguished.

3.2.1 Known epicentral intensity value, but isolated or sporadic IDPs available

This can often be the case of significant ancient earthquakes, for which handed-down information at the epicenter is sufficient to assign an intensity value, but information lacks when getting away from it. In this case the "I_0 strategy" is adopted (Fig. 4), i.e. the depth is fixed a priori (see Sect. 3.3) and 144 individual magnitudes (M_i) are computed for the considered earthquake using the selected IPEs, the a priori depth and the epicentral intensity value.

3.2.2 Unknown epicentral intensity value

This is generally the case of extremely poorly known earthquakes, for which we dispose of several felt testimonies, but the associated information and the descriptions of the effects induced by the earthquake are not sufficient to assign a quantitative value to any data point (i.e. absence of quantified IDP). Nonetheless, based on expert judgment, SisFrance provides an epicenter location for these earthquakes. Following the methodology illustrated in Sect. 2.1, an intensity value is assigned to "felt" testimonies, based on the year they derive from. We therefore dispose of a "felt intensity value" and a "felt radius", this latter being defined as the mean hypocentral distance of felt testimony locations. In this case the "Felt strategy" (Fig. 4) is adopted, i.e. the depth is fixed a priori similarly to the previous case and 144 individual magnitudes (M_i) are computed using the selected IPEs.

3.3 Earthquake magnitude and depth

For all the cases identified in the previous sections, the earthquake magnitude is obtained as the arithmetic mean of all individual magnitudes (M_i), while the earthquake depth, when computed, is obtained as the geometric mean of individual depths (H_i). Uncertainties associated to individual magnitudes and depths, as well as the total uncertainty associated to the final earthquake magnitude and depth values are computed following the methodology described in Sect. 6.

4 Simultaneous inversion of earthquake magnitude and depth

4.1 Methodology

Baumont et al. (2017) describe the attenuation of macroseismic intensity as function of the hypocentral distance and the magnitude as follows:

$$I_j = c_1 + c_2 \cdot M_W + \beta \cdot \log_{10}\left(R_j\right) + \gamma_{region} \cdot R_j \tag{4}$$

where similarly to Eq. (1), I_j and $R_j = \sqrt{D_j^2 + H^2}$ are the intensity and the mean hypocentral radius of the jth isoseismal, defined consistently to the considered metric (i.e. R_{AVG}, R_{OBS}, R_{P50}, R_{P84}, R_{F50} or R_{F84}). D_j are jth the epicentral radii and H the focal depth. M_W is the moment magnitude of the earthquake and c_1, c_2, β and γ_{region} are the IPE coefficients determined by Baumont et al. (2017) consistently to each metric. Using the Tarantola (2005) notation, our forward problem can be formulated as follows in a condensed form:

$$d = g(m) \tag{5}$$

where d is the vector of the observable parameters (i.e. the intensities of each isoseismal, defined according to a given metric and including the epicentral intensity value provided by SisFrance), g is the operator and m is the vector of model parameters (the unknowns M_W and H). The operator g being a non-linear function of m, we face a problem of non-linear least squares minimization. To solve it we assume that $g(m)$ is quasi-linear in the region of the model-data space of significant posterior probability density and we apply an iterative Quasi-Newton procedure. Following Tarantola (2005), at each iteration the perturbation of the model parameters is given by:

$$m_{n+1} = m_n - \left(G_n^t C_D^{-1} G_n + C_M^{-1}\right)^{-1} \left(G_n^t C_D^{-1} (d_n - d_{obs}) + C_M^{-1}\left(m_n - m_{prior}\right)\right) \tag{6}$$

where $d_n = g(m_n)$ is the vector composed of the isoseismal intensities I_j computed with the current model, m_n is the vector composed of the seismological parameters M_W and H at iteration n and G_n is the matrix composed of the partial derivatives of I with respect to the current values of the seismological parameters:

$$\begin{aligned} &(G_n)^j_\alpha = \left(\partial g^j / \partial m^\alpha\right)_{m_n}: \\ &\left(\frac{\partial g^j}{\partial H}\right)_{m_n} = \frac{\beta \cdot H}{\left(D_j^2 + H^2\right) \cdot \ln(10)} + \frac{\gamma \cdot H}{\sqrt{D_j^2 + H^2}} \\ &\left(\frac{\partial g^j}{\partial M_W}\right)_{m_n} = c_2 \end{aligned} \tag{7}$$

C_D and C_M are the data and model covariance matrices, respectively. C_D represents observational uncertainties as well as model uncertainties. Making the assumption of Gaussian uncertainties for both, modelization and observations, $C_D = C_d + C_T$. C_d is the data covariance matrix, a diagonal matrix with $(C_d)_{jj} = (\sigma_d)_j^2$ where $j = 1,\ldots$ is the number of isoseismals for the considered metric (including the epicentral value provided by SisFrance) and $(\sigma_d)_j$ are the weighted standard deviations associated to each isoseismal. C_T is the covariance matrix representing the modelization uncertainties, a diagonal matrix with $(C_T)_{jj} = (\sigma_{IPE})_j^2$. Standard deviations associated to IPEs (σ_{IPE}) are provided by Baumont et al. (2017). C_M is also a diagonal matrix with $(C_M)_{kk} = (\sigma_M)_k^2$ where $k = 1, 2$ is the number of parameters (i.e. M_W and H) and $(\sigma_M)_k$ are defined a priori.

As already described in Sect. 2.2, since the metrics used in this work are based on an intensity binning, the standard deviations associated to the isoseismals are originally expressed as weighted standard deviations of the isoseismal radii. The standard deviations on the intensity value of each isoseismal $(\sigma_d)_j$ are then obtained from these weighted standard deviations using the apparent slope of the decay of intensity with the log distance, measured on the isoseismal radii above completeness for the considered event.

In practice, to avoid giving higher importance to less represented isoseismals (isoseismals built on the basis of a smaller number of IDPs might have lower standard deviations than others based on a larger number of IDPs), a minimum value for the standard deviation of the isoseismal radii is imposed before obtaining the $(\sigma_d)_j$. This minimum value is equal to $1/\sqrt{W_{QC}} \cdot 1/\sqrt{N_j}$ where N_j is the number of IDPs in the jth isoseismal.

As mentioned above, the epicentral intensity provided by SisFrance is considered as an additional isoseismal, with a standard deviation defined on the basis of the quality index assigned by SisFrance, as follows: $\sigma_{I_0,A} = 0.250$, $\sigma_{I_0,K} = 0.250$, $\sigma_{I_0,B} = 0.375$, $\sigma_{I_0,C} = 0.500$, $\sigma_{I_0,E} = 0.750$ for quality indexes equal to A, K, B, C or E, respectively.

The algorithm is initiated at a starting point corresponding to the seismological parameters that would be obtained using an I_0 strategy for the considered metric. The iterative procedure is stopped when maximum likelihood point m_{ML} is reached, i.e. when the ratio between each model parameter at iteration n and at iteration $n-1$ is smaller or equal than a tolerance value fixed at 1 ± 0.0001. To avoid the danger of falling in a local minimum, the algorithm is reinitiated five times at different randomly chosen magnitude and depth starting couples. The solution presenting the minimum misfit function is selected. For each computation, the misfit function is computed as follows:

$$\begin{aligned} 2S(m) &= \|g(m) - d_{obs}\|_D^2 + \left\|m - m_{prior}\right\|_M^2 = (g(m) - d_{obs})^t C_D^{-1} (g(m) - d_{obs}) \\ &\quad + \left(m - m_{prior}\right)^t C_M^{-1} \left(m - m_{prior}\right) \end{aligned} \tag{8}$$

In average, ten iterations are performed to conveniently approach the point m_{ML}. Then, the a posteriori covariance operator can be estimated as:

$$\widetilde{C}_M \cong \left(G^t C_D^{-1} G + C_M^{-1}\right) = C_M - C_M G^t \left(G C_M G^t + C_D\right)^{-1} G C_M \tag{9}$$

where this time, G are the partial derivatives taken at the convergence point: $(G_n)^j_\alpha = \left(\partial g^j / \partial m^\alpha\right)_{m_{ML}}$. Following Tarantola (2005), we interpret the square roots of the diagonal elements of the posterior covariance operator $\widetilde{C_M}$ as "uncertainty bars" on the posterior values of the model parameters M_W and H (see Sect. 6).

4.2 Application to well described earthquakes, example of the 1866 Brenne earthquake

As shown in Fig. 3, the macroseismic field for the 1866 Brenne earthquake is well represented and well distributed in space, with isoseismals considered complete down to intensity IV MSK (Fig. 3c). This earthquake is a typical example of a SisFrance event for which the weighted least squares (WLSQ) inversion scheme described in Sect. 4.1 succeeds in determining robust, joint estimates of magnitude and depth parameters. Figure 5 show an example of an IPE fitting the isoseismal radii defined according to the different metrics (R_{AVG}, R_{OBS}, R_{P50}, R_{P84}, R_{F50}, R_{F84}).

Figure 6 shows the synthesis of all the solutions obtained by simultaneous magnitude and depth WLSQ inversion. As described before, each solution is representative of one of

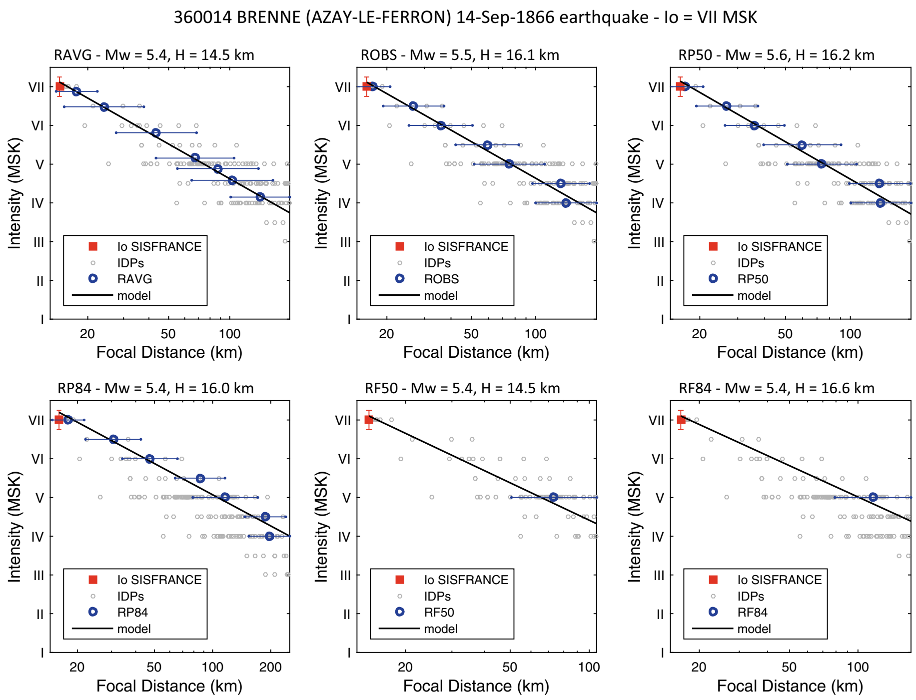

Fig. 5 Plots representing the fit of one IPE (making reference to the work of Baumont et al. (2017): 2.2.2.0 model for low the attenuation domain, calibration subset Dc30N5) to the isoseismal radii defined according to the six metrics (starting from the *upper left panel* and following clock-wise: R_{AVG}, R_{OBS}, R_{P50}, R_{F84}, R_{F50} and R_{P84} metrics). *Black line* IPE; *blue circles* isoseismal radii with associated weighted standard deviations; *red square* epicentral intensity provided by the SisFrance database with associated uncertainty; *grey light circles* IDPs

the 144 selected models. Although some dispersion of the results can be observed, the solutions obtained from the proposed inversion scheme are consistent to one another. We consider that the observed dispersion is representative of the uncertainty we would expect for the seismological parameters of an earthquake of the XIX Century. Indeed, the total uncertainty on the earthquake magnitude value is of 0.3 units, while the total uncertainty on the logarithm ($\log_{10}$) of the earthquake depth value is equal to 0.10. Detailed description of the uncertainty quantification procedure will be given in Sect. 6.

4.3 Application to offshore or foreign earthquakes, the example of the 1804 English Channel earthquake

Figure 7 shows the example of the September 1804, English Channel earthquake. Based on expert judgment, SisFrance provides a location for the epicenter of this earthquake, located off the coast of Saint-Malo. Obviously, no information within the epicentral area is available, while a few far field IDPs, as well as two felt testimonies are provided by SisFrance for this event (Fig. 7). Given the small number of IDPs available, no intensity of completeness can be estimated on this dataset. In order to exploit at maximum all the information available, prior to running the inversion, the value of IV MSK is assigned to the two felt testimonies available, according to the criteria described in Sect. 2.2.

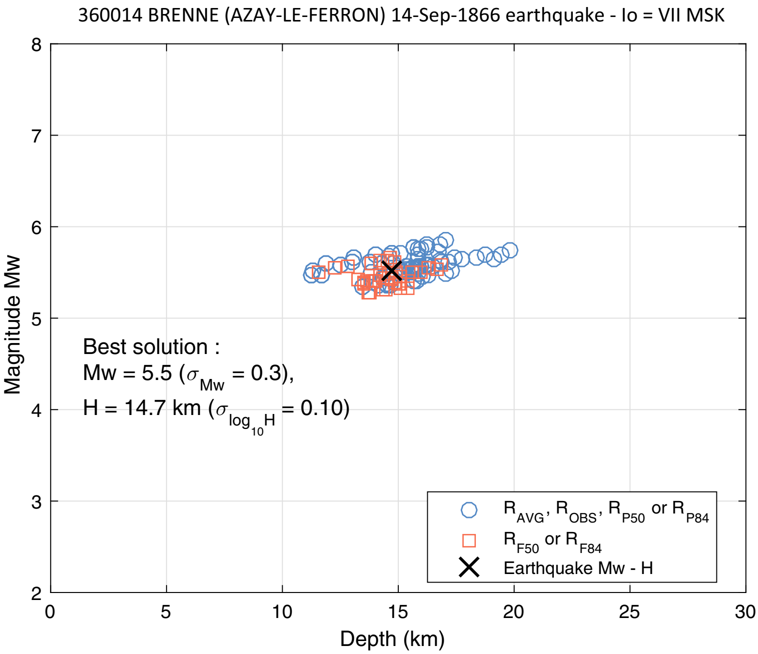

Fig. 6 Synthesis of magnitude and depth results obtained for the 1866 Brenne earthquake using all selected IPEs in the complete ET scheme described in Sect. 3. *Blue circles* WLSQ inversion solutions for R_{AVG}, R_{OBS}, R_{P50} and R_{P84} metrics; *Red squares* WLSQ inversion solutions for R_{F50} and R_{F84} metrics; *black cross* best solution, obtained as the arithmetic mean of all magnitude solutions and the geometric mean of all depth solutions

Following the inversion scheme described in Sect. 3, the "Simplified ET" (Fig. 4) is applied for this earthquake, i.e. for each selected IPE, the simultaneous inversion of magnitude and depth is based on the epicentral intensity value provided by SisFrance and on the isoseismal defined according to the R_{F50} and R_{F84} metrics. As mentioned above, these two metrics are representative of far-field reliable information rather than describing the decay of the macroseismic intensity with respect to the distance.

Figure 8 shows an example of an IPE fitting the isoseismals defined according to the R_{F50} and the R_{F84} metrics.

Figure 9 shows the synthesis of all solutions obtained by simultaneous magnitude and depth WLSQ inversion for the four selected IPE families, the six calibration earthquake subsets and both metrics, i.e. $4 \times 6 \times 2 = 48$ solutions. In this case, the dispersion of the results in terms of depths is larger than that of the Brenne earthquake (Fig. 6). Keeping in mind that the considered earthquake is located offshore and is described only by few far field macroseismic observations, we consider this dispersion as a realistic representation of the uncertainty associated to the depth of an earthquake under these conditions. Lesser scatter is obtained for the magnitude results, which is a parameter easier to constrain than depth, especially when far field information is available. The total uncertainty on the earthquake magnitude value is of 0.6 units, while the total uncertainty on the logarithm ($\log_{10}$) of the earthquake depth value is equal to 0.18. Detailed description of the uncertainty quantification procedure will be given in Sect. 6.

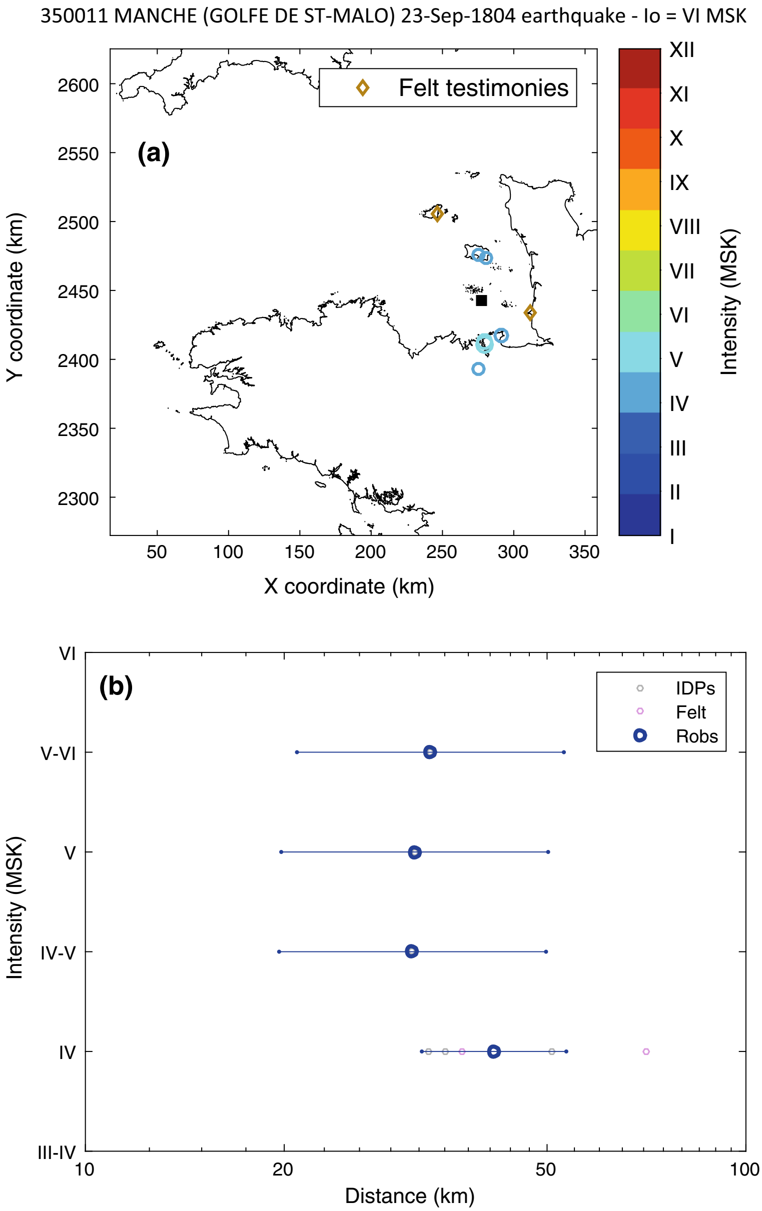

Fig. 7 The 1804 earthquake in the English Channel (SisFrance reference number: 350011); **a** IDPs map, *black square* earthquake epicentral location; *color circles* IDPs (*circle colors* and *sizes* are related to the intensity value, following the *colorbar* on the *right side*; *gold diamonds* felt testimonies. **b** Decay of intensity with distance, *light grey circles* IDPs; *light purple circles* felt testimonies converted into quantified intensity values; *blue circles* isoseismal radii obtained using the R_{OBS} intensity binning strategy (the whole dataset is considered); *blue lines* standard deviation associated to each isoseismal radius

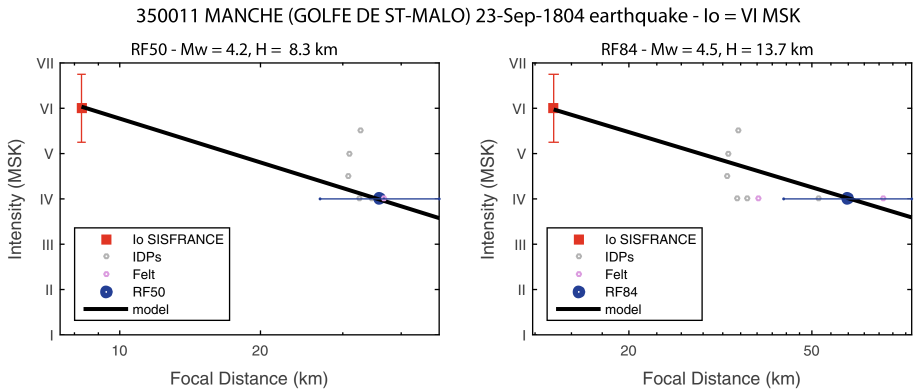

Fig. 8 Fit of one IPE (making reference to the work by Baumont et al. (2017): 2.2.1.0 model for low the attenuation domain, calibration subset Dc50N7) to the isoseismal radius defined according to the two metrics (R_{F50}, R_{F84} in the *left* and *right panel*, respectively, see explanation in the text). *Black line* IPE; *blue circles* isoseismal radius with associated weighted standard deviations; *red square* epicentral intensity provided by the SisFrance database with associated uncertainty; *light grey circles* IDPs; *light purple circles* felt testimonies

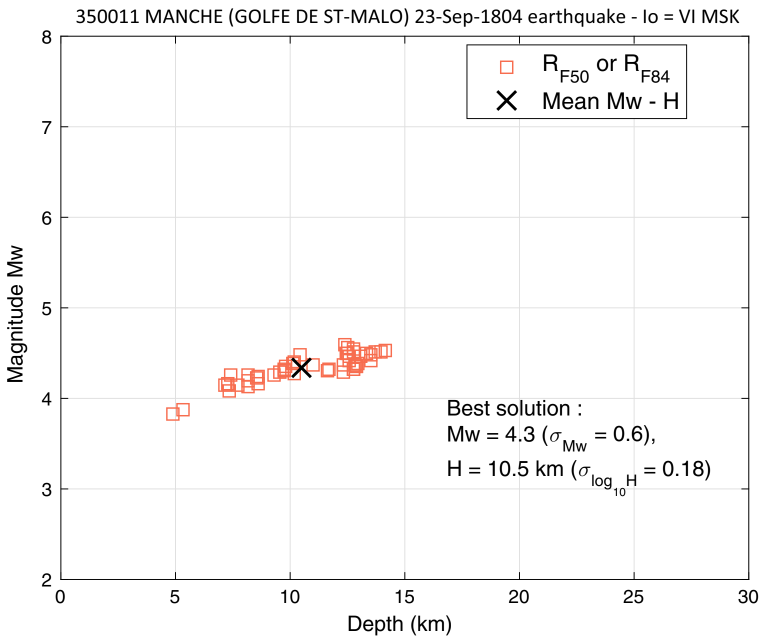

Fig. 9 Synthesis of magnitude and depth results obtained for the 1804 English Channel earthquake, using all selected IPEs in the simplified ET scheme described in Sect. 3. *Red squares* WLSQ inversion solutions for R_{F50} and R_{F84} metrics; *black cross* best solution, obtained as the arithmetic mean of all magnitude solutions and the geometric mean of all depth solutions

5 Magnitude and depth estimation for poorly known earthquakes

As explained in Sect. 3.2, in the case of poorly known earthquakes, i.e. the cases of single/sporadic (if any) quantified IDP available, the depth must be determined a priori using complementary information. Manchuel et al. (2017), divide the French metropolitan territory and surrounding areas into the following nine regions: the French Alps, the Pyrenees, the Armorican Massif, the European Cenozioc Rift System (ECRIS), the Provence, the Tricastin region, the Hainaut, the Atlantic Ocean zone and the rest of France (see Fig. 3 in Manchuel et al. 2017). Each region is considered homogeneous in terms of seismogenic depth properties. In each of them, Manchuel et al. (2017) build a regional depth distribution using the earthquake depths computed using the WLSQ inversion proposed in the present paper (i.e. the depth of the earthquakes falling in the cases described in Sect. 3.1). These authors suggest then to select the median of this, assumed to be representative of the regional seismogenic depth, distribution as the "preferred value" to be used as a priori depth value for poorly described earthquakes in the considered region.

5.1 "I_0 strategy": the example of the 1509 Manosque earthquake

Figure 10 illustrates the case of the 1509 Manosque earthquake. A single description of the earthquake effects allowing to assign an intensity value has been collected by SisFrance. The epicenter is then located by SisFrance on this same locality. A "felt" testimony is also provided for this earthquake (Fig. 10a). Nevertheless, the epicentral intensity being available, the "I_0 strategy" is applied (see the flowchart in Fig. 4). Following the regionalization scheme proposed by Manchuel et al. (2017) the 1509 Manosque earthquake is located within the Provence region. The earthquake depth is then fixed a priori to the value of 6 km, with an associated uncertainty $\sigma_{\log10(h)} = 0.20$ (see Table 1 in Manchuel

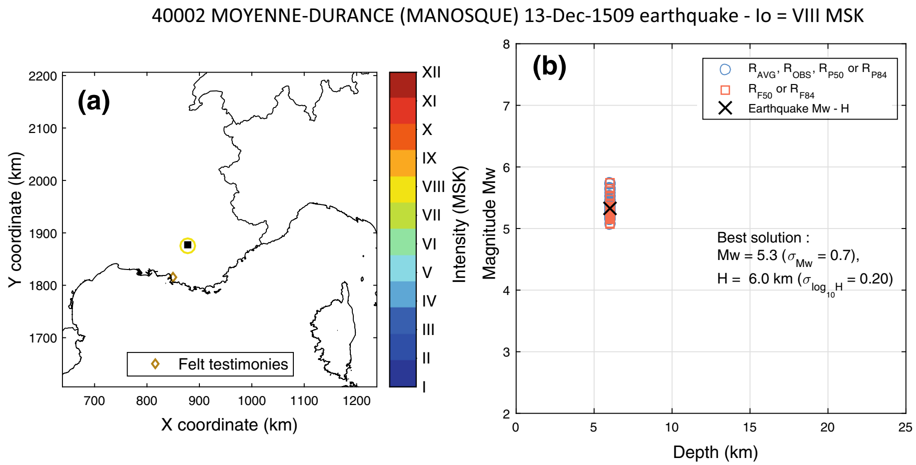

Fig. 10 The 1509 Manosque earthquake (SisFrance reference number: 40002); **a** IDPs map: *black square* earthquake epicentral location; *colored circles* IDPs (*circle color* and *size* is related to the intensity value, following the *colorbar* on the *right side*); *gold diamonds* felt testimonies. **b** Synthesis of results obtained using Eq. (4) expressed as function of Mw using all selected IPEs; *blue circles* results for R_{AVG}, R_{OBS}, R_{P50} and R_{P84} metrics; *red squares* results for R_{F50} and R_{F84} metrics; *black cross* best solution, obtained as the arithmetic mean of all magnitude results

et al. 2017). For each IPE, the magnitude is then computed inverting Eq. (4) as function of M_W, where I_j and R_j are the epicentral intensity (VIII MSK) and the depth fixed a priori, respectively.

The magnitude solutions obtained using the 144 models are summarized in Fig. 10b. Since the depth is fixed a priori, only the variability of the magnitude results as function of the adopted IPEs can be appreciated. Finally, the earthquake magnitude is computed as the arithmetic mean of all individual magnitude solutions, i.e. $M_W = 5.3$. The uncertainty associated to this value is computed following the methodology detailed in Sect. 6.2 and is equal to 0.7.

5.2 Felt strategy: the example of the 1852 Albertville earthquake

Figure 11 illustrates the 1852 Albertville earthquake example. Only seven "felt" testimonies have been collected by SisFrance for this event, together with four "non-felt" testimonies. No epicentral intensity value is therefore provided by SisFrance. Nevertheless, based on expert judgement, Sisfrance provides a location for the epicenter of this earthquake (Fig. 11a). Following the procedure explained in Sect. 2.2, the value of IV MSK is assigned to the "felt" testimonies, and the "Felt strategy" (Fig. 4) is applied. Following the seismogenic depth regionalization scheme proposed by Manchuel et al. the 1852 Albertville earthquake is located within the Alps region. The earthquake depth is then fixed a priori to the value of 7 km, with an associated uncertainty $\sigma_{\log10(h)} = 0.24$ (see Table 1 in Manchuel et al. 2017). For each IPE, the magnitude is then computed inverting Eq. (4) as function of M_W, where I_j and R_j are the quantified felt intensity (IV MSK) and the felt radius (computed as the mean hypocentral distance of "felt" testimonies), respectively.

The magnitude solutions obtained using the 144 models are summarized in Fig. 11b. Since the depth is fixed a priori, only the variability of the magnitude results as function of the adopted IPEs can be appreciated. Finally, the earthquake magnitude is computed as the arithmetic mean of all individual magnitude solutions, i.e. $M_W = 2.9$. The uncertainty

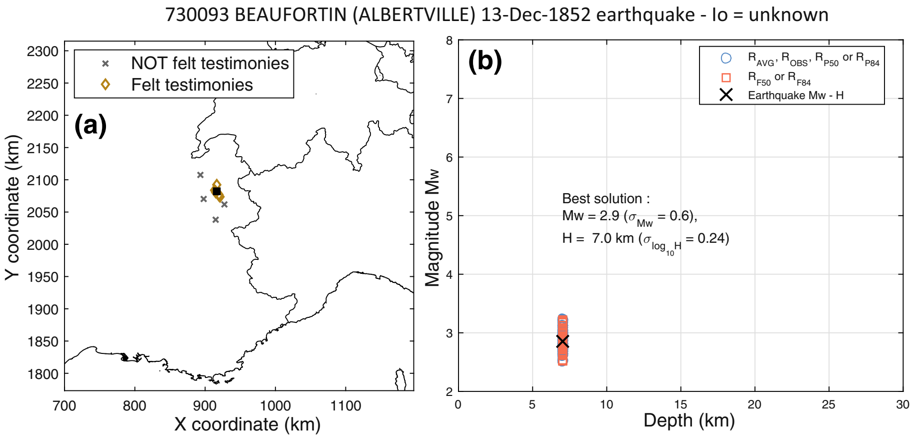

Fig. 11 The 1852 Albertville earthquake (SisFrance reference number: 730093). **a** Map of "felt", "not-felt" testimonies (*gold diamonds* and *black crosses*, respectively) and epicentral location (*black square*). **b** Synthesis of results obtained using Eq. (4) inverted as function of Mw with all selected IPEs; *blue circles*: results for R_{AVG}, R_{OBS}, R_{P50} and R_{P84} metrics; *red squares* results for R_{F50} and R_{F84} metrics; *black cross* best solution, obtained as the arithmetic mean of all magnitude results

associated to this value is computed following the methodology detailed in Sect. 6.2 and is equal to 0.6. Please, note that the magnitude of this earthquake is slightly below the validity range of the Baumont et al. (2017) IPEs. From a perspective of completeness, the seismological parameters of these small earthquakes can be obtained following the same scheme as larger ones. For sake of clarity, however, in the seismic catalog presented by Manchuel et al. (2017) they are assigned a specific flag.

6 The treatment of the uncertainties

6.1 Uncertainties associated to Mw and depths from maximum likelihood inversion scheme

The maximum likelihood inversion scheme described in the previous sections allows us to take into account the uncertainties associated to the estimation of seismological parameters of historical earthquakes, i.e. arising from both, the uncertainties affecting the macroseismic data, and the "inaccuracy" of the IPE used in the computations. Furthermore, the adopted ET overall approach, also allows to characterize the epistemic uncertainty associated to the IPE selection process, choice that usually carries a subjectivity component. Indeed, given the number of IPEs that globally fit the data equally well in a mathematical sense (Baumont et al. 2017), the final choice of the "one" IPE to be used in the seismological parameter computation for French historical earthquakes would have been rather arbitrary. Moreover, as shown by Baumont et al. (2017), the calibration of IPEs demonstrates to be sensitive to both, the chosen parametrization and the calibration dataset. In light of this, the use of a set of models rather than a single one, allows us to take into account this epistemic uncertainty.

As discussed in Sect. 4.1, when macroseismic information available allows performing the WLSQ inversion following either, the "Complete" or the "Simplified ET", the square roots of the diagonal elements of the posterior covariance operator $\widetilde{C}_M$ are interpreted as "uncertainty bars" on the posterior values of the individual model parameters M_{Wi} and H_i. The "Complete" (resp "Simplified") ET therefore produces N solution couples, where N equals 144 (resp 48). The final M_W value is obtained as the arithmetic mean of the N individual magnitude solutions. Similarly, the final H value is obtained as the geometric mean of the N individual depth solutions. As mentioned before, for each magnitude solution, the WLSQ inversion scheme produces an associated uncertainty value, $UncM_{Wi}$. Similarly, the WLSQ inversion produces an uncertainty value associated to each depth solution, $UncH_i$. These uncertainties include (1) the effect of the uncertainties carried by the macroseismic data used in the inversion, (2) the variability of the ith IPE, and (3) the way the ith IPE model fit the data. However, the variability in the magnitude and depth couples in the space of the solutions also needs to be considered in the total uncertainty associated to the final earthquake and depth solutions. Therefore, the two uncertainty components, i.e. the one coming from the macroseismic data, the models and the inversion procedure, and the second being representative of the variability of the obtained solutions, are combined before being associated to the final M_W and H values. The former component is computed as the mean of the N $UncM_{Wi}$ and $UncH_i$, respectively. The latter component is expressed, as the standard deviation of the M_{Wi} distribution for the magnitude, and as the standard deviation of the $\log_{10}H_i$ distribution for the depth.

The two components of the uncertainty are then combined in the following way. Concerning the uncertainty associated to the final magnitude value: a quadratic combination is employed, as follows:

$$Unc_{M_W} = \sigma_{M_W} = \sqrt{\left(mean\left(Unc_{M_{W,i}}\right)\right)^2 + \left(std\left(M_{W,i}\right)\right)^2} \tag{10}$$

On the other hand, a Monte Carlo approach is applied to combine the two components of the uncertainty associated to the final depth value. Both, the posterior probability density function from the depth inversion solutions (defined as a normal distribution centered on the earthquake depth value and having standard deviation equal to the average of standard deviations associated to individual depth solutions), and the distribution of depth solutions issued from the different ET branches (log-normal distribution), are randomly sampled (100,000 random samples from each distribution). The total uncertainty associated to the final depth value (expressed in the log10 scale), is defined as half the difference between the 84th and the 16th percentiles of the sample distribution.

Going back to the examples given in Sects. 4.2 and 4.3, the uncertainties associated to the magnitude and depth estimates are computed as follows.

1. September 1866 Brenne earthquake:
   $mean\left(Unc_{M_{W,i}}\right) = 0.28$, $std\left(M_{W,i}\right) = 0.13$, which give (applying Eq. 10): $Unc_{M_W} = \sigma_{M_W} = 0.31$.
   $mean(Unc_{H_i}) = 6.34$ km, $std(\log_{10} H_i) = 0.05$, which give (applying the Monte Carlo approach described above): $Unc_{\log_{10} H} = \sigma_{\log_{10} H} = 0.10$.
2. September 1804 English Channel earthquake:
   $mean\left(Unc_{M_{W,i}}\right) = 0.61$, $std\left(M_{W,i}\right) = 0.18$, which give (applying Eq. 10): $Unc_{M_W} = \sigma_{M_W} = 0.64$.
   $mean(Unc_{H_i}) = 7.30$ km, $std(\log_{10} H_i) = 0.12$, which give (applying the Monte Carlo approach described above): $Unc_{\log_{10} H} = \sigma_{\log_{10} H} = 0.18$.

6.2 Uncertainties associated to Mw and H estimates for poorly known earthquakes

As described in Sects. 3.2 and 5, when either, the "I_0", or the "Felt" Strategies are applied, the earthquake depth must be fixed a priori. Manchuel et al. (2017) define the a priori depth value as the median of the regional depth distribution, where the depths used to build the distribution are the solutions of the WLSQ inversion proposed in the present paper, i.e. the depths computed for better described historical earthquakes in the considered region. The uncertainties on the a priori depth values are defined as half the difference between the 84th and the 16th percentiles of the distribution in the logarithmic scale, as follows:

$$\sigma_{\log_{10} H_r^*} = \frac{1}{2}\left(P_{84}(\log_{10} H_r) - P_{16}(\log_{10} H_r)\right) \tag{11}$$

where $H_r^* = P_{50}(H_r)$ is the a priori depth value assigned to region r. The associated uncertainty is defined in the logarithmic ($\log_{10}$) scale to be consistent with the observed depth distributions (Manchuel et al. 2017).

Then, for each IPE, the magnitude value is simply computed by inverting Eq. (4) as $M_W = f(I, R)$. For the "I_0 strategy", the epicentral intensity I_0 is used as I value and the a priori depth H is used as R value. For the "Felt strategy", the quantified felt intensity is used as I and the felt radius is used as R value. For each magnitude solution, the associated uncertainty is then computed using the partial derivatives method, as follows:

$$\Delta M_W = \left|\frac{\partial f}{\partial I}\right|\Delta I + \left|\frac{\partial f}{\partial H}\right|\Delta H \tag{12}$$

Therefore, in the case of the "I_0 strategy":

$$\Delta M_{Wi} = \sigma_{M_{Wi}} = \frac{1}{c_{2i}} \cdot \sigma_{EMPEi} + \left(-\frac{\beta_i}{c_{2i}} \cdot \frac{1}{H_r^* \cdot \ln(10)} - \frac{\gamma_i}{c_{2i}}\right) \cdot 10^{Unc_{\log_{10}H_r^*}} \tag{13}$$

and, for the Felt strategy:

$$\Delta M_{Wi} = \sigma_{M_{Wi}} = \frac{1}{c_{2i}} \cdot \sigma_{EMPEi} + \frac{H_r^*}{\sqrt{D_F^2 + H_r^{*2}}} \cdot \left(-\frac{\beta_i}{c_{2_i}} \cdot \frac{1}{\sqrt{D_F^2 + H_r^{*2}} \cdot \ln(10)} - \frac{\gamma_i}{c_{2i}}\right) \cdot 10^{Unc_{\log_{10}H_r^*}} \tag{14}$$

where H_r^* is the a priori depth, common for all earthquakes in a given region, D_F is the felt radius and σ_{EMPEi} is the standard deviation associated to the ith IPE.

7 Discussion and conclusions

In this work we developed a procedure aimed at determining seismological parameters (moment magnitude M_W and depth) for historical earthquakes, based on an ET approach. Depending on the macroseismic information available for the considered earthquake, different cases are identified. For sufficiently well described earthquakes (i.e. when at least two isoseismals can be defined), a simultaneous inversion of M_W and depth is performed, based on a weighted least square maximum likelihood (WLSQ) inversion scheme (Tarantola 2005). For poorly described earthquakes, i.e. when either, only the epicentral intensity, or only sporadic (if any) quantified intensity observations are available, the depth needs to be fixed a priori (regional depth values for the French territory are proposed by Manchuel et al. companion paper) and the magnitude is computed accordingly (by simply inverting Eq. 4) using the epicentral intensity or the Felt isoseismal, respectively. These are the so-called "I_0" and "Felt" strategies, respectively.

In all cases, the uncertainties related to both, the macroseismic data and the model (Intensity Prediction Equation–IPE), are fully taken into account and propagated to the results. In the case of WLSQ inversion, the square roots of the diagonal elements of the posterior covariance matrix are interpreted as uncertainty bars on the seismological parameters. For the "I_0" and 'Felt' strategies, the uncertainty on the a priori depth value is defined as half the difference between the 84th and the 16th percentiles of the regional depth distribution, in the logarithmic scale (see Manchuel et al. for details). The uncertainty on the M_W magnitude is then obtained by propagating the uncertainties on the a priori depth and on the IPEs using the partial derivatives methodology.

In addition to the M_W and depth uncertainties related to the macroseismic data and modelization uncertainties, the ET approach also allows to take into account the epistemic uncertainty related to the choice of the IPE(s) to be used in the computations. Indeed, the methodology we propose in this work allows the operator to consider several different IPEs, each of them representing a branch of the ET. In the present work, four different families of IPEs have been considered, calibrated on the basis of six different earthquake calibration sets and formulated for six different metrics (see Baumont et al. companion paper). It gives a total of 144 IPEs for the "Complete ET" and 48 IPEs for the "Simplified ET" (only the R_{F50} and the R_{F84} metrics are retained in this latter case) that globally suit the French context equally well. The final M_W and depth values are defined as the mean (arithmetic and geometric mean, respectively) of all M_{Wi} and H_i obtained using the different IPEs, and the variability of the results [expressed as standard deviation of M_{Wi} and of $\log_{10}(H)$] is interpreted as epistemic uncertainty, i.e. the variability of the results related to the lack of information, which prevents us from defining the "true" model.

The two components of the uncertainty, i.e. the one coming from the data and the modelization uncertainties, and the variability induced in the results by the use of different IPEs, are then combined to give the total uncertainty associated to the final M_W and depth values. We use a quadratic combination of the two components to determine the total uncertainty associated to the magnitude value, while we use a Monte Carlo approach to obtain the total uncertainty associated to the depth value, expressed as the logarithm ($\log_{10}$) of depth.

The described approach is applied in the companion paper (Manchuel et al. 2017) to estimate M_W and depth of French historical earthquakes, using the whole SisFrance (EDF, IRSN, BRGM) macroseismic database.

Note that, given the inherent uncertainty in the evaluation of the intensity based on very few observations, the reader should be aware of the limitations entailed by the use of the I_0 or Felt strategies on the estimated parameters. In spite of these limitations, however, exploiting the scarce information available is the only possible option to take into account poorly described earthquakes in seismic hazard studies. The uncertainties associated to the parameters estimated using these strategies, which are systematically larger than those associated to estimates obtained through either a "Complete" or a "Simplified" ET (see the catalogue provided as supplementary material to the Manchuel et al. 2017 paper), reflect these limitations. The use of all the selected IPEs also in the case of poorly known earthquakes supports our confidence in the robustness of the obtained results.

A major improvement to the proposed methodology that will be considered in future works is to take into account the uncertainty on the epicentral location provided by SisFrance. Indeed, in the present work, only the uncertainty on the epicentral intensity value, those related to the quality of individual intensity observations, as well as those related to the selection and fitting to the data of IPEs are taken into account. However, as shown by Cecic et al. (1996), the procedure for epicentral parameter determination can be uncertain and variable among different experts. While we expect the uncertainty on the epicentral location to have a negligible effect on the magnitude and depth estimates for well determined epicentres (although some influence on the uncertainty associated to these values is likely to appear), this is most likely not the case for poorly described ancient events with only sporadic intensity data. Indeed, in the SisFrance database, ancient earthquakes described by single (if any) intensity data points, are placed on the locality where the earthquake testimony is collected. However, for scarcely populated areas, the testimony of the earthquake effects might actually be located at tens of kilometres from the real epicenter. This can have a significant impact on the central value of the magnitude (the depth being fixed a priori). Finally, for earthquakes either, located offshore, or presenting an unbalanced scarce macroseismic field, we expect the uncertainty on the epicentral location to have a significant effect on the uncertainty associated to the magnitude and depth estimates, but this effect could be less important on their central values.

To further improve the proposed ET approach, a Bayesian updating technique can be used in the future, for each earthquake, to update the weight of each IPE in the inversion as function of the fit to the data.

Acknowledgements This work was carried out within the SIGMA project. It strongly benefited from the comments and the discussions held within the catalogue work group and the SIGMA scientific committee meetings. The comments of the anonymous reviewers allowed to improve the original version of the manuscript. We would like to thank Fabrice Cotton for interesting and constructive discussions. Merlin Keller and Irméla Zentner provided us with very useful suggestions for improvements. We also acknowledge the researchers from the Dieudonne mathematics laboratory of Nice for the exchanges we had during the initial phase of the work. Finally we would like to thank the data-providers and the seismological observatories for their efforts in collecting and making available earthquake data.

References

Ambraseys NN (1985) Intensity attenuation and magnitude-intensity relationships for Northwestern European earthquakes. Earthquake Eng Struct Dynam 13:733–788

Bakun WH (2006) MMI attenuation and historical earthquakes in the basin and range province of Western North America. Bull Seismol Soc Am 96(6):2206–2220

Bakun WH, Scotti O (2006) Regional intensity attenuation models for France and the estimation of magnitude and location of historical earthquakes. Geophys J Int 164:596–610

Bakun WH, Wentworth CM (1997) Estimating earthquake location and magnitude from seismic intensity data. Bull Seismol Soc Am 87(6):1502–1521

Bakun WH, Johnston AC, Hopper MG (2003) Estimating locations and magnitudes of earthquakes in Eastern North America from modified Mercalli intensities. Bull Seismol Soc Am 93(1):190–202

Baumont D, Scotti O (2011) The French parametric earthquake catalogue (FPEC) based on the best events of the SisFrance macroseismic database-version 1.1. IRSN/DEI/2011-012

Baumont D, Manchuel K, Traversa P, Durouchoux C, Nayman E, Ameri G (2017) Empirical intensity attenuation models calibrated in Mw for metropolitan France. Bull Earthq Eng (**this issue**)

Beauval C, Yepes H, Bakun WH, Egred J, Alvarado A, Singaucho JC (2010) Locations and magnitudes of historical earthquakes in the Sierra of Ecuador (1587–1996). Geophys J Int 181(3):1613–1633

Burton PW, McGonicle R, Nelson G, Musson RMW (1985) Macroseismic focal depth and intensity attenuation for British earthquakes. Earthquake engneering in Britain. Telford, London, pp 91–110

Cecic I, Musson RMW, Stucchi M (1996) Do seismologists agree upon epicentre determination from macroseismic data? A survey of ESC working group ‘macroseismology’. Ann Geophys 39(5):1013–1027

Galanopulos AG (1961) On magnitude determination by using macroseismic data. Ann Geofis 14:225–253

Gasperini P, Ferrari G (2000) Deriving numerical estimates from descriptive information: the computation of earthquake parameters. Ann Geofis 43:729–746

Gasperini P, Vannucci G, Tripone D, Boschi E (2010) The location and sizing of historical earthquakes using the attenuation of macroseismic intensity with distance. Bull Seismol Soc Am 100(5A):2035–2066

Hinzen KG, Oemisch M (2001) Location and magnitude from seismic intensity data of recent and historic earthquakes in the Northern Rhine area, Central Europe. Bull Seismol Soc Am 91(1):40–56

Johnston AC (1996) Seismic moment assessment of earthquakes in stable continental regions-II. Historical seismicity. Geophys J Int 125:639–678

Kövesligethy R (1907) Seismischer Stärkegrad und Intensität der Beben. Gerlands Beitr Geoph 8:24–103

Levret A, Backe JC, Cushing M (1994) Atlas of macroseismic maps for French earthquakes with their principal characteristics. Nat Hazards 10:19–46

Locati M, Rovida A, Albini P, Stucchi M (2014) The AHEAD portal: a gateway to european historical earthquake data. Seismol Resour Lett 85(3):727–734

Manchuel K, Traversa P, Baumont D, Cara M, Nayman E, Durouchoux C (2017) The French seismic CATalogue (FCAT-17). Bull Earthq Eng (**this issue**)

Mayor J, Traversa P, Margerin L, Calvet M (2017) Tomography of crustal seismic attenuation in metropolitan France: implications for seismicity analysis. Bull Earthq Eng. doi:10.1007/s10518-017-0124-8

Musson RMW (1996) Determination of parameters for historical British earthquakes. Ann Geofis 39:1041–1047

Scotti O, Levret A, Hernandez B (1999) Détermination des caractéristiques des séismes pour les études d'aléa sismique: comparaison des évaluations macrosismiques et instrumentales, 5ème Colloque National de l'AFPS, ENS Cachan (France), pp 42–47

Sibol MS, Bollinger GA, Birch JB (1987) Estimations of magnitudes in central and eastern North America using intensity and felt area. Bull Seismol Soc Am 77:1635–1654

Sponheuer W (1960) Methoden zur Herdtiefenbesttimmung in der Makroseismik, Freiberger Forschungsthefte C88. Akademie Verlag, Berlin, p 117

Stucchi M, Rovida A, Gomez Capera AA, Alexandre P, Camelbeeck T, Demircioglu MB, Gasperini P, Kouskouna V, Musson RMW, Radulian M, Sesetyan K, Vilanova S, Baumont D, Bungum H, Fäh D, Lenhardt W, Makropoulos K, Martinez Solares JM, Scotti O, Živcic M, Albini P, Batllo J, Papaioannou C, Tatevossian R, Locati M, Meletti C, Viganò D, Giardini D (2013) The SHARE European earthquake catalogue (SHEEC) 1000–1899. J Seismol 17(2):524–544

Tarantola A (2005) Inverse problem theory and methods for model parameter estimation. SIAM, Philadelphia, p 342

Toppozada TR (1975) Earthquake magnitude as a function of intensity data in California and Western Nevada. Bull Seismol Soc Am 65:1223–1238