Bulletin of the Seismological Society of America, Vol. 107, No. 3, pp. 1136–1146, June 2017, doi: 10.1785/0120160204

On the Epicenter Determination of Historical Earthquakes Attested to by Sparse Intensity Data Sets

by Päivi Mäntyniemi, Ruben E. Tatevossian, and Roman N. Vakarchuk

Abstract Historical earthquakes have to be parameterized for seismic-hazard analyses, although there may be only a few intensity assignments available for them. We studied epicenter determination for 18 million synthetic samples of 3–11 intensity data points (IDPs). The IDP distributions corresponded to earthquakes that occurred offshore, close to the coast, or onshore. We assumed an ordinal variable, an attenuation relationship, and a point source. The attenuation relationship was utilized to encompass every IDP of a sample using a lower and upper radius that corresponded to the respective intensity. The epicenter must fit all the intensity rings simultaneously. The successes and failures of epicenter determination were monitored for a fixed magnitude and depth. We investigated where the epicenter was found, its uncertainty, and its uniqueness. Small location uncertainties may be obtained for the smallest samples but increasing the sample size led to a larger proportion of small uncertainties provided that intensities were error-free. A large range of intensities in the sample, a short distance to the true epicenter, and, to a lesser extent, a small azimuthal gap were indicators of a good solution. A location uncertainty of 20 km and smaller is realistic in many cases, but uncertainties of 5 km are extremely seldom occurrences. The proportion of good locations was reduced when the intensities were erroneous. Epicenters were determined for the earthquakes of 26 April 1458 in central Italy, 14 July 1765 in Sweden, and 23 December 1875 in the eastern United States.

Introduction

Parametric earthquake catalogs (PECs) are important inputs for seismic-hazard analyses. Knowledge of long-term seismicity has a direct bearing on insights into current seismic hazards, thus PECs typically cover the noninstrumental and instrumental eras. The end users of PECs may be unaware of the fact that deriving parameters for earthquakes stemming from the different eras entail entirely different procedures. For historical earthquakes, seismic intensity values (hereafter, intensities) are estimated on the basis of documented evidence, and the earthquake magnitude, epicenter, and possibly also the depth are derived from intensity data. Written documents and earthquake parameters may appear incompatible, but parameters are customarily derived from sets of intensity data when compiling PECs.

Concern has been expressed over the lack of standardization in the determination of seismic parameters for historical earthquakes. Stucchi (1994) pointed out some pitfalls in the parameterization: Entries of PECs look homogeneous, but may be based on inadequate analysis of data and unsystematic assessments. Cecić *et al.* (1996) showed that seismologists sometimes establish their own approaches to derive epicenter coordinates for historical earthquakes. The differences between individual solutions become more pronounced when the sets of intensity data are sparse or irregularly distributed. Castelli and Monachesi (1996) derived reasonable alternative epicenter locations for historical earthquakes in central Italy. The epicenter coordinates carried forward into the PEC may be selected subjectively. Formalized, independent methods for assessing macroseismic earthquake parameters in a consistent way were developed by Bakun and Wentworth (1997, 1999), Gasperini *et al.* (1999, 2010), and Musson (2009). These approaches are a prerequisite for processing intensity data for large international initiatives such as Musson and Jiménez's macroseismic estimation of earthquake parameters (see Data and Resources, Deliverable D3).

Uncertainties must be associated with the parameters of historical earthquake(s) as a prerequisite for studies of regional seismotectonics and seismic hazard. The uncertainties given in the PEC may not only be useless, but actually misleading due to the potential bias introduced from assessments by individual researchers. Tatevossian and Mäntyniemi (2014) proposed past earthquake scenarios to derive parameter uncertainty when the sets of intensity data are sparse. The first-order task is to design scenarios that fit the data and associate them with a probability using expert judgment. The second-order task is to assess the uncertainties

with each scenario. Thus, the uncertainty of the parametric solution becomes discrete instead of continuous, and several solutions are available for one historical earthquake in the PEC, which makes the selection more transparent.

Additional information is sometimes required for the analysis of historical earthquakes. When assumptions about the integrity of the documents and inference from regional seismotectonics are used to expand the existing textual information for a set of earthquake parameters, data and conjecture can become mixed and the result may have very little to do with what actually happened in the past (Musson, 1998). Additional information can be imported in a more formalized manner. A rather poorly documented historical earthquake in a given region can be compared with a well-documented, so-called model earthquake from the same region. If many data points of the two events are alike, there may be reason to suggest that they could also have similar parameters. For example, Gutdeutsch and Hammerl (1999) compared the 25 January 1348 Villach earthquake with the well-documented 6 May 1976 Friuli (Italy) earthquake, and found that the 1976 event can be used as a model for the 1348 earthquake. Houng and Hong (2013) used the instrumental seismicity record to create a probability density function that is utilized to locate historical earthquakes. In this case, instrumental seismicity can be understood as model seismicity. Although the model earthquake concept can be helpful, it is obviously difficult to apply it to very small datasets. If the seismicity in a given region is low or poorly documented, there may not be many candidates available to use as model earthquakes.

Sparse data are a limitation when compiling PECs, but the adequacy of data has not been studied in much detail. Gasperini *et al.* (2010) excluded all earthquakes that had fewer than 20 intensity data points (IDPs). Bakun *et al.* (2011) compared epicenter and magnitude determinations in Italy using earthquakes with many intensity assignments, and pointed out that more work was needed to establish how the parameters can be determined if few intensity assignments were available.

We investigated epicenter determination in the present study using sparse sets of IDPs. Different epicenter determinations may fit the data for geometrical reasons, and consequently different expert judgments may have equal validity. We used 18 million samples of synthetic IDPs to obtain statistical insight into the issue of accurate epicenter determination. The uncertainties associated with epicenters of historical earthquakes can be large (Musson, 1989), and it is difficult to quantify them using real data. The successes and failures of epicenter determination can be monitored using synthetic data for a fixed magnitude and depth. We studied where the epicenter was found, its uncertainty and uniqueness, and the respective data properties. Case histories on the earthquakes of 26 April 1458 in central Italy, 14 July 1765 in Sweden, and 23 December 1875 in the eastern United States are presented to validate the approach. The inverse problem was solved for different combinations of input intensities, magnitudes, depths, and attenuation coefficients. The location accuracy attainable for sparse sets of intensity data is discussed.

Determination of Macroseismic Epicenter

We focus on earthquakes that can be modeled as point sources. The inverse problem is solved for randomly selected samples and a minimum number of assumptions (an ordinal variable, attenuation equation). This allows us to monitor the successes and failures of epicenter determination by utilizing only the information of the IDP configurations.

Many candidate epicenters that produce the given intensity of an IDP for a given magnitude and focal depth can be found within the ring that encompasses it. The upper and lower radius of the ring is calculated using the intensity attenuation equation. The epicenter cannot be determined using a single IDP. If, however, there are more IDPs with corresponding rings, the true epicenter fits the intensities at all localities simultaneously, provided that intensity assignments are not contaminated by errors (Fig. 1).

Two criteria were adopted to investigate the epicenter determination. First, the fixed epicenter has to be found within the area constrained by the maximum number of overlapping rings. Each point within the area of overlap is the epicenter with equal probability. All rings will overlap when intensities are free from error. If there is error in the intensities, not all of the rings will overlap, and the area where the maximum number of rings overlap can be selected as a candidate epicenter (Fig. 2a). In such cases, however, the rings may overlap purely by chance at a point that is not the true epicenter, which will produce a false epicenter (Fig. 2b). This problem cannot be resolved without additional information. The size of the area constrained by the maximum number of overlapping rings corresponds to the uncertainty. The area may be of irregular shape.

The second criterion is the uniqueness of the epicenter. The rings may overlap at more than one place. For example, two possible epicenter scenarios S1 and S2 may emerge (Fig. 1). This can happen for geometrical reasons. In such cases, the total area of overlap corresponds to the location uncertainty. The square root of the area can alternatively be used to handle smaller numerical values. In the following, the location uncertainty refers to the square root.

Drawing intersecting circles surrounding observation points resembles classic instrumental epicenter determination using three seismic stations and radii equal to the epicentral distances calculated from the *S*–*P*-wave arrival times. If seismic-wave velocities are not known exactly, then the instrumental exercise also involves rings rather than circles surrounding the observation points.

Tatevossian *et al.* (2013) and Radziminovich (2014) used circles to encompass IDPs related to historical earthquakes for which only sparse sets of intensity data are available. Intensity is an ordinal variable; therefore the utilization of rings instead of circles is advocated. In cases of real data, neither the magnitude nor the epicenter is known, thus different possible combinations of magnitude and epicenter are fitted to the available IDPs. The input intensities, magnitudes, and depths can be allowed to float. The investigation

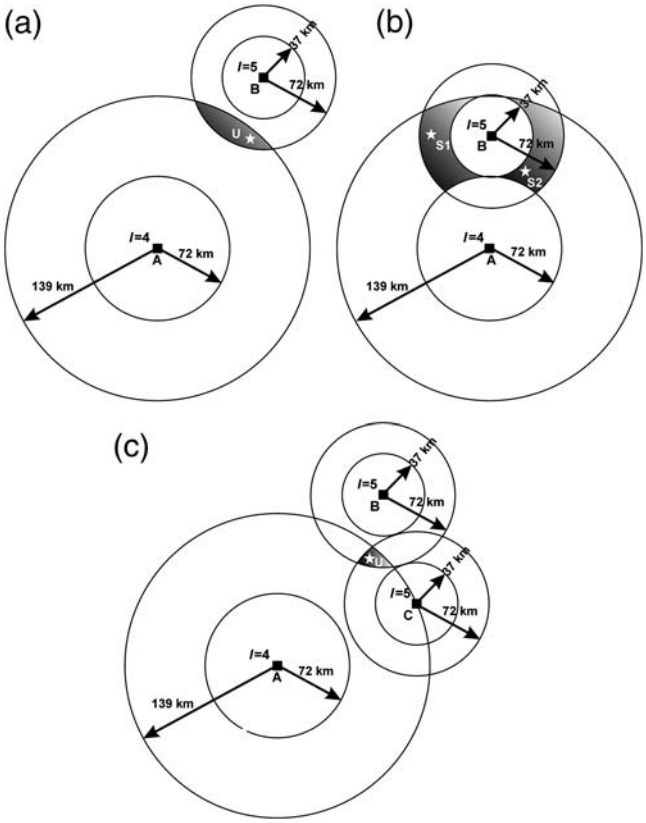

Figure 1. Epicenter determination for sparse sets of intensity data. The black squares demark localities A, B, and C with intensities 4, 5, and 5, respectively. The arrows show the lower and upper radii that encompass the localities that were computed using an intensity attenuation relationship. The shaded areas constrained by the overlapping rings contain all possible epicenters that produce the corresponding intensities, and their sizes correspond to the location uncertainty.

can be focused on a specified range of input parameters to limit the number of inversion processes.

Generation of Synthetic Data for Stability Tests

We generated synthetic macroseismic fields by placing the earthquake epicenter in the middle of an area with sides measuring 5°. The focal depth was fixed at 5 km, and the earthquake magnitude at M_L 5. We generated coordinates of 500 localities in the area randomly (Fig. 3a). Most of the localities can be found within the distance range of 100–275 km from the epicenter. The shortest distance was 9.7 km, and the farthest was 385 km.

We created two sets of synthetic intensities. First, the intensity I at locality i was computed using an intensity attenuation equation and rounded to the nearest integer; these data are referred to as the rounded intensities. The Shebalin equation (Kondorskaya and Shebalin, 1982) we used is

$$I_i = 1.5M_L - 3.5 \times \log \sqrt{R_i^2 + H^2} + 3, \qquad (1)$$

in which, $i = 1, \ldots, 500$, M_L is the local magnitude, R_i the epicentral distance to locality i, and H the focal depth, both in kilometers, and 3.5 the attenuation coefficient. Intensities are given on the Medvedev–Sponheuer–Karnik (MSK) scale.

The data quality was simulated by a second dataset. An error of +1 unit was added to intensities at localities 376–500, −1 to intensities 1–125, and intensities 126–375 were kept error-free. This dataset is referred to as the intensities with error. The errors can result from different practices or mistakes made in the assignment of intensity, or from site effects that amplify or lower the intensities for sites at the same distance to the epicenter. Anomalously high or low intensities often stand out from large intensity datasets, but are less obvious in sparse datasets. The biases in intensity data should be taken into account when deriving intensity attenuation relationships (Bakun and Scotti, 2006), but the topic is beyond the scope of the present study.

The distributions of the calculated intensities show that there are fewer intensities 2 than 3, which indicates that most intensities 2 are located outside of the map frame (Fig. 3b). Intensity $I = 2$ is difficult to determine, and rarely available for historical earthquakes. The distribution of intensities with error is smoother, because the number of intensities $I = 3$ decreased, and the range of intensities became wider.

We created one million samples for each number of IDPs. Synthetic samples with 3–11 IDPs were selected randomly from the 500 fixed localities. The smallest sample size was 3, because it is not possible to pick up a million different samples with only 2 IDPs out of 500 localities. The azimuthal gap, the shortest distance to the epicenter, and the intensity range of the 18 million synthetic data samples were calculated (Table 1).

The high number of simulated samples of IDPs covers a wide range of macroseismic fields that correspond to earthquakes that were offshore, close to the coast, or onshore. For example, a sample with an azimuthal gap of 180° can be taken to represent a coastline. A sample that includes the true maximum intensity can be taken to originate from newspaper reports that tend to state the extreme effects (Ringdal *et al.*, 1978). Original documents can be lost through fires, wars, neglect, etc. over time, regardless of their content. The incomplete macroseismic fields represent the information that survived from the past to the modern seismologist.

Results

Epicenter Determination for Rounded Intensities

The fixed epicenter was found within the area that was constrained by the maximum number of overlapping rings using all 9 million samples of the rounded intensities. The obtained range of uncertainties, however, is wide and the solutions are not always unique. Table 2 shows that the mean values of the location uncertainty decreased systematically as the sample size increased. The mean is about 113 km for sample size 3, and about 25 km for sample size 11. The

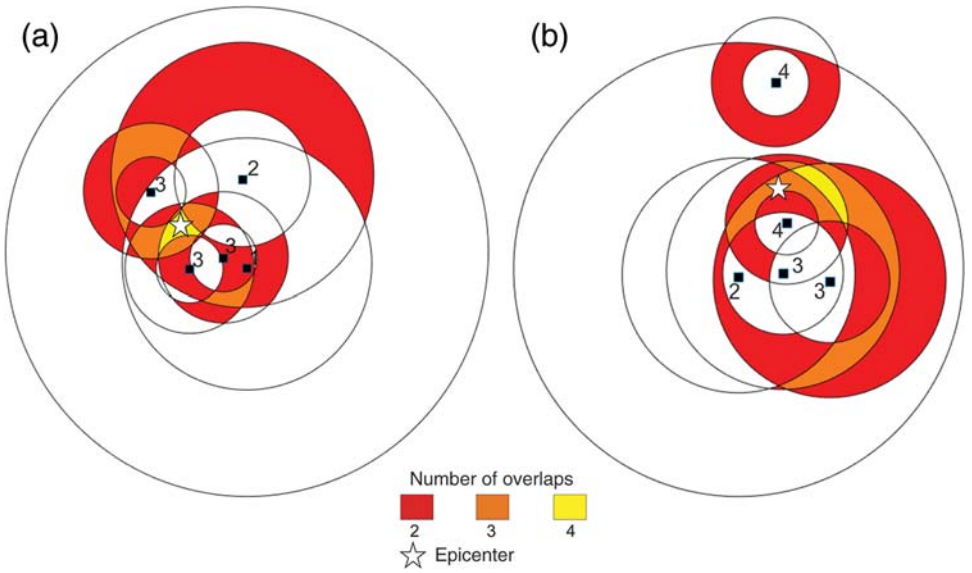

Figure 2. The criterion for the epicenter. Squares depict localities, and the numbers signify the corresponding intensities. Colors mark places where two, three, or four rings overlap. (a) The maximum number of overlapping rings is four, and the epicenter can be found in the area of overlap. (b) The maximum number of overlapping rings is four, but the epicenter does not coincide with the area constrained by the overlapping rings, producing a false solution.

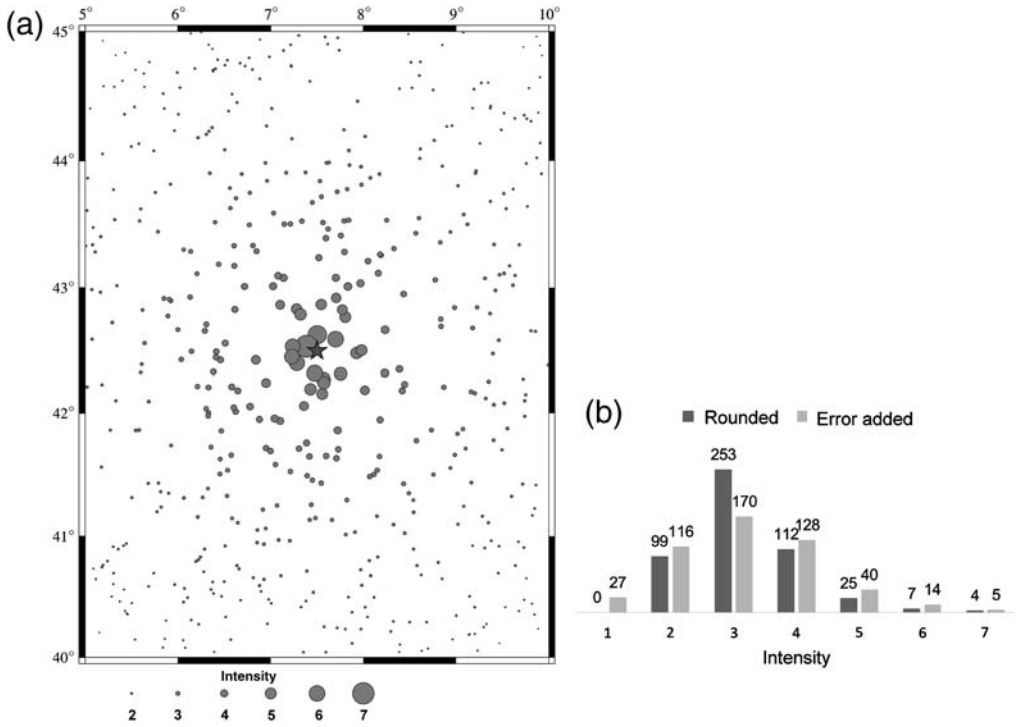

Figure 3. (a) A map of the 500 localities generated randomly and used to obtain samples of intensity data points (IDPs) randomly. The star demarks the fixed epicenter. (b) The numbers of synthetic intensities created for 500 localities. The dark columns depict the intensities that were rounded to the closest integer, and the light columns represent intensities containing errors.

uncertainty distributions have long tails. The skewness was at its maximum 1.71 in case of 3 IDPs, and in the 0.79–0.85 range in cases of 6–11 IDPs. A more confined distribution indicates that the proportion of small location uncertainties increased. About 2.8% of the uncertainties in the category of 4 IDPs were at 20 km or less (Fig. 4a), whereas in case of 9 IDPs the proportion was 27.8% (Fig. 4b).

The maximum number of overlapping rings equaled the sample size. The solution was not always unique in all 9 million cases, but the proportion of nonunique solutions was reduced

Table 1
Properties of the Synthetic Data Samples

	IDP	Minimum Value	Mean	Standard Deviation	Maximum Value	Median	Skewness
AG (°)	3	120.11	230.13	53.03	359.92	223.72	0.31
	4	91.47	201.07	49.80	358.30	193.75	0.51
	5	75.25	180.44	46.32	354.41	174.35	0.59
	6	63.70	165.07	43.30	346.88	159.43	0.62
	7	55.45	153.17	40.89	345.58	148.42	0.64
	8	51.63	143.30	38.69	334.43	138.70	0.64
	9	49.84	135.35	37.02	329.43	131.21	0.63
	10	42.27	128.46	35.48	318.50	124.53	0.63
	11	42.04	122.61	34.17	312.02	118.63	0.62
CD (km)	3	9.70	120.93	57.92	345.94	114.34	0.42
	4	9.70	106.03	51.27	340.04	100.86	0.48
	5	9.70	95.61	46.45	315.05	90.95	0.50
	6	9.70	87.91	42.63	292.69	83.48	0.49
	7	9.70	81.88	30.76	275.46	77.92	0.48
	8	9.70	76.92	37.49	275.87	74.06	0.46
	9	9.70	72.84	35.53	294.77	70.92	0.45
	10	9.70	69.38	33.94	259.31	68.23	0.43
	11	9.70	66.40	32.55	246.15	65.87	0.42
IR	3R*	0	1.39	0.98	5	1	0.90
	4R	0	1.71	1.00	5	2	0.81
	5R	0	1.96	1.02	5	2	0.76
	6R	0	2.15	1.01	5	2	0.72
	7R	0	2.31	1.01	5	2	0.71
	8R	0	2.45	1.01	5	2	0.69
	9R	0	2.57	1.01	5	2	0.68
	10R	0	2.67	1.01	5	2	0.66
	11R	0	2.76	1.00	5	3	0.64
	3E*	0	1.92	1.12	6	2	0.62
	4E	0	2.34	1.12	6	2	0.52
	5E	0	2.65	1.12	6	3	0.46
	6E	0	2.89	1.10	6	3	0.42
	7E	0	3.09	1.09	6	3	0.38
	8E	0	3.25	1.09	6	3	0.34
	9E	0	3.40	1.08	6	3	0.30
	10E	0	3.52	1.07	6	3	0.27
	11E	0	3.64	1.06	6	4	0.23

IDP, intensity data point; AG, azimuthal gap; CD, the shortest distance to the epicenter; IR, the range of intensities in the sample.

*The letters R and E stand for intensities rounded to the closest integer and for intensities with error, respectively.

with an increasing number of IDPs. For example, in the category of 3 IDPs, 18.4% of the solutions were nonunique, that is, two or more scenarios fitted the data, whereas in the 10 IDPs category only 2% of the solutions were nonunique.

Effect of Data Properties for Rounded Intensities

The properties of synthetic samples can be compared with the obtained location uncertainty. In the case of 5 IDPs, the azimuthal gap of a sample did not seem to be a strong indicator of small location uncertainty (Fig. 5a), although the uncertainties decreased slightly with decreasing azimuthal gap. About 7.4% of the uncertainties were found at 20 km or less when the azimuthal gap was 120° or less, and this proportion dropped to 3.1% when the azimuthal gap was between 300° and 360°.

Table 2
Location Uncertainty (in Kilometers) Parameters for the Synthetic Dataset Comprising Rounded Intensities

IDP	Minimum Value	Mean	Standard Deviation	Maximum Value	Median	Skewness
3	1.28	112.99	59.71	782.87	105.64	1.71
4	1.28	77.84	39.47	743.99	74.27	1.34
5	0.95	59.18	29.29	634.42	55.26	1.01
6	0.84	47.87	23.57	330.67	45.06	0.85
7	0.84	40.33	19.98	302.82	38.31	0.82
8	0.84	34.74	17.36	212.82	32.66	0.79
9	0.84	30.74	15.46	345.97	29.48	0.82
10	0.84	27.49	13.87	139.31	25.64	0.82
11	0.84	24.93	12.69	129.28	22.99	0.84

The closer at least one IDP of a sample was located to the epicenter, the larger the proportion of small uncertainties for the sample size 9 IDPs (Fig. 5b). When the shortest distance was at 25 km or less, all uncertainties were found at 50 km or less, and 73.9% of these were found within 20 km or less. The proportion of rather good solutions increased for cases with the shortest distance above 250 km, but there are only five of them, so that case was statistically unrepresentative.

A large range of intensities in a sample seems to imply a larger proportion of small uncertainties. In the case of 6 IDPs and an intensity range equal to 5, all uncertainties were found at 50 km or less, and 52.2% of these were 20 km or less. When the intensity range was 1, only about 4% of the uncertainties were found at 20 km or less (Fig. 5c).

The summaries for the rounded intensities show that small location uncertainties may also be obtained in small samples, but a larger sample size leads to a larger proportion of small uncertainties (Fig. 6). Although the proportions increase, the absolute numbers of cases do not always increase. For example, there was a low chance for the azimuthal gap to be large in cases of 8–11 IDPs; consequently, the respective proportions were not considered representative (Fig. 6a). A short distance to the epicenter indicates a larger proportion of good epicenters for all sample sizes (Fig. 6b). The range of intensities is a good indicator of small epicenter uncertainty. If the range of intensities in the sample is 5, the proportion of uncertainties for 20 km or less was found to be 20.8% for 3 IDPs, but 79.1% for 10 IDPs (Fig. 6c).

The Role of the Azimuthal Gap

The azimuthal gap tells us how evenly the available IDPs are distributed, whereas the range of intensities corresponds to the distances between the localities. The range of intensities and the shortest distance to the epicenter are related in the sense that if all intensities in a sample located far from the epicenter are equal, they are probably more low rather than high intensities. It is reasonable that the chance

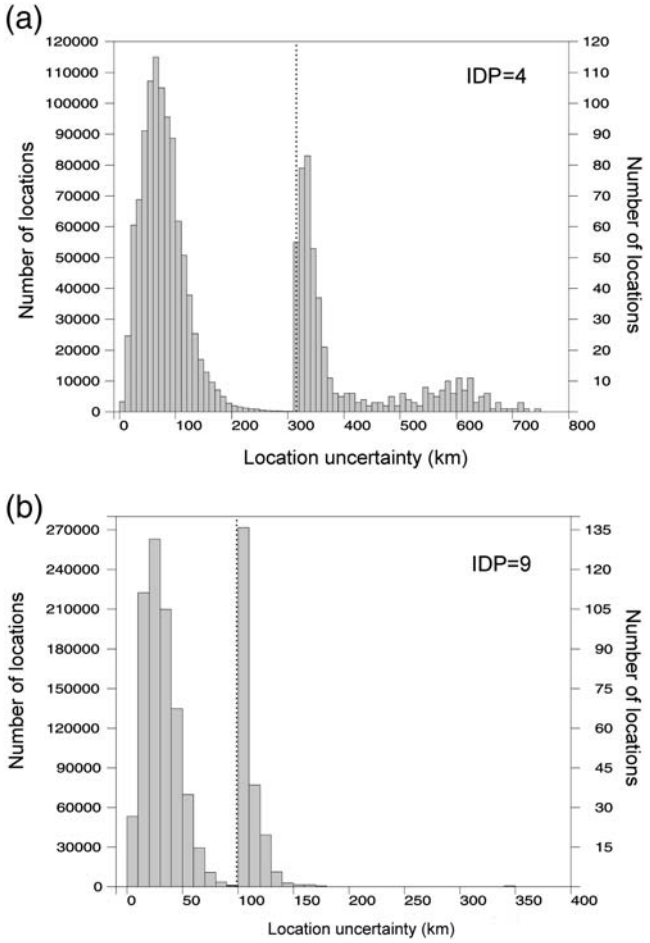

Figure 4. Distribution of the location uncertainty in the case of (a) 4 and (b) 9 synthetic IDPs. The largest uncertainties are 744 and 346 km, respectively. The long tails of the uncertainty distributions have been enhanced by the two *y*-axis scales at different sides of the dotted line. This better illustrates the number of uncertainties that have very large values. The total number of locations is one million for both cases.

of a small location uncertainty increased when a larger part of the affected area was covered. A heavily clustered sample with a narrow range of intensities cannot yield a small location uncertainty or a unique solution. The largest uncertainties typically resulted from samples that are composed of low intensities $I = 2$ and $I = 3$.

Epicenter Determination for Intensities with Error

The maximum number of overlapping rings did not always equal the sample size. However, the larger the sample size, the bigger the proportion of cases in which the epicenter was found in the area of overlap. This finding held true for 87.6% of epicenters in the case of 3 IDPs, and for 99.9% of epicenters in cases of 10 and 11 IDPs (Table 3). However, many formally acceptable epicenters were associated with impractically large uncertainties. The distributions of location uncertainty were more skewed than those found for the rounded intensities, and the skewness increased with the sample size. The largest location uncertainties were typically found for samples with erroneous intensities.

Increasing the sample size had only a weak influence on the proportion of small uncertainties. The clear patterns observed for rounded intensities, shown in Figures 5 and 6, deteriorated significantly. For intensity range 5 and 3 IDPs, 3.7% of the location uncertainties were 20 km or less, and for 11 IDPs the corresponding figure was 27.9%. For shortest distance to the epicenter equal to 25 km or less, 6.0% of the uncertainties were found to be 20 km or less for samples of 3 IDPSs, and 27.9% for 11 IDPs. The corresponding proportions for rounded intensities were 20.6% for 3 IDPs and 83.1% for 11 IDPs. The chance of a small uncertainty was significantly reduced with the quality of intensities. When intensities contain errors, it may also happen by chance that the rings overlap in an area that does not contain the true epicenter, or that the epicenter might be found within the area constrained by overlapping rings related to erroneous intensities.

Figure 7 shows the influence of intensity errors on the proportion of cases in which the epicenter was found within the area of overlapping rings. The proportion of intensities with errors in the sample reduced the proportion of such locations rather dramatically. More than one-third (38.7%) of the locations succeeded for 4 IDPs (3 of them without error, i.e., 75%); in the 7 IDPs category (6 without error, i.e., 86%), the proportion was 63.3 and in case of 10 IDPs (9 without error, i.e., 90%), the proportion was 74.1%.

Validation Tests Performed on Three Case Histories

The 26 April 1458 Earthquake in Central Italy

We investigated the determination of the epicenter for the earthquake of 26 April 1458 in Val Tiberina, central Italy. The data were retrieved from the 2011 version of the intensity database of damaging earthquakes in the Italian region edited by Locati, Camassi, and Stucchi (see Data and Resources). There were five IDPs: Cittá di Castello (VIII–IX), Sansepolcro (VIII–IX), Montone (VII–VIII), Gubbio (V), and Perugia (IV–V). The intensities were on the Mercalli–Cancani–Sieberg (MCS) scale. The moment magnitude given was M_w 5.78 ± 0.34. The attenuation equation (1) was used. The finding for attenuation coefficient 4 was similar to that of the Italian attenuation equation derived by Pasolini *et al.* (2008). The intensity degrees between MSK assumed by equation (1) and the MCS scale were found to be similar (Musson *et al.*, 2010).

Intensities were allowed to float within the reported range. For example, the intensity at Cittá di Castello was taken to be either VIII or IX. Thus, there were a total of 768 combinations of intensity. Focal depths between 5 and 15 km in increments of 2 km, and moment magnitudes from 5 to 6 in increments of 0.1 magnitude units were used to limit the number of iterations. Values 3.5, 3.75, or 4 were given to the

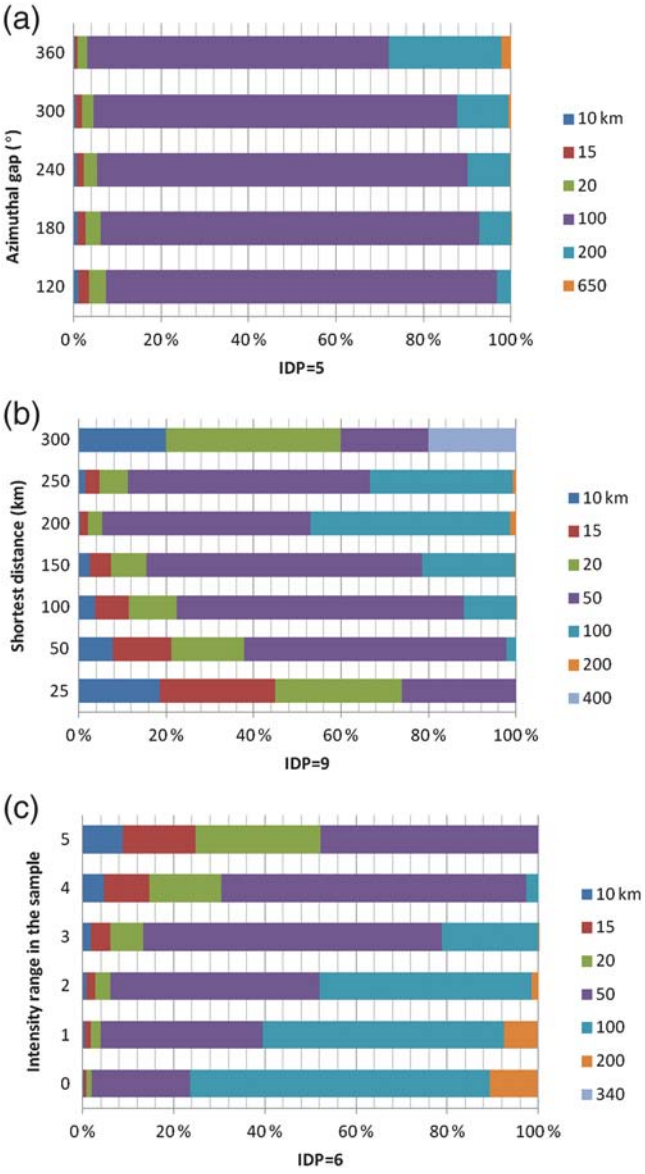

Figure 5. Influence of synthetic data features on the location uncertainty. The intensities are rounded off to the nearest integer. In all cases, the x axis is the proportion of locations that belong to the uncertainty class that is indicated by color: blue is an uncertainty less than 10 km, red is an uncertainty above 10 km and below 15 km, and so on. The largest classes have different uppermost values in the three cases. (a) The y axis is the azimuthal gap in the case of 5 IDPs. The first class is an azimuthal gap 120° or less, the next class is azimuthal gap above 120° and below 180°, and so on. (b) The y axis is the shortest distance to the epicenter in case of 9 IDPs. The first class is the shortest distance 25 km or less, 50 for the shortest distance above 25 and below 50 km, and so on. (c) The y axis is the intensity range in the sample in the case of 6 IDPs.

attenuation coefficient of equation (1). There were 633 solutions. Figure 8 illustrates the epicenter determinations related to magnitudes M_w 5.8, 5.9, and 6.0 for the attenuation coefficient equal to 4 and at a focal depth of 5 km. They are located 5, 9, and 12 km, respectively, from the epicenter placed at Cittá di Castello according to the Italian database.

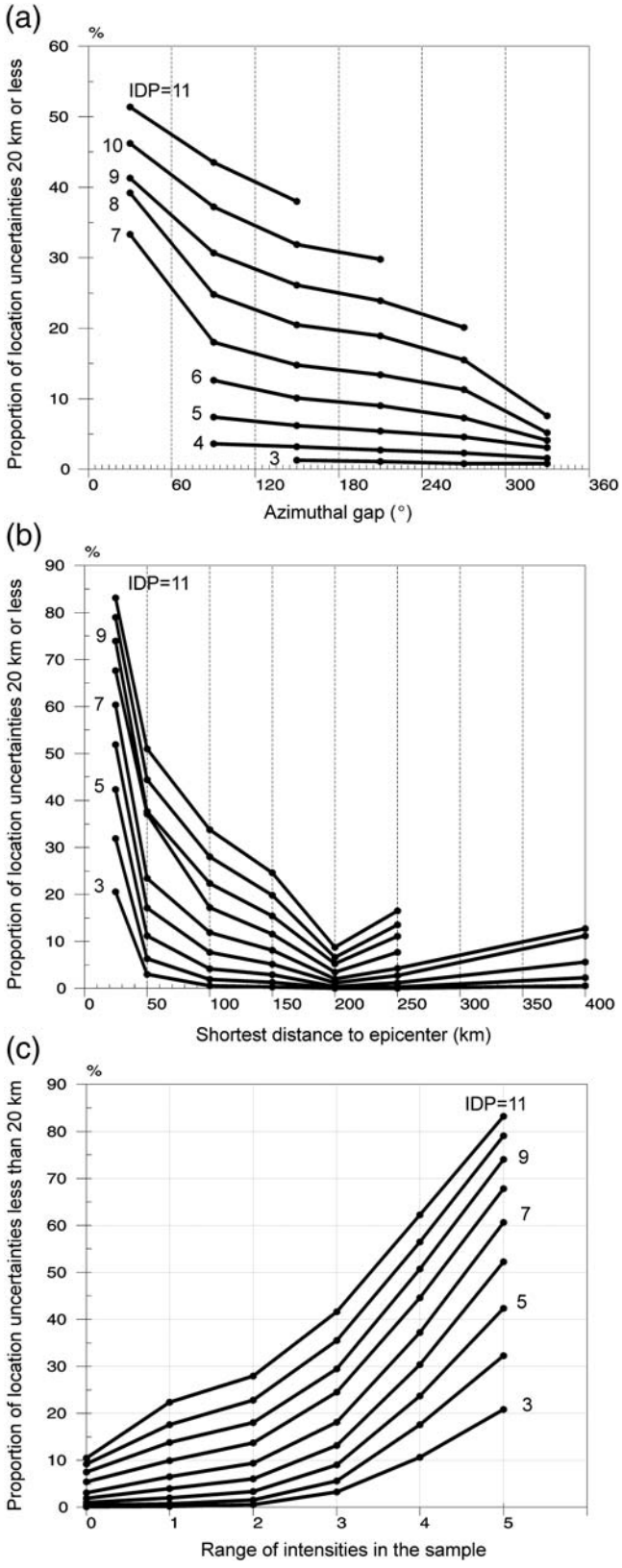

Figure 6. Summaries of the influence of data features on the location uncertainty for different sample sizes of rounded synthetic intensities: (a) azimuthal gap (°), (b) shortest distance to the epicenter (km), and (c) range of intensities in the sample. The intensities do not contain error.

Table 3
Location Uncertainty (in Kilometers) Parameters for the Synthetic Dataset Comprising Intensities with Error

IDP	Minimum Value	Mean	Standard Deviation	Maximum Value	Median	Skewness	Number (Locations)*
3	5.22	355.61	277.78	1055.42	252.75	1.25	876,024
4	1.58	277.41	250.36	1055.42	187.99	1.73	938,652
5	0.95	217.29	218.07	1055.42	150.31	2.28	969,473
6	0.84	171.39	184.15	1055.42	116.41	2.91	984,930
7	0.84	137.05	152.01	1055.42	91.43	3.62	992,577
8	0.84	111.80	125.52	1055.42	80.20	4.42	996,372
9	0.84	92.61	101.03	1055.42	71.59	5.29	998,181
10	0.84	78.68	82.34	1055.42	65.50	6.14	999,096
11	0.84	68.07	67.07	1055.42	60.89	6.97	999,542

*This is the number of successful locations.

The 14 July 1765 Earthquake in Sweden

The epicenter was determined for the earthquake of 14 July 1765 that occurred in the Swedish realm in northern Europe. The six IDPs were Skellefteå (V), Piteå (IV), Luleå (III), Pyhäjoki (III), Kalajoki (III), and Tornio (I) from Mäntyniemi (2012). The intensities were on the European Macroseismic scale, according to Musson *et al.* (2010) similar to MSK intensities assumed by equation (1). The attenuation coefficient was taken to be 3 for the shield region. The five felt intensities were allowed to float within ±1 units, and the epicenter was found for the reported intensities (Fig. 9). It was located 50 km northwest of the epicenter given by Ahjos and Uski (1992). The corresponding magnitude was M_L 4.0. For faster attenuation, coefficient value 3.5, the magnitude increased to M_L 4.6, and the epicenter uncertainty decreased (Fig. 9). Depths from 5 to 10 km fit these epicenters. The main population centers were located along the coast, so it cannot be inferred how far inland the area of perceptibility extended. We preferred the slower attenuation and smaller magnitude M_L 4.0. The earthquake was reportedly not felt in Tornio, which indicates that it was not a particularly large-magnitude event.

The 23 December 1875 Earthquake in the Eastern United States

We also investigated the epicenter determination for the earthquake of 23 December 1875 in the eastern United States. An intensity dataset of eight IDPs was available at the U.S. Earthquake Intensity Database of the National Oceanic and Atmospheric Administration (see Data and Resources). The intensity attenuation equation by Bakun *et al.* (2003) was used:

$$\mathrm{MMI} = 1.41 + 1.68 \times \mathbf{M} - 0.00345 \times \Delta - 2.08 \log(\Delta), \quad (2)$$

in which MMI is modified Mercalli intensity, **M** is moment magnitude, and Δ is the distance from the epicenter (in kilometers).

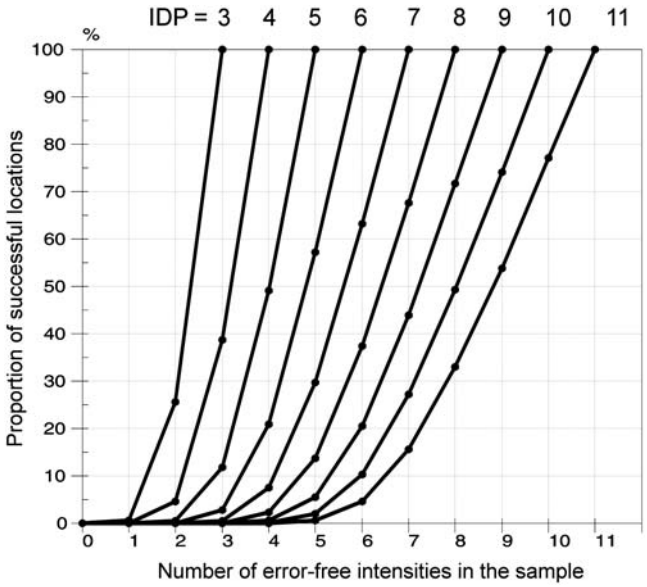

Figure 7. The proportion of successful locations (*y* axis) versus the number of error-free intensities in the sample (*x* axis). Successful is defined as when the true epicenter is found within the area of overlapping rings. The nine curves refer to samples with 3–11 IDPs.

The database does not give a magnitude for this earthquake. Moment magnitudes between 5.8 and 6.4, with increments of 0.2 units, were therefore used as inputs. The eight intensities were allowed to float within ±1 units, which gave 2187 intensity combinations. There was no solution obtained using the initial intensities. There were 57 solutions for modified intensities that corresponded to magnitudes between M_w 5.8 and 6.2. Three solutions were picked out (Fig. 10). The first epicenter corresponded to the minimum number of modifications (3) to the initial intensities. The other two illustrated solutions corresponded to cases with the reported maximum intensity VII taken to be the true maximum.

The three illustrated epicenters differ from the database epicenter by 220, 100, and 110 km, respectively. The solutions 2 and 3 are located within an area of larger intensities, but epicenter 1 is located southward. There were three initial intensity values equal to VI: Manakin (epicentral distance from the database epicenter equal to 70 km), Richmond (160 km), and Wilmington (376 km). The value from the Wilmington data may follow from soil amplification and explain some of the misfit.

Discussion and Conclusions

The information of IDPs consists of ordinal intensities, the configurations of the corresponding localities, and the distances between these localities (Mäntyniemi *et al.*, 2014). There is a lack of information for small numbers of IDPs that allow several equally defensible epicenter determinations to be made. We performed extensive tests of synthetic IDPs to find out how the epicenter can be determined without importing any specific working assumptions, calibration events, or model earthquakes.

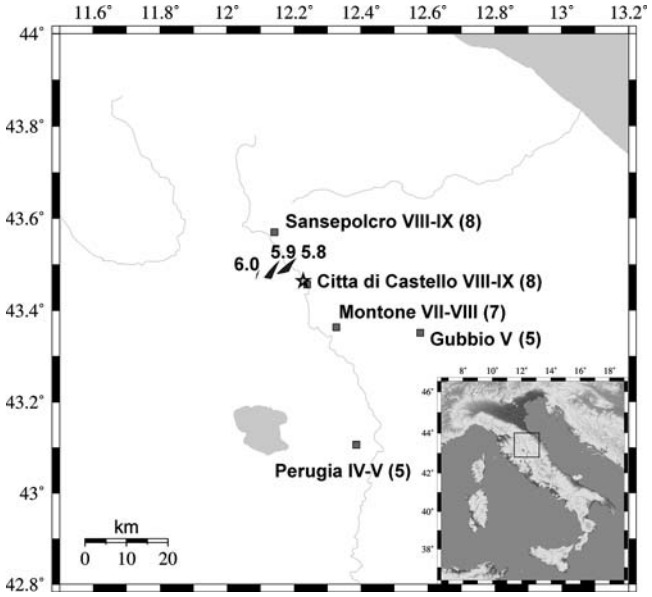

Figure 8. The 26 April 1458 earthquake in central Italy. The squares represent the five localities available. The star is the epicenter according to the Italian DBMI11 database. The roman numerals are the initial intensities, and the numbers in the brackets are intensities that fit the shaded epicenter determinations related to magnitudes M_w 5.8, 5.9, and 6.0.

The quality of the solution can be characterized statistically from the data alone, but as to which alternative epicenter is the most probable cannot be directly inferred. With error-free intensities, a large range of intensities in the sample, a shortest distance to the true epicenter, and to a lesser extent a small azimuthal gap are good indicators of a likely epicenter location. The proportion of small uncertainties increased when at least one good indicator was fulfilled. Summaries presented in Figure 6 give estimates as to the chance of getting a small location uncertainty. The values of the indicators can be calculated for trial epicenters. The interactive tool facilitates the design of scenarios of a past earthquake, which makes the analysis more transparent. Different combinations of input data can be used to illustrate which intensity values contribute to the success or failure of the solution (Figs. 8–10). Additional information may be imported after the computations, because expert knowledge based on local conditions may help reduce the number of potential solutions. For example, information about site conditions may help identify anomalously high or low intensities in the available dataset.

Szeliga *et al.* (2010) used the method developed by Bakun and Wentworth (1997, 1999) to derive magnitudes and epicenters for historical Indian earthquakes, some of which had a moment magnitude above 8. They concluded that location accuracies were probably no better than 50 km, and a dataset with fewer than 10 localities cannot give reliable locations. Szeliga *et al.* (2010) also pointed out the importance of having a large range of intensity observations for each earthquake.

The present results show that different properties of the data influence the location uncertainty. It is impossible to give a single specific numerical value for each of the historical earthquakes. The simulations showed that location accuracies of 20 km or less were attainable in many of the cases that involved fewer than 10 IDPs (Figs. 5 and 6). Park and Hong (2016) determined epicenter and magnitude for historical earthquakes from intensity data by choosing the set of parameters with the highest probability. Their synthetic tests also found that parameters are better for short epicentral distances and improve with an increasing number of IDPs.

Bondár *et al.* (2004) investigated location accuracy using local networks of seismic stations. Their simulations showed that conditions for the crustal events that were located with 5 km accuracy or less with a 95% confidence level had the following characteristics: they were located using at least 10 stations, all within 250 km, the azimuthal gap was less than 110°, and at least one station was located within 30 km of the epicenter. Those conditions are similar to what we found in the present study (Figs. 5 and 6), and they are governed by geometry.

Gasperini *et al.* (2010) used earthquakes that had at least 20 assigned IDPs. They found that in fewer than 5% of cases the distance between the macroseismic and corresponding instrumental epicenter was in excess of 30 km. About 55%–60% of the distances were 10 km or less, and about 20%–25% of them were 5 km or less. Their results and ours support the notion that the proportion of good epicenters increases with an increasing number of IDPs.

In the present simulations, obtaining a location accuracy of 5 km or less was an extremely rare event. For sample sizes of 3–11 IDPs and an intensity range of zero, the chance was one in a million at best. For intensity ranges of ≥ 1, the chance of obtaining location uncertainty of 5 km increased with the number of IDPs. In the case of 3 IDPs, fewer than 20 locations had an accuracy of 5 km or less. In the case of 6 IDPs, the number had increased to ~200, whereas for 11 IDPs there were 3258 such locations at best (in case of intensity range 5).

Our results imply that attempting epicenter determination using reliable intensities rather than a larger number of IDPs with less reliable intensities may be the advisable approach to take, because a single outlier in a sample decreases the proportion of successful locations (Fig. 7). Small location uncertainties can be obtained using small samples, but the likelihood of small uncertainty increases with the sample size. Very poor individual epicenter determinations may be obtained even if the sample size is large. Alternatively, individual good locations can occur even when the sample size is small. A location uncertainty of 20 km or less is realistic in many cases. If the intensities are reliable, sample sizes ranging from 3–5 IDPs would definitely be sparse data, 6–8 IDPs have a much-improved chance of a good location, and 9–11 IDPs are rather ample sample sizes.

Data and Resources

The seismic intensity data for the 1458 earthquake in Italy are available in the intensity database of damaging earthquakes in the Italian region by Istituto Nazionale di Geofisica

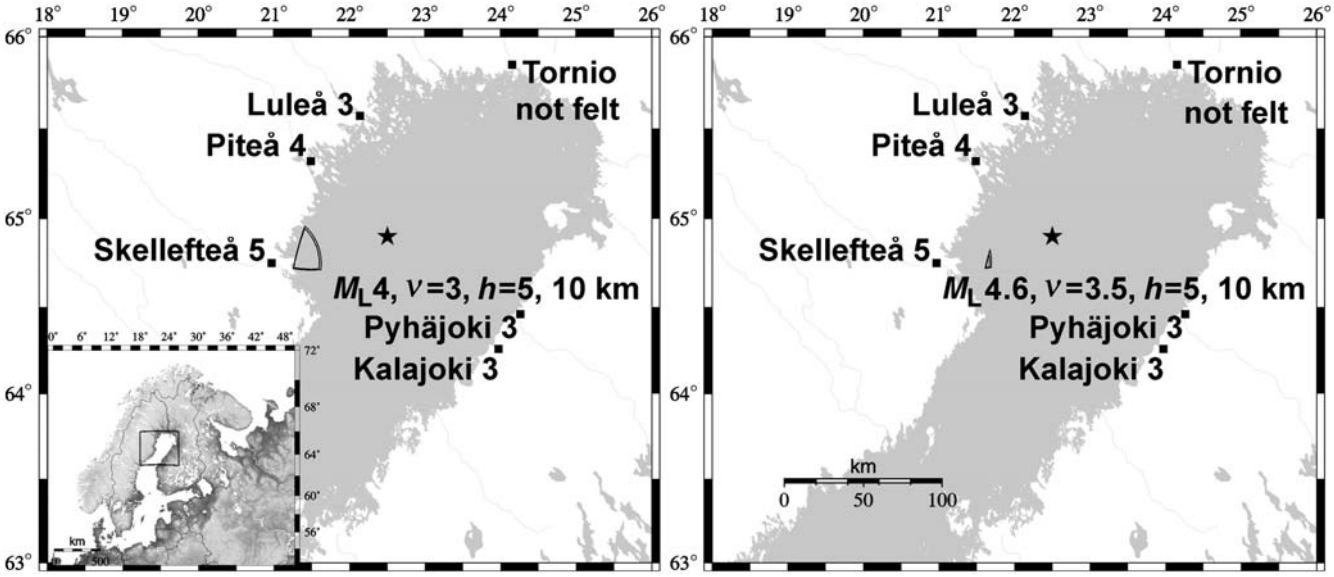

Figure 9. The 14 June 1765 earthquake in Sweden, northern Europe. The squares denote the six localities with reported intensities that also fit the epicenter determinations. The star demarks the epicenter from the parametric earthquake catalog by Ahjos and Uski (1992). Increasing the attenuation coefficient value ν from 3 to 3.5 increased the magnitude from M_L 4.0 to 4.6. Depths from 5 to 10 km fit the illustrated epicenters.

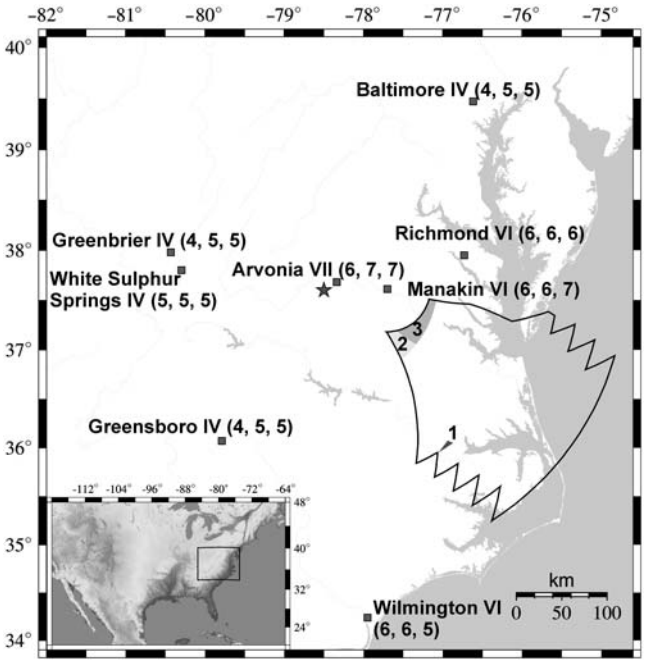

Figure 10. The 23 December 1875 earthquake in the eastern United States. Squares represent the eight localities. The star is the epicenter given in the U.S. Earthquake Intensity Database. The roman numerals are the initial intensities, and the numbers in the brackets are intensities that fit the three illustrated epicenter determinations. The black outline bounds the location of all 57 possible solutions.

e Vulcanologia, Milano, Bologna, at http://emidius.mi.ingv.it/DBMI11 (last accessed June 2016).

The data for the 1875 eastern United States earthquake are available at the U.S. Earthquake Intensity Database of the National Centers for Environmental Information of the National Oceanic and Atmospheric Administration, at https://www.ngdc.noaa.gov/nndc/struts/form?t=101650&s=35&d=35 (last accessed November 2016). The Network of Research Infrastructures for European Seismology (NERIES) NA4 module Deliverable D3 is available at http://emidius.mi.ingv.it/neries_NA4/deliverables.php (last accessed June 2016). Figures 3a and 8–10 were prepared using the Generic Mapping Tools software package by Wessel and Smith (1998).

Acknowledgments

The authors are grateful for the travel grants based on the contractual bilateral cooperation between the Academy of Finland and the Russian Academy of Sciences. This study was partly supported by the Russian Foundation for Basic Research project number 14-05-00258. We acknowledge with appreciation an anonymous reviewer for constructive comments, and Alisdair Mclean and David Whipp for revising the English.

References

Ahjos, T., and M. Uski (1992). Earthquakes in northern Europe in 1375–1989, *Tectonophysics* **207,** 1–23.

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

Bakun, W. H., and C. M. Wentworth (1999). Erratum to estimating earthquake location and magnitude from seismic intensity data, *Bull. Seismol. Soc. Am.* **89,** 557.

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

Bakun, W. H., A. Gómez Capera, and M. Stucchi (2011). Epistemic uncertainty in the location and magnitude of earthquakes in Italy from macroseismic data, *Bull. Seismol. Soc. Am.* **101,** 2712–2725.

Bakun, W. H., A. C. Johnston, and M. G. Hopper (2003). Estimating locations and magnitudes of earthquakes in eastern North America from modified Mercalli intensities, *Bull. Seismol. Soc. Am.* **93,** 190–202.

Bondár, I., S. C. Myers, E. R. Engdahl, and E. A. Bergman (2004). Epicentre accuracy based on seismic network criteria, *Geophys. J. Int.* **156,** 483–496.

Castelli, V., and G. Monachesi (1996). Problems of reliability in earthquake parameters determination from historical records, *Ann. Geophys.* **39,** 1029–1040.

Cecić, I., R. M. W. Musson, and M. Stucchi (1996). Do seismologists agree upon epicentre determination from macroseismic data? A survey of ESC Working Group 'Macroseismology', *Ann. Geophys.* **39,** 1013–1027.

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

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

Gutdeutsch, R., and C. Hammerl (1999). An uncertainty parameter of historical earthquakes—The record threshold, *J. Seismol.* **3,** 351–362.

Houng, S. E., and T.-K. Hong (2013). Probabilistic analysis of the Korean historical earthquake records, *Bull. Seismol. Soc. Am.* **103,** 2782–2796.

Kondorskaya, N. V., and N. V. Shebalin (Editors) (1982). *New Catalog of Strong Earthquakes in the Territory of the Soviet Union from Ancient Times till 1977*, NOAA, Boulder, Colorado.

Mäntyniemi, P. (2012). Intraplate seismicity and seismic hazard: The Gulf of Bothnia area in northern Europe revisited, in *Earthquake Research and Analysis—New Frontiers in Seismology*, S. D'Amico (Editor), InTechOpen, Rijeka, Croatia, 283–298.

Mäntyniemi, P., R. E. Tatevossian, and T. N. Tatevossian (2014). Uncertain historical earthquakes and seismic hazard: Theoretical and practical considerations, *Geomat. Nat. Hazards Risk* **5,** 1–6.

Musson, R. M. W. (1989). Accuracy of historical earthquake locations in Britain, *Geol. Mag.* **126,** 685–689.

Musson, R. M. W. (1998). Inference and assumption in historical seismology, *Surv. Geophys.* **19,** 189–203.

Musson, R. M. W. (2009). MEEP 2.0 user guide, *Br. Geol. Surv. Open Rept. OR/09/045*, 22 pp.

Musson, R. M. W., G. Grünthal, and M. Stucchi (2010). The comparison of macroseismic intensity scales, *J. Seismol.* **14,** 413–428.

Park, S., and T.-K. Hong (2016). Joint determination of event epicenter and magnitude from seismic intensities, *Bull. Seismol. Soc. Am.* **106,** 499–511.

Pasolini, C., D. Albarello, P. Gasperini, V. D'Amico, and B. Lolli (2008). The attenuation of seismic intensity in Italy, part II: Modeling and validation, *Bull. Seismol. Soc. Am.* **98,** 692–708.

Radziminovich, Y. B. (2014). November 13, 1898 earthquake in Western Transbaikalia: A white spot in seismic history of East Siberia, *J. Seismol.* **18,** 93–107.

Ringdal, F., E. S. Husebye, O. A. Sandvin, and A. Christoffersson (1978). Statistical test theory in the analysis of macroseismic questionnaires, *Tectonophysics* **49,** 161–170.

Stucchi, M (1994). Recommendations for the compilation of a European parametric earthquake catalogue, with special reference to historical records, in *Historical Investigation of European Earthquakes, vol. 2, Materials of the CEC Project Review of Historical Seismicity in Europe*, P. Albini and A. Moroni (Editors), Istituto di Ricerca sul Rischio Sismico, Milano, Italy, 181–190.

Szeliga, W., S. Hough, S. Martin, and R. Bilham (2010). Intensity, magnitude, location, and attenuation in India for felt earthquakes since 1762, *Bull. Seismol. Soc. Am.* **100,** 570–584.

Tatevossian, R. E., and P. Mäntyniemi (2014). Earthquake scenarios: A practical way to handle alternative solutions to historical earthquakes and to increase the transparency of seismic hazard assessment, *Nat. Hazards* **72,** 549–564.

Tatevossian, R. E., T. N. Tatevossian, and P. Mäntyniemi (2013). Earthquake activity in Finland and the Russian North in December 1758: Rare reports and their interpretation, *Ann. Geophys.* **56,** 1–18.

Wessel, P., and W. H. F. Smith (1998). New, improved version of the Generic Mapping Tools released, *Eos Trans. AGU* **79,** 579.

Institute of Seismology
Department of Geosciences and Geography
University of Helsinki
P.O. Box 68
FI-00014 Helsinki, Finland
paivi.mantyniemi@helsinki.fi
(P.M.)

Institute of Physics of the Earth
Russian Academy of Sciences
B. Gruzinskaya 10
Moscow 123242, Russia
(R.E.T., R.N.V.)

Manuscript received 30 June 2016;
Published Online 14 March 2017