



NEtwork of Research Infrastructures for European Seismology

Report

Macroseismic estimation of earthquake parameters

Activity:	Distributed Archive of Historical Earthquake Data
Activity number:	NA4
Deliverable:	Procedures for macroseismic estimation of earthquake parameters
Deliverable number:	D3
Responsible activity leader:	RMW Musson
Responsible participant:	BGS
Author:	RMW Musson and MJ Jiménez

**Sixth Framework Programme
EC project number: 026130**

Foreword

This report is a deliverable in the NERIES project, module NA-4, and deals with the selection of a suitable method for determining earthquake parameters for historical earthquakes from intensity data points. This is issue 2.2, dated 6 November 2008.

Acknowledgements

The authors would like to thank the other members of the NA-4 module for their assistance in the discussions of this topic, especially in the context of the Workshop meeting in Edinburgh, 26-27 February 2007, and also the NERIES General Assembly at Utrecht, 30 June-2 July 2008. The historical seismology team of INGV Milano worked hard on testing and evaluating the first version of this system, work which proved invaluable in developing the final version. Thanks also to Paolo Gasperini for advice regarding the inner workings of BOXER.

Contents

Foreword	i
Acknowledgements	i
Contents	i
1 Introduction	1
2 Issues in macroseismic parameter estimation	1
2.1 Isoleismals and IDPs	1
2.2 Epicentre	3
2.3 Magnitude	5
2.4 Depth.....	7
3 Testing methods for macroseismic parameter estimation in a historical context	8
4 Developing a solution for NERIES NA-4	18
4.1 Principles	18
4.2 Problems and solutions	19
4.3 Stages in the application of the method	20
4.4 Estimating uncertainties.....	21
5 Software operation	23
6 Calibration	27
7 Earthquakes with very few IDPs	30
8 Earthquakes with no IDPs	31
9 Offshore earthquakes	34
10 Conclusions	35
References	36

FIGURES

Figure 1 - Example intensity map	2
Figure 2 - Effect of depth on isoseismals.....	7
Figure 3 - Test data set #1.....	10
Figure 4 - Test data set #2.....	11
Figure 5 - Test data set #3.....	12
Figure 6 - Test data set #4.....	13
Figure 7 - Test data set #5.....	14
Figure 8 - Test data set #6.....	15
Figure 9 - Test data set #7.....	16
Figure 10 - Earthquake of 26 November 1972.....	27
Figure 11 - Computed magnitudes for the calibration events	30
Figure 12 - Magnitudes for single IDP cases: Italy	31
Figure 13 - Isoseismal map of the earthquake of 15 August 1926.....	33
Figure 14 - Earthquake of 6 September 2002 (detail).....	35

TABLES

Table 1 - Test data sets	9
Table 2 - Absolute results of the method comparison.....	17
Table 3 - Summary of result rankings for seven cases	18
Table 4 - Example of magnitude uncertainty	22

1 Introduction

The derivation of earthquake parameters from macroseismic (intensity) data is an inveterate problem. Yet for earthquakes in the pre-instrumental period (roughly, before 1900) intensity data points (IDPs) are the only form of numerical data available to the seismologist. In order to produce a numerate, consistent catalogue of historical earthquakes that can be combined in a compatible way with modern instrumental data requires some system for estimating what instrumental parameters would have been obtained had seismometers been in operation.

Successive catalogue authors have had to deal with this problem as they saw fit; but as most earthquake catalogues have been compiled as national initiatives, one finds that one type of method has been used in one country, something else in another, and so on. This leads to obvious problems of inconsistency when it comes to studies that need to transcend national borders.

A major aim of the NA-4 module of the European Framework project NERIES is to produce a catalogue of European earthquakes before 1900 in which there is the greatest possible level of internal consistency in the determination of earthquake parameters. This means the use of uniform procedures for determining earthquake parameters over the whole of Europe. Finding suitable procedures that can be used for this is a difficult task, and is the subject of this report.

The parameters to be determined are essentially the location and the size of each earthquake. Precisely how one defines location in this context is arguable – one speaks of the “macroseismic epicentre”, but this is not necessarily exactly the same as an epicentre in the sense of the surface projection of the point where an earthquake rupture initiates. Where an earthquake rupture is large, while the distribution of high intensities may delineate its extent, there is no possibility to determine the initiating point – and probably not a lot of interest in doing so either. The co-ordinates that will be used will be those of the centre point of the rupture; something approximating to the focus from which the seismic energy radiated. For this reason, the term barycentre is sometimes preferred (Cecić et al 1996).

From the point of view of seismic hazard, arguably such a point is of greater interest, in terms of reconstructing the seismic field. Since earthquakes occur in three dimensions, as well as the latitude and longitude co-ordinates, one needs also some sort of depth of focus. This is evidently more meaningful for smaller earthquakes of limited rupture dimensions.

“Size” has to be considered here to mean “magnitude”; whereas many earlier historical earthquake catalogues were content to use epicentral intensity, I_0 , as a size measure, for modern applications, and for consistency with modern data sets, this is not enough, and magnitude, preferably moment magnitude, M_w , has to be estimated.

2 Issues in macroseismic parameter estimation

The following discussion was prepared for the NERIES NA-4 Workshop in Edinburgh, 26-27 February 2007, and appears here in an edited version.

2.1 ISOSEISMALS AND IDPS

One basic issue to be addressed at the outset is whether or not isoseismals are used. If so, there is an additional stage to analysing an earthquake of preparing isoseismals, which, of course, brings additional difficulties. It's well known that there are as many ways of drawing isoseismals for a given data set as there are seismologists. On the other hand, this is partly due to different seismologists having different ideas about how much smoothing to apply. If one first establishes some ground rules as to how isoseismals are to be drawn, possibly much inter-seismologist variability could be avoided. For instance, there are various automatic systems that

can be used. Schenková et al. (2007) present quite reasonable results from the kriging of intensity data to produce “objective” isoseismals.

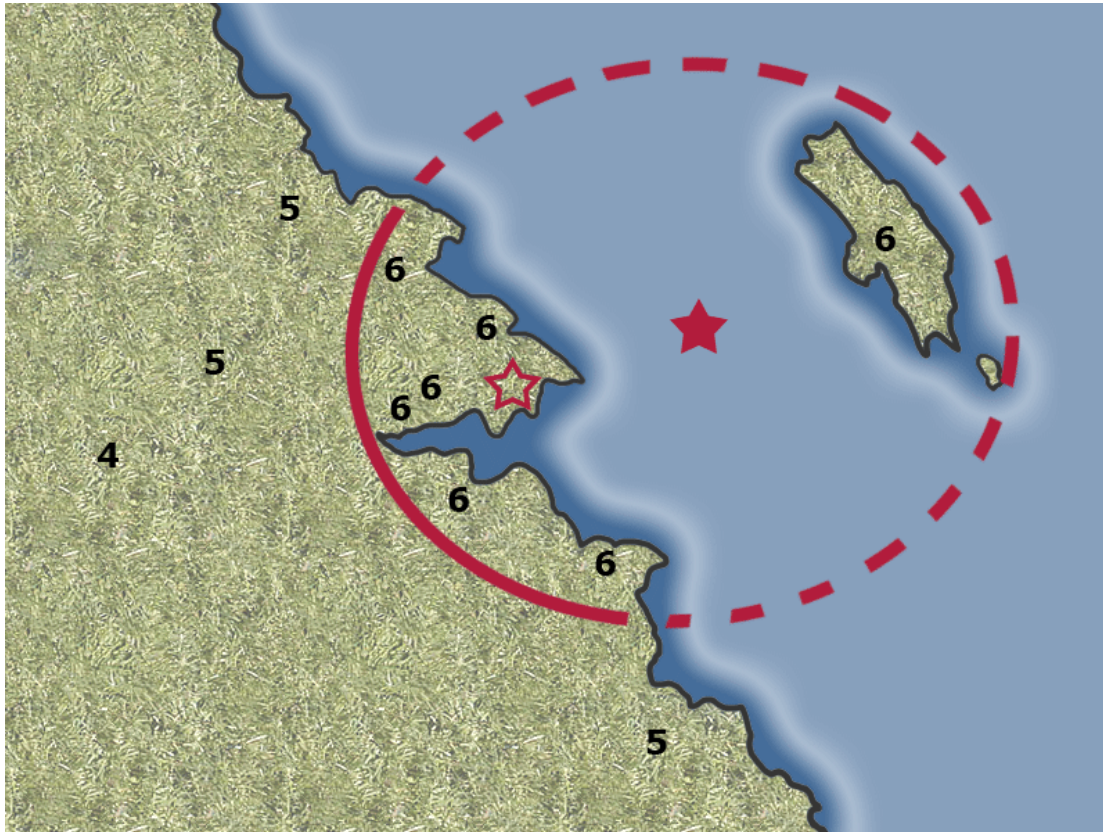


Figure 1 - Example intensity map

The reason for using isoseismals is this: they can reduce or eliminate distortions due to population distribution. Consider Figure 1, in which an offshore earthquake produces its highest intensities at several places on the mainland and in one island location. A seismologist, knowing that the irregular distribution is due to topography, can estimate an isoseismal as shown, which is likely to be a better expression of reality than any processing of the individual data points, in which the clustering due to population factors will distort results. From an isoseismal one could deduce an epicentre in the position of the solid star; from data points one would be likely to obtain the position of the open star.

One possible approach to regularise isoseismal construction is to do some preliminary processing by means of gridding – overlaying (say) a 10 km grid and taking the mode of observations within each square. This would then give a new pseudo-IDP at the centre of the square which would be used in subsequent processing. This was employed by Principia (1982).

Note also that a related problem occurs in intensity attenuation studies. Suppose one wishes to solve for

$$I = a + b M + c \ln R \quad (1)$$

Because intensity values are integer, if one had a very well-behaved data set with concentric rings of intensity 6, 5, 4, one would effectively have a step function with distance from the epicentre. The mean distance for intensity 5 would actually be halfway between the outermost 6 and the innermost 4, and it would be this distance that would be predicted by equation (1) if the coefficients were based on IDPs.

In reality, even though intensity is integer so far as one can assign it, it is true that at least notionally one can imagine that there is a gradation within the annulus of intensity 5 from “only just 5” to “almost 6”. If equation (1) is determined from isoseismals, it will correctly predict “only just 5”, or “5.0”, at the distance where intensity 5 gives way to intensity 4.

2.2 EPICENTRE

There are quite a few ways to locate earthquakes from macroseismic data, some of which are so similar they can be grouped together.

Not considered here are methods that are doubtful or impractical for historical data sets, such as triangulation on reported time, estimated direction of shaking, orientation of cracks in buildings and so on.

2.2.1 Highest intensity

This is possibly the oldest macroseismic method in existence. The epicentre is the centre of the area of highest intensity. *Exactly* how one defines this is another matter. It could be, for instance:

- The geometric mean of IDPs equal to I_{max} .
- The geometric mean of IDPs equal to I_{max} and I_{max-1} .
- The geometric mean of IDPs equal to I_{max} and I_{max-1} weighted in favour of I_{max} .
- The centroid of the highest isoseismal.
- The mean of the centroids of the two highest isoseismals.
- The mean of the centroid of the highest isoseismal and the mean of IDPs equal to I_{max} (assuming the highest isoseismal has a value of I_{max-1}).

And so on. One can make many permutations. This is one of the main objections to this sort of method; there is not much reason to prefer one of these to another, except perhaps that in cases like Figure 1, isoseismals may be preferable to IDPs. Usually, the results given by the different options are very similar.

It is possible to extend the method from the highest intensities to lower ones: below I_{max-1} . It could be argued, at least in some cases, that the centroid of the area of general observation is related to the epicentral location, and the location of the highest intensities is chiefly related to surface effects. One could imagine such a case with an earthquake occurring in a mountain area near to a broad river valley. The situation in Figure 1 is another possible case for looking at lower intensities.

On the other hand, the wide felt area of an earthquake can be skewed by a variety of factors such as variations in attenuation, directionality, and so on.

In cases where a data set consists of very few IDPs (or only one) then a highest intensity approach is the only possible method to use. There is a danger, though, that this can give the erroneous impression that earthquakes happen directly below historically important cities. Hence, for instance, maps of historical earthquakes in Italy that seem to show an active centre directly under Venice. This will be returned to later.

2.2.2 Boxer

The Boxer method, now well-established, is to a large degree an extension of the highest-intensity technique (Gasperini et al.1999). In this method, one takes first the IDPs equal to I_{max} ; if these are less than four, one adds the data from successively lower intensity levels until at least four data points are available. These are then trimmed by removing the most outlying 25% of the data, and the mean is taken of the remaining points.

Boxer is aimed not so much at determining epicentre as determining the surface projection of the fault plane; the epicentre can be any point within this. Clearly, for large earthquakes the concept of epicentre is a marginal one. The technique is still based on the distribution of the highest intensity values, and therefore should still suffer from similar problems regarding population distribution. It's hard to see how the method could cope with any earthquake that was offshore, or distinguish such an event.

Also, for earthquakes of lesser magnitude where the rupture length may only be a few tens of km, it's not clear that the orientation of the area of highest intensity is necessarily related to the orientation of the fault plane. It could be due to radiation pattern, for instance.

Thus, although Boxer has produced some good results for modern earthquakes where the true fault parameters are known, it is questionable how reliable it is for historical data sets, especially for moderate-sized earthquakes.

2.2.3 *Minimising residuals*

This method is generally called the "Bakun-Wentworth" method, but it was described by Peruzza previously (in 1992). However, as formulated by Bakun and Wentworth (1997), it expands on what was proposed by Peruzza (1992). The method of Papazachos (1992) is also similar.

The principle, as set out by Peruzza (1992), is as follows. If one considers the epicentre as the spot on the surface from which the energy of an earthquake appears to radiate, then, taking some attenuation relationship of the form

$$I_0 - I = f_n [r] \quad (2)$$

(where r is distance), then this should fit the data best if the distance is measured from the true epicentre rather than from some other point. Thus, given some form of equation (2) considered to be regionally appropriate, one can take a test epicentral position, calculate the expected intensity for each IDP location, record the residual from the true observed value, and then sum all the residuals. Repeat this over a grid of possible locations, and one defines a probability distribution function (pdf) of the epicentral location; the place with the minimum sum of residuals is the epicentre.

The addition provided by Bakun and Wentworth (1997) is that instead of using equation (2) they use equation (1), and during the grid search, they treat magnitude as an unknown variable (instead of using a known I_0 value). This results in two pdfs, one for epicentre and one for magnitude. One surface shows the probability that the epicentre is in any location; the other shows, assuming the epicentre is in a given location, what the preferred magnitude must be. Thus one adopts a solution based on expert judgement in the area of highest probability. For large earthquakes, this may be moderated by consideration of available tectonic structures.

The main limitation of this is that the coefficients in equation (1) may be uncertain. One therefore needs to calibrate the method for any particular region – which leads to another question as to how small a large a region the calibration should be performed for. The principle of always choosing the solution that minimises the magnitude could also be questioned.

2.2.4 *Pairwise comparison*

This name is used here to distinguish a class of methods derived from a proposal by Shumila, presented at the ESC General Assembly in Athens in 1994. The basis of the method, which leads to a simple processing algorithm, is as follows. Consider any two IDPs in a data set. If one is higher than the other, then logically, one expects the epicentre to be closer to the higher intensity point. Therefore all possible locations that are closer to the higher IDP are preferable to all locations that are closer to the lower one. This applies to all possible pairs of IDPs. Therefore the most likely epicentral location is the point that maximises closeness to the higher IDP in any pair. The method was never published.

The principle is attractively elegant, and unlike the residual-based methods, requires no additional input in the form of attenuation equations. In practice, it can give unstable results with poor data sets, but for large modern data sets it has yielded extremely accurate solutions (within a few km of the instrumental location) even for offshore events.

2.2.5 *Expert judgement*

The final method considered here is one in which the seismologist chooses a location based on expert judgement, using any evidence or reasoning. Consider the following example. A historical earthquake is described as being strong with slight damage in Venice. There are no

other reports. Six days later, there is a report of a small earthquake in Belluno. The seismologist reasons as follows: “It is unlikely that a strong earthquake had an epicentre close to Venice. The event following is probably an aftershock. Aftershocks should be located close to the epicentre of the main shock, so the epicentre of the first event should be near Belluno”. This is a macroseismic location, but it draws on several types of information, including the seismologist’s knowledge of the area and ideas about typical earthquake behaviour.

Such an approach has the advantage of taking into account the maximum of information, and the disadvantage of being subjective and not perfectly repeatable. However, as Veneziano (1995) has written in a different context, “Since the human mind is very effective in processing complex information structures ... the results of judgemental methods may actually be more accurate than those produced by formal ... rules”.

Clearly, it would be difficult to apply this method at a European level. It requires local knowledge for each area, and would not be a common standard across the whole catalogue.

2.3 MAGNITUDE

There are two basic approaches to estimating magnitude from macroseismic data: using maximum intensity or some measure of area. Both maximum intensity and area (especially the former) are subject to variability primarily due to focal depth and/or geology.

2.3.1 *Maximum intensity*

Here we consider that maximum intensity can stand in as a measure of magnitude. It should be understood that in this case, maximum intensity could also be epicentral intensity variously defined, including a decimal interpolation of the intensity field to the point considered to be the epicentre. Thus the approaches considered in this section should be considered classes with variants according to the actual definition of “maximum intensity” applied. (Furthermore, it’s hard to see that there is much in the way of objective reasons for preferring one definition to another).

In the simplest form, one can use the following:

$$M = a + b I_{\max} \quad (3)$$

Such an equation is liable to produce high residuals when fitted to data. There are a number of problems. First, I_{\max} is heavily dependent on focal depth. Thus a high I_{\max} value may simply indicate a smallish earthquake at very shallow depth. Reasoning that *most* seismicity in a region is of a certain depth, therefore *all* seismicity is the same depth (and therefore that depth can be discounted for a poorly-documented historical event) is clearly unsound.

An extreme case is the 15 January 1865 earthquake in NW England, which wrecked one village and was barely perceptible 10 km away (Musson 1998a). There may be other similar cases elsewhere that have not been recognised in the historical catalogue because people have assumed that high intensity must mean a large earthquake.

Secondly, the highest intensity may be partly due to soil conditions. This was probably an influence in the 1865 case.

Thirdly, the highest reported intensity may be nowhere near the epicentre. This is particularly the case for earthquakes with poor data sets, close to either mountains or the coast.

The first problem can be overcome up to a point by including depth in the calculation, as in:

$$M = a + b I_{\max} + c \ln h \quad (4)$$

Of course, this assumes that depth can be determined (see next section) which is frequently not the case. An informal and unpublished test on UK data found that magnitudes determined using equation (3) had an uncertainty of ± 0.7 , whereas if equation (4) was used, this could be improved to ± 0.4 .

The soil condition problem could possibly be tackled by making some correction factor; but this is something that would be likely to be arbitrary and difficult to apply.

The third problem seems to be more or less insoluble. Given enough data, one could extrapolate a maximum intensity, but this is likely to be very uncertain.

Rather than using equation (3), which assumes a linear relationship between M and I_{max} , an alternative approach that has been tried in Italy (Cavallini and Rebez 1996) is to treat each I_0 value separately and observe the distribution of magnitude values for earthquakes with this I_{max} . Thus for 42 British earthquakes with I_{max} of 6, the mean magnitude is 4.1 and the median is 4.4 ML. However, the range is from 2.3-5.4 ML and the standard deviation on the mean is ± 0.9 .

2.3.2 *Felt area*

Instead of looking at the peak intensity value, one can instead look at the total felt area, or the area felt at some given intensity value. As a general rule, magnitude is linearly related to the log of the area at some intensity value. Ideally this should be as low an intensity as possible; the lower the intensity used, the less sensitive the result is to depth (see Figure 2). Using the area of intensity 3 (or 4 for larger events) can thus give quite robust estimates of magnitude, on which neither depth nor soil effects nor population distribution have a great influence.

Because magnitude is related to the log of the felt area, one can have large inaccuracies in the area without much impact on the magnitude. Suppose one takes the approximation:

$$M = \log A \quad (5)$$

(where A is total felt area) and consider an earthquake in a coastal region. The onshore felt area is 22,000 sq km. The offshore felt area is uncertain; by rough extrapolation of isoseismals one obtains an additional 10,000 sq km for a total area of 32,000 sq km, which equates to a magnitude of 4.5. Now, suppose that the estimate of the offshore area was wrong by a factor of two, and the total felt area should be 42,000 sq km. The change in the magnitude is only +0.1. Thus for poorly constrained historical data sets this method can be rather suitable. Radius can be used rather than area.

It is possible to define several equations: for the area of intensity 3, 4, 5 and so on. There is a question as to which to use for events with several isoseismals. Although the lower ones are better because they are insensitive to depth, they may be less well defined. In Ambraseys (1985), all available isoseismals are used.

In cases where isoseismals are not drawn, area or radius needs to be defined in some way with regard to the IDP set.

A question arises as to cases where there are only very few data points of high intensity and nothing else (for instance, the only information about an earthquake is that in two villages the intensity was 8). One conclusion would be that this method cannot be used, since the area of intensity 8 is not at all a robust indicator, and is, in any case, very poorly defined. When the data are very poor, the most one can do, probably, is to try and distinguish very roughly if this is a small or a large earthquake. Intensity 8 does not necessarily indicate high magnitude. One can analyse as follows: suppose this was a large earthquake with a true total felt area of some tens of thousands of sq km, centred near the two surviving IDPs. Is it reasonable to suppose, given the historical circumstances, that no reports would survive at lower intensity from all the other towns and cities in the area that would have been affected if this had been a large earthquake? On such reasoning, one can estimate very roughly the limits of the felt area that might be credible for this event, given the historical and geographical context, and this may at least impose limits or approximate values on the estimated magnitude.

2.3.3 *Attenuation-based methods*

This is simply the Bakun-Wentworth (1997) method as described above in section 2.2.3.

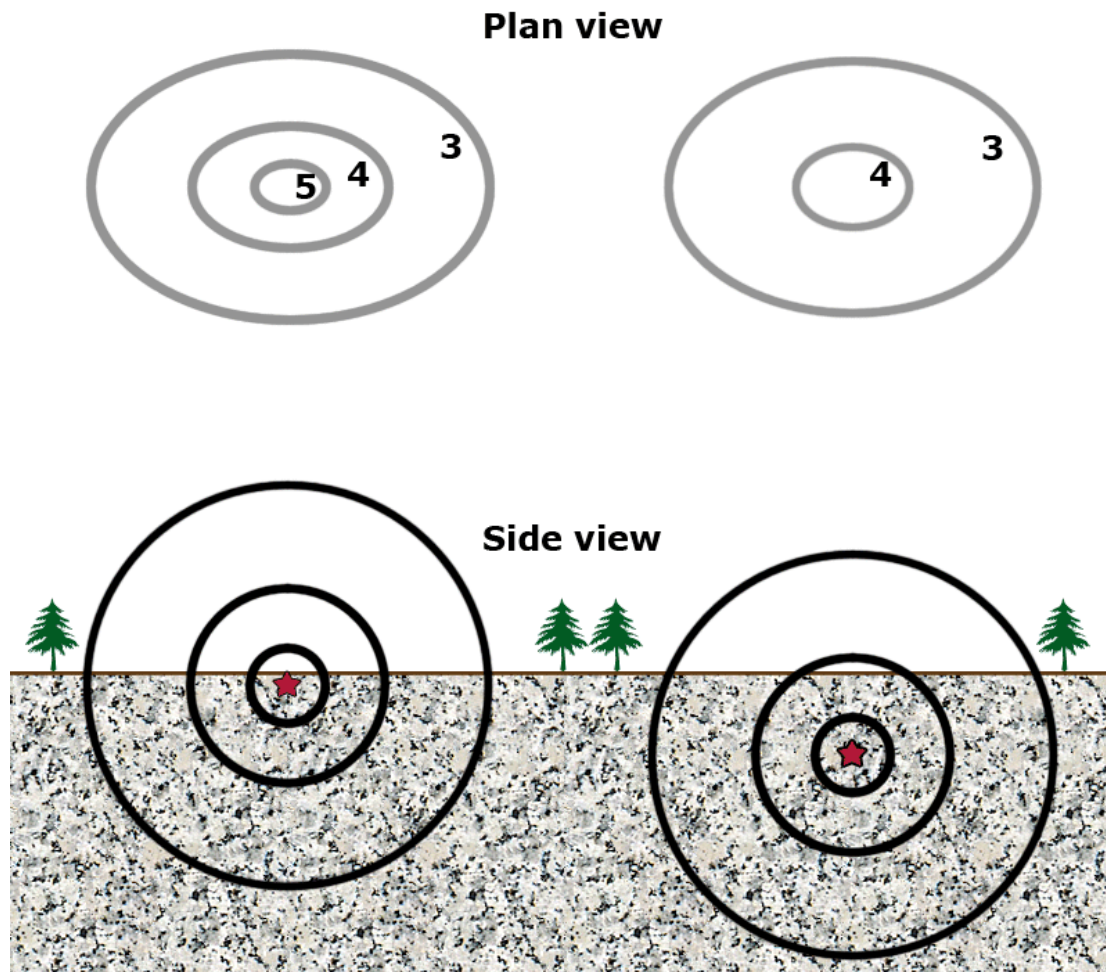


Figure 2 - Effect of depth on isoseismals

2.4 DEPTH

There are various methods for estimating depth from intensity data, attributed variously to Kövesligethy (1906), Blake (1941) and Sponheuer (1960), but all similar and ultimately deriving from the work of Kövesligethy (1906). The principle is evident in Figure 2: for the same magnitude event (same total area) a shallow event will tend to have a pattern of tightly-spaced contours of high intensity around the epicentre, whereas a deeper event will have lower epicentral intensity and broader spacing of contours.

The same pattern can be discerned in the distribution of IDPs even if one doesn't actually construct isoseismals; thus in the context of the SIRENE project (Levret et al. 1994) the Kövesligethy method was used on the basic IDPs.

This method clearly assumes that source dimensions are insubstantial. It is appropriate for use with small earthquakes, but not with large ones that fracture the surface. For large surface-rupturing earthquakes it is doubtful whether the concept of depth is of any value; hence the tendency for ground motion models to use Joyner-Boore distance (the closest horizontal distance to the vertical projection of the rupture) instead of hypocentral distance.

The model can be described by the following equation:

$$I_0 - I_i = 3 \log (r_i/h) + 3 \alpha \log e (r_i - h) \quad (6)$$

In this, r_i is the radius of the isoseismal for intensity I_i , h is depth, e is Euler's constant and α is an anelastic attenuation coefficient.

This can be solved by computing residuals for all possible values of I_0 and h , and plot these as a surface (Burton et al. 1985). I_0 can be considered to be unknown but constrained; it may be fractionally higher than observed I_{max} but not unreasonably so.

When used with isoseismals, this method requires no assumption of an epicentre, since one can simply take isoseismal areas and estimate the radius for an equivalent circular area. When used with IDPs, once again there is a potential bias due to population distribution. An average distance to IDPs of intensity I could be used, given an adopted epicentre, but some recalibration would be needed to allow for the fact that the mean radius to IDPs of intensity I is going to be less than the isoseismal radius.

Calibration raises two other issues. Firstly, the value of 3 in equation (6) is given as a scaling constant, but some workers have chosen it as a variable to be found (usually denoted as k). Secondly, one might ask what the parameter α represents. It can be taken to be a regional parameter expressing some crustal quality, to be found by optimising using a regional aggregate of data sets. This is the view of Karník (1969), who gives suggested values of α for different parts of Europe. But there exist studies (e.g. Zsiros 1996) in which α is solved for separately for each earthquake, which clearly implies some other interpretation of what it is meant to represent physically.

3 Testing methods for macroseismic parameter estimation in a historical context

At the NA-4 Workshop in February 2007, all the methods described above were discussed. It was resolved that some form of test should be made to see which methods performed best in the context of typical historical cases. The problem of testing methods against instrumental data is that generally, for an event for which instrumental data are available, most likely the macroseismic data will be of relatively good quality. When the methods are used in earnest, they will be applied to historical intensity data sets the quality of which may be very poor.

To overcome this problem, it was decided to fake some historical data sets. A small selection of earthquakes was compiled for which intensity and instrumental data were available: three moderate to large earthquakes from Greece and two much smaller ones from the UK, so that the sample encompassed both interplate and intraplate seismicity. These data sets were then degraded by hand by removing IDPs, especially those from small or poorly-connected places, in an attempt to construct data sets that represented the sort of information that might have been available had the earthquakes in question occurred in previous centuries. Two of the events were degraded twice in this way, to create a poor and a very poor version of the data set. The two UK events included some IDPs with no actual intensity value ("Felt"), and one of these had an offshore epicentre.

The resulting seven test data sets are shown in Table 1 and Figures 3-9

Set #	Date	Location	Lat	Lon	Mw	# IDPs
1	1915/07/08	Ionian Islands	38.50	20.62	6.7	16
2	1915/07/08	Ionian Islands	38.50	20.62	6.7	10
3	1981/03/04	Attica	38.18	23.23	6.3	19
4	1984/10/25	Northern Greece	40.20	21.75	5.5	7
5	2001/10/28	Eastern England	52.84	-0.85	3.4	77
6	2001/10/28	Eastern England	52.84	-0.85	3.4	11
7	1999/03/04	Western Scotland	55.40	-5.24	3.2	31

Table 1 - Test data sets

The analyses of these data sets were performed by members of the NA-4 module according to five different methods. These were: the Peruzza method (by BGS), the Shumila method (by BGS), the B&W method (by both INGV and ETHZ, using slightly different approaches), the Boxer method (by INGV) and the Papazachos method (by ITSAK). The first two methods were used for epicentre only; the others were used for epicentre and magnitude. The absolute errors are given in Table 2, but for comparison it is probably more useful to consider the rankings, since otherwise a large residual in one test case could have an excessive effect on the overall results, and some cases are “easier” to resolve than others – Set #7 being particularly awkward.

The magnitudes of the UK earthquakes are consistently over-estimated, which is not surprising given that the methods were not specifically calibrated for UK conditions. The results do not show one method consistently out-performing the others. It’s interesting to note, for instance, that in Cases 5 and 6 (the same data set degraded to different extents to represent firstly, an 19th century data set, and secondly, and 18th century one), Boxer gives the best result for Case 5 and the second worst for Case 6. To arrive at some sort of summary of results, the ranks were summed across the seven test cases (so that a low score is best) and the results are presented in Table 3.

These results show that the B&W method as performed by INGV gave much the best magnitudes, but did not do nearly so well at computing epicentres. The B&W method as performed by ETHZ did best for epicentres, but the Peruzza method was almost equal.

It would probably not be good to over-interpret these results, given that none of the testers were able to make regional calibrations in advance of the experiment.

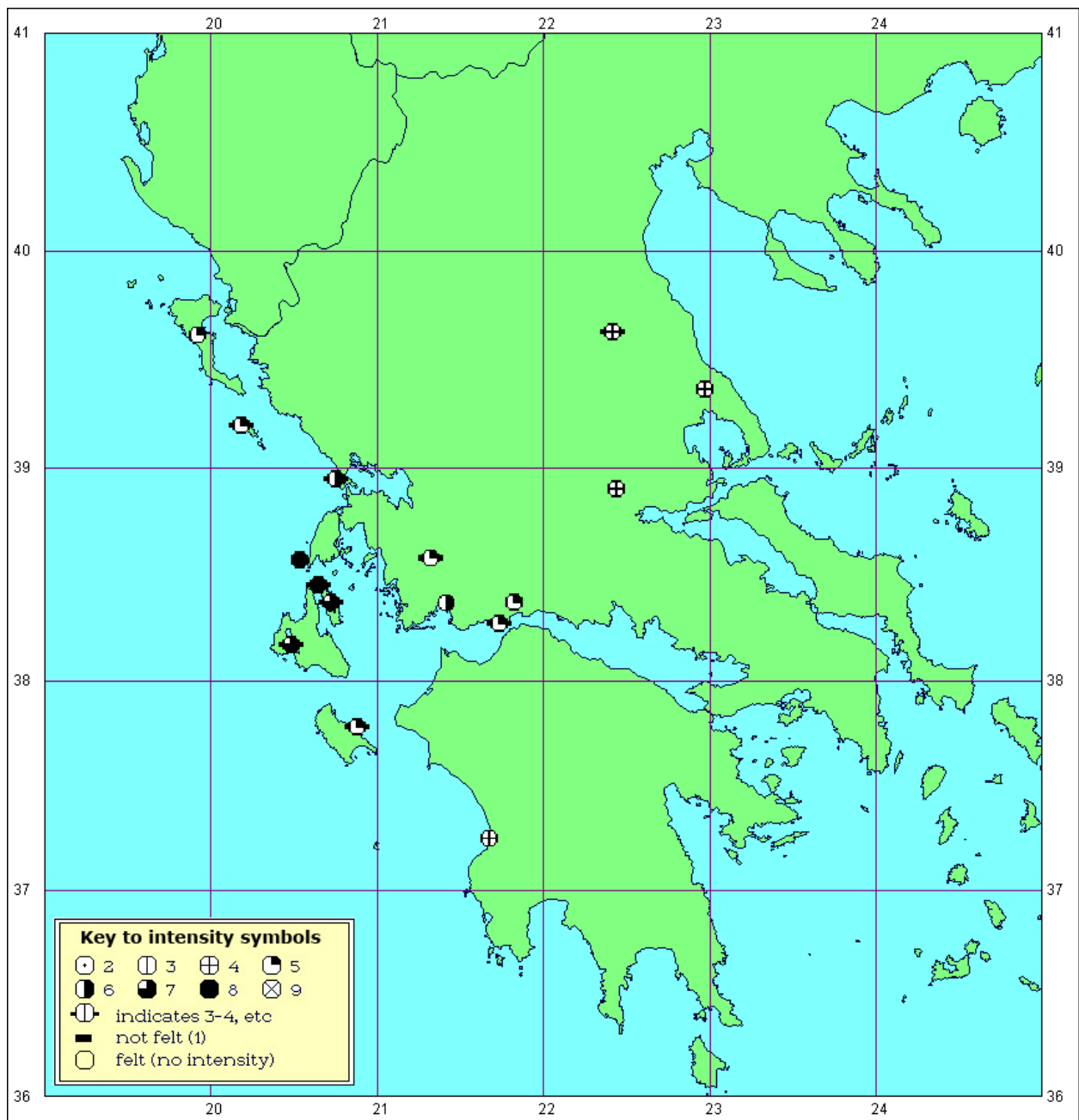


Figure 3 - Test data set #1

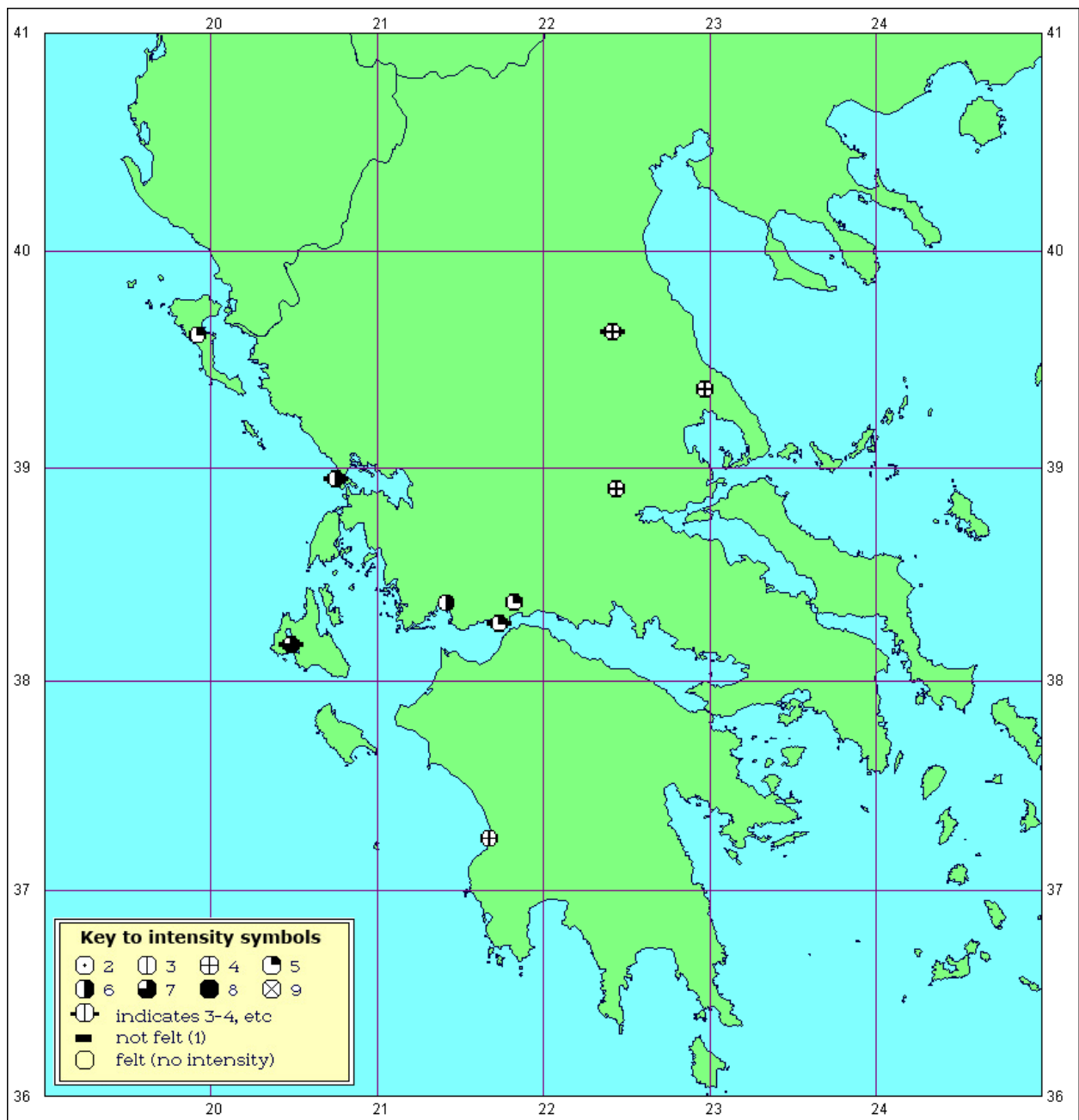


Figure 4 - Test data set #2

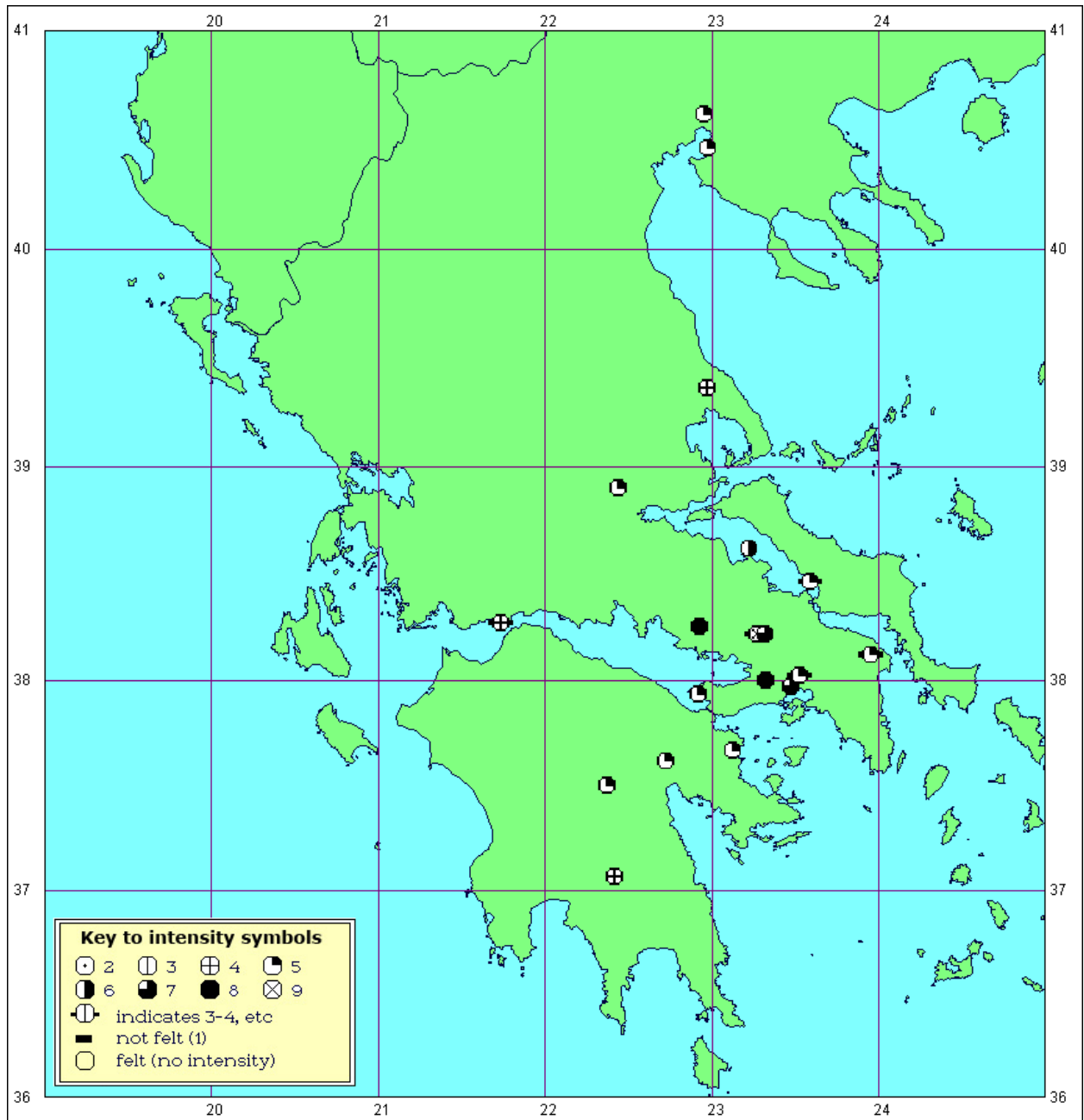


Figure 5 - Test data set #3

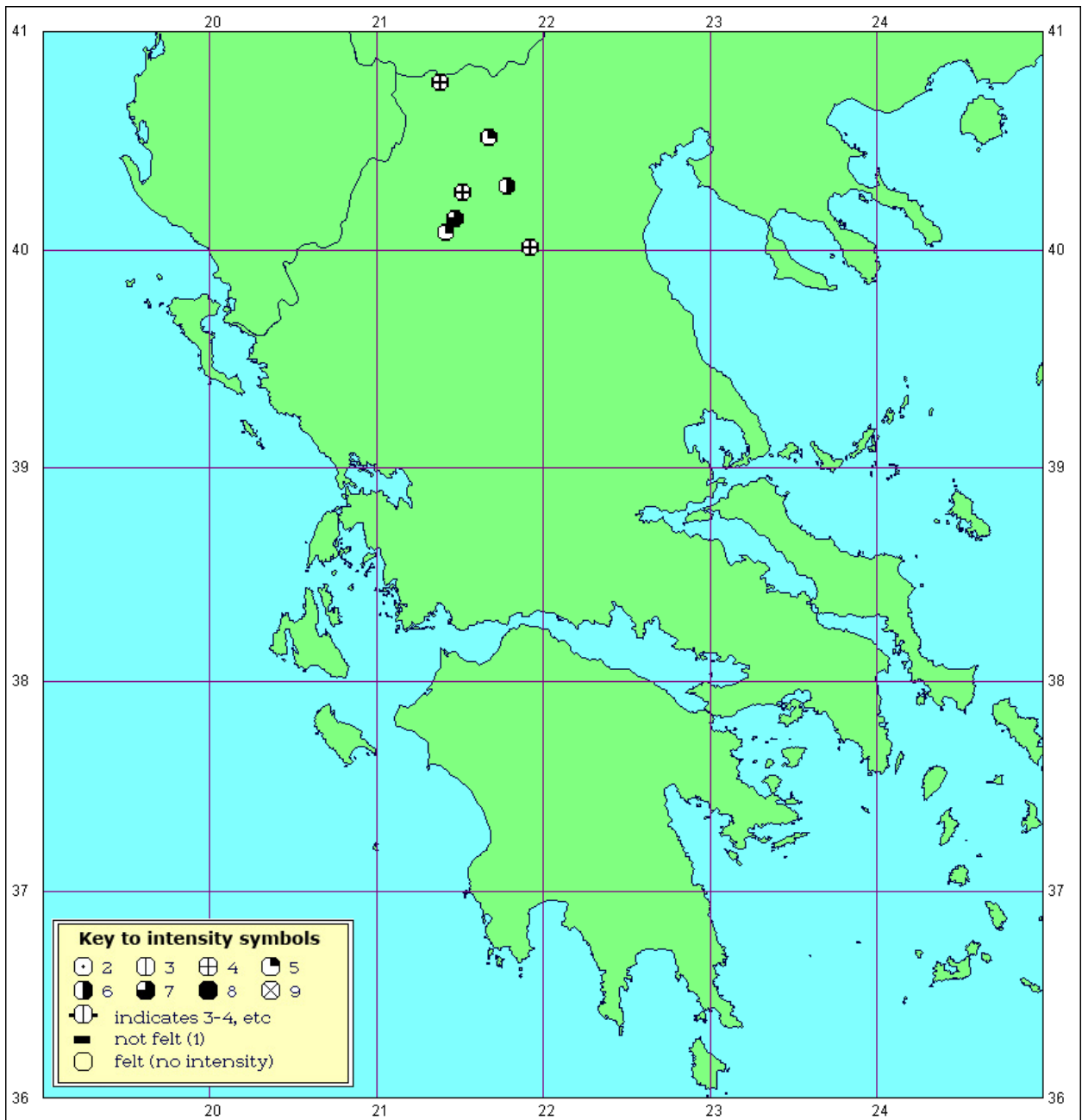


Figure 6 - Test data set #4

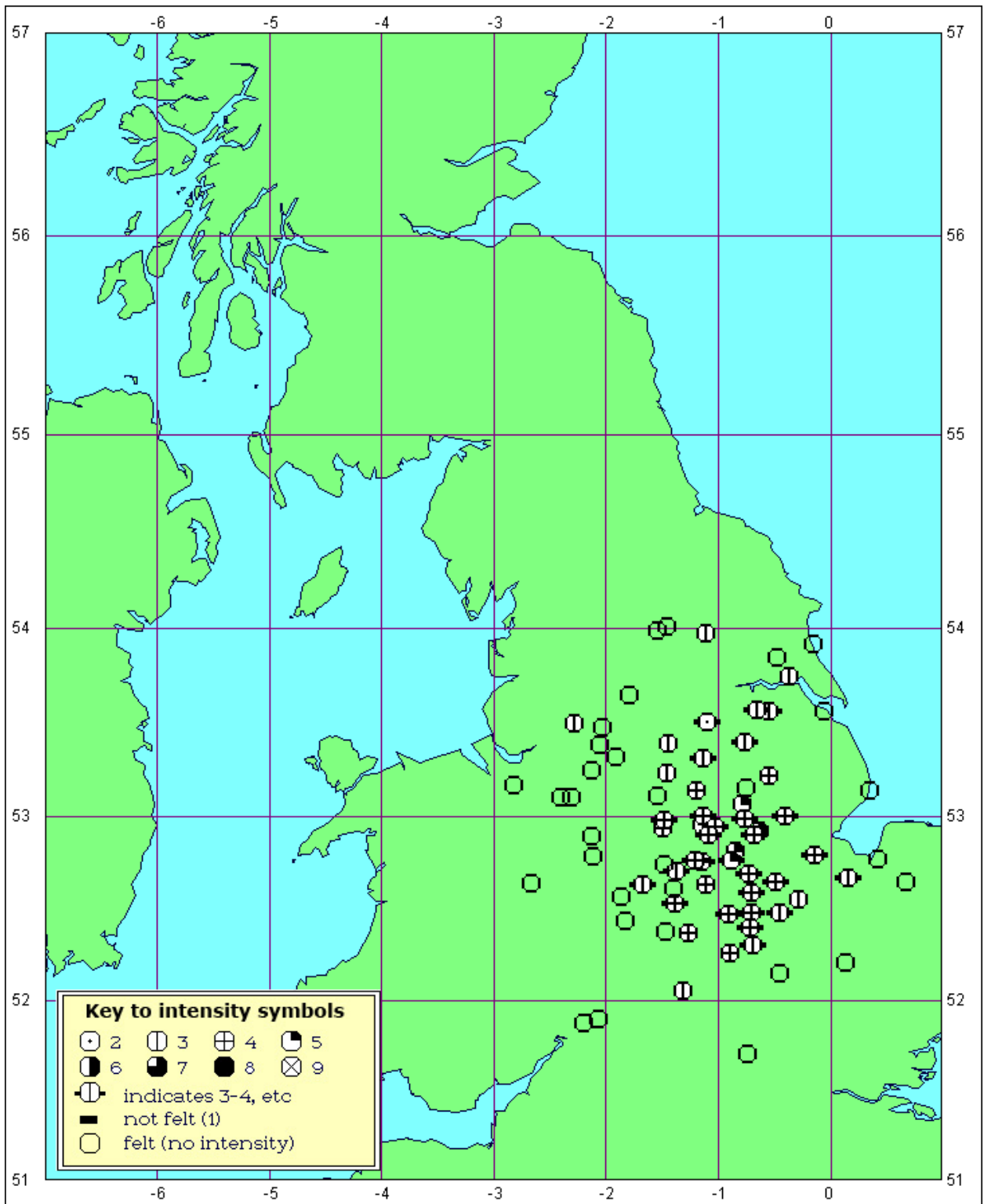


Figure 7 - Test data set #5

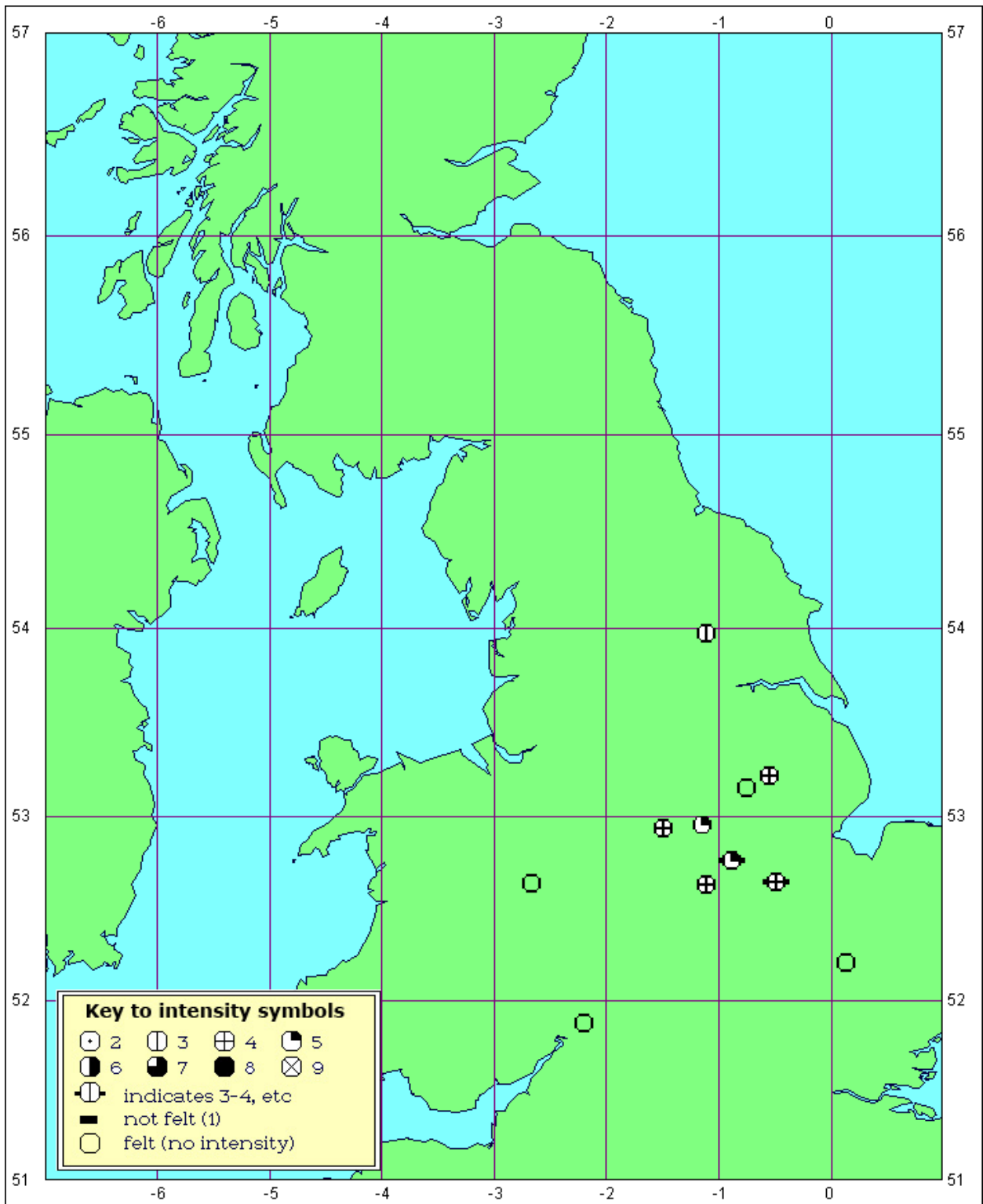


Figure 8 - Test data set #6

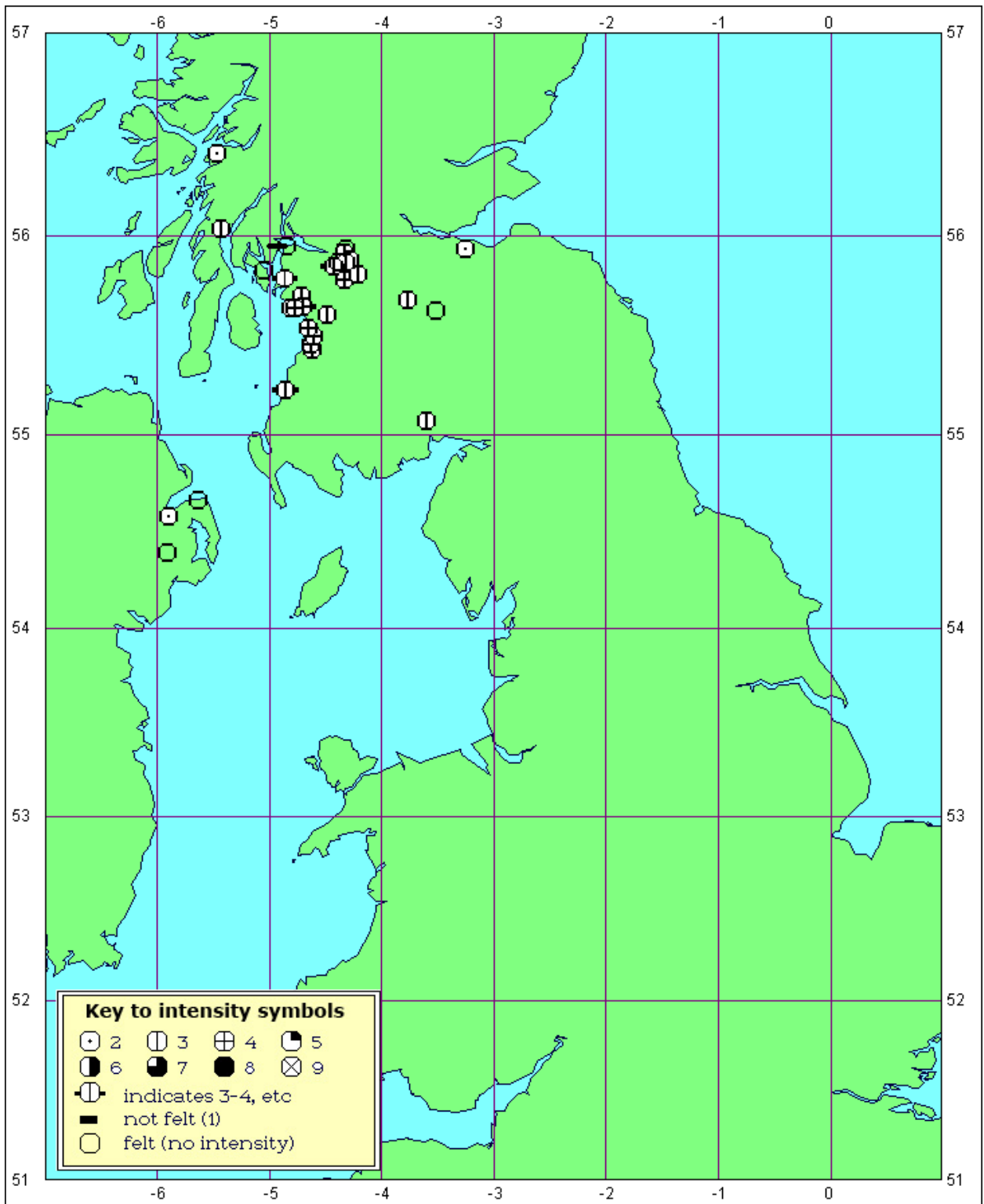


Figure 9 - Test data set #7

Case 1	Method	Distance error	Magnitude error	km rank	Mw rank
	Peruzza	9.0		3	
	Shumila	7.5		2	
	B&W (INGV)	13.0	0.0	4	1
	B&W (ETHZ)	5.3	-0.3	1	3
	Boxer	18.8	-0.9	6	4
	Papazachos	15.8	0.2	5	2
Case 2	Method	Distance error	Magnitude error	km rank	Mw rank
	Peruzza	14.7		2	
	Shumila	52.8		6	
	B&W (INGV)	48.4	0.3	5	1
	B&W (ETHZ)	38.4	-0.3	4	3
	Boxer	6.6	-1.8	1	4
	Papazachos	22.1	0.3	3	1
Case 3	Method	Distance error	Magnitude error	km rank	Mw rank
	Peruzza	4.5		2	
	Shumila	7.8		5	
	B&W (INGV)	5.2	0.0	3	1
	B&W (ETHZ)	4.2	0.2	1	2
	Boxer	5.6	-0.5	4	4
	Papazachos	10.4	0.3	6	3
Case 4	Method	Distance error	Magnitude error	km rank	Mw rank
	Peruzza	19.1		2	
	Shumila	41.7		6	
	B&W (INGV)	21.7	-0.5	3	2
	B&W (ETHZ)	22.5	-0.7	4	3
	Boxer	10.7	-0.7	1	3
	Papazachos	31.8	0.4	5	1
Case 5	Method	Distance error	Magnitude error	km rank	Mw rank
	Peruzza	9.0		4	
	Shumila	7.5		3	
	B&W (INGV)	7.3	1.2	2	1
	B&W (ETHZ)	9.0	1.4	4	3
	Boxer	6.8	1.2	1	1
	Papazachos	7.7	2.0	6	4
Case 6	Method	Distance error	Magnitude error	km rank	Mw rank
	Peruzza	3.4		1	
	Shumila	9.5		4	
	B&W (INGV)	7.8	1.3	3	1
	B&W (ETHZ)	7.5	1.4	2	2
	Boxer	18.0	1.4	5	2
	Papazachos	19.3	2.1	6	4
Case 7	Method	Distance error	Magnitude error	km rank	Mw rank
	Peruzza	52.1		4	
	Shumila	40.6		2	
	B&W (INGV)	62.3	0.9	6	2
	B&W (ETHZ)	34.7	0.8	1	1
	Boxer	45.1	1.2	3	3
	Papazachos	52.8	1.9	5	4

Table 2 - Absolute results of the method comparison

Distance								
Method	1	2	3	4	5	6	7	Total
Peruzza	3	2	2	2	4	1	4	18
Shumila	2	6	5	6	3	4	2	28
B&W (INGV)	4	5	3	3	2	3	6	26
B&W (ETHZ)	1	4	1	4	4	2	1	17
Boxer	6	1	4	1	1	5	3	21
Papazachos	5	3	6	5	6	6	5	36

Magnitude								
Method	1	2	3	4	5	6	7	Total
B&W (INGV)	1	1	1	2	1	1	2	9
B&W (ETHZ)	3	3	2	3	3	2	1	17
Boxer	4	4	4	3	1	2	3	21
Papazachos	2	1	3	1	4	4	4	19

Table 3 - Summary of result rankings for seven cases

4 Developing a solution for NERIES NA-4

4.1 PRINCIPLES

The declared aim of the NA-4 module is to devise procedures that can be used in a consistent way across Europe, to produce homogeneous results. This aim imposes some definite constraints on what methods can be used. Any solution has to be capable of being applied in a systematic way, with the minimum of interpretation. Thus, whatever the advantages of using isoseismals, for practical purposes such an approach must be rejected in favour of using only the IDP files.

A system like Boxer, where a simple software operation can be used to process IDP data files, is highly attractive from a practical point of view, if it can also produce results of sufficient quality.

Secondly, methods that require extensive calibration to work are problematic. One of the points discussed at the February 2007 Workshop was what degree of regionalisation is appropriate. Is it adequate to treat Europe as two regions, interplate and intraplate? Or should intraplate Europe be divided into the Fennoscandian Shield and the rest of NW Europe? Or should regionalisation be pursued to a national level? Or are there significant differences even within different countries? These are questions difficult to resolve.

This is a drawback with the B&W method, even though it performed well in the test, since it depends a lot on calibration exercises. However, a similar approach can be adopted. The Peruzza method performed well in the test exercises, and has the advantage that, at least in theory, no previous calibration is required. Furthermore, since it uses the Kövesligethy model, it is easy to integrate depth determination into the procedure, something not addressed at all by the B&W method. Instead of simply trying to optimise a two-dimensional epicentre against a macroseismic field, the exercise is conducted in three dimensions to obtain the optimal hypocentre. This then leaves magnitude determination to a second step.

We concluded that the best approach to magnitude determination should be based on physical principles, and here the work of Frankel (1994) is extremely helpful. In this, the radius of perceptibility is interpreted in terms of measurable physical parameters such as

shear-wave velocity and the dominant frequency at which weak earthquake shaking is perceived by humans:

$$M = n \log (A/\pi) + (2m / (2.3 \pi^{0.5})) A^{0.5} + C \quad (7)$$

where n is geometrical spreading (taken to be 0.5) and

$$m = (\pi f)/(Q\beta) \quad (8)$$

where f is the predominant frequency of earthquake motion at the limit of the felt area (believed to be 3 Hz), Q is the shear wave attenuation, and β is the shear-wave velocity (3.5 km/s). A is the felt area in km, and C is a scaling constant.

By such procedures, the limit of observation, in other words (for practical purposes) isoseismal 3 EMS, can be related directly to the moment magnitude using physical parameters, of which the regionally-varying one is the crustal attenuation parameter Q , which is often known from other studies, so the appropriate value should already be to hand. There is however, also the constant C in equation (7), which was found by Frankel (1994) to be 1.74 for stable continental crust and 2.53 for California. This appears to be a physical scaling factor and requires regional adjustment.

Thus far, the problem of regionalisation is largely bypassed. If one happens to know, for instance, that Q is higher in one region than another, the appropriate value can be used for any earthquake in that region. If Q is not known at all, it is not too difficult to make an estimate based on known values from tectonically similar regions. The remaining issue is the appropriate value for C ; it was hoped at the outset of this study that a few standard values could be proposed; however, this proved optimistic.

However, Frankel's method deals only with the limit of perceptibility, and this is often poorly known. In countries like the UK, where earthquakes are rare and therefore newsworthy, low intensities are reported quite consistently in source materials after around 1700. In other countries where earthquakes are more common, only higher intensities may be considered worth reporting. Also, in any country, for early historical periods it is likely that reports are highly restricted to damaging intensities.

To solve this, it is necessary to be able to relate isoseismal distribution to the overall size of the felt area. This can be done as follows. Given that the Kövesligethy model has already been used to optimise the depth and epicentre, this model also has something to say about the expected spacing of isoseismals. From equation (6) it follows that

$$(3 \log (R_i/h) + 3 \alpha \log e (R_i - h)) - (3 \log (R_{i+1}/h) + 3 \alpha \log e (R_{i+1} - h)) = 1 \quad (9)$$

- thus, given the radius of the expected felt area, derived from a hypothetical magnitude, one can compute the expected radius of any higher isoseismals, given the depth, which is already known. Therefore: suppose a case where data are only available for intensities 5 and 6 EMS; for a range of magnitudes, one can estimate first, the expected radius of isoseismal 3 EMS from the Frankel model, and from that, the expected radii of higher isoseismals using the Kövesligethy model – including those for intensity 5 and 6. The magnitude that gives the best match to the observed radii (least squares of residuals) can be considered to be the optimal magnitude determination for the earthquake.

4.2 PROBLEMS AND SOLUTIONS

Several problems remain. The first is that the use in the discussion above of terms like “limit of the felt area” and “isoseismal 3” looks in the direction of isoseismals, when it was already declared that the use of isoseismals is not practical in NERIES. Use of the median was at first considered, but this is inconsistent with the principle of an isoseismal being more or less a limiting distance. For practical purposes, we therefore consider the “isoseismal radius” for intensity I to be the 84-percentile of the distribution of epicentral distances of points of intensity I . For the purpose of determining this, intensity values of 5-6 are considered to be 5

(etc). With poor data sets it is possible for (e.g.) the radius for intensity 5 to be less than the radius for intensity 6. In such a case, the radius for 5 would be set to that determined for 6. This can happen in data sets where some lower intensity value is very under-represented in the data set.

The second problem relates to the situation already discussed where the epicentre may be offshore, or in a remote area, and the observed maximum intensity therefore underestimates what would have been observed at the epicentre, leading to possible misestimates in applying the Kövesligethy model. To counter this, optimisation needs to include a range of possible “real” (even fractional) maximum intensities up to some credible limit. The “credible limit” is something that the user must decide on for any particular earthquake.

The third problem relates to the common case where only one IDP exists. It was suggested earlier that these cases may be best decided by expert judgement, but failing this, a conventional approach has to be defined. We therefore propose that in cases where a single IDP exists, the location of the IDP is taken as the epicentre, *faux de mieux*, and it is assumed that the radius of the intensity corresponding to the solitary value is 3 km, with a default depth of 10 km. One can then match this radius to those associated with different magnitudes using the process described above.

In equations (6) and (9) there is a constant 3, which represents the ratio of successive isoseismals as a function of the intervals in the construction of the intensity scale itself, related to the correspondence between the degrees of the scale and ground motion amplitudes. This can be considered also to be a variable – in the analogous work of Blake (1941) it is treated as a variable that may take values over a suggested range of 3 to 6. So while 3.0 is the expected value according to Sponheuer (1960), other values may obtain depending on the scale used and individual practices. So it may be necessary to treat this variable (written as K) as something to be solved for.

One of the features of the Kövesligethy model is (fairly obviously) as depth increases, for the same I_{max} , the affected area also increases (since an earthquake has to be larger if it produces the same I_{max} value despite an increase in depth). Thus in the case of a data set with a few high value IDPs and many IDPs of medium or low intensity, a good fit can sometimes be found by a notional epicentre at the edge of the felt area and a large focal depth. Such a combination will produce high residuals near the epicentre, but these will be offset by generally low residuals at lower intensities. Measures to counteract this will be described in the next section.

Of course, in some cases such an interpretation may be correct, for instance in the case of a large offshore earthquake where all the intensity data is at some distance from the actual epicentre.

4.3 STAGES IN THE APPLICATION OF THE METHOD

In the interests of efficiency, a search routine to determine the epicentre is better than a brute force testing of every possible position. The strategy adopted is as follows.

Firstly, an initial estimate of the epicentral position is made on the basis of the centroid of the high intensity points. The procedure adopted follows Gasparini et al (1999) to some degree, in that the centroid is based on a data set trimmed of outliers, approximately 25% of points being eliminated. An initial centroid estimate is made based on all high intensity points (the definition of “high” starting with the highest value points and adding progressively the next highest until an adequate number of points is available). The data points furthest from this initial centroid are trimmed, and then the centroid of the remaining points is calculated. This position, the centroid solution, is then used as a starting position for a search procedure. If one takes this position as 0,0 in a kilometre grid system, the points to be considered at each step of the search are -R,-R -R,0 -R,R 0,-R 0,0 0,R R,-R R,0 and R,R. The initial value of R is taken to be 64 km. The RMS is calculated for each of these nine points; the one with the

lowest RMS is taken to be 0,0 for the next step of the search, with R reduced by 50%. This is repeated so that R is successively 64, 32, 16, 8, 4, 2, 1 and 0.5 km. The point with the lowest RMS at the final iteration is taken to be the best-fit epicentre.

The RMS is calculated for each point by fitting the Kövesligethy model to the IDP set. For points in the macroseismic field distant from the high intensities it is not appropriate to fit this using the actual I_0 of the earthquake, so for each test epicentre the nearest three IDPs are examined, and the highest of the three is used as the I_0 for that position. It is then allowed to vary upwards by the margin previously specified. At each pass, the I_0 above the observed intensity is optimised. At this stage, depth is held fixed at a default value supplied by the user.

The estimation is made using a normalisation routine, in order that the solution is not dominated by low intensity observations, which may be far more numerous. The contribution of each IDP is weighted according to the number of observations for that intensity, so that all intensities are given equal weight, and then a further slight weighting is applied so that higher intensities are actually given more importance.

Having solved for epicentre, the next stage is to compute the expected intensity radii for each magnitude between 3.0 and 8.5. These expected radii are compared to the observed radii, and the magnitude with the lowest RMS is taken to be the result.

How the comparison is made is open to some possibilities. The obvious way to compare predicted and observed intensity radii is to take the square of the difference between the radii in km. A possible drawback to this is that the radii of lower intensities will have more impact simply because the numbers are larger. Other solutions are possible – for instance, to take the absolute difference in radii as a proportion of the observed radius. Several schemes were investigated, but in practice did not appear to give better results than a straightforward RMS approach.

Depth is also calculated using the intensity radii rather than the IDPs, as this was found to be more robust. Given the epicentre, a joint optimisation on depth and notional I_0 value is performed, subject to the user-specified limit on how much the I_0 may exceed the observed I_{max} . This calculation is done before the magnitude calculation, so that the magnitude can be estimated using hypocentral distances.

Magnitude, depth and I_0 are calculated in the same way for both the centroid solution and the attenuation solution.

4.4 ESTIMATING UNCERTAINTIES

Another issue is that of estimating the uncertainty in the results. In the case of magnitude, for instance, one is used to seeing a magnitude given as 4.5 ± 0.2 . What this means is that the overall magnitude is the mean of a number of station magnitudes, and the standard deviation of the mean is 0.2. If it is demanded that macroseismic magnitudes be written the same way, then a problem arises – as described above, a macroseismic magnitude is a best-fit value. It is not the mean of several values, and cannot have a standard deviation. It is possible to have a goodness-of-fit estimate, but this is not, strictly speaking, an uncertainty. There are several possibilities.

Firstly, the overall RMS may be low or high. If high, it means that the data do not fit the theoretical model very well, which implies that the magnitude may not be very certain.

Secondly, the RMS distribution may not constrain the magnitude well – in other words, a range of magnitudes give approximately the same RMS. In such a case, the uncertainty in the final value is much greater than cases where the best-fit magnitude gives a substantially lower RMS than other possible values.

Thirdly, one may have a well-constrained value with low RMS which is nevertheless wrong. This could be because the system has been poorly calibrated. It could also be due to an anomalous data set; and anomalous cases cannot be avoided entirely.

Equally, of course, an estimate may appear to be poorly constrained and yet is completely accurate. In general, one tends to look to instrumental determinations of parameters as the “correct” answers that the macroseismic estimates need to match. In fact, instrumental values are themselves uncertain, and there are many cases where macroseismic estimates may be much more accurate than the instrumental solutions

The solution adopted, after some experimentation, is based on the degree to which the distribution of RMS values over the parameter space of interest is steeply or shallowly curved in the region of the minimum value, a steep curve indicating a well-constrained value. What measure is taken is always going to be arbitrary; the decisions adopted here were based on empirical tests on a variety of earthquakes from the Italian database, comparing the RMS distributions with instrumental parameters and other macroseismic estimates made by members of the INGV Milano team (Meletti, Stucchi and Gomez Capera 2008 *pers. comm.*).

For the uncertainty in the epicentre, the approach taken is based on the relative difference in RMS value at various distance ranges. This exploits the search method already described. At each step in the search, the proportional difference between the highest and lowest RMS is recorded. As the search area gradually contracts around the overall RMS minimum at each iteration, this difference gets less and less. Some empirical tests showed that the distance between the best-fit epicentre and the instrumental epicentre generally was associated with a difference in RMS of a factor of 2. For the purposes of this method, the distance interval associated with an RMS factor of 2 is thus considered to be the uncertainty. This is achieved by interpolation. So for instance, if at Step 4 (delta = 8 km) the worst RMS is 2.5 times the best RMS, and at Step 5 (delta = 4 km) the factor is 1.5, the uncertainty would be considered to be ± 6 km.

Examining magnitude RMS distributions, it appeared that the same approach gave reasonable results. We thus define here the uncertainty in the magnitude as the increase or decrease in the best-fit magnitude value necessary to double the RMS score. This is illustrated in Table 4.

<i>Magnitude</i>	RMS
4.5	5.89
4.6	4.71
4.7	3.84
4.8	2.35
4.9	4.02
5.0	4.99
5.1	6.05

Table 4 - Example of magnitude uncertainty

For a result as shown in Table 4, the solution would be given as 4.8 ± 0.2 .

The method described above was coded into an application called MEEP (Macroseismic Estimation of Earthquake Parameters). The operation of this program is described in the next section.

Matelica	43.256	13.009	4
Pedaso	43.097	13.841	4
Perugia	43.106	12.386	4
San Severino Marche	43.229	13.177	4
Spoleto	42.732	12.736	4
Tortoreto	42.803	13.914	4
Castel del Monte	42.364	13.727	3
Guidonia Montecelio (Guidonia)	41.992	12.722	3
Carapelle Calvisio	42.298	13.685	2
Pesaro	43.905	12.905	2

Where PPPPP indicates where the latitude value is to be found, LLLLLL the longitude, VVVVV the intensity value, UUUUUUUUUUUUUUUUUUUUU the place name, and the other codes are for comment and quality factor, neither used in this file. “F” in the intensity position would indicate “felt” with no intensity capable of being assigned. These points are ignored in the analysis. The file above is the 26 November 1972 Montefortino earthquake, taken from the Italian database (INGV 2008).

In addition, a file called Consts.dat is used. This will look as follows:

```
Margin for Io above observed value.....:0.5
Regional Q value.....:300.0
Scaling factor C.....:2.09
Default depth value.....:10.0
Regional alpha.....:0.005
Iseismal K factor.....:3.9
Intensity file quality factor threshold.:1
Frequency of human perception (Hz).....:3.0
Geometric spreading (N).....:0.5
```

The parameters in this file are ordered in such a way that those most likely needing modification in different cases are at the top. The last three probably never need to be changed, and are only included at all for the purpose of possible experimentation.

The first line sets the permissible increase in I_0 above the observed maximum intensity, for use when fitting the attenuation model. The program will test possible values in increments of 0.1 of an intensity unit. For onshore earthquakes where there is a good density of IDPs in the epicentral region, the margin should be low, 0.5 to 1.0 or even as low as 0.1. For earthquakes believed to be offshore, the margin may be higher. This is the only parameter that may need to be adjusted on a quake-by-quake basis.

The second line is the regional Q factor for Lg waves at 3.0 Hz, which can be taken from the literature, or estimated (see next section).

The third line is the scaling factor C from Frankel's (1994) perception model; this was found by Frankel (1994) to be 1.74 for stable continental regions (SCR) and 2.53 for California. Suitable European values need to be investigated (see next section).

The next line is the initial depth value to be used.

The fifth line is the regional value of alpha in the Kövesligethy equation. Suitable values are given in Karník (1969), or can be found by a joint optimisation study. The fifth line relates to the scaling of isoseismals; calibration is discussed in the next section.

The following lines will not normally be adjusted.

The sixth line is for cases where the IDP file includes some quality factor where higher values indicate poorer quality data (see Musson 1998c) – IDPs with quality factors higher than the threshold value will not be used. For files without quality factors, this value is unimportant – all data will be used. In most cases this can be ignored.

The remaining two lines are from Frankel (1994), are self explanatory, and should not need changing.

When run from the DOS prompt, the program prompts for a file name. It expects, but does not require, files to have the suffix .int (in line with the WIZMAP II analysis program). If the file has this extension, the extension need not be specified. Thus if the file is called 16611001.int, one could type simply 16611001 as the filename. If it were called Testfile.dat, the full filename would be required. The Consts.dat file needs to be in the same directory.

The program now asks if the user already knows the epicentre. This is for cases where the user may have already fixed the epicentre by judgement from a poor data set, and wishes only to estimate the magnitude. If this option is required, the user may type "y" or "Y" in response, and the program will prompt for the latitude and longitude of the epicentre that has been selected.

The output then looks as follows:

```
**** MEEP v1.6 ****
```

```
Data file name.....:19721126
```

```
Do you know the epicentre?...:n
```

```
Total number of datapoints      =    58
```

```
Quality threshold                =     1
```

```
Number of usable points          =    58
```

```
Number of intensity points read =    58
```

```
*** SUMMARY OF DATA ***
```

```
# of points of intensity         8 =     2
```

```
# of points of intensity        7-8 =     6
```

```
# of points of intensity         7 =     3
```

```
# of points of intensity        6-7 =     3
```

```
# of points of intensity         6 =    13
```

```
# of points of intensity        5-6 =     1
```

```
# of points of intensity         5 =    13
```

```
# of points of intensity        4-5 =     7
```

```
# of points of intensity         4 =     6
```

```
# of points of intensity        3-4 =     0
```

```
# of points of intensity         3 =     2
```

```
# of points of intensity        2-3 =     0
```

```
# of points of intensity         2 =     2
```

```
# of points felt/no intensity    =     0
```

```
# of points not felt             =     0
```

```
***      END OF SUMMARY      ***
```

```
Maximum intensity (Imax)        =     8
```

```
Number of Imax points           =     2
```

```
Second highest intensity        =    7-8
```

```
Number of 2nd highest I points =     6
```

```
Centroid solution =  42.980  13.533
```

```
Step 1 Delta = 64.0 RMS = 74.144 ( 42.980 13.533) 8.0 / 8.4
```

```
Step 2 Delta = 32.0 RMS = 74.144 ( 42.980 13.533) 8.0 / 8.4
```

```
Step 3 Delta = 16.0 RMS = 74.144 ( 42.980 13.533) 8.0 / 8.4
```

```
Step 4 Delta =  8.0 RMS = 70.829 ( 43.003 13.440) 8.0 / 8.4
```

```
Step 5 Delta =  4.0 RMS = 69.383 ( 42.991 13.487) 8.0 / 8.4
```

```
Step 6 Delta =  2.0 RMS = 68.951 ( 42.981 13.456) 8.0 / 8.4
```

```
Step 7 Delta =  1.0 RMS = 68.824 ( 42.978 13.468) 8.0 / 8.4
```

```
Step 8 Delta =  0.5 RMS = 68.813 ( 42.982 13.470) 8.0 / 8.4
```

```
Attenuation solution =  42.982  13.470 (+-  4.1 km)
```

```
Notional Io                    =    8.4
```

```
For intensity 8 on 2 points, effective distance = 25.3 km
```

```
For intensity 7 on 9 points, effective distance = 33.1 km
```

```

For intensity 6 on 16 points, effective distance = 41.6 km
For intensity 5 on 14 points, effective distance = 52.3 km
For intensity 4 on 13 points, effective distance = 80.2 km
For intensity 3 on 2 points, effective distance = 135.7 km

```

```

----- RESULTS -----
Centroid:
Epicentre at 42.980 13.533 M = 5.3 (+- 0.4) h = 7 km (Imax = 8.4)
Attenuation:
Epicentre at 42.982 13.470 M = 5.2 (+- 0.4) h = 7 km (Imax = 8.4)

Locational uncertainty +- 4.1 km
Stop - Program terminated.

```

This is the result from processing the Montefortino file shown above. The IDPs are plotted as Figure 10.

The first part of the output summarises the intensity data in terms of the number of points for different values, including points recorded only as “felt” or “not felt/ intensity 1” (which are not present in this file, and not used anyway).

The program first computes the centroid solution, and prints this. This is the point used as the starting point for the grid search for the attenuation method. The lines beginning “Step” show the progress of the grid search over successive iterations at progressively shorter search radii (Delta) as the program converges on the location with the least misfit. This is the RMS value, based on the average intensity disparity. The values in brackets are the current best estimate of the epicentre.

The two following values at the end of each line are the base I_0 value for this trial and the optimised I_0 value. The base I_0 is the largest of the nearest three IDPs, and the optimised value is the fractionally higher value that minimises the residuals.

The final attenuation solution is then printed; the epicentre is given as latitude and longitude and uncertainty in km. In this example, the actual distance between the estimated epicentre and the instrumental one is about 2 km (see Figure 10). The fractional I_0 associated with the chosen epicentre is printed.

The program then lists the “effective distances”, which are the equivalents of isoseismal radii, being the 84-percentile distances for each intensity values, rounding down uncertain values, so 3-4 counts as 3. Intensity 2, if present, is not used.

Finally, there is a section of results. Two determinations are given, one based on the centroid, the other based on the attenuation method. For each, the latitude and longitude is printed, the magnitude and uncertainty on the magnitude, the depth, and the fractional I_0 value. The epicentral uncertainty for the attenuation solution is then printed again (uncertainty for the centroid epicentre is undefined).

The instrumental magnitude for this event is 5.3 Mw.

The program also produces a separate output file (which in this case would be called 19761126_sol.txt) which contains additional output from the stages through which the parameters were optimised. This can be useful for determining how a particular solution was arrived at.

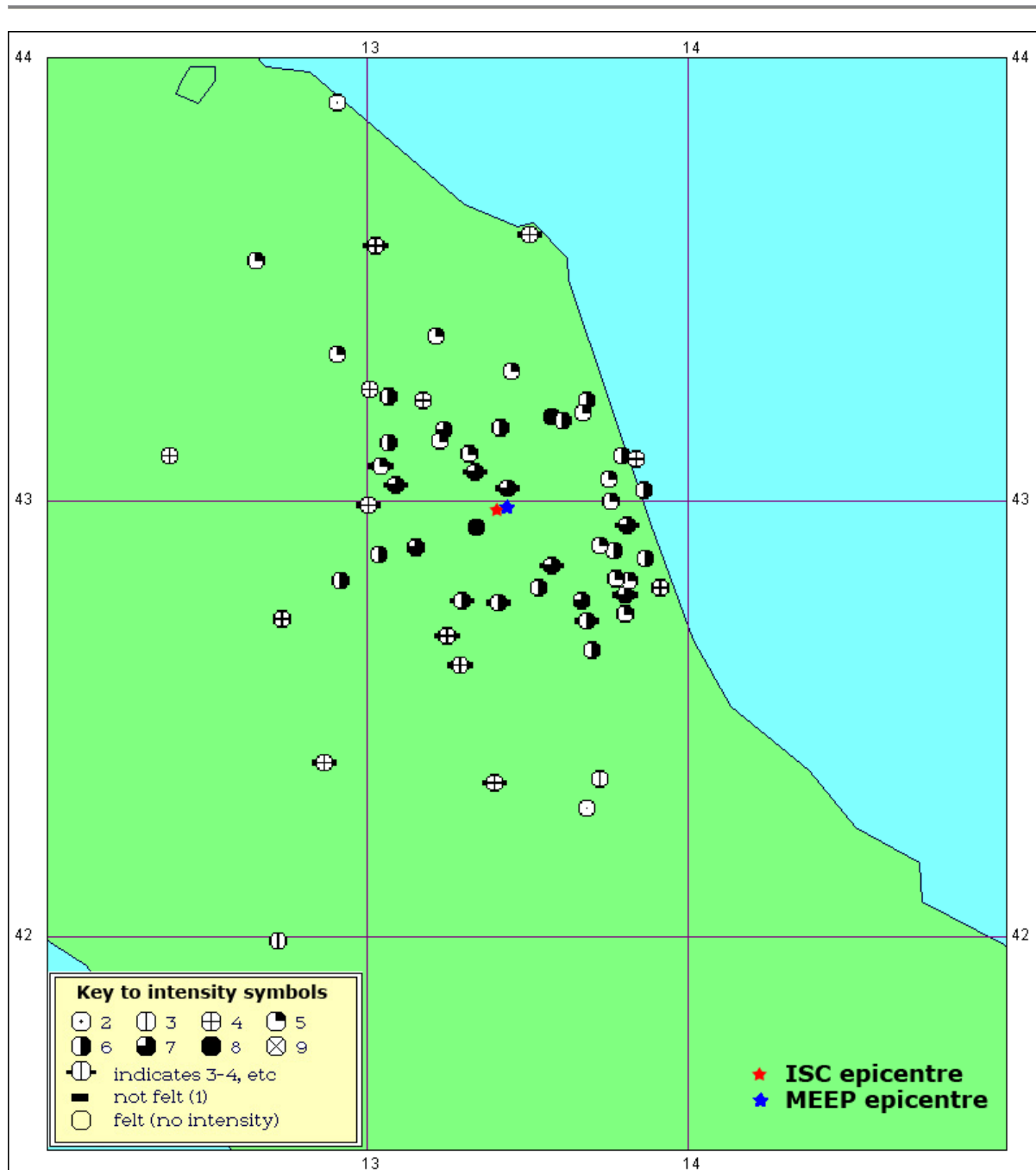


Figure 10 - Earthquake of 26 November 1972

6 Calibration

Since the issue of calibration cannot be completely avoided, the intention here is to make it as straightforward as possible. A companion program, CALIMEEP, is provided. This section of the report describes the stages of preparing a calibration, using Italy as an example.

First, a calibration data set of events needs to be prepared. These will be drawn from available earthquakes that have both IDPs and instrumental parameters. The selection of events will depend on what is available, but ideally should cover the expected range of magnitudes evenly, and not be dominated by small magnitude events. Also, it is as well not to include obviously anomalous events.

For Italy, 28 events were selected, varying in magnitude from 3.7 to 7.3 Mw. Each IDP set was allocated a separate file, in the format given in Section 5, and with the name in the format YYYYMMDD.int. An additional file has to be prepared with the name List.txt. This looks as follows:

```
19081228 38.00 15.50 7.3 10
19801123 40.76 15.31 6.9 10
19300723 41.10 15.70 6.7 10
19200907 44.30 10.30 6.5 10
19190629 44.00 11.50 6.3 10
19840507 41.76 13.90 5.9 10
19431003 42.90 13.70 5.8 10
20021101 41.74 14.84 5.7 10
19840429 43.20 12.59 5.7 10
19840511 41.77 13.89 5.5 10
19710715 44.78 10.29 5.4 10
20030914 44.26 11.38 5.3 10
19580624 42.35 13.43 5.2 10
20041124 45.69 10.52 5.1 10
19800614 41.82 13.66 5.0 10
20000821 44.77 8.43 4.9 10
20030126 43.90 11.93 4.7 10
20031230 41.64 14.85 4.6 10
20051215 42.74 12.76 4.5 10
20020418 40.58 15.55 4.4 10
20021113 45.65 10.14 4.3 10
20041209 42.79 13.79 4.2 10
20050521 40.99 14.52 4.1 10
20050418 44.72 9.35 4.0 10
20050413 44.69 9.33 3.9 10
20041204 45.94 12.00 3.8 10
20050301 41.67 14.87 3.7 10
```

Each line represents one earthquake, and contains the following information: the date, in the format YYYYMMDD (this serves also to specify the filename, as the suffix ".int" is understood); the latitude, longitude, magnitude and depth. Since in this case individual depths were not available, all depths were set to 10 km. The list above serves also to indicate which events were chosen for the calibration. They are listed in order of descending magnitude (this is not required, but is useful when examining the results).

Next, one needs to select an appropriate value of α for the Kövesligethy model. This was taken to be 0.005 (Karník 1969). An initial value for Q (shear-wave attenuation at 3 Hz) also needs to be chosen. Evidently, for all of Italy, this is going to vary, so for general Italian use an average value is needed. The results are not sensitive to small changes in this parameter, so rounded values are sufficient. Following Bianco (2002) a value of 300 is used as an initial trial.

The program CALIMEEP is now run from the DOS prompt in the same directory as contains List.txt and all the IDP files. The output looks like this:

```
Value for Q.....:300
Value for Alpha.....:0.005
For K = 3.00 and C = 1.97 misfit is 2.68
For K = 3.10 and C = 1.99 misfit is 2.72
For K = 3.20 and C = 2.00 misfit is 2.71
For K = 3.30 and C = 2.01 misfit is 2.95
For K = 3.40 and C = 2.04 misfit is 2.99
For K = 3.50 and C = 2.07 misfit is 2.94
For K = 3.60 and C = 2.04 misfit is 2.74
For K = 3.70 and C = 2.05 misfit is 2.69
For K = 3.80 and C = 2.08 misfit is 2.62
```

```

For K = 3.90 and C = 2.09 misfit is 2.33
For K = 4.00 and C = 2.11 misfit is 2.50
For K = 4.10 and C = 2.12 misfit is 2.71
For K = 4.20 and C = 2.14 misfit is 2.83
For K = 4.30 and C = 2.16 misfit is 2.87
For K = 4.40 and C = 2.17 misfit is 2.67
For K = 4.50 and C = 2.19 misfit is 2.61
For K = 4.60 and C = 2.20 misfit is 2.58
For K = 4.70 and C = 2.21 misfit is 2.51
For K = 4.80 and C = 2.23 misfit is 2.61
For K = 4.90 and C = 2.23 misfit is 2.72
For K = 5.00 and C = 2.24 misfit is 2.64
For K = 5.10 and C = 2.26 misfit is 2.47
For K = 5.20 and C = 2.26 misfit is 2.56
For K = 5.30 and C = 2.28 misfit is 2.71
For K = 5.40 and C = 2.29 misfit is 2.99
For K = 5.50 and C = 2.29 misfit is 2.97
For K = 5.60 and C = 2.31 misfit is 2.86
For K = 5.70 and C = 2.31 misfit is 3.00
For K = 5.80 and C = 2.32 misfit is 2.88
For K = 5.90 and C = 2.32 misfit is 2.93
For K = 6.00 and C = 2.33 misfit is 3.13

```

```

Inst Macr Magnitude

```

```

7.3 7.5
6.9 6.9
6.7 6.9
6.5 6.4
6.3 5.8
5.9 5.9
5.8 5.8
5.7 5.3
5.7 5.0
5.5 5.7
5.4 5.8
5.3 5.1
5.2 4.8
5.1 5.7
5.0 4.7
4.9 5.0
4.7 4.6
4.6 4.6
4.5 4.6
4.4 4.3
4.3 4.8
4.2 4.2
4.1 4.3
4.0 4.3
3.9 3.9
3.8 4.0
3.7 3.9

```

```

Value for K is 3.90 ; value for C is 2.09 (misfit = 2.33)

```

The first section shows successive tests of values for K , over a much wider range than is strictly required – the text above has been edited to show only the results from 3.0 to 6.0, but in fact 1.5 to 10.0 are tested; the results for each value are printed largely to show that the program is running.

The next section shows the instrumental and macroseismic magnitudes for the events in the calibration set, using the best fit parameters. A quick inspection of this shows that there are no biases – values are just as accurate at high magnitudes as at low magnitudes (see Figure 11).

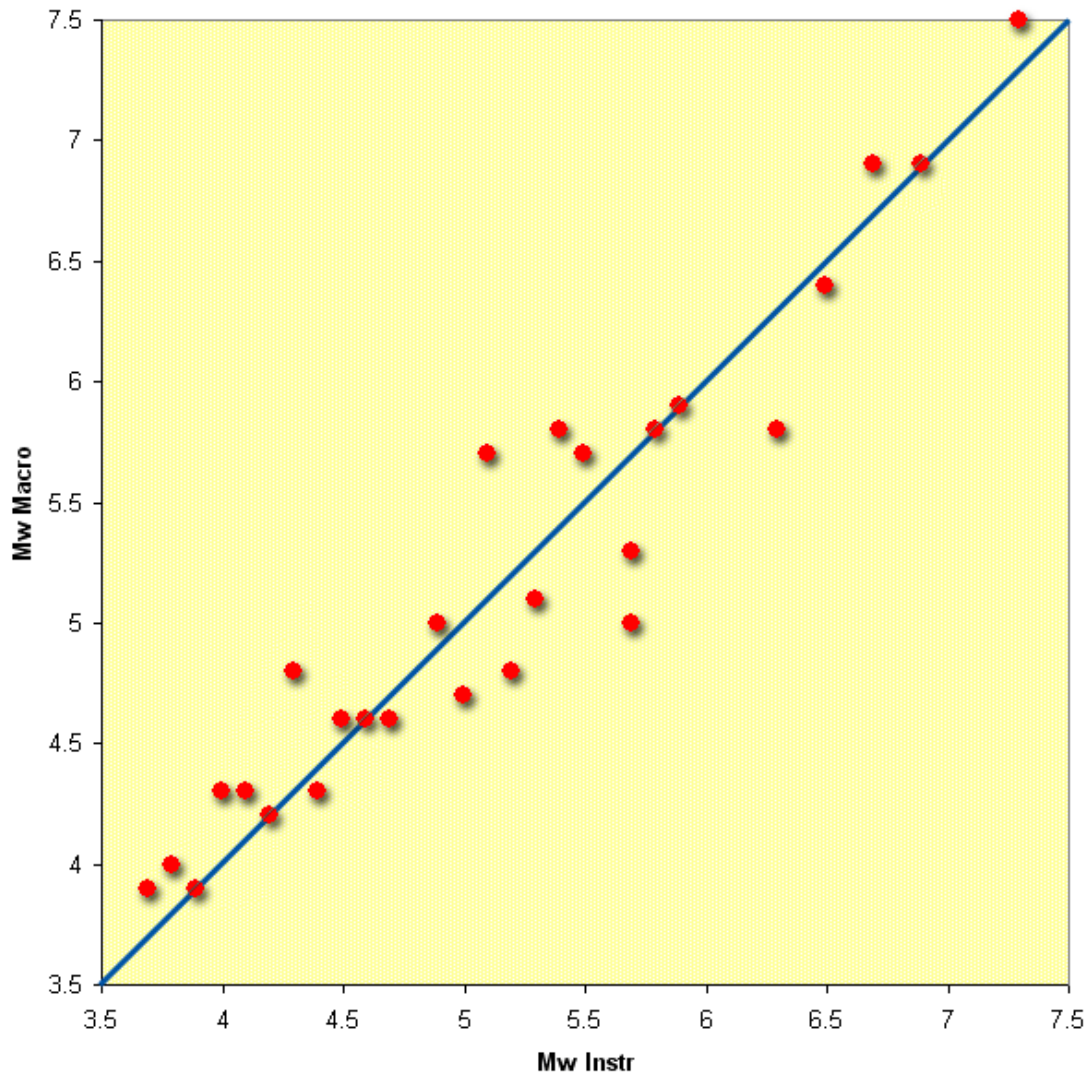


Figure 11 - Computed magnitudes for the calibration events

Finally, the best values are printed with the total misfit (i.e. the RMS for the whole data set).

To check the suitability of Q , the program can be run again, varying the value. The RMS for $Q=300$ is 2.33. For values of 290 and 310, RMS is 2.45 and 2.43, so 300 is a suitable value.

7 Earthquakes with very few IDPs

As previously mentioned, if an earthquake has only one IDP, this is necessarily the epicentre, with an assumed radius of 3 km for the epicentral intensity taken for the magnitude calculations. This means by implication that, for the same calibration, all earthquakes with only one IDP and the same I_0 value will have the same magnitude.

As an example, Figure 12 shows the magnitudes associated with different possible I_0 values for Italy using the calibration results shown previously.

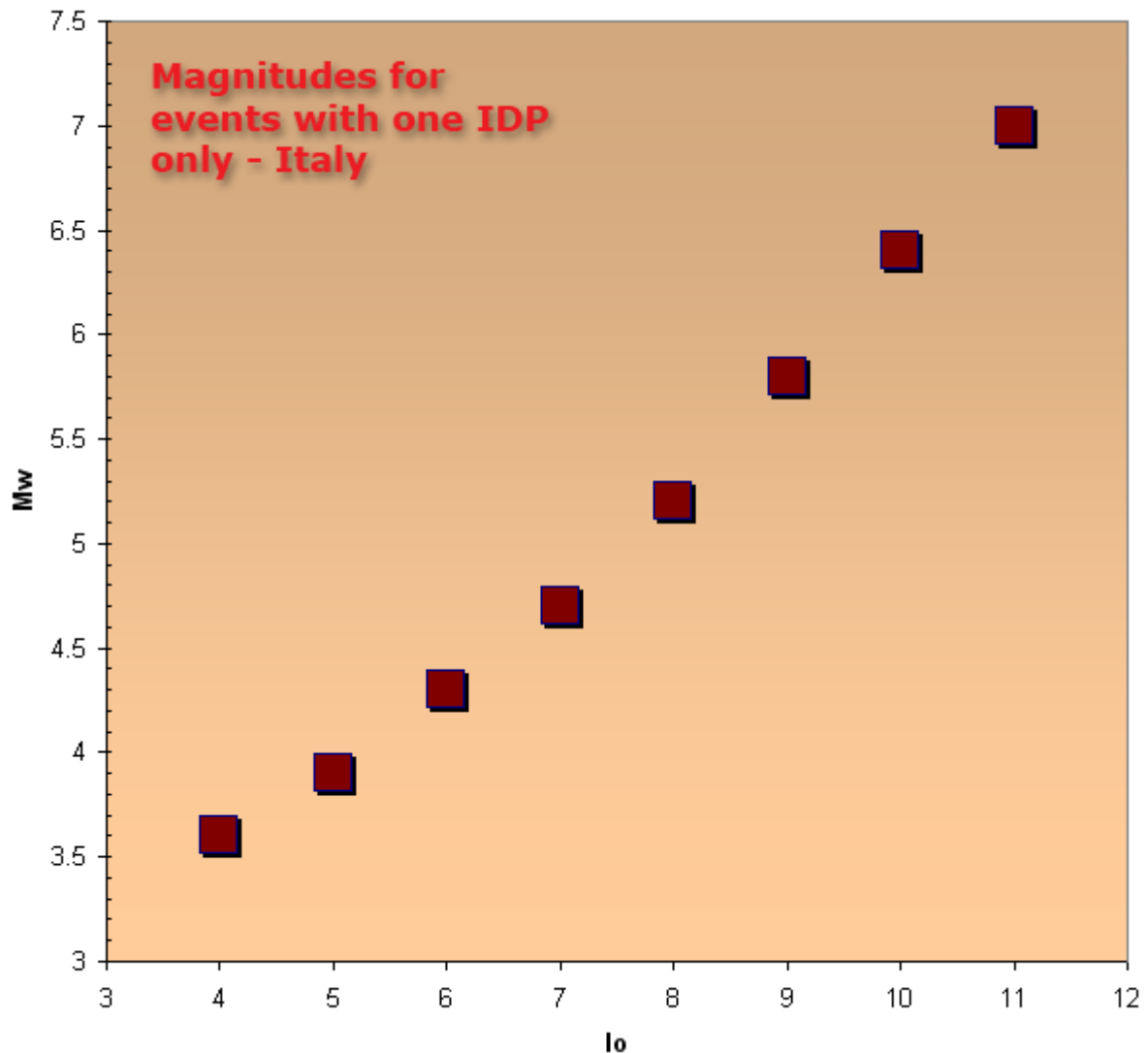


Figure 12 - Magnitudes for single IDP cases: Italy

For earthquakes with very few IDPs but more than one, the program does attempt to find a solution, but the location will clearly be very unstable. For a two-point case, many possible solutions will give the same fit, and the final choice will depend more or less randomly on effects produced by the granularity of the search (i.e. the discrete increments in depth and maximum intensity) having very tiny differences in the RMS value. In cases where the epicentre looks doubtful, the centroid position can be taken instead (and the magnitude recalculated by forcing this epicentre).

8 Earthquakes with no IDPs

The procedures described above deal with historical earthquakes that have some IDPs available, and give a solution even in cases where only one IDP is known. It is also possible to calculate a magnitude for an earthquake with only one IDP which is obviously not the epicentre. Thus, for instance, given an earthquake felt in Venice at intensity 4, if one assumes an epicentre in the vicinity of 46.2 13.1, one obtains a magnitude of 5.8 M_w , which is quite believable.

One may have data files with no intensity values (only indications of “felt” or “damage”). It would have been possible to design MEEP to operate even with such files as these, by

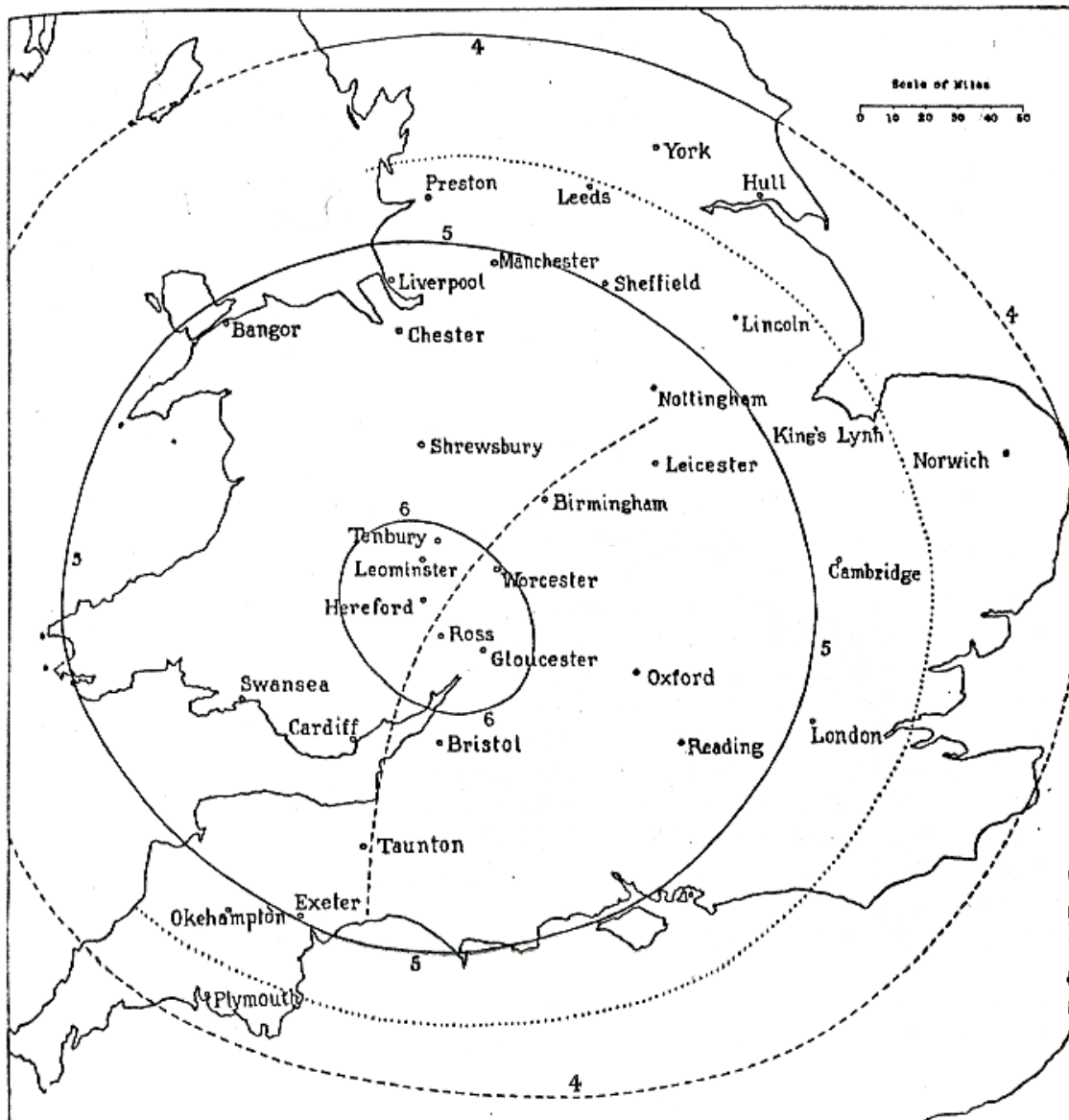


Figure 13 - Isoseismal map of the earthquake of 15 August 1926

This map is based on an extensive set of questionnaire data, collected by Charles Davison, all of which was destroyed after his death in 1941 by his widow. What can be made of it? The first problem is that the intensity values given are in the Davison Scale, a variant on the Rossi-Forel Scale (Davison 1900). In fact, this eccentric scale is such that an exact correspondence with normal Rossi-Forel values cannot be taken for granted.

If this map were all that existed, one might, in determining parameters by hand, make an estimate of the epicentre from the centroid of the innermost isoseismal, and estimate the magnitude from the approximate isoseismal radii. This could be duplicated using MEEP by taking a series of points around the perimeter of each isoseismal, assigning them an intensity using some correspondence of scales, and then treating these as if they were intensity observations.

Alternatively, since a number of towns are marked on the map, one could assume that these towns refer to observations and can be treated as IDPs with an intensity value taken from the enclosing isoseismal. There is no guarantee that this assumption is correct – the towns may have been added to the map by Davison merely for geographical reference – there is no way of knowing.

Davison's method of drawing isoseismals is discussed by Principia (1982), based on data presented in Davison (1899). Their conclusion is that Davison tended to draw an isoseismal N such that it bounded, with a smooth curve, all points of intensity N , including outliers. In which case, it would be unlikely if the procedure of generating artificial IDPs to trace the isoseismals would give a result corresponding to reality.

This is shown to be the case by a revaluation of the earthquake by Musson et al. (1984) from newspaper reports and other contemporary material, in which it is shown that Figure 13 completely misrepresents the distribution of effects from this earthquake, even to the extent that an epicentre based on Figure 13 would be seriously in error. The true epicentre is near the northern limit of the isoseismal 6 in Figure 13.

This is an example where a published isoseismal map is untrustworthy and known to be so. There will be many other examples of published maps that are trustworthy – and many cases where the trustworthiness is unknown. Different cases need to be taken on their merits.

An alternative is to assign no parameters at all, as in the case of catalogues such as Musson (1994), where most medieval earthquakes are represented only by a date, a location name, and sometimes an approximate I_{max} value. Caution should be taken not to over-interpret very poor data (Musson 1998d). A report that merely states, "This year was a great earthquake in England" really cannot be given parameters beyond a date. For very early earthquakes, omitting parameters simply limits the date range for which the catalogue is usable for hazard analysis. It would not be a good idea to omit parameters for an earthquake such as the 15 August 1926 event shown in Figure 13, if Davison's (1927) map was all that one had. But any solution adopted for forcing parameters out of Figure 13 (such as choosing artificial IDPs in such a way as to try and compensate for the probable exaggeration of area and intensity) would not necessarily be exportable to any other case.

9 Offshore earthquakes

In estimating the RMS values for the intensity field, the program is essentially calculating a goodness-of-fit surface and selecting the minimum point. In tests, the search algorithm used here (with the successive steps of decreasing distance) appears to work well, finding the minimum value efficiently, compared to a brute force approach in which every possible point over a wide area is tested. Typically the brute force approach gave the same result, but took much longer.

One problem that does not admit of an easy solution is determining the maximum allowed intensity above the nearest observed intensity. This is critical for offshore events. The higher the notional intensity that could be hypothesised at the offshore epicentre, the further from shore the epicentre can be.

As an example, consider the Palermo earthquake of 6 September 2002 (Figure 14). In this figure, successive determinations have made of the epicentre in which the I_0 value has been allowed to be higher than the nearest observed intensity by +0.5, +1.0, +1.5, +2.0, + 2.5 and +3.0 degrees. As can be seen, once the margin is increased above +1.0, the epicentre steadily moves further out to sea, until at a margin of +3.0 degrees, it is around 6 km from the instrumental position. Also, the further offshore the epicentre, the larger the magnitude will be.

For a historical event, there is no clear way to assess the best margin to use other than by expert judgement. There will also be cases where the distribution of IDPs shows a similar pattern to Figure 14, but where the empty space is not sea, but a foreign country from which data were not collected. In that case, one would be much less inclined to allow a high margin for the I_0 value.

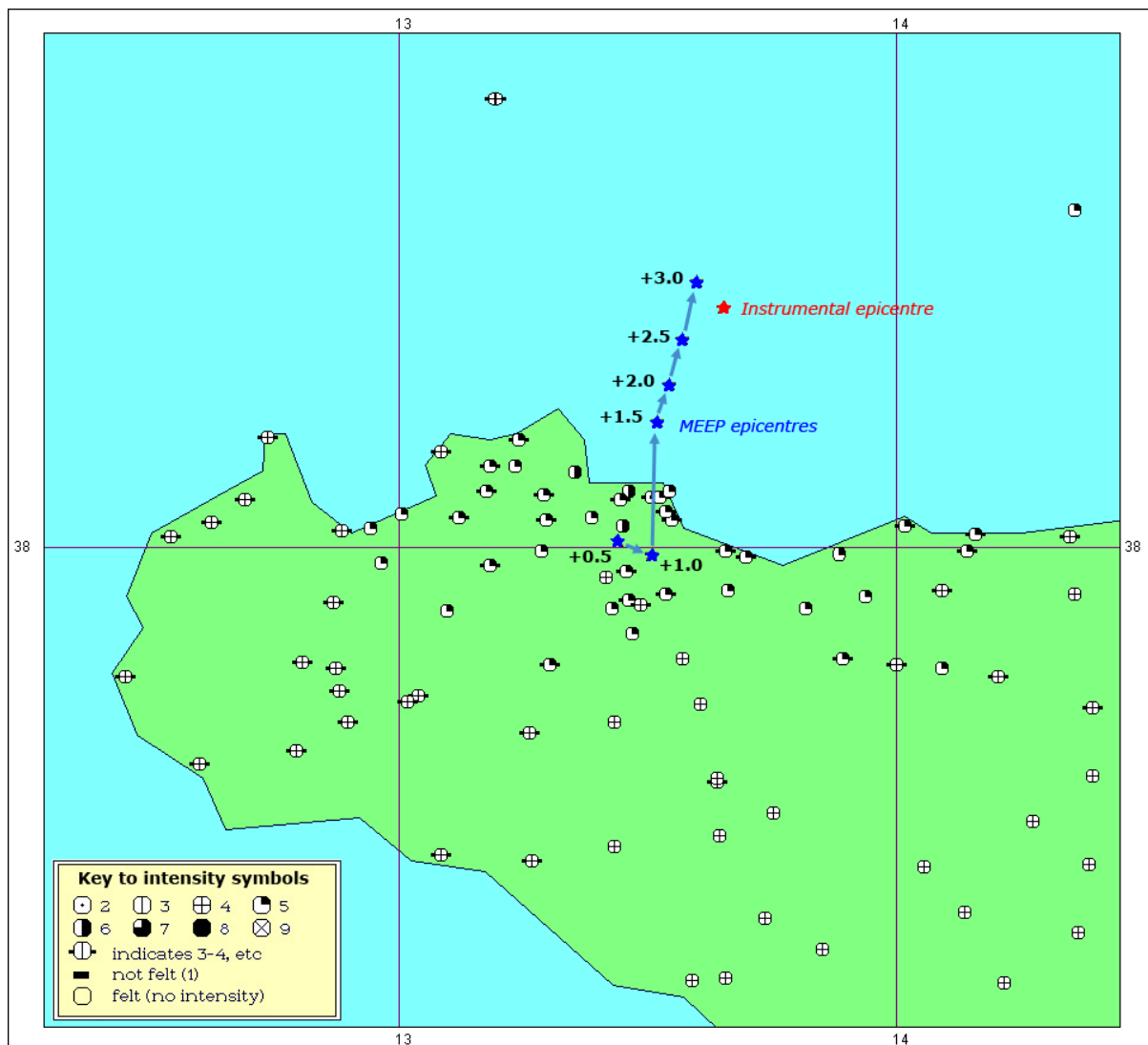


Figure 14 - Earthquake of 6 September 2002 (detail)

There is also a tendency for the algorithm to deal with difficult cases by depressing the depth, which decreases disparities in cases where intensities are irregular. Depths of 30 km (the maximum allowed) may indicate problem cases rather than genuinely base-of-the crust seismicity.

10 Conclusions

It was hoped initially that a method could be provided that would require no initial calibration. This can be realised for location, but not for magnitude. One possible solution that was looked at was to define macroseismic magnitude in some way that was entirely independent of instrumental magnitude. Thus one could simply define a parameter M_i using equation (5), such that any two earthquakes with the same felt area would have the same magnitude, anywhere in Europe. "Magnitude" would then be another way of writing felt area. The issue of relating macroseismic magnitude to instrumental magnitude would then simply be deferred to regional correlations between M_i and M_w . A similar approach could be taken by using MEEP with the same Consts.dat file for all Europe.

However, since ultimately users will want to have M_w magnitudes available, it is better to attempt to arrive at macroseismic M_w at the outset, which MEEP allows one to do. Decisions are needed, however, to arrive at suitable parameters for the Consts.dat file. Only one of

these is earthquake-specific: the notional increase allowed between observed I_{max} and that used in the optimised location. This will inevitably be different for, say, an onshore and an offshore event. It is because this is the parameter that most often needs adjusting that it appears first in the file. In the case of an offshore event, increasing this value allows for epicentral positions further offshore, with by implication, larger magnitudes.

The last three lines in the file, we believe, should never require changing. They are provided to allow experimentation. The alpha value is the only value besides the maximum intensity increase that affects the location, and for this, regional values are known from Karnik (1969).

This leaves the parameters that affect magnitude: regional Q , Frankel's (1994) constant C , and the scaling parameter K . A procedure has been described for determining these for any region using a set of IDP files for instrumentally-recorded events, in a straightforward way. The definition of region is left to the user; in the examples given in this report, average values were obtained for all Italy; these are open to improvement for regions within Italy.

There is also a possible issue in the case of large earthquakes affecting tectonically disparate regimes – attenuating more rapidly in one direction than another. One possible solution would be to prune out IDPs that are tectonically inconsistent (or use the centroid position as a fall-back).

The program is intended to be as robust as possible, and to be usable for events onshore and offshore, and also for earthquakes with very little data, even as little as a single IDP. That said, no system can give good results 100% of the time, and there can still be problems with anomalous or malformed data sets – for instance, cases where intensities are over-estimated, or more than one earthquake has been run together, or any other problem.

Such problem cases are usually flagged up with high uncertainties in the location, often with maximum depths as well; increasing the depth minimises the relative discrepancies in hypocentral distances, and so reduces the degree of misfit in cases of “odd” intensity distributions. Odd distributions, of course, may be due to the earthquake actually being relatively deep. Unrealistically high magnitudes can be a hint that the original IDPs might require reassessment.

Nevertheless, although the epicentral uncertainty tends to be a good measure of accuracy, there is no “uncertainty in the uncertainty”, and there may be cases where, for instance, very limited data sets provide a solution that converges well – but since it is dependent on very few points, which may be themselves inaccurate, the uncertainty may give a false impression that the results are accurate, whereas the real message is that the results are strongly convergent from the available data.

Thus the system needs to be used intelligently. Unusual results can be audited by looking at the `-sol.txt` file, which traces the route to both the epicentre and the magnitude.

These cautions aside, it is hoped that this package represents an adaptable and useful set of procedures for determining consistent earthquake parameters from macroseismic data, especially for historical earthquakes.

References

AMBRASEYS, N N. 1985. Intensity-attenuation and magnitude-intensity relationships for northwest European earthquakes. *Earthquake Engineering and Structural Dynamics*, Vol. 13, 733-778.

BAKUN, W H, and WENTWORTH, C M. 1997. Estimating earthquake location and magnitude from seismic intensity data. *Bulletin of the Seismological Society of America*, Vol. 87, 1502-1521.

-
- BIANCO, F, DEL PEZZO, E, CASTELLANO, M, IBÁÑEZ, M J, and DI LUCCIO, F. 2002. Separation of intrinsic and scattering seismic attenuation in the southern Apennine zone Italy. *Geophysical Journal International*, Vol. 150, 10-22.
- BLAKE, A. 1941. On the estimation of focal depth from macroseismic data. *Bulletin of the Seismological Society of America*, Vol. 31, 225-231.
- BURTON, P W, MCGONIGLE, R, NEILSON, G, and MUSSON, R M W. 1985. Macroseismic focal depth and intensity attenuation for British earthquakes. 91-110 in *Earthquake engineering in Britain*. (London: Telford.)
- CAVALLINI, F, and REBEZ, A. 1996. Representing earthquake intensity-magnitude relationship with a nonlinear function. *Bulletin of the Seismological Society of America*, Vol. 86, 73-78.
- CECIĆ, I, MUSSON, R M W, and STUCCHI, M. 1996. Do seismologists agree upon epicentre determination from macroseismic data? A survey of ESC Working Group ""Macroseismology"". *Annali di Geofisica*, Vol. 39, 1013-1027.
- DAVISON, C. 1899. *The Hereford earthquake of December 17, 1896*. (Birmingham: Cornish Brothers.)
- DAVISON, C. 1900. Scales of seismic intensity. *Philosophical Magazine*, Vol. 50, 44-53.
- DAVISON, C. 1927. The Hereford earthquake of 15th August, 1926. *Geological Magazine*, Vol. 64, 162-167.
- FRANKEL, A. 1994. Implications of felt area-magnitude relations for earthquake scaling and the average frequency of perceptible ground motion. *Bulletin of the Seismological Society of America*, Vol. 84, 462-465.
- GASPERINI, P, BERNARDINI, F, VALENSISE, G, and BOSCHI, E. 1999. Defining seismogenic sources from historical earthquakes felt reports. *Bulletin of the Seismological Society of America*, Vol. 89, 94-110.
- KARNÍK, V. 1969. *Seismicity of the European area*. (Dordrecht: Reidel Publishing Company.)
- KÖVESLIGETHY, R D. 1906. A makroszeizmikus rengések feldolgozása. *Mathematikai és Természettudományi Értesítő*, Vol. 24, 349-368.
- LEVRET, A, BACKE, J C, and CUSHING, M. 1994. Atlas of macroseismic maps for French earthquakes with their principal characteristics. *Natural Hazards*, Vol. 10, 19-46.
- MUSSON, R M W. 1998a. The Barrow-in-Furness earthquake of 15 February 1865: Liquefaction from a very small magnitude event. *Pure and Applied Geophysics*, Vol. 152, 733-745.
- MUSSON, R M W. 1998b. A self-parsing file format for earthquake catalogue and data files. *Seismological Research Letters*, Vol. 69, 248-250.
- MUSSON, R M W. 1998c. Intensity assignments from historical earthquake data: issues of certainty and quality. *Annali di Geofisica*, Vol. 41, 79-91.

-
- MUSSON, R M W. 1998d. Inference and assumption in historical seismology. *Surveys in Geophysics*, Vol. 19, 189-203.
- MUSSON, R M W, NEILSON, G, and BURTON, P W. 1984. Macroseismic reports on historical British earthquakes VIII: South Wales. *British Geological Survey Report*, 233.
- PAPAZACHOS, B C. 1992. Anisotropy radiation modelling of macroseismic intensities for estimation of attenuation structure of the upper crust in Greece. *Pure and Applied Geophysics*, Vol. 138, 445-469.
- PERUZZA, L. 1992. Procedure of macroseismic epicentre evaluation for seismic hazard purposes. *Proceedings of the XXIII General Assembly of the ESC, Prague*, Vol. 2, 434-437.
- PRINCIPIA MECHANICA LTD. 1982. *British earthquakes*. (Cambridge: PML.)
- SCHENKOVÁ, Z, SCHENK, V, KALOGERAS, I, PICHL, R, KOTTNAUER, P, PAPATSIMBA, C, and PANOPOULOU, G. 2007. Iseismal maps drawing by the kriging method. *Journal of Seismology*, Vol. 11, 345-353.
- SPONHEUER, W. 1960. Methoden zur Herdtiefenbestimmung in der Makroseismik. in *Freiberger Forschungshefte C88*. (Berlin: Akademie Verlag.)
- VENEZIANO, D. 1995. Uncertainty and expert opinion in geologic hazards of the Earth, *Engineers and Education: A symposium in Honor of Robert V Whitman*, 7-8 October 1994, Cambridge MA, Massachusetts Institute of Technology.
- ZSÏROS, T. 1996. Macroseismic focal depth and intensity attenuation in the Carpathian Region. *Acta Geodetica Geophysica Hungarica*, Vol. 31, 115-125.