A mixed automatic-manual seismic catalog for Central-Eastern Italy: analysis of homogeneity

Marco Cattaneo[1,*], Massimo Frapiccini[1], Chiara Ladina[1], Simone Marzorati[1], Giancarlo Monachesi[1]

[1] *Istituto Nazionale di Geofisica e Vulcanologia, Ancona, Italy*

ABSTRACT

A comparison between pickings and locations obtained by automatic and manual procedures in the analysis of the seismicity of Central-Eastern Italy is presented. In a first step we compared automatic and manual pickings, demonstrating that in many cases the adopted algorithm, after some tuning, is able to reproduce both the timing and the weight assignment of a human operator. The comparison of automatic and manual locations allowed to demonstrate that, when the automatic procedure is able to reach a solution stable from the statistical point of view, these locations are comparable with the manual ones within the estimated error limits. Once established these reliability criteria, we began to produce a mixed automatic-manual catalog: the events located by the automatic procedure with estimated errors below the selected thresholds (2 km in horizontal and 3 km in vertical) were directly introduced in the catalog, other events were revised by a human operator. In this way more than 64% of the events did not need human intervention, allowing to correctly manage also a period of increased seismicity, characterized by more than 4000 events per month: in total 121894 events were located with good accuracy in a time period of less than 7 years (August 2009 - April 2016). In a last step, a further control of the reliability of the whole procedure was performed, by manually analyzing all the events occurred in the last month of the analyzed period and classified as reliable by the automatic procedure: two expert seismologists interpreted these events, and the comparison demonstrated that the differences between the automatic and manual pickings and locations are slightly larger, but comparable with the differences between two human operators. As further checks, an analysis of the distribution of the depth estimates on the whole catalog demonstrated that data from the manual or the automatic part are nearly indistinguishable for the central, better monitored area; furthermore the automatic system demonstrated to be able to correctly locate also quarry blasts, with a reasonable estimate of the depth of these very critical events. Finally, a quick look at the geographical and depth distribution of the seismicity summarized in the catalog is presented; also in this case the main result is the good overlap of automatic and manual locations, at least for the well-monitored areas.

1. Introduction - Reasons for a mixed catalog

The routine work of analyzing the seismicity recorded by a seismic network by means of manual pickings performed by expert analysts can become heavy and time-consuming, particularly when dealing with dense networks and areas characterized by a rather high rate of seismicity. Also to overcome these difficulties, in the last years many algorithms of automatic pickings both for P and for S phases have been developed [e.g. Spallarossa et al. 2014, Ross et al. 2016] and relevant quotations). These techniques allow to obtain reliable automatic pickings, and thus automatic locations, for many events, but can fail in particular circumstances: low signal-to-noise ratio in some stations, superposition of closely located nearly simultaneous events, stations characterized by non-stationary noise producing many false pickings. In these cases, the intervention of a human operator can allow the correct location of these events. This mixed approach has been adopted for the routine analysis of the seismicity in Central-Eastern Italy; in this work we want to analyze the possible dis-homogeneity in the locations introduced by this technique. It is worth noting that a mixed approach (automatic first location and successive manual review) is a quite common procedure, and is adopted e.g. by the INGV national monitoring system [Mazza et al. 2011]; the originality of our approach is that, when the automatic analysis obtains a stable result, the manual revision is completely skipped.

A possible inconvenience of the automatic processing procedures is the mis-location inside the network of external events. This is particularly true for networks of limited extension, and dealing with P phases only: in these cases, sometimes the fake inner location could also be characterized by low error estimates. In our case, the extension of the network, and the systematic search of

S waves, limits the risk of such behavior. In our experience, the external events, when not correctly recognized, are located nevertheless outside the network, at the distance limit of the Pn waves from the edges of the network itself, and are usually associated with high estimated location errors, so that they are sent to the manual revision phase.

The seismic monitoring of Central-Eastern Italy has been gradually improved during the last 20 years. After the 1997 Colfiorito seismic sequence the INGV national seismic network (RSN) changed its configuration in the area, anticipating the process of re-design of the network geometry that interested the whole network starting on 2002 [Amato et al. 2006]. In the same years, the regional networks of Marche, Umbria and Abruzzo improved their instrumentation too, and, more important still, started a process of more efficient data sharing both between each other and with the RSN, under the coordination of INGV. Some results of this process of data integration are presented in De Luca et al. [2009]. Starting in 2009, the network geometry and the procedures of data managements changed again: the whole regional network of the Marches, and part of those of Umbria and Abruzzo, where merged in the RSN [D'Alema et al. 2011, Monachesi et al. 2013]; moreover, the beginning of the TABOO experiment led to the installation of a dense multi-parametric network in the Altotiberina area [Chiaraluce et al. 2014]. Obviously the continuous improvement of the network led to a progressive reduction of the detection threshold, so that the obtained catalog cannot be considered as completely homogeneous. The aim of this paper is just to analyze if the adoption of a mixed approach introduced a further, artificial level of dis-homogeneity.

At the beginning of 2016, about 100 stations in the area 42.4-44.2;11.0-14.0 are centralized in the INGV Office in Ancona and managed in an homogeneous way. Most of these data are shared between the acquisition systems of Ancona and the central monitoring system of Rome, where they are fully integrated in the RSN, allowing the real-time control of Italian seismicity [Mazza et al. 2011]. Figure 1 shows the position of the stations, superimposed to the map of recent seismicity (2009-2016) and to the main geographical features of the area. A description of the tectonic framework is beyond the scope of this paper [see e.g. Carannante et al. 2013, for a description of the same area].

Besides a real time control of the seismicity, adopting the same tools utilized by RSN in Rome, the data centralized in Ancona are subjected to a more detailed off-line analysis procedure.

In particular, the continuously registered waveforms are analyzed by means of a STA/LTA algorithm
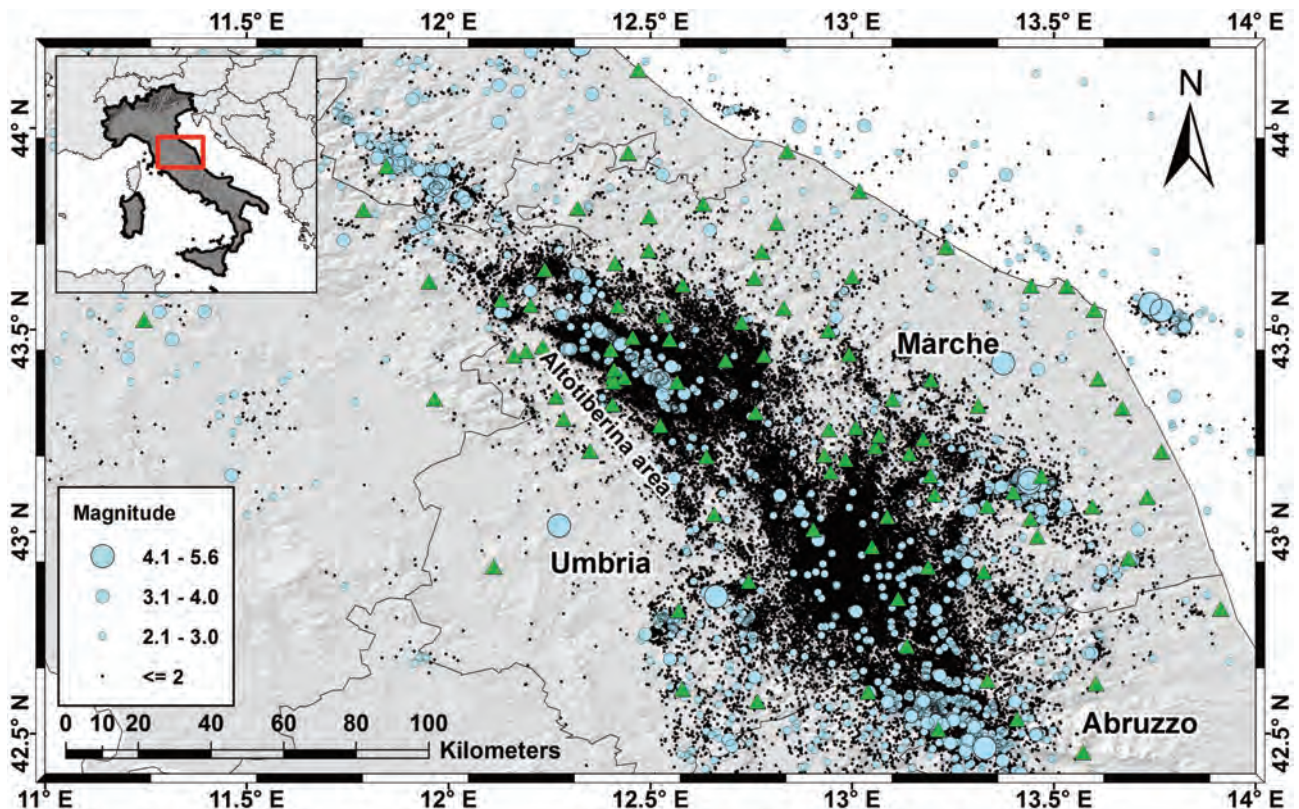

Figure 1. Map of the study area, reporting the located seismicity in the period 2009-2016 (size of the circles proportional to the magnitude), the stations position (green triangles) and the main geographical elements quoted in the text.

on the band-pass filtered signals, joined to a coincidence control based on the definition of sub-networks of key stations. More in detail, the signal is filtered between 2 and 15 Hz, STA and LTA windows are 1 and 100 s long respectively; the trigger is declared for ratio great then 5 with duration at least of 3 s; the duration is stopped when ratio is reduced below 2. Between the 100+ stations, just 41, characterized by low background noise and low occurrence of transient noise, are selected, and divided in 10 groups of closely located stations (less than about 60 km). The coincidence control is thus performed on these groups of stations on 10 s long windows overlapped by 50 %. The crossing of the coincidence threshold is controlled by weights assigned at each single component of seismic stations, based on their level of noise, to enhance the contribution of better stations: this allows to maintain low STA/LTA ratios (and thus high sensitivity) without increasing too much the probability of false events declaration due to coincidences happening by chance.

In this way, even detecting events down to magnitude around -1, the number of false detections is very low (usually smaller then 5%, see Figure 2 below). Obviously, magnitude -1 is not the catalog completeness limit, neither the average detection threshold of the network: Marzorati and Cattaneo [2016] demonstrated that the detection threshold of this network is quite strongly space- and time-dependent, spanning from values around -1 for the best monitored areas to values around 1.5 for the most marginal ones.

Up to May 2013, each detected earthquake was manually treated by human operators: all the steps of phase picking, earthquake location and magnitude computation were preformed by means of interactive programs. Nevertheless, in the meantime automatic procedures were developed, mainly devoted to furnish nearly-real time results for the civil protection purposes.

The continuous improvement of the detection capability of the seismic network, connected with a significant increase of the seismicity of the area during 2013, led to over-charge the manpower devoted to the seismic analysis. Figure 2 shows the number of detected triggers and of located events for each month in the period August 2009 - April 2016. It can be observed that, following a gradual increase of the mean number, with fluctuations linked to peculiar sequences, from April 2013 onwards we observed a strong increase of seismicity, reaching values between 6000 and 7000 events/month in the period December 2013 - April 2014. It is also evident the small difference between the number of detected triggers and located events, confirming that the mean number of false events is very low.

At the same time, the improvement of the automatic picking and location procedures led to obtain locations and magnitude estimates that were often comparable with the manual ones.

Obviously the automatic procedure cannot always reach the same results as a human expert; for this reason we decided to adopt a mixed approach, in which the human operator intervenes whenever the automatic procedure fails to reach a satisfying result.
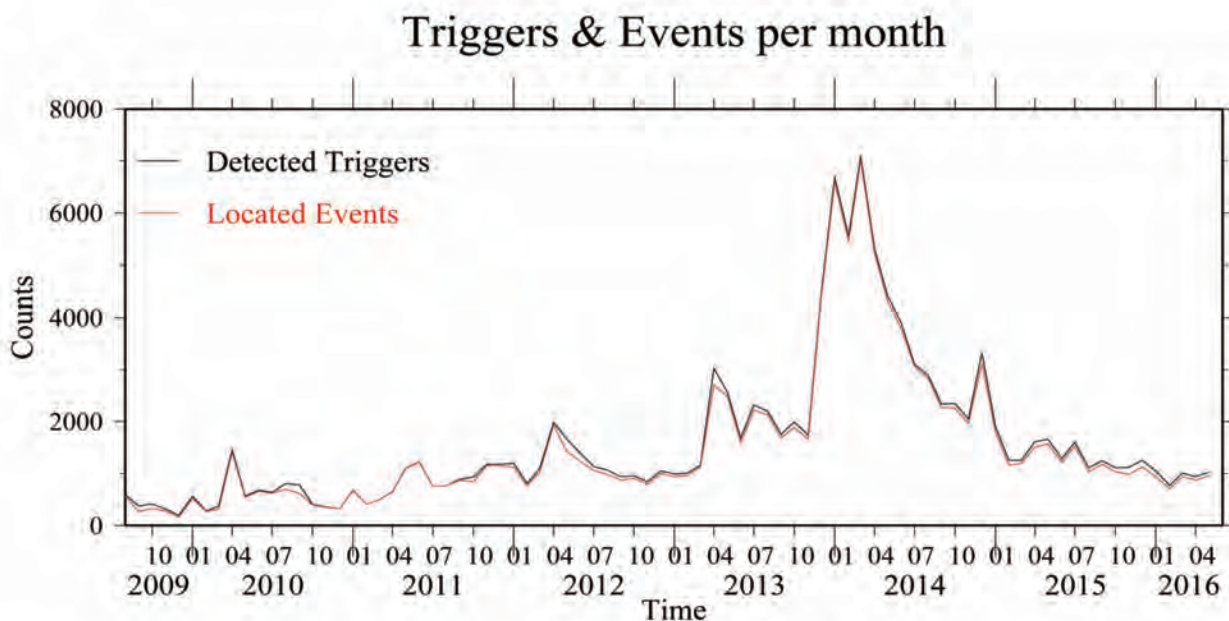

Figure 2. Number of triggers produced by the detection procedure, and of the located events, in the analyzed period (August 2009 - April 2016). Note the strong increase of seismicity from the end of 2013 to the beginning of 2015.

2. Comparison between manual and automatic results

Before attempting this approach, we had to be sure that the automatic procedure was able to mimic the output of an human operator. The core of the automatic system is the RSNI-Picker [Spallarossa et al. 2014, Scafidi et al. 2016]. This picker is based on the Akaike information criterion [AIC, Akaike 1974], but this technique is inserted in an iterative procedure, that allows to strongly reduce the occurrence of false pickings. Moreover, a user-calibrated procedure allows to assign to each picking a weight that should mimic the weight assignment of a human operator. The adopted weight code is based on the so-called "hypo-family" code [Lee and Lahr 1972]: the best quality data have assigned weight 0 (full weight), while increasing uncertainty is represented by increasing code (from 1 to 4), meaning decreasing weight (¾, ½, ¼, no weight in the original Hypo71 code). In our application, 0 code means uncertainty of the order of 0.01-0.03 s, 1 between 0.03 and 0.06 s, 2 between 0.06 and 0.2 s, 3 between 0.2 and 0.5 s, 4 larger than 0.5 s.

In order to compare the picking performances of the automatic system with those of a human operator, we analyzed 2 months' worth of data (January and February 2013, 1578 events) both with the automatic procedure and with the usual manual procedure. For the earthquakes for which the automatic procedure was able to furnish a stable location, we compared the P and S pickings with those obtained by a human operator; data are divided on the basis of the automatically assigned weight (Figures 3 a-b). It is quite evident that for P phases most of the automatic pickings fall in the 0 weight class (or, better, most of the other pickings were discarded by the procedure), and that most of them differ from the manual pickings for very small amounts. More in detail, 87.5% of the automatic P picking with 0 weight show differences below 0.03 s with respect to the manual ones (the nominal upper limit of the 0 weight class); taking into account that also the manual pickings are affected by some uncertainty, we could consider a reading potentially correct if the difference lies below 0.05 s, in which case the percentage is 93.9%.while the percentage grows to 96.6% for differences below 0.1 s, i.e. readings not clearly wrong.

The results of these comparisons between automatic and manual pickings and weights were used to re-calibrate the weight assignment procedure, adopted in the routine procedure starting in June 2013.

On the basis of these pickings, earthquakes are located by using the Hypoellipse code [Lahr 1999]. The adopted propagation model is based on the 1-D model obtained by De Luca et al. [2009], with station correc-
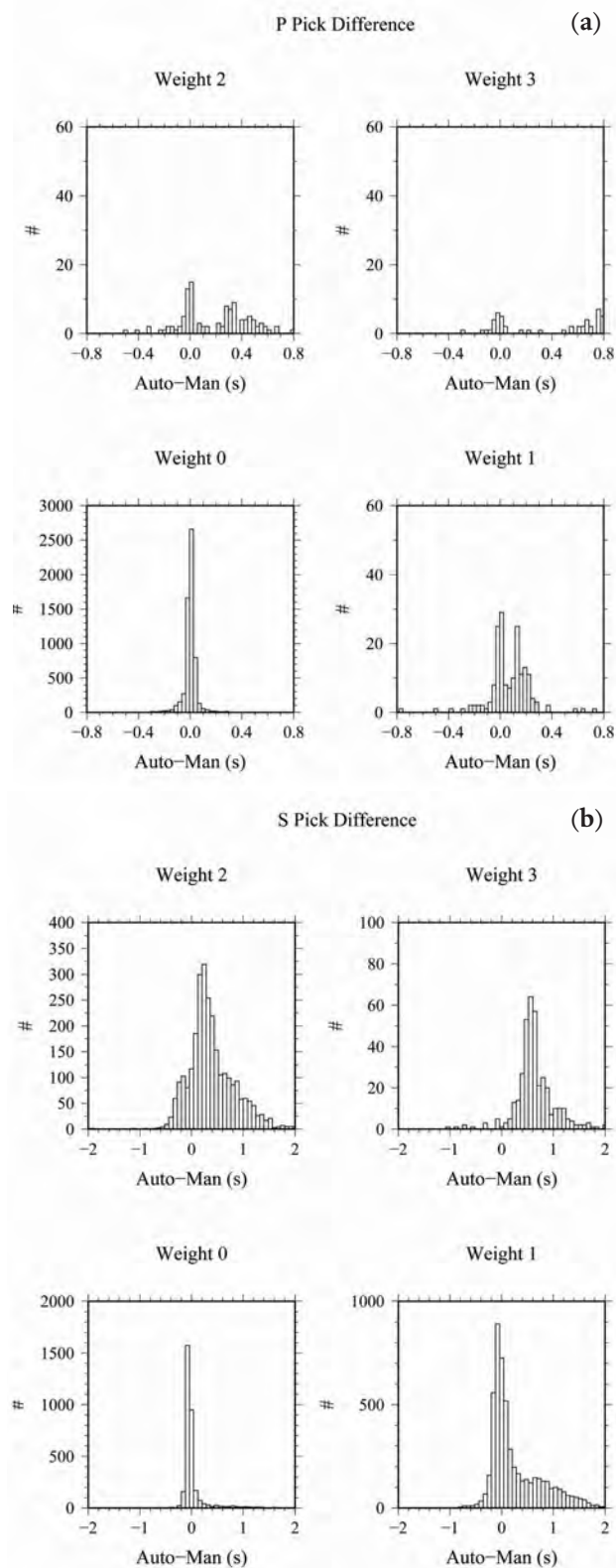

Figure 3. Differences between automatic and manual pickings in the period January - February 2013. a): P pickings; b): S pickings. Weights are related to Hypo71 quality classes.

tions re-computed for the stations installed after the publication of that paper. It is worth noting that the code allows to use the so-called Jeffreys weighting [Jeffreys 1973]: data producing residuals that fall in the non-Gaussian tail of the error distribution are progressively
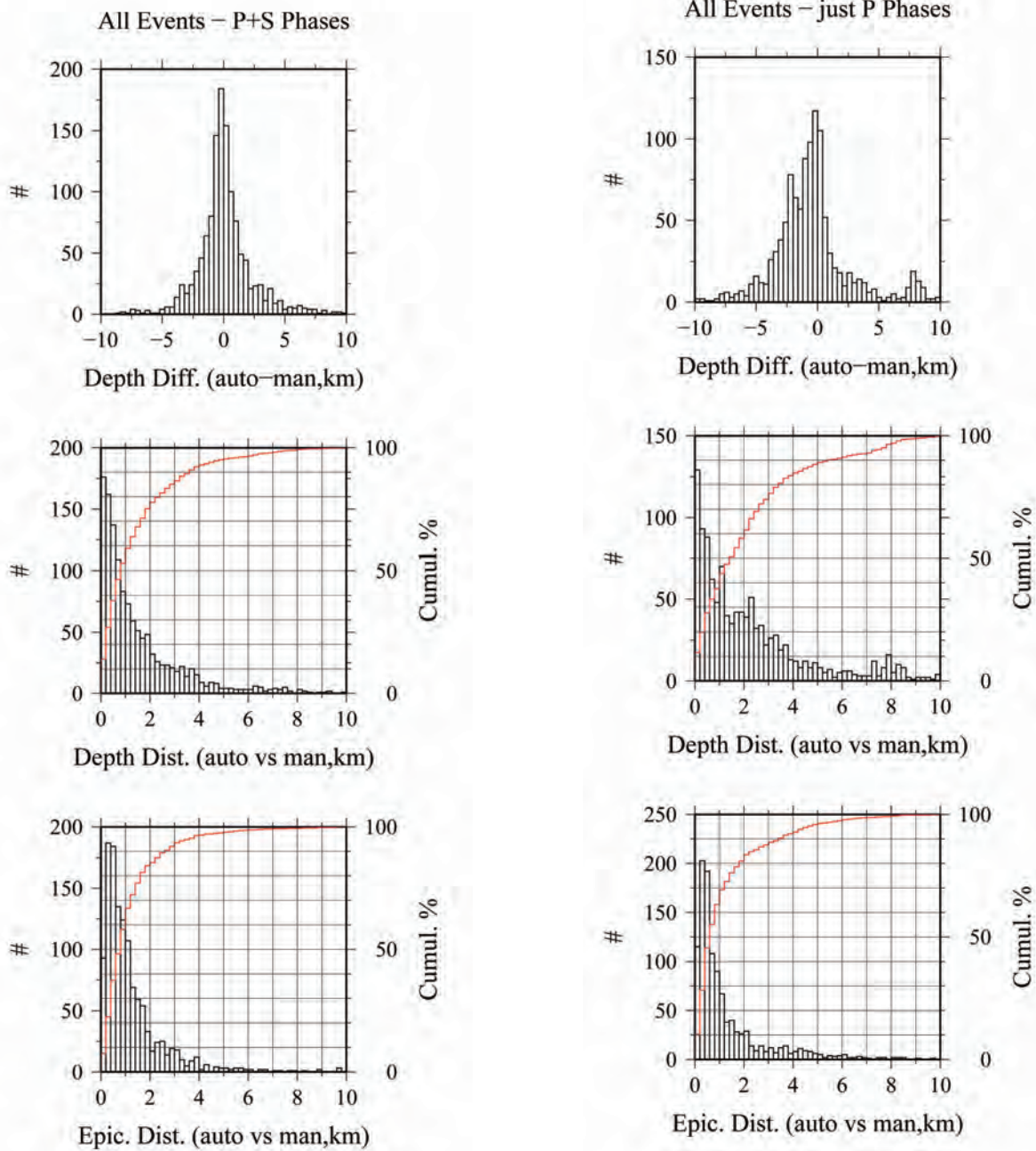
Figure 4. Comparison between automatic and manual locations, period January-February 2013, all data.

Figure 5. Comparison between automatic and manual locations, period January-February 2013, using just P pickings.

down-weighted during the iterations, and often completely cut off at the end. This is particularly important when dealing with automatic pickings, in which a small percentage of data could show errors completely outside the mean statistical distribution, and that could strongly influence the location procedure.

As already stated, locations are computed using both pickings from the automatic procedures and pickings obtained by an expert seismologist revision of waveforms. A comparison of the whole dataset of automatic location (1226 earthquakes, Figure 4) shows that more than 90% of the locations show horizontal differences

below 3 km and vertical differences below 4 km.

It can be noted that the good coherency of the automatically estimated depths with respect to the reference ones is mainly due to the reliability of the S-wave pickings. Indeed, if we compare the results obtained by using only the P-wave pickings (Figure 5), the distribution of horizontal errors is just slightly enlarged towards higher values, while the vertical error is strongly affected: in this case less than 80% of the locations show a vertical difference below 4 km.

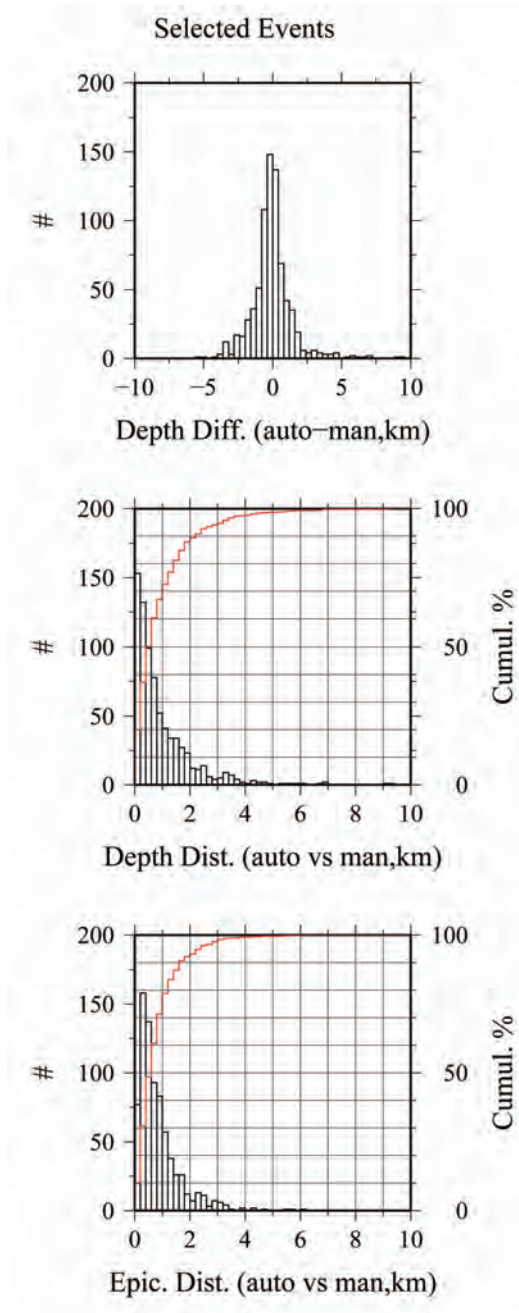The location program (hypoellipse) furnishes also an estimate of the location error; the complete de-

Figure 6. Comparison between automatic and manual locations, period January-February 2013, selection with ERH < 2 km and ERZ < 3 km.

scription is represented by the 68% joint confidence ellipsoid, but this is often synthesized in the horizontal error (erh), related to the maximum elongation of the projection of the ellipsoid on the surface, and the vertical error (erz), related the maximum vertical elongation of the ellipsoid [Lahr 1999]. Obviously these values cannot represent the "true" location error, but only a statistical estimate, based on the estimated standard error of arrival times, the weight code assigned to each arrival time, and, for each station, the partial derivatives of travel time with respect to latitude, longitude and depth.

We can thus select, from the automatic locations,

only the events showing estimated errors below a particular threshold; adopting 2 km as ERH threshold and 3 km as ERZ threshold, we obtain a sub-set of 759 more stable locations (48% of the analyzed triggers). The comparison of these locations with those obtained by the human revision is shown in Figure 6; it is quite evident that the difference distribution is more concentrated towards low values; more than 90% of such events show horizontal differences below 2 km, and depth differences below 3 km. Obviously the human expert location cannot be assumed as the "true" location; the most interesting result at this stage is that the automatic procedure is able to mimic the behavior of the human procedure, as regards the location accuracy, within the estimated error in most cases.

3. The mixed catalog and a-posteriori control

On the basis of the previously shown results, we decided to assume as reliable the automatic locations showing estimated horizontal error below 2 km and vertical error below 3 km. As an exception, events showing an automatic local magnitude above 2.5 were in any case revised by a human operator, also to pick polarities for focal mechanism computation. From June 2013 till April 2016 this choice allowed to promote as reliable (and thus insert directly in the catalog) 51783 automatic locations, out of a total of 80296 events. This means that more than 64% of the events were located with the requested accuracy by the automatic procedure. This rate is significantly higher than that observed in the first two months of analysis (48%). There can be two reasons for this result: on one side, the first analysis has been used to re-calibrate the weight assignment scheme of the procedure, so that it is likely that the subsequent application of this scheme improved the location capability of the system; on the other side, starting in August 2013, a rich seismic sequence interested the Altotiberina Valley area, where the seismic network is more dense, and thus the probability of a good quality location is higher.

In order to further verify the reliability of the whole automatic procedure, for the last month of analysis (April 2016) the data that were expected to be taken from the automatic procedure, on the basis of the selection criteria, were also analyzed by two human operators (the two seismologists that usually perform the control of the low-quality data). A first comparison is referred to the pickings: in Figure 7 the difference of the automatic pickings with respect to those of one of the two human operators is represented, both for P and S waves. It is quite evident that for P pickings (3581 readings) the distribution of differences is quite symmetrical with respect to
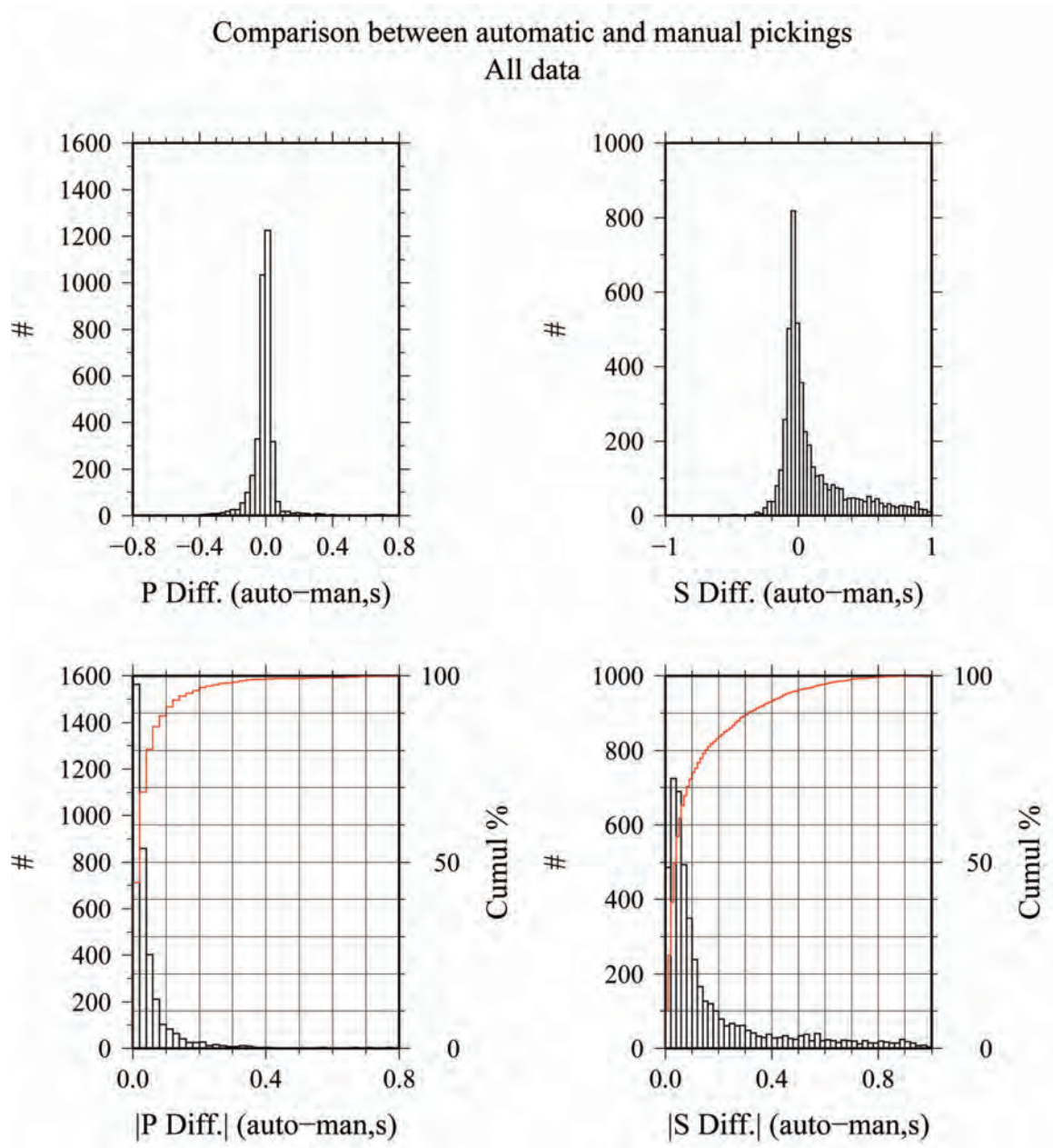
Figure 7. Comparison between automatic and manual pickings, period April 2016, all data.

the 0 line, and that about 90% of the pickings show differences below 0.1 s; on the other side, for S phases (4844 readings) the distribution has an asymmetrical tail (prevailing positive differences), and just about 83% of the pickings show differences below 0.2 s, confirming the higher criticality of the S picking. But, if we restrict the analysis to just the clearer readings (assigned weights, in the Hypo code, of 0 or 1 both from the human and the automatic operators, Figure 8), results are quite different: for P phases (3161 readings) the distribution is still narrower, and about 94% of the picking show differences below 0.1 s. For S phases (2771 readings) the asymmetry appears strongly reduced, and about 92% of the differences lie below 0.2 s.

If on the contrary we compare the pickings of the

two human operators, results are rather different: looking at the whole dataset (5695 P and 5689 s readings, Figure 9), the distribution of differences is rather symmetrical with respect to the 0 line, with a scatter higher for S phases, as expected: about 92% of P phases show differences below 0.1 s and 97% below 0.2s, while for S phases we have to enlarge the interval to 0.2 and 0.4s in order to obtain 93% and 98%, respectively. If we restrict the analysis to the best quality data (weights 0 and 1, 3969 P and 3379 S readings, Figure 10), the trend is the same observed for the automatic data, with a stronger narrowing of the distribution towards low values: more than 98% of the pickings show differences below 0.1 s for P phases and below 0.2 s for S phases.
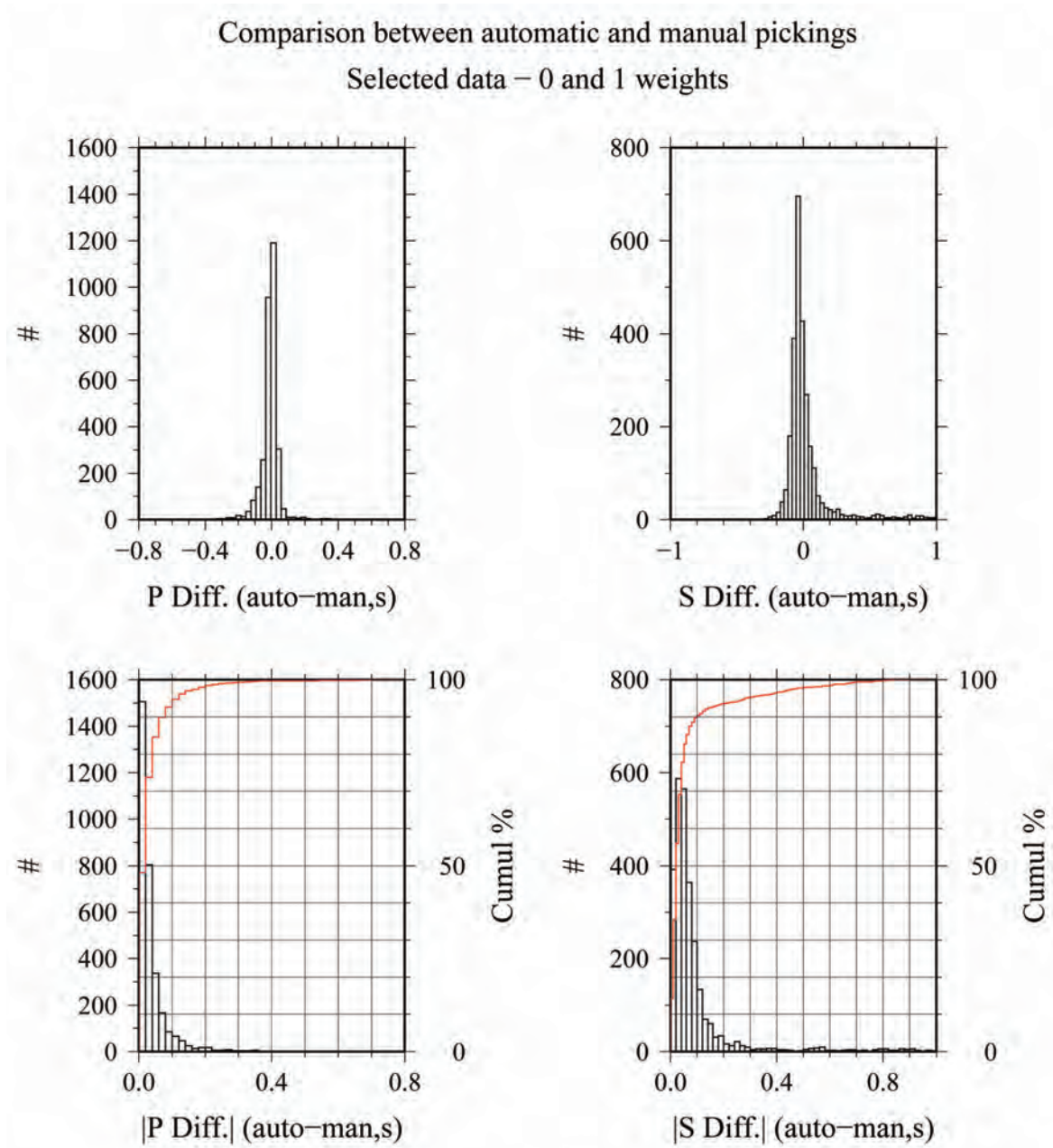
In summary, the automatic picker behaves quite dif-

Figure 8. Comparison between automatic and manual pickings, period April 2016, Hypo weights 0 and 1.

ferently from a human operator mainly for low-quality S readings, while for P readings and for high-quality S readings differences from a human operator are just slightly higher than differences between two expert human operators.

As the following step, we can see how these picking differences influence the location procedure: Figure 11a shows the distances between the automatic locations and those obtained by one of the human operators; for the depth differences, both the absolute values of the depth differences and the difference itself are shown, in order to verify the asymmetry of the distribution. It can be observed that about 94% of the locations show distances below 2 km in the epicentral position, and about 95% below 3 km in depth difference (2 and 3 km were

the adopted thresholds of estimated error for the selection of the automatic locations). If we restrict the analysis to the really best located events, mainly on the basis of the number of phases (> 12), of the maximum azimuthal gap of the recording stations ($< 200°$), of the rms of the residuals (< 0.3 s) and of the distance of the closest station (< 20 km), the results are just slightly better (Figure 11b): about 95% of events lie within 2 km as horizontal distance, and 97% within 3 km in depth. This means that just in a very few cases the location, even with good statistical parameters, is somehow intrinsically unstable, also beyond the estimated errors.

The same comparison can be made between the locations by the two human operators (Figure 12a for the whole dataset and 12b for the dataset selected on the ba-
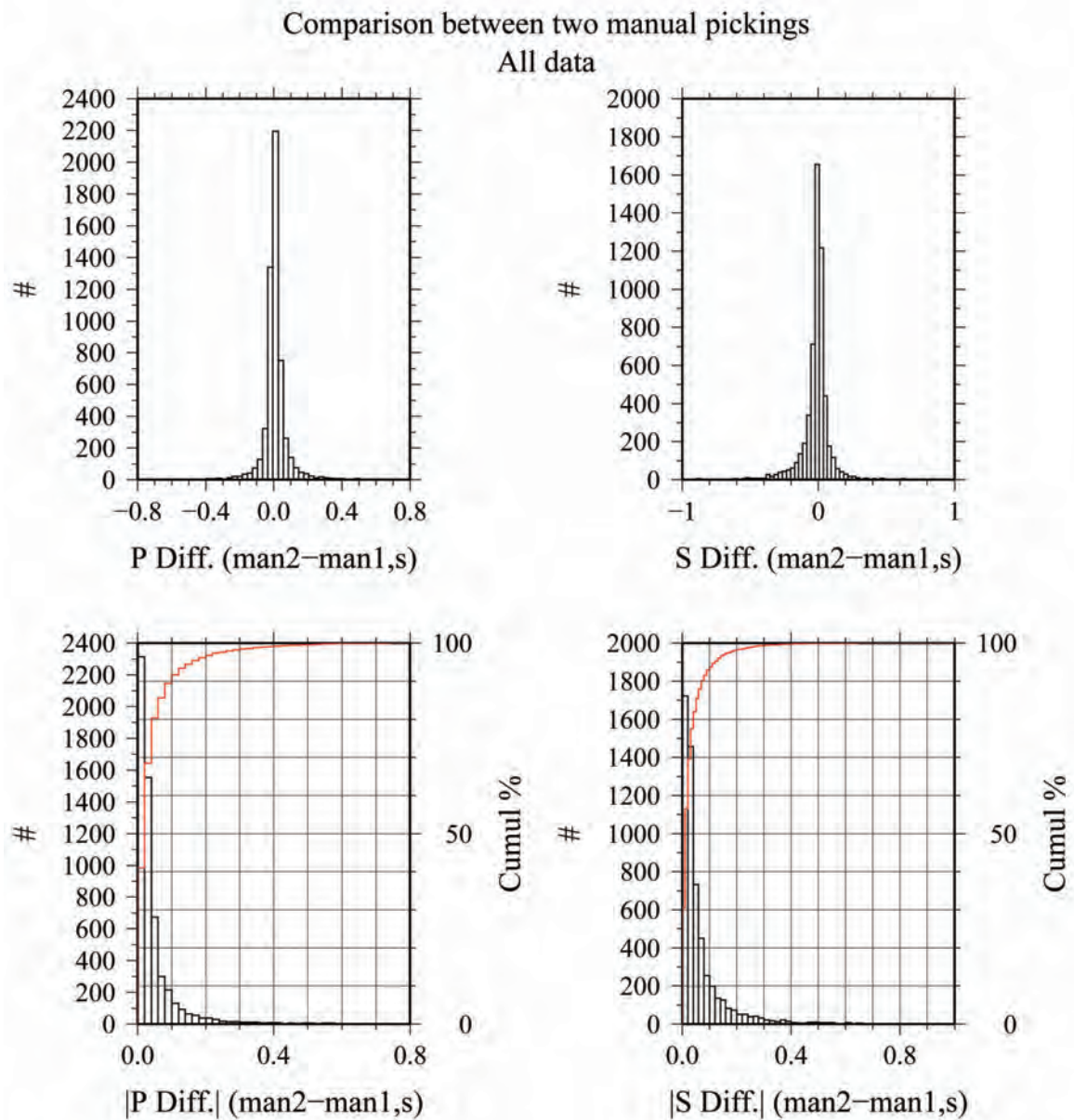
Figure 9. Comparison between two manual pickings, period April 2016, all data.

sis of the same quality criteria): in this cases very few events of the whole dataset show differences above the maximum estimated error, while in the restricted dataset only one event shows an horizontal distance slightly above 2 km and a depth distance above 4 km.

4. Stability of the depth estimate

As already stated, the most critical parameter of the location procedure is the focal depth; in particular, in areas where the station inter-distance is larger, the depth estimate becomes more critical, and more strictly dependent on the availability of good-quality S wave readings also for stations rather far from the epicenter. In these cases, the limits of the automatic picker can re-

duce the capability of a correct depth estimate. On the contrary, in areas where the station inter-distance is smaller, it is more probable that the automatic procedure can be able to compute depth estimates with good precision.

In order to verify this a-priori hypothesis, we performed a statistical analysis of the geographical depth distribution in our catalog. At first, we analyzed the whole catalog: we selected the more stable solution ($erh < 2$ km, $erz < 3$ km), we subdivided the analyzed area in a grid sized $0.05°$ both in latitude and in longitude, and we computed the depth interval containing at least 80% of the events located within each cell. Figure 13a shows a map view of the depth of the top and bottom of this layer. It can be observed that for a quite large subset of the analyzed area a rather narrow definition of the main
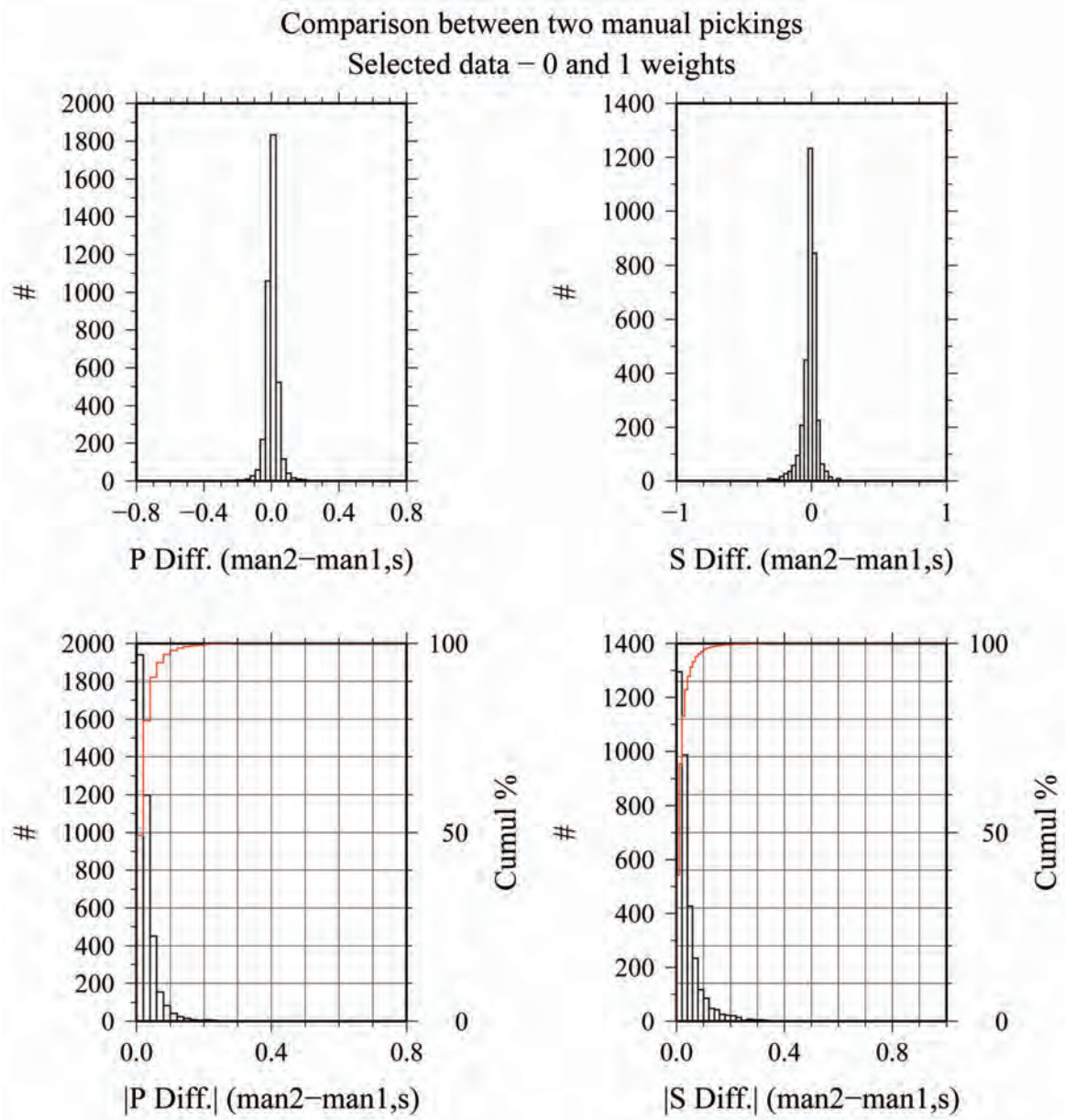
Figure 10. Comparison between two manual pickings, period April 2016, Hypo weights 0 and 1.

seismogenic layer is possible; in particular some specific trends, already described e.g. in de Luca et al. [2009] and in Carannante et al. [2013] can be recognized. A detailed discussion of this result is beyond the scope of this paper; here we just want to compare this result with the analogous obtained by using just the automatic locations present in the catalog (Figure 13b).

It is evident that the well-resolved area is more limited, as expected: only the region monitored by the denser part of the seismic network can produce stable automatic locations. Within this area, the main features of the depth distribution appear well reproduced by the automatic catalog with respect to the complete one. The main difference appears in the area around the point 43°18' - 12°20': in this area, the complete catalog shows a patch of maximum depth around 16-20 km,

while the automatic catalog does not show anything similar. In this case, a possible explanation can be found in the different time period covered by the two catalogs: the complete one spans a time interval of nearly 7 years (August 2009 - April 2016), the automatic one covers about 3 years only (June 2013- April 2016). The anomalously deep seismicity observed in this area (beneath the Tiber valley close to Umbertide) is rather episodic: it was recognized during a temporary experiment with a dense seismic network in this area during 2000-2001 [Piccinini et al. 2003], and has been again observed in the following years, but without any regularity. The complete catalog shows 47 such events, enough in order to influence the depth statistics; on the contrary the automatic catalog shows just 14 such events, not sufficient to influence the seismogenic layer estimate.
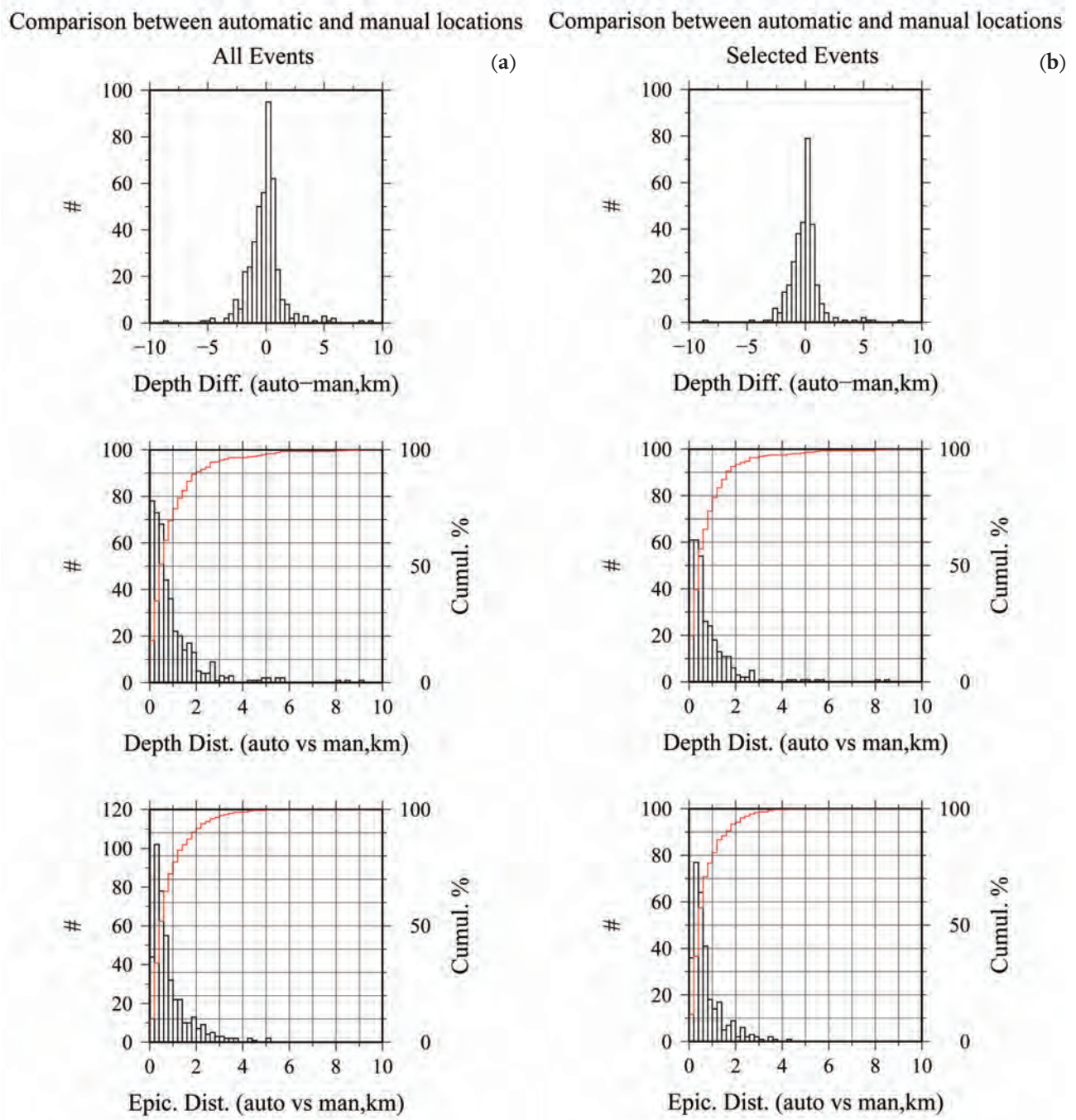
Figure 11. Comparison between automatic and manual locations, period April 2016; a): all data; b) selected dataset (see text for details).

As a general issue, the comparison of the depth distribution in different time spans could be quite critical, being not assured the stationarity of the phenomenon. In this case, dealing mainly with low-level, probably background seismicity, we are quite confident of the correctness of this assumption, and our results seem to confirm it.

5. Further control of automatic location quality: quarry blasts location

The above presented analysis is based on data relevant to the so-called "tectonic earthquakes". However, in this area different sources of non-tectonic earthquakes have been also recognized [Cattaneo et al. 2014]. The routine procedure is able to identify these events, on the basis of preliminary locations and waveforms similarity. The events can thus be subtracted from the seismic catalog, while their data are stored, so that it is possible to use them for a further quality control.

The use of quarry blasts for verifying the reliability of the whole location procedure (phase picking, propagation models, location programs, …) has been already presented [see e.g. Cattaneo et al. 1999, or more recently Viganò et al. 2015].
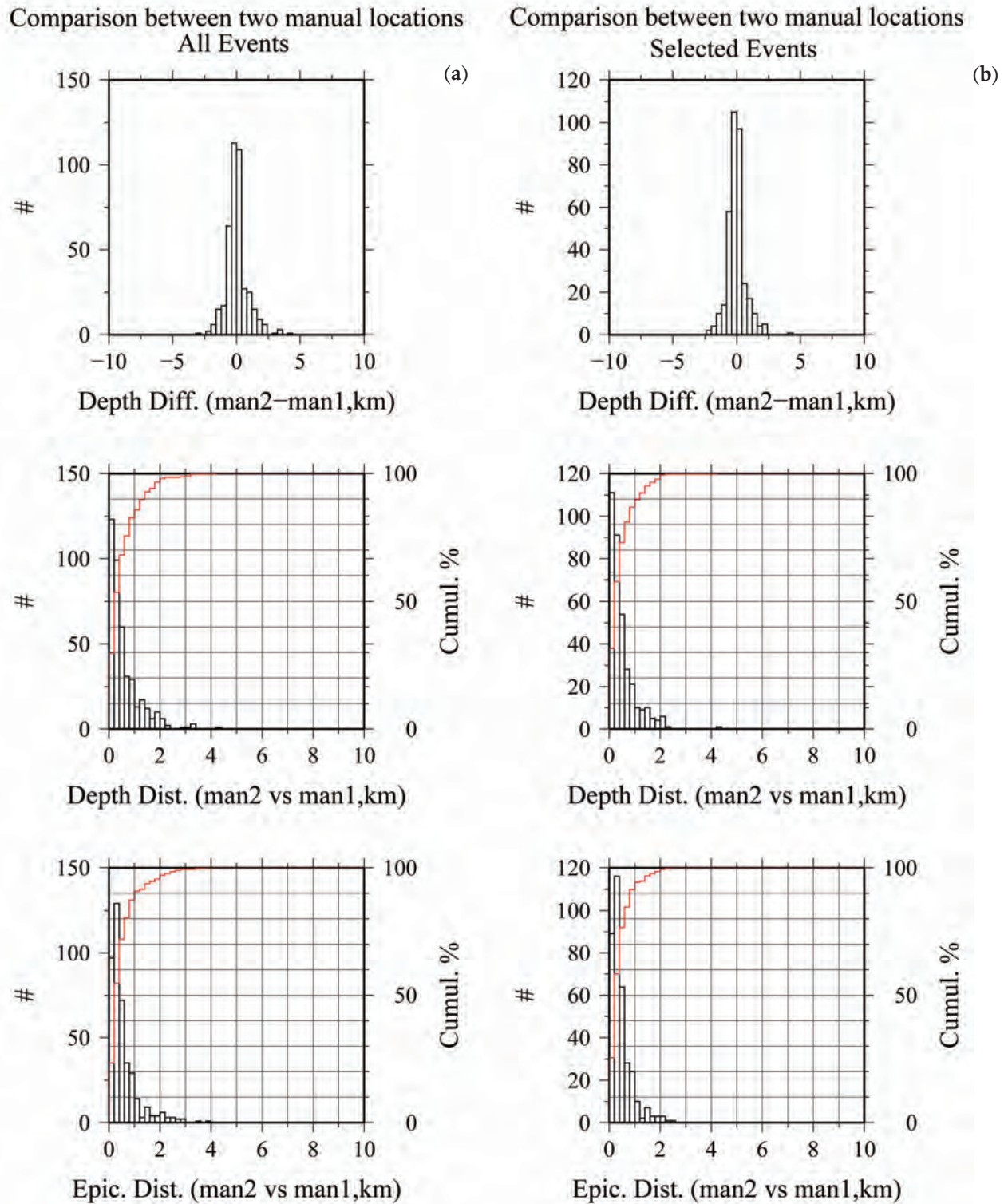
Figure 12. Comparison between two manual locations, period April 2016; a): all data; b) selected dataset (see text for details).

It can be assumed that the location of quarry blasts is more critical than that of an earthquake of the same equivalent magnitude: on one hand, a shot does not generate strong direct S waves, but what we record in the far field is mainly composed by converted waves; moreover it is more critical to verify the location of a very shallow event is than of a deeper one, both for geometrical reasons and on account of the higher level of

heterogeneity that we should expect in the shallowest layers, that for a deeper earthquake are crossed by the rays just close to the stations (and the possible inconsistency with respect to the propagation model is usually included in a station correction term), while for a quarry blast are crossed also near the source. In this sense, we can assume a quarry blast location as a "worst case scenario" for our location procedures. On the
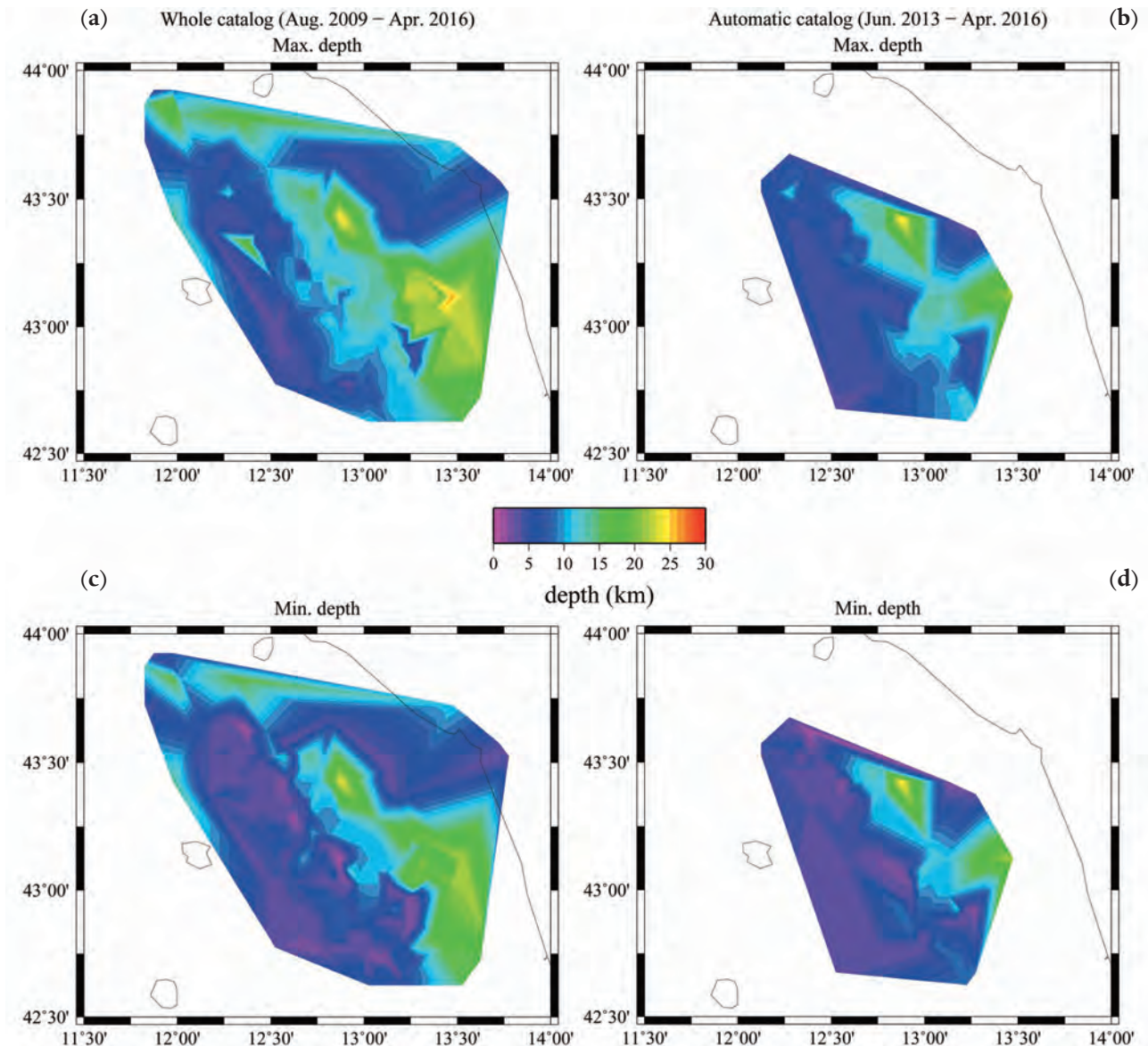
Figure 13. Map of top and bottom of the main seismogenic layer (containing 80% of the seismicity for each cell of a grid); a) all data; b) just automatically located data.

other side, for quarry blasts we know with good precision the location, so that they represent an interesting tool in order to verify the reliability of our estimates.

We thus selected 491 quarry blasts located by the automatic procedure in the period June 2013-April 2016 with the quality criteria already discussed in the previous paragraphs, and compared their locations with the location of the closest active quarry. The difference in horizontal location and in depth are presented in Figure 14a; it can be noticed that about 90% of the events show horizontal and vertical errors below the adopted error thresholds (2 km and 3 km, respectively). Still better results are obtained if we reduce the analysis to the quarries within the best monitored area (we just excluded quarries south of 43.00N and north of 43.55N), and we select the locations characterized by at least one station within 15 km; in this case (Figure 14b), about

95% of the locations lie within 2 km from the quarry, and 99% within 3 km. As regards depth, about 93% show depth estimates shallower than 3 km, and more than 99% shallower than 4 km.

Taking into account the already discussed criticality of the quarry blast locations, we can thus be rather confident that the automatic procedure is able to correctly locate even low-magnitude very shallow events, with a good accuracy on the horizontal coordinates and with the capability at least to restrict the depth estimate to the shallowest layers.

9. Preliminary analysis of the mixed catalog

In summary, the analysis procedure above presented allowed to locate, in the period August 2009-April 2016, 117430 tectonic earthquakes and 4464
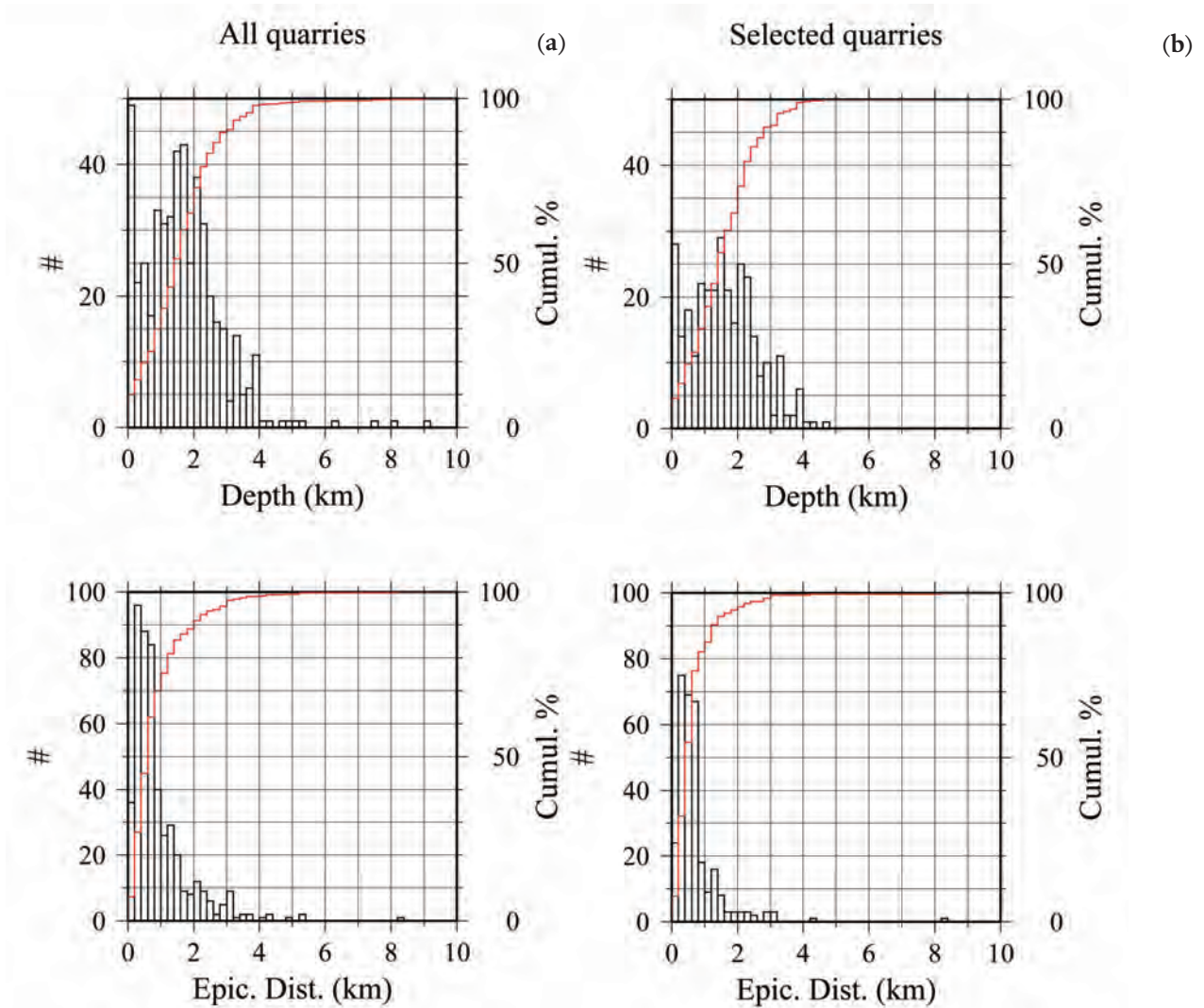
Figure 14. Statistics of the distances of automatically located quarry blasts and the barycenter of the closest quarry, and of the estimated quarry blast depth; a) all data; b) data from quarries located in the best resolved area.

non-tectonic earthquakes in the selected area. These locations are generated by the use of 1373646 P phases (824470 manual readings and 549176 automatic pickings) and 1668915 S phases (797326 manual readings and 871589 automatic pickings).

The excess of S phases with respect to P phases, mainly for the automatic readings, could appear surprising, but it must be taken into account that, for low magnitude events recorded at distances larger than the focal depth, often the signal to noise ratio is higher for S waves than for P waves, and thus the automatic picker finds an easier job for S phases.

This number of events can be considered rather high, taking into account the moderate seismicity of this part of Italy and the lack of events with magnitude higher than 4.9 in the analyzed area and time period, that could generate strong sequences (see e.g. the Mw 6.1 L'Aquila earthquake of April 2009, that generated at least 64000 aftershocks in few months, Valoroso et al., 2013). In our opinion, this catalog could represent an interesting

tool in order to analyze the so-called "background seismicity". The magnitude distribution of the events in the catalog is presented in Figure 15: 87.4% of the located events shows a local magnitude below 1.0, and 40.8% below 0.0. A quite similar distribution is obtained by analyzing just the automatic locations (top histogram): 92.2% below 1.0 and 47.1% below 0.0. It is thus demonstrated the capability of the automatic picker to correctly manage events of even very low magnitude, if they are located in an area with a dense enough network.

A quick look to the main characteristics of the seismicity depicted by this catalog can be furnished by some cross-sections perpendicular to the main Apennines trend (Figure 16). Both in the map view and in the cross-sections, the manual locations are reported as black dots and the automatic ones as red dots. These sections can be compared with those presented in De Luca et al. [2009] and in Carannante et al. [2013]; the main difference is that in this case a very simplified location model is adopted and automatic locations are
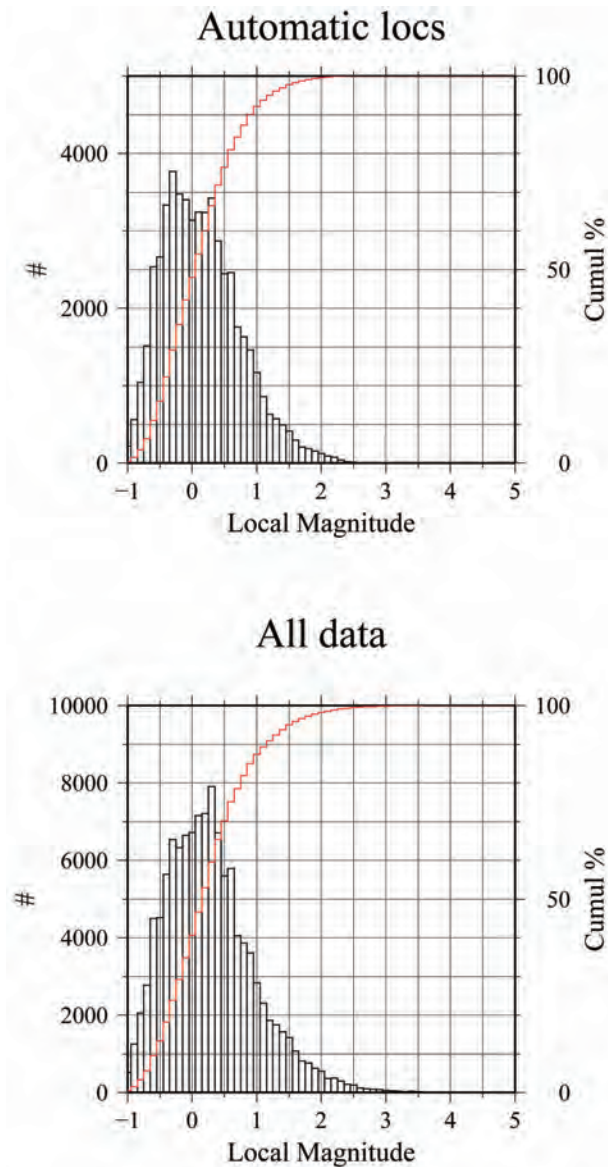
Figure 15. Magnitude distribution of the events in the whole catalog (August 2009 - April 2016). Top histogram: just automatic locations. Bottom histogram: all data.

also included, but on the other side the number of located events is significantly higher. The main purpose of this analysis at this stage is to check the coherency of the data of this mixed catalog with the main results of the previous, more accurate works.

Starting from NW, two main trends are recognizable as regards the depth distribution: the shallower part shows a mean deepening from SW to NE, while the deeper part shows an opposite trend.

The NE-deepening trend of the shallow part is very evident from section 5 to section 8; here the so-called Altotiberina fault represents a quite sharp limit to the distribution of seismicity in depth [Piccinini et al. 2003, Chiaraluce et al. 2007], with few exceptions: the most evident is the already quoted seismicity of the Umbertide area, recognizable in sections 6 and 7 at depths between 15 and 20 km. As anticipated, the

deeper part shows an opposite trend, going from about 20 km beneath the Adriatic coast down to about 60 km beneath the Apennines chain. This trend is very evident for sections 1-11, less evident but present till section 15, and hardly recognizable in the SE-most sections. This seismicity can be associated with the well-known sinking of the Adria lithosphere beneath the Apennines chain [Selvaggi and Amato 1992].

As already described in Carannante et al. [2013], the most complex depth distribution is related to the central sections (from 11 to 15). Here the NE-ward deepening of the shallow seismicity appears less evident, and a deeper seismogenic layer is present starting at the NE bound of this seismicity, concentrated between 15 and 25 km.

A detailed analysis of this seismicity is beyond the scope of this paper: here we just wanted to be confirmed in the fact that our routine locations show a very good coherence with the previous results, and that the automatic part of our catalog completely overlaps the manual part.

Indeed, the main trends of the depth distribution are well represented also by the automatic locations (red dots) mainly in the central part of the central sections, as expected. We already noticed that the automatic system performs at its best where the seismic network is denser. On the contrary, for the more external locations (SW limits of section 15-18 and sections 19-20), the contribution of automatic locations is strongly reduced. Another exception can be noticed for the seismicity at sea, very evident at the NE limit of sections 11-12, mainly related to the seismic sequence of July-August 2013 following a Mw 4.9 earthquake: in this case, given the criticality of locations, characterized by a large azimuthal gap of stations, we decided to manually interpret all the events.

7. Conclusions

This study was aimed to convince first of all ourselves that a mixed automatic-manual approach can produce an homogeneous catalog of events. Indeed, automatic procedures have already demonstrated their ability to achieve very accurate pickings and locations also for very large datasets [see e.g. Valoroso et al. 2013 or Scafidi et al. 2016], but usually in these cases no effort is made to maintain the whole completeness of the analyzed catalog: events not fulfilling the quality criteria are simply discarded from the analysis.

On the contrary, when dealing with the long-term monitoring of micro-seismicity in an area, every effort should be made to maintain the completeness and the
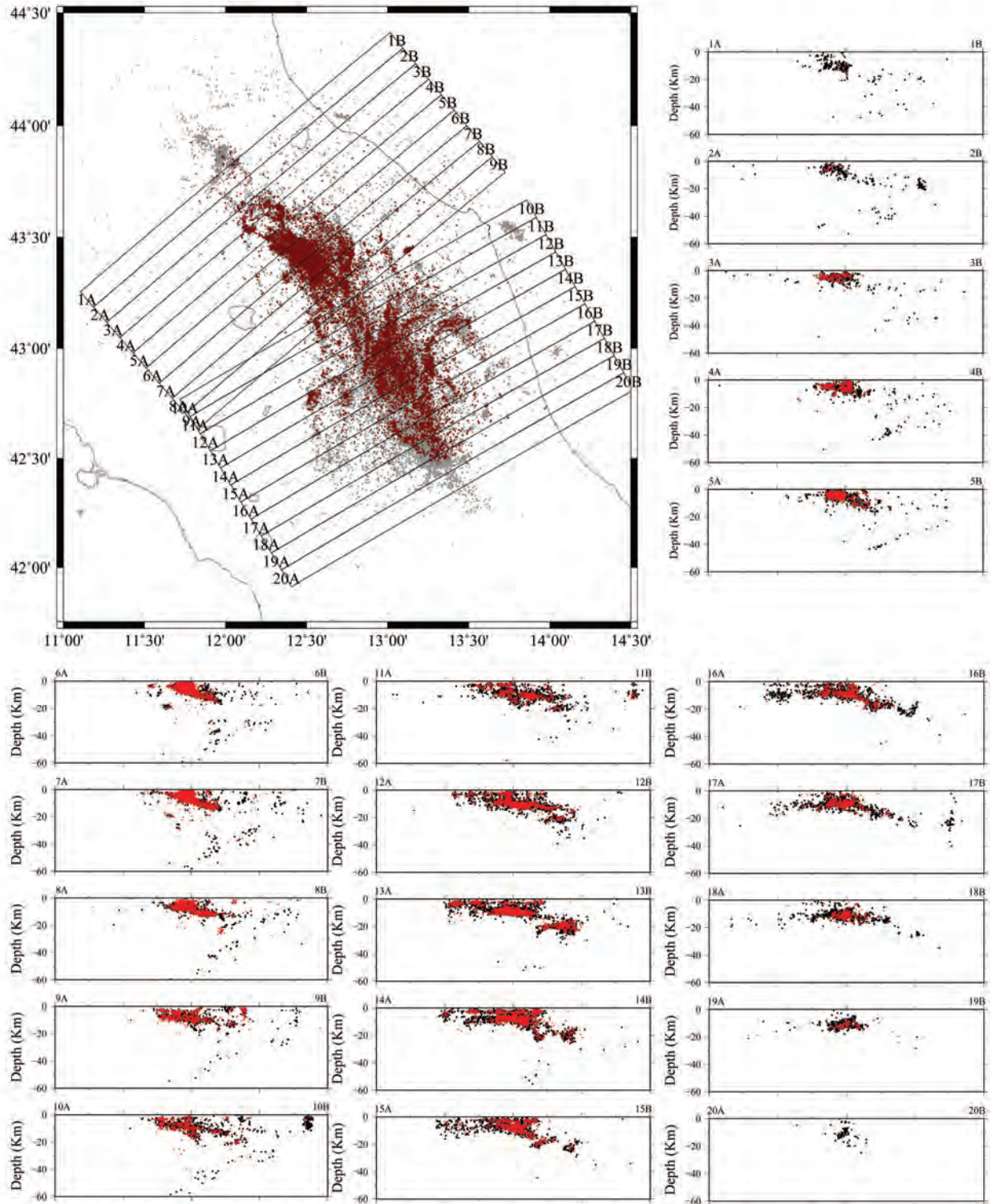
Figure 16. Map view and anti-Apennines cross-sections of the seismicity synthesized by the catalog (August 2009- April 2016). Black dots: manual locations; red dots: automatic locations.

homogeneity of the location quality for all the events fulfilling some pre-defined detection criteria. In our opinion, the most efficient tool to achieve these requirements is the mixed approach, in which human intervention is required when the automatic procedure is unable to reach the target quality thresholds..

In a first step, pickings from the adopted automatic procedure were compared with pickings derived from the manual revision activity; as a by-product, this comparison allowed to re-calibrate the weight assignment of the automatic procedure, in order to better mimic the human assignment, following the procedure described in Spallarossa et al [2014]. The comparison of the locations obtained from these pickings allowed to

verify the correctness of the statistical errors associated to these locations; in particular it was possible to verify that, when the automatic procedure is able to obtain a location stable from a statistical point of view, and thus with estimated error below reasonable thresholds (that we arbitrarily fixed in 2 km for the horizontal error, and 3 km for the vertical one), the difference with respect to the human location is nearly always smaller than the estimated error.

Based on these criteria, the mixed catalog started to be populated from June 2013, and is continuously updated. For a better characterization of the automatic pickings and locations, a further comparison has been carried out for data recorded in the last month of the analyzed time window (April 2016); the same data were picked by two different expert seismologists, in order to verify also the somehow intrinsic instability of our estimates. This analysis demonstrated that the automatic system produced differences slightly larger than, but comparable to the differences between the two human operators, thus confirming the reliability of the procedure.

Also the comparison of the geographical depth distribution, and the analysis of the automatically located quarry blasts, confirmed the overall stability of our locations.

This procedure is now routinely applied to the seismic monitoring of Central-Eastern Italy; in our opinion the produced mixed dataset (both pickings and locations) can be used as a whole for any seismological analysis in this area.

Data and sharing resources

Some Figures were created using the GMT software package [Wessel et al. 2013].

The Catalog of tectonic events in Central-Eastern Italy (August 2009 - April 2016) is available in electronic Supplements of the paper.

Record Format:

1-12: Event ID in the RSM database
14-92: event location, hypo71 format
14-30: origin time
32-39: latitude
41-49: longitude
51-56: depth (km)
60-63: local magnitude
4-66: number of phases with non-zero weight in the location
68-70: maximum azimuthal gap
71-75: distance of the 3-rd closest station
76-80: root mean square of residuals of the final location
82-85: erh, max. horizontal projection of the error ellipsoid as computed by Hypoellipse (km)
87-90: erz, max. vertical projection of the error elip-

soid as computed by Hypoellipse (km)
92: summary location quality, as assigned by Hypoellipse (A: best, D: worst)
94: flag for the adopted location; A: Automatic, M: Manual

References

Akaike H. (1974). A new look at the statistical model identification. IEEE Transactions on Automatic Control, 19(6), 716-723, doi: 10.1109/TAC.1974.1100705

Amato A., L. Badiali, M. Cattaneo, A. Delladio, F. Doumaz and F. Mele (2006). The real-time earthquake monitoring system in Italy. Geosciences-Revue du BRGM.

Carannante S., G. Monachesi, M. Cattaneo, A. Amato, C. Chiarabba (2013). Deep structure and tectonics of the northern-central Apennines as seen by regional-scale tomography and 3-D located earthquakes. J. Geophys. Res. 118(10), 5391-5403. DOI: 10.1002/jgrb.50371

Cattaneo M., P. Augliera, S. Parolai, D. Spallarossa (1999). Anomalously deep earthquakes in north-western Italy. J. Seismology, 3, 421-435.

Cattaneo M., E. Caffagni, S. Carannante, E. D'Alema, M. Frapiccini, C. Ladina, S. Marzorati and G. Monachesi (2014). A catalogue of non-tectonic earthquakes in central-eastern Italy. Ann. Geophys., 57/3, DOI: 10.4401/ag-6434.

Chiaraluce L., C. Chiarabba, C. Collettini and M. Cocco (2007). Architecture and mechanics of an active low-angle normal fault: Alto Tiberina Fault, northern Apennines, Italy. Journal of Geophysical Research, v112, B10310, doi:10.1029/2007JB005015.

Chiaraluce L., A. Amato, S. Carannante, V. Castelli, M. Cattaneo, M. Cocco, C. Collettini, E. D'Alema, R. Di Stefano, D. Latorre, S. Marzorati, F. Mirabella, G. Monachesi, D. Piccinini, A. Nardi, A. Piersanti, S. Stramondo and L. Valoroso (2014). The Alto Tiberina Near Fault Observatory (northern Apennines, Italy). Ann. Geophys, 57/3, doi:10.4401/ag-6426.

D'Alema E., M. Cattaneo, M. Frapiccini, S. Marzorati, G. Monachesi and M. Ferretti (2011). Rete Sismometrica Marchigiana e sua integrazione con la RSN e con la Rete AVT. Miscellanea INGV, 2011_10, ISSN 2039-6651.

De Luca G., M. Cattaneo, G. Monachesi and A, Amato (2009). Seismicity in the Umbria-Marche region from the integration of national and regional seismic networks. Tectonophysics, 476(1), 219-231. doi: 10.1016/j.tecto.2008.11.032.

Jeffreys H. (1973). On travel times in seismology. Collected Papers of Sir Harold Jeffreys on Geophysics and Other Sciences, H. Jeffreys (editor), Vol 2, Gordon and Brench Sci. Publ., London, 36-120.

Lahr, J.C. (1999, revised 2012). HYPOELLIPSE: a computer program for determining local earthquake hypocentral parameters, magnitude, and first-motion pattern. U.S. Geological Survey Open-File Report, 99-23,version 1.1, 119 and software, available at http://pubs.usgs.gov/of/1999/ofr-99-0023/.

Lee W.H. And J.C. Lahr (1972). HYPO71: a computer program for determining hypocenter, magnitude, and first motion pattern of local earthquakes. US Geological Survey Open-File Report 72-224.

Marzorati S. and M. Cattaneo (2016). Stima automatica della magnitudo minima rilevabile dalla rete sismica RESIICO. Quaderni di Geofisica, 2016-136.

Mazza S., A. Basili, A. Bono, V. Lauciani, C. Marcocci, A. Mandiello, F. Mele, S. Pintore, M. Quintiliani, L. Scognamiglio and G. Selvaggi (2011). L'acquisizione e l'analisi dei dati sismici al CNT. Miscellanea INGV, 2011_10, 108-112.

Monachesi G., M. Cattaneo, C. Ladina, S. Marzorati, E. D'Alema, M. Frapiccini, S. Carannante, M. Ferretti, M. Sebastianelli, A. Delladio and G. Selvaggi (2013). Esperienze di monitoraggio integrato: il caso della Rete sismometrica dell'Italia centro orientale e dei suoi servizi. Quaderni di Geofisica, 2013/106, ISSN: 1590-2595.

Piccinini D., M. Cattaneo, C. Chiarabba, L. Chiaraluce, M. Demartin, M. Di Bona, M. Moretti, G. Selvaggi, P. Augliera, D. Spallarossa (2003). A microseismic study in a low seismicity area in Italy: the Città di Castello 2000-2001 experiment. Ann. Geophysics, 46/6, DOI:10.4401/ag-3476.

Ross Z.E., M.C. White, F.L. Vernon and Y. Ben-Zion (2016). An Improved Algorithm for Real-Time S-Wave picking with Application to the (Augmented) ANZA Network in southern California. Bull. Seism. Soc. Am., DOI: 10.1785/0120150230.

Scafidi D., D. Spallarossa, C. Turino, G. Ferretti and A. Viganò (2016). Automatic P and S-Wave Local Earthquake Tomography: Testing Performance of the Automatic Phase-Picker Engine "RSNI-Picker". Bull. Seism. Soc. Am., doi: 10.1785/0120150084.

Selvaggi G and A. Amato (1992). Subcrustal Earthquakes in the Northern Apennines (Italy): Evidence for a Still Active Subduction? Geophys. Res. Letters, 19,21, 2127-2130.

Spallarossa D., G. Ferretti, D. Scafidi, C. Turino and M. Pasta (2014). Performance of the RSNI-Picker. Seism. Res. Lett., 85/6, doi:10.1785/0220130136.

Valoroso L., L. Chiaraluce, D. Piccinini, R. Di Stefano, D. Schaff and F. Waldhauser (2013): Radiography of a normal fault system by 64,000 high-precision earthquake locations: The 2009 L'Aquila (central Italy) case study. Journal of Geophysical Research, 118/3, 1156-1176, DOI: 10.1002/jgrb.50130.

Viganò A., D. Scafidi, G. Ranalli, S. Martin, B. Della Vedova and D. Spallarossa (2015). Earthquake relocations, crustal rheology, and active deformation in the central-eastern Alps (N Italy). Tectonophysiscs, 661, 81-98. http://dx.doi.org/19.1016/j.tecto.2015.08.017.

Wessel P., W.H.F. Smith, R. Scharroo, J. Luis and F. Wobbe (2013). Generic mapping tools: improved version released. EOS Trans. Am. Geophys. Union, 94:409-410. doi:10.1002/2013EO450001.

*Corresponding author: Marco Cattaneo,
Istituto Nazionale di Geofisica e Vulcanologia, Centro Nazionale Terremoti, Ancona, Italy;
email: marco.cattaneo@ingv.it.