Seismological Research Letters

A complete automatic procedure to compile reliable seismic catalogues and travel-time and strong motion parameters datasets

--Manuscript Draft--

Manuscript Number:	SRL-D-18-00257R2
Full Title:	A complete automatic procedure to compile reliable seismic catalogues and travel-time and strong motion parameters datasets
Article Type:	Article - Regular Section
Corresponding Author:	Davide Scafidi, Ph.D. Universita degli Studi di Genova Genoa, GE ITALY
Corresponding Author Secondary Information:	
Corresponding Author's Institution:	Universita degli Studi di Genova
Corresponding Author's Secondary Institution:	
First Author:	Davide Scafidi, Ph.D.
First Author Secondary Information:	
Order of Authors:	Davide Scafidi, Ph.D.
	Daniele Spallarossa, Prof.
	Gabriele Ferretti, Ph.D.
	Simone Barani, Ph.D.
	Barbara Castello
	Lucia Margheriti
Order of Authors Secondary Information:	
Manuscript Region of Origin:	ITALY
Suggested Reviewers:	
Opposed Reviewers:	

A complete automatic procedure to compile reliable seismic catalogues and travel-time and strong motion parameters datasets

Davide Scafidi(1), Daniele Spallarossa(1), Gabriele Ferretti(1), Simone Barani(1), Barbara Castello(2), Lucia Margheriti(2)

(1) University of Genoa, DISTAV (Dipartimento di Scienze della Terra dell'Ambiente e della Vita), Viale Benedetto XV 5, 16132 Genoa, Italy

(2) INGV, Istituto Nazionale di Geofisica e Vulcanologia, via di Vigna Murata 605, 00143 Rome, Italy

Keywords: Automatic seismic codes, automatic seismic processing, automatic seismic catalogue, automatic strong-motion parameters calculation

Abstract

The compilation of reliable and complete seismic catalogues represents a fundamental issue for most studies in seismology. Nowadays, the availability of an ever-increasing number of stations and, therefore, the huge amount of recordings to be processed and analyzed require a lot of effort in terms of man-hours. In the present work, we present a fully automatic procedure for compiling seismic catalogues starting from continuous recordings. The procedure relies on a multi-step algorithm that includes event detection tool, automatic P- and S-phase picker, hypocenter locator, and magnitude and strong motion parameter calculator. This automatic procedure is applied for compiling seismic catalogues for two real-world usage scenarios starting from the open-access waveform database provided by EIDA (European Integrated Data Archive). The first scenario concerns the monitoring of the seismicity of North-western Italy, the second one concerns the analysis of the data recorded during the first month of the 2016 sequence in Central Italy. The comparison between reference manually revised and automatic seismic catalogues points out negligible differences in terms of both P- and S-phase pickings, hypocentral coordinates, and local magnitude values, thus showing the overall reliability of the procedure. The ability of the proposed automatic procedure in detecting and locating very low-magnitude events is prominent to compile automatic catalogues characterized by a magnitude of completeness significantly lower than that of reference manual catalogues.

Introduction

Dense seismic networks made of an ever-increasing number of stations are becoming very common all around the World, providing a lot of experimental data for a variety of seismic applications such as seismic monitoring (Grigoli et al. 2017; Benz 2017; Moretti et al., 2016; Scafidi et al. 2015), high-resolution imaging of Earth interior (AlpArray 2014), induced micro-seismicity analyses (Benz et al. 2015; Ellsworth 2013), and so on. The raw data provided by these seismic networks (e.g. continuous recordings of seismic signals) are generally worldwide accessible through dedicated web portals. As an example, through the European Integrated Data Archive (EIDA, https://www.orfeus-eu.org/data/eida) or the Incorporated Research Institutions for Seismology Data Management Center (IRIS-DMC) it is possible to easily download continuous recordings provided by most seismic stations installed all around the world. The collected amount of raw data can often exceed many Gigabytes per day. In order to fully take the advantage of this huge amount of experimental data, seismologists have to analyze all recordings for extracting seismic events, identifying seismic phase arrival times, locating earthquakes, calculating magnitude and shaking parameters (i.e. Peak Ground Acceleration values), or deriving any other parameters useful for their researches. Such analyses are very time-consuming and, often, they can not be carried out without the aid of automatic procedures able to process raw data (e.g., seismic recordings) and extract from them the parameters of interest. Obviously, the employed automatic procedures have to guarantee both appropriate level of reliability, providing data comparable to those derived from manual analyses, and short computing time. For example, a reliable automatic picker is an algorithm capable of providing, in a short time, P- and S-phase arrival times similar to those derived from a visual inspection made by an expert seismologist.

The automatic picker presented by Spallarossa et al., (2014) and Scafidi et al. (2018), called "RSNI-Picker$_2$", has been proven to be able to process huge amount of data in a short time, to detect reliable P- and S-phase arrival times, and to accurately locate events (Scafidi et al., 2016). Starting from this picker engine, in the present study, we present a completely automatic procedure that, starting from a database of raw seismic waveforms (i.e., downloaded from EIDA), is able to recognize and extract time windows containing potential earthquakes (or explosions), to detect P- and S-phases arrival times, to locate earthquakes, and to calculate magnitude and some strong-motion parameters (such as Peak Ground Acceleration, PGA, Peak Ground Velocity, PGV, Peak Ground Displacement, PGD, Housner Intensities, IH, and Spectral Accelerations, SA). The procedure produces, in a completely automatic way, a seismic catalogue. The automatically derived parameters are saved in textual files (ASCII format) and stored in a PostgreSQL database too.

The performances of the automatic procedure proposed in the present paper is evaluated comparing the outputs (e.g., P- and S-phase picks, hypocenter coordinates, magnitude values) to those published in reference catalogues compiled through the accurate manual revision done by INGV (Istituto Nazionale di Geofisica e Vulcanologia) seismologists. In particular, we tested the procedure considering two different scenarios. In the first test, the procedure is applied to North-western Italy in order to compile a regional seismic catalogue from January to July, 2017. In the second test, the continuous recordings provided by a very dense network installed after the M_W 6.0 "Amatrice" earthquake (occurred on 24 August, 2016 in Central Italy) are processed in order to study the evolution of the sequence during the first month.

Automatic procedure description

The proposed automatic procedure for compiling seismic catalogues works on a dataset of raw seismic waveforms (e.g., daily continuous seismic recordings). It consists of 4 main steps (Figure 1) that are (i) detection of single station triggers, (ii) trigger association and event detection, (iii) extraction of time windows including potential earthquake (or explosion), (iv) picking of P- and S-arrival times, event location, and strong motion parameter calculation.

The software package, called CASP (Complete Automatic Seismic Processor), consists of four modules written in standard “C” programming language to optimize portability across a range of computer architectures and calculation speed.

Besides continuous waveform dataset, the procedure needs as input the meta-data of the seismic stations providing the waveforms (e.g., geographical position, instrument response parameters). Generally, the meta-data may be easily downloaded from the same web portal used for retrieving the waveforms. They can be linked with waveform itself in the full-SEED formats, or they can be taken alone in various standard formats (station XML, response file, dataless, poles and zeroes file). All the cited metadata formats are supported by the presented procedure.

The output of CASP is a complete dataset where P- and S-phase arrival times and location parameters (hypocentral coordinates and magnitude) with associated uncertainties, and strong motion data (PGA, PGV, IH, SA) are listed for every earthquake.

The CASP procedure is equipped by a “starting” module that automatically downloads from a selected web portal (e.g., EIDA database) all the continuous recordings of a list of seismic stations and creates a dataset where the waveforms are organized in files, one for each station. In the default setting, the waveforms to retrieve are expected to be in Mini-SEED format, according to conventional standard (e.g., EIDA or IRIS data format). The “starting” module can be skipped if a dataset of Mini-SEED files is already available.

The four main steps of the automatic procedure are summarized below.

Step 1 (“RSNI-Trigger” module): for each station separately, the vertical component of the continuous recording is processed in order to perform trigger calculation. In details, the signal is band-pass filtered and the STA/LTA ratio (i.e., the ratio of the short-term average over the long-term average) is calculated. When the ratio exceeds a threshold value, a single station trigger is declared. The final trigger time is then determined as the minimum of the Akaike Information Criterion (AIC) function (Akaike 1974) computed within a signal window around the trigger identified by STA/LTA analysis. The use of the AIC-based algorithm allows to overcome typical STA/LTA analysis errors. For example, in case of a low-magnitude event recorded by a seismic station very close to the hypocenter, the STA/LTA tool can place the trigger time in correspondence of the S-wave arrival time. The subsequent application of the AIC algorithm allows to move the trigger time to the right position, in correspondence of the P onset. The output of the “RSNI-Trigger” module is a list of trigger times for each station. This module is driven by a set of parameters that has to be empirically calibrated for each station as a function of background noise at station, sensor type, and so on. These parameters are (i) the corner frequencies of the band-pass filter, (ii) the short- and long-term average constants, (iii) the STA/LTA ratio threshold, (iv) the search window length for the AIC function calculation, and (v) the minimum time-interval permitted between two consecutive triggers. Of note, it is convenient to set the “RSNI-Trigger” module in order to be very sensitive and to not loose any potential seismic trigger because the following modules of the CASP procedure are able to identify triggers not related to seismic events.

Step 2 (“RSNI-Detect” module): the lists of the single station trigger times are read and analyzed in order to identify the signal windows including potential events (e.g., earthquakes and explosions). The module “RSNI-Detect” has been engineered to be easily and effectively portable to different application cases, from small local seismic networks to large and dense regional ones. The event detection algorithm may work considering a set of sub-networks, which may overlap, designed to maximize the detection capability. The usage of sub-networks also optimizes the detection of earthquakes with similar origin time but occurred in different areas. The “RSNI-Detect” module requires an input configuration file containing the number of sub-networks, and the list of stations belonging to each sub-network. The detector algorithm is based on a coincident system (that works independently on each sub-network) defining the number of data channels (e.g., vertical component) which must have triggered coincidentally within a coincidence window in order to declare the start of a potential event. The module makes a further check on events detected considering each sub-network in order to avoid the multiple identification of a same event. This check is based on the comparison among earthquake locations computed using the trigger times as P-phase picks by the NonLinLoc (Lomax et al. 2000, 2009) software (with a configuration optimized for the study area).

As in the previous step, the parameters controlling the “RSNI-Detect” module (i.e., sub-network geometries, number of stations which must be triggering coincidentally, coincidence window length) can be set to have a very low detection threshold. This way, it is possible to recognize very low-magnitude earthquakes and, generally, many false events too. All false events will be effectively recognized and discarded during the last step of the automatic procedure (RSNI-Picker$_2$ module).

Step 3 (“RSNI-Extract” module): the seismograms of the recognized potential events are extracted from the dataset collecting all the continuous recordings and converted in SAC format. Such extraction is controlled by some user-definable parameters that are (i) the pre-event time (i.e., number of seconds before the event origin time as derived from the preliminary location obtained in the previous step), (ii) the window length to extract, and (iii) the extraction radius (the maximum epicenter-station distance, derived from the same preliminary location, allowed for the data extraction of a station).

Step 4 (“RSNI-Picker$_2$” module): the extracted seismograms relevant to each recognized event are processed through the “RSNI-Picker$_2$” software (Scafidi et al. 2018) in order to pick P- and S-phase arrival times, locate the earthquake, compute magnitude and strong motion parameters (e.g., PGA, PGV, PGD, IH, SA).

As described in detail in Spallarossa *et al.* (2014) and Scafidi et al. (2018), the “RSNI-Picker$_2$” is an iterative procedure for automatically identifying phase arrival times through AIC algorithm (Akaike 1974) where pick identification is checked and refined based on locations computed at each iteration. Such iterative structure allows an optimized recognition of false or imprecise picks. The picking algorithm implements the NonLinLoc probability-based locator code (Lomax et al. 2000) for locating events by using a 1D or 3D velocity model. Local magnitude (M_L; Richter 1935, 1958) and strong motion parameters calculations are performed by a multi-thread algorithm that allows to overcome problems due to waveform saturation and/or distortion. “RSNI-Picker$_2$” provides a quality estimate for each calculated parameter such as quality weight of automatic picks, standard quality parameters of locations and uncertainty of the M_L values. The “RSNI-Picker$_2$” is also equipped with a tool for identifying out-of-network events such as teleseisms or regional earthquakes through an appropriate spectral analysis. Moreover, as shown by Scafidi et al. (2018) that deeply tested the performances of the “RSNI-Picker$_2$”, all false events are recognized by this algorithm since they are not locatable or located with a high level of uncertainty. The RSNI-Picker$_2$ module is controlled by several parameters driving, for example, the picker engine (e.g., filtering, pick validation), and the location and magnitude computation tool (e.g., velocity model, attenuation coefficients). Such parameters have to be carefully and empirically defined by the user as a function of the kind of application (e.g. seismicity of the study area, network geometry, seismic signal quality).

All processed data are stored in a PostgreSQL database from which they can be easily extracted through SQL queries according to own needs.

The automatic procedure described above is summarized in the flowchart of figure 1.

Comparison between manual and automatic catalogues

In order to assess the reliability of the outcomes of the CASP procedure, we have compared automatically derived catalogues with reference ones. The last have been provided by INGV and they have been compiled through a careful manual revision made by expert seismologists. In the present paper, we present two different application of the automatic procedure considering two typical scenarios of use in seismology. In the first, the CASP procedure is tested analyzing the seismicity of a wide area covered by a regional network, while, in the second, it is applied to study the evolution of a seismic sequence in a small area and recorded by a dense network. The comparison between reference and automatic catalogues has been made considering differences among P- and S-phase picks (e.g., number of picked phases and arrival times), earthquake locations (e.g., epicentral coordinates and depth), and local magnitude values.

1) Regional seismicity case study: seismic catalogue of North-western Italy and surrounding areas

The first case study is aimed at testing the capability of the automatic procedure to compile a complete and reliable seismic catalogue of a wide region. The seismically active area of the North-western Italy (Figure 2) has been considered for this test, considering a period of 7 months starting from January 1, 2017 to July 31, 2017. The area is 450 x 450 km wide (box in figure 2) and its seismicity is monitored through 302 permanent seismic stations (triangles in figure 2). The daily continuous recordings of all the seismic stations have been automatically downloaded from the EIDA database (contributed from the Rete Sismica IV (Italian Seismic Network, INGV 2006) and Rete Sismica GU (Regional Seismic Network of North-western Italy, University of Genoa, 1967), using the “starting” module, and analyzed through the CASP procedure in order to compile the seismic catalogue.

In order to optimize the event detection process (step 1 and 2), 23 different sub networks have been defined considering the seismicity of the area and the network geometry (Figure 3). In a wide area like the one here taken into account, the correct definition of these sub-networks is of crucial importance mainly to correctly detect events with nearly coincident origin time, but occurred in different places. In particular, 13 sub-networks are designed to detect the seismicity of the Alpine region in the northern and western parts of the area, 5 sub-networks cover the area spanning from the gulf of Genoa to the northern Apennines, 4 sub-networks extend in the Po Plain and foothill areas, and 1 sub-network is specifically designed for detecting offshore earthquakes in the Ligurian Sea.

3186 earthquakes have been detected and processed by the automatic procedure during the 7 months considered in this test and a total amount of 26558 P-wave and 23911 S-wave arrival times have been automatically detected. The automatic catalogue carried out by the CASP procedure has been compared to the Italian Seismic Bulletin published by INGV. In the same period and in the same area the Italian Seismic Bulletin reports 943 earthquakes, localized through overall 9108 P-wave and 5129 S-wave arrival times, taking into account only the stations considered in this test (Figure 1). Figure 4 shows the differences between manual and automatic P- (panel a) and S-phase arrival times (panel b), and between manual and automatic locations in terms of epicentral position (panel c), focal depth (panel d), and magnitude values (panel e).

Regarding P-phases, the common manual and automatic readings are 7156. The picking rate (that is the ratio between the total number of common automatic and manual readings to all available manual ones) is about 79%. The distribution of the differences between the reference arrival times and the automatic ones (Figure 4, panel a) shows an average value of 0.00 s with a standard deviation of 0.16 s. The 92% of differences are within 0.1 s and the 95% are within 0.2 s.

Regarding the S-phase picks, the common readings are 4769. The picking rate is about 99%. The distribution of the differences shows an average value of 0.05 s with a standard deviation of 0.26 s (Figure 4, panel b). The 82% of the differences are within 0.2 s and the 93% are within 0.5 s. The automatic S-phase readings tend, in average, to slightly anticipate the reference ones.

The comparison between manual and automatic data in terms of arrival times confirms what already shown by Scafidi et al., 2016, that is that the automatic procedure is suitable to create a catalogue of arrival times with a level of accuracy suitable to be employed in a local tomographic study.

Looking at the comparison in terms of earthquake locations, we have to keep in mind that some differences could be also due to the different velocity model and locating methodology used by the automatic and the manual procedures. The automatic procedure adopts the probabilistic approach of NonLinLoc with a three-dimensional velocity models for P- and S-wave based on local tomographic studies centered on the tested area (Scafidi et al. 2009, Scafidi and Solarino 2012). The manual revised locations are calculated using the methodology described in Battelli et al., 2013, based on a 1D velocity model with 3 layers, with a fixed P-wave to S-wave velocity ratio.

All earthquakes listed in the manual bulletin for the area under study have been detected and located by the CASP procedure. The average difference between manual and automatic epicentral locations is equal to 2.93 km with a standard deviation of 2.59 km and the 90% of the differences are within 6.17 km (Figure 4, panel c). Considering depth differences (that, generally, are more influenced by the difference in the location procedure), the average difference value is 0.89 km with a standard deviation of 4.90 km, and the 90% of data are within 6.02 km. On average, the automatic focal depths are slightly shallower than the manual ones.

Regarding the magnitude, the average value of differences between reference and automatic values stays within 0.3 (that is the average estimated error of magnitude calculation), with a standard deviation of 0.2.

It is worth noting that, beyond the common events, the automatic procedure has detected 2243 more events and a manual visual inspection on such events confirmed that they are real earthquakes. The automatic catalogue reports, for the area under study, about 3 times more earthquakes. Figure 5 shows the depth and magnitude distribution of all events listed in the automatic catalogue (light gray) versus those of all events in the INGV bulletin (dark gray). Most earthquakes added by the automatic procedure are characterized by small magnitudes, also lower than 0, that were not present in the reference bulletin. As a consequence, the magnitude of completeness of the automatic catalogue results significantly lowered.

The CASP procedure took an average of 2.4 hours of computation times to process a complete day of data on a standard workstation equipped with an Intel Core i7-7700 CPU using 4 parallel threads.

2) Seismic sequence case study: catalogue of aftershocks following the 24 August 2016 "Amatrice" earthquake

The second case study is aimed at testing the capability of the automatic procedure to compile a catalogue during an intense seismic crisis. To this end, the data recorded by the dense network installed after the M_W 6.0 "Amatrice" earthquake has been taken into account for the period August 24, 2016 to September 30, 2016. The daily continuous recordings of all available stations have been automatically downloaded from the EIDA database. With respect to the previous test, the considered area is smaller, 150 X 150 km wide (Figure 6), while the average number of events per day is considerably higher (13465 earthquakes in 38 days, as reported in the INGV bulletin).

Figure 6 shows the earthquake distribution reported in the reference catalogue compiled by INGV seismologists. Of note, that this reference bulletin is still preliminary, since data have not been fully revised. The epicentral locations (circles, top panel), and the depth and magnitude distributions (bottom panels) are reported.

To automatically compile the catalogue, the seismic stations have been grouped into 3 sub-networks (Figure 7).

The differences between manual and automatic revised catalogue are shown in Figure 8. The picking rate for P-phases is about 76% (the common automatic and manual readings are about 144000) while for S-phases is about 90% (more than 133000 are the common readings). The average value of P-phase differences (i.e., the difference between manual and automatic P arrival times) is equal to 0.06 s with a standard deviation of 0.31 s. About 89% of arrival time differences are lower than 0.1 s and about 92% are within 0.2 s. For S-phase, the average difference is 0.07 s (with a standard deviation of 0.40 s), about 86% of differences are lower than 0.2 s and about 94% are within 0.5 s. Similarly to the previous test, the automatic S-phase readings tend, in average, to slightly anticipate the reference ones.

Regarding the reference and automatic epicentral location differences (Figure 7), the average value is 1.15 km, with a standard deviation of 2.04 km, and the 90% of epicentral differences are still lower than 1.84 km. The focal depths computed by the CASP procedure are slightly shallower than the manual ones. The average difference between manual and automatic focal depths is 3.14 km with a standard deviation of 2.57 km, and the 90% of depth differences are within 5.93 km.

The automatically computed magnitude values are generally lower than the manual ones. The average magnitude difference is about 0.5 with a standard deviation of 0.30 (that is within the average estimated error of magnitude calculation). Probably, this effect can be also ascribed to the focal depth differences, previously observed; the manually revised earthquakes are deeper than the automatically derived ones and, as a consequence, their magnitude values tend to exceed the automatic ones.

More than 95% of the events reported in the reference catalogue have been detected and localized by the automatic procedure. As in the previous test, beyond the earthquakes listed in both reference and automatic catalogue, the CASP procedure has recognized and located about 35000 more seismic events. Figure 9 shows the depth and magnitude distribution of all earthquakes listed in the automatic catalogue (light gray) versus those of all events in the INGV bulletin (dark gray). The events recognized by the CASP procedure only are in average characterized by small magnitudes.

Figure 10 shows the temporal distribution of the number of events per day comparing the reference and the automatic catalogues. It is evident that, during an intense seismic crisis like the one taken into account here, the automatic procedure allows to averagely process (detect and localize) about 3.7 times more earthquakes than manual revision. Therefore, the automatic catalogue presents a magnitude of completeness significantly lower than the reference one.

To process all data of the most seismically active day with more than 2000 detected events, the CASP procedure took less than 8 hours of computation times on a standard workstation equipped with an Intel Core i7-7700 CPU using 4 parallel threads. The CASP procedure has also proved to be able to detect and locate up to more than 2 events occurring in a minute.

Conclusions

In this paper, we have presented a completely automatic procedure to produce a complete seismic dataset (P- and S-wave arrival times, location and strong-motion parameters). This procedure uses continuous recordings provided by a network of seismic stations (for example, downloadable from an open-access database such as EIDA) to detect and locate earthquakes. The procedure is driven by a set of user-definable parameters that can be opportunely calibrated according to network geometry and seismic features of the target area. The picker engine employed into the CASP procedure is the "RSNI-Picker$_2$" software proposed by Scafidi et al. (2018). The comparison of automatic catalogues compiled by CASP with reference ones derived from the manual revision of the data made by INGV seismologists, has shown the effectiveness and reliability of the proposed procedure.

As a result in both application tests taken into account in the present study, the most differences (90% of data) among manual and automatic readings fall within 0.1 s for P-phases and 0.3 s for S-phases with an average picking rate of about 77% and 94%, for P- and S-phase respectively.

Moreover, the automatically derived epicenters fall within few kilometers from the reference ones. The focal depths computed by the CASP procedure showed an almost systematic shift with respect to the reference ones even if the average differences are about 0.70 and 3.14 km for the two tested scenarios. These differences are presumably due to the usage of different location methods (e.g., difference of location algorithm and velocity model) into automatic and manual processing.

Besides these encouraging results, it is worth stressing on the capability of the CASP procedure in detecting and locating very low-magnitude earthquakes, leading to catalogue characterized by a magnitude of completeness lower than that shown by reference bulletins. This advantage is extremely evident when processing data during an intense seismic crisis characterized by a large number of event per day. In such conditions, the very short processing time needed by the CASP procedure allows a better recognition and a more complete analysis of the seismicity.

Following the previous considerations, our procedure appears a promising tool for automatically compiling seismic datasets both in the context of seismic monitoring at a regional scale and during seismic sequences. The reliability and accuracy of the automatically derived data showed that the seismic datasets compiled through the CASP procedure can be confidently used in several application fields in seismology or engineering-seismology such as seismic monitoring, tomographic studies, seismotectonic analyses, seismic hazard assessments, and ground motion prediction equation estimates.

Currently, the CASP software is effectively used in various scientific projects and operational environments. It is used by the University of Genoa for the automatic seismic monitoring of the North-western Italy and also for the automatic calculation of strong-motion parameters at dam sites for the “Provincia Autonoma di Trento”. For both applications, the software automatically processes data streams in real time. CASP has also been used as a state-of-the-art technique to develop a comprehensive high-resolution earthquake catalog in the international scientific project “The Central Apennines Earthquake Sequence Under a New Microscope” (http://gotw.nerc.ac.uk/list_full.asp?pcode=NE%2FR000794%2F1&cookieConsent=A), led by NERC British Geological Survey in partnership with USGS (United States Geological Survey), INGV, University of Bristol, EPOS (European Plate Observing System), Stanford University and Columbia University. Moreover, the research of Bindi et al., 2018, focused on Ground-Motion Prediction Equations calibration in Central Italy, has been based on a high resolution earthquake catalog created by the CASP software.

Data and resources

Seismic waveform data used in this study were collected by RSNI (Regional Seismic network of North-western Italy, GU international code; University of Genoa, 1967) managed by the Università degli Studi di Genova, and by the European Integrated Data Archive (EIDA) of the Observatories & Research Facilities for European Seismology (ORFEUS). Reference manual bulletins were taken from INGV(National Institute of Geophysics and Vulcanology). Most figures were prepared using the Generic Mapping Tools software package (http://gmt.soest.hawaii.edu/).

Researchers interested in the CASP procedure are welcomed to contact the authors for any scientific project or for a free trial.

Acknowledgments

The authors thank SRL editor Zhigang Peng, associate editor Erol Kalkan, and an anonymous reviewer for helpful comments and considerations.

References

Akaike, H. (1974). Markovian representation of stochastic processes and its application to the analysis of autoregressive moving average process. Annals of the Institute of Statistical Mathematics 26 (1), 363–387.

AlpArray Seismic Network (2014). Eastern Alpine Seismic Investigation (EASI) - AlpArray Complimentary Experiment. AlpArray Working Group. Other/Seismic Network. doi:10.12686/alparray/xt_2014

Battelli, P., Arcoraci L., Berardi M., Castellano C., Marchetti A., Mele F., Nardi A., Rossi A. (2013). Italian Seismic Bulletin 2010, Quaderni di Geofisica, anno 2013 N. 115, ISSN 1590-2595.

Benz, H. (2017). Building a National Seismic Monitoring Center: NEIC from 2000 to the Present, Seismological Research Letters, Volume 88, Number 2B, 457-461.

Benz, H. M., N. D. McMahon, R. C. Aster, D. E. McNamara, and D. B. Harris (2015). Hundreds of earthquakes per day: The 2014 Guthrie, Oklahoma, earthquake sequence, Geophys. Res. Lett. 86, no. 5, 1318–1325.

Bindi, D., D. Spallarossa, M. Picozzi, D. Scafidi, F. Cotton (2018). Impact of Magnitude Selection on Aleatory Variability Associated with Ground- Motion Prediction Equations: Part I—Local, Energy, and Moment Magnitude Calibration and Stress- Drop Variability in Central Italy, Bulletin of the Seismological Society of America, 108 (3A), 1427–1442, doi: 10.1785/0120170356.

Ellsworth, W. L. (2013). Injection-induced earthquakes. Science 341, no. 6142, doi: 10.1126/science.1225942

Grigoli, F., S. Cesca, E. Priolo, A. P. Rinaldi, J. F. Clinton, T. A. Stabile, B. Dost, M. G. Fernandez, S. Wiemer, and T. Dahm (2017). Current challenges in monitoring, discrimination, and management of induced seismicity related to underground industrial activities: A European perspective, Rev. Geophys., 55, doi:10.1002/2016RG000542.

INGV Seismological Data Centre. (2006, January 1). Rete Sismica Nazionale (RSN). Istituto Nazionale di Geofisica e Vulcanologia (INGV), Italy. https://doi.org/10.13127/sd/x0fxnh7qfy

Lomax, A., J. Virieux, P. Volant, and C. Berge-Thierry (2000). Probabilistic earthquake location in 3D and layered models: introduction of a Metropolis–Gibbs method and comparison with linear locations. In: Advances in seismic event location, ed. C. H. Thurber and N. Rabinowitz, 101–134. Dordrecht and Boston: Kluwer Academic Publishers.

Lomax, A., A. Michelini, and A. Curtis (2009). Earthquake Location, Direct, Global-Search Methods. In: Encyclopedia of Complexity and Systems Science, ed. R. A. Meyers, 2449–2473. Springer-Verlag, New York.

Moretti, M., Pondrelli, S., Margheriti, L., Abruzzese, L., Anselmi, M., Arroucau, P., et al. (2016). SISMIKO: Emergency network deployment and data sharing for the 2016 central italy seismic sequence, Annals of Geophysics, 59, https://doi.org/10.4401/ag-7212.

Richter, C. F. (1935). An instrumental earthquake magnitude scale, Bull. Seism. Soc. Am. 25, 1–31.

Richter, C. F. (1958). Elementary Seismology, Freeman, San Francisco, 578 pp.

Scafidi, D., Viganò A., Ferretti G., and Spallarossa D. (2018). Robust picking and accurate location with RSNI-Picker$_2$: real-time automatic monitoring of earthquakes and non-tectonic events, Seismological Research Letters, 89 (4), pp. 1478–1487, doi: https://doi.org/10.1785/0220170206.

Scafidi, D., Spallarossa D., Turino C., Ferretti G. and Viganò A. (2016). Automatic P- and S-Wave Local Earthquake Tomography: Testing Performance of the Automatic Phase-Picker Engine “RSNI-Picker”, Bulletin of the Seismological Society of America, Vol. 106, No. 2, pp. 526–536, doi: 10.1785/0120150084.

Scafidi, D., S. Barani, R. De Ferrari, G. Ferretti, M. Pasta, M. Pavan, D. Spallarossa, and C. Turino

(2015). Seismicity of Northwestern Italy during the last 30 years, J. Seismol. 19, 201–218.

Scafidi, D., and S. Solarino (2012). Can local earthquake tomography settle the matter about subduction in the Northern and Central Apennines? Response from a new high resolution P velocity and Vp/Vs ratio 3-D model. Tectonophysics 554–557, 63–73; doi:10.1016/j.tecto.2012.06.007.

Scafidi, D., S. Solarino, and C. Eva (2009). P wave seismic velocity and Vp/Vs ratio beneath the Italian peninsula from local earthquake tomography. Tectonophysics 465, 1–23; doi:10.1016/j.tecto.2008.07.013.

Spallarossa, D., G. Ferretti, D. Scafidi, C. Turino, and M. Pasta (2014). Performance of the RSNI-Picker, Seismol. Res. Lett. 85, 1243–1254.

University of Genoa (1967). Regional Seismic Network of North-western Italy. International Federation of Digital Seismograph Networks. Other/Seismic Network. 10.7914/SN/GU.

Figure captions

Figure 1: flowchart of the CASP procedure for compiling, in a fully automatic way, seismic datasets.

Figure 2: map of the selected test area (rectangular box) for the regional seismicity case study. The distribution of seismic stations (triangles) and the location of earthquakes (circles) taken from the INGV manual database for the time period going from 2017-01-01 to 2017-07-31 are also shown.

Figure 3: sub-networks considered for compiling the catalogue of the North-western Italy seismicity through the CASP procedure.

Figure 4: comparison between reference and automatic catalogue. Differences between reference arrival times (manually picked) and concurrent automatic ones, for P- (a panel) and S-phase (b-panel); Differences between reference location and concurrent automatic ones in terms of epicentral differences (c panel), focal depth differences (d-panel), and magnitude differences (e panel).

Figure 5: magnitude and depth distribution of the earthquakes listed in the automatic catalogue (light gray), and of earthquakes of the reference INGV bulletin (dark gray), for the north-western Italy case study.

Figure 6: map of the selected test area (rectangular box) for the seismic sequence case study. The distribution of seismic stations (triangles) and the location of earthquakes (circles) taken from the INGV manual database for the time period going from 2016-08-24 to 2016-09-30 are also shown

Figure 7: sub-networks considered for compiling the catalogue of the seismic sequence through the CASP procedure.

Figure 8: comparison between reference and automatic catalogue. Differences between reference arrival times (manually picked) and concurrent automatic ones, for P- (a panel) and S-phase (b-

panel); Differences between reference location and concurrent automatic ones in terms of epicentral differences (c panel), focal depth differences (d-panel), and magnitude differences (e panel).

Figure 9: magnitude and depth distribution of the earthquakes listed in the automatic catalogue (light gray), and of earthquakes of the reference INGV bulletin (dark gray), for the seismic sequence case study.

Figure 10: number of events per day as detected by the automatic procedure compared to the number of events per day revised in the manual reference catalogue, for the seismic sequence case study.

Authors full mailing address

Davide Scafidi : scafidi@dipteris.unige.it
Daniele Spallarossa : daniele@dipteris.unige.it
Gabriele Ferretti : ferretti@dipteris.unige.it
Simone Barani : barani@dipteris.unige.it
Barbara Castello : barbara.castello@ingv.it
Lucia Margheriti : lucia.margheriti@ingv.it

Figure 1

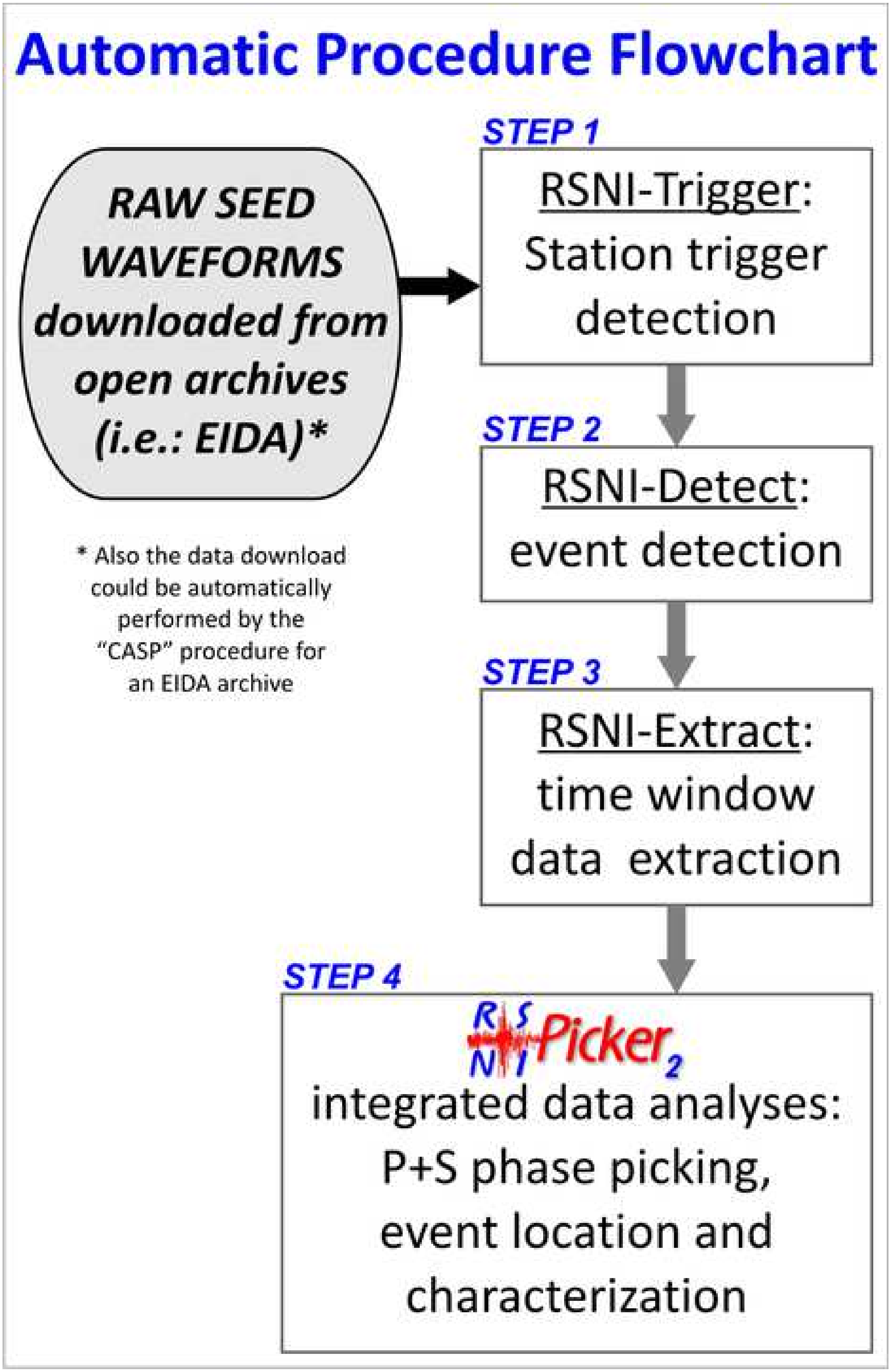

Automatic Procedure Flowchart
RAW SEED WAVEFORMS downloaded from open archives (i.e.: EIDA)*
* Also the data download could be automatically performed by the "CASP" procedure for an EIDA archive
STEP 1
RSNI-Trigger: Station trigger detection
STEP 2
RSNI-Detect: event detection
STEP 3
RSNI-Extract: time window data extraction
STEP 4
RSNI-Picker2
integrated data analyses: P+S phase picking, event location and characterization

Figure 2

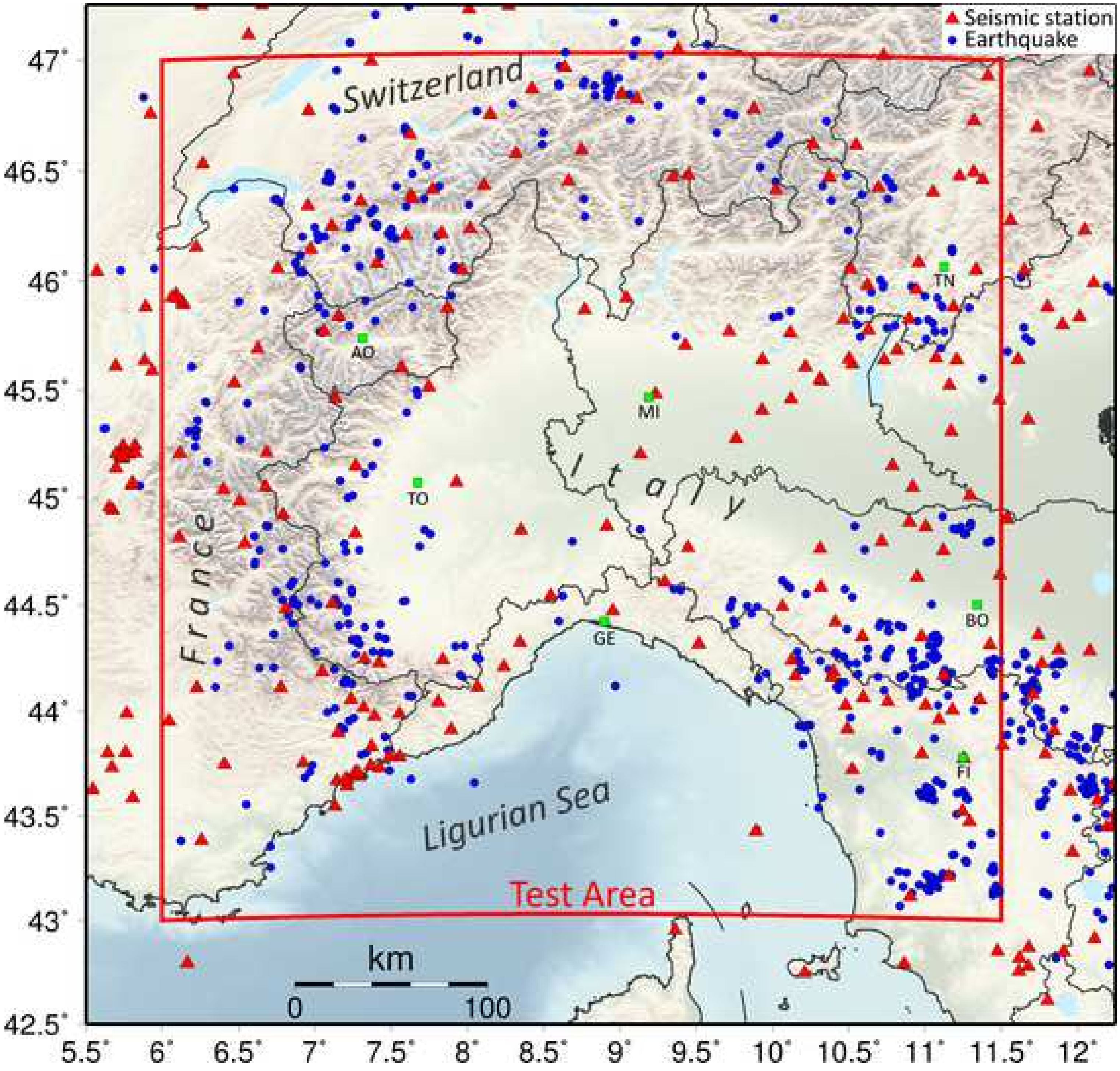

▲ Seismic station
● Earthquake
Switzerland
France
Italy
Ligurian Sea
Test Area
AO
TO
MI
GE
TN
BO
FI
km
0
100
47°
46.5°
46°
45.5°
45°
44.5°
44°
43.5°
43°
42.5°
5.5°
6°
6.5°
7°
7.5°
8°
8.5°
9°
9.5°
10°
10.5°
11°
11.5°
12°

Figure 3

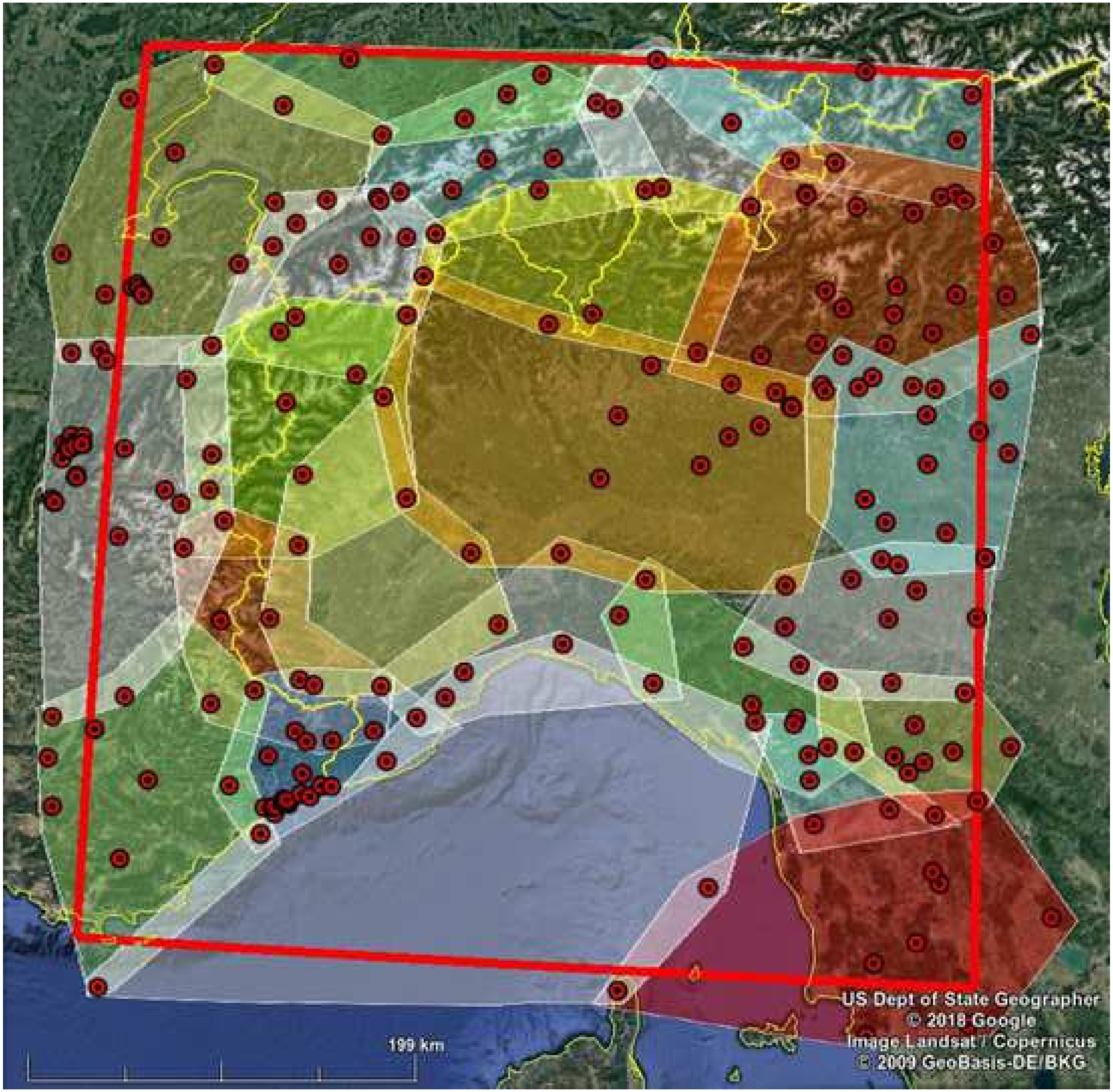

199 km
US Dept of State Geographer
© 2018 Google
Image Landsat / Copernicus
© 2009 GeoBasis-DE/BKG

Figure 4

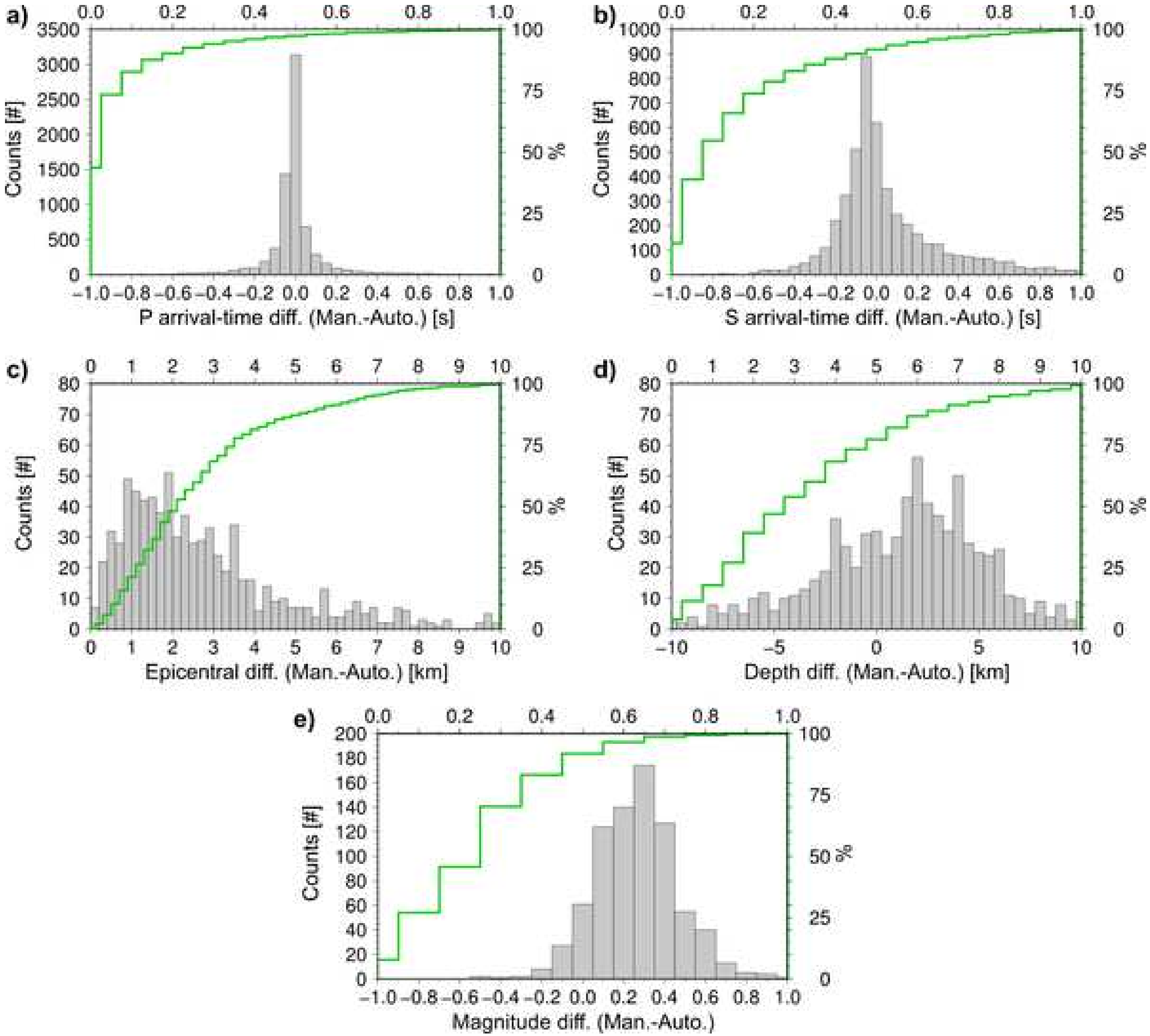

a)
0.0 0.2 0.4 0.6 0.8 1.0
3500 3000 2500 2000 1500 1000 500 0
Counts [#]
100 75 50 25 0
%
−1.0 −0.8 −0.6 −0.4 −0.2 0.0 0.2 0.4 0.6 0.8 1.0
P arrival-time diff. (Man.-Auto.) [s]
b)
0.0 0.2 0.4 0.6 0.8 1.0
1000 900 800 700 600 500 400 300 200 100 0
Counts [#]
100 75 50 25 0
%
−1.0 −0.8 −0.6 −0.4 −0.2 0.0 0.2 0.4 0.6 0.8 1.0
S arrival-time diff. (Man.-Auto.) [s]
c)
0 1 2 3 4 5 6 7 8 9 10
80 70 60 50 40 30 20 10 0
Counts [#]
100 75 50 25 0
%
0 1 2 3 4 5 6 7 8 9 10
Epicentral diff. (Man.-Auto.) [km]
d)
0 1 2 3 4 5 6 7 8 9 10
80 70 60 50 40 30 20 10 0
Counts [#]
100 75 50 25 0
%
−10 −5 0 5 10
Depth diff. (Man.-Auto.) [km]
e)
0.0 0.2 0.4 0.6 0.8 1.0
200 180 160 140 120 100 80 60 40 20 0
Counts [#]
100 75 50 25 0
%
−1.0 −0.8 −0.6 −0.4 −0.2 0.0 0.2 0.4 0.6 0.8 1.0
Magnitude diff. (Man.-Auto.)

Figure 5

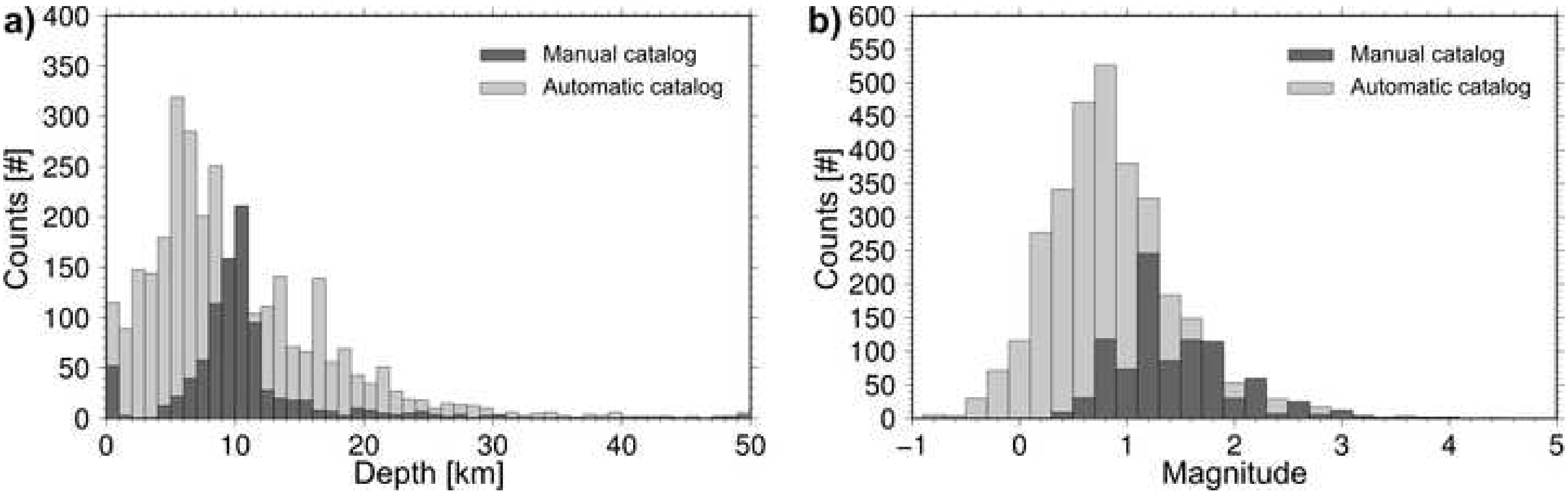

a)
400
350
300
250
200
150
100
50
0
Counts [#]
0
10
20
30
40
50
Depth [km]
Manual catalog
Automatic catalog
b)
600
550
500
450
400
350
300
250
200
150
100
50
0
Counts [#]
−1
0
1
2
3
4
5
Magnitude
Manual catalog
Automatic catalog

Figure 6

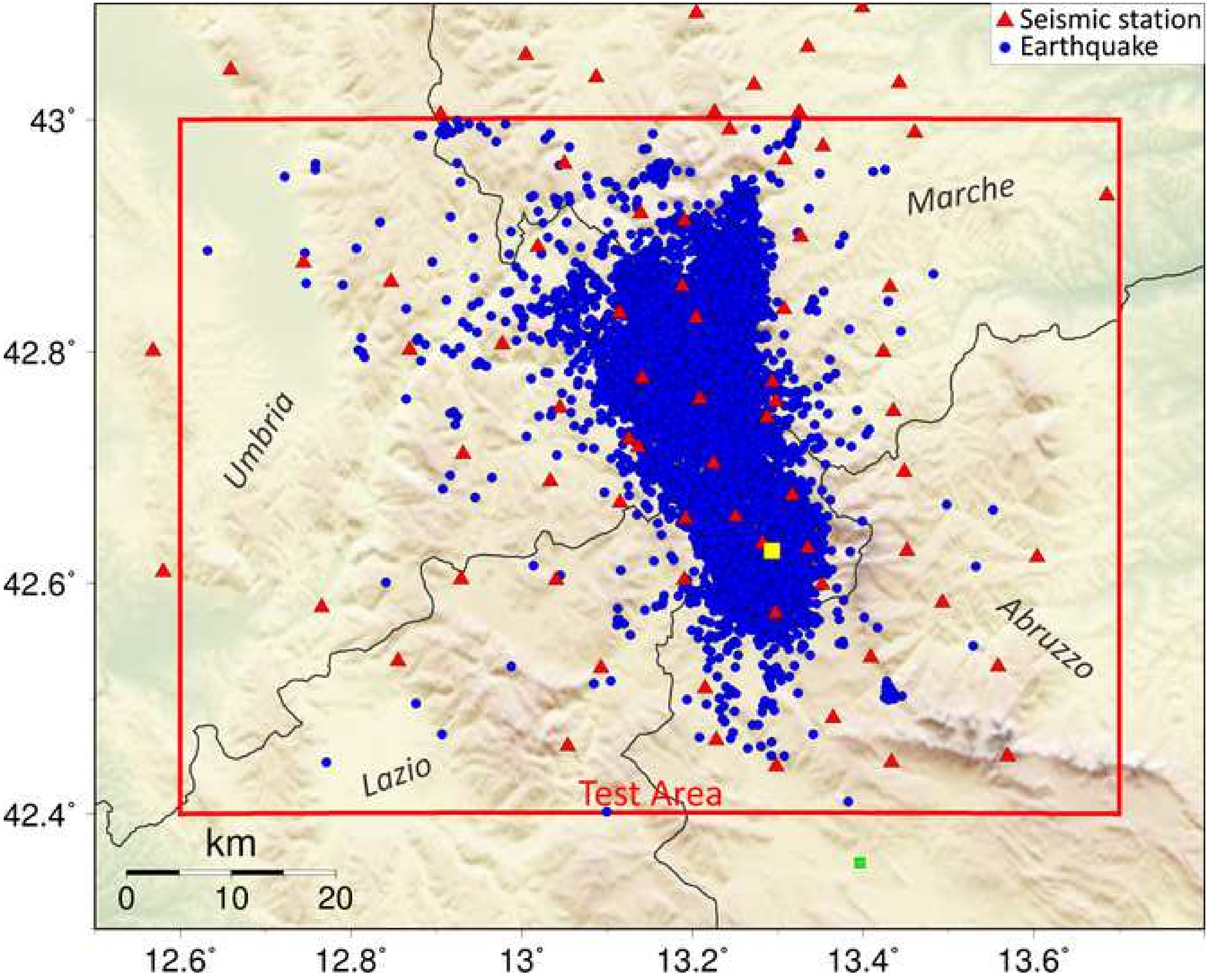
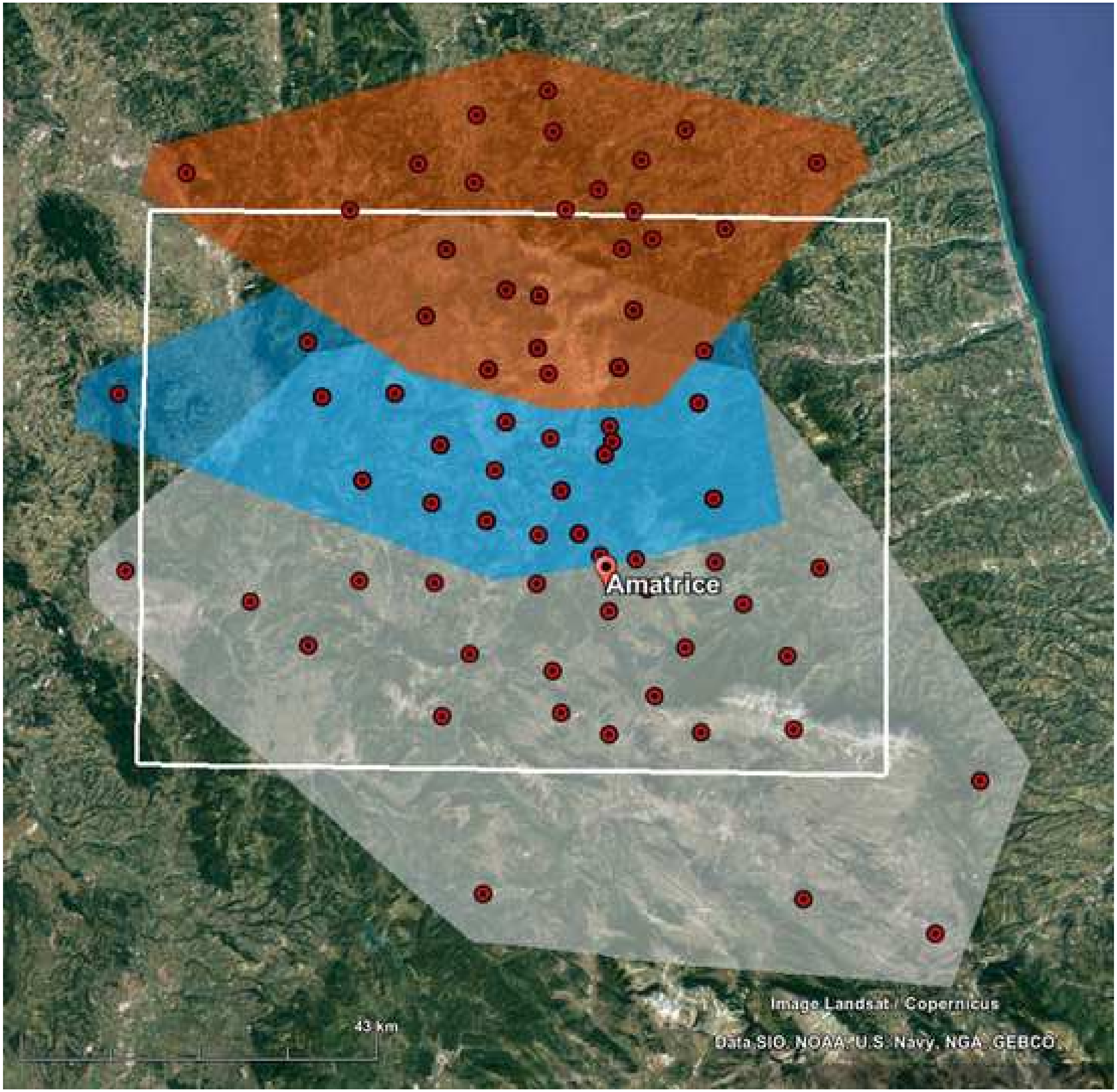

▲ Seismic station
● Earthquake
43°
42.8°
42.6°
42.4°
12.6°
12.8°
13°
13.2°
13.4°
13.6°
Marche
Umbria
Abruzzo
Lazio
Test Area
km
0
10
20

Figure 7


Amatrice
Image Landsat / Copernicus
Data SIO, NOAA, U.S. Navy, NGA, GEBCO
43 km

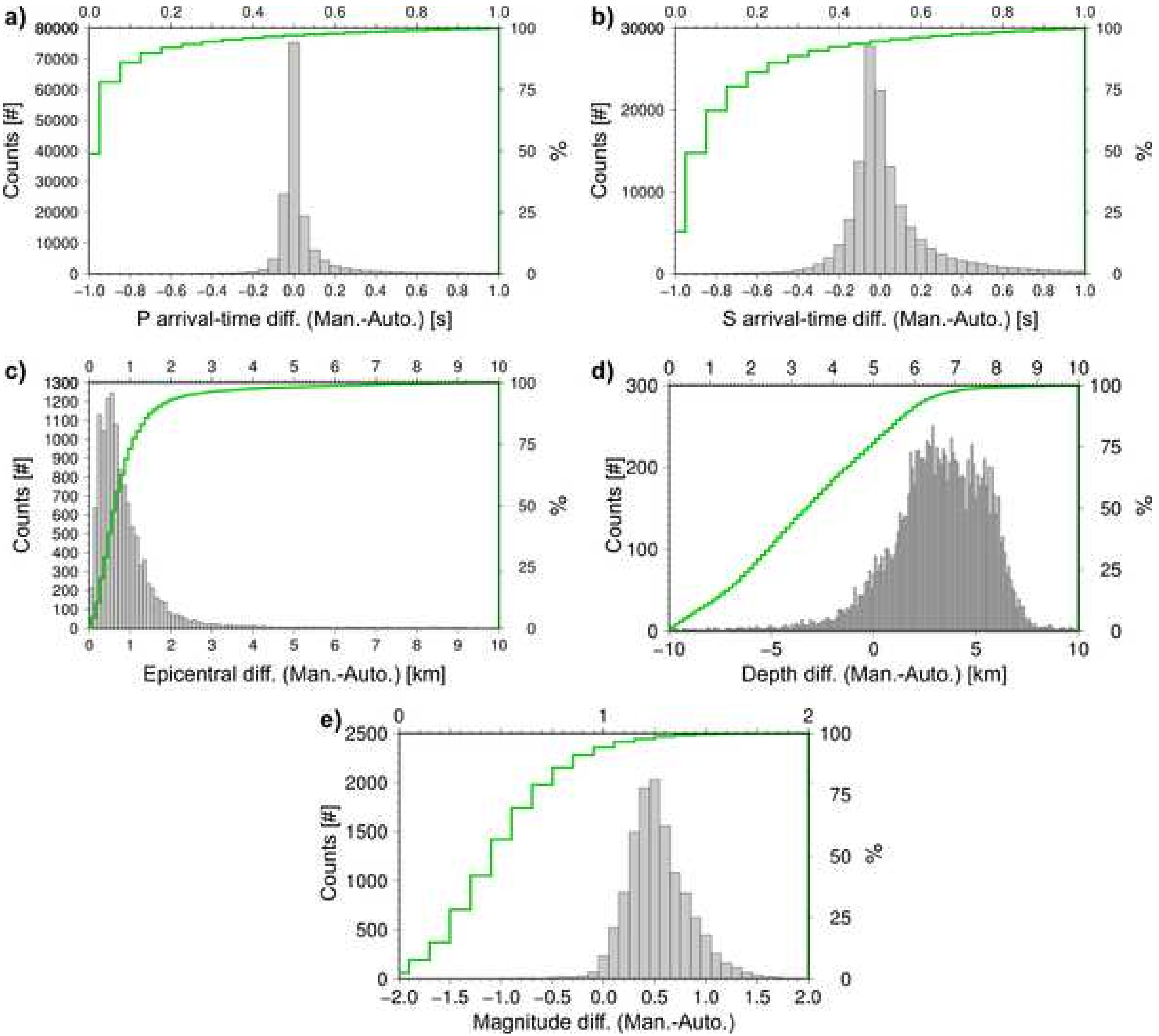

a)
Counts [#]
%
P arrival-time diff. (Man.-Auto.) [s]
b)
S arrival-time diff. (Man.-Auto.) [s]
c)
Epicentral diff. (Man.-Auto.) [km]
d)
Depth diff. (Man.-Auto.) [km]
e)
Magnitude diff. (Man.-Auto.)

Figure 9

Click here to access/download;Figure;Fig9.tif

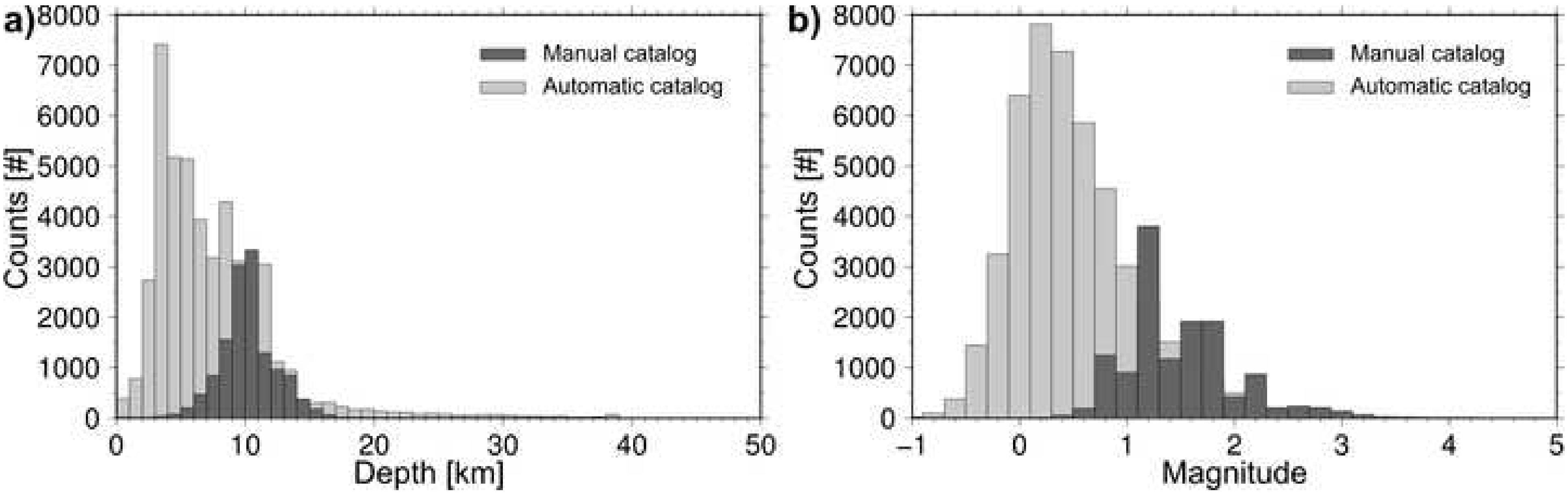
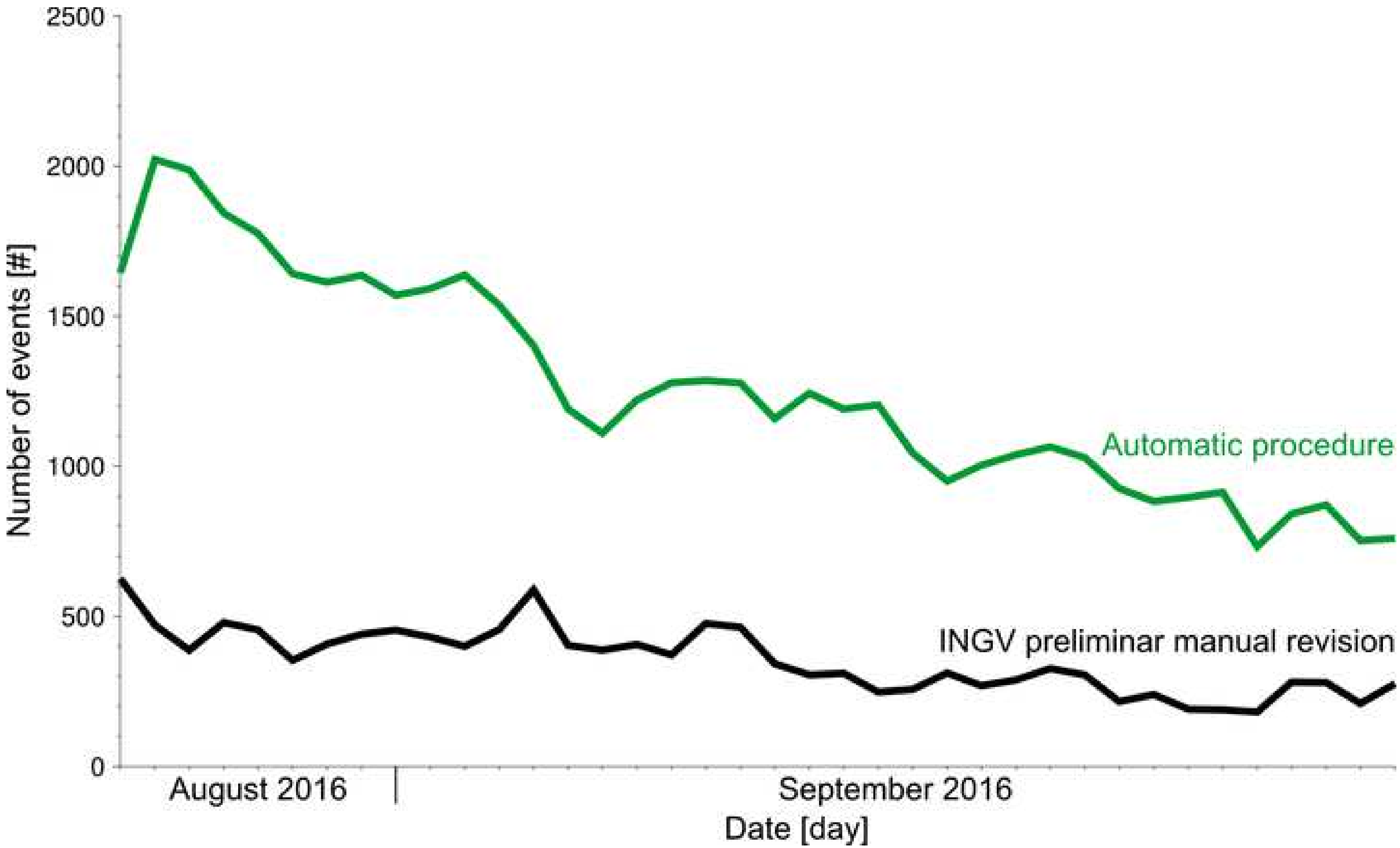


Figure 10


2500
2000
1500
1000
500
0
Number of events [#]
Automatic procedure
INGV preliminar manual revision
August 2016
September 2016
Date [day]