Manuscript prepared for Nat. Hazards Earth Syst. Sci.

with version 2014/09/16 7.15 Copernicus papers of the LaTeX class copernicus.cls.

Date: 27 July 2015

# Appraising the Early-est earthquake monitoring system for tsunami alerting at the Italian candidate Tsunami Service Provider

Fabrizio Bernardi[1], Anthony Lomax[2], Alberto Michelini[1], Valentino Lauciani[1], Alessio Piatanesi[1], and Stefano Lorito[1]

[1]Istituto Nazionale di Geofisica e Vulcanologia, Via di Vigna Murata 605, 00143 Roma, Italy

[2]ALomax scientific, Allé du Micocoulier 161, 06370 Mouans-Sartoux, France

*Correspondence to:* Fabrizio Bernardi (fabrizio.bernardi@ingv.it)

**Abstract.** In this paper we present and discuss the performance of the procedure for earthquake location and characterization implemented in the Italian candidate Tsunami Service Provider at INGV in Roma. Following the ICG/NEAMTWS guidelines, the first tsunami warning messages are based only on seismic information, i.e. epicenter location, hypocenter depth and magnitude, which are automatically computed by the software Early-est. Early-est is a package for rapid location and seismic/tsunamigenic characterization of earthquakes. The Early-est software package operates on offline-event or continuous-realtime seismic waveform data to perform trace processing and picking, and, at a regular *report interval*, phase association, event detection, hypocenter location, and event characterization. Early-est also provides $mb$, $M_{wp}$ and $M_{wpd}$ magnitude estimations. $mb$ magnitudes are preferred for events with $M_{wp} \lesssim 5.8$, while $M_{wpd}$ are valid for events with $M_{wp} \gtrsim 7.2$. In this paper we present the earthquake parameters computed by Early-est from the beginning of 2012 till the end of December 2014 at global scale for events with magnitude $M \geq 5.5$, and the detection timeline. We compare the earthquake parameters automatically computed by Early-est with the same

parameters listed into reference catalogs. Such reference catalogs are manually revised/verified by scientists. The Goal of this work is to test the accuracy and reliability of the fully automatic locations provided by Early-est. From our analysis the epicenter location, hypocenter depth and magnitude parameters do not differ significantly from the values in the reference catalogs. Both magnitudes $mb$ and $M_{wp}$ show differences with respect to the reference catalogs. We thus derived correction functions in order to minimize the differences and correct biases between our values and the ones of the reference catalogs. Particularly relevant is correction of the $M_{wp}$ distance dependency, since this magnitude refers to the larger and probably tsunamigenic earthquakes. $M_{wp}$ values at stations with epicentral distance $\Delta \lesssim 30^o$ are significantly overestimated with respect the CMT-global solutions, whereas $M_{wp}$ values at stations with epicentral distance $\Delta \gtrsim 90^o$ are slightly underestimated. After applying such distance correction the $M_{wp}$ provided by Early-est differs from CMT-global catalog values of about $\delta M_{wp} \approx 0.0 \mp 0.2$. Early-est continuously acquires time series data and updates the earthquake source parameters. Our analysis shows that the epicenter coordinates and the magnitude values converge within less than 10 minutes (5 minutes in the Mediterranean area) toward the stable values. Our analysis shows that we can compute $M_{wp}$ magnitudes that do not suffer of the short epicentral distance dependency overestimation, and we can provide robust and reliable earthquake source parameters to compile tsunami warning message within less than about 15 minutes after event origin time.

## 1 Introduction

Tsunamis may produce dangerous coastal flooding and inundations accompanied by powerful currents which can cause significant damage and casualties. A tsunami may be generated when a large or great earthquake occurs in oceans or inland close to the coast. When such earthquakes occur, a tsunami warning should be issued to alert national authorities and emergency management officials to take actions for the entire tsunami hazard zone such as evacuating the population or to securing

critical facilities such as nuclear power plants. With advance evacuating planning and well-informaed communities, tsunami warnings could be also sent directly to the population.

Reliable tsunami warnings should be disseminated as fast as possible in order to be effective also for the coastal areas very close to the earthquake source, since a tsunami may arrive at these areas within the first minutes after the event origin time. Populations exposed to tsunami hazard in the near field of the source, however, should be aware that the time between warning issuance and tsunami impact may be too short to escape the tsunami, that warning may arrive even after the tsunami, or the system may be subject to failure, for several reasons. Hence, the population should know how to self-evacuate relying, when present, on natural warnings, such as strong and/or unusually long shaking, ocean withdrawal, anomalously rising tide, roaring sounds from the ocean, etc.

To provide the earliest possible alerts initial warnings from regional tsunami warning systems are normally based only on seismic information. Thus, fast, precise and reliable earthquake source parameters like epicenter coordinates, hypocenter depth and magnitude are crucial for seismologically based tsunami early warning procedures. This is particularly important in the Mediterranean Sea, where the tsunami wave travel times between source regions and coast lines are short and dedicated deep-sea instruments, such as the DART buoys (http://nctr.pmel.noaa.gov/Dart/), are not in place.

The Istituto Nazionale di Geofisica e Vulcanologia (INGV) in Italy is a candidate Tsunami Service Provider (cTSP) in the framework of ICG/NEAMTWS (NEAMTWS, 2011), which is the tsunami early warning and mitigation system established by IOC/UNESCO for the North-eastern Atlantic, the Mediterranean and connected seas. For this reason, the Centro Allerta Tsunami (CAT) (tsunami alert centre in Italian), has been established at the INGV headquarter in Rome at the end of 2013. The CAT mission is to implement and maintain a 24/7 service along with the ordinary seismic surveillance of the national territory, and to work towards a Probabilistic Seismic Hazard Assessment (PSHA) for the italian coasts, that is a tsunami hazard map for seismically-induced tsunamis (Basili et al., 2013). CAT-INGV started operations on a 24/7 basis as cTSP in October 2014. Monthly communication tests are performed towards national authorities, subscriber IOC member states and

other institutions, such as the DG-ECHO Emergency Response Coordination Center in Brussels. In the NEAM region there are three other cTSPs in operation: CENALT in France, NOA in Greece, and KOERI in Turkey. IPMA, in Portugal, should begin operations soon. Each of these cTSP's has its specific competence source areas within the NEAM region.

At the national level, INGV is responsible for issuing messages to the Civil Protection authority, which is presently responsible for alert dissemination. INGV also maintains the national seismic network and exchanges seismic data in real time with a number of international seismic data providers. The Istituto Superiore per la Protezione e Ricerca Ambientale (ISPRA) maintains the national sea level network and provides real time data to INGV monitoring room. The implemented tsunami warning procedure uses the *Early-est* software developed by Lomax and Michelini (2009a, b, 2011, 2012) to rapidly detect, locate and determine the magnitude for large to great regional and teleseismic earthquakes.

The purpose of this paper is to analyze the Early-est performances regarding past events, in order to evaluate its reliability for the near-real time tsunami warnings disseminated by the INGV, and eventually tune the procedure as a whole.

INGV cTSP follows the ICG/NEAMTWS guidelines. ICG/NEAMTWS rules establish that a cTSP must disseminate a tsunami message, with warning levels that depend on location, magnitude and depth of the earthquake according to a decision matrix, for all earthquakes with magnitudes $M \geq 5.5$ in their zone of competence. Messages are sent for earthquakes that are large and shallow enough, and occurring in sea areas or inland but sufficiently close to the coast to possibly generate a tsunami. INGV is responsible for the earthquake and tsunami source zone extending from the Gibraltar Strait in the west, to Marmara and Levantine Seas to the east.

The seismicity in the Mediterranean region is moderate to high but includes also $M8+$ earthquakes in the past, that generated significant tsunamis (Maramai et al., 2014; Lorito et al., 2015). It is difficult to assess if $M9$-class earthquakes might occur, can not be excluded (Kagan and Jackson, 2013). Even if tsunamigenic earthquakes are likely to occur, their time recurrence intervals are how-

ever quite long (Koravos et al., 2003; Jenny et al., 2004; Bungum and Lindholm, 2007); moreover, the Mediterranean Sea is a relatively small area, and earthquakes with $M \geq 5.5$ do not occur very frequently: the Global CMT catalogs (Dziewonski et al., 1981; Ekström et al., 2012) includes about 125 earthquakes with $M_w \geq 5.5$ within the Mediterranean area, which implies an occurrence rate of $\approx 30$ each 10 years. Early-est has been now running for several years, but only since the beginning of March 2012 has its current major version release been online and its solutions could be systematically archived. Thus we do have few events to analyze for tuning our tsunami alert procedure (Table 1). For this reason we perform our analysis using all worldwide occurred earthquakes located by Early-est since March 2012. To perform the analysis and tune our procedure we proceed by comparing the epicenters, the hypocenter depths and magnitudes estimation provided fully automatically by Early-est with the same parameters provided by other agencies taken as reference. Such agencies provide manually validated/revised locations and magnitude estimations for earthquakes at global scale.

This paper is structured as follows. In the next section, we give a brief overview of the Early-est algorithm. In section 3 we describe the dataset used in our analysis. In the following three sections we then analyze and compare the earthquake source parameters provided by Early-est with the ones provided by the reference agencies; first the epicenter location (section 4), then the hypocenter depth (section 5) and last the magnitude (section 6). In section 7 we will analyze the speed performances of Early-est with respect to the location and the magnitude parameters in order to set the timeline of our automatic tsunami warning procedure. Lastly, we present the discussions and conclusions.

## 2 Early-est Algorithm Description

Early-est is a software package for rapid location and seismic/tsunamigenic characterization of earthquakes. The Early-est software package operates on offline-event or continuous-realtime seismic waveform data to perform trace processing and picking, and, at a regular *report interval*, phase association, event detection, hypocenter location, and event characterization. This characterization

(Table A.1) includes mb and $M_{wp}$ magnitudes, the determination of apparent rupture duration, $T_0$, large earthquake magnitude, $M_{wpd}$, and assessment of tsunamigenic potential using $T_d$ and $T_{50}Ex$, as described in Lomax and Michelini (2009a, b, 2011). The Early-est program reads mini-seed data packets from file or a SeedLink server (http://ds.iris.edu/ds/nodes/dmc/services/seedlink, http://www.seiscomp3.org/wiki, doc/applications/seedlink), respectively, and passes each packet to a *trace-processing* module. The program also calls an *associate/locate* - reporting module at regular *reporting intervals* (e.g. after all data is read for mini-seed; every 1 min for SeedLink). The Early-est software maintains a persistent *pick list* for the current *reporting window* (e.g. the last hour before real-time) and an *event list* for a specified archive interval (e.g. the last 10 days). The *pick list* is updated continuously as picking and trace processing are applied to new data packets. The *event list* is updated at each *reporting interval* as new event locations are found or previous locations are deleted. At each *reporting interval* the *associate/locate* module processes the current *pick list* from scratch, without making use of previous associations or location information from the *event list*; this memory-less procedure simplifies the *associate/locate* module and makes it very robust with respect to changes in the *pick list*, but increases the computational load. To reduce this load, the persistence of association and location information for well located events is currently being added to Early-est.

### 2.1 Trace-processing module

The *trace-processing* module processes each new data packet passed by the Early-est program. This processing includes channel identification, quality control, filtering for picking, picking and further filtering and pre-processing as required for seismic and tsunamigenic event characterization (Table A.1).

Picking in Early-est is performed by *FilterPicker* (Lomax and Michelini, 2012; Vassallo et al., 2012), a general purpose, broad-band, phase detector and picker which is applicable to real-time seismic monitoring and earthquake early-warning. *FilterPicker* uses an efficient algorithm which operates stably on continuous, real-time, broadband signals, avoids excessive picking during large

events, and produces onset timing, realistic timing uncertainty, onset polarity and amplitude information. In practice, it operates on a pre-defined number of frequency bands by generating a set of band-passed time-series with different center frequencies. Characteristic functions are determined for each frequency band and a pick is declared if and when, within a window of predefined time width, the integral of the maximum of the characteristic functions exceeds a pre-defined threshold.

After picking on each new data packet, for each pick in the *pick list* for the current packet channel, the *trace-processing* module applies various analyses on the channel data and updates values needed for event characterization. Recursive, time-domain algorithms are used for all filtering and other time-series processing.

### 2.2 Associate/locate - reporting module

The Early-est *associate/locate - reporting* module calls an *oct-tree associate/locate* module with the current *pick list*, and then the *reporting module* which determines event characterization results and generates graphical and alpha-numeric reporting output. The *oct-tree associate/locate* module efficiently and robustly associates picks, and detects and locates seismic events over the whole Earth from 0 to $700km$ depth using the efficient, non-linearized, probabilistic and global, oct-tree importance-sampling search (Lomax et al., 2001, 2009). See Appendix A for more details.

The Early-est *reporting* module processes the current *pick list* and *event list* to determine event characterization results (Table A.1) and generate graphical, alpha-numeric, XML, HTML and other reporting output for events, picks, stations, etc. An e-mail or other alert message can be generated for each event with magnitudes or tsunamigenic potential exceeding pre-set thresholds. Figure A.1 shows the main graphical display of Early-est, which summarizes the evolving trace-processing, associate/locate and event characterization results in real-time.

## 3 Dataset

The Early-est catalog (EEc in this paper) includes fully automatic and unrevised location and magnitude estimations for 5449 events from around the globe recorded at regional and teleseismic distance with magnitude $M \gtrsim 5.0$. The current major version release of Early-est has been running since the beginning of March 2012. Our analysis will use locations and magnitudes for events occurred from the beginning of March 2012 till end of December 2014; At the beginning of March 2012 Early-est was using about 300 seismic broadband stations. The number of stations has been continuously increasing, and at the end of September 2014 the Early-est software has using a virtual station network of 494 stations (Figure 1).

We use the following as *reference catalogs*: i) the catalog provided by GEOFON project of the Deutches GeoForschungsZentrum (Gc in this paper, http://geofon.gfz-potsdam.de/eqinfo/form.php); ii) the catalog provided by the U.S. National Earthquake Information Center (Nc in this paper); iii) the catalog provided by the EMSC-CSEM (Cc in this paper) (http://www.emsc-csem.org/Earthquake), (Godey et al., 2007); iv) the catalog provided by Global CMT project (CMTc in this work) (Dziewonski et al., 1981; Ekström et al., 2012); v) and the catalog provided by the Pacific Tsunami Warning Center (Pc in this paper) provided to the authors of this paper by the courtesy of Barry Hirshorn by the Pacific Tsunami Warning Center (http://ptwc.weather.gov). The CMTc and the Pc will be used specifically to compare and assets the $M_{wp}$ and $M_{wpd}$ magnitudes.

The above mentioned observatories and centers provide manually verified and/or revised earthquakes source parameters for different time periods. Table 2 summarizes the catalogs abbreviations and time windows for each catalog used in this work. The ICG/NEAMTWS guidelines indicate that tsunami warning must be disseminated for all events in the Mediterranean and Northen-eastern Atlantic regions with $M \geq 5.5$. For this reason, although Early-est locate events with magnitude $M \gtrsim 5.0$, our analysis will focus only on worldwide earthquakes with magnitude $M \geq 5.5$.

## 4 Epicenter location

In this section we use the three reference catalogs Nc, Gc, and Cc and the Early-est catalog EEc.

We first build three couples with the three reference catalogs (Gc-Cc, Cc-Nc and Gc-Nc) and we compute the distance between the epicenter coordinates for each earthquake listed in both catalogs of each couple.

The top panel in figure 2 shows the histograms representing the distributions of the location differences in each couple of the reference catalogs. The $M \geq 5.5$ earthquakes are generally located with a mean distance differences smaller than $\overline{\delta\Delta_{ref}} \lesssim 20 \mp 25[km]$; almost $95\%$ of all earthquakes are located with distance differences $\delta\Delta_{ref} \lesssim 50[km]$. We did not find evidences for geographical and/or tectonic dependence of this uncertainty.

We then compare the epicenter coordinates between the earthquakes listed into the EEc and each of the three reference catalogs (Figure 2, bottom panels), i.e. we build the couples EEc-Cc, EEc-Nc and EEc-Gc. The histograms show that the epicenter location differences between the EEc and the reference catalogs $\delta\Delta_{EEc}$ are similar to the differences plotted on the top panels. The mean location differences between the EEc and the reference catalogs is about $\overline{\delta\Delta_{EEc}} \lesssim 20 \mp 20[km]$ and $95\%$ of all events into the dataset show differences $\delta\Delta_{EEc} \lesssim 45[km]$.

Generally our analysis showed that earthquakes with $M \geq 5.5$ can be located, when using seismic data form global networks, with a empirical uncertainty, defined as the mean location difference with respect the reference catalogs, of about $\nu \approx 20 \mp 25[km]$.

## 5 Hypocenter depth

In this section we proceed as described in the section above: we use the three reference catalogs Nc, Gc, and Cc and the Early-est catalog EEc and to build the catalog couples used in the previous section 4. We then compute the depth difference between the hypocenters for each earthquake listed in both catalogs of each couple.

Figure 3 (top panels) shows the histograms representing the distribution of the depth differences in each couple of the reference catalogs. The hypocenter depth estimation for earthquakes with magnitude $M \geq 5.5$ listed in global catalogs is generally well resolved: the mean and standard deviations difference are $\overline{\delta Z_{ref}} \approx 0 \mp 25[km]$ for all catalog couples. We did not found evidences for geographical and/or tectonic dependence of these differences.

We then compare the hypocenter depths between the EEc and each of the three reference catalogs (Figure 3 bottom panels, couples EEc-Cc, EEc-Nc and EEc-Gc). The bottom panels show that the hypocenter depth estimation differences between the Early-est catalog and the reference catalogs do not differ significantly: the mean difference distributions are about $\overline{\delta Z_{EEc}} \approx 0 \mp 30[km]$.

Generally our analysis showed that hypocenter depth of earthquakes with $M \geq 5.5$ can be precisely estimate, when using seismic data form global networks, with a empirical uncertainty, of about $\nu \approx 00 \mp 30[km]$.

## 6 Magnitude

Early-est provides three different types of magnitude: $mb$, $M_{wp}$ and $M_{wpd}$ (Lomax and Michelini, 2011) and then automatically decides each minute which magnitude type is the most significant following the rules in table 3. The criteria to assign the *best magnitude* listed in Table 3 follow two simple principles: i) a minimum number of observations is required to obtain reliable magnitude estimations, and ii) magnitude types are reliable within magnitude ranges. Following Lomax and Michelini (2009a, b, 2011) we set the validity range $5.8 \leq M_{wp} < 7.2$ for the *best magnitude*; $mb$ is assigned to *best magnitude* when $M_{wp} < 5.8$ and $M_{wpd}$ is assigned to *best magnitude* when $M_{wp} > 7.2$ In this work we compare the Early-est magnitude types $M_{wp}$ and $M_{wpd}$ with respect to the reference magnitude types $M_{wp}$ and $M_w$ Since the ICG/NEAMTWS guidelines prescribe that for earthquakes with depth $Z \leq 100[km]$ a standard general warning should be delivered only for events with $M \geq 5.5$, and no action shall be taken for smaller magnitudes, we analyze in this section

only the magnitude comparisons for events with $Z \leq 100[km]$.

As in sections 4 and 5 we first compare the magnitudes provided by the reference catalogs. Then, we compare the magnitudes provided by Early-est with the magnitudes listed in the reference catalogs. First we will compare all *best magnitude* (i.e. $mb$ or $M_{wp}$) together, considering only the couple between catalogs where the magnitude types are identical (Figure 4). This comparison will provide a general overview on how the *best magnitude* of Early-est matches with the magnitude of the reference catalogs.

Figure 4 shows the distribution of the magnitude differences $\delta M^{EEc} = M^{EEc} - M^{ref}$ between the values of the EEc and the ones of the reference catalogs.

When comparing the Early-est magnitudes with the magnitudes of the two reference catalogs (center and right panels of figure 4), Early-est seems to overestimate the magnitudes of about $\overline{\delta M^{EEc}} \approx 0.1 \mp 0.2$. The percentiles show that more than $10\%$ of the magnitudes provided by Early-est differ significantly from the magnitude provided by the reference catalogs. The overestimation and the wider distribution appear to be homogeneously distributed among all magnitude ranges.

In the next subsections we will analyze more in details the magnitude values for each single magnitude type $mb$ and $M_{wp}$ separately.

### 6.1 $mb$

In this subsection we compare the $mb^{EE}$ magnitudes provided by Early-est with respect the $mb$ magnitudes provided by Neic ($mb^{Nc}$) and EMSC ($mb^{Cc}$). We use the $mb^{EEc}$ only when Early-Est assigns *best magnitude* $= mb$ following the rules of Table 3.

Figure 5 shows the $mb^{EEc}$ with respect to the $mb^{Nc}$ (top left panel) and with respect to the $mb^{Cc}$ (top right panel). These two plots show scattered and sparse distributed values, which are coherent with the magnitude differences of the histograms in figure 4 c) and figure 4 d). The mean $\overline{\delta mb}$

indicates that the catalogs are coherent, but the standard deviation and the percentiles point out that the $mb^{EEc}$ can be significantly underestimated or overestimated with respect $mb^{Nc}$ and $mb^{Cc}$.

In order to correct such scattered and sparse distributions we computed a linear regression function for each panel (thick dashed lines on the top panels). These functions are computed for $f_1 = mb^{EEc} \rightarrow mb^{Nc}$ and for $f_2 = mb^{EEc} \rightarrow mb^{Cc}$ respectively - the constant $a$ and $b$ of the linear function are showed in the left upper corners of figure 4 a) and figure 4 b). We then applied the regression functions $f_1$ and $f_2$ to the $mb^{EEc}$ values and we recompute the differences (third row of histograms). Both new distributions have mean values close to 0 and smaller standard deviation and percentiles with respect the original ones.

The two functions appear similar but show different $a$ and $b$ constants. In order to test if such differences are significant, we applied the first function $f_1$, derived for $mb^{EEc} \rightarrow mb^{Nc}$, and we computed the differences with respect the $mb^{Cc}$ values. Second we applied function $f_2$ derived for $mb^{EE} \rightarrow mb^{Cc}$ and computed the residuals with respect the $mb^{Nc}$ values. Applying these corrections we obtain two new difference distributions $\delta mb_{Cc}^{EEc \rightarrow Nc}$ and $\delta mb_{Nc}^{EE \rightarrow Cc}$ (bottom left and right panels). The distributions $\delta mb^{EE \rightarrow Nc}$ and the $\delta mb_{Cc}^{EEc \rightarrow Nc}$, and the distributions $\delta mb^{EE \rightarrow Cc}$ and the $\delta mb_{Nc}^{EE \rightarrow Cc}$ as well, appear to be significantly different. We performed a t-test between $\delta mb^{EE \rightarrow Nc}$ and the $\delta mb_{Cc}^{EEc \rightarrow Nc}$ distribution and between $\delta mb_{Cc}^{EEc \rightarrow Nc}$ and $\delta mb_{Nc}^{EEc \rightarrow Cc}$. The null hypotheses $H_o$ is rejected at more than $95\%$.

From the percentiles of the corrected distributions, particularly on the left side, we observe that the regression function $f_1$, when applied, produces a narrower magnitude difference distribution with respect the function $f_2$.

Generally, after applying the linear corrections, the resulting $mb^{EE}$ uncertainty ($\nu \approx 0.00 \mp 0.14$) with respect the reference catalogs is coherent with the overall magnitude uncertainty between the two reference catalogs (figure 4, left panel).

### 6.2 $M_{wp}$

As a reference, we first compare the magnitudes $M_{wp}^{Pc}$ values provided by the Pacific Tsunami Warning Center (PTWC) using the correction of Whitmore et al. (2002) with the $M_w^{CMTc}$ of the CMT-Harvard catalog (figure 6). The magnitudes compare well with a mean difference $\mu = 0.04 \mp 0.19$ for events with magnitude about $M_{wp} \lesssim 7.0 - 7.5$. For larger events, the magnitudes $M_{wp}^{Pc}$ begin to overestimate with respect to the $M_w^{CMTc}$.

We now compare the magnitudes $M_{wp}^{EEc}$ with the $M_w^{CMTc}$ (figure 7). The $M_{wp}^{EEc}$ magnitudes appear to be significantly overestimated ( $> 0.2$ magnitude unit), for earthquakes with $M_w^{CMT} \leq 6.5$.

$M_{wp}$ is based on the far-field approximation to the P-wave displacement due to a double couple point source (Tsuboi et al., 1995), thus we should consider that $M_{wp}$ computed in the near field may result biased. In fact Hirshorn et al. (2012) showed that single station $M_{wp}$ values measured at stations at epicentral distances $\Delta \leq 15^o$ have positive residuals with respect the Harvard centroid moment tensor $M_w$. Nevertheless, our procedure is built to obtain reliable $M_{wp}$ estimates as fast as possible, thus we aim to also use $M_{wp}$ measured from stations close to the epicenter.

To test if our $M_{wp}^{EEc}$ values may be dependent as a function of the distance between station and epicenter, we plotted the station residuals at each station for each event with respect the epicenter distance (Figure 8). Station residuals are defined as $\delta M_{wp}^i = M_{wp}^{EEc,i} - M_w^{CMTc}$, where $i$ indicate the $M_{wp}$ values measured at each station.

Figure 8 top left shows the residuals $\delta M_{wp}^i$ (grey dots) for all events with hypocenter depth $\leq 100[km]$ plotted with respect the epicentral distance in degrees. From these residuals we compute the regression function (dashed line in figure 8):

$$f(\Delta) = -1.32e^{-6} \cdot \Delta^3 + 2.40e^{-4} \cdot \Delta^2 - 0.0146 \cdot \Delta + 0.314 \quad (1)$$

Figure 8 and equation 1 show, that the $\delta M_{wp}^{i}$ are overestimated for distances $\Delta \lesssim 30^{o}$ and slightly underestimate for distances $\Delta \gtrsim 90^{o}$. After applying the regression function $f(\Delta)$ to the station values, the distance dependency of $M_{wp}^{i}$ is removed (Figure 8 top right panel).

The distance dependency of the measured $M_{wp}^{EEc,i}$ at each station results in a general overestimation of the $M_{wp}^{EEc}$ with respect the $M_{w}^{CMTc}$ (Figure 7 bottom left). The overestimation of $M_{wp}^{EEc}$ could of course be removed using only $M_{wp}$ measured at stations with epicentral distance $30^{o} \leq \Delta \leq 90^{o}$. Nevertheless Early-est is designed to provide automatic magnitude estimations within few minutes after event origin time in order to disseminate early tsunami warnings. Thus the closer stations are relevant and must be used.

For this reason we apply the equation 1 to remove the distance dependency of the measured $M_{wp}^{EEc,i}$ and we then recompute the magnitude events $M_{wp,corr}^{EEc}$ . To recompute the $M_{wp,corr}^{EEc}$ we follow the Early-est procedure: we trim off stations with $M_{wp}^{EEc,i} < 10^{th}$ percentile and with $M_{wp}^{EEc,i} > 10^{th}$ percentile. The event magnitude is $M_{wp} = 50^{th}$ percentile of the remaining values. The histogram of Figure 8 bottom right shows the corrected magnitude differences $\delta M_{wp,corr}^{EEc}$. The right shift of the original magnitude differences distribution (Figure 8 bottom left) is corrected. The resulting magnitude $M_{wp}^{EEc}$ uncertainty with respect to the $M_{w}^{CMTc}$ is $\delta M_{wp} = 0.0 \mp 0.2$, which is consistent with the uncertainty of the $M_{wp}$ provided by the with PTWC with respect the global CMT-Harvard catalog. .

## 7 Speed performance and tsunami warning alert timeline

In the previous section we analyzed the final epicenter location, hypocenter depth and magnitude values provided by Early-est, i.e. the values obtained about 20 minutes to one hour after event origin time. A tsunami alert however, is meaningful when delivered within a short time after event origin time and with reliable earthquake source parameters. In order to plan the timeline procedure at the CAT-INGV, we want to know how fast the earthquake source parameters computed by Early-est converge toward a stable values.

We thus first analyze how fast Early-est provides a first automatic location, and second how fast the epicenter coordinates and the magnitudes stabilize toward the stable values.

The histogram in figure 9 shows the delay time after event origin time when a first automatic location of Early-est becomes available. We generally have to wait at least two minutes in order to have a first automatic solution; within 7 and 10 minutes after event origin time about $95\%$ and $100\%$ respectively of all earthquakes are located. At global scale a large number of earthquakes are located along the oceanic ridges and trenches, which are far away from most of the seismic stations. In the Mediterranean area the distances between earthquake source and seismic stations are generally shorter than at global scale. Table 1 lists the 12 events with magnitude $M \geq 5.5$ that occurred in the Mediterranean area between March 2012 and the end of December 2014. These 12 events do not form a reliable statistic, but from Table 1 we may reasonably expect to locate an event in the Mediterranean area with magnitude $M \geq 5.5$ within 2-3 minutes after event origin time.

Figure 10 shows how fast a first location (top panel) and magnitude (bottom panel) stabilizes towards the final and stable values.

Both panels indicate that for most of the events the epicenter coordinates and magnitudes within the first 8-10 minutes after the first available location may be considered stable and significantly close to the final values, since the magnitudes are $\mu+\sigma \leq 0.2$ and the epicenter locations are $\mu+\sigma \leq 10[km]$ respectively.

The CAT-INGV uses the earthquake source parameters provided by Early-est to compile the tsunami warning message to be disseminated to the civil authorities. The mission of the CAT is to provide tsunami warnings for earthquakes with $M \geq 5.5$ which occur in the Mediterranean reagion according to the ICG/NEAMTWS guidelines.

Based on the speed performances of Early-Est on computing reliable earthquake source parameters (figure 10) and on the minimum delay time after event origin time to localize an event in the Mediterranean (table 1), we planed a timeline and actions that allow the seismologist at the CAT-

INGV to verify and distribute reliable tsunami warning messages within the very short but reasonable save time interval after event origin time.

Based on the figure 10 we decided to automatically compile a tsunami warning alert message always for the $2^{nd}$, the $5^{th}$ and the $8^{th}$ locations available after the first location. Considering that the first location in the Mediterranean area may be available within 2-3 minutes after event origin time, the $2^{th}$, the $5^{th}$ and the $8^{th}$ locations may be available between about 5, 8 and 11 minutes after event origin time. Therefore, in case of an earthquake in the Mediterranean area, the continuos monitoring of Early-est provides information to the seismologists for issuing tsunami warnings. Based on figure 10 and table 1, such procedure may be executed within about $\approx 15$ minutes after event origin time. The messages are delivered via fax, Thus the messages reach the authorities within seconds to a few minutes after sending.

## 8 Discussions and final remarks

Early-est is able to provide first location within about 7 minutes from origin time for almost $95\%$ of all worldwide earthquakes. In the Mediterranean area, where the epicentral distance between earthquake and seismic station is smaller, we may expect a first automatic location within 2-3 minutes after event origin time. Generally within less than 10 minutes after the first location, the estimations converge to a stable values.

From our analysis the automatic locations and source depth estimates provided by Early-est for global $M \geq 5.5$ earthquakes are robust and reliable; in fact the epicenter source parameters estimates by Early-est are coherent with the epicenter source parameters provided after manual revision/validation by other agencies (NEIC, GFZ and CSEM-EMSC) that locate earthquakes at global scale.

Generally our analysis showed that earthquakes with $M \geq 5.5$ can be located, when using seismic data form global networks, with a empirical uncertainty, defined as the mean location difference with respect the reference catalogs, of about $\nu \approx 20 \mp 25[km]$. The location provided by Early-est show

differences with respect the locations of the reference catalogs, that are comparable to the location differences between the reference catalogs ($\nu \approx 20 \mp 25[km]$).

A similar conclusion is valid for the mean Early-est focal depth difference for global $M \geq 5.5$ earthquakes is about $\nu \approx 0 \mp 25[km]$, which is also coherent with the focal depth differences between the reference catalogs.

Early-est uses only a sub set of all worldwide, public, real-time stations, and sometimes the available number of stations may be reduced because of latencies, does not seems to affect the quality of the estimated epicenter coordinates and hypocenter depth.

The magnitude is a key earthquake parameter to determine the tsunami alert level (see section 1). The decision matrix defined by the NEAMTWS (2011) sets the tsunami warning level on the basis of the magnitude, hypocenter depth and of the distance between the epicenter and the coastal forecast points. The automatic magnitudes $mb$ and $M_{wp}$ provided by Early-est show differences with respect the used reference values that in some cases may be significant in the context of the tsunami warning.

The magnitudes $mb$ provided by Early-est compare well with the $mb$ values provided by reference agencies from the point of view of the mean differences, but show sparse and scattered distributions that can be larger than $\mp 0.3$ units of magnitude. Such sparse distribution can be corrected by increasing the signal-to-noise ratio threshold for the $mb$ station values. On the other hand a higher signal-to-noise ratio threshold may reduce the number of station readings, and would require more stations to obtain a reliable $mb$ value. This would result into a slower magnitude estimation, which may affect the efficiency and the speed required for tsunami warnings dissemination. A linear correction of the computed $mb$ values produces indeed a reduction of the standard deviation to about $\mp 0.15$ units of magnitude. Both corrections $f_1$ and $f_2$ allow to avoid large magnitude over- and underestimations. The correction function $f_1$ shows slightly more narrow distributions than correction function $f_2$.

Nevertheless the magnitude $mb$ starts to saturate from magnitude $mb \gtrsim 6.0$ and for this reason Early-est does not use $mb$ when $M_{wp} \geq 5.8$. Thus, $mb$ values apply to earthquakes which are not generally expected to be tsunamigenic.

The Early-est magnitude $M_{wp}$ values are reliable when computed using only stations with epicentral distance $30^o \leq \Delta \leq 90^o$. As expected (Tsuboi et al., 1995; Hirshorn et al., 2012), single stations $M_{wp}^i$ measurements at distance $\Delta \leq 30^o$ are significantly overestimated (Figure 8). The observed distance dependent bias at each station results in a general overestimation of the final $M_{wp}$ (Figure 7). Early-est is designed to provide automatic magnitude estimation within few minutes after event origin time in order to disseminate early tsunami warning, thus the closer stations are relevant and must be used. For this reason we prefer to correct the station $M_{wp}$ values to remove the overestimation of the single station $M_{wp}$ values at istance $\Delta \leq 30^o$, instead of introducing a minimum distance cut off.

Since the assignment rules for the *best magnitude* depends on the number of station measuring reliable $mb$, $M_{wp}$ and on $M_{wpd}$ and the magnitude value for each ones (table 3) , the assigned *best magnitude* may vary between $mb$, $M_{wp}$ and $M_{wpd}$ at each run. This is particularly true within the first minutes after event origin time, when the number of available waveforms may still be small, and the magnitudes values may not be stable yet (figure 10). The linear correction for $mb$ and the distance dependent correction for $M_{wp}$ will thus produce a stable and reliable *best magnitude* useful for seismologically based tsunami early warning procedures.

The CAT-INGV provides seismologically based tsunami early warning when earthquakes with magnitude $M \geq 5.5$ occurs in the Mediterranean area. Such tsunami warning messages are based on the fully automatically location and magnitude estimations provided by the Early-est software. The analysis of a data-set of three years of worldwide earthquakes, showed that Early-est is a robust, reliable and efficient software for automatic real-time earthquake source parameter estimation, which provides reliable and robust location parameters and magnitude estimations within few minutes after event origin time.

*Acknowledgements.* We thank the two referees P. Roudil and F. Haslinger for their review that help us to improve the paper. The magnitude $M_{wp}$ parameters of the Pc used in this paper were provided to the authors of this paper by the courtesy of Barry Hirshorn by the Pacific Tsunami Warning Center. We used broadband seismograms recorded by the Global Seismic Network obtained from the IRIS DMC and from NEIC. This work has been funded by the Italian Flagship Project RITMARE, by the EU FP7 project NERA (262330), and by project AS- TARTE (Assessment, Strategy And Risk Reduction for Tsunamis in Europe) FP7-ENV2013 6.4-3, grant 603839. The Early-est software was is fruiting of Italian DPC attachment B2. Figures are produced with GMT (Wessel and Smith, 1995) and *python matplotlib*.

## References


Basili, R., Tiberti, M. M., Kastelic, V., Romano, F., Piatanesi, A., Selva, J., and Lorito, S.: Integrating geologic fault data into tsunami hazard studies, Nat. Hazards Earth Syst. Sci., 13, 1025–1050, doi:10.5194/nhess-13-1025-2013, 2013.

Bormann, P. and Saul, J.: The new IASPEI standard broadband magnitude $mb$, Seism. Res. Lett., 79, 698–705, doi:10.1785/gssrl.79.5.698, 2008.

Bormann, P. and Saul, J.: Earthquake magnitude, in Encyclopedia of Complexity and Systems Science, ed. Meyers, A., Springer, New York, doi:10.1007/978-0-387-30440-3_151, 2009.

Bungum, H. and Lindholm, C.: Tsunamigenic Seismic Sources in the North Sea, the Norwegian Continental Margin and the Norwegian-Greenland Sea, Tech. rep., NORSAR, 2007.

Dziewonski, A. M., Chou, T.-A., and Woodhouse, J. H.: Determination of earthquake source parameters from waveform data for studies of global and regional seismicity, J. geophys. Res., 86, 2825–2852, 1981.

Ekström, G., Nettles, M., and Dziewonski, A. M.: The global CMT project 2004-2010: Centroid-moment tensors for 13,017 earthquakes, Phys. Earth planet. Int., 200, 1–9, doi:doi:10.1016/j.pepi.2012.04.002, 2012.

Godey, S., Bossu, R., Guilbert, J., and Mazet-Roux, G.: The Euro-Mediterranean Bulletin: A comprehensive seismological Bulletin at regional scale, Seism. Res. Lett., 77, 460–474, 2007.

Hardebeck, J. L. and Shearer, P. M.: A new method for determining first-motion focal mechanisms, Bull. Seism. Soc. Am., 92, 2264–2276, 2002.

Hirshorn, B., Weinstein, S., and Tsuboi, S.: On the Application of $M_{wp}$ in the Near Field and the March 11, 2011 Tohoku Earthquake, Pure and Applied Geophysics, 170, 975–991, doi:10.1007/s00024-012-0495-3, 2012.

Jenny, S., Goes, S., Giardini, D., and Kahle, H.-G.: Earthquake recurrence parameters from seismic and geodetic strain rates in the eastern Mediterranean, Geophys. J. Int., 157, doi:10.1111/j.1365–246X.2004.02 261.x, 2004.

Kagan, Y. Y. and Jackson, D. D.: ohoku Earthquake: A Surprise?, Bull. Seism. Soc. Am., 103, 1181–1194, doi:10.1785/0120120110, 2013.

Kennett, B. L. N., Engdahl, E. R., and Buland, R.: Constraints on seismic velocities in the Earth from travel times, Geophys. J. Int., 122, 108–124, 1995.

Koravos, G. C., Main, I. G., Tsapanos, T. M., and Musson, R. M. W.: Maximum earthquake magnitudes in the Aegean area constrained by tectonic moment release rates, GJI, 152, 94–112, 2003.

Lomax, A. and Michelini, A.: $M_{wpd}$: A Duration-Amplitude Procedure for Rapid Determination of Earthquake Magnitude and Tsunamigenic Potential from P Waveforms, Geophys. J. Int., 176, 200–214, doi:10.1111/j.1365-246X.2008.03974.x, 2009a.

Lomax, A. and Michelini, A.: Tsunami early warning using earthquake rupture duration, Geophys. Res. Lett., 36, L09 306, doi:10.1029/2009GL037223, 2009b.

Lomax, A. and Michelini, A.: Tsunami early warning using earthquake rupture duration and P-wave dominant period: the importance of length and depth of faulting, Geophys. J. Int., 185, 283–291, doi:10.1111/j.1365-246X.2010.04916.x, 2011.

Lomax, A. and Michelini, A.: Tsunami early warning within 5 minutes, Pure and Applied Geophysics, 170, 1385–1395, doi:10.1007/s00024-012-0512-6, 2012.

Lomax, A., , and Curtis, A.: Fast, probabilistic earthquake location in 3D models using oct-tree importance sampling, Geophys. Res. Abstr., 3, 955, 2001.

Lomax, A., Michelini, A., and Curtis, A.: Earthquake Location, Direct, Global-SearchMethods, Encyclopedia of Complexity and Systems Science, part. 5, pp. 2449–2473, doi:10.1007/978-0-387-30440-3_150, 2009.

Lorito, S., Selva, J., Basili, R., Tiberti, M. M., and Piatanesi, A.: Probabilistic hazard for seismically induced tsunamis: accuracy and feasibility of inundation maps, Geophys. J. Int., 200, 574–588, doi:10.1093/gji/ggu408, 2015.

Maramai, A., Brizuela, B., and Graziani, L.: The Euro-Mediterranean Tsunami Catalogue, Annals of Geophysics, 57, S0435, doi:10.4401/ag-6437, 2014.

NEAMTWS: Interim Operational Users Guide for the Tsunami Early Warning and Mitigation System in the North-eastern Atlantic, the Mediterranean and Connected Seas (NEAMTWS). Version 2.0, ICG/NEAMTWS-VIII, 2011.

Tsuboi, S., Takano, K. A., and Yamanaka, Y.: Rapid determination of $M_w$ from broadband P waveforms, Bull. Seism. Soc. Am., 85, 606–613, 1995.

Tsuboi, S., Takano, K. A., and Yamanaka, Y.: Application of $M_{wp}$ to deep and teleseismic earthquakes, Bull. Seism. Soc. Am., 89, 1345–1351, 1999.

Vassallo, M., Satriano, C., and Lomax, A.: Automatic picker developments and optimization: A strategy for improving the performances of automatic phase pickers, Seism. Res. Lett., 83, 541–554, doi:10.1785/gssrl.83.3.541, 2012.

Wessel, P. and Smith, W. H. F.: New version of the generic mapping tool released, Trans. Am. Geophys. Union, p. 329, 1995.

Whitmore, P. M., Tsuboi, S., Hirshorn, B., and Sokolowski, T. J.: Magnitude-dependent correction for Mwp, Science of Tsunami Hazards, Science of Tsunami Hazards, 20, 187–192, 2002.

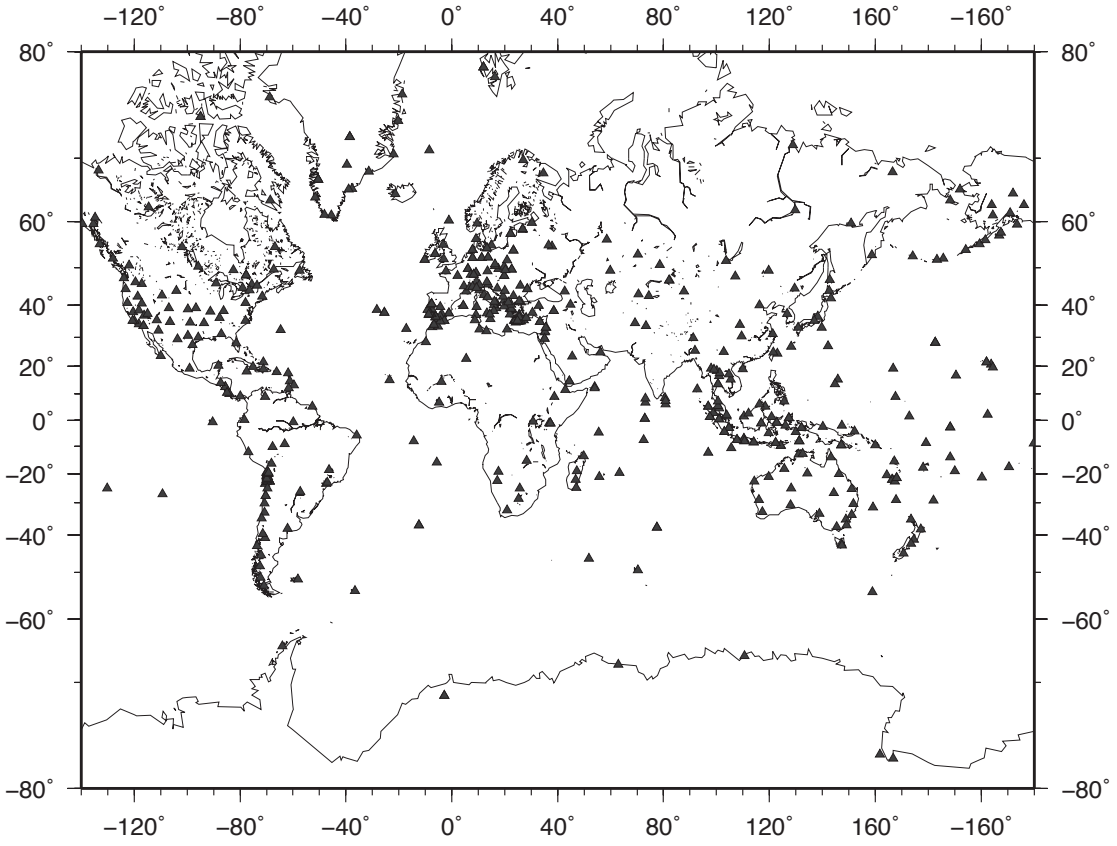



Figure 1: Global map with the 494 seismic broadband stations used by Early-est. The list is updated at the end of September 2014. The stations belong to the following networks: AK, AT, AU, BK, BL, CH, CI, CN, CU, CX, CZ, DK, FR, GB, GE, GT, HL, HT, IC, II, IM, IN, IP, IU, IV, IW, JP, KZ, LB, LX, MN, MS, MY, ND, NN, NO, NZ, PL, PM, PS, TM, TT, US, UW, WM. The network codes are assigned by the International Federation of Digital Seismograph Networks (FDSN) archive. When working in the real-time, latencies in the data stream and/or connection problems may occur, reducing the number of waveform available for location and magnitude estimation.

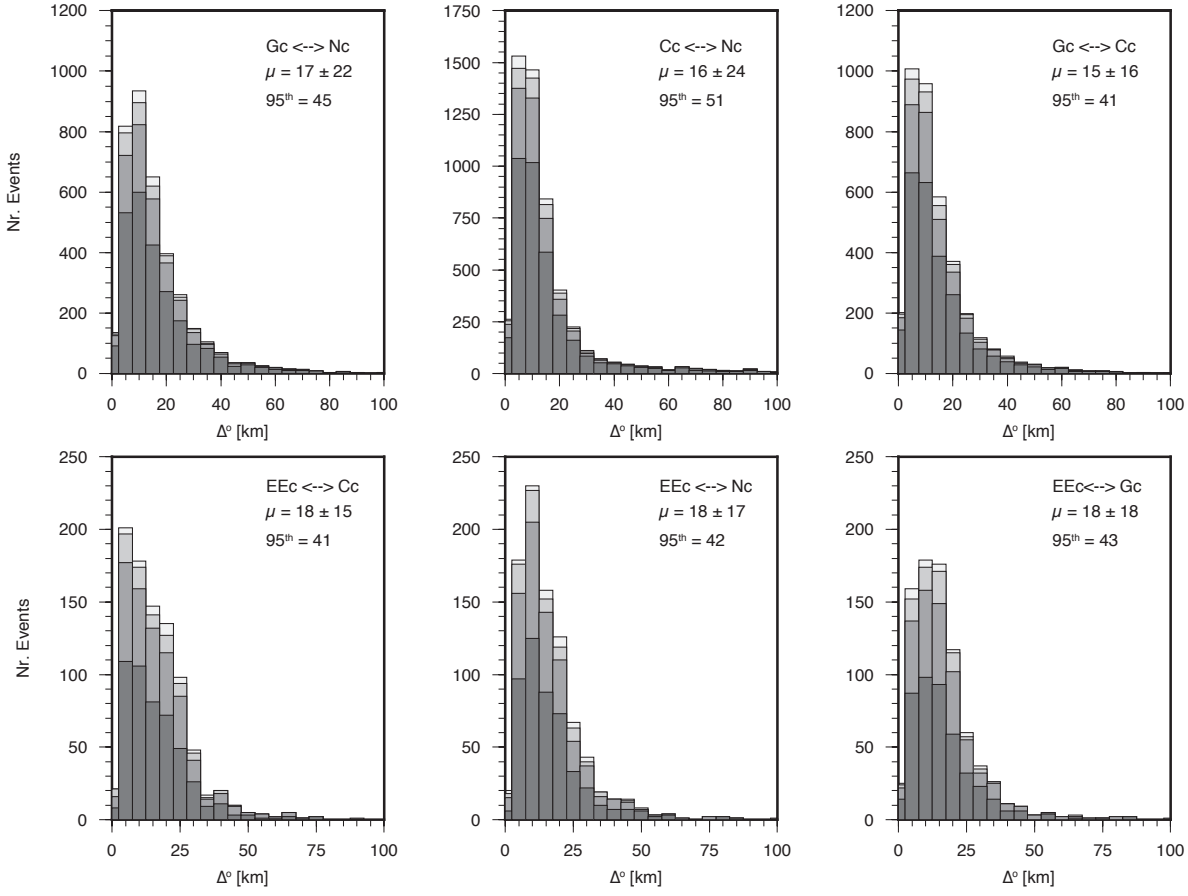



Figure 2: Epicenter location difference distributions for the events listed in the reference and in the Early-est catalogs. The epicenter location difference is expressed in kilometers on the x-axis; the vertical axis refers to the number of events for each bin; the bins are $5[km]$ each. The top panels show the location difference between the locations of the three reference catalogs Nc, Gc, and Cc . The bottom panels show the location difference between Early-est and the reference catalogs. The gray color scale and magnitude ranges: dark grey $5.5 \leq M < 6$, middle dark grey $6.0 \leq M < 6.5$, middle light grey $6.5 \leq M < 7.0$, light grey. The mean and the standard deviation and the 95% percentiles for the entire dates (i.e. regardless to the magnitude) are indicated on the top right hand of each panel.

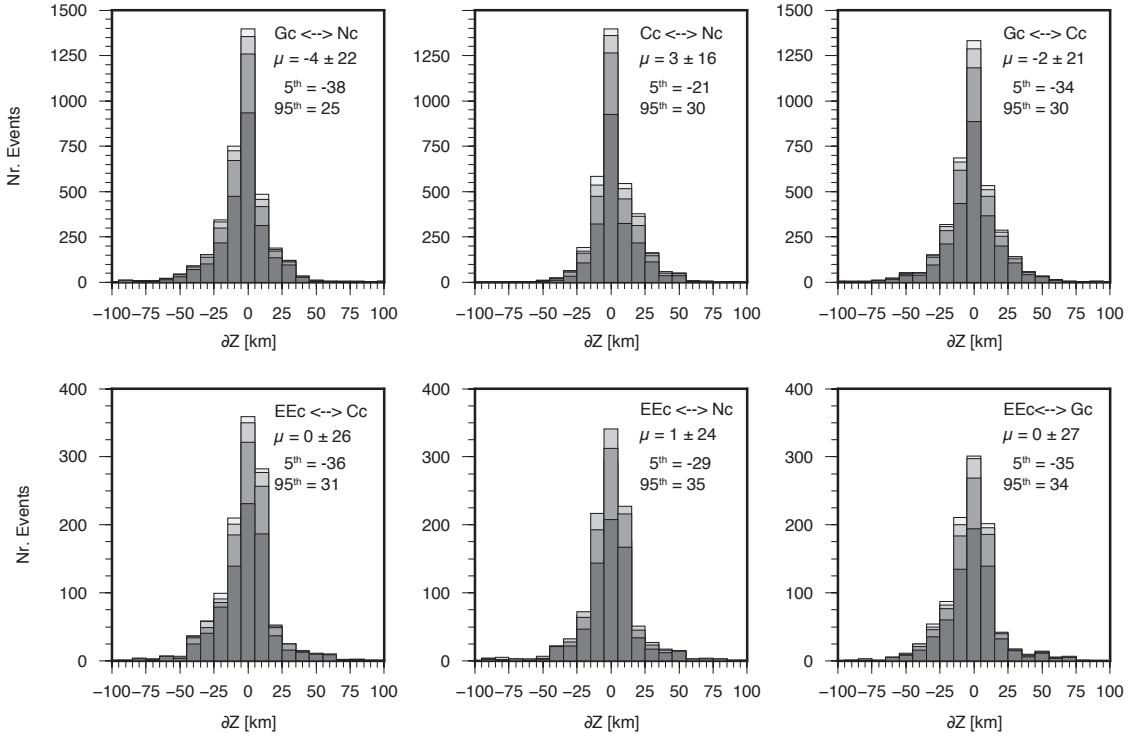



Figure 3: Hypocenter depth difference distributions for the events listed in the reference and in the Early-est catalogs. The hypocenter depth difference is expressed in kilometers on the x-axis; the vertical axis refers to the number of events for each bin; the bins are $5[km]$ each. The top panels show the hypocenter depth difference distribution between the locations of the three reference catalogs Nc, Gc, and Cc . The bottom panels show the hypocenter depth difference between Early-est and the reference catalogs. The gray color scale and magnitude ranges: dark grey $5.5 \leq M < 6$, middle dark grey $6.0 \leq M < 6.5$, middle light grey $6.5 \leq M < 7.0$, light grey. The mean and the standard deviation and the 95% percentiles for the entire dates (i.e. regardless to the magnitude) are indicated on the top right hand of each panel.

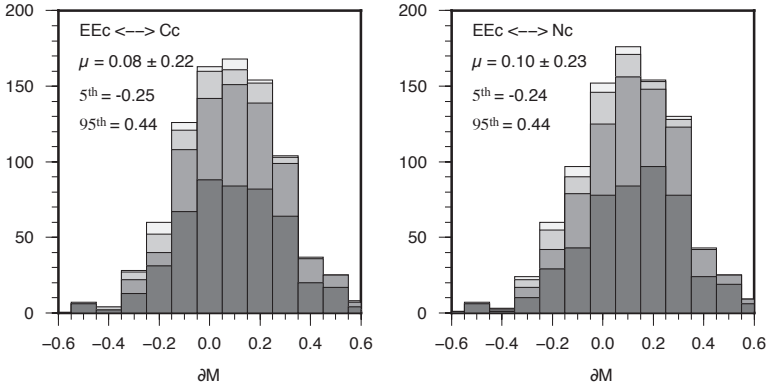



Figure 4: Magnitude difference distributions for the events listed in the EEc catalog with respect the two Ec and Cc reference catalogs. Differences are computed only when the same magnitude type is provided for the same event into the two compared catalogs. The magnitude difference is on the x-axis; the vertical axis refers to the number of events for each bin; the bins are $0.1$ magnitude each. The color scale refers to the same magnitude ranges as in figure 3 and figure 2 and not to the magnitude type. The gray color scale and magnitude ranges: dark grey $5.5 \leq M < 6$, middle dark grey $6.0 \leq M < 6.5$, middle light grey $6.5 \leq M < 7.0$, light grey. The mean and the standard deviation and the 95% percentiles for the entire dates (i.e. regardless to the magnitude) are indicated on the top right hand of each panel.

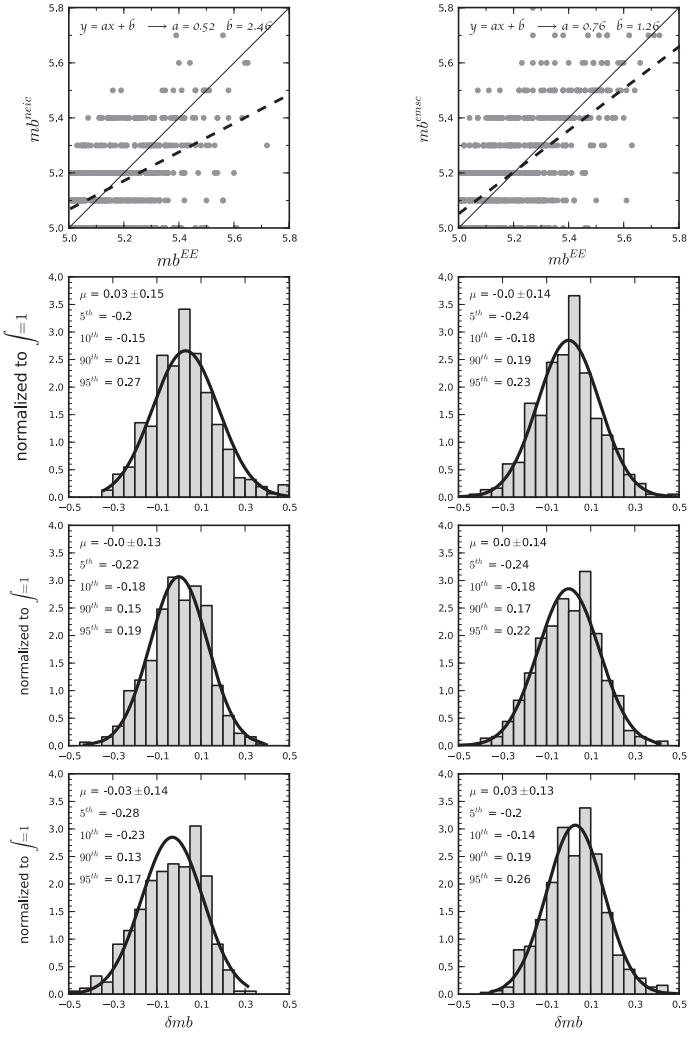



Figure 5: Magnitude $mb$ differences between the Early-est catalog and the reference catalogs (Nc on the left and Cc on the right). Top row panels a) and b): magnitudes $mb$ comparison between the Early-est values (x-axis) and the reference catalog values (y-axis). The dashed lines refer to the linear regression functions; the $a$ and $b$ constant are indicated on the left upper corner; the thin black line refers to the $1:1$ proportion. $2^{nd}$ row panels c) and d): magnitude $mb$ difference distribution; the bins are 0.05 magnitude units wide each. The black line refers to the theoretical distribution derived from measured mean $\mu$ and standard deviation $\sigma$ with $\int = 1$. $3^{nd}$ row panels e) and f): as in $2^{nd}$ row panels but after applying the correction function showed in top panels to the Early-est $mb$. $4^{nd}$ row panels g) and h): as in $3^{nd}$ row panels, but on the left panel apply the EEc-Cc derived correction; on the right panel apply the EEc-Nc derived correction.

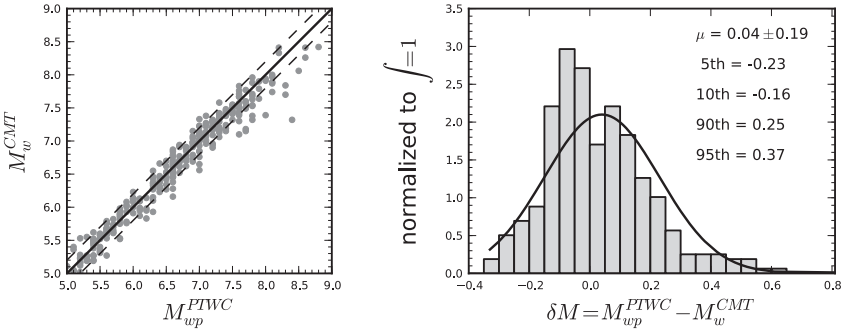



Figure 6: Comparison between the $M_{wp}$ magnitudes computed by the Pacific Tsunami Warning Center (PTWC) with the $M_w$ magnitudes from CMT-Harvard catalog. Plot on the left side: dot: magnitudes values; continuos line: 1:1 ratio; dashed lines: $\mp 0.2$ uncertainty. The histogram on the right side show the $\delta M_{wp} - M_w$ distribution. Mean, standard deviation and percentiles are indicated on the top right hand of the right panel. The bins are 0.05 magnitude wide each.

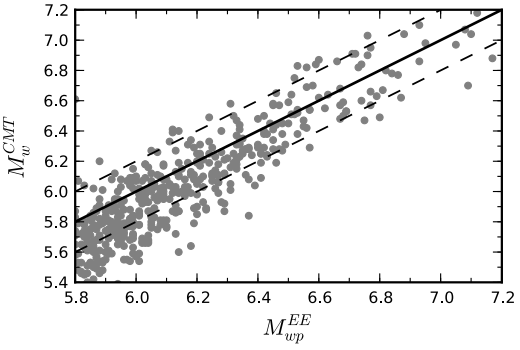



Figure 7: Early-est magnitudes $M_{wp}$ compared with respect to the CMT-Harvard $M_w$ of CMT-Harvard catalog. continuos line: 1:1 ratio; dashed lines: $\mp 0.2$ uncertainty.

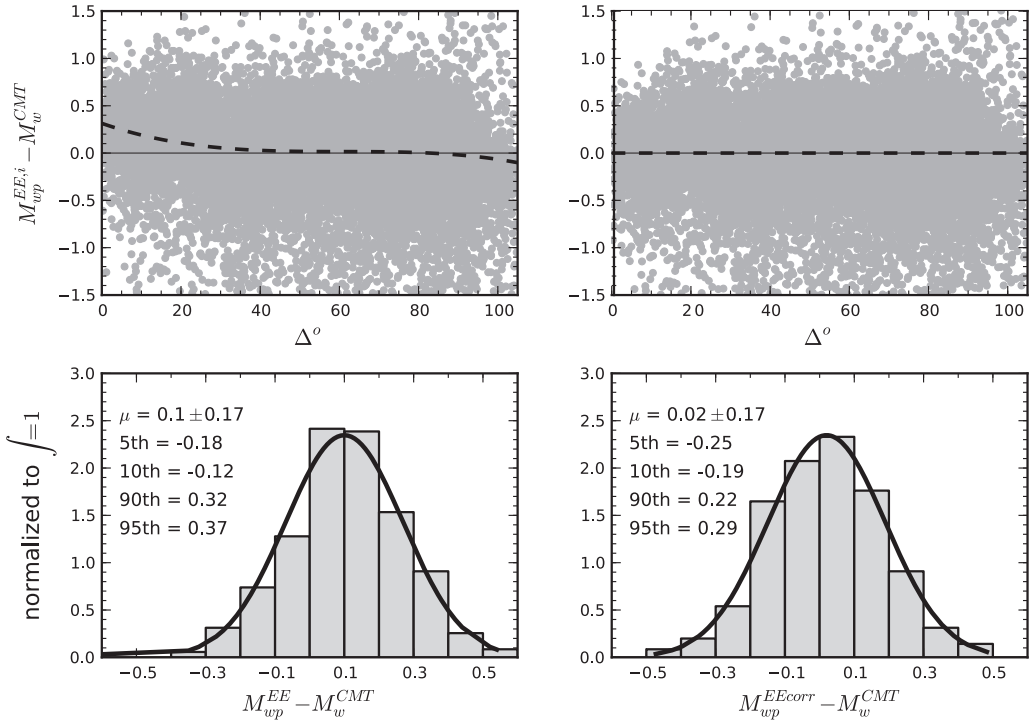



Figure 8: Epicentral distance dependence of the $M_{wp}$ for events with hypocentral depth $\leq 100[km]$. Top left panel: station residuals $\delta M_{wp}^i = M_{wp}^{EE,i} - M_w^{CMT}$ (grey dots) plotted with respect the epicentral distance in degree; the dashed line represents a $3^{th}$ degree polynomial regression function (equation 1), which best fit the data. Top right panel: station residuals$\delta M_{wp}^i = M_{wp}^{EEcorr,i} - M_w^{CMT}$ (grey dots) after applying the regression function (equation 1), plotted with respect the epicentral distance in degree; the dashed line is a $3^{th}$ degree polynomial regression function, which fest fit the corrected residuals with respect the distance. Bottom left panel: event magnitude difference $\Delta M_{wp}$ distribution before the distance correction. These distribution reflect Figure 7; mean, standard deviation and percentiles are indicated on the left of the histogram; bins are 0.5 magnitude wide each; the black solid line refers to theoretical distribution with $\int = 1$. Bottom right panel: event magnitude difference $\Delta M_{wp}^{corr}$ distribution after the distance correction using equation 1.

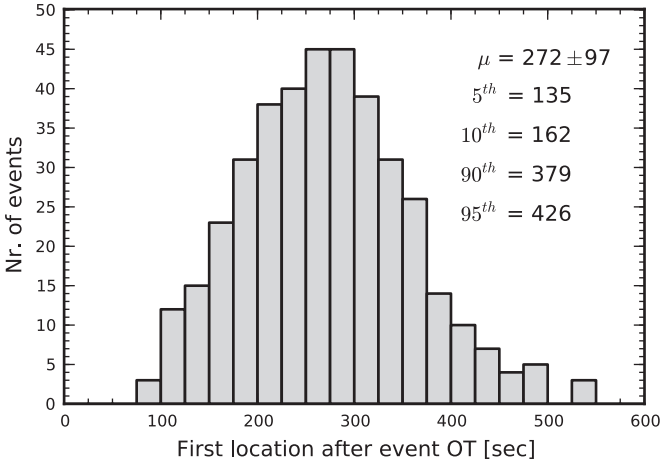



Figure 9: Early-est first location performance. This figure shows how fast a first location for global events is available through Early-est. The bins (25 seconds wide) on the x-axis refers to the seconds after event origin time at which a first location is available. On the right top panel the mean, the standard deviation and 4 representative percentiles are indicated.

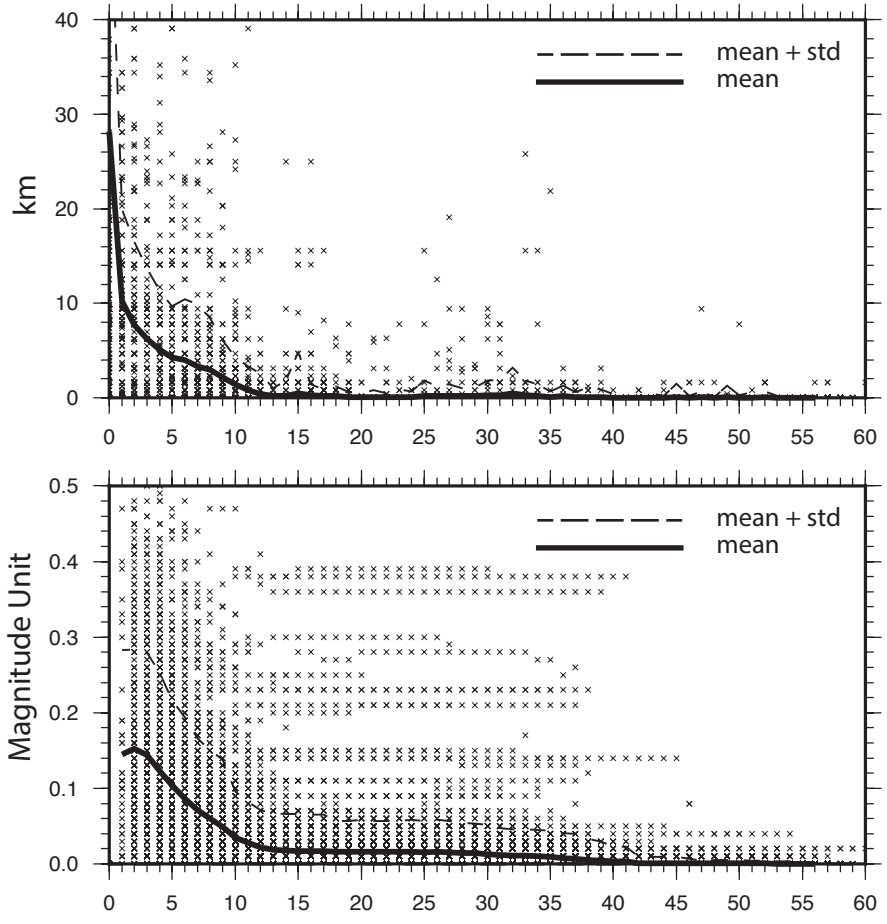



Figure 10: Early-est location and magnitude estimation stability performances. This figure shows how fast a first location (top panel) and magnitude (bottom panel) estimations evolves towards stable values. Top panel: for each run we compute the distance in kilometers between the current epicenter and the epicenter of the last location. Bottom panel: for each run we compute the absolute magnitude difference between the current magnitude and the final magnitude. In this panel, most of the magnitudes are available 2 minutes after event origin time, since often the first automatic location may not provide a magnitude value.The magnitude refers to the 'best' magnitude decided by Early-est (Table 3) at each run. In both panels difference values (black cross) are plotted on the y-axis with respect the minutes after the first location (0 value at the x-axis). The black line is the mean value computed for each minute and the dashed line the mean plus the standard deviation.

Table 1: List of earthquakes occurred in the Mediterranean area located by Early-est with $M \geq 5.5$ between March 2012 and December 2014. For each event we listed the computed event origin time, epicenter coordinates, hypocenter depth, the maximum $68\%$ confidence error in $xyz$-space in kilometers, the preferred magnitude ($mb$, $M_{wp}$ or $M_{wpd}$), a reference magnitude, when the first Early-est location were available in seconds after the event origin time, and when the magnitude stabilize in minutes after the first location available. The magnitude is stable when the difference with respect the final magnitude is $\leq \mp 0.2$.

| Nr | Date | Time | lat. | lon. | Depth | $\delta(xyz)$ | $Mag^{best}$ | $Mag^{ref}$ | First loc. | First mag. |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 2012-06-10 | 12:44:15 | 36.36 | 28.93 | 19.7 | 4.3 | $M_{wp} = 6.1$ | $M_w^{CMT} = 6.1$ | 167 | 10 |
| 2 | 2012-09-12 | 03:27:43 | 34.77 | 24.08 | 10.0 | 5.1 | $mb = 5.7$ | $mb^{Nc} = 5.4$ | 201 | 7 |
| 3 | 2013-01-08 | 14:16:09 | 39.62 | 25.49 | 10.1 | 4.2 | $M_{wp} = 5.7$ | $M_w^{CMT} = 5.7$ | 174 | 3 |
| 4 | 2013-06-15 | 16:11:02 | 34.51 | 24.99 | 15.4 | 5.4 | $M_{wp} = 6.4$ | $M_w^{CMT} = 6.3$ | 181 | 2 |
| 5 | 2013-06-16 | 21:39:07 | 34.51 | 25.00 | 18.6 | 4.8 | $M_{wp} = 6.1$ | $M_w^{CMT} = 6.0$ | 117 | 3 |
| 6 | 2013-10-12 | 13:11:51 | 35.52 | 23.30 | 11.5 | 5.2 | $M_{wp} = 6.6$ | $M_w^{CMT} = 6.8$ | 194 | 2 |
| 7 | 2013-12-28 | 15:21:06 | 36.04 | 31.30 | 56.8 | 8.5 | $M_{wp} = 6.0$ | $M_w^{CMT} = 5.9$ | 358 | 5 |
| 8 | 2014-01-26 | 18:45:10 | 38.29 | 20.38 | 19.8 | 2.5 | $mb = 5.2$ | $M_w^{Nc} = 5.4$ | 115 | 3 |
| 9 | 2014-02-03 | 03:08:46 | 38.25 | 20.40 | 10.1 | 2.3 | $M_{wp} = 6.1$ | $M_w^{CMT} = 6.0$ | 77 | 7 |
| 10 | 2014-04-04 | 20:08:07 | 37.26 | 23.71 | 115.9 | 2.2 | $mb = 5.5$ | $M_w^{CMT} = 5.6$ | 119 | 6 |
| 11 | 2014-05-24 | 09:25:03 | 40.23 | 25.34 | 10.1 | 4.8 | $M_{wp} = 6.6$ | $M_w^{CMT} = 6.9$ | 124 | 7 |
| 12 | 2014-08-29 | 03:45:06 | 36.75 | 23.67 | 81.2 | 2.7 | $M_{wp} = 5.8$ | $M_w^{CMT} = 5.8$ | 119 | 4 |

Table 2: Global earthquake catalogs used for the analysis in this work. For each catalog we indicated the begin and end time of the time window of the dataset included into this work. Catalog abbreviation used into this paper is between brackets in the first column.

| Catalog | Begin | End | type |
|---|---|---|---|
| Early-est (EEc) | 03-2012 | 12-2014 | automatic |
| Neic (Nc) | 01-2004 | 12-2014 | revised |
| Gfz (Gc) | 06-2006 | 12-2014 | revised |
| CSEM (Cc) | 10-2004 | 12-2014 | revised |
| PTWC (Pc) | 2013-12-28 | 06-2014 | revised |
| CMT-Harvard (CMT) | 1976 | 010-2014 | revised |

Table 3: This table summarize the rules used by Early-est to define the *best magnitude* (i.e.: the most significative magnitude type) for each earthquake. Each location run Early-est computes $mb$, $M_{wp}$, $M_{wpd}$. The magnitude $mb$ is computed using the $30s$ time window or the apparent source duration $T_o$ as time window if $T_o < 30s$ and the IASPEI WWSSN-SP response for convolution. The magnitude $M_{wp}$ is scaled to the largest of the first two maxima on integrated displacement within the window from $t_P$ to $t_P + T_o$ time or $120s$ after $t_P$, where $t_P$ is the $P$-arrival time, whichever window is the shortest. The magnitude $M_{wpd}$ (*duration-amplitude*), which can be viewed as an extension of the $M_{wp}$ moment-magnitude algorithm, is computed following the $M_{wp}$ procedure and corrections described into Lomax and Michelini (2012).

| Best magnitude | #[1] | Magnitude range[2] |
|---|---|---|
| $M_{wpd}$ | $\geq 6$ | $M_{wp} \geq 7.2$ |
| $M_{wp}$ | $\geq 6$ | $5.8 \leq M_{wp} < 7.2$ |
| $mb$ | $\geq 6$ | $M_{wp} < 5.8$ |

[1]: Number of recording stations with good signal-to-noise ratio and reliable amplitude reading
[2]: Magnitude range validity

**Appendix A: Oct-tree associate/locate module**

The oct-tree associate/locate module (Figure A.2) efficiently and robustly associates picks, and detects and locates seismic events over the whole Earth from 0 to 700km depth using the efficient, non-linearized, probabilistic and global, oct-tree importance-sampling search (Lomax et al., 2001, 2009). The objective function for the oct-tree search is a probability function, $P(\mathbf{x})$, based on stacking of implicit origin-times for each pick for each potential source $\mathbf{x}_{test}$: given a seismic wave velocity model (currently ak135 (Kennett et al., 1995)), a pick time $t_{\mathbf{p}}$ at a seismic station, and assuming a seismic phase type that may have produced the pick, the phase travel-time from source $\mathbf{x}_{test}$ to station $T_{\mathbf{x}}$ can be calculated and thus the implicit origin-time $T_0$ for the source and phase can be determined by back projection (e.g., $T_0 = T_{\mathbf{p}} - T_{\mathbf{x}}$). The set of stacks of $T_0$ for all picks forms a histogram over potential origin-times for a source at $\mathbf{x}_{test}$. If the maximum histogram value exceeds a specified threshold, and if the associated picks for the maximum pass tests on amplitudes and station distributions, then $P(\mathbf{x}_{test})$ is retained to drive further the oct-tree search to find a maximum $\mathbf{x}_{max} = max[P(\mathbf{x})]$ and define a seismic event at $\mathbf{x}_{max}$ and associated picks.

The oct-tree search is *direct* and non-linearized – it does not involve linearization of the equations relating the pick times to the source location, and is global and *probabilistic* – it samples throughout the prior probability density function (PDF) for the seismic location problem. The search uses an initial, coarse, regular grid-search followed by recursive, octal sub-division and sampling of cells in three-dimensional, latitude/longitude/depth space to generate a cascaded, oct-tree structure of sampled cells. The oct-tree search produces approximate *importance-sampling* - the spatial density of sampled cells follows the objective function $P$.

For each latitude/longitude/depth cell of volume v visited by the oct-tree search, a histogram-like stack over implicit origin-times for first-arrival, P phases (currently Pg, P, Pdiff, PKPdf) for all picks in the *pick list* is constructed. Each origin-time value $T_0$ is assigned a distance and pick-quality weighted amplitude A between 0 and 1.0, and an uncertainty $\sigma$ determined by the sum of half the maximum travel-time range across the cell volume with the travel-time and pick uncertainties. Each

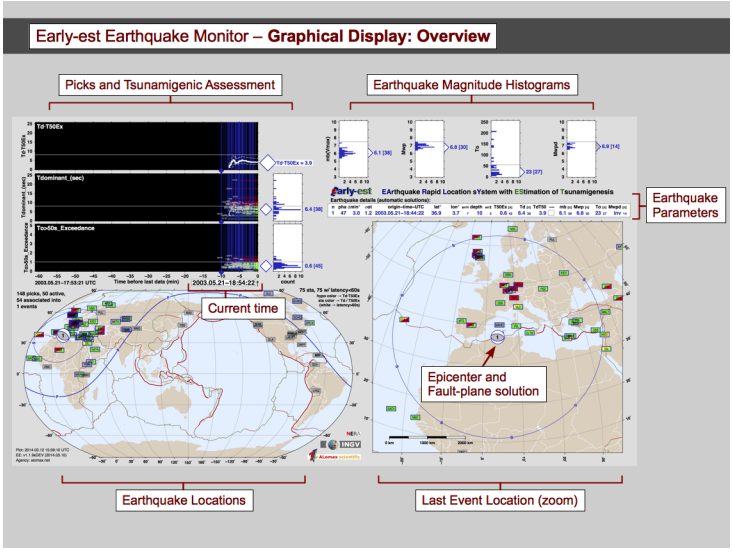
implicit origin-time is included in the origin-time stack with amplitude A using two step-function time-limits at $T_0 \mp \sigma$ inserted in time order. After all picks have been processed, the maximum of the origin-time stack is found by a systematic scan over the available time-limits; the use of step-function time-limits and time ordering makes this scan very fast. All picks whose origin time-limits overlap the stack maximum time are flagged as associated. The stack value, combined with the variance of the implicit origin-times from all associate picks, is converted to a probability, $P(\mathbf{x}, v)$. If the maximum stack value exceeds a specified threshold (currently 4.5), and if the associated picks for the maximum pass tests on amplitude attenuation, and station distance and azimuth distributions, then $P(\mathbf{x}, v)$ is stored for use in the progression of the oct-tree search. If any of these conditions are not met, then the *oct-tree associate/locate* module returns, with a flag that no event has been associated. $P(\mathbf{x}, v)$ represents the relative probability that an event is located within a cell of volume $v$ at position $\mathbf{x}$.

The oct-tree search to associate / locate is paused when the subdivided cells reach an adaptively determined, minimum size (e.g. $\leq 5km$ for a location constrained by regional to globally distributed stations, $\leq 1km$ for a location constrained by locally distributed stations); at this pause uncertainty measures (e.g. PDF scatter samples) on the association stage are generated. The oct-tree search and cell subdivision is then continued for a fixed number of samples (currently about 4600) to obtain a refined, precise location by fixing the associated phases to those corresponding to the maximum of the $P(\mathbf{x}, v)$ found in the association stage. The fixing of the associated phases is necessary for small cell sizes since a decreasing cell volume combined with the step-function limits on origin-time leads to a continuous reduction in $P(\mathbf{x}, v)$ values and eventual instability and non-convergence of the oct-tree search near and at the optimal source location. The precise oct-tree results provide uncertainty measures (e.g. PDF scatter samples, uncertainty ellipsoid) on the location.

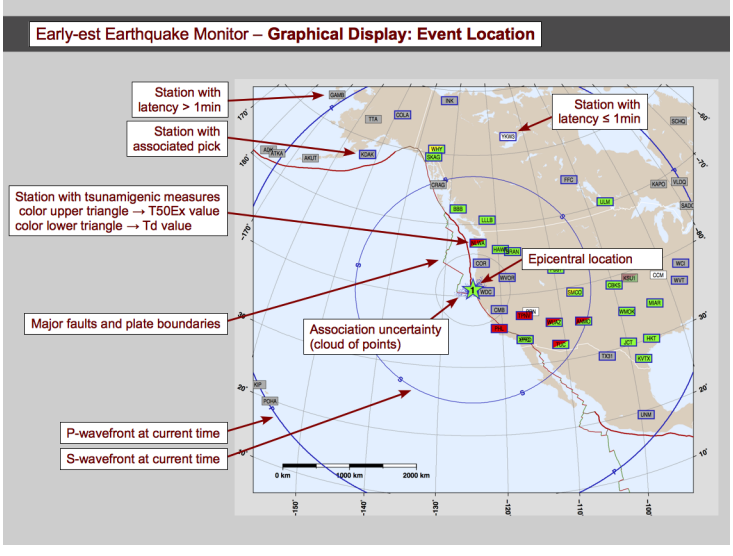
When the *oct-tree associate/locate* module returns an event, the associated picks for this event are masked in the *pick list* and the *oct-tree associate/locate* module is called again using the remaining, non-associated picks, until no further events are returned. Thus multiple events can be associated

and located within a report interval, and, in general, the events are identified in order of the number of associated picks and better location constraint.

Early-est runs the *oct-tree associate/locate* module every 1 minute using all picks from the past hour, without knowledge of or preserving information from previously associations and event locations. This procedure makes Early-est relatively simple algorithmically and robust with regards to changes in the set of available picks and the number of associated picks defining locations. In particular, this procedure allows early stage locations with few associated picks to easily move in space or origin time, or to split into multiple events, or to be absorbed into other events, or to disappear as more pick data becomes available. However, this procedure is inefficient for later stage event locations which are defined by a larger number of associated picks, e.g. more than 10-20 picks, since such locations are very unlikely to change; much processing effort is repeated each minute to re-obtain a previous result. This inefficiency can be problematic after large earthquakes, when the repeated re-processing of hundreds of picks from a mainshock and large aftershock can cause Early-est to fall behind real-time.




(a) Screen overview




(b) Location map

Figure A.1: main graphical display of Early-est

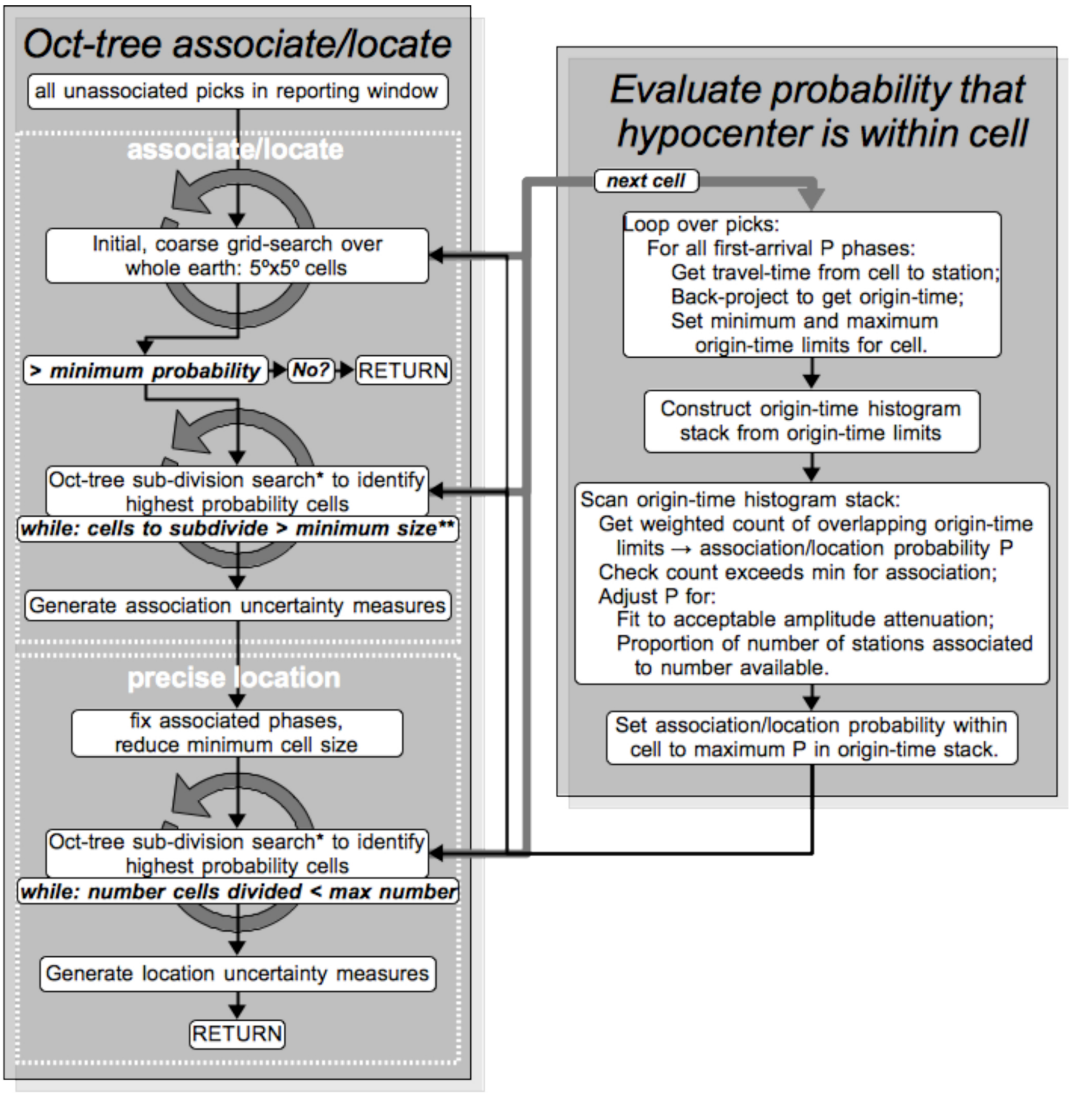



Figure A.2: Early-est Associate/Locate Flow-Diagram:
*Cell division is performed at a fixed cell size for a specified number of cells or until no cell available to divide; the fixed cell size is then reduced and cell division continued.
**minimum size is adaptively reduced in proportion to number of associated stations near epicenter.

Table A.1: Early-est parameter specifications

| *Measure* | *References* | *Description, modifications* |
|---|---|---|
| $Td$ | (Lomax and Michelini, 2011) | Max. dominant period smoothed over $5s$ in window from $T_{\mathbf{P}}$ to $T_{\mathbf{P}} + 55$. |
| $T_{50}Ex$ | (Lomax and Michelini, 2011) | $T_{50}$ Excedance, modified as follow:<br>1. Reduced $T_{50}Ex$ minimum distance to $5^o$ |
| $Td \cdot T_{50}Ex$ | (Lomax and Michelini, 2011) | Period-duration discriminant for tsunami potential, modified as follows:<br>1. Reduced $Td \cdot T_{50}Ex$ minimum distance to $5^o$ |
| $T_o$ | (Lomax and Michelini, 2009a, b, 2011) | High-frequency, apparent source-duration, modified as follows:<br>1. Removed smoothing window width of $10s$ from To for short durations; ,applied with a linear ramp from $10 \rightarrow 0s$ for initial durations of $20 \rightarrow 60s$, minimum duration is highest frequency in HF stream ($0.2s$).<br>2. Reduced $T_o$ minimum distance to $5^o$<br>3. Added reference of $T_o$ duration to $S$ arrival time ($T_s$) if raw duration end time $T_{o,end}$ is after $T_s$ (e.g. if $T_{o,end} > T_s + (T_s - T_p)/3$ then $T_o = T_{o,end} - T_s$). |
| $mb(V_{max})$ | (Bormann and Saul, 2008, 2009) | $mb$ body wave magnitude using $V_{max}$ formulation:<br>1. Apply to BRB velocity a recursive, time-domain filter that implements the WWSSN-SP displacement response<br>2. WWSSN-SP displacement response from Working Group on Magnitudes (Magnitude WG) of the International Association of Seismology and Physics of the Earth's Interior (IASPEI) Commission on Seismological Observation and Interpretation (CoSOI) 2011<br>This filter is applied to the BRB velocity, so effectively gives: integrate $\rightarrow$ simulate the WWSSN-SP response $\rightarrow$ differentiate, without doing the integration and differentiation.<br>1. Measure $V_{max}$ - the peak from $T_p$ to the lesser of $T_p + T_o$ or $T_p + 30s$<br>2. Apply: $mb(V_{max}) = log_{10}(V_{max}/2\pi) + Q(\Delta, h)$ |
| $M_{wp}$ | (Tsuboi et al., 1995, 1999) | $M_{wp}$ magnitude, modified as follows:<br>1. Applied from $T_p$ to the lesser of $T_p + T_o$ or $T_p + 120s$. |
| $M_{wpd}(RT)$ | (Lomax and Michelini, 2009a) | $M_{wpd}$ duration-amplitude, large earthquake magnitude, modified as follows to allow simple and robust real-time application without event type determination:<br>1. Use constant $k = 4.213e19$; PREM depth correction; no geometrical spreading or attenuation corrections.<br>2. Moment correction applied to all event types if $T_o > 80s$<br>3. Moment correction applied to all event types if $T_o > 80s$<br>4. Reduced $M_{wpd}$ minimum distance to $5^o$ (Stable since added reference of $T_o$ duration to $T_s$). |
| Focal mech. | | P-arrival, first-motion focal mechanism using the HASH program. |
| Focal mech. | (Hardebeck and Shearer, 2002) | Probabilistic, P-arrival, first-motion and amplitude focal mechanism algorithm (fmamp). Uses oct-tree search; solution quality based on wighted distribution (quasi-pdf) of P and T axis. (Note: Under development; not included yet in Early-est distribution.) |