Optimal, real-time earthquake location for early warning

Claudio Satriano

RISSC-Lab, Dipartimento di Scienze Fisiche, Università di Napoli Federico II

Anthony Lomax

Anthony Lomax Scientific Software, Mouans-Sartoux, France

Aldo Zollo

RISSC-Lab, Dipartimento di Scienze Fisiche, Università di Napoli Federico II

Abstract

An effective early warning system must be capable of estimating the location and size of a potentially destructive earthquake within a few seconds after the event is first detected.

In this work we present an evolutionary, real-time location technique, based on the equal differential time (EDT) formulation and on a probabilistic approach for hypocenter estimation. The algorithm, at each time step, relies on the information coming from triggered arrivals and not yet triggered station. With just one recorded arrival, hypocentral position can be constrained by the Voronoi cell associated to the first triggering station. As time passes and more triggers become available, the evolutionary location converges to a standard EDT location.

We performed synthetic location tests using the actual geometry of the ISNet (Irpinia Seismic Network, Southern Italy) in order to evaluate the accuracy of the algorithm and its robustness in the presence of outliers.

Introduction

Destructive S and surface waves from a large earthquake can take several tens of seconds to travel from the earthquake source region to distant populated areas and sensitive infrastructure. If there is a seismological monitoring network in the source region, modern seismological analysis methods and communications systems allow characterization of the event and the issuing of alarm messages within seconds, leaving tens of seconds for mitigating actions to be taken. This procedure is known as early warning. For example, for an earthquake in the Irpinia region of Southern Italy there is a delay of about 25-30 sec before the first energetic S wave trains arrive at Naples at about 80-100 km distance. With an early warning system, alarm messages could be sent to critical sites in Naples 20 or more seconds before strong shaking commences.

The characterization of an earthquake includes, most importantly, estimates of its location and size [*Zollo et al.*, this issue]. Here we are concerned with obtaining the most constraint possible on the location of the event hypocenter as time passes after event detection. This constraint is expressed as a probability density function (*pdf*) for the hypocenter location in 3D space. This time-evolving, probabilistic, optimal location information will form a critical part of early warning messages, allowing actions to be taken based on the range of likely source distances and directions as estimated at the time of each message [*Iervolino et al.*, this issue].

There are many approaches to standard earthquake location, which is performed when all the phase arrival times for an event are available. Our optimal, real-time location methodology is based on the equal differential-time formulation (EDT) of *Font et al.* [2004] and *Lomax* [2005] for standard earthquake location. EDT is a generalization of the master-station method [*Zhou*, 1994] and the "method of hyperbolas" cited by *Milne* [1886]. The EDT location is given by the maximum of a stack over quasi-hyperbolic surfaces, on each of which the difference in calculated travel-time to a pair of stations is equal to the difference in observed arrival times for the two stations. The EDT location determination is independent of origin time and reduces to a 3D search over latitude, longitude and depth. Because it uses a stack, EDT is highly robust in the presence of outliers in the data [*Lomax*, 2005]. This robustness is critical for the present problem, since we will often work with small numbers of data and may have outlier data such as false triggers and misidentified picks due to energetic, later phases.

Previous work on earthquake location for early warning includes several novel approaches to gain constraint on the location at an earlier time and with fewer observations than for standard earthquake location. Horiuchi et al. (2005) combine standard L_2–norm event location, EDT location on quasi-hyperbolic surfaces, and the information from not yet arrived data to constrain the event location beginning when there are triggered arrivals from two stations. The two arrivals define a hyperbolic surface on which the event can be located. A largest volume which may contain the hypocenter is bounded by EDT surfaces constructed using the current time (t_{now}) as a substitute for future, unknown arrival times at the stations which have not yet recorded arrivals. This volume shrinks as T_{now} progresses, even if no further stations record an arrival. Rydelek and Pujol (2004) apply the approach of Horiuchi et al. (2005) for the case of only two stations triggered.

Method

We assume that a seismic network has known sets of operational and non-operational stations, that when an earthquake occurs *P* wave arrival picks will become available from some of the operational stations, and that there may be outlier picks which are not due the *P* arrival. Our methodology is related to that of Horiuchi et al. (2005), which we extend and generalize by a) starting the location procedure after only one station has triggered, b) using the EDT approach throughout to incorporate the triggered arrivals and the not yet triggered stations, c) estimating the hypocenter probabilistically as a *pdf* instead of as a point, and d) applying a full, global search for each update of the location estimate.

When a first station S_n triggers with an arrival at $t_n = t_{\text{now}}$, we can already place useful limits on a *pdf* volume that is likely to contain the hypocenter. These limits are given by conditional EDT surfaces on which the *P* travel time to the first triggering station $tt_n(\mathbf{x})$ is equal to the travel-time to each of the operational but not yet triggered stations, $tt_l(\mathbf{x})$, $l \neq n$. In the case of a homogeneous medium, the hypocentral *pdf* volume is the Voronoi cell around the first recording station defined by the perpendicular bisector surfaces with each of the immediate neighbor stations (Figure 1b).

As the current time t_{now} progresses we gain the additional information that the not yet triggered stations can only trigger with $t_l > t_{\text{now}}$. Thus the *pdf* volume is bounded by conditional EDT surfaces that satisfy the inequality $tt_l(\mathbf{x}) - tt_n(\mathbf{x}) < t_{\text{now}} - t_n, \ l \neq n$. This hypocentral *pdf* volume will be smaller than the previous *pdf* volume estimate since the updated, conditional EDT surfaces tend to fold towards and around the first triggered station (Figure 1c).

When the second and later stations trigger, we construct standard, true EDT surfaces between each pair S_l, S_m of the triggered stations using the equality $tt_l(\mathbf{x}) - tt_m(\mathbf{x}) = t_l - t_m, \ l \neq m$. These EDT surfaces are stacked with the volume defined by the not yet triggered stations, as described above, to form the current hypocentral *pdf* volume (Figure 1d-f). In practice, all EDT surfaces are given a finite width by including errors in the arrival time picking and the travel-time calculation.

As more stations trigger, the number of not yet triggered stations becomes small, and the stacked true EDT surfaces and volumes bounded by conditional EDT surfaces converge towards the hypocentral *pdf* volume that is obtained with standard EDT location using the full set of data from all operational stations.

If there are a small number of outlier data, the final hypocentral *pdf* volume will usually give an unbiased estimate of the hypocentral location, as with standard EDT location. However, if one or more of the first arrival times is an outlier, the early estimates of the hypocentral *pdf* volume may be biased. If N_{out} is the number of outlier data, the bias should be significantly reduced after about $4+N_{out}$ arrivals have been obtained, and should be further reduces as the solution converges towards a standard EDT location.

Algorithm

We consider a network of N stations (S_0, …, S_N), a gridded search volume V containing the network, and the travel times from each station to each grid point in V computed for a given velocity model. If S_n is the first station to trigger, we search for grid points (i, j, k) in V where the following system of differential time inequalities is satisfied:

$$(tt_l - tt_n)_{i,j,k} \geq \delta t_{n,l}\,; l \neq n\,, \tag{1}$$

where tt_i is the travel time from the grid point (i, j, k) to the station S_i and δt is the time interval between the arrival time at station Sn and the latest time for which we have information from station S_l:

$$\delta t_{n,l} = t_{\text{now}} - d_k - t_n\,, \tag{2}$$

where t_{now} is the current clock time and d_l is the delay time for receiving information from station S_l.

The system (1) defines the volume where the hypocenter may be located given that, at current time t_{now}, only the station S_n has triggered. In the case of a homogeneous medium and all $\delta t_{n,l} = 0$ (*i.e.*, $t_{\text{now}} = t_n$ and $d_l = 0$), (1) defines the *Voronoi cell* for the station S_n relative to the positions of the other operational stations. For each inequality in (1), we define a value $p_{n,l}$ which is 1 if the inequality is satisfied and 0 if not. Then we sum the $p_{n,l}$ for each station l at each grid point, obtaining a non-normalized probability density $P(i, j, k)$, where $P(i, j, k)=N-1$ for grid points where all the inequalities are satisfied and a value less than $N-1$ elsewhere.

When an additional station triggers, we re-evaluate the system (1) for all pairs of triggered stations S_n and all not yet triggered stations S_l. Next, we search for grid points where the following equation is satisfied:

$$\left|(tt_l - tt_m) - (t_l - t_m)\right|_{i,j,k} \leq \sigma; l \neq m\,, \tag{3}$$

where S_l and S_m are triggered stations and σ gives the uncertainty in the arrival time picking and the travel-time calculation. This is the standard EDT equation.

We define a value $q_{l,m}$ which is 1 if the inequality (3) is satisfied and 0 otherwise. We sum the $q_{l,m}$ with the $p_{n,l}$ obtained from the re-evaluation of (1) to obtain a new $P(i, j, k)$. The maximum value of

P is

$$P_{\max} = (N - n_T)n_T + n_T(n_T - 1)/2, \tag{4}$$

where n_T is the number of stations that have triggered. The first term in (4) counts the number of inequalities from (1) and the second term the number from (3).

Starting from P, we define a value:

$$Q(i,j,k) = \left(\frac{P(i,j,k)}{P_{\max}} \right)^N, \tag{5}$$

which can be taken as the relative probability density (with value between 0 and 1) for the given grid cell to contains the hypocenter.

We calculate an updated value for $Q(i, j, k)$ when a new station triggers or after a predetermined time interval, whichever is earlier. Then, an alarm message can be sent including information on the current constraint on the hypocentral location. This information may include, for example, the grid point where $Q(i, j, k)$ is greatest, or an uncertainty on the hypocentral location given by the largest horizontal and vertical distances between cells where $Q(i, j, k) > \alpha Q$ and α is a constant < 1. For message recipients at specific localities, the hypocentral location and uncertainty message might be provided as the likely epicentral distance range to the locality.

Location tests

In order to evaluate the accuracy and the robustness of the location technique, we conducted several synthetic tests using the geometry of the ISNet (Irpinia Seismic Network, Southern Italy) [*Weber et al.*, this issue] and a 1D Vp model for the region (table 1) with a constant Vp/Vs of 1.68.

Depth (km)	**Vp (km /s)**
0.0	2.0
1.0	3.2
2.5	4.5
15.0	6.2
35.0	7.4
40.0	8.0

Our first test considers a shallow earthquake, occurring at the center of the network at a depth of 1 km. The event is located using only P triggers. Each panel in Figure 2 represents the projection along three orthogonal planes passing through the true hypocenter of the earthquake location probability density $Q(i,j,k)$. The first snapshot is taken when the first station, ST24, triggers (T=0); the constraint on the earthquake location is given by the volume defined by equation (1), there is no constrain on depth. After 1 second, station ST25 triggers; the location is now constrained by the previously defined volume (which has been collapsing around station ST24) and the EDT surface defined by equation (3). After 2 seconds, 4 stations have triggered and the location is already well constrained for early warning purposes.

Figure 3 shows a location performed using only P triggers for an earthquake occurring outside the network at a depth of 10 km. At T=0 the maximum probability volume is bounded only towards the network. After 1 second, two more stations trigger and the volume is bounded in all directions. As time evolves, the constraint on the location volume improves, but it retains an elongated shape because, for events outside the network, the event distance is poorly determined. The depth is only constrained by a lower limit, but this depth bounds includes the true value.

Recently in Italy there have been large earthquakes characterized by multiple event ruptures and intense seismic activity related to foreshocks and aftershocks (*i.e.*, Friuli, 1976 [*Zollo et al.*, 1997]; Irpinia, 1980 [*Bernard and Zollo*, 1989]; Umbria-Marche, 1997 [*Amato et al.*, 1998]). The major instrumental event in the Irpinia region, the Mw=6.9, 1980 earthquake, had multiple sub-events with three main shocks occurring within about 20 seconds of each other. It is, therefore, important to check how our evolutionary location method performs when two or more events occur close in time.

We made a synthetic test for two events occurring at different places within the Irpinia Seismic Network with origin times separated by 3 seconds, using both P and S picks for location (Figure 4a). If an S pick from the first event (S1) comes after a P pick from the second event (P2), we assume a probability of 20% for the triggering system to erroneously interpret P2 as S1. For instance, ST13 is the first station recording the second event, but it does not trigger correctly. The first trigger comes from station ST25, biasing the hypocenter estimation in the very beginning of the location process (Figures 4b, 4c). The bias is however strongly reduced after 1 sec circa, as soon as new stations trigger consistently. There are other misinterpreted picks at T = 4.6s (ST13, S1 as S2), 8.7s (ST14, S1 as P2), 8.8s (ST04, S1 as S2), and 12.5s (ST02, S1 as P2). These outliers, however, do not influence significantly the quality of location, since this is already constrained by a large number of consistent picks (Figure 4c). The hypocentral position is correctly estimated, with no more strong oscillations, after about 2 seconds form the first trigger, while the depth is properly identified after 4-5 seconds. For both the events, the uncertainty on x and y components becomes smaller than 1 km after about 4 seconds from the first trigger, while the uncertainty on z is lower than 2 km after 5 seconds.

Discussion

We have presented a real time evolutionary location technique based on the equal differential-time (EDT) approach which makes it very robust in presence of outliers. At each time step, this algorithm makes use of information from triggered arrivals and not yet triggered stations. Constraint on the hypocenter location is obtained as soon as the first station has triggered and is updated at fixed time intervals or when a new station triggers.

The hypocenter location is estimated as a probability density function defined in a gridded search volume. This makes it easy to incorporate the location results into a decision system for seismic early warning. Such a system can base a decision rule on the evaluation of the probability that a certain ground motion intensity measurement (IM), like PGA or PGV, exceeds a given threshold [*Iervolino et al.*, this issue]. The probability density function for IM is calculated from the evaluation of the *hazard integral*:

$$\hat{f}_{IM}(im) = \int_{M}\int_{R} f_{IM|M,R}(im|m,r)\,\hat{f}_{M}(m)\,\hat{f}_{R}(r)\;dm\;dr$$

where $f_{IM|M,R}$ is an attenuation law, f_M is the *pdf* for the magnitude (estimated in real-time [*Zollo et al.*, this issue]), and f_R is the *pdf* for the hypocentral (or epicentral) distance, which can be obtained directly from our location technique.

Synthetic location tests show that a good accuracy, very close to standard "off-line" algorithms, is achieved after 4-5 seconds. The test on two quasi-simultaneous events demonstrates that, as long as the triggering system has a good detecting capability, the two locations can be handled as separate processes and wrong picks treated as outliers, whose bias is strongly reduced when several consistent arrivals are available.

References

Amato, A., et al. (1998), The 1997 Umbria-Marche, Italy, earthquake sequence: a first look at the main shocks and aftershocks, *Geophys. Res. Lett., 25(15)*, 2861–2864.

Bernard, P., and A. Zollo (1989), The Irpinia (Italy) 1980 earthquake: Detailed analysis of a complex normal faulting, *J. Geophys. Res., 94(B2)*, 1631–1647.

Font, Y., H. Kao, S. Lallemand, C.-S. Liu, and L.-Y. Chiao (2004), Hypocentral determination offshore Eastern Taiwan using the Maximum Intersection method. *Geophys. J. Int.*, *158*, 655-675.

Horiuchi, S., H. Negishi, K. Abe, A. Kamimura, and Y. Fujinawa (2005), An Automatic Processing System for Broadcasting Earthquake Alarms, *Bull. Seism. Soc. Am., 95*, 708-718.

Iervolino I., V. Convertito, M. Giorgio, G. Manfredi and A. Zollo (2006). The Crywolf Issue in Seismic Early Warning Applications for the Campanian Region. In *Seismic Early Warning*, Eds. P. Gasparini, G. Manfredi and J. Szchau, Springer-Verlag, XX-XX.

Lomax, A. (2005), A Reanalysis of the Hypocentral Location and Related Observations for the Great 1906 California Earthquake, *Bull. Seism. Soc. Am.*, *95*, 861-877.

Milne, J. (1886), *Earthquakes and Other Earth Movements*, Appelton, New York, 361pp.

Rydelek, P., and J. Pujol (2004), Real-time seismic warning with a 2-station subarray, *Bull. Seism. Soc. Am., 94*, 1546-1550.

Weber, E., G. Iannaccone, A. Zollo, A. Bobbio, L. Cantore, M. Corciulo, V. Convertito, M. Di Crosta, L. Elia, A. Emolo, C. Martino, A. Romeo, C. Satriano (2006). Development and testing of an advanced monitoring infrastructure (ISNet) for seismic early-warning applications in the Campania region of southern Italy. In *Seismic Early Warning*, Eds. P. Gasparini, G. Manfredi and J. Szchau , Springer-Verlag, XX-XX.

Zhou, H. (1994), Rapid 3-D hypocentral determination using a master station method, *J. Geophys., Res., 99*, 15439-15455.

Zollo, A. and M. Lancieri (2006). Real-time estimation of earthquake magnitude for seismic early warning. In *Seismic Early Warning*, Eds. P. Gasparini, G. Manfredi and J. Szchau, Springer-Verlag, XX-XX.

Zollo, A., A. Bobbio, A. Emolo, A. Herrero and G. De Natale (1997). Modelling of ground acceleration in the near source range: the case of 1976, Friuli earthquake (M = 6.5), northern Italy, *Journal of Seismology, 1(4)*, 305-319.

Figure Captions

Figure 1

Evolutionary earthquake location algorithm.

(a) Given a seismic network with known sets of operational and non-operational stations, we can a priori define the Voronoi cell associated to each station.

(b) When the first station triggers, we can define a volume that is likely to contain the hypocenter limited by the "conditional" EDT surfaces on which the P travel time to the first triggering station is equal to the travel-time to each of the operational but not yet triggered stations.

(c) As time progresses, we gain additional information from the stations have not yet triggered: the EDT surfaces move towards and bend around the first triggering station, and the hypocenter volume shrinks.

(d) When the second station triggers, we can define a "true" EDT surface and the actual hypocenter is likely to be at the intersection between this surface and the previous defined volume (which keeps shrinking).

(e) When a third station triggers, we can define two more "true" EDT surfaces, increasing the constraint on hypocenter position.

(f) As more stations trigger, the location converges to the standard EDT location.

Figure 2

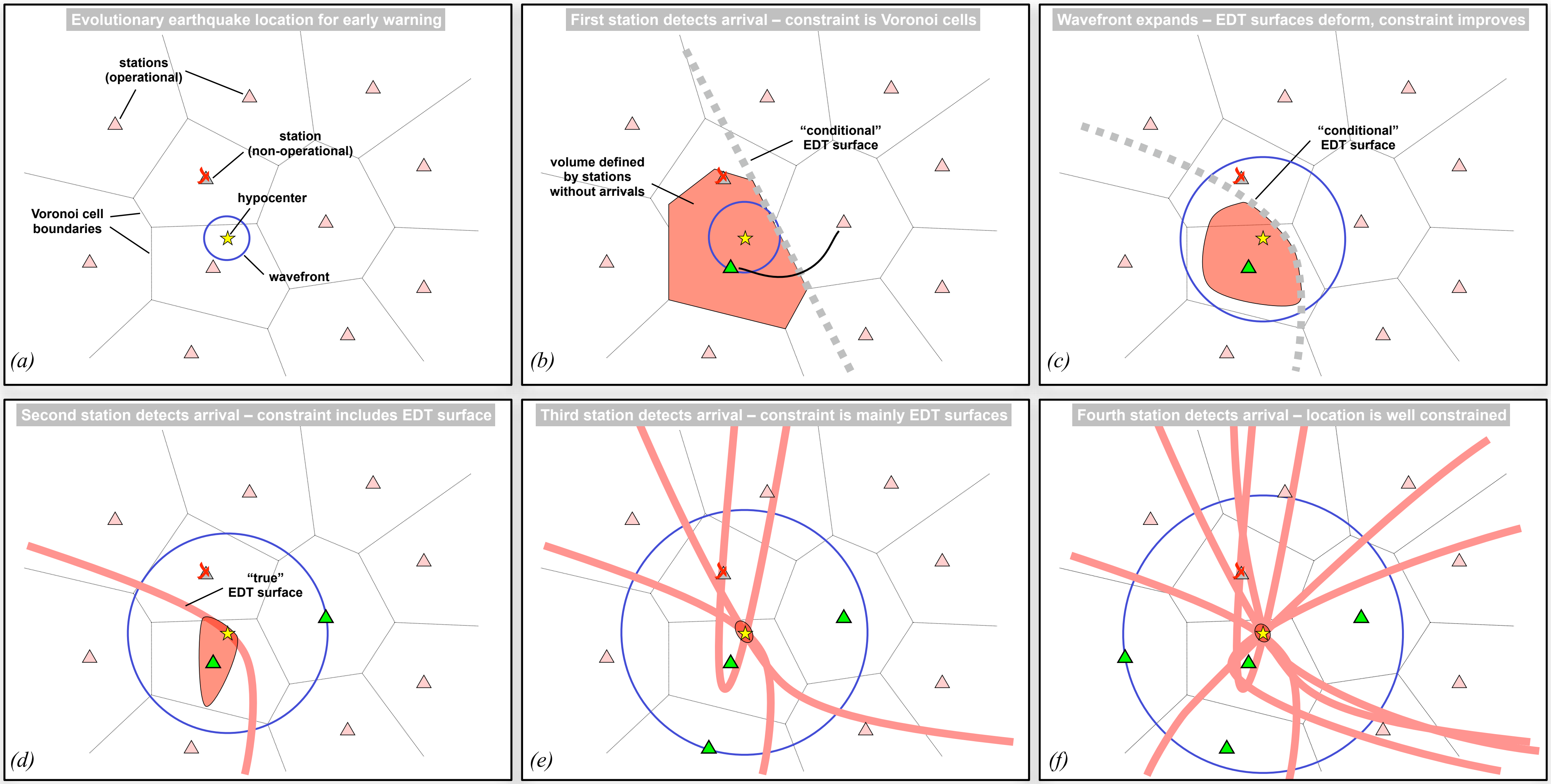
Location test for an event occurring at the center of the ISNet network. The probability function is projected along three planes passing through the true hypocenter (identified with a star). T=0sec is the time at which the first station triggers. For each snapshot, stations which have triggered are marked with a circle. Location is performed using only P picks.

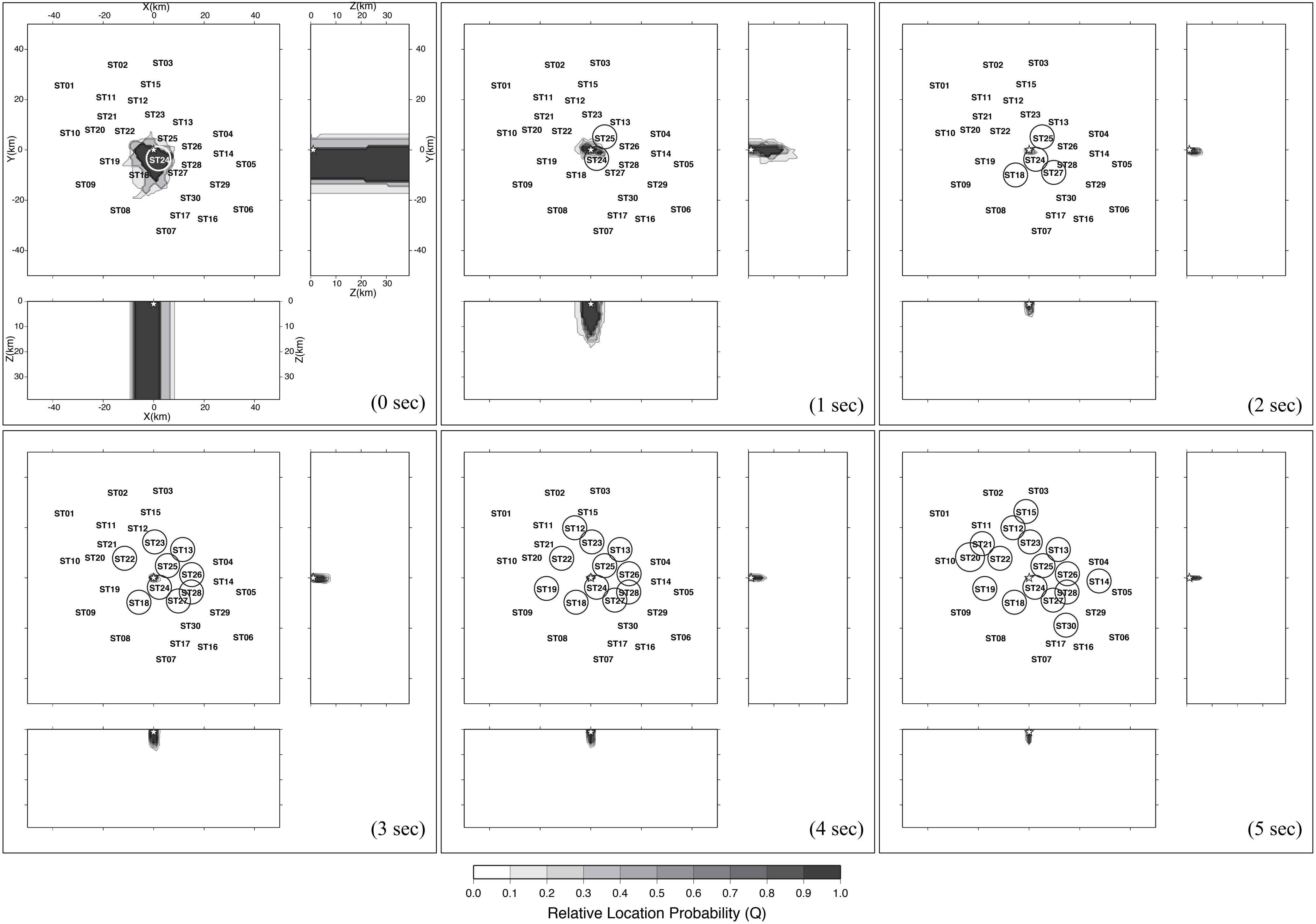
Figure 3

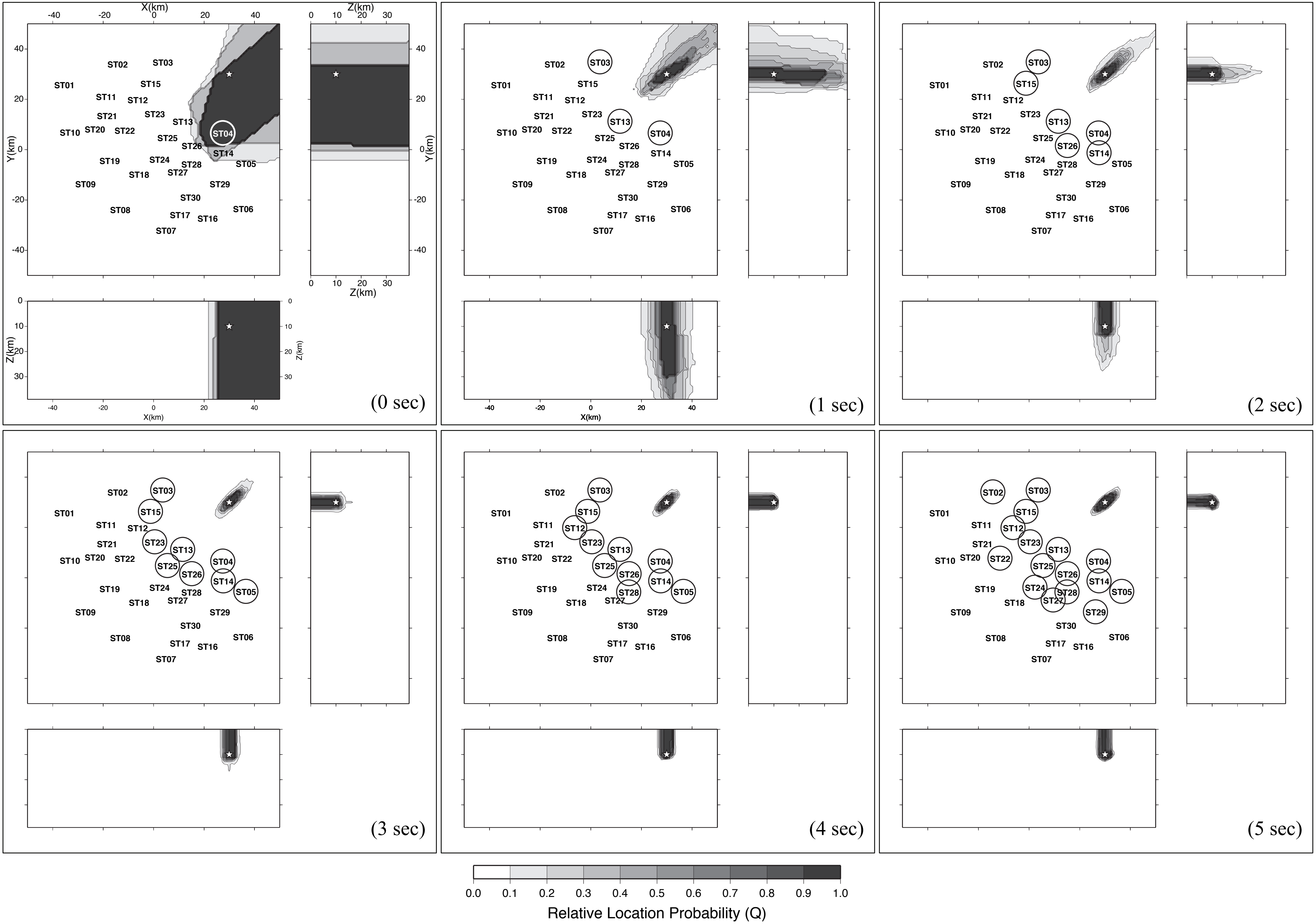
Location test for an event occurring outside the network (see fig. 2 for notes).

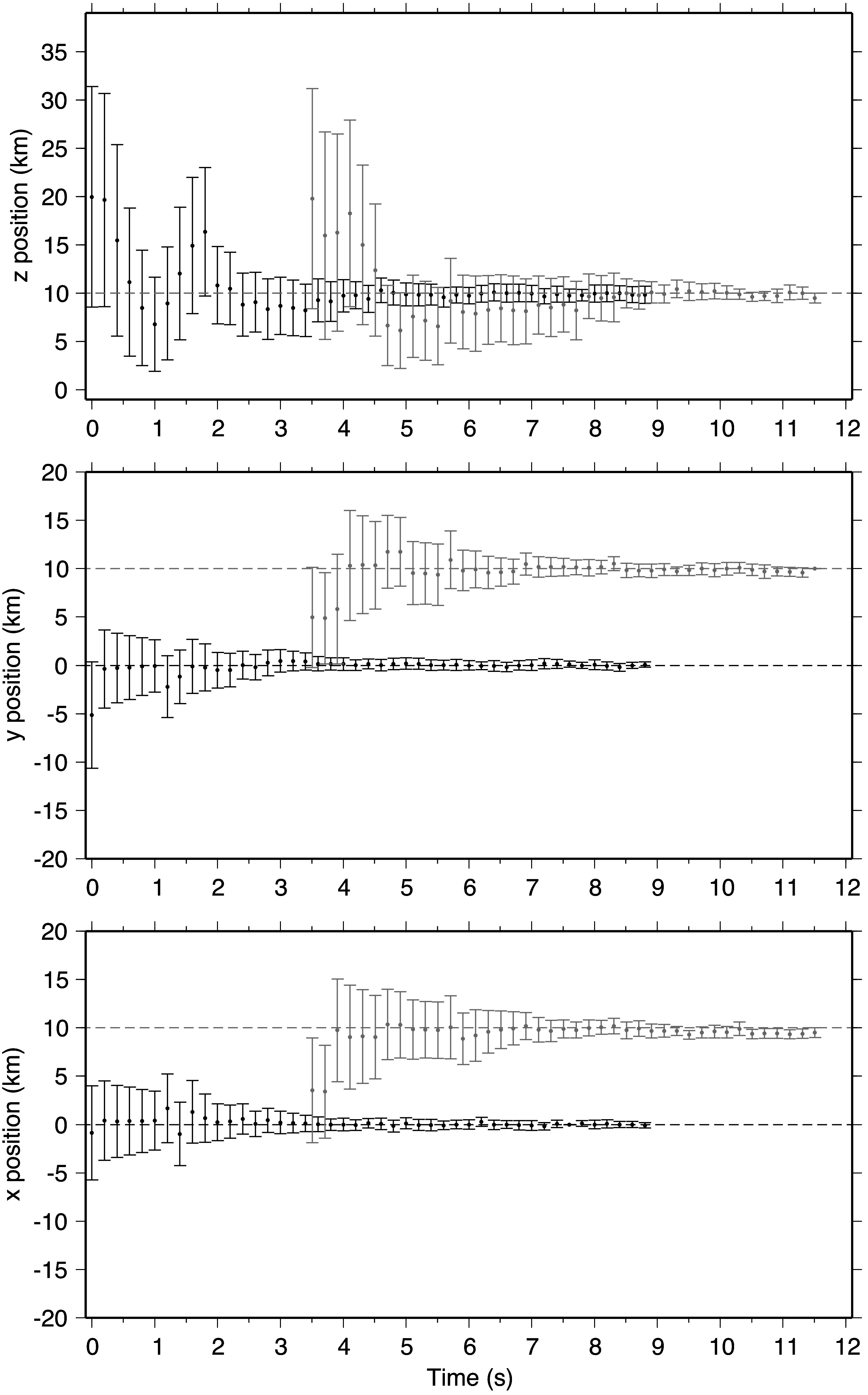
Figure 4

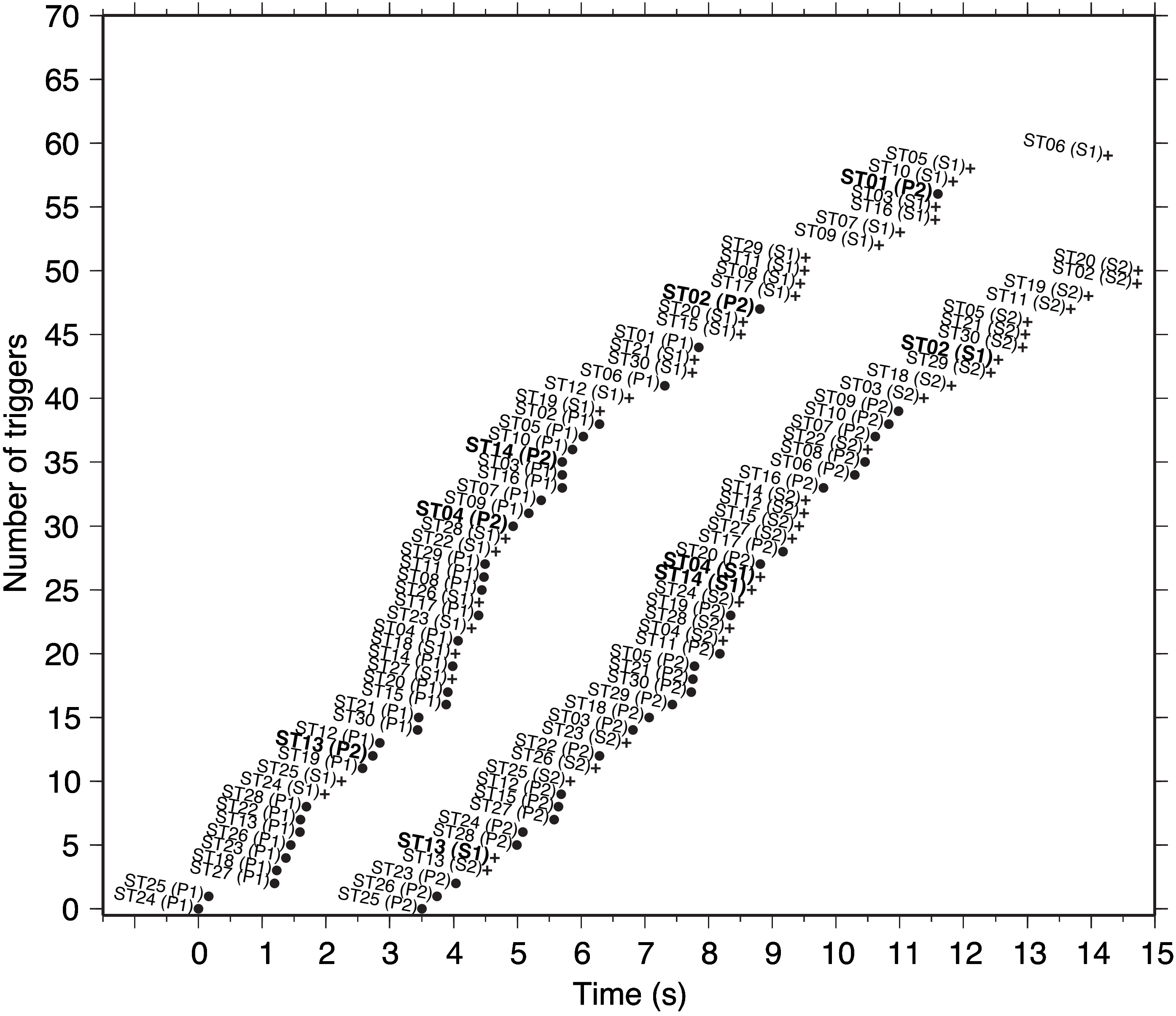
Location test for two events occurring at different places within the Irpinia Seismic Network with origin times separated by 3 seconds. T=0sec is the time at which first station triggers. (a) Actual position of the first (black star) and the second (gray star) hypocenter. (b) Mean value and standard deviation for location along the three axes for the first (black bars) and the second (gray bars) event. The dashed lines represent the true values. (c) Triggering sequence for the first (upper sequence) and the second (lower sequence) event. P triggers are marked with dots, S triggers with stars. Misinterpreted arrivals are evidenced in bold.


Evolutionary earthquake location for early warning
stations
(operational)
station
(non-operational)
hypocenter
Voronoi cell
boundaries
wavefront
(a)
First station detects arrival – constraint is Voronoi cells
"conditional"
EDT surface
volume defined
by stations
without arrivals
(b)
Wavefront expands – EDT surfaces deform, constraint improves
"conditional"
EDT surface
(c)
Second station detects arrival – constraint includes EDT surface
"true"
EDT surface
(d)
Third station detects arrival – constraint is mainly EDT surfaces
(e)
Fourth station detects arrival – location is well constrained
(f)


X(km)
Y(km)
Z(km)
-40
-20
0
20
40
10
30
ST01
ST02
ST03
ST04
ST05
ST06
ST07
ST08
ST09
ST10
ST11
ST12
ST13
ST14
ST15
ST16
ST17
ST18
ST19
ST20
ST21
ST22
ST23
ST24
ST25
ST26
ST27
ST28
ST29
ST30
(0 sec)
(1 sec)
(2 sec)
(3 sec)
(4 sec)
(5 sec)
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1.0
Relative Location Probability (Q)


X(km)
Y(km)
Z(km)
-40
-20
0
20
40
10
30
ST01
ST02
ST03
ST04
ST05
ST06
ST07
ST08
ST09
ST10
ST11
ST12
ST13
ST14
ST15
ST16
ST17
ST18
ST19
ST20
ST21
ST22
ST23
ST24
ST25
ST26
ST27
ST28
ST29
ST30
(0 sec)
(1 sec)
(2 sec)
(3 sec)
(4 sec)
(5 sec)
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1.0
Relative Location Probability (Q)


z position (km)
35
30
25
20
15
10
5
0
y position (km)
20
15
10
5
0
-5
-10
-15
-20
x position (km)
20
15
10
5
0
-5
-10
-15
-20
0
1
2
3
4
5
6
7
8
9
10
11
12
Time (s)


Number of triggers
0
5
10
15
20
25
30
35
40
45
50
55
60
65
70
Time (s)
0
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
ST24 (P1)
ST25 (P1)
ST27 (P1)
ST18 (P1)
ST23 (P1)
ST26 (P1)
ST13 (P1)
ST22 (P1)
ST28 (P1)
ST24 (S1)
ST25 (S1)
ST19 (P1)
ST13 (P2)
ST12 (P1)
ST30 (P1)
ST21 (P1)
ST15 (P1)
ST20 (P1)
ST27 (S1)
ST14 (P1)
ST18 (S1)
ST04 (P1)
ST23 (S1)
ST17 (P1)
ST26 (S1)
ST08 (P1)
ST11 (P1)
ST29 (P1)
ST22 (S1)
ST28 (S1)
ST04 (P2)
ST09 (P1)
ST07 (P1)
ST16 (P1)
ST03 (P1)
ST14 (P2)
ST10 (P1)
ST05 (P1)
ST02 (P1)
ST19 (S1)
ST12 (S1)
ST06 (P1)
ST30 (S1)
ST21 (S1)
ST01 (P1)
ST15 (S1)
ST20 (S1)
ST02 (P2)
ST17 (S1)
ST08 (S1)
ST11 (S1)
ST29 (S1)
ST09 (S1)
ST07 (S1)
ST16 (S1)
ST03 (S1)
ST01 (P2)
ST10 (S1)
ST05 (S1)
ST06 (S1)
ST25 (P2)
ST26 (P2)
ST23 (P2)
ST13 (S2)
ST13 (S1)
ST28 (P2)
ST24 (P2)
ST27 (P2)
ST15 (P2)
ST12 (P2)
ST25 (S2)
ST26 (S2)
ST22 (P2)
ST23 (S2)
ST03 (P2)
ST18 (P2)
ST29 (P2)
ST30 (P2)
ST21 (P2)
ST05 (P2)
ST11 (P2)
ST04 (S2)
ST28 (S2)
ST19 (P2)
ST24 (S2)
ST14 (S1)
ST04 (S1)
ST20 (P2)
ST17 (P2)
ST27 (S2)
ST15 (S2)
ST12 (S2)
ST14 (S2)
ST16 (P2)
ST06 (P2)
ST08 (P2)
ST22 (S2)
ST07 (P2)
ST10 (P2)
ST09 (P2)
ST03 (S2)
ST18 (S2)
ST29 (S2)
ST02 (S1)
ST30 (S2)
ST21 (S2)
ST05 (S2)
ST11 (S2)
ST19 (S2)
ST02 (S2)
ST20 (S2)