First European Conference on Earthquake Engineering and Seismology
(a joint event of the 13th ECEE & 30th General Assembly of the ESC)
Geneva, Switzerland, 3-8 September 2006
Paper Number: 810

IMPROVED EARTHQUAKE LOCATIONS IN GREECE USING THE DD ALGORITHM AND A 3D VELOCITY MODEL

Odysseus GALANIS[1], Constantinos PAPAZACHOS[2], Emmanuel SCORDILIS[3], and Panagiotis HATZIDIMITRIOU[4]

SUMMARY

Accurate earthquake locations are necessary for seismicity, seismic hazard, seismotectonics, seismic tomography and other basic research and applications. As the existing catalogs are not, in most cases, accurate enough, they need to be revised using more advanced earthquake location algorithms such as the double-difference (DD) earthquake location algorithm. We located ~100000 earthquakes which occurred in Greece and the surrounding areas (33-43N, 18-30E) during a period of 23 years (1981-2003). The data used was a list of P and S wave arrival picks compiled by combining data from the archives of a regional seismic network operating since 1981, from phase picks contributed to the International Seismological Center by neighbouring networks, as well as data from temporary local networks. The earthquake hypocenters were relocated using the Double-Difference earthquake location algorithm whenever this was possible. The original, freely distributed implementation of the algorithm [Waldhauser, 2001] was altered, in order to use a three-dimensional seismic wave velocity model of the area which has been determined by earlier tomographic studies. In cases where double-difference location was not possible because the event is spatially isolated or the pick errors are too high, the events were located with the conventional absolute location method (Geiger's method). The resulting earthquake catalog consists of events divided into two categories, according to the method used for the location (double difference location or absolute location). The new catalog, especially the double-difference located part, reveals seismicity patterns otherwise invisible or blurred by the absolute location errors.

1. INTRODUCTION

1.1 The Double-Difference Method for Earthquake Location

According to [Aki and Richards, 1980], earthquake location by using arrival-time data is by far the oldest inverse problem studied in seismology. According to [Lomnitz, 2006], on the other hand, "we do not know how to locate earthquakes". From these, and from the fact that location is the first step in every research related with earthquakes, it follows that earthquake location is neither a tangential nor a trivial problem for modern seismology.

The standard method for locating earthquakes, most commonly used by seismological networks, is Geiger's method [Geiger, 1910]. This consists of iteratively solving a linear system of equations for the four source

[1] Geophysical Laboratory, School of Geology, Aristotle University of Thessaloniki, P.O. Box 352-1, GR-54 124 Thessaloniki, Greece
Email : ogalanis@geo.auth.gr
[2] Geophysical Laboratory, School of Geology, Aristotle University of Thessaloniki, P.O. Box 352-1, GR-54 124 Thessaloniki, Greece
Email : kpapaza@geo.auth.gr
[3] Geophysical Laboratory, School of Geology, Aristotle University of Thessaloniki, P.O. Box 352-1, GR-54 124 Thessaloniki, Greece
Email : manolis@geo.auth.gr
[4] Geophysical Laboratory, School of Geology, Aristotle University of Thessaloniki, P.O. Box 352-1, GR-54 124 Thessaloniki, Greece
Email : chdimitr@geo.auth.gr

parameters (three spatial coordinates and origin time). More specifically, an a priori velocity model (typically one consisting of layers of constant velocity) is used to calculate the expected arrival times $T_i(\mathbf{a^0})$ from a trial source with parameters $\mathbf{a^0}$, to a number of stations. The differences of these times with the observed arrival times t_i, $d_i(\mathbf{a^0})=t_i-T_i(\mathbf{a^0})$ are used to deduce the corrections that have to be applied to the trial source parameters, towards the source parameters $\hat{\mathbf{a}}$ which minimize these differences or travel time residuals. It is assumed that these corrections are, at a first approach, linearly related to the travel time residuals. This can be formulated as a system of linear equations

$$\mathbf{d} = \mathbf{Gm} \tag{1}$$

where $\mathbf{d}$ is a vector with components d_i, $\mathbf{m}=\mathbf{a^0}-\hat{\mathbf{a}}$ is the vector source parameters' corrections and $\mathbf{G}$ is a matrix with elements

$$G_{ij} = \frac{\partial d_i}{\partial a_j}(\mathbf{a^0}) \tag{2}$$

which serves to quantify the effect of a change in source to the corresponding change in arrival time. Equation 1 is solved iteratively until $\mathbf{a^0}$ converges to $\hat{\mathbf{a}}$ to a desirable degree of precision. Starting in the 1960s there have been many implementations of Geiger's algorithm in the form of computer programs. Some of them are HYPOLAYR [Eaton, 1969], HYPO71 [Lee and Lahr, 1972], HYPOINVERSE [Klein, 1985] and HYPOELLIPSE [Lahr, 1980].

The Double-Difference (DD) method, introduced by [Waldhauser and Ellsworth, 2000] is also based in solving a system of equations similar to equation 1. However, the quantity to minimize is not the travel time difference but the double difference

$$d_k = (t_k^i - t_k^j) - (T_k^i(\mathbf{a^i}) - T_k^j(\mathbf{a^j})) \tag{3}$$

where, conserving the same notation as the previous equations, an observation d_k is the difference between two differences (hence the name of the method). These are the difference of observed travel times t_k^i and t_k^j for two events i and j, at the same station, and the difference of the calculated travel times T_k^i and T_k^j for the same two events at the same station. In this case, the components of $\mathbf{m}$ are the corrections to all source parameters of all the events being located. In addition, for equation 1 to have any meaning, the components of $\mathbf{G}$ must become

$$G_{kl} = \frac{\partial d_k}{\partial a_l} \tag{4}$$

for all observations d_k and all source parameters a_l. In the double-difference case, the matrix $\mathbf{G}$ is very sparse, as a change in a source parameter does not affect observations of double-differences between travel times from all other sources. Some of the advantages of the method follow directly from equation 3. Firstly, there are far more observations, as for N events there can be, in the best case, up to $N(N-1)/2$ couples of events. Secondly, the term $(T_k^i-T_k^j)$ is only affected by the values of the seismic wave velocity in the region between the locations of the two events and not by the velocity values along the whole path to the station. Lastly, equation 3 implies that the exact values of t_k^i and t_k^j are irrelevant as long as their difference is known. This allows for the use of differential arrival times, like those obtained by waveform cross-correlation or any other method of aligning seismic signals. The DD method has been implemented as a freely available computer program under the name HYPODD [Waldhauser, 2001].

1.2 Area of Study

The area studied is Greece and the surrounding regions, and more specifically the rectangle with latitudes 33°N-43°N and longitudes 18°E-30°E. The active tectonics status of the area is an immediate result of its placement in a boundary region between tectonic plates (figure 1, [Papazachos et al., 1997]). The most prominent large-scale tectonic features are the Hellenic arc, where the subduction of the Eastern Mediterranean Lithosphere beneath the Aegean Microplate takes place, the western part of the North Anatolian Fault Zone along which the

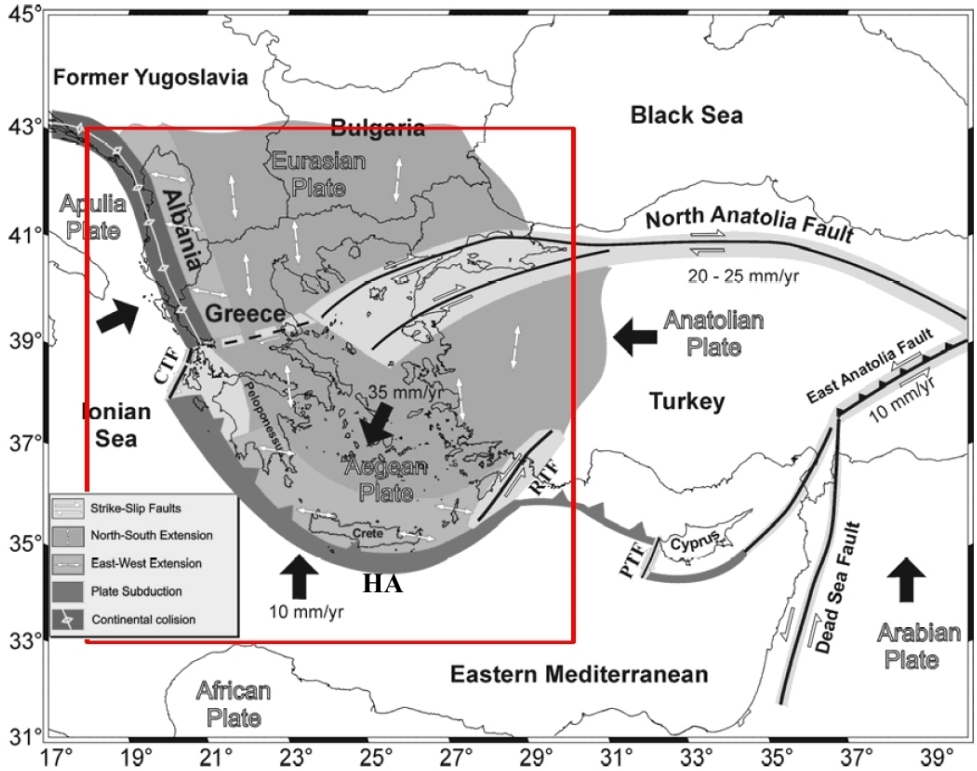

Figure 1: Tectonic Plate boundaries and motions at the area of interest (modified from Papazachos et al. [1997]). The Hellenic Arc (HA), North Anatolia Fault Zone and Cephalonia Transform Fault (CTF) are marked. The red rectangle defines the area studied in this work.

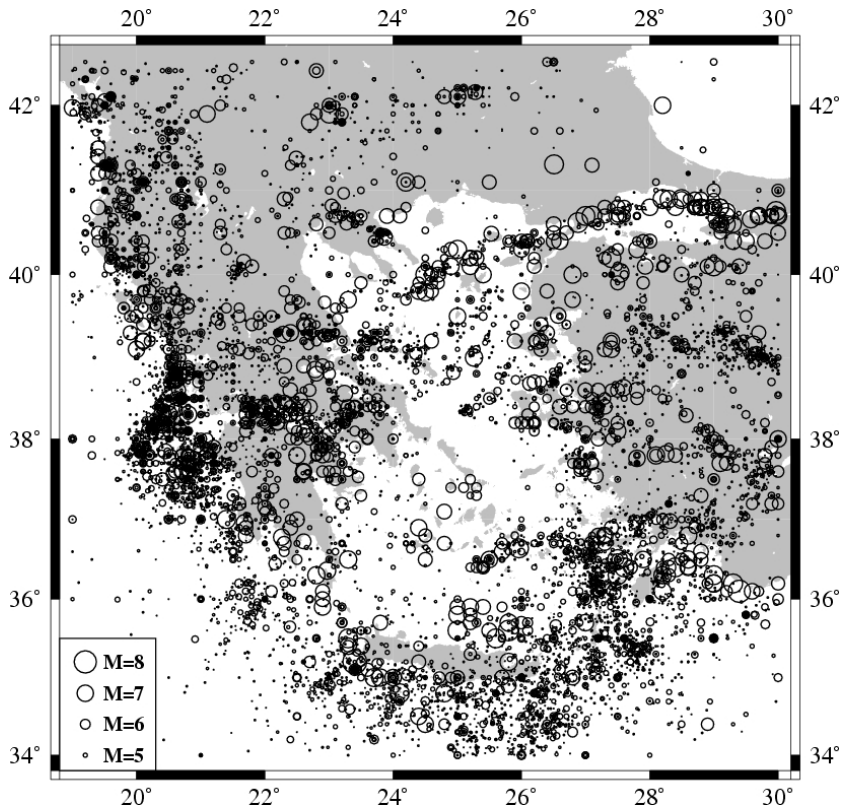

Figure 2: Seismicity of the region. The spatial extent of this data set is slightly smaller than the one used in this work. It consists of all known earthquakes of magnitude M≥6.0 since 550BC, M≥4.9 since 1911 and M≥4.5 since 1950 and up to 2005AD [Papazachos et al. 2000] and [Papazachos et al. 2006]. Note the concentration of epicentres along the main large-scale tectonic features.

Anatolian Plate moves westwards, and the Cephalonia Transform Fault at the northwest end of the Hellenic Arc. Shallow earthquakes occur throughout all of this area, with a maximum at the Ionian Sea Islands, which have the highest level of seismicity in the western part of the Eurasian-Melanesian Zone. Intermediate depth earthquakes are limited to the South Aegean Sea where they define a Benioff zone. The focal depth of these earthquakes can reach down to 180 Km [Papazachos and Papazachou, 1989]. The epicenters of three sets of known events for the period 550BC – 2005AD [Papazachos et al., 2000, Papazachos et al., 2006] are plotted in figure 2. The three sets are all known earthquakes of magnitude M≥6.0 since 550BC, M≥4.9 since 1911 and M≥4.5 since 1950 and up to 2005AD.

2. DATA

The data set used was body wave arrival times at stations in Greece and the surrounding area. The bulk of the data was retrieved from the archives of the telemetric network of the Aristotle University of Thessaloniki. This network has been operating continuously since the beginning of 1981. A bulletin which contains phase arrival information is published monthly [Geophysical Laboratory, Aristotle University of Thessaloniki 2006]. Data from this bulletin was merged with data from a number of temporary local networks operating in Greece from 1981 to 1995. Finally, phases reported to the International Seismological Centre were also added [ISC 2006].
The merging of data was accomplished by starting with an initial catalog and then progressively adding data from the other data sources. The initial catalog of events and phases was one which had been compiled by the Geophysical Laboratory of the Aristotle University of Thessaloniki [Papazachos et al., 2000]. As this covered the period from 1981 to 1995 only, the ISC bulletin was used to extend this data set to the end of 2003. This choice for the initial catalog proved to be quite successful, as at the end of the merging procedure relatively few additions had been performed. After redundant (duplicate) phases and events were deleted and the phase catalog was sorted by origin time, a first removal of outliers was done based on theoretical travel times using a simple three layer velocity model [Panagiotopoulos, 1984]. All phases with an absolute residual of 10s or larger, in the case of P waves and 15s or larger, in the case of S waves were removed. Phases corresponding to epicentral distances longer than 10 degrees were also removed.

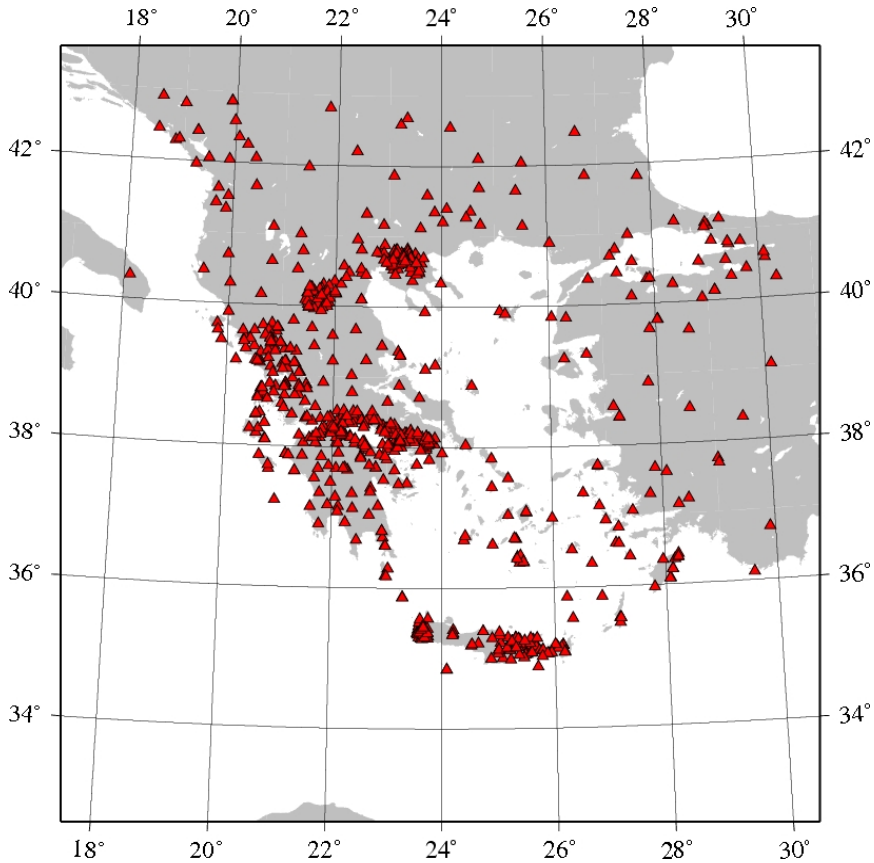

Figure 3: Stations used in the present study

Adding phases and events to the initial catalog was performed in a semiautomatic manner. The software developed for this purpose, works on phases assumed by the data provider to originate from the same event. The group of phases is assigned to the event of the initial catalog which fits best in terms of expected arrival times. In the case one of the phases could be attributed to two or more events, the user is prompted to decide for the whole group or for each phase separately. In the case in which no event in the initial catalog fits the phase group sufficiently, the event parameters provided are adopted and a new event is added to the initial catalog.

As three dimensional ray tracing is only possible if both the source and the station lie inside the volume in which the seismic wave velocity is defined, a large number of phases was not used in the DD relocation procedure (although these phases were used for location with the standard method). The stations which were used are shown in figure 3, and the events which were used are shown in figure 4. In the end, the data set used for DD-relocation consisted of 1 568 400 arrival times from 126 318 earthquakes at 711 stations.

The 3-D seismic wave velocity model used was based on the model of [Papazachos, 1994] and [Papazachos and Nolet, 1997].This velocity model was patched with smaller scale, more precise models when possible. This was done for Albania and the surrounding area [Papazachos et al., 2006], the Serbomacedonian Massif in Northern Greece [Papazachos, 1998], and an area including the North Aegean, the eastern part of Northern Greece and the Southern part of Bulgaria [Papazachos and Scordilis, 1998]. The model was extended to Central and Eastern Mediterranean, based on known Moho discontinuity depths [Papazachos and Karagianni, 2006], and an average one-dimensional velocity model for the area [Comninakis and Papazachos, 1976]. The volume in which the P and S wave velocity was defined is a parallelepiped with dimensions 1050Km from E to W, 1150Km from S to N and 200Km from its base to its top. The vertical axis is parallel to the Earth's radius which intersects the surface at a point with geographical coordinates 37.8°N, 24.5°E, and the two horizontal axes are oriented to the North and East at the same point.

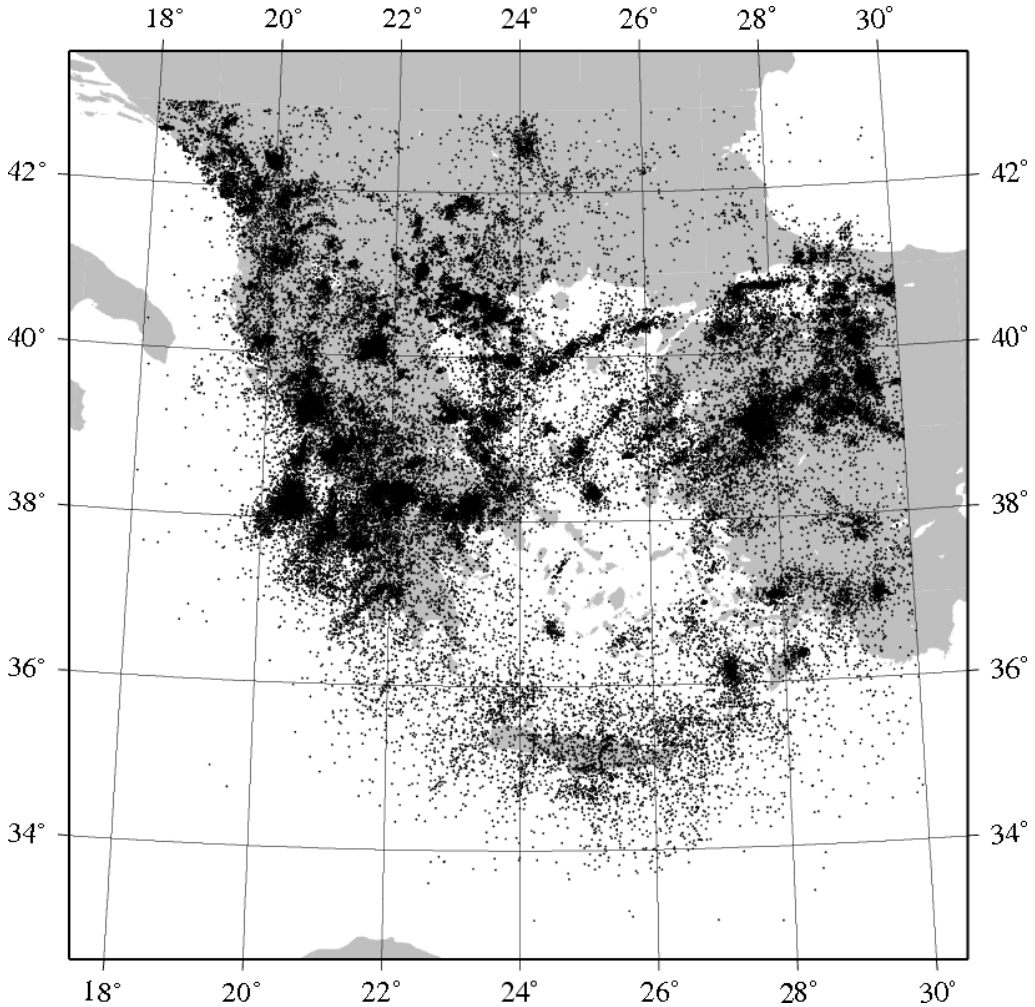

Figure 4: Events used in the present study

3. METHOD

3.1 Three Dimensional Ray Tracing

Ray tracing is necessary for evaluating theoretical travel times (T in equation 3) and the partial derivatives of the double differences with respect to source parameter corrections (the non-zero elements of **G** in equation 4). The 3-D ray tracing was performed in two steps. The first step is finding the fastest path in a regular grid as described in detail in [Moser, 1991]. The volume in which the velocity is defined, was divided in cells with dimensions 10Km × 10Km × 2 Km, the latter dimension being the depth. Except for the regular nodes at the vertices of cells, additional nodes are defined at the location of the stations. It is then possible to calculate travel times along straight lines joining two nodes. This allows for finding the shortest path, consisting of segments joining nodes, between any two nodes, using Dijkstra's algorithm [Dijkstra, 1959]. Thus the shortest path from the station to every other node is calculated. In fact, Dijkstra's algorithm calculates the shortest path to every node anyway, even if the shortest path to just one node is needed.

After an approximate shortest path is found, it can be optimised using ray bending as described in [Aki and Richards, 1980] and [Moser et al., 1992]. In brief, the ray bending procedure perturbs the ray to the direction which minimizes the travel-time derivative. Having an approximate initial path from the discrete grid, guarantees that the absolute minimum, instead of a local minimum, of the travel time function will be reached. Once the ray has been calculated, finding the partial derivatives G_{kl} is trivial as these depend only on the seismic wave velocity and the direction of the ray at the origin.

Taking advantage of the fact that earthquakes do not occur in all of the volume of interest, but only in specific parts of it, ray bending is performed only for nodes surrounding cells which actually contain earthquake foci. Then the travel time and partial derivatives at any point inside the cell are estimated by linear interpolation from the 8 nodes at the vertices of the cell. Cubic interpolation is also possible, and it requires using values from the nodes of the 26 cells which are in contact to the cell of interest by sharing a face, edge or vertex. By calculating the rays from the significant nodes only, for each station, and storing the results before starting the location process, 3-D ray tracing becomes efficient in terms of computation time. If, during the course of earthquake relocation, a focus shifts inside a cell with unknown travel times and derivatives at some of its vertices, then the appropriate rays are calculated and the travel times and derivatives are stored for future use.

3.2 Inversion of Travel Times

The method used to invert the travel time data was the Double Difference (DD) algorithm [Waldhauser and Ellsworth, 2000]. The original algorithm had to be modified in order to use travel times and partial derivatives estimated by 3D ray tracing. Also, the algorithm had to additionally check if an earthquake focus shifts outside the volume in which the velocity model is defined. Finally, a modification was made concerning the inversion process itself. [Waldhauser, 2001] suggests the empirical rule than when using damped least squares, the amount of damping should be such that the condition number of the linear system is between 40 and 80. As locating tens of thousands of earthquakes involves a large number of earthquake clusters with condition numbers ranging many orders of magnitude, a single value of damping, which is the only option in the original software, would be inappropriate. Instead, the modified program, adjusts by trial and error the amount of damping for each cluster, so that the condition numbers of the resulting linear systems always have reasonable values (typically < 1000).

The clustering parameters used are chosen so that as few events as possible end up not being members of a cluster. On the other hand, if the links between events are not strong enough, i.e. there are not enough common phases (recordings at the same station) for a couple of events, the linear system may become ill-conditioned. Moreover, if the inter-event distance is too long, the assumption that the spatial fluctuations of velocity are relatively unimportant (see section 1.1) is no longer valid. Taking these factors into account, the clustering parameters used were a maximum distance of 10Km between event couples, with at least 8 common phases for each couple. A visualization of the division of events into clusters is shown in figure 5.

After a first set of iterations in which the data is not weighed at all, there is a second set of iterations during which the observations (double-differences) are weighed by a factor which depends on their deviation from their expected values. As a result of this, the events are relocated using all available data in the first iterations, and then the locations are fine-tuned using the best data only.

4. RESULTS

Out of 126 318 events, 49 958 were clustered and given as input to the DD procedure. DD relocation was successfully completed for 48 803 of them. The remaining events are those for which relocation could not be completed because their foci shifted out of the volume for which the velocity model is defined, or they lost their connection to other events when some DD observations were rejected as outliers. A map of all clustered events before and after relocation is shown in figure 6 (left-hand side). When compared to figures 2 and 4, the two distributions of epicentres exhibit a sharper image because spatially isolated earthquakes have been omitted. It should be kept in mind though that they are not representative of the seismicity of the area. When seen in detail (figure 6, right-hand side), the relocated events seem to delineate linear structures more clearly than the unrelocated events. This is even clearer when examining the 3-D distribution of seismicity in profiles, as is seen in the example presented in figure 6 (right-hand side) for one of the main branches of the North Anatolia Trough. Examination of the depth distribution along cross-section A-B, shows that not only the horizontal spatial clustering of hypocenters is more "compact", delineating a nearly-vertical strike-slip fault but also that the depth extent of the main earthquake cluster is limited between 0-20km, whereas in the original distribution a large number of events extended to depths up to 30km. Furthermore, the large number of events exhibiting and artificial depth of 0 km in the original cross-section (due to the limitations of the location procedure and/or velocity model used) have been relocated to larger "normal" depths.

In areas where earthquakes are not tightly clustered, at least in this data set which covers a relatively short period, the distribution of foci is virtually unresolved. For example in the South Aegean the seismicity is at least as high as the rest of the area, but for most events the DD-relocation procedure fails. As a result of this there is also little or no improvement in the locations of the Benioff Zone events.

The events which were not relocated successfully with the DD method were located using Geiger's algorithm with all available phases and the mean 1-D velocity model for the area. The new catalog consists of 126 549 events, out of which 48 803 are marked as 'D-class' (DD-located) and 77746 as 'G-class' (Geiger-located). The magnitudes (for events which have an estimated magnitude) are directly inherited from the catalogs of [Papazachos et al. 2000] and [Papazachos et al. 2006].

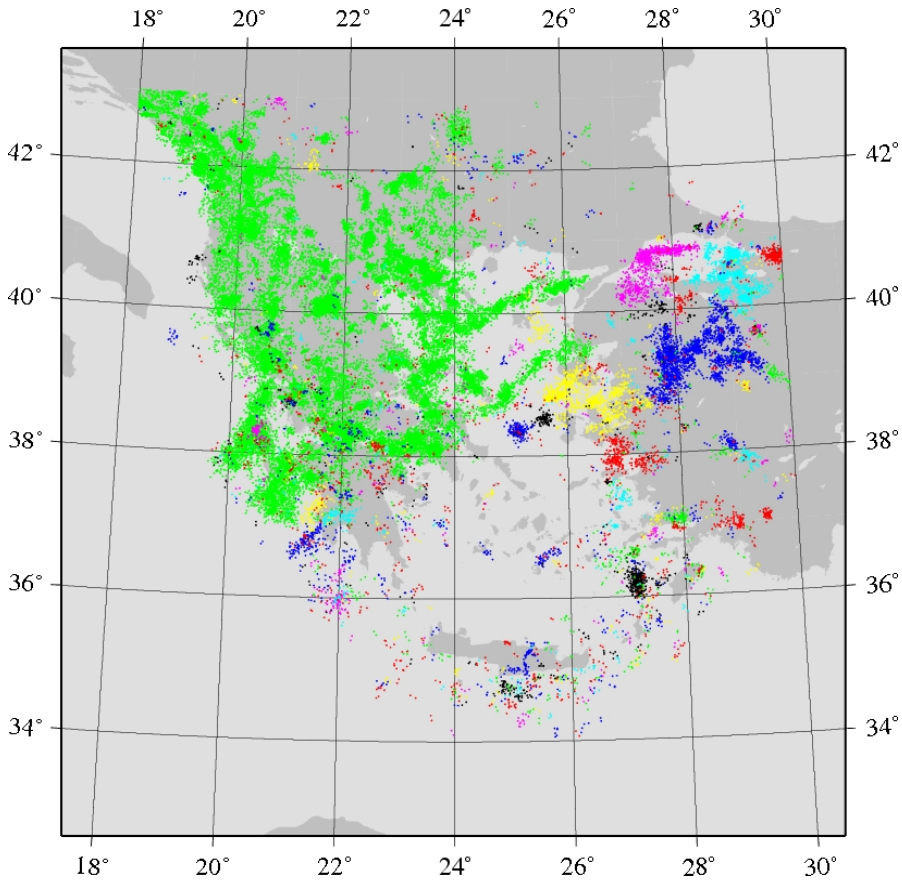

Figure 5: Event clustering. Different colours denote different clusters.

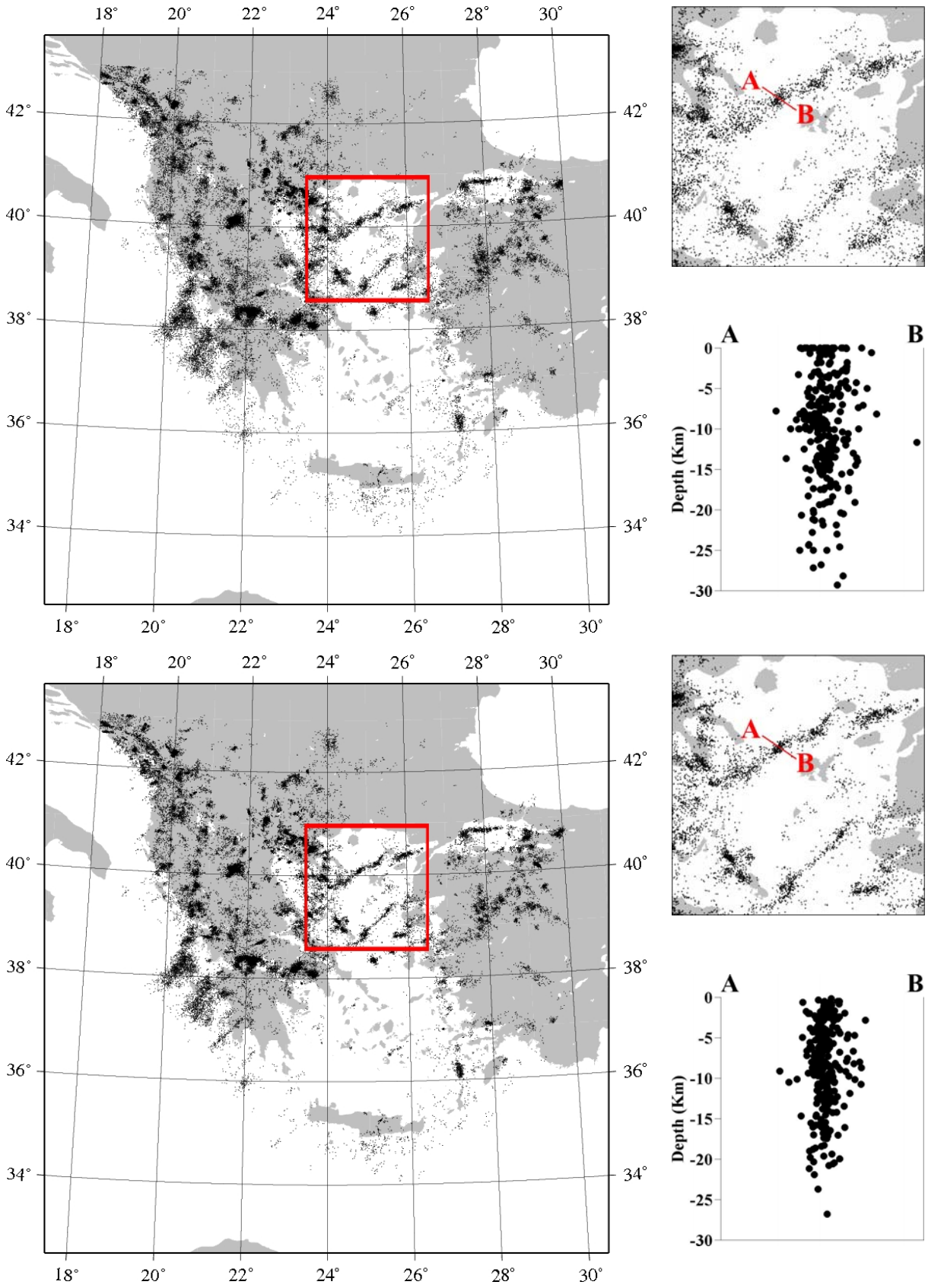

Figure 6: Relocation of clustered events. Catalog locations are shown on the top and DD-located events on the bottom. On the right of the large maps there is a detailed view of the area in the red rectangle. On the top it is in map view and on the bottom in a profile perpendicular to the main tectonic feature.

5. CONCLUSIONS

Double difference location can significantly improve the accuracy of the focal parameters of earthquakes. Its accuracy depends mainly, like absolute earthquake location, on the distribution of the seismological stations, and the quality of phase picks. In the present work we demonstrate that the incorporation of double difference techniques in combination with a 3-D velocity model can result in a significant improvement of the instrumental-period catalog and the delineation of active structures which are “blurred” by the traditional location process.

We intend to further develop the obtained results by improving the quality of phase picks in the future with the use of waveform cross-correlation like [Poupinet et al., 1984] or other techniques for aligning waveforms like [Du et al., 2004]. Additional improvement of the final catalogue will be implemented by the use of a three-dimensional velocity model also for the location of events with Geiger’s method, as in that case velocity fluctuations are much more critical than in the double-difference case. Finally, a quantitative assessment of location errors for the specific data set would make it useful for practical purposes.

ACKNOWLEDGEMENTS

The GMT software [Wessel and Smith, 1995] was used to generate most of the maps of this study. This work has been supported by the Greek Ministry of Education HERACLITUS (EPEAEK) project (Res. Comm. AUTH project 21761).

6. REFERENCES

Aki, K. and Richards, P.G. (1980), Quantitative Seismology, Theory and Methods, vol. II, *W.H. Freeman and Company*, San Francisco

Comninakis, P.E. and Papazachos, B.C. (1976), A note on the crustal structure of the eastern Mediterranean, *Annali di Geofisica*, 29, 59-63

Dijkstra, E.W. (1959), A Note on Two Problems in Connexion with Graphs, *Numerische Mathematik*, 1, 269-271

Du, W.-X. , Thurber, C.H. and Eberhart-Phillips, D. (2004), Earthquake Relocation Using Cross-Correlation Time Delay Estimates Verified with the Bispectrum Method, *Bulletin of the Seismological Society of America*, 94, 856-866

Eaton, J. P. (1969), HYPOLAYR, a computer program for determining hypocenters of local earthquakes in an earth consisting of uniform flat layers over a half space, *U.S. Geological Survey Open-File Report*, 155 p.

Geiger, L. (1910), Herdbestimmung bei Erdbeben ans den Ankunftzeiten, *Königliche Gesellschaft der Wissenschaften zu Göttingen*, 4, 331-349

Geophysical Laboratory, Aristotle University of Thessaloniki (2006), Seismicity Bulletin, http://lemnos.geo.auth.gr/the_seisnet/WEBSITE_2005/bulletins_en.html, *Geophysical Laboratory, Aristotle University of Thessaloniki*, Greece

International Seismological Centre (2006), On-line Bulletin, http://www.isc.ac.uk/Bull, *International Seismological Centre*, Thatcham, United Kingdom

Klein, F. W. (1985), User's guide to HYPOINVERSE, a program for VAX and Professional 350 computers to solve for earthquake locations, *U.S. Geological Survey Open-File Report 85-515*, 53 p.

Lahr, J. C. (1980), HYPOELLIPSE/MULTICS: A computer program for determining local earthquake hypocentral parameters, magnitude, and first motion pattern, *U.S. Geological Survey Open-File Report 80-59*, 59 p.

Lee, W. H. K. and Lahr, J. A. (1972), HYPO71: a computer program for determining hypocenter, magnitude and first motion pattern of local earthquakes, *U.S. Geological Survey Open-File Report*, 100 p.

Lomnitz, C. (2006), Three Theorems of Earthquake Location, *Bulletin of the Seismological Society of America*, 96, 306-312

Moser, T.J. (1991), Shortest path calculation of seismic rays, *Geophysics*, 56, 59-67

Moser, T.J., Nolet, G. and Snieder, R. (1992), Ray Bending Revisited, *Bulletin of the Seismological Society of America*, 82, 259-288

Panagiotopoulos, D.G. (1984), Travel time curves and crustal structure in the southern Balkan region, *Ph.D. Thesis, University of Tessaloniki*, 159 p. (in Greek).

Papazachos, B.C. and Papazachou, C.B. (1997), The Earthquakes of Greece, *Ziti Publications*, Thessaloniki

Papazachos, B. C., Papadimitriou, E. E., Kiratzi, A. A. and Papazachos, C. B. (1997), The stress field in the Aegean area as deduced from fault plane solutions of shallow earthquakes, *IASPEI 29th General Assembly*, Thessaloniki, 18-28 August 1997

Papazachos, B.C., Comninakis, P.E., Karakaisis, G.F., Karakostas, B.G., Papaioannou, Ch.A., Papazachos, C.B. and Scordilis E.M. (2000), A catalogue of earthquakes in Greece and surrounding area for the period 550BC-1999, *Publ. Geophys. Laboratory, University of Thessaloniki*, Greece

Papazachos, B.C., Comninakis, P.E., Scordilis, E.M., Karakaisis, G.F. and Papazachos C.B. (2006). A catalogue of earthquakes in the Mediterranean and surrounding area for the period 1901 - 2005, *Publ. Geophys. Laboratory, University of Thessaloniki*, Greece

Papazachos, C.B. (1994), Structure of the Crust and Upper Mantle in Southeast Europe by inversion of seismic and gravimetric data, *Ph.D. Thesis, University of Thessaloniki*, 208 p. (in Greek)

Papazachos, C. B. and Nolet G. (1997), P and S deep velocity structure of the Hellenic area obtained by robust nonlinear inversion of travel times, *Journal of Geophysical Research*, 102, 8349-8367

Papazachos, C. B. (1998), Crustal P- and S-velocity structure of the Serbomacedonian Massif (Northern Greece) obtained by non-linear inversion of traveltimes, *Geophysical Journal International*, 134, 25-39

Papazachos C. B. and Scordilis, E. M. (1998), Crustal structure of the Rhodope and surrounding area obtained by non-linear inversion of P and S travel times and its tectonic implication, *Acta Vulcanologica*, 10, 339-345

Papazachos, C. B. and Karagianni, E. E. (2006), Creating a Moho depth model for the broader Aegean area using results from various geophysical studies: Problems and implications for active tectonics, *European Geosciences Union General Assembly 2006*, Vienna, Austria, 02 – 07 April 2006, (*abstract*)

Papazachos C. B., Scordilis, E. M. and Peci V. (2006), Crustal and upper mantle P- and S- velocity structure of the southern Adriatic-Eurasia collision obtained by non-linear inversion of regional travel times, *Geophysical Journal International*, (submitted).

Poupinet, G., Ellsworth, W. L. and Fréchet, J. (1984), Monitoring velocity variations in the crust using earthquake doublets: an application to the Calaveras fault, California, *Journal of Geophysical Research*, 89, 5719–5731

Waldhauser, F. and Ellsworth, W.L. (2000), A Double-Difference Earthquake Location Algorithm: Method and Application to the Northern Hayward Fault, California, *Bulletin of the Seismological Society of America*, 90, 1353-1368

Waldhauser, F. (2001), hypoDD - A Program to Compute Double Difference Hypocenter Locations, *U.S. Geological Survey Open-File Report 01-113*, 25 p.

Wessel, P. and Smith, W. (1995), New version of the Generic Mapping Tools, *EOS, Transactions, American Geophysical Union*, 76-329