NORSAR Scientific Report No. 2-2007

Semiannual Technical Summary

1 January - 30 June 2007

Frode Ringdal (ed.)

Kjeller, August 2007

6.2 Application of array-based waveform correlation techniques to the detection of the 2003 Lefkada Island, Greece, aftershock sequence focusing on the very small aperture TRISAR array

6.2.1 Introduction

The Ionian Islands region, depicted in Fig. 6.2.1, is the most seismically active area in Greece. Its most prominent geodynamic feature is the Cephalonia Transform Fault Zone (CTFZ), which terminates the Hellenic subduction zone and comprises of two main segments, the southern Cephalonia segment - CS, and the northern Lefkada segment - LS (e.g., Louvari et al., 1999).

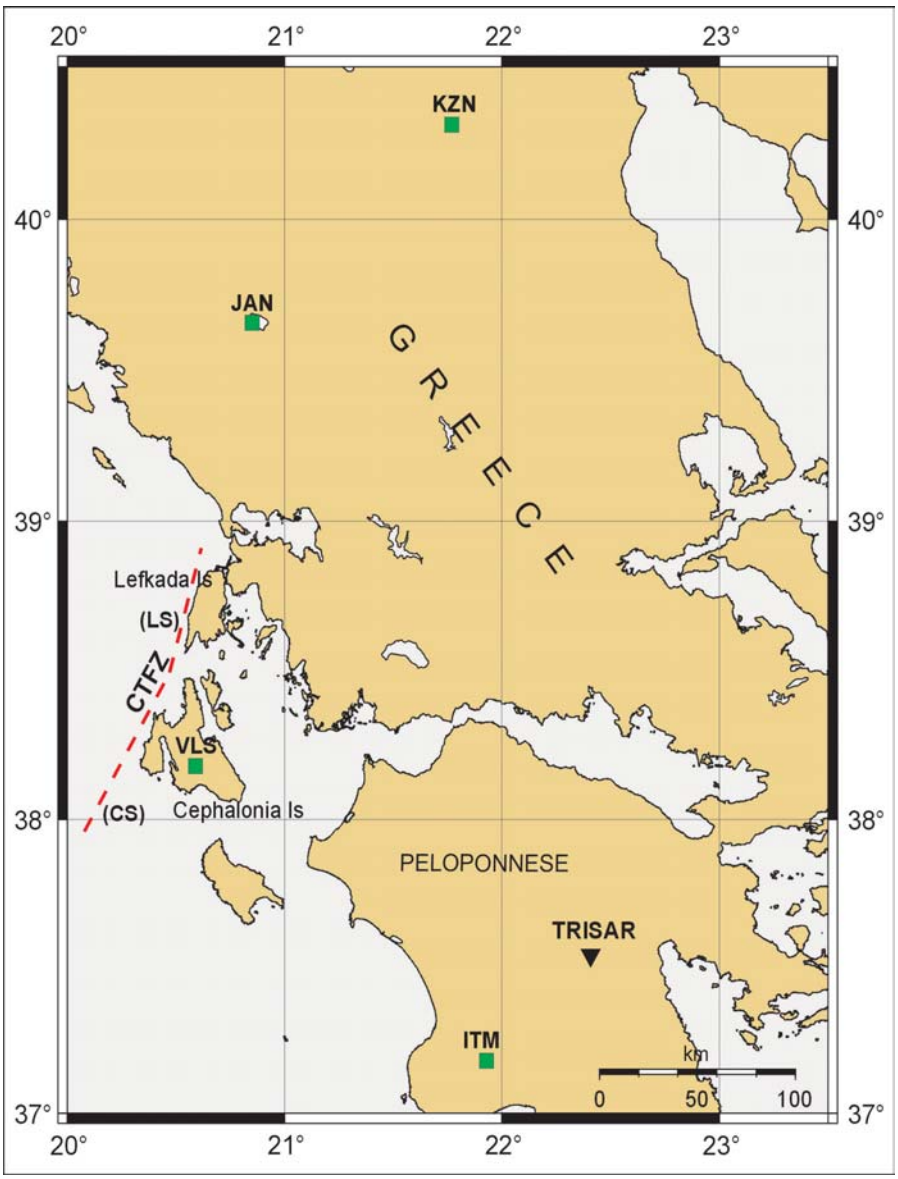

Fig. 6.2.1. Map of the Ionian Islands region and location of the Tripoli Seismic Array (TRISAR - black inverted triangle) and the GI-NOA stations used in this study (green squares). The two segments (CS and LS) of the Cephalonia Transform Fault Zone (CTFZ) are depicted with a dashed, red line.

A strong earthquake of Mw = 6.2 occurred on the northern part of the CTFZ Lefkada segment, off the NW coast of Lefkada Island, on 14th August 2003. The mainshock was followed by a vast number of aftershocks, distributed along the Lefkada Island coastline and extending southwards to the northern coasts of Cephalonia Island (e.g., Karakostas et al., 2004).

The first two days of this activity were recorded by the very small-aperture Tripoli Seismic Array (TRISAR), which is located in central Peloponnese, southern Greece. TRISAR is a 3-component, 4-site array, operated by the Seismological Laboratory of the University of Athens (Pirli et al., 2004). Three short-period instruments form an almost equilateral triangle with side length of the order of 250 m, while a reference broadband station is situated in the middle of this deployment. Routine TRISAR data processing (Pirli, 2005) involves automatic event detection and location using the DP, EP and RONAPP algorithms developed at NORSAR (Fyen, 1987;1989; Mykkeltveit and Bungum, 1984), the results being reviewed by an analyst.

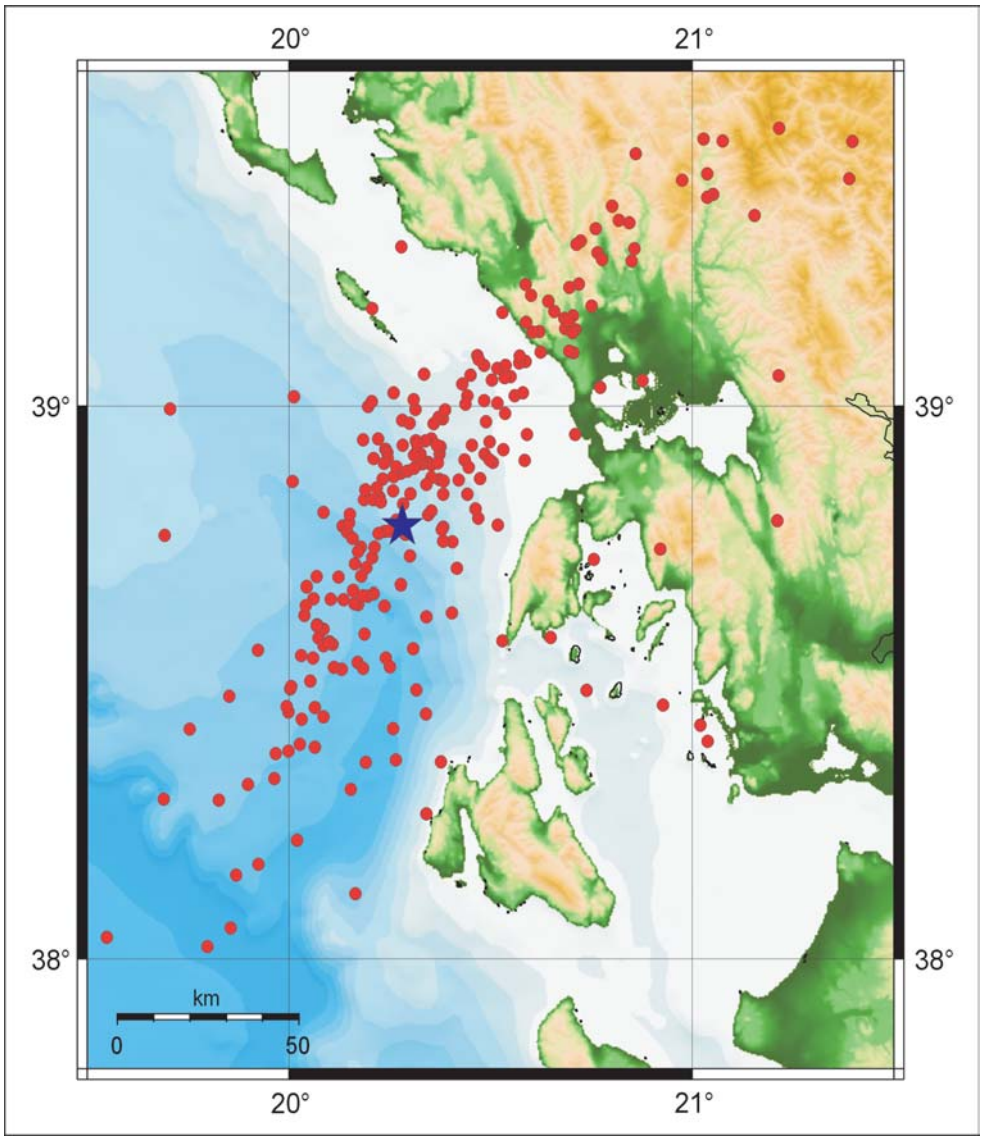

Fig. 6.2.2. Map of single-station TRISAR epicenter location estimates for 254 events of the 2003 Lefkada sequence. The location estimate for the mainshock epicenter is denoted with a blue star.

In the case of the 2003 Lefkada activity, which is located approximately 200 km from TRISAR, single-array locations were obtained for a total of 254 events: a much larger number than has been recorded and analyzed by other networks/agencies. The corresponding single-station epicenter location estimates presented in Fig. 6.2.2 resulted after careful manual review of initial automatic solutions. Even so, the obtained location estimate distribution exhibits significantly larger scatter than regional and local network solutions, due to the limited TRISAR slowness vector resolution and the larger residuals observed in the area of the Ionian Islands (Pirli, 2005; Pirli et al., 2007).

Array-based waveform correlation techniques have been applied to a subset of 244 Lefkada events, which resulted after discarding problematic data, to investigate event spatio-temporal clustering during the initial stages of the seismic sequence. Results obtained are compared with those for single, 3-component local and regional stations, operated by the Geodynamics Institute of the National Observatory of Athens (GI-NOA), and will constitute the basis for a future event relocation exercise applying relative location methods.

Moreover, further interest lies in the assessment of the applicability of full-waveform matching techniques to detect seismicity distributed over a large area.

6.2.2 Method

The 244 Lefkada events recorded by TRISAR are characterized by large variations in signal amplitude and SNR. The provisional epicenter location estimates (Fig. 6.2.2) are distributed over a far larger geographical region than could be anticipated for events resulting in highly correlating waveforms at regional distances (c.f. Geller and Mueller, 1980).

One by one, each event was assigned as a master event with a template waveform being extracted for each available individual trace. For each master event, the template waveforms were correlated against a target time-window of data surrounding each of the other events. Fig. 6.2.3 (top) shows an example of the correlation procedure, with template waveforms coloured blue and corresponding targets black. All waveforms were filtered in a frequency band providing optimal SNR prior to the correlation.

The correlation coefficient channels for the individual traces were then stacked to provide an array correlation beam, as shown in red in the lower part of Fig. 6.2.3 (c.f. Gibbons and Ringdal, 2006). Note that, even on this very small aperture array, the waveforms recorded at the different sites are sufficiently dissimilar for a significant improvement to be made in the correlation coefficient SNR by the stacking process. Note also that correlation coefficient traces from all channels, both vertical and horizontal, are included in the stack.

The array correlation coefficients between the 244 events are displayed in a similarity matrix in Fig. 6.2.4. This matrix is approximately symmetric with asymmetry resulting only from the definitions of the time-windows employed. In the typical cases, the correlation coefficient c_{ij} between events *i* and *j* was simply taken to be the maximum of the two coefficients c_{ij} and c_{ji}. Indeed, a large discrepancy between the measurements of c_{ij} and c_{ji} is a clear indication that the correlation coefficients are not providing an indication of similarity between the same segments of waveform. One such case is event #108, where the asymmetry in many correlation coefficient values was ascribed to a contamination of the waveform template by an unrelated signal.

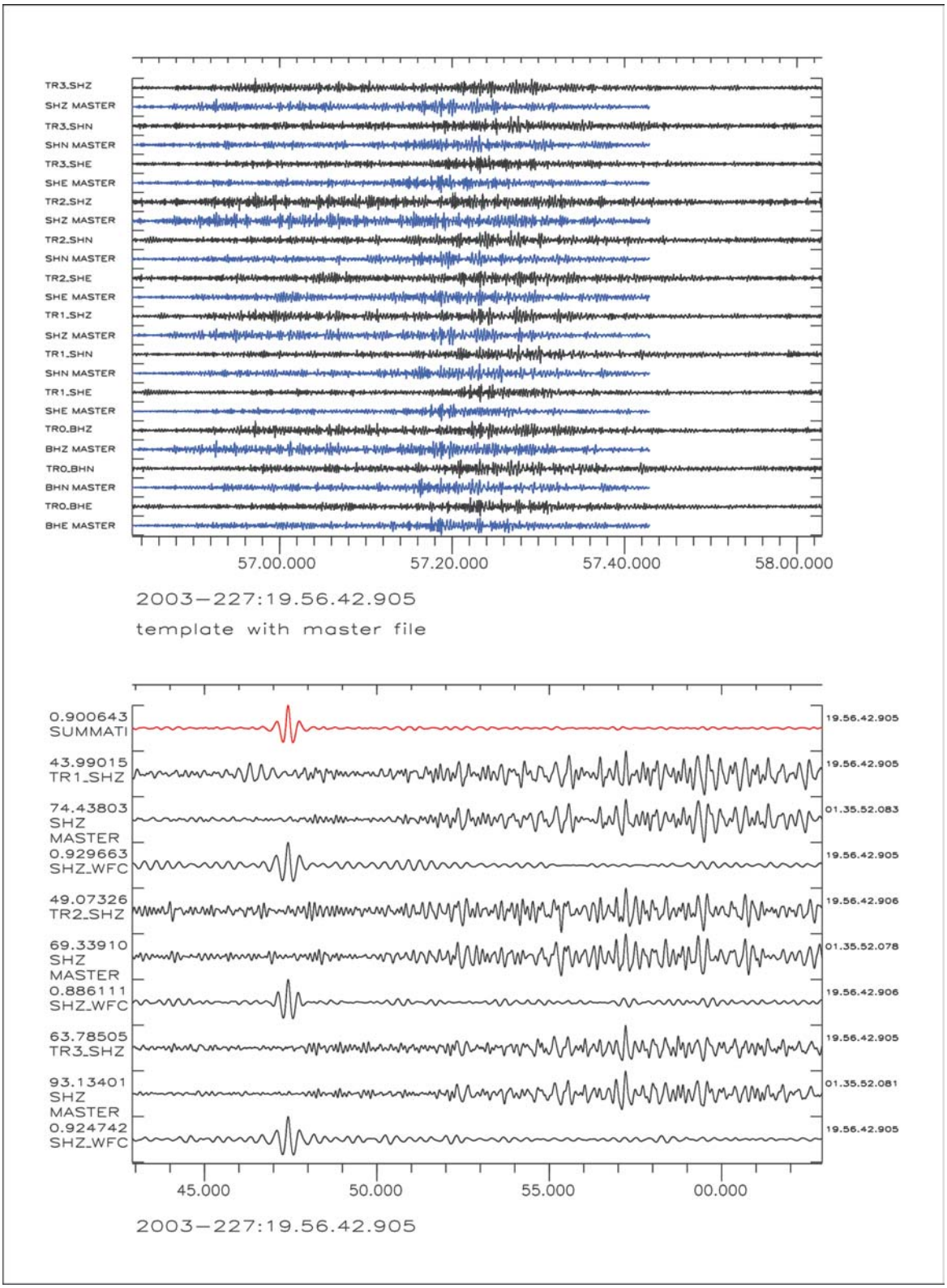

Fig. 6.2.3. Example of the waveform correlation procedure using the 3 component data of all TRISAR elements. The master event waveforms (2003-227:01:35) are coloured blue, while the slightly longer time-windows for the target event (2003-227:19:56) are coloured black (top). Each component of each array element is correlated separately (correlation trace: WFC), the final result being the correlation beam, coloured red (bottom).

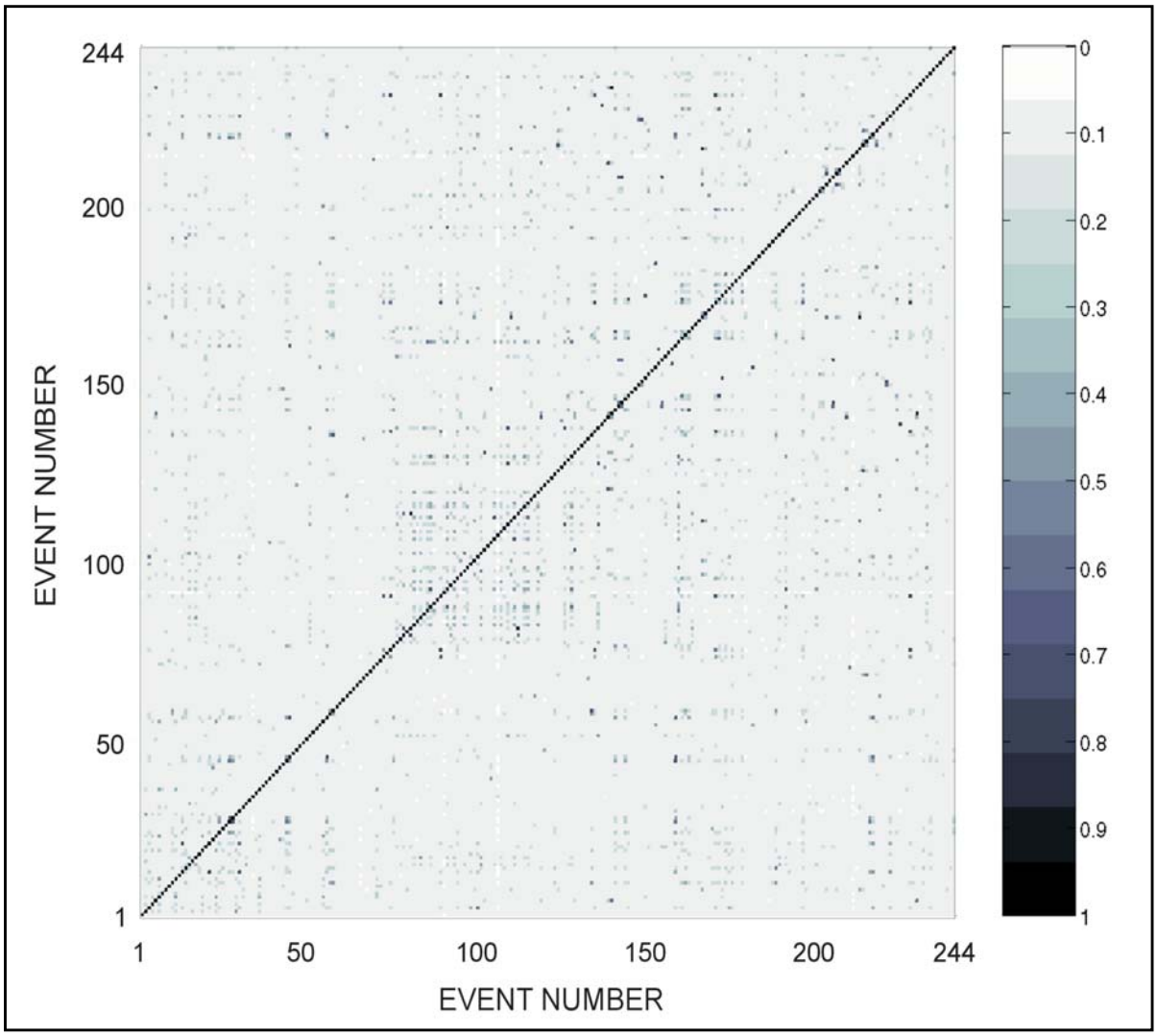

Fig. 6.2.4. Similarity matrix for the 244 Lefkada events analyzed in this study. The colour scale represents waveform correlation coefficient values. The few asymmetric features that can be observed (e.g., event #108) are attributed to overlapping successive aftershocks within the same correlation window.

In order to assess the clustering properties of the events we seek a distance function (i.e. a dissimilarity matrix). An intuitive distance function can be constructed from the correlation coefficients using

$$d_{ij} = \frac{1}{2}\left(1 - \frac{c_{ij}}{2}\right)$$

where c_{ij} and d_{ij} are respectively the fully-normalized maximum array correlation coefficient and the distance function between events *i* and *j*. However, there are problems associated with this representation. For example, the correlation coefficient is dependent upon many factors such as the signal to noise ratio and the presence of interfering signals (two events which are exactly co-located may result in a lower correlation coefficient than two events with considerable separation). In addition, the distances d_{ij} will not in general constitute a Euclidean dis-

tance matrix. The reader is referred to Saber (1984) for details. In the current application, we used the **cmdscale** function of the commercial package MATLAB to construct a Euclidean dissimilarity matrix from our observations, and additional MATLAB routines to perform the subsequent cluster analysis.

The associations between the 244 events are displayed in Fig. 6.2.5 via a dendrogram constructed using the Ward linkage method for hierarchical cluster analysis (Ward, 1963). Using a cut-off distance of 1.1, ten clusters can be identified and are represented using different colours.

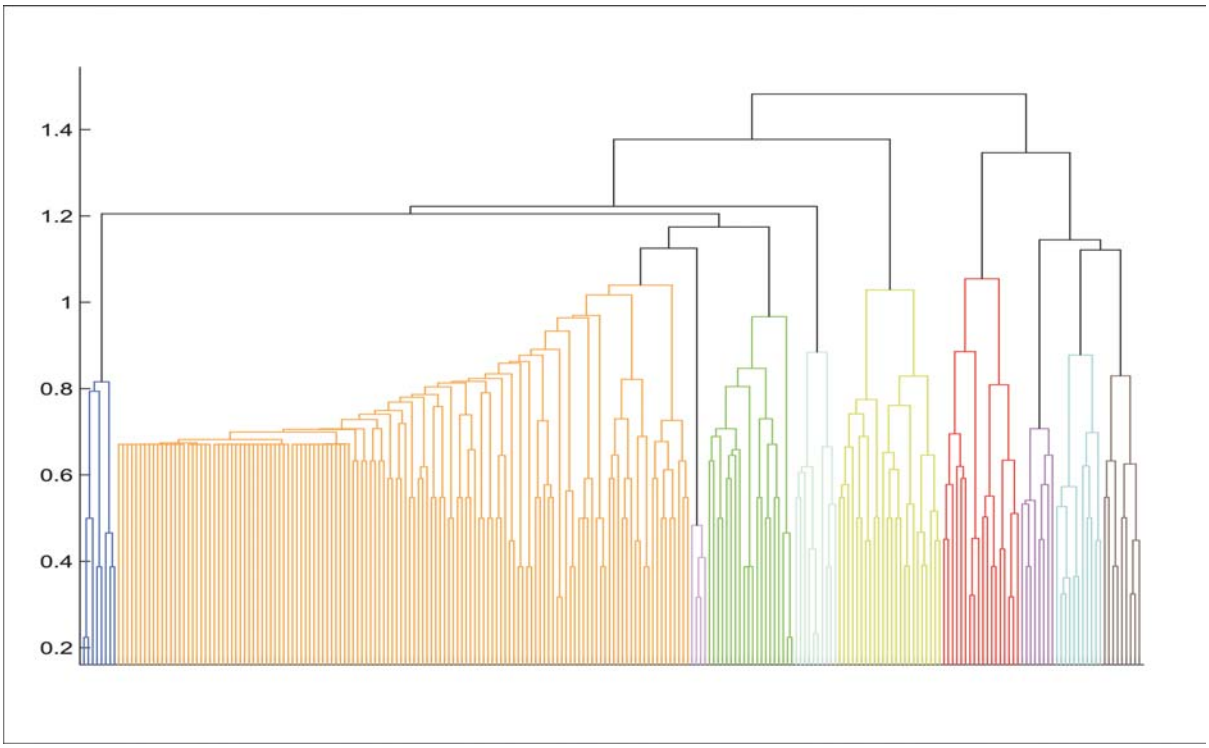

Fig. 6.2.5. Ward linkage dendrogram for the 244 Lefkada events recorded by TRISAR. Using the 1.1 distance value as cut-off level, 10 clusters are identified, represented by different colours.

6.2.3 Discussion

As expected for such a large aftershock area, the correlation coefficient resulting between the waveforms from two randomly chosen events is relatively low (with the given instrumentation and parameters chosen, correlation coefficient values of approximately 0.1 were commonly obtained - this value being quite typical for correlations of randomly chosen data segments). However, there are smaller groups of events characterized by higher degrees of waveform similarity.

Taking into consideration the size of the dataset, and to avoid chaining effects and obtain a more intuitive image of event clustering, Ward linkage was the preferred method for the construction of the dendrogram exhibited in Fig. 6.2.5. Ward clustering is an agglomerative clustering technique that assumes that the total sum of squared deviations of every point from the mean of its cluster represents the loss of information which results from the grouping of individuals into clusters. During each step, the combination of every possible pair of clusters is

considered, the resulting clusters being those whose combination exhibits the minimum increase in the error sum of squares.

The cut-off level for cluster identification from the dendrogram of Fig. 6.2.5 was set to the distance value of 1.1, maintaining a balance between the rather low level of minimum similarity and the mean value. According to this, ten event clusters have been identified, noted on the dendrogram with the usage of different colours.

Most of the resulting clusters are populated by a small number of events, reflecting the high diversity of waveforms expected over such a large aftershock area. The largest cluster, consisting of 65 events includes the mainshock, which exhibits relatively low correlation coefficients with the associated aftershocks.

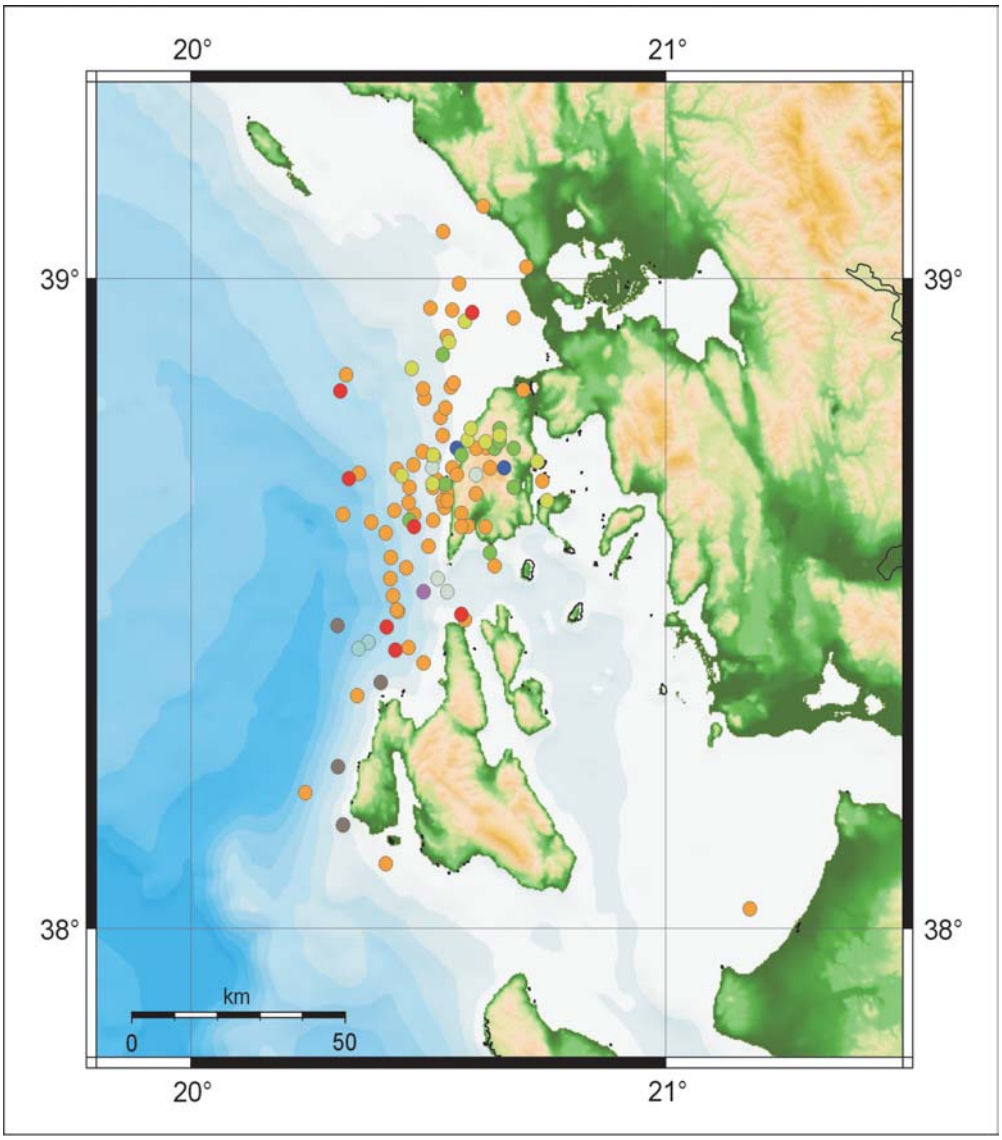

Fig. 6.2.6. Epicenter map for the subset of 108 Lefkada events as located by the ISC. The different colours correspond to those used in the dendrogram of Fig. 6.2.5 to discriminate between the obtained event clusters. Cluster #3 is not represented on this map.

Due to the rather large scatter of TRISAR location estimates apparent in the epicenter map of Fig. 6.2.2, a subset was composed according to the available 108 locations reviewed by the ISC and published in the On-Line Bulletin, to investigate the estimated spatial distribution of the clusters obtained. Fig. 6.2.6 shows a map of these reviewed epicenter solutions, events being sorted according to the clusters of Fig. 6.2.5 by using the same colours as in the dendrogram. The only cluster not represented on the map is cluster #3. Large spatial scatter is observed for events belonging to the same clusters, indicating that epicenter location estimates are inadequately constrained. This suggests that further research, involving the accurate relocation of the sequence with the application of relative location techniques, may provide better insight to the mechanisms controlling the evolution of this sequence.

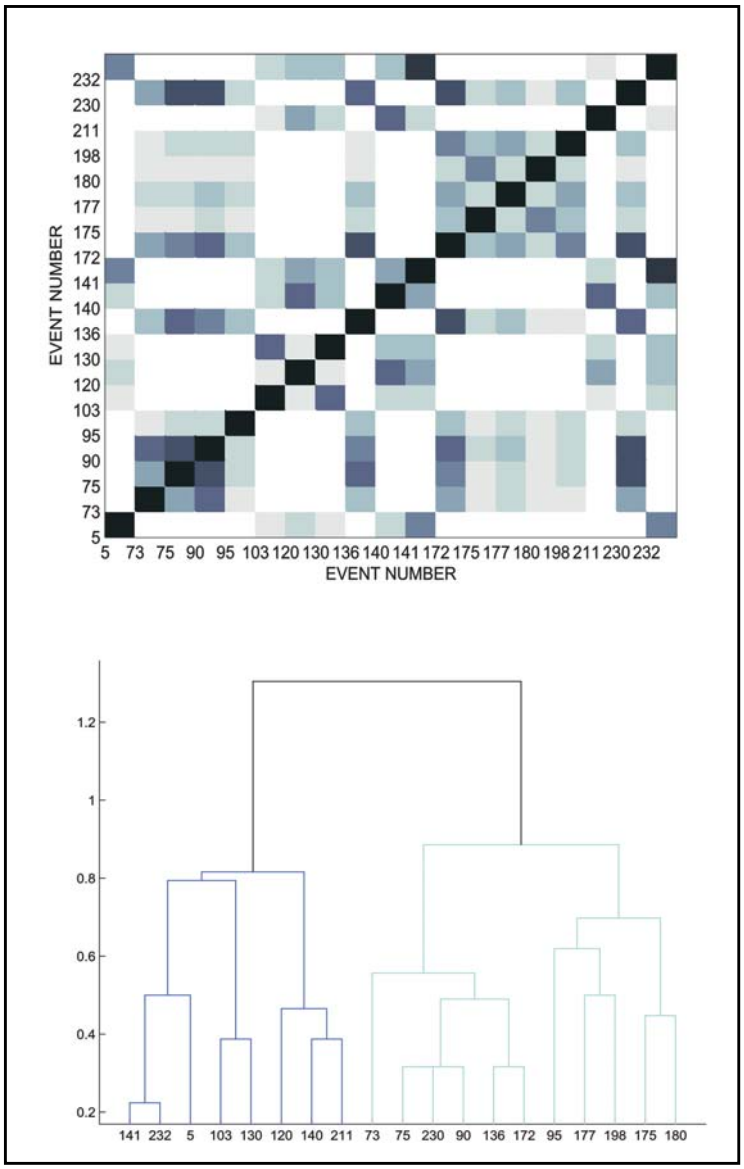

Fig. 6.2.7. Similarity matrix (top) and Ward linkage dendrogram (bottom) for the events belonging to clusters #1 and #9 of the original dataset, as recorded by TRISAR. The matrix colourscale is the same as in Fig. 6.2.4. The 2 clusters are fully consistent with those of Fig. 6.2.5, thereby noted with the same colour.

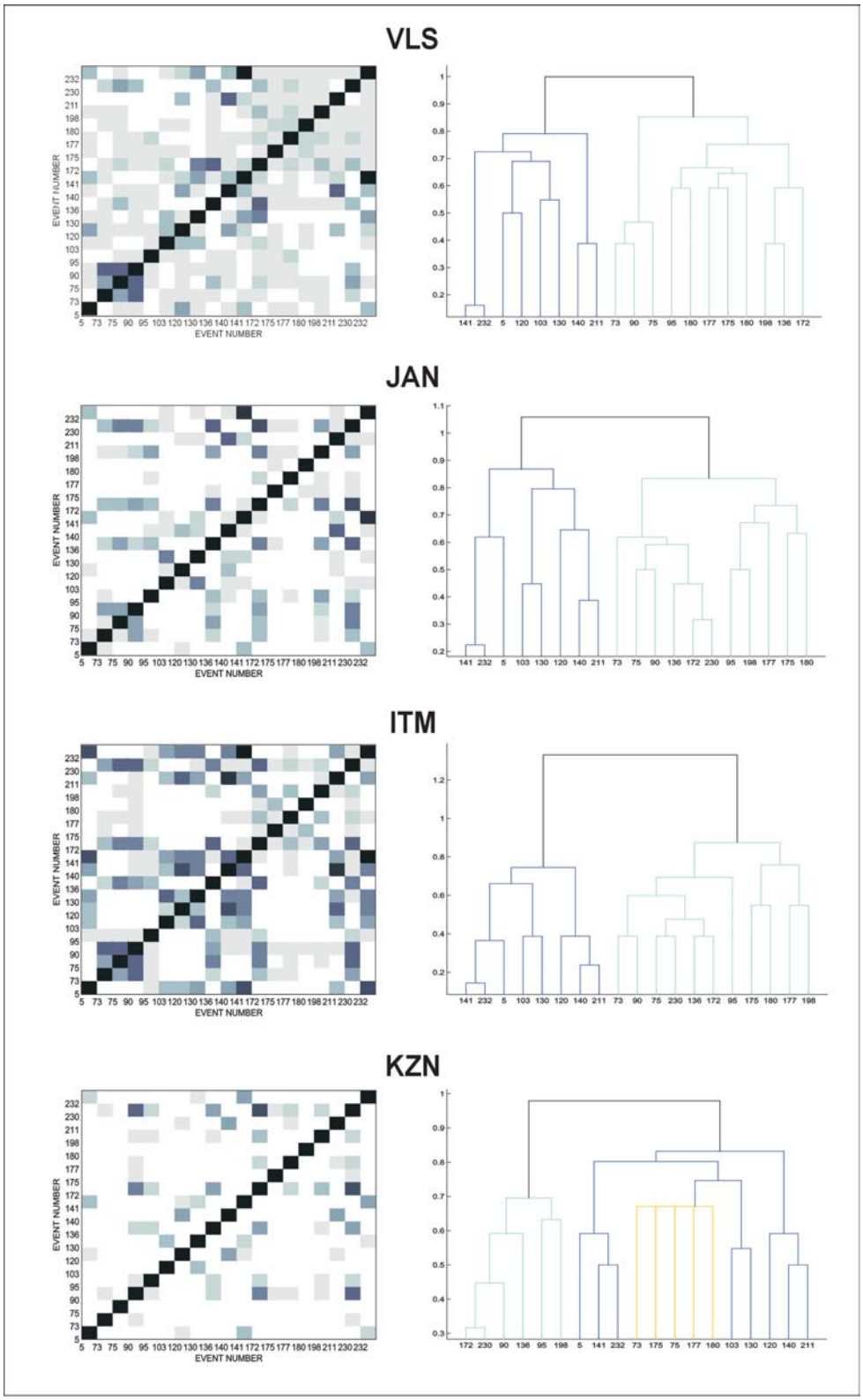

Fig. 6.2.8. Similarity matrices and Ward linkage dendrograms for the events belonging to clusters #1 and #9 (Fig. 6.2.5) for the four GI-NOA single, 3C stations used in this study (VLS, JAN, ITM and KZN), depicted in Fig. 6.2.1. Matrix colourscale is the same as in Fig. 6.2.4. Colouring of resulting clusters corresponds to that of Fig. 6.2.5. In the case of KZN, misplaced events are depicted in yellow.

In order to assess the validity of the clustering results, two clusters (#1 and #9) were selected to repeat the waveform correlation procedure using data from four single, 3-component GI-NOA stations (VLS, JAN, ITM, and KZN), located at local and regional distances (see Fig. 6.2.1). The clusters were selected so that there is a clear separation between them, according to the dendrogram of Fig. 6.2.5, and they are populated by events with varying SNR levels.

Fig. 6.2.7 (top) displays the resulting similarity matrix for TRISAR, while Fig. 6.2.8 (left) shows the matrix for the four GI-NOA single stations for the two selected clusters. The main features of the similarity matrices appear consistent in all cases. The similarity level varies depending on SNR and path effects, however the most similar events and groups can be easily identified in all cases.

Corresponding Ward linkage dendrograms were constructed for TRISAR (Fig. 6.2.7 - bottom) and the used GI-NOA stations (Fig. 6.2.8). In all cases except for station KZN the two initial clusters are clearly separated and the most similar events within each cluster are linked together, as for example in the case of event #141 and event #232. For intermediate similarity levels, the associations between the different events change for each station. Regarding station KZN, which is characterized by rather low SNR levels and is situated in the most diverse geotectonic environment with respect to all other stations, the low levels of correlation affect significantly the linkage results, assigning the events coloured with yellow (Fig. 6.2.8 - bottom right) to the wrong cluster.

6.2.4 Conclusions

Array-based waveform correlation techniques were applied to the first two days of the Lefkada aftershock sequence. The observed degrees of waveform similarity are consistent with the large extent of the aftershock area and the great diversity of associated waveforms. One limitation of the method applied here appears to be its sensitivity to the time windows used for the template and target waveforms. The cause of these difficulties is that template and target time-windows are defined for single events based upon location estimates; an iterative scheme to modify window definitions according to a matched filter detector would presumably improve the situation.

The mainshock does not appear to correlate highly with any aftershocks, belonging to a larger group of events loosely linked together. This can be attributed to the different rupture process that is associated with the mainshock.

In most cases, according to the available bulletins, events populating the same clusters appear scattered on both segments of the CTFZ and even in areas lying outside the fault zone. This suggests that the location estimates used are poorly constrained. Indeed, some events in the ISC On-Line Bulletin with a large separation between epicenters were verified manually to produce highly similar waveforms suggesting a far smaller distance between epicenters than the bulletins suggest (c.f. Geller and Mueller, 1980).

The obtained cluster pattern appears to be independent of the recording station, supporting the validity of the results. The waveform similarity suggests the possibility of obtaining accurate relocation estimates (see, for example, Richards et al., 2006) which, in turn, may be used in the future to explore further the seismicity patterns and characteristics of this seismic sequence. Following a relocation, it would also be interesting to investigate the relation between the obtained event clusters and estimated focal mechanisms.

Acknowledgements

N. Melis kindly provided the Lefkada sequence data from the seismographic network of the Geodynamics Institute of the National Observatory of Athens. Maps included in this study were created using the Generic Mapping Tools software (Wessel and Smith, 1991, 1998).

Myrto Pirli
Steven J. Gibbons
Johannes Schweitzer

References

Fyen, J. (1987): Improvements and modifications. NORSAR Sci. Rep. **2-86/87**, NORSAR, Kjeller, Norway.

Fyen, J. (1989): Event Processor program package. NORSAR Sci. Rep. **2-88/89**, NORSAR, Kjeller, Norway.

Geller, R.J. and Mueller, C.S. (1980): Four similar earthquakes in Central California. Geophys. Res. Lett., **7**, 821-824.

Gibbons, S.J. and Ringdal, F. (2006): The detection of low magnitude seismic events using array-based waveform correlation. Geophys. J. Int., **165**, 149-166.

International Seismological Centre (2001): On-line Bulletin. http://www.isc.ac.uk, Internatl. Seis. Cent., Thatcham, United Kingdom.

Karakostas, V.G., Papadimitriou, E.E. and Papazachos, C.B. (2004): Properties of the 2003 Lefkada, Ionian Islands, Greece, earthquake seismic sequence and seismicity triggering. Bull. Seism. Soc. Am., **94**, 1976-1981.

Louvari, E., Kiratzi, A.A. and Papazachos, B.C. (1999): The Cephalonia Transform Fault and its extension to western Lefkada Island (Greece). Tectonophysics, **308**, 223-236.

Mykkeltveit, S. and Bungum, H. (1984): Processing of regional seismic events using data from small-aperture arrays. Bull. Seism. Soc. Am., **74**, 2313-2333.

Pirli, M. (2005): A contribution to earthquake location in Greece with the use of seismic arrays. PhD Thesis, University of Athens, Athens, Greece, 256 p. (in Greek).

Pirli, M., Voulgaris, N., Alexopoulos, J. and Makropoulos, K. (2004): Installation and preliminary results from a small aperture seismic array in Tripoli, Greece. Bull. Geol. Soc. Greece, **XXXVI (3)**, 1499-1508.

Pirli, M., Pirlis, E. and Voulgaris, N. (2007): Mislocation vectors for the Tripoli Seismic Array, Greece, and structural effect implications from backazimuth and slowness residual analysis. Bull. Geol. Soc. Greece, **XXXX (3)**, 1234-1245.

Richards, P. G., Waldhauser, F., Schaff, D., Kim, W.-Y. (2006): The applicability of modern methods of earthquake location. Pure Appl. Geophys., **163**, 351-372.

Saber, G. A. F. (1984): Multivariate Observations, Wiley, Hoboken, New Jersey.

Ward, J.H. (1963): Hierarchical grouping to optimize an objective function. J. Am. Statist. Assoc., **58**, 236-244.

Wessel, P. and Smith, W.H.F. (1991): Free software helps map and display data. EOS Trans. Am. Geoph. Union, **72**, 441, 445-446.

Wessel, P. and Smith, W.H.F. (1998): New, improved version of Generic Mapping Tools released. EOS Trans. Am. Geoph. Union, **79**, 579.