

Earthquake-explosion discrimination using waveform cross-correlation technique for mines in southeast of Tehran

A. Kahbasi · A. Moradi

Received: 18 April 2015 / Accepted: 30 November 2015 / Published online: 7 December 2015
© Springer Science+Business Media Dordrecht 2015

Abstract The presence of man-made explosions in a seismic catalogue leads to errors in statistical analyses of seismicity. Recently, the need to monitor man-made explosions used for mining, road excavating, and other constructional applications has become a demanding challenge for the seismologists. In this way, we gain new insight into the cross-correlation technique and conduct this approach to discriminate explosions from seismic datasets. Following this, improved P-wave arrival times are used for more precise relocation. In this study, the waveform cross-correlation technique provides a reliable means for discriminating explosions which have cross-correlation coefficients (CC) of 0.6 or greater with their own corresponding stacked waveforms. The results illustrate that approximately 80 % of seismicity of southeast of Tehran, recorded by the Iranian Seismological Center (IRSC), includes events which have cross-correlation coefficients of ≥ 0.6 with their corresponding stacked waveforms. Furthermore, with improved P-wave arrival time, there is a better chance to relocate explosions precisely in the region under study.

Keywords Cross-correlation · Earthquake · Explosion · Relocation

1 Introduction

Tehran is a densely populated megacity and seismically significant located on the southern flank of the Alborz collision zone, which has been encircled by some major faults comprising the Mosha, North Tehran, North and South Ray, Parchin, Pishva, Garmsar, and Kahrizak faults. Tehran has experienced no large earthquake in the vicinity of these main active faults over 170 years. Apart from this point, we do not have a complete and comprehensive seismic catalogue for the Tehran region; however, some destructive historical earthquakes are associated with the Tehran's active faults (e.g., see Ambraseys and Melville 1982; Berberian 1994).

Seismic catalogues often include man-made contamination, for instance quarry explosions and marine shots. Seismologists assessing seismicity of a region are frequently confronted with the challenge of identifying and excluding quarry blasts from seismic catalogues (Wyss and Martyrosian 1998). A contaminated seismic catalogue is a major potential source of errors and falsifies seismicity rates in a region under investigation (Zúñiga and Wiemer 1999). Quarry blasts usually show an artificially high *b*-value that leads to misinterpretation; hence, network operators strive to identify quarry blasts. However, identifying such events in a seismic catalogue is not a simple and straightforward task (Giardini et al. 2004).

It is also worth mentioning that Tehran has been surrounded by tremendous mining activities. Several mining activities occur daily in southeast of Tehran, particularly around Bibi Shahrbanoo Mountains situated

A. Kahbasi · A. Moradi (✉)
Earth Physics Department, Institute of Geophysics, University of
Tehran, Post Box 14155-6466, Tehran, Iran
e-mail: asmoradi@ut.ac.ir

at the end of the Parchin fault. The activities are frequently recorded as seismic events which cause contamination of a seismic catalogue as well as inaccurate monitoring of microseismic events as a seismic precursor. Since explicit information about blasts' occurrence and characteristics fails to be obtained through collaboration with the mine operators, the current study tries to employ a simple technique to instantly identify and discriminate explosions from earthquakes over the region under investigation.

2 Data and method

In this study, procedures are applied to two types of waveform datasets recorded by two different networks. Since the maximum magnitude for recorded explosions in the region is 2.5, the used data have a magnitudes less than 2.5 and consist of events recorded in the Tehran City Seismic Network (TCSN) and Iranian Seismological Center (IRSC) permanent networks from June 2006 to January 2013. Figure 1 represents the location of the seismic stations as well as the tectonic map of the region. Data acquisition was performed at sampling rates of 50 and 75 samples per second for the IRSC and TCSN networks, respectively.

2.1 Tehran City Seismic Network

The first network is the TCSN operated by Tehran Disaster Management and Mitigation Organization (TDMMO; <http://www.tdmno.ir/>). The TCSN consists of 13 three-component short-period seismographs established in June 2004 to monitor the stress field and seismicity of Tehran city and its vicinity. The first step in data provision includes visual discrimination of explosions from earthquakes recorded by the TCSN network. Visual inspection was carried out through recognition of waveforms recorded by stations near the mines which are devoid of S-wave and by dispersed Rg phase of waveforms recorded by stations located away from the mines (Fig. 2). Båth (1975) introduced using the presence or absence of the dispersed Rg as a discriminant to distinguish explosions from earthquakes. The Rg phase as a result of shallow events exhibits predominantly large amplitude on the vertical components of stations. In addition, the seismic wave velocity would be between 2.98 and 3.12 km/s at periods between 8 and 12 s (Press and Ewing 1952). Based on the absence of S-wave and the presence of dispersed Rg phase, 8 earthquakes and 606 mine explosions were visually discriminated to create a waveform dataset for the TCSN network across the studied region. Records of such stations and their Z-components were chosen. Afterwards, the original data

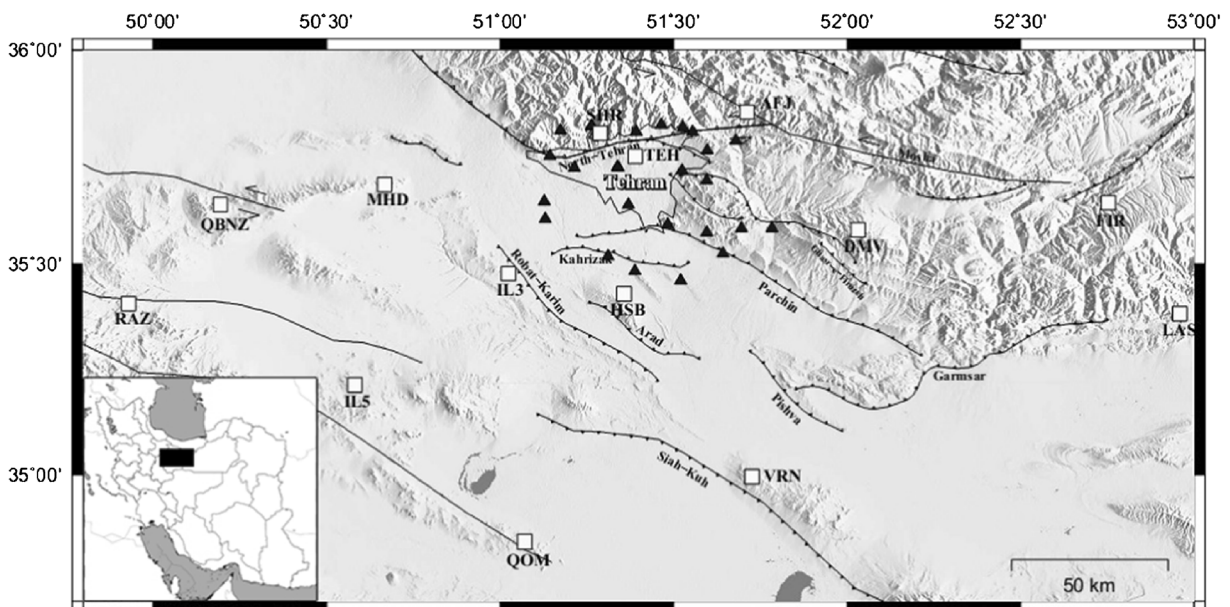
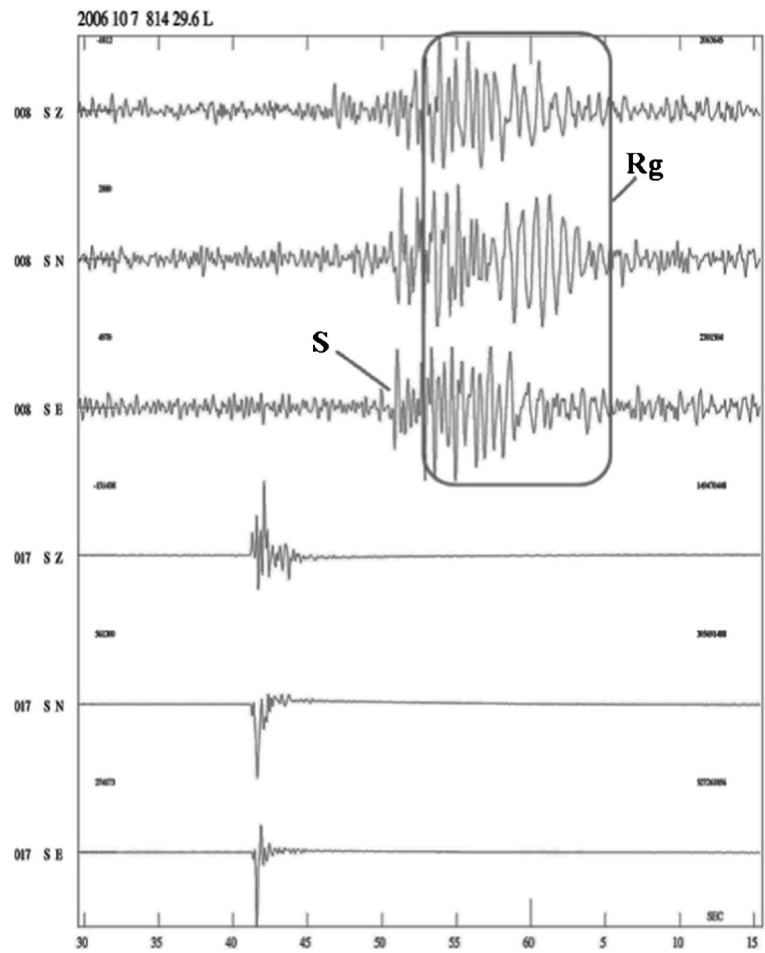


Fig. 1 Seismic stations whose records were used in this study. The *white squares* are the IRSC stations, and the *black triangles* are the TDMMO stations operating in the region

Fig. 2 Three components of two explosions recorded by stations located away from the mines (*top*) and near the mines (*bottom*). For waveforms recorded by stations located away from the mines, the dispersed Rg is an indicator to recognize explosions, and for stations located near the mines, explosion waveforms are devoid of S-wave (Yaminifard and Moradi 2012)



were converted to the SAC format (<http://www.llnl.gov/sac>) and each header was then updated by the P-wave arrival time of the TCSN network's seismic catalogue. To create reference events dataset for the TCSN network, we stacked several determinate explosions discriminated visually for each station to produce TCSN stacked dataset.

2.2 Iranian Seismological Center

The second seismic network, IRSC (<http://irsc.ut.ac.ir>), is the national seismic network operated by the Institute of Geophysics, University of Tehran (IGUT; <http://geophysics.ut.ac.ir/En/>) consisting of 105 seismic stations across the country. In order to choose recorded events from the IRSC network database, we utilized a statistical criterion suggested by Wiemer and Baer (2000) to determine the explosion-prone region. This algorithm has been defined based on the fact that quarry

blasts are occurring during daytime hours. High ratio (greater than 1.5) indicates quarry-rich areas (Wiemer and Baer 2000). Figure 3 illustrates the ratio of daytime and nighttime events (Rq). Seismic catalogue of the IRSC network was used to reveal the quarry-rich area. Based on obtained results, the region was limited from 51.42° E to 51.62° E and from 35.56° N to 35.76° N, an overall area of approximately 400 km². Eventually, we chose events limited by the mentioned area and repeated data processing steps for data recorded by the IRSC network. The number of waveforms was 338 events in southeast of Tehran, recorded at vertical-component stations of the IRSC network, from 2006 to 2013. Similarly, we needed to provide a dataset as an event reference for the IRSC network. We made the stacked waveform dataset for the IRSC network on the basis of common explosions between the TCSN and IRSC datasets. Figure 4 exhibits an instance of how the stacked waveform dataset was provided.

Fig. 3 Map of the ratio between daytime and nightly events in the studied area. Explosion-prone area was determined by the seismic catalogue of the IRSC network. The ratio is a statistical criterion suggested by Wiemer and Baer (2000). The *black triangles* are the Tehran City Seismic Network (TCSN) stations, and the *white squares* are the Institute of Geophysics, University of Tehran (IGUT), stations. The *black square* is the region under investigation

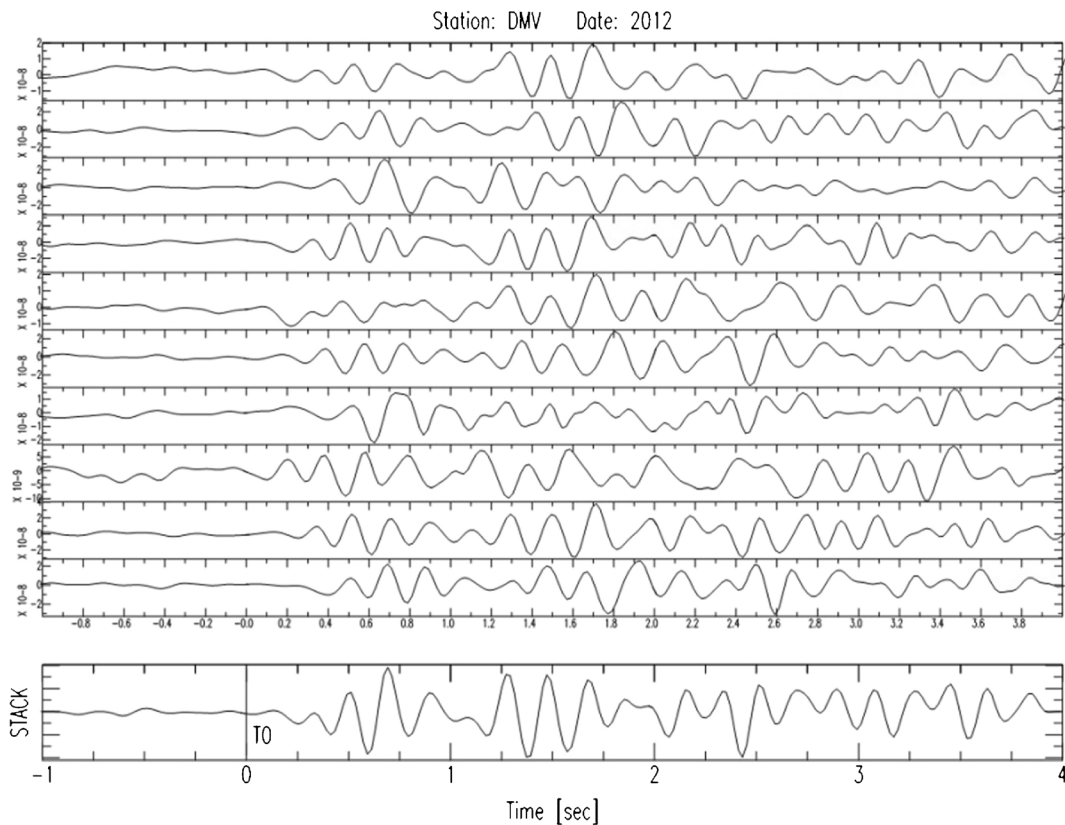
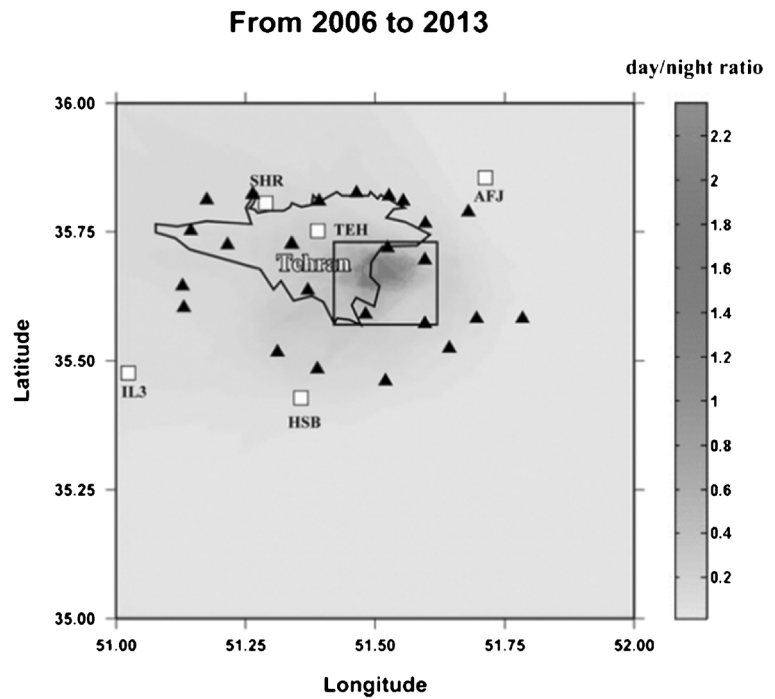


Fig. 4 An example of stacked waveform for DMV station in 2012 (*bottom*) to form a reference trace for cross-correlation using 10 well-recorded explosion waveforms (*top*)

2.3 Cross-correlation

Similar waveforms are produced when events are co-located and share the same ray paths between source and receivers (Schaff and Waldhauser 2005). However, waveforms can be similar due to trade-off between the path and source, even if events are not co-located. Thus, in a small area, the cross-correlation method can be a reliable technique to find similar events.

Several studies have been established to reveal the utilization of combined waveform cross-correlation and relative location techniques (Deichmann and Garcia-Fernandez 1992; Shearer 1997; Waldhauser and Ellsworth 2000; Waldhauser et al. 1999; Schaff et al. 2002). In this article, a semi-structured approach was chosen to identify explosions from a seismic catalogue and to improve P-wave arrival times using the cross-correlation technique, simultaneously. Furthermore, the double-difference method was applied over excluded explosions to relocate precisely.

As far as the cross-correlation technique is concerned, it is computed by fixing a window around P-wave in the time domain. Also, both time series must have the same sample rate. The most significant parameters in cross-correlation approach are cross-correlation coefficient (CC) as a threshold value for similarity and delay time (τ). The parameters utilized in the method are illustrated in Fig. 5.

The major factors of discrimination are defined by two parameters, a cross-correlation threshold (CC) and a

window length. Understanding of how CC depends on these factors is substantial. To exemplify, shorter window lengths would lead to higher CCs than longer window lengths for the same waveforms. According to statistics, a CC threshold of 0.7 for a window length of 1 s is equivalent to a CC threshold of 0.6 for a window length of 2 s. Computing correlation at different window lengths provides appropriate CC (Schaff and Waldhauser 2005).

For the TCSN network's database, a window length of 1 s around P-wave (0.2 s before P and 0.8 after that) and a threshold of 0.6 were determined on a trial-and-error basis to increase the number of similar waveforms correlated with their own corresponding stacked waveforms, considering that events of the TCSN network have been visually discriminated beforehand. Admittedly, CC=0.7 is the most robust threshold to discriminate explosions, but there again, CC=0.6 level or even lower is sometimes a useful threshold in the case that superimposed random environmental noise exists.

Five hundred eighty-nine events were identified by cross-correlation analysis of the vertical recording components of the TCSN network's database. A total of 583 events out of 589 events were common between visual method and cross-correlation analysis (roughly 96 %).

We pursued an alternative way to choose a proper window length and threshold regarding the lack of distinct explosions' reference database for the IRSC network's database. As it can be seen from Fig. 6, the variation of CCs as a function of window lengths was

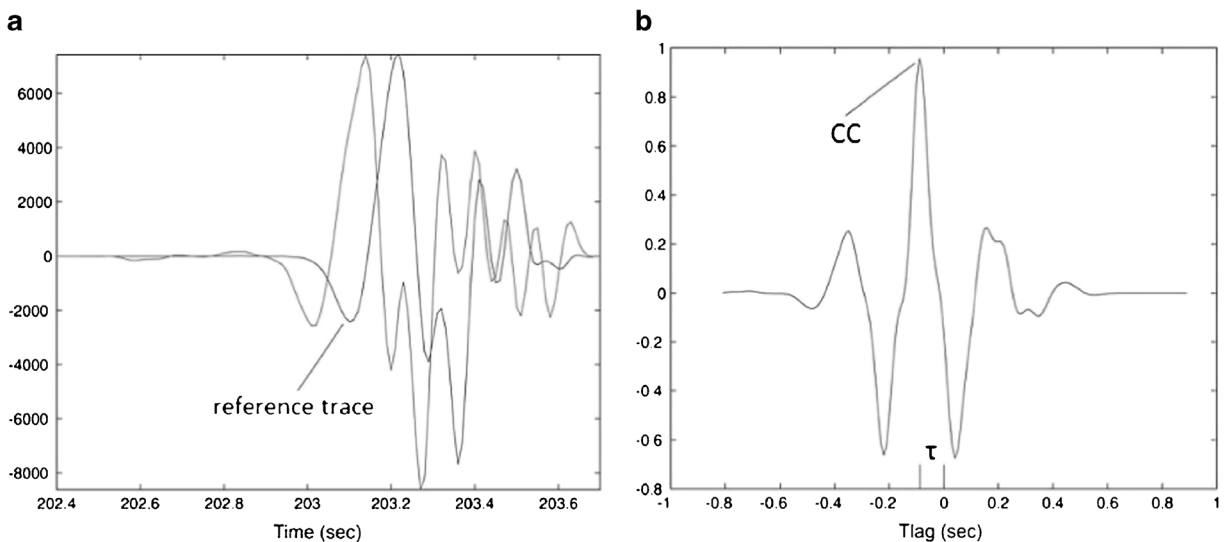


Fig. 5 **a** Overlapping a trace and a master event. **b** Illustration of two principle parameters of the cross-correlation approach, namely cross-correlation coefficient (CC) and delay time (τ)

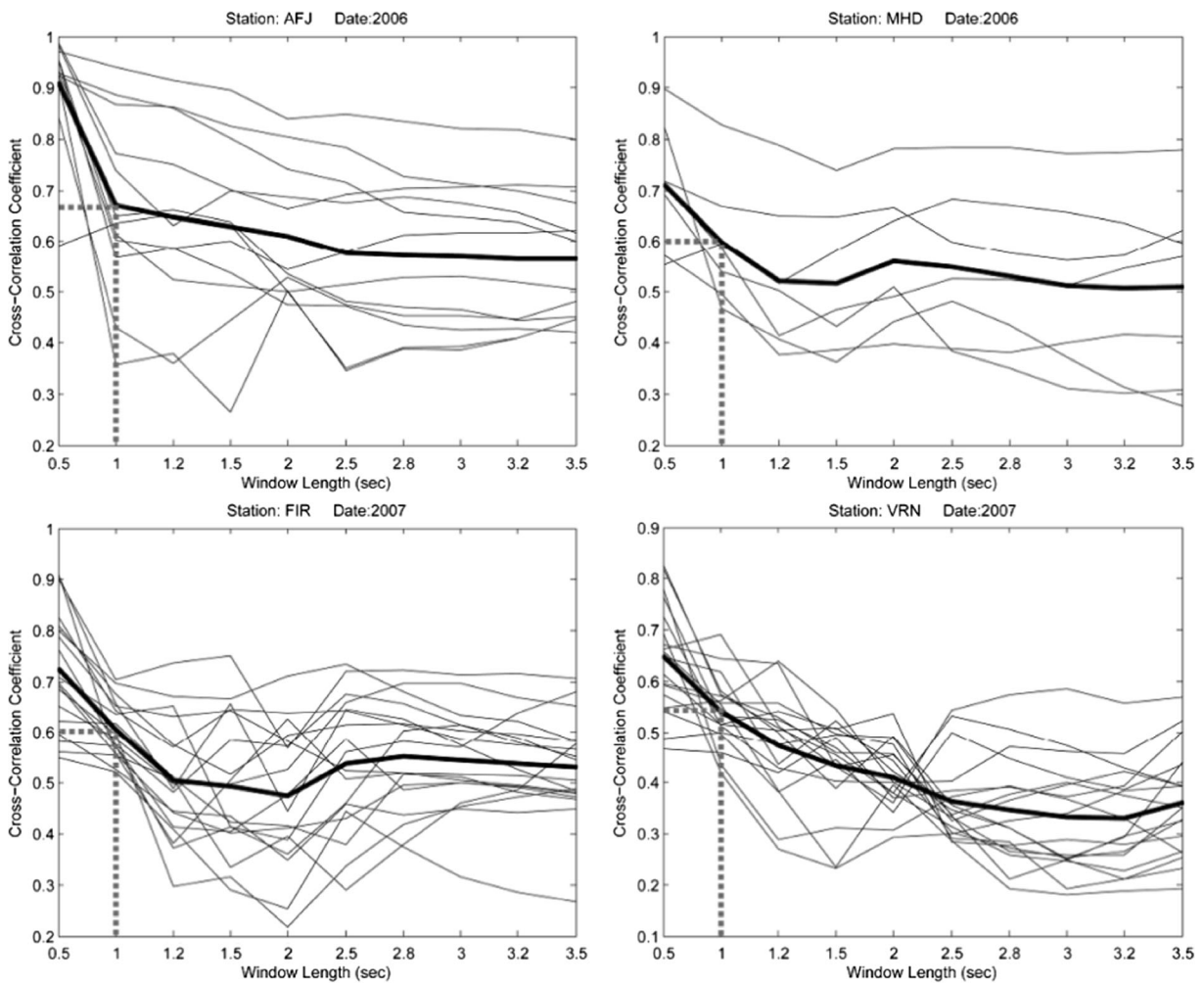


Fig. 6 Cross-correlation coefficients versus different window lengths for the IRSC network's database. The average has been shown by the *bold gray line*

provided by the waveforms utilized in stacking part to choose a threshold and a window length with a high discrimination rate.

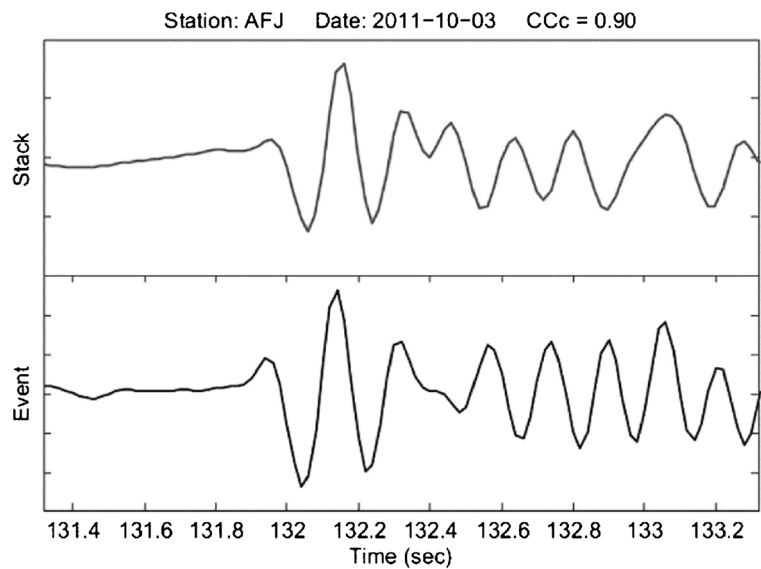
Thus, a threshold of 0.6 and a window length of 1 s were attributed to the procedure of cross-correlation. To take an example, Fig. 7 exhibits $CC=0.90$ for the window length-included P-wave. As a consequence, 274 events (just over 80 %) out of 338 events recorded by the IRSC network were discriminated as explosions in the studied area.

A histogram of the number of events as a function of hours of a day shows an obvious peak during daytime hours in the region where explosions occur (Wiemer and Baer 2000). Figure 8 shows the histogram of the number of events per hour before discrimination (Fig. 8a) and after that (Fig. 8b). The histogram is based on the IRSC

seismicity catalogue from 2006 to 2013, and the catalogue contains a larger area which consists of mining activities in southeast of Tehran (longitude $51\text{--}52^\circ$, latitude $35\text{--}36^\circ$). It is apparent from Fig. 8b that the remarkable peak has not been diminished, which means that there are still other mining activities in the vicinity of Tehran city which have influence on the seismic catalogue.

CC matrix can be used to illustrate the waveform similarity calculated from the cross-correlation analysis (Fig. 9). The CC matrix is a way to visualize the similarity between each waveform pair. As was mentioned, the similarity of two waveforms is quantitatively represented in the cross-correlation coefficient. The cross-correlation coefficients form a symmetric matrix with $N \times N$ elements, where N is number of waveforms from

Fig. 7 Illustration of the cross-correlation between a stacked waveform and an event. The figure shows $CC=0.90$ for the window length



different sources and received at the same station. In accordance with the obtained results from CC matrices, in some cases, the CC values decrease because of the breakdown in waveform similarity with increase of random environmental noises or decrease of the signal-to-noise ratio and different explosive charges.

Figure 10 displays the spatial variability of CC matrices for 2006. At near stations, such as the AFJ station, most explosions have high CC values. In contrast, at other stations, such as the RAZ station, it can be seen that waveform similarity is low. This

variation among stations imposed a challenge for determining fixed CC coefficients across the region. The challenge demonstrates how much near stations are important.

In this research, in view of the fact that the cross-correlation technique was employed to exclude explosions from a seismic catalogue, implementation of this approach led to optimization of the explosions' location over the studied region. Figure 11 shows examples of the improved P-wave arrival time observed at station AFJ after application of cross-correlation. Since

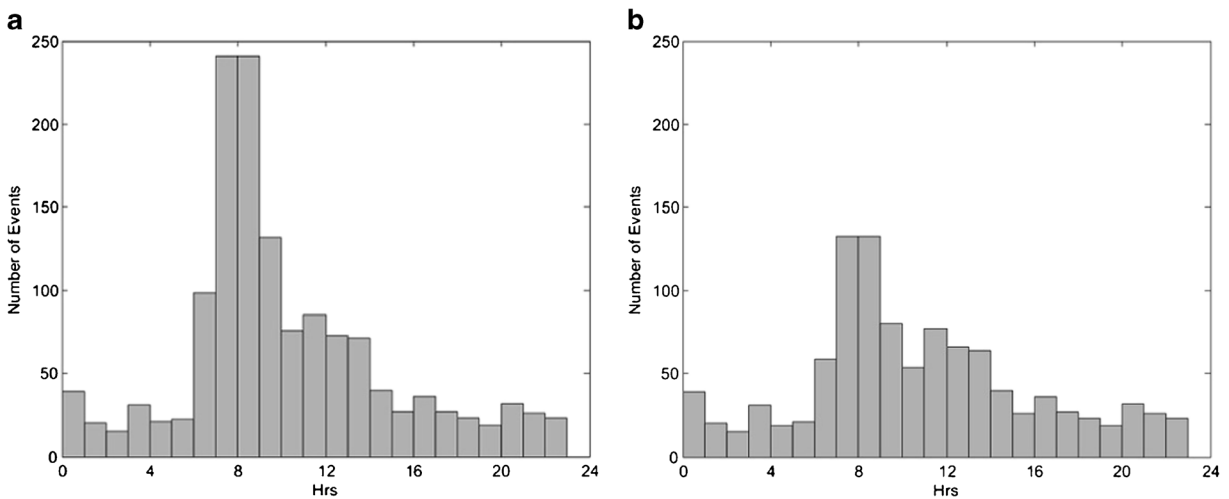
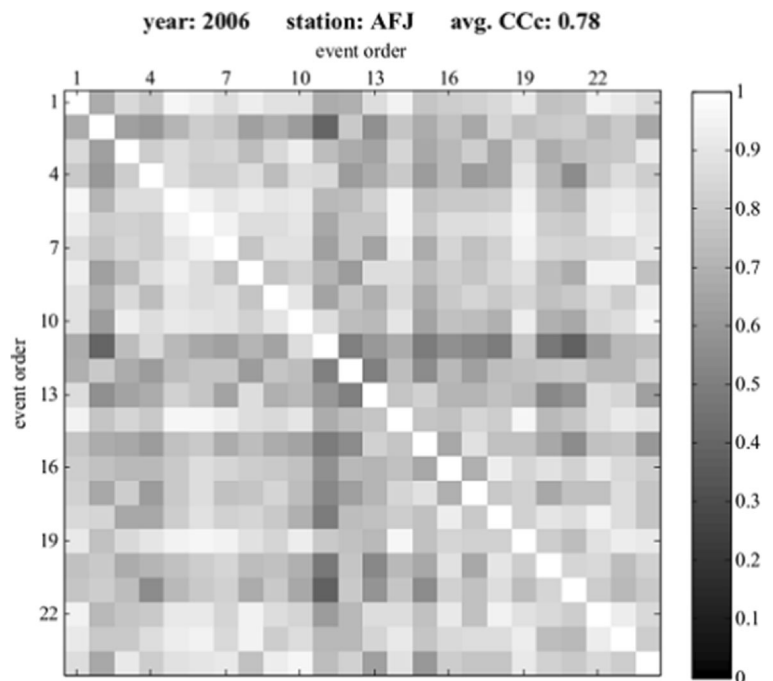


Fig. 8 The comparison between the number of events per hour, **a** before discrimination of explosions and **b** after that. The histogram is based on the IRSC seismicity catalogue from 2006 to 2013, and the catalogue contains a larger area which consists of mining

activities in southeast of Tehran (longitude 51–52°, latitude: 35–36°). Figure 8b reveals that the remarkable peak has not been diminished, which means the Tehran city is still under impact of other mining activities in its vicinity

Fig. 9 CC matrix reveals a similarity between two events. For instance, the first event is more similar to the 22th events recorded by the IRSC network



operators often fail to accurately pick too small recorded events, picking errors may occur. Figure 11a represents an example which has been re-picked by cross-correlation and has been improved approximately 0.2 s. Although the insignificant improvement seems to be negligible, 0.2 s is pretty remarkable to improve

the location of events in a small area. Figure 11b illustrates that the cross-correlation technique is under profound impact of random environmental noises and it may be impossible to emphatically assert that the P_C time has been improved by the cross-correlation technique.

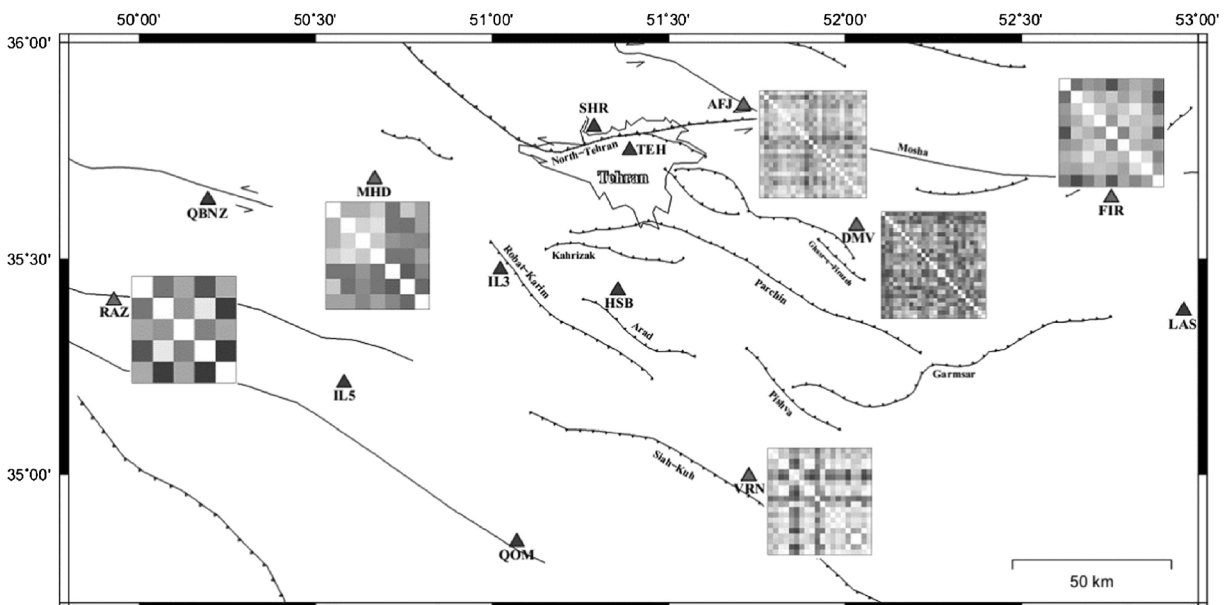


Fig. 10 Spatial variability of CC matrices among the IRSC network's stations for 2006. At near stations, most explosions have high CC values

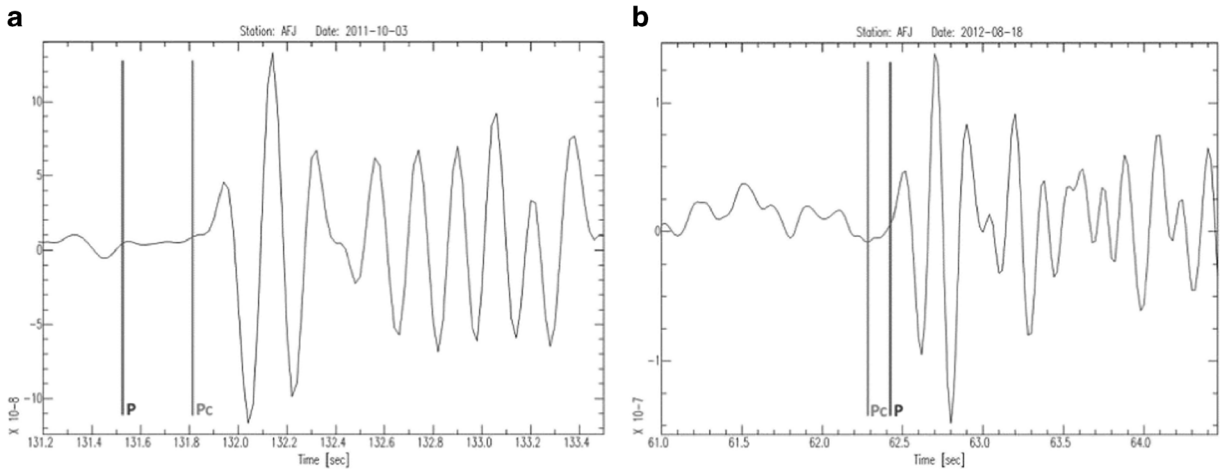


Fig. 11 Two examples of aligned waves which have been obtained from the cross-correlation technique. *P* is the manual pick before cross-correlation, and *P_C* is the optimized pick after cross-correlation. **a** Utilizing cross-correlation analysis leads to improvement in P-wave arrival time in the IRSC database. **b** Cross-

correlation technique is under profound impact of random environmental noises, and it is not possible to emphatically assert whether the *P_C* time has been improved by cross-correlation technique or not

2.4 Relocation

Location estimates often suffer from three types of errors. An inappropriate velocity model is the first source of error which can be reduced by employing relative location methods. The second is operator errors that can be handled by waveform cross-correlation technique. The CC method comparatively reduces picking errors through improving the P-wave arrival time. Inappropriate distribution of stations across a region is the last but not the least type of error which can be decreased using a remarkably dense seismic network.

Consequently, the double-difference method was employed for discriminated explosion in order to improve the explosions’ location. The approach was taken in the hypoDD algorithm introduced by Waldhauser and Ellsworth (2000) to obtain the best location by converting the arrival times into time differences between common phases of different earthquakes recorded at the same station. Meanwhile, errors related to the velocity model can be reduced by travel-time differences. One also should not overlook the fact that the number of stations is also an important factor for accuracy of the events’ location. Therefore, we combined

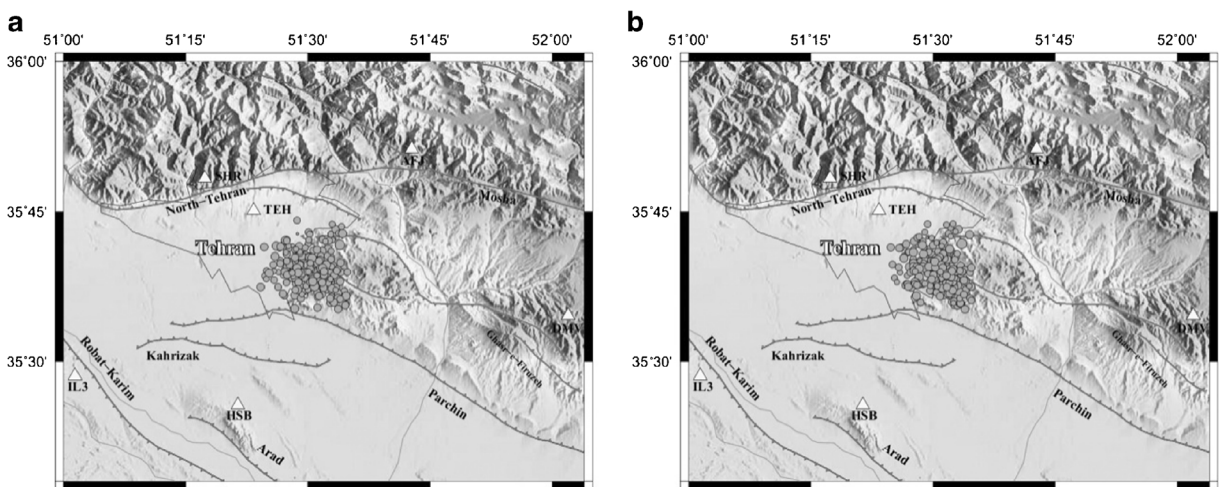


Fig. 12 Illustration of **a** discriminated explosions and **b** located explosions using the double-difference method during 2006 to 2013 in southeast of Tehran

two different datasets. Eventually, we employed the double-difference method to relocate discriminated explosions (Fig. 12). Higher density of relocated explosions reflects that the cross-correlation method has been precisely executed across the studied region.

3 Discussion and conclusion

This study resulted in the wealth of the cross-correlation technique, measured for events recorded by two types of networks (TCSN and IGUT) from June 2006 to January 2013 to discriminate earthquakes and explosions for locally recorded data in mines in southeast of Tehran. Explosions recorded by the TCSN network have been successfully discriminated (nearly 96 %) by a cross-correlation coefficients (CC) of 0.6 and a window length of 1 s. Furthermore, the results showed that approximately 80 % of seismicity recorded by the IRSC (IGUT) network across the studied area entails events that have cross-correlation coefficients of ≥ 0.6 with corresponding stacked waveforms, which means that the recorded events are explosions. Finally, discriminated explosions were relocated by the double-difference method to improve the location of explosions. Densely relocated explosions show that the cross-correlation method has been properly employed over the studied region.

We see that distribution of stations is also important, since stations near the mines are needed to visually discriminate the explosions. In case that near stations are unavailable, the cross-correlation technique, which is based on waveforms, is considered as a reliable means to discriminate quarry blasts through stacked datasets. The obtained results confirm the reliability of the methodology in identifying mining areas through CC matrices; however, the existence of remarkable peak during daytime means that there are still other mining activities in the vicinity of Tehran city which have influence on the seismic catalogue.

Identifying and discriminating explosions are significant for earthquake monitoring in the southeast of Tehran, since the region is characterized by discontinuous seismic activity with shallow, large-magnitude earthquakes and with apparent long recurrence periods. We highly recommend the use of the CC technique in the vicinity of Tehran city to discriminate the rest of the mining areas and to precisely monitor seismicity of Tehran city. Proper CC coefficient and window length as well as stacked datasets can be obtained for each mine

in order to discriminate explosions from the seismic catalogue by cross-correlation analysis.

Acknowledgments We are grateful to the Iranian Seismological Center, Institute of Geophysics, University of Tehran (IRSC, IGUT), and Tehran Disaster Management and Mitigation Organization (TDMMO) for providing the used data in this investigation. Additionally, a word of appreciation must be profoundly given to Dr. Farzam Yaminifard for his invaluable assistance in this study. We would like to couch our heartfelt gratitude to the anonymous reviewers for their constructive suggestions which helped further improved the manuscript.

References

- Ambraseys NN, Melville CP (1982) A history of Persian earthquakes. Cambridge Earth Science Series, Cambridge University Press, London
- Báth M (1975) Short-period Rayleigh waves from near-surface events. *Phys Earth Planet Inter* 10:369–376
- Berberian M (1994) Natural hazards and the first earthquake catalog of Iran, vol. 1: historical hazards in Iran prior 1900, I.I.E.E.S. report
- Deichmann N, Garcia-Fernandez M (1992) Rupture geometry from high-precision relative hypocentre locations of micro-earthquake ruptures. *Geophys J Int* 110:501–517
- Giardini D, Wiemer S, Fäh et al. (2004) Seismic hazard assessment of Switzerland. Report, Swiss Seismological Service, ETH Zurich
- Press F, Ewing M (1952) Two slow surface waves across North America. *Bull Seism Soc Am* 42:219–228
- Schaff DP, Waldhauser F (2005) Waveform cross-correlation-based differential travel-time measurements at the northern California seismic network. *Bull Seism Soc Am* 95(6):2446–2461. doi:10.1785/0120040221
- Schaff DP, Bokelmann GHR, Beroza GC, Waldhauser F, Ellsworth WL (2002) High resolution image of Calaveras Fault seismicity. *J Geophys Res* 107:2186. doi:10.1029/2001JB000633
- Shearer PM (1997) Improving local earthquake locations using the L1 norm and waveform cross correlation: application to the Whittier Narrows, California, aftershock sequence. *J Geophys Res* 102:8269–8283
- Waldhauser F, Ellsworth WL (2000) A double-difference earthquake location algorithm: method and application to the northern Hayward fault, California. *Bull Seismol Soc Am* 90:1353–1368
- Waldhauser F, Ellsworth WL, Cole A (1999) Slip-parallel seismic lineations along the northern Hayward fault, California. *Geophys Res Lett* 26:3525–3528
- Wiemer S, Baer M (2000) Mapping and removing quarry blast events from seismicity catalogs. *Bull Seism Soc Am* 90(2):525–530
- Wyss M, Martyrosian AH (1998) Seismic quiescence before the M7, 1988, Spitak earthquake, Armenia. *Geophys J Int* 124:329–340
- Yaminifard F, Moradi A (2012) Crustal velocity structure beneath Tehran based on teleseismic and mining explosion data recorded by Tehran City Seismic Network (TCSN). *J Earth Space Phys* 37(3):59–69
- Zúñiga FR, Wiemer S (1999) Seismicity patterns: are they always related to natural causes? *Pure Appl Geophys* 155:713–726