ORIGINAL ARTICLE

Development of declustered processed earthquake accelerogram database for the Iranian Plateau: including near-field record categorization

Ali Khansefid (ID) · **Ali Bakhshi** · **Anooshiravan Ansari**

Abstract In this paper, a comprehensive accelerogram database of the Iranian plateau containing 3585 data with all three components is gathered. The raw data are processed by the wavelet-based denoising method, and results are compared with the contaminated data. All the data are classified into mainshock and aftershock categories using the time and spatial window method. Afterward, the data are categorized into the pulse-like and non-pulse-like events based on the detection of velocity pulse in any of horizontal and/or vertical directions. Eventually, among 3585 data, the ones with an average shear wave velocity of top 30 m of subsurface soil profile are selected and their important ground motion parameters such as peak ground acceleration, velocity, and displacement, Arias intensity, ground motion duration, and acceleration response spectra are studied. It is observed that the intensity of vertical component of the accelerograms is almost at the same level as the strong horizontal direction or even more, in some cases. Near-field events cause more severe responses than far-field earthquakes. Generally, the mainshocks responses are slightly tougher than the aftershocks. By com-

paring the acceleration response spectra of the accelerograms with the ones obtained from the building design codes, the necessity of revising the existing standard spectra is revealed for considering the effects of mainshocks, aftershocks, near-field effects, as well as separate vertical component of the ground motions directly. Finally, the presented database, as a powerful tool, will give an interesting opportunity to the scientists and engineers who are enthusiastic to work on the Iranian plateau accelerograms.

1 Introduction

One of the most challenging parts of any structural time history analysis is the uncertainties that arise from the record-to-record variability, which basically occurred due to the lack of appropriate accelerogram database. If a strong database exists for a certain location, then it will be possible to overcome this problem via two different approaches which are far beyond the scope of this paper, namely, probabilistic methods using the existing or artificially generated accelerograms (Hu et al. 2018; Khansefid and Bakhshi 2019; Khansefid et al. 2019), or selecting the most appropriate accelerograms from the proposed database for any further deterministic time history analysis (Baker et al. 2011). As a basis of both of these methods, it is initially necessary to prepare a comprehensive and reliable database of processed accelerograms. Recently,

A. Khansefid (✉) · A. Bakhshi
Department of Civil Engineering, Sharif University of Technology, Tehran, Iran
e-mail: khansefidali@gmail.com

A. Ansari
International Institute of Earthquake Engineering and Seismology, Tehran, Iran

this task is being initiated worldwide. However, utill now, limited works have been conducted on the recorded accelerograms of the Iranian plateau.

By reviewing the past works done on the accelerograms recorded in the Iranian plateau, there has been some sporadic research in different fields that have used a limited number of accelerograms of the Iranian database. As an early work, the Building and Housing Research Center ((BHRC) (1997) of Iran reported basic accelerogram data of their accelerograph network. Many projects where conducted by different researchers (Nowroozi 2005; Ghasemi et al. 2009; Motaghi and Ghods 2012; Saffari et al. 2012) to introduce the ground motion prediction equations by using the existing incomprehensive accelerogram database. Zare and Sinaiean (2014) worked on a set of accelerograms to improve the criteria (H/V method) of soil classification. Akkar et al. (2014) prepared a comprehensive accelerogram database for the Middle East region for using in the future production of predictive models that contains a small part of the Iranian database. Recently, Zafarani and Soghrat (2017) prepared a dataset of Iranian recorded time history, containing 461 earthquakes, and reported some seismological information of records such as epicentral distance, moment magnitudes, peak ground accelerations, and etc. As it is seen, in most of these research works, only a small set of accelerograms are used to model different kinds of earthquake-related models. However, in the last one, the database itself and earthquake accelerograms properties for a set of considerable number of events are presented.

In this paper, an attempt is made to prepare a more comprehensive and also categorized database of the Iranian accelerograms.[1] In this regard, the most complete official data are collected. Then, from the existing data, the ones including all three directions of accelerograms are extracted. The adaptive wavelet denoising process, baseline adjustment, and high-pass filtering are implemented on all accelerograms in all directions. Afterward, the recorded accelerograms are separated into the foreshock, mainshock, and aftershock categories. In addition, they are categorized based on the existence of the directivity effect of the near-field earthquake. These kinds of categorization are done for the first time, in this scale, for the Iranian accelerogram database. Finally, for each earthquake set, all major parameters of accelerograms are

extracted and compared with the raw data as well as data proposed in the Iranian national seismic standard ((BHRC) 2014). It is observed that the correction method makes it possible to not only use the strong ground motions with a high-intensity level for future analysis, but also the weaker ones. Interestingly, the results show that the aftershocks can induce accelerations at the same level as the mainshocks or even higher. Moreover, in contrary to the general belief in the seismic design standards, in many cases, the peak value of the vertical component of accelerograms are higher than horizontal ones. Also, the effect of forward directivity in the accelerograms increases the peak values of recorded time histories significantly. Finally, by comparing the acceleration response spectra of the database, as the most important engineering parameter of any earthquake, with the suggested one by the national code, the necessity of a review of the existing version to consider the mainshock-aftershock as well as near-field and far-field categorization is highlighted.

2 Database

The quality of the available database plays a key role in the study of any natural phenomenon. There are some credible international sources for earthquake accelerograms. However, they contain a limited number of accelerograms corresponding to the events that occurred in the Iranian plateau. In addition, these limited data are only for well-known strong ground motions like Tabas, Manjil, Bam, and etc., and do not contain the ones with less intensity. As an official center, there is only a national institute called BHRC which prepares raw accelerograms. They started to record the accelerograms from the year 1973 after installing the first generation of analog accelerographs (SSA-1 analog devices). During the last decades, more than 10,000 accelerograms have been recorded with analog and digital instruments. The data correspond to the geographical latitude of 25.0° to 41.0° and longitudinal coordinates of 43.0 to 64.0. These recorded data include the information such as time of occurrence, epicentral information, station location, acceleration time histories in three orthogonal directions (if any exist), and moment magnitude of the earthquakes. The latter has been presented in different scales such as M_b, M_s, M_L, and M_w. With respect to the saturation problem of the first three scales, reported magnitudes of all recorded accelerograms are converted to the M_w by using

---

[1] Two hundred number of accelerograms data from the whole database is published for public as a supplementary file. The rest of the data are accessible only via email request from corresponding author

Karimiparidari et al. (2013) equations generated for the Iranian plateau. After reviewing the database, it is declared that many of these data suffer from insufficiency. In other words, all of the information including magnitude, source to site distance, focal depth, station average shear wave velocity of top 30 m of a subsurface soil profile, and 3-component accelerograms of event do not exist simultaneously. Hence, after omitting these incomplete data, a total number of 3585 records remain, each of which subtends 3-component accelerograms. The spatial distribution of the accelerogram data on the Iranian plateau and their corresponding magnitude are illustrated in Fig. 1.

The frequency of four important seismological information of existing accelerogram database is shown in Fig. 2. The most of accelerograms (more than 90%) are related to the events with the magnitude less than 6.0. Almost 70% of the accelerograms are recorded at the stations with a distance (epicentral distance) smaller than 50 km which implies that many of them are the candidate of being near-field accelerogram. The reported focal depth of accelerograms varies between 0.3 and 90 km. Nearly 65% of the data are pertaining to the earthquakes with the focal depth of fewer than 25 km. In other words, the database contains many shallow crustal earthquakes. The average shear wave velocity ($V_{S,30}$) of the accelerograms are reported only for 1424 data out of all 3585 accelerograms due to the lack of information in the seismic station. Among 985 official stations, only in 477 cases the average shear wave velocity is measured (less than 50%) (BHRC 2017). Moreover, only two events are recorded on the very soft soil ($V_{S,30}$ less than 185 m/s) based on the
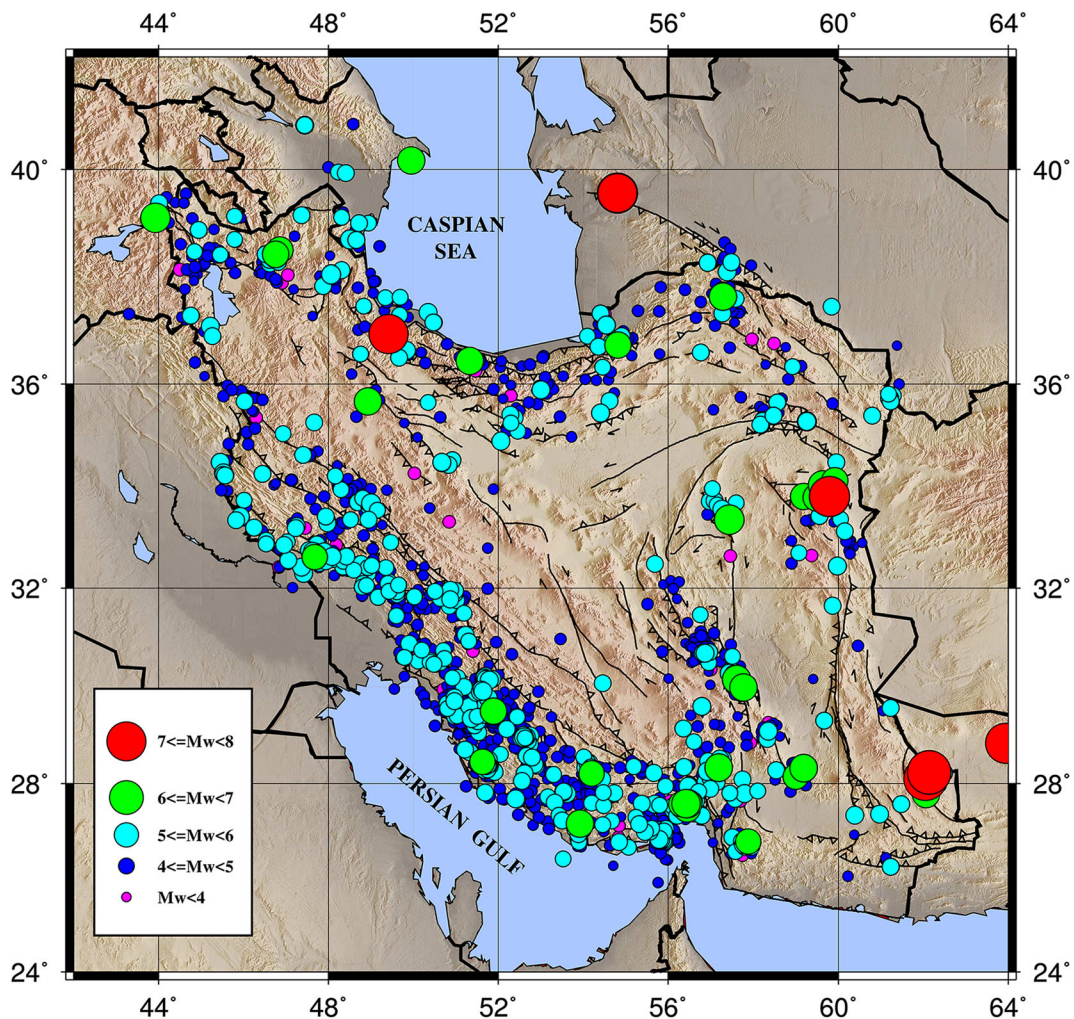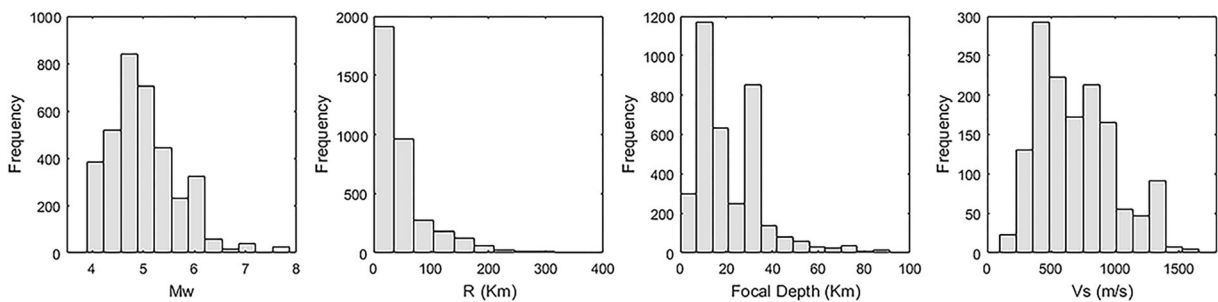

Fig. 1 Epicentral location and moment magnitude distribution of the earthquakes in the Iranian accelerogram database

Fig. 2 Frequency of important seismological information of the Iranian accelerogram database

Iranian national seismic standard (BHRC 2014). Exactly 65% of acceleration time series are recorded on the soils with the $V_{S,30}$ between 185 and 800 m/s, and the rest of data corresponds to the hard rock soil condition.

3 Denoising method

The accelerograms data provided by the BHRC are raw, many of them, more or less, are contaminated with the noises. This can be very problematic insofar as the case of Bam in 2003. Only 5 out of 24 accelerograms recorded by the BHRC were usable after denoising by traditional methods (Zare and Hamzehloo 2005). Hence, identifying and omitting the noises from the accelerograms is a very challenging and long lasting problem since the noises declared in the signals are from known and/or unknown sources.

There are different processing methods (Boore and Bommer 2005) to decompose the noises from the contaminated accelerograms. The commonly used method is to apply a frequency filter on the record, based on which the record will be adjusted by band-pass filtering and baseline adjustment. This method suffers from some major defects (Ansari et al. 2007) including inability to identify and omit the non-stationary noises, high sensitivity of the final processed signal properties to the filter parameters, and lack of accuracy in the case of identifying the noise of signals that are recorded with the analog instrument in which their natural frequency overlaps with the earthquake excitation frequency range. Therefore, this method is not suitable for denoising the accelerograms, especially the ones that are recorded with analog devices. In order to overcome this problem, Galiana-Merino et al. (2003) presented a wavelet transform-based method which was initially proposed by Donoho and Johnstone (1994, 1995), Donoho 1995, Donoho et al. (1995) in other fields. Ansari et al. (2007, 2010) suggested a modified version of this method as a two-

step adaptive wavelet correction method, capable of adjusting highly contaminated signals.

In addition to the wavelet modification, the baseline adjustment is done according to the procedure described by Iwan et al. (1985). However, in some cases, it is seen that even by using this combined method there is a long period noise on the signals identifiable in velocity and displacement time histories of weaker accelerograms. Hereupon, after all, a high-pass Acasual Butterworth filter is applied to remove these undesired signals.

In this research, to evaluate the effect of mother wavelet function, three different functions are considered including Symlet, Daubechies, and Coiflet. In baseline correction, the parameters t_1 and t_2 of the Iwan et al. (1985) method are selected as a time of the first and last passage of the accelerogram from 0.05 g. The Butterworth filter order is set to be 8 to filter the noises sharply, and its corner frequency is selected as a maximum of 0.1 Hz and the value obtained from the broadly applicable method described by Boore and Bommer (2005).

The above procedure is implemented on all 3585 accelerograms of the Iranian database in all three directions. Here, as a sample, the result of processing of an event with only the Symlet mother wavelet is presented in the following, whereas the results of all the accelerograms with different mother wavelet functions will be described in the next parts. As a sample, the Tasooj event is considered with the BHRC ID of "5852/01" recorded at the station of Tasooj ($V_{S,30}$ = 709 m/s) with the epicentral distance of 8.5 km by the SSA-2 accelerograph. This event is a mainshock with a moment magnitude (M_w) of 4.9.

The recorded accelerogram and its integrations (velocity and displacement time series) are presented in Fig. 3. The most important thing is that the acceleration time series do not show sensible noise visually. However, it is clearly seen in the velocity and displacement time histories that the signal is contaminated with noises insofar as they seem to be unphysical. Therefore, to
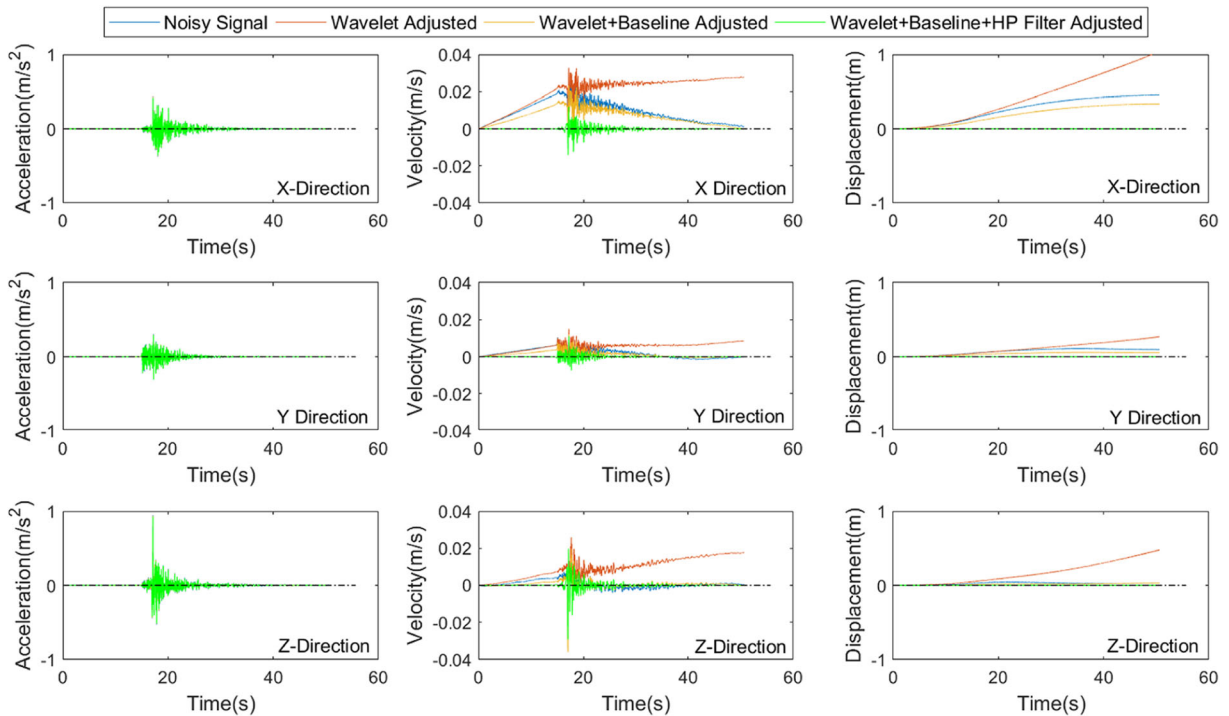
Fig. 3 Acceleration, velocity, and displacement time series of Tasooj ground motion including noisy and processed signals with different methods

make the signal applicable for engineering purposes, different adjustment methods are used such as wavelet correction, wavelet correction plus baseline adjustment, and wavelet correction plus baseline adjustment plus high-pass filtering. Among all, the latter works well and leads to the physical velocity time series which tend

toward zero at the end of the record. Moreover, the displacement signal is significantly reduced, which is a bit difficult to say that it is the real displacement time history of the ground, because of the sensitivity of final result to the baseline adjustment and filtering parameter selection (Boore 2001).

Fig. 4 Mean value of PGA, PGV, and PGD of all 3585 Iranian earthquakes before and after adjustment by three separate mother wavelet functions
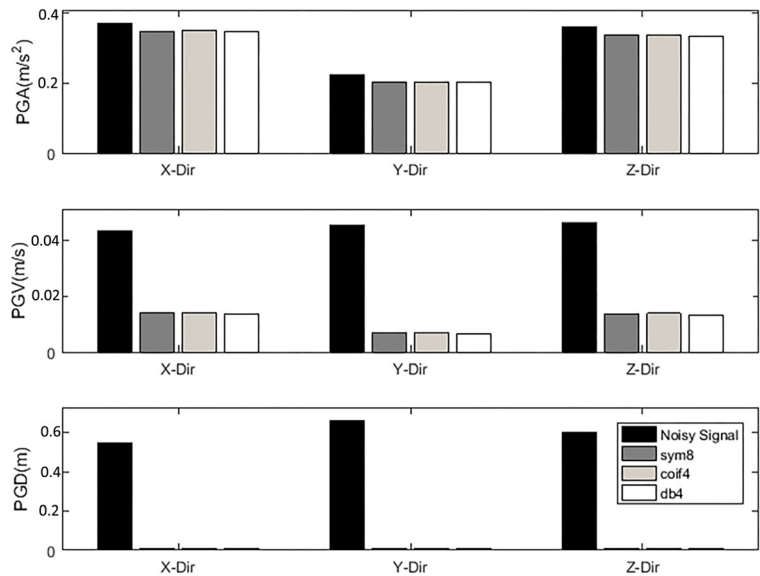
This result is only a sample from all 3585 data that shows the effectiveness of this accelerogram adjustment method. This procedure is repeated for all records and some of the most important ground motion parameters will be discussed in the following paragraphs. As a first item, the peak ground motion acceleration (PGA), velocity (PGV), and displacement (PGD) are presented (Fig. 4).

The first thing that is clearly seen in Fig. 4 is the independency of peak ground motion acceleration, velocity, and displacement of earthquakes to the selection of the mother wavelet function, at least for these three selected functions. Moreover, after denoising the signals in all directions and for all mentioned ground motion parameters, the peak values are reduced. Among all of them, PGA, PGV, and PGD show the least changes respectively equal to 8, 76, and 95%. The main reason for facing these completely different levels of changes in peak values of ground motion intensities of recorded accelerograms can be explained by focusing on the type of existing noise in the recorded accelerograms. The most seen noises in the database records are long period signals that mostly affect the velocity and displacement responses (see Fig. 3 as a sample of this noise). It should also be noted that in the case of displacement responses, the results of the processed signals are somehow questionable since the existing baseline adjustment methods

cannot exactly find the real residual displacement (Boore 2001).

The processed database prepares a robust tool to find the probability density function of the peak values of recorded ground motions leading to an overall perspective on the probability of confronting a recorded accelerogram with a specified PGA, PGV, or PGD as well as site-to-source distance (R) in the Iranian plateau. As it is illustrated in Fig. 5, two sets of results are shown. Figure 5a–c presents that among different commonly used probability density functions including Lognormal, Weibull, Exponential, and Gamma, the first one is the best estimator for the PGA, PGV, and PGD of the Iranian plateau recorded accelerograms at all recording stations determined by maximum negative Log likelihood method. The mean and standard deviation of the obtained probability density functions for PGA, PGV, and PGD are 0.275 and 0.092 m/s^2, 0.0100 and 0.0001 m/s, as well as 0.0096 and 0.0003 m; respectively. In Fig. 5d–f, the joint probability density functions of facing a recorded accelerogram with a specific site-to-source distance (R), and PGA, PGV, or PGD are presented. It is indicated that by increase in the distance (almost after 70 km) of the recording station to the seismic source, the probability of confronting considerable values for PGA (almost greater than 0.05 g), PGV
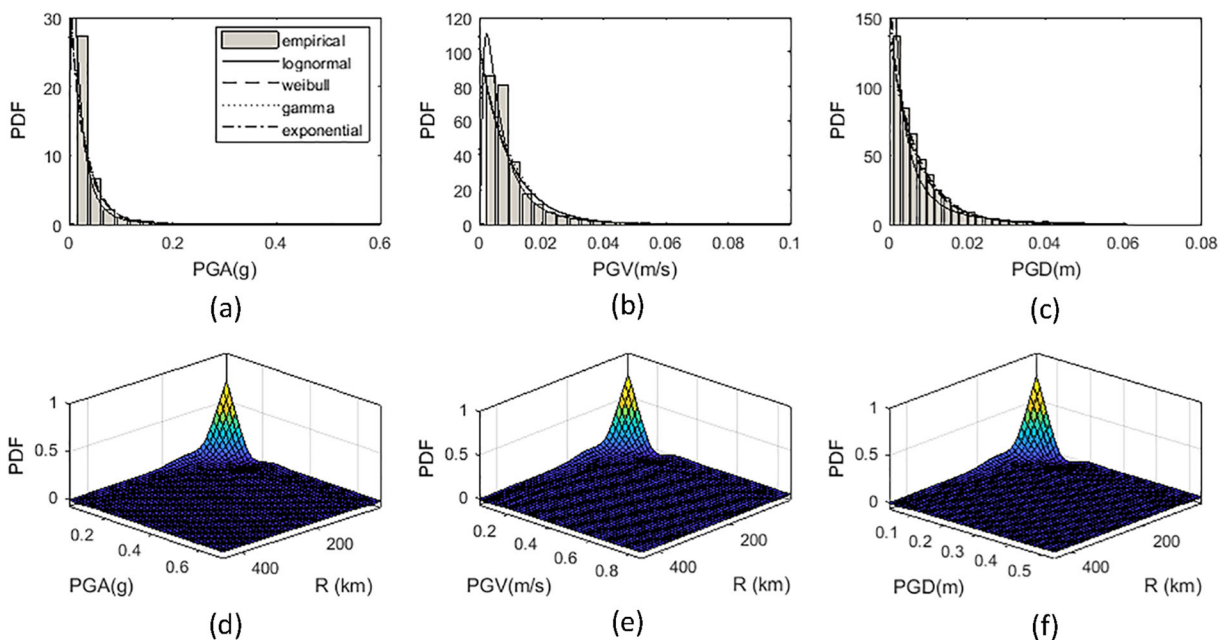
Fig. 5 Probability density functions of **a** PGA, **b** PGV, and **c** PGD of recorded accelerograms, and joint probability density functions of recorded accelerograms for the specific **d** PGA and R, **e** PGV and R, and **f** PGD and R of the Iranian plateau accelerograms
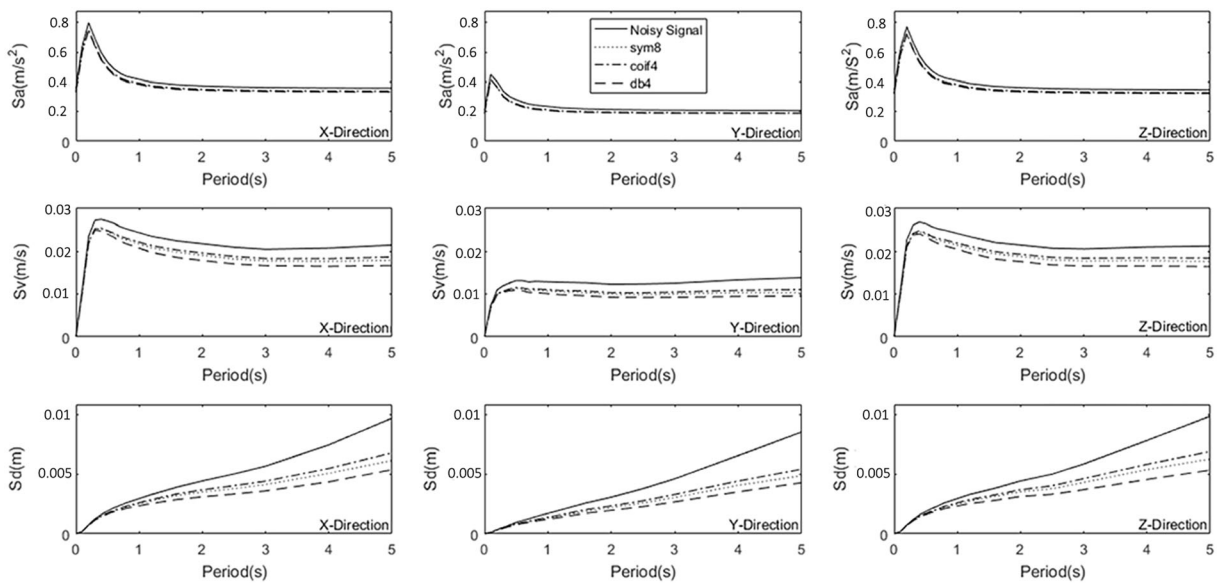
Fig. 6 Comparison of acceleration, velocity, and displacement response spectra of noisy and adjusted accelerograms for all 3585 records of database for all 3 components

(almost greater than 0.1 m/s), and PGD (almost greater than 0.05 m) decreases nearly to zero.

The last set of important parameters, which can lighten the space of accelerogram denoising process, is the response spectra of the events including acceleration, velocity, and displacement responses. These spectra are very valuable from the engineering point of view, especially the first one. As it is observed in Fig. 6, the acceleration response spectra of the noisy and processed signals do not differ significantly. However, the velocity and displacement response spectra experience more severe changes, especially in the higher periods. In other words, this type of denoising method is applicable to the common type of buildings designed based on the acceleration response spectrum. Whereas, in case of structures sensitive to the ground motion displacement, with the high natural period of vibration, usage of the result of this denoising method is not recommended. By considering the results of both Figs. 3 and 6, it is declared that by eliminating the low-frequency noises, which are the commonly observed error types in the recorded data, the more intensive discrepancy in the velocity and displacement spectrum of the noisy and processed signal is expected. As another observation, it is confirmed once more that the result of the processed accelerogram does not depend on the selection of the mother wavelet function tangibly.

4 Declustering

There are several works in the scientific world (Wang et al. 2004; Ruiz-Garcia 2012; Hu et al. 2018) that deal with aftershocks accelerograms, their declustering procedure, features, and simulations. However, this important subject has not been addressed in the literature for the Iranian plateau with the emphasis on studying the mainshocks and aftershocks accelerogram separately and comparing their properties with each other. In this study, this issue is going to be dealt with. Therefore, the first step is to decluster the earthquake records to the mainshocks and aftershocks. Mainshocks can be distinguished from aftershocks in the earthquake catalogs through different procedures (Gardner and Knopoff 1974, Grünthal 1985, Reasenberg 1985, Molchan and Dmitrieva 1992, Zaliapin et al. 2008, Baiesi and Paczuski 2004). Among all, in accordance with the Khansefid and Bakhshi (2018) suggestion, the space and time window method of Grünthal (1985) is adopted for declustering the catalog which simply separates both foreshock and aftershocks from the mainshock events without having any information about the fault shape, its mechanism, and etc. Based on this method, all records of the database are initially descendingly sorted according to their magnitudes. Afterward, events are controlled by the following time-spatial relationships to find their aftershocks and foreshocks:

$$T_{\text{foreshock}}(M_{\text{W}}) = \begin{cases} \exp\left(-4.77 + \sqrt{0.62 + 17.32 M_{\text{w}}}\right) & M_{\text{w}} < 7.8 \\ \exp(6.44 + 0.055 M_{\text{w}}) & \text{Otherwise} \end{cases}$$
$$(1)$$

$$T_{\text{aftershock}}(M_{\text{W}}) = \begin{cases} \exp\left(-3.95 + \sqrt{0.62 + 17.32 M_{\text{w}}}\right) & M_{\text{w}} < 6.6 \\ \exp(6.44 + 0.055 M_{\text{w}}) & \text{Otherwise} \end{cases}$$
$$(2)$$

$$R(M_{\text{W}}) = \exp\left(1.77 + \sqrt{0.037 + 1.02 M_{\text{w}}}\right) \quad (3)$$

where R is the spatial window in km, M_{w} is moment magnitude, and $T_{\text{aftershock}}$ and $T_{\text{foreshock}}$ are the time windows of aftershocks and foreshocks in days, respectively. In order to calculate the hypocentral distance between any two separate events, the Haversine function (Abramowitz and Stegun 1964) is implemented.

Application of this procedure leads to the 1690 mainshocks, 1529 aftershocks, and 366 foreshocks.

5 Identification of the pulse-like accelerograms

The next important classification that has never been done in this scale for the Iranian earthquakes is the pulse-like and non-pulse-like classification. In other words, this classification indicates whether an accelerogram includes forward directivity effect or not. There are many different methods that quantitatively identify pulse-like ground motions from the non-pulse-like accelerograms such as the ones presented by Mavroeidis and Papageorgiou (2003), Somerville (2003), Fu and Menun (2004), Baker (2007), and Zhai et al. (2013). In this research, Baker's (2007) method is used which works based on the wavelet decomposition. In this way, the velocity record of any signal is divided into the extracted pulse and the remaining residual signal. Afterward, the record will be classified as pulse-like or non-pulse-like in accordance with the Pulse Index value as follows:

$$P_{\text{I}} = \frac{1}{1 + \exp\left(-23.3 + 14.6(PGV\_Ratio) + 20.5(\text{Energy\_Ratio})\right)} \quad (4)$$

where PGV_Ratio is the ratio of PGV of the residual velocity record divided by the original one, and $Energy\_Ratio$ is the ratio of energy of the residual signal divided by the original one. P_{I} is a pulse indicator with a value between 0 and 1. Any signal can be classified as a pulse-like provided that P_{I} is higher than 0.85 and non-pulse-like if the P_{I} is obtained smaller than 0.15. In the case of value between 0.15 and 0.85, the record is labeled as ambiguous.

In this research, in order to classify the records of the existing database, all three components of accelerograms are taken into account, i.e., if the velocity pulse is observed at least in one component of the record, it is considered as a pulse-like accelerogram. By implementing this procedure on the whole database, it is found that there are 907 pulse-like, 1864 non-pulse-like, and 814 ambiguous records. As a sample of identification, the results of the Andimeshk event with a BHRC ID of 1527 occurred on 1994/07/31, and were recorded at Andimeshk station with M_{w} equal to 5.2, the epicenter to a station distance of 4 km, and focal depth of 48 km is presented in Fig. 7.

As it is seen, there is a tangible pulse in the velocity record of the Andimeshk earthquake in both strong horizontal and vertical components illustrated in Fig. 7. This observation implies that the fault mechanism is not simply strike-slip and there might be an additional movement in the vertical direction. It is also evident that by extracting the pulse from the velocity record, the residual signal sounds like the far-field velocity time history.

By implementing this method on all 3585 processed accelerograms, it is concluded that the period of velocity pulses of the pulse-like events varies almost from 0.24 to 3.50 s. During this process, the pulse properties were susceptible to be affected by the filter properties, if the corner period of the filter (a period for which the signal frequency content with the period greater than this value is influenced by the filtering process, i.e., reverse of corner frequency) matched with the signal pulse period. In Fig. 8, the corner period of the frequency filter obtained (Boore and Bommer 2005) and applied to each recorded accelerogram as well as the corresponding pulse period of signals are illustrated. Any data located on the red line or on the left side of the line is a candidate to confront some undesired frequency content changes in its pulse properties, because its corner period is smaller than or equal to the pulse period. As it is indicated in Fig. 8, the number of these type of data in comparison to the whole database is negligible. Therefore, the pulse property results for the whole database are not affected by the filtering process. Finally, among the pulse-like data, in 408 of them, the pulse is only observed in the horizontal direction, and in 292 cases the pulse appears in the vertical one, where for 207 events the velocity pulse is seen in both horizontal and vertical directions.

The probability density functions of the pulse characteristics (pulse period, T_{P}, and amplitude, V_{P}) are illustrated in Fig. 9. As it is seen, among the different well-known probability density functions namely

Horizontal Direction

Vertical Direction
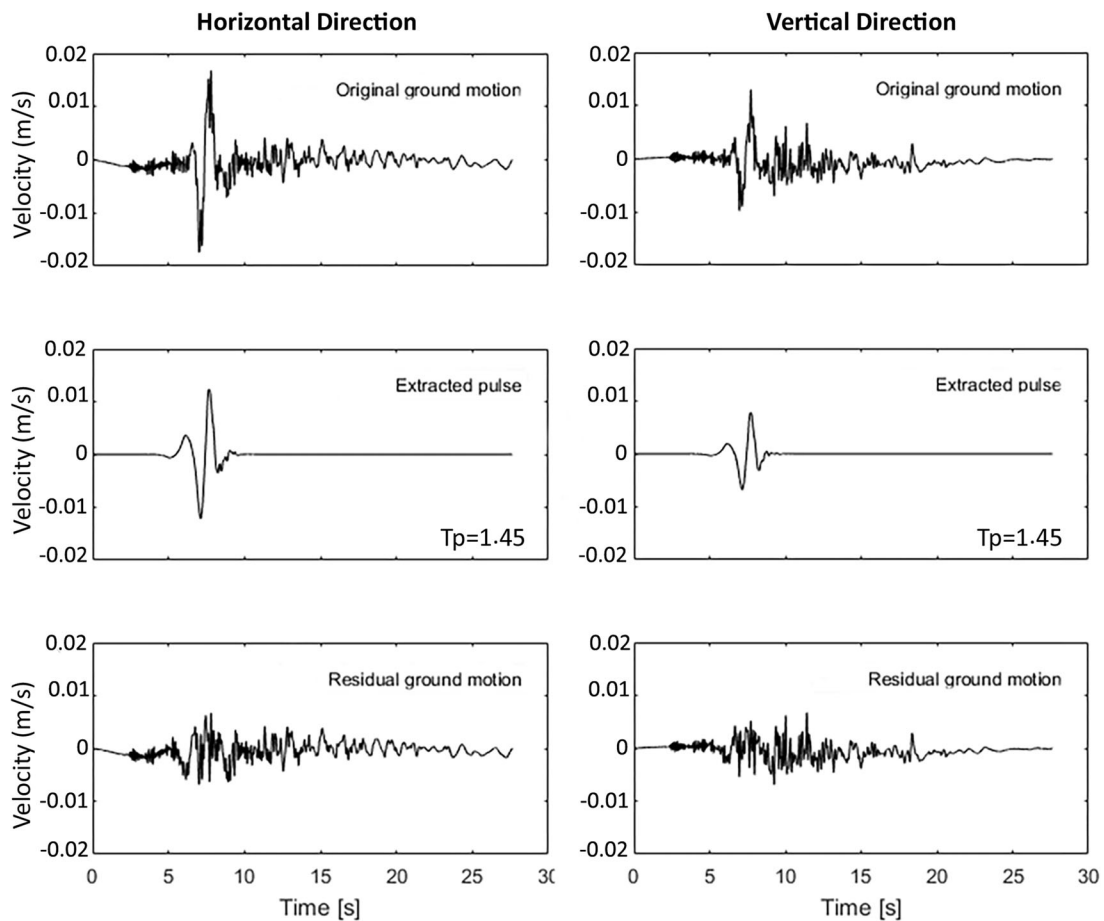
Fig. 7 Velocity time history of the adjusted Andimeshk event with its extracted pulse and the residual time series

Lognormal, Weibull, Inverse-Gaussian, Exponential, Gamma, and Log-logistic, by using the maximum negative Log likelihood method, the first one gives a good estimation for the pulse amplitude probability distribution with the mean of 0.017 m/s and the standard deviation of 0.0004 m/s. On the other hand, the Inverse-Gaussian probability distribution with a mean of 0.591 s and the standard deviation of 0.173 s is an acceptable approximation of the pulse period probability density function.

6 Ground motion parameters of accelerograms

In this part, important features of all accelerograms of the Iranian database are statistically studied for near-field and far-field as well as mainshock and aftershock categories. In this part, to achieve more helpful results, among all 3585 3-component accelerograms, only the ones with the site average shear wave velocity are considered for the evaluation. Therefore, the data number in this part is reduced to the 1424 recorded accelerograms. These data are separated with respect to their soil type and PGA based on the Iranian national seismic code (BHRC 2014) which is very close to those defined in Eurocode 8 (2004). In accordance with this standard, four different soil types are defined regarding to the site average shear wave velocity of top 30 m of subsurface soil profile, including soil type 1 ($V_{S30} \geq$ 750 m/s), soil type 2 ($375 \leq V_{S30} < 750$ m/s), soil type 3 ($185 \leq V_{S30} < 375$ m/s), and soil type 4 ($V_{S30} <$ 185 m/s). The categorization scheme of the database is illustrated in Fig. 10. It should be mentioned that in this diagram, the total number of data, 968, differs from the previously mentioned one, since some of the accelerograms in the shrunk database are categorized as an ambiguous pulse-like record or foreshocks. In addition, the database only includes soil types 1, 2,

Fig. 8 Pulse period of any signal
of pulse-like data (raw
accelerograms) and the corner
period of applied frequency filter
for processing the raw form of
signal
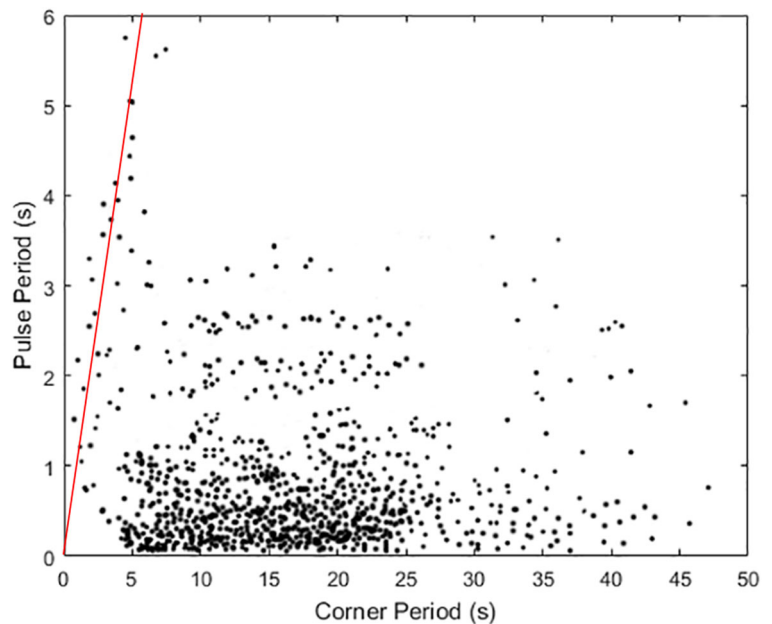
and 3. In other words, in the entire database, only two data exist for soil type 4 which are excluded from the database.

In Fig. 10, as it is expected, the number of aftershocks (507) is larger than the mainshocks (461). In addition, from another stance, most parts of the recorded accelerograms are far-field (641), almost by 2.0 times of the near-field accelerograms (327). Moreover, most of the acceleration time series are recorded on soil types 1 and 2 while there is sufficient data seized on the soil

type 3. These categorized data will be used to evaluate the most important parameters of earthquake accelerograms from the engineering point of view.

Peak ground values of the earthquake accelerograms are one of the widely used ground motion parameters by engineers. These parameters are illustrated for different categories of the Iranian accelerogram in Fig. 11. In this diagram, the effects of event magnitude, site-to-source distance, velocity pulse, aftershocks, and soil condition (V_{S30}) are taken into account. Firstly, as it is seen, the
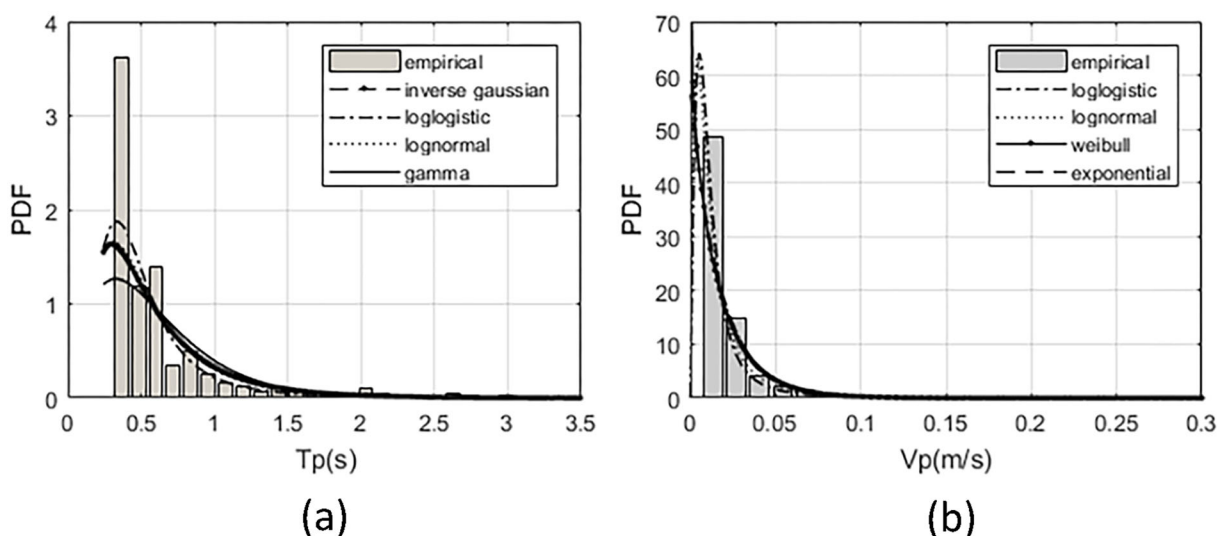
Fig. 9 Probability density functions of pulse properties of the Iranian pulse-like accelerograms **a** pulse period, T_P, and **b** pulse amplitude, V_P
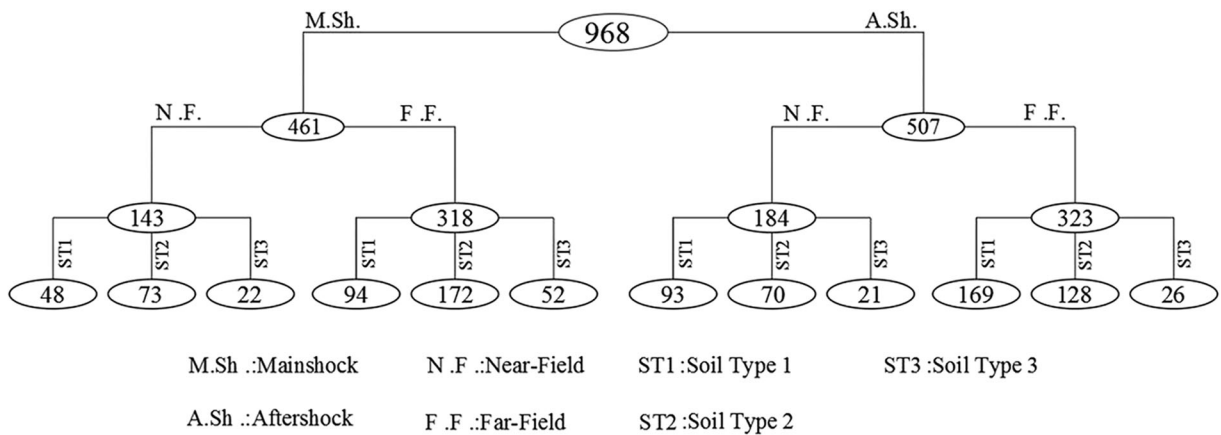
Fig. 10 Number of Iranian accelerograms categorized to the mainshock-aftershock as well as near-field and far-field by considering their recording station soil type
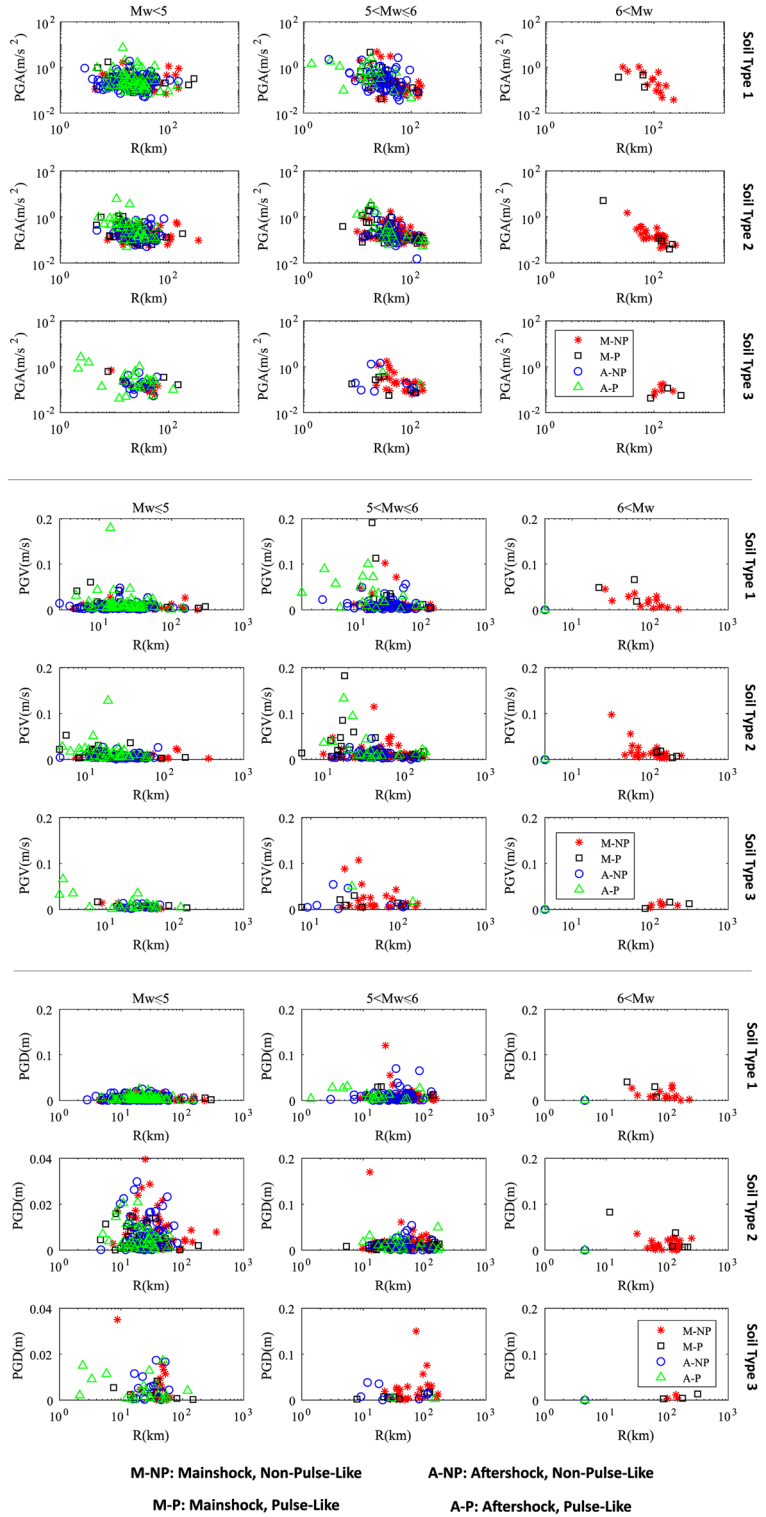
PGA of recorded accelerograms decreases significantly by the increase of the site-to-source distance (almost after 70 km) in all soil types, independent of whether they are relevant to the mainshocks or aftershocks, or to the pulse-like or non-pulse-like accelerograms. Additionally, on average, by the increase of the moment magnitude from less than 5 to the range of 5 to 6, for any specific site-to-source distance, the PGA, PGV, and PGD of recorded signals increase 52, 76, and 41%, respectively. Due to the lack of enough data for the $M_w > 6$, it is not reliable to compare the peak values. Moreover, generally, the change in the site condition from hard rock (soil type 1) to the moderate stiff soil (soil type 3), for any specific site-to-source distance, leads to the decrease of PGA, PGV, and PGD up to 25, 12, and 18%, respectively. Next important observation from the data relates to the velocity pulse effect. On average, the velocity pulse considerably affects the PGA, PGV, and PGD of the recorded earthquake signal, insofar as it increases the mean of PGA (54%) of recorded data (for all distances and magnitude) from 0.28 to 0.43 m/s^2, PGV (70%) from 0.10 to 0.17 m/s, and PGD (57%) from 0.007 to 0.011 m. These values prove the importance of consideration of the velocity pulse effect which is not addressed in many building design codes like the Iranian national seismic code (BHRC 2014).

As another interesting observation in Fig. 11, the average of PGAs of recorded data on soil types 1 to 3 are 0.42, 0.28, and 0.24 m/s^2 for the mainshocks accelerograms, while these values are 0.39, 0.29, and 0.28 m/s^2 for the aftershocks, respectively. These results show that the aftershocks can have considerable

intensity comparable to the mainshocks. This is different from the common idea about the mainshocks and aftershocks which believes the latter is weaker than the first one. This may be due to consideration of moment magnitude as an effective measure of intensity. While, the important and effective parameter on buildings' damages under earthquake excitation is the input accelerogram which is the combination of source effect (magnitude), path effect (distance), and site effect (site soil condition), e.g., there can be many events with large moment magnitude and high focal depth that lead to a small value of PGA. On the other hand, aftershocks may occur in closer locations with different focal depths, therefore causing larger PGA in compare to their corresponding mainshock. However, in this inference, it should be considered that all the aftershocks and mainshocks are not necessarily recorded at the same station. In other words, in limited cases, it is possible that a mainshock event in the database with a higher magnitude was recorded at stations located far from the seismic source, while its following aftershocks with lesser magnitude recorded at stations closer to the seismic source. Therefore, it may lead to higher PGA for aftershocks.

To investigate more about the aftershock's PGA and its relation with the mainshock's PGA, the PGA ratio of the aftershocks to the corresponding mainshock is calculated for all recorded accelerograms and the cumulative density function (CDF) of this parameter is obtained and illustrated in Fig. 12. It is clearly seen that the probability of facing an aftershock with the PGA less the corresponding mainshock is 42%. Moreover, with the low probability level (3%), it is possible to confront

Fig. 11 Peak ground motion parameters of earthquakes of the Iranian database

M-NP: Mainshock, Non-Pulse-Like   A-NP: Aftershock, Non-Pulse-Like

M-P: Mainshock, Pulse-Like   A-P: Aftershock, Pulse-Like

an aftershock with PGA four times larger than its corresponding mainshock. This totally confirms the

previous results about the mean PGA value of the mainshocks and aftershocks. This observation can be
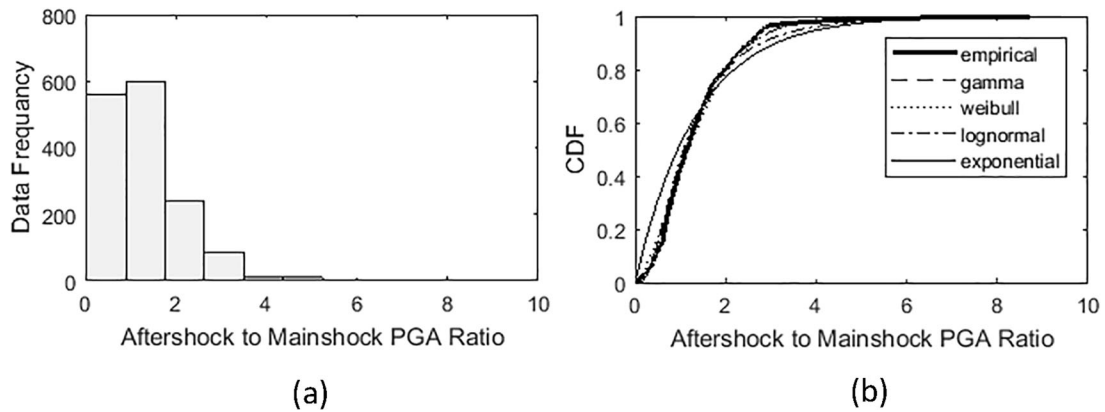
Fig. 12 Aftershock to mainshock PGA ratio of the Iranian accelerogram database, **a** frequency of data, **b** cumulative density function

due to the closer occurrence location of the aftershocks to the recording station with respect to the corresponding mainshock, or the different underlying soil layer properties that the waves propagate along with it. Additionally, among well-known standard probability distribution functions including the Lognormal, Weibull, Exponential, and Gamma functions, by implementing the maximum negative Log likelihood method, the first one can estimate this empirical CDF very well with the mean of 1.47 and standard deviation 1.32.

Next, the vertical to horizontal PGA ratio, which has been addressed by many researchers (Bozorgnia and Campbell 2016; Haji-Soltani et al. 2017; Gülerce et al. 2017) in the previous works, is assessed from the probabilistic perspective for all recorded data of the databank. In this regard, the PGA ratio of vertical direction to horizontal one is calculated for all data, and then the corresponding CDF of this parameter is obtained and

illustrated in Fig. 13. Interestingly, the probability of facing an earthquake with the vertical PGA smaller than the horizontal one is 62%, which means that it is considerably probable (equal to 38%) to face a recorded earthquake with higher vertical PGA in comparison to the horizontal one. Additionally, again among the previously defined standard probability distribution functions, the Lognormal distribution with the mean of 0.99 and the standard deviation of 0.43 is the best estimator.

The relationship between PGA and PGV of the recorded data, especially for the pulse-like events, is an important issue from the engineering perspective which can lead to interesting results. This relationship for different types of records, namely pulse-like, non-pulse-like, and ambiguous, are established and indicated in Fig. 14 by applying the least square curve fitting method as an initial perspective for the engineers/scientists who are interested to use these accelerograms for the further
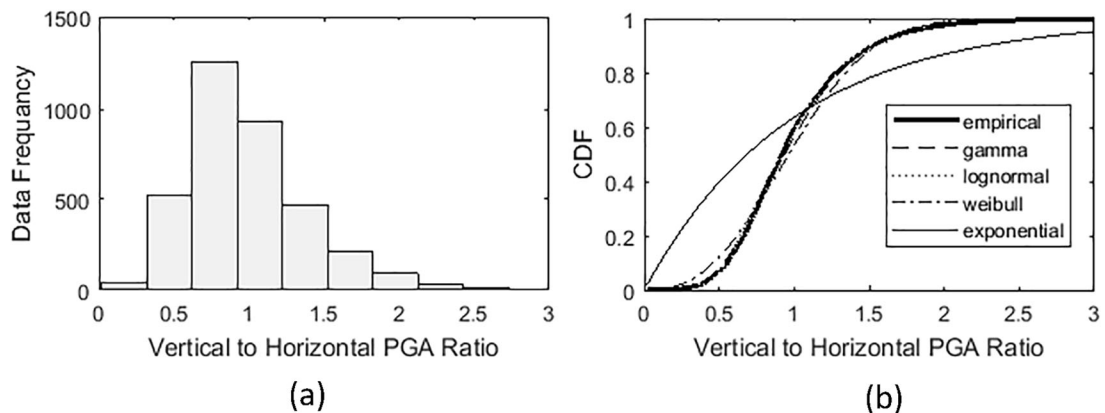
Fig. 13 Vertical to horizontal PGA ratio of the Iranian accelerogram database, **a** frequency of date, **b** cumulative density function
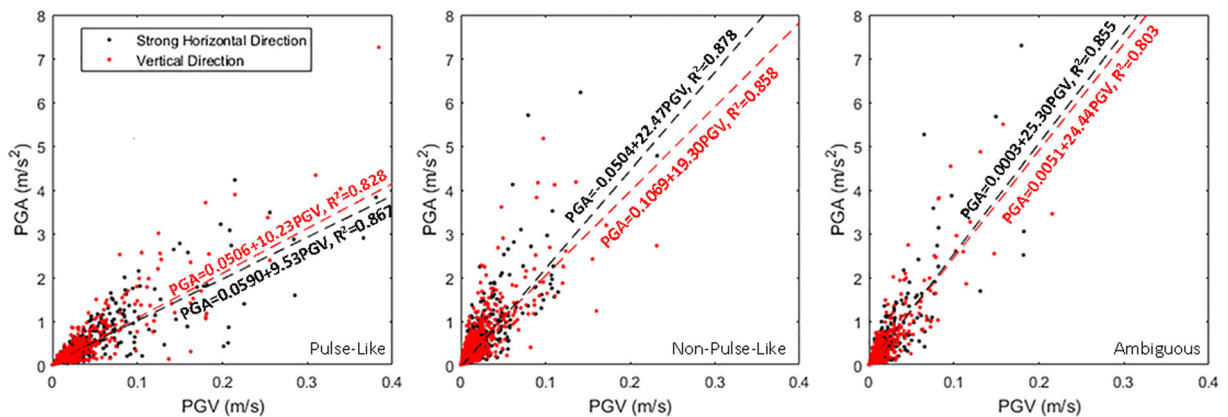
**Fig. 14** PGA-PGV relationship of the Iranian accelerogram database

structural dynamic time history analysis. It is seen that these two ground motion parameters increase with each other linearly; however, for a larger value of parameters, they tend to diverge from the straight line. In addition, the relationship does not show dependency on the direction of signal. In other words, fitted linear equations to the horizontal and vertical components are close to each other. In the case of the pulse-like event, the slope of fitted line is less than the other two record types, which means the higher value of PGV in near-field signals. Interestingly, in contrary to the ambiguous and non-pulse-like records, the slope of horizontal component of pulse-like events is higher than the vertical ones. It implies that the pulses in the horizontal plane are more intensive than those recorded in the vertical direction. Moreover, the PGA-PGV relationships of the ambiguous records are more similar to the non-pulse-like data.

In addition to the ground motion recorded time history peak values, some other ground motion parameters exist that are important to identify the properties of earthquakes such as Arias intensity and ground motion duration. In Fig. 15, the average of Arias intensity of accelerograms in the strong horizontal direction is illustrated for different moment magnitudes, site-to-source distances, and soil types. It is seen that by the increase in the site-to-source distance (more than 60 km), the Arias intensity decreases considerably. Also, signals related to the earthquake events with the moment magnitude between 5 and 6 show about 26% higher Arias intensity than the ones with the magnitudes less than 5. Moreover, on average, the Arias intensity of accelerograms recorded on the softer site (soil type 3) is 28% lesser than recorded events on the hard rock, while these differences are more significant in the case of events with

smaller magnitude. In addition, generally, pulse-like earthquakes Arias intensity is 30% larger than the non-pulse-like signals. Besides, averagely, the mainshocks Arias intensity is 35% higher than the aftershock ones, especially in the case of accelerograms with the higher site-to-source distance, which implies the higher level of energy existing in the mainshock accelerograms.

The next important parameter is the ground motion duration of the accelerograms evaluated based on the Trifunac and Brady method (Trifunac and Brady 1975). This is one of the key parameters in the selection of accelerograms for the nonlinear dynamic analysis of structures (ASCE7-16, 2016, Khansefid and Bakhshi 2017). As it is observed in Fig. 16, the larger the site-to-source distance, the higher the strong motion duration is. Additionally, the average duration of events with larger magnitude (6 or above) is about 84% higher than the ones with the magnitude of 5 or less. Moreover, the average of strong motion duration for soil types 1, 2, and 3 are 7.0, 10.1, and 11.2 s, respectively, which implies that in the softer soil layers, it takes more for the earthquake waves to release a major part of their energy. It is also interesting to mention that the mean value of pulse-like strong motion duration (7.03 s.) is almost 32% shorter than non-pulse-like ones (10.38 s.). This is due to the fact that in the pulse-like events, most of the event energy is discharged during the shorter time by a destructive velocity pulse. Finally, on average, the aftershocks duration (8.16 s.) is 25% less than the mainshocks (10.6 s.), which is due to the lower amount of energy released in the aftershock events.

In most structural design codes, there is no appropriate classification for mainshocks and aftershocks, as well as near-field (pulse-like) and far-field (non-pulse-like) events. Herein, the average acceleration response

M-NP: Mainshock, Non-Pulse-Like   A-NP: Aftershock, Non-Pulse-Like

M-P: Mainshock, Pulse-Like   A-P: Aftershock, Pulse-Like
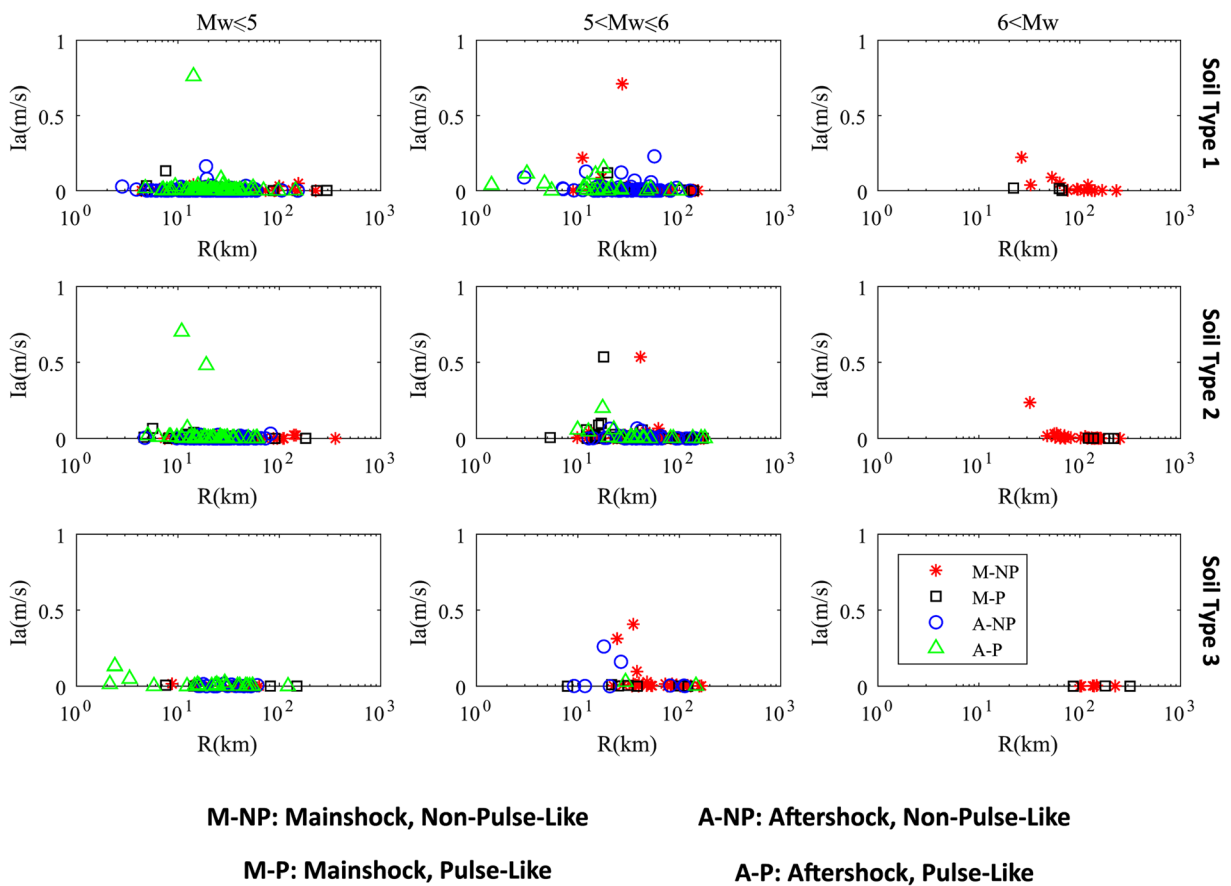
Fig. 15 Arias intensity of the Iranian accelerogram database

spectra of earthquakes are illustrated and discussed for different categories. The site soil condition and the seismicity level of the records are considered. Soil type classification was mentioned before, and in accordance with the Iranian seismic code (BHRC 2014) for the seismicity level there are four different categories based on the peak ground acceleration of the events: low (PGA = 0.20 g), intermediate (PGA = 0.25 g), high (PGA = 0.30 g), and very high (PGA = 0.35 g).

In Fig. 17, the average of 5% damped acceleration response spectra of the Iranian earthquake database are plotted for all three principal directions of motion. For pulse-like recorded accelerogram, the principal direction in the horizontal plane is found based on maximizing the energy of the signal. In the case of non-pulse-like records, the relationship suggested by Rezaeian and Kiureghian (2012) is used to rotate the signals. These spectra, in the principal directions, are separately drawn for different seismicity levels and site soil classes. In addition, the near-field directivity effect and mainshock-

aftershock classification are taken into account. Furthermore, for any of the classes, the corresponding standard seismic acceleration spectrum is plotted to be compared with the results of database accelerograms. Due to the lack of data, in some categories, there is no data to calculate the acceleration response spectrum.

As it is observed, the standard spectra are not suitably matched with the obtained spectra from the accelerogram database. For the high seismicity level, the standard acceleration spectrum is lower than the database ones in many cases. In other seismicity levels, this is somehow reverse, i.e., the standard spectra are above the calculated results in the low to medium period range whereas in the long periods they are underestimated. The next remarkable observation is the importance of considering earthquakes in the vertical direction. It is observed that they are nearly at the same level of the strong horizontal direction response both for pulse-like and non-pulse-like events. This is against the general attitude that believes the vertical direction accelerograms are weaker than the horizontal ones
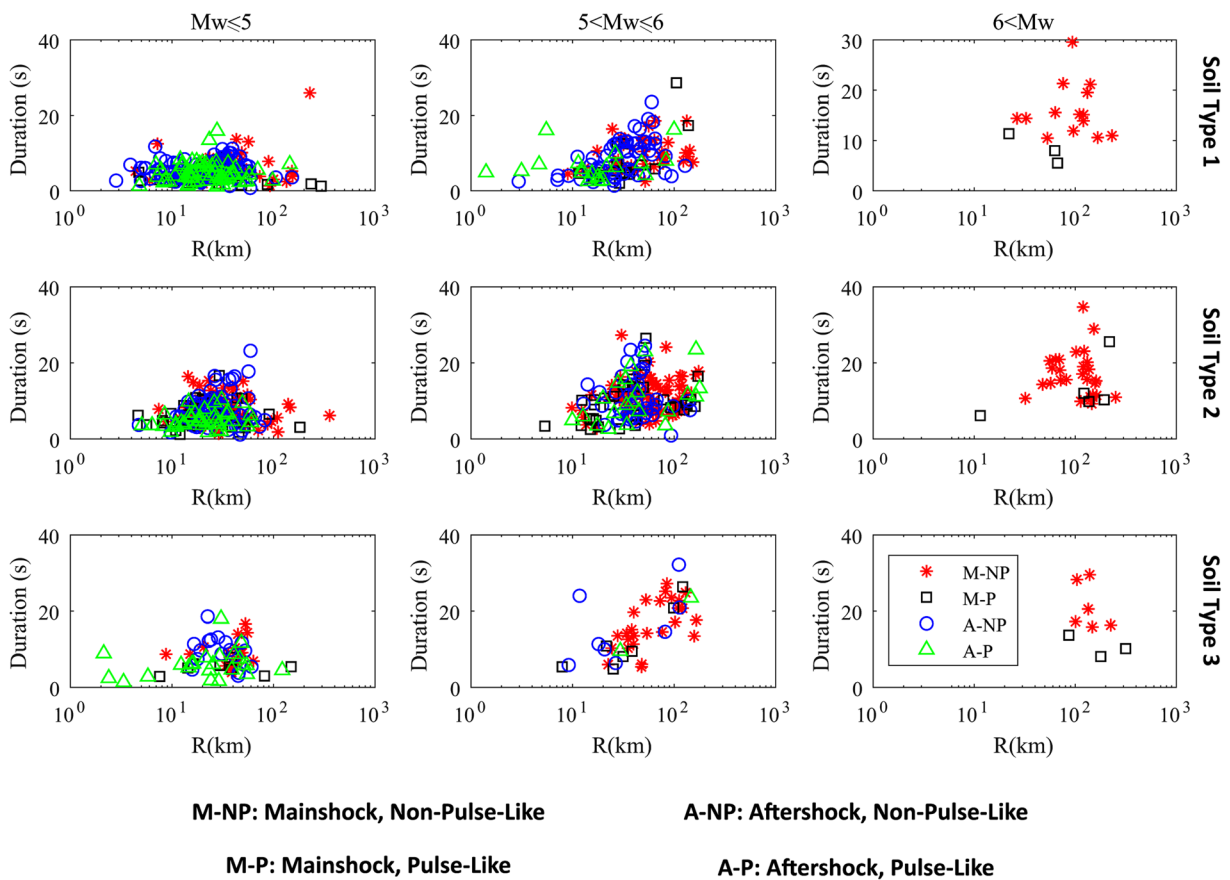
Fig. 16 Ground motion duration of the Iranian accelerogram database

(especially for the far-field event) and it is not necessary to consider them in many cases. As an example, only for the far-field event in some national and international codes (Office of Deputy for Strategic Supervision 2014, ASCE 41-13 2013), it is suggested to multiply the horizontal spectrum by $\frac{2}{3}$ to find the vertical spectrum. However, this guide is totally violated by results from the existing accelerogram database of Iran. As another important result, for some soil types and seismicity levels, pulse-like mainshocks create higher acceleration response, especially in the higher period range. In some cases, the non-pulse-like mainshocks, and also in some others the aftershocks, produce a high level of acceleration spectra. Therefore, it is important to consider the mainshock-aftershock classification as well as pulse-like and non-pulse-like categories to suggest a standard acceleration spectrum.

As a final discussion, we compare some important features of the whole database presented in this research to that of the proposed or used in the other articles for the Iranian plateau. Six different databases provided by

Ghasemi et al. (2009), Ghodrati Amiri et al. (2009), Saffari et al. (2012), Yazdani and Kowsari (2013), Zafarani and Soghrat (2017), and Soghrat and Ziyaeifar (2017) are considered. The important parameters of each database are extracted and shown in Table 1. It is seen that among the different databases, the one presented by Zafarani and Soghrat (2017) is more comprehensive than the other previous research works. In comparison with their database, our suggested one considers the wider time span, bigger magnitude range, and also a higher number of three-components records. In addition to these advantages, here in this research, the mainshock-aftershock classification as well as pulse-like and non-pulse-like identification of the Iranian accelerogram database is presented for the first time in this scale. Therefore, this database can be used for further analysis to generate different types of ground motion prediction equations, structural dynamic time history analysis, and any other type of studies in the field of earthquake engineering with considering proposed categorization.

Fig. 17 Acceleration response spectra of the Iranian accelerograms for near-field and far-field, as well as mainshock and aftershock categories
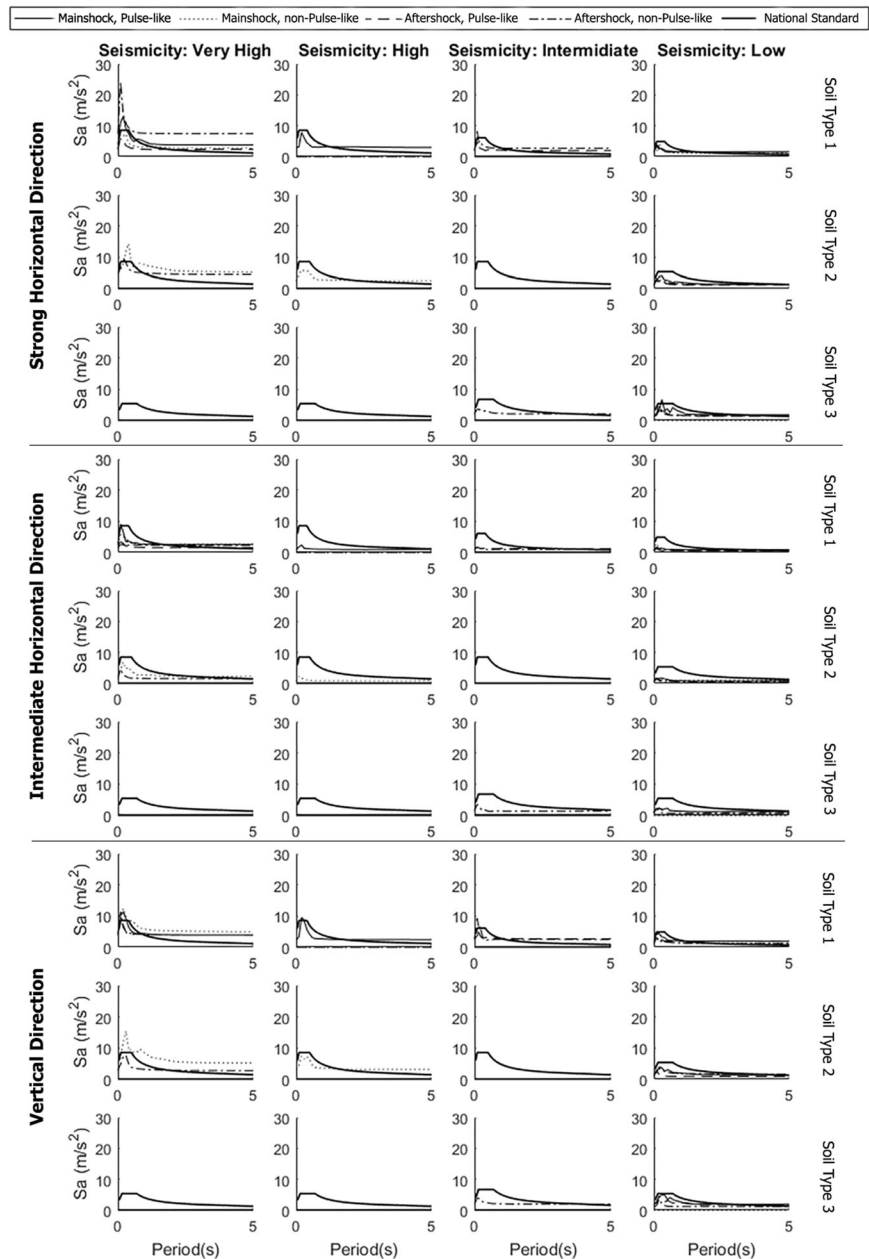
7 Concluding remarks

In this article, it is attempted to propose a comprehensive database of Iranian accelerograms and evaluate their corresponding statistical properties from the engineering point of view. The official database of Iranian accelerogram is collected. All raw and contaminated accelerograms are processed by a new combined correction method including two-step wavelet correction, baseline adjustment, and high-pass filtering. It was revealed that the correction

process does not affect the acceleration time history and response spectra tangibly. However, the velocity time history and response spectra, as well as the displacement ones faced significant changes. Moreover, it is observed that the selection of different mother wavelet functions does not have a considerable effect on the final result of the processed signals.

After processing the accelerograms, all data are declustered to the mainshock-aftershock sequences. Thereupon, the accelerograms are categorized based on the

Table 1 General information of the important properties of the different proposed databases of Iranian accelerograms

Research	Recording time interval	Components	Number of records	Soil type	Source to site distance (Km)	Magnitude
Ghasemi	1977–2007	2-horizontal components	999	Soil and rock	$7 < R < 469$	$5.0 < M_w < 7.4$
Saffari	1975–2008	2-horizontal components	351	Types 1,2, and 3[a]	$5 < R < 135$	$5.0 < M_w < 7.3$
Zafarani	1975–2013	3-components	2286	Types 1,2, 3, 4 and unknown[a]	$0.2 < R < 600$	$3.9 < M_w < 7.3$
Ghodrati	1973–2007	2-horizontal components	725	Soil and rock	$5 < R < 200$	$4.0 < M_S < 7.7$
Yazdani	1978–2008	2-horizontal components	111	Soil and rock	$R < 200$	$5.0 < M_w < 7.4$
Soghrat	1990–2012	3-components	325	Types 1,2, 3, and 4[a]	$R < 300$	$4.1 < M_w < 7.3$
This research	1973–2016	3-components	3585	Types 1,2, 3, 4, and unknown[a]	$0.3 < R < 480$	$3.9 < M_w < 7.7$

[a] The soil classification is based on the Iranian national seismic code (BHRC 2014)

existence of forward directivity effect in the velocity time history of processed signal in any of three dimensions to the pulse-like and non-pulse-like accelerograms. The general results are summarized below:

– A high level of vertical ground motion responses (PGA, PGV, PGD, Arias Intensity, and acceleration response spectra) which is almost equal to the horizontal strong motion direction or even more for all categories is observed with the high probability level. It shows the necessity of considering the vertical direction of earthquakes in the Iranian plateau for the structural analysis directly, not by multiplying the horizontal direction spectrum by $\frac{2}{3}$.

– PGA, PGV, PGD, and Arias intensity of near-field (pulse-like) records are far beyond the far-field (non-pulse-like) events that reveal the importance of separating near-field from far-field events. However, their duration is less, which is due to the forward directivity phenomenon.

– As it is expected, aftershocks recorded accelerograms have slightly less average PGV, PGD, arias Intensity, and duration than the mainshocks. However, their PGA can be larger than the mainshocks with a considerable probability level.

– The existing acceleration response spectrum of the Iranian national standard slightly is not a perfect representative for the Iranian accelerogram database. It suffers from a lack of appropriate classification to encounter the near-field and far-field events as well as mainshocks and aftershocks.

In the end, in this article, a comprehensive database of Iranian processed accelerograms is presented which is categorized for the first time in this scale to the mainshock-aftershock as well as near-field and far-field events. This database provides a good opportunity for future study on earthquakes themselves, as well as the structures under different types of earthquake excitation.

References

Abramowitz M, Stegun IA (1964) Handbook of mathematical functions: with formulas, graphs, and mathematical tables. Dover Publications, New York

Akkar S, Kale Ö, Ansari A, Durgaryan R, Gündoğan AA, Hamzehloo H, Harmandar E, Tsereteli N, Waseem M, Yazjeen T, Yilmaz MT (2014) EMME strong-motion database serving for predictive model selection to EMME ground motion logic tree application, 2nd edn. European Conference on Earthquake Engineering and Seismology, Istanbul, Turkey

Ansari A, Noorzad A, Zare M (2007) Application of wavelet multi-resolution analysis for correction of seismic acceleration records. J Geophys Eng 4(4):362–377

Ansari A, Noorzad A, Zafarani H, Vahidifard H (2010) Correction of highly noisy strong motion records using a modified wavelet de-noising method. Soil Dyn Earthq Eng 30(11): 1168–1181

Baiesi M, Paczuski M (2004) Scale free networks of earthquakes and aftershocks. Phys Rev E 69(6):066106

Baker JW (2007) Quantitative classification of near-fault ground motions using wavelet analysis. Bull Seismol Soc Am 97(5): 1486–1501

Baker JW, Lin T, Shahi SK, Jayaram N (2011) PEER 11/03: new ground motion selection procedures and selected motions for the PEER transportation research program. Pacific Earthquake Engineering Research Center, Berkeley, USA

Boore DM (2001) Effect of baseline corrections on displacements and response spectra for several recordings of the 1999 Chi-Chi, Taiwan, earthquake. Bull Seismol Soc Am 91(5):1199–1211

Boore DM, Bommer JJ (2005) Processing of strong-motion accelerograms: needs, options and consequences. Soil Dyn Earthq Eng 25(2):93–115

Bozorgnia Y, Campbell KW (2016) Ground motion model for the vertical-to-horizontal (V/H) ratios of PGA, PGV, and response spectra. Earthquake Spectra 32(2):951–978

Building, Housing and Urban Development Research Center (BHRC) (1997) Basic accelerograms data of the Iranian Accelerographs Network. BHRC, Tehran

Building, Housing and Urban Development Research Center (BHRC) (2014) Iranian code of practice for seismic resistant design of building. BHRC, Tehran

Building, Housing and Urban Development Research Center (BHRC) (2017) Iran Strong Motion Network, http://smd.bhrc.ac.ir/Portal/en/Search/Stations. accessed 14 December 2017

Donoho DL (1995) Denoising by soft-thresholding. IEEE Trans Inf Theory 41(3):613–627

Donoho DL, Johnstone IM (1994) Ideal spatial adaptation by wavelet shrinkage. Biometrika 81(3):425–455

Donoho DL, Johnstone IM (1995) Adapting to unknown smoothness via wavelet shrinkage. J Am Stat Assoc 90(432):1200–1224

Donoho DL, Johnstone IM, Kerkyacharian G, Picard D (1995) Wavelet shrinkage: asymptopia. J R Stat Soc 57(2):301–369

European Committee for Standardization (2004) Eurocode 8: design of structures for earthquake resistance—part 1: general rules, seismic actions and rules for buildings

Fu Q, Menun C (2004) Seismic-environment-based simulation of near-fault ground motions. 13th World Conference on Earthquake Engineering, Vancouver, Canada

Galiana-Merino JJ, Rosa-Herranz J, Giner J, Molina S, Botella F (2003) De-noising of short-period seismograms by wavelet packet transform. Bull Seismol Soc Am 93(6):2554–2564

Gardner JK, Knopoff L (1974) Is the sequence of earthquakes in southern California, with aftershocks removed, Poissonian? Bull Seismol Soc Am 64(5):1363–1367

Ghasemi H, Zare M, Fukushima Y, Koketsu K (2009) An empirical spectral ground-motion model for Iran. J Seismol 13(4):499–515

Ghodrati Amiri G, Khorasani M, Mirza-Hessabi R, Razavian-Amrei SA (2009) Ground-motion prediction equations of spectral ordinates and arias intensity for Iran. J Earthq Eng 14(1):1–29

Grünthal G (1985) The up-dated earthquake catalogue for the German Democratic Republic and adjacent areas-statistical data characteristics and conclusions for hazard assessment. 3rd International symposium on the analysis of seismicity and seismic risk, Liblice Castle, Czechoslovakia

Gülerce Z, Kamai R, Abrahamson NA, Silva WJ (2017) Ground motion prediction equations for the vertical ground motion component based on the nga-w2 database. Earthquake Spectra 33(2):499–528

Haji-Soltani A, Pezeshk SH, Malekmohammadi M, Zandieh A (2017) A study of vertical-to-horizontal ratio of earthquake components in the gulf coast region. Bull Seismol Soc Am 107(5):2055–2066

Hu S, Gardoni P, Xu L (2018) Stochastic procedure for the simulation of synthetic main shock-aftershock ground motion sequences. Earthq Eng Struct Dyn 47(11):2275–2296

Iwan WD, Moser MA, Peng CY (1985) Some observations on strong-motion earthquake measurement using a digital accelerograph. Bull Seismol Soc Am 75(5):1225–1246

Karimiparidari S, Zare M, Memarian H, Kijko A (2013) Iranian earthquakes, a uniform catalog with moment magnitude. J Seismol 17(3):897–911

Khansefid A, Bakhshi A (2017) Strengthening an existing industrial building by optimally designed passive dampers under seismic and service loads. 4th International conference on smart monitoring, assessment and rehabilitation of civil structures (SMAR 2017), Zurich, Switzerland

Khansefid A, Bakhshi A (2018) Statistical evaluation and probabilistic modeling of aftershock sequences of Iranian plateau. J Seismol 22(5):1249–1261

Khansefid A, Bakhshi A (2019) Advanced two-step integrated optimization of actively controlled nonlinear structure under mainshock–aftershock sequences. J Vib Control 25(4):748–762

Khansefid A, Bakhshi A, Ansari A (2019) Empirical predictive model for generating synthetic non-stationary stochastic accelerogram of the Iranian plateau: including far- and near-field effects as well as mainshock and aftershock categorization. Bull Earthq Eng, https://doi.org/10.1007/s10518-019-00624-1

Mavroeidis GP, Papageorgiou AS (2003) A mathematical representation of near-fault ground motions. Bull Seismol Soc Am 93(3):1099–1131

Molchan GM, Dmitrieva OE (1992) Aftershock identification: methods and new approaches. Geophys J Int 109(3):501–516

Motaghi K, Ghods A (2012) Attenuation of ground-motion spectral amplitudes and its variations across the central Alborz mountains. Bull Seismol Soc Am 102(4):1417–1428

Nowroozi A (2005) Attenuation relations for peak horizontal and vertical accelerations of earthquake ground motion in Iran: a preliminary analysis. J Seismol Earthquake Eng 7(2):109–128

Office of Deputy for Strategic Supervision (2014) No.626: guideline for seismic hazard analysis. Islamic Republic of Iran Vice Presidency for Strategic Planning and Supervision, Tehran

Reasenberg P (1985) Second-order moment of central California seismicity, 1969-1982. J Geophys Res 90(B7):5479–5495

Rezaeian S, Kiureghian AD (2012) Simulation of orthogonal horizontal ground motion components for specified earthquake and site characteristics. Earthq Eng Struct Dyn 41(2):335–353

Ruiz-Garcia J (2012) Mainshock-aftershock ground motion features and their influence in building's seismic response. J Earthq Eng 16(5):719–737

Saffari H, Kuwata Y, Takada S, Mahdavian A (2012) Updated PGA, PGV, and spectral acceleration attenuation relationship for Iran. Earthquake Spectra 28(1):257–276

Soghrat MR, Ziyaeifar M (2017) Ground motion prediction equations for horizontal and vertical components of acceleration in Northern Iran. J Seismol 201(1):99–125

Somerville PG (2003) Magnitude scaling of the near-fault rupture directivity pulse. Phys Earth Planet Inter 137(1–4):201–212

Structural Engineering Institute (2013) ASCE 41-13: seismic rehabilitation of existing buildings. American Society of Civil Engineers (ASCE), Virginia

Structural Engineering Institute (2016) ASCE 7-16: minimum design loads and associated criteria for buildings and other structures. American Society of Civil Engineers (ASCE), Virginia

Trifunac MD, Brady AD (1975) A study on the duration of strong earthquake ground motion. Bull Seismol Soc Am 65(3):581–626

Wang GQ, Boore DM, Igel H, Zhou XY (2004) Comparisons of ground motions from five aftershocks of the 1999 Chi-Chi, Taiwan, earthquake with empirical predictions largely based on data from California. Bull Seismol Soc Am 94(6):2198–2212

Yazdani A, Kowsari M (2013) Earthquake ground-motion prediction equations for northern Iran. Nat Hazards 69(3):1877–1894

Zafarani H, Soghrat MR (2017) A selected dataset of the Iranian strong motion records. Nat Hazards 86(3):1307–1332

Zaliapin I, Gebrielov A, Keilis-Borok V, Wong H (2008) Aftershock identification. Phys Rev Lett 101(1):1–4

Zare M, Hamzehloo H (2005) Strong ground-motion measurements during the 2003 Bam, Iran Earthquake. Earthquake Spectra 21(S1):165–179

Zare M, Sinaiean F (2014) Site effects and classification of Iran accelerographic stations. Geodyn Res Int Bull 1(2):15–23

Zhai C, Chang Z, Li S, Chen Z, Xie L (2013) Quantitative identification of near-fault pulse-like ground motions based on energy. Bull Seismol Soc Am 103(5):2591–2603