Earthquake record processing algorithms to form a strong motion database

V A Mironov¹,², S A Peretokin²,³ and K V Simonov¹

¹ Institute of Computational Modeling SB RAS, 50/44 Akademgorodok, Krasnoyarsk, Russian Federation
² Krasnoyarsk Branch of the Federal Research Center for Information and Computational Technologies, 53 Mira Pr, Krasnoyarsk, Russian Federation
³ Schmidt Institute of Physics of the Earth RAS, 10/1 Bolshaya Gruzinskaya, Moscow, Russian Federation

E-mail: vasya-kun@mail.ru

Abstract. This study is devoted to the development of algorithms and software for earthquake record processing. The algorithms are based on the methodology used by the Pacific Earthquake Engineering Research Center for the implementation of the scientific project NGA-West2. The purpose of processing is to determine reliable values of ground acceleration and other parameters of earthquakes from the available records of velocity time series. To analyze the operation of the algorithms, earthquake records (simultaneously recorded time series of acceleration and velocity) taken from the European Rapid Raw Strong-Motion database were used. The developed algorithms and the implemented software will allow in the future to form a database of strong motions for building regional attenuation models on the territory of the Russian Federation.

1. Introduction
Seismic hazard assessment of the construction site is an integral part of the complex of geotechnical surveys in the design of critical object. An important task when performing probabilistic seismic hazard analysis (PSHA) is the construction of reasonable regional earthquake source zone model and ground motion prediction equation (GMPE). These models are based on regional databases that include seismotectonic and seismological data.

Modern GMPEs predict the physical characteristics of ground motion (vibration amplitudes and response spectra) which depend on the distance to the earthquake source, focal mechanism, magnitude, soil characteristics at the site, and a number of other factors. At the same time, the correctness of the obtained GMPEs and, therefore, the results of PSHA depend on the quality, completeness and reliability of the initial data.

In modern GMPEs, for example, Abrahamson et al. 2014 [1], Campbell and Bozorgnia 2014 [2], Chiou and Youngs 2014 [3], distinguish the following regions: Taiwan, China, Turkey, Japan, Italy. For these regions, the GMPE has its own set of coefficients. There are country-specific GMPE (for example, Switzerland). Often to perform PSHA a logical tree is used where different GMPEs are taken into account.

Thus, in order to obtain correct seismic hazard assessments there is a critical need to build new regional GMPEs or improve old ones. To accomplish this task, a database of strong motions must be
formed that meets a set of requirements. The database should include not only the original time series records, but also processed time series records of accelerations, velocities, displacements and their peak values, acceleration response spectra, filter parameters, earthquake parameters (coordinates, magnitude, distances to the recording station, etc.). This work is devoted to algorithms for earthquake record processing algorithms to form a strong motion database. The main stages of the algorithms are based on the methodology used by the Pacific Earthquake Engineering Research Center for the implementation of the scientific project NGA-West2 [4]. On the territory of the Russian Federation, the overwhelming majority of seismic stations record velocity time series. Taking into account the peculiarities of registering earthquakes on the territory of the Russian Federation, the processing stages were corrected and implemented in MATLAB.

2. Earthquake recording processing scheme
Earthquake recording processing scheme for further use in constructing the GMPE consists of the following stages:

- Converting original earthquake records to editable format.
- Baseline correction.
- Removing the instrument response.
- Selecting a signal window.
- Determining the corner frequencies for a Butterworth band-pass filter.
- Filter Waveform.
- Calculating acceleration and displacement time series.
- Additional correction (trend removal).
- Calculation of 5% damped acceleration response spectrum and RotD50 of accelerations.

First, the original three-component velocity records are read. Baseline correction is performed by subtracting the average of the entire waveform record from each element in the time series.

Removing the instrument response is as follows. The frequency response is plotted from the poles, zeros, gain and sensitivity of the recording instrument. After that, the lowest frequency is selected where a flat response is observed (or the frequency specified by the user taking into account the research tasks). To avoid data loss, we divide this frequency by 1.25 [5]. The resulting frequency is the corner frequency of the high-pass pre-filtering. Further, each element of the time series is divided by the station gain, and then a high-pass filter is applied (5-pole acausal Butterworth filter) [6]. Before filtering, zeros were padded to the left and right of the time series [7], after filtering, these zeros are removed. Thus, the original data is converted into physical quantities in which the low frequency noise has been removed.

A signal window is selected for each record. It should include the noise before the earthquake and the earthquake itself. The length of the noise recording is chosen, if possible, to be at least 1 minute. After, the noise window and the earthquake window are determined. For each window, the Fourier amplitude spectrum and its smoothed function are plotted, calculated following the algorithm of Konno and Ohmachi [8]. The ratios of the smoothed signal-to-noise functions are calculated from the specified frequencies. At high frequencies, the signal-to-noise calculation starts at 80% of the Nyquist frequency. At low frequencies, the calculation of the signal-to-noise ratio begins with the corner frequency determined earlier. Frequencies where the signal-to-noise ratio is greater than 3 are taken as corner frequencies for band-pass filtering. As a result, a band-pass 5-pole acausal Butterworth filter is applied with preliminary addition and subsequent removal of zeros on the right and left.

Then, the computation of the time series of displacements (accelerations) is performed by integrating (differentiating) in the frequency domain of the filtered velocity time series using the fast Fourier transform. Zeros were padded to avoid trends when obtaining displacement time series [7]. If the trends are still present, then a sixth order polynomial is selected for displacement time series. In this case, the coefficients of the zero and first order are equal to 0. The resulting polynomial is twice
differentiated and subtracted from the acceleration time series. From the corrected accelerations, we obtain the velocities and displacements by integrating [7].

Finally, the calculation of the response spectra of the corrected acceleration time series is performed. Also, the calculation of the geometric mean response spectrum from two horizontal components (RotD50) is performed, which does not depend on the azimuth of the earthquake epicenter relative to the recording station [9]. The calculation of RotD50 is due to the fact that modern GMPEs are focused precisely on obtaining the geometric mean response spectra.

If the initial data is acceleration time series, then in the presented algorithm, the velocities and displacements are found by differentiating the accelerations. The rest of the stages in the proposed scheme remain the same.

3. Verification of algorithms using simultaneously recorded time series of acceleration and velocity

To analyze and verify the operation of the algorithms, earthquake records (simultaneously recorded time series of acceleration and velocity) taken from the European Rapid Raw Strong-Motion database were used [10]. Table 1 shows the characteristics of the two selected earthquakes (events) and the parameters of the stations.

Table 1. Characteristics of earthquakes and stations.

| Event Number | 1 | 2 |
|--------------|---|---|
| Event ID     | 20210327_0000078 20200424_0000178 | |
| Event time (UTC) | 27.03.2021 13:47 | 24.04.2020 22:04 |
| Event magnitude | 5.5 (mw) | 4.9 (mb) |
| Event depth (km) | 5 | 22 |
| Event latitude (°) | 42.67 | 45.9 |
| Event longitude (°) | 16.37 | 27.47 |
| Network code | IV | GE |
| Station code | APRC | TIRR |
| Station epicentral distance (km) | 122.21 | 176.54 |
| Station latitude (°) | 42.67 | 45.9 |
| Station longitude (°) | 16.37 | 27.47 |
| Station elevation (m) | 672 | 77 |

Figure 1. Velocity time series of the horizontal component N-S for event 1.
Figure 1 shows the velocity time series of the horizontal component N-S for event 1 after baseline correction and removing the instrument response (initial data is a velocity time series). In the figure, the vertical line shows the separation between the noise window and the earthquake window.

Figure 2 shows the procedure for determining corner frequencies from earthquake and noise windows. It can be seen from the figure that the noise level is significantly lower than the earthquake level. According to the proposed algorithm, the corner frequency in the low frequency domain is 0.08 Hz, the corner frequency in the high frequency domain is 40 Hz. After filtering, time series of acceleration and displacement were obtained.

**Figure 2.** Procedure for determining corner frequencies.

Figure 3 shows the acceleration time series. Figure 4 shows the displacement time series. Figure 5 shows the acceleration response spectra of two horizontal components obtained from the velocity time series of event 1. The figure also shows the geometric mean acceleration response spectrum RotD50.

**Figure 3.** Acceleration time series of the horizontal component N-S for event 1.
A similar processing scheme was performed for all components of the velocity time series of earthquake 1 and earthquake 2. Also, the proposed algorithms were applied to all components of the acceleration time series of earthquake 1 and earthquake 2. Table 2 presents the values of the peak ground acceleration for the initial waveforms of acceleration (PGA-measurement) and for accelerations obtained using velocity time series (PGA-modeling).

**Table 2. Comparison of the PGA of two earthquakes.**

| Event Number | 1       | 2       |
|--------------|---------|---------|
| PGA-measurement Z component (ms²) | 0.0206  | 0.0125  |
| PGA-measurement N-S component (ms²) | 0.0195  | 0.0125  |
| PGA-modeling N-S component (ms²) | 0.0466  | 0.0172  |
| PGA-modeling E-W component (ms²) | 0.0462  | 0.0184  |
| RotD50(PGA-measurement) (ms²)    | 0.0388  | 0.0191  |
| RotD50(PGA-modeling) (ms²)      | 0.0378  | 0.0176  |

Table 3 shows the peak ground velocity values for the initial velocity waveforms (PGV-measurement) and for velocities obtained using acceleration time series algorithms (PGV-modeling). Figure 6 shows the response spectra of three components of event 1: model (obtained from velocity time series) and measured (taken from the European Rapid Raw Strong-Motion database).
Table 3. Comparison of the PGV of two earthquakes.

| Event Number | 1   | 2   |
|--------------|-----|-----|
| PGV-measurement Z component (ms⁻¹) | 0.0020 | 0.0004 |
| PGV-modeling Z component (ms⁻¹) | 0.0019 | 0.0004 |
| PGV-measurement N-S component (ms⁻¹) | 0.0030 | 0.0006 |
| PGV-modeling N-S component (ms⁻¹) | 0.0030 | 0.0006 |
| PGV-measurement E-W component (ms⁻¹) | 0.0029 | 0.0005 |
| PGV-modeling E-W component (ms⁻¹) | 0.0028 | 0.0006 |

Figure 6. Comparison of the response spectra of the three components (event 1).
From tables 2 and 3 it can be seen that the model (obtained during processing) and measured values of PGA and PGV are similar, the maximum deviation is 6% and 4%, respectively. In figure 6 the maximum deviation of the response spectra is observed in the domain of the corner frequency boundaries during filtering. In the engineering domain of spectral periods from 0.1 to 3 s, the maximum deviation does not exceed 3%. The observed differences in the obtained data can be caused by the following reasons. First, not strictly accurate characteristics of the recording equipment or the effect of local noise on recording. Second, using different filtering techniques or corner frequencies. The latter reason well explains the differences in the response spectra in figure 6. In general, the data obtained after processing the earthquake records show good convergence. Therefore, the developed earthquake record processing algorithms can be applied to form a strong motion database.

4. Conclusion

The work is devoted to earthquake record processing algorithms to form a strong motion database. The main algorithms are presented which are based on the methodology used by the Pacific Earthquake Engineering Research Center for the implementation of the scientific project NGA-West2.

The implemented process of earthquake record processing is automated as much as possible. The system of criteria and standards minimizes the influence of the human factor (expert opinion) on the final result. The main stages of processing are the transition from initial data to physical quantities, filtering, integration and differentiation of time series, obtaining acceleration response spectra, obtaining PGA and PGV. In the proposed processing scheme, the initial data can be both time series of velocities and time series of accelerations.

To verify and analyze the operation of the algorithms, earthquake records taken from the European Rapid Raw Strong-Motion database were used. Two recent earthquakes of magnitude 5.5 and 4.9 were considered. The main selection criterion was the presence of simultaneously recorded time series of acceleration and velocity at the registration site. In addition, it is necessary that the description of the characteristics of the equipment (zeros, poles, gain and sensitivity) was free access.

Comparison of the results of the algorithms, namely, obtaining accelerations from velocities and vice versa, showed the following. The maximum deviation is 6% and 4%. The maximum deviation of the response spectra is observed in the domain of the corner frequency boundaries during filtering. In the engineering domain of spectral periods from 0.1 to 3 s, the maximum deviation does not exceed 3%. Thus, developed earthquake record processing algorithms show a high convergence of the results. Therefore, these algorithms can be used in the formation of strong motion database in order to further build regional GMPEs.

References
[1] Abrahamson N A, Silva W J and Kamai R 2014 Summary of the ASK14 ground motion relation for active crustal regions Earthquake Spectra 30 pp 1025–55
[2] Campbell K W and Bozorgnia Y 2014 NGA-West2 ground motion model for the average horizontal components of PGA, PGV, and 5%-damped linear acceleration response spectra Earthquake Spectra 30 pp 1087–115
[3] Chiou B S J and Youngs R R 2014 Update of the Chiou and Youngs NGA model for the average horizontal component of peak ground motion and response spectra Earthquake Spectra 30 pp 1117–53
[4] Bozorgnia Y et al. 2014 NGA-West2 research project Earthquake Spectra 30 pp 973–87
[5] Abrahamson N A and Silva W J 1997 Empirical response spectral attenuation relations for shallow crustal earthquakes Seismological Research Letters 68 pp 94-127
[6] Rennolet S B, Morgan P. Moschetti M P, Thompson E M and Yeck W L 2018 A flatfile of ground motion intensity measurements from induced earthquakes in Oklahoma and Kansas Earthquake Spectra 30 pp 1–20
[7] Boore D M 2005 On pads and filters: processing strong-motion data BULL. SEISMOLOG Soc AMER 95 pp 745–50
[8] Konno K and Ohmachi T 1998 Ground-motion characteristics estimated from spectral ratio between horizontal and vertical components of microtremor *Bull Seismol Soc Amer* **88** pp 228–41

[9] Boore D M, Watson-Lamprey J and Abrahamson N A 2006 Orientation-independent measures of ground motion *Bull Seismol Soc Amer* **96** pp 1502–11

[10] Cauzzi C et al. 2016 Introducing the European rapid raw strong-motion database *Seismological Research Letters* **87** pp 977-86