

17<sup>th</sup> World Conference on Earthquake Engineering, 17WCEE Sendai, Japan - September 13th to 18th 2020

## AUTOMATIC SEARCH ALGORITHM FOR LOW CUT-OFF FREQUENCY FOR FILTERING STRONG-MOTION RECORDS

Haiying Yu<sup>(1)</sup>, Da Zhu<sup>(2)</sup>, Baofeng Zhou<sup>(3)</sup>, Xinsheng Ma<sup>(4)</sup>, Wenxi Ma<sup>(5)</sup>

- <sup>(1)</sup> Professor, Key Laboratory of Earthquake Engineering and Engineering Vibration, Institute of Engineering Mechanics, China Earthquake Administration, haiyingyu@126.com
- <sup>(2)</sup> Graduate student, Key Laboratory of Earthquake Engineering and Engineering Vibration, Institute of Engineering Mechanics, China Earthquake Administration, gonglisuozcd@gmail.com
- <sup>(3)</sup> Graduate student, Key Laboratory of Earthquake Engineering and Engineering Vibration, Institute of Engineering Mechanics, China Earthquake Administration, zbf166@126.com
- (4) Graduate student, Key Laboratory of Earthquake Engineering and Engineering Vibration, Institute of Engineering Mechanics, China Earthquake Administration, xinshenghit@163.com
- <sup>(5)</sup> Graduate student, Key Laboratory of Earthquake Engineering and Engineering Vibration, Institute of Engineering Mechanics, China Earthquake Administration, calvinmwx@163.com

## Abstract

Aiming at the low efficiency of traditional methods to determine the low cut-off frequency in strong-motion records processing, we propose an automatic search model to determine the cut-off frequency of low-frequency filters and use the loss function in a statistical learning method to determine the end condition of the automatic search model flow. Based on strong-motion records obtained from the principal shocks and aftershocks of the 2008 Wenchuan earthquake and 2013 Lushan earthquake in China, we compare the computational result of the automatic search algorithm and the low cut-off frequency identified by the traditional method, analyze the errors of the automatic search algorithm, and propose theoretical principles and applicable conditions of their use. Experimental results indicate that the automatic search algorithm we proposed improves computational efficiency significantly compared to the traditional method. Particularly, our automatic search algorithm is suitable for processing strong-motion records in batches.

Keywords: strong-motion records; filter; low cut-off frequency; automatic search algorithm; loss function



17<sup>th</sup> World Conference on Earthquake Engineering, 17WCEE Sendai, Japan - September 13th to 18th 2020

### 1. Introduction

Strong-motion records are fundamental in earthquake engineering and engineering seismology; they provide high-quality scientific data for seismic fortification, rapid reporting, early warning of seismic intensity and strong-motion observation for major projects. However, due to the influence of environmental vibration interference and so on, strong-motion records inevitably mix with noise. Therefore, it is important to study how to identify noise in strong-motion records. Whether in the past or the future, frequency field denoising is a major denoising technology [1]. Thus, research on filtering technology plays an important role in strong-motion recording and processing.

The selection of filters is one of the core issues in strong-motion recording and processing. Filters vary for different strong-motion recordings and processing mechanisms. The United States Geological Survey (USGS) uses a bilateral Butterworth non-causal filter, the California Strong-motion Instrumentation Program (CSMIP) uses an Ormsby or low-pass Butterworth filter as a non-causal filter, and CSMIP uses causal filters as well. For strong-motion event data, Zhang recommend to use an IIR non-causal filter; for real-time flow data, Zhang recommend to use an IIR causal filter; in terms of the selection of the order of an IIR filter for strong-motion records obtained from Lushan earthquake, Zhang suggests to use the fourth order [2]. Accordingly, in this paper, we use a Butterworth bilateral non-causal filter and the fourth order as our choice.

For digital strong-motion records, the low cut-off frequency and high cut-off frequency both are important parameters in terms of filtering technology. According to the Nyquist's Sampling Theorem, the high-frequency strong-motion acceleration cut-off frequency of the maximum sampling frequency should not be more than 1/2, and the minimum low-frequency cut-off should not be less than twice the length of the last record. CSMIP usually chooses 23 Hz as the cut-off frequency of the high-frequency filter, 25 Hz as the termination frequency, and 0.05-0.07 Hz as the cut-off for low-frequency filtering. Syun'itiro Omote et al. determined the cut-off frequency according to the characteristics of the long period section of the pseudo-velocity spectrum (damping ratio of 5%) [3]. Kenzo Toki studied the strong-motion records obtained by the SMAC-type strong-motion instrument, and the cut-off frequency was determined by the 1/f spectrum [4]. According to Xie Lili and Zhou Baofeng, the high cut-off frequency is very small for strong-motion records to influence displacement of baseline shift, accounting for a seismic design that focused more on the frequency range, with a high cut-off frequency of 35 Hz [5][6].

The low cut-off frequency has a great impact on the displacement baseline offset of strong-motion records. Existing methodology for determining the cut-off frequency, such as empirical and pseudo-velocity spectrum and source spectrum and Fourier amplitude spectrum, cannot filter low-frequency noise effectively and accurately. A good method is to filter after the integral velocity and displacement effect to determine the rationality of the low cut-off frequency, i.e. the traditional method, but its efficiency is still low. Therefore, it is not suitable for China strong-motion network center and massive strong-motion records. In this paper, an automatic search algorithm is proposed to solve the problem of low efficiency in traditional method. However, the result obtained by the traditional method is the standard for the accuracy of the automatic search algorithm.

## 2. Fundamental Data

We use the strong-motion acceleration records obtained from the main shocks and aftershocks of the 2008 Wenchuan earthquake and 2013 Lushan earthquake [7][8]. From the two earthquakes, the national strong-motion network obtained a large number of complete strong-motion records. The strong-motion records collected from the Wenchuan and Lushan earthquake acceleration records are selected as the basis for our study. And the criteria for their selection are as follows: (1) triaxial peak ground acceleration must be at least 20 cm/s/s; (2) strong-motion records with singular waveform features [9][10] are avoided.

According to the criteria, we selected in total of 609 strong-motion acceleration records included both main shocks records and aftershocks records originated from strong-motion data collected from 2008

The 17th World Conference on Earthquake Engineering

17<sup>th</sup> World Conference on Earthquake Engineering, 17WCEE Sendai, Japan - September 13th to 18th 2020



Wenchuan earthquake [7][11] and 2014 Lushan earthquake [8][12]. Fig.1 demonstrates the statistical relationship between magnitude and quantity of these strong-motion records.



Fig. 1 Quantity of strong-motion records grouped by magnitude

## 3. Automatic Search Algorithm for Low Cut-off Frequency

### 3.1 Principle

Traditional method determines the low cut-off frequency by means of observing the displacement and time curve of filtered strong-motion records. Under the circumstance that there is no permanent displacement at the seismic station, the mean value of the tail of the reasonably filtered displacement is close to zero, and the tail-fitting line maintains a horizontal state. Analogous to the traditional method, we propose an automatic search algorithm for low cut-off frequency, as follows:

(1) Set a series of frequency points and store them into a one-dimensional array, i.e. *LPS*, and then sort *LPS* in an ascending order so that we can use the sorted array as the search range of the low cut-off frequency. In addition, the more intensive the frequency points are, the more accurate the computational result is, but it also adds the computational cost.

(2) Take out a frequency point from the LPS in turn and filtered it out. Because the frequency points stored in the LPS have been sorted, the low-frequency information of the strong-motion records can be preserved as much as possible.

(3) Compute the displacement according to the filtered acceleration of the strong-motion records.

(4) Determine whether the mean value and the absolute value of the slope of the tail of the displacement are lower than  $1/n_1$  and  $1/n_2$  of the peak ground displacement respectively. If both conditions are satisfied, then the current low-cut frequency point acts as the candidate; otherwise, jump to step (2). Additionally,  $n_1$  and  $n_2$  are positive integers. In terms of traditional method, the range of  $n_1$  is from 1 to 5 and the range of  $n_2$  is from 50 to 800.

(5) Compare the candidate frequency with the lower limit of the low cut-off frequency determined by Nyquist's Theorem and take larger one among them as the final result and output the low cut-off frequency.

The flowchart of automatic search algorithm model is illustrated in Fig.2.

3.2 Determination of parameters

Based on existing theory and research, we propose the following on the parameters involved in the automatic search model:

(1) 97 frequency points are extracted from the interval between 0.04 and 1 Hz as the search range of the low cut-off frequency.

(2) The bidirectional Butterworth non-causal filter is adopted, using fourth-order filtering.

17<sup>th</sup> World Conference on Earthquake Engineering, 17WCEE Sendai, Japan - September 13th to 18th 2020



(3) The high cut-off frequency of the filter is 35 Hz.

(4) The tail of the displacement is defined as the end of the strong-motion record, which accounts for 1/4 of the total length of the record.



Fig. 2 Flowchart of automatic search algorithm model

The main factor affecting the low cut-off frequency of the traditional method is the subjective judgment of the displacement of the ground-motion records, so the end condition of the automatic search process is the main factor affecting its accuracy. From the algorithm model flowchart as illustrated in Fig.2, the end condition to determine the model flow of the algorithm is to determine the two important parameters  $n_1$  and  $n_2$ . Based on the theory of statistical learning [13], we make the following assumptions:

(1) All of the strong-motion records are stored in X.

(2) The automatic search algorithm model is called a decision function denoted by  $f_n(X)$ , and  $n = (n_1, n_2)^T$ , where T denotes the transposition operation. The decision function  $f_n(X)$  expresses utilization parameters *n* to calculate the low cut-off frequency of each strong-motion record.

(3) According to the traditional method, the low cut-off frequency of the strong-motion records is determined in the  $\mathbf{Y}$ .

(4) X,  $f_n(X)$ , and Y are column vectors.

After making the above assumptions, we must develop the quality criterion of value judgment n, so that the loss function is introduced into the algorithm to measure the quality of the algorithm. The loss function is a non-negative real function, written as  $L(Y, f_n(X))$ . There are many kinds of loss function; in this paper we use

$$L(Y, f_n(X)) = |Y - f_n(X)| \tag{1}$$

The smaller the loss function, the better the parameters used by the decision function.  $n_1$  varies from 1 to 5 and takes a test point at intervals of 0.2 in this paper;  $n_2$  varies from 50 to 800, with test points at intervals of 10. The loss function calculated at all test points is shown in Fig.3. The value of the loss function reaches the minimum of 0.58 when  $n_1$ =4 and  $n_1$ =440. Therefore, the end conditions of the automatic search algorithm are as follows: the tail mean of the displacement is less than 1/4 of the corresponding displacement peak, and the absolute value of the tail linear regression slope is less than 1/440 of the corresponding displacement peak value.



17<sup>th</sup> World Conference on Earthquake Engineering, 17WCEE Sendai, Japan - September 13th to 18th 2020



Fig. 3 Loss function

### 3.3 Calculation example

Based on the low cut-off frequency determined by the traditional method as the reference standard, and based on 609 strong-motion accelerations recorded by the main and aftershocks of the two earthquakes in Wenchuan and Lushan, the low cut-off frequency is calculated by the automatic search algorithm. The results of traditional method and the automatic search algorithm are shown in Fig.4.



#### Fig. 4 Results of traditional method and automatic search algorithm

Taking the record identified by 051YAL130420113402 as an example, the low cut-off frequency is determined according to the automatic search algorithm. The duration of the recording is 80.625 second. According to the Nyquist's Sampling Theorem, the low cut-off frequency of the record is 0.025 Hz. The statistical curves of the acceleration, PGD statistical curve (in the figure, fPGD is PGD after different frequency filtering, and rPGD is an unfiltered PGD), tail mean statistical curve of the displacement, post-shift time history of the filter, and slope of the tail linear regression are shown in Fig.5.

According to the traditional method, the velocity and the displacement time history of the filter are drawn with different low cut-off frequencies, as shown in Fig.6. The figure shows that the influence of the low cut-off frequency and velocity is very small, but it has a great influence on the displacement. Through the observation of the tail of the displacement time history, the time history of displacement is reasonable. For record identified by 051YAL130420113402, the cut-off low frequency should be 0.08 Hz, which is consistent with the 0.08 Hz determined by the frequency-search method.

#### The 17th World Conference on Earthquake Engineering

17<sup>th</sup> World Conference on Earthquake Engineering, 17WCEE Sendai, Japan - September 13th to 18th 2020





Fig.5 The frequency search of the record identified by 051YAL130420113402

From Fig.4, we can see that whether the traditional method or the automatic search algorithm determines the low cut-off frequency, for most strong-motion records it is about 0.1 Hz, and the results obtained by the automatic search algorithm are close to the traditional results. By analyzing Fig.5 and Fig.6, it is feasible to determine the low cut-off frequency and the computational efficiency by an automatic search algorithm. In actual mass strong-motion records processing, the efficiency of the automatic search algorithm is significantly higher than that of the traditional method, and a large amount of manpower is saved. In principle, it can be said that only the end conditions of the algorithm can be adjusted to the strong-motion records of different earthquakes.



17WCE

2020

The 17th World Conference on Earthquake Engineering

17<sup>th</sup> World Conference on Earthquake Engineering, 17WCEE Sendai, Japan - September 13th to 18th 2020



Fig. 6 The filtered velocity and displacement time-history curve of the record identified by 051YAL130420113402

3.4 Error statistics and analysis

The automatic search algorithm is mainly developed on the basis of Zhang's research [2] on the low cut-off frequency, and the resultant Zhang's algorithm. The main difference between the automatic search algorithm and the Zhang's algorithm is the definition of the data tail and the end conditions of the algorithm. According to Zhang's research [2], the tail is defined as the length of 10 seconds at the end of the ground-motion record; The end condition of the algorithm is:

(1) the mean of tail 10s (the tail segment with duration of 10 seconds at the end of the displacement and time curve) is lower than 0.1 cm;

(2) the mean of tail 10s is lower than 1/10 of the mean of tail 10s after filtering with 0.04 Hz;

(3) the mean of tail 10s is lower than 1/10 of the peak of the corresponding displacement and time curve.

To evaluate the effectiveness of the automatic search algorithm, we used the automatic search algorithm and Zhang's algorithm to process 609 strong-motion records and determine the low cut-off frequency of each record. To compare the accuracy with the traditional method, the automatic search algorithm and Zhang's algorithm were used to make the difference between the low cut-off frequency obtained by the traditional method, and the error statistical results are obtained. The error statistics of the automatic search algorithm and the Zhang's algorithm are shown in Fig.7, and the error statistics table is shown in Table 1. It can be seen that, compared with the traditional methods, the low cut-off frequency error of the automatic search algorithm is basically within 0.03 Hz.



(a) Statistics of low-frequency cut-off frequency errors in soil records

17WCE

2020

The 17th World Conference on Earthquake Engineering

17<sup>th</sup> World Conference on Earthquake Engineering, 17WCEE Sendai, Japan - September 13th to 18th 2020







(c) Error statistics of low-frequency cut-off frequency for all records

Fig. 7 Automatic determination of low-frequency cut-off frequency error statistics

Upper limit of error (Hz)	Percentage of automatic search algorithm	Percentage of Zhang's algorithm
0.01	53.20%	26.60%
0.02	73.56%	35.14%
0.03	83.42%	42.69%
0.04	94.25%	50.57%
0.05	96.39%	53.69%
0.06	97.70%	55.01%
0.07	98.52%	57.96%

Fahla 1	Automotio	datamaination	of low from	non out out off fro	anon ary amon statistics
гаріет	Amomane	determination	or row-rea	пепсу сш-он те	allency error statistics
1 4010 1	1 Ideomatic	acterimitation	01 10 11 1100	active out out the	quelle y ell'or bluiblieb

3.5 Applicable conditions of automatic search algorithm

From Fig.7, we can see that the error values of two strong-motion records for low cut-off frequency are above 0.2 Hz. The two records are identified by 051AXT080512142801 and 051CDZ130420080203 respectively, and their acceleration time histories are shown in Fig.8.



Fig. 8 Acceleration time-history curves of the records identified by 051AXT080512142801 and 051CDZ130420080203

For the strong-motion record identified by 051AXT080512142801, according to Yu Haiying's study, which indicates that the effect of the permanent displacement occurred recorded strong motion station by the Wenchuan earthquake, 8 earthquakes in Wenchuan in 21 groups of near field strong motion stations are a permanent displacement in different degrees, due to the limited space, not a list of more than 21 group of

## 8

17<sup>th</sup> World Conference on Earthquake Engineering, 17WCEE Sendai, Japan - September 13th to 18th 2020



three component near field strong-motion records of integral calculation and filtering results [11]. As Fig.9 shows on the left, using the automatic search algorithm to determine the cut-off frequency of low frequency of the strong-motion records band-pass filter (low cut-off frequency 0.29 Hz), two times the integral of the displacement time history would be lost permanent-displacement information. The strong-motion recorded in the near field of a large earthquake can filter out the low-frequency components of the real ground motion [14]. Therefore, not only the strong-motion records are not suitable for filtering, but any strong-motion records with permanent displacement are not suitable for filtering the zero-line correction. That is, it is not suitable for an automatic search algorithm to determine the low cut-off frequency.



Fig. 9 Displacement time-history curves of the records identified by 051AXT080512142801 and 051CDZ130420080203

For strong-motion record identified by 051CDZ130420080203, the data of this record are incomplete from the right diagram of Fig.8. The maximum value of 10 s before recording is 0.055 cm/s<sup>2</sup>, and the maximum value of 10 s is 5.190 cm/s<sup>2</sup>. The difference between ground motion before and after an earthquake is quite different, so it is estimated that this record is short of the last segment. From Fig.9, we can see that by using the right graph, we can use the low cut-off frequency determined by the automatic search algorithm to pass the band-pass filter (low cut-off frequency 0.29 Hz) for the strong-motion records, and the mean value of the last part of the displacement time history obtained by the two integrals is larger. Therefore, the incomplete strong-vibration records are not suitable for the use of an automatic search algorithm to determine the low cut-off frequency.

In conclusion, in addition to avoiding strong-motion records with singular waveform characteristics, the automatic search algorithm can accurately determine the low cut-off frequency when strong-motion records are required to meet the following conditions:

- (1) There is no permanent displacement at the station where the strong-vibration records are recorded;
- (2) The strong-motion record is complete without the absence of the end segment.

### 4. Conclusion

In this paper, by analyzing the principles of selecting low cut-off frequency and taking the strong-motion records collected from the 2008 Wenchuan earthquake and 2013 Lushan earthquake in China as an example, we propose an automatic search algorithm model to determine the low cut-off frequency for filtering strong-motion records and the end condition of the model based on statistical learning methodology. The low cut-off frequency determined by the automatic search algorithm is very close to that computed by the traditional method. Currently, in terms of the methodologies to determine the low cut-off frequency of strong-motion records, there is no standard and only the traditional method as reference, but the algorithm that we proposed in this paper provides a new quantitative way to automatically determine the low cut-off frequency for filtering strong-motion records. And we draw some conclusions as follows:

(1) The accuracy of the low cut-off frequency determined by the automatic search algorithm is close to that of the traditional method, and the computational efficiency is greatly improved compare to the traditional method. And it is extremely suitable for processing a mass amount of strong-motion records in batches.

17<sup>th</sup> World Conference on Earthquake Engineering, 17WCEE Sendai, Japan - September 13th to 18th 2020



(2) The algorithm is suitable for integral strong-motion records that do not include permanent displacement and the records are collected from strong-motion seismic stations.

(3) In principle, the algorithm can be applied to strong-motion records of any different earthquakes by adjusting its end condition.

## 5. Acknowledgements

This research is supported by the National Key Research and Development Plan (No. 2018YFC1504002-03), the National Natural Science Foundation of High-speed Railway Joint Fund (No. U1534202) and the Natural Science Foundation of Heilongjiang Province of P.R. China (No. E2015070).

## 6. References

- [1] Zhang JH, Lü N, Tian LY, Lu WZ, Zhong L (2006): An overview of the methods and techniques for seismic data noise attenuation. Progress in Geophysics, 21(2), 546-553.
- [2] Zhang TY (2016): Study on data processing technology based on strong-motion records of Lushan earthquake. Institute of Engineering Mechanics, China Earthquake Administration, Harbin, China.
- [3] Syun'itiro O, Tokiharu O, Shigeto H, Tomonori I (1988): Data processing method for acceleration records and its application results. Proceedings of 2<sup>nd</sup> Workshop on Processing of Seismic Strong-motion Records, Tokyo, Japan.
- [4] Kenzo T, Sumio S (1988): A method for correction of accelerograms. Proceedings of 2<sup>nd</sup> Workshop on Processing of Seismic Strong-motion Records, Tokyo, Japan.
- [5] Xie LL, Li SB, Qian QK (1983): Some features of current procedure for strong-motion data processing and analysis in China. Earthquake Engineering and Engineering Vibration, 3(1), 1-14.
- [6] Zhou BF (2012): Some key issues on the strong motion observation. Institute of Engineering Mechanics, China Earthquake Administration, Harbin, China.
- [7] Department of Earthquake Disaster Prevention, China Earthquake Administration (2008): Uncorrected acceleration records from the *Ms* 8.0 Wenchuan earthquake. Report on Strong Motion Records in China, set 12, vol. 1. Seismological Press, Beijing, China.
- [8] China Strong Motion Network Centre (2014): Uncorrected acceleration records from the *Ms* 7.0 Lushan earthquake and aftershocks. Report on Strong Motion Records in China, set 17, vol. 1. Seismological Press, Beijing, China.
- [9] Zhou BF, Song TS, Yu HY, Xie LL (2014): The study on the wizard wave in the strong-motion record. Earthquake Engineering and Engineering Dynamics, 34(S1), 93-99.
- [10] Zhou BF, Wen RZ, Xie LL (2014): The preliminary study on the "spike" in strong-motion records. China Civil Engineering Journal, 47(S2), 295-299.
- [11] Yu HY, Jiang WX, Xie QC, Yang YQ, Cheng X, Yang J (2009): Baseline correction of digital strong-motion records in near-field. Journal of Earthquake Engineering and Engineering Vibration, 29(6), 1-12.
- [12] Yu HY, Zhou BF, Yang C, Jiang WX, Xie QC, Zhao XF, Xu SM (2014): Preliminary analysis on strong-motion records from the *Ms* 7.0 Lushan earthquake and aftershocks. Earthquake Engineering and Engineering Dynamics, 34(S1), 153-160.
- [13] Li H (2019): Statistical Learning Method. Tsinghua University Press, 2<sup>nd</sup> edition.
- [14] Yu HY, Wang D, Yang YQ, Xie QC, Jiang WX, Zhou BF (2009): The preliminary analysis of strong ground motion records from the *Ms* 8.0 Wenchuan earthquake. Journal of Earthquake Engineering and Engineering Vibration, 29(1): 1-13.