Bulletin of the Seismological Society of America, Vol. 95, No. 2, pp. 745–750, April 2005, doi: 10.1785/0120040160

On Pads and Filters: Processing Strong-Motion Data

by David M. Boore

Abstract Processing of strong-motion data in many cases can be as straightforward as filtering the acceleration time series and integrating to obtain velocity and displacement. To avoid the introduction of spurious low-frequency noise in quantities derived from the filtered accelerations, however, care must be taken to append zero pads of adequate length to the beginning and end of the segment of recorded data. These padded sections of the filtered acceleration need to be retained when deriving velocities, displacements, Fourier spectra, and response spectra. In addition, these padded and filtered sections should also be included in the time series used in the dynamic analysis of structures and soils to ensure compatibility with the filtered accelerations.

Introduction

As illustrated in Figure 1, low-cut filters are an effective way of removing the low-frequency noise that is present in many, if not most, analog and digital strong-motion recordings (Trifunac, 1971; Boore *et al.*, 2002). This low-frequency noise usually appears as drifts in the displacements derived from double integration of acceleration, making it difficult to determine the true peak displacement of the ground motion. But whether the filtering is done in the time or in the frequency domain, assumptions are needed about the time series outside of the time range of recorded values. Usually the time series is padded with zeros before filtering (in frequency-domain filtering using the fast Fourier transform [FFT], added zeros are usually needed so that the number of points in the time series is a power of two). If the duration of the pad is not long enough or if some scheme other than adding zeros is used to extend the record, there will be distortion in the filtered record due to what Press *et al.* (1992) refer to as "wrap-around pollution." On the other hand, even if the zero pads are long enough to accommodate the filter transients, removal of those padded portions of the filtered data before dissemination of the data can lead users to think that the disseminated accelerations, velocities, and displacements are incompatible, in that the latter two might not be the same as those derived by simple integration of the processed acceleration, assuming zero initial conditions. (The processed data distributed by both the National Strong-Motion Program [NSMP] of the U.S. Geological Survey (USGS) and the California Strong-Motion Instrumentation Program [CSMIP] of the California Geological Survey [CGS] have been stripped of the padded portions of the records.)

The response spectra computed from the pad-stripped processed data can also be incompatible with the response spectra from the processed data computed using the padded and filtered time series (Malhotra, 2001). In these cases the Fourier amplitude spectrum may not tend toward zero at low frequencies, with the consequence that the velocity and displacement may show low-frequency drifts that should not be present after low-cut filtering (Shyam-Sunder and Connor, 1982). This has required some data-processing procedures to be overly complicated, with low-cut filtering of the velocity and the displacement. Converse and Brady (1992) introduced a simplified procedure that used padding of the acceleration time series and integration of the padded and filtered acceleration time series (similar to a procedure advocated earlier by Shyam-Sunder and Connor, 1982). This is a simple and robust procedure for processing both analog and digital accelerograms, but the more complicated procedures are still commonly used. The purpose of this article is to draw attention to the need for pads and to emphasize the simplicity of strong-motion data processing if adequate zero pads are included.

Pad Requirements

Allowance must be made for the duration of the impulse response of the filter. I have heard the argument that because no one knows what happened on either side of the recorded segment of data, processing procedures should make no assumptions about the time series outside of the recorded data segment. This argument overlooks the fact that all filtering procedures make assumptions about the time series outside of the recorded segment. Time-domain filtering programs assume that the time series is zero on either side of the segment of data being filtered, and frequency-domain filtering using a discrete Fourier transform, as in the FFT, assumes that the time series is periodic, with period equal to the length of the data segment extended with zeros or truncated to a number of points equal to a power of two. Often the assumptions about the data outside the recorded segment are made implicitly within the computer programs and are

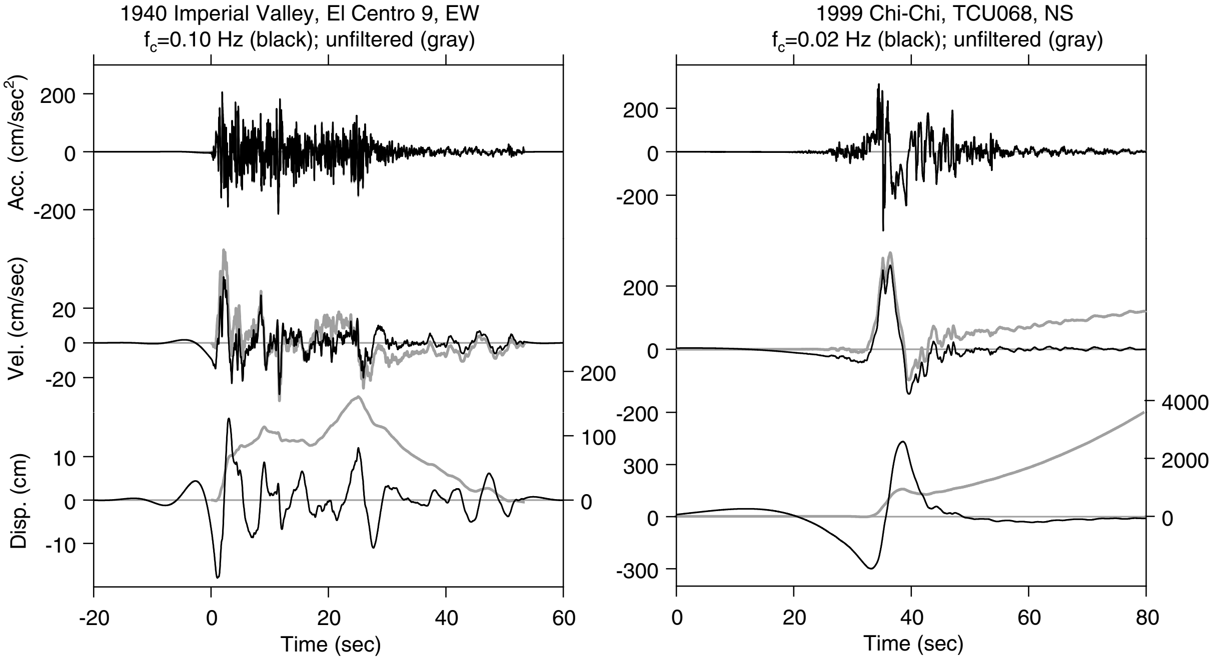

Figure 1. Displacements derived from analog (1940 El Centro) and digital recordings (1999 Chi-Chi). Shown are displacements and velocities from unfiltered and filtered accelerations (using two passes of a fourth-order Butterworth filter with corner frequencies as indicated). The displacement axis labels for the unfiltered motions are given on the right side of the graphs. The return of the unfiltered displacement for the 1940 recording to values near zero at the end of the motion is coincidental and unusual; in most cases, the displacements from unfiltered accelerations wander far from zero at the end of the record (as it does for the 1999 recording).

not visible to the practitioner. This can have differing consequences for time- and frequency-domain filtering. For time-domain processing, low-frequency noise may be reintroduced into the quantities derived from the filtered acceleration time series (such as velocities, displacements, and response spectra) if the portions of the filtered acceleration time series that were padded with zeros are ignored when deriving the other quantities (this produces apparent incompatibility in the various measures of ground motion). For frequency-domain processing the filter response at one end of the data segment will be influenced by data at the other end (because of circular convolution). This wrap-around pollution can be particularly severe if there is a discontinuity in the motion when the end wraps around to the beginning. It is much better practice to extend explicitly the data segment before filtering. Trifunac (1971) argued that a natural assumption is that the time series is zero on either side of the data segment. Hudson (1979) extended the time series by reflecting the data segment about itself on either end. Although it eliminates the discontinuity problem, it is not a physically realistic assumption and leads to distortion near the ends of the data segment (this procedure is still in use in routine processing; e.g., Todorosvka and Lee [2004]). Extending the record with a sufficient length of zeros and retaining these zero-padded sections in all analysis avoids compatibility problems. The procedure also reduces distortion problems, to the extent that the actual ground motion was near zero on either side of the data segment.

The length of the filter transient depends on the order of the filter and on the filter corner. Some examples of filter responses in the frequency and time domain are shown in Figure 2. (All examples in this article use two passes of a Butterworth filter in the time domain, one in the forward and one in the reverse direction, to produce a filter with zero-phase shift; this acausal filter is preferable to a causal filter, as discussed in Boore and Akkar [2003].) Converse and Brady (1992) recommend that the pad length be determined from the following formula:

$$T_{\mathrm{zpad}} = 1.5n/f_c , \tag{1}$$

where T_{zpad} is the total length of zeros to be added to the record, n (an integer) is the order of the Butterworth filter, and f_c is the filter corner frequency. The duration of zeros given by this equation are shown in Figure 3 as a function of filter order for three values of filter corner frequency. (If the record is a digital recording with noise-free pre-event samples, this portion of the record can count as part of the required pad length.) Usually half the zeros are added before the data and half after the data. To avoid truncation effects, tapers near the beginning and end of the data can be applied, but according to Converse and Brady (1992), it is preferable

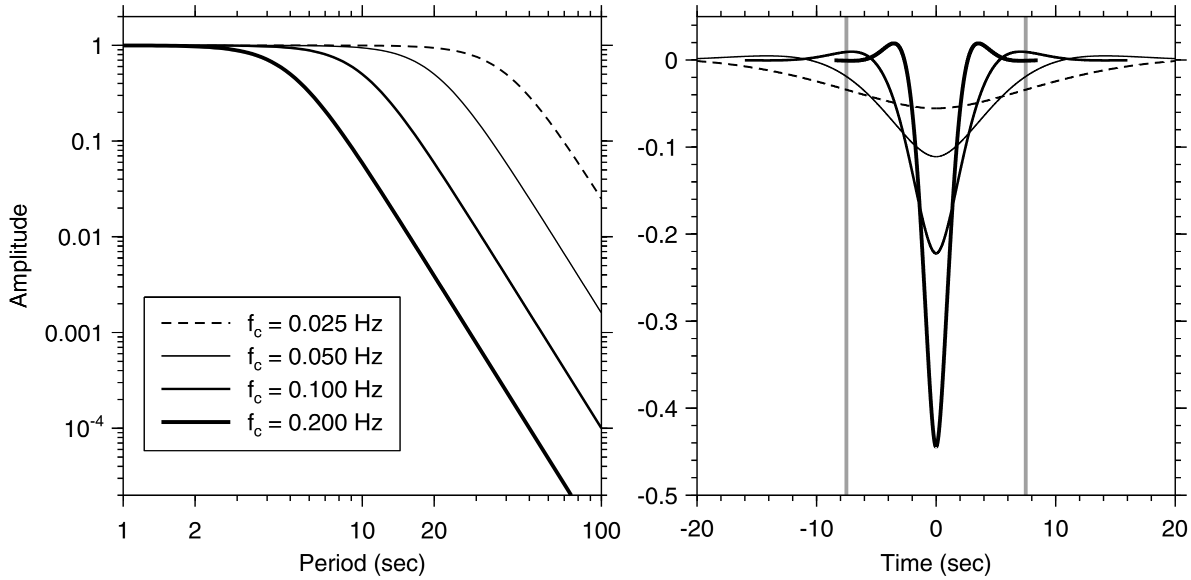

Figure 2. The Fourier amplitude and the time-domain response of a filtered impulse centered at $t = 0$ sec with amplitude of 200, sampled at 200 samples per second. The filtering was done in the time domain with two passes of a second-order Butterworth filter. The filter frequencies (f_c) were 0.025, 0.05, 0.10, and 0.2 Hz. The time-domain impulse responses are made up of the original impulse minus the response of a high-cut filter; only the latter is shown in the figure, plotted at a greatly expanded amplitude scale to show details. The gray vertical lines at ± 7.5 sec on the right-hand graph encompass the value of T_{zpad} given by equation (1) for $f_c = 0.20$ Hz and $n = 2$. The legend for corner frequencies (f_c) applies to both graphs. The time-domain impulse responses for higher-order filters would have more oscillations.

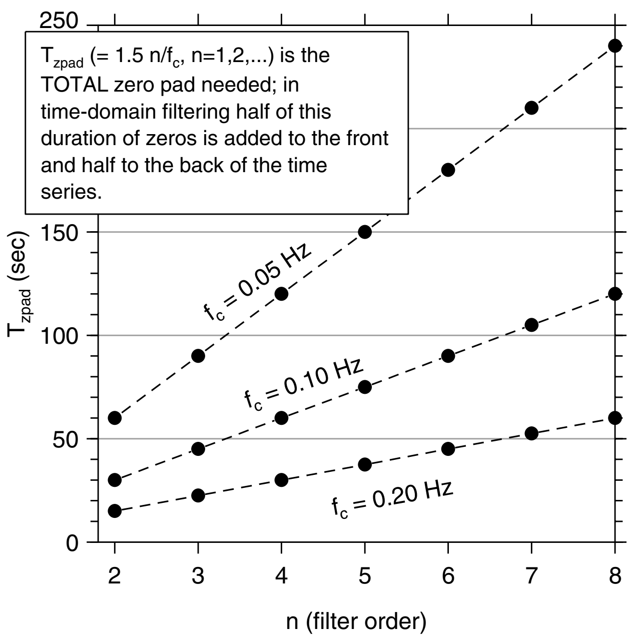

Figure 3. Examples of the total length of the time-domain zero pad recommended by Converse and Brady (1992) to allow for the filter response in two-pass (acausal), nth-order Butterworth filters (these pads are needed regardless of whether the filtering is done in the time or frequency domain).

to search inward from either end of the data segment for the first zero crossing, replacing the data from the beginning and end to the first zero crossings with zeros (to preserve the original length of the data segment); the zero pads are then added to either side of the data segment. This procedure needs modification, however, if a significant portion of the acceleration time series near the beginning or end of the data segment is above or below zero. In such cases a prefiltering baseline correction may be required.

Compatibility Issues

The processed strong-motion data (time series and spectra) provided by the NSMP of the USGS include the effects of filter transients. The results distributed to the public, however, have had the padded portions of the processed time series removed (this is also true of the CSMIP of the CGS). There are a number of reasons that the padded portions have been stripped off of the disseminated records, including file size (the padded portions can be long and have small amplitudes for the acceleration time series), aesthetics (the pre- and postrecorded portions transients can be quite large on the displacement time series and look nothing like true ground motion), and to avoid the incorrect interpretation that the transient motions before and after the recorded data represent actual ground motion. Computer space is no longer an issue, however, and even though they clearly do not rep-

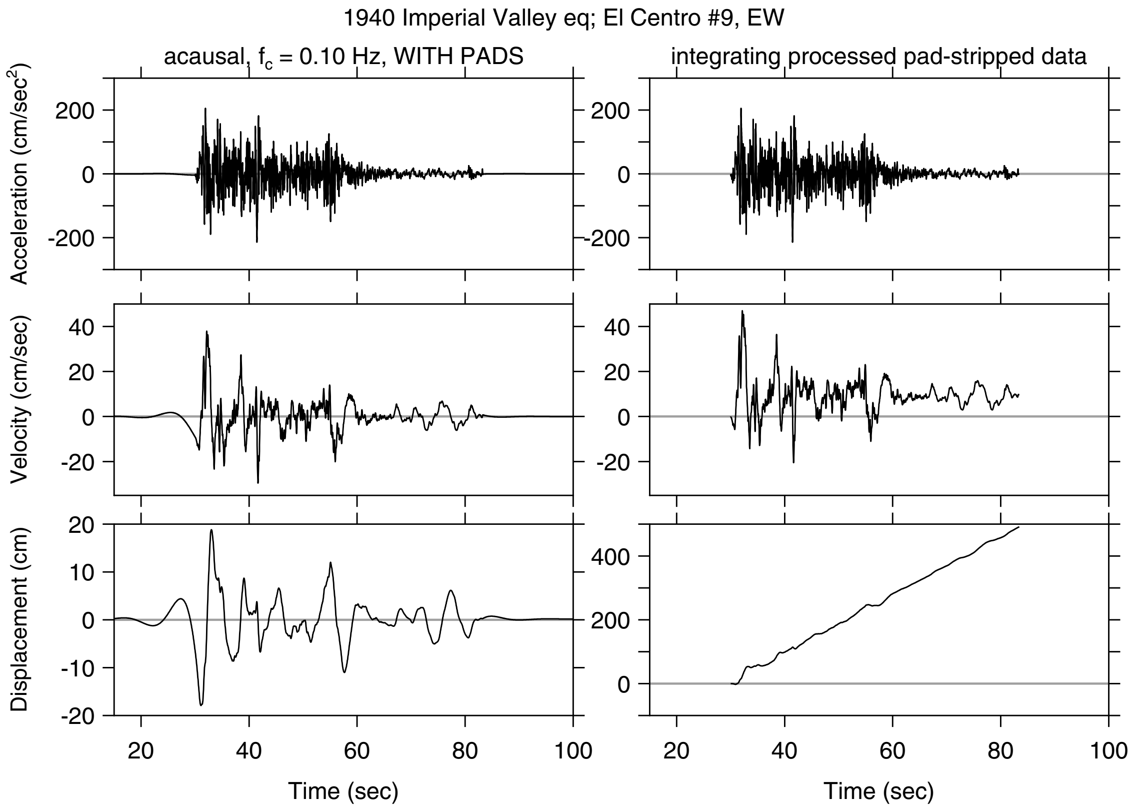

Figure 4. Accelerations, velocities, and displacements derived from the east–west component accelerations recorded at El Centro station 9 during the 1940 Imperial Valley earthquake, illustrating the incompatibility of the processed data that do not include the padded portions of the processed data. This is a particularly egregious example, but many records share the general features shown here. The unprocessed data are from Seekins *et al.* (1992).

resent the ground motion before and after the recorded segment of data, the transients are a necessary mathematical consequence of the filtering operation. Most importantly, removing the padded portions results in an apparent incompatibility between the distributed velocities, displacements, and spectra with those that would be computed from the distributed acceleration time series (e.g., figure 14 in Boore and Akkar, 2003).

An example of this incompatibility is shown in Figure 4. The left panel shows the results of filtering a padded recording (using equation 1 to determine the length of pads) and integration of the filtered acceleration time series without removing the padded portions. The filtered acceleration contains pre- and post-data-segment filter transients, but they are small in amplitude. They become obvious, however, on integration, particularly for the pre-recorded-motion filter transient (the post-recorded-motion transient is less obvious because the motion had subsided to small values by the end of the recorded data segment). If the low-amplitude filter transient in the acceleration time series had been ignored in deriving the velocity and displacements (by stripping off the pre-data-segment portion of the filtered acceleration), low-frequency noise would have apparently been reintroduced into the results (as shown in the right panel of Fig. 4). Although appropriate choices of initial velocity and displacement could have removed the low-frequency trends, this would not have been easy to do for computations of Fourier spectra or response spectra (for the latter, different initial conditions would be needed for each oscillator period). As a result, both types of spectra will show increases of low-frequency amplitudes that should not be present in low-cut filtered data. This is shown in Figure 5.

Comparison with Alternative Processing Schemes

Trifunac (1971) was the first to advocate low-cut filtering as a way of removing low-frequency drifts in the velocities and displacement time series obtained from integration of recorded acceleration time series. Before his article, the drifts were removed using procedures that involved fitting polynomials to the data. The processing method proposed by Trifunac (1971) was quite complicated, however, as shown by the flowchart in his figure 1. The method contained in Converse and Brady (1992) is much simpler. It includes adding zero pads to the acceleration time series, filtering, and integration of the filtered acceleration time series. No

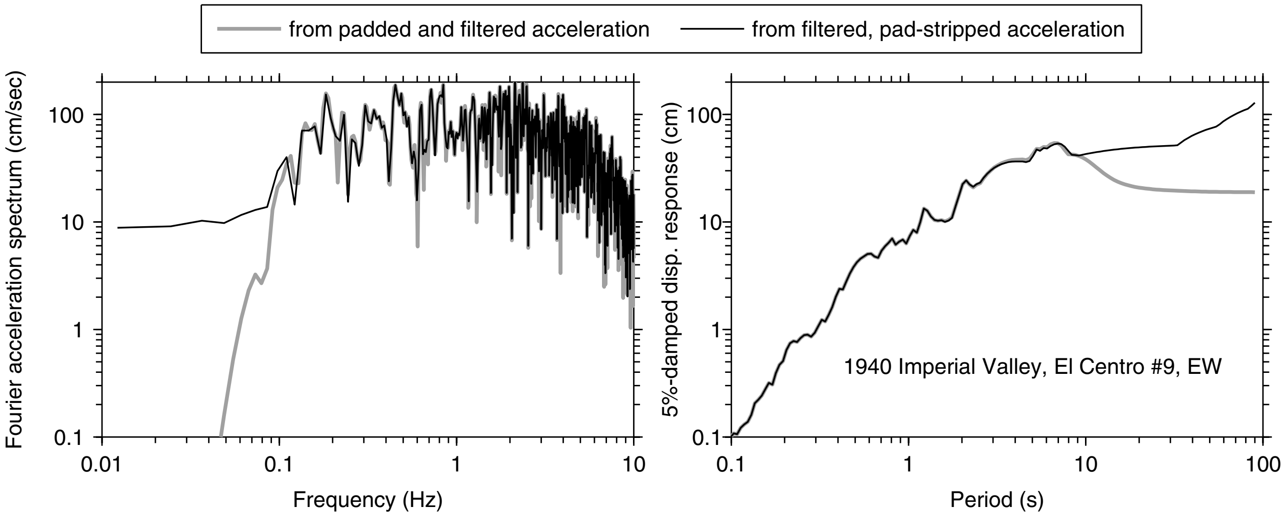

Figure 5. Fourier and response spectra computed for the El Centro station 9 recording of the 1940 Imperial Valley earthquake, from the filtered acceleration before and after removal of the zero-padded portion.

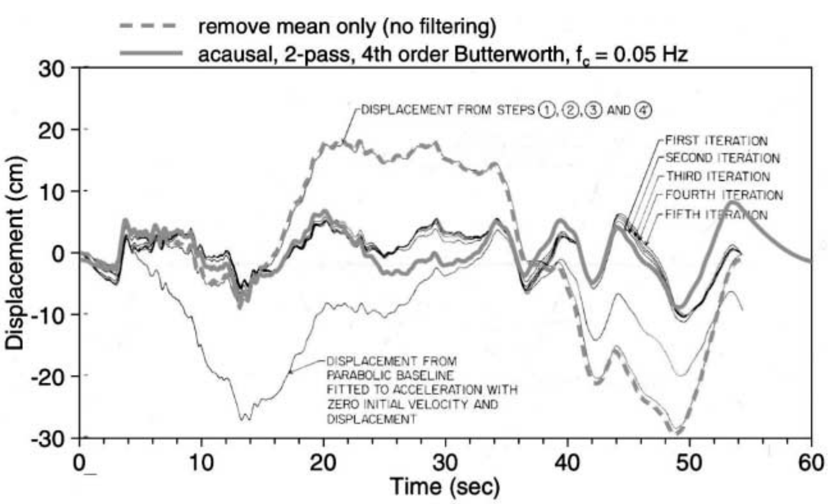

Figure 6. Comparison of the Converse and Brady processing (gray) with that of Trifunac (1971), for the N21°E component of the motion at Taft from the 1952 Kern County earthquake. The unprocessed data came from Seekins *et al.* (1992). The base figure is a scanned image of figure 6 from Trifunac (1971).

baseline corrections or determinations of initial conditions are required. The key to the simplicity of the Converse and Brady (1992) method is the use of zero pads of adequate length in the processing. The method yields results similar to those from the more complicated method, as shown in Figure 6.

Chiu (1997) introduced a method for processing digital recordings that was simpler than the commonly used "Volume II" processing method requiring filtering, integration, and then filtering again (Lee and Trifunac, 1984). Chiu's method is based on low-cut filtering, but unfortunately it is somewhat more complicated than need be because it requires removing a linear trend inexplicably appearing in the displacement times series obtained from integration of the low-cut filtered acceleration time series. Such a trend should not exist if the low-cut filtering was done properly. Figure 7 shows an application of the Converse and Brady (1992) processing procedure to the data used in Chiu's figure 5. The low-frequency distortion from the unprocessed data (top graph, Fig. 7) is removed without the appearance of a linear trend in displacement (bottom graph, Fig. 7). As shown in the legend in Figure 7, the processing steps included padding the acceleration time series, filtering, and integrating twice.

Discussion

In many cases the low-frequency noise present in strong-motion data can be eliminated by low-cut filtering. In those cases, if zero pads of adequate length are appended to the beginning and end of the recorded data segment before filtering, and if these padded sections are retained for subsequent processing (including single and double integration

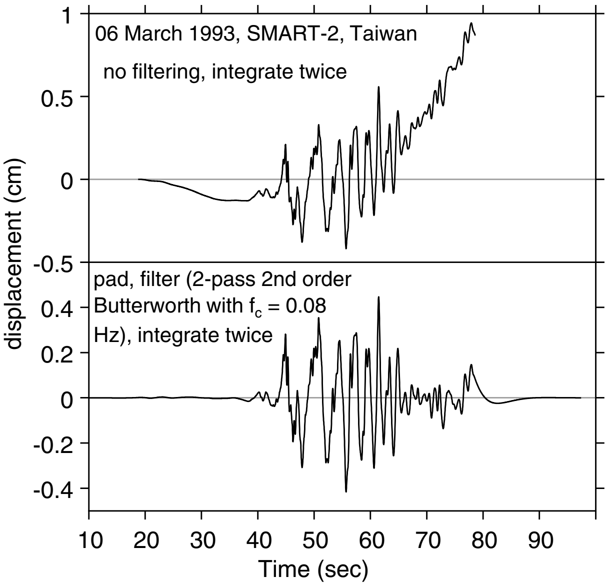

Figure 7. A digital acceleration time series obtained from the SMART-2 array in Taiwan, processed as indicated in the legends of the two graphs; for both graphs the mean of the pre-*P*-wave portion of the record was removed from the whole record before further processing. This record was used in figure 5 of Chiu (1997) to illustrate a method of baseline correction for digital recordings.

to obtain velocity and displacement), then no further corrections are needed.

The ability of filtering alone to remove low-frequency distortions clearly depends on the nature of the distortions. Figures 20 and 21 in Boore *et al.* (2002) give an example where prefiltering baseline correction was needed, in this case a series of baseline offsets, to avoid distortions in the derived velocity and displacement waveforms. In my experience this is the exception rather than the rule.

If the padded sections are not retained in further processing, in general there will be an incompatibility of results derived from the filtered accelerations (including velocity and displacement time series and Fourier and response spectra) compared with the results obtained when the padded sections are retained. This incompatibility is an issue only if users do further processing of the data provided by the responsible agencies. If only elastic response spectra or peak ground motions provided by the agencies are used no further processing would be required. The incompatibility problem may arise, however, when the distributed (and pad-stripped) acceleration time series are used in structural or site response calculations. This problem is particularly important when the resonant periods of the structures are long or when there is significant softening of the structures due to nonlinear effects, thus shifting the resonant periods to longer period. The obvious solution to this incompatibility is to provide the complete processed time series, including the padded portions. For completeness, the interval corresponding to the original segment of data should also be provided.

Acknowledgments

I thank Hung-Chie Chiu for providing the SMART-2 data used in the last figure and Julian Bommer, Vladimir Graizer, Praveen K. Malhotra, Tony Shakal, and Chris Stephens for reviewing the article.

References

Boore, D. M., and S. Akkar (2003). Effect of causal and acausal filters on elastic and inelastic response spectra, *Earthquake Eng. Struct. Dyn.* **32,** 1729–1748.

Boore, D. M., C. D. Stephens, and W. B. Joyner (2002). Comments on baseline correction of digital strong-motion data: Examples from the 1999 Hector Mine, California, earthquake, *Bull. Seism. Soc. Am.* **92,** 1543–1560.

Chiu, H.-C. (1997). Stable baseline correction of digital strong-motion data, *Bull. Seism. Soc. Am.* **87,** 932–944.

Converse, A. M., and A. G. Brady (1992). BAP: basic strong-motion accelerogram processing software, version 1.0, *U.S. Geol. Surv. Open-File Rept. 92-296A*, 174 pp.

Hudson, D. E. (1979). *Reading and Interpreting Strong Motion Accelerograms*, Vol. **1,** Earthquake Engineering Research Institute Monograph, Berkeley, California, 112 pp.

Lee, V. W., and M. D. Trifunac (1984). Current developments in data processing of strong motion accelerograms, Rept. 84-01, 99 pp., Department of Civil Engineering, University of Southern California, Los Angeles.

Malhotra, P. K. (2001). Response spectrum of incompatible acceleration, velocity, and displacement time histories, *Earthquake Eng. Struct. Dyn.* **30,** 279–286.

Press, W. H., S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery (1992). *Numerical Recipes in FORTRAN: The Art of Scientific Computing*, Cambridge University Press, Cambridge, 963 pp.

Seekins, L. C., A. G. Brady, C. Carpenter, and N. Brown (1992). Digitized strong-motion accelerograms of North and Central American earthquakes 1933–1986, *U.S. Geol. Surv. Digital Data Series DDS-7.*

Shyam-Sunder, S., and J. J. Connor (1982). A new procedure for processing strong-motion earthquake signals, *Bull. Seism. Soc. Am.* **72,** 643–661.

Todorovska, M. I., and W. Lee (2004). Data processing and recording at University of Southern California, in *Proc. Workshop on Strong-Motion Record Processing*, Consortium of Organizations for Strong-Motion Observation Systems (COSMOS), May 26–27, 2004, Richmond, California.

Trifunac, M. D. (1971). Zero baseline correction of strong-motion accelerograms, *Bull. Seism. Soc. Am.* **61,** 1201–1211.

U.S. Geological Survey, MS 977
345 Middlefield Road
Menlo Park, California 94025
boore@usgs.gov

Manuscript received 16 August 2004.