

A Compatible Baseline Correction Algorithm for Strong-Motion Data

Hung-Chie Chiu *

Institute of Earth Sciences, Academia Sinica, Taipei, Taiwan

Received 13 July 2011, accepted 25 October 2011

ABSTRACT

In physics, acceleration, velocity, and displacement should be convertible with each other. However, many strong-motion data do not meet this requirement; the double integration of a disseminated acceleration might not be the same as the corresponding disseminated displacement. This data incompatibility influences not only on the waveform but also on the derived terms from acceleration, such as response spectra. This can become a serious problem in the calculation of a nonlinear response (Pecknold and Riddell 1978, 1979). We show that the non-zero initial value of waveforms is the direct source of the data incompatibility, and propose a numerical algorithm to solve the problem by adding a prefix acceleration impulse. We suggest a polynomial function of order of three as the impulse function. The coefficients of this polynomial function can be determined by initial acceleration, velocity and displacement which can be obtained by routine data processing. Numerical tests show this added impulse can effectively remove the data incompatibility and cause negligible effects on waveforms and response spectra.

Key words: Ground acceleration, Baseline correction, Data incompatibility, Strong motion, Response spectra

Citation: Chiu, H. C., 2012: A compatible baseline correction algorithm for strong-motion data. *Terr. Atmos. Ocean. Sci.*, 23, 171-180, doi: 10.3319/TAO.2011.10.25.01(T)

1. INTRODUCTION

For ground motion, in principle, acceleration, velocity, and displacement should be interchangeable. The integration of an acceleration waveform results in a velocity time history while displacement can be obtained by the integration of the velocity waveform or a double integration of the acceleration waveform. However, this basic principle of physics is not valid for many disseminated data from various strong-motion data providers. Some examples are shown in Fig. 1. From the top to the bottom there are three sets of three-component waveforms disseminated by the Berkeley Digital Seismic Network (BDSN), the National Strong-Motion Program of the US Geological Survey (USGS), and the California Strong-Motion Instrumentation Program of California Geological Survey (CGS). The waveforms in columns 1 and 2 are disseminated acceleration and displacement waveforms provided by The Consortium of Organizations for Strong-Motion Observation Systems (COSMOS) data center. The displacements in the third column

are obtained by taking double integration of the acceleration waveforms in column 1 and assuming zero initial velocity and displacement. All waveforms in column 3 are different from the corresponding waveforms in column 2 and show a significant linear baseline drift.

This incompatibility not only appears in displacement waveforms, but also affects the terms derived from acceleration, such as the long-period portion of the response spectra (Malhotra 2001). Pecknold and Riddell (1978, 1979) also found that incompatible acceleration might cause a low frequency distortion in the calculation of structural response and that effect becomes more serious in nonlinear cases (Pecknold and Riddell 1978).

Several possible causes have been reported to be responsible for this incompatibility. Pecknold and Riddell (1978, 1979) attributed the data incompatibility to an incomplete (truncated) record in analog-type data. Chiu (1997) showed that the linear-trend baseline error found in the displacement waveform is due to existing non-zero initial velocity. Boore (2005) pointed out that wrap-around pollution (Press et al. 1992), and pad-stripped data are two other effects that might also cause incompatibility.

* Corresponding author
E-mail: chiu@earth.sinica.edu.tw

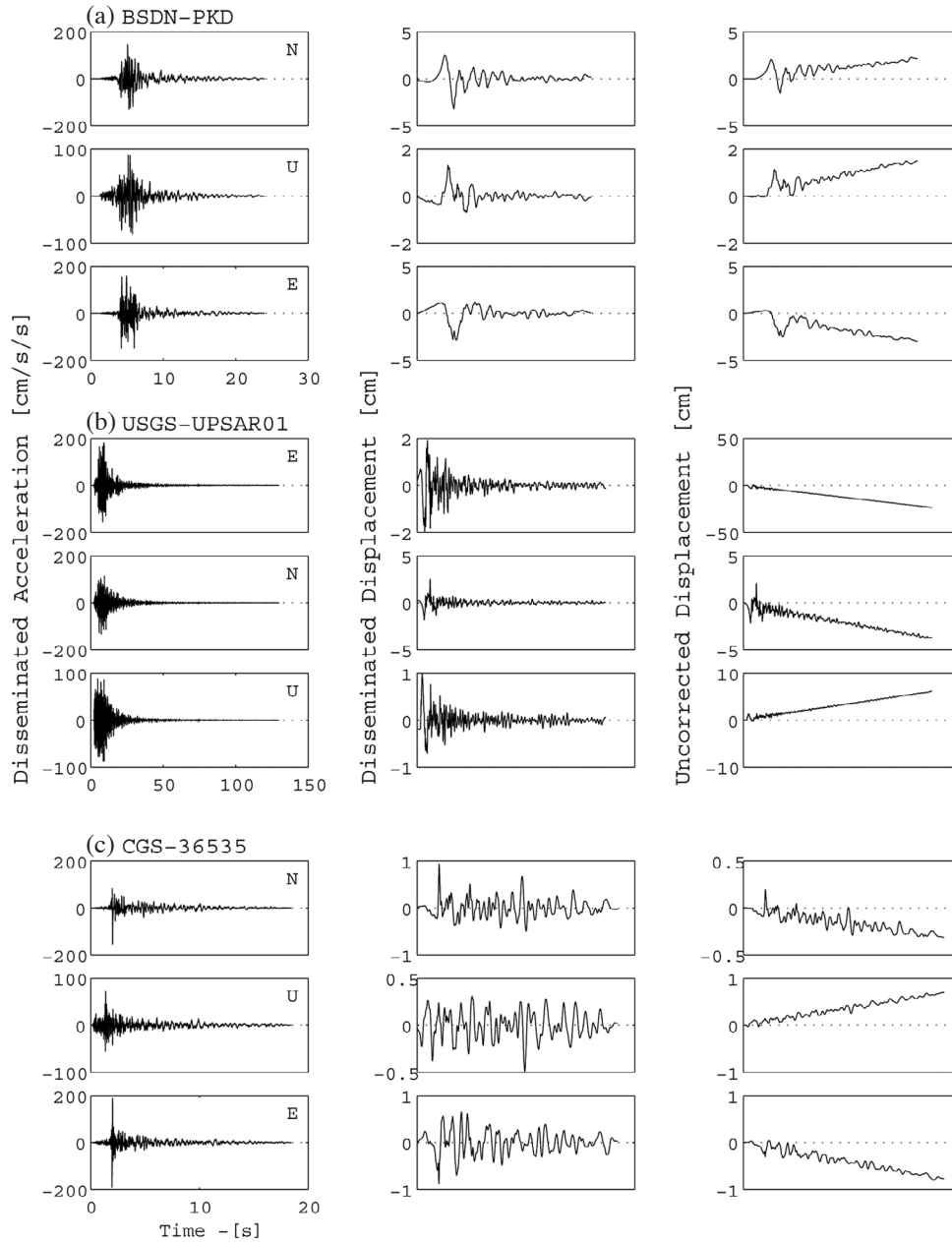


Fig. 1. Three sets of the 2004 Parkfield earthquake data with incompatible waveforms. Three columns of waveforms from left to the right are the disseminated acceleration, disseminated displacements and the uncorrected displacements obtained by taking a double integration of acceleration waveforms in column 1.

In order to remove or minimize the effects of the data incompatibility found in disseminated strong-motion waveforms, Pecknold and Riddell (1978, 1979) suggested adding a prefix acceleration impulse. They defined the impulse as a three-term function which corresponded to the initial acceleration, velocity and displacement respectively. In their approach each term is multiplied by a third order influence function. This influence function is complicated and lacks a clear physical meaning. Boore (2005) suggested adding adequate zero pads before filtering and retaining the padded sections in subsequent processing. Adding zero pads

only reduce the effects of the non-zero initial value and it requires a long pad to effectively reduce the waveform incompatibility. To reduce the effect of incompatibility on the long-period response spectra, Malhotra (2001) combined equilibrium equations for both stiff and flexible systems to obtain a new equation that could work well in calculating response spectra for both a stiff, high-frequency system and a flexible, low-frequency system. Malhotra's approach only reduces the effects of waveform incompatibility on the calculation of response spectra and did nothing with regard to removing data incompatibility in waveforms. In this study,

we show that the non-zero initial velocity is the direct source of data incompatibility and propose a numerical algorithm to solve this problem.

2. CAUSES OF DATA INCOMPATIBILITY IN STRONG-MOTION DATA

The pattern of the baseline drift shown in the waveforms of Fig. 1 is a vital clue for understanding the causes of data incompatibility. The nearly linear baseline drift implies that the major baseline drift in these waveforms results from a non-zero initial velocity. This non-zero initial value is very common in disseminated strong-motion data because of random noise (Chiu 1997) and the waveform truncation of triggered-type recordings (Pecknold and Riddell 1978, 1979; Chiu 1997). However, a non-zero initial acceleration doesn't cause a significant baseline drift in velocity waveforms because most disseminated acceleration waveforms have been high-pass filtered which forces both ends of the integrated velocity waveform to be zero. If the initial ground velocity is not zero, the integrated velocity waveform will produce a constant drift from the true ground velocity and result in a displacement waveform with a linear trend by taking another integration of this velocity waveform.

A large non-zero initial value happens very often in an incomplete data which missing first P waves or both first P and S waves. This type of data is found largely in analog data from triggered-type recordings where the data was truncated due to the mechanical delay of the starter. In some cases, the truncated primary P wave can be found in digital data when the recorder was triggered by later arrivals and the pre-event memory was insufficiently long. Fortunately, data truncation is no longer a problem with modern digital recorders as these problems can be solved by selecting the proper pre-event memory. Even for the digital data, the non-

zero initial value is still true because of the existing of background noise in all records. However, these initial values are at the level of background noise which is much smaller than that in the truncated waveforms.

The truncation of waveform is often found at the tail in a strong-motion record. Although the design of modern trigger-type recorders allows the recording to be sustained for an extended time period (post-event time) after the ground shaking drops below the trigger level, the ground motion in most cases does not totally cease shaking prior to the end of record. In general, ground motion in the tail of a recording is larger than that in the pre-event portion because additional seismic waves are added to background noise. For example, for a set of digital accelerations, velocity and displacement waveforms are shown in Fig. 2a and their enlarged waveforms of the first and last 12 seconds are given in Fig. 2b and c respectively. Apparently, the amplitudes of waveforms in Fig. 2c are much larger than that of the corresponding waveforms in Fig. 2b.

Although non-zero values are found in both ends of the record, data incompatibility is mainly due to the non-zero initial value. To show the effects of initial velocity on the final displacement offset, a numerical test was performed and the results are shown in Fig. 3. All six displacement traces in this figure are derived from the same disseminated acceleration waveform by double integrating the waveform. However, different starting points were selected for making the numerical integration. The starting time t_i (in seconds) for each trace is marked at the end of the trace. All six starting times are within the pre-event recording where ground acceleration is below $0.3 \text{ cm s}^{-1} \text{ s}^{-1}$ and velocity less than 0.03 cm s^{-1} . The results indicate that the significant baseline offset and large variation in baseline are mainly due to the small differences in initial values.

To show the relationship between the initial values and the final offset in displacement, a similar test was carried

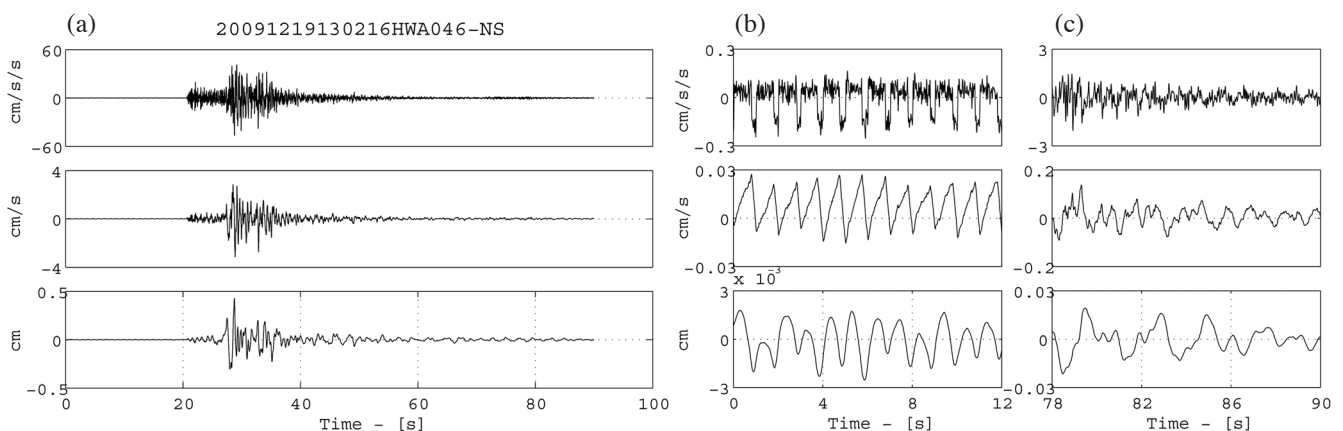


Fig. 2. One of the typical digital accelerations and its velocity and displacement waveforms are shown in (a). The enlargement of their first 12-second records is shown in (b), and the enlargement of the last 12-second records is shown in (c).

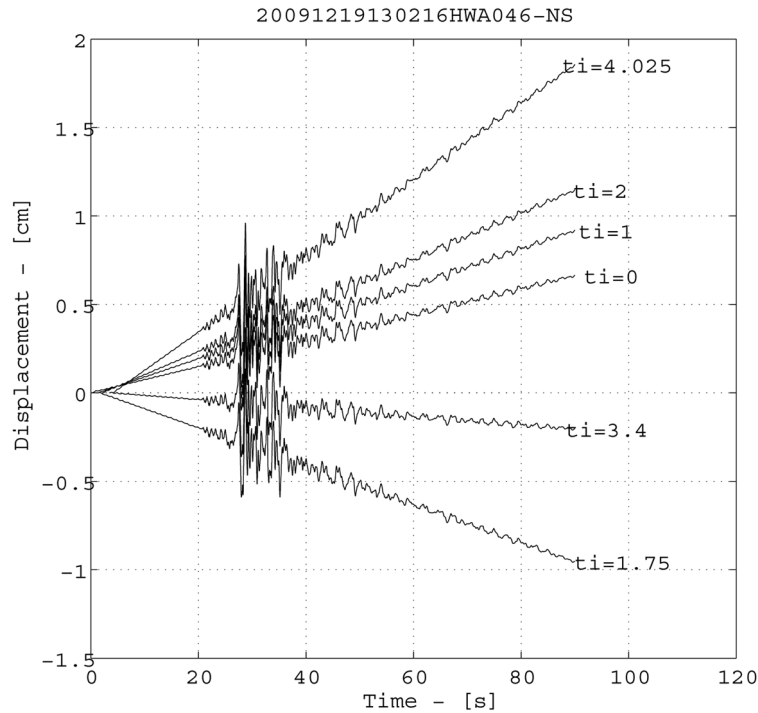


Fig. 3. Displacement waveforms obtained by double integration of the same acceleration waveforms as those in Fig. 2 with various starting points at t_i , and the corresponding t_i (in seconds) is marked at the tail of each trace.

out, but with an increased number of measurements and a change in plotting format. The calculations were for each 0.1 second increasing from the starting time from 0 to 12 seconds, giving a total of 121 measurements. Results are shown in the left portion of Fig. 4. It can be seen that the final displacement offset is proportional to the initial velocity, suggesting that the initial velocity is the only major source of displacement offsets.

A similar test for measuring the effect of a truncated tail is shown in the right portion of Fig. 4. Keeping the same starting point ($t_i = 0$) and cutting the tail in lengths from 0 to 12 seconds for each 0.1-second step, again we measured the final displacement for each sample. Instead of measuring the initial velocity in the left plot, the final velocity was measured after cutting data points at the tail of a record. As shown in the plot, the final displacement is almost constant for all 121 samples and the constant value depends on the initial velocity while the small variation in final displacement is due to the small difference in data length and ground motion variation at the end point. This result shows that the truncation at the end of a record does not contribute much to the final offset of the incompatible displacement waveform.

To summarize the causes of data incompatibility discussed above, results suggest that the incompatibility of disseminated strong-motion data is primarily due to non-zero initial velocity. This conclusion implies that removing or minimizing the waveform incompatibility can be done by

modifying these initial values to be zero or close to zero.

3. METHOD

Since the incompatibility comes from the non-zero initial value of the disseminated data, we can either drop data at the beginning of a record to let the new initial value to be close to zero or add a prefix impulse to have a zero initial value. Since selecting a new initial value close to zero only reduces the effect of incompatibility, we choose the second method for removing the data incompatibility by adding a prefix impulse. In the following discussion, a set of processed data based on the 1997 algorithm (Chiu 1997) is used for demonstrating the proposed method. The basic principle underlying this method is adding a prefix impulse to the disseminated acceleration waveforms so that these waveforms have zero initial values. If the initial velocity becomes zero, corrections in the 1997 Algorithm for removing the effects of initial velocity become unnecessary and the data automatically become compatible.

To distinguish various waveforms that appear in the following discussion, the corrected acceleration, velocity, and displacement arising from traditional methods are named as disseminated data, whilst the integration and double integration of the disseminated acceleration are named as uncorrected velocity and displacement. The waveforms after the correction, using the proposed method, are called corrected waveforms.

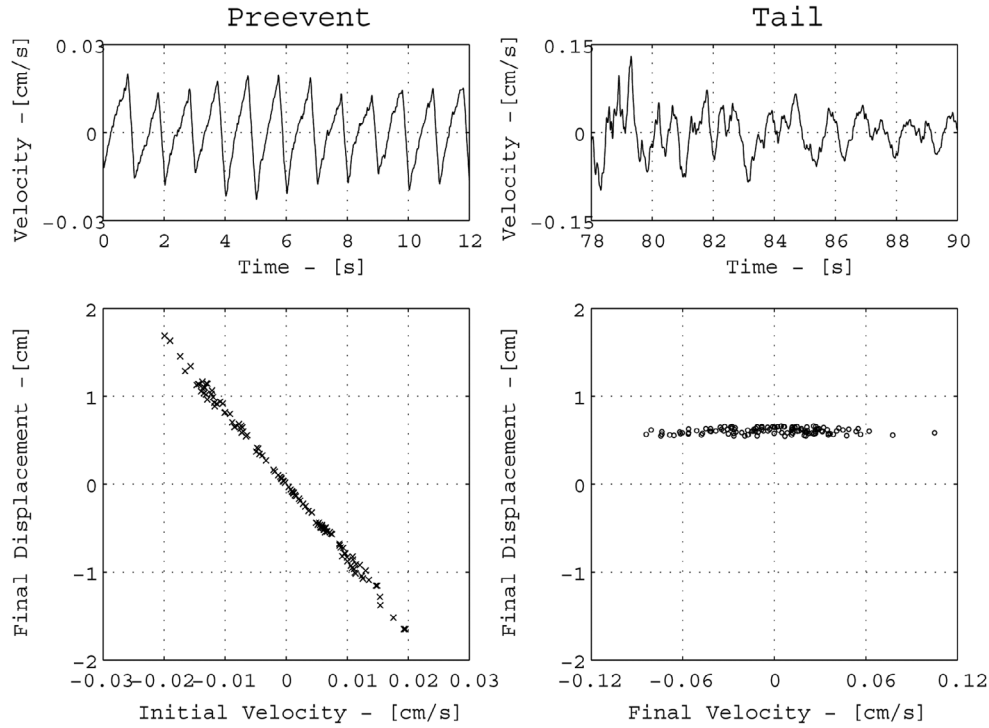


Fig. 4. Relationship between the final displacement and initial velocity (bottom left). The starting point to calculate the displacement is selected for every 0.1 of a second from 0 to 12 s and the corresponding velocity waveform is given in the top left. The relationship between the final displacement and final velocity are shown in the bottom right. The last data point in each calculation of the displacement is selected for every 0.1 of a second from 78 to 90 s and the corresponding velocity waveform is given in the top right.

Selecting the acceleration impulse in this study uses a more general and more straightforward approach than that suggested by Pecknold and Riddell (1979). We chose the impulse to satisfy the following two conditions:

- (a) The initial values of the impulse acceleration $I_a(0)$, the impulse velocity $I_v(0)$ and the impulse displacement $I_d(0)$ are all zero.
- (b) If l is the impulse length, the final values of the acceleration impulse $I_a(l)$, velocity impulse $I_v(l)$ and displacement impulse $I_d(l)$ should be equal to the corresponding initial values of the disseminated acceleration a_0 , velocity v_0 and displacement d_0 .

In general, these two conditions have six unknowns and require a set of six linear equations to provide a unique solution. However, the number of unknowns can be reduced to three if we properly select the base functions and let them satisfy condition (a). Many combinations of linear equations can satisfy this requirement. Any function and its integration and double integration that have zero initial values can be selected as the base functions. If t is the time variable, then any power function of t belongs to this family. However, a high-order power function will result in a large-amplitude and a complicated impulse function, which is undesirable. In this study, we suggest selecting a third order polynomial equation as the acceleration impulse:

$$I_a = et + ft^2 + gt^3 \quad (1)$$

where e, f and g are three unknowns to be determined. Integrating and double integrating I_a by assuming $I_v(0) = 0$ and $I_d(0) = 0$ result in:

$$I_v = \frac{1}{2}et^2 + \frac{1}{3}ft^3 + \frac{1}{4}gt^4 \quad (2)$$

$$I_d = \frac{1}{6}et^3 + \frac{1}{12}ft^4 + \frac{1}{20}gt^5 \quad (3)$$

These three equations automatically satisfy condition (a). If we apply condition (b) to these equations, we obtain:

$$\begin{cases} a_0 = el + ft^2 + gl^3 \\ v_0 = \frac{e}{2}l^2 + \frac{f}{3}l^3 + \frac{g}{4}l^4 \\ d_0 = \frac{e}{6}l^3 + \frac{f}{12}l^4 + \frac{g}{20}l^5 \end{cases} \quad (4)$$

For a given impulse length l , this system of equations has a unique solution for coefficients e, f and g . After solving this

system of equations, we can substitute these three coefficients back to (1) and give the required impulse acceleration.

4. RESULTS AND DISCUSSION

Since the proposed method is built on the 1997 Algorithm, it was expected to work well for those waveforms processed using the 1997 Algorithm. Two sets of three component data, as shown in Fig. 5, were selected to show the performance of the proposed method. The disseminated displacement waveforms corrected using the 1997 Algorithm are shown in the first column while the corresponding uncorrected and corrected displacements are shown in the second and third columns. The first (Fig. 5a) set of data has a smaller baseline drift while the second (Fig. 5b) set of data shows significant baseline drifts. Regardless of the initial values, both sets of the corrected displacements match the disseminated waveforms equally well.

Next, the 2004 Parkfield strong-motion data recorded at USGS station 1797 and processed by COSMOS were used as another example to demonstrate the proposed method. Although there is no visual difference between the disseminated and uncorrected velocity waveforms (Fig. 6a), the significant baseline drift appears in the uncorrected

displacement waveforms (Fig. 6b). On the other hand, both disseminated and corrected data have identical velocity (Fig. 6a) and displacement (Fig. 6b) waveforms, except that the corrected data have a prefix impulse.

To show what had been modified on the corrected acceleration waveform and the effects of correction on the velocity and displacement waveforms, we plot the first 1.5 seconds ($0 \sim 1.5$ s) of three types (disseminated, uncorrected and corrected) waveforms and the 1-second prefix ($-1 \sim 0$ s) added in the proposed method as shown in Fig. 7. In the first row are three-component acceleration waveforms. There is no difference among these three types of waveforms except the 1-second acceleration impulse in the corrected acceleration waveforms. The shape and amplitude of these impulses are different between components because they were determined by the initial values of acceleration, velocity and displacement of the corresponding component. A comparison of velocity waveforms is shown in second row. Again, the disseminated and corrected waveforms are the same except the later has a velocity impulse. The disseminated and the corrected waveforms are still identical in displacement except that the latter has a prefix while the difference between these two types of waveform and the uncorrected displacement waveform is an extra linear trend. Overall,

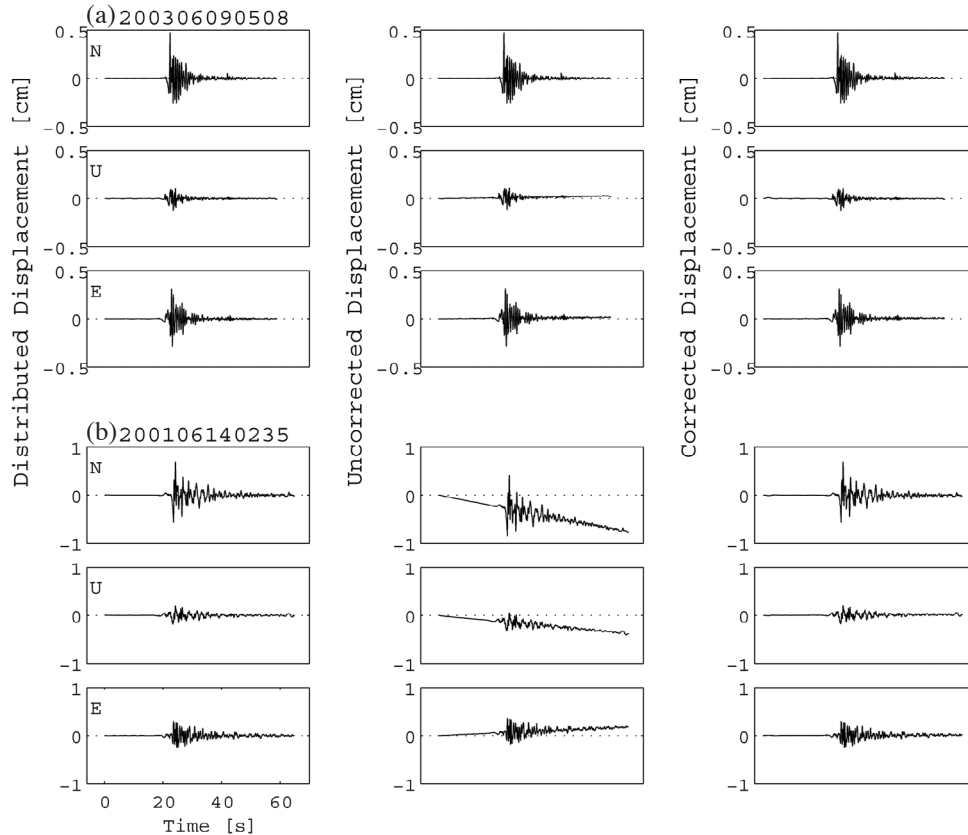


Fig. 5. Two sets of the typical digital strong-motion data recorded at station ILA066 of the Taiwan TSMIP strong-motion network are selected to demonstrate the performance of the proposed methods. The displacement waveforms in the first column are corrected using the 1997 Algorithm. The displacement waveforms without and with compatibility correction are shown in columns 2 and 3 respectively.

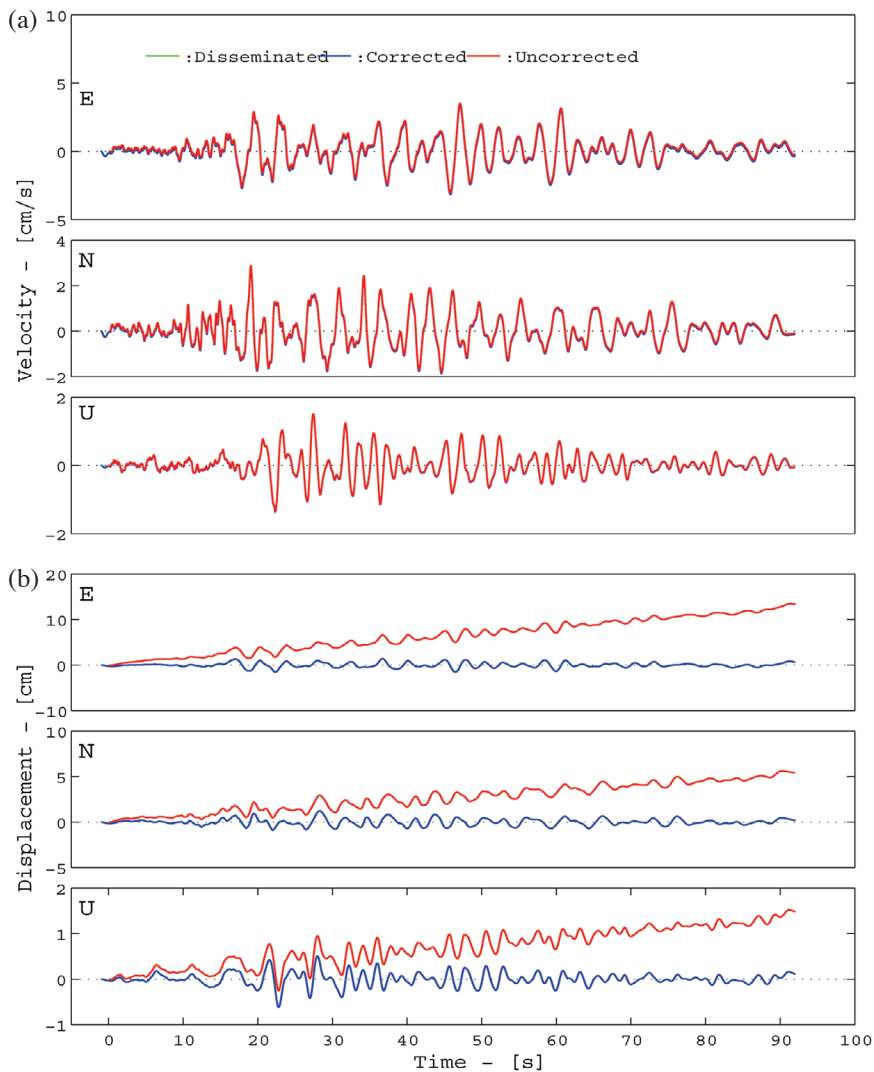


Fig. 6. A set of three-component strong-motion data set recorded at USGS station 1797 in the 2004 Parkfield Earthquake is selected to show the differences in velocity (a) and displacement (b) waveforms among three types of data. The positive direction of the waveform is marked on the left upper corner. Each plot includes three types of data which are disseminated, uncorrected and corrected data. The uncorrected data is obtained by taking integration or a double integration of the disseminated acceleration.

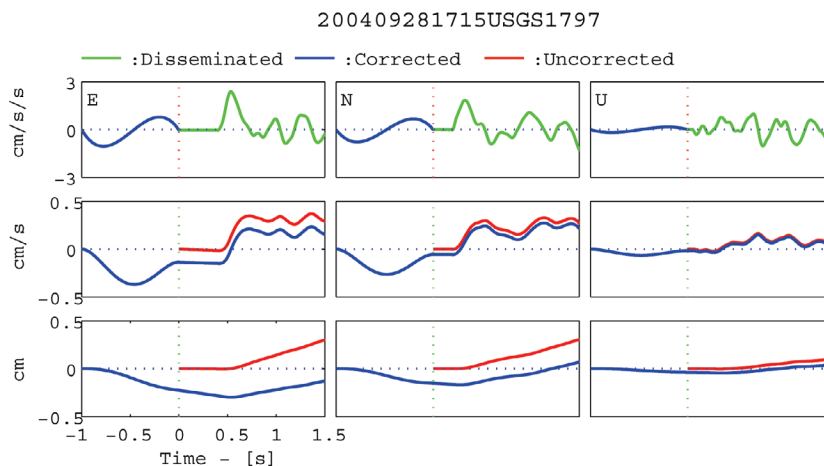


Fig. 7. Comparisons of the first 1.5 s waveforms among the disseminated (thick line), uncorrected and corrected data. Three rows from top to bottom are acceleration, velocity and displacement waveforms and the three columns from left to the right are EW, NS, and vertical components.

the proposed method is an exact correction algorithm. The proposed method adds a prefix impulse to the acceleration waveform so that the end point of impulse matches the initial value of disseminated waveform while the remainder of the waveforms stays the same.

Since an extra impulse is added to the disseminated acceleration waveform for removing data incompatibility, it

is better to keep the amplitude of this impulse as smaller as possible. According to Eq. (4), ' l ' is the only variable that can change the amplitude of impulse. In the following test, we will examine the effect of the impulse length on the amplitude of impulse.

Impulse waveforms for three cases with impulse lengths of 1, 2 and 5 s are given in Figs. 8a, b, and c re-

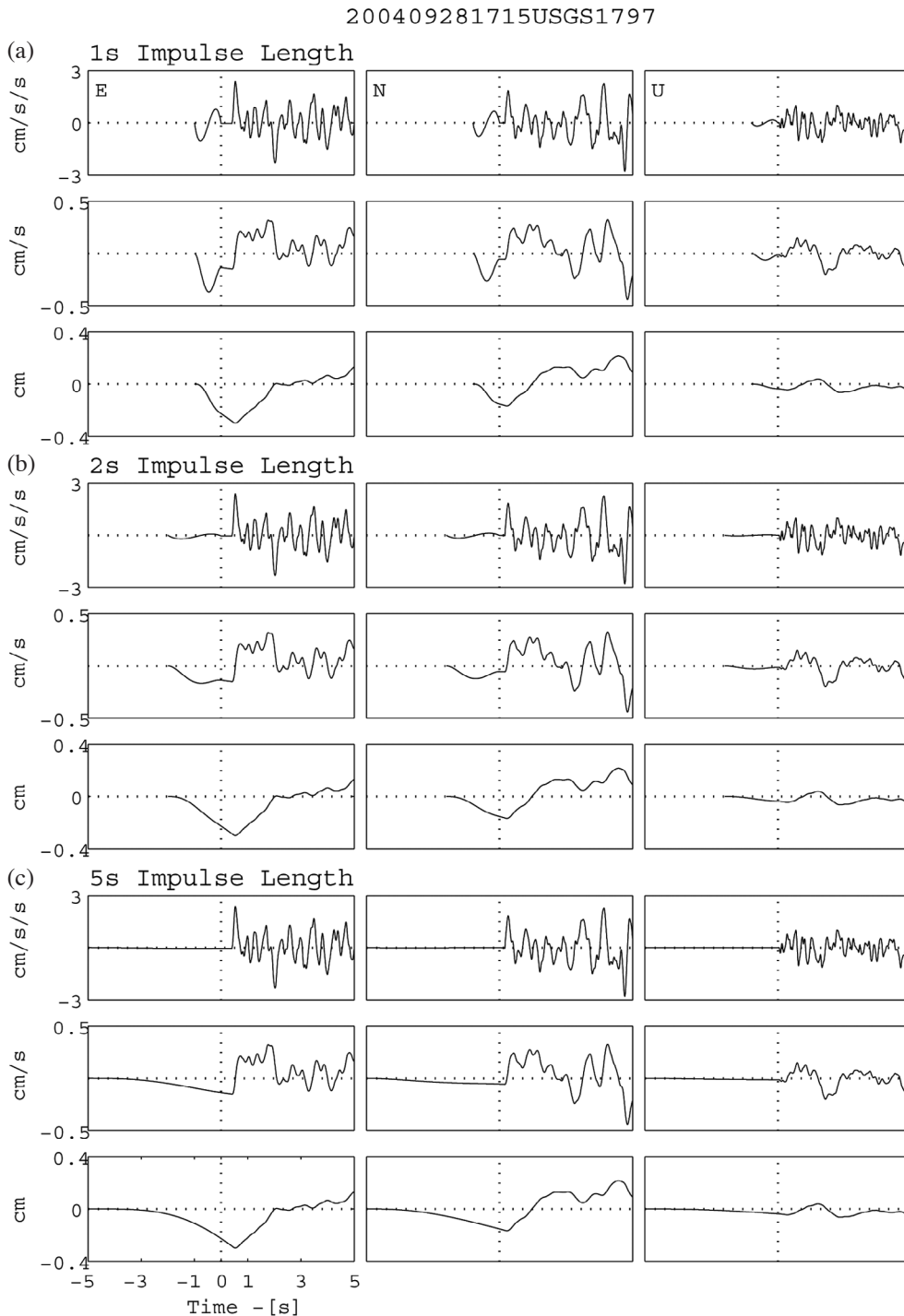


Fig. 8. The effects of the impulse length on the shape and amplitude of prefix impulse. This figure has a similar format to that of Fig. 7. The impulse lengths in (a), (b), and (c) are 1, 2 and 5 s, respectively.

spectively. Results show that the increasing impulse length not only reduces amplitude but also changes the shape of impulse. The amplitude reduction is very fast as the length of the impulse increases. For impulse lengths of 1, 2 and 5 s, the peak amplitude of the impulse in the E component are 1.0586, 0.2087 and 0.0468 $\text{cm s}^{-1} \text{s}^{-1}$, respectively. Thus, selecting a longer length for the prefix impulse can effectively reduce its amplitude.

Finally, the effect of a prefix impulse on response spectra should be checked because it is of great concern in engineering applications. Since our corrections do not have major modifications on the acceleration waveform, it is expected that the proposed correction has a limited effect on the response spectra. A numerical verification is given in Fig. 9. Three components of the disseminated and corrected acceleration are plotted in the top three seismic traces and the 5% pseudo velocity spectra are shown in the bottom of Fig. 9. Each panel in the figure has two traces. Except for the small difference in the long period, these spectra are almost identical for most periods. As expected, the effect

of impulse on the calculation of response spectra can be ignored.

5. CONCLUSIONS

This study has shown that data incompatibility, found in many of the disseminated strong-motion data, is mainly due to the non-zero initial velocity and the proposed method can effectively remove the data incompatibility by adding a prefix impulse in acceleration waveforms. In general, there are several ways to construct an acceleration impulse for removing incompatibility. However, a polynomial function of an order of 3 is good and simple enough for our purpose. The coefficients of this polynomial function depend on the initial acceleration, velocity, and displacement which can be obtained by routine data processing.

The proposed method does not change the waveforms of the disseminated data. A specific prefix impulse is added to the disseminated acceleration so that the velocity and displacement waveforms, obtained by the integration and

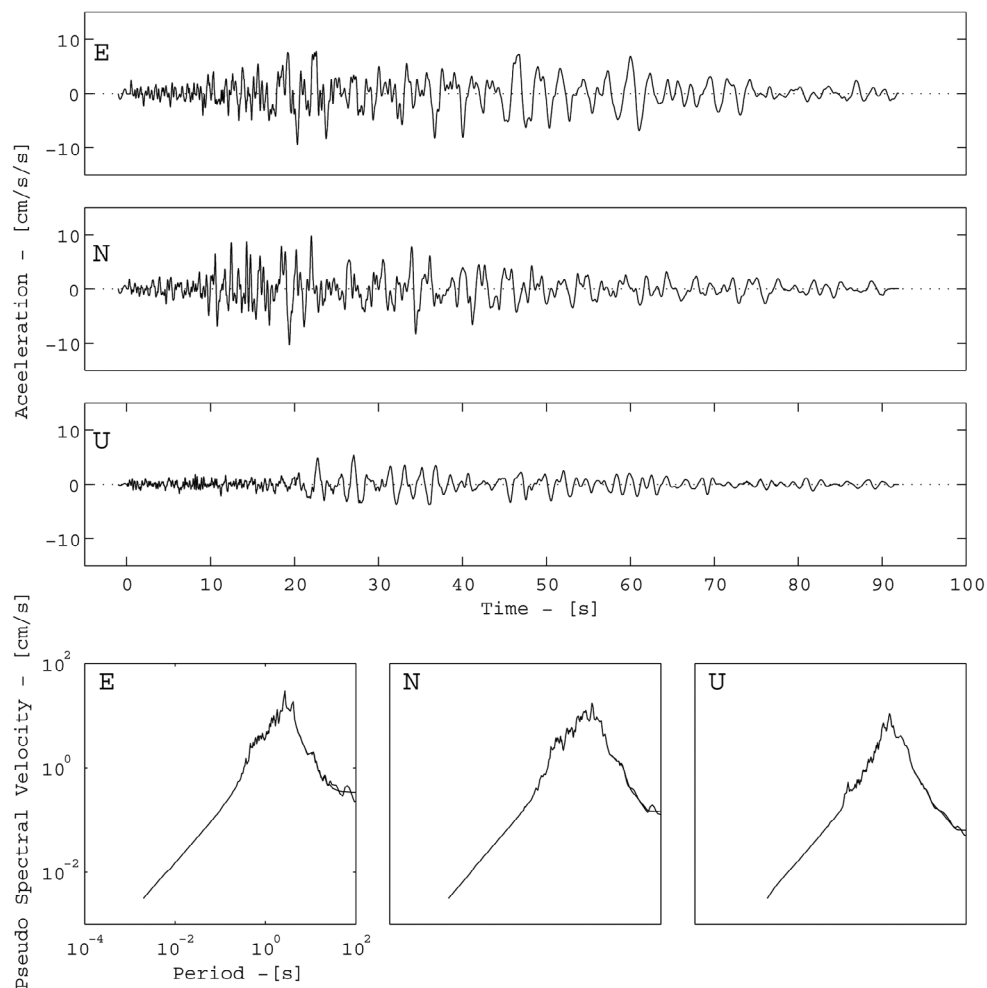


Fig. 9. Comparisons of three-component waveforms and their 5% pseudo spectral velocities between the disseminated and the corrected accelerations using the proposed method.

double integration respectively, are exactly the same with the disseminated waveforms after the prefix impulse. Since the prefix impulse is much shorter and much smaller than the seismic signal, the effects of this extra impulse on both the waveform and response spectra are negligible.

Acknowledgements We thank two anonymous reviewers for their helpful comments. This research was supported by Academia Sinica and the National Science Council (under the Grant NSC 99-2116-M-001-023), Taiwan, Republic of China. The waveform data was provided by the Central Weather Bureau in Taiwan and COSMOS (<http://www.cosmos-eq.org/>) in USA, The support from these organizations is gratefully acknowledged.

REFERENCES

- Boore, D. M., 2005: On pads and filters: Processing strong-motion data. *Bull. Seismol. Soc. Am.*, **95**, 745-750, doi: 10.1785/0120040160. [[Link](#)]
- Chiu, H. C., 1997: Stable baseline correction of digital strong-motion data. *Bull. Seismol. Soc. Am.*, **87**, 932-944.
- Malhotra, P. K., 2001: Response spectrum of incompatible acceleration, velocity and displacement histories. *Earthq. Eng. Struct. Dyn.*, **30**, 279-286, doi: 10.1002/1096-9845(200102)30:2<279::AID-EQE11>3.0.CO;2-Q. [[Link](#)]
- Pecknold, D. A. and R. Riddell, 1978: Effect of initial base motion on response spectra. *J. Eng. Mech. Div., ASCE*, **104**, 485-491.
- Pecknold, D. A. and R. Riddell, 1979: Closure to discussion on effect of initial base motion on response spectra. *J. Eng. Mech. Div., ASCE*, **105**, 1057-1060.
- 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, 933 pp.