Empirical baseline correction of strong-motion data

by Rongjiang Wang, last update June 2025

Previous studies have shown that the co-seismic displacement at near-fault stations can be retrieved by double integrating strong-motion records after an empirical baseline correction. Generally, the strong-motion derived displacement is considerably less accurate than that measured using, for example, the GNSS and InSAR techniques. However, it is useful for rapid earthquake source studies, especially when the GNSS network has an insufficient coverage of the earthquake epicentral area. So far, it is normally necessary to filter the strong-motion based waveform data with a bandpass filter before they are used for kinematic source inversions. After a successful baseline correction, the bandpass filter can be replaced by a lowpass filter, so that the low-frequency content of waveform including the derived static displacement, which provides strong constraints on the spatial distribution of earthquake fault slips, can be used.

Most empirical baseline corrections are performed either manually or automatically using a bilinear approach (e.g., Iwan et al. 1985, Boore 2001, Wu and Wu 2007, Wang et al. 2011). In most cases, raw strong-motion accelerogram include all a pre-seismic baseline offset, which can be easily removed. During the period of strong ground shaking, further baseline variation is generally generated by (1) the perturbation of the earth's gravity field, (2) the free-air effect on the vertical component and (3) the effect of ground tilt on the two horizontal components. Occasionally, also the vertical component can be affected by the ground tilt because it is not really installed vertically, particularly in case of borehole strong-motion sensors. After the strong ground shaking is over, the co-seismic baseline variation is going to be stabilized to a post-seismic permanent baseline offset. The latter can be easily estimated by the slop of post-seismic trend of the velocity seismogram integrated from the raw accelerogram. In principle, the co-seismic baseline variation cannot be corrected completely because it is not separable from the ground acceleration (the equivalence principle). However, its accumulated effect can be estimated because it results in a post-seismic offset of the uncorrected velocity seismogram. Therefore, the bilinear approach is aimed to correct an average of baseline shift during the strong ground shaking period followed by a post-seismic permanent baseline offset. The whole baseline correction on the velocity seismogram is represented by two connected line segments. Various manual or automatic schemes based on this approach differ only in the way how to select the start and stop time of the co-seismic baseline shift.

Here we present an updated scheme of the approach developed by Wang et al. (2011). To further minimize the uncertainty of this approach, we replace the bilinear correction by a natural curve correction obtained through iterative smoothing the uncorrected velocity seismogram. The new scheme consists of the following 3 steps.

Step 1. Integrate raw accelerograms to the uncorrected velocity seismograms after a pre-seismic baseline correction

- Assume a raw accelerogram $a_{raw}(t)$ is given for time window $t \in [t_0, t_{end}]$ with the known P wave arrival t_{pre} within the time window. In practice, we suggest a pre-seismic window $t_{pre} - t_0$ between 5 s and 30 s and a large enough signal window $t_{end} - t_{pre}$.
- We estimate the pre-seismic baseline offset

$$\Delta a_{pre} = \frac{1}{t_{pre} - t_0} \int_{t_0}^{t_{pre}} a_{raw}(\tau) d\tau, \tag{1}$$

  and remove it from the raw accelerogram to get an accelerogram only including seismically induced baseline errors,

$$a_0(t) = a_{raw}(t) - \Delta a_{pre}. \tag{2}$$

- To estimate the time when the co-seismic baseline shift is stabilized to a constant post-seismic offset, we introduce function

$$E(t) = \int_0^t |a_0(\tau)| d\tau, \tag{3}$$

and time t_γ satisfying $E(t_\gamma) = \gamma E(t_{end})$, and assume that $t_{pst} = t_{\gamma=85\%}$ can be regarded as the time when the co-seismic baseline shift has been stabilized.

- Integrate $a_0(t)$ to velocity seismogram,

$$v_0(t) = \int_0^t a_0(\tau) d\tau. \tag{4}$$

Step 2. Estimate post-seismic baseline shift and the starting velocity correction

- Calculate the post-seismic linear trend of $v_0(t)$ via least-squares regression

$$f(t) = v_{pst} + \frac{v_{end} - v_{pst}}{t_{end} - t_{pst}} (t - t_{pst}), \tag{5}$$

where v_{pst} and v_{end} are the start and end values of $f(t)$ at $t = t_{pst}$ and t_{end}, respectively.

- Define another function

$$g(t) = \begin{cases} 0, & t_0 \le t \le t_{pre}, \\ v_0(t), & t_{pre} < t < t_{pst}, \\ w_0(t), & t_{pst} \le t \le t_{end}, \end{cases} \tag{6}$$

as the starting correction curve, where function $w_0(t)$ is a weighted average of $v_0(t)$ and $f(t)$, i.e., the sum of right-tapered $v_0(t)$ and left-tapered $f(t)$, e.g.,

$$w_0(t) = v_0(t) \cos^2 \left[\frac{\pi(t - t_{pst})}{2(t_{end} - t_{pst})} \right] + f(t) \sin^2 \left[\frac{\pi(t - t_{pst})}{2(t_{end} - t_{pst})} \right]. \tag{7}$$

Step 3. Get final velocity correction via iterative smoothing

- Smooth $g(t)$ iteratively using a small moving window, but fixing $g(t_{pre}) = 0$ and $g(t_{end}) = v_{end}$,

$$g(t) \coloneqq \begin{cases} 0, & t_0 \le t \le t_{pre}, \\ \dfrac{1}{2\Delta t} \displaystyle\int_{t-\Delta t}^{t+\Delta t} g(\tau) d\tau, & t_{pre} < t < t_{end}, \\ v_{end}, & t = t_{end}, \end{cases} \tag{8}$$

where $a \coloneqq b$ means updating a by b. The smoothing process is terminated when $g(t)$ appears no extremum after t_{pst} and at most only one extremum before t_{pst}. So $g(t)$ becomes a smooth and, in most cases, monotone curve.

- Set the velocity correction curve $v_{err}(t) = 0$ for $t_0 \le t \le t_{pre}$ and $v_{err}(t) = g(t)$ for $t_{pst} \le t \le t_{end}$. For the remaining co-seismic period $t_{pre} < t < t_{pst}$, $v_{err}(t)$ is obtained by modifying $g(t)$. In the case that the co-seismic and post-seismic baseline shift have the same sign, i.e., $g(t_{pst}) \cdot [g(t_{end}) - g(t_{pst})] \ge 0$, and the former is smaller than the latter, i.e., $\left| \frac{g(t_{pst})}{t_{pst} - t_{pre}} \right| < \left| \frac{g(t_{end}) - g(t_{pst})}{t_{end} - t_{pst}} \right|$, we shift t_{pre} rightward to $max[t_{pre}, (t_{pre} + 2t_{fzc})/3]$, where t_{fzc} is the time of zero-cross of the extrapolated post-seismic trend $f(t)$. The modified function $g_m(t)$ should (a) be smooth and monotone, (b) start with 0 at $t = t_{pre}$ and equal $g(t_{pst})$ at $t = t_{pst}$, and (c) best fit $g(t)$ for $t \in (t_{pst}, t_{pst})$ in the least-squares sense. Finally, the velocity-based baseline error is estimated by

$$v_{err}(t) = \begin{cases} 0, & t_0 \le t \le t_{pre}, \\ g_m(t), & t_{pre} < t < t_{pst}, \\ g(t), & t_{pst} \le t \le t_{end}. \end{cases} \tag{9}$$

- Make the baseline correction on the velocity seismogram

$$v(t) = v_0(t) - v_{err}(t) \tag{10}$$

and integrate it to the corrected displacement seismogram

$$u(t) = \int_0^t v(\tau)d\tau. \tag{11}$$

References

Boore, D. M. (2001). Effect of baseline corrections on displacement and response spectra for several recordings of the 1999 Chi-Chi, Taiwan, earthquake. *Bull. Seismol. Soc. Am. 91*, 1199-1211.

Iwan, W., M. Moser, and C. Peng (1985). Some observations on strong-motion earthquake measurement using a digital acceleration. *Bull. Seismol. Soc. Am. 75*, 1225-1246.

Wang, R., B. Schurr, C. Milkereit, Zh. Shao, and M. Jin (2011). An improved automatic scheme for empirical baseline correction of digital strong-motion records. *Bull. Seismol. Soc. Am. 101*, 2029-2044.

Wu, Y., and C. Wu (2007). Approximate recovery of coseismic deformation from Taiwan strong-motion records. *J. Seismol. 11*, 159–170.

User's manual of FORTRAN code "smblc2023"

1. The rar file "smblc2023-code+input.rar" includes the source code, the executable code for Windows system and an application example (Figs. 1 and 2). Users of Linux system can compile the code with, e.g., "gfortran":

```
......
......>cd SourceCode
....../SourceCode> gfortran -o smblc2023 *.f -O3
......
```

2. "smblc2023" reads all parameters from the input file. Users can copy the example input file, which is self-explanatory, and modify the parameters, but do not change their formats. All text lines starting with "#" are documentations, which will be ignored during reading.

3. Before starting "smblc2023", please generate a data folder like ".../SMData/" in the example and put all raw strong-motion data there, one ascii file for each station including 3 columns for the 3 components, respectively. Additionally, the data folder should include the "StationInfo.dat" file, as given in the example, for the necessary information about the earthquake and the raw strong-motion data.

4. All outputs are stored in another existing folder like ".../Output/" in the example.

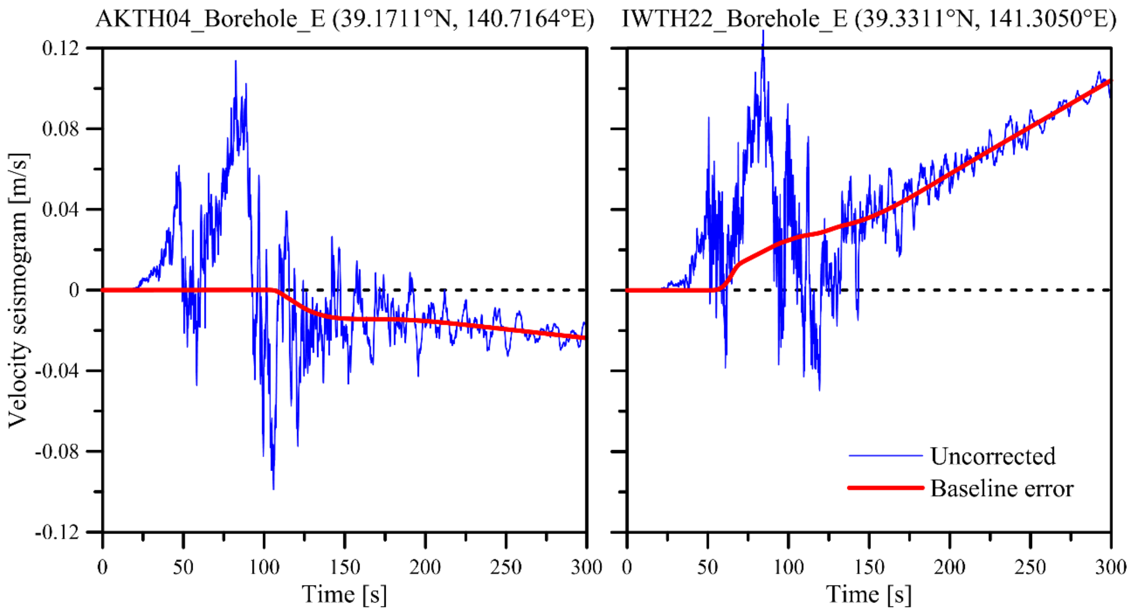

Figure 1. Two examples of the empirical baseline correction for strong-motion derived velocity seismograms of the 2011 Mw9.1 Tohoku earthquake, recorded by the Kik-Net borehole sensors.

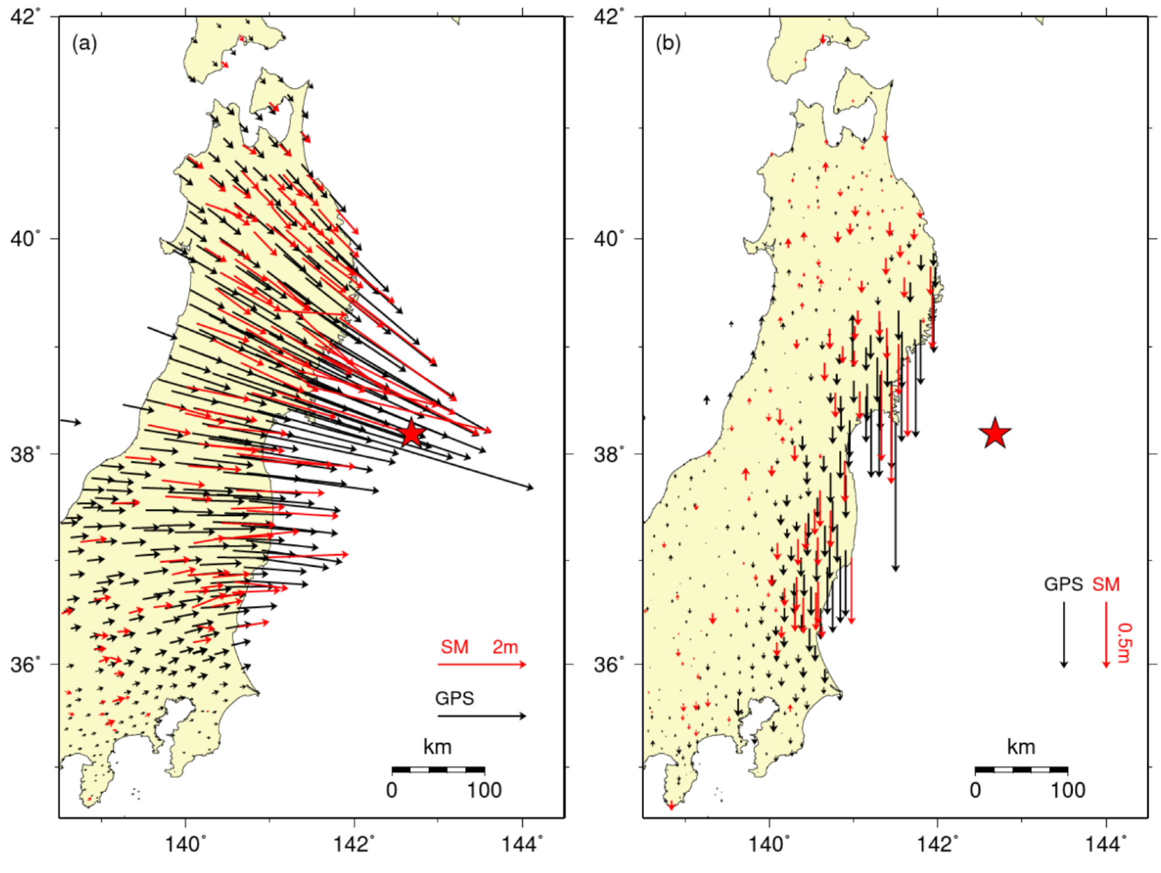

Figure 2. Static horizontal (a) and vertical (b) displacements of the 2011 Mw9.1 Tohoku earthquake, derived from the Kik-Net strong-motion records after the empirical baseline correction, in comparison with the GPS data.