

On robust and reliable automated baseline corrections for strong motion seismology

Diego Melgar,¹ Yehuda Bock,¹ Dominga Sanchez,² and Brendan W. Crowell¹

Received 28 November 2012; revised 14 February 2013; accepted 14 February 2013; published 31 March 2013.

[1] Computation of displacements from strong motion inertial sensors is to date an open problem. Two distinct methodologies have been proposed to solve it. One involves baseline corrections determined from the inertial data themselves and the other a combination with other geophysical sensors such as GPS. Here we analyze a proposed automated baseline correction algorithm using only accelerometer data and compare it to the results from the real-time combination of strong motion and GPS data. The analysis is performed on 48 collocated GPS and accelerometers in Japan that recorded the 2011 Mw 9.0 Tohoku-oki earthquake. We study the time and frequency domain behavior of both methodologies. We find that the error incurred from automated baseline corrections that rely on seismic data alone is complex and can be large in both the time and frequency domains of interest in seismological and engineering applications. The GPS/accelerometer combination has no such problems and can adequately recover broadband strong motion displacements for this event. The problems and ambiguities with baseline corrections and the success of the GPS/accelerometer combination lead us to advocate for instrument collocations as opposed to automated baseline correction algorithms for accelerometers.

Citation: Melgar, D., Y. Bock, D. Sanchez, and B. W. Crowell (2013), On robust and reliable automated baseline corrections for strong motion seismology, *J. Geophys. Res. Solid Earth*, 118, 1177–1187, doi:10.1002/jgrb.50135.

1. Introduction

[2] The range of motions produced by the seismic source is broad both in frequency and dynamic range. It is well known that no one sensor can capture all signals of interest to seismology and earthquake engineering [Havskov and Alguacil, 2002]. Seismologists typically rely on seismometers, whose response is related to the velocity of the ground, for measurement of small amplitude signals (*weak motion*), but for large amplitude signals, these sensitive instruments saturate or clip. In this case, strong motion sensors, whose response is related to the acceleration of the ground and have lower gains, are preferred. Modern observatory grade accelerometers rely on the force feedback principle and can measure motions as small as 1 nm at 1 Hz and 100 nm at 0.1 Hz and accelerations of up to 4g. Furthermore, their frequency response is flat from DC to 50–200 Hz [Havskov and Alguacil, 2002].

[3] Thus, in principle, there should be no difficulty in integrating a strong motion record to velocity and displacement. This is not the case; in practice, the simple integration of an accelerogram produces unphysical velocity and displacement waveforms that grow unbounded as time progresses. Computation of broadband displacements from strong motion recordings is a thoroughly studied procedure that is fraught with many known problems and has no known single solution. We use the term *broadband displacement* in the same sense as Bock *et al.* [2011], i.e., a strong motion displacement waveform that captures both transient phenomena (waves) and permanent or static deformation, a recording reliable down to DC.

[4] The problems associated with the double integration of accelerometer recordings have been comprehensively studied, and many sources of error have been suggested: numerical error in the integration procedure, mechanical hysteresis, cross-axis sensitivity, and unresolved rotational motions [Graizer, 1979; Iwan *et al.*, 1985; Boore, 1999, 2001; Boore *et al.*, 2002; Smyth and Wu, 2007]. It is generally assumed that small offsets are introduced in the acceleration time series; upon integration, these *baseline offsets* produce the linear and quadratic trends observed in the velocity and displacement time series, respectively. Many possible sources have been invoked as the source of these offsets with unresolved rotational motion increasingly considered the main error [Graizer, 2006; Pillet and Virieux, 2007]. Motion is described by six degrees of freedom, three translations, and three rotations. Accelerometers are incapable of discerning between rotational and translational motions; thus, rotational

All supporting information may be found in the online version of this article.

¹Cecil H. & Ida M. Green Institute of Geophysics and Planetary Physics, Scripps Institution of Oceanography, La Jolla, California, USA.

²Jacobs School of Engineering, University of California San Diego, La Jolla, California, USA.

Corresponding author: D. Melgar, Cecil H. & Ida M. Green Institute of Geophysics and Planetary Physics, Scripps Institution of Oceanography, 9500 Gillman Dr., La Jolla, CA 92093, USA. (dmelgarm@ucsd.edu)

©2013. American Geophysical Union. All Rights Reserved.
2169-9313/13/10.1002/jgrb.50135

motions are recorded as spurious translations. Effectively, this results in a change of the baseline of the accelerometer, even if by a small amount, leading to unphysical drifts in the singly integrated velocity waveforms and doubly integrated displacement waveforms.

[5] Many correction schemes, collectively known as baseline corrections, have been proposed over time to deal with this problem. Rotational motions become more prevalent close to the source and at long periods [Graizer, 2006]; thus, the simplest baseline correction scheme is a high-pass filter [Boore and Bommer, 2005]. This leads to accurate recovery of the mid to high frequency part of the displacement record but suppresses completely long period information such as the static offset.

[6] To ameliorate this, a number of more elaborate correction schemes exist [Boore and Bommer, 2005] that rely on function fitting to the singly integrated velocity time series. The most reliable scheme that routinely produces plausible displacement waveforms (which include a measure of the static offset) is described in Boore [1999, 2001] and is a modification of the scheme proposed by Iwan *et al.* [1985] and henceforth referred to as the Boore-Iwan or BI correction scheme. In this method, a piecewise linear function is fit to the uncorrected velocity time series; the slope of each straight line segment represents an acceleration step which is then subtracted from the original acceleration data. This baseline-corrected acceleration record is subsequently integrated to velocity and displacement. If the intervals for fitting the linear functions to the velocity data are selected appropriately, this algorithm will produce waveforms that look plausible. They will contain both permanent and transient motions. The difficulty then lies in determining what these appropriate time intervals are from the data themselves. As discussed by Boore [1999, 2001], this is an ambiguous process. To diminish this uncertainty, subsequent research has focused on determining plausible times for the fits and then grid searching for waveforms that most resembles a ramp or step function [Wu and Wu, 2007; Chao *et al.*, 2010; Wang *et al.*, 2011].

[7] Indeed, there is ample interest in access to such real-time broadband displacements. Static offsets have been used for rapid source dimension computations [Perez-Campos *et al.*, 2013; Crowell *et al.*, 2013] and centroid moment tensor (CMT) determination of large sources [Melgar *et al.*, 2012; Crowell *et al.*, 2013]. It has been demonstrated as well that such algorithms can provide rapid descriptions of the source in a timelier and complete manner than traditional seismic data. Static slip inversions can also be computed rapidly with such data [Crowell *et al.*, 2012, 2013; Wright *et al.*, 2012] and ingested into rapid models of the ensuing tsunami [Ohta *et al.*, 2012]. Except for select locales like the U.S. and Japan, Global Positioning System (GPS) networks with enough density to adequately capture large events are not as prevalent as strong motion deployments; thus, extracting reliable long period information that includes the static field, from accelerometer data alone, would be very valuable. Furthermore, inversions with regional high-rate GPS data that model both the static and dynamic components of the waveform have been demonstrated for CMT determination [O'Toole *et al.*, 2012, 2013] and kinematic slip inversions [Ji *et al.*, 2004; Ammon *et al.*, 2011]. However, the low sampling rates of GPS (~1–5 Hz) and its high noise levels compared to seismic data make full

waveform inversions from GPS data alone suspect. Hence, such inversions are usually leveraged with teleseismic data. This is unsuitable for rapid analysis of sources at regional scales, and it makes access to broadband displacements, as would ostensibly be produced from baseline corrections, very appealing.

[8] In the following, we present an analysis of the current state of the art in automated correction of strong motion data. We compare the results of such automated corrections with those obtained from collocated GPS and strong motion instruments and analyze their time and frequency domain characteristics. Previous studies such as Boore *et al.* [2002] and Wang *et al.* [2011, 2013] assert that baseline correction is usually only of importance for the estimation of static offsets. It is also assumed that the measurement error incurred from baseline offsets is generally not of importance at frequencies of engineering interest. For most structures, this is considered to be the range of frequencies higher than 0.1 Hz but could be useful at longer periods for large structures such as long span bridges. In the following work, we will test these assumptions. We ascertain not only the effect that baseline corrections have on the waveforms but present as well a detailed analysis of the frequency characteristics of the results. We seek to evaluate the reliability of strong motion displacements at frequencies of both seismological and engineering interest.

2. Automatic Baseline Correction

[9] In order for a baseline correction to be considered as automatic, it must be objective. There must exist some rules or set of rules to select the baseline correction parameters from the data themselves without soliciting operator interaction. Two distinct methods have been applied in solving this problem; the first relies on accelerometer data alone and the second on external information from other geophysical sensors, mainly GPS.

2.1. Automated Corrections From Accelerometer Data Alone

[10] Figure 1 is a plot of what is traditionally observed from zeroth-order corrected (only the pre-event mean is removed from the record) integration of an accelerogram. A linear drift is apparent in the latter part of the velocity time series, and a parabolic drift in the displacement overshadowed any useful information. The BI correction scheme is as follows: consider time t_i as some initial time for correction and t_f as the time at which the record ends, a least squares straight line fit is obtained from time t_1 , which lies between t_i and t_f , to the end of the record as

$$v_f(t) = v_0 + a_f(t); t \in (t_1, t_f). \quad (1)$$

Subsequently, another straight line is fit from t_i to t_1 with the constraint that velocity be zero at the start of the record and that the final velocity after baseline correction averages zero. These constraints are satisfied if [Boore, 1999]

$$a_m = \frac{v_f(t_1)}{t_1 - t_i} \quad (2)$$

Then, the acceleration baseline, a_m , is subtracted from the uncorrected record from times t_i to t_1 and the baseline, a_f ,

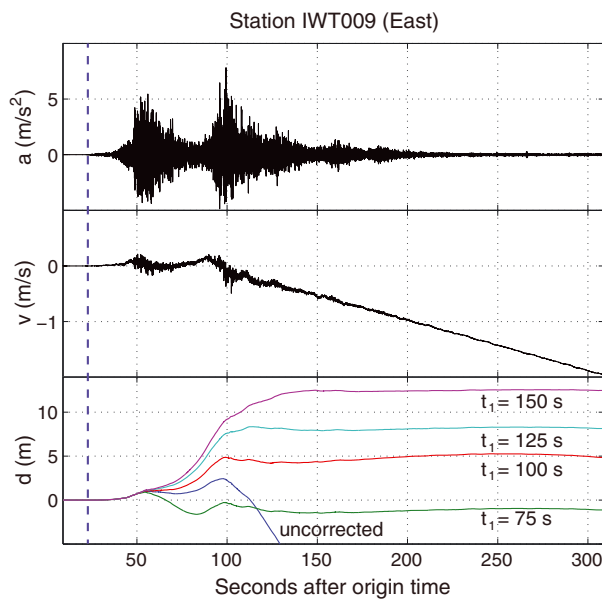


Figure 1. Example of integration of an uncorrected accelerogram as well as the result of applying the BI correction scheme. The initial correction time (t_i) is held fixed at the P-wave arrival, and the intermediate correction time (t_1) is allowed to vary from 75 to 150 s. The dashed blue line indicates the P-wave arrival.

is subtracted from the uncorrected record for times t_i to t_f . The record is then integrated to velocity and displacement. This scheme produces waveforms that look plausible; they contain both transient and permanent motions. However, an ambiguity lies in the selection of times t_i and t_f . As has been amply discussed by Boore [1999, 2001], this ambiguity is not easily resolved without external information, and each investigator relies on subjective judgment to ascertain what looks best. Figure 1 illustrates such an example where the same waveform has been baseline corrected for several values of t_1 while holding t_i fixed at the P-wave arrival with results that vary wildly. If the waveform is complex, as in this example which has two distinct pulses of very strong shaking, then more baselines might need to be subtracted. However, if there is already ample ambiguity in the simple determination of the two baselines a_m and a_f , the problem is compounded with the introduction of more baselines.

[11] The correction times t_i and t_f can and do vary for each station-event pair and even for different channels at the same station during the same event. Practically, this means that analysis of broadband displacements from baseline-corrected accelerometer records is inherently ambiguous and complicated for real-time seismological applications or across large networks both for real-time and post-processing purposes.

[12] Research into automated baseline correction from accelerometer data alone has focused on variations of the BI scheme. Wu and Wu [2007] proposed a variant in which the times for each linear segment of the baseline correction are determined by a grid search such that the resulting time series best matches a ramp function. Chao et al. [2010] elaborated on the formulation of Wu and Wu [2007] by adding an extra restriction that the times be selected after certain threshold values of acceleration energy have been accrued. Both studies compare their results to static offsets

determined from GPS and find that their estimates are somewhat similar. They still maintain some significant differences though, and no analysis on the adequacy of the remaining part of the waveform is performed. It is implicitly assumed that if the static field is well fit, then the rest of the waveform will be reliable as well.

[13] An important advance in automatic baseline correction is presented in Wang et al. [2011] who present another variation of the BI bilinear scheme that performs better than those discussed thus far. They develop some simple rules for determination of the interval of possible correction times based on analysis of the uncorrected acceleration and displacement waveforms. From an analysis of the time at which the peak ground acceleration occurs and the time of last zero crossing in the uncorrected displacement, they determine bounds for the grid search of baseline correction times. They then perform the grid search among these possible correction times and fit, via a non-linear regression, a step function to all possible waveforms. An optimum correction (the one that best fits this step function) is then selected. Unlike previous studies, they compare their results not only to measured static offsets but to observed 1 Hz GPS data; they do this for a single station. They find for that one station an error of ~20% in the static field estimation but a very good agreement between the corrected displacement and the GPS for the first 200s of the waveform.

[14] In a follow-up study, Wang et al. [2013] apply their methodology to accelerometer records for the 2011 Mw 9.0 Tohoku-oki earthquake. They obtain reasonable estimates of the static field for many stations in the KiK-net network but also find numerous outliers. They then develop a simplified scheme to screen the outliers by excluding coseismic offsets that deviate more than 15° from the predictions of a static slip inversion. Furthermore, they compare their automatic corrections for selected borehole sensors in the KiK-net network with nearby high-rate GPS stations with mixed results. They find that while parts of the waveforms might be a good match, the static estimates can be in error by a significant amount. To ameliorate this, they then propose to use the static field from nearby GPS stations as a constraint in the correction procedure. From the pool of all candidate baseline corrections, they select the one that fits a step function of amplitude given by the static field. It is noteworthy that Wang et al. [2013] do not provide baseline-corrected solutions for the sister strong motion network K-net. They readily acknowledge that K-net stations, which generally have less favorable site responses [Tsuda et al., 2006], are not well modeled by this automatic approach.

[15] In the following pages, we will use the methodology of Wang et al. [2011] for our automatic baseline corrections (ABC) from accelerometer data alone as we consider it to be the most robust one to date. We add one important modification that we have found improves the performance of the algorithm. To discriminate which baseline correction best fits the step function that was fit from non-linear regression, we employ the L_1 -norm as opposed to the traditional L_2 -norm.

2.2. Multi-sensor Platforms

[16] It is increasingly obvious that while improvements can be made in the recovery of broadband displacements, devising a methodology that can function well under any conditions using the accelerometer data alone is perhaps

not realistic. It would be desirable to have strong motion rotational sensors to quantify the contribution of the three rotational degrees of freedom and thus completely recover ground motion [Trifunac and Todorovska, 2001]; however, such sensors are not yet available.

[17] Thus far, the most reliable external source of information has been GPS. Verbosity of the raw GPS data exerts pressure on telemetry requirements and limits the sampling rate of the instrument usually to 1–5 Hz, although recordings as high as 50 Hz have been made [Genrich and Bock, 2006; Bock et al., 2011] and higher sampling observations are becoming more prevalent [Avallone et al., 2011, 2012]. This slow sample rate imposes a high frequency limit on the displacements that can be resolved and introduces aliasing effects [Smalley, 2009]. Additionally, noise levels are of the order of ~1–2 cm in the horizontal direction and ~5–10 cm in the vertical direction [Genrich and Bock, 2006; Bock et al., 2011; Wright et al., 2012] which are sensibly higher than what can be obtained with adequately corrected accelerometer data. Furthermore, GPS measures displacements in an absolute reference frame and does not suffer from baseline offsets. Thus, both data sets are complementary and alleviate each other's weaknesses. Nevertheless, given their different site selection requirements, both sets of networks have developed independently and are seldom collocated.

[18] Nikolaidis et al. [2001] recognized the utility of using collocated 30 s sampled GPS as a long period constraint in the recovery of ground motion, and indeed, the approach taken by Wang et al. [2013] in using the static field constraint is a refinement of this initial effort. Emore et al. [2007] demonstrated with data from the 2003 Mw 8.3 Tokachi-oki earthquake that collocated 1 Hz GPS, and 100–200 Hz accelerometer data could effectively be combined by solving an inverse problem. They used the central difference approximation to compute the derivatives of the GPS time series to solve for a series of acceleration steps that minimize the difference between GPS and accelerometer data to within some pre-established tolerance. This is, again, a sophisticated version of the BI scheme with multiple baselines being fit. A similar approach is demonstrated in Wang et al. [2013] where a series of sinusoidal functions is subtracted from the acceleration time series until the difference between the GPS and accelerometer derived displacement agrees within some tolerance.

[19] However, for both real-time and post-processing operations, we prefer the integrated GPS+accelerometer Kalman filter approach demonstrated numerically by Smyth and Wu [2007], first implemented by Kogan et al. [2008] and refined and thoroughly studied by Bock et al. [2011]. The Kalman filter algorithm is attractive over the aforementioned methodologies because it operates on a sample-by-sample basis (although introduction of a Kalman smoother induces a lag [Bock et al., 2011]) such that one need not wait for the full waveform to accrue before processing the data. Furthermore, it does not require one to assume any mechanism by which baseline offsets are introduced, as in the previously cited works. Rather, the Kalman filter simply estimates velocity and displacement from the time series themselves based on the statistical properties of the pre-event noise and the simple assumption that between accelerometer samples, the acceleration is constant. When applying any Kalman filter, the selection of the weights between data types

(Kalman gain) is critical. A high weight on the physics and inputs (the strong motion data) will yield data that resemble the integrated accelerations. Too high a weight on the observations (GPS positions) will produce waveforms that artificially resemble the GPS displacements. Bock et al. [2011] demonstrated that an objective and automatic way of determining these weights is by computing the variance from windows of pre-event noise, and indeed, this is the approach we take here. We then employ a system variance multiplier $Kq = 100$ for all stations and all channels. This value was demonstrated as optimum in Bock et al. [2011] and has been routinely employed in our data processing, i.e., Crowell et al. [2013] and Geng et al. [2013]. Shaketable testing by Bock et al. [2011] at the University of California San Diego's Large High Performance Outdoor Shaketable (LHPOST) compared the output of the filter, with the Kalman gain determined from pre-event noise and with the variance multiplier, with the deterministic input into the LHPOST for four earthquake simulations. It was determined that the RMS agreement between filter and input was of millimeter level. The frequency domain performance of the Kalman filter was also tested during the LHPOST experiment. It was shown that the filter was capable of reproducing the frequency content from DC to the accelerometer Nyquist frequency.

[20] The Kalman filter approach has been shown to improve static offset estimates and yield more reliable source models [Crowell et al., 2013]. Additionally, waveforms from the Kalman filter have enough dynamic range to measure small amplitude body waves for large [Bock et al., 2011; Crowell et al., 2013] and small events [Geng et al., 2013] at local and regional distances. Crowell et al. [2013] utilizes the versatility of these waveforms and demonstrates their potential utility for earthquake early warning using scaling of the peak amplitude of the P wave.

[21] Thus, in the aggregate of previous studies, it has been recognized that no one instrument is capable of discerning the full spectrum of strong ground motions and that multi-instrument platforms are necessary for adequate observation. In the following, we will employ Kalman filtered displacements (KD) obtained from collocated GPS and accelerometers.

3. Comparison Between ABC and KD Waveforms

3.1. Method

[22] One of the most relevant data sets for strong motions seismology in recent times is the set of records for the 2011 Mw 9.0 Tohoku-oki earthquake. For this event, we process data from 816 GPS stations throughout Honshu and Hokkaido Islands (Figure 2). We process the GPS data in a simulated real-time mode using the method of instantaneous positioning [Bock et al., 2000, 2011]. We triangulate the GPS network using a Delaunay triangulation scheme, and individual triangles are processed independently for relative station positions. The triangles are combined through a real-time network adjustment Crowell et al. [2009], and station positions are referenced to station 0848 on Hokkaido Island, 900 km northwest of the hypocenter (Figure 2). We identified 139 collocated GPS/accelerometer station throughout Japan. We define a collocation as a pair of instruments separated by 1500 m or less. Accelerometers in Japan function in triggered mode and many of those stations did not trigger in time.

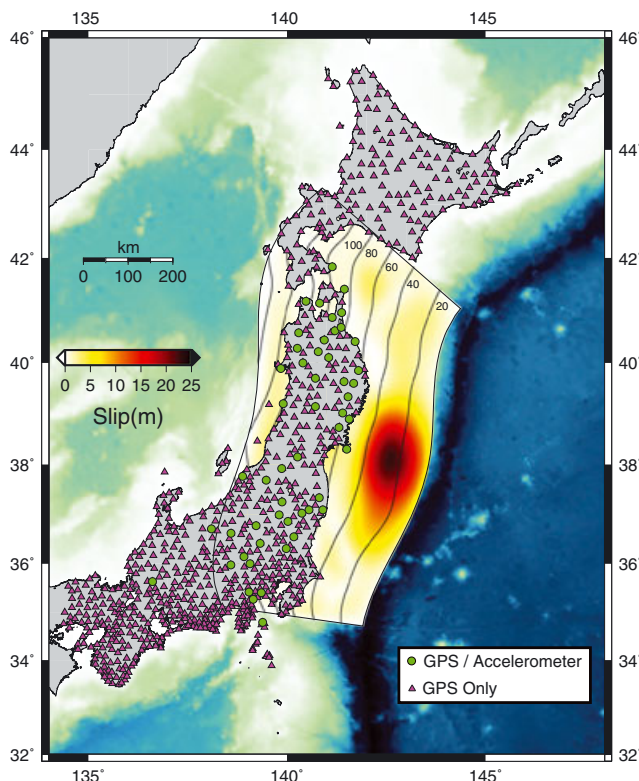


Figure 2. Station map depicting GPS only and GPS/accelerometer stations. Also included is the static slip inversion computed from real-time GPS by *Crowell et al.* [2013] and the iso-depth contours of the slab geometry determined in *Hayes et al.* [2012].

Furthermore, given the large duration of the source, many of those instruments terminated recording while significant shaking was ongoing. Thus, we culled the data set to 48 stations (Figure 2) that contain at least 2 s of pre-event noise and enough record to discern static offsets in the corrected data (a table with station locations and accelerometer to GPS separations is in Supplement S1 in the auxiliary material). Of these 48 stations, 40 are K-net accelerometer sites and 8 are KiK-net accelerometer sites. An interactive map of Japanese GPS and accelerometer stations as well as coseismic displacements is available at the GPS Explorer data portal (<http://geoapp03.ucsd.edu/gridsphere/gridsphere?cid=Honshu>).

[23] We processed the three channel data for those 48 stations using an in-house program, *baselineK*, that replicates the method of *Wang et al.* [2011] and which we make freely available. After removing the pre-event mean of the acceleration data, they are ingested into the program which itself requires no operator interaction, simulating an actual real-time scenario.

[24] Our correction code also allows one to constrain the solution with some pre-established value of the static field [*Wang et al.*, 2013]. Again, simulating a real-time scenario, we extracted the coseismic offset from the 1 Hz GPS data at these collocated stations using the trailing variance method described in *Melgar et al.* [2012] and employed these values as constraints in the correction. We also processed the data according to the Kalman filter method (codes *kalmand* and *kalmans*, also freely available) from

Bock et al. [2011] with the additional modification described in *Crowell et al.* [2013] and *Geng et al.* [2013] of an adaptive system variance.

[25] For all three data sets, the automatic corrections, the static field constrained corrections, and the Kalman filter, we computed the power spectral density using the multitaper method [*Percival and Walden*, 1993]. We also computed 5% damped displacement response spectra by a finite difference solution to the single degree of freedom oscillator; that code (*respspec*) is provided as well.

3.2. Time Domain Analysis

[26] Figure 3 illustrates the results of ABC waveforms for two stations where no static field constraint has been applied. K-net station AOM011 is 296 km from the event centroid and is an example of a good correction by the automated procedure. When compared to the KD waveform, we see that the static offsets roughly match as does the shaking for all three components. In contrast, K-net station GNM004, 430 km from the centroid, is an example of a failed correction. The displacement converges to the wrong value for the static field, and as a result, the shaking is also miscomputed. These two stations are the extremes of the behavior that we observe at the remaining stations. The plots for all stations are in Supplement S2.

[27] As mentioned before, we can apply the constraint that the ABC waveform converges to the static offset measured from a nearby GPS station. Figure 4 compares the result of applying such a constraint for station K-net station AOM010, 323 km from the centroid. This station is representative of the breadth of behavior we observe at all stations (all waveforms with the static constraint are also in Supplement S2). The vertical channel fit to the KD waveform is greatly improved by adding the constraint. The east component is mildly improved but does not reach the level defined by the KD

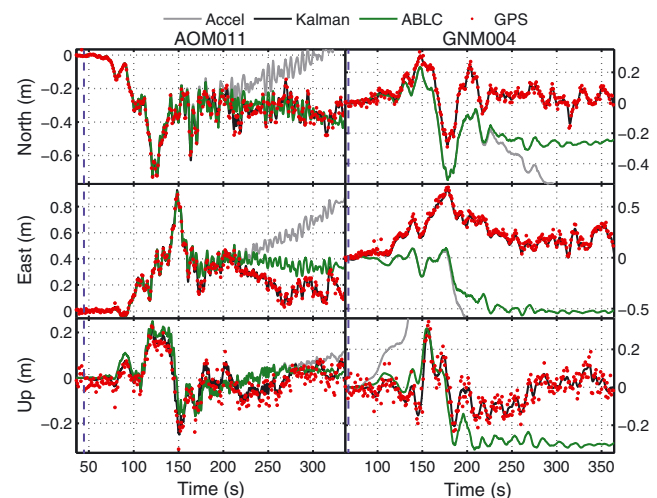


Figure 3. Results of applying the automatic baseline correction of *Wang et al.* [2011] with no static field constraints. The comparison is made to the Kalman filter, uncorrected accelerometer integration, and 1 Hz GPS. K-net station AOM011 is 296 km from the centroid and 897 m from GPS station 539, and K-net station GNM004 is 430 km from the centroid and 103 m from GPS station 591. The dashed blue line indicates the P-wave arrival.

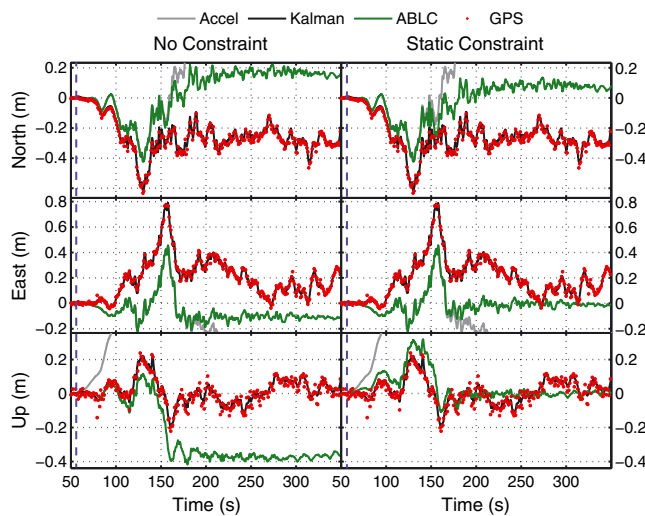


Figure 4. Comparison between the ABC waveforms with and without the static field constraint. The results are for K-net station AOM010 323 km from the centroid and 1360 m from GPS station 537 and compared to the Kalman filter, uncorrected accelerometer integration, and 1 Hz GPS. The dashed blue line indicates the P-wave arrival.

waveform. The improvement in the north component is only marginal, and a large difference (~ 20 cm) remains in the static field between ABC and KD waveforms. Recall that we determine the optimum correction to be the one that best fits a step function of amplitude given by the static field. What these results indicate is that there might not be a combination of baseline correction times that provides a sufficient improvement to the displacement computation.

[28] For both the constrained and unconstrained solutions, there are records that compare favorably with the KD waveforms (Figures 3 and 4; Supplement S2). However, given the large amplitude motions present at most of these stations, it is difficult to judge by simple visual inspection the adequacy of this comparison. Figure 5 presents the records ordered by distance to the centroid with the difference between the KD waveforms and the ABC waveforms for all 48 stations. We find that the difference is smallest in the earlier parts of the waveform. Additionally, we see that the static field constraint improves the fit to the KD waveforms, but large differences of several tens of centimeters remain. Also, notice that the stations closest to the centroid are the ones with the poorest fits that even with the addition of the static field constraint improve little.

[29] Inspection of the KD waveforms (Figures 3 and 4; Supplement S2) shows that the Kalman filter method corrects the three components of motion at all 48 stations successfully. We decimate the KD and ABC static field constrained waveforms to the GPS sample times and compute the RMS of the difference between both sets of waveforms and the GPS (Figure 6). In the north direction, the RMS for the ABC waveforms ranges from 4.3 cm to 150 cm; for the KD waveforms, the range is from 0.6 cm to 5 cm. In the east direction, the ABC waveforms have RMS values from 6 cm to 320 cm, and the KD waveforms, from 1 cm to 10 cm. In the vertical direction, the ABC waveforms have RMS values that range from 4 cm to 43 cm, and the KD waveforms, from 2 cm to 10 cm. For all stations on all channels, we

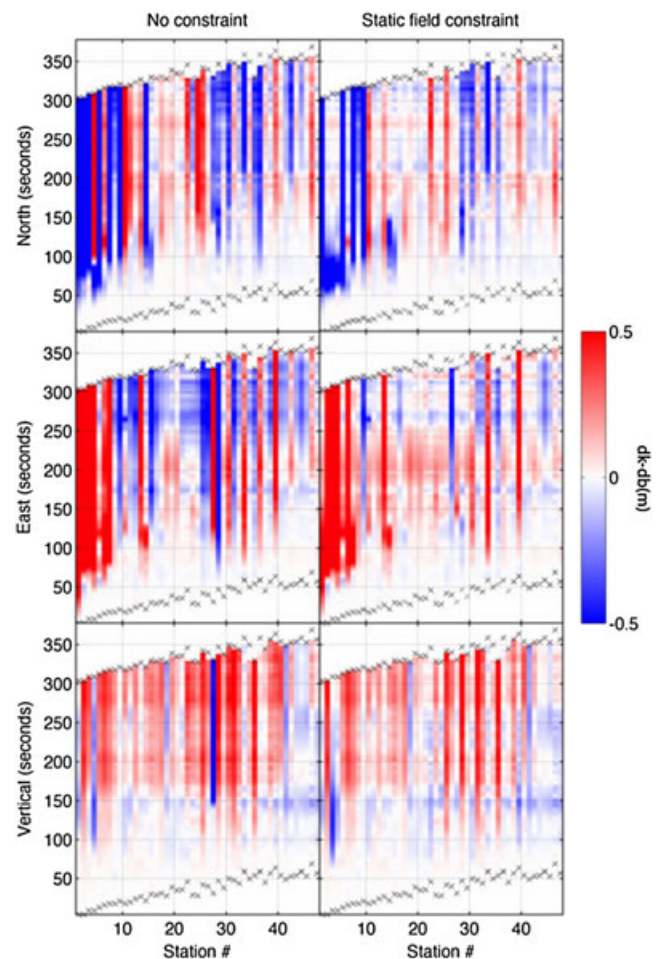


Figure 5. Difference between the Kalman filtered displacements (dk) and automatic baseline corrected (db) waveforms for both constrained and unconstrained solutions. The stations are ordered by increasing distance to the centroid, and the crosses denote the start and end of each waveform.

find that the KD waveforms have a smaller RMS than the ABC waveforms.

3.3. Frequency Domain Analysis

[30] It remains unclear in the literature at which frequencies baseline corrections introduce errors into measurements. This is of consequence for source observations and strong motion analysis. Also, it is generally assumed that baseline corrections have no effect at frequencies of engineering interest. In the following is a frequency domain analysis where we compare ABC and KD waveforms.

[31] Figure 7 is an example of the power spectral density (PSD) for ABC, KD, uncorrected accelerometer integration, and GPS data for the north component of K-net station AOM027, 363 km from the centroid. These results are for the unconstrained baseline correction. The results are consistent with Bock *et al.* [2011], namely that the accelerometer overestimates long period content and that GPS overestimates the spectral content at the higher end of its frequency band. Furthermore, one can observe the desirable characteristics of the KD waveforms; they follow the GPS spectra at low frequencies and transition to the accelerometer at high frequencies. It is also interesting to note that the ABC

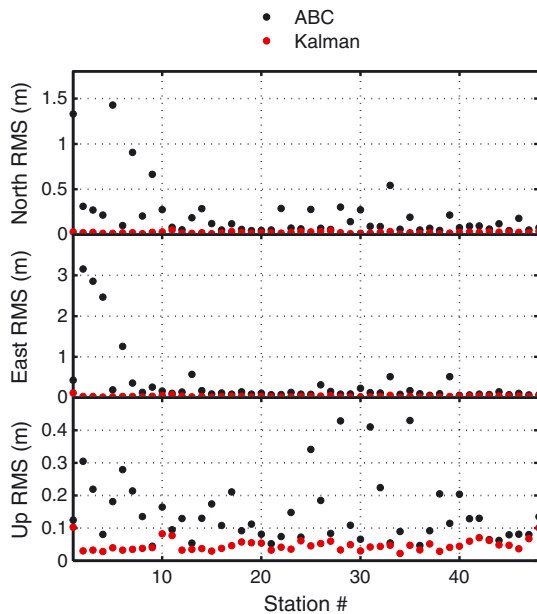


Figure 6. RMS of the difference between the ABC with static field constraint and GPS waveforms and the KD and GPS waveforms for all stations. The stations are ordered by increasing distance from the centroid.

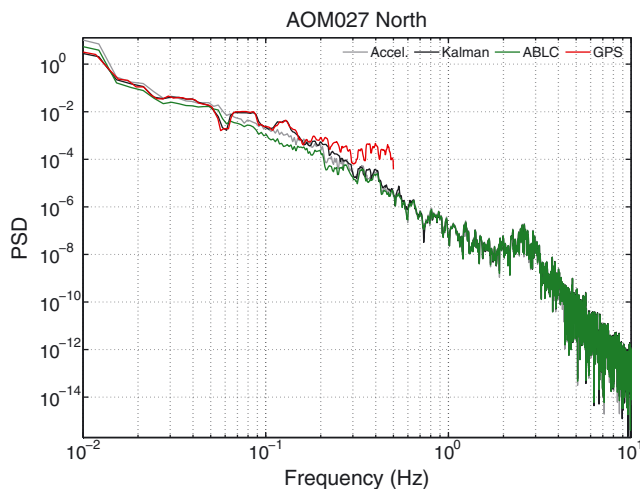


Figure 7. Power spectral densities for the north displacements at K-net station AOM027, 363 km from the centroid and 28 m from GPS station 535. These results are for the unconstrained baseline correction.

waveform has different frequency contents than the KD waveforms at frequencies at least as high as 0.5 Hz and that the ABC waveform does not agree with the GPS at any part of the long period spectrum. The spectra for all waveforms unconstrained and constrained are in Supplement S3.

[32] Figure 8 is a similar analysis for the displacement response spectra (spectra for all stations is in Supplement S4). The results are for the east component of K-net station MYG001 138 km from the centroid. The ABC solution is unconstrained. In line with the PSD analysis, we find that the KD waveform tracks the GPS result at long periods and

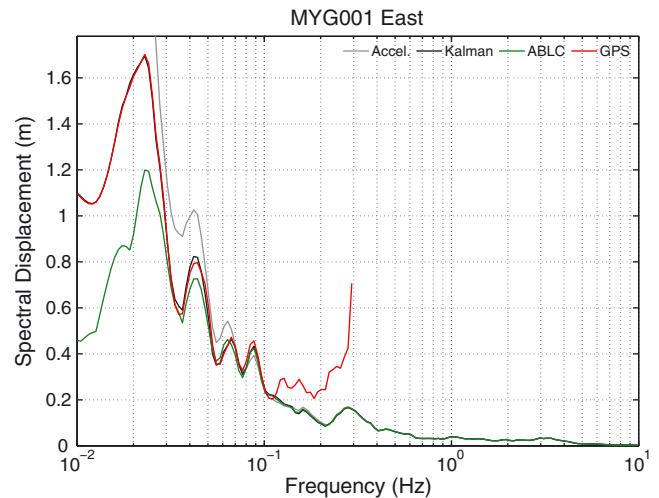


Figure 8. 5% damped displacement response spectra for the east component of K-net station MYG001 138 km from the centroid and 393 m from GPS station 172. These results are for the unconstrained baseline correction.

the uncorrected accelerometer at high frequencies. Similarly, we find that the uncorrected accelerometer spectrum differs from the KD result at periods longer than ~ 11 s while the ABC spectrum begins to diverge from the KD spectrum at around ~ 13 s. It is also important to note that for periods shorter than 10 s, the GPS only spectrum is unreliable.

[33] To better understand the relationship between the ABC and KD PSDs, we computed the ratio between the KD PSDs and ABC PSDs for all stations. This is shown in logarithmic scale in Figure 9 for both the constrained and unconstrained corrections and ordered by distance to the centroid. Several features are of interest, first that the discrepancy between the KD and ABC spectra is larger for stations close to the source. At long periods, there is mixed behavior between over- and underestimation of the spectra by the ABC results in the horizontal components. However, for the vertical component, we find that the ABC spectra consistently overestimate the frequency content. It is also important to note that the differences in spectral content are of consequence for frequencies as high as 0.5 Hz. We also find that addition of the static field constraint provides only a marginal improvement in the estimation at long periods. Addition of this constraint leads to little to no improvement in the recovery of the mid-range frequencies.

[34] Figure 10 is a similar plot for the response spectra; here we present the difference between the KD and ABC derived spectral displacements for both constrained and unconstrained results. Similar to what we described for the PSD comparison, we find a better fit by the ABC spectra in the vertical direction than in the horizontal directions. Additionally, we find little improvement in the spectral displacement estimates from addition of the static field constraint. We also find that for the response spectra, the largest differences occur at periods longer than 8–9 s.

4. Discussion

[35] The implications of the results shown here are far reaching. Even though GPS has higher noise levels and

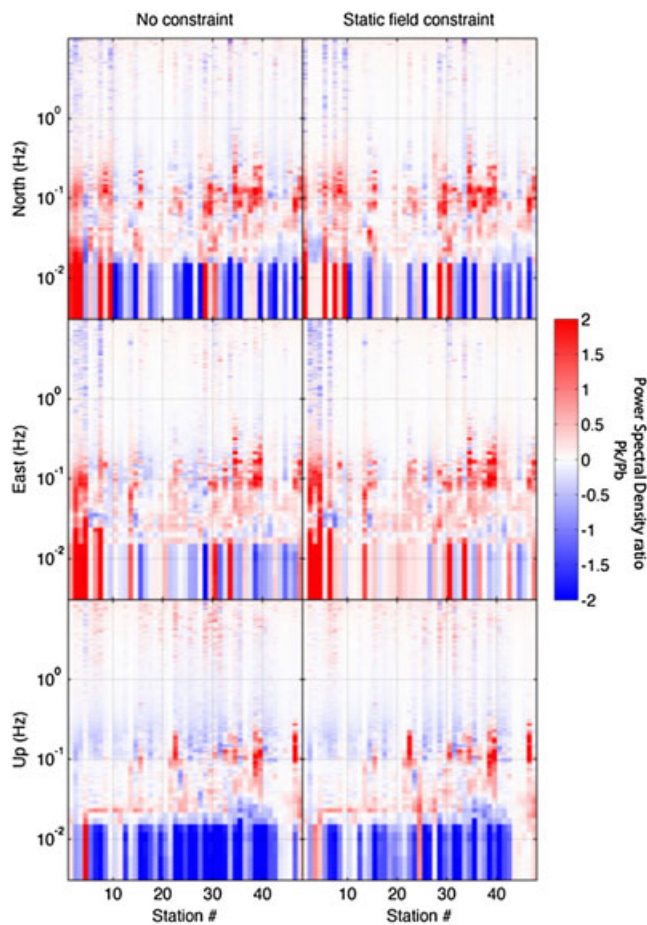


Figure 9. Logarithm of the ratio between the Kalman filtered (Pk) and automatic baseline corrected (Pb) power spectra for all stations ordered by increasing distance to the centroid. Results are for both constrained and unconstrained waveforms.

aliasing effects are present, whatever method is chosen to compute the displacements (KD or ABC), it must agree broadly with the GPS time series. Figures 3–6 and Supplement S2 demonstrate this and show that with the Kalman filter, we achieve accurate recovery of both permanent and transient motions while serious difficulties arise when employing the ABC method. We assert that this verifies that the Kalman filter methodology described in *Bock et al.* [2011] has no substantive problems in resolving broadband displacements for the Mw 9.0 Tohoku-oki earthquake. That we have done so using the K-net acceleration data is important. As acknowledged by *Wang et al.* [2013], recovery of displacements from those records is difficult because of the large baseline offsets, ostensibly introduced because of unfavorable site conditions. The incorporation of real-time GPS with the Kalman filter resolves these offsets with little difficulty.

[36] However, we recognize that collocation of GPS and strong motion sensors is still the exception rather than the norm, and baseline-corrected displacements from accelerometer data alone are important. We have evaluated the suitability of automated baseline correction procedures in both the time and frequency domain. We independently verified that

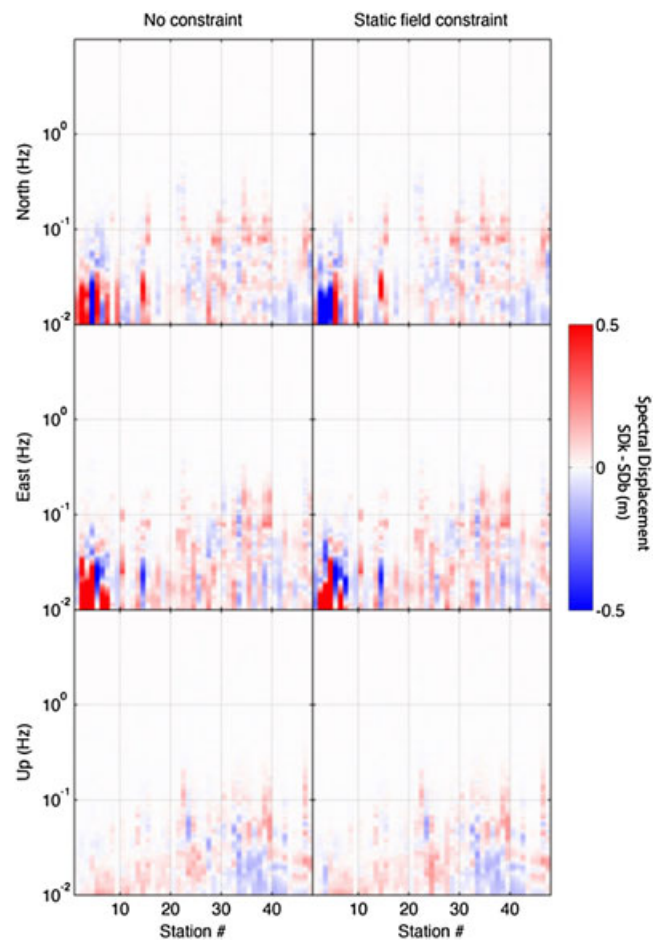


Figure 10. Difference between the Kalman filtered (SDk) and automatic baseline corrected (SDb) 5% damped response spectra for all stations ordered by increasing distance to the centroid. Results are for both constrained and unconstrained waveforms.

automatic baseline correction schemes such as that of *Wang et al.* [2011] can indeed produce in some cases waveforms that accurately capture both static and dynamic motions. However, for the subset of 48 stations used in this research, we found that there can be very large differences in the estimated static offsets when compared to actual offsets measured by the real-time GPS (Figure 11). This is important because as discussed in Section 1, one of the most fertile current areas of research for rapid earthquake response is modeling of the source with static offsets. If the offsets determined from the ABC schemes are unreliable, their utility for rapid response will be very limited. *Wang et al.* [2013] proposed detecting outliers in the static field estimation by comparison with synthetics determined from a static slip inversion. This necessitates some important assumptions, for example, for the 2011 Tohoku-oki event, that the earthquake has ruptured the mega-thrust and that the mechanism is reverse faulting. These assumptions are critical and will not always hold up. Outer rise normal faulting events can be sizeable and are not uncommon. Large strike-slip events close to the trench such as the Mw 8.6 strike-slip event off Sumatra, Indonesia, on 11 April 2012 [*Satriano et al.*, 2012] will be significantly mis-modeled.

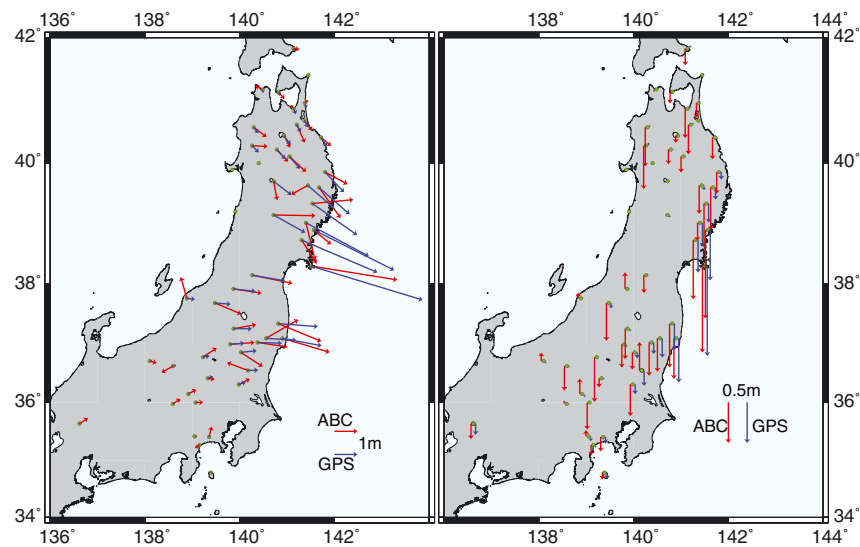


Figure 11. Comparison between static offsets determined from the automated baseline corrected procedure (ABC) and the real-time GPS.

Furthermore, quiescent borehole stations such as those of the KiK-net deployment used in *Wang et al.* [2013] are not the norm. Rather, stations of heterogeneous site response like those of the K-net deployment are more common.

[37] We have identified important discrepancies between ABC and KD waveforms. The automatic corrections can converge to solutions that look plausible but as shown in Figure 3 can be very far from the correct solution. Even with addition of the static field constraint, the baseline offsets might be so large that the simple bilinear scheme is not enough to correct the solution. But the most troubling case is shown in Figure 12. This presents the north component of K-net station IWT009 157 km from the centroid and corrected with the static offset constraint. The plot shows that the ABC waveform converges to the correct static value, but the shaking is miscomputed by as much as 0.5 m. This type of behavior is not uncommon in the corrections (Figure 5, Supplement S2). Thus, the assumption that convergence to the static field value ensures reliability of

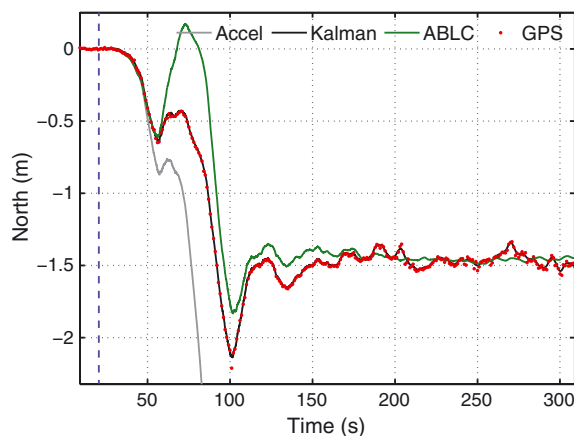


Figure 12. North displacement waveforms for station K-net station IWT009 157 km from the centroid and 758 m from GPS station 911. The dashed blue line indicates the P-wave arrival.

the waveform is not routinely satisfied. It is possible for a correct static field to be computed but to be in error by large amounts in the dynamic component of the waveform.

[38] We too find that addition of the static field constrained as proposed by *Wang et al.* [2013] improves the time domain solution. Nonetheless, we have also shown that ABC waveforms even when they converge to the correct static offset value can still be in error in regard to the transient component of the waveform (Figure 12, Supplement S2).

[39] The frequency domain analysis also shows some interesting behavior. We have once again demonstrated the performance of the Kalman filter displacements, following the GPS spectra at long periods and the accelerometer spectra at higher frequencies. We have also shown that the ABC PSD can be miscomputed at frequencies as high as ~ 0.5 Hz. Furthermore, we have demonstrated that addition of the static field constraint does little to improve the spectral recovery.

[40] With the response spectra, we have found a similar pattern except that the ABC spectra are in error at lower frequencies, usually at periods longer than 8–9 s. These discrepancies in both PSD and response spectra are well within the frequencies that are of interest to engineering seismology, especially for the PSD estimation. Furthermore, the frequency domain error incurred by the ABC waveforms is not simple. Sometimes it is an overestimate and sometimes an underestimate. This has relevance for seismological applications that seek to use baseline-corrected waveforms for strong ground motion analysis, source studies, and response. This is true as well for engineering purposes. *Boore et al.* [2002] had already shown that this might be the case but found that for most stations, the bias was introduced at much longer periods. However, that conclusion was reached from accelerometer data alone, and no comparison was made to the frequency content of high-rate GPS or combined data, which was unavailable at the time.

[41] Tangentially, we have demonstrated that the useful frequency content of GPS only waveforms is also limited to periods longer than ~ 10 s. Higher-frequency studies from

high-rate data will be impeded by overestimates of the spectral content. This was observed previously in [Bock *et al.*, 2011] and can be in part resolved by faster sampling by the receiver [Genrich and Bock, 2006; Smalley, 2009]. The performance of GPS positions can potentially improve with better GPS processing techniques [Geng *et al.*, 2013] and noise reducing spatial and temporal error filters [Hung and Rau, 2013], although applying them in real time can be cumbersome. However, Geng *et al.* [2013] found that the remaining errors in KD solutions for small earthquakes can not easily be associated with traditional GPS error sources such as multipath.

[42] Assuming that baseline offsets once introduced remain constant throughout the record seems to be insufficient to explain the scope of behavior in the observed baseline-corrected waveforms. If this were the case, BI-type correction algorithms would always find a suitable set of correction times to remove the undesirable errors. Rather, it seems that baseline offsets are introduced in a time-varying fashion. There is increasing evidence that rotational motions account for most of the baseline offsets [Graizer, 2006; Pillet and Virieux, 2007]; thus, aspiring to design an algorithm that can accurately and objectively correct acceleration from acceleration data alone is a frustrating proposition. There is no way to simply distinguish between rotations and translations with enough sensitivity from the inertial sensor itself. Furthermore, although the incorporation of outside constraints, such as the static field, improves matters, it does not solve the problem. For a simple waveform with short duration and small accelerations, incorporating the static constraint might suffice. However, for the Tohoku-oki waveforms, which have complex shapes, this simple constraint is insufficient. Rotational motions happen continuously throughout the record and introduce baseline offsets of different amplitudes in a continuous fashion [Graizer, 2006], not in the simple piecewise manner in which we have traditionally tried to model them. Thus, constraining the end of the record is not enough, and there is no way to objectively determine offsets that happen earlier in the record

[43] The aggregate of results and analysis we show here is by no means proof by exhaustion of the unsuitability of automatic baseline corrections that rely on accelerometer data alone. Other schemes might yet be developed that provide better results. However, we have shown that the corrections in this particular case are suspect for static field determination: that they might miscompute the dynamic component of the seismogram and that the errors are incurred at frequencies well within what is of interest to seismology and engineering. Furthermore, we have shown that broadband displacements computed from collocated GPS and accelerometers via a Kalman filter have none of those problems.

[44] Collocations are still not the norm, and deploying or upgrading a network to collocated status can be a costly proposition. One possible solution is discussed in Bock *et al.* [2011] and Wang *et al.* [2013] where it was demonstrated that low-cost micro electro-mechanical sensors (MEMS) are increasing in sensitivity to the point where they are suitable for regional monitoring. MEMS accelerometers are a fraction of the cost of observatory grade accelerometers. Further limitations arise because GPS requires good sky visibility, limiting deployments inside structures and curbing the scope of engineering applications of the GPS/

accelerometer combination. However, given the current limitations of automated baseline correction procedures that rely on inertial sensors alone, as outlined in this paper, the superior performance of collocated processing techniques, and the availability of cheaper MEMS sensors, we advocate that if broadband displacements are a priority target of a particular network, then retrofitting of existing geodetic infrastructure with accelerometers should be given a high priority. We find it unlikely that displacements from accelerometer data alone will reach the robustness or reliability required for real-time and rapid observations.

5. Conclusions

[45] We have shown that automatic baseline correction schemes from accelerometer data alone can, for selected stations, provide useful displacement information. However, the static offsets determined by such automated schemes are often unreliable. Furthermore, adding a static field constraint to the baseline correction procedure does not guarantee convergence to the correct value and even if the correct static offset is obtained; other parts of the waveform might be in error by large amounts. The objective estimation of baseline-corrected displacements, even when incorporating reliable and independent information such as the static field estimated from GPS, remains elusive. We have demonstrated that such a scheme, which eliminates the need for subjective judgment, is still fraught with problems. We tested the assumption that baseline corrections introduce errors only at long periods outside what is of engineering interest and find it does not hold for this particular data set. Analysis of the power spectral densities and displacement response spectra demonstrates that the frequency domain error incurred by the baseline correction is well within the frequency range of interest to both seismology and engineering. The Kalman filter method, which combines high-rate GPS and accelerometer data, is not affected by these problems. We demonstrated that even for the K-net data set which displays behavior associated with large baseline offsets, the Kalman filter solution routinely produces reliable broadband displacements.

[46] **Acknowledgments.** We would like to thank the National Research Institute for Earth Science and Disaster Prevention for access to K-NET and KiK-net data and the Geospatial Information Authority of Japan for GEONET GPS data. This paper was funded by NASA AIST-11 grant NNX09A167G, NASA ROSES grant no. NNX12AK24G, and the first author's Earth and Space Science Fellowship (NASA grant no. NNX12AN55H). We would like to acknowledge Thomas O'Toole, Rongjiang Wang, and Masumi Yamada for their helpful comments and discussion. Reviews by Bob Smalley and an anonymous reviewer were very valuable in improving the content and clarity of the manuscript. The codes used in this work can be downloaded from <http://igpppublic.ucsd.edu/~dmelgarm/Code>, and the processed data are available upon request.

References

- Ammon, C. J., T. Lay, H. Kanamori, and M. Cleveland (2011), A rupture model of the 2011 off the Pacific coast of Tohoku earthquake, *Earth Planets Space*, 63, 693–696, doi:10.5047/eps.2011.05.015.
- Avallone, A., et al. (2011), Very high rate (10 Hz) GPS seismology for moderate-magnitude earthquakes: The case of the Mw 6.3 L'Aquila (central Italy) event, *J. Geophys. Res.*, 116, B02305, doi:10.1029/2010JB007834.
- Avallone, A., et al. (2012), High-rate (1 Hz to 20 Hz) GPS coseismic dynamic displacements carried out during the Emilia 2012 seismic sequence, *Annals Geophys.*, 55(4), 773–779, doi: 10.4401/ag-6162.

- Bock, Y., R. Nikolaidis, P. J. de Jonge, and M. Bevis (2000), Instantaneous geodetic positioning at medium distances with the Global Positioning System, *J. Geophys. Res.*, **105**, 28,233–28,253.
- Bock, Y., D. Melgar, and B. W. Crowell (2011), Real-time strong-motion broadband displacements from collocated GPS and accelerometers, *Bull. Seism. Soc. Am.*, **101**(6), 2904–2925, doi:10.1785/0120110007.
- Boore, D. M. (1999), Effect of baseline corrections on response spectra for two recordings of the 1999 Chi-Chi, Taiwan, earthquake, *U.S. Geol. Surv. Open-File Rept.* 99–545, 37.
- Boore, D. M. (2001), Effect of baseline corrections on displacements and response spectra for several recordings of the 1999 Chi-Chi, Taiwan, earthquake, *Bull. Seism. Soc. Am.*, **91**, 1199–1211, doi:10.1785/0120000703.
- Boore, D. M., and J. J. Bommer (2005), Processing of strong-motion accelerograms: Needs, options and consequences, *Soil Dyn. & Earth. Eng.*, **25**, 93–115, doi:10.1016/j.soildyn.2004.10.007.
- 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. Seismol. Soc. Amer.*, **92**, 1543–1560, doi:10.1785/0120000926.
- Chao, W. A., Y. M. Wu, and L. Zhao (2010), An automatic scheme for baseline correction of strong motion records in coseismic deformation determination, *J. Seismol.*, **14**, 495–504, doi:10.1007/s10950-009-9178-7.
- Crowell, B. W., Y. Bock, and M. B. Squibb (2009), Demonstration of earthquake early warning using total displacement waveforms from real-time GPS networks, *Seismol. Res. Lett.*, **80**, 772–782, doi:10.1785/gssrl.80.5.772.
- Crowell, B. W., Y. Bock, and D. Melgar (2012), Real-time inversion of GPS data for finite fault modeling and rapid hazard assessment, *Geophys. Res. Lett.*, **39**, L09305, doi:10.1029/2012GL051318.
- Crowell, B. W., D. Melgar, Y. Bock, and J. S. Haase (2013), Earthquake early warning with seismogeodesy: Example of the 2011 Mw 9.0 Tohoku-oki earthquake, *Geophys. Res. Lett.*, submitted.
- Emore, G., J. Haase, K. Choi, K. M. Larson, and A. Yamagiwa (2007), Recovering seismic displacements through combined use of 1-Hz GPS and strong-motion accelerometers, *Bull. Seismol. Soc. Am.*, **97**, 357–378, doi:10.1785/0120060153.
- Geng, J., Y. Bock, D. Melgar, B. W. Crowell, and J. Haase (2013), A new seismogeodetic approach applied to GPS and accelerometer observations of the 2012 Brawley seismic swarm: Implications for earthquake early warning, *Geochem. Geophys. Geosys.*, in review.
- Genrich, J. F., and Y. Bock (2006), Instantaneous geodetic positioning with 10–50 Hz GPS measurements: Noise characteristics and implications for monitoring networks, *J. Geophys. Res.*, **111**, B03403, doi:10.1029/2005JB003617.
- Graizer, V. M. (1979), Determination of the true ground displacement by using strong motion records, *Izv. Phys. Solid Earth*, **15**, 875–885.
- Graizer, V. M. (2006), Tilts in strong ground motion, *Bull. Seism. Soc. Am.*, **96**(6), 2090–2102, doi:10.1785/0120060065.
- Havskov, J., and G. Alguacil (2002), *Instrumentation in Earthquake Seismology*, Springer 1st ed., 313 pp., Dordrecht, Netherlands.
- Hayes, G. P., D. J. Wald, and R. L. Johnson (2012), Slab1.0: A three-dimensional model of global subduction zone geometries, *J. Geophys. Res.*, **117**, B01302, doi:10.1029/2011JB008524.
- Hung, H. K., and R. J. Rau (2013), Surface waves of the 2011 Tohoku earthquake: Observations of Taiwan's dense high-rate GPS network, *J. Geophys. Res.*, doi:10.1029/2012JB009689, in press.
- Iwan, W. D., M. A. Mooser, and C. Y. Peng (1985), Some observations on strong motion earthquake measurements using a digital accelerometer, *Bull. Seism. Soc. Am.*, **75**, 1225–1246.
- Ji, C., K. M. Larson, Y. Tan, K. W. Hudnut, and K. Choi (2004), Slip history of the 2003 San Simeon earthquake constrained by combining 1-Hz GPS, strong motion, and teleseismic data, *Geophys. Res. Lett.*, **31**, L17608, doi:10.1029/2004GL020448.
- Kogan, M. G., W. Y. Kim, Y. Bock and A. W. Smyth (2008), Load response on a large suspension bridge during the NYC marathon revealed by GPS and accelerometers, *Seism. Res. Lett.*, **79**(1), 12–19, doi:10.1785/gssrl.79.1.12.
- Melgar, D., Y. Bock and B. W. Crowell (2012), Real-time centroid moment tensor determination for large earthquakes from local and regional displacement records, *Geophys. J. Int.*, **188**, 703–718, doi:10.1111/j.1365-246X.2011.05297.x.
- Nikolaidis, R., Y. Bock, P. J. de Jonge, P. Shearer, D. C. Agnew, and M. Van Domselaar (2001), Seismic wave observations with the Global Positioning System, *J. Geophys. Res.*, **106**, 21,897–21,916, doi:10.1029/2001JB000329.
- Ohta, Y., et al. (2012), Quasi real-time fault model estimation for near-field tsunami forecasting based on RTK-GPS analysis: Application to the 2011 Tohoku-oki earthquake (Mw 9.0), *J. Geophys. Res.*, **117**, B02311, doi:10.1029/2011JB008750.
- O'Toole, T. B., A. P. Valentine, and J. H. Woodhouse (2012), Centroid-moment tensor inversions using high-rate GPS waveforms, *Geophys. J. Int.*, **191**, 257–270, doi:10.1111/j.1365-246X.2012.05608.x.
- O'Toole, T. B., A. P. Valentine, and J. H. Woodhouse (2013), Earthquake source parameters from GPS-measured static displacements with potential for real-time application, *Geophys. Res. Lett.*, **40**, 60–65, doi:10.1029/2012GL054209.
- Percival, D. B., and A. T. Walden (1993), *Spectral Analysis for Physical Applications: Multitaper and Conventional Univariate Techniques*, 612 pp., Cambridge University Press, UK.
- Perez-Campos, X., D. Melgar, S. K. Singh, V. Cruz-Atienza, A. Iglesias, and V. Hjorleifsdottir (2013), Determination of tsunamigenic potential of a scenario earthquake in the Guerrero seismic gap along the Mexican subduction zone, *Seism. Res. Lett.*, doi:10.1785/0220120156.
- Pillet, R., and J. Virieux (2007), The effects of seismic rotations on inertial sensors, *Geophys. J. Int.*, **171**, 1314–1323, doi:10.1111/j.1365-246X.2007.03617.x.
- Satriano, C., E. Kiraly, P. Bernard, and J.-P. Vilotte (2012), The 2012 Mw 8.6 Sumatra earthquake: Evidence of westward sequential seismic ruptures associated to the reactivation of a N-S ocean fabric, *Geophys. Res. Lett.*, **39**, L15302, doi:10.1029/2012GL052387.
- Smalley, R. (2009), High-rate GPS: How high do we need to go? *Seism. Res. Lett.*, **80**(6), 1054–1061, doi:10.1785/gssrl.80.6.1054.
- Smyth, A., and M. Wu (2007), Multi-rate Kalman filtering for the data fusion of displacement and acceleration response measurements in dynamic system monitoring, *Mech. Syst. Signal Process.*, **21**, 706–723, doi:10.1016/j.ymssp.2006.03.005.
- Trifunac, M. D., and M. I. Todorovska (2001), A note on the useable dynamic range of accelerographs recording translation, *Soil Dyn. Earth. Eng.*, **21**, 275–286, doi:10.1016/S0267-7261(01)00014-8.
- Tsuda, K., J. Steidl, R. Archuleta, and D. Assimaki (2006), Site-response estimation for the 2003 Miyagi-Oki earthquake sequence considering nonlinear site response, *Bull. Seism. Soc. Am.*, **96**(4A), 1474–1482, doi:10.1785/0120050160.
- Wang, R., B. Schurr, C. Milkereit, Z. Shao, and M. Jin (2011), An improved automatic scheme for empirical baseline correction of digital strong motion records, *Bull. Seism. Soc. Am.*, **101**(5), 2029–2044, doi:10.1785/0120110039.
- Wang, R., S. Parolai, M. Ge, M. Jin, T. R. Walter, and J. Zschau (2013), The 2011 Mw 9.0 Tohoku earthquake: Comparison of GPS and strong motion data, *Bull. Seism. Soc. Am.*, doi:10.1785/0120110264, in press.
- Wright, T. J., N. Houlie, M. Hildyard, and T. Iwabuchi (2012), Realtime reliable magnitudes for large earthquakes from 1 Hz GPS precise point positioning: The 2011 Tohoku-Oki (Japan) earthquake, *Geophys. Res. Lett.*, **39**, L12302, doi:10.1029/2012GL051894.
- Wu, Y. M., and C. F. Wu (2007), Approximate recovery of coseismic deformation from Taiwan strong-motion records, *J. Seismol.*, **11**, 159–170, doi:10.1007/s10950-006-9043-x.