Bedford, J., Bevis, M. (2018): Greedy Automatic Signal Decomposition and Its Application to Daily GPS Time Series. - *Journal of Geophysical Research*, *123*, 8, pp. 6992—7003.

DOI: http://doi.org/10.1029/2017JB014765

Journal of Geophysical Research: Solid Earth

RESEARCH ARTICLE
10.1029/2017JB014765

Greedy Automatic Signal Decomposition and Its Application to Daily GPS Time Series

Jonathan Bedford[1] **and Michael Bevis**[2]

[1]Deutsches GeoForschungsZentrum Telegrafenberg, Potsdam, Germany, [2]School of Earth Sciences, Ohio State University, Columbus, OH, USA

Key Points:
- The Greedy Automatic Signal Decomposition algorithm (GrAtSiD) is a fully automatic trend estimator for GPS time series
- A variety of transient signal shapes can be approximated by one or two multitransient functions; steps are modeled with Heaviside functions
- GrAtSiD decomposes signal into three parts: seasonal, noise, and the sum of secular and transient motion

Supporting Information:
- Supporting Information S1
- Movie S1
- Movie S2
- Movie S3
- Movie S4

Correspondence to:
J. Bedford,
jbed@gfz-potsdam.de

Citation:
Bedford, J., & Bevis, M. (2018). Greedy automatic signal decomposition and its application to daily GPS time series. *Journal of Geophysical Research: Solid Earth*, *123*, 6992–7003. https://doi.org/10.1029/2017JB014765

Received 24 JUL 2017
Accepted 17 JUL 2018
Accepted article online 29 JUL 2018
Published online 15 AUG 2018

Abstract The recognition of transient motion in terrestrial continuous Global Positioning System (GPS) time series implies the knowledge of certain time functions that we assume to be ever present in the time series. By assuming that the permanent time functions are the long-term secular velocity of the Earth and the seasonal oscillations, we define the total remaining signal as transient motion. Here we adopt the multitransient as a versatile function for modeling transient motion over a range of time scales. We define the multitransient as the sum of two or more transient decaying functions with different characteristic time scales and identical onset times. We then demonstrate the greedy approach to fitting the time series by using a minimum number of multitransients (sparse functions) in addition to the permanent time functions in a linear regression. The Greedy Automatic Signal Decomposition algorithm decomposes the signal into three parts: (1) background seasonal motion, (2) secular and transient motion, and (3) a residual (noise). We describe the greedy algorithm with synthetic examples before demonstrating its application to time series of daily GPS solutions. The implementation of the multitransient allows for a more realistic plate-trajectory model, whereby a full range of transient signal time scales, from short-duration slow slip to longer-duration processes such as postseismic slab accelerations or postseismic decay, can all be estimated with the same function. Since Greedy Automatic Signal Decomposition algorithm automatically estimates trend, its application to a GPS network allows for the common mode filter to be applied seamlessly.

Plain Language Summary Over the past few decades, earthquake scientists have been increasing the deployment of continuous GPS stations. This is because high precision time series of surface motions are great for testing the hypotheses of earthquake physics. There is now a wealth of stations (over 15,000 worldwide) but unfortunately no quick way of separating the expected, background plate motion from the unexpected, unusual transient signal. In this study, we present a method aimed at solving this very problem. Adapting an algorithm used in the fields of statistics and electrical engineering called "Greedy Optimization" we show how, by assuming that transient signals are rare occurrences in the time series, it is possible to separate them from the expected signals in a computationally efficient way. We believe that this method, if developed to run even faster (for example with parallelization), will pave the way for rapid analysis of unexpected signals on a spatial scale that has previously been impossible. This opens the door to many new exciting discoveries into how the earth is moving in response to plate-tectonics and weather. Furthermore, we conclude that this method has the potential to be applied to a wide variety of time series data.

1. Introduction

The proliferation of continuous Global Positioning System (cGPS) stations deployed for tectonic monitoring has created a situation where we now have the challenge of uncovering the extent of transient signals in the time series. Such extractions of the suspected transient signal typically entail the visual inspection of the time series after removing or otherwise accounting for seasonal and secular behavior. This workflow is slow and prone to human error when the data sets are very large. Therefore, considerable research has been geared toward automatically separating transient signals of interest: In these studies, the cGPS signal has been decomposed or filtered by a variety of approaches such as singular value decomposition (e.g., principal component analysis/independent component analysis, PCA/ICA; Gualandi et al., 2016; Kositsky & Avouac, 2009), multichannel singular spectrum analysis (Walwer et al., 2016), linear regression (Riel et al., 2014), and Kalman filtering of state vectors (e.g., the network inversion filter; Segall & Matthews, 1997). For regression, onset times of sudden displacements (due to earthquakes, antenna changes, or reference frame shifts) are typically fed into the algorithm either by invoking *jump times* for Heaviside steps or jumps in the time

series or by splitting the time series analysis before and after large, nearly instantaneous displacements or jumps. Such a priori specification of steps in the data is relatively straightforward for many artificial displacements (e.g., antenna changes) but is more difficult in the event of an earthquake since the cutoff distances and magnitudes must be predefined, resulting in steps being invoked where none exist in the time series, and vice versa. The onset of a decay function must also be assigned for larger seismic events to account for the decaying shape of the postseismic signal. For very large events the origin time is well known, but smaller, local earthquakes can also produce transients, and these events may be harder to detect. Such unexpected transients, along with episodic tremor and slip episodes, are individually emergent phenomena and can be identified with detection algorithms applied to networks where they are known or suspected to exist. Typically, the detection of these emergent transients involves the indentification of a spatially coherent anomaly (e.g., Bartlow et al., 2011). Of the various approaches to estimating or retrieving transient behavior, linear regression (e.g., Riel et al., 2014) is the only one that implies an expected signal. The expected *background* signal, in this case, is the extended trajectory model (ETM) of Bevis and Brown (2014), which we reformulate here to use exponential rather than logarithmic transients and also to eliminate the explicit use of a reference time:

$$x(t) = mt + d\sum_{k=1}^{n_k}\left[s_k \sin(\omega_k t) + c_k \cos(\omega_k t)\right] + \sum_{j=1}^{n_j} b_j H(t - t_j) + \sum_{i=1}^{n_i} a_i\left(1 - e^{(-(t-t_i)/T_i)}\right) + \xi(t) \quad (1)$$

where displacement, x, as a function of time, t, is described using a series of secular, seasonal, step, and decay basis functions. We note that, in our experience, the results obtained by invoking equal numbers of exponential or logarithmic transients are so similar that in most contexts it makes little difference which form is chosen. But in either case, the individual transient functions are evaluated only for $t > t_i$. We also note that equation (1) is slightly less general than the ETM of Bevis and Brown (2014) since we assume that the displacement trend component of motion is a constant velocity trend, that is, linear in time, rather than polynomial in time (which would allow, e.g., a quadratic trend implying a constant and persistent acceleration).

The secular velocity is represented by gradient m and constant d. The seasonal oscillations are represented by coefficients s_k and c_k where the number of frequencies, n_k, is normally two, and ω_1 and ω_2 are angular frequencies associated with annual and semiannual periods. H represents the Heaviside functions with coefficients b_j at n_j number of jump times t_j specified a priori. Some of these jumps will have decays associated with them, with coefficients a_i and decay constants T_i at n_i onset locations, and with t_i being the decay onset time. ξ (t) is the noise in the data and is here assumed to be normally distributed. In Riel et al. (2014), the unexpected (transient) signal is modeled as a series of B-splines in linear combination with equation (1) and is retrieved by minimizing the number of transient functions that, in addition to the expected functions, solve the optimization problem:

$$\text{minimize}\left\{\|d - p\|_2 + \lambda_a\|C_a\|_2 + \lambda_b\|C_b\|_1\right\} \quad (2)$$

where vectors d and p contain the data and model predictions, vectors C_a and C_b contain the model coefficients for the permanent and sparse functions, and λ_a and λ_b are the regularization weights. The second and third terms in equation (2) are the regularization of the inversion, and therefore, the B-spline model parameters are regularized in the L1 sense. L1 regularization promotes sparsity in the solution vector (C_b), which means that the weighting of the L1 regularization determines the number of nonzero coefficients in the C_b vector.

In this study we approximate the linear regression approach of Riel et al. (2014) but with a different form of the transient (unexpected) time functions in the time series and with a greedy approximation to the L1 regularized optimization. It will be shown that the time series can be fit with permanent functions of seasonal and secular in addition to as few transient functions as possible. Importantly, we will show that there is no need to assume the onset of any steps or decays in the time series. Accordingly, we define the transient signal as being everything other than the seasonal, secular, and noise, therefore modifying the ETM of equation (1) to the following form:

$$x(t) = mt + d + \sum_{k=1}^{n_k}\left[s_k \sin(\omega_k t) + c_k \cos(\omega_k t)\right] + \sum_{j=1}^{n_j} b_j H(t - t_j) + \sum_{r=1}^{n_r}\sum_{i=1}^{n_i}\left[A_i\left(1 - e^{(-(t-t_r)/T_i)}\right)\right] + \xi(t) \quad (3)$$

where $n_r + n_j$ transient functions model the steps, decays, and other transient signal shapes. Transient signal

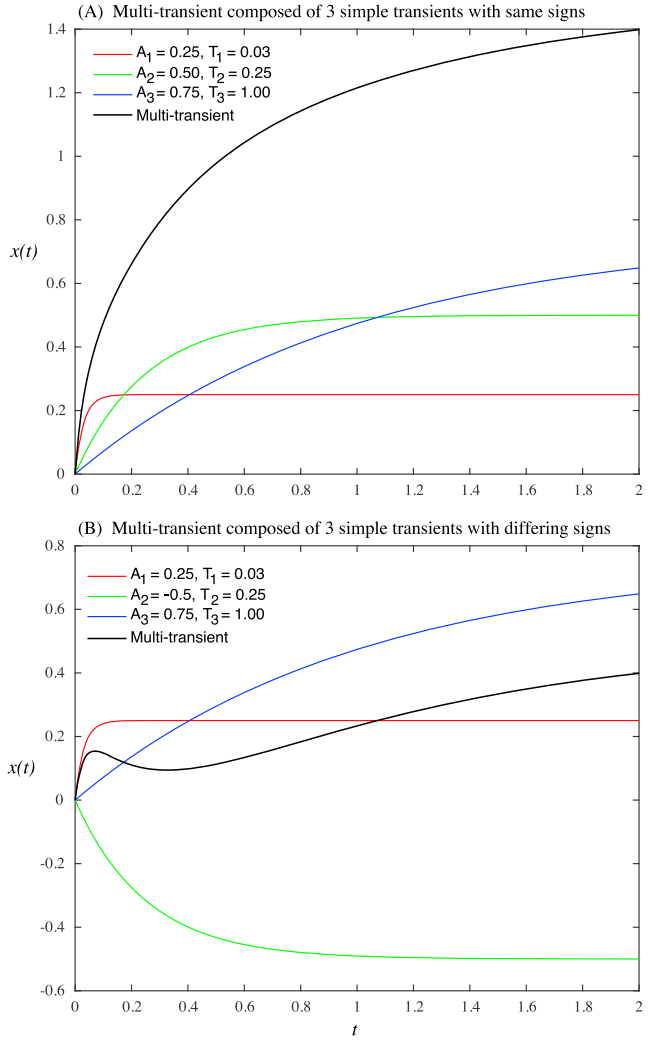

Figure 1. (a) Colored exponential decay functions with time constants and coefficients indicated by the legend. Black function shows the summation of these three original functions. All coefficients have the same sign. (b) Same as (a) except that the coefficients have different signs, therefore producing reversals of motion in the decaying time series.

might be composed of earthquake or equipment related sudden displacements, slow slip, postseismic decay, or other noise such as anomalous atmospheric conditions or precipitation loading. In the formulation of equation (3), the step-like displacements are modeled with the Heaviside function and the more gradual transients with the multitransient. We define a multitransient as the sum of two or more simple exponential decay functions with the same onset time but different characteristic delay times. Weiss et al. (2016) and Loveless and Meade (2016) have used this exponential form of the multitransient to improve their ability to estimate interseismic velocities in Bolivia and Japan, respectively. In this study we have specifically chosen exponential functions because a sum of exponentials reaches an asymptotic value, a feature that allows for finite motions of various transient signals to be approximated. Here the multitransient with coefficients, A_i, is a linear combination of n_i exponential decays with identical starting time, t_r, and a predetermined selection of n_i decay constants T_i. Figure 1 shows how individual exponential functions combine to produce particular multitransients. A constant sign for the coefficients, A_i, produces a decay similar to that of a single exponential decay function, whereas variation in the signs can produce a decaying function with reversals in sense of motion. The ability of the multitransient to produce both smoothly and bumpily decaying time series means that it is a versatile enough function for approximating a variety of transient signal shapes and time scales. In the synthetic examples it will be demonstrated that by combining just a couple of these multitransients in series, it is possible to approximate a wide variety of signal shapes. We will first introduce the Greedy Automatic Signal Decomposition algorithm (GrAtSiD) that is designed to recover a minimum number of multitransient functions that fit the time series in linear combination with the permanent time functions. Results will then be shown for the synthetic case and real data examples on individual cGPS stations and application to a Global Navigation Satellite System network.

2. Method

2.1. The Greedy Algorithm

L1 regularization, often called the *Lasso* or sparse regularization, typically requires a large dictionary of sparse functions through which the inversion procedure will search until it converges on a solution that requires a minimal number of these functions (Hastie et al., 2015). Such L1 regularization to the optimization problem is significantly slower than L2 regularization, and therefore, the sparse dictionaries should be constructed conservatively if solving with convex optimization approaches (e.g., Boyd & Vandenberghe, 2004; Grant et al., 2008). When the number of required sparse functions in the model space exceeds reasonable computational limitations, globally optimal sparse solutions can be approximated with iterative methods such as coordinate descent (e.g., Tseng, 2001), orthogonal matching pursuit (e.g., Cai & Wang, 2011), or greedy algorithms (see Needell et al., 2008). In the simplest terms, a greedy algorithm works by iteratively selecting a sparse basis function from the sparse dictionary that can best fit the residual of the current iteration. This procedure requires many computationally cheap intermediate inversions or correlations to be performed in each iteration, as the remaining sparse functions are first sifted through, before the best new candidate sparse functions from this initial stage are tested alongside existing sparse and permanent basis functions in the regression of the full data. Such greedy algorithms can be designed to satisfactorily converge given an appropriate rejection criteria for the candidate sparse basis functions: After some number of iterations, of the remaining sparse basis functions that have not already been incorporated into the regression of the full data, none are able to improve the fit of the regression above a certain predefined threshold and at this point the optimization has converged.

In our application (additional details with figures and movies included in the supporting information), at each iteration of the algorithm we fit the residual with many different combinations of two transient onsets (TOs) to find two new candidate TOs. The two candidate TOs are the combination that produce the best fit to the current residual. In the forthcoming examples, each TO is either a matrix of three sparse basis functions representing the multitransient or a Heaviside function and therefore the design matrices for the tested combinations of two TOs has four or six columns at the length of the data vector.

Note that the number and values of decay constants must be predefined as hyperparameters. In all following examples, we use the values 10^3, 10^2, and 10^1 as decay constants since they allow for the estimation of a variety of transient shapes. The matrix that horizontally concatenates all the possible TOs is called the dictionary of sparse functions. The initial residual is found by fitting a straight line (secular basis function) to the original time series and taking the difference. Subsequent residuals depend on which TOs are deemed necessary after a combinatorial exploration of the current candidate TOs alongside the permanent (seasonal and secular) basis functions.

Each iteration of the Greedy algorithm will perform thousands of very fast inversions. These inversions take the form:

$$m = \left(G^T G + I.\varepsilon^2\right)^{-1} G^T d \tag{4}$$

where m is the vector of coefficients for the particular inversion, G represents the basis functions, and d represents the signal currently being fit. The Tikhonov damping (identity matrix I) is weighted by the same amount, ε^2, in all inversions. Therefore, our greedy algorithm uses L2 regularization despite being designed to mimic the L1 regularization of sparse basis functions. The procedure for rejecting candidate TOs aims to balance the improvement of fit to the data with minimizing the number of TOs. Therefore, on each iteration, we select the smallest list of TOs that improves the fit above a certain predefined threshold and reject the TOs not in that list. If no list of TOs can exceed this threshold, the two newest candidate TOs are rejected. Rejected candidate TOs are placed into a quarantine that prevents them being searched for in the subsequent iteration. If new TOs are added to the solution, a swap-out test is performed whereby we swap the type of transient (either Heaviside or Multitransient) within a predefined window spanning before and after each TO. The swap-out stage continues until neither swapping the transient type nor adjusting the onset time can improve the fit of the regression. An iteration is completed following either the rejection of both new candidate TOs or the completion of the swap-out phase. Convergence is decided when the list of TOs at the end of each iteration is the same over a predefined number of iterations (e.g., 5). Before starting the next iteration, the updated residual is calculated as the difference between the full data and the prediction of the permanent and accepted sparse basis functions. Gaussian-distributed random noise with a standard deviation of the updated residual is added to the residual before the next iteration. Both addition of random noise and the quarantining of rejected basis functions are measures taken to mitigate convergence to local minima. After convergence, the algorithm can be repeated from the starting residual, or alternatively, a fraction of the converged list of TOs can be randomly removed and the starting residual calculated from that point. After a predefined number of convergences, the extracted signals can be analyzed together to gauge a statistical measure of uncertainty in the transient signal. In the following synthetic and data examples we run the algorithm for 10 convergences. After each convergence we remove half of the TOs at random. For each converged extracted transient signal, we can define its similarity to the other extracted signals using the measure of similarity, S:

$$S_i = \frac{1}{\sum_{j=1}^{n} \|M_i - M_j\|2} (i \neq j) \tag{5}$$

where the similarity score for each converged series, i, is the sum of norms of the difference in n transients after transient time series M_j has been tilted and offset to minimize its difference to transient time series M_i. In the synthetic examples shown in this study, we display the transient with highest similarity to the others.

The hyperparameters of the algorithm include the magnitude of the Tikhonov damping (ε^2 in equation (4)), decay constants and number of transients in each multitransient, misfit improvement threshold for accepting new TOs, window size of the swap-out stage, and the number of convergences. A full list of hyperparameters and their effect on the algorithm is listed in the supporting information. In choosing hyperparameters, one

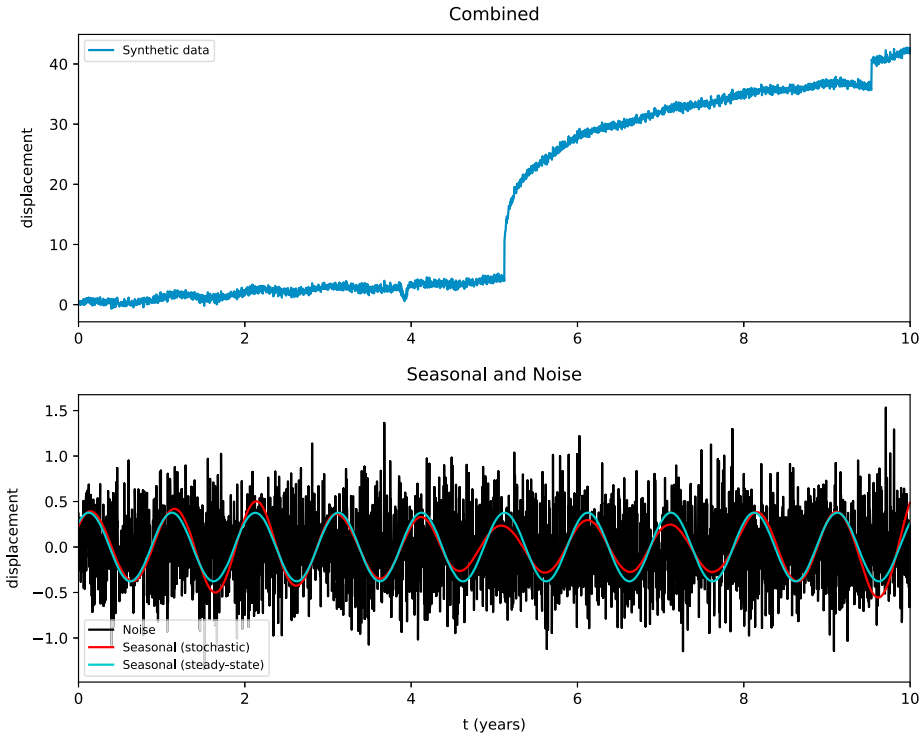

Figure 2. The top panel shows the synthetic signal complete with random Gaussian distributed noise. The signal is the sum of a linear secular motion, two steps, arctangent functions, Gaussian functions, stochastic seasonal oscillations, and Gaussian distributed noise. The lower panel shows the background and stochastic seasonal oscillations and the Gaussian distributed noise. With the Greedy Automatic Signal Decomposition algorithm, we aim to recover the background seasonal and Gaussian noise to leave behind the secular and transient motions.

must consider available computational resources, expected shape of the transient signal, and the balance between overfitting and underfitting the data.

In both real and synthetic applications of this algorithm, the model parameter for the secular term can converge to a quite unrealistic value. This is because the long wavelength of the secular term can trade off with the transient signal, especially if the solution is supported by a TO early on in the time series. This problem is something that can be expected since some GPS time series, upon visual inspection, seem to have a nonconstant linear velocity. Furthermore, such changes in apparent secular velocity have been shown in recent literature (Heki & Mitsui, 2013; Loveless & Meade, 2016; Melnick et al., 2017). Therefore, we can think of GrAtSiD as decomposing the signal into three parts: (1) background (steady state) annual and semiannual oscillations, (2) secular and transient motion represented by the secular, Heavisides, and multitransient terms of the regression, and (3) the remaining residual (noise). Accordingly, by subtracting the residual from the original time series, we are left with an estimation of trend.

3. Results

3.1. Synthetic Case

Figure 2 shows a synthetic cGPS time series consisting of secular, seasonal, and transient (unexpected) motions. This particular synthetic signal is similar to what is observed at active plate boundaries, where earthquakes, decays, and other transient motions are thought to accompany the background secular and seasonal motions. The unexpected signals in this case are represented by some step (Heaviside), decay, arctangent, and Gaussian functions (plotted in isolation in Figure 3). We additionally consider unexpected seasonal motions that arise in nature due to anomalous climactic effects by means of a stochastic contribution to the seasonal (see equation (12) of Davis et al., 2012). Finally, we add Gaussian distributed random noise to the combined signal. Figure 3 shows the extracted transient and seasonal signals and the overall fit of the linear regression to the synthetic data for the transient extraction with highest similarity

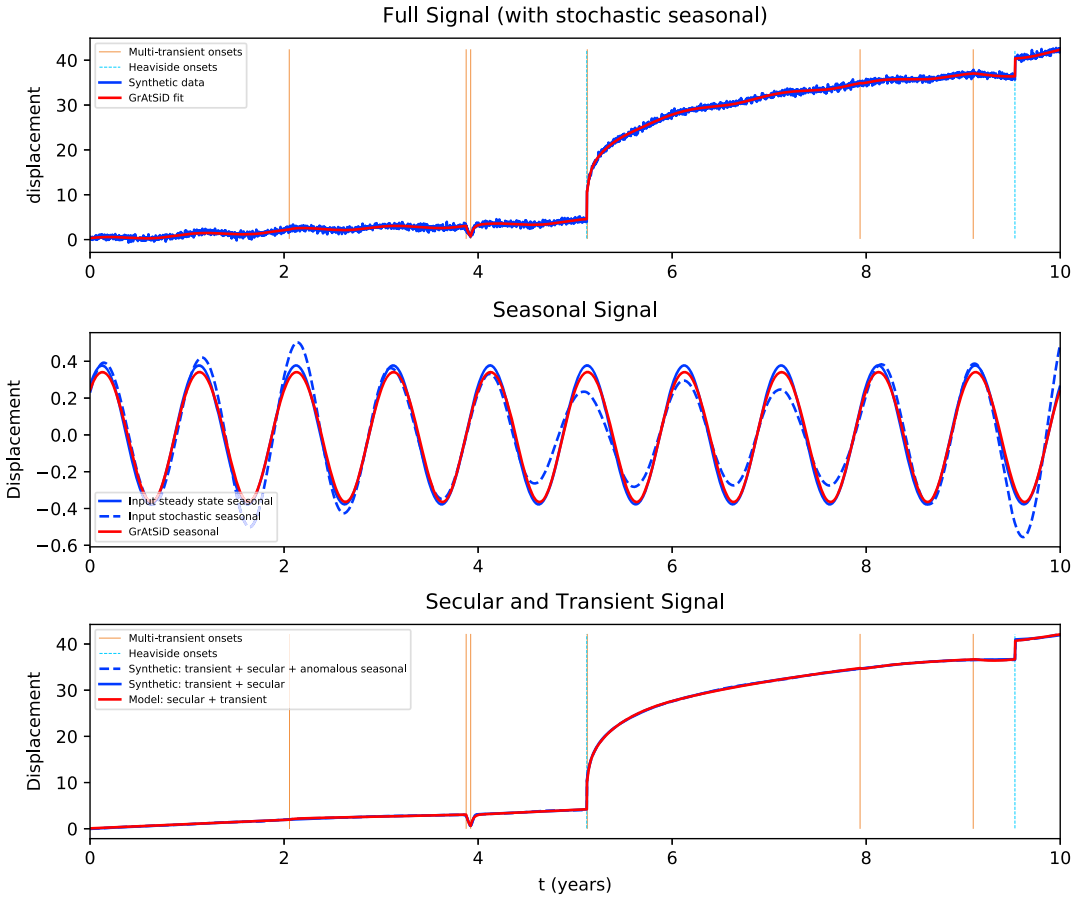

Figure 3. (top) The synthetic data and the fit from GrAtSiD are shown along with the Heaviside and multitransient onset times. This panel demonstrates how GrAtSiD functions as a trend estimator. (middle) Input and extracted synthetic seasonal signals. (bottom) The input synthetic secular and transient signal and the GrAtSiD fit. The extracted transient has been detrended and shifted to minimize residual to the known input transient. GrAtSiD = Greedy Automatic Signal Decomposition algorithm.

score out of 15 converged transient signals. Note that, in this example, the GrAtSiD algorithm is run blind—it is not given any information about the onset time of steps. In Figure 3 we see a decent recovery of both the input background seasonal and transient signals as well as the successful discrimination between sudden and more gradual transients. In recovering the steady state seasonal in the presence of seasonal variability, the algorithm is effectively seeing the anomalous seasonal motion as part of the transient motion that can be fit with the multitransient functions. For most gradual transient shapes, such as the one occurring at t = 4 years, we see that typically one TO or two TOs in series are able to approximate the signal. Figure 4 shows the residual for each iteration along with the best new candidate TOs. We see that the transients tend to be identified in order of magnitude, with subsequent TOs being less likely to be retained in the converged upon list of TOs (the algorithm is fully illustrated in supporting information section S2 and with Movies S1–S3). The performance of the algorithm in the presence of a perfectly steady state seasonal signal can be found in the supporting information. The supporting information contains two more examples of synthetic transient signal recovery (supporting information Figures S1–S11).

3.1.1. Real Data—Single Station Approach

We also applied the greedy algorithm to daily solutions of cGPS for two real stations EMAT and GISB. EMAT was processed as outlined in Melnick et al. (2017), while GISB was processed according to the Nevada Geodetic Lab procedures (see http://geodesy.unr.edu/gps/ngl.acn.txt). Additionally, EMAT was filtered for common mode noise, as outlined in the next section (stations used for common mode filtering shown in supporting information Figure S14). Figure 5 shows the data and estimated signals after detrending. Again, we show the model corresponding to the transient signal with the highest similarity score out of 10 converged transient signals. The algorithm succeeds in extracting the transient signal spotted by human

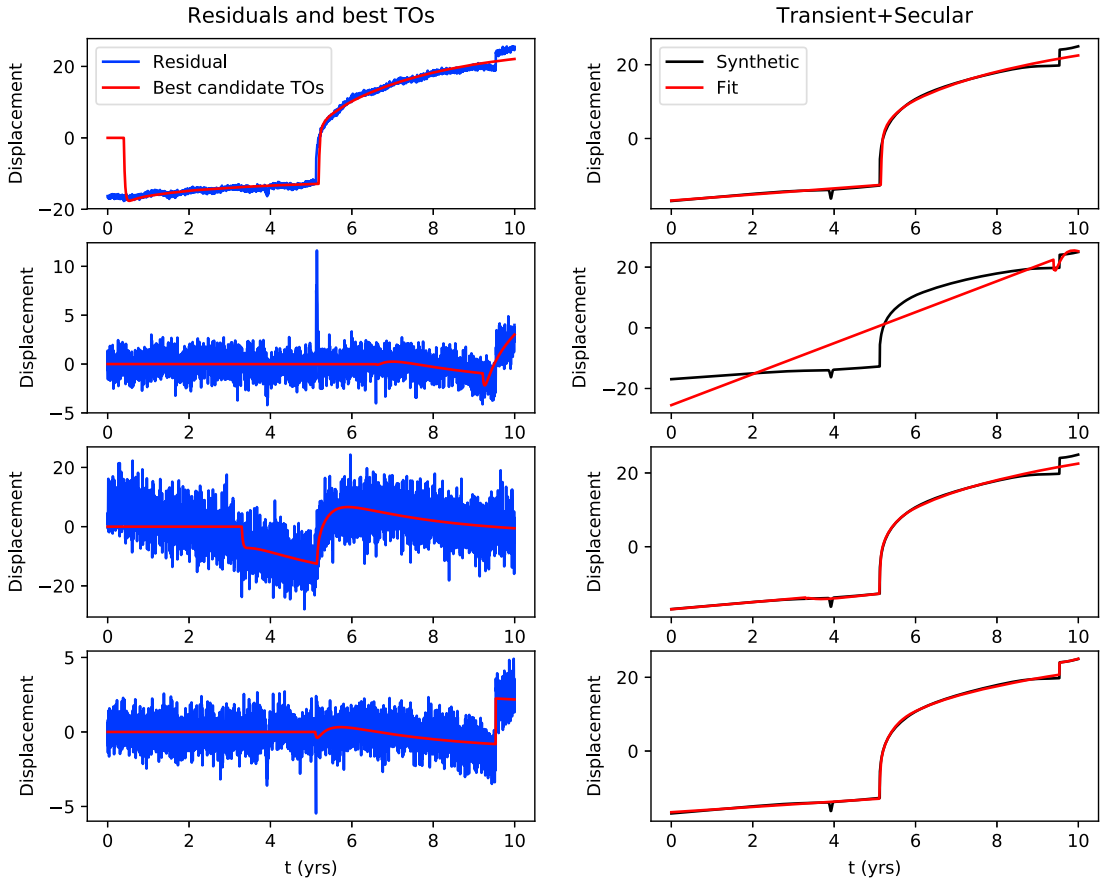

Figure 4. Descending the left column shows the combination of two TOs that best fits the current residual for consecutive iterations. Descending the right column shows the estimation of the synthetic transient signal for consecutive iterations. TO = transient onset.

eye in the published cases of slow slip (GISB; Wallace et al., 2016) and long-term velocity changes (EMAT; Melnick et al., 2017). Since it is a regression-based approach, GrAtSiD is easily able to handle missing data in the time series. We suggest only the removal of outliers from the time series before running GrAtSiD, since other denoising measures such as median filtering can smooth over transients in certain scenarios.

3.1.2. Real Data—Common Mode Noise Reduction With a Network Approach

Following convergence, GrAtSiD has decomposed the data into three parts: (1) the background seasonal signal, (2) the sum of secular and transient functions, and (3) the remaining residual (i.e., the signal minus parts 1 and 2). Therefore, the remaining residual contains the true noise in addition apparent noise arising from the underfitting or overfitting of the regression. Wdowinski et al. (1997) demonstrated how the noise in daily time series of cGPS networks is correlated between stations and can therefore be diminished by averaging and subtracting. This common mode filter (CMF) approach only works when one has a good estimate on the background trend in the time series. Rather conveniently, GrAtSiD automatically recovers this background trend (sum of parts one and two), and therefore, the CMF can be easily applied to time series in a network following the application of GrAtSiD. Figure 6 shows time series at a station in the GEONET cGPS network in Japan (see supporting information for network map and location of the time series shown). GrAtSiD was applied blind on all time series of daily F3 cGPS solutions, and the median of the residuals for each day was removed from the time series. These results demonstrate how GrAtSiD allows for a very effective noise reduction via the CMF.

The supporting information includes a movie (Movie S4) for the background seasonal signal of the Japanese network extracted by running GrAtSiD on the common mode filtered solutions. With a station-by-station approach to the decomposition, we recover a fairly spatiotemporally coherent background seasonal signal, even with a blind application of the algorithm and one convergence per time series. Indeed, GrAtSiD could be modified to enforce some spatial coherency (e.g., as shown in Riel et al., 2014). Alternatively, spatially

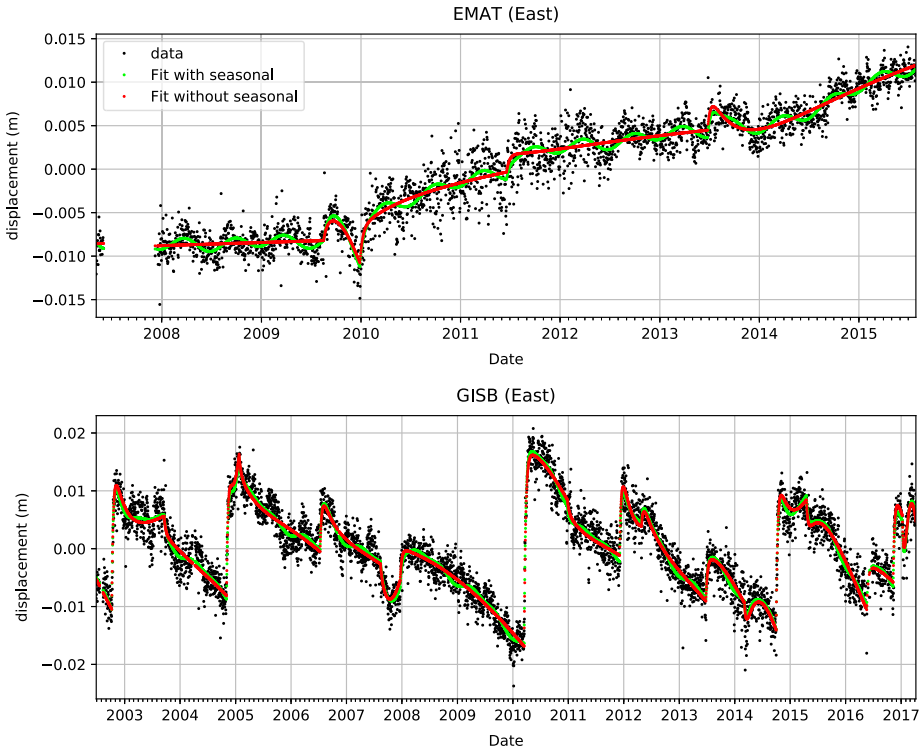

Figure 5. Top panel shows the data (black) for the manually detrended east component of station EMAT in Chile (see map in supporting information Figure S14 for location). The GrAtSiD model is shown including the secular, transient functions and background seasonal (green) and without the seasonal (red). We can see that GrAtSiD is able to model the change in interseismic velocity that this station undergoes (see Melnick et al., 2017). The bottom panel shows the east component of station GISB in New Zealand (also presented in Wallace et al., 2016). Colors are the same as panel above. We see that GrAtSiD is able to automatically model the slow-slip-related motions. Note, that both examples are for GrAtSiD run blind (without steps in the data being predefined). GrAtSiD = Greedy Automatic Signal Decomposition algorithm.

coherent transient motions might be found by applying PCA/ICA (Gualandi et al., 2016; Kositsky & Avouac, 2009) to the data minus the noise extracted by GrAtSiD (i.e., applying PCA/ICA on the trend of the data as estimated by GrAtSiD). Our choice, in this study, to solve for transient functions on a station-by-station basis, allows for very local transients to manifest themselves in the model. Such local transients might be controlled by site specific conditions such as soil/monument instability (see Williams et al., 2004) or

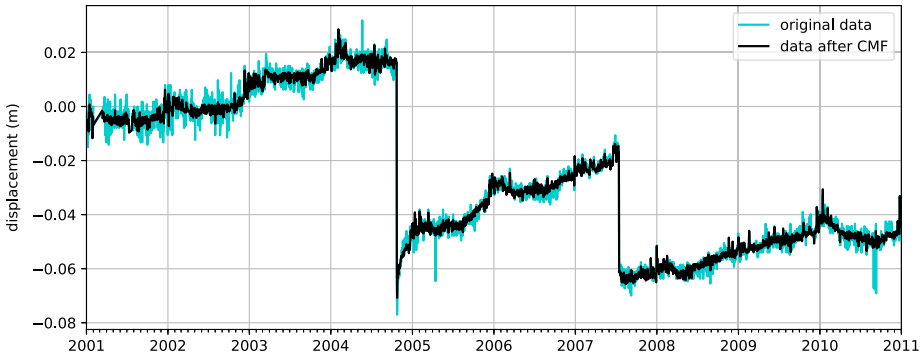

Figure 6. Top panel shows the daily continuous time series for station Ojiya in the Japanese network (location shown on Figure S15 in the supporting information). Processed F3 solution shown in light blue. Time series after first common mode filter shown in black. Red line is the fit to secular, transient, and background seasonal motion on the common mode filtered data (black). Bottom panel shows the original time series (same as light blue in top panel) and red shows the removal of the common mode noise and seasonal from GrAtSiD run on the common mode filtered data. Note that GrAtSiD was run blind to any known steps in the data. GrAtSiD = Greedy Automatic Signal Decomposition algorithm; CMF = common mode filter.

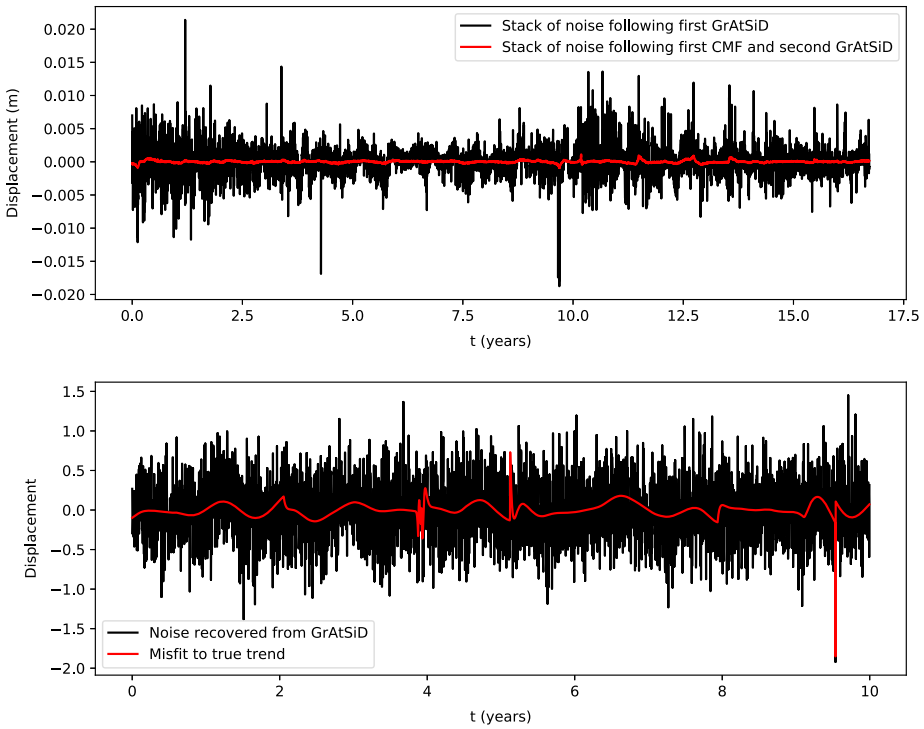

Figure 7. Top panel: Black time series shows the median of residuals following the first GrAtSiD application to the Japanese continuous Global Positioning System network. Red time series shows the median of residuals following common mode filtering and application of GrAtSiD again. Bottom panel: Black time series shows the residual recovered from running GrAtSiD on the synthetic time series (Figure 2). Red shows the misfit of the gratsid recovered trend (synthetic data minus the recovered residual) to the actual trend (data minus the synthetic random noise). GrAtSiD = Greedy Automatic Signal Decomposition algorithm; CMF = common mode filter.

human-controlled depletion of a nearby water resource. By enforcing spatial coherency, local transients would increasingly be mapped into noise with increasing weight and spatial extent of the enforced coherency. The choice of an appropriate spatial extent and weighing of coherency would depend on the expected spatial wavelength of transient signals and the spatial density of measurements.

4. Discussion

An ongoing aim of the tectonic GPS research community is to be able to both extract and detect transient motion in the GPS (see Lohman & Murray, 2013). Algorithmic transient detection aims to identify periods in time during which there occurs an acceleration and sustained velocity that is statistically distinct from the background (expected) velocity (e.g., Crowell et al., 2016). Accordingly, we do not suggest that GrAtSiD is detecting transient motion. Rather, more specifically, it is extracting transient motion, whereby the very low frequency component of the transient signal trades off heavily with the secular velocity. Given this trade-off, we suggest that the secular signal can be estimated to leave behind a relative transient, R:

$$R(t) = d(t) - p_B(t) - s(t) - \xi(t) \tag{6}$$

where t is time, d is the original time series, p_B is the prediction from a simple background tectonic velocity model, and s and ξ are the seasonal and noise isolated from the GrAtSiD algorithm. Such a background tectonic velocity model might be a backslip prediction (e.g., Savage, 1983) based on a uniform locking pattern across the largest known nearby faults with a velocity estimated from plate velocity models.

From time series presented in this paper it is clear that the ability to extract a transient signal is dependent on the noise level, the magnitude of the transient, and the shape of its function. This result is obvious yet is not quantified within this study. Supporting information Figures S6–S11 show the full suites of transient signals extracted upon each convergence for the synthetic examples (Figure 3 and S1–S5). Supporting information

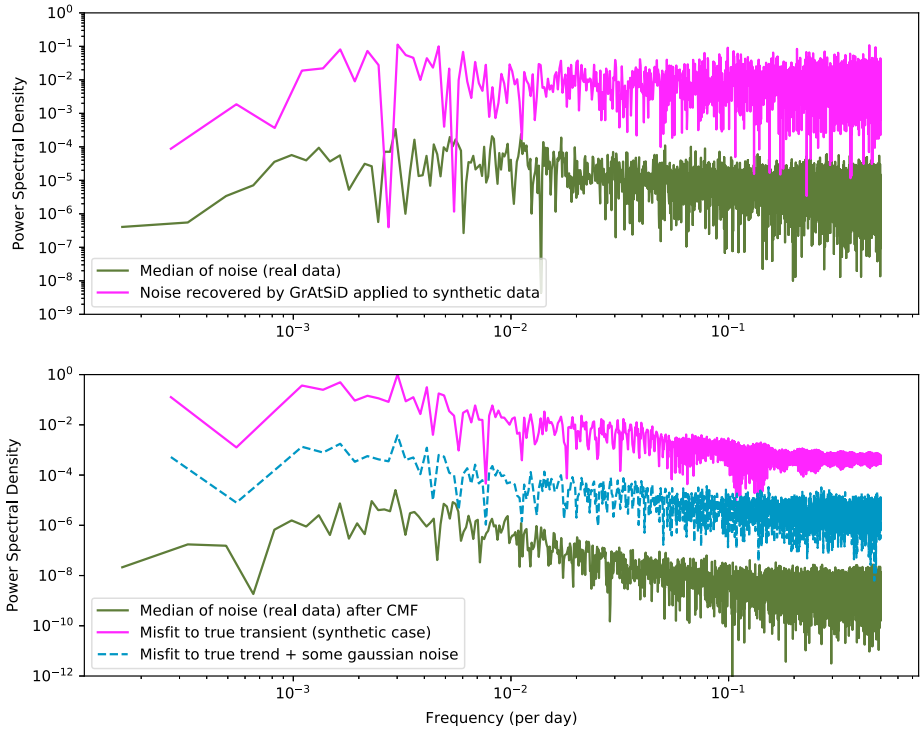

Figure 8. Top panel shows the power spectral densities of the black time series of Figure 7. Bottom panel shows the power spectral densities of the red time series of Figure 7. All power spectral densities have been shifted along the *y* axis for clarity. Pink and dark green represent synthetic and real data respectively. In the bottom panel, the dashed blue power spectral density represents the synthetic misfit with Gaussian noise added (standard deviation of added white noise is 10% of the standard deviation of the white noise in the synthetic time series). GrAtSiD = Greedy Automatic Signal Decomposition algorithm; CMF = common mode filter.

section S1 outlines a basic investigation into the uncertainty of the extracted transient signal, but this analysis is limited by our low number of convergences (in this case 10) for each synthetic signal. Accordingly, we suggest that future work could involve performance benchmarking of the algorithm in terms of investigating how transient signal recovery is impacted by noise, transient signal shape, and transient signal magnitude. Development of the code, such as parallelization, would allow for extensive hyperparameter testing that is currently too expensive (see details in supporting information). Furthermore, we have only shown examples of the application to one directional component of a three-component GPS time series. The algorithm could also be readily modified so that the multitransient functions exist in 3-D space.

Another important consideration in assessing the error of the decomposed signal is the characteristic noise of the time series. During the synthetic testing of the algorithm, noise added to the time series was considered to be Gaussian. Previous studies of the noise in cGPS have suggested that GPS time series contain significant amounts of colored noise such as flicker and random walk (e.g., Dmitrieva et al., 2015; Langbein, 2008). In this study, we can compare the noise extracted from the GEONET solutions with the noise extracted from the synthetic examples. Figure 7 shows the median noise of the east component of the GEONET solutions as recovered from the first and second runs of GrAtSiD. The data have been common mode filtered after the first run. We see that there is some colored noise in both solutions, but with considerably more high-frequency component in the first residuals from the first GrAtSiD run. Similarly, Figure 7 shows the final noise recovered from running GrAtSiD on the synthetic data (time series in Figures 2 and 3) and the misfit of the sum of secular and transient functions to the true secular and transient motions (where true transient includes also the anomalous seasonal signal). In presenting these synthetic and real signals on the same plot we aim to frame our understanding of noise in the data in terms of underfitting and overfitting of the data. In the lower panel, the misfit to the true signal is showing the model's failure to recreate the true shape of transient motions. Similarly, the second stack of the residuals in the top panel is showing us the systematic

failure of the regression to fit the transient shapes in the data in Japan. The higher contribution of the lower frequencies in this second residual stack is showing us that there is a transient signal in the time series that is common to most stations in the Japanese network. Whether or not this is explained by tectonics or processing is unclear, but such motion is systematically underfit or overfit by the regression. Figure 8 shows the power spectral densities (PSDs) for the time series shown in Figure 7. Here we see that the first stack of the noise in the real data has a similar PSD to the residual recovered from synthetic data with Gaussian noise. The addition of a small amount of Gaussian noise (with standard deviation 10% of that added to the original synthetic time series) to the true misfit in the synthetic case results in a PSD similar to that of the median noise in the CMF filtered real data. In this case, adding a small amount of Gaussian noise to the true misfit is simulating the inability of the CMF to remove all noise in the real data. Therefore, with the assumption of only Gaussian noise in the synthetic GPS time series we are able to recreate PSDs comparable to real data examples. This is not to say that other noise such as flicker and random walk are not present in the real data, rather that the assumption of a Gaussian noise (implied by the application of a least squares regression) is not too unreasonable.

As mentioned in section 1, standard regression approaches to the GPS signal separation problem often specify functions for a list of assumed earthquake jumps and times of equipment or processing strategy related shifts in the time series. The algorithm presented in this paper can easily be modified to include Heaviside functions at predefined times corresponding to known shifts in the time series.

5. Conclusions

We have shown that, by using a minimum number of versatile multitransient functions, we are able to approximate signal shapes with a range of characteristic time scales. Such signal might include very low frequency postseismic decay, subtle subduction velocity changes, or very high frequency steps such as elastic responses to earthquakes. By using these multitransients as sparse dictionary basis functions in a greedy optimization algorithm, we have shown that, with a priori knowledge on the periodic wavelengths present in the signal, we are able to successfully extract transients from realistic synthetic time series. We need not assume any steps in the signal beforehand (e.g., from known nearby earthquakes and artificial equipment or processing related steps). Therefore, GrAtSiD is an automatic trend estimator (e.g., Blewitt et al., 2016). The greedy algorithm allows us to explore a vast dictionary of sparse transient functions and offers an alternative to using L1 regularized optimization solvers.

GrAtSiD separates the signal into three components: (1) the background seasonal signal, (2) the sum of secular and transient functions, and (3) the remaining residual. The modeled secular and transient trade off significantly and are therefore better treated as one component of the decomposition. Transient motion can be invoked by detrending this component according to an assumption of the constant background tectonic activity. Such a grouping of secular and transient motions can be convenient for modeling nonsteady state interseismic motions that are increasingly apparent in the cGPS record at multiple time scales. We envisage future developments of the GrAtSiD such as a parallelization of the code to facilitate faster convergences and a comprehensive investigation of hyperparameter choices. The algorithm shown in this paper has been developed for the application to daily GPS time series but could be modified to extract unexpected signals from nay time series where the periodicities are well known.

Acknowledgments

J. B. is grateful for funding provided by German Science Foundation (DFG) grant MO-2310/3. J. B. appreciates useful discussions with GFZ colleagues, in particular, Marcos Moreno and Chrisoph Sens-Schoenfelder. For Japanese GPS data contact the GSI Japan. GISB time series was downloaded from the Nevada Geodetic Laboratory website. Accordingly, we acknowledge the New Zealand GeoNet project and its sponsors EQC, GNS Science, and LINZ for providing data used in this study. South American time series were processed by Zhiguo Deng at GFZ Potsdam. For access to EMAT time series and stations used to make the CMF for EMAT, contact Jonathan Bedford: jbed@gfz-potsdam.de. We would like to thank two anonymous reviewers for their comments that contributed to the improvement of this manuscript.

References

Bartlow, N. M., Miyazaki, S. I., Bradley, A. M., &, Segall, P. (2011). Space-time correlation of slip and tremor during the 2009 Cascadia slow slip event. *Geophysical Research Letters*, *38*, L18309. https://doi.org/10.1029/2011GL048714

Bevis, M., & Brown, A. (2014). Trajectory models and reference frames for crustal motion geodesy. *Journal of Geodesy*, *88*(3), 283–311. https://doi.org/10.1007/s00190-013-0685-5

Blewitt, G., Kreemer, C., Hammond, W. C., & Gazeaux, J. (2016). MIDAS robust trend estimator for accurate GPS station velocities without step detection. *Journal of Geophysical Research: Solid Earth*, *12*, 2054–2068. https://doi.org/10.1002/2015JB012552

Boyd, S., & Vandenberghe, L. (2004). *Convex optimization*. Los Angeles: Cambridge University Press. https://doi.org/10.1017/CBO9780511804441

Cai, T. T., & Wang, L. (2011). Orthogonal matching pursuit for sparse signal recovery with noise. *IEEE Transactions on Information Theory*, *57*(7), 4680–4688. https://doi.org/10.1109/TIT.2011.2146090

Crowell, B. W., Bock, Y., & Liu, Z. (2016). Single-station automated detection of transient deformation in GPS time series with the relative strength index: A case study of Cascadian slow slip. *Journal of Geophysical Research: Solid Earth*, *121*, 9077–9094. https://doi.org/10.1002/2016JB013542

Davis, J. L., Wernicke, B. P., & Tamisiea, M. E. (2012). On seasonal signals in geodetic time series. *Journal of Geophysical Research*, *117*, B01403. https://doi.org/10.1029/2011JB008690

Dmitrieva, K., Segall, P., & DeMets, C. (2015). Network-based estimation of time-dependent noise in GPS position time series. *Journal of Geodesy*, *89*(6), 591–606. https://doi.org/10.1007/s00190-015-0801-9

Grant, M., Boyd, S. and Ye, Y., 2008. CVX: Matlab software for disciplined convex programming.

Gualandi, A., Serpelloni, E., & Belardinelli, M. E. (2016). Blind source separation problem in GPS time series. *Journal of Geodesy*, *90*(4), 323–341. https://doi.org/10.1007/s00190-015-0875-4

Hastie, T., Tibshirani, R., & Wainwright, M. (2015). *Statistical learning with sparsity*. Boca Raton, FL: CRC Press. https://doi.org/10.1201/b18401

Heki, K., & Mitsui, Y. (2013). Accelerated pacific plate subduction following interplate thrust earthquakes at the Japan trench. *Earth and Planetary Science Letters*, *363*, 44–49. https://doi.org/10.1016/j.epsl.2012.12.031

Kositsky, A. P., & Avouac, J. P. (2009). Inverting geodetic time-series with a principal component analysis-based inversion method (PCAIM). *Journal of Geophysical Research*, *115*, B03401. https://doi.org/10.1029/2009JB006535

Langbein, J. (2008). Noise in GPS displacement measurements from Southern California and Southern Nevada. *Journal of Geophysical Research*, *113*, B05405. https://doi.org/10.1029/2007JB005247

Lohman, R. B., & Murray, J. R. (2013). The SCEC geodetic transient-detection validation exercise. *Seismological Research Letters*, *84*(3), 419–425. https://doi.org/10.1785/0220130041

Loveless, J. P., & Meade, B. J. (2016). Two decades of spatiotemporal variations in subduction zone coupling offshore Japan. *Earth and Planetary Science Letters*, *436*, 19–30. https://doi.org/10.1016/j.epsl.2015.12.033

Melnick, D., Moreno, M., Quinteros, J., Baez, J. C., Deng, Z., Li, S., & Oncken, O. (2017). The super-interseismic phase of the megathrust earthquake cycle in Chile. *Geophysical Research Letters*, *44*(2), 784–791. https://doi.org/10.1002/2016GL071845

Needell, D., Tropp, J., & Vershynin, R. (2008, October). Greedy signal recovery review. In *Signals, systems and computers, 2008 42nd Asilomar Conference on* (pp. 1048-1050). IEEE.

Riel, B., Simons, M., Agram, P., & Zhan, Z. (2014). Detecting transient signals in geodetic time series using sparse estimation techniques. *Journal of Geophysical Research: Solid Earth*, *119*, 5140–5160. https://doi.org/10.1002/2014JB011077

Savage, J. C. (1983). A dislocation model of strain accumulation and release at a subduction zone. *Journal of Geophysical Research: Solid Earth*, *88*(B6), 4984–4996. https://doi.org/10.1029/JB088iB06p04984

Segall, P., & Matthews, M. (1997). Time dependent inversion of geodetic data. *Journal of Geophysical Research*, *102*, 22–391. https://doi.org/10.1029/97JB01795

Tseng, P. (2001). Convergence of a block coordinate descent method for nondifferentiable minimization. *Journal of Optimization Theory and Applications*, *109*(3), 475–494. https://doi.org/10.1023/A:1017501703105

Wallace, L. M., Webb, S. C., Ito, Y., Mochizuki, K., Hino, R., Henrys, S., et al. (2016). Slow slip near the trench at the Hikurangi subduction zone, New Zealand. *Science*, *352*(6286), 701–704. https://doi.org/10.1126/science.aaf2349

Walwer, D., Calais, E., & Ghil, M. (2016). Data-adaptive detection of transient deformation in geodetic networks. *Journal of Geophysical Research: Solid Earth*, *121*, 2129–2152. https://doi.org/10.1002/2015JB012424

Wdowinski, S., Bock, Y., Zhang, J., Fang, P., & Genrich, J. (1997). Southern California permanent GPS geodetic array: Spatial filtering of daily positions for estimating coseismic and postseismic displacements induced by the 1992 Landers earthquake. *Journal of Geophysical Research*, *102*(B8), 18,057–18,070.

Weiss, J. R., Brooks, B. A., Foster, J. H., Bevis, M., Echalar, A., Caccamise, D., et al. (2016). Isolating active orogenic wedge deformation in the southern Subandes of Bolivia. *Journal of Geophysical Research: Solid Earth*, *121*, 6192–6218. https://doi.org/10.1002/2016JB013145

Williams, S. D., Bock, Y., Fang, P., Jamason, P., Nikolaidis, R. M., Prawirodirdjo, L., et al. (2004). Error analysis of continuous GPS position time series. *Journal of Geophysical Research*, *109*, B03412. https://doi.org/10.1029/2003JB002741