Transient Signal Detection
Using GPS Position Time Series

by

Kang Hyeun Ji

B.S. Earth Science Education
Seoul National University, 2002

Submitted to the Department of Earth, Atmospheric, and Planetary Sciences
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY
September 2011

Author.....
Department of Earth, Atmospheric, and Planetary Sciences
July 20, 2011

Certified by.....
Thomas A. Herring
Professor of Geophysics
Thesis Supervisor

Accepted by.....
Maria T. Zuber
E. A. Griswold Professor of Geophysics
Head, Department of Earth, Atmospheric, and Planetary Sciences

Transient Signal Detection Using GPS Position Time Series

by

Kang Hyeun Ji

Submitted to the Department of Earth, Atmospheric, and Planetary Sciences
on July 20, 2011 in partial fulfillment of the requirements for the degree of
Doctor of Philosophy

ABSTRACT

Continuously operating Global Positioning System (GPS) networks record station position changes with millimeter-level accuracy and have revealed transient deformations on various spatial and temporal scales. However, the transient deformation may not be easily identified from the position time series because of low signal-to-noise ratios (SNR), correlated noise in space and time and large number of sites in a network. As a systematic detection method, we use state estimation based on Kalman filtering and principal component analysis (PCA). State estimation improves the SNR in the time domain by estimating secular and transient motions and reducing the level of both white and colored noise. PCA improves the SNR in space domain by accounting for the coherence of transient motions between nearby sites. Synthetic tests show that the method is capable of detecting transient signals embedded in noisy data but complex signals (e.g., large-scale signals in space and time, multiple and/or propagating signals) are difficult to detect and interpret. We demonstrate the detection capability with two known signals in the Los Angeles basin, California: far-field coseismic offsets associated with the 1999 Hector Mine earthquake and locally-observed hydrologic deformation due to heavy rainfall in winter 2004–2005 in San Gabriel Valley. We applied the method to the daily GPS data from the Plate Boundary Observatory (PBO) network in Alaska and in the Washington State section of the Cascadia subduction zone. We have detected a transient signal whose maximum displacement is ~9 mm in the horizontal and ~11 mm in the vertical at Akutan volcano, Alaska, during the first half of 2008 and two previously unrecognized small slow slip events with average surface displacements less than 2 mm, which was thought to be below current GPS resolution. The detection method improves the SNR and therefore provides higher resolution for detecting weak transient signals, and it can be used as a routine monitoring system.

Thesis supervisor: Thomas A. Herring
Title: Professor of Geophysics

Acknowledgements

This thesis would not have been possible without many friends and colleagues who have supported and encouraged me for many years.

First of all, I thank my advisor Thomas Herring for many interesting discussions, continuous support and incredible patience. I thank my thesis committee, Bradford Hager, Carl Wunsch, and Brendan Meade for their insightful comments, interesting questions, and criticism. I would also like to thank Robert King and Simon McClusky for answering my questions about GPS processing.

I used the FAKENET code developed by Duncan Agnew in Chapter 2 and tremor data provided by Aaron Wech in Chapter 4. The thesis has been benefitted from the transient signal detection exercise supported by the Southern California Earthquake Center (SCEC). I would especially thank Rowena Lohman and Jessica Murray-Moraleda for their leadership of the SCEC exercise.

I thank Vicki McKenna, Carol Sprague, and Theresa Macloon for their administrative support and Linda Meinke for her technical support including many hardware and software problems.

My deep gratitude goes to my former and current officemates: Ping Wang, Andrea Llenos, Kyle Bradley, Grant Farmer, Nasruddin Nazerali, Junlun Li, Abdulaziz AlMuhaidib, Hussam Busfar, Yulia Agaramokova, Sedar Sahin, and Martina Coccia. I also thank other friends in EAPS: Chang Li, Chin-Wu Chen, Huajian Yao, Scott Burdick, Krystle Catalli, Einat Lev, Sudhish Bakku, Rosalee Lamm, Jiangning Lu, Ben Black, Frank Centinello, Achraf Koulali Idrissi, Noa Bechor, and Michael Floyd.

I have been fortunate to have great Korean friends: Sang-Heon (Dan) Shim, Sang-Jin Ryu, Dong-Chul Kim, Taeho Shin, Gunwoo Lee, Jun Woo Choi, Sang-Wan Kim, Sang-Won Byun, Won-Yong Lee, Sang-Il Lee, Sang-Jin Lyu, Yong-Sang Choi, Eunjee Lee, Min Sub Sim, and Jong-Mi Lee. I especially thank Byeongju Jung and Kyung-Chul Park for delicious Korean style BBQ, great wine, and exciting sea fishing.

I thank my undergraduate advisor Professor Byung-Doo Kwon who introduced me geophysics and taught me basic and firm background knowledge.

My family has always been supportive of my work. My endless love goes to all my family.

This study was supported by National Science Foundation grant EAR-0734947-03, National Aeronautics and Space Administration grant NNX009AK68G, and Southern California Earthquake Center grant EAR-052922.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Statements of the thesis

We use state estimation and principal component analysis (PCA) to detect anomalous transient signals from Global Positioning System (GPS) position time series. The method focuses on improvement of the signal-to-noise ratio (SNR) of GPS time series in both space and time. We demonstrate with applications to synthetic and real data sets the method is capable of detecting relatively weak transient signals that might not be apparent in the noisy raw data.

1.2 High-precision GPS geodesy

GPS is a global navigation satellite system (GNSS) that provides users with their location and time information. GPS was originally designed for military purposes, but it has become a useful tool for a variety of civilian applications from car navigation to

surveying. As a geodetic method, the precision of GPS is currently at the millimeter level and has benefited from, among other things, improvements in satellite and receiver/antenna technology, elaborate models for reducing errors, continuous GPS sites distributed globally and regionally, international collaborations such as the International GNSS service (IGS) [*Beutler et al.*, 1994a]. GPS is now routinely used for studying a wide range of surface deformation processes such as plate motions [e.g., *Kreemer et al.*, 2003; *Prawirodirdjo and Bock*, 2004], interseismic strain accumulation [e.g., *Ruegg et al.*, 2009; *Prawirodirdjo et al.*, 2010], earthquake deformation [e.g., *Reilinger et al.*, 2000; *Hudnut et al.*, 2002], dynamic ground motions [e.g., *Larson et al.*, 2003; *Bock et al.*, 2004], slow slip events in subduction zones [e.g., *Dragert et al.*, 2001; *Miller et al.*, 2002; *Ozawa et al.*, 2002], volcano deformation [e.g., *Cervelli et al.*, 2002; *Chang et al.*, 2007; *Freymueller and Kaufman*, 2010], hydrologic deformation [e.g., *King et al.*, 2007; *Tregoning et al.*, 2009].

With the positions of the satellites (e.g., from broadcast ephemeris or from more precise IGS orbits for geodetic applications), one can obtain the receiver position estimates from the pseudoranges or the carrier phase or both. The satellites transmit two codes (C/A code and P code) that modulate two carriers at frequencies 1575.42 MHz (L1) and 1227.60 MHz (L2) and L5 on some newer satellites. Pseudoranges can be obtained from measuring time differences between transmission and reception of the time-tagged signal. The term "pseudo" is used because of the satellite and receiver clock errors. Precise positioning can be done with the phase differences between the incoming carrier and a receiver-generated carrier. The phase measurements are more precise (millimeter level) than the coded pseudorange measurements (meter level). Details of

GPS principles can be found in books [e.g., *Parkinson and Spilker*, 1996; *Hofmann-Wellenhof et al.*, 2001; *Leick*, 2004].

High-precision geodetic GPS has been accomplished by the improvements of the system hardware and software. The main features of the high-precision GPS are:

(1) The current GPS constellation (31 active satellites in December 2010) ensures at least four or more satellites in view with good sky coverage and guarantees measurement anytime at all weather conditions and anywhere on surface or near surface, provided that there is no obstruction along the signal path.

(2) Geodetic GPS receivers acquire dual-frequency (L1 and L2) carrier phase data. The frequency-dependent ionospheric delay errors can be largely eliminated by combining dual frequency data [e.g., *Bock et al.*, 1986; *Dong and Bock*, 1989].

(3) Antennas are designed to suppress signal multipath (i.e., the interference between the direct and the reflected signal). Choke ring antennas (covered by radome) are often used for geodetic applications.

(4) Since the broadcast ephemeris is not precise enough for geodetic purposes, IGS provides precise GPS satellite orbits with the accuracy of ~ 2.5 cm. Details can be found at http://igscb.jps.nasa.gov/index.html.

(5) Double differencing (DD) and precise point positioning (PPP) are commonly used techniques for GPS data analysis. The DD eliminates satellite-specific errors by differencing data between receivers, and receiver-specific errors by differencing data between satellites [e.g., *Bock et al.*, 1986; *Schaffrin and Bock*, 1988]. Given precise orbits and clocks determined from a global network (e.g., IGS), the PPP estimates receiver-

specific parameters of individual sites and produces results similar in precision to the simultaneous analysis of all data [*Zumberge et al.*, 1997].

(6) Various types of errors can be modeled with additional parameters in GPS geodetic inversions. The models, mostly originated from earlier techniques such as very long baseline interferometry (VLBI) and satellite laser ranging (SLR), include GPS satellite orbits [e.g., *Beutler et al.*, 1994b; *Springer et al.*, 1999a], the satellite yaw attitude [e.g., *Bar-Sever*, 1996], solar radiation pressure forcing on the satellite [e.g., *Fliegel et al.*, 1992; *Springer et al.*, 1999b], higher-order ionospheric effect [e.g., *Bassiri and Hajj*, 1993; *Fritsche et al.*, 2005; *Hernandez-Pajares et al.*, 2007; *Hoque and Jakowski*, 2008; *Petrie et al.*, 2010], atmospheric delays [e.g., *Niell*, 1996; *Chen and Herring*, 1997; *Boehm et al.*, 2006b; *Boehm et al.*, 2006a; *Tregoning and Herring*, 2006], Earth rotation variations [e.g., *Yoder et al.*, 1981; *Barnes et al.*, 1983; *Lindqwister et al.*, 1992; *Herring and Dong*, 1994; *Mathews et al.*, 2002; *Gross et al.*, 2003], solid Earth and ocean tide loading [e.g., *Scherneck*, 1991; *Watson et al.*, 2006], atmospheric pressure loading [e.g., *van Dam et al.*, 1994; *Tregoning and van Dam*, 2005], antenna phase center variations [e.g., *Wu et al.*, 1993; *Schmid and Rothacher*, 2003; *Schmid et al.*, 2007], multipath [e.g., *Elosequi et al.*, 1995; *Park et al.*, 2004]. Model parameters that cannot be predicted accurately must be estimated from GPS (and other space-geodetic) data.

(7) High-precision geodetic software packages estimate model parameters (e.g., receiver coordinates, satellite orbits, tropospheric bias, receiver clock bias, earth orientation parameters, reference frame transformation parameters, etc.), edit data (e.g., cycle slips and outliers in the carrier phase observables) [*Blewitt*, 1990], and resolve integer ambiguity in the initial phase measurements (i.e., unknown initial number of

wavelengths) [*Blewitt*, 1989; *Dong and Bock*, 1989]. Three software packages are widely used for geophysical research purposes: BERNESE by Astronomical Institute, University of Bern, Switzerland [*Beutler et al.*, 2007]; GAMIT/GLOBK by Massachusetts Institute of Technology, USA [*Herring et al.*, 2009]; GIPSY/OASIS II by Jet Propulsion Laboratory, USA [*Webb and Zumberge*, 1997]. These packages are, to a large degree, automated so that even non-expert users can easily process GPS data.

(8) GPS modernization (e.g., new civilian signals L2C and L5 at 1176.45 MHz and planned launches of Block IIF and Block IIIA satellites) will further improve the precision and accuracy of 3-dimensional positioning.

In addition to the capability of high-precision long-term stable 3-dimensional positioning, GPS has replaced other geodetic techniques because GPS receivers and antennas are inexpensive, portable and do not require visibility between sites. Prior to the advent of GPS, VLBI and SLR were used to make deformation measurements for baselines greater than 100 km [*Herring*, 1999]. Both techniques have high precision but they require large facilities and have high costs so that the number of sites around the world is limited. For shorter baselines, conventional surveying methods (i.e., trilateration, triangulation, and spirit leveling) can be used but they require visibility between sites and suffer from atmospheric refraction errors. GPS can also be complementary to strainmeters that provide better sensitivity to strain but poor stability and poor spatial coverage, and Interferometric Synthetic Aperture Radar (InSAR) that provides extensive spatial coverage but poor temporal resolution.

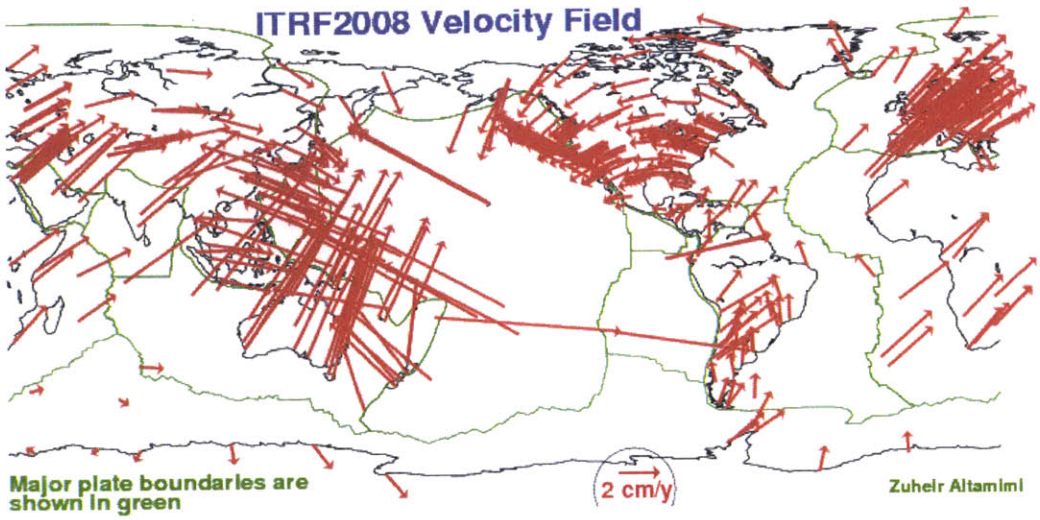


Figure 1-1. Global velocity field, relative to ITRF2008, measured at the stations of the IGS network (418 stations and 350 active stations as of January 2011). The global IGS network not only provides a global reference system such as ITRF, but also can be used to study the plate tectonics and a wide variety of deformation processes. Figure from http://itrf.ensg.ign.fr/ITRF_solutions/2008/

As the price of GPS receivers has decreased, GPS stations can be deployed with modest budgets, and continuously operating GPS networks established. The IGS network is a global network with over 400 GPS stations (Figure 1-1). The IGS network is used in geophysical studies such as the International Terrestrial Reference Frame (ITRF) [*Altamimi et al.*, 2007], Earth rotation variations, satellite orbit determination, atmosphere monitoring, changes in sea level and ice sheets, global plate motion (Figure 1-1), etc. A number of GPS stations are also installed around active deformation regions. Examples of such regional networks having over 1000 GPS stations are the Plate Boundary Observatory (PBO) in western US (Figure 1-2) and the GPS Earth Observatory Network (GEONET) in Japan (Figure 1-3). The regional GPS networks provide dense spatial and temporal sampling of the surface deformation field and allow continuous monitoring of a wide range of deformation processes with better precision and accuracy than was

previously possible by campaign-style GPS field works. Figure 1-4 shows a typical PBO station with a deep drilled braced monument, which is designed to support long-term stability of the monument.

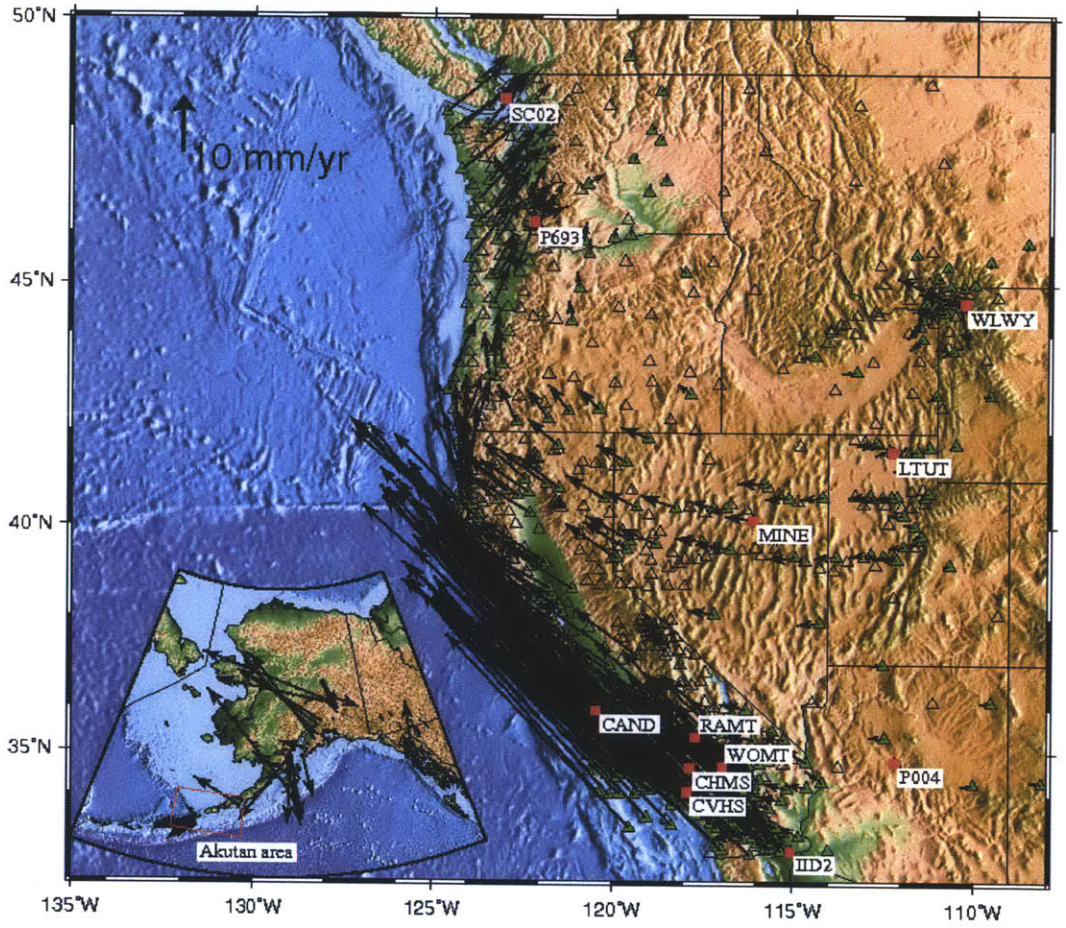


Figure 1-2. Horizontal velocities measured at the stations (green triangles) in the Plate Boundary Observatory (PBO) GPS network. The stations with open triangles were not used in the velocity calculation. The velocities are relative to the Stable North America Reference Frame (SNARF; for details, see the SNARF working group webpage at http://www.unavco.org/research_science/workinggroups_projects/snarf/snarf.htm). The velocity field shows north-west motion in California due to the relative motion between the North America plate and the Pacific plate and northeast motion in Cascadia due to the subduction coupling of the Juan de Fuca, Explorer and Gorda plates with the North American Plate. The motions of Alaskan sites (inset plot) is largely due to the subduction of the Pacific plate and clockwise rotating Bering plate [*Cross and Freymueller*, 2008]. The GPS data obtained from the labeled stations and stations in the Akutan area will be given in the later figures.

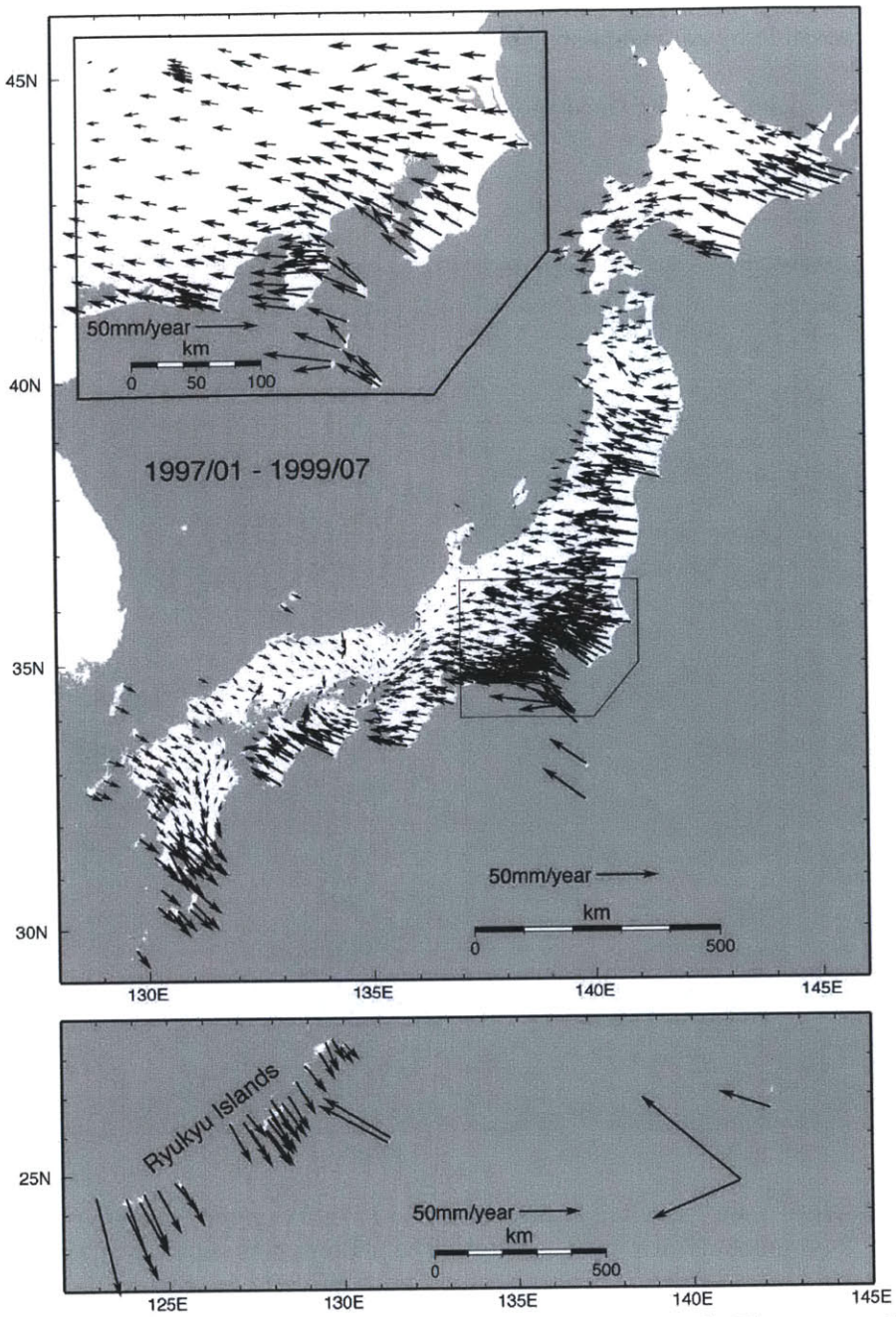


Figure 1-3. Horizontal velocities of the stations in the GPS Earth Observatory Network (GEONET) in Japan. The velocities are relative to the stable part of the Eurasian plate. Northwestward motion is primarily due to the subduction of the Philippine Sea plate and the Pacific plate beneath the Japanese Islands. The southwester Japan including Ryukyu island arc shows trench-ward motion maybe due to back-arc spreading associated with weak or no subduction coupling. Figure from *Sagiya* [2004].

Figure 1-4. Photo of the PBO station CAND installed at Parkfield, California, in May 1999 as part of the Southern California Integrated GPS Network (SCIGN) for monitoring strain accumulation across the San Andreas Fault and deformation associated with Parkfield earthquakes. The antenna with radome (left front) is mounted on the deep drilled braced monument that consists of 5 legs (one is a vertical leg braced the others) placed into drilled holes (~10 m below the surface) and welded together. The shed houses GPS and telemetry equipment (right back). The solar panel for power supply is shown on the roof of the shed and antennas for telemetry next to the shed. Photo from http://pboweb.unavco.org/shared/scripts/stations/?checkkey=CAND

1.3 GPS position time series

A GPS station in a continuous network measures surface motion and produces a time series of position changes in each coordinate direction (often north, east and up) relative to a reference frame. The temporal resolution is often one day, but higher sampling rates can also be obtained [e.g., *Larson et al.*, 2003; *Bock et al.*, 2004; *Langbein and Bock*, 2004; *Larson et al.*, 2007]. With rough repeatability at the level of 1 mm horizontal and 3 mm vertical, daily GPS position time series show various types of signals driven

tectonically or non-tectonically, as well as noise correlated in both space and time. The high precision and dense sampling enable the GPS position time series to be an important and primary resource for monitoring and modeling temporal and spatial variations of station positions in all phases of seismic, volcanic and hydrologic cycles. In this section, we show several types of signals frequently appeared in daily GPS time series and noise characteristics of GPS time series and discuss issues related to transient signal detection.

1.3.1 Signals in GPS position time series

Daily GPS position time series typically show secular velocities, repeating signals (often seasonal), offsets, outliers and data gaps (Figure 1-5). The secular velocities, particularly in horizontal components, are mostly due to tectonic plate motion and strain accumulation at the plate boundary (Figure 1-2). The estimates of linear rates can be biased if other types of signals are not accounted for [*Blewitt and Lavallee*, 2002; *Williams*, 2003b; a]. Repeating signals often have seasonal periods (i.e., annual and semiannual). The sources of the seasonal variations may include surface mass redistributions (e.g., ocean tide loading, atmospheric loading, non-tidal oceanic mass, groundwater loading, etc.), errors in models (e.g., satellite orbital models, atmospheric models, phase center variation models, etc.), bedrock thermal expansion, thermal noise of the antenna, multipath, snow cover on the antenna, etc. [*Dong et al.*, 2002]. There exist aliasing effects; short-period (diurnal and semi-diurnal) signals can alias into signals with seasonal periods [*Penna and Stewart*, 2003; *Stewart et al.*, 2005; *Penna et al.*, 2007]. It is unlikely that these signals can be accurately represented by perfect sinusoids with constant amplitude [*Bennett*, 2008]. GPS time series often have offsets due to changes of

equipment (e.g., receiver, antenna, monument, or radome; see Figure 1-5c), changes in processing strategy (e.g., models, elevation cutoff angle, reference frame, etc.), and earthquakes (see Figure 1-6a) [*Williams*, 2003a].

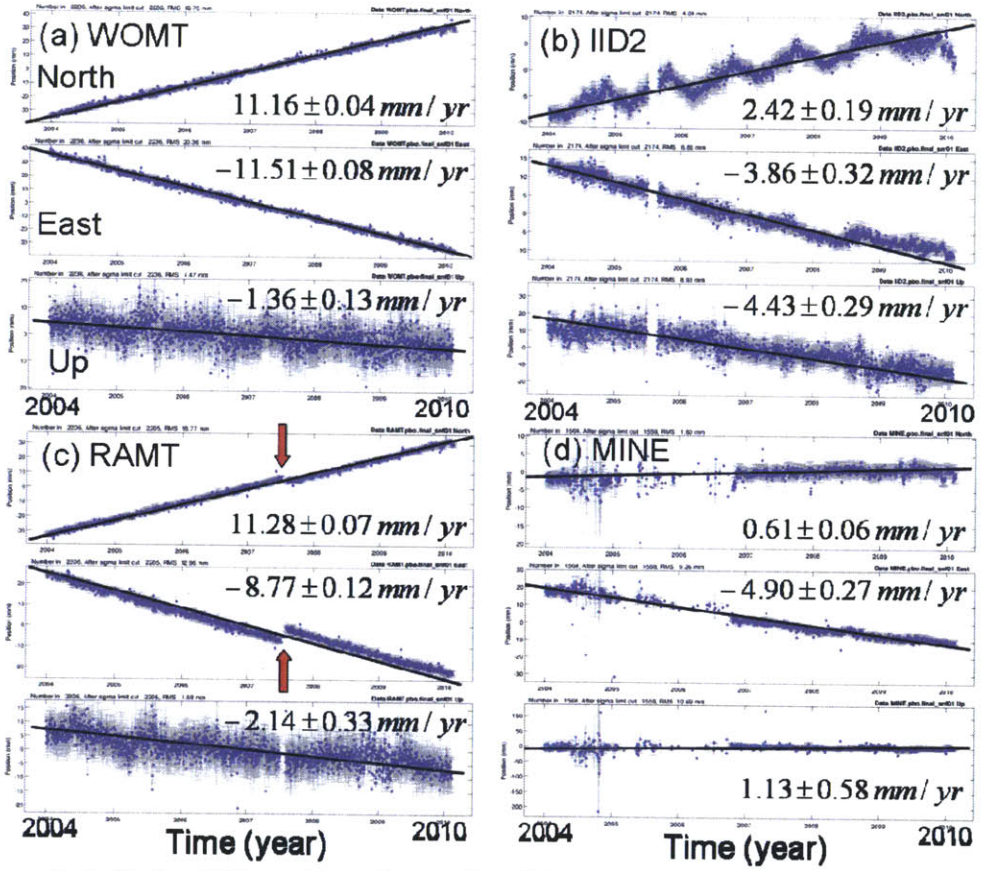


Figure 1-5. Daily GPS position time series (blue dots) with 1-sigma uncertainties or formal errors (gray error bars) measured at the stations (a) WOMT, (b) IID2, (c) RAMT, and (d) MINE (see Figure 1-2 for the site locations). As shown by the line fits, the most dominant signal is the secular velocity due to the North America plate motion particularly in horizontal components. The causes of the vertical velocities are less obvious. In addition, the time series often show repeating signals in (b), offsets indicated by red arrows in (c) and outliers and data gaps in (d). The offset in (c) was caused by antenna change. The uncertainties of the velocity estimates were computed assuming white noise.

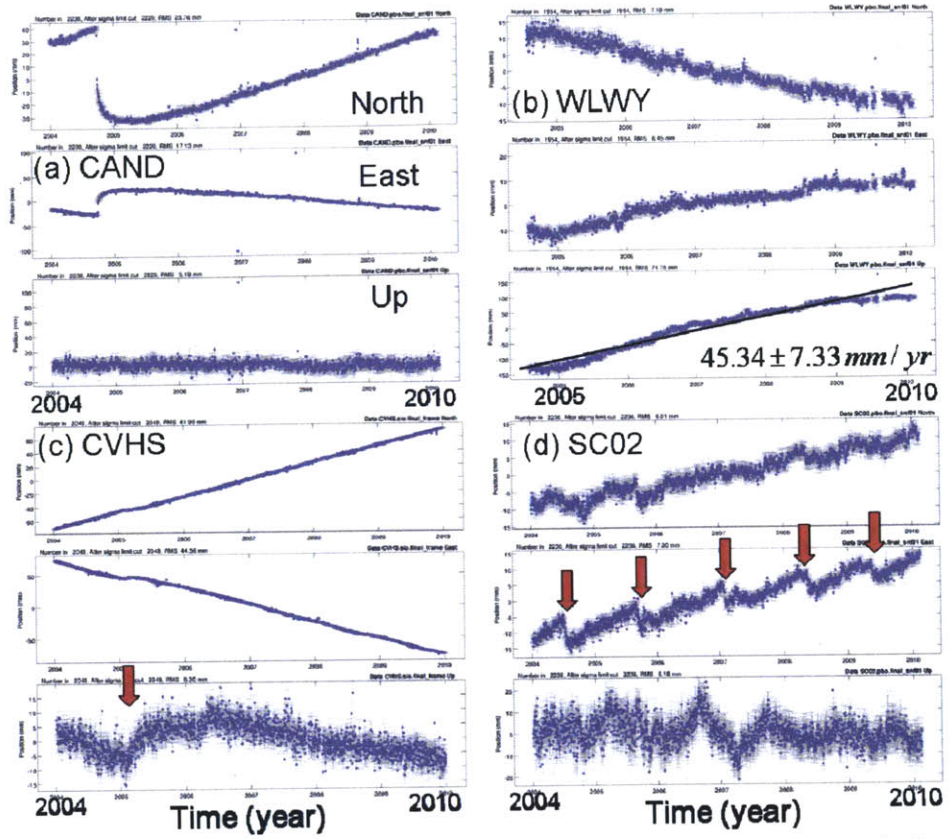


Figure 1-6. Daily GPS position time series (blue dots) with 1-sigma uncertainties (gray error bars) measured at the stations (a) CAND, (b) WLWY, (c) CVHS, and (d) SC02 (see Figure 1-2 for the site locations). The time series show well-known tectonic signals: (a) coseismic and postseismic deformation following the 2004 Mw 6.0 Parkfield, California, earthquake, dominant in horizontal components, (b) magmatic deformation (e.g., massive uplift of ~45 mm/year) at Yellowstone caldera, (c) hydrologic deformation in the San Gabriel valley, California, aquifer system due to heavy rainfall in winter 2004-2005 originally reported by *King et al.* [2007], and (d) episodic slow slip events within the Cascadia subduction zone, dominantly in east component, at about 14-month period.

In addition to the normally appearing signals, GPS time series may also contain transient signals from various tectonic sources. Figure 1-6 shows four examples of transient signals: (a) coseismic and postseismic deformation following the Parkfield earthquake [e.g., *Freed*, 2007; *Savage and Langbein*, 2008], (b) magmatic uplift at Yellowstone caldera, Wyoming [e.g., *Wicks et al.*, 2006; *Chang et al.*, 2007], (c) hydrologic deformation in the San Gabriel valley, California [*King et al.*, 2007], and (d)

the slow slip events within the Cascadia subduction zone, Washington [e.g., *Dragert et al.*, 2001; *Miller et al.*, 2002; *Szeliga et al.*, 2008]. These tectonic signals have also been recorded at multiple sites distributed in the region (see Figure 1-14 and Figure 1-15). Those signals are easy to detect from the time series because of their large magnitudes (i.e., high SNR).

When GPS time series are detrended, that is, the estimates of steady motions (e.g., secular velocities and annual and semiannual signals) are removed, the residual time series may reveal transient signals not apparent in the raw time series. Figure 1-7 shows changes in the rate of uplift of WLWY at Yellowstone caldera and the horizontal deformation associated with the vertical signal of CVHS at San Gabriel valley aquifer. These features are not easily identified in the raw time series (Figure 1-6b and c). In other words, removing steady signals is critical for the detection of small transient signals.

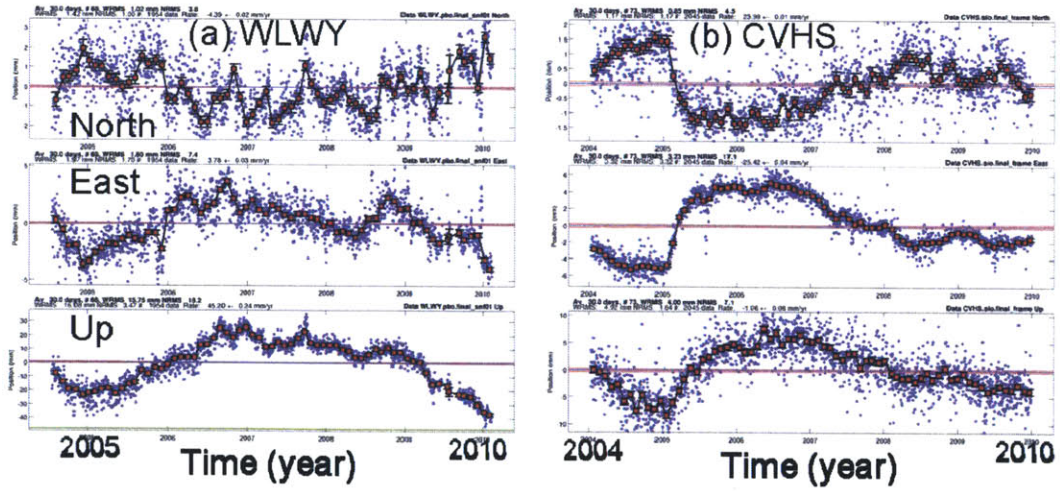


Figure 1-7. Residual time series (blue dots) of the stations (a) WLWY and (b) CVHS (see Figure 1-2 for the site locations and Figure 1-6 for the daily GPS position time series). The residual time series were obtained by removing the estimates of secular velocities and annual and semiannual sinusoids from the daily time series. Also shown are 30-day averages (red squares) with 1-sigma uncertainties (black error bars). The residual time series show features of position changes that are not clear from the daily time series; the rate of uplift at WLWY has decreased since 2009 and the horizontal components of CVHS also contain the hydrologic signal associated with the vertical.

GPS time series can also be affected by non-tectonic signals. For example, Figure 1-8 shows (a) the effects of antenna malfunction, (b) changes in vegetation near the antenna, (c) snow cover on the antenna in winter season, and (d) possible water vapor effect in summer season. These signals are mostly seasonal. Figure 1-9 shows three photos of the growth of a tree near the station CHMS (Figure 1-8b). The antenna malfunction and tree effects can be removed by changing antennas and trimming nearby trees, respectively. The snow effects often appear at high-latitude (e.g., Alaskan sites) and high-altitude sites (e.g., those on top of mountains).

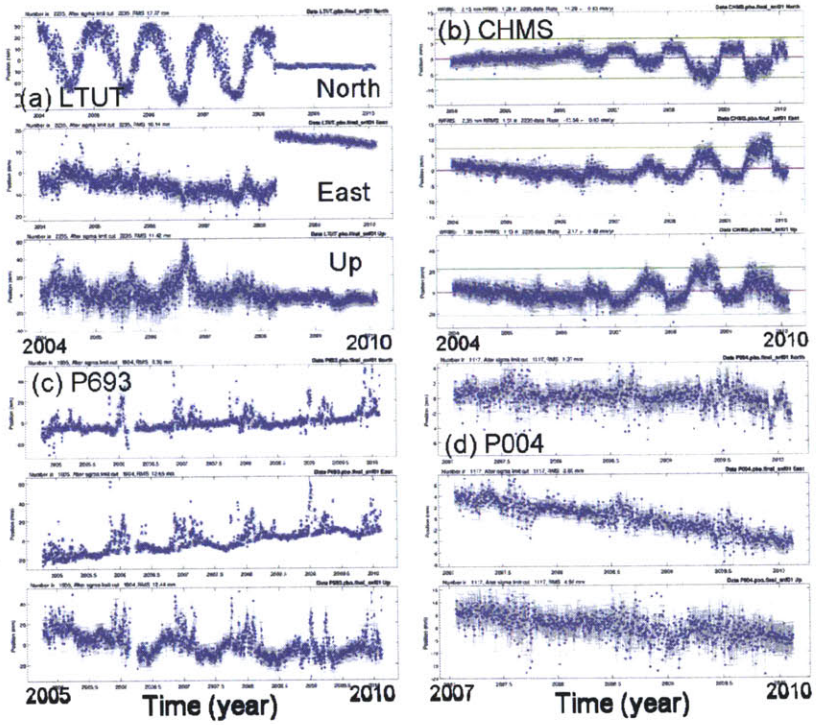


Figure 1-8. Daily GPS position time series (blue dots) with 1-sigma uncertainties (gray error bars) measured at the stations (a) LTUT, (b) CHMS with secular velocities removed, (c) P693, and (d) P004 (see Figure 1-2 for the site locations). The time series show non-tectonic (mostly seasonal) signals: (a) antenna malfunction, (b) effect of a tree growing near the antenna, (c) snow cover on the antenna in winter season, and (d) possible water vapor effect in summer season. After changing antenna of LTUT, the massive repeating signal in north component disappeared. The repeating and growing signal at CHMS is correlated with seasonal cycle and the growth of the tree (see Figure 1-9). The station P693 is located at Mount St. Helens at an altitude of ~2100 meters.

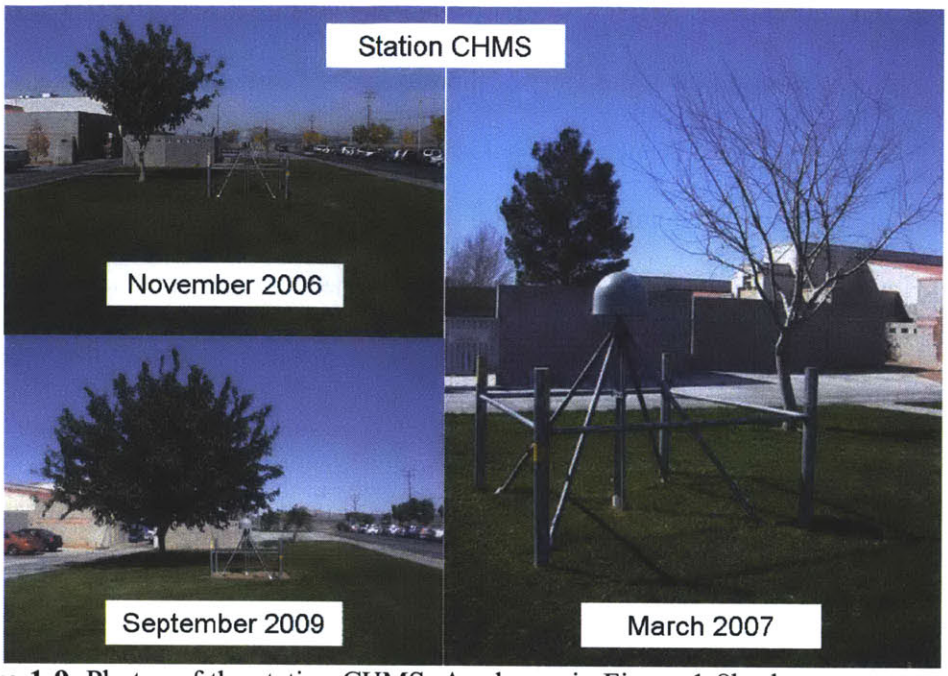


Figure 1-9. Photos of the station CHMS. As shown in Figure 1-8b, the measurements at this site are affected by the seasonal cycle and the growth of a nearby tree. Photo from http://pboweb.unavco.org/shared/scripts/stations/?checkkey=CHMS&sec=photos&activity=maintenance

1.3.2 Noise in GPS position time series

Noise in GPS position time series is correlated in time [e.g., *Zhang et al.*, 1997; *Mao et al.*, 1999; *Williams et al.*, 2004; *Langbein*, 2008] and space [e.g., *Wdowinski et al.*, 1997; *Dong et al.*, 2006] rather than simply independent observations. If correlated noise is neglected in the least squares estimation of site velocity (i.e., only with white noise assumed), the velocity uncertainties are unrealistically small [*Johnson and Agnew*, 1995; *Williams*, 2003b]. For transient signal detection, small transient signals cannot be simply discriminated from correlated noise. For example, Figure 1-10a shows a small curvature in the east component of the station WOMT. The curvature represents the temporal pattern of the postseismic relaxation observed in the far field because the station WOMT is located at ~86 km northwest of the epicenter of the Landers earthquake [e.g., *Bock et*

al., 1993] and ~60 km west of the epicenter of the Hector Mine earthquake [e.g., *Hudnut et al.*, 2002]. Only through the proximity of the site to these earthquakes and much larger signals at sites nearer the epicenters is it possible to recognize this as a transient because of small amplitude and temporal correlation in GPS time series. Figure 1-10b shows two distinct position changes at the station IID2: ~6 mm in north during 2009 and ~ 4 mm in east during 2008. However, we do not have information yet to determine if those signals are transient signals or correlated noise. For the proper interpretation of the estimated site velocities and/or detected transient signals, it is critical to understand the stochastic nature of GPS position time series.

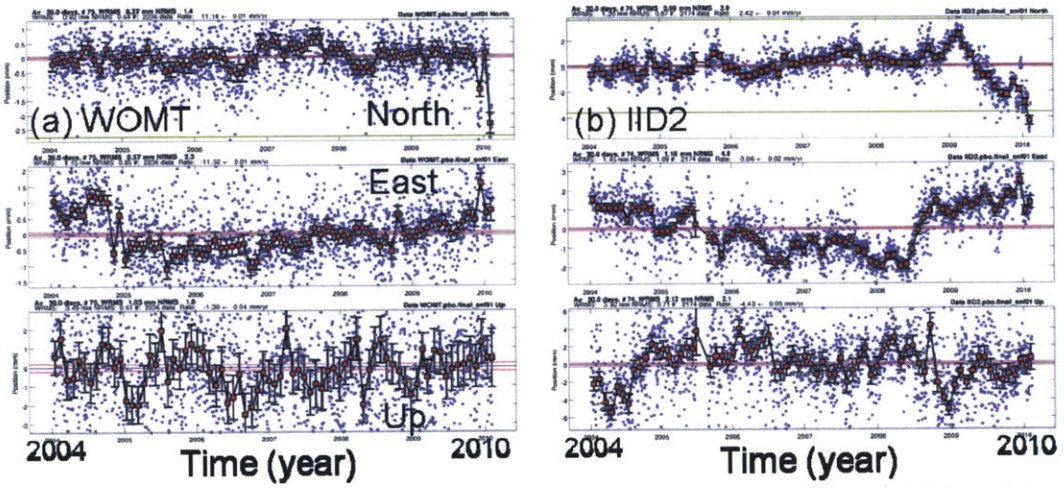


Figure 1-10. Residual time series (blue dots) of the stations (a) WOMT and (b) IID2 (see Figure 1-2 for the site locations and Figure 1-5 for the daily GPS position time series). Also shown are 30-day averages (red squares) with 1-sigma uncertainties (black error bars). The residual time series show temporal correlation rather than white noise. (a) The east component of WOMT exhibits a long-term curvature that possibly results from the postseismic deformation following the 1992 Mw 7.3 Landers earthquake and the 1999 Mw 7.1 Hector Mine earthquake. The station WOMT is one of Mojave Desert sites ~86 km northwest of the Landers epicenter and ~60 km west of the Hector Mine epicenter. (b) The horizontal components of IID2 show two distinct position changes compared to the background temporal correlation: ~6 mm in the north component during 2009 and ~ 4 mm in the east component during 2008. The vertical component does not show any distinct position changes. The two position changes seem to be never or weakly related because of non-overlapped duration. To determine if the two position changes are coherent transient signals, other sites around this site are required to examine.

Although sources of correlated noise and their contribution to daily position data have not been clearly understood, a variety of sources can corrupt the data with different spatial and temporal scales. Possible sources can be grouped into those that are common to sites in a network and those that are site-specific. The former includes satellite orbit errors, reference frame errors (e.g., Earth orientation) and atmospheric delay errors [*Wdowinski et al.*, 1997; *Williams et al.*, 2004; *Dong et al.*, 2006], and these errors are responsible for spatially correlated noise over the network. Site-specific errors includes errors correlated with observation geometry between a receiver and satellites (e.g., satellite sky coverage and track orientation), multipath, antenna phase center variations, and instability of geodetic monuments. The constellation geometry can produce different noise levels among the three components (i.e., highest level in vertical components and lowest level in north components [e.g., *Mao et al.*, 1999]). Random motion of geodetic monuments has been identified as a random-walk process (see below) from other geodetic measurements (e.g., strainmeter data [*Wyatt*, 1982; *Wyatt*, 1989], and two-color electronic distance measurement (EDM) data [*Langbein and Johnson*, 1997; *Langbein*, 2004]), complicating the detection of tectonic transient signals.

As with many physical phenomena, noise in geodetic data has been modeled in the frequency domain as a power law process with power spectral density (PSD) of the form $f^{-\kappa}$ where f is frequency and κ is the spectral index [*Mandelbrot and Van Ness*, 1968; *Mandelbrot*, 1983]. For large κ, the process becomes smoother and has more power at lower frequencies [*Zhang et al.*, 1997]. Geophysical phenomena often approximate processes with $1 < \kappa < 3$ [*Agnew*, 1992] termed "fractal random walk" including classical Brownian motion ("random walk") with $\kappa = 2$. These processes are

nonstationary. Stationary processes have $-1 < \kappa < 1$ termed "fractal white noise" including classical white noise with $\kappa = 0$ (i.e., flat spectrum independent of frequency). Processes with $\kappa = 1$, called flicker noise, are commonly observed in various dynamical processes such as sunspot variability, the wobble of the Earth about its axis, undersea currents, and uncertainties in time measured by atomic clocks [*Williams et al.*, 2004].

Noise in GPS time series has been estimated by several approaches such as autocorrelation analysis, spectral analysis, and maximum likelihood estimation (MLE) [e.g., *Zhang et al.*, 1997; *Mao et al.*, 1999; *Williams et al.*, 2004; *Langbein*, 2008]. For example, *Williams et al.* [2004] found in global solutions by MLE analysis that white plus flicker noise can best describe the noise characteristics. In regional solutions, the spectral index is more varied than in the global solutions, possibly reflecting a mixture of local effects (e.g., differences in atmospheric effects from region to region, different types of geodetic monuments, and other local conditions). Furthermore, there are differences in noise content between different analysis centers and different analysis strategies. *Langbein* [2008] found, from 236 GPS sites with data range between 3.5 and 10 years measured in Southern California and Southern Nevada, that either flicker or random-walk noise can describe about half of the variance of the time series. However, there is no dominant noise model (see Table 1-1); the remaining time series can be best categorized as either a combination of flicker and random-walk; power law noise; first-order Gauss-Markov plus random-walk noise; or power law plus broadband, seasonal noise. From all of the time series examined, *Langbein* [2008] provided nominal amplitudes of flicker and random-walk noise (see Table 1-2) for direct use of the noise models without intensive estimation of model parameters.

Table 1-1. Best noise model for SCIGN and SBAR Data[a]

Noise model[b]	Percentage of All Sites		
	North	East	Vertical
FL	32	27	41
RW	21	35	18
PL	11	11	6
FLRW	20	16	21
FOGMRW	8	3	9
BPPL	8	9	5

[a]Table from *Langbein* [2008]; SCIGN: Southern California Integrated GPS Network; SBAR: Southern part of the Basin and Range Geodetic Network

[b]FL: flicker noise, RW: random-walk noise, PL: power law noise, FLRW: a combination of flicker and random-walk noise, FOGMRW: a combination of first-order Gauss-Markov plus random-walk noise, and BPPL: a combination of band-pass filtered and power law noise

Table 1-2. Nominal values for flicker and random-walk model of noise[a]

Component	Parameter	Units	Median	Inter-Quartile Range
North	White noise	mm	0.55	0.44 0.76
	Flicker noise	$mm/a^{0.25}$	0.9	0.4 1.6
	Random-walk noise	$mm/a^{0.5}$	0.5	0.0 1.4
East	White noise	mm	0.67	0.55 0.86
	Flicker noise	$mm/a^{0.25}$	0.9	0.3 1.6
	Random-walk noise	$mm/a^{0.5}$	0.7	0.0 1.4
Up	White noise	mm	2.35	2.02 2.90
	Flicker noise	$mm/a^{0.25}$	3.6	2.3 5.9
	Random-walk noise	$mm/a^{0.5}$	1.5	0.0 2.9

[a]Table from *Langbein* [2008]

Motion of geodetic monuments with respect to the deeper crust has been described by a random-walk process [*Johnson and Agnew*, 1995]. The monument instability can be caused by processes such as precipitation loading, tidal strains, and weathering effects on rock and soil [*Johnson and Agnew*, 1995; *Langbein and Johnson*, 1997]. To reduce the monument motion or mitigate the weathering effect on surface, monuments are deeply (~10 m) anchored, laterally braced, and isolated from the surface [*Wyatt*, 1989]. The deep drilled braced monuments are found to be more stable than other types of monuments [*Williams et al.*, 2004] so that they have been used for continuous GPS networks such as the Southern California Integrated GPS Network (SCIGN) and PBO (see Figures 1-4 and 1-9). *Langbein* [2008] argued that the major factor of the monument stability is the location of the GPS site; sites located in the dry deserts (e.g., Basin and Range, Nevada) with low erosion rates and low humidity will perform better than those located near areas of active groundwater or oil pumping (e.g., Los Angeles basin, California).

Analyses of GPS data yield uncertainties of the position estimates (gray error bars in Figure 1-5). These are often called "formal errors" that are normally calculated from noise in the phase data sampled at a given rate assuming white noise in the phase measurements. In GPS time series analysis, formal errors are often used to represent approximate white noise contribution to the variation of the position. However, temporal correlation in the phase data raises the scale issue in formal errors, depending on statistics and sampling rate used in the GPS geodetic inversion. We re-estimate formal errors to represent white noise in GPS time series in Appendices A and H with different approaches.

Spatially correlated noise is often referred to as "common mode error" (CME) because multiple sites in a regional network exhibit similar temporal patterns (Figure 1-11). The spatial correlation could be due to orbital, reference frame and large-scale atmospheric errors [*Wdowinski et al.*, 1997; *Williams et al.*, 2004; *Dong et al.*, 2006] that commonly affect the sites in a network. The definition of CME is still an open question in terms of its spatial extent. Figure 1-12 shows correlation between residual time series from two global solutions as a function of station separation. The correlation gradually decreases from about 0.8 at around 0.1° (10 km) to 0.5 at around 10° (1000 km) and drops sharply to zero at around 40°. Depending on the size of the network, the CME may or may not be considered as uniform.

The CME can be mitigated by several approaches. Assuming spatial uniformity of the CME, the regional filtering proposed by *Wdowinski et al.* [1997] estimates the CME by computing a (weighted) mean from the position residuals of selected reference sites for each coordinate and at each epoch. A spatiotemporal filtering approach, PCA and Karhunen-Loeve expansion proposed by *Dong et al.* [2006], estimates the CME allowing for non-uniform spatial distribution (Figure 1-11). Although not being appropriate for larger networks (e.g., the entire PBO network), the spatial uniformity of the CME works well for regional networks such as the Southern California Integrated GPS Network (SCIGN) [*Dong et al.*, 2006], implying that the network size is smaller than the CME wavelength. Another approach to the reduction of the CME is a reference frame transformation (i.e., daily Helmert transformation) to a regional reference frame defined by the positions and velocities of a set of stable sites selected for the transformation (see Chapter 3).

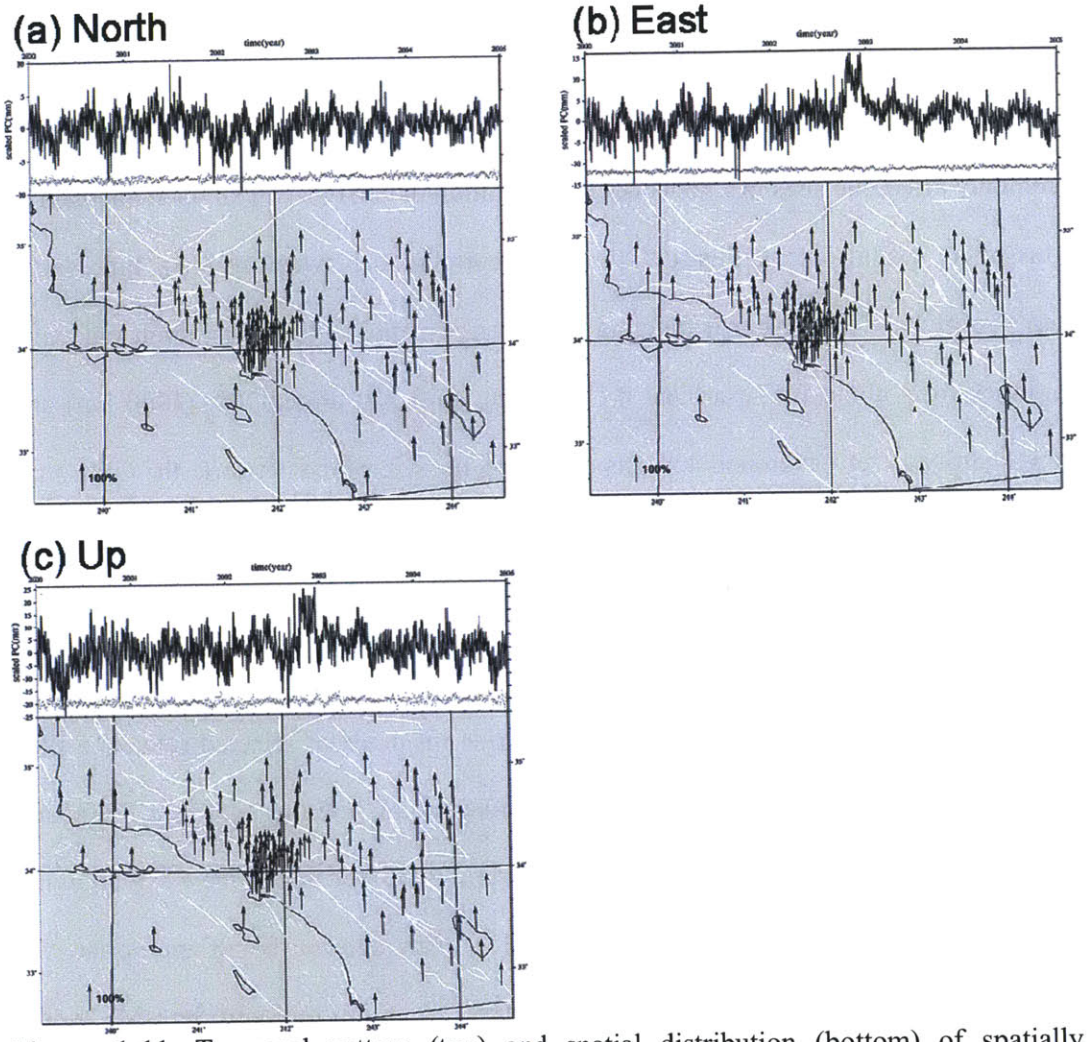


Figure 1-11. Temporal pattern (top) and spatial distribution (bottom) of spatially correlated noise contained in the measurements from the Southern California Integrated GPS Network (SCIGN) for the period 2000 to 2004. The spatiotemporal pattern was obtained by principal component analysis of the residual time series. The vertical arrows indicate the scaled amplitude of the temporal pattern in each component at each site. The scale is shown lower left of the figure. The temporal patterns are correlated in time. The amplitude distribution looks uniform in space, which may indicate that the SCIGN network is smaller than the CME wavelength. Figures from *Dong et al.* [2006].

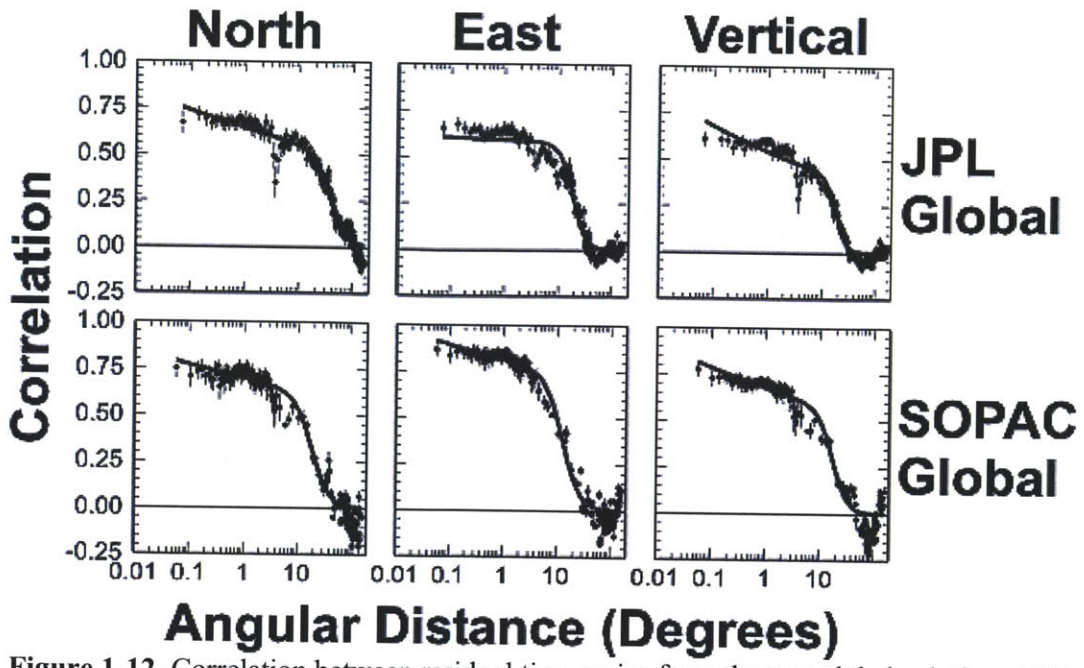


Figure 1-12. Correlation between residual time series from the two global solutions (JPL: Jet Propulsion Laboratory and SOPAC: Scripps Orbit and Permanent Array Center) as a function of station separation in degrees. Individual points are the weighted mean (and its 3-sigma error) of 100 individual correlation estimates. Weighting is based on the number of points in the series used to calculate the correlation. Solid line is an arbitrary fit to the data that assumes the correlation tends to zero at large distances and incorporates a break in slope at an unspecified angle. Figure from *Williams et al.* [2004].

1.3.3 Spatial coherence of transient signals

Transient tectonic signals can often be discriminated from temporally correlated noise or non-tectonic signals because tectonic signals often occur coherently in space. Here spatial coherence means that multiple sites experience a temporal pattern simultaneously (see Figure 1-13). Spatial coherence can help determine if a temporal pattern is a tectonic transient signal. For example, the long-term curvature shown in Figure 1-10a was determined as a temporal pattern due to postseismic relaxation because it was recorded in multiple sites, with similar pattern but different amplitudes, in the Mojave Desert and in Nevada [*Freed et al.*, 2007] (see Figure 1-13a).

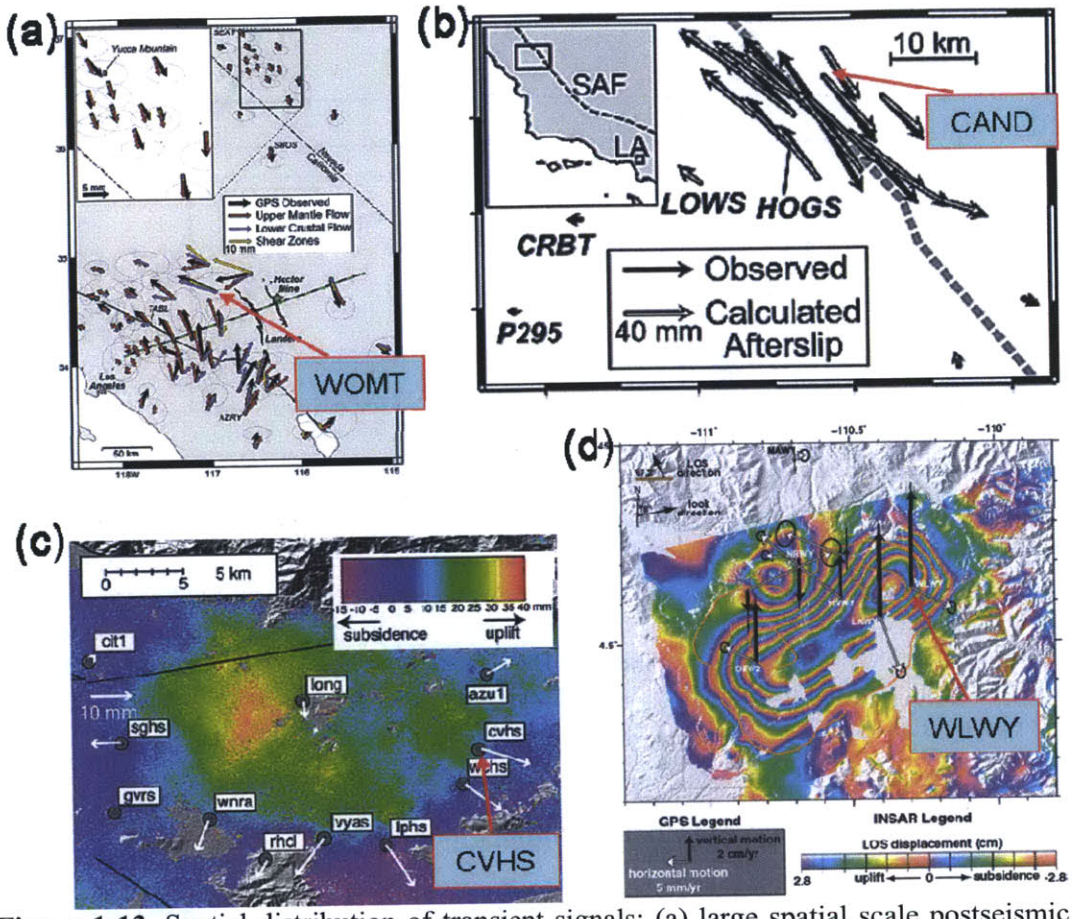


Figure 1-13. Spatial distribution of transient signals: (a) large spatial scale postseismic deformation following the 1999 Hector Mine earthquake, (b) postseismic deformation following the 2004 Parkfield earthquake, (c) hydrologic deformation in the San Gabriel valley, and (d) magmatic deformation at Yellowstone caldera. The spatial patterns are systematic and coherent. The earthquake-related deformation patterns are strike-slip due to (a) deep mantle flow and (b) afterslip. The fluid-related deformation patterns (with color fringes from InSAR measurements) indicate inflations due to (c) recharging of aquifers after heavy rainfall in winter 2004–2005 and (d) magma intrusion beneath the caldera. The raw time series and residuals of the four stations in the box can be found in Figures 1-5, 1-6, 1-7 and 1-10. Figures from (a) *Freed et al.* [2007], (b) *Freed* [2007], (c) *King et al.* [2007], and (d) *Chang et al.* [2007].

Transient signal detection is difficult with a single station. As an example, Figure 1-10b shows that the station IID2 exhibits two distinct position changes in the residual time series: ~6 mm in the north component during 2009 and ~ 4 mm in the east component during 2008. The residual time series of P500 and P509, the two closest stations to IID2, do not clearly reveal the corresponding position changes (Figure 1-14a and b). Furthermore, the differences of the residuals between P500 and IID2, between P509 and IID2, and between P500 and P509 confirm that the two position changes are specific at IID2 (Figure 1-14c~e). If the position changes are driven tectonically, the spatial extent of the coherent signal would be smaller than the shortest separation (~ 25 km between P500 and IID2) or IID2 is located near the boundary of the deformed area (this cannot be tested because IID2 is located at the edge of the PBO array (see Figure 1-2)). Other types of measurements, if available, would be useful for signal identification.

It is interesting to see that the nearby stations P500 and P509 show similar residual patterns in north and vertical components (Figure 1-14a and b). One might regard the patterns as coherent tectonic signals because two sites are only ~22 km apart and experience the patterns at the same time. However, such a quick decision could be wrong because of spatially correlated noise driven non-tectonically.

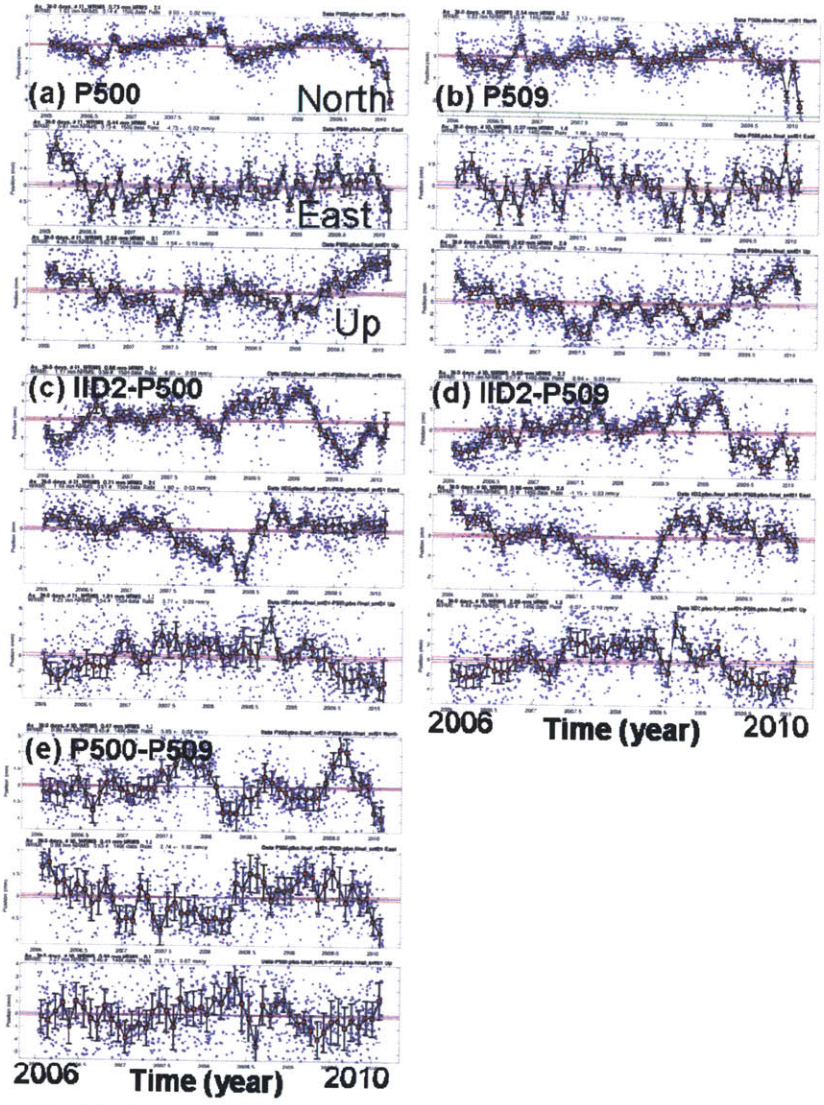


Figure 1-14. Residual time series of the stations (a) P500 and (b) P509 and differences of the residual time series (c) between P500 and IID2, (d) between P509 and IID2, and (e) between P500 and P509. Also shown are 30-day averages (red squares) with 1-sigma uncertainties (black error bars). The stations P500 and P509, the two closest stations to the station IID2, are located ~25 km west and ~ 32 km northwest of the station IID2, respectively. The residual time series do not clearly show the position changes associated with the two distinct position changes shown in Figure 1-10b. The differences between IID2 and the nearby stations show that the identified position changes were not cancelled so that they are specific at the station IID2, while the differences between P500 and P509 do not exhibit the position changes associated with those of IID2. Note that the stations P500 and P509 (~22 km apart along the north-south direction approximately) exhibit similar residual patterns particularly in north and vertical components. It is also not easy to determine if the temporal patterns are related to a coherent tectonic deformation or spatially correlated noise.

1.4 Transient signal detection

GPS time series provide various kinds of transient signals driven tectonically or non-tectonically and also contain noise that can be correlated in space and time. Among various signals, tectonically-driven transient signals are of interest for crustal deformation studies. In this section, we discuss why transient signal detection is not easy, how a transient signal is defined, and how a transient signal can be detected.

Transient signal detection is challenging for a number of reasons. Firstly, the signals are often weak and part of the algorithm development here is to obtain reliable detections of the smallest signals. Large signals such as those due to big earthquakes can be easily detected directly from actual data, but weaker signals may go unnoticed. For example, Figure 1-15 shows a set of (detrended) time series from the area of Akutan Island, Alaska. As shown in Chapter 3, the data set actually contains a transient signal whose maximum horizontal displacement is ~ 9 mm which developed over a period of ~6 months, but the signal is not apparent in the time series. Secondly, GPS time series contain correlated noise. When the SNR is low, one could incorrectly identify such correlated noise as a transient signal. The Akutan data set also suffers from the correlated noise (Figure 1-15). Thirdly, models for transient signals are not generally available. Parametric estimation can often be too restrictive (i.e., the form of the transient signal must be assumed). Fourthly, GPS measurements suffer from outliers and missing data. Lastly, large GPS networks make it time-consuming to visually inspect all GPS time series. For these reasons, a desirable detection algorithm improves the SNR, accounts for space-time correlation structure, allows nonparametric and/or stochastic approaches, and is reliable and efficient.

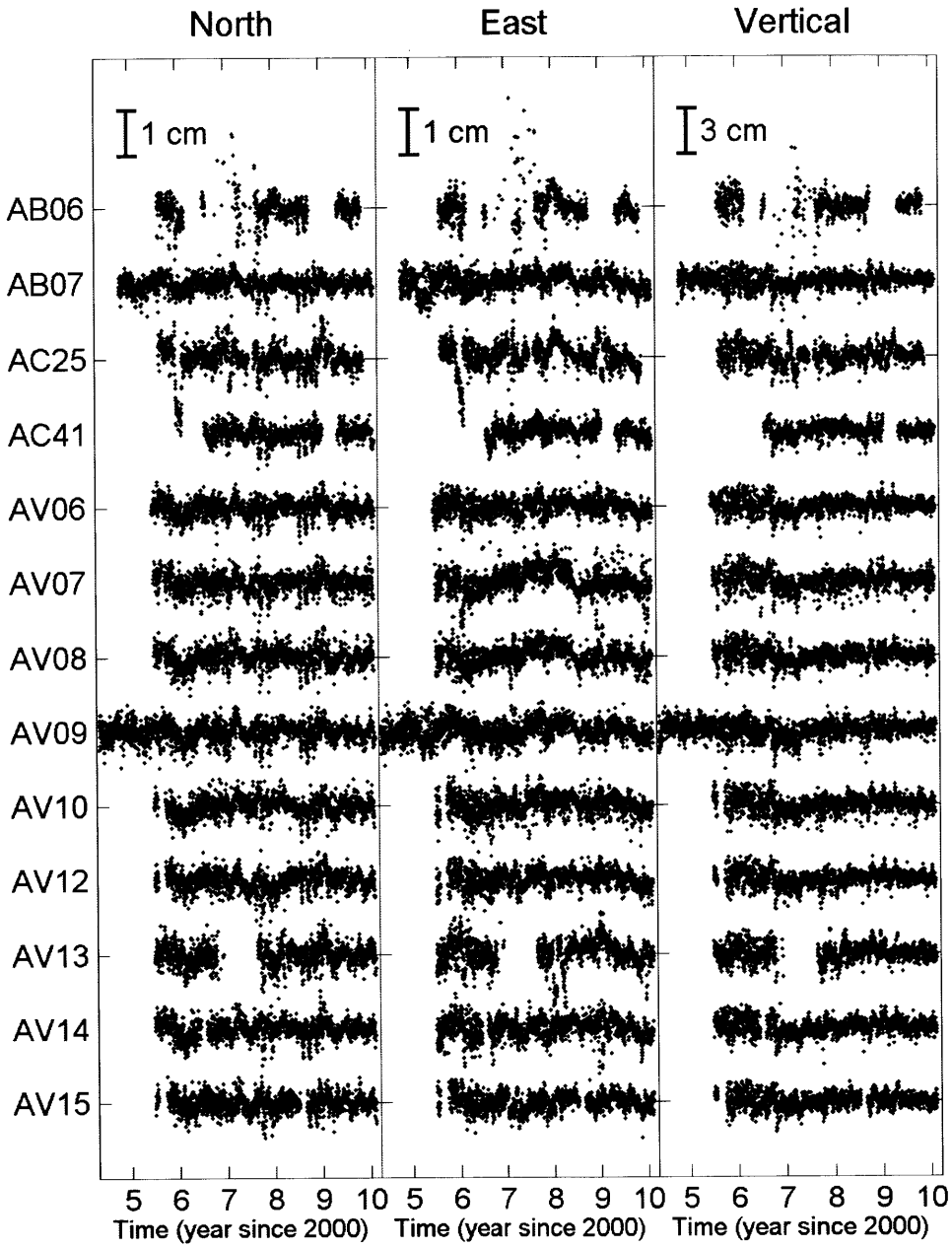


Figure 1-15. Residual time series (dots) of 13 stations in the area of Akutan Island, Alaska (see Figure 1-2 for the location of Akutan Island). Vertical axis indicates station IDs with scales in upper left for each component and horizontal axis denotes time in year since 2000. Formal uncertainties are not shown for clarity. A common noisy pattern can be seen in north component around late 2008, suggesting spatially correlated noise. When viewing page edge-on, more common patterns can be found in each component. These time series actually contain the transient signal we discuss in Chapter 3, but spatial and temporal extent of the signal cannot be easily identified from the time series because of low signal-to-noise ratio and correlated noise in space and time.

A transient signal we want to detect is defined by the following characteristics. Firstly, the temporal pattern is different from steady motions (e.g., secular velocities and periodic signals) making it a "transient" signal. However, the term "steady" depends on a given data span. If the Yellowstone data are given only for 2009, the uplift can be removed as a steady motion (see Figure 1-7a). Secondly, the temporal pattern is also different from an assumed stochastic noise model. In this study, a first-order Gauss-Markov (FOGM) process is selected to describe the temporal correlation because it can be easily implemented in Kalman filters and because when multiple processes are added the spectrum of the resulting power can mimic the spectrums of many other noise models (but not all) [*Langbein*, 2008] (see Appendix A for more details). Thirdly, a transient signal is coherent in space (see Figure 1-13), which can distinguish tectonically-driven signals from non-tectonic ones. However, a problem may arise when a signal is not coherent, for example, slow slip events that are propagating in space and time (Figure 1-16). Finally, a transient signal is localized in space so that it can be distinguished from spatially correlated noise represented by the uniformly distributed CME (see Figure 1-11). However, the locality of transient signals and uniformity of the CME are network-dependent; a transient signal possibly extends over a small network and the CME is regarded as a function of the wavelengths of various systematic errors [*Dong et al.*, 2006].

Note from Figure 1-10a that the postseismic deformation signal in the east component at WOMT is weak relative to random noise. The 30-day averages better represent the curvature pattern than the residual time series. Improving SNR by reducing the noise level (averaging in this case) is an important and primary process for detecting transient signals with noisy measurements.

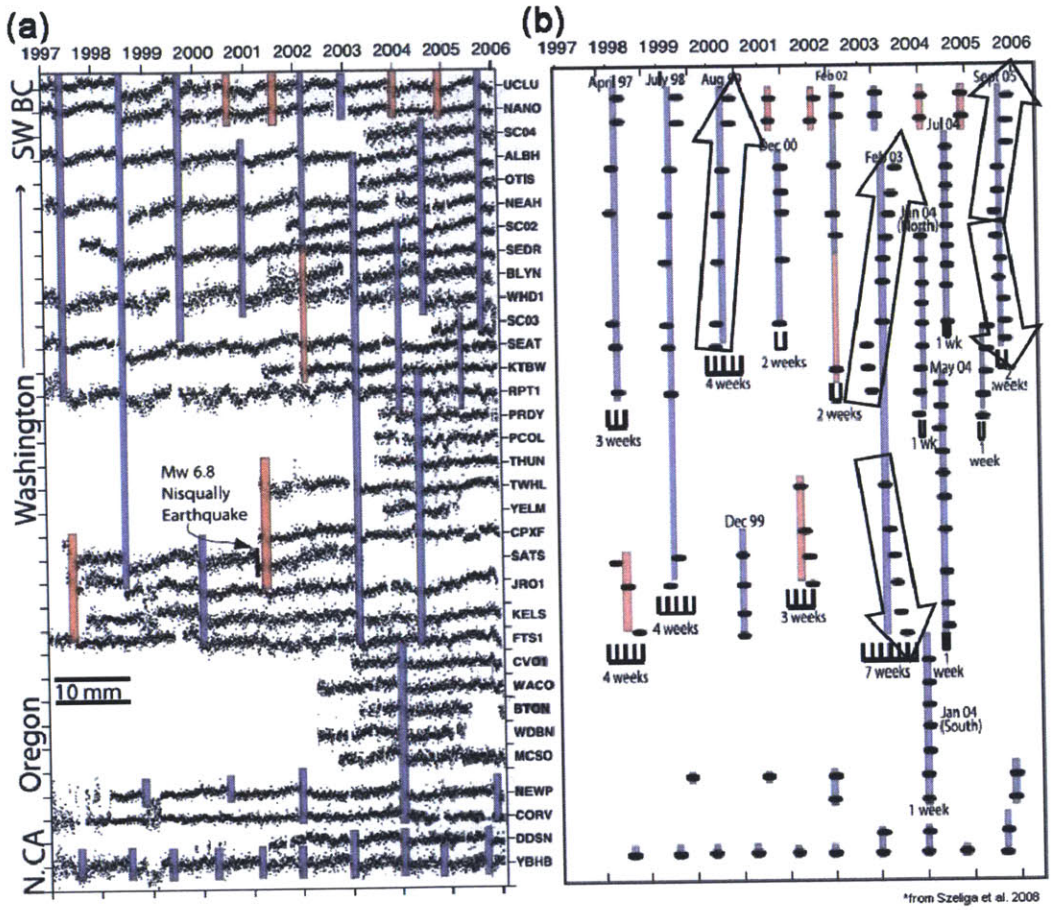


Figure 1-16. Slow slip events measured at GPS sites in the Cascadia subduction zone. (a) Time series of the east component stacked in an increasing order of latitude in vertical. The vertical bars indicate slip events well recorded by GPS and confirmed by tremor (blue) and less than four stations and not confirmed by tremor (red). (b) Temporal evolution and spatial propagation of slow slip events. Most events last 3–4 weeks, propagate at average rates of 10–15 km/day both unidirectionally and bidirectionally as indicated by arrows. Typical durations are 3–4 weeks having the shortest of 1 week and the longest of 7 weeks. N. CA: northern California; SW BC: southwestern British Columbia. Figure from *Szeliga et al.* [2008].

As a transient signal detector, we propose a spatiotemporal filtering approach by combining state estimation in the Kalman filter formulations and principal component analysis (PCA). State estimation improves the SNR in the time domain by reducing the random noise level and PCA improves the SNR in the space domain by accounting for the spatial coherence of transient signals. The enhanced SNR allows us to detect transient signals that are not apparent in data.

A smoother reduces white noise, estimates a state vector and its uncertainty, and interpolates missing data. A state vector typically includes a secular rate, annual and semiannual sinusoids, offsets due to known receiver/antenna changes, transient signals if parameterization is possible, and a FOGM process. The FOGM process represents temporally correlated noise in the time series and, if any, unmodeled or mismodeled transient signals. Figure 1-17 shows state estimation results from the station CAND. The residuals and the FOGM estimates do not show any long-term curvature so that the log parameterization with 1-day decay time well represents postseismic deformation following the 2004 Parkfield earthquake. State estimation interpolates missing data whose uncertainties tend to be larger than those of nearby existing data (Figure 1-17c). The larger interpolation uncertainties allow us to assess the impact of interpolation on the signal detection by PCA.

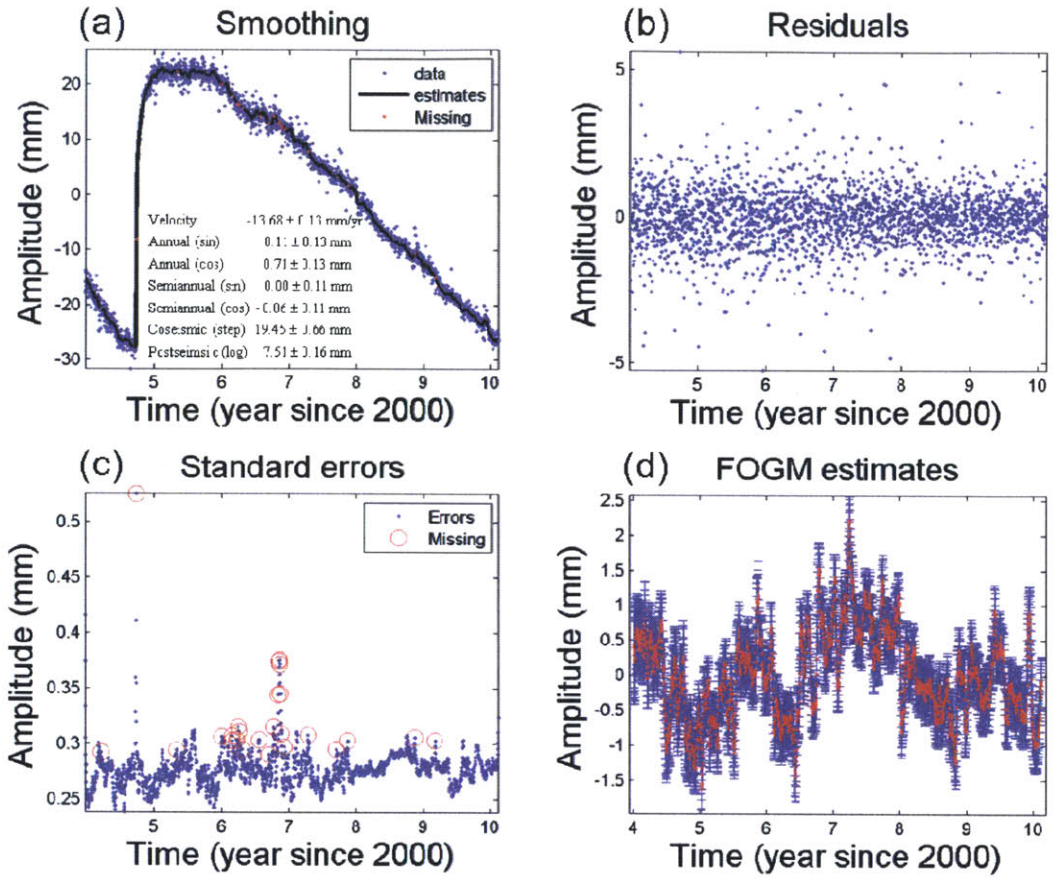


Figure 1-17. State estimation results from the station CAND (see Figure 1-2 for the site locations and Figure 1-6 for the daily GPS position time series). (a) State estimates (black curve) of daily GPS time series (blue dots). The state vector includes an initial offset, a secular rate, annual and semiannual sinusoids, coseismic offsets and postseismic log parameterization with 1-day decay time associated with the 2004 Parkfield earthquake, and a first-order Gauss-Markov (FOGM) process. Missing data are interpolated (red dots). (b) Residuals after removing the state estimates. Residuals look white with the scatter decreasing in time. (c) Standard errors of the state estimates (blue dots). Results that have been interpolated due to missing data are enclosed by red circles. The interpolation uncertainties tend to be larger than those of nearby existing data. (d) FOGM estimates (red dots) and their uncertainties (blue error bars). The FOGM estimates represent remaining temporal correlation.

PCA decomposes the FOGM estimates and any modeled transient signals of interest into principal components (PCs) for the temporal variations and sample eigenvectors for the spatial distribution of the PCs. Coherent signals often, depending on the SNR, appear in a few higher-ranked PCs. The significance of the PCA results can be tested by chi-square statistics, lag differences (i.e., differences between points in a PC at a given time lag), and visual inspection if a signal is systematic, coherent, and localized.

To show how PCA is used for detecting a transient signal, a set of synthetic time series were generated by adding an arctangent signal at 10 out of 100 FOGM time series. The sample eigenvectors show no local group of significant elements and the PCs exhibit oscillations (Figure 1-18). Such oscillations whose frequencies increase with the ranking of PCs are due to temporal correlation in the time series projecting into the PCs. After removing the oscillations from the PCs, lag differences of the residual PCs show peaks for lags of 10 to 90 around time 200 ~ 400 (Figure 1-19). If we choose the data time interval between 200 and 400, a significant local group of the first sample eigenvector elements appears around index 15 and the first PC looks similar to the signal (Figure 1-20). One can further reduce the data dimension in space for further enhancing the SNR. Details can be found in Appendix D.

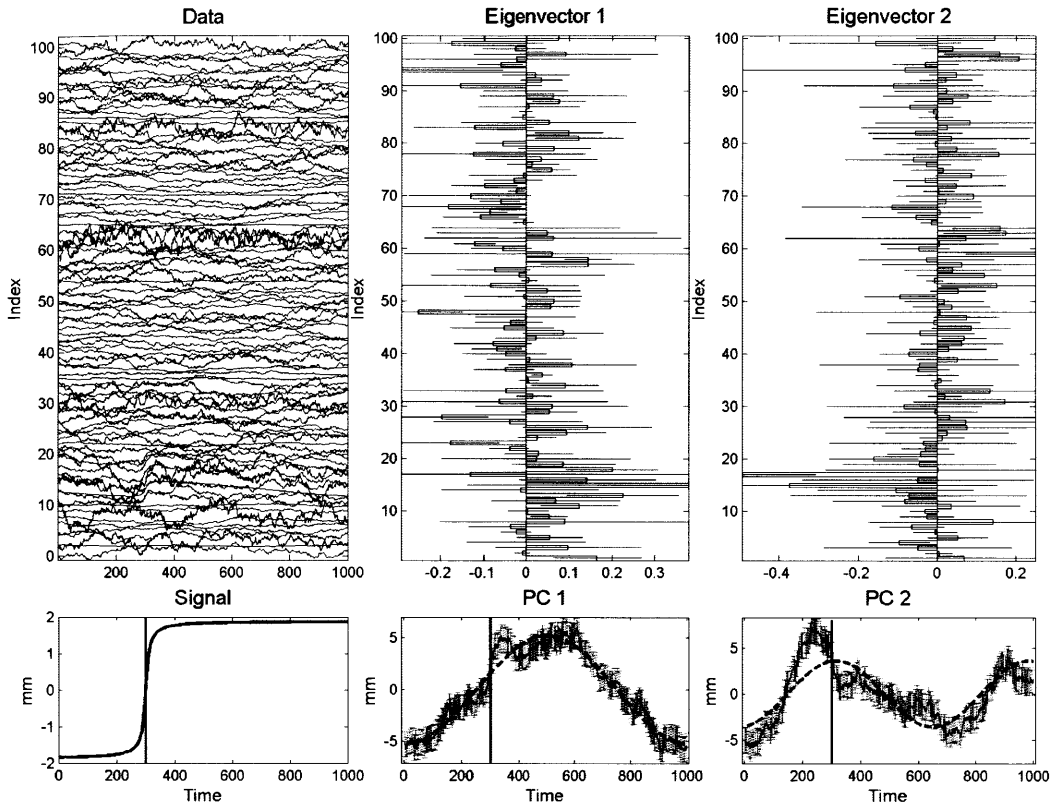


Figure 1-18. PCA results from synthetic data that were generated by adding a signal, $1.2 \times \arctan[(t - t_0)/10]$ where time t_0 is 300 (vertical lines), to 10 consecutive time series out of the 100 FOGM samples with uniformly distributed correlation times (1 ~2000 days) and variances (0 ~ 5 mm^2). The average FOGM sample variance is 0.75 ± 0.83 mm^2 with minimum of 0.01 mm^2 and maximum of 4.89 mm^2. The signal power is 2.63 mm^2 and the average noise variance of the 10 components is 1.53 ± 1.49 mm^2; the signal is not apparent in the original data. Linear trends were removed before PCA. Data time series are shifted by 1 mm so that the vertical axis indicates amplitudes as well as data indices. All errors represent 1-sigma uncertainties. There is no significant eigenvector element (horizontal bars) with respect to 3-sigma uncertainty and sinusoidal functions (dashed curves) give a good fit to the PCs, suggesting that the PCs are mostly affected by time-correlated noise. Slight misfits between time 200 and 400 can be found, particularly in the second PC, but it is not clear at this stage if the misfits are driven by a signal. Analyses of this class of problem are given in Appendix D.

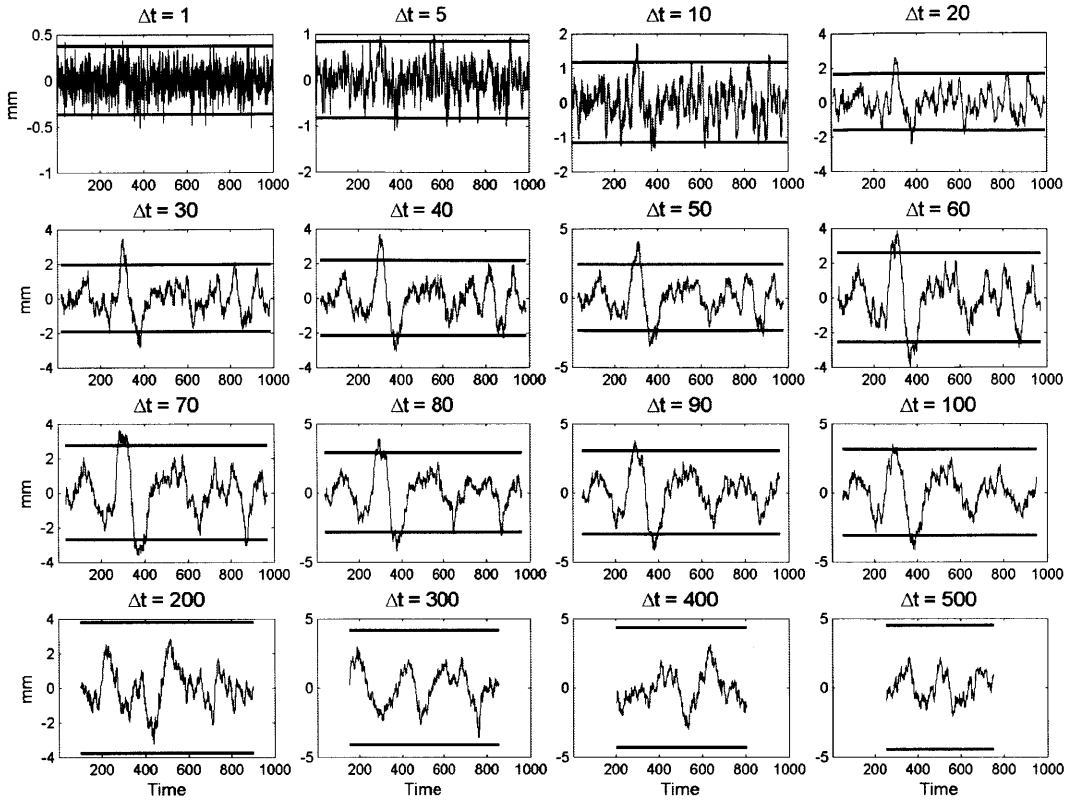


Figure 1-19. Lag (Δt) differences of the residual principal component (PC) from the sinusoidal fit to the first PC shown in Figure 1-18. The horizontal lines show the 3-sigma uncertainties of the lag differences obtained from the first-order Gauss-Markov variance estimates. For $10 \sim 90$ lags, peaks larger than the threshold appear around time $200 \sim 400$, which indicates that a transient event may exist for the time interval.

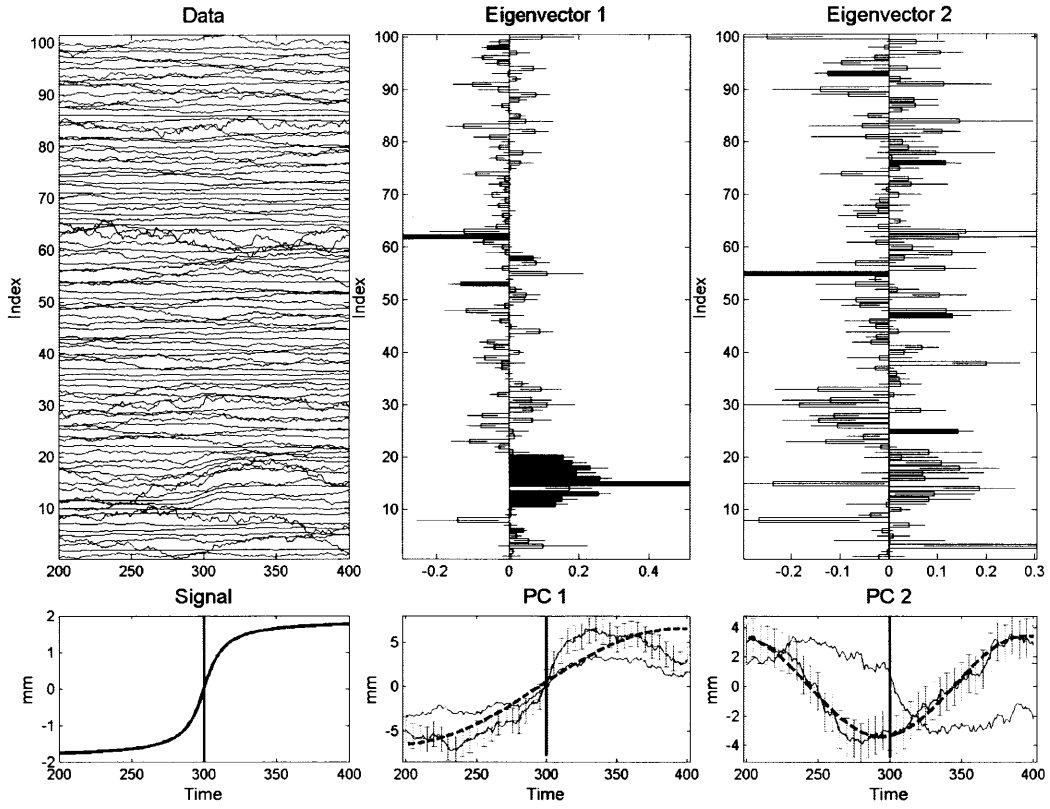


Figure 1-20. Same as in Figure 1-18 but, as identified by the lag differences in Figure 1-19, the time interval between 200 and 400 was selected for PCA. The first sample eigenvector shows a local group of significant elements around index 15 where the locality means the closeness of data indices in the one-dimensional case. The first principal component (PC) looks similar to the signal. The sinusoidal fit (dashed curve) is not good for the first PC ($\chi^2_{dof} = 1.8$). The second PC seems to be driven by noise based on the random spatial pattern and a good sinusoidal fit ($\chi^2_{dof} = 0.1$). The true signal is located at indices from 11 to 20. Index 14 of the first sample eigenvector is not significant. When PCA used data only with indices from 11 to 20, the index 14 became significant (see Appendix D). The curves without error bars are the PCs shown in Figure 1-18. Both PCs exhibit the signal (the sign of the second PC is flipped) so that the first PC is weaker than the first PC from data between 200 and 400.

1.5 Thesis outline

In Chapter 2, state estimation and PCA are applied to the synthetic data sets from the transient detection exercise, supported by the Southern California Earthquake Center (SCEC) since 2008 [*Murray-Moraleda and Lohman*, 2010]. In the exercise, participants have applied and tested their detection methodologies to synthetic GPS time series that, in some cases, contained signals due to transient fault slip but were contaminated with temporally and spatially correlated noise. Techniques developed by participants successfully identified large slip signals in the data, but they are still incapable of detecting more subtle signals. Further details on the exercise can be found at http://www.scec.org/research/projects/TransientDetection/. In the chapter, we discuss the detection capability of the method based on our results from the exercise.

In Chapter 3, the detection algorithm is applied to the PBO daily GPS position time series measured in Alaska. We have detected a transient signal that occurred during the first half of 2008 at Akutan volcano, ~1224 km southwest of Anchorage, in the Aleutian arc. The signal can be modeled as a volcanic inflation event. The signal cannot be easily identified in the original position time series (see Figure 1-15) because of the low SNR and highly correlated noise. We demonstrate that the state estimation and PCA is capable of enhancing the SNR of weak transient signals from a noisy and large data set. A short version of this chapter can be found in *Ji and Herring* [2011].

In Chapter 4, the detection algorithm is applied to the PBO daily GPS position data measured in the Washington State section of the Cascadia subduction zone. We have detected two previously unrecognized small slow slip events (SSEs) independently of tremor information. The spatial patterns of the small SSEs are similar to well-known

major SSEs (see Figure 1-16) but in a smaller area where tremors are also observed. The average displacements are less than 2 mm, which was thought below GPS resolution.

The remaining part of the thesis includes thesis conclusions and future work in Chapter 5 and Appendices: improving signal-to-noise ratio of GPS position time series by state estimation (Appendix A), state estimation: a filter/smoother approach (Appendix B), grid search for FOGM parameters (Appendix C), detecting transient signals using GPS measurements by PCA (Appendix D) which includes applications with real data in Los Angeles basin, California, formulations and properties of PCA (Appendix E), properties of sample eigenvalues and eigenvectors from independent Gaussian variables (Appendix F), issues related to Monte Carlo simulation for PCA statistics (Appendix G), and formal error scaling for realistic error statistics (Appendix H).

Chapter 2

Synthetic tests in the SCEC transient detection exercise

Abstract

This chapter discusses the strengths and weaknesses of the detection algorithm based on the results from the synthetic data sets used in the Southern California Earthquake Center (SCEC) transient detection exercise. The algorithm improves signal-to-noise ratios (SNR) by removing spatially correlated noise or common mode error (CME) by reference frame transformation, estimating temporally correlated noise by state estimation, and detecting spatially coherent transient signals by principal component analysis (PCA). Uncertainties of PCA results allow us to infer the significance of transient signals. However, the algorithm is not capable of detecting weak and complex signals such as those distributed broadly in space and time and those recorded by small number of stations. Large-scale signals are difficult to separate from CME in space and/or steady motions in time. Large numbers of stations, particularly in the near field, will increase possibility of detection even when a signal is weak. Multiple and/or propagating signals can be detected if SNR is high, but they appear in multiple principal components, making interpretation complicated. The more complex signals expected in the future phases of the SCEC exercise will allow us to evaluate and improve the detection capability of the algorithm.

2.1. Introduction

The Southern California Earthquake Center (SCEC; http://www.scec.org/) is a community of researchers at over 60 institutions worldwide, headquartered at the University of Southern California and funded by the National Science Foundation and the U.S. Geological Survey. SCEC has made efforts to understand earthquakes and mitigate earthquake risk in Southern California and elsewhere. The broad objective of SCEC's Tectonic Geodesy disciplinary activities is to make a variety of geodetic data collected in Southern California available and to use the measurements for addressing problems on deformation processes affecting the region. One of the topics of interest for fulfilling the objective is to "develop a geodetic network processing system that will detect anomalous strain transients" [*Murray-Moraleda and Lohman*, 2010].

SCEC has supported an ongoing research project, SCEC transient detection exercise. As of January 2011, three phases of testing have been completed since 2009. All Phases used synthetic Global Positioning System (GPS) daily position time series and Phase III included a set of real GPS data measured in Southern California. In some cases, synthetic data contained signals due to transient fault slip but were contaminated with noise. Participants applied their detection methods to the test data sets and reported back on any transient signals they detected. Eleven groups participated in the recent Phase III exercise. The detection techniques generally fall into several categories such as principal component analysis (PCA), Kalman filtering, or strain rate mapping. Most algorithms, including ours, identified large signals but did not detect subtle and complex signals [*Murray-Moraleda and Lohman*, 2010]. Details are available at the detection exercise webpage (http://www.scec.org/research/projects/TransientDetection/).

In the SCEC exercise, researchers can evaluate and improve their own method because data characteristics are known ex post facto. In this chapter, we present our detection procedure and discuss related issues based on the test results. We will not consider Phase I results here because in this phase the signals are large and obvious even in the time series. Analyses of the test results have led to improvements in the detection method and we discuss reprocessed results here.

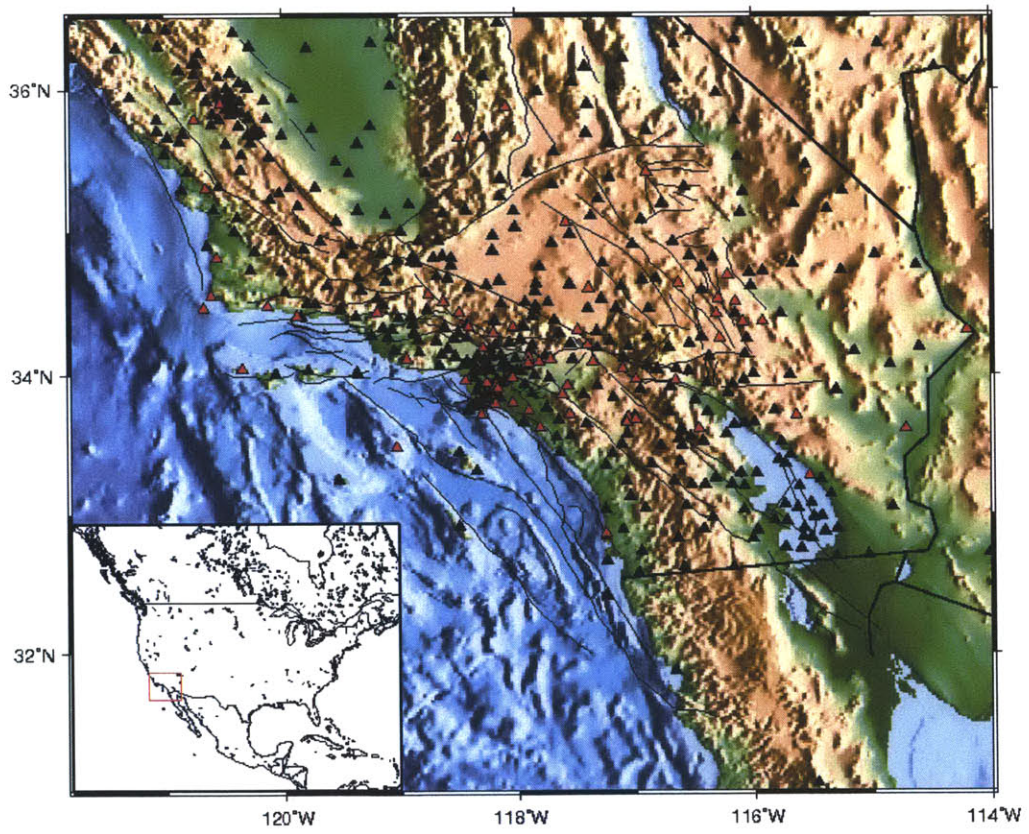


Figure 2-1. Total 492 GPS sites (triangles) in Southern California used in the SCEC detection exercise. The light black lines are faults from the SCEC Community Fault model version 3.0.5 and the thick black lines are borders between US states and between US and Mexico. Also shown are the 64 sites (red triangles) used as reference frame sites for stabilizing reference frame in Phase IIA Set 002307. The reference frame sites were selected as those having velocity uncertainty less than a given threshold and distributed, to some extent, uniformly over the network.

2.2. SCEC synthetic data sets

The synthetic data sets were generated by the FAKENET code [*Agnew*, 2009]. The code simulates real GPS position time series, including secular motion from the Scripps Orbit and Permanent Array Center (http://sopac.ucsd.edu/) velocities, missing data, random noise (i.e., white, random-walk, and flicker noise), common mode errors, seasonal noise with distance-dependent amplitude, and transient signals due to slip on faults from the SCEC Community Fault Model (http://structure.harvard.edu/cfm/). Transient signals may not be apparent in the time series depending on the assigned amplitudes of noise and signals.

Phase II consisted of three groups (labeled A, B and C), each of which has six data sets. Each data set contains synthetic daily position time series from 438 stations located in Southern California (latitude 31.0°N ~ 36.5°N and longitude 121.8°W ~ 114.0°W) and time interval between 1 January 2000 and 1 January 2009. The minimum and maximum sample sizes are 5 and 3242, respectively, and the average sample size is about 2042 ± 964. Missing data and maximum missing interval are, on average, ~11% and ~5%, respectively. Phase III had 7 data sets from 492 stations in the same region (Figure 2-1) and one year longer time span up to 1 January 2010. The minimum and maximum sample sizes are 5 and 3604, respectively, and the average sample size is about 2171 ± 1083. Missing data and maximum missing interval are, on average, ~11% and ~4%, respectively. One of the 7 data sets is a real one.

Most of the signal types are thrust and strike-slip faulting with or without propagation. Phase III data sets contain multiple signals which are more subtle and complex including large-scale flow and aquifer signal. Table 2-1 summarizes the

characteristics of signals with signal time interval and maximum horizontal and vertical displacements.

Table 2-1. Information of signals included in the test data sets[a]

Phase	Set	Signal (Detection)[b]	Area	Duration (year)[c]	Max. Disp. (mm)[d]		Type[e]	# of Figures
					H	V		
IIA	001253	Y(N)	Oakridge	1.31	3.71	5.65	T, P	
	002227	N	-	-	-	-	-	
	002307	Y(N)	Temecula	6.63	4.62	1.85	S	1, 4, 5
	003841	N	-	-	-	-	-	
	003923	Y(Y)	San Gabriel	3.61	98.09	148.77	T, P	2, 3
	010157	Y(Y)	Compton	2.27	20.64	18.10	T, P	15
		Y(Y)	Compton	4.01	5.06	9.50	T	
IIB	003638	N	-	-	-	-	-	
	003756	Y(N)	San Joaquin	0.02	4.99	4.50	T	13
	010016	N	-	-	-	-	-	
	010056	Y(Y)	Whittier	0.12	23.34	37.11	T	8, 18, 19
	010731	Y(N)	San Gabriel	0.94	8.26	19.21	T, P	
		Y(Y)	San Gabriel	0.92	10.14	6.66	T, P	
	011017	Y(N)	Coachella	5.11	5.89	0.95	S, P	
IIC	000643	Y(Y)	Mill Creek	2.37	25.27	28.18	T	
	001213	N	-	-	-	-	-	
	001417	Y(N)	Anza	6.36	5.57	1.54	S, P	
	003717	N	-	-	-	-	-	6(a)
	004608	Y(Y)	Anza	4.53	9.07	1.53	S, P	17
	010935	N	-	-	-	-	-	
III	Set A	Y(Y)	Santa Monica	1.46	4.58	7.53	T	6(b), 7, 12
		Y(N)	Elsinore fault	5.49	1.52	0.03	S	
	Set B	Real	-	-	-	-	-	
	Set C	Y(N)	Large-scale flow	8.71	7.86	0.04	F	14
		Y(Y)	Garnet Hill	1.32	10.22	18.99	T, P	16(a), 21
		Y(N)	San Bernardino	0.41	4.97	5.13	A	16(b)
	Set D	Y(Y)	Santa Monica	1.46	4.77	8.03	T	
		Y(N)	Elsinore fault	0.11	1.53	0.04	S	
	Set E	Y(N)	San Andreas	4.01	2.09	0.30	S, P	
		Y(Y)	San Andreas	1.16	4.46	0.66	S	9, 10, 11
	Set F	Y(Y)	Santa Monica	0.16	9.84	15.11	T	
		Y(N)	Elsinore fault	1.37	2.69	1.47	S	
	Set G	Y(Y)	Santa Monica	6.20	6.46	11.18	T	
		Y(N)	Elsinore fault	4.44	2.17	0.05	S	

[a]Modified and extended from the table posted at
http://scec.usc.edu/planning/Phase_II_true_signal,_Groups_A,_B,_C
[b]Y=Yes, N=No.
[c]The signal duration is based on the fault slip used in the FAKENET code. Slip rates in most cases are not uniform over the duration. For example, the second signal in Set 010157 mostly occurred for first a few days, not for the whole duration of 4.01 years. The slip rate in the Set 002307 increases beginning in early 2002 and lasting for ~6.6 years.
[d]Max. Disp. means the maximum displacement. H = Horizontal, V = Vertical.
[e]T = Thrust, S = strike slip, P = Propagation, F = Flow, A = Aquifer inflation.

2.3 Detection procedure

The detection algorithm uses state estimation and PCA (see Appendices for technical details). State estimation improves the signal-to-noise ratio (SNR) in the time domain by reducing the random noise, while PCA improves the SNR in the space domain by accounting for the spatial coherence of transient signals. The output of state estimation is used as the input of PCA. In this section, we show the detection procedure with examples from the synthetic data sets.

We used the smoother approach to construct an estimator. The state vector included an initial offset, secular velocity, annual and semi-annual sinusoids and a first-order Gauss-Markov (FOGM) process. The FOGM process accounts for not only temporally correlated noise but also transient signals if any. The most time-consuming part of the analysis is the estimation of the FOGM parameters (i.e., correlation time and variance). We estimated the parameters by minimizing innovation-based cost functions obtained from Kalman filtering (see Appendix A). On a mid-speed PC, the estimation took about 4~5 hours per data set. Figure 2-2 shows an example of state estimation results. The FOGM estimates well represent the signal around 2004. The state estimates are tabulated in Table 2-2 except for the time-varying FOGM state estimates.

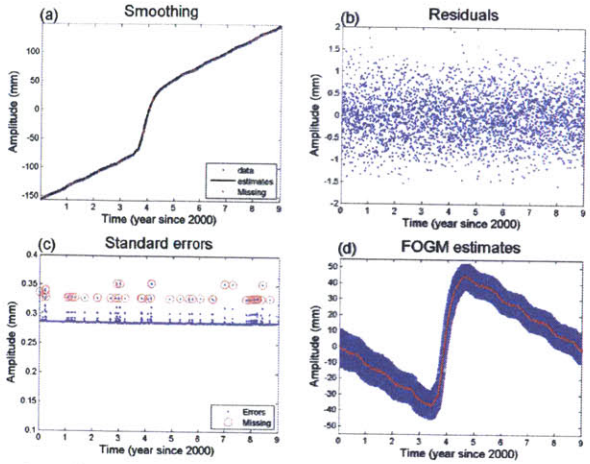


Figure 2-2. State estimation results from the north component of the station JPLM in Phase IIA Set 003923. (a) State estimates (black curve) of daily GPS time series (blue dots; not clearly seen from the figure). A large signal is shown around 2004. (b) Residuals after removing the state estimates. Residuals look white with standard error of ~0.54 mm. (c) Standard errors of the state estimates (blue dots). Values enclosed by red circles are those of missing data interpolation and they are larger than those of nearby existing data. (d) FOGM state estimates (red dots) with uncertainties (blue error bars). The FOGM estimates mainly represent the transient slip. The small oscillation is due to the leakage of the large signal power into the annual and semiannual sinusoids.

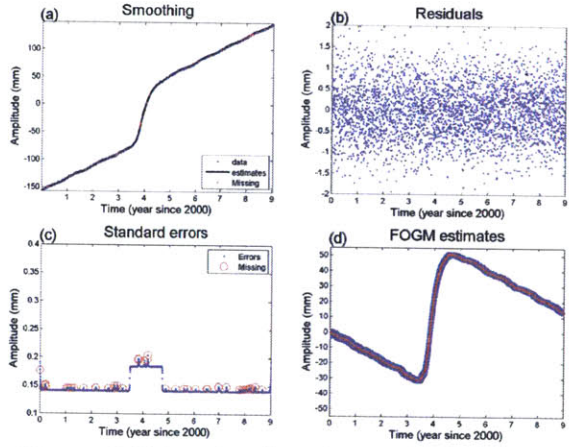


Figure 2-3. Same as Figure 2-2 but the time interval between 2003.496 and 2004.749 was not included in the FOGM parameter estimation. This reduces standard errors from ~0.29 mm to ~0.15 mm and also the uncertainties of FOGM state estimates from ~8.45 mm to ~1.74 mm on average. On the other hand, the root mean square error of the residuals slightly increases from 0.54 to 0.65 mm, which indicates more smooth state estimates due to smaller FOGM variance estimates. For the smoother to track the large slip, we increased the process variances in the time interval by a factor of three. As shown in (c), the standard errors in the time interval become somewhat larger. The small oscillation shown in Figure 2-2 has been mitigated.

The FOGM process is used to account for both transient signals and temporally correlated noise in the time series. Due to different temporal characteristics, however, transient signals can affect the estimation of correlated noise and vice versa. Suppose that a short and large signal exists in the time series with small background correlated noise. The signal will cause the FOGM variance and the FOGM state errors to be larger than expected. On the contrary, the small background noise can make FOGM state estimates poorly track the signal. When a detected transient signal is large as in Figure 2-2, we re-estimate the FOGM parameters excluding data within the signal time interval. The FOGM variances and state errors decrease and become consistent with background time-correlated noise (Figure 2-3 and Table 2-2). Furthermore, we re-estimate the FOGM state by increasing the process variance within the signal time interval so that the FOGM state estimates can track the signal properly. Depending on how much the process variance is increased, however, the level of smoothing can be degraded with larger state errors.

Table 2-2. State estimation results of the station JPLM in Phase IIA 003923 data set

		All data used[a]			Data between 2003.496 and 2004.749 NOT used[b]		
Component		North	East	Up	North	East	Up
Secular rate (mm/year)		33.80±1.33	-23.48±1.25	-4.64±2.20	32.15±0.30	-23.27±0.49	-4.76±1.24
Annual (mm)	Sin	1.80±0.41	0.44±0.84	-4.86±0.76	1.09±0.11	0.34±0.53	-4.49±0.68
	Cos	-0.77±0.41	1.14±0.84	-4.37±0.76	-0.67±0.11	0.99±0.54	-4.36±0.69
Semiannual (mm)	Sin	0.18±0.21	0.37±0.42	0.29±0.39	0.09±0.06	0.43±0.27	0.31±0.35
	Cos	0.10±0.21	0.06±0.42	-0.19±0.39	0.05±0.06	0.01±0.27	-0.02±0.35
FOGM parameters[c]	Tau (day)	3288	563	2276	2346	285	868
	Var (mm^2)	135.77	98.20	318.80	5.85	18.97	89.55
RMSE (mm)		0.54	0.54	2.64	0.65	0.53	2.64
FOGM error (mm)		8.45	5.68	12.72	1.74	2.18	6.08

[a]See Figure 2-2.
[b]See Figure 2-3.
[c]FOGM: first-order Gauss-Markov process, Tau: correlation time, Var: variance.

Each synthetic data set contains common temporal patterns in all three coordinates (Figure 2-4). The common patterns, often called common mode errors (CME) [*Wdowinski et al.*, 1997; *Dong et al.*, 2006], were removed by translation, rotation and scaling of the network (i.e., a seven-parameter Helmert transformation). The transformation is with a set of reference frame sites (see Figure 2-1). The transformation generally performs more robustly with large numbers of uniformly distributed reference frame sites. We chose the sites whose velocity uncertainty less than 0.1 mm/yr (indicating low noise and long duration) in the horizontal and 0.15 mm/yr in the vertical when velocity uncertainties are estimated accounting for correlated noise. We excluded those sites within the transient deforming region once the region was identified. The number of reference frame sites varies from 45 to 99 for Phase II and from 76 to 146 for Phase III. Figure 2-5 shows that the CME is mostly mitigated after the transformation.

Based on the variance-covariance structure, PCA decomposes the FOGM state estimates into principal components (PCs) for the temporal variation and sample eigenvectors for the spatial distribution of the PCs. Horizontal and vertical components were separately analyzed. Uncertainties of PCs were obtained by propagating FOGM state uncertainties plus FOGM variance estimates along the principal axes represented by sample eigenvectors. Uncertainties of sample eigenvectors were obtained from an asymptotic distribution with effective sample size to account for temporal correlation in the FOGM state estimates. When data are temporally correlated, lag differences (i.e., differences between points of a PC with a time lag) can have smaller uncertainties than the PCs have so that lag differences may reveal transient signals which are not apparent in the PCs. Appendix D discusses the uncertainty estimation of PCA and lag differences.

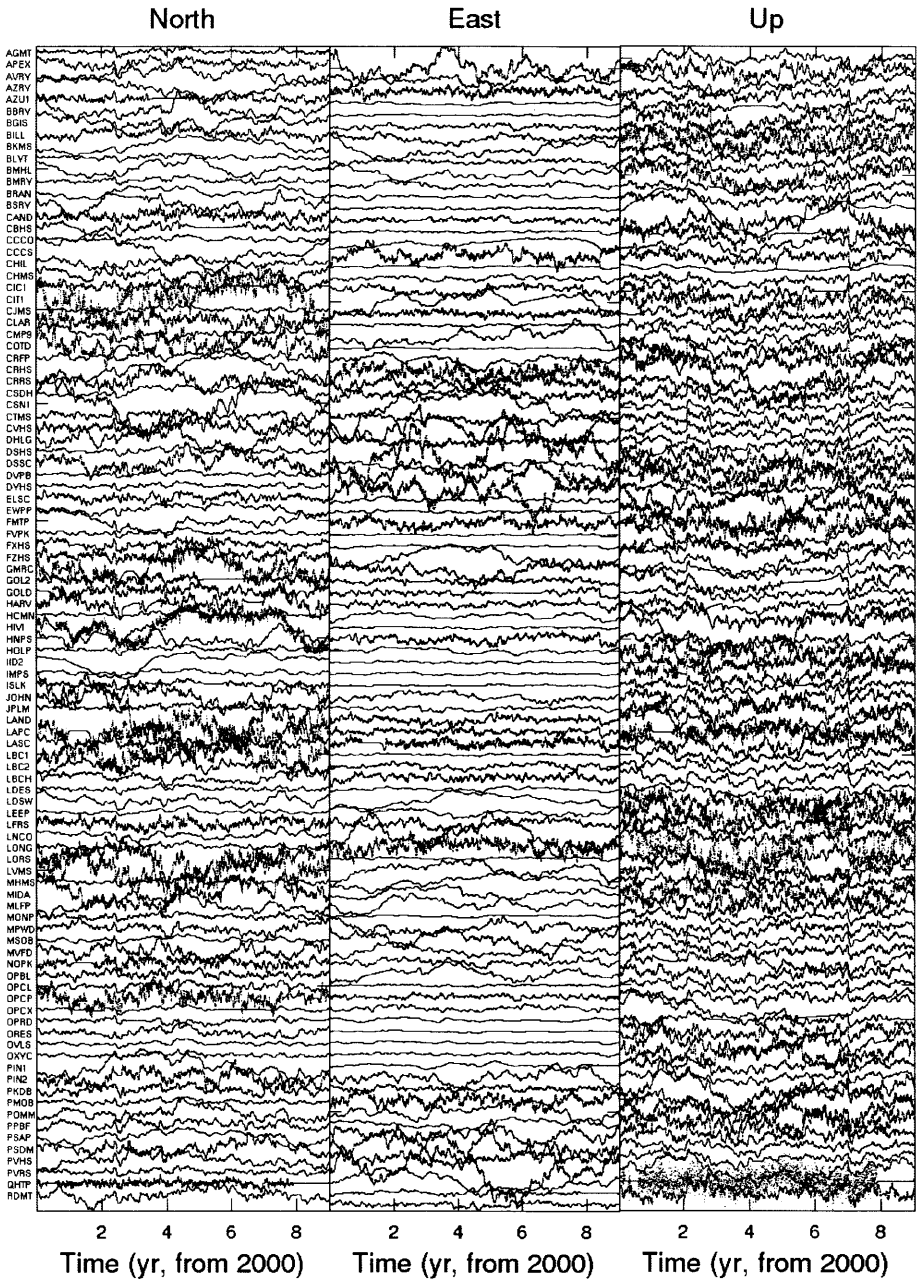


Figure 2-4. A stack of FOGM state estimates from Phase IIA Set 002307. Only 100 stations are shown with full-time (9 years) data span. Horizontal components are scaled by a factor of three relative to the vertical scale. The plot provides an overview of temporal variation in a large network. Common patterns between sites can be found in each component such as those at time 2.5 in north and at times 2.0, 2.5, and 7.0 in vertical (best seen by viewing page edge-on). The data set actually contains a transient signal but the signal is not apparent because of low signal-to-noise ratio and correlated noise in space and time. We were not able to detect the transient signal.

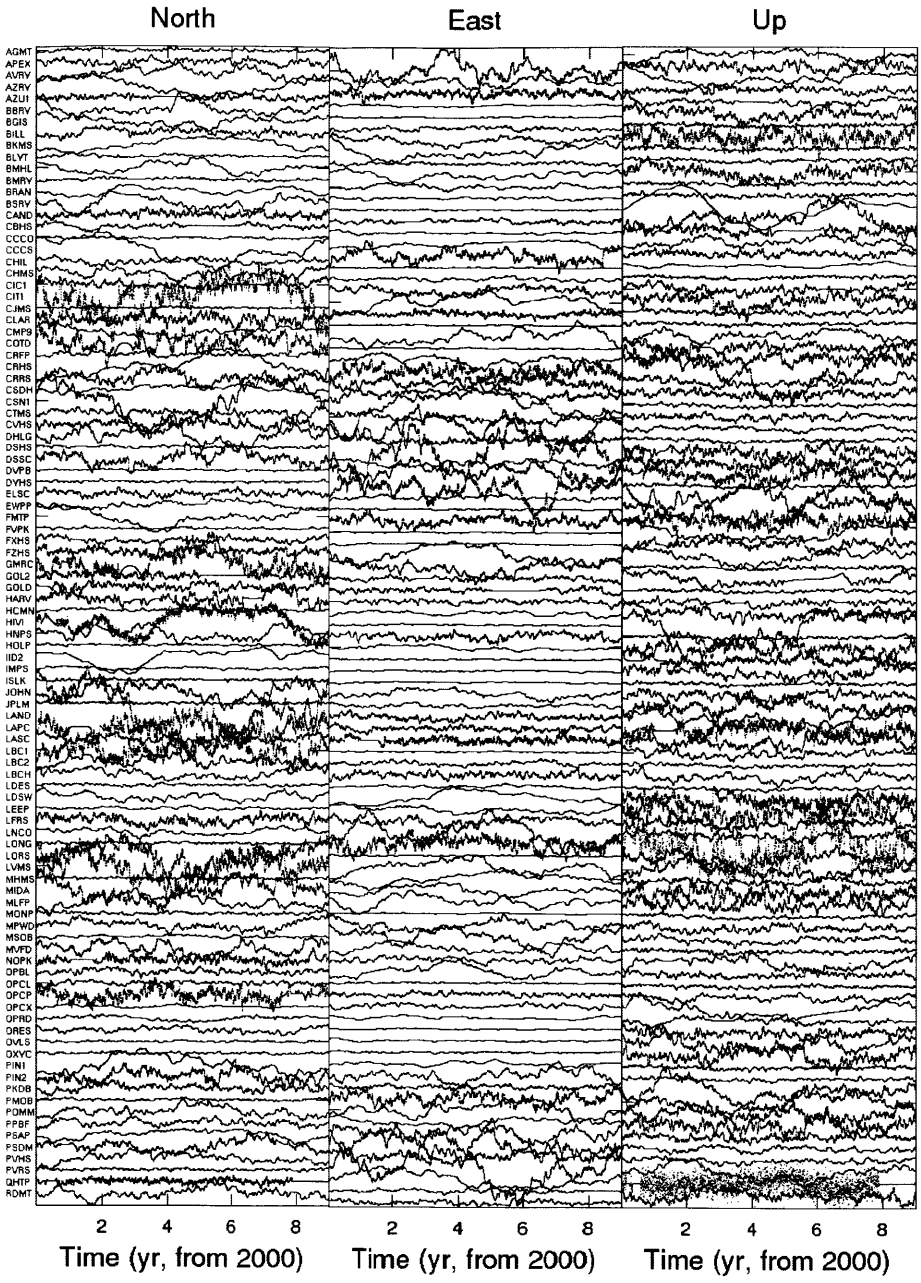


Figure 2-5. Same as in Figure 2-4 but after reference frame transformation. The common temporal patterns have been mostly disappeared.

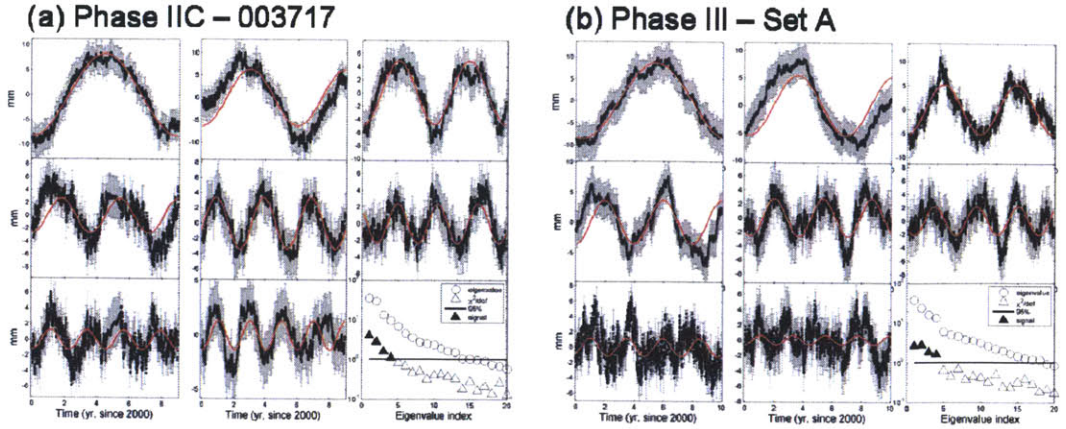


Figure 2-6. First 8 horizontal principal components (PCs; black dots) with 1-sigma uncertainties (gray error bars) and the fits of sinusoids (red curves). The lower-right subplots show first 20 sample eigenvalues (circles), chi-squares per degree-of-freedom (triangles) and 95% limit of the chi-squares statistics (solid horizontal line). Solid triangles indicate the chi-square values larger than the upper limit. The PCs exhibit oscillations with frequencies increasing with the PC ranking. The lowest frequency is one cycle per data span due to the removal of secular rates. Although the chi-square statistics show that the first four PCs are significant, the oscillations are due to the temporal correlation of the FOGM state estimates rather than due to transient signals. The data set in (b) actually contains a transient signal, but the dominant oscillations prevent the signal from being identified in the PCs. The effect of temporal correlation can be reduced by removing the sinusoidal fits from the PCs. The last two PCs in (b) show less clear oscillations because a few stations have large amplitudes in these components.

PCs are affected by temporal correlation in the FOGM state estimates [e.g., *Vautard and Ghil*, 1989; *Allen and Smith*, 1994; *Allen and Robertson*, 1996; *Allen and Smith*, 1996]. Most of the data sets, except for those with large signals, show oscillations in PCs with frequencies increasing with the PC ranking (Figure 2-6). Such oscillations could make transient signals unclear if the SNR is not sufficiently high. In this case, the effect of temporal correlation can be reduced by fitting and removing sinusoids from the PCs (see Appendix D). The residual PCs are used in the lag difference calculation.

The size of data sets (~450 stations and ~9-year time interval) is too big for PCA to detect a weak transient signal. As the data size increases in space and/or time, noise contributes to PCA more than a localized transient signal does. As a result, PCA with a

large data set often shows random spatial patterns and low-frequency temporal oscillations in PCs (Figure 2-7). To reduce the noise contribution, we divided a data set into several subsets. For example, we divided the Southern California into four regions and used data within a 2-year moving window (Figure 2-7). The number and the size of subsets are flexible to change. When a transient signal is suspected, a subset can be chosen with appropriate size to remove unnecessary data and concentrate on the suspected signal.

Some of stations possibly have a temporal pattern which has large amplitude and is not coherent with nearby stations. The large-amplitude temporal pattern may mask a transient signal (Figure 2-8a). The variance of one component can be larger than the total variance of the signal at other sites. It thus appears on the largest eigenvalue component. Figure 2-8b shows that a transient signal has been revealed when large amplitude stations are removed.

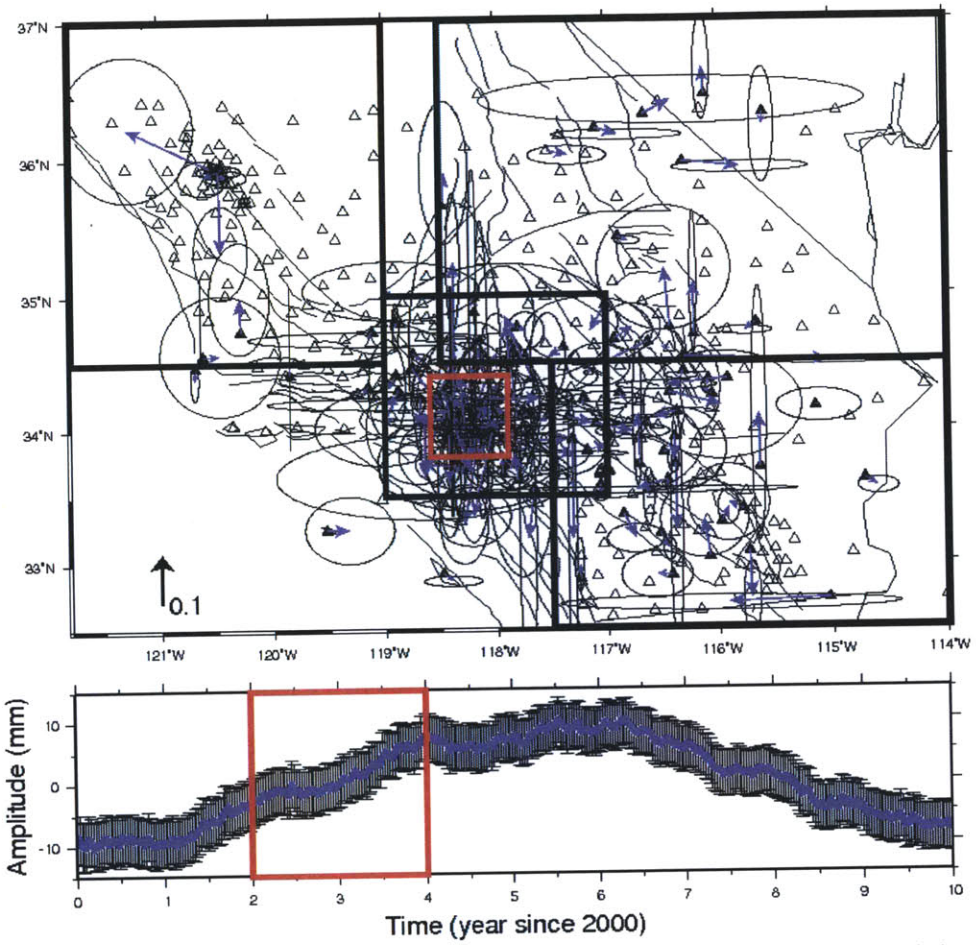


Figure 2-7. Random spatial pattern and low-frequency temporal oscillation of the first principal component from Phase III Set A. The patterns indicate dominant noise contribution. To reduce the large-scale noise effects, the data size should be decreased in both space and time (squares). If a transient signal is suspected, the windows can be more confined to the region and the time around the signal. A transient signal was detected within the red squares (see Figure 2-12) but it is not seen from the large-scale PCA results. Open triangles indicate those stations that do not span the 10-year time interval. Error ellipses and bars indicate 1-sigma uncertainties. The black arrow in the lower-left corner indicates the spatial length scale of 0.1. Actual displacement at each station is the product of the spatial amplitude and the temporal amplitude. For example, the black arrow indicate ~0.9 mm between 2000.0 and 2003.0, ~1.7 mm between 2000.0 and 2006.0, or ~0.1 mm between 2000.0 and 2010.0.

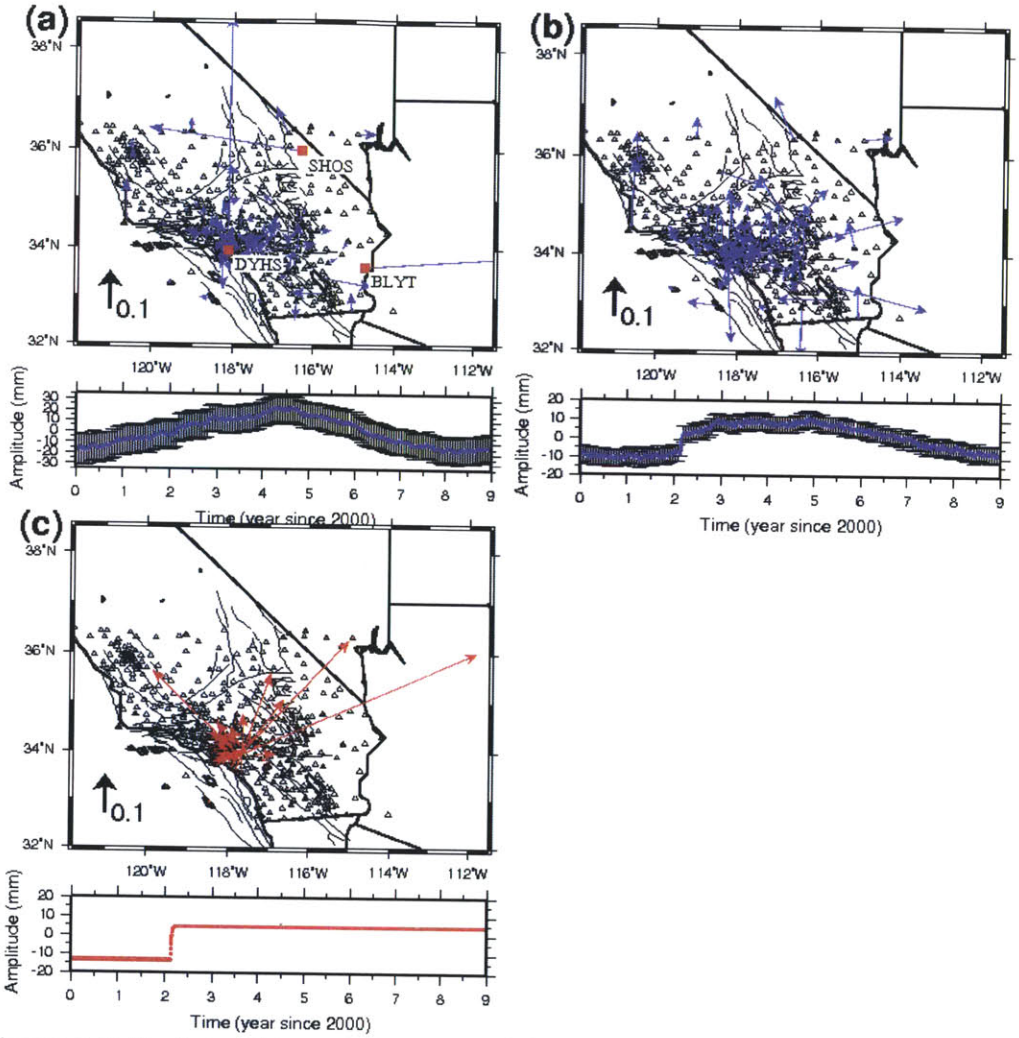


Figure 2-8. Spatial and temporal patterns of the first principal component from Phase IIB Set 010056. (a) The spatial pattern shows three stations (BLYT, DYHS and SHOS) whose amplitudes are anomalously large because of large site-specific time-correlated noise. The temporal pattern shows a low-frequency oscillation, indicating dominant noise effect. (b) After removing the stations from the data set, a localized systematic spatial pattern has been revealed at latitude ~34°N and longitude ~117.5°W. Furthermore, the temporal pattern shows a fast slip in early 2002. (c) Spatial and temporal patterns of the true signal (red arrows and dots). This signal cannot be clearly identified in (a) due to the large amplitude stations. For clear identification, the data size must be reduced in space and time (see Figure 2-18). Error ellipses in space are not shown for clarity.

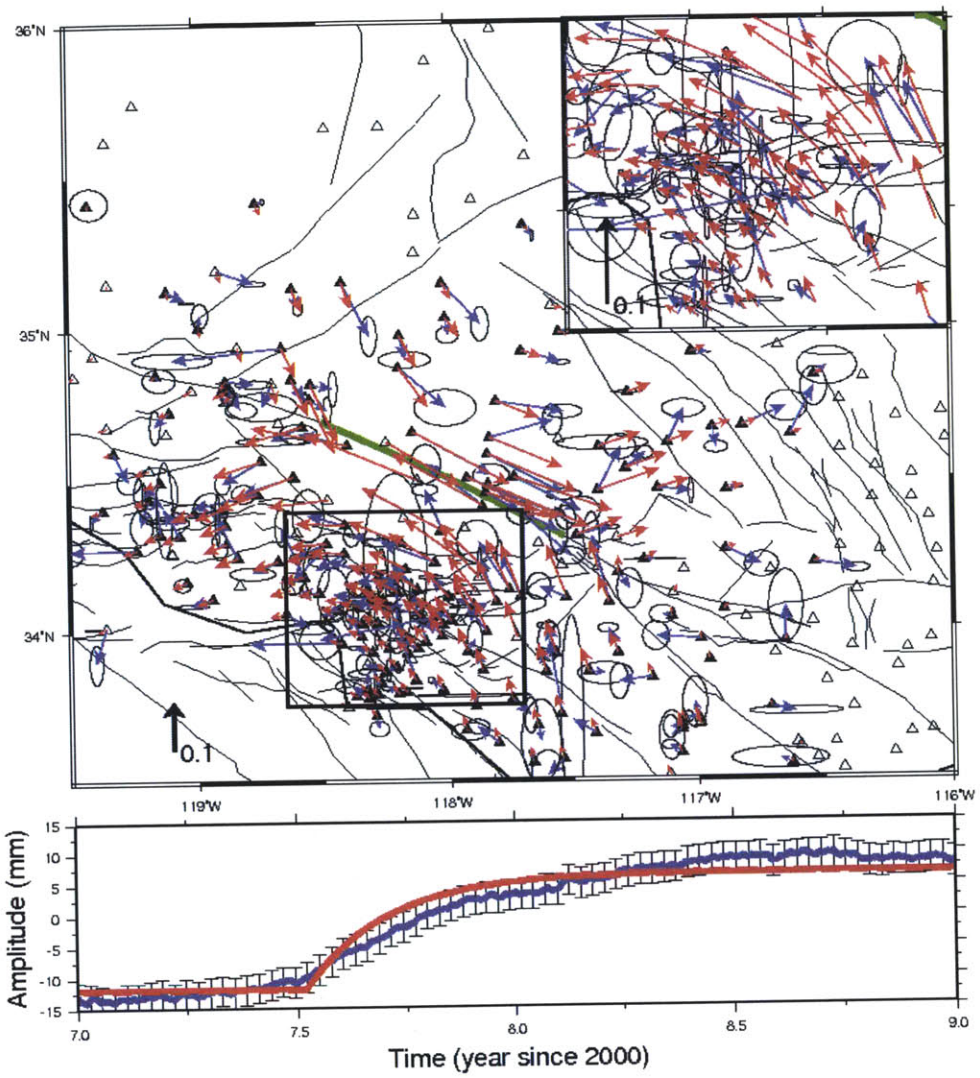


Figure 2-9. Spatial and temporal patterns (blue arrows and curve) of the first principal component (PC) from Phase III Set E. Error ellipses and bars indicate 1-sigma uncertainty. The black arrow is a scale of amplitude 0.1. The true signal (red arrows and curve) is a right-lateral strike slip on a vertical fault (green line) and the PC represents the signal. Although the maximum horizontal displacement of the signal is small (~4.46 mm), PCA detected the signal because the signal has recorded at a number of stations. The inset plot in the upper right zooms in the area surrounded by the black square.

2.4 Results

The algorithm successfully detected large signals obvious in the first PC (and one or more PCs when signals are propagating) but detection was not easy in low SNR cases. Figure 2-9 shows a low SNR case where the true signal has the maximum horizontal displacement of ~4.46 mm, which is the minimum among the signals detected in the test exercise. The first PC represents the true signal in both space and time but it is also represented by a sinusoidal oscillation (Figure 2-10), which may indicate noise contribution to the PC. Nevertheless, we determine the first PC as a transient signal because of its systematic and localized spatial pattern (Figure 2-9) and an anomalous time interval indicated by lag differences of the PC (Figure 2-11).

Transient signals were mainly detected in the horizontal components because they are less noisy than the vertical components. As an example, Phase III Set A contains a thrust-type signal whose maximum horizontal displacement (~4.58 mm) is less than the vertical displacement (~7.53 mm), but the horizontal components represent the true signal better than the vertical components (Figure 2-12). Nevertheless, vertical components can be useful to support a transient signal detected in horizontal components.

The detection algorithm was not successful at identifying some of the weak signals (see Table 2-1). The smallest horizontal displacement among the detected signals is ~4.46 mm in Phase III Set E (Figure 2-9). However, some of larger signals were not detected due to signal complexity. Figure 2-13 shows an undetected case of a larger signal; a transient signal with maximum horizontal displacement of ~4.99 mm is measured by only two stations TRAK and FVPK with sufficient amplitudes. Although the spatial pattern at the stations is consistent with the true signal, the stations were

treated as those with anomalously large amplitude. Furthermore, the temporal pattern does not represent the fast slip. If the signal was measured at more near-field stations, PCA would have detected the signal as done in the minimum displacement case shown in Figure 2-9.

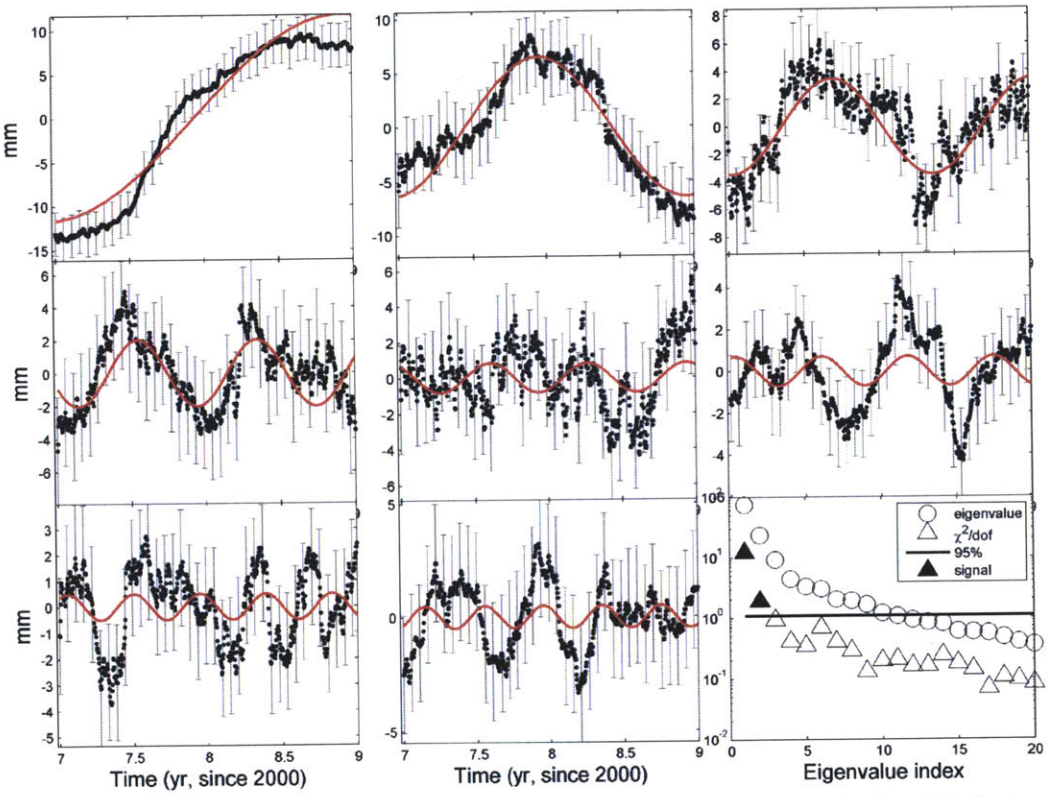


Figure 2-10. First 8 horizontal principal components (PCs; black dots) with 1-sigma uncertainties (gray error bars) from Phase III Set E. The chi-square test shows that the first PC is significant and the second PC is marginally significant (lower-right subplot). However, the two PCs are well represented by sinusoidal oscillations (red curves), suggesting noise contribution to the PCs. Nevertheless, the temporal pattern of the first PC is determined as a transient signal because the first PC shows a systematic and localized spatial pattern (see Figure 2-9). On the other hand, noise probably dominates the second PC because of its random spatial distribution.

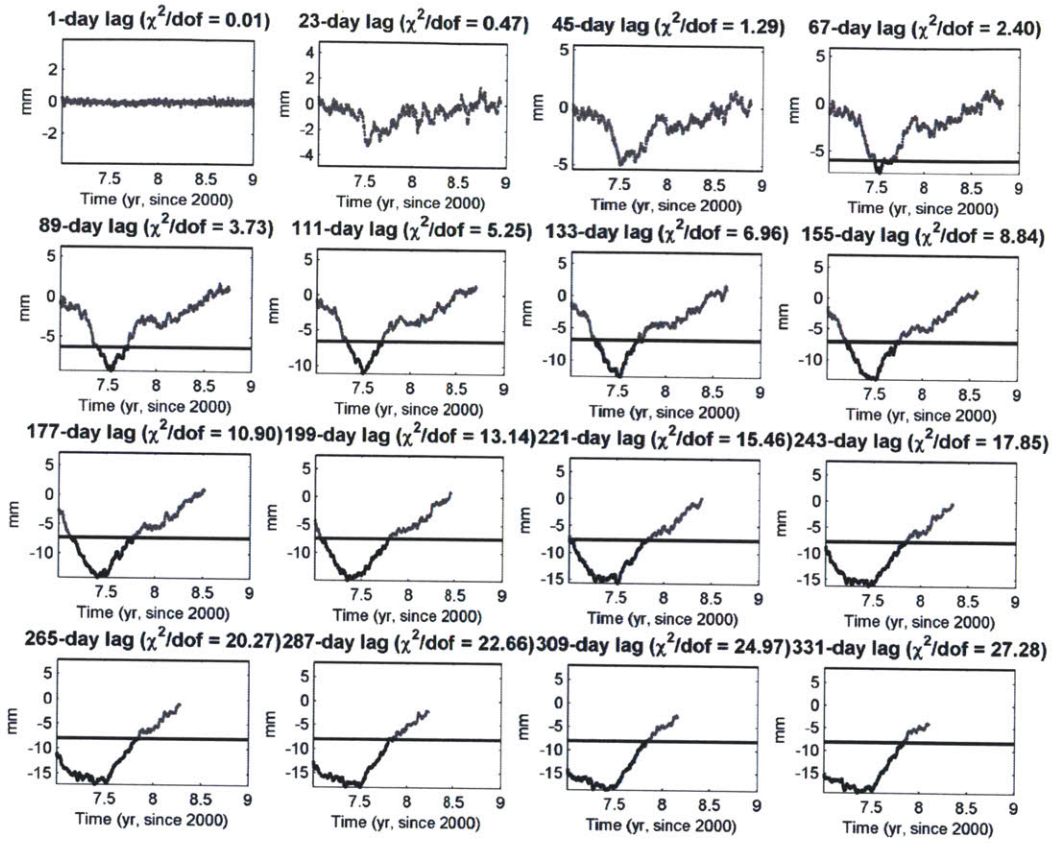


Figure 2-11. Lag differences of the first principal component (PC) shown in Figure 2-10 with given time lags shown on the top of each subplot. The epochs of the differences are associated with the early epoch of the two differenced points. The horizontal lines show the 3-sigma uncertainty of the lag differences. The lag differences outside the 3-sigma range are indicated by black dots. For the lag of 67 days and more, a time interval around 2007.5 is anomalous in terms of the 3-sigma uncertainty with chi-squares per degrees of freedom of 2.4 and larger. The true signal starts at 2007.52 and ends at 2008.68.

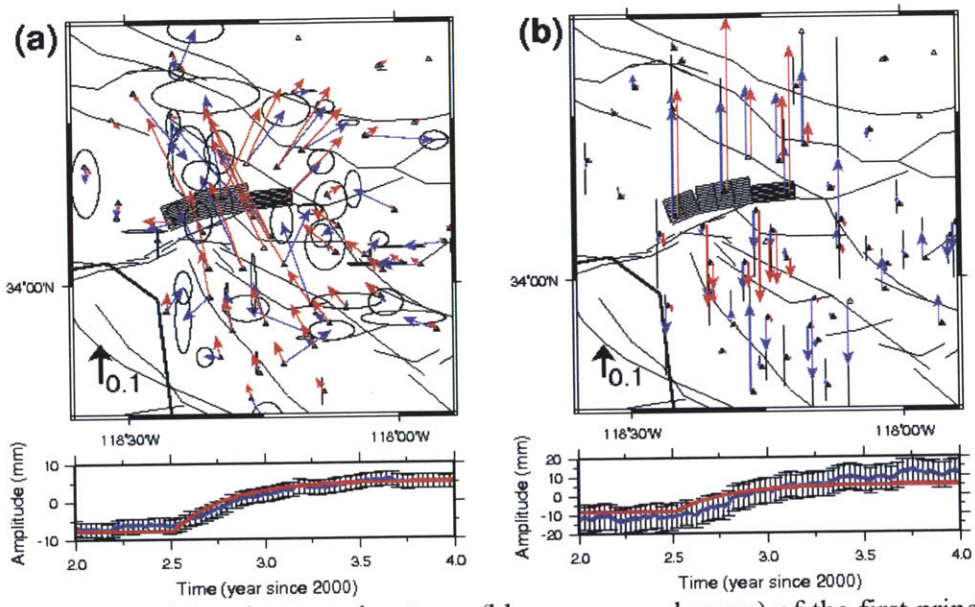


Figure 2-12. Spatial and temporal patterns (blue arrows and curve) of the first principal component (PC) from (a) the horizontal and (b) the vertical components in Phase III Set A. Error ellipses and bars indicate 1-sigma uncertainty. The black arrow is a scale of amplitude 0.1. Black squares show the fault plane projected onto the surface. The horizontal components represent the patterns of the true signal (red arrows and curve) better than the vertical components. A larger scale map is shown in Figure 2-7.

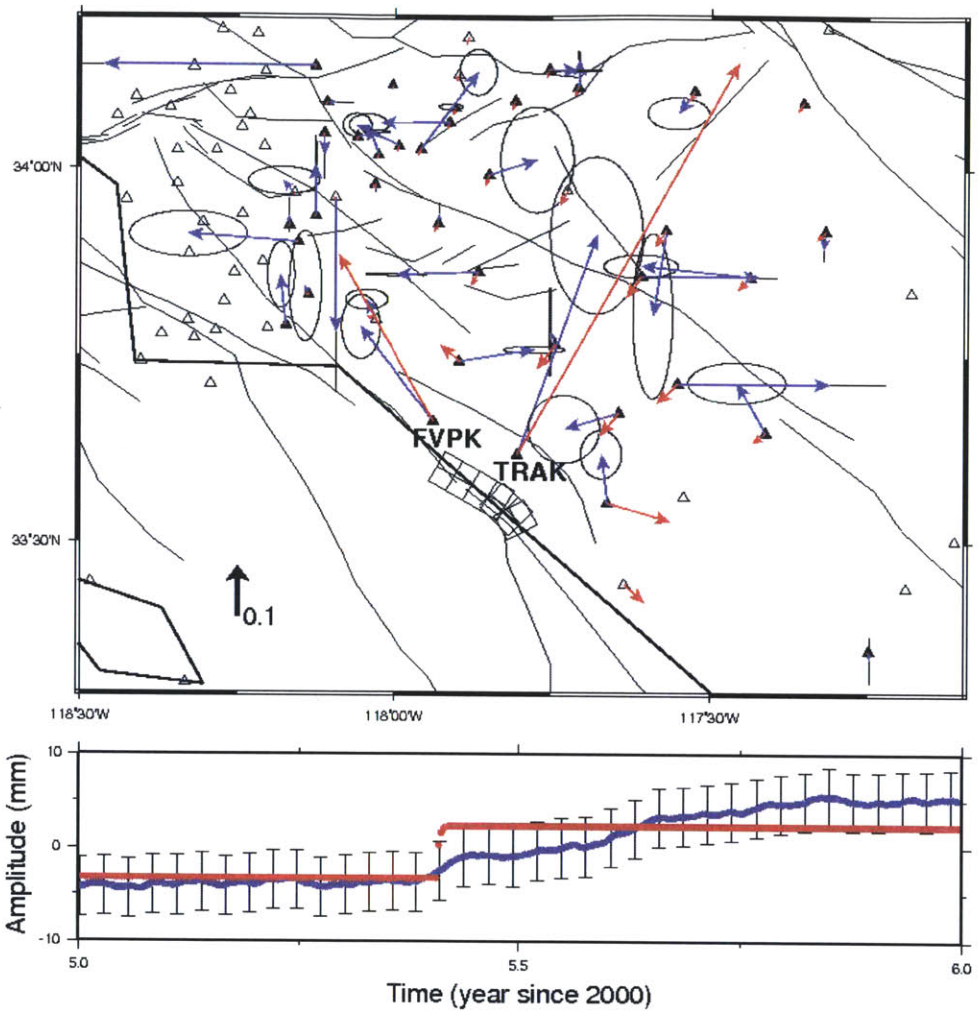


Figure 2-13. Spatial and temporal patterns (blue arrows and curve) of the first principal component (PC) from Phase IIB Set 003756. The true signal (red arrows and curve) is a thrust event on a fault plane dipping ~23° (black squares). Although the two stations TRAK and FVPK are consistent with the true signal pattern, the overall noisy pattern leads the two stations to be locally anomalous stations with large amplitudes. The fast slip at epoch 2005.41 cannot be seen from the PC temporal pattern. For the signal to be detected, more number of stations should be required particularly in the near field. Error ellipses and bars indicate 1-sigma uncertainty. The black arrow is a scale of amplitude 0.1.

2.5 Issues related to transient signal detection

In addition to the low SNR, the signal detection by PCA can be limited due to signal complexities such as large spatial scale signals, multiple signals, propagating signals, and signals with missing data. Furthermore, PCA uncertainty is a critical factor for the detection capability. In the following subsections, we discuss issues on the detection of the complex signals and on the PCA uncertainties.

2.5.1 Large-scale signals

This study defines a transient signal as being localized in a given space. Spatial locality can separate a transient signal from CME over a network. However, the locality is relative to the size of data used in PCA. When spatial extent of data is smaller than a transient signal, the separation will be difficult. Figure 2-14 shows a large-scale signal in space, a uniform flow to the southwest over the network. The spatial pattern of the first PC roughly represents the true signal, but the detection failed because the PC pattern is not localized. Moreover, the temporal pattern of the PC shows an oscillation that can be thought of as the contribution of temporally correlated noise. If adjusted by a linear rate (green curve in Figure 2-14), the PC temporal pattern approximates the true signal. To some extent, PCA revealed the signal in space and time.

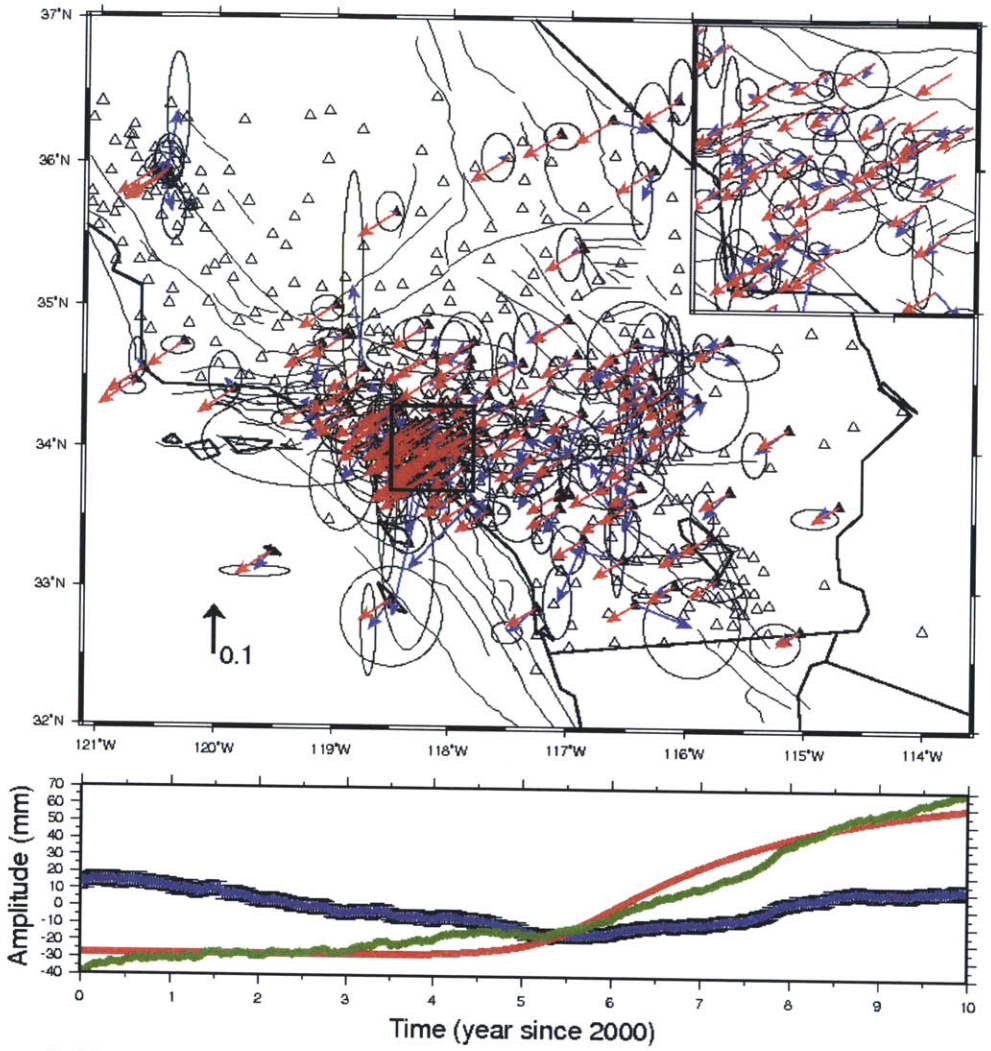


Figure 2-14. Spatial and temporal patterns (blue arrows and curve) of the first principal component (PC) from Phase III Set C. The true signal (red arrows and curve) is a large-scale flow to the southwest with the maximum horizontal displacement of ~7.86 mm. The temporal extent of the signal is relatively long (~8.71 years but ~5 years for majority of the signal). Although the PC roughly represents the flow in space, the detection failed in the test exercise because the PC results in non-localized spatial pattern and a low-frequency temporal oscillation. The oscillation has a frequency of 1 cycle per data span because the secular rate was removed. The PC fit to the signal (green curve) indicates that the PC temporal pattern represented the true signal but it was rotated by the secular rate estimation. The inset plot zooms in the area surrounded by the black square.

2.5.2 Multiple signals

Some of the data sets contain multiple signals (see Table 2-1). Multiple signals can be detected in some cases. Figure 2-15, for example, shows the case where multiple signals are distinct in time. Two thrust-faulting signals had occurred for 2.27 years since 2001.61 and for 4.01 years since 2003.06. Although the signal time intervals are overlapped, the majority of the signals are separated and distinct. When PCA used data between 2001.0 and 2004.0, the two signals appeared in multiple PCs (Figure 2-15a and b). The spatial pattern of the second PC is different from that of a thrust faulting, complicating interpretation. When PCA used data between 2001.0 and 2003.0 for the larger signal and between 2002.75 and 2003.5 for the smaller one, the first PC well represents most of the spatial and temporal pattern of each signal (Figure 2-15c and d). Signal interpretation becomes simpler.

When multiple signals exist, the larger signals can mask the smaller ones. The Phase III Set C is a good example. In addition to the large-scale flow (Figure 2-14), the data set contains thrust faulting and inflation in a horizontal opening aquifer. The thrust signal was detected because of its high SNR, while the aquifer signal was not detected because of contamination by the large-scale flow (Figure 2-16). The temporal pattern of the first PC is almost linear. The differences between the first PC and true signal (black arrows in Figure 2-16b) roughly indicate southwestward motion as in the large-scale flow. Without the large-scale signal, the aquifer signal might have been detected because its horizontal displacement of ~4.97 mm is not so small compared with other detected signals (see Table 2-1) and because the signal was recorded at multiple stations.

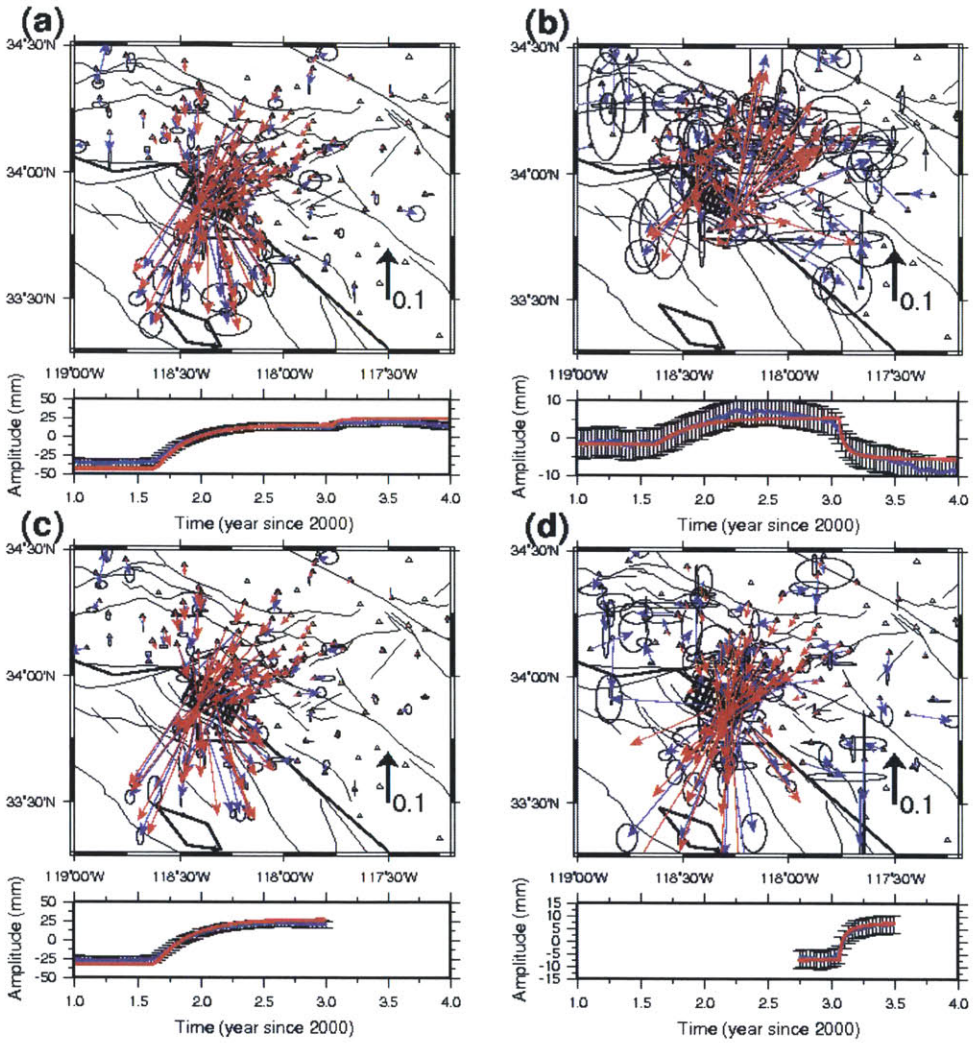


Figure 2-15. Spatial and temporal patterns (blue arrows and curves) from Phase IIC Set 010157: (a) the first principal component (PC) and (b) the second PC, (c) the first PC between 2001.00 and 2003.00, and (d) the first PC between 2002.75 and 2003.5. The data set has two trust-type signals (red arrows and curves) on a fault dipping ~17° (black squares). When the two signals are included in PCA, they appear at multiple PCs (a and b). The second PC looks different from the thrust faulting. Since the two signals occurred at distinct time interval, each signal can be treated in PCA separately (c and d). The thrust-type signals clearly appear at the first PC, making interpretation simple.

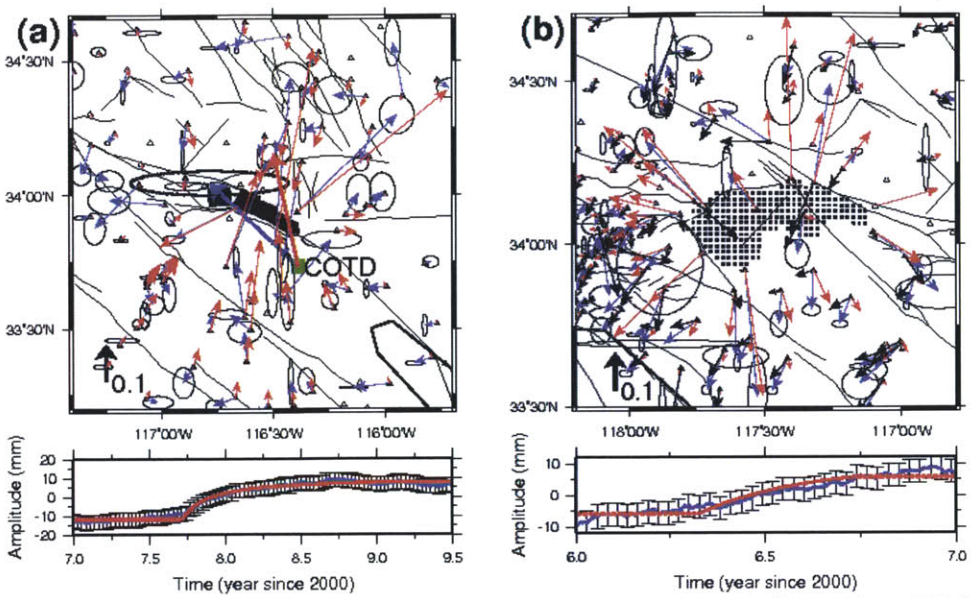


Figure 2-16. Spatial and temporal patterns (blue arrows and curves) from Phase III Set C in which there are three types of signals: (a) a thrust faulting on a fault dipping ~55° (black patches), (b) inflation in a horizontal opening aquifer (black patches), and large-scale flow at depth shown in Figure 2-14. PCA results from the true signal are shown by red arrows and curves. The thrust signal was detected because of its large amplitude; the maximum horizontal displacement is ~10.22 mm larger than that of the large-scale flow. On the other hand, the aquifer signal was not detected because of the low SNR and contamination by the large-scale flow. The differences between PC and true signal (black arrows) roughly indicate southwest direction that is consistent with the large-scale flow. Furthermore, the PC temporal pattern in (b) is almost linear so that it is difficult to determine if the temporal pattern is due to the large-scale flow or the aquifer signal. Note that the east component of station COTD (green square) is inconsistent and has larger uncertainty (thick ellipse) relative to the north component. This is due to large temporally correlated noise in the east component (see Figure 2-21).

2.5.3 Propagating signals

When multiple stations experience a transient signal simultaneously (i.e., coherent signal), the signal will dominantly appear at the first PC if the SNR is high. The signal coherence can distinguish tectonic signals from non-tectonic signals that often occur incoherently. However, all transient signals are not always coherent. An example is signal propagation in space and time. Here propagating signals mean the surface deformation due to slip propagation along a fault so that each station observes the surface deformation at

different epochs and possibly different time intervals. The signal propagation can be more complicated depending on the fault geometry, propagation velocity and direction, network configuration and sampling rate. PCA is capable of detecting a propagating signal but the signal appears in multiple PCs and the multiples complicate the interpretation (Figure 2-17).

2.5.4 Missing data

Since PCA takes data measured simultaneously, a large data gap makes a considerable amount of data directly unusable. Missing data interpolation retains measurements that otherwise will not be used in PCA. Figure 2-18 shows that the signal can be approximately recovered from the resulting PCs even when the stations OEOC and CNPP have data missing during the signal time interval. The temporal pattern of the first PC is slightly over-smoothed during the interval of the signal due to the interpolation. As the FOGM correlation time becomes longer relative to a data gap, interpolation tends to be linear (Appendix B). The minimum FOGM correlation time of the two stations is 156 days larger than the missing gap of ~43 days. The linearly interpolated FOGM estimates produce somewhat smooth PC. Figure 2-19 shows that the reconstructed time series from the first PC approximate the true signal, even when the signal was missing.

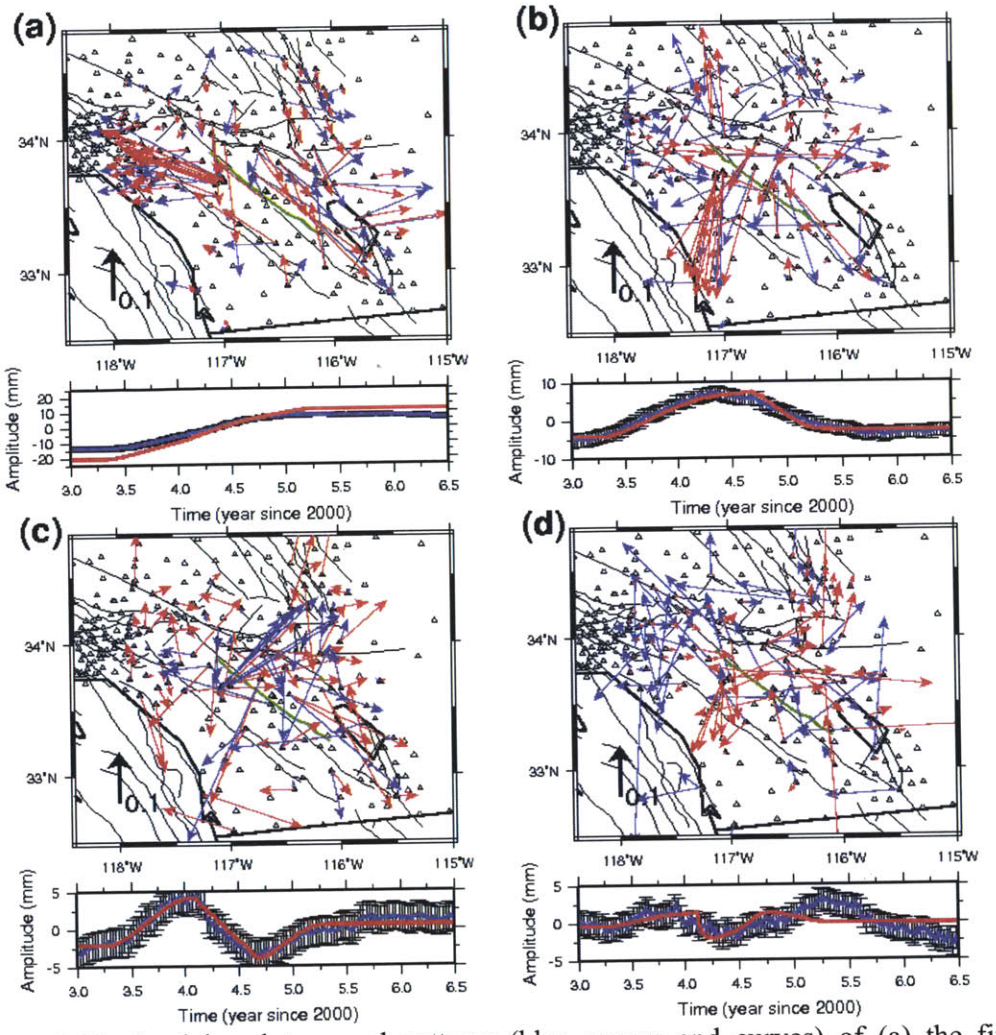


Figure 2-17. Spatial and temporal patterns (blue arrows and curves) of (a) the first principal component (PC), (b) the second PC, (c) the third PC, and (d) the fourth PC from Phase IIC Set 004608. Error ellipses are not shown for clarity. The red arrows and curves were generated by PCA of the true signal that is a propagating right-lateral slip on a vertical fault (green line). The slip started at ~33.5°N and ~116.5°W and propagated along the fault plane bidirectionally. PCA can detect the propagating signal but the signal is spread over multiple PCs so that the interpretation is complicate. The first PC shows a right-lateral strike slip as of the true signal, but lower-ranked PCs shows noisy spatial pattern. The temporal patterns of the second and the third PCs are still similar to that of the true signal.

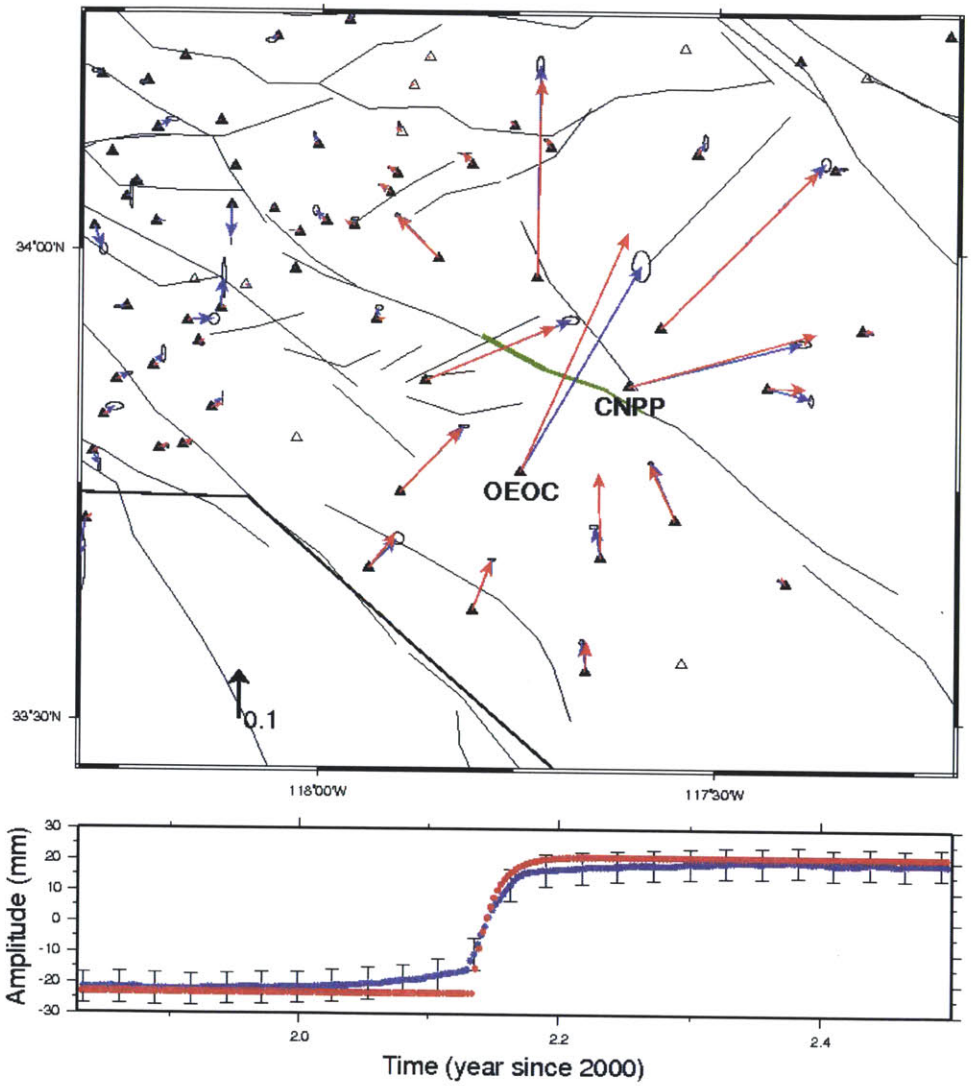


Figure 2-18. Spatial and temporal patterns (blue arrows and curve) of the first principal component (PC) from Phase IIB Set 010056. The true signal (red arrows and curve) is a thrust faulting on a fault dipping ~80° (green line). The spatial pattern of the PC well represents the true signal. The temporal pattern of the PC is slightly over-smoothed because the stations OEOC (see Figure 2-19) and CNPP do not have data during the signal time interval.

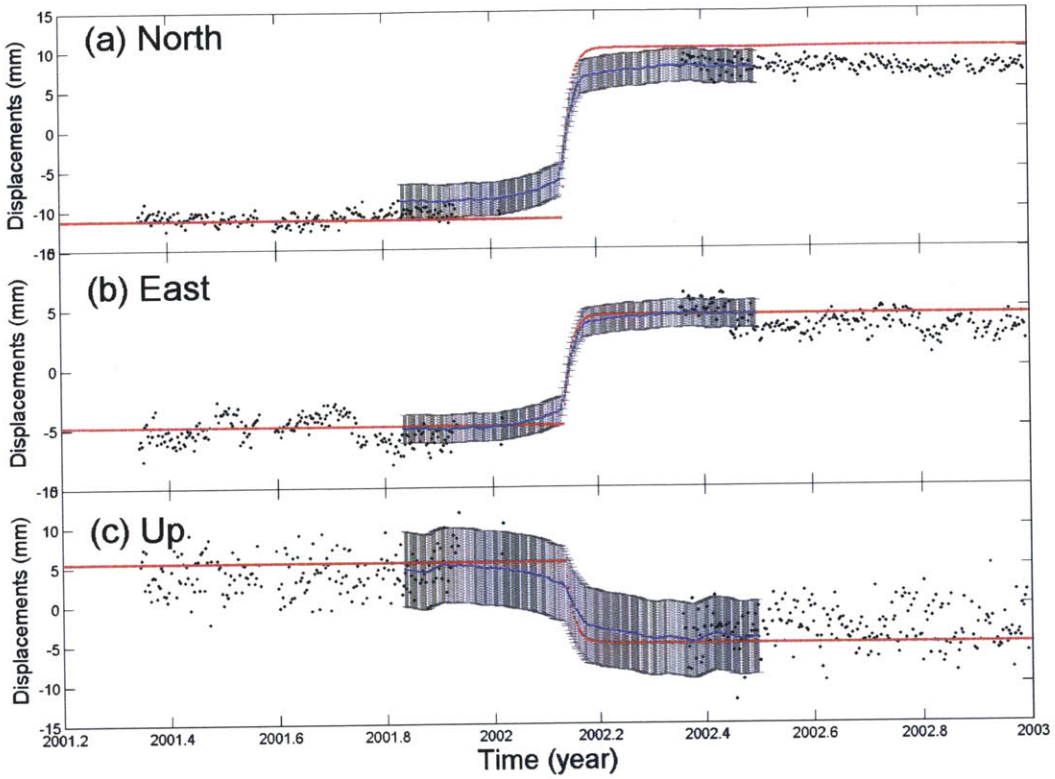


Figure 2-19. Detrended data time series (black dots), reconstructed time series from the first principal component (PC; blue dots) with uncertainties (gray error bars), and true signal (red) at station OEOC from Phase IIB Set 010056. The detrended time series were obtained by removing state estimates except for the first-order Gauss-Markov state estimates. Data in the signal time interval are not available. The reconstructed time series were obtained from the first PC multiplied by the first sample eigenvector element associated with the station OEOC. PCA was done with data between epochs 2001.83 and 2002.50 to increase the SNR (see Figure 2-18). PCA can estimate the missing data by considering spatial coherence of the signal.

2.5.5 Uncertainty estimates

GPS analysis packages provide statistics of position uncertainties, often called formal errors. The formal errors can be used in the Kalman filter as white noise measurement uncertainties. The state uncertainties can be readily obtained from the Kalman filter formulations. The uncertainty of temporally correlated noise can be described by the

FOGM variances, which are estimated by minimizing a cost function based on innovations sequence (see Appendix A).

The synthetic data sets did not provide formal error statistics. We assumed the typical formal errors of 1 mm in horizontal and 3 mm in vertical and scaled the assumed values by the method described in Appendix A. Figure 2-20 shows the estimates of white noise amplitudes in terms of the true values from Phase IIA data sets. The mean true amplitudes are 0.71, 0.58, and 2.42 mm in north, east and vertical components, respectively. The mean estimates are close to the true value (i.e., 0.94, 0.80, and 2.52 mm in north, east and vertical, respectively). The estimates and the true values are positively correlated but the correlation coefficients are somewhat broadly distributed (0.34 ~ 0.83). One possible explanation of such misfits is that random-walk and flicker noise were used for simulating time-correlated noise but the FOGM process was used for estimating the noise. The misfits can also be caused by short data length and missing data.

Both the state estimation uncertainty and the FOGM variance estimates are considered for PCA uncertainty estimation (see Appendix D). The resulting PCA uncertainties are used to infer the significance of the PCA results. Figure 2-16a shows that the east component of the station COTD is inconsistent with the true signal. Such inconsistency is due to large time-correlated noise interacting with the true signal constructively during the signal time interval (Figure 2-21). This component has large spatial uncertainty (Figure 2-16a) because the FOGM variance was estimated from the whole time series, not from those used in PCA. The large spatial uncertainty suggests that the east component is not significantly different from zero (green curves in Figure 2-21).

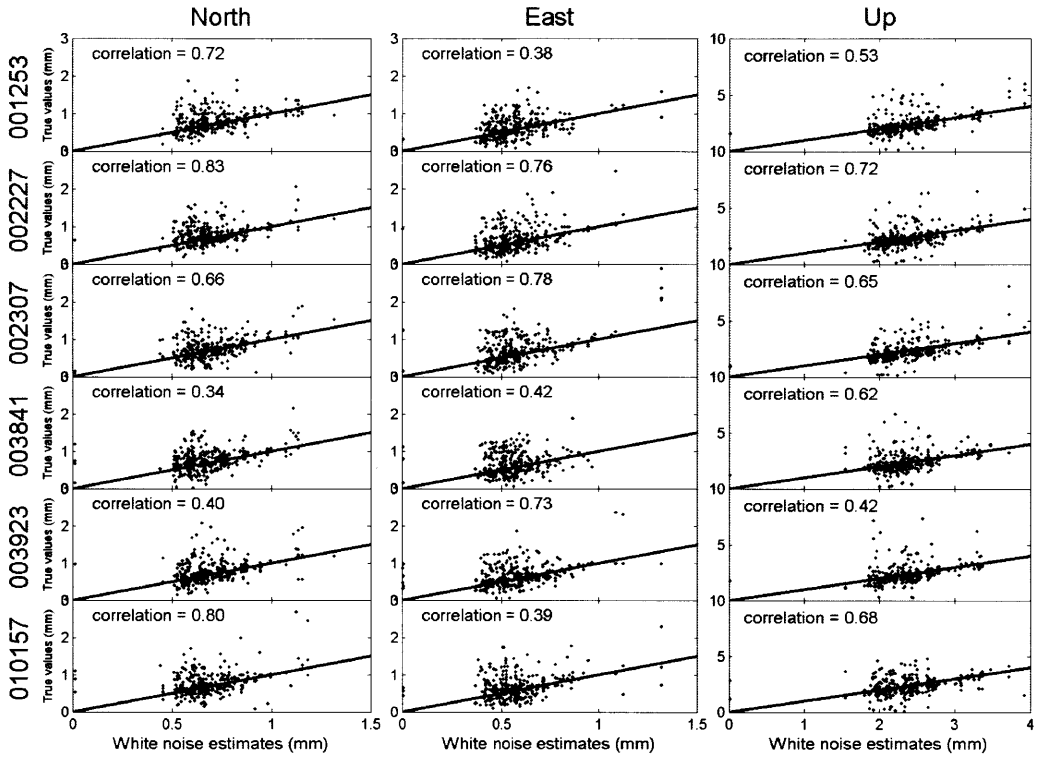


Figure 2-20. The estimates of white noise amplitudes vs. true values in all sets of Phase IIA. Stations with data length less than 1000 and missing data larger than 20% were not included in the calculation. The solid lines have a slope of one so that all values must lie along the line if the estimates are correct. There are positive correlation with correlation coefficients between 0.34 and 0.83. The white noise estimation was mostly affected by temporally correlated noise (simulated by random-walk and flicker processes) and incorrect noise model (i.e., first-order Gauss-Markov process).

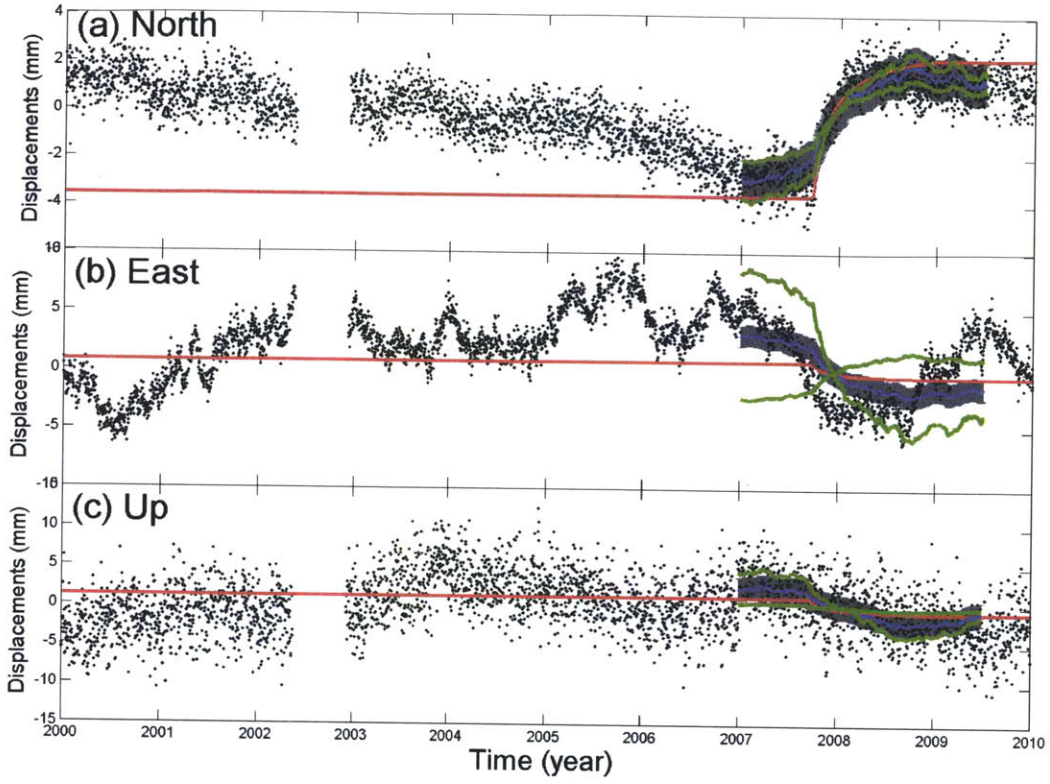


Figure 2-21. Similar to Figure 2-19 but for the station COTD from Phase III Set C. Also shown are the 3-sigma ranges of the reconstructed time series (green curves), which were obtained from the first PC multiplied by the maximum and minimum sample eigenvector associated with 3-sigma uncertainties. The east component shows a motion larger than the true signal. The large motion arises from the highly correlated noise interacting with the true signal constructively. The uncertainty of the sample eigenvector (green curves and also see Figure 2-16a) can explain the large variation.

2.6 Summary

Detection capability depends on many factors such as the SNR, temporal correlation, spatial extent and sampling of signals, and the signal complexity. A measure of detection threshold cannot be generally obtained. Nevertheless, the SCEC synthetic data sets have been useful for evaluating and improving the detection method. Here we summarize strengths and weaknesses of the method learned from the test exercise.

1. Common mode errors can be mitigated by translation, rotation and scaling of the network. Stable sites were selected as reference frame sites but those sites within the transient deforming region were excluded.

2. Although random-walk and flicker noise were used for temporally correlated noise in the synthetic data sets, the FOGM state estimates can represent the temporally correlated noise as well as transient signals. Furthermore, FOGM variance estimates account for the uncertainties due to temporally correlated noise.

3. If a large signal exists, the FOGM state estimates can be over-smoothed and the FOGM variance overestimated. This can be partly solved by increasing process noise variance in the signal time interval and excluding the interval for the FOGM variance.

4. PCA considers spatial coherence of a transient signal. As more number of stations record a coherent signal, PCA tends to reveal the signal in first a few PCs.

5. Some of stations have anomalously large spatial amplitudes compared with nearby stations, mostly due to site-specific noise. Such stations should be removed before PCA because they dominate in PCs (often in the first PC).

6. PCs tend to have low-frequency oscillations when data contain long temporal correlation. The correlation effect can be mitigated by fitting and removing oscillations with sinusoids with frequency increasing with PC ranking or by reducing data dimension to smaller size in space and time so as to focus on the transient interval.

7. Uncertainties of PCA results are obtained from the state estimation errors and the FOGM variance estimates. Based on the uncertainties, a transient signal can be identified by chi-square test and lag differences of PCs.

8. Large-scale signals in space cannot be easily detected because they are difficult to separate from common mode errors.

9. Multiple signals can be detected if the SNRs are high, but they appear at multiple PCs, making interpretation complicated. If multiple signals are not separated in space and time, PCA will produce mixed effect or larger signal can mask weaker signal.

10. As in the case of multiple signals, PCA can detect propagating signals but they also appear at multiple PCs, making interpretation of PCA results complicated.

11. Missing data interpolation allows PCA to retain more data in space and time and thus increase the possibility of signal detection.

12. The estimates of white noise amplitudes are close to the true value in the mean sense. Misfits may be due to incorrect noise model; random-walk and flicker noise was used in simulation but the FOGM process was used in estimation.

Acknowledgements

This study used the FAKENET program provided by Duncan Agnew.

Chapter 3

Transient Inflation at Akutan Volcano, Alaska, During Early 2008[†]

Abstract

Continuous Global Positioning System (GPS) networks record station position changes with millimeter-level accuracy and have revealed transient deformations on various spatial and temporal scales. However, the transient deformation may not be easily identified from the position time series because of the large number of sites in a network, low signal-to-noise ratios (SNR) and correlated noise in space and time. Here we apply state estimation and principal component analysis to the daily GPS position time series measured in Alaska sites of the Plate Boundary Observatory network. Our algorithm detects a transient signal, whose maximum displacement is ~9 mm in horizontal and ~11 mm in vertical, that occurred at Akutan volcano during the first half of 2008. A simple Mogi source inversion suggests inflation at shallow depth (~3.9 km) beneath the volcano. Although the detection was not easy because the signal was aseismic, non-eruptive and weak (not apparent in raw daily time series), our detection method improves the SNR and therefore provides higher resolution for detecting the transient signal.

[†] Parts of this chapter were published as: Ji, K. H., and T. A. Herring (2011), Transient signal detection using GPS measurements: Transient inflation at Akutan volcano, Alaska, during early 2008, *Geophys. Res. Lett.*, 38, L06307, doi:10.1029/2011GL046904.

3.1 Introduction

Continuous Global Positioning System (GPS) networks, e.g., Plate Boundary Observatory (PBO; http://pbo.unavco.org/), provide data that allow a broad temporal and spatial spectrum of millimeter-level surface deformations to be studied. GPS position time series have revealed transient deformation due to earthquakes, volcanic activities, groundwater transport, and other processes. The large size of the networks, however, makes it time consuming to visually inspect each GPS time series to identify transient signals. Furthermore, GPS time series suffer from spatially [e.g., *Wdowinski et al.*, 1997; *Dong et al.*, 2006] and temporally correlated noise [e.g., *Williams et al.*, 2004; *Langbein*, 2008] that may mask transient signals depending on signal-to-noise ratios (SNR) of the transient. A transient signal detector needs to handle data from large networks while accounting for the correlated noise.

We apply our detection algorithm to the PBO daily GPS position data collected in Alaska (Figure 3-1). The algorithm is based on state estimation in a Kalman filter formulation and principal component analysis (PCA) (see Appendices for technical details). State estimation improves the SNR in the time domain by estimating secular velocity, seasonal sinusoids, time-correlated noise, and transient signals if any. PCA improves the SNR in the space domain by accounting for the spatial coherence of transient signals. PCA transforms the complex spatiotemporal structure of data into uncorrelated principal components (PCs) for the temporal variations and orthogonal sample eigenvectors for the spatial distributions. Due to the simple space-time separation and other advantages, PCA has been widely used in crustal deformation studies [e.g., *Savage and Langbein*, 2008; *Savage and Svarc*, 2009; *Kositsky and Avouac*, 2010].

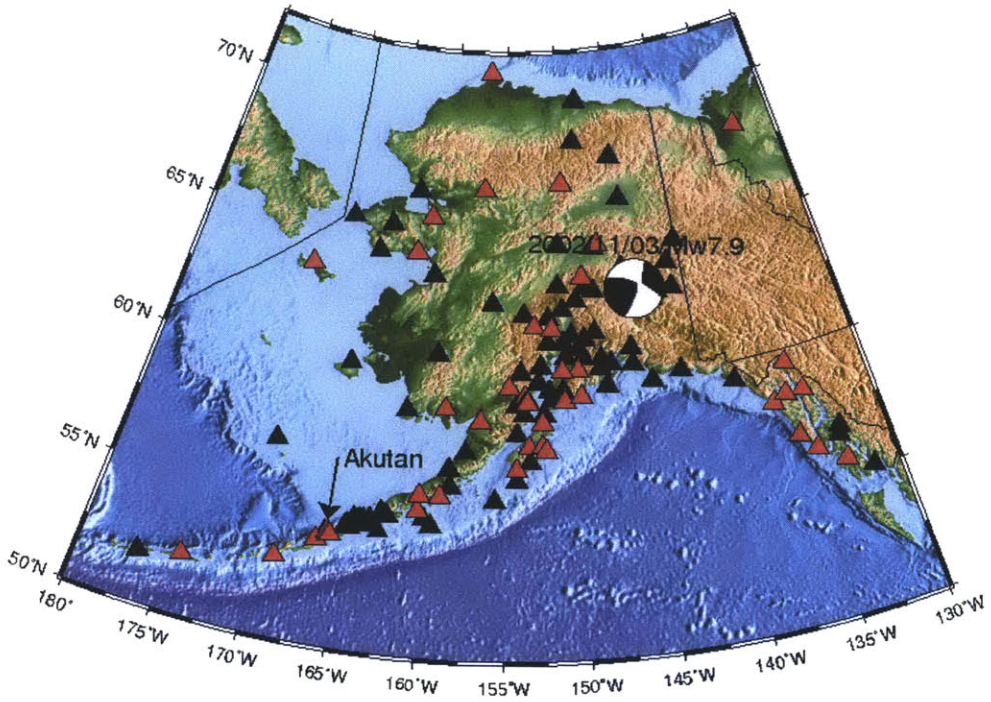


Figure 3-1. Total 151 PBO GPS sites (triangles) within the range of latitude 50°N ~ 75° and longitude 180°W ~ 130°W. The 48 sites (red triangles) are the reference frame sites used for the reference frame transformation. The distribution of the reference frame sites is relatively uniform over the network. Postseismic deformation following the 2002 Mw 7.9 Denali earthquake (focal mechanism) appears in the data set.

The algorithm has detected a transient signal that occurred at Akutan volcano, ~1224 km southwest of Anchorage, during the first half of 2008. The transient signal is consistent with a volcanic inflation source. The signal, however, is not easy to identify in the original data (Figure 3-2) because of the low SNR and highly correlated noise.

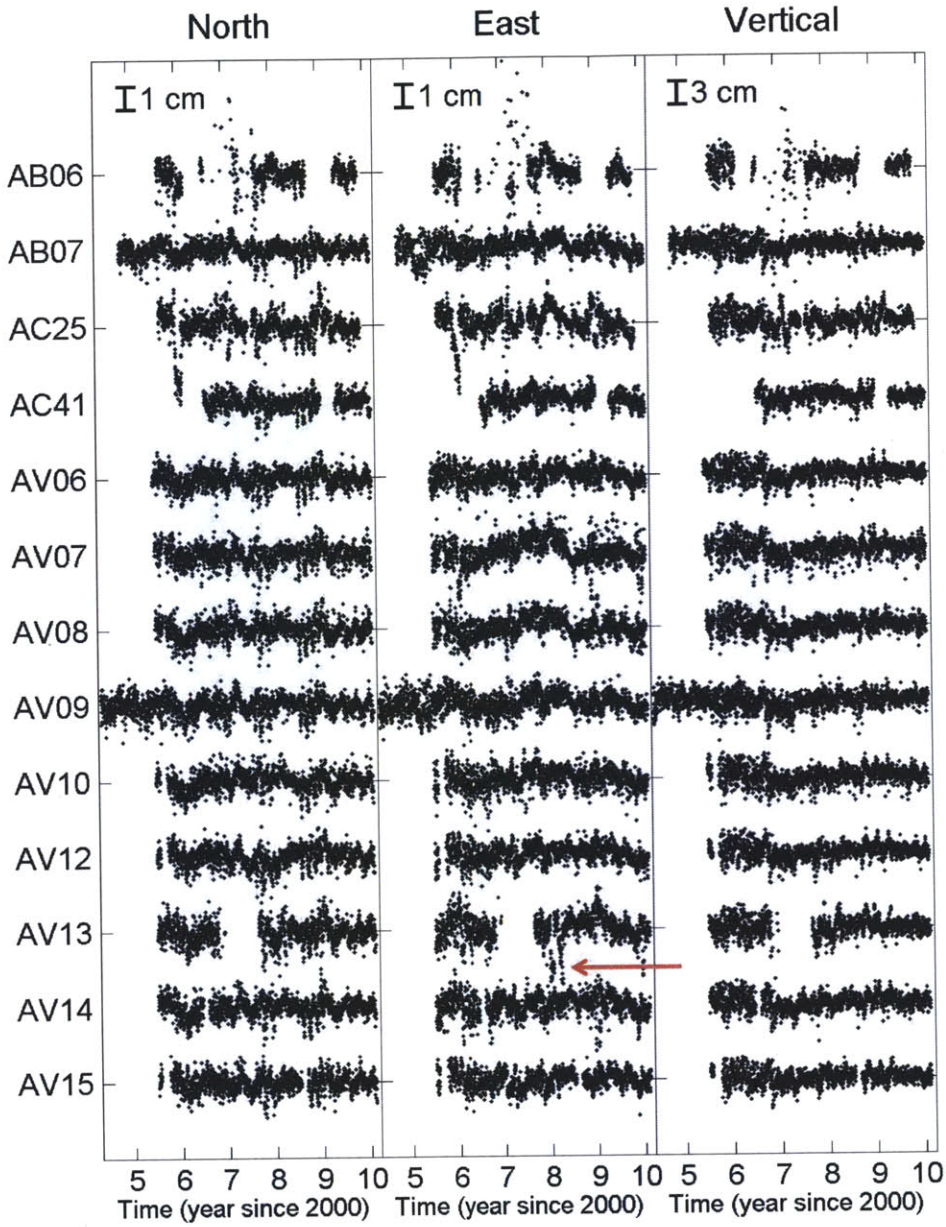


Figure 3-2. GPS time series around the area of Akutan volcano after removing linear rates and annual and semi-annual sinusoids. Formal uncertainties are not shown for clarity. They are typically 1 mm horizontal and 3 mm vertical. The time series are spatially and temporally correlated. Common patterns between the sites (e.g., around late 2008 in north component) can be found in each component, but they exist at most Alaskan sites (see Figure 3-3a). These time series actually contain the transient signal we detected, but spatial and temporal extent of the signal cannot be easily seen from the time series because of low signal-to-noise ratio and correlated noise in space and time. The station AV13 was not included in PCA because of the outliers (indicated by the red arrow) in the east component.

3.2 Data and Method

The PBO GPS analysis consists of two stages (details can be found at http://pboweb.unavco.org/?pageid=101). First, two analysis centers, Central Washington University using GIPSY/OASIS software [*Webb and Zumberge*, 1997] and New Mexico Institute of Mining and Technology using GAMIT/GLOBK software [*Herring et al.*, 2009], process phase and pseudorange data to estimate station coordinates averaged over 24-hour durations. The Analysis Center Coordinator (Massachusetts Institute of Technology) combines the results from the first stage and aligns them to the PBO realization of the Stable North America Reference Frame (SNARF; see http://www.unavco.org/research_science/workinggroups_projects/snarf/snarf.htm). Final solutions are available with 14-20 day latency due to the availability of GPS orbits. The standard PBO GPS level-2 products are available from the PBO Archives. We used daily position time series of 151 Alaska sites between 1 January 2004 and 13 February 2010. The sample sizes are between 220 and 2235 with an average of 1065 ± 540.

Our detection algorithm uses a smoother and PCA. The state vector in the smoother includes a secular rate, annual and semi-annual sinusoids, and a first-order Gauss-Markov (FOGM) process used to account for temporally correlated noise and also transient signals, if any (see Appendix A). The estimates of the FOGM stochastic process are used as the input of the PCA. The FOGM process is not necessarily a physical process for both temporally correlated noise and transient signals. Other noise models have been reported such as random-walk, flicker, or power-law noise [e.g., *Williams et al.*, 2004; *Langbein*, 2008]. The FOGM process provides a stochastic model that can accommodate correlated noise and signals, and it can be easily implemented in Kalman

filters. When multiple FOGM processes are combined, they can approximate the spectral shape of many, but not all, noise processes [*Langbein*, 2008].

PCA decomposes the FOGM state estimates into principal components (PCs) for the time history and sample eigenvectors for spatial distribution (see Appendix D). We do not include the linear and sinusoidal components in the PCA. PCs have the unit of length and sample eigenvectors are dimensionless. Sample eigenvalues are the variances of PCs; larger sample eigenvalues imply more temporal variations in PCs. The PC uncertainties are obtained by error propagation. The sample eigenvector uncertainties are an asymptotic result derived with an effective sample size that accounts for temporal correlation. These uncertainties allow us to infer the significance of the PCA results using a conventional χ^2 test. When a PC contains a transient signal, chi-square per degree of freedom (χ^2_{dof}) of the PC will be larger than that expected due to Gaussian noise in the estimates. Significant deviations from random behavior can also be tested with lag differences of a PC i.e., differences between points of a PC at a given time lag.

3.3 Spatially correlated noise

GPS time series suffer from spatially correlated noise (Figure 3-2); often referred to as "common mode error" (CME) that may result from orbital, reference frame and large-scale atmospheric errors [*Wdowinski et al.*, 1997; *Williams et al.*, 2004; *Dong et al.*, 2006]. Discrimination between signal and noise in space can be made by the fact that tectonic sources produce localized and systematic pattern in space, while the CME has a relatively uniform pattern over large areas.

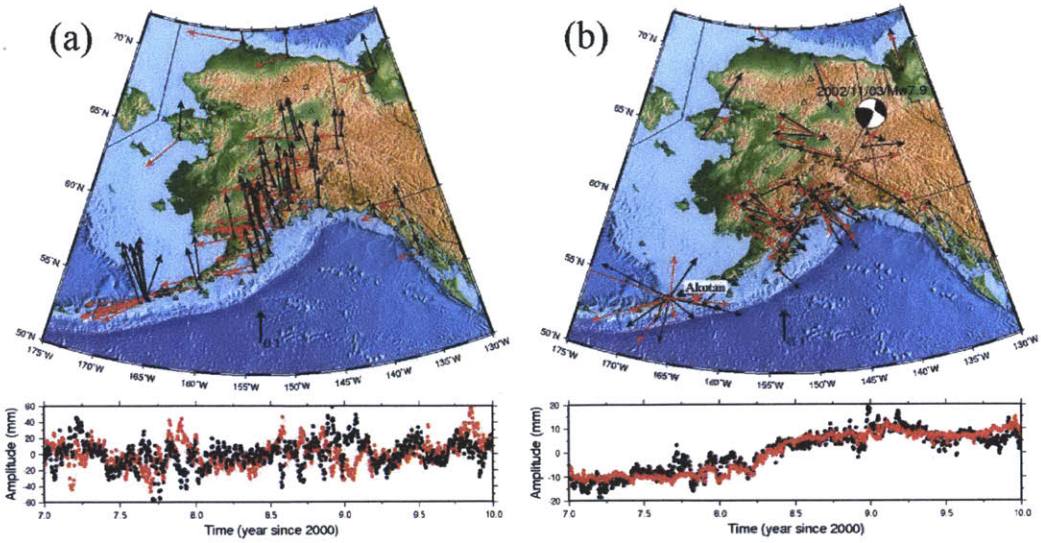


Figure 3-3. (top) Spatial and (bottom) temporal patterns of (a) the first PC (black; eigenvalue 28.4% and $\chi^2_{dof} = 19.5$) and the second PC (red; eigenvalue 26.8% and $\chi^2_{dof} = 14.4$) from the horizontal components of 58 sites (solid triangles) relative to SNARF and (b) the third PC (black; eigenvalue 7.0% and $\chi^2_{dof} = 3.1$) relative to SNARF and the first PC (red; eigenvalue 20.8% and $\chi^2_{dof} = 3.5$) relative to a regional reference frame. The 3-year time interval was selected for uniform site distribution and more samples in time. The two PCs in (a) show relatively uniform spatial patterns and random temporal patterns (the common noisy pattern in north component around late 2008 shown in Figure 3-2 appears in the first PC), while the two PCs in (b) show non-uniform spatial patterns. The PC in a regional frame provides clearer temporal pattern and reveals a signal early 2008. The scale arrows with length of 0.1 indicate 10 % of the PC amplitude. The PCs have units of millimeters. Actual displacements are the product of spatial and temporal amplitudes. The focal mechanism of the 2002 Denali earthquake is shown in (b).

As was found by *Dong et al.* [2006] for the sites in southern California, we found that the Alaskan sites also contain spatially correlated noise (Figure 3-3a). The largest two horizontal PCs show two orthogonal CMEs whose temporal patterns are neither purely random nor purely systematic and whose spatial patterns look uniform over the network. Even though the CME is a function of the wavelengths of various systematic errors [*Dong et al.*, 2006], the spatial uniformity of the CME is a valid assumption for the Alaska network, implying that the network size is smaller than the CME wavelength.

The CME may mask a transient signal if the SNR is low. As more number of sites are used in PCA, the CME more contributes to the variance [*Dong et al.*, 2006] such that a localized transient signal can be weaker than the CME in the PC space. To suppress the CME, one may simply ignore the PCs containing the CME and examine the next PC. Figure 3-3b shows that the third PC exhibits non-uniform spatial pattern but the temporal pattern is still noisy. Due to the orthogonality in PCA, the third PC may be contaminated by the first two PCs and vice versa.

Another approach to suppressing the CME is a reference frame transformation by translation, rotation and scaling of the network (i.e., a seven-parameter Helmert transformation). We transformed data in SNARF into those in a regional frame that was realized by 5 ~ 48 reference frame sites at each epoch, depending on the availability of data and the solution quality, whose spatial distribution is relatively uniform (see Figure 3-1). The velocity uncertainties of these sites are less than 0.45 mm/yr in horizontal and 0.9 mm/yr in vertical when velocity uncertainties are estimated accounting for correlated noise. The first PC shows a non-uniform spatial pattern and exposes a transient event starting in early 2008 and ending by mid-2008 (Figure 3-3b). For further investigation, we will use these data with the CME removed through a reference frame transformation.

For the vertical component, the first PC with respect to SNARF also contains large spatially correlated noise (Figure 3-4). The first PC shows that all sites move in the same direction and with similar size. After the reference frame transformation, the resulting first PC shows a non-uniform spatial distribution; the south-western sites are in the opposite direction to the other sites. Since the temporal pattern is noisy ($\chi^2_{dof} = 1.7$), it is not clear if there exists a vertical transient signal associated with the horizontal signal.

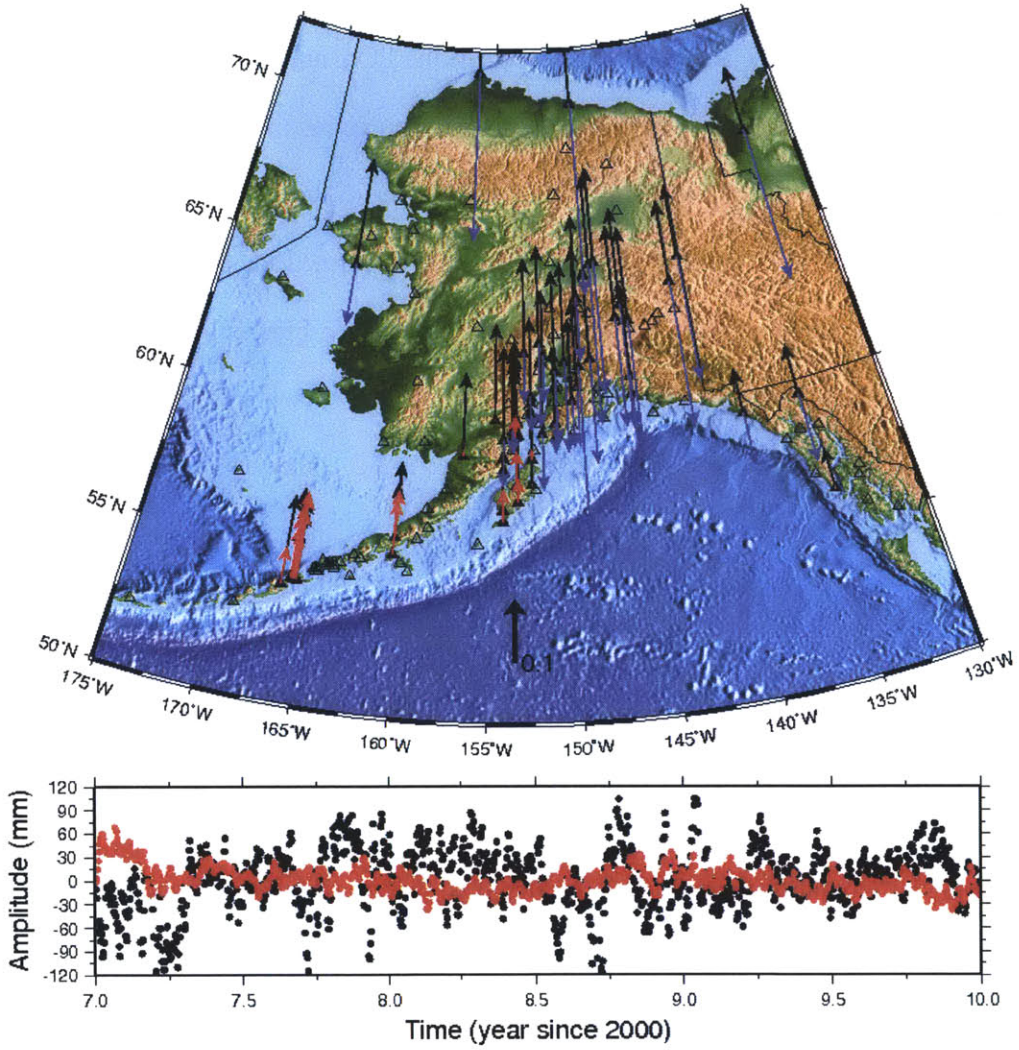


Figure 3-4. (top) Spatial and (bottom) temporal patterns of the first PC from the vertical components with respect to SNARF (black; eigenvalue 55.1% and $\chi^2_{dof} = 14.9$) and to a regional reference frame (in colors; eigenvalue 21.0% and $\chi^2_{dof} = 1.7$). The SNARF PC is uniform in space and noisy in time, suggesting common mode error (CME). The regional PC shows opposite motion between south-western sites (red) and the others (blue). Since the PC does not exhibit any systematic temporal pattern (red dots), the opposite sense of the vertical motion might be another type of CME rather than a transient signal.

3.4 Signal identification and significance

We have detected a transient signal during early 2008, but the location of the signal is not clear. The signal can be isolated by including in the PCA only those sites and the time interval that experiences the target signal. Because the maximum amplitude of the spatial pattern in Figure 3-3b occurred at the station AV07, one of the Akutan volcano sites, and sites in this region also show large amplitudes, we performed PCA with 7 sites on the Akutan Island (see Figure 3-5 for site location and also see Figure 3-2 for the raw time series). We used a one-year time interval starting at 1 October 2007 that spans the detected signal. The station AV13 was not included in PCA because of the outliers at the beginning of 2008 in the east component (see Figure 3-2), that are believed due to snow on the antenna radome. Figure 3-5a shows that the spatial and temporal patterns of the resulting first PC were similar to those shown in Figure 3-3b but with improved SNR (i.e., increase in χ^2_{dof} from 3.5 to 4.3). The sample eigenvalue also increased from 20.8% to 64.5% of the total variance of the data set, resulting in a more focused signal. The maximum horizontal displacement of the signal is ~9 mm, which is too small to detect by satellite radar interferometry (~3.5 mm line-of-sight displacement).

We were not able to identify from the vertical PC any signal related to the horizontal signal (Figure 3-5b). The vertical PC shows that all sites appear to move in the same direction with relatively uniform amplitudes, and the temporal pattern has a χ^2 value consistent with simply random noise ($\chi^2_{dof} = 1.0$). To examine if the identified transient signal also exists in vertical, we first fit an arctangent function to the horizontal PC (see Figure 3-5a and Table 3-1) and then obtained displacements by fitting the estimated arctangent function to the original time series with other parameters (see Figure

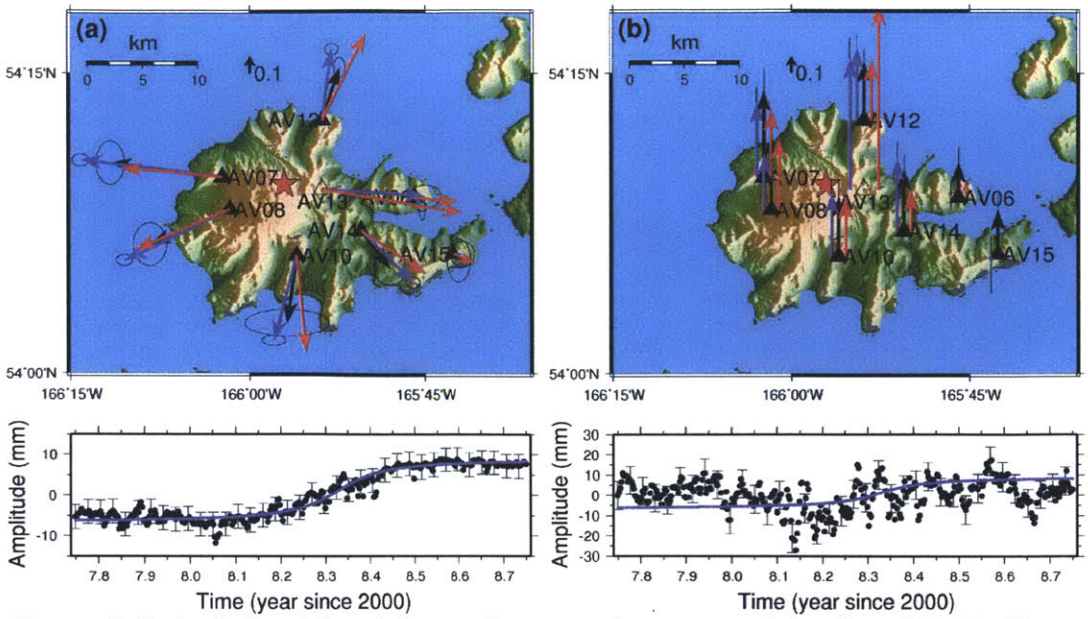


Figure 3-5. (top) Spatial and (bottom) temporal patterns of the first PC (black), an arctangent fit to the horizontal PC (blue curve), amplitudes from the fits of the estimated arctangent function to the original time series (blue arrows), and a Mogi source inversion (red arrows and red star) in (a) the horizontal and (b) the vertical components. The blue curve in (b) is the same as the blue curve in (a), not the fit to the vertical PC. All errors represent 1-sigma uncertainties. In the horizontal component, the three types of spatial patterns indicate outward expansion from the volcano center. On the other hand, the signal is not clear in the vertical PC (eigenvalue 52.2% and $\chi^2_{dof} = 1.0$), but the arctangent fits and the model prediction demonstrate the existence of the signal with amplitude (maximum ~11 mm at AV13) concentrated in the near field.

Table 3-1. Model parameters from the fit of $f(t)$ to the first PC from Akutan sites[a]

β_1 (mm)	β_2 (year)	β_3 (year)	β_4 (mm)
5.0836 ± 0.2665	2008.3387 ± 0.0093	0.0747 ± 0.0148	1.1347 ± 0.2232

[a] $f(t) = \beta_1 \tan^{-1}(\frac{t-\beta_2}{\beta_3}) + \beta_4$ where t is in year.

3-6 and Table 3-2). The horizontal displacements are similar to those from PCA but with slightly larger amplitudes and smaller uncertainties (Figure 3-5a). The vertical displacements are more concentrated in the near field (Figure 3-5b), which demonstrates that the vertical signal exists (also see Figure 3-6).

Table 3-2. GPS time series inversion results

		Rates	Offsets[b]	Annual (mm)		Semiannual (mm)		Arctangent[c]
		(mm/yr)	(mm)	Sine	Cosine	Sine	Cosine	
	N	-2.42(0.11)[a]	-0.59(0.43)	1.03(0.09)	-0.22(0.08)	0.94(0.08)	-0.15(0.08)	-0.05(0.02)
AV06	E	-4.03(0.14)	1.15(0.56)	-2.66(0.10)	1.18(0.10)	-0.69(0.10)	0.22(0.10)	0.04(0.02)
	U	5.03(0.36)	-3.35(1.34)	0.94(0.28)	-1.60(0.28)	0.40(0.28)	1.00(0.28)	0.01(0.06)
	N	-3.15(0.14)	-1.11(0.58)	-0.01(0.11)	0.57(0.11)	1.29(0.11)	-0.59(0.11)	0.10(0.02)
AV07	E	-2.71(0.19)	-4.04(0.78)	-1.61(0.14)	1.34(0.15)	-0.60(0.14)	-0.06(0.14)	-0.80(0.03)
	U	3.24(0.39)	3.63(1.64)	0.54(0.30)	-0.96(0.31)	-0.10(0.30)	0.09(0.30)	0.46(0.07)
	N	-0.79(0.13)	-2.37(0.55)	1.28(0.10)	0.58(0.10)	1.16(0.10)	-0.27(0.10)	-0.29(0.02)
AV08	E	-3.03(0.18)	-3.46(0.77)	-2.17(0.14)	1.08(0.14)	-0.44(0.14)	0.03(0.14)	-0.60(0.03)
	U	3.54(0.43)	2.87(1.77)	0.29(0.33)	-0.15(0.33)	0.02(0.33)	0.21(0.33)	0.34(0.07)
	N	-1.96(0.12)	-3.17(0.71)	1.16(0.09)	0.16(0.09)	0.72(0.09)	-0.62(0.09)	-0.49(0.02)
AV10	E	-3.84(0.26)	-0.56(1.38)	-2.97(0.21)	1.09(0.21)	-0.14(0.21)	0.43(0.21)	-0.13(0.05)
	U	4.99(0.34)	-3.31(2.02)	1.55(0.26)	-1.44(0.26)	0.72(0.26)	1.27(0.26)	0.39(0.06)
	N	-0.82(0.13)	3.68(0.75)	0.40(0.10)	0.23(0.10)	1.22(0.10)	0.25(0.10)	0.41(0.02)
AV12	E	-2.62(0.16)	1.97(1.04)	-2.68(0.12)	0.69(0.12)	-0.86(0.12)	-0.14(0.12)	0.05(0.03)
	U	5.25(0.35)	0.59(2.04)	1.65(0.27)	-1.43(0.27)	0.38(0.27)	0.57(0.26)	0.44(0.06)
	N	-2.65(0.34)	-0.17(0.67) -2.79(4.66) 2.68(4.64)	1.26(0.18)	0.47(0.18)	1.42(0.17)	0.71(0.17)	-0.03(0.04)
AV13	E	-5.16(0.48)	5.60(0.97) 6.54(6.18) -3.43(6.16)	-2.13(0.25)	1.27(0.26)	-0.73(0.25)	-0.49(0.25)	0.60(0.06)
	U	5.30(0.72)	5.58(1.56) 15.93(10.90) -15.62(10.87)	1.59(0.37)	-1.14(0.39)	0.51(0.38)	0.87(0.37)	0.80(0.09)
	N	-1.84(0.15)	-2.63(0.44) -1.23(0.28)	0.88(0.08)	0.76(0.09)	1.08(0.08)	-0.64(0.09)	-0.29(0.02)
AV14	E	-4.88(0.18)	1.91(0.59) 2.13(0.35)	-2.56(0.11)	1.02(0.11)	-0.72(0.10)	0.41(0.11)	0.30(0.03)
	U	3.32(0.48)	-1.46(1.36) 2.09(0.89)	1.88(0.27)	-1.67(0.28)	-0.11(0.27)	2.14(0.27)	0.44(0.07)
	N	-2.55(0.27)	0.15(1.04) -0.93(0.42) -0.41(0.63)	0.69(0.15)	0.28(0.14)	1.81(0.14)	0.11(0.14)	-0.00(0.05)
AV15	E	-5.49(0.27)	-3.50(1.24) -3.43(0.41) -1.12(0.61)	-2.58(0.14)	1.42(0.14)	-0.59(0.14)	-0.02(0.14)	0.06(0.05)
	U	4.31(0.65)	-5.37(2.66) 4.85(0.99) 1.38(1.48)	1.83(0.34)	-1.13(0.34)	-0.07(0.34)	0.36(0.33)	-0.06(0.12)

[a]-2.42(0.11) = -2.42 ± 0.11 where the uncertainties were obtained by propagating formal errors and FOGM variance estimates.

[b]Offsets occurred in 2005/09/24 at all sites, 2007/08/22 and 2007/08/26 at AV13, 2006/07/26 at AV14, 2007/01/04 and 2008/08/25 at AV15. Multiple offsets appear in chronological order. Most offsets are due to receiver/antenna changes except for 2007/01/04 offset which was visually detected.

[c]The arctangent model can be found in Table 3-1 with $\beta_4 = 0$ and the fits to the GPS time series are shown in Figure 3-6. The amplitude estimates are unitless.

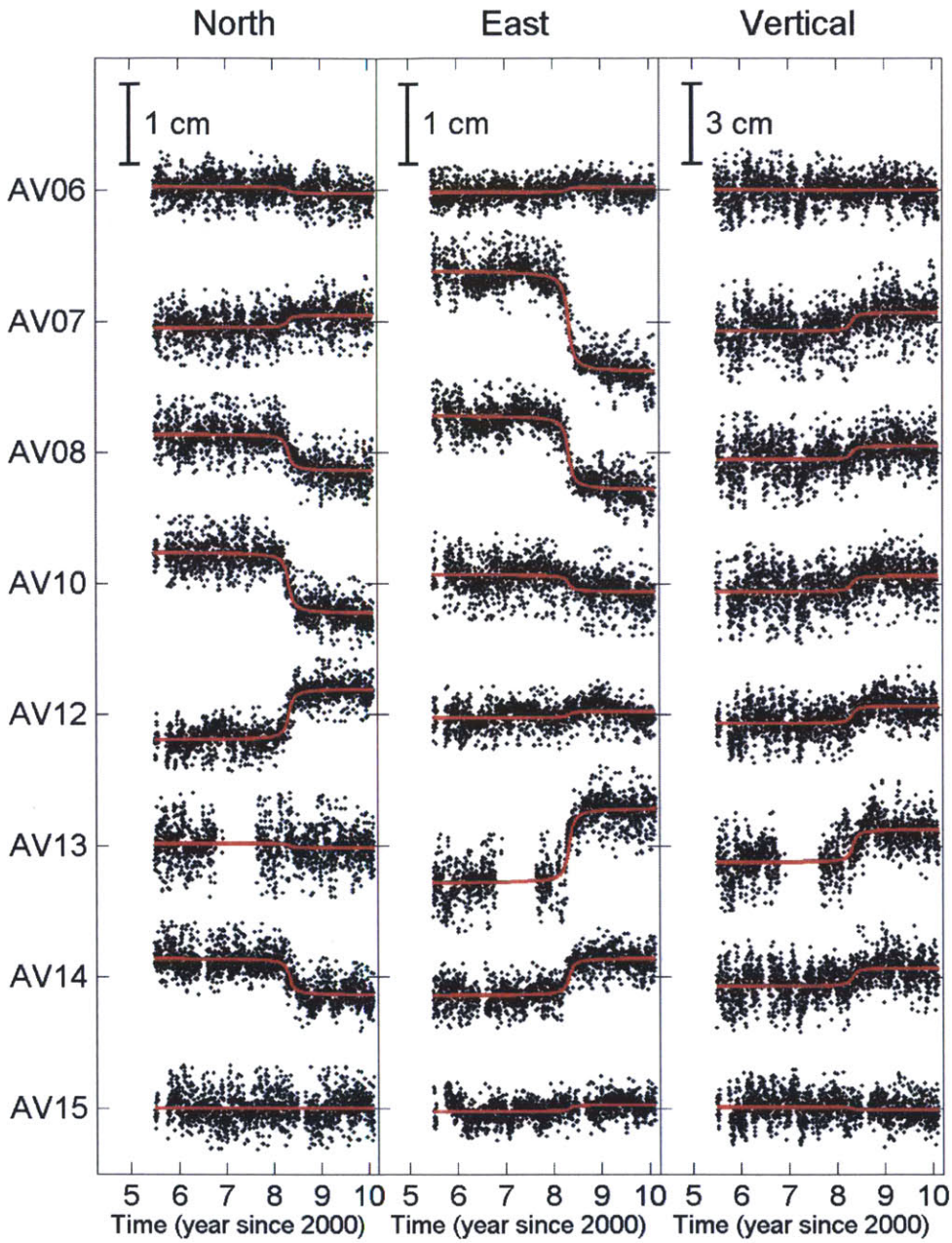


Figure 3-6. GPS time series fits of the arctangent function (red curves) obtained from the horizontal PC (see Table 3-1 for parameter estimates). Secular rates, annual and semiannual sinusoids, and offsets are removed from the original time series. The remaining time series (black dots) are clearer than those shown in Figure 3-2 because of the reduction of spatially correlated noise by the reference frame transformation and also more rigorous outlier detection due to the arctangent model for transient signal. The arctangent fits show that the vertical component also exhibits the transient signal identified from the horizontal PC. The displacements inferred from the arctangent fits between 2008.1 and 2008.5 are shown by blue arrows in Figure 3-5.

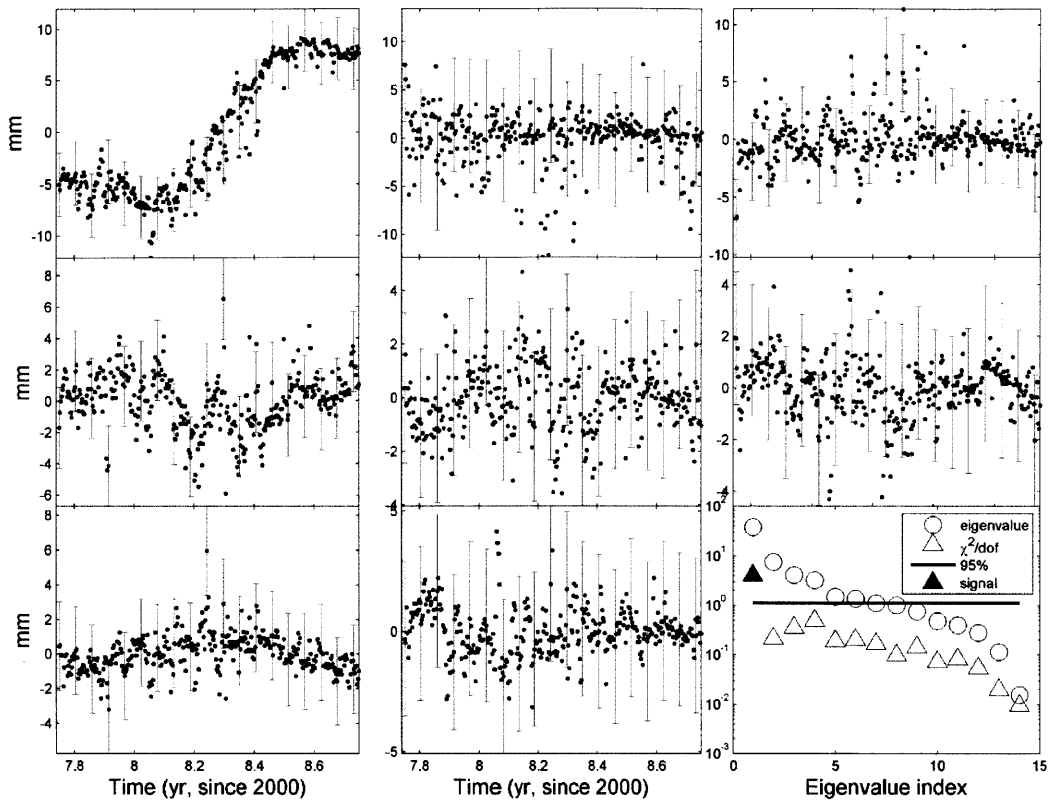


Figure 3-7. Top eight principal components (PCs) of total 14 PCs from the horizontal components of the Akutan volcano sites. The 1-sigma PC uncertainties are drawn every 20 points. The first PC (upper-left) shows a systematic temporal pattern, while the others look random. Also shown in the lower-right subplot are the sample eigenvalues (circles) and chi-squares per degree of freedom (χ^2_{dof} ; triangles). The first PC is the only component (solid triangle) whose $\chi^2_{dof} = 4.3$ is larger than the 95% upper limit of 1.124 (horizontal line).

The significance of the transient signal can be determined in several ways. Firstly, the first horizontal PC is the only PC having a systematic temporal pattern (Figure 3-7). Secondly, the PC explains ~65% of the total data variance. Thirdly, $\chi^2_{dof} = 4.3$ of the PC is larger than the 95% upper limit expected if the signal were simply random noise (Figure 3-7). Fourthly, lag differences of the PC show that early 2008 contains significant anomalies for lags of more than 100 days (Figure 3-8). Fifthly, the baseline between

stations AV07 and AV14 shows dominant motion in east-west direction, and the baseline between stations AV12 and AV10 shows dominant motion in north-south direction (Figure 3-9), which is consistent with the radial horizontal pattern. Finally, the systematic radial pattern shows spatial coherence of the signal.

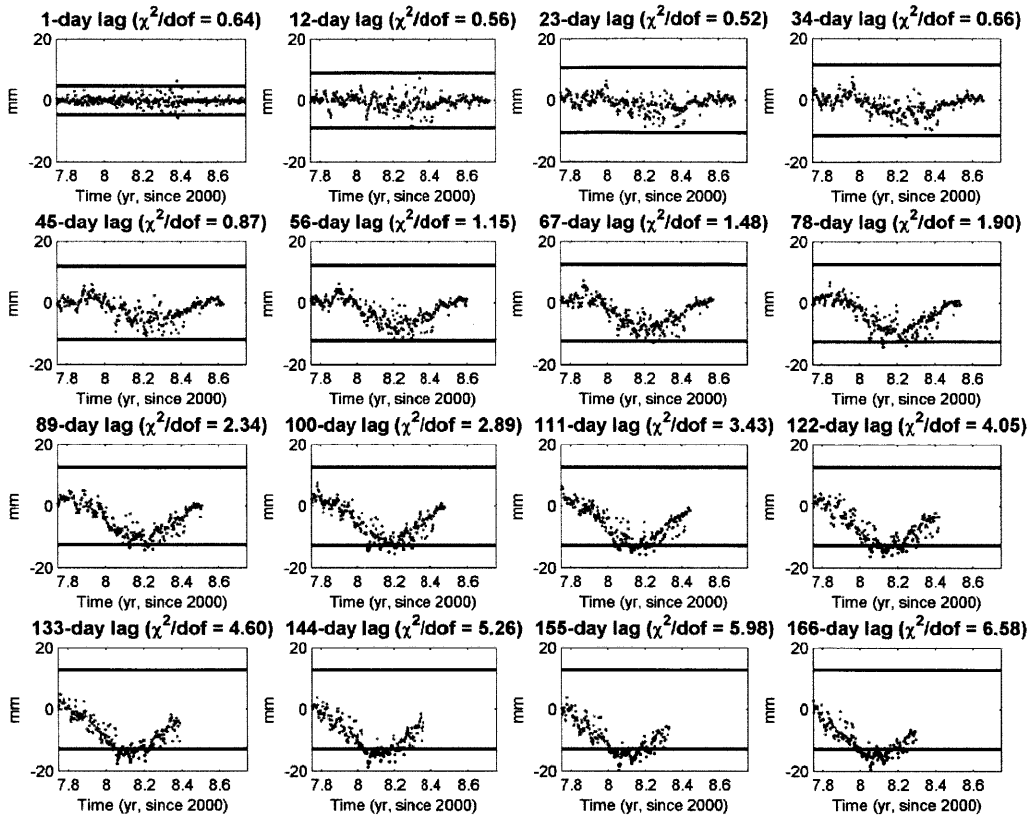


Figure 3-8. Lag differences of the first principal component (PC), i.e., differences between points in the first PC at given time lags. A total of 16 cases are shown with different time lags and chi-squares per degree of freedom (shown at the top of each subplot). The horizontal line indicates 3-sigma uncertainty. For the time lags larger than 100 days, the time period of early 2008 contains anomalies (black dots) larger than the 3-sigma uncertainty. For shorter time lags, on the other hand, the lag differences are not significant (gray dots) with respect to the 3-sigma uncertainty.

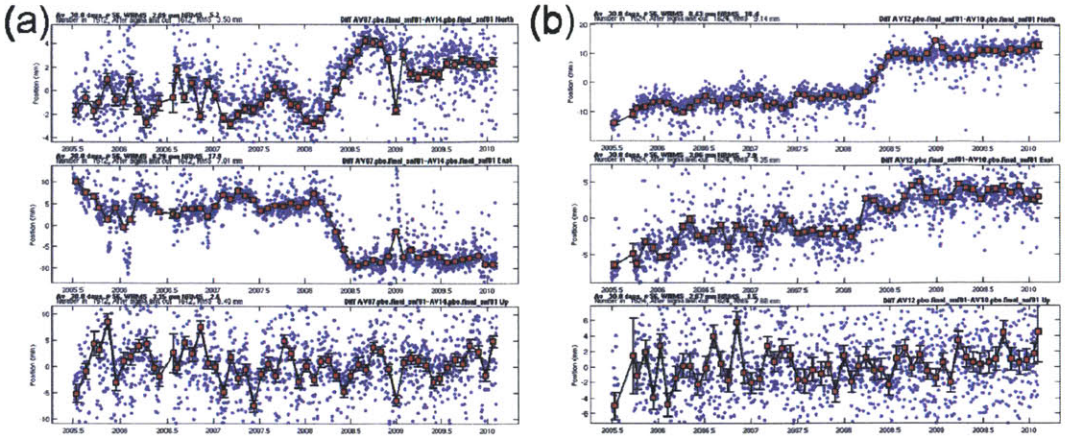


Figure 3-9. Three components (north, east and up from the top) of baselines between stations (a) AV07 and AV14, and (b) AV12 and AV10 (see Figure 3-5 for site location). The baselines were calculated from the positions relative to SNARF. The scale of vertical axes is different in each component. During the first half of 2008, the relative displacements are dominant (a) in east component and (b) in north component, which is consistent with the radially expanding pattern of the identified transient signal. For clarity, we calculated 30-day averages (red squares) with uncertainties. The vertical component shows no relative motion between the stations.

One may argue from Figure 3-3b that the PC temporal pattern exists everywhere, not only at Akutan Island. However, the signal in other area is not real. For example, Figure 3-3b shows a strike-slip type of pattern in the area affected by the 2002 M7.9 Denali earthquake [e.g., *Hreinsdottir et al.*, 2003; *Freed et al.*, 2006]. When we performed PCA with only the sites near the 2002 Denali earthquake epicenter area and one-year time interval in which the detected signal concentrates, the resulting first PC ($\chi^2_{dof} = 0.3$) shows a steady linear temporal pattern and nearly random spatial pattern (Figure 3-10). Furthermore, the larger error statistics in both space and time indicate considerable noise contribution to PCA. This suggests that the Denali area is not responsible for the detected signal starting in early 2008.

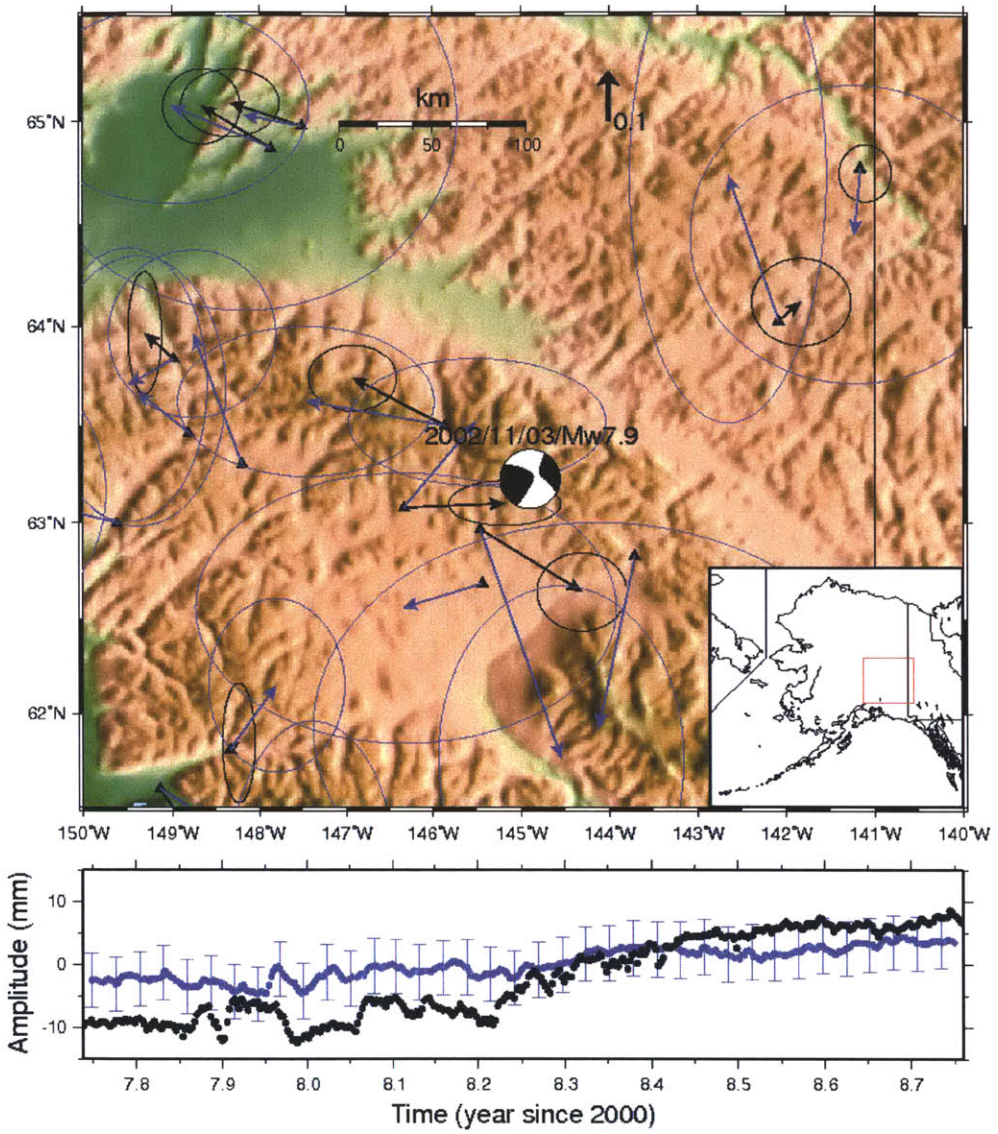


Figure 3-10. (top) Spatial and (bottom) temporal patterns of the first PC (blue; eigenvalue 24.1% and $\chi^2_{dof} = 0.3$ (low SNR)) from 15 sites near the 2002 Denali earthquake epicenter (focal mechanism) and the first PC shown in Figure 3-3b (black). Five more sites are available for PCA with smaller data dimension. PCA with only Denali sites show spatial and temporal patterns different from those in large-scale PCA results. As a result, the identified signal (black time series) is not real in the Denali area. The linear temporal pattern (blue time series) may be a part of the postseismic deformation following the Denali earthquake or local linearity in time remained after secular rate estimation.

3.5 Discussion

Akutan volcano is one of the most active volcanoes in the Aleutian arc. The most recent eruption was a series of small steam and ash emissions from March to May 1992 [*Lu et al.*, 2000]. The identified transient signal may be driven by volcanic sources (e.g., inflation or pressurization of magma bodies beneath the volcano). We used a Mogi point source [*Mogi*, 1958] to model the 3-dimensional surface displacements obtained from the arctangent parameterization (Figure 3-5 and Table 3-3). The estimated Mogi source locates near the volcano summit at a shallow depth (~3.9 km below sea level) and explains most of the identified signal. The model prediction indicates that the uncertainties from PCA are more realistic than those of the arctangent fits. The uncertainties of the arctangent fits are too small because they are based on a white noise assumption whereas the PCA uncertainties account for temporally correlated noise. The station AV13 was not included in the source inversion. Its predicted displacements differ ~2.5 mm in horizontal and ~4 mm in vertical, which are consistent with their uncertainties. If the station AV13 is included in the source inversion, the horizontal and vertical differences decrease to ~0.7 mm and ~1.3 mm, respectively, and the source locates slightly deeper (~4.3 km; see Table 3-3).

Table 3-3. Mogi source inversion results

AV13	Latitude (degree)	Longitude (degree)	Depth (km)	Volume change ($\times 10^6$ m^3)
Not included	54.157 ± 0.009	-165.952 ± 0.015	3.89 ± 1.67	2.1 ± 0.6
Included	54.157 ± 0.008	-165.952 ± 0.012	4.32 ± 1.29	2.0 ± 0.6

Using satellite radar interferometry, *Lu et al.* [2000] mapped ground cracks of more than 60 cm uplift on the northwest flank of the volcano associated with the March 1996 earthquake swarm and constrained the deformation as a composite source: shallow dike intrusion for the deformation associated with the ground cracks and a deep Mogi source for volcano-wide inflation. The Mogi source is deeper (~13 km) than our estimates (~3.9 km). Based on geodetic measurements, the source depth at nearby volcanoes has been inferred to be at less than 10 km, for example, ~3 km at Okmok volcano [*Biggs et al.*, 2010], ~7 km at Makushin volcano [*Lu et al.*, 2002a], ~6 km at Westdahl volcano [*Lu et al.*, 2003], and <10 km at Shishaldin volcano [*Moran et al.*, 2006].

The 2008 inflation signal was aseismic. The inflation was not associated with increase in sesimicity [*Dixon and Stihler*, 2009]. The number of located earthquakes in 2008 was twice that in 2007, but at a similar level of seismicity prior to 2007. Furthermore, there was only one earthquake of magnitude 2 recorded while the inflation was ongoing [see *Dixon and Stihler*, 2009, Figure A21].

The 2008 inflation event did not culminate in an eruption, and there were no inflation or deflation events in 2007 or in 2009. The 2008 inflation may be an episode that accumulates magma in a shallow reservoir. Many volcanoes inflate in an episodic fashion [e.g., *Lu et al.*, 2002b; *Fournier et al.*, 2009]. When a threshold is reached through a series of episodes, magma will eventually erupt. The inflation associated with the March 1996 earthquake swarm can be regarded as an episode. Continuous GPS monitoring of volcano deformation can constrain the characteristics of the eruption cycle.

3.6 Conclusions

We have detected a transient signal at Akutan volcano during the first half of 2008. A simple Mogi source at a shallow depth (~3.9 km below sea level) can explain most of the displacement field. The event detected could in part of an inflationary sequence that will eventually lead to an eruption. There is no evidence for deflation of the volcano during the time span of the data analyzed here (2005-2010). The aseismic and non-eruptive inflation signal is difficult to detect in raw time series because of low SNR, the large size of the network and correlated noise in space and time. State estimation and PCA improved the SNR and successfully revealed the transient signal imbedded in the original time series. Our detection algorithm can provide higher resolution for accurately detecting transient signals and the resulting solutions can be further used for modeling and interpreting the transient deformation.

Chapter 4

Detection of small slow slip events in Cascadia

Abstract

Slow slip events (SSEs) have occurred in subduction zones at roughly regular intervals and are accompanied by seismic tremor bursts. Since smaller tremors have been observed between major episodes, smaller SSEs are expected to exist. However, SSEs with surface displacements of 3 mm or less are below current Global Positioning System (GPS) resolution. Here we use state estimation and principal component analysis (PCA) for detecting small SSEs from GPS position time series. Secular and transient motions as well as time-correlated noise are estimated by using a smoother based on Kalman filtering. PCA accounts for the coherence of SSEs at multiple stations and extracts variance-maximizing patterns from the estimates of transient motions. We applied the method to the GPS data collected in Washington State section of the Cascadia subduction zone and have detected two previously unrecognized small SSEs with average surface displacements less than 2 mm. The small SSEs are also accompanied by small tremors. The method improves the signal-to-noise ratio in both space and time and makes the level of GPS limitation lower. This method can be used for the GPS data in southwest Japan where short-term SSEs are small and have been observed only with tiltmeters.

4.1 Introduction

Episodic tremor and slip (ETS) events [after *Rogers and Dragert*, 2003] refer to transient signals in which non-volcanic tremors [e.g., *Obara*, 2002] and slow slip events (SSEs) [e.g., *Hirose et al.*, 1999; *Dragert et al.*, 2001; *Miller et al.*, 2002] accompany each other with regular recurrence intervals along subduction zones [*Schwartz and Rokosky*, 2007, and references therein]. Tremors are low-amplitude and low-frequency (1-10 Hz lower than those of small earthquakes) seismic signals that have noise-like, emergent waveforms whose envelopes are coherent between sites even separated by hundreds of kilometers, and last for a few minutes to days. SSEs, on the other hand, occur in the direction opposite to the plate convergence with duration and recurrence interval from days to years and with surface displacements from millimeters to centimeters. ETS events provide insight into the understanding of strain accumulation and release and thus seismic hazards in subduction zones.

All SSEs observed by Global Positioning System (GPS) have been associated with major tremor episodes but many smaller tremors occur more frequently without resolvable GPS measurements of SSEs [*Szeliga et al.*, 2008; *Aguiar et al.*, 2009]. In northern Cascadia, smaller tremors accommodate approximately 45% of the total duration of the tremors observed from 2006 through 2009 [*Wech et al.*, 2010]. Based on borehole strainmeter and (less obvious) GPS observation in northern Cascadia in November 2006, *Wang et al.* [2008] showed that shorter-duration tremor episodes accompanied an SSE in much smaller area and for shorter duration than other well-known major ETS events, suggesting that smaller tremors accompany SSEs but at levels below current GPS resolution [*Aguiar et al.*, 2009; *Wech et al.*, 2010].

Several search techniques for SSEs have been developed. For individual GPS time series, SSEs can be estimated in a least squares sense by fitting a hyperbolic tangent function in a moving time window [*Brudzinski and Allen*, 2007; *Holtkamp and Brudzinski*, 2010] or by fitting an offset at an epoch inferred from a Gaussian wavelet transformation [*Szeliga et al.*, 2008]. These techniques are useful for precise timing and displacement once a SSE has been identified, but they do not account for spatial coherence of SSEs between GPS sites in a systematic way. *Kimura et al.* [2011] used an automated detection algorithm based on the Network Inversion Filter [*Segall and Matthews*, 1997] for tilt data in Shikoku, Japan. The method accounts for the coherence of slip history and provides statistical evaluations for identification of SSEs, but it requires fault geometry for source parameter inversion.

We have developed a transient signal detection method: a combination of state estimation based on smoothing and principal component analysis (PCA) [*Ji and Herring*, 2011]. Based purely on observations without any source-specific model, the method improves the signal-to-noise ratio in time by reducing the level of both white and colored noise and in space by accounting for the spatial coherence of SSEs. We apply our detection algorithm to the daily GPS position data collected in the Plate Boundary Observatory (PBO) network located in the Washington State section of the Cascadia subduction zone in western US (Figure 4-1). Other than major ETS events, the algorithm has detected three small SSEs that are also coincident with tremors; one is a part of a major ETS event and the other two have not been previously recognized.

4.2 Data and Method

We used 7-year daily GPS data (1 January 2004 ~ 29 January 2011) measured at 58 stations between 46°N ~ 49°N and 120°W ~ 125°W (Figure 4-1). The sites on Mount St. Helens (~46.2°N and ~122.2°W) were not considered because they are noisy. The sample sizes are varying between 880 and 2586 with an average of 1826 ±634. The PBO time series are relative to the Stable North America Reference Frame (SNARF). To reduce possible common mode errors [*Wdowinski et al.*, 1997; *Dong et al.*, 2006], we re-stabilized the reference frame by transforming data in SNARF into those in a regional frame realized by 79 stable reference frame sites (Figure 4-1). The reference sites were chosen by those far from the subduction margin to avoid ETS events (i.e., stations east of 122°W) and those with velocity uncertainties less than 0.15 mm/yr in horizontal and 0.45 mm/yr in vertical when correlated noise is accounted for in the velocity estimation.

Our detection algorithm consists of a smoother based on Kalman filtering and PCA (see Appendices A and D and also see *Ji and Herring* [2011]). We apply a smoother to individual GPS time series with a state vector whose elements are a secular rate, annual and semi-annual sinusoids, and a first-order Gauss-Markov (FOGM) process. The FOGM process is used to account for temporally correlated noise [*Zhang et al.*, 1997; *Mao et al.*, 1999; *Williams et al.*, 2004; *Langbein*, 2008] and any transient signals such as SSEs. PCA extracts uncorrelated and variance-maximizing patterns (i.e., principal components (PCs) for temporal variations and sample eigenvectors for spatial distributions) from the FOGM state estimates. For the detection of SSEs, we use only horizontal components (i.e., north and east components) in PCA.

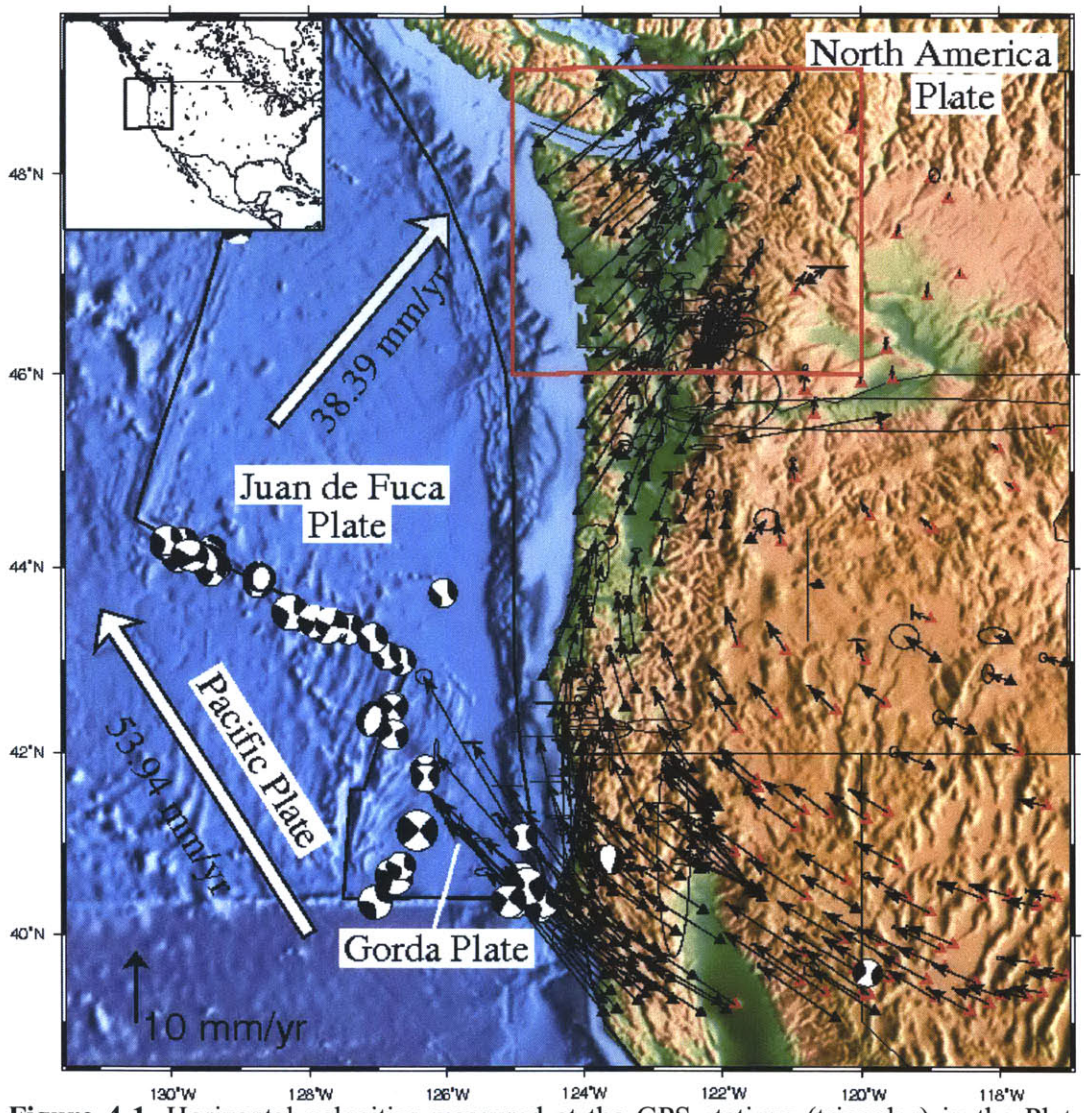


Figure 4-1. Horizontal velocities measured at the GPS stations (triangles) in the Plate Boundary Observatory network. The velocities are relative to the Stable North America Reference Frame (SNARF). The velocity field shows northwest motion below latitude ~42°N due to the relative motion between the North America plate and the Pacific plate and northeast motion in Cascadia due to the subduction coupling of the Juan de Fuca and Gorda plates with the North American Plate. The red box indicates the study area. The red triangles shows the stations used in the reference frame stabilization. The plate velocity is from MORVEL [*DeMets et al.*, 2010].

The significance of PCA results can be tested in several ways (see Appendix D). In time domain, chi-square tests can be performed for the overall temporal variation of a

PC based on the estimates of the PC uncertainty. For local temporal variations, lag-differences of a PC (i.e., differences between points of a PC at a given time lag) can be compared with their uncertainties. In space domain, the significance of sample eigenvectors can be tested with their uncertainty estimates. Furthermore, a systematic spatial pattern of a sample eigenvector must be attained (e.g., westward motion opposite to interseismic loading). Such conditions are often satisfied with the PC having the largest sample eigenvalue (denoted here as the first PC).

The method initially identified nine major SSEs (Figure 4-2) which have occurred for about 3 ~ 8 weeks (longer duration is due to migration of the SSEs) with maximum displacements in a range of 5 ~ 8 mm. The large SSEs and the strain accumulation between the SSEs could mask smaller events. We can reduce the effect of the large SSEs by removing data associated with the SSEs and including an offset for each SSE in the state vector. The resulting FOGM state estimates may reveal smaller SSEs. For example, the December 2005 event was weakly observed at the station SC02 compared to other major events (Figure 4-3a) because the station is located at the northern margin of the event. After removing the data associated with the six large SSEs, the event becomes the largest one among the remaining westward displacements (Figure 4-3b). Smaller westward motions can be seen but one cannot determine if they are real without considering coherence with neighboring sites. PCA accounts for the coherence based on the variance-covariance structure of the FOGM state estimates from multiple sites.

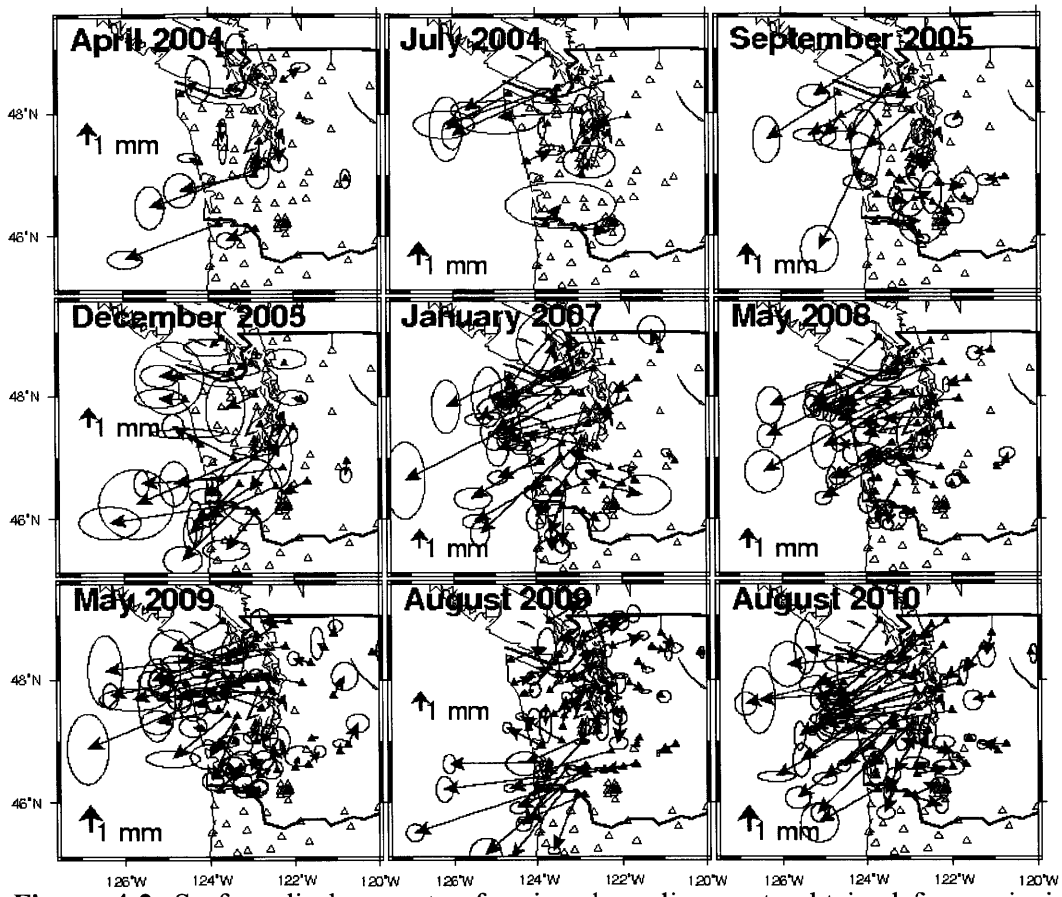


Figure 4-2. Surface displacements of major slow slip events obtained from principal component analysis (PCA). The directions of the events are roughly opposite to the plate convergence. Three events (April 2004, December 2005 and August 2009) dominate in southern Washington and also northern Oregon (sites not included in this study), while the others dominate mostly in northern Washington. Open triangles indicates PBO sites not used in PCA. Error ellipses represent 1-sigma uncertainties. Spatial resolution improves with time as site density increases.

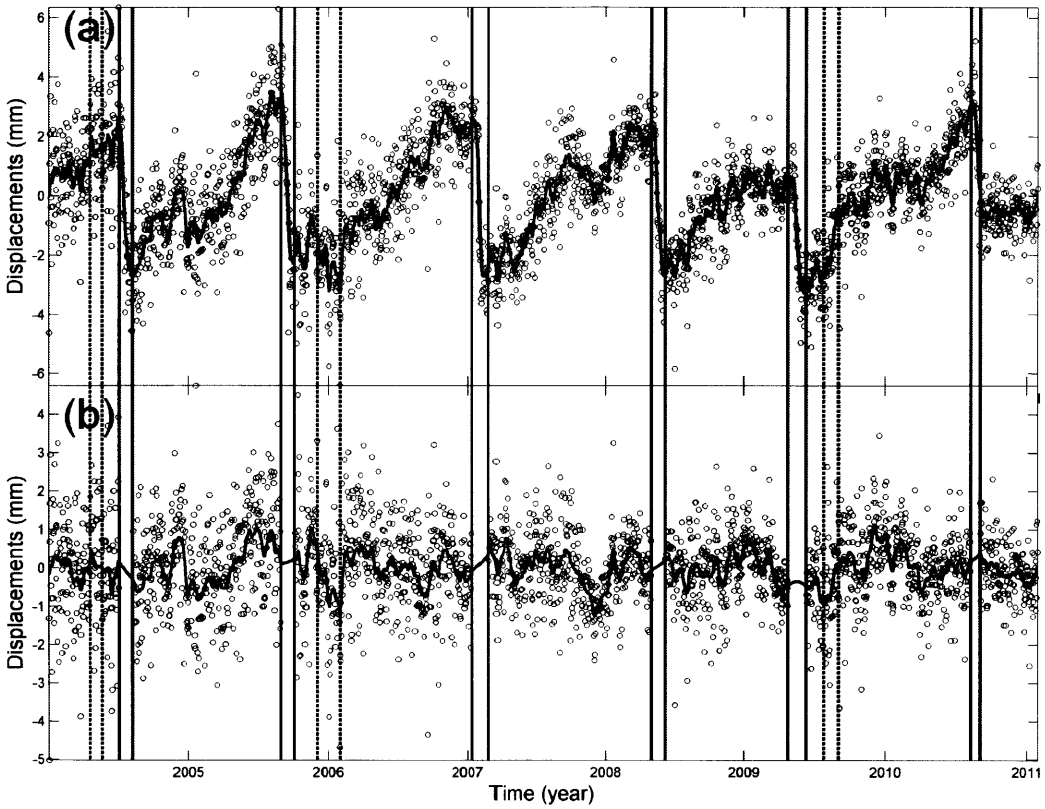


Figure 4-3. Detrended time series (circles) and the estimates of the first-order Gauss-Markov (FOGM) states (black lines) at the east (or longitudinal) component of the site SC02 (48.546°N and 123.008°W) in northern Washington. (a) Six major slow slip events (SSEs) are clearly seen between the time intervals delineated by vertical lines at about 14-month recurrence interval [*Szeliga et al.*, 2008]. Three SSEs delineated by vertical dashed lines are weak or absent at this site because they are dominant in southern Washington. The large SSEs and also strain accumulation between the SSEs may mask small events. (b) The FOGM state was re-estimated by removing data between the time intervals of the large SSEs and estimating linear trends, annual and semiannual sinusoids, and six offsets at each epoch in the middle of the event time intervals. The December 2005 event now shows the largest westward motion. Smaller westward displacements can be seen with amplitude of 2 mm or less (e.g., those around epochs of 2004.9, 2007.8, 2009.2, and 2010.2), but their significance cannot be determined without considering coherence with nearby stations.

PCA depends on the size of data in both space and time. When the data size is large, more amount of noise influence the PCA results. When the data size is small, PCA may detect SSEs partially or miss them completely. In this study, we divided the study area into northern and southern Washington (i.e., 47.5°N ~ 49.0°N and 46.0°N ~ 47.5°N) not only for smaller data set but also for possible segmentation of SSEs between the two subareas [*Brudzinski and Allen*, 2007]. For each subarea, PCA used data in a three-month time window moving at one-month interval. If an anomalous signal is detected, the data size is refined (possibly merged into the other subarea) and PCA is repeated. In this way, we can maximize the capability of the signal detection.

4.3 Results

We have detected a few small transient signals that resemble SSEs, but it is hard to claim that they are SSEs because of low amplitudes; maximum displacements are mostly less than 3 mm. Here we show three cases that are believed to be ETS events. Table 4-1 summarizes the information of the three cases.

Table 4-1. Information of the small signals detected by principal component analysis[a].

	Centroid time[b] (years)	Duration (years)	# of sites[c]	Maximum displacement (mm)	Average displacement (mm)	Sample eigenvalue (%)	χ^2_{dof} [d]
Case 1	2007.375	0.25	14	2.9	1.8	73	4.1
Case 2	2007.825	0.15	16	1.9	1.2	60	2.6
Case 3	2009.150	0.15	13	3.6	1.5	67	2.7

[a]Time and displacements were approximately obtained. See Figures 4-4, 4-5, and 4-8 for spatial pattern of the signals.
[b]Time in the middle of signal interval.
[c]Number of sites used in PCA.
[d]Chi-square per degrees-of-freedom larger than $\chi^2_{dof} \approx 1.2$ at 5% significance level for all the cases, which means that the temporal pattern cannot be simply explained by random noise.

In the first case, a small transient signal occurred in southern Washington for about three months between April and June 2007 (red vectors in Figure 4-4). The signal time interval was determined significant by the lag-differences of the PC temporal pattern with respect to 3-sigma uncertainty. The spatial pattern described by the sample eigenvector of the first PC shows that all sites move seaward similar to typical SSEs. In fact, the signal represents a northern part of the April 2007 event accompanied by significant tremor bursts mostly in northern Oregon (Figure 4-4). The maximum displacement of the event is ~7 mm, while the detected signal shows only ~3 mm so that it was not initially detected as those shown in Figure 4-2. This demonstrates that the method is capable of detecting small events having displacements of ~3 mm based on the coherence between 14 stations.

In the second case, an anomalous signal has been detected in northern Washington (Figure 4-5). The spatial pattern shows southwest motions similar to major SSEs in this region, particularly at those sites in the Puget Sound area (i.e., 47.5°N ~ 49.0°N and 122.0°W ~ 123.5°W). The signal lasted for about two months between October and November 2007. However, based on the lag-difference calculation, the signal is only marginally significant, which may be due to the small displacement (at most ~ 2 mm at the station ALBH) compared to the background noise.

In order to see if the PC temporal pattern exists in raw data, we first made the PC pattern smooth with an arctangent function (blue curve in Figure 4-5). Note that a part of the arctangent function produces a good fit to the PC temporal pattern (i.e., the function is not in the middle of the PC temporal pattern). We constructed a model to cover outside of the signal time interval by extending the value at the first epoch for the earlier times and

the value at the last epoch for the later times. Then we estimated the amplitude of the model as well as a linear rate from the raw data which only span from 1 August 2007 to 1 January 2008. The resulting spatial pattern is similar to the PC spatial pattern in northern Washington, although stations in southern Washington are noisy (Figure 4-5). In spite of small amplitudes, the model fit suggests that the PC temporal pattern exists in the raw data (Figure 4-6), not an artifact from PCA.

It may be possible that the detected temporal pattern is a common response to network-wide error sources, often called common mode errors, which could be unresolved after the reference frame re-stabilization. However, far-field stations do not exhibit such temporal pattern (Figure 4-7). This suggests that the signal is not a common mode error but a spatially localized signal.

Detection of small SSEs was done independently of the tremor information. We examined the tremor data [*Wech and Creager*, 2008; *Wech et al.*, 2009] to see if there is a spatial and temporal correlation between the identified signal and tremors. Anomalous tremors occurred between 2007.75 and 2007.8 for the maximum of ~12 hours/day (Figure 4-5), which is far above the median value (~3.5 hours/day) from all tremor data available in this region. The tremors were located in the Puget Sound area where the signal shows westward motion determined significant by PCA. Although the tremor activity did not cover the whole time interval of the signal, the association of tremors with the signal supports that the signal represents an ETS event.

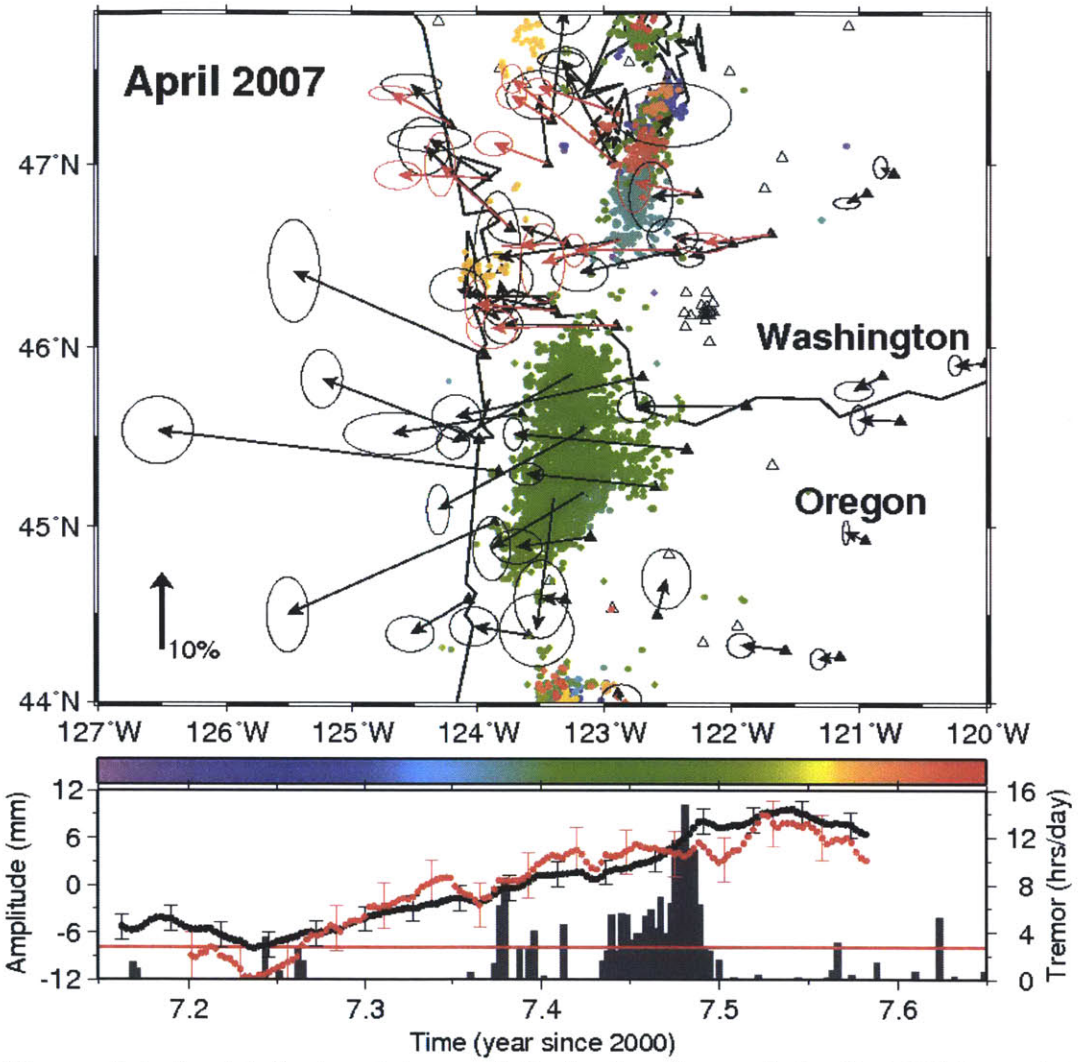


Figure 4-4. Spatial (top) and temporal (bottom) patterns of the first PC from data between 44.0°N and 47.5°N (black vectors and dots) and from data between 46.0°N and 47.5°N (red vectors and dots). The PCA results from the large-scale data represent a major slow slip events occurred mostly in northern Oregon with maximum displacement of ~7 mm. The event was not initially detected with only sites in Washington State because of small amplitude (maximum of ~3 mm). Also shown are the time history (histogram) and locations (colored dots with horizontal color scale) of the tremors [*Wech and Creager*, 2008; *Wech et al.*, 2009]. The red horizontal line indicates the median value of tremors (~3.5 hrs/day) between 1 January 2006 and 22 May 2011. The tremors are well correlated with the SSE in space but do not cover the whole time interval of the SSE. Error ellipses and bars a 1-sigma uncertainty. The red vectors are significant at 5% significance level. The error bars are drawn at every 10 days for simplicity. The black vertical arrow indicates 10% of the displacement obtained from the temporal pattern.

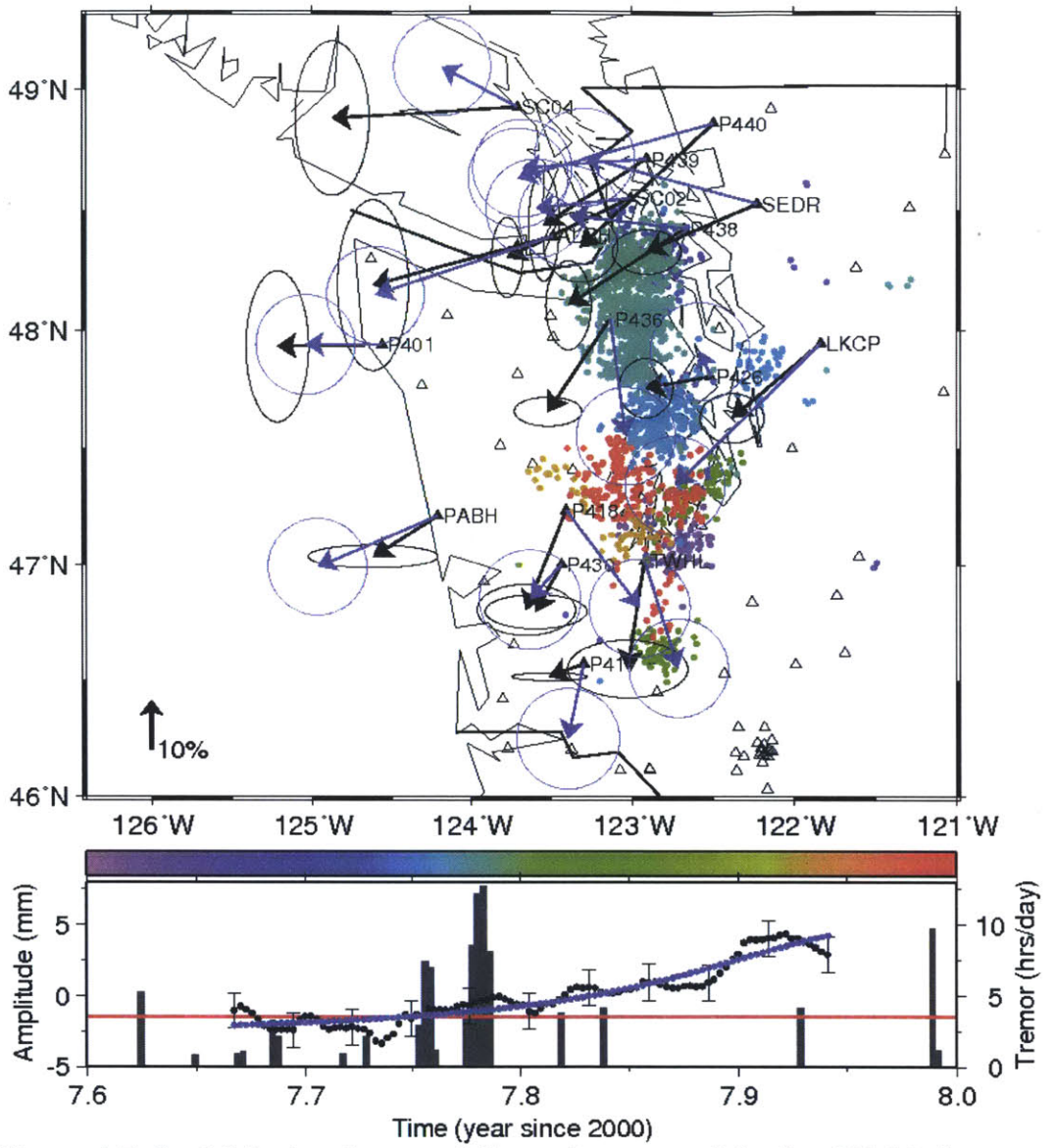


Figure 4-5. Spatial (top) and temporal (bottom) patterns of the first PC (black vectors and dots). The black vectors are significant at 5% significance level and show a westward motion with maximum displacement of ~2 mm. Also shown is an arctangent fit to the PC temporal pattern (blue dots), which is used to fit raw data (see Figure 4-6) and shows similar spatial pattern in northern Washington (blue vectors). Stations in southern Washington are less obvious with respect to their uncertainties. Tremors are anomalous between 2007.75 and 2007.80 (histogram); two peaks of tremors lasted 4 and 5 days with a 4-day interval and mostly located in the Puget Sound area (colored dots with horizontal color scale). The red horizontal line indicates the median value of tremors (~3.5 hrs/day) between 1 January 2006 and 22 May 2011. Error ellipses and bars indicate 1-sigma uncertainty. The black vertical arrow indicates 10% of the displacement obtained from the temporal pattern.

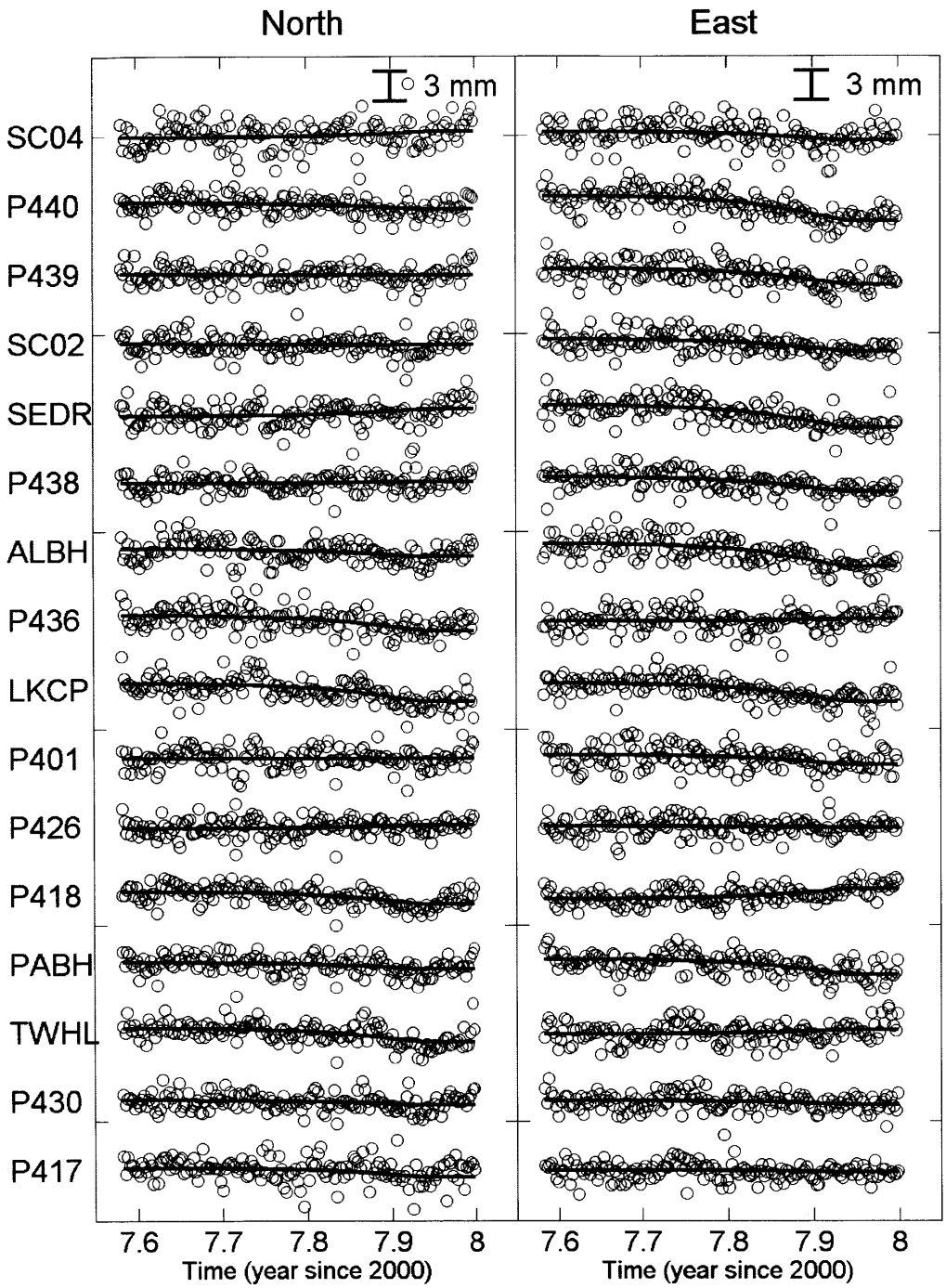


Figure 4-6. GPS time series (circles) and the arctangent fits (curves) obtained from the PC temporal pattern as shown in Figure 4-5. Linear rates are simultaneously estimated and removed. The station IDs are in descending order of latitude from top to bottom. In spite of small amplitudes, westward motion can be seen particularly at northern sites. The displacements inferred from the model fits are shown by blue vectors in Figure 4-5.

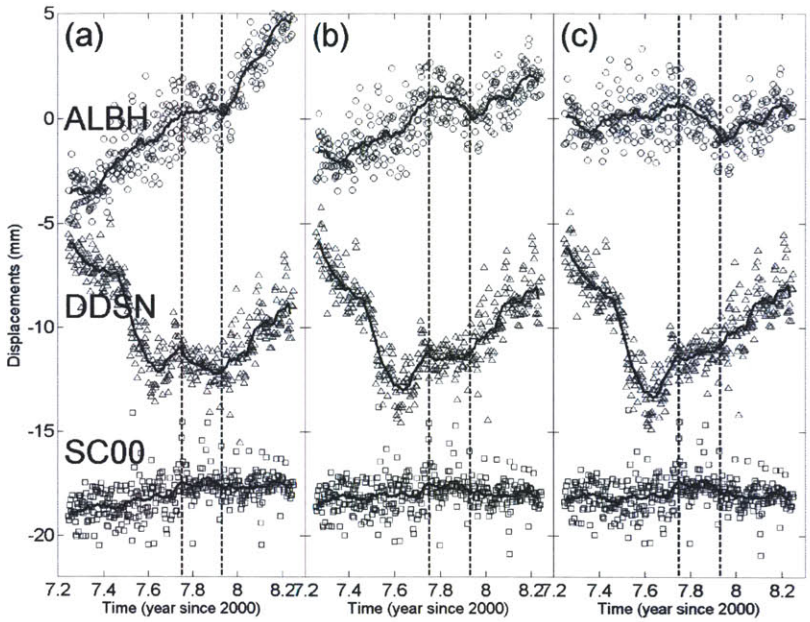


Figure 4-7. Time series of the east component of the three stations: ALBH (circles), DDSN (triangles) and SC00 (squares). DDSN is located about 600 km south of ALBH and SC00 is located about 250 km southeast of ALBH. The vertical lines delineate the time interval of the identified signal. The solid lines are values of moving averages with 30-day window. (a) Raw data time series are shown. A rough linear pattern is shown in ALBH and SC00. (b) Time series are detrended by removing a rate that is computed from the whole time interval including major SSEs. (c) Time series are detrended by removing a rate that is computed from the whole time interval but with major SSEs removed and estimated by offsets as shown in Figure 4-3b. Among the three solutions, the case (c) best show the small transient signal in ALBH. However, the far-field stations do not show the signal, which means that the signal is not spatially-correlated and network-wide error.

In the third case, an anomalous signal occurred in northern Washington for about two months between February and March 2009 (Figure 4-8). As in the second case, the lag-differences showed marginal significance of the signal due to the small displacement; the maximum displacement is ~4 mm at station P435, but the other stations moved less than 2 mm. There are two peaks of tremors for this signal. When compared to the May 2009 event, the smaller tremors occurred in smaller area and in shorter interval (see Figure 4-8). From the PC temporal pattern, however, it is not clear if the identified signal consists of two smaller signals corresponding to the two peaks of tremors.

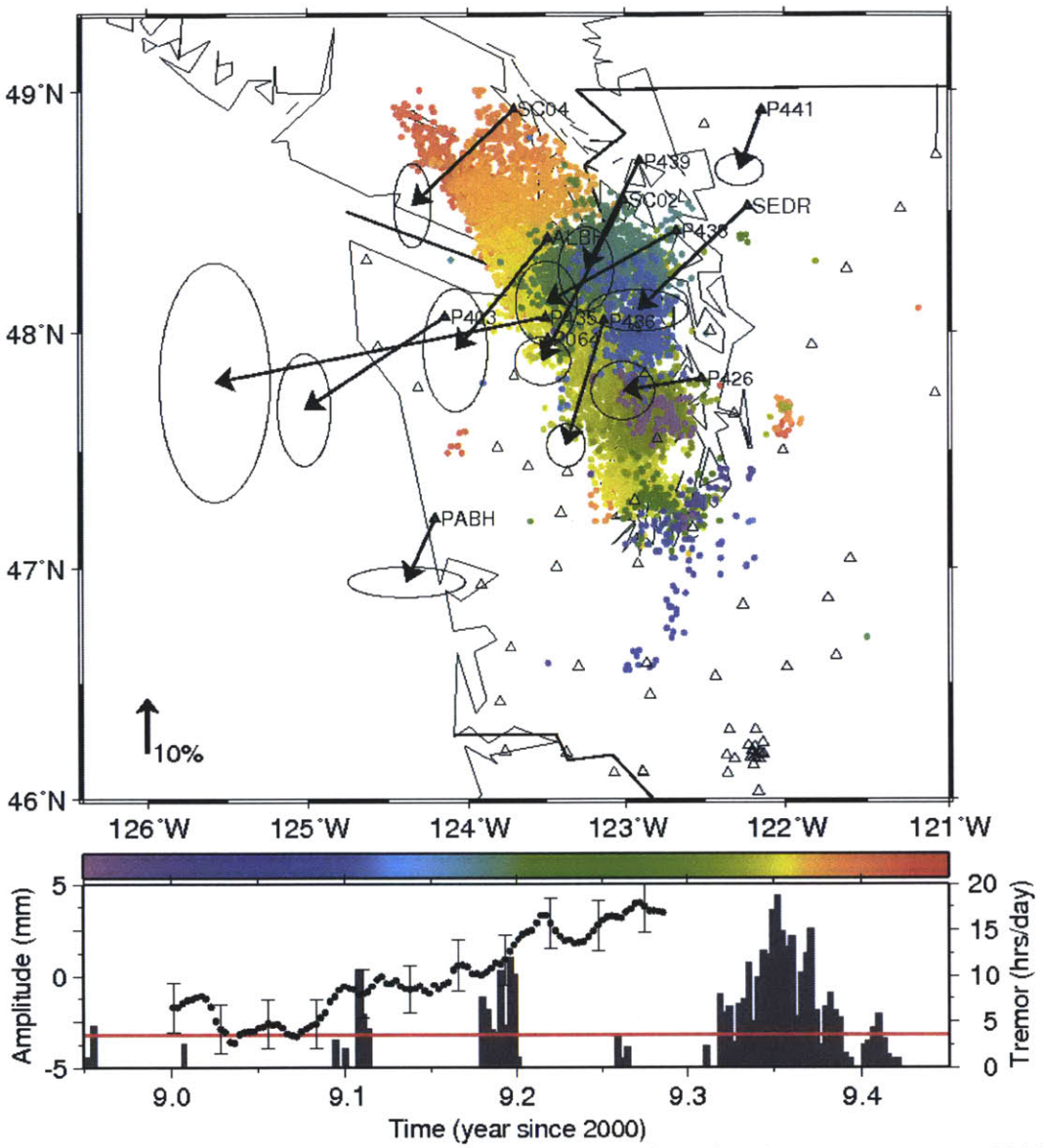


Figure 4-8. Similar to Figure 4-5 but with PCA results from data between 1 January 2009 and 15 April 2009. The signal time interval is roughly from 2009.075 to 2009.225 with maximum displacement of ~4 mm at station P435. The displacements of the other stations are less than 2 mm. The spatial direction is slightly further to south when compared to that shown in Figure 4-5, but the uncertainties of a few sites reflect that north-south direction is noisier. Two peaks of tremors lasted 3 and 9 days with about 3-week interval with maximum rate of ~12 hours/day. The tremors occurred in a smaller area and for a short time when compared to the tremors associated with the May 2009 event (tremors later than 2009.3; see Figure 4-2 for surface displacement pattern).

4.4 Discussion

The capability of detecting small SSEs is sensitive to how steady motions (e.g., linear rates and seasonal signals) are estimated and removed. For example, the raw data of the station ALBH show a temporary halt rather than a reversal in east-west direction (Figure 4-7a). When the time series is detrended by removing a rate computed from the whole data including the major SSEs, a slight westward motion can be seen (Figure 4-7b) but this motion was not able to be detected by PCA due to still low amplitude. When the time series is detrended by removing a rate computed from the whole data but with major SSEs removed and estimated by offsets as shown in Figure 4-3b, the small transient signal can be seen more clearly (Figure 4-7c) and it was detected by PCA (Figure 4-5).

According to the linear relationship between seismic moment and slip duration [*Ide et al.*, 2007; *Aguiar et al.*, 2009], small ETS events should have shorter duration than the major ETS events. The tremors reflect this relationship (see Figure 4-8), while the PCA-inferred small SSEs appears to have longer duration than the tremors (Figures 4-5 and 4-8). Even in the case of major SSEs, tremor data only cover a part of the PC temporal pattern (Figure 4-4). This may imply that the PC temporal pattern is not appropriate to precisely determine the duration of the SSEs, particularly when the magnitude of the SSEs is low. Nevertheless, spatial and temporal correlation between the small SSEs and the tremors suggests that they are related phenomena.

We have detected a few more signals that resemble SSEs in space (e.g., those around 2004.98, 2006.84 and 2008.57 in northern Washington and around 2006.66 in southern Washington). They were determined significant by PCA statistics, but we could not relate them with tremors (available since 2006 in this study). The case occurred in

2006.66 matched with tremors for three days with a rate of only ~6 hours/day slightly larger than the normal tremor activity (~3.5 hours/day).

In some cases, we have seen that spatial and temporal patterns are similar to SSEs but they fail the significance test. For example, Figure 4-9 shows westward motions similar to SSEs in this region but the pattern is not significant at the 5% level. Furthermore, the displacements are mostly less than 1 mm, which is at the typical noise level of horizontal components of PBO sites. This shows that the PCA uncertainties are reasonably estimated. Although anomalous tremors existed for the time interval, the relation with the signal is unclear because of the small amplitude.

ETS events have been also observed in southwest Japan, lasting a few days or a week and recurring with intervals of 2 ~ 6 months [*Obara et al.*, 2004; *Hirose and Obara*, 2005; *Obara and Hirose*, 2006]. Since the SSEs are too small to be detected by GPS (long-term SSEs of durations from months to years with centimeters of surface displacements have been detected by GPS [*Hirose et al.*, 1999; *Hirose and Obara*, 2005]), the SSEs have been recorded by high-sensitivity tiltmeters. The surface displacements inferred from a dislocation model based on tilt changes are less than 2 mm [*Obara et al.*, 2004], which is thought to be limitation of GPS resolution [*Wech et al.*, 2010]. In this study, we used GPS measurements at multiple stations and have identified small SSEs with average surface displacements of less than 2 mm (see Table 4-1).

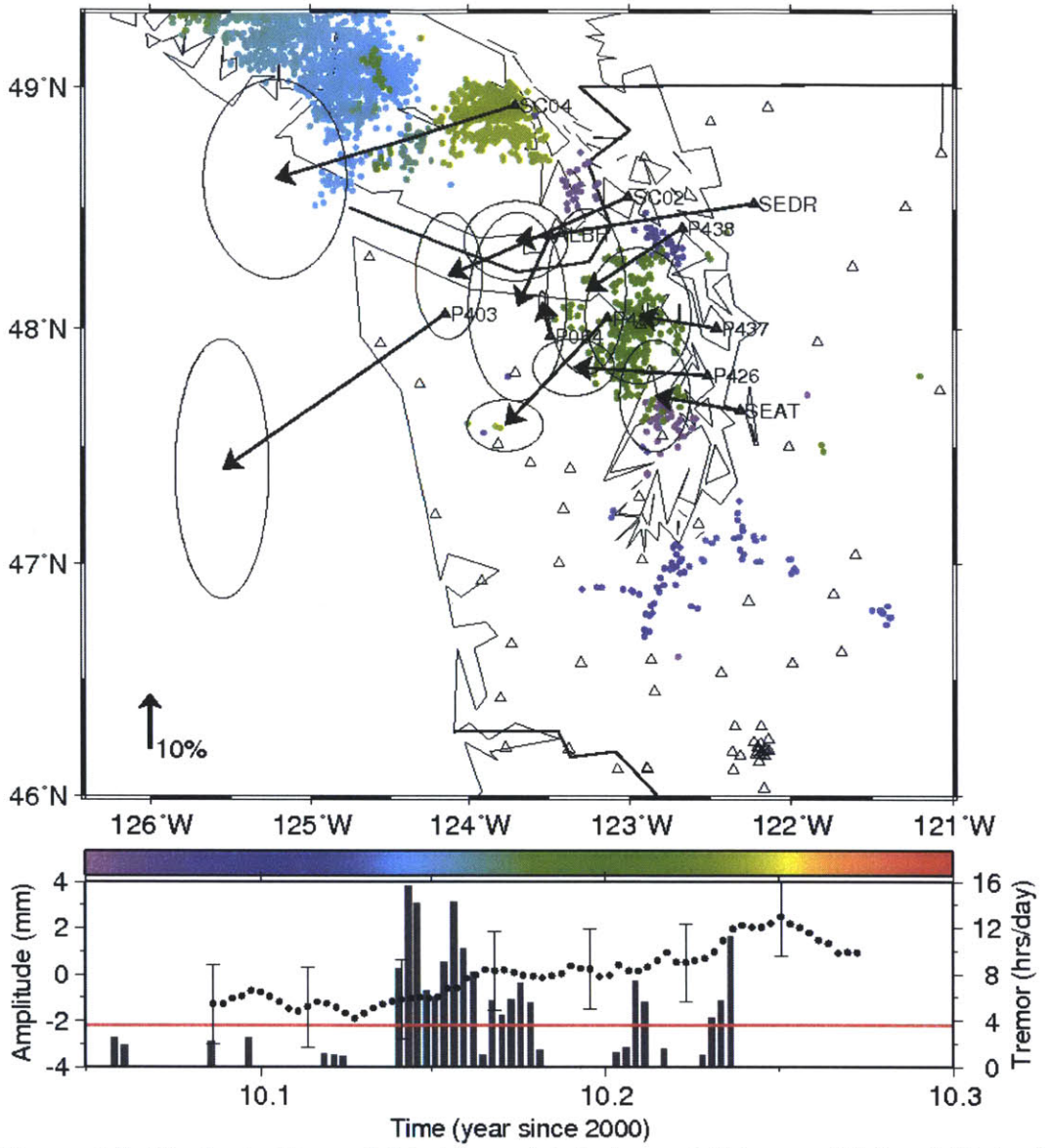


Figure 4-9. Similar to Figure 4-8 but from data between 1 February 2010 and 10 April 2010. The signal time interval is roughly from 2010.125 to 2010.25 with displacements of ~1 mm at stations P403 and SC04 and less than 1 mm at the other stations. The PC temporal pattern is not significant relative to the uncertainties ($\chi^2_{dof} = 0.5$). The spatial pattern looks similar to other events, but the pattern is not significant at 5% significance level. The error ellipses indicate 1-sigma uncertainty. The tremors around 2010.15 are not likely to be related with the signal because they occurred in Vancouver Island where the PBO network does not cover. The two peaks of tremors after 2010.2 occurred near the station SC04 and in the Puget Sound area, but the contribution of tremors is not clear.

Major ETS events are observed to release strain energy accumulating due to the plate coupling in Cascadia subduction zone [*Aguiar et al.*, 2009; *Chapman and Melbourne*, 2009]. At about 14-month period, major ETS events in northern Cascadia have 2–3 cm of fault slip (Mw = 6.3 ~ 6.8) along the plate interface at depth of 25-40 km [*Szeliga et al.*, 2008], only accounting for 45–65% of the plate convergence rate of 4 cm/year. Smaller ETS events may contribute to moment release as much as major ETS events, and thus they are important to characterize the process of strain accumulation and release and assess seismic hazards of future megathrust earthquakes.

4.5 Conclusions

We applied our detection method to the PBO GPS data collected in Washington State section of the Cascadia subduction zone and have detected two previously unrecognized small SSEs independently of tremor information. The spatial patterns of the small SSEs are similar to well-known major SSEs but in a smaller area where tremors are also observed. The average displacements are less than 2 mm, which was recognized as below GPS resolution. The durations of the SSEs are about 2 months longer than the duration of tremors, but the slip duration is not precisely determined from PCA-inferred temporal pattern due to low signal-to-noise ratio. We have also detected some more signals that resemble SSEs but PCA statistics indicated that they are not significant. The detection of small SSEs supports the claim that the small tremors also accompany SSEs.

Acknowledgements

This study used tremor data provided by Aaron Wech.

Chapter 5
Conclusions and future work

5.1 Conclusions

We use state estimation and principal component analysis (PCA) for detecting anomalous transient signals from Global Positioning System (GPS) position time series. The method focuses on improvement of the signal-to-noise ratio (SNR) of GPS time series in both space and time. We have demonstrated with applications to synthetic and real data sets that the method is capable of detecting relatively weak transient signals that might not be apparent in the noisy raw data.

For state estimation in this study, a smoother based on a forward-backward Kalman filters is used. As a stochastic state, a first-order Gauss-Markov (FOGM) process is used to account for temporally correlated noise and any transient signals in GPS position time series. To implement the FOGM process in a Kalman filter, the correlation time and the variance of the process can be estimated by minimizing innovation-based cost functions. Although depending on sample size, parameter correlation, relative

magnitude of colored and white noise, and the SNR, state estimation can improve the SNR by passing low-frequency power but attenuating high-frequency power.

PCA decomposes the estimates of FOGM state into uncorrelated and variance-maximizing patterns (i.e., principal components (PCs) for temporal variation and sample eigenvectors for spatial distribution). Temporal correlation in the FOGM state estimates can produce large-amplitude low-frequency oscillations in high-ranked PCs. The effect of temporal correlation can be reduced by removing sinusoidal oscillations from PCs. For statistical inference on PCA results, we approximately obtain PC uncertainties by propagating errors and sample eigenvector uncertainties by using asymptotic distributions with effective sample size to account for temporal correlation in PCs. These uncertainties allow us to identify transient signals in space from sample eigenvector distribution and in time from lag differences of PCs or residual PCs with sinusoidal oscillations removed.

We tested the detection capability of the method with synthetic data used in the transient detection exercise supported by the Southern California Earthquake Center. A measure of detection threshold cannot be generally obtained. Detection capability depends on many factors such as the SNR, temporal correlation, spatial extent and sampling of signals, and the signal complexity. Although random-walk and flicker noise were used for temporally correlated noise in the synthetic data, the FOGM state estimates can represent the temporally correlated noise and also transient signals. As more number of stations record a transient signal that is coherent between neighboring sites, PCA tends to reveal the signal in first a few PCs. Large-scale signals in space, however, cannot be easily detected because they are difficult to separate from network-wide common mode errors. Multiple and/or propagating signals can be detected if the SNR is high, but they

appear at multiple PCs, making interpretation complicated. Missing data interpolation by state estimation allows PCA to retain more data in space and time and thus increases the possibility of signal detection.

We also demonstrate the detection capability with two known signals in the Los Angeles basin, California: the coseismic deformation associated with the 1999 M7.1 Hector Mine earthquake and the hydrologic expansion of the San Gabriel Valley aquifer system due to heavy rainfall in winter 2004–2005. Although the two signals were previously recognized with GPS, the strength of the PCA-based signal detection can be readily shown because the coseismic offsets were detected with only far-field sites having maximum surface displacement of ~ 6 mm and the hydrologic signal occurred at only 11 sites but was detected from 71 sites available for the time interval of the signal.

We applied our detection method to the daily GPS position time series from the Plate Boundary Observatory (PBO) network in Alaska. We have detected a transient signal at Akutan volcano during the first half of 2008. A simple Mogi source at a shallow depth (~3.9 km below sea level) can explain most of the displacement field. The event detected could in part of an inflationary sequence that will eventually lead to an eruption. The aseismic and non-eruptive inflation signal is difficult to detect in raw time series because of low SNR, the large size of the network and correlated noise in space and time. State estimation and PCA improved the SNR and successfully revealed the signal.

We also applied the detection method to the PBO GPS data collected in Washington State section of the Cascadia subduction zone and have detected two previously unrecognized small slow slip events (SSEs) independently of tremor information. The spatial patterns of the small SSEs are similar to well-known major SSEs

but in a smaller area where tremors are also observed. The average displacements are less than 2 mm, which was thought below GPS resolution. The detection of small SSEs can improve our understanding of the process of strain accumulation and release in subduction zones and help evaluate seismic hazards of future megathrust earthquakes.

5.2 Future work

The detection algorithm can be further developed for a continuously running automated near-real-time detection system. For this, the algorithm must be run efficiently with little user-intervention and with rigorous detection criteria. Currently, state estimation is automated because FOGM process parameters are estimated from measurements. However, some steps of PCA have been done manually. For example, whether a signal is systematic or coherent between nearby sites (e.g., strike-slip or dip-slip faulting, expansion or contraction, uplift or subsidence, etc) has been visually determined from sample eigenvector distributions. A measure of signal coherence is difficult to obtain in general. If a transient deformation pattern is available or approximately expected, one can use the pattern to evaluate if an observed transient signal is similar to the pattern. Another example is that local outliers (i.e., sites with large amplitude but not coherent with nearby sites) have been manually removed from the data set. By defining a local outlier whose eigenvector element must be larger than a given value and whose correlation with nearby sites must be less than a given number, local outliers can be removed in an automatic fashion, but it was not always successful.

For near real-time processing algorithm, the detection method can be simply applied to a moving window with a given duration (e.g., half or a year) and at a given

interval (e.g., weekly) as new data become available. Using parallel computing many different events can be monitored simultaneously. The time series noise models (i.e., FOGM process in this study), which can take considerable time to estimate, would not be necessarily updated every time the algorithm runs unless something changes dramatically at a site. The IGS-final-orbit products have 2-3 week latency and the IGS-rapid-orbit products have 24-hour latency in general.

In this study, the detection algorithm uses daily GPS position time series. Since the algorithm can be readily used if input data are time series at multiple locations (i.e., multivariate time series), it can be adapted to use or incorporate other types of geodetic data such as interferometric synthetic aperture radar (InSAR) data that provide dense spatial coverage and sensitivity to vertical deformation and strainmeter data that provide a higher sampling rate and sensitivity to small signals than GPS. However, noise characteristics of these data are less well understood than those of GPS data. The FOGM processes may be applicable because they can mimic the spectrums of many other noise models (but not all) when multiple FOGM processes are added.

We have applied the detection algorithm to the entire PBO network from Alaska to southern California and this study includes detection results in three regions: Alaska (Chapter 3), northern Cascadia (Chapter 4), and Los Angeles basin (Appendix D). We have also detected some more transient signals, for example, recent deceleration of uplift from 2006 to June 2010 at Yellowstone [*Chang et al.*, 2010], the effect of snow coverage on top of antenna of stations on Mount St. Helens, inflation during 2007 and deflation for 2009-2010 at Long Valley Caldera with a few millimeters of surface displacement, which is small compared to previous episodes (e.g., maximum uplift during the 2002–2003

episode is ~35 ± 8 mm, about 1/3 the magnitude of the 1997–1998 episode [*Feng and Newman*, 2009]). Since the PBO network covers the western margin of US from Alaska to southern California, there would be many places where transient deformation has occurred and may go unnoticed. Furthermore, there are more GPS networks (e.g., GPS Earth Observatory Network (GEONET) in Japan) to which our detection algorithm can also be applied for anomalous transient signal detection.

Finally, GPS modernization (e.g., new civilian signals L2C and L5 at 1176.45 MHZ and planned launches of Block IIF and Block IIIA satellites) will further improve the precision and accuracy of 3-dimensional positioning. Furthermore, other global navigation satellite systems (GNSS; e.g., European Galileo, Russian GLONASS, Chinese Compass (being developed)) will also improve the positioning. Incorporating such improvements in the detection method leads to more precise and more accurate detection of transient signals.

Appendix A

Improving signal-to-noise ratio of GPS position time series by state estimation

Abstract

The signal-to-noise ratio (SNR) of GPS position time series can be improved by a smoother based on the Kalman filter formulations. However, if the noise characteristics are not well determined the state estimation will be inaccurate and result in large estimation errors. Here we develop an algorithm to determine the noise process. We use a first-order Gauss-Markov (FOGM) process, characterized by a correlation time and a variance, to account for temporal correlation in GPS time series including unmodeled transient deformation signals and colored noise. The FOGM correlation time and the variance are estimated by minimizing two cost functions obtained from the innovations sequence (i.e., differences between measurements and predictions). The measurement white noise variance can then be estimated by scaling its assumed value (formal errors are often used) by the ratio between the FOGM variance estimates from the two cost functions. The method is not sensitive to correlation between the noise parameters because they are independently estimated.

A.1 Introduction

Global Positioning System (GPS) stations in a continuously operating network (e.g., Plate Boundary Observatory in the western United States and GPS Earth Observation Network in Japan) measure surface motion and produce a time series of position changes in each coordinate direction (usually displacements in the north, east and up direction) relative to a reference frame. With repeatabilities at the level of 1 mm in horizontal and 3 mm in vertical, daily GPS position time series have revealed transient motions due to seismic, volcanic or hydrologic activities. However, weak transient signals are often difficult to detect because of temporally correlated noise in GPS position time series. A desirable property of a signal detector is one which increases signal-to-noise ratio (SNR). Among various techniques for this purpose, we improve the SNR of daily GPS position time series by state estimation using a smoother based on a forward-backward running Kalman filter approach [e.g., *Gelb*, 1974; *Anderson and Moore*, 1979; *Brown and Hwang*, 1997]. The state estimation is briefly summarized in Appendix B.

The state vector typically, but not necessarily, includes a secular rate, annual and semi-annual sinusoids and offsets due to receiver/antenna changes. Temporally correlated noise can be characterized by stochastic models such as random-walk, flicker, power-law noise [e.g., *Zhang et al.*, 1997; *Mao et al.*, 1999; *Williams et al.*, 2004; *Beavan*, 2005; *Langbein*, 2008]. We use a first-order Gauss-Markov (FOGM) process to describe temporally correlated noise and transient signals if any. This does not necessarily mean that this process is a physical process for temporally correlated noise and transient signals, but we use it here as a convenient and flexible representation of temporal correlation in GPS position time series. Transient signals can be described by deterministic states if

proper models are available (e.g., step functions for coseismic offsets and exponential or logarithmic parameterization for postseismic deformation). Even in this case, however, a stochastic state is required to account for temporally correlated noise, model misfits and other transient signals that might exist.

The main objective of this study is to estimate the FOGM parameters (i.e., correlation time and variance) and measurement white noise variance. Poor specification of these parameters may result in large estimation errors. The parameters of noise processes in GPS time series have been estimated by maximum likelihood estimation (MLE) and power spectral density (PSD) methods in the geodetic literature [e.g., *Mao et al.*, 1999; *Williams et al.*, 2004; *Langbein*, 2008]. The MLE method has a few advantages over the PSD method (e.g., it can handle uneven samples and missing data and it can estimate both tectonic and noise parameters), but it is computationally expensive.

An alternative approach to the estimation of the noise parameters is to modify or adapt the Kalman filter to the actual measurements, which is known as adaptive Kalman filtering. An adaptive filter readjusts noise statistics based on the measurements [*Gelb*, 1974]. Most techniques use the innovations sequence defined as a sequence of differences between measurements and predictions [e.g., *Jazwinski*, 1969; *Mehra*, 1970; *Myers and Tapley*, 1976; *Mohamed and Schwarz*, 1999; *Yang and Gao*, 2006]. The innovations sequence conveys new information, inferred from the process noise models, brought by the current measurement given all the past measurements. The new information is a basis for estimating the noise statistics and thus for improving Kalman filter performance.

Our method also uses the innovations sequence because it can be efficiently done by Kalman filtering. However, our approach has some advantages over conventional adaptive approaches. Firstly, most adaptive schemes estimate the noise statistics on-the-fly as a new measurement is available that could cause instability of the filter [*Myers and Tapley*, 1976; *Blanchet et al.*, 1997], while our approach is stable because it uses all the innovations available (i.e., post-processing). Secondly, most adaptive schemes estimate stochastic noise statistics (Q_k and R_k), while we also estimate the stochastic state transition (Φ_k) because it is a function of the FOGM correlation time. Thirdly, most adaptive schemes simultaneously estimate the noise statistics, while we estimate each of the FOGM parameters independently, mainly due to parameter correlation, by minimizing two different cost functions constructed by the innovations sequence. Our method provides robust estimates of the noise parameters and thus improves the SNR.

A.2 First-order Gauss-Markov (FOGM) process

A FOGM process is selected as a stochastic state because it can be easily implemented in Kalman filters and because its spectrum, when multiple FOGM processes are added, can mimic the spectrums of many other noise models (but not all) [*Langbein*, 2008]. A FOGM process can be expressed by a simple differential equation, while flicker and power-law noise cannot. The FOGM spectrum is, in a log scale, flat at low frequencies and decreases with slope of -2 at high frequencies as do random-walk processes. The spectrum is controlled by a variance (σ^2_{fogm}) for the overall power and by a correlation time (τ) for the transition from flat to the decreasing power. At a corner

frequency, $f_c = 1/\pi\tau$, extrapolation of the flat spectrum intersects the slope for high frequencies. As a correlation time is close to zero, a FOGM process approaches to a white process (i.e., a flat spectrum). When a correlation time goes to infinity, a FOGM process becomes a random-walk process.

The FOGM state transition (Φ_k) and the one-step-ahead process noise variance (Q_k) are related to the two FOGM parameters (τ and σ^2). A FOGM autocovariance function is given by $\rho(\Delta t) = \sigma^2 \exp\left(-|\Delta t|/\tau\right)$ where Δt is a time interval and where $|\bullet|$ indicates the absolute value. Let x_k and w_k be a FOGM state and process noise with variance Q_k at time t_k, respectively. After multiplying both sides of the state transition equation $x_{k+1} = \Phi_k x_k + w_k$ by x_k, taking the expectation, and using the fact that w_k is uncorrelated with x_k, we have the Yule-Walker equation $\rho(\Delta t_k) = \Phi_k \rho(0)$ where $\Delta t_k = t_{k+1} - t_k$ and thus $\Phi_k = \exp\left(-\Delta t_k/\tau\right)$. When the correlation time increases (i.e., closer to random-walk process), the state transition approaches to one. On the other hand, when the correlation time decreases (i.e., closer to white noise), the state transition approaches to zero. When multiplying both sides of $x_{k+1} = \Phi_k x_k + w_k$ by x_{k+1}, we obtain the process noise variance $Q_k = \sigma^2\left(1 - \exp\left(-2\Delta t_k/\tau\right)\right)$. Before running a smoother, the state transition and the process noise variance must be specified. Equivalently, the FOGM parameters must be known or estimated, which will be discussed in the next section.

A.3 Innovation-based FOGM parameter estimation

The FOGM parameters can be estimated by using the innovations sequence from the Kalman filter running forward in time. The innovations sequence is a sequence of differences between measurements and predictions from the Kalman filter. In this section, we examine the properties of the innovations sequence and estimate the FOGM parameters by minimizing cost functions constructed from the innovations sequence.

A.3.1 Innovations sequence

Given measurements y_k where $k = 1,2,\cdots,n$, an innovation sequence ν_k at time t_k is defined as $\nu_k = y_k - H_k \hat{x}_{k|k-1}$ where H_k is the design or measurement matrix and $\hat{x}_{k|k-1}$ denotes the predicted state vector at t_k based on all of the measurements up to t_{k-1}. For an optimal Kalman filter (i.e., the system and noise parameters are exactly known), the innovation variance is given by

$$C_k = H_k P_{k|k-1} H_k^T + R_k \tag{A-1}$$

where $P_{k|k-1}$ is an error variance of $\hat{x}_{k|k-1}$ and is obtained by $P_{k|k-1} = \Phi P_{k-1|k-1} \Phi^T + Q$ in the prediction step (see Equation B-4 in Appendix B). The innovations sequence is statistically equivalent to the measurements in a linear system and conveys new information brought by the current measurement given all the past measurements [*Kailath*, 1968]. The innovation variance sequence gives weights to the new information via the Kalman gain (see Equation B-5 in Appendix B). Note that the innovation variance is not a function of actual measurements. Given a state-space model, the innovation variance can be directly determined regardless of actual measurements.

For an optimal filter, the innovations sequence is a Gaussian white sequence with zero mean and variance given in Equation (A-1) [*Kailath*, 1968]. Since v_i and v_j for $i \neq j$ are independent, individual innovations bring new information from measurements. As a diagnosis for the optimality, the whiteness can be tested by checking if lagged autocovariances of innovations are less than a given threshold [*Mehra*, 1970]. In theory, the lagged autocovariance sequence must be zero for an optimal filter, provided that the model is linear and the process and measurement errors are independent [*Daley*, 1992]. Large lagged autocovariances may indicate poorly constructed state-space model.

In suboptimal cases, the innovations sequence is not white and its variance is different from the innovation variance given in Equation (A-1) [*Anderson and Moore*, 1979]. Instead sample innovation variance can be considered:

$$\hat{C} = \frac{1}{n}\sum_{k=1}^{n} v_k^2 \tag{A-2}$$

where all the available innovations are used. Most adaptive schemes estimate the innovation variance (say $\hat{C}_k$) using the most recent innovations within a moving fixed-interval time window. The values of Q_k and R_k are adapted based on the values of $\hat{C}_k$. For example, $\hat{R}_k = \hat{C}_k - H_k P_{k|k-1} H_k^T$ from Equation (A-1) [e.g., *Mehra*, 1970; *Myers and Tapley*, 1976; *Mohamed and Schwarz*, 1999]. Small size of the moving window, however, possibly causes the adaptive filters unstable and $\hat{R}_k$ negative. Equation (A-2) assumes that an innovations sequence is stationary and thus our estimation technique is stable.

We investigate the effect of suboptimal filter on the innovations sequence by comparing the (time-averaged) innovation variance in Equation (A-1) with the sample innovation variance in Equation (A-2). FOGM plus white samples were generated with

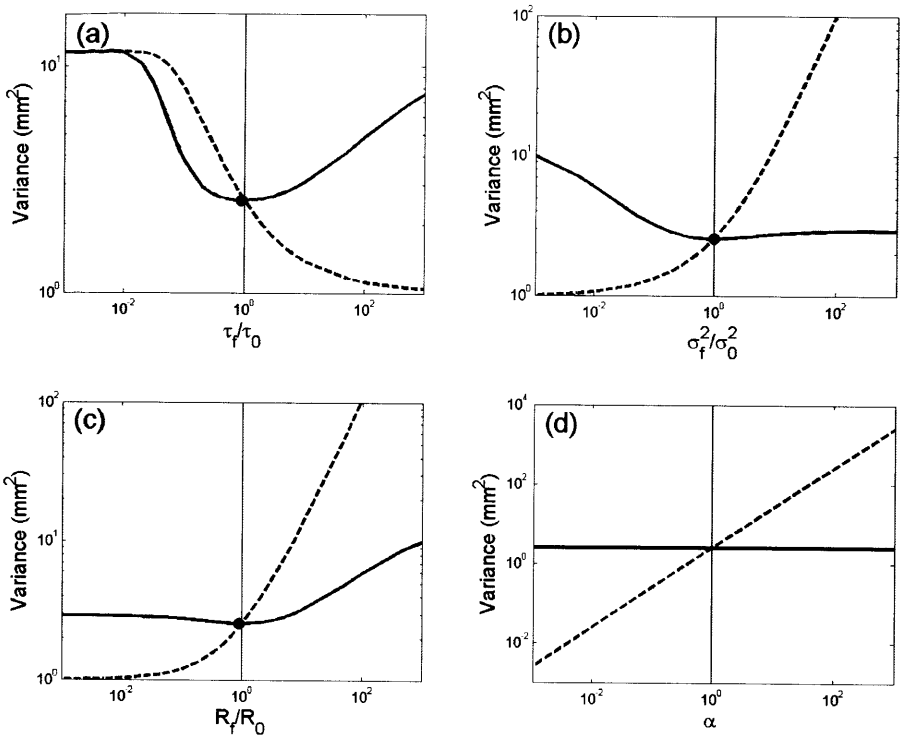


Figure A-1. (dashed) Innovation variance and (solid) sample innovation variance from FOGM samples with $\tau_0 = 20$ days, $\sigma_0^2 = 10.5$ mm^2 (equivalently, $Q_0 = 1$ mm^2) plus white samples with $R_0 = 1$ mm^2. The parameters τ, σ^2 and R are, respectively, the FOGM correlation time, FOGM variance and white noise variance with subscript f for the filter and subscript 0 for the samples. The sample size (T) is 1000 (i.e., $T/\tau_0 = 50$). The sample innovation variances were averaged over 1000 realizations. All axes are in a log scale. (a) ~ (c) Only one parameter is varying as shown in horizontal axes, while the other two parameters are fixed to the true values. (d) The parameter α is a scale factor such that $\sigma_f^2 = \alpha\sigma_0^2$ and $R_f = \alpha R_0$. The sample innovation variance is equal to the innovation variance when the filter parameters are equal to the true parameters (vertical lines), that is, the filter is optimal. The cases (b) ~ (d) indicate that if $\tau_f = \tau_0$ the sample innovation variances depend on the ratio σ_f^2/R_f, not on the individual values. In particular, the case (d) shows that the sample variances are those from the optimal filter whenever $\sigma_f^2/R_f = \sigma_0^2/R_0$ if $\tau_f = \tau_0$. The minimum values of the sample innovation variances (sold circles) in (a) ~ (c) are associated with the true parameter, but it is not true that the minimum occurs only if the filter is optimal because of the dependence on σ_f^2/R_f.

given parameters (τ_0, σ_0^2 and R_0), while Kalman filters were constructed with a range of filter parameters (τ_f, σ_f^2 and R_f). We assume here that the white noise variance is time invariant (i.e., $R_k = R$ for all k). The sample innovation variance is equal to the innovation variance and becomes the minimum if filter parameters are equal to the true parameters (i.e., optimal filter) (Figure A-1). The innovation variance, on the other hand, is a monotonic decreasing function of τ_f (Figure A-1a) and a monotonic increasing function of σ_f^2 and R_f (Figure A-1b and c), irrespective of the optimal condition because it does not depend on measurements.

The sample innovation variance depends on the ratio between the FOGM variance and the white noise variance (σ_f^2/R_f) when τ_f is fixed. The sample innovation variance becomes larger when $\sigma_f^2 < \sigma_0^2$ (Figure A-1b) or when $R_f > R_0$ (Figure A-1c), indicating dependence on the ratio not the individual values of σ_f^2 and R_f. If $\tau_f = \tau_0$, the sample innovation variance becomes the minimum when $\sigma_f^2/R_f = \sigma_0^2/R_0$ (Figure A-1d). It is difficult to distinguish between too large a value of σ_f^2 and too small a value of R_f, and vice versa [*Daley*, 1992]. The algorithm we develop below solves this problem.

A.3.2 Minimization of cost functions

The FOGM parameters are estimated by minimizing cost functions based on innovations sequence. We test three types of cost functions. Firstly, we consider the sample innovation variance $\hat{C}$ given in Equation (A-2). If the filter is optimal, $\hat{C}$ approaches the minimum asymptotically [*Anderson and Moore*, 1979] (the reverse is not

true as shown in Figure A-1d). Secondly, assuming Gaussian statistics, the maximum likelihood condition is to minimize a cost function [e.g., *Blanchet et al.*, 1997; *Mohamed and Schwarz*, 1999]

$$L = \frac{1}{n}\sum_{k=1}^{n}\log(\det(C_k)) + \frac{1}{n}\sum_{k=1}^{n}\frac{\nu_k^2}{C_k} \tag{A-3}$$

where $\det(C_k)$ indicates the determinant of C_k. The first term in the right hand side has its unit of log of variance, while the second term is dimensionless. Unit change does not affect the minimization because constants are added to or subtracted from L in a log space. Finally, since the innovations sequence from an optimal filter is white Gaussian with variance C_k, we minimize a cost function

$$X = \left|\frac{1}{n}\sum_{k=1}^{n}\frac{\nu_k^2}{C_k} - 1\right|. \tag{A-4}$$

For an optimal filter, the first term in the right hand side becomes the chi-square per degrees of freedom whose expectation is one. Therefore, minimizing X is equivalent to finding a condition that makes the first term close to one. The cost function $\hat{C}$ does not depend on C_k, while the others do. Thus $\hat{C}$ is expected to have different properties.

Figure A-2 shows that the minimum values of the cost functions occur at or near the optimal condition when only one parameter is varying with the other parameters fixed to the true values. Note, from Figure A-2d, that the cost function $\hat{C}$ has no minimum because of its dependence on σ_f^2/R_f, while the cost functions L and X have the minimum at the optimal case. This suggests that L and X have dependence on σ_f^2 and R_f rather than on the ratio σ_f^2/R_f.

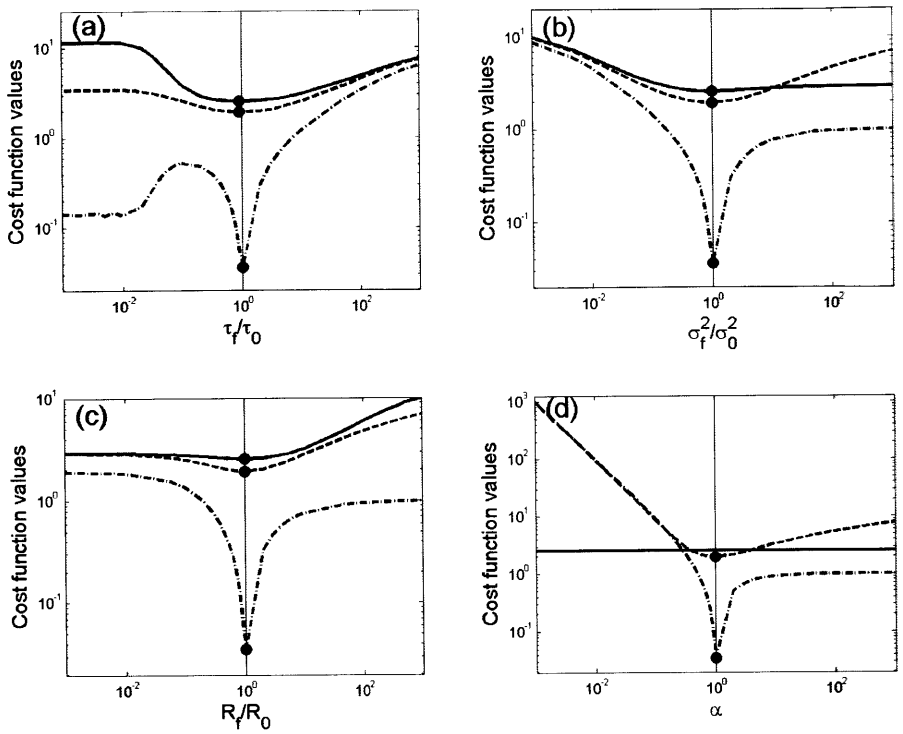


Figure A-2. Same as in Figure A-1 but vertical axis represents the values of cost functions $\hat{C}$ (solid curves; the same shown in Figure A-1), L (dashed curves) and X (dash-dotted curves). The cost functions were averaged over 1000 realizations. The average minimum values (solid circles) of the cost functions are at or near the optimal case (vertical lines). The cost function $\hat{C}$ in (d) has no minimum because of its dependence on σ_f^2/R_f, while L and X have the minimum at the optimal case, suggesting that they have dependence on the individual values of σ_f^2 and R_f rather than on the ratio σ_f^2/R_f.

When the FOGM parameters are varied, they are positively correlated in terms of the minimum cost. For the test cases assuming $R_f = R_0$, the correlation coefficients are 0.59 for $\hat{C}$, 0.88 for L, and 0.92 for X (Figure A-3). The positive correlation can be inferred from $Q_0 \approx 2\Delta t\sigma_f^2/\tau_f$ if $2\Delta t/\tau_0 << 1$ ($2\Delta t/\tau_0 = 0.1$ in the test cases). Since Q_0 and Δt are fixed and positive, τ_f and σ_f^2 are positively correlated. The slope of the distribution seems to be a function of $Q_0/2\Delta t$ (= 0.5 in the test cases).

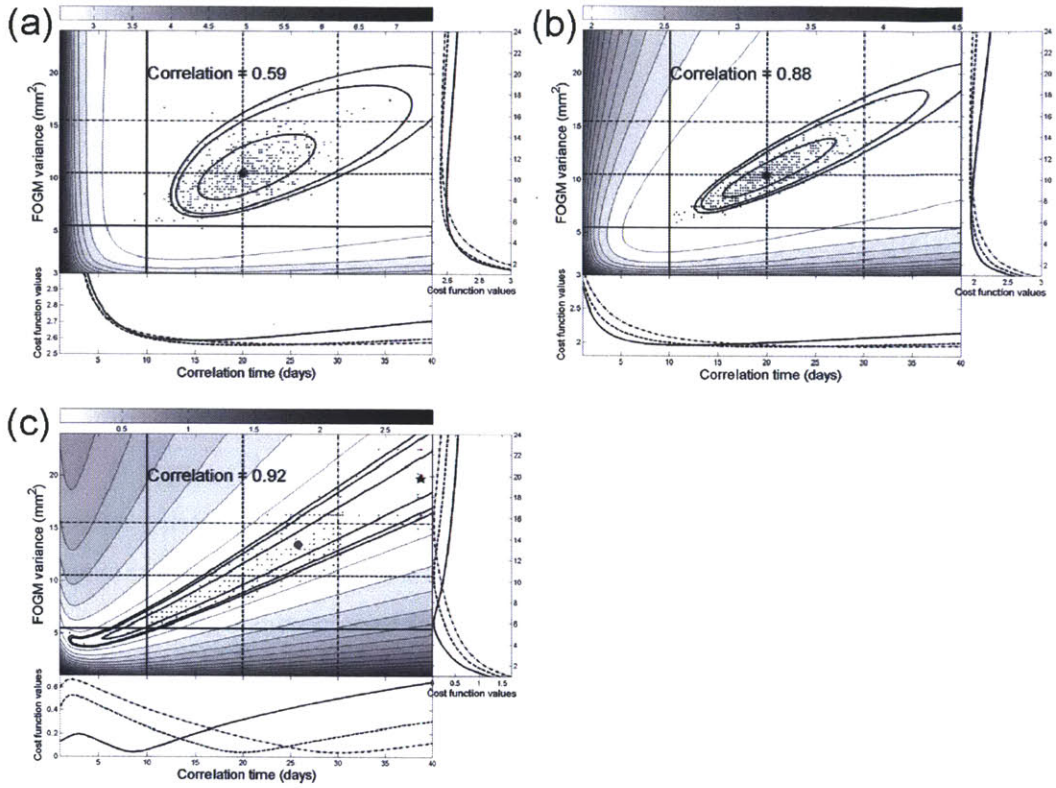


Figure A-3. Minimum values (dots) of the cost function (a) $\hat{C}$, (b) L, and (c) X on a grid of FOGM parameters from 500 realizations. True parameters are given in Figure A-1. The white noise variance used in the filter is fixed to the true value. The distribution (contours with gray scale at the top) was obtained by averaging the cost values from 500 realizations. Three thick contours enclose 50%, 90%, and 95% of the minimum values. Also shown are horizontal and vertical lines along which cross sections are drawn with the same type of line. The FOGM parameters are positively correlated in terms of the minimization of the cost functions. The average of the minimum values (gray circle) and the minimum of the distribution (star) are close to the true values except for the case (c), in which the minimum of the distribution (star) locates at the boundary of the parameter space. This indicates that the minimum may be outside of the given parameter space and far from the true values. As a result, the average of the minimum values (gray circle) is biased in this two-variable case. In the single-variable case, on the other hand, the minimum values can be well recovered at the true values (see Figure A-2).

The minimum values of the cost functions $\hat{C}$ and L, in the mean sense, coincide with the true values (see Figure A-3a and b), and thus the minimization of $\hat{C}$ or L provides a way to the FOGM parameter estimation in this two-variable case. However, the minimum of the cost function X locates at the boundary of the given parameter space (Figure A-3c), which indicates that the minimum may be outside of the space possibly due to high parameter correlation. For this reason, we will not consider the cost function X. When one parameter is fixed, X can be used (see Figure A-2).

The distribution of the minimum cost depends on the values of T/τ_0, Q_0/R_0 and R_f where T denotes the total data span. Let us assume $R_f = R_0$ here (Cases of $R_f \neq R_0$ will be discussed in the next section). As expected from large sample statistics, the minimum $\hat{C}$ and L are more compactly distributed for large T/τ_0 (Figure A-4), indicating consistency. When T/τ_0 increases from 3 to 50, the parameter correlation also increases from 0.22 to 0.59 for $\hat{C}$ and from 0.77 to 0.88 for L. The orientation of the distribution does not depend on T/τ_0.

The distribution of the minimum $\hat{C}$ rotates counterclockwise as Q_0/R_0 increases (Figure A-5). When $Q_0/R_0 = 0.01$, the FOGM variance can be estimated better than the correlation time. When $Q_0/R_0 = 10$, the correlation time can be better estimated. The slope of the distribution is a function of $Q_0/2\Delta t R_0$. For the minimum L, the slope of the distribution remains the same for $Q_0/R_0 > 1$ approximately (Figure A-5). The distribution becomes more compact along the axis of the FOGM variance as Q_0/R_0 further increases; the FOGM variance is better estimated than the correlation time for large Q_0/R_0.

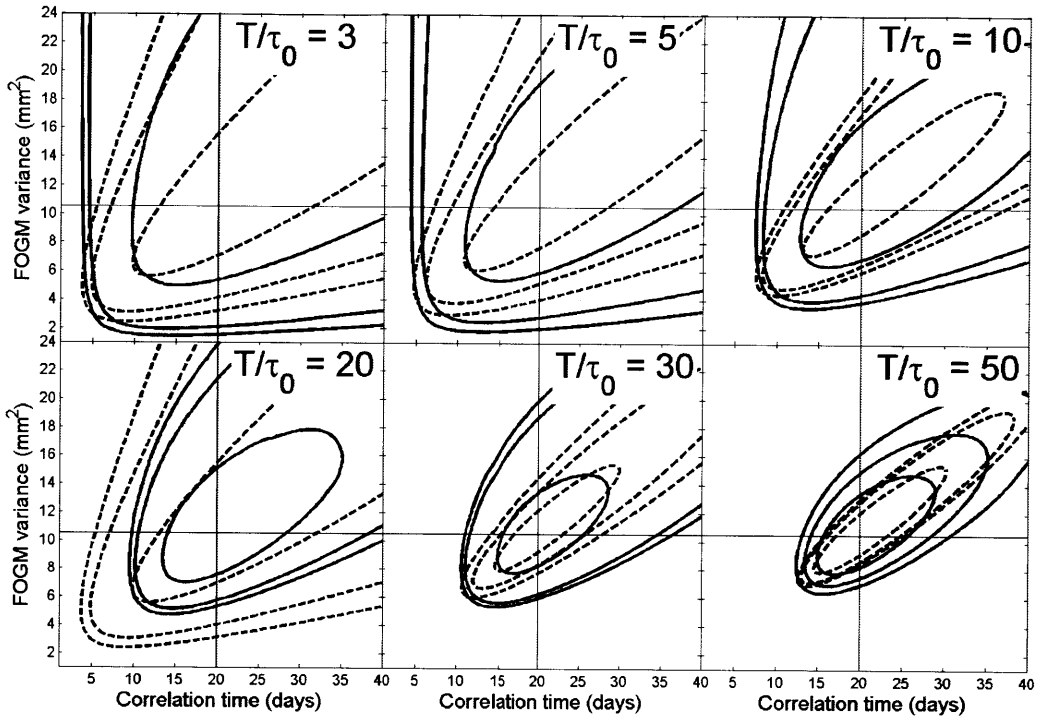


Figure A-4. Distributions of minimum values of the cost functions $\hat{C}$ (solid contours) and L (dashed contours) in the FOGM parameter space for different values of T/τ_0 where T is the total data span and τ_0 is the true FOGM correlation time (20 days indicated by vertical lines). The true FOGM variance is 10.5 mm^2 (horizontal lines) and white noise variance is 1 mm^2 (equivalently, $Q_0/R_0 = 1$). Three contours enclose 50%, 90%, and 95% of the minimum values. As T/τ_0 increases, the distribution becomes more compact and has higher correlation.

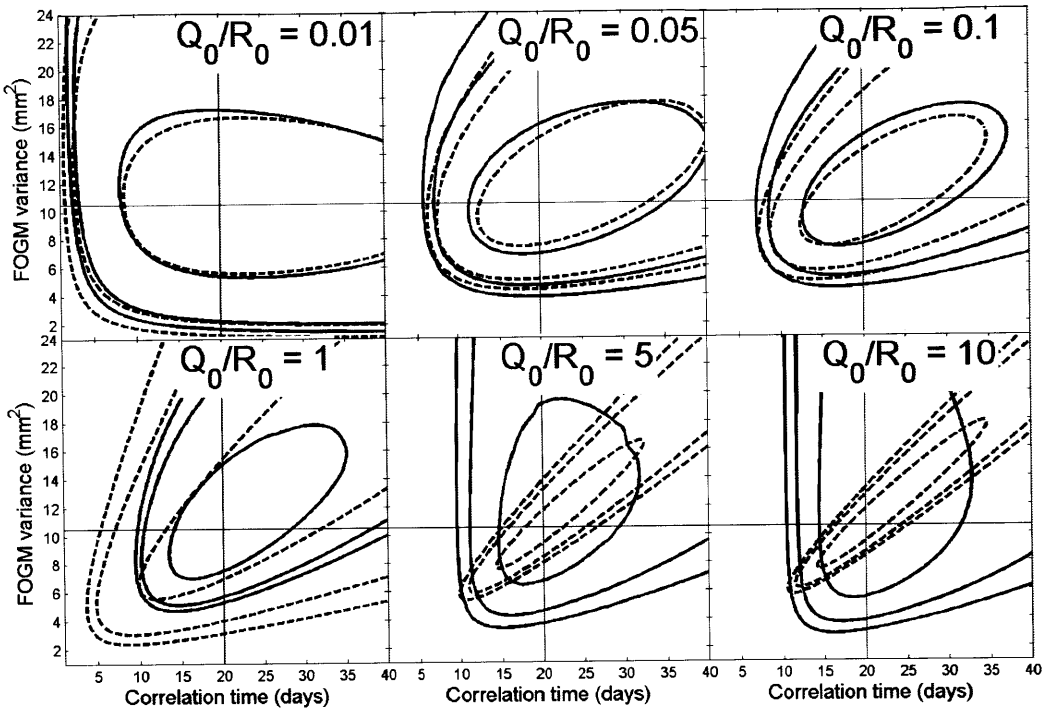


Figure A-5. Similar to Figure A-4 but the measurement noise variance R_0 changes according to the values of Q_0/R_0 given in each plot. True FOGM parameters are the same as those in Figure A-4. The data span T is 400 days so that $T/\tau_0 = 20$. As Q_0/R_0 increases, the distribution of $\hat{C}$ (solid contours) rotates counterclockwise, indicating changes in correlation between the parameters. The distribution of L (dashed contours) looks similar to that of $\hat{C}$ for $Q_0/R_0 < 1$, but the slope of the distribution remains the same for $Q_0/R_0 > 1$ approximately. Instead, the distribution becomes more compact particularly along the vertical direction, that is, the FOGM variance will be better estimated than the correlation time.

A.4 Estimation of measurement noise variance

In daily GPS position time series, measurement white noise is relatively better known (~1 mm in horizontal and ~3 mm in vertical) than the parameters of the colored noise. Formal errors are often used to approximate measurement noise variance. Formal errors are normally calculated from noise in the phase measurements at a given rate under the white Gaussian assumption. However, temporal correlation in the phase data raises the scale issue in formal errors, depending on statistics and sampling rate used in the GPS geodetic inversion. In this section, we examine how incorrectly specified measurement noise variance impacts on the FOGM parameter estimates and estimate the measurement noise variance based on the findings.

The FOGM parameter estimation is affected by incorrectly assumed measurement noise variance (i.e., $R_f \neq R_0$). Figure A-6 shows that the FOGM parameter estimates depend on the values of $R_f = \alpha R_0$ where $\alpha = 0.5$, 1 and 2. The FOGM variance from $\hat{C}$ is sensitive to variations in R_f (Figure A-6a), while the correlation time from L is sensitive (Figure A-6b). For $\hat{C}$, the state transition is relatively better determined because it depends only on the correlation time, but the one-step-ahead process noise variance cannot be independently separated from the measurement noise variance, which is consistent with the dependence of $\hat{C}$ on σ_f^2/R_f (see Section A.3.2). For L, on the other hand, both state transition and process noise variance cannot be independently separated from the measurement noise variance because they depend on the correlation time which is sensitive to the variation of R_f more than the FOGM variance is.

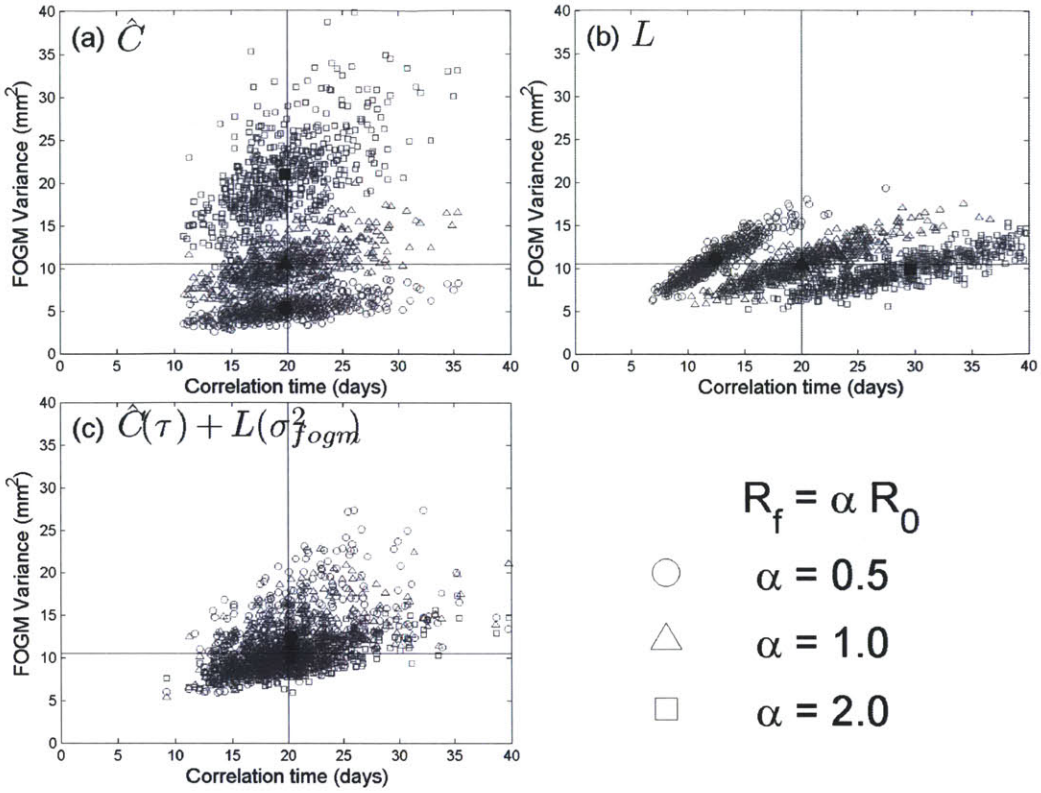


Figure A-6. Estimates of FOGM parameters from minimizing cost functions (a) $\hat{C}$, (b) L, and (c) the FOGM correlation time (τ) from $\hat{C}$ and the FOGM variance (σ^2_{fogm}) from L, with $R_f = \alpha R_0$ where R_f and R_0 are the measurement noise variance used for the filter and for the sample generation, respectively, and α is given in the lower right side. Total 500 synthetic time series were generated with true parameters indicated by horizontal and vertical lines (numbers are given in Figure A-1). The minimum values were found by using a 2-dimensional 3-point grid search technique (see Appendix C). The average parameter estimates are shown in boldface. The FOGM variance is sensitive to R_f for the minimum $\hat{C}$, while the correlation time is sensitive to R_f for the minimum L. In other words, the correlation time can be better estimated by minimizing $\hat{C}$, while the FOGM variance can be better estimated by minimizing L. The case (c) shows that the strong dependence of parameter estimates on R_f disappears when the correlation time and the FOGM variance are independently estimated from $\hat{C}$ and L, respectively.

As can be seen in Figure A-6, the correlation time can be better estimated by minimizing $\hat{C}$, while the FOGM variance can be better estimated by minimizing L. If the correlation time and the FOGM variance are independently estimated by minimizing

$\hat{C}$ and L, respectively, the dependence of parameter estimates on R_f can be reduced (Figure A-6c). Thus one can recover the FOGM parameters not severely affected by incorrect measurement noise variance.

The measurement noise variance can be scaled by the ratio between the FOGM variances from the two cost functions. Figure A-6a shows that the FOGM variance from $\hat{C}$ becomes, on average, a half of the true value when $\alpha = 0.5$ and twice as large as the true value when $\alpha = 2$; the FOGM variance is proportional to R_f. If the FOGM variance from $\hat{C}$ is different from that from L, say two times larger, then R_f can also be thought two times larger. We compared the FOGM variance estimates from $\hat{C}$ (denoted by σ_C^2) and from L (denoted by σ_L^2) shown in Figure A-6. The two FOGM variances approximately show a linear trend with slope α, although slightly off the trend for larger FOGM variances (Figure A-7a). When $\alpha = 0.5$, the ratio ($\sigma_C^2/\sigma_L^2 = 0.45 \pm 0.11$) is not significantly different from α (Figure A-7b). The measurement noise variance can be readily scaled as in $\hat{R} = R_f \times \sigma_L^2/\sigma_C^2$. Since $\sigma_L^2/\hat{R} = \sigma_C^2/R_f$, the cost function $\hat{C}$ is still the minimum because of its dependence on the ratio σ_f^2/R_f (see Figure A-2d).

A.5 Frequency response of a smoother

The frequency response of a smoother is similar to that of a low-pass filter because a smoother estimates low-frequency data trend and reduces high-frequency noise based on a state-space model. In this section, we evaluate the effect of different parameter settings on frequency response of smoothers to FOGM plus white samples. We describe

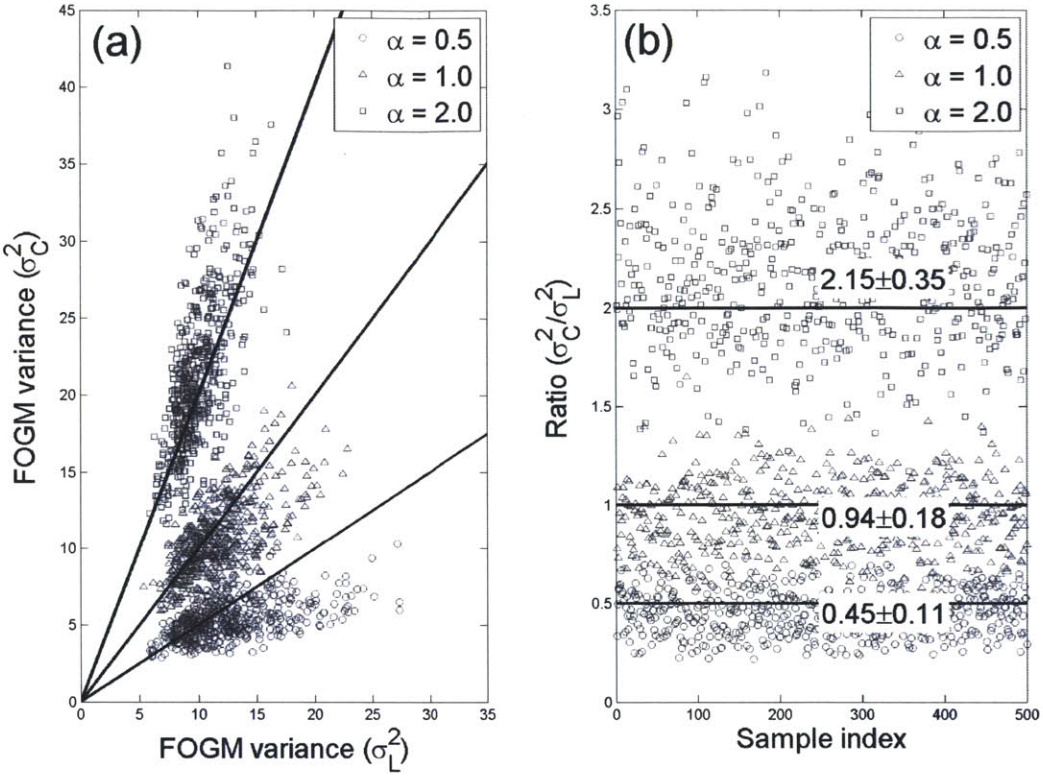


Figure A-7. (a) FOGM variance estimates σ_C^2 from minimum $\hat{C}$ vs. σ_L^2 from minimum L calculated in Figure A-6d. The solid lines have slopes of α used for $R_f = \alpha R_0$. The two FOGM variances follow the linear trend with slope α, although slightly off the trend for larger FOGM variances. (b) σ_C^2/σ_L^2 vs. sample index (total 500 synthetic time series). The horizontal lines indicate the values of α. Also shown are the mean values of the ratios with standard errors. The mean values are not significantly different from the values of α, which suggests that the ratio σ_C^2/σ_L^2 can provide a scale factor (the inverse of α) for correcting the (incorrectly) assumed measurement noise variance (i.e., $\hat{R} = R_f \times \sigma_L^2/\sigma_C^2$).

the frequency response of smoothers in terms of the power spectral density (PSD) of state estimates and transfer functions that describe the characteristics of smoothers and the relationships between the input and the output of smoothers, respectively.

When smoothers were applied to FOGM plus white samples, smoothers pass low-frequency power but attenuate high-frequency power depending on the values of Q_0/R_0

and T/τ_0 (Figure A-8). The FOGM state can be better estimated when both ratios are large. As Q_0/R_0 increases, the spectrum of the smoother output is closer to the true FOGM spectrum and powers at higher frequencies can be recovered, indicating higher cutoff frequency. As expected, low-frequency power can be recovered for larger T/τ_0.

A transfer function of a smoother shows the relation between the input and the output of a smoother as a function of frequency. Here we consider a transfer function as a linear mapping of Fourier transform of the FOGM samples (excluding white noise samples) to the Fourier transform of the smoother output. Since transfer functions are complex-valued functions in general, they can be decomposed into gain and phase functions that show, respectively, the amplitude relationship and phase difference between the input and the output. The overall gain pattern is similar to that of low-pass filters with cutoff frequency depending on Q_0/R_0 and T/τ_0 (Figure A-9). Powers in higher frequencies are more attenuated as Q_0/R_0 decreases. Low-frequency powers are also attenuated for small Q_0/R_0, which means that smoothers are not the same as typical low-pass filters. For a coherence measure, we consider standard errors of phases over 1000 realizations. The standard errors are smaller at low frequencies (Figure A-10), indicating that powers at low frequencies tend to be coherently recovered relative to higher frequencies. For larger Q_0/R_0, the coherence extends to higher frequencies.

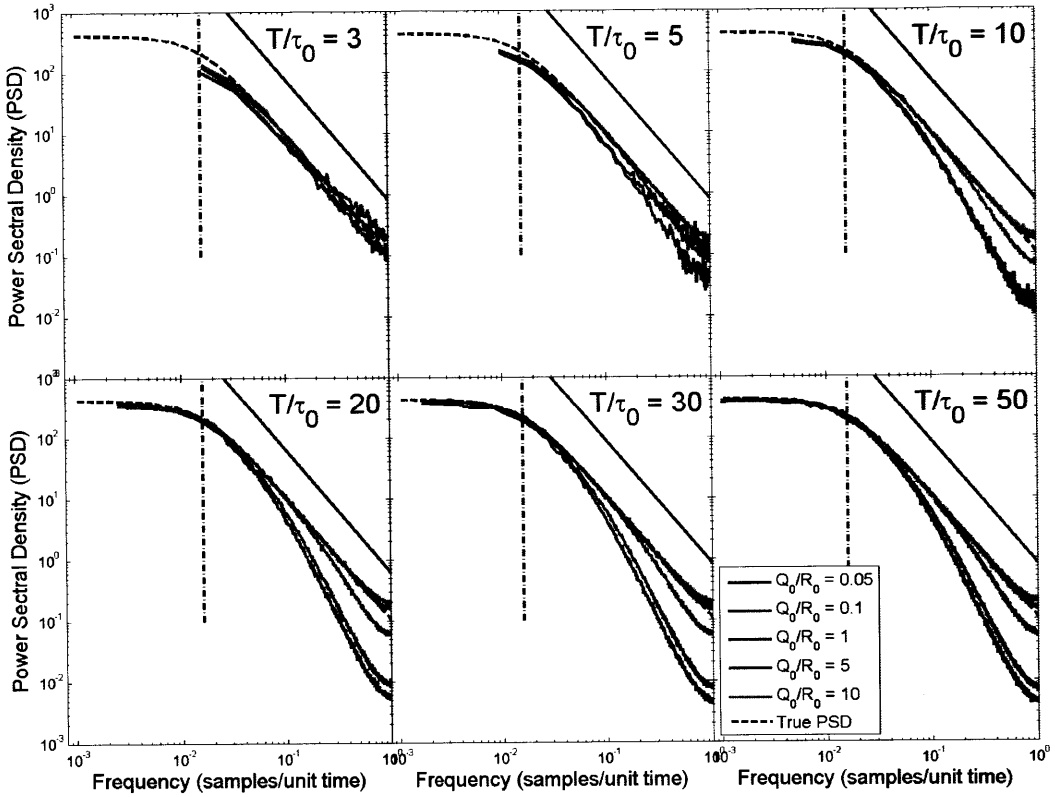


Figure A-8. Power spectral density (PSD) of smoother outputs. Synthetic data were generated with a FOGM process having $\tau_0 = 20$ days and $\sigma_0^2 = 10.5$ mm^2 (equivalently, $Q_0 = 1$ mm^2) plus a white process with zero mean and variance R_0 given by Q_0/R_0. The data span T is given by T/τ_0. The PSD's were averaged over 1000 realizations. The true FOGM spectrum (dashed curves) is flat at low frequencies and has a slope of -2 (solid lines) at high frequencies as do the random-walk processes. The transition from flat to the decreasing power is defined by the corner frequency (dash-dotted vertical lines) given by $f_c = 1/\pi\tau$ (samples/unit time). In general, a smoother loses more powers at higher frequencies. As Q_0/R_0 increases, the PSD is closer to the true spectrum even at high frequencies. As expected, low-frequency power can be recovered for larger T/τ_0.

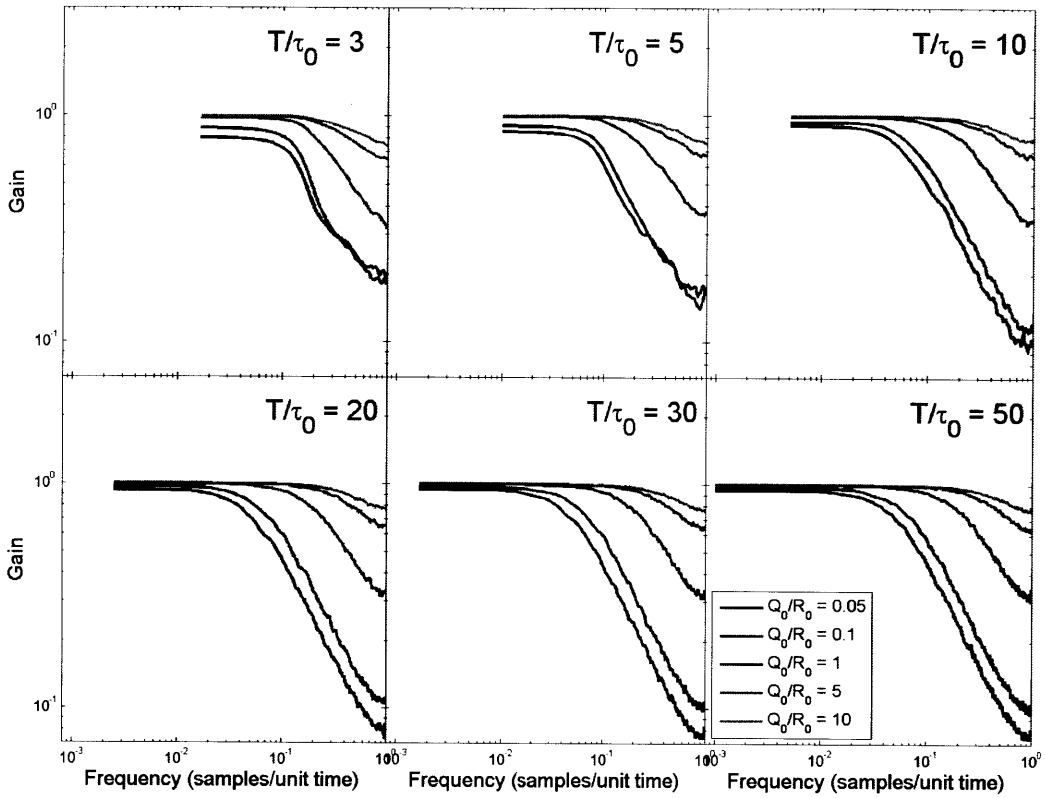


Figure A-9. Gain functions of smoothers used in Figure A-8. Transfer functions were estimated by using Welch's overlapped segment averaging (WOSA) with block size of 50% of the total sample size (T) and 65% overlap with a data taper of zeroth-order discrete prolate spheroidal sequence (dpss) with NW = 4. The dpss is a sequence of length N that has the largest possible concentration of its energy in the frequency interval between –W and W [*Percival and Walden*, 1993]. The gain functions were averaged over 1000 realizations. Similar to low-pass filters, the smoothers pass low-frequency powers but attenuate powers with frequencies higher than the cutoff frequency depending on values of Q_0/R_0 and T/τ_0 . As the value of Q_0/R_0 decreases, powers in higher frequencies are more attenuated. It can be seen for small Q_0/R_0 that low-frequency powers are also attenuated, which is unlikely in low-pass filters.

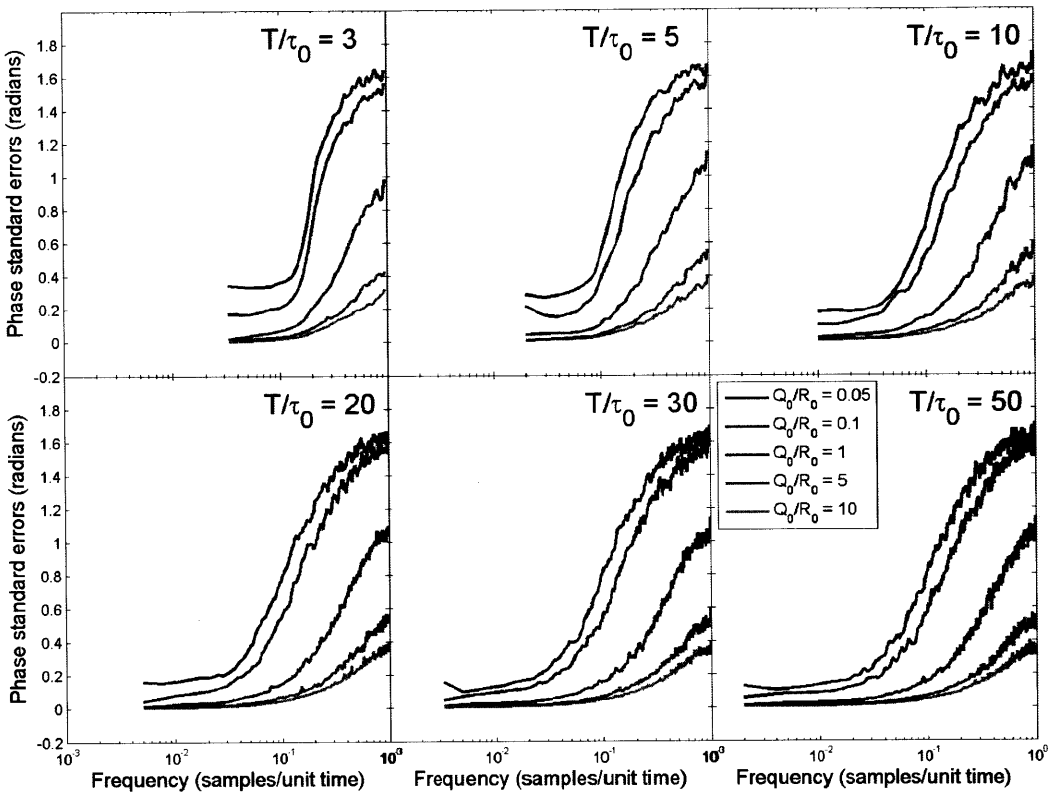


Figure A-10. Standard errors of phase differences between the input and the output of smoothers used in Figure A-8. The phase differences are arguments of transfer functions between $-\pi$ and π. Small standard errors indicate coherence between the input and the output, while large standard errors indicate incoherence. Low frequencies are coherent relative to high frequencies. For large Q_0/R_0, the coherence extends to higher frequencies compared to the cases of small Q_0/R_0.

A.6 Discussion

Given a state-space model, Equation (A-1) provides the innovation variance at each epoch. The innovation variance does not depend on actual measurements but only on the given model. Thus, in suboptimal cases, the innovation variance may not well represent the random properties of the innovations sequence. In most adaptive schemes, the sample innovation variance given in Equation (A-2) is often used as an estimator of the innovation variance. If we replace the innovation variance in Equations (A-3) and (A-4) with the sample innovation variance, the minimization of L is equivalent to the minimization of $\hat{C}$ and X is always zero.

A suboptimal filter may produce the optimal innovation variance due to parameter correlation. The optimal innovation variance underbounds the errors in estimating innovations [*Anderson and Moore*, 1979]. In other words, the sample innovation variance from a suboptimal filter is greater than or equal to that from an optimal filter. The equality is due to the correlation between FOGM parameters or between process (Q) and measurement noise variances (R).

The minimization of the cost functions needs a direct grid search because the minimum variance formulation suffers severe analytical difficulties [*Mohamed and Schwarz*, 1999]. However, the grid search is not efficient for large data sets. For efficiency, a three-point grid search technique [*Kim*, 1997] can be used in a two-dimensional FOGM parameter space. Detailed description of the method can be found in Appendix C. The grid search is similar to another type of adaptive estimation called parallel processing techniques or multiple model adaptive estimation [e.g., *Anderson and Moore*, 1979; *Mohamed and Schwarz*, 1999]. The techniques use a parallel bank of

Kalman filters with parameters on a finite number of grid points. From multiple filters, the techniques usually find a weighted sum of conditional state estimates, while our approach simply finds the FOGM parameters that minimize the cost functions.

Measurement noise is defined to be white Gaussian with zero mean and variance R_k at time t_k and it is often represented by so-called "formal errors" that depend on statistics and sampling used in the GPS geodetic inversion. However, there is a scale issue in formal errors because of temporal correlation of phase observations. The formal errors can be initially used for a smoother and then measurement noise variance can be estimated by the method discussed in Section A.4. Another way is to readjust formal errors such that formal errors become consistent with random data scatter. For example, formal errors can be scaled by comparing the scatter of one-day differences of data with the formal errors (see Appendix H). The one-day differencing reduces the effect of temporal correlation.

This study assumes that the FOGM parameters are constant in time. The assumption, however, is not appropriate when a large transient signal exists in a time interval. Here transient signals are defined as those having temporal characteristics different from time-correlated noise and often originating from earthquakes, volcano activities, groundwater transport, etc. Since transient signals and time-correlated noise are estimated by a single FOGM state, transient signals can affect the estimation of time-correlated noise and vice versa if proper models for transient signals are not available. Suppose that a GPS time series contains a small background noise and a large transient slip in a short time interval. The background noise variance will be estimated somewhat larger due to the large slip and the FOGM state estimates will poorly track the slip due to

the small noise. If the slip interval is known a priori, such misfits can be mitigated by increasing the filter process variance (Q_f) within the slip interval. The innovations in the slip interval are not included in the minimization of cost functions so that the FOGM parameters are estimated from only the background noise. This approach does not require models for transient signals.

GPS time series contain outliers and data with large uncertainties. Assuming Gaussian statistics, outliers are defined as data having absolute values larger than a cutoff value (e.g., five times larger than the RMS error of the residuals from state estimation only with deterministic states). Model misfits tend to make good data appear to be outliers. Initially a large cutoff value is used. Large uncertainty data are defined as data having formal errors larger than a cutoff value (e.g., five times larger than the median value of formal errors). Outliers and large uncertainty data can be removed iteratively until no such data are detected.

When the cost functions are calculated, some of the innovations (particularly at the beginning of the filter run) may appear to be outliers due to inadequate state model and initial conditions. Assuming Gaussian statistics, we define outliers as innovations whose absolute values are larger than a cutoff value (e.g., five times the standard error of the innovation sequence). Outliers are ignored in the calculation of the cost functions.

Since PCA takes data measured simultaneously, a large data gap makes a considerable amount of data directly unusable. Missing data interpolation retains measured data that otherwise would be ignored in PCA. Examples of interpolations are cubic splines [*Aoki and Scholz*, 2003], a three-point Lagrangian interpolation [*Dong et al.*, 2006], spatial averaging [*Savage and Thatcher*, 1992; *Dong et al.*, 2006], an iterative use

of PCA [*Dong et al.*, 2006], and Kalman filtering [*Kawamura and Yamaoka*, 2006]. Kalman filters estimate missing data and their uncertainties based on stochastic nature of data (Appendix B). The uncertainty increases as the interpolation is far away from the known data. The interpolation uncertainties allow us to assess the impact of the interpolation on PCA.

A.7 Conclusions

The signal-to-noise ratio (SNR) of GPS position time series can be improved by state estimation. A FOGM process is used to account for temporal correlation in GPS position time series. The FOGM correlation time can be estimated by minimizing $\hat{C}$ (i.e., the minimum sample innovation variance) and the FOGM variance by minimizing L (i.e., the maximum likelihood). The measurement noise variance is obtained by scaling its assumed value by the ratio of the FOGM variance from the minimum L to that from the minimum $\hat{C}$. Although depending on sample size, parameter correlation, relative magnitude of colored and white noise, and the SNR, a smoother based on a forward-backward Kalman filtering can improve the SNR of GPS position time series by passing low-frequency power (colored noise and possibly transient signals) by attenuating high-frequency power (mostly white noise).

Appendix B

State estimation: a filter/smoother approach

The state estimation is summarized here in a filter/smoother formalism and details can be found in general textbooks [e.g., *Gelb*, 1974; *Anderson and Moore*, 1979; *Brown and Hwang*, 1997]. Consider a linear, discrete-time, finite-dimensional system. Let $\mathbf{x}_k$ be a state vector at time t_k where $k = 1,2,\cdots,n$ and n is the total number of measurements. The state vector evolves in time according to the system dynamics represented by

$$\mathbf{x}_{k+1} = \mathbf{\Phi}_k \mathbf{x}_k + \mathbf{w}_k \tag{B-1}$$

where $\mathbf{\Phi}_k$ is the state transition matrix for the time step from t_k to t_{k+1} and $\mathbf{w}_k$ is the process noise vector assumed to be white Gaussian with zero mean and a covariance matrix $\mathbf{Q}_k$. The state-vector elements with zero process noise are said to be deterministic. Other elements in the state vector with process noise are called stochastic. Stationary

stochastic processes are not required here. The measurements $\mathbf{y}_k$ are modeled by a linear combination of the state-vector elements with a random perturbation

$$\mathbf{y}_k = \mathbf{H}_k \mathbf{x}_k + \mathbf{v}_k \tag{B-2}$$

where $\mathbf{H}_k$ is the design or measurement matrix and $\mathbf{v}_k$ is the measurement noise vector assumed to be white Gaussian with zero mean and a covariance matrix $\mathbf{R}_k$. The process and measurement noise ($\mathbf{w}_k$ and $\mathbf{v}_k$) are assumed to be uncorrelated. Optimality requires that the system model ($\mathbf{\Phi}_k$ and $\mathbf{H}_k$) and noise statistics ($\mathbf{Q}_k$ and $\mathbf{R}_k$) are exactly known. Unmodeled or mismodeled dynamics will degrade the filter performance and can even lead to a divergence of the filter.

The Kalman filtering problem is to determine the estimates of one-step predicted and filtered state vector and the associated error covariance matrices. Given the state-space equations (B-1) and (B-2), noise statistics $\mathbf{Q}_k$ and $\mathbf{R}_k$, and initial values of the state vector and its variance-covariance matrix, two steps are involved sequentially:

Prediction

$$\hat{\mathbf{x}}_{k+1|k} = \mathbf{\Phi}_k \hat{\mathbf{x}}_{k|k} \tag{B-3}$$

$$\mathbf{P}_{k+1|k} = \mathbf{\Phi}_k \mathbf{P}_{k|k} \mathbf{\Phi}_k^{\mathrm{T}} + \mathbf{Q}_k \tag{B-4}$$

Update

$$\mathbf{K}_{k+1} = \mathbf{P}_{k+1|k} \mathbf{H}_{k+1}^{\mathrm{T}} \left(\mathbf{H}_{k+1} \mathbf{P}_{k+1|k} \mathbf{H}_{k+1}^{\mathrm{T}} + \mathbf{R}_{k+1} \right)^{-1} \tag{B-5}$$

$$\hat{\mathbf{x}}_{k+1|k+1} = \hat{\mathbf{x}}_{k+1|k} + \mathbf{K}_{k+1} \left(\mathbf{y}_{k+1} - \mathbf{H}_{k+1} \hat{\mathbf{x}}_{k+1|k} \right) \tag{B-6}$$

$$\mathbf{P}_{k+1|k+1} = \left(\mathbf{I} - \mathbf{K}_{k+1} \mathbf{H}_{k+1} \right) \mathbf{P}_{k+1|k} \tag{B-7}$$

where $\mathbf{K}_{k+1}$ is the Kalman gain at t_{k+1}, $\hat{\mathbf{x}}_{i|j}$ denotes a state vector estimated at t_i based on all of the measurements up to t_j, and $\mathbf{P}_{i|j}$ is an error covariance matrix of $\hat{\mathbf{x}}_{i|j}$. In Equation (B-3), the state vector is predicted according to a given state-space model. When measurements become available, Equation (B-6) provides a linear combination of the model prediction and the data, weighted by the Kalman gain that minimizes the mean square error, to give an improved estimate of the state vector, which is used for the prediction at the next epoch. The error covariance matrices are obtained in Equations (B-4) and (B-7). Note that $\mathbf{P}_{k+1|k}$, $\mathbf{K}_{k+1}$, and $\mathbf{P}_{k+1|k+1}$ are independent of the measurements so that they can be computed before the filter is actually run (generally not true in nonlinear filtering problems). The Kalman filter operates in a recursive mode; it needs only the estimates at the present time for the estimates in the future, independently of those in the past. For unique solutions, the filter starts with an initial condition (or a priori constraint) on both the state vector and its covariance matrix.

Note that the terms inside parentheses in Equations (B-6) and (B-5) are called an innovation at time t_{k+1} and its covariance matrix, respectively. The innovations were first used by Wiener [*Kailath*, 1968] and have been widely used in the engineering literature particularly for adaptive Kalman filtering to estimate the noise statistics (i.e., $\mathbf{Q}_k$ and $\mathbf{R}_k$). See Appendix A for more details.

In smoothing problems, measurements later than the current epoch can be used in the state estimation. With more measurements, smoothers perform better than filters (i.e., smaller error variances) depending on the signal-to-noise ratio and state-space model dynamics. Smoothers improve the estimates of stochastic state-vector elements, while

final filtered solutions are the same as the smoothed estimates for deterministic state-vector elements. Here we employ a fixed-interval smoother based on the Fraser-Potter forward-backward filter approach [*Fraser and Potter*, 1969; *Gelb*, 1974; *Brown and Hwang*, 1997]. *Herring et al.* [1990] provided simple formulas for this smoother:

$$\mathbf{K}_k^s = \mathbf{P}_{k|k}^f (\mathbf{P}_{k|k+1}^b + \mathbf{P}_{k|k}^f)^{-1} \tag{B-8}$$

$$\hat{\mathbf{x}}_k^s = \hat{\mathbf{x}}_{k|k}^f + \mathbf{K}_k^s (\hat{\mathbf{x}}_{k|k+1}^b - \hat{\mathbf{x}}_{k|k}^f) \tag{B-9}$$

$$\mathbf{P}_k^s = \mathbf{P}_{k|k}^f - \mathbf{K}_k^s \mathbf{P}_{k|k}^f \tag{B-10}$$

where $\mathbf{K}_k^s$ is the smoother gain matrix and the superscripts s, f, and b indicate the solutions from smoothing, forward and backward filtering, respectively. Smooth estimates are obtained by linearly combining the estimates from forward and backward running Kalman filters. The updated solutions from the forward filter and the predicted solutions from the backward filter are combined so that measurements are not taken twice.

For the backward running filter, the prediction equations (B-3) and (B-4) must be modified due to the state transition reversal [*Brown and Hwang*, 1997; *Wunsch*, 2006]:

$$\hat{\mathbf{x}}_{k|k+1}^b = \mathbf{\Phi}_k^{-1} \hat{\mathbf{x}}_{k+1|k+1}^b \tag{B-11}$$

$$\mathbf{P}_{k|k+1}^b = \mathbf{\Phi}_k^{-1} (\mathbf{P}_{k+1|k+1}^b + \mathbf{Q}_k) \mathbf{\Phi}_k^{-1^{\mathrm{T}}} \tag{B-12}$$

The matrices $\mathbf{\Phi}_k$ and $\mathbf{Q}_k$ are those used in the forward filter for the time step from t_k to t_{k+1}. The transition matrix must be invertible by definition.

For unique solutions, both forward and backward filters have initial values of the state vector and its uncertainty. One would use weak constraints of $\hat{\mathbf{x}}_{1|0}^f = \mathbf{0}$ and $\mathbf{P}_{1|0}^f = u_f \mathbf{I}$ where u_f is a number large enough not to affect parameter estimates but not so large as

to cause significant numerical rounding errors [*Herring et al.*, 1990]. For the backward filter, the initial condition must satisfy $\mathbf{P}_n^s = \mathbf{P}_{n|n}^f$ because the last solution from the forward filter has already included all the measurements. From Equations (B-10) and (B-8), the condition requires infinite $\mathbf{P}_{n|n+1}^b$ unless an a posteriori constraint matrix is applied. We put $\hat{\mathbf{x}}_{n|n+1}^b = \hat{\mathbf{x}}_{n|n}^f$ and $\mathbf{P}_{n|n+1}^b = u_b\mathbf{D}$ where $\mathbf{D}$ is a diagonal matrix with elements as diagonals of $\mathbf{P}_{n|n}^f$ and u_b is a large number to the extent without numerical rounding errors. We put 10^6 for u_f and u_b when data are in millimeter.

Kalman filters can estimate missing data and their uncertainties based on the state-space model and the stochastic nature of data. Missing data interpolation can be done by performing only prediction steps or zeros with large uncertainties are placed at the location of data missing so that the Kalman gain in Equation (B-5) becomes close to zero and the update step can be skipped. To see how the interpolation works, four synthetic time series were generated with different correlation times and with different lengths of data gaps (Figure B-1). The interpolated data decays to the mean from both ends of a gap. The shape of interpolation is concave up/down when end-points are both positive/negative or an inflected curve when end-points have different sign. If the correlation time is small compared to the length of data gap, the decay is fast and interpolations rapidly approach to the mean (a steady state). With a longer correlation time, interpolations become linear. The uncertainties grow as the interpolations are farther from the known data and reach a steady state (σ_{fogm}) if the filter has a small correlation time compared to the length of the gap (e.g., the last two gaps in Figure B-1).

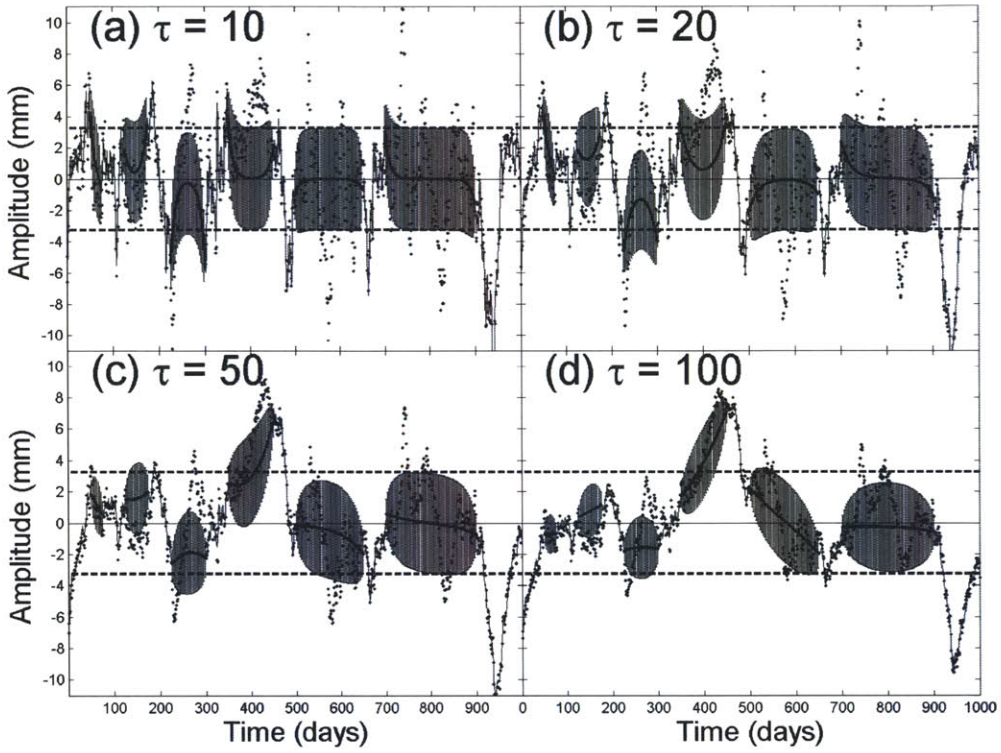


Figure B-1. Interpolation of missing data. Synthetic data (dots) were generated with a first-order Gauss-Markov noise process with correlation time (τ) given in the plot and the variance σ^2_{fogm} = 10.5 mm^2 (i.e., Q = 1 mm^2 for τ = 20 days) plus a white process with σ^2_w = 1 mm^2. All of the time series have the same realization of random numbers. There are six data gaps with intervals of 25, 50, 75, 100, 150, and 200 days from right to left. Shown are the smooth solutions (thin curves), the interpolated solutions (thick curves) and their uncertainties (error bars drawn only for interpolated data at every third points). The horizontal dashed lines indicate $\pm \sigma_{fogm}$. The average of the time series is zero.

Appendix C

Grid search for FOGM parameters

The optimality of state estimation requires that a state-space model is completely known. In most practical situations, however, Kalman filters are obviously suboptimal because the system and noise parameters are unknown or known only approximately. The FOGM parameters can be found by minimizing the innovation-based cost functions as discussed in Appendix A. Here we use a grid search technique to find the minimum cost.

An accurate search would need very fine grids but it is inefficient for a large amount of data. Instead we use a two-dimensional three-point grid search technique. One-dimensional version of the technique can be found in *Kim* [1997]. The technique starts with a given initial search space. Gridding the space 3-by-3, we find a point associated with the minimum cost from the nine points. Then a new search space can be constructed having its center at the minimum point and its boundary determined as small as possible. This procedure will be iteratively done until satisfying specific stopping criteria.

We set the initial space between 1 day and the data time span for the correlation time and between zero and the mean square error of data for the variance. If the initial space does not include the global minimum, the solution may be found at or near a boundary of the space. Even when the initial space contains the global minimum, the search may converge too quickly to get the global minimum. The parameter correlation may result in a local minimum. To avoid a local minimum, we reset the search space once the search has converged such that the center of the reset space is at the converged solution. The search-reset procedure is repeated until satisfying specific stopping criteria.

Whether the search is reset or terminated cannot be determined in a general way; it depend on various factors such as data characteristics (e.g., sample size, outliers, missing data, sampling rate, etc), appropriateness or adequacy of the state-space model, and trade-off between accuracy and efficiency. For a large data set, such factors cannot be handled properly for individual time series. Here we provide a few empirical conditions that can be applied uniformly to a data set. However, the conditions are not sufficient or necessary and the upper and lower bounds may need trial and errors to be set.

We call a session the search steps between resets. The following conditions are in a descending order of priority: (1) The total number of searches must be specified to eliminate possible infinite loop. (2) The first search step in a session will always continue to the second step to escape from first-run stop. (3) If the FOGM variance is greater than a given upper bound (e.g., a few times more than the mean square error), the search will stop. (4) If absolute or percent differences between the FOGM parameters in a search step and in the last session are less than given lower bounds, the search will stop. (5) If either of the FOGM parameters are less than a given lower bound, the search will be reset.

With the same priority, if the search space for either FOGM parameter in a session is less than a given lower bound (e.g., a few times smaller than the initial space in the current session), the search will be reset. (6) If both FOGM parameters are the same as those in the last search step, the search will stop. (7) If the FOGM parameters are greater than given upper bounds, the parameters will be fixed at the upper bounds and the search will stop. (8) If absolute difference between the minimum costs in a current search step and in the last session is less than a given lower bound, the search will be reset.

Appendix D

Detecting transient signals in GPS measurements using principal component analysis

Abstract

We propose a method for detecting anomalous transient signals in Global Positioning System (GPS) measurements. The method consists of state estimation based on Kalman filtering and principal component analysis (PCA). State estimation reduces measurement white noise, accounts for time-correlated noise, and interpolates missing data. First-order Gauss-Markov (FOGM) processes are used to model the stochastic nature of GPS time series. PCA decomposes the FOGM state estimates into principal components (PCs) for temporal variation and sample eigenvectors for spatial distribution. When there are long period temporal correlation in the FOGM state estimates, the PCA generates low-frequency and large-amplitude oscillations in higher-ranked PCs with the frequencies of the oscillations proportional to the PC ranking. We minimize the effect of temporal correlation by removing the sinusoidal oscillations from PCs. Uncertainties in the PCA estimates are approximated by propagating errors for PCs and by using asymptotic distributions with effective sample size for sample eigenvectors. Together with the uncertainties, one can identify transient signals in space from sample eigenvector distribution and in time from lag differences of PCs or residuals from the fits of sinusoidal oscillations. We demonstrate the detection capability of the algorithm with application to the data measured in the Los Angeles basin, California.

D.1 Introduction

Continuously operating Global Positioning System (GPS) networks (e.g., Plate Boundary Observatory in the western US and GPS Earth Observatory Network in Japan) record daily station position changes with an accuracy of a few millimeters. Such dense observations in space and time enable us to identify transient motions such as postseismic deformation [e.g., *Bock et al.*, 1993; *Hudnut et al.*, 2002], slow slip events in subduction zones [e.g., *Dragert et al.*, 2001; *Miller et al.*, 2002; *Ozawa et al.*, 2002], volcanic deformation [e.g., *Dixon et al.*, 1997; *Cervelli et al.*, 2002], and hydrologic deformation [e.g., *King et al.*, 2007], and other processes. Large transient motions such as those due to big earthquakes can be easily detected directly from actual measurements, while weaker motions may go unnoticed due to noise in GPS position time series. Furthermore, the large GPS networks make it time consuming to visually inspect all GPS time series without a priori information. One of the priority objectives of the Southern California Earthquake Center is to "develop a geodetic network processing system that will detect anomalous strain transients" [*Murray-Moraleda and Lohman*, 2010].

The definitions of signals and noise vary with context. There will always be a nonzero probability of misidentifying noise as signal and vice versa [*Allen and Smith*, 1996]. For GPS position time series, a major factor increasing the probability of the misidentification is spatially and temporally correlated noise. Temporally correlated noise has been characterized by stochastic processes such as flicker, random-walk, power-law processes [e.g., *Zhang et al.*, 1997; *Mao et al.*, 1999; *Williams et al.*, 2004; *Langbein*, 2008]. Spatially correlated noise is often referred to as common mode errors in a regional network [e.g., *Wdowinski et al.*, 1997; *Dong et al.*, 2006]. Here we define transient

signals, regardless of whether they are driven tectonically or not, by processes whose temporal evolution differs from an assumed background time-correlated noise model and whose spatial distribution is locally concentrated and systematic rather than random.

In addition to the correlated noise and large network size, detecting small transient signals is challenging for several reasons. Firstly, GPS time series are corrupted by measurement noise defined as white Gaussian. In the perspective of signal detection, the signal-to-noise ratio (SNR) is obviously important and must be improved. Secondly, a general model does not exist for all transient signals. Parametric estimation of transient signals can often be too restrictive (i.e., the form of the transient signal must be assumed). Thirdly, GPS measurements suffer from outliers and missing data. For these reasons, a desirable detection algorithm improves the SNR, accounts for space-time correlation structure, allows nonparametric and/or stochastic approaches, and is reliable and efficient.

The Network Inversion Filter (NIF) [*Segall and Matthews*, 1997; *McGuire and Segall*, 2003] can be used as a signal detector. The NIF estimates the distributions of fault slip and slip rates from surface measurements, accounting for spatially coherent signals, random motions of station monuments, and spatial and temporal smoothing of the slip-rate distribution. The NIF requires fault geometry and assumes homogeneous, isotropic and elastic deformation. Inaccurate fault geometry or inelastic deformation may lead to unreliable fault slip estimates. An approach to avoid source-specific models in the NIF was recently proposed by *Ohtani et al.* [2010], called Network Strain Filter. Similarly our method is based purely on measurements without any source-specific model.

We propose a spatiotemporal filtering approach, state estimation based on Kalman filtering (Appendix A) and principal component analysis (PCA; see Appendix E).

State estimation enhances the SNR in time domain by reducing noise level based on a given state-space model. On the other hand, PCA enhances the SNR in space domain by considering spatial coherence of a transient signal. In addition to a secular rate, annual and semi-annual sinusoids, and offsets, the Kalman filter state vector includes a first-order Gauss-Markov (FOGM) process for temporally correlated noise and any transient signals. PCA decomposes the FOGM state estimates into uncorrelated and variance-maximizing patterns represented by principal components (PCs) in time and sample eigenvectors in space. Such a simple space-time representation with improved SNR enables us to interpret the complex correlation structure of data in a simple way and may reveal transient signals that may not be apparent in actual measurements.

Details of state estimation are described in Appendix A. This study addresses two main issues for use of PCA as a signal detector. Firstly, PCA possibly suffers from the temporal correlation in the FOGM state estimates. As temporal correlation increases, higher-ranked PCs tend to exhibit large-amplitude low-frequency oscillations that could mask weak transient signals. Secondly, uncertainties of PCs and sample eigenvectors are difficult to obtain. Analytic solutions are limited to independent normality and time invariance [e.g., *Anderson*, 2003; *Muirhead*, 2005]. Alternatively, the Monte Carlo method has been frequently used for the uncertainty estimation in crustal deformation studies [e.g., *Savage*, 1988; 1995a; *Smith et al.*, 2007], but the method also has its own weaknesses such as unknown mean values of variables, unknown population PCs, sign switching, and computation at expense (see Appendix G).

In Section D.2, the effect of temporal correlation on PCA is discussed and the uncertainties of the PCA results are obtained approximately. In Section D.3, we attempt

to reduce the effect of temporal correlation by fitting sinusoids to the oscillatory patterns, and identify transient signals by comparing lag differences of PCs or PC residuals with assigned uncertainties. We illustrate detection procedures and demonstrate detection capability with synthetic data in Section D.4 and with real data in Section D.5. Section D.6 and D.7 are discussions and conclusions, respectively.

D.2 Principal component analysis

PCA in this study uses the FOGM state estimates, not actual measurements as done in other crustal deformation studies. Since white noise has been reduced in the state estimation, PCA may provide higher SNR. However, PCA could produce large oscillatory patterns in high-ranked PCs due to temporal correlation in the FOGM estimates, even when raw data do not have the patterns. For detecting weak signals, such oscillations must be reduced. Furthermore, uncertainties are necessary to assess the significance of PCA results, even though PCA is a simple matrix algebra and does not require any assumption of probability distribution [*Jolliffe*, 2002]. In this section, we examine the effect of temporal correlation of the FOGM state estimates on PCA results and obtain PCA uncertainties approximately.

D.2.1 Effect of temporal correlation on PCA

Temporal correlation is one of the major difficulties in signal detection. As correlation time increases relative to the data span, data power tends to be concentrated on high-ranked PCs and the PCs possibly exhibit low-frequency oscillations (Figure D-1). The low-frequency oscillations have been studied in climate research in the formulation of

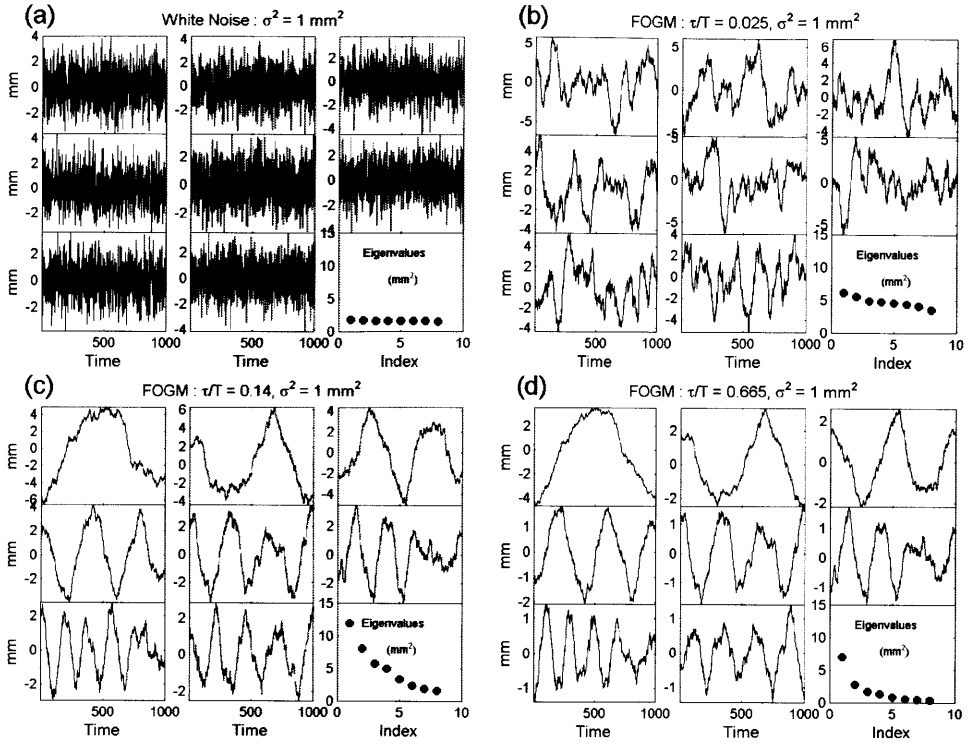


Figure D-1. Top eight PCs from synthetic data from (a) a white process and (b ~ d) first-order Gauss-Markov (FOGM) processes with variance (σ^2) and correlation time (τ) increasing from (b) to (d). In all cases, population eigenvalues are equal to 1 mm^2 (i.e., degenerate). Each data set consists of 1000 samples (T) and 100 variables. Linear trends were removed from the synthetic data before PCA. The ranking of PCs decreases from left to right and from top to bottom. As the correlation time increases, PCs tend to exhibit oscillations and data power represented by sample eigenvalues (solid circles) is mostly concentrated on higher-ranked PCs. The sample eigenvalues in (d) are smaller than those in (c) because sample variance decreases as temporal correlation increases.

singular spectrum analysis [e.g., *Vautard and Ghil*, 1989; *Allen and Smith*, 1994; *Allen and Robertson*, 1996; *Allen and Smith*, 1996]. Even though autoregressive processes of order one (AR(1) including FOGM processes) do not support oscillations, the PCs from AR(1) may exhibit oscillations [*Allen and Smith*, 1996]. The oscillations could mask transient signals if the SNR is not sufficiently high. *Allen and Smith* [1996] demonstrated that an oscillation in climate data cannot be evaluated without considering the low-frequency oscillations from purely red spectra.

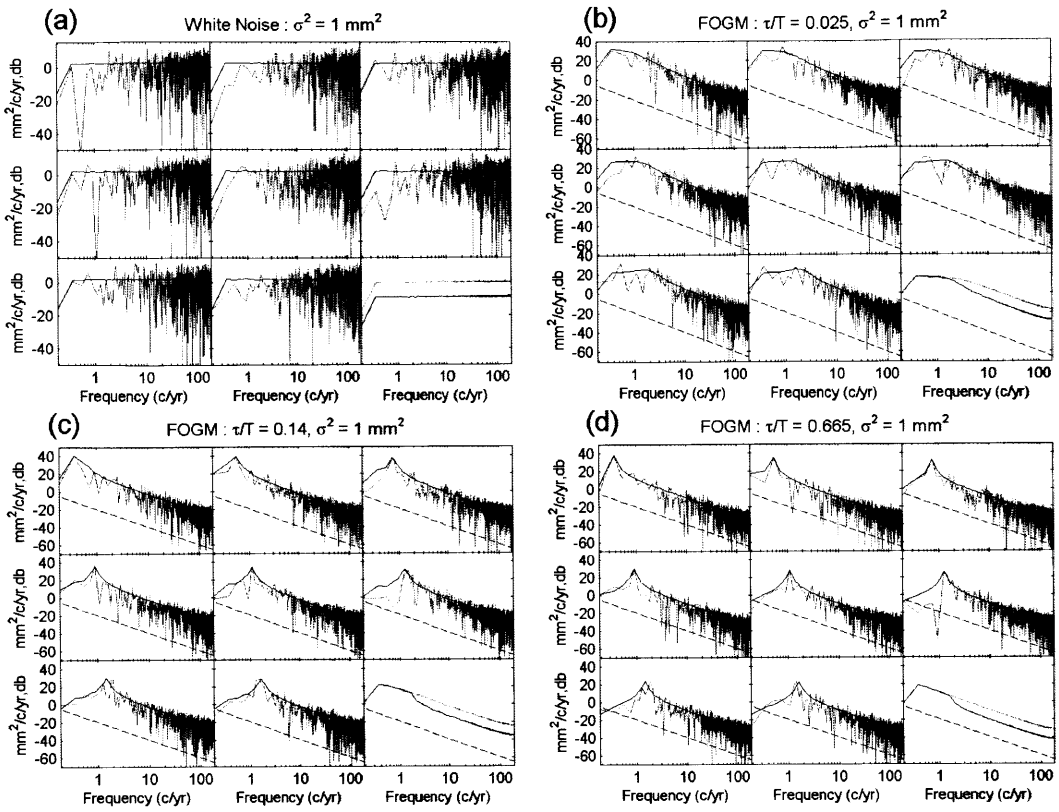


Figure D-2. Periodograms (gray) of the eight PCs shown in Figure D-1. Also shown are the ensemble averages (black) of periodograms from 1000 realizations. The dashed lines indicate the slope of -2 (i.e., the slope of the random-walk process in the log space). Powers at the lowest Fourier frequency are small because linear trends were removed before PCA. (a ~ b) The spectra of the PCs are similar to those of data, while (c ~ d) a spectral peak in each PC appears at a particular frequency proportional to the ranking of PCs. The right-bottom frames in each case show ensemble average of periodograms from data (gray) and from data reconstructed by the eight PCs (black). As temporal correlation increases, low-frequency data powers are well represented by higher-ranked PCs.

Although the amplitudes of the oscillations are varying in time, the frequencies seem to be regular and increase by about 0.5 cycle per data span as PC rank increases; the first PC oscillates with 1.0 cycle per data span, the second PC with 1.5 cycle per data span, and so forth (Figure D-1). For strong correlation, Figure D-2 shows that the periodogram of the j th PC has a spectral peak approximately at an angular frequency

$$\omega_j = \frac{2\pi}{T} \times \frac{(j+\theta)}{2} \tag{D-1}$$

where T is the total data span, $\theta = 1$ if linear trends are removed and $\theta = 0$ if not. If PCA includes data in a shorter time interval while linear trends are estimated for a longer time interval, then $\theta = 0$ because of local linearity. If periodic signals are estimated and removed, the corresponding oscillations will have low power.

The power concentration or the artificial variance compression [*Allen and Smith*, 1996] is due to effective degeneracy (see Appendix F for the white noise case). The effective degeneracy refers to the case when the separation of population eigenvalues is comparable to or smaller than the sampling variation so that sample eigenvectors are unstably determined [*North et al.*, 1982]. In this case, sample eigenvalues are sensitive to the variation in sample covariances and have relatively wide distribution around the population eigenvalue. The PC ranking process always assigns larger sample eigenvalues to higher-ranked PCs so that powers are compressed into higher-ranked PCs. In colored noise cases, variations in sample covariances are more substantial than in the white noise case, so is the power concentration.

Sample eigenvalues can decrease as temporal correlation increases because sample means of FOGM state estimates are removed before PCA to account for the relative variations between variables. For AR(1) with variance σ^2 and correlation time τ, the expected sample variance about the sample mean is given by *Heino et al.* [2000]:

$$\left\langle s^2 \right\rangle = \sigma^2 \frac{T}{T-1}\left(1 - \frac{1+r}{T(1-r)} + \frac{2r\left(1-r^T\right)}{T^2(1-r)^2}\right) \tag{D-2}$$

where r is the autocorrelation at 1-day lag ($|r| = |\exp(-1/\tau)| < 1$). If the process is positively correlated, the sample variance is smaller than the variance, which can be

Table D-1. Values of τ/T associated with maximum sample eigenvalues for first ten components

Component index (j)	$p = 50$ [a]	$p = 100$	$p = 200$	Third-order polynomial[b]
1	0.1381(72)[c]	0.1359(64)	0.1487(60)	0.1388
2	0.1043(33)	0.1043(75)	0.1095(31)	0.1056
3	0.0732(59)	0.0753(60)	0.0751(44)	0.0803
4	0.0561(48)	0.0610(37)	0.0616(26)	0.0619
5	0.0426(22)	0.0495(25)	0.0509(26)	0.0490
6	0.0345(21)	0.0419(23)	0.0427(09)	0.0405
7	0.0289(19)	0.0367(15)	0.0384(12)	0.0351
8	0.0245(20)	0.0314(12)	0.0331(08)	0.0316
9	0.0203(17)	0.0272(08)	0.0305(11)	0.0287
10	0.0182(07)	0.0253(19)	0.0291(05)	0.0251

[a] p is the total number of components.

[b] $f(j) = aj^3 + bj^2 + cj + d$ where j is the component index and the coefficient estimates are $a = -0.0002 \pm 0.0001$, $b = 0.0053 \pm 0.0014$, $c = -0.0476 \pm 0.0081$, and $d = 0.1814 \pm 0.0146$.

[c] 0.1381(72) = 0.1381 ± 0.0072. The mean values and standard errors were obtained from the scatter of 1000 realizations.

substantial for short data and strong correlation. As temporal correlation increases, the sample variance reduction will be larger than the power concentration such that sample eigenvalues start decreasing at a point depending on the PC ranking and the total number of components. Table D-1 shows the values of τ/T associated with maximum sample eigenvalues for some cases. We approximate those values by fitting a third-order polynomial (Table D-1). In the next section we will use the expected sample variance and the polynomial estimates for corrections for PCA uncertainty estimation.

D.2.2 Uncertainties of PCA results

For statistical inferences on PCA results, distributions of sample eigenvalues and eigenvectors can be derived analytically and also asymptotically [e.g., *Anderson*, 2003; *Muirhead*, 2005] (Appendix F), but both approaches are limited to independent normality

and time invariance. The Monte Carlo method (see Appendix G) can readily employ not only time-varying means and variances but also temporal correlations, but it also has its own weaknesses (e.g., unknown mean of variables, unknown population PCs, sign switching, and expensive computation). Here we obtain the uncertainties by propagating errors for PCs and by modifying asymptotic result for sample eigenvectors.

The uncertainty of the j th PC at the i th epoch can be obtained by projecting noise variances onto the principal axes represented by sample eigenvectors:

$$\hat{e}_{ij}^2 = \mathbf{a}_j^{\mathrm{T}} \boldsymbol{\Sigma}_i \mathbf{a}_j \tag{D-3}$$

where $\mathbf{a}_j$ is the j th sample eigenvector and $\boldsymbol{\Sigma}_i$ is a $p \times p$ matrix whose h th diagonal element is the sum of white noise variance from the FOGM state uncertainty at i th epoch and the expected sample variance ($\left\langle s_h^2 \right\rangle$) from Equation (D-2) with FOGM process variance and correlation time estimated by the method proposed in Appendix A. Off-diagonal terms are zero because noise between variables is assumed to be independent.

Uncertainties of sample eigenvectors can be obtained from the asymptotic result, Equation (F-2) in Appendix F, by replacing population eigenvalues and eigenvectors with sample counterparts and sample size (n) with effective sample size (n_{eff}):

$$\mathbf{V}_j = \frac{l_j}{n_{eff}} \sum_{\substack{k=1 \\ k \neq j}}^{p} \frac{l_k}{(l_k - l_j)^2} \mathbf{a}_k \mathbf{a}_k^{\mathrm{T}} \tag{D-4}$$

where $\mathbf{V}_j$ is the variance-covariance matrix of the j th sample eigenvector $\mathbf{a}_j$ and l_j is the j th sample eigenvalue. The denominator in the summation accounts for the separation between sample eigenvalues and the square avoids negative uncertainties. The

effective sample size can be derived by dividing the variance by the expected variance of sample mean [*Thiebaux and Zwiers*, 1984]:

$$n_{eff} = n \times \left[\sum_{\Delta t=-(n-1)}^{+(n-1)} \left(1 - \frac{|\Delta t|}{n} \right) \times r(\Delta t) \right]^{-1} \tag{D-5}$$

where $r(\Delta t)$ is the autocorrelation function of lag Δt. For AR(1) processes with correlation time τ, $r(\Delta t) = \exp\left(-|\Delta t|/\tau\right)$. For high correlation, the effective sample size is too small and the resulting uncertainty becomes too large. In this case, the effective sample size is fixed by replacing τ with the third-order polynomial estimates (Table D-1) when the value of τ/T is larger than the estimates at which sample eigenvalues start decreasing. Since each variable in PCA (i.e., FOGM state estimates) possibly has different correlation time, an average correlation time (weighted by FOGM variances) can be used for the autocorrelation function in Equation (D-5).

Equation (D-3) is not strictly valid because sample eigenvectors are random and can be unstable due to effective degeneracy (Appendix F). We examined with synthetic data how effective degeneracy affects the uncertainty estimation. Unlike the PC uncertainties from the Monte Carlo method, the PC uncertainties from Equation (D-3) are equal for all components because all the variables have the same expected FOGM variance (i.e., $\mathbf{\Sigma}_i = \langle s^2 \rangle \mathbf{I}$ for all i) (Figure D-3). The propagated PC uncertainties of higher-ranked PCs tend to be underestimated, while those of lower-ranked PCs tend to be overestimated. Furthermore, Equation (D-4) provides somewhat larger variances because the largest sample eigenvalue is positively biased due to the degeneracy (Figure D-4). The larger variances do not have any impact on signal detection because the sample eigenvectors are completely uncertain. Note that the uncertainties of sample eigenvectors

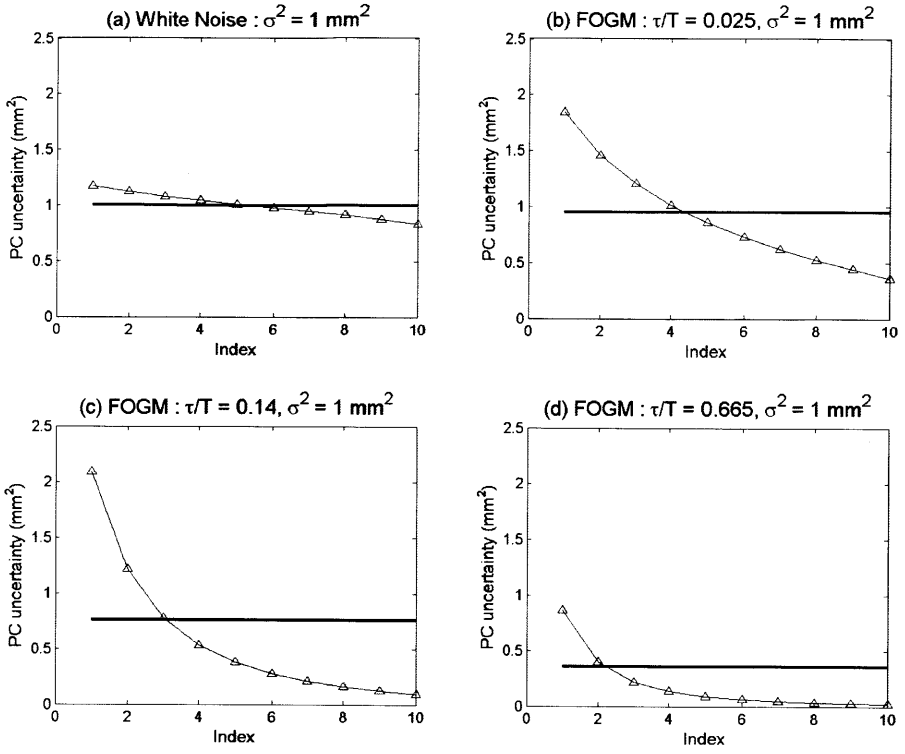


Figure D-3. Uncertainties of PCs from the Monte Carlo simulation (triangles) averaged over 10,000 realizations in time and uncertainties from the error propagation (thick horizontal lines). Synthetic data were generated from (a) a white process and (b ~ d) FOGM processes with variance (σ^2) and correlation time (τ) given in each plot. Each data set has 10 variables with 1000 samples (T). Linear trends were removed before PCA. All population eigenvalues are equal to 1 mm^2 (i.e., degenerate). The ensemble PC uncertainties account for the power concentration due to degeneracy, while the propagated uncertainties are all equal for all components. As temporal correlation increases, the propagated uncertainties decrease to account for the sample variance reduction.

are severely underestimated if the actual sample size is used. The effective sample size in Equation (D-5) makes the uncertainties comparable to those from the simulations.

When effective degeneracy is negligible, errors in Equations (D-3) and (D-4) are small. To illustrate this, we added an arctangent sequence to the FOGM samples used in the previous test such that the first population eigenvalue is widely separated from the

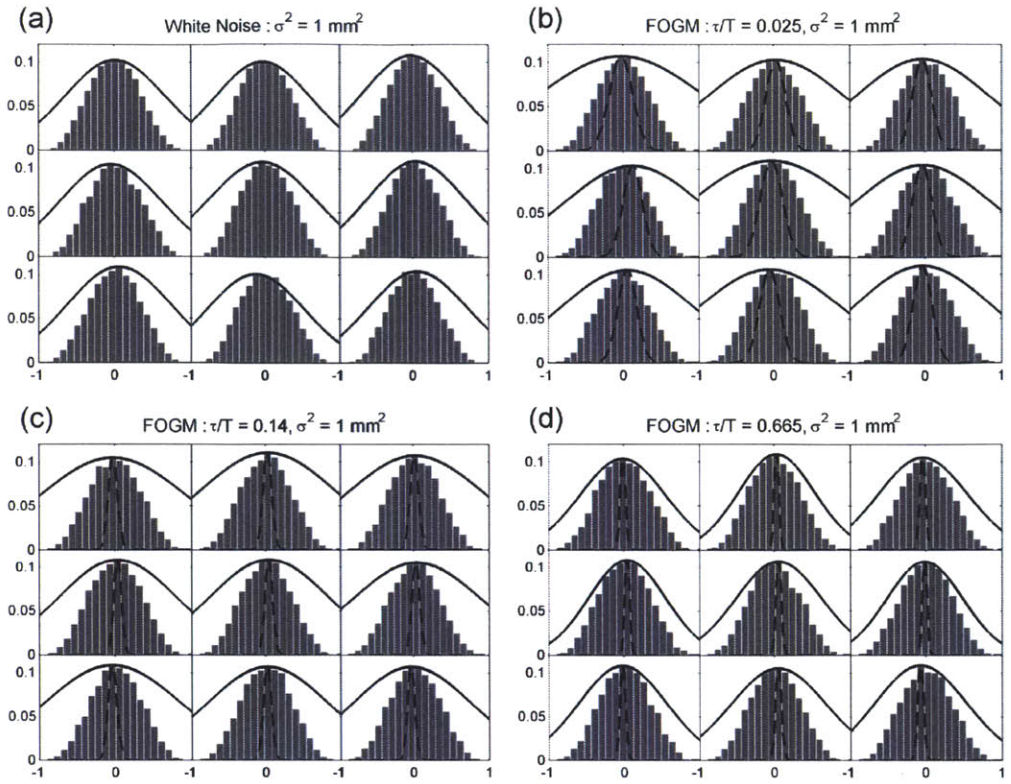


Figure D-4. Distributions of nine elements of the first sample eigenvector from the synthetic data sets used in Figure D-3. The distributions were derived from the Monte Carlo simulations (histograms), Equation (D-4) using the effective sample size (solid curves) and using the actual sample size (dashed curves). The horizontal and vertical axes represent sample eigenvector elements and relative frequencies, respectively. All the sample eigenvectors are uniformly distributed with respect to orthogonal rotations (i.e., the conditional Haar invariance distribution [*Anderson*, 2003]) because population eigenvalues are all equal to 1 mm^2 (i.e., degenerate). The asymptotic distributions are somewhat large due to the power concentration. The effective sample size makes the distributions comparable to those from the simulations, while the distributions using the actual sample size tend to be severely underestimated. (c ~ d) The effective sample sizes are fixed to 4.2 calculated from the third-order polynomial (Table D-1), otherwise they become too small.

others. The first PC uncertainties are close to those from the simulations but those of the other components are biased due to the effective degeneracy (Figure D-5). The asymptotic distributions of the first sample eigenvector are comparable to the simulated ones (Figure D-6).

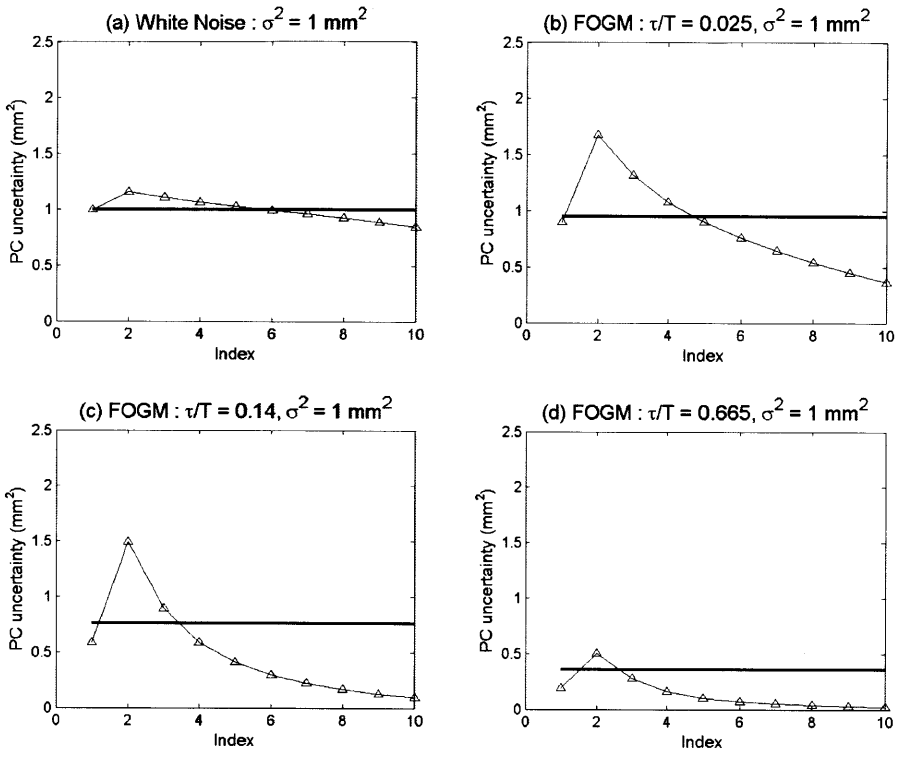


Figure D-5. Same as in Figure D-3 except for the addition of an arctangent sequence to the FOGM processes. Since the addition makes the first population eigenvalue much larger than the others (i.e., effective degeneracy is negligible), the propagated uncertainties (horizontal lines) are close to the simulated ones (triangles). The other nine components are biased due to the effective degeneracy.

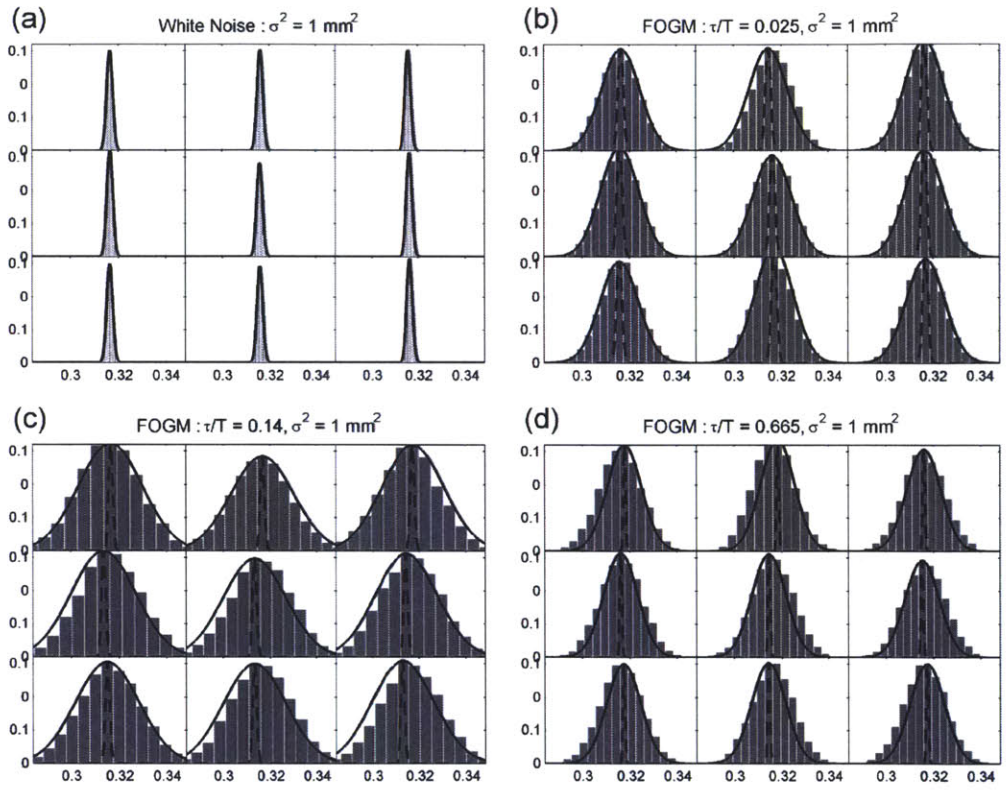


Figure D-6. Same as in Figure D-4 except for the addition of an arctangent sequence to the FOGM processes. The asymptotic distributions (solid curves) are comparable to the simulated ones (histograms) because effective degeneracy is negligible for the first component. All sample eigenvector elements have the same average value (~0.316) because the arctangent was added to variables with the same amplitude. The distributions calculated with actual sample size (dashed curves) have very small variances.

D.3 Signal identification

The PCA uncertainties help us identify transient signals in time and space. The uncertainties of sample eigenvectors from Equation (D-4) can be directly used to identify where a transient signal occurs. However, the uncertainties of PCs from Equation (D-3) may not be directly used because of large-amplitude and low-frequency oscillations in higher-ranked PCs. The PC uncertainties do not account for the power concentration so that the oscillations can be falsely identified as transient signals. In this study, an oscillation in the j th PC can be reduced by fitting sine and cosine functions with angular frequency ω_j given in Equation (D-1) and temporal significance can be assessed by lag differences of the residual PCs (with sinusoids removed) and their uncertainties.

Oscillations in PCs are not necessarily due to temporal correlation but they could represent actual oscillatory patterns. If an oscillation of a PC is locally concentrated and systematic in space, it is likely to be a transient signal. In this case, lag differences of the residual PC will yield no information about the signal. For safe determination, both PCs and residual PCs should be considered in the lag difference calculations.

The objective of lag differencing is to find time intervals that exhibit temporal variations larger than expected. Due to temporal correlation, uncertainties of lag differences can be smaller than those of PCs. In the case of daily GPS time series, a time lag (Δt) is limited to positive integers less than the number of samples. The lag difference of the j th component at the i th epoch is given by $d_{ij}^{\Delta t} = \tilde{z}_{ij} - \tilde{z}_{i+\Delta t,j}$ where $\tilde{z}_{ij}$ can be either the j th residual PC or the j th PC if it shows a systematic and localized spatial pattern. The corresponding uncertainty is given by

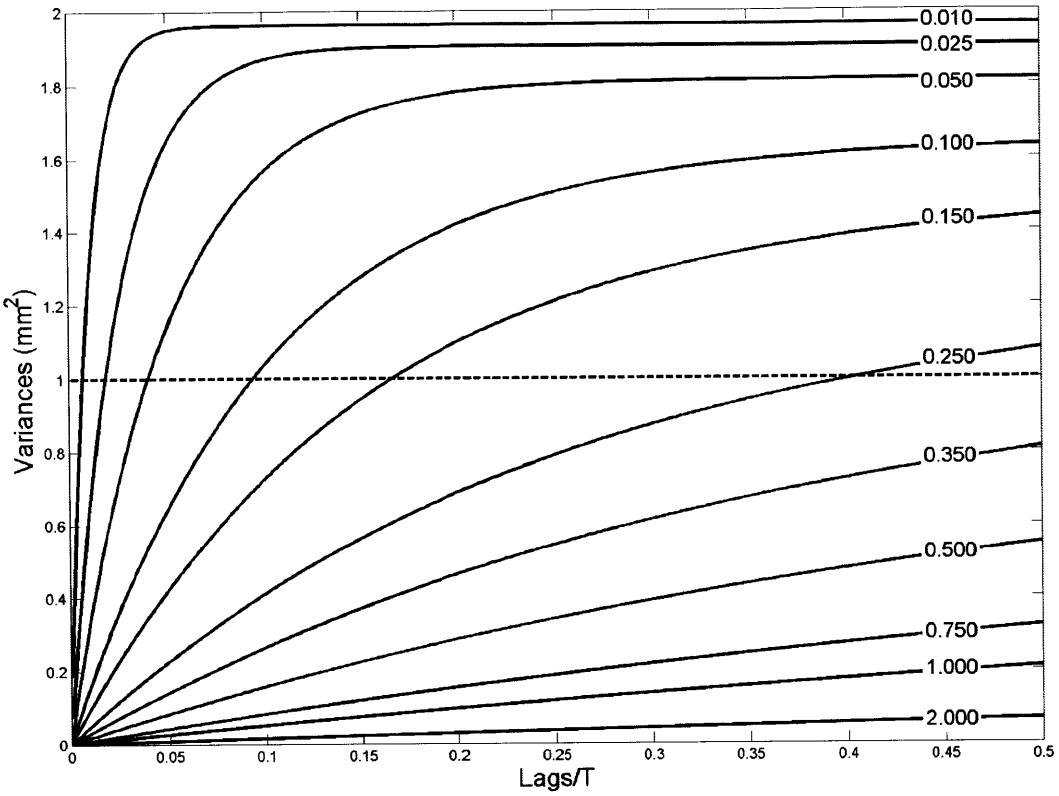


Figure D-7. Uncertainties of lag differences as a function of lags (Δt) scaled by the total time span ($T = 1000$) having different correlation times (τ). The numbers on each curve are the values of τ/T. A one-component case is considered with FOGM variance of 1 mm^2 (dashed horizontal line) for all epochs. For each case, Equation (D-6) becomes $\delta^{\Delta t} = \langle s^2 \rangle \times (2 - 2\exp(-\Delta t/\tau))$ for all epochs. The uncertainties are much smaller than the non-differenced values for long correlation times and/or short lags, while they become larger for short correlation time and/or long lags.

$$\delta_{ij}^{\Delta t} = \hat{e}_{ij}^2 + \hat{e}_{i+\Delta t,j}^2 - 2v_j^{\Delta t} \tag{D-6}$$

where the first two terms in the right hand side can be obtained from Equation (D-3) and $v_j^{\Delta t}$ is the contribution of FOGM autocovariances along the j th principal axis. The autocovariance of the h th variable is given by $c_h^{\Delta t} = \langle s_h^2 \rangle \exp(-\Delta t/\tau_h)$ where $\langle s_h^2 \rangle$ is from Equation (D-2). We have $v_j^{\Delta t} = \mathbf{a}_j^{\mathrm{T}} \mathbf{C}^{\Delta t} \mathbf{a}_j$ where $\mathbf{C}^{\Delta t}$ is a $p \times p$ diagonal matrix

whose h th diagonal element is $c_h^{\Delta t}$. Figure D-7 shows that when correlation times are long and/or lags are short, the uncertainties of lag differences are much smaller than those of non-differenced values, making signal detection more robust. For short correlation times and/or long lags, however, the lag differences have variances larger than the non-differenced values, making them less useful for detecting transient signals.

An anomalous time interval indentified by lag differencing does not necessarily indicate a transient signal. Final decision must be made by investigating whether the spatial distribution of PCs is locally concentrated and systematic in space (e.g., strike-slip or dip-slip faulting, expansion or contraction, uplift or subsidence, etc).

D.4 Numerical examples

Here we test the detection method with synthetic data. We first consider a pure noise case where synthetic time series were generated from FOGM processes with uniformly distributed correlation times and variances (Figure D-8). As expected, the first and the second PCs show large oscillatory patterns; the chi-square per degrees of freedom (χ^2_{dof}) is 10.1 for PC1 and 4.9 for PC2. The PCs are well fitted by sinusoids with frequencies in Equation (D-1) with $\theta = 1$ ($\chi^2_{dof} = 1.1$ for the residual PC1 and $\chi^2_{dof} = 0.5$ for the residual PC2). The sample eigenvectors show no significant localized spatial pattern. All the results suggest that the oscillations in PCs are due to time-correlated noise.

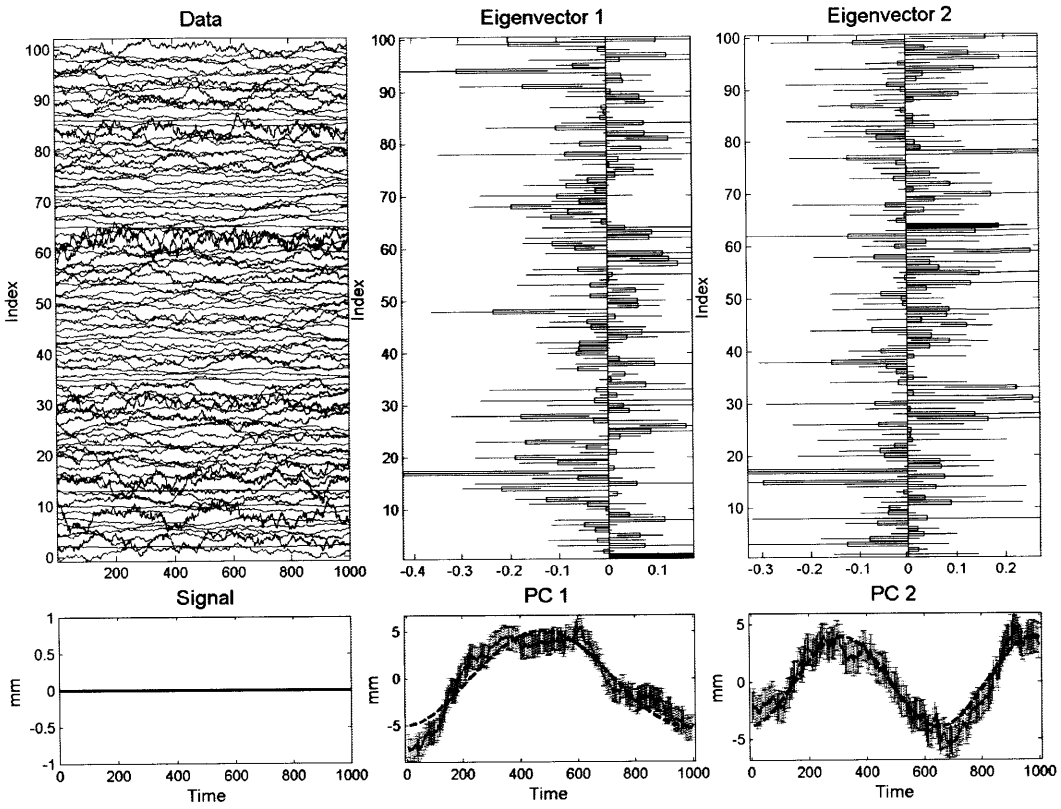


Figure D-8. Plots of synthetic data, signal (zeros), top two sample eigenvectors (horizontal bars) and PCs. The data were generated from 100 FOGM processes with uniformly distributed correlation time (1 ~ 2000 days) and variance (0 ~ 5 mm^2). The average sample variance is 0.75 ± 0.83 mm^2 with minimum of 0.01 mm^2 and maximum of 4.89 mm^2. Linear trends were removed before PCA. Data time series are shifted by 1 mm so that the vertical axis indicates amplitudes as well as data indices. All errors represent 1-sigma uncertainties. Solid bars (one in each component) indicate the eigenvector elements larger than 3-sigma uncertainty. There is no local group of significant eigenvector elements where the locality means the closeness of data indices in the one-dimensional case. The PCs are well fitted by sinusoids (dashed curve) with the frequency in Equation (D-1). PCA results give no clear indication of transient signals.

False positives may appear in lag differences (e.g., the peaks around time 150 in Figure D-9) probably due to incorrect PC uncertainties and FOGM parameter estimates or by random chance. False positives will incorrectly detect noise as a signal. To determine if the anomalies in lag differences are false positives, we chose the time interval between 1 and 300 for the data used in Figure D-8. PCA did not show any

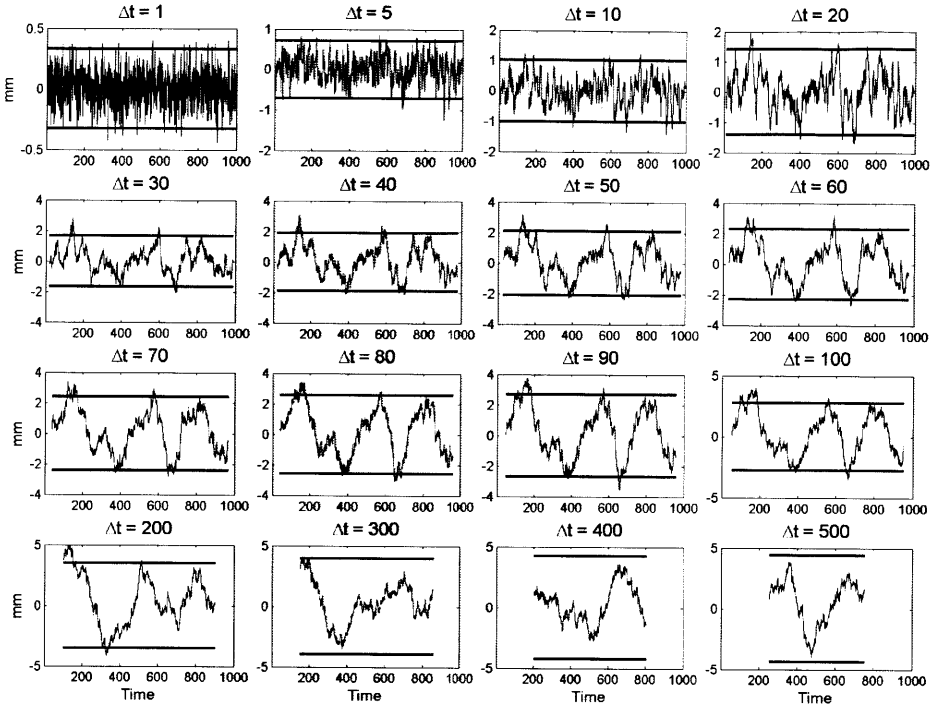


Figure D-9. Lag (Δt) differences of the residual PC1 from the sinusoidal fit shown in Figure D-8. The uncertainties from Equation (D-6) (3-sigma; horizontal lines) are asymptotically increasing as the time lag increases. Anomalies occur around time 150 for Δt =20 ~ 200 (e.g., ~10% of the total number of samples for Δt =100). To see if a transient signal exists, PCA can be performed with data around the time (e.g., between 1 and 300).

anomalous transient signal around time 150 (Figure D-10); the anomalies are false positives. Such determination can be hardly made with the full time interval because of high noise level.

We added an arctangent sequence as a signal to 10 consecutive time series out of 100 FOGM time series used in the previous test. The signal amplitude was chosen so as not to be apparent in the data (Figure D-11). As in the previous case, the PCs show dominant oscillatory patterns (χ^2_{dof} = 9.8 for PC1 and χ^2_{dof} = 5.1 for PC2) and the sample eigenvectors show no significant local group of data. However, the sinusoidal fitting shows slight misfits in the PCs between time 200 and 400, particularly in PC2. Lag-

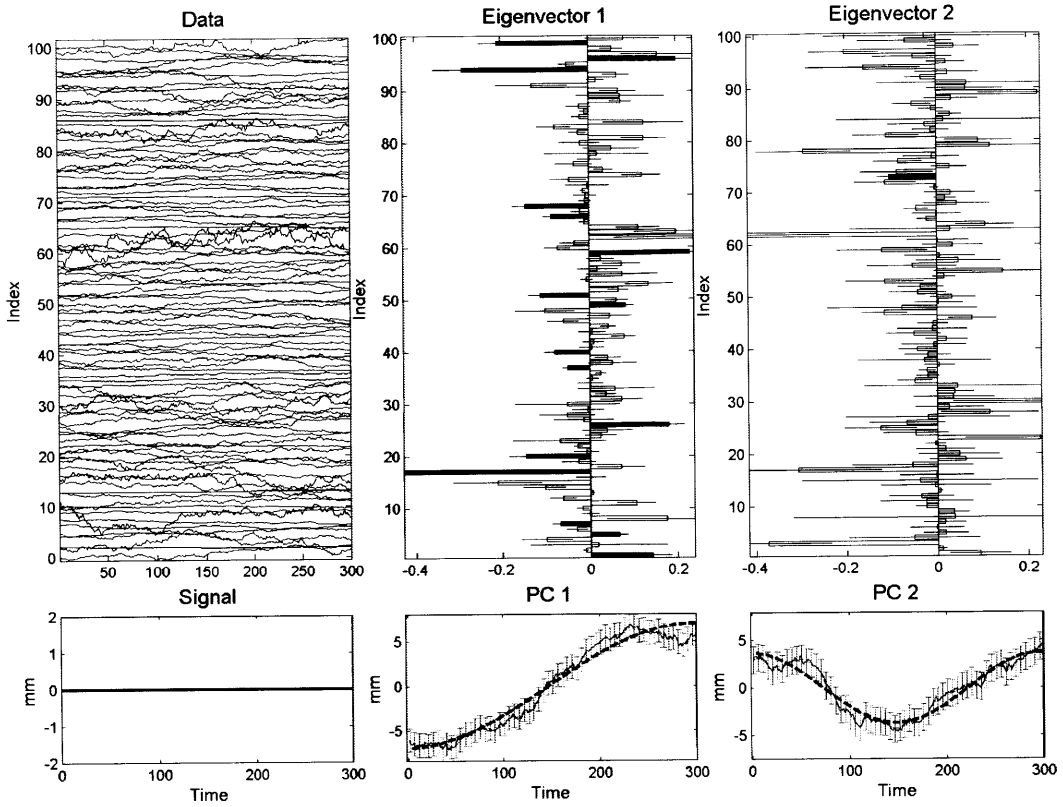


Figure D-10. Same as in Figure D-8 but with time interval between 1 and 300. The data seem to be homogeneous (i.e., no anomalous transient) around time 150. The sample eigenvectors have some significant elements (solid bars) with respect to 3-sigma uncertainty but they are not localized in space. Sinusoidal functions (dashed curves) give a good fit to the PCs (χ^2_{dof} = 0.5 for PC1 and χ^2_{dof} = 0.3 for PC2). The random spatial pattern and the oscillations suggest that a transient signal would not occur in this time interval. Note the fit of the first PC shows half cycle (i.e., Equation (D-1) with $\theta = 0$) because the linear trend was estimated and removed from the entire time series, not to the subinterval of interest here.

differences suggests that a transient signal may exist around the time (Figure D-12). When we chose the data between time 200 and 400, PCA revealed a clear spatial pattern (Figure D-13). When we further select the data between indices 11 and 20, the first sample eigenvector clearly shows that the signal exists relatively uniformly and in the same direction (Figure D-14). The true signal is identically distributed between indices 11 and 20 centered at time 300.

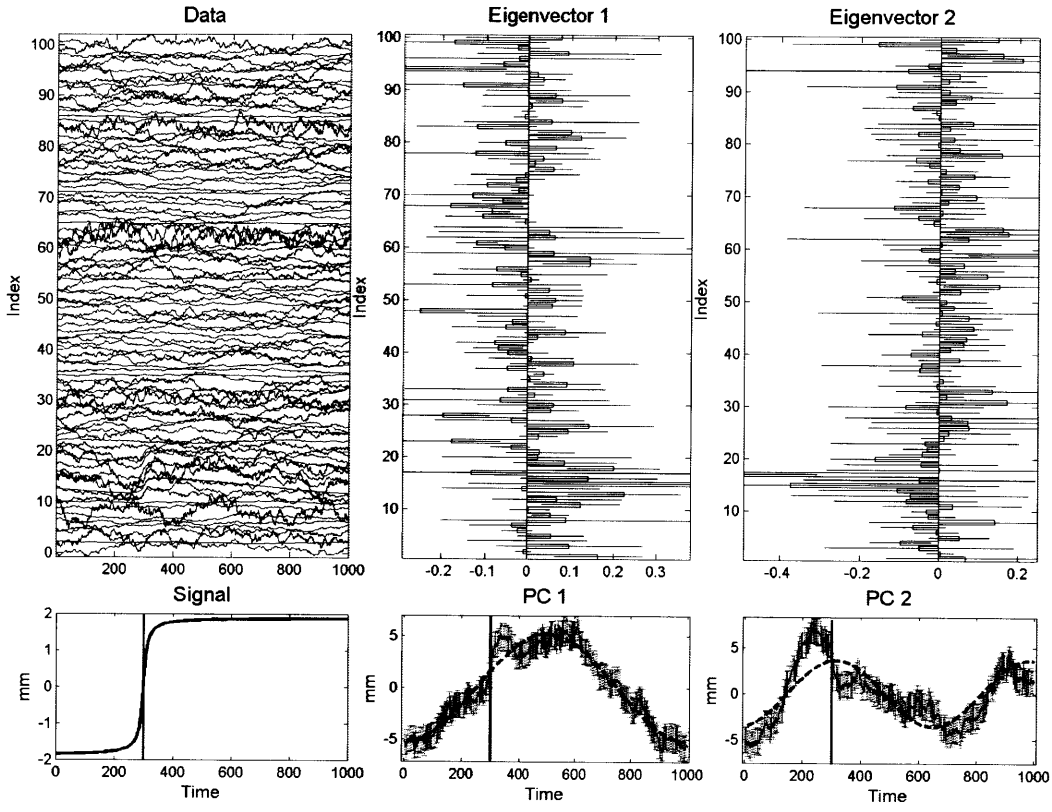


Figure D-11. Same as in Figure D-8 but a transient signal, $1.2 \times \arctan[(t - t_0)/10]$ where time t_0 is 300 (vertical lines), was added to 10 consecutive time series out of the 100 FOGM time series used in Figure D-8. The signal power or variance is 2.63 mm^2 and the average noise variance of the 10 components is 1.53 ± 1.49 mm^2; the signal is not apparent in the original data. There is no significant eigenvector element and sinusoidal functions (dashed curves) give a good fit to the PCs, suggesting that the PCs are mostly affected by time-correlated noise. Slight misfits between time 200 and 400 can be found, particularly in the second PC, but it is not clear at this stage if the misfits are driven by a signal.

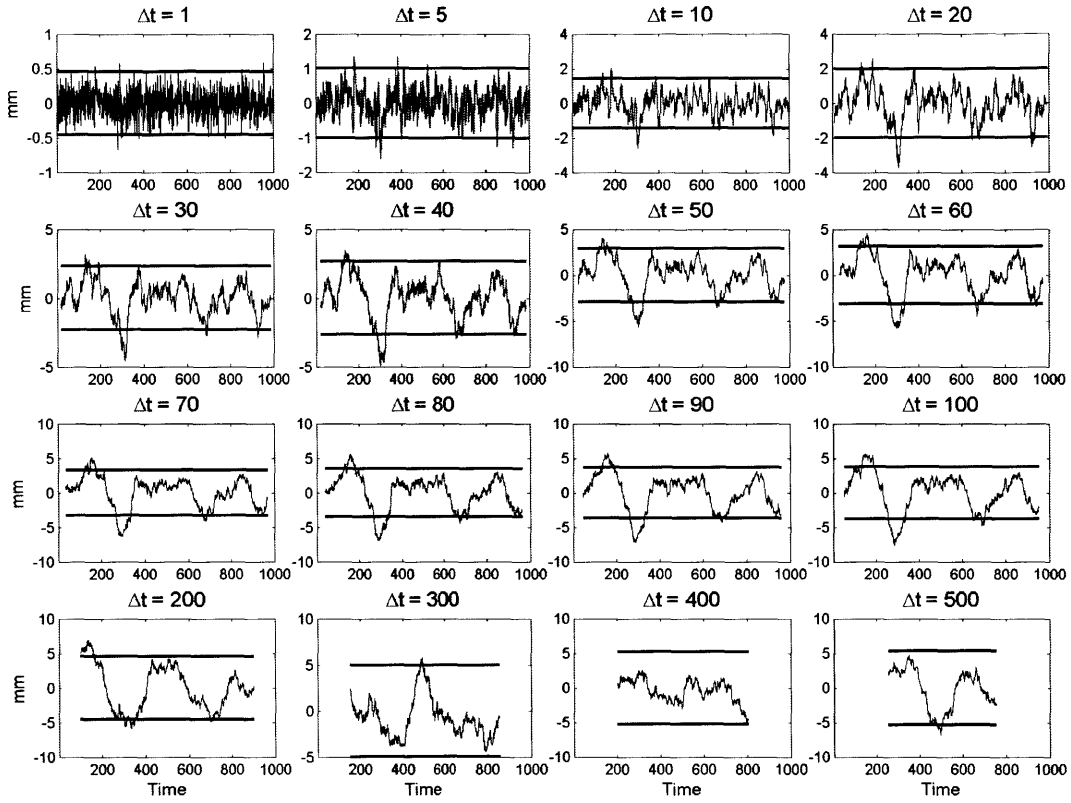


Figure D-12. Same as in Figure D-9 but lag differences of the residual PC2 from the sinusoidal fit shown in Figure D-11. For 30 ~ 100 lags, the time period of 200 ~ 400 is significant with respect to the 3-sigma uncertainties. The number of anomalies is 15.9% at 100 lags.

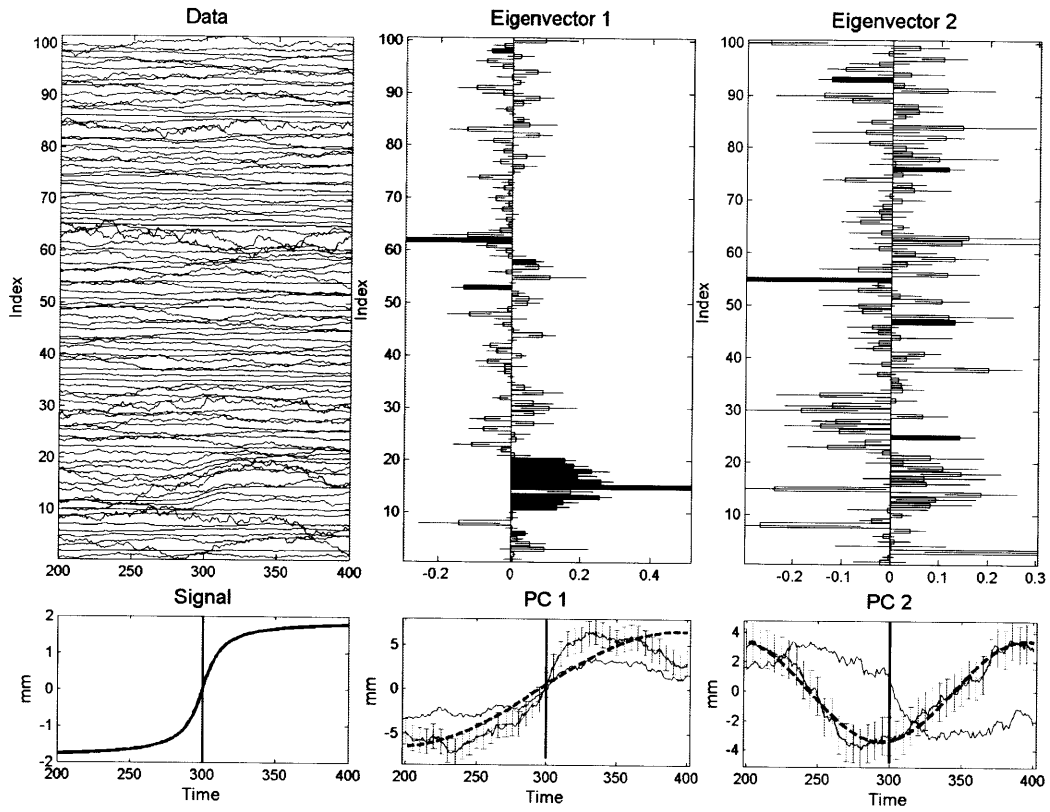


Figure D-13. Same as in Figure D-11 but from the data between time 200 and 400. The first sample eigenvector shows a local group of significant elements around index 15. The first PC looks similar to the signal. The sinusoidal fit (dashed curve) is not good for the first PC ($\chi^2_{dof} = 1.8$). The second PC seems to be driven by noise based on the random spatial pattern and a good sinusoidal fit ($\chi^2_{dof} = 0.1$). The true signal is located at indices from 11 to 20. Index 14 of the first sample eigenvector is not significant. When PCA used data only with indices from 11 to 20, the index 14 became significant (see Figure D-14). The curves without error bars are the PCs shown in Figure D-11. Both PCs exhibit the signal (the sign of the second PC is flipped) so that the first PC is weaker than the first PC from data between 200 and 400.

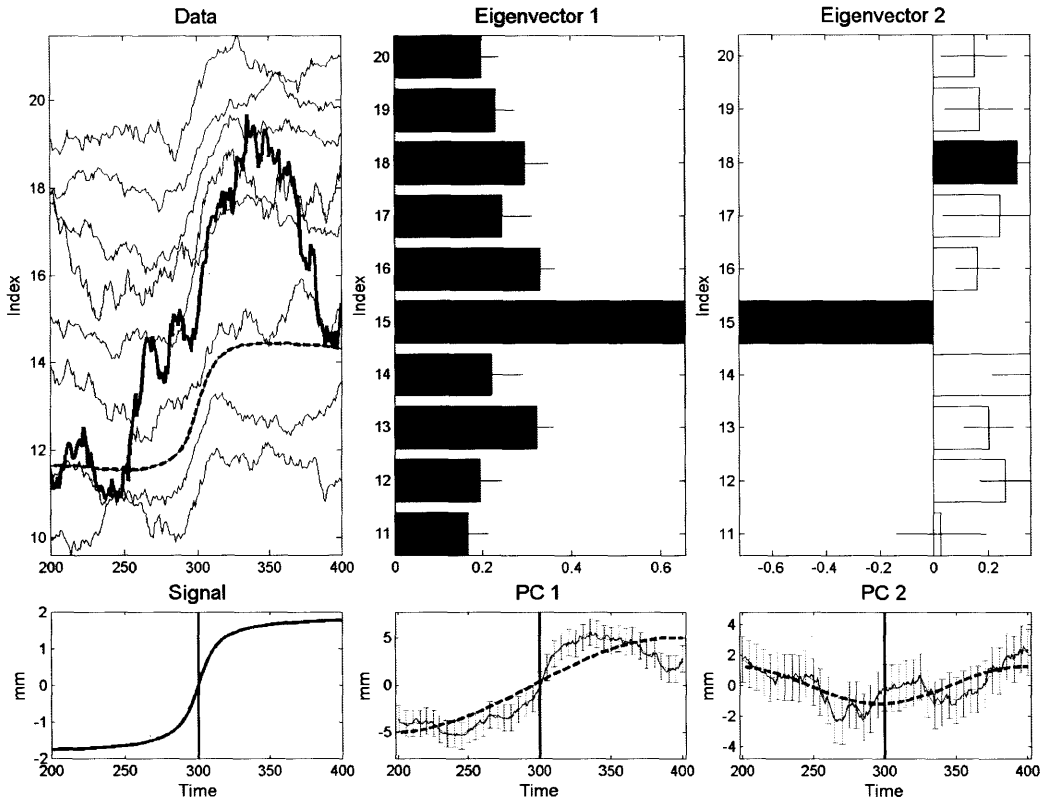


Figure D-14. Same as in Figure D-13 but here the spatial dimension was selected with indices between 11 and 20. Now the temporal variation in the signal can be seen from the data section. The first sample eigenvector clearly shows that the signal exists with indices between 11 and 20 with relatively uniform amplitude and same direction (i.e., same sign). The actual signal is identically distributed into the data with those indices. The element of the first eigenvector at index 14 becomes significant with respect to 3-sigma uncertainty. Larger amplitudes at index 15 result from the amplification of the signal by temporally correlated noise (thick line in the data section). The first PC looks similar to the signal, while the second PC represents local effects particularly due to the data with index 15. The data with index 13 (dashed curve in the data section) is smooth because of its small FOGM noise variance (~0.02 mm^2).

D.5 Data analysis: Los Angeles basin

We applied the algorithm to the daily position data measured at 90 sites of the Southern California Integrated GPS Network (SCIGN) and the Plate Boundary Observatory (PBO) in the area of the Los Angeles basin (33°30′N ~ 34°30′N and 117°30′W ~ 118°30′W). The maximum data span is 16 years (1 January 1994 ~ 31 December 2009) and the sample sizes are between 414 and 5788 with an average of 3518 ± 1265. To remove spatially correlated noise [*Wdowinski et al.*, 1997; *Dong et al.*, 2006], we performed a reference frame transformation by translation, rotation and scaling of the network (i.e., seven-parameter Helmert transformation). We chose 184 reference frame sites over the southern California whose velocity uncertainties are less than 0.1 mm/yr in horizontal and 0.3 mm/yr in vertical when temporally correlated noise are accounted for.

For each GPS time series, we applied the forward-backward Kalman filtering to estimate initial position (1 state), secular velocities (1 state), annual and semi-annual sinusoids (4 states) and FOGM process (1 state). Assuming no a priori information, the 7 states can be generally applied for GPS time series. Seasonal components are important because groundwater-related seasonal deformation has been observed in the Santa Ana basin [*Bawden et al.*, 2001; *Watson et al.*, 2002; *Lanari et al.*, 2004; *Argus et al.*, 2005]. Next we applied PCA to the FOGM state estimates which represent deviations from steady motions including time-correlated noise and transient signals if any. The 16-year time interval was divided into 1-year intervals for PCA to increase the SNR. When a transient signal was detected, the time interval was refined to more focus on the signal.

As examples, we show here two transient signals that were identified: the coseismic deformation associated with the 1999 M7.1 Hector Mine earthquake [e.g.,

Hurst et al., 2000; *Simons et al.*, 2002] (Figure D-15) and the hydrologic expansion of the San Gabriel Valley, California, aquifer system [*King et al.*, 2007] (Figure D-16). In both cases, the first PC from the horizontal components is the only component that is statistically significant (χ^2_{dof} = 10.8 for the coseismic signal and χ^2_{dof} = 12.0 for the hydrologic signal). Lag differences show significant anomalies around 1999.8 for the coseismic signal and 2005.0 for the hydrologic signal. The PCs and lag differences from the vertical components, however, did not show any significant signal in both cases (χ^2_{dof} = 0.6 for the coseismic signal and χ^2_{dof} = 1.3 for the hydrologic signal).

Although the two signals were previously recognized with GPS, the strength of the PCA-based signal detection can be readily shown. The epicenter of the Hector Mine earthquake (34.594°N and 116.271°W) is about 190 km ENE away from Los Angeles. With these far-field stations, PCA detected the weak coseismic offset having maximum displacement of ~6 mm, suggesting detection capability for a weak signal. The hydrologic signal occurred at only 11 sites but was detected from 71 sites available for the time interval of the signal, suggesting detection capability for a large GPS network.

The detection of the earthquake allows us to refine the state-space model. We included an offset at the earthquake epoch in the state vector and used the offset estimates in PCA. The resulting temporal pattern shows the discontinuity due to the earthquake, unlike the smooth PC from the FOGM state estimates (Figure D-15). The displacement field can be described by a dislocation model [*Okada*, 1985] from an inversion including displacements close to the epicenter, but the directions tend to be off as farther away from the epicenter (Figure D-15). One may further include a relaxation term in the state vector for postseismic deformation that occurs throughout southern California and into Nevada

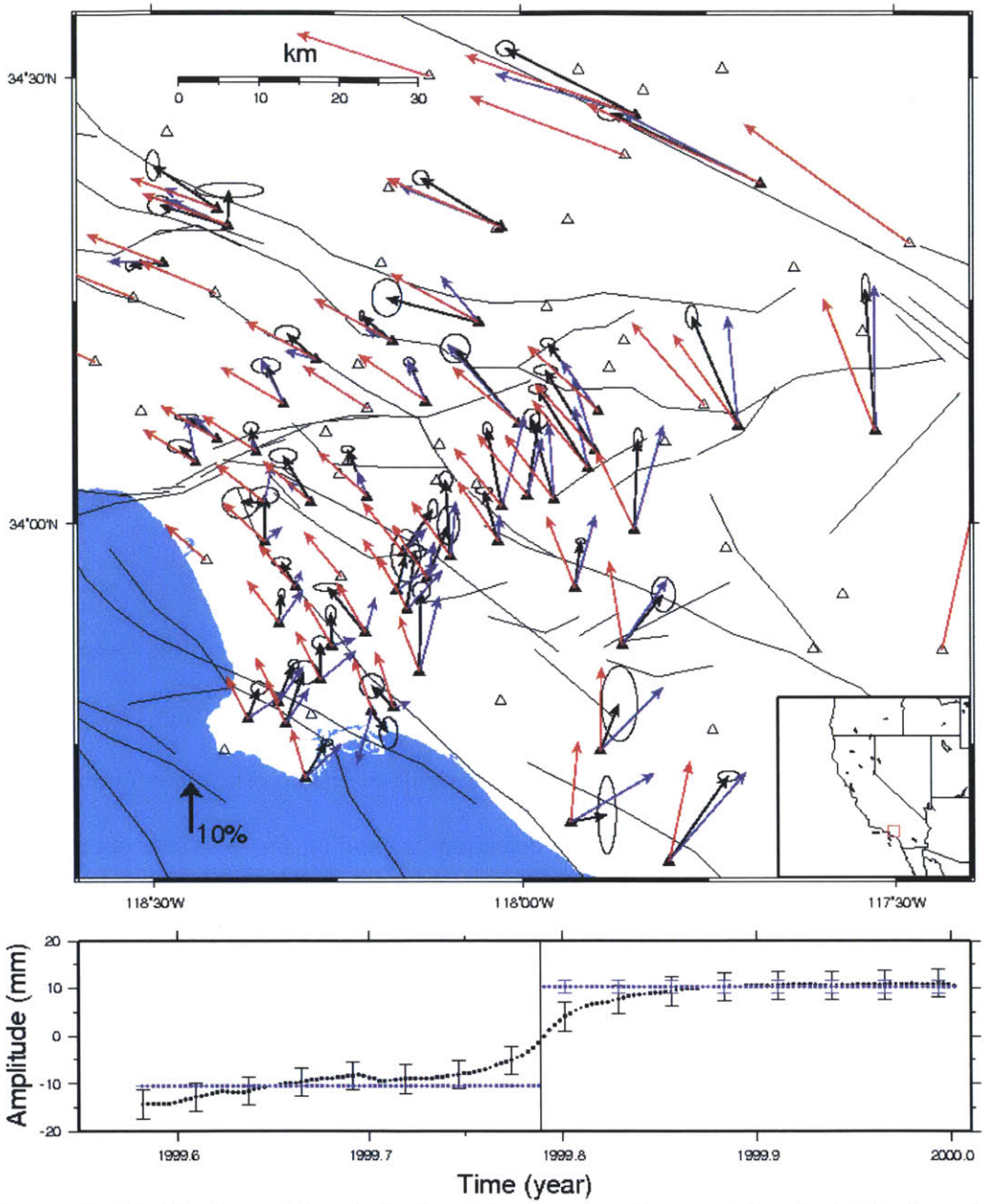


Figure D-15. (Top) spatial and (bottom) temporal patterns of the first PC from the horizontal components of the FOGM state estimates. The patterns from 52 sites (solid triangles) show the coseismic deformation associated with the 1999 M7.1 Hector Mine earthquake. The temporal pattern (black dots) is smooth around the epoch of the earthquake (vertical line in the time plot) that is clearly shown by the PC from the state estimates of coseismic offsets (blue vectors and dots). The displacements from a dislocation model (red vectors) show similar magnitude but the directions tend to be off as farther away from the epicenter. The error ellipses and bars (drawn at every 10 days for clarity) indicate 1-sigma uncertainty. The vertical black arrow indicates 10% amplitude of the PCs and it is equivalent to 2.1 mm for blue and red vectors.

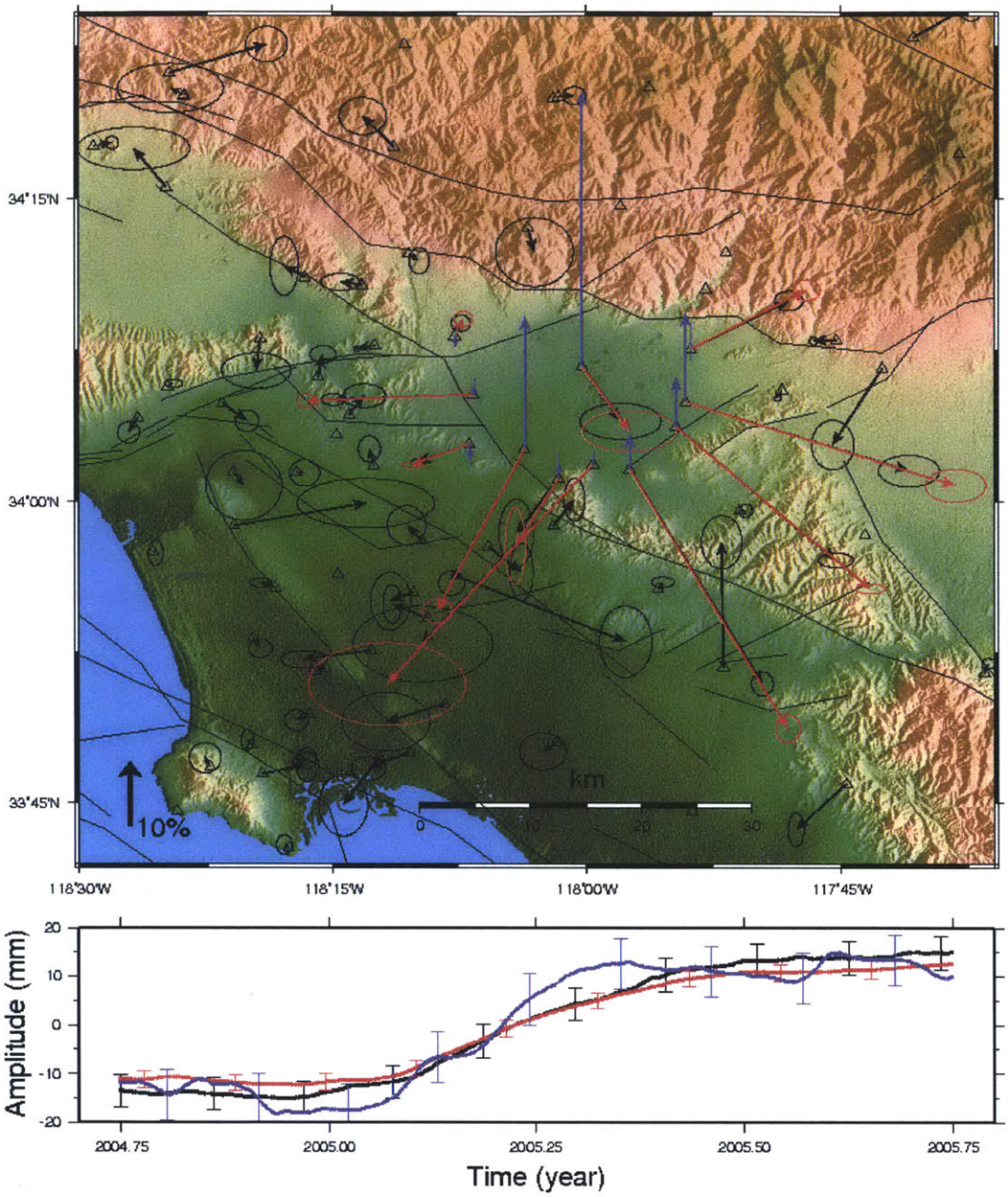


Figure D-16. (Top) spatial and (bottom) temporal patterns of the first PC from the horizontal components of the FOGM state estimates. The patterns from 71 sites (solid triangles) show the hydrologic expansion of the San Gabriel Valley aquifer system. The PCA results from only 11 near-field sites show almost the same pattern but with smaller PC uncertainties (red vectors and curve). Also shown are the results from the vertical components (blue vectors and curve). The vertical PC shows the duration of the event between 2005.0 and 2005.4, while the horizontal PC shows slightly longer duration. The error ellipses and bars (drawn at every 40 days for clarity) indicate 1-sigma uncertainty. The vertical black arrow indicates 10% (5% for vertical) amplitude of the PCs.

[*Freed et al.*, 2007]. However, the postseismic deformation was not detected by PCA because of low SNR at these far-field stations. Near-field stations with large amplitudes may help PCA detect the postseismic deformation in the far field.

The hydrologic signal occurred at the San Gabriel Valley (Figure D-16). The spatial pattern at the sites outside of the valley looks random. To increase the SNR, we chose 11 sites located in the valley and re-estimated the FOGM parameters without considering data before 2002 due to large noise and data between 2005 and 2006 due to the identified signal. In this way, we can remove the signal contribution to the PC uncertainties; the average of the PC uncertainties decreases from 3.4 mm (black error bars in Figure D-16) to 1.7 mm (blue error bars). The smaller uncertainties as well as focused area make the vertical PC become significant ($\chi^2_{dof} = 5.4$; see Figure D-16). The vertical PC shows a temporal pattern similar to the horizontal PC and the lag differences support the significance of the signal. The vertical PC shows the duration of the event between 2005.0 and 2005.4 slightly shorter than that from the horizontal PC. The vertical PC has one dominant site (station LONG at the center of the valley) so that the duration may reflect a local response.

D.6 Discussion

Coherence and orthogonality of PCA results may lead to incorrect interpretation. *Savage* [1988] defined a coherent signal that has its spatial distribution fixed in time and its temporal evolution fixed in space. PCs are orthogonal to each other by definition. Signals from different types of sources or signals propagating in space and time tend to be distributed into several PCs to satisfy the coherence and orthogonality. A physical signal

may be represented by combining two or more PCs without imposing the orthogonality [*Savage and Thatcher*, 1992; *Savage*, 1995b], or by rotating PCs within a specified time interval [*Kawamura and Yamaoka*, 2006].

One of the main objectives of PCA is to reduce data dimension; a few PCs can be selected to cover most of variations in the data set without serious loss of information. *Preisendorfer* [1988] categorized selection rules for significant PCs into three groups. Firstly, space-map selection rules use PC coefficients to examine whether the spatial distribution of a PC is similar to the distribution of a physical signal. For example, *Savage* [1995a] tested if sample eigenvectors represent a sinusoidal pattern as a function of azimuthal angles in horizontal space. But the physical distribution must be specified a priori. Secondly, dominant-variance selection rules concern relative importance of PCs based on sample eigenvalues (e.g., the cumulative percentage of sample eigenvalues and Rule N described in Appendix G). The rules have been widely used but perform poorly when data are highly correlated. Finally, time-history selection rules (e.g., periodograms or autocorrelations of PCs) retain PCs whose temporal pattern is not random. Lag differencing belongs to this group.

PCA is sensitive to local outliers. Here a local outlier is a site having a large temporal pattern (often dominant in the first PC) that is not coherent with nearby sites. The local outliers can be simply ignored in PCA by excluding them in the data set. Another way is to use a sample correlation matrix, called the Karhunen-Loeve expansion [*Dong et al.*, 2006], that is useful when variables are in different units or types and when variances are widely different [*Jolliffe*, 2002]. However, sample covariance matrices are preferred here because interpretations of PCA results are easier. For example, sample

eigenvalues directly represent the total power in each PC. In general, the two types of matrices produce different solutions because PCA is not invariant under linear transformations of variables [*Muirhead*, 2005].

The signal detection depends on a number of factors related to data (e.g., SNR, data size, missing data, outliers, large uncertainty data, etc), state estimation (e.g., state-space models, stochastic noise parameters, a priori constraints, etc), and PCA (e.g., temporal correlation, effective degeneracy, rank deficiency (not considered in this study), etc). It is almost impossible to account for all the factors correctly. However, evidence for a transient signal can be evaluated with various types of products from the algorithm: plots of FOGM state estimates (e.g., data plot in Figure D-11), temporal and spatial patterns of PCs, sample eigenvalues, uncertainties of PCA solutions, and lag differences of PCs (or residual PCs).

Transient signals detected by the algorithm can be further used in geophysical modeling approaches. For example, data reconstructed by the PCs containing transient signals can be used for parametric modeling, finite element modeling, or inversion techniques for fault slip (e.g., the Network Inversion Filter [*Segall and Matthews*, 1997; *McGuire and Segall*, 2003] and the PCA-based inversion [*Kositsky and Avouac*, 2010]). Furthermore, our algorithm can utilize other types of geodetic data (e.g., data from creep meters and strain meters) if the noise characteristics of the data are reasonably known.

D.7 Conclusions

We propose an algorithm, which combines state estimation and PCA, for detecting transient signals from continuous GPS measurements. State estimation accounts for temporal correlation in GPS time series by including a FOGM process in a state vector and PCA extracts variance-maximizing patterns in space and time from the FOGM estimates. Temporal correlation in the FOGM state estimates, however, can produce large-amplitude low-frequency oscillations in high-ranked PCs. We reduce the effect of temporal correlation by removing sinusoidal oscillations from PCs. We approximately obtain PC uncertainties by propagating errors and sample eigenvector uncertainties by using asymptotic distributions with effective sample size. These uncertainties allow us to identify transient signals in space from sample eigenvector distribution and in time from lag differences of the PCs or the residual PCs with sinusoidal oscillations removed. We demonstrate with synthetic and real data that the algorithm provides a simple space-time representation of highly correlated data with improved SNR, which is capable of detecting transient signals that may not be apparent in actual data.

Appendix E
Formulations and properties of principal component analysis

We summarize the formulations and properties of principal component analysis (PCA). More details can be found in general textbooks [e.g., *Preisendorfer*, 1988; *von Storch and Zwiers*, 1999; *Jolliffe*, 2002; *Jackson*, 2003]. PCA has been widely used in geodetic studies [e.g., *Savage*, 1988; *Savage and Thatcher*, 1992; *Savage*, 1995b; a; *Aoki and Scholz*, 2003; *Dong et al.*, 2006; *Kawamura and Yamaoka*, 2006; *Smith et al.*, 2007; *Savage and Langbein*, 2008; *Savage and Svarc*, 2009; *Kositsky and Avouac*, 2010].

We form an $n \times p$ data matrix $\mathbf{X}$ whose element x_{ij} is the i th sample of the j th variable. For GPS time series, samples are tagged with epochs and variables are direction coordinates (e.g. north, east, and vertical). Thus n and p correspond to temporal and spatial dimensions, respectively. Elements in a row vector share common epoch. To describe the variability, samples consist of deviations from the time average of each variable. In the main text and Appendix D, samples are the first-order Gauss-Markov

state estimates that describe temporally correlated noise and transient signals if any with reduced white noise.

PCA requires the concept of eigenvalues and eigenvectors. We define a $p \times p$ sample covariance matrix as $\mathbf{S} = \mathbf{X}^{\mathrm{T}}\mathbf{X}/dof$ where dof and T denote degrees of freedom and matrix transpose, respectively. The sample covariance matrix can be decomposed as $\mathbf{S} = \mathbf{ALA}^{\mathrm{T}}$ where $\mathbf{A}$ is a $p \times p$ orthonormal sample eigenvector matrix and $\mathbf{L}$ is a $p \times p$ diagonal sample eigenvalue matrix. The pairs of sample eigenvalues and sample eigenvectors are ordered in the sense of decreasing sample eigenvalues. If $n > p$ (assumed throughout the thesis), a sample covariance matrix is positive definite. Sample eigenvalues are distinct with probability one due to random sampling variations [*Muirhead*, 2005].

The sample principal components (PCs) are defined by an orthonormal linear transformation [*Jolliffe*, 2002]

$$\mathbf{Z} = \mathbf{XA} \tag{E-1}$$

where $\mathbf{Z}$ is an $n \times p$ sample PC matrix whose element z_{ij} is a value of the j th sample PC at the i th epoch. Since $\mathbf{Z}^{\mathrm{T}}\mathbf{Z}/(n-1) = \mathbf{L}$, sample PCs are uncorrelated and have power (or energy) equal to sample eigenvalues. The larger sample eigenvalues imply the more temporal variations in the sample PCs.

The data matrix can be exactly reconstructed by

$$\mathbf{X} = \mathbf{ZA}^{\mathrm{T}} \tag{E-2}$$

The element a_{hj} of $\mathbf{A}$ is the j th PC coefficient associated with the h th variable; the sample eigenvectors reflect the amount of power from each PC at each stations (in GPS

case). In other words, sample PCs represent common temporal patterns that are distributed in space according to sample eigenvectors. The separation of space and time provides a simplified interpretation of a complicated spatiotemporal multivariate system. Sample PCs have the units of original variables and sample eigenvectors are unitless.

When a correlated signal is present in a data set, its temporal variation can be mostly covered by the sample PCs associated with first a few sample eigenvalues. Synthesized by Equation (E-2) with such sample PCs, dimension-reduced data approximate to the original data with the minimum Euclidean norm between the data and the approximation [*Jolliffe*, 2002]. Similarly, the spatial dimension can be reduced by choosing a subset of the original variables with relatively large sample eigenvectors. In this context, the dimension reduction is equivalent to the noise reduction. There exist many rules to determine how many PCs and variables should be retained [e.g., *Preisendorfer*, 1988; *Jolliffe*, 2002], but they are mostly subjective [*Jolliffe*, 2002]. While the selection of PCs with large sample eigenvalues is useful for data compression, the eigenvalue rank-order does not necessarily indicate statistical or physical significance of PCs [*Allen and Smith*, 1996].

Appendix F

Eigenvalues and eigenvectors from independent Gaussian variables

Here we summarize the statistical properties of sample eigenvalues and sample eigenvectors from independent Gaussian variables. Suppose that a $n \times p$ data matrix $\mathbf{X}$ consists of $x_{ih} \sim N(\mu_{ih}, \sigma_{ih}^2)$ where $N(a,b)$ denotes a Gaussian distribution with mean a and variance b, and n and p denotes the number of samples and variables, respectively. For simplicity, we assume $\sigma_{ih}^2 = \sigma_h^2$ for all i. A $p \times p$ diagonal matrix $\mathbf{\Sigma}$ denotes the noise covariance matrix having σ_h^2 in the h th diagonal. Each variable is centered at its sample mean. Then a sample covariance matrix, $\mathbf{S} = \mathbf{X}^{\mathrm{T}}\mathbf{X}/(n-1)$, has the noncentral Wishart distribution if the mean μ_{ih} is varying in time, otherwise it has the central Wishart distribution [*Anderson*, 1963; 2003; *Muirhead*, 2005]. In the central Wishart case, $\mathbf{S}$ is an unbiased estimator of $\mathbf{\Sigma}$. The following discussion is restricted to the central case. The noncentral case can be handled by Monte Carlo simulations (Appendix G).

Muirhead [2005] analytically derived the joint distribution of sample eigenvalues by integrating the Wishart density with respect to sample eigenvectors and marginal

distributions by integrating the joint distributions with respect to sample eigenvalues not under consideration. It involves a two-matrix arguments hypergeometric function with an infinite series expansion with respect to zonal polynomials. In a simple case of $\mathbf{\Sigma} = \sigma^2 \mathbf{I}_p$ (often called the null case) where $\mathbf{I}_p$ is an identity matrix of size p, the joint distribution consists of elementary functions, but exact marginal distributions are available only for the largest and the smallest sample eigenvalues.

Due to the complexity of exact distributions, asymptotic distributions have been developed [e.g., *Anderson*, 1963; *North et al.*, 1982; *Anderson*, 2003; *Muirhead*, 2005]. Let l_j and $\mathbf{a}_j$ be the j th sample eigenvalue and sample eigenvector, respectively, where $j = 1,2,\cdots,p$. The population counterparts are denoted by λ_j and $\boldsymbol{\alpha}_j$. Assuming that all population eigenvalues are distinct, *Jolliffe* [2002] summarized the asymptotic results as

$$l_j \sim N(\lambda_j, \frac{2\lambda_j^2}{n-1}) \tag{F-1}$$

$$\mathbf{a}_j \sim N(\boldsymbol{\alpha}_j, \frac{\lambda_j}{n-1} \sum_{\substack{k=1 \\ k \neq j}}^{p} \frac{\lambda_k}{(\lambda_k - \lambda_j)^2} \boldsymbol{\alpha}_k \boldsymbol{\alpha}_k^{\mathrm{T}}). \tag{F-2}$$

Furthermore, a sample eigenvalue is independent of the others and of all sample eigenvectors, and an element of a sample eigenvector is correlated with the other elements of the entire sample eigenvectors due to the orthogonality. The asymptotic results suggest that l_j and $\mathbf{a}_j$ are consistent and unbiased estimators. Statistical inferences based on the asymptotic distributions [e.g., *Jolliffe*, 2002; *Jackson*, 2003], however, are limited by the underlying assumptions: distinct population eigenvalues, time-invariant mean and noise, independent Gaussian noise, and large sample size.

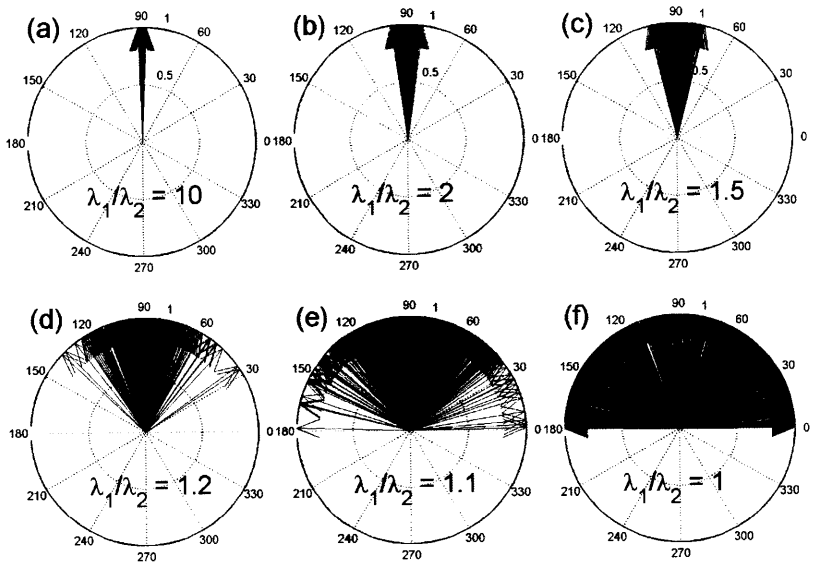


Figure F-1. Distributions of the sample eigenvectors obtained from two independent Gaussian variables with zero mean and variance λ_j (equivalently, the j th population eigenvalue). The distributions were generated from 1000 realizations. The number of samples is 1000 for each realization. $\lambda_2 = 1$ and λ_1 is given according to the values of λ_1/λ_2 given in each plot. Only one of the two sample eigenvectors is shown. As the eigenvalues approach being equal, the eigenvectors are completely uncertain.

Sample eigenvalues are biased when sample size is small. *von Storch and Hannoschock* [1985] attempted to correct the biases by using Monte Carlo simulations, jackknifing, and asymptotic results by *Lawley* [1956]. However, the corrections increase the uncertainty of sample eigenvalues. The Lawley correction possibly yields negative sample eigenvalues and is not appropriate for closely spaced sample eigenvalues.

The variance-covariance matrix in Equation (F-2) depends on the separation between population eigenvalues. Sample eigenvectors associated with widely separated population eigenvalues will have small uncertainties, while those with equal population eigenvalues (i.e., degenerate) will be completely uncertain (Figure F-1f) and have a uniform distribution of orthogonal rotations, called the conditional Haar invariant

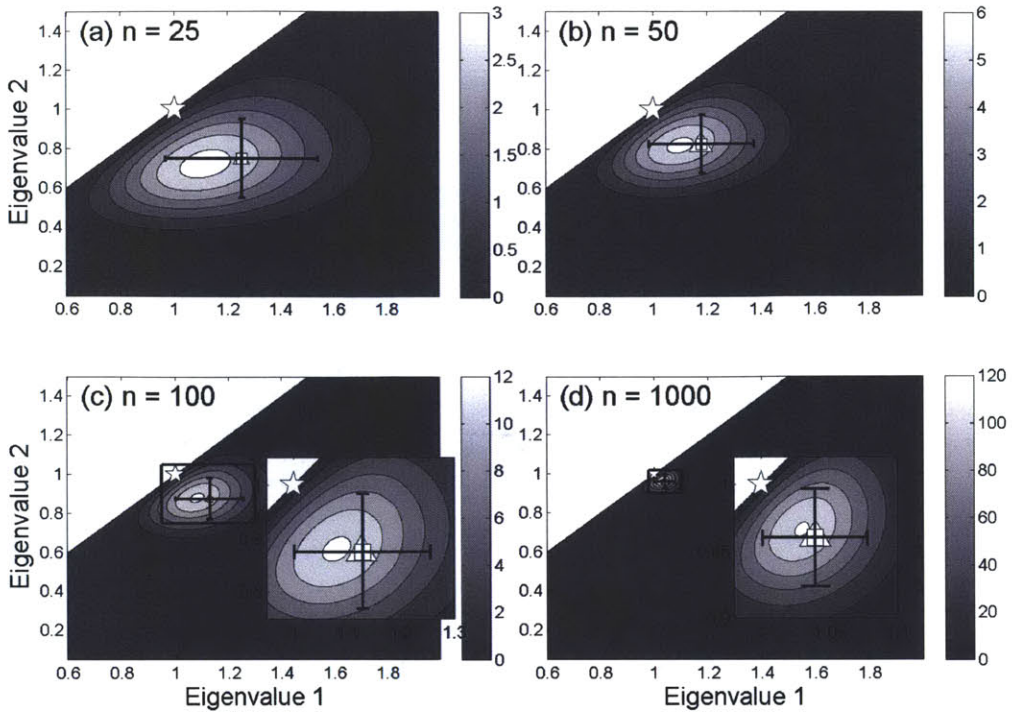


Figure F-2. Joint distributions of the sample eigenvalues from two independent Gaussian variables with zero mean and variance (or population eigenvalue) of one (stars). The number of samples is denoted by n. The numeric probability densities (contours) were obtained from the analytic expression given by *Muirhead* [2005]. Expectations of the sample eigenvalues (triangles) were calculated by the numerical integration of the joint distribution. Ensemble averages of the sample eigenvalues (squares) and standard deviations (horizontal and vertical error bars) were obtained from 1000 simulations. The upper–left white part is not defined due to the descending order of the sample eigenvalues. Zoom plots of the square boxes are shown in (c) and (d). The sample eigenvalues are biased and asymptotically approach the true value.

distribution (conditioned on fixed sign) [*Anderson*, 1963; 2003]. When the separation of population eigenvalues is comparable to or smaller than sampling variation, referred to as "effective degeneracy" [*North et al.*, 1982], sample eigenvectors tend to be unstable (see Figure F-1d~e). Effectively degenerate population eigenvalues are not distinguishable by the sample eigenvalues, and the corresponding sample eigenvectors are a mixture of the population eigenvectors [*North et al.*, 1982; *Quadrelli et al.*, 2005; *Hannachi et al.*, 2007].

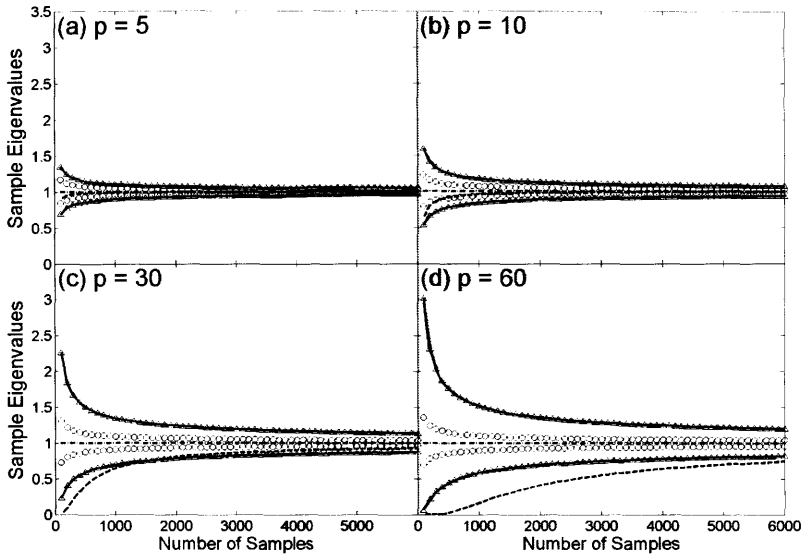


Figure F-3. Plots of the largest (triangles above one) and the smallest sample eigenvalues (triangles below one) from p number of independent Gaussian variables with zero mean and variance (or population eigenvalue) of one. Also shown are sample variances of the maximum (circles above one) and the minimum values (circles below one), the total sample variances divided by p (dash-dotted lines), and the generalized sample variances (dashed curves). The total sample variances are close to one for all the sample size, while the generalized sample variances (dashed curves) are significantly varying, particularly when p is large and the sample size is small. Least-squares fits (solid curves) to the sample eigenvalues are based on the model given in Equation (F-3), and the parameter estimates for selected values of p can be found in Table F-1. All the quantities are the ensemble averages from m number of random simulations. For computational efficiency, the values of m were chosen to be $n \times m = 10^7$ where n is the number of samples.

In the degenerate case, larger sample eigenvalues are positively biased and smaller ones are negatively biased [*von Storch and Hannoschock*, 1985], sometimes called the artificial variance compression [*Allen and Smith*, 1996]. The bias is due to the ranking of sample eigenvalues and sampling variations in $\mathbf{S}$. In the two-dimensional case shown in Figure F-2, sample eigenvalues must be placed in a region where the first sample eigenvalue is greater than the second. We measured sampling variations in $\mathbf{S}$ by the generalized sample variance (i.e., determinant of $\mathbf{S}$), the total sample variance (i.e., trace of $\mathbf{S}$), and the maximum and minimum sample variances, among which only the

generalized sample variance accounts for sample covariances (i.e., off diagonals of $\mathbf{S}$). The generalized sample variance, on average, differs from its true value more than the other quantities do (Figure F-3), which indicates that the biases of sample eigenvalues are mostly due to sampling variation in sample covariances. The bias becomes smaller with larger sample size and smaller number of variables (Figure F-3), which is expected from large sample statistics and less number of sample covariances, respectively.

When some of population eigenvalues are equal, they can be grouped with equal values. Each group is asymptotically independent of the other groups [*Anderson*, 1963; 2003]. If a group has only one member, Equations (F-1) and (F-2) are asymptotically valid. If a group has multiple members, the sample eigenvalues are biased and the sample eigenvectors have the conditional Haar invariant distribution. In general, however, there is no clear indication of which population eigenvalues are equal or effectively degenerate.

Here we provide an empirical model for the sample eigenvalues associated with equal population eigenvalues (i.e., $\lambda_j = \sigma^2$ for all j). How much sample eigenvalues are biased depends on σ^2, n, p, and the ranking of sample eigenvalues. The model is applied only to the largest and the smallest sample eigenvalues (Figure F-3). Let $\bar{l}_j$ be the ensemble average of the j th sample eigenvalues. The model function is given by

$$\bar{l}_j(n, \beta_{1j}, \beta_{2j} \mid \sigma^2, p) = \sigma^2 (e^{\beta_{1j}\sqrt{p/n}} + \beta_{2j}\sqrt{p/n}) \tag{F-3}$$

for $j = 1$ or p. The form of $f(x \mid y)$ denotes a function of x given y. The estimates of the parameters β_1 and β_2 are shown in Table F-1 for given p. Linear interpolation can be used to approximate parameters for p not in the table and $\bar{l}_j$ for $j = 2, \cdots, p-1$. For large n, $\bar{l}_j \approx \sigma^2 \left(1 + (\beta_{1j} + \beta_{2j})\sqrt{p/n}\right)$ so that the two parameters cannot be well separated.

The model can be used to account for the sample eigenvalue bias with computational efficiency avoiding a number of simulations.

When $n \leq p$, the rank of **S** is no longer full and the distribution of **S** does not exist. Since the rank of **S** is $n-1$, there are $n-1$ nonzero sample eigenvalues. Based on the fact that the total sample variance is equal to the sum of sample eigenvalues and approximately equal to the sum of population eigenvalues (see Figure F-3), we have $\sum_{j=1}^{p} \bar{l}_j = \sum_{j=1}^{n-1} \bar{l}_j \approx p\sigma^2$. As n decreases with p fixed or as p increases with n fixed, the values of nonzero $\bar{l}_j$ increase to satisfy the above equalities (Figure F-4). We also modeled the sequence of the largest sample eigenvalue using Equation (F-3) and tabulated parameter estimates in Table F-1 for given p.

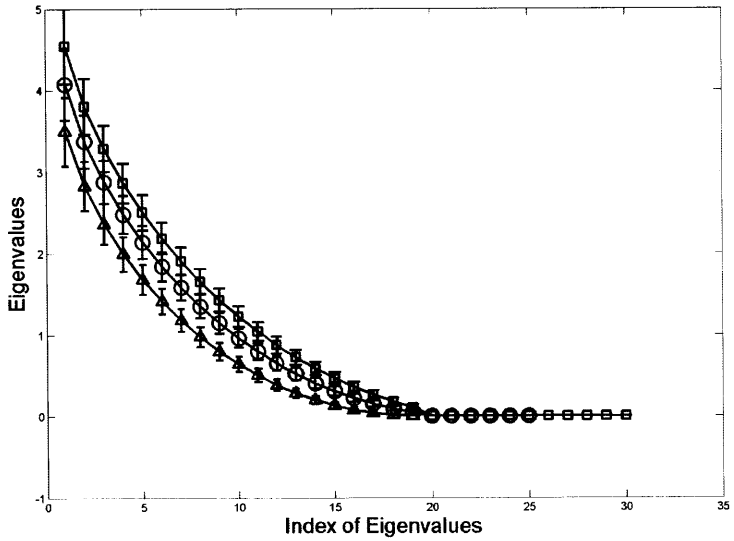


Figure F-4. Ensemble averages of sample eigenvalues when the numbers of variables are $p = 19$ (triangles), $p = 25$ (circles) and $p = 30$ (squares). The number of samples is $n = 20$ so that the case of $p = 19$ has a full rank, while the other cases are rank deficient; the sample eigenvalues with indices larger than 19 become zero. The population eigenvalue is $\lambda_j = 1$ for $j = 1,2,\cdots,p$. Error bars approximately represent the 1-sigma range of the scatter of synthetic sample eigenvalues from 50,000 realizations.

Table F-1. Parameter estimates for the largest and the smallest sample eigenvalue sequences[a]

p	$n > p$				$n \leq p$	
	$\hat{\beta}_{11}$	$\hat{\beta}_{21}$	$\hat{\beta}_{1p}$	$\hat{\beta}_{2p}$	$\hat{\beta}_{11}$	$\hat{\beta}_{21}$
4	0.775(6)[b]	0.583(8)	0.705(6)	-2.062(8)	-	-
6	0.895(3)	0.640(4)	0.867(3)	-2.392(4)	-	-
8	0.940(2)	0.694(3)	0.937(2)	-2.549(3)	-	-
10	0.994(3)	0.698(4)	0.992(3)	-2.662(3)	-	-
12	1.011(3)	0.727(4)	1.015(2)	-2.721(3)	-	-
14	1.024(3)	0.747(4)	1.030(2)	-2.762(3)	-	-
16	1.032(3)	0.767(4)	1.038(2)	-2.789(4)	-	-
18	1.035(3)	0.787(4)	1.044(2)	-2.809(4)	-	-
20	1.065(3)	0.767(4)	1.070(2)	-2.855(4)	0.919(7)	1.09(2)
22	1.065(3)	0.784(4)	1.073(3)	-2.867(4)	0.916(7)	1.12(2)
24	1.065(3)	0.798(5)	1.074(3)	-2.876(4)	0.909(7)	1.17(2)
26	1.065(3)	0.811(5)	1.075(3)	-2.882(4)	0.908(5)	1.19(2)
28	1.063(3)	0.824(5)	1.075(3)	-2.888(4)	0.903(6)	1.22(2)
30	1.084(3)	0.807(5)	1.093(3)	-2.917(4)	0.900(6)	1.25(2)
35	1.080(3)	0.832(5)	1.091(3)	-2.924(5)	0.889(6)	1.31(2)
40	1.095(3)	0.828(5)	1.104(3)	-2.948(4)	0.878(7)	1.37(3)
45	1.090(3)	0.848(5)	1.100(3)	-2.949(5)	0.868(7)	1.42(3)
50	1.102(3)	0.843(5)	1.110(3)	-2.966(5)	0.859(7)	1.47(3)
55	1.096(3)	0.861(5)	1.106(3)	-2.964(5)	0.849(7)	1.52(3)
60	1.104(3)	0.858(5)	1.114(3)	-2.978(5)	0.841(7)	1.56(3)
70	1.107(3)	0.867(5)	1.115(3)	-2.985(5)	0.825(7)	1.63(3)
80	1.110(3)	0.875(5)	1.117(3)	-2.991(5)	0.811(7)	1.70(4)
90	1.109(3)	0.884(5)	1.118(3)	-2.995(5)	0.798(7)	1.77(4)
100	1.110(3)	0.890(6)	1.118(3)	-2.997(5)	0.785(7)	1.83(4)
120	1.112(3)	0.901(5)	1.119(3)	-3.000(5)	0.763(7)	1.93(4)
140	1.111(3)	0.910(6)	1.119(3)	-3.001(5)	0.744(6)	2.03(4)
160	1.112(3)	0.918(6)	1.118(3)	-3.001(5)	0.727(6)	2.12(5)
180	1.111(3)	0.925(6)	1.117(3)	-3.001(5)	0.711(6)	2.20(5)
200	1.110(3)	0.932(5)	1.117(3)	-3.000(5)	0.697(6)	2.27(5)

[a] The model is given by Equation (F-3) and parameters were estimated with $\sigma^2 = 1$.
[b] $0.775(6) = 0.775 \pm 0.06$. The errors are from t-test with 0.05 level.

Appendix G

Issues related to Monte Carlo simulation for PCA statistics

The Monte Carlo simulation is based on repeated random sampling to obtain approximate results when deterministic approaches to exact solutions are not possible. The method has been frequently used in crustal deformation studies for PCA uncertainty estimation [e.g., *Savage*, 1988; 1995a; *Smith et al.*, 2007]. Here we briefly discuss the method and issues related to the statistics of principal component analysis (PCA).

Suppose that samples consist of time-varying mean μ_{ih} and noise ε_{ih} from a realizable stochastic process for $i = 1,2,\cdots,n$ and $h = 1,2,\cdots,p$ where n and p are the number of samples and variables, respectively. After applying PCA to each of a number of synthetic data sets, PCA uncertainties can be obtained from the scatter of PCA solutions. For example, the uncertainties of principal components (PCs) are given by

$$\tilde{e}_{ij}^2 = \frac{1}{m}\sum_{k=1}^{m}(z_{ijk} - \zeta_{ij})^2 \qquad \text{(G-1)}$$

where z_{ijk} is the value of the j th PC at the i th epoch from the k th realization, ζ_{ij} is the population PC associated with z_{ij}, and m is the total number of realizations. The uncertainties of sample eigenvectors and sample eigenvalues can be similarly obtained.

Several aspects weaken the method. Firstly, μ_{ih} and ζ_{ij} are generally unknown. One can replace μ_{ih} with samples and ζ_{ij} with ensemble averages of PCs, but the results are possibly biased. Secondly, the sign of a PC is arbitrary and possibly switching during repeated simulations [*Jolliffe*, 2002]. The PC uncertainties from Equation (G-1) will have systematic temporal patterns without correction of sign switching. The sign can be fixed at an epoch t_i when z_{ij} is far away from zero not affected by sampling variation. Thirdly, the method is time-consuming for large data size when reliable results are needed.

When population eigenvalues are close to each other with respect to random noise, termed effective degeneracy (Appendix F), the sample eigenvectors and sample PCs are, respectively, mixtures of the effectively degenerate population eigenvectors and population PCs, rather than one-to-one correspondence. In this case, the PC uncertainties from Equation (G-1) tend to have systematic temporal patterns (Figure G-1) even when noise characteristics are invariant in time, and sample eigenvectors become unstable (Figure G-2). The effective degeneracy may be indicated by sample eigenvalues close to each other, but it cannot be clearly determined without accurate knowledge of noise properties. Temporal patterns in PC uncertainties may warn us about the effective degeneracy as well as the sign switching and inaccurate values of μ_{ih} and ζ_{ij}.

The Monte Carlo method has also been used to evaluate the significance of PCs relative to assumed background noise [e.g. *Preisendorfer*, 1988; *Savage*, 1988; 1995a;

Allen and Smith, 1996; *Smith et al.*, 2007]. For example, Rule N by *Preisendorfer* [1988] is based on a null hypothesis that samples come from an independent Gaussian population with zero mean and variance σ_{ih}^2. The null hypothesis can be tested based on the scatter of a number of ensemble eigenvalues from the independent Gaussian population. If sample eigenvalues are above an upper limit associated with a given significance level, the test rejects the null hypothesis. This method can be extended to realizable stochastic noise processes [e.g., *Allen and Smith*, 1996]. For higher-ranked PCs, however, ensemble eigenvalues tend to be biased positively and may increase the probability of incorrectly accepting the null hypothesis. Furthermore, if the noise variances are spread widely, the largest ensemble eigenvalue does not necessarily represent the noise level of the first PC from actual data.

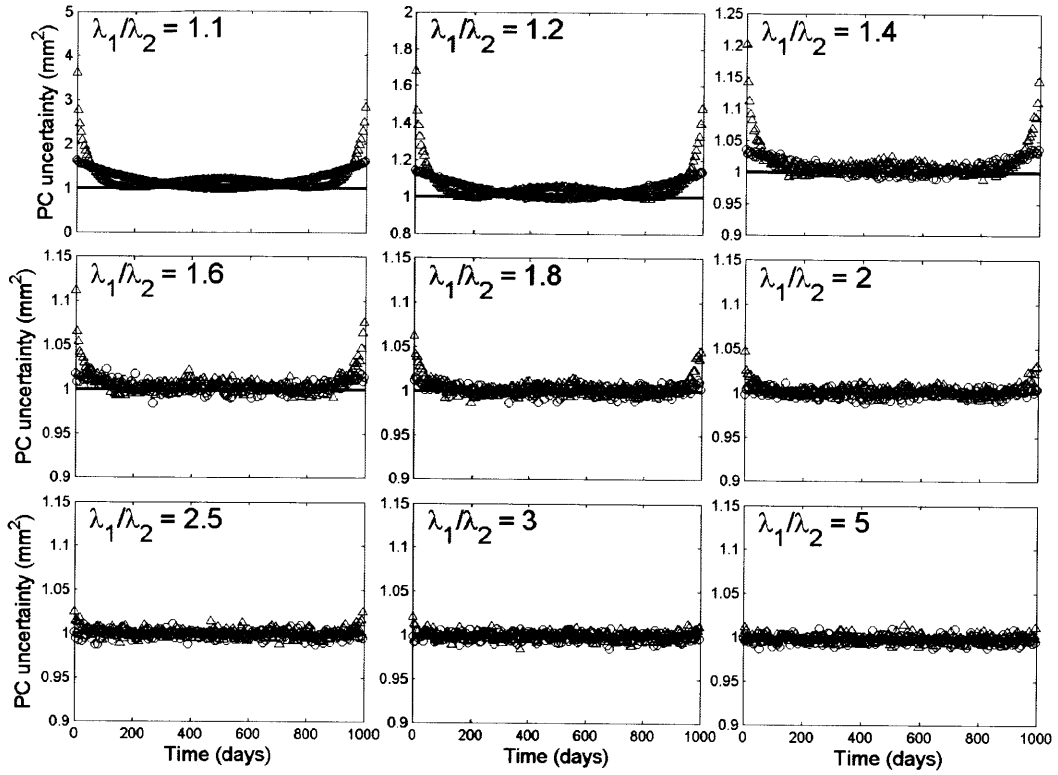


Figure G-1. Uncertainties of principal components (PCs) obtained from the Monte Carlo simulations with 100,000 realizations. Two Gaussian variables have 1000 samples with time-varying mean (the associated Legendre functions of degree one with signal power α_j^2 where $j = 1, 2$) and variance of 1 mm^2 for all time (horizontal lines). Since they are orthogonal, the Legendre functions are also population PCs with population eigenvalues $\lambda_j = \alpha_j^2 + 1$ mm^2. In each case, $\lambda_1 = 10$ mm^2 (i.e., $\alpha_1^2 = 9$ mm^2) and λ_2 were chosen according to the values of λ_1/λ_2 given in each plot. As λ_1/λ_2 gets close to one, the uncertainties of the first PC (triangles) and the second PC (circles) tend to have time-dependent systematic patterns in spite of identical noise characteristics in time. When $\lambda_1/\lambda_2 = 1.1$, the space between signal powers is only about 1 mm^2 which is comparable to the noise level. As λ_1/λ_2 increases, the temporal patterns in PC uncertainties become less pronounced and are close to 1 mm^2. Uncertainties are drawn at every five days. Sign switching was corrected by fitting population PCs to sample PCs in a least squares sense.

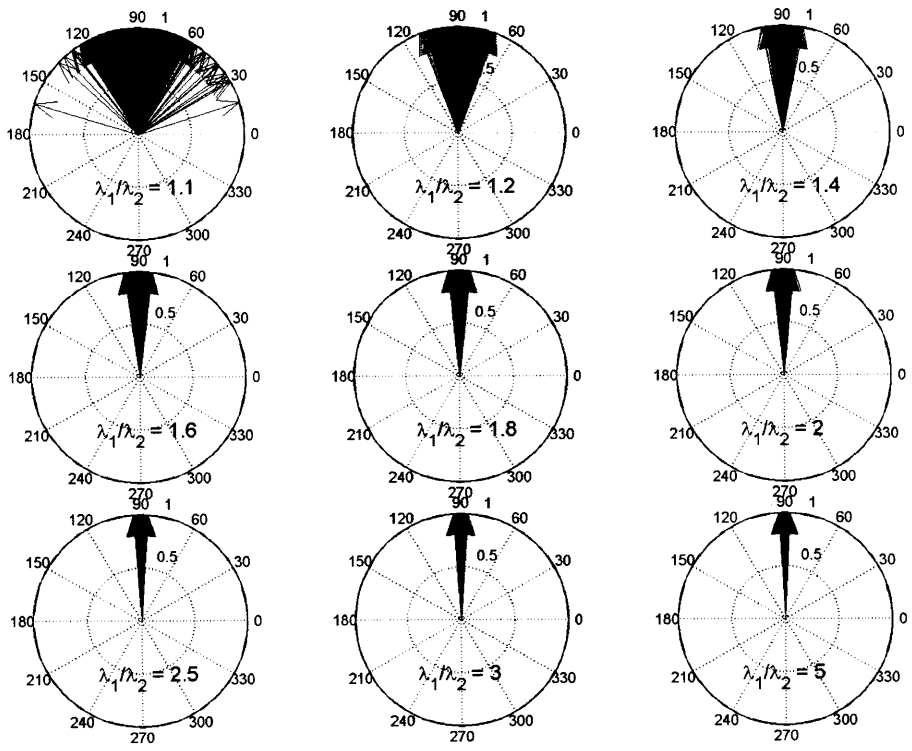


Figure G-2. Distributions of sample eigenvectors from the samples used in Figure G-1. In each case, the sample eigenvector of the first PC is shown. As λ_1/λ_2 is close to one, the sample eigenvectors become unstable.

Appendix H

Formal error scaling

White noise in GPS time series is assumed to be represented by formal errors calculated from noise in the phase measurements sampled at a given rate. However, temporal correlation in the phase data raises the scale issue in formal errors, depending on statistics and sampling rate used in the GPS geodetic inversion. Here we obtain scale factors by comparing formal errors with one-day differences of time series. The resulting scaled formal errors become more consistent with short-term data scatter.

Suppose that $y_k = \mu_k + \varepsilon_k$ at time t_k where y_k is a measurement, μ_k represents data trends (including time-dependent noise), and ε_k is an independent Gaussian noise with variance σ_k^2 given by the formal error of y_k. Assuming $\mu_k \approx \mu_{k-1}$ that would be reasonable for such a short time interval as one day (no missing data assumed), we obtain one-day differences by $dy_k = y_k - y_{k-1} \approx \varepsilon_k - \varepsilon_{k-1}$ and their errors by $d\sigma_k^2 = \sigma_k^2 + \sigma_{k-1}^2$. The sequence of one-day differences has no correlation in time except for one-day time lag having negative correlation of $-\sigma_k^2$ between dy_{k+1} and dy_k. Assuming that dy_k is a white Gaussian sequence for all k, we define a weighted-mean-square error as

$$d\sigma_{wrms}^2 = \sum_{k=2}^{n}\left(\frac{dy_k - \overline{dy}}{d\sigma_k}\right)^2 \Bigg/ \sum_{k=2}^{n}\left(\frac{1}{d\sigma_k}\right)^2 \qquad \text{(H-1)}$$

where $\overline{dy}$ is the sample average of dy_k and n is the number of samples. Under the null hypothesis that dy_k is independent Gaussian with mean $\overline{dy}$ and variance $d\sigma_k^2$, the numerator in Equation (H-1) has the chi square distribution with $n-2$ degrees of freedom where the reduction by two is due to differencing and mean estimated. Note that $\left\langle \chi^2_{n-2} \right\rangle = n-2$ where $\langle x \rangle$ denotes the expectation of x. If we rewrite Equation (H-1) as

$$d\sigma^2_{wrms} = \left(\frac{1}{n-2} \sum_{k=2}^{n} \left(\frac{dy_k - \overline{dy}}{d\sigma_k} \right)^2 \right) \times \left(\frac{n-2}{\sum_{k=2}^{n} (1/d\sigma_k)^2} \right) \equiv \alpha^2 \times d\overline{\sigma}^2 \qquad \text{(H-2)}$$

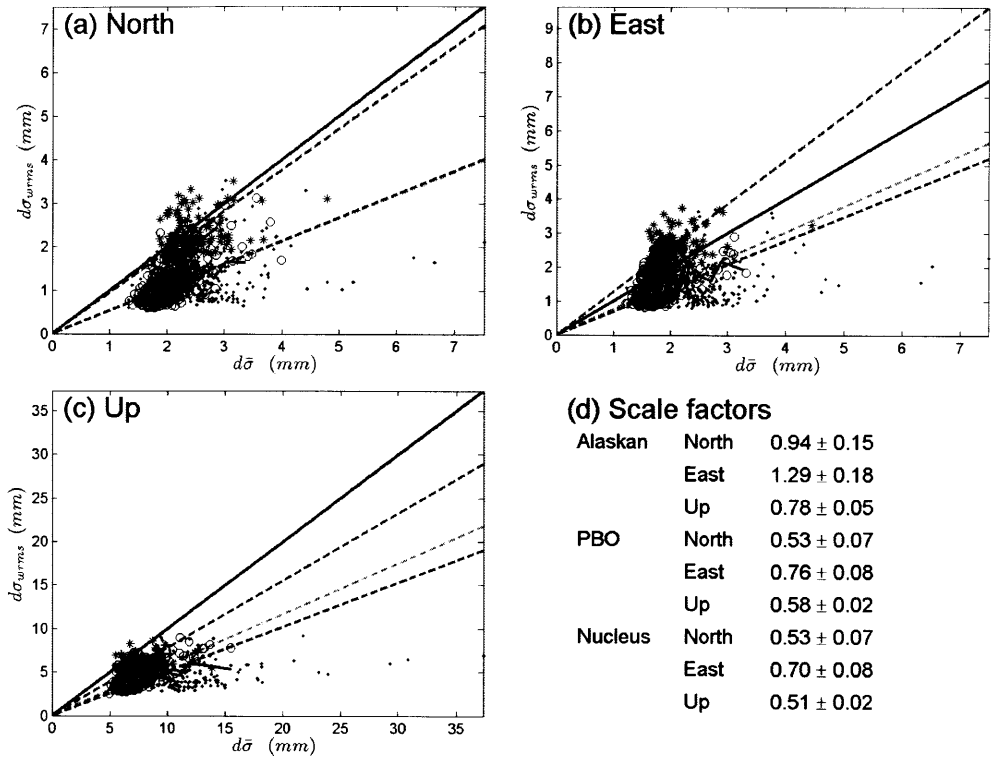
where α^2 and $d\overline{\sigma}^2$ are accordingly defined, we directly have $\langle \alpha \rangle = 1$ under the null hypothesis. This means that the formal errors are consistent with short-term data scatter represented by one-day differences. If the null hypothesis is not true, the expectation of α will be different from one. For consistency, formal errors are scaled by α uniformly in time (i.e., $\hat{\sigma}_k^2 = (\alpha\sigma_k)^2$ for all k).

We calculated the scale factors for the stations in the Plate Boundary Observatory (PBO) network with maximum observation interval from 1 January 2004 to 13 February 2010. The north, east and vertical components were treated separately. For reliable estimation, we assume that a set of GPS sites has experienced the same level of scaling. We divided total 1315 PBO sites into three groups: 151 Alaskan sites (those with latitude from +50° to +70° and longitude from -180° to -130°), 755 PBO sites (those with station IDs that begin with "P" followed by three digits), and 409 nucleus sites (the remaining stations used) that are parts of already existing GPS networks. The Alaskan sites may reflect error characteristics at high latitude and the other two groups may show

differences between newly installed PBO sites and old sites. We removed stations (~17% of Alaskan sites, ~8% of PBO sites and ~4% of nucleus sites) which sample size is less than 100 and which values of $d\sigma_{wrms}$ is larger than 4 mm in horizontal and 10 mm in vertical. After stacking one-day differences for each group and each coordinate, we calculated the values of $d\sigma_{wrms}$ and $d\bar{\sigma}$ per 10,000 differences. The scale factors were obtained by straight-line fit to the scatter of such values (Figure H-1).

The scale factors for Alaskan sites are closer to one, which means that the formal errors represent short-term data scatter relatively well and the sampling rate used for phase data would be appropriate for the white noise assumption. On the other hand, the scale factors for the other two groups are smaller than one, about a half particularly in north and vertical, indicating that the formal errors are too large. The larger scale factors for Alaskan sites may reflect latitude-dependent noise characteristics. The similar scale factors between the PBO and nucleus sites show no significant difference in formal errors between newly installed PBO stations and pre-existing stations. The scale factors for east component are larger than that for north component, which may imply incomplete integer-cycle phase ambiguity resolution in the GPS data analysis.

While temporal uniformity of the scale factors would be reasonable because formal errors are often of similar size in time, spatial uniformity would not be appropriate for the entire PBO network; scale factors of the Alaskan sites are larger than those of the PBO and nucleus sites. If we further divided PBO and nucleus sites into several groups with respect to geographic areas, the scale factors might have been different from each other. The spatial extent of the uniform scaling and the sources of the non-uniformity appeared in Alaskan sites are interesting topics in the error analysis of GPS data.



(d) Scale factors

Alaskan	North	0.94 ± 0.15
	East	1.29 ± 0.18
	Up	0.78 ± 0.05
PBO	North	0.53 ± 0.07
	East	0.76 ± 0.08
	Up	0.58 ± 0.02
Nucleus	North	0.53 ± 0.07
	East	0.70 ± 0.08
	Up	0.51 ± 0.02

Figure H-1. Estimation of scale factors from the scatter of $d\sigma_{wrms}$ in terms of $d\bar{\sigma}$ defined in (F-2). Each subplot shows the values of $d\sigma_{wrms}$ for Alaskan stations (asterisks), PBO stations (circles), and nucleus stations (dots). Also shown are the values of $d\sigma_{wrms}$ per 10,000 stacked one-day differences and the straight-line fits for the Alaskan (blue), the PBO (green), and the nucleus stations (red). The estimates of the scale factor α from straight-line fits are shown in (d). The black solid line in each subplot is the line for $\alpha = 1$.

Bibliography

Agnew, D. C. (1992), The Time-Domain Behavior of Power-Law Noises, *Geophys. Res. Lett.*, *19(4)*, 333-336, doi:10.1029/91GL02832.

Agnew, D. C. (2009), *FAKENET User Manual*, Scripps Institution of Oceanography, La Jolla, California.

Aguiar, A. C., T. I. Melbourne, and C. W. Scrivner (2009), Moment release rate of Cascadia tremor constrained by GPS, *J. Geophys. Res.*, *114*, B00A05, doi:10.1029/2008JB005909.

Allen, M. R., and L. A. Smith (1994), Investigating the origins and significance of low-frequency modes of climate variability, *Geophys. Res. Lett.*, *21(10)*, 883-886, doi:10.1029/94GL00978.

Allen, M. R., and A. W. Robertson (1996), Distinguishing modulated oscillations from coloured noise in multivariate datasets, *Clim Dynam*, *12(11)*, 775-784, doi:10.1007/s003820050142.

Allen, M. R., and L. A. Smith (1996), Monte Carlo SSA: Detecting irregular oscillations in the presence of colored noise, *J. Climate*, *9*, 12, 3373-3404.

Altamimi, Z., X. Collilieux, J. Legrand, B. Garayt, and C. Boucher (2007), ITRF2005: A new release of the International Terrestrial Reference Frame based on time series of station positions and earth orientation parameters, *J. Geophys. Res.*, *112*, B09401, doi:10.1029/2007JB004949.

Anderson, B. D. O., and J. B. Moore (1979), *Optimal filtering*, 357 pp., Prentice-Hall, Englewood Cliffs, N.J.

Anderson, T. W. (1963), Asymptotic theory for principal component analysis, *Ann. Math. Stat.*, *34*, 122-148.

Anderson, T. W. (2003), *An introduction to multivariate statistical analysis*, 3rd ed., 721 pp., Wiley-Interscience, Hoboken, N.J.

Aoki, Y., and C. H. Scholz (2003), Vertical deformation of the Japanese islands, 1996-1999, *J. Geophys. Res.*, *108*, B5, 2257, doi:10.1029/2002JB002129.

Argus, D. F., M. B. Heflin, G. Peltzer, F. Crampe, and F. H. Webb (2005), Interseismic strain accumulation and anthropogenic motion in metropolitan Los Angeles, *J. Geophys. Res.*, *110*, B04401, doi:10.1029/2003JB002934.

Bar-Sever, Y. E. (1996), A new model for GPS yaw attitude, *J. Geod.*, *70(11)*, 714-723, doi:10.1007/BF00867149.

Barnes, R. T. H., R. Hide, A. A. White, and C. A. Wilson (1983), Atmospheric Angular-Momentum Fluctuations, Length-of-Day Changes and Polar Motion, *Proc. R. Soc. Lond.*, *387(1792)*, 31-73, doi:10.1098/rspa.1983.0050.

Bassiri, S., and G. A. Hajj (1993), Higher-order ionospheric effects on the global positioning system observables and means of modeling them, *Manuscr. Geod.*, *18*, 280-289.

Bawden, G. W., W. Thatcher, R. S. Stein, K. W. Hudnut, and G. Peltzer (2001), Tectonic contraction across Los Angeles after removal of groundwater pumping effects, *Nature*, *412*, 6849, 812-815, doi:10.1038/35090558.

Beavan, J. (2005), Noise properties of continuous GPS data from concrete pillar geodetic monuments in New Zealand and comparison with data from U.S. deep drilled braced monuments, *J. Geophys. Res.*, *110*, B08410, doi:10.1029/2005JB003642.

Bennett, R. A. (2008), Instantaneous deformation from continuous GPS: contributions from quasi-periodic loads, *Geophys. J. Int.*, *174(3)*, 1052-1064, doi:10.1111/j.1365-246X.2008.03846.x.

Beutler, G., I. I. Mueller, and R. E. Neilan (1994a), The International GPS Service for Geodynamics: Development and start of official service on January 1, 1994, *Bull. Geodesique*, *68*, 39-70.

Beutler, G., E. Brockmann, W. Gurtner, U. Hugentobler, L. Mervart, and M. Rothacher (1994b), Extended orbit modeling techniques at the CODE processing center of the International GPS Service for geodynamics (IGS): theory and initial results, *Manuscr. Geod.*, *19*, 367-386.

Beutler, G., et al. (2007), *Bernese GPS Software Version 5.0*, Astronomical Institute, University of Bern, Bern, Switzerland.

Biggs, J., Z. Lu, T. Fournier, and J. Freymueller (2010), Magma flux at Okmok Volcano, Alaska, from a joint inversion of continuous GPS, campaign GPS, and interferometric synthetic aperture radar, *J. Geophys. Res.*, *115*, B12401, doi:10.1029/2010JB007577.

Blanchet, I., C. Frankignoul, and M. A. Cane (1997), A Comparison of Adaptive Kalman Filters for a Tropical Pacific Ocean Model, *Mon. Wea. Rev.*, *125*, 1, 40-58

Blewitt, G. (1989), Carrier Phase Ambiguity Resolution for the Global Positioning System Applied to Geodetic Baselines up to 2000 km, *J. Geophys. Res.*, *94(B8)*, 10187-10203, doi:10.1029/JB094iB08p10187.

Blewitt, G. (1990), An automatic editing algorithm for GPS data, *Geophys. Res. Lett.*, *17(3)*, 199-202, doi:10.1029/GL017i003p00199.

Blewitt, G., and D. Lavallee (2002), Effect of annual signals on geodetic velocity, *J. Geophys. Res.*, *107(B7)*, 2145, doi:10.1029/2001JB000570.

Bock, Y., L. Prawirodirdjo, and T. I. Melbourne (2004), Detection of arbitrarily large dynamic ground motions with a dense high-rate GPS network, *Geophys. Res. Lett.*, *31*, L06604, doi:10.1029/2003GL019150.

Bock, Y., S. A. Gourevitch, C. C. Counselman III, R. W. King, and R. I. Abbot (1986), Interferometric analysis of GPS phase observation, *Manuscr. Geod.*, *11*, 282-288.

Bock, Y., et al. (1993), Detection of crustal deformation from the Landers earthquake sequence using continuous geodetic measurements, *Nature*, *361*, 337-340, doi:10.1038/361337a0.

Boehm, J., B. Werl, and H. Schuh (2006a), Troposphere mapping functions for GPS and very long baseline interferometry from European Centre for Medium-Range Weather Forecasts operational analysis data, *J. Geophys. Res.*, *111*, B02406, doi:10.1029/2005JB003629.

Boehm, J., A. Niell, P. Tregoning, and H. Schuh (2006b), Global Mapping Function (GMF): A new empirical mapping function based on numerical weather model data, *Geophys. Res. Lett.*, *33*, L07304, doi:10.1029/2005GL025546.

Brown, R. G., and P. Y. C. Hwang (1997), *Introduction to random signals and applied Kalman filtering : with MATLAB exercises and solutions*, 3rd ed., 484 pp., Wiley, New York.

Brudzinski, M. R., and R. M. Allen (2007), Segmentation in episodic tremor and slip all along Cascadia, *Geology*, *35*, 10, 907-910, doi:10.1130/G23740A.1.

Cervelli, P., P. Segall, K. Johnson, M. Lisowski, and A. Miklius (2002), Sudden aseismic fault slip on the south flank of Kilauea volcano, *Nature*, *415*, 1014-1018, doi:10.1038/4151014a.

Chang, W. L., R. B. Smith, J. Farrell, and C. M. Puskas (2010), An extraordinary episode of Yellowstone caldera uplift, 2004-2010, from GPS and InSAR observations, *Geophys. Res. Lett.*, *37*, L23302, doi:10.1029/2010GL045451.

Chang, W. L., R. B. Smith, C. Wicks, J. M. Farrell, and C. M. Puskas (2007), Accelerated uplift and magmatic intrusion of the Yellowstone caldera, 2004 to 2006, *Science, 318(5852)*, 952-956, doi:10.1126/science.1146842.

Chapman, J. S., and T. I. Melbourne (2009), Future Cascadia megathrust rupture delineated by episodic tremor and slip, *Geophys. Res. Lett., 36*, L22301, doi:10.1029/2009GL040465.

Chen, G., and T. A. Herring (1997), Effects of atmospheric azimuthal asymmetry on the analysis of space geodetic data, *J. Geophys. Res., 102(B9)*, 20489-20502, doi:10.1029/97JB01739.

Cross, R. S., and J. T. Freymueller (2008), Evidence for and implications of a Bering plate based on geodetic measurements from the Aleutians and western Alaska, *J. Geophys. Res., 113*, B07405, doi:10.1029/2007JB005136.

Daley, R. (1992), The Lagged Innovation Covariance - A Performance Diagnostic for Atmospheric Data Assimilation, *Mon. Wea. Rev., 120*, 1, 178-196

DeMets, C., R. G. Gordon, and D. F. Argus (2010), Geologically current plate motions, *Geophys. J. Int., 181*, 1, 1-80, doi:10.1111/j.1365-246X.2009.04491.x.

Dixon, J. P., and S. D. Stihler (2009), Catalog of earthquake hypocenters at Alaskan volcanoes: January 1 through December 31, 2008, *U.S. Geological Survay Data Series 467*, 88 p.

Dixon, T. H., A. L. Mao, M. Bursik, M. Heflin, J. Langbein, R. Stein, and F. Webb (1997), Continuous monitoring of surface deformation at Long Valley Caldera, California, with GPS, *J. Geophys. Res., 102(B6)*, 12017-12034, doi:10.1029/96JB03902.

Dong, D., P. Fang, Y. Bock, M. K. Cheng, and S. Miyazaki (2002), Anatomy of apparent seasonal variations from GPS-derived site position time series, *J. Geophys. Res., 107(B4)*, 2075, doi:10.1029/2001JB000573.

Dong, D., P. Fang, Y. Bock, F. Webb, L. Prawirodirdjo, S. Kedar, and P. Jamason (2006), Spatiotemporal filtering using principal component analysis and Karhunen-Loeve expansion approaches for regional GPS network analysis, *J. Geophys. Res., 111*, B03405, doi:10.1029/2005JB003806.

Dong, D. N., and Y. Bock (1989), Global Positioning System Network Analysis with Phase Ambiguity Resolution Applied to Crustal Deformation Studies in California, *J. Geophys. Res., 94(B4)*, 3949-3966, doi:10.1029/JB094iB04p03949.

Dragert, H., K. L. Wang, and T. S. James (2001), A silent slip event on the deeper Cascadia subduction interface, *Science, 292(5521)*, 1525-1528, doi:10.1126/science.1060152.

Elosequi, P., J. L. Davis, R. T. K. Jaldehag, J. M. Johansson, A. E. Niell, and I. I. Shapiro (1995), Geodesy using the Global Positioning System: The effects of signal scattering on estimates of site position, *J. Geophys. Res.*, *100(B6)*, 9921-9934, doi:10.1029/95JB00868.

Feng, L. J., and A. V. Newman (2009), Constraints on continued episodic inflation at Long Valley Caldera, based on seismic and geodetic observations, *J. Geophys. Res.*, *114*, B06403, doi:10.1029/2008JB006240.

Fliegel, H. F., T. E. Gallini, and E. R. Swift (1992), Global Positioning System Radiation Force Model for Geodetic Applications, *J. Geophys. Res.*, *97(B1)*, 559-568, doi:10.1029/91JB02564.

Fournier, T., J. Freymueller, and P. Cervelli (2009), Tracking magma volume recovery at Okmok volcano using GPS and an unscented Kalman filter, *J. Geophys. Res.*, *114*, B02405, doi:10.1029/2008JB005837.

Fraser, D. C., and J. E. Potter (1969), The Optimun Linear Smoother as a Combination of Two Optimum Linear Filters, *IEEE Trans. Automat. Contr.*, *Ac14*, 4, 387-390

Freed, A. M. (2007), Afterslip (and only afterslip) following the 2004 Parkfield, California, earthquake, *Geophys. Res. Lett.*, *34*, L06312, doi:10.1029/2006GL029155.

Freed, A. M., R. Burgmann, and T. Herring (2007), Far-reaching transient motions after Mojave earthquakes require broad mantle flow beneath a strong crust, *Geophys. Res. Lett.*, *34*, L19302, doi:10.1029/2007GL030959.

Freed, A. M., R. Burgmann, E. Calais, J. Freymueller, and S. Hreinsdottir (2006), Implications of deformation following the 2002 Denali, Alaska, earthquake for postseismic relaxation processes and lithospheric rheology, *J. Geophys. Res.*, *111*, B01401, doi:10.1029/2005JB003894.

Freymueller, J. T., and A. M. Kaufman (2010), Changes in the magma system during the 2008 eruption of Okmok volcano, Alaska, based on GPS measurements, *J. Geophys. Res.*, *115*, B12415, doi:10.1029/2010JB007716.

Fritsche, M., R. Dietrich, C. Knofel, A. Rulke, S. Vey, M. Rothacher, and P. Steigenberger (2005), Impact of higher-order ionospheric terms on GPS estimates, *Geophys. Res. Lett.*, *32*, L23311, doi:10.1029/2005GL024342.

Gelb, A. (1974), *Applied optimal estimation*, 374 pp., M.I.T. Press, Cambridge, Mass.

Gross, R. S., I. Fukumori, and D. Menemenlis (2003), Atmospheric and oceanic excitation of the Earth's wobbles during 1980-2000, *J. Geophys. Res.*, *108(B8)*, 2370, doi:10.1029/2002JB002143.

Hannachi, A., I. T. Jolliffe, and D. B. Stephenson (2007), Empirical orthogonal functions and related techniques in atmospheric science: A review, *Int. J. Climatol.*, *27(9)*, 1119-1152, doi:10.1002/joc.1499.

Heino, M., J. Ripa, and V. Kaitala (2000), Extinction risk under coloured environmental noise, *Ecography*, *23*, 2, 177-184

Hernandez-Pajares, M., J. M. Juan, J. Sanz, and R. Orus (2007), Second-order ionospheric term in GPS: Implementation and impact on geodetic estimates, *J. Geophys. Res.*, *112*, B08417, doi:10.1029/2006JB004707.

Herring, T. A. (1999), Geodetic applications of GPS, *Proc. IEEE*, *87(1)*, 92-110, doi:10.1109/5.736344.

Herring, T. A., and D. N. Dong (1994), Measurement of Diurnal and Semidiurnal Rotational Variations and Tidal Parameters of Earth, *J. Geophys. Res.*, *99(B9)*, 18051-18071, doi:10.1029/94JB00341.

Herring, T. A., J. L. Davis, and Shapiro, II (1990), Geodesy by Radio Interferometry - the Application of Kalman Filtering to the Analysis of Very Long Base-Line Interferometry Data, *J. Geophys. Res.*, *95*, 12561-12581., doi:10.1029/JB095iB08p12561.

Herring, T. A., R. W. King, and S. C. McClusky (2009), *Introduction to GAMIT/GLOBK Release 10.35*, Mass. Instit. Tech.

Hirose, H., and K. Obara (2005), Repeating short- and long-term slow slip events with deep tremor activity around the Bungo channel region, southwest Japan, *Earth Planets Space*, *57*, 961-972.

Hirose, H., K. Hirahara, F. Kimata, N. Fujii, and S. Miyazaki (1999), A slow thrust slip event following the two 1996 Hyuganada earthquakes beneath the Bungo Channel, southwest Japan, *Geophys. Res. Lett.*, *26(21)*, 3237-3240, doi:10.1029/1999GL010999.

Hofmann-Wellenhof, B., H. Lichtenegger, and J. Collins (2001), *Global Positioning System : Theory and Practice*, 5th, rev. ed., 382 pp., Springer-Verlag, New York.

Holtkamp, S., and M. R. Brudzinski (2010), Determination of slow slip episodes and strain accumulation along the Cascadia margin, *J. Geophys. Res.*, *115*, B00A17, doi:10.1029/2008JB006058.

Hoque, M. M., and N. Jakowski (2008), Estimate of higher order ionospheric errors in GNSS positioning, *Radio Sci*, *43*, RS5008, doi:10.1029/2007RS003817.

Hreinsdottir, S., J. T. Freymueller, H. J. Fletcher, C. F. Larsen, and R. Burgmann (2003), Coseismic slip distribution of the 2002 Mw7.9 Denali fault earthquake, Alaska,

determined from GPS measurements, *Geophys. Res. Lett., 30(13)*, 1670, doi:10.1029/2003GL017447.

Hudnut, K. W., N. E. King, J. E. Galetzka, K. F. Stark, J. A. Behr, A. Aspiotes, S. van Wyk, R. Moffitt, S. Dockter, and F. Wyatt (2002), Continuous GPS observations of postseismic deformation following the 16 October 1999 Hector Mine, California, earthquake (Mw 7.1), *Bull. Seism. Soc. Am., 92(4)*, 1403-1422, doi:10.1785/0120000912.

Hurst, K. J., D. F. Argus, A. Donnellan, M. B. Heflin, D. C. Jefferson, G. A. Lyzenga, J. W. Parker, M. Smith, F. H. Webb, and J. F. Zumberge (2000), The coseismic geodetic signature of the 1999 Hector Mine Earthquake, *Geophys. Res. Lett., 27*, 17, 2733-2736, doi:10.1029/2000GL011841.

Ide, S., G. C. Beroza, D. R. Shelly, and T. Uchide (2007), A scaling law for slow earthquakes, *Nature, 447*, 76-79, doi:10.1038/nature05780.

Jackson, J. E. (2003), *A user's guide to principal components*, 569 pp., Wiley, Hoboken, N.J.

Jazwinski, A. H. (1969), Adaptive Filtering, *Automatica, 5*, 4, 475-485

Ji, K. H., and T. A. Herring (2011), Transient signal detection using GPS measurements: Transient inflation at Akutan volcano, Alaska, during early 2008, *Geophys. Res. Lett., 38*, L06307, doi:10.1029/2011GL046904.

Johnson, H. O., and D. C. Agnew (1995), Monument Motion and Measurements of Crustal Velocities, *Geophys. Res. Lett., 22*, 2905-2908, doi:10.1029/95GL02661.

Jolliffe, I. T. (2002), *Principal component analysis*, 2nd ed., 487 pp., Springer, New York.

Kailath, T. (1968), An innovations approach to least-squares estimation - Part I: Linear filtering in additive white noise, *IEEE Trans. Automat. Contr., AC-13*, 646-655.

Kawamura, M., and K. Yamaoka (2006), Spatiotemporal characteristics of the displacement field revealed with principal component analysis and the mode-rotation technique, *Tectonophysics, 419(1-4)*, 55-73, doi:10.1016/j.tecto.2006.03.019.

Kim, J. (1997), Iterated Grid Search Algorithm on Unimodal Criteria, Ph. D. thesis, Department of Statistics, 116 pp, Virginia Polytechnic Institute and State University, Blacksburg, Virginia.

Kimura, T., K. Obara, H. Kimura, and H. Hirose (2011), Automated detection of slow slip events within the Nankai subduction zone, *Geophys. Res. Lett., 38*, L01311, doi:10.1029/2010GL045899.

King, N. E., et al. (2007), Space geodetic observation of expansion of the San Gabriel Valley, California, aquifer system, during heavy rainfall in winter 2004-2005, *J. Geophys. Res.*, *112*, B03409, doi:10.1029/2006JB004448.

Kositsky, A. P., and J.-P. Avouac (2010), Inverting geodetic time series with a principal component analysis-based inversion method, *J. Geophys. Res.*, *115*, B03401, doi:10.1029/2009JB006535.

Kreemer, C., W. E. Holt, and A. J. Haines (2003), An integrated global model of present-day plate motions and plate boundary deformation, *Geophys. J. Int.*, *154(1)*, 8-34, doi:10.1046/j.1365-246X.2003.01917.x.

Lanari, R., P. Lundgren, M. Manzo, and F. Casu (2004), Satellite radar interferometry time series analysis of surface deformation for Los Angeles, California, *Geophys. Res. Lett.*, *31*, L23613, doi:10.1029/2004GL021294.

Langbein, J. (2004), Noise in two-color electronic distance meter measurements revisited, *J. Geophys. Res.*, *109*, B04406, doi:10.1029/2003JB002819.

Langbein, J. (2008), Noise in GPS displacement measurements from Southern California and Southern Nevada, *J. Geophys. Res.*, *113*, B05405, doi:10.1029/2007JB005247.

Langbein, J., and H. Johnson (1997), Correlated errors in geodetic time series: Implications for time-dependent deformation, *J. Geophys. Res.*, *102*, 591-603, doi:10.1029/96JB02945.

Langbein, J., and Y. Bock (2004), High-rate real-time GPS network at Parkfield: Utility for detecting fault slip and seismic displacements, *Geophys. Res. Lett.*, *31*, L15S20, doi:10.1029/2003GL019408.

Larson, K. M., P. Bodin, and J. Gomberg (2003), Using 1-Hz GPS data to measure deformations caused by the Denali fault earthquake, *Science*, *300(5624)*, 1421-1424, doi:10.1126/science.1084531.

Larson, K. M., A. Bilich, and P. Axelrad (2007), Improving the precision of high-rate GPS, *J. Geophys. Res.*, *112*, B05422, doi:10.1029/2006JB004367.

Lawley, D. N. (1956), Tests of significance for the latent roots of covariance and correlation matrices, *Biometrika*, *43*, 128-136.

Leick, A. (2004), *GPS satellite surveying*, 3rd ed., 435 pp., John Wiley & sons, Inc., Hoboken, NJ.

Lindqwister, U. J., A. P. Freedman, and G. Blewitt (1992), Daily estimates of the Earths pole position with the global positioning system, *Geophys. Res. Lett.*, *19(9)*, 845-848, doi:10.1029/92GL00576.

Lu, Z., C. Wicks, J. A. Power, and D. Dzurisin (2000), Ground deformation associated with the March 1996 earthquake swarm at Akutan volcano, Alaska, revealed by satellite radar interferometry, *J. Geophys. Res.*, *105(B9)*, 21483-21495, doi:10.1029/2000JB900200.

Lu, Z., J. A. Power, V. S. McConnell, C. Wicks, and D. Dzurisin (2002a), Preeruptive inflation and surface interferometric coherence characteristics revealed by satellite radar interferometry at Makushin Volcano, Alaska: 1993-2000, *J. Geophys. Res.*, *107*, B11, doi:10.1029/2001JB000970.

Lu, Z., T. Masterlark, D. Dzurisin, R. Rykhus, and C. Wicks (2003), Magma supply dynamics at Westdahl volcano, Alaska, modeled from satellite radar interferometry, *J. Geophys. Res.*, *108*, B7, doi:10.1029/2002JB002311.

Lu, Z., C. Wicks, D. Dzurisin, J. A. Power, S. C. Moran, and W. Thatcher (2002b), Magmatic inflation at a dormant stratovolcano: 1996-1998 activity at Mount Peulik volcano, Alaska, revealed by satellite radar interferometry, *J. Geophys. Res.*, *107(B7)*, 2134, doi:10.1029/2001JB000471.

Mandelbrot, B. (1983), *The fractal geometry of nature*, 468 pp., W.H. Freeman, San Francisco.

Mandelbrot, B., and J. W. Van Ness (1968), Fractional Brownian motions, fractional noises, and applications, *Siam Rev*, *10*, 4, 422-439

Mao, A. L., C. G. A. Harrison, and T. H. Dixon (1999), Noise in GPS coordinate time series, *J. Geophys. Res.*, *104(B2)*, 2797-2816, doi:10.1029/1998JB900033.

Mathews, P. M., T. A. Herring, and B. A. Buffett (2002), Modeling of nutation and precession: New nutation series for nonrigid Earth and insights into the Earth's interior, *J. Geophys. Res.*, *107(B4)*, 2068, doi:10.1029/2001JB000390.

McGuire, J. J., and P. Segall (2003), Imaging of aseismic fault slip transients recorded by dense geodetic networks, *Geophys. J. Int.*, *155(3)*, 778-788, doi:10.1111/j.1365-246X.2003.02022.x.

Mehra, R. K. (1970), On the identification of variances and adaptive Kalman filtering, *IEEE Trans. Automat. Contr.*, *AC-15*, 175-184.

Miller, M. M., T. Melbourne, D. J. Johnson, and W. Q. Sumner (2002), Periodic slow earthquakes from the Cascadia subduction zone, *Science*, *295(5564)*, 2423, doi:10.1126/science.1071193.

Mogi, K. (1958), Relations between the eruptions of various volcanoes and the deformation of the ground surface around them, *Bull. Earth. Res. Inst. Univ. Tokyo*, *36*, 99-134.

Mohamed, A. H., and K. P. Schwarz (1999), Adaptive Kalman filtering for INS GPS, *J. Geod.*, *73*, 4, 193-203, doi:10.1007/s001900050236.

Moran, S. C., O. Kwoun, T. Masterlark, and Z. Lu (2006), On the absence of InSAR-detected volcano deformation spanning the 1995-1996 and 1999 eruptions of Shishaldin Volcano, Alaska, *J Volcanol Geoth Res*, *150*, 1-3, 119-131

Muirhead, R. J. (2005), *Aspects of multivariate statistical theory*, 2nd ed., 712 pp., Wiley, New York.

Murray-Moraleda, J., and R. Lohman (2010), Workshop targets development of geodetic transient detection methods, *Eos Trans. AGU*, *91(6)*, doi:10.1029/2010EO060008.

Myers, K. A., and B. D. Tapley (1976), Adaptive Sequential Estimation with Unknown Noise Statistics, *IEEE Trans. Automat. Contr.*, *21*, 4, 520-523

Niell, A. E. (1996), Global mapping functions for the atmosphere delay at radio wavelengths, *J. Geophys. Res.*, *101(B2)*, 3227-3246, doi:10.1029/95JB03048.

North, G. R., T. L. Bell, R. F. Cahalan, and F. J. Moeng (1982), Sampling Errors in the Estimation of Empirical Orthogonal Functions, *Mon. Wea. Rev.*, *110*, 7, 699-706.

Obara, K. (2002), Nonvolcanic deep tremor associated with subduction in southwest Japan, *Science*, *296*, 5573, 1679-1681, doi:10.1126/science.1070378.

Obara, K., and H. Hirose (2006), Non-volcanic deep low-frequency tremors accompanying slow slips in the southwest Japan subduction zone, *Tectonophysics*, *417*, 33-51, doi:10.1016/j.tecto.2005.04.013.

Obara, K., H. Hirose, F. Yamamizu, and K. Kasahara (2004), Episodic slow slip events accompanied by non-volcanic tremors in southwest Japan subduction zone, *Geophys. Res. Lett.*, *31*, L23602, doi:10.1029/2004GL020848.

Ohtani, R., J. J. McGuire, and P. Segall (2010), Network strain filter: A new tool for monitoring and detecting transient deformation signals in GPS arrays, *J. Geophys. Res.*, *115*, B12418, doi:10.1029/2010JB007442.

Okada, Y. (1985), Surface Deformation Due to Shear and Tensile Faults in a Half-Space, *Bull. Seism. Soc. Am.*, *75*, 4, 1135-1154

Ozawa, S., M. Murakami, M. Kaidzu, T. Tada, T. Sagiya, Y. Hatanaka, H. Yarai, and T. Nishimura (2002), Detection and monitoring of ongoing aseismic slip in the Tokai region, Central Japan, *Science, 298(5595)*, 1009-1012, doi:10.1126/science.1076780.

Park, K. D., R. S. Nerem, M. S. Schenewerk, and J. L. Davis (2004), Site-specific multipath characteristics of global IGS and CORS GPS sites, *J. Geod., 77(12)*, 799-803, doi:10.1007/s00190-003-0359-9.

Parkinson, B. W., and J. J. Spilker (1996), *The global positioning system : theory and applications*, American Institute of Aeronautics and Astronautics, Washington, DC.

Penna, N. T., and M. P. Stewart (2003), Aliased tidal signatures in continuous GPS height time series, *Geophys. Res. Lett., 30(23)*, 2184, doi:10.1029/2003GL018828.

Penna, N. T., M. A. King, and M. P. Stewart (2007), GPS height time series: Short-period origins of spurious long-period signals, *J. Geophys. Res., 112*, B02402, doi:10.1029/2005JB004047.

Percival, D. B., and A. T. Walden (1993), *Spectral analysis for physical applications : multitaper and conventional univariate techniques*, xxvii, 583 p. pp., Cambridge University Press, Cambridge ; New York, NY, USA.

Petrie, E. J., M. A. King, P. Moore, and D. A. Lavallee (2010), Higher-order ionospheric effects on the GPS reference frame and velocities, *J. Geophys. Res., 115*, B03417, doi:10.1029/2009JB006677.

Prawirodirdjo, L., and Y. Bock (2004), Instantaneous global plate motion model from 12 years of continuous GPS observations, *J. Geophys. Res., 109*, B08405, doi:10.1029/2003JB002944.

Prawirodirdjo, L., R. McCaffrey, C. D. Chadwell, Y. Bock, and C. Subarya (2010), Geodetic observations of an earthquake cycle at the Sumatra subduction zone: Role of interseismic strain segmentation, *J. Geophys. Res., 115*, B03414, doi:10.1029/2008JB006139.

Preisendorfer, R. W. (1988), *Principal component analysis in meteorology and oceanography*, 425 pp., Elsevier, New York.

Quadrelli, R., C. S. Bretherton, and J. M. Wallace (2005), On sampling errors in empirical orthogonal functions, *J. Climate, 18(17)*, 3704-3710, doi:10.1175/JCLI3500.1.

Reilinger, R. E., et al. (2000), Coseismic and postseismic fault slip for the 17 August 1999, M=7.5, Izmit, Turkey earthquake, *Science, 289(5484)*, 1519-1524, doi:10.1126/science.289.5484.1519.

Rogers, G., and H. Dragert (2003), Episodic tremor and slip on the Cascadia subduction zone: The chatter of silent slip, *Science*, *300*, 5627, 1942-1943, doi:10.1126/science.1084783.

Ruegg, J. C., A. Rudloff, C. Vigny, R. Madariaga, J. B. de Chabalier, J. Campos, E. Kausel, S. Barrientos, and D. Dimitrov (2009), Interseismic strain accumulation measured by GPS in the seismic gap between Constitucion and Concepcion in Chile, *Phys. Earth Planet. Inter.*, *175(1-2)*, 78-85, doi:10.1016/j.pepi.2008.02.015.

Sagiya, T. (2004), A decade of GEONET: 1994-2003 - The continuous GPS observation in Japan and its impact on earthquake studies, *Earth Planets Space*, *56*, 8, xxix-xli.

Savage, J. C. (1988), Principal Component Analysis of Geodetically Measured Deformation in Long Valley Caldera, Eastern California, 1983-1987, *J. Geophys. Res.*, *93(B11)*, 13297-13305, doi:10.1029/JB093iB11p13297.

Savage, J. C. (1995a), Principal Component Analysis of Interseismic Deformation in Southern California, *J. Geophys. Res.*, *100(B7)*, 12691-12701, doi:10.1029/95JB01155.

Savage, J. C. (1995b), Interseismic Uplift at the Nankai Subduction Zone, Southwest Japan, 1951-1990, *J. Geophys. Res.*, *100(B4)*, 6339-6350, doi:10.1029/95JB00242.

Savage, J. C., and W. Thatcher (1992), Interseismic Deformation at the Nankai Trough, Japan, Subduction Zone, *J. Geophys. Res.*, *97(B7)*, 11117-11135, doi:10.1029/92JB00810.

Savage, J. C., and J. Langbein (2008), Postearthquake relaxation after the 2004 M6 Parkfield, California earthquake and rate-and-state friction, *J. Geophys. Res.*, *113*, B10407, doi:10.1029/2008JB005723.

Savage, J. C., and J. L. Svarc (2009), Postseismic relaxation following the 1992 M7.3 Landers and 1999 M7.1 Hector Mine earthquakes, southern California, *J. Geophys. Res.*, *114*, B01401, doi:10.1029/2008JB005938.

Schaffrin, B., and Y. Bock (1988), A unified scheme for processing GPS phase observations, *Bull. Geodesique*, *62(2)*, 142-160, doi:10.1007/BF02519222.

Scherneck, H. G. (1991), A parametrized solid earth tide model and ocean tide loading effects for global geodetic baseline measurements, *Geophys. J. Int.*, *106(3)*, 677-694, doi:10.1111/j.1365-246X.1991.tb06339.x.

Schmid, R., and M. Rothacher (2003), Estimation of elevation-dependent satellite antenna phase center variations of GPS satellites, *J. Geod.*, *77(7-8)*, 440-446, doi:10.1007/s00190-003-0339-0.

Schmid, R., P. Steigenberger, G. Gendt, M. Ge, and M. Rothacher (2007), Generation of a consistent absolute phase-center correction model for GPS receiver and satellite antennas, *J. Geod.*, *81(12)*, 781-798, doi:10.1007/s00190-007-0148-y.

Schwartz, S. Y., and J. M. Rokosky (2007), Slow slip events and seismic tremor at circum-pacific subduction zones, *Rev. Geophys.*, *45*, RG3004, doi:10.1029/2006RG000208.

Segall, P., and M. Matthews (1997), Time dependent inversion of geodetic data, *J. Geophys. Res.*, *102(B10)*, 22391-22409, doi:10.1029/97JB01795.

Simons, M., Y. Fialko, and L. Rivera (2002), Coseismic deformation from the 1999 M-w 7.1 Hector Mine, California, earthquake as inferred from InSAR and GPS observations, *Bull. Seism. Soc. Am.*, *92*, 4, 1390-1402, doi:10.1785/0120000933.

Smith, E. G. C., T. D. Williams, and D. J. Darby (2007), Principal component analysis and modeling of the subsidence of the shoreline of Lake Taupo, New Zealand, 1983-1999: Evidence for dewatering of a magmatic intrusion?, *J. Geophys. Res.*, *112*, B08406, doi:10.1029/2006JG004652.

Springer, T. A., G. Beutler, and M. Rothacher (1999a), Improving the orbit estimates of GPS satellites, *J. Geod.*, *73(3)*, 147-157, doi:10.1007/s001900050230.

Springer, T. A., G. Beutler, and M. Rothacher (1999b), A New Solar Radiation Pressure Model for GPS Satellites, *GPS Solutions*, *2(3)*, 50-62, doi:10.1007/PL00012757.

Stewart, M. P., N. T. Penna, and D. D. Lichti (2005), Investigating the propagation mechanism of unmodelled systematic errors on coordinate time series estimated using least squares, *J. Geod.*, *79(8)*, 479-489, doi:10.1007/s00190-005-0478-6.

Szeliga, W., T. Melbourne, M. Santillan, and M. Miller (2008), GPS constraints on 34 slow slip events within the Cascadia subduction zone, 1997-2005, *J. Geophys. Res.*, *113*, B04404, doi:10.1029/2007JB004948.

Thiebaux, H. J., and F. W. Zwiers (1984), The Interpretation and Estimation of Effective Sample-Size, *J. Clim. Appl. Meteorol.*, *23*, 5, 800-811

Tregoning, P., and T. van Dam (2005), Atmospheric pressure loading corrections applied to GPS data at the observation level, *Geophys. Res. Lett.*, *32*, L22310, doi:10.1029/2005GL024104.

Tregoning, P., and T. A. Herring (2006), Impact of a priori zenith hydrostatic delay errors on GPS estimates of station heights and zenith total delays, *Geophys. Res. Lett.*, *33*, L23303, doi:10.1029/2006GL027706.

Tregoning, P., C. Watson, G. Ramillien, H. McQueen, and J. Zhang (2009), Detecting hydrologic deformation using GRACE and GPS, *Geophys. Res. Lett.*, *36*, L15401, doi:10.1029/2009GL038718.

van Dam, T. M., G. Blewitt, and M. B. Heflin (1994), Atmospheric-Pressure Loading Effects on Global Positioning System Coordinate Determinations, *J. Geophys. Res.*, *99(B12)*, 23,939-923,950, doi:10.1029/94JB02122.

Vautard, R., and M. Ghil (1989), Singular Spectrum Analysis in Nonlinear Dynamics, with Applications to Paleoclimatic Time-Series, *Physica D*, *35(3)*, 395-424, doi:10.1016/0167-2789(89)90077-8.

von Storch, H., and G. Hannoschock (1985), Statistical Aspects of Estimated Principal Vectors (EOFs) Based on Small Sample Sizes, *J. Clim. Appl. Meteorol.*, *24*, 7, 716-724.

von Storch, H., and F. W. Zwiers (1999), *Statistical analysis in climate research*, 484 pp., Cambridge University Press, Cambridge.

Wang, K. L., H. Dragert, H. Kao, and E. Roeloffs (2008), Characterizing an "uncharacteristic" ETS event in northern Cascadia, *Geophys. Res. Lett.*, *35*, L15303, doi:10.1029/2088GL034415.

Watson, C., P. Tregoning, and R. Coleman (2006), Impact of solid Earth tide models on GPS coordinate and tropospheric time series, *Geophys. Res. Lett.*, *33*, L08306, doi:10.1029/2005GL025538.

Watson, K. M., Y. Bock, and D. T. Sandwell (2002), Satellite interferometric observations of displacements associated with seasonal groundwater in the Los Angeles basin, *J. Geophys. Res.*, *107*, (B4), 2074, doi:10.1029/2001JB000470

Wdowinski, S., Y. Bock, J. Zhang, P. Fang, and J. Genrich (1997), Southern California Permanent GPS Geodetic Array: Spatial filtering of daily positions for estimating coseismic and postseismic displacements induced by the 1992 Landers earthquake, *J. Geophys. Res.*, *102(B8)*, 18057-18070, doi:10.1029/97JB01378.

Webb, F., and J. Zumberge (1997), *An Introduction to GIPSY/OASIS-II, Precision Software for the Analysis of Data from the Global Positioning System*, Jet Propulsion Laboratory.

Wech, A. G., and K. C. Creager (2008), Automated detection and location of Cascadia tremor, *Geophys. Res. Lett.*, *35*, L20302, doi:10.1029/2008GL035458.

Wech, A. G., K. C. Creager, and T. I. Melbourne (2009), Seismic and geodetic constraints on Cascadia slow slip, *J. Geophys. Res.*, *114*, B10316, doi:10.1029/2008JB006090.

Wech, A. G., K. C. Creager, H. Houston, and J. E. Vidale (2010), An earthquake-like magnitude-frequency distribution of slow slip in northern Cascadia, *Geophys. Res. Lett.*, *37*, L22310, doi:10.1029/2010GL044881.

Wicks, C. W., W. Thatcher, D. Dzurisin, and J. Svarc (2006), Uplift, thermal unrest and magma intrusion at Yellowstone caldera, *Nature*, *440*, 72-75, doi:10.1038/nature04507.

Williams, S. D. P. (2003a), Offsets in Global Positioning System time series, *J. Geophys. Res.*, *108(B6)*, 2310, doi:10.1029/2002JB002156.

Williams, S. D. P. (2003b), The effect of coloured noise on the uncertainties of rates estimated from geodetic time series, *J. Geod.*, *76(9-10)*, 483-494, doi:10.1007/s00190-002-0283-4.

Williams, S. D. P., Y. Bock, P. Fang, P. Jamason, R. M. Nikolaidis, L. Prawirodirdjo, M. Miller, and D. J. Johnson (2004), Error analysis of continuous GPS position time series, *J. Geophys. Res.*, *109*, B03412, doi:10.1029/2003JB002741.

Wu, J. T., S. C. Wu, G. A. Hajj, W. Bertiger, and S. M. Lichten (1993), Effects of antenna orientation on GPS carrier phase, *Manuscr. Geod.*, *18*, 91-98.

Wunsch, C. (2006), *Discrete inverse and state estimation problems : with geophysical fluid applications*, 371 pp., Cambridge University Press, Cambridge.

Wyatt, F. (1982), Displacement of Surface Monuments: Horizontal Motion, *J. Geophys. Res.*, *87*, 979-989, doi:10.1029/JB087iB02p00979.

Wyatt, F. K. (1989), Displacement of Surface Monuments: Vertical Motion, *J. Geophys. Res.*, *94*, 1655-1664, doi:10.1029/JB094iB02p01655.

Yang, Y., and W. Gao (2006), An optimal adaptive Kalman filter, *J. Geod.*, *80*, 4, 177-183, doi:10.1007/s00190-006-0041-0

Yoder, C. F., J. G. Williams, and M. E. Parke (1981), Tidal Variations of Earth Rotation, *J. Geophys. Res.*, *86(B2)*, 881-891, doi:10.1029/JB086iB02p00881.

Zhang, J., Y. Bock, H. Johnson, P. Fang, S. Williams, J. Genrich, S. Wdowinski, and J. Behr (1997), Southern California Permanent GPS Geodetic Array: Error analysis of daily position estimates and site velocities, *J. Geophys. Res.*, *102(B8)*, 18035-18055, doi:10.1029/97JB01380.

Zumberge, J. F., M. B. Heflin, D. C. Jefferson, M. M. Watkins, and F. H. Webb (1997), Precise point positioning for the efficient and robust analysis of GPS data from large networks, *J. Geophys. Res.*, *102(B3)*, 5005-5017, doi:10.1029/96JB03860.