Monitoring of the Long Valley Caldera inflation episode from a principal component analysis-based inversion method applied to InSAR and EDM Data

Yu-nung Nina Lin[1], Andrew Kositsky[1], and Jean-Philippe Avouac[1]
1. Devision of Geological and Planetary Sciences, California Institute of Technology

Abstract

In this study we apply the principal component analysis-based inversion method (PCAIM, Kositsky2009) on multiple sources of geodetic datasets. PCAIM can be used to extract from the original data the signals that are coherent both in time and in space. Because tectonic and non-tectonic signals have different properties of coherency, they contribute different amounts to each component. By inverting and summing the components with the largest portion of tectonic signal separately with known Greens functions and discarding those that appear to primarily be noise, we can rebuild the InSAR time series with primarily tectonic signals.
We test this method in the Long Valley Caldera area which experienced multiple inflation episodes since the 80's. We focus on the 1997-98 episode, using 24 ERS scenes and 65 interferograms. To give the decomposition more temporal resolution and continuity, we add 8 two-color electronic distance meter (EDM) time series which have dense temporal sampling rates to carry out joint decomposition and inversion. The result shows that the first principal component contains most of the inflation signals with a clear pulse in 1997-98. A direct inversion using Mogi's model shows the inflation of magma near 11 km in depth. This study proves the capability of PCAIM to model and interpret InSAR time series and perform true joint inversions between multiple data source.

Techinical Overview of PCAIM

We place a geodetic dataset with identical sampling epochs in a $m \times n$ matrix X_0 where each row corresponds to one time-series from the whole dataset. We assume that the medium is elastic, and therefore the relationship between the source that generates the signal and the response at the surface is linear. Surface displacement then obey:

$$X_0 = G_\alpha L$$

where G_α denotes the Green's functions relating surface displacement X_0 with the source L at depth. We then apply the principal component decomposition on the data matrix X_0 and get:

$$X_r = U_r S_r V_r^t$$

where U is the spatial function, V is the time function, S is the principal value and r is the number of principal components necessary to fit the data within uncertainties. In this decomposition, each individual component corresponds to a linear combination of the contributions from various sources and not to a particular, identifiable physical source, although the various components can be recombined to extract the contribution of particular sources. To determine how many components are needed to represent the original data, we select the number of components so that the reduced chi square X^2_{red} of $X - X_r$ is approximately equal to one, i.e.

$$X^2_{red} = \frac{1}{N-r(n+m+1)} \sum_{i=1}^{m} \sum_{j=1}^{n} \frac{(X(i,j)-X_r(i,j))^2}{\sigma(i,j)^2} \approx 1$$

where N is the total number of data. We can also use F-test to determine the appropriate number of components to be selected:

$$F = \frac{X_k^2 - X_{k+1}^2}{X_{k+1}^2} \frac{N-p_{k+1}}{p_{k+1} p_k}$$

Once we determine the appropriate number of components to use in the decomposition, we can carry out the inversion on each of the spatial function:

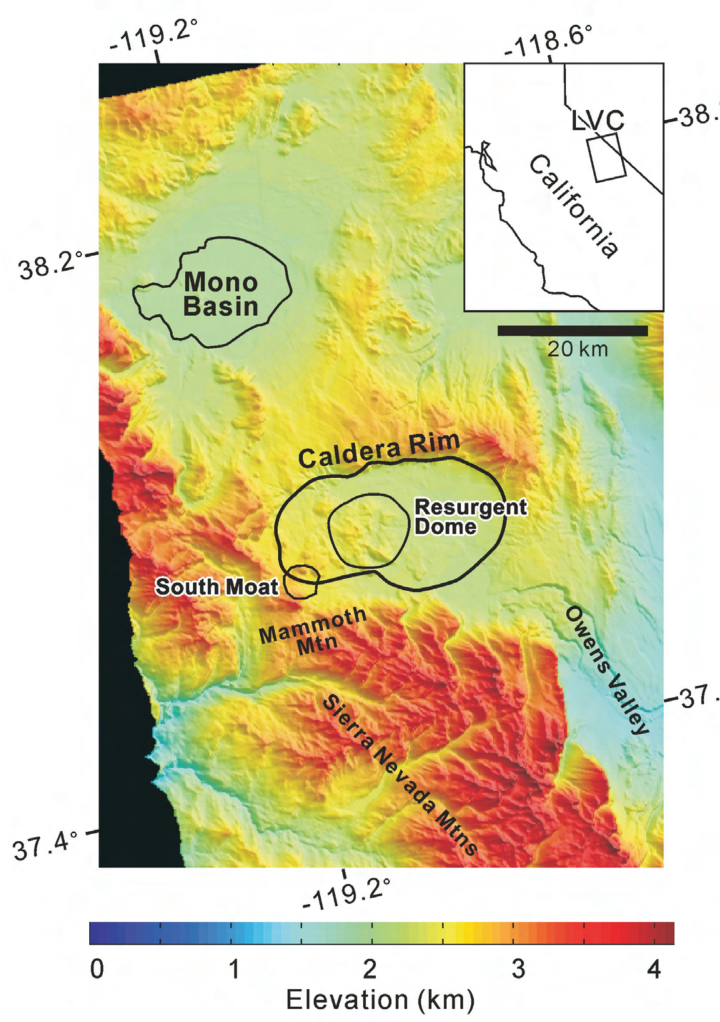
$$G_\alpha L_r = U_r$$

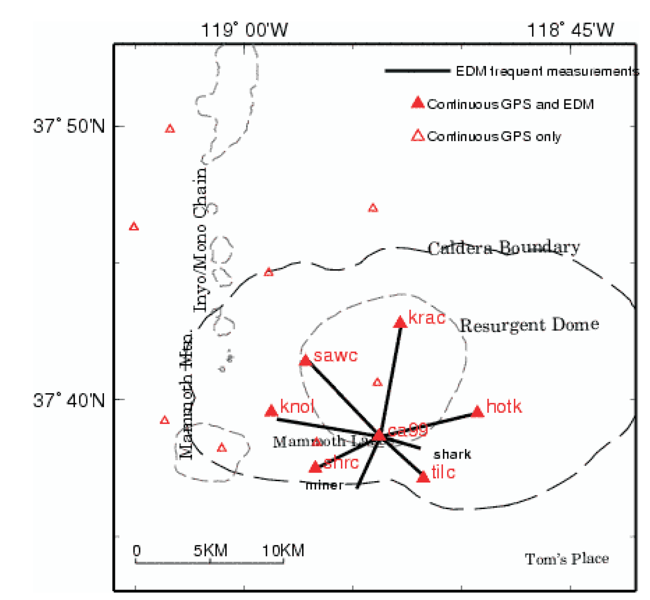
After solving for the source model for each component, we can sum up the contribution from all the components and derive the time-series for the source.

Inversion by Joining Datasets of Different Sampling Epochs

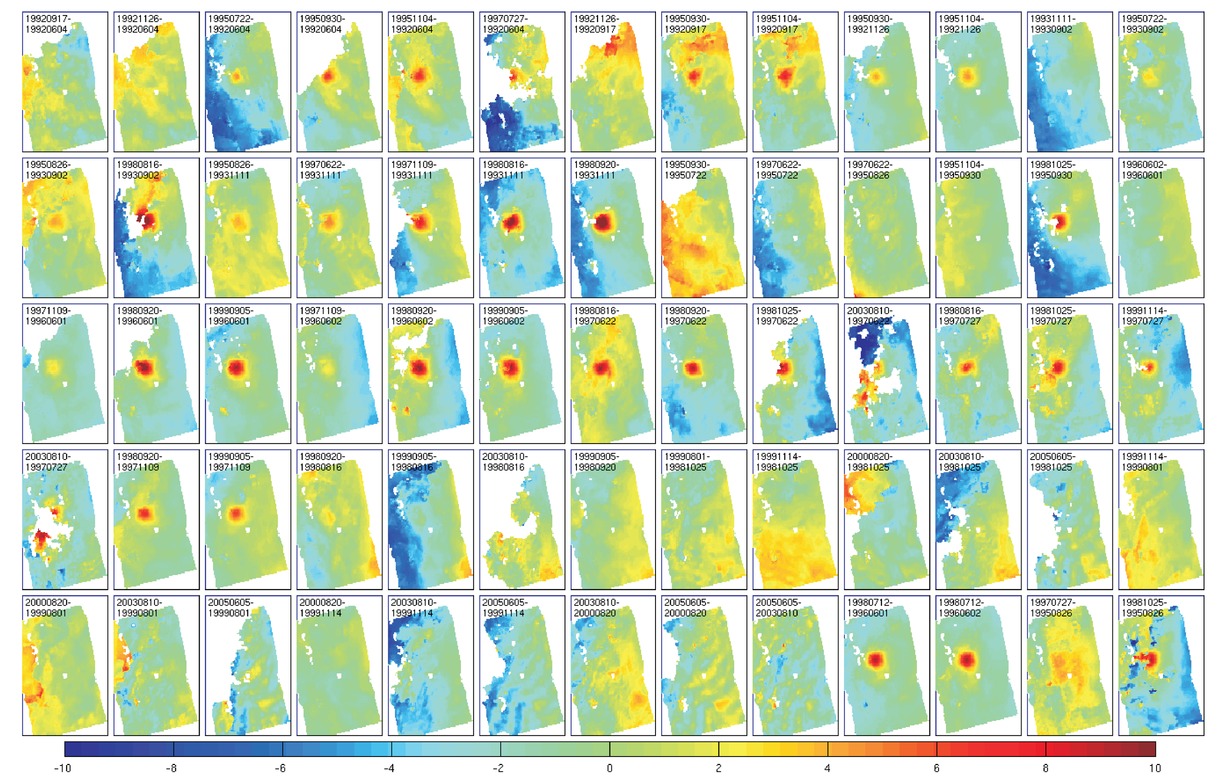
Now we have two geodetic datasets with very different sampling epochs, for example InSAR vs. GPS. If we combine them together into one big data matrix X_0, the matrix will be zero-padded in many entries due to the low temporal sampling rate of SAR imagery. We have adaped our method by including a data error matrix W_0. Whenever there is no data entry in the X_0 matrix, we set its error to be `Inf`. Then we use a more sophisticated decomposition developed by *Srebro and Jaakkola*[2003]. With this adaptation, we can take advantage of the high spatial sampling rate of InSAR data and high temporal sampling rate of GPS (or other continuous geodetic measurement) data to invert for the time-series of source function.

In this study, we select the Long Valley Caldera to test this joint inversion method. The Long Valley Caldera has experienced a large inflation event between mid-April and late June 1997. This episode is one of the five inflation events through the past 30 years. During the 1997 episode, it first showed an exponential growth increase in mid-April, and an exponential growth decay in late November 1997, cumulating in ~10 cm of uplift [Newman et al., 2001; Hill 2003]. We use 65 interferograms and carry out the simplest version of SBAS time-series [Berardino et al., 2002]. We also including the two-color electronic distance measurement (EDM) data which has daily record since 1984 [Langbein 2003]. Next we show the work flow of combining datasets for PCAIM analysis.

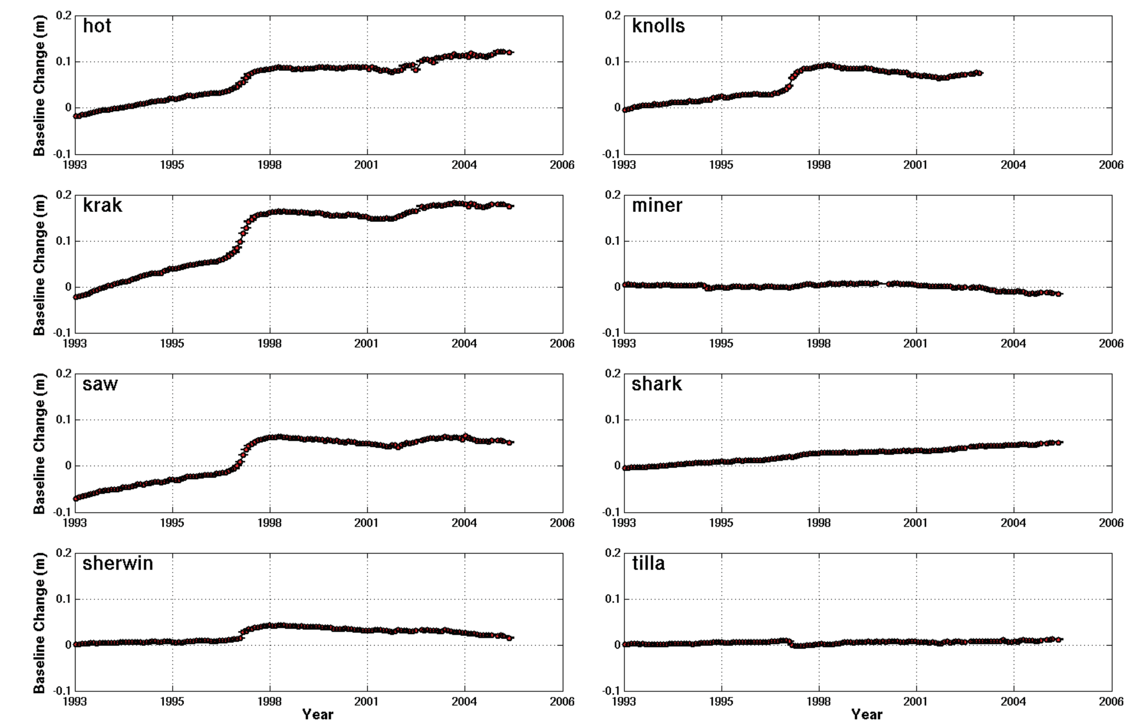




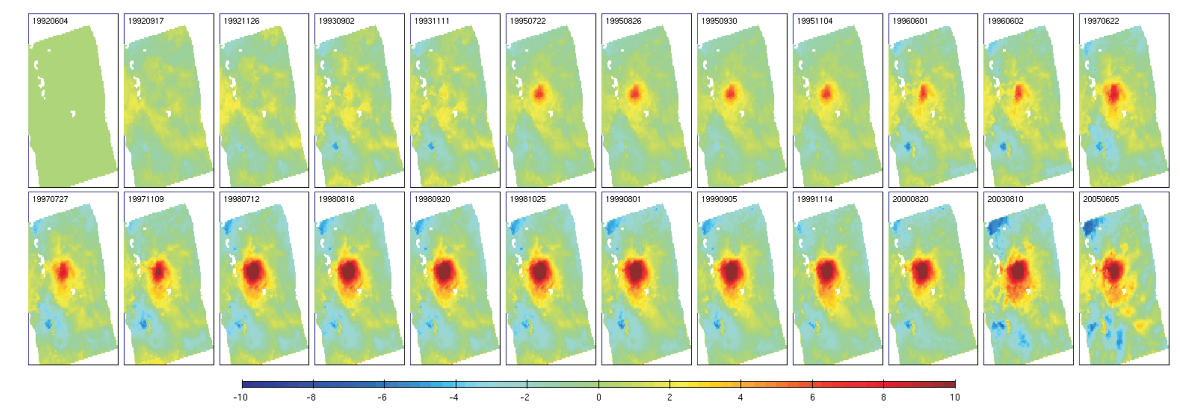
InSAR Original Dataset



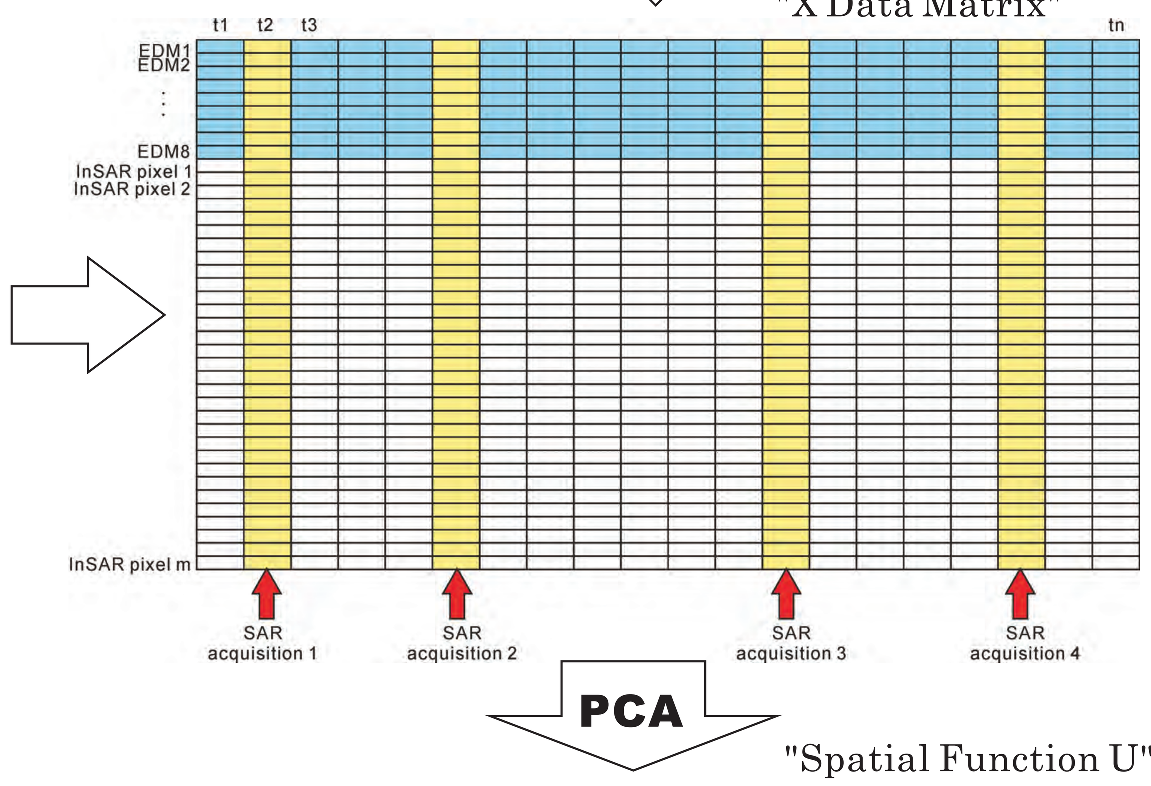
Electronic Distance Measurements (EDM)



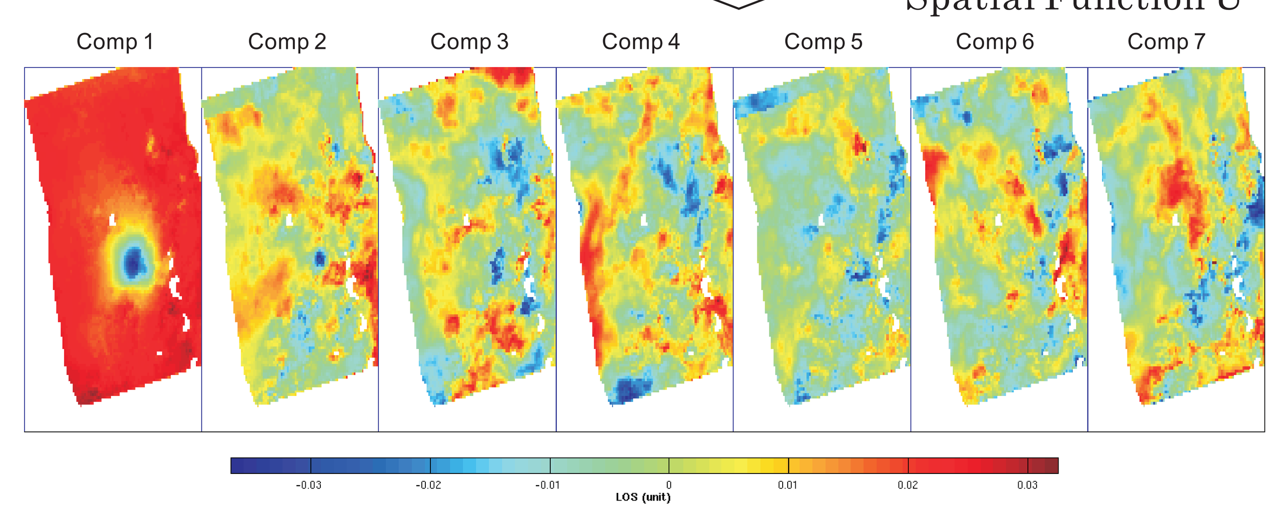
Simple SBAS Time Series



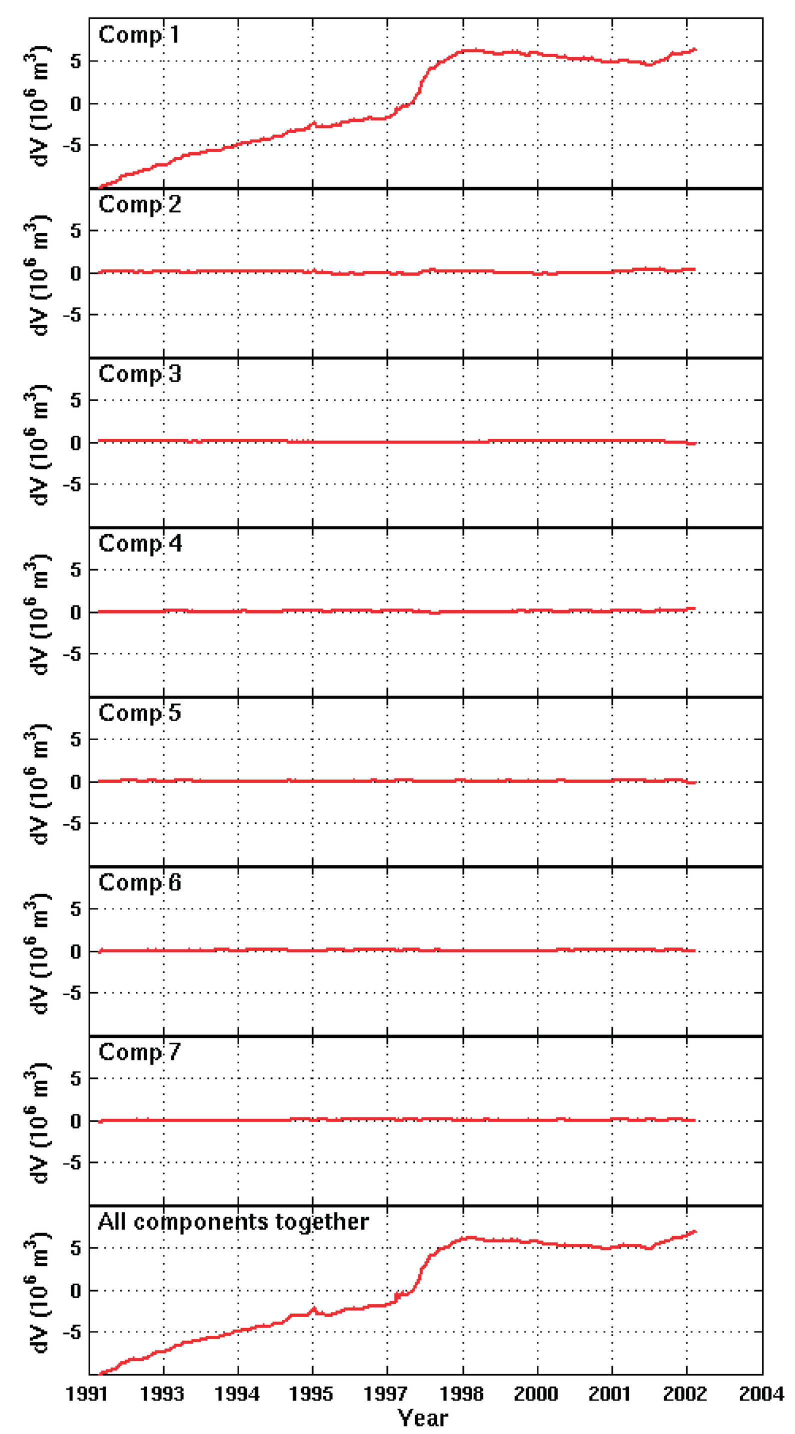
"X Data Matrix"



"Spatial Function U"

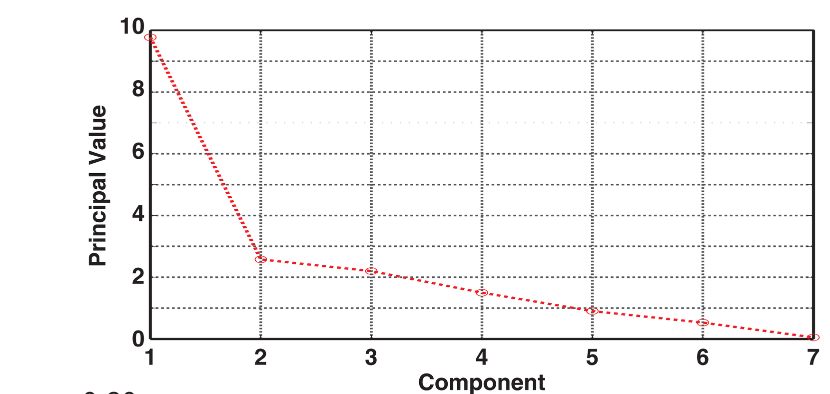


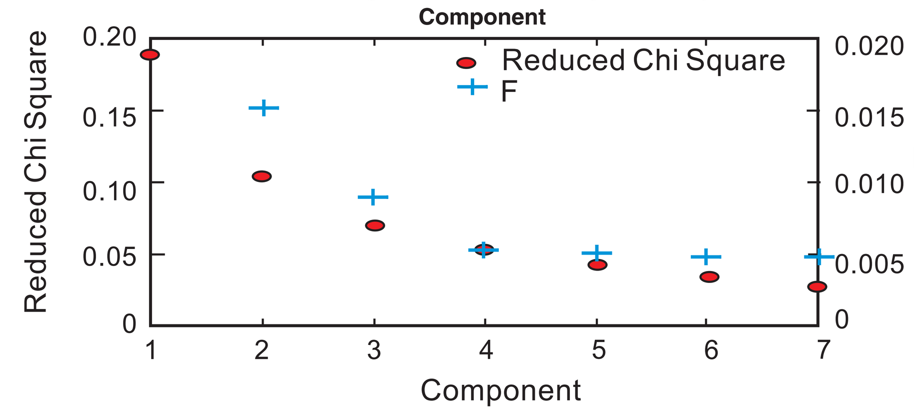
Long Valley Magmatic Inflation
Source Depth = 11 km

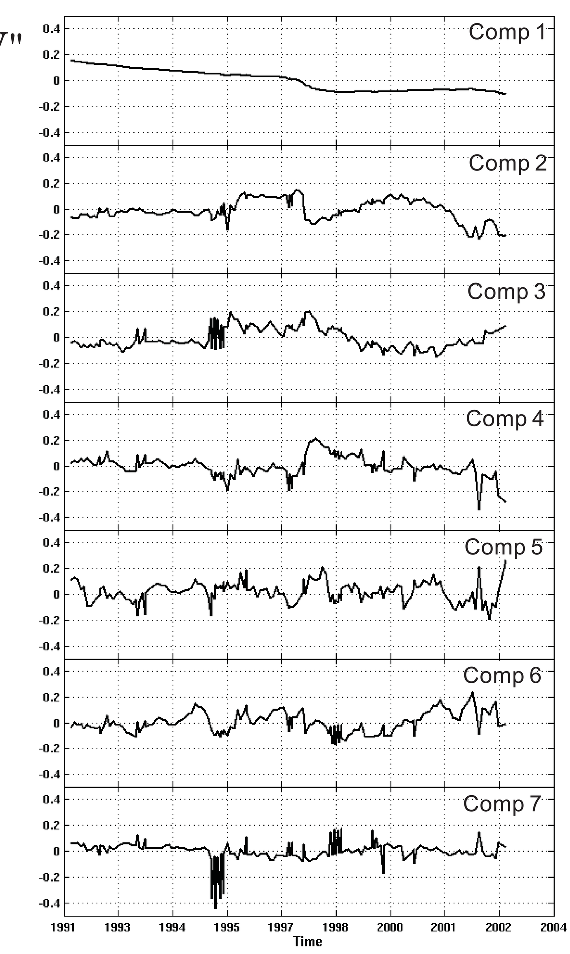


Inversion

"Temporal Function V"







Summary

For this preliminary test, the magmatic inflation signals seem mostly separated into the first component. The rest of the components look more or less random. It is still possible that there may be some tectonic signals within component 2 to 7, but because their amplitude is not big enough, they can be easily mixed with tropospheric signals. In that sense PCAIM may not be able to solve minor tectonic events. However, PCAIM does provide us with a fast and relatively simple way to analyze and run inversion on the time series, with all decent spatial resolution.

Another interesting test in the future would be to recombine the spatial function derived from InSAR data and compare with independent water vapor mappings. This method may serve as a complementary way to correct for tropospheric signals for other modeling purposes.

References

Hill, D., J. Langbein, and S. Prejean, Relations between seismicity and deformation during unrest in Long Valley Caldera, California, from 1995 through 1999, *J. Volcanol. Geotherm. Res.*, *127*, 175–193, doi:10.1016/S0377-0273(03)00169-0, 2003.

Kositsky, A., and J.-P. Avouac, Inverting geodetic time-series with a principal component analysis-based inversion method (PCAIM), *J. Geophys. Res.*, in press, 2009.

Langbeinn, J., Deformation of the Long Valley caldera, California: Inferences from measurements from 1988 to 2001, *J. Volcanol. Geotherm. Res.*, *127*, 247–267, doi:10.1016/S0377V0273(03)00172-0, 2003.

Newman, A., T. Dixon, G. Ofoegbu, and J. Dixon, Geodetic and seismic constrains on recent activity at Long Valley Caldera, California: Evidence for viscoelastic rheology, *J. Volcanol. Geotherm. Res.*, *105*, 183–206, doi:10.1016/S0377-0273(00)00255-9, 2001.