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

Yunung Nina Lin, Andrew Kositsky, and Jean-Philippe Avouac
1. Division of Geological and Planetary Sciences, California Institute of Technology

Abstract

In this study, we apply the principal component analysis-based inversion method (PCAIM, Kot-
sky et al., 2009) in multiple sources of geodetic datasets. PCAIM can be used to extract from the original dataset the signals that are coherent both in time and in space. Because tectonic and magmatic 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 Green's functions and discarding those that appear to primarily be noise, we can rebuild the InSAR times series with primarily tectonic signals. We test this method in the Long Valley Caldera area which experienced multiple inflation episodes since the 1990s. We focus on the 1997-1998 episode, using 24 InSAR images and 65 interferograms to give the decomposition more temporal resolution and continuity. We find a tectonic component that has dense temporal sampling rates to carry out joint deconvolution and inversion. The result shows that the first principal component contains most of the inflation signals with a clear pulse in 1997-1998. 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 times series and perform true joint inversions between multiple data sources.

Technical Overview of PCAIM

We place a geodetic dataset with identical sampling epochs in a N × m matrix X₀ where each row corresponds to one timeseries from the whole dataset. We assume that the medium is static, and therefore the relationship between the source that generates the signal and the response at the surface is linear. Surface displacement then obey:

\[ \mathbf{X}_0 = \mathbf{G}_0 \mathbf{U} \]

where \( \mathbf{G}_0 \) denotes the Green's functions relating surface displacement \( \mathbf{X}_0 \) with the source \( \mathbf{U} \) at depth. We then apply the principal component decomposition on the data matrix \( \mathbf{X}_0 \) and get:

\[ \mathbf{X}_0 = \mathbf{U} \mathbf{V} \mathbf{S} \]

where \( \mathbf{U} \) is the spatial function, \( \mathbf{V} \) is the time function, \( \mathbf{S} \) is the source vector, and \( n \) is the number of principal components necessary to fit the data within uncertainty. 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 \( \chi^2_{red} \) of \( \mathbf{X}_0 = \mathbf{U} \mathbf{V} \mathbf{S} \) is approximately equal to one, i.e.,

\[ \chi^2_{red} = \frac{\sum (\mathbf{X}_0 - \mathbf{U} \mathbf{V} \mathbf{S})^2}{\sum (\mathbf{X}_0 - \mathbf{U} \mathbf{V} \mathbf{S})^2} = 1 \]

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

\[ F = \frac{\mathbf{X}_0 - \mathbf{U} \mathbf{V} \mathbf{S}}{\sum (\mathbf{X}_0 - \mathbf{U} \mathbf{V} \mathbf{S})^2} \]

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

\[ \mathbf{G}_j \mathbf{S}_j = \mathbf{U}_j \]

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

Inversion by Joining Dataset of Different Sampling Epochs

Now we have two geodetic datasets with very different sampling epochs, for example InSAR vs. EDM. If we combine them together into one big data matrix \( \mathbf{X}_0 \), the matrix will be zero-padded in many entries due to the low temporal sampling rate of SAR imagery. We have adapted our method by including a data error matrix \( \mathbf{W}_0 \). Whenever there is no data entry in the \( \mathbf{X}_0 \) matrix, we set its error to be high. Then we use a more sophisticated decomposition developed by Sterk et al. (2009). With this adaptation, we can take advantage of the high spatial sampling rate of InSAR data and high temporal sampling rate of EDM data to invert for the timeseries of source function.

In this study, we use 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 over the past 30 years. During the 1997 event, it first showed an exponentially increasing strain, and an exponential stress decay in late November 1997, culminating in ~10 cm of uplift (Newman et al., 2001; Hill 2003). We use 65 interferograms and carry out the simplest version of SHAS timeseries (Bachman et al., 2002). We also include the two-electronic distance measurement (EDM) data which has been conducted since 1984 (Langlois, 2005). Next we show the workflow of combining datasets for PCAIM analysis.

InSAR Original Dataset

Electronic Distance Measurements (EDM)

"X Data Matrix"

"Spatial Function V"

"Temporal Function U"

Summary

For this preliminary test, the tectonic inflation signals are 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 amplitudes are not big enough, they can be easily mixed with topographic signal. In that case, 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 times series, with all distinct spatial resolution. Another interesting test in the future would be to reconsider the spatial function derived from InSAR data and compare with independent laser survey measurements. This method may serve as a complementary way to correct for topographic signals for other modeling purposes.

References


