[1] Jet Propulsion Laboratory
California Institute of Technology

Variational Bayesian Independent Component Analysis for InSAR displacement time series with application to Central Valley, California

Adriano Gualandi[1,2,*] and Zhen Liu[1]

* adriano.geolandi@gmail.com

On-line meeting, 4-8 May 2020

[2] Caltech

EGU General Assembly 2020

EGU2020-19567

Outline

- Unsupervised learning for InSAR time series analysis:
  Independent Component Analysis (ICA)
  - S-mode vs T-mode
  - FastICA vs Variational Bayesian ICA (vbICA)

- Test cases:
  - Synthetic #1: linear mix of 2 Mogi sources
  - Synthetic #2: linear mix of 1 Mogi source and 1 fault creeping signal

- Real cases:
  - Central Valley
  - Central San Andreas Fault

- Conclusions

Multivariate approach to time series analysis

Data matrix, *X*, contains spatio-temporal information

$$X_{M\times T} = \begin{pmatrix} x_{11} & \cdots & x_{1T} \\ \vdots & \ddots & \vdots \\ x_{M1} & \cdots & x_{MT} \end{pmatrix}$$

T-mode
Statistically independent temporal evolution
Chaussard & Farr, *GRL*, 2019

$$\mathbf{x_j} \in \mathbb{R}^T, \qquad j = 1, \ldots, M \qquad \text{random variable}$$

$$X_{T\times M} = \begin{pmatrix} x_{11} & \cdots & x_{1M} \\ \vdots & \ddots & \vdots \\ x_{T1} & \cdots & x_{TM} \end{pmatrix} = \begin{pmatrix} \boldsymbol{x}_1 \\ \vdots \\ \boldsymbol{x}_T \end{pmatrix}$$

S-mode
Statistically independent spatial distributions
Ebmeier, *JGR*, 2016
Gaddes et al., *JGR*, 2018

$$\mathbf{x_t} \in \mathbb{R}^M, \qquad t = 1, \ldots, T \qquad \text{random variable}$$

FastICA vs vbICA

$$X_{cen} \cong AS + N$$

A: mixing matrix
S: sources
N: Gaussian noise

FastICA (Hyvärinen and Oja, 1997)
Cumulant based approximation: deviations from Gaussianity up to the 4th order are used as a proxy for statistical independence

What if sources with multimodal distribution?

vbICA (Choudrey and Roberts, 2003)
Modeling approach which uses a Mix of Gaussians (MoG) to express the sources

Hidden variables & unknown parameters

$$\mathbf{W} = \{\mathbf{A}, \mathbf{\Lambda}, \mathbf{S}, \mathbf{q}, \boldsymbol{\theta}\}$$

$$p'(\mathbf{W}) = p'(\mathbf{\Lambda})p'(\mathbf{A})p'(\mathbf{S} \mid \mathbf{q})p'(\mathbf{q})p'(\boldsymbol{\theta})$$

Variational approach

$$p(W|X) = \frac{p(X|W)p(W)}{p(X)}$$ Bayes theorem

$$KL[p'(W)||p(W|X)] \equiv \int_W p'(W)\ln\frac{p'(W)}{p(W|X)}\,dW =$$

$$= \int_W p'(W)\ln\frac{p'(W)p(X)}{p(X|W)p(W)}\,dW =$$

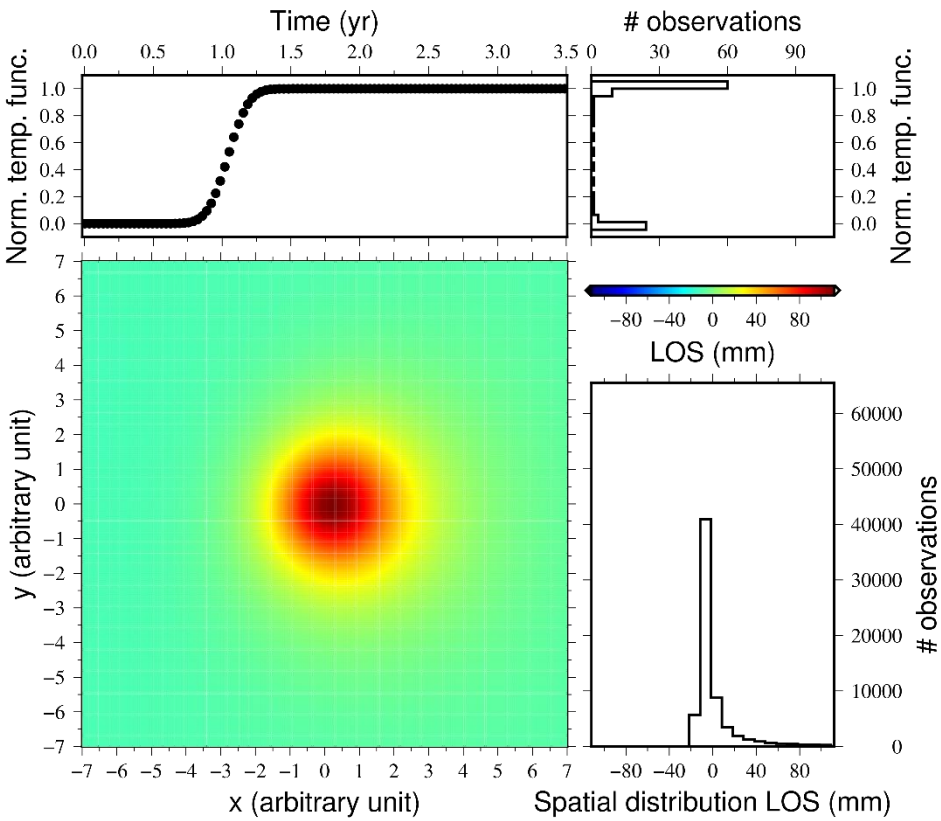
$$= -\int_W p'(W)\ln\frac{p(X|W)p(W)}{p'(W)}\,dW + \int_W p'(W)\ln p(X)\,dW$$

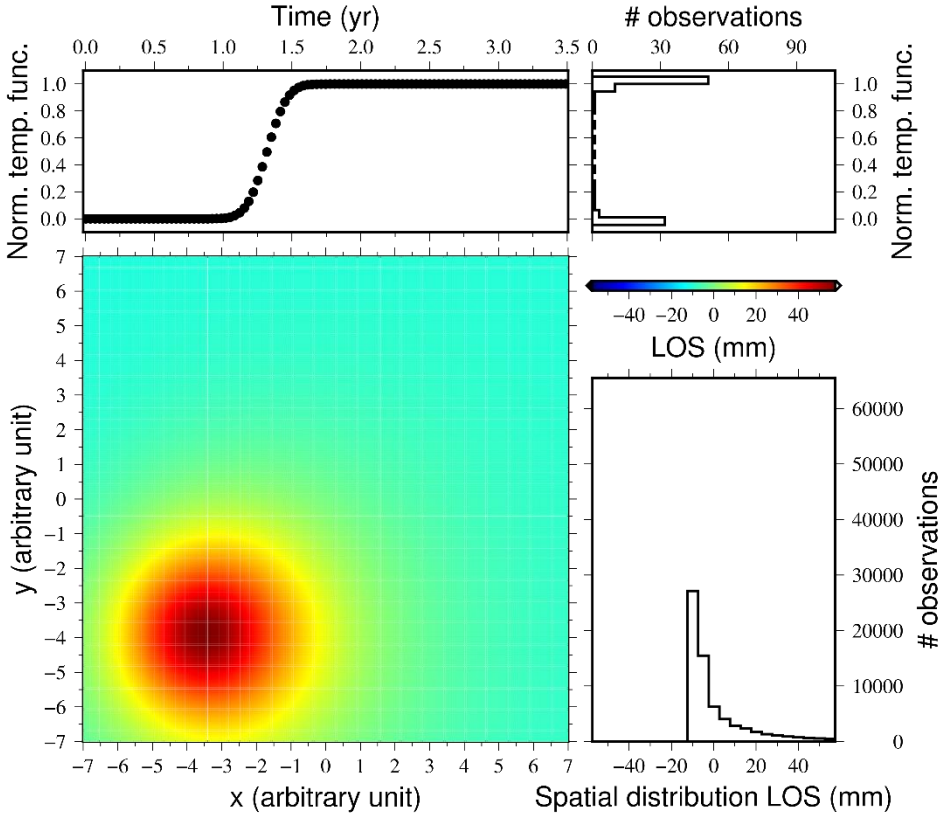
$$-NFE \qquad \ln p(X)$$

$$\Rightarrow \quad \ln p(X) = NFE + KL[p'(W)||p(W|X)]$$

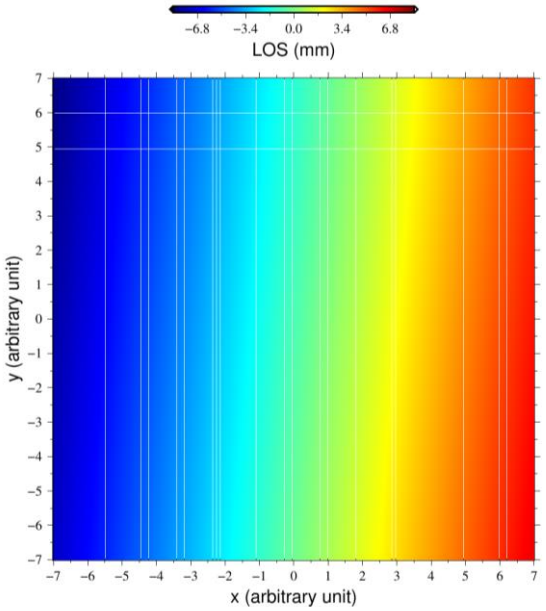
Synthetic #1

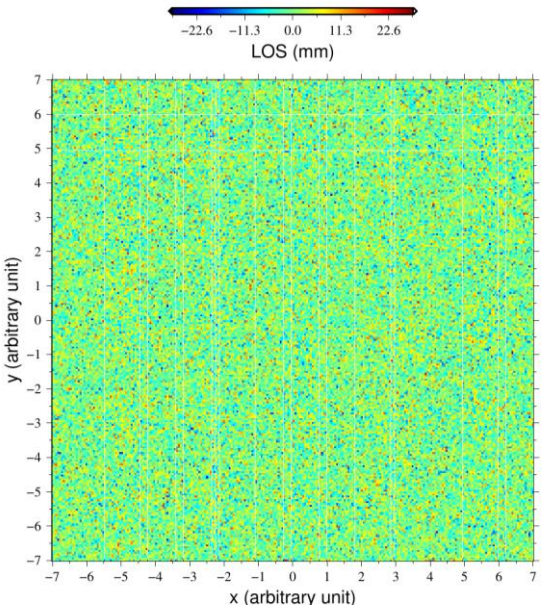
True sources

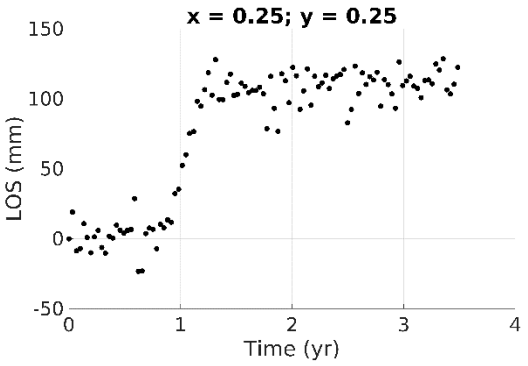


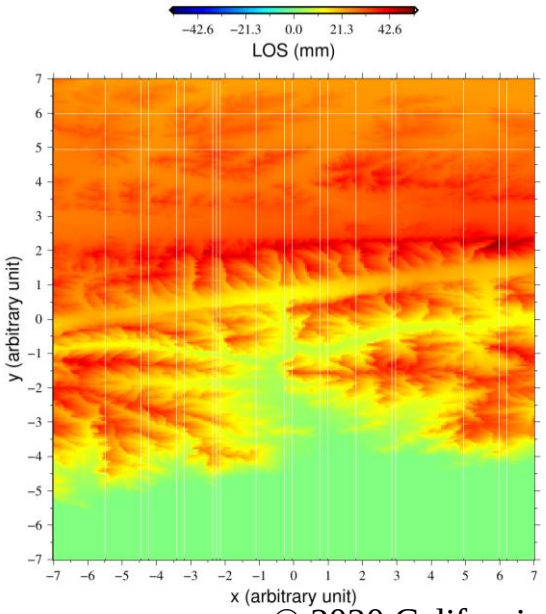


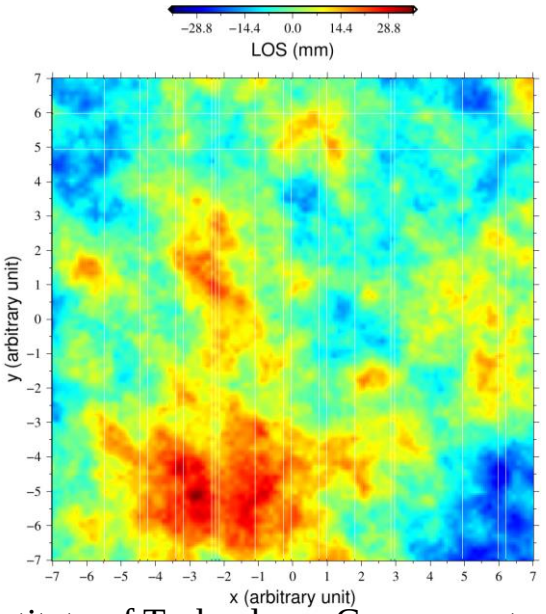
Noise

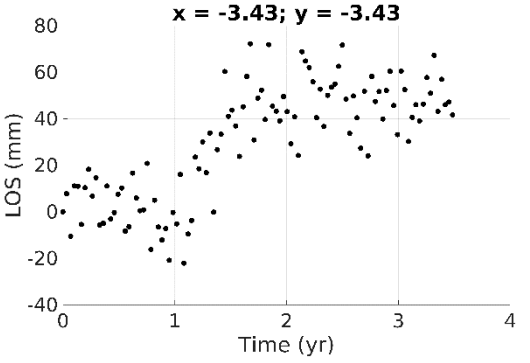




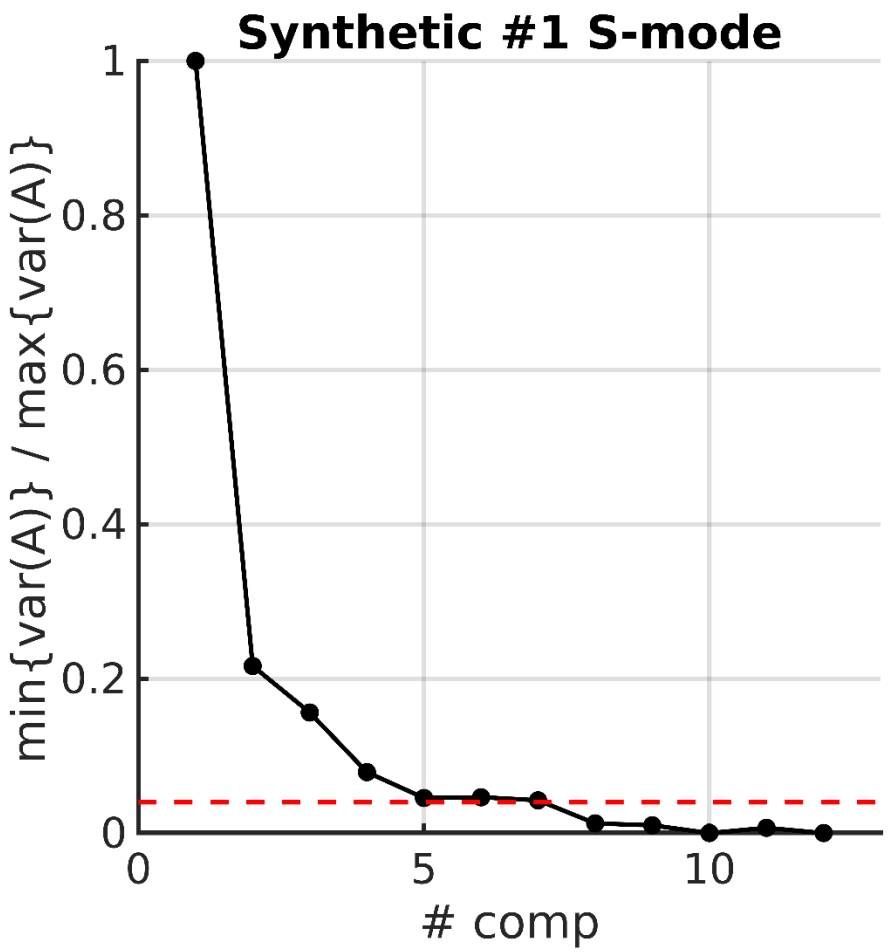


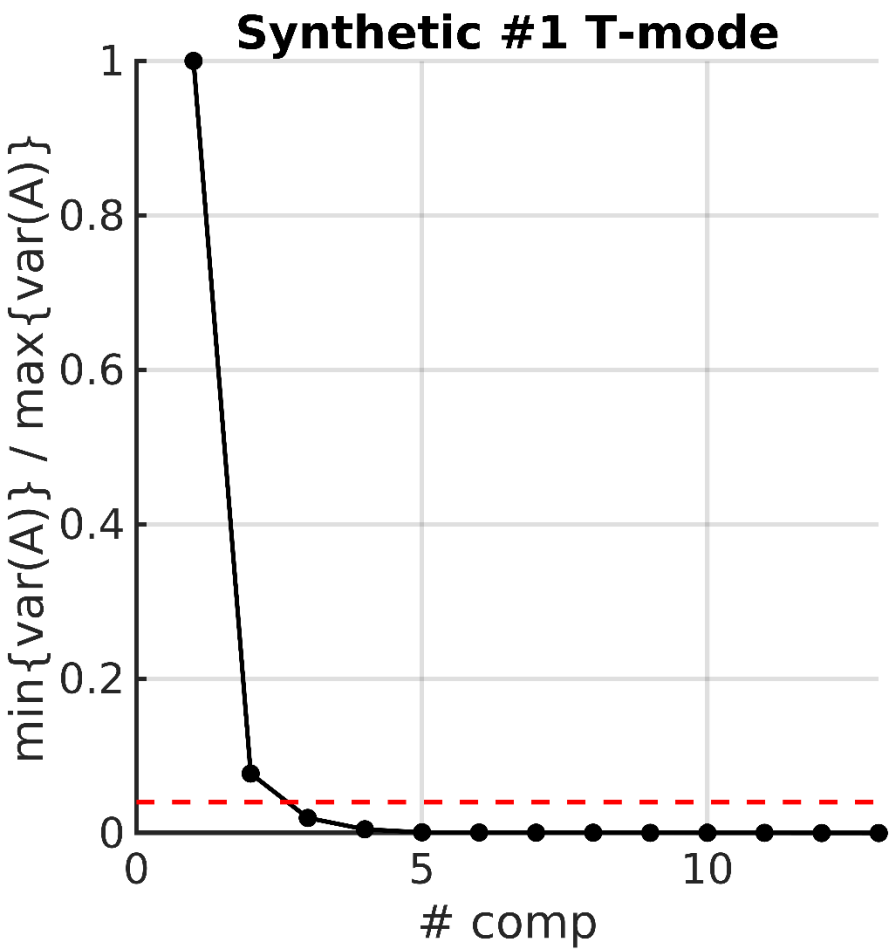






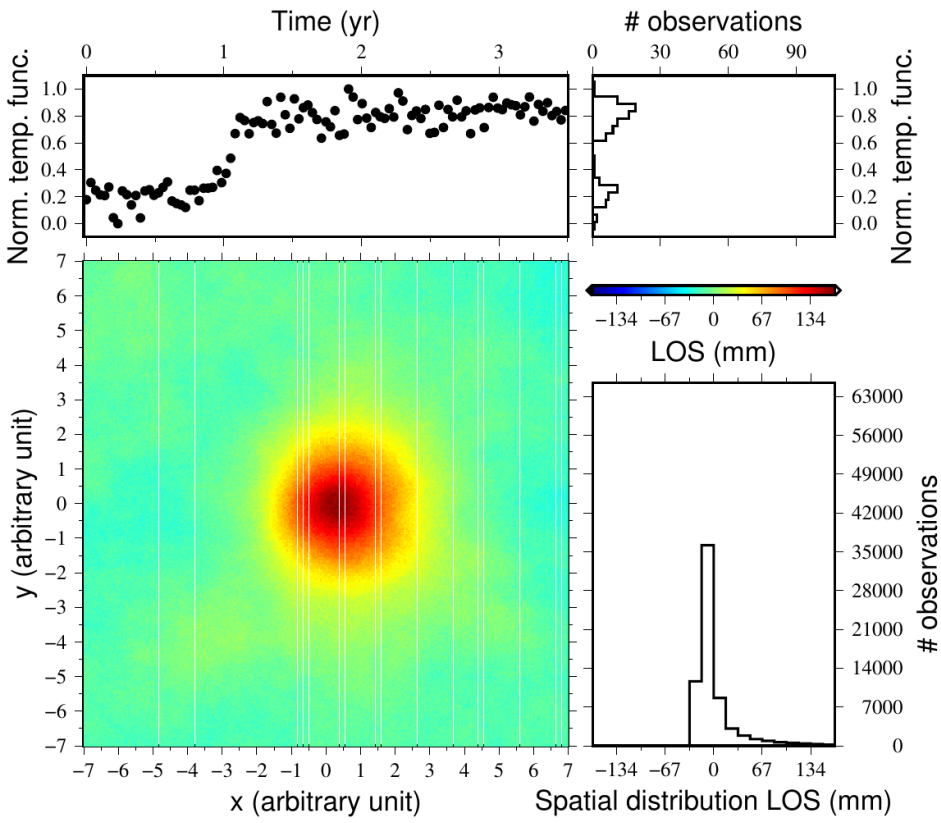
How many components?

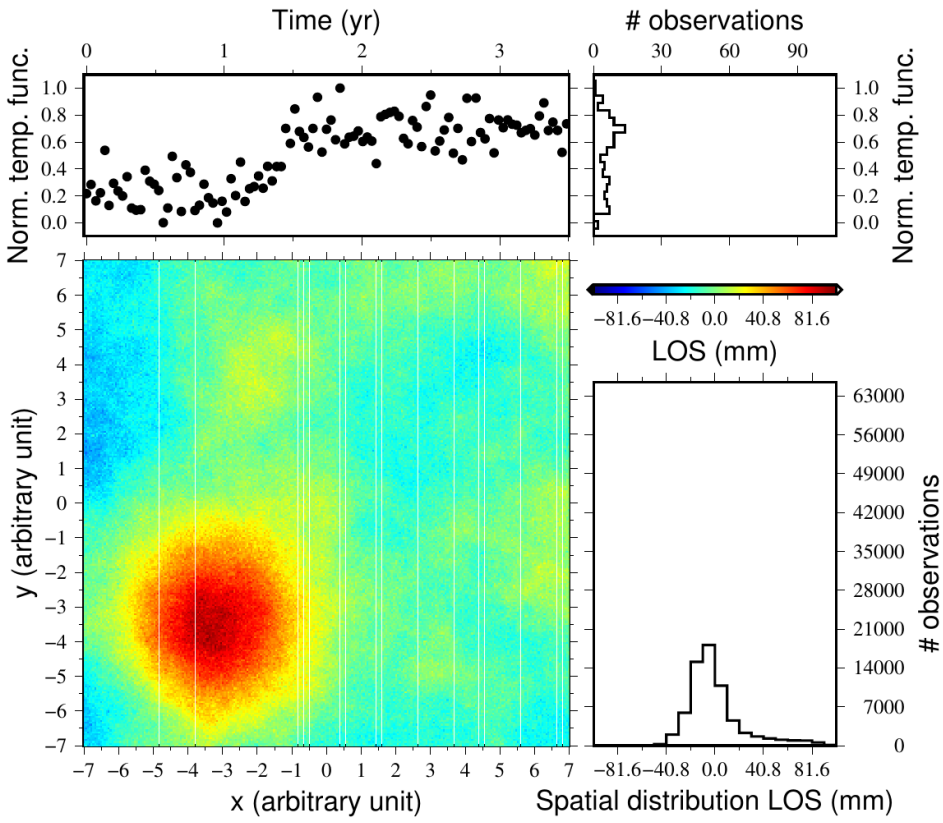




Synthetic #1: S-mode

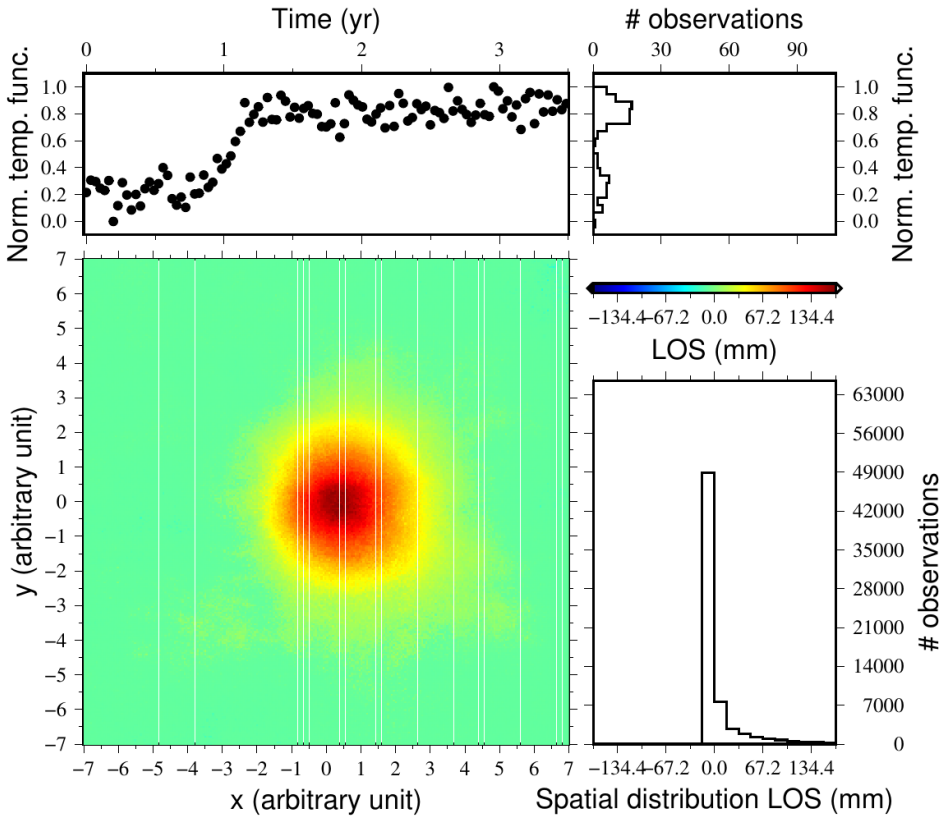
FastICA

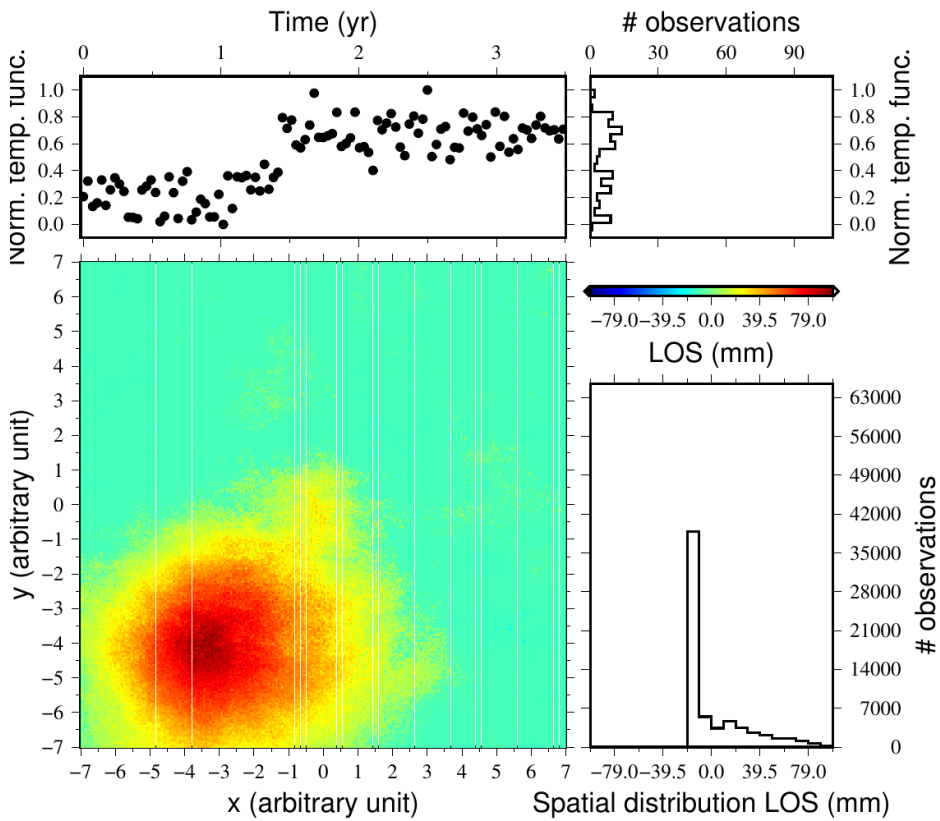




Synthetic #1: S-mode

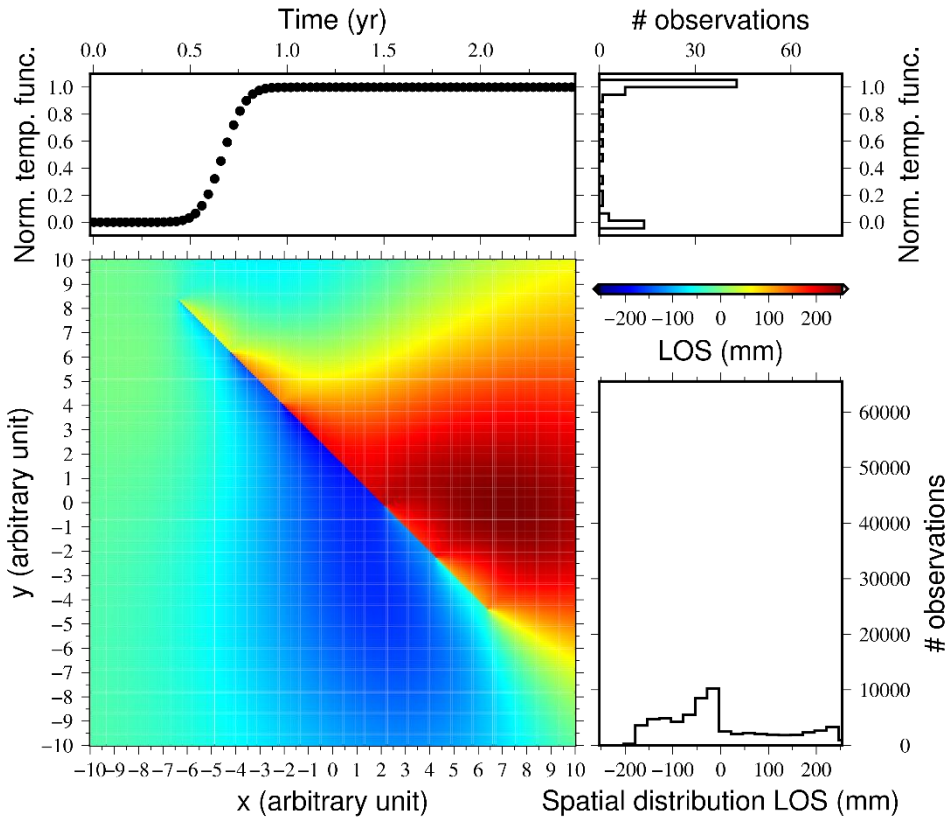
vbICA

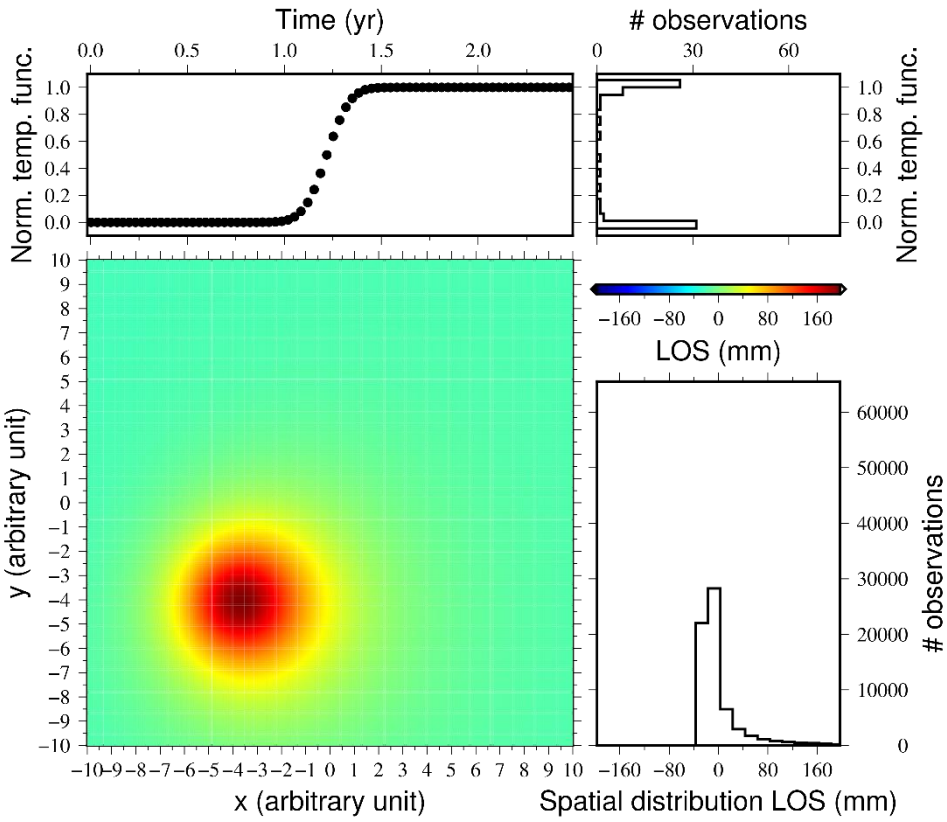




Synthetic #2

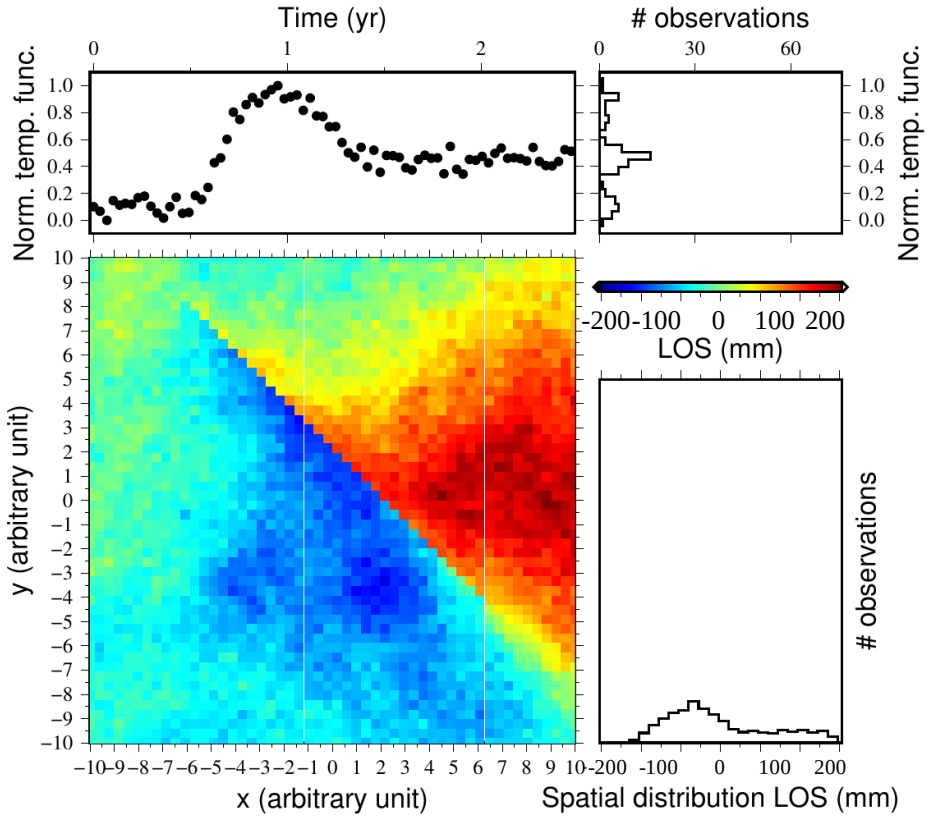
True sources

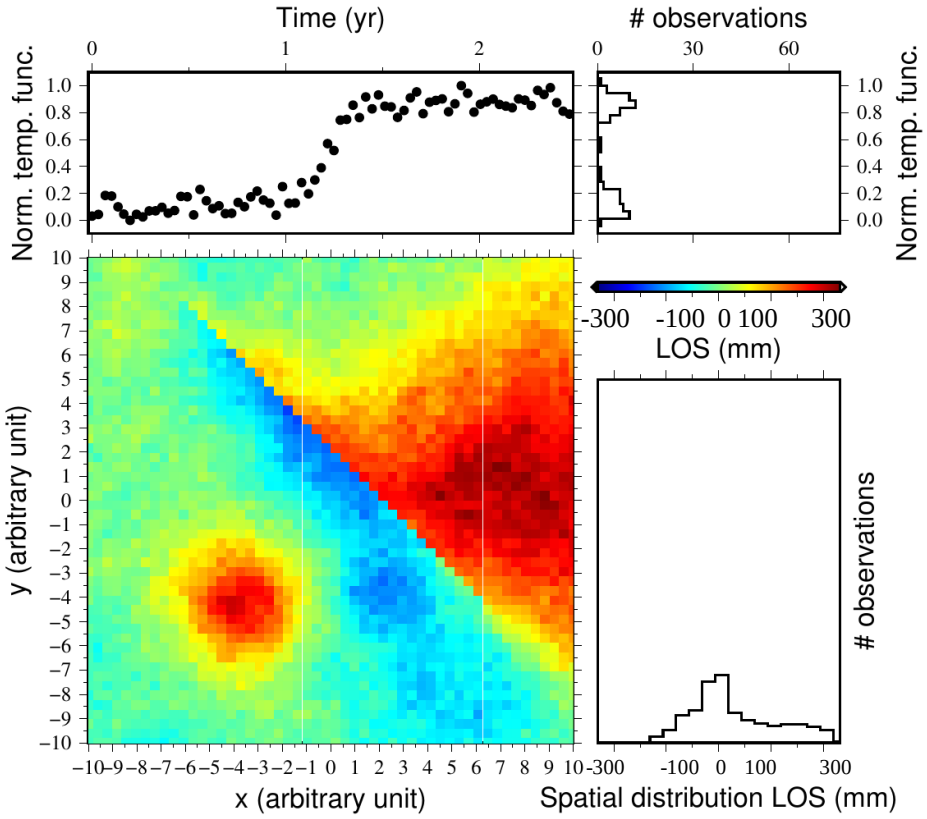




Synthetic #2: T-mode

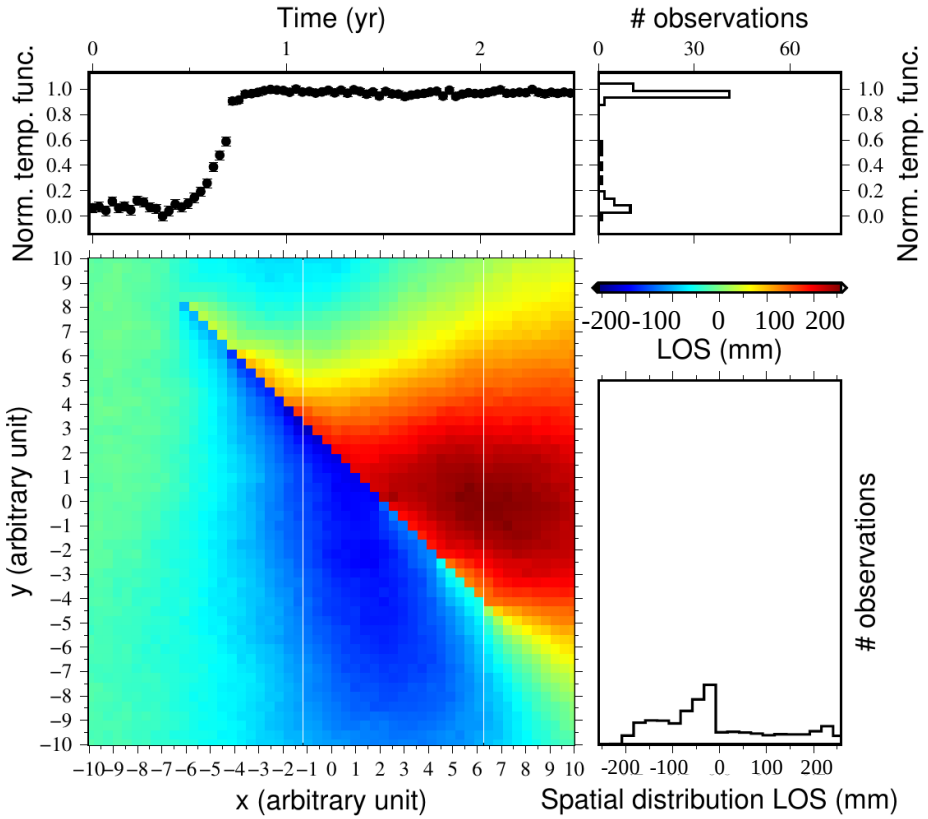
FastICA

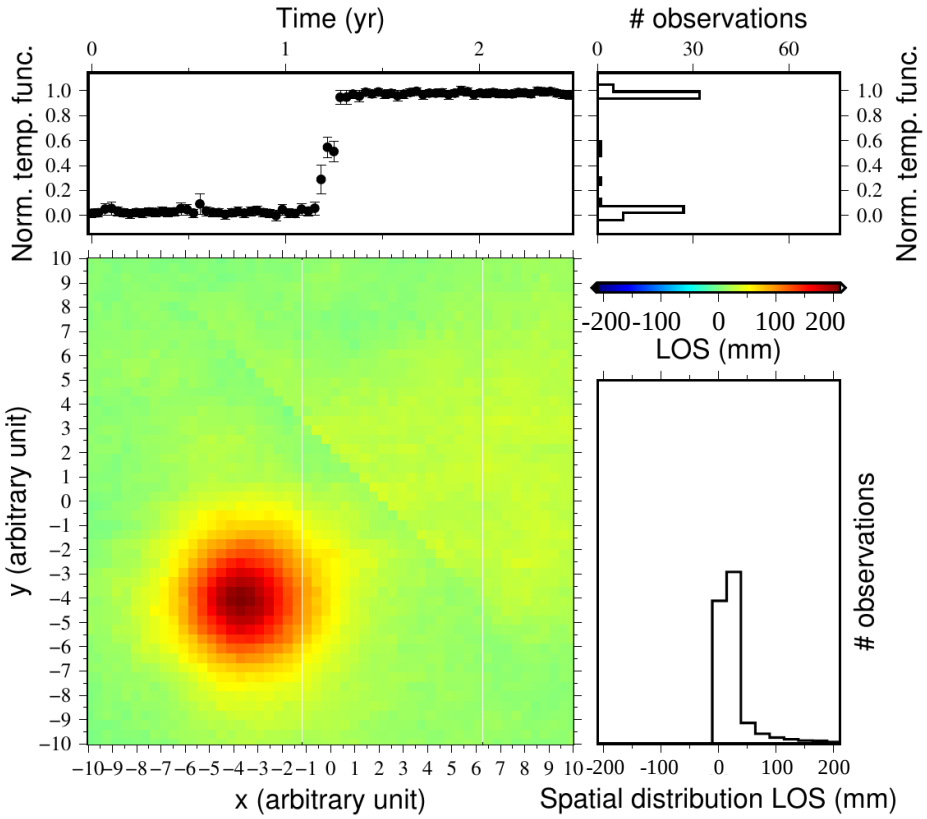




Synthetic #2: T-mode

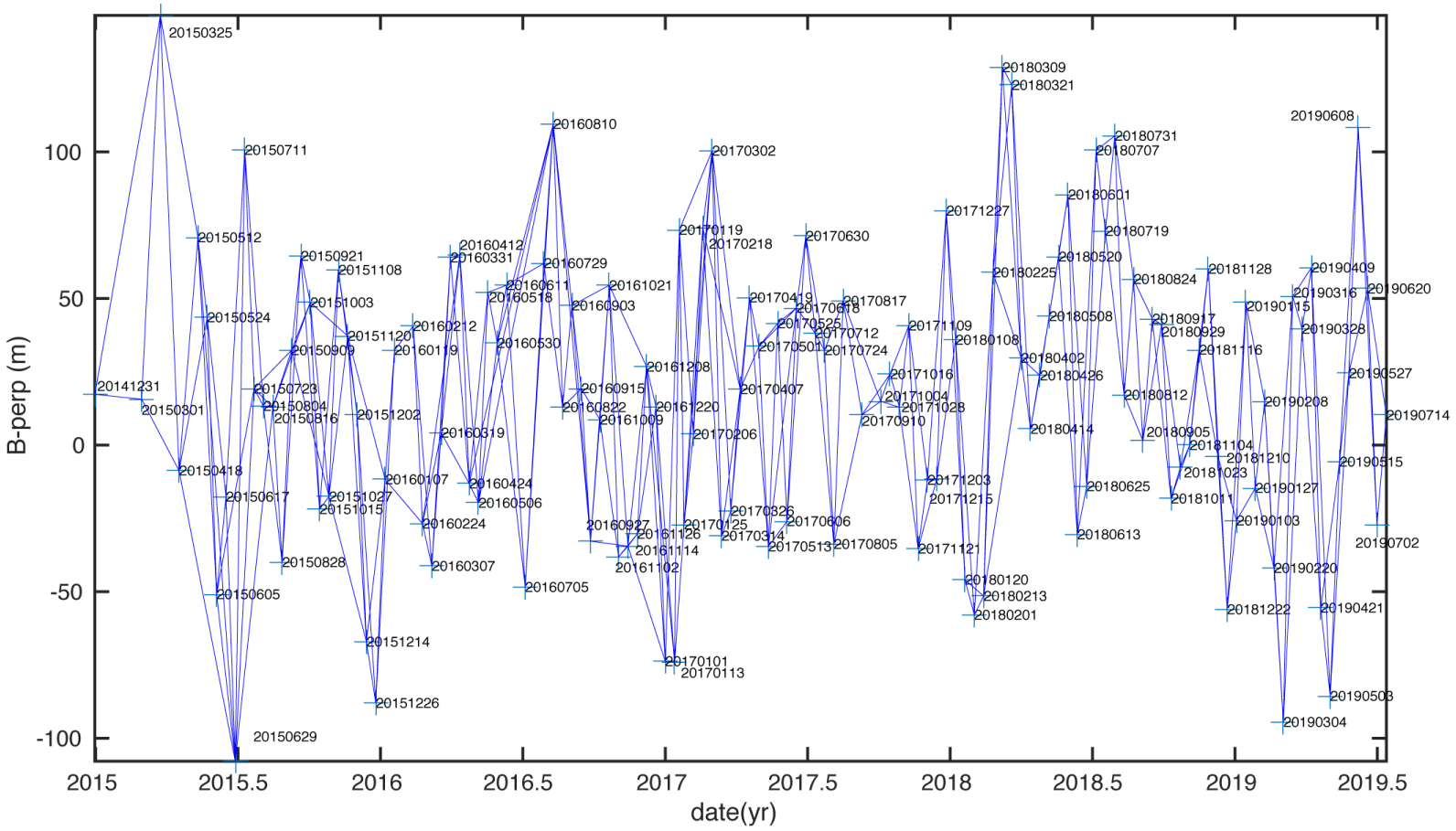
vbICA





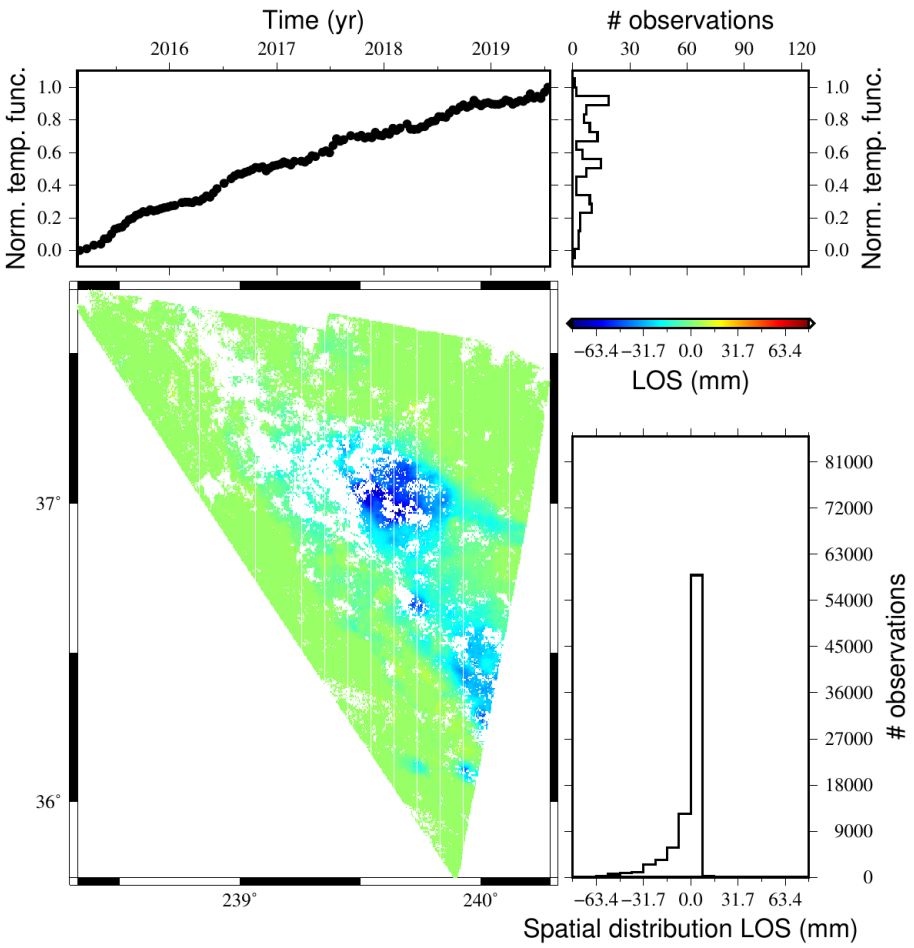
Real case scenarios

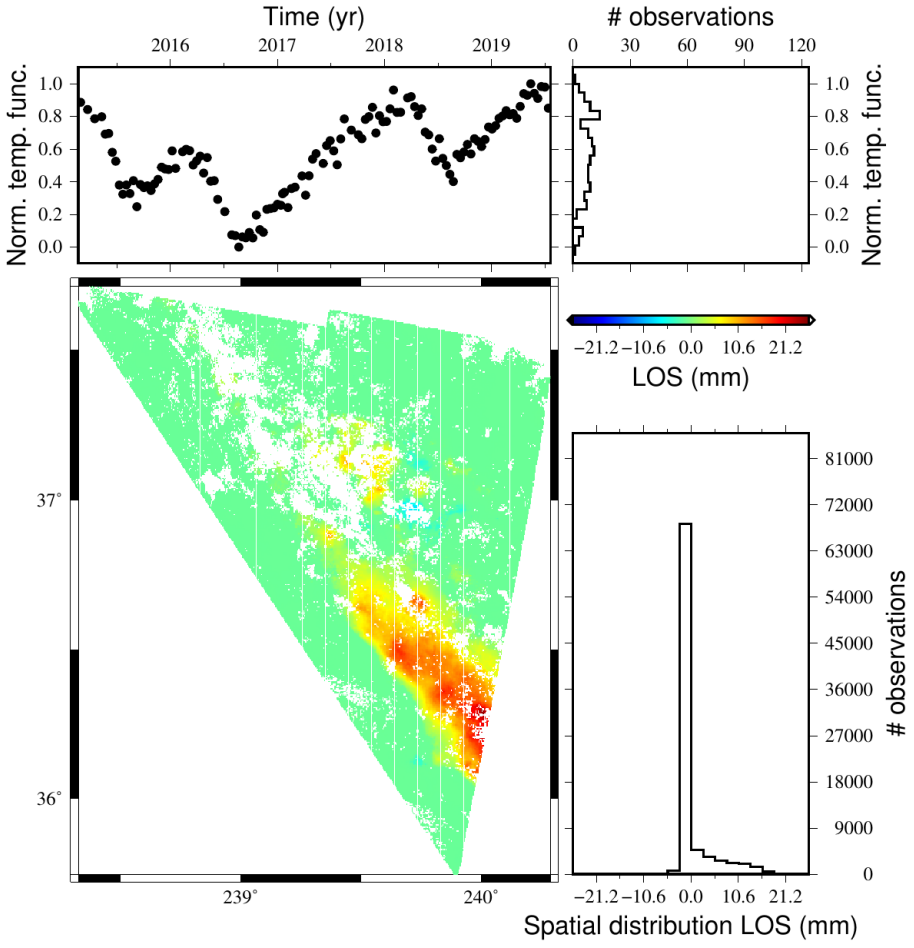
Satellite	Orbit direction	Track	Frame	# of SAR images	# of interferometry pairs	Time span	Incidence angle
S-1	descending	42	469,474	129	287	2014/12/31 2019/07/14	30.7°-45.9°



Central Valley

vbICA

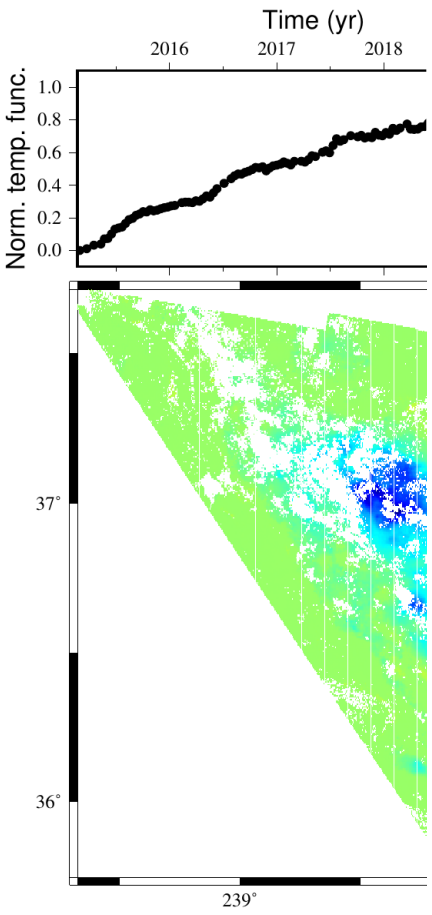




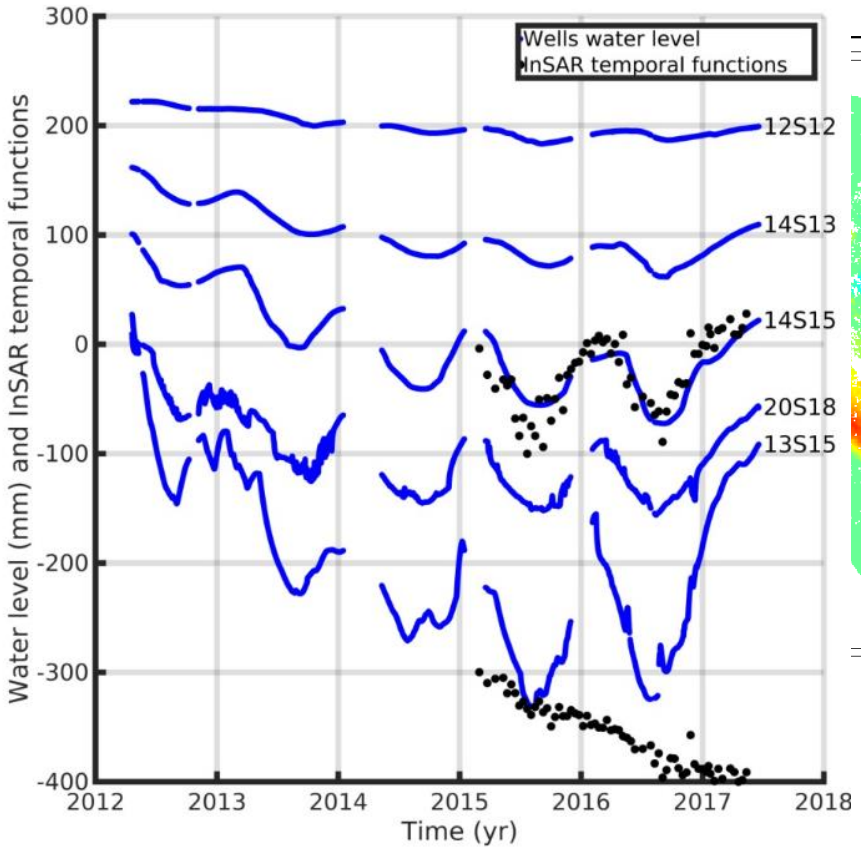
Central Valley

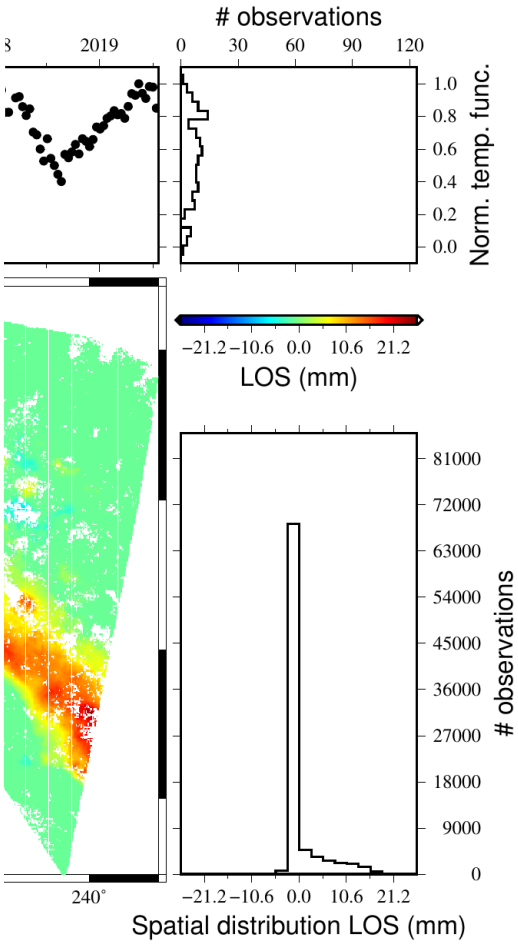
vbICA

separation eases estimate of local storage coefficients



Inelastic



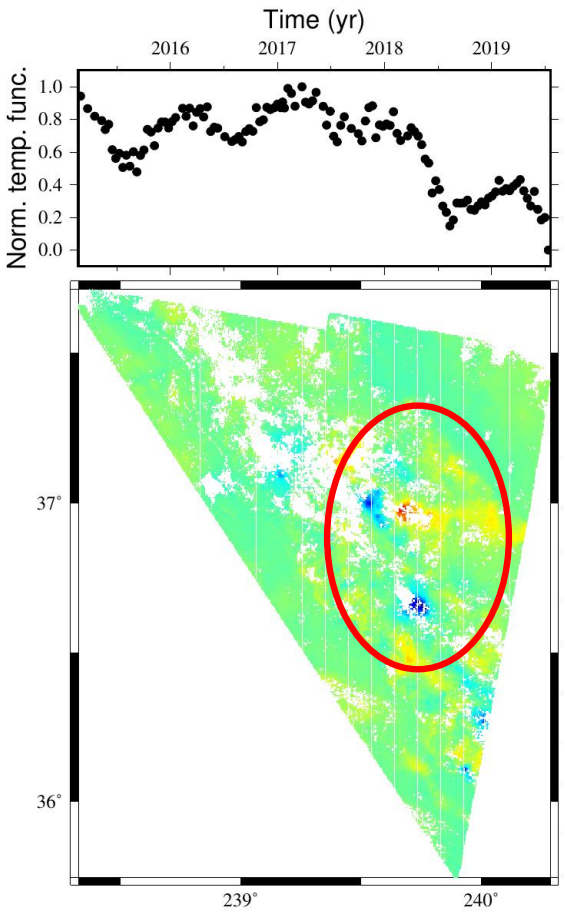


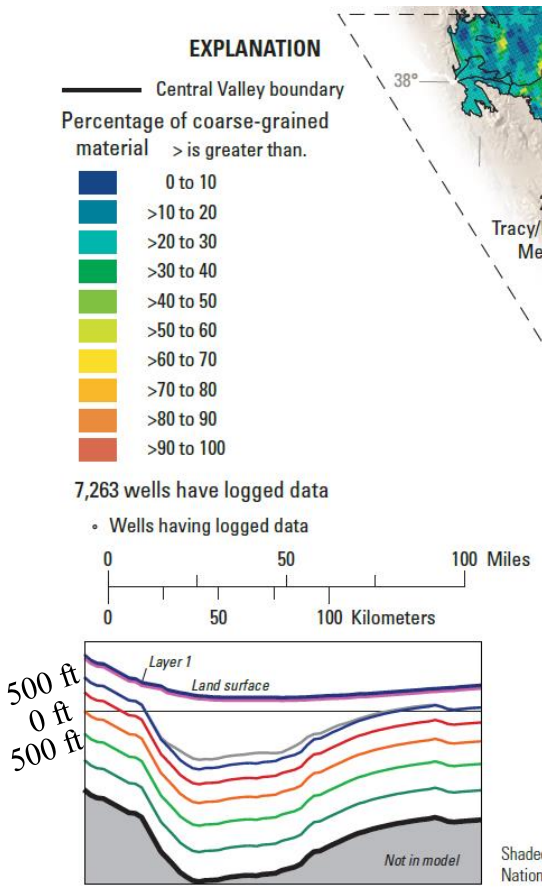
Elastic

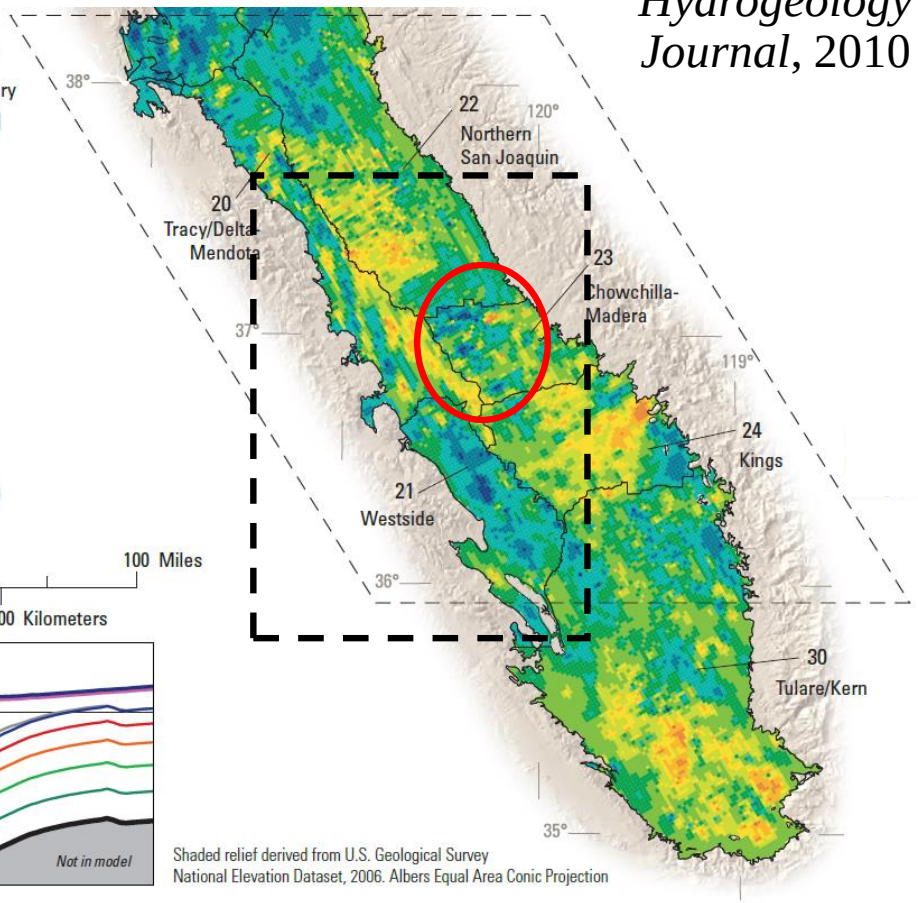
Central Valley

vbICA

Faunt et al., *Hydrogeology Journal*, 2010





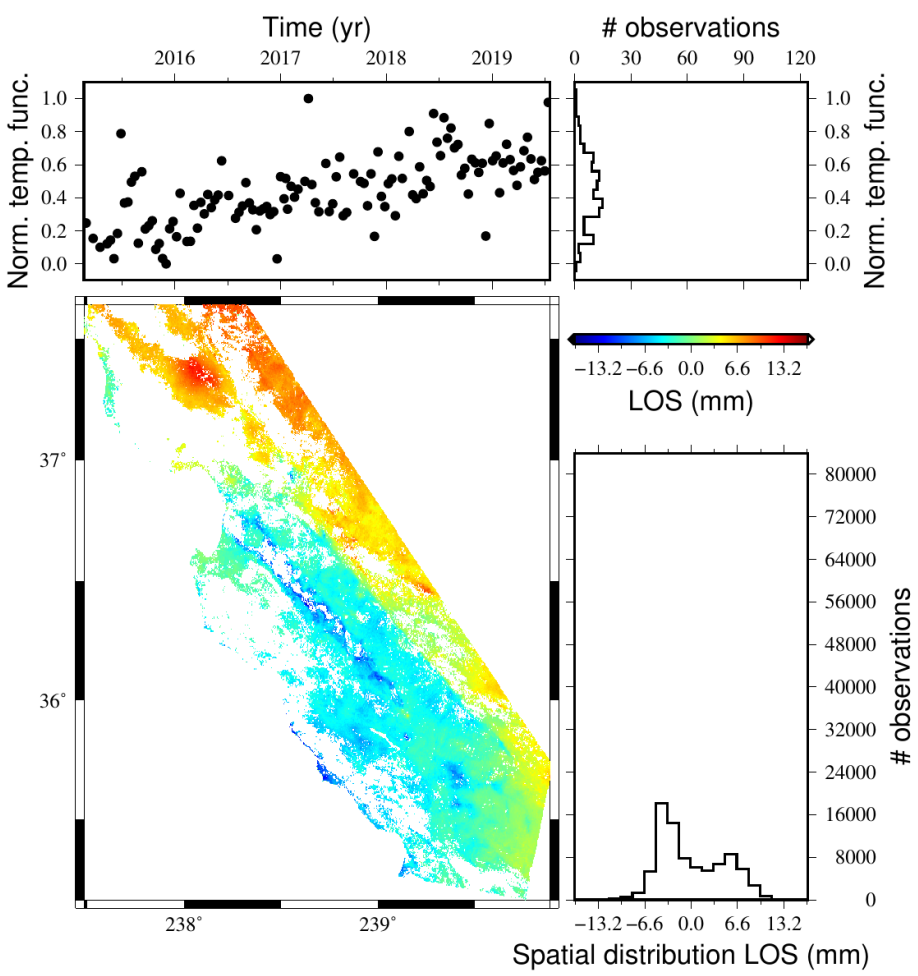


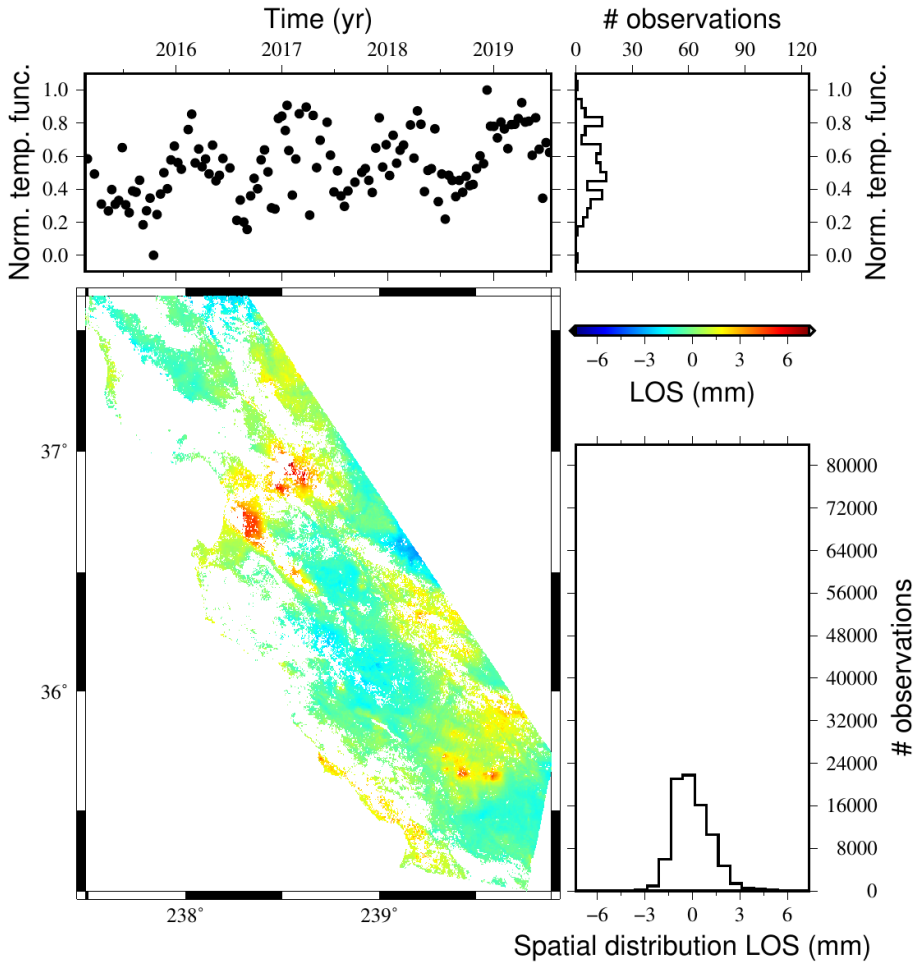
Elastic shallow response?

Low % of coarse grained material at <10 m

Central San Andreas Fault

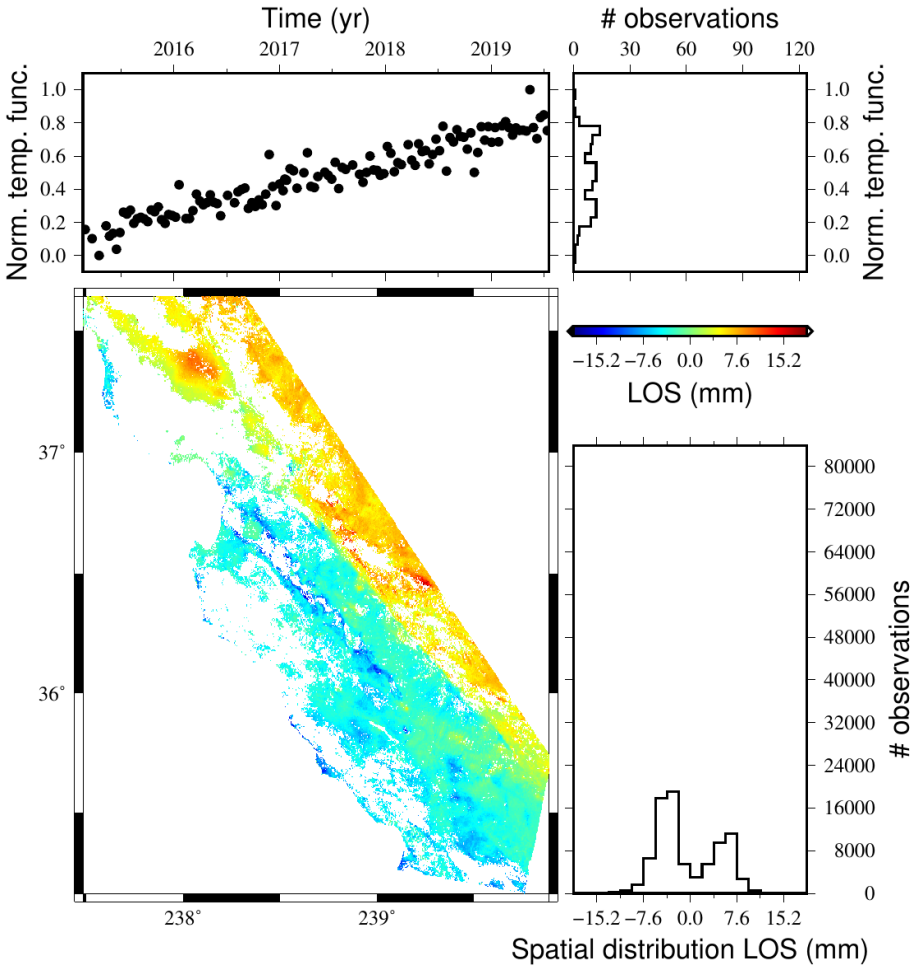
FastICA

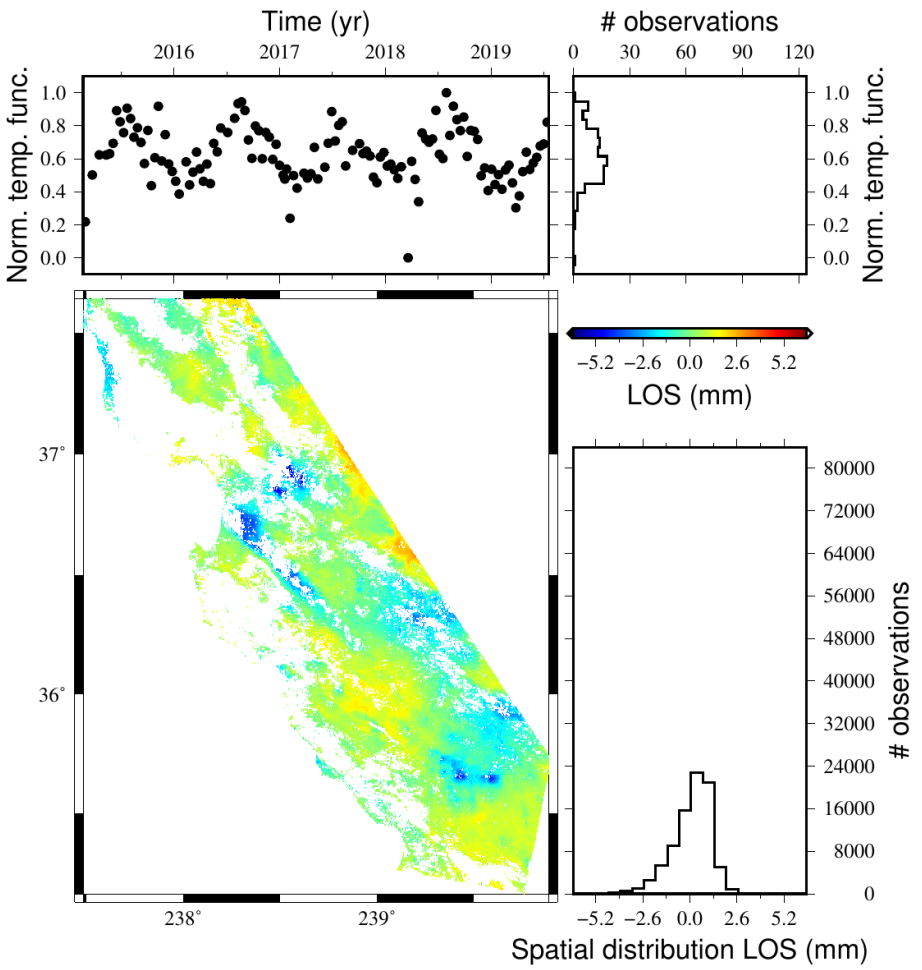




Central San Andreas Fault

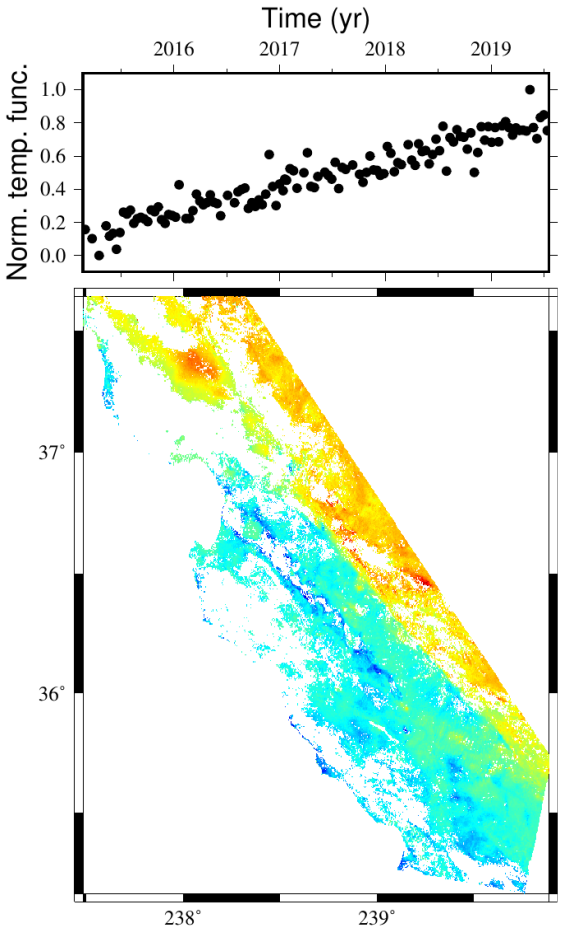
vbICA

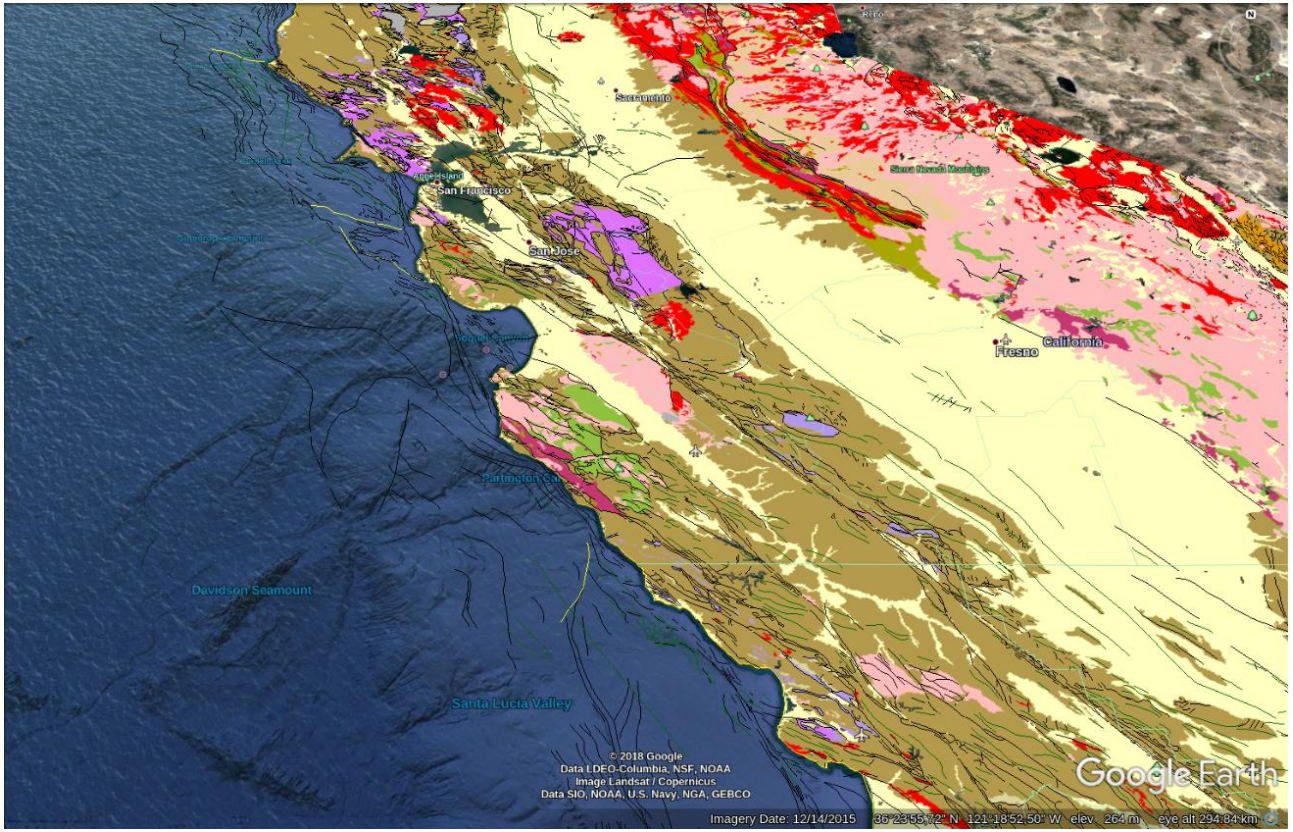




Central San Andreas Fault

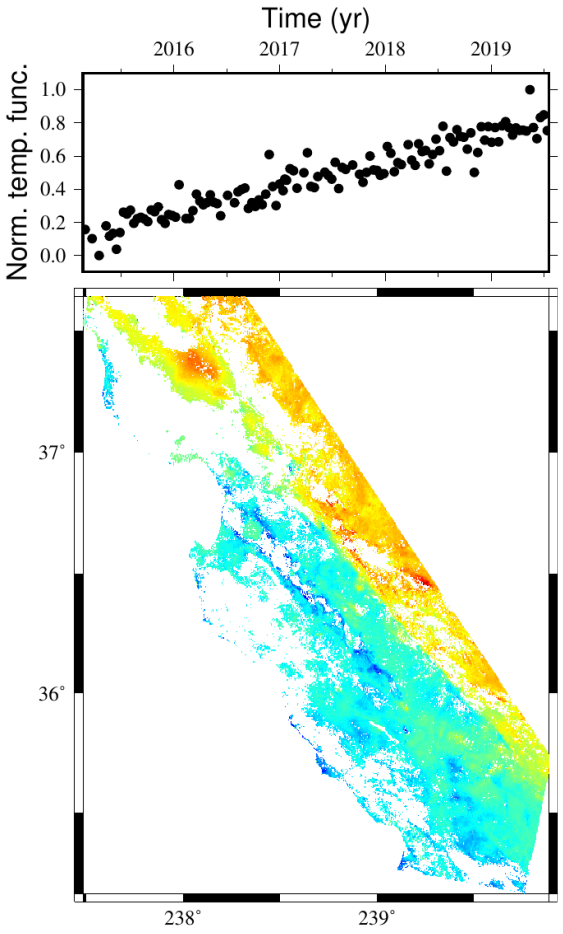
vbICA

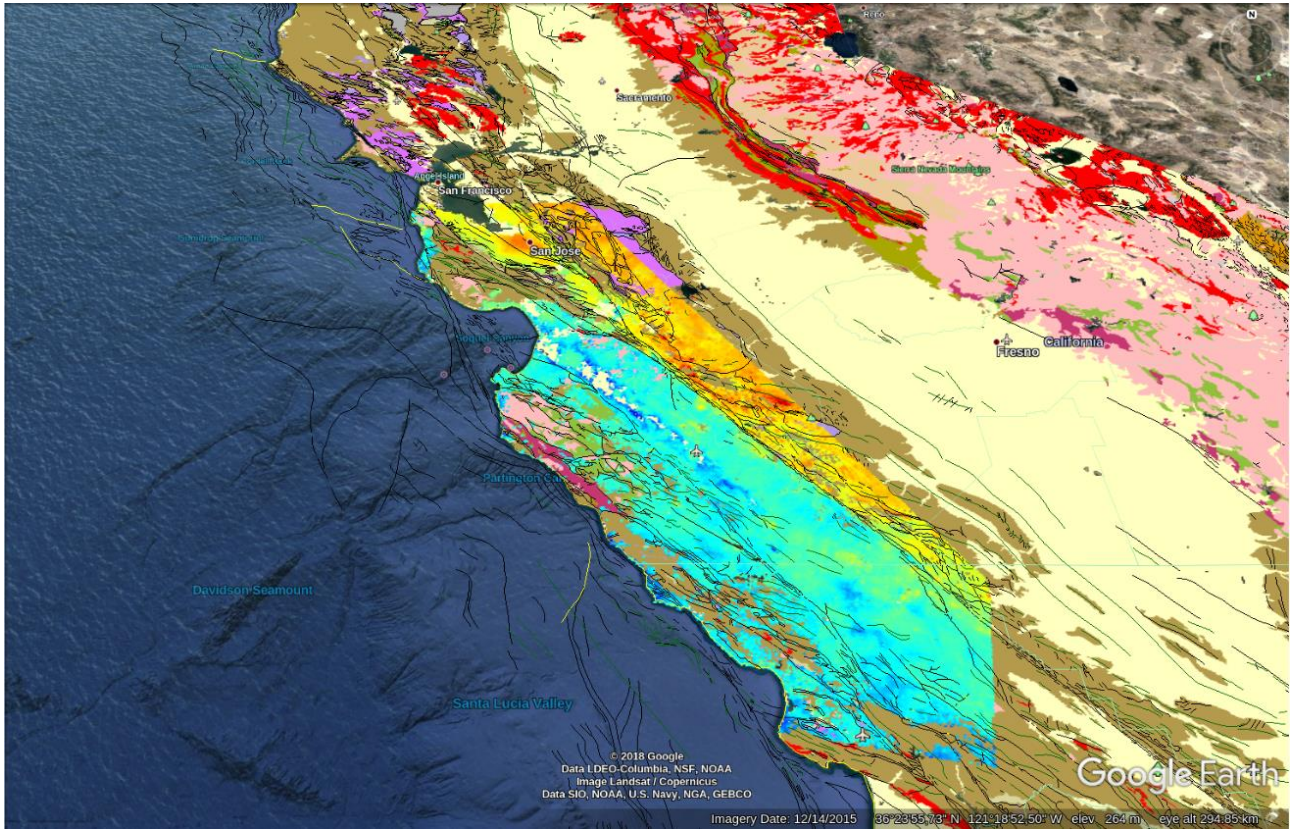




Central San Andreas Fault

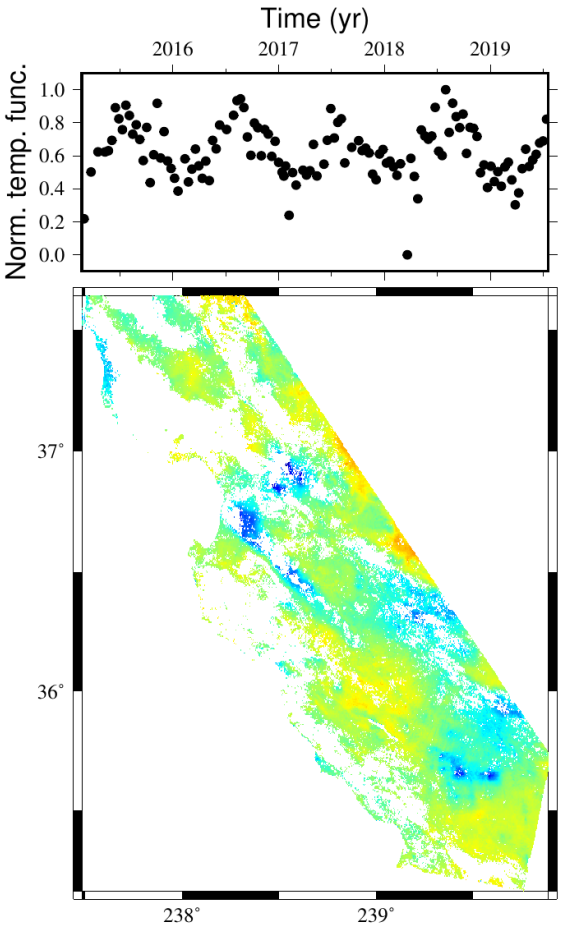
vbICA

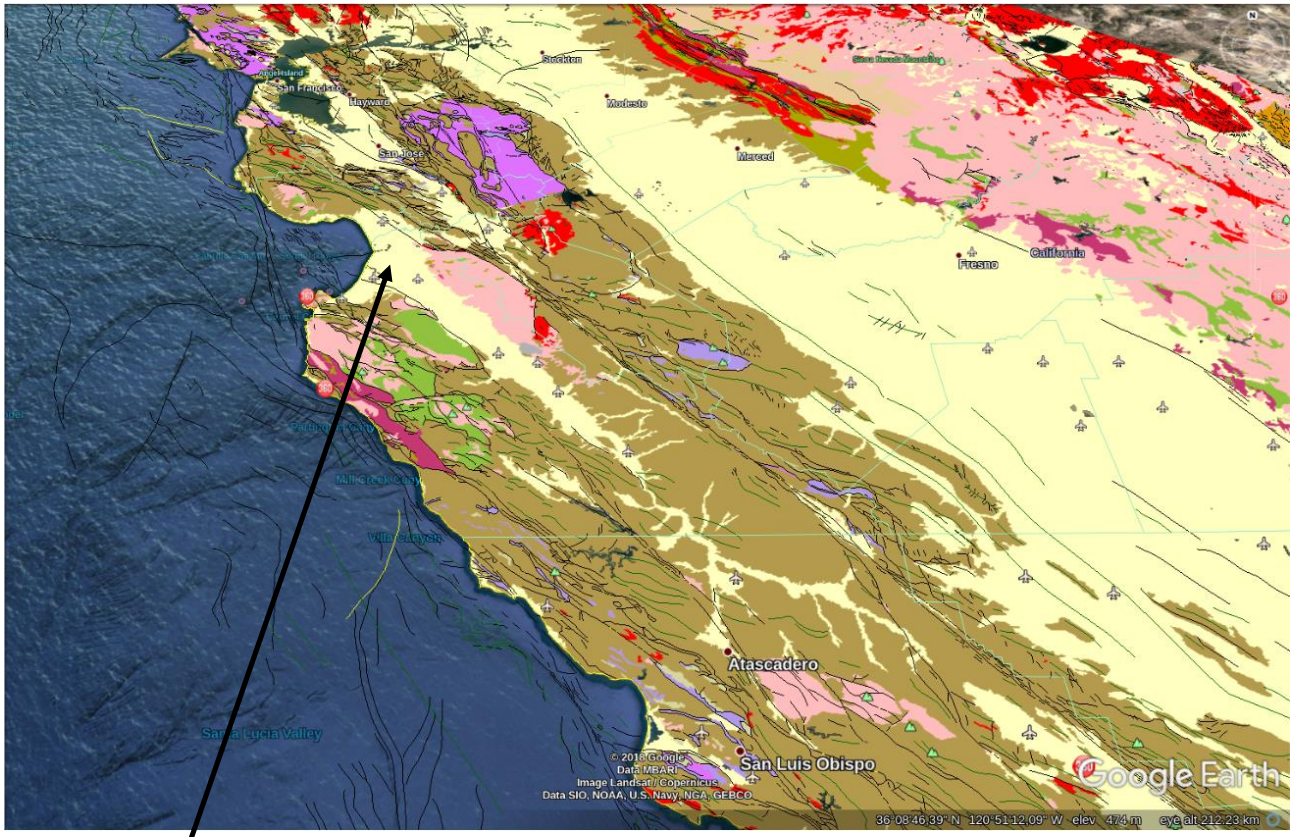




Central San Andreas Fault

vbICA

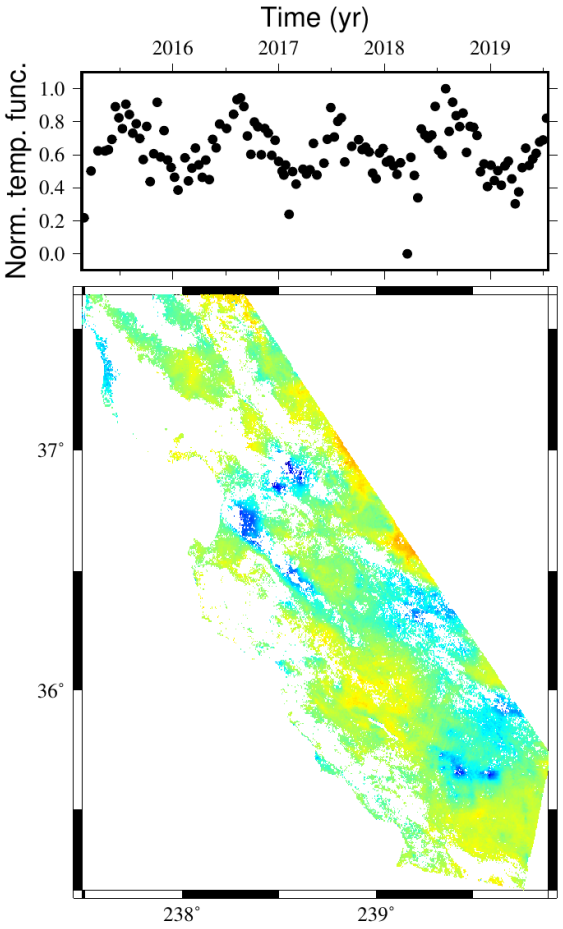


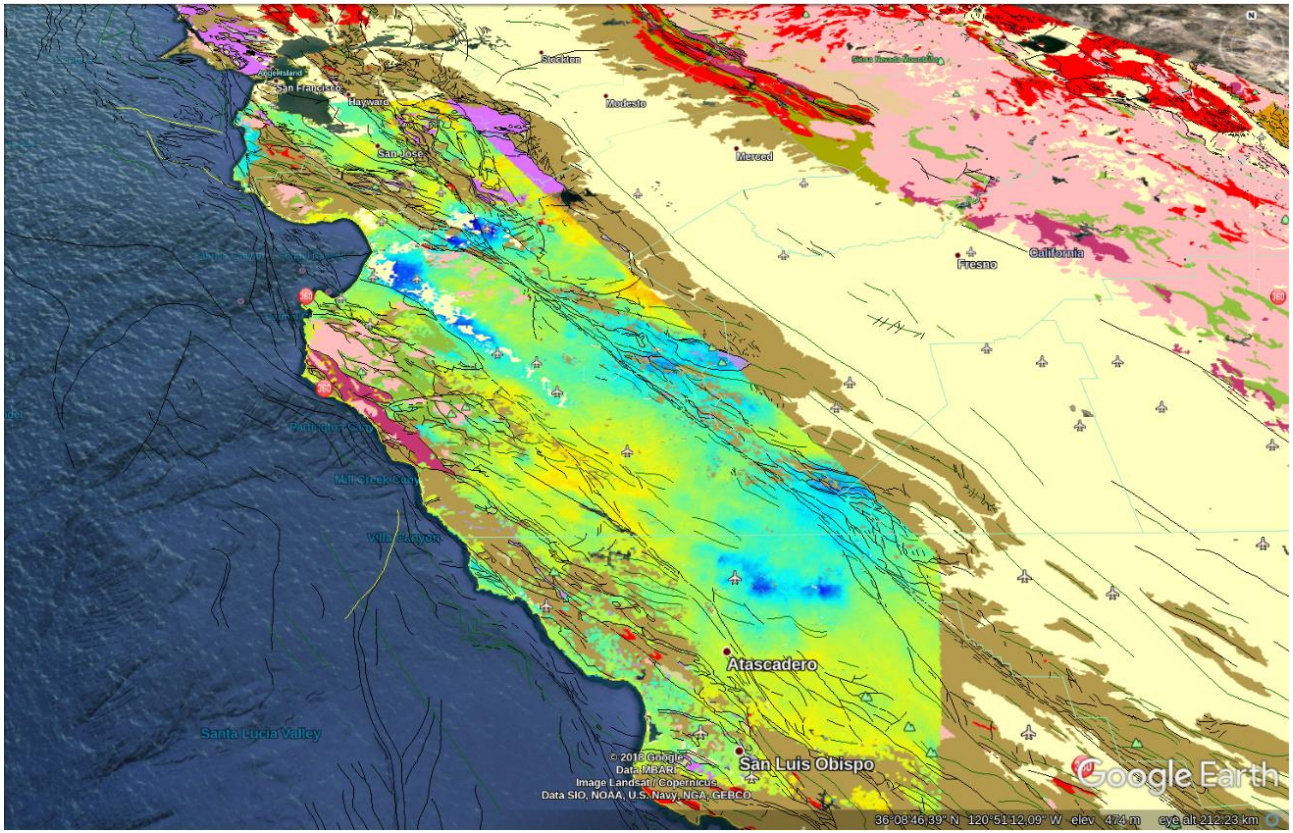


Quaternary alluvial basin

Central San Andreas Fault

vbICA





Conclusions

- vbICA outperforms FastICA for InSAR-like time series analysis

- Not always S-mode works best: if sources are separated in time, T-mode seems to work better

- Whitening pre-processing is crucial to be able to separate sources in our synthetic tests

- Separation of elastic and inelastic responses, and potentially of deep and shallow responses, in Central Valley

- Better characterization of tectonic signal on Central San Andreas Fault and seasonal matches quaternary alluvial basin

Future work

- Attempt automatization of the procedure

adriano.geolandi@gmail.com