

# Journal of Geophysical Research: Solid Earth

## RESEARCH ARTICLE

10.1029/2018JB016210

### Key Points:

- Independent component analysis (ICA) is the best of a suite of blind signal separation methods for use with InSAR data
- Independent component analysis can be used to automatically isolate signals of geophysical interest during the 2015 Wolf Volcano eruption
- Through monitoring the changes in the signals recovered by ICA, signs of unrest at a volcano can be automatically detected

### Supporting Information:

- Supporting Information S1

### Correspondence to:

M. E. Gaddes,  
 eemeg@leeds.ac.uk

### Citation:

Gaddes, M. E., Hooper, A.,  
 Bagnardi, M., Inman, H., & Albino, F.  
 (2018). Blind signal separation  
 methods for InSAR: The potential to  
 automatically detect and monitor  
 signals of volcanic deformation.  
*Journal of Geophysical Research:  
 Solid Earth*, 123, 10,226–10,251.  
<https://doi.org/10.1029/2018JB016210>

Received 11 JUN 2018

Accepted 23 OCT 2018

Accepted article online 1 NOV 2018

Published online 16 NOV 2018

## Blind Signal Separation Methods for InSAR: The Potential to Automatically Detect and Monitor Signals of Volcanic Deformation

M. E. Gaddes<sup>1</sup> , A. Hooper<sup>1</sup> , M. Bagnardi<sup>2,3</sup> , H. Inman<sup>4,5</sup>, and F. Albino<sup>6</sup> 

<sup>1</sup>COMET, University of Leeds, Leeds, UK, <sup>2</sup>Formerly at COMET, University of Leeds, Leeds, UK, <sup>3</sup>Now at Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA, USA, <sup>4</sup>Formerly at School of Earth and Environment, University of Leeds, Leeds, UK, <sup>5</sup>Now at Shearwater Geoservices, Pasadena, CA, USA, <sup>6</sup>School of Earth Sciences, University of Bristol, Bristol, UK

**Abstract** There are some 1,500 volcanoes with the potential to erupt, but most are not instrumentally monitored. However, routine acquisition by the Sentinel-1 satellites now fulfills the requirements needed for interferometric synthetic aperture radar (InSAR) to progress from a retrospective analysis tool to one used for near-real-time monitoring globally. However, global monitoring produces vast quantities of data, and consequently, an automatic detection algorithm is therefore required that is able to identify signs of new deformation, or changes in rate, in a time series of interferograms. On the basis that much of the signal contained in a time series of interferograms can be considered as a linear mixture of several latent sources, we explore the use of blind source separation methods to address this issue. We consider principal component analysis and independent component analysis (ICA) which have previously been applied to InSAR data and nonnegative matrix factorization which has not. Our systematic analysis of the three methods shows ICA to be best suited for most applications with InSAR data. However, care must be taken in the dimension reduction step of ICA not to remove important smaller magnitude signals. We apply ICA to the 2015 Wolf Volcano eruption (Galapagos Archipelago, Ecuador) and automatically isolate three signals, which are broadly similar to those manually identified in other studies. Finally, we develop a prototype detection algorithm based on ICA to identify the onset of the eruption.

**Plain Language Summary** Through monitoring volcanoes, we can detect signs that they may be likely to erupt in the future, yet many of the world's volcanoes remain unmonitored. However, the latest generation of RADAR satellites produces regular measurements of ground displacement around the majority of the world's volcanoes, and these measurements can be used to detect signs of unrest at many previously unmonitored volcanoes. As the volume of data produced in this undertaking is vast, an automatic detection algorithm is required to detect these signals of interest. In this work, we present methods to separate signals indicative of volcanic unrest from other nuisance signals and to then monitor the geophysically important signals to detect if a volcano has entered a period of unrest.

### 1. Introduction

The world's ~1,500 subaerial active volcanoes (Siebert & Simkin, 2013) pose a diverse set of geophysical hazards, which range from those such as pyroclastic flows, which impact people living near the volcano, to those such as ash clouds, which can impact people across the globe. Data gathered from methods or instruments such as seismology, tiltmeters, strain gauges, leveling lines, satellite-derived thermal imagery, satellite-derived gas emissions, satellite-derived deformation measurements, and geochemical analysis of fumarolic gases can be used to identify the signs of volcanic unrest that are indicative of a certain hazard (Sparks et al., 2012, and references therein). However, the majority of the world's volcanoes are not routinely monitored (Sparks et al., 2012), and the vast majority of the methods previously mentioned would be prohibitively expensive to extend to all volcanoes, due to the requirement for more instruments to be acquired and deployed.

The ability of interferometric synthetic aperture radar (InSAR) to measure ground displacements allows deformation of both a volcano's flanks and the area surrounding it to be constrained (e.g., Pinel et al., 2014, and references therein), and these measurements of displacement can be a valuable indicator of volcanic unrest that may lead to an eruption (Biggs et al., 2014; Ebmeier et al., 2018). Therefore, a satellite with a suitable

acquisition strategy could provide the measurements that would allow monitoring of all of the world's subaerial volcanoes.

Until recently, InSAR was not suited to real-time global monitoring as synthetic aperture radar (SAR) satellites in operation before 2015 rarely routinely acquired images over areas of geophysical interest, and data were generally not available until days after it was acquired. However, the European Space Agency's most recent SAR satellites (the Sentinel-1 constellation) have addressed the limitations previously listed and have allowed InSAR to evolve from a retrospective analysis tool into one that can be used for near-real-time monitoring. Other advances that make this possible include the construction of new processing facilities to automatically and rapidly create interferograms for large parts of the globe (e.g., González et al., 2016; Meyer et al., 2016) and new methods for rapidly updating a time series when new interferograms are created (e.g., Spaans & Hooper, 2016).

However, using data derived from the Sentinel-1 satellites to monitor the world's subaerial active volcanoes produces too many interferograms for them to be manually inspected for signs of unrest. Consequently, we present results of work to construct an algorithm to automatically detect signs of unrest in a time series of interferograms over a given volcano. To avoid the time-consuming nature of labeling data (e.g., manually inspecting interferograms and assigning information such as *contains deformation*), our algorithm must be an unsupervised one (i.e., one that is trained on unlabeled data). One class of unsupervised learning algorithms is termed blind signal separation (BSS) methods, and one of these (independent component analysis [ICA]) has been utilized to isolate signals of geophysical interest in an InSAR time series by Ebmeier (2016). Therefore, we aim to construct an algorithm based on the BSS premise that latent sources of interest can be extracted from a time series of interferograms.

However, as ICA is one of a suite of BSS methods, we endeavor to ascertain which of the methods is best suited to InSAR data. We do this by first introducing the fundamentals of several BSS techniques (section 2), explaining our novel method for applying a new BSS method to InSAR data (section 3), and comparing the results of applying several BSS methods to a synthetic data set (section 4). In section 5 we validate that the chosen method works with real data through the use of two time series of Sentinel-1 interferograms, before describing and demonstrating our prototype automatic detection algorithm in section 6.

## 2. BSS

### 2.1. Interferograms as Mixtures of Signals

Interferograms consist of measurements of amplitude and phase at pixel locations. After correcting for geometric terms, the phase consists of contributions from several sources:

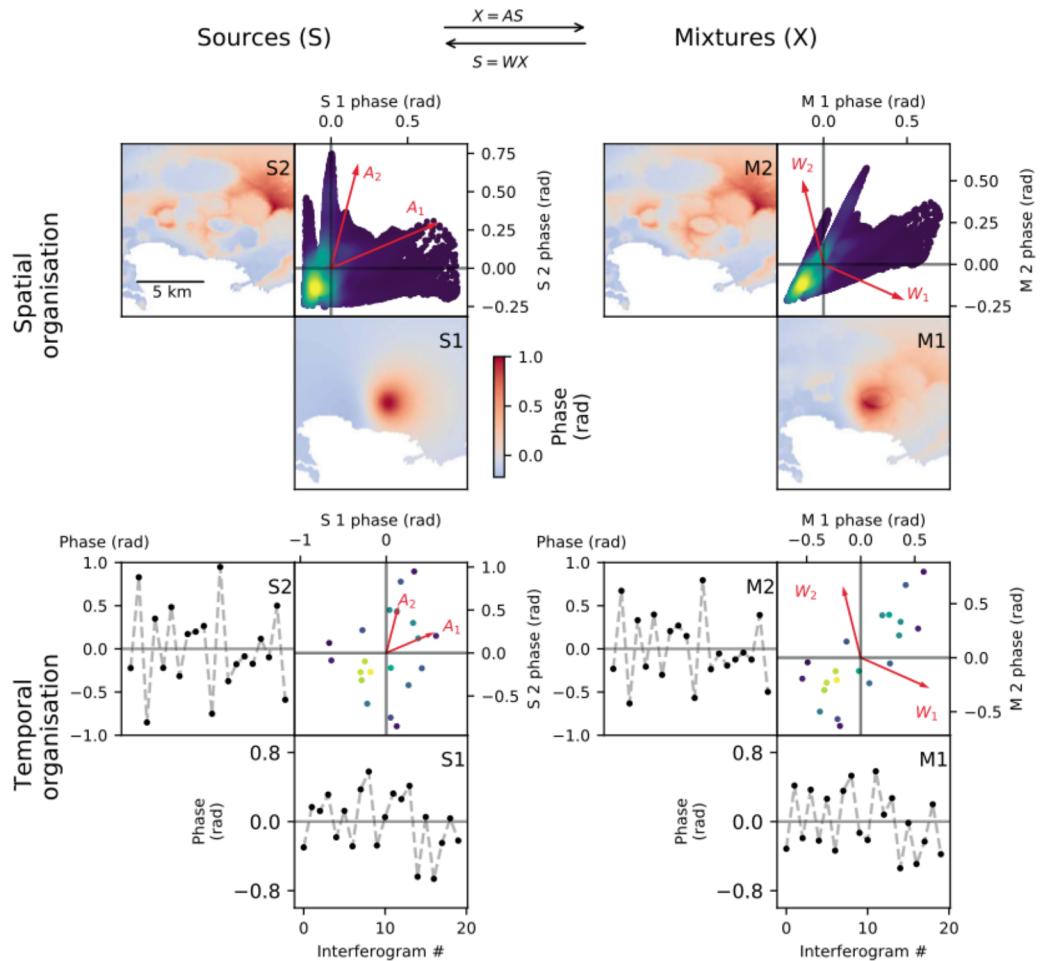
$$\varphi = W \{ \varphi_{\text{def}} + \varphi_{\text{orb}} + \varphi_{\text{atm}} + \Delta\varphi_{\theta} + \varphi_N \} \quad (1)$$

where  $\varphi_{\text{def}}$  is the phase change due to deformation of the ground surface,  $\varphi_{\text{orb}}$  is the phase due to errors in the location of the satellite at each acquisition,  $\varphi_{\text{atm}}$  is the phase change due to changes in the atmospheric delay,  $\Delta\varphi_{\theta}$  is the phase due to misestimation of the look angle,  $\varphi_N$  is the phase noise, and  $W$  is a wrapping operator that results in the phase lying between  $-\pi$  and  $\pi$  (Hooper et al., 2012). In geophysical applications  $\varphi_{\text{def}}$  is usually the signal of interest and a suite of methods exist to reduce the contributions from other terms. Signals that are considered to be dominated by deformation have been attributed to a variety of volcanic processes, including preeruptive inflation of a magma chamber, subsidence due to flank loading by new material, subsidence due to cooling of a magma body below a volcano, and subsidence due to changes in a volcano's geothermal system (Ebmeier et al., 2018).

If multiple latent signals combine in unknown quantities to form an interferogram, recovering the original signals can be viewed as a BSS problem (Jutten & Herault, 1991). In this class of problems, the observed mixed signals are assumed to be generated using the following mixing model:

$$\mathbf{X} = \mathbf{AS} \quad (2)$$

where using the standard nomenclature of BSS literature,  $\mathbf{X}$  contains the mixtures as row vectors,  $\mathbf{S}$  contains the unknown sources as row vectors, and  $\mathbf{A}$  is the unknown mixing matrix that combines varying amounts of the sources to create each mixture. In this work, we refer to scalars as lower case letters (e.g.,  $a$ ), row vectors



**Figure 1.** Linear mixing for spatially and temporally organized data. (top row) Spatial organization, in which a deformation signal ( $S_1$ ) is mixed with a topographically correlated signal ( $S_2$ , generated from the topography of Campi Flegrei, Italy) through reprojecting the data in the directions  $A_1$  and  $A_2$ , to produce two mixtures,  $M_1$  and  $M_2$ . The sources can be recovered from the mixtures by reprojecting the data in the directions  $W_1$  and  $W_2$ . In this architecture, the mixture space has as many dimensions as there are interferograms and as many points in the space as there are pixels in the interferograms. (bottom row) Temporal organization, in which a pixel with a phase change due to deformation over 20 epochs ( $S_1$ ) is mixed with a pixel with a phase change due to a topographically correlated atmospheric signal ( $S_2$ ) in a similar manner to the spatial case to produce two pixels ( $M_1$  and  $M_2$ ) that exhibit both deformation and atmospheric signals. In this architecture, the mixture space has as many dimensions as the interferograms have pixels and as many points in this space as there are interferograms. Progression from blue to yellow is used to indicate areas of high point density in these (and any following) scatter plots.

as bold case lower letters (e.g.,  $\mathbf{a}$ ), and matrices as bold upper case letters (e.g.,  $\mathbf{A}$ ). The sources,  $\mathbf{S}$ , can be recovered if we could calculate the unmixing matrix,  $\mathbf{W}$

$$\mathbf{S} = \mathbf{W}\mathbf{X} \quad (3)$$

where

$$\mathbf{A} = \mathbf{W}^{-1} \quad (4)$$

The preceding description of linear mixing can also be expressed in terms of Euclidean geometry and is demonstrated in Figure 1. The  $m$  variables measured at  $n$  time points can be considered as points in an  $mD$  space (e.g., two  $10^5$  pixel interferograms would be  $10^5$  points in a 2-D space). If the sources are stored in a data matrix ( $\mathbf{S}$ ) with each row containing a new variable, each column of this matrix is a  $mD$  position vector determining that observations point in the space. Matrix multiplication of these sources ( $\mathbf{S}$ ) by a mixing matrix ( $\mathbf{A}$ ) is equivalent to the inner product of the row vectors of  $\mathbf{A}$  and the column vectors of  $\mathbf{S}$ .

Consequently, the row vectors of  $\mathbf{A}$  can be considered as the axes defining a new subspace that the data ( $\mathbf{S}$ ) are projected into. The unmixing process ( $\mathbf{S} = \mathbf{WX}$ ) can be considered in a similar way, with the rows of  $\mathbf{W}$  containing the basis vectors required to recover the sources. The goal of BSS is to find the basis vectors required to recover each source from the mixtures (i.e., the rows of  $\mathbf{W}$ ). However, it must be noted that the sign of the sources that are recovered by BSS methods (the rows of  $\mathbf{S}$ ) remains ambiguous, as the opposing sign may be present in the column of the mixing matrix that controls the strength of a given source in each mixture.

Linear mixing can be complicated by differences in the number of mixtures relative to the number of latent sources. In the simplest case, the number of mixtures is equal to the number of latent sources and the mixing and unmixing matrices ( $\mathbf{A}$  and  $\mathbf{W}$ ) are square. However, in what is termed the overcomplete case (underdetermined in inverse theory terminology) there are more latent sources than mixtures, and in the undercomplete case there are more mixtures than latent sources (Amari, 1999). A time series of InSAR data at a subaerial volcano is likely to consist of tens to hundreds of interferograms but of substantially fewer latent sources (such as deformation and a topographically correlated atmospheric phase screen [APS]), and we therefore expect it to be undercomplete.

In this line of reasoning we do not expect the unique turbulent APS present in each interferogram to be recoverable as sources. An attempt to recover the turbulent APSs would increase the number of sources to more than the number of mixtures (a significant change) and shift the problem from the undercomplete to overcomplete case. The sparse nature of the time course for each turbulent APS and the huge increase in the number of sources that would have to be sought lead us to discount the turbulent APSs as sources and instead treat them as noise terms in the undercomplete case. However, other atmospheric signals such as those due to atmospheric pressure gradients associated with weather systems or the rain shadow effect are not unique, and we may therefore expect to recover them as components. Methods to correct for APSs are also routinely applied to InSAR data, but as these can introduce new and erroneous signals, we apply BSS to uncorrected time series.

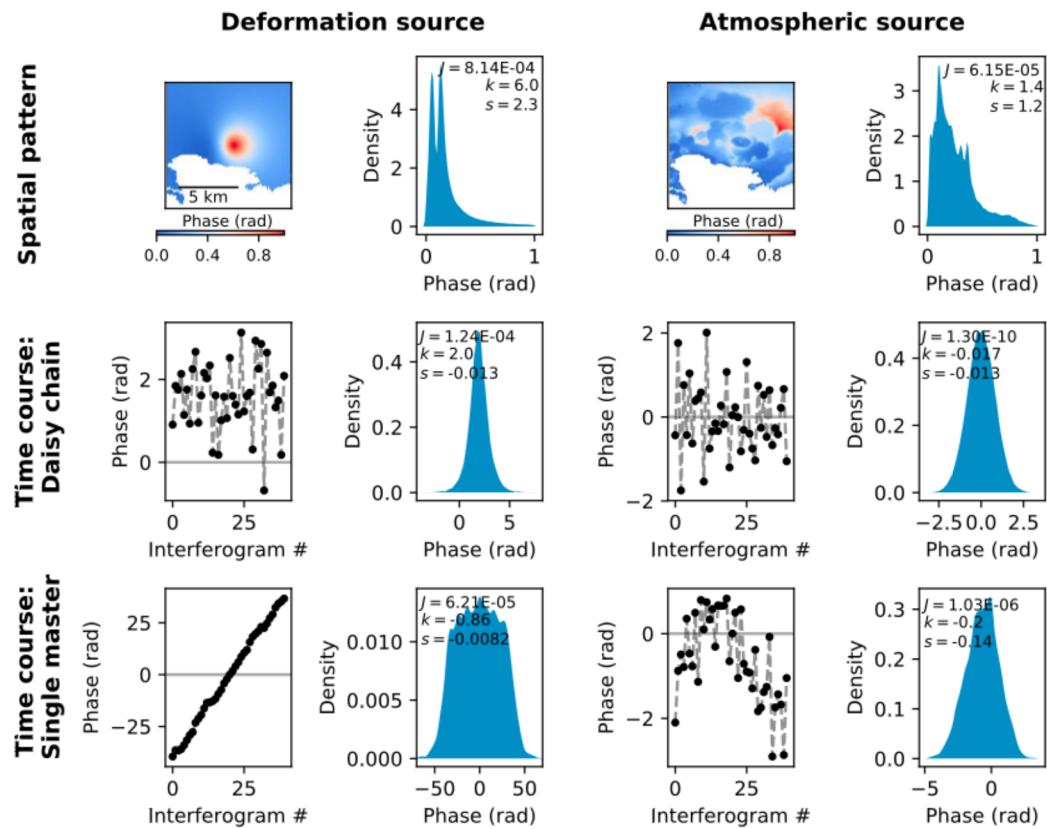
BSS for the undercomplete case requires extra processing when certain algorithms are used, and these caveats are discussed for each method. To implement BSS, we have investigated nonnegative matrix factorization (NMF), principal component analysis (PCA), and ICA. These methods are introduced in the following subsections, before our novel application of NMF to InSAR data is demonstrated in section 3.

## 2.2. Organizing and Synthesizing Data

Before considering how to apply NMF, PCA, and ICA, we must first transform a time series of interferograms into a form that these methods can be applied to. The three methods consider the statistics of multiple observations of several variables and consequently do not require the spatial (or temporal) relationships between pixels (or interferograms) to be conserved. Therefore, the information contained within the time series can be converted to row (or column) vectors, providing that this reorganization is performed consistently. We also refer to the time history of a spatial map (which can be as small as one pixel or more commonly as large as a latent source) as a time course, in the style of BSS literature.

How these row vectors are formed has important implications, and consequently, the two different approaches are termed as architecture I and architecture II in BSS literature (Bartlett et al., 2002). However, this nomenclature is opaque when applied to InSAR data and we instead refer to architecture I as spatial organization, as when ICA is applied to architecture I data, the recovered latent sources are spatially independent. When ICA is applied to architecture II, data the discovered latent sources are temporally independent, so we instead refer to architecture II as temporal organization.

In spatial organization, each image contains multiple realizations of a single random variable, with the number of random variables being equal to the number of images, while in temporal organization, each pixel is a random variable, with as many observations as there are images. In this work, we adhere to the conventions of BSS literature and place variables as rows in our data matrix, and each observation of these variables occupies a new column. Therefore, for a time series of  $t$  interferograms each of  $p$  pixels, our data matrix would be  $t \times p$  for spatial organization and  $p \times t$  for temporal organization. In Euclidean space, a time series organized spatially is likely to consist of  $10^1\text{--}10^2$  interferograms and requires a space of equal dimensions, while for temporal organization, an interferogram is likely to consist of  $10^3\text{--}10^7$  pixels and requires a space of very high dimensionality. Figure 1 demonstrates the difference of spatial and temporal organization, and how the (un)mixing vectors described in the previous section can be interpreted. As the number of interferograms

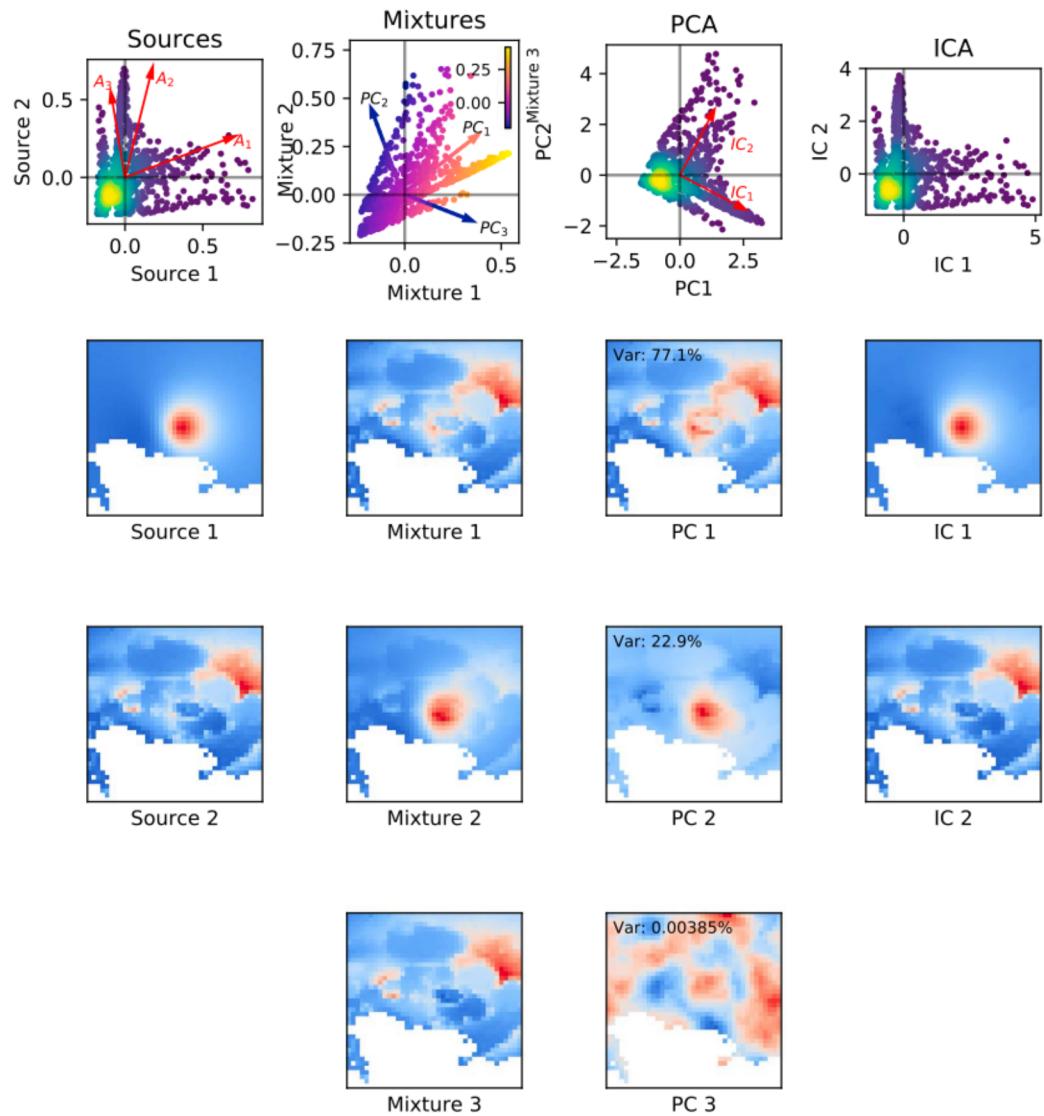


**Figure 2.** Sources and estimated probability density function (PDFs) for synthetic deformation source (left) and atmospheric source (right). (top row) The spatial pattern of the two sources, corresponding estimated PDF (Gaussian kernel density estimate) and measures of non-Gaussianity (introduced in section 2.4) of sources, where  $J$  is approximate negentropy (0 for Gaussian),  $k$  the excess kurtosis (0 for Gaussian), and  $s$  the skewness (0 for Gaussian). Campi Flegrei was used for this synthetic example, with the Gulf of Pozzuoli visible as the masked (white) pixels in the lower part of the image. The spatial pattern of the deformation source is highly non-Gaussian for all measures, while the atmospheric source is closer to Gaussian. (middle row) Temporal nature of the synthetic sources when a daisy chain of interferograms is created. The estimate of the PDF is generated from 100 synthetic time courses each of 40 interferograms, but for clarity only the first time course is shown in the left-hand plots. (bottom row) Temporal nature of the synthetic sources when a single master is used to create the interferograms. The deformation time course appears non-Gaussian as it approximates a uniform distribution, while the atmospheric source is again closer to Gaussian. Note that the non-Gaussian nature of the temporal signals is a result of how they were synthesized and that the temporal trend in the atmospheric source is due to the sinusoidal term described in the text, which is intended to approximate seasonal changes in the signal's strength.

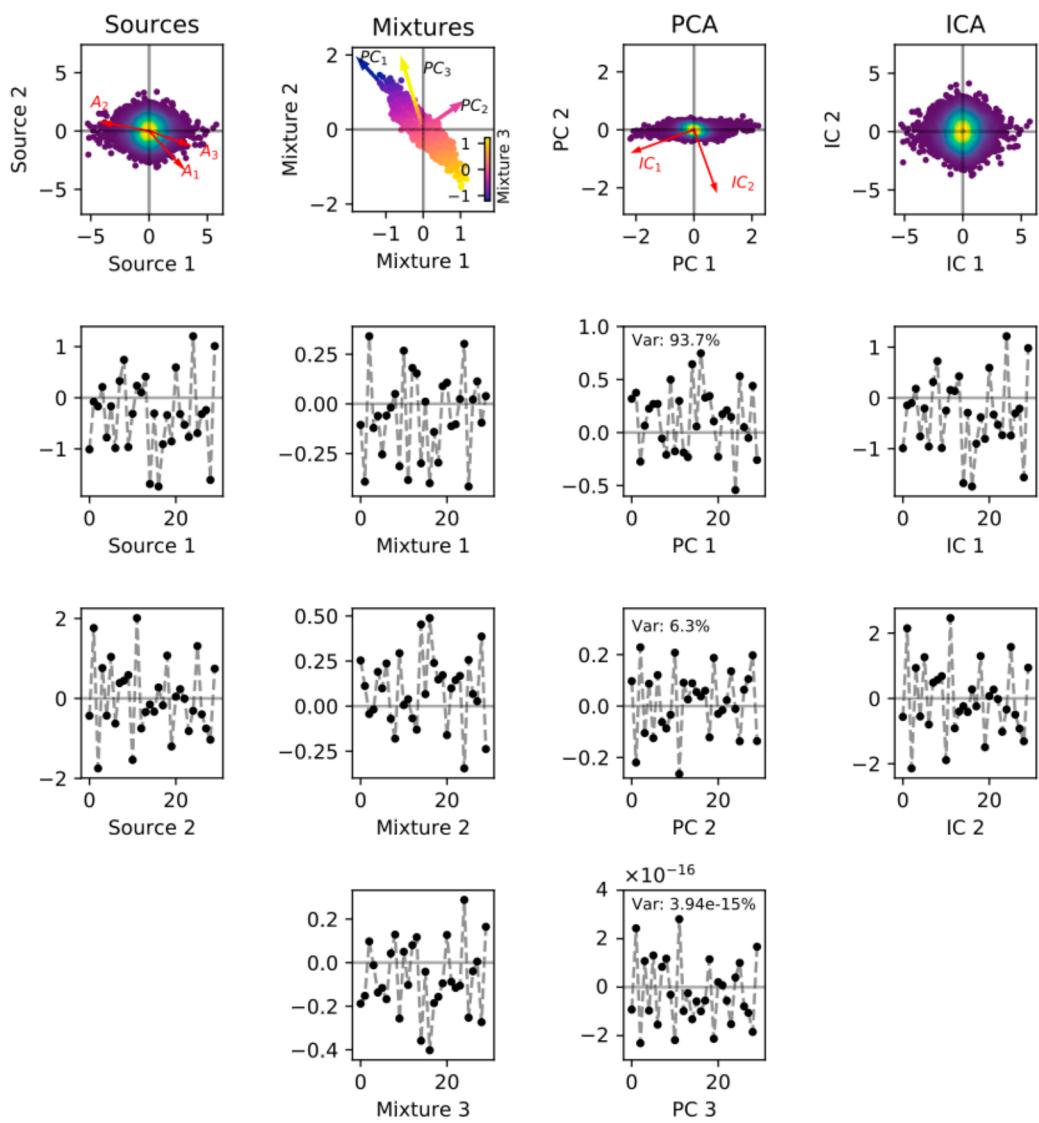
controls the number of data points when the data are organized temporally, a hypothetical time series of only three interferograms would provide a very sparsely populated space in which to perform the analysis and would be equivalent to the scatter plots in Figure 1 having only three data points.

To both introduce and compare PCA, ICA, and NMF, we generate a synthetic time series using equation (2), with one matrix (either  $\mathbf{A}$  or  $\mathbf{S}$ ) containing the spatial patterns of the two synthetic InSAR signals, while the other matrix contains the strengths (termed time courses) with which each spatial pattern contributes to each interferogram. To generate spatially organized data, we postulate that the spatial maps are statistically independent and so place these in  $\mathbf{S}$ , while for temporally organized data we postulate that the time courses are statistically independent and so place these in  $\mathbf{S}$ . We choose our two synthetic InSAR signals to be ground deformation due to a volcanic process, and apparent ground movement due to a topographically correlated APS as separation of signals of this type has been shown to be important (e.g., Delacourt et al., 1998).

The spatial pattern of the two synthetic signals is shown in Figure 2 and, as these are sources when the data are organized spatially, provides insights into the challenge of recovering them. The area corresponds to Campi Flegrei (Italy), with the subaerial caldera walls apparent in the atmospheric signal. The surface deformation is



**Figure 3.** (column 1) A scatter plot comparing the values for each pixel in the two sources (row 1) and the two sources (rows 2 to 3). These are mixed with the three mixing vectors of  $\mathbf{A}$  (plotted in red). As per the previous scatter plots, yellow indicates areas of high point density. (column 2) Scatter plot of the three mixtures (top) and the three mixtures. As the data lie on a plane in the 3-D space, color is used to for the third axis. The three 3-D unmixing vectors found by PCA are plotted with the color indicating their vertical component.  $PC_1$  and  $PC_2$  progress through the color spectrum of the vertical axis at the same rate as the data points and consequently lie in the plane. However,  $PC_3$  progresses through the color spectrum more rapidly as it lies perpendicular to the plane. (column 3) The three components recovered by PCA and the variance in each direction. As  $PC_3$  can be seen to correspond to noise, only components 1 and 2 are retained for use by ICA. The upper scatter plot shows  $PC_1$  and  $PC_2$  rescaled in such a way that the variance in each direction is 1 (i.e., the data are whitened). The unmixing vectors found by ICA are shown in red. (column 4) The two sources recovered by ICA, which can be seen to be a good approximation of the original sources. Note that due to the ambiguity of the sign of the recovered sources found by BSS methods, some of the recovered sources are sign-flipped versions of the original. However, to aid in quick analysis of the recovered sources, we remove any obvious sign flips from figures within this work. PCA = principal component analysis; ICA = independent component analysis; BSS = blind signal separation; PC = principal component; IC = independent component.



**Figure 4.** (column 1) The two temporal sources that are mixed with the three mixing vectors of  $\mathbf{A}$  (plotted in red). As per the previous scatter plots, yellow indicates areas of high point density. (column 2) Scatter plot of the three mixtures (top), the unmixing vectors found by PCA, and the three mixtures. These show the same features as those discussed in detail in Figure 3, but the salient point remains that the points lie on a plane, with  $PC_1$  and  $PC_2$  in the plane and  $PC_3$  orthogonal to it. (column 3) The three components recovered by PCA and the variance in each direction.  $PC_2$  can be seen to recover source 2 well, but  $PC_1$  is a poor approximation of source 1. As component 3 can be seen to correspond to noise, only components 1 and 2 are retained for use by ICA. The unmixing vectors found by ICA are shown in red. (column 4) The two sources recovered by ICA. Visual inspection shows these to be a good approximation of the original sources. PCA = principal component analysis; ICA = independent component analysis; PC = principal component; IC = independent component.

created by modeling the inflation of a point source in an elastic half space (Mogi, 1958) as this has been used to successfully model observations of deformation at Campi Flegrei (Lundgren et al., 2001) and is strongly non-Gaussian by all the measures utilized. The spatial pattern for the topographically correlated delay is synthesized by assuming a linear relationship between phase and altitude for each pixel (calculated using the Shuttle Radar Topography Mission (SRTM) 30-m Digital Elevation Model (DEM) Farr et al., 2007).

The temporal nature of the two synthetic signals is also shown in Figure 2 and, as these are sources when the data are temporally organized, provides insights into the challenge of recovering them. The time course for the deformation is generated to approximate a period of inflation at a volcano by varying in strength around

a mean value that is above zero. The time course is generated by drawing values from a hyperbolic secant distribution with mean of 0.1 and variance of 1. This distribution is chosen as its excess kurtosis ( $k = 2$ ) is similar to that found during analysis of the size of changes in displacement over 6-day intervals at a variety of Global Positioning System (GPS) stations ( $k = 2.4$ ) that experienced deformation due to volcanic unrest, using a method similar to that described in Liu et al. (2018) (see supporting information for further details).

The time course for the topographically correlated tropospheric phase delay is synthesized as a sinusoidal wave (wavelength: 1 year, amplitude: 4 rad/km) combined with Gaussian noise as this closely resembles the temporal evolution of the delay/elevation ratio at Colima Volcano measured by Pinel et al. (2011). While this function may not be applicable to all volcanoes, it provides a challenging example for our synthetic tests as when a daisy chain of interferograms is formed, the long-term sinusoidal trend is removed, and only the synthetic Gaussian noise remains.

In section 2.1 we also addressed why the signal introduced by the turbulent APS could only be treated as a noise term (in contrast to recovering the turbulent APS for each interferogram). For synthetic tests, we generate turbulent APSs as spatially correlated noise and then difference pairs of these to make signals that would be expected in either a single master or daisy chain time series. As the turbulent APS acts as noise and complicates the recovery of latent sources, we do not include it in the examples presented in this section. However, in the more complex synthetic tests performed in section 4, a turbulent APS is included. Figures 3 and 4 show examples of spatially organized and temporally organized synthetic data and the results of applying ICA and PCA to them.

### 2.3. PCA

PCA (also termed the Karhunen-Loéve expansion Karhunen, 1947; the Hotelling transform, Hotelling, 1933; and empirical orthogonal functions, Lorenz, 1956) has been applied to InSAR data by several authors, such as Kositsky and Avouac (2010), Rudolph et al. (2013), and Chaussard et al. (2014) to isolate signals of geophysical interest. The first principal component (PC) is the direction in which the maximum amount of variance in the mean-centered mixtures can be explained; the second seeks the same result given that it is orthogonal to the first and so on. The data can then be projected in the basis defined by the PCs, with the result that the data are now uncorrelated. The PCs can be found in a variety of ways (e.g., singular value decomposition) but are routinely found by calculating the eigenvectors and eigenvalues of the mean-centered data's covariance matrix. Once the eigenvectors have been calculated, the change of basis can be achieved using the following:

$$\mathbf{S}_{\text{reco}} = \mathbf{E}^T \mathbf{X} \quad (5)$$

where  $\mathbf{S}_{\text{reco}}$  are the reconstructions of the sources,  $E$  is a matrix of eigenvectors as columns (termed  $W$  in equation 3), and  $\mathbf{X}$  is the matrix of mixtures. A consequence of PCA requiring the data to be first mean centered is that a reference pixel need not be chosen for the time series as all the interferograms are adjusted so that when the data are spatially organized each interferogram has a mean of zero and when temporally organized the time history for each pixel has a mean of zero. The strength of PCA lies in the nature that the first PCs (or eigenvectors) contain the majority of the variance of the data, and so by discarding the later components, a large proportion of the variance of the data can be expressed in relatively few dimensions. However, due to the orthogonality of the PCs, it is apparent that should the rows of  $A$  not be orthogonal, PCA cannot separate the two sources fully. Consequently, in some fields ICA (Comon, 1994; Jutten & Herault, 1991) is preferred to PCA.

PCA can be applied to both spatially organized data (to find uncorrelated images) and temporally organized data (to find uncorrelated time courses). However, when performing PCA on images (such as interferograms) organized temporally (i.e., where each pixel is a variable), many data sets will have more dimensions than points populating the space (e.g., for a time series of 19 interferograms, each of 10,000 pixels, this would result in 19 points in 10,000 dimensions). Consequently, both the calculations of the covariance matrix (of size  $10,000 \times 10,000$  in this example) and its eigendecomposition become very computationally expensive. Previous applications of PCA to temporally organized interferograms by Ebmeier (2016) have circumvented this by spatially downsampling the data and so reducing the number of pixels and therefore dimensions. However, we instead calculate the PCs of these sparsely populated high-dimensional spaces using the PCA *compact trick* (Solem, 2012), which is computationally efficient as it uses the constraint that when the number of dimensions is greater than the number of samples,  $s$ , there will only be  $s - 1$  eigenvectors.

Minimal considerations of whether the data are complete or undercomplete are required when applying PCA. In the complete case, all the eigenvectors are retained and  $\mathbf{E}^T$  is square. In the noiseless undercomplete case

with  $m$  mixtures and  $s$  latent sources, the observations of the mixtures will lie on an  $sD$  hyperplane (demonstrated in Figure 3) and the last  $m - s$  eigenvalues will be zero, indicating that the corresponding eigenvalues can be discarded (reducing  $\mathbf{E}^T$  to an  $s \times m$  matrix). In the noisy undercomplete case, the data create a volume of equal dimension to the space and all eigenvalues are nonzero, but further interpretation depends on the signal-to-noise ratio (SNR). In the case that the SNR is high enough, the first  $s$  eigenvalues are significantly larger than the remaining  $m - s$  and the number of sources can be identified, but as the SNR decreases, the difference between the two sets of eigenvalues decreases until separation and constraint of the number of sources becomes difficult. The difficulty in choosing the number of components to retain is demonstrated in McKeown et al. (1998), where a small but interesting signal was identified in a very low eigenvector. Spatially small signals contained within large (250 km wide) Sentinel-1 interferograms may contribute little to the overall variance of the time series and be at risk of omission due to occurring in a low-ranked eigenvector yet may be of geophysical interest. We expect the strength of a signal, the proportion of an image that it covers, and the number of interferograms that it features in to determine how highly the signal is ranked within the eigenvectors extracted by PCA. Consequently, we postulate that in order to maximize the chance of a signal of a given strength and spatial size to be retained in the highest eigenvectors, the interferograms should be cropped to the smallest practical area of interest around a volcano. Given that deformation sources are commonly offset from a volcanic center (Ebmeier et al., 2018) by tens of kilometres, the smallest practical area is likely to be around 20–50 km in size but to remain a parameter that is tuned for each application of a BSS method to a volcanic center.

#### 2.4. ICA

ICA provides a new basis for the mixtures such that they are no longer statistically dependent (i.e., the independent components, or sources, have been recovered) and is the dominant method to apply BSS (Stone, 2002). In a similar fashion to PCA, it has been applied to InSAR data by Ebmeier (2016) in order to isolate signals of volcanic interest. ICA can be used to identify latent sources as these usually have probability density functions (PDFs) that are less Gaussian than mixtures. This is due to the central limit theorem, which stipulates that by summing several independent sources, the resulting mixture has a more Gaussian PDF than any of the constituent sources (Hyvärinen & Oja, 2000). ICA algorithms measure the Gaussianity of linear combinations of the mixtures, before adjusting this mixture in order to create a signal that has maximum non-Gaussianity and is therefore likely to be one of the original sources. This point also highlights the limitation that ICA algorithms cannot be applied to Gaussian signals (or random variables), as the latent sources would not be more non-Gaussian than the mixtures. The wealth of successful applications of ICA to BSS problems can be used to justify the expectation that it would outperform PCA as two physical processes that are unrelated (such as deformation at a volcano and atmospheric delay) are likely to be statistically independent, as opposed to merely uncorrelated. However, signals such as deformation and atmospheric delay at a stratovolcano may be physically unrelated but are likely to be spatially similar. We explore the search for spatially uncorrelated and independent sources at a stratovolcano in more detail in section 4.5.

Several measures of the non-Gaussianity of random variables exist and have been used by various ICA algorithms. Kurtosis provides one of the simplest measures to implement and measure how *peaked* a distribution is. As Gaussian random variables have a kurtosis of 3, it is common for kurtosis to instead refer to excess kurtosis, which is a measure in which all values are reduced by 3, shifting the kurtosis of a Gaussian distribution to 0. A random variable with a high kurtosis (i.e.,  $k > 0$ ) has a peaked or *spiky* PDF with many values closely grouped together and long tails, while a random variable with a low kurtosis (i.e.,  $k < 0$ ) has a very broad PDF (such as a uniform distribution, with a kurtosis of  $-1.2$ ). Gradient descent can be used to maximize the non-Gaussianity of the signals to be recovered, but the more complex fixed-point iteration presented as FastICA in Hyvärinen and Oja (1997) has two advantages in that it has been shown to converge more quickly and also does not require a learning rate to be chosen.

While kurtosis is an intuitive and computationally efficient way of measuring the non-Gaussian nature of a random variable, it is not robust and can be heavily influenced by outliers (Hyvärinen, 1999). Therefore, differential entropy is used as a measure of non-Gaussianity in more recent versions of the FastICA algorithm (Hyvärinen, 1999). Differential entropy is a measure of the information that a variable conveys and is largest for more unpredictable (or random) variables and lowest for more predictable ones (in the case of a coin that always landed on heads, the entropy would be 0). For a collection of random variables with equal variance, the maximum differential entropy is achieved by that with a Gaussian distribution (Hyvärinen et al., 2001). Therefore, it can be used as a measure of how Gaussian a random variable is. This can be simplified by defining

a new quantity, negentropy, which is 0 for a Gaussian random variable and always nonnegative. However, as differential entropy (and so negentropy) requires the PDF of a variable to be known, it remains difficult to measure efficiently. Therefore, approximations for differential entropy have been developed for the FastICA algorithm, but a full discussion of these is outside the scope of this paper (see, e.g., Hyvärinen et al., 1998, 2001). As per using kurtosis, gradient descent can be used to maximize the approximated negentropy, but a faster and more robust fixed-point algorithm is presented in a newer version of FastICA in Hyvärinen (1999).

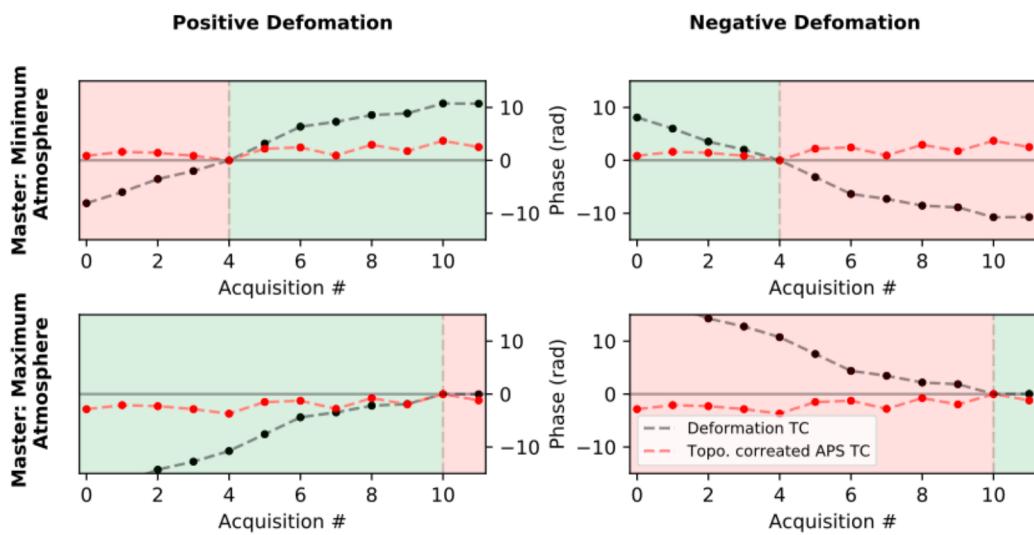
As per PCA, ICA can be applied to both spatially and temporally organized data. When ICA is applied to spatially organized data, spatially independent sources and unconstrained time courses are recovered, and the method is referred to as spatial ICA or sICA. When ICA is applied to temporally organized data, independent time courses and unconstrained spatial sources are recovered, and the method is referred to as temporal ICA or tICA. The independent sources that ICA seeks to recover are similar to the uncorrelated sources that PCA seeks to recover, but as independence is a more robust measure than the uncorrelatedness which PCA seeks (discussed further in section S3), we may expect the sources recovered by ICA to be more faithful reconstructions of the latent sources that generated the data. ICA has been applied across a variety of fields, ranging from tICA on speech data (Bell & Sejnowski, 1995), tICA on electroencephalographic data (Makeig et al., 1996), sICA on functional magnetic resonance imaging data (McKeown & Makeig, 1998), and sICA on facial images (Bartlett et al., 2002). More recently, sICA and tICA have also been performed on geophysical data using the FastICA algorithm (Chaussard et al., 2017; Cohen-Weber et al., 2018; Ebmeier, 2016; Frappart et al., 2010).

Unlike PCA, applying ICA to undercomplete data requires extra considerations. FastICA can only operate with square mixing and unmixing matrices, and while this makes them very suitable for the complete case, it makes their application to the undercomplete case more challenging. An example of performing ICA with rectangular matrices is presented in Porrill and Stone (1998) and termed undercomplete ICA, while the error-gated Hebbian rule proposed by Isomura and Toyoizumi (2016) retains square mixing and unmixing matrices but is capable of recovering repeated versions of the original sources if the data are undercomplete. However, we choose to implement the commonly used method of dimension reduction, as the undercomplete ICA method presented in Porrill and Stone (1998) has not been applied to data similar to geophysical data, and the error-gated Hebbian rule ICA algorithm of Isomura and Toyoizumi (2016) requires tuning of a learning rate, which may prohibit the automating of the implementation of ICA that is required for ICA to be used in an automatic detection algorithm.

Dimension reduction seeks to compress the data by expressing it using a new and smaller set of variables which, in the ideal case, are still able to convey the essential features of the data. Using this as a preprocessing step for ICA allows us to reduce the number of input variables to equal the number of sources that we wish to recover and then allows ICA to be performed on the new lower dimension data (i.e., where  $\mathbf{A}$  and  $\mathbf{W}$  are now square). PCA is commonly used for this process but includes the important caveat that some signals may be discarded in dimensions that were thought of as unimportant (McKeown et al., 1998). Most ICA algorithms require the data to be whitened (or sphered) prior to ingestion as this greatly simplifies the unmixing matrix as it changes from a full rank to orthogonal matrix (Hyvärinen & Oja, 1997), and as PCA can be used to perform whitening, it is usually incorporated into the dimension reduction step. As tICA requires temporally organized data to be whitened, we again use the compact trick (discussed in section 2.3) to allow tPCA to be performed without downsampling the data.

## 2.5. PCA and ICA Examples

To demonstrate the functioning of PCA and ICA, we present a low-dimension example of undercomplete data (three mixtures of two sources). While these are of significantly lower dimension than would be encountered for real data, they remain useful examples as they allow for the data and (un)mixing vectors to be plotted clearly and the role of PCA as a preprocessing step for ICA to be demonstrated. For the time series to be of such low dimension that it can be plotted in 3-D, spatially organized data must consist of three interferograms, while temporally organized data must consist of three pixels (in a similar fashion to the 2-D case shown in Figure 1). While three mixtures could be generated from more than two sources, this is the opposite of the undercomplete case that we expect to encounter with InSAR data. Therefore, we use two sources and a  $3 \times 2$  mixing matrix ( $\mathbf{A}$ ) to generate undercomplete data. To generate spatially organized data, the spatial pattern of the two sources (deformation and topographically correlated delay) is placed as rows in  $\mathbf{S}$ , and the time courses placed as columns in  $\mathbf{A}$  (see equation 2), while for temporal data, the opposite is performed.



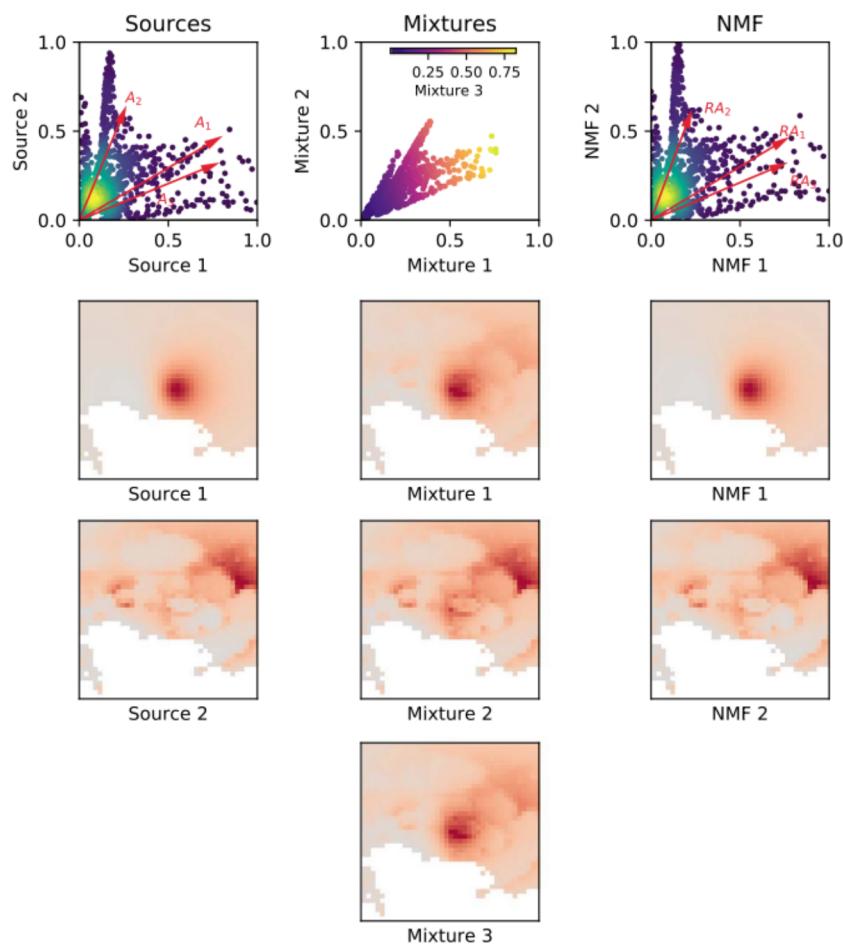
**Figure 5.** Possible scenarios for application of NMF to single master InSAR time series. (top left) Uplift (black points) and a topographically correlated APS (red points) with the master interferogram on the acquisition with the minimum atmosphere (number 4). Data after the master (green shading) are nonnegative and can be used, while data before the master contain both positive and negative data and cannot be used (red shading). (top right) As in top left but with subsidence. Data before the master data are nonnegative. (bottom left) Uplift and a master chosen for the maximum atmosphere (number 10). The data before the master are nonpositive and can be utilized by NMF providing a trivial sign flip is performed. (bottom right) Subsidence and a master chosen for the maximum atmosphere. A sign flip of the data after the master facilitates application of NMF. NMF = nonnegative matrix factorization; InSAR = interferometric synthetic aperture radar; APS = atmospheric phase screen; TC = time course.

Figure 3 shows the Euclidean representation of mixing sources to create spatially organized data and the results of using PCA and ICA to attempt to recover the sources. As only two sources and a negligible amount of noise (visible as PC3 in Figure 3) are used to generate the data, the data lie on a plane in the 3-D space of the three mixtures. The first two PCs lie in this plane and due to the orthogonality constraints imposed on the unmixing vectors of PCA, the third lies perpendicular to the plane. This results in the variance accounted for by the third PC being approximately zero, and we are therefore able to conclude that the three mixtures were generated by two sources. To perform ICA, we rescale the data projected in the direction of the first two PCs such that the variance in each direction is one (i.e., the data are whitened) and then seek two unmixing vectors in this 2-D space. Should noise be added to the above case, the three mixtures would no longer lie in a plane and would instead create a volume within the 3-D space. If the signal is of significantly larger magnitude than the noise, identifying the plane in which the sources lie remains possible, and ICA is able to recover the sources accurately. However, as more noise is added, the plane that the sources creates becomes harder to identify using PCA, and if the vectors found by PCA are not aligned correctly (i.e., the first two lie in this plane), some signal may be present in the third PC. When this component is discarded, the signal it contained is also lost, and ICA may therefore fail to recover the original sources faithfully.

Figure 4 shows the Euclidean representation of mixing sources to create temporally organized data and the results of using PCA and ICA to attempt to recover the sources. As this scenario is limited to three pixels, these do not have any spatial meaning, so the elements of  $\mathbf{A}$  are chosen from uniform distributions  $-0.5$  to  $+0.5$ . In a similar manner to the spatial example discussed in the previous paragraph, the very low variance of the third PC ( $1.85 \times 10^{-15}$  % of the total variance) indicates that it is numerical noise and should not be retained for use by ICA. While comparison of the original and recovered sources is not as clear as for the spatial patterns, examination shows that the sources recovered by PCA are visibly different to the original sources while those recovered by ICA are not.

## 2.6. NMF

NMF developed from positive matrix factorization (Paatero & Tapper, 1994) and factorizes a nonnegative data matrix of mixtures,  $\mathbf{X}$ , into two nonnegative matrices,  $\mathbf{A}$  and  $\mathbf{S}$  (often termed  $\mathbf{W}$  and  $\mathbf{H}$  in NMF literature). NMF became well known when Lee and Seung (2000) showed that a collection of 2,500 facial images, each of

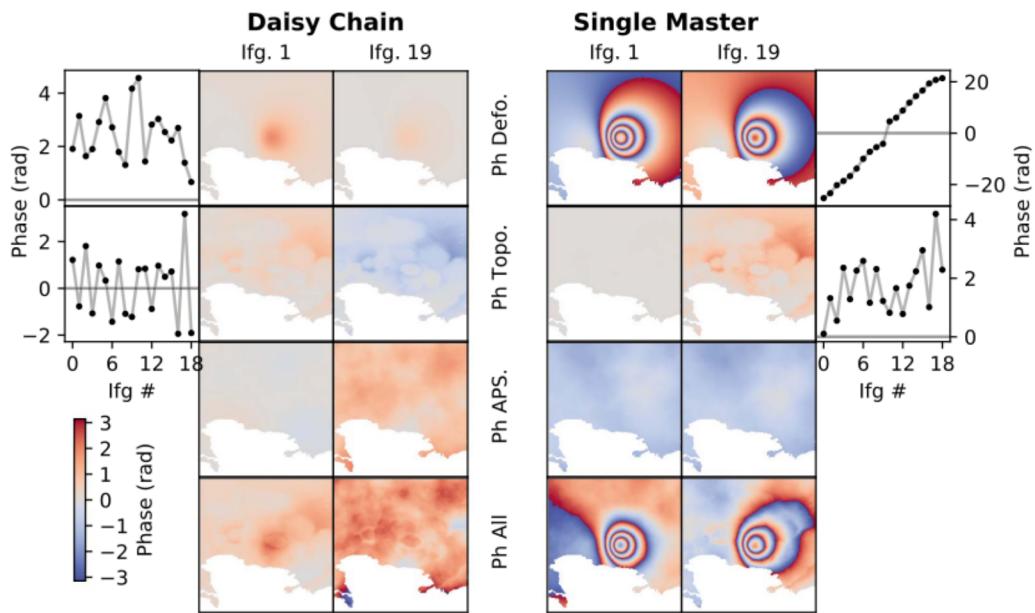


**Figure 6.** Results of application of NMF to spatially organized data. (column 1) To create three mixtures from two sources, three 2-D vectors are required (rows of  $\mathbf{A}$ ). (column 2) Scatter plot of the three mixtures. The points form a plane in the 3-D space which allow them to be visualized adequately using color for the third axis. (column 3) The two recovered sources and mixing vectors, showing the near exact recovery of the original sources. NMF = nonnegative matrix factorization.

361 pixels, could be expressed as 49 sources that were easily interpretable to a human observer (i.e., the sources corresponded to parts of a face, such as eyes or a nose). In terms of Euclidean geometry, this corresponds to a 361-dimensional space populated by 2,500 points being condensed to a 49-dimensional space and corresponds with the underdetermined case described in section 2.1. However, to our knowledge, NMF has not been applied to InSAR data.

The multiplicative update rules of Lee and Seung (2000) find the local minimum of an objective function that measures the misfit between  $\mathbf{X}$  and  $\mathbf{A} \times \mathbf{S}$ . If the mixtures are linear combinations of a smaller number of latent sources, gradual identification of these sources (and their mixing matrix) occurs as the objective function reduces. In the case of noiseless data, the objective function may approach 0.

However, in the complete case (i.e., the number of sources and mixtures are equal) any positive and non-linearly dependent vectors can be used to fit the data exactly, providing the values of  $\mathbf{S}$  that are adjusted correctly by the algorithm. Consequently, the original sources are not recovered, and the data are instead reconstructed from near-random permutations of the latent sources. This situation can be demonstrated by considering observations of three mixtures that are fit when the rows of  $\mathbf{A}$  contain vectors in the direction of the space's coordinates (i.e.,  $[1, 0, 0]$  etc.), and the three recovered sources would be the three mixtures. This arrangement would be able to reconstruct the data exactly but would not recover meaningful or accurate sources. Therefore, unlike PCA and ICA, NMF can only be applied to undercomplete data.



**Figure 7.** Comparison of a synthetic time series of interferograms arranged as either a daisy chain (left) or relative to a single master (right) formed from 20 acquisitions. The top row shows the signal due do inflation of a Mogi source (and its associated time course), the second due to a topographically correlated atmospheric signal (and time course), the third due to a turbulent atmospheric phase screen, and the fourth due to the combination of these. In the single master time series, the master date was chosen to be when the topographically correlated atmospheric signal was a minimum, in order to ensure it remains positive in all the interferograms and nonnegative matrix factorization can be used. The phase is wrapped to between  $-\pi$  and  $\pi$  as the scale varies significantly between the two cases. In the daisy chain case, the signal due to the Mogi source is of comparable magnitude to the other signals but in the single master case dwarfs them. Figure 8 shows the results of applying the suite of blind signal separation techniques to each case.

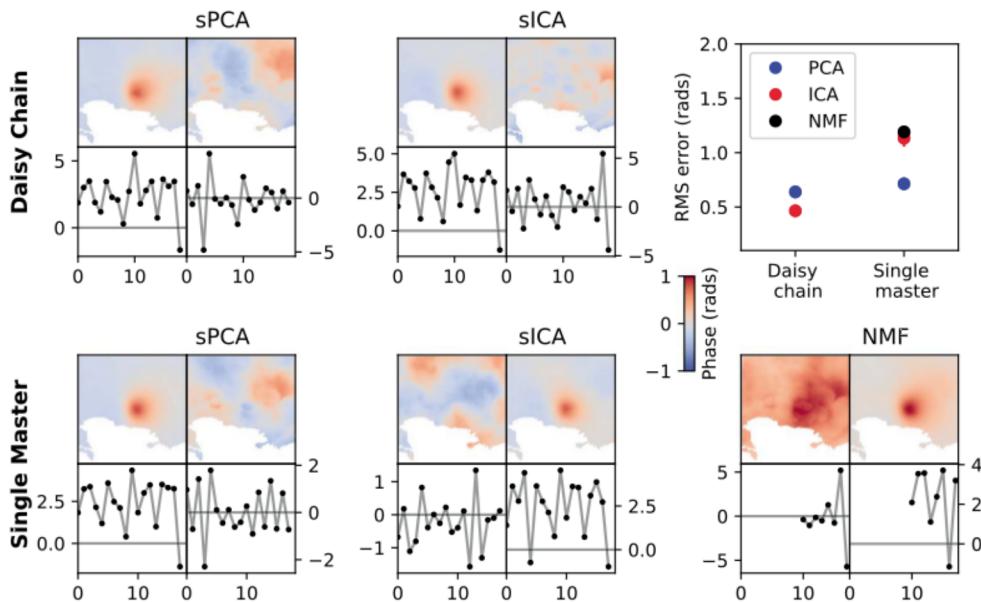
### 3. Application of NMF to InSAR Time Series

As NMF cannot be used on data that contains negative values, we instead construct a time series of interferograms relative to a single master. This master image is chosen to be the date on which either the strongest/weakest topographically correlated atmosphere occurs as this ensures that this signal is either entirely positive/entirely negative. We synthesize a volcanic signal that is steadily inflating/deflating at a varying rate, to ensure that all the data after/before the master are positive. NMF can then be applied to regions of the time series in which both signals are positive or, by applying a trivial sign flip, both signals are negative. Figure 5 demonstrates these possible scenarios which, while somewhat limited, we believe remain useful. Dates on which strong topographically correlated APS signals occur may be estimated from outside deforming regions, but more complex methods are required for cases in which the deformation is not of a constant sign before or after the master image, which are outside the remit of this initial study.

Figure 6 shows the results of applying NMF to spatially organized undercomplete 3-D data that correspond to any of the green regions of Figure 5. Very slight correlations remain between the two sources due to incomplete separation, but NMF can be seen to recover the sources well (mean residual per pixel: 0.0001). For brevity, we do not show the results of applying the same process to temporally organized data as it is very similar to Figure 4. The addition of noise in the form of a turbulent APS and any associated negative values is addressed in section 4.1 through adding small values to all of the pixels within an interferograms (or time points within a time courses if the data are temporally organized).

### 4. Comparison of PCA, ICA, and NMF

To determine which of the BSS methods is most suited to isolating signals of geophysical interest, we construct a collection of differing time series in which we vary whether (1) the interferograms are constructed as a daisy chain or relative to a single master, (2) the number of sources recovered, (3) the length of the time series, and (4) the strength of the turbulent APS. These time series consist of three types of synthetic signals aimed to



**Figure 8.** The results of applying PCA, ICA, and NMF to the spatially organized version of the data (i.e., sPCA and sICA) shown in Figure 7. To ease interpretation and comparison of the recovered time courses, the single master cases are differentiated to produce the equivalent daisy chain time courses. Application of the BSS methods to temporally organized data (i.e., tPCA and tICA) produced very poor results and so are omitted for clarity, but further figures (e.g., Figures 9–11) show these methods. The time courses recovered by NMF (lower right) only show the signals after the master date as these signals are predominantly positive but are generally very poor reconstructions of the synthetic sources. In the daisy chain case, both PCA and ICA are seen to recover both the sources well, while in the single master case PCA and ICA are seen to recover the deformation source well and the topographically correlated APS poorly. PCA = principal component analysis; ICA = independent component analysis; NMF = nonnegative matrix factorization; sPCA = spatially organized principal component analysis; sICA = spatially organized independent component analysis; BSS = blind signal separation; APS = atmospheric phase screen.

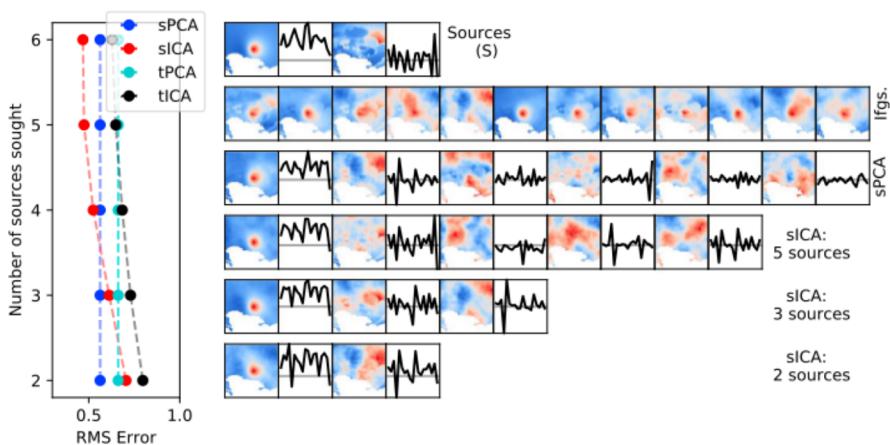
mimic those seen in real data, with one corresponding to a deformation source, a second to a topographically correlated APS, and a third to a turbulent APS. While this is not a comprehensive list of types of APS seen in InSAR data, it contains enough complexity for illuminating experiments with BSS methods to be performed.

To evaluate the fidelity of the sources recovered by a BSS method, we first multiply each time course by its associated spatial map to construct a time series of only that signal. The mean residual per pixel is then calculated between each synthesized source and each recovered source, before finally calculating the total residual when different recovered sources are used to match each synthesized source (i.e., one recovered source cannot be used to fit both synthesized sources).

#### 4.1. Single Master Versus Daisy Chain Time Series

When creating interferograms from SAR acquisitions, the interferograms can be constructed to show the signal between two acquisitions with a short spatial or temporal baseline or between each acquisition to a single master acquisition (see Hooper et al., 2012, for a more complete description). In the case that spatial baselines remain small, interferograms can be created between acquisitions with the shortest temporal baselines, making what is often termed a daisy chain of interferograms. Figure 7 shows how the same signals manifest themselves when constructed using either the daisy chain or single master approach, while Figure 8 shows the results of using the BSS methods described previously on each case.

For the single master case, NMF can be seen to recover the sources poorly as even though the deformation source is successfully isolated in one recovered source; the second recovered source contains a mixture of the two synthetic sources. This incorrect separation creates a high residual (root-mean-square [RMS] error: ~1.2 rad). When applied to single master data, sPCA and sICA are also able to recover the deformation signal accurately and the topographically correlated APS poorly, producing a moderate to high overall error (RMS error: ~0.8 and ~1.1 rad, respectively). This result seems plausible as when the data are organized relative to a single master, the deformation signal becomes around an order of magnitude larger than the topographically



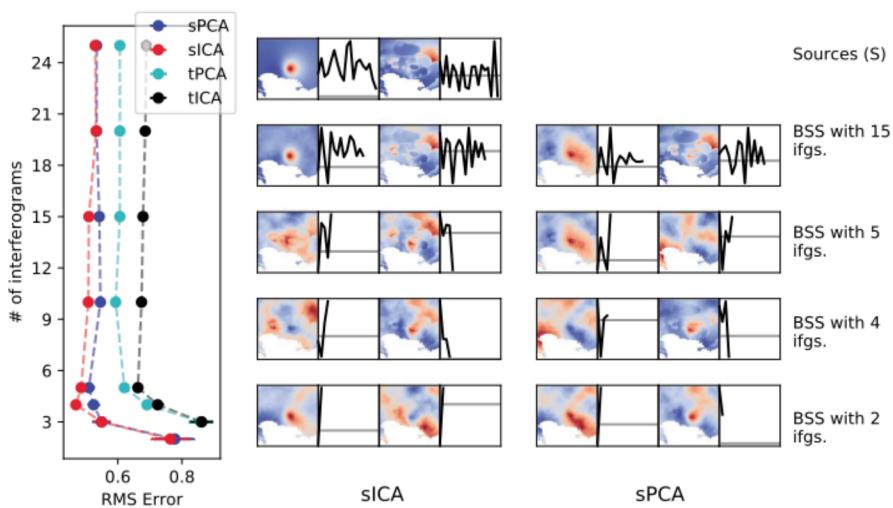
**Figure 9.** The results of sPCA, sICA, tPCA, and tICA applied to a time series of 25 daisy chain interferograms when the number of recovered sources is varied. The RMS residual between each synthesized and recovered case is shown on the left after averaging the results over 50 synthetic time series, while the right-hand section shows the results from one particular time series. The two sources (top row) each have variances of 1 and are mixed with a turbulent atmospheric signal also of variance 1 to produce the time series of 25 interferograms in row 2. The third row shows the six largest sPCA components, with the first showing elements of the synthesized Mogi source and the fourth showing a mixture of elements of the topographically correlated APS and noise created by the turbulent APS. The remaining three rows show the results of sICA when a decreasing number of sources are sought. As ICA does not place sources in a significant order (unlike PCA), the recovered sources most like the synthesized sources have been placed in the first two columns for clarity. sICA performs approximately equally well when five or six sources are recovered, which we attribute to being a consequence of part of the topographically correlated APS residing in PC4 and PC5 and only being accessible to the sPCA = spatially organized principal component analysis; sICA = spatially organized independent component analysis; tPCA = temporally organized principal component analysis; tICA = temporally organized independent component analysis; RMS = root-mean-square; PCA = principal component analysis; ICA = independent component analysis; APS = atmospheric phase screen.

correlated APS signal (see the time courses on the upper right of Figure 7). In the daisy chain case, both synthetic signals have a comparable magnitude and are recovered well, producing a low overall error (RMS error:  $\sim 0.5$  rad). We therefore conclude that through organizing the data in a daisy chain both atmospheric and deformation signals retain comparable variances and are more accurately recovered than in the single master case. Consequently, we select this method as optimal and use it in the following scenarios.

#### 4.2. Number of Sources Sought

In section 2.1 we introduced the undercomplete nature of InSAR data and discussed the importance of PCA as a preprocessing step for ICA. In the noiseless case introduced in Figure 3, using PCA to reduce the dimensionality of the data to equal to the number of sources was trivial as the sudden change in variance between the second and third components suggested that the data formed an almost flat 2-D feature in the 3-D space. In the example shown in Figure 9, the data are very undercomplete (25 interferograms from two sources) and contain significant noise (in the form of a turbulent APS), which presents a more realistic challenge for PCA and ICA.

As the number of sources that PCA recovers does not vary, it is only applied once to produce one set of recovered sources and time courses. In contrast to the data used in the previous section, we increase the variance of the turbulent atmospheric signal by 10% to visibly relegate the topographically correlated APS to the third and fourth PC (shown in Figure 9). Consequently, when sICA is performed to recover two sources, the signal contained in the third and fourth PCA components is discarded and sICA can only recover the deformation source, producing a high RMS residual. However, as sICA is performed on more of the first sPCA sources, the signal contained in the higher sPCA sources is available to the FastICA algorithm and the fidelity of the recovered sources increases (producing a lower residual). A consequence of this is that the ICA algorithm then also recovers some sources that are just noise (a turbulent atmospheric signal), but these can potentially be separated from those of interest either by eye or using clustering methods (Ebmeier, 2016). Figure 9 shows that the minimum residual is found when three extra sources are sought.



**Figure 10.** The results of sPCA, sICA, tPCA, and tICA applied to time series of interferograms of increasing length. The mean residual per pixel in the differing cases is shown on the left after averaging the results over 50 synthetic time series (with error bars showing the variance), while the right-hand section shows the results from one particular time series. The top row shows the two synthesized sources and their associated time courses, while the second to fifth rows show the results of sPCA and sICA applied to time series of different lengths. Performance of all the BSS methods increases as the time series increases to around five interferograms in length before then remaining constant. In the time series depicted on the right-hand side, the results when using longer time series show some aspects of the turbulent APS signal relegating the topographically correlated APS to below the second component. sPCA = spatially organized principal component analysis; sICA = spatially organized independent component analysis; tPCA = temporally organized principal component analysis; tICA = temporally organized independent component analysis; RMS = root-mean-square; BSS = blind signal separation; APS = atmospheric phase screen.

When applying ICA to real world examples, the problem becomes more complex as the number of sources is not necessarily known. However, this example shows that for ICA to perform optimally, the number of sources need be known only approximately so that the number to be recovered can be set correspondingly higher.

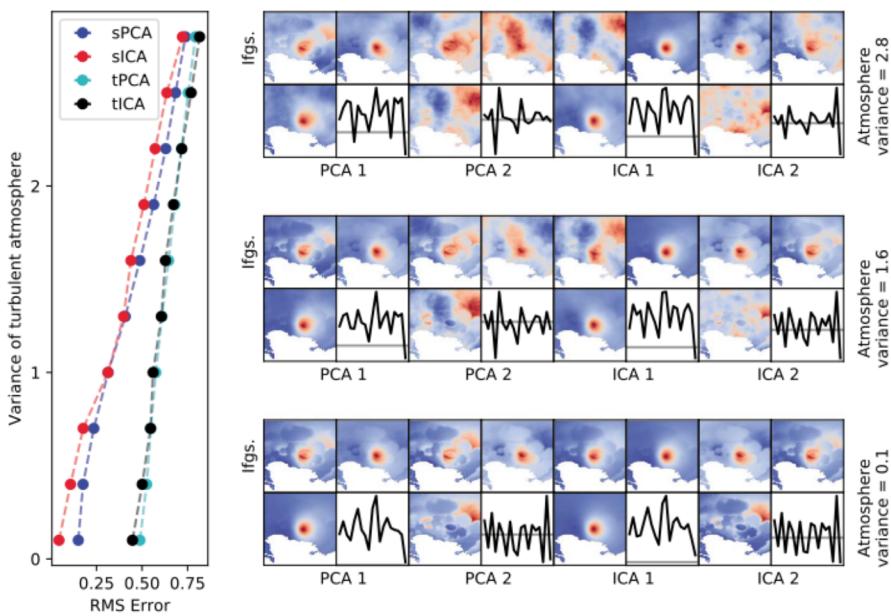
#### 4.3. Length of Time Series

The length of a time series of interferograms over a volcanic center can vary between a single interferogram spanning an event and a time series spanning several years. While the case of a single interferogram is not suitable for BSS methods to be applied, identification of the most suitable BSS method for time series of different lengths is important for an automatic detection algorithm. Consequently, in a similar fashion to the previous sections, we construct a suite of synthetic time series and crop these to different lengths to analyze the performance of different methods. In light of the findings of previous sections, we construct the time series as a daisy chain of interferograms and set the FastICA algorithm to retrieve four sources (i.e., two more than the two used to create the time series).

Figure 10 shows the results of applying sPCA, sICA, tPCA, and tICA to these time series. As per the previous experiments, application of PCA and ICA to temporally organized data produces poor results (mean residuals of 0.6–0.2, respectively), while to spatially organized data produces significantly improved results (mean residuals of 0.1–0.03, respectively). sICA is seen to outperform sPCA at all time series lengths, though the difference decreases in magnitude as the time series increases in length.

#### 4.4. Strength of Turbulent Atmosphere Contribution

In contrast to a topographically correlated atmospheric signal, a turbulent atmospheric signal cannot be retrieved by PCA, ICA, or NMF and consequently appears as noise in the time series. To determine whether PCA or ICA is best suited to dealing with time series with a strong turbulent atmospheric signal, we synthesize a suite of time series with differing strengths of turbulent atmospheric signal. In light of the findings of the previous section, we again arrange the time series as daisy chain of interferograms and seek four sources with the FastICA algorithm.



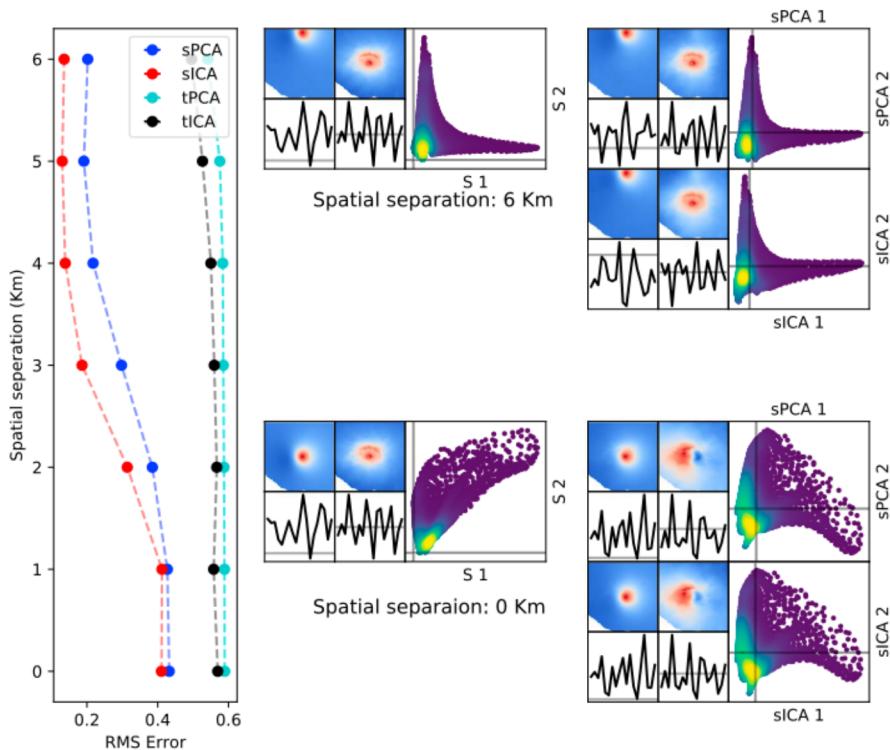
**Figure 11.** The results of sPCA, sICA, tPCA, and tICA applied to a suite of time series with different strengths of turbulent APS. The strength of the turbulent APS signal is quantified in terms of its variance which was set to 1 for the previous examples. The mean residual per pixel in the differing cases is shown on the left after averaging the results over 50 synthetic time series, while the right-hand side shows three cases for differing strengths of turbulent APS. For each case, the top row shows the time series of interferograms (showing the differing contribution of the turbulent APS), while the second row shows the sources and associated time courses recovered by sPCA and sICA. sPCA = spatially organized principal component analysis; tPCA = temporally organized principal component analysis; sICA = spatially organized independent component analysis; tICA = temporally organized independent component analysis; APS = atmospheric phase screen.

Figure 11 shows the results of applying sPCA, sICA, tPCA, and tICA to these time series. As per the previous experiments, the methods that are applied to spatially organized data produce results with significantly lower residuals. Of these, sICA is seen to outperform sPCA in the majority of cases, although there are certain strengths of atmospheric noise at which the results are comparable.

#### 4.5. Spatial Independence of Sources

When using PCA and ICA to recover spatial maps (i.e., images) as sources, we are seeking sources that are either uncorrelated or statistically independent. Consequently, a key assumption of ICA is that the latent sources are statistically independent (or uncorrelated for PCA), which may not be the case for signals that may be encountered at certain volcanic centers. Figure 12 shows the results of performing PCA and ICA in a similar manner as described in the preceding sections but instead of at Campi Flegrei, on Mount Vesuvius (a ~1,000-m-high stratovolcano ~15 km east of Naples). In such a case, the topographically correlated APS is dominated by the signal produced by the cone of the stratovolcano, and if the synthetic deformation source is located under this cone, the two signals lie in the same location. When considering the pixels of the interferograms, those at the top of the cone are now likely have high values in both the deformation and topographically correlated APS, while those at the edges are likely to have low values in the two sources. Consequently, the two sources are now no longer statistically independent, and we would not expect PCA or ICA to be able to separate them. This result is seen in Figure 12, which shows that as the two synthetic sources are brought closer together, they cease to be statically independent and the results of sPCA and sICA decrease in quality.

Figure 12 also shows the results of performing tPCA and tICA on the data with nonstatistically independent spatial maps. The results of these methods remain generally constant at all spatial separations, which we interpret as being due to PCA and the FastICA algorithm finding time courses that remain statistically independent regardless of changes to the spatial sources. However, in accordance with the preceding sections, the results of tPCA/tICA remain poor and are worse than those found by sICA when the sources are not statistically independent.



**Figure 12.** The results of sPCA, sICA, tPCA, and tICA applied to a suite of time series in which the overlap of the two sources is varied. The left-hand section shows the mean residual per pixel decreasing for sPCA and sICA as the spatial separation is increased and changing little for tPCA and tICA. The upper right-hand section shows the results with 6 km of separation. The two sources can be seen not to overlap, and no correlations are seen in the scatter plot comparing the two sources. The sources recovered by sPCA and sICA can be seen to be good reconstructions, and the associated scatter plots show the same lack of correlation as the original sources. The lower right-hand section shows the results with 0 km of separation. The two sources overlap and cause many of the pixels to be correlated (as demonstrated in the scatter plot). The sources recovered by sICA are not good reconstructions as the scatter plot shows that the FastICA algorithm has sought sources that are statistically independent (and therefore uncorrelated as seen in the scatter plot with the majority of the points now lying along the coordinate axis). The results from sPCA are similar, except with one source a negative version of the source found by sICA (causing the data in the scatter plot to be mirrored around the y axis). sPCA = spatially organized principal component analysis; sICA = spatially organized independent component analysis; tPCA = temporally organized principal component analysis; tICA = temporally organized independent component analysis; RMS = root-mean-square.

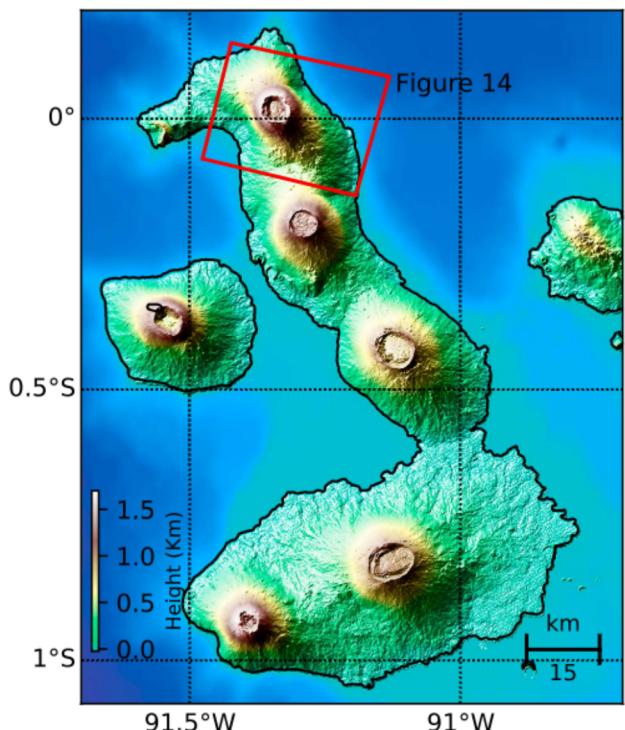
#### 4.6. Comparison Conclusions and Method Limitations

From the experiments carried out in the previous four subsections, we conclude from the suite of BSS methods studied that sICA is the most suited to use with InSAR time series. When sICA is used, performance is optimal when the number of sources sought is set to be slightly larger than the number expected to exist (such as recovering five sources when two are postulated to exist) and when interferograms are constructed to minimize the temporal baselines, such as through creating a daisy chain of interferograms.

However, limitations in the application of sICA remain. The performance of the algorithm reduces significantly when the spatial statistical independence of the sources is reduced, such as may happen at a stratovolcano where a topographically correlated APS and a broad deformation signal may be spatially similar. This issue is explored further through application of sICA to real data at a stratovolcano in section 5.2. The performance of sICA is also limited in cases where noise from the turbulent atmosphere may dwarf signals of geophysical interest, but we do not discover any thresholds at which sICA fails and instead see a gradual degradation of the accuracy of the recovered sources as the signal to noise ratio decreases.

#### 5. Application to Real Data

To further explore the ability of sICA to recover latent signals from a time series of interferograms, we present results from its application to two time series. The first spans the 2015 eruption at Wolf volcano



**Figure 13.** A shaded relief DEM of the western isles of the Galapagos Archipelago, with the calderas of six shield volcanoes visible as *upturned soup bowls*. Figure 14 focuses on Wolf Volcano, and the extent of the figure is depicted by the red region. Topography is taken from the SRTM DEM (Farr et al., 2007) and bathymetry from the GEBCO\_2014 Grid (GEBCO, 2015).

1998). Figure 14 shows a subset of this time series focussed on the eruptive period, with pixels with a postfiltering coherence of  $< 0.8$  masked. The dominant features of this time series are range increase for the caldera floor (interpreted as predominantly subsidence) and range decrease for the area around the circumferential fissure (likely to be eastward motion). However, interferogram eight also shows a broader signal (more visible in the wrapped interferogram) that corresponds to a signal attributed to the deeper magma chamber by Xu et al. (2016).

Figure 15 shows the results of applying sICA to the time series. As around 3–4 geophysical signals are expected, we set the FastICA algorithm to recover six components (in light of the results of Section 4.2). Visual inspection of the higher order PCs (6–20) suggests that these contain only turbulent atmospheric signals which are not at risk of containing important geophysical signals. sICA is able to recover the spatial pattern and time courses of the caldera floor subsidence, movement of the region surrounding the circumferential dike, and broad subsidence around the volcano. Recombination of the time courses and spatial patterns allows the time series to be reconstructed with minimal atmospheric signals.

### 5.2. Etna Volcano

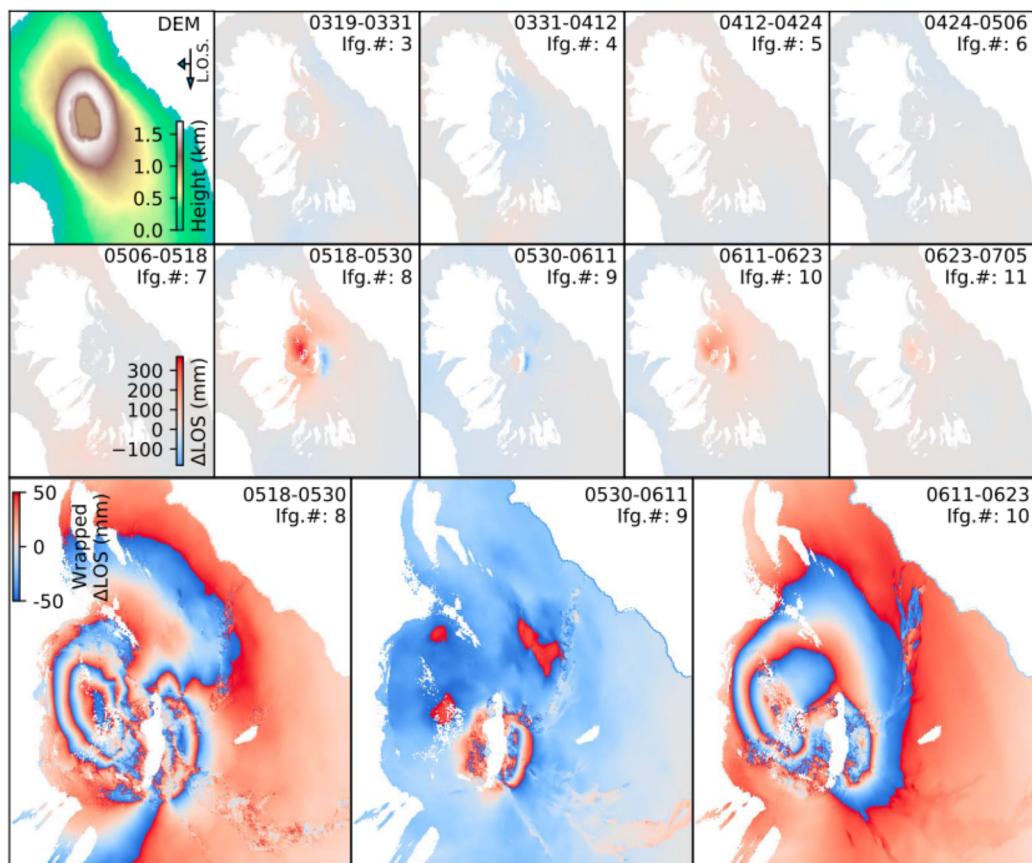
Mount Etna is a stratovolcano located on the eastern edge of the island of Sicily. It composes of over  $\sim 3,300$  m of elevation between its summit and eastern flanks and is considered to be one of the world's two most monitored volcanoes (González & Palano, 2014). InSAR has been used to measure deformation centered under the cone of Etna during the later portion of the 1991–1993 eruption (Massonnet et al., 1995), but a portion of the signals attributed to deformation by this study were later revised to be due to a topographically correlated APS (Delacourt et al., 1998). Both the deformation and APS described in these studies were centered under the topographic expression of this volcano, and as they are therefore unlikely to be spatially independent, we expect application of sICA at Etna to be challenging. Subsequent geodetic studies at Etna have also measured other deformation processes that may be recoverable by sICA, such as eastward movement of sections

(Galapagos archipelago, Ecuador) and was chosen as we further develop the use of sICA in an automatic detection algorithm that is able to detect the onset of this eruption. The second is centered on Mount Etna and was chosen as an example of a stratovolcano at which application of sICA is likely to be problematic, due to the considerations of spatial independence discussed in section 4.5.

#### 5.1. Wolf Volcano

Several existing studies present detailed results of modeling the observed surface deformations (e.g., Novellis et al., 2017; Xu et al., 2016) but we instead focus on the ability of sICA to automatically isolate the signals discussed in these papers. A detailed schematic of the timing of Sentinel-1 acquisitions and the two phases of the eruption is presented in Novellis et al. (2017), but we include the salient features and an overview map (Figure 13) here. The first phase involved the opening of a circumferential fissure on 25 May (2015) on the southeastern caldera rim. This fissure produced two lava flows down the southeastern flanks of the volcano (Venzke, 2015), but by 2 June activity ceased (Bernard et al., 2015). The second phase initiated around 11 October and involved an intracaldera fissure with lava flows that covered the caldera floor (Bernard et al., 2015). The surface deformation associated with these events was attributed to two dykes (one circumferential and one intracaldera) and two magma chambers ( $\sim 1$  and  $\sim 5$  km below sea level) by Xu et al. (2016) and to one dyke and one shallow magma chamber ( $\sim 1.5$  km below the caldera floor) by Novellis et al. (2017).

We formed a time series of 20 daisy chain descending Sentinel-1 interferograms covering the Isabela and Fernandina Islands in the western Galapagos archipelago from 13 December 2014 to 21 October 2015. The unwrapped interferograms were formed using LiCSAR (González et al., 2016) and include filtering with a Goldstein filter (Goldstein & Werner,

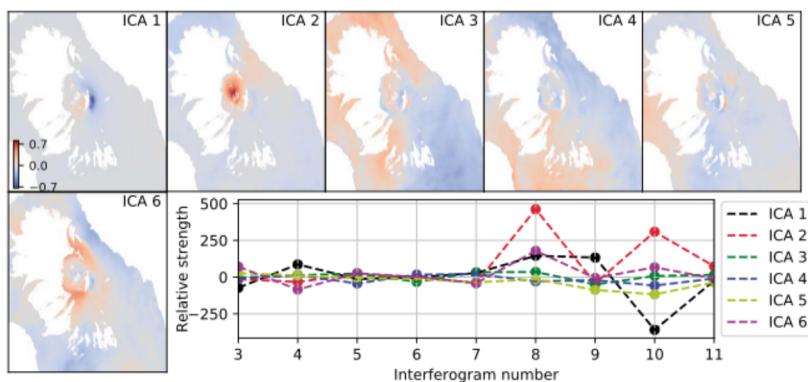


**Figure 14.** (top left) DEM showing the northern section of the island of Isabella (Galapagos Archipelago) and the cone of Wolf Volcano with the satellite line-of-sight (L.O.S.) vector in the top right corner. Interferograms 3–11: A subset of the Sentinel-1 time series spanning the period of unrest that occurred in May and June of 2015. Numbers in the top left of each interferogram depict the dates of the two acquisitions that the image spans (mmdd-mmdd); pixels with an average coherence  $< 0.8$  are masked (predominantly removing the pixels on the western vegetated slopes of Wolf), and line-of-sight (LOS) change is measured in millimeter, with positive values indicating an increase in range (corresponding to subsidence of the ground). (bottom panels) Interferograms 8–10 rewrapped to the interval  $-50$  to  $50$  mm in order to highlight more subtle features of the deformation pattern (such as the broad deflation signal in interferogram 8).

of the faulted eastern flank (Solaro et al., 2010) and westward movement of the western flank (Aloisi et al., 2007; Lundgren & Rosen, 2003)

We utilized a time series of Sentinel-1 interferograms that were formed as a preliminary result of work to use the LiCSAR processor (González et al., 2016) to automatically form interferograms at all of the world's volcanoes that have been active during the Holocene. Data storage and processing were performed at the Climate, Environment, and Monitoring from Space facility, and as some images are not yet available at this facility, our time series is split into two distinct sections (3 September 2016 to 8 March 2017 and 14 January 2018 to 2 May 2018).

However, as we expect these two time series to contain the same signals, we are able to perform sICA on them as a single time series. In contrast to the Galapagos time series used in the preceding section, the Etna time series contains interferograms created between each acquisition and around three following it, which created a network of overlapping interferograms. Pixels with a coherence below 0.7 in any of the interferograms were masked throughout the entire time series of 76 interferograms (which are shown in Figure S2). The time courses recovered by sICA describe the strength that each component was used in each of the 76 interferograms, and a simple least squares inversion was performed to invert for the strength that each component was used in a daisy chain of 28 interferograms linking the acquisition dates in the manner described by Lundgren et al. (2001).

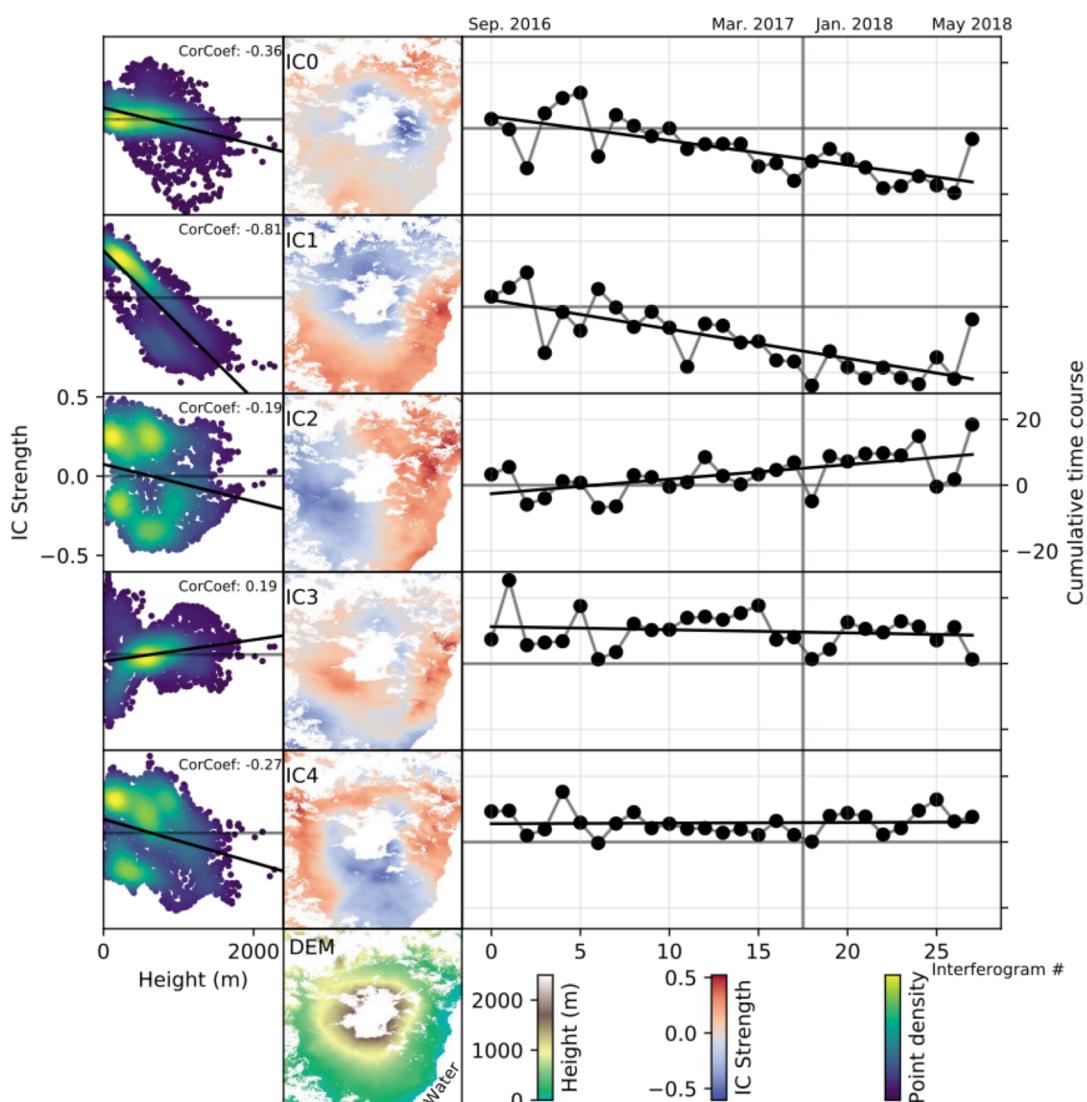


**Figure 15.** Results of sICA applied to the time series shown in Figure 14, showing the six components recovered and the strength of each one throughout a subset of the time series (lower right). We interpret components 1, 2, and 6 as representing deformation and the remainder as representing atmospheric signals. Component 1 appears to capture the signal near to the circumferential fissure, component 2 the subsidence of the caldera floor, and component 6 the broad subsidence associated with the deeper chamber. The remaining signals (3–5) contain traces of the other signals (such as the circumferential dike signal), but we interpret them as containing predominantly atmospheric signals.  
ICA = independent component analysis

Figure 16 shows the independent components (ICs) recovered by sICA, their strengths through the daisy chain of 28 interferograms, and a comparison between each IC and the DEM. We interpret the spatial pattern of IC0 as capturing eastward movement of the densely faulted eastern flank of the volcano. The cumulative nature of this component's time course is also indicative of it capturing deformation, but the final value attained by the cumulative time course remains low when compared to the changes at each time step, which is exemplified by the large change seen in the last data point. Inspection of the interferograms used in this analysis (Figure S2) shows the penultimate interferograms that contain a broad negative signal, which is likely to have caused this IC to have been used to attempt to fit it. Given the short nature of this time series (~ 12 months), the confidence in this measurement has the potential to be improved upon through the use of longer time series, but this remains beyond the scope of this paper. We also interpret IC2 as capturing broader east-west spreading of the volcano due to both the spatial pattern seen and the cumulative nature of the time course. In contrast to IC0-2, the time courses of IC3 and IC4 do not show any cumulative motion, and we conclude that these capture purely atmospheric signals.

When the correlations between phase and elevation are considered for each IC, only IC1 is seen to exhibit a strong linear relationship, which we interpret as suggesting that it may be capturing a topographically correlated APS. However, inspection of its cumulative time course shows that, with the exception of the last data point, an approximately linear increase occurs, which we interpret as suggesting that the IC is capturing broad, volcano wide deformation of the type first measured by Massonnet et al. (1995). Through use of a longer time series, we envisage that trends in the cumulative time course would become clearer, and should the cumulative signal continue to return toward zero as it does in the last data point, we would be more confident that the signal is solely capturing a topographically correlated APS. However, in the case that cumulative motion continues, we would be more confident that the IC contained a signal due to deformation but could not rule out a contribution from a topographically correlated APS.

From our initial analysis, we conclude that when applying sICA at stratovolcanoes, some signals of geophysical interest may be isolated, but separation of a broad inflation/deflation signal that is centered under the cone from a topographically correlated APS may not be possible. However, as we seek to use sICA as the foundation for an automatic detection algorithm, these results do not dissuade us. Through isolating the combination of broad deformation and a topographic APS to one IC, we are still able to characterize the baseline behavior of this combined signal through analysis of the IC's time course. Taking the example of a stratovolcano in which both a topographically correlated APS and a gradual inflation signal are isolated into one IC, any change in the rate of inflation would lead to a change in the time course of the IC, which we could then seek to flag as an indicator of the volcano entering a period of unrest. However, further development of this remains beyond the remit of this paper.

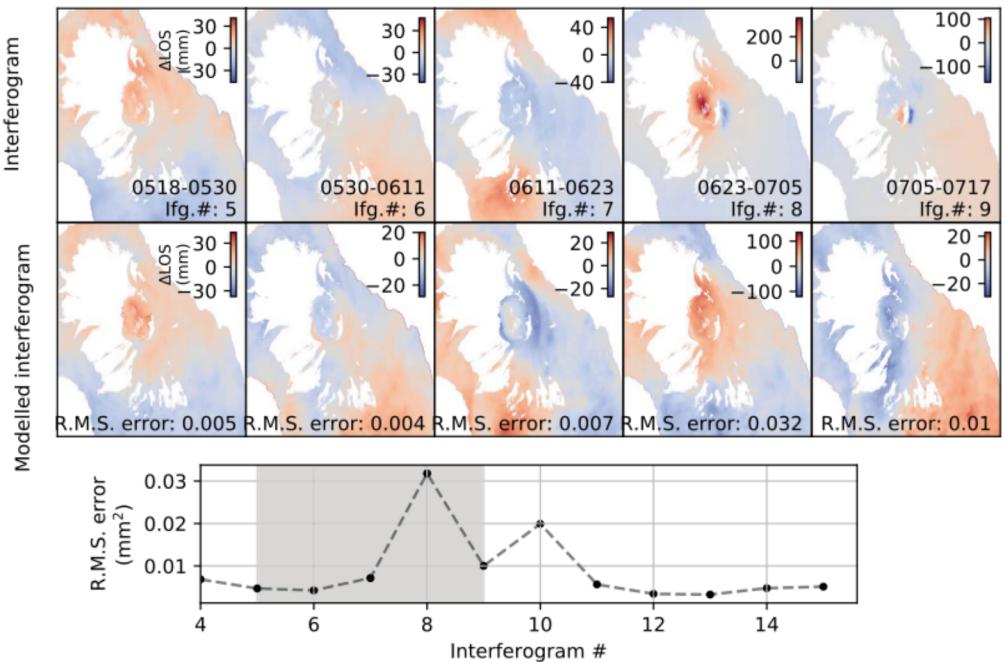


**Figure 16.** Column two shows the five independent components (ICs) recovered by sICA at Etna and the DEM covering the area processed. Pixels in the area marked *water* do not have a useable radar return, while the remaining white areas of the DEM are masked due to low coherence. Column 1 shows the results of plotting each IC against the DEM, with IC1's graph showing the clearest linear relationship which we interpret as being a result of IC1 capturing a topographically correlated APS. Column 3 shows the cumulative time courses, with IC1-3 showing cumulative motion throughout the two time periods that the data span.

## 6. Use of SICA in a Monitoring Algorithm

Our initial hypothesis was that the signals present in an InSAR time series at a volcanic center could be expressed as a linear combination of a small number of latent signals and that through isolation of these signals we could implement a way to monitor signs of unrest at a volcano. The results of the preceding section and those of Ebmeier (2016) have demonstrated the validity of this approach, and we now present results of a prototype monitoring algorithm that incorporates sICA.

Tools such as LiCSAR (González et al., 2016) are now producing time series of Sentinel-1 data at almost all of the world's active volcanoes. With routine acquisitions, a daisy chain of interferograms is quick to build to the minimum six images required for sICA to be applied in the manner described in the preceding section. sICA can be applied to these six images to determine six latent components that characterize the atmospheric and geophysical signals for a period of steady state activity at the volcano, which may or may not include background deformation. When the next daisy chain interferogram is added to the time series, we perform a simple least squares inversion to fit this image using a combination of the learned components, before



**Figure 17.** Results of our prototype monitoring algorithm. The algorithm attempts to reconstruct a given interferogram (top row) using a linear mixture (middle row) of six independent components determined from the first six interferograms. The mean of the absolute value of the residual is then calculated and plotted for each new acquisition (bottom row). The five interferograms in the top and middle rows correspond to the shaded region. Before the period of unrest, the residual is seen to hover around  $\sim 0.005 \text{ mm}^2$  as the reconstructions approximate the originals well, but with the introduction of a new signal the residual increases approximately sixfold.

calculating the mean absolute residual for each pixel between the actual and recreated data. We postulate that if no new deformation source is present, the mean absolute residual (henceforth referred to as residual) will be low, as no new signals are present. However, should the volcano enter a period of unrest leading to a new deformation source, the residual is likely to increase, which can be used to flag potential activity. Alternatively, if a background deformation source changes rate, this will lead to an uncharacteristic contribution from the component in which it lies, which can also serve as a flag.

Figure 17 shows the results of applying this algorithm to the Wolf time series presented in the previous section. Interferograms from before the period of unrest can be fit using the sICA components with a residual of  $\sim 1 \text{ mm}^2$ , yet when new signals are encountered during the period of unrest (such as subsidence of the caldera floor), the residual increases markedly and provides a clear flag that the volcano has entered a period of unrest. After this abates, the residual returns to pre-unrest levels.

## 7. Discussion

From our comparison of NMF, (t/s)PCA, and (t/s)ICA we have found that in almost all of the synthetic tests performed, sICA is the method most capable of recovering latent signals from a time series of interferograms. In application to real data (the Sentinel-1 Galapagos time series), the results obtained using sICA were highly plausible as the spatial patterns of the recovered sources agreed with signals identified as being due to geophysical processes in other studies.

This result provides justification for the construction of an automatic monitoring algorithm based on sICA, as it facilitates processes such as isolation of a signal of interest or separation of geophysical signals from atmospheric ones.

The recovery of temporally independent time courses initially appears the most attractive approach as we are confident that the geophysical processes of interest at a volcanic site are temporally statistically independent from atmospheric processes yet is hindered by two constraints: First, the results presented in Pinel et al. (2011) suggest that a topographically correlated APS signal will be Gaussian in time (rendering it unrecoverable by

ICA in most cases) and second, (and more importantly) the large number of pixels observed at relatively few times produces a data matrix which is unsuited to ICA.

In contrast, spatially organized data utilize the transpose of the temporal data and are well suited for the FastICA algorithm. However, statistical independence of the spatial nature of the sources is not to be expected in some cases (and in section 4.5 we show that violation of this assumption does affect the fidelity of the sources recovered), and while many sources of geophysical interest are likely to be non-Gaussian, this is not necessarily true for atmospheric signals. Despite these trade-offs, we show that in the majority of the synthetic cases considered, sICA outperforms all the other methods considered and in all but the synthetic stratovolcano case is able to recover useful latent signals. The vulnerability of sICA to signals that is not independent remains problematic for some applications, but as we seek only to characterize baseline behavior for our automatic detection algorithm, we believe that sICA can still be used for our goals. Through application of the method to other data sets in the future, we expect that more information on the importance of these two potential limitations will come to light.

Scope remains for the refinement of sICA with InSAR data. A key part of most ICA algorithms is how the non-Gaussian nature of a signal is measured and various approaches for this exist. In this study, we use a measurement similar to kurtosis, but inspection of estimates of the PDFs of signals (e.g., Figure 2) suggests that for many signals (such as a Mogi source), they may be more clearly identified through using a measure such as skewness. This has been implemented for medical imaging data by Stone et al. (2002) and may be applicable for InSAR data.

ICA also requires the same pixels to be used throughout the time series. However, as the number of coherent pixels changes between interferograms, our method of masking all pixels with an average coherence  $< 0.8$  does not make use of all the information available. A more complex strategy to incorporate the information in pixels that are only coherent in some interferograms remains beyond the scope of this work but may allow for more subtle signals to be recovered with sICA.

## 8. Conclusion

Our study suggests that sICA is the most suited BSS method for use with an InSAR time series at a volcanic center. This is shown through synthetic tests and application to a time series of Sentinel-1 data that spans the 2015 Wolf eruption in which three signals of geophysical interest were isolated. However, aspects of the FastICA algorithm appear suitable for fine tuning to further increase its suitability for use with InSAR data. We introduce a simple algorithm that incorporates sICA to detect when a volcano enters a period of unrest and demonstrate that it would have automatically identified the May 2015 eruption at Wolf volcano. Building from this point, future work on the automatic detection algorithm could allow for identification of different types of unrest (e.g., caused by the acceleration of a previously steady state process) and form an integral part of a system to automatically monitor all of the world's subaerial volcanoes.

## Acknowledgments

The two sets of Copernicus Sentinel data (covering Wolf and Etna volcanoes) were acquired by the European Space Agency (ESA) and were obtained by the authors of this study from the Alaska Satellite Facility archive (<https://www.asf.alaska.edu/>). MEG is supported by the LiCS grant. Figures were prepared in Matplotlib (Hunter, 2007). The Python implementation of FastICA used to perform ICA was modified from Varoquaux et al. (2012). We are also grateful to editor Paul Tregoning, associate editor Emma Hill, and two anonymous reviewers for their helpful and constructive reviews.

## References

- Aloisi, M., Amore, M., Mattia, M., & Patane, D. (2007). Faulting on the western flank of Mt. Etna and magma intrusions in the shallow crust. *Terra Nova*, 19(1), 89–94. <https://doi.org/10.1111/j.1365-3121.2006.00724.x>
- Amari, S. (1999). Natural gradient learning for over- and under-complete bases in ICA natural gradient learning for over- and under-complete bases in ICA. *Neural Computation*, 11(8), 1875–1883. <https://doi.org/10.1162/08997699300015990>
- Bartlett, M. S., Movellan, J., & Sejnowski, T. J. (2002). Face recognition by independent component analysis face recognition by independent component analysis. *IEEE Transactions on Neural Networks*, 13(6), 486–504.
- Bell, A. J., & Sejnowski, T. J. (1995). Information-maximization approach to blind separation and blind deconvolution information-maximization approach to blind separation and blind deconvolution. *Neural Computation*, 11(9)(1994), 1129–1159.
- Bernard, B., Ramon, P., Wright, H., Guevara, A., Hidalgo, S., Pacheco, D., & Vasconez, F. (2015). Preliminary Results on the 2015 Eruption of Wolf Volcano, Isabela Island, Galápagos: Chronology, Dispersion of the Volcanic Products, and Insight into the Eruptive Dynamics. San Francisco, CA: Abstract V31B-3022 presented at 2015 Fall Meeting, AGU.
- Biggs, J., Ebmeier, S. K., Aspinall, W. P., Lu, Z., Pritchard, M. E., Sparks, R. S. J., & Mather, T. A. (2014). Global link between deformation and volcanic eruption quantified by satellite imagery. Global link between deformation and volcanic eruption quantified by satellite imagery. *Nature Communications*, 5, 3471. <https://doi.org/10.1038/ncomms4471>
- Chaussard, E., Bürgmann, R., Shirzaei, M., Fielding, E. J., & Baker, B. (2014). Predictability of hydraulic head changes and characterization of aquifer-system and fault properties from InSAR-derived ground deformation. *Journal of Geophysical Research: Solid Earth*, 119, 6572–6590. <https://doi.org/10.1002/2014JB011266>
- Chaussard, E., Milillo, P., Bürgmann, R., Perissin, D., Fielding, E. J., & Baker, B. (2017). Remote sensing of ground deformation for monitoring groundwater management practices: Application to the Santa Clara Valley during the 2012–2015 California drought. *Journal of Geophysical Research: Solid Earth*, 122, 8566–8582. <https://doi.org/10.1002/2017JB014676>

- Cohen-Weber, J., Burgmann, R., Chaussard, E., Giannico, C., & Ferretti, A. (2018). Spatiotemporal patterns of precipitation-modulated landslide deformation from independent component analysis of InSAR time series. *Geophysical Journal International*, 45, 1878–1887. <https://doi.org/10.1002/2017GL075950>
- Comon, P. (1994). Independent component analysis. A new concept? *Signal Processing*, 36(3), 287–314. [https://doi.org/10.1016/0165-1684\(94\)90029-9](https://doi.org/10.1016/0165-1684(94)90029-9)
- Delacourt, C., Briole, P., & Achache, J. A. (1998). Tropospheric corrections of SAR interferograms with strong topography. Application to Etna. *Geophysical Research Letters*, 25(15), 2849–2852. <https://doi.org/10.1029/98GL02112>
- Ebmeier, S. K. (2016). Application of independent component analysis to multi-temporal InSAR data with volcanic case studies. *Journal of Geophysical Research: Solid Earth*, 121, 8970–8986. <https://doi.org/10.1002/2016JB013765>
- Ebmeier, S. K., Andrews, B. J., Araya, M. C., Arnold, D. W. D., Biggs, J., Cooper, C., & Williamson, J. L. (2018). Synthesis of global satellite observations of magmatic and volcanic deformation: Implications for volcano monitoring & the lateral extent of magmatic domains. *Journal of Applied Volcanology*, 7, 1–26. <https://doi.org/10.1186/s13617-018-0071-3>
- Farr, T., Rosen, P., Caro, E., Crippen, R., Duren, R., Hensley, S., et al. (2007). The shuttle radar topography mission. *Reviews of Geophysics*, 45, RG2004. <https://doi.org/10.1029/2005RG000183>
- Frappart, F., Ramillien, G., Maisongrande, P., & Bonnet, M. P. (2010). Denoising satellite gravity signals by independent component analysis. *IEEE Geoscience and Remote Sensing Letters*, 7(3), 421–425. <https://doi.org/10.1109/LGRS.2009.2037837>
- GEBCO (2015). The GEBCO 2014 Grid. version 20150318. Retrieved 2018-02-16, from www.gebco.net
- Goldstein, R. M., & Werner, C. L. (1998). Radar interferogram filtering for geophysical applications. *Geophysical Research Letters*, 25(21), 4035–4038. <https://doi.org/10.1029/1998GL900033>
- González, P. J., & Palano, M. (2014). Mt. Etna 2001 eruption: New insights into the magmatic feeding system and the mechanical response of the western flank from a detailed geodetic dataset. *Journal of Volcanology and Geothermal Research*, 274, 108–121. <https://doi.org/10.1016/j.jvolgeores.2014.02.001>
- González, P. J., Walters, R. J., Hatton, E. L., Spaans, K., McDougall, A., Hooper, A., & Wright, T. J. (2016). LiCSAR: Tools for automated generation of Sentinel-1 frame interferograms. AGU Fall meeting.
- Hooper, A., Bekart, D., Spaans, K., & Arkan, M. (2012). Recent advances in SAR interferometry time series analysis for measuring crustal deformation. *Tectonophysics*, 514–517, 1–13. <https://doi.org/10.1016/j.tecto.2011.10.013>
- Hotelling, H. (1933). Analysis of a complex of statistical variables into principal components. *Journal of Educational Psychology*, 24(7), 498–520.
- Hunter, J. D. (2007). Matplotlib: A 2D graphics environment. *Computing In Science & Engineering*, 9(3), 90–95.
- Hyvärinen, A. (1999). Fast and robust fixed-point algorithms for independent component analysis. *IEEE Transactions on Neural Networks*, 10(3), 626–634. <https://doi.org/10.1109/72.761722>
- Hyvärinen, A., Karhunen, J., & Oja, E. (2001). *Independent Component Analysis* (Vol. 21). Hoboken, NJ: John Wiley and Sons. <https://doi.org/10.1002/0471221317>
- Hyvärinen, A., & Oja, E. (1997). A fast fixed-point algorithm for independent component analysis. *Neural Computation*, 9(7), 1483–1492. <https://doi.org/10.1162/neco.1997.9.7.1483>
- Hyvärinen, A., & Oja, E. (2000). Independent component analysis: Algorithms and applications. *Neural Networks*, 13(4–5), 411–430. [https://doi.org/10.1016/S0893-6080\(00\)00026-5](https://doi.org/10.1016/S0893-6080(00)00026-5)
- Hyvärinen, A., Oja, E., Hoyer, P., & Hurri, J. (1998). Image feature extraction by sparse coding and independent component analysis. In *Proc. 14th Int. Conf. on Pattern Recognition (ICPR)* (pp. 1268–1273). <https://doi.org/10.1109/ICPR.1998.711932>
- Isomura, T., & Toyozumi, T. (2016). A local learning rule for independent component analysis. *Scientific Reports*, 6, 28073. <https://doi.org/10.1038/srep28073>
- Jutten, C., & Herault, J. (1991). Blind separation of sources. Part I: An adaptive algorithm based on neuromimetic architecture. *Signal Processing*, 24, 1–10. [https://doi.org/10.1016/0165-1684\(91\)90079-X](https://doi.org/10.1016/0165-1684(91)90079-X)
- Karhunen, K. (1947). *Über Lineare Methoden in der Wahrscheinlichkeitsrechnung* (Vol. 37). Universitat Helsinki: SANA.
- Kositsky, A. P., & Avouac, J. P. (2010). Inverting geodetic time series with a principal component analysis-based inversion method. *Journal of Geophysical Research*, 115, B03401. <https://doi.org/10.1029/2009JB006535>
- Lee, D. D., & Seung, H. S. (2000). Learning the parts of objects by non-negative matrix factorization. *Nature*, 401, 788–791.
- Liu, B., King, M., & Dai, W. (2018). Common mode error in Antarctic GPS coordinate time-series on its effect on bedrock-uplift estimates. *Geophysical Journal International*, 214, 1652–1664. <https://doi.org/10.1093/gji/ggy217>
- Lorenz, E. N. (1956). Empirical orthogonal functions and statistical weather prediction (*Scientific Report No. 1, Statistical Forecasting Project*). Cambridge, MA: Massachusetts Institute of Technology.
- Lundgren, P., & Rosen, P. A. (2003). Source model for the 2001 flank eruption of Mt. Etna volcano. *Geophysical Research Letters*, 30(7), 1388. <https://doi.org/10.1029/2002GL016774>
- Lundgren, P., Usai, S., Sansosti, E., Lanari, R., Tesauro, M., Fornaro, G., & Berardino, P. (2001). Modeling surface deformation observed with synthetic aperture radar interferometry at Campi Flegrei caldera. *Journal of Geophysical Research*, 106(B9), 19,355–19,366. <https://doi.org/10.1029/2001JB000194>
- Makeig, S., Bell, A. J., Jung, T. P., & Sejnowski, T. J. (1996). Independent component analysis of electroencephalographic data. *Advances in Neural Information Processing Systems*, 8, 145–151. <https://doi.org/10.1109/ICOSP.2002.1180091>
- Massonnet, D., Briole, P., & Amaud, A. (1995). Deflation of Mount Etna monitored by spaceborne radar interferometry. *Nature*, 375(6532), 567–570. <https://doi.org/10.1038/375567a0>
- McKeown, M. J., Jung, T. P., Makeig, S., Brown, G., Kindermann, S. S., Lee, T. W., & Sejnowski, T. J. (1998). Spatially independent activity patterns in functional MRI data during the Stroop color-naming task. *Proceedings of the National Academy of Sciences of the United States of America*, 95(3), 803–810. <https://doi.org/10.1073/pnas.95.3.803>
- McKeown, M., & Makeig, S. (1998). Analysis of fMRI data by blind separation into independent spatial components. *Human Brain Mapping*, 6, 160–188.
- Meyer, F. J., Webley, P. W., Dehn, J., Arko, S. A., McAlpin, D. B., & Gong, W. (2016). The SARVIEWS project: Automated SAR processing in support of operational near real-time volcano monitoring. AGU Fall meeting.
- Mogi, K. (1958). Relations between the eruptions of various volcanoes and the deformations of the ground surfaces around them. *Bulletin of the Earthquake Research Institute*, 36, 99–134. <https://doi.org/10.1016/j.pnlg.2004.04.016>
- Novellis, V. D., Castaldo, R., Luca, C. D., Pepe, S., Zinno, I., Casu, F., & Solaro, G. (2017). Source modelling of the 2015 Wolf volcano (Galápagos) eruption inferred from Sentinel-1 A DinSAR deformation maps and pre-eruptive ENVISAT time series. *Journal of Volcanology and Geothermal Research*, 344, 246–256. <https://doi.org/10.1016/j.jvolgeores.2017.05.013>

- Paatero, P., & Tapper, U. (1994). Positive matrix factorization—A nonnegative factor model with optimal utilization of error estimates of data values. *Environmetrics*, 5, 111–126. <https://doi.org/10.1002/env.3170050203>
- Pinel, V., Hooper, A., De la Cruz-Reyna, S., Reyes-Davila, G., Doin, M. P., & Bascou, P. (2011). The challenging retrieval of the displacement field from InSAR data for andesitic stratovolcanoes: Case study of Popocatepetl and Colima Volcano, Mexico. *Journal of Volcanology and Geothermal Research*, 200(1-2), 49–61. <https://doi.org/10.1016/j.jvolgeores.2010.12.002>
- Pinel, V., Poland, M., & Hooper, A. (2014). Volcanology: Lessons learned from synthetic aperture radar imagery. *Journal of Volcanology and Geothermal Research*, 289, 81–113. <https://doi.org/10.1016/j.jvolgeores.2014.10.010>
- Porrill, J., & Stone, J. J. V. (1998). Undercomplete independent component analysis for signal separation and dimension reduction. 1–10.
- Rudolph, M. L., Shirzaei, M., Manga, M., & Fukushima, Y. (2013). Evolution and future of the Lusi mud eruption inferred from ground deformation. *Geophysical Research Letters*, 40, 1089–1092. <https://doi.org/10.1002/grl.50189>
- Siebert, L., & Simkin, T. (2013). Volcanoes of the world: An illustrated catalog of Holocene volcanoes and their eruptions. Retrieved from [www.volcano.si.edu](http://www.volcano.si.edu)
- Solaro, G., Acocella, V., Pepe, S., Ruch, J., Neri, M., & Sansosti, E. (2010). Anatomy of an unstable volcano from InSAR: Multiple processes affecting flank instability at Mt. Etna, 1994–2008. *Journal of Geophysical Research*, 115, B10405. <https://doi.org/10.1029/2009JB00820>
- Solem, J. E. (2012). Programming computer vision with Python.
- Spaans, K., & Hooper, A. (2016). InSAR processing for volcano monitoring and other near-real time applications. *Journal of Geophysical Research: Solid Earth*, 121, 2947–2960. <https://doi.org/10.1002/2015JB012752>
- Sparks, R. S. J., Biggs, J., & Neuberg, J. W. (2012). Monitoring volcanoes. *Science*, 335(6074), 1310–1311. <https://doi.org/10.1126/science.1219485>
- Stone, J. (2002). Independent component analysis: An introduction. *Trends in Cognitive Sciences*, 6(2), 59–64. [https://doi.org/10.1016/S1364-6613\(00\)01813-1](https://doi.org/10.1016/S1364-6613(00)01813-1)
- Stone, J., Porrill, J., Porter, N., & Wilkinson, I. (2002). Spatiotemporal independent component analysis of event-related fMRI data using skewed probability density functions. *Neuroimage*, 15(2), 407–421. <https://doi.org/10.1006/nimg.2001.0986>
- Varoquaux, G., Lafaye de Micheaux, P., & van der Walt, S. (2012). Python implementation of the fast ICA algorithms. Retrieved from <https://github.com/GaelVaroquaux/canica/blob/master/canica/algorithms/fastica.py>
- Venzke, E. E. (2015). Global Volcanism Program, 2016. Report on Wolf (Ecuador). Bulletin of the Global Volcanism Network 41:10. Smithsonian Institution.
- Xu, W., Jónsson, S., Ruch, J., & Aoki, Y. (2016). The 2015 Wolf volcano (Galápagos) eruption studied using Sentinel-1 and ALOS-2 data. *Geophysical Research Letters*, 43, 9573–9580. <https://doi.org/10.1002/2016GL069820>