EUROVOLC

European Network of Observatories and Research Infrastructure for Volcanology

Deliverable Report

D9.2: Rapid joint application of Sentinel-1 and GNSS data

<table>
<thead>
<tr>
<th>WP:</th>
<th>Volcano pre-eruptive unrest detection schemes</th>
</tr>
</thead>
<tbody>
<tr>
<td>WP number:</td>
<td>WP9, Task 9.4</td>
</tr>
<tr>
<td>WP Leader:</td>
<td>Dublin Institute for Advanced Studies</td>
</tr>
<tr>
<td>Responsible task leader:</td>
<td>University of Leeds</td>
</tr>
<tr>
<td>Responsible Partner:</td>
<td>University of Leeds</td>
</tr>
<tr>
<td>Authors:</td>
<td>Daniel Juncu, Milan Lazecky, Marco Bagnardi, Matthew Gaddes, Andrew Hooper</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Type of Deliverable:</th>
<th>Report [x]</th>
<th>Demonstrator [ ]</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Prototype  [ ]</td>
<td>Other [x]</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Dissemination level:</th>
<th>Public [x]</th>
<th>Restricted Designated Group [ ]</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Prog. Participants [ ]</td>
<td>Confidential (consortium) [ ]</td>
</tr>
</tbody>
</table>

Programme: H2020
Project number: 731070
CONTENTS

Summary ........................................................................................................................................1
Introduction .................................................................................................................................... 2
1. Automated processing of Sentinel-1 interferograms through the LiCSAR system .................3
2. Automatic Detection of Volcano Deformation ......................................................................... 5
3. GBIS: Software for Bayesian inversion of Surface Deformation Data for Rapid Estimates
   of Source Parameters and Uncertainties .................................................................................. 9
References .................................................................................................................................... 12
Summary

This report summarizes the work that we have done to take the next step to making InSAR data a real-time volcano monitoring tool. To carry out joint rapid inversion of InSAR and GNSS, first we need to rapidly process the InSAR data. We achieve this using the LiCSAR system that provides automatically processed interferograms of Sentinel-1 images, and into which we have integrated the coverage of most European volcanoes. To analyse these data we have developed a machine learning algorithm to automatically detect new deformation in the InSAR data. Testing our approach on real data from an episode of pre-eruptive volcanic unrest shows the capability to detect changes in deformation and flag them as a potential sign of an impending eruption. Finally, we have developed a software, GBIS, to rapidly and jointly invert InSAR and GNSS data using Bayesian inference and estimate properties of sources of volcanic deformation, which we have made freely available online.
Introduction

Geodetic measurements are an important tool that volcano observatories have at their disposal to study volcanoes. To be truly useful to monitor pre-eruptive volcanic unrest progress needed to be made regarding the acquisition, processing and analysis of satellite based geodetic data.

In the past, continuous GNSS measurements were the only real-time geodetic monitoring tool, providing daily (or above) measurement frequency but spatially sparse observations because continuous GNSS networks are costly. InSAR data are complimentary to GNSS measurements because of their greater spatial density but acquisitions were not freely available and for that reason end users had to acquire satellite images and process them themselves. An important change happened with the Sentinel-1 mission, launched by the European Space Agency in 2014, of which data are freely available online. A team at the University of Leeds created an automatized processing chain (see Section 1) that at the point of writing covers acquisitions from most of the European volcanoes and data are made available through an online portal (https://comet.nerc.ac.uk/COMET-LiCS-portal).

With the new abundance of easily available deformation data new challenges arise for data analysis. It becomes increasingly necessary to look at automated analysis tools for large InSAR datasets. Such a tool can be a blind signal separation algorithm (Section 2), that can be used to distinguish different deformation sources from each other as well as from noise sources and to flag changes in activity.

Additionally, we have created a software (GBIS, Section 3) to invert surface displacements for source parameters using static deformation models and made it freely available and in order to facilitate the work of monitoring and research agencies.
1 Automated processing of Sentinel-1 interferograms through the LiC-SAR system

The LiCSAR Sentinel-1 processing infrastructure was initially designed to capture tectonic movements on a large scale with high resolution along plate boundaries, in order to create maps of tectonic strain. Since the infrastructure already exists it is easily expanded over areas of interest for other fields such as monitoring of volcano deformation. At the time of writing, most of the most active European volcanic areas are covered, such as the volcanoes of Iceland, Italy, and the Spanish Canary Islands.

The LiCSAR system currently uses ESA’s Sentinel-1 wideswath (IWS) SLC data to generate large-scale interferograms multilooked to achieve resolution of around 100x100 m per pixel. The output spatial units are called frames, which we have defined globally. Frames are defined as a collection of Sentinel-1 bursts imaged during the satellite’s pass within a given orbital track. Usually frames consist of 39 bursts (13 bursts within each of three observation swaths) — such a standard frame covers an area of around 220x250 km. While frames are defined globally across the Earth, LiCSAR is performing the InSAR analysis on a selection of frames over active tectonic and volcanic areas (e.g. Alpine-Himalayan Belt, East African Volcanic Rift etc.). Frames covering active volcanic areas are treated in prioritised manner, receiving updates continuously. The update frequency is getting increased depending on the development of technical possibilities.

The LiCSAR processing chain uses various open-source tools but the core processing functionality is based on mature commercial software for processing SAR data, GAMMA. For geocoding and topographic phase screen generation, NASA’s SRTM DEM is used, though some frames are prepared using other models (Aster GDEM, JAXA ALOS World 3D or DLR TanDEM-X WorldDEM).

The basic InSAR products generated by LiCSAR are: unfiltered wrapped and filtered unwrapped interferograms, coherence maps and complementary specialized frame images (including incidence angle map or ENU files needed for motion vector extraction). All products are georeferenced to the WGS-84 geographic coordinate system. The InSAR products are distributed in geotiff format containing original values as well as generated RGB bitmaps for a quick visualization. Some products of a special interest (e.g. co-seismic interferograms) are converted into Google Earth KML files. The InSAR products are shared through both the custom COMET LiCS products web portal and the European Plate Observing System (EPOS) system, with NERC’s (the UK’s Natural Environment Research Council) Data Repository to be holding the data products in near future.

It takes around 2 hours to make the InSAR products available online since the start of the processing. There are various optimisation strategies existing or being developed leading to generate and provide the InSAR outputs as early as possible. This is necessary especially in case of monitoring active volcanoes. Here, the generated interferograms are processed by machine learning techniques with the aim to detect surface displacements on monitored volcanoes. In order to achieve fast results, priority areas of interest are identified and their related Sentinel-1 data are directly downloaded through a Relay Hub, once they are downlinked from the satellite to the ESA facility and processed towards Level 1 SLC data (several hours after the physical acquisition). LiCSAR algorithms are also tracking the physical existence of new
data through Copernicus Scihub service and uses it for data backfilling as an alternative data source. NASA’s Copernicus data mirror (at the Alaska Satellite Facility) is also being used as an alternative data source.

Once a new acquisition arrives, it is decomposed into pre-defined burst units and registered in a local database that handles burst and frame definitions. Images including bursts that form a given frame are extracted and merged into frame images. These are co-registered to a primary frame image (a master image) that was set during the initialization of the frame, beforehand. The co-registration process (including spectral diversity and other necessary corrections) has the longest processing runtime - around 1 hour per a frame image. Therefore, the LiCSAR system stores the generated lookup table for faster reprocessing of the image in the future, if needed. Once co-registered, the interferograms are formed combining the new image with three chronologically previous ones. This way is suitable for interpretation and for further use of the interferograms in multitemporal InSAR processing methods based on small baselines strategy (e.g. the small baseline approach currently implemented into the custom LiCS-BAS chain). The interferogram unwrapping is performed using an optimized snaphu approach. After the processing, results are shared through the website in the form of georeferenced TIFF files and preview bitmap rasters. Currently the website resides at https://comet.nerc.ac.uk/COMET-LiCS-portal, Figure 1.1 shows a screenshot of the web portal.

![Figure 1.1: Screenshot of COMET LiCS Sentinel-1 InSAR portal web map of processed frames containing LiCS InSAR products.](image)

Future goals for LiCSAR monitoring of volcanoes include expanding the coverage to active volcanic regions worldwide, including highly active European overseas territories such Réunion, as well as increased resolution and near real-time monitoring for data from volcanoes that are deemed high-interest. Volcanic islands that are not covered by the IWS mode will be monitored using Sentinel-1’s StripMap mode.
2 Automatic Detection of Volcano Deformation

2.1 Introduction

With increasing abundance of rapidly available InSAR data and hundreds of active volcanoes to be monitored, there is a requirement to move towards automated monitoring and deformation detection. Blind signal separation (BSS) methods are a class of unsupervised learning algorithms that can be used to isolate latent signals from InSAR data time series containing overlapping signals from several contributing sources.

We have investigated three different types of BSS algorithms in order to identify the one that best suited for dealing with InSAR timeseries (Gaddes et al., 2018). Based on these results, an algorithm was built that is able to automatically detect signs of volcanic unrest in a time series of interferograms (Gaddes et al., 2019).

2.2 Blind signal separation methods

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

\[ \phi = W \{ \phi_{\text{def}} + \phi_{\text{orb}} + \phi_{\text{atm}} + \Delta\phi_{\theta} + \phi_{N} \} \]  

(2.1)

where \( \phi_{\text{def}} \) is the phase change due to deformation of the ground surface, \( \phi_{\text{orb}} \) is the phase due to errors in the location of the satellite at each acquisition, \( \phi_{\text{atm}} \) is the phase change due to changes in the atmospheric delay, \( \Delta\phi_{\theta} \) is the phase due to misestimation of the look angle, \( \phi_{N} \) 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 \( \phi_{\text{def}} \) 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 and Herault, 1991). Gaddes et al. (2018) consider three different types of BSS methods, namely principal component analysis (PCA), independent component analysis (ICA) and nonnegative matrix factorization.

PCA and ICA are methods that both attempt to separate multivariate signals into a set of subcomponents. In PCA this set is formed by orthogonal components where the primary component is that which has the largest variance in the mean-centered mixtures and each subsequent component has the largest variance and is orthogonal to the preceding components. ICA decomposed the signal into constituents under the assumption that each component has a non-Gaussian probability distribution. This is based on the central limit theorem, which stipulates that by summing several independent non-Gaussian sources, the
resulting mixture has a more Gaussian PDF than any of the constituent sources (Hyvärinen and Oja, 2000). The NMF method is a different approach from the first two, here a nonnegative data matrix of mixtures is factorised into two nonnegative matrixes so that the difference between the data matrix and the constituent matrices is minimized.

In a processed InSAR dataset generally the main components to be identified are the deformation signal, $\phi_{\text{def}}$ (the signal of interest in volcano monitoring) and the principal InSAR noise source due to variations in atmospheric delay $\phi_{\text{atm}}$.

2.3 Automatic detection

Gaddes et al. (2019) built upon the preceding study and incorporated the ICA algorithm into an automatic detection algorithm for changes in volcanic deformation. The threefold approach consists of 1) isolating signals of geophysical interest from interferograms, 2) using the components learned in stage 1 to characterise the baseline data and 3) ingestion of new interferograms to determine whether the signals present have deviated strongly enough from those in the baseline data to warrant flagging the volcano as having entered a period of unrest. The algorithm was tested on synthetic datasets as well as real data from the Sierra Negra volcano (Galapagos, Ecuador).

2.4 Results

The independent component analysis algorithm, or rather, the derived spatially organized independent component analysis algorithm (sICA) proved to be the best suited BSS method of the tested varieties. Applied to data from the 2015 eruption of Wolf volcano (Galapagos Islands, Ecuador), the algorithm is able to separate three distinct deformation signals as can be seen in Figure 2.1. The results show that the sICA algorithm may be a good candidate to be used in volcano monitoring.

The results of applying the machine learning algorithm to data from the Sierra Negra volcano are discussed in detail in Gaddes et al. (2019). Figure 2.2 shows the results of applying our automatic detection algorithm to the time series. The most striking feature is the flagging of the time course of the isolated signal IC1, as indicated by the orange colouring of the points, due to the rate of inflation increasing. Closer to the eruption, other time courses also exhibit unusual behaviour which is flagged as a sign of unrest (e.g. the time course of IC3 from interferogram onwards), and may be due to processes such as slip on the intra-caldera faults causing slight changes in the shape of the uplift pattern, which then requires different use of the baseline components during the inversion step. Automatic detection of the new large signals associated with the onset of the eruption captured in interferogram 97 is achieved through the inability of the learned components to fit these new signals, which causes both measures of the residual to increase rapidly. The results show that the presented automatic detection algorithm is able to detect signs of volcanic unrest due to both a change in rate of a pre-existing signal, and the emergence of a new signal and that it has the potential to be a very useful new volcano monitoring tool.
Figure 2.1: Results of sICA applied to the time series shown in Gaddes et al. (2018), 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 from Gaddes et al. (2018).
Figure 2.2: The results of applying our automatic detection algorithm to a time series of Sentinel-1 interferograms which cover the final \(\sim 3.5\) years of inflation before the 2018 eruption of Sierra Negra (the anomalous points associated with the co-eruptive interferogram can be seen on the right hand y-axis). Roughly every five interferograms are shown, but some liberty is taken to ensure those of particular interest are visible (e.g. 55, 56 and 97). The components are initially used in a similar fashion before and after the switch to the ingestion phase (marked by the black vertical line), before more pronounced deformation from around interferogram 65 causes IC1 to be flagged as having deviated significantly from the baseline data. The residual when the final co-eruptive interferogram is fitted is an order of magnitude larger than seen previously, and is omitted from the RMS residual plot for clarity. The two high values of RMS residual for interferograms 55 and 56 are due to a strong atmospheric signal in the acquisition common to the two. Figure from Gaddes et al. (2019).
3 GBIS: Software for Bayesian inversion of Surface Deformation Data for Rapid Estimates of Source Parameters and Uncertainties

3.1 Introduction

Inverting geodetic data for deformation sources using elastic models is a standard problem in the field of volcano deformation. A well established approach is using the method of Bayesian inference, a stochastic method that allows for using data errors and prior knowledge of the studied system to estimate probability distributions of model parameters. While this method is commonly used, only few freely available software packages exist that run Bayesian inference out-of-the-box on GNSS and InSAR data, incorporating the estimation of InSAR data errors and let the user produce rapid estimates of model parameter estimates and their uncertainties. For volcano observatories, that often lack resources to undertake their own software development work it is very helpful if software is made freely accessible to them that addresses monitoring needs, easily used and maintained by a different institution. To this end, Bagnardi and Hooper (2018) have developed a Bayesian inference algorithm and created a MATLAB code called Geodetic Bayesian Inversion Software (GBIS) that is available here. For more detail, the reader is referred to Bagnardi and Hooper (2018).

3.2 Methods

Bayesian inference in the context of modelling of geodetic data refers to inferring the probability of parameter values of (commonly) physical models that links sources of deformation in the subsurface to geodetic observations at the surface, given the information contained in those observations. The method is based on Bayes’ rule which states that the probability of a set of model parameters \( p(m|d) \), given the observations is dependant on the prior (i.e., before the observations are considered) knowledge \( p(m) \) about parameter values, the likelihood of the data \( p(d|m) \) (essentially, the difference between the actual observations and a set of "predicted" observations based on the given set of model parameters) and the marginal likelihood \( p(d) \) which is independent of the model parameters (Gelman et al., 2014). In equation form Bayes’s rule reads

\[
p(m|d) = \frac{p(m) p(d|m)}{p(d)}.
\]

There are various ways to sample from the posterior probability distribution \( p(m|d) \), a common approach being through so called Markov Chain simulation in which the parameter space is explored by testing sequences of random values within said space (Gelman et al., 2014). A particularly popular algorithm within this group is the Metropolis algorithm (Metropolis et al., 1953), in which random samples \( m' \) are drawn from the parameter space and either accepted or rejected based on the ratio of the probability of \( m' \), \( p(m'|d) \) to that of the previously accepted parameter \( p(m^{t-1}|d) \), the ratio \( r \) of which can be written as

\[
r = \frac{p(m'|d)}{p(m^{t-1}|d)}.
\]

A sample \( m' \) will be accepted with a probability of \( \min(r,1) \). The ratio in Equation 3.2 is convenient.
as when inserting Equation 3.1 into Equation 3.2, the denominator will always cancel out because, as
mentioned above, it is independent of the value of \( m \). Additionally, in the context of volcano deformation
(or pretty much any crustal deformation problem), not much prior information is known about model pa-
rameters and only uninformative, uniform prior probability distributions, that establish parameter bounds
but nothing more, are used. This way, the \( p(m) \) terms cancel out as well and Equation 3.2 simplifies to

\[
 r = \frac{p(d|m')}{p(d|m_{t-1})}.
\]  

(3.3)

This means that in order to sample from the posterior distribution, the only requirement is that nothing
but the weighted residuals (the difference between the actual and model-predicted observations, weighted
by the data uncertainties) needs to be calculated for each randomly proposed model parameter value
\( m' \) in order to decide whether a sample is rejected or accepted, and the sequence of accepted samples
approximates the wanted posterior distribution \( p(m|d) \).

Within GBIS, the parameter space of each parameter \( m \) is explored by a random walk with steps drawn
from a uniform distribution with limited step size, such that

\[
m' = m_{t-1} + a_n \Delta m,
\]  

(3.4)

where \( a_n \) is a random draw from the uniform distribution \( U(-1,1) \). It is common in deformation mod-
elling to constrain parameter values for inversions by limiting bounds based on prior assumptions. This
is implemented by replacing every value outside of the parameter bounds by a value that is within the
bounds, but with the same absolute distance to the bound as the replaced value:

\[
m^* = 2UB - m' \quad \text{if} \quad m' > UB
\]

\[
m^* = 2LB - m' \quad \text{if} \quad m' < LB
\]  

(3.5)

In order to maximize efficiency in exploring the parameter space, the random walk step size \( \Delta m \) needs
to be tuned so that it will yield optimal convergence speed. This is done by tracking the acceptance
rate adjusting it towards the optimal acceptance rate of around 20\%, which is done at regular, custom
intervals.

An overview of the algorithm is given in Figure 3.1.

3.3 Usage instructions

The software package contains a manual which gives detailed instructions on how to run the code. The
main requirements are properly formatted files containing GNSS and/or InSAR data. GNSS files need to
contain for each station: longitude, latitude, horizontal North-South displacement, North-South displace-
ment error, horizontal East-West displacement, East-West displacement error, vertical displacement, ver-
tical displacement error.
Figure 3.1: Schematic representation of the proposed Bayesian inversion approach, including the Markov chain Monte Carlo method-Metropolis-Hastings iterative algorithm. Each step is fully described in the text. $b$ is a random value from a uniform distribution within the range $[0,1]$. Note that for $i < 20,000$, we perform a sensitivity test (see section 2.2) every $N_s$ iterations $j$ to tune the step size $\Delta m^j$. GNSS = global navigation satellite system, InSAR = interferometric synthetic aperture radar. Figure from Bagnardi and Hooper (2018).

The InSAR data must contain for each data point: longitude, latitude, phase, incidence angle, as well as the heading angle of the satellite. InSAR uncertainties do not need need to be given as input, the full variance-covariance matrix will be simulated on-the-fly (with user interaction).
References


