 doi:10.20944/preprints202206.0097.v1

Technical Note

Assessing the accuracy of deformation time series using an automated InSAR pipeline

Kelly M. Olsen [1]*, Matthew T. Calef [1] and Piyush S. Agram [1]

[1] Descartes Labs, 1607 Paseo de Peralta, Suite B, Santa Fe, NM, 87501, USA
* Correspondence: kelly.olsen@descarteslabs.com

Abstract: InSAR and associated analytic methods can measure surface deformation from low earth orbit with a claimed accuracy of centimeters to millimeters. The realized accuracy depends on the area being measured and on the choice of analytic method, suggesting one choose a method in response to the area being measured. Here we consider a specific fixed analytic method and compare the results it produces to measurements gathered from other means in a variety of settings. In particular we compare Sentinel-1 InSAR with GPS at the Kilauea volcano around the 2018 eruption, with GPS in the city of Arica, Chile, and with public survey data at a decommissioned tailings mine. In addition, we compare two independent Sentinel-1 InSAR analyses for a railway station in Oslo, Norway. Our goal is estimate the accuracy of a fully automated Sentinel-1 InSAR pipeline in various settings. Our conclusions are that centimeter level accuracy is a reasonable claim in many, but not all settings, and that accuracy is typically not lost by using an automated pipeline, instead of hand-selecting and tuning parameters.

Keywords: InSAR; deformation; validation; GPS; consistency

1. Introduction

Space-based InSAR can measure the relative line-of-sight (LOS) deformation history of an area of interest (AOI) with a nominal accuracy centimeters to millimeters [1]. The actual accuracy depends on the AOI and time period measured, as well as the choice of analytic method. Historically InSAR analyses were performed on demand where an InSAR practitioner could assess results and modify the analysis in response to the estimated quality of results; and even iterate if needed. While this process allows for more control over the quality of the derived deformation results, it is time-consuming, computationally expensive, and may introduce biases. With the advent of inexpensive scalable computing resources, it is now possible to run InSAR continuously over large areas, e.g. the InSAR Norway project [2] and European Ground Monitoring Service (EGMS) [3]. These large-scale InSAR data-sets can be extremely valuable in understanding geological process, monitoring municipalities, and assessing the health of large man-made structures. Further, these data-sets allow one to develop an alerting mechanism for areas of concern. With this as motivation, we consider the question: can InSAR results generated as part of an autonomous pipeline maintain high levels of accuracy, similar to their more manual counterparts?

Previous studies have attempted to quantify the accuracy of InSAR measurements mainly using GPS and/or leveling data. The studies report a wide range in accuracy, from sub-millimeter to 10+ centimeter accuracy, depending on the deformation retrieval method, the location of study, and the data source for comparison [4–14]. These analyses have typically been run as an one-off analyses, with manually tuned processing to eliminate unwrapping errors [15], or to vary parameters and thresholds to obtain the best results following an assessment of the initial results.

"Blind" experiments have also been performed, such as [5], to try and eliminate human-derived biases. However, in the case of [5], the experimental set up itself was complex and

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 7 June 2022 doi:10.20944/preprints202206.0097.v1

costly (two pairs of corner reflectors installed, one pair fixed and the other on a platform that was shifted horizontally and vertically with sub-millimeter accuracy.) While impressive, this doesn't reflect how InSAR would be deployed at scale, nor the range of types of AOIs on which InSAR would be used.

In this study, we run a single InSAR method over several AOIs where GPS or survey data are available to use as sources of independent ground-motion comparisons. In addition, we compare two relatively independent InSAR analysis of the same area to each other with the goal of assessing consistency within our pipeline. In all cases, we do not tune our analysis per AOI, as our goal is to discern an estimate of the accuracy of an automated pipeline-based InSAR analysis. Our InSAR automated pipeline contains no influences or outside knowledge about the deformation signal, making the pipeline "blind" to the deformation. An automated pipeline can save costs associated with computational resources and human time, as well as eliminate biases that may appear when hand-selecting parameters. While this work does not provide a complete characterization of the accuracy of a fully automated InSAR system, it does provide a valuable insight into the accuracy of an automated deformation retrieval pipeline and highlights some areas of possible limitations. We show that our accuracy levels are similar to those in many prior published results, indicating that an automated InSAR pipeline is likely sufficient for many monitoring use cases.

2. Methods

2.1. InSAR Methods

Our InSAR pipeline starts with Sentinel-1 [16] and begins by geocoding complex SAR imagery on to a UTM grid that has a 10m north-south spacing and a 2.5m east-west spacing [17]. We do not use enhanced spectral diversity (ESD), as our analyses included in this manuscript all fit within a single Sentinel-1 burst footprint. The anisotropy in the UTM grid spacing aims to capture the corresponding anisotropy in the range and azimuth resolutions in IW-TOPS collections [18]. By geocoding to a well known geospatial coordinate system, we can use the Descartes Labs computing platform [19], which allows us to bring to bear scalable compute resources.

For each AOI we extract a stack of co-registered complex SAR scenes. Based on the size of the input AOI we may downsample these input complex data. If we downsample, we do so first in the east direction until the spacing is the same in the east and north directions; any further downsampling is done in both the east and north directions. Following the method described in [20], we obtain a rank-one approximation to the sample correlation matrix per pixel within our stack of co-registered complex SAR images. In forming the SCM we use a kernel that is determined by any anisotropy in the spacing of the input complex SAR data. We consider the agreement of the phase of this rank-one approximation with the phase of the SCM as the temporal coherence (cf [21]). We select pixels that exceed a temporal coherence of 0.6. If there is any remaining anisotropy in the spacing of these "good" pixels, we downsample among them to generate a point cloud with an isotropic spacing. The phase of the complex vector for the rank-one approximation to the SCM becomes the per-epoch input phase for what follows. In the context of terminology laid out in [3], the scatterers identified by our method can be classified as a combination of high signal-to-clutter ratio (SCR) permanent scatterer pixels and distributed scatterers.

We form a Delaunay Triangulation of the selected pixels and consider the histories of differential phase across the links in the triangulation. Following [22], we employ an extended minimum cost flow (EMCF) algorithm. The first step in this algorithm aims to try to ensure historical consistency in the differential link-phase. To do this one considers each SAR collection as a "configuration" point in a two-dimensional space. One coordinate is the time of the collection, the other coordinate is the component of the orbit that is perpendicular to the LOS. These two coordinates are in different units forcing one to choose how to scale them. A natural approach is to estimate the decorrelation time and the decorrelation baseline and scale them both so a value of, say, 1 indicates decorrelation

 doi:10.20944/preprints202206.0097.v1

in either direction. However, the orbit control for Sentinel-1 is sufficiently good that this approach leads to a set of configuration points that is "stretched out" in time. The next step is to form a second Delaunay Triangulation of this set of points, this "stretched out" configuration point-set leads to a Delaunay Triangulation where only three links, or two triangles, intersect most time slices. In this case the dual graph has no cycles, limiting the ability to establish consistency. In response to this we scale the configuration points-set so that the standard deviations of the time and baseline coordinates are the same before forming the Delaunay Triangulation. Using a minimum cost flow algorithm [23], we unwrap the differential phase history for each link. Our approach of processing links first is consistent with unwrapping methods listed in [3] and can be interpreted as another variant.

We then use the unwrapped differential phase history of the links to spatially unwrap interferograms with the same minimum cost flow method. We almost always unwrap interferograms formed from the phase differences from successive epochs for the following reason: if one uses a redundant interferogram network, one has to resolve possible inconsistencies in the unwrapped interferograms. Any continuous function for recovering per-epoch data from a redundant interferogram network, e.g. least-squares, will have the effect of averaging any unwrapping errors into the final result. Because minimum cost flow is an integer programming method the unwrapping "errors" are discrete, and such averaging will make these errors less-pronounced, which makes these errors harder to detect later. Instead, we search for a minimum spanning tree of the graph of Delaunay Triangulation of the configuration point-set. We use a cost function based on a heuristic estimation of decorrelation along links in configuration space. Again, because of the highly controlled orbit, the highest cost is temporal decorrelation and the minimum spanning tree forms interferograms connecting successive epochs. At this point we remove the component of the recovered deformation that projects onto the perpendicular baseline, as this is likely caused by DEM error.

The only part of our pipeline that isn't, so far as we know, described in the literature is our link-based unwrapping error detector. In this method we look at the differential unwrapped phase history across each link. If, between successive epochs, the unwrapped phase jumps by more than π and if that jump is more than 3σ from the median of the epoch-to-epoch phase changes, we assess a correction of 2π. Corrections from one epoch to the next are applied by solving the least squares problem to establish point corrections from link corrections. We do this in an iterative fashion where we assess and correct erroneous phase jumps in the time history, stopping when we detect no further errors.

For the current work we do not attempt to remove atmospheric phase screens (APS) or apply any form of temporal smoothing. APS can be a significant and unwanted contribution to the final retrieved deformation and ought to be removed in production settings. There are several reasons we do not attempt to address APS in the current work. First, in almost all cases the problem of APS removal can be considered independently of, and after, the unwrapping problem. Second, APS removal methods that ingest ancillary weather data would be prohibitive to run at scale leaving methods that are largely based on temporal smoothing, e.g. [24] or spatio-temporal filtering in combination with a polynomial signal model, e.g [3] both of which can have the effect of smoothing over undetected unwrapping errors. Because our goal is to highlight the accuracy of a fully-automated deformation retrieval method, we present data without any APS removal method applied.

In addition we assess uncertainty of deformation due to estimated phase noise using the method described in [25]. The presented pipeline does not use any temporal or spatial smoothing aside from the kernel used when forming the SCM, and we perform no additional unwrapping-error correction apart from our automated link-based unwrapping error detector. We consider this deformation retrieval "blind" in that the analysis parameters, except spatial down-sampling, are fixed and no hand-tuning was performed. Any down-sampling is chosen in a deterministic way based only on AOI size to reduce computation time.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 7 June 2022 doi:10.20944/preprints202206.0097.v1

2.2. Comparison Methods

Global Positioning System (GPS) data is one of the Global Navigation Satellite Systems (GNSS) that uses an array of satellites to provide latitude, longitude, and altitude within a few centimeters to meters [26]. For our comparisons with GPS, we project the GPS-derived three-dimensional motion onto the satellite's LOS incidence vector. One of our comparisons relies on publicly available survey data of *total station measurements* [27]. The survey data present horizontal displacement as stereograms, which we translated to east and north by hand. Uncertainty by component for GPS and correlation between uncertainties are reported with each datum in the GPS data sets. The survey data report a vertical uncertainty of 10mm and a horizontal uncertainty of 25mm, no correlation data were given so we treat the uncertainties as uncorrelated.

InSAR-derived deformation is relative in both time and space, requiring pinning down two degrees of freedom. For InSAR analyses where we have external data sources at two points (e.g. 2+ GPS stations or leveling measurements), we reference the InSAR data to one of the points and compare that to the difference of the external data at the two points. It is common to reference InSAR results to the first epoch. One could reference the external data to the datum closest in time to the first InSAR epoch. This has the drawback that any noise or error at the first epoch of either the InSAR results or the external data will contribute a fixed offset for the rest of the data. Instead we use the median of the difference of the two as a fixed offset. This reflects a use case where InSAR is used to assess change over time with an understanding that the absolute value will depend on the time window. This also means that the results we present are "best possible" in some regards.

3. Results

3.1. Kilauea, HI

We compare our InSAR results to five GPS stations, located around and within the area that exhibited a sharp increase in subsidence at the end of April 2018. We choose the GPS station MANE for our reference point (Fig. 1), as it is in an area that is largely non-deforming relative to the volcano, and compare to the other four GPS stations (CNPK, KOSM, OUTL, and MALU). As noted above, we calculate the line-of-sight deformation for each GPS point, then subtract the MANE deformation signal from each of the other four GPS signals. Similarly, our InSAR deformation signal is generated by subtracting the line-of-sight deformation at the InSAR point closest to MANE from the points closest InSAR points to each of the other GPS stations. Note that, in this case, we compute an offset based on the median difference in the InSAR and GPS data only for epochs before the obvious subsidence. In this case there is a clear event, whose change we would like to track.

We show a comparison of GPS and InSAR results in Figure 2. The error bars on the difference show uncertainty and are based on our ensemble methods, the reported GPS uncertainty for each station, and, for lack of better information, the assumption that uncertainties between GPS stations and between our ensemble uncertainty are independent. While two of the stations comparisons (CNPK-MANE and OUTL-MANE) show high agreement, the other two have show differences on order of ~10 cm beginning in late May, following the initial rapid collapse of the summit. However, average RMS GPS/InSAR error was still low at each of the sites, with an average of 6.3 cm 1. The differences in the GPS and InSAR time series signals are likely due to unwrapping errors given the rapid subsidence and vegetation. We have not smoothed the GPS signal, which may contain errors, e.g. the GPS signal increases sharply before decreasing. APS is also likely a factor here – differences in spacing between MANE and the other GPS stations range from ~5.2-8.2 km, and differences in elevation from ~6-130 m, which can increase atmospheric differences.

For this AOI we used imagery collected from swath 2 of orbital track 124. We sampled the geocoded data at 20m intervals. We formed the SCM using a kernel whose full-width half-max was six pixels. In this example we further downsampled in the east direction by a factor of four simply to reduce the number of pixels.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 7 June 2022 doi:10.20944/preprints202206.0097.v1

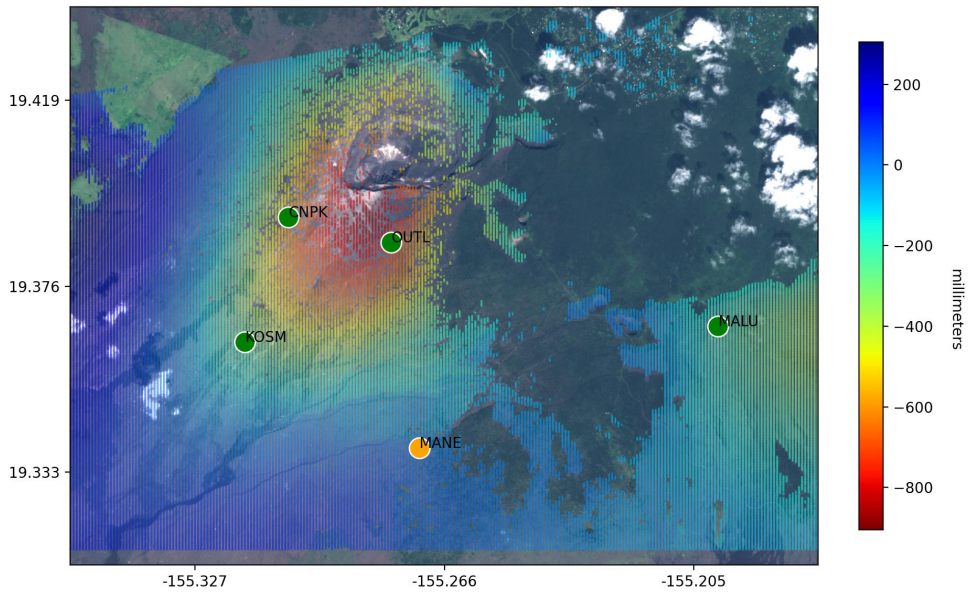

Figure 1. Unfiltered cumulative deformation over 2018 in mm, overlaid on optical data, showing the reference (orange) point and test (green) points for the Kilauea Volcano AOI.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 7 June 2022 doi:10.20944/preprints202206.0097.v1

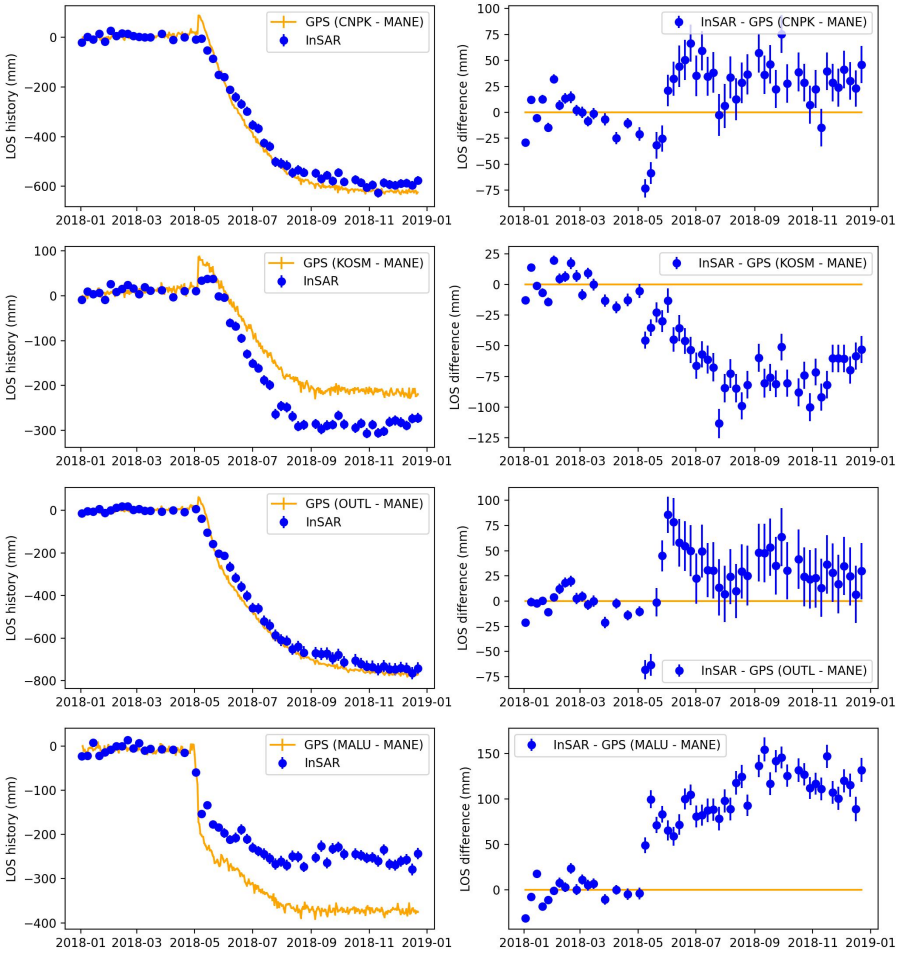

Figure 2. Deformation histories as measured with InSAR and GPS (left) and their difference (right) for the Kilauea Volcano AOI.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 7 June 2022 doi:10.20944/preprints202206.0097.v1

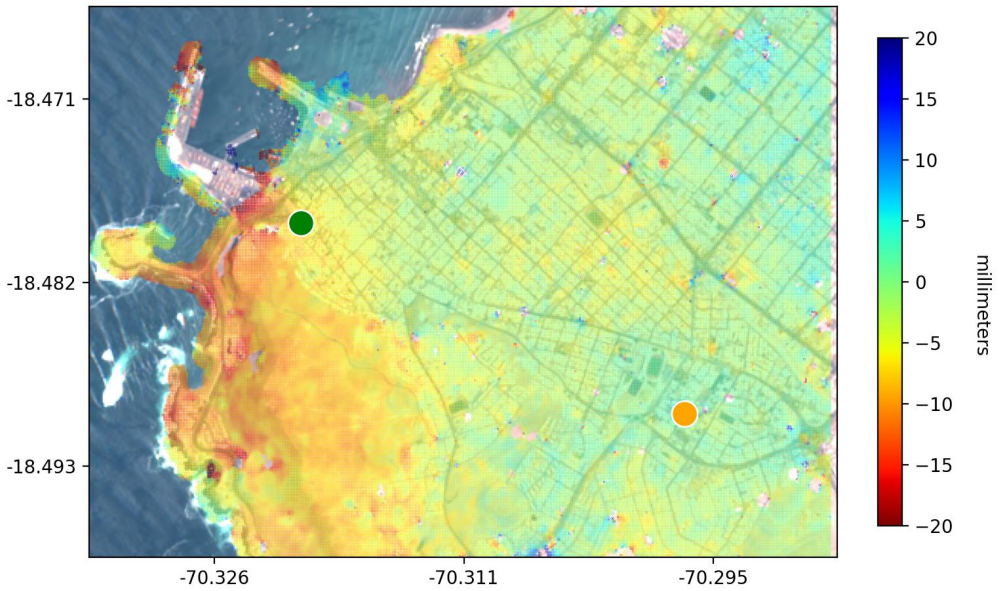

Figure 3. Unfiltered cumulative deformation over 2021 in mm, overlaid on optical data, showing the reference point at the UTAR GPS station (orange dot) and test point at the BN15 GPS station (green dot) for the Arica Chile AOI.

Table 1. Average RMS error in cm for each comparison.

Comparison	RMS error (cm)
InSAR-GPS (CNPK-MANE)	3.55
InSAR-GPS (KOSM-MANE)	5.88
InSAR-GPS (OUTL-MANE)	4.07
InSAR-GPS (MALU-MANE)	11.69
InSAR-GPS Kilaeua average	**6.29**
InSAR-GPS (UTAR-BN15)	0.18
InSAR-leveling (BD2-BD1)	0.57
InSAR-leveling (BD3-BD1)	0.82
InSAR-leveling (BD4-BD1)	0.81
InSAR-leveling (BD5-BD1)	0.71
InSAR-leveling (BD6-BD1)	0.57
InSAR-leveling (BD7-BD1)	0.38
InSAR-leveling average	**0.64**
InSAR-InSAR (Oslo)	0.0957

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 7 June 2022 doi:10.20944/preprints202206.0097.v1

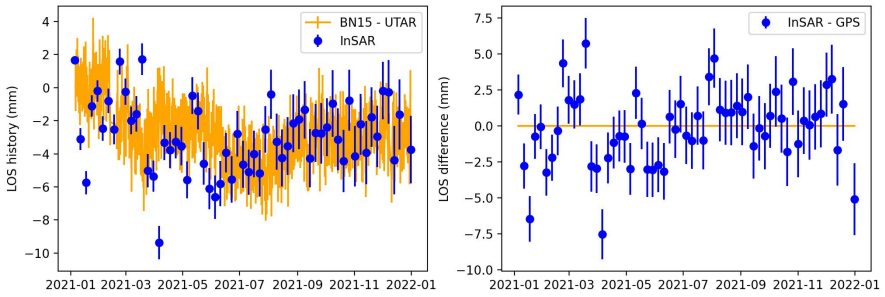

Figure 4. Deformation histories as measured with InSAR and GPS (left) and their difference (right) for the Arica AOI.

3.2. Arica, Chile

The next AOI we considered is in Arica, Chile, which contains two GPS stations, BN15 and UTAR shown in Fig. 3. These stations offer a useful comparison because there is very little deformation over this time period (2021-2022), which allows us to test how well we can measure small-scale changes, unlike Kilaeaua where the cumulative relative deformation was almost a meter in places. For this comparison, we reference our InSAR deformation to the location of the UTAR station and report the relative motion between UTAR and BN15 stations for both our InSAR results and the GPS data.

Fig. 4 shows the difference between the spatially referenced InSAR results and GPS derived deformation. As with the Kilauea AOI, we use error bars to represent the uncertainty in the difference between the InSAR and GPS results, and have once again assumed that uncertainties at BN15 and UTAR are independent from each other and from the InSAR ensemble standard deviation.

The test and reference points are a little over a mile apart, so APS is likely a contributor, but even without addressing APS, InSAR and GPS agree to within a centimeter. The mean and median differences in magnitude of the signals are both ~2 mm. Once again, the GPS measurements have also not been smoothed or time-averaged in any way, so some of the differences may be due to errors in the GPS measurements as well as errors in the InSAR measurements.

We used swath 3 of orbital track 54 for this example. We used the input geocoded data with an east-west spacing of 2.5m and a north-south spacing of 10m. We used a kernel for the SCM with a full-width half-max of 6 pixels in the north-south direction and 22 in the east-west direction. After selecting pixels whose temporal coherence was greater than 0.6, we decimated in the east-west direction by 4. This provides posted results every 10m.

3.3. Bose Lake Dam

The Bose Lake Dam AOI is roughly 100 miles northeast of Vancouver, Canada. This AOI allows us to compare InSAR results to intermittent total survey measurements that are publicly available in the Teck report found online [27]. The Bose Lake Dam has seven pins, BD1 through BD7, along the crest of the dam. This is shown in Figure 5. Between July of 2016 and January of 2020 there were six total-station survey measurements that included all the pins. We reference our InSAR results to BD1 and report the relative motion between the other pins and BD1.

The InSAR and survey results and their differences are shown in Fig. 6. Winter conditions (e.g. snow, heavy rain) provided unusable SAR imagery during the months of December-April, so these months have been excluded from the analysis. This is automatically detected by looking at seasonal trends in coherence. Additionally, our ensemble results showed that the estimated uncertainty grew substantially during the winter months.

As noted, the report from which we derived these survey measurements [27] listed the vertical uncertainty at 10mm and the horizontal uncertainty at 25mm. The difference

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 7 June 2022 doi:10.20944/preprints202206.0097.v1

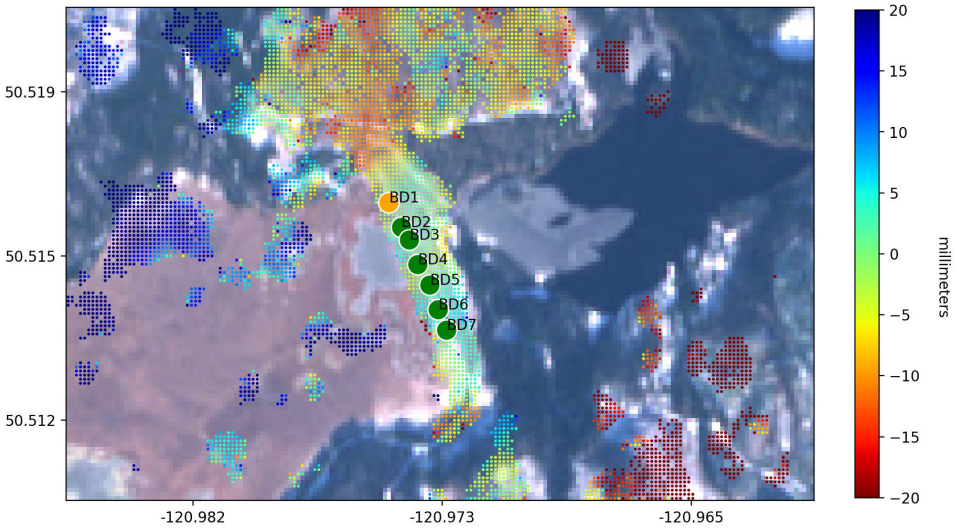

Figure 5. Unfiltered cumulative deformation over from July 2016 through January of 2021 in mm, overlaid on optical data. The reference point is the BD1 pin (orange point) and the test points are bins BD2 through BD7 (green points).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 7 June 2022 doi:10.20944/preprints202206.0097.v1

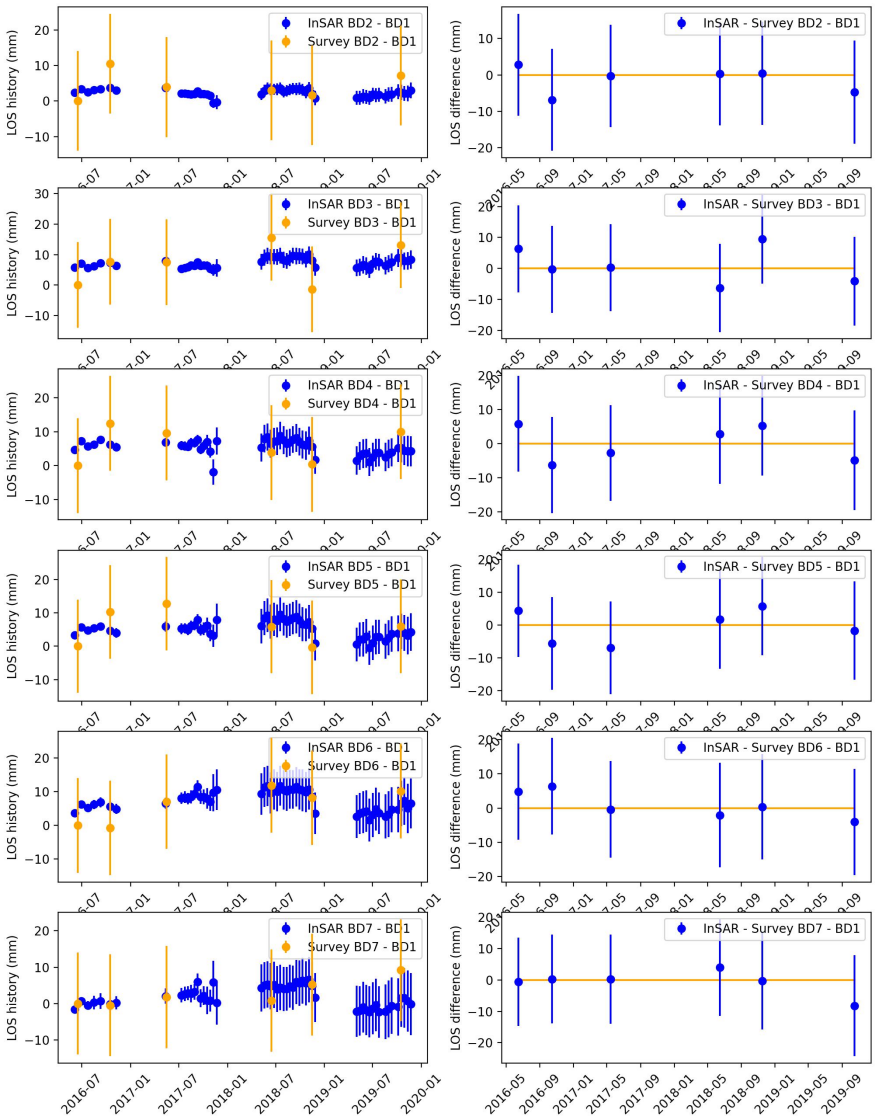

Figure 6. Deformation histories as measured with InSAR and surveys (left) and their difference (right) for the Bose Lake Dam AOI. The rows are the differences in deformation between BD2 through BD7 and BD1.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 7 June 2022 doi:10.20944/preprints202206.0097.v1

between the InSAR and survey measurements is less than a cm, and for all points well within the estimated uncertainty on the difference.

We used swath 3 of orbit track 115 in this example. The kernel, geocoded data sampling and decimation were the same as those used in the Arica example.

3.4. Oslo

The polar orbits of Sentinel-1 lead to overlapping orbit tracks over Northern Europe. The railyard in Oslo is covered by swath 3 of orbital track 44 and by swath 2 of orbital track 146, which are both ascending tracks. Additionally the railyard is covered by two descending tracks: swath 2 of orbital track 139 and swath 3 of orbital track 66.

When one has LOS deformation results from InSAR based on stacks of ascending and descending collections, one can compute motion in the plane spanned by the two incidence vectors. For polar orbits, this plane is spanned by a vector that is nearly vertical, and a vector that is nearly east.

With this, we compute the vertical motion derived from analyses based on data collected from orbital tracks 44 and 139 and refer to the result as Z_1. We call the result of the same analysis applied to data from orbital tracks 146 and 66, Z_2. Fig. 7 shows the cumulative vertical deformation between May 2021 and May 2022.

Figure 8 shows the deformation histories for the two different analyses and their difference. For this pair of points the difference in the analyses largely stays within 2mm, but is often outside the uncertainty suggested by the ensemble analysis. Our intuition is that, due to the proximity of the test and reference points, APS contributions will be minimal and that the differences beyond our estimated error are due to consistent differences in the scattering properties due to slight differences in the imaging geometries.

We chose the test and reference points to reflect a best-possible measurement. Both points are in a connected region of high quality points imaging metal structures, and both points are close together. This is not representative of all measurements we might want to make with these data, and there are a number of questions one can address with these data. The simplest is how much do the results from these two analyses differ for all common points in this area? We show this difference in the top left and lower right plots of Fig. 9. The top left plot shows the histogram of differences in deformation at the last epoch. The blue curve in the lower right plot is standard deviation of the differences by epoch. It increases with time as InSAR derived deformation at an epoch depends on data from all prior epochs. The standard deviation is less than 4mm for all epochs. Note that there is linear temporal interpolation because the data sets do not provide data for the same epochs, and this leads to the first difference being not identically zero. It is important to bear in mind the obvious, that the standard deviation reflects properties of a distribution and does not imply any bound on the values within; the maximum observed difference between these two data sets was 38mm.

The top right plot is the histogram of values $R = (Z_1 - Z_2)/\sigma_{Z_1-Z_2}$, where $\sigma_{Z_1-Z_2}$ is derived from our ensemble methods. This histogram is for all epochs and all common points, the standard deviation is roughly 1.5. The orange curve in the lower right plot is the standard deviation of these data per epoch. If our ensemble methods captured the true uncertainty this quantity would be one. The lower left plot is a scatter plot of $(Z_1 - Z_2)$ vs $\sigma_{Z_1-Z_2}$. This lower left plot, and the orange curve in the lower right plot indicates the degree to which the uncertainty derived from ensemble methods reflects the degree to which the two analyses are likely to disagree.

The kernel, geocoded data sampling and decimation were the same as those used in the Arica and Bose-Lake cases.

4. Discussion

Overall, we find good agreement between our deformation signals, the GPS time series, and the leveling measurements, with accuracy ranging from ~2-10mm in areas where we believe there were no unwrapping errors, to ~8-14cm where unwrapping errors

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 7 June 2022 doi:10.20944/preprints202206.0097.v1

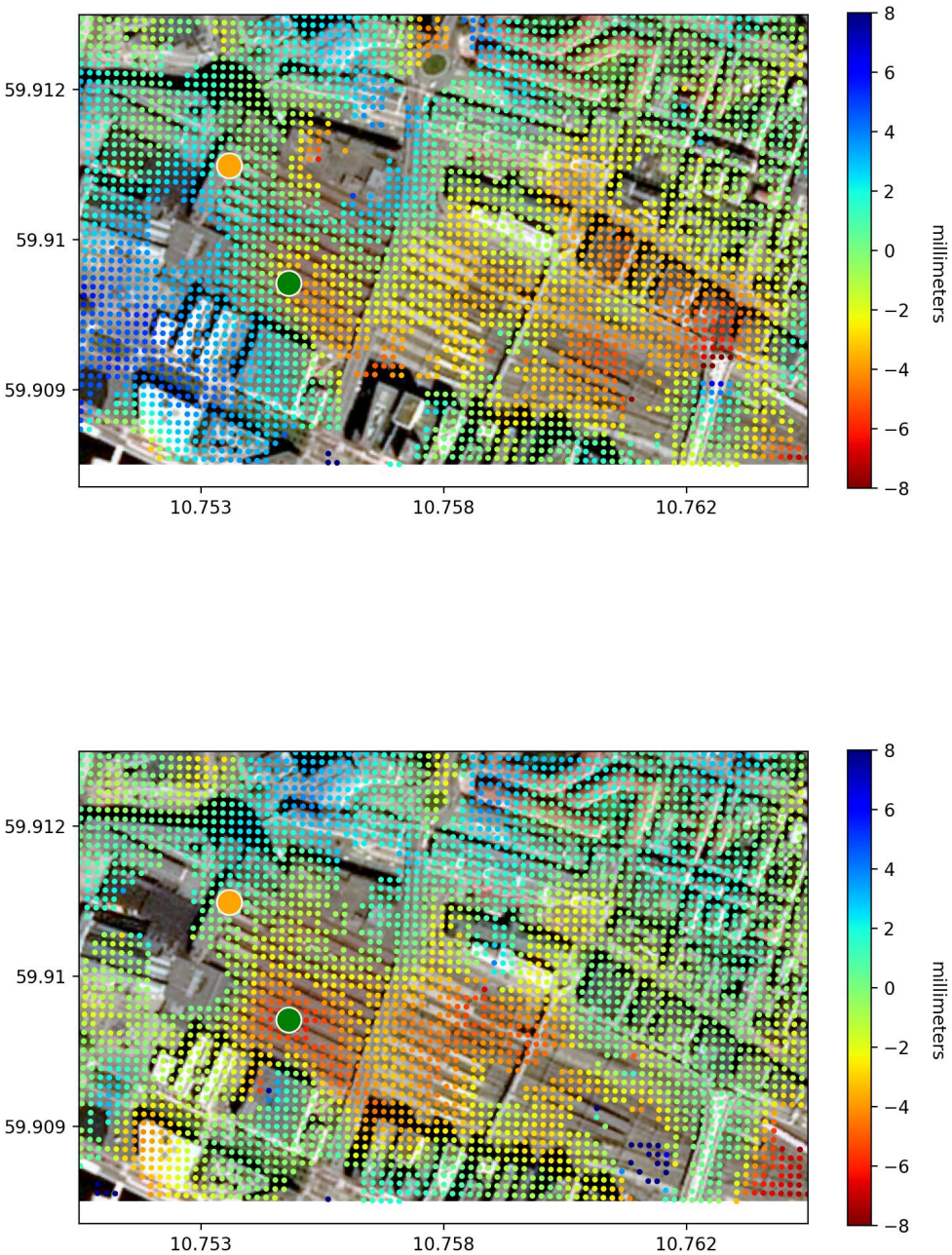

Figure 7. Unfiltered cumulative vertical deformation from May of 2021 to May of 2022 in mm overlaid on optical data, with the reference (orange) and test (green) points. The top shows the 44/139 results and the bottom show the 146/66 results.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 7 June 2022 doi:10.20944/preprints202206.0097.v1

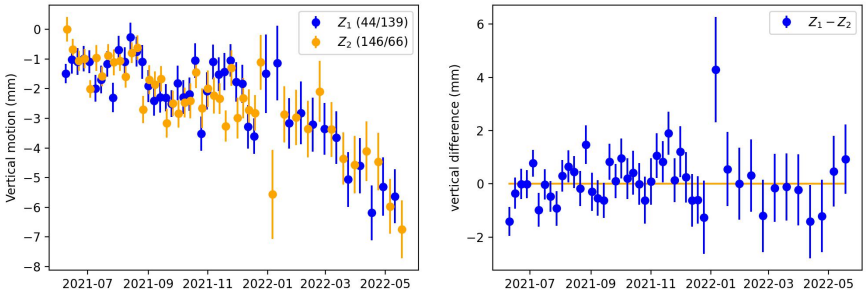

Figure 8. Vertical deformation histories as measured with two different InSAR analyses (left) and their difference (right) for the Oslo rail yard.

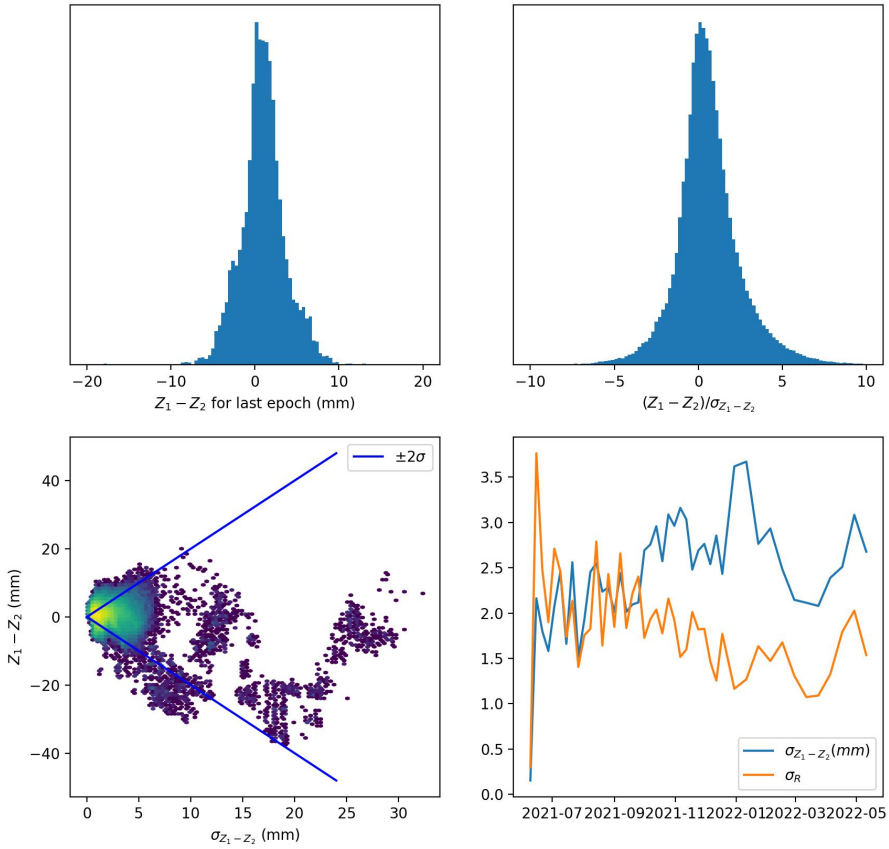

Figure 9. Top left shows a histogram of differences of $Z_1 - Z_2$ for the last epoch. Top right show a histogram of quotients of the difference of corresponding measurements and the estimated uncertainty in the difference. Bottom left shows a scatter plot of estimated uncertainty in the difference on the x-axis and the observed difference on the y-axis. Lines showing $\pm 2\sigma$ are shown in blue for reference. Bottom right is the standard deviation for $Z_1 - Z_2$ in mm per epoch (blue), and the standard deviation of R per epoch (orange) where $R = (Z_1 - Z_2)/\sigma_{Z_1 - Z_2}$ and $\sigma_{Z_1 - Z_2}$ is derived from our ensemble methods.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 7 June 2022 doi:10.20944/preprints202206.0097.v1

may have been present (Kilauea). These differences between between measurements are on par with others that have been published in the literature when compared to GPS and leveling measurements [4,6–14], which indicates that we do not see a loss of accuracy when using an automated pipeline.

We believe that APS and unwrapping errors are likely the largest contributors to inaccuracies for InSAR measurements. In regards to the Kilauea AOI, studies have shown that atmospheric delays increase both with distance between the test and reference points, and with difference in elevation between the test and reference points [28,29]. The GPS stations at this AOI are placed far apart and at different elevations along the volcano, leading to increased influence of APS between the stations. Further, we suspect that there were persistent unwrapping errors due to a combination of phase noise and the sheer amount of deformation. However, our differences in accuracy for Kilauea are similar to those reported in [6] without APS removed for 12-day temporal baselines. [6] reports average RMS GPS/InSAR errors for each site between 2.7 and 8.1 cm, with an average of 6.3 cm. Our average RMS GPS/InSAR errors for the four GPS pairs are ~6.3 cm as well. Improving accuracy here would likely require a more sophisticated unwrapping error detector and would require removing the effects of APS. Traditionally, such discrepancies would be addressed during reprocessing by possible use of *a priori* information from complementary data sources or an expert operator intimately familiar with this geographic location. We also believe that this particular AOI or any AOI with meter-scale rapid deformation events like volcanic eruptions or earthquakes would be challenging for most time-series InSAR methods, including those in [3], as these are often designed and tuned to observe slowly changing trends in time and space. While we note that average errors are high for this AOI and the total deformation was both overestimated and underestimated at points, the automated analysis still identified the eruption event and can serve as a monitoring alert.

Finally, it is possible that GPS itself may have inaccuracies, particularly in the vertical solutions, e.g., the jump in the difference between the KOSM and MANE GPS data at Kilauea in May of 2018.

At Arica and Bose Lake, which we would consider typical AOIs exhibiting slowly varying trends in space and time, agreement between GPS and InSAR was sufficient for us to conclude that there were likely no unwrapping errors, and that APS may have been responsible for the large outlier in Arica in April 2021. Modest APS effects such as this would be reduced with a simple scheme such as that presented in [24].

Our results for the Oslo AOI compare two deformation histories derived from our pipeline. Each of these histories are based on a pair of input stacks of data collected from ascending and descending passes. These results indicate the degree of consistency in our pipeline across different input stacks of Sentinel-1 imagery. For the selected test and reference points agreement is within a few mm. For all points in the area the agreement is largely within a cm.

The error bars in the differences between InSAR-derived deformation histories and other sources of deformation histories are based on our ensemble methods, which attempt to estimate uncertainty in deformation history due to local phase noise. Often the observed differences exceed these error bars suggesting there are sources of difference beyond local phase noise.

Overall, from these experiments, we believe that while an automated InSAR pipeline does save human and compute time, computational costs, and allows for us to monitor more areas and obtain quick results, it does not eliminate the need for interpretation.

5. Conclusions

This work presents comparisons of surface deformation measurements from InSAR compared to surface deformation measurements derived from GPS, total station survey measurements, and with other InSAR results. The InSAR algorithm was not changed between these different analyses, and the only parameters that were changed from AOI to

AOI were resolution and seasonal exclusion of unusable data, both of which can be done in a deterministic way.

The goal of this work was to assess the accuracy and consistency of an InSAR pipeline that was run autonomously. The uncertainty analysis presented in this work did not consider uncertainty from APS, nor did it include any spatial or temporal post-filtering steps, consequently the estimated uncertainties should be interpreted as upper bounds on what is possible. Our conclusion is that while accuracy is going to depend on context, for measurements within half a km of the reference point and for locations that are not deforming too rapidly, a cm is a reasonable estimate for our choice of pipeline parameters and thresholds and that a mm is possible, although unlikely without human intervention or tuning and at the expense of spatial coverage. Further, corroborating data such as GPS and survey tools don't support that level of accuracy themselves. For certain settings, e.g. rapid meter-scale deformation at the volcano at Kilauea, completely autonomous Sentinel-1 InSAR is likely to provide temporally resolved trends, but accuracy will be approximately of 10cm.

Author Contributions: Conceptualization, K.M.O., M.T.C., and P.S.A.; methodology,K.M.O., M.T.C., and P.S.A.; investigation, K.M.O.; data curation, P.S.A. and M.T.C., software, K.M.O., M.T.C., and P.S.A.; validation, K.M.O. and M.T.C.; writing—original draft preparation, M.T.C., K.M.O., and P.S.A.; writing—review and editing, K.M.O., M.T.C., and P.S.A.; visualization, M.T.C. and K.M.O.; All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The single-look complex (SLC) Sentinel-1 imagery and associated metadata are available at the Alaska Satellite Facility's Vertex Portal here: https://search.asf.alaska.edu/. GPS data used as a comparison can be found at the Nevada Geodetic Laboratory's website here: http://geodesy.unr.edu/. The Teck survey data can be found in the *Instrumentation Plots* section of this document: https://www.teck.com/media/2019-Dam-Safety-Inspection-for-Bethlehem-Tailings-Storage-Facility.pdf accessed May 12, 2022.

Acknowledgments: The authors would like to thank Mike Warren and Scott Arko for maintaining Descartes Labs Sentinel-1 data pipeline and for helpful discussions relating to InSAR.

Conflicts of Interest: The authors assert that they have no conflicts of interest.

References

1. Rosen, P.A.; Hensley, S.; Joughin, I.R.; Li, F.K.; Madsen, S.N.; Rodriguez, E.; Goldstein, R.M. Synthetic aperture radar interferometry. *Proceedings of the IEEE* **2000**, *88*, 333–382.
2. Dehls, J.F.; Larsen, Y.; Marinkovic, P.; Lauknes, T.R.; Stødle, D.; Moldestad, D.A. INSAR. No: A National Insar Deformation Mapping/Monitoring Service In Norway–From Concept To Operations. In Proceedings of the IGARSS 2019-2019 IEEE International Geoscience and Remote Sensing Symposium. IEEE, 2019, pp. 5461–5464.
3. Ferretti, A.; Passera, E.; Capes, R. End-to-end implementation and operation of the European Ground Motion Service (EGMS): Algorithm Theoretical Basis Document. Technical Report EGMS-D3-ALG-SC1-2.0-006, European Environment Agency, 2021.
4. Shanker, P.; Casu, F.; Zebker, H.A.; Lanari, R. Comparison of persistent scatterers and small baseline time-series InSAR results: A case study of the San Francisco bay area. *IEEE Geoscience and Remote Sensing Letters* **2011**, *8*, 592–596.
5. Ferretti, A.; Savio, G.; Barzaghi, R.; Borghi, A.; Musazzi, S.; Novali, F.; Prati, C.; Rocca, F. Submillimeter accuracy of InSAR time series: Experimental validation. *IEEE Transactions on Geoscience and Remote Sensing* **2007**, *45*, 1142–1153.
6. Zebker, H. Accuracy of a model-free algorithm for temporal InSAR tropospheric correction. *Remote Sensing* **2021**, *13*, 409.
7. Lee, I.; Chang, H.C.; Ge, L. GPS campaigns for validation of InSAR derived DEMs. *Journal of Global Positioning Systems* **2005**, *4*, 82–87.
8. Jiang, J.; Lohman, R.B. Coherence-guided InSAR deformation analysis in the presence of ongoing land surface changes in the Imperial Valley, California. *Remote Sensing of Environment* **2021**, *253*, 112160.
9. Manzo, M.; Fialko, Y.; Casu, F.; Pepe, A.; Lanari, R. A quantitative assessment of DInSAR measurements of interseismic deformation: the Southern San Andreas Fault case study. *Pure and Applied Geophysics* **2012**, *169*, 1463–1482.
10. Armaş, I.; Gheorghe, M.; Lendvai, A.M.; Dumitru, P.D.; Bădescu, O.; Călin, A. InSAR validation based on GNSS measurements in Bucharest. *International Journal of Remote Sensing* **2016**, *37*, 5565–5580.
11. Casu, F.; Manzo, M.; Lanari, R. A quantitative assessment of the SBAS algorithm performance for surface deformation retrieval from DInSAR data. *Remote Sensing of Environment* **2006**, *102*, 195–210.

12. Yang, K.; Yan, L.; Huang, G.; Chen, C.; Wu, Z. Monitoring building deformation with InSAR: Experiments and validation. *Sensors* **2016**, *16*, 2182.
13. Marinkovic, P.; Ketelaar, G.; van Leijen, F.; Hanssen, R. InSAR quality control: Analysis of five years of corner reflector time series. In Proceedings of the Proceedings of Fringe 2007 Workshop (ESA SP-649), Frascati, Italy, 2007, Vol. 26, p. 30.
14. Luo, Q.; Zhou, G.; Perissin, D. Monitoring of subsidence along Jingjin inter-city railway with high-resolution TerraSAR-X MT-InSAR analysis. *Remote Sensing* **2017**, *9*, 717.
15. Goldstein, R.M.; Zebker, H.A.; Werner, C.L. Satellite radar interferometry: Two-dimensional phase unwrapping. *Radio science* **1988**, *23*, 713–720.
16. Torres, R.; Snoeij, P.; Geudtner, D.; Bibby, D.; Davidson, M.; Attema, E.; Potin, P.; Rommen, B.; Floury, N.; Brown, M.; et al. GMES Sentinel-1 mission. *Remote Sensing of Environment* **2012**, *120*, 9 – 24. The Sentinel Missions - New Opportunities for Science, https://doi.org/https://doi.org/10.1016/j.rse.2011.05.028.
17. Zebker, H.A. User-Friendly InSAR Data Products: Fast and Simple Timeseries Processing. *IEEE Geoscience and Remote Sensing Letters* **2017**, *14*, 2122–2126.
18. Meta, A.; Mittermayer, J.; Prats, P.; Scheiber, R.; Steinbrecher, U. TOPS Imaging With TerraSAR-X: Mode Design and Performance Analysis. *IEEE Transactions on Geoscience and Remote Sensing* **2010**, *48*, 759–769.
19. Beneke, C.M.; Skillman, S.; Warren, M.S.; Kelton, T.; Brumby, S.P.; Chartrand, R.; Mathis, M. A Platform for Scalable Satellite and Geospatial Data Analysis. In Proceedings of the AGU Fall Meeting Abstracts, 2017, Vol. 2017, pp. IN32C–04.
20. Guarnieri, A.M.; Tebaldini, S. On the exploitation of target statistics for SAR interferometry applications. *IEEE Transactions on Geoscience and Remote Sensing* **2008**, *46*, 3436–3443.
21. Ferretti, A.; Fumagalli, A.; Novali, F.; Prati, C.; Rocca, F.; Rucci, A. A new algorithm for processing interferometric data-stacks: SqueeSAR. *IEEE transactions on geoscience and remote sensing* **2011**, *49*, 3460–3470.
22. Pepe, A.; Lanari, R. On the extension of the minimum cost flow algorithm for phase unwrapping of multitemporal differential SAR interferograms. *IEEE Transactions on Geoscience and remote sensing* **2006**, *44*, 2374–2383.
23. Perron, L. Operations research and constraint programming at google. In Proceedings of the International Conference on Principles and Practice of Constraint Programming. Springer, 2011, pp. 2–2.
24. Tymofyeyeva, E.; Fialko, Y. Mitigation of atmospheric phase delays in InSAR data, with application to the eastern California shear zone. *Journal of Geophysical Research: Solid Earth* **2015**, *120*, 5952–5963.
25. Olsen, K.M.; Calef, M.T.; Agram, P.S. Contextual accuracy assessments for InSAR methods using synthetic data. In review.
26. Hofmann-Wellenhof, B.; Lichtenegger, H.; Wasle, E. *GNSS–global navigation satellite systems: GPS, GLONASS, Galileo, and more*; Springer Science & Business Media, 2007.
27. Teck Highland Valley Copper Partnership, 2019 Dam Safety Inspection Report. https://www.teck.com/media/2019-Dam-Safety-Inspection-for-Bethlehem-Tailings-Storage-Facility.pdf. Accessed: 2022-05-12.
28. Goldstein, R. Atmospheric limitations to repeat-track radar interferometry. *Geophysical research letters* **1995**, *22*, 2517–2520.
29. Fornaro, G.; D'Agostino, N.; Giuliani, R.; Noviello, C.; Reale, D.; Verde, S. Assimilation of GPS-derived atmospheric propagation delay in DInSAR data processing. *IEEE journal of selected topics in applied Earth observations and remote sensing* **2014**, *8*, 784–799.