UNIVERSITY OF CALIFORNIA SAN DIEGO

Comparing GNSS and InSAR Velocities in Preparation for NISAR

A thesis submitted in partial satisfaction of the
requirements for the degree Master of Science

in

Earth Sciences

by

Amy Whetter

Committee in charge:

Professor Adrian Borsa, Chair
Professor Matthias Morzfeld
Professor David Sandwell

2022

This thesis of Amy Whetter is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

University of California San Diego

2020

Contents

List of Figures

List of Tables

ACKNOWLEDGEMENTS

There are several people that contributed to making my two years at Scripps so positive. First, I would like to thank Adrian Borsa for advising me. He provided excellent guidance while also allowing me to take ownership of my work. I would like to acknowledge David Sandwell and Matti Morzfeld for agreeing to sit on my committee. Next, I'd like to thank the entire NISAR Solid Earth Science Team for their collaboration, answering my questions and helping troubleshoot MintPy. I would also like to thank NGS and their advanced degree program for providing me with this opportunity and support throughout the program. Thanks to the entire 2020 IGPP cohort for constant support, community, and late night homework help. Finally, thanks to my support team outside of Scripps, my family, for constant encouragement, my partner, for listening to me complain and keeping me fed, and my triathlon team, for stress relieving workouts.

ABSTRACT OF THE THESIS

Comparing GNSS and InSAR Derived Ground Motion Velocities in Preparation for NISAR

by

Amy Whetter

Master of Science in Earth Sciences, 2020

Professor Adrian Borsa, Chair

A key element in the preparation for NISAR's launch (see Section 1.2) in 2023 is developing a methodology for validating NISAR-estimated displacements against GNSS displacements. In collaboration with the NISAR Solid Earth Science team I developed a Jupyter notebook workflow to download and generate timeseries of Sentinel-1 interferograms and GNSS daily solutions, project their displacements into radar line-of-sight (LOS), and derive velocity trends for each GNSS station in the given study site region and its collocated InSAR pixel. These velocity trends are then compared to asses the agreement of Interferometric SAR (InSAR) with GNSS. The effect of varying the following parameters were investigated: the GNSS station velocity approximation model, the reference station selected, various applied InSAR corrections, and the number of InSAR pixels averaged to form the InSAR collocated

velocity. Analysis was conducted at two study sites, the Mojave Desert in California and a section of the Central Valley in California. The Mojave study site did not meet the validation requirement set by the NISAR team while the Central Valley did. Another key element of this project is providing future users with a clear and simple workflow to compare GNSS and InSAR data. The workflow discussed in this thesis has been handed off to the Solid Earth Science team to maintain and improve the user accessibility by moving the workflow into a cloud based environment.

Chapter 1

Introduction

1.1 Motivation

Understanding ground motion is an important and longstanding challenge. Historically, ground motion has been studied using continuously operating Global Navigation Satellite System (GNSS) stations on permanent monuments. GNSS only gives sparse point measurements of deformation, which is a limitation as many important deformation processes (e.g., earthquakes, groundwater extraction) are highly variable in space. Although GNSS has high temporal sampling, the current low spatial density of GNSS networks is not adequate for resolving strong spatial gradients in surface deformation. Relying solely on GNSS at its current spatial density can thus lead to a misleading conclusions about the nature of deformation processes.

A way to increase GNSS resolution in areas of scientific interest is to densify the GNSS station network; however, this solution is typically not feasible due to high cost (Wei et al., 2010). InSAR (Interferometric Synthetic Aperture Radar), which observes surface deformation over large areas (250 km swaths) at high resolution (100 m) and moderate temporal sampling (6-12 days), provides a means of addressing shortfalls in GNSS spatial sampling. Since it is an imaging technique, InSAR captures deformation that is missed by GNSS, at the cost of higher noise due to atmospheric and other effects (Bekaert et al., 2015; Gray et al., 2000).

In this thesis, I examine GNSS and InSAR estimates of surface deformation over

two regions (one stable, one actively deforming) in California, both to understand how the observations compare, and to develop algorithms that will be used for the validation of NASA's upcoming NISAR mission (see Section 1.2). For the stable region, I use an area of the Mojave Desert in California where I expect minimal surface motion over the analysis period. Active regions are defined by the NISAR mission as locations where the Global Strain Rate Model (Blewitt et al., 2018) indicates strain in excess of 20 ns/y or 1 mm/y over 50 km (Algorithm Theoretical Basis Document, 2022). For the active region, I use the Southern Central Valley, California, which features some of the highest rates of vertical deformation on the planet. Subsidence rates can be up to 131 mm/yr (Levy et al., 2020).

Since groundwater pumping is the source of most Central Valley subsidence (Famiglietti et al., 2011), the Valley is a location of both scientific and socioeconomic interest. In fact, one of the most important reasons to characterize surface motion is for monitoring groundwater withdrawal. Groundwater is extracted from aquifers all over the world to meet agriculture and urban water demand, and its use is a widespread resource management issue (NISAR User Handbook, 2018). Aquifers can permanently deform when groundwater is withdrawn unsustainably, losing their capacity for future storage (Smith et al., 2017). As water shortages increase due to a changing climate, losing existing storage capacity is of great societal concern.

Monitoring storage capacity has become particularly pertinent in California since the introduction of the Sustainable Groundwater Management Act (SGMA BMP 6), which was enacted to reduce over-extraction of groundwater and to balance pumping with recharge within groundwater basins. Permanent deformation can be avoided by sustainable, low-impact groundwater extraction, which monitoring can help achieve. InSAR has the potential to play a key role in this monitoring because of its broad spatial coverage and high sampling rate (e.g., 12-day repeat data collection in the case

of Sentinel-1A and NISAR), creating a robust timeseries of surface motion that can in turn be used to infer groundwater storage changes and better manage groundwater supply (NISAR User Handbook, 2018).

Groundwater over-extraction not only has the potential to reduce storage capacity, but the accompanying surface subsidence can also lead to infrastructure damage. Land subsidence in the Sacramento River Delta has increased the threat of flooding, and the sinking of Central Valley canals has reduced water flow downstream Escriva-Bou et al. (2020). Similarly, dams become less effective as their storage capacity is reduced. Roads and railways can suffer damage from subsidence. In all these cases, InSAR can play an important role in understanding surface motion, mitigating damages, and saving resources.

1.2 NISAR Mission and Validation Requirements

The NASA-ISRO Synthetic Aperture Radar (NISAR) is an upcoming Earth orbiting satellite mission that is expected to launch in early 2023. The National Aeronautics and Space Administration (NASA) and the Indian Space Research Organisation (ISRO) have put significant time and resources into planning and preparing for this mission. It will be the first radar satellite mission to collect data in two different wavelengths, L-band and S- band, and it will have a repeat orbit of 12 days. Multiple wavelengths and low repeat time will create a tremendous amount of new data that can be used for a variety of scientific applications (NISAR Mission Concept).

The NISAR mission team has outlined all that needs to be accomplished before launch. One of the NISAR science team's current priorities is creating the workflow that will be used to calibrate and validate InSAR observations, and to demonstrate that these observations meet mission requirements. The work described in this thesis directly

supports these initiatives, and I have worked in close collaboration with the Solid Earth Science team.

NISAR's Solid Earth Science (SES) Mission requirements are outlined in the Solid Earth Algorithm Theoretical Basis Document (ATBD). Three types of surface deformation are addressed in the SES ATBD: transient motion, coseismic motion from large earthquakes, and secular deformation. Each has its own validation requirements. My work focuses on the secular deformation, which is the most relevant in the Central Valley.

The ATBD secular deformation validation requirement states,

> "The NISAR project shall measure at least two components of the spatially and temporally averaged relative vector velocities over active regions of Earth's land surface with accuracy of 2 mm/yr or better, over length scales 0.1 km $\leq$ L $\leq$ 50 km" (Algorithm Theoretical Basis Document, 2022)"

For this thesis, I analyze only one component of vector velocity in two locations as a proof-of-concept for NISAR validation, with the understanding that actual validation will use two look directions and many calibration areas. NISAR uses different validation approaches to satisfy this requirement, based on whether or not independent reference data is available. The work in this thesis contributes to validation approach #1, which assumes that a network of GNSS stations exists within the calibration/validation region to provide independent displacement measurements. This approach requires GNSS station displacements to be projected into Line of Site (LOS) velocity estimates so that they can be compared to the InSAR LOS displacements.

The reason GNSS is used as a reference for InSAR is that GNSS displacements are derived from continuous observations of up to dozens of GNSS satellites over 24-hour periods using stable monuments. This means that GNSS estimates of surface motion can take advantage of varying observational geometry and models of atmospheric delay

to minimize uncertainty. This is unlike InSAR, whose underlying images are acquired from much more limited geometric and temporal sampling, and thus are more likely to be impacted by effects that are unrelated to actual Earth deformation.

Chapter 2

Data Collection and Processing

The work presented in this thesis uses collocated InSAR and GNSS data at two separate locations in California. This chapter discusses how these data were collected and processed.

2.1 Study Sites

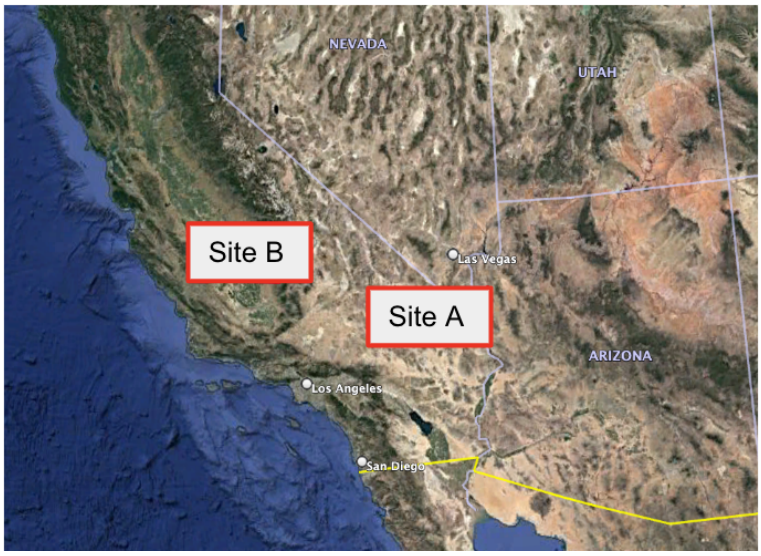

Figure 2.1: Study site location map. Each study site in shown by a white box outlined in red. Site A, the control, is located in California's Mojave Desert. Site B is located in California's Central Valley and has significant vertical motion.

Two study sites were selected, Figure 2.1. Site A covers a portion of the Mojave Desert in California. This site is used as a control site as it currently shows minimal ground motion and has simple terrain, seen in Figure 2.2. The Site A bounding box

coordinates are 34.66, -116.62, 35.60, -114.39 (S,N,W,E), covering a 110 km by 220 km area. Observational data for Site A includes 244 Sentinel-1 interferograms, from 71 SAR acquisitions spanning temporal baselines of approximately 1 week to 1 year, spanning the period 2018/01/01 to 2020/01/01. We also use North/East/Up (N/E/U) displacements from 27 GNSS stations over the same period. During InSAR processing the raw SAR SLC (single look complex) images where spatially filtered using a multi-look filter with 19px in azimuth and 7px in range.

Figure 2.2: Inside the red boundary shows Site A and reference station P626

Site B covers a swath of the southern Central Valley and a portion of the adjacent Sierra Nevada. The site includes areas of significant subsidence and features diverse terrain, with high relief to the west and east of a broad flat valley floor, seen in Figure 2.3. This site was selected to see how InSAR performs under high-deformation conditions. The bounding box coordinates are 35.77, 36.75, -120.61, -118.06 (S,N,W,E), which covers an area of 100 km by 200 km. Data for this site includes 203 Sentinel-1 interferograms, from 70 SAR acquisitions with roughly the same temporal baselines as for Site A, spanning the same two year time period as for Site A (2018/01/01 to 2020/08/01), along with N/E/U displacements from 56 GNSS stations. A two year timeframe was used to help mitigate the effect of high displacement gradients and atmospheric noise in the interferograms.

Figure 2.3: Inside the red boundary shows Site B and reference station P467

2.2 Data Collection

The NISAR Science Team is testing NISAR validation algorithms using synthetic aperture radar data from Sentinel-1, as the NISAR mission will not be able to generate data until after launch. Sentinel-1 uses C-Band radar, which has a shorter wavelength than NISAR's primary L-Band instrument. Sentinel-1 is less susceptible to ionospheric noise than NISAR (Bekaert et al., 2015; Gray et al., 2000), however the Sentinel-1 mission still serves as a good NISAR analogue. The Sentinel-1 repeat time is nominally 12 days, but was 6 days when both the 1-A and 1-B satellites were flying. Similarly, NISAR plans a 12-day repeat cycle with ascending and descending passes giving 6-day average sampling (NISAR User Handbook, 2018).

For the validation described in this thesis, two Sentinel-1 datasets were used: Frame 475 on Path 173 (descending) for the Mojave, and Frames 471-473 on Path 144 (descending) for the Central Valley. A total of 447 geocoded unwrapped (GUNW) interferograms for these two sites were prepared by the NISAR project, staged at the Alaska Satellite Facility (ASF), and downloaded in my workfow for validation testing. Note that it takes under a minute to download an interferogram, which is a small fraction of the time to process that interferogram locally. GUNW products are extremely easy

to use, since they are projected from radar azimuth/range coordinates onto a uniform X/Y pixel grid that can easily be converted to lat/lon. We initially processed and downloaded a subset of these interferograms using ASF's Vertex tool, but switched to the NISAR-project-generated products to test production workflows.

By way of context, Vertex is a web-based user interface created by ASF to provide SAR data discovery and on-demand processing for standard products (e.g., radiometrically terrain corrected SAR backscatter; InSAR line-of-sight displacement). Prior to the availability of on-demand processing, InSAR users had to use powerful, but complex, processing tools such as GMTSAR (Sandwell et al., 2011) and ISCE (Rosen et al., 2012). ASF developed Vertex's on-demand services as part of its Getting Ready for NISAR (GRFN) project, which pioneered the development of services allowing non-expert users to make their own interferograms (Garron et al., 2019). Both project-generated and user-generated interferograms are staged on Vertex under "S1 InSAR (beta)" category of data products.

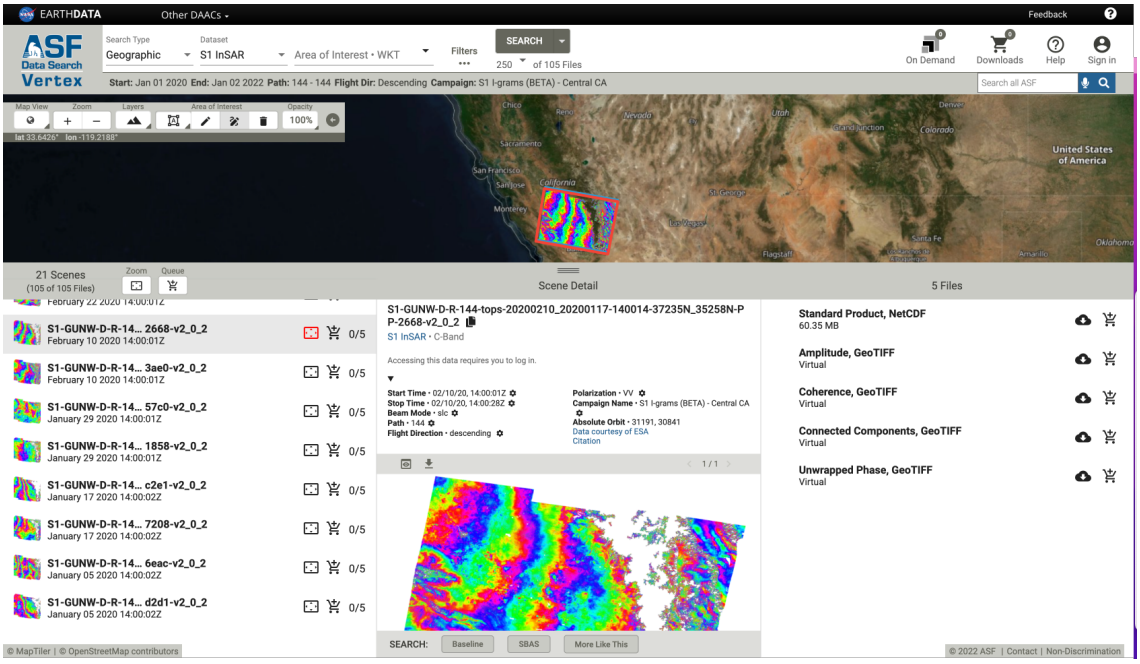

Figure 2.4: Vertex interface, showing ASF-archived GUNW interferograms for Path 144 over our Central Valley study site.

Raw data from permanent GNSS networks are processed by several analysis centers (ACs), each of which uses different processing software and parameters. The standard AC data product is a daily timeseries of North/East/Up (NEU) positions for each GNSS station, provided in the IGS14 reference frame. For this analysis I used University of Nevada Reno's (UNR) daily position solutions, rotated into the Sentinel-1 radar line of sight (see Section 3.2 below). Every week, UNR updates the daily positions for approximately 10,000 stations (Blewitt et al., 2018). While UNR is a popular data source, this analysis could be redone with data from any of the processing centers.

In addition to position timeseries, analysis centers such as UNR typically produce GNSS station velocity estimates. However, we cannot use these estimates for validating the NISAR secular requirement, since AC velocities are typically estimated over the entire station lifespan, while we need estimates for the (typically shorter) period spanned by the available InSAR observations. Additional information on InSAR and GNSS velocity estimation is included in section 3.2.

We note that while the work in this thesis was produced and executed on a laptop Jupyter computing environment, the actual validation workflow for NISAR will be run entirely in the cloud. My secular requirement Jupyter notebook is currently being transferred to OpenSARLabs, a cloud service that sits next to ASF's cloud archive and which will allow users to use InSAR data for validation without the need for slow data downloads.

2.3 InSAR Timeseries Processing Overview

There are two main techniques for creating displacement timeseries from InSAR: Persistent Scatterer InSAR (PS-InSAR), and Small BAseline Subset InSAR (SBAS). Each technique has pros and cons, giving different utility for each method. Persistent

scatterers are points that dominate the SAR backscatter, i.e., that show up as very bright points in the SAR image (Ferretti et al., 2000). These points commonly appear within cities, often from reflections off the edges of buildings or from unintentional corner reflectors. Due to their high amplitudes relative to other nearby SAR scatterers, point scatterers are associated with phase and amplitude stability over time and thus do not decorrelate between SAR images. PS-InSAR does not support phase unwrapping (the process converting phase to useful units by removing complete 2π phase cycles at each point of the interferometric phase image (Werner et al., 2002)), since deformation estimates are only available for a sparse set of PS points. Instead, the SAR phases at individual points are differenced and a least squares adjustment is done with respect to a reference point. This method is best for one dimensional motion with relatively low spatial gradients in displacement, and is not the methodology used in this analysis.

SBAS is a technique (Yunjun et al., 2019) that combines many high coherence interferograms into a robust timeseries of surface deformation. SBAS selects a subset of available interferograms with short spatial and moderate temporal baselines to increase the coherence of deformation information, especially in natural environments. This method typically uses a large number of overlapping interferograms, which can appear redundant but actually acts as a robust mechanism to catch and mitigate errors and reduce noise. However, more interferograms requires higher computational efforts. Another drawback of SBAS relative to PS-InSAR is that adding new data into SBAS timeseries requires redoing the entire timeseries inversion rather than just adding new acquisitions at the end. Additionally, SBAS only works with no temporal gaps in the interferogram network, although "bad" images can be sidestepped by forming interferograms that span the impacted epoch. A major benefit of SBAS is that it provides continuous spatial coverage where unwrapping is successful, which is better for complex geophysical signals like surface deformation from earthquakes and groundwater pump-

ing.

For this analysis, I use SBAS because I have access to a large number of interferograms with good correlation and unwrapping, which means I do not have to rely on a small number of persistent scatters to estimate deformation.

Chapter 3

Methods

This chapter presents the software used and analysis workflow I wrote to get the InSAR and GNSS data sets into radar line-of-sight (LOS) velocities so they could be jointly analyzed.

3.1 Third-Party Software for Data Download and Timeseries Analysis

There are several software packages available for InSAR Timeseries analysis. For persistent scatterer analysis, some examples are Traditional PS-InSAR, StaMPS, dePSI, and Coherent Target InSAR. For SBAS, some examples are Traditional SBAS InSAR, StamPS SBAS InSAR, GIAnT, and MintPy. The NISAR mission selected MintPy for requirement validation, and since a primary motivation for this project is to assist the NISAR mission, I use MintPy as well. MintPy (Miami INSAR time-series software in Python, (**?**) was originally developed at the University of Miami, but is now maintained by NASA's JPL as an open source project and is available on GitHub. MintPy has multiple capabilities. It can be used to look at the displacement of a single point, a transect, or an entire area, with application to earthquakes, volcanoes, subsidence and more. As I describe below, I use MintPy to analyze deformation over entire areas, and I take advantage of several of MintPy's processing features.

In addition to MintPy, I use JPL's open-source ARIA-tools python package to

download interferograms from the Alaska Satellite Facility (ASF) and to reformat these datasets for MintPy to read. ARIA-tools is well-documented and available on GitHub (ARIA GitHub).

MintPy uses SBAS to derive the timeseries. In MintPy, the input interferogram is already multilooked and filtered, thus, the noise is low already. Additionally, simple SBAS without any extra constraints is unbiased when the network is fully connected, thus, no smoothing is needed in the time domain (Yunjun et al., 2019).

3.2 Jupyter Notebook Workflow

This section describes the Python workflow I created to do the InSAR/GNSS validation. I wrote this workflow in the form of two Jupyter notebooks, one of which implemented the NISAR validation algorithm from start to end, and the other of which I used to perform the tests that I describe in Chapter 5. In order to use this notebook, it is necessary to have already installed MintPy and ARIA-tools. I recommend setting up an ATBD python environment as described on the NISAR Solid Earth ATBD GitLab (NISAR GitLab). Additionally, you must have an ASF account. A summery of the below steps can be found in Figure 3.1.

Step 1: Setting up Workflow Parameters

The following user-defined variables set the location and time period of the validation analysis. Theoretically, this analysis can be run anywhere on Earth, but it does require pre-made interferograms and GNSS stations to be available within the specified analysis region:

i. project directory

ii. project name

iii. download region (in S,N,W,E format)

iv. analysis region (in S,N,W,E format)

v. download start date

vi. download end date

Step 2: Download Interferograms

Because the NISAR mission will produce interferograms as a standard data product (NISAR User Handbook, 2018) and will produce all interferograms needed for validation, this workflow does not perform interferogram processing. Instead, it uses ariaDownload.py, a script within ARIA-tools, to query the ASF archive for any existing interferograms within the spatial and temporal analysis bounds. This script downloads available interferograms into a product folder within the project directory. Depending on the number of interferograms available within the region between the start and end dates, this step can take several hours.

Phase unwrapping is done using Statistical-cost, Network-flow Algorithm for Phase Unwrapping (SNAPHU), which is widely used in InSAR processing software. MintPy and ARIA-tools have build-in processing to minimize unwrapping errors. (ASF Phase Unwrapping Documentation)

Step 3: Prepare Interferograms

To set up the downloaded data, the workflow uses the ariaTSetup.py script from ARIA-tools. This step first confines all interferograms in the product directory to the analysis region, although all our files should already meet this requirement. Next, it downloads and applies the Global Self-consistent Hierarchical High-resolution Shorelines (GSHHS, (Wessel and Smith, 1996)) water mask, which removes all pixels over water. Finally (via the –croptounion option), it crops interferograms to the analysis region. Once run it creates the output directories within the overall project directory that contain

information about the coherence, imaging geometry, and the inferferogram stack. This information is used in later processing steps.

Step 4: MintPy Configuration

MintPy uses a configuration file to set most of its runtime parameters. The configuration file must be updated to set the *mintpy.load.incAngleFile* and *mintpy.load.azAngleFile* parameters to the .vrt files in the incidentAngle and azimuthAngle folders mentioned above. Custom configuration options include excluding specific interferograms and setting minimum coherence limits. The configuration file also includes parameters for controlling the application of various corrections to the InSAR data (e.g., tropospheric correction, topography correction, and either linear or quadratic deramping). To apply topographic and troposphere corrections requires setting up accounts with OpenTopography and Copernicus to obtain an API key for programmatic access to their archives. MintPy will need these keys to download correction data (Yunjun et al., 2019). Note, MintPy's naming convention is "topography" correction, however, this correction is really an elevation-dependent troposphere correction.

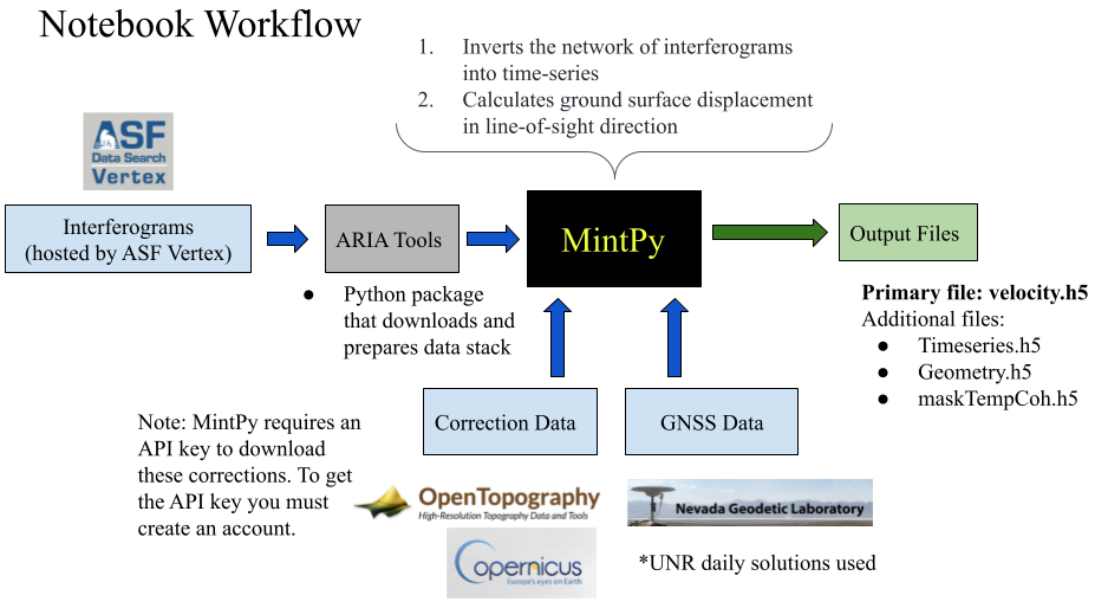

Figure 3.1: Flow diagram showing the jupyter notebook workflow.

Step 5: Generating an InSAR Velocity File

Next I run smallbaselineApp.py in MintPy, which performs the following steps. Note that each of these steps can be run individually for debugging purposes, or all at once.

i. **load_data**: load all interferograms prepared by ARIA-tools.

ii. **modify_network**: runs quality control and drops poor interferograms.

iii. **reference_point**: reference all deformation to a single point within the analysis region; the reference point can be user-specified, but MintPy selects it by default.

iv. **invert_network**: inverts the network of interferograms.

v. **correct_troposphere, deramp, correct_topography**: optional corrections.

vi. **velocity**: creates a velocity.h5 file containing the velocity for each pixel of the interferogram stack. By default MintPy finds the best-fitting line to all the interferogram offsets, but MintPy can simultaneously estimate seasonal

terms and offsets to improve the linear fit.

Step 6: Get GNSS Positions and Create LOS Timeseries

The next step in the workflow uses MintPy to query the University of Nevada Reno (UNR) archive for all the GNSS stations in the analysis region, as defined from metadata on InSAR coverage in the velocity.h5 file. For each GNSS station, MintPy downloads the daily NEU (North/East/Up) position solutions from UNR.

Traditionally, GNSS displacements are published in three components: North, East, and Up (N/E/U). InSAR displacements most commonly come in a single Line-of-Sight (LOS) component that is directed from the SAR satellite down to the ground. To compare these two datasets we must first get them into the same coordinates. It is possible to convert InSAR LOS to N/E/U, but this requires 3 or more look directions (e.g., data from both the ascending and descending paths) or other assumptions to derive the correct relationships between N/E/U component. As we only use data from the descending satellite track, we must instead convert the GNSS displacements to LOS. This is done using the InSAR viewing geometry for the analysis region to reproject the daily positions into the radar line-of-sight (LOS). For this conversion the geometry of the first interferogram in the stack is used. However, within that scene the specific geometry for each GNSS station is used, as the geometry changes throughout the scene. MintPy was not originally programmed to handle GNSS timeseries, but this functionality was added to support the NISAR mission validation. My workflow performs basic quality control to remove stations that do not meet a temporal coverage threshold of 70% during the study time period.

Step 7: Estimate GNSS LOS Velocities

I use MintPy to estimate LOS velocities at each GNSS station in the analysis region. For a default model, I uses a linear term with a single annual sinusoid. The annual periodic

term is important to include in regions where there are significant seasonal effects, like in California, but it does not suit all stations and all regions. Figure 3.2 shows how the velocity model is implemented within the notebook, with double periodic terms and steps included for stations with complex seasonal behavior and/or uncorrected timeseries offsets.

An offset is a jump in the timeseries, typically from a change in an equipment, as seen in Figure 3.3 around 2019-10. Three stations within the two study regions have offsets: AZBH (on 9/19/2019) and BEPK and P573 (on 7/6/2019). Initially, I flagged and removed stations with offsets, but that resulted in the loss of GNSS data needed for validation. To better address offsets, I included step functions in the model, placing the step at offsets dates above (which I determined from examining the timeseries).

I used a special customization for CAFP and CRCN, which are unusual in that a single annual sinusoid did not approximate their seasonal displacements. A double sinusoid, with one-year and half-year periodicity, fit the displacement data the best.

```
if site_id == 'AZBH':
    model = {
        'polynomial' : 1,
        'periodic'   : [1.0],
        'step'       : ['20190919']
    }
elif site_id == 'BEPK' or site_id == 'P573':
    model = {
        'polynomial' : 1,
        'periodic'   : [1.0],
        'step'       : ['20190706']
    }

elif site_id == 'CAFP'or site_id == 'CRCN':
    model = {
        'polynomial' : 1,
        'periodic'   : [2]
    }
else:
    model = {
        'polynomial' : 1,
        'periodic'   : [1],
    }
```

Figure 3.2: Final model used to approximate GNSS station velocities.

Figure 3.3 highlights the bias in estimated velocity that can occur when offsets are not accounted for. The original linear-only model (green line) poorly fits the displacement data (blue dots) and exhibits a positive LOS velocity. The modified model with sinusoidal and offset terms (orange line) fits the displacement data well, and has a negative LOS velocity.

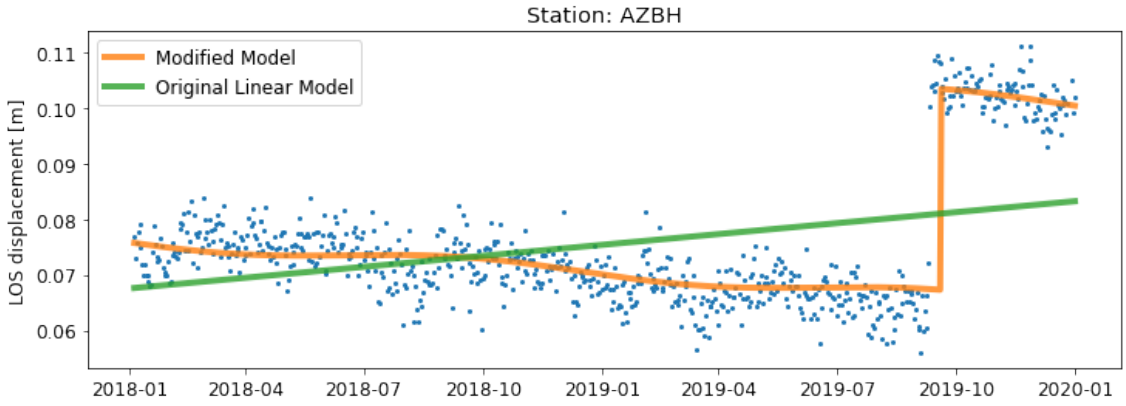

Figure 3.3: In blue show the daily GNSS solutions projected into the LOS. In green is the velocity trendline used for the initial results. In orange is the velocity trendline after modifying the model to include the offset on 2019/09/19.

Figure 3.4, which shows a station with significant seasonality in displacement, highlights how a model with an annual periodic term (orange) fits the GNSS positions (in blue) better than the linear-only model (green). In this example, the resulting velocities are not too different, but you can imagine a case where they could be if the timeseries started and ended in opposite seasons (with a peak on one side and trough on the other).

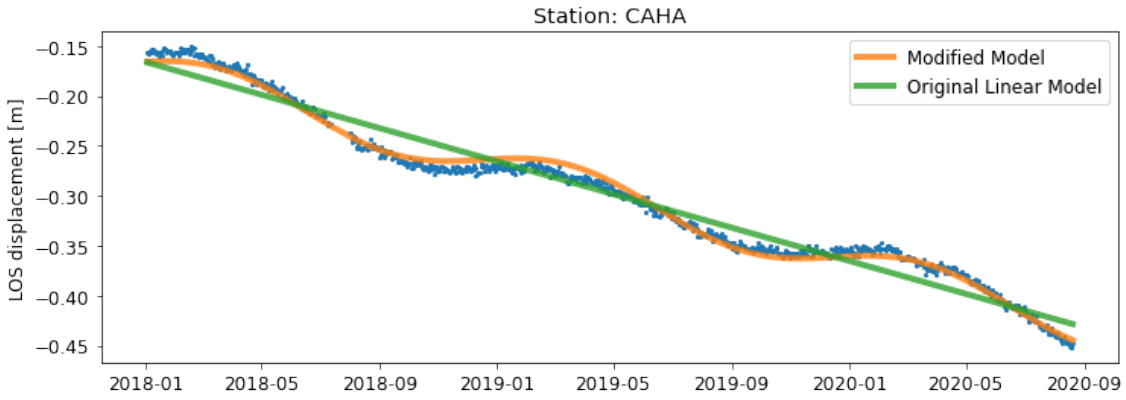

Figure 3.4: In blue show the daily GNSS solutions projected into the LOS. In green is the velocity trendline used for the initial results. In Orange is the velocity trendline after modifying the model to account for the seasonal trend in the data.

Step 8: Reference GNSS and InSAR LOS velocities to Common Frame

To compare InSAR and GNSS velocities across the analysis region requires that both are in the same frame of reference. While GNSS positions (and thus velocities) are provided in an absolute reference frame (e.g., IGS08), InSAR displacements and velocities are provided in a relative frame of reference based on whatever location is chosen as the InSAR reference point in Step 5. However, the original S1 data are in ITRF14.

In this step, the workflow defines a single GNSS station to be a common reference point for both GNSS and InSAR. The GNSS LOS velocity for this station is subtracted from that of all other GNSS velocites, and the InSAR velocity for this pixel is subtracted from that of all other pixels. While this step only removes offsets between the InSAR and GNSS velocity fields, it is still important. Identifying stations located away from deformation processes (i.e., away from subsidence/uplift) to avoid introducing a large local bias, I chose station P626 for Site A in the Mojave and P467 for site B in the Central Valley. Choosing a difference alignment method such as GInSAR (Neely et al., 2020) could improve results and will be discussed Section 6.1.

Step 9: Averaging InSAR Pixel Velocities

InSAR data are provided on a uniform x/y grid for analysis, where each grid point is called a "pixel," and each pixel has dimensions of about 100m on a side. To mitigate the impact of noise or missing data in individual pixels, I took the median velocity of multiple pixels to form our InSAR velocity for comparison. Specifically, I took the median of all InSAR pixels within a 11 px by 11 px window (approximately 1km by 1km) around a GNSS station location to be the InSAR velocity for that location. The 11 px by 11 px (1km by 1km) window represented a trade-off between minimizing noise and introducing error from using too large an area. In practice, my analysis in Chapter 5 shows that the value of the averaging radius has little impact on the InSAR velocity estimate.

After these steps are complete in the workflow, the InSASR and GNSS velocity datasets are ready for analysis.

Chapter 4

Results

This chapter presents the initial results for each study site.

4.1 Site A: Mojave

Plotting the GNSS station velocities with InSAR velocities gives a quick visual way to compare these two datasets. In Figure 4.1 we see that none of the GNSS stations exhibit large velocity discrepancies with respect to the background InSAR velocities.

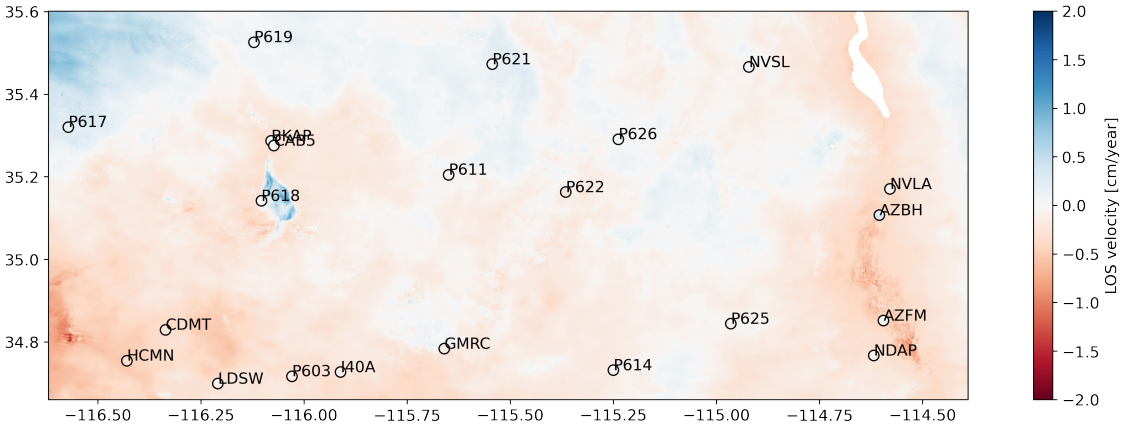

Figure 4.1: InSAR velocity map of Site A. GNSS station velocities are overlaid, shown within the circular marker and colored to the same scale as InSAR velocities. Reference station is P626, which is roughly in the center of the study region.

Figure 4.1 also serves as a check on the validity of the InSAR velocity data, which has the spatial pattern we might expect for this area. On the right, we see the Colorado river masked out and some subsidence (red) in the agricultural regions bordering the Colorado river to the south. Slightly left of center we see a small circular area of uplift. This is

the oasis of Mara Soda Lake, a dry lake bed next to Soda Springs which appears to have differential motion relative to the desert around it. These features give confidence that the InSAR velocities are revealing actual Earth deformation. Overall, however, Site A is fairly stable, which is what we and expected.

Another way to look at these results is in a histogram of the velocities. On the left of Figure 4.2 the InSAR velocities (in blue) are plotted with the GNSS velocities overlaid (in yellow). On the right are the residuals (GNSS minus InSAR) between the datasets at each GNSS station. While the InSAR and GNSS velocity distributions have roughly the same spread (standard deviations of 0.19 cm/y and 0.13 cm/y, respectively), they have different mean values (-0.18 cm/y and -0.05 cm/y), which results in a positively biased residual distribution. Overall, both sets of velocities exhibit small velocities overall, with the differences between them consistent with small amounts of additive noise.

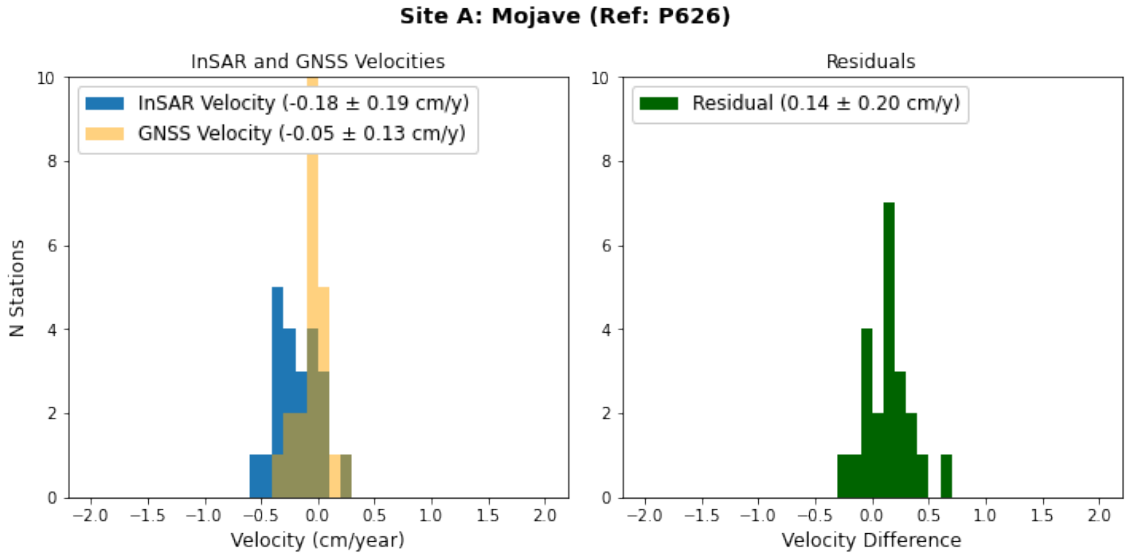

Figure 4.2: Left: the Mojave velocities and standard deviation with P626 as the reference station. Right: GNSS minus InSAR residuals with standard deviation.

4.2 Site B: Central Valley

The Central Valley velocities are shown in Figure 4.3. This analysis region has far more decorrelation than does the Mojave, as indicated by the white patches in the center (agricultural fields) and right (high mountains). Both the InSAR and the GNSS velocities show a large subsidence bowl in the Central Valley, as expected from previous studies (Faunt et al., 2016; Levy et al., 2020) and known high rates of groundwater pumping in the region. The broad range of velocities across Site B is important to the NISAR team for testing the validation workflow. Visually, it appears there is good agreement between GNSS and InSAR, including in areas of the highest subsidence (dark red).

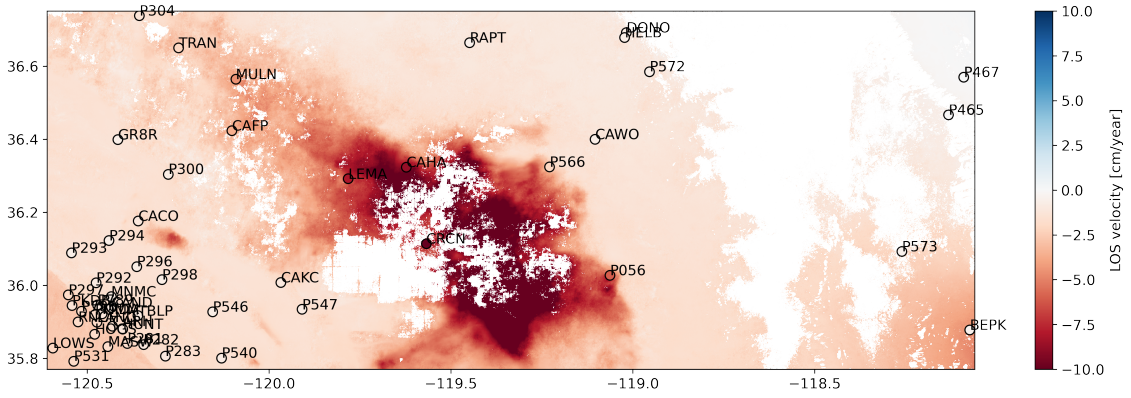

Figure 4.3: InSAR velocity map of site B with GNSS station velocities overlaid, shown within the circular marker and colored to the same scale as InSAR velocities. Reference station is P467, to the far right of the study region.

Next we look at the histogram of the velocities and and residuals for the Central Valley (Figure 4.4). The range of velocity values is significantly larger than in the Mojave, and this is reflected in the much larger standard deviations of the InSAR (1.6 cm/y) and GNSS (2.1 cm/y) distributions. As in the Mojave, the InSAR velocities are biased negatively compared to GNSS, which results in the positive bias in the (GNSS minus InSAR) residuals. Unlike in the Mojave, where the residual standard deviation is

only slightly smaller than that of either InSAR or GNSS, in the Central Valley, the much smaller standard deviation of the residual suggests that the GNSS and InSAR velocities are primarily reflecting common Earth deformation processes (rather than noise).

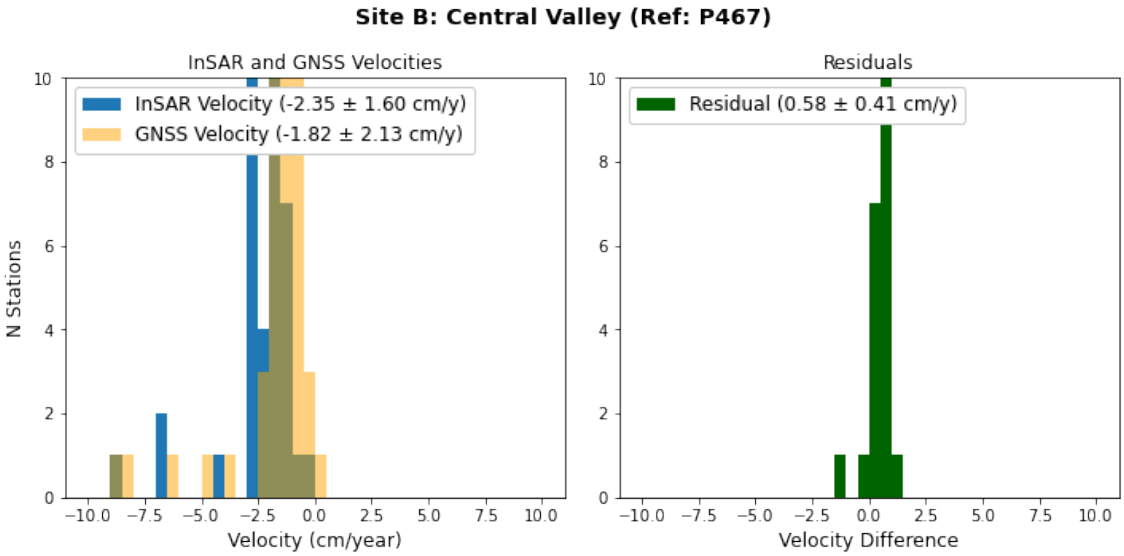

Figure 4.4: Left: the Central Valley velocities and standard deviation with P467 as the reference station. Right: GNSS minus InSAR residuals with standard deviation

Chapter 5

Discussion

After reviewing my initial results, I explored the effect of making different choices in the analysis. I experimented with changing the GNSS reference station, the spatial averaging of InSAR velocity, and applying different corrections to the InSAR dataset.

5.1 Reference Station Selection

I tested different reference stations to see how important the reference station selection was and how much it would effect the InSAR/GNSS comparison. I compared three stations to the originally chosen stations in each analysis region (P626 in the Mojave, P467 in the Central Valley). The first station had the closest to average velocity. The second station was the one whose velocity was closest to zero. The third was a noisy or large-offset station. My intention behind testing stations with velocities near average and closest to zero was to see if the selection the reference station could be automated, rather than via manual user input as it is done now.

Figure 5.1 shows histograms for the InSAR and GNSS velocity distributions for the Mojave. As expected, the shape of the velocity distributions do not change significantly between plots, but the location (i.e., mean velocity value) does. Further, the first three distributions are all very similar, which indicates that the reference station selection is not as critical as I first assumed. However, I recommend continuing to manually select reference stations, using these metrics to help narrow the selection. Additionally, I recommend looking at the timeseries plots for the station before finalizing the selection.

Looking at the timeseries will help prevent selecting a noisy station whose velocity may not reflect that of other stations around it. An example of such a station is station AZBH, on the far right of Figure 5.1, which gives a reference velocity that is not a good match to the InSAR data.

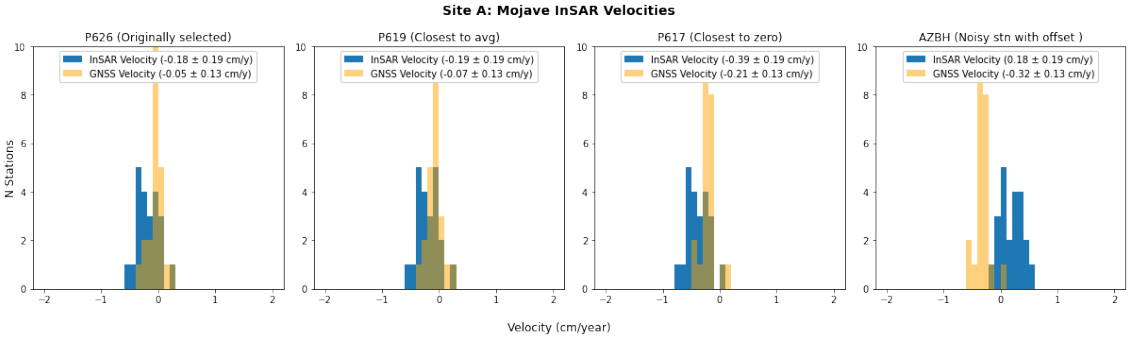

Figure 5.1: InSAR (blue) and GNSS (yellow) velocity distributions for Site A. InSAR velocities remain the same throughout the four graphs while the GNSS velocities shift as different reference stations are applied.

In site B, we see a similar result but with more noise in the GNSS and a larger spread of velocity values. Overall the four reference stations show similar agreement with the InSAR velocities, although LOWS and MIDA show the best match with InSAR velocities (residual medians of 0.27 cm/y and -0.02 cm/y, respectively, compared to 0.71 cm/y for the original reference station P467).

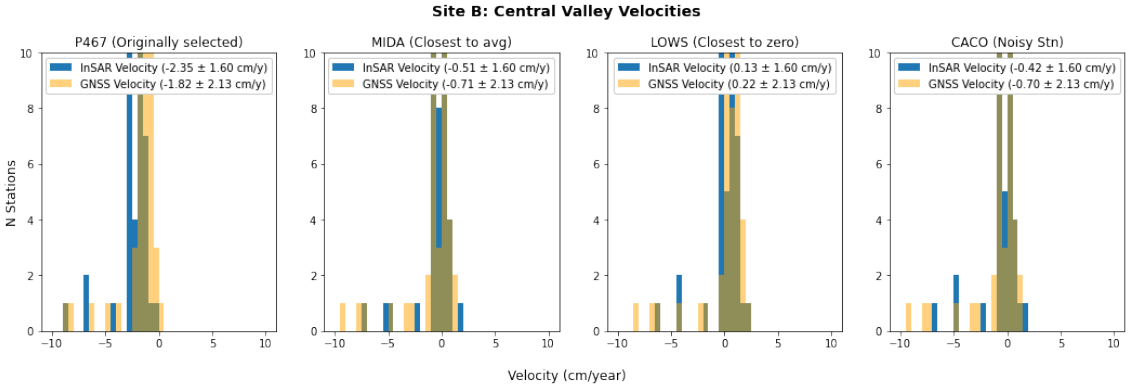

Figure 5.2: InSAR (blue) and GNSS (yellow) velocity distributions for site B, following the convention in Figure 5.1. All four reference stations are reasonable choices.

5.2 Impact of InSAR Corrections

The original InSAR data has no corrections applied. I tested the effects of applying various corrections available in MintPy, for the purpose of identifying which ones are most likely to improve the InSAR/GNSS comparison. A summary of results for each correction is shown in Table 5.1 (Mojave) and Table 5.2 (Central Valley). These tables include the standard deviations of InSAR velocities and the (GNSS minus InSAR) residual velocities, and the best fit line and Pearson correlation coefficient between the two sets of velocities.

Figure 5.3 (Mojave) and Figure 5.4 (Central Valley) visually summarize my findings. Each subplot shows GNSS velocities on the x-axis and uncorrected InSAR velocities on the y-axis, where each point represents a GNSS station location. The diagonal black line indicates perfect 1:1 agreement between GNSS and InSAR, and ideally all points would fall along this line. The colored line in each subplot is the result of a linear fit between GNSS and InSAR velocities. In the legend, I report the slope and intercept of this fit, along with their corresponding uncertainties. The Pearson correlation coefficient, which I report in the tables, is used to see how close the data are to a line. A value of 1 means the data fall perfectly on a line, while a value of 0 means the data are randomly scattered in relation to each other.

With no corrections, the slope of the linear relationship between GNSS and InSAR is 0.323 and the Pearson correlation coefficient is 0.213. This represents poor correlation of InSAR and GNSS, which is what we might expect in a region like the Mojave with little actual deformation. I tested several corrections by editing the project configuration file as explained in Step 4 of Section 3.2. The tropospheric correction (middle-top subplot in Figure 5.3) greatly improved the comparison, yielding a slope of 0.671 and a Pearson coefficient of 0.540. This is an expected results as most InSAR

error comes from the troposphere (Bekaert et al., 2015). The topography correction (middle subplot) gave the best result, with a slope of 0.834 and Pearson coefficient of 0.544. The shapes of the topography and troposphere correction scatterplots are similar, as expected given that the topographic correction is primarily adjusting for radar path length through the troposphere.

Applying linear deramping (top-right plot) changed the scatterplot significantly, but did not improve the linear relationship between GNSS and INSAR, giving a slope of 0.277 and Pearson coefficient of 0.318. The quadratic deramp also did not improve the result, givine a slope of 0.227 and (an identical) Pearson coefficient of 0.227. It is unclear if the deramp worsened the results because the ramp in the velocity map was due to actual deformation observed in both GNSS and InSAR, or if there is an error with how this correction is applied in MintPy. This concern has been brought to the NISAR Solid Earth team for further examination and discussion (see Section 6.1). A combination of corrections was tested and can be seen in Table 5.1, but none of these performed as well as the topographic and tropospheric corrections alone.

While the topography correction is the best, none of the corrections yield a particularly high linear correlation between GNSS and InSAR. As mentioned earlier, the velocity values themselves are low and the range of velocities is so small that any error is magnified. Additionally, some of these corrections make the correlation between the InSAR data and GNSS data worse rather than improving it. This is a very important take-away, since it demands that future users of this workflow must apply these corrections with caution and thorough examination.

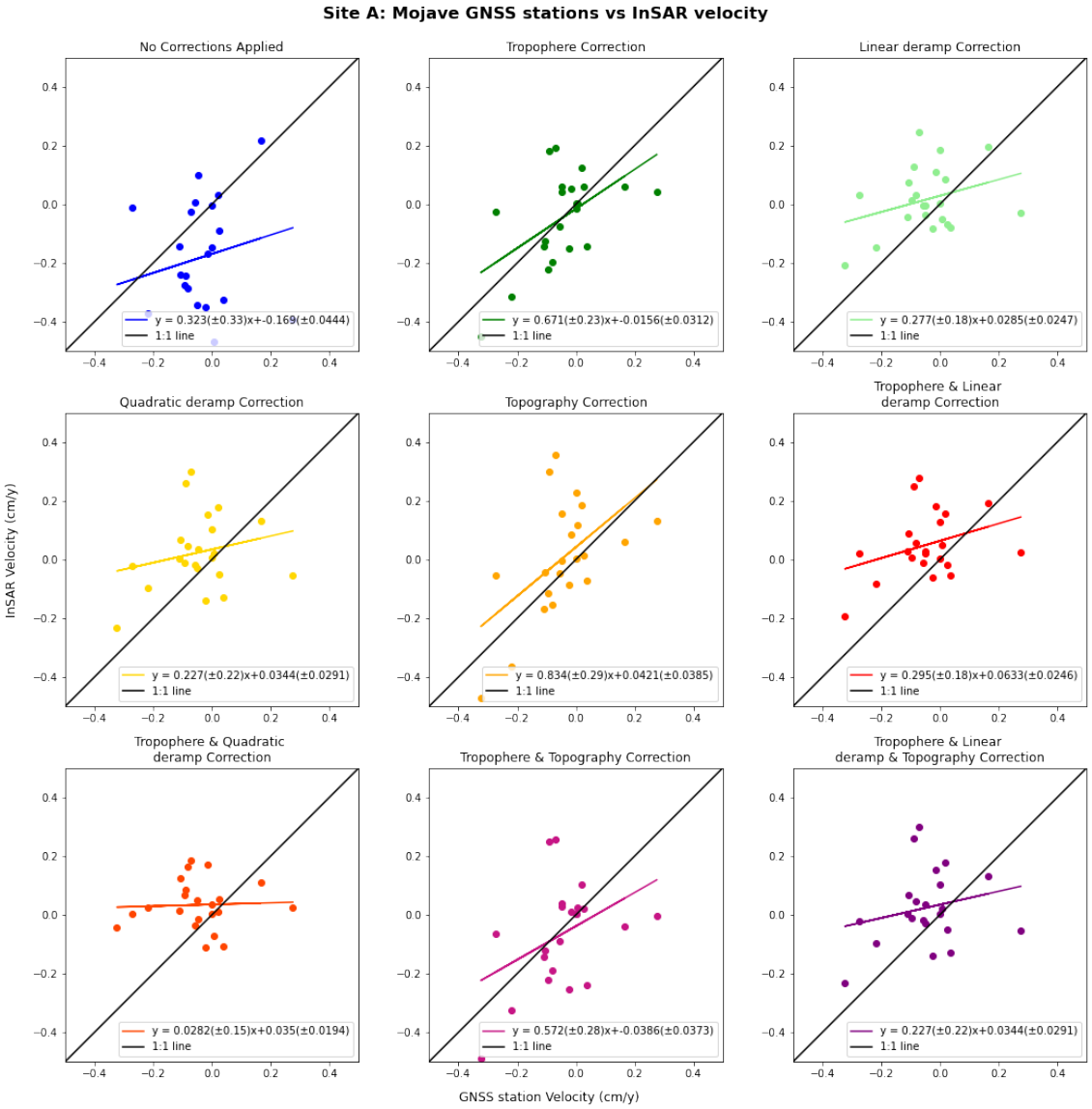

Figure 5.3: Each subplot contains a different correction or correction combination applied to the InSAR data. GNSS stations are plotted with their GNSS velocity along the x-axis and InSAR velocity along the y-axis. Overlaid is the line of best fit, the ideal or perfect fit is shown in black with a 1:1 relation between the datasets.

Table 5.1: Complete statistical list for each correction and correction combinations plotted in 5.3.

Mojave Correction Statistics

	InSAR std	Residual std	Best fit Equation	Pearson Corr Coef
No Correction	0.190281	0.204350	y = 0.323(±0.33)x+-0.169(±0.0444)	0.212811
Troposheric	0.155381	0.137063	y = 0.671(±0.23)x+-0.0156(±0.0312)	0.540779
Linear Deramp	0.108981	0.137446	y = 0.277(±0.18)x+0.0285(±0.0247)	0.318015
Quadratic Deramp	0.125195	0.155681	y = 0.227(±0.22)x+0.0344(±0.0291)	0.227485
Topography	0.191955	0.162333	y = 0.834(±0.29)x+0.0421(±0.0385)	0.544548
Linear Deramp & Troposheric	0.109362	0.135686	y = 0.295(±0.18)x+0.0633(±0.0246)	0.337491
Quadatic Deramp & Troposheric	0.081393	0.146422	y = 0.0282(±0.15)x+0.035(±0.0194)	0.043391
Troposheric & Topography	0.171784	0.165076	y = 0.572(±0.28)x+-0.0386(±0.0373)	0.417189
All (w/linear)	0.125195	0.155681	y = 0.227(±0.22)x+0.0344(±0.0291)	0.227485

For Site B in the Central Valley, with no corrections the line of best fit gives a slope of 0.863 and a Pearson coefficient of 0.978. This is much better than any result for Site A. One reason for this could be that there is a much larger range of velocities in the Central Valley. This makes the Pearson coefficient harder to compare between the two analysis regions, since a few really large values are driving the high Pearson coefficient values.

The tropospheric correction has a slope of 0.904 and Pearson coefficient of 0.986. Just as in site A, this improves the relationship between the two datasets. The topography correction gives a slope of 0.89 and coefficient of 0.982, which is also an improvement from no correction. The linear deramp has a slope of 0.61 and coefficient of 0.941. The quadratic deramp has a slope of 0.613 and coefficient of 0.779. Again, neither deramping correction improves the results, and both must be used with caution. The full lineup of best fit equations can be found in Table 5.2 as well as the correction combination results.

For the Central Valley, applying InSAR corrections does not significantly improve the comparison between GNSS and InSAR. While not including corrections is probably

fine in regions with large deformation, my results do not indicate they need to be added as standard practice.

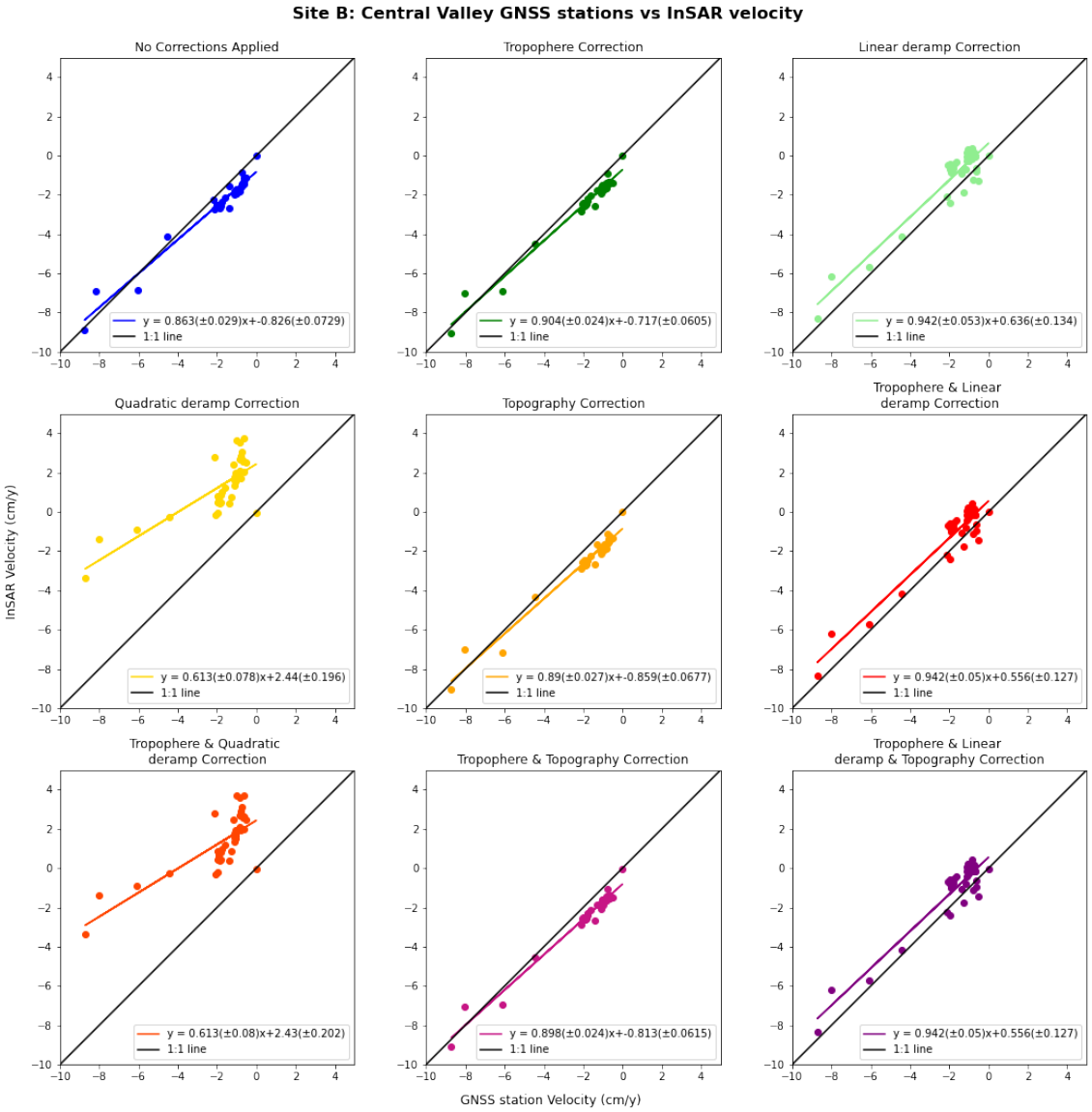

Figure 5.4: Same as Figure 5.3, but for Site B in the Central Valley.

Table 5.2: Same as Table 5.2, but for Site B in the Central Valley.

Central Valley Correction Statistics

	InSAR std	Residual std	Best fit Equation	Pearson Corr Coef
No Correction	1.600237	0.413967	y = 0.863(±0.029)x+-0.826(±0.0729)	0.978394
Troposheric	1.648528	0.322950	y = 0.904(±0.024)x+-0.717(±0.0605)	0.986173
Linear Deramp	1.797864	0.613430	y = 0.942(±0.053)x+0.636(±0.134)	0.941795
Quadratic Deramp	1.412962	1.126205	y = 0.613(±0.078)x+2.44(±0.196)	0.779543
Topography	1.629446	0.364271	y = 0.89(±0.027)x+-0.859(±0.0677)	0.982200
Linear Deramp & Troposheric	1.788925	0.582755	y = 0.942(±0.05)x+0.556(±0.127)	0.947220
Quadatic Deramp & Troposheric	1.431409	1.148563	y = 0.613(±0.08)x+2.43(±0.202)	0.769810
Troposheric & Topography	1.637847	0.333322	y = 0.898(±0.024)x+-0.813(±0.0615)	0.985488
All (w/linear)	1.788925	0.582755	y = 0.942(±0.05)x+0.556(±0.127)	0.947220

5.3 Spatial Averaging of InSAR Velocities

I conducted follow-up analysis on the size of the spatial averaging region for InSAR velocity estimates, to determine how much averaging is optimal for comparison with GNSS station velocity observations. In my workflow, I parameterize the averaging region in terms of the radius in InSAR pixels around a measurement point. This is done after MintPy has estimated a velocity value for each pixel and solely for the purpose of comparing to GNSS stations.

The InSAR velocities shown in Figures 4.1 and 4.3 were created using a 11px by 11px (1km) median smoothing window, as explained in Step 10 of section 3.2. To examine the impact of median window size on the velocity estimate, I compared raw 100m-resolution MintPy velocities against velocities smoothed using median windows with widths of 11px (1km) and 101 px (10km).

Examining Figure 5.5 and Figure 5.6, there is very little difference between the raw velocity estimate and either of the smoothed velocity fields. The residual plots in the right panels of those figures show that changing spatial averaging has much less of

an impact than I initially assumed. It is a very surprising finding that there is almost no residual difference between the 101 px and 1 px plots, which indicates that the velocity measurements are really robust. However, for stations near the edge of the study site the median window may extend beyond the analysis region, which means that large windows may not equally sample velocities on all sides of the station. For that reason, and for speed of processing, I recommend continuing to use the 11 px median window.

In addition to the results shown in the figures here, I also performed the GNSS/ InSAR velocity comparison from 4.1 and 4.2 using different median smoothing of InSAR velocities. These results were almost identical to those obtained using the raw InSAR velocities.

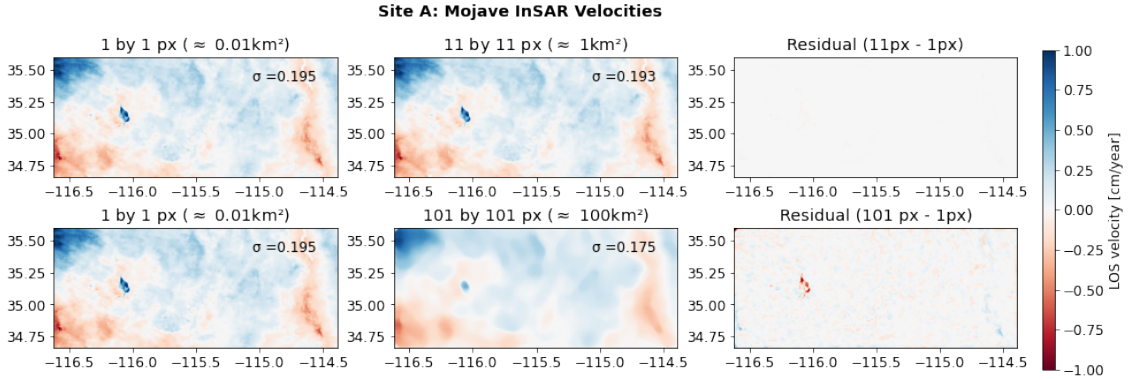

Figure 5.5: Site A InSAR velocities with a median filter applied. Left shows the raw MintPy velocity estimate. Middle shows a 11px (1km) and 101px (10km) filter. Right shows the difference between the raw and filtered images.

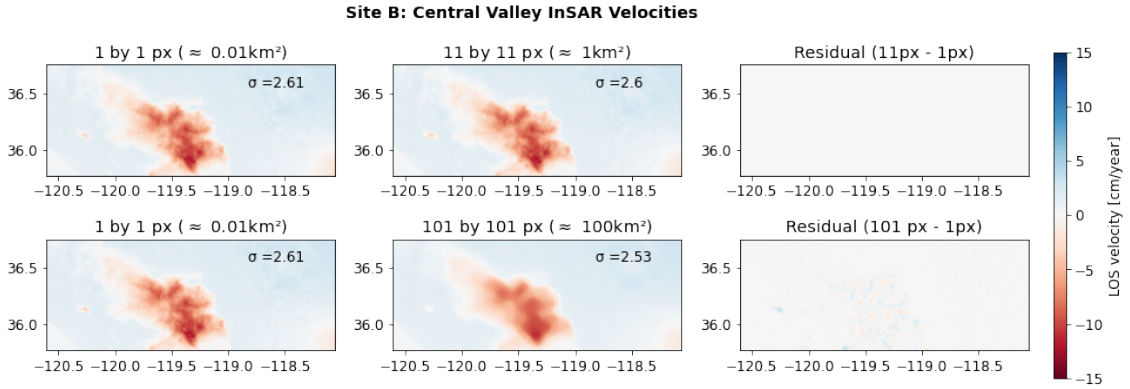

Figure 5.6: Same as Figure 5.5, but for Site B in the Central Valley.

5.4 Meeting NISAR's Validation Requirement

As this work is meant to support NISAR validation, it is important to check if measured velocities meet the required accuracy of 2 mm/y or better over length scales of 0.1 km $\leq$ L $\leq$ 50 km. Since the requirement is written in terms of baseline distances, the validation is performed on relative velocities between various points in the analysis region and not on absolute velocities. The validation still uses GNSS for reference, but instead of comparing GNSS velocities directly with InSAR velocities as I did above, it compares pairwise differences of GNSS and InSAR velocities between GNSS station locations.

Specifically, I calculate baseline distances for all pairs of GNSS station locations in the analysis region. I difference the GNSS LOS velocities and the InSAR velocities for each pair, then take the difference of the differences (a double difference). Since we assume the GNSS velocities to be correct, this double difference represents the assumed error in InSAR velocities over the various baseline distances in the dataset.

In order to assess whether the InSAR velocity error meets NISAR's accuracy requirements, the NISAR Solid Earth Team groups the baselines into 5-km-wide bins and:

> "assumes that the residuals between cGNSS and InSAR follow a Gaussian distribution. For each distance bin, if the fraction of residuals lying below the bin threshold value is more than 0.683 (i.e., one standard deviation), we judge the derived secular deformation rate... to pass the corresponding requirement. (Algorithm Theoretical Basis Document, 2022)"

The ATBD document additionally specifies that only stations within 50 km of each other need to satisfy the velocity requirement. Figure 5.7 shows the double differences for all station pairs in the Mojave analysis region with baselines of less than 50 km. The red line indicates the requirement threshold accuracy of 2mm/yr.

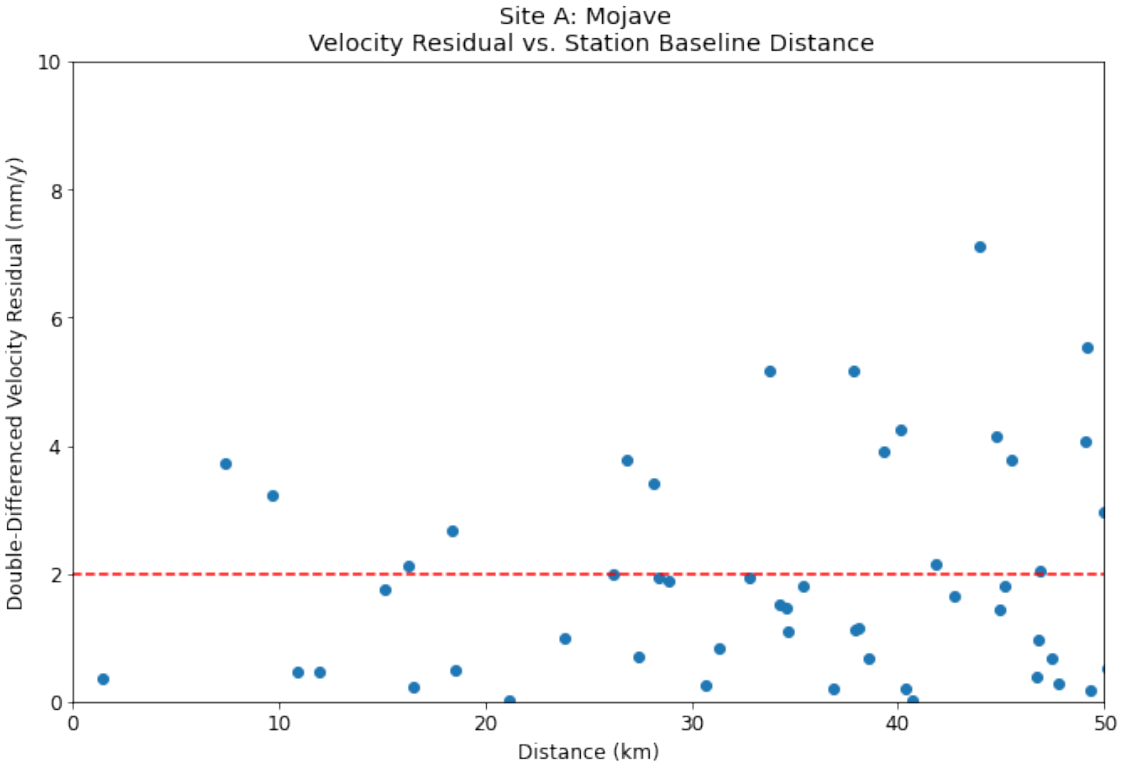

Figure 5.7: The x-axis is the distance between GNSS stations in the Mojave analysis region. On the y-axis is the absolute value of the double-differenced velocity residual. To meet the NISAR requirements for any given distance interval, 68% of the points within that interval must fall below the 2 mm/yr threshold shown in red

Table 5.3 highlights if the secular requirement is met or not. The table columns show results for each of the 5-km-wide distance bins. The first row shows the total number of double differences in the bin, and the second row shows the number of double differences under the 2 mm/y threshold. The third row shows the percentage of differences that meet the requirement, and the forth shows if the InSAR velocities in a given bin pass (true) or fail (green) to meet requirement. Site A just fails the requirement, with 67.5% of all points lying below the 2 mm/y threshold. I think this is mostly due to the limited data within the 50km range.

Table 5.3: Site A pass/fail validation table. Row 1 shows the total stations in a given distance bin, Row 2 shows the number of stations that meet the requirement, Row 3 shows the percent that pass the 2mm/yr threshold where red indicates failure and green indicates passing.

	0.10-5.09	5.09-10.08	10.08-15.07	15.07-20.06	20.06-25.05	25.05-30.04	30.04-35.03	35.03-40.02	40.02-45.01	45.01-50.00	total
20180105-20200101	1	2	2	5	2	6	7	7	8	11	40

	0.10-5.09	5.09-10.08	10.08-15.07	15.07-20.06	20.06-25.05	25.05-30.04	30.04-35.03	35.03-40.02	40.02-45.01	45.01-50.00	total
20180105-20200101	1	0	2	3	2	4	6	5	4	6	27

	0.10-5.09	5.09-10.08	10.08-15.07	15.07-20.06	20.06-25.05	25.05-30.04	30.04-35.03	35.03-40.02	40.02-45.01	45.01-50.00	total
20180105-20200101	1.000000	0.000000	1.000000	0.600000	1.000000	0.666667	0.857143	0.714286	0.500000	0.545455	0.675000

	0.10-5.09	5.09-10.08	10.08-15.07	15.07-20.06	20.06-25.05	25.05-30.04	30.04-35.03	35.03-40.02	40.02-45.01	45.01-50.00	total
20180105-20200101	true	false	true	false	true	false	true	true	false	false	false

```
This velocity dataset does not pass the requirement.
```

Site B in the Central Valley passes the NISAR validation requirement, with 69.9% of the double differences at baselines $\leq$ 50 km lying below the 2 mm/y requirement. This is a significant result that demonstrates that InSAR data hold up in actively deforming regions. It is unexpected that Site B performed better than site A. This could simply be due to the fact that there are significantly more stations within 50 km of each other in the Central Valley, as seen in Figure 5.8 and Figure 5.4.

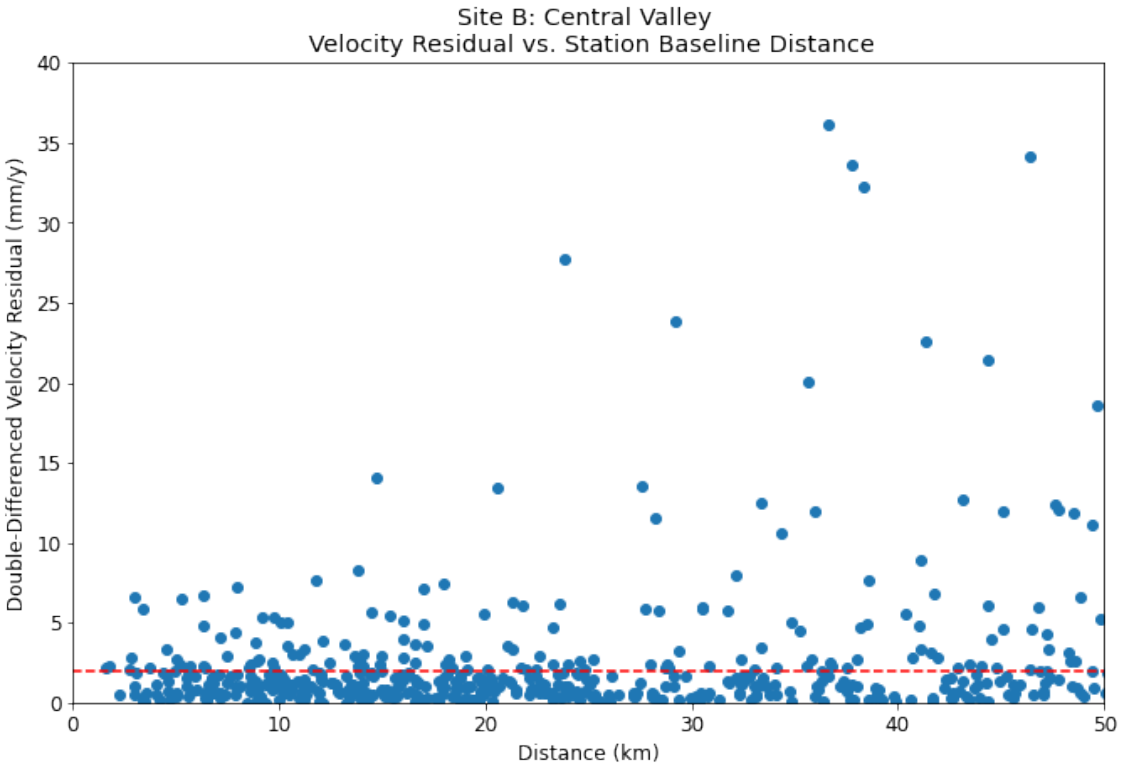

Figure 5.8: Same as Figure 5.7, but for the Central Valley. There are many more double differences in this figure, especially at short distances (due to the concentration of stations around Parkfield at the southwest corner of Figure 4.3)

Table 5.4: Same as Table 5.3, but for Site B in the Central Valley.

	0.10-5.09	5.09-10.08	10.08-15.07	15.07-20.06	20.06-25.05	25.05-30.04	30.04-35.03	35.03-40.02	40.02-45.01	45.01-50.00	total
20180103-20201007	28	73	75	74	58	40	42	38	37	38	465

	0.10-5.09	5.09-10.08	10.08-15.07	15.07-20.06	20.06-25.05	25.05-30.04	30.04-35.03	35.03-40.02	40.02-45.01	45.01-50.00	total
20180103-20201007	19	50	55	57	42	29	30	23	20	19	325

	0.10-5.09	5.09-10.08	10.08-15.07	15.07-20.06	20.06-25.05	25.05-30.04	30.04-35.03	35.03-40.02	40.02-45.01	45.01-50.00	total
20180103-20201007	0.678571	0.684932	0.733333	0.770270	0.724138	0.725000	0.714286	0.605263	0.540541	0.500000	0.698925

	0.10-5.09	5.09-10.08	10.08-15.07	15.07-20.06	20.06-25.05	25.05-30.04	30.04-35.03	35.03-40.02	40.02-45.01	45.01-50.00	total
20180103-20201007	false	true	true	true	true	true	true	false	false	false	true

```
This velocity dataset passes the requirement.
```

Chapter 6

Recommendations and Limitations

6.1 Recommendations for MintPy Updates

MintPy has extensive built-in functionality, however some are quiet buried. A great deal of my time on this project was spent reading through the code to find the right function I needed to call. More examples and documentation would allow users to better understand this tool. Further documentation is being developed and will be released by the NISAR mission. While developing this notebook, significant effort went into making it as user friendly as possible and ease of use was a top consideration throughout this process.

I also recommend changing how the deramp correction is applied. This correction is currently being done to the entire stack while it would be much better to do to each interferogram before the stack is built. This issue has been brought to the attention to the developers and is being worked on.

Another place MintPy could be improved is by implementing a different method to align the InSAR data to the GNSS. Rather than using a single reference station you could align to the entire network of GNSS stations as done in the GInSAR method (Neely et al., 2020).

6.2 Limitations

A limitation to using this workflow is the high start-up cost. There are many detailed steps required to set up the Python environment required to run this notebook. While there is documentation, it is easy to get stuck.

Data availability is another limitation. For this workflow to work there has to be preexisting processed interferograms. As can be seen in the ASF Vertex archive of premade Sentinel-1 interferograms (the "S1 InSAR beta" products), these data are currently limited. ASF is currently working to expand their interferogram archive, which will increase the functionality of the workflow I have developed. The NISAR mission will produce interferograms as a standard product, so I anticipate this will not be a major problem for NISAR.

GNSS data are limited as well. While the study sites selected for this work have significant GNSS station coverage, that is not always the case. In an area of poor or no GNSS coverage but small expected deformation, it is possible to analyze the distribution of InSAR velocities with respect to zero velocity, but it is better to have the actual GNSS velocites for comparison.

InSAR signals contain noise, which can be significant and effect velocity validation results, especially over short timeframes. However, using many interferograms over long periods, velocity errors become negligible. Unfortunately, lengthening the analysis timeframe is not always an easy solution, since it will increase workflow run time in the notebook and require more storage space for additional interferograms.

Notebook speed is the another limitation. Downloading the InSAR data from vertex can takes several hours and single steps within the notebook can take several minutes. This issue has been discussed and addressed by the NISAR Solid Earth Science team. To solve this limitation, they have been migrating the calibration and validation

notebooks to the cloud, using ASF's openSARLabs resources.

Finally, this notebook works best when run once for a single result. It does not work well when running multiple tests, such as I did when testing processing parameters for Chapter 5. File management and naming rapidly become tedious and difficult, as there is not a great way to save output velocities from one test to another. For the purpose of this work, I exported each test as a csv file and had a separate notebook to compare results from different tests.

Chapter 7

Conclusions

This work directly helped the NISAR Solid Earth Science team meet their calibration and validation proof-of-concept goals. In order to use this workflow, some prior knowledge of GNSS is helpful and I highly recommend plotting the stations daily LOS positions, manually flagging any offsets times that appear, and incorporating offsets into the velocity estimation model. Additionally, I recommend adding a troposphere correction to interferograms, especially in areas of high topographic relief.

The workflow described in this thesis will continue to be developed and will serve as a public resource to support the NISAR mission and the use of NISAR data. As mission preparation continues, my validation results in the Mojave and Central Valley should provide valuable information about what to expect with NISAR validation. The Central Valley study site met the velocity requirement, while the Mojave site just barely missed this requirement, and further work will determine if longer analysis periods or the switch from Sentinel-1's C-Band observations to NISAR's L-Band observations will make a difference.

After this analysis, I do not think InSAR should be used in place of GNSS, but rather to supplement it. Using InSAR to supplement GNSS data helps shed light on continuous deformation that might be missing from GNSS data alone. While some stations match incredibly well, and the requirement was met in the Central Valley I believe the velocity requirement set a fairly low bar to pass. Additionally, I do not think InSAR is absolutely reliable, as there are some station locations that show significant velocity differences.

This work could be improved if corner reflectors were at the GNSS sites. This would make the results more robust. As it is, we are assuming the signal we see is displacement, however there are a lot of different signals that could be causing error in both the GNSS and InSAR displacements. By using corner reflectors, the signal would be so strong, other signals become arbitrary and would remove any ambiguity. However, doing this analysis at corner reflectors would require doing point-scatterer InSAR inversion rather than a SBAS inversion, which is not currently possible with this workflow process.

References

D.P.S. Bekaert, R.J. Walters, T.J. Wright, A.J. Hooper, and D.J. Parker. Statistical Comparison of InSAR Tropospheric Correction Techniques. *Remote Sensing of Environment*, 170:40–47, 2015. ISSN 0034-4257. doi: https://doi.org/10.1016/j.rse.2015.08.035. URL `https://www.sciencedirect.com/science/article/pii/S0034425715301231`.

Geoffrey Blewitt, William C. Hammond, and Corné Kreemer. Harnessing the GPS Data Explosion for Interdisciplinary Science. *EOS*, September 2018. URL `https://eos.org/science-updates/harnessing-the-gps-data-explosion-for-interdisciplinary-science`.

Alvar Escriva-Bou, Jelena Jezdimirovic, and Ellen Hanak. Sinking Lands, Damaged Infrastructure: Will Better Groundwater Management End Subsidence? *Public Policy Institute of California*, May 2020. URL `https://www.ppic.org/blog/sinking-lands-damaged-infrastructure-will-better-groundwater-management-end-subsidence/`.

Alaska Satellite Facility. How to Phase Unwrap an Interferogram. `https://asf.alaska.edu/how-to/data-recipes/phase-unwrap-an-interferogram/`.

J. Famiglietti, Min-Hui Lo, S. Ho, James Bethune, K. Anderson, Tajdarul Syed, S. Swenson, Caroline De Linage, and Matthew Rodell. Satellites Measure Recent Rates of Groundwater Depletion in California's Central Valley. *Geophysical Research Letters - GEOPHYS RES LETT*, 38, 02 2011. doi: 10.1029/2010GL046442.

Claudia C Faunt, Michelle Sneed, Jon Traum, and Justin T. Brandt. Water Availability and Land Subsidence in the Central Valley, California, USA. *Hydrogeology Journal*, 24(3):675–684, 2016.

Alessandro Ferretti, Claudio Prati, and Fabio Rocca. Analysis of Permanent Scatterers in SAR interferometry. *IEEE Trans. Geosci. Remote Sens*, 39:761 – 763 vol.2, 2000. doi: 10.1109/IGARSS.2000.861695.

Jessica Garron, Chris Stoner, Andrew Johnston, and Scott Arko. Sentinel-1 Interferometry from the Cloud to the Scientist. `https://asf.alaska.edu/wp-content/uploads/2019/03/grfn_agu_2017v2.pptx.pdf`, 2019.

A. Laurence Gray, Karim E. Mattar, and George Sofko. Influence of ionospheric electron density fluctuations on satellite radar interferometry. *Geophysical Research Letters*, 27:1451–1454, 2000. doi: https://doi.org/10.1029/2000GL000016.

NASA Jet Propulsion Laboratory. ARIA-tools. `https://github.com/aria-tools/ARIA-tools`. GitHub.

NASA Jet Propulsion Laboratory. NISAR Science Algorithms Solid earth ATBD. https://gitlab.com/nisar-science-algorithms/solid-earth/ATBD, 2013.

Morgan Levy, Wesley Neely, Adrian Borsa, and Jennifer Burney. Fine-Scale Spatiotemporal Variation in Subsidence Across California's San joaquin valley explained by groundwater demand. *Environmental Research Letters*, 15, 10 2020. doi: 10.1088/1748-9326/abb55c.

NASA. *Algorithm Theoretical Basis Document: Algorithms to Validate NISAR L2 Coseismic, Transient and Secular Displacement Requirements*. NISAR Solid Earth Team.

NASA. NISAR Mission Concept. URL https://nisar.jpl.nasa.gov/. NASA Jet Propulsion Laboratory California Institute of Technology.

NASA. *NASA-ISRO SAR (NISAR) Mission Science Users' Handbook*. NISAR Science Team, 2018.

Wesley R. Neely, Adrian A. Borsa, and Francesca Silverii. GInSAR: A cGPS Correction for Enhanced InSAR Time Series. *IEEE Transactions on Geoscience and Remote Sensing*, 58(1):136–146, 2020. doi: 10.1109/TGRS.2019.2934118.

California Department of Water Resources Sustainable Groundwater Management Program. BPM 6 Sustainable Management Criteria, November 2017. URL https://water.ca.gov/-/media/DWR-Website/Web-Pages/Programs/Groundwater-Management/Sustainable-Groundwater-Management/Best-Management-Practices-and-Guidance-Documents/Files/BMP-6-Sustainable-Management-Criteria-DRAFT_ay_19.pdf.

Paul Rosen, Eric Gurrola, Gian Sacco, and Howard Zebker. The InSAR Scientific Computing Environment. 2012.

David Sandwell, Rob Mellors, Xiaopeng Tong, Matt Wei, and Paul Wessel. Open Radar Interferometry Software for Mapping Surface Deformation. 92, 07 2011. doi: 10.1029/2011EO280002.

Ryan G Smith, Rosemary Knight, Chen J, Reeves J A, Zebker H A, Farr T, and Liu Z. Estimating the Permanent Loss of Groundwater Storage in the Southern San Joaquin Valley, California. *Water Resources Research*, 53:2133–2148, 2017. doi: doi: 10.1002/2016WR019861.

Meng Wei, David Sandwell, and Bridget Smith-Konter. Optimal Combination of InSAR and GPS for Measuring Interseismic Crustal deformation. *Advances in Space Research*, 46(2):236–249, 2010. ISSN 0273-1177. doi: https://doi.org/10.1016/j.asr.2010.03.013. URL https://www.sciencedirect.com/science/article/pii/S0273117710001845. GNSS Remote Sensing-1.

Charles L. Werner, Urs Wegmüller, and Tazio Strozzi. Processing Strategies for Phase Unwrapping for InSAR Applications. 2002. URL https://www.gamma-rs.ch/uploads/media/2002-4_PhaseUnwrapping.pdf.

Paul Wessel and Walter Smith. A Global, Self-Consistent, Hierarchical, High-Resolution Shoreline Database. *Journal of Geophysical Research*, 101:8741–8743, 04 1996. doi: 10.1029/96JB00104.

Zhang Yunjun, Heresh Fattahi, and Falk Amelung. Small Baseline InSAR Time Series Analysis: Unwrapping Error Correction and Noise Reduction. *Computers Geosciences*, 133:104331, 2019. ISSN 0098-3004. doi: https://doi.org/10.1016/j.cageo.2019.104331. URL https://www.sciencedirect.com/science/article/pii/S0098300419304194.