Posted on 26 Nov 2022 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.1002/essoar.10503476.1 — This a preprint and has not been peer reviewed. Data may be preliminary.

Comparison of InSAR time series generation techniques as part of the collaborative GeoSCIFramework research project

Brianna Corsa[1], Kristy Tiampo[1], Krisztina Kelevitz[2], Scott Baker[3], and Charles Meertens[3]

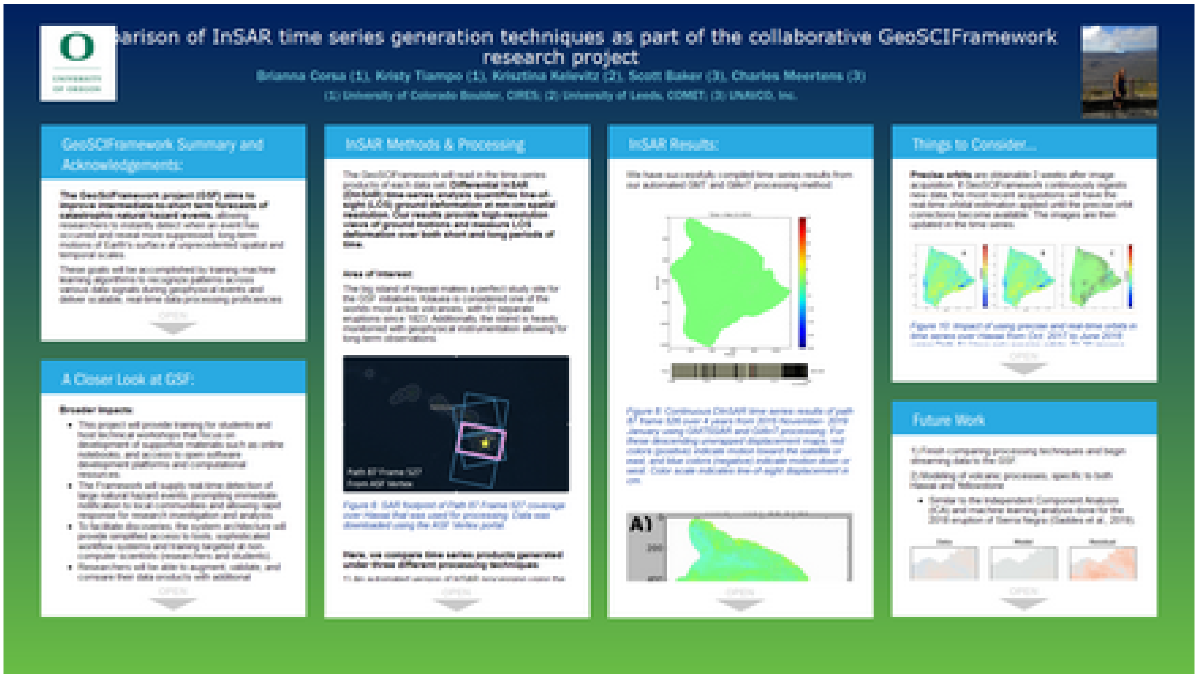
[1]University of Colorado Boulder, CIRES
[2]University of Leeds, COMET
[3]UNAVCO, Inc.

November 26, 2022

Abstract

The GeoSciFramework project (GSF), funded by the NSF Office of Advanced Cyberinfrastructure and NSF EarthCube programs, aims to improve intermediate-to-short term forecasts of catastrophic natural hazard events, allowing researchers to instantly detect when an event has occurred and reveal more suppressed, long-term motions of Earth's surface at unprecedented spatial and temporal scales. These goals will be accomplished by training machine learning algorithms to recognize patterns across various data signals during geophysical events and deliver scalable, real-time data processing proficiencies for time series generation. The algorithm will employ an advanced convolutional neural network method wherein spatio-temporal analyses are informed both by physics-based models and continuous datasets, including Interferometric Synthetic Aperture Radar (InSAR), seismic, GNSS, tide gauge, and gas-emission data. The project architecture accommodates increasingly large datasets by implementing similar software packages already proven to support internet searches and intelligence gathering. This talk will focus primarily on the Differential InSAR (DInSAR) time-series analysis component, which quantifies line-of-sight (LOS) ground deformation at mm-cm spatial resolution. Here, we compare time series products generated under three different processing techniques. The first, an automated version of InSAR processing using the small baseline subset (SBAS) method performed in parallel on systems such as Generic Mapping Tool SAR (GMT5SAR) and the Generic InSAR Analysis Toolbox (GIAnT). The second method will resemble the first but will implement different processing systems for performance comparison using the InSAR Scientific Computing Environment (ISCE) and the Miami InSAR Time Series Software in Python (MintPy). The final strategy, developed by Drs. Zheng and Zebker from Stanford University, concentrates on the topographic phase component of the SAR signal so that simple cross multiplication returns an observation sequence of interferograms in geographic coordinates [Zebker, 2017]. Our results provide high-resolution views of ground motions and measure LOS deformation over both short and long periods of time.

Comparison of InSAR time series generation techniques as part of the collaborative GeoSCIFramework research project



Brianna Corsa (1), Kristy Tiampo (1), Krisztina Kelevitz (2), Scott Baker (3), Charles Meertens (3)

(1) University of Colorado Boulder, CIRES; (2) University of Leeds, COMET; (3) UNAVCO, Inc.

PRESENTED AT:

GEOSCIFRAMEWORK SUMMARY AND ACKNOWLEDGEMENTS:

The GeoSciFramework project (GSF) aims to improve intermediate-to-short term forecasts of catastrophic natural hazard events, allowing researchers to instantly detect when an event has occurred and reveal more suppressed, long-term motions of Earth's surface at unprecedented spatial and temporal scales.

These goals will be accomplished by training machine learning algorithms to recognize patterns across various data signals during geophysical events and deliver scalable, real-time data processing proficiencies for time series generation.

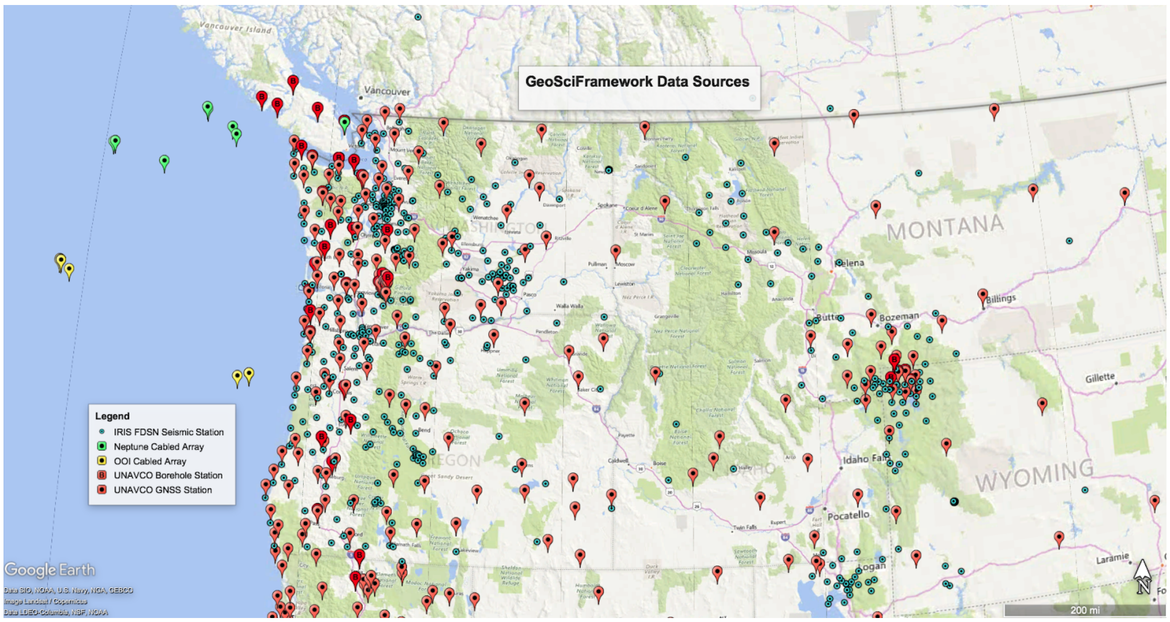


Figure 1: Map showing some of the data sources that GeoSCIFramework will ingest, process, and stream including onshore and offshore seismometers, seismograph stations, borehole strainmeters, and GPS/GNSS stations.

The algorithm will employ an advanced convolutional neural network method wherein spatio-temporal analyses are informed both by physics-based models and continuous datasets, including Interferometric Synthetic Aperture Radar (InSAR), seismic, GNSS, tide gauge, and gas-emission data.

The project architecture accommodates increasingly large datasets by implementing similar software packages already proven to support internet searches and intelligence gathering. The analysis of geophysical data on a global extent is a petabyte-scale Big Data problem that will be addressed using NSF XSEDE resources.

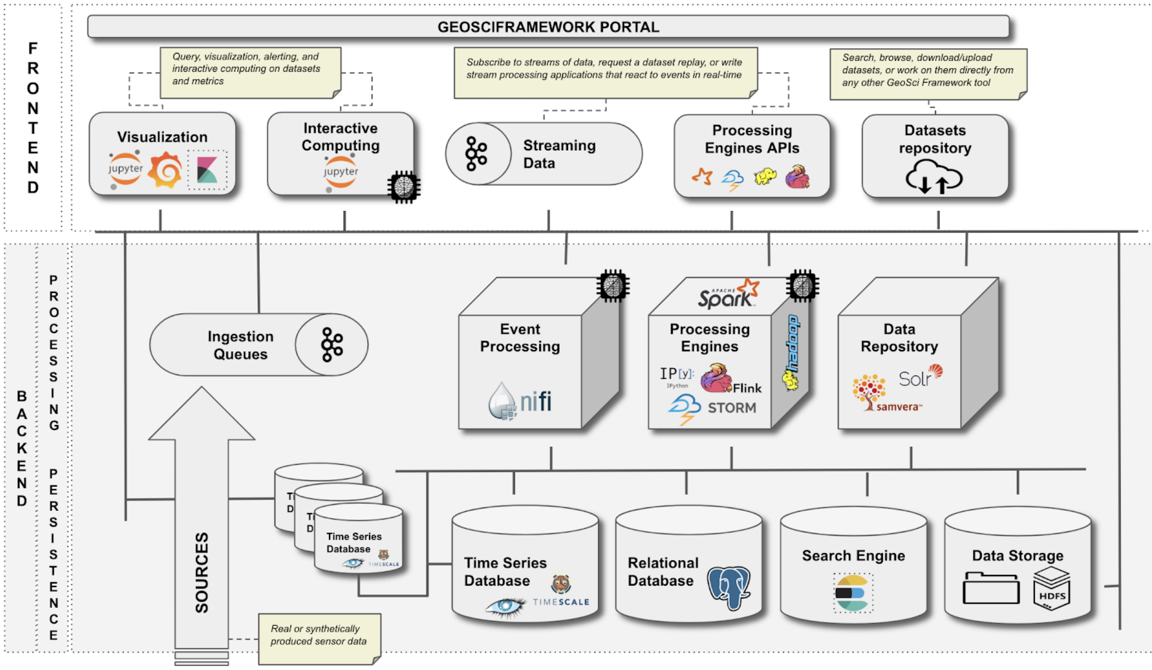


Figure 2: GeoSCIFramework Project Architecture. At the top, "Frontend" depicts the different user-facing components that the GSF provides: from visualization and interactive computing tools to streaming data, static datasets, and even direct access to advanced analytics frameworks. At bottom, "Backend" shows the different components that take care of the heavy lifting: ingesting, curating and consolidating data, computational processing, and dataset management.

GeoSciFramework is funded through the National Science Foundation 2018 Solicitation: NSF 18-531 Cyberinfrastructure for Sustained Scientific Innovation (CSSI) and through NSF EarthCube programs

Participating Institutions: UNAVCO/GAGE, *[NSF Award # 1835791]*; University of Colorado (CIRES) *[NSF Award # 1835566];* University of Oregon *[NSF Award # 1835661]*; and Rutgers University (Ocean Observatories Initiative - OOI) *[NSF Award # 1835692]*

Collaborating Institutions: IRIS/SAGE, University of Texas Arlington (TACC/XSEDE)

A CLOSER LOOK AT GSF:

Broader Impacts:

- This project will provide training for students and host technical workshops that focus on development of supportive materials such as online notebooks, and access to open software development platforms and computational resources.
- The Framework will supply real-time detection of large natural hazard events, prompting immediate notification to local communities and allowing rapid response for research investigation and analysis.
- To facilitate discoveries, the system architecture will provide simplified access to tools, sophisticated workflow systems and training targeted at non-computer scientists (researchers and students).
- Researchers will be able to augment, validate, and compare their data products with additional monitoring data.
- Scientists working on data integration, such as satellite radar with in-situ ground deformation measurements, will have easy and open access to multi-data real-time platforms.

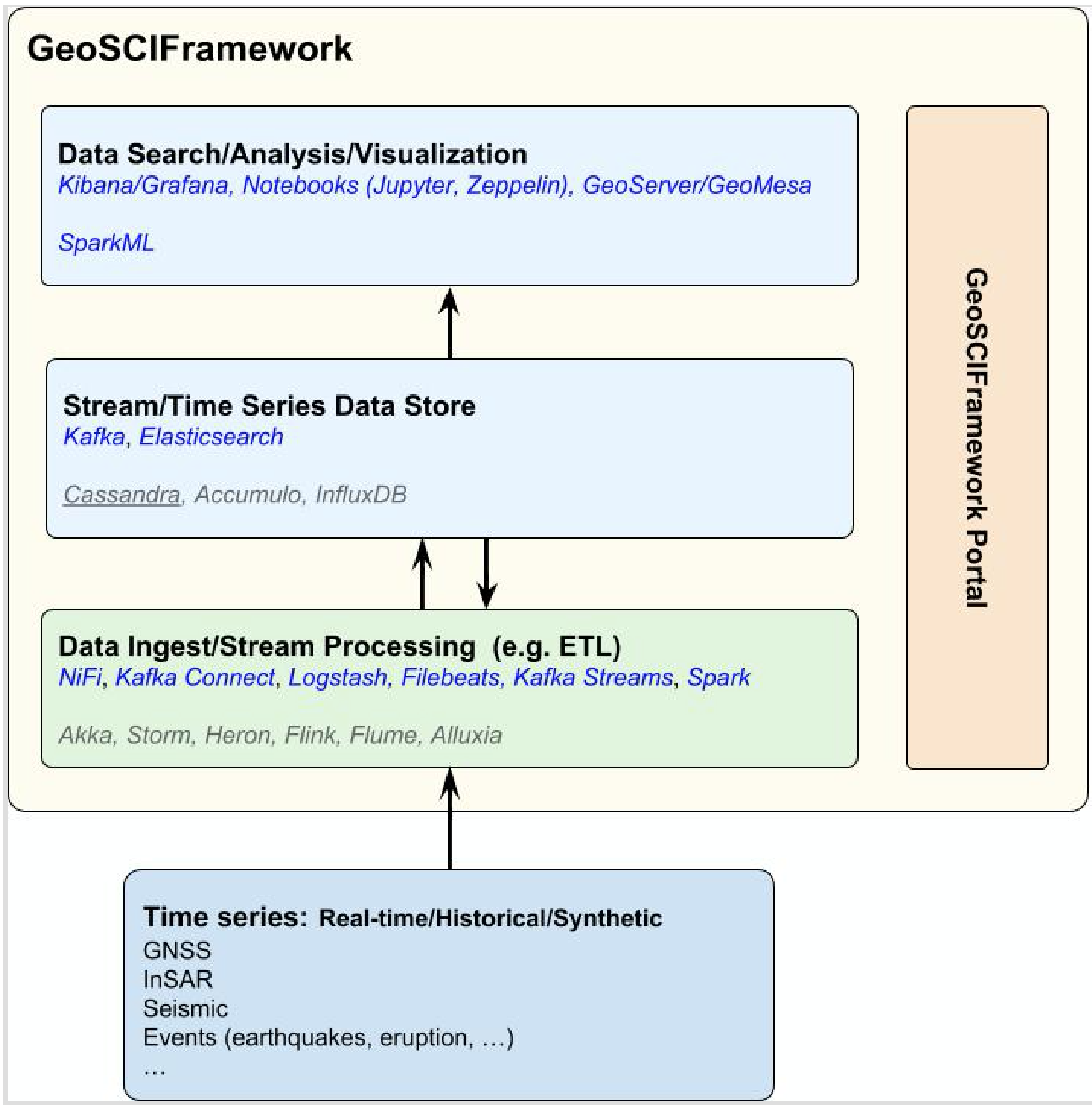


Figure 3: The flowchart illustrates the four main elements comprising the GeoSCIFramework where data is acquired, analysed and presented to users. 1) Data Ingest/Stream Processing layer is the point of entry for time series streamed from different sensor networks. 2) The Stream/Time Series Data Store layer is responsible for storing raw and curating time series in a format suitable for analytics. 3) Data Search/Analysis/Visualization layer manages computing resources, analytics and machine learning framework. 4) GeoSCIFramework Portal is the user-facing layer.

Progress of Other Data Sets in GSF:

Seismic/Tide Gauge:

The goal is to predict or forecast shaking intensity and tsunami heights caused by earthquakes. Dr. Diego Melgar and his team at the University of Oregon simulate magnitude 7-9 earthquakes in Chile, (termed fakequakes), and integrate them with real event data in order to train the machine learning algorithms.

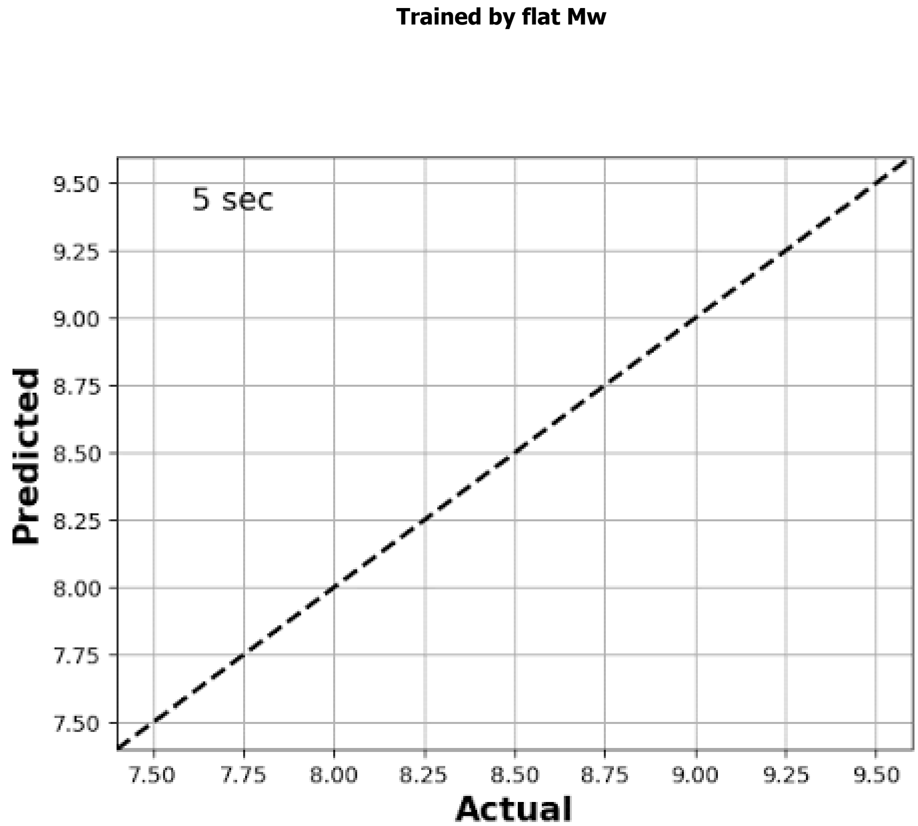


Figure 4: Initial performance tests of machine learning algorithms that were trained on both synthetic and real data sets show highly correlated prediction processing for magnitude 7-9 earthquakes.

GPS/GNSS:

Drs. Dave Mencin and Kathleen Hodgkinson at UNAVCO head the GPS/GNSS division of GSF.

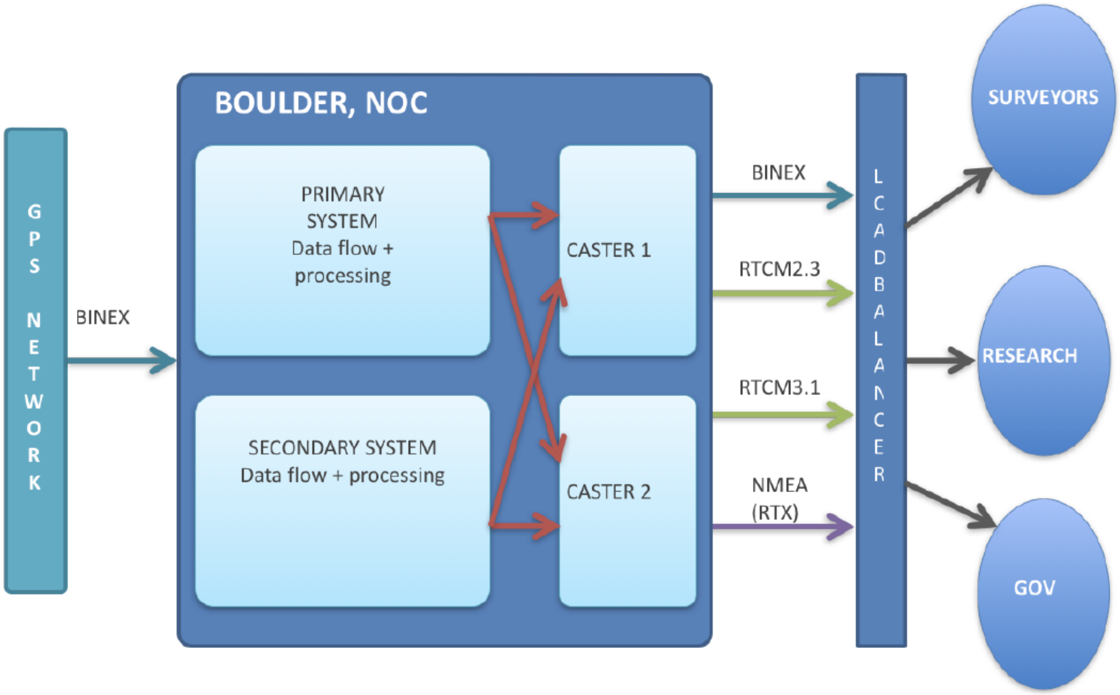


Figure 5: Schematic of cloud based, redundant GNSS data flow, data processing and data distribution used to support real-time GNSS in the GeoSciFramework. Chosen protocols (BINEX, RTCM, NTRIP) are standards used by the community.

For more information, please visit our website here: GeoSCIFramework (http:// https://www.unavco.org/projects/major-projects/earthcube/geosciframework/geosciframework.html)

INSAR METHODS & PROCESSING

The GeoSCIFramework will read in the time-series products of each data set. **Differential InSAR (DInSAR) time-series analysis quantifies line-of-sight (LOS) ground deformation at mm-cm spatial resolution. Our results provide high-resolution views of ground motions and measure LOS deformation over both short and long periods of time.**

Area of Interest:

The big island of Hawaii makes a perfect study site for the GSF initiatives. Kilauea is considered one of the worlds most active volcanoes, with 61 separate eruptions since 1823. Additionally, the island is heavily monitorred with geophysical instrumentation allowing for long-term observations.

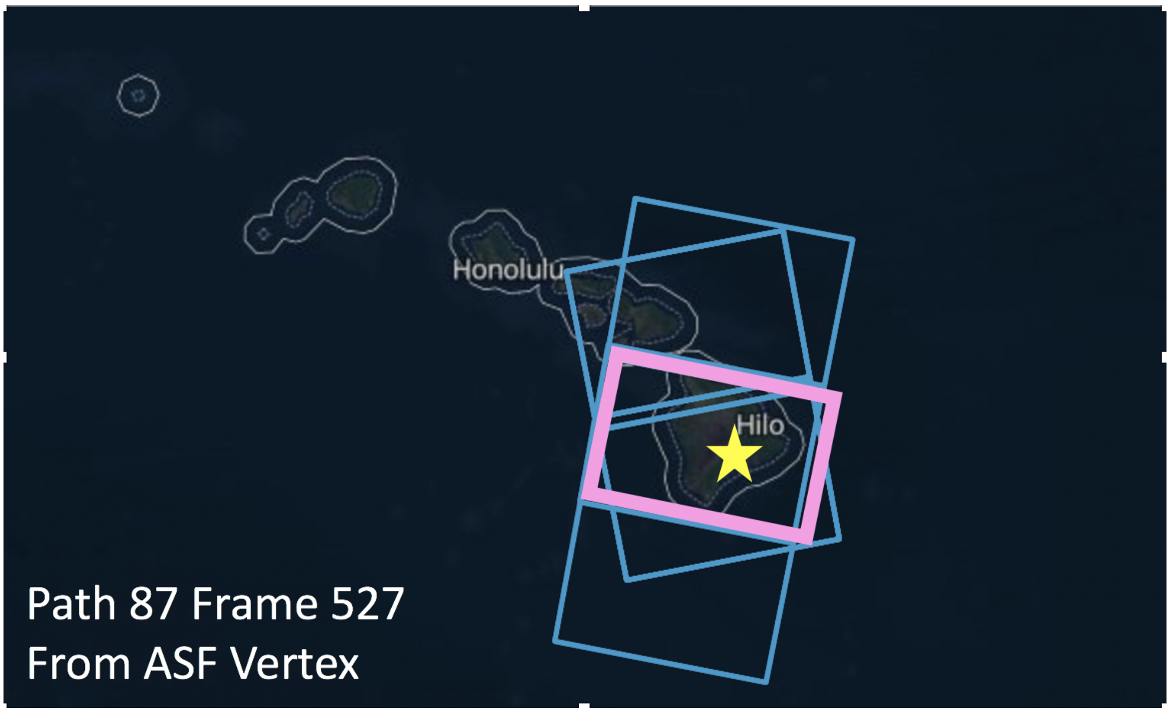


Figure 6: SAR footprint of Path 87 Frame 527 coverage over Hawaii that was used for processing. Data was downloaded using the ASF Vertex portal.

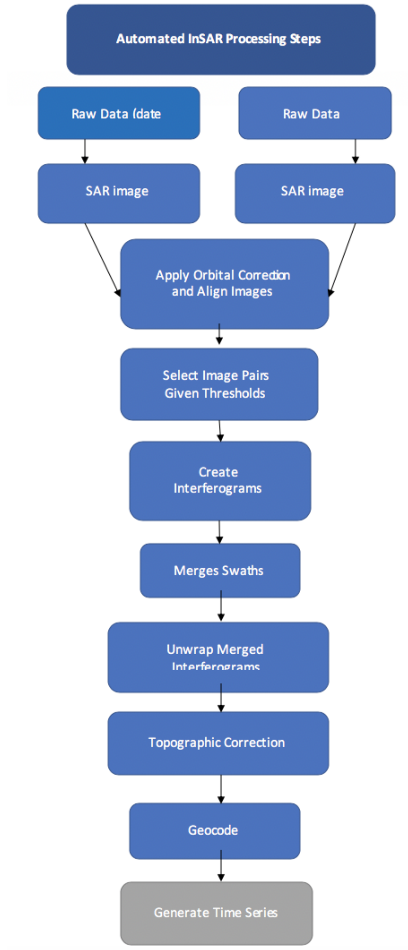
Here, we compare time series products generated under three different processing techniques:

1) An automated version of InSAR processing using the small baseline subset (SBAS) method performed in parallel on systems such as Generic Mapping Tool SAR (**GMT5SAR**) and the Generic InSAR Analysis Toolbox (**GIAnT**).

2) For performance comparison, we use the InSAR Scientific Computing Environment (**ISCE**) and the Miami InSAR Time Series Software in Python (**MintPy**).

3) The final strategy, developed by Drs. Zheng and Zebker from Stanford University, concentrates on the topographic phase component of the SAR signal so that simple cross multiplication returns an observation sequence of interferograms in geographic coordinates [Zebker, 2017].

A)



Automated InSAR Processing Steps
Raw Data (date
Raw Data
SAR image
SAR image
Apply Orbital Correction and Align Images
Select Image Pairs Given Thresholds
Create Interferograms
Merges Swaths
Unwrap Merged Interferograms
Topographic Correction
Geocode
Generate Time Series

B)

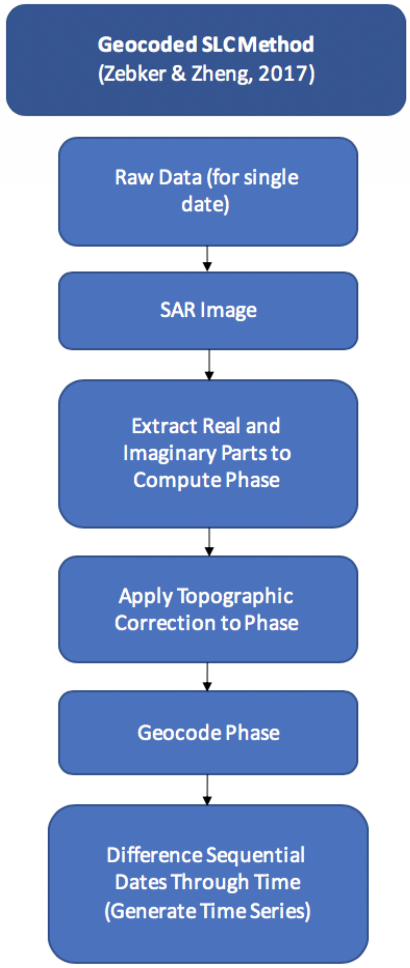


Figure 7: Flow chart in A) represents the typical processing steps that an InSAR program, such as GMT5SAR or ISCE, follows. These utilize both the phase and amplitude components of the SAR signal. Time series are generated using secondary programs such as GIAnT or MintPy. The flow chart in B) depicts the new theory proposed by Drs. Zheng and Zebker, which only utilizes the topographic phase to detect surface deformation between two acquisitions and is capable of directly building time series. This method is less computationally expensive and produces results in much less time.

INSAR RESULTS:

We have successfully compiled time series results from our automated GMT and GIAnT processing method.

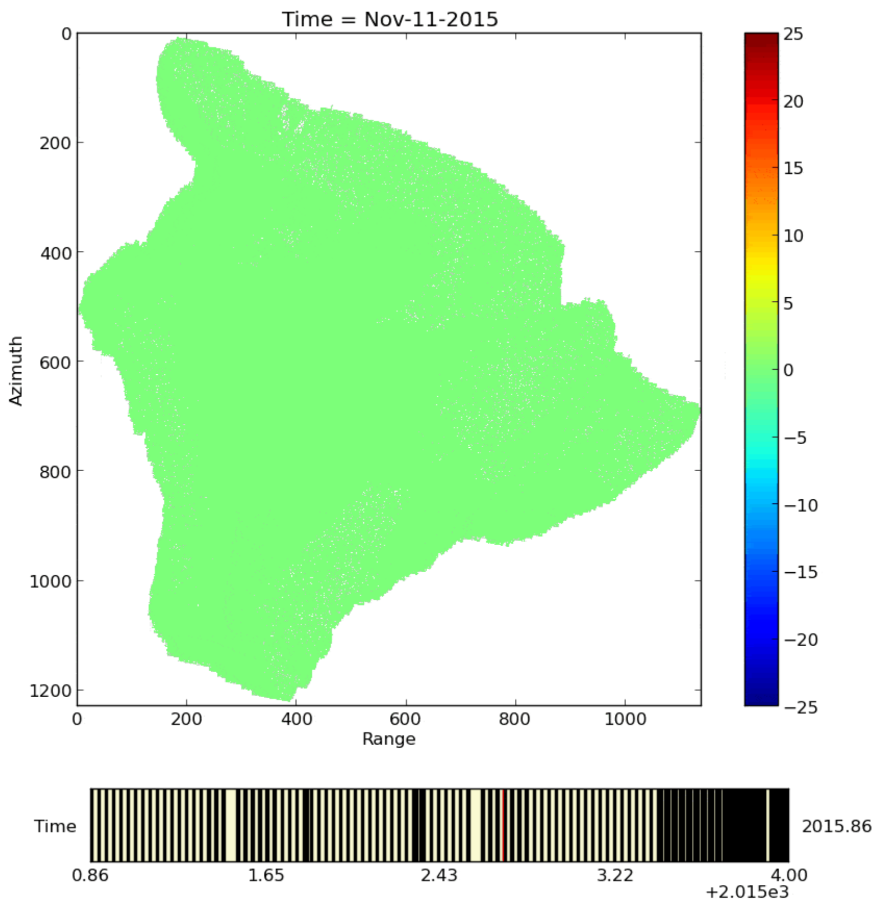
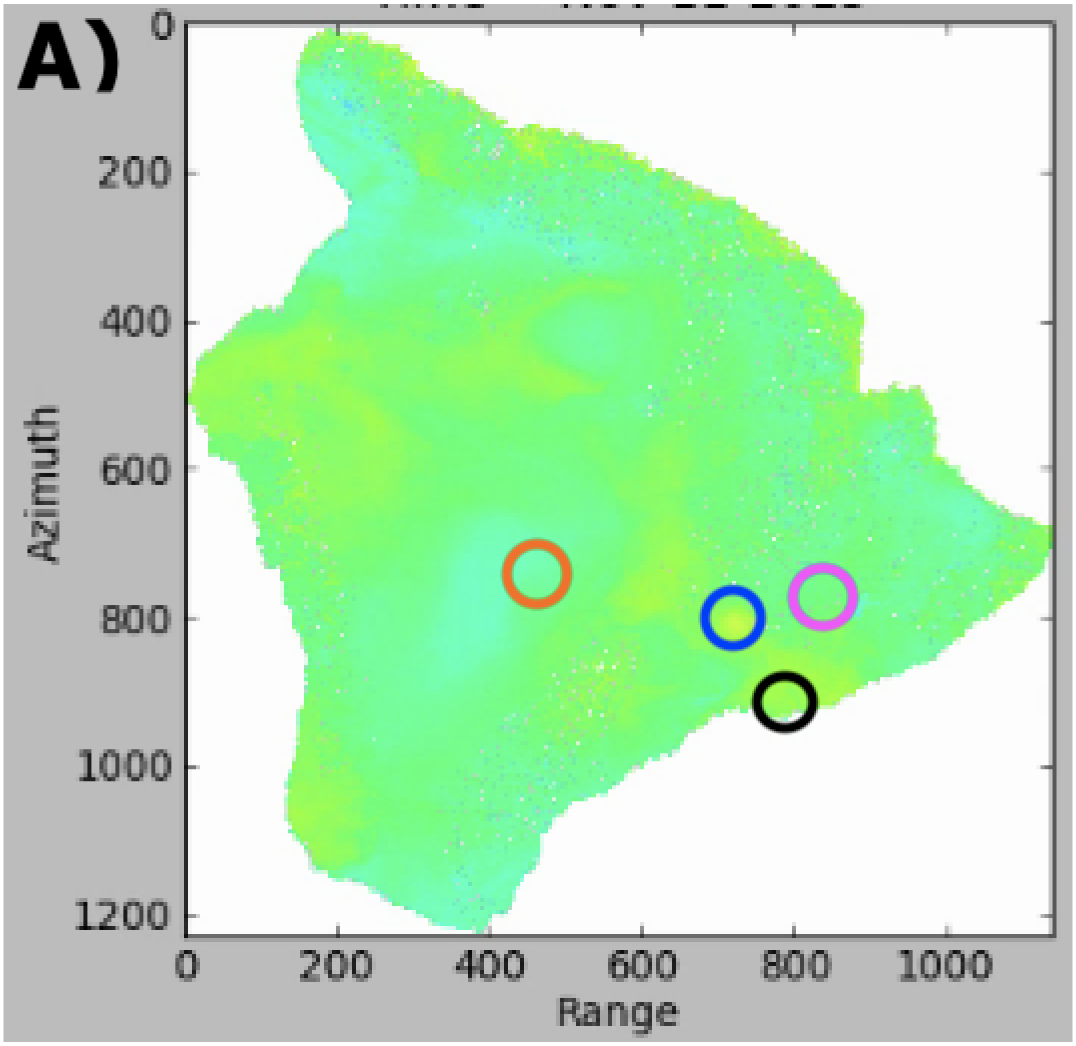
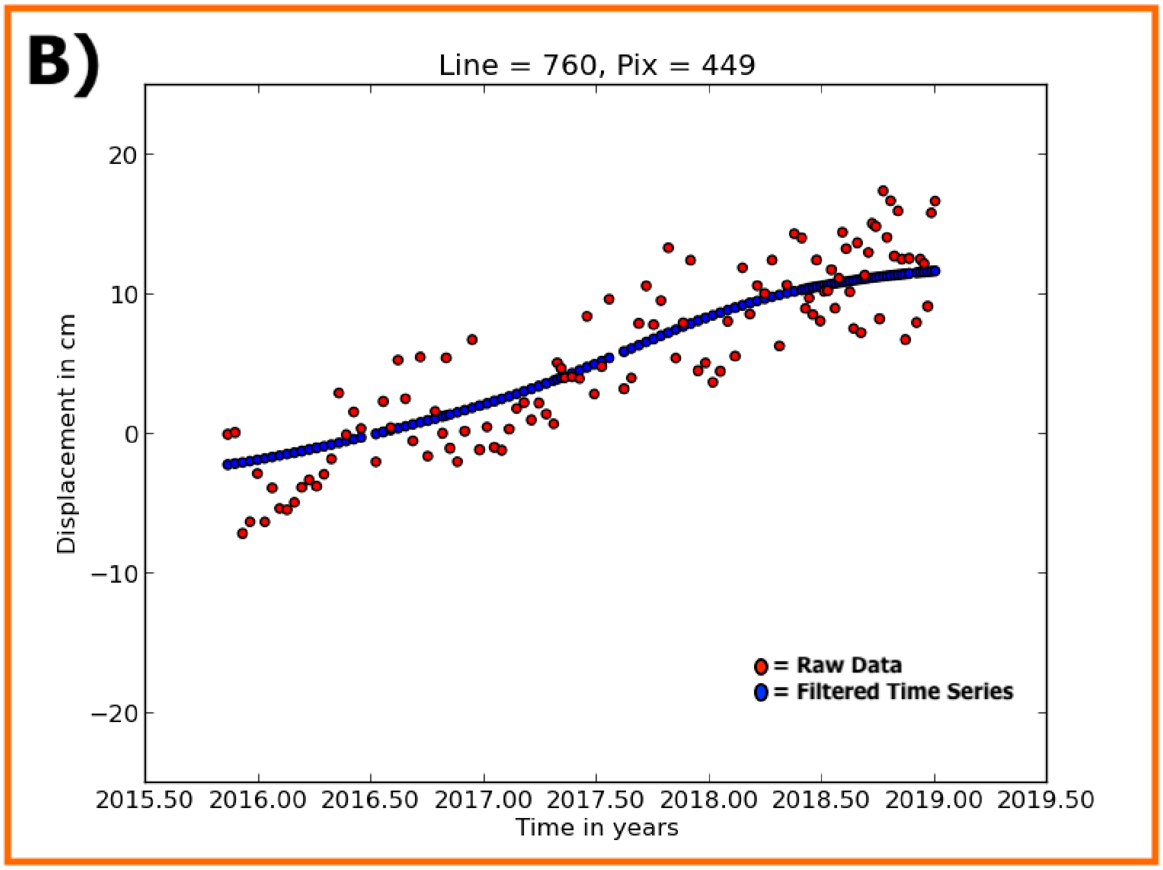
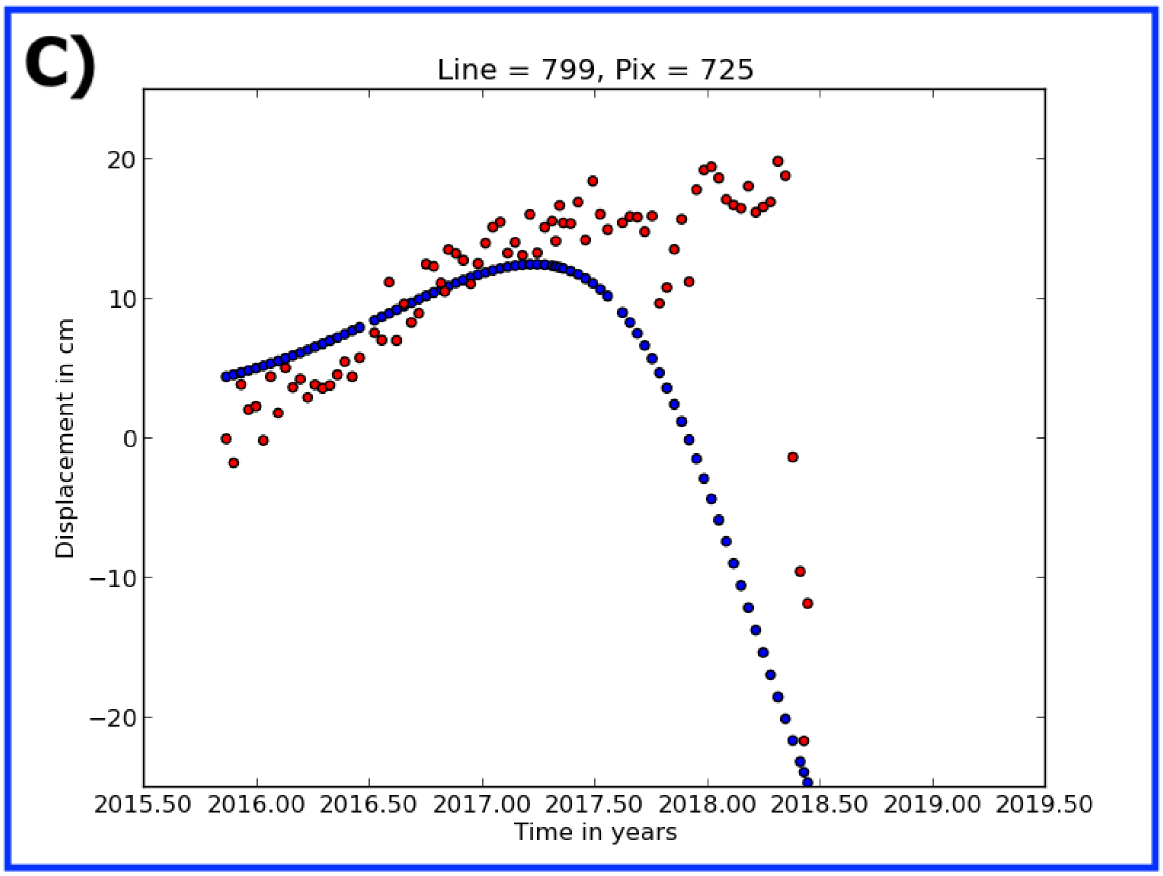
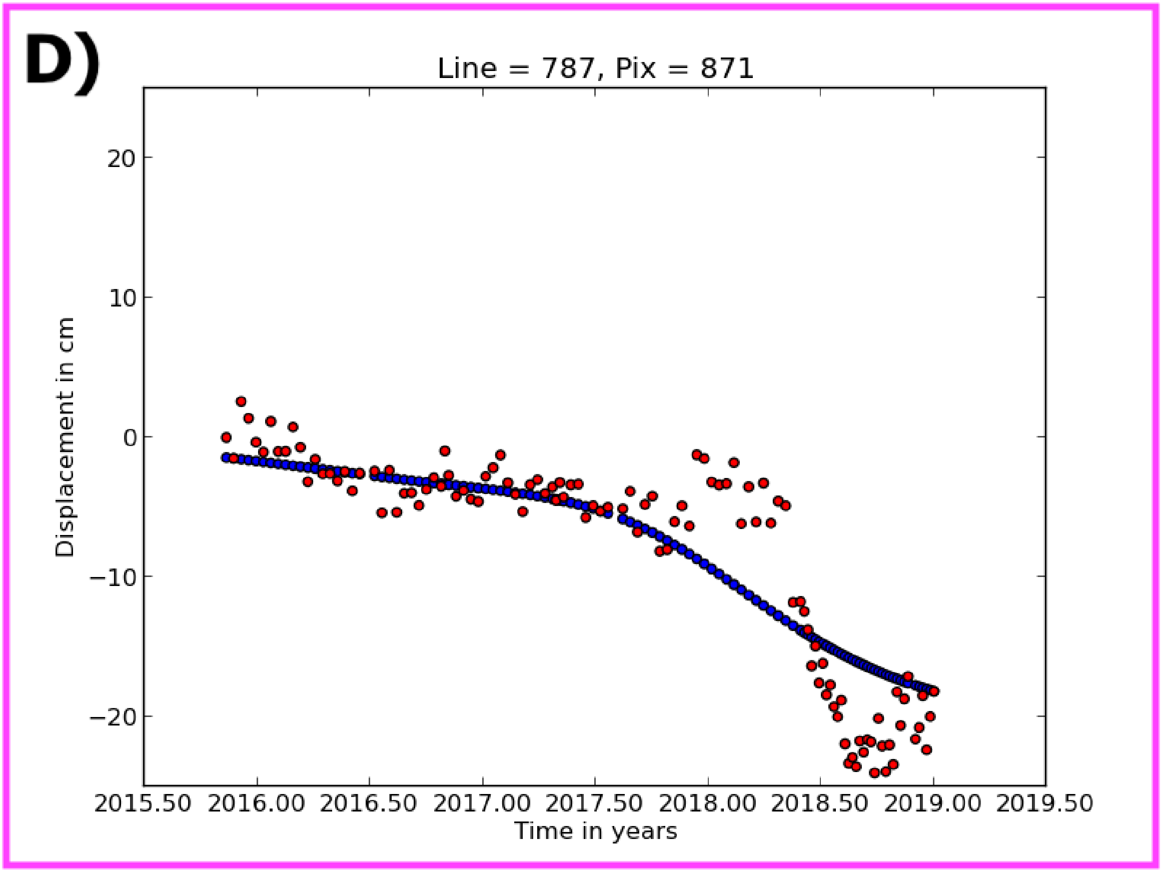


Figure 8: Continuous DInSAR time series results of path 87 frame 526 over 4 years from 2015 November- 2019 January using GMT5SAR and GIAnT processing. For these descending unwrapped displacement maps, red colors (positive) indicate motion toward the satellite or east, and blue colors (negative) indicate motion down or west. Color scale indicates line-of-sight displacement in cm.

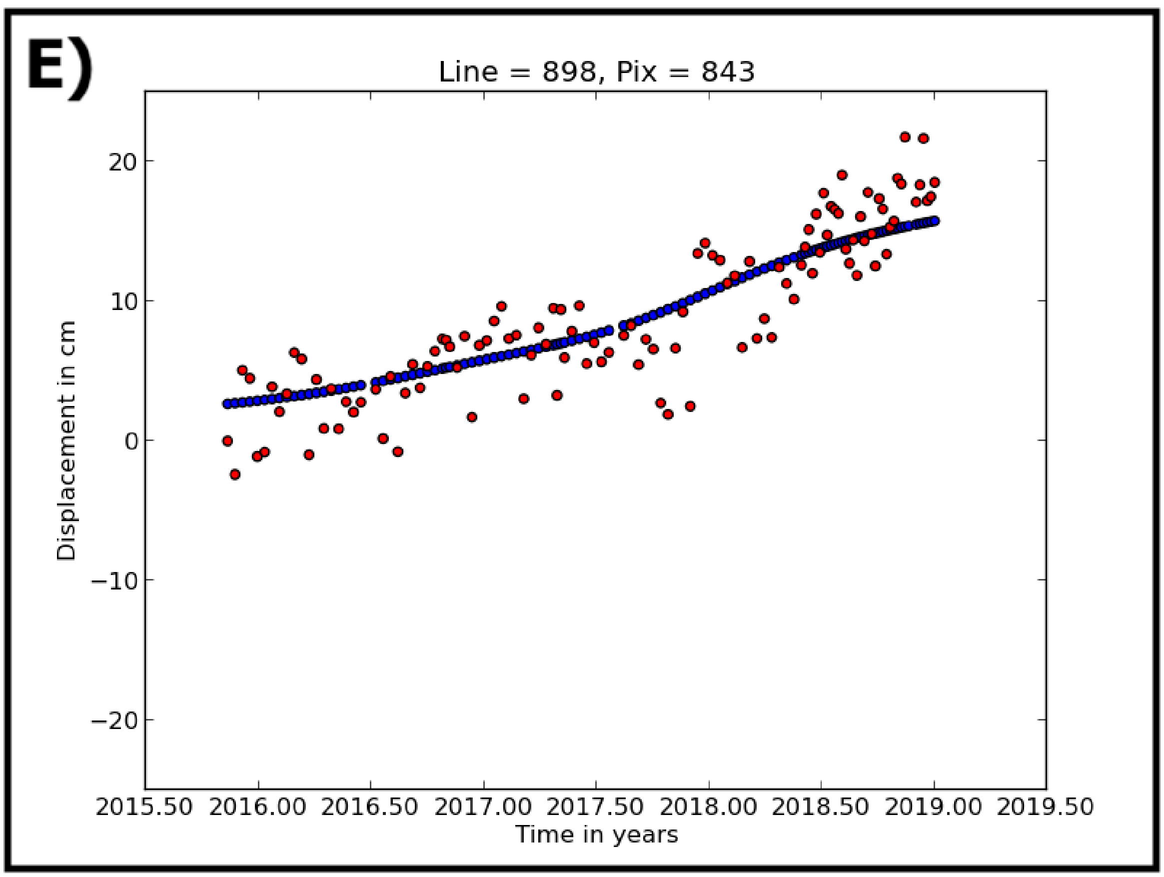

A)
Azimuth
Range
0
200
400
600
800
1000
1200


B)
Line = 760, Pix = 449
Displacement in cm
Time in years
= Raw Data
= Filtered Time Series
20
10
0
−10
−20
2015.50
2016.00
2016.50
2017.00
2017.50
2018.00
2018.50
2019.00
2019.50


C)
Line = 799, Pix = 725
Displacement in cm
Time in years
20
10
0
−10
−20
2015.50
2016.00
2016.50
2017.00
2017.50
2018.00
2018.50
2019.00
2019.50



D)
Line = 787, Pix = 871
Displacement in cm
20
10
0
−10
−20
2015.50
2016.00
2016.50
2017.00
2017.50
2018.00
2018.50
2019.00
2019.50
Time in years



E)
Line = 898, Pix = 843
Displacement in cm
20
10
0
−10
−20
2015.50
2016.00
2016.50
2017.00
2017.50
2018.00
2018.50
2019.00
2019.50
Time in years

Figure 9: Time series over A) regions that had the greatest displacement including B) Mauna Loa volcanic summit [outlined in orange], C) Kilauea volcano [outlined in blue], D) the Kilauea East Rift Zone [outlined in pink] and E) Pu'U'o'O volcanic flow field [outlined in black]. Graphs show raw displacement data in red and the filtered time series data in blue. B) Mauna Loa experiences a gradual uplift of 14 cm over the 4 year time period, likely due to the injection of magma beneath the island. The region begins to stabilize once the May 4th eruption begins. C) Over 19 cm of inflation occurs directly under Kilauea volcano from the build up of magma over November 2015 leading up to the eruption on May 4th, 2018. The sudden expulsion of that magma caused more than 40 cm of subsidence at the crater. D) After the eruption began, the East Rift Zone also sunk by more than 15 cm. E) A lot of lava surfaced or was deposited at the Pu'U'o'O Flow Field along the east edge of the island where large quantities entered into the ocean. The ground here was raised by over 12 cm.

THINGS TO CONSIDER...

Precise orbits are obtainable 2 weeks after image acquisition. If GeoSCIFramework continuously ingests new data, the most recent acquisitions will have the real-time orbital estimation applied until the precise orbit corrections become available. The images are then updated in the time series.

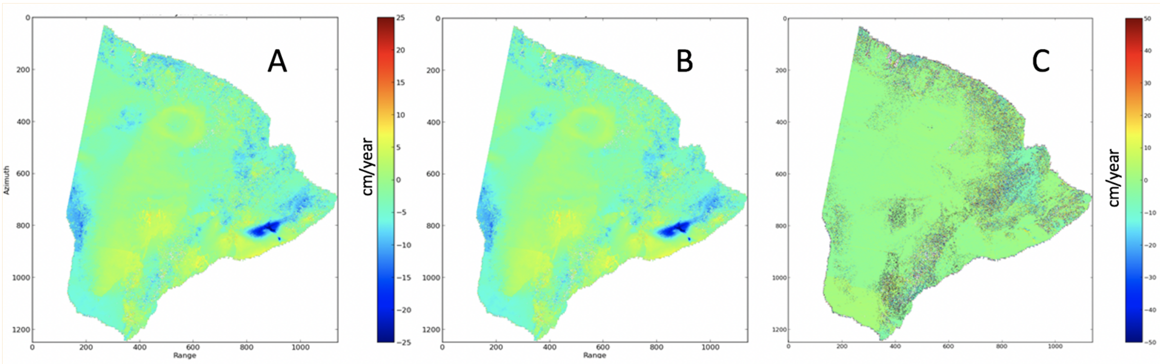


Figure 10: Impact of using precise and real-time orbits in time series over Hawaii from Oct. 2017 to June 2018 using Path A) Uses only precise orbits. B) 39 images using precise orbits and 6 using real-time C) The difference between A and B. (Note change in scale)

The atmosphere has a significant impact on SAR imagery. Water vapor, for example, causes delay in the phase signal, often distorting DInSAR products. To minimize error in our time series, we apply an **atmospheric calibration,** using the Generic Atmospheric Correction Online Service (GACOS) model.

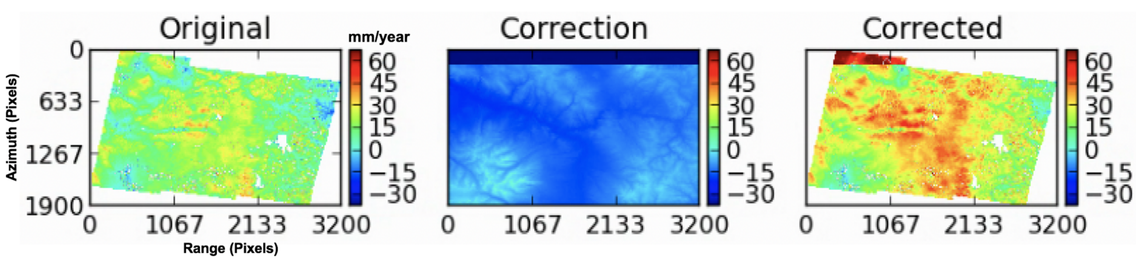


Figure 11: Atmospheric correction in SAR interferograms from GACOS over Yellowstone National Park. The unwrapped interferogram spans dates 09/02/17-10/20/17. This is a more extreme case that demonstrates how DInSAR results can be significantly improved when applying an atmospheric correction model.

Our team has also considered optimizing Digital Elevation Model (**DEM) resolution** and quality for InSAR processing by comparing results from a 10- and 30-m DEM. When using a higher-resolution DEM, InSAR processing can take exponentially more time to complete. Higher-resolution DEMs can also expose more details within the SAR image, producing more precise results. This trade-off is one worth considering for GSF.

FUTURE WORK

1) Finish comparing processing techniques and begin streaming data to the GSF.

2) Modeling of volcanic processes, specific to both Hawaii and Yellowstone

- Similar to the Independent Component Analysis (ICA) and machine learning analysis done for the 2018 eruption of Sierra Negra (Gaddes et al., 2019).

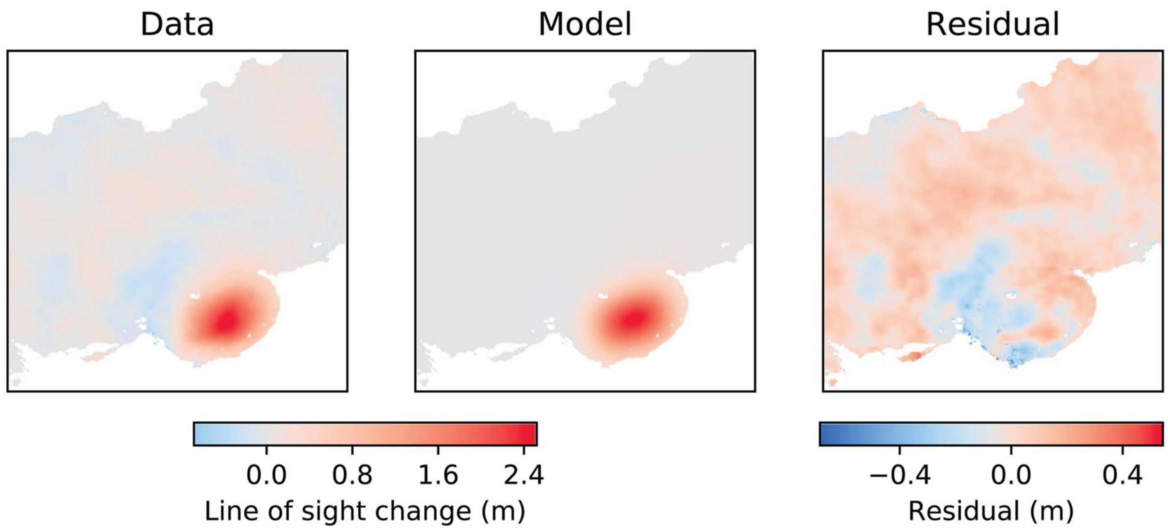


Figure 12: Figure and caption taken from Gaddes et al., 2019: Data: The signal contained in IC1 throughout the Sentinel−1 time series, showing ~2.4 m of motion toward the satellite during the Sentinel−1 time series. Model: The result of our optimal forward model, which treats the magma chamber as a 6.2×3.7−km2 rectangular dislocation at a depth of ~2.0 km. Residual: The misfit between our model and the data, which is dominated by a mottled pattern across the majority of the scene which independent component analysis is unable to remove from IC1 and our model is unable to fit.

3) Integrating additional data sets and time series into the streaming analysis.

- Ability to visualize time series, and compare signals from major geophysical events, with other data sets, such as seismic and GNSS time series.

4) Machine learning analysis of integrated time series.

Please contact **Brianna.Corsa@Colorado.edu** for questions or inquiries.