**Land Subsidence Data Documentation**

# Preface

The production of this data relies on an open-source analysis package developed by [Morishita et al. (2020)](https://doi.org/10.3390/rs12030424) called LiCSBAS. Their abstract provided an excellent overview of the analysis package:

“*For the past five years, the 2-satellite Sentinel-1 constellation has provided abundant and useful Synthetic Aperture Radar (SAR) data, which have the potential to reveal global ground surface deformation at high spatial and temporal resolutions. However, for most users, fully exploiting the large amount of associated data is challenging, especially over wide areas.*

*To help address this challenge, we have developed LiCSBAS, an open-source SAR interferometry (InSAR) time series analysis package that integrates with the automated Sentinel-1 InSAR processor (LiCSAR). LiCSBAS utilizes freely available LiCSAR products, and users can save processing time and disk space while obtaining the results of InSAR time series analysis.*

*In the LiCSBAS processing scheme, interferograms with many unwrapping errors are automatically identified by loop closure and removed. Reliable time series and velocities are derived with the aid of masking using several noise indices. The easy implementation of atmospheric corrections to reduce noise is achieved with the Generic Atmospheric Correction Online Service for InSAR (GACOS).*

*Using case studies in southern Tohoku and the Echigo Plain, Japan, we demonstrate that LiCSBAS applied to LiCSAR products can detect both large-scale (>100 km) and localized (~km) relative displacements with an accuracy of <1 cm/epoch and ~2 mm/yr. We detect displacements with different temporal characteristics, including linear, periodic, and episodic, in Niigata, Ojiya, and Sanjo City, respectively. LiCSBAS and LiCSAR products facilitate greater exploitation of globally available and abundant SAR datasets and enhance their applications for scientific research and societal benefit.*”

The Result and Discussion sections mirror those from [Harintaka et al. (2024)](https://doi.org/10.1080/10106049.2024.2364726), as they conducted a study similar to ours, suitable for comparison/validation purposes.

# LiCSBAS Workflow for Land Subsidence Analysis in DKI Jakarta

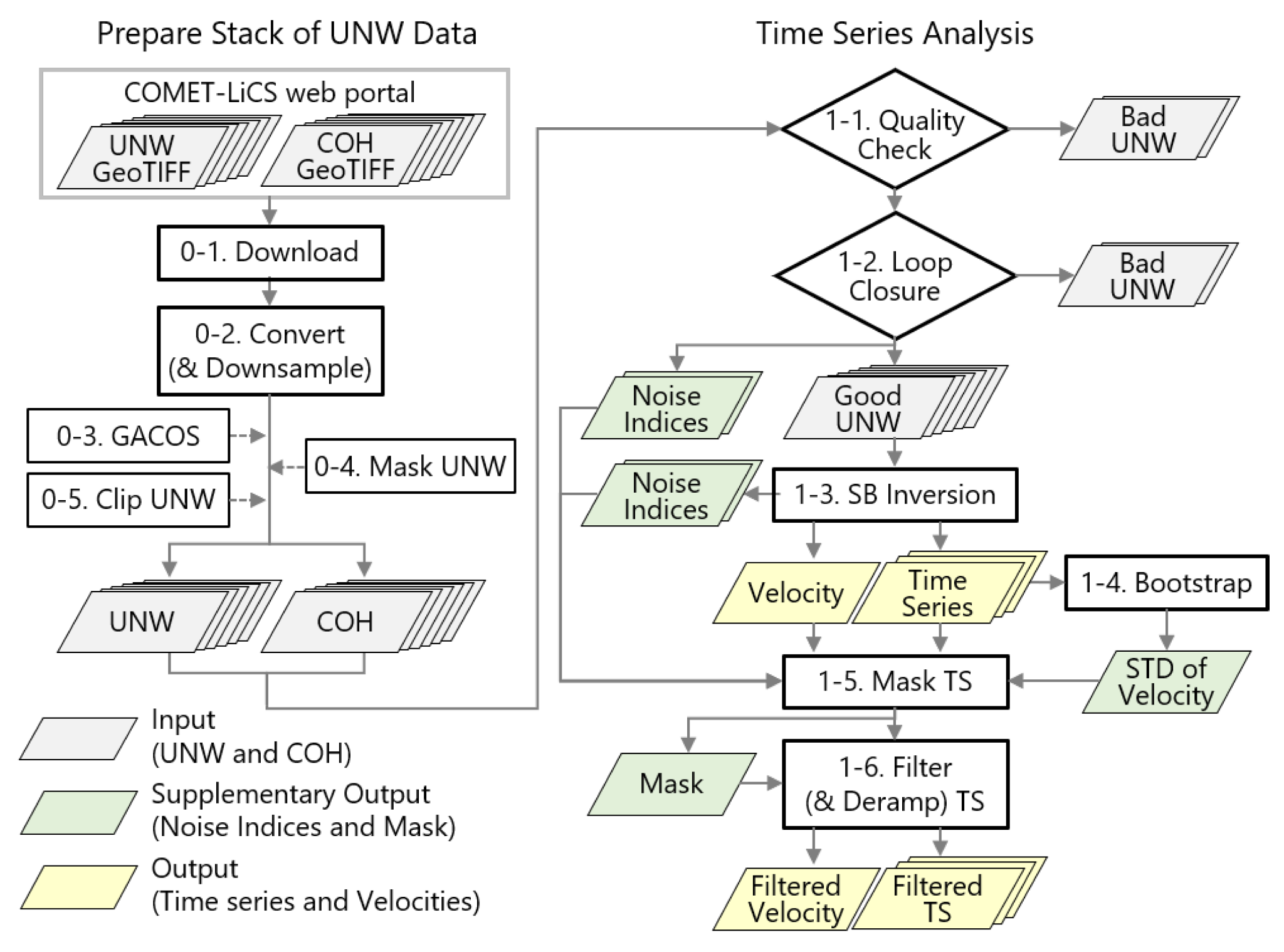


Figure 1. LiCSBAS Workflow

The LiCSBAS workflow is divided into two main steps: data preparation and time series analysis. It is designed to be customisable, allowing users to tailor this workflow to better suit their analysis in their respective study areas. The utilisation and customisation of the workflow are described as follows[[1]](#footnote-0):

## Data Preparation

### Download LiCSAR Products

This processing step automatically downloads GeoTIFF files of unwrapped interferograms (unw) and coherence (cc) from the COMET-LiCS web portal for a specified LiCSAR frame (typically 250 × 250 km), covering both ascending and descending tracks. These time-series datasets, usually available from late 2014 onward, are provided in a consistent geographic reference. The download also includes key metadata such as line-of-sight (LOS) vectors[[2]](#footnote-1) and perpendicular baselines.

As our research requires data from 2016 to 2024, 20160101 (January 1, 2016) and 20241231 (December 31, 2024) are selected as temporal boundaries for processing. The COMET-LiCS web portal shows only ascending data[[3]](#footnote-2) that covers the 2016 to 2024 requirement. Descending data, however, is only available from 2016 to 2023.

In this step, the GeoTIFF files of unw and cc of DKI Jakarta (Frame ID: 098A\_09673\_121312) are downloaded from the COMET-LiCS web portal.

### Convert GeoTIFF

The package converts downloaded unw and cc GeoTIFF files into float32 and uint8 formats, respectively, for efficient time-series processing. Optional downsampling (multilooking) via boxcar averaging reduces file size and speeds up processing while preserving sufficient detail when full ~100 m resolution is not necessary. Null output is returned if over half the pixels in a boxcar window are null.

The downloaded GeoTIFF files, unw and cc, are converted to float32 and uint8 format, respectively. The GeoTIFF files’ format is converted without downsampling since the spatial resolution of the data is already 100 m. Further downsampling will only reduce the accuracy since we want the final output of 30 m resolution (achieved through kriging interpolation using the LiCSBAS output).

### Tropospheric Noise Correction Using GACOS (Optional)

To reduce tropospheric noise in InSAR time series, LiCSBAS supports correction using GACOS delay maps, which are derived from European Centre for Medium-Range Weather Forecasts (ECMWF) weather data and available globally in near real-time. These corrections help improve the accuracy of displacement time series and average velocities, especially before spatiotemporal filtering. However, since tropospheric delays can sometimes correlate with deformation or vary over time, users should verify the correction’s effectiveness.

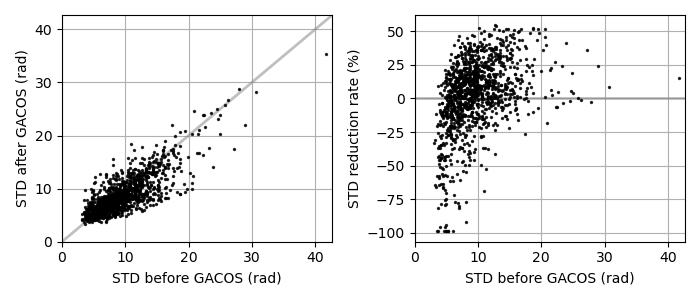


Figure 2. Correlation diagrams of the standard deviation (STD) of unwrapped phases in the 1,203 interferograms before and after the GACOS correction. Grey lines denote no change in STD.

GACOS correction is an optional step in this workflow. Thus, to determine its significance, we experimented with the workflow with and without the GACOS correction. The experiment revealed that GACOS correction does not yield a meaningful result, as the standard deviation (STD) of unwrapped phases for each entire interferogram increases (not decreases) by 2.6% on average and 5.9% on median, indicating minimal tropospheric noise mitigation[[4]](#footnote-3). The absolute difference of velocity between with and without GACOS correction is also minimal, with both average and STD of 0.2 mm. Therefore, this step is skipped in the final workflow to capture more data points in the final result.

### Mask Interferograms (Optional)

This optional step masks unwanted areas, such as isolated islands or low-coherence zones, in the unwrapped data stack to avoid regions with unwrapping errors. Masking helps improve the accuracy of network refinement in Step b.ii.

DKI Jakarta is mostly covered by built-up areas (e.g., buildings), which minimizes the unwrapping error ([Braun & Veci, 2021](https://step.esa.int/docs/tutorials/S1TBX%20TOPSAR%20Interferometry%20with%20Sentinel-1%20Tutorial_v2.pdf)). However, the northern area contains a relatively large vegetation area and borders the sea, resulting in many unwrapping errors ([Braun & Veci, 2021](https://step.esa.int/docs/tutorials/S1TBX%20TOPSAR%20Interferometry%20with%20Sentinel-1%20Tutorial_v2.pdf)). Despite having many errors, we cannot mask out the northern area since it is our area of interest. This step is then skipped in the final workflow.

### Clip Interferograms (Optional)

This step clips a defined rectangular area from the unw and cc data to focus the analysis on the area of interest. Clipping reduces file size and speeds up processing, while also potentially improving loop closure (Step b.ii.) by excluding irrelevant or error-prone areas.

The DKI Jakarta bounding box coordinates are used to clip the original image.

## Time Series Analysis

### Quality Check

In this processing step, LiCSBAS checks the quality of unwrapped interferograms by assessing average coherence and the percentage of valid pixels. Interferograms with severe decorrelation (e.g., from vegetation, snow, or missing bursts in Sentinel-1 SLC data) are flagged as bad data. These low-quality interferograms can negatively impact the time-series analysis and are excluded.

All parameters in this step are left to their default settings.

### Network Refinement by Loop Closure

LiCSBAS identifies unwrapped interferograms with significant unwrapping errors using loop closure analysis. If all loops involving a given interferogram show large root mean square (RMS) phase residuals, that interferogram is flagged as bad and removed from further processing. This conservative, interferogram-level filtering avoids false corrections that pixel-level methods might introduce. A preliminary reference point is then selected[[5]](#footnote-4).

All parameters in this step are left to their default settings.

### Small Baseline Network Inversion

LiCSBAS performs SB (Small Baseline) inversion using the new small baselines subset (NSBAS) method to estimate displacement time series and mean velocity from a network of unwrapped interferograms. The inversion solves a system of equations ()[[6]](#footnote-5) that links interferogram measurements to incremental displacements. To handle gaps in the network (due to decorrelation or unwrapping issues), a temporal constraint assuming linear displacement is added to improve realism. This constraint helps bridge disconnected network segments, though results over gaps must be interpreted cautiously. After inversion, a new stable reference point is selected as the pixel with the smallest RMS deviation and the fewest data gaps.

All parameters in this step are left in their default settings.

### Estimate the Standard Deviation of the Velocity by Bootstrap

LiCSBAS estimates the standard deviation (STD) of the mean velocity using the percentile bootstrap method. This involves resampling the displacement time series 100 times, recalculating velocities for each sample, and then deriving the STD from the ensemble. This helps assess the reliability of the velocity estimate, as a higher STD suggests noisy or nonlinear signals. Additionally, spatio-temporal consistency (STC) checks are performed to evaluate stability further.

All parameters in this step are left in their default settings.

### Mask Noisy Pixels in the Time Series

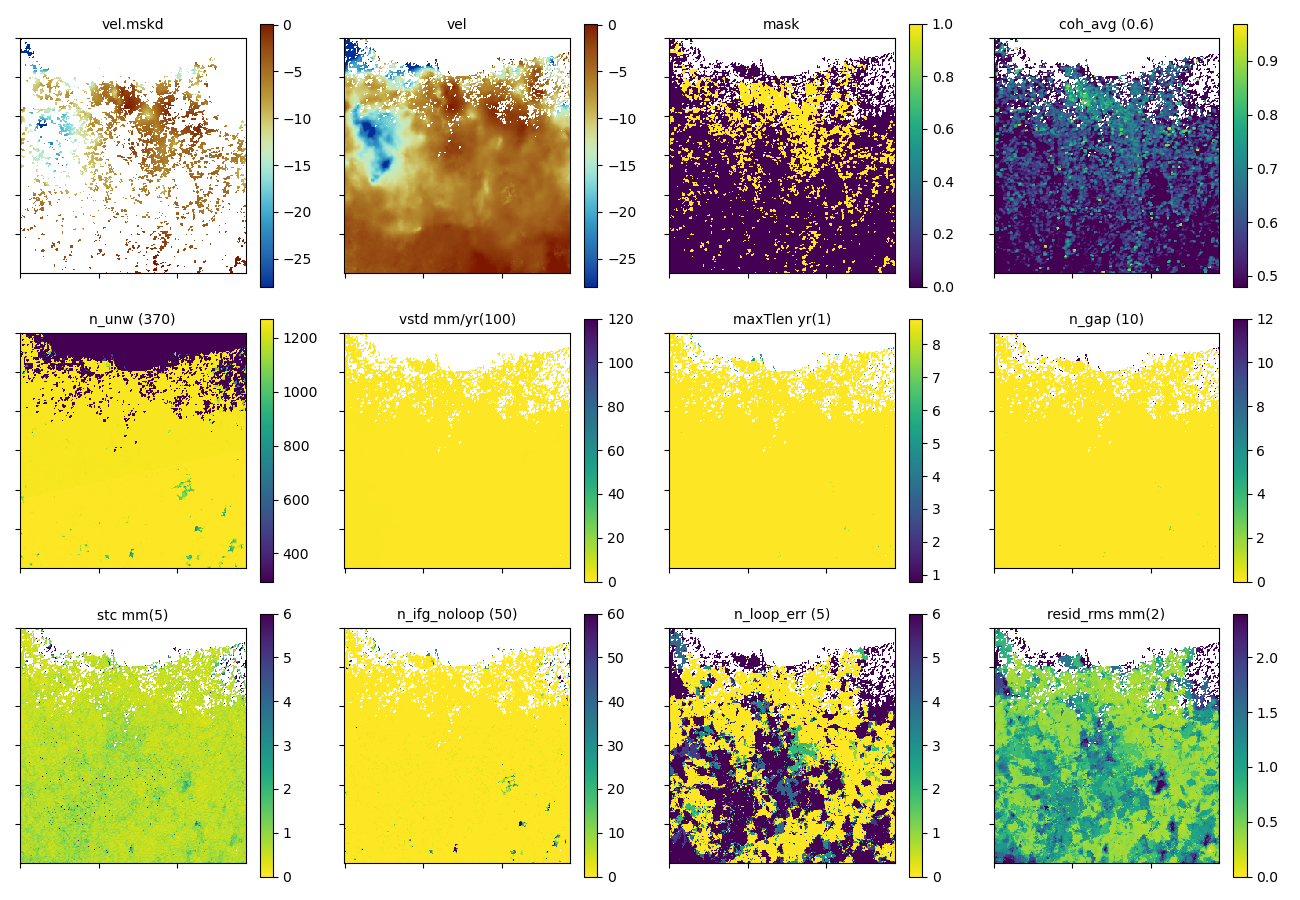


Figure 3. Noise indices for the full image

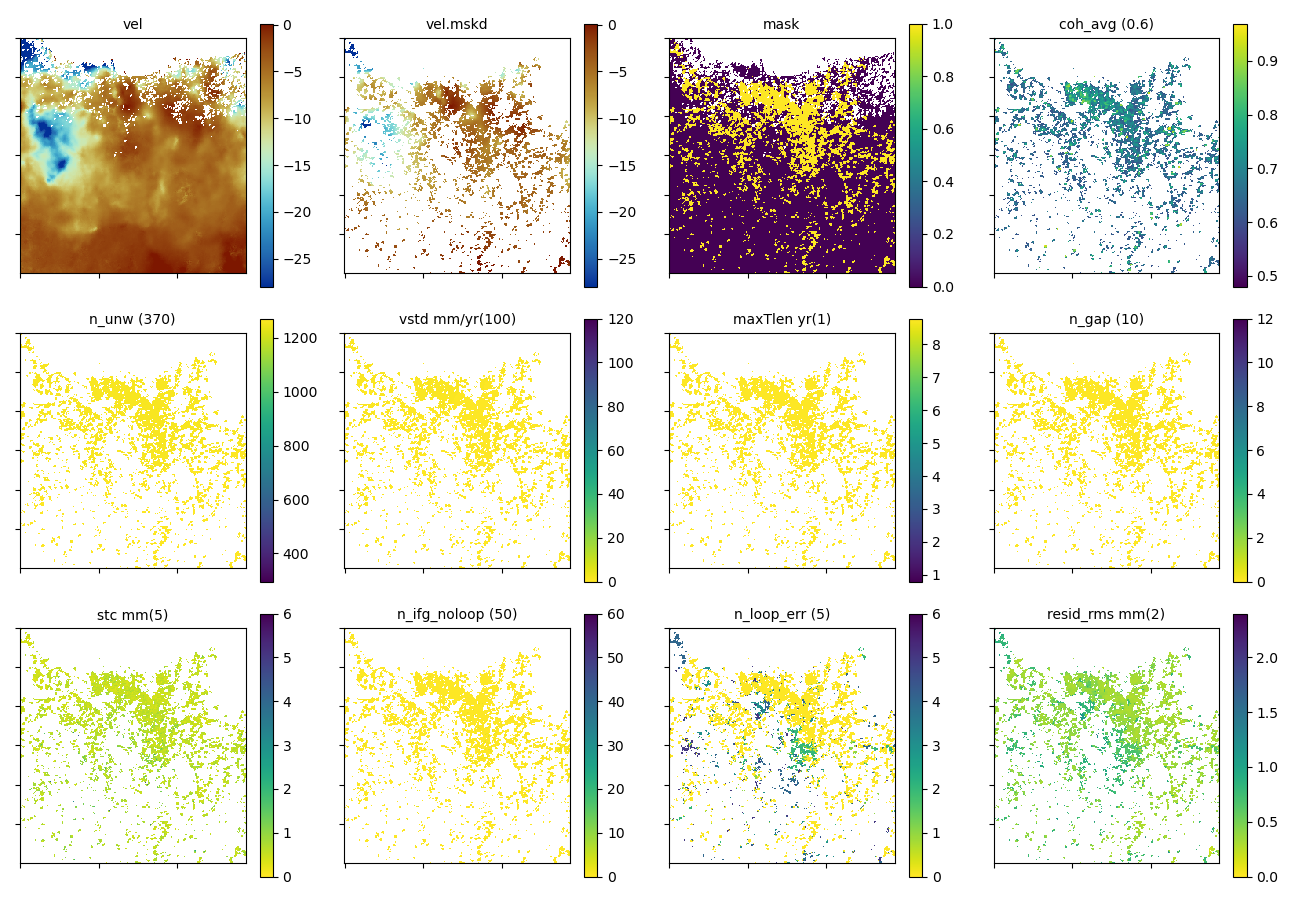


Figure 4. Noise indices for the clipped image

To ensure that only reliable displacement time series are used, LiCSBAS masks "bad" pixels based on several noise indices (e.g., coherence, velocity STD, network gaps, STC, etc.). If any index exceeds a defined threshold, the pixel is excluded. Thresholds can be manually adjusted for stricter or looser masking. This process helps filter out unreliable pixels caused by unwrapping errors or network gaps, improving the quality of the final velocity and time series outputs.

In this processing step, noise indices are observed to determine the optimum threshold values. Coherence is a common metric to consider when determining pixel quality in InSAR processing, but different studies use different thresholds. The European Space Agency (ESA), producer of the Sentinel-1 SAR data used by LiCSBAS, mentions that a good coherence is >0.6 ([Braun & Veci, 2021](https://step.esa.int/docs/tutorials/S1TBX%20TOPSAR%20Interferometry%20with%20Sentinel-1%20Tutorial_v2.pdf)). This is in agreement with our resulting noise indices, as Figure 3 shows that areas with average coherence of <0.6 contain pixels with higher unclosed loops after network refinement (*n\_loop\_err*) and RMS of residuals in the SB inversion (*resid\_rms\_mm*). Other parameters, such as velocity STD, network gaps, etc., show only a small number of bad-quality pixels, which are sufficiently handled by their respective default threshold values, as shown in Figure 4. Thus, the coherence threshold was set to 0.6 while others are left to their default settings to ensure high-quality pixels as output.

### Spatiotemporal Filtering of Time Series

LiCSBAS applies a spatio-temporal filter (high-pass in time, low-pass in space using Gaussian kernels) to reduce spatially correlated, temporally uncorrelated noise (e.g., tropospheric, ionospheric, orbital). Optionally, deramping (linear, bilinear, or quadratic) and elevation-based corrections can also be applied. The filter helps preserve short-term or nonlinear displacements (e.g., seasonal signals), especially with high-frequency data like Sentinel-1. It is also worth noting that overly strong filtering may suppress real signals. After filtering, a stable reference point is re-evaluated.

All parameters in this step are left in their default settings.

# Result

## Ground Deformation in DKI Jakarta

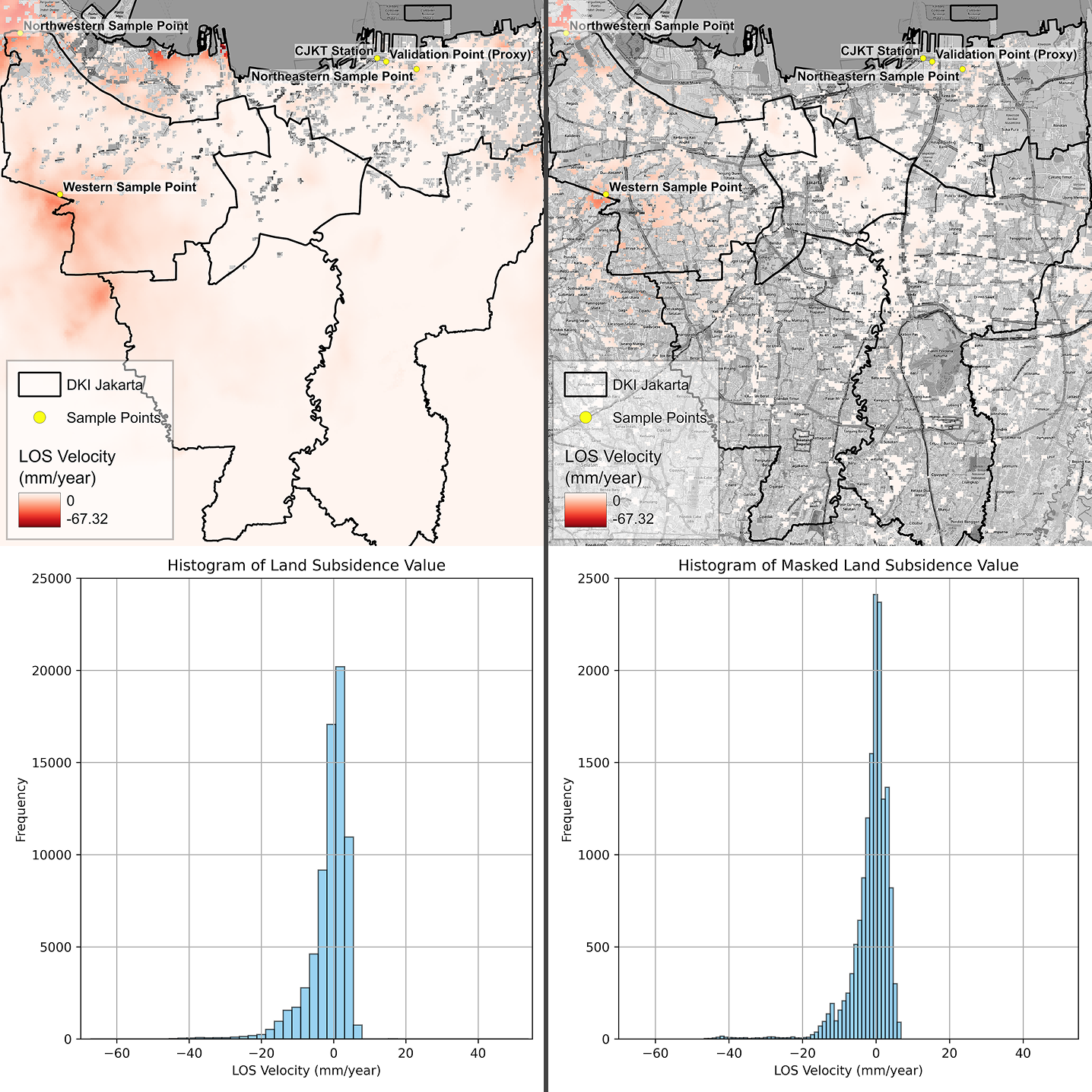


Figure 5. Land subsidence (left) and masked land subsidence (right) in DKI Jakarta and its histogram. Positive values are capped at 0 in both maps

The processing pipeline of LiCSBAS revealed the LOS velocity of ground deformation in DKI Jakarta with a spatial resolution of ~100 m. The deformation is concentrated in three main locations in DKI Jakarta: western, northwestern, and northeastern regions. The distribution of the LOS velocity[[7]](#footnote-6) and its histogram are shown in Figure 5, with the left figure being the land subsidence value before applying the quality mask, while the right figure is after masking. The minimum value (maximum subsidence) in DKI Jakarta is -29.86 mm/year (masked). The western and northwestern hotspots are situated in a densely populated area and a coastal area (warehouses and construction area of Pantai Indah Kapuk (PIK) 2), respectively, along the border with Tangerang City, Banten Province. The northeastern region, on the other hand, is located near the Port of Tanjung Priok.



Figure 6. Subsidence trend for western and northwestern hotspot regions

The subsidence from the three hotspots indicates two different trends, as shown in Figure 6. The western region has been experiencing rapid subsidence since 2016 and has slowed down significantly since 2019. Furthermore, the northeastern and northwestern regions are experiencing linear deformation with no significant signs of slowing down, albeit with drastically different velocities.

## Validation with Ground Measurements and Other Studies

Despite InSAR offering a larger extent for land subsidence analysis, the result should always be validated with ground measurements. One such ground measurement station (named CJKT station) in DKI Jakarta is maintained by *Badan Informasi Geospasial* (Geospatial Information Agency) of Indonesia as part of a network of continuously operating reference stations (CORS) called InaCORS, providing ground deformation measurements using global navigation satellite system (GNSS). The CJKT station is located in Tanjung Priok, but the LiCSBAS output does not have any data points in the station’s coordinates. Therefore, despite not directly intersecting the CORS station, we utilised the closest data point to the station (~561 m apart) as a proxy (shown in Figure 5). [Susilo et al. (2023)](https://doi.org/10.1038/s41597-023-02274-0) provided the station’s measurement data from 2011 to 2021. To match both the LiCSBAS’s and the station’s measurements, we only used data from 2016 to 2021 for this validation. Additionally, we removed some data points from the station’s measurements with vertical displacement uncertainty within the top 20th percentile and only used this cleaned data for our analysis unless stated otherwise.

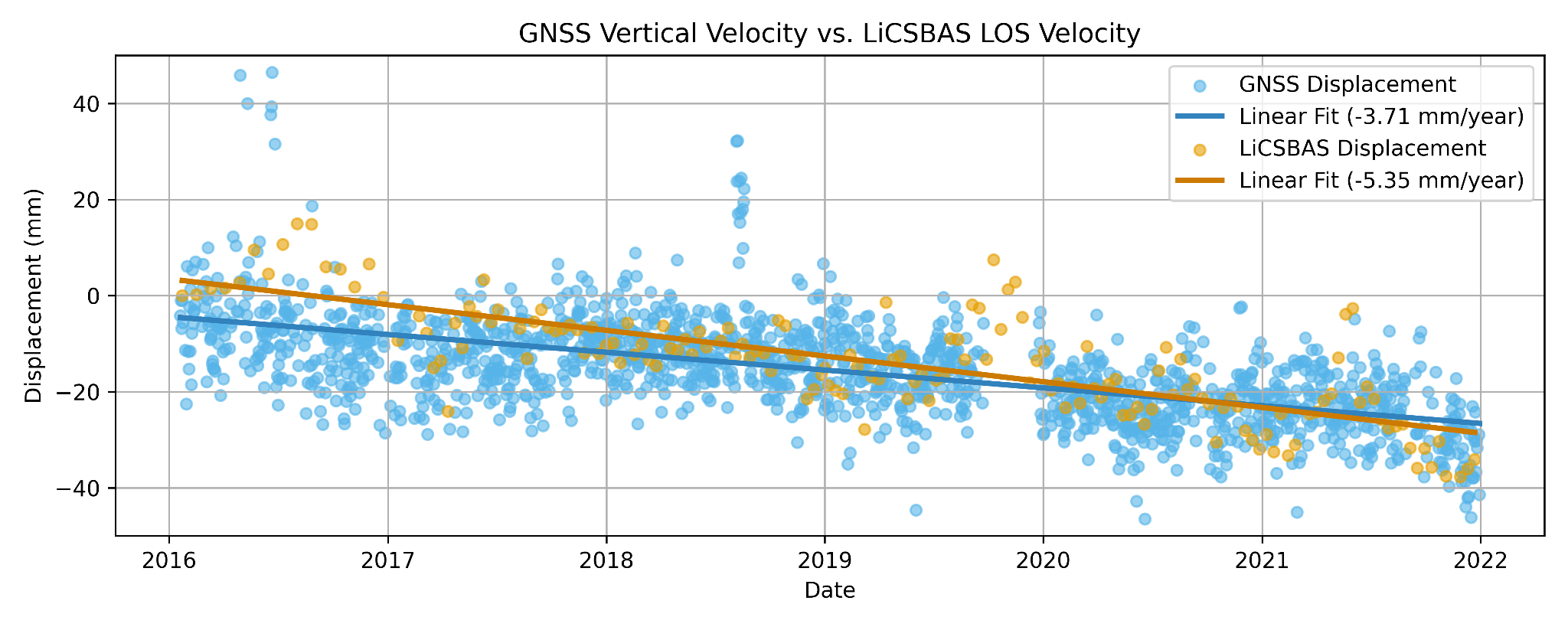


Figure 7. Linear fit of GNSS and LiCSBAS measurements

Both the LiCSBAS’s and the station’s measurements show a matching linear subsidence. To calculate the velocity from both datasets, linear regression is utilised to find the slope of the best fitting line (i.e., annual velocity) from LiCSBAS and CORS measurements. Our LOS velocity overestimates the vertical velocity from GNSS measurements with an absolute difference of 1.64 mm/year, as shown in Figure 7. To calculate the precision and accuracy of the LiCSBAS result, Harintaka et al. (2024) suggested the use of STD and RMSE. Compared against CJKT measurements, the difference between the standard deviation of both measurements is 0.78 mm with a root-mean-squared error (RMSE)[[8]](#footnote-7) of 9.40 mm.

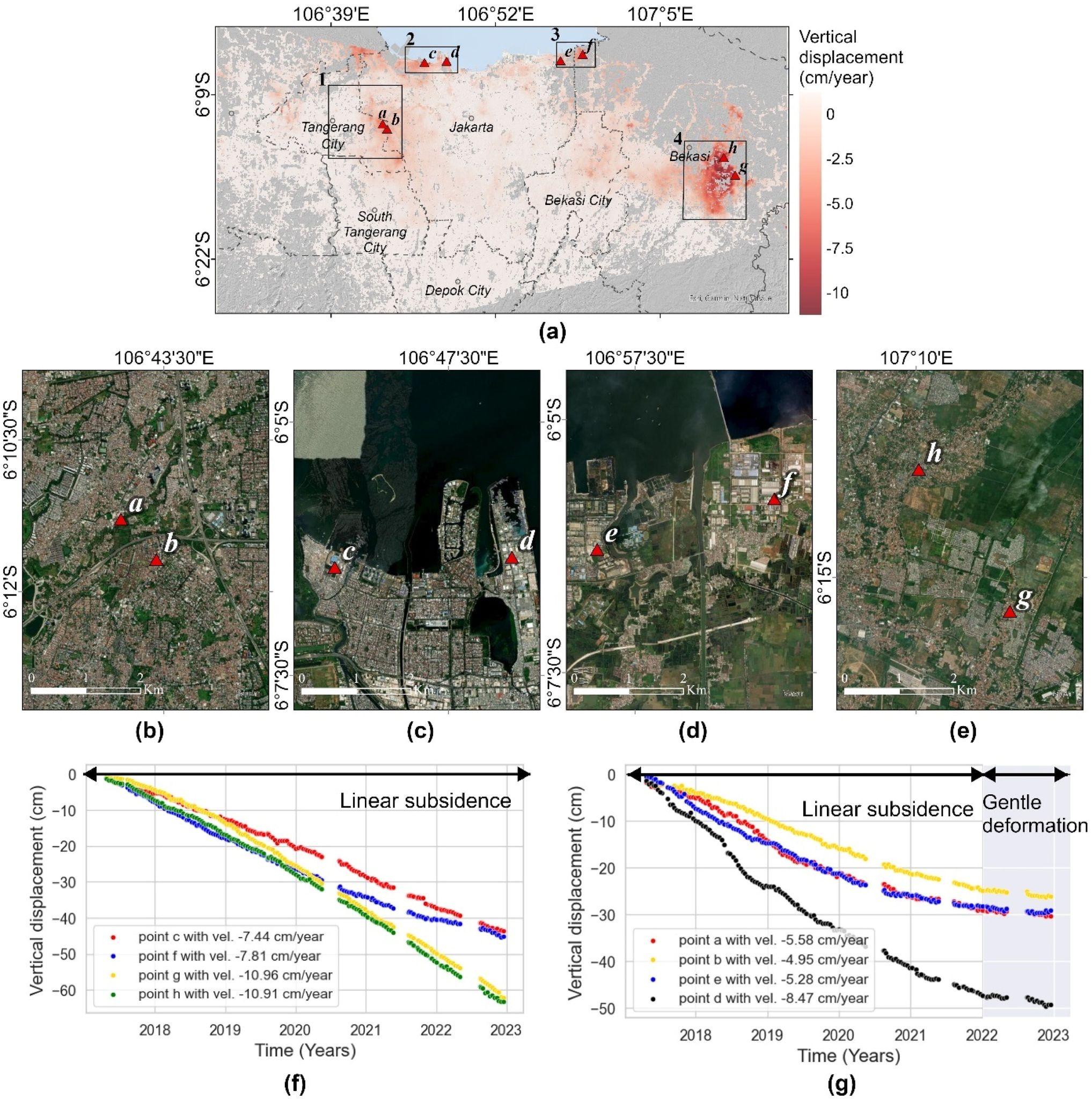


Figure 10. Subsidence trends from Harintaka et al. (2024)

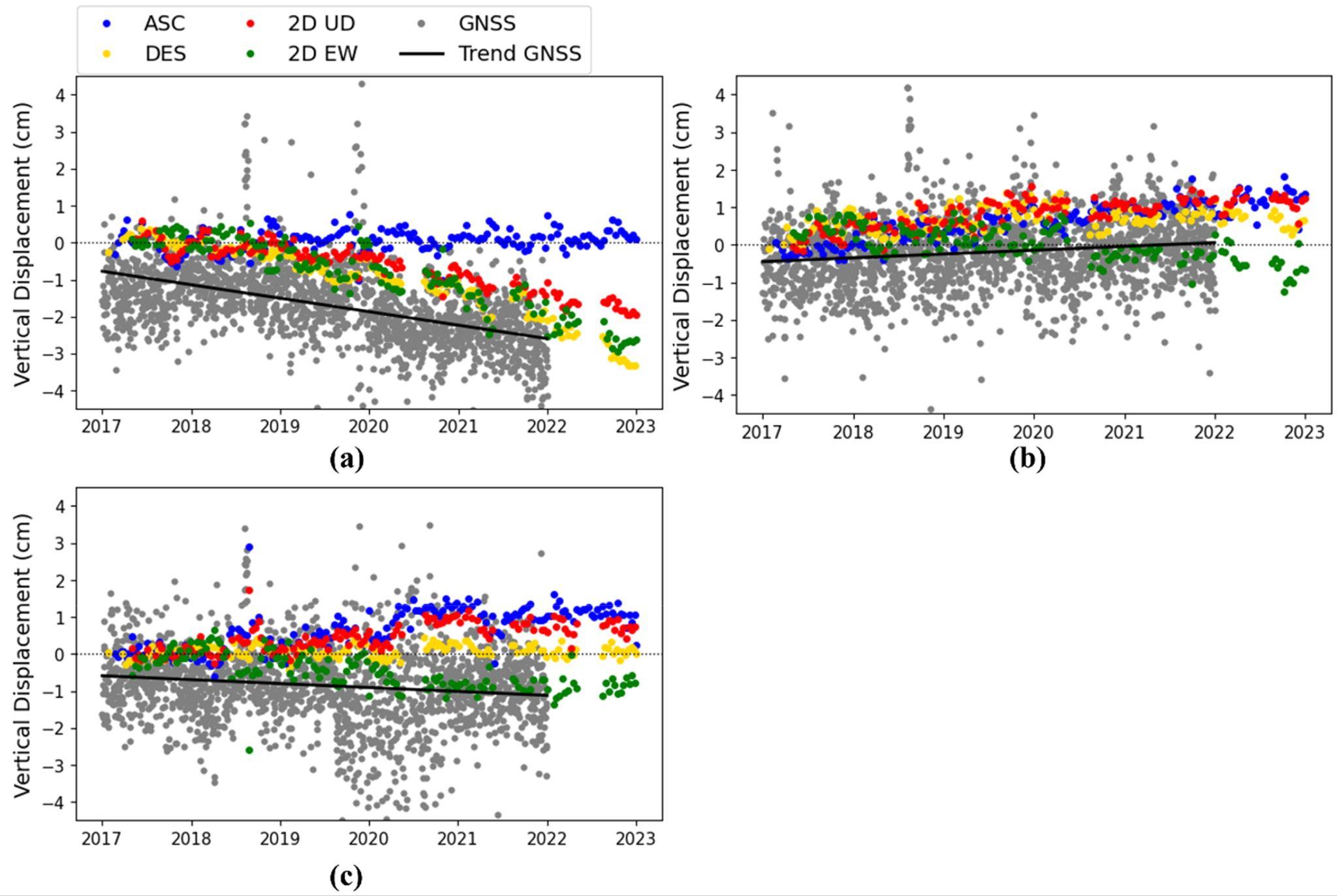


Figure 11. InSAR validation with GNSS measurements in CJKT (a), CBTU (b), and CTGR (c) stations from Harintaka et al. (2024)

Harintaka et al. (2024) investigated the land subsidence phenomenon in the Greater Jakarta Area using InSAR to determine vertical velocity with data spanning from 2017 to 2022 (results shown in Figure 10). Compared to our analysis, the subsidence hotspots in DKI Jakarta are located within the same three regions, with the addition of an extra hotspot in the north. However, the northern hotspot is not included in our final result due to pixels there not passing our quality filter, as shown in Figure 5, and not because it is unnoticed.

The authors also analyzed the subsidence trends, with one region (Figure 10; point a and b) where we also have the data (i.e., the western hotspot). The comparison between the subsidence trend from our study with Harintaka et al. (2024) shows that the trend from the western region is matching, showing a slowing subsidence. Furthermore, when compared to the CJKT station measurements, the LOS displacement trend of the ascending data from Harintaka et al. (2024) shows a disagreement (shown in Figure 11 (a; blue points)), while ours shows an agreement with both CJKT and their vertical displacements.

# Discussion

## Reliability of LiCSBAS Measurement

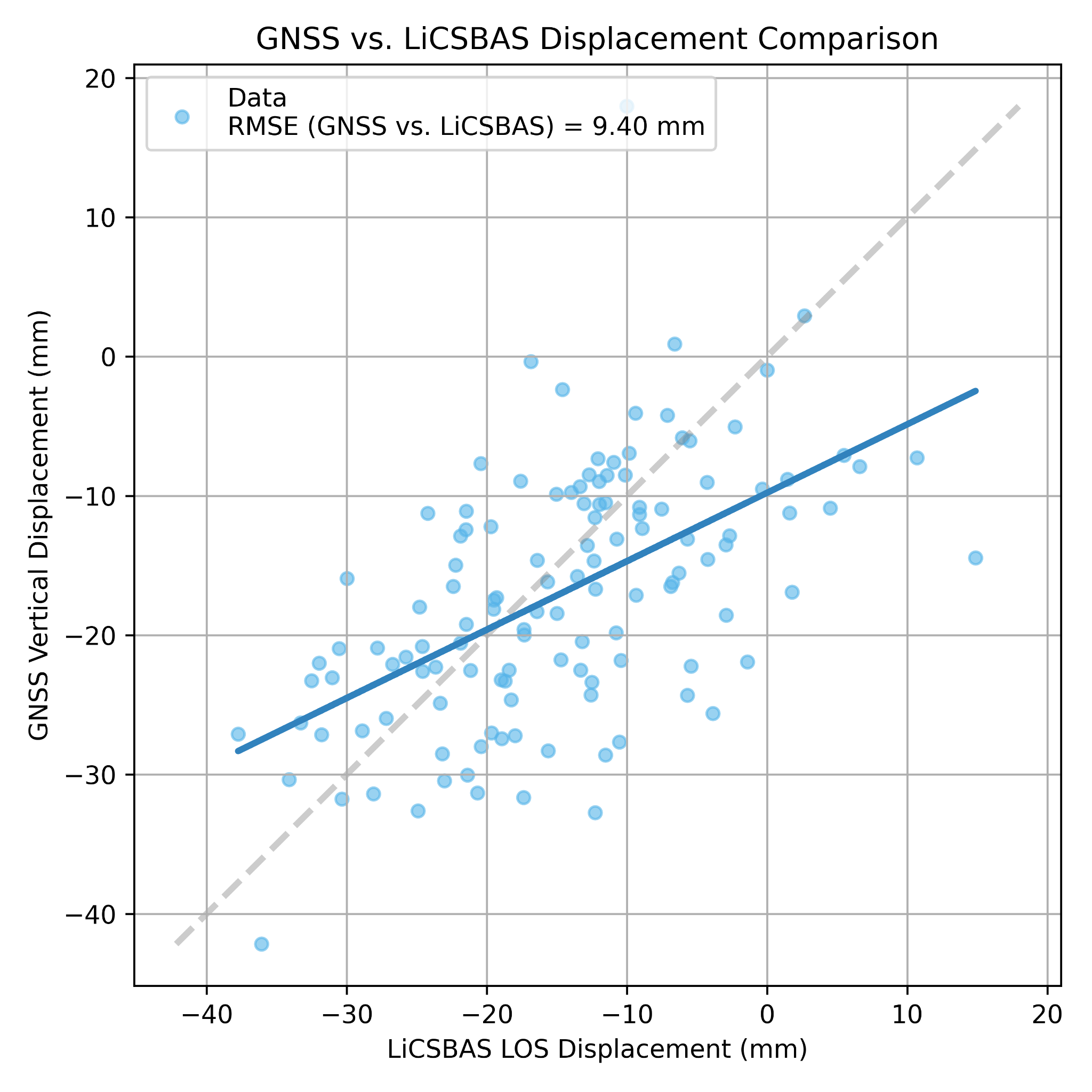


Figure 13. Per epoch comparison of GNSS vertical displacement and LiCSBAS LOS displacement

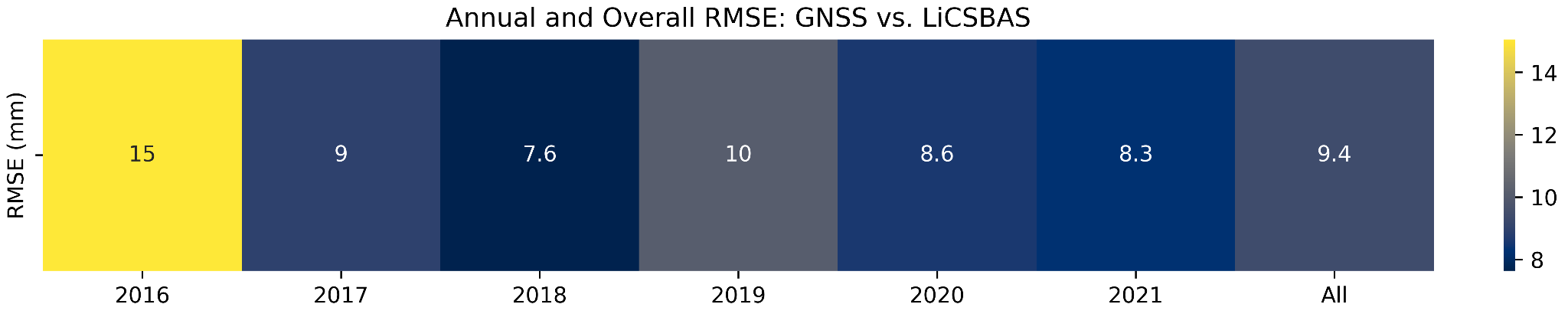


Figure 14. Annual and overall RMSE of LiCSBAS measurements against GNSS measurements

Since we do not have adequate descending data, the result of the processing pipeline cannot be decomposed from LOS displacement/velocity to vertical displacement/velocity. Without decomposition, the displacement value would not be as accurate (Harintaka et al., 2024). In a study of land subsidence with InSAR in Jakarta, [Ng et al. (2012)](https://doi.org/10.1016/j.jag.2012.01.018) found that a horizontal movement to the east of 10 mm/year could cause an error of 8 mm/year in vertical direction. Both values are technically incomparable since vertical displacement/velocity is more accurate than its counterpart because it is free from horizontal movement. In addition, the validation pixel we used is distanced from the CJKT station which means it does not capture displacement exactly at the area measured by the CJKT station.

Despite the odds against the LOS results, the precision and accuracy are surprisingly good when compared to both GNSS measurements and another similar study, as elaborated in Section C.b. When compared to GNSS measurements epoch-to-epoch, the LiCSBAS output appears to underestimate the vertical displacement value (Figure 13), despite the resulting LOS velocity overestimating the vertical velocity (Figure 7). Compared against GNSS measurements, the precision and accuracy of our results, measured by STD and RMSE, has an overall difference in STD of <1 mm and an overall RMSE of <1 cm. The RMSE per year is shown in Figure 14. The source of error varies, but it could be attributed to horizontal displacement, atmospheric condition, and orbital errors (Harintaka et al., 2024; Morishita et al., 2020). The subsidence trend is also matching when compared to another similar study on a region where we also have data points from LiCSBAS. This goes to show that the LiCSBAS analysis package handled the processing really well, even when only producing LOS displacement/velocity.

## Reliability of GNSS Measurement

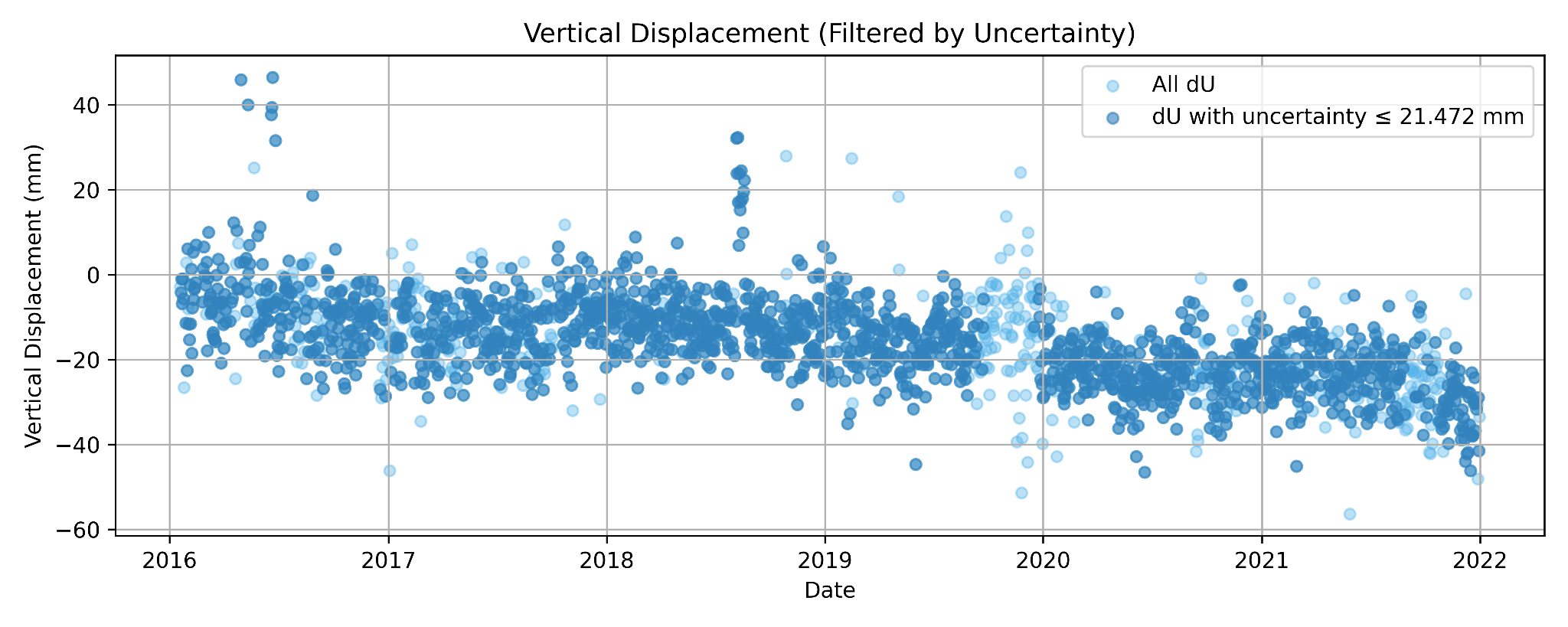


Figure 15. Unfiltered and filtered GNSS measurements

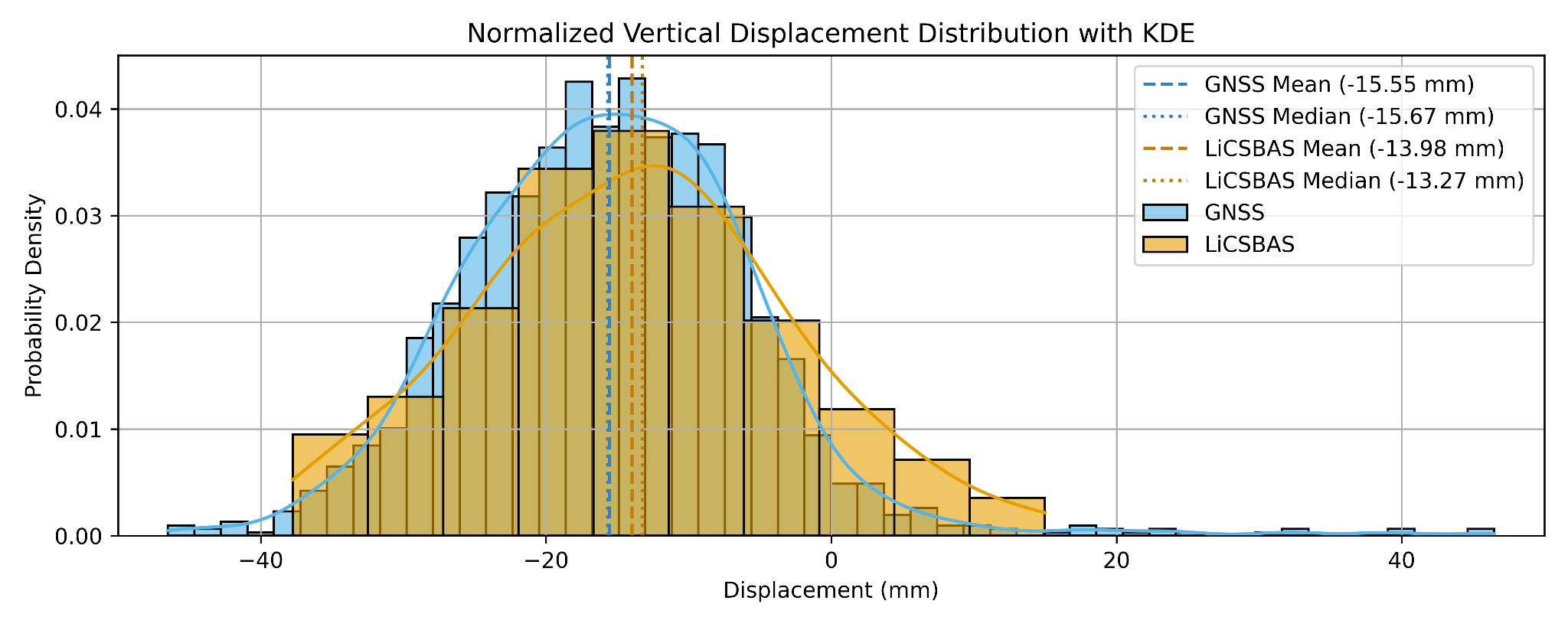


Figure 16. Histogram of GNSS and LiCSBAS measurements

Even as a ground measurement station, data provided by the CJKT station is not without errors. For starters, the STD of vertical displacement high, reaching 10.23 mm. This could be attributed to the location of the station that is surrounded with tall buildings, contributing to obstruction and signal multipath[[9]](#footnote-8) (Harintaka et al., 2024). Before clearing out the data within the top 20% vertical uncertainty, the value reaches an average of 18.92 mm and STD of 6.69 mm (Susilo et al., 2023). After clearing it up, the average and STD of vertical uncertainty significantly reduced to 16.76 mm and 2.86 mm, respectively (filtered and unfiltered data shown in Figure 15). We can also see that a lot of measurements in the late 2019 has some of the highest vertical uncertainty. Nevertheless, the LOS displacement and vertical displacement shows an overall agreement, as shown in Figure 16.

1. More details regarding each processing step can be found in [Morishita et al. (2020)](https://doi.org/10.3390/rs12030424) and their respective codes in [Morishita (2023/2025)](https://github.com/yumorishita/LiCSBAS2) [↑](#footnote-ref-0)
2. More on LOS vectors: <https://satelliteblog.cgg.com/insar-line-of-sight-explained/>, <https://site.tre-altamira.com/insar/>. [↑](#footnote-ref-1)
3. Ascending and descending data reflect the orbit of the satellite when capturing the SAR data, whether it is orbiting from the north pole to the south pole or vice versa. Without adequate descending data, we cannot calculate vertical displacement/velocity and can only get LOS displacement/velocity. [↑](#footnote-ref-2)
4. The STD decreased from 9.6 rad to 8.9 rad on average and from 8.9 rad to 8.1 rad on median by the GACOS correction [↑](#footnote-ref-3)
5. This is the pixel with valid, error-free unwrapped data across all interferograms and the lowest loop phase RMS, used for the next inversion step. [↑](#footnote-ref-4)
6. Suppose that we have a stack of *M*-unwrapped interferograms 𝐝 = [𝑑1,…, 𝑑𝑀]T produced from *N* images acquired at (𝑡0, …,𝑡𝑁−1), *N*-1 incremental displacement vector 𝐦 = [𝑚1,…,𝑚𝑁−1]T (i.e., *mi* is the incremental displacement between time *ti*-1 and ti) can be derived by solving ***d* = *Gm*** where ***G*** is a *M*×(*N*-1) design matrix of zeros and ones describing the relationship between the network of the interferograms and incremental displacements, considering that the unwrapped interferogram (i.e., displacement between two acquisitions) is the sum of corresponding successive incremental displacements. [↑](#footnote-ref-5)
7. LOS velocity is to be differentiated with vertical and horizontal velocity. LOS velocity comprised of movements along the vertical and horizontal axis, with negative values inferring ground deformation velocity in down-east directions and positive values in up-west directions. Due to descending data being insufficient in 2024, we cannot do LOS decomposition to get separate vertical and horizontal velocity. [↑](#footnote-ref-6)
8. The RMSE is calculated with both measurements from the same date (epoch). [↑](#footnote-ref-7)
9. More on multipath error: <https://www.sbg-systems.com/glossary/multipath-error/> [↑](#footnote-ref-8)