

High-Resolution Subsurface Mapping Using SBAS-InSAR and Multilayer Compaction Monitoring Wells for Land Subsidence Mitigation

Author One* Author Two* Author Three[†] Author Four[‡]

July 21, 2025

Abstract

Land subsidence from excessive groundwater extraction has long affected the Choushui River Fluvial Plain (CRFP) in Taiwan, increasing flood risks and threatening infrastructure like the Taiwan High Speed Rail (THSR). A network of multilayer compaction monitoring wells (MLCWs) aids in understanding subsidence mechanisms, but their high installation costs limit dense deployment. This study integrates SBAS-InSAR surface deformation data with MLCW measurements using geographically weighted regression (GWR) to map layer-wise compaction regionally. This approach enables cost-effective, high-resolution monitoring to inform groundwater management and infrastructure protection.

Keywords: land subsidence; mitigation; layer-wise compaction

Contents

1	Introduction	1
2	Study Area Background	1
3	Data Sets	2
3.1	Multilayer Compaction	2
3.2	Cumulative vertical displacement	2

*Department of X, University A, Country. author.one@example.com

[†]Department of Y, University B, Country. author.three@example.com

[‡]Department of Z, University C, Country. author.four@example.com

4	Methodology	4
4.1	Geographically Temporally Weighted Regression	4
4.1.1	Model Calibration	4
4.1.2	Posterior Uncertainty Assessment	6
4.2	Curve Fitting	7
5	Results	7
6	Discussion	7
7	Conclusion	7

List of Figures

1	Study Area	1
2	Cross-section	2
3	Comparison of InSAR results and leveling data. (a) The results from InSAR processing. (b) The ground truth data from leveling surveys.	4

List of Tables

1	Summary of the Sentinel-1A datasets used in this study.	3
---	---	---

1 Introduction

2 Study Area Background

The CRFP belongs to the western coastal region of central Taiwan. CRFP has an area of approximately 2000 km^2 , which belongs to the regions of Changhua, Yunlin, and the northern part of the Chiayi counties, with a surface elevation ranging from 0 m to 100 – 150 m above sea level. The CRFP boundary is shaped by Douliu Hill and Bagua Tableland on the eastern side, the Wu River to the north, the Beigang River to the south, and the Taiwan Strait to the west (Figure 1). The hydrogeological structures of the study area are divided into three sections from the east to the west, namely proximal fan, middle fan, and distal fan [Survey, 1999]. Each section was comprised of various sedimentary materials, with the average grain sizes decreasing from hilly regions to coastal areas. The sedimentary materials in the CRFP are divided into four primary groups, grading from very coarse grains to very fine grains: gravel, coarse sand, fine sand, and clay or silt [Survey, 1999]. These materials are weathered products of rock formations located in the upstream watershed of the study area, such as slate, quartzite, shale, sandstone, and mudstone [Liu et al., 2004, Hung et al., 2015]. The borehole profiles suggest that gravel and coarse sand are primarily present at the proximal fan and part of the middle fan, whereas the distal fan mainly witnesses fine-grain materials, including fine sand, clay, and silt (Figure 2).



Figure 1: Study Area, Cross Section, and Borehole



Figure 2: Cross-section showing the distribution of sedimentary materials along profile AA'.

3 Data Sets

3.1 Multilayer Compaction

A multilayer compaction monitoring well (MLCW) is a specialized borehole extensometer that captures the subtle subsurface compaction by reading measurements at magnetic rings, strategically installed at boundaries between significant aquifers, or transitions between fine and coarse sedimentary materials, as defined by the Geological Survey and Mining Management Agency (GSMMA). The depth of each MLCW often extends up to 300 m, with 21 to 26 magnetic rings anchored throughout the profile. Based on hydrogeological properties, aquifer units at each well are determined, each containing a number of corresponding magnetic rings, providing measurements of aquifer-specific compaction. The installation and measurement approaches of the MLCWs have been comprehensively described by [Hung et al., 2021]. In this study, 29 MLCWs were employed, with monthly data collected from April 2016 to November 2021.

3.2 Cumulative vertical displacement

This study derived cumulative vertical displacement from an 8-year dataset of Sentinel-1 SAR images (2016 - 2024), acquired from both ascending and descending orbits; the image properties are summarized in Table 1. First, the images from each acquisition orbit were processed separately applying the *hyp3-isce2* plugin, part of the Hybrid Pluggable Processing Pipeline (HyP3) [Hogenson et al., 2025]. For interferogram formation, each SAR image was paired with up to four subsequent consecutive images to minimize temporal decorrelation. Next, each image pair was analyzed using the InSAR Scientific Computing Environment 2 (ISCE2) TOPS workflow, which sequentially performs burst-level coregistration of Single Look Complex (SLC) images, interferogram generation, waterbody masking and topographic phase correction with the Copernicus GLO-30 DEM, phase unwrapping via the SNAPHU algorithm, and geocoding to produce unwrapped interferograms suitable for further time-series InSAR analysis. A detailed description of the ISCE2 workflow is available in [Rosen et al., 2012, Yagüe-Martínez et al., 2016, Fattahi et al., 2017].

Table 1: Summary of the Sentinel-1A datasets used in this study.

Parameters	Ascending	Descending
Relative Orbit (Path)	69	105
Acquisition Period	4/2016 – 11/2021	
Number of Images	266	264
Acquisition Mode	Interferometric Wide (IW)	
Polarization	VV	
Incidence Angles	32° – 38°	38° – 43°
Satellite Headings	347.63°	192.37°

Subsequently, a time-series analysis was conducted on the unwrapped interferograms from ascending and descending orbits separately using the small baseline subset approach, provided by Miami InSAR Time-series software (MintPy) [Yunjun et al., 2019]. The valid interferogram network was formed through a two-stage selection. The minimum spanning tree (MST) [Perissin and Wang, 2012] first identified the most coherent interferograms to connect all SAR images, by using the inverse of the average spatial coherence of all interferograms as weight. After this, any interferograms not included in the MST were excluded if their average spatial coherence was lower than 0.3. After network formation, the optimal values of the interferometric phase timeseries were estimated through a network inversion using the inverse of the phase variance as weight [Tough et al., 1995]. The temporal coherence, as a product of the network inversion process, was utilized to evaluate the reliability of the estimated value at each pixel [Pepe and Lanari, 2006]. Pixels with temporal coherence (γ_{temp}) below 0.65 were then masked from further processing. Regarding atmospheric correction, the tropospheric delay components were removed based on the empirical linear relationship between InSAR phase delay and elevation [Doin et al., 2009]. All interferograms from both ascending and descending orbits were referenced to a single stable point, which was located far from the subsiding area (Figure 3a). Finally, the line-of-sight deformation (d_{LOS}) was decomposed into the east-west (d_{E-W}) and vertical (d_U) components, assuming the deformation in north-south direction (d_{N-S}) was negligible. The decomposition employed the incidence angles (θ) and satellite heading angles (α) from both ascending (asc) and descending ($desc$) images [Hanssen, 2001]:

$$\begin{bmatrix} d_U \\ d_{E-W} \end{bmatrix} = \begin{bmatrix} \cos(\theta^{asc}) & \sin(\theta^{asc}) \cos(\alpha^{asc}) \\ \cos(\theta^{desc}) & \sin(\theta^{desc}) \cos(\alpha^{desc}) \end{bmatrix}^{-1} \begin{bmatrix} d_{LOS}^{asc} \\ d_{LOS}^{desc} \end{bmatrix}. \quad (1)$$

Validation of the vertical displacements was performed against precise leveling survey, with benchmarks locations shown in Figure 1. First, reliable measurement points ($\gamma_{temp} \geq 0.65$) within a 200 m radius of each benchmark were selected. Average velocities were then derived for these points from both the InSAR and annual leveling survey time series, calculated as the slope of the best-fitting line to their respective displacement time series. Such a comparison

69 based on rates was required due to the different temporal sampling intervals of the two datasets
 70 Figure 3.

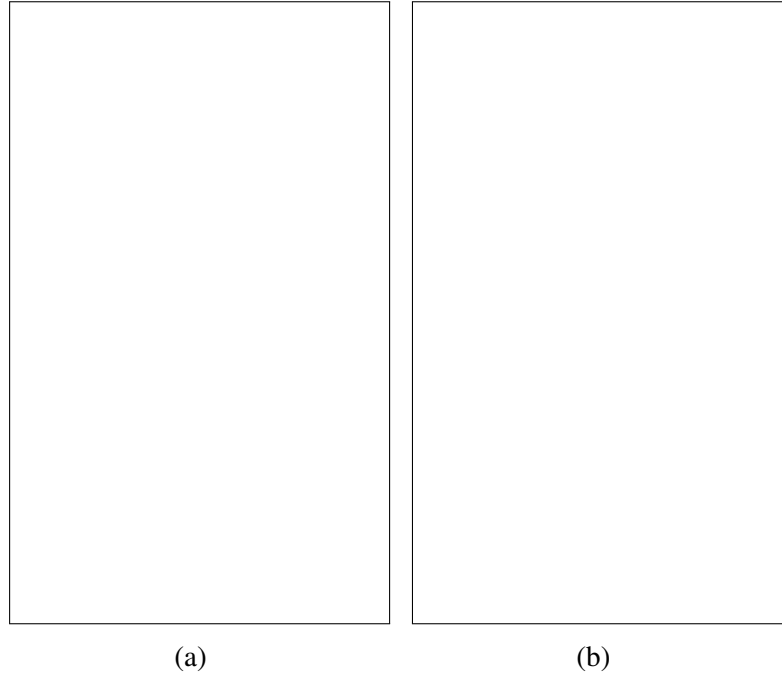


Figure 3: Comparison of InSAR results and leveling data. (a) The results from InSAR processing.
 (b) The ground truth data from leveling surveys.

71 4 Methodology

72 4.1 Geographically Temporally Weighted Regression

73 4.1.1 Model Calibration

74 Before discussing the GTWR model, it is useful to review the fundamental concepts of
 75 ordinary least squares (OLS) regression and geographically weighted regression (GWR). The
 76 OLS models the relationship between a dependent variable and one or more predictors, under
 77 the assumption that the relationships being modeled are spatially stationary, expressed as:

$$y_i = \sum_{k=0}^p \beta_k x_{ik} + \varepsilon_i \quad (2)$$

78 where y_i denotes the dependent variable at the i -th location; x_{ik} represents the independent
 79 variables (or predictors), with $x_{i0} = 1$ for the intercept β_0 ; β_k indicates the regression coefficient
 80 corresponding to the k -th predictor; ε_i is the random error at the i -th location, and p is the
 81 number of independent variables (excluding the intercept). The OLS method assigns a single set
 82 of coefficients to all observations regardless of their geographical locations, which deliberately
 83 neglects potential spatial variation and can cause model misspecification [Fotheringham et al.](#)
 84 [\[2017\]](#). The GWR approach allows the regression coefficients of the predictors to vary at

different locations, thereby capturing local variations [Brunsdon et al., 1996, Fotheringham, 2002]. The GWR is formulated as follows:

$$y_i = \sum_{k=0}^p \beta_k(u_i, v_i) x_{ik} + \varepsilon_i \quad (3)$$

where the coefficients appear similar to those in an OLS model. However, the key difference is that the GWR coefficients β_k are indexed by the spatial coordinates (u_i, v_i) of the i -th point. Equation 3 indicates that the relationships between dependent and independent variables are spatially nonstationary, which provides a more realistic model than the global regression, which assumes spatially constant relationships. However, the environmental quantities not only exhibit spatial but also temporal relationships. Therefore, Huang et al. [2010] introduced the GTWR model to account for spatiotemporal nonstationarity. The general form of the GTWR model can be expressed as:

$$y_i = \sum_{k=0}^p \beta_k(u_i, v_i, t_i) x_{ik} + \varepsilon_i \quad (4)$$

The GTWR terms mirror those in GWR, adding the temporal dimension t_i . The GTWR model is calibrated based on the assumption that observations closer to the i -th point in spatiotemporal coordinate system have greater influence on the estimation of $\beta_k(u_i, v_i, t_i)$ than more distant ones. The estimated parameters $\hat{\beta}(u_i, v_i, t_i)$ can be obtained by:

$$\hat{\beta}(u_i, v_i, t_i) = (\mathbf{X}^T \mathbf{W}(u_i, v_i, t_i) \mathbf{X})^{-1} \mathbf{X}^T \mathbf{W}(u_i, v_i, t_i) \mathbf{y} \quad (5)$$

where $\mathbf{W}(u_i, v_i, t_i)$ is a diagonal matrix with elements denoting the spatiotemporal weights assigned to the observations associated with the i -th point. The weights are usually obtained through a kernel function that implement either a fixed bandwidth (using specific distance thresholds) or an adaptive bandwidth (seeking an optimal number of nearest observations), as described by Páez et al. [2002]. A fixed bandwidth can introduce estimation bias dependent on observation density. More specifically, in regions with dense observations, an excessive number of neighbors might lead to over-smoothing, while an insufficient number is more likely to cause uncertain estimates. The adaptive bandwidth ensures each regression point is informed by exactly the same number of neighbors, regardless of data density. Due to the sparse distribution of the MLCW stations across the study area, this study employed the adaptive bi-square kernel function, defined as:

$$W_{ij} = \begin{cases} [1 - (d_{ij}^{ST}/h_i)^2]^2, & \text{if } d_{ij}^{ST} < h_i \\ 0, & \text{otherwise} \end{cases} \quad (6)$$

where d_{ij}^{ST} denotes the spatiotemporal distances between location i and j , and h_i is the adaptive bandwidth. Since spatial and temporal distances are measured in different units and exhibit different scale effects, [Wu et al. \[2014\]](#) proposed an integrated formulation, expressed as:

$$\begin{cases} d_{ij}^{ST} = \lambda d_{ij}^S + (1 - \lambda) d_{ij}^T + 2\sqrt{\lambda(1 - \lambda)} d_{ij}^S d_{ij}^T \cos(\xi), & t_j < t_i \\ d_{ij}^{ST} = \infty, & t_j > t_i \end{cases} \quad (7)$$

where t_i and t_j are sampled times at locations i and j ; λ and $\xi \in [0, \pi]$ are adjustment parameters. Noticeably, Equation 7 implies that the GTWR model will simplify to GWR (eq. 3) when $\lambda = 1$, and space-time interaction effects reach their maximum when $\xi = 0$. In practice, the bandwidths h_i , λ , and ξ can be determined and optimized using a corrected version of the Akaike Information Criterion (AICc) [[Hurvich et al., 2002](#)], defined as:

$$\text{AICc} = 2n \log_e(\hat{\sigma}) + n \log_e(2\pi) + n \left\{ \frac{n + \text{tr}(\mathbf{S})}{n - 2 - \text{tr}(\mathbf{S})} \right\} \quad (8)$$

where n is the sample size, $\hat{\sigma}$ is the estimated standard deviation of the error term, and $\text{tr}(\mathbf{S})$ denotes the trace of the hat matrix \mathbf{S} . The hat matrix \mathbf{S} is defined as:

$$\mathbf{S} = \{S_{ij}\}, \text{ where } S_{ij} = \mathbf{x}_i^T (\mathbf{X}^T \mathbf{W}(u_i, v_i, t_i) \mathbf{X})^{-1} \mathbf{X}^T \mathbf{W}(u_i, v_i, t_i) \quad (9)$$

where \mathbf{x}_i^T represents the values of independent variables at the i -th point. The hat matrix \mathbf{S} transforms observed values into fitted values through spatiotemporal weightings, as in $\hat{\mathbf{y}} = \mathbf{S}\mathbf{y}$, with its trace representing the effective number of parameters in the model [Hoaglin and Welsh \[1978\]](#). The detailed information of these components can be found in [Fotheringham \[2002\]](#).

4.1.2 Posterior Uncertainty Assessment

When the optimal bandwidth is obtained through the calibration process, the fitted value $\hat{\mathbf{y}}$ at any unsampled location can be calculated using the hat matrix relationship as previously mentioned. The prediction error variance for each fitted value $\hat{\mathbf{y}}$ is defined as:

$$\sigma_i^2 = \hat{\sigma}^2 \left(1 + \mathbf{x}_i^T (\mathbf{X}^T \mathbf{W}(u_i, v_i, t_i) \mathbf{X})^{-1} \mathbf{x}_i \right) \quad (10)$$

where $\hat{\sigma}^2$ represents the overall model error variance, which is defined as:

$$\hat{\sigma}^2 = \frac{\sum_{i=1}^N (y_i - \hat{y}_i)^2}{n - 2\text{tr}(\mathbf{S}) + \text{tr}(\mathbf{S}^T \mathbf{S})} \quad (11)$$

The first term in Equation 10 accounts for the intrinsic model residual variance, while the second term assesses the additional uncertainty that arises when estimating local relationships

from limited neighboring observations.

4.2 Curve Fitting

5 Results

6 Discussion

7 Conclusion

Acknowledgements

Acknowledge any funding, institutional support, or personal contributions that helped your research.

References

- Chris Brunsdon, A. Stewart Fotheringham, and Martin E. Charlton. Geographically weighted regression: A method for exploring spatial nonstationarity. *Geographical Analysis*, 28(4): 281–298, 1996. doi: <https://doi.org/10.1111/j.1538-4632.1996.tb00936.x>. URL <https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1538-4632.1996.tb00936.x>.
- M.-P. Doin, C. Lasserre, G. Peltzer, O. Cavalié, and C. Doubre. Corrections of stratified tropospheric delays in sar interferometry: Validation with global atmospheric models. *Journal of Applied Geophysics*, 69(1):35–50, 2009. ISSN 0926-9851. doi: <https://doi.org/10.1016/j.jappgeo.2009.03.010>. URL <https://www.sciencedirect.com/science/article/pii/S0926985109000603>. Advances in SAR Interferometry from the 2007 Fringe Workshop.
- Heresh Fattahi, Piyush Agram, and Mark Simons. A network-based enhanced spectral diversity approach for tops time-series analysis. *IEEE Transactions on Geoscience and Remote Sensing*, 55(2):777–786, 2017. doi: 10.1109/TGRS.2016.2614925.
- A. Stewart Fotheringham, Wenbai Yang, and Wei Kang. Multiscale geographically weighted regression (mgwr). *Annals of the American Association of Geographers*, 107(6):1247–1265, 2017. doi: 10.1080/24694452.2017.1352480. URL <https://doi.org/10.1080/24694452.2017.1352480>.
- AS Fotheringham. Geographically weighted regression: the analysis of spatially varying relationships. *John Wiley and Sons Ltd*, 2002.
- Ramon F. Hanssen. *Radar Interferometry: Data Interpretation and Error Analysis*, volume 2 of *Remote Sensing and Digital Image Processing*. Springer Dordrecht, 1 edition, 2001. ISBN 978-0-7923-6945-5. doi: 10.1007/0-306-47633-9. URL <https://doi.org/10.1007/0-306-47633-9>.

- David C Hoaglin and Roy E Welsch. The hat matrix in regression and anova. *The American Statistician*, 32(1):17–22, 1978.
- Kirk Hogenson, Heidi Kristenson, Joseph Kennedy, Andrew Johnston, James Rine, Thomas Logan, Jiang Zhu, Forrest Williams, Jake Herrmann, Jacquelyn Smale, and Franz Meyer. Hybrid pluggable processing pipeline (hyp3): A cloud-native infrastructure for generic processing of sar data, May 2025. URL <https://doi.org/10.5281/zenodo.15498989>.
- Bo Huang, Bo Wu, and Michael Barry. Geographically and temporally weighted regression for modeling spatio-temporal variation in house prices. *International journal of geographical information science*, 24(3):383–401, 2010.
- W. C. Hung, C. Wang, C. Hwang, Y. A. Chen, H. C. Chiu, and S. H. Lin. Multiple sensors applied to monitor land subsidence in central taiwan. *Proc. IAHS*, 372:385–391, 2015. ISSN 2199-899X. doi: 10.5194/piahs-372-385-2015. URL <https://piahs.copernicus.org/articles/372/385/2015/>.
- Wei-Chia Hung, Cheinway Hwang, Michelle Sneed, Yi-An Chen, Chi-Hua Chu, and Shao-Hung Lin. Measuring and interpreting multilayer aquifer-system compactions for a sustainable groundwater-system development. *Water Resources Research*, 57(4):e2020WR028194, 2021. doi: <https://doi.org/10.1029/2020WR028194>. URL <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2020WR028194>. e2020WR028194 2020WR028194.
- Clifford M. Hurvich, Jeffrey S. Simonoff, and Chih-Ling Tsai. Smoothing parameter selection in nonparametric regression using an improved akaike information criterion. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 60(2):271–293, 01 2002. ISSN 1369-7412. doi: 10.1111/1467-9868.00125. URL <https://doi.org/10.1111/1467-9868.00125>.
- Chih-Hsi Liu, Yii-Wen Pan, Jyh-Jong Liao, Chen-Tair Huang, and Shoung Ouyang. Characterization of land subsidence in the choshui river alluvial fan, taiwan. *Environmental Geology*, 45(8):1154–1166, 2004. ISSN 1432-0495. doi: 10.1007/s00254-004-0983-6. URL <https://doi.org/10.1007/s00254-004-0983-6>.
- A. Pepe and R. Lanari. On the extension of the minimum cost flow algorithm for phase unwrapping of multitemporal differential sar interferograms. *IEEE Transactions on Geoscience and Remote Sensing*, 44(9):2374–2383, 2006. doi: 10.1109/TGRS.2006.873207.
- Daniele Perissin and Teng Wang. Repeat-pass sar interferometry with partially coherent targets. *IEEE Transactions on Geoscience and Remote Sensing*, 50(1):271–280, 2012. doi: 10.1109/TGRS.2011.2160644.

- Antonio Páez, Takashi Uchida, and Kazuaki Miyamoto. A general framework for estimation and inference of geographically weighted regression models: 1. location-specific kernel bandwidths and a test for locational heterogeneity. *Environment and Planning A*, 34(4): 733–754, 2002. doi: 10.1068/a34110. URL <https://doi.org/10.1068/a34110>.
- Paul A. Rosen, Eric Gurrola, Gian Franco Sacco, and Howard Zebker. The insar scientific computing environment. In *EUSAR 2012; 9th European Conference on Synthetic Aperture Radar*, pages 730–733, 2012.
- Central Geological Survey. Project of groundwater monitoring network in taiwan during first stage: Research report of choushui river alluvial fan. Report, 1999.
- J. A. Tough, D. Blacknell, and S. Quegan. A statistical description of polarimetric and interferometric synthetic aperture radar data. *Proceedings of the Royal Society of London. Series A: Mathematical and Physical Sciences*, 449(1937):567–589, 1995. doi: 10.1098/rspa.1995.0059. URL <https://royalsocietypublishing.org/doi/abs/10.1098/rspa.1995.0059>.
- Bo Wu, Rongrong Li, and Bo Huang. A geographically and temporally weighted autoregressive model with application to housing prices. *International Journal of Geographical Information Science*, 28(5):1186–1204, 2014. doi: 10.1080/13658816.2013.878463. URL <https://doi.org/10.1080/13658816.2013.878463>.
- Néstor Yagüe-Martínez, Pau Prats-Iraola, Fernando Rodríguez González, Ramon Brcic, Robert Shau, Dirk Geudtner, Michael Eineder, and Richard Bamler. Interferometric processing of sentinel-1 tops data. *IEEE Transactions on Geoscience and Remote Sensing*, 54(4):2220–2234, 2016. doi: 10.1109/TGRS.2015.2497902.
- 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: 10.1016/j.cageo.2019.104331. URL <https://www.sciencedirect.com/science/article/pii/S0098300419304194>.