A Multivariate Multi-step Deep Learning Framework for InSAR Displacement Prediction

Maria Maia[1(✉)], Luís Miguel Matos[1,2], Luís Magalhães[1,3], Joaquim Tinoco[4], and Steffan Davies[5]

[1] ALGORITMI/LASI Research Center – Department Information Systems, Minho University, Guimarães, Portugal
pg55904@alunos.uminho.pt, {luis.matos,lmagalhaes}@dsi.uminho.pt
[2] EPMQ, CCG ZGDV Institute, Guimarães, Portugal
[3] CVIG, CCG ZGDV Institute, Guimarães, Portugal
[4] ISISE, ARISE, Department of Civil Engineering, University of Minho, Guimarães, Portugal
jtinoco@civil.uminho.pt
[5] Spotlite-AETHRA, Rua Pedro Nunes, IPN, C 3030-199 Coimbra, Portugal
steffan@spotlitedata.com

Abstract. Accurate forecasting of slope displacements is critical for landslide risk management and the development of reliable early warning systems. This study investigates the application of three deep learning architectures: Temporal Convolutional Network (TCN), Patch Time Series Transformer (PatchTST), and Temporal Fusion Transformer (TFT) for multi-step forecasting of InSAR-derived slope displacements across multiple prediction horizons. The proposed framework adopts a joint multivariate representation that simultaneously predicts displacements across all monitored points, incorporating spatial context through neighbouring measurements and daily precipitation as an environmental covariate. Models are evaluated on both Persistent Scatterer (PS) and Small Baseline Subset (SBAS) datasets, considering horizontal and vertical displacement components independently. Results show that all three models achieve R^2 values consistently above 0.75, with TFT demonstrating the strongest overall performance on PS data and PatchTST reaching the highest R^2 of 0.91 on the SBAS vertical component at $H = 5$. The TCN, while generally less competitive at longer horizons, outperforms attention-based models on the SBAS horizontal component, suggesting that its convolutional inductive bias is better suited to signals exhibiting rapid directional reversals. These findings highlight the complementary strengths of the investigated architectures and establish a foundation for future integration of additional environmental covariates and more rigorous validation strategies.

Keywords: Slope Displacement Prediction · InSAR · Deep Learning · Network Level Scalability

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2027
O. Gervasi et al. (Eds.): ICCSA 2026, LNCS 16769, pp. 639–650, 2027.
https://doi.org/10.1007/978-3-032-30494-0_40

1 Introduction

Slope instabilities, including landslides, rock falls, and debris flows, pose significant hazards in both natural and engineered slopes across all geographic regions of the world. These movements can be triggered by natural factors such as intense rainfall, rapid snowmelt, earthquakes, and volcanic eruptions, as well as by human activities such as excavation and road construction [1]. The resulting hazards are responsible for thousands of deaths and injuries, as well as monetary losses [2], highlighting the need for effective investigation, monitoring, and remediation measures.

Advanced remote sensing techniques, including Interferometric Synthetic Aperture Radar (InSAR) together with Machine Learning (ML) approaches, have significantly improved the detection, monitoring and forecasting of slope displacements [3,4]. InSAR techniques such as PS-InSAR [5] and SBAS-InSAR [6] enable systematic, repeated observations of ground deformation over large areas with millimetric precision, supporting both rapid post-event mapping and long-term monitoring of slow-moving slopes [7].

Building upon these techniques, ML, and more specifically Deep Learning (DL), have emerged as powerful tools to further enhance displacement forecasting. Given that InSAR derived measurements are inherently acquired as multitemporal time series, DL architectures designed for sequential data have shown strong capability in capturing the complex and nonlinear temporal patterns of slope displacements. The combination of these approaches has therefore opened new perspectives for the development of more accurate and reliable early warning systems, contributing to improved landslide risk management at both local and regional scales [8].

Motivated by the limited exploration of advanced DL architectures in the geotechnical domain, this study investigates three models with distinct and complementary strengths: the Temporal Fusion Transformer (TFT) [9], adopted for its ability to handle heterogeneous multivariate inputs while providing interpretable insights through its attention mechanism; the Patch Time Series Transformer (PatchTST) [10], selected for its patch-based tokenisation that captures local temporal semantics and attends to longer historical windows with reduced computational cost; and the Temporal Convolutional Network (TCN) [11], included as a strong convolutional baseline, leveraging dilated causal convolutions for efficient modelling of long-range temporal dependencies. All three models are evaluated under a unified multi-step forecasting framework across multiple prediction horizons, considering all monitored points simultaneously, where the spatial context of neighbouring points at a given temporal *lag* is incorporated as input features.

The remainder of this paper is organized as follows: Sect. 2 presents a review of the related work on DL approaches for landslide displacement forecasting, Sect. 3 describes the proposed methodology, Sect. 4 presents the experimental results and Sect. 5 draws the main conclusions, while outlining directions for future work.

2 Related Work

The application of DL to the forecasting of displacement derived from InSAR data has grown substantially over the past decade.

Building on this foundation, Mirmazloumi et al. [12] coupled point-scale LSTM forecasting with a spatial early alert detector that considers adjacent measurement points, demonstrating that spatial context reduces false alarm rates in mining collapse scenarios. Wang et al. [13] proposed the Spacetimeformer network, which jointly attends to temporal and spatial relationships across InSAR points via a dual global-local self-attention mechanism, outperforming LSTM and vanilla Transformer baselines, though at the cost of extensive site-specific preprocessing that constrains scalability. Moualla et al. [14] addressed irregular Sentinel-1 temporal sampling, showing that TFT achieves the best trade-off between accuracy and computational cost, but treating each monitoring point independently without exploiting spatial structure.

Taken together, these studies reveal a clear tension: approaches that model spatial dependencies tend to require regular temporal sampling and site-specific preprocessing, while methods designed for irregular time series treat each point independently. Additionally, spatial context is often incorporated heuristically through neighbourhood averaging [12], rather than through learned spatiotemporal representations.

The present study addresses this gap by investigating three architecturally distinct deep learning models under a unified multi-step forecasting framework that simultaneously predicts displacements across all monitored points and incorporates spatial context through neighbour measurements at a given temporal lag as explicit input features. This work extends the preliminary investigation of Owusu-Ansah et al. [15], who applied MR and ANN to InSAR vertical displacement time series from highway slopes using a sliding-window strategy, achieving $R^2 = 0.93$ but identifying the need for additional input variables and more robust model optimization.

3 Materials and Methods

3.1 Data

InSAR Displacement Data. The data used in this study were derived from Sentinel-1 acquisitions processed via PS-InSAR and SBAS-InSAR techniques. The resulting datasets are organized into two resolution categories: Persistent Scatterer (PS) and Small Baseline Subset (SBAS), each containing separate files for horizontal and vertical displacement components, further stratified by monitored element: slope, pavement and unclassified area of interest (AOI). Each file comprises 88 columns, including longitude, latitude and the displacement time series spanning from 5 June 2023 to 16 November 2025. The number of measurement points per file varies with both resolution and element type, as summarised in Table 1.

Table 1. Dimensions of the IP3 InSAR datasets.

Category	PS rows	SBAS rows
Slope	38	34
Pavement	92	89
AOI	933	671

Precipitation Data. Precipitation data were obtained from the PDIR-Now product, available through the CHRS Data Portal [16]. Daily accumulated precipitation values (mm) were retrieved in NetCDF format for the period coinciding with the InSAR time series. Since the PDIR-Now grid resolution does not align directly with the InSAR measurement point coordinates, a nearest-neighbour spatial matching procedure was applied to associate each point with its closest grid cell. The resulting data were then restructured into a tabular format and subsequently merged with the corresponding InSAR displacement files on the basis of matching spatial coordinates, yielding a unified dataset ready for model input.

3.2 Methods

All models share a common input representation and training pipeline. The displacement time series from all monitored points are first reorganised into a sliding-window format, extracting sequences of $K = 4$ consecutive observations as input and the following H observations as prediction targets, a configuration empirically determined in [15] as the best-performing baseline. This windowing is applied independently to each point and each forecasting horizon $H \in \{1, 5, 10\}$, corresponding to prediction windows of approximately 12, 60 and 120 days, respectively. Rather than treating each monitoring point independently, the framework constructs a joint multivariate representation capturing the spatial context of the entire network, with neighbouring points selected based on geographic proximity to the point of interest and modelled through direct correlations without distance-based weighting. At each time step K, the displacement values of all monitored points are stacked into a single vector and the corresponding precipitation values are concatenated alongside them, producing a combined feature vector of dimension $n_{\text{points}} \times 2$ per time step. This structure is repeated across the $K = 4$ lag steps, resulting in 3D input arrays of shape $(\ n_{\text{samples}},\ 4,\ n_{\text{points}} \times 2)$. The target arrays are constructed analogously, stacking the H future displacement values for all points into a flat output vector of dimension $H \times n_{\text{points}}$. Model validation is assessed through 5-fold cross-validation without shuffling, preserving the temporal ordering of the data, and the best-performing fold is retrained on the full dataset to produce final predictions.

The three main architectures investigated are the TCN, TFT, and PatchTST. The TCN [11] replaces recurrence with stacked dilated causal convolutions,

enabling the capture of long-range temporal dependencies while maintaining computational efficiency and avoiding the vanishing gradient problems of recurrent models. The architecture comprises a single TCN block with $F = 64$ filters, a kernel size of 3, and four dilation levels $\{1, 2, 4, 8\}$, with causal padding and skip connections. A dropout rate of 0.1 is applied within the block, and the final representation is passed to a dense output layer with linear activation. Training is conducted with the Adam optimiser and MSE loss, with early stopping over a maximum of 100 epochs and a batch size of 32. The TFT [9] combines a multi-head self-attention mechanism with position-wise feed-forward layers and residual connections, offering interpretable multi-horizon forecasting through its attention weights. The input sequence is first projected to a latent dimension of $d_{\text{model}} = 64$, followed by a multi-head attention layer with $n_{\text{heads}} = 4$ and a feed-forward network of dimension 128 with ReLU activation. Residual connections and layer normalisation are applied after both sub-layers, and a global average pooling operation aggregates the temporal representations before a dropout layer and the final output projection. The PatchTST [10] extends this Transformer encoder by introducing a patch-based tokenisation of the input sequence prior to the attention mechanism: a 1D convolutional layer with patch length 2 and $F = 64$ filters segments the input into overlapping sub-series patches reducing the effective sequence length while preserving local temporal structure. This design is particularly advantageous for longer horizons, as the patching step reduces the quadratic cost of self-attention. Both Transformer-based models are trained under the same protocol as the TCN.

The predictive performance of all models is assessed using three metrics: the coefficient of determination (R^2), Mean Absolute Error (MAE) and Normalised MAE (NMAE), defined as:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \tag{1}$$

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \tag{2}$$

$$\text{NMAE} = \frac{\text{MAE}}{\max(y) - \min(y)} \tag{3}$$

where y_i denotes the observed displacement value, $\hat{y}_i$ the corresponding predicted value, $\bar{y}$ the mean of observed values, and n the total number of samples.

4 Results

Figure 1 presents the time series comparison between observed and predicted vertical displacements for a representative PS monitoring point for $H = 1$[1].

[1] Time series comparisons (Figs. 1, 3, 5 and 7) are shown for $H = 1$ only, as qualitative behaviour is consistent across all horizons.

All three models closely track the observed signal, capturing both the gradual negative trend and the more abrupt displacement events characteristic of the mid-2024 period. The TFT and PatchTST predictions exhibit smoother trajectories compared to the TCN, which tends to follow higher-frequency fluctuations more closely, though at the cost of occasional over-responses to short-term noise.

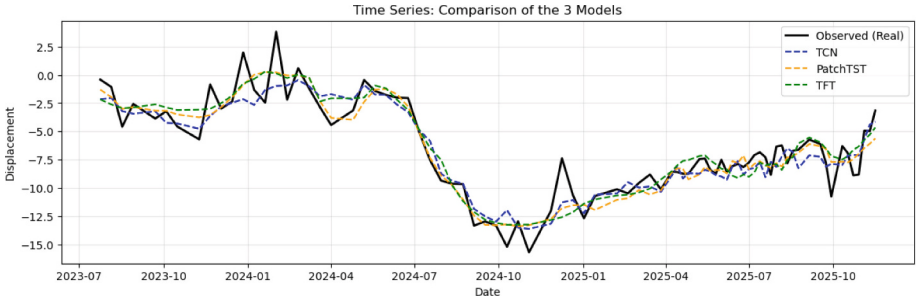

Fig. 1. Time series comparison of observed and predicted PS vertical displacements for $H = 1$.

The quantitative evaluation across all three horizons is summarised in Fig. 2. The TFT consistently achieves the lowest MAE and NMAE values across all forecasting horizons, while also attaining the highest R^2 scores, particularly at $H = 5$ and $H = 10$ where it reaches approximately 0.88 and 0.78, respectively. PatchTST performs comparably to TFT at $H = 1$ but shows a slight degradation at longer horizons. The TCN, while competitive at $H = 1$ ($R^2 \approx 0.85$), exhibits a more pronounced decline in performance as the horizon extends, suggesting reduced capacity to capture multi-step displacement dynamics relative to the attention-based architectures.

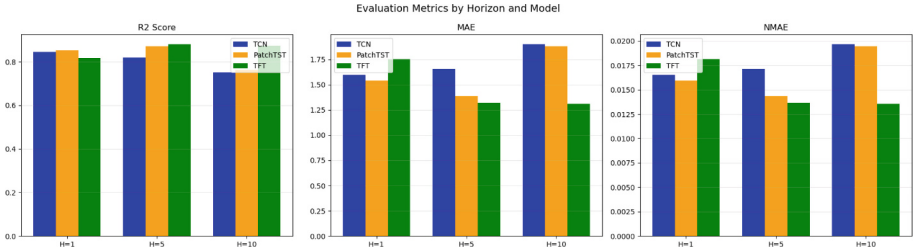

Fig. 2. Quantitative evaluation across all horizons for PS vertical displacements.

The time series comparison for PS horizontal displacement for $H = 1$, shown in Fig. 3, reveals a broadly similar pattern to the vertical component, with a pronounced negative trend developing from early 2024 and partial recovery towards

the end of the monitoring period. However, the horizontal signal exhibits greater high-frequency variability and more abrupt short-term fluctuations, posing a more challenging forecasting target. The TFT and PatchTST models produce smoother predictions that capture the dominant trend effectively, while the TCN tends to lag slightly behind rapid directional changes, particularly during the transition period between mid-2023 and early 2024.

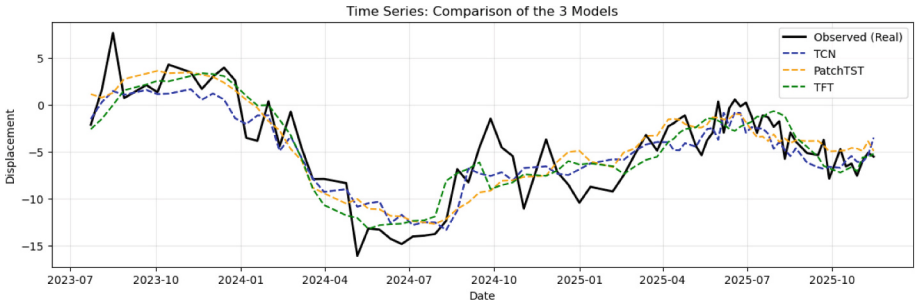

Fig. 3. Time series comparison of observed and predicted PS horizontal displacements for $H = 1$.

The evaluation metrics presented in Fig. 4 confirm that the horizontal component is generally harder to predict than the vertical, as reflected by slightly higher MAE and NMAE values across all models and horizons. Nonetheless, the ranking of models remains consistent: TFT achieves the lowest error and highest R^2 across all horizons, followed closely by PatchTST, with the TCN exhibiting higher MAE values. A notable observation is that the TCN's R^2 score remains relatively stable across horizons (≈ 0.85), suggesting that while its absolute error is higher, its ability to explain variance in the horizontal displacement does not degrade as sharply with increasing horizon as observed in the vertical component.

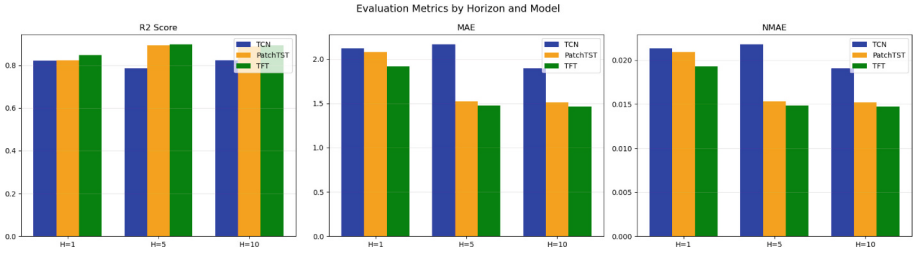

Fig. 4. Quantitative evaluation across all horizons for PS horizontal displacements.

The SBAS vertical displacement time series for $H = 1$, shown in Fig. 5, reveals a more pronounced and sustained negative trend than its PS counterpart, with a sharp acceleration in displacement occurring around mid-to-late

2024, reaching values below −25 mm before a partial recovery in early 2025. This abrupt event represents the most challenging feature of the entire dataset, and the models' ability to anticipate it varies considerably. The PatchTST and TFT predictions track the gradual pre-event trend well but underestimate the magnitude of the rapid displacement episode, smoothing over the most extreme values. The TCN, by contrast, exhibits a more erratic response during this period, with predictions diverging more significantly from the observed signal, suggesting reduced robustness to sudden regime changes.

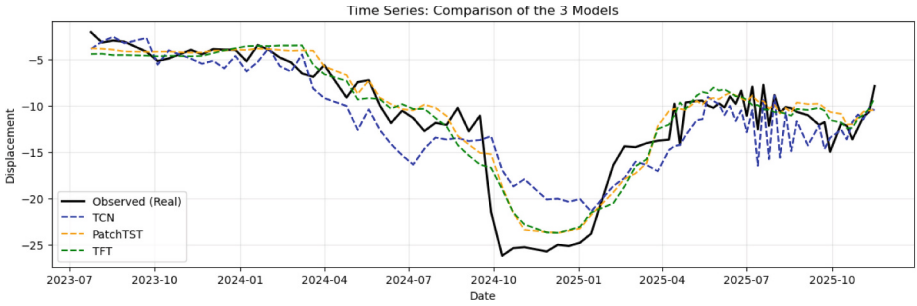

Fig. 5. Time series comparison of observed and predicted SBAS vertical displacements for $H = 1$.

The quantitative metrics in Fig. 6 reflect this increased difficulty. While R^2 values remain broadly comparable to the PS results ($\approx 0.80 - 0.90$ across models and horizons), the absolute MAE values are higher, with the TCN reaching ≈ 2.3 at $H = 1$ compared to ≈ 1.8 for TFT. A particularly noteworthy result is that PatchTST achieves its best relative performance at H = 5, surpassing TFT in R^2 score, suggesting that its patch-based tokenisation may be especially advantageous for capturing medium-range dynamics. At $H = 10$, all models show some degradation, though TFT and PatchTST maintain lower error levels than the TCN.

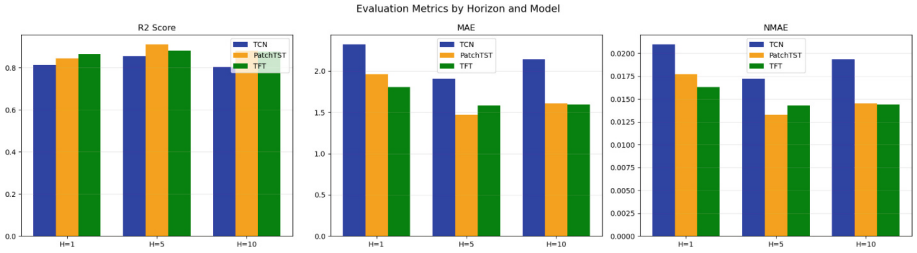

Fig. 6. Quantitative evaluation across all horizons for SBAS vertical displacements.

The SBAS horizontal displacement for $H = 1$, shown in Fig. 7, presents a distinct behavioural pattern relative to the other components, characterised by an initial phase of moderate positive displacements followed by a sharp negative excursion around mid-to-late 2024 and a strong positive recovery trend through 2025. This signal exhibits the highest degree of non-stationarity across all four datasets, with both the direction and rate of displacement reversing markedly within the monitoring period. Despite this complexity, TFT and PatchTST broadly capture the dominant trend transitions, though they tend to underestimate peak values during the most dynamic phases. The TCN, while tracking the general trajectory, shows greater variability in its predictions and occasional phase misalignment during abrupt reversals.

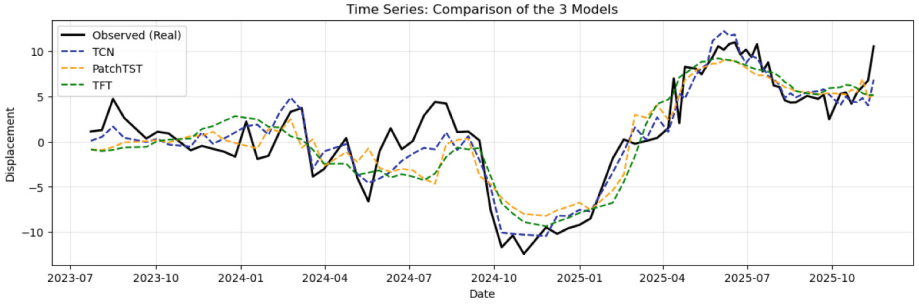

Fig. 7. Time series comparison of observed and predicted SBAS horizontal displacements for $H = 1$.

The evaluation metrics presented in Fig. 8 highlight an important reversal in model rankings for this component: the TCN achieves the highest R^2 scores at all horizons (≈ 0.90 at $H = 1$ and $H = 10$), and the lowest MAE at $H = 1$ (≈ 1.58), while PatchTST and TFT exhibit notably higher errors at the shortest horizon, with MAE values reaching ≈ 2.08. This suggests that the convolutional inductive bias of the TCN may be better suited to capturing the more oscillatory and rapidly reversing dynamics of the SBAS horizontal signal, whereas the attention-based models.

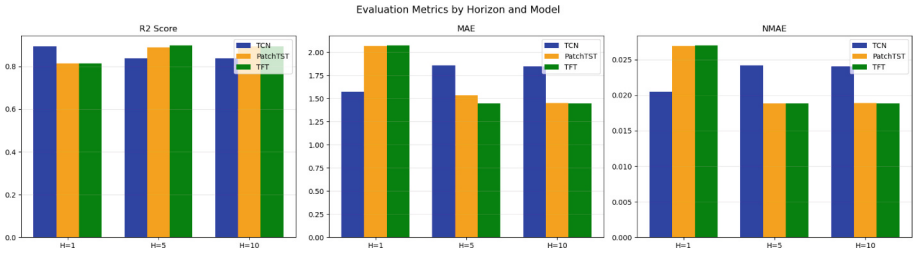

Fig. 8. Quantitative evaluation across all horizons for SBAS horizontal displacements.

5 Conclusions

This study investigated the application of three architectures: TCN, PatchTST and TFT for multi-step forecasting of InSAR-derived slope displacements across PS and SBAS datasets for both vertical and horizontal components. All three models demonstrated reasonable predictive capability, achieving R^2 values consistently above 0.75 across all configurations, with TFT and PatchTST generally outperforming the TCN at medium and long forecasting horizons. The TFT exhibited the strongest overall performance on the PS datasets, while PatchTST achieved the highest R^2 of all models across all configurations on the SBAS vertical component at $H = 5$ (0.91). The TCN, despite being generally less competitive, proved to be the best-performing model on the SBAS horizontal component at $H = 1$, suggesting that its convolutional inductive bias may be advantageous for signals characterised by rapid directional reversals, where the smoothing tendency of attention-based models becomes a limitation. The joint multivariate representation across all monitored points, combined with the inclusion of precipitation as an environmental covariate, represents a methodological contribution relative to purely univariate approaches. However, some limitations should be acknowledged, including the short input window of $K = 4$ that may constrain the models' ability to capture longer-period seasonal or climatic cycles that influence displacement dynamics observations and the preliminary nature of the hyperparameter optimisation. Future work will focus on the integration of additional environmental covariates and on the adoption of more rigorous model validation strategies, along with more exhaustive hyperparameter optimisation and the incorporation of explainability techniques to better assess the generalisation capacity and interpretability of the proposed frameworks.

Acknowledgments. This work was funded by the ANI (Agência Nacional de Inovação) through the R&D project "WISE - Wireless Transportation Infrastructure Safety Evaluation", project number 14383, operation code COMPETE2030-FEDER-00574700, cofunded by the Innovation and Digital Transition Program (COMPETE 2030), by the PORTUGAL 2030 and by the European Union. Also, this work is financed by national funds through FCT–Foundation for Science and Technology, under grant agreement CEECINST/00018/2021/CP2806/CT0006 attributed to the 2nd author. This work was also supported by FCT/MCTES under the R&D Unit Institute for Sustainability and Innovation in Structural Engineering (ISISE), under the references UID/4029/2025 (https://doi.org/10.54499/UID/04029/2025), UID/PRR/04029/2025 (https://doi.org/10.54499/UID/PRR/04029/2025) and UID/PRR2/04029/2025 (https://doi.org/10.54499/UID/PRR2/04029/2025), and under the Associate Laboratory Advanced Production and Intelligent Systems ARISE under reference LA/P/0112/2020 (https://doi.org/10.54499/LA/P/0112/2020).

References

1. Schuster, R.L.: Socioeconomic significance of landslides. In: Turner, A.K., Schuster, R.L. (eds.) Landslides - Investigation and Mitigation. Transportation Research Board Special Report 247, pp. 12–35. National Academy Press, Washington, DC (1996)
2. Petley, D.N.: Landslides and engineered slopes: protecting society through improved understanding. In: Proceedings of the 11th International Symposium on Landslides and 2nd North American Symposium on Landslides, pp. 3–13. CRC Press, Taylor and Francis Group, Boca Raton (2012)
3. Casagli, N., Cigna, F., Bianchini, S., et al.: Landslide mapping and monitoring by using radar and optical remote sensing: examples from the EC-FP7 project SAFER. Remote Sens. Appl. Soc. Envir. **4**, 92–108 (2016). https://doi.org/10.1016/j.rsase.2016.07.001
4. Nava, L., Carraro, E., Reyes-Carmona, C., et al.: Landslide displacement forecasting using deep learning and monitoring data across selected sites. Landslides **20**(10), 2111–2129 (2023). https://doi.org/10.1007/s10346-023-02104-9
5. Ferretti, A., Prati, C., Rocca, F.: Permanent scatterers in SAR interferometry. IEEE Trans. Geosci. Remote Sens. **39**(1), 8–20 (2001)
6. Berardino, P., Fornaro, G., Lanari, R., Sansosti, E.: A new algorithm for surface deformation monitoring based on small baseline differential SAR interferometry. IEEE Trans. Geosci. Remote Sens. **40**(11), 2375–2383 (2002)
7. Novellino, A., Pennington, C., Leeming, K., et al.: Mapping landslides from space: a review. Landslides **21**(5), 1041–1052 (2024). https://doi.org/10.1007/s10346-024-02215-x
8. Ebrahim, K.M.P., Fares, A., Faris, N., Zayed, T.: Exploring time series models for landslide prediction: a literature review. Geoenvironmental Disasters **11**(25), 1–28 (2024). https://doi.org/10.1186/s40677-024-00288-3
9. Lim, B., Arık, S.Ö., Loeff, N., Pfister, T.: Temporal Fusion Transformers for interpretable multi-horizon time series forecasting. Int. J. Forecast. **37**(4), 1748–1764 (2021). https://doi.org/10.1016/j.ijforecast.2021.03.012
10. Nie, Y., Nguyen, N.H., Sinthong, P., Kalagnanam, J.: A Time Series is Worth 64 Words: Long-term Forecasting with Transformers. In: 11th International Conference on Learning Representations, ICLR 2023 (2023). https://doi.org/10.48550/arXiv.2211.14730
11. Bai, s., kolter, j.z., koltun, v.: an empirical Evaluation of Generic Convolutional and Recurrent Networks for Sequence Modeling. arXiv preprint arXiv:1803.01271 (2018). https://doi.org/10.48550/arXiv.1803.01271
12. Mirmazloumi, S.M., Wassie, Y., Nava, L., Cuevas-González, M., Crosetto, M., Monserrat, O.: InSAR time series and LSTM model to support early warning detection tools of ground instabilities: mining site case studies. Bull. Eng. Geol. Environ. **82**, 374 (2023). https://doi.org/10.1007/s10064-023-03388-w
13. Wang, J., et al.: Spatiotemporal mechanism-based Spacetimeformer network for InSAR deformation prediction and identification of retrogressive thaw slumps in the Chumar River Basin. Remote Sens. **16**(11), 1891 (2024). https://doi.org/10.3390/rs16111891
14. Moualla, L., Rucci, A., Naletto, G., Anantrasirichai, N., Da Deppo, V.: Hybrid GIS-transformer approach for forecasting Sentinel-1 displacement time series. Remote Sens. **17**(14), 2382 (2025). https://doi.org/10.3390/rs17142382

15. Owusu-Ansah, D., Tinoco, J., Matos, J., Cabral, R.: Combining slope satellite image analysis and artificial intelligence algorithms for highway service level assessment. In: Proceeding XVIII ECSMGE 2024, pp. 1269–1272. Taylor & Francis, London (2024). https://doi.org/10.1201/9781003431749-230
16. CHRS Data Portal, University of California, Irvine. https://chrsdata.eng.uci.edu/. Accessed 16/Mar/2026