Taylor & Francis
Taylor & Francis Group

GIScience & Remote Sensing

ISSN: 1548-1603 (Print) 1943-7226 (Online) Journal homepage: www.tandfonline.com/journals/tgrs20

Analyzing dominant driving factors of ground deformation in mining area via multimodal feature fusion and explainable modeling

Maoqi Liu , Marzieh Baes , Mahdi Motagh , Qian Song , Yao Sun , Zelong Guo , Tao Li & Sichun Long

To cite this article: Maoqi Liu , Marzieh Baes , Mahdi Motagh , Qian Song , Yao Sun , Zelong Guo , Tao Li & Sichun Long (2026) Analyzing dominant driving factors of ground deformation in mining area via multimodal feature fusion and explainable modeling, GIScience & Remote Sensing, 63:1, 2700121, DOI: 10.1080/15481603.2026.2700121

To link to this article: https://doi.org/10.1080/15481603.2026.2700121

GISCIENCE & REMOTE SENSING
2026, VOL. 63, NO. 1, 2700121
https://doi.org/10.1080/15481603.2026.2700121

Taylor & Francis by informa

METHOD

OPEN ACCESS Check for updates

Analyzing dominant driving factors of ground deformation in mining area via multimodal feature fusion and explainable modeling

Maoqi Liu[a,b,c,d], Marzieh Baes[e], Mahdi Motagh[b,f], Qian Song[b], Yao Sun[g], Zelong Guo[b,h], Tao Li[i] and Sichun Long[a,c,d]

[a]School of Resources & Environment and Safety Engineering, Hunan University of Science and Technology, Xiangtan, China; [b]Department of Geodesy, Section of Remote Sensing and Geoinformatics, GFZ Helmholtz Centre for Geosciences, Potsdam, Germany; [c]School of Earth Sciences and Spatial Information Engineering, Hunan University of Science and Technology, Xiangtan, China; [d]Sanya Institute of Hunan University of Science and Technology, Sanya, Hainan, China; [e]Department of Geosystems, Section of Subsurface Process Modelling, GFZ Helmholtz Centre for Geosciences, Potsdam, Germany; [f]Institute for Photogrammetry and GeoInformation, Leibniz University Hannover, Hannover, Germany; [g]Remote Sensing Technology Institute, German Aerospace Center, Weßling, Germany; [h]School of Remote Sensing Science and Technology, Aerospace Information Technology University, Jinan, China; [i]GNSS Research Center, Wuhan university, Wuhan, China

ABSTRACT

Ground deformation in areas of natural resource extraction results from the interplay of multiple anthropogenic and hydrogeological factors. Identifying the main factors and their relationship to deformation is crucial for mitigating disasters when extreme events happen. Here, we show how multimodal artificial intelligence models can disentangle the effects of dynamic and static features related to deformation in the Rhineland coalfield in Germany. We systematically investigate the synergistic effects of multi-source data by integrating static environmental features, including mining areas, urban areas, fault lines, farmland, and geological distributions, with dynamic features, such as surface deformation time-series and annual groundwater-level observations, and organising them into 12 nested multimodal subsets. Using mainstream deep learning modules, including Convolutional Neural Networks, Graph Convolutional Networks, and Transformer, we build six distinct network architectures and evaluate them under comparable settings through progressive feature fusion and feature ablation. The results show that multimodal fusion consistently improves predictive accuracy compared to a model driven only by deformation data, indicating that static spatial distribution information enhances surface deformation prediction when incorporated as contextual background. Within the controlled feature-space design, feature ablation provides stable channel-level comparisons, whereas SHAP remains a useful complementary post hoc attribution method. At the regional scale, geology and groundwater emerge as the most important predictors of the overall deformation pattern, which is consistent with the interpretation that geological conditions shape the hydrogeological setting and the spatial expression of groundwater fluctuations. At the local scale, fault information becomes the most important predictor within spatially confined high-impact zones, because structural discontinuities concentrate deformation gradients and groundwater-related responses. This distinction between regional background control and local structural triggering provides a data-driven basis for assessing deformation susceptibility and mitigating hazards in mining areas.

ARTICLE HISTORY
Received 18 February 2026
Accepted 3 July 2026

KEYWORDS
Disaster mitigation; susceptibility assessment; interpretability; multimodal model

1. Introduction

Ground surface deformation hazard can lead to damage to buildings, transportation infrastructure, or other public facilities. In addition to deformation caused by natural processes, increasing concern about impacts of human activities has been raised by researchers, particularly those associated with natural resource extraction like mining, drilling for oil and gas and pumping groundwater (Fang et al. 2022; Hager et al. 2021; Herrera-García et al. 2021; Koks et al. 2019). Owing to its extended temporal scale and widespread spatial

CONTACT Sichun Long sclong@hnust.edu.cn
Supplemental data for this article can be accessed online at https://doi.org/10.1080/15481603.2026.2700121.

impact, such deformation adversely affects ecosystems, human health, and socio-economic systems. These impacts not only increase financial burdens on governmental authorities but also pose significant threats to the daily safety of residents in nearby areas (Hudson-Edwards et al. 2024; Sonter et al. 2020). Moreover, previous studies indicate that human activities not only worsen follow-up disasters but also increase the total impact of extreme events (Hill et al. 2024). Such activities, especially natural resource extraction, destabilise the original strata, causing sediment displacement and hydrological disturbances (Gill and Malamud 2017; Hill et al. 2024). These disturbances can, in turn, lead to secondary disasters such as induced seismicity, land subsidence, and flooding (Atkinson, Eaton, and Igonin 2020; Moein et al. 2023). To mitigate or prevent these disasters, it is crucial to clarify the causal relationships between triggering factors and disasters, and accordingly, take steps to minimise their impact scientifically. In regions affected by resource extraction, this requires a comprehensive analysis of the drivers of surface deformation, a clear understanding of deformation patterns, and the development of reliable susceptibility assessment strategies to enhance public safety and reduce disaster-related losses.

Researchers have examined geohazards associated with natural resource extraction, including deformation linked to groundwater withdrawal (Ao et al. 2024; Haghshenas Haghighi and Motagh 2024; Lees and Knight 2024) and surface deformation due to mineral resource extraction (Dong et al. 2025; Huning et al. 2024). Specifically for open-pit mining areas, literature shows that multiple interacting drivers control ground deformation. These are often categorised into: (1) Geological and geotechnical factors, such as lithology, geological structures (e.g. faults, joints), and geomechanical properties (Witkowski et al. 2024; Zhang, Zhang, and Ma 2022); (2) Hydrogeological factors, particularly groundwater level changes and pore pressure alterations caused by large-scale mine dewatering (Moser et al. 2024; Scoular et al. 2022); and (3) Anthropogenic and spatial context, including mining geometry, land-use distribution (e.g. mining areas, urban zones), and related surface activities (Chen et al. 2015; Tao, Zhang, and Du 2024). This complex interaction of static (e.g. geology, land use) and dynamic (e.g. dewatering) factors requires integrative analysis. Consequently, researchers have proposed various approaches tailored to different cases and objectives, ranging from geomechanical modelling (Cai et al. 2023; Chen et al. 2022; Ma et al. 2023) to data-driven methods (Guo et al. 2024; Xiong, Kong, and Song 2024), investigating deformation processes across different sites and objectives. In recent years, the rapid advancement of artificial intelligence has highlighted the powerful inferential capabilities of deep learning. Concurrently, the increasing availability of satellite observations has further propelled data-driven methods in geoscience research, leading to many significant achievements, especially in studies of geological hazards (Dong et al. 2024; Yao et al. 2023; Bountos et al. 2021; Kim, Jung, and Lu 2025; Xing et al. 2024). These data-driven studies commonly rely on deep learning methods to extract spatial features from remote sensing data for deformation prediction.

Despite these advances, current data-driven modelling studies remain constrained, especially in multimodal spatiotemporal regression. Many studies assess a narrow set of feature combinations or a single architecture setting, which makes it difficult to separate the marginal value of adding a modality from architecture-specific effects in regression (Bergen et al. 2019; Youssef et al. 2023). In the context of geohazard research, physics-based models often face limitations in spatial scale and parameter reliability, particularly in numerical simulation methods (Chen et al. 2022). While promising, data-driven approaches also exhibit key limitations. Studies often evaluate a limited range of architectures, and many models do not explicitly account for multimodal inputs and complex spatiotemporal dependencies in geohazard processes, which can reduce stability and generalisation (Peng et al. 2024; Radman, Akhoondzadeh, and Hosseiny 2021; Smith and Majumdar 2020). In addition, available observations often only partially represent multisource drivers and coupled spatiotemporal relationships. For instance, in regression tasks based on Interferometric Synthetic Aperture Radar (InSAR) time-series observations, most studies either directly use the raw displacement data as training inputs (Anantrasirichai et al. 2020; Yao et al. 2023), or select only a subset of high-coherence points to construct the target dataset (Long et al. 2023; Zhou et al. 2025). These practices can miss high-dimensional dependencies and nonlinear interactions that influence deformation patterns, which reduces model reliability. Although deep learning models possess strong inference capabilities in regression tasks, these issues in data-driven modelling also compromise their interpretability, particularly for deep architectures (Shen et al. 2023). When models are trained solely on InSAR observations, it becomes difficult to infer the underlying causal factors of surface deformation or hazard events. In current geohazard research, model interpretability is typically addressed through post-hoc interpretability methods

(Zhu et al. 2025), such as Shapley Additive exPlanations (SHAP) (Davydzenka, Tahmasebi, and Shokri 2024; Jiang et al. 2024), which compute the average marginal contribution of each feature across possible feature combinations and assign corresponding importance scores (Lundberg and Lee 2017). However, most SHAP applications in this domain focus on classification tasks (Bhosale et al. 2024; Ma et al. 2025). Even in those cases where multiple factors are included, the data used are typically unimodal, either dynamic time-series data or static environmental spatial features, thus limiting the method's ability to handle complex multimodal interactions. Relatively few studies in geohazard research have used multimodal datasets for data-driven modelling and interpreted the model, particularly for regression tasks, even though such approaches may mitigate the constraints of single-modality data present in many data-driven methods (Gerard, Zhao, and Sullivan 2023; Ren et al. 2025).

To address these gaps, we developed a controlled multimodal regression framework that integrates InSAR deformation time series with auxiliary hydrogeological and environmental factors to quantify the contribution of each factor. We selected the Rhenish coalfield in North Rhine-Westphalia, Germany, as the study area, which is the largest open-pit coal mining region in Europe. Since the mid-1950s, large-scale lignite extraction has substantially reshaped the regional landscape. Building on surface deformation time series, we constructed datasets that contain different combinations of modal factors and designed a series of deep learning architectures to train these datasets, thereby supporting interpretability analysis in a multimodal regression setting. Using these datasets and deep learning architectures, we evaluated the influence of different factors on surface deformation prediction at the model-cluster level.

To the best of our knowledge, existing research on surface deformation patterns under multimodal factor conditions has rarely analysed the influence of primary driving factors on disasters at the model-cluster level. Here, we quantify factor effects from the model-cluster level by coupling controlled feature space subsets with six deep learning architectures with input-layer adaptations to accommodate different channel combinations. We then perform multimodal regression ablation using the best-performing architecture from comprehensive evaluations, thereby deciphering the global importance of factors. The contributions of this paper are as follows: (1) Progressive feature fusion under comparable inputs demonstrates multimodal gains, and feature ablation quantifies the marginal contribution of each modality to surface deformation prediction. (2) Compared feature ablation with SHAP in a multimodal spatiotemporal regression setting to clarify their distinct roles in global feature interpretation. (3) By converting the global weights derived from ablation into a spatial weighting scheme, we linked the modelling results to surface deformation susceptibility assessment.

2. Study area and data

2.1. Experiment area

The Rhenish coalfield lies in the southwestern part of North Rhine-Westphalia, Germany (Figure 1a). It is the largest lignite mining district in Europe, covering approximately 2,500 km^2, and has long played a central role in regional energy production (Gerwin et al. 2023). The study area includes the three major open-pit mines in this coalfield, Hambach, Garzweiler, and Inden (Figure 1b). The Hambach mine lies between Jülich and Kerpen. Open-pit extraction began there in 1978, and it has since become one of the region's highest-producing and deepest mines. Garzweiler mine has a longer mining history, and lignite extraction in its coalfield extends back for more than a century. Inden mine, which lies between the Düren and Aachen regions, forms the western part of the coalfield's mining system and supplies coal to the nearby Weisweiler power plant. Together, these three mines represent the principal lignite extraction system in the Rhenish coalfield and capture the regional-scale interactions among mining activity, groundwater decline, and surface deformation. Investigating surface deformation jointly controlled by mining activity, groundwater decline, and other factors provides an important scientific basis for local deformation mitigation, monitoring, and susceptibility assessment.

InSAR observations, obtained from the European Ground Motion Service (EGMS), reveal significant surface deformation in the Rhineland coalfield (Ritushree, Baes, and Motagh 2026). The region severely affected by ground deformation includes three main open-pit coal mines - Hambach, Garzweiler, and Inden, as well as areas such as near Bergheim, which previously hosted the now-inactive Bergheim and Fortuna

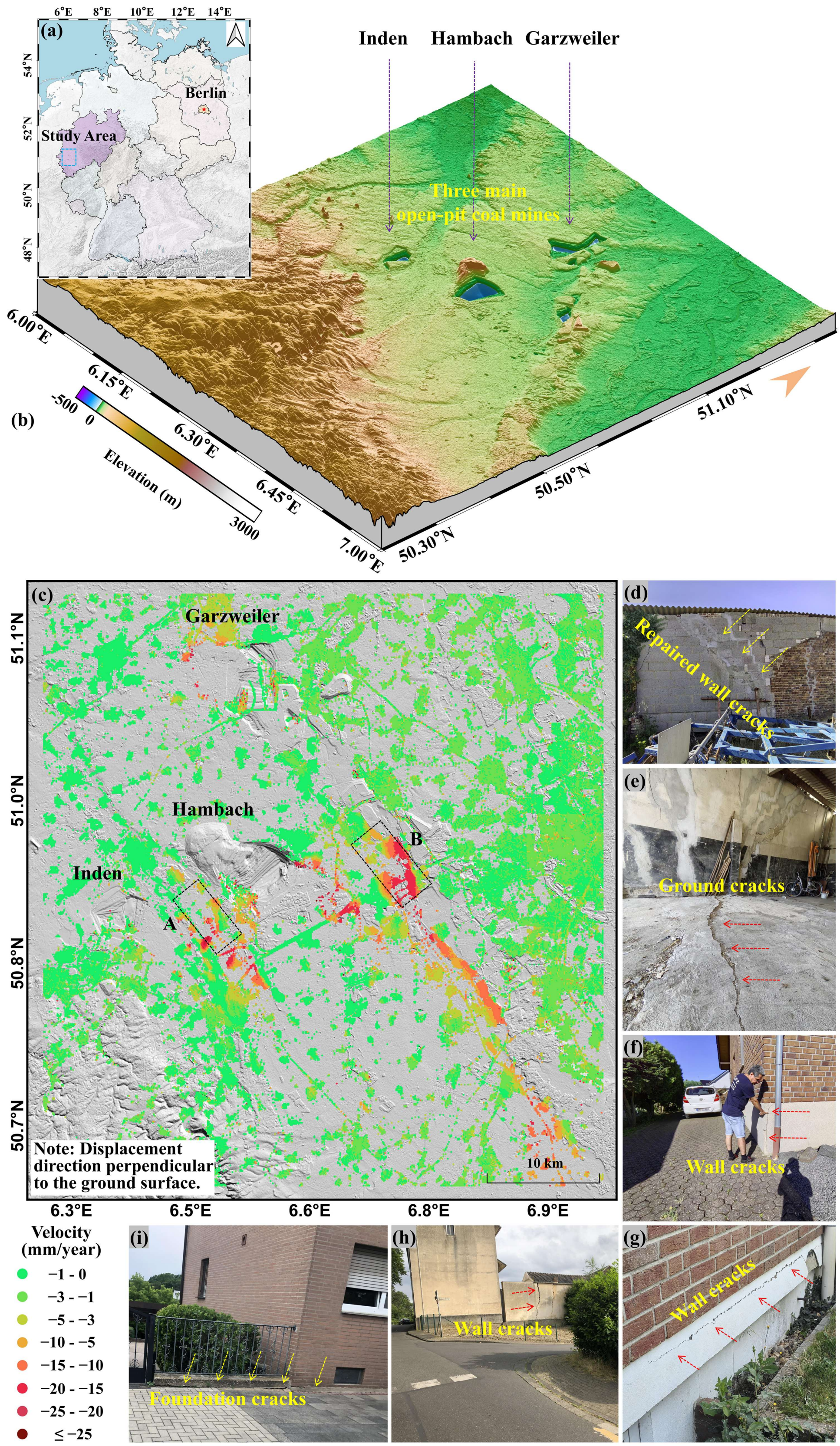

Figure 1. Overview and Deformation of the study area. (a) Geographical location of the study area, and (b) Distribution of the three main active Open-pit mines in the area. (c) InSAR observations from EGMS show the average deformation velocity for 01/2016—12/2021 in this area. (d–i) Exemplary ground pictures of damages in the study areas. (d–g) Corresponds to damages in the region A, and (h-i) to the region B (Photo Credit: Maoqi Liu & Mahdi Motagh, June 2024).

open-pit mines. Field investigations were conducted on buildings and infrastructure in areas A and B to verify the results of the subsequent deformation susceptibility assessment. The results showed that buildings and infrastructure in both areas sustained significant damage, coinciding with areas of substantial subsidence detected by InSAR. Figure 1(d–i) illustrates examples of structural damage observed in the deformation hotspots identified by InSAR. Specifically, (d-g) shows evidence of building damage in the surface deformation zone of a building located east of the Hambach mine (area A in Figure 1c), while (h) and (i) show photographs of damage observed near Bergheim (area B in Figure 1c).

2.2. Data

This dataset includes dynamic observations of surface deformation and groundwater levels, together with static layers that describe geological and anthropogenic factors. Although other processes, such as aquifer geometry, soil thickness, building loads, and tectonic activity, can locally modulate land subsidence under specific hydrogeological and human settings (Navarro-Hernández et al. 2023), data-driven geospatial modelling remains fundamentally constrained by the availability of high-quality, spatially consistent, and observationally supported features that can be integrated uniformly across the entire study area. We therefore restricted the analysis to a subset of factors that simultaneously satisfy three stringent criteria. First, they maintain a well-established first-order mechanical relationship with mining-related surface deformation. Second, they provide complete spatial coverage and sufficient data quality across the whole study area. Third, they can be accurately co-registered with the InSAR observations on the same spatiotemporal grid. This selection prioritises the primary controlling factors that are directly observable and suitable for robust regional-scale modelling. Table 1 provides a detailed summary of the dataset, including its data sources, resolution, and accuracy specifications.

2.2.1. Dynamic data

In this study, surface deformation refers to land-surface displacement, and we represent it using the vertical displacement component provided by the EGMS Level-3 Ortho product. We therefore use InSAR-derived vertical displacement observations as the observational representation of surface deformation in the study area, rather than treating raw SAR measurements as the deformation variable itself. Unlike single-geometry Line-of-Sight (LOS) measurements, which mix vertical and horizontal motion, the Ortho product estimates the vertical component by combining calibrated ascending and descending Sentinel-1 time series referenced to a geodetic GNSS model. We use this vertical component as the target variable because it directly represents the observed subsidence and uplift relevant to the present analysis. The EGMS validation report indicates a velocity precision of better than 1 mm/year for the vertical component (Costantini et al. 2021; Crosetto et al. 2026).

We collected groundwater-level time-series data from 292 monitoring stations in the study area, covering the period from 1968 to 2021. To align as closely as possible with the observation period for groundwater levels, we selected the InSAR observation window from 2016 to 2021 on EGMS. For the observation period, the coefficient of variation (CV) was computed at each station to quantify the spatial heterogeneity of groundwater fluctuation across the region (see Figure 2a). The results indicate moderate variability in groundwater levels in parts of the study area.

Table 1. Summary of dynamic and static datasets used in this study and hydrogeological processes.

Category	Dataset	Source	Time Period	Time/Spatial Resolution	Accuracy/Note
Dynamic	Surface Displacement (Vertical Component)	EGMS Level-3 Ortho	01/ 2016–12/2021	6 days/100 m	Velocity precision < 1 mm/yr (Vertical).
	Groundwater Levels	Erftverband Company upon request	01/ 1968–12/2021	30 days/292 Monitoring Stations	Interpolated to 100 m grid.
Static	Mining/Urban/Agriculture	OpenStreetMap	2024	\	Approximately 5 m (Mean), (Marco et al. 2012)
	Geological Formations	GD NRW	Static	\	\
	Fault Lines	GD NRW	Static	\	\

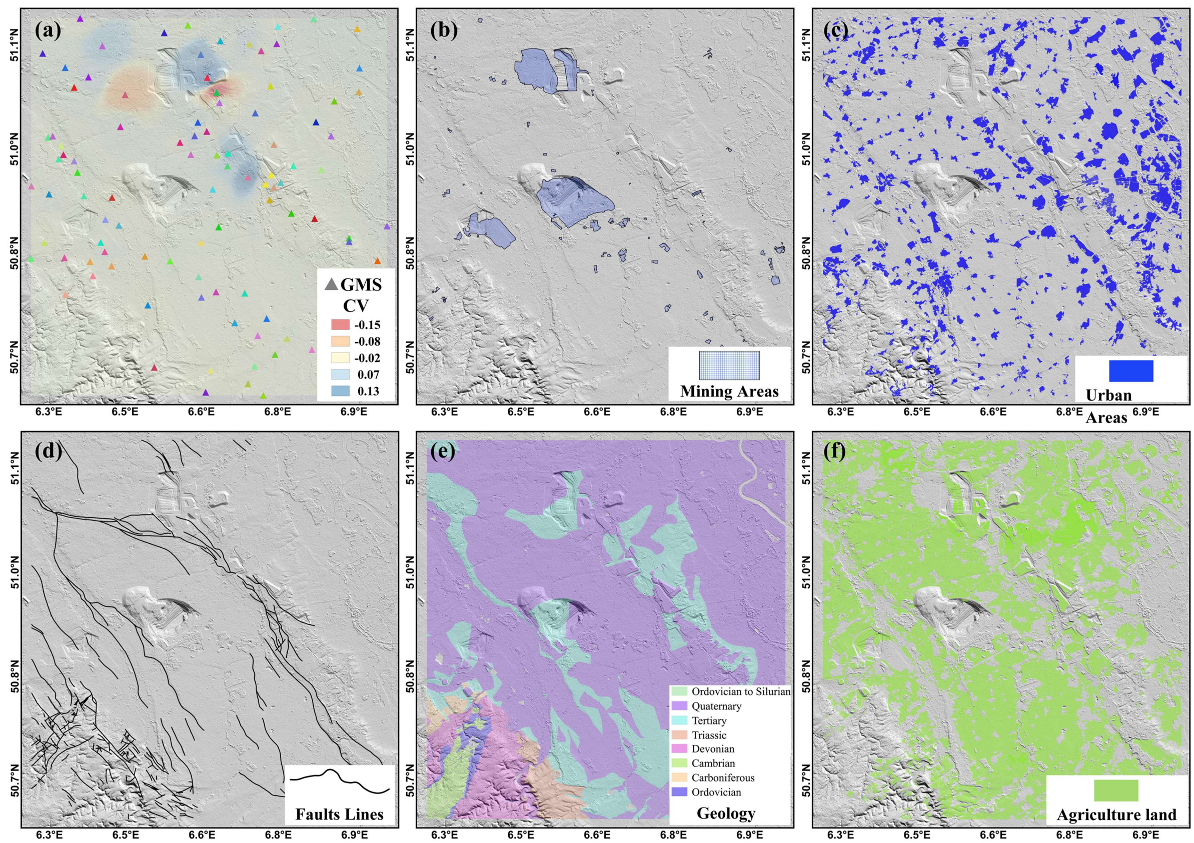

Figure 2. Data overview. (a) Groundwater level variability, 2016 to 2021. Figure (a) shows the location of the groundwater level monitoring station (GMS) and the coefficient of variation (CV), which is used to quantify spatial heterogeneity in groundwater fluctuation across the region. A larger absolute value of CV corresponds to more intense fluctuations. (b) Mining area distribution, (c) Urban distribution, the blue part represents urban area. (d) Fault lines, (e) Geological formations, (f) Agricultural land distribution, the green part represents agricultural land.

2.2.2. Static data

We included five static spatial factors that influence surface stability in the study area: mining areas, urban settlements, agricultural land, mapped faults, and geological formations. We extracted boundaries for mining, urban, and agricultural areas from OpenStreetMap due to its high update frequency, which ensures an accurate depiction of the current surface features in our study area (see Figure 2b, c, and f). We obtained geological and fault data from the Geological Survey of North Rhine - Westphalia (GD NRW), Germany (Geological Survey of North Rhine - Westphalia 2015) and rasterised it to 100 m for modelling (see Figure 2b and e). All static layers were unified in the EGMS coordinate system and grid resolution to facilitate subsequent modelling work.

3. Methodology

3.1. Spatiotemporal dataset construction

We assembled a unified multimodal spatiotemporal data stack by integrating EGMS vertical-displacement time series with groundwater observations and static spatial layers representing environmental, anthropogenic, and geological factors. For the InSAR-derived deformation data, we used the valid EGMS observation points at each acquisition date as the sampling locations of surface deformation. These observations provided the temporal records of the deformation field for subsequent modelling. Groundwater data collected from multiple monitoring stations provided time series that reflect subsurface hydrodynamic changes. Static spatial layers, including fault lines, urban areas, mining areas, agricultural land, and geological formations, were incorporated to evaluate the contribution of each auxiliary factor to surface deformation prediction. Based on these data and the research objectives, we developed a spatiotemporal harmonisation preprocessing framework to ensure temporal and spatial alignment across all modalities, while preserving key information required for subsequent modelling, including the high-coherence point mask, channel indices, and feature-specific metadata for each input channel.

The workflow of data processing framework we used here is illustrated in Figure 3. To unify multi-source data onto a common spatiotemporal grid, the time window of the InSAR data (January 2016 to December 2021) was used as the reference temporal framework. Then groundwater level time series were temporally interpolated using bidirectional linear interpolation (Arisoy et al. 2015): original observations were treated as anchor points and linear interpolation was performed both forward and backward between each pair of adjacent anchors. This approach ensured that endpoint values remained consistent with original measurements while achieving smooth transitions in between.

Subsequently, we interpolated the InSAR observations and groundwater measurements separately onto a common 100 m grid. We defined this grid to match the spatial resolution of the EGMS Level-3 Ortho product and to ensure spatial consistency during multimodal data fusion. For the InSAR channel, we applied the inverse distance weighting (IDW) method (Plésiat et al. 2024) to the valid EGMS observation points at each acquisition date to generate gridded vertical-displacement fields over the study area. For the groundwater channel, we applied the same method to measurements from the 292 monitoring wells. We then resampled the static spatial layers, including fault lines, geological units, mining areas, urban areas, and agricultural land, to the same grid. This procedure produced spatially aligned and temporally synchronised dynamic and static inputs for subsequent modelling.

Building on the dataset fusion framework, a controlled feature-space experimental design was implemented to construct a multimodal feature space combining the InSAR data (displacement channel - D) with six auxiliary features: groundwater level (W), mining area distribution (M), urban distribution (B), fault distribution (F), geological distribution (G), and agricultural land distribution (A). Twelve feature subset combinations were designed (12 datasets). The first seven (D, DW, DWM, DWMB, DWMBF, DWMBFG,

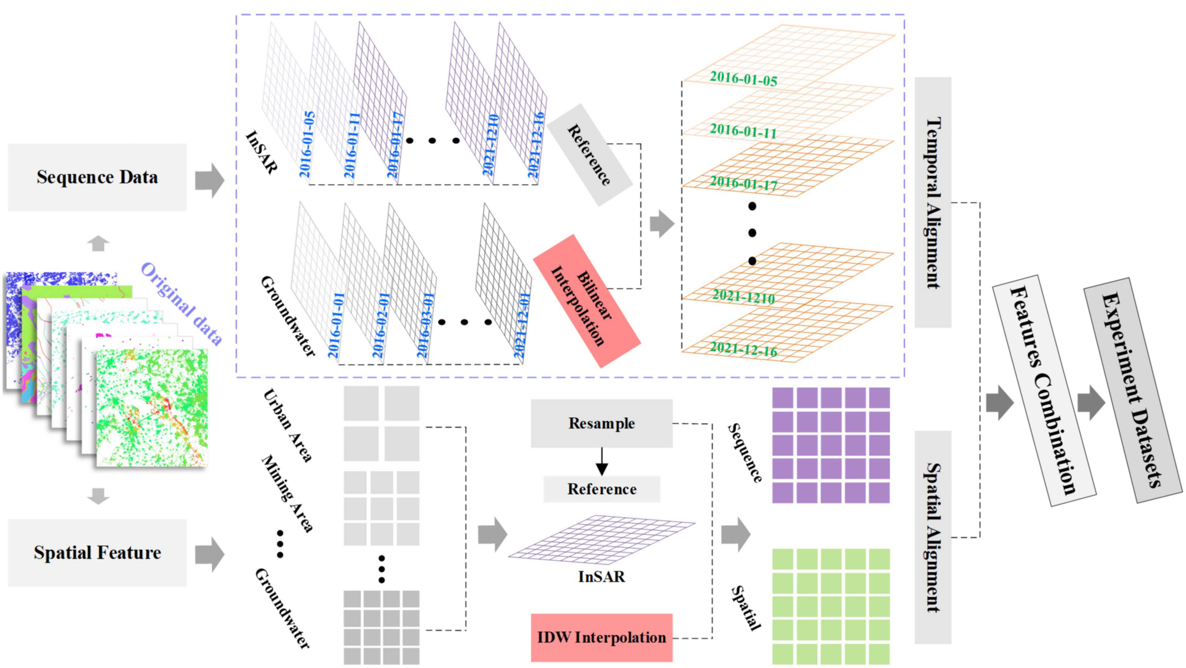

Figure 3. Overview of the data processing workflow. The workflow begins by categorising the raw data into two types: sequence data (including InSAR and groundwater time series) and spatial features (which include static environmental and geological layers and one reference frame from the sequence data). We first perform temporal alignment (TA) on the sequence data. Using the InSAR time series as the temporal reference, bidirectional linear interpolation is applied to align groundwater observations with the InSAR timeline. Next, we construct the spatial feature layers, which integrate both spatial and sequence-derived inputs. We then apply spatial alignment (SA) using inverse distance weighting (IDW), with the InSAR spatial resolution serving as the reference.

DWMBFGA) were used in progressive feature fusion experiments to evaluate how multimodal data synergistically enhance deformation modelling. The remaining five subsets (DMBFGA, DWBFGA, DWMFGA, DWMBGA, DWMBFA), along with DWMBFG from the fusion experiments, were used in feature ablation experiments to quantify the model's sensitivity to each of the auxiliary factors. This nested design isolates incremental modality value under strict subset comparability, which supports a direct comparison between fusion gains and ablation-based attributions. The two subsequent sections provide detailed descriptions of model architecture and the progressive feature fusion and feature ablation experiments used in this study.

3.2. Model architecture

Surface deformation is typically controlled by complex nonlinear relationships and pronounced spatiotemporal dependencies involving human activities, hydrogeological conditions, and the environmental and geological background. Previous studies show that spatial dependence and spatial heterogeneity are fundamental properties of geospatial phenomena, and that spatial context itself carries essential information, especially in the modelling of environmental and Earth observation data, where it cannot be ignored or excessively compressed (Goodchild and Li 2021; Koldasbayeva et al. 2024). When traditional machine learning methods are used for spatial field prediction, researchers usually transform the original spatial fields into statistical measures through feature engineering before applying regression or classification models. Although this strategy provides compact inputs, it often weakens within-patch spatial configuration, neighbourhood relationships, and multiscale spatial context, thereby limiting the model's ability to represent complex spatiotemporal structure (Reichstein, Camps-Valls, and Stevens 2019; Shen 2018). By contrast, deep learning allows spatial context on regular grids to enter the model directly as input features and can automatically learn hierarchical spatial patterns and temporal dependencies from spatiotemporal observations. It is therefore better suited to mining multimodal information under complex spatiotemporal structure, particularly for image-like or field-like data in remote sensing monitoring scenarios (Li et al. 2020; Reichstein, Camps-Valls, and Stevens 2019; Yu and Ma 2021). To characterise the spatiotemporal interactions among variables in the multimodal dataset and quantify the influence of different spatiotemporal factors on surface deformation prediction, we used deep learning to model the relationships between dynamic spatiotemporal features, such as surface deformation and groundwater level, and static spatial distributions.

Given that multimodal datasets contain both temporal sequences of deformation and spatially distributed features, deep learning-based regression tasks allow for modular architecture design, where spatial feature extractors and temporal dependency models can be integrated to jointly capture spatiotemporal patterns. In this study, surface deformation prediction was formulated as a regression task, with deep learning models trained to learn the mapping between multimodal inputs and future deformation values. In implementation, the multi-channel data at each time step were treated as a "multi-channel image". To effectively model these data, we employed foundational convolutional architectures (e.g. CNN (Alzubaidi et al. 2021) and ResNet (He et al. 2016), graph convolutional networks (GCN) (Li et al. 2019; Zhang, Cui, and Zhu 2020), and temporal modelling modules including LSTM (Greff et al. 2016) and Transformer (Vaswani et al. 2017). Building upon these components, we built six deep learning architectures of increasing complexity: CNN+LSTM, CNN+Transformer, GCN+LSTM, GCN+Transformer, ResNet+LSTM, and ResNet +Transformer. These architectures are trained on the datasets, with CNN, ResNet, and GCN components designed to extract the spatial features at each time point, while LSTM and Transformer modules capture temporal dependencies. All models were optimised to support variable numbers of input channels at the input layer.

The above architectures were first individually trained on the initial seven datasets described in Section 3.1 to validate the synergistic enhancement effect of multimodal data and to identify the best-performing architecture. Subsequently, the best-performing architecture was trained on the remaining five datasets to evaluate feature importance under the feature ablation framework. To ensure that the time-series data division and normalisation adhered to existing studies (Ismail Fawaz et al. 2019; Lim and Zohren 2021) and avoided data leakage during training, the datasets were divided strictly in chronological order. The training set accounted for 80% of the data (from 5 January 2016 to 4 October 2020), with the validation and test sets each comprising 10%. Normalisation parameters (channel-wise mean and standard deviation) were calculated solely from the training set and applied to the whole dataset. The validation set was used for tuning hyperparameters of the model during training, while the test set was an independent set for evaluating the trained model.

Moreover, several less complex models were analysed, encompassing Ridge Regression, Random Forest Regression, and Gradient Boosting Regression (see S1 in the supplementary material) to provide a comprehensive overview of the experimental findings. However, as our main objective was to preserve intra-patch spatial information for multimodal prediction and the subsequent ablation analysis, we focused the main experiments on deep learning architectures.

3.3. Progressive feature fusion and feature ablation strategies

To quantify the contribution of environmental and geological factors embedded in multimodal data to surface deformation prediction, this study proposed a dual-track experimental framework that combines progressive feature fusion with feature ablation. We also used statistical significance testing to improve the robustness and interpretability of the feature-importance analysis. Progressive fusion tests whether additional modalities yield consistent performance gains, and feature ablation measures the performance drop after removing one auxiliary factor from a comparable input setting. For each frame of the multimodal dataset, static spatial features (including environmental and geological factors) are concatenated with dynamic deformation signals along the channel dimension and jointly fed into the feature extraction network. Figure 4a illustrates this process using the ResNet+Transformer architecture as an example, and the same principle can also be applied to CNN- or GCN-based designs. During convolution, local spatial patterns are extracted while simultaneously perceiving both the static background and dynamic variations at the same location, thereby learning fused representations of static and dynamic information. The Transformer then processes these representations for temporal modelling. Based on this, all architectures were used to conduct feature fusion and feature ablation experiments (see Figure 4b, c).

The progressive feature fusion strategy is designed to incrementally incorporate auxiliary factors on top of the InSAR channel, which serves as the main factor. As shown in Figure 4b, all 6 architectures described in Section 3.2 were trained and evaluated on datasets ranging from one to seven channels using a progressive feature fusion strategy to investigate the relationship between model performance and the number of input features. To evaluate model performance, five evaluation metrics were chosen: Root Mean Squared Error (RMSE), Relative Root Mean Squared Error (RRMSE), Huber Robust Loss Function (Huber_Loss), Median Absolute Error (Median_AE), and the Coefficient of Determination (R^2). If the model performance exhibited a statistically significant positive correlation with an increasing number of channels, it could support that the progressive fusion strategy effectively captures the synergistic enhancement effects of multi-source data. Additionally, this outcome indirectly highlights the contribution of static spatial information to surface deformation prediction.

Although SHAP provides an intuitive and mathematically grounded framework for post hoc feature attribution, its global summaries can be highly sensitive to the aggregation rule when applied to multimodal spatiotemporal regression. In our setting, static spatial layers were repeated across all time steps, so temporal aggregation may dilute their contributions and shift the global ranking toward dynamic variables. In addition, static factors may interact with dynamic variables such as groundwater during prediction. When SHAP values are aggregated across space and time, the attributions assigned to these related channels may depend on feature interactions and aggregation choices, which can reduce the stability of global rankings (Bilodeau et al. 2024; Krell et al. 2023). Feature ablation provides a complementary, performance-based comparison. It estimates the change in predictive performance after removing one auxiliary factor while keeping the other training settings unchanged, which allows this method to summarise model dependence at the global channel level under the current design. We therefore used SHAP as a complementary attribution analysis and treated feature ablation as the primary tool for global channel-level comparison under controlled feature subsets. The feature ablation strategy uses the best-performing model, selected from 42 candidates trained across seven datasets and six architectures, as the reference model. We denote its performance as E_c, where c denotes the channel configuration. We then compare this reference performance with that of comparable models retrained with fewer data channels. This is done in a

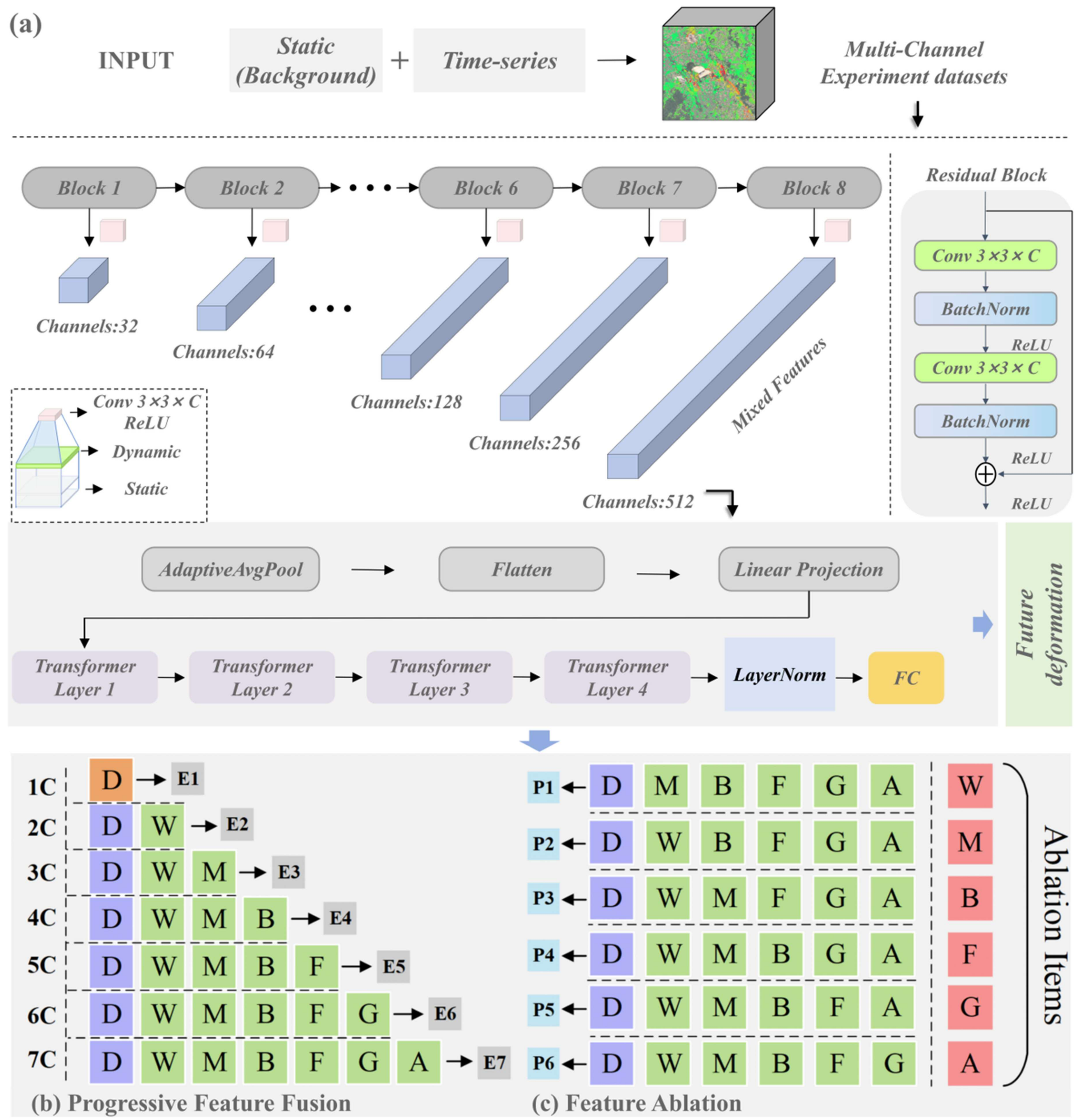

Figure 4. (a) ResNet+Transformer architecture static feature processing, (b) Progressive feature fusion, (c) feature ablation strategies. The pink rectangle in (a) represents the convolution in the residual block, and the blue rectangle represents the features obtained by the residual block. The change in their length indicates the gradual increase in the dimension of the extracted features. The figure in the dashed box shows that the convolution simultaneously encodes dynamic deformation information and spatial features. The final features are linearly projected and then fed into the Transformer layer for time series modelling. In (b) and (c) E1–E7 and P1–P6 represent the experimental evaluation metrics for the progressive feature fusion and the feature ablation, respectively. C1–C7 represent the dataset channel numbers. The abbreviations D, W, M, B, F, G, and A refer to input features: D = deformation (InSAR data), W = groundwater, M = mining areas, B = buildings (urban areas), F = fault lines, G = geology, and A = agricultural areas. In progressive feature fusion, features are added incrementally to assess their cumulative impact on model performance. In the feature ablation experiment, one auxiliary feature channel was removed from the 7-channel input in each experiment, while the remaining channels remained unchanged. The model was then retrained under the same training settings to evaluate the marginal contribution of the feature to the prediction performance. The displacement channel D was always retained, and the 1-channel model was used as the baseline.

sequential removal manner, where one specific (i-th) dataset is removed at one time (its model performance denoted as P_i) (Figure 4c). We define the marginal performance drop caused by removing feature i as:

$$\Delta_i = |E_c - P_i| \tag{1}$$

where a larger Δ_i value indicates greater importance of feature i to the model's prediction. To obtain a stable global ranking under multiple metrics, we repeated each ablation experiment R times and extended Equation (1) to Kevaluation metrics. Let $S_j^{(r)}$ denote the j-th metric for the full model at repeat r, and let $S_{i,j}^{(r)}$ denote the same metric after removing feature i. We first compute the mean performance drop for each feature and each metric:

$$\Delta_{i,j} = |\bar{S}_{i,j} - \bar{S}_j|, \quad \bar{S}_j = \frac{1}{R}\sum_{r=1}^{R} S_j^{(r)}, \quad \bar{S}_{i,j} = \frac{1}{R}\sum_{r=1}^{R} S_{i,j}^{(r)} \tag{2}$$

We then normalise the metric specific drops across features:

$$W_{i,j} = \frac{\Delta_{i,j}}{\sum_{i=1}^{m}\Delta_{i,j}} \tag{3}$$

Finally, we aggregate the normalised weights across metrics to obtain a single global importance score:

$$w_i = \frac{1}{K}\sum_{j=1}^{K} W_{i,j} \tag{4}$$

To avoid the interference of multicollinearity, a variance inflation factor (VIF) test is conducted to ensure that there is no multicollinearity among all experimental factors. This reduces the risk that the ablation-based importance scores are dominated by strong linear collinearity among inputs and improves the stability of the comparative interpretation. This study unifies multimodal fusion and interpretability within an experimental framework centred on a controlled feature space. We constructed twelve nested feature subsets and compared their training performance across six architectural classes. Progressive fusion validated synergistic gains, while ablation experiments quantified the marginal contributions of individual factors. Statistical tests and collinearity control ensured the reliability of attribution. Figure 5 provides an overview of this process.

4. Experiments and results

4.1. Results of progressive feature fusion

We trained models using the six different deep learning architectures (i.e. CNN+LSTM, CNN+Transformer, GCN+LSTM, GCN+Transformer, ResNet+LSTM, and ResNet+Transformer) on the 7 datasets described in Section 3.1. All the datasets were divided into a training set (80%), a validation set (10%), and a test set (10%). Table S2 (See the supplementary materials) provides detailed parameter configurations for model training.

To ensure the reliability and comparability of the results, each model architecture is assigned a fixed set of hyperparameters for the progressive feature fusion experiments. To each architecture, the same configuration is applied consistently across all seven datasets with incrementally added features. Each training scenario is retrained 10 times under identical training conditions (e.g. weight initialisation, optimiser settings), to mitigate the influence of randomness and enhance statistical robustness. Figure 6 shows the performance of CNN+LSTM, ResNet+Transformer, and GCN+Transformer trained on different datasets (the remaining experimental results are shown in Figure S1 in the supplementary materials). The performance of all architectures improves with the increase in the number of datasets, indicating that combining environmental and geological factors with time-varying data can improve the accuracy of regional surface deformation time-series prediction. While some minor fluctuations occur between certain datasets, the overall trend shows a decrease in RMSE. Among the architectures, CNN+LSTM and GCN +Transformer exhibited the most consistent performance, with narrow 95% confidence intervals (see Figure 6). While GCN+LSTM also had tight confidence intervals, ranging from 4C to 6C, its model performance declined. Thus, we opted for CNN+LSTM and GCN+Transformer in the subsequent analysis.

We used paired t-tests and Spearman rank correlation coefficients to assess the relationship between model performance and the number of input channels. The paired t-tests examined local differences between adjacent channel configurations, whereas the Spearman analysis evaluated the overall monotonic trend as the number of channels increased. Supplementary Tables S3 and S4 report the paired t-test

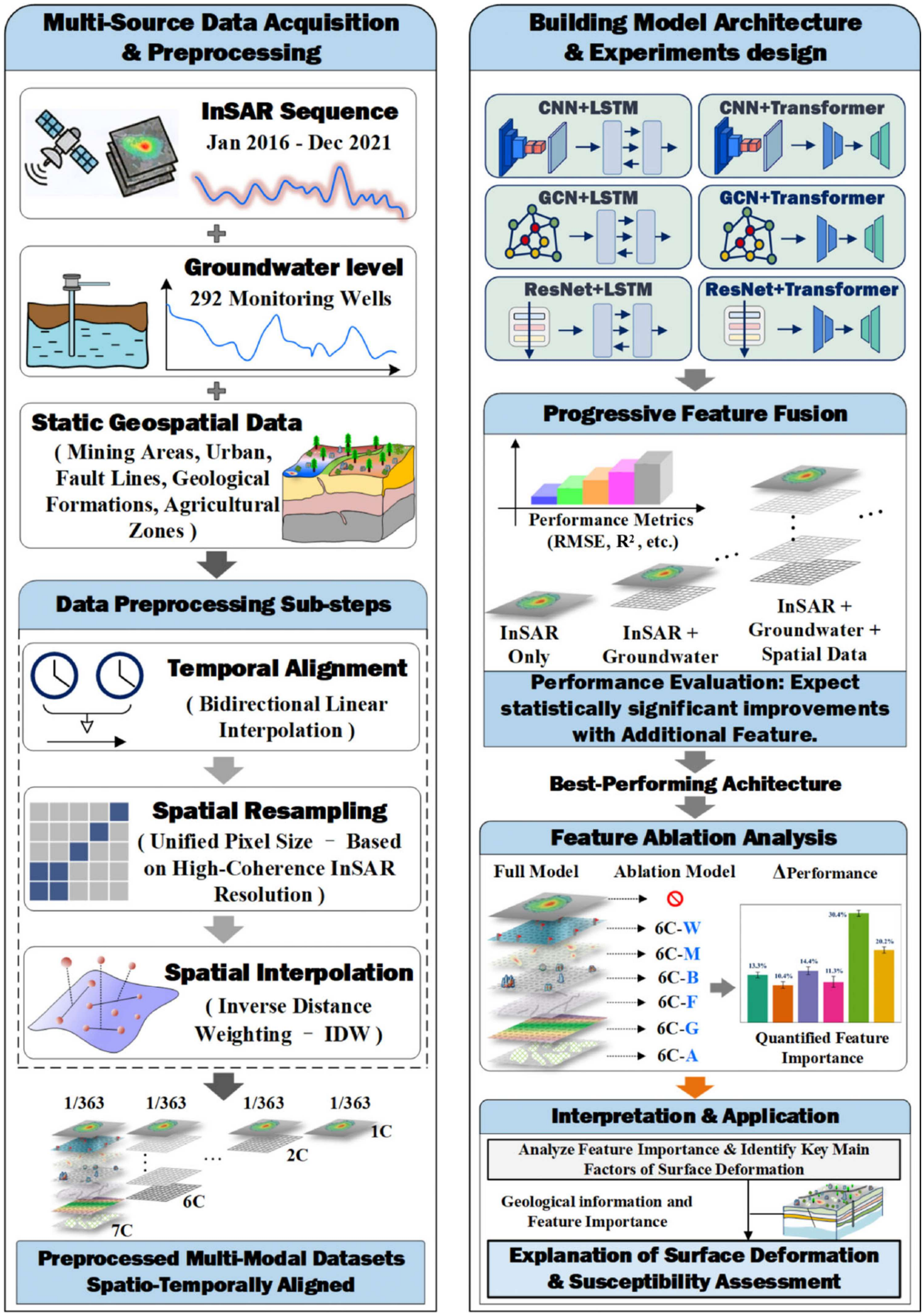

Figure 5. Methodology flowchart. Data is first collected and preprocessed, as illustrated in the left section of the diagram, with each step depicted schematically. This preprocessing yields a designed multimodal dataset. Subsequently, experiments are conducted using the constructed framework, as depicted in the right-hand section. This includes feature fusion experiments, ablation studies, feature importance assessments, and ultimately integrates geological information analysis to evaluate rationality. Feature importance will be utilised for deformation susceptibility assessment. Schematic diagrams are provided for each step to facilitate comprehension.

statistics, corresponding *p*-values, and paired effect sizes for comparisons between adjacent channel configurations, as well as the Spearman rank correlation results for the overall trend. For CNN+LSTM and GCN+Transformer, the paired t-tests showed statistically significant differences between most adjacent channel groups, except for the transition from four to five channels in GCN+Transformer. Figure 7 presents

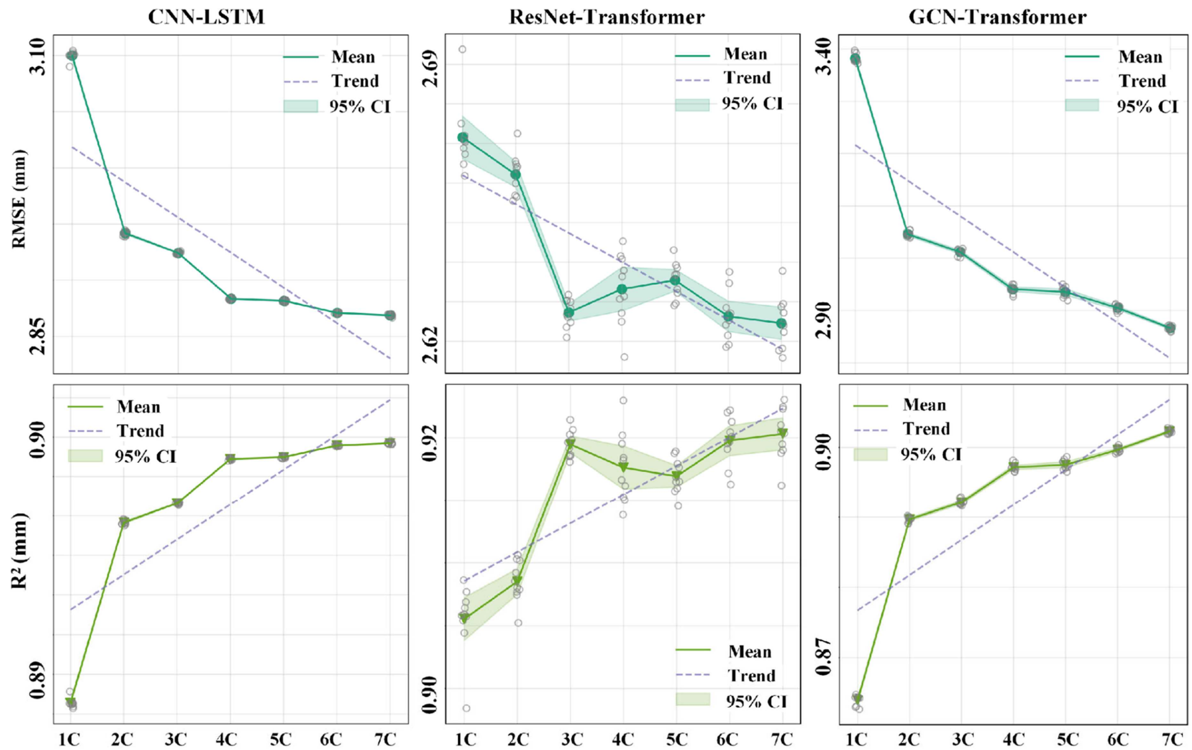

Figure 6. Model performance evaluation for progressive feature fusion experiment. The rows and columns represent model performance for different architectures and evaluation metrics, respectively. 1C, 2C, ..., and 7C refer to datasets with increasing numbers of input channels, corresponding to the following combinations: D, DW, DWM, DWMB, DWMBF, DWMBFG, DWMBFGA. In each subplot, the mean (solid line), trend (dashed line) and Confidence interval (CI—shading area) are derived from 10 repeated training runs of each experiment.

the paired t-test results based on R^2. Overall, these results support an increase in model performance with the number of modalities, although the marginal gains vary across architectures and channel transitions.

Beyond statistical significance, the paired effect sizes reported in Table S3 help quantify the magnitude of performance changes between adjacent channel configurations, whereas the absolute differences on the original metric scale indicate their practical relevance. For CNN+LSTM and GCN+Transformer, most adjacent transitions were associated with clear performance gains, although the magnitude of these gains varied across stages. On the original metric scale, the cumulative improvement from 1C to 7C was substantial, particularly for GCN+Transformer. For CNN+LSTM, RMSE decreased by 7.46%, and GCN+Transformer RMSE decreased by 15.25%. These results indicate that the gains from multimodal fusion were not only statistically significant but also supported by effect size estimates and practically meaningful improvements on the original evaluation scale. By contrast, the transition from 4C (DWMB) to 5C (DWMBF), which introduced the fault channel, reduced RMSE by only about 0.18% for GCN+Transformer, and the other evaluation metrics showed a more mixed pattern. Under the current configuration, CNN+LSTM still showed a significant improvement over the same transition, whereas the fault channel provided only limited incremental benefit for GCN+Transformer at this stage. This pattern suggests that the additional information carried by the fault channel can have practical value for improving predictive accuracy, but that value depends on the architecture.

The Spearman correlation analysis further reveals the overall trend: although some models do not exhibit statistically significant performance improvements between adjacent channel configurations in *t*-test, the global Spearman test results across evaluation metrics consistently show that model performance improves as the amount of input information increases ($p < 0.001$, Table S3). Compared to CNN+LSTM, where all metrics yield Spearman coefficients in the range of $0.884 < |\rho| < 0.988$ ($p < 0.001$), the GCN+Transformer architecture demonstrates smaller fluctuations across all evaluation metrics ($0.941 < |\rho| < 0.977$, $p < 0.001$), indicating greater

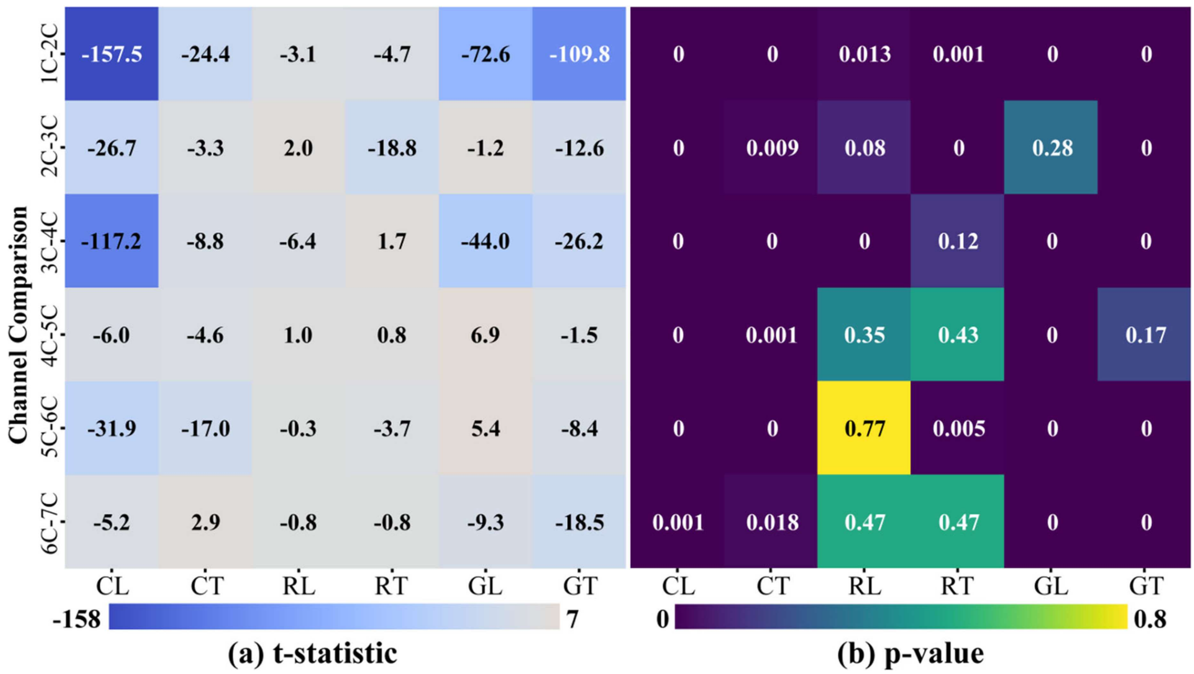

Figure 7. Validation of progressive feature fusion results. (a) t-statistic, (b) *p*-value. The x-axis in (a) and (b) refers to different architectures - CL, CT, RL, RT, GL, and GT stand for CNN+LSTM, CNN+Transformer, ResNet+LSTM, ResNet+Transformer, GCN +LSTM, and GCN+Transformer, respectively. Y axis represents various adjacent channel combinations. 1C, 2C, ..., and 7C have the same meaning as the explanation in Figure 6. For the t-statistic value related to R^2, $t < 0$ indicates that the high channel's model performance is better than the low channel's model; conversely, if $t > 0$, the reverse is true.

robustness in performance. In addition, GCN+Transformer achieves a fourfold improvement in training efficiency compared to CNN+LSTM.

The aggregate statistical results confirm the effectiveness of the progressive feature fusion strategy in surface deformation prediction (Bodnar et al. 2025). As the number of input channels increases from 1 to 7, the model's error metrics decrease significantly, and the coefficient of determination (R^2) increases markedly. Although minor performance fluctuations emerge between certain adjacent channel configurations in architectures such as CNN+Transformer, ResNet+LSTM, ResNet+Transformer, and GCN+LSTM, the overall results strongly support the synergistic benefits of multimodal data fusion. Finally, given the large-scale nature of the dataset and the planned implementation of extensive hyperparameter tuning and feature ablation experiments, computational efficiency emerges as a critical practical consideration. Taking into account model performance, sensitivity to additional information, potential for further optimisation, and computational scalability, the GCN+Transformer architecture is selected as the optimal model for this study.

4.2. Global feature importance results based on feature ablation study

This study proposed the use of a feature ablation strategy to quantify feature importance by comparing the performance of the best-performing model, based on GCN+Transformer architecture, with all 7 channels to that of models where individual features were removed. This approach preserves the underlying coupling among features, enabling a more accurate assessment of each feature's independent contribution. To ensure the validity of the ablation experiment results, we further evaluated the degree of multicollinearity among features, thereby minimising potential bias arising from feature redundancy or linear dependence. Specifically, the Variance Inflation Factor (VIF) was calculated for each feature to identify possible high-collinearity issues. A VIF value below 4 is generally considered indicative of no multicollinearity (Marta et al. 2023; Seidler et al. 2025).

In Figure 8a, the diagonal of the lower triangle displays the VIF values for each feature, while the off-diagonal entries represent pairwise correlation coefficients. All VIF values are well below the threshold (all < 2), indicating an absence of serious multicollinearity. Additionally, the correlation coefficients among features do not reveal strong linear dependencies. Together, these results confirm that the features maintain a high degree of relative independence within the model. The low multicollinearity ensures that feature importance can be reliably assessed under the ablation strategy, thereby enhancing the statistical credibility of the evaluation.

To evaluate global feature importance, we applied a feature ablation strategy using the GCN+Transformer architecture, which was identified as the best-performing architecture in Section 4.1. Specifically, for the complete 7-channel dataset, we systematically remove one auxiliary feature at a time, retrain the model, and evaluate its performance. To ensure robustness, we repeated each experiment ten times. We compute the global importance weights using the metric-specific performance drop, normalisation, and multi-metric aggregation defined in Section 3.3, (1) to (4). As shown in Figure 8b, the weights of the auxiliary features are: groundwater (W) 13.3%, mining area distribution (M) 10.4%, urban distribution (B) 14.4%, fault distribution (F) 11.3%, geological distribution (G) 30.4%, and agricultural land distribution (A) 20.2%. Among these, geological distribution (G) causes the largest decrease in model performance when ablated, indicating that it is the most influential auxiliary feature in this predictive setting.

A particularly noteworthy and perhaps counterintuitive finding is that static geological distribution (G) emerged as more important than the dynamic groundwater factor (W). This result does not conflict with the hydrogeological interpretation that groundwater variation is closely linked to subsidence. Instead, it highlights the critical influence of subsurface geological conditions on the overall deformation process (Motagh et al. 2017). The model does not treat features as isolated linear predictors. Rather, the Graph Convolutional Network (GCN) component first learns to interpret static features (such as G, M, and F) as a foundational background map that defines the local physical properties and susceptibility of the terrain. Subsequently, the model learns a conditional dynamic function in which the influence of groundwater variation (W) on subsidence (D) is largely modulated by these local geological conditions. Section 5.2 further discusses the reliability of this finding.

4.3. Model noise and temporal-gap sensitivity tests

To improve the reliability of the results and the subsequent analyses, we performed noise and temporal-gap sensitivity analyses for all models presented in this section. For noise sensitivity, we added Gaussian noise to the test data (Hendrycks and Dietterich 2019). For temporal-gap sensitivity, we applied a continuous block-masking strategy with end-value imputation (Che et al. 2018; Nie 2022). We restricted the sensitivity analysis to the dynamic variables because these channels contain measurement noise or temporal gaps and therefore dominate the random uncertainty in the multimodal inputs. By contrast,

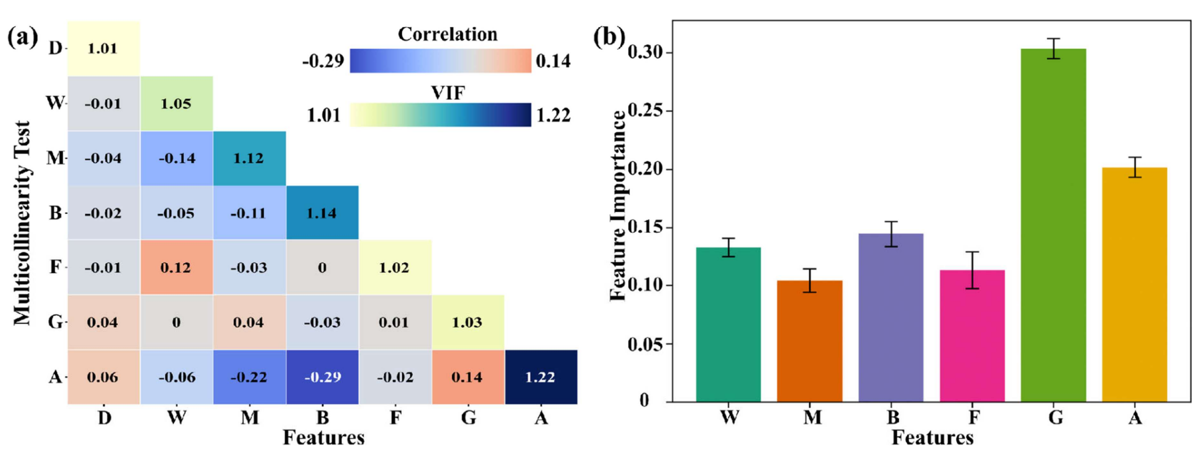

Figure 8. (a) Multicollinearity Test, (b) Feature Importance. In (a), correlation coefficients and Variance Inflation Factors (VIF) are shown in the off- diagonal and diagonal of the triangle, respectively. The abbreviations D, W, M, B, F, G, and A refer to input features: D = deformation (InSAR data), W = groundwater, M = mining areas, B = buildings (urban areas), F = fault lines, G = geology, and A = agricultural areas.

uncertainty in the static features mainly arises from classification errors with spatial structure, and adding Gaussian noise would create configurations that are physically unrealistic. Nevertheless, the feature-ablation experiments indirectly assessed the effect of uncertainty in the static data, because removing each static factor provides an upper bound on the model's dependence on that layer.

The results of the dynamic-channel noise sensitivity tests are presented in Supplementary Figures S2a and S2b. Based on the precision of the D-channel data, 1 mm/year, and its standard deviation of 5.84 mm, we defined six noise levels, 10%, 20%, 30%, 40%, 50%, and 60%. This range spans values close to the precision of the EGMS vertical component as well as extreme cases that substantially exceed it, allowing a systematic assessment of model robustness to different levels of noise perturbation in the D-channel. We applied the same noise levels to the W-channel to evaluate its robustness. The results for noise added to the D-channel, shown in Supplementary Figure S2a, indicate that the performance of the seven models changed only slightly as the noise level increased, demonstrating a degree of noise stability. The results for the W-channel, shown in Supplementary Figure S2b, reveal even smaller performance changes than those observed for the D-channel as the noise level increased, again indicating model stability under noise perturbation.

To test sensitivity to missing temporal data, we masked K time steps in the dynamic input sequences during inference and imputed the missing values using the last observation carried forward method. Supplementary Figures S2c and S2d show the results for the displacement channel D and the groundwater channel W, respectively. Across all seven models, the mean RMSE for the D-channel under normal conditions was 2.90 mm. When the most recent one to five time steps were missing, the mean RMSE increased to 3.23, 3.93, 4.01, 4.10, and 4.19 mm, respectively. The loss of a single time step increased RMSE by about 11%, indicating that the models retained some tolerance to isolated data gaps. As the number of consecutive missing steps increased from 2 to 5, RMSE rose by about 35% to 44% relative to normal conditions, and the corresponding R^2 decreased from 0.90 to about 0.80. These results indicate substantially greater sensitivity to severe data gaps. Notably, even under severe missing-data conditions, RMSE remained below the standard deviation, suggesting that the models still retained predictive capability despite substantial observational gaps. The analysis of the W-channel yielded comparable results.

5. Discussion

We conducted progressive feature-fusion experiments and applied significance tests to the results, confirming the synergistic enhancement effect of multimodal data in surface deformation modelling. By using controlled subsets of the feature space, we linked fusion gains to ablation-based attribution under comparable input conditions and summarised factor effects at the model-cluster level rather than from a single trained instance. On this basis, we used the optimal architecture and best model configuration to quantify the global importance of auxiliary factors in the prediction process through feature ablation. This section focuses on evaluating the importance of auxiliary features in multimodal surface deformation models and comparing our results with previous studies. We first applied SHAP to examine the performance of post hoc attribution methods in a multimodal setting and to identify their limitations in interpreting multimodal data. We then evaluated the reliability of the global feature importance derived from feature ablation. In addition, we explored a pixel-difference filtering approach to analyse local, non-global feature importance and finally assessed whether the ablation-based global weights can support surface deformation susceptibility mapping in this study.

5.1. Feature importance analysis using SHAP

In most existing studies, researchers determine global feature importance with SHAP, especially in classification tasks. Comparable studies in regression remain relatively limited, particularly for multimodal datasets. Compared with our approach, a SHAP-based method requires only one model trained on the multi-channel dataset to compute global feature importance. To match the characteristics of our experimental models, we used DeepExplainer (Lundberg and Lee 2017; Shrikumar, Greenside, and Kundaje 2017), an implementation of SHAP for deep neural networks that approximates Shapley values to derive global feature importance.

Because our goal was to examine how auxiliary factors influence deformation modelling, we computed SHAP values for the six auxiliary factors while holding the deformation-history channel fixed at its background mean. We then summarised channel importance by averaging the absolute attributions across all time steps and pixels. The results show that groundwater, mining-area distribution, urban areas, fault lines, agricultural areas, and geological structure account for 95.95%, 0.58%, 0.52%, 0.26%, 1.70%, and 1.35% of the total weight, respectively. Under this aggregation rule, groundwater receives substantially higher global importance than the static spatial layers. This ranking differs from the global importance pattern derived from the feature-ablation experiments.

One possible explanation for this difference is that SHAP produces element-level attributions, which require additional aggregation across space and time to obtain global scores. As a result, SHAP-based global channel weights may depend on the background set, feature representation, and spatiotemporal aggregation rule (Lundberg et al. 2020). Previous studies also show that under correlated inputs and complex representations, such global summaries can be sensitive to the aggregation strategy and feature dependence structure (Kumar et al. 2020; Lim et al. 2021). In our setting, static layers repeat over time and usually show sparse spatial distributions. Global mean aggregation may therefore dilute their contributions and shift the ranking toward continuous dynamic fields, such as groundwater (Krell et al. 2023).

Feature ablation addresses the same issue from a different perspective. It quantifies the change in predictive performance after removing one auxiliary factor while keeping the other training settings unchanged. This design allows it to evaluate channel-level predictive dependence effectively under the current controlled framework and provides directly comparable estimates of the contribution of each auxiliary factor to model performance in this modelling scenario. However, feature ablation is more computationally demanding because it requires repeated model retraining, whereas SHAP achieves higher computational efficiency with only one trained model and remains suitable for fine-grained post hoc attribution. The two methods, therefore, provide complementary perspectives on feature importance in multimodal spatiotemporal modelling.

For quantitative comparison, Supplementary Table S5 lists the weights of the six auxiliary factors derived from SHAP and from feature ablation. We quantified the agreement between the two global importance rankings with rank-based consistency metrics. The Spearman rank correlation ρ = 0.37, and the Kendall rank correlation was τ = 0.20, which indicates weak agreement between the two global rankings. These results suggest that SHAP-based and ablation-based global feature importance capture different aspects of model behaviour in this multimodal setting, and that the interpretation of global factor importance depends on the explanatory framework used (Kumar et al. 2020; Lim et al. 2021).

5.2. Global feature importance result analysis using feature ablation

Since the feature ablation strategy identified geological formations as the most relevant feature in deformation prediction, here we further analyse the relationship between geological information and surface displacement monitoring data in the study area. Figure 9 presents the geological overview of the area. Specifically, Figure 9a illustrates the stratigraphic units within the study area alongside their corresponding lithological proportions, with the permeability values for each lithology annotated. Figure 9b depicts the overall distribution of surface-exposed stratigraphic units within the region (representing the G factor in progressive feature fusion and ablation experiments), and Figure 9c provides the lithological composition of these stratigraphic units. Figure 9h displays a local geological cross-section of the area. Additionally, to more intuitively demonstrate the relationship between surface deformation and geological information, we superimposed the displacement rates from 2016 to 2021 onto the stratigraphic map (Figure 9c). By comparing the stratigraphic data and surface deformation data in Figure 9, we observe that areas exhibiting higher surface deformation rates are predominantly concentrated within Quaternary strata (9b), which dominate the exposed surface in this region (9a). Detailed analysis of the lithological characteristics of Quaternary strata indicates that the highest deformation rates occur primarily in sand and sand-gravel deposits with relatively high permeability (9c). Concurrently, Figure 9h provides a regional geological cross-section, supplementing vertical stratigraphic information. Examining the Hambach and Bergheim areas annotated in Figure 9h, regions with thicker Quaternary deposits also exhibit characteristics of high deformation rates.

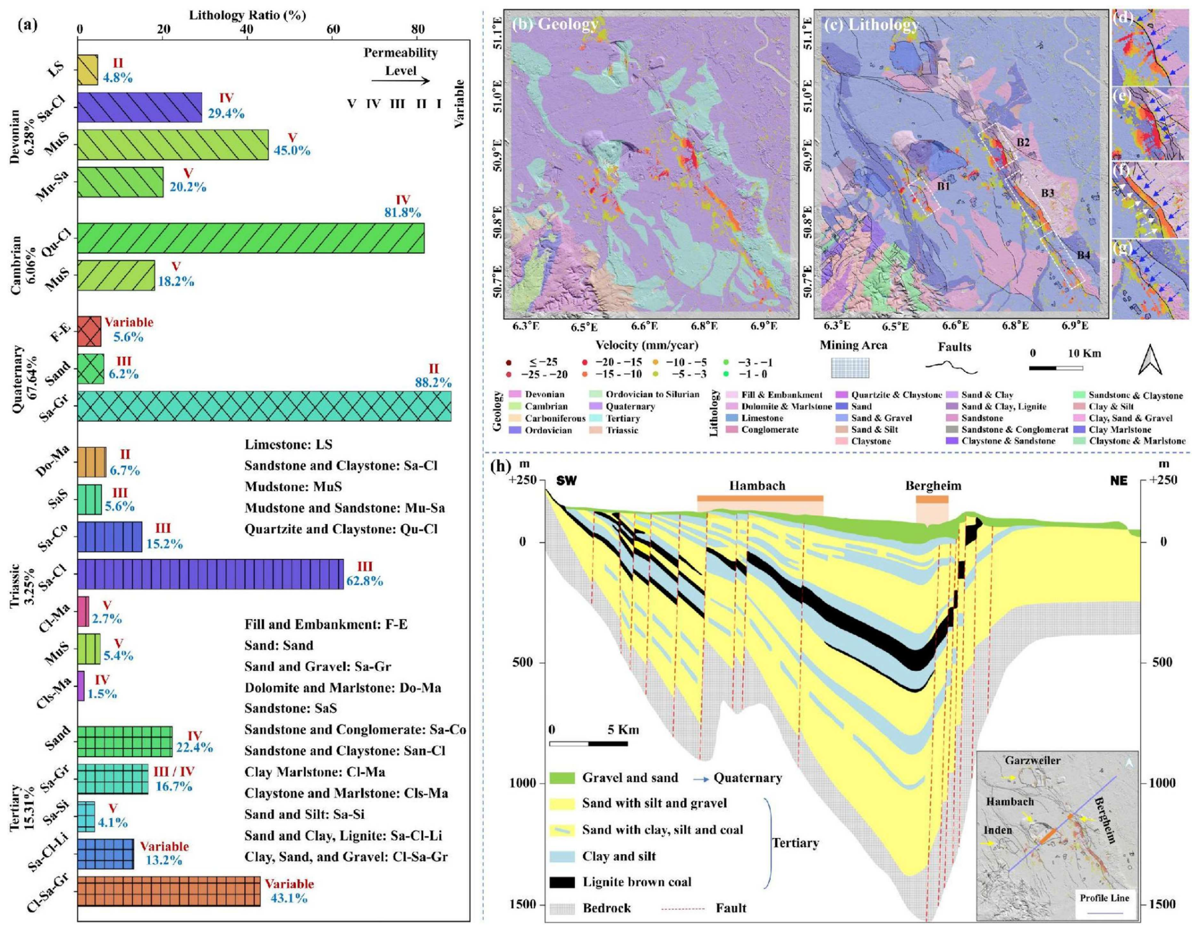

Figure 9. Regional geological overview. (a) Statistical summary of regional geological information. The permeability is categorised into different levels, ranging from V to I. The term "Variable" indicates that the permeability varies within the defined categories. The X-axis displays different strata and rock types; different fill patterns represent various strata, while colours distinguish between rock types. The percentage values shown for each stratum and above each histogram indicate the proportion of that stratum in the entire study area and the proportion of a specific rock type within that particular stratum. (b) Regional geological map. (c) Regional lithological map. InSAR deformation data is overlaid on both (a) and (b), with (c) additionally showing faults (black lines) and mining area boundaries. (d) to (g) are enlarged views of white dashed-line boxes B1 to B4, illustrating deformation features near faults. (h) Regional geological cross-section.

5.2.1. Primary and driving factor analysis

Taken together, the map patterns in Figure 9b and c and the profile context in Figure 9h suggest a stratigraphy-controlled susceptibility template, where thick unconsolidated units provide the deformable layer and groundwater-driven pore-pressure changes provide the transient forcing. This template helps explain why global factor importance emphasises geological setting. This pattern underscores the strong influence of hydro-mechanical properties on surface deformation, which is spatially heterogeneous and closely tied to the geological setting (McDermott and Kolditz 2006). In the study area, Quaternary (67.64%) and Tertiary (15.31%) deposits make up the majority of the subsurface and are primarily composed of highly permeable sediments such as sandstone and gravel (Figure 9a). These unconsolidated materials are more susceptible to compaction and deformation under changes in pore pressure. In contrast, the Devonian (6.28%) and Cambrian (6.06%) strata consist of low-permeability and more rigid lithologies, such as mudstone and quartzite, which are less prone to deformation. This unique geological distribution (G), coupled with groundwater fluctuations, plays a critical role in shaping deformation patterns. High-permeability strata, such as Quaternary sand and gravel, exhibit a rapid hydraulic response to groundwater level fluctuations (Dochartaigh, Bonsor, and Bricker 2017). According to Terzaghi's effective stress principle ($\sigma' = \sigma - u$, where σ' is effective stress, σ is total stress, and u is pore water pressure), abrupt changes in pore water pressure within these formations cause redistributions of effective stress, which in turn lead to elastic or plastic deformation. For example, groundwater extraction accelerates the dissipation of pore water pressure, increases effective stress, and triggers subsidence dominated by compaction (Galloway, Jones, and Ingebritsen 1999). Conversely, rapid aquifer recharge can cause short-term elastic rebound. However, the limited compressibility of coarse-grained sediments generally constrains the magnitude of such deformation, resulting in subsidence that predominantly follows a long-term trend.

In addition to their volumetric dominance, Quaternary high-permeability sediments significantly contribute to deformation due to both their sensitivity to groundwater level fluctuations and widespread spatial distribution. Their hydraulic responsiveness amplifies surface displacement, especially in zones of active groundwater extraction or recharge. Meanwhile, the low-permeability Devonian and Cambrian strata introduce spatial heterogeneity by impeding groundwater flow and localising pressure changes. The Triassic, though covering only 3.25% of the area, exhibits a wide range of permeability—from high to very low (see Figure 9a). This variability makes it a locally influential unit in deformation modelling. Despite its limited areal extent, the Triassic's heterogeneous hydro-mechanical behaviour adds complexity to the overall deformation field, thereby increasing the weight of geological distribution (G) in predictive models. Together, these findings underscore the central role of geological heterogeneity in shaping surface deformation patterns.

5.2.2. Feature-level rationality analysis

Based on the geological background analysis above, we interpret groundwater level variation as a physically plausible process closely linked to surface deformation in this study area. To explain the feature importance rankings derived from the feature ablation strategy, we discuss the results from two perspectives: hydrogeological consistency and influence scope.

From a hydrogeological interpretation perspective, groundwater variability in the study area is closely associated with multiple spatial factors. Therefore, it is physically plausible that some spatial factors receive higher ablation weights than groundwater in this predictive setting, especially when they provide persistent spatial context that conditions how groundwater-related signals are expressed in the model (Frye, Rowat, and Feige 2020; Yu et al. 2020). From the influence scope perspective, we further examined the rationality of the weight ranking among the spatial features themselves and the groundwater factors. We performed a statistical analysis of the distribution of the spatial features, finding that agricultural land, urban areas, faults, and mining areas account for 59.22%, 23.78%, 5.69%, and 4.97% of the total pixels, respectively. Comparing these results, we found that the weight ranking in the ablation strategy is broadly consistent with the spatial extent of these features, which provides supportive evidence for the observed importance hierarchy.

Agricultural land and urban areas regulate groundwater levels through groundwater extraction, exerting widespread impacts on surface deformation. By contrast, mining areas and faults exhibit more localised effects on groundwater levels, with smaller influence scopes. As shown in Figure 9c and the zoomed-in views in Figure 9d to Figure 9g, surface deformation primarily occurs in areas where faults and mining activities are distributed. A closer inspection suggests that faults act as structural discontinuities that compartmentalise the deformation field, limiting cross-fault displacement transfer and producing deformation gradients across the fault trace. Across the white dashed boxes B1 to B4 in Figure 9c, deformation concentrates preferentially on one side of the mapped fault traces, which indicates a sharp deformation gradient across the structural boundary. Such asymmetry is consistent with fault-zone compartmentalisation that limits lateral pressure diffusion and localises compaction within fault-bounded units, rather than a purely smooth regional trend (Motagh et al. 2007). Collectively, the influence scope of agricultural land and urban areas on groundwater regulation exceeds that of mining areas and faults. Consequently, groundwater exerts a broader impact on surface deformation than either faults or mining areas. Therefore, analysing the factors' influence scopes supports the higher weight assigned to groundwater compared to mining areas and faults.

Taken together, the feature importance ranking derived from the ablation experiments is broadly consistent with the geological setting, hydrogeological conditions, and spatial distribution of the input factors across the study area. This global ranking primarily reflects the factors that govern the broad spatial pattern of deformation at the scale of the study area, with geological and groundwater-related information providing persistent background control at the regional scale. Nevertheless, the ablation-derived weights quantify predictive importance under the current model and data setting, rather than providing stand-alone evidence of strict physical causality. Therefore, any physical interpretation of these rankings still requires support from hydrogeological and geomechanical evidence, including geological setting, permeability contrasts, groundwater variation, and effective-stress reasoning. This interpretation also needs to account for spatial heterogeneity, because deformation gradients concentrate near structural boundaries and dominant controls may vary locally. In the next section, we further investigate this scale dependence through an analysis of local feature importance in areas exhibiting spatially concentrated deformation responses.

5.3. Non-global feature importance analysis based on feature ablation strategy

The study above assessed the global impacts of various factors on surface deformation. However, as shown in Figure 2, groundwater level changes, fault lines and mining area distributions, and the spatial distribution of high-rate surface deformation regions (Figure 1c) exhibit pronounced spatial heterogeneity. The uneven distribution of these elements suggests that the influence of each factor may vary significantly across the study area. Thus, analysing the local impacts of these factors on surface deformation is of critical importance.

To explore these localised effects, we examine the spatial variation in the model's predictions. We observe that the pixel-level differences between the predictions of the full model (7C) and the ablated models (6C) exhibit distinct spatial patterns. We hypothesise that regions with larger absolute differences ($|P_{7C}$ - $P_{6C\_i}|$, where P represents the predicted deformation and i is the removed feature) correspond to areas where the excluded feature has a significant influence on the model's predictions. Conversely, smaller differences suggest lower local sensitivity to that feature. Based on this premise, we propose a pixel-difference filtering method for local importance analysis. For each ablation scenario, we compute the pixel-level absolute difference maps between the 7C model's predictions and the corresponding 6C model (with one feature removed). To highlight regions most affected by the feature removal, potentially indicating higher sensitivity or instability - we apply median filtering to these difference maps.

By incrementally increasing the filtering threshold, we can identify pixels that exhibit the most significant predictive differences, as illustrated in Figure 10(a). The underlying principle is that median filtering removes isolated outlier pixels with minor differences while preserving contiguous regions with consistently large differences. These retained regions may reflect areas where the removed feature had a strong local impact. To quantify this local importance, we compare the proportion of pixels remaining after applying a consistent filtering threshold across all ablation difference maps. A higher number of retained pixels for a given feature suggests that its removal led to substantial changes in predictions across a broader area. This suggests that the feature plays a more influential role in shaping deformation predictions in those regions. This data-driven approach, grounded in empirical model behaviour, highlights the spatial dimension of feature influence and provides a complementary perspective to the global analysis.

Applying our proposed local influence analysis method reveals the spatial importance of each factor, as shown in Figure 10b. At a threshold of 0, the number of affected pixels after removing any single factor is identical across all ablation scenarios, establishing a uniform baseline for comparison. However, starting from a threshold of 0.2, significant differences emerge in the proportion of pixels affected by the removal of different factors. As the filtering threshold increases, these proportions change notably, effectively isolating regions with larger local impacts. The statistical results reveal that the model without the fault factor (6C-F) results in the highest number of affected pixels across all the threshold values. This finding strongly suggests that, within the study area, the fault distribution exerts the most extensive and significant local influence on surface deformation predictions. This marked local impact can be partially attributed to the structural hydrogeological role of faults, which can serve as preferential pathways for groundwater flow. At the same time, fault zones can also reduce cross-fault connectivity, so they can amplify along-structure responses while maintaining deformation contrasts across the structural boundary (Bense et al. 2013). Within their zones of influence, faults can dominate groundwater behaviour, overriding the broader lithologic control (geological factor, G) on groundwater levels (W) by focusing pressure gradients into

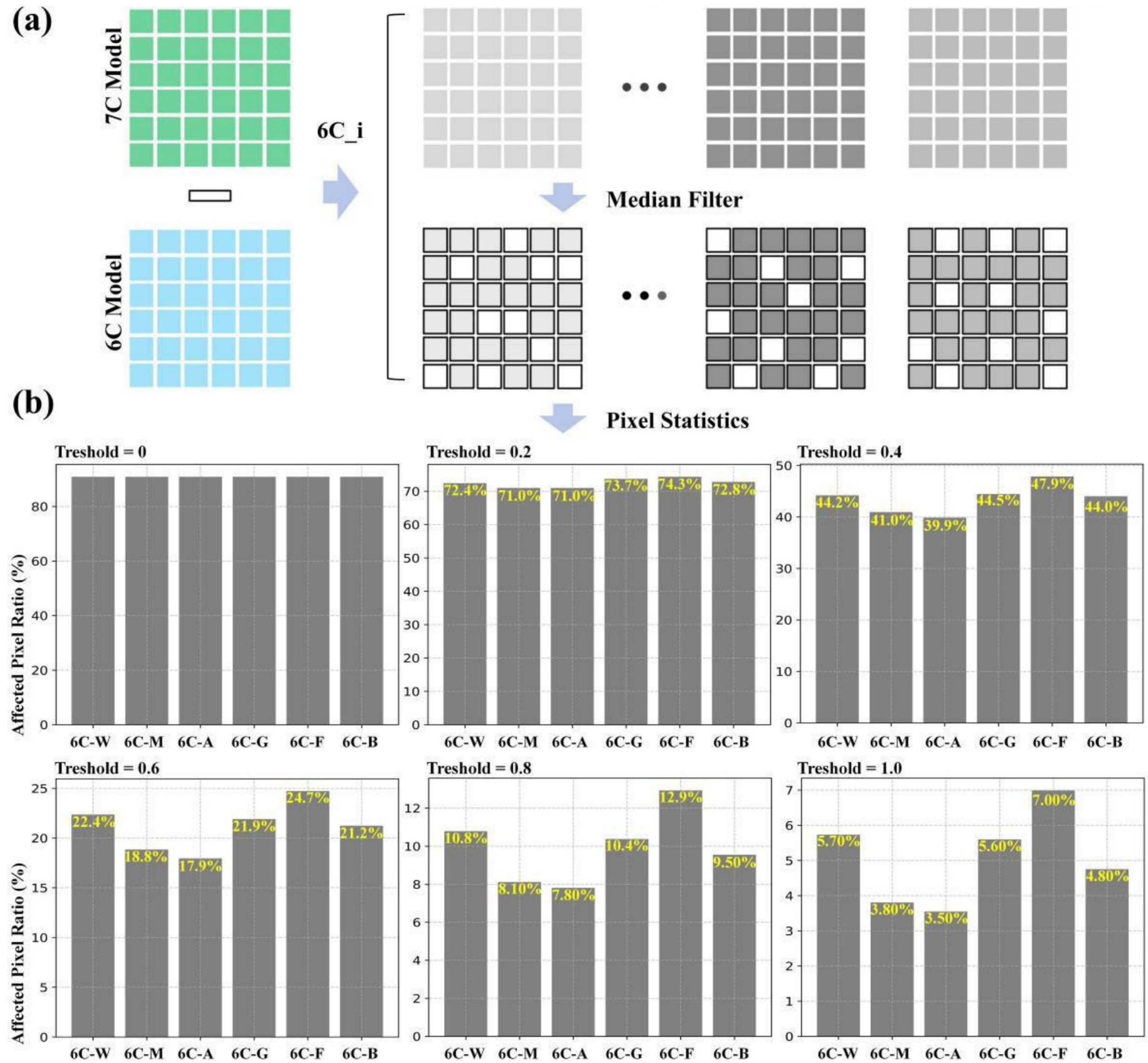

Figure 10. Results of Non-Global Feature Importance Analysis. (a) Schematic diagram of pixel difference filtering, (b) Importance of local features based on pixel difference filtering. In all histograms, the x axes indicate the feature which is removed (6C_i, where i is the excluded feature). The histograms display the percentage of affected areas by the removed feature for different thresholds.

fault-influenced corridors (Chesley et al. 2021). The geological background provided in Figure 9h serves as evidence for this interpretation.

Interestingly, the relative local importance ranking of geological (G) and groundwater (W) factors exhibits a threshold-dependent pattern. At lower thresholds (0.2 and 0.4), which capture broader regions of moderate impact, the proportion of affected pixels resulting from the removal of the geological factor exceeds that of the groundwater factor ablation (G>W). In contrast, at higher thresholds (0.6, 0.8, and 1.0), which isolate smaller regions with high-intensity impact, the removal of groundwater data affects a greater proportion of pixels than the removal of geological information (W>G).

Taking the ablation of the fault feature as an example, this scale-dependent behaviour is further visualised in Figure 11. At the same time, we conducted the same experiments for other features and found that, as the threshold increases, the pixel variation pattern overall resembles the case of ablating the fault factor. Among these, the case of ablating the groundwater factor shows the most similarity, followed by the geological factors. Our findings highlight the strong role of the fault factor in the non-global feature importance analysis, indicating that groundwater- and geology-related information becomes particularly influential in faulted areas. This scale-dependent pattern is consistent with an interpretation in which

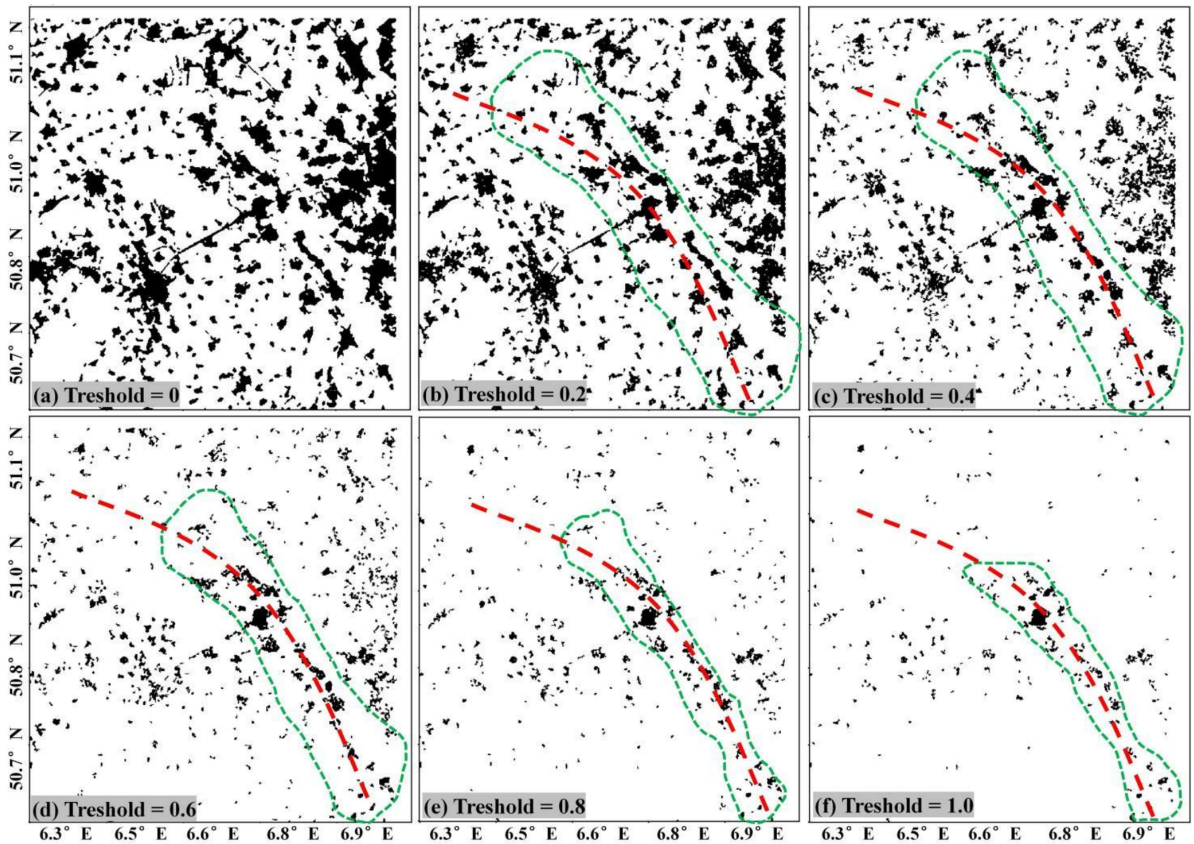

Figure 11. Spatial distribution of pixel difference filtering results. Subplots (a) to (f) illustrate the spatial distribution of pixel differences under increasing threshold conditions following the ablation of fault features, and the red dashed line shows the distribution of the fault line. The green dotted box highlights regions of significant predictive change. As the threshold increases, the affected areas contract and progressively concentrate around the fault line, indicating more pronounced deformation in the vicinity of these structures.

geology sets the background susceptibility, while fault-controlled connectivity and gradients concentrate groundwater-related stress changes into localised deformation patches.

Based on these observations, a hierarchy of local importance can be established that varies with the spatial scale or intensity of the observed impacts. For broader regions characterised by lower thresholds, the relative importance of factors influencing surface deformation follows the order: fault lines (F) > geological formations (G) > groundwater level changes (W) > urban areas (B) > mining areas (M) > agricultural areas (A). In contrast, for more localised and high-intensity impact regions, identified at higher thresholds, the ranking shifts to: fault lines (F) > groundwater level changes (W) > geological formations (G) > urban areas (B) > mining areas (M) > agricultural areas (A).

These findings provide valuable insights for the study area. The pronounced local influence of faults highlighted by our model analysis is consistent with the interpretation that faults may act as conduits that amplify groundwater responses in some parts of the study area. As such, development planning should prioritise the avoidance of fault zones. For existing infrastructures located near faults or within high-permeability strata, enhanced monitoring is recommended, especially following heavy rainfall events, when faults may induce sudden changes in groundwater level. This approach can help mitigate the risk of deformation-related damage.

5.4. Susceptibility mapping based on global feature importance

We further constructed a surface deformation susceptibility score for the study area using the global feature-importance weights derived from feature ablation. We then converted the six auxiliary factors identified in Section 4.2 into spatial predictor layers with a 100 m grid and integrated these layers using a weighted linear combination. We calculated the susceptibility score at pixel (x,y) as:

$$R(x, y) = \sum_{i=1}^{6} w_i F_i(x, y) \tag{5}$$

where w_i denotes the normalised global importance weight of factor i derived in Section 4.2, and $F_i(x, y)$ denotes the normalised spatial value of that factor at pixel (x,y). This formulation follows the weighted linear combination framework commonly used in spatial overlay analysis (Malczewski 2006).

For groundwater, we used the groundwater-level change rate extracted from the time-series data as the groundwater-derived predictor layer. We represented mining, urban, fault, and agricultural areas as binary predictor layers and normalised them before integration. We reclassified the geological layer into ordered susceptibility classes based on regional lithological composition and permeability contrasts. We then normalised the geological layer and included it in the weighted summation (Bundesanstalt für and and Rohstoffe 2019; Schäfer et al. 2005). Table S6 details how each factor was converted into a spatial susceptibility input. Table S7 summarises the reclassification scheme for the geological layer. Finally, we divided the continuous susceptibility index into four classes using the Jenks natural breaks method for map visualisation.

Figure 12a shows the susceptibility rank. Areas surrounding the coal mining districts fall into the high-susceptibility or moderate-susceptibility classes, whereas areas near the mapped fault traces fall into the very high-susceptibility class. Within the high-susceptibility class, field investigation in Area A, located southwest of the Hambach mine (Figure 12a), identified severe structural damage (Figure 12f–i). Area B, located near the former Bergheim and Fortuna open-pit mines, also falls into the high-susceptibility class, and field observations confirmed building damage in this area (Figure 12b–e). Both field-investigated areas, Area A and Area B, belong to the high-susceptibility class and exhibit documented structural damage. These observations provide qualitative field-based support for the spatial plausibility of the susceptibility assessment.

From an application perspective, the weights derived from the ablation analysis provide a model-based basis for prioritising deformation screening. Specifically, the higher weights assigned to geological and land-use related factors support regional susceptibility zoning. The groundwater-derived layer can guide dynamic monitoring in deformation-prone areas. Fault and mining-area layers further help identify priority zones for denser InSAR monitoring, field investigation, or infrastructure inspection.

5.5. Limitations

Our primary evaluation objective was to predict future surface deformation within the same study area from historical observations. We therefore adopted a strict chronological split in the main experiments and computed normalisation statistics from the training set only. This design prevented future information from entering the training process and controlled the bias introduced by normalisation based on the full dataset. This setting is consistent with studies that assess temporal extrapolation by holding out future periods, and the resulting performance primarily serves future prediction within the same system rather than spatial transfer evaluation (Wunsch, Liesch, and Broda 2022). However, because our data underwent temporal alignment, spatial interpolation, and unified gridding, spatial dependence may persist among neighbouring pixels and local patches. As a result, although the chronological split suits the temporal extrapolation objective of this study, the test results may still provide an optimistic estimate of the true generalisation ability, especially in evaluation settings that require stronger spatial independence (Koldasbayeva et al. 2024; Ploton et al. 2020).

To define the scope of applicability of the chronological-split results, we conducted block-based spatial split tests on the GCN + Transformer architecture under two input settings, full-feature input and InSAR-only input, and repeated the experiments 10 times with different random seeds and block configurations. The results show that absolute model performance declines under spatial splitting relative to chronological splitting, which indicates that the original chronological-split results may be optimistic because spatial dependence remains in the data. This finding also suggests that the performance gains and their statistical

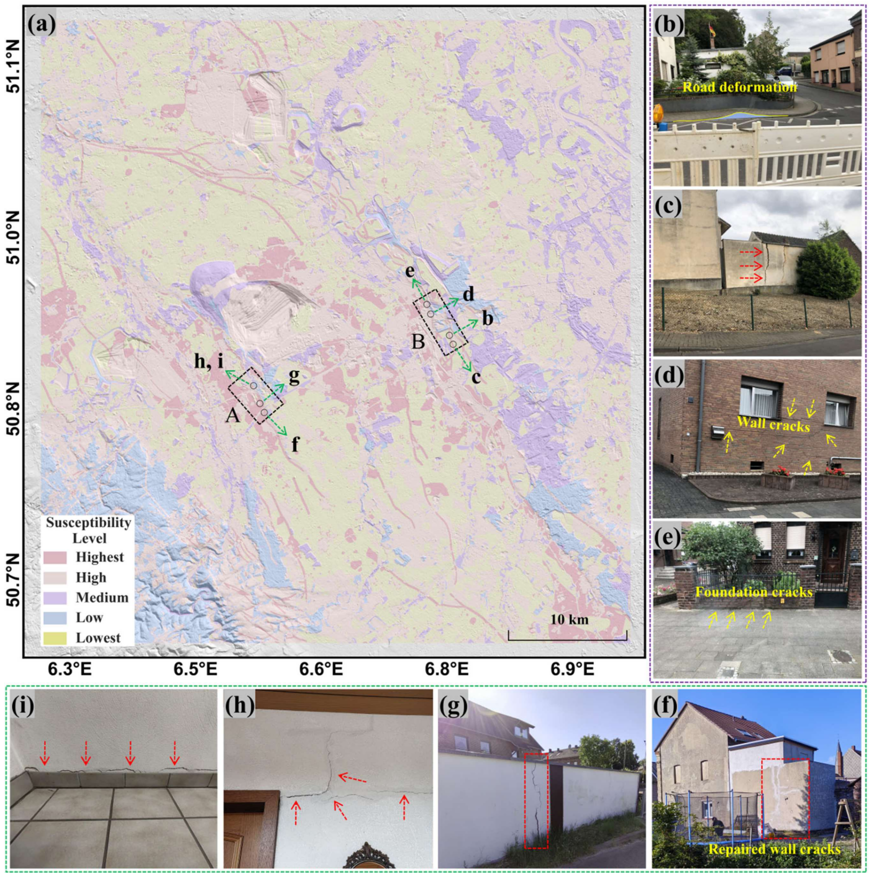

Figure 12. Global deformation susceptibility map for the study area. (a) Surface deformation susceptibility assessment map. (b–e) Field observations of damage from Area B, near the former open-pit mines. (f–i) Field observations of damage from Area A, close to the Hambach open-pit mine (Photo credit: Maoqi Liu & Mahdi Motagh, June 2024). Areas A and B are located within high-susceptibility zones in (a), consistent with the field observations. Circles in (a) mark the locations of the documented damaged structures.

significance reported under chronological splitting should be interpreted as results for temporal extrapolation within the study area, rather than as evidence of generalisation under spatially independent conditions.

Even so, across the 10 spatial-split experiments, the full-feature model still showed a consistent relative advantage over the InSAR-only model on major metrics such as RMSE, which supports the overall gain of multimodal input over unimodal input. For example, across the 10 repeated spatial-split experiments, the mean RMSE decreased from 5.758 ± 0.906 to 5.614 ± 0.827, and the full-feature model outperformed the InSAR-only model in all 10 runs for RMSE, R^2, and RRMSE, as reported in Supplementary Table S8. This result indicates that the benefit of multimodal information does not arise entirely from the relatively easier test set under chronological splitting. Therefore, although the spatial-split tests suggest that the original chronological-split evaluation may be optimistic, the basic conclusion that multimodal input provides a real benefit over unimodal input still holds. Future work should combine stricter spatial buffer splitting, independent validation at the level of raw observations, and stronger physics-constrained modelling to improve the conservativeness of the evaluation and the robustness of the interpretation.

A further limitation concerns that the susceptibility assessment in this study currently relies only on supporting evidence from damaged buildings and infrastructure in the two field-surveyed areas, Area A and Area B. These observations indicate spatial agreement between high-susceptibility zones and locally documented damage, but their spatial coverage remains limited, and this validation is primarily qualitative. They therefore do not constitute a systematic validation of the full susceptibility map. Future work should incorporate validation data that are independent of the weight-construction process to enable a systematic assessment.

6. Conclusion

By developing a multimodal deep learning framework that integrates dynamic monitoring data with static environmental information, we quantified how auxiliary factors affect surface deformation prediction in the study area. We employed a controlled nested feature subset design comprising twelve datasets, utilising these datasets to conduct progressive feature fusion experiments and feature ablation experiments on the six established architectures, respectively. Progressive fusion results indicate that additional auxiliary factors generally reduce error metrics and enhance the R^2 values of the architectures. Based on statistical analysis from progressive feature fusion experiments, we selected the GCN + Transformer architecture for feature ablation experiments and subsequent global importance analysis. Concurrently, we compared the global feature-importance rankings derived from SHAP and feature ablation, and found weak agreement between the two approaches. This difference suggests that, in the present multimodal spatiotemporal setting, SHAP-based and ablation-based global summaries capture different aspects of model behaviour. SHAP remains useful for efficient post hoc attribution from a single trained model. In contrast, feature ablation quantifies channel-level predictive dependence under controlled retraining with matched training settings, although at substantially higher computational cost. We therefore used feature ablation as the main basis for channel-level global comparison in this study, while treating SHAP as a complementary attribution analysis.

The global feature importance weights derived from the comprehensive ablation experiment are as follows: geological features account for 30.4%, agricultural areas for 20.2%, urban areas for 14.4%, groundwater for 13.3%, fault lines for 11.3%, and mining areas for 10.4%. Based on the weighted ranking and the geological background analysis, geological information is the most influential auxiliary feature, and groundwater variation is the most influential dynamic feature in this predictive setting. Combined with the regional geological background and Terzaghi's effective stress principle, this pattern is consistent with an interpretation in which groundwater level fluctuations are closely linked to deformation, while geological conditions modulate the spatial expression of that response through contrasts in rock permeability. This mechanism explains why the Geology feature holds the highest proportion in the global weighting. From a hydrogeological perspective, groundwater-level variations in the study area are associated with multiple spatial factors. The influence of the auxiliary factors on model predictions may therefore be expressed in part through their spatial associations with groundwater fluctuations and their distribution patterns. However, local importance analysis based on pixel difference filtering reveals that fault factors correspond to the highest proportion of affected pixels across all thresholds, indicating faults contribute most significantly to deformation prediction within locally high-impact zones. This phenomenon is partly attributable to the hydrogeological structural role of faults, as they serve as preferential pathways for groundwater flow and dominate groundwater behaviour within their influence zones, thereby amplifying local deformation responses.

Finally, we translated the global weights into regional surface deformation susceptibility assessment maps, with field observations from two high-susceptibility areas corroborating the patterns depicted. Future work should validate this controlled feature space and model clustering workflow across different regions and hazard types, while strengthening physical constraints beyond prediction dependency when interpreting attribution results.

Acknowledgements

This study was supported by the National Natural Science Foundation of China (grant Nos. 42530710, 42377453, and 42530236), the Science and Technology Innovation Programme of Hunan Province (grant Nos. 2021RC4037 and 2023JJ30235). This work was also partially supported by the Federal Ministry of Economic Affairs and Climate Action (BMWK) under the project SARKI4Tagebaufolgen. Maoqi Liu acknowledges the China Scholarship Council (grant No. 202308430215) that financed his research stay at GFZ. The authors thank Thomas Walter, Yuankun Xu, Wandi Wang, and Dibakar Kamalini Ritushree from GFZ for their helpful discussions and two anonymous reviewers for their valuable suggestions that helped improve the quality of the paper.

Author contributions

CRediT: **Maoqi Liu:** Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing; **Marzieh Baes:** Investigation, Resources, Writing – review & editing; **Mahdi Motagh:** Conceptualization, Investigation, Resources, Writing – review & editing; **Qian Song:** Methodology, Writing – review & editing; **Yao Sun:** Methodology, Writing – review & editing; **Zelong Guo:** Writing – review & editing; **Tao Li:** Conceptualization, Methodology, Writing – review & editing; **Sichun Long:** Conceptualization, Funding acquisition, Methodology, Supervision, Writing – review & editing.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This research was supported by the Science and Technology Innovation Programme of Hunan Province (Grant nos. 2021RC4037; 2023JJ30235), the National Natural Science Foundation of China (Grant nos. 42530710; 42377453; 42530236), the China Scholarship Council (Grant no. 202308430215).

Data availability statement

The experimental dataset (Maoqi et al. 2025) and the corresponding source code for this study are openly accessible via Figshare (https://doi.org/10.6084/m9.figshare.30074560) and GitHub (https://github.com/AomomoCaptain/Data_Fusion), respectively.

References

Alzubaidi, L., J. Zhang, A. J. Humaidi, A. Al-Dujaili, Y. Duan, O. Al-Shamma, J. Santamaría, M. A. Fadhel, M. Al-Amidie, and L. Farhan. 2021. "Review of Deep Learning: Concepts, CNN Architectures, Challenges, Applications, Future Directions." *Journal of Big Data* 8 (1): 53. https://doi.org/10.1186/s40537-021-00444-8.

Anantrasirichai, N., J. Biggs, K. Kelevitz, Z. Sadeghi, T. Wright, J. Thompson, A. M. Achim, and D. Bull. 2020. "Detecting Ground Deformation in the Built Environment Using Sparse Satellite InSAR Data with a Convolutional Neural Network." *IEEE Transactions on Geoscience and Remote Sensing* 59 (4): 2940–2950. https://doi.org/10.1109/TGRS.2020.3018315.

Ao, Z., X. Hu, S. Tao, X. Hu, G. Wang, M. Li, X. Liang, et al. 2024. "A National-Scale Assessment of Land Subsidence in China's Major Cities." *Science (New York, N.Y.)* 384 (6693): 301–306. https://doi.org/10.1126/science.adl4366.

Arisoy, E., A. Sethy, B. Ramabhadran, and S. Chen. 2015. Bidirectional Recurrent Neural Network Language Models for Automatic Speech Recognition. In *2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP*. 5421–5425. IEEE. https://doi.org/10.1109/ICASSP.2015.7179007

Atkinson, G. M., D. W. Eaton, and N. Igonin. 2020. "Developments in Understanding Seismicity Triggered By Hydraulic Fracturing." *Nature Reviews Earth & Environment* 1 (5): 264–277. https://doi.org/10.1038/s43017-020-0049-7.

Bense, V. F., T. Gleeson, S. E. Loveless, O. Bour, and J. Scibek. 2013. "Fault Zone Hydrogeology." *Earth-Science Reviews* 127: 171–192. https://doi.org/10.1016/j.earscirev.2013.09.008.

Bergen, K. J., P. A. Johnson, M. V. de Hoop, and G. C. Beroza. 2019. "Machine Learning for Data-Driven Discovery in Solid Earth Geoscience." *Science (New York, N.Y.)* 363 (6433): eaau0323. https://doi.org/10.1126/science.aau0323.

Bhosale, V., O. Kadam, H. Pendurkar, S. Gosavi, and S. Padelkar. 2024. Algorithmic Insights into Mapping Landslide Susceptibility: a Random Forest Approach in High-Risk Zones of Uttarakhand. In *IGARSS 2024 - 2024 IEEE International Geoscience and Remote Sensing Symposium* 10197–10202.IEEE. https://doi.org/10.1109/IGARSS53475.2024.10641963.

Bilodeau, B., N. Jaques, P. W. Koh, and B. Kim. 2024. "Impossibility Theorems for Feature Attribution." *Proceedings of the National Academy of Sciences* 121 (2): e2304406120. https://doi.org/10.1073/pnas.2304406120.

Bodnar, C., W. P. Bruinsma, A. Lucic, M. Stanley, A. Allen, J. Brandstetter, P. Garvan, et al. 2025. "A Foundation Model for the Earth System." *Nature (London)* 641: 1–8. https://doi.org/10.1038/s41586-025-09005-y.

Bountos, N. I., I. Papoutsis, D. Michail, and N. Anantrasirichai. 2021. "Self-Supervised Contrastive Learning for Volcanic Unrest Detection." *IEEE Geoscience and Remote Sensing Letters* 19: 1–5. https://doi.org/10.1109/LGRS.2021.3104506.

Bundesanstalt für, G., and Rohstoffe 2019. Hydrogeological Map of Germany 1:250,000 (HÜK250). https://gdk.gdi-de.org/geonetwork/srv/api/records/61ac4628-6b62-48c6-89b8-46270819f0d6

Cai, Y., Y. Jin, Z. Wang, T. Chen, Y. Wang, W. Kong, W. Xiao, X. Li, X. Lian, and H. Hu. 2023. "A Review of Monitoring, Calculation, and Simulation Methods for Ground Subsidence Induced By Coal Mining." *International Journal of Coal Science & Technology* 10 (1): 32. https://doi.org/10.1007/s40789-023-00595-4.

Che, Z., S. Purushotham, K. Cho, D. Sontag, and Y. Liu. 2018. "Recurrent Neural Networks for Multivariate Time Series with Missing Values." *Scientific reports* 8 (1): 6085. https://doi.org/10.1038/s41598-018-24271-9.

Chen, B., H. Gong, X. Li, K. Lei, M. Gao, C. Zhou, and Y. Ke. 2015. "Spatial–Temporal Evolution Patterns of Land Subsidence with Different Situation of Space Utilization." *Natural Hazards* 77 (3): 1765–1783. https://doi.org/10.1007/s11069-015-1674-1.

Chen, J., Z. Dai, S. Dong, X. Zhang, G. Sun, J. Wu, R. Ershadnia, S. Yin, and M. R. Soltanian. 2022. "Integration of Deep Learning and Information Theory for Designing Monitoring Networks in Heterogeneous Aquifer Systems." *Water Resources Research* 58 (10): e2022WR032429. https://doi.org/10.1029/2022WR032429.

Chesley, C., S. Naif, K. Key, and D. Bassett. 2021. "Fluid-Rich Subducting Topography Generates Anomalous Forearc Porosity." *Nature (London)* 595 (7866): 255–260. https://doi.org/10.1038/s41586-021-03619-8.

Costantini, M., F. Minati, F. Trillo, A. Ferretti, F. Novali, E. Passera, J. Dehls, et al. 2021. European Ground Motion Service (EGMS). In *2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS*. 3293–3296. IEEE. https://doi.org/10.1109/IGARSS47720.2021.9553562

Crosetto, M., M. Cuevas-González, M. S. Mróz, D. Moldestad, F. Raspini, N. Casagli, L. Bateson, et al. 2026. "European Ground Motion Service: a Decade of Sentinel-1 Observations." *Remote Sensing of Environment* 339: 115389. https://doi.org/10.1016/j.rse.2026.115389.

Davydzenka, T., P. Tahmasebi, and N. Shokri. 2024. "Unveiling the Global Extent of Land Subsidence: the Sinking Crisis." *Geophysical Research Letters* 51 (4): e2023GL104497. https://doi.org/10.1029/2023GL104497.

Dochartaigh, B. Ó, H. Bonsor, and S. Bricker. 2017. "Improving Understanding of Shallow Urban Groundwater: the Quaternary Groundwater System in Glasgow, UK." *Earth and Environmental Science Transactions of the Royal Society of Edinburgh* 108 (2-3): 155–172. https://doi.org/10.1017/S1755691018000385.

Dong, D., D. Ming, M. Li, H. Xu, Y. Wei, and M. Huang. 2025. "A Novel Knowledge-Based Multi-Modal Semi-Supervised Framework for 3D Change Detection and Mining Volume Estimation in Open-Pit Mines Using GF7 Satellite Images." *ISPRS Journal of Photogrammetry and Remote Sensing* 228: 519–536. https://doi.org/10.1016/j.isprsjprs.2025.07.034.

Dong, F., H. Yin, W. Cheng, C. Zhang, D. Zhang, H. Ding, C. Lu, and Y. Wang. 2024. "Quantitative Prediction Model and Prewarning System of Water Yield Capacity (WYC) from Coal Seam Roof Based on Deep Learning and Joint Advanced Detection." *Energy* 290: 130200. https://doi.org/10.1016/j.energy.2023.130200.

Fang, J., R. J. Nicholls, S. Brown, D. Lincke, J. Hinkel, A. T. Vafeidis, S. Du, Q. Zhao, M. Liu, and P. Shi. 2022. "Benefits of Subsidence Control for Coastal Flooding in China." *Nature Communications* 13 (1): 6946. https://doi.org/10.1038/s41467-022-34525-w.

Frye, C., C. Rowat, and I. Feige. 2020. "Asymmetric Shapley Values: Incorporating Causal Knowledge into Model-Agnostic Explainability." *Advances in Neural Information Processing Systems* 33: 1229–1239.

Galloway, D. L., D. R. Jones, and S. E. Ingebritsen. 1999. "Land Subsidence in the United States, Reston, VA: U.S. Geological Survey Circular 1182. eds. U.S. Geological Survey. https://doi.org/10.3133/cir1182.

Geological Survey of North Rhine - Westphalia. 2015. *Geological Survey of North Rhine - Westphalia*. International Federation of Digital Seismograph Networks, [Data set]. https://doi.org/10.7914/pz2d-8r35

Gerard, S., Y. Zhao, and J. Sullivan. 2023. "Wildfirespreadts: a Dataset of Multi-Modal Time Series for Wildfire Spread Prediction." *Advances in Neural Information Processing Systems* 36: 74515–74529. https://doi.org/10.52202/075280-3258.

Gerwin, W., T. Raab, K. Birkhofer, C. Hinz, P. Letmathe, M. Leuchner, M. Roß-Nickoll, et al. 2023. "Perspectives of Lignite Post-Mining Landscapes under Changing Environmental Conditions: What Can We Learn from a Comparison between the Rhenish and Lusatian Region in Germany?" *Environmental Sciences Europe* 35 (2023): 36. https://doi.org/10.1186/s12302-023-00738-z.

Gill, J. C., and B. D. Malamud. 2017. "Anthropogenic Processes, Natural Hazards, and Interactions in a Multi-Hazard Framework." *Earth-Science Reviews* 166: 246–269. https://doi.org/10.1016/j.earscirev.2017.01.002.

Greff, K., R. K. Srivastava, J. Koutník, B. R. Steunebrink, and J. Schmidhuber. 2016. "LSTM: a Search Space Odyssey." *IEEE Transactions on Neural Networks and Learning Systems* 28 (10): 2222–2232. https://doi.org/10.1109/TNNLS.2016.2582924.

Goodchild, Michael F., and W. Li. 2021. "Replication across space and time must be weak in the social and environmental sciences." *Proceedings of the National Academy of Sciences* 118(35). https://doi.org/10.1073/pnas.2015759118.

Guo, F., N. Zhang, X. Feng, Z. Xie, and Y. Li. 2024. "Autonomous Prediction of Rock Deformation in Fault Zones of Coal Roadways Using Supervised Machine Learning." *Tunnelling and Underground Space Technology* 147: 105724. https://doi.org/10.1016/j.tust.2024.105724.

Hager, B. H., J. Dieterich, C. Frohlich, R. Juanes, S. Mantica, J. H. Shaw, F. Bottazzi, et al. 2021. "A Process-Based Approach to Understanding and Managing Triggered Seismicity." *Nature (London)* 595 (7869): 684–689. https://doi.org/10.1038/s41586-021-03668-z.

Haghshenas Haghighi, M., and M. Motagh. 2024. "Uncovering the Impacts of Depleting Aquifers: a Remote Sensing Analysis of Land Subsidence in Iran." *Science Advances* 10 (19): eadk3039. https://doi.org/10.1126/sciadv.adk3039.

He, K., X. Zhang, S. Ren, and J. Sun. 2016. Identity Mappings in Deep Residual Networks. In. *Computer Vision – ECCV 2016*. B. Leibe, J. Matas, N. Sebe and M. Welling, eds. 9908. 630–645, Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-46493-0_38

Hendrycks, D., and Dietterich, T. 2019. Benchmarking neural network robustness to common corruptions and perturbations arXiv preprint arXiv:1903.12261. https://doi.org/10.48550/arXiv.1903.12261.

Herrera-García, G., P. Ezquerro, R. Tomás, M. Béjar-Pizarro, J. López-Vinielles, M. Rossi, R. M. Mateos, et al. 2021. "Mapping the Global Threat of Land Subsidence." *Science (New York, N.Y.)* 371 (6524): 34–36. https://doi.org/10.1126/science.abb8549.

Hill, E. M., J. W. McCaughey, A. D. Switzer, D. Lallemant, Y. Wang, and S. Sathiakumar. 2024. "Human Amplification of Secondary Earthquake Hazards through Environmental Modifications." *Nature Reviews Earth & Environment* 5 (6): 463–476. https://doi.org/10.1038/s43017-024-00551-z.

Hudson-Edwards, K. A., D. Kemp, L. A. Torres-Cruz, M. G. Macklin, P. A. Brewer, J. R. Owen, D. M. Franks, E. Marquis, and C. J. Thomas. 2024. "Tailings Storage Facilities, Failures and Disaster Risk." *Nature Reviews Earth & Environment* 5 (9): 612–630. https://doi.org/10.1038/s43017-024-00576-4.

Huning, L. S., C. A. Love, H. Anjileli, F. Vahedifard, Y. Zhao, P. L. Chaffe, K. Cooper, et al. 2024. "Global Land Subsidence: Impact of Climate Extremes and Human Activities." *Reviews of Geophysics* 62 (4): e2023RG000817. https://doi.org/10.1029/2023RG000817.

Ismail Fawaz, H., G. Forestier, J. Weber, L. Idoumghar, and P. A. Muller. 2019. "Deep Learning for Time Series Classification: a Review." *Data Mining and Knowledge Discovery* 33 (4): 917–963. https://doi.org/10.1007/s10618-019-00619-1.

Jiang, Z., C. Zhao, X. Liu, G. Shi, M. Yan, Q Zhang, and J. Peng. 2024. "The Regional Differentiation on the Spatial Distribution and Influencing Factors of Potential Landslides Across the Entire Loess Plateau, China Based on InSAR and Sub-Region XGBoost-SHAP Model." *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*. https://doi.org/10.1109/JSTARS.2024.3504713.

Kim, J., H. S. Jung, and Z. Lu. 2025. "Ground Surface Displacement Measurement from Sar Imagery Using Deep Learning." *Remote Sensing of Environment* 318: 114577. https://doi.org/10.1016/j.rse.2024.114577.

Koks, E. E., J. Rozenberg, C. Zorn, M. Tariverdi, M. Vousdoukas, S. A. Fraser, J. W. Hall, and S. Hallegatte. 2019. "A Global Multi-Hazard Risk Analysis of Road and Railway Infrastructure Assets." *Nature Communications* 10 (1): 2677. https://doi.org/10.1038/s41467-019-10442-3.

Koldasbayeva, D., P. Tregubova, M. Gasanov, A. Zaytsev, A. Petrovskaia, and E. Burnaev. 2024. "Challenges in Data-Driven Geospatial Modeling for Environmental Research and Practice." *Nature Communication* 15: 10700. https://doi.org/10.1038/s41467-024-55240-8.

Krell, E., H. Kamangir, W. Collins, S. A. King, and P. Tissot. 2023. "Aggregation Strategies to Improve XAI for Geoscience Models That Use Correlated, High-Dimensional Rasters." *Environmental Data Science* 2: e45. https://doi.org/10.1017/eds.2023.39.

Kumar, I. E., S. Venkatasubramanian, C. Scheidegger, and S. Friedler. 2020. "Problems with Shapley-Value-Based Explanations as Feature Importance Measures, *International conference on machine learning*. 5491–5500. PMLR. https://doi.org/10.48550/arXiv.2002.11097.

Lees, M., and R. Knight. 2024. "Quantification of Record-Breaking Subsidence in California's San Joaquin Valley." *Communications Earth & Environment* 5 (1): 677. https://doi.org/10.1038/s43247-024-01778-w.

Li, G., M. Muller, A. Thabet, and B. Ghanem. 2019. "DeepGCNs: Can GCNs Go as Deep as CNNs?" *In Proceedings of the IEEE/CVF International Conference on Computer Vision*, 9267–9276. https://doi.org/10.1109/ICCV.2019.00936.

Li, X., B. Liu, G. Zheng, Y. Ren, S. Zhang, L. Gao, and F. Wang. 2020. "Deep-Learning-Based Information Mining from Ocean Remote-Sensing Imagery." *National Science Review* 7 (10): 1584–1605. https://doi.org/10.1093/nsr/nwaa047.

Lim, B., and S. Zohren. 2021. "Time-Series Forecasting with Deep Learning: a Survey." *Philosophical Transactions of the Royal Society A* 379 (2194): 20200209. https://doi.org/10.1098/rsta.2020.0209.

Lim, B., S. Ö Arık, N. Loeff, and T. Pfister. 2021. "Temporal Fusion Transformers for Interpretable Multi-Horizon Time Series Forecasting." *International Journal of Forecasting* 37 (4): 1748–1764. https://doi.org/10.1016/j.ijforecast.2021.03.012.

Long, S., M. Liu, C. Xiong, T. Li, W. Wu, H. Ding, L. Zhang, C. Zhu, and S. Lu. 2023. "Research on Prediction of Surface Deformation in Mining Areas Based on TPE-Optimized Integrated Models and Multi-Temporal Insar." *Remote Sensing* 15 (23): 5546. https://doi.org/10.3390/rs15235546.

Lundberg, S. M., and S. I. Lee. 2017. "A Unified Approach to Interpreting Model Predictions." *Advances in Neural Information Processing Systems* 30: 4765–4774.

Lundberg, S. M., G. Erion, H. Chen, A. DeGrave, J. M. Prutkin, B. Nair, R. Katz, J. Himmelfarb, N. Bansal, and S. Lee. 2020. "From Local Explanations to Global Understanding with Explainable AI for Trees." *Nature Machine Intelligence* 2(2020): 56–67. https://doi.org/10.1038/s42256-019-0138-9.

Ma, S., H. Qiu, D. Yang, J. Wang, Y. Zhu, B. Tang, K. Sun, and M. Cao. 2023. "Surface Multi-Hazard Effect of Underground Coal Mining." *Landslides* 20 (1): 39–52. https://doi.org/10.1007/s10346-022-01961-0.

Ma, P., L. Chen, C. Yu, Q. Zhu, Y. Ding, Z. Wu, H. Li, C. Tian, and X. Fan. 2025. "Dynamic Landslide Susceptibility Mapping over Last Three Decades to Uncover Variations in Landslide Causation in Subtropical Urban Mountainous Areas." *Remote Sensing of Environment* 326: 114800. https://doi.org/10.1016/j.rse.2025.114800.

Malczewski, J. 2006. "GIS-Based Multicriteria Decision Analysis: a Survey of the Literature." *International Journal of Geographical Information Science* 20 (7): 703–726. https://doi.org/10.1080/13658810600661508.

Maoqi, L., Baes, M., Motagh, M., Long, S., Song, Q., and Sun, Y., et al. 2025. "Dynamic and Static data in the paper Analyzing Dominant Driving Factors of Ground Deformation in Mining Area via Multimodal Feature Fusion and Explainable Modeling". figshare. Dataset. https://doi.org/10.6084/m9.figshare.30074560.

Marco, H., C. Amelunxen, P. Neis, and A. Zipf. 2012. Comparative Spatial Analysis of Positional Accuracy of OpenStreetMap and Proprietary Geodata. In *GI_Forum. Geospatial Crossroads.* T Jekel, A. Car, J. Strobl and A. Griesebner, eds. 24–33. Wichmann. https://gispoint.de/fileadmin/user_upload/paper_gis_open/537521013.pdf.

Marta, S., A. Zimmer, M. Caccianiga, M. Gobbi, R. Ambrosini, R. S. Azzoni, F. Gili, et al. 2023. "Heterogeneous Changes of Soil Microclimate in High Mountains and Glacier Forelands." *Nature Communications* 14 (1): 5306. https://doi.org/10.1038/s41467-023-41063-6.

McDermott, C., and O. Kolditz. 2006. "Geomechanical Model for Fracture Deformation under Hydraulic, Mechanical and Thermal Loads." *Hydrogeology Journal* 14 (4): 485–498. https://doi.org/10.1007/s10040-005-0455-4.

Moein, M. J., C. Langenbruch, R. Schultz, F. Grigoli, W. L. Ellsworth, R. Wang, A. P. Rinaldi, and S. Shapiro. 2023. "The Physical Mechanisms of Induced Earthquakes." *Nature Reviews Earth & Environment* 4 (12): 847–863. https://doi.org/10.1038/s43017-023-00497-8.

Moser, B., P. G. Cook, A. D. Miller, S. Dogramaci, and I. Wallis. 2024. "The Hydraulic Evolution of Groundwater-Fed Pit Lakes after Mine Closure." *Groundwater* 62 (6): 889–903. https://doi.org/10.1111/gwat.13419.

Motagh, M., Y. Djamour, T. R. Walter, H. U. Wetzel, J. Zschau, and S. Arabi. 2007. "Land Subsidence in Mashhad Valley, Northeast Iran: Results from InSAR, Levelling and GPS." *Geophysical Journal International* 168 (2): 518–526. https://doi.org/10.1111/j.1365-246X.2006.03246.x.

Motagh, M., R. Shamshiri, M. H. Haghighi, H. U. Wetzel, B. Akbari, H. Nahavandchi, M. Haghshenas Haghighi, S. Roessner, and S. Arabi. 2017. "Quantifying Groundwater Exploitation Induced Subsidence in the Rafsanjan Plain, Southeastern Iran, Using InSAR Time-Series and in Situ Measurements." *Engineering Geology* 218: 134–151. https://doi.org/10.1016/j.enggeo.2017.01.011.

Navarro-Hernández, M. I., R. Tomás, J. Valdes-Abellan, G. Bru, P. Ezquerro, C. Guardiola-Albert, A. Elçi, et al. 2023. "Monitoring Land Subsidence Induced By Tectonic Activity and Groundwater Extraction in the Eastern Gediz River Basin (Türkiye) Using Sentinel-1 Observations." *Engineering Geology* 327: 107343. https://doi.org/10.1016/j.enggeo.2023.107343.

Nie, Y. 2022. A Time Series is Worth 64Words: Long-term Forecasting with Transformers. arXiv preprint arXiv:2211.14730. https://doi.org/10.48550/arXiv.2211.14730.

Peng, M., M. Motagh, Z. Lu, Z. Xia, Z. Guo, C. Zhao, and Q. Liu. 2024. "Characterization and Prediction of InSAR-Derived Ground Motion with ICA-Assisted LSTM Model." *Remote Sensing of Environment* 301: 113923. https://doi.org/10.1016/j.rse.2023.113923.

Plésiat, É., R. J. Dunn, M. G. Donat, and C. Kadow. 2024. "Artificial Intelligence Reveals Past Climate Extremes By Reconstructing Historical Records." *Nature Communications* 15 (1): 9191. https://doi.org/10.1038/s41467-024-53464-2.

Ploton, P., F. Mortier, M. Réjou-Méchain, N. Barbier, N. Picard, V. Rossi, C. Dormann, et al. 2020. "Spatial Validation Reveals Poor Predictive Performance of Large-Scale Ecological Mapping Models." Nature Communication 11: 4540. https://doi.org/10.1038/s41467-020-18321-y.

Radman, A., M. Akhoondzadeh, and B. Hosseiny. 2021. "Integrating InSAR and Deep-Learning for Modeling and Predicting Subsidence over the Adjacent Area of Lake Urmia, Iran." *GIScience & Remote Sensing* 58 (8): 1413–1433. https://doi.org/10.1080/15481603.2021.1991689.

Reichstein, M., G. Camps-Valls, B. Stevens, M. Jung, J. Denzler, N. Carvalhais, and Prabhat. 2019. "Deep Learning and Process Understanding for Data-Driven Earth System Science." *Nature (London)* 566: 195–204. https://doi.org/10.1038/s41586-019-0912-1.

Ren, R., Z. Wang, C. Yang, J. Liu, R. Jiang, Y. Zhou, and B. He. 2025. "Enhancing Robotic Skill Acquisition with Multimodal Sensory Data: a Novel Dataset for Kitchen Tasks." *Scientific Data* 12 (1): 476. https://doi.org/10.1038/s41597-025-04798-z.

Ritushree, D. K., M. Baes, and M. Motagh. 2026. "Hydrological and Tectonic Linkage to Subsidence Risk in the Rhineland Coalfields, Germany: Integrating Machine Learning with Remote Sensing and Geospatial Data." *Engineering Geology* 369: 108751. https://doi.org/10.1016/j.enggeo.2026.108751.

Schäfer, A., T. Utescher, M. Klett, and M. Valdivia-Manchego. 2005. "The Cenozoic Lower Rhine Basin, Rifting, Sedimentation, and Cyclic Stratigraphy." *International Journal of Earth Sciences* 94 (4): 621–639. https://doi.org/10.1007/s00531-005-0499-7.

Scoular, J., R. Ghail, P. Mason, and J. Lawrence. 2022. "Are Measured InSAR Displacements a Function of the Chosen Processing Method?" *Quarterly Journal of Engineering Geology and Hydrogeology* 55 (4): qjegh2022-049. https://doi.org/10.1144/qjegh2022-049.

Seidler, V., E. C. Utazi, A. B. Finaret, S. Luckeneder, G. Zens, M. Bodarenko, A. W. Smith, S. E. K. Bradley, A. J. Tatem, and P. Webb. 2025. "Subnational Variations in the Quality of Household Survey Data in Sub-Saharan Africa." *Nature Communications* 16 (1): 3771. https://doi.org/10.1038/s41467-025-58776-5.

Shen, C. 2018. "A Transdisciplinary Review of Deep Learning Research and Its Relevance for Water Resources Scientists." *Water Resources Research* 54.11(2018): 8558–8593. https://doi.org/10.1029/2018WR022643.

Shen, C., A. P. Appling, P. Gentine, T. Bandai, H. Gupta, A. Tartakovsky, M. Baity-Jesi, et al. 2023. "Differentiable Modelling to Unify Machine Learning and Physical Models for Geosciences." *Nature Reviews Earth & Environment* 4 (8): 552–567. https://doi.org/10.1038/s43017-023-00450-9.

Shrikumar, A., P. Greenside, and A. Kundaje. 2017. "Learning Important Features through Propagating Activation Differences, In *Proceedings of the 34th International Conference on Machine Learning*. 3145–3153. PMLR.

Smith, R. G., and S. Majumdar. 2020. "Groundwater Storage Loss Associated with Land Subsidence in Western United States Mapped Using Machine Learning." *Water Resources Research* 56 (7): e2019WR026621. https://doi.org/10.1029/2019WR026621.

Sonter, L. J., M. C. Dade, J. E. Watson, and R. K. Valenta. 2020. "Renewable Energy Production Will Exacerbate Mining Threats to Biodiversity." *Nature Communications* 11 (1): 4174. https://doi.org/10.1038/s41467-020-17928-5.

Tao, Y., R. Zhang, and H. Du. 2024. "Failure Prediction of Open-Pit Mine Landslides Containing Complex Geological Structures Using the Inverse Velocity Method." *Water* 16 (3): 430. https://doi.org/10.3390/w16030430.

Vaswani, A., N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, et al. 2017. "Attention Is All You Need." *Advances in Neural Information Processing Systems* 30: 5998–6008.

Witkowski, W. T., M. Łucka, A. Guzy, H. Sudhaus, A. Barańska, and R. Hejmanowski. 2024. "Impact of Mining-Induced Seismicity on Land Subsidence Occurrence." *Remote Sensing of Environment* 301: 113934. https://doi.org/10.1016/j.rse.2023.113934.

Wunsch, A., T. Liesch, and S. Broda. 2022. "Deep Learning Shows Declining Groundwater Levels in Germany Until 2100 Due to Climate Change." *Nature Communications* 13 (1): 1221. https://doi.org/10.1038/s41467-022-28770-2.

Xing, J., J. Zhang, J. Li, Y. Gao, S. Du, C. Zhang, and Y. Wang. 2024. "CTMNet: Enhanced Open-Pit Mine Extraction and Change Detection with a Hybrid CNN-Transformer Multitask Network." *IEEE Transactions on Geoscience and Remote Sensing* 62: 1–19. https://doi.org/10.1109/TGRS.2024.3492715.

Xiong, Y., D. Kong, and G. Song. 2024. "Research Hotspots and Development Trends of Green Coal Mining: Exploring the Path to Sustainable Development of Coal Mines." *Resources Policy* 92: 105039. https://doi.org/10.1016/j.resourpol.2024.105039.

Yao, S., Y. He, L. Zhang, W. Yang, Y. Chen, Q. Sun, Z. Zhao, and S. Cao. 2023. "A ConvLSTM Neural Network Model for Spatiotemporal Prediction of Mining Area Surface Deformation Based on SBAS-InSAR Monitoring Data." *IEEE Transactions on Geoscience and Remote Sensing* 61: 1–22. https://doi.org/10.1109/TGRS.2023.3236510.

Youssef, K., K. Shao, S. Moon, and L. Bouchard. 2023. "Landslide Susceptibility Modeling By Interpretable Neural Network." *Communications Earth & Environment* 4: 162. 2023. https://doi.org/10.1038/s43247-023-00806-5.

Yu, S., and J. Ma. 2021. "Deep Learning for Geophysics: Current and Future Trends." *Reviews of Geophysics* 59 (3): e2021RG000742. https://doi.org/10.1029/2021RG000742.

Yu, K., X. Guo, L. Liu, J. Li, H. Wang, Z. Ling, and X. Wu. 2020. "Causality-Based Feature Selection: Methods and Evaluations." *ACM Computing Surveys (CSUR)* 53 (5): 1–36. https://doi.org/10.1145/3409382.

Zhang, Z., P. Cui, and W. Zhu. 2020. "Deep Learning on Graphs: a Survey." *IEEE Transactions on Knowledge and Data Engineering* 34 (1): 249–270. https://doi.org/10.1109/TKDE.2020.2981333.

Zhang, Y., J. Zhang, and J. Ma. 2022. "Stability Analysis of a Steep Rock Slope in a Large Open-Pit Mine in a High-Intensity Area: a Case Study of the Yejiagou Boron Iron Mine." *Geofluids* 2022 (1): 9113173. https://doi.org/10.1155/2022/9113173.

Zhou, C., M. Ye, Z. Xia, W. Wang, C. Luo, and J. P. Muller. 2025. "An Interpretable Attention-Based Deep Learning Method for Landslide Prediction Based on Multi-Temporal InSAR Time Series: a Case Study of Xinpu Landslide in the TGRA." *Remote Sensing of Environment* 318: 114580. https://doi.org/10.1016/j.rse.2024.114580.

Zhu, J., S. Miao, R. Ying, and P. Li. 2025. "Towards Unveiling Sensitive and Decisive Patterns in Explainable AI with a Case Study in Geometric Deep Learning." *Nature Machine Intelligence* 7: 1–13. https://doi.org/10.1038/s42256-025-00998-9.