Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

A Graph Attention Framework with Kinematic Constraints for Network-Based GNSS Time Series Prediction

Siyu Zhu
School of Transportation Engineering, East China Jiaotong University, Nanchang 330013, China
https://orcid.org/0009-0009-4349-100X

Gaël Kermarrec
Institute of Meteorology and Climatology, Leibniz University Hannover, Germany

Xiaoxing He
Jiangxi Province Key Laboratory of Water Ecological Conservation in Headwater Regions, Jiangxi University of Science and Technology, Ganzhou 341000, China

Shengdao Wang
Division of Geodetic Science, School of Earth Sciences, The Ohio State University, Columbus, Ohio 43210, USA

Rui Fernandes
Institute Dom Luiz, University of Beira Interior, 6201-001 Covilhã, Portugal

Zhengkai Huang
zhkhuang@whu.edu.cn

School of Transportation Engineering, East China Jiaotong University, Nanchang 330013, China

Jean-Philippe Montillet
Institute Dom Luiz, University of Beira Interior, 6201-001 Covilhã, Portugal

Yu Zhou
Jiangxi Province Key Laboratory of Water Ecological Conservation in Headwater Regions, Jiangxi University of Science and Technology, Ganzhou 341000, China

---

Research Article

Keywords: GNSS time series prediction, DA-GAT-BiLSTM, Direction Loss, KNN station graph, Bayesian optimization

Posted Date: March 3rd, 2026

DOI: https://doi.org/10.21203/rs.3.rs-9008300/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: The authors declare no competing interests.

A Graph Attention Framework with Kinematic Constraints for Network-Based GNSS Time Series Prediction

Siyu Zhu[1], Gaël Kermarrec[2], Xiaoxing He[3], Shengdao Wang[4], Rui Fernandes[5]

Zhengkai Huang[1,*], Jean-Philippe Montillet[5], Yu Zhou[3]

[1] School of Transportation Engineering, East China Jiaotong University, Nanchang 330013, China
[2] Institute of Meteorology and Climatology, Leibniz University Hannover, Germany
[3] Jiangxi Province Key Laboratory of Water Ecological Conservation in Headwater Regions, Jiangxi University of Science and Technology, Ganzhou 341000, China
[4] Division of Geodetic Science, School of Earth Sciences, The Ohio State University, Columbus, Ohio 43210, USA
[5] Institute Dom Luiz, University of Beira Interior, 6201-001 Covilhã, Portugal
* Corresponding authors: zhkhuang@whu.edu.cn(Z.H).

Abstract: Predicting GNSS displacement time series accurately is challenging due to complex spatial dependencies, multi-source noise, and hyperparameter tuning difficulties. Based on this, We propose a Bayesian-optimized DA-GAT-BiLSTM model with a kinematically-motivated Direction Loss. The framework combines a KNN-based geodesic station graph with graph attention for spatial encoding, BiLSTM for temporal modelling, and a direction-aware head for coupled E/N/U prediction. A kinematically-motivated Direction Loss enforces smoothness and cross-component consistency, with Optuna-based optimization for efficient hyperparameter selection, aiming to enable a novel network-based prediction through inter-station information sharing. Using data from 100 stations worldwide (2000–2024), spatial autocorrelation (Moran's I = 0.550/0.625/0.585 for E/N/U) supports a sparse kNN geodesic prior (k = 10), enabling DA-GAT-BiLSTM to deliver millimetre-level E/N/U prediction (R^2 = 0.948/0.964/0.937; mean residuals -0.11/-0.04/0.14 mm for E/N/U). Ablations confirm that graph attention is critical (MAE +72–76%; R^2 −9--14% without GAT), the direction-aware head improves robustness—especially vertically (MAE +40–45% without DA)—and the kinematically-motivated Direction Loss outperforms MSE on high-quality stations (MAE −23–37%, RMSE −24–38%, R^2 +0.09–0.14). Optuna achieves a lower validation loss (0.104 in 12.5 h) than grid (0.145 in 48.5 h) or random search (0.118) Comparison with VARIMA, GCN-LSTM, and GAT-Transformer shows that DA-GAT-BiLSTM achieves lower prediction errors across displacement components, with performance differences corresponding to the presence of adaptive spatial attention, bidirectional temporal encoding, and direction-aware cross-component modelling. The proposed framework offers a reliable tool for crustal deformation monitoring and early-warning applications.

Keywords: GNSS time series prediction, DA-GAT-BiLSTM, Direction Loss, KNN station graph, Bayesian optimization

1. Introduction

Understanding how tectonic stress accumulates and releases across fault systems requires precise, long-term observations of ground motion. High-precision GNSS networks now provide three-dimensional displacement measurements at thousands of stations worldwide, generating dense time series that capture everything from slow tectonic creep to sudden co-seismic jumps (Kreemer et al., 1998; He et al., 2025; Sielfeld et al., 2019). These data have proven invaluable for earthquake early warning, hazard assessment, and infrastructure monitoring (Uzel et al., 2013; Ducrocq et al., 2021).

Yet extracting meaningful predictions from GNSS time series is far from straightforward. Crustal motion exhibits strong spatial correlations—neighbouring stations often share a common tectonic driver—but also displays highly heterogeneous local behaviour due to differences in geology, fault geometry, and loading conditions (Li et al., 2022; Zhu et al., 2023; Peng et al., 2025). The signals themselves are contaminated by instrumental noise, atmospheric delays, and seasonal hydrological loading, making it difficult to separate genuine tectonic motion from spurious fluctuations (Ruhl et al., 2017; Santamaría-Gómez et al., 2011; Zhou et al., 2025).

Traditional approaches to this problem have limitations. Spatial interpolation techniques like kriging are linked with computational costs at large scales and cannot easily adapt to complex, anisotropic dependencies (Du & Haider, 2025; Barbosa et al., 2017). Classical time-series models such as ARIMA handle temporal patterns well but ignore the spatial structure of station networks entirely (Althobaiti, 2025). In fact, a closely related limitation is that many practical prediction pipelines still adopt a single-station formulation—predicting each station (and often each component) independently—even when dense networks are available. This prevents models from borrowing strength from spatially correlated deformation signals across nearby stations, and can yield spatially inconsistent forecasts at the network level; moreover, the "train-and-tune per station" workflow scales poorly as the number of stations grows (Alevizakou et al., 2018; Gao et al., 2022; Chen et al., 2023). Physics-based numerical simulations, while interpretable, depend heavily on simplifying assumptions and idealized boundary conditions that rarely hold in practice (Xiong et al., 2022).

Machine learning offers a promising alternative, and recent work has explored neural networks for displacement prediction with some success (Wang et al., 2021; Gade et al., 2021; Wang et al., 2024). However, most existing methods still rely on hand-crafted features and heuristic spatial weighting schemes, making it hard to capture the full complexity of station interactions(Wang et al.,2025;Abdalla et al.,2024). Moreover, the three displacement components exhibit fundamentally different characteristics: horizontal motions (east and north) typically display stronger seasonal loading patterns and inter-component correlation driven by regional tectonic strain, whereas vertical displacements are more susceptible to local hydrological effects and atmospheric noise. They also tend to treat the three displacement components independently, ignoring the physical coupling between horizontal and vertical motion. In addition, uncertainty quantification remains largely absent, and hyperparameter tuning is often performed through computationally expensive trial-and-error (Lee and Seok, 2021; Meenakshi et al., 2022; Tabar & Sisman, 2025).

In this paper, we propose a different approach. Our DA-GAT-BiLSTM framework models the GNSS station network with graph attention to learn data-driven spatial dependencies, while bidirectional recurrent layers capture temporal dynamics. We further incorporate a direction-aware mechanism to account for the distinct yet coupled behaviour of horizontal and vertical displacement components. To improve physical plausibility, we introduce a kinematically-motivated loss that promotes smooth, consistent motion across time and components. Finally, Optuna-based Bayesian optimization enables efficient, systematic hyperparameter tuning.

The remainder of this paper is organized as follows. Section 2 introduces the GNSS time-series data sources and the preprocessing procedures. Section 3 details the proposed spatial network-based prediction framework, including the spatiotemporal modelling architecture, the kinematically-motivated loss design, and the Bayesian hyperparameter optimization strategy. Section 4 reports the experimental results and comparisons with baseline methods. Section 5 discusses the main findings, summarizes the conclusions, and outlines directions for future work.

2. Data source and preprocessing

2.1. GNSS Time Series Data

We obtained displacement time series from the Nevada Geodetic Laboratory (NGL), which provides standardized, multi-source fusion products derived from global GNSS networks (Blewitt et al., 2018). The dataset incorporates observations from the Network of the Americas (NOTA), Japan's GEONET, and numerous continuously operating reference station networks contributed by international partners. Processing follows NGL's latest workflow based on GipsyX v2.3 software, with solutions aligned to ITRF2020. Precise orbit and clock products come from NASA's Jet Propulsion Laboratory, and ocean tide-loading corrections use the FES2014b model (Bertiger et al., 2020).

From this database, we selected 100 stations for analysis. Rather than imposing a strictly uniform spatial distribution, we prioritized data quality: stations needed long, nearly continuous records from 2000 to 2024 without major gaps, equipment changes, or unstable coordinate behaviour. In densely instrumented regions like North America and East Asia, we thinned the network using distance-based criteria to avoid excessive clustering while retaining the most complete records. In sparsely covered areas such as South America and Asia, we kept all stations meeting our quality thresholds. The resulting network spans diverse tectonic environments (plate boundaries, active deformation zones, and stable cratonic interiors) (see Figure 1).

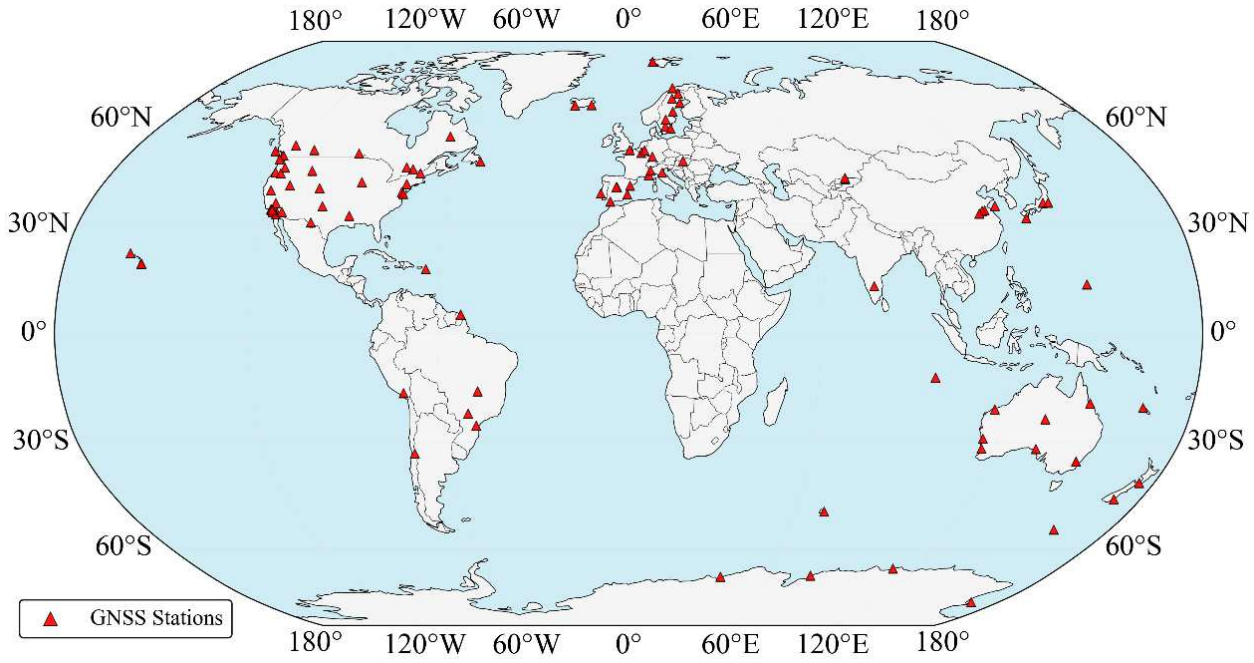


Fig.1. Spatial distribution of 100 GNSS stations, highlighted in red triangles.

2.2 Data preprocessing

Before modeling, outliers were identified using an IQR-based rule applied to the displacement values (per station and per component), with thresholds estimated from the corrected training split(Swami et al., 2023). Data gaps were handled only to maintain a regular time index for batching: Interpolation was used selectively for short interruptions to ensure continuity without over-smoothing longer interruptions(Zhao et al., 2026). Importantly, the train/validation/test split was defined chronologically before any parameter-estimating preprocessing (e.g., outlier thresholds, harmonic coefficients, or normalization/statistical summaries), and these quantities were estimated on the training period only and then applied to validation/test.

Figure 2 shows raw and processed time series for station 0OSK on Sweden's Baltic coast. The raw data contain intermittent high-amplitude excursions, especially in the vertical component. The shaded background and the vertical dashed lines indicate the dataset split into training (2000–2015), validation (2015–2020), and test (2020–2024) periods. The right panels quantify short-timescale variability by plotting the rolling standard deviation (Rolling STD, in mm) through time for the raw (dashed) and processed (solid) series, thereby highlighting changes in noise level while preserving the secular deformation trends across all splits. Overall, the processed series exhibits consistently lower Rolling STD than the raw series, indicating reduced short-term noise; the reductions are summarized in the annotations as noise decreasing by 9.6% for the east component, 9.7% for the north component, and 6.1% for the up component.

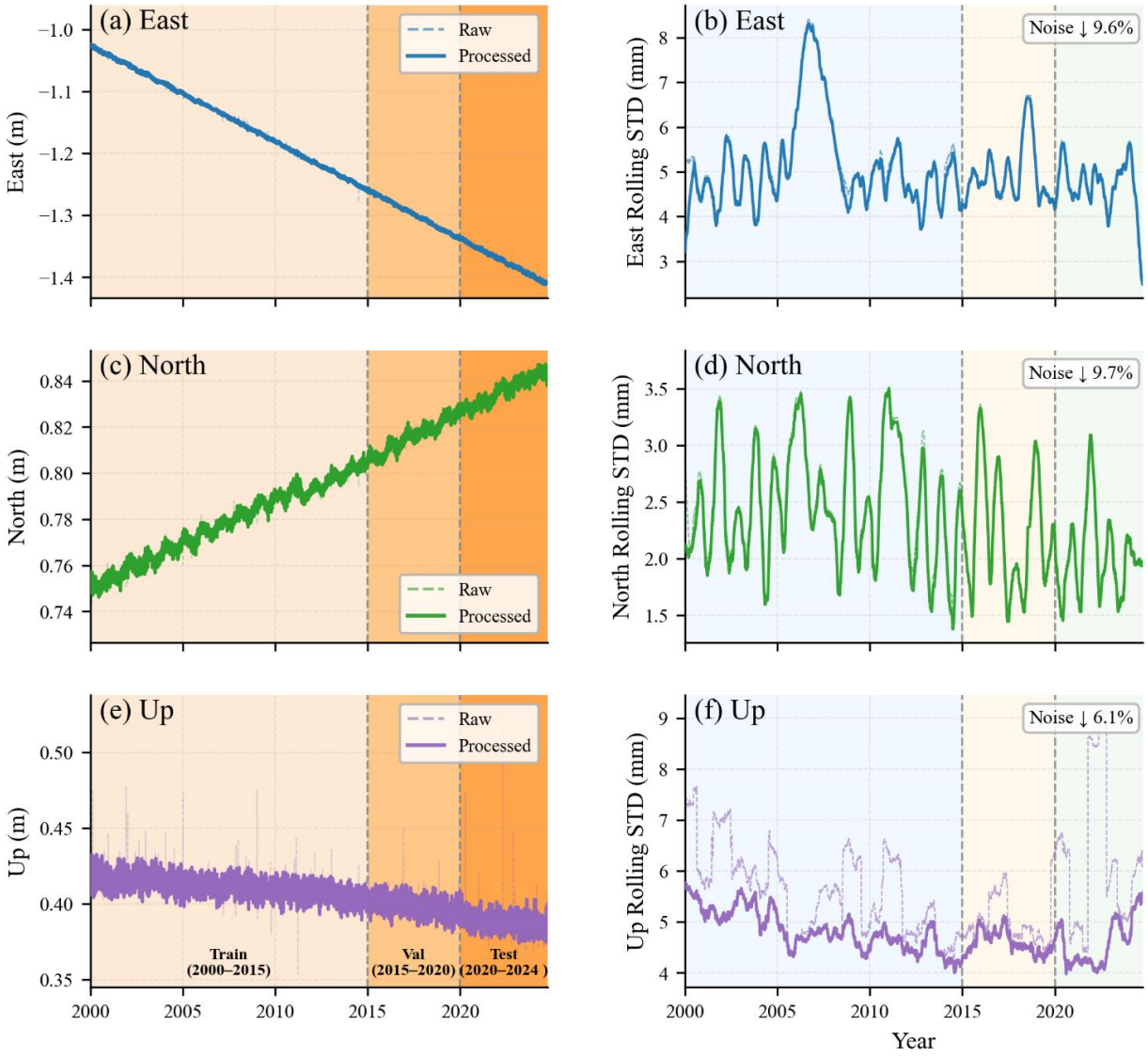


Fig.2. Preprocessing effects and rolling-noise comparison of three-component GNSS displacement time series at station 0OSK (2000-2024; chronological split with 2020-2024 held out for testing).

3 Spatial Network-Based Framework for GNSS Time Series Prediction

3.1. Modelling Framework Overview

Based on pre-processed data, our prediction framework integrates spatial network topology with temporal dynamics through a graph neural architecture (Fig.3). A k-nearest-neighbour station graph encodes inter-station proximity via edge weights based on haversine distances. Input features—temporal descriptors (rolling statistics, velocity/acceleration, trends, seasonal encodings) and spatial attributes (clustering, coordinates)—pass through projection layers into multi-head GAT with residual connections, followed by BiLSTM for temporal

dependencies.(Deng et al., 2025). Temporal attention refines sequence representations, and a direction-aware head generates separate 7-day forecasts for East, North, and Up displacements.

This graph-informed spatiotemporal architecture incorporates interrelated design considerations shaping its spatiotemporal learning:

(1) To account for spatial correlations among neighboring stations sharing common tectonic drivers, we design GAT layers that propagate information across the network through adaptive attention, dynamically weighting neighboring contributions.

(2) Recognizing distinct horizontal versus vertical motion characteristics, we introduce a direction-aware head employing dedicated encoders with gated aggregation to preserve directional signatures.

(3) To encourage physical plausibility beyond standard regression metrics, we incorporate a composite loss integrating velocity consistency, temporal smoothness, drift correction, and cross-directional covariance alongside prediction error.

(4) Given that hyperparameter tuning remains labor-intensive, we employ Bayesian optimization (Optuna/TPE) to automatically tune GAT depth, attention heads, LSTM capacity, and loss weights, balancing spatial-temporal-physical objectives.

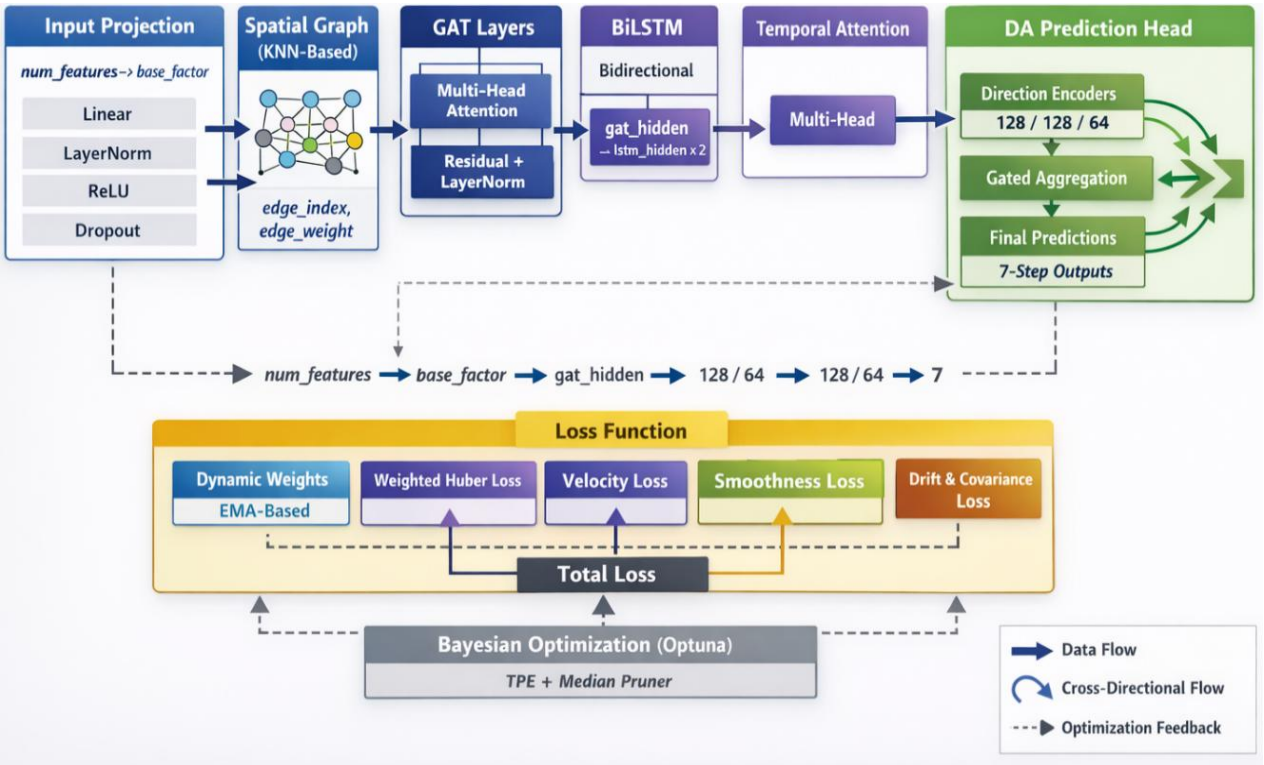


Fig.3. Overview of the DA-GAT-BiLSTM Framework for Network based GNSS Displacement Prediction

3.2 Spatial Graph Construction for Network-Scale Prediction

Distant stations tend to reflect site-specific geological processes, whereas proximal stations are more governed by shared tectonic forcing. A key limitation of existing GNSS displacement prediction approaches is their treatment of each station as an independent prediction unit, thereby neglecting the spatial dependence and autocorrelation among stations inherent to crustal deformation (Broelemann et al., 2017). To address this bottleneck, we construct a spatial graph that enables information sharing across stations; weight sharing reduces the number of model parameters, while explicitly incorporating spatial contextual information into both training and prediction.

3.2.1 K-Nearest Neighbour Graph Construction

We model the GNSS network as a graph G = (V, E), where each node represents one station and edges connect spatially proximate stations. For N stations with coordinates (latitude, longitude), we construct edges using k-nearest neighbour connectivity(Duan et al., 2026).

To establish edge connectivity, we apply k-nearest neighbour selection based on geodesic distance. The Haversine distance between stations i and j accounts for Earth's spherical geometry:

$$\mathrm{d}_{ij} = 2R\arcsin\left(\sqrt{\sin^2\left(\frac{\Phi_j-\Phi_i}{2}\right)+\cos\left(\Phi_i\right)\cos\left(\Phi_j\right)\sin^2\left(\frac{\lambda_j-\lambda_i}{2}\right)}\right) \tag{1}$$

Where R = 6371 km, φ denotes latitude in radians, and λ denotes longitude in radians.

We implement this calculation using scikit-learn's Nearest Neighbors class with ball tree algorithm and haversine metric. For each station i, we identify k nearest neighbors and create directed edges to these neighbors. This process generates approximately k edges per node, yielding total edge count near kN for N stations. If distance filtering eliminates all edges for a station, we construct connections to ensure graph connectivity.
Edge weights encode spatial proximity. For each edge (i, j) connecting station i to neighbor j, we define weight as:

$$w_{ij} = \frac{1}{d_{ij}+\varepsilon} \tag{2}$$

where $\varepsilon = 10^{-6}$ prevents division by zero. This inverse-distance weighting assigns higher influence to closer neighbors and lower influence to distant neighbors. The weights remain constant during training since station positions are fixed. The resulting graph structure serves as input to graph attention layers that aggregate spatial information.

3.2.2 Spatial Autocorrelation Validation

We quantify spatial dependencies using Moran's I statistic. For displacement component x, Moran's I is:

$$I = \frac{N}{\sum_i \sum_j w_{ij}} \frac{\sum_i \sum_j w_{ij} \left(x_i - \bar{x}\right)\left(x_j - \bar{x}\right)}{\sum_i \left(x_i - \bar{x}\right)^2} \tag{3}$$

where N is the number of stations, w_{ij} are edge weights from the KNN graph, and $\bar{x}$ is the mean displacement.

Values of I range from -1 to 1. Positive values indicate spatial clustering (similar values concentrate in space), values near zero suggest spatial independence, and negative values indicate spatial dispersion(Z. Wang et al., 2026).

3.3. DA-GAT-BiLSTM Architecture

We ingest spatiotemporal observations organised along four axes—batch, station, time, and feature—and construct a KNN-based graph to specify candidate inter-station connections via the edge index. A stack of GATv2 layers then performs spatial aggregation, progressively enriching each station's representation from local neighbourhood cues to broader network-level dependencies. The spatially encoded signals are passed to a bidirectional LSTM, where forward and backward recurrent streams jointly capture temporal dependencies in both directions; residual connections are used to stabilise optimisation and preserve information flow. We further apply multi-head attention over the temporal representations to upweight salient time steps. Finally, a direction-aware module routes the shared features into direction-specific pathways and prediction heads, producing three coordinated outputs corresponding to the east, north, and vertical components, thereby improving directional consistency. (Figure 4).

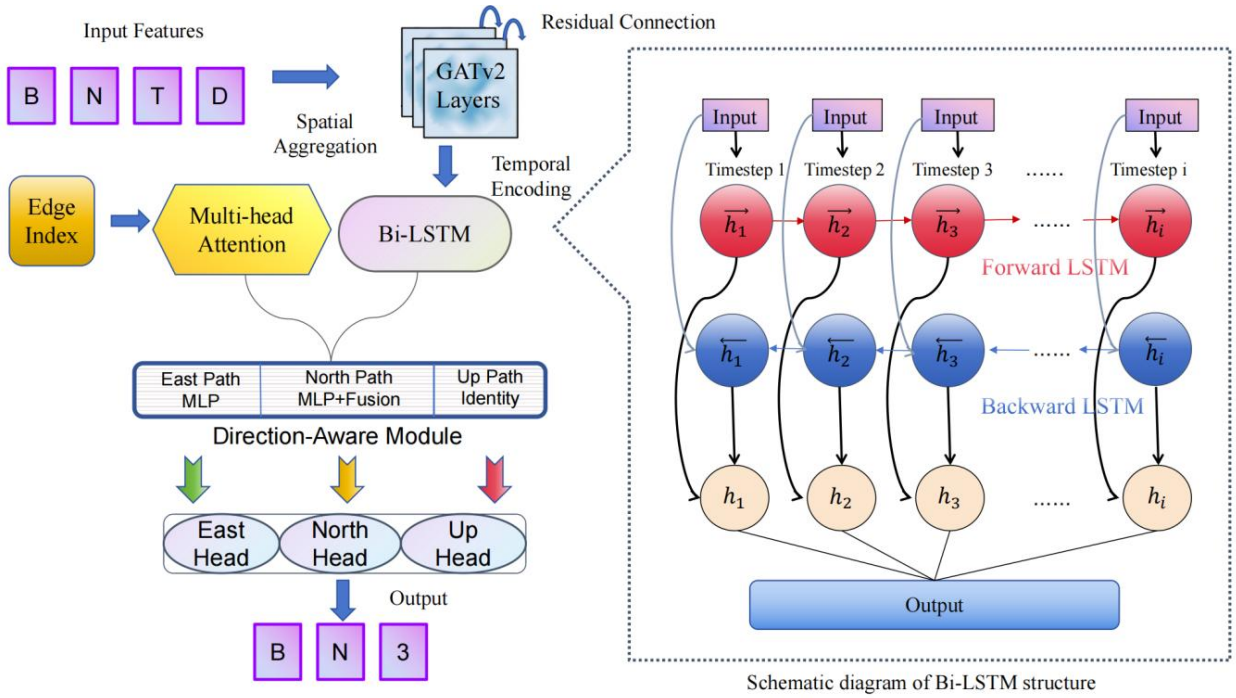


Fig. 4. Architecture of the DA-GATv2-BiLSTM Model

3.3.1 Graph Attention Networks (GAT)

Based on spatial graph, we use a three-layer cascaded GATv2Conv architecture to model intricate spatial dependencies. This design captures multi-scale spatial interactions and improves the model's ability to process complex, heterogeneous geological environments. A dynamic attention-weighting mechanism is incorporated, and the node representation update at the k-th GAT layer is defined as:

$$h_i^{(k)} = \Big\|_{m=1}^{M} \sigma\left(\sum_{j \in N_i} \alpha_{ij}^{(m)} W^{(m)} h_j^{k-1}\right) \tag{4}$$

$$\alpha_{ij}^{(m)} = soft\max_j\left(Leaky\,\mathrm{Re}LU\left(\alpha^{(m)T}\left[W^{(m)}h_i \Big\| W^{(m)}h_j\right]\right)\right) \tag{5}$$

where a is the learnable attention parameter, || represents multi-head attention concatenation, W is the learnable weight matrix, σ is the SILU activation function, and $\alpha_{ij}^{(m)}$ is the normalized attention weight from the m-th attention head.

The first GATv2Conv layer extracts local spatial features through dynamic attention; the second layer aggregates higher-order neighbourhood information to enlarge the receptive field; and the third layer integrates global spatial relationships, outputting node embeddings with enhanced representational capacity. Through this three-layer stacked GAT, the model generates spatially aware node embeddings $H_{spatial} \in \mathbb{R}^{NxDgat}$ (where N denotes the number of stations).

Throughout the process, Layer Normalization and residual connections are applied at each layer to stabilize training:

$$H^{(k)} = \Phi\left(H^{(k-1)}, A\right) + \rho\left(H^{(k-1)}\right) \tag{6}$$

3.3.2 Bidirectional Long Short-Term Memory (BiLSTM)

The spatially embedded representation $H_{spatial}$ feeds into a bidirectional LSTM that models sequential data to capture long-term dependencies within the displacement field. The core of the LSTM unit lies in its input gate i_t, forget gate f_t, output gate o_t, and memory cell C_t. The bidirectional LSTM processes the time series in both forward and reverse directions, then concatenates the outputs:

$$\overrightarrow{h_t} = \overrightarrow{LSTM}\left(x_1, x_2, ..., x_t\right) \tag{7}$$

$$\overleftarrow{h_t} = \overleftarrow{LSTM}\left(x_t, x_{t-1}, ..., x_1\right) \tag{8}$$

$$h_t = \left[\overrightarrow{h_t}; \overleftarrow{h_t}\right] \tag{9}$$

where $\overrightarrow{h}_t$ and $\overleftarrow{h}_t$ represent the hidden states of the forward and backward LSTMs at time step t, and h_t is the concatenated output. The final temporal feature is obtained by:

$$H_{(temp)} = \left[\overrightarrow{h_t}; \overleftarrow{h_t}\right] \in R^{N \times 2H_{lstm}} \tag{10}$$

We employ bidirectional processing to encode long-range dependencies within the historical input window by reading the window in both directions and concatenating the hidden states. Because the entire input window precedes the forecast timestamp, this bidirectional encoding does not use observations beyond the prediction horizon; it serves as a richer representation of the available history. A multi-head self-attention mechanism then enhances the model's sensitivity to the most informative time steps:

$$Attention\left(Q, K, V\right) = soft\max\left(\frac{QK^T}{\sqrt{d_k}}\right)V \tag{11}$$

3.3.3 Direction-Aware Prediction Head

Horizontal and vertical displacement components behave quite differently: the east and north signals typically show seasonal loading effects and mutual correlation, while the vertical component tends to be noisier with pronounced long-term trends. We handle this asymmetry via a direction-aware graph attention module that learns edge-wise directional weights from relative station coordinates and uses gated cross-direction refinement to adaptively modulate message passing across components.

To achieve adaptive modelling of coupling relationships among components, we designed a Direction-Aware (DA) attention gating mechanism embedded in the LSTM hidden-state update. Let h_t^E, h_t^N, and h_t^U denote the hidden state vectors of the east, north, and vertical components at time step t. The cross-directional interaction gate g_{ij} controls information flow from direction j to direction i:

$$g_{i,j} = \sigma\left(W_{ij}\left[h_t^j; h_{t-1}^j\right] + b_{ij}\right) \tag{12}$$

where W_{ij} and b_{ij} are learnable parameters, [;] denotes vector concatenation, and σ is the Sigmoid activation. The enhanced hidden state $\tilde{h}_t^i$ for direction i, after integrating information from other directions, is computed as:

$$h_t^i = h_t^i + \sum_{j \neq i} g_{ij} \cdot \left(V_{ij} h_t^i\right) \tag{13}$$

This mechanism operates in a purely data-driven manner without prior physical assumptions, enabling the model to automatically learn dynamic spatiotemporal dependencies among directions. Finally, predictions for the three components are produced through independent two-layer perceptron heads:

$$y_d = W_d^{(2)} \cdot LayerNorm\left(SiLU\left(W_d^{(1)} f_d + b_d^{(1)}\right)\right) + b_d^{(2)} \tag{14}$$

3.4. Kinematically-motivated Loss Function

Standard mean squared error treats all prediction errors equally, ignoring the physical relationships between displacement components and their temporal evolution. We designed Direction Loss to incorporate domain knowledge explicitly. The loss combines several terms:

$$L = L_{Huber} + \beta L_{drift} + \gamma L_{\text{cov}} + \delta L_{smooth} \tag{15}$$

$$L_{Huber}^{(d)} = \frac{1}{NT} \sum_{i=1}^{N} \sum_{t=1}^{T} \begin{cases} \frac{1}{2}\left(y_{i,t}^{(d)} - y_{i,t}^{(d)}\right)^2 & |y - y| \le a_d \\ a_d\left(|y - y| - \frac{1}{2} a_d\right) & otherwise \end{cases} \tag{16}$$

where the Huber loss L_{huber} imposes robust constraints on prediction error along direction d; the drift penalty $L^d{}_{ri} f_t$ enforces consistency of displacement changes between consecutive time steps; the covariance constraint L^{cov} minimizes error covariance across directions to suppress spurious correlations; and the smoothness constraint L_{smooth} regularizes predicted curves via second-order differences.

A dynamic weighting mechanism adaptively adjusts directional weights based on historical error trends:

$$w_d = \frac{\exp(z_d)}{\sum_d \exp\left(z_d\right)} \times \frac{1}{\varepsilon_d + \varepsilon_d} \times \eta_d \tag{17}$$

Direction Loss introduces kinematically-motivated constraints to enhance the model's capability in representing anisotropic responses and nonlinear dynamic processes of crustal motion.

3.5. Bayesian Hyperparameter Optimization

Deep learning models involve many hyperparameters—learning rates, dropout rates, layer sizes, scaling factors—and finding good combinations through grid search quickly becomes prohibitive. We adopted Bayesian optimization using the Optuna framework with Tree-structured Parzen Estimator (TPE) sampling (Amini et al., 2025). The algorithm builds a probabilistic model of how hyperparameters affect validation loss, then uses an Expected Improvement criterion to select promising configurations:

$$EI_t(x) = E_{p(f(x)|D_{1:t})}\left[\max\left(f(x) - f\left(x^+\right), 0\right)\right] \tag{18}$$

where f(x) represents the validation loss function, x^+ denotes the current best hyperparameter combination, and $D_{1:t}$ is the historical trial data distribution. This guided search converges much faster than exhaustive enumeration while achieving equal or better performance.

4. Results

4.1 Spatial structure of station motions

To characterise spatial coupling, we represent the 100 stations as a distance-weighted proximity network on the globe (Fig. 1d). The construction of this spatial graph requires specification of two parameters: the number of nearest neighbors (k) and the maximum separation distance. To determine the optimal k value, we conducted a systematic sensitivity analysis across $k \in \{5, 8, 10, 12, 15, 20\}$ (Table 1).

The results reveal a performance optimum at k=10. Prediction error measured by RMSE decreases from 2.775 mm at k=5 to 2.489 mm at k=10, then increases to 2.713 mm at k=15 and 3.426 mm at k=20. The corresponding R^2 values follow an inverse pattern, reaching a maximum of 0.852 at k=10. Network topology metrics show that k=10 provides 8.24 effective neighbors per station on average, with a network density of 0.824 and mean edge distance of 902.6 km. Smaller k values (k=5: 4.77 effective neighbors, 695.97 km mean distance) limit spatial information propagation, while larger k values (k=15: 11.14 effective neighbors, 1077.97 km mean distance; k=20: 13.68 effective neighbors, 1233.99 km mean distance) introduce longer-range connections that increase computational cost (1.627× and 2.197× relative to k=10, respectively) without performance gains. The k=8–12 range exhibits RMSE variation within 0.052 mm, indicating stability around the optimum.

Table1. Sensitivity analysis of the number of nearest neighbors (k) in spatial graph construction

K	Avg neighbors	Network density	Avg edge distance(km)	RMSE (mm)	MAE (mm)	R^2	Relative time(h)
5	4.77	0.954	695.974	2.775	2.109	0.812	0.435
8	7.02	0.875	844.044	2.524	1.918	0.845	0.765

10	8.24	0.824	902.596	2.489	1.892	0.852	1.134
12	9.44	0.786	969.127	2.541	1.931	0.841	1.244
15	11.14	0.743	1077.967	2.713	2.062	0.805	1.627
20	13.68	0.684	1233.994	3.426	2.604	0.784	2.197

Based on this analysis, each station is linked to its 10 nearest neighbours, subject to a maximum separation of 3000 km, producing a sparse but connected graph that prioritizes local neighbourhood interactions (Fig. 1d). Each station is linked to its 10 nearest neighbours, subject to a maximum separation of 3000 km, producing a sparse but connected graph that prioritises local neighbourhood interactions. The map view reveals regionally coherent neighbourhoods (for example across western North America, northern Europe and Australia), where dense short-range links dominate, while a smaller number of longer edges connect geographically isolated stations and bridge gaps between sparsely sampled regions. This structure provides an explicit representation of inter-station adjacency for network-level spatial analyses.

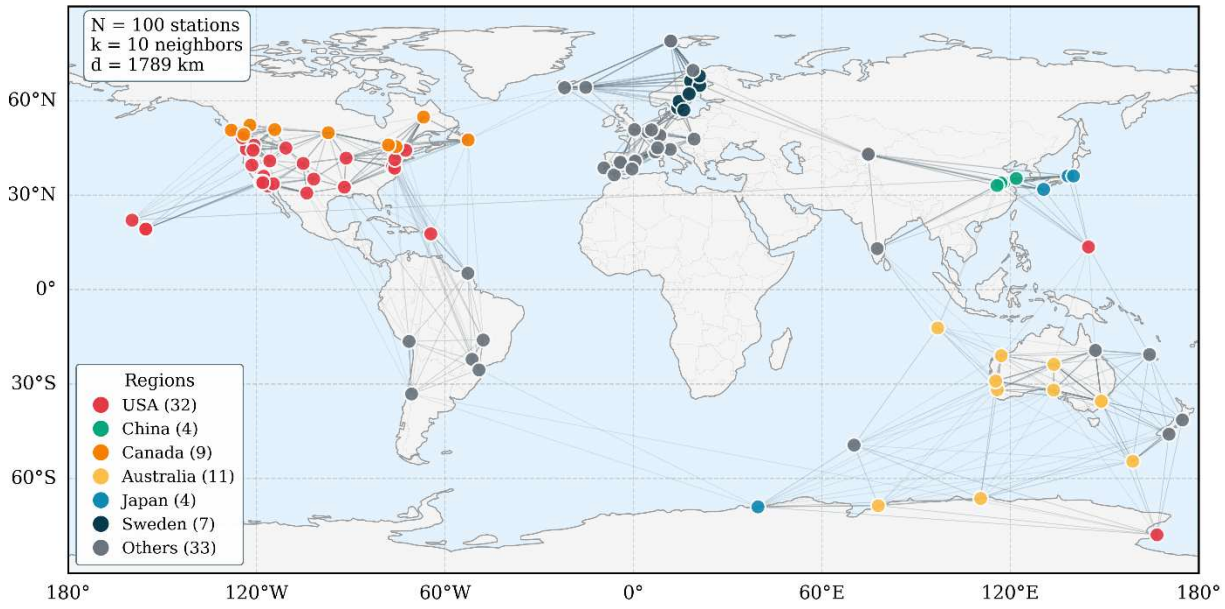


Fig.5. Spatial connectivity of the global GNSS network

Spatial dependence was then quantified using Moran's I (Fig.6 a–c). All three displacement components show moderate positive autocorrelation (East: I = 0.550; North: I = 0.625; Up: I = 0.585; n = 100), as reflected by the positive slopes in the Moran scatterplots between standardized displacement and spatial lag. Horizontal components exhibit stronger spatial coupling than the vertical component, consistent with their more coherent regional deformation signatures, whereas the vertical component is comparatively more heterogeneous and noise-sensitive.

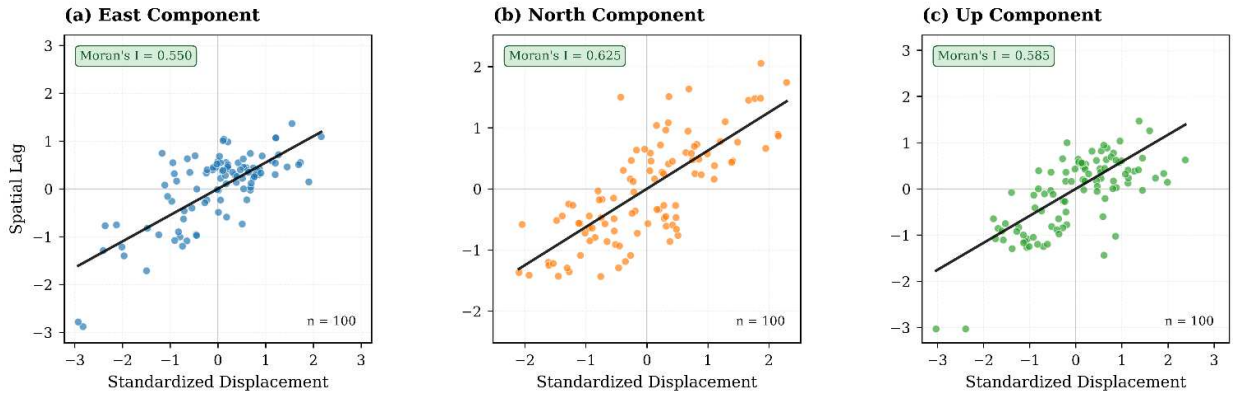


Fig.6. Moran's I of GNSS displacement components

4.2 Overall Model Performance

Prediction Accuracy. We evaluated DA-GAT-BiLSTM on the held-out test period from 2020 to 2024. As illustrated in Figure 7, the model demonstrates reasonable predictive performance across all three displacement components. The left panels show that predicted trajectories (dashed lines) generally follow the observed displacement curves (solid lines) for East, North, and Up components. The right panels display the residual distributions with Gaussian kernel density estimation, revealing relatively small mean errors: μ = -0.11 mm for East, μ = -0.04 mm for North, and μ = 0.14 mm for Up. The standard deviations are 2.51 mm, 2.18 mm, and 4.11 mm, respectively, indicating millimeter-level prediction accuracy with relatively low systematic bias. The approximately symmetric residual distributions suggest reasonable model reliability.

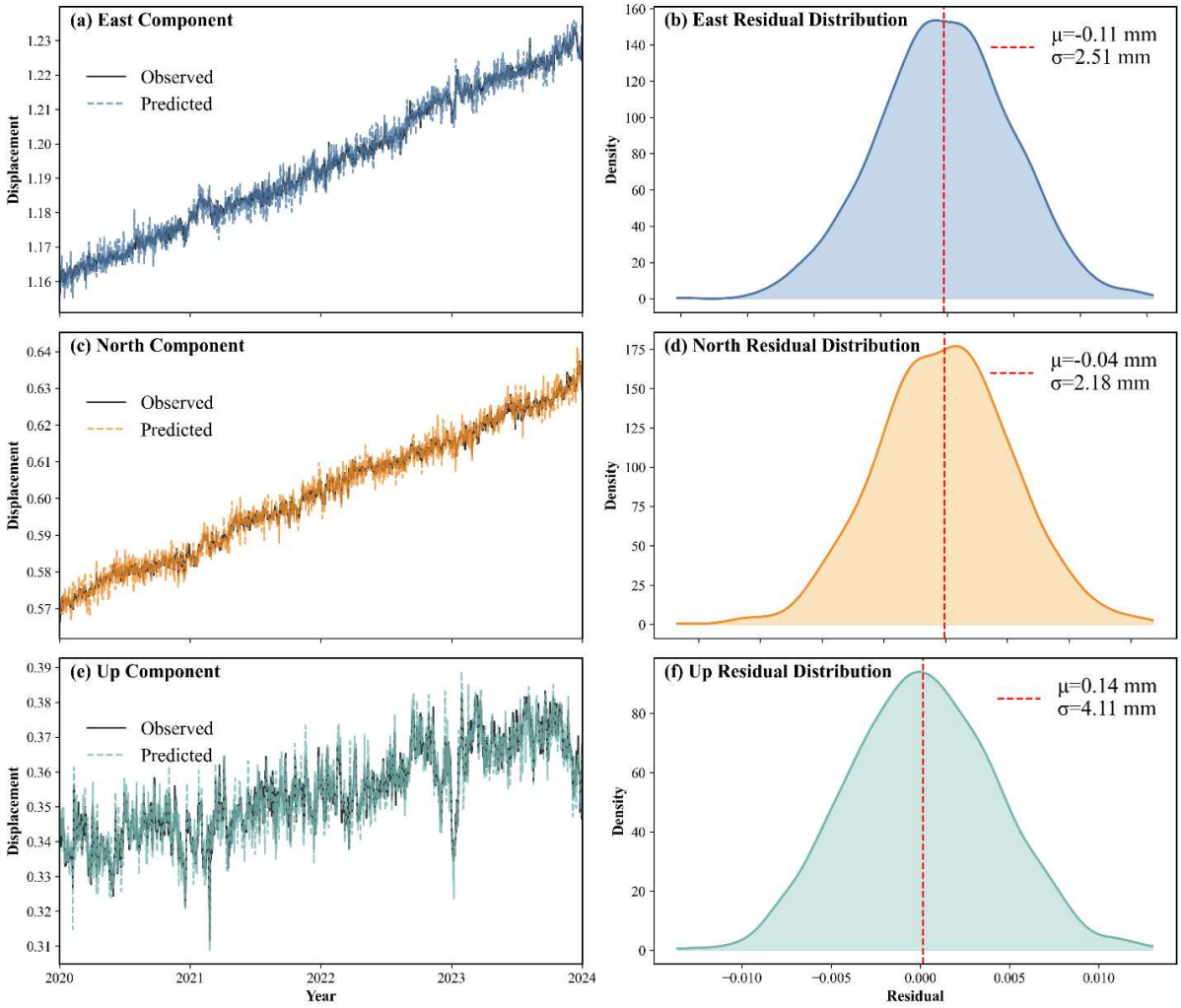


Fig. 7. Analysis of Predicted Trajectories and Error Distributions in Three Displacement Components Prediction

Figure 8 presents scatter plots of predicted versus observed values. All three components show linear relationships, with R^2 exceeding 0.93: specifically 0.964 for North, 0.948 for East, and 0.937 for Vertical. Data points cluster tightly around the ideal y=x line, with horizontal components showing particularly close agreement. The vertical component displays slightly more scatter, as expected given its lower signal-to-noise ratio, but performance remains excellent.

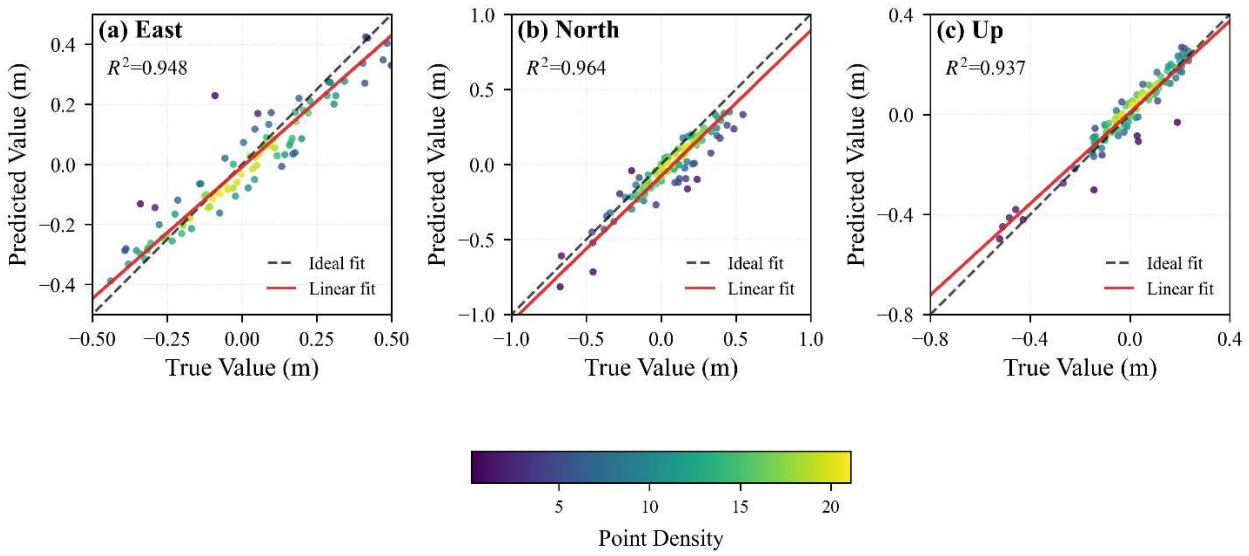


Fig.8. Evaluation of R²Values Across Three Displacement Components

Sensitivity to Data Quality and Window Length. Model accuracy depends on data completeness (Figure 9). Stations with less than 80% coverage show elevated MAE, while those exceeding 90% coverage achieve notably better predictions. This pattern reflects the intuitive expectation that more complete records provide richer training signal.

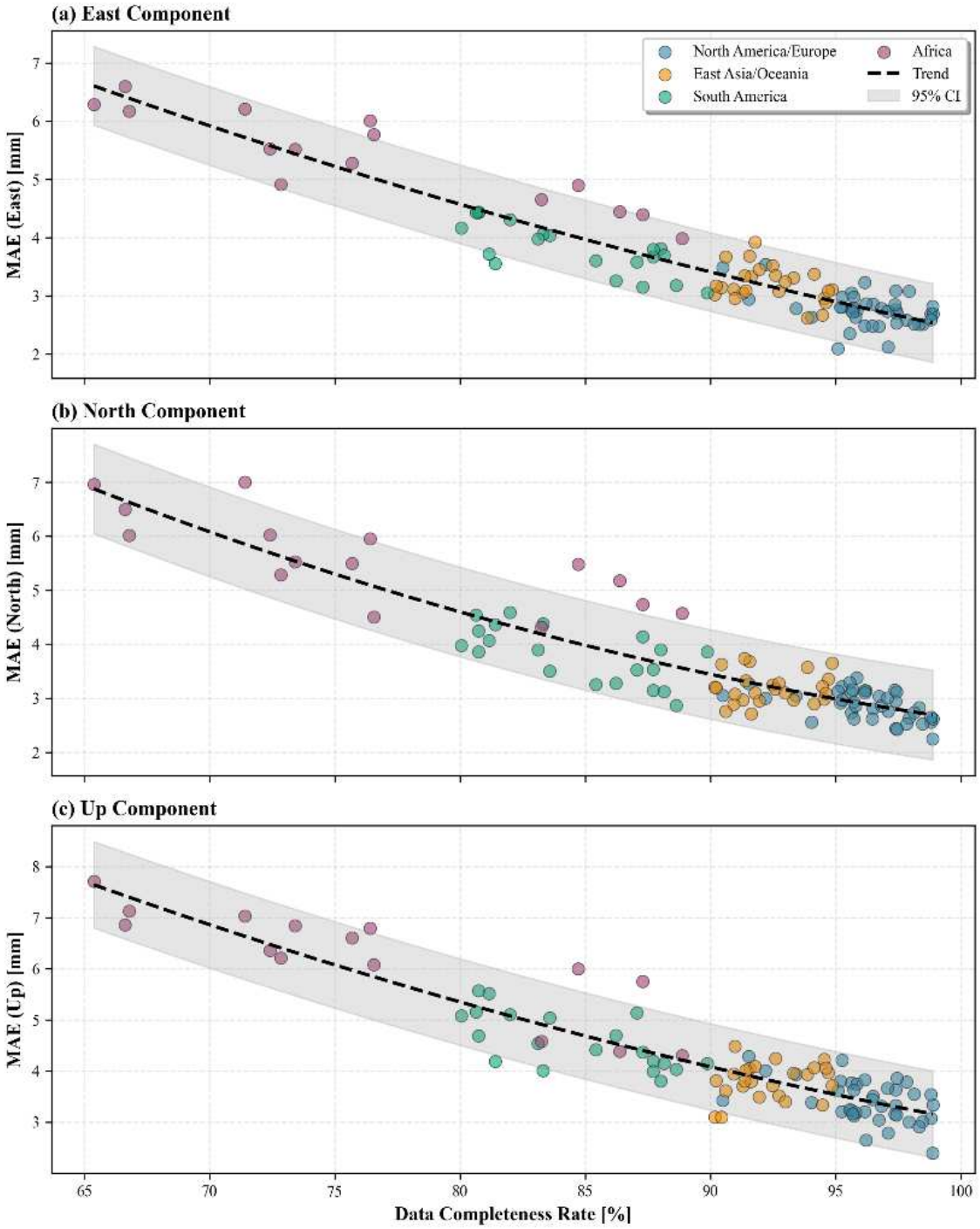


Fig.9. Impact of Data Completeness on MAE for East, North, and Up Components Across Regions

The choice of input window length also matters (Figure 10). Both MAE and RMSE decrease as the window extends, reaching minima around 60 days. Longer windows provide diminishing returns and eventually degrade performance, likely due to noise accumulation or overfitting. The 60-day window strikes a practical balance between capturing sufficient historical context and

maintaining computational efficiency. Training time and model complexity both grow substantially with longer windows.

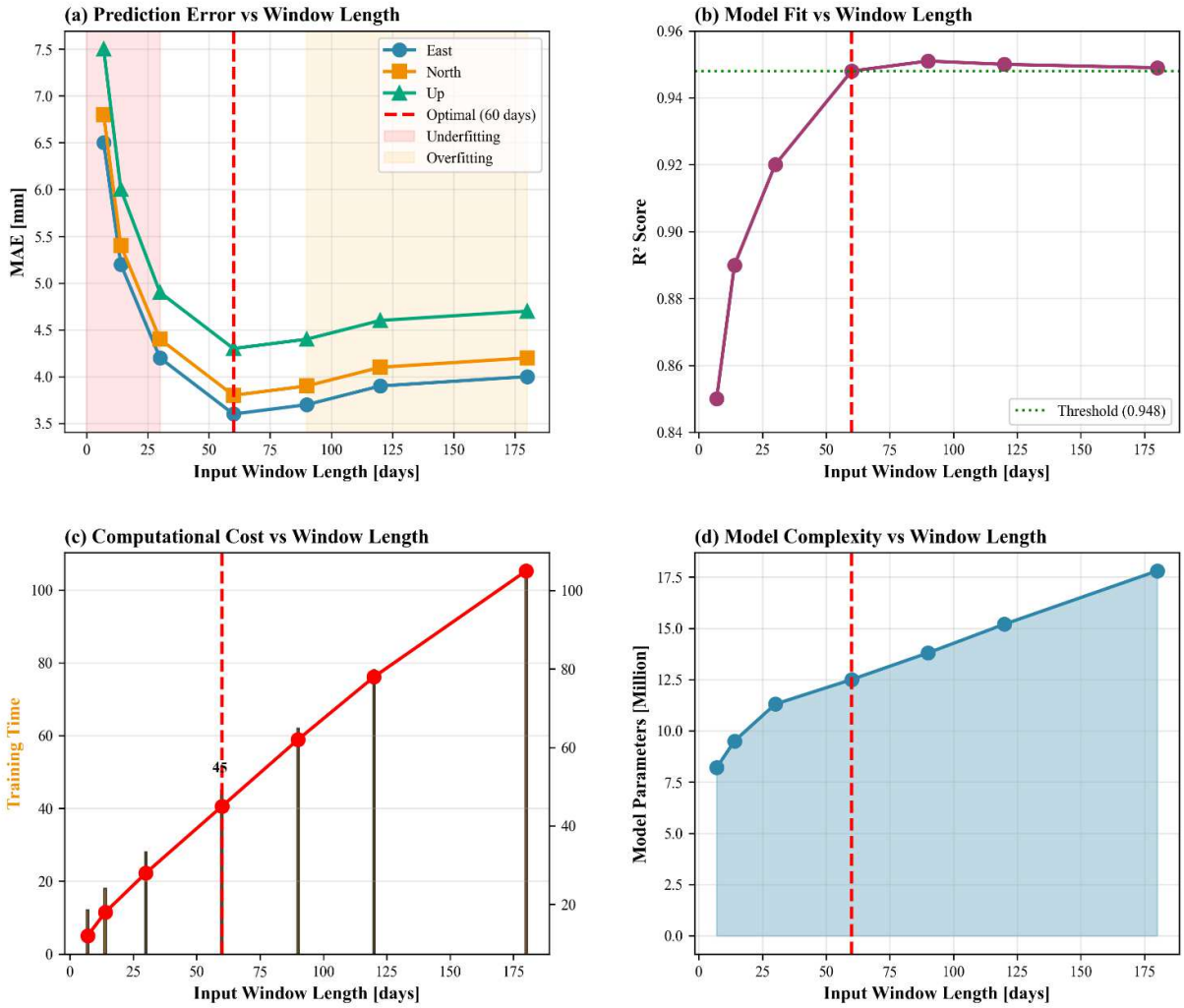


Fig.10.Effect of Input Window Length on Prediction Error, Model Fit, Computational Cost, and Model Complexity

4.3. Ablation Study

To understand how each component contributes to overall performance, we systematically removed individual modules and measured the degradation (Figure 11).

Contribution of network-based spatial modeling (GAT) Removing the GAT module leads to the most severe performance degradation among all ablations: MAE increases by 72–76% and R^2 decreases by 0.09–0.12 across the three components. This result highlights that explicit network-based spatial aggregation is central for GNSS displacement prediction at scale. In the full model, the GAT layers propagate information through the station graph and adaptively weight neighboring stations, allowing the predictor to borrow strength from spatially correlated deformation signals that share common tectonic drivers. Once GAT is removed, the model

effectively loses its ability to exploit these inter-station dependencies; consequently, forecasts become more susceptible to local noise and station-specific idiosyncrasies, and the learned dynamics cannot maintain the same level of spatial coherence. The consistent degradation across East–North–Up indicates that spatial context is not merely beneficial for one direction, but rather provides a network-level constraint that stabilizes prediction for all components.

Contribution of the direction-aware prediction head (DA) Ablating the direction-aware mechanism yields the second-largest drop in performance: MAE increases by 40–45%, with the strongest impact on the Up component. This pattern is consistent with the known asymmetry of GNSS motion characteristics: horizontal components often exhibit stronger mutual correlation and clearer seasonal signatures, whereas the vertical component is typically noisier and more trend-dominated. The direction-aware head in the full model is designed to preserve these directional signatures through component-specific pathways and gated cross-direction interaction, so that shared spatiotemporal features are selectively exchanged rather than indiscriminately mixed. When this mechanism is removed, the network must rely on a less structured representation that tends to blur the distinct statistical properties of horizontal versus vertical motion. As a result, the predictor shows a pronounced loss of fidelity for the Up component, suggesting that direction-aware disentanglement and controlled cross-component fusion are necessary to handle anisotropic noise and heterogeneous signal composition in ENU GNSS prediction.

Contribution of bidirectional temporal encoding (BiLSTM) Replacing bidirectional with unidirectional LSTM increases MAE by 31-42% across components and reduces R^2 by 0.04-0.05. The unidirectional variant summarizes the input window using only a forward recurrence and cannot exploit information from both early and late positions within the available historical window when forming its representation, which weakens long-range dependency capture and degrades forecast-trajectory consistency.

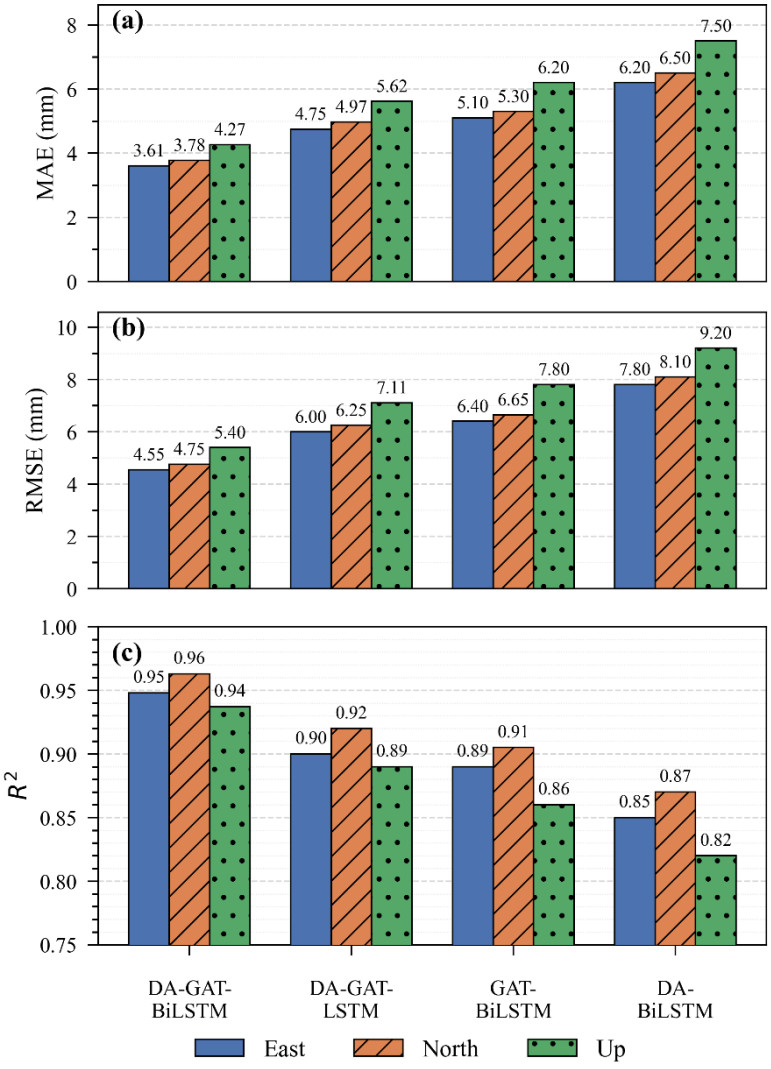


Fig.11.Ablation Study of DA-GAT-BiLSTM Variants for Positioning Performance

Overall, the ablation results demonstrate that all three components contribute meaningfully to the final performance, but with different roles: GAT-based spatial aggregation provides the dominant gain, while direction-aware modeling and bidirectional temporal encoding offer complementary improvements that are particularly important for handling cross-component heterogeneity and stabilizing temporal dynamics. Their combination yields a more complete representation of crustal deformation processes than any single module in isolation.

4.4. Impact of Direction Loss

Figure 12 compares displacement reconstructions from models trained with conventional MSE versus our Direction Loss. The MSE-trained model captures broad trends reasonably well but produces noticeable short-period oscillations, with the largest deviations occurring near peaks and troughs—particularly in the vertical component. Direction Loss training substantially attenuates

these high-frequency artifacts, yielding predictions that track the reference signal more faithfully across all components.

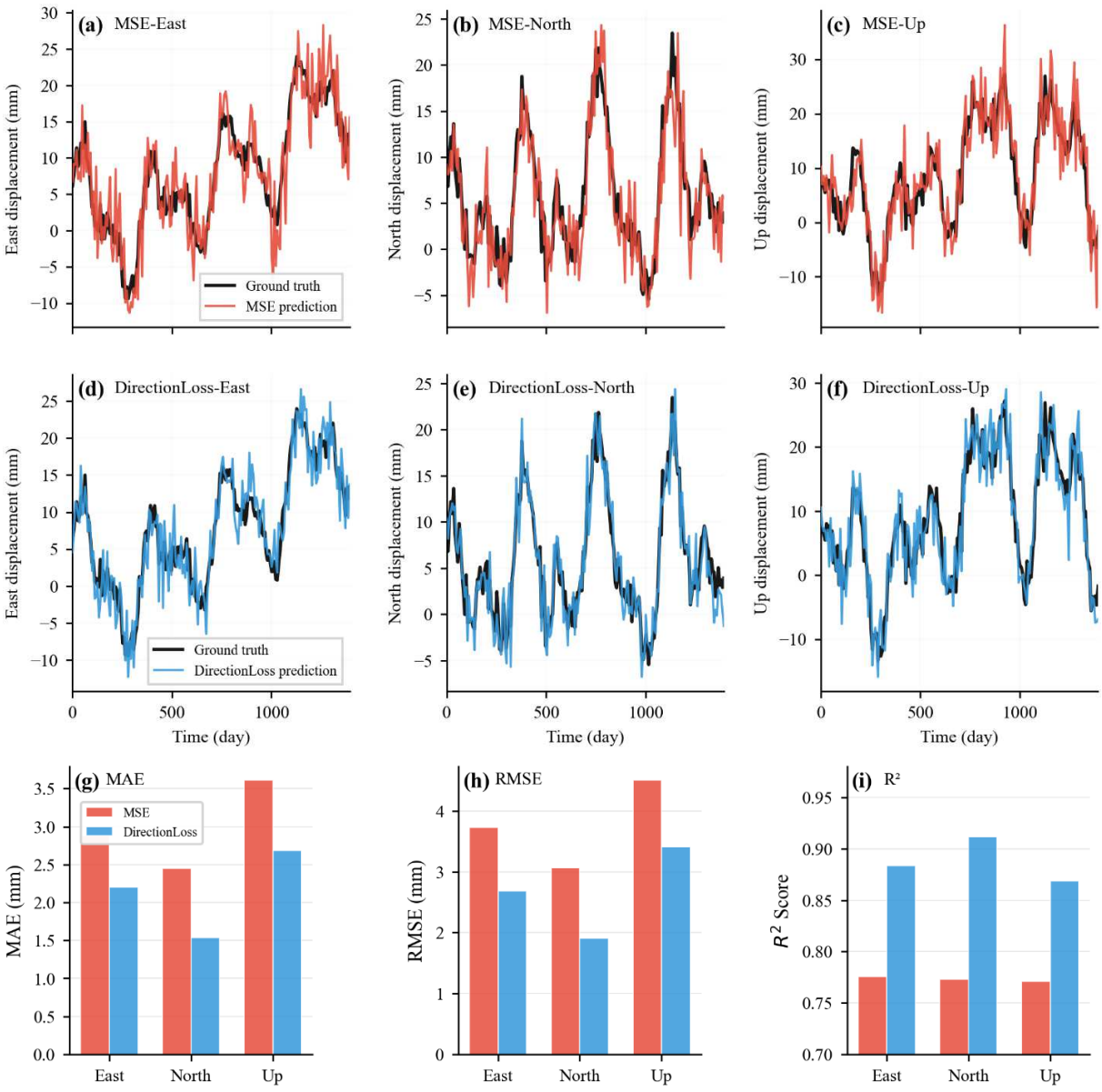


Fig. 12. Benchmarking a Direction-Aware Loss Function Against MSE for Multi-Component Displacement Prediction

Table 2 quantifies these improvements. Direction Loss reduces MAE by 23–37% and RMSE by 24–38% compared to MSE, while R^2 increases by 9–14 percentage points. Residual bias remains near zero in both cases, but the standard deviation of residuals drops from roughly 3.0–4.5 mm under MSE to about 1.9–3.4 mm under Direction Loss. The improvement stems primarily from reduced residual dispersion rather than a shift in mean error—exactly what we would hope from a kinematically-motivated constraint.

Table 2. Comparative performance of MSE and Direction Loss

Direction	Model	MAE (mm)	RMSE (mm)	R^2 Score	Bias (mm)	Std of residual	MAE (↓)	RMSE (↓)	R^2 Score (↑)
East	MSE	2.8827	3.7275	0.786	0.086	3.7265	22.57%	27.88%	12.42%
	Direction Loss	2.232	2.6882	0.883	-0.027	2.688			
North	MSE	2.4503	3.0634	0.773	-0.4551	3.0294	37.31%	37.62%	17.93%
	Direction Loss	1.536	1.9108	0.912	-0.1458	1.9052			
Up	MSE	3.6097	4.5034	0.769	0.2266	4.4977	25.53%	24.27%	12.95%
	Direction Loss	2.688	3.4105	0.869	0.0157	3.4105			

4.5. Hyperparameter Optimization

Table 3 compares three optimization strategies: grid search, random search, and Bayesian optimization with TPE. The Bayesian approach found the lowest validation loss (0.104) in just 12.5 hours—about a quarter of the time required by grid search (48.5 hours) and half that of random search (24 hours). Both conventional methods produced higher losses (0.145 and 0.118 respectively), representing 39% and 13.5% degradation relative to the Bayesian optimum. These results confirm that intelligent, guided search outperforms exhaustive enumeration in high-dimensional hyperparameter spaces.

Table 3. Comparison of Best-Performing Configurations from HPO Methods

Optimization Method	Base Factor	Gat Hidden	Lstm Hidden	Learning Rate	Dropout Rate	Val Loss	Total Optimization Time(Hours)
Grid Search	8	32	64	0.01	0.2	0.145	48.5
	8	64	128	0.001	0.3	0.132	(Fixed Budget)
	…	…	…	…	…	…	
Random Search	16	128	256	0.005	0.1	0.121	24.0
	32	64	512	0.0005	0.4	0.118	(Fixed Budget)

	…	…	…	…	…	…	
Bayesian Optimization(TPE)	64	256	512	0.002	0.25	0.104	12.5

4.6. Comparison with Baseline Methods

We benchmarked DA-GAT-BiLSTM against three representative baselines that progressively incorporate different components of our architecture:

(1) VARIMA: A traditional statistical baseline that models cross-station dependencies through multivariate time series analysis, capturing inter-station correlations implicitly via covariance structures rather than explicit spatial graphs. This model lacks all three core components of our framework: spatial graph topology , bidirectional temporal modeming, and direction-aware prediction heads.

(2) GCN-LSTM: A graph-based neural baseline that employs standard Graph Convolutional Networks (GCN) to capture spatial dependencies through k-nearest neighbour topology, followed by unidirectional LSTM for sequential temporal encoding. The model uses three independent fully connected layers to predict displacement components separately. This baseline integrates explicit spatial graphs but lacks adaptive attention mechanisms, bidirectional temporal context, and cross-component coupling.

(3) GAT-Transformer: An advanced architecture combining multi-head Graph Attention Networks (GATv2Conv) for adaptive spatial feature aggregation with Transformer encoders for parallel temporal modelling via self-attention. Despite these sophisticated components, this model employs independent prediction heads for each displacement direction without modelling their inherent correlations. This baseline includes adaptive spatial attention and temporal encoding, but lacks direction-aware cross-component modelling.

All deep learning baselines were trained using the same Direction Loss function, Bayesian hyperparameter optimization strategy, and data preprocessing pipeline as our proposed DA-GAT-BiLSTM to ensure fair comparison.

Error Distributions. Figure 13 shows box plot distributions of prediction errors. VARIMA exhibits median errors deviating from zero by -0.01 to 0.02 in horizontal components, with interquartile ranges spanning ±0.05. GCN-LSTM shows median values near zero but retains dispersion in the vertical component where outliers extend beyond ±0.04. GAT-Transformer achieves tighter distributions in East and North directions (IQR ≈ ±0.01), with larger spread in the Up component. DA-GAT-BiLSTM produces error distributions with median errors within ±0.005 of zero and IQR < ±0.01 across all three components.

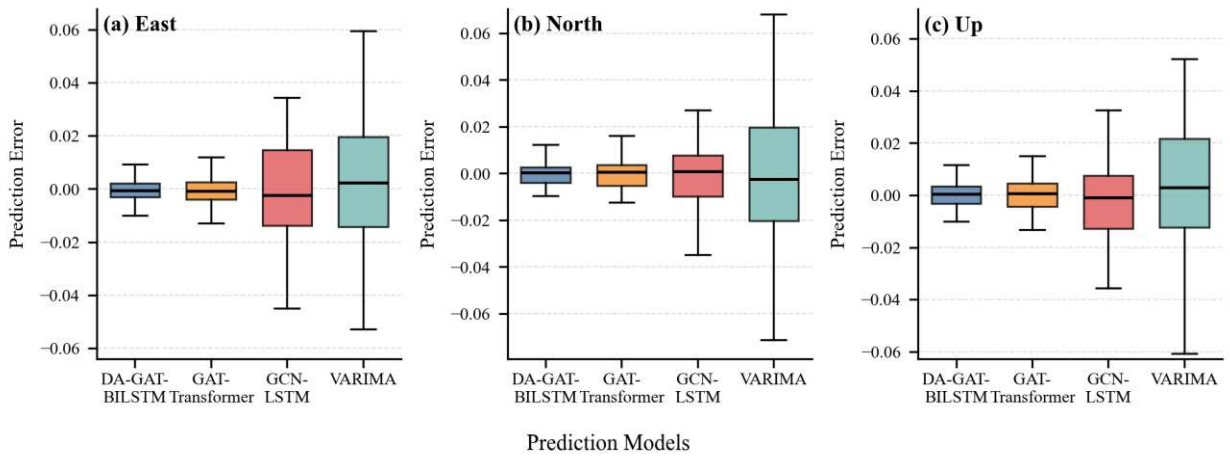


Fig.13. Comparative Analysis of Prediction Error Distributions Across Multiple Models and Spatial Dimension

Quantitative Metrics. Figure 14 presents violin plots of prediction error distributions. DA-GAT-BiLSTM achieves the best overall performance: eastward and northward MAE values and RMSE values both under 10 mm, and error distributions are highly concentrated and symmetric.

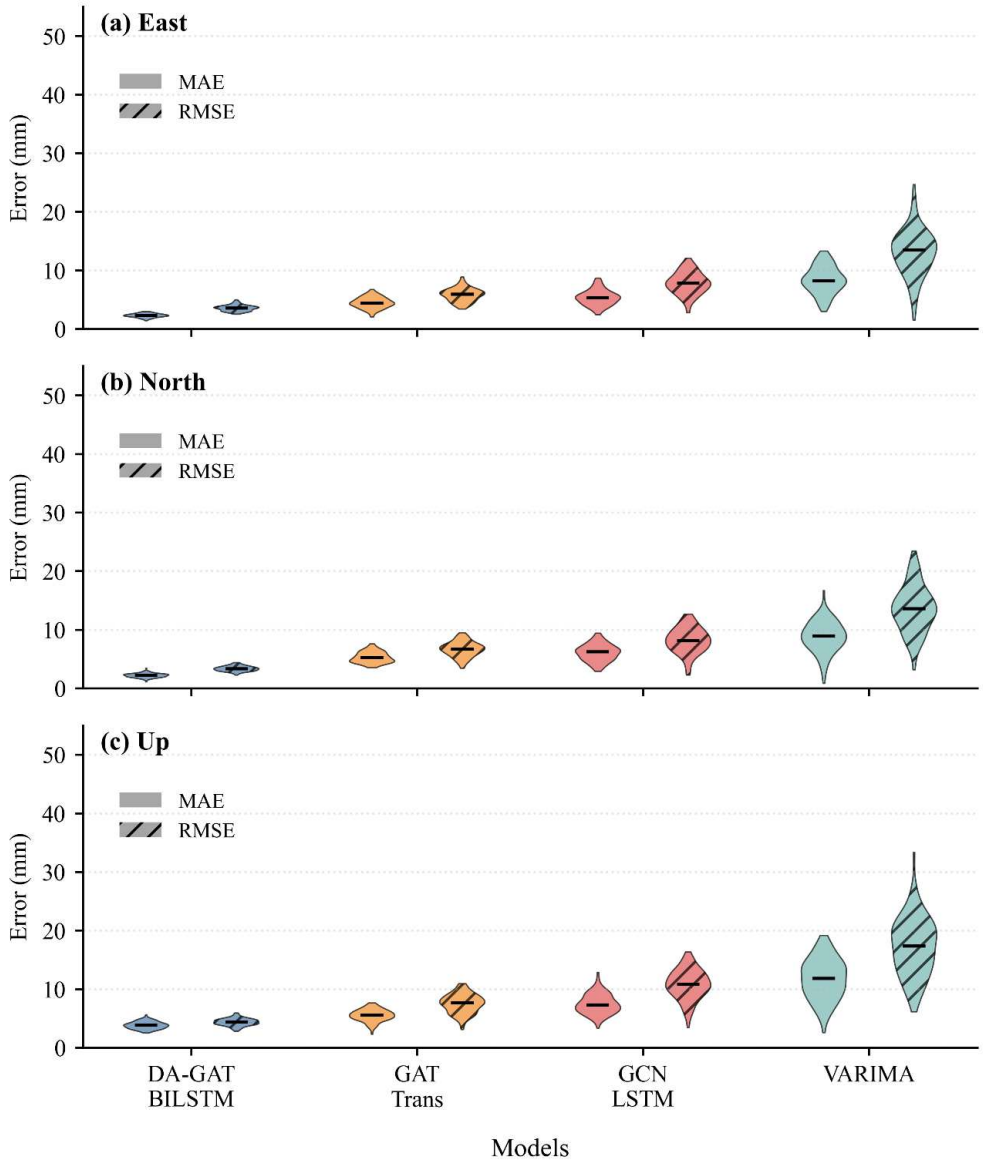


Fig.14. Comparative Methodological Assessment of Forecast Stability Across Geospatial Components

Table 4 further quantifies each model's time-series forecasting capability on the 2020–2024 test set by using MAE and RMSE across the three displacement components. DA-GAT-BiLSTM achieves MAE ranging from 2.18 to 3.85 mm and RMSE from 3.29 to 4.61 mm. GAT-Transformer shows MAE approximately 65-90% higher in horizontal components and RMSE values reaching 6.61-8.33 mm. GCN-LSTM exhibits MAE in the 5-7 mm range with RMSE exceeding 10 mm in the vertical component. VARIMA produces MAE above 8 mm across all components, with RMSE reaching 17.23 mm for the Up direction.

Table 4. Comparative Assessment of Error Metrics for Multi-Model ENU Geodetic Displacement Prediction

Models	East		North		Up	
	MAE (mm)	**RMSE (mm)**	**MAE (mm)**	**RMSE (mm)**	**MAE (mm)**	**RMSE (mm)**
DA-GAT-BILSTM	2.34	3.52	2.18	3.29	3.85	4.61
GAT-Transformer	4.45	5.8	5.05	6.61	6.35	8.33
GCN-LSTM	5.28	7.84	5.71	8.35	7.15	10.58
VARIMA	8.47	12.68	9.13	13.58	11.76	17.23

The results indicate that models lacking direction-aware prediction heads (GAT-Transformer) show moderate error increases, while models without adaptive spatial attention or bidirectional temporal encoding (GCN-LSTM, VARIMA) exhibit larger degradation in prediction accuracy.

5. Discussion and Conclusions

This work addresses challenges in GNSS displacement time-series prediction by introducing a KNN-based geodesic station graph with graph-attention learning for network based joint prediction, designing a spatiotemporal architecture that integrates graph-based spatial encoding with BiLSTM temporal modeling and a direction-aware prediction head to account for coupled E/N/U component behaviors, developing a Direction Loss function to constrain trajectory smoothness and cross-component consistency, and employing Bayesian optimization to reduce hyperparameter tuning costs.

Experimental results indicate that:

(1) **Network based joint prediction:** Inter-station spatial autocorrelation (Moran's I = 0.550/0.625/0.585 for E/N/U; n = 100) supports using a kNN geodesic graph prior (k = 10). DA-GAT-BiLSTM achieves mean residuals of −0.11/−0.04/0.14 mm and R^2 of 0.948/0.963/0.937 for E/N/U on the 2020–2024 test period.

(2) **Spatial graph & GAT:** Ablating GAT increases MAE by 72–76% and reduces R^2 by 0.09–0.12, indicating that graph-attention mechanisms contribute to modeling inter-station dependencies.

(3) **Direction-aware head (DA):** Removing the direction-aware head increases MAE by 40–45%, with larger increases in the vertical component, suggesting that cross-component modeling provides benefits for displacement prediction.

(4) **Direction Loss:** The Direction Loss reduces MAE by 23–37% and RMSE by 24–38% with R^2 gains of 0.09–0.14 compared to MSE loss on high-quality stations.

(5) **Hyperparameter optimization:** Bayesian optimization reaches validation loss of 0.104 in 12.5 hours compared to 0.145 (48.5 hours) for grid search and 0.118 for random search.

(6) **Model comparison:** DA-GAT-BiLSTM achieves lower errors than baseline methods across all displacement components. GAT-Transformer shows moderate error increases, while GCN-LSTM and VARIMA exhibit larger prediction errors. The performance gradient corresponds to the integration of adaptive spatial attention, bidirectional temporal encoding, and direction-aware cross-component modeling in the framework.

Funding: This work was supported by National Natural Science Foundation China (42364002), Outstanding Youth Fund of Jiangxi Natural Science Foundation (20252BAC220015), and Jiangxi Provincial Postgraduate Innovation Special Fund Project (YC2025-S524).

Code Availability: The computer code is available at:
https://github.com/SpaceGeodesyLab/DA-GAT-BiLSTM

Author Contributions: Conceptualization, Methodology, Writing—original draft, S.Z. and X.H; Supervision, Project administration, Writing—review & editing, G.K, S.W, R.F and J.M; Investigation, Data curation, Validation, Y.Z. All authors have read and agreed to the published version of the manuscript.

Data Availability: GNSS time series were obtained from the Nevada Geodetic Laboratory (http://geodesy.unr.edu/gps_timeseries). This study used data from 100 globally distributed stations.

Conflicts of Interest: The authors declare no competing interests.

References

Abdalla, A., Shami, S., Shahriari, M. A., & Azar, M. K. (2024). Estimation of land displacement in east baton rouge parish, louisiana, using InSAR: Comparisons with GNSS and machine learning models. The Egyptian Journal of Remote Sensing and Space Sciences, 27(2), 204–215. https://doi.org/10.1016/j.ejrs.2024.02.008

Alevizakou, E.-G., Siolas, G., & Pantazis, G. (2018). Short-Term and Long-Term Prediction for the ENU Point Position Changing by Using Artificial Neural Networks. ISPRS International Journal of Geo-Information, 7(3), 86. doi:10.3390/ijgi7030086

Althobaiti, Z. F. (2025). Improved prediction of carbon dioxide emissions using a hybrid SSA ARIMA model based on annual time series data in Bahrain. Scientific Reports, 15(1), 25699.

Amini, S., Vannieuwenhuyse, I., & Morales-Hernández, A. (2025). Constrained bayesian optimization: A review. IEEE Access, 13, 1581–1593.

Barbosa, A. R., Ribeiro, F. L. A., & Neves, L. A. C. (2017). Influence of earthquake ground-motion duration on damage estimation: Application to steel moment resisting frames. Earthquake Engineering and Structural Dynamics, 46(1), 27–49.

Bertiger, W., et al. (2020). GipsyX/RTGx, a new tool set for space geodetic operations and research. Advances in Space Research, 66(3), 469–489.

Blewitt, G., Hammond, W.C., & Kreemer, C. (2018). Harnessing the GPS data explosion for interdisciplinary science. Eos, 99.

Broelemann, K., Jiang, X., Mukherjee, S., & Chowdhury, A. S. (2017). Variants of k-regular nearest neighbor graph and their construction. Information Processing Letters, 126, 18–23. https://doi.org/10.1016/j.ipl.2017.06.001

Chen, H., Lu, T., Huang, J., He, X., Yu, K., Sun, X., Ma, X., & Huang, Z. (2023). An Improved VMD-LSTM Model for Time-Varying GNSS Time Series Prediction with Temporally Correlated Noise. Remote Sensing, 15(14), 3694. doi:10.3390/rs15143694

Chen, Y., Xia, J., Yu, C., & Chen, B. (2022). Editorial: InSAR crustal deformation monitoring, modeling and error analysis. Frontiers in Environmental Science, 10, 1009492.

Deng, L., Xiao, Y., Chen, Q., Peng, W., Li, Z., Chen, H., & Wu, Z. (2025). Detection and interpretation of the time-varying seasonal signals in China with multi-geodetic measurements. Geodesy and Geodynamics, 16(1), 42–54. https://doi.org/10.1016/j.geog.2024.05.005

Du, Y., & Haider, M. (2025). Transit ridership in Toronto and COVID-19: Statistical and spatial analysis of ridership changes during the pandemic. Transportation Research Procedia, 82, 342–355.

Duan, Y., Liang, M., Li, Y., Gao, R., Chen, J., Chen, Z. S., & Wang, H. (2026). Big data fusion-driven geospatial knowledge graph construction method for sustainable smart cities. Sustainable Cities and Society, 136, 107027. https://doi.org/10.1016/j.scs.2025.107027

Ducrocq, C., et al. (2021). Inflation-deflation episodes in the Hengill and Hrómundartindur volcanic complexes, SW Iceland. Frontiers in Earth Science, 9, 725109.

Gade, M., Nayek, P. S., & Dhanya, J. (2021). A new neural network-based prediction model for Newmark's sliding displacements. Bulletin of Engineering Geology and the Environment, 80(1), 385–397.

Gao, W., Li, Z., Chen, Q., Jiang, W., & Feng, Y. (2022). Modelling and prediction of GNSS time series using GBDT, LSTM and SVM machine learning approaches. Journal of Geodesy, 96, 71. doi:10.1007/s00190-022-01662-5

He, X., et al. (2017). Review of current GPS methodologies for producing accurate time series and their error sources. Journal of Geodynamics, 106, 12–29.

He, X., et al. (2025). Performance analysis of Geodetic Monuments on GNSS time series noise property and velocity uncertainty estimation. Advances in Space Research.

Jeuland, M. A., et al. (2022). Impact evaluation of water infrastructure investments: Methods, challenges and demonstration from a large-scale urban improvement in Jordan. Human Society.

Kim, M., Lee, S., & Choi, S.-M. (2024). DEformer: Dual embedded transformer for multivariate time series prediction. IEEE Access, 12, 153851–153858.

Kreemer, C., Govers, R., Furlong, K. P., & Holt, W. E. (1998). Plate boundary deformation between the Pacific and North America in the explorer region. Tectonophysics, 293(3–4), 225–238.

Li, S., Yang, R., & Zhan, X. (2024). Characterization of multi-band GNSS multipath in urban canyons using the ENU ray-tracing method. GPS Solutions, 28(1), 49.

Li, X., & Peng, X. (2025). Mean squared error representative points of Pareto distributions and their estimation. Entropy, 27(3), 249.

Li, X., Wang, X., & Chen, Y. (2022). InSAR atmospheric delay correction model integrated from multi-source data based on VCE. Remote Sensing, 14(17), 4329.

Lim, B., & Zohren, S. (2021). Time-series prediction with deep learning: A survey. Philosophical Transactions of the Royal Society A, 379(2194), 20200209.

Peng, M., Wu, Q., & Liu, L. (2025). Geometric design of a seismic array for local geological structure velocity tomography in the reservoir area of lower Jinsha River. Journal of Geophysics and Engineering, 22(5), 1475–1490.

Ruhl, C.J., Melgar, D., Grapenthin, R., & Allen, R.M. (2017). The value of real-time GNSS to earthquake early warning. Geophysical Research Letters, 44(16), 8311–8319.

Santamaría-Gómez, A., et al. (2011). Correlated errors in GPS position time series: Implications for velocity estimates. Journal of Geophysical Research: Solid Earth, 116(B1).

Sielfeld, G., Lange, D., & Cembrano, J. (2019). Intra-arc crustal seismicity: Seismotectonic implications for the southern Andes volcanic zone, Chile. Tectonics, 38(2), 552–578.

Swami, R., Dave, M., & Ranga, V. (2023). IQR-based approach for DDoS detection and mitigation in SDN. Defence Technology, 25, 76–87. https://doi.org/10.1016/j.dt.2022.10.006

Tabar, M. E., & Sisman, Y. (2025). Deep learning-based modeling and prediction of GNSS time series: A comparative analysis of adaptive optimization algorithms. Advances in Space Research, 76(4), 2086–2103.

Uzel, T., et al. (2013). Monitoring the tectonic plate movements in Turkey based on the national continuous GNSS network. Arabian Journal of Geosciences, 6(9), 3573–3580.

Wang, J., Dong, J., Zhou, H., & Dong, X. (2025). Population-based meta-heuristic optimization algorithm booster: An evolutionary and learning competition scheme. Neurocomputing, 643, 130405. https://doi.org/10.1016/j.neucom.2025.130405

Wang, J., et al. (2021). A new multi-scale sliding window LSTM framework (MSSW-LSTM): A case study for GNSS time-series prediction. Remote Sensing, 13(16), 3328.

Wang, Z., Chen, Y., Zhang, Y., & Ouyang, H. (2026). A novel evaluation method for the distribution of fibers within interacting fiber bundles based on moran's I. Thin-Walled Structures, 218, 114111. https://doi.org/10.1016/j.tws.2025.114111

Xiong, S., et al. (2022). Physical model tests and numerical modeling of stabilizing mechanism of portal double-row piles in landslides with interbedded weak and hard bedrock. Bulletin of Engineering Geology and the Environment, 81(3), 101.

Zhang, H., Li, Y., Zhu, T., Han, C., Wen, S., & Xiang, Y. (2025). A scalable and semi-supervised approach for ethereum fraud detection using graph neural networks and K-

nearest neighbors clustering. Journal of Information Security and Applications, 95, 104255. https://doi.org/10.1016/j.jisa.2025.104255

Zhang, S., Li, J., Zhao, L., Zeng, A., Ming, F., Liu, N., Chen, X., Feng, Z., & Wang, H. (2025). gCMEbox: A MATLAB toolbox for extracting and analyzing common-mode errors from GNSS time series. Advances in Space Research, 75(1), 497–514. https://doi.org/10.1016/j.asr.2024.09.067

Zhao, P., Shen, Z., Cao, D., Lin, Z., Chen, Z., & Hu, W. (2026). Missing data-aware robust electrical load prediction based on hierarchical downsampling-upsampling spatiotemporal graph network. Applied Energy, 405, 127209. https://doi.org/10.1016/j.apenergy.2025.127209

Zhou, Y., et al. (2025). An improved ICEEMDAN-MPA-GRU model for GNSS height time series prediction with weighted quality evaluation index. GPS Solutions, 29(3), 113.

Zhu, H., et al. (2023). Characterizing hydrological droughts within three watersheds in Yunnan, China from GNSS-inferred terrestrial water storage changes constrained by GRACE data. Geophysical Journal International, 235(2), 1581–1599.