Real-Time Attitude Prediction and Dynamic Monitoring of Super-Tall Buildings Using TQWT-TCN-LSTM-GAM Integrated with GNSS Multi-Antenna Systems

Jian Wang, Yongbo Lai, Xinyi Mao

School of Geomatics and Urban Spatial Information, Beijing University of Civil Engineering and Architecture, Beijing 102616, China - wangjian@bucea.edu.cn, laiyongbo0806@163.com, maoxinyihanzifen@163.com

Keywords: Super-Tall Buildings, GNSS Multi-Antenna, TQWT, TCN-LSTM-GAM, Attitude Prediction, Structural Health Monitoring.

Abstract

The safe operation of supertall buildings urgently requires real-time, high-precision attitude monitoring. However, under the coupled influence of complex operational loads and extreme environmental factors such as wind, earthquakes, temperature, and humidity, GNSS-based deformation monitoring systems often face challenges such as discontinuous data acquisition, fluctuating accuracy, and strong nonlinearity. These issues hinder the ability to meet the demands for real-time, high-precision, and high-frequency structural health monitoring of supertall buildings. To address these challenges, this paper proposes an attitude monitoring method that integrates Tunable Q-factor Wavelet Transform (TQWT) with a Temporal Convolutional Network–Long Short-Term Memory (TCN-LSTM) hybrid neural network model. The model adopts a serial architecture to significantly enhance the capability of processing attitude information. Additionally, a Global Attention Mechanism (GAM) is incorporated to improve the model's sensitivity, responsiveness, and representational accuracy for local anomalies and non-stationary features, thereby enabling real-time and high-precision monitoring of torsional deformations in supertall buildings. Based on this approach, a spatiotemporal prediction model for building attitude was developed. Validation experiments using a GNSS multi-antenna system demonstrated that the proposed TQWT-TCN-LSTM-GAM model achieves significantly higher prediction accuracy under complex environmental conditions compared to traditional neural network and machine learning methods.

1. Introduction

1.1 Background and Challenges of GNSS Multi-Antenna Attitude Monitoring in supertall buildings

The structural safety of supertall buildings has become a critical issue under long-term operational loads and extreme environmental conditions such as strong winds, earthquakes, temperature variations, and humidity. Accurate and real-time monitoring of structural attitude, especially torsional deformation, is essential for ensuring safety and serviceability throughout the building's lifecycle.

Multi-antenna Global Navigation Satellite System (GNSS) technology has shown promise in attitude determination by using carrier phase difference observations to calculate high-precision relative vectors between a reference antenna and multiple slave antennas (Liu and Ou, 2003). When antennas are strategically deployed at key locations of a super high-rise structure, the system can effectively monitor building attitude changes. GNSS-based attitude determination methods are typically divided into direct attitude solutions and least squares-based approaches, depending on antenna configuration (Chen et al., 2012). While both methods show comparable accuracy under ideal conditions (Zhang et al., 2016), their performance in complex environments is highly sensitive to observational errors, such as atmospheric delays and multipath interference, as well as external factors including antenna layout, baseline length, and the number of antennas (Ballal and Bleakley, 2014). To enhance short-baseline ambiguity resolution, Teunissen (1999) introduced a baseline-constrained approach for reliable ambiguity fixing. Zhang et al. (2020) designed a single-receiver GNSS system supporting multiple antennas and achieved static attitude accuracy better than 0.1° by fixing single-difference ambiguities. In dynamic shipborne tests, Wei et al. (2022) proposed a baseline-length-weighted least squares method that improved roll angle accuracy by 13%, providing new insights for GNSS-based systems in complex motion platforms. Nevertheless, applications of GNSS multi-antenna systems in supertall buildings remain rare. Unlike mobile platforms, such buildings are subject to quasi-static deformations and vibration responses under compound loads. In such scenarios, GNSS-based monitoring systems are often hindered by discontinuous data, accuracy fluctuations, and strong nonlinearity (Liu et al., 2025), making it difficult to meet the demands of high-frequency, high-precision real-time health monitoring.

To overcome these challenges, this study introduces deep learning techniques to enhance GNSS-based attitude monitoring. The aim is to achieve real-time and high-precision prediction of torsional deformation, enabling comprehensive structural monitoring over the entire service life of supertall buildings.

1.2 Related Work: Deep Learning in Structural Deformation Monitoring

Recent studies have demonstrated the effectiveness of deep learning models in modeling nonlinear and non-stationary time series data in various structural monitoring scenarios. Wu (2025) proposed a TCN-RBF hybrid architecture based on wavelet packet decomposition for bridge settlement prediction. The approach decouples the data into frequency components, where Temporal Convolutional Networks (TCN) and Radial Basis Function (RBF) networks separately model low-frequency trends and high-frequency fluctuations. The model significantly improved prediction accuracy on the Equator Bridge dataset. Lü et al. (2025) developed a three-stage architecture integrating Variational Mode Decomposition (VMD), Atom Search Optimization (AOA), and Bidirectional LSTM (BiLSTM) to enhance dam deformation forecasting, achieving optimal performance across all evaluation metrics. In landslide displacement forecasting, Wang et al. (2021) introduced a Mutual Information-IPSO-LSTM framework, combining environmental factor selection via mutual information with LSTM optimization through Improved Particle Swarm Optimization (IPSO), achieving sub-centimeter-level RMSE. For ship attitude forecasting, Ma Chao (2023) designed a VMD-TCN composite model, where each intrinsic mode function (IMF) decomposed by VMD is modeled separately by TCNs, enabling high-accuracy short-term motion prediction. Furthermore, Ma et al. (2025) developed a robust VMD-TCN-LSTM-NGO architecture combining TCN and LSTM with NGO-based parameter tuning. Validated under sea state 5 using the KCS benchmark model, the framework achieved stable 10-second predictions for heave and pitch, outperforming single models in robustness.

While hybrid models are used in dams, bridges, and ships, their application to supertall building monitoring remains unexplored. This study fills this gap by introducing a novel TQWT-TCN-LSTM-GAM hybrid model, designed specifically for real-time, high-precision prediction of torsional deformation in supertall buildings based on GNSS multi-antenna time series data.

2. Methodology

2.1 GNSS-Based Multi-Antenna Attitude Determination

2.1.1 Coordinate System Definitions: In GNSS-based multi-antenna attitude determination, the key objective is to establish the rotational relationship between the body-fixed coordinate frame defined by the rigid platform (e.g., a supertall building) and the local-level coordinate frame, often referred to as the East–North–Up (ENU) system. This orientation relationship, known as the attitude, is described via a sequence of coordinate transformations and can ultimately be expressed in the Earth-Centered, Earth-Fixed (ECEF) frame. The resulting attitude parameters include yaw, pitch, and roll angles.

2.1.2 Attitude Angle Definition: The attitude of a rigid body—such as a super high-rise building—is defined as the orientation of its body-fixed coordinate system (B-frame) relative to the local-level coordinate system (L-frame / ENU frame). This orientation relationship is fully described by three Euler angles: yaw, pitch, and roll, which together represent the complete three-dimensional (3D) attitude of the structure. The definitions of these attitude angles and their corresponding rotational axes are summarized in Table 1.

Attitude Angle	Description of Rotation	Rotation Axis in Body Frame
Yaw	Angle between the body Y-axis and geographic North	Rotation about the Z-axis
Pitch	Angle between the body Z-axis and the local Up direction	Rotation about the X-axis
Roll	Angle between the body X-axis and geographic East	Rotation about the Y-axis

Table 1. Definitions of Attitude Angles

2.1.3 Direct Method for Multi-Antenna GNSS Attitude Determination: In the direct solution algorithm of GNSS multi-antenna attitude determination, the spatial orientation of a rigid body (such as a super high-rise building) is geometrically derived from baseline vectors formed between multiple GNSS receivers mounted on the structure. Let receivers 1 and 2 define the baseline vector B_{12}, and receivers 1 and 3 define another baseline vector B_{13}. Based on these vectors, the body-fixed coordinate system (B-frame) is constructed as follows:

- The origin is located at the phase center of the master antenna (receiver 1).
- The Y-axis (pitch axis) is defined to be parallel to the baseline vector B_{12}.
- The Z-axis (yaw axis) is the normal vector perpendicular to the plane defined by the three antenna phase centers, i.e., orthogonal to the plane formed by B_{12} and B_{13}.
- The X-axis (roll axis) is perpendicular to the plane defined by the Y-axis and Z-axis, and its direction is determined according to the right-hand rule.

By solving the orientation based on the baseline vectors B_{12} and B_{13}, real-time attitude monitoring of the super high-rise structure can be achieved. The yaw angle γ, pitch angle α, and roll angle β are computed using Equations 1 to 3:

$$\gamma = \arctan\left(\frac{E_{12}}{N_{12}}\right) \tag{1}$$

where E_{12} = East component of the baseline vector from receiver 1 to receiver 2
N_{12} = North component of the baseline vector from receiver 1 to receiver 2

$$\alpha = \arctan\left(\frac{U_{12}}{\sqrt{(E_{12})^2 + (N_{12})^2}}\right) \tag{2}$$

where U_{12} = Up component of the baseline vector from receiver 1 to receiver 2

$$\beta = -\arctan\left(\frac{E_{13}\cos\gamma - N_{13}\sin\gamma}{-E_{13}\sin\alpha\sin\gamma - N_{13}\sin\alpha\cos\gamma + U_{13}\cos\alpha}\right) \tag{3}$$

where E_{13} = East component of the baseline vector from receiver 1 to receiver 3
N_{13} = North component of the baseline vector from receiver 1 to receiver 3
U_{13} = Up component of the baseline vector from receiver 1 to receiver 3

In essence, the attitude angles represent the relative angular displacements required to align the body-fixed coordinate system with the ENU reference frame. The transformation involves three sequential rotations around the body X-axis (roll), Y-axis (pitch), and Z-axis (yaw), corresponding to the standard Z-Y-X Euler rotation convention.

2.1.4 Coordinate Frame Transformations: To perform accurate attitude estimation, it is essential to transform GNSS-observed coordinates and baseline vectors across three coordinate systems: the ECEF frame, the local-level ENU frame, and the body-fixed coordinate system.

The first step is the transformation from the ECEF coordinate system to the local-level ENU coordinate system.

The baseline vector X_E obtained from GNSS differential positioning is initially expressed in the ECEF coordinate system. By applying a coordinate transformation matrix, this vector can be converted into the local ENU coordinate system, yielding the corresponding vector X_L, such that: $X_L = RX_E$. The rotation matrix R is calculated as shown in Equation 4.

$$R = \begin{bmatrix} -\sin\lambda\cos\varphi & -\sin\varphi & -\cos\lambda\cos\varphi \\ -\sin\lambda\sin\varphi & \cos\varphi & -\cos\lambda\sin\varphi \\ -\cos\lambda & 0 & -\sin\lambda \end{bmatrix} \tag{4}$$

where λ = Longitude of the observation site

φ = Latitude of the observation site

The second step is the transformation between the ENU coordinate system and the B-frame.

Both the ENU coordinate system and the body-fixed coordinate system used in this study are Cartesian coordinate systems. The transformation from the local-level ENU frame to the body frame is performed through a sequence of Euler rotations. Assuming that the two frames share the same origin, each elemental rotation can be represented by a rotation matrix. The rotations about the X-, Y-, and Z-axes are denoted by the matrices $R_1(\alpha)$, $R_2(\beta)$, and $R_3(\gamma)$, respectively, as defined in Equation 5 to 7. To transform the ENU coordinate system into the body-fixed frame, a sequential rotation about the Z-, X-, and Y-axes by angles γ, α, and β, respectively, is applied. The resulting transformation matrix is given in Equation 8. Conversely, to transform the body-fixed coordinate system back into the ENU frame, a sequential rotation about the Y-, X-, and Z-axes by angles β, α, and γ, respectively, is applied. The corresponding transformation matrix is shown in Equation 9.

$$R_1(\alpha) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\alpha & \sin\alpha \\ 0 & -\sin\alpha & \cos\alpha \end{bmatrix} \tag{5}$$

$$R_2(\beta) = \begin{bmatrix} \cos\beta & 0 & -\sin\beta \\ 0 & 1 & 0 \\ \sin\beta & 0 & \cos\beta \end{bmatrix} \tag{6}$$

$$R_3(\gamma) = \begin{bmatrix} \cos\gamma & \sin\gamma & 0 \\ -\sin\gamma & \cos\gamma & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{7}$$

$$R_L^B = R_2(\beta)R_1(\alpha)R_3(\gamma) \tag{8}$$

$$R_B^L = R_3(\gamma)R_1(\alpha)R_2(\beta) \tag{9}$$

where α= pitch angle

β = roll angle

γ = yaw angle

Therefore, if a vector is represented as X_B in the body coordinate system and as X_L in the ENU coordinate system, their transformation relationship can be expressed as follows:

$$X_L = R_B^L X_B, \quad X_B = R_L^B X_L \tag{10}$$

2.2 Tunable Q-factor Wavelet Transform（TQWT）

In the attitude monitoring of supertall buildings, GNSS-based multi-antenna signals are often subject to strong nonlinearity, non-stationarity, and significant noise interference due to the coupling effects of wind loads, seismic excitations, operational disturbances, and ambient temperature and humidity. These disturbances may lead to short-term fluctuations and obscure underlying structural deformation trends, thereby reducing the accuracy and stability of downstream prediction models. To address this issue, this study introduces the Tunable Q-Factor Wavelet Transform (TQWT) as a signal preprocessing module, which performs multi-scale decomposition and noise suppression on GNSS attitude sequences. This enhances the signal’s stationarity and structural expressiveness.

2.2.1 Multi-Scale Denoising and Modeling Concept: Unlike traditional wavelet transforms with fixed Q-factors, TQWT allows dynamic adjustment of its frequency decomposition structure to match the oscillatory characteristics of the input signal. By tuning the Q-factor (which controls oscillatory sensitivity), the redundancy factor r (which affects frequency resolution), and the decomposition level J (which determines the depth of analysis), TQWT can effectively retain low-frequency structural trends while remaining sensitive to high-frequency anomalies.

The following meteorology-adaptive TQWT parameter tuning strategy is applied in this study to address the challenges posed by variable environmental disturbances.

(1)Strong Wind Conditions: Q = 1.9 – 2.2, r = 3, J = 10, to enhance separation of low-frequency wind-induced vibrations.
(2)Rainy Weather: Q = 2.1 – 2.5, r = 4, J = 12, for suppressing broadband noise caused by raindrop impacts.
(3)Temperature Fluctuations: Q = 2.0, r = 3, J = 9, to balance thermal expansion/contraction effects with full-spectrum feature retention.

This adaptive strategy enables precise feature extraction and noise reduction across diverse weather conditions by aligning decomposition parameters with the spectral characteristics of structural and environmental responses.

2.2.2 Integration into the Hybrid Model: In the proposed TQWT-TCN-LSTM-GAM model, the TQWT module serves as the initial preprocessing stage. It outputs multi-scale attitude sub-signals with reduced noise and clearer structural trends, which facilitate subsequent TCN-LSTM modules in capturing temporal dependencies and enhance the Global Attention Mechanism (GAM) in perceiving local anomalies and sudden changes. By incorporating TQWT into the model, the overall monitoring system achieves significantly improved sensitivity and robustness to key deformation patterns in GNSS signals, enabling high-precision, real-time prediction of torsional deformation in supertall buildings.

2.3 Deep Sequence Modeling Using TCN-LSTM-GAM

2.3.1 Temporal Convolutional Network (TCN) Model: The pose time-series data preprocessed by TQWT is first input into the multi-scale Temporal Convolutional Network (TCN). The TCN module specializes in extracting high-frequency vibration features from ultra-high-rise building poses, as illustrated in Figure 1.

The TCN module adopts a multi-layer residual block structure connected in series. The core of each block consists of causal dilated convolutions (kernel length K=3, dilation factors constrained only to d=1 and d=2), which strictly maintain sequence length and causality through asymmetric zero-padding.

The deliberate constraint to small dilation factors ensures the convolutional receptive field tightly focuses on adjacent timesteps (d=1 covers 3 points, d=2 covers 5 points). This design significantly enhances the module's ability to capture local dynamics, such as transient peaks induced by wind loads and sudden phase transitions, while simultaneously avoiding redundant calculations associated with long-term dependencies. Within the residual pathway, 1×1 convolutions align channel dimensions and stabilize gradient propagation, ensuring the stackability of deep modules. A key synergistic design involves the direct concatenation of the high-frequency features output by the TCN with the low-frequency components derived from TQWT decomposition along the channel dimension. This forms a complementary feature flow integrating local vibrations and global trends, providing structured input for subsequent long-term modeling by the lightweight LSTM.

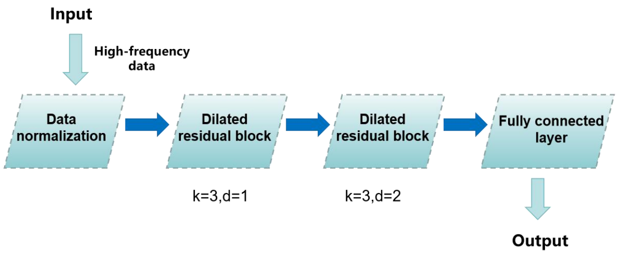

Figure 1. Architecture of the Temporal Convolutional Network (TCN) Model

2.3.2 Long Short-Term Memory (LSTM) Model: Long Short-Term Memory (LSTM) networks are a specialized type of Recurrent Neural Network (RNN) (Liu et al., 2025). Traditional RNNs suffer from gradient explosion and vanishing issues when processing long-term sequential data. To address this, LSTM introduces three gating units and a memory cell (Zhao et al., 2024) on the basis of standard RNNs. This architecture enables effective capture of long-term dependencies in ultra-high-rise building attitude variations. Consequently, LSTM is exceptionally well-suited for predicting long-term attitude sequences in super-tall structures. The internal structure of an LSTM neural unit is illustrated in Figure 2.

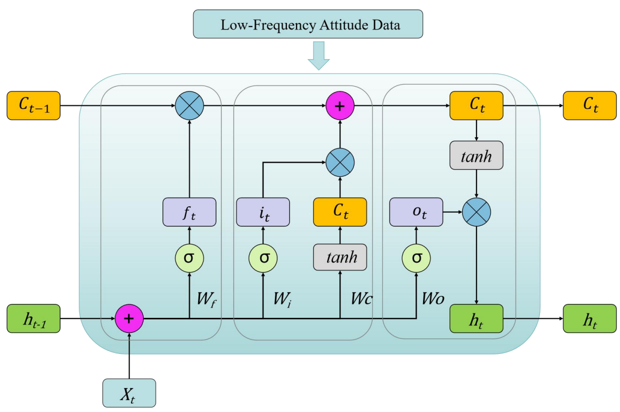

Figure 2. Internal Structure of the LSTM Neural Unit

2.3.3 Global Attention Mechanism (GAM) Model: To enhance the model's perception and representation capabilities for localized abrupt attitude changes and non-stationary vibration characteristics in ultra-high-rise structures under extreme environmental loads (e.g., strong winds, earthquakes), we introduce the Temporally Adaptive Global Attention Mechanism (GAM) at the feature fusion layer. This mechanism processes multi-scale temporal features extracted by TCN and LSTM through the following steps for dynamic enhancement of critical information:

(1) Attention Weight Generation: Applies 1D convolutional operations along the temporal dimension to adaptively learn feature importance at each timestep, generating normalized attention weights via a Sigmoid function.
(2) Dynamic Feature Enhancement: Uses element-wise multiplication to amplify features during abrupt events. (e.g., high-frequency vibration peaks, anomalous fluctuations).
(3) Residual Fusion Design: Incorporates skip connections to transmit raw attitude data or shallow features to the output layer, where they undergo weighted fusion with GAM-enhanced features. This design preserves low-frequency trend information (e.g., thermal deformation, long-term settlement) from the original sequence while strengthening transient feature extraction, preventing attenuation of long-period signals in deep networks.

The GAM mechanism enables "extreme-environment focusing" capability, substantially improving representation accuracy for non-stationary transient vibration features. Its synergy with residual architecture achieves decoupled enhancement and complementary integration of high-frequency local details and low-frequency global trends, providing critical assurance for high-precision attitude prediction in complex environments.

2.3.4 Model Fusion Structure and Training Strategy: The TQWT-TCN-LSTM-GAM model, which integrates TQWT, TCN-LSTM, and GAM modules, constitutes the proposed framework for attitude prediction in ultra-high-rise buildings (see Figure 3). In the TQWT-TCN-LSTM-GAM hybrid neural network model, the Tunable Q-factor Wavelet Transform (TQWT) reduces the nonlinearity of deformation attitudes in ultra-high-rise structures. The preprocessed attitude time-series data is first fed into a multi-scale Temporal Convolutional Network (TCN) module. This module employs a multi-layer cascaded residual block structure with varying dilation rates (d=1, d=2) to achieve cross-temporal-scale feature extraction.

The TCN specializes in capturing local dynamic features within sequential data, making it particularly suitable for identifying high-frequency vibration modes induced by factors such as wind loads in super-tall buildings.

To address the TCN's limitations in modeling long-term dependencies, its output features are concatenated with low-frequency data along the channel dimension and passed to a lightweight Long Short-Term Memory (LSTM) module. The LSTM further models low-frequency features and global trend variations, thereby compensating for the TCN's deficiencies in long-term dependency modeling.

At the terminal stage of the model, a Temporal Adaptive Global Attention Mechanism (GAM) is introduced. This module processes the temporal features output by the LSTM and applies a 1D convolutional operation along the time dimension to generate an attention weight vector with the same length as the original sequence. These weights are normalized via a Sigmoid function to dynamically quantify the importance of features at each timestep. The GAM adaptively focuses on critical vibration characteristics during abrupt events (e.g., high-frequency peaks under extreme wind conditions). By assigning higher weights to these key timesteps, it significantly enhances the model's responsiveness and representational accuracy for local anomalies and non-stationary features. Simultaneously, the model incorporates skip connections that directly transmit preprocessed raw attitude sequences or shallow features to the fusion layer. This effectively preserves low-frequency trend information from the original time series, preventing the loss of critical long-cycle information during deep feature extraction in complex networks. Consequently, the model enhances its multi-scale representational capacity for differentiating between vibration characteristics (high-frequency/local) and trend variations (low-frequency/global).

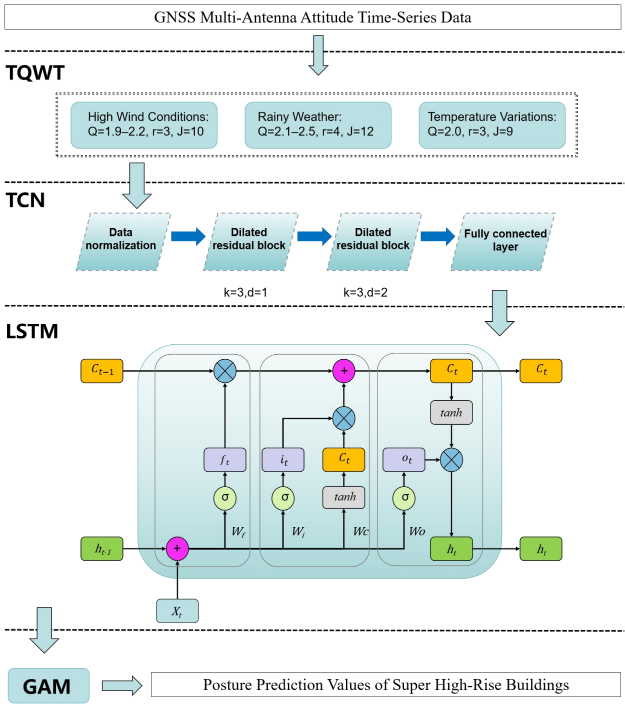

Figure 3. Forecasting framework of the TQWT-TCN-LSTM-GAM hybrid architecture

Finally, the feature maps dynamically enhanced by GAM attention weights are fused at the element-wise level with the original/shallow features delivered via skip connections. The integrated composite features are consolidated through fully connected layers and regressed to output high-precision predictions of ultra-high-rise building attitude changes. This enables real-time, high-accuracy attitude monitoring for torsional deformation surveillance in super-tall structures.

2.4 Accuracy Evaluation Metrics

To evaluate the performance of the prediction model, error analysis was conducted on the test set using the Mean Absolute Error (MAE), Root Mean Square Error (RMSE), Coefficient of Determination (R^2), and Symmetric Mean Absolute Percentage Error (SMAPE). The expression is shown in Equation 11.

$$\begin{cases} RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i-\hat{y}_i)^2} \\ R^2 = 1-\frac{\sum_{i=1}^{n}(y_i-\hat{y}_i)^2}{\sum_{i=1}^{n}(y_i-\bar{y}_i)^2} \\ MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i-\hat{y}_i| \\ SMAPE = \frac{1}{n}\sum_{i=1}^{n}\frac{|\hat{y}_i-y_i|}{(|\hat{y}_i|+|y_i|)/2}\times 100\% \end{cases} \tag{11}$$

where y_i = true value
$\hat{y}_i$ = predicted value
$\bar{y}_i$ = mean of true values
n = sample size

RMSE reflects the magnitude of overall prediction errors, with lower values indicating higher stability. R^2 measures the model's ability to explain the variance in observed data, where values closer to 1 denote better fitting. MAE quantifies the average absolute deviation between predicted and true values, serving as a direct indicator of prediction accuracy. SMAPE, an improved version of MAPE, uses the sum of predicted and actual values in the denominator, providing a robust measure of relative error. These metrics jointly assess the predictive accuracy and reliability of different models under the same dataset.

3. Experiments and Results

3.1 Experimental Setup

The device used in this study is a Windows 11 system with the following basic configuration: Intel(R) Core(TM) i5-8300H CPU @ 2.30GHz, GPU: NVIDIA GeForce GTX1650 4GB, and RAM: 8G. The Python version used is Python 3.9.12.

The key hyperparameters of the deep learning network used in this experiment are summarized in Table 2.

Parameter	Setting
Dropout rate	0.1
Epoch	100
Batch size（train）	256
Batch size（test）	256
Learning rate	0.0005
Optimizer	Adam
Loss function	MSE
TCN layers	2

Table 2. Hyperparameter settings of the TQWT-TCN-LSTM-GAM model used in this study.

3.2 Data Acquisition and Preprocessing

3.2.1 Data Acquisition: The experimental data used in this study were collected from the GNSS-based structural health monitoring system deployed on a supertall building located in Tianjin, China. The configuration of the GNSS multi-antenna layout is illustrated in Figure 4.

https://doi.org/10.5194/isprs-archives-XLVIII-4-W14-2025-297-2025 | © Author(s) 2025. CC BY 4.0 License.

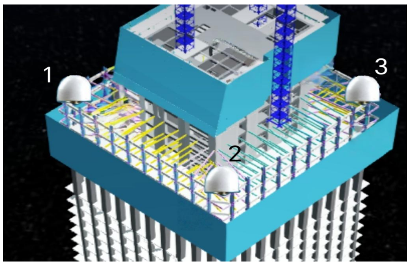

Figure 4. GNSS Multi-Antenna Deployment Scheme on the Supertall Building

3.2.2 TQWT Denoising Results Analysis: Based on the baseline vectors B_{12} and B_{13} constructed from the receivers, the pitch and heading attitude angles were derived through solution calculations. The heading and pitch angles obtained from the baseline vector B_{12} solution are shown in Figure 5, while those derived from the baseline vector B_{13} solution are presented in Figure 6.

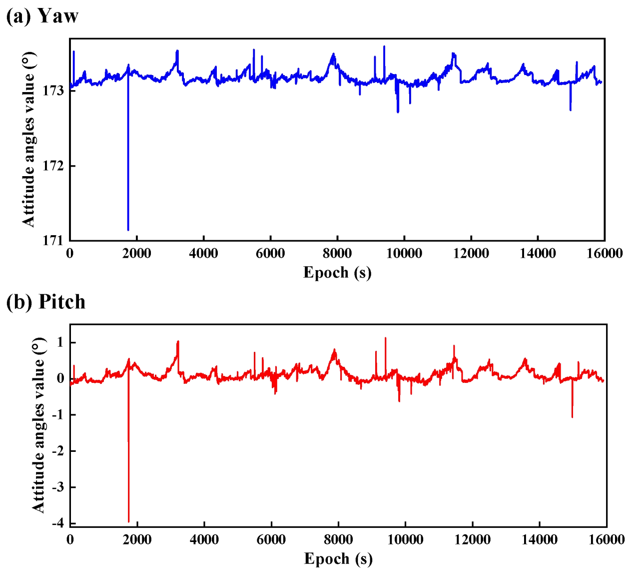

Figure 5. Attitude angles derived from baseline vector B_{12} in the local horizontal coordinate system. (a) Yaw; (b) Pitch.

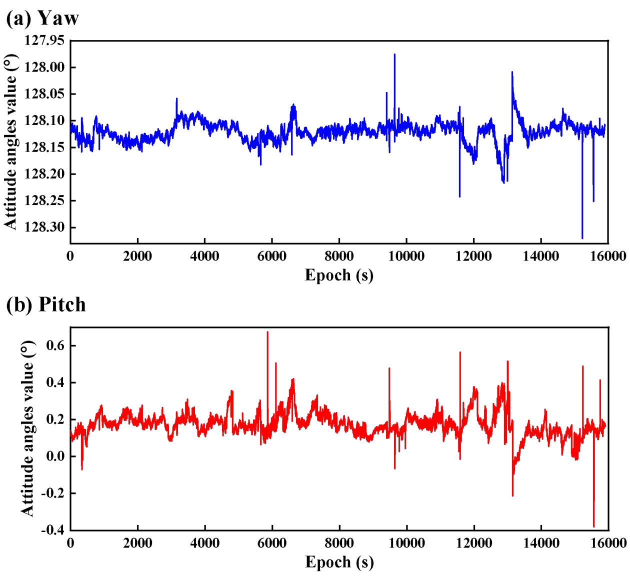

Figure 6. Attitude angles derived from baseline vector B_{13} in the local horizontal coordinate system. (a) Yaw; (b) Pitch.

Since the dual-antenna GNSS direct attitude determination method can only resolve two-dimensional attitudes (pitch and heading angles), this study integrates the solutions from both baselines to reconstruct a complete three-dimensional attitude time series for the super-tall building. To address the pronounced nonlinear characteristics of the attitude data, the Tunable Q-factor Wavelet Transform (TQWT) was introduced during the preprocessing stage, establishing a denoising model that integrates trend separation and adaptive filtering. Figure 7 compares TQWT-processed and original signals.

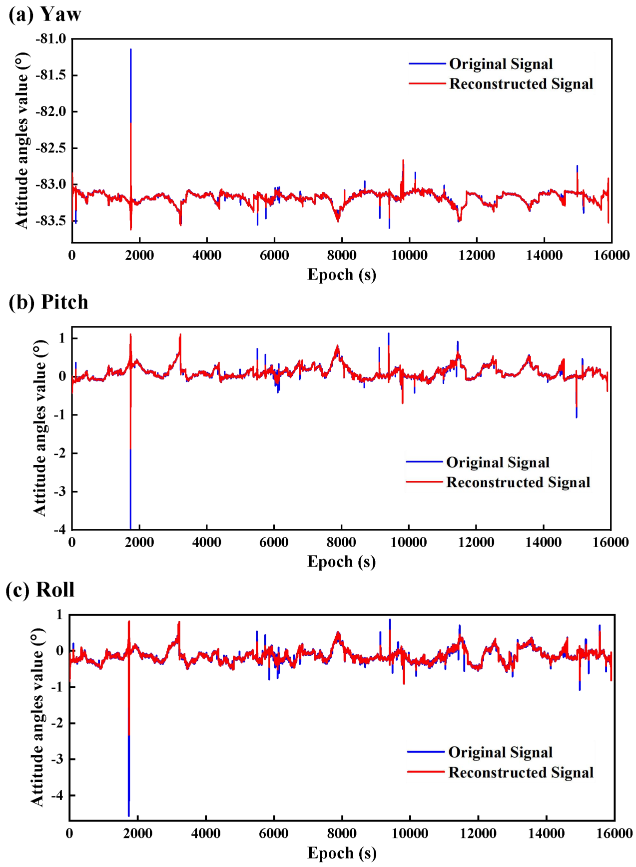

Figure 7. Comparison Between Original Signal and Reconstructed Signal. (a) Yaw; (b) Pitch; (c)Roll.

3.3 TQWT-TCN-LSTM-GAM Model Accuracy Evaluation Results

In the proposed hybrid model TQWT-TCN-LSTM-GAM, the GNSS attitude time-series data are first preprocessed using TQWT to suppress noise and enhance signal stationarity. The processed data are then input into the TCN-LSTM-GAM network for accurate prediction of high-rise building attitude angles.

To comprehensively evaluate the model's performance, four widely used metrics were adopted: Root Mean Square Error (RMSE), Mean Absolute Error (MAE), Symmetric Mean Absolute Percentage Error (SMAPE), and the coefficient of determination (R^2). These indicators provide both absolute and relative assessments of prediction accuracy.

As shown in Table 3, the TQWT-TCN-LSTM-GAM model demonstrates excellent predictive performance across all three attitude components—Yaw, Pitch, and Roll. The low MAE and RMSE values indicate minimal prediction error, while the high R^2 values (>92%) confirm strong consistency between predicted and actual attitude angles. In particular, the model achieved the highest accuracy in Yaw angle prediction, with an MAE of only 0.0069 degrees and an R^2 of 94.21%, showcasing its effectiveness in capturing subtle attitude variations.

These results collectively validate the robustness and accuracy of the proposed hybrid model in attitude monitoring applications under complex conditions commonly encountered in supertall building environments.

	MAE (Degree)	RMSE (Degree)	SMAPE (%)	R^2 (%)
Yaw	0.0069	0.0197	0.59%	94.21%
Pitch	0.0173	0.0483	1.12%	92.32%
Roll	0.0212	0.0518	1.19%	92.16%

Table 3. Analysis of Attitude Angle Accuracy Indicators Based on the TQWT-TCN-LSTM-GAM Model

3.4 Comparative Experimental Results

To further validate the superiority of the proposed TQWT-TCN-LSTM-GAM model, a set of comparative experiments was conducted using three representative attitude time series: Yaw, Pitch, and Roll. Four models were evaluated in this study: Temporal Convolutional Network (TCN), Long Short-Term Memory network (LSTM), eXtreme Gradient Boosting (XGBoost), and the proposed hybrid model.

The prediction results of TCN, LSTM, XGBoost, and the proposed TQWT-TCN-LSTM-GAM model are compared in Figure 8, where all models are shown to capture the general trend of attitude variations to varying degrees. However, a closer inspection of the blue-shaded region in Figure 8 reveals significant differences in prediction accuracy. Regardless of whether the time series undergoes gentle fluctuations or abrupt changes, the TQWT-TCN-LSTM-GAM model consistently produces predictions that closely match the ground truth.

Compared with the baseline models, the hybrid model not only achieves higher accuracy in smooth segments but also demonstrates strong robustness during periods of rapid attitude fluctuation. This is primarily attributed to three key mechanisms:

(1)TQWT effectively eliminates high-frequency noise while preserving essential signal characteristics, improving the signal quality fed into the model;
(2)TCN is capable of effectively capturing local temporal patterns and short-term fluctuations through stacked causal and dilated convolutional layers, making it well-suited for modeling high-resolution variations in attitude signals;
(3)LSTM excels at learning long-term dependencies in time series, ensuring that contextual information is retained over extended time horizons and contributing to the stability of attitude prediction;
(4)GAM (Global Attention Mechanism) dynamically assigns importance weights to different features and time steps, allowing the model to focus on globally relevant patterns and improving its interpretability and generalization.

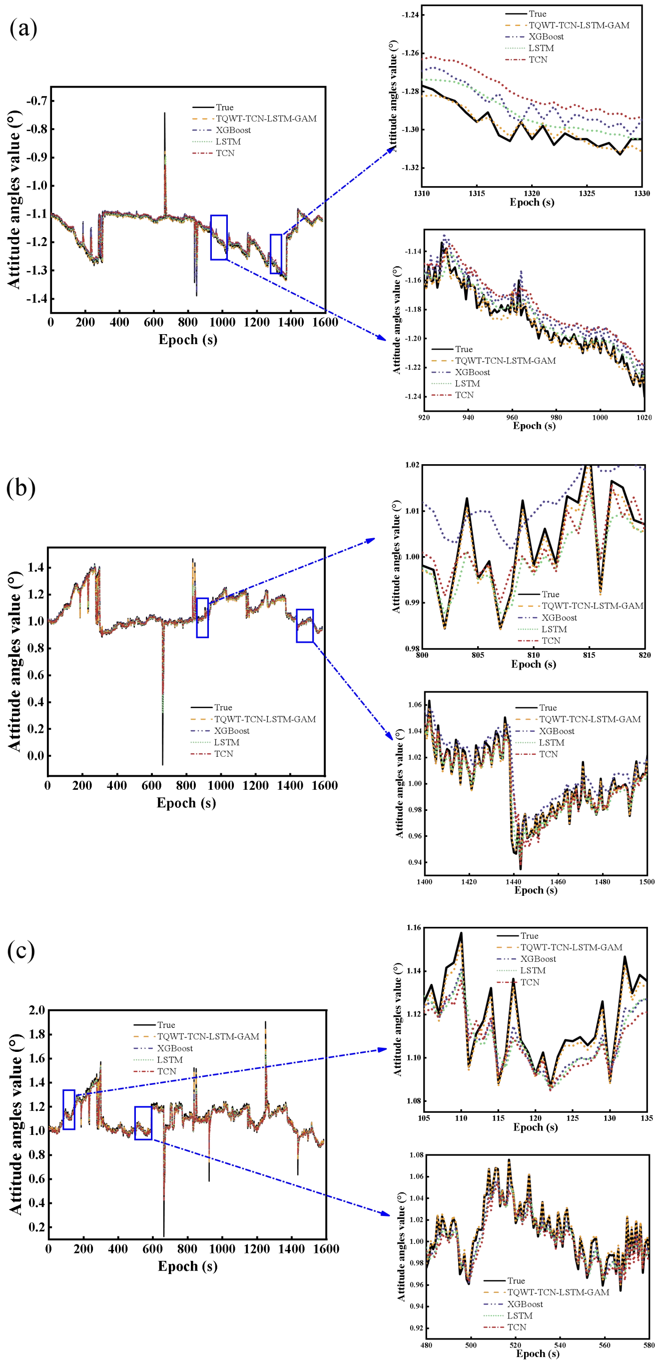

Figure 8. Comparison of Attitude Angle Prediction Results Using Different Models (TCN, LSTM, XGBoost, and TQWT-TCN-LSTM-GAM) (a) Yaw; (b) Pitch; (c)Roll.

Overall, the experiment results confirm that the proposed TQWT-TCN-LSTM-GAM model significantly outperforms traditional deep learning and machine learning models in both prediction accuracy and generalization performance. It is particularly well-suited for attitude monitoring in complex environments typical of supertall buildings.

4. Conclusions

In response to the demand for real-time, high-precision, and high-frequency health monitoring of torsional deformation in supertall buildings under complex environmental conditions, this study proposes a short-term attitude prediction method that integrates GNSS multi-antenna measurements with a deep learning framework. The core of the method lies in a hybrid neural network model based on TQWT and a TCN-LSTM architecture, further enhanced by the introduction of a Global Attention Mechanism (GAM). This design significantly improves the model's ability to capture local anomalies and characterize non-stationary dynamic features, enabling accurate spatiotemporal prediction of building attitude.

Experiments were conducted using attitude angle time series—specifically yaw, pitch, and roll—as prediction targets. Model performance was quantitatively evaluated using four metrics: RMSE, MAE, SMAPE, and R^2. Results demonstrate that the proposed TQWT-TCN-LSTM-GAM model consistently outperforms baseline models including LSTM, TCN, and XGBoost across all evaluation metrics.

The model offers practical potential for structural health monitoring. It shows promising application prospects in real-time deformation monitoring of supertall buildings, particularly under dynamic and uncertain environmental conditions.

References

Ballal, T., Bleakley, C.J., 2014. GNSS instantaneous ambiguity resolution and attitude determination exploiting the receiver antenna configuration. *IEEE Trans. Aerosp. Electron. Syst.*, 50(3), 2061–2069.

Chen, Y., Tang, Y., Zhou, W., et al., 2012. Error analysis of triple-antenna GPS attitude determination. *Global Positioning System*, 37(3), 16–18. (in Chinese)

Chen, T.Q., Guestrin, C., 2016. XGBoost: A scalable tree boosting system. In: *Proc. 22nd ACM SIGKDD Int. Conf. Knowledge Discovery and Data Mining*, 785–794.

Liu, G.Y., Ou, J.K., 2003. GPS single-epoch orientation and attitude determination algorithm and accuracy analysis. *Geomatics and Information Science of Wuhan University*, 28(6), 732–735. (in Chinese)

Liu, F., Xin, C., Guo, Y., et al., 2025. Water quality prediction based on LSTM and XGBoost fusion model. *Hydropower*, 1–11. (in Chinese)

Lu, G.X., Chen, B., Zhang, M.X., et al., 2025. Dam deformation prediction model based on VMD-AOA-BiLSTM. *Hydropower and Energy Science*, (7), 148–152. https://doi.org/10.20040/j.cnki.1000-7709.2025.20241707 (in Chinese)

Ma, C., 2023. Research on deep learning-based ship attitude prediction technology. Southeast University. https://doi.org/10.27014/d.cnki.gdnau.2023.001342 (in Chinese)

Ma, Y.Y., Jiang, Z.Q., Wang, X.M., et al., 2025. A ship motion prediction method based on TCN-LSTM optimized by NGO parameters. *Shipbuilding of China*, 66(1), 187–197. (in Chinese)

Teunissen, P., 1999. The Least-Squares Ambiguity Decorrelation Adjustment: A method for fast GPS ambiguity estimation. *Journal of Geodesy*, 73, 587–593.

Wang, L., Xu, H., Shu, B., et al., 2021. Multi-source data fusion for landslide monitoring using mutual information and IPSO-LSTM. *Geomatics and Information Science of Wuhan University*, 46(10), 1478–1488. https://doi.org/10.13203/j.whugis20210131 (in Chinese)

Wei, Y.S., He, K.F., Liu, D.X., et al., 2022. GNSS multi-antenna attitude determination method based on baseline length weighting. *Telemetry and Telecontrol*, 43(3), 18–23. (in Chinese)

Wu, C.C., 2025. Application of TCN-RBF neural network based on wavelet packet decomposition in bridge settlement prediction. *Beijing Surveying and Mapping*, 39(1), 105–110. https://doi.org/10.19580/j.cnki.1007-3000.2025.01.017 (in Chinese)

Zhang, F.Z., Chai, Y.J., Chai, H., et al., 2016. Accuracy analysis of two GNSS multi-antenna attitude determination methods. *Journal of Inertial Technology*, 24(1), 30–35. (in Chinese)

Zhang, P., Zhao, Y., Lin, H., et al., 2020. A novel GNSS attitude determination method based on primary baseline switching for a multi-antenna platform. *Remote Sensing*, 12(5), 747.

Zhao, Y.P., Huang, W., Zhang, J.F., 2024. Lithium battery remaining useful life prediction based on hybrid scale health factors using LSTM-Transformer model. *Electronic Measurement Technology*, 47(11), 112–122.