Bao *et al.*
Intelligent Marine Technology and Systems (2026) 4:16
https://doi.org/10.1007/s44295-026-00105-7

imts Intelligent Marine Technology and Systems

RESEARCH PAPER Open Access

Geomagnetic field modeling with dual-architecture neural networks

Zhijin Bao[1], Junhui Xing[1,2*], Haowei Xu[1], Jiayi Wei[1] and Chong Xu[3]

Abstract

Traditional geomagnetic field modeling requires balancing computational efficiency with the resolution of complex dynamics. A dual-architecture deep learning framework is introduced as a high-efficiency surrogate for global modeling. A sinusoidal representation network (SIREN) captures the continuous, quasi-static internal field, and a transformer with an 'adaptive recursion' mechanism models the dynamic external field. The recursive mechanism adjusts computational depth according to geomagnetic activity, optimizing efficiency and accuracy. Trained on CHAOS-8.4 data, the framework reproduces the global field with high fidelity. Under identical CPU hardware conditions, the SIREN-based surrogate model achieved an inference speed-up of approximately 1100 times compared to traditional implementations. Native graphics processing unit (GPU) acceleration further enables large scale parallel computations without additional engineering effort. This study validated the feasibility of deep learning as a robust real-time surrogate for complex geophysical simulations.

Keywords Geomagnetic field modeling, Deep learning, Sinusoidal representation network, Recursive transformer, Surrogate modeling

1 Introduction

The geomagnetic field is a fundamental property of the Earth that shields the biosphere from solar wind and cosmic radiation and serves as a critical reference for navigation and geophysical exploration (Merrill et al. 1998). Accurate modeling is therefore crucial for investigating Earth's deep interior dynamics and monitoring near-Earth space weather. The geomagnetic field represents a superposition of contributions of sources: the dominant, slowly varying external component generated by the geodynamo in the fluid outer core, and highly dynamic external components arising from solar wind and the magnetosphere-ionosphere system (Hulot et al. 2010).

Traditionally, geomagnetic field modeling has relied on spherical harmonic analysis (Upendran et al. 2022) for the internal field, as exemplified by the International Geomagnetic Reference Field (IGRF) (Kloss and Beggan 2025) and the high-resolution CHAOS model series (Kloss et al. 2026). External perturbations are typically addressed using computationally intensive magnetohydrodynamic (MHD) simulations (Tóth et al. 2005) or parameterized empirical indices. Physics-based models provide interpretability but incur high computational costs and often fail to resolve the high-frequency nonlinear dynamics during geomagnetic storms. Empirical models, while efficient, may overlook complex local variations.

Recent advances in the 'AI for Science' paradigm have revolutionized geophysics, introducing data-driven approaches that balance accuracy and efficiency (Bergen et al. 2019; Camporeale 2019). Deep learning techniques now permeate the data processing pipeline. In

*Correspondence:
Junhui Xing
junhuixing@ouc.edu.cn
[1] Key Laboratory of Submarine Geosciences and Prospecting Techniques, MOE and College of Marine Geosciences, Ocean University of China, Qingdao 266100, China
[2] Key Laboratory of Polar Geology and Marine Mineral Resources (China University of Geosciences Beijing), Ministry of Education, Beijing 100083, China
[3] Department of Earth Sciences, University of Oxford, Oxford OX1 2JD, UK

Springer

preprocessing, machine learning algorithms outperform traditional interpolation for reconstructing missing observations (Liu et al. 2019) and denoising multisource magnetic signals (Li et al. 2023; Yang 2024), thereby ensuring high-quality inputs for subsequent tasks.

For internal field studies, emphasis has shifted from pure simulations to inversion and separation. Novel frameworks such as current loop-based U-Net model segmentation inference (CLUMSI) (Kuslits et al. 2024) utilize U-Net architectures to invert equivalent current loops in the core, whereas spherical deep learning methods (Otzen 2022) achieved separation of core and lithospheric fields by learning nonlinear spatial features beyond the capability of standard spherical harmonics. In the secular variation domain, long short-term memory (LSTM) networks have demonstrated significant improvements over linear extrapolation methods for regional trend prediction (Li and Chen 2025).

Applications of deep learning are even more extensive for modeling highly dynamic external fields and space weather events. For global index prediction (e.g., the disturbance storm time (Dst) index and the symmetric H-component (SYM-H) index), recurrent neural networks (RNNs) such as LSTM and GRU have become standard tools, outperforming traditional baselines (Wintoft and Wik 2021). Recent advances focus on addressing the limitations of naive deep learning models, such as the 'time lag' problem (Laperre et al. 2018), through hybrid architectures like empirical mode decomposition-LSTM (EMD-LSTM) (Zhang et al. 2023) and specialized 'storm-time' models (Ren et al. 2023) designed to capture extreme events. Furthermore, the field is moving towards rigorous uncertainty quantification (Tasistro-Hart et al. 2021; Conde et al. 2023) and early warning systems based on solar source imaging (Pelkum Donahue and Inceoglu 2024; Sakpal 2024). On a global scale, models such as Deep leArninG Geomagnetic pErtuRbation (DAGGER) (Upendran et al. 2022) and electron precipitation flux deep neural networks (EPFN) (Chen et al. 2023) have demonstrated the feasibility of rapid grid-free global forecasting of geomagnetic perturbations and particle precipitation.

Despite these advancements, a unified, efficient, and physics-aware framework for global geomagnetic field modeling remains a significant challenge. Existing deep learning approaches primarily emphasize localized time-series prediction or source-specific isolation, with limited attention to full-field reconstruction. Moreover, standard neural networks, such as multilayer perceptrons (MLPs) and convolutional neural networks (CNNs) often struggle to represent the continuous, derivative-smooth nature of potential fields (internal fields) while simultaneously capturing the multi-scale temporal dynamics of external disturbances.

To address these gaps, a dual-architecture framework is introduced to capture the distinct spatiotemporal characteristics of geomagnetic sources observed on Earth's surface. The contributions of this work lie in aligning neural representations with the specific nature of the signal. For the quasi-static and spatially continuous internal field, a sinusoidal representation network (SIREN) is used. Unlike grid-based methods limited by discrete resolutions or standard networks using piecewise linear activations (e.g., rectified linear unit, ReLU), SIREN utilizes periodic activations to achieve a mesh-free, continuous field representation. This enables high-fidelity reconstruction of complex global spatial structures without reliance on spherical harmonic truncation. For the highly dynamic external field, a transformer with an 'adaptive recursion' mechanism is designed. This mechanism dynamically adjusts the computational depth based on the intensity of geomagnetic activity. This resolves the inefficiency of fixed-depth models during quiet periods while ensuring sufficient feature extraction during storm events, effectively balancing computational efficiency with prediction accuracy for surface disturbances. The integration of these specialized architectures establishes an efficient surrogate model, validating the feasibility of deep learning as a real-time alternative to traditional physical simulations for surface-level full-field reconstruction.

2 Methods

2.1 Data

ChaosMagPy (Kloss 2024) is a Python package used to compute the CHAOS geomagnetic field model. Magnetic field datasets utilized in this study were generated with ChaosMagPy v0.15 + CHAOS-8.4, in the geographic coordinate system. Two distinct datasets were constructed to train the internal and external field models independently.

2.1.1 Internal field data

The internal field model training set comprised two components. The first was a benchmark grid dataset generated annually on July 1 with a spatial resolution of 1°, enabling the model to capture the macroscopic static structure of the global magnetic field. The second component consisted of random spatiotemporal samples with 45000 points drawn every 30 days. Of these, 70% were uniformly distributed across the Earth's surface, whereas the remaining 30% were concentrated near the geomagnetic poles and the South Atlantic anomaly (SAA) to enhance the model performance in critical regions.

The resulting training set exhibited a multiscale distribution aligned with the residual learning framework. Annual grid data provided a spatially complete but temporally sparse structural backbone representing the static long-wavelength components of the internal field. At this baseline, high-frequency random samples overlay crucial temporal dynamics. By tasking the model with learning only the residuals—the difference between the true field and the static baseline—its learning strategically focuses on capturing the complex, time-varying behavior that is most prominent in scientifically critical regions.

2.1.2 External field data

The ionospheric and magnetospheric field data were treated differently. The external magnetospheric field is static in the geocentric solar magnetospheric (GSM) coordinate system with variation in the geographic coordinate system determined solely by geomagnetic dipole axis orientation. In contrast, the internal magnetospheric field requires the ring current index (RC-index) (Conde et al. 2023) for its calculation, in addition to the dipole axis orientation. The ionospheric field is considerably more complex, as its parameters are influenced not only by solar wind and the interplanetary magnetic field (IMF), but also by the magnetic local time (MLT).

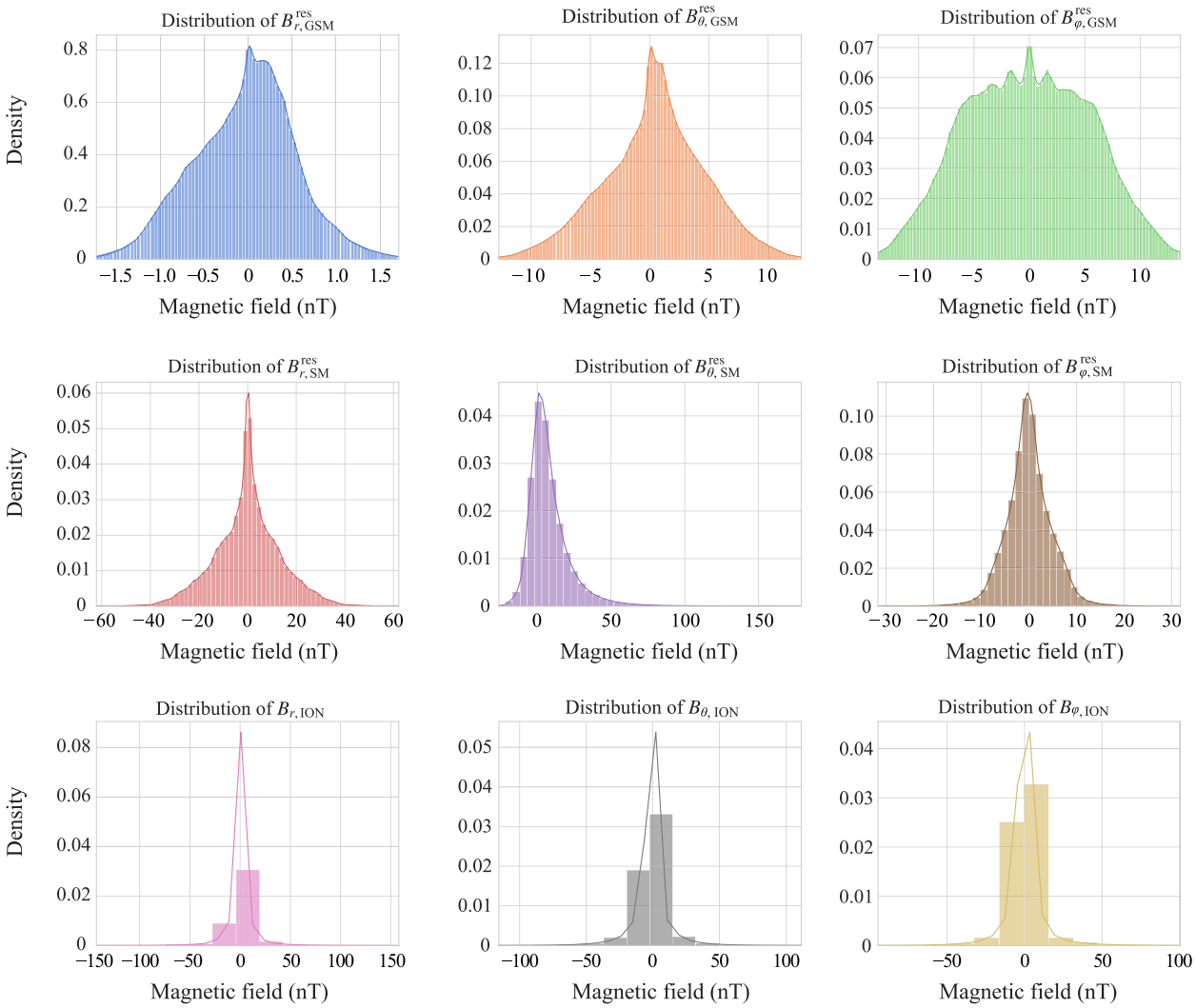

Fig. 1 Distribution of external magnetic field components. The figure is arranged in a 3×3 grid, with columns representing the radial (B_r), polar (B_θ), and azimuthal (B_φ) components, respectively. Rows correspond to the GSM, SM, and ionospheric (ION) coordinate systems. Each subplot displays a normalized histogram (density) and a kernel density estimation (KDE) curve for the respective target variable. Note that the GSM and SM components are presented as residuals (B^{res}), representing the deviation from a baseline model, while the ION components are direct target values

Figure 1 illustrates the statistical properties of the nine target variables for neural network training. Residual components in GSM and solar magnetic (SM) coordinates (top and middle rows) are strongly centralized around zero and exhibit quasi-symmetric, heavy-tailed characteristics resembling a Laplacian or Student's t-distribution more than a pure Gaussian. This centralization indicates that for most timestamps, the baseline model provides a reasonable approximation, and the network's primary task is to learn these relatively small, symmetric deviations.

Parameters were selected for their direct influence on near-Earth space conditions. The $F_{10.7}$ index (the solar radiation flux with a wavelength of 10.7 cm) quantifies solar extreme ultraviolet (EUV) radiation, controlling the conductivity and density of the ionosphere, thus serving as a key driver of the ionospheric magnetic field. The IMF, particularly its north–south component (B_z), governs the energy transfer from solar wind to the magnetosphere through magnetic reconnection, making it the primary trigger for geomagnetic storms and a crucial input for modeling the magnetospheric (GSM) field. The RC-index provides a global measure of ring current intensity, which is essential to accurately model the symmetric disturbance (SM) field.

For the external field model, pre-training was performed on a dataset with fixed solar activity parameters ($F_{10.7}$: 100 sfu, IMF_y: 0 nT, IMF_z: 1 nT, V_x (solar wind velocity x-component): 400 km/s, RC-index: −10 nT). Random sampling was avoided to prevent overrepresentation of geomagnetically quiet periods. Instead, a balanced, stratified sampling strategy was implemented, categorizing historical data into three activity levels based on the planetary K-index (Kp) index: quiet ($\mathrm{Kp} \leq 20$), active ($20 < \mathrm{Kp} \leq 40$), and storm ($\mathrm{Kp} > 40$), with spatial sampling biased towards uniform, auroral, and SAA regions differently for each level. The data were then allocated for model fine-tuning in a 15:8:5 ratio, oversampling rare storm events. Notably, we assigned a 10% sampling weight to the SAA for both the quiet and active periods. This is primarily because the weakening of the magnetic field in the SAA stems from the long-term evolution of the core field, making its short-term response to solar wind disturbances relatively independent. Therefore, it is necessary to maintain a consistent focus on this region across activity levels.

2.1.3 RC-index, IMF, and $F_{10.7}$ data

RC-index data were provided by the National Space Institute, Technical University of Denmark (DTU Space) (Finlay et al. 2015). This index monitors variations, derived from a global network of geomagnetic observatories and combined with precise magnetic field separation from the CHAOS model. Data for the IMF and solar wind parameters were sourced from the 1-min resolution OMNI (King and Papitashvili 2005) dataset, accessed via the OMNI web interface of NASA's Space Physics Data Facility (SPDF). The $F_{10.7}$ solar radio flux index data were obtained from the Dominion Radio Astrophysical Observatory in Penticton, Canada (Tapping 2013) and retrieved through the LASP Interactive Solar Irradiance Data Center (LISIRD) interface at the Laboratory for Atmospheric and Space Physics (LASP), University of Colorado, Boulder.

As shown in Fig. S1 (see the Supplementary Material), on January 10th, the B_z component turned sharply southward, coinciding with a strong negative excursion of the RC-index indicating the onset of the main phase. Throughout this period, the $F_{10.7}$ index remained stable at approximately 74 solar flux units (sfu), indicating relatively low solar activity.

The $F_{10.7}$ index exhibits a right-skewed distribution with a mean of 112.1 sfu. Both IMF B_y and B_z components are symmetrically centered near zero, with B_z showing a more concentrated peak (Fig. S2). Additionally, the RC-index presents a heavily left-skewed distribution with a long negative tail (mean = −13.1 nT), which characterizes the episodic occurrence of geomagnetic storms.

2.2 Model architecture

2.2.1 SIREN

The core of SIREN (Sitzmann et al. 2020) comprises a series of specialized MLP layers. This paradigm, which represents signals as continuous functions, has demonstrated exceptional fidelity in reconstruction tasks for

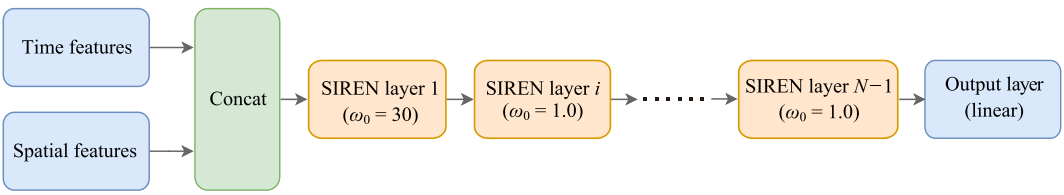

Fig. 2 Architecture of the SIREN model

signals, such as images and 3D shapes (Fig. 2). Consequently, it is naturally suited for modeling spatially continuous and smooth physical fields such as the Earth's internal magnetic field.

Specifically, the network interprets the model input as coordinates in a continuous domain and learns the mapping from these coordinates to the target physical quantities. Unlike conventional MLPs that use activation functions like ReLU, each layer Φ_i in a SIREN employs the sine function as its activation:

$$x_{i+1} = \sin(\omega_0 * (W_i * x_i + b_i)), \quad (1)$$

where W_i and b_i are the weights and bias of the i layer, and the critical parameter ω_0 controls the frequency of the sinusoidal activation. Sitzmann et al. (2020) noted that proper weight initialization is crucial for ensuring stability during training and effectively representing signal details. For the first layer, weights ω_0 are sampled from a uniform distribution $U(-\frac{1}{d_{\text{in}}}, \frac{1}{d_{\text{in}}})$; for subsequent layers, weights W_i are sampled from $U(-\frac{\sqrt{\frac{6}{d_{\text{in}}}}}{\omega_0}, \frac{\sqrt{\frac{6}{d_{\text{in}}}}}{\omega_0})$, where d_{in} is the input dimensionality of the layer. This initialization scheme ensures that the distribution of activations remains consistent as they propagate through the network.

The model implementation employs concatenated spatiotemporal coordinates as input. These coordinates are first mapped into a feature space through an initial SIREN layer ($\omega_0 = 30.0$). Subsequent standard SIREN layers ($\omega_0 = 1.0$) progressively build a complex representation of the target field. A final layer without activation maps hidden features to three components of the geomagnetic field $(B_r, B_\theta, B_\varphi)$.

Spatial extrapolation capability of the model was evaluated using a specialized generalization test to evaluate. In this test, a contiguous geographical region (0° to 30°E) was designated as a holdout validation set, ensuring that the model never encountered data from this area during training. This rigorous test was intended to evaluate the performance of the data-driven model in extrapolation tasks.

After validating the potential of our architecture, a final practical model was developed. For maximum accuracy across all regions, a conventional random sampling approach was adopted, reserving 15% of global data, including the 0° to 30°E region for validation.

2.2.2 Transformer

This model is based on the transformer framework proposed by Vaswani et al. (2017), a framework whose effectiveness, along with its variants such as bidirectional encoder representations from transformers (Devlin et al. 2019), has been demonstrated in handling complex sequential tasks, but incorporates several modifications. Its core innovation is the introduction of a dynamic computation depth mechanism inspired by a universal transformer and adaptive computation time (Graves 2016). Termed 'adaptive recursion', this mechanism adjusts processing steps according to signal complexity. As illustrated in Fig. 3, the architecture comprises three stages: feature embedding, recursive transformer computation, and multitask prediction heads.

The dynamic depth mechanism reflects a physical hypothesis during geomagnetically quiet periods, external field structures remain relatively simple, requiring shallow computations. Conversely, during magnetic storms, systems, such as ring- and field-aligned currents, become exceptionally complex, necessitating deep feature extraction and nonlinear mapping. Recursive routing operationalizes this intuition, allocating fewer resources to simple samples while invoking deeper layers from complex storm events. This approach improves computational efficiency and interpretability without reducing performance. Construction of the external field model involved a simplifying assumption modulation effect on external field coordinate systems arising from long-term core field evolution were neglected. Although the CHAOS model accounts for such modulations (e.g., GSM/SM to geographic coordinate system and the definition of quasi-dipole coordinates) these subtle long timescales variations were excluded.

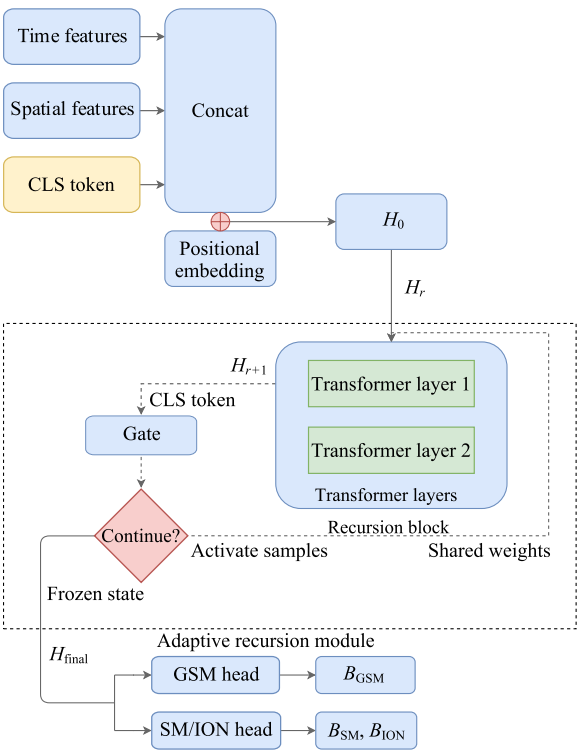

Fig. 3 Architecture of the feature fusion transformer (FFT) model for external field modeling

Feature embedding: Feature fusion strategy was employed to effectively integrate heterogeneous data sources. As shown in the stage of Fig. 3, distinct drivers, scalar solar wind parameters (time-varying) and coordinate vectors (spatial), were mapped onto a unified high-dimensional latent space via separate embedding layers. These are concatenated using a learnable classification (CLS) token to form the initial sequence. This fusion allows the model to learn the nonlinear coupling between the solar drivers and local geomagnetic responses. The input sequence H_0, is formed by concatenating the CLS tokens, time embedding, and spatial embedding. A learnable positional encoding P is then added to this sequence to distinguish between different token sources.

Adaptive recursion: Although inspired by the universal transformer, the recursive architecture was specifically adapted to address the variable complexity of geomagnetic storms.

The core of the model is a computation module with a maximum recursion depth of $R_{\max}$. In each round of recursion (from $r = 0$ to $R_{\max-1}$), the current hidden state, H_r, passes through a transformer block composed of L standard transformer encoder layers.

At the end of each recursion round ($r < R_{\max-1}$), a halting gate based on the CLS token representation (h_{CLS}, r) calculates the continuation probability p_r:

$$p_r = \sigma(W_r * h_{\mathrm{CLS}}, r + b_r), \tag{2}$$

where σ is the Sigmoid function. Only if $p_r > 0.5$ will the sample proceed to the next recursion round, $r + 1$. This mechanism allows simple samples to terminate their computations early, after a shallow recursion, whereas complex samples can leverage deeper computations to extract more intricate features. This allowed the model to dynamically allocate computational resources using shallow recursions for quasi-static quiet periods and deeper recursions for highly nonlinear storm onsets.

Multi-task prediction head: After a dynamic number of recursive computations (up to the maximum of R), CLS token (h_{CLS}, final) was obtained. A key architectural contribution is the use of two independent prediction heads to decouple source contributions, enabling identification of distinct response patterns of different field sources to external drivers: GSM head—a two-layer MLP that predicts the three components of the GSM magnetic field (B_r, B_θ, B_φ); SM/ION head—a larger two-layer MLP for joint prediction of six components of the SM and ION fields.

This design effectively isolates the gradients of distinct physical sources, preventing the dominant near-Earth fields from overshadowing the subtle signals of the distant magnetosphere during training.

2.2.3 Experiment settings

Training and validation loss curves for the internal field model (SIREN) and external field model (transformer-based), were monitored to assess optimization stability and potential overfitting (Fig. 4). Hyperparameters governing this process are listed in Table 1.

As illustrated in Fig. 4, both networks exhibit a stable convergence trajectory. The initial rapid decrease in loss can be attributed to the linear warm-up phase (15 epochs for SIREN, 5 for the transformer), followed by a smooth decline facilitated by the cosine decay learning rate scheduler.

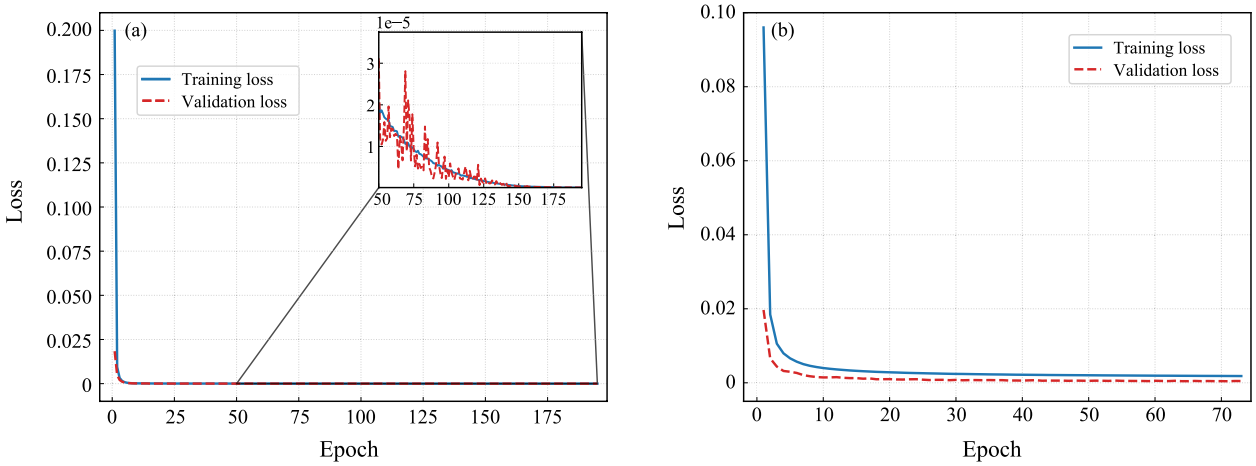

Fig. 4 Training and validation loss curves using smooth *L*1 loss. (**a**) Internal field model (SIREN) training process. The inset plot highlights the convergence details in the later stages, showing that the validation loss closely follows the training loss without diverging. (**b**) External field model (transformer-based FFT) training process. Both models utilized a cosine decay scheduler with linear warmup and were terminated via early stopping at epoch 195 and epoch 73, respectively, ensuring optimal convergence and stability

Table 1 Detailed hyperparameters and training configuration for the proposed models

Category	Parameter	Internal field model (SIREN)	External field model (FFT)
Optimization	Optimizer	AdamW	AdamW
	Peak learning rate	4×10^{-4}	4×10^{-5}
	Weight decay	1×10^{-5}	1×10^{-5}
	Batch size	32768	8192
	Loss function	Smooth $L1$ loss	Smooth $L1$ loss
	Precision	Automatic mixed precision (AMP)	Automatic mixed precision (AMP)
Scheduling	Scheduler type	Cosine decay with linear warmup	Cosine decay with linear warmup
	Warmup epochs	15	5
	Max epochs	200	120
	Early stopping patience	25 epochs	10 epochs
Architecture	Hidden dimension	384	1024 (embedding dimension)
	Network depth	8 layers	4 recursions (2 layers)
	Feedforward dimension	N/A	3072
	Attention heads	N/A	16
	Dropout rate	0	0.1
	Activation function	Sine	Gaussian error linear unit

N/A, not applicable

Specifically, for the SIREN model (Fig. 4a), the training was extended to 195 epochs. Despite the deep 8-layer architecture, both training and validation loss plateaued without exhibiting signs of overfitting. This stability is reinforced by the inset plot, which shows the validation error remains consistently low, even in later epochs.

The transformer-based external field model (Fig. 4b) converged more rapidly, triggering an early stopping mechanism (patience = 10) in epoch 73. The alignment between the training and validation curves in both figures demonstrates that the regularization techniques effectively prevented overfitting, ensuring that the models generalized well to the unseen data.

3 Results

3.1 Internal field

3.1.1 Overall

The total field was plotted on a 1° grid for January 1, 2024, with results presented in Fig. S3. On a large scale, the model demonstrates excellent consistency with the ground truth.

The dataset was computed on a 1° grid for July 1 of each year from 2000 to 2024. Spatial generalization of the internal field model was evaluated through two scenarios: 1) interpolation performance on a randomly sampled validation set; and 2) extrapolation performance over a contiguous geographical region (0°–30°E) excluded from training.

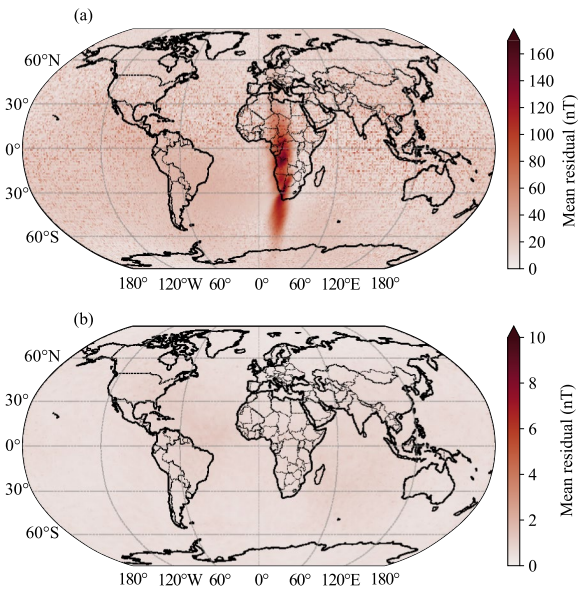

Fig. 5 Global distribution of annual mean total residuals for the internal field models (averaged over 2000–2024). (**a**) Validation of the spatial-split model. The model was trained with data excluded from the 0°–30°E longitudinal band. The prominent dark strip reveals the limited extrapolation capability of the network in the unseen region, with significant biases concentrated over Africa and the South Atlantic (residuals > 160 nT). (**b**) Performance of the final model trained on the complete dataset. It demonstrates high reconstruction accuracy globally, with a mean residual of 0.61 nT and a root mean square error (RMSE) of 0.70 nT. Different color scales were used for (**a**) (0–160 nT) and (**b**) (0–10 nT) to enhance visualization of residual distributions

Table 2 Statistical summary of prediction residuals (nT) for the spatial-split model and the final model on the test set

Model	Count	ME	STD	Min.	25th percentile	50th percentile	75th percentile	Max.	RMSE
Spatial-split model	1633525	22.13	16.65	0.090	12.78	18.20	26.83	362.86	27.69
Final model	1633525	0.61	0.34	0.004	0.37	0.55	0.77	4.64	0.70

Count, the total number of data points; ME, mean error; STD, standard deviation; min., the smallest value in the dataset; 25th percentile, the first quartile; 50th percentile, the middle value; 75th percentile, the third quartile; max., the largest value in the dataset

Figure 5 shows near perfect interpolation performance, with residual means approaching zero, and standard deviations below 0.53 nT, confirming high accuracy and stability. In contrast extrapolation to the unseen geographical region produced systematic degradation. All magnetic field components exhibit a significant systematic bias of up to −11 nT, accompanied by a more than 20-fold increase in the standard deviation of the error. Furthermore, extreme prediction failures exceeding 200 nT were observed within the test region.

The systematic discrepancy of the SIREN model in the spatial generalization test clearly reveals a key limitation of the current approach. As a purely data-driven model, it exhibits exceptional interpolation capabilities within regions covered by the training data. However, when confronted with unseen contiguous geographical areas, the lack of intrinsic physical constraints means that the model cannot guarantee that its predictions adhere to physical laws, leading to a degradation in the generalization performance.

Table 2 presents a comprehensive statistical comparison of the predicted residuals. The spatial split model yielded a high root mean square error (RMSE) of 27.69 nT and a maximum error exceeding 360 nT, quantitatively confirming its inability to extrapolate to unseen regions. In contrast, the final model maintains a low mean error (0.61 nT) and an RMSE of 0.70 nT, demonstrating robust global reconstruction capabilities.

3.1.2 Time series

To examine the performance of the model over a long time series, four representative locations were selected: Beijing (116.41°E, 39.90°N), the Equator (0°, 0°), the South Geomagnetic Pole (136.46°E, 63.91°S) (2015), and the North Geomagnetic Pole (157.39°E, 86.17°N) (2015). For each location, we plotted the temporal evolution of the residuals.

Figure 6 presents a comparative analysis of the spatiotemporal generalization capability of the model stemming from different data-partitioning strategies. The upper panels for all four locations consistently demonstrate that the predictions of the final model (red dashed line) are visually indistinguishable from the ground truth of the CHAOS-8.4 model (black line), accurately capturing the secular variation over the 25-year period.

Residual plots (bottom panels) provide critical validation. The spatial-split model (blue line) was trained on a dataset partitioned by geographic location to evaluate generalization to entirely unseen regions. This outcome validates the hypothesis that models trained on geographically restricted data cannot extrapolate unique secular variation patterns of unobserved areas.

In contrast, the final model, trained on a representative random sampling of spatiotemporal points, demonstrated robust generalization. Its residuals (red dashed line) remain stable and tightly centered around zero, with fluctuations typically within ±1 nT. For instance, at the Equator, the mean residual for the B_r component was dramatically reduced from 11.56 nT in the spatial-split experiment to just 0.06 nT in the final model. This powerful contrast confirms that a training strategy based on representative spatiotemporal sampling is essential to build a truly global and predictive model.

3.2 External field

3.2.1 Overall

The global residuals were calculated at 24 specific time points throughout 2015 (days 8th and 22nd day of each month at 12:00 UT). As shown in Table 3, the overall RMSE for the total field components are 0.59 nT, 0.71 nT, and 0.41 nT, respectively.

The primary systematic errors are concentrated in the ionospheric field and, consequently, in the total external field, where ionospheric contributions dominate. These errors appear as geographically dependent banded structures. In the B_θ component map, examples include an error band along the magnetic equator, a tilted residual pattern extending northwest to southeast in the mid-latitudes, and an annular residual pattern near the magnetic poles.

To assess spatial response in the equatorial electrojet (EEJ) region, high-resolution meridional scans (0.1° steps) were conducted at longitudes 0° (African sector) and 285° (South American sector). To maximize EEJ intensity, input drivers were fixed to active conditions ($F_{10.7}$ = 150 sfu, V_x = 450 km/s, $B_z = -5$ nT), with local time (LT) normalized to 12:00 noon. Figure S4a reveals that the peaks of the prediction residuals (ΔB_θ) shift longitudinally, perfectly aligning with the magnetic dip equator (vertical dashed lines) derived from the CHAOS-8.4 model. Figure S4b shows

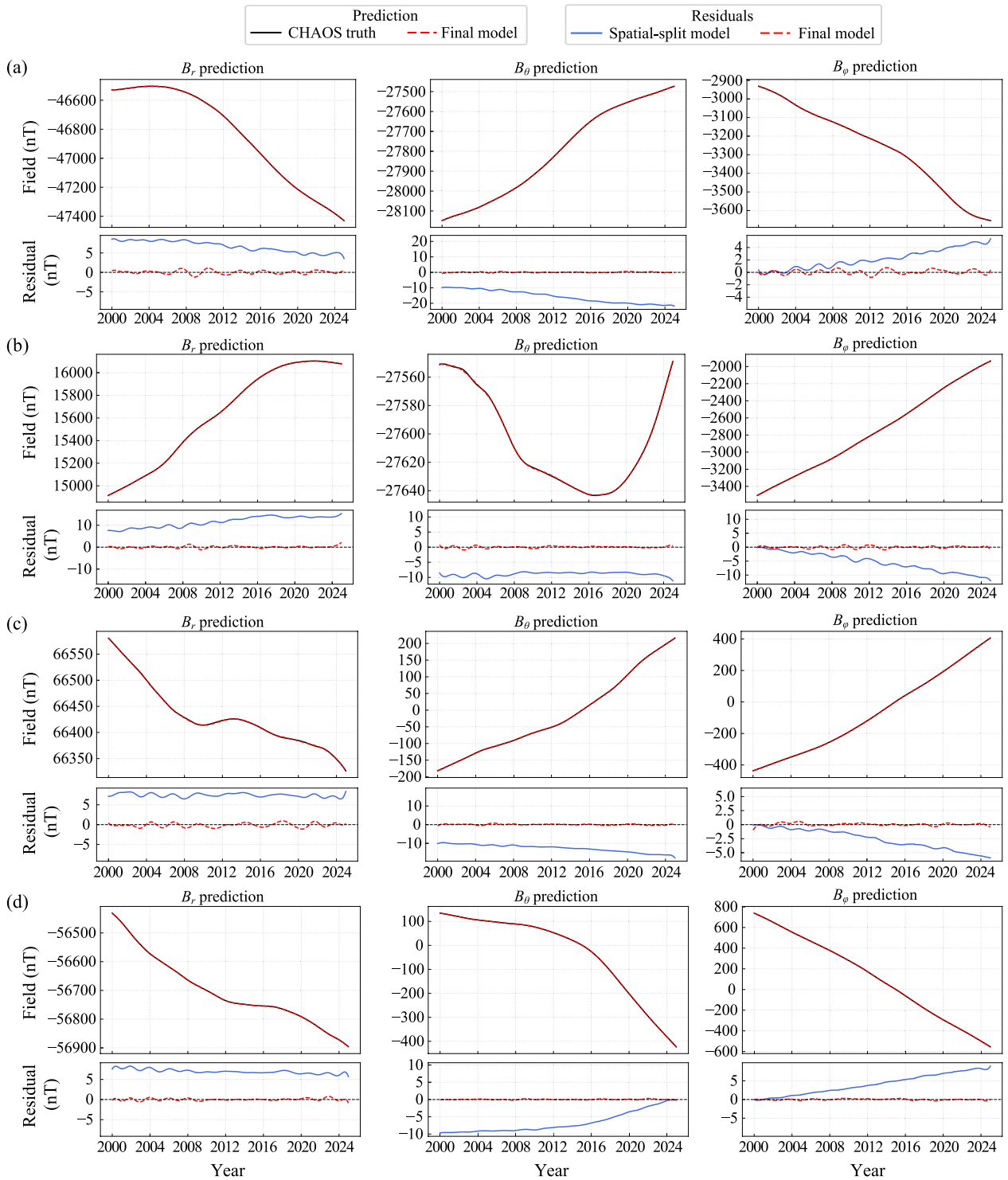

Fig. 6 Time series analysis of the internal field model predictions at four key locations from 2000 to 2024: (**a**) Beijing, (**b**) the Equator, (**c**) the South Geomagnetic Pole, and (**d**) the North Geomagnetic Pole. Each location is organized into three columns corresponding to the radial (B_r), polar (B_θ), and azimuthal (B_φ) components. The upper panel in each column compares the prediction of our final model (red dashed line) against the CHAOS-8.4 model truth (black line). The lower panel displays the residuals (prediction − truth) for both the final model (red dashed line) and a baseline spatial-split model (blue line)

Table 3 Global RMSE for the external field

Field	B_r	B_θ	B_φ
GSM	0.02	0.07	0.10
SM	0.14	0.46	0.14
ION	0.58	0.53	0.37
Total	0.59	0.71	0.41

the corresponding reference profiles, confirming that the physical displacement of the EEJ structure was the primary source of these structural errors.

The global distribution of the mean error for the external field is evaluated based on annual values from the deep learning model (Fig. 7). As demonstrated, for both GSM and SM fields, the systematic error remains nearly zero across most of the globe. This confirms that the model's learning of key physical processes, such as magnetopause, ring, and tail currents, is highly accurate and unbiased.

The structural nature of the model residuals on a global scale is shown in Fig. S5. The latitudinal distribution of the residual peaks was neither random nor symmetric about the geographic equator; instead, it formed a sinusoidal trajectory that was phase-locked with the magnetic dip equator. This precise tracking of the 'error ridge' along the physical magnetic equator indicates that while the model attempts to capture the EEJ morphology within the geographic coordinate system, the lack of explicit magnetic coordinate constraints leads to a persistent systematic misalignment in the spatial location of the predicted EEJ. This phenomenon provides definitive evidence that feature engineering based solely on geographic coordinates is insufficient for resolving geomagnetically controlled ionospheric current systems, necessitating the inclusion of magnetic coordinates to eliminate structural errors.

3.2.2 Time series

Time-series validation of the surrogate model's ability to capture rapid temporal variations was performed at Beijing station from March 23 to March 26, 2023, a period characterized by severe geomagnetic activity. Reference data were the theoretical field values synthesized by the CHAOS-8.4 model for the specified spatiotemporal coordinates, rather than direct instrumental observations. This setup enabled assessment of approximation precision relative to a physical model. As illustrated in Figs. 8 and 9, the baseline Phase 1 model (dashed blue lines) represents the initial stage of the two-phase training framework, pre-trained using only static and periodic features, such as spatial location and seasonal cycles. Owing to the lack of dynamic space weather driver inputs, such as solar wind and ring current indices, this baseline model fails to reproduce rapid fluctuations and instead generates a smooth background field.

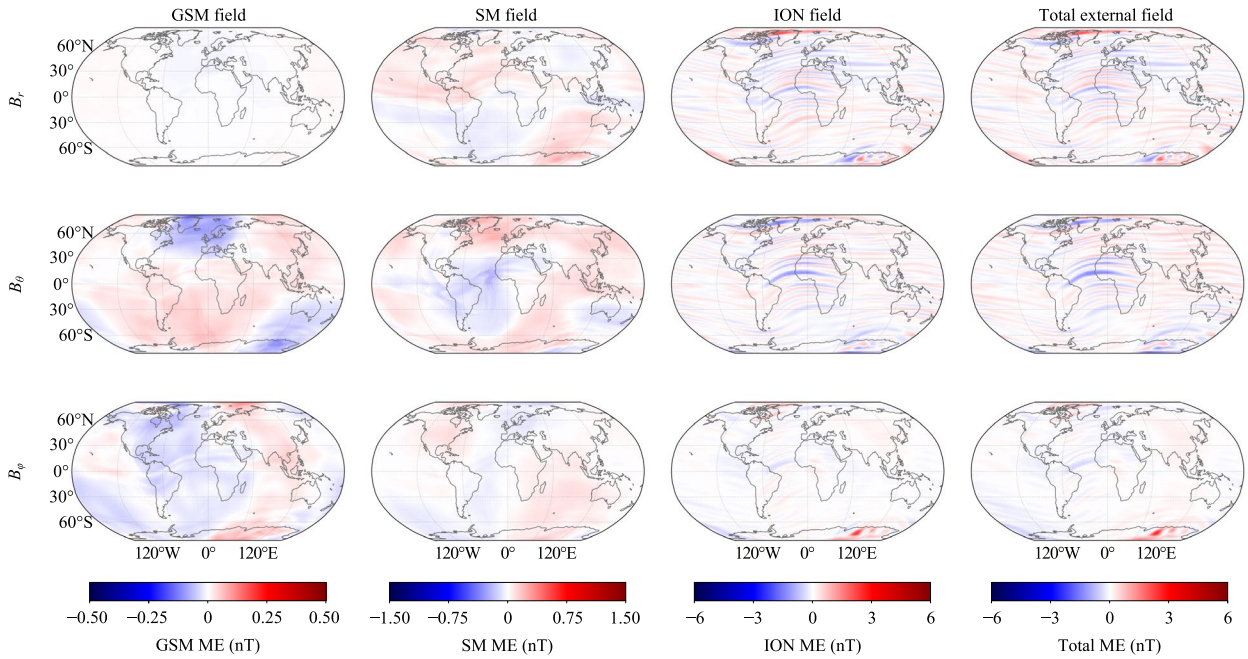

Fig. 7 Global distribution of the mean error (ME) for the external field. The maps are based on annual values from the deep learning model. For both GSM (distant magnetosphere) and SM (near-Earth magnetosphere) fields, the systematic error remains nearly zero across most of the globe, indicated by the predominance of near-white coloration

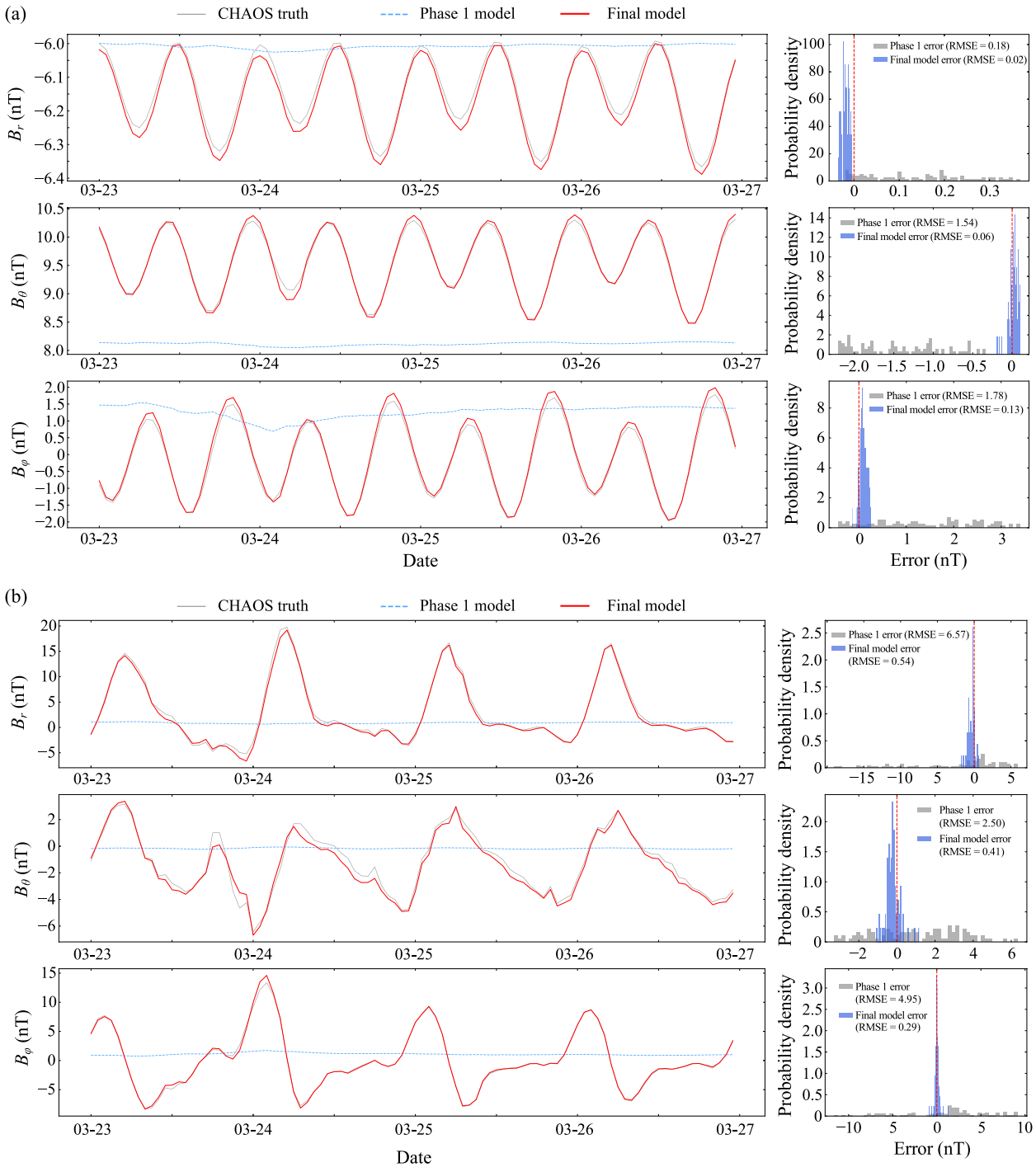

Fig. 8 Time-series validation and error analysis of external field components at Beijing (March 23–26, 2023). The panels display (from top to bottom) the magnetic field components (B_r, B_θ, B_φ) for the (**a**) GSM, (**b**) ionospheric fields. Left panels: time-series comparisons show that the final model (solid red lines) achieves high consistency with the CHAOS-8.4 ground truth (solid gray lines), significantly outperforming the phase 1 baseline (dashed blue lines). Right panels: error distribution histograms for the corresponding components (red dashed lines indicate zero bias), highlighting the substantial reduction in systematic bias and RMSE achieved by the final model

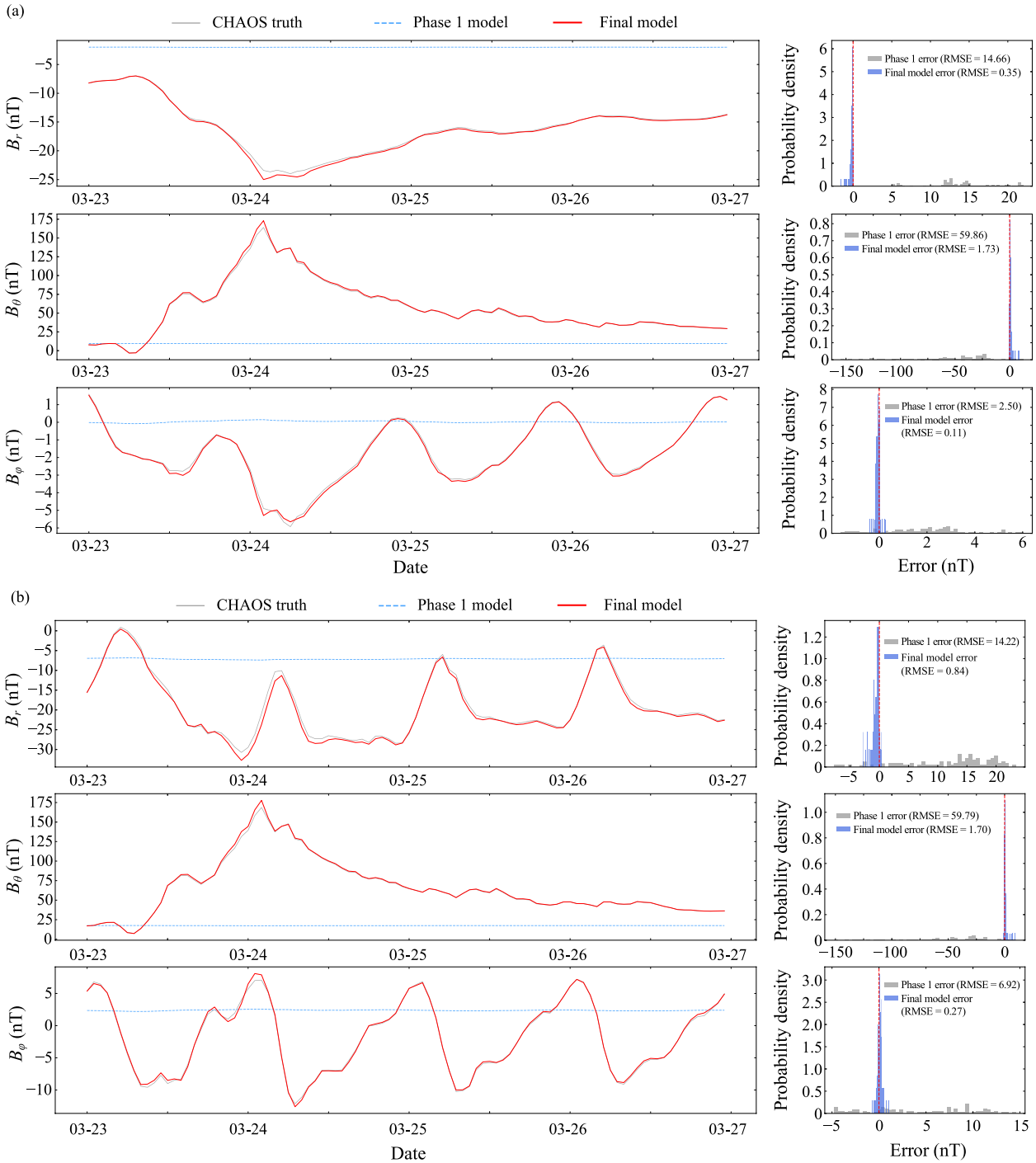

Fig. 9 Same as Fig. 8, but for (**a**) SM and (**b**) total external

4 Discussion

4.1 Efficiency improvements

This section aims to quantitatively evaluate the computational efficiency advantages of the proposed deep learning models (SIREN and transformer) over the traditional physical model (CHAOS-8.4, as implemented in ChaosMagPy). The assessment focused exclusively on forward inference and excluded training costs. Tests were conducted on an AMD Ryzen 9900×CPU with a DDR5 6000 cl28 24 GB×2 memory.

The ChaosMagPy baseline was heavily optimized for CPU execution. Porting the core spherical harmonic

algorithms into a GPU environment requires substantial structural rewrites that fall outside the scope of this study. Consequently, GPU-based performance metrics reported for the deep learning models are not intended as direct comparison with ChaosMagPy but rather as demonstration of the inherent scalability of deep learning frameworks on modern parallel computing hardware.

The SIREN model achieved substantial acceleration in internal field computations. Single-point evaluation (batch size = 1) was approximately 1110 times faster than that of ChaosMagPy (Fig. S6) underscoring its suitability for low-latency applications. As the batch size increased, ChaosMagPy's vectorized calculations became more efficient, causing the relative speedup factor of SIREN to decrease. Nevertheless, SIREN maintained a significant advantage, remaining about 2.3 times faster even at a batch size of 10000.

Conversely, the transformer model for the external field task was slower than the optimized batched ChaosMagPy implementation on the CPU (Fig. S6). The architectural complexity of the transformer, necessary for capturing external field physics, introduced significant computational overhead that outweighs direct calculations on a CPU. Earlier claims of speedup were artifact of flawed benchmarking. These results underscore that for CPU-bound batch-processing workflows, the traditional physics-based model for the external field remains more efficient.

The performance of the SIREN model scales significantly with the parallel processing power of the GPU (Fig. S7). The speedup factor increased with size, exceeding 1100 times initially and reaching an extraordinary 12700× at a batch size of 100000. This demonstrated that architecture is highly effective for large-scale parallel computations.

The transformer model also transitioned from CPU liability to GPU assets, though with a more nuanced speedup profile. Its peak relative gain exceeded 11.5× at a batch size = 1, but this factor declined with larger batches stabilizing at approximately 2.3 times. This reduction occurred because the CPU-based ChaosMagPy baseline became significantly more efficient per point owing to vectorization, thereby reducing the relative performance margin of the GPU. Despite this diminishing relative gain, the transformer on a GPU consistently maintained a notable absolute performance advantage, making it a high-performance alternative until it encountered out-of-memory (OOM) limitations around a batch size of 10000.

Overall, the dual-architecture approach offers hardware and task dependent computational improvements. The SIREN model provides a universally applicable speed-up for internal field modeling, making it ideal for both real-time single-point queries on a CPU and massive-scale computations on a GPU. The transformer model, which is not a viable replacement for CPU-based external field calculations, is a powerful and highly efficient solution when deployed on a GPU. It offers the most substantial relative gains for low-latency, small-batch tasks while still providing a consistent, albeit smaller, speedup for large-batch processing. This highlights the critical importance of co-designing neural network architectures with target hardware platforms and workload characteristics to maximize scientific and operational impacts (Krizhevsky et al. 2012).

4.2 Analysis of residuals versus physical drivers

As a system transitions from a linear quasi-static quiet state to a highly nonlinear dynamic storm phase, predictability declines sharply. This section examines the relationship between the model residuals and two external drivers, the RC-index and the solar $F_{10.7}$ radio flux index, to evaluate its predictive fidelity. A binned statistical analysis was applied to an independent dataset, calculating RMSE across parameter intervals.

The relationship between model error in the radial component of the SM field (B_r^{SM}) and the RC-index was evaluated, with results plotted in Fig. S8. The model's RMSE is only 0.21 nT during geomagnetically quiet periods (-20 nT < RC-index ≤ 0 nT). As geomagnetic activity intensifies, the RMSE systematically increases, rising to 0.28 nT during moderately active periods (-50 nT < RC-index ≤ -20 nT) and reaching 1.78 nT during major magnetic storms (-500 nT < RC-index ≤ -150 nT). This trend aligns with physical expectations. Intense geomagnetic storms are typically initiated by magnetic reconnection between the southward IMF and Earth's magnetic field, leading to substantial energy injection into the magnetosphere. This energy enhances the ring current, making spatiotemporal variations in the magnetic field more complex and nonlinear (Dungey 1961).

A critical question is whether the observed error increase simply reflects larger signal perturbations during storm (i.e., larger data gradients) or indicates genuine degradation in model performance. This distinction can be addressed by comparing the error growth rates with the signal growth rates. From quiet periods to severe storms, the RMSE of the model increases by a factor of approximately 8.6 (1.79/0.21). Over the same transition, the magnitude of the RC-index signal increased by a factor of 10–50 (from 10–20 nT to 150–500 nT). The slower growth rate of model error relative to the physical signal demonstrates that model maintains a high signal-to-noise ratio even under extreme geomagnetic storms, confirming its robustness to extreme space weather events.

The solar $F_{10.7}$ index serves as a crucial proxy for solar EUV radiation, the primary driver of ionospheric

ionization. Therefore, analyzing the relationship between the model's error in the ionospheric field and the $F_{10.7}$ index directly validates the model's ability to capture the physics of Sun-Earth coupling.

The relationship between error in azimuthal ionospheric field component $B_{\varphi}^{\mathrm{ION}}$ and the $F_{10.7}$ index is illustrated in Fig. S9. The positive correlation indicates that model error scales with solar activity, confirming that the model effectively leverages this input feature to adjust its predictions of the ionospheric magnetic field strength.

4.3 Comparison with traditional spherical harmonic approach

In this study, the CHAOS-8.4 model functioned both as the ground truth for training and as a state-of-the-art traditional modeling benchmark based on spherical harmonic analysis (SHA). A comprehensive comparison was conducted between the proposed deep learning framework and the traditional SHA method, focusing on accuracy and computational efficiency.

In terms of accuracy, the internal field model achieves an RMSE of 0.70 nT (Table 2), and the external field model maintains a global RMSE of approximately 0.41 nT (Table 3). These residuals are negligible compared with the total field intensity, indicating that our model successfully reproduces the high-fidelity reconstruction capabilities of the traditional SHA method.

Efficiency gains are more pronounced. As detailed in Section 4.1, the SIREN-based internal field model achieved a speedup of approximately 1100 times compared with the CPU-based SHA implementation (ChaosMagPy). Similarly, the transformer-based external field model demonstrated substantial GPU acceleration for large-batch processing. This comparison highlights the core advantage of the proposed framework: it retains the precision of traditional physical–mathematical models while eliminating computational bottlenecks, positioning it a superior candidate for real-time applications. To contextualize the performance of the deep learning surrogate, a detailed statistical breakdown of the prediction residuals across multiple International Real-time Magnetic Observatory Network (INTERMAGNET) validation stations is presented in Table 4. Distinguishing static spatial representation from dynamic temporal tracking is crucial when evaluating global geomagnetic field models. As shown in Table 4, the traditional CHAOS-8.4 model often yields lower RMSE at specific stations (e.g., Alice Springs (ASP) and Abisko (ABK)), primarily due to the mean bias metric. The physics-based CHAOS-8.4 model explicitly incorporates high-resolution static lithospheric field components, which results in lower systematic baseline offsets. In contrast, the deep learning model, constrained by grid resolution and coordinate representations, exhibited slightly higher static biases for certain localized anomalies.

However, the primary objective of an external field module is to capture the highly nonlinear and rapid temporal variations driven by solar activity. The STD provides a more rigorous metric for dynamic tracking capability from static background offsets. As demonstrated in Table 4 and corroborated by the bar charts in Fig. 10, the deep learning model achieves highly competitive STD values across most stations. In the global aggregate (all stations), it maintains nearly identical STD for the *Y*-component (202.47 nT vs. 202.48 nT) and outperforms the CHAOS-8.4 baseline in the *Z*-component (387.71 nT vs. 407.51 nT). Furthermore, the *R* approaching 1.0 for the globally aggregated data confirms that the morphological evolution of the predicted field is strongly correlated with the ground truth. This statistical evidence aligns with the time-series validation shown in Figs. 8 and 10, where deep learning predictions accurately track complex high-frequency fluctuations during geomagnetic storms once the static offsets are corrected. To further validate the model's performance during geomagnetically active windows, we compared our predictions against 1-day resolution ground observations (Fig. 11). The tight alignment between the prediction curve and the observed data points confirms that the model accurately reproduces geomagnetic dynamics and short-term disturbances.

4.4 Clarification on practicality and efficiency

Although the framework depends on source-separated data from CHAOS-8.4 as a surrogate model, this dependency does not limit its operational practicality. Its primary value lies in extreme computational acceleration, enabling real-time applications that are impossible using the original physics-based model.

As quantified in Section 4.1, the proposed SIREN architecture achieves a speedup of over 1000 times on a CPU and up to 12700 times on a GPU compared to the traditional CHAOS model (see Figs. S6 and S7). This transforms high-fidelity field calculations from a computational burden into negligible latency. Although the external field transformer incurs CPU overhead, it effectively leverages GPU parallelism for large-scale batch processing.

Thus, the practicality of the dual framework is not in generating new physical models from raw data (inversion), but in deploying existing high-fidelity models on resource-constrained platforms (e.g., embedded navigation systems) and enabling massive-scale simulations with minimal time cost.

Table 4 Statistical comparison of model performance. Quantitative comparison of the proposed deep learning model (ours) to the CHAOS-8.4 model

Station	Component	*N*	RMSE (ours)	RMSE (CHAOS)	STD (ours)	STD (CHAOS)	Mean bias (ours)	Mean bias (CHAOS)	*R* (ours)
ABG	*X*	812	70.07	87.86	12.05	3.39	−69.02	−87.80	0.29
	Y	812	412.75	409.25	3.00	2.55	412.73	409.25	0.97
	Z	812	302.25	376.97	2.06	2.55	302.25	376.96	1.00
ARS	*X*	796	572.03	430.47	8.89	4.72	571.96	430.45	0.79
	Y	796	192.46	195.71	6.34	5.95	−192.35	−195.62	0.64
	Z	796	150.25	182.31	4.39	2.94	150.18	182.29	0.97
ASP	*X*	807	248.29	155.85	10.44	3.91	248.07	155.80	0.41
	Y	807	6.12	4.10	3.41	3.05	−5.08	−2.73	0.90
	Z	807	467.06	395.96	3.35	1.32	467.05	395.96	0.48
ABK	*X*	759	462.56	358.09	49.22	48.92	459.93	354.74	0.33
	Y	759	89.38	81.00	16.21	16.17	87.90	79.37	0.69
	Z	759	92.71	84.16	27.64	28.14	88.50	79.32	0.37
API	*X*	844	68.37	68.93	11.76	3.99	−67.35	−68.82	0.41
	Y	844	197.42	199.91	3.96	3.16	197.38	199.88	0.59
	Z	844	641.30	695.52	3.39	2.73	−641.29	−695.52	0.76
All	*X*	4018	345.83	261.48	265.23	213.20	221.92	151.38	1.00
	Y	4018	226.87	225.99	202.47	202.48	102.35	100.38	1.00
	Z	4018	393.40	412.00	387.71	407.51	66.65	60.71	1.00

STD, standard deviation; *R*, Pearson correlation coefficient; ABG, Alibag; ARS, Arthurs Seat; API, Apia

4.5 Summary

In summary, residuals analysis indicated that the proposed deep learning model achieved high overall accuracy while maintaining strong physical consistency in response to key geophysical parameters, such as RC-index and $F_{10.7}$. These findings highlight the scientific reliability and utility of the model for geophysical research.

Several limitations remain. The internal field model exhibits systematic bias when extrapolated to contiguous geographical regions not included in training, revealing limited spatial generalization. Additionally, the model is constrained to the Earth's surface (radius $R = 6371.2$ km) and does not account for the radial variation in the magnetic field across altitudes. These limitations define clear directions for future research.

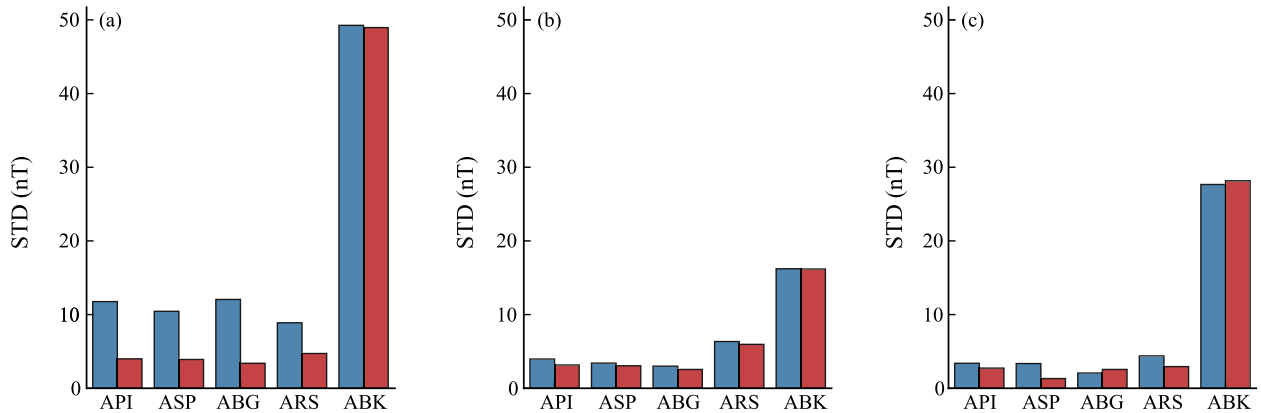

Fig. 10 Comparison of prediction precision between the proposed deep learning model and the CHAOS-8.4 model. The labels on the *x*-axis represent the standard IAGA codes (British Geological Survey 2025) for the geomagnetic observatories used in this study. The bar chart displays the standard deviation of prediction errors (STD) for the (**a**) *X*, (**b**) *Y*, and (**c**) *Z* geomagnetic components across five INTERMAGNET validation stations. The STD metric is used here to evaluate the model's precision in capturing temporal dynamics while excluding the influence of static crustal biases. Blue bars represent our deep learning surrogate model, and red bars represent the physics-based CHAOS-8 model. Lower values indicate higher precision. The deep learning model demonstrates competitive performance, achieving comparable STD values than CHAOS-8 in most cases (e.g., global *Z*-component), confirming its robustness in tracking dynamic field variations

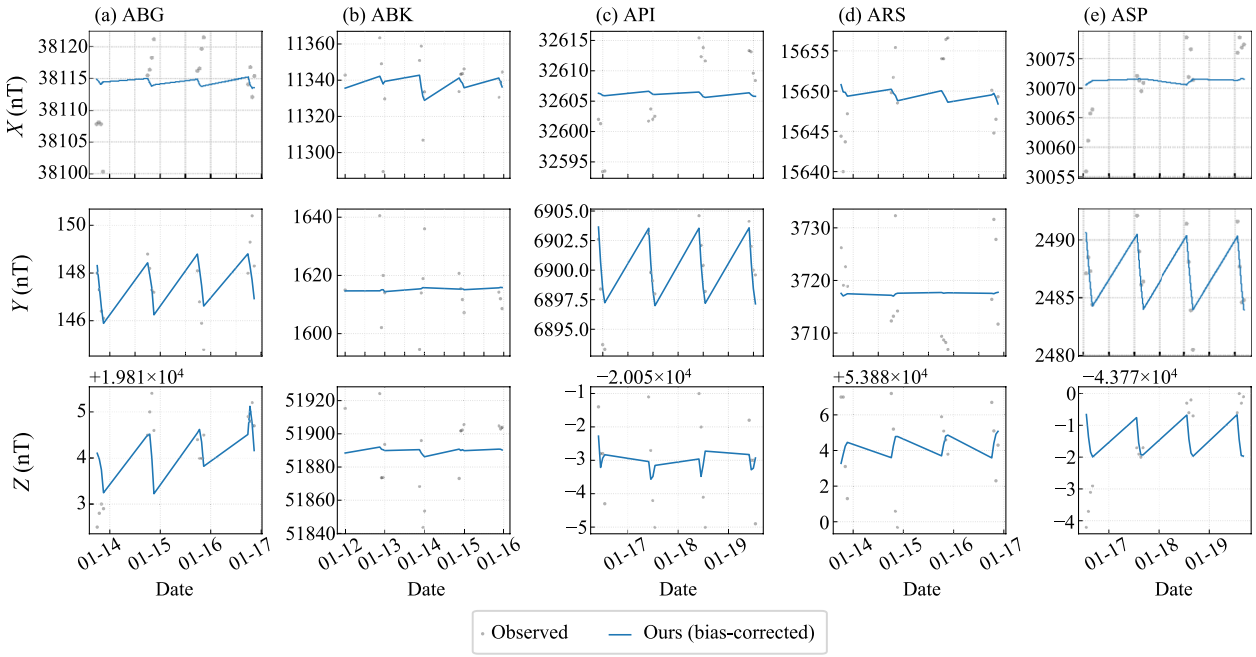

Fig.11 Time-series validation against ground observations. Comparison of model predictions with 1-min resolution observed data during a geomagnetically active window. Gray scatter points represent the observed measurements, displayed with transparency to show data density. The Blue solid line indicates our model's prediction after applying a static bias correction (meaning removal). The tight alignment between the prediction curve and the observed data points confirms that the model accurately reproduces complex geomagnetic dynamics and short-term disturbances

Future research should proceed along four primary avenues. First, physics-informed neural networks will be incorporated by embedding physical constraints, such as the divergence-free condition of the magnetic field ($\nabla \cdot B = 0$), as a regularization term in the loss function. Given that SIREN's smooth and analytically differentiable nature of SIREN has already shown exceptional performance in solving partial differential equations, suggesting that integration with physics-informed neural networks will enhance the physical consistency and spatial generalization of the model. Second, input will be extended to three-dimensional spatial coordinates to construct a complete 3D structural model of the geomagnetic field. Third, the external field model, currently simplified relative to CHAOS-8.4 (particularly for the outer magnetosphere in the GSM framework), will be updated by referencing more sophisticated models, such as the Tsyganenko model. Fourth, optimization techniques including model quantization and pruning will be explored to enable efficient deployment on edge-computing devices and online forecasting platforms.

5 Conclusions

A dual-architecture deep learning framework was developed by tailoring specific neural architectures to the distinct physical characteristics of field sources. Key findings include:

1) Internal field modeling: SIREN provided superior continuous representation of the quasi-static internal field. This mesh-free approach achieved an inference speedup of approximately 1100 times compared with traditional spherical harmonic implementations on CPUs, effectively resolving the bottleneck for real-time global navigation applications.
2) External field modeling: The proposed transformer, which incorporates a novel adaptive recursion mechanism, successfully captured the highly dynamic external field. Dynamic depth allocation enabled autonomous balancing of computational costs with signal complexity, ensuring a high-fidelity reconstruction of nonlinear storm-time responses while avoiding unnecessary overhead during quiet periods.

3) Framework validation: Results validate the potential of deep learning as a robust surrogate for complex geophysical simulations. The established framework establishes a scalable paradigm for future space weather monitoring systems and supports integration of multi-source physical constraints into real-time forecasting pipelines.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1007/s44295-026-00105-7.

Supplementary Material 1.

Acknowledgements
This work was supported by the Open Foundation Project of the Key Laboratory of Polar Geology and Marine Mineral Resources (China University of Geosciences, Beijing), Ministry of Education (No. PGMR-2025-101), and the National Natural Science Foundation of China (No. 42076224).
We acknowledge the NASA/GSFC Space Physics Data Facility for the OMNI data, and the PIs of the ACE, Wind, IMP-8, and Geotail missions. $F_{10.7}$ data were obtained via LISIRD, provided by the Dominion Radio Astrophysical Observatory (NRC Canada). We thank the World Data Center for Geomagnetism (Edinburgh) and the institutes operating the Alibag, Apia, Arti, Abisko, and Alice Springs observatories for providing magnetic data. Finally, we thank DTU Space, particularly C. C. Finlay and N. Olsen, for providing the RC-index and the CHAOS model.

Additional information
Edited by: Lin Gao.

Authors' contributions
Zhijin Bao: conceptualization, methodology, software, validation, formal analysis, writing–original draft, and visualization. Junhui Xing: conceptualization, resources, writing–review and editing, supervision, project administration. Haowei Xu: data curation, investigation, validation. Jiayi Wei: software, visualization, writing–review and editing. Chong Xu: conceptualization, software.

Data availability
The datasets generated and analyzed in the current study and codes used in this study are available from the corresponding author upon reasonable request.
The datasets analyzed during the current study are available in the following public repositories: CHAOS-7/8 Model and RC-index: Provided by DTU Space, available at http://www.spacecenter.dk/files/magnetic-models/CHAOS-7/ (accessed via the ChaosMagPy Python package https://doi.org/10.5281/zenodo.3352398). OMNI data (IMF and Solar Wind): Sourced from NASA's Space Physics Data Facility (SPDF) OMNIWeb service at https://omniweb.gsfc.nasa.gov/. $F_{10.7}$ index: Provided by the Dominion Radio Astrophysical Observatory (DRAO) and accessed via the LASP Interactive Solar Irradiance Data Center (LISIRD) at https://lasp.colorado.edu/lisird/.

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
Informed consent for publication was obtained from all participants.

Competing interests
The authors declare that they have no conflict of interests. Junhui Xing is one of the Editorial Board Members, but he was not involved in the journal's review of, or decision related to, this manuscript.

Received: 13 December 2025 Revised: 17 April 2026 Accepted: 27 April 2026
Published online: 21 May 2026

References

Bergen KJ, Johnson PA, de Hoop MV, Beroza GC (2019) Machine learning for data-driven discovery in solid Earth geoscience. Science 363(6433):eaau0323. https://doi.org/10.1126/science.aau0323
British Geological Survey (2025) World Data Centre for Geomagnetism, Edinburgh. Retrieved from https://wdc-dataportal.bgs.ac.uk/
Camporeale E (2019) The challenge of machine learning in space weather: nowcasting and forecasting. Space Weather 17(8):1166–1207. https://doi.org/10.1029/2018SW002061
Chen Z, Tian H, Li HM, Tang RX, Ouyang ZH, Deng XH (2023) The study on the global evolution of energetic electron precipitation during geomagnetic storm based on deep learning algorithm. J Geophys Res-Space Phys 128(3):e2022JA030974. https://doi.org/10.1029/2022JA030974
Conde D, Castillo FL, Escobar C, García C, García JE, Sanz V et al (2023) Forecasting geomagnetic storm disturbances and their uncertainties using deep learning. Space Weather 21(11):e2023SW003474. https://doi.org/10.1029/2023SW003474
Devlin J, Chang MW, Lee K, Toutanova K (2019) BERT: pre-training of deep bidirectional transformers for language understanding. In: Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers). Association for Computational Linguistics, pp 4171–4186. https://aclanthology.org/N19-1423/
Dungey JW (1961) Interplanetary magnetic field and the auroral zones. Phys Rev Lett 6(2):47–48. https://doi.org/10.1103/PhysRevLett.6.47
Finlay CC, Olsen N, Tøffner-Clausen L (2015) DTU candidate field models for IGRF-12 and the CHAOS-5 geomagnetic field model. Earth Planets Space 67(1):114. https://doi.org/10.1186/s40623-015-0274-3
Graves A (2016) Adaptive computation time for recurrent neural networks. Preprint at arXiv:1603.08983
Hulot G, Balogh A, Christensen UR, Constable CG, Mandea M, Olsen N (2010) The Earth's magnetic field in the space age: an introduction to terrestrial magnetism. Space Sci Rev 155(1–4):1–7. https://doi.org/10.1007/s11214-010-9703-6
King JH, Papitashvili NE (2005) Solar wind spatial scales in and comparisons of hourly wind and ACE plasma and magnetic field data. J Geophys Res-Space Phys 110(A2):A02104. https://doi.org/10.1029/2004JA010649
Kloss C (2024) ChaosMagPy: Python package for the CHAOS geomagnetic field model. Zenodo. https://doi.org/10.5281/zenodo.16091680
Kloss C, Beggan C (2025) The 14th generation of the international geomagnetic reference field. In: EGU General Assembly Conference Abstracts. EGU25-3566. https://doi.org/10.5194/egusphere-egu25-3566
Kloss C, Finlay CC, Olsen N, Tøffner-Clausen L, Gillet N, Grayver A (2026) The CHAOS-8 geomagnetic field model. Earth Planets Space 78:21. https://doi.org/10.1186/s40623-025-02352-0
Krizhevsky A, Sutskever I, Hinton GE (2012) ImageNet classification with deep convolutional neural networks. In: Advances in Neural Information Processing Systems 25 (NIPS 2012). NIPS, pp 1–9. https://papers.nips.cc/paper_files/paper/2012/hash/c399862d3b9d6b76c8436e924a68c45b-Abstract.html
Kuslits L, Horváth A, Wesztergom V, Beggan C, Rubóczki T, Prácser E et al (2024) An alternate representation of the geomagnetic core field obtained using machine learning. Earth Planets Space 76(1):77. https://doi.org/10.1186/s40623-024-02024-5
Laperre B, Amaya J, Lapenta G (2018) Prediction of the Dst-index using a long short-term memory network. https://www.iihe.ac.be/~dldecker/Doc/Abstracts/AstroGeoPlasma/Oral/abstract_Brecht_Laperre.pdf
Li G, Zhou XH, Chen CJ, Xu LA, Zhou F, Shi FS et al (2023) Multitype geomagnetic noise removal via an improved U-Net deep learning network. IEEE Trans Geosci Remote Sens 61:5916512. https://doi.org/10.1109/TGRS.2023.3307422
Li J, Chen B (2025) The secular variation prediction method of geomagnetic field in Chinese mainland based on long short-term memory neural network. Acta Seismol Sin 47(3):390–409. https://doi.org/10.11939/jass.20240013

Liu H, Liu Z, Liu S, Liu YH, Bin JC, Shi F et al (2019) A nonlinear regression application via machine learning techniques for geomagnetic data reconstruction processing. IEEE Trans Geosci Remote Sens 57(1):128–140. https://doi.org/10.1109/TGRS.2018.2852632

Merrill RT, McElhinny MW, McFadden PL (1998) The magnetic field of the earth: paleomagnetism, the core, and the deep mantle, vol 63. Academic Press

Otzen M (2022) Geostatistical simulation and deep learning in geomagnetism. PhD thesis, Technical University of Denmark, Kgs. Lyngby

Pelkum Donahue K, Inceoglu F (2024) Forecasting solar flares with a transformer network. Front Astron Space Sci 10:1298609. https://doi.org/10.3389/fspas.2023.1298609

Ren XD, Yang PX, Mei DK, Liu H, Xu GZ, Dong Y (2023) Global ionospheric TEC forecasting for geomagnetic storm time using a deep learning-based multi-model ensemble method. Space Weather 21(3):e2022SW003231. https://doi.org/10.1029/2022SW003231

Sakpal S (2024) Prediction of space weather events through analysis of active region magnetograms using convolutional neural network. Preprint at arXiv:2405.02545

Sitzmann V, Martel JNP, Bergman AW, Lindell DB, Wetzstein G (2020) Implicit neural representations with periodic activation functions. In: 34th Conference on Neural Information Processing Systems (NeurIPS). NIPS, pp 7462–7473

Tapping KF (2013) The 10.7 cm solar radio flux ($F_{10.7}$). Space Weather 11(7):394–406. https://doi.org/10.1002/swe.20064

Tasistro-Hart A, Grayver A, Kuvshinov A (2021) Probabilistic geomagnetic storm forecasting via deep learning. J Geophys Res-Space Phys 126(1):e2020JA028228. https://doi.org/10.1029/2020JA028228

Tóth G, Sokolov IV, Gombosi TI, Chesney DR, Clauer CR, De Zeeuw DL et al (2005) Space weather modeling framework: a new tool for the space science community. J Geophys Res-Space Physics 110(A12):A12226. https://doi.org/10.1029/2005JA011126

Upendran V, Tigas P, Ferdousi B, Bloch T, Cheung MC, Ganju S et al (2022) Global geomagnetic perturbation forecasting using deep learning. Space Weather 20(6):e2022SW003045. https://doi.org/10.1029/2022SW003045

Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez AN et al (2017) Attention is all you need. In: 31st Conference on Neural Information Processing Systems (NIPS 2017). NIPS, pp 1–11. https://proceedings.neurips.cc/paper_files/paper/2017/file/3f5ee243547dee91fbd053c1c4a845aa-Paper.pdf

Wintoft P, Wik M (2021) Exploring three recurrent neural network architectures for geomagnetic predictions. Front Astron Space Sci 8:664483. https://doi.org/10.3389/fspas.2021.664483

Yang J (2024) Research on geomagnetic signal denoising in geomagnetic navigation based on deep learning. Master thesis, Chang'an University, Xi'an

Zhang JY, Feng Y, Zhang JX, Li YJ (2023) The short time prediction of the Dst index based on the long-short time memory and empirical mode decomposition–long-short time memory models. Appl Sci 13(21):11824. https://doi.org/10.3390/app132111824

Publisher's Note