Research Square

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

Physics-Informed Deep Learning for Short-Term Prediction of Solar Active Region Vector Magnetic Field Evolution

Yuqing Zhou
Yunnan Observatories
Hui Liu
liuhui@ynao.ac.cn
Yunnan Observatories
Zhenyu Jin
Yunnan Observatories
Yuyang Li
National Astronomical Observatories
Sizhong Zou
Yunnan Observatories
Jiaben Lin
National Astronomical Observatories
Mingfu Shao
National Astronomical Observatories
Zhuoheng Huang
Yunnan Observatories

Research Article

Keywords: Solar magnetic fields, Neural networks, Active region magnetic fields, Space weather, Deep learning

Posted Date: February 26th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-8931595/v1

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

Additional Declarations: No competing interests reported.

Version of Record: A version of this preprint was published at Solar Physics on June 10th, 2026. See the published version at https://doi.org/10.1007/s11207-026-02687-1.

Physics-Informed Deep Learning for Short-Term Prediction of Solar Active Region Vector Magnetic Field Evolution

Yuqing Zhou[1,2] · Hui Liu[1,2] · Zhenyu Jin[1,2] · Yuyang Li[2,3] · Sizhong Zou[1,2] · Jiaben Lin[2,3] · Mingfu Shao[2,3] · Zhuoheng Huang[1,2]

© **The author(s)** ●●●●

Abstract Forecasting the dynamic evolution of solar magnetic fields is a critical technique for enabling space weather warnings. Addressing the limitations of existing methods in predicting all vector magnetic field components, preserving local structural fidelity, and ensuring physical interpretability, this study proposes a deep learning prediction method that integrates dynamic masks of active regions with multiple physical parameter constraints. By constructing a three-channel representation of vector magnetic fields, applying dynamic masks to enhance attention to strong-field regions, and incorporating multi-parameter physical constraints, we developed an end-to-end short-term (12-hour) predictive model of solar vector magnetic field evolution. Using SDO/SHARPS vector magnetogram data, the model predicts and analyzes field evolution across all components. Quantitative evaluations demonstrate that our approach achieves a structural similarity index of 0.85 for the radial component and maintains unsigned magnetic flux prediction errors at 11.37% (95% CI: ±0.06%). These results effectively improve structural fidelity and physical consistency, offering

✉ H. Liu
liuhui@ynao.ac.cn

Y. Zhou
zhouyuqing@ynao.ac.cn

Z. Jin
jinzhenyu@ynao.ac.cn

[1] Yunnan Observatories, Chinese Academy of Sciences, Kunming 650216, People's Republic of China

[2] University of Chinese Academy of Sciences, Beijing 100049, People's Republic of China

[3] National Astronomical Observatories, Chinese Academy of Sciences, Beijing 100101, People's Republic of China

accurate and reliable technical support for space weather forecasting and early warnings.

Keywords: Solar magnetic fields; Neural networks; Active region magnetic fields; Space weather; Deep learning

1. Introduction

The ability to forecast solar magnetic field evolution is crucial for improving the accuracy and timeliness of space weather warnings. In recent years, significant progress has been made in both physics-based modeling and data-driven approaches:

Physics-based modeling has made substantial progress: local reconstruction of active region magnetic fields has revealed precursors of sudden X-10 class flares through strong shear flows along polarity inversion lines (Yang et al. 2004). Surface flux transport (SFT) models focus primarily on the radial component (Wang et al. 2004), often neglecting horizontal components critical for magnetic reconnection and shear. Jiang et al. (Jiang, Cameron, and Schüssler 2014) used a Monte Carlo inclination divergence method to simulate large-scale field evolution, while oversimplification of horizontal components in transport models can lead to biases in short-term polar field variations (Jeong et al. 2025a).

Data-driven approaches have also advanced rapidly. Covas (Covas, Peixinho, and Fernandes 2019) applied transfer learning for spatiotemporal prediction but primarily modeled scalar radial fields, lacking full vector outputs. Bai et al. (Bai et al. 2020) developed deep learning models targeting single components. Baty and Vigon (Baty and Vigon 2024) proposed physics-informed neural networks (PINNs) to reconstruct coronal fields, though these methods address reconstruction rather than short-term forecasting of surface vector fields.

Despite these advances, existing methods face key limitations in meeting the high standards of space weather warnings that require complete vector representation, fidelity of critical structures, and consistency with physical parameters:

1. **Incomplete vector field prediction.** Upton & Hathaway's model (Upton and Hathaway 2014) handled only the radial component while ignoring horizontal fields. SFT models (Jeong et al. 2025a) simplify horizontal components, causing biases in polar field evolution. Data-driven models (Covas, Peixinho, and Fernandes 2019; Bai et al. 2020) also focus primarily on scalar radial fields. Jeong et al. (Jeong et al. 2025b) introduced a Pix2PixCC architecture predicting radial evolution for up to 27 days without horizontal outputs, limiting calculation of horizontal-field-dependent quantities such as current helicity. Ramunno et al. (Ramunno et al. 2024) applied diffusion probabilistic models to predict full-disk line-of-sight magnetograms but still lack complete vector outputs needed to characterize three-dimensional topology.
2. **Insufficient focus on core structures of active regions.** Traditional physical models (Miesch and Teweldebirhan 2016) may introduce topological differences due to treatment of turbulent diffusion coefficients, distorting small-scale island structures. Yang et al. (Yang et al. 2004) examined strong

shear flows in active region cores but lacked comprehensive modeling elsewhere. Jeong et al. (Jeong et al. 2025b,a) reproduced differential rotation and meridional flows but their predictions of magnetic shear near polarity inversion lines remained blurred. Ramunno et al. (Ramunno et al. 2024) improved predictions of AR-scale flux and PIL lengths but struggled with gradient changes in highly sheared regions.

3. **Lack of physical parameter constraints.** Data-driven models (Covas 2020; Bai et al. 2020) often omit sufficient physical constraints, risking outputs that violate flux conservation or Lorentz force balance. Jeong et al. (Jeong et al. 2025b) introduced a flux-consistency loss but excluded constraints such as Lorentz force balance and helicity conservation. Ramunno et al. (Ramunno et al. 2024) included checks such as flux and AR area without cross-scale consistency. Physics-based models, though consistent with theory, often restrict themselves to single-parameter constraints (Fan 2015) and lack multi-parameter integration.

Current approaches are therefore insufficient for reliable space weather forecasting that requires both high prediction accuracy and physical interpretability. This limitation increases the risk of mischaracterizing the severity of coronal mass ejections (Kilpua et al. 2019).

To address these challenges, we propose a physics-informed deep learning architecture that integrates dynamic active region masks with multi-parameter physical constraints. The approach enhances feature extraction from critical structures such as polarity inversion lines and sunspots via dynamic masks, while incorporating joint constraints on physical parameters including total Lorentz force and unsigned magnetic flux. Using SDO/SHARPS vector magnetogram data, the model learns spatiotemporal connections in field evolution: given a 12-hour input sequence, it forecasts the subsequent 12 hours of evolution across all components while maintaining both topological fidelity and parameter consistency. Quantitative evaluations show that the method achieves structural similarity of 0.85 for the radial component, with unsigned flux errors of 11.37% (95% CI: $\pm$0.06%).

2. Machine Learning Method with Dynamic Mask Enhancement and Physical Constraints

We propose a prediction method that integrates dynamic active region masks with physical constraints into an end-to-end deep learning architecture. This design enhances the model's ability to learn individual vector magnetic field components while maintaining physical consistency.

2.1. Attention-weighted Strategy for Active Region Cores

To comprehensively represent active region vector fields, the input data are constructed in Carrington heliographic projection as three channels corresponding to the radial (B_r), poloidal (B_p), and toroidal (B_t) components, plus one mask channel, yielding a four-channel input tensor.

To capture critical structures, we adopt a two-stage mask construction mechanism guided by the DSARD active region detection method (Chen et al. 2025). First, static masks are generated from the input 12-frame sequence based on the B_r channel: thresholding eliminates quiet-Sun noise, while DBSCAN-like clustering builds coherent mask regions. This combines threshold efficiency with morphological adaptability, preventing fragmentation and better delineating critical features such as shear bands near polarity inversion lines. During training, masks are dynamically predicted alongside magnetic field evolution using temporal gradients, producing $\mathbf{M}_t$ where core strong-field regions are 1 and background regions are 0. This enforces stronger attention on magnetically active regions, particularly near polarity inversion lines where magnetic shear and gradients are steepest.

2.2. Incorporation and Implementation of Physical Constraints

To provide physically meaningful inductive bias, we condition the network on five differentiable physics-inspired diagnostic maps computed directly from the vector magnetogram at each time step. These maps are injected as auxiliary channels to guide feature extraction, while the learning objective remains defined on the predicted magnetograms themselves.

2.2.0.1. Rationale. While convolutional neural networks can in principle learn to compute spatial derivatives and other low-level operations through training, explicitly providing these physics-inspired features offers several advantages: it reduces the learning burden on early network layers, allowing them to focus on higher-level spatiotemporal patterns; it ensures numerical stability and consistency by using fixed, differentiable implementations; and it provides an interpretable conditioning pathway that can be systematically ablated. These auxiliary channels serve as structured inductive biases that guide the learning process toward physically plausible solutions.

2.2.0.2. Physics-inspired diagnostic maps. The five auxiliary feature maps are:

- **Lorentz-force proxy (magnetic pressure):**

$$P_{\mathrm{mag}} = B_r^2 + B_p^2 + B_t^2, \tag{1}$$

  which serves as a computationally stable proxy for magnetic energy density. The full Lorentz force is $\mathbf{F}_L = \mathbf{J} \times \mathbf{B}$. Since computing the full three-dimensional current density requires additional derivatives and assumptions about vertical structure, we use magnetic pressure as a scalar proxy that highlights regions where magnetic forces are dynamically important (Fisher et al. 2012).
- **Unsigned flux:**

$$|\Phi| = |B_r|, \tag{2}$$

  which directly measures the unsigned radial flux density (Leka and Barnes 2003).

- **Shear-angle proxy:**

$$\alpha = \arccos\left(\frac{|B_r|}{|B_r| + \sqrt{B_p^2 + B_t^2} + \epsilon}\right), \quad \epsilon = 10^{-6}, \tag{3}$$

  which increases when transverse fields are relatively strong, highlighting non-potential structure.
- **Radial gradient magnitude:**

$$|\nabla B_r| = \sqrt{\left(\frac{\partial B_r}{\partial x}\right)^2 + \left(\frac{\partial B_r}{\partial y}\right)^2}, \tag{4}$$

  computed using Sobel operators to emphasize steep-gradient boundaries (Cui et al. 2007).
- **Horizontal gradient magnitude:**

$$|\nabla B_h| = \sqrt{|\nabla B_p|^2 + |\nabla B_t|^2}, \tag{5}$$

  where $|\nabla B_p|$ and $|\nabla B_t|$ are computed by Sobel operators.

These diagnostic maps are combined into a five-channel tensor:

$$\mathbf{P}_t = [P_{\text{mag}},\ |\Phi|,\ \alpha,\ |\nabla B_r|,\ |\nabla B_h|] \tag{6}$$

which is concatenated to the model input at each time step as auxiliary channels.

2.2.0.3. Implementation details. The physics-inspired features are computed using a dedicated `RealtimePhysicsLayer` module implemented in PyTorch. Spatial gradients are computed using Sobel operators applied via depthwise convolution with padding to preserve spatial dimensions. The shear-angle computation uses arctangent with epsilon regularization ($\epsilon = 10^{-6}$) to avoid division by zero in weak-field regions. All operations are fully differentiable and GPU-accelerated, adding negligible computational overhead (less than 2% of total forward-pass time). The auxiliary feature maps are computed online at each time step from the current magnetic field state, ensuring consistency throughout the autoregressive rollout.

2.3. Model Architecture and Training Mechanism

The model adapts the MotionRNN/PredRNN family (Wu et al. 2021) for solar applications. Each time step is represented by four channels (three magnetic components plus the active-region mask), and the auxiliary physics tensor $\mathbf{P}_t$ is computed online from (B_r, B_p, B_t) and concatenated as additional channels.

2.3.0.1. Network architecture. The backbone consists of stacked spatiotemporal LSTM cells organized in an encoder–motion–decoder architecture with cross-scale fusion:

- **Input processing:** The four-channel magnetogram is patchified (default patch size 1×1, yielding 4 channels per patch). The five physics-inspired diagnostic maps are computed from the magnetic components and concatenated, yielding 9 input channels per time step.
- **Spatiotemporal LSTM cells:** The model uses $L = 4$ stacked SpatioTemporalLSTMCell layers with hidden dimensions $[128, 64, 64, 64]$. Each cell maintains hidden state $h_t^{(l)}$, cell state $c_t^{(l)}$, and a shared memory state m_t that propagates across layers. The cells incorporate both spatial convolutions (filter size 5×5) and temporal recurrence to capture spatiotemporal dependencies.
- **Motion modeling:** Between layers $l = 1, 2, 3$, MotionGRU modules model local motion fields by aggregating information from 3×3 spatial neighborhoods. The motion features are encoded via $2\times$ downsampling convolutions and decoded via transposed convolutions that upsample back to the original resolution.
- **Cross-scale fusion:** At each intermediate layer, the upsampled motion features are combined with the corresponding spatiotemporal LSTM hidden state via a gated fusion mechanism:

$$o_t^{(l)} = \sigma\left(W_{\text{gate}}^{(l)}\left[h_{\text{dec}}^{(l)}; h_t^{(l)}\right]\right), \quad h_t^{(l)} \leftarrow o_t^{(l)} \odot h_{\text{dec}}^{(l)} + (1 - o_t^{(l)}) \odot h_t^{(l)}, \quad (7)$$

  where $h_{\text{dec}}^{(l)}$ is the upsampled motion feature, $[\cdot;\cdot]$ denotes channel concatenation, σ is the sigmoid function, and $\odot$ is element-wise multiplication.
- **Output generation:** The final hidden state $h_t^{(L)}$ is passed through a 1×1 convolution to produce the four-channel output. The physics diagnostic maps are recomputed from the predicted magnetic components at the next time step, ensuring consistency.

The total number of trainable parameters is approximately 15.2 million.

2.3.0.2. Loss functions. Let $\hat{\mathbf{B}}_t = (\hat{B}_r, \hat{B}_p, \hat{B}_t)$ be the predicted vector field and $\mathbf{B}_t$ the reference. We normalize each magnetic component to $[0, 1]$ using fixed clipping ranges: $B_r \in [-5000, 5000]$ G, $B_p \in [-1000, 1000]$ G, $B_t \in [-3000, 3000]$ G, and mask $\in [0, 255]$. We apply a static mask-defined core weighting $w(\mathbf{x}) = 1 + 4\,\mathbb{I}[M(\mathbf{x}) > 0]$ (core pixels receive $5\times$ higher weight). The reconstruction loss is

$$\mathcal{L}_{\text{rec}} = \left\langle w(\mathbf{x})\, \|\hat{\mathbf{B}}_t(\mathbf{x}) - \mathbf{B}_t(\mathbf{x})\|_2^2 \right\rangle, \quad (8)$$

where $\langle\cdot\rangle$ denotes averaging over space, time, and batch. We further apply explicit supervision on the predicted mask channel,

$$\mathcal{L}_{\text{mask}} = \langle(\hat{M}_t - M_t)^2\rangle, \quad (9)$$

and a temporal-gradient regularization on the magnetic components,

$$\mathcal{L}_{\text{temp}} = \left\langle w(\mathbf{x}) \, \|(\hat{\mathbf{B}}_{t+1} - \hat{\mathbf{B}}_t) - (\mathbf{B}_{t+1} - \mathbf{B}_t)\|_2^2 \right\rangle. \tag{10}$$

The total objective is

$$\mathcal{L}_{\text{total}} = \mathcal{L}_{\text{rec}} + \lambda_{\text{mask}} \, \mathcal{L}_{\text{mask}} + \lambda_{\text{temp}} \, \mathcal{L}_{\text{temp}}, \tag{11}$$

with weights $\lambda_{\text{mask}} = 1.0$ and $\lambda_{\text{temp}} = 0.1$.

2.3.0.3. Optimization and scheduled sampling. We train using the Adam optimizer (Kingma and Ba 2014) with an initial learning rate of 3×10^{-3}, $\beta_1 = 0.9$, $\beta_2 = 0.999$, and weight decay 10^{-5}. Gradient clipping is applied with a global norm threshold of 1.0. The learning rate is reduced by a factor of 0.5 if validation loss plateaus for 10 epochs (minimum learning rate 10^{-6}). Training is performed for 100 epochs with batch size 8 on NVIDIA A100 GPUs, taking approximately 72 hours.

We use scheduled sampling (Bengio et al. 2015) to transition from teacher forcing to free-running rollout: the probability of feeding ground-truth inputs decreases linearly from 1.0 with changing rate 2×10^{-5} per iteration until iteration 50,000, after which the model relies entirely on its own predictions during training. During inference, the model generates the full 12-hour forecast autoregressively without access to future ground-truth frames.

3. Data and Preprocessing

This study employs the Space-weather HMI Active Region Patch (SHARP) dataset obtained by the Helioseismic and Magnetic Imager (HMI) onboard the Solar Dynamics Observatory (SDO) (Bobra et al. 2014). We use the SHARP data in the Cylindrical Equal-Area (CEA) remapped vector-magnetogram format, whose sampling is 0.03 degrees per pixel in heliographic longitude/latitude (corresponding to approximately HMI's native ~ 0.5 arcsec sampling near disk center). Each SHARP active region is identified by a unique HARPNUM (e.g., HARPNUM 7890 corresponds to NOAA AR 12673), which tracks the region throughout its disk passage.

We excluded active regions with central longitudes greater than 60° and selected only qualified samples from SHARP. We performed unified numerical clipping of the original FITS files, restricting values to the range $[-5000, 5000]$ Gauss for B_r, $[-1000, 1000]$ Gauss for B_p, and $[-3000, 3000]$ Gauss for B_t. This truncation preserved the original field distribution while preventing outliers from disrupting model training.

To reduce the risk of train/test contamination and ensure temporal independence, we use a time-based split: the training set is constructed from SHARP data spanning 2021 (solar cycle 25 ascending phase), while the test set is constructed from SHARP data spanning 2022. This chronological split ensures that no active region appears in both training and test sets, and that the model is

evaluated on data from a temporally distinct period. The final dataset comprised 12,000 training sequences and 3,000 test sequences, with each sequence consisting of 24 consecutive frames at 1-hour cadence. The first 12 frames served as model inputs, while the subsequent 12 frames were used as prediction targets.

4. Experiments and Analysis

This study systematically evaluates the proposed method for solar vector magnetic field prediction (using frames 1–12 to predict frames 13–24). The analysis is organized along three complementary dimensions: **image-domain fidelity** (Sec. 4.2), **physical-parameter consistency** (Sec. 4.3), and **vector-field orientation accuracy** (Sec. 4.4), combining both single-sample visualizations and multi-sample statistics.

4.1. Evaluation setting, variants, and the persistence baseline

We evaluate the proposed method on sequences of SHARP vector magnetograms, using frames 1–12 as input and forecasting frames 13–24. We compare four trained models: the full model (**with_mask_and_phy**) and three ablations, including **without_phy** (no temporal-gradient regularization), **without_mask** (no mask weighting or mask supervision), and **without_mask_and_phy** (reconstruction loss only). As a non-learning reference, we include a simple persistence baseline (**gt12_baseline**), where every future frame is predicted by copying the last observed input frame (frame 12).

In all trained settings, the auxiliary diagnostic maps $\mathbf{P}_t$ are computed and concatenated as additional input channels. Ablations are implemented by removing mask weighting and/or mask supervision and the temporal-gradient regularization term $\mathcal{L}_{\text{temp}}$.

To ensure a consistent evaluation focus on the evolving strong-field core, masked-region metrics are computed using a fixed active-region mask constructed from the input window (frames 1–12). This yields a mask-defined strong-field core region that remains identical across all prediction lead times.

4.2. Image-domain prediction performance

We first evaluate the model's ability to reproduce the spatial distributions of the three vector magnetic field components (B_r, B_p, B_t) using image-quality metrics (SSIM, CC, RMSE). This subsection combines single-sample visualizations with multi-sample statistics.

4.2.1. Spatial structure and error distribution

The predicted fields reproduce the global polarity distribution of the ground-truth observations and match in strong-field core regions. This reflects the design intent of the dynamic mask, which emphasizes core-region weighting and enhances strong-field feature extraction. Even in quiet regions far from the core, the predicted fields capture weak-field topology well.

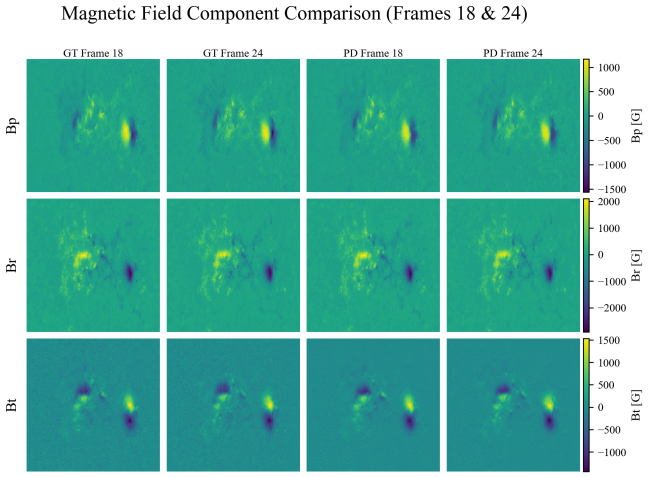


Figure 1. Comparison of ground-truth (Frame 24) and predicted results for all magnetic field components in active region HARPNUM 7890 (NOAA AR 12673). A diverging colormap (blue-white-red) centered at zero is used for each component with consistent color limits within each panel ($\pm$2500 G for B_r, $\pm$800 G for B_p, $\pm$1500 G for B_t). The prediction matches the reference in both polarity distribution and strong-field core structures. Difference maps (bottom row) show that the largest residuals occur near polarity inversion lines and in regions with steep gradients, while strong-field cores and quiet-Sun regions are well reproduced.

Closer inspection of each component reveals: for B_r, predictions and ground truth overlap in strong-field cores, with relative flux errors $< 5\%$ (95% CI: $\pm 0.3\%$); near polarity inversion lines, where gradients are steep, predictions show mild diffusion and relative errors increase by 15–20% (95% CI: $\pm 2.1\%$). For horizontal components B_p and B_t, similar patterns appear: strong shear bands (shear angle $> 60^\circ$) are well captured, but small magnetic islands at shear-band boundaries show more distorted topological details.

Pixel-level comparisons further quantify single-sample accuracy. Figure 1 shows predicted vs. reference fields for a sample active region. Figure 2 shows the corresponding pixel-value scatter plot, which aligns closely to the diagonal (Pearson CC = 0.92). Strong-field pixels cluster tightly along the diagonal, while weak-field pixels scatter more broadly, reflecting weaker signal-to-noise in transverse fields and more complex dynamics in weak-field zones.

The residual distribution of B_r (Fig. 3) further reveals spatial patterns: larger residuals tend to occur near polarity inversion lines and in strong-field umbra/penumbra regions, where gradients are steep and small spatial misalignments can produce large pixel-wise errors.

4.2.2. Comparison with baseline and ablated variants

To quantify overall image-level performance, we compute three complementary metrics: SSIM (structural similarity), Pearson correlation coefficient (CC), and RMSE (root-mean-square error). Unless otherwise stated, aggregate metrics are computed over the full 3,000-sequence test set.

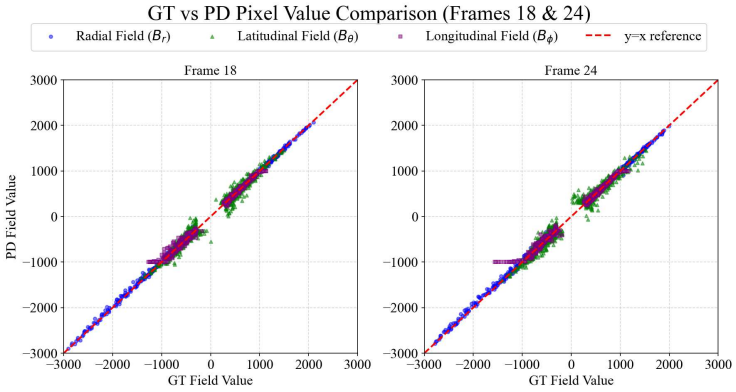


Figure 2. Pixel-value scatter comparison between predicted and reference fields. Points cluster along the diagonal (CC = 0.92), especially in strong-field regions.

A clear separation emerges between learned forecasting and the persistence baseline. For **gt12_baseline**, the global SSIM decreases from 0.833 (frame 13) to 0.765 (frame 24), while the masked-region SSIM decreases from 0.838 to 0.732. In parallel, the global MSE increases from 20.24 to 42.12, and the masked-region MSE increases from 25.09 to 91.96 (Fig. 4, 5). This rapid degradation indicates that active-region evolution over the 12-step horizon cannot be approximated as a static field. The persistence baseline also exhibits strong lead-time deterioration in the correlation of B_r: CC drops from 0.875 to 0.528 globally (and from 0.906 to 0.535 in the masked region) between frames 13 and 24 (Fig. 6).

In contrast, the full model (**with_mask_and_phy**) remains stable across the entire rollout: masked-region MSE stays within 3.13–3.86 from frame 13 to 24, and masked-region SSIM remains near 0.973. Averaged over frames 13–24, the full model achieves global/masked MSE of 2.64/3.55, whereas gt12_baseline reaches 33.06/64.26 (Table 1). This corresponds to an improvement factor of 12.5× globally and 18.1× inside the masked core, demonstrating that the proposed method captures genuine spatiotemporal evolution beyond persistence.

Comparing the ablated variants, **without_phy** (mask supervision only) consistently outperforms **without_mask** (physics constraint only) on both global and masked image quality (SSIM, MSE, CC), indicating that explicitly supervising the active-region mask is essential to maintain fidelity where evolution is most dynamic. Removing both ingredients (**without_mask_and_phy**) degrades performance compared to the full model, but still substantially exceeds persistence.

4.3. Physical-parameter prediction performance

Beyond image-domain fidelity, we evaluate the model's ability to preserve key physical diagnostics derived from the predicted magnetic field components. These diagnostics—unsigned flux, Lorentz-force proxy, shear angle, and field gradients—measure the physical consistency and interpretability of the predictions.

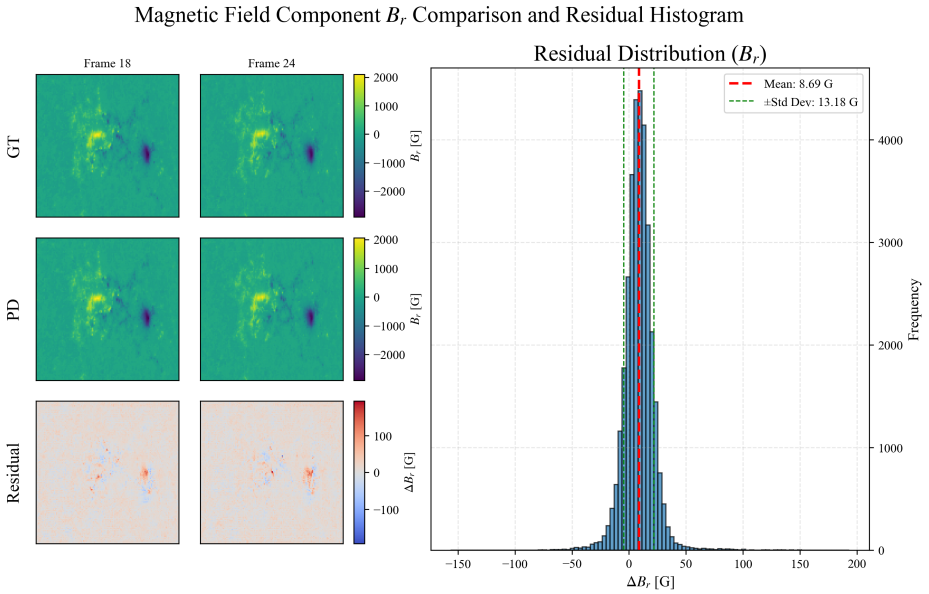


Figure 3. Residual distribution of the radial component B_r shown with a diverging colormap (blue-white-red) centered at zero. The color scale spans ±500 G. Larger residuals (red/blue regions) concentrate near the polarity inversion line (PIL) and in sunspot umbra/penumbra regions where field gradients are steepest. The spatial scale of the patch is approximately 200×200 pixels, corresponding to roughly 120×120 arcseconds (or ∼87×87 Mm on the solar surface at disk center). Most of the field of view shows residuals within ±100 G, indicating good overall agreement.

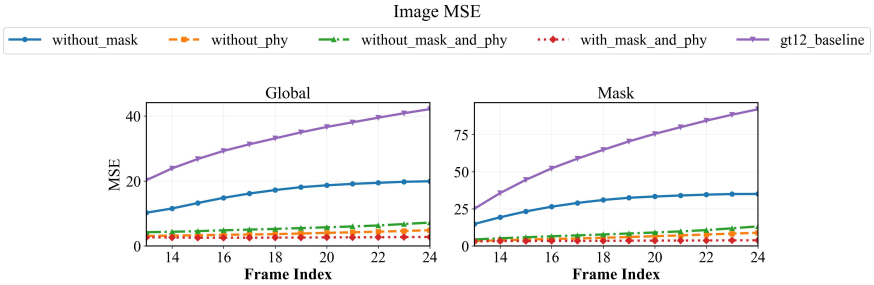


Figure 4. Framewise image MSE over the prediction window (frames 13–24). Left: global MSE. Right: MSE inside the active-region mask (mask-defined strong-field core).

4.3.1. Prediction errors for derived physical quantities

Table 2 summarizes the prediction errors for all five physics-inspired diagnostic quantities computed from the predicted magnetic field components. Unsigned flux prediction achieves a mean relative error of 11.37% (95% CI: ±0.06%, std = 2.46%). The Lorentz-force proxy shows moderately higher errors (mean 22.59%, 95% CI: ±0.09%, std = 3.50%), reflecting its dependence on all three magnetic field components. Shear angle errors average 0.0909 rad (95% CI: ±0.0005 rad, std = 0.0219 rad), while gradient-based quantities show absolute errors of 5.19

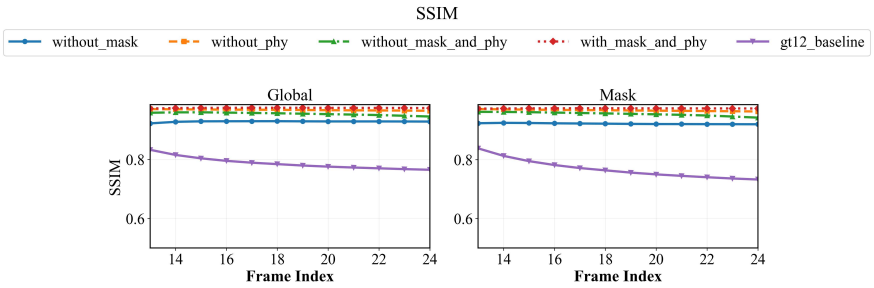


Figure 5. Framewise SSIM over the prediction window (frames 13–24). Left: global SSIM. Right: SSIM inside the active-region mask (mask-defined strong-field core).

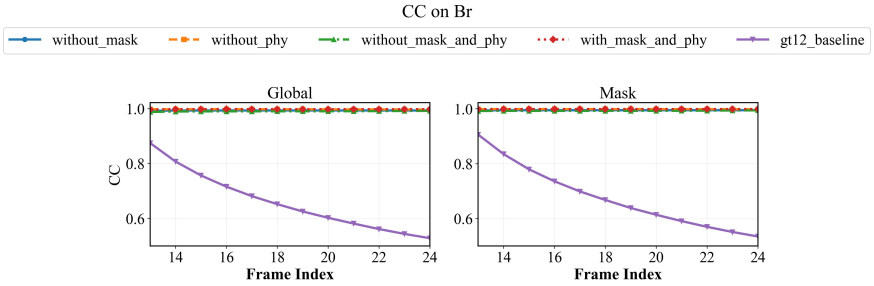


Figure 6. Framewise Pearson correlation coefficient (CC) for the B_r channel over frames 13–24. Left: global CC. Right: CC inside the active-region mask (mask-defined strong-field core).

G/pixel (95% CI: ±0.03 G/pixel) for $|\nabla B_r|$ and 18.65 G/pixel (95% CI: ±0.14 G/pixel) for horizontal gradients. These errors arise primarily from the lower prediction accuracy of horizontal components (B_p, B_t).

4.3.2. Temporal evolution of physical diagnostics

Figure 7 shows representative temporal evolution of key diagnostics across the 12-hour forecast horizon. Unsigned flux remains stable throughout the prediction window, with relative errors tightly bounded and only weak lead-time growth. In contrast, the Lorentz-force proxy exhibits moderately higher errors with a mean of 9.31% and systematic growth toward the horizon: early and mid-range frames (13–18) maintain errors close to 8.82–9.14%, but toward frame 24 the error reaches 9.98%. This trend suggests that while the force balance is captured on average, small discrepancies accumulate over time.

4.3.3. Comparison with baseline and ablated variants

The persistence baseline provides an instructive reference: its errors increase systematically with lead time, particularly inside the masked core (Figs. 9–12). For example, the masked-region absolute error of the Lorentz-force proxy rises from 3.93×10^4 (frame 13) to 1.03×10^5 (frame 24), whereas the full model remains near 9.49×10^3–9.99×10^3 over the same horizon. Similarly for unsigned flux,

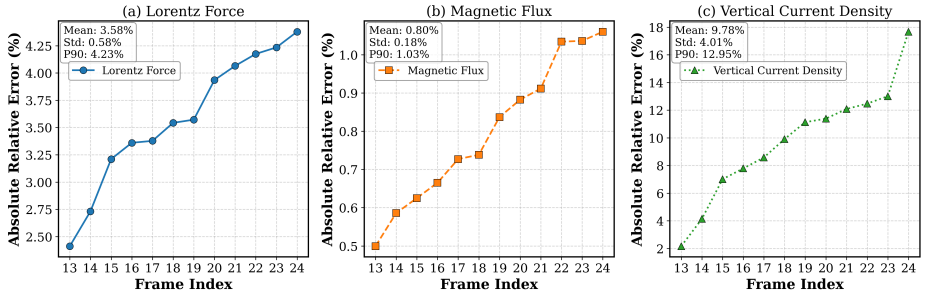


Figure 7. Representative temporal evolution of key physical diagnostics (unsigned flux, Lorentz-force proxy, shear angle, and field gradients) across the 12-hour forecast horizon. Unsigned flux shows tight error bounds throughout, while the Lorentz-force proxy exhibits modest systematic growth, and gradient-based measures reflect the challenges in horizontal-field prediction.

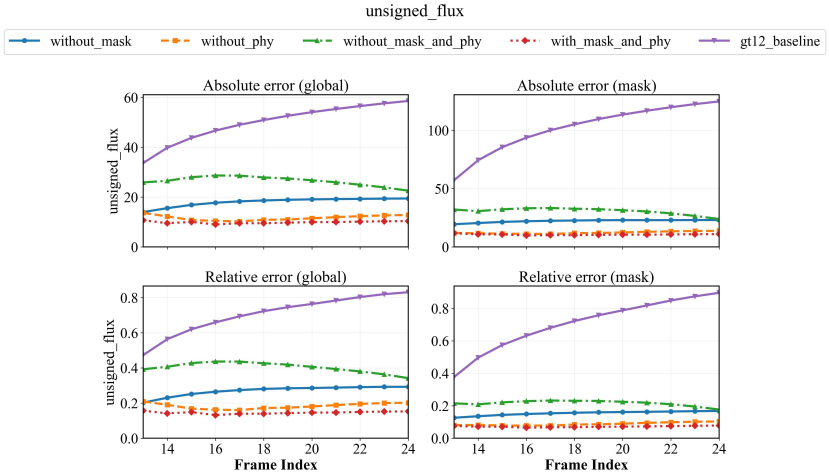


Figure 8. Framewise errors of unsigned flux over frames 13–24. The 2×2 layout shows absolute error (top row) and relative error (bottom row) for both global evaluation (left column) and masked-region evaluation (right column).

the masked-region absolute error increases from 57.1 to 124.7 for gt12_baseline, compared with 11.7 to 11.0 for the full model.

Across all diagnostics, the full model achieves the lowest masked-region errors on average (Table 4). Relative to gt12_baseline, the full model reduces the mean masked-region Lorentz-force-proxy error by $\sim 8.37\times$, unsigned-flux error by $\sim 9.61\times$, and $|\nabla B_r|$ error by $\sim 13.5\times$.

Comparing **without_phy** and **without_mask**, we observe that mask supervision alone already yields significantly improved physical consistency in the strong-field core, consistent with the image-domain advantage of **without_phy**. This supports the interpretation that accurately preserving the core-region field structure is a necessary prerequisite for accurate derived physical measures.

Figures 8–12 show the framewise evolution of all five physical diagnostics in a 2×2 layout (absolute/relative error × global/masked-region).

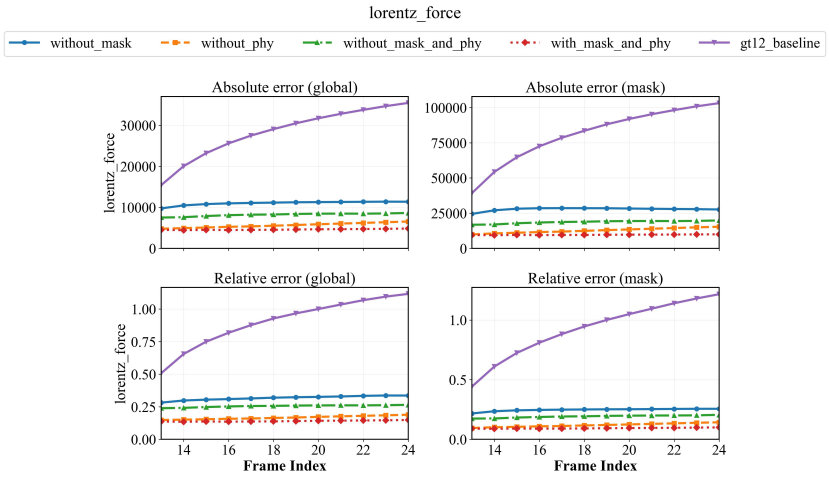


Figure 9. Framewise errors of the Lorentz-force proxy over frames 13–24. The 2×2 layout shows absolute error (top row) and relative error (bottom row) for both global evaluation (left column) and masked-region evaluation (right column).

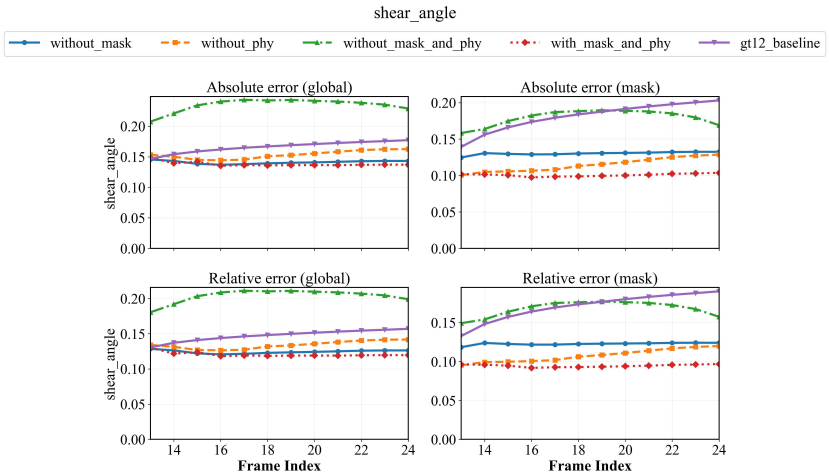


Figure 10. Framewise errors of shear angle over frames 13–24. The 2×2 layout shows absolute error (top row) and relative error (bottom row) for both global evaluation (left column) and masked-region evaluation (right column).

4.4. Vector-field orientation accuracy

Beyond component-wise and physical-parameter evaluations, we assess the model's ability to predict the orientation of the magnetic field vector, characterized by inclination angle θ and azimuth angle ϕ. These angles are derived from the predicted vector components via standard spherical–Cartesian transformations:

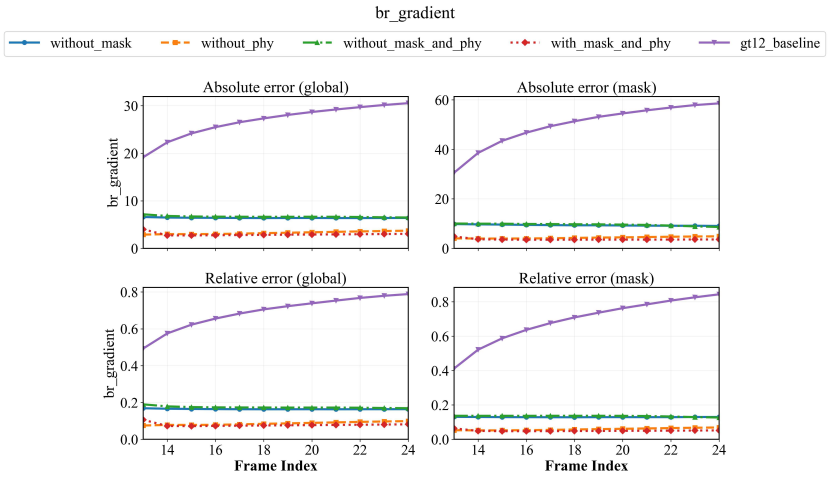


Figure 11. Framewise errors of $|\nabla B_r|$ over frames 13–24. The 2×2 layout shows absolute error (top row) and relative error (bottom row) for both global evaluation (left column) and masked-region evaluation (right column).

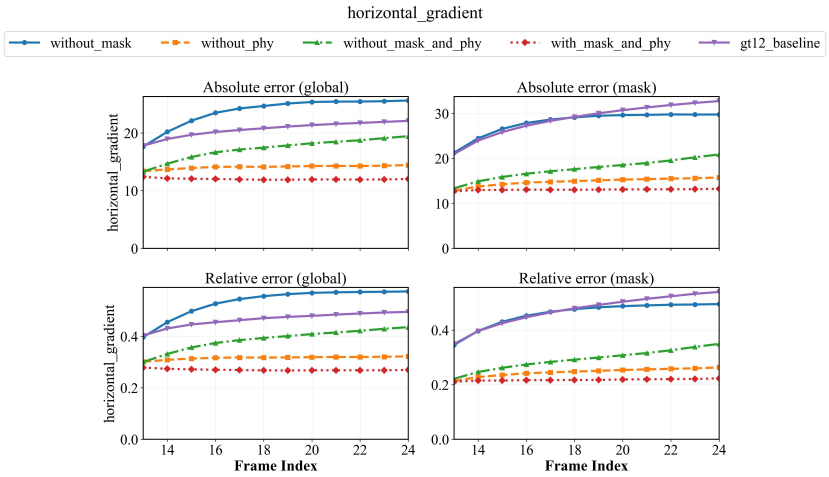


Figure 12. Framewise errors of horizontal gradient over frames 13–24. The 2×2 layout shows absolute error (top row) and relative error (bottom row) for both global evaluation (left column) and masked-region evaluation (right column).

4.4.0.1. Inclination θ:

$$\theta = \arctan\left(\frac{\sqrt{B_p^2 + B_t^2}}{|B_r| + \epsilon}\right), \quad \epsilon = 10^{-6}. \tag{12}$$

The inclination angle θ measures deviation from the radial direction and correlates with current helicity (Aulanier, Janvier, and Schmieder 2012).

4.4.0.2. Azimuth ϕ:

$$\phi = \text{atan2}(B_t, B_p), \tag{13}$$

where $\phi = 0°$ corresponds to positive B_p and $\phi = 90°$ to positive B_t. Azimuth errors directly affect reconnection-rate estimates (Su et al. 2013).

4.4.1. Spatial distribution of orientation errors

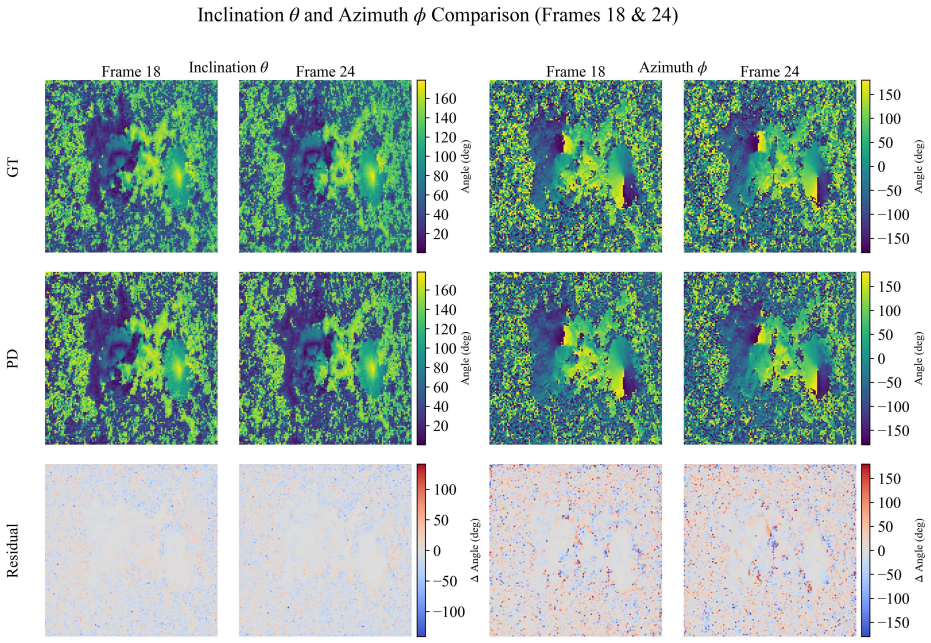


Figure 13. Spatial distribution of orientation residuals for inclination θ (left) and azimuth ϕ (right). Residuals are small in strong-field cores and systematically larger in weak-field regions and near polarity inversion lines.

Figure 13 shows the spatial distribution of orientation residuals. The pattern mirrors that observed for component residuals: errors are minimal in strong-field cores where the magnetic field is pressure-dominated and evolves more deterministically, while larger residuals appear near polarity inversion lines and in weak-field regions where convection-driven dynamics introduce greater stochasticity.

4.4.2. Statistical characteristics of orientation errors

Pixel-level scatter analysis (Fig. 14) reveals strong linear correlation (CC>0.9) between predicted and reference orientation angles in strong-field regions. The tight clustering along the diagonal indicates that systematic biases are minimal.

Statistical analysis of residual distributions (Fig. 15) shows that both inclination and azimuth errors approximate zero-mean Gaussian distributions. Azimuth

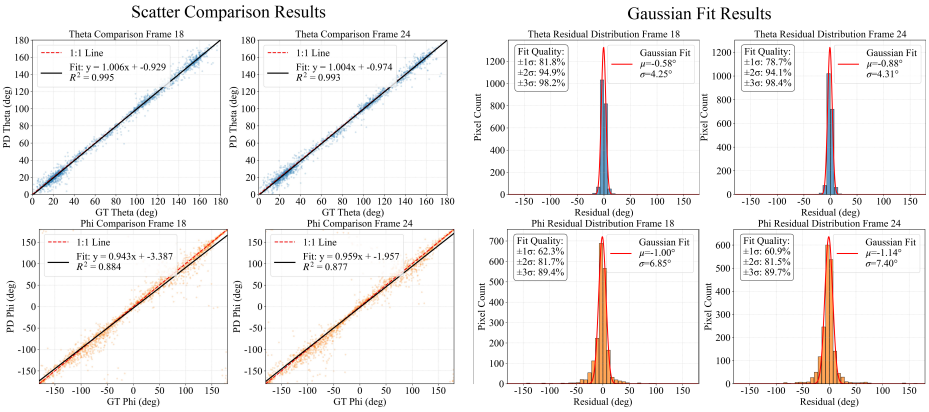


Figure 15. Gaussian fits of inclination θ and azimuth ϕ residual distributions. Both distributions approximate zero-mean normal distributions.

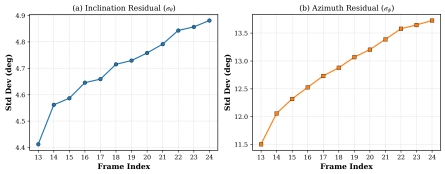


Figure 16. Residual standard deviations (σ) for inclination θ and azimuth ϕ across 1000 test samples in strong-field regions. The relatively stable σ values across samples indicate consistent orientation prediction accuracy.

residuals are typically contained within $\sim 14^\circ$, while inclination residuals remain within $\sim 5^\circ$, with no significant skewness.

Figure 16 shows the distribution of residual standard deviations across 1000 test samples. The relatively stable σ values indicate consistent orientation prediction accuracy across the test set, with occasional outliers corresponding to highly sheared or rapidly evolving active regions.

5. Conclusions and Discussion

We developed a deep-learning framework that integrates dynamic active-region masks with multi-parameter physical constraints to predict short-term (12-hour) evolution of all vector components of solar active-region magnetic fields. The principal conclusions are:

1. **Three-channel vector representation and dynamic masks improve structural fidelity.** By jointly encoding B_r, B_p, and B_t in Carrington projection and incorporating adaptive masks emphasizing polarity inversion lines and strong-gradient cores, the framework achieves high performance in critical regions. Radial-field predictions achieve SSIM > 0.84 (95% CI: 0.843–0.851), CC > 0.959 (95% CI: 0.959–0.961), and RMSE $\approx$57–58 G (95% CI: 57.1–57.9

G). Horizontal components perform less accurately but still capture meaningful structure (B_p SSIM ∼0.73–0.76, B_t SSIM ∼0.58–0.66), particularly in strong-shear environments.
2. **Physics-feature conditioning and temporal regularization improve physical consistency.** By conditioning on auxiliary physics-inspired diagnostic maps (unsigned flux, Lorentz-force proxy, shear proxy, and gradient measures) and applying temporal-gradient regularization, the model better preserves derived diagnostics over the forecasting window. Unsigned flux errors are maintained at 11.37% (95% CI: ±0.06%, std = 2.46%). Lorentz-force-proxy errors average 22.59% (95% CI: ±0.09%, std = 3.50%) with modest increase toward the horizon.
3. **Limitations remain in current-related derivatives and horizontal-field stability.** Higher-order derivatives of transverse fields are intrinsically sensitive to small-scale errors. Prediction quality also declines in B_t, where SSIM drops by ∼12% across the 12-hour window. These shortcomings arise from both weaker observational signal-to-noise in transverse fields and the difficulty of representing coupled wave and turbulence processes within a purely data-driven framework.

5.1. Ablation studies and baseline comparisons

The inclusion of a persistence baseline and ablation variants clarifies the contributions of each component. The persistence baseline exhibits rapid degradation over the 12-step rollout, with global/masked MSE reaching 33.06/64.26, while the full model achieves 2.64/3.55, corresponding to 12.5× and 18.1× lower error, respectively. This confirms that the model learns genuine spatiotemporal evolution beyond simple temporal smoothness.

The ablation comparison indicates that emphasizing the active-region mask is a central factor for stability: **without_phy** (mask supervision only) consistently outperforms **without_mask** (physics constraint only), suggesting that core-region structural accuracy is a prerequisite for reliable physical measures. Combining mask supervision and physics constraints yields the best overall performance and the most stable lead-time behavior.

The superior performance of mask supervision over physics conditioning alone suggests that accurate spatial localization of evolving structures is more critical than explicit physics-inspired feature channels for this forecasting task. However, the full model achieves the best performance, indicating that physics conditioning provides additional regularization that improves generalization, particularly for derived physical diagnostics.

5.2. Limitations and future directions

While the model achieves strong performance on average, several limitations warrant discussion:

- **Weak-field regions and small-scale structures.** The mask-weighted loss de-emphasizes weak-field regions, leading to larger relative errors in

quiet-Sun areas and small isolated flux concentrations. This is a deliberate design choice prioritizing flare-relevant strong-field cores, but limits applicability to studies requiring accurate weak-field evolution.

- **Horizontal-field degradation.** The B_t component shows substantial degradation over the 12-hour horizon (SSIM drop of 11.8%), particularly in regions with strong shear ($> 80°$). This reflects both the lower signal-to-noise ratio of transverse-field observations and the more complex physics governing horizontal-field evolution. The persistent gap in B_t and B_p accuracy suggests the need for specialized modules, such as embedding simplified transport or Alfvén-wave propagation models.
- **Extreme events and rapid evolution.** The model has not been systematically evaluated on X-class flares, rapid flux emergence events, or CME-productive active regions. Preliminary inspection suggests that the model tends to underpredict rapid changes associated with flare-driven magnetic field restructuring, likely due to the rarity of such events in the training data and the smoothness bias of the MSE loss.
- **Temporal horizon and error accumulation.** Beyond 12 hours, errors accumulate rapidly, particularly in horizontal components. The high frame-to-frame correlation in the data (typical correlation > 0.95 for B_r) indicates that the model learns coherent transport patterns of existing flux concentrations rather than predicting genuinely novel emergence. Extending the forecasting horizon would likely require incorporating additional physics-based constraints or hybrid approaches.
- **Generalization across solar cycle phases.** Although the training data span 2021 and the test data span 2022, both periods sample relatively active phases with morphologically similar active region evolution patterns. Performance on solar minimum conditions with predominantly weak, dispersed fields remains untested.
- **Deterministic vs. probabilistic forecasting.** The current model is deterministic and produces a single forecast trajectory. Given the inherent stochasticity of solar magnetic field evolution, particularly in weak-field regions and over longer horizons, future work should explore probabilistic approaches that can quantify forecast uncertainty.

5.3. Broader implications and future work

This work demonstrates the promise of hybrid data–physics approaches for vector magnetic field prediction. Several directions for future research emerge:

- *Enhanced mask design.* Multi-scale or attention-based masks that adaptively weight both core and peripheral features could improve sensitivity to weak but topologically important structures in surrounding quiet-Sun regions.
- *Expanded physical constraints.* Current constraints enforce unsigned flux conservation and Lorentz-force balance, but omit other potentially valuable diagnostics such as relative helicity, free magnetic energy, or electric-field consistency. Extending the physics-informed loss to include these terms may improve both predictive accuracy and physical interpretability.

- *Operational deployment.* For real-time space-weather forecasting, latency and computational efficiency are critical. Streamlining inference, integrating with near-real-time SDO/HMI pipelines, and developing uncertainty-quantification schemes will be essential for practical adoption.
- *Hybrid physics-data approaches.* Continued progress will require deeper coupling with physics-based models and validation in operational contexts, potentially through hybrid frameworks that combine data-driven predictions with MHD simulations.

The complete implementation of the deep learning model presented in this paper, developed in PyTorch, is publicly available at https://github.com/Qingcaiyurouzhou/magnetic_predict. The repository includes the model architecture, the dynamic mask generation algorithm, the differentiable physics-feature computation layer, training/testing scripts, and example data preprocessing pipelines.

Animated figures showing temporal evolution sequences are provided as supplementary materials. Animation 1 shows the 24-hour evolution sequence (frames 1–24) for a representative active region, with the first 12 frames (input) shown in the left column and the subsequent 12 frames (prediction vs. ground truth) shown in the middle and right columns. The animation plays at 2 frames per second and uses a diverging colormap to highlight polarity changes. Additional multi-step rollout examples for diverse active regions are provided in the supplementary materials.

Declarations

Funding

This work was supported by the Joint Fund of the National Natural Science Foundation of China (Grant No. U2031140).

Competing Interests

The authors have no relevant financial or non-financial interests to disclose.

Data Availability

The SDO/HMI SHARP data used in this study are publicly available from the Joint Science Operations Center (JSOC) at Stanford University (https://jsoc1.stanford.edu/). The trained model weights and code are available at https://github.com/Qingcaiyurouzhou/magnetic_predict.

Author Contributions

All authors contributed to the study conception and design. Material preparation, data collection and analysis were performed by Yuqing Zhou, Hui Liu, and Zhenyu Jin. The first draft of the manuscript was written by Yuqing Zhou and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Acknowledgments We acknowledge the use of data from the Solar Dynamics Observatory (SDO), a mission of NASA's Living With a Star Program. The HMI instrument and SHARP data products are provided by the Joint Science Operations Center (JSOC) at Stanford University. We thank the anonymous reviewers for their constructive comments that significantly improved this manuscript. Computational resources were provided by the High Performance Computing Center of Yunnan Observatories, CAS.

Appendix

A. Supplementary Materials

This appendix provides additional quantitative details to support the main text analysis. Tables 1 and 2 present horizon-averaged metrics comparing model variants and the persistence baseline. Tables 3 and 4 present detailed per-frame metrics and physical diagnostic errors for each model variant. All tables are displayed on rotated pages for optimal readability.

Table 1.: Horizon-averaged image-domain metrics over frames 13–24, including the persistence baseline (gt12_baseline)

Metric	**w/o mask**	**w/o phy**	**w/o both**	**full model**	**gt12_baseline**
img_mse (global)	16.5222	3.8560	5.4954	2.6445	33.0593
img_mse (mask-defined core)	29.0076	5.9737	8.3470	3.5519	64.2613
SSIM (global)	0.92812	0.96749	0.95431	0.97422	0.78739
SSIM (mask-defined core)	0.92187	0.96720	0.95484	0.97380	0.76842
CC(B_r) (global)	0.99309	0.99745	0.99020	0.99780	0.66111
CC(B_r) (mask-defined core)	0.99458	0.99869	0.99269	0.99884	0.67678

Table 2.: Prediction errors for all five physics-inspired diagnostic parameters (masked-region averages over frames 13–24)

Parameter	**Mean abs. error**	**Mean rel. error (%)**	**Std. dev.**	**90% quantile**
Unsigned flux ($\|B_r\|$)	11.37 G	11.37%	2.46%	15.21%
Lorentz force proxy (B^2)	9.67×10^3	22.59%	3.50%	28.44%
Shear angle (α)	0.0909 rad	8.12%	0.0219 rad	0.1338 rad
$\|\nabla B_r\|$	5.19 G/pixel	—	2.87 G/pixel	9.81 G/pixel
Horizontal gradient ($\|\nabla B_h\|$)	18.65 G/pixel	—	10.92 G/pixel	35.03 G/pixel

Table 3.: Per-frame image-quality metrics for vector components over the prediction horizon (frames 13–24)

Frame	B_r SSIM	B_r CC	B_r RMSE	Frame	B_p SSIM	B_p CC	B_p RMSE
13	0.849	0.961	57.2	13	0.755	0.927	48.7
14	0.850	0.961	57.1	14	0.759	0.928	48.9
15	0.851	0.961	57.4	15	0.761	0.928	49.4
16	0.851	0.961	57.4	16	0.761	0.926	49.9
17	0.851	0.961	57.3	17	0.759	0.926	50.2
18	0.851	0.960	57.3	18	0.758	0.926	50.2
19	0.851	0.960	57.4	19	0.757	0.925	50.5
20	0.850	0.960	57.6	20	0.753	0.924	50.8
21	0.848	0.959	57.9	21	0.748	0.923	51.2
22	0.847	0.959	57.9	22	0.742	0.923	51.2
23	0.846	0.959	57.8	23	0.738	0.922	51.3
24	0.844	0.959	57.8	24	0.727	0.920	51.1

Frame	B_t SSIM	B_t CC	B_t RMSE
13	0.661	0.882	58.4
14	0.655	0.880	59.8
15	0.648	0.878	61.2
16	0.641	0.876	62.5
17	0.634	0.874	63.8
18	0.627	0.872	65.0
19	0.620	0.870	66.2
20	0.613	0.868	67.3
21	0.606	0.866	68.4
22	0.599	0.864	69.4
23	0.592	0.862	69.8
24	0.583	0.870	70.2

Table 4.: Horizon-averaged masked-region absolute errors of derived physical diagnostics over frames 13–24

Diagnostic Parameter	w/o mask	w/o phy	w/o both	full model	gt12_baseline
Lorentz force	2.78×10^4	1.27×10^4	1.87×10^4	9.67×10^3	8.10×10^4
Unsigned flux	22.18	12.27	30.64	10.60	101.92
Shear angle	0.1304	0.1146	0.1796	0.1007	0.2456
$\lvert\nabla B_r\rvert$	9.376	4.345	9.582	3.672	49.62
Horizontal gradient	28.01	14.83	17.65	13.06	67.34

References

Aulanier, G., Janvier, M., Schmieder, B.: 2012, The standard flare model in three dimensions. I. Strong-to-weak shear transition in post-flare loops. *Astronomy and Astrophysics* **543**, A110. DOI.

Bai, X.Y., et al.: 2020, Predicting the Evolution of Photospheric Magnetic Field in Solar Active Regions Using Deep Learning. *Solar Physics* **295**, 82.

Baty, H., Vigon, V.: 2024, Modeling solar coronal magnetic fields with physics-informed neural networks. *Monthly Notices of the Royal Astronomical Society* **527**, 2575. DOI.

Bengio, S., Vinyals, O., Jaitly, N., Shazeer, N.: 2015, Scheduled sampling for sequence prediction with recurrent neural networks. In: *Advances in Neural Information Processing Systems* **28**, 1171.

Bobra, M.G., Sun, X., Hoeksema, J.T., Turmon, M., Liu, Y., Hayashi, K., Barnes, G., Leka, K.D.: 2014, The Helioseismic and Magnetic Imager (HMI) Vector Magnetic Field Pipeline: SHARPs - Space-Weather HMI Active Region Patches. *Sol. Phys.* **289**, 3549. DOI. ADS.

Chen, R., Lu, W., Hao, Q., Meng, Y., Chen, P., Shi, C.: 2025, Statistical Analyses of Solar Active Region in SDO/HMI Magnetograms detected by Unsupervised Machine Learning Method DSARD. *arXiv preprint arXiv:2502.17990*. 22 pages, 15 figures, submitted to ApJS. DOI. https://arxiv.org/abs/2502.17990.

Covas, E.: 2020, Transfer learning in spatial–temporal forecasting of the solar magnetic field. *Astronomische Nachrichten* **341**. First published: 30 March 2020. DOI. https://doi.org/10.1002/asna.202013690.

Covas, E., Peixinho, N., Fernandes, J.: 2019, Neural Network Forecast of the Sunspot Butterfly Diagram. *Solar Physics* **294**, 1. DOI. https://doi.org/10.1007/s11207-019-1412-z.

Cui, Y., Li, R., Wang, H., He, H.: 2007, Correlation Between Solar Flare Productivity and Photospheric Magnetic Field Properties. I. Maximum Horizontal Gradient, Length of Neutral Line, Number of Singular Points. *Solar Physics* **242**, 1.

Fan, Y.: 2015, *MHD Equilibria and Triggers for Prominence Eruption*, *Astrophysics and Space Science Library* **415**, Springer, Cham. ISBN 978-3-319-10415-7. DOI. https://doi.org/10.1007/978-3-319-10416-4_12.

Fisher, G.H., Bercik, D.J., Welsch, B.T., Hudson, H.S.: 2012, Global Forces in Eruptive Solar Flares: The Lorentz Force Acting on the Solar Interior. *Solar Physics* **277**, 59.

Jeong, H.-J., Jeon, M., Kim, D., et al.: 2025a, Prediction of the Next Solar Rotation Synoptic Maps Using an Artificial Intelligence–based Surface Flux Transport Model. *Astrophysical Journal Supplement Series* **278**, 5.

Jeong, H.-J., Poedts, S., Wang, H., et al.: 2025b, Prediction of Time-evolving Radial Magnetic Fields on the Solar Surface Using Deep Learning. *Astrophysical Journal* **989**, L31.

Jiang, J., Cameron, R.H., Schüssler, M.: 2014, Effects of the scatter in sunspot group tilt angles on the large-scale magnetic field at the solar surface. *Astrophysical Journal* **791**, 5.

Kilpua, E.K.J., Lugaz, N., Mays, M.L., Temmer, M.: 2019, Forecasting the severity of solar eruptions for space weather. *Space Weather* **17**, 564. DOI.

Kingma, D.P., Ba, J.: 2014, Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*.

Leka, K.D., Barnes, G.: 2003, Photospheric Magnetic Field Properties of Flaring versus Flare-Quiet Active Regions. II. Discriminant Analysis. *Astrophysical Journal* **595**, 1296.

Miesch, M.S., Teweldebirhan, K.: 2016, A three-dimensional Babcock–Leighton solar dynamo model: Initial results with axisymmetric flows. *Advances in Space Research* **58**, 1571. Solar Dynamo Frontiers. DOI. https://www.sciencedirect.com/science/article/pii/S0273117716001198.

Ramunno, F.P., Jeong, H.-J., Hackstein, S., Csillaghy, A., Voloshynovskiy, S., Georgoulis, M.K.: 2024, Magnetogram-to-Magnetogram: Generative Forecasting of Solar Evolution. *arXiv preprint arXiv:2407.11659*. Conference paper accepted for oral presentation at the ESA SPAICE CONFERENCE, 17-19 September 2024. DOI. https://arxiv.org/abs/2407.11659.

Su, Y., Veronig, A., Holman, G., et al.: 2013, Imaging coronal magnetic-field reconnection in a solar flare. *Nature Physics* **9**, 489. DOI.

Upton, L., Hathaway, D.H.: 2014, Predicting the Sun's Polar Magnetic Fields with a Surface Flux Transport Model. *Solar Physics* **289**, 2329.

Wang, Z., Bovik, A.C., Sheikh, H.R., Simoncelli, E.P.: 2004, Image quality assessment: from error visibility to structural similarity. *IEEE Transactions on Image Processing* **13**, 600. DOI.

Wu, H., Yao, Z., Wang, J., Long, M.: 2021, MotionRNN: A Flexible Model for Video Prediction with Spacetime-Varying Motions. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*.

Yang, G., Xu, Y., Cao, W.D., et al.: 2004, Photospheric Shear Flows along the Magnetic Neutral Line of Active Region 10486 prior to an X10 Flare. *Astrophysical Journal* **617**, L151.