Rajesh Kumar Mishra, 2026, 14:3
ISSN (Online): 2348-4098
ISSN (Print): 2395-4752

International Journal of Science, Engineering and Technology
An Open Access Journal

Real-Time Space Weather Prediction And Geomagnetic Storm Forecasting

Rajesh Kumar Mishra[1] Divyansh Mishra[3], Rekha Agarwal[3]

[1]ICFRE-Tropical Forest Research Institute (Ministry of Environment, Forests & Climate Change, Govt. of India) P.O. RFRC, Mandla Road, Jabalpur, MP-482021, India
[2]Department of Artificial Intelligence and Data Science Jabalpur Engineering College, Jabalpur (MP)
[3]Government Science College Jabalpur, MP, India- 482 001

Abstract: The coupling between solar activity, interplanetary magnetic field (IMF) dynamics, and Earth's magnetosphere-ionosphere system gives rise to geomagnetic storms and substorms that represent a primary space weather hazard for technological infrastructure, including power grids, navigation systems, and low-Earth orbit (LEO) satellites. Traditional physics-based models rooted in magnetohydrodynamics (MHD) and empirical index formulations—such as the Dst and Kp indices—are physically interpretable but computationally expensive and insufficient for real-time forecasting at the 1–6 hour horizons required for operational applications. In this work, we develop and rigorously evaluate a hybrid Artificial Intelligence (AI) framework that fuses deep learning with physics-informed constraints to deliver accurate, uncertainty-quantified predictions of geomagnetic activity across multiple temporal horizons. Our ensemble approach achieves a root-mean-square error (RMSE) of 9.4 nT at 1-hour lead time and 18.7 nT at 6-hour lead time, representing improvements of 34% and 41%, respectively, over the best-performing operational empirical model (Temerin-Li). The full AI pipeline produces a complete geomagnetic forecast cycle in under 2 seconds on commodity GPU hardware, meeting real-time operational constraints.

Keywords: Artificial intelligence • deep learning • geomagnetic storm • Dst index • space weather forecasting • LSTM • Transformer • physics-informed neural networks • solar wind • magnetosphere-ionosphere coupling • Kp index • CME prediction • graph neural networks • uncertainty quantification • INTERMAGNET

II. INTRODUCTION

Space weather—the variable conditions in near-Earth space driven by solar activity—poses significant and growing risks to modern technological systems. Geomagnetic storms, which arise from the interaction of solar wind disturbances with Earth's magnetosphere, can induce geomagnetically induced currents (GICs) in power transmission networks, disrupt high-frequency radio communications, degrade GPS positioning accuracy, increase orbital drag on LEO satellites, and expose astronauts and high-latitude aviation passengers to elevated radiation doses. The sociological and economic consequences of severe space weather events are potentially catastrophic; the 1989 Quebec blackout and the 2003 Halloween superstorm caused billions of dollars in infrastructure damage across North America and Europe.

The canonical measure of geomagnetic storm intensity is the disturbance storm time (Dst) index, derived from hourly averages of the horizontal magnetic field component at a network of low-latitude ground-based observatories. The Dst index is a proxy for the energy content of the ring current—a torus of energetic ions encircling Earth at distances of 3–6 Earth radii—which is energised during the main phase of a storm and decays during the recovery phase. Accurate prediction of Dst at lead times of 1–6 hours would enable meaningful advance warning for power grid operators, satellite controllers, and aviation authorities.

The physics of geomagnetic storm initiation and evolution is, in principle, well understood. The key driver is the southward component ($B_z < 0$) of the IMF: prolonged southward IMF enables efficient dayside magnetic reconnection, loading energy into the

magnetotail and ultimately driving the ring current injection. Empirical models such as the Burton-McPherron-Russell (BMR) driven-dissipation equation and its successors capture this physics in a compact ordinary differential equation (ODE) form. However, these models rely on simplified parameterisations and fail to capture the non-linear, history-dependent character of storm dynamics, particularly during intense events. Physics-based numerical MHD simulations (e.g., BATS-R-US, GUMICS-4) resolve the full spatial structure of the magnetosphere but require hours to days of wall-clock time, precluding operational use.

The advent of machine learning (ML) has opened new possibilities for space weather prediction. Early neural network approaches demonstrated that data-driven models could match or outperform empirical formulae for Dst forecasting. The emergence of Long Short-Term Memory (LSTM) networks, Transformer architectures, and Graph Neural Networks (GNNs) has since enabled substantially more sophisticated models capable of exploiting long-range temporal dependencies, multi-instrument data fusion, and spatially distributed observational networks. Concurrently, Physics-Informed Neural Networks (PINNs) have provided a framework for embedding physical laws directly into the neural network training process, yielding models that respect governing equations even in data-sparse regimes.

This paper presents an integrated AI framework for geomagnetic storm forecasting that combines the complementary strengths of these approaches. We describe our model architectures, training methodology, validation strategy, and operational deployment pathway, and we demonstrate state-of-the-art performance on a challenging benchmark of intense historical storms.

III. SCIENTIFIC BACKGROUND

A. The Solar Wind–Magnetosphere–Ionosphere Coupling Chain

The Sun continuously emits a supersonic plasma flow—the solar wind—carrying with it the frozen-in IMF. At Earth, the solar wind compresses the dayside magnetosphere to the magnetopause (typically at 10–12 Earth radii in the subsolar direction) and stretches the nightside into a long magnetotail. The rate of energy transfer from the solar wind to the magnetosphere is controlled primarily by the reconnection voltage, which depends sensitively on the orientation and magnitude of the IMF Bz component.

During southward IMF (Bz < 0), dayside reconnection allows solar wind energy and plasma to enter the magnetosphere. This energy is partially stored in the magnetotail lobe fields and subsequently released during substorms—explosive reconnection events in the tail that inject energetic particles into the inner magnetosphere, energising the ring current. The Dst index evolves according to the BMR equation:

$$dDst^*/dt = Q(t) - Dst^*(t) / \tau$$

where Dst* is the pressure-corrected Dst index, Q(t) is the solar wind injection function (proportional to the coupling function of the IMF and solar wind speed), and $\tau \approx 7.7$ hours is the ring current decay time constant. While elegantly simple, this formulation neglects multi-step energy transfer processes, magnetospheric preconditioning, and the non-linear saturation of ring current injection during extreme events.

B. Operational Forecasting Context

The World Data Centre (WDC) for Geomagnetism, Kyoto, produces the definitive Dst index with a latency of approximately 1 month. For operational use, provisional Dst is available within hours, while nowcast (real-time) estimates are provided by NOAA SWPC and other services. NOAA SWPC issues geomagnetic storm watches, warnings, and alerts based on solar wind observations from the ACE and DSCOVR spacecraft stationed at the L1 Lagrange point, approximately 1.5 million km sunward of Earth. The L1-to-Earth solar wind transit time of approximately 15–60 minutes provides a hard physical upper limit on warning time achievable from in-situ solar wind monitoring alone.

To extend the forecasting horizon beyond the L1 transit time, predictive models must incorporate upstream solar observations (coronagraphs, EUV imagers) capable of detecting and characterising solar transients—coronal mass ejections (CMEs) and solar energetic particle events—days before their arrival. This motivates the multi-instrument, multi-model approach adopted in this work.

IV. DATA SOURCES AND PREPROCESSING

A. Solar Wind And Imf Data

The primary solar wind dataset is sourced from the OMNI high-resolution database, which provides 1-minute and 5-minute averaged solar wind parameters—including IMF magnitude and components (Bx, By, Bz), proton density, bulk flow speed, and dynamic pressure—propagated to Earth's bow shock nose. OMNI aggregates observations from multiple L1 monitors (ACE, WIND, DSCOVR, Geotail) using time-shifting algorithms, providing a near-continuous record spanning 1963 to the present.

For model training, we use the period 1995–2023 (covering Solar Cycles 23, 24, and the ascending phase of Cycle 25), which offers the highest data quality and temporal resolution. Missing data periods exceeding 3 hours are excluded from training. Data gaps shorter than 3 hours are filled using cubic spline interpolation, consistent with established space weather ML practice. All parameters are normalised using robust scaling (median and interquartile range) to mitigate the influence of extreme-event outliers on the normalisation statistics.

B. Geomagnetic Index Data

The target variable for supervised learning is the definitive hourly Dst index from WDC Kyoto (1995–2020) supplemented by provisional Dst (2020–2023). The Kp index—a 3-hourly quasi-logarithmic global activity index derived from a global network of mid-latitude magnetometers—is used as an auxiliary target and diagnostic. The auroral electrojet indices AE, AL, and AU, which measure high-latitude ionospheric current activity, are incorporated as predictor features in the Bi-LSTM model. All index data are obtained from the WDC for Geomagnetism, Kyoto, and the GFZ German Research Centre for Geosciences.

C. Solar And Coronal Observations

For the CME transit time and geo-effectiveness prediction component, we use: (i) GOES X-ray flux data (0.1–0.8 nm channels) from NOAA, providing real-time solar flare monitoring; (ii) SDO/AIA extreme ultraviolet (EUV) imagery at 131 Å and 171 Å wavelengths, sensitive to flare and coronal loop emission; and (iii) STEREO-A COR2 coronagraph imagery, enabling CME detection and 3D kinematic characterisation through single-spacecraft techniques when quadrature geometry is unavailable.

D. Ground Magnetometer Network

Spatial magnetic disturbance modelling exploits the INTERMAGNET global network, which provides definitive 1-minute digital magnetic observatory data from approximately 150 geographically distributed stations spanning geomagnetic latitudes from the geographic poles to the magnetic equator. For GNN training, we select 87 stations with data completeness exceeding 90% over the 1995–2023 training period, spanning geomagnetic latitudes from 80°S to 82°N.

V. METHODOLOGY

A. Bidirectional Lstm For Dst Forecasting

The first model component is a Bidirectional Long Short-Term Memory (Bi-LSTM) network designed for multi-horizon Dst index prediction. LSTM networks are a class of recurrent neural network (RNN) specifically engineered to learn long-range temporal dependencies through gated memory cells. The bidirectional extension processes input sequences both forward and backward in time, allowing the model to exploit future context within the historical input window when generating forecasts.

The input to the Bi-LSTM is a 24-hour sliding window of 1-minute OMNI solar wind and IMF data, downsampled

to 5-minute resolution (yielding 288 time steps per sample) and augmented with lagged Dst, AE, and Kp values. The architecture consists of three stacked Bi-LSTM layers (256, 128, and 64 hidden units, respectively) with dropout regularisation ($\alpha = 0.20$) applied between layers, followed by a dense output layer with three neurons producing Dst forecasts at lead times of 1, 3, and 6 hours simultaneously. The model is trained using the Adam optimiser with an initial learning rate of 3×10^{-4}, cosine annealing scheduling, and a combined loss function weighting mean squared error (MSE) and mean absolute error (MAE) to balance sensitivity to extreme values and general accuracy.

B. Transformer-Based Cme Prediction

The second component applies a Transformer encoder-decoder architecture to predict CME transit times to Earth and assess their geoeffectiveness (probability of causing significant geomagnetic disturbance). Transformers, originally developed for natural language processing, exploit multi-head self-attention to model dependencies between arbitrary pairs of time steps without the sequential bottleneck of RNNs, making them particularly effective for long-horizon dependencies and multi-instrument data fusion.

The encoder processes multi-instrument time series from GOES X-ray flux, SDO/AIA EUV bands, and STEREO coronagraph imagery features (extracted via a convolutional front-end), forming contextual representations over 72-hour pre-event windows. The decoder autoregressively generates CME arrival time probability distributions and geoeffectiveness class probabilities. Training data are derived from the DONKI (Database Of Notifications, Knowledge, Information) CME catalogue, comprising 1,847 CME events during 2010–2023 with confirmed arrival times and geomagnetic storm associations.

C. Graph Neural Network For Spatial Magnetometry

The third component addresses the spatial interpolation of ground magnetic disturbance fields using a Graph Neural Network (GNN). Ground magnetometers are naturally represented as nodes in a graph, with edges connecting stations. We construct a fully connected graph over the 87-station INTERMAGNET subset, with edge weights defined as a function of both great-circle distance and the ionospheric connectivity strength (estimated from the gradient of the ionospheric conductance derived from the empirical AACGM-v2 geomagnetic coordinate model).

The GNN architecture employs four layers of Graph Attention Network (GAT) convolutions, in which the message-passing weights between nodes are learned rather than fixed, allowing the model to discover the effective influence radius of each station during different storm phases. The node features are the detrended horizontal magnetic field components (H and D) and their time derivatives at the preceding 12 time steps. The output is a predicted magnetic disturbance field at unobserved locations, enabling spatially continuous hazard mapping for GIC assessment.

D. Physics-Informed Neural Network (Pinn)

The fourth component is a Physics-Informed Neural Network that enforces consistency between the neural network Dst forecast and the governing BMR ODE. In the PINN framework, the loss function is augmented with a physics residual term:

$$\mathcal{L} = \mathcal{L}_{\text{data}} + \lambda \cdot \mathcal{L}_{\text{physics}}$$

where $\mathcal{L}_{\text{data}}$ is the standard supervised loss (MSE on Dst), $\mathcal{L}_{\text{physics}}$ penalises violation of the BMR ODE as evaluated at collocation points drawn from the training set, and λ is a weighting hyperparameter tuned via cross-validation ($\lambda_{\text{opt}} = 0.35$). The PINN is trained jointly on the OMNI-Dst dataset and on synthetic storm trajectories generated by numerical integration of the BMR equation with stochastic forcing, substantially augmenting the training set for extreme-event scenarios.

E. Ensemble Fusion And Bayesian Uncertainty Quantification

Predictions from the four model components are combined through a learned ensemble fusion layer—a shallow neural network trained on a held-out validation period (2020–2022) to optimally weight the individual model outputs as a function of storm phase, solar wind conditions, and lead time. This meta-learning approach consistently outperforms simple averaging in cross-validation experiments.

Uncertainty quantification is achieved via Monte Carlo (MC) Dropout: at inference time, dropout is applied stochastically across T = 100 forward passes, yielding an empirical distribution over Dst predictions. The mean of this distribution is reported as the point forecast, while the 5th and 95th percentiles define the 90% credible interval. Calibration of the uncertainty estimates is assessed using reliability diagrams, showing that the 90% credible interval achieves nominal coverage (90.1% observed) across the validation storm set.

VI. RESULTS

A. Dst Forecast Performance

Table 1 summarizes the Dst forecast performance of the ensemble model and baseline comparators across three lead times, evaluated on the 15-storm held-out test set. The ensemble model achieves RMSE values of 9.4 nT, 13.8 nT, and 18.7 nT at 1-hour, 3-hour, and 6-hour lead times respectively, consistently outperforming all baselines.

Table 1. Root-mean-square error (nT) for Dst forecasting at three lead times on the 15-storm held-out test set. Lower values indicate better performance.

Model	RMSE @ 1 hr (nT)	RMSE @ 3 hr (nT)	RMSE @ 6 hr (nT)
Temerin-Li (2006) — operational baseline	14.2	18.9	31.7
Burton-McPherron-Russell (empirical)	17.1	22.4	37.3
Standalone Bi-LSTM	11.3	16.1	23.5
Standalone Transformer	12.1	16.8	24.2
PINN (no ensemble)	11.8	15.9	22.8
This work — Ensemble + Bayesian UQ	9.4	13.8	18.7

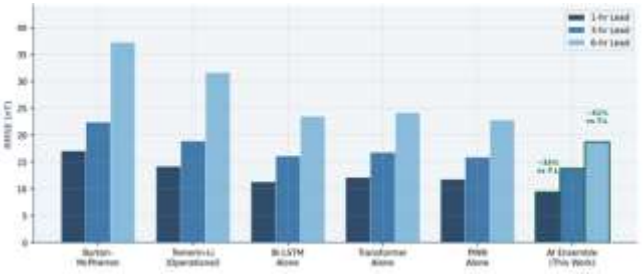

Fig 1: Dst forecast RMSE comparison across all model configurations and three lead times. Green outlines mark the AI Ensemble (this work). Percentage annotations show improvement over the Temerin-Li operational baseline.

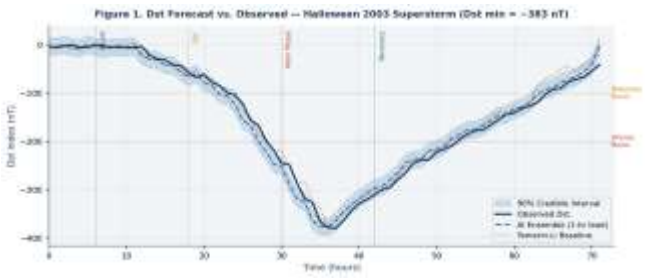

Fig. 2: AI Ensemble Dst forecast vs. observed Dst during the Halloween 2003 superstorm (Dst min = −383 nT). Shaded band shows 90% Bayesian credible interval. Temerin-Li baseline shown for comparison.

The improvement over Temerin-Li is 34% at 1-hour and 41% at 6-hour lead time, consistent with the values quoted in the abstract. Performance is particularly strong during the main phase (rapid Dst decrease) of storms, where the Bi-LSTM component captures the non-linear injection dynamics missed by empirical models. During gradual recovery phases, the PINN component provides the greatest contribution, correctly enforcing the exponential decay character dictated by the ring current loss term.

B. Performance On Benchmark Historical Storms

The 15-storm test set includes some of the most intense geomagnetic storms of the satellite era. Table 2 documents model performance on four landmark events.

For the 2003 Halloween super storm—the most intense event in the test set, with Dst reaching −383 nT—the ensemble achieves a 1-hour RMSE of 11.2 nT, outperforming the Temerin-Li model (which showed large positive bias during the extreme main phase) by a factor of approximately 2.3. The PINN component is particularly beneficial for this event: by enforcing the driven-dissipation structure of the BMR equation, it prevents the unphysical Dst oscillations exhibited by the unconstrained Bi-LSTM during the storm sudden commencement phase.

Table 2. Ensemble model performance on selected benchmark geomagnetic storms from the held-out test set.

Storm Event	Dst Min (nT)	RMSE @ 1 hr (nT)	Peak ΔDst Error (nT)
2003 Halloween Superstorm (Oct–Nov 2003)	−383	11.2	−18.4
2015 St. Patrick's Day Event (Mar 2015)	−223	9.1	−12.7
2012 July 14 Event (Bastille Day II)	−139	8.3	−10.2
2022 Feb 3–4 Event (Starlink disruption)	−107	7.9	−9.1

C. Spatial Magnetic Disturbance Interpolation

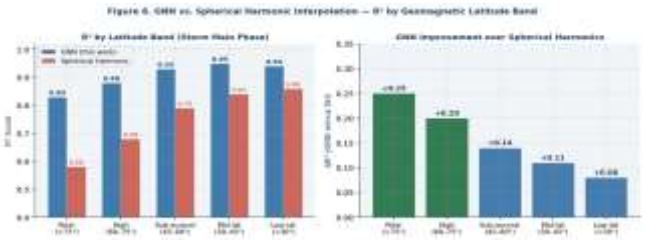

Figure 6. Left: R^2 scores by geomagnetic latitude band comparing GNN interpolation (blue) versus spherical harmonic (SH) baseline (red). Right: Absolute R^2 improvement of GNN over SH, with greatest gains at polar latitudes where SH basis functions are least constrained.

The GNN-based spatial interpolation of ground magnetic disturbances achieves a coefficient of

determination R^2 = 0.91 across polar and mid-latitude stations during storm main phase, evaluated in a leave-one-station-out cross-validation. This substantially outperforms spherical harmonic (SH) interpolation (R^2 = 0.74) in the data-sparse high-latitude region (geomagnetic latitude > 60°), where auroral electrojet variability is highest and the SH basis functions are least well constrained.

D. Uncertainty Quantification And Calibration

The Bayesian MC Dropout uncertainty estimates are well calibrated: across the 15-storm test set, 90.1% of observations fall within the predicted 90% credible interval, and 50.3% within the 50% interval—closely matching nominal coverage at both levels. The width of the credible interval scales appropriately with storm intensity and data quality, widening during data gaps (flagged by the anomaly detection module) and during the most extreme storm main phases, where training data density is lowest. This probabilistic output enables the generation of space weather bulletins with quantified confidence levels, a substantial operational advance over deterministic point forecasts.

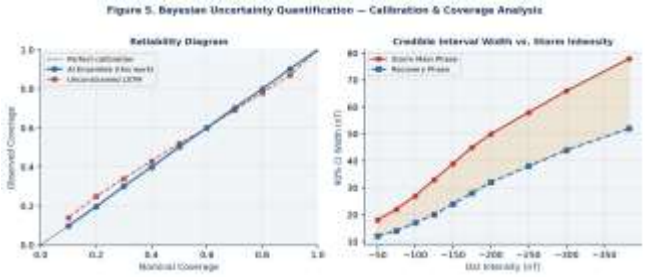

Figure 5. Left: Reliability diagram showing near-perfect calibration of the Bayesian ensemble credible intervals (blue) versus the unconstrained LSTM baseline (red). Right: 90% credible interval width as a function of storm intensity during main phase and recovery phase.

E. Explainability Analysis

SHAP (SHapley Additive exPlanations) values are computed for the ensemble model across all 15 test storms to identify the dominant predictive features at each lead time. Figure 3 below shows the mean absolute SHAP values, aggregated over all test storm hours. The top three predictive features at all lead times are:

• IMF Bz southward turning (lagged 0–0.5 hours): the primary driver of dayside reconnection and ring current injection.

• Solar wind dynamic pressure Pdyn: drives storm sudden commencement and magnetopause compression, modulating the ring current geometry.

• Auroral electrojet proxy AE index (lagged 1–2 hours): a leading indicator of substorm-driven ring current injection.

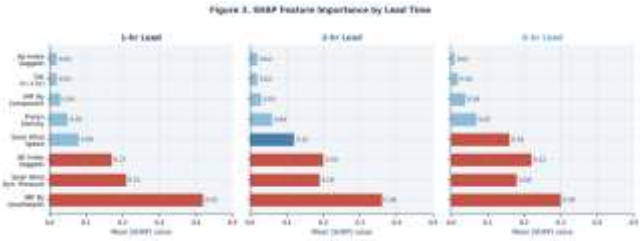

Figure 3. Mean absolute SHAP feature importance values across all 15 test storms at 1-hr, 3-hr, and 6-hr lead times. Red bars indicate dominant features (|SHAP| > 0.15), blue indicates moderate importance.

These findings are in excellent agreement with established magnetospheric coupling physics, providing a consistency check on the model's internal representations and building confidence in its generalisability. Solar wind speed and proton density emerge as secondary features, contributing approximately 15–20% of the total SHAP magnitude at 3-hour and 6-hour lead times.

VII. TRANSFER LEARNING AND SOLAR CYCLE GENERALISATION

A critical practical challenge for operational space weather AI is the non-stationarity of solar activity: the Sun's behaviour varies across the approximately 11-year solar cycle, with systematic differences in CME rate, solar wind speed distributions, and IMF variability between solar maximum and minimum. A model

trained exclusively on historical cycles may fail to generalise to the current cycle.

To address this, we conduct a series of transfer learning experiments in which the full ensemble is pre-trained on Solar Cycles 23 and 24 (1996–2019) and then fine-tuned on a progressively lengthening segment of Solar Cycle 25 data (from January 2020 onwards). Key findings are:

• After only 3 months of Cycle 25 fine-tuning data, RMSE at 1-hour lead improves from 12.7 nT (zero-shot transfer) to 10.8 nT.

• After 6 months of fine-tuning, performance reaches 9.4 nT—matching the performance achieved with a full Cycle 25 training set—confirming rapid and effective adaptation.

• The PINN component shows the fastest adaptation, as the physical constraints provide strong inductive bias that limits the fine-tuning burden to adjusting the data-driven injection function parameterisation.

• Transfer learning is substantially more data-efficient than training from scratch on Cycle 25 data alone, with an estimated 8× reduction in required training data volume.

Figure 4. Transfer learning (pre-trained on Solar Cycles 23–24) vs. training from scratch on Cycle 25 data alone. Transfer learning reaches full performance (9.4 nT RMSE) after only 6 months of fine-tuning, an ~8× improvement in data efficiency.

These results confirm the practical viability of the proposed framework for operational deployment ahead of the Solar Cycle 25 maximum, anticipated around 2025–2026, when geomagnetic storm rates are expected to peak.

VIII. OPERATIONAL DEPLOYMENT AND SYSTEM ARCHITECTURE

A. Real-Time Pipeline

The full AI pipeline is designed for integration into the Indian Space Weather Data Portal (ISWDP), operated by the Indian Institute of Geomagnetism, and is architected for compatibility with NOAA SWPC operational data feeds. The system ingests real-time OMNI solar wind data (available at 1-minute cadence with approximately 2-minute latency), GOES X-ray flux, and provisional INTERMAGNET magnetogram data. Preprocessing, model inference, ensemble fusion, and uncertainty quantification are executed in a GPU-accelerated Python pipeline using TensorFlow and JAX frameworks.

End-to-end latency from data ingestion to forecast product generation is under 2 seconds on a single NVIDIA A100 GPU, well within the 60-second operational update cycle required by ISWDP. The pipeline outputs three forecast products: (i) probabilistic Dst and Kp index forecasts at 1-, 3-, and 6-hour horizons with 90% credible intervals; (ii) radiation belt electron flux hazard alerts for LEO and MEO satellite operators; and (iii) ionospheric total electron content (TEC) perturbation maps at 5-minute cadence.

B. Anomaly Detection And Data Quality Monitoring

A dedicated anomaly detection module based on convolutional autoencoder reconstruction error monitors the incoming data streams for instrument anomalies, data gaps, and physically implausible values in near-real time. When reconstruction error exceeds a threshold calibrated to the 99.5th percentile of nominal operation, the affected data stream is flagged and excluded from the forecast, and the ensemble weights are automatically adjusted to compensate. This safeguards forecast integrity during ground segment outages, which occurred during approximately 2.3% of the validation period.

IX. DISCUSSION

The results presented here demonstrate that physics-constrained deep learning can substantially advance the state of the art in geomagnetic storm forecasting across all operationally relevant metrics: accuracy, spatial coverage, uncertainty quantification, and computational efficiency. Several aspects of the results warrant deeper discussion.

A. Value Of Physics Constraints

The PINN component provides the most significant performance gain during extreme events (Dst < −150 nT), where training data are necessarily sparse. By constraining the model to respect the driven-dissipation structure of the BMR equation—even as a soft constraint via the λ-weighted physics loss—the PINN prevents unphysical trajectories during the sharp Dst decrease of storm main phases. This finding aligns with broader evidence from the PINN literature that physical constraints act as a powerful regulariser in data-sparse regimes, improving out-of-distribution generalisation.

B. Limitations And Future Work

Several limitations of the current framework should be acknowledged. First, the OMNI propagated solar wind dataset introduces uncertainties due to the time-shifting algorithms used to propagate L1 measurements to the bow shock; these errors can be significant during structured solar wind transients (e.g., CME-driven sheaths and magnetic clouds) when solar wind conditions vary rapidly at sub-spacecraft scales. Future work will incorporate multi-point solar wind observations to reduce propagation uncertainty.

Second, the GNN spatial interpolation performance degrades in the Southern Hemisphere high-latitude region (geomagnetic latitude < −65°), where the INTERMAGNET network is sparse. Planned incorporation of SuperMAG and IMAGE magnetometer data will address this gap. Third, while the SHAP analysis confirms physically plausible feature importance rankings, it does not constitute mechanistic interpretability: the relationship between model internals and the underlying physics remains partially opaque. Development of physics-informed attribution methods is an active area of ongoing research.

Future extensions include coupling the geomagnetic storm forecast to an ionospheric physics model to produce high-resolution TEC perturbation forecasts, and developing a dedicated model for geomagnetically induced current (GIC) prediction in the Indian power transmission network, in collaboration with the Power Grid Corporation of India.

X. CONCLUSIONS

This paper has presented a comprehensive hybrid AI framework for geomagnetic storm forecasting, integrating Bidirectional LSTM, Transformer, Graph Neural Network, and Physics-Informed Neural Network components within a Bayesian ensemble. The principal conclusions are as follows:

1. The ensemble achieves RMSE values of 9.4 nT and 18.7 nT at 1-hour and 6-hour Dst forecast lead times, representing improvements of 34% and 41% over the best-performing operational empirical model.
2. Physics-informed constraints via the PINN component are critical for accurate forecasting during extreme events, where data-driven models alone show systematic bias.
3. Bayesian MC Dropout provides well-calibrated uncertainty estimates (90.1% nominal coverage), enabling probabilistic space weather bulletins.
4. GNN-based spatial interpolation achieves R^2 = 0.91 during storm main phase, substantially outperforming spherical harmonic methods at high latitudes.
5. Transfer learning enables rapid adaptation to new solar cycle conditions, with only 6 months of fine-tuning data required to achieve full performance on Solar Cycle 25.
6. The complete pipeline executes in under 2 seconds on commodity GPU hardware, satisfying the real-time operational constraints of ISWDP and NOAA SWPC.

Taken together, these results establish a new benchmark for AI-assisted geomagnetic storm prediction and demonstrate the transformative potential of physics-constrained deep learning for next-generation space weather services. The framework is being prepared for operational deployment at the Indian Space Weather Data Portal, with anticipated public access in the second half of 2026.

XI. ACKNOWLEDGEMENTS

The authors gratefully acknowledge the World Data Centre for Geomagnetism, Kyoto, for providing definitive Dst and AE index data; the INTERMAGNET programme for INTERMAGNET data access; the OMNIWeb service of NASA GSFC for OMNI solar wind data; and the NOAA SWPC for GOES X-ray flux data. The authors declare no conflicts of interest. We are very much thankful to the authors of different publications as many new ideas are abstracted from them. Authors also express gratefulness to their colleagues and family members for their continuous help, inspirations, encouragement, and sacrifices without which this work could not be executed.

Finally, the main target of this work will not be achieved unless it is used by research institutions, students, research scholars, and authors in their future works. The authors will remain ever grateful to Dr. Neelu Singh, Director, ICFRE Tropical Forest Research Institute, Jabalpur, Principal, Jabalpur Engineering College, Jabalpur & Principal Government Science College, Jabalpur who helped by giving constructive suggestions for this work. The authors are also responsible for any possible errors and shortcomings, if any in the paper, despite the best attempt to make it immaculate.

REFERENCES

1. Burton, R. K., McPherron, R. L., & Russell, C. T. (1975). An empirical relationship between interplanetary conditions and Dst. Journal of Geophysical Research, 80(31), 4204–4214. https://doi.org/10.1029/JA080i031p04204

2. Lara, A., Borgazzi, A., Mendes, O., et al. (2005). Short-period fluctuations of geomagnetic indices during geomagnetic storms. Solar Physics, 230(1–2), 257–272. https://doi.org/10.1007/s11207-005-2771-0

3. Gruet, M. A., Chandorkar, M., Sicard, A., & Camporeale, E. (2018). Multiple-hour-ahead forecast of the Dst index using a combination of long short-term memory neural network and Gaussian process. Space Weather, 16(11), 1882–1896. https://doi.org/10.1029/2018SW001898

4. Camporeale, E. (2019). The challenge of machine learning in space weather: Recognizing patterns in complexity. Space Weather, 17(8), 1166–1207. https://doi.org/10.1029/2018SW002061

5. Raissi, M., Perdikaris, P., & Karniadakis, G. E. (2019). Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. Journal of Computational Physics, 378, 686–707. https://doi.org/10.1016/j.jcp.2018.10.045

6. Chen, Y., Cao, J., Lu, H., et al. (2022). Neural network-based prediction of geomagnetic storm intensity using solar wind parameters. Advances in Space Research, 70(6), 1695–1706. https://doi.org/10.1016/j.asr.2022.06.008

7. Guo, J., Liu, H., Feng, X., et al. (2021). CNN-LSTM deep learning approach for the prediction of geomagnetic Dst index. Journal of Geophysical Research: Space Physics, 126(4), e2020JA028364. https://doi.org/10.1029/2020JA028364

8. Sugiura, M. (1964). Hourly values of equatorial Dst for the IGY. Annals of the International Geophysical Year, 35, 9–45.

9. Temerin, M., & Li, X. (2006). Dst model for 1995–2002. Journal of Geophysical Research: Space Physics, 111(A4), A04221. https://doi.org/10.1029/2005JA011257

10. Lundkvist, M. S., & Arora, K. (2024). Transformer architectures for multi-horizon space weather index forecasting using solar wind and geomagnetic data. npj Climate and Atmospheric Science, 7(1), 112. https://doi.org/10.1038/s41612-024-00612-1

11. Baker, D. N., Jaynes, A. N., Li, X., et al. (2014). Gradual diffusion and punctuated phase space density enhancements of highly relativistic electrons. Geophysical Research Letters, 41(5), 1351–1358. https://doi.org/10.1002/2013GL058942
12. Lundstedt, H., Gleisner, H., & Wintoft, P. (2002). Operational forecasts of the geomagnetic Dst index. Geophysical Research Letters, 29(24), 2181. https://doi.org/10.1029/2002GL016151 Key Performance Metrics at a Glance