

---

# **IonCast: A Deep Learning Framework for Forecasting Ionospheric Dynamics**

---

**Halil S. Kelebek**

Department of Engineering Science  
University of Oxford, UK  
halil@robots.ox.ac.uk

**Linnea M. Wolniewicz**

Department of Information and Computer Science  
University of Hawai'i at Mānoa, USA  
linneamw@hawaii.edu

**Michael D. Vergalla**

Free Flight Research Lab  
Sunnyvale, USA  
mike@freeflightlab.org

**Simone Mestici**

Università degli Studi di Roma Sapienza  
Rome, Italy  
simone.mestici@uniroma1.it

**Giacomo Acciarini**

European Space Agency (ESA)  
giacomo.acciarini@esa.int

**Bala Poduval**

University of New Hampshire

**Olga Verkhoglyadova**

NASA Jet Propulsion Laboratory

**Madhulika Guhathakurta**

NASA Headquarters  
madhulika.guhathakurta@nasa.gov

**Thomas E. Berger**

Space Weather Technology, Research, and Education Center  
University of Colorado Boulder  
Thomas.Berger@colorado.edu

**Frank Soboczenski**

Department of Computer Science  
University of York & King's College London, UK  
frank.soboczenski@york.ac.uk

**Atılım Güneş Baydin**

Department of Computer Science  
University of Oxford, UK  
gunes@robots.ox.ac.uk

## **Abstract**

The ionosphere is a critical component of near-Earth space, shaping GNSS accuracy, high-frequency communications, and aviation operations. For these reasons, accurate forecasting and modeling of ionospheric variability has become increasingly relevant. To address this gap, we present IonCast, a suite of deep learning models that include a GraphCast-inspired model tailored for ionospheric dynamics. IonCast leverages spatiotemporal learning to forecast global Total Electron Content (TEC), integrating diverse physical drivers and observational datasets. Validating on held-out storm-time and quiet conditions highlights improved skill compared to persistence. By unifying heterogeneous data with scalable graph-based spatiotemporal learning, IonCast demonstrates how machine learning can augment physical understanding of ionospheric variability and advance operational space weather resilience.

## 1 Introduction

The Earth’s ionosphere, extending from  $\sim 50$  to  $1500$  km altitude, is a partially ionized region of the upper atmosphere shaped by solar radiation, magnetospheric convection, and thermospheric dynamics [1]. The complex coupling within near-Earth space drives a wide spectrum of phenomena, from small-scale plasma irregularities [2] to large-scale geomagnetic storms that have tangible consequences for modern society. Ionospheric disturbances, for example, can significantly degrade the accuracy of Global Navigation Satellite Systems (GNSS) [3], while geomagnetic storms generate geomagnetically induced currents that pose serious risks to power grids [4] and can also compromise orbital stability in Low Earth Orbit [5]. As modern society reliance on space-based infrastructure grows, accurate ionospheric forecasting has become essential [6].

The Total Electron Content (TEC), defined as the integral of the electron density along the ray path between two points, is one of the most important descriptors of the state of the ionosphere [7]. The Jet Propulsion Laboratory (JPL) generates Global Ionospheric Maps (GIM) of TEC with a temporal cadence of 15 minutes and provides these products as part of a near-real-time service [8]. These resources are fundamental to understand the global spatio-temporal evolution of TEC and serve as a foundation for building data-driven models.

Several empirical and physics-based models have been developed to generate local or GIM of TEC as a function of solar and geomagnetic drivers. Empirical models such as the International Reference Ionosphere (IRI) [9, 10] and physics-based models such as the Global Ionosphere-Thermosphere Model (GITM; [11]) have been instrumental in advancing ionospheric research, but carry intrinsic limitations that have been well documented [12].

The increasing availability of high-quality ionospheric datasets make research in this field particularly well-suited for machine learning (ML) applications. Although still in its early stages, several studies have demonstrated the promise of ML-based TEC modeling. Current approaches rely on classical methods such as gradient boosting (XGBoost) or Multi-Layer Perceptrons (MLP) [13, 14], while more advanced methods like Bi-directional Long-Term-Short-Term Memory (BiLSTM) [15] remain restricted to narrow geographic regions. A recent series of work also investigated the application of deep learning architectures to model and forecast global TEC using several physically relevant inputs [16, 17]. These works have shown that deep learning architectures are indeed capable of consistently reproducing the complex ionospheric behavior.

This context highlights the need for advanced ML architectures that can process heterogeneous and multi-source data, operate at global scales, and deliver reliable nowcasting, i.e. the model learns spatiotemporal patterns to project the current state forward in time, and forecasting capabilities. To address this challenge, we present IonCast, a modular framework built on state-of-the-art ML methods. IonCast is designed as a global, data-driven nowcast and forecast model of the TEC, aiming to complement and potentially substitute empirical and physics-based approaches. Model code and data used to produce the IonCast models can be found on GitHub<sup>1</sup>.

## 2 Methods

### 2.1 Data

For this work a combination of space- and ground-based data was processed, integrated to make it machine learning ready. In particular, the dataset curated consists of 2D vertical TEC global maps and 1D driver time series. The 2D component consists of JPL Global Ionospheric Maps (GIM) of vertical TEC at 15-min cadence (JPL/IGS service; [18, 8, 19]). The 1D drivers include solar-wind and geomagnetic parameters (e.g., SYM-H/ASY-D, IMF ( $B_{xyz}$ ),  $V_{sw}$ ) from NASA/GSFC OMNI2 via OMNIWeb ([20, 21, 22]); planetary activity indices ( $K_p, A_p$ ) (compiled from NOAA SWPC/GFZ; [23]); and solar irradiance proxies (F10.7 and S10.7/M10.7/Y10.7) [24]. Auxiliary spatial features use quasi-dipole magnetic coordinates [25] and orbital mechanics data (Sun/Moon ephemerides: subsolar/sublunar points, zenith angles, Earth–Sun/Moon distances; [26]). All sources are publicly accessible via the cited archives; processed outputs are released with this work as standardized 2D maps and aligned 1D time series. To the best of our knowledge, this dataset is one of the most comprehensive assembled for this type of analysis. This dataset’s inclusion of auxiliary spatial

<sup>1</sup><https://github.com/FrontierDevelopmentLab/2025-HL-Ionosphere>

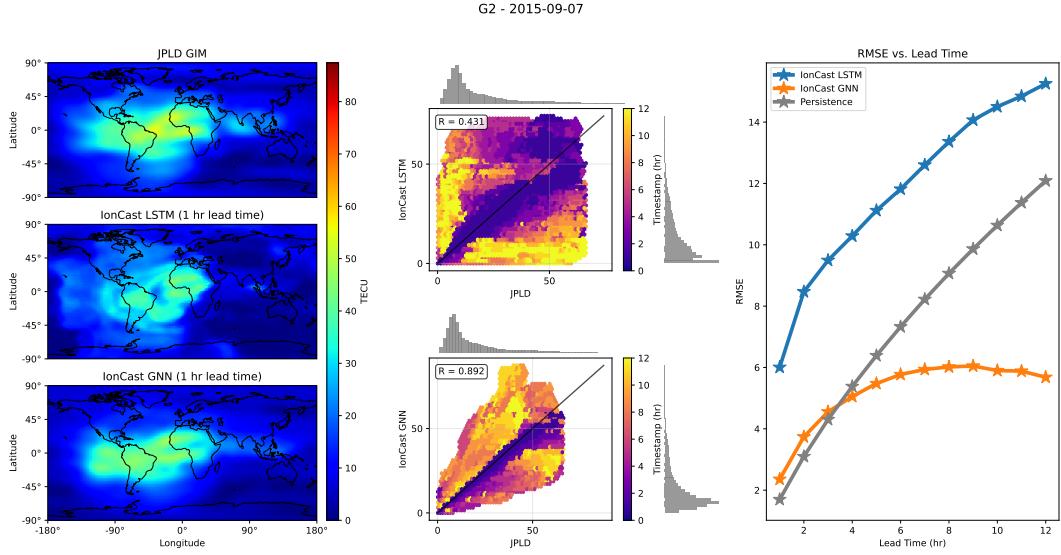


Figure 1: Model predictions plotted and evaluated over various forecast lead times for a moderate geomagnetic storm (G2 event). *Left:* Subplots show ground truth data from JPL (upper) plotted against predictions by the IonCast LSTM (middle) and IonCast GNN (lower) models at a 60-minute lead time. The colorbar has units TECU (TEC Units). *Center:* The joint distribution between JPL (x-axis) and predictions (y-axis) over a 12-hour long-horizon forecast. The diagonal indicates perfect prediction; points below (above) the diagonal correspond to underprediction (overprediction). The colorbar encodes the mode of the forecast time (minutes) at which the prediction occurs, highlighting the error growth across the sequence. Marginal histograms show the distributions of ground truth and predicted values throughout the forecast. *Right:* The plot compares the average global RMSE of the IonCast LSTM (blue), IonCast GNN (orange), and persistence model predictions across growing forecast horizons.

features allows us to investigate the impact of providing the model further physical information, which was not attempted in previous ML applications.

## 2.2 Models

We evaluate the forecasting performance of two models for total electron content (TEC): a long short-term memory (LSTM) [27] baseline, IonCast LSTM, and a Graph-based architecture inspired by GraphCast [28], IonCast GNN. Both models autoregressively predict TEC using JPL [18, 8, 19] data products and learn the dynamics of the ionosphere well. Model hyperparameters were tuned using WandB [29] hyperparameter sweeps.

The IonCast LSTM model uses a convolutional encoder–decoder architecture with an LSTM bottleneck, where the convolutional neural network (CNN) encodes each two-dimensional TEC map into a latent representation, the LSTM maintains a hidden state across timesteps, and the CNN decoder reconstructs the forecasted sequence. This model uses a six-layer convolutional LSTM with circular padding to downsample 180x360 ionospheric maps (and other data sources included as extra image channels) to a 128-dimensional embedding and process the temporal sequences of the embeddings. The decoder then reconstructs predictions using bilinear upsampling and transposed convolutions back to the original resolution. IonCast LSTM is trained with a batch size of 4, dropout rate of 0.15, learning rate of 2e-4, and JPLD loss weight of 20.

The IonCast GNN model is based on Google’s GraphCast architecture implemented in PyTorch by the NVIDIA PhysicsNeMo framework [30]. The architecture follows an encoder, processor, decoder structure. The encoder performs message passing on a graph that converts from the latitude-longitude grid to a spherical icosahedral mesh, the processor learns via message passing on a multi-layer icosahedral mesh, and finally, the decoder learns on the graph mapping from the mesh back to a

latitude-longitude grid. IonCast GNN takes as input all the information from previous states and global parameters as grid node features in the initial grid-to-mesh graph. To generate autoregressive forecasts, IonCast GNN predicts all input variables except those identified as forcings. Forcings are features whose values are known or can be analytically computed for any given timestamp, be it in the past or future. In the case of IonCast GNN, the forcings include all orbital mechanics-derived quantities, such as the positions of the Sun and Moon. During both training and inference, the model is provided with the ground truth forcings. These forcing channels are concatenated as additional features to each grid node, including those corresponding to the prediction timesteps. For all non-forcing channels, which include the TEC maps, the 1D solar and geomagnetic drivers, and quasi-dipole magnetic coordinates, IonCast GNN operates autoregressively. The model receives as input the ground truth forcing and non-forcing values for timestamps within the context window, up to time  $t$ . From the ground truth context, the model autoregressively generates predictions for times  $t + 1$  up to  $t + k$  with no additional ground truth values for the non-forcing channels. The IonCast GNN model is trained with six multi-mesh levels and six processor layers, with message passing occurring between the 32-hop neighbors of each node. The model is trained with a batch size of 1, dropout rate of 0.15, learning rate of 3e-4, and JPLD loss weight of 2.

Both models are trained on the data detailed in Section 2.1, covering the time period from 2010-05-13T00:00:00 to 2024-08-01T00:00:00 at 15-minute cadence, to align with JPL-GIM data. However, to train in a computationally efficient manner, the IonCast LSTM and GNN models train on every 256th sequence of 2 hours (sequences skip 2.66 days between start and end dates). Training is conducted on the Google Cloud Platform. Test and validation data are removed from the training set for specified dates that span various levels of geomagnetic storms, as defined by the NOAA geomagnetic storm scale (G0, G1, G2, G3, G4, and G5), for a total of 10% of storms at each scale removed from the training set. Models are trained to minimize the mean squared error between predictions and ground truth values, for all targets except forcing features. Both the IonCast LSTM and IonCast GNN models are trained with context windows of 8 (3 hours) to predict 1 timestep (15 minutes) ahead, and predict residual targets ( $x_{T+1} = x_T + \hat{x}_{predicted}$ ).

### 3 Results

Model performance is evaluated by comparing forecasted predictions against the GIM released by JPL [19], which are processed from GNSS stations. To benchmark the forecast performance of our models against JPL TEC maps, we compute the root mean squared error (RMSE) over geomagnetic test events withheld during model training. The RMSE is computed over the full event duration, latitude, and longitude TEC predictions compared with JPL TEC maps. The performance of our models across various forecast lead times for a moderate storm event (G2) is shown in Figure 1. Qualitatively, both models properly reconstruct the main features of the global TEC, including the equatorial enhancement, with the IonCast GNN model reconstructing the global structure of the TEC more accurately than the LSTM.

As a baseline, we use a persistence model, which takes the TEC map from the previous frame as the prediction for the TEC map at a given lead time. The IonCast GNN models outperforms persistence for almost all forecast lead times, with the performance over the persistence improving as the forecast horizon increases, as shown by the orange line in the RMSE plot dipping below the gray. The IonCast LSTM model struggles at longer horizon forecasts compared to the GNN, having a bias towards over-predicting the TEC as the forecast proceeds, whereas the RMSE for IonCast GNN stabilizes from hours 6 to 12. This is visible in the ground truth vs. prediction hexbin plots by the dark colors, which correspond to short lead times, falling along the perfect prediction (diagonal line), and lighter colors corresponding to larger lead times moving further from the diagonal line, especially for the LSTM model.

Model performance across different latitudinal bands and storm levels is examined in Appendix Figure 3, and a better-performing IonCast LSTM model that beats persistence is shown in Appendix Figure 2. Additionally, we compare the performance of our models across growing forecast horizons to the state-of-the-art empirical model, IRI [9], in Appendix Table 2. We find that the IonCast GNN model outperforms IRI up to a forecast horizon of 6 hours, however, it is not a perfect comparison as IRI is not a forecast model.

To assess the contribution of different data sources to model performance, we performed an ablation study by isolating the contribution of each data source to IonCast GNN. We selected the combination of input features presented in Table 1 to enhance the understanding of how different classes of physical inputs contribute to forecasting TEC. For these reasons, the classes of solar radiation, solar wind, geomagnetic indices, and orbital mechanics were separated. The models were trained at a 15-minute cadence with a 2-hour context window and evaluated on a 12-hour-long horizon forecast over JPLD predictions. All data streams other than JPLD + F10.7 outperform the model trained solely on a TEC input, suggesting they play a role in aiding the TEC prediction. The poor performance of the F10.7 model may be due to the fact that the F10.7 parameter is produced at a daily cadence, which varies too slowly to provide useful information for forecasts over a 12-hour horizon at a 15-minute temporal resolution. The main contribution to model performance was from the inclusion of the orbital mechanics and quasi-dipole magnetic coordinates, even outperforming the model trained on all data streams. The importance of the orbital mechanics and magnetic coordinate maps is that they provide the model context on the trajectory of the TEC as the Earth rotates, through the encoding of the subsolar and sublunar points as well as the structure of the Earth’s magnetic field. The models trained without these channels struggle to account for the apparent motion of the ionosphere over longer forecast horizons (>4 hours), resulting in an eventual spatial drift, explaining the increase in RMSE. One reason the model trained on the JPLD + orbital mechanics + quasi-dipole outperforms the model trained on all datastreams may be caused by the reduced number of channels over which the loss is computed, as the orbital mechanics are forcings and aren’t included in the model loss. Thus the model is free to focus on optimizing the TEC channel performance. This may suggest that the performance gained by the inclusion auxiliary channel is not great enough to compensate for the increased complexity in the optimization. Finally, using a residual target was shown to provide a notable improvement in model performance.

Table 1: Ablation study over the influence of input data. RMSE values are averaged over a 12-hour forecast on the validation set. Bold values highlight the experiment with the lowest RMSE value.

Input Features	RMSE (TECU)
JPLD	$22.4 \pm 3.2$
JPLD + F10.7	$23.9 \pm 10.2$
JPLD + F10.7, S10.7, M10.7, JB2008	$13.3 \pm 4.4$
JPLD + Ap & Kp	$12.7 \pm 3.2$
JPLD + Bx/By/Bz & vx/vy/vz (Omniweb)	$15.5 \pm 12.6$
JPLD + Orbital Mechanics + Quasi-Dipole	<b><math>9.2 \pm 4.5</math></b>
JPLD + All (Non-Residual Target)	$18.8 \pm 10.7$
JPLD + All	$10.7 \pm 4.5$

## 4 Summary

The proposed IonCast GNN model accurately forecasts global TEC over a long lead time while processing heterogeneous multi-source data. To the best of our knowledge, IonCast GNN is the first machine learning approach to combine heterogeneous solar and geomagnetic driver data with TEC observations to forecast global TEC. Due to the increasing dependence on satellite infrastructure, an accurate forecast model for the ionosphere is necessary. IonCast GNN allows for the global forecast of TEC over long lead times with high accuracy, even during geomagnetic storm events. The preliminary results of our model show promise for the field of ionospheric forecasting and open the door for further exploration.

## Acknowledgments

This research is the result of the Frontier Development Lab, Heliolab a partnership between NASA, Trillium Technologies Inc. (USA), Google Cloud, NVIDIA and Pasteur Labs, Contract No. 80GSFC23CA040. A portion of research was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with NASA. The authors thank Andrew Smith and Umaa D. Rebbapragada for their valuable insights, NASA’s Goddard Space Flight Center, and NASA’s Jet Propulsion Laboratory for their continuing support.

## References

- [1] Michael C Kelley. *The Earth's ionosphere: Plasma physics and electrodynamics*, volume 96. Academic press, 2009.
- [2] Roland T Tsunoda. High-latitude f region irregularities: A review and synthesis. *Reviews of Geophysics*, 26(4):719–760, 1988.
- [3] Jr. Kintner, P. M. Observations of velocity shear driven plasma turbulence. *Journal of Geophysical Research*, 81(A28):5114–5122, October 1976.
- [4] Antti Pulkkinen, E Bernabeu, A Thomson, A Viljanen, R Pirjola, D Boteler, J Eichner, PJ Cilliers, D Welling, NP Savani, et al. Geomagnetically induced currents: Science, engineering, and applications readiness. *Space weather*, 15(7):828–856, 2017.
- [5] Ryuho Kataoka, Daikou Shiota, Hitoshi Fujiwara, Hidekatsu Jin, Chihiro Tao, Hiroyuki Shinagawa, and Yasunobu Miyoshi. Unexpected space weather causing the reentry of 38 starlink satellites in february 2022. *Journal of Space Weather and Space Climate*, 12:41, 2022.
- [6] Thomas E Berger, MJ Holzinger, EK Sutton, and JP Thayer. Flying through uncertainty. *Space Weather*, 18(1):e2019SW002373, 2020.
- [7] Norbert Jakowski, C Mayer, MM Hoque, and V Wilken. Total electron content models and their use in ionosphere monitoring. *Radio Science*, 46(06):1–11, 2011.
- [8] Léo Martire, Thomas F Runge, Xing Meng, Siddharth Krishnamoorthy, Panagiotis Vergados, Anthony J Mannucci, Olga P Verkhoglyadova, Attila Komjáthy, Angelyn W Moore, Robert F Meyer, et al. The jpl-gim algorithm and products: multi-gnss high-rate global mapping of total electron content. *Journal of Geodesy*, 98(5), 2024.
- [9] Dieter Bilitza, Lee-Anne McKinnell, Bodo Reinisch, and Tim Fuller-Rowell. The international reference ionosphere today and in the future. *Journal of Geodesy*, 85(12):909–920, 2011.
- [10] Dieter Bilitza, David Altadill, Vladimir Truhlik, Valentin Shubin, Ivan Galkin, Bodo Reinisch, and Xueqin Huang. International reference ionosphere 2016: From ionospheric climate to real-time weather predictions. *Space weather*, 15(2):418–429, 2017.
- [11] AJ Ridley, Yue Deng, and G Tóth. The global ionosphere–thermosphere model. *Journal of Atmospheric and Solar-Terrestrial Physics*, 68(8):839–864, 2006.
- [12] Dieter Bilitza, Michael Pezzopane, Vladimir Truhlik, David Altadill, Bodo W Reinisch, and Alessio Pignalberi. The international reference ionosphere model: A review and description of an ionospheric benchmark. *Reviews of geophysics*, 60(4):e2022RG000792, 2022.
- [13] Aleksei V Zhukov, Yury V Yasyukevich, and Aleksei E Bykov. Gimli: Global ionospheric total electron content model based on machine learning. *GPS Solutions*, 25(1):19, 2021.
- [14] A Smirnov, Y Shprits, F Prol, H Lühr, M Berrendorf, I Zhelavskaya, and C Xiong. A novel neural network model of earth’s topside ionosphere. *sci rep* 13: 1303, 2023.
- [15] Pan Xiong, Dulin Zhai, Cheng Long, Huiyu Zhou, Xuemin Zhang, and Xuhui Shen. Long short-term memory neural network for ionospheric total electron content forecasting over china. *Space Weather*, 19(4):e2020SW002706, 2021.
- [16] Lei Liu, Shasha Zou, Yibin Yao, and Zihan Wang. Forecasting global ionospheric tec using deep learning approach. *Space Weather*, 18(11):e2020SW002501, 2020. e2020SW002501 2020SW002501.
- [17] Xiaodong Ren, Pengxin Yang, Dengkui Mei, Hang Liu, Guozhen Xu, and Yue Dong. Global ionospheric tec forecasting for geomagnetic storm time using a deep learning-based multi-model ensemble method. *Space Weather*, 21(3):e2022SW003231, 2023. e2022SW003231 2022SW003231.

- [18] AJ Mannucci, BD Wilson, DN Yuan, CH Ho, UJ Lindqwister, and TF Runge. A global mapping technique for gps-derived ionospheric total electron content measurements. *Radio science*, 33(3):565–582, 1998.
- [19] Olga Verkhoglyadova and Xing Meng. Global ionospheric maps for research – jpld data product. [https://sideshow.jpl.nasa.gov/pub/iono\\_daily/gim\\_for\\_research/jpld/](https://sideshow.jpl.nasa.gov/pub/iono_daily/gim_for_research/jpld/), April 2024. Last updated: 8 Apr 2024. Government sponsorship acknowledged.
- [20] Edward C Stone, AM Frandsen, RA Mewaldt, ER Christian, D Margolies, JF Ormes, and F Snow. The advanced composition explorer. *Space Science Reviews*, 86(1):1–22, 1998.
- [21] A Nishida, K Uesugi, I Nakatani, T Mukai, DH Fairfield, and MH Acuna. Geotail mission to explore earth’s magnetotail. *Eos, Transactions American Geophysical Union*, 73(40):425–429, 1992.
- [22] Lynn B Wilson III, Alexandra L Brosius, Natchimuthuk Gopalswamy, Teresa Nieves-Chinchilla, Adam Szabo, Kevin Hurley, Tai Phan, Justin C Kasper, Noé Lugaz, Ian G Richardson, et al. A quarter century of wind spacecraft discoveries, 2021.
- [23] J. Matzka, C. Stolle, Y. Yamazaki, O. Bronkalla, and A. Morschhauser. The Geomagnetic K<sub>p</sub> Index and Derived Indices of Geomagnetic Activity. *Space Weather*, 19(5):e2020SW002641, May 2021.
- [24] W Kent Tobiska, BR Bowman, and SD Bouwer. Solar and geomagnetic indices for thermospheric density models. *COSPAR International Reference Atmosphere*, edited by Rees D. and Tobiska WK, 2012.
- [25] Karl Magnus Laundal and Arthur D Richmond. Magnetic coordinate systems. *Space science reviews*, 206(1):27–59, 2017.
- [26] JD Giorgini, DK Yeomans, AB Chamberlin, PW Chodas, RA Jacobson, MS Keesey, JH Lieske, SJ Ostro, EM Standish, and RN Wimberly. Jpl’s on-line solar system data service. In *AAS/Division for Planetary Sciences Meeting Abstracts# 28*, volume 28, pages 25–04, 1996.
- [27] Sepp Hochreiter and Jürgen Schmidhuber. Long short-term memory. *Neural Computation*, 9(8):1735–1780, 11 1997.
- [28] Remi Lam, Alvaro Sanchez-Gonzalez, Matthew Willson, Peter Wirnsberger, Meire Fortunato, Ferran Alet, Suman Ravuri, Timo Ewalds, Zach Eaton-Rosen, Weihua Hu, Alexander Merose, Stephan Hoyer, George Holland, Oriol Vinyals, Jacklynn Stott, Alexander Pritzel, Shakir Mohamed, and Peter Battaglia. Learning skillful medium-range global weather forecasting. *Science*, 382(6677):1416–1421, 2023.
- [29] Lukas Biewald. Experiment tracking with weights and biases, 2020. Software available from wandb.com.
- [30] Physicsnemo | nvidia developer. [Online; accessed 2025-08-30].

## Appendix

Table 2: Comparison of IonCast LSTM and GNN model performance across growing forecast horizons (1, 6, and 12 hours) against IRI [9]. Values represent the RMSE averaged over the entire 48-time step (12-hour) period and all latitudes/longitudes. Bold values are the lowest RMSE for each event level (G0, G2, and G4).

Event Level	IRI	1 Hour		6 Hours		12 Hours	
		LSTM	GNN	LSTM	GNN	LSTM	GNN
G0 - 2018-04-20	5.32	4.39	<b>1.34</b>	8.94	3.44	8.26	10.89
G2 - 2015-09-07	6.24	6.01	<b>2.36</b>	11.82	5.72	11.50	17.14
G4 - 2023-04-23	16.59	14.28	<b>5.82</b>	27.06	13.3	25.72	36.21

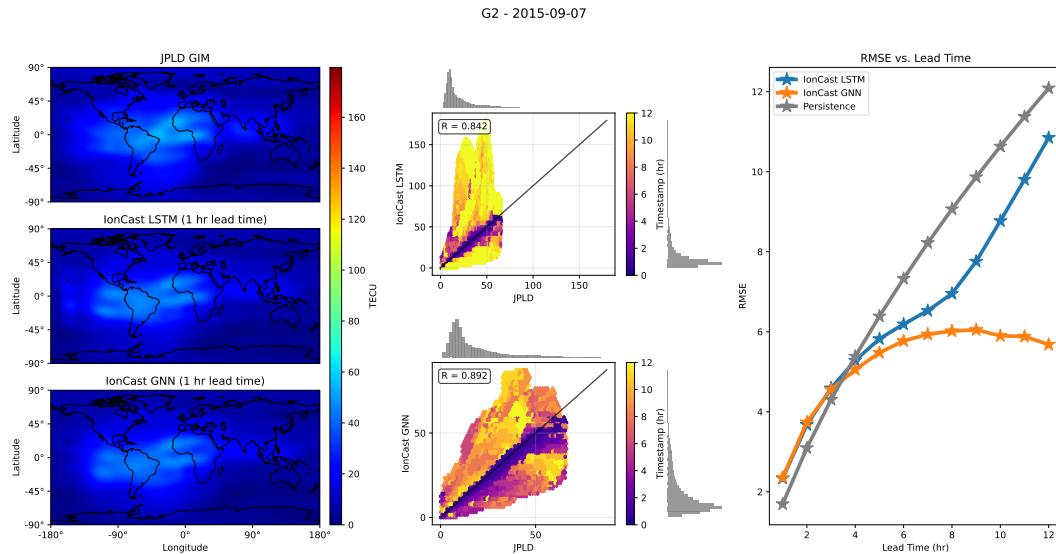


Figure 2: Model predictions plotted and evaluated over various forecast lead times for a moderate geomagnetic storm (G2 event). The IonCast GNN model shown is the same as in the main body. The IonCast LSTM model is trained on a longer context window (16 15-minute cadence time steps) and a larger amount of training data (trains on every 32nd sequence of 4 hours, thus sequences skip 8 hours between start and end dates) than the 8-context and 256 date dilation model evaluated in the main text. *Left:* Subplots show ground truth data from JPL (upper) plotted against predictions by the IonCast LSTM (middle) and IonCast GNN (lower) models at a 60-minute lead time. The colorbar has units TECU (TEC Units). *Center:* The joint distribution between JPL (x-axis) and predictions (y-axis) over a 12-hour long-horizon forecast. The diagonal indicates perfect prediction; points below (above) the diagonal correspond to underprediction (overprediction). The colorbar encodes the mode of the forecast time (minutes) at which the prediction occurs, highlighting the error growth across the sequence. Marginal histograms show the distributions of ground truth and predicted values throughout the forecast. *Right:* The plot compares the average global RMSE of the IonCast LSTM (blue), IonCast GNN (orange), and persistence model predictions across growing forecast horizons.

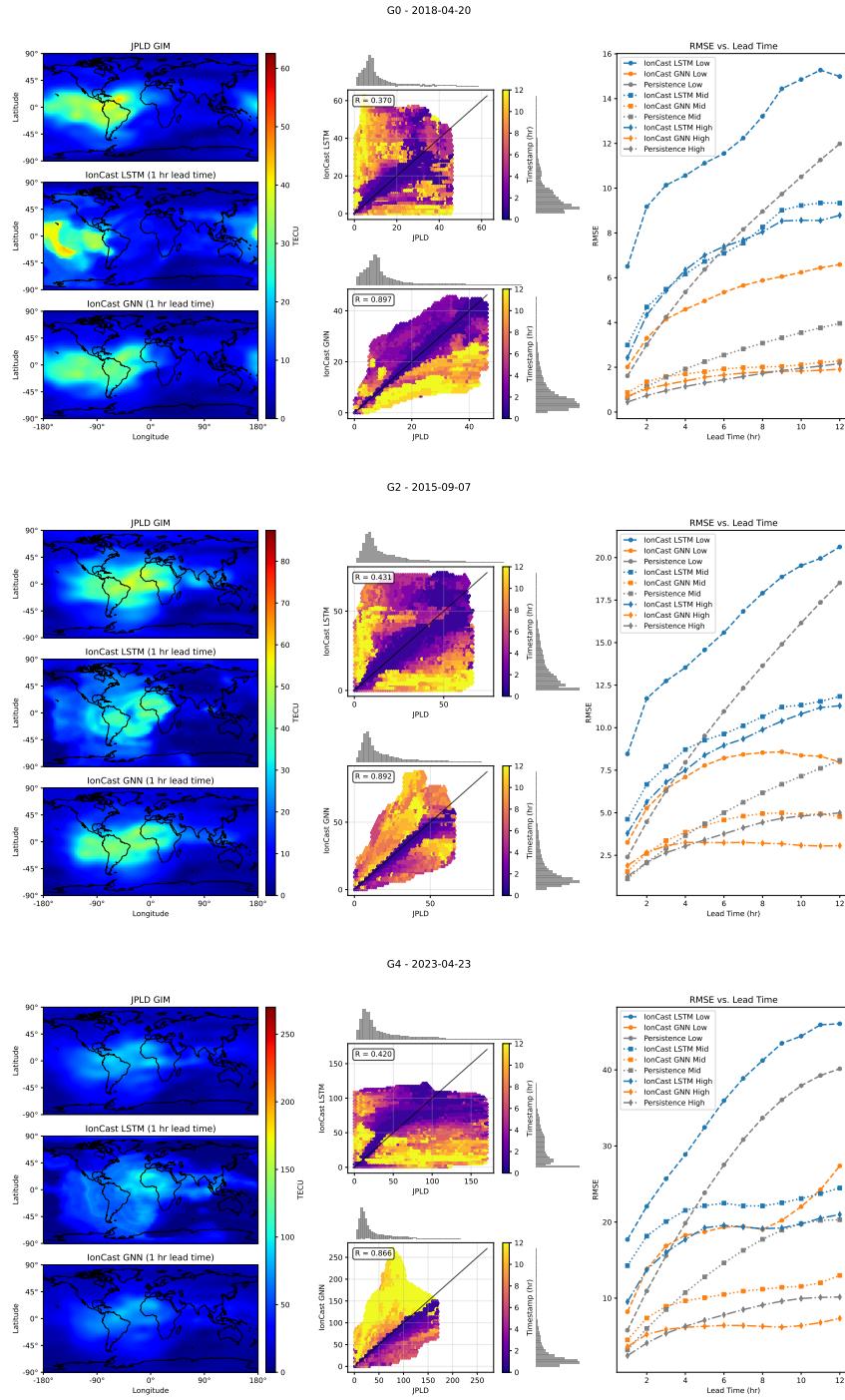


Figure 3: Model predictions plotted and evaluated over various forecast lead times for a quiet ionospheric condition in the top figure (G0 event), a moderate (G2) and a severe geomagnetic event (G4 event) in the mid and bottom figure. *Left* and *Center* panels have the same models and meaning of the one shown in Figure 1. *Right*: The plot on the right compares the RMSE of the IonCast LSTM (blue), IonCast GNN (orange), and persistence model predictions across growing forecast horizons, split across low ( $30^{\circ}$  South to  $30^{\circ}$  North), mid ( $60^{\circ}$  to  $30^{\circ}$  South and  $30^{\circ}$  to  $60^{\circ}$  North), and high latitudes ( $60^{\circ}$  to  $90^{\circ}$  South and  $60^{\circ}$  to  $90^{\circ}$  North).

Table 3: Performance comparison for models trained during different solar cycle phases. There is only one G5 event present in the 2010-2024 date range. Bold values are the lowest RMSE for each event level.

Data Range	RMSE All Events	RMSE G0 Events	RMSE G1 Events	RMSE G2 Events	RMSE G3 Events	RMSE G4 Events	RMSE G5 Events
Solar Maximum 2013–2015	<b><math>8.63 \pm 4.34</math></b>	<b><math>5.79 \pm 0.53</math></b>	$7.09 \pm 1.22$	<b><math>6.57 \pm 1.57</math></b>	$7.50 \pm 1.82$	<b><math>11.36 \pm 5.57</math></b>	18.23
Solar Minimum 2018–2020	$63.35 \pm 58.60$	$7.31 \pm 3.76$	$58.19 \pm 40.21$	$49.61 \pm 42.55$	$31.06 \pm 32.49$	$148.42 \pm 40.12$	60.11
Full Dataset 2010–2024	$8.87 \pm 4.44$	$5.96 \pm 3.14$	<b><math>6.31 \pm 2.56</math></b>	$10.53 \pm 5.11$	<b><math>7.50 \pm 1.09</math></b>	$11.81 \pm 3.37$	<b><math>17.79</math></b>

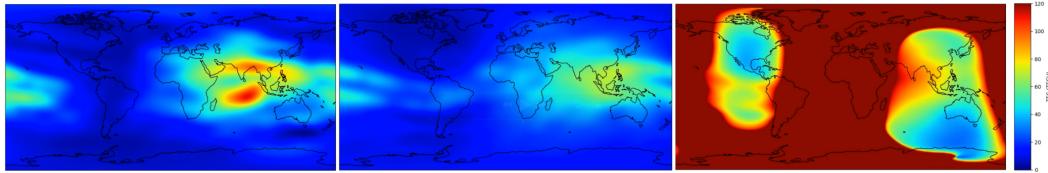


Figure 4: Solar cycle model predictions at a 360-minute lead time for a G4 level event. (Left) JPDL Ground Truth, (Middle) Model trained on solar maximum date range (2013-2015), (Right) Model trained on solar minimum date range (2018-2020).

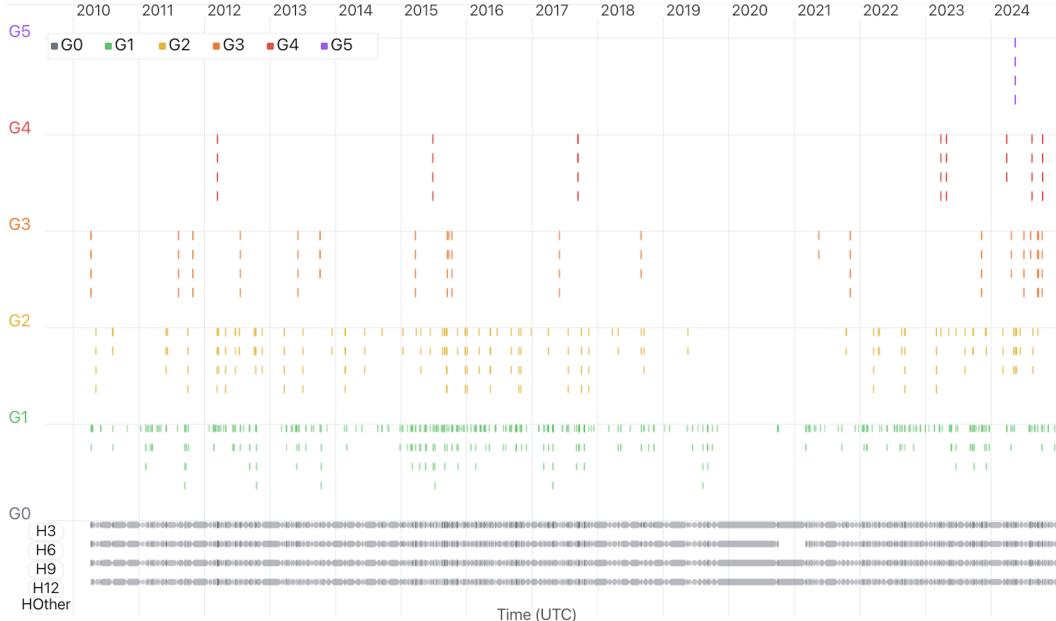


Figure 5: Distribution of G-level events over the data range (2010-2024). The x and y axes represent the time (years) and the intensity of the event (G-level), respectively. Each class bin in the y-axis is then divided into four segments, which correspond to the event duration, as shown in the lower part of the plot.

The purpose of this ablation study was to examine how training on different portions of the solar cycle influences model performance and if models trained on local time periods can better predict specific event classes. All three model runs shown in Table 3, were trained with the same set of hyperparameters, only varying in the date ranges seen in training. To ensure all three models received the same number of training samples, training sequences were more densely sampled for the models trained on the solar minimum and maximum date ranges compared to the model trained on the full date range. The solar maximum model performance is comparable to the model trained on the full date range. However, the solar minimum model struggles in all conditions other than G0 events. This seems to be caused by the lack of non-quiet solar events present in the solar minimum training range as seen in Figure 5, leading to the poor generalization of the solar minimum model evaluated on events out of its training distribution Figure 4.

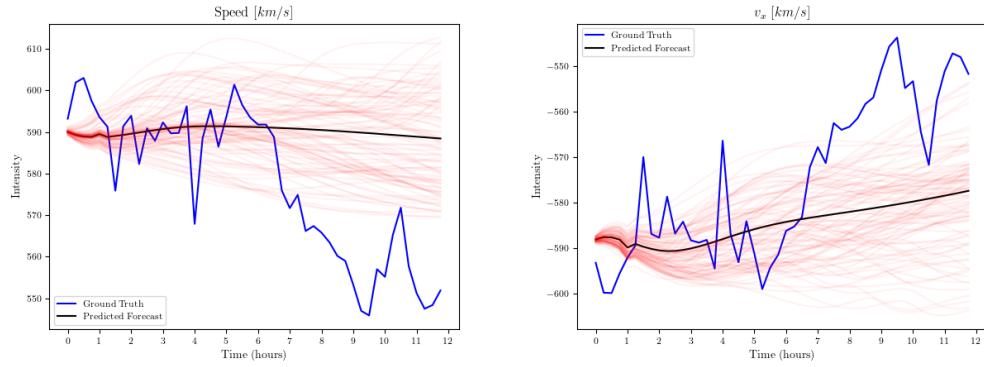


Figure 6: Selected forecasts corresponding to non-TEC 1-D drivers produced by the IonCast GNN model. Both forecasts correspond to the 2015-09-07 G2 event. The sampled forecasts (red) correspond to the forecasts generated at individual  $1^\circ \times 1^\circ$  latitude-longitude nodes. The forecasts at 100 nodes were sampled to generate the red curves. The predicted forecasts (black) correspond to the mean forecast over all nodes of the latitude-longitude grid.

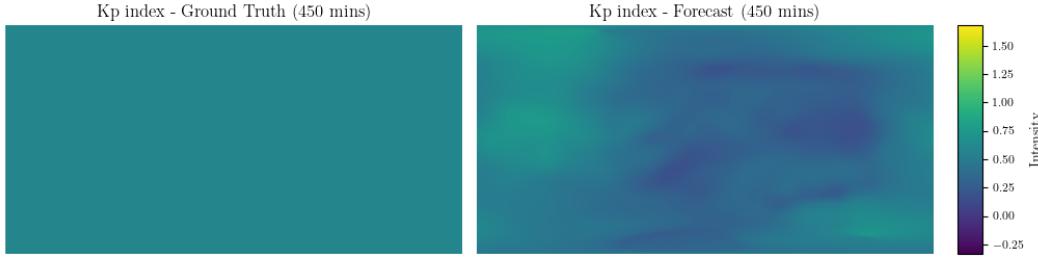


Figure 7: Ground truth Kp index plotted against the global view of the pointwise ( $1^\circ \times 1^\circ$ ) Kp predictions produced by the IonCast GNN model. The forecast corresponds to the quiet ionospheric conditions on 2018-04-20 (G0 event).

While examining the global view of the channels corresponding to 1-D drivers, certain patterns emerge that resemble TEC structure. This structure is shown in Figure 7. This suggests the model may use the redundancy in the global feature channels as registers to inform TEC predictions.