University of Reading

Beyond mean solar wind conditions: turbulence-aware forecasting of the AE index

Article

Published Version

Creative Commons: Attribution 4.0 (CC-BY)

Open Access

Waters, C. L. ORCID: https://orcid.org/0000-0002-5208-773X, Chen, C. H. K. and Owens, M. J. ORCID: https://orcid.org/0000-0003-2061-2453 (2026) Beyond mean solar wind conditions: turbulence-aware forecasting of the AE index. Space Weather, 24 (7). e2026SW005094. ISSN 1542-7390 doi: 10.1029/2026SW005094 Available at https://centaur.reading.ac.uk/131145/

It is advisable to refer to the publisher's version if you intend to cite from the work. See Guidance on citing.

To link to this article DOI: http://dx.doi.org/10.1029/2026SW005094

Publisher: American Geophysical Union

All outputs in CentAUR are protected by Intellectual Property Rights law, including copyright law. Copyright and IPR is retained by the creators or other copyright holders. Terms and conditions for use of this material are defined in the End User Agreement.

www.reading.ac.uk/centaur

University of Reading

CentAUR

Central Archive at the University of Reading

Reading's research outputs online

Check for updates

Space Weather

RESEARCH ARTICLE
10.1029/2026SW005094

Beyond Mean Solar Wind Conditions: Turbulence-Aware Forecasting of the AE Index

Cara L. Waters[1], **Christopher H. K. Chen**[1], **and Mathew J. Owens**[2]

[1]Department of Physics & Astronomy, Queen Mary University of London, London, UK, [2]Department of Meteorology, University of Reading, Reading, UK

Key Points:

- Including solar wind turbulence parameters improves short-term forecasts of auroral electrojet index compared to models using average conditions alone
- Improved forecasts reduce false alarms and better capture extreme geomagnetic events, supporting more reliable operational decision making
- Turbulence-aware forecasts provide insight into how solar wind variability influences energy transfer in Earth's space environment

Correspondence to:
C. L. Waters,
c.waters@qmul.ac.uk

Citation:
Waters, C. L., Chen, C. H. K., & Owens, M. J. (2026). Beyond mean solar wind conditions: Turbulence-aware forecasting of the AE index. *Space Weather*, *24*, e2026SW005094. https://doi.org/10.1029/2026SW005094

Received 27 MAR 2026
Accepted 11 JUL 2026

Author Contributions:
Conceptualization: Cara L. Waters, Christopher H. K. Chen
Data curation: Cara L. Waters
Formal analysis: Cara L. Waters
Funding acquisition: Christopher H. K. Chen
Methodology: Cara L. Waters, Christopher H. K. Chen, Mathew J. Owens
Software: Cara L. Waters
Validation: Cara L. Waters
Visualization: Cara L. Waters
Writing – original draft: Cara L. Waters
Writing – review & editing: Cara L. Waters, Christopher H. K. Chen, Mathew J. Owens

© 2026 The Author(s). Space Weather published by Wiley Periodicals LLC on behalf of American Geophysical Union. This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

Abstract

The auroral electrojet (AE) index is a key indicator of high latitude geomagnetic activity and is widely used in operational space weather monitoring, yet forecasting AE from upstream solar wind conditions remains challenging due to nonlinear coupling, internal magnetospheric dynamics, and multiscale variability. We test whether incorporating solar wind turbulence improves short timescale AE forecasts beyond models based only on mean solar wind and interplanetary magnetic field parameters. Two gradient boosted decision tree (XGBoost) models are developed using near-Earth solar wind observations: a baseline model using standard mean parameters and a turbulence-aware model that additionally includes measures of fluctuation amplitude, intermittency, and Alfvénic structure. Both models achieve peak performance at short lead times, with correlations exceeding 0.8 at 60 min. However, while the baseline model exhibits a clear skill peak at 75 min, the turbulence-aware model maintains comparable skill across 60–90 min horizons, indicating reduced degradation with lead time. The turbulence-aware model also provides consistent improvements over both the baseline and persistence and, critically, improves forecast robustness for high-impact events. Cost–loss analysis shows that, for the baseline model, economic value decreases systematically with increasing AE threshold and the range of cost–loss ratios yielding positive value narrows. In contrast, the turbulence-aware model maintains an approximately constant zero-value cost–loss threshold across all event levels, indicating stable economic usefulness even for extreme AE conditions. This demonstrates that turbulence provides complementary, scale-dependent information beyond mean solar wind parameters, improving both forecast performance and decision-relevant value for operational space weather applications.

Plain Language Summary

The auroral electrojet (AE) index is a measure of electrical currents flowing in Earth's upper atmosphere near the poles and is widely used to monitor space weather conditions that can affect satellites, power systems, and radio communications. Forecasting AE is challenging because Earth's space environment depends not only on the average conditions of the solar wind—a stream of charged particles from the Sun—but also on how turbulent and variable that solar wind is. In this study, we investigated whether including information about solar wind turbulence improves short-term forecasts of AE. We compared two machine learning models: one that uses only average solar wind and magnetic field conditions, and another that also includes simple measures describing solar wind turbulence. We found that the model including turbulence performs better overall, especially during periods of strong geomagnetic activity. Importantly, it is better at forecasting extreme AE events and produces fewer false alarms, which is important for space weather forecasting and decision making. These results show that accounting for solar wind turbulence can improve the reliability of geomagnetic activity forecasts without making prediction models more complex, helping users respond more effectively to space weather hazards.

1. Introduction

The auroral electrojet (AE) index is a widely used proxy for high latitude geomagnetic activity and substorm-related ionospheric current systems. It is constructed from the range of the horizontal magnetic perturbation (east to west) observed at a network of auroral-zone ground magnetometers, providing a global measure of electrojet intensity and its variability on minute to hour timescales (Davis & Sugiura, 1966; Mayaud, 1980). Because AE responds rapidly and strongly to changes in solar wind driving and reflects the efficiency of magnetosphere–ionosphere coupling, it has long been used as a diagnostic of energy transfer in the geospace system and as an indicator of disturbed near-Earth space conditions for operational forecasting.

15427390, 2026, 7, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2026SW005094 by NICE, National Institute for Health and Care Excellence, Wiley Online Library on [03/08/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Forecasting AE from upstream solar wind conditions is challenging for two closely related reasons. First, the coupling between the solar wind and magnetosphere is highly nonlinear and depends on multiple interacting drivers, including the interplanetary magnetic field (IMF) orientation and strength, solar wind speed, density, and dynamic pressure. At the most basic level, southward IMF enables magnetic reconnection at the dayside magnetopause, opening flux and driving large-scale convection in the Dungey cycle framework (Dungey, 1961; Kan & Lee, 1979). This conceptual picture underpins many empirical descriptions of solar wind–magnetosphere coupling and has motivated the development of a wide range of empirical coupling functions and energy input proxies designed to condense the geoeffective solar wind driving into a small number of physically motivated parameters (e.g., Newell et al., 2007; Perreault & Akasofu, 1978). While such functions capture broad statistical relationships, they necessarily simplify the underlying dynamics and do not uniquely represent the pathways by which energy is transferred, stored, and released within the coupled magnetosphere–ionosphere system. Comparative studies have further shown that the performance of coupling functions depends on activity level, timescale, and the specific geomagnetic response under consideration (Finch & Lockwood, 2007; Lockwood, 2022).

Second, geomagnetic activity is not controlled solely by slowly varying mean solar wind conditions. Instead, it is also sensitive to fluctuations, intermittency, and multiscale structure in the solar wind and IMF, which can modulate reconnection rates, plasma transport, and energy conversion processes across a wide range of temporal and spatial scales (Borovsky, 2003; Jankovičovà et al., 2008; Lyons et al., 2003). As a result, two intervals with similar mean solar wind properties can produce different geomagnetic responses, highlighting the importance of variability and turbulence in solar wind–magnetosphere coupling. This variability is reflected not only in the solar wind but also in AE itself, which exhibits intermittent, scale-dependent behavior and signatures of both directly driven and internally generated dynamics (Consolini, 2005; Freeman et al., 2000; Hnat et al., 2003).

The potential importance of solar wind turbulence in driving magnetospheric dynamics has been recognized since early theoretical work. Initial hypotheses proposed that solar wind–magnetosphere interaction could be described in terms of a viscous-like coupling, in which turbulent interactions at the magnetopause transfer momentum and energy into the magnetosphere, driving large-scale convection (Axford & Hines, 1961). Order-of-magnitude estimates suggested that such viscous energy input could be comparable to the energy dissipation associated with geomagnetic storms, and this mechanism was invoked to explain features such as the extended length of the magnetotail and asymmetries in particle trapping boundaries (Axford, 1964).

Subsequent observational and numerical studies have shown that solar wind plasma transport across the magnetopause can occur through multiple boundary processes, including Kelvin-Helmholtz instability and associated vortex-induced mixing along the flanks, which can facilitate the entry of magnetosheath plasma even under northward IMF conditions (Hasegawa et al., 2004; Nykyri & Otto, 2001). In addition, turbulence generated upstream of the bow shock and within the magnetosheath can modify the structure and variability of the magnetopause, influencing the efficiency and intermittency of energy transfer into the magnetosphere (e.g., Zimbardo et al., 2010). In this modern view, viscous-like interaction and turbulence-mediated transport are understood as contributing pathways within a broader coupling system.

Observational studies have since provided further evidence linking solar wind fluctuations to geomagnetic activity. B_z Fluctuations at low frequencies have been shown to impact geomagnetic activity strongly during both strongly northward and southward IMF (Osmane et al., 2015). Alfvénic turbulence in fast solar wind streams has been shown to act as an effective driver of auroral activity, particularly during solar minimum, even under relatively weak mean solar wind forcing (D'Amicis et al., 2009). During solar maximum, when solar wind fluctuations are more strongly influenced by advected coronal structures, this relationship weakens, suggesting a change in the dominant coupling regime (D'Amicis et al., 2010). More generally, statistical analyses of solar wind fluctuations have demonstrated that while skewness remains close to zero across scales and does not strongly distinguish between quiet and active intervals, kurtosis increases significantly at smaller scales and during disturbed periods, reflecting enhanced intermittency (Jankovičovà et al., 2008). This supports the idea that non-Gaussian, intermittent fluctuations may modulate the efficiency of solar wind–magnetosphere coupling, beyond what is captured by mean solar wind parameters alone (Borovsky, 2003). Measures characterizing the Alfvénic nature of the solar wind, including cross helicity and residual energy, have also been shown to vary systematically with geomagnetic activity and solar cycle phase, further supporting a link between turbulence properties and geoeffectiveness (D'Amicis et al., 2007, 2011). A recent review has reinforced this overall picture, arguing that

15427390, 2026, 7, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2026SW005094 by NICE, National Institute for Health and Care Excellence, Wiley Online Library on [03/08/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

turbulence properties in the upstream solar wind can contribute significantly to geomagnetic response, especially at high latitudes (D'Amicis et al., 2020).

Despite decades of study, solar wind–magnetosphere coupling remains incompletely understood. Several key physical processes are still insufficiently characterized for predictive modeling, including the detailed control of reconnection, plasma entry and mass loading pathways, and nonlinear feedbacks and saturation effects within the coupled system (Borovsky, 2021a). These limitations pose a challenge for space weather forecasting, particularly for quantities such as AE that respond rapidly to both large-scale driving and smaller-scale variability.

In this context, data-driven approaches provide a complementary avenue for improving geomagnetic activity forecasts. Modern machine learning methods can represent nonlinear dependencies without prescribing a fixed functional form and can integrate information from a large number of correlated input variables. A substantial body of work has therefore explored data-driven approaches to AE prediction, including early feed-forward and recurrent neural networks, dynamic neural networks, and nonlinear system identification techniques (Gleisner & Lundstedt, 2001; Gu et al., 2019; Pallocchia et al., 2008; Takalo & Timonen, 1997; Weigel et al., 1999). These studies demonstrate that meaningful predictive skill can be obtained from L1/near-Earth solar wind measurements, but they also highlight persistent limitations: forecast skill can vary strongly between quiet and active intervals, and models trained primarily on mean or coarsely average solar wind parameters often struggle to capture rapid intensifications and sharp gradients. More broadly, machine learning methods are increasingly central to space weather nowcasting and forecasting, offering flexible nonlinear function approximation while raising new questions about robustness, generalization, and physical interpretability (Camporeale, 2019).

Previous studies of AE forecasting span several generations of modeling approaches and provide quantitative context for evaluating current performance. Early empirical, linear, and neural network-based models demonstrated that statistically significant prediction skill is achievable, particularly at short lead times, but also revealed clear limitations. Dynamic and recurrent neural network models driven by upstream solar wind parameters typically achieved correlation coefficients of order 0.6–0.7 for lead times of tens of minutes to around 1 hour, with systematic underestimation of peak AE values and reduced skill during highly disturbed intervals (Gleisner & Lundstedt, 2001; Takalo & Timonen, 1997; Weigel et al., 1999). Empirical solar wind-driven models have reported higher correlations at short lead times, with correlations approaching 0.85–0.9 for 10 min averaged AE and root mean square errors of order ∼100 nT, though performance degrades as geomagnetic activity increases (Luo et al., 2013). More recent data-driven and system identification approaches have improved overall skill at hour-scale lead times, with correlations commonly reported in the range of ∼0.85–0.88 and prediction efficiencies of ∼0.7–0.75, while still exhibiting increasing error and bias during extreme events (Gu et al., 2019). Deep learning models developed in the past few years report comparable headline correlations for similar lead times, although error metrics are often reported for normalized quantities, complicating direct comparison with earlier studies, but can achieve correlations of greater than 0.85 (Zou et al., 2024). Taken together, these results suggest that, for solar wind-driven AE prediction, forecast skill has improved incrementally over time but appears to approach an upper bound set by internal magnetospheric dynamics, nonlinear coupling processes, and the influence of multiscale variability (Bargatze et al., 1985; Camporeale, 2019). This motivates continued investigation into whether additional physically meaningful descriptors of the upstream solar wind—beyond mean parameters alone—can provide further gains in forecast accuracy and robustness.

Comparable efforts have targeted other electrojet and auroral indices, reflecting both the diversity of available ground-based products and differing operational needs. Forecasting of the westward electrojet (AL) index has been addressed using neural network approaches (Amariutei & Ganushkina, 2012), while more recent work has applied machine learning to the SuperMAG SML index (Maimaiti et al., 2019), including using an imbalanced regression neural network (Chu et al., 2025), which predicts SML from solar wind inputs without relying on past index values and is specifically designed to recover strong to extreme events that conventional mean squared error training tends to underestimate. Regional indices derived from dense magnetometer networks, such as the IMAGE IL/IU/IE indices, have likewise been used to capture the spatio-temporal structure of high latitude activity (e.g., Tanskanen, 2009; Ventriglia et al., 2025). Across these indices, however, the upstream solar wind is typically represented through mean parameters or empirical coupling functions; the explicit turbulent nature of the driving flow is rarely incorporated as a predictor.

In this study we test the hypothesis that explicitly including turbulence information in the upstream drivers improves short-timescale AE forecasts beyond what is achievable with mean solar wind parameters alone. We

10.1029/2026SW005094

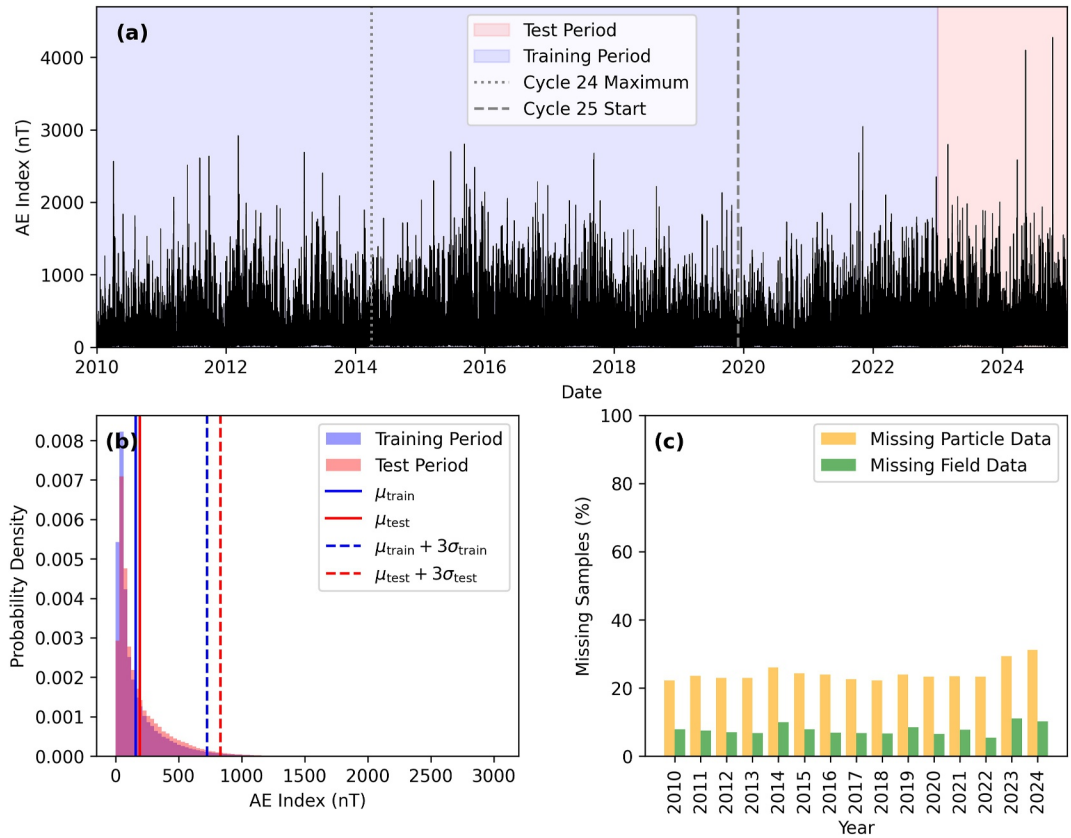

Figure 1. An overview of the period of OMNI data and auroral electrojet (AE) index used, with (a) the AE index over the entire period, with the training period highlighted in blue, the test period highlighted in red, and the maximum of solar cycle 24 and start of solar cycle 25 indicated by vertical lines, (b) histograms of AE index over both test periods, with means and means plus three standard deviations indicated, and (c) the percentage of missing data for both field and particle data per year.

build two gradient boosted decision tree models using XGBoost (T. Chen et al., 2026) to forecast AE from time-shifted near-Earth solar wind conditions. The first model uses a standard set of mean solar wind and IMF predictors (e.g., V, n, $|\mathbf{B}|$, IMF components). The second model augments these predictors with explicit turbulence descriptors computed from higher-cadence fluctuations over rolling windows, intended to capture the amplitude, intermittency, and Alfvénic structure of the solar wind (e.g., fluctuation levels with $\delta\mathbf{B}$ and $\delta\mathbf{V}$, and dimensionless measures motivated by MHD turbulence such as cross helicity and residual energy). Such quantities have been argued to modulate geoeffectivity and coupling efficiency even under comparable mean driving by altering the occurrence and effectiveness of reconnection and by injecting variability across magnetospheric response times (Borovsky, 2003; D'Amicis et al., 2020; Jankovičová et al., 2008). By comparing these models, we determine the added value of turbulence-aware models for AE prediction.

Section 2 describes the data set used for this study. The methods, both construction of the machine learning model and the parameters used, are described in Section 3, along with the evaluation methods. Section 4 uses these evaluation methods to determine the additional forecasting skill provided by a turbulence-aware model, and Section 5 uses interpretation methods to determine the physical drivers of this. Finally, Section 6 discusses the benefit for forecast users, Section 7 discusses these findings and future paths for space weather forecasts, and Section 8 concludes.

2. Data

The data set used spans January 2010 to December 2024, covering the majority of solar cycle 24 and the rising phase of solar cycle 25 (Figure 1a). This interval includes extended quiet periods and multiple episodes of extreme geomagnetic activity, providing a wide range of solar wind driving conditions.

15427390, 2026, 7, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2026SW005094 by NICE, National Institute for Health and Care Excellence, Wiley Online Library on [05/08/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

15427390, 2026, 7, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2026SW005094 by NICE, National Institute for Health and Care Excellence, Wiley Online Library on [03/08/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Solar wind plasma and IMF data are taken from the 1-min OMNI data set (Papitashvili & King, 2020), which combines measurements from multiple upstream spacecraft and propagates them to the Earth's bow shock nose. This propagation introduces a timing uncertainty, typically of order a few minutes, arising from the assumptions inherent in the phase front propagation technique (Case & Wild, 2012; King & Papitashvili, 2005; Weimer et al., 2003). This uncertainty is small compared with the timescales over which the fluctuations we examine evolve and the forecast lead time, and we therefore do not expect it to materially affect our results. We note that, because the prediction is keyed to conditions at the bow shock while the driving solar wind is observed upstream, there is a genuine lead time between input and predicted response. In this sense this constitutes short term forecasting rather than nowcasting, as the propagation delay represents a real predictive horizon. The variables used are the GSM IMF components (B_x, B_y, B_z), magnetic field strength $|\mathbf{B}|$, GSE solar wind velocity components (V_x, V_y, V_z), and proton number density n.

Geomagnetic activity is characterized using the 1-min AE index, a widely used proxy for AE intensity and substorm activity (Davis & Sugiura, 1966). The AE distribution is strongly non-Gaussian with a pronounced high-activity tail (Figure 1b), motivating conditional evaluation later in the analysis. It is worth noting that AE index is derived only from northern hemisphere data, and that asymmetries in the geomagnetic field mean the southern hemisphere response may not be identical.

The data are split chronologically to ensure temporal independence, with January 2010–December 2022 used for training and January 2023–December 2024 used for testing. This allows evaluation under previously unseen conditions during the rising phase of solar cycle 25 after training on a full solar cycle. Magnetic field data coverage exceeds 90% over the full interval, while plasma moments exhibit larger data gaps (Figure 1c). Short gaps of up to 3 minutes are forward-interpolated to preserved short timescale continuity, while longer gaps remain missing. No interpolation is applied to the AE index, which is used only where observed.

3. Methods

3.1. Forecast Formulation and Feature Construction

The AE index is treated as the prediction target, with forecasts defined at fixed lead times $\tau \in \{30, 45, 60, 75, 90, 120, 180\}$ minutes. All input features at time t are used to predict $AE(t + \tau)$, ensuring a strictly causal formulation.

To capture the multiscale nature of solar wind driving, rolling statistics are computed over multiple window lengths. For bulk solar wind parameters, rolling means are calculated over windows of 5–120 min. For turbulence diagnostics, statistics are computed over windows of 5–60 min. A minimum data availability criterion of 50% within each window is imposed to avoid spurious values due to sparse sampling.

Turbulence features are derived from fluctuations defined relative to the rolling mean background. From these fluctuations, RMS amplitudes, higher order moments (skewness and kurtosis), and physically motivated turbulence parameters are computed. These include the normalised cross helicity and residual energy (e.g., C. H. K. Chen, 2016; Schekochihin, 2022),

$$\sigma_c(W) = \frac{2\langle \delta\mathbf{v} \cdot \delta\mathbf{b} \rangle_W}{\langle |\delta\mathbf{v}|^2 \rangle_W + \langle |\delta\mathbf{b}|^2 \rangle_W}, \tag{1}$$

$$\sigma_r(W) = \frac{\langle |\delta\mathbf{v}|^2 \rangle_W - \langle |\delta\mathbf{b}|^2 \rangle_W}{\langle |\delta\mathbf{v}|^2 \rangle_W + \langle |\delta\mathbf{b}|^2 \rangle_W}, \tag{2}$$

which both make use of the magnetic field normalized by the local Alfvén speed,

$$\mathbf{b}(t) = \frac{\mathbf{B}(t)}{\sqrt{\mu_0 m_p n(t)}}, \tag{3}$$

where n is the proton number density, m_p is the proton mass, and μ_0 is the permeability of free space. σ_c of ± 1 indicates unidirectional Alfvénic fluctuations, whereas 0 indicates balanced turbulence, and $\sigma_r < 0$ indicates

magnetically dominated turbulence, whereas $\sigma_r > 0$ indicates kinetically dominated turbulence. The final parameter derived is the magnetic compressibility,

$$C_B(W) = \frac{\langle \delta |\mathbf{B}|^2 \rangle_W}{\langle \delta B_x^2 + \delta B_y^2 + \delta B_z^2 \rangle_W}, \tag{4}$$

which quantifies the fraction of magnetic energy fluctuation in compressive modes. Together, these quantities describe the amplitude, intermittency, geometry, and energy balance of solar wind turbulence.

3.2. Forecast Models

Forecasts are produced using gradient-boosted decision trees implemented with XGBoost (T. Chen et al., 2026). In gradient boosting, an ensemble model is constructed by sequentially adding many shallow decision trees, where each new tree is trained to reduce the prediction errors (residuals) of the existing ensemble. This stage-wise optimisation corresponds to performing gradient descent in function space on a specified loss function, here the squared error objective. The final forecast is obtained as the weighted sum of the predictions from all trees, allowing the model to represent complex nonlinear relationships and interactions between input variables.

XGBoost (Extreme Gradient Boosting) is an efficient and regularized implementation of this approach. It incorporates shrinkage (learning rate control), tree depth constraints, and explicit penalty terms on model complexity to reduce overfitting, while enabling fast parallel training on large data sets. These properties make gradient-boosted trees well suited to space weather forecasting problems, where the solar wind–magnetosphere coupling exhibits nonlinear responses, threshold behavior, and interactions across multiple temporal scales. In addition, tree-based models naturally accommodate features with different units and dynamic ranges without requiring strong distributional assumptions, and permit post-hoc interpretation through feature attribution methods.

These properties come with limitations worth noting. Tree ensembles cannot extrapolate beyond the range of target values present in the training data, since each prediction is an average over training samples within a leaf; the rarest and most extreme AE intensifications are therefore prone to systematic under-prediction. Gradient-boosted trees also have no intrinsic representation of temporal structure and treat each sample independently, so any sequential information must be provided explicitly through engineered predictors, such as the rolling window diagnostics used in this study.

Models are trained using a squared-error objective with regularization to limit overfitting. Three model configurations are considered:

1. A baseline model using only large scale solar wind parameters and AE history;
2. A turbulence-aware model that augments the baseline with solar wind turbulence diagnostics; and
3. A noise-control model in which turbulence features are replaced by statistically matched random surrogates, to isolate the contribution of providing a greater number of features alone.

3.3. Training, Validation, and Optimization

All models are trained using a strictly time-ordered procedure to reflect the real forecasting setting. In particular, data from later times are never used to inform predictions at earlier times. This prevents "information leakage," in which knowledge of the future could artificially improve model performance.

Model settings (hyperparameters), such as the complexity of the decision trees and the rate at which the model learns from new information, are automatically tuned using the Optuna optimisation framework (Akiba et al., 2019). To do this, the training interval is divided into two parts: an initial portion used to fit the model, and a final 20% subset used for validation. Model performance on this validation subset provides an objective way to select hyperparameters that generalize well to unseen data.

During training, early stopping is applied. This means that model fitting is halted if performance on the validation data no longer improves, helping to avoid overfitting, where the model learns noise or event-specific details rather than robust physical relationships.

15427390, 2026, 7, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2026SW005094 by NICE, National Institute for Health and Care Excellence, Wiley Online Library on [03/08/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

15427390, 2026, 7, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2026SW005094 by NICE, National Institute for Health and Care Excellence, Wiley Online Library on [03/08/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Table 1
Tuned 60-Minute Forecast Horizon Model Hyperparameters

Parameter	Base model	Turbulence-aware model
η	0.0182	0.0118
Max. depth	4	5
Subsample	0.423	0.373
Column sample by tree	0.830	0.956
Min. child weight	0.152	6.48
γ	3.66	0.140
λ	0.492	0.669

After the optimal hyperparameters are identified, the final models are retrained using the full training interval to maximize the information available for learning before evaluation on independent test data. Automated hyperparameter tuning frameworks such as Optuna (Akiba et al., 2019), together with validation monitoring and early stopping, are widely used and represent established best practice in modern machine learning workflows.

The final hyperparameters for the base and turbulence-aware models forecasting with a 60 min horizon obtained from tuning are shown in Table 1. The learning rate η sets the shrinkage applied to each tree's contribution; smaller values slow the learning, requiring more trees but typically improving generalization. The maximum depth limits how many splits each tree can make, bounding the order of feature interactions the model can represent and acting as a primary control on model complexity. Subsample is the fraction of training rows drawn at random for each tree, while column sample by tree is the fraction of features made available to each tree; both introduce stochasticity that reduces overfitting and decorrelates the ensemble. The minimum child weight sets the smallest total instance weight permitted in a leaf, so that larger values prevent the model from forming leaves that fit only a few samples. The parameter γ is the minimum loss reduction required to make a further split, so higher values yield more conservative, less complex trees. Finally, γ is the L2 regularization term on the leaf weights, penalizing large leaf values to further discourage overfitting.

3.4. Model Evaluation and Explainability

Model performance is evaluated on the test interval using RMSE, mean absolute error (MAE), R^2, and Pearson correlation. A persistence forecast $AE(t)$ is used as a baseline, and a persistence skill score is computed to quantify improvement,

$$S = 1 - \frac{\mathrm{MSE}_{\mathrm{model}}}{\mathrm{MSE}_{\mathrm{persist}}}, \tag{5}$$

where $S > 0$ indicates improvement over persistence.

Model behavior is interpreted using Shapley Additive Explanations (SHAP) (S. Lundberg & Lee, 2017). Global feature importance is quantified using mean absolute SHAP values, while dependence plots are used to examine nonlinear responses. Because solar wind features are strongly correlated, a correlation-aware SHAP formulation is employed (Aas et al., 2020), preserving empirical feature dependencies and yielding physically meaningful attributions.

SHAP values are computed on a stratified subset of the test data to ensure representation across geomagnetic activity levels. Because the AE distribution is heavily weighted toward quiet time values, uniform random sampling of the test set would yield few high activity intervals and provide poor coverage of the large AE regime that is of greatest interest. To address this, the test set is partitioned into bins defined by quantiles of the observed AE distribution, with finer bins at higher activity levels, and a fixed number of samples is drawn at random without replacement from each bin. The quantile ranges with their sample numbers are as follows: Q0–50 with 1,500 min, Q50–60 with 1,000 min, Q60–80 with 800 min, Q80–90 with 500 min, and Q90–100 with 300 min. This over represents active intervals relative to their natural frequency while retaining a substantial quiet time population, so that the resulting attributions are not dominated by low activity samples and the model's behavior during disturbed conditions is adequately characterized.

3.5. Cost/Loss Analysis

To evaluate whether the turbulence-aware forecast adds forecast value in comparison to the mean-values forecast, the application to operational space weather forecasting must be considered. This is carried out using cost/loss analysis as described in Richardson (2011) and Owens and Riley (2017), taking into account different thresholds of AE index at which mitigating action would be taken.

An event is defined in relation to a threshold T,

$$E_t = \begin{cases} 1, & AE_{\text{obs}}(t) \geq T, \\ 0, & \text{otherwise.} \end{cases} \tag{6}$$

As the forecast is deterministic (has no probability associated with a prediction), the probability of exceeding a threshold is

$$p_t = \begin{cases} 1, & \hat{AE}(t) \geq T, \\ 0, & \text{otherwise.} \end{cases} \tag{7}$$

The event rate for climatology is used to compare the benefit added by a model. In this case, this is taken as the long-term fraction of time that the chosen AE threshold is exceeded,

$$o = \mathbb{P}(E_t = 1) = \frac{1}{N}\sum_{t=1}^{N} E_t. \tag{8}$$

For a user with a cost/loss ratio $r = C/L \in [0,1]$, where C is the cost of mitigating for an event and L is the loss associated with an event, they will take mitigating action ($A_t = 1$) when $p_t = 1$ and not ($A_t = 0$) when $p_t = 0$. This results in expenses associated with the forecast, climatology, and a perfect prediction,

$$E_{\text{fcst}} = C\sum_{t=1}^{N} A_t + L\sum_{t=1}^{N} E_t(1 - A_t), \tag{9}$$

$$E_{\text{clim}} = \begin{cases} CN, & o \geq r, \\ L\sum_{t=1}^{N} E_t, & o < r, \end{cases} \tag{10}$$

$$E_{\text{perf}} = C\sum_{t=1}^{N} E_t. \tag{11}$$

The potential economic value of the forecast can then be found from these,

$$V = \frac{E_{\text{clim}} - E_{\text{fcst}}}{E_{\text{clim}} - E_{\text{perf}}}, \tag{12}$$

which only depends on the ratio of cost and loss r. This allows for comparison of the potential economic value between different forecast users. $V = 1$ indicates a perfect forecast, $V = 0$ indicates no improvement over climatology, and $V < 0$ indicates that a forecast is worse than climatology when considering economic benefit.

4. AE Prediction Performance

A commonly used performance metric for assessing prediction of AE index is the correlation coefficient between the predicted and observed AE index. However, to compare between different target lags, this will almost always show a peak at the shortest lag due to information decay. Instead, the correlation is considered along with skill versus persistence, as shown in Figure 2a. For the base model, the skill peaks at a 75 min lag window. In contrast, for the turbulence-aware model, the skill is sustained across multiple forecast horizons and is approximately equal for forecast horizons of 60, 75 and 90 min, indicating improved robustness to increasing lead time. As in both models the correlation is greatest at the shorter horizon, the models predicting 60 min into the future are chosen for further analysis. It is worth noting that the model replacing turbulence parameters with noise performs approximately equally to the base model, demonstrating that the improvement arises from physically meaningful structure in the turbulence parameters rather than simply increasing the number of inputs. Overall at 60 min, the correlation coefficient exceeds 0.8, only slightly lower than values reported for more complex neural network methods.

15427390, 2026, 7, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2026SW005094 by NICE, National Institute for Health and Care Excellence, Wiley Online Library on [03/08/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

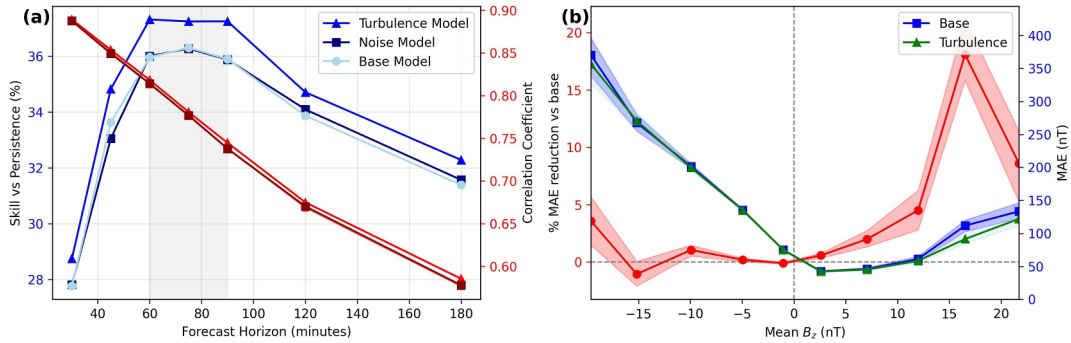

Figure 2. (a) Comparison of the performance of each of the base, turbulence and noise models for each of the chosen forecast horizons, with skill versus persistence shown in blue and correlation coefficient in red, with triangular markers for the turbulence-aware model, squares for the noise model, and circles for the base model. The range of horizons giving high quality predictions of auroral electrojet index are shown by a gray shaded region. (b) Comparison of the mean absolute error between the base (blue squares) and turbulence (green triangles) models for different values of B_z averaged over the previous hour, with the binned percentage reduction by the turbulence model shown in red along with the range of values for each bin.

While short-term AE persistence contributes substantially to forecast skill, the turbulence-aware model provides a consistent and measurable improvement over both persistence and the base model. This demonstrates that the model is exploiting structured information in the upstream solar wind beyond simple autoregressive behavior, and that turbulence descriptors provide additional predictive power for short timescale geomagnetic activity.

Motivated by the different effects of solar wind driving under southward and northward IMF, the improvement made by the turbulence-aware model is considered for different values of B_z in GSM coordinates (Figure 2b). For southward B_z, there is little change to the MAE between the two models, except for a small improvement at very large southward IMF. However, under northward IMF, the greater the magnitude of B_z, the greater the percentage improvement in MAE obtained by the turbulence model. This suggests that turbulence plays a greater role in modulating the effect of solar wind driving on AE index under northward IMF.

The improvement provided by the turbulence-aware model can also be seen to vary with the magnitude of AE index (Figure 3a). When plotting the observed versus predicted AE index, the turbulence model is more successful at following the $AE_{obs} = AE_{pred}$ line. The base model falls below this line for the higher AE values, indicating that it is overpredicting for more extreme events. This could result in the turbulence-aware model reducing false alarm rates for extreme space weather events. In Figure 3b, this is shown further by binning the AE index into bands by percentile. For lower values of AE index, the MAE is approximately the same for both

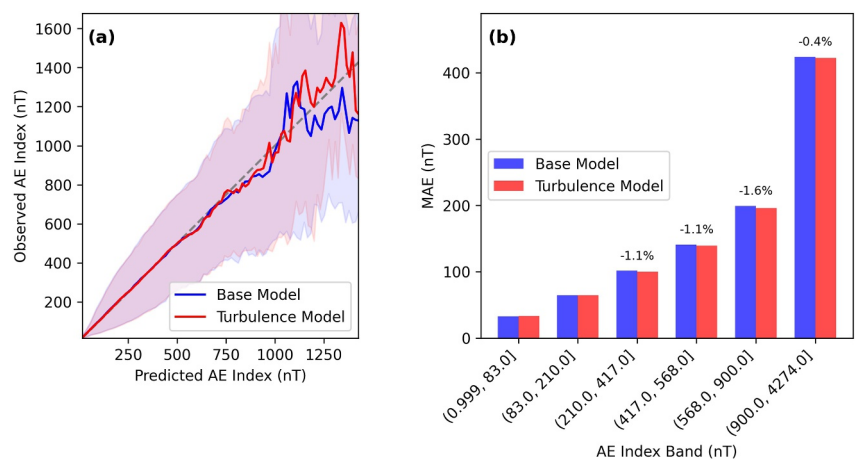

Figure 3. (a) Observed versus predicted auroral electrojet (AE) index by the base model (blue) and turbulence model (red), binned in predicted AE index, with $AE_{obs} = AE_{pred}$ shown as a gray dashed line. The associated 10th and 90th percentiles are shown for each model as a confidence interval. (b) Average mean absolute error for both the base (blue) and turbulence (red) models for bands of AE index, defined by the 0th, 50th, 75th, 90th, 95th, 99th, and 100th percentiles as their edges.

15427390, 2026, 7, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2026SW005094 by NICE, National Institute for Health and Care Excellence, Wiley Online Library on [03/08/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

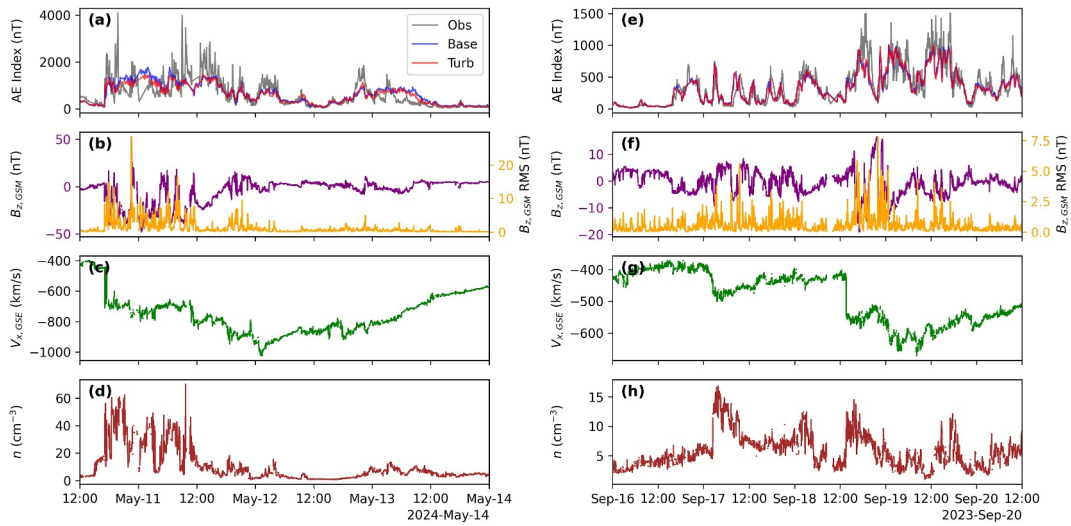

Figure 4. Two example space weather events with (a, e) observed auroral electrojet (AE) index in gray, predicted (by the base model) AE index in blue, and predicted (by the turbulence model) AE index in red, (b, f) z_{GSM} component of magnetic field $B_{z,GSM}$ in nT in purple and RMS fluctuations in this component $\delta B_{z,rms}$ in yellow, (c, g) x_{GSE} component of solar wind velocity $V_{x,GSE}$ (km/s) in green, and (d, h) proton number density n (cm^{-3}) in brown, for (a–d) 11 to 14 May 2024 and (e–h) 16 to 20 September 2023.

models. However, in the higher bands (75th percentile and above), the MAE is reduced by the turbulence model, by over 1% between the 75th and 99th percentiles, and by 0.4% above the 99th percentile. This indicates that overall the turbulence-aware model can predict more extreme events effectively, up until those that are the most extreme, where it does not give as much of an improvement. This is expected, as the most extreme space weather events are likely to have additional physical effects at play which are not described adequately by the training data, as the test data contains the two largest spikes in AE of the whole data period.

This can be seen by considering example space weather events. Figure 4 shows two events, with one being the extreme storm of May 2024 (Figures 4a–4d), and the other a more typical event from September 2023 (Figures 4e–4h). For May 2024, the timing of increases and decreases in AE index is predicted well (Figure 4a). However, the magnitude of AE index is consistently underestimated. This is most likely due to the additional physical effects known to be occurring during this event due to multiple CME impacts (e.g., MacTaggart et al., 2025; Tulasi Ram et al., 2024) and additional flips in B_y (Ohtani et al., 2025), causing a strongly driven and compressed magnetopause (Beedle et al., 2025) and leading to additional ionospheric structure (e.g., Astafyeva et al., 2025; Kuai et al., 2025). The multiple impacts can be seen in B_z in Figure 4b, often with large $\delta B_{z,rms}$ associated with them, along with solar wind speed increases (Figure 4c) and density enhancements (Figure 4d).

September 2023 provides a view of a much more "typical" event. Figure 4e shows that the predicted AE index corresponds much more closely in magnitude here, along with the timing of increases in decreases being predicted well. This event had a mixture of northward and southward IMF, with high $\delta B_{z,rms}$ corresponding to additional structures within the solar wind (Figure 4f). This event also exhibits slow wind until around 12:00 on September 18th, before a sharp increase in wind speed with the impact of a structure (Figure 4g), and shows density enhancements corresponding to these structures (Figure 4h). Overall this demonstrates that the prediction works well for a more commonly occurring space weather event.

As previously discussed, comparing absolute error in prediction between events does not give the full picture due to their relative magnitudes in AE index. However, for these two events, we instead normalize by the median AE index to compare the root mean squared error. For May 2024, this median is 406 nT, and for September 2023, this is 302 nT. In May 2024, the base model normalized RMSE is 0.99, and the turbulence-aware model normalized RMSE is 0.94, resulting in a reduction of 0.05 times the median AE index. However, in September 2023, the base normalized RMSE is 0.52, and the turbulence-aware model normalized RMSE is 0.52, staying largely the same. This corresponds with it being forecasts for the most extreme events which are improved by using a turbulence-aware model, as discussed in Figure 3.

15427390, 2026, 7, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2026SW005094 by NICE, National Institute for Health and Care Excellence, Wiley Online Library on [03/08/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

10.1029/2026SW005094

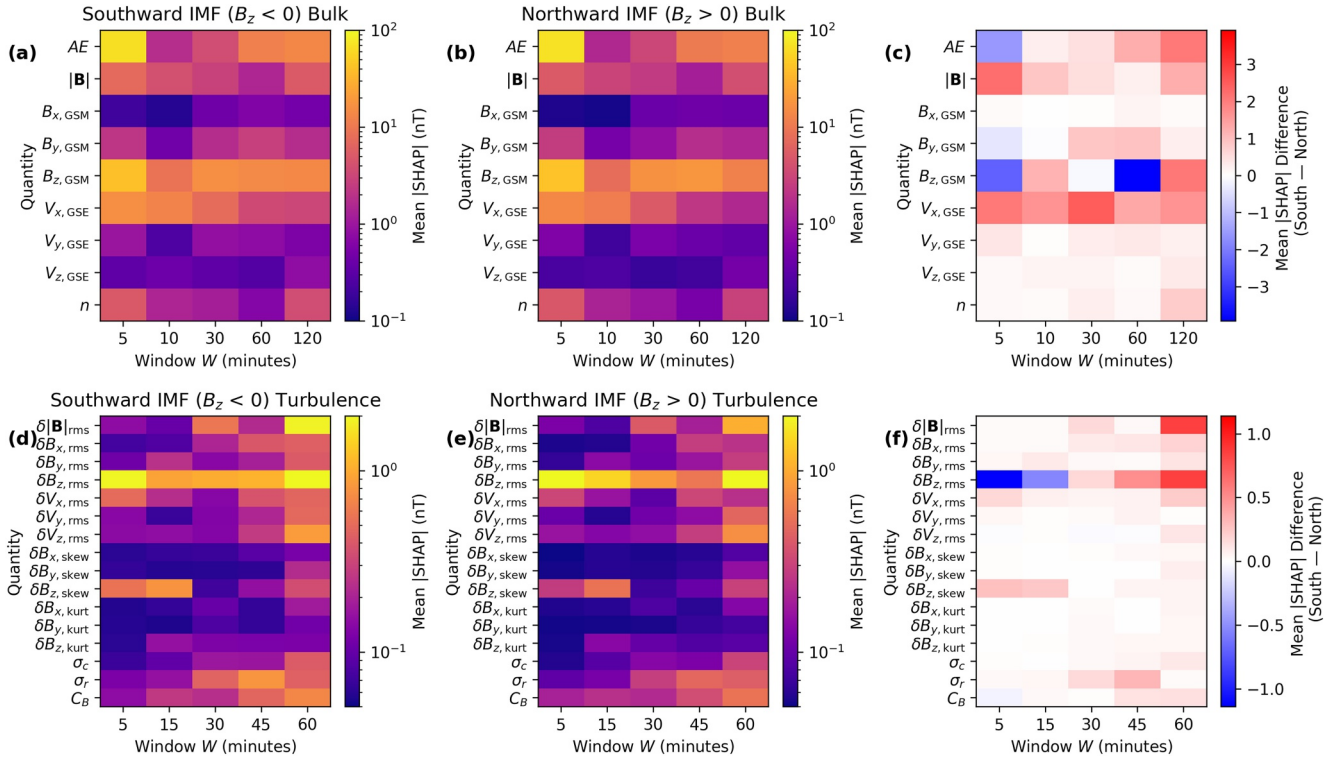

Figure 5. (a, b, d, and e) Mean of the magnitudes of the Shapley Additive Explanations values in the turbulence-aware model for each of the features shown as a heatmap, organized by physical variable and window W, for (a) bulk parameters under southward interplanetary magnetic field (IMF), (b) bulk parameters under northward IMF, (d) turbulence parameters under southward IMF, and (e) turbulence parameters under northward IMF. Panel (c) shows the difference in these values (southward–northward) for the bulk parameters, and (f) for the turbulence parameters.

Overall, explicitly including turbulence parameters within the upstream solar wind data set to predict AE index can improve the prediction skill and reduce errors in prediction of extreme space weather events. The percentage error reduction is greatest under northward IMF, suggesting that turbulence influences the effects of solar wind driving on AE index under northward IMF more strongly than under southward IMF. Although the percentage reduction in error is modest, the improvement is systematic, physically interpretable, and concentrated in regimes where baseline models perform worst—namely under northward IMF and during elevated AE activity—and therefore has disproportionate value for both scientific understanding and operational decision making.

5. Parameter Dependence

5.1. Average Contributions

Correlation-aware SHAP is used to quantify the mean contribution of each input variable to the turbulence-aware AE prediction. Contributions are separated by IMF orientation (Figure 5), reflecting differences in model behavior under northward and southward conditions.

For bulk parameters (Figures 5a and 5b), the dominant predictor is the 5-min mean AE index, capturing short-term persistence of AE activity (Akasofu, 1964; McPherron, 1970). Intermediate windows (10–30 min) add little additional information, while contributions increase again at 60–120 min, indicating sensitivity to longer time-scale clustering of geomagnetic activity (Pulkkinen et al., 2006; Tanskanen, 2009). The next most important parameter is the 5-min mean $B_{z,\mathrm{GSM}}$, consistent with its role in controlling dayside reconnection and energy input (Dungey, 1961; Gonzalez et al., 1994). Solar wind speed $V_{x,\mathrm{GSE}}$ contributes primarily at short lags, with diminishing importance at longer times due to autocorrelation and reduced relevance for longer term prediction. Other bulk parameters, including $|\mathbf{B}|$ and n, provide secondary contributions, while B_x, B_y, and transverse velocity components have minimal impact.

15427390, 2026, 7, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2026SW005094 by NICE, National Institute for Health and Care Excellence, Wiley Online Library on [03/08/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

15427390, 2026, 7, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2026SW005094 by NICE, National Institute for Health and Care Excellence, Wiley Online Library on [03/08/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

The IMF dependent difference (Figure 5c) show a clear separation in controlling mechanisms. Under northward IMF, AE persistence and short timescale B_z variations are more important, indicating a more gradual and variability-driven response. Under southward IMF, solar wind speed and $|\mathbf{B}|$ become more influential, consistent with stronger driving and enhanced coupling efficiency. Longer timescale AE history (120 min) is also more important under southward IMF, reflecting sustained or recurrent activity during prolonged driving.

For turbulence parameters, the dominant contributions arise from magnetic field fluctuations and their higher order structure (Figures 5d and 5e). RMS fluctuations in $B_{z,\mathrm{GSM}}$ are the most important turbulence predictors across multiple timescales, with strong contributions at both short (5 min) and long (60 min) windows. This indicates sensitivity to both rapid IMF variability and longer timescale fluctuations controlling coupling efficiency. RMS fluctuations in $|\mathbf{B}|$ at 60 min are similarly important, highlighting the role of turbulent magnetic energy. Higher order statistics also contribute, with skewness of B_z fluctuations important at short timescales, consistent with the impact of intermittent southward excursions.

At longer timescales, parameters describing turbulence structure become more important. Residual energy σ_r contributes for windows $\geq$30 min, indicating sensitivity to magnetic–kinetic energy imbalance, while compressibility C_B becomes important at 45–60 min, suggesting that fluctuation geometry influences coupling. Velocity fluctuations play a secondary role, with only weak contributions at long timescales.

The IMF dependence of turbulence contributions is systematic. Short timescale B_z fluctuations are more important under northward IMF, indicating enhanced sensitivity to directional variability when coupling is weak. In contrast, longer timescale fluctuations, magnetic energy content, and structural parameters (σ_r, C_B) are more important under southward IMF, reflecting cumulative turbulent driving under strong coupling. Overall, these results show that turbulence influences AE through different mechanisms depending on IMF orientation: short timescale variability dominates under weak driving, while longer timescale amplitude and structure control the response under strong driving.

5.2. Behavior of Bulk Features

Figures 6a, 6d, 6e, and 6f shows the SHAP dependence of the most influential bulk solar wind and IMF parameters. At short lead times (5–60 min), the model response is dominated by IMF B_z, solar wind speed V_x, and magnetic field magnitude $|\mathbf{B}|$, consistent with established solar wind–magnetosphere coupling physics (Borovsky, 2003; Luo et al., 2013; Newell et al., 2007). At longer lead times (120 min), these dependencies weaken or change sign, reflecting the increasing influence of internal magnetospheric dynamics and the reduced direct control of upstream conditions (Weigel et al., 2003).

IMF B_z (Figure 6a) shows the expected strong sensitivity to southward conditions at short lead times, with increasingly negative B_z producing large positive contributions to AE. For strong southward IMF ($B_z \lesssim -10$ nT), the response saturates, indicating that additional increases in driving do not produce proportionally larger AE. This plateau is consistent with known nonlinear limits in magnetospheric coupling, including ionospheric conductance constraints and feedback processes (Borovsky, 2021b; Myllys et al., 2016; Siscoe et al., 2002). As lead time increases, the magnitude of the B_z dependence decreases, and by 120 min the relationship reverses, with strong southward IMF associated with reduced AE. This reflects the transition from directly driven behavior to phase-dependent evolution, where strong driving is followed by recovery or unloading (Borovsky, 2021b; Weigel et al., 2003).

Solar wind speed V_x (Figure 6d) exhibits a monotonic dependence at short lead times, with faster flows producing larger positive contributions, consistent with enhanced convective electric field driving (Luo et al., 2013; Newell et al., 2007). This dependence weakens steadily with increasing lead time and is negligible by 120 min, indicating that solar wind speed primarily controls short-term energy input rather than longer term AE evolution.

Proton density n (Figure 6e) has a weaker and more transient influence. At short lead times, increasing density produces positive contributions, consistent with enhanced dynamic pressure and magnetopause compression. However, this dependence collapses at longer lead times and becomes non-monotonic, indicating limited predictive value beyond short timescales (Borovsky, 2021b; Weigel et al., 2003).

Magnetic field magnitude $|\mathbf{B}|$ (Figure 6f) shows a threshold-like response at short lead times, with low values associated with reduced AE and a transition to positive contributions above ~5–10 nT. At higher field strengths,

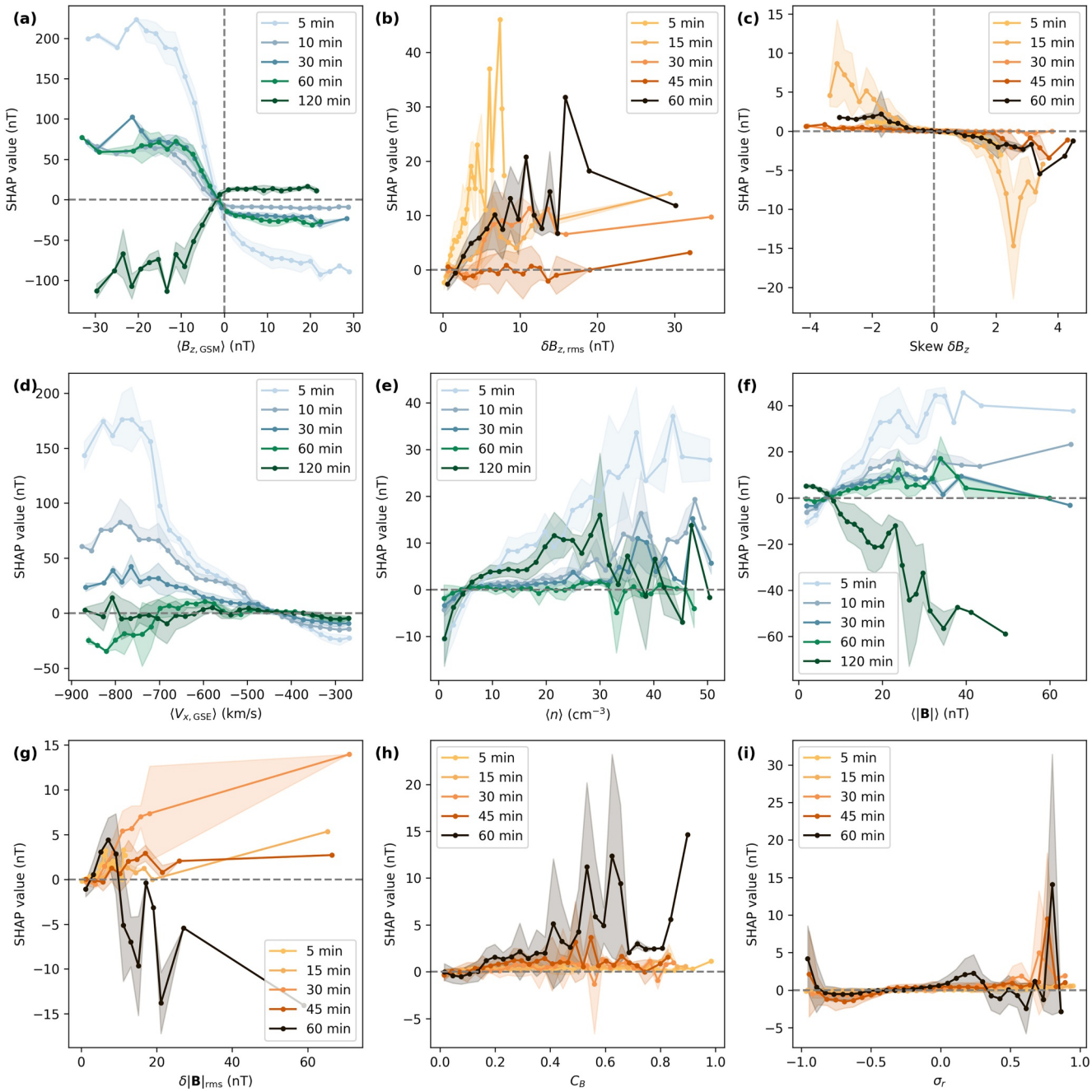

Figure 6. Median values of Shapley Additive Explanations value for bins of (a) z component of magnetic field $B_{z,\mathrm{GSM}}$, (b) RMS fluctuations of B_z $\delta B_{z,\mathrm{rms}}$, (c) skew of δB_z, (d) x component of velocity $V_{x,\mathrm{GSE}}$, (e) proton number density n, (f) magnetic field strength $|\mathbf{B}|$, (g) RMS fluctuations of field strength $\delta|\mathbf{B}|_{\mathrm{rms}}$, (h) magnetic compressibility C_B, and (i) residual energy σ_r. Bulk parameters (a, d, e, f) are shown windowed over 5, 10, 30, 60 and 120 min, and turbulence parameters (b, c, g, h, i) are shown windowed over 5, 15, 30, 45 and 60 min. The median values $\pm\sigma$ are shown.

the response flattens, indicating diminishing sensitivity under strong driving. As with B_z, the dependence weakens and reverses at longer lead times, reflecting the reduced predictive role of instantaneous upstream conditions and the increasing importance of internal dynamics.

Overall, these results show that AE predictability at short lead times is controlled by IMF orientation and solar wind driving in a manner consistent with established coupling physics. At longer lead times, the influence of these drivers diminishes and, in some cases, reverses, highlighting the transition to internally governed magnetospheric dynamics and setting a fundamental limit on solar wind-driven predictability.

15427390, 2026, 7, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2026SW005094 by NICE, National Institute for Health and Care Excellence, Wiley Online Library on [03/08/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

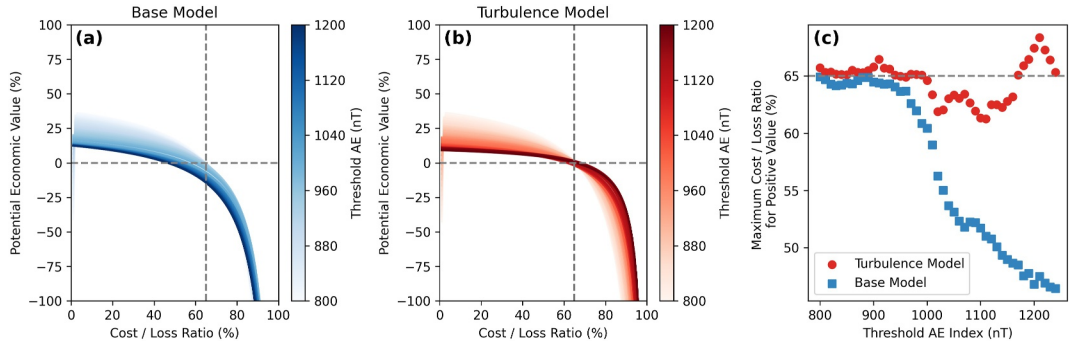

Figure 7. Potential economic value V against cost/loss ratio r for the base model in (a) and turbulence model in (b), for a range of thresholds of auroral electrojet (AE) between 800 and 1,200 nT. (c) The x-intercept of each of the curves plotted against the threshold AE index for the turbulence model (red circles) and the base model (blue squares).

5.3. Behavior of Turbulence Features

Figures 6b, 6c, and 6g–6i shows the SHAP dependence of turbulence parameters, demonstrating how fluctuation amplitude, intermittency, and structure influence the predicted AE index. The dominant turbulence signal arises from magnetic field variability. RMS fluctuations in B_z exhibit a strong positive contribution at short timescales (5 min), with continued but weaker influence at longer windows (15–60 min). This indicates that both rapid directional variability and longer timescale fluctuations in IMF orientation enhance geomagnetic response, consistent with the central role of B_z in controlling coupling.

In contrast, RMS fluctuations in the total field $|\mathbf{B}|$ show a weaker and more scale-dependent effect. While intermediate timescales (∼30 min) show a modest positive contribution, large amplitude fluctuations at longer timescales (60 min) produce negative SHAP values, indicating that strong magnetic variability does not necessarily translate into increased geomagnetic activity. This suggests a regime in which increased fluctuation amplitude reflects less efficient or more disorganised coupling rather than stronger driving.

Higher order and structural turbulence parameters exhibit more selective behavior. The skewness of B_z fluctuations is most influential at intermediate timescales (∼15 min), with negative skewness (intermittent southward excursions) producing positive contributions and positive skewness producing negative contributions. This highlights the importance of intermittent, asymmetric fluctuations rather than purely Gaussian variability. At shorter and longer windows, the influence of skewness is weak, indicating that intermittency is most effectively captured at intermediate timescales.

Parameters describing turbulence structure and energy partition become important at longer timescales. Residual energy σ_r shows a strong positive contribution as it approaches unity, indicating sensitivity to magnetic–kinetic imbalance in the fluctuations. Compressibility C_B also becomes important at 45–60 min, with peak contributions at moderate values (∼0.5) and reduced influence at very high compressibility, suggesting that moderately structured turbulence is most effective at enhancing geomagnetic response.

Overall, these results show that turbulence contributes to AE prediction through distinct physical pathways across scales. Short timescale variability in B_z enhances immediate coupling, intermittency at intermediate scales modulates the occurrence of geoeffective fluctuations, and longer timescale properties such as energy partition and fluctuation geometry influence the cumulative efficiency of energy transfer. Importantly, these effects persist even when mean IMF conditions are included, indicating that turbulence provides complementary predictive information beyond large scale solar wind driving.

6. Model Benefits (Cost/Loss Analysis)

Figures 7a and 7b present the potential economic value of the AE forecasts as a function of cost-loss ratio for a range of AE index thresholds. Following the cost-loss framework defined in Section 3, the potential economic value V is shown as a percentage, where 100% denotes the value of a perfect forecast and 0% denotes no improvement over climatology. Negative values indicate a forecast less useful than climatology. Recall that V depends only on the user's cost-loss ratio $r = C/L$, so each curve describes how the forecast's value varies across

15427390, 2026, 7, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2026SW005094 by NICE, National Institute for Health and Care Excellence, Wiley Online Library on [03/08/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

users with different relative costs of taking mitigating action. The cost-loss ratio at which a curve crosses zero therefore marks the boundary beyond which climatology is the better basis for action, and the span of r over which the curve is positive indicates the breath of users for whom the forecast is beneficially beneficial.

For the base model, the potential economic value systematically decreases as the AE threshold increases, with the value curves shifting downward and the range of cost-loss ratios for which the forecast provides positive value narrowing. In particular, the cost-loss ratio at which the potential value falls to zero decreases monotonically with increasing threshold, indicating that the base model becomes progressively less useful for decision making as the focus shifts to more extreme AE events. This behavior reflects a deterioration in the model's ability to provide reliable discrimination between events and non-events at higher thresholds, consistent with increasing false alarms or missed events under more stringent conditions.

In contrast, the turbulence-augmented model exhibits markedly different behavior. While the overall magnitude of the potential value varies with threshold, the cost-loss ratio at which the value drops to zero remains approximately constant across all AE thresholds considered. This indicates that the turbulence model retains economic usefulness over a similar range of decision maker preferences, even as the event definition becomes more extreme. Figure 7c highlights this distinction: the zero-value intercept for the turbulence model is largely invariant, whereas the intercept for the base model decreases steadily with increasing threshold. These results demonstrate that incorporating turbulence parameters stabilizes forecast value across event severities, implying improved robustness of the forecast for high impact, low probability events. From a decision making perspective, this suggests that the turbulence model maintains practical utility for mitigation strategies targeting extreme geomagnetic activity, whereas the base model's usefulness rapidly degrades as event severity increases.

7. Discussion

This study shows that explicitly incorporating turbulence-related descriptors into a data driven AE forecasting model produces systematic and physically interpretable improvements over a baseline model that uses only mean solar wind and IMF parameters. Although the absolute reduction in global error metrics is modest, the improvements are concentrated in regimes where baseline performance is weakest, including periods of northward IMF and elevated AE activity. Importantly, these gains extend beyond average error reduction to operationally relevant measures, such as reduced false alarms and improve cost-loss performance for high impact events.

Interpretation of both bulk and turbulence features using SHAP is broadly consistent with established solar wind–magnetosphere coupling physics. Short-lead forecasts are dominated by IMF orientation, solar wind speed, and magnetic field strength, while turbulence descriptors provide additional information related to variability, intermittency, and Alfvénic structure that is not captured by mean parameters alone. At longer lead times, the reduced influence of upstream drivers and the emergence of sign reversals reflect increasing control by internal magnetospheric dynamics rather than deficiencies in the model.

The improved performance of the turbulence-aware model supports the view that solar wind fluctuations play an active role in regulating magnetospheric energy transfer, rather than acting solely as noise around a mean driving state. Turbulence parameters encode information about the structure and coherence of the upstream plasma, which can influence reconnection efficiency, plasma transport, and the timing of energy release in the magnetosphere.

Previous observational studies have shown that Alfvénic turbulence can enhance auroral activity even under weak mean driving, while intermittent fluctuations may locally increase reconnection rates. The present results are consistent with this picture: turbulence descriptors retain predictive influence even when mean IMF strength and orientation are included, indicating that they encode complementary information rather than simply acting as proxies for stronger large scale driving. The fact that improvements are most pronounced under northward IMF further supports the interpretation that variability becomes particularly important when steady reconnection is suppressed.

Forecast improvements are strongly regime dependent. At short lead times, turbulence-aware forecasts better capture rapid AE enhancements and elevated activity levels, while at longer lead times the additional information becomes less influential as internal magnetospheric dynamics dominate. This behavior is consistent with previous work showing that AE predictability based on upstream solar wind measurements is fundamentally limited at multi-hour timescales.

15427390, 2026, 7, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2026SW005094 by NICE, National Institute for Health and Care Excellence, Wiley Online Library on [03/08/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

15427390, 2026, 7, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2026SW005094 by NICE, National Institute for Health and Care Excellence, Wiley Online Library on [05/08/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

The plateauing or sign reversal of some SHAP dependencies at longer horizons should therefore not be interpreted as unphysical behavior, but rather as a reflection of phase-dependent statistics in a nonlinear, history dependent system. Strong upstream driving is often associated with earlier AE enhancement; at later times, the conditional expectation may favor declining activity as the system unloads stored energy. These results reinforce the existence of an effective upper bound on AE predictability using near-Earth solar wind measurements alone. Turbulence information can move performance closer to this bound but cannot eliminate it.

The SHAP-based interpretation of bulk solar wind and IMF parameters aligns with established coupling physics. The strong short lead sensitivity to southward IMF reflects reconnection-controlled energy transfer, while the plateauing response under strong southward IMF is consistent with nonlinear saturation effects related to conductance limitations and magnetosphere–ionosphere feedbacks. The dominant influence of solar wind speed and magnetic field strength at short lead times is similarly consistent with their role in settings the convective electric field.

At longer lead times, the weakening and occasional sign reversal of these dependencies reflects the increasing importance of internal dynamics and event timing. These effects highlight the limitations of marginal feature interpretations in an autocorrelated system and underscore the need to interpret SHAP results as statistical attributions rather than direct causal relationships.

From an operational perspective, the results demonstrate that modest improvements in average forecast error can translate into meaningful gains when evaluated using decision-relevant metrics. The turbulence-aware model consistently improves cost-loss performance, particularly for higher thresholds and cost-loss ratios, indicating reduced false alarms and improved detection of impactful events. These findings suggest that incorporating physically motivated descriptors of variability may be a productive strategy for improving space weather forecasts, complementing advances in machine learning methodology.

Several limitations should be noted. The analysis is restricted to near-Earth solar wind measurements, limiting lead time and excluding explicit information about the internal magnetospheric state. Incorporating proxies for magnetospheric preconditioning may further improve forecasts, particularly at longer horizons. In addition, the turbulence descriptors have been derived from finite-resolution measurements and may not fully capture physical processes occurring below this resolution; future work could investigate the effects of higher resolution data. This work also made use of relatively simple machine learning models in order to demonstrate the benefit of adding turbulence-derived parameters, however, prediction skill could be improved further by combining this with more sophisticated models.

Other future work could explore adaptive windowing strategies, integrate turbulence diagnostics with state-space models, or apply similar approaches to other geomagnetic indices. More broadly, combining physically interpretable turbulence measures with data driven forecasting frameworks offers a promising path toward improved performance and understanding in space weather prediction.

8. Conclusions

This study has examined whether the inclusion of explicit solar wind turbulence descriptors improves short-term forecasts of the AE index within a machine learning framework. Two XGBoost models were developed and evaluated: a baseline model using mean solar wind parameters, and a turbulence-aware model augmented with physically motivated measured of variability.

The turbulence-aware model consistently outperforms the baseline across a range of lead times and evaluation metrics. It also outperforms a noise surrogate model, showing that the improvement is not just due to a greater number of parameters. While improvements in global error statistics are modest, they are systematic and most pronounced in regimes where baseline forecasts perform poorly, including under northward IMF conditions and during elevated AE activity. Importantly, these gains translate into improved operational performance, as demonstrated by reduced false alarms and improved cost-loss outcomes for higher AE thresholds.

Analysis of feature attributions confirms that bulk solar wind parameters dominate short-lead AE predictability, while turbulence descriptors provide complementary information that enhances forecast robustness. At longer lead times, forecast skill degrades for both models, consistent with increasing influence of internal magnetospheric dynamics and known limits to predictability based on upstream solar wind measurements alone.

15427390, 2026, 7, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2026SW005094 by NICE, National Institute for Health and Care Excellence, Wiley Online Library on [03/08/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Overall, these results demonstrate that incorporating physically interpretable descriptors of solar wind variability can yield meaningful improvements in geomagnetic activity forecasts without increasing model complexity. This approach provides a practical pathway for enhancing space weather prediction systems and motivates further exploration of turbulence-aware, physics-informed data driven models for other geomagnetic indices.

Conflict of Interest

The authors declare no conflicts of interest relevant to this study.

Availability Statement

The baseline, turbulence-aware and noise-surrogate AE prediction models, along with instructions on how to run them, can be found at https://github.com/carawaters/TurbulenceSpaceWeatherForecast (Waters, 2026). We acknowledge use of NASA/GSFC's Space Physics Data Facility's OMNIWeb service, and OMNI data (Papitashvili & King, 2020). The AE index was obtained from the World Data Center for Geomagnetism, Kyoto (https://wdc.kugi.kyoto-u.ac.jp/), as detailed by Davis and Sugiura (1966). Data was accessed using PySPEDAS, an implementation of the Space Physics Environment Data Analysis Software (SPEDAS) framework in Python (Grimes et al., 2022) (https://pyspedas.readthedocs.io/en/latest/), version 2.0.6 released 14th January 2026. The XGBoost model was implemented using the XGBoost Python package (T. Chen et al., 2026) (https://xgboost.readthedocs.io/en/stable/), version 3.1.3 released 9th January 2026, and was optimised using Optuna (Akiba et al., 2019) (https://github.com/optuna/optuna), version 4.6.0 released 20th November 2025. Model interpretation used SHAP (SHapley Additive exPlanations) (S. M. Lundberg & Lee, 2017) (https://shap.readthedocs.io/en/latest/index.html), version 0.50.0 released 11th November 2025, with the Correlated SHAP implementation (https://github.com/Fraunhofer-SCAI/corr_shap), based on the method described by Aas et al. (2020), version 0.0.2 released 14th November 2023.

Acknowledgments
CLW and CHKC are supported by UKRI Future Leaders Fellowship MR/W007657/1. CHKC is also supported by STFC Consolidated Grant ST/X000974/1. MJO is part-funded by STFC Grant ST/V000497/1 and NERC Grant NE/Y001052/1. This research utilised Queen Mary's Apocrita HPC facility, supported by QMUL Research-IT (http://doi.org/10.5281/zenodo.438045). The authors thank A. Smith for discussions about correlation-aware feature attribution.

References

Aas, K., Jullum, M., & Løland, A. (2020). Explaining individual predictions when features are dependent: More accurate approximations to Shapley values. Retrieved from https://arxiv.org/abs/1903.10464

Akasofu, S.-I. (1964). The development of the auroral substorm. *Planetary and Space Science*, *12*(4), 273–282. https://doi.org/10.1016/0032-0633(64)90151-5

Akiba, T., Sano, S., Yanase, T., Ohta, T., & Koyama, M. (2019). Optuna: A next-generation hyperparameter optimization framework. In *Proceedings of the 25th acm sigkdd international conference on knowledge discovery and data mining* (pp. 2623–2631). Association for Computing Machinery. https://doi.org/10.1145/3292500.3330701

Amariutei, O. A., & Ganushkina, N. Y. (2012). On the prediction of the auroral westward electrojet index. *Annals of Geophysics*, *30*(5), 841–847. https://doi.org/10.5194/angeo-30-841-2012

Astafyeva, E., Maletckii, B., Ouar, I. D., Förster, M., Themens, D. R., Huba, J. D., et al. (2025). An extraordinary dayside negative ionospheric storm of 11 May 2024 and total electron content (TEC) vortices. JGR. *Space Physics*, *130*(12), e2025JA034571. https://doi.org/10.1029/2025JA034571

Axford, W. I. (1964). Viscous interaction between the solar wind and the Earth's magnetosphere. *Planetary and Space Science*, *12*(1), 45–53. https://doi.org/10.1016/0032-0633(64)90067-4

Axford, W. I., & Hines, C. O. (1961). A unifying theory of high-latitude geophysical phenomena and geomagnetic storms. *Canadian Journal of Physics*, *39*, 1433–1464. https://doi.org/10.1139/p61-172

Bargatze, L. F., Baker, D. N., McPherron, R. L., & Hones, E. W. (1985). Magnetospheric impulse response for many levels of geomagnetic activity. *Journal of Geophysical Research*, *90*(A7), 6387–6394. https://doi.org/10.1029/JA090iA07p06387

Beedle, J. M. H., Genestreti, K. J., Shuster, J. R., Rice, R. C., Fuselier, S. A., Phan, T. D., et al. (2025). MMS observations of a compressed, strongly driven magnetopause during the 2024 Mother's Day storm. *Geophysical Research Letters*, *52*(24), e2025GL118368. https://doi.org/10.1029/2025GL118368

Borovsky, J. E. (2003). The strong relationship between the solar wind and the AE index: An analysis of data from 1998–2002. *Journal of Geophysical Research*, *108*(A7), 1284. https://doi.org/10.1029/2002JA009601

Borovsky, J. E. (2021a). Is our understanding of solar-wind/magnetosphere coupling satisfactory? *Frontiers in Astronomy and Space Sciences*, *8*, 634073. https://doi.org/10.3389/fspas.2021.634073

Borovsky, J. E. (2021b). On the saturation (or not) of geomagnetic indices. *Frontiers in Astronomy and Space Sciences*, *8*, 740811. https://doi.org/10.3389/fspas.2021.740811

Camporeale, E. (2019). The challenge of machine learning in space weather: Nowcasting and forecasting. *Space Weather*, *17*(8), 1166–1207. https://doi.org/10.1029/2018SW002061

Case, N. A., & Wild, J. A. (2012). A statistical comparison of solar wind propagation delays derived from multispacecraft techniques. *Journal of Geophysical Research: Space Physics*, *117*(A2). https://doi.org/10.1029/2011JA016946

Chen, C. H. K. (2016). Recent progress in astrophysical plasma turbulence from solar wind observations. *Journal of Plasma Physics*, *82*(6), 535820602. https://doi.org/10.1017/S0022377816001124

Chen, T., He, T., Benesty, M., Khotilovich, V., Tang, Y., Cho, H., et al. (2026). xgboost: Extreme gradient boosting [Computer software manual]. Retrieved from https://github.com/dmlc/xgboost(Pythonpackageversion3.1.3

Chu, X., Jia, L., McPherron, R., Li, X., & Bortnik, J. (2025). Imbalanced regression artificial neural network model for auroral electrojet indices (Iranna): Can we predict strong events? *Space Weather*, *23*(5), e2024SW004236. https://doi.org/10.1029/2024SW004236

Consolini, G. (2005). Local intermittency measure analysis of the AE index. *Geophysical Research Letters*, *32*, L04101. https://doi.org/10.1029/2004GL022063

D'Amicis, R., Bruno, R., & Bavassano, B. (2007). Is geomagnetic activity driven by solar wind turbulence? *Geophysical Research Letters*, *34*(5), L05108. https://doi.org/10.1029/2006GL028896

D'Amicis, R., Bruno, R., & Bavassano, B. (2009). Alfvénic turbulence in high speed solar wind streams as a driver for auroral activity. *Journal of Atmospheric and Solar-Terrestrial Physics*, *71*(10–11), 1014–1022. https://doi.org/10.1016/j.jastp.2008.05.002

D'Amicis, R., Bruno, R., & Bavassano, B. (2010). Geomagnetic activity driven by solar wind turbulence. *Advances in Space Research*, *46*(4), 514–520. https://doi.org/10.1016/j.asr.2009.08.031

D'Amicis, R., Bruno, R., & Bavassano, B. (2011). Response of the geomagnetic activity to solar wind turbulence during solar cycle 23. *Journal of Atmospheric and Solar-Terrestrial Physics*, *73*(5–6), 653–657. https://doi.org/10.1016/j.jastp.2011.01.012

D'Amicis, R., Telloni, D., & Bruno, R. (2020). The effect of solar-wind turbulence on magnetospheric activity. *Frontiers in Physics*, *8*, 604857. https://doi.org/10.3389/fphy.2020.604857

Davis, T. N., & Sugiura, M. (1966). Auroral electrojet activity index AE and its universal time variations. *Journal of Geophysical Research*, *71*(3), 785–801. https://doi.org/10.1029/JZ071i003p00785

Dungey, J. W. (1961). Interplanetary magnetic field and the auroral zones. *Physical Review Letters*, *6*(2), 47–48. https://doi.org/10.1103/PhysRevLett.6.47

Finch, I. D., & Lockwood, M. (2007). Solar wind–magnetosphere coupling functions on timescales of 1 day to 1 year. *Annales Geophysicae*, *25*(2), 495–506. https://doi.org/10.5194/angeo-25-495-2007

Freeman, M. P., Watkins, N. W., & Riley, D. J. (2000). Evidence for a solar wind origin of the power law burst lifetime distribution of the AE indices. *Geophysical Research Letters*, *27*(8), 1087–1090. https://doi.org/10.1029/1999GL010742

Gleisner, H., & Lundstedt, H. (2001). Auroral electrojet predictions with dynamic neural networks. *Journal of Geophysical Research*, *106*(A11), 24541–24550. https://doi.org/10.1029/2001JA900046

Gonzalez, W. D., Joselyn, J. A., Kamide, Y., Kroehl, H. W., Rostoker, G., Tsurutani, B. T., & Vasyliunas, V. M. (1994). What is a geomagnetic storm? *Journal of Geophysical Research*, *99*(A4), 5771–5792. https://doi.org/10.1029/93JA02867

Grimes, E. W., Harter, B., Hatzigeorgiu, N., Drozdov, A., Lewis, J. W., Angelopoulos, V., et al. (2022). The space physics environment data analaysis system in python. *Frontiers in Astronomy and Space Sciences*, *9*, 1020815. https://doi.org/10.3389/fspas.2022.1020815

Gu, Y., Wei, H.-L., Boynton, R. J., Walker, S. N., & Balikhin, M. A. (2019). System identification and data-driven forecasting of AE index and prediction uncertainty analysis using a new cloud-NARX model. *Journal of Geophysical Research: Space Physics*, *124*(1), 248–263. https://doi.org/10.1029/2018JA025957

Hasegawa, H., Fujimoto, M., Phan, T. D., Rème, H., Balogh, A., Dunlop, M. W., et al. (2004). Transport of solar wind into Earth's magnetosphere through rolled-up kelvin-helmholtz vortices. *Nature*, *430*(7001), 755–758. https://doi.org/10.1038/nature02799

Hnat, B., Chapman, S. C., & Rowlands, G. (2003). Scaling and a Fokker–Planck model for fluctuations in geomagnetic indices. *Geophysical Research Letters*, *30*, 2174. https://doi.org/10.1029/2003GL018146

Jankovičovà, D., Vörös, Z., & Šimkanin, J. (2008). The influence of solar wind turbulence on geomagnetic activity. *Nonlinear Processes in Geophysics*, *15*(1), 53–59. https://doi.org/10.5194/npg-15-53-2008

Kan, J. R., & Lee, L. C. (1979). Energy coupling function and solar wind–magnetosphere dynamo. *Geophysical Research Letters*, *6*(7), 577–580. https://doi.org/10.1029/GL006i007p00577

King, J. H., & Papitashvili, N. E. (2005). Solar wind spatial scales in and comparisons of hourly wind and ace plasma and magnetic field data. *Journal of Geophysical Research: Space Physics*, *110*(A2). https://doi.org/10.1029/2004JA010649

Kuai, J., Zhang, X., Zhong, J., Liu, L., Huang, F., Yue, X., et al. (2025). Prominent ionospheric responses throughout the multiple different altitudes from bottomside to topside ionosphere in the Asian-Australian sector during the May 2024 mother's day storm. *Journal of Geophysical Research: Space Physics*, *130*(12), e2025JA034867. https://doi.org/10.1029/2025JA034867

Lockwood, M. (2022). Solar wind–magnetosphere coupling functions: Pitfalls, limitations, and applications. *Space Weather*, *20*(2), e2021SW002989. https://doi.org/10.1029/2021SW002989

Lundberg, S., & Lee, S.-I. (2017). A unified approach to interpreting model predictions. Retrieved from https://arxiv.org/abs/1705.07874

Lundberg, S. M., & Lee, S.-I. (2017). *A unified approach to interpreting model predictions*. Curran Associates, Inc. Retrieved from http://papers.nips.cc/paper/7062-a-unified-approach-to-interpreting-model-predictions.pdf

Luo, B., Li, X., Temerin, M., & Liu, S. (2013). Prediction of the AU, AL, and AE indices using solar wind parameters. *Journal of Geophysical Research: Space Physics*, *118*(12), 7683–7694. https://doi.org/10.1002/2013JA019188

Lyons, L. R., Liu, S., Ruohoniemi, J. M., Solovyev, S. I., & Samson, J. C. (2003). Observations of dayside convection reduction leading to substorm onset. *Journal of Geophysical Research: Space Physics*, *108*(A3). https://doi.org/10.1029/2002JA009670

MacTaggart, D., Williams, T., & Aslam, O. P. M. (2025). The magnetic topology of AR13664 leading to its first halo CME. *Journal of Geophysical Research: Space Physics*, *130*(4), e2024JA033462. https://doi.org/10.1029/2024JA033462

Maimaiti, M., Kunduri, B., Ruohoniemi, J. M., Baker, J. B. H., & House, L. L. (2019). A deep learning-based approach to forecast the onset of magnetic substorms. *Space Weather*, *17*(11), 1534–1552. https://doi.org/10.1029/2019SW002251

Mayaud, P.-N. (1980). Derivation, meaning, and use of geomagnetic indices. *American Geophysical Union*, *22*. https://doi.org/10.1029/GM022

McPherron, R. L. (1970). Growth phase of magnetospheric substorms. *Journal of Geophysical Research*, *75*(28), 5592–5599. https://doi.org/10.1029/JA075i028p05592

Myllys, M., Kilpua, E. K. J., Pulkkinen, T. I., Tanskanen, E. I., Osmane, A., Pitkänen, T., & Hietala, H. (2016). Solar wind–magnetosphere coupling efficiency during ejecta- and sheath-driven geomagnetic storms. *Journal of Geophysical Research: Space Physics*, *121*, 4378–4396. https://doi.org/10.1002/2016JA022407

Newell, P. T., Sotirelis, T., Liou, K., Meng, C.-I., & Rich, F. J. (2007). A nearly universal solar wind–magnetosphere coupling function inferred from 10 magnetospheric state variables. *Journal of Geophysical Research*, *112*(A1), A01206. https://doi.org/10.1029/2006JA012015

Nykyri, K., & Otto, A. (2001). Plasma transport at the magnetospheric boundary due to reconnection in Kelvin–Helmholtz vortices. *Geophysical Research Letters*, *28*(18), 3565–3568. https://doi.org/10.1029/2001GL013239

Ohtani, S., Zou, Y., Merkin, V. G., Wiltberger, M., Pham, K. H., Raptis, S., et al. (2025). Ground magnetic response to an extraordinary IMF By flip during the May 2024 storm: Travel time from the magnetosheath to dayside high latitudes. *Journal of Geophysical Research: Space Physics*, *130*(5), e2024JA033691. https://doi.org/10.1029/2024JA033691

15427390, 2026, 7, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2026SW005094 by NICE, National Institute for Health and Care Excellence, Wiley Online Library on [03/08/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Osmane, A., Dimmock, A. P., Naderpour, R., Pulkkinen, T. I., & Nykyri, K. (2015). The impact of solar wind ULF B_Z fluctuations on geomagnetic activity for viscous timescales during strongly northward and southward IMF. *Journal of Geophysical Research: Space Physics*, *120*(11), 9307–9322. https://doi.org/10.1002/2015JA021505

Owens, M. J., & Riley, P. (2017). Probabilistic solar wind forecasting using large ensembles of near-sun conditions with a simple one-dimensional "upwind" scheme. *Space Weather*, *15*(11), 1461–1474. https://doi.org/10.1002/2017SW001679

Pallocchia, G., Amata, E., Consolini, G., Marcucci, M. F., & Bertello, I. (2008). AE index forecast at different time scales through an ANN algorithm based on L1 IMF and plasma measurements. *Journal of Atmospheric and Solar-Terrestrial Physics*, *70*(2–4), 663–668. https://doi.org/10.1016/j.jastp.2007.08.038

Papitashvili, N. E., & King, J. H. (2020). OMNI 1-min data [Dataset]. *NASA Space Physics Data Facility*. https://doi.org/10.48322/45bb-8792

Perreault, P., & Akasofu, S.-I. (1978). A study of geomagnetic storms. *Geophysical Journal International*, *54*(3), 547–573. https://doi.org/10.1111/j.1365-246X.1978.tb05494.x

Pulkkinen, A., Klimas, A., Vassiliadis, D., & Uritsky, V. (2006). Role of stochastic fluctuations in the magnetosphere-ionosphere system: A stochastic model for the AE index variations. *Journal of Geophysical Research*, *111*(A10). https://doi.org/10.1029/2006JA011661

Richardson, D. S. (2011). Economic value and skill. In I. T. Jolliffe & D. B. Stephenson (Eds.), *Forecast verification: A practitioner's guide in atmospheric science* (2nd ed., pp. 167–184). Wiley. https://doi.org/10.1002/9781119960003.ch9

Schekochihin, A. A. (2022). MHD turbulence: A biased review. *Journal of Plasma Physics*, *88*(5), 155880501. https://doi.org/10.1017/S0022377822000721

Siscoe, G. L., Crooker, N. U., & Siebert, K. D. (2002). Transpolar potential saturation: Roles of region 1 current system and solar wind ram pressure. *Journal of Geophysical Research*, *107*(A10). https://doi.org/10.1029/2001JA009176

Takalo, J., & Timonen, J. (1997). Neural network prediction of AE data. *Geophysical Research Letters*, *24*(19), 2403–2406. https://doi.org/10.1029/97GL02457

Tanskanen, E. I. (2009). A comprehensive high-throughput analysis of substorms observed by image magnetometer network: Years 1993–2003 examined. *Journal of Geophysical Research*, *114*(A5). https://doi.org/10.1029/2008JA013682

Tulasi Ram, S., Veenadhari, B., Dimri, A. P., Bulusu, J., Bagiya, M., Gurubaran, S., et al. (2024). Super-intense geomagnetic storm on 10–11 May 2024: Possible mechanisms and impacts. *Space Weather*, *22*(12), e2024SW004126. https://doi.org/10.1029/2024SW004126

Ventriglia, V., Guerra, M., Cesaroni, C., Spogli, L., Altadill, D., Segarra, A., et al. (2025). An explainable machine learning model for large-scale travelling ionospheric disturbances forecasting. *Journal of Space Weather and Space Climate*, *15*(25), 25. https://doi.org/10.1051/swsc/2025020

Waters, C. L. (2026). TurbulenceSpaceWeatherForecast [Software]. Retrieved from https://github.com/carawaters/TurbulenceSpaceWeatherForecastdoi:10.5281/zenodo.19259477

Weigel, R. S., Horton, W., & Tajima, T. (1999). Forecasting the AE index using neural networks. *Geophysical Research Letters*, *26*, 1353–1356. https://doi.org/10.1029/1999GL900280

Weigel, R. S., Vassiliadis, D., Klimas, A. J., & Baker, D. N. (2003). Solar wind coupling to and predictability of ground magnetic field variations. *Journal of Geophysical Research*, *108*. https://doi.org/10.1029/2002JA009627

Weimer, D. R., Ober, D. M., Maynard, N. C., Collier, M. R., McComas, D. J., Ness, N. F., et al. (2003). Predicting interplanetary magnetic field (IMF) propagation delay times using the minimum variance technique. *Journal of Geophysical Research: Space Physics*, *108*(A1). https://doi.org/10.1029/2002JA009405

Zimbardo, G., Greco, A., Sorriso-Valvo, L., Perri, S., Vörös, Z., Aburjania, G., et al. (2010). Magnetic turbulence in the geospace environment. *Space Science Reviews*, *156*(1–4), 89–134. https://doi.org/10.1007/s11214-010-9692-5

Zou, Z., Huang, H., Zuo, P., Ni, B., San, W., Yuan, Q., et al. (2024). A forecast model of geomagnetic indices from the solar wind fluids observations based on long short-term memory neural network. *Physics of Fluids*, *36*(2), 026616. https://doi.org/10.1063/5.0196284

15427390, 2026, 7, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2026SW005094 by NICE, National Institute for Health and Care Excellence, Wiley Online Library on [03/08/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License