https://doi.org/10.5194/egusphere-2026-6859
Preprint. Discussion started: 19 November 2026
© Author(s) 2026. CC BY 4.0 License.

Predictability of IRI-2020 topside electron temperature errors over the African sector: nonlinear dynamics and out-of-sample skill from Swarm Langmuir-probe observations

Ozuem V. Edward[1], Oladipo E. Abe[2], Tunde D. Akinnubi[1], Blessing Ngozi O. Ezeoke[3]

[1]*Department of Physics, Anchor University Lagos, Nigeria*

[2]*Space Physics Research Laboratory, Department of Physics, Federal University of Oye-Ekiti, P. M. B. 373, Oye Ekiti, Ekiti State, Nigeria.*

[3]*Department of Physics, Lead City University, Ibadan, Nigeria*

Correspondence: Ozuem V. Edward (eozuem@gmail.com)

Abstract

The residual between the electron temperature measured by the Swarm A and B Langmuir probes and the IRI-2020 topside electron temperature (TBT-2012 option) was analysed along satellite passes over the African low-latitude sector (20°W–55°E, dip latitude within ±40°) in a geomagnetically quiet window (February 2024) and a disturbed window (late April to May 2024, including the superstorm of 10–12 May), and predicted out of sample from February to May 2024. Three instrumental effects were identified and handled before the analysis: a relaxation after each 128-s probe sweep, the absence of the radar-based calibration over much of the daytime anomaly, and switches of the night-side temperature between two states separated by 600–900 K that are not reflected in the quality flags; 452 of 627 passes in the two windows were free of such switches. Delay embedding, recurrence quantification and comparison with 100 iterative amplitude-adjusted Fourier-transform surrogates show that the determinism of the residual exceeds that of a matched linear process in 396 of the 452 passes. The excess is largest on the quiet day side (median 5.3 surrogate standard deviations) and smaller under storm forcing

https://doi.org/10.5194/egusphere-2026-6859
Preprint. Discussion started: 19 November 2026
© Author(s) 2026. CC BY 4.0 License.



(3.3; $p = 6.2 \times 10^{-8}$), whereas raw determinism increases. Applying the radar-based calibration changes determinism by 0.14–0.32 of the between-pass interquartile range, and replacing TBT-2012 by the Bil-1985 option by 0.15–0.63, so the invariants are attributable mainly to the IRI error in its default configuration. Surrogate-supported along-track horizons of about 330–740 km were found for 51 passes. In the prediction experiment, IRI-2020 alone had an RMSE of 559 K, dominated by an offset that a 60-day climatology of recent passes reduced to 271 K; relative to this climatology, a gradient-boosted tree ensemble using the residual of previous days improved the mean-square error by 0.14 (95 % interval 0.02–0.23), and a nowcast also using the measured ion density by 0.39 (0.27–0.48). Day-side passes with a larger nonlinear excess were predicted less well.

https://doi.org/10.5194/egusphere-2026-6859
Preprint. Discussion started: 19 November 2026
© Author(s) 2026. CC BY 4.0 License.

1 Introduction

The electron temperature, T_e, is a principal state variable of the topside ionosphere. It controls the plasma scale height, the rates of temperature-dependent recombination and the partition of energy between electrons, ions and neutrals, and it enters directly into the interpretation of radio-propagation and in situ plasma measurements (Schunk and Nagy, 2009). The International Reference Ionosphere (IRI) is the internationally recognised empirical standard for ionospheric plasma parameters, and in its 2020 version it specifies the topside electron temperature through the TBT-2012 model of Truhlik et al. (2012), with the older Bil-1985 formulation retained as an option (Bilitza et al., 2022). Both are climatological descriptions: they represent the average dependence of T_e on altitude, latitude, local time, season and solar activity, and by construction they cannot follow the day-to-day and storm-time variability of the topside thermal structure.

Comparisons of IRI with in situ topside measurements quantify this limitation. Pignalberi et al. (2016) compared IRI with early Swarm Langmuir-probe observations during the St. Patrick's Day storm of 2015, and Pignalberi et al. (2021) showed, using Swarm, incoherent-scatter radar observations and IRI, that the median difference between the Swarm and radar electron temperatures, 300–400 K before calibration, is reduced to near zero by the calibration, and that IRI does not reproduce the high-temperature tail of the observed distribution. The calibration that removes most of this offset was derived by Lomidze et al. (2018) from coincident radar and radio-occultation data. Further Swarm studies have described the correlation between electron density and temperature in the topside (Pignalberi et al., 2024a; Yan et al., 2022) and the solar-activity dependence of T_e that IRI does not capture (Pignalberi et al., 2024b). At low latitudes, the electron temperature is organised by the equatorial plasma fountain and the equatorial ionisation anomaly (EIA): the morning overshoot of T_e is enhanced by downward plasma drift (Oyama et al., 1996) and varies with solar flux (Stolle et al., 2011), and during the May 2024 superstorm Swarm recorded electron temperature increases of the order of 1500 K in the EIA trough (Shreedevi et al., 2025). These studies establish that the IRI electron temperature error is large and structured at low latitudes. They do not establish whether that error is dynamically

https://doi.org/10.5194/egusphere-2026-6859
Preprint. Discussion started: 19 November 2026
© Author(s) 2026. CC BY 4.0 License.

deterministic, how its deterministic content is organised across the anomaly, or whether it can be predicted out of sample.

The first of these questions was addressed for total electron content in a study of the African equatorial ionosphere (Edward, 2026). That work formed the residual between GNSS-derived TEC and IRI-2020 at six stations, reconstructed its phase space by delay embedding (Takens, 1981) with the delay and dimension chosen from the average mutual information (Fraser and Swinney, 1986) and the false-nearest-neighbour fraction (Kennel et al., 1992), and characterised it by recurrence quantification analysis (RQA; Eckmann et al., 1987; Zbilut and Webber, 1992; Marwan et al., 2007), by the maximal Lyapunov exponent (Rosenstein et al., 1993), and by comparison with iterative amplitude-adjusted Fourier-transform (IAAFT) surrogates (Schreiber and Schmitz, 1996). A central element of that study was an attribution test showing that a slowly varying contaminant of the residual, the plasmaspheric content above the IRI ceiling, perturbed the dynamical invariants only by small and explicitly bounded amounts. Related nonlinear analyses of ionospheric and geomagnetic time series have been reported by Consolini et al. (2020) and Materassi et al. (2023).

The present study transfers this framework from TEC to the topside electron temperature measured in situ by the Langmuir probes of the European Space Agency Swarm constellation (Friis-Christensen et al., 2006; Knudsen et al., 2017), over the African low-latitude sector, for the same quiet (February 2024) and disturbed (late April to May 2024, including the superstorm of 10–12 May) windows. Three differences from the TEC case shape the design. First, a satellite does not provide a fixed-station time series; the analysis is therefore carried out along track, pass by pass, and the Lyapunov horizon becomes an along-track scale rather than a time in hours. Second, the slowly varying contaminant of the residual is now instrumental: it is the calibration offset of the Langmuir-probe temperature relative to incoherent-scatter radars, which Swarm distributes as a separate correction. Third, because each pass crosses the whole anomaly, both EIA crests are sampled, and the dip-latitude organisation can be resolved within each pass rather than across a fixed station network.

The study is organised around five questions. (i) Does the IRI-2020 topside T_e residual over Africa carry deterministic structure in excess of a linear stochastic process with the same spectrum and amplitude distribution? (ii) Are the dynamical invariants governed by the IRI error

https://doi.org/10.5194/egusphere-2026-6859
Preprint. Discussion started: 19 November 2026
© Author(s) 2026. CC BY 4.0 License.

rather than by the instrumental calibration offset or by the choice of the IRI topside T_e option? (iii) How is the deterministic content organised in dip latitude and local time, and how does it change under storm forcing? (iv) Can the residual be predicted out of sample, so that a data-driven correction of IRI-2020 T_e becomes possible, and how does the skill depend on geomagnetic activity? (v) Is the predictive skill of a pass related to its dynamical invariants?

2 Attribution framework for the electron-temperature residual

Let $\delta T_e(t)$ denote the residual formed along a satellite pass by subtracting the IRI-2020 electron temperature from the Langmuir-probe value at the satellite position. Writing the measured temperature as the sum of a part reproducible in principle by the model, the IRI prediction, an instrumental offset $B(t)$ relative to the true temperature, and observational noise $\eta(t)$, the residual decomposes as

$$\delta T_e(t) = T_e^{\mathrm{LP}}(t) - T_e^{\mathrm{IRI}}(t) = E(t) + B(t) + \eta(t), \qquad (1)$$

where $E(t)$ is the IRI prediction error, the object of interest. The delay-coordinate reconstruction with dimension m and delay τ is

$$\boldsymbol{X}(t) = [\delta T_e(t),\ \delta T_e(t+\tau),\ \ldots,\ \delta T_e(t+(m-1)\tau)], \qquad (2)$$

and by linearity it separates into the reconstructions of the three components. The offset $B(t)$ plays the role that the plasmaspheric content played in the TEC study (Edward, 2026). If its increments over one embedding span $(m-1)\tau$ are bounded by a small quantity ϵ_B, each component of its reconstruction differs from $B(t)$ by at most ϵ_B, so that $\|\boldsymbol{X}_B(t) - B(t)\,\mathbf{1}_m\|_\infty \leq \epsilon_B$, where $\mathbf{1}_m$ is the main diagonal of the embedding space. The offset then translates the reconstructed trajectory along that diagonal without deforming its internal geometry, and the triangle inequality bounds its effect on every recurrence distance,

$$\left|\, \left\|\boldsymbol{X}(t_i) - \boldsymbol{X}(t_j)\right\|_\infty - \left\|\boldsymbol{X}_{E+\eta}(t_i) - \boldsymbol{X}_{E+\eta}(t_j)\right\|_\infty \right| \leq |B(t_i) - B(t_j)| + 2\epsilon_B. \qquad (3)$$

The derivation is identical to that of Edward (2026) with the plasmaspheric term replaced by the instrumental term. Whether the bound is small is an empirical matter, and it is tested directly in Sect. 4.3 by recomputing the invariants with and without the radar-based calibration of Lomidze et al. (2018). Two features distinguish the present case. The calibration correction depends on the

https://doi.org/10.5194/egusphere-2026-6859
Preprint. Discussion started: 19 November 2026
© Author(s) 2026. CC BY 4.0 License.

measured temperature and density, so it is not a pure additive constant along a pass. It is also undefined where the plasma density exceeds the range of the calibration data, which on Swarm is the case over most of the daytime anomaly crests. The calibration test is therefore available only for the passes on which the correction is defined throughout, which are predominantly night-side passes. A second, model-side control replaces the TBT-2012 topside electron temperature of IRI-2020 by the Bil-1985 option; this changes T_e^{IRI} in a known direction and measures how much of the residual's dynamical structure is a property of the particular IRI formulation.

3 Data and methods

3.1 Swarm Langmuir-probe electron temperature

Electron temperatures were taken from the Swarm Level-1b Langmuir-probe product (EFIx_LP_1B, baseline 0701), which provides the electron temperature, the ion density, the spacecraft potential and quality flags at 2 Hz (Knudsen et al., 2017). Swarm A, at a mean altitude of about 470 km, and Swarm B, at about 510 km, were used; Swarm C flies alongside Swarm A and was not used. Files were obtained for every day from 1 January to 25 May 2024. The nonlinear diagnostics use the quiet window of 1–29 February 2024 and the storm window of 27 April to 25 May 2024; the remaining days serve only as training history for the prediction experiment of Sect. 3.6.

Samples were retained when the temperature flag indicated nominal high-gain data without a detected spike, the probe-source flag indicated high-gain data, the Te-related error bits of the detailed flag word were clear, and the temperature lay between 300 and 10 000 K. The Langmuir probes perform a voltage sweep every 128 s. Inspection of the 2-Hz data showed that each sweep is followed by a temperature excess of several tens of kelvin that relaxes over a few seconds, a periodic artefact that would otherwise add spurious recurrences. Samples within 0.75 s of each sweep were removed, and the relaxation was removed sweep by sweep by fitting the samples 0.5–25 s after the sweep with a constant, a linear term and an exponential of common e-folding time, and subtracting the fitted exponential. Over the 292 satellite-days, the fitted e-folding time had a median of 5.5 s (interquartile range 3.0–12.0 s, the upper quartile coinciding with the upper end of the search grid) and the fitted amplitude a median of 40 K (interquartile range 31–55 K). After the correction, the median Te in the interval 8–20 s after a sweep still differed from the

https://doi.org/10.5194/egusphere-2026-6859
Preprint. Discussion started: 19 November 2026
© Author(s) 2026. CC BY 4.0 License.

median in the 20 s before it by a median absolute amount of 28 K (interquartile range 19–35 K), with a median signed difference of 4 K. Part of this difference is the along-track change of the temperature over about 30 s, and part is an instrumental step at the sweep. The residual sweep signature is a periodic, approximately linear component of the series, so it is preserved by the IAAFT surrogates of Sect. 3.4 and does not by itself raise the standardised excess of determinism; its possible influence on the raw invariants is examined by the controls of Sect. 3.5. The 2-Hz samples were averaged onto a uniform 1-s grid, gaps of up to 10 s were filled by linear interpolation, and isolated spikes were removed with a Hampel filter of half-width 5 s and a threshold of three scaled median absolute deviations. Inspection of the night-side passes revealed a further instrumental feature that is not represented in the quality flags. On both satellites the night-side temperature alternates between two levels: the pass medians form two clusters, centred near 1500–1650 K and near 2150–2250 K, separated at 1870 K for Swarm A and 1910 K for Swarm B by the Otsu criterion. The level can change from one pass to the next, about 95 min apart, and also within a pass, where it switches at a sweep and persists for one or several 128-s sweep cycles before returning. The ion density, the spacecraft potential and the flags do not change at these transitions; on Swarm B the ratio of the electron-derived to the ion-derived density does (about 0.8 in the upper state and about 0.3 in the lower state on the nights examined). Changes of 600–900 K between successive orbits under quiet conditions, without any accompanying change of density, are difficult to reconcile with a geophysical origin, so the bistability is treated here as instrumental; its cause could not be established from the distributed product, and we found no description of it in the validation literature (Catapano et al., 2022). Within a pass, a state switch appears as a step of several hundred kelvin at a sweep, far larger than the median absolute step of 28 K at ordinary sweeps. Every pass was therefore tested for switches: a pass was classified as switching when the difference between the median temperature 8–20 s after any sweep and the median in the 20 s before it exceeded 300 K in magnitude, or, for night passes, where the 60-s median temperature changes by more than 400 K between consecutive minutes (the night-side distribution of these changes is bimodal, with smooth passes below 200 K and switching passes above 500 K). The primary analysis uses only passes without switches. A pass that remains in one state throughout carries the state as a constant offset, which by Eq. (3) does not alter the recurrence structure, the Lyapunov exponent or the surrogate comparison. The same analysis with all passes is retained as a sensitivity case. The radar-based

https://doi.org/10.5194/egusphere-2026-6859
Preprint. Discussion started: 19 November 2026
© Author(s) 2026. CC BY 4.0 License.

calibration (Lomidze et al., 2018) is distributed as a separate correction; the calibrated temperature $T_e^{\mathrm{LP}} + \Delta T_e^{\mathrm{cal}}$ was formed only for the calibration test, and the uncalibrated temperature is used everywhere else because the correction is undefined over much of the daytime anomaly.

3.2 IRI-2020 electron temperature

IRI-2020 (Bilitza et al., 2022) was evaluated at the geodetic position, altitude and time of every 1-s sample, with the TBT-2012 topside electron temperature (Truhlik et al., 2012), the default of IRI-2020, as the primary configuration and the Bil-1985 option as the model-side control. The IRI-STORM correction was active; it modifies the F2-peak density but not the topside temperature formulations. The same IRI-2020 code and coefficient files were used as in Edward (2026).

3.3 Pass construction and geomagnetic context

A pass is a continuous segment of a satellite orbit inside the box bounded by 20°W and 55°E and by 45°S and 45°N geographic latitude, trimmed to the longest contiguous stretch with dip latitude between −40° and +40°. Dip latitude was computed from the IGRF inclination I at the sub-satellite point for epoch 2024.2, as $\lambda_{\mathrm{dip}} = \arctan(\tan I/2)$, the convention of Edward (2026). The trimming removes the subauroral edge of the southern end of the box, where storm-time temperatures reach several thousand kelvin and would otherwise dominate the phase-space diameter, and the region near the South Atlantic Anomaly where dip latitude is not monotonic along the track. Passes shorter than 900 s or with less than 90 % valid samples were rejected. Each pass was assigned the solar local time of its dip-equator crossing and classified as day (08–16 h), night (20–04 h), or twilight (the remaining hours), and was assigned the planetary Kp index of its central time (Matzka et al., 2021). The 30-min Hp30 index (Yamazaki et al., 2022) and the daily F10.7 flux were used as drivers in the prediction experiment.

Of the passes over the box in the 146 days, 1428 satisfied these criteria (88 were too short after trimming, 276 had too few valid samples and 114 contained a gap longer than 10 s); 627 of them fall in the two analysis windows, and 452 of these are free of state switches. The switch test removed passes unevenly: 35 of 76 quiet-window night passes and 68 of 95 storm-window night passes, but only 4–26 passes in each of the other window and local-time classes. Table 1 lists

https://doi.org/10.5194/egusphere-2026-6859
Preprint. Discussion started: 19 November 2026
© Author(s) 2026. CC BY 4.0 License.

them by window, local-time class and satellite, and Fig. 1 shows the geometry, the geomagnetic activity and the local times sampled. Because the orbital planes of Swarm A and Swarm B precess slowly in local time, the local-time classes are not sampled evenly: in the quiet window Swarm A provided the day and night passes and Swarm B the dawn and dusk passes, whereas in the storm window the roles were reversed. Local-time class, satellite altitude and window are therefore partly confounded, and comparisons between windows are made within a local-time class. The radar-based calibration was defined over the whole pass on all 76 quiet-window night passes but on only 10 of the 95 storm-window night passes, so the calibration test is informative mainly for quiet conditions.

3.4 Nonlinear diagnostics

For every pass, the delay τ was fixed at the first minimum of the average mutual information (Fraser and Swinney, 1986), searched up to 300 s, and the dimension m at the first value for which the false-nearest-neighbour fraction (Kennel et al., 1992) fell below 1 %. Because the 1-s Langmuir-probe temperature contains measurement noise, the false-nearest-neighbour fraction of some passes levels off at a few per cent instead of falling below 1 %; for those passes m was taken as the first dimension at which the fraction stopped decreasing by more than one percentage point, and the rule used is recorded for every pass. Embedding spans longer than one third of the pass were flagged.

The recurrence matrix was built with the Chebyshev norm,

$$R_{ij} = \Theta\left(\varepsilon - \left\|\boldsymbol{X}_i - \boldsymbol{X}_j\right\|_\infty\right), \qquad (4)$$

with the threshold ε set to 10 % of the maximum phase-space diameter and the line of identity excluded. Determinism is the fraction of recurrence points that lie on diagonal lines of length at least $l_{\min} = 2$,

$$\mathrm{DET} = \frac{\sum_{l \geq l_{\min}} l\, P(l)}{\sum_{l \geq 1} l\, P(l)}, \qquad (5)$$

where $P(l)$ is the histogram of diagonal-line lengths, and laminarity (LAM), trapping time (TT), diagonal-line entropy (ENTR) and recurrence rate (RR) were computed as in Marwan et al. (2007). To resolve the dip-latitude organisation inside a pass, determinism was also

https://doi.org/10.5194/egusphere-2026-6859
Preprint. Discussion started: 19 November 2026
© Author(s) 2026. CC BY 4.0 License.

decomposed by rows of the recurrence matrix. For $l_{\min} = 2$ a recurrence point lies on a diagonal line if and only if one of its diagonal neighbours is also a recurrence point, so that

$$\mathrm{DET}_b = \frac{\sum_{i\in b} \sum_j R_{ij} \left(R_{i-1,j-1} \vee R_{i+1,j+1}\right)}{\sum_{i\in b} \sum_j R_{ij}}, \qquad (6)$$

where b is the set of delay vectors whose central sample falls in a given 5° dip-latitude bin. Summed over all rows, Eq. (6) reproduces Eq. (5) exactly.

The maximal Lyapunov exponent λ_1 was estimated by the method of Rosenstein et al. (1993) as the slope of the mean logarithmic divergence of nearest neighbours, excluding temporal neighbours closer than the first zero of the autocorrelation function:

$$y(k) = \langle \ln d_j(k) \rangle, \qquad \lambda_1 = \frac{\mathrm{d}y}{\mathrm{d}(k\,\Delta t)}, \qquad (7)$$

with $\Delta t = 1$ s. The along-track horizon was expressed as $1/\lambda_1$ in seconds, as the ground distance v_g/λ_1 travelled in that time, where v_g is the mean ground-track speed of the pass, and as the corresponding change in dip latitude.

Statistical significance was assessed against 100 IAAFT surrogates per pass (Schreiber and Schmitz, 1996), which preserve the power spectrum and amplitude distribution of the residual while destroying nonlinear phase structure. Following Marwan et al. (2007) and Edward (2026), the threshold of the observed series was held fixed for its surrogates. Determinism is reported both as a raw value and as the standardised excess

$$z_{\mathrm{DET}} = \frac{\mathrm{DET}_{\mathrm{obs}} - \overline{\mathrm{DET}}_{\mathrm{surr}}}{\sigma_{\mathrm{surr}}}, \qquad (8)$$

and the null hypothesis was rejected at the 5 % level when fewer than 5 % of the surrogates reached the observed value. The same comparison was made for each dip-latitude bin of Eq. (6) and for λ_1; a horizon is called statistically supported when $\lambda_1 > 0$ and the surrogate test for λ_1 rejects the null hypothesis. The IAAFT null assumes stationarity, which a pass crossing the anomaly does not strictly satisfy; this caveat is shared with the TEC study, where the diurnal cycle plays the same role.

https://doi.org/10.5194/egusphere-2026-6859
Preprint. Discussion started: 19 November 2026
© Author(s) 2026. CC BY 4.0 License.

3.5 Sensitivity and robustness protocol

Five controls were applied to every pass. (E1a) The IRI topside temperature option was changed from TBT-2012 to Bil-1985 with the embedding fixed. (E1b) On passes where the radar-based calibration is defined throughout, the calibrated temperature replaced the uncalibrated one with the embedding fixed. (E2) The surrogate test was repeated with the alternative convention in which each surrogate receives its own threshold. (E3) The invariants were recomputed on the first difference of the residual, which suppresses any slowly varying term, and on the residual block-averaged to a 3-s cadence, which reduces the measurement noise, each with its own embedding. (E4) A common embedding, the median m and τ of all passes, replaced the per-pass embedding. The perturbations from E1a and E1b are compared with the spread of the reference invariants between passes of the same group, in the manner of the altitude-bound test of Edward (2026).

3.6 Out-of-sample prediction of the residual

The prediction experiment asks whether δT_e, and hence the corrected temperature $T_e^{\mathrm{IRI}} + \widehat{\delta T_e}$, can be predicted for passes not used in fitting. It is forward-chained: every pass of day d is predicted by models that use only passes from days $d - 60$ to $d - 1$, and the evaluation covers every day from 1 February to 25 May 2024, so that the first evaluated day has 31 days of history. Driver indices enter causally, as the last completed Hp30 and Kp values before each sample, the maximum Hp30 over the preceding 24 h, the F10.7 flux of the previous day and the trailing 81-day mean F10.7. Seven predictors of δT_e were compared. IRI alone ($\widehat{\delta T_e} = 0$) is the baseline. BIAS7 is the median residual of the same satellite and orbital node over the preceding seven days. CLIM60 is the median residual of the preceding 60 days in bins of 2.5° dip latitude, 1 h local time and satellite, with coarser bins used where a fine bin has fewer than 20 samples. PER1 is the median residual profile, in 1° dip-latitude bins, of the passes of the same satellite and node on the previous day. TREE-D, TREE-H and TREE-HN are gradient-boosted regression-tree ensembles (Friedman, 2001),

$$F_M(\boldsymbol{x}) = F_0 + \nu \sum_{k=1}^{M} h_k(\boldsymbol{x}), \qquad (9)$$

with $M = 300$ trees of at most 63 splits and at least 200 samples per leaf, learning rate $\nu = 0.05$, and predictors binned into 64 quantile bins, all fixed before any evaluation. TREE-D uses

https://doi.org/10.5194/egusphere-2026-6859
Preprint. Discussion started: 19 November 2026
© Author(s) 2026. CC BY 4.0 License.

geophysical and driver features only: dip and geographic latitude, longitude, altitude, the sine and cosine of local time, satellite, orbital node, the IRI temperature and density at the sample, and the five driver indices. TREE-H adds the PER1 and BIAS7 values as features. TREE-HN further adds the concurrently measured ion density; it is a nowcast that requires the in situ density and is reported separately for that reason. The tree ensembles were retrained every seven days on every fifteenth 1-s sample of the preceding 60 days, a thinning that removes most of the redundancy between adjacent along-track samples.

Skill was measured by the root-mean-square error (RMSE), the mean absolute error, the mean error, and the mean-square-error skill score relative to a reference,

$$SS = 1 - \frac{\mathrm{MSE}_{\mathrm{model}}}{\mathrm{MSE}_{\mathrm{ref}}}, \qquad (10)$$

with IRI and CLIM60 as references. Confidence intervals were obtained from 1000 day-block bootstrap resamples (Künsch, 1989), daily mean-square errors were compared with the Wilcoxon signed-rank test (Wilcoxon, 1945), and the false-discovery rate across all tests was controlled by the procedure of Benjamini and Hochberg (1995). Results are reported by window, local-time class, Kp class of the verifying pass ($Kp < 3$, $3 \leq Kp < 5$, $Kp \geq 5$) and dip latitude.

4 Results and discussion

The observation geometry and the geomagnetic context are shown in Fig. 1, and the passes retained after quality control are listed in Table 1. Kp does not exceed 4.3 during any retained pass of the quiet window, whereas the storm window contains the superstorm of 10–12 May 2024, with Kp reaching 9 (Fig. 1b). Each empirical result below is followed by its physical interpretation and by an assessment of what it does and does not establish. Unless stated otherwise, the statistics refer to the 452 passes free of Langmuir-probe state switches; the corresponding results for all 627 passes are summarised in Sect. 4.7.

Table 1. Passes retained for the nonlinear analysis (free of Langmuir-probe temperature state switches), by window, local-time class and satellite: number of passes, number on which the radar-based calibration is defined throughout, median altitude, median solar local time of the dip-

https://doi.org/10.5194/egusphere-2026-6859
Preprint. Discussion started: 19 November 2026
© Author(s) 2026. CC BY 4.0 License.

equator crossing, median and maximum Kp, median of the pass-mean and pass-RMS residual, and median mean calibration offset (calibrated minus uncalibrated temperature) where defined.

Window	LT class	Sat.	Passes	Passes with calibration	Altitude (km)	SLT (h)	Kp median	Kp max	Mean δT_e (K)	RMS δT_e (K)	Calibration offset (K)
Quiet	Day	A	51	0	479	14.7	1.3	3.7	27	154	–
Quiet	Night	A	41	41	478	2.9	0.7	3.3	1171	1172	−460
Quiet	Dawn	A	2	2	479	4.1	1.0	1.0	1140	1144	−493
Quiet	Dawn	B	87	86	512	6.4	1.0	4.3	676	762	−478
Quiet	Dusk	A	2	0	479	16.1	1.5	2.0	158	303	–
Quiet	Dusk	B	72	0	512	18.3	1.0	3.7	388	423	–
Storm	Day	A	8	8	475	8.1	1.3	2.0	603	637	−280
Storm	Day	B	56	10	509	10.8	2.0	9.0	246	290	−188
Storm	Night	A	4	0	474	20.0	4.0	4.3	289	297	–
Storm	Night	B	23	3	510	22.3	1.3	7.7	484	489	−430
Storm	Dawn	A	61	55	474	6.9	1.7	8.3	564	616	−204
Storm	Dusk	A	45	1	473	18.6	2.0	8.7	268	304	−197

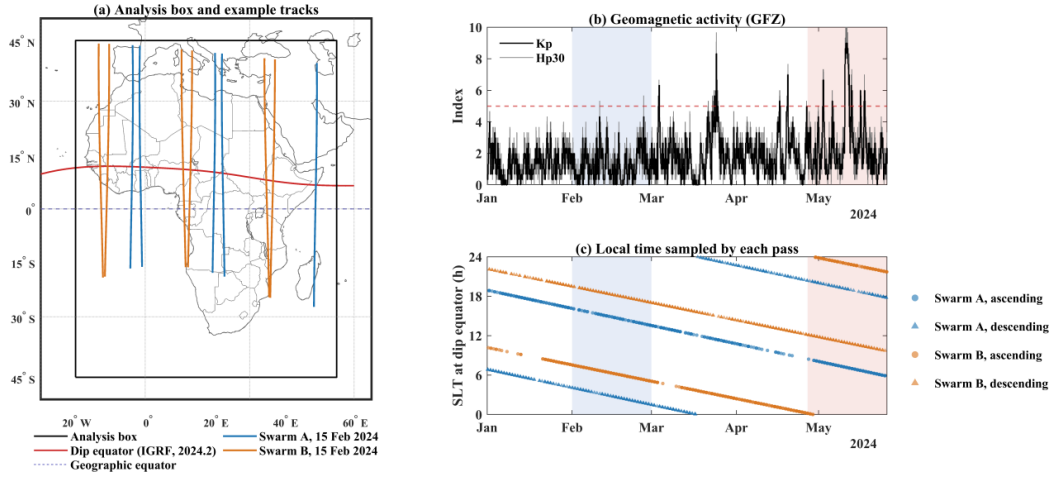


Figure 1. Observation geometry and geomagnetic context. (a) The analysis box over Africa (20°W–55°E, 45°S–45°N), the IGRF dip equator for epoch 2024.2, the geographic equator, and the Swarm A and Swarm B tracks inside the box on 15 February 2024 before trimming to |dip latitude| ≤ 40°. (b) The 3-hourly Kp index and the 30-min Hp30 index from 1 January to 25 May 2024, with the quiet (blue) and storm (red) windows shaded and Kp = 5 marked by the dashed line. (c) Solar local time at the dip-equator crossing of every pass, for the ascending and descending nodes of both satellites.

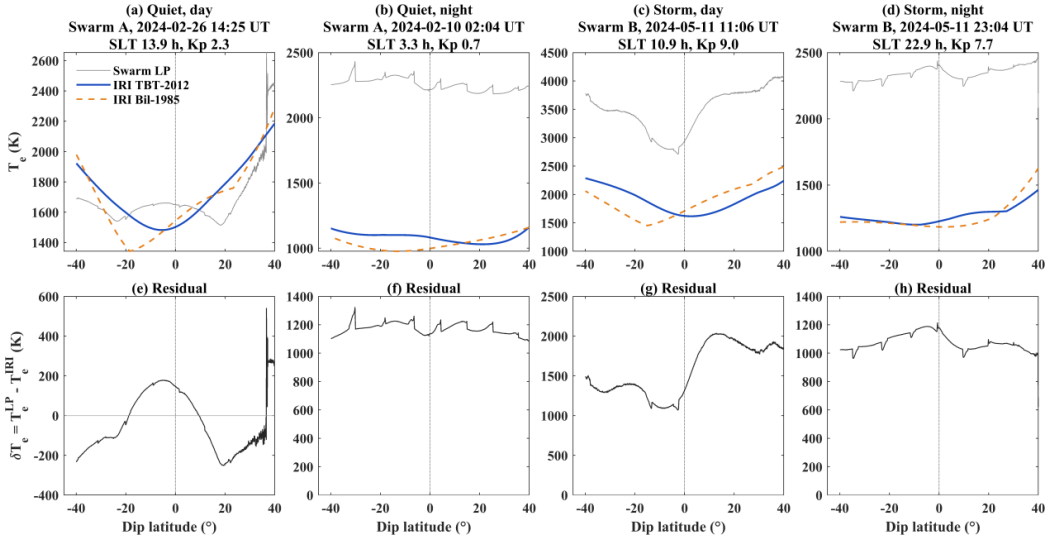
https://doi.org/10.5194/egusphere-2026-6859
Preprint. Discussion started: 19 November 2026
© Author(s) 2026. CC BY 4.0 License.

4.1 Character of the residual

Four representative passes are shown in Fig. 2 and the median residual as a function of dip latitude in Fig. 3. On the day side in the quiet window the residual is small on average: the median pass-mean residual is 27 K and the median pass RMS residual 154 K (Table 1), and its median changes sign along the pass: it is positive between about 22°S and 8°N dip latitude, with a maximum of about 135 K near 6°S, and negative between about 10°N and 28°N, with a minimum of about −145 K near 16°N, and south of about 25°S (Fig. 3a). This structure is consistent with the anticorrelation of electron temperature and density in the daytime topside, where enhanced density over the crests increases electron cooling (Schunk and Nagy, 2009; Pignalberi et al., 2024a; Yan et al., 2022), and with a climatological model that represents this anticorrelation only on average. In the storm window the day-side residual is larger, with a median pass RMS of 324 K against 154 K in the quiet window (rank-sum $p = 2.9 \times 10^{-11}$), and positive over most of the pass, consistent with the storm-time increases of topside electron temperature reported for the same event by Shreedevi et al. (2025).

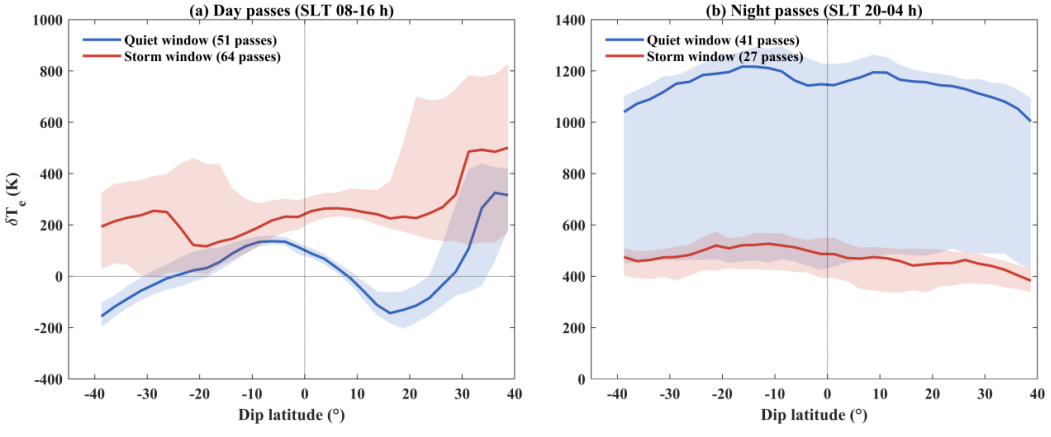
On the night side the residual is dominated by an offset: the median pass-mean residual is 1171 K in the quiet window and 476 K in the storm window (Table 1). This difference is not primarily geophysical. Among the retained night passes, those in the lower Langmuir-probe state have a median pass-mean residual of 459 K in the quiet window and 455 K in the storm window, whereas those in the upper state have 1185 K and 1045 K, and the quiet window happens to contain mostly upper-state passes (25 of 41) and the storm window mostly lower-state passes (24 of 27). The radar-based calibration, where defined, lowers the night-side temperatures by a median of 460 K in the quiet window (Table 1), of the size reported by Lomidze et al. (2018) and Pignalberi et al. (2021); applied to lower-state passes it would bring the residual close to zero, but it cannot by itself remove the 600–900 K difference between the two states. The absolute night-side residual therefore cannot be attributed to IRI with confidence, and the analysis below relies on quantities that are insensitive to a constant offset within a pass.

https://doi.org/10.5194/egusphere-2026-6859
Preprint. Discussion started: 19 November 2026
© Author(s) 2026. CC BY 4.0 License.



355

356 Figure 2. Representative passes. Top row: Langmuir-probe electron temperature (grey), IRI-2020 electron
357 temperature with the TBT-2012 option (blue) and with the Bil-1985 option (orange dashed) against dip
358 latitude. Bottom row: the residual δTe = Te(LP) − Te(IRI, TBT-2012). (a, e) quiet window, day; (b, f)
359 quiet window, night; (c, g) storm window, day; (d, h) storm window, night. The quiet passes are those
360 with the median RMS residual of their group and the storm passes those with the highest Kp of their
361 group.



362

363 Figure 3. Median residual (line) and interquartile range (shading) against dip latitude for all retained
364 passes, in 2.5° bins: (a) day passes (solar local time 08–16 h) and (b) night passes (20–04 h), for the quiet
365 (blue) and storm (red) windows.

https://doi.org/10.5194/egusphere-2026-6859
Preprint. Discussion started: 19 November 2026
© Author(s) 2026. CC BY 4.0 License.

4.2 Phase-space reconstruction

The embedding parameters are summarised in Table 2 and Fig. 4. The delay ranges from 1 to 74 s and the dimension from 3 to 10, with medians of 18–29 s and 3–5 in the different groups. The false-nearest-neighbour fraction fell below 1 % for 388 of the 452 passes; for the remaining 64, most of them night passes, it levelled off at a few per cent and the saturation rule of Sect. 3.4 was used. All but one embedding span lie within one third of the pass. The small dimensions indicate that the along-track variation of the residual over a pass of about 1000 s is captured by a few delay coordinates, while the delays, which correspond to about 130–205 km of ground track at the median ground-track speed of 7.07 km s^{-1}, reflect the scale over which the residual decorrelates. The delay is longest on quiet nights (median 29 s), where the residual varies most smoothly. Phase portraits and recurrence plots of the four representative passes are shown in Fig. 5.

Table 2. Embedding parameters by window and local-time class: median and range of the dimension m and the delay τ, number of passes for which the false-nearest-neighbour fraction fell below 1 % or saturated, and number of passes whose embedding span does not exceed one third of the pass.

Window	LT class	Passes	m median (range)	τ median (range) (s)	FNN < 1 %	FNN saturation	Span ≤ N/3
Quiet	Day	51	4 (3–7)	18 (2–47)	24	27	51
Quiet	Night	41	5 (3–9)	29 (11–61)	32	9	41
Quiet	Dawn	89	4 (3–8)	21 (1–38)	88	1	89
Quiet	Dusk	74	4 (3–9)	22 (1–74)	69	5	74
Storm	Day	64	4 (3–8)	24 (1–59)	61	3	64
Storm	Night	27	4 (3–10)	13 (2–70)	12	15	27
Storm	Dawn	61	3 (3–9)	25 (1–56)	61	0	61
Storm	Dusk	45	4 (3–9)	19 (1–56)	41	4	44

https://doi.org/10.5194/egusphere-2026-6859
Preprint. Discussion started: 19 November 2026
© Author(s) 2026. CC BY 4.0 License.

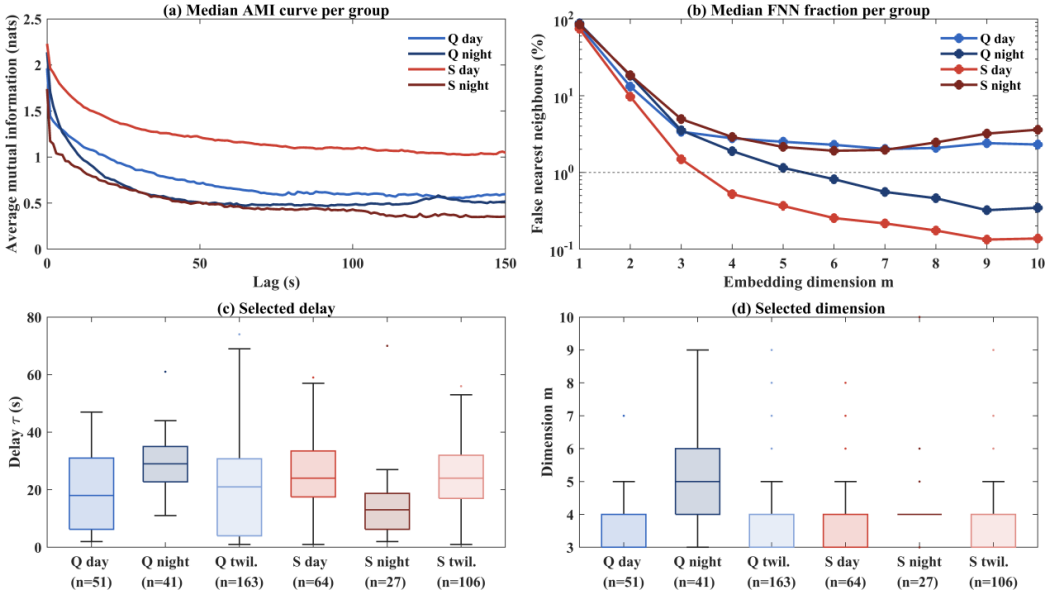


Figure 4. Embedding parameters. (a) Median average mutual information against lag and (b) median false-nearest-neighbour fraction against embedding dimension, for the day and night groups of both windows; the dashed line marks 1 %. (c) Selected delay and (d) selected dimension per pass by group (Q, quiet window; S, storm window; twil., dawn and dusk passes).

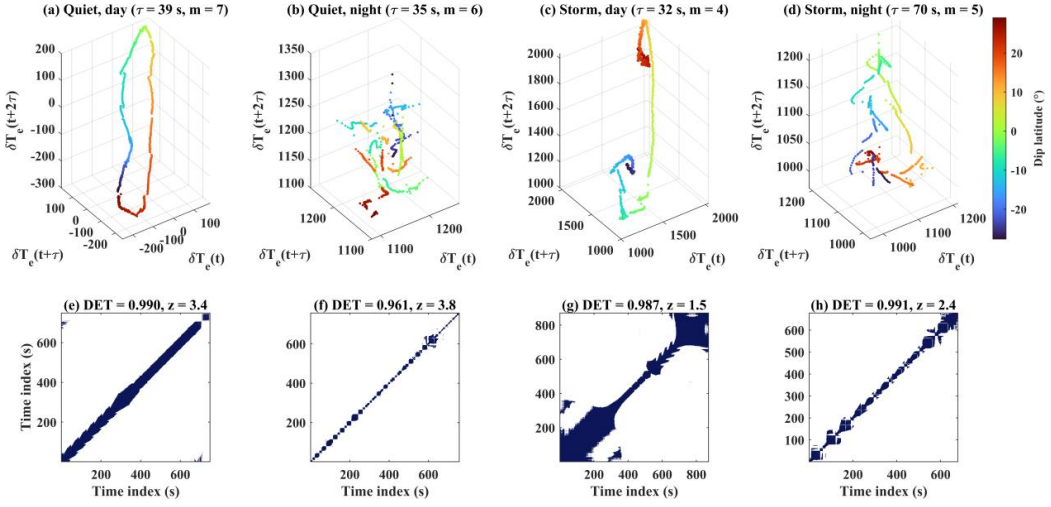


Figure 5. Phase portraits (top row) and recurrence plots (bottom row) of the residual for the four passes of Fig. 2, using the pass-specific delay τ and dimension m; the portraits show δTe(t), δTe(t + τ) and δTe(t + 2τ) coloured by dip latitude, and the recurrence plots use the Chebyshev norm with a threshold of 10 % of the maximum phase-space diameter.

https://doi.org/10.5194/egusphere-2026-6859
Preprint. Discussion started: 19 November 2026
© Author(s) 2026. CC BY 4.0 License.

4.3 Attribution tests: instrumental calibration and IRI temperature option

The attribution tests bear directly on whether the invariants describe the IRI error rather than the instrument or the particular IRI formulation (Table 3, Fig. 6). Replacing the TBT-2012 topside electron temperature by the Bil-1985 option, with the embedding fixed, changed the determinism by a median of 0.007 over all retained passes and by at most 0.20, and the group medians of the change lie between 0.15 and 0.63 of the between-pass interquartile range of determinism in the same group; for the quiet-window dusk passes, whose between-pass spread is very small (interquartile range 0.0025), the ratio is 1.36. The corresponding change of the Lyapunov exponent has group medians of 0.25–0.76 of the between-pass interquartile range. Applying the radar-based calibration, on the 206 retained passes where it is defined throughout, changed the determinism by a median of 0.006 and by at most 0.14, with group medians of 0.14–0.32 of the between-pass interquartile range, and changed the Lyapunov exponent by 0.10–0.23 of its interquartile range, in the groups with more than three such passes.

Two conclusions follow. First, the instrumental calibration, the analogue here of the plasmaspheric term in the TEC study, perturbs the invariants by a fraction of the between-pass structure, consistent with the bound of Eq. (3); the perturbation is, however, larger than its TEC counterpart, for which the median perturbation was about one twentieth of the between-station spread (Edward, 2026). Second, the choice of IRI topside temperature model matters more than the calibration: for some passes the change of determinism is comparable to the spread between passes, so the invariants reported below describe the residual with respect to IRI-2020 in its default TBT-2012 configuration and should not be transferred to the Bil-1985 option without recomputation. Because the calibration is defined over only 3 of the 27 retained storm-window night passes, the calibration test is informative mainly for quiet conditions.

https://doi.org/10.5194/egusphere-2026-6859
Preprint. Discussion started: 19 November 2026
© Author(s) 2026. CC BY 4.0 License.

Table 3. Attribution tests. For each window and local-time class, the median and maximum absolute change of determinism and of the Lyapunov exponent when the IRI topside temperature option is changed from TBT-2012 to Bil-1985, and when the radar-based Langmuir-probe calibration is applied, with the embedding fixed, compared with the between-pass interquartile range of the reference quantity in the same group.

Window	LT class	Control	Passes	|ΔDET| median (max)	Between-pass IQR of DET	Ratio	$\|\Delta\lambda_1\|$ median (max) (10^{-3} s^{-1})	Between-pass IQR of λ_1 (10^{-3} s^{-1})	Ratio
Quiet	Day	IRI Bil-1985	51	0.009 (0.049)	0.038	0.25	3.57 (15.63)	14.25	0.25
Quiet	Night	IRI Bil-1985	41	0.009 (0.202)	0.062	0.15	2.19 (14.09)	6.15	0.36
Quiet	Night	LP calibration	41	0.010 (0.140)	0.062	0.17	1.44 (8.03)	6.15	0.23
Quiet	Dawn	IRI Bil-1985	89	0.009 (0.137)	0.018	0.47	3.06 (23.99)	11.29	0.27
Quiet	Dawn	LP calibration	88	0.006 (0.057)	0.018	0.32	1.12 (20.10)	11.29	0.10
Quiet	Dusk	IRI Bil-1985	74	0.003 (0.076)	0.003	1.36	3.56 (22.70)	10.14	0.35
Storm	Day	IRI Bil-1985	64	0.004 (0.039)	0.014	0.32	4.19 (54.78)	7.55	0.55
Storm	Day	LP calibration	18	0.002 (0.019)	0.014	0.14	1.12 (10.04)	7.55	0.15
Storm	Night	IRI Bil-1985	27	0.014 (0.078)	0.054	0.25	4.57 (23.65)	13.38	0.34
Storm	Night	LP calibration	3	0.006 (0.007)	0.054	0.11	2.56 (2.62)	13.38	0.19
Storm	Dawn	IRI Bil-1985	61	0.013 (0.093)	0.021	0.63	4.77 (21.28)	6.28	0.76
Storm	Dawn	LP calibration	55	0.004 (0.056)	0.021	0.17	1.14 (12.37)	6.28	0.18
Storm	Dusk	IRI Bil-1985	45	0.008 (0.095)	0.022	0.36	4.99 (29.12)	7.40	0.67
Storm	Dusk	LP calibration	1	0.012 (0.012)	0.022	0.55	0.92 (0.92)	7.40	0.12

https://doi.org/10.5194/egusphere-2026-6859
Preprint. Discussion started: 19 November 2026
© Author(s) 2026. CC BY 4.0 License.

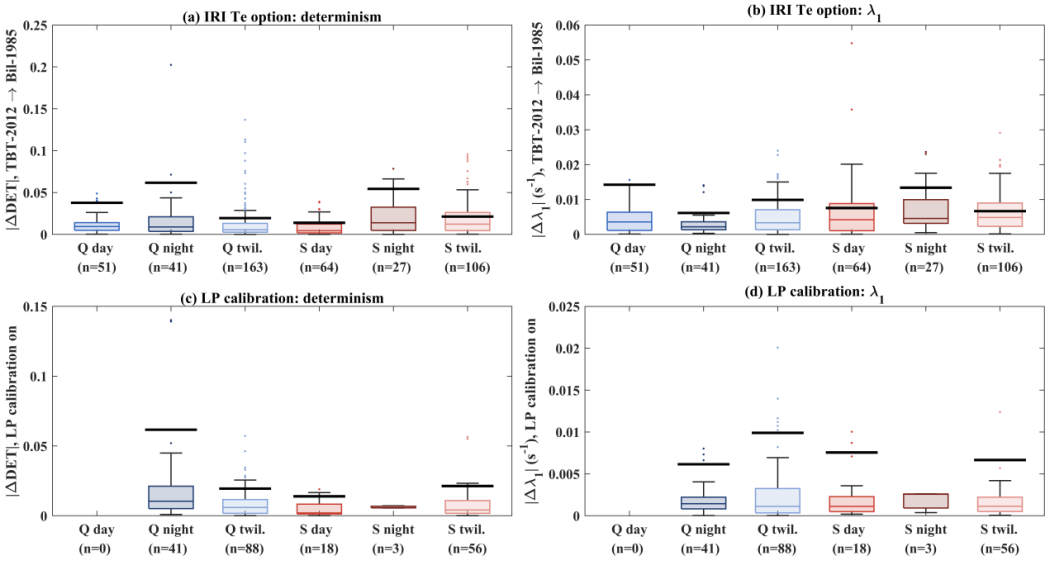


Figure 6. Attribution tests. Absolute change of (a, c) determinism and (b, d) the maximal Lyapunov exponent when (a, b) the IRI topside electron temperature option is changed from TBT-2012 to Bil-1985 and (c, d) the radar-based Langmuir-probe calibration is applied, with the embedding fixed. Black bars mark the between-pass interquartile range of the reference quantity in each group with at least four passes.

4.4 Recurrence quantification and surrogate testing

Raw determinism is high in every group, with group medians of 0.938–0.998 (Table 4, Fig. 7), and it does not discriminate between conditions in a physically interpretable way. On the day side, determinism is higher in the storm window than in the quiet window (medians 0.992 and 0.963, rank-sum $p = 3.1 \times 10^{-11}$, Table 5), whereas the standardised excess over the IAAFT surrogates moves in the opposite direction (Table 6, Fig. 8): the median z_{DET} is 5.31 in the quiet window and 3.30 in the storm window ($p = 6.2 \times 10^{-8}$), and across day passes z_{DET} decreases weakly with Kp (Spearman $r_s = -0.19$, $p = 0.042$). As in the TEC study (Edward, 2026), a higher raw determinism under storm forcing reflects stronger, smoother large-scale structure in the residual rather than a larger nonlinear content, and the surrogate excess, not raw determinism, is the quantity to interpret.

The surrogate test rejects the linear stochastic null hypothesis for 396 of the 452 passes with the fixed-threshold convention and for 323 with the adaptive convention. Rejection is nearly universal on the day side (51 of 51 quiet and 62 of 64 storm passes) and frequent on the night

https://doi.org/10.5194/egusphere-2026-6859
Preprint. Discussion started: 19 November 2026
© Author(s) 2026. CC BY 4.0 License.

side (36 of 41 and 25 of 27), but less frequent at dawn, where it holds for 63 of 89 quiet and 46 of 61 storm passes with the fixed threshold and for only 27 and 30 with the adaptive threshold (Table 6). The median z_{DET} is lowest in the twilight classes (2.22–2.69). On the night side, the quiet and storm medians (3.78 and 4.76) do not differ significantly ($p = 0.054$), and within the lower probe state the storm-window median is the larger (5.15 against 2.84); the night-side comparison is, however, limited by the small number of retained storm-window night passes. The residual therefore contains deterministic structure beyond its linear correlation in most passes, most clearly on the quiet day side, and storm forcing reduces the day-side excess without removing it.

Table 4. Recurrence quantification measures by window and local-time class (medians, with the interquartile range of determinism).

Window	LT class	Passes	DET median (IQR)	LAM	TT	ENTR	RR
Quiet	Day	51	0.963 (0.941–0.979)	0.977	29.4	2.87	0.157
Quiet	Night	41	0.962 (0.928–0.989)	0.978	10.5	2.91	0.065
Quiet	Dawn	89	0.982 (0.969–0.987)	0.986	37.6	3.54	0.166
Quiet	Dusk	74	0.998 (0.997–0.999)	0.999	140.6	5.01	0.480
Storm	Day	64	0.992 (0.983–0.997)	0.995	45.1	3.79	0.184
Storm	Night	27	0.938 (0.917–0.972)	0.965	20.5	2.84	0.217
Storm	Dawn	61	0.977 (0.965–0.986)	0.983	28.8	3.43	0.139
Storm	Dusk	45	0.984 (0.971–0.994)	0.990	37.0	3.65	0.236

https://doi.org/10.5194/egusphere-2026-6859
Preprint. Discussion started: 19 November 2026
© Author(s) 2026. CC BY 4.0 License.

Table 5. Quiet–storm comparisons within each local-time class (Wilcoxon rank-sum test) and Spearman correlations with the Kp index of the pass.

Test	LT class	Quantity	n	Median quiet	Median storm	r_s	p
Rank-sum, quiet vs storm	Day	DET	115	0.963	0.992	–	3.1×10^{-11}
Spearman vs Kp	Day	DET	115	–	–	0.22	0.016
Rank-sum, quiet vs storm	Day	z_{DET}	115	5.306	3.297	–	6.2×10^{-8}
Spearman vs Kp	Day	z_{DET}	115	–	–	−0.19	0.042
Rank-sum, quiet vs storm	Day	λ_1	115	0.013	0.006	–	3.9×10^{-6}
Spearman vs Kp	Day	λ_1	115	–	–	−0.16	0.090
Rank-sum, quiet vs storm	Day	RMS δT_e	115	154.300	323.650	–	2.9×10^{-11}
Spearman vs Kp	Day	RMS δT_e	115	–	–	0.03	0.719
Rank-sum, quiet vs storm	Night	DET	68	0.962	0.938	–	0.126
Spearman vs Kp	Night	DET	68	–	–	−0.02	0.842
Rank-sum, quiet vs storm	Night	z_{DET}	68	3.784	4.758	–	0.054
Spearman vs Kp	Night	z_{DET}	68	–	–	−0.18	0.131
Rank-sum, quiet vs storm	Night	λ_1	68	0.005	0.011	–	2.1×10^{-4}
Spearman vs Kp	Night	λ_1	68	–	–	−0.04	0.745
Rank-sum, quiet vs storm	Night	RMS δT_e	68	1171.600	478.300	–	1.7×10^{-4}
Spearman vs Kp	Night	RMS δT_e	68	–	–	−0.43	2.6×10^{-4}
Rank-sum, quiet vs storm	Dawn	DET	150	0.982	0.977	–	0.217
Spearman vs Kp	Dawn	DET	150	–	–	−0.12	0.142
Rank-sum, quiet vs storm	Dawn	z_{DET}	150	2.224	2.525	–	0.191
Spearman vs Kp	Dawn	z_{DET}	150	–	–	0.05	0.547
Rank-sum, quiet vs storm	Dawn	λ_1	150	0.011	0.009	–	0.032
Spearman vs Kp	Dawn	λ_1	150	–	–	−0.02	0.825
Rank-sum, quiet vs storm	Dawn	RMS δT_e	150	774.900	616.500	–	0.013
Spearman vs Kp	Dawn	RMS δT_e	150	–	–	−0.07	0.407
Rank-sum, quiet vs storm	Dusk	DET	119	0.998	0.984	–	3.7×10^{-10}
Spearman vs Kp	Dusk	DET	119	–	–	−0.40	6.1×10^{-6}
Rank-sum, quiet vs storm	Dusk	z_{DET}	119	2.608	2.695	–	0.359
Spearman vs Kp	Dusk	z_{DET}	119	–	–	0.14	0.142
Rank-sum, quiet vs storm	Dusk	λ_1	119	0.008	0.009	–	0.574
Spearman vs Kp	Dusk	λ_1	119	–	–	0.16	0.074
Rank-sum, quiet vs storm	Dusk	RMS δT_e	119	422.100	304.200	–	1.4×10^{-7}
Spearman vs Kp	Dusk	RMS δT_e	119	–	–	−0.27	0.003

https://doi.org/10.5194/egusphere-2026-6859
Preprint. Discussion started: 19 November 2026
© Author(s) 2026. CC BY 4.0 License.

Table 6. Surrogate test: median and interquartile range of the standardised determinism excess z_{DET}, number of passes for which the null hypothesis is rejected at the 5 % level with the fixed-threshold and adaptive-threshold conventions, and median standardised excess of laminarity.

Window	LT class	Passes	z_{DET} median (IQR)	Significant (fixed ε)	z_{DET} median, adaptive ε	Significant (adaptive ε)	z_{LAM} median
Quiet	Day	51	5.31 (3.60 to 8.11)	51	2.24	49	5.45
Quiet	Night	41	3.78 (2.74 to 5.01)	36	1.92	32	3.64
Quiet	Dawn	89	2.22 (1.34 to 2.77)	63	0.75	27	1.91
Quiet	Dusk	74	2.61 (1.74 to 3.21)	73	1.20	74	2.57
Storm	Day	64	3.30 (2.66 to 3.87)	62	2.02	56	3.32
Storm	Night	27	4.76 (3.10 to 7.83)	25	1.92	23	5.12
Storm	Dawn	61	2.52 (1.28 to 3.19)	46	0.99	30	2.49
Storm	Dusk	45	2.69 (2.06 to 3.70)	40	1.23	32	2.80

https://doi.org/10.5194/egusphere-2026-6859
Preprint. Discussion started: 19 November 2026
© Author(s) 2026. CC BY 4.0 License.

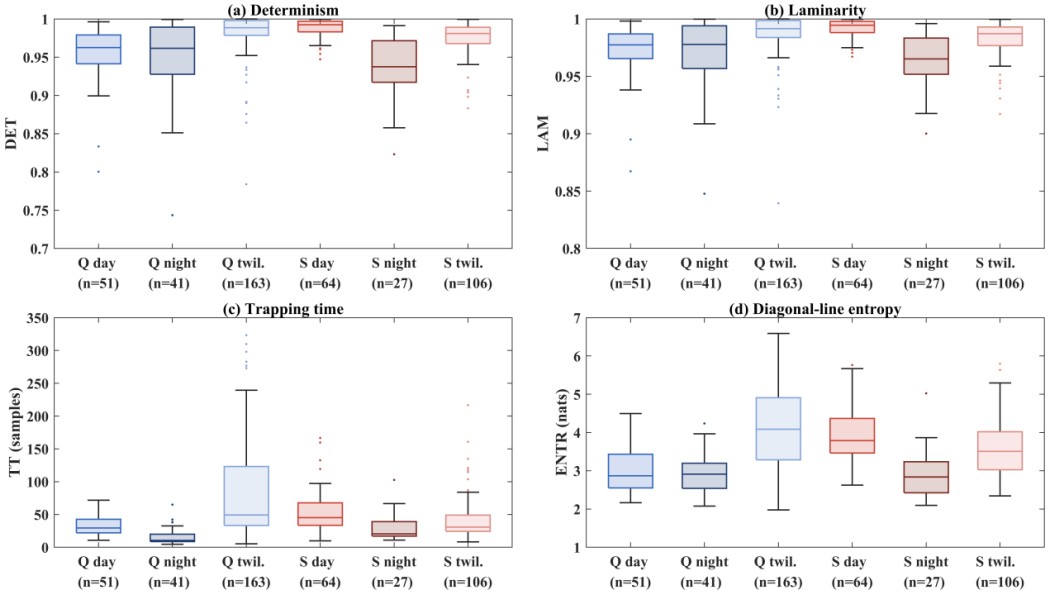


Figure 7. Recurrence quantification measures per pass by group: (a) determinism, (b) laminarity, (c) trapping time and (d) diagonal-line entropy. Group labels as in Fig. 4, with the number of passes in brackets.

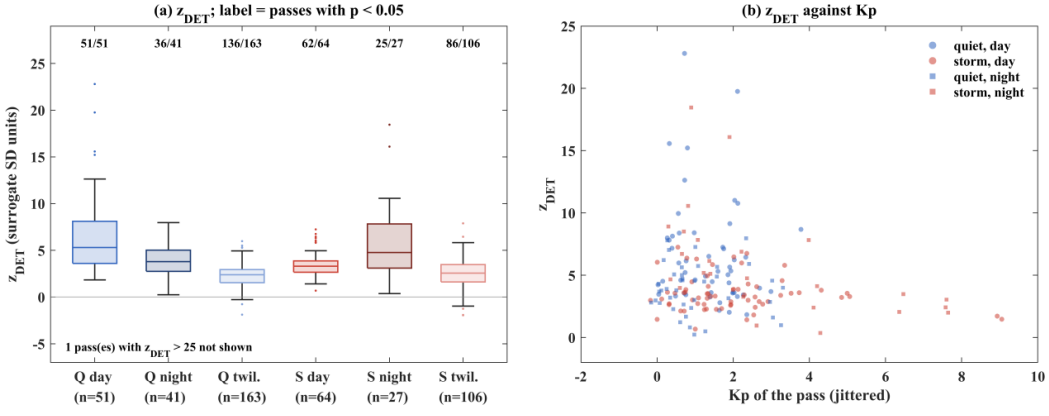


Figure 8. Surrogate testing. (a) Standardised excess of determinism over 100 IAAFT surrogates by group, with the number of passes for which the null hypothesis is rejected at the 5 % level above each box. (b) The excess against the Kp index of the pass for day and night passes of both windows; Kp is jittered horizontally for legibility. Passes outside the plotted range are counted in panel (a).

https://doi.org/10.5194/egusphere-2026-6859
Preprint. Discussion started: 19 November 2026
© Author(s) 2026. CC BY 4.0 License.

4.5 Dip-latitude organisation inside the passes

The row-resolved determinism of Eq. (6) shows how the excess is distributed along the pass (Fig. 9, Table 7). On the quiet day side the median local excess is nearly uniform between 37°S and 17°N dip latitude (1.7–2.1) and declines towards the northern end of the pass (1.0–1.4 poleward of 22°N), and its rank correlation with absolute dip latitude is not significant (r_s = −0.17, p = 0.53). The clearest organisations are found in the twilight sectors and on the storm day side. On quiet-window dawn passes the local excess is lowest between the dip equator and 17°N (0.3–0.4) and highest poleward of 27° in both hemispheres (0.9), giving r_s = 0.78 (p = 3.5 × 10^{-4}) between absolute dip latitude and median local excess. On quiet-window dusk passes and on storm-window day passes the gradient has the opposite sign (r_s = −0.69, p = 0.003, and −0.65, p = 0.006), with the larger excess at low dip latitudes. On storm-window night passes the local excess falls to 0.4–0.7 within 5° of the dip equator against 1.6–2.1 over the flanks, although the correlation with absolute dip latitude is not significant over the full range (r_s = 0.39, p = 0.13). Because each pass covers both hemispheres, the northern and southern medians can be compared directly; they agree to within 0.4 in every group (Table 7).

These patterns are consistent with the deterministic content of the IRI error being organised by the equatorial electrodynamics that shape the anomaly, since the sense of the gradient changes between dawn, dusk and storm-time day conditions, in which the vertical plasma drift and the anomaly development differ (Balan et al., 2018). They do not by themselves identify a mechanism, and the local excess is noisier than the whole-pass excess because each 5° bin contains only about 60 delay vectors.

https://doi.org/10.5194/egusphere-2026-6859
Preprint. Discussion started: 19 November 2026
© Author(s) 2026. CC BY 4.0 License.

Table 7. Dip-latitude organisation of the local determinism excess: Spearman rank correlation between absolute dip latitude and the median local z_{DET} over the 5° bins with at least three passes, and the median local excess in the northern and southern halves of the pass.

Window	LT class	Dip bins	r_S (|dip|, local z_{DET})	p	Median local z_{DET}, north	Median local z_{DET}, south
Quiet	Day	16	−0.17	0.525	1.61	1.95
Quiet	Night	16	0.39	0.135	1.38	1.31
Quiet	Dawn	16	0.78	3.5×10^{-4}	0.55	0.77
Quiet	Dusk	16	−0.69	0.003	0.68	0.79
Storm	Day	16	−0.65	0.006	1.27	1.12
Storm	Night	16	0.39	0.135	1.71	1.66
Storm	Dawn	16	0.20	0.455	0.98	0.79
Storm	Dusk	16	−0.40	0.122	0.92	0.91

https://doi.org/10.5194/egusphere-2026-6859
Preprint. Discussion started: 19 November 2026
© Author(s) 2026. CC BY 4.0 License.

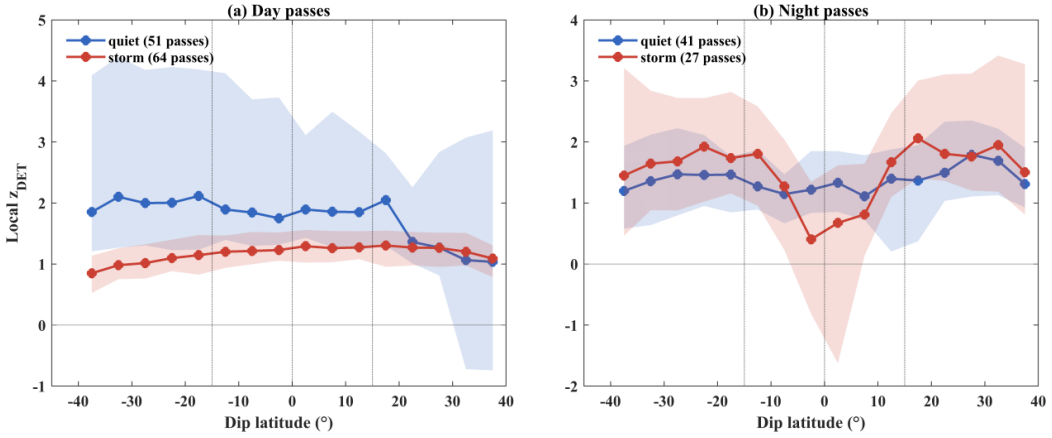


Figure 9. Dip-latitude organisation of the determinism excess. Median (line) and interquartile range (shading) of the local excess of Eq. (6) in 5° dip-latitude bins for (a) day and (b) night passes of the quiet (blue) and storm (red) windows. Dotted vertical lines mark the dip equator and ±15° dip latitude.

4.6 Lyapunov exponents and along-track horizons

The maximal Lyapunov exponent is positive for 438 of the 452 passes, with group medians between 0.0054 and 0.0126 s^{-1} (Table 8, Fig. 10). It is largest on the quiet day side (median 0.0126 s^{-1}) and smaller on the storm day side (0.0062 s^{-1}, $p = 3.9 \times 10^{-6}$), whereas on the night side it is smaller in the quiet window (0.0054 s^{-1}) than in the storm window (0.0106 s^{-1}, $p = 2.1 \times 10^{-4}$); within the lower probe state the night-side medians are 0.0030 and 0.0112 s^{-1}. Only 51 exponents exceed those of their surrogates at the 5 % level, and these are concentrated in three groups: 24 of the 51 quiet day passes, 8 of the 41 quiet night passes and 12 of the 27 storm night passes. The surrogate-supported along-track horizons $1/\lambda_1$ have medians of 46 s on the quiet day side, 104 s on quiet nights and 47 s on storm nights, corresponding to about 330, 740 and 340 km of ground track, or 3.7°, 8.4° and 4.0° of dip latitude. In the other groups the exponent cannot be distinguished from that of a linear process with the same spectrum.

In physical terms, two points of the residual that are close in the reconstructed state space at one point of the pass remain close for only tens of seconds of flight, a few degrees of dip latitude, before their separation grows at the rate λ_1. This is a statement about the along-track geometry of the IRI error at a given time, not about its persistence in time; the temporal predictability of the residual is assessed directly by the prediction experiment of Sect. 4.8.

https://doi.org/10.5194/egusphere-2026-6859
Preprint. Discussion started: 19 November 2026
© Author(s) 2026. CC BY 4.0 License.

Table 8. Maximal Lyapunov exponent: median and interquartile range, number of passes for which the exponent exceeds that of the surrogates at the 5 % level, and median and range of the supported along-track horizons, expressed in seconds, in kilometres of ground track and in degrees of dip latitude.

Window	LT class	Passes	λ_1 median (IQR) (10^{-3} s^{-1})	Supported	Horizon median (range) (s)	Horizon median (km)	Horizon median (° dip)
Quiet	Day	51	12.62 (7.57–21.81)	24	46 (28–215)	329	3.7
Quiet	Night	41	5.38 (2.54–8.69)	8	104 (56–313)	742	8.4
Quiet	Dawn	89	10.91 (7.08–18.38)	1	63 (63–63)	445	5.3
Quiet	Dusk	74	8.18 (4.22–14.36)	1	24 (24–24)	172	1.8
Storm	Day	64	6.22 (2.27–9.82)	1	27 (27–27)	193	2.1
Storm	Night	27	10.56 (6.51–19.90)	12	47 (36–106)	336	4.0
Storm	Dawn	61	9.00 (6.91–13.18)	2	65 (58–71)	461	5.2
Storm	Dusk	45	9.35 (6.14–13.54)	2	37 (28–46)	264	3.0

https://doi.org/10.5194/egusphere-2026-6859
Preprint. Discussion started: 19 November 2026
© Author(s) 2026. CC BY 4.0 License.

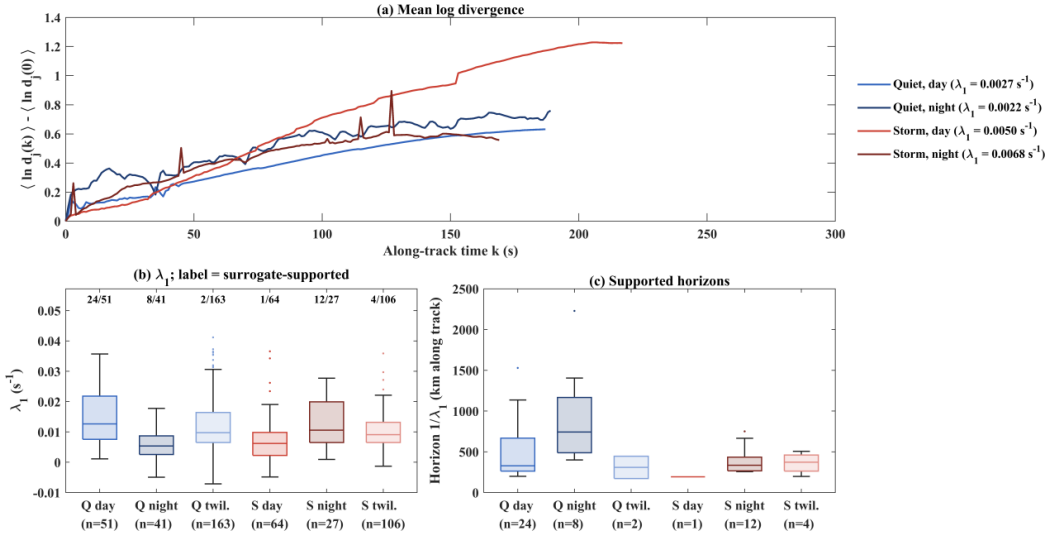


Figure 10. Lyapunov analysis. (a) Mean logarithmic divergence of nearest neighbours against along-track time for the four passes of Fig. 2, offset to zero at k = 0. (b) Maximal Lyapunov exponent by group, with the number of exponents exceeding those of the surrogates at the 5 % level above each box. (c) Supported along-track horizons 1/λ1 expressed as ground distance.

4.7 Robustness of the invariants

The controls of Sect. 3.5 are summarised in Table 9 and Fig. 11. Imposing a common embedding (m = 4, τ = 24 s, the medians of all passes) preserves the ranking of determinism between passes well (r_s = 0.68–0.96 in the different groups), and block-averaging to a 3-s cadence preserves it moderately (r_s = 0.36–0.76). The first difference of the residual changes determinism substantially, and on dawn passes, whose residual is smooth, it reduces the median from 0.98 to 0.29–0.47; the rank correlation with the reference determinism is weak (r_s from −0.20 to 0.47). This is expected, because differencing removes the large-scale structure that dominates the recurrence plots and amplifies the measurement noise, but it shows that raw determinism is not a stable property of the residual under pre-filtering. The Lyapunov exponent is the least robust quantity: its rank correlation with the reference under the three controls ranges from −0.31 to 0.72. The determinism excess and its organisation are therefore the more reliable results of this

https://doi.org/10.5194/egusphere-2026-6859
Preprint. Discussion started: 19 November 2026
© Author(s) 2026. CC BY 4.0 License.

study, and the Lyapunov exponents and horizons are reported as indicative values for the groups in which they are supported by the surrogate test.

Repeating the statistics with all 627 passes, including those with state switches, leaves the day-side and twilight results essentially unchanged: the quiet and storm day-side medians of z_{DET} become 5.13 and 3.31 ($p = 1.0 \times 10^{-8}$), and those of λ_1 0.0133 and 0.0056 s^{-1}. The night side is affected: the median z_{DET} falls to 2.79 in the quiet window and 2.19 in the storm window, the quiet–storm difference of λ_1 is no longer significant ($p = 0.076$), and the number of supported night-side horizons rises to 13 and 20. The state switches thus act on the night-side invariants in the way expected from Eq. (3) for a contaminant that is not slowly varying, and they would have been misattributed to the IRI error without the quality control.

https://doi.org/10.5194/egusphere-2026-6859
Preprint. Discussion started: 19 November 2026
© Author(s) 2026. CC BY 4.0 License.

Table 9. Robustness controls: median determinism and Lyapunov exponent under the first-difference, 3-s cadence and common-embedding controls compared with the reference, and the Spearman rank correlation between control and reference across passes.

Window	LT class	Control	Passes	DET median: control / reference	r_s DET	λ_1 median: control / reference (10^{-3} s^{-1})	r_s λ_1
Quiet	Day	First difference	51	0.927 / 0.963	−0.20	7.28 / 12.62	−0.12
Quiet	Day	3-s cadence	51	0.977 / 0.963	0.76	6.51 / 12.62	0.27
Quiet	Day	Common embedding	51	0.957 / 0.963	0.87	9.85 / 12.62	−0.20
Quiet	Night	First difference	41	0.962 / 0.962	0.30	10.47 / 5.38	0.26
Quiet	Night	3-s cadence	41	0.921 / 0.962	0.74	3.37 / 5.38	0.33
Quiet	Night	Common embedding	41	0.968 / 0.962	0.95	9.06 / 5.38	0.26
Quiet	Dawn	First difference	89	0.293 / 0.982	0.18	4.75 / 10.91	−0.31
Quiet	Dawn	3-s cadence	89	0.992 / 0.982	0.36	6.33 / 10.91	0.46
Quiet	Dawn	Common embedding	89	0.977 / 0.982	0.68	7.73 / 10.91	0.27
Quiet	Dusk	First difference	74	0.987 / 0.998	0.35	5.28 / 8.18	−0.13
Quiet	Dusk	3-s cadence	74	0.997 / 0.998	0.53	3.42 / 8.18	0.15
Quiet	Dusk	Common embedding	74	0.998 / 0.998	0.85	6.89 / 8.18	0.14
Storm	Day	First difference	64	0.976 / 0.992	0.47	4.58 / 6.22	0.01
Storm	Day	3-s cadence	64	0.994 / 0.992	0.67	3.44 / 6.22	0.61
Storm	Day	Common embedding	64	0.990 / 0.992	0.92	7.46 / 6.22	0.72
Storm	Night	First difference	27	0.961 / 0.938	0.09	9.55 / 10.56	0.12
Storm	Night	3-s cadence	27	0.924 / 0.938	0.67	6.27 / 10.56	0.03
Storm	Night	Common embedding	27	0.928 / 0.938	0.92	4.75 / 10.56	0.18
Storm	Dawn	First difference	61	0.475 / 0.977	0.03	4.55 / 9.00	−0.25
Storm	Dawn	3-s cadence	61	0.989 / 0.977	0.42	7.23 / 9.00	0.48
Storm	Dawn	Common embedding	61	0.977 / 0.977	0.96	8.64 / 9.00	0.50
Storm	Dusk	First difference	45	0.897 / 0.984	0.15	6.29 / 9.35	0.40
Storm	Dusk	3-s cadence	45	0.993 / 0.984	0.48	3.94 / 9.35	0.16
Storm	Dusk	Common embedding	45	0.983 / 0.984	0.70	7.22 / 9.35	0.23

https://doi.org/10.5194/egusphere-2026-6859
Preprint. Discussion started: 19 November 2026
© Author(s) 2026. CC BY 4.0 License.

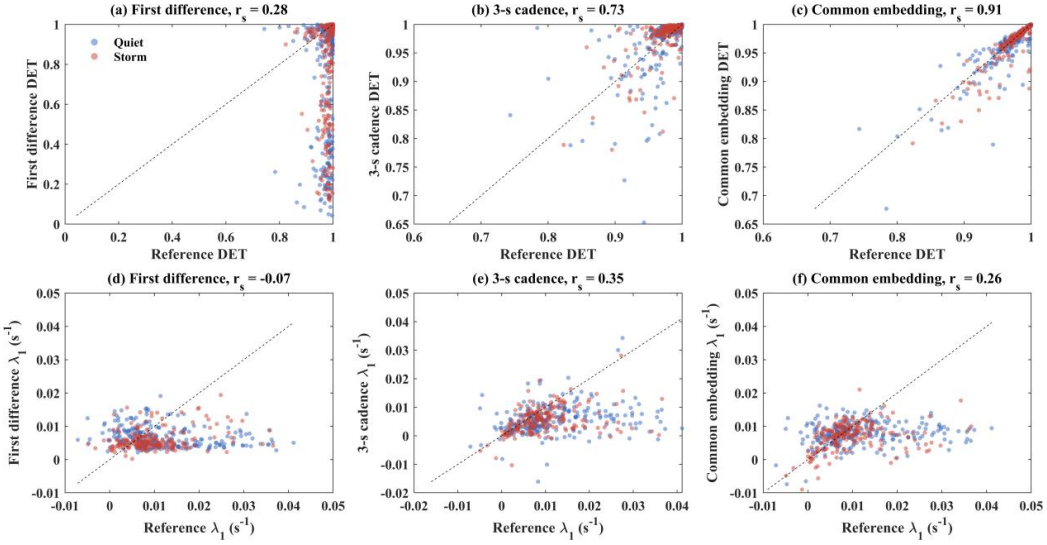


Figure 11. Robustness controls. Determinism (top row) and maximal Lyapunov exponent (bottom row) of each pass under (a, d) the first difference of the residual, (b, e) block-averaging to a 3-s cadence and (c, f) a common embedding (m = 4, τ = 24 s), against the reference values; the Spearman rank correlation over all passes is given in each title and the dashed line is the identity.

4.8 Out-of-sample prediction of the residual

The prediction experiment verified 845 passes free of state switches on 115 days between 1 February and 25 May 2024 (Table 10, Fig. 12). IRI-2020 alone has an RMSE of 559 K, and its mean error of −415 K shows that most of this is an offset between the model and the uncalibrated Langmuir-probe temperature, largest at night (mean error −681 K) and smallest on the day side (−169 K). Every model that uses past residuals removes most of this offset. The 60-day climatology CLIM60 reduces the RMSE to 271 K, a mean-square-error skill of 0.76 relative to IRI, and the previous-day profile PER1 (276 K), the driver-only tree ensemble TREE-D (276 K) and the 7-day bias BIAS7 (300 K) perform similarly or worse. Because a correction of the mean offset is already provided by the simplest benchmarks, CLIM60 rather than IRI is the reference against which additional skill is judged.

Against CLIM60, the tree ensemble with residual history, TREE-H, reduces the RMSE to 252 K, a skill of 0.14 with a 95 % day-block bootstrap interval of 0.02–0.23 (Table 11); the daily mean-square errors do not, however, differ significantly from those of CLIM60 by the signed-rank test ($p = 0.11$, 0.15 after false-discovery-rate control, Table 12). The gain is concentrated at night

https://doi.org/10.5194/egusphere-2026-6859
Preprint. Discussion started: 19 November 2026
© Author(s) 2026. CC BY 4.0 License.

(skill 0.32, interval 0.11–0.51) and on the storm-window day side (0.29, interval 0.08–0.46), while on the quiet-window day side TREE-H is worse than CLIM60 (−0.71, interval −1.63 to −0.06). The predictor importance of TREE-H is dominated by the previous-day profile (normalised importance 0.84) followed by the 7-day bias (0.04), with all geophysical and driver features below 0.02 (Sect. 4.9), so the skill of TREE-H derives mainly from persistence of the residual from one day to the next, combined in a way that adapts to local time and dip latitude. The nowcast variant TREE-HN, which in addition uses the concurrently measured ion density, reduces the RMSE to 212 K, a skill of 0.39 (interval 0.27–0.48) relative to CLIM60, significant in the daily comparison ($p = 1.3 \times 10^{-5}$, 3.9×10^{-5} after control); its gain is largest on the day side (0.42) and in the twilight sectors (0.43–0.50). This is consistent with the strong dependence of the topside electron temperature on the local density (Pignalberi et al., 2024a; Yan et al., 2022), which the climatological IRI temperature formulation does not use, but it requires a density measurement at the time and place of the prediction and is therefore a nowcast rather than a forecast.

The dependence on geomagnetic activity is shown in Fig. 13a and Table 13. For passes with Kp below 3 (711 passes), CLIM60 reaches 235 K and TREE-H adds a skill of only 0.07. For $3 \leq Kp < 5$ (102 passes) the RMSE of CLIM60 rises to 335 K and the skill of TREE-H increases to 0.36. For $Kp \geq 5$ (32 passes on 11 days) all models degrade: IRI reaches 795 K, CLIM60 607 K, TREE-H 562 K (skill 0.14) and TREE-HN 397 K (0.57). Along the pass, the errors of all models increase towards the northern end of the storm-window passes and peak near the dip equator in the quiet window, where the IRI error itself is largest (Fig. 13b, c). During the superstorm of 10–12 May (Fig. 14), the observed pass-mean residual rose on 11 May to between about 800 and 1600 K, and the pass RMSE of IRI, CLIM60 and PER1 reached about 800–1600 K; none of the models that rely on previous days anticipated the increase, and TREE-H reduced the error only partly. TREE-HN, which uses the storm-time density, reduced the pass RMSE on that day to about 450–800 K. The prediction of the residual therefore retains useful skill relative to climatology in moderately disturbed conditions, but the onset of a severe storm is not predicted by any model that uses only past residuals and the geomagnetic indices available at the time.

https://doi.org/10.5194/egusphere-2026-6859
Preprint. Discussion started: 19 November 2026
© Author(s) 2026. CC BY 4.0 License.

Repeating the experiment with all passes, including those with state switches, gives skills relative to CLIM60 of 0.19 for TREE-H and 0.46 for TREE-HN, close to the values above; the main effect of the state switches is to inflate the night-side errors of every model.

Table 10. Out-of-sample prediction by window and local-time class: number of verified passes and days, RMSE of IRI-2020 alone, and RMSE of each model with its mean-square-error skill relative to IRI and relative to CLIM60.

Window	LT class	Passes	Days	IRI RMSE (K)	BIAS7 RMSE (SS vs IRI; SS vs CLIM60)	CLIM60 RMSE (SS vs IRI; SS vs CLIM60)	PER1 RMSE (SS vs IRI; SS vs CLIM60)	TREE-D RMSE (SS vs IRI; SS vs CLIM60)	TREE-H RMSE (SS vs IRI; SS vs CLIM60)	TREE-HN RMSE (SS vs IRI; SS vs CLIM60)
Quiet	All	255	29	657	313 (0.77; −0.52)	254 (0.85; 0.00)	223 (0.89; 0.23)	255 (0.85; −0.01)	221 (0.89; 0.24)	186 (0.92; 0.46)
Quiet	Day	51	25	194	183 (0.11; −0.46)	151 (0.39; 0.00)	166 (0.27; −0.20)	158 (0.34; −0.08)	198 (−0.04; −0.71)	171 (0.22; −0.28)
Quiet	Night	41	21	977	371 (0.86; 0.08)	387 (0.84; 0.00)	254 (0.93; 0.57)	390 (0.84; −0.01)	250 (0.93; 0.58)	243 (0.94; 0.61)
Storm	All	197	29	549	392 (0.49; −0.09)	376 (0.53; 0.00)	422 (0.41; −0.26)	356 (0.58; 0.11)	345 (0.61; 0.16)	253 (0.79; 0.55)
Storm	Day	64	27	501	399 (0.36; 0.12)	425 (0.28; 0.00)	455 (0.17; −0.15)	386 (0.41; 0.18)	359 (0.49; 0.29)	249 (0.75; 0.66)
Storm	Night	27	11	548	459 (0.30; −0.68)	354 (0.58; 0.00)	458 (0.30; −0.67)	259 (0.78; 0.47)	263 (0.77; 0.45)	263 (0.77; 0.45)
Other	All	393	57	490	229 (0.78; −0.14)	214 (0.81; 0.00)	206 (0.82; 0.07)	240 (0.76; −0.26)	213 (0.81; 0.01)	205 (0.82; 0.08)
Other	Day	193	55	228	166 (0.47; −0.15)	155 (0.54; 0.00)	168 (0.46; −0.18)	199 (0.24; −0.65)	178 (0.39; −0.32)	154 (0.54; 0.01)
Other	Night	149	49	717	298 (0.83; −0.18)	275 (0.85; 0.00)	245 (0.88; 0.21)	296 (0.83; −0.16)	255 (0.87; 0.14)	264 (0.86; 0.08)
All	All	845	115	559	300 (0.71; −0.22)	271 (0.76; 0.00)	276 (0.76; −0.04)	276 (0.76; −0.03)	252 (0.80; 0.14)	212 (0.86; 0.39)
All	Day	308	107	301	236 (0.39; 0.01)	237 (0.38; 0.00)	256 (0.28; −0.16)	244 (0.34; −0.06)	230 (0.42; 0.06)	181 (0.64; 0.42)
All	Night	217	81	756	337 (0.80; −0.18)	309 (0.83; 0.00)	282 (0.86; 0.17)	312 (0.83; −0.01)	255 (0.89; 0.32)	260 (0.88; 0.30)

https://doi.org/10.5194/egusphere-2026-6859
Preprint. Discussion started: 19 November 2026
© Author(s) 2026. CC BY 4.0 License.

620 **Table 11.** Mean-square-error skill relative to IRI-2020 with 95 % day-block bootstrap confidence
621 intervals, for the quiet window, the storm window and the full evaluation period.

Window	BIAS7 SS vs IRI [95 % CI]	CLIM60 SS vs IRI [95 % CI]	PER1 SS vs IRI [95 % CI]	TREE-D SS vs IRI [95 % CI]	TREE-H SS vs IRI [95 % CI]	TREE-HN SS vs IRI [95 % CI]
Quiet	0.77 [0.71, 0.83]	0.85 [0.80, 0.89]	0.89 [0.82, 0.93]	0.85 [0.80, 0.89]	0.89 [0.85, 0.92]	0.92 [0.89, 0.95]
Storm	0.49 [0.36, 0.65]	0.53 [0.42, 0.66]	0.41 [0.23, 0.62]	0.58 [0.49, 0.69]	0.61 [0.52, 0.71]	0.79 [0.75, 0.84]
All	0.71 [0.65, 0.77]	0.76 [0.71, 0.82]	0.76 [0.67, 0.83]	0.76 [0.70, 0.80]	0.80 [0.74, 0.84]	0.86 [0.83, 0.88]

622

623

https://doi.org/10.5194/egusphere-2026-6859
Preprint. Discussion started: 19 November 2026
© Author(s) 2026. CC BY 4.0 License.

Table 12. Wilcoxon signed-rank tests of the daily mean-square errors of each model against IRI-2020 and against CLIM60, with Benjamini–Hochberg false-discovery-rate control over all tests.

Window	LT class	Model	Reference	Days	Median daily ΔMSE (K^2)	p	p (FDR)
Quiet	All	BIAS7	IRI	29	−328187	4.8×10^{-6}	1.8×10^{-5}
Quiet	All	BIAS7	CLIM60	29	34222	0.001	0.002
Quiet	All	CLIM60	IRI	29	−405338	5.3×10^{-6}	1.8×10^{-5}
Quiet	All	PER1	IRI	29	−431193	2.6×10^{-6}	1.1×10^{-5}
Quiet	All	PER1	CLIM60	29	−10271	0.053	0.080
Quiet	All	TREE-D	IRI	29	−394036	5.3×10^{-6}	1.8×10^{-5}
Quiet	All	TREE-D	CLIM60	29	5584	0.581	0.670
Quiet	All	TREE-H	IRI	29	−416175	3.2×10^{-6}	1.3×10^{-5}
Quiet	All	TREE-H	CLIM60	29	−9241	0.112	0.154
Quiet	All	TREE-HN	IRI	29	−436484	2.8×10^{-6}	1.2×10^{-5}
Quiet	All	TREE-HN	CLIM60	29	−20353	0.002	0.003
Storm	All	BIAS7	IRI	29	−120135	2.6×10^{-6}	1.1×10^{-5}
Storm	All	BIAS7	CLIM60	29	6977	0.469	0.555
Storm	All	CLIM60	IRI	29	−136059	1.5×10^{-5}	4.2×10^{-5}
Storm	All	PER1	IRI	29	−115683	6.0×10^{-5}	1.5×10^{-4}
Storm	All	PER1	CLIM60	29	−3052	0.804	0.846
Storm	All	TREE-D	IRI	29	−130006	2.6×10^{-6}	1.1×10^{-5}
Storm	All	TREE-D	CLIM60	29	−7615	0.127	0.172
Storm	All	TREE-H	IRI	29	−147697	2.8×10^{-6}	1.2×10^{-5}
Storm	All	TREE-H	CLIM60	29	−19420	0.064	0.095
Storm	All	TREE-HN	IRI	29	−186288	2.6×10^{-6}	1.1×10^{-5}
Storm	All	TREE-HN	CLIM60	29	−49612	5.0×10^{-5}	1.3×10^{-4}
All	All	BIAS7	IRI	115	−159012	2.1×10^{-20}	1.7×10^{-18}
All	All	BIAS7	CLIM60	115	4932	1.2×10^{-4}	2.8×10^{-4}
All	All	CLIM60	IRI	115	−160424	1.5×10^{-19}	6.6×10^{-18}
All	All	PER1	IRI	115	−157361	2.8×10^{-19}	1.0×10^{-17}
All	All	PER1	CLIM60	115	−2203	0.185	0.247
All	All	TREE-D	IRI	115	−161969	2.3×10^{-20}	1.7×10^{-18}
All	All	TREE-D	CLIM60	115	2861	0.249	0.322
All	All	TREE-H	IRI	115	−168137	3.7×10^{-20}	2.1×10^{-18}
All	All	TREE-H	CLIM60	115	−784	0.109	0.152
All	All	TREE-HN	IRI	115	−186288	1.8×10^{-20}	1.7×10^{-18}
All	All	TREE-HN	CLIM60	115	−7661	1.3×10^{-5}	3.9×10^{-5}

Table 13. Prediction RMSE and skill relative to IRI-2020 by Kp class of the verifying pass.

Kp class	LT class	Passes	Days	IRI RMSE (K)	BIAS7 RMSE (SS vs IRI)	CLIM60 RMSE (SS vs IRI)	PER1 RMSE (SS vs IRI)	TREE-D RMSE (SS vs IRI)	TREE-H RMSE (SS vs IRI)	TREE-HN RMSE (SS vs IRI)
Kp<3	All	711	108	543	268 (0.76)	235 (0.81)	237 (0.81)	243 (0.80)	226 (0.83)	198 (0.87)
Kp<3	Day	259	95	266	199 (0.44)	190 (0.49)	201 (0.43)	213 (0.36)	195 (0.46)	164 (0.62)
Kp<3	Night	170	69	757	305 (0.84)	281 (0.86)	242 (0.90)	294 (0.85)	244 (0.90)	249 (0.89)
3<=Kp<5	All	102	34	578	328 (0.68)	335 (0.66)	317 (0.70)	332 (0.67)	268 (0.78)	227 (0.85)
3<=Kp<5	Day	43	20	340	281 (0.32)	305 (0.19)	372 (−0.20)	296 (0.24)	282 (0.31)	210 (0.62)
3<=Kp<5	Night	38	20	740	344 (0.78)	356 (0.77)	286 (0.85)	333 (0.80)	252 (0.88)	254 (0.88)
Kp>=5	All	32	11	795	660 (0.31)	607 (0.42)	673 (0.28)	586 (0.46)	562 (0.50)	397 (0.75)
Kp>=5	Day	6	4	890	778 (0.24)	824 (0.14)	794 (0.20)	694 (0.39)	721 (0.34)	450 (0.75)
Kp>=5	Night	9	5	794	691 (0.24)	539 (0.54)	688 (0.25)	489 (0.62)	425 (0.71)	430 (0.71)

https://doi.org/10.5194/egusphere-2026-6859
Preprint. Discussion started: 19 November 2026
© Author(s) 2026. CC BY 4.0 License.

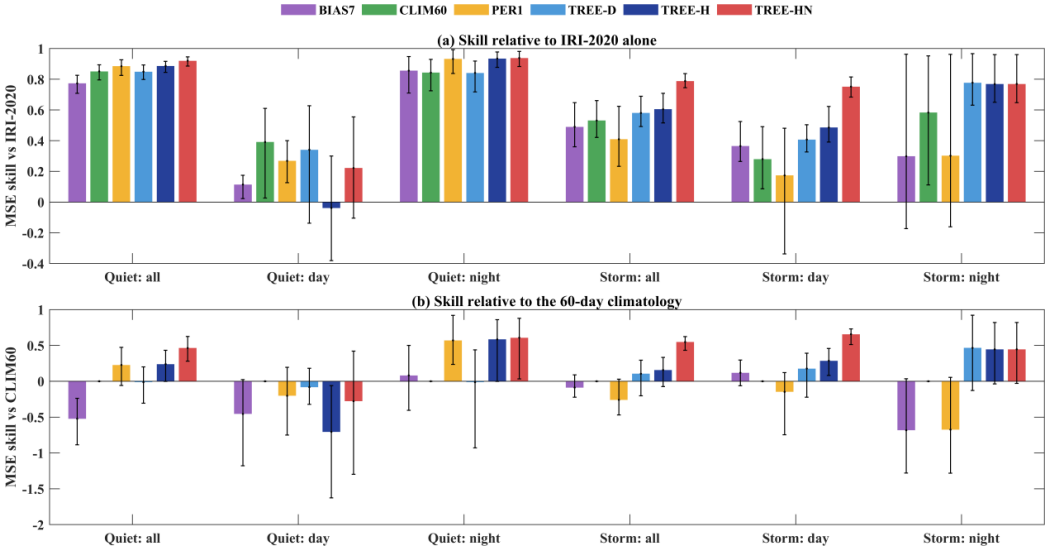


Figure 12. Out-of-sample skill of the residual models: mean-square-error skill (a) relative to IRI-2020 alone and (b) relative to the 60-day climatology CLIM60, for all, day and night passes of the quiet and storm windows. Error bars are 95 % day-block bootstrap intervals.

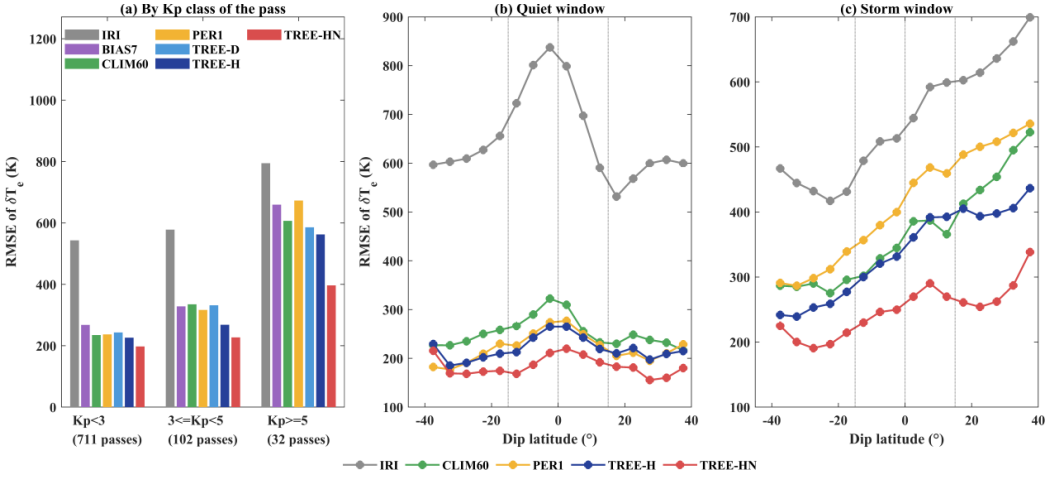


Figure 13. Prediction error by geomagnetic activity and dip latitude. (a) RMSE of each model by Kp class of the verifying pass over the full evaluation period. (b, c) RMSE in 5° dip-latitude bins in the (b) quiet and (c) storm windows; dotted lines mark the dip equator and ±15° dip latitude.

https://doi.org/10.5194/egusphere-2026-6859
Preprint. Discussion started: 19 November 2026
© Author(s) 2026. CC BY 4.0 License.

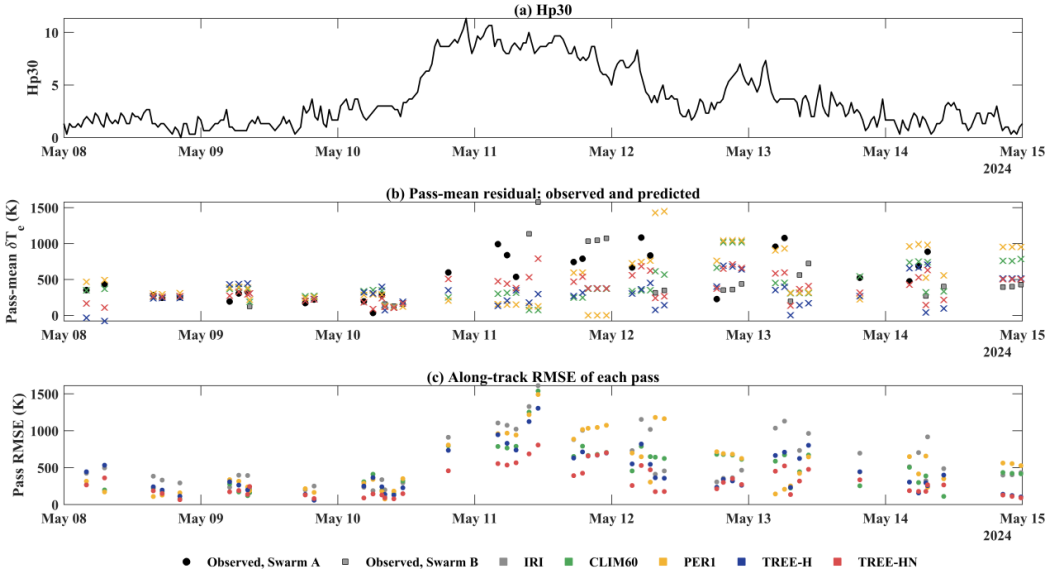


Figure 14. The superstorm of May 2024. (a) The Hp30 index, 8–15 May 2024. (b) Pass-mean residual observed by Swarm A (black circles) and Swarm B (grey squares) and predicted by CLIM60, PER1, TREE-H and TREE-HN (crosses, colours as in the legend). (c) Along-track RMSE of each pass for IRI-2020 and the four models (dots).

4.9 Relation between predictive skill and the dynamical invariants

The predictor importance of TREE-H is shown in Fig. 15a, and the pass-level skill of TREE-H relative to IRI is compared with the invariants of the same pass in Fig. 15b and Table 14. Over all 452 diagnosed passes the correlation with z_{DET} is weak and negative ($r_s = -0.14$, $p = 0.003$) and that with the skill relative to CLIM60 is absent ($r_s = 0.00$, $p = 0.95$). On the day side the relation is clearer: passes with a larger surrogate excess are predicted less well, both relative to IRI ($r_s = -0.35$, $p < 0.001$) and relative to CLIM60 ($r_s = -0.25$, $p = 0.008$), and a larger Lyapunov exponent is also associated with a lower skill relative to CLIM60 ($r_s = -0.22$, $p = 0.020$). On the night side no correlation is significant, apart from a positive association between z_{DET} and skill relative to IRI on the 27 storm-window night passes ($r_s = 0.47$, $p = 0.015$).

The day-side result has a direct interpretation. The tree ensembles predict each sample from drivers and from the residual of previous days at the same local time and dip latitude; they do not use the along-track trajectory of the current pass. Structure that is deterministic along the track but changes from day to day, which is what a large surrogate excess with a short along-track

https://doi.org/10.5194/egusphere-2026-6859
Preprint. Discussion started: 19 November 2026
© Author(s) 2026. CC BY 4.0 License.

horizon describes, is therefore not captured by such a model and appears as error. Exploiting it would require a model that conditions on the earlier part of the same pass, which is possible operationally only in a nowcast setting.

Table 14. Spearman correlation across passes between the dynamical invariants (z_{DET}, DET, λ_1) and the mean-square-error skill of TREE-H relative to IRI-2020 and relative to CLIM60.

Window	LT class	Invariant	Passes	r_s with skill vs IRI	p	r_s with skill vs CLIM60	p
Quiet	All	z_{DET}	255	−0.23	1.9×10^{-4}	−0.06	0.351
Quiet	All	DET	255	0.24	1.4×10^{-4}	−0.13	0.045
Quiet	All	λ_1	255	−0.26	2.0×10^{-5}	−0.03	0.591
Quiet	Day	z_{DET}	51	−0.13	0.362	0.21	0.138
Quiet	Day	DET	51	0.27	0.054	−0.15	0.308
Quiet	Day	λ_1	51	−0.18	0.204	−0.06	0.694
Quiet	Night	z_{DET}	41	−0.01	0.969	0.14	0.386
Quiet	Night	DET	41	−0.09	0.583	−0.11	0.505
Quiet	Night	λ_1	41	−0.28	0.076	0.12	0.471
Storm	All	z_{DET}	197	0.11	0.121	0.11	0.125
Storm	All	DET	197	−0.45	4.7×10^{-11}	0.05	0.506
Storm	All	λ_1	197	0.30	1.6×10^{-5}	−0.06	0.387
Storm	Day	z_{DET}	64	−0.13	0.311	−0.17	0.167
Storm	Day	DET	64	−0.23	0.072	−0.21	0.090
Storm	Day	λ_1	64	0.21	0.096	−0.01	0.962
Storm	Night	z_{DET}	27	0.47	0.015	−0.05	0.818
Storm	Night	DET	27	−0.43	0.026	0.01	0.945
Storm	Night	λ_1	27	0.28	0.156	0.05	0.818
All	All	z_{DET}	452	−0.14	0.003	0.00	0.950
All	All	DET	452	0.05	0.247	−0.07	0.153
All	All	λ_1	452	−0.02	0.650	−0.07	0.128
All	Day	z_{DET}	115	−0.35	1.2×10^{-4}	−0.25	0.008
All	Day	DET	115	0.26	0.005	0.13	0.167
All	Day	λ_1	115	−0.17	0.064	−0.22	0.020
All	Night	z_{DET}	68	0.07	0.568	0.10	0.396
All	Night	DET	68	−0.07	0.592	−0.08	0.522
All	Night	λ_1	68	−0.18	0.145	0.15	0.229

https://doi.org/10.5194/egusphere-2026-6859
Preprint. Discussion started: 19 November 2026
© Author(s) 2026. CC BY 4.0 License.

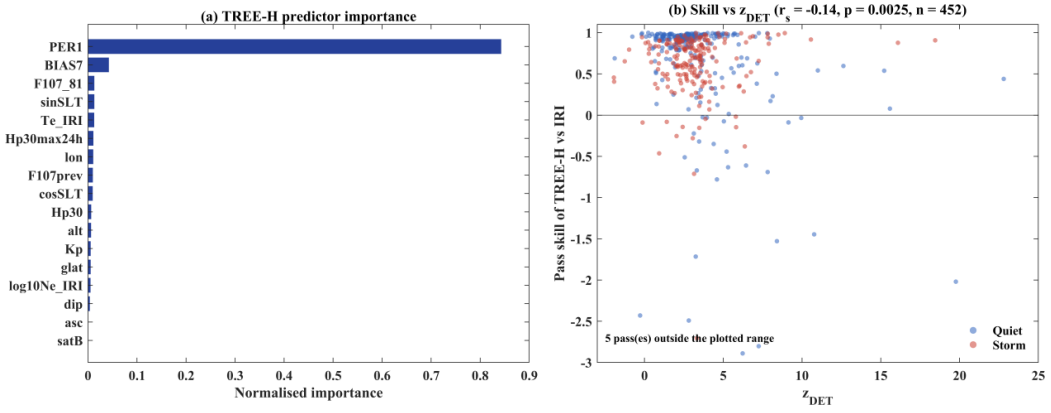


Figure 15. (a) Normalised predictor importance of TREE-H averaged over the weekly retraining dates. (b) Pass-level mean-square-error skill of TREE-H relative to IRI-2020 against the determinism excess z_DET of the same pass, for the quiet (blue) and storm (red) windows; the Spearman correlation over all diagnosed passes is given in the title, and passes outside the plotted range are counted in the panel.

4.10 Limitations

The conclusions are conditioned by the observing system and by the design. (i) The reference is an in situ measurement whose instrumental behaviour is only partly known. The radar-based calibration is undefined over much of the daytime anomaly, so the uncalibrated temperature had to be used throughout; the night-side temperature shows a two-state behaviour not described in the product documentation, which was handled by excluding passes with switches and which leaves the absolute night-side residual uncertain by several hundred kelvin; and a residual step of order 10–30 K remains at the 128-s probe sweeps after correction. (ii) The analysis is along track: the invariants and horizons describe the spatial structure of the residual over a pass of about 1000 s, not its evolution in time at a fixed location, and the IAAFT null hypothesis assumes a stationarity that a pass crossing the anomaly does not strictly satisfy. (iii) The local-time sampling is set by the slow precession of the two orbital planes, so local time, satellite altitude and window are partly confounded, and twilight passes come mostly from one satellite in each window. (iv) The two windows lie in one year near solar maximum, and the storm window is dominated by a single exceptional event. (v) The surrogate ensembles have 100 members, and the Lyapunov exponent is not robust to pre-filtering or to the embedding convention. (vi) The prediction experiment covers one season of one year, and its targets are on the measurement scale of the uncalibrated Langmuir probe, so the corrected temperature inherits the probe's calibration offset

https://doi.org/10.5194/egusphere-2026-6859
Preprint. Discussion started: 19 November 2026
© Author(s) 2026. CC BY 4.0 License.

relative to incoherent-scatter radars. Natural extensions are a recalibrated Swarm temperature product with the state behaviour resolved, longer records covering the declining phase of the solar cycle, comparison with other longitude sectors, and independent validation against ground-based radar temperatures.

https://doi.org/10.5194/egusphere-2026-6859
Preprint. Discussion started: 19 November 2026
© Author(s) 2026. CC BY 4.0 License.

5 Conclusions

The residual between the Swarm Langmuir-probe electron temperature and the IRI-2020 topside electron temperature was analysed along 452 quality-controlled passes over the African low-latitude sector in a quiet (February 2024) and a disturbed (late April to May 2024) window, and predicted out of sample for 845 passes between February and May 2024. The main findings are as follows.

1. Data quality. Before any dynamical interpretation, the in situ reference required three corrections: removal of a relaxation of several tens of kelvin after each 128-s probe sweep, use of the uncalibrated temperature because the radar-based calibration is undefined over much of the daytime anomaly, and exclusion of passes containing switches between two night-side temperature states separated by 600–900 K. The two-state behaviour is not reflected in the quality flags; without its exclusion it would have altered the night-side invariants.

2. Deterministic structure. The determinism of the residual exceeded that of 100 IAAFT surrogates for 396 of the 452 passes. The excess is largest on the quiet day side (median z_{DET} 5.3) and is reduced on the storm-window day side (3.3), while raw determinism moves in the opposite direction; as for TEC, the surrogate excess and not raw determinism measures the nonlinear content of the IRI error.

3. Attribution. The radar-based calibration of the probe, the analogue of the plasmaspheric term in the TEC case, changes the determinism by a median of 0.14–0.32 of the between-pass interquartile range, and the choice of IRI topside temperature option (TBT-2012 or Bil-1985) by 0.15–0.63, with individual passes affected more strongly. The invariants are therefore attributable mainly to the IRI error, but they are specific to the default TBT-2012 configuration.

4. Organisation. The deterministic excess is distributed differently along the pass in different local-time and storm conditions, with the lowest values near the dip equator on quiet dawn passes and on storm-window night passes and the opposite gradient on quiet dusk

https://doi.org/10.5194/egusphere-2026-6859
Preprint. Discussion started: 19 November 2026
© Author(s) 2026. CC BY 4.0 License.

and storm-window day passes; the northern and southern halves of the passes give similar values.

5. Predictability. Surrogate-supported Lyapunov exponents were found for 51 passes, mostly on the quiet day side and at night, with along-track horizons of about 330–740 km (3.7–8.4° of dip latitude); the exponent is not robust to pre-filtering or to the embedding convention and is reported as indicative. In time, the residual is predictable mainly through persistence: relative to a 60-day climatology, a tree ensemble using the residual of previous days improves the mean-square error by 0.14 (95 % interval 0.02–0.23), and a nowcast that also uses the measured density by 0.39 (0.27–0.48). Day-side passes with a larger nonlinear excess are predicted less well.

6. Operational implication. Most of the IRI-2020 error relative to the Swarm probe is a slowly varying offset that a running climatology of recent passes removes; additional skill comes from day-to-day persistence and, much more, from concurrent density information, while the onset of a severe storm is not predicted by models that use only past residuals and the available geomagnetic indices. Any correction derived in this way inherits the calibration offset of the probe relative to incoherent-scatter radars.

Code and data availability

Swarm Level-1b Langmuir-probe data (EFIx_LP_1B, baseline 0701) are distributed by the European Space Agency at https://swarm-diss.eo.esa.int. The Kp, Hp30 and F10.7 indices are provided by the GFZ German Research Centre for Geosciences (Matzka et al., 2021; Yamazaki et al., 2022). The IRI-2020 model is available from https://irimodel.org. The derived residual series and the diagnostic and prediction results are available from the corresponding author on reasonable request. The processing, analysis and plotting code is available from the corresponding author on request.

https://doi.org/10.5194/egusphere-2026-6859
Preprint. Discussion started: 19 November 2026
© Author(s) 2026. CC BY 4.0 License.

Author contributions

OVE conceived the study, carried out the data processing, analysis and figure preparation, and wrote the manuscript. OEA contributed to the methodology and data validation and supervised the research. TDA assisted with formal analysis and model validation. BNOE assisted with study conceptualisation and provided critical reviews of the manuscript. All authors contributed to writing – review and editing

Competing interests

The authors declare that they have no conflict of interest.

Acknowledgements

The authors thank the European Space Agency for the Swarm Langmuir-probe data, the GFZ German Research Centre for Geosciences for the Kp, Hp30 and F10.7 indices, and the International Reference Ionosphere working group for the IRI-2020 model. The authors used Claude (Anthropic), a large language model, to assist in writing and testing the data-processing, analysis and plotting code and in drafting and editing the manuscript text. The authors reviewed and edited the code, results and text and take full responsibility for the content of this publication.

https://doi.org/10.5194/egusphere-2026-6859
Preprint. Discussion started: 19 November 2026
© Author(s) 2026. CC BY 4.0 License.

References

Balan, N., Liu, L., and Le, H.: A brief review of equatorial ionization anomaly and ionospheric irregularities, Earth Planet. Phys., 2, 257–275, https://doi.org/10.26464/epp2018025, 2018.

Benjamini, Y. and Hochberg, Y.: Controlling the false discovery rate: a practical and powerful approach to multiple testing, J. R. Stat. Soc. Ser. B, 57, 289–300, https://doi.org/10.1111/j.2517-6161.1995.tb02031.x, 1995.

Bilitza, D., Pezzopane, M., Truhlik, V., Altadill, D., Reinisch, B. W., and Pignalberi, A.: The International Reference Ionosphere model: a review and description of an ionospheric benchmark, Rev. Geophys., 60, e2022RG000792, https://doi.org/10.1029/2022RG000792, 2022.

Catapano, F., Buchert, S., Qamili, E., Nilsson, T., Bouffard, J., Siemes, C., Coco, I., D'Amicis, R., Tøffner-Clausen, L., Trenchi, L., Holmdahl Olsen, P. E., and Stromme, A.: Swarm Langmuir probes' data quality validation and future improvements, Geosci. Instrum. Method. Data Syst., 11, 149–162, https://doi.org/10.5194/gi-11-149-2022, 2022.

Consolini, G., De Michelis, P., Alberti, T., Giannattasio, F., Coco, I., Tozzi, R., and Chang, T. T. S.: On the multifractal features of low-frequency magnetic field fluctuations in the field-aligned current ionospheric polar regions: Swarm observations, J. Geophys. Res. Space Phys., 125, e2019JA027429, https://doi.org/10.1029/2019JA027429, 2020.

Eckmann, J.-P., Kamphorst, S. O., and Ruelle, D.: Recurrence plots of dynamical systems, Europhys. Lett., 4, 973–977, https://doi.org/10.1209/0295-5075/4/9/004, 1987.

Edward, O. V.: Nonlinear dynamical structure of IRI-2020 total electron content prediction errors over the African equatorial ionosphere: recurrence quantification and Lyapunov analysis of GNSS-derived residuals during quiet and storm conditions, Adv. Space Res., https://doi.org/10.1016/j.asr.2026.09.092, 2026.

Fraser, A. M. and Swinney, H. L.: Independent coordinates for strange attractors from mutual information, Phys. Rev. A, 33, 1134–1140, https://doi.org/10.1103/PhysRevA.33.1134, 1986.

Friedman, J. H.: Greedy function approximation: a gradient boosting machine, Ann. Stat., 29, 1189–1232, https://doi.org/10.1214/aos/1013203451, 2001.

https://doi.org/10.5194/egusphere-2026-6859
Preprint. Discussion started: 19 November 2026
© Author(s) 2026. CC BY 4.0 License.



784 Friis-Christensen, E., Lühr, H., and Hulot, G.: Swarm: a constellation to study the Earth's
785 magnetic field, Earth Planets Space, 58, 351–358, https://doi.org/10.1186/BF03351933, 2006.

786 Kennel, M. B., Brown, R., and Abarbanel, H. D. I.: Determining embedding dimension for phase-
787 space reconstruction using a geometrical construction, Phys. Rev. A, 45, 3403–3411,
788 https://doi.org/10.1103/PhysRevA.45.3403, 1992.

789 Knudsen, D. J., Burchill, J. K., Buchert, S. C., Eriksson, A. I., Gill, R., Wahlund, J.-E., Åhlen, L.,
790 Smith, M., and Moffat, B.: Thermal ion imagers and Langmuir probes in the Swarm electric field
791 instruments, J. Geophys. Res. Space Phys., 122, 2655–2673,
792 https://doi.org/10.1002/2016JA022571, 2017.

793 Künsch, H. R.: The jackknife and the bootstrap for general stationary observations, Ann. Stat.,
794 17, 1217–1241, https://doi.org/10.1214/aos/1176347265, 1989.

795 Lomidze, L., Knudsen, D. J., Burchill, J., Kouznetsov, A., and Buchert, S. C.: Calibration and
796 validation of Swarm plasma densities and electron temperatures using ground-based radars and
797 satellite radio occultation measurements, Radio Sci., 53, 15–36,
798 https://doi.org/10.1002/2017RS006415, 2018.

799 Marwan, N., Romano, M. C., Thiel, M., and Kurths, J.: Recurrence plots for the analysis of
800 complex systems, Phys. Rep., 438, 237–329, https://doi.org/10.1016/j.physrep.2006.11.001,
801 2007.

802 Materassi, M., Alberti, T., Migoya-Orué, Y., Radicella, S. M., and Consolini, G.: Chaos and
803 predictability in ionospheric time series, Entropy, 25, 368, https://doi.org/10.3390/e25020368,
804 2023.

805 Matzka, J., Stolle, C., Yamazaki, Y., Bronkalla, O., and Morschhauser, A.: The geomagnetic Kp
806 index and derived indices of geomagnetic activity, Space Weather, 19, e2020SW002641,
807 https://doi.org/10.1029/2020SW002641, 2021.

808 Oyama, K.-I., Balan, N., Watanabe, S., Takahashi, T., Isoda, F., Bailey, G. J., and Oya, H.:
809 Morning overshoot of Te enhanced by downward plasma drift in the equatorial topside
810 ionosphere, J. Geomagn. Geoelectr., 48, 959–966, https://doi.org/10.5636/jgg.48.959, 1996.

https://doi.org/10.5194/egusphere-2026-6859
Preprint. Discussion started: 19 November 2026
© Author(s) 2026. CC BY 4.0 License.

Pignalberi, A., Pezzopane, M., Tozzi, R., De Michelis, P., and Coco, I.: Comparison between IRI and preliminary Swarm Langmuir probe measurements during the St. Patrick storm period, Earth Planets Space, 68, 93, https://doi.org/10.1186/s40623-016-0466-5, 2016.

Pignalberi, A., Giannattasio, F., Truhlik, V., Coco, I., Pezzopane, M., Consolini, G., De Michelis, P., and Tozzi, R.: On the electron temperature in the topside ionosphere as seen by Swarm satellites, incoherent scatter radars, and the International Reference Ionosphere model, Remote Sens., 13, 4077, https://doi.org/10.3390/rs13204077, 2021.

Pignalberi, A., Giannattasio, F., Truhlik, V., Coco, I., Pezzopane, M., and Alberti, T.: Investigating the main features of the correlation between electron density and temperature in the topside ionosphere through Swarm satellites data, J. Geophys. Res. Space Phys., 129, e2023JA032201, https://doi.org/10.1029/2023JA032201, 2024a.

Pignalberi, A., Truhlik, V., Giannattasio, F., Coco, I., and Pezzopane, M.: Mid- and high-latitude electron temperature dependence on solar activity in the topside ionosphere through the Swarm B satellite observations and the International Reference Ionosphere model, Atmosphere, 15, 490, https://doi.org/10.3390/atmos15040490, 2024b.

Rosenstein, M. T., Collins, J. J., and De Luca, C. J.: A practical method for calculating largest Lyapunov exponents from small data sets, Physica D, 65, 117–134, https://doi.org/10.1016/0167-2789(93)90009-P, 1993.

Schreiber, T. and Schmitz, A.: Improved surrogate data for nonlinearity tests, Phys. Rev. Lett., 77, 635–638, https://doi.org/10.1103/PhysRevLett.77.635, 1996.

Schunk, R. and Nagy, A.: Ionospheres: Physics, Plasma Physics, and Chemistry, 2nd edn., Cambridge University Press, Cambridge, https://doi.org/10.1017/CBO9780511635342, 2009.

Shreedevi, P. R., Nair, A. S., Miyoshi, Y., Buchert, S. C., Otsuka, Y., Shinbori, A., Krishnan, L. G., Perwitasari, S., and Nishioka, M.: Superstorm-driven electron temperature (Te) anomalies in the topside low-latitude ionosphere: role of the equatorial ionization anomaly (EIA), Earth Planets Space, 77, 163, https://doi.org/10.1186/s40623-025-02282-x, 2025.

https://doi.org/10.5194/egusphere-2026-6859
Preprint. Discussion started: 19 November 2026
© Author(s) 2026. CC BY 4.0 License.



837 Stolle, C., Liu, H., Truhlík, V., Lühr, H., and Richards, P. G.: Solar flux variation of the electron
838 temperature morning overshoot in the equatorial F region, J. Geophys. Res., 116, A04308,
839 https://doi.org/10.1029/2010JA016235, 2011.

840 Takens, F.: Detecting strange attractors in turbulence, in: Dynamical Systems and Turbulence,
841 Warwick 1980, edited by: Rand, D. A. and Young, L.-S., Lecture Notes in Mathematics, 898,
842 Springer, Berlin, 366–381, https://doi.org/10.1007/BFb0091924, 1981.

843 Truhlik, V., Bilitza, D., and Triskova, L.: A new global empirical model of the electron
844 temperature with the inclusion of the solar activity variations for IRI, Earth Planets Space, 64,
845 531–543, https://doi.org/10.5047/eps.2011.10.016, 2012.

846 Wilcoxon, F.: Individual comparisons by ranking methods, Biom. Bull., 1, 80–83,
847 https://doi.org/10.2307/3001968, 1945.

848 Yamazaki, Y., Matzka, J., Stolle, C., Kervalishvili, G., Rauberg, J., Bronkalla, O., Morschhauser,
849 A., Bruinsma, S., Shprits, Y. Y., and Jackson, D. R.: Geomagnetic activity index Hpo, Geophys.
850 Res. Lett., 49, e2022GL098860, https://doi.org/10.1029/2022GL098860, 2022.

851 Yan, R., Xiong, C., Zhima, Z., Shen, X., Liu, D., Liu, C., Guan, Y., Zhu, K., Zheng, L., and Lv,
852 F.: Correlation between Ne and Te around 14:00 LT in the topside ionosphere observed by CSES,
853 Swarm and CHAMP satellites, Front. Earth Sci., 10, 860234,
854 https://doi.org/10.3389/feart.2022.860234, 2022.