Three-Dimensional Structure and Spatiotemporal Evolution of Ionospheric Disturbances During the 19–22 January 2026 Geomagnetic Storm

Supplementary Information

This Supplementary Information presents ten additional diagnostic figures, seven supplementary animations, and the mathematical specification and evaluation protocol for the proposed framework. Supplementary Figures S1–S10 retain the supplied image data, with captions reconciled against their actual panels. Movies S1–S7 are ordered by their printed epochs; the PDF contains representative frames, and the complete animations accompany the manuscript as GIF files.

The supplied diagnostics include retrospective event comparisons and controlled prediction experiments. Their reported statistics are distinguished from the prospective MP-PET protocol in Notes S7–S18. Raw scoring arrays, complete execution records, and independent electron-density profiles were not supplied with these graphics. The additional prediction diagnostics therefore do not establish the absolute accuracy or unique altitude allocation of the reconstructed density field. Product-specific time bases, amplitude scales, support masks, and preprocessing settings are retained throughout.

Contents

S1 Observational coverage and geometric support . . . 3
S2 Reference information and held-out prediction diagnostics . . . 5
S3 Altitude structure and sampling geometry . . . 8
S4 Spatial, temporal, and cross-altitude diagnostics . . . 12
S5 Column cancellation and processing sensitivity . . . 17
S6 Supplementary animations . . . 20
Scope of the technical supplement . . . 28
S7 Slant observations and the disturbance state . . . 28
S8 Fixed spatial representation . . . 31
S9 Multiple propagation fields and structured evolution . . . 32
S10 Duplication-controlled joint objective . . . 34
S11 Profiled propagation estimation . . . 36
S12 Identifiability after accounting for competing explanations . . . 39
S13 Amplitude, phase, and interference response . . . 41
S14 Model selection, validation, and reported quantities . . . 42
S15 Numerical specification and reconstruction procedure . . . 47
S16 Characteristic integration and operator bounds . . . 49
S17 Profile sensitivity and inner-accuracy control . . . 50
S18 Reproducibility, coverage, and interpretation . . . 52

Supplementary Note S1. Observational coverage and geometric support

S1.1. Observing geometry, temporal availability, and spatial support

Figure S1. The supplied network inventory contains 2,198 receiver identifiers across the four-day interval, with 1,597 and 1,590 receivers in W1 and W2, respectively. The seven-receiver difference is approximately 0.44% of the W1 count, but similar counts do not imply identical spatial or angular sampling. The two 30-min windows cover one hour in total, and their support maps use the same 350-km level and percentage scale. The ray-coverage and local-geometry diagnostics differ visibly across altitude, demonstrating that the presence of intersecting rays alone does not establish adequate local geometry. The index panel annotates SYM-H at −257 nT at 11:52 and AE at 4,366 nT at 11:33 on 20 January. These annotations are retained as properties of the supplied series; they do not replace the separately cited event chronology in the main manuscript. The precise support thresholds and index-product processing remain matters of source provenance, rather than quantities recoverable from this image.

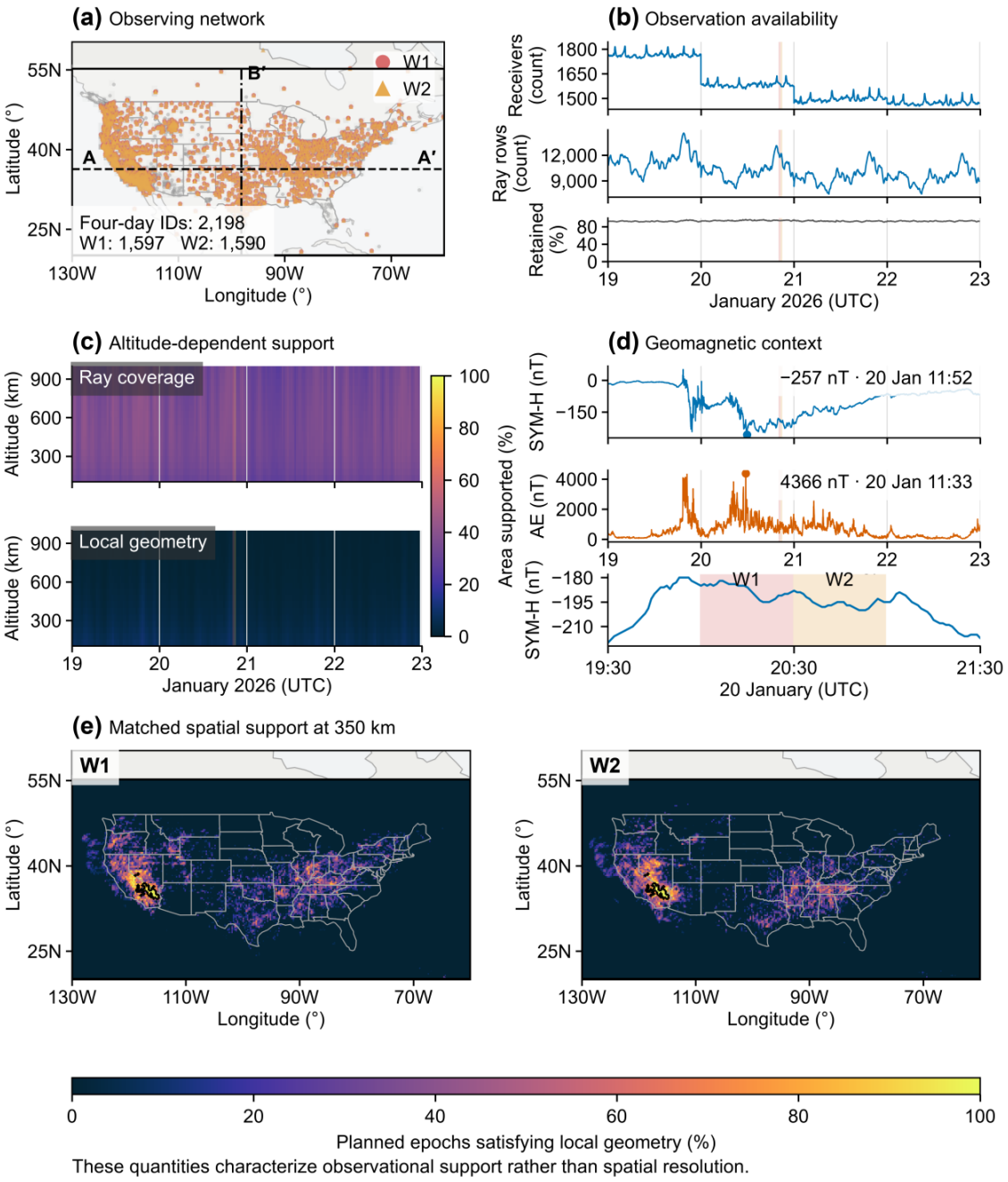


Figure S1: Geomagnetic context and observational support for the 19–22 January 2026 North American case study. (a) Receiver distribution, the indicated reconstruction domain, station subsets for windows W1 and W2, and the A–A′ and B–B′ transects. (b) Active-receiver counts, usable ray rows, and retained-observation fractions. (c) Altitude-dependent fractions of the domain satisfying the source ray-coverage and local-geometric-support criteria. (d) The supplied SYM-H and AE records and an expanded SYM-H view of two adjacent 30-min windows on 20 January: W1, 19:59:42–20:29:42 UTC, and W2, 20:29:42–20:59:42 UTC. (e) Spatial support at 350 km, expressed as the percentage of planned epochs satisfying the local-geometry criterion; the common contour identifies cells meeting the source criterion in both windows. The annotated index extrema are those of this supplied series. All displayed times in this figure are UTC. Missing observations are distinguished from zero support. These diagnostics characterize observational support and do not measure spatial resolving power.

Supplementary Note S2. Reference information and held-out prediction diagnostics

S2.1. Reference information in ionospheric-increment prediction

Figure S2. The controlled benchmark reports reference skills of 4.31%, 0.84%, and 2.68% at 30-, 120-, and 300-s leads, with respective 95% intervals of 3.78–4.84%, 0.70–0.99%, and 2.21–3.18%. The corresponding scoring populations contain 8,064, 7,794, and 7,253 intervals per realization and therefore differ across leads. Mismatched references yield negative skill. The source protocol specifies 16 training, eight development, and 24 test realizations; 14 reference stations, four development targets, and six final held-out targets have mutually exclusive roles. Simulations preserve the real ray geometry, 30-s sampling, missing observations, and arc labels, with a 5-mm geometry-free phase endpoint-noise standard deviation and the induced differencing correlation. AR32 was selected from AR0/1/2/4/8/16/32 on the development set, providing 930 s of history. For the actual event, the overall MSE-reduction intervals and the hybrid baseline-comparison intervals cross zero. The reported positive simulation skills therefore should not be generalized into an established event-wide MSE advantage or an independently validated three-dimensional density reconstruction.

S2.2. Held-out prediction diagnostics and observation geometry

Figure S3. The six cluster counts are 5,170, 5,193, 5,070, 5,266, 5,204, and 5,219, summing to the 31,122 observations in the pooled comparison. Every cluster is annotated with 100% prediction coverage, which describes the availability of predictions within this selected sample rather than network-wide availability. Five cluster point estimates favor three-dimensional projection, whereas Cluster 4 has a small positive difference; the pooled interval includes zero. The elevation-bin counts, 2,396, 5,707, 5,977, 6,476, 6,224, and 4,342, independently sum to the same sample size. Displayed three-dimensional MAE decreases from approximately 1.1 TECU near 25° elevation to approximately 0.2 TECU near 78°. In the representative arc, both predictions substantially overestimate the first negative excursion, despite improved agreement later in the window. The near-zero pooled bias therefore does not imply uniformly small local errors, and the image alone does not establish a statistically resolved pooled advantage.

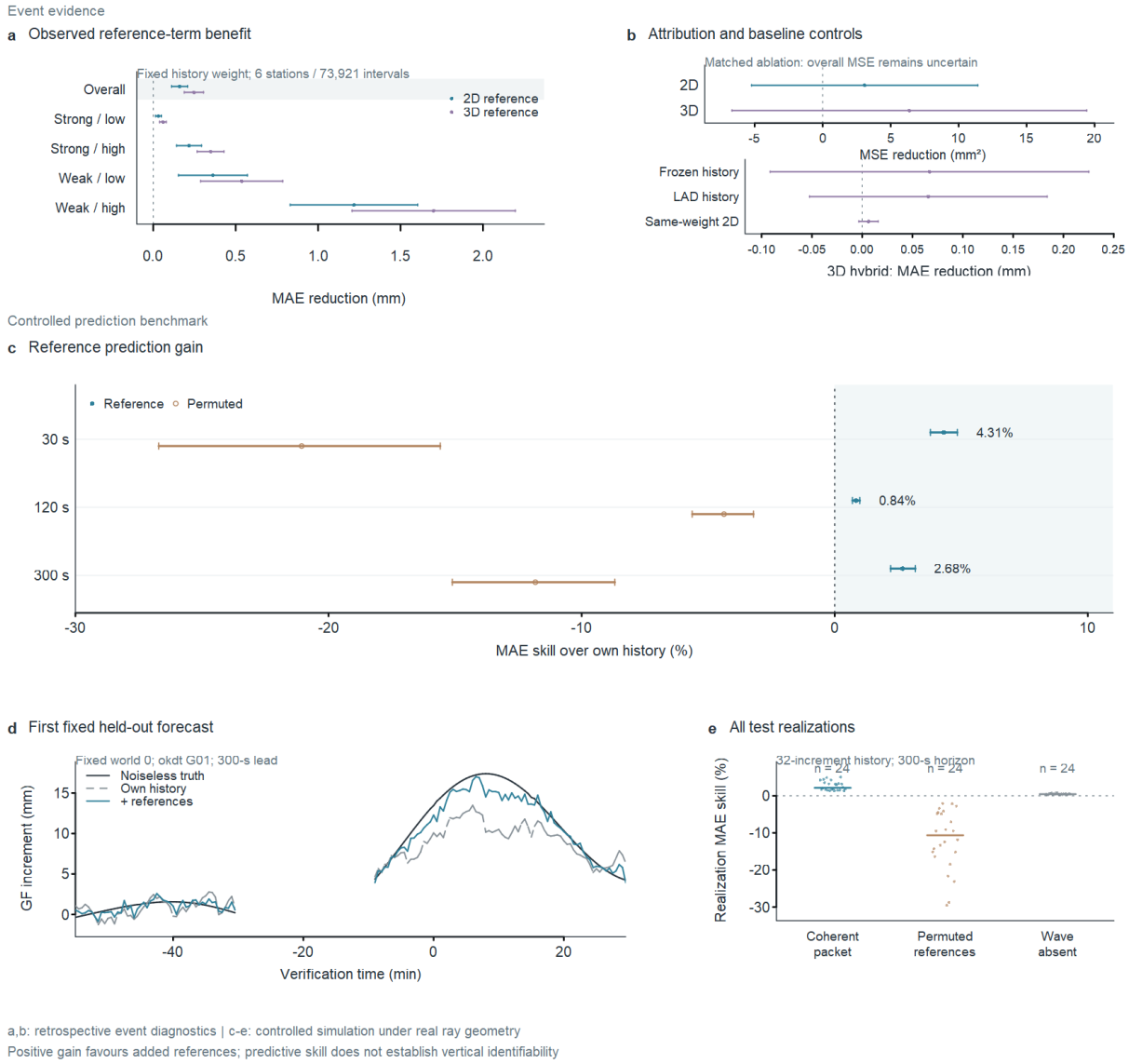


Figure S2: Contribution of independent reference observations to GNSS ionospheric-increment prediction. (a) Retrospective event MAE reductions for two-dimensional and three-dimensional reference terms relative to matched history-shrinkage baselines. (b) The corresponding MSE reductions (upper panel) and three-dimensional hybrid MAE reductions relative to frozen-history, development-fitted least-absolute-deviation (LAD) history, and same-weight two-dimensional baselines (lower panel). The event comparison uses 73,921 common scoring intervals from six held-out target stations with equal station weights. Its error bars are source-reported approximate simultaneous 95% intervals, calculated separately by comparison family using synchronized 10-min block resampling within nine windows. (c) Controlled-simulation MAE skill for matched and mismatched reference realizations at 30-, 120-, and 300-s leads; error bars are pointwise percentile 95% intervals from paired realization resampling. (d) The prespecified first test realization at station okdt, satellite G01, with a 300-s lead; verification time is relative to 20:30 GPST. (e) Skills for all 24 test realizations at 300 s; short bars mark medians, and the wave-absent control retains common background variability. The forecast target is the final 30-s increment ending at $t_0 + h$, rather than the cumulative change over the lead. Positive MAE skill, $100[1 - \overline{\mathrm{MAE}}_{\mathrm{reference}}/\overline{\mathrm{MAE}}_{\mathrm{history}}]$, favors the added references. Predictive skill does not by itself establish vertical identifiability.

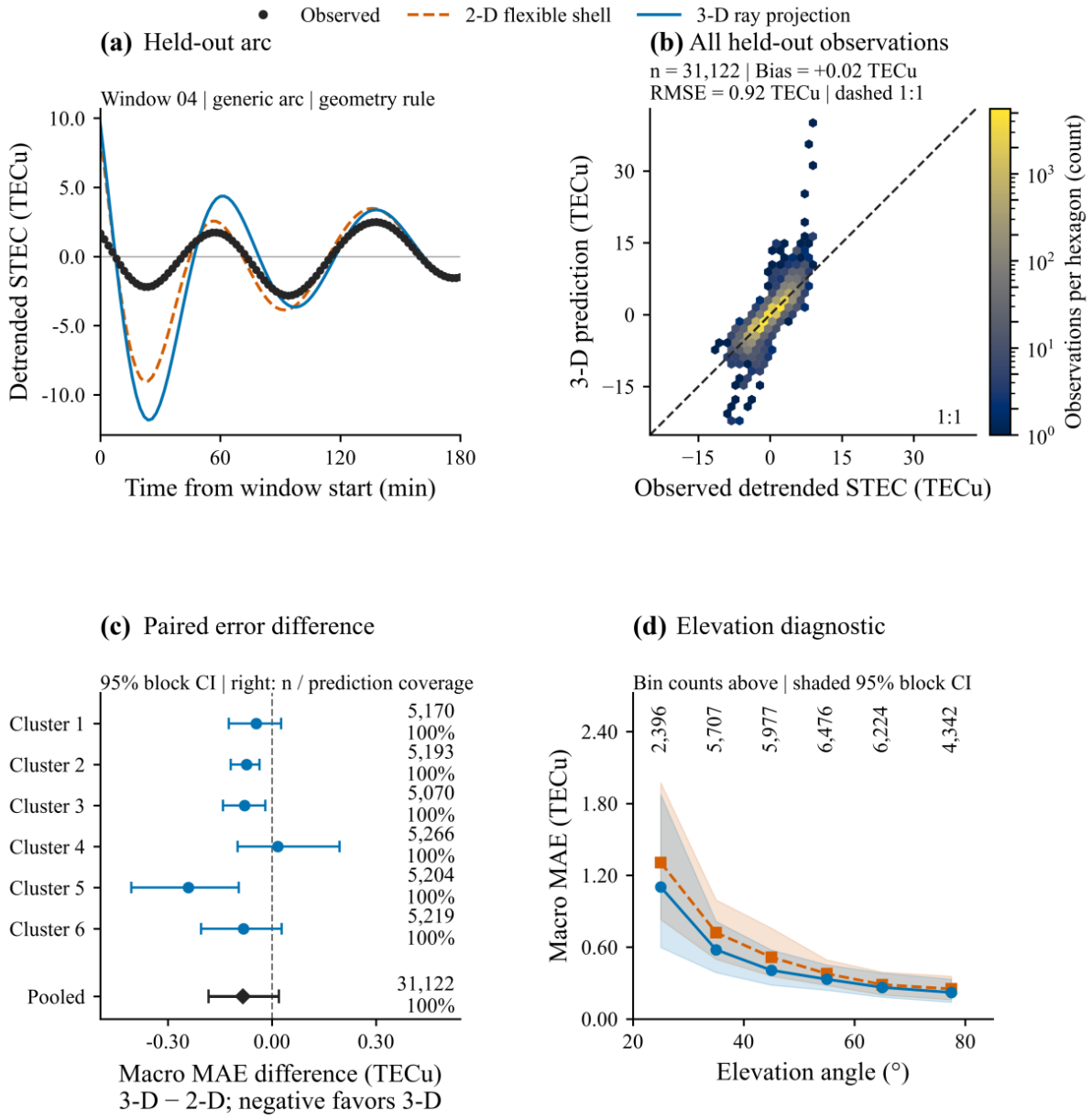


Figure S3: Prediction diagnostics for the observations labeled as held out in the supplied analysis. (a) Observed detrended STEC and predictions from a two-dimensional flexible-shell representation and three-dimensional ray projection for the displayed Window 04 arc. (b) Observed versus three-dimensional predicted detrended STEC for 31,122 samples; hexagon colors show sample counts on a logarithmic scale, and the dashed line denotes one-to-one agreement. The source annotates a +0.02-TECU bias and 0.92-TECU RMSE. (c) Paired macro-MAE differences, three-dimensional minus two-dimensional, for six clusters and the pooled sample; negative differences favor the three-dimensional representation. Error bars denote source-reported 95% block confidence intervals, and the annotations give sample counts and prediction coverage. (d) Macro MAE versus elevation angle, with bin counts and 95% block intervals. These are diagnostics of the supplied held-out prediction products; station identities, arc membership, selection independence, and the exact block definition require the corresponding execution records.

Supplementary Note S3. Altitude structure and sampling geometry

The epoch identifiers S1–S4 in Figs. S4–S8 designate selected source times and are independent of the supplementary object numbers. Figures S4 and S5 reproduce source-directory times. Figures S6 and S7 instead print the source's adopted GPS-time-minus-18-s conversion. These conventions are retained explicitly, rather than treating labels that differ by 18 s as identical UTC annotations.

S3.1. Selected altitude slices and two section geometries

Figure S4. The matrix contains 28 horizontal maps and eight sections, for 36 panels in total. Its horizontal maps cover seven successive 50-km bands from 150 to 500 km, while the section views extend down to 100 km. The source timestamps span 37 h 47 min, with successive separations of 11 h 51 min, 17 h 10 min, and 8 h 46 min. The S2 scale covers $\pm 5 \times 10^{10}$ electrons m^{-3}, whereas the S1, S3, and S4 limits are three times larger at $\pm 1.5 \times 10^{11}$ electrons m^{-3}. Similar colors therefore cannot be read as equal residual amplitudes across all four columns. Broad alternating-sign structures occur in several lower bands, with reduced displayed contrast toward 450–500 km in many locations. The sectional geometries show different apparent inclinations within the same source field; their correspondence does not establish field-aligned propagation. These amplitudes describe this supplementary product and do not define a rescaling of the main-manuscript figures.

S3.2. Vertical and magnetic-field-line section comparison

Figure S5. The eight panels compare two sampling geometries over the same 400-km height interval and an approximately 3,400-km reference path. In S3, a broad positive feature occupies much of the approximately 700–1,500-km interval below about 350 km in both representations, although its boundaries and apparent inclination change with the sampling geometry. S1 and S2 contain alternating positive and negative features, while extensive gaps remain in the more distant portions of several sections. Because the S2 display limit is a factor of three smaller than that of the other rows, visual contrast alone cannot rank the four epochs by physical amplitude. The comparison describes geometric continuity within the supplied filtered product. It provides no independent estimate of 50-km vertical resolving power, no vertical phase speed, and no evidence that the disturbed plasma actually moves along the drawn reference magnetic field lines.

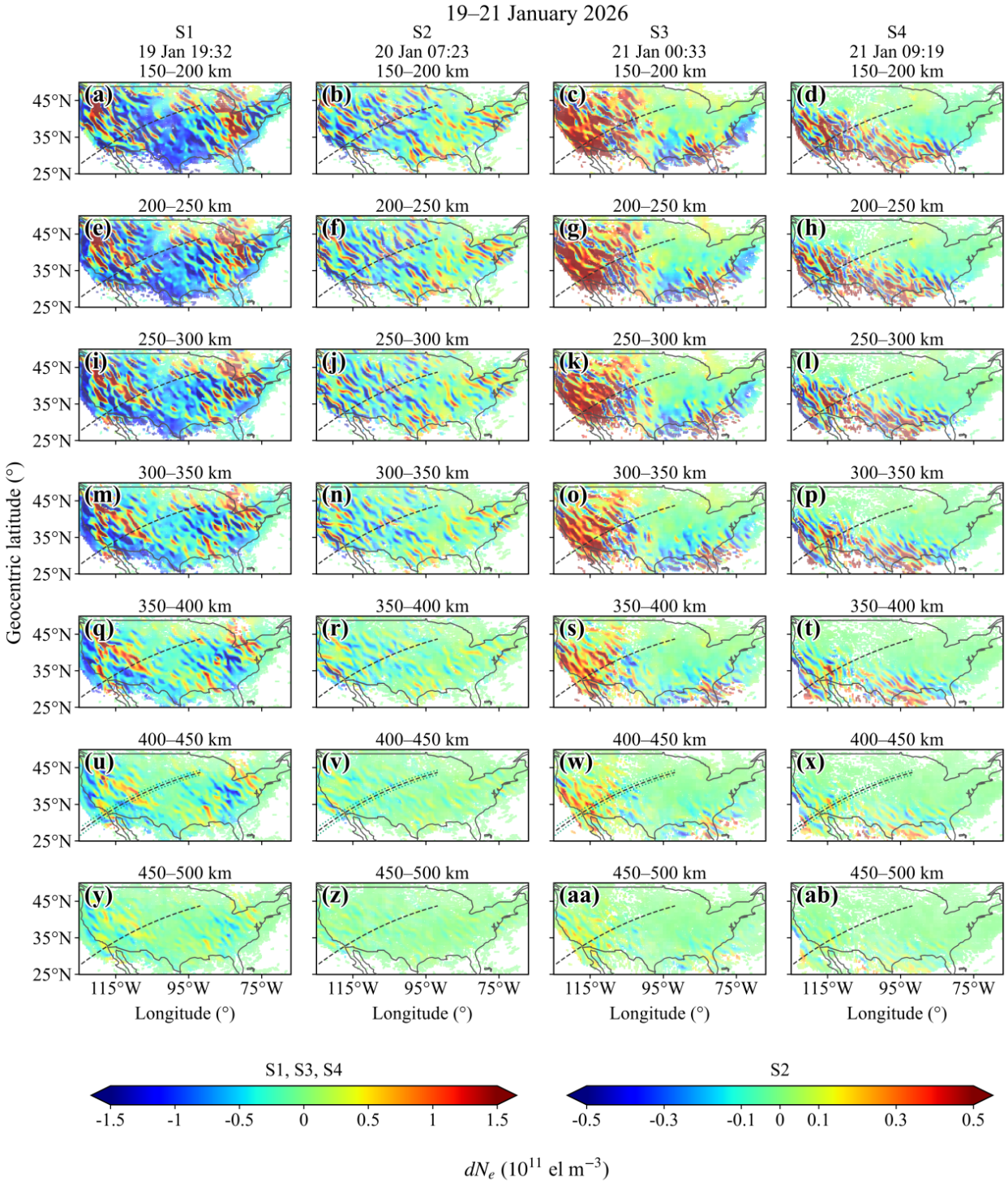


Figure S4: Selected altitude slices at S1–S4. Panels (a)–(ab) show seven 50-km bands from 150 to 500 km. The source directory timestamps appear above the columns. The two color bars apply to S1/S3/S4 and S2, respectively. The section panels and full caption continue on the next page.

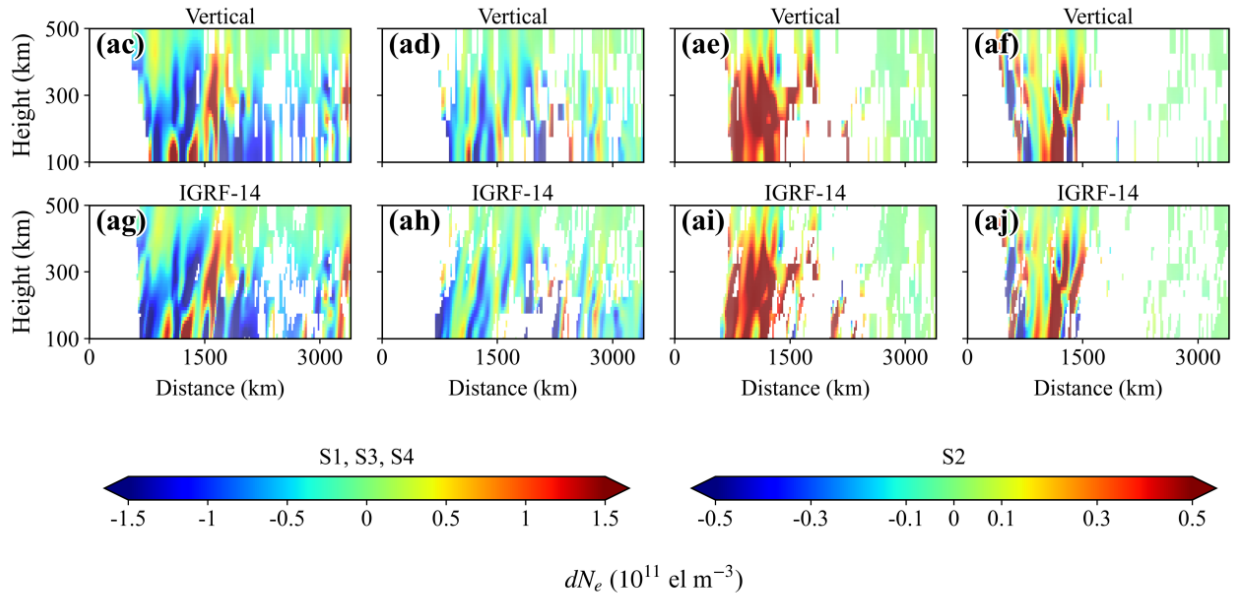


Figure S4 (continued): Height-dependent electron-density residual structure at four source-selected epochs: S1, 19 January 2026 at 19:32; S2, 20 January at 07:23; S3, 21 January at 00:33; and S4, 21 January at 09:19. Time labels are reproduced from the source directories. Panels (a)–(ab) show seven 50-km horizontal bands spanning 150–500 km. Panels (ac)–(af) show vertical sections, and panels (ag)–(aj) show sections sampled along IGRF-14 reference field lines anchored to the same geographic path at 300 km; both section types span 100–500 km. Black dashed lines mark the geographic path, and green dotted lines in the 400–450-km maps mark field-line sampling positions. Colors show the source dN_e in 10^{11} electrons m^{-3}: the displayed limits are ± 1.5 for S1, S3, and S4, and ± 0.5 for S2. The supplied processing description specifies anisotropic edge-preserving spatial filtering without vertical smoothing or amplitude gain; fading represents geometric support and processing sensitivity. White areas retain missing values. Different column scales must be respected when comparing amplitudes. S1–S4 identify selected epochs, not independently verified storm-phase boundaries.

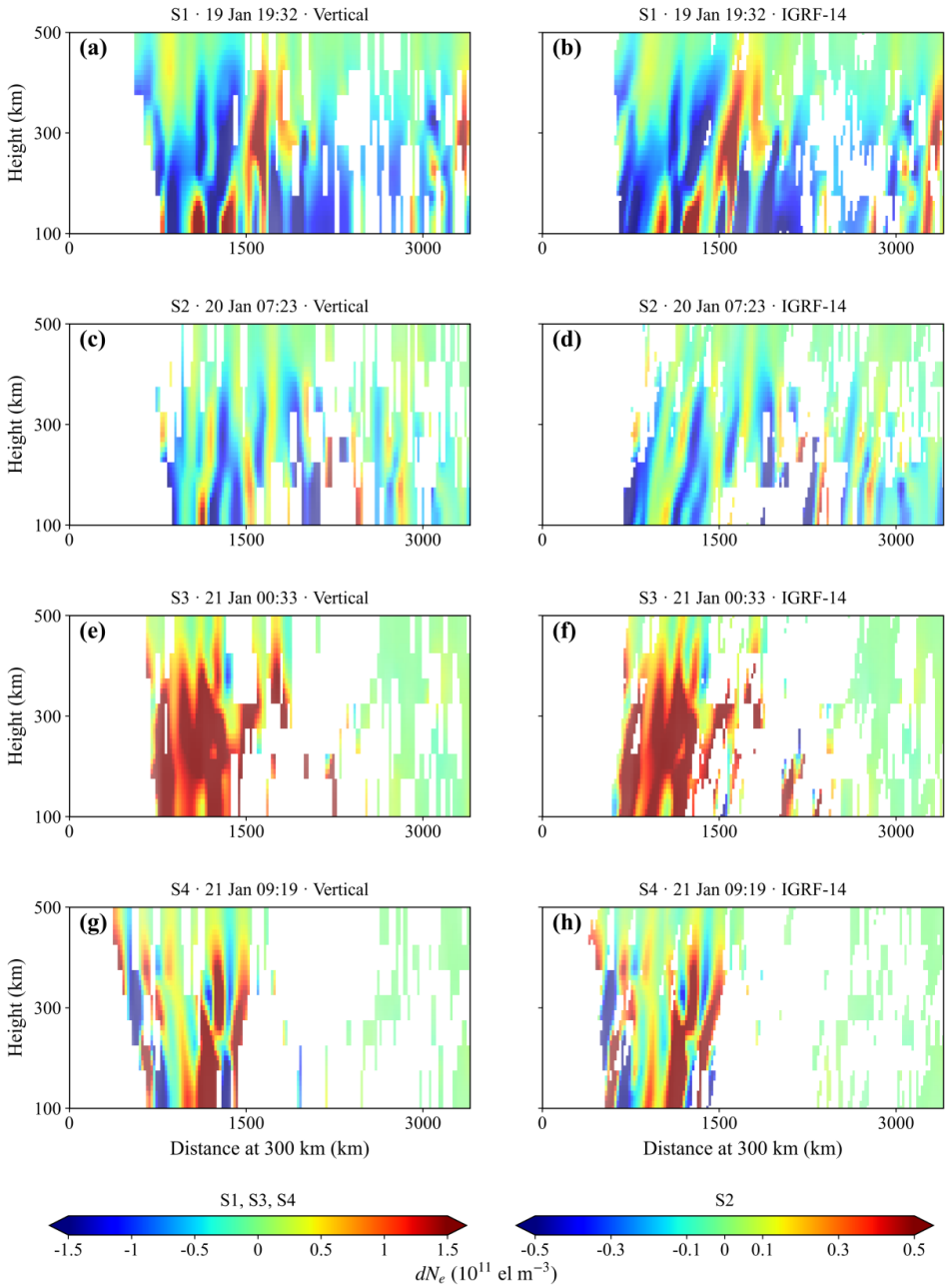


Figure S5: Vertical and magnetic-field-line sampling of the same filtered residual field at the four epochs S1–S4 defined in Fig. S4. Panels (a), (c), (e), and (g) show vertical sections along the shared geographic path; panels (b), (d), (f), and (h) show the corresponding IGRF-14 field-line sections anchored to that path at 300 km. The abscissa is distance along the 300-km reference path, and the ordinate is radial height from 100 to 500 km. Native height bins are 50 km thick; the source description specifies interpolation only between adjacent valid nodes, no bridging of missing values, and use of the outermost-bin representatives over 100–125 and 475–500 km. The displayed scales are $\pm 1.5 \times 10^{11}$ electrons m^{-3} for S1, S3, and S4 and $\pm 0.5 \times 10^{11}$ electrons m^{-3} for S2. Fading follows the source support/sensitivity convention. Interpolation adds no independent vertical information, and field-line sampling does not itself demonstrate field-aligned structuring or propagation.

Supplementary Note S4. Spatial, temporal, and cross-altitude diagnostics

S4.1. Spatial and time-frequency diagnostics in the 400–450 km layer

Figure S6. The S1 diagnostics resolve the same 50-km layer in geographic, distance–time, and period–time coordinates. The six-hour display is centered on 19 January 19:31:42 UTC, and the A–B section extends approximately 3600 km at the 425-km reference altitude. Alternating positive and negative features locally reach the $\pm 0.8 \times 10^{11}$ electrons m^{-3} display levels. The 120-min diagnostic interval isolates a shorter portion of this evolving field, while the wavelet panel explicitly separates the 10–60-min analysis band from longer periods. The visually strongest power occupies approximately 60–80 min, with part of that enhancement within the marked edge region. Consequently, the color maximum cannot by itself identify a well-constrained 10–60-min wave packet. The maps and sections establish a spatially heterogeneous perturbation field with temporally varying spectral content; an identified propagating feature would additionally require consistent phase tracking beyond the section geometry and its observation-support mask.

S4.2. Four-window comparison of horizontal structure and local wavelet power

Figure S7. The four 360-min windows compare the same 400–450-km layer and A–B path on a common density scale. S1 and S3 contain extensive alternating structures near the $\pm 8 \times 10^{10}$ electrons m^{-3} display levels, whereas S2 has substantially lower contrast over much of the section. S4 combines localized western structure with relatively weak perturbations farther along the path. The row-specific wavelet scales differ materially: their labeled ticks include 6 and 12 in S1, 0.08 and 0.16 in S2, 4 and 8 in S3, and 0.4 and 0.8 in S4, in the displayed power units. Equal colors therefore do not indicate equal power between windows. S3 additionally crosses a source-day boundary that divides the local wavelet analysis into two separately processed intervals. The comparison supports differences in regional contrast and local spectral expression, while the available panels do not establish a common phase trajectory across the four windows or a continuous wave spanning the S3 processing boundary.

S4.3. Local altitude-time structure and band-limited waveforms

Figure S8. The local altitude–time display samples eight 50-km layers over 100–500 km, while the lower panels compare three upper-layer centers separated by 50 km. The common density color scale makes the marked differences between S1–S4 directly visible, but the band-limited curves use different ordinate ranges and must be read in physical units. For example, the colored 375-km waveform in S2 reaches approximately $\pm 0.4 \times 10^{11}$ electrons m^{-3}, whereas S4 includes excursions approaching -1.1×10^{11} and $+0.8 \times 10^{11}$ electrons m^{-3}. Several S1 excursions exceed the upper-panel 0.8×10^{11} color level in magnitude, illustrating why a saturated heat-map color cannot specify the underlying waveform amplitude. In S3, the day boundary and excluded edge intervals interrupt the inferentially usable series. The displayed differences quantify height-dependent residual amplitudes and temporal organization at a fixed location; the plots alone do not establish vertical phase propagation, transport, or an altitude-independent oscillation period.

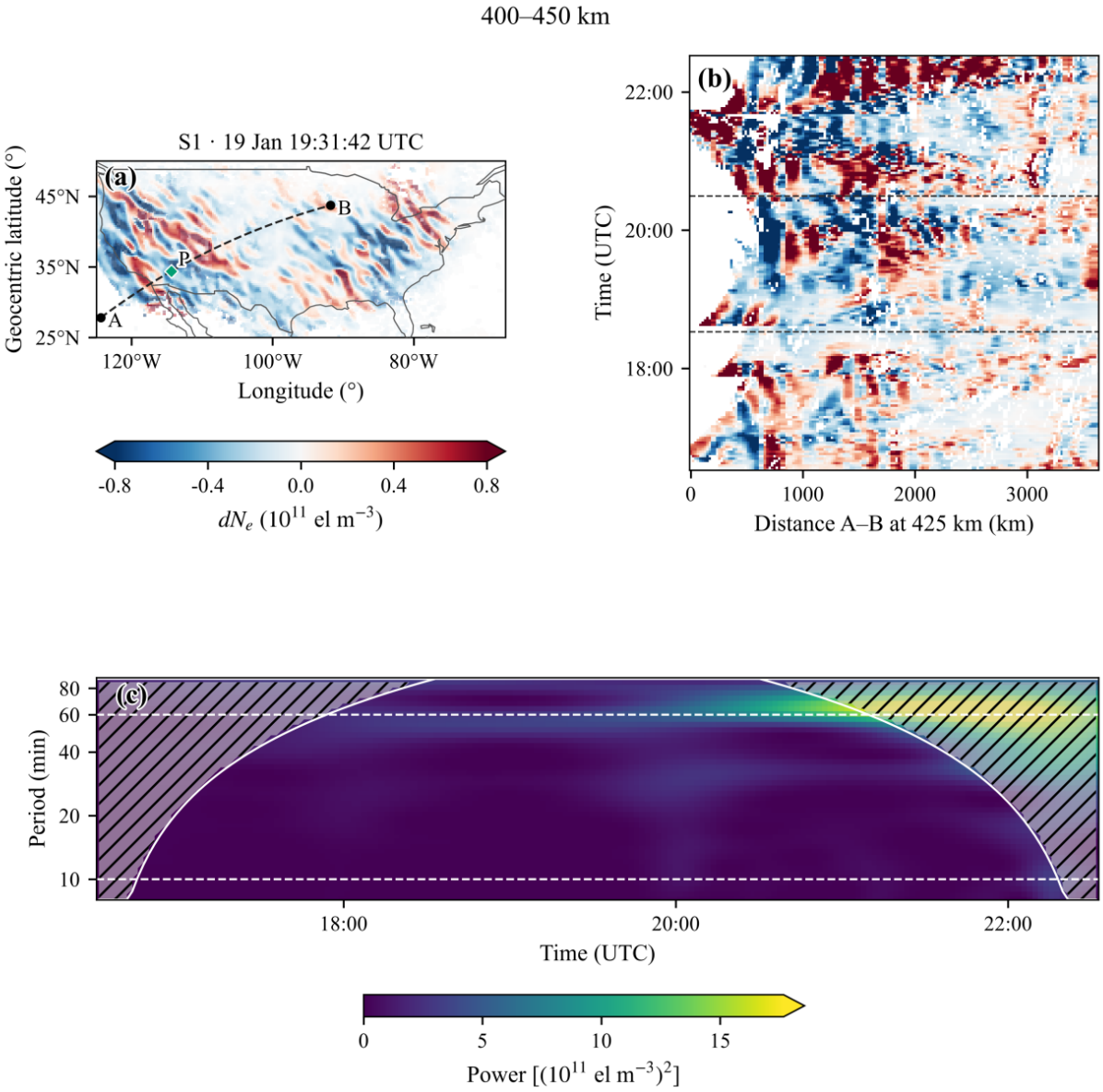


Figure S6: Spatial and time–frequency diagnostics of the 400–450 km electron-density perturbation layer. (a) Filtered dN_e at 19:31:42 UTC on 19 January 2026 (S1); the dashed A–B path defines the spatial section, and P marks the local wavelet-analysis position. Color fading represents the contemporaneous receiver-ray support. (b) Distance–time section along A–B, with distance evaluated on the 425-km spherical shell; horizontal dashed lines delimit the prescribed 120-min diagnostic interval. (c) Morlet-6 wavelet power at P. Dashed horizontal lines delimit the 10–60-min analysis band; hatched regions identify edge-affected power and are excluded from quantitative interpretation. The map and distance–time display use a zero-centered linear scale of $\pm 0.8 \times 10^{11}$ electrons m^{-3}; color-bar extensions denote values outside the displayed range. Wavelet power is expressed in (10^{11} electrons m^{-3})^2. Missing values remain blank. UTC follows the source's adopted GPS-time-minus-18-s conversion. These diagnostics describe spatial structure and temporal power; they do not independently establish horizontal propagation.

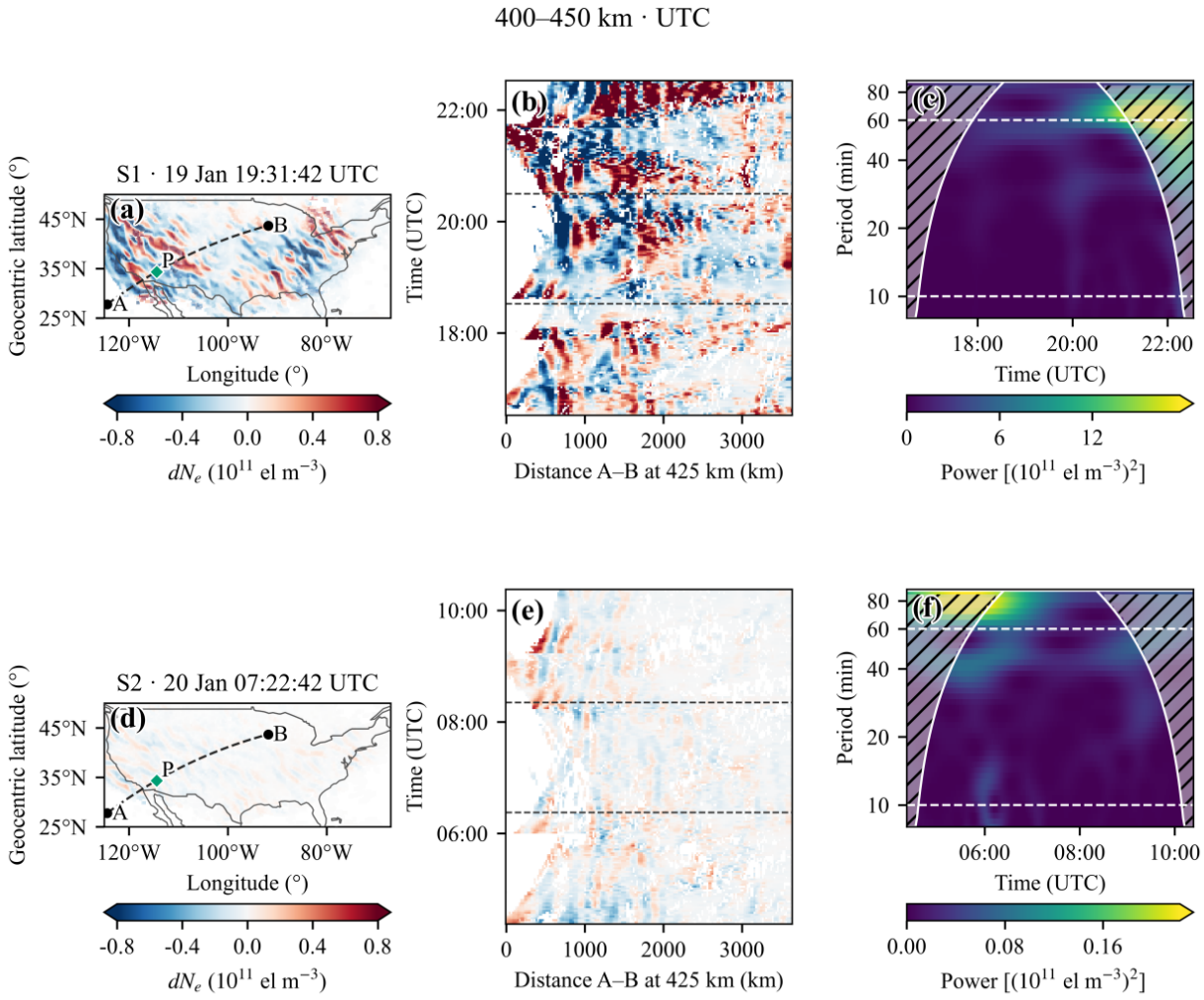


Figure S7: Comparison of 425-km disturbance diagnostics across the S1–S4 windows. Rows correspond to 19 January 19:31:42, 20 January 07:22:42, 21 January 00:32:42, and 21 January 09:18:42 UTC, respectively. Each window extends 180 min before and after its selected center. The columns show the central-time 400–450-km perturbation map, the A–B distance–time section, and Morlet-6 wavelet power at P. P is selected using valid temporal support and remains fixed within each window. Maps and distance–time panels share the $\pm 0.8 \times 10^{11}$ electrons m^{-3} scale; wavelet-power scales are labeled separately for each row. Dashed horizontal lines in the sections delimit the prescribed diagnostic subwindows, while the wavelet panels mark the 10–60-min period band. Dotted lines identify source-day boundaries. S3 is analyzed separately on either side of the boundary, without joining phase across it; hatched wavelet edge regions remain visible. Missing regions are blank, and reduced ray support is represented by color fading. UTC follows the adopted GPS-time-minus-18-s conversion. The present page shows S1 and S2, panels (a)–(f); S3 and S4 continue on the next page.

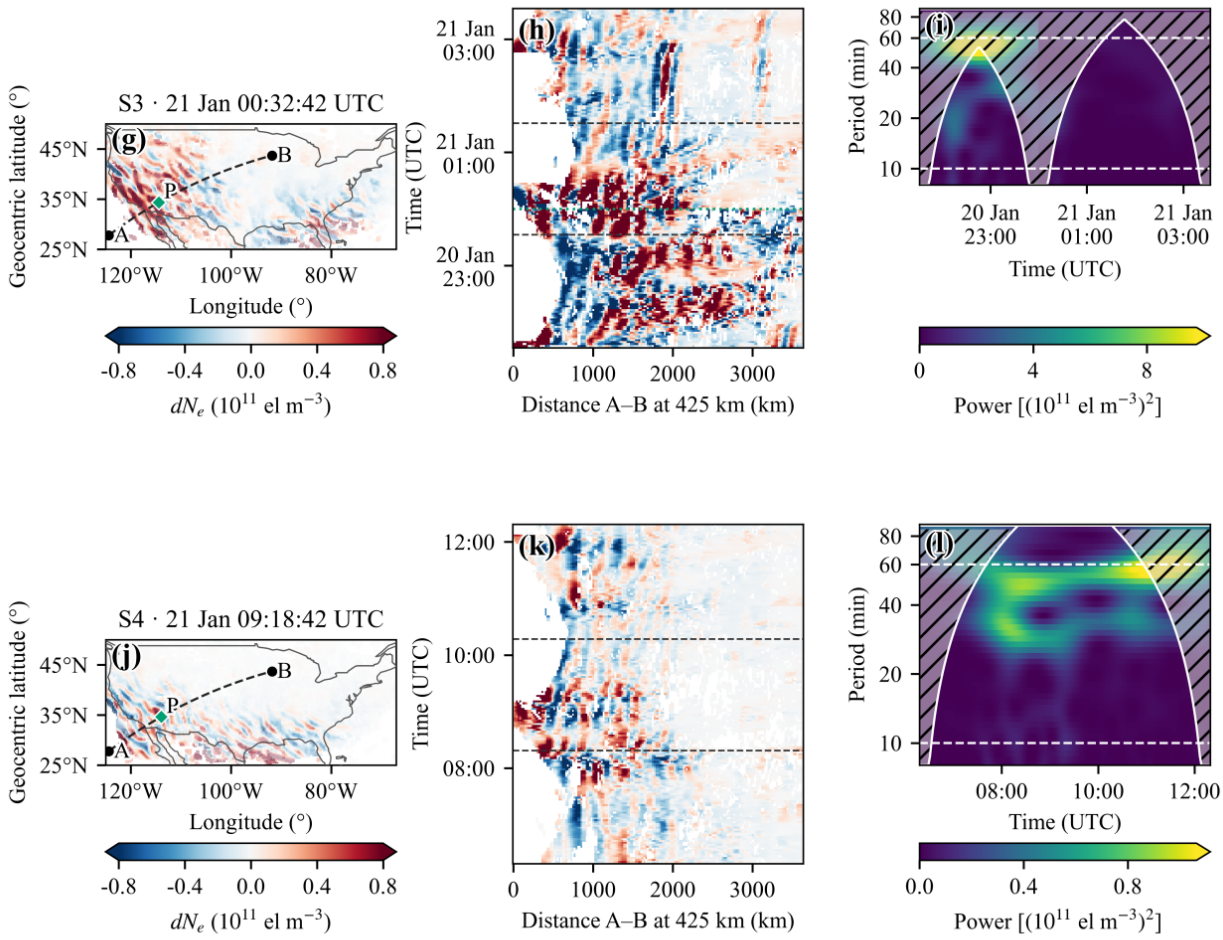


Figure S7 (continued): Four-window diagnostics continued: S3 and S4, panels (g)–(l). The density scale remains $\pm 0.8 \times 10^{11}$ electrons m^{-3}, while the wavelet-power bars are specific to each row. The S3 day boundary divides the local wavelet analysis into separate intervals. All conventions are as defined on the preceding page.

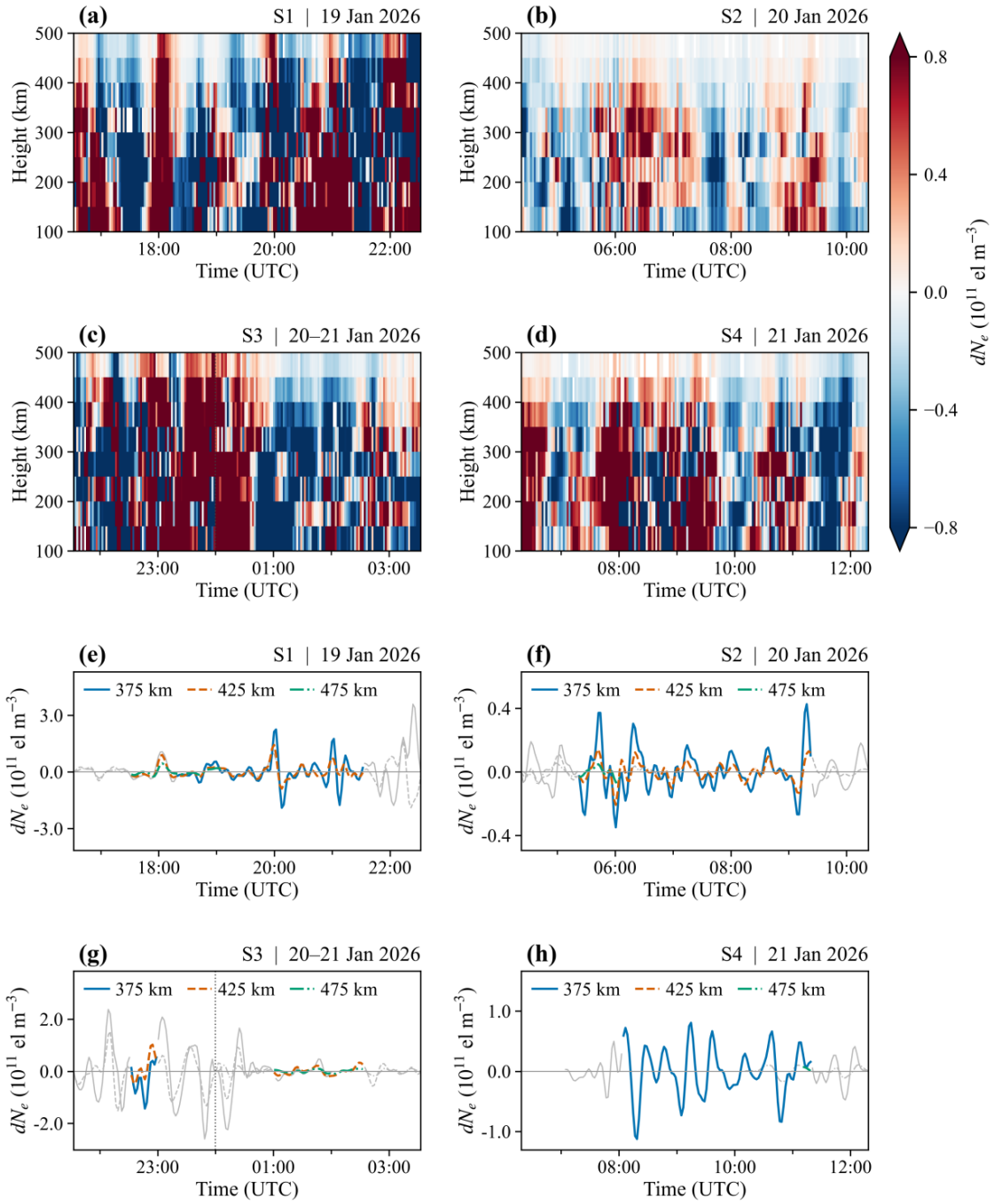


Figure S8: Cross-altitude disturbance structure at 114.375°W, 34.375°N. Panels (a)–(d) show the S1–S4 altitude–time distributions of dN_e over 100–500 km, using the native 50-km altitude bins and the supplied structure-preserving spatial filtering. The common zero-centered color scale is $\pm 0.8 \times 10^{11}$ electrons m^{-3}; triangular extensions indicate values beyond the displayed range, and missing values remain blank. Panels (e)–(h) show the corresponding 10–60-min band-limited series at 375, 425, and 475 km, in the original density units and without normalization or vertical offsets. Colored shading represents sensitivity to the spatial filtering, while gray curves denote filter-edge intervals excluded from inference. Processing is performed separately on the two sides of each source-day boundary; the boundary is marked for S3. The waveform ordinates differ between windows. The 50-km spacing specifies altitude sampling and does not establish independent vertical resolving power.

Supplementary Note S5. Column cancellation and processing sensitivity

S5.1. Column cancellation and the reported value of altitude-resolved information

Figure S9. The two selected profiles demonstrate that similar signed column changes can correspond to substantially different altitude structure. Their net columns differ by only 0.079 TECU, whereas the absolute column for B is approximately 3.92 times that for A. The reported signed and absolute values imply cancellation of approximately 20.9% and 75.8%, respectively. Profile B combines -0.505×10^{11} electrons m^{-3} in the 100–250-km interval with $+0.237 \times 10^{11}$ electrons m^{-3} over 700–1000 km, so the small net column conceals large opposite-sign contributions. In the reported held-out comparison, paired MAE differences are +0.045, +0.088, and +2.357 TECU for the strong two-dimensional, two-layer, and column-ablation alternatives, respectively. These are source-reported predictive contrasts on 52,500 targets from four receivers. They describe the displayed evaluation rather than establish absolute density accuracy. The filter-response control additionally shows that temporal processing can introduce opposite-sign lobes, which must be considered when interpreting the cancellation diagnostic.

S5.2. Four-day evolution and processing sensitivity of conditional stage summaries

Figure S10. The source-defined stage comparison uses two main-stage, four re-enhancement, and 24 recovery samples. Under the $H120$ baseline, VTEC-residual RMS is lower by 0.74 TECU in E than M and by 0.63 TECU in R than M, with 6.04% common spatial support. Restricting the analysis to strict common support gives corresponding differences of −0.78 and −0.65 TECU. The contrast magnitude changes with detrending: the $H90$ cases give −0.48 and −0.40 TECU, whereas $H150$ gives −1.05 and −0.93 TECU. Thirty-minute aggregation reduces these differences to −0.15 and −0.05 TECU. The displayed common-area fractions range from 0.40% for fixed network B to 15.60% for the 1° aggregation grid; the 30-min row reports 0.00%. These contrasts demonstrate sensitivity to processing and observational support. The conditional-center estimates remain near 249 km, but they do not identify a physical layer height, and the plotted intervals cannot be assigned an unstated confidence level.

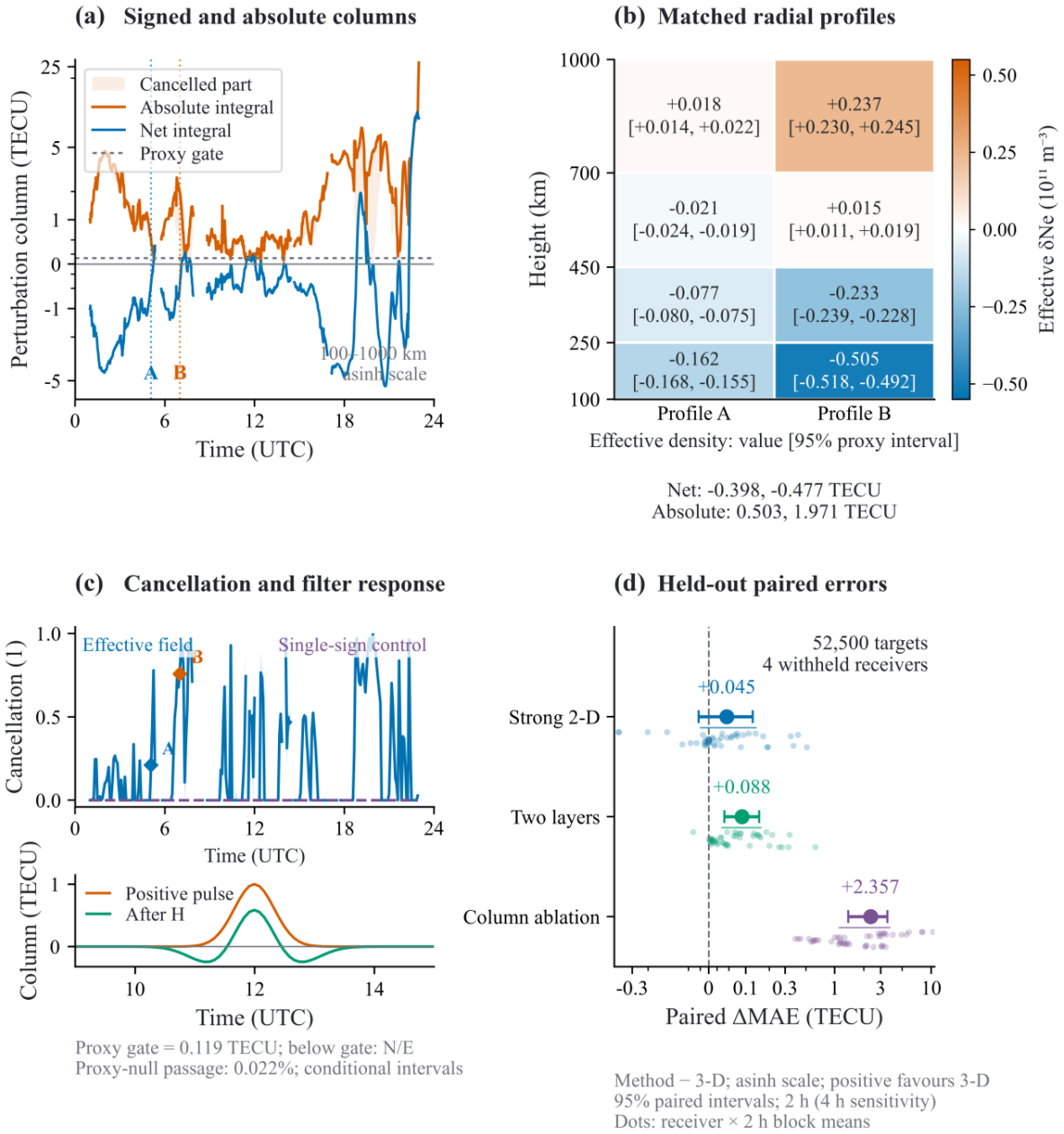


Figure S9: Column cancellation and the reported predictive contribution of altitude-resolved information. (a) Signed and absolute perturbation-column integrals over the same valid 100–1000-km height interval, displayed on an inverse-hyperbolic-sine scale. The shaded difference represents cancellation, and A and B mark the selected profiles. (b) Effective residual densities over 100–250, 250–450, 450–700, and 700–1000 km for profiles A and B; labels give the values and the source's 95% proxy intervals in 10^{11} electrons m^{-3}. Their signed columns are −0.398 and −0.477 TECU, and their absolute columns are 0.503 and 1.971 TECU. (c) Cancellation diagnostic, single-sign control, and the response of a positive pulse to the depicted temporal operator H. The proxy gate is 0.119 TECU; subthreshold cases are not evaluated, and the reported proxy-null passage is 0.022%. (d) Paired MAE differences, defined as comparison method minus three-dimensional estimate, for 52,500 targets from four withheld receivers. Positive differences favor the three-dimensional representation. Points show receiver-by-2-h block means; the displayed 95% paired intervals use 2-h blocking, with a 4-h sensitivity comparison. The source does not specify the calendar date of panels (a)–(c), and proxy intervals are not independently established density-accuracy intervals.

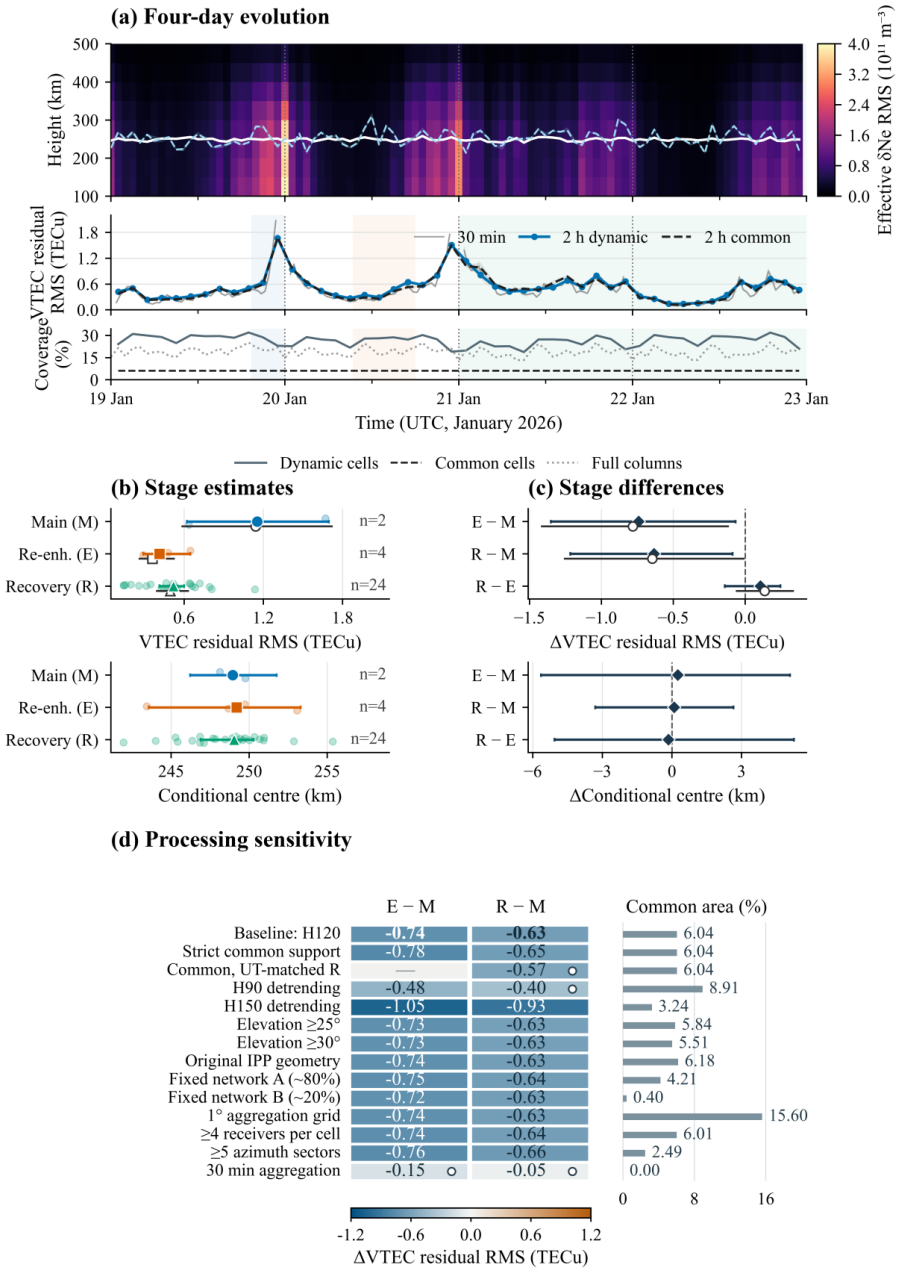


Figure S10: Four-day evolution and processing sensitivity of the supplied event diagnostics. (a) Effective residual-density RMS over 100–500 km during 19–22 January 2026, VTEC-residual RMS under 30-min and 2-h aggregation, and coverage fractions for dynamically supported cells, common cells, and complete columns. The overlaid altitude curves are conditional center diagnostics. (b) Source-defined main (M), re-enhancement (E), and recovery (R) summaries for VTEC-residual RMS and conditional center, with annotated sample counts $n = 2$, 4, and 24. (c) Pairwise stage differences for the same quantities. (d) VTEC-residual RMS differences under alternative support masks, detrending settings, elevation thresholds, geometry, station subsets, grid aggregation, and coverage criteria; the adjacent bars report the common area. The baseline is labeled $H120$, and the $H90$ and $H150$ cases are retained as source processing labels. Figure-level uncertainty marks are reproduced as supplied; their complete interval definitions and the precise stage boundaries are not documented in the accompanying caption. Conditional center is not interpreted as F2-peak height or a plasma-displacement measurement.

Supplementary Note S6. Supplementary animations

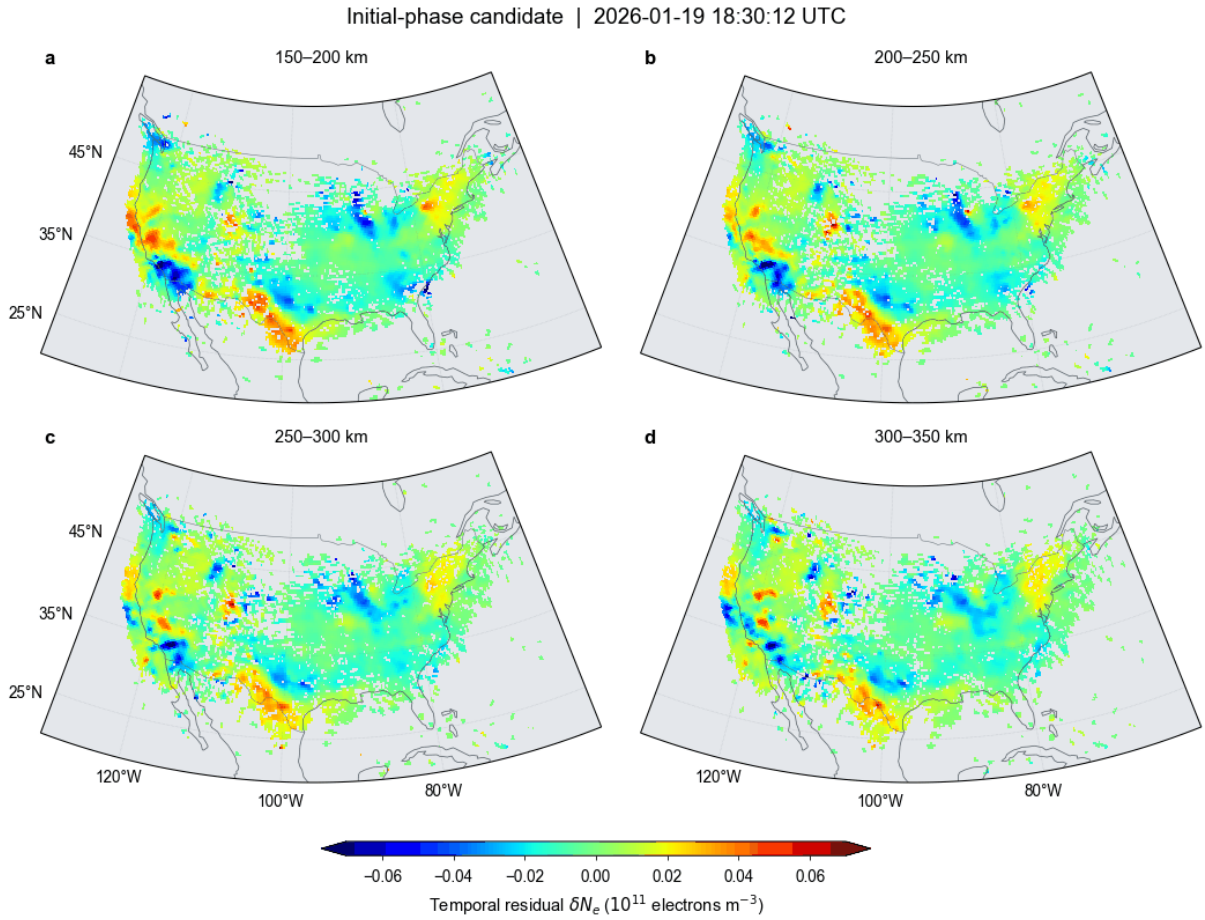
The seven animations preserve the original GIF frames and playback settings. Movies S1, S2, S3, and S5 display electron-density residuals in four 50-km layers; Movies S4, S6, and S7 display regional dVTEC. They are numbered by the actual dates and times printed in their frames. Similarly named source files need not describe the same epoch or physical quantity.

Each file contains 31 frames spanning 15 min, with nominal 30-s spacing between displayed epochs. The 120-ms frame duration gives a 3.72-s playback cycle and is unrelated to physical wave speed. This display spacing does not establish the cadence of an archived tomographic solution. The dVTEC frames explicitly report a 30° elevation cutoff, whereas the proposed reconstruction protocol uses 35°; those processing conventions are retained separately.

The PDF provides the middle frame of each sequence and its caption. The complete animations are supplied as `movies/Movie_S1.gif` through `Movie_S7.gif`. They can be opened directly, through the accompanying GIF links, or from the offline index `movies/index.html`. UTC is printed on the density sequences; the dVTEC sequences use UT as labeled in their source frames, without an additional time conversion here.

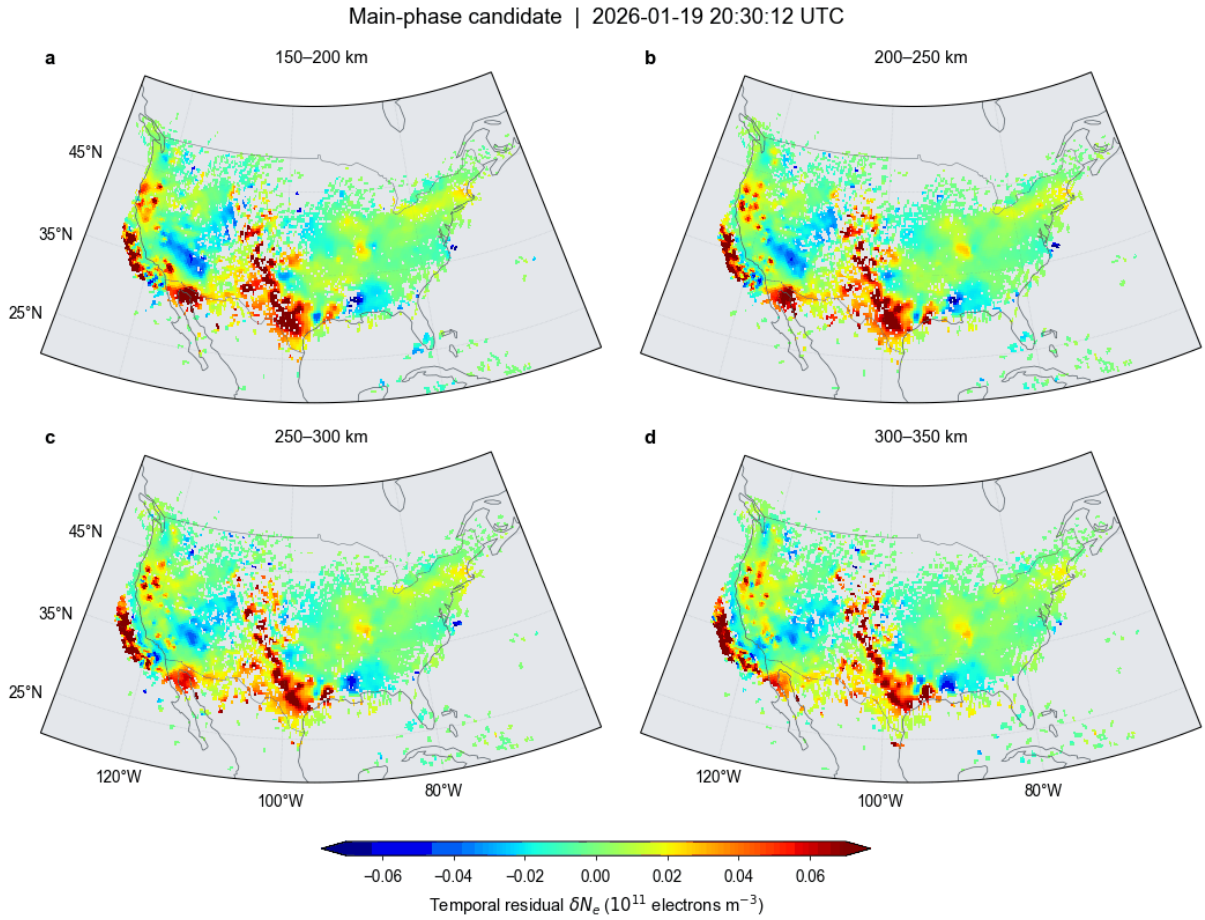
Table S1: Supplementary animation inventory. All dates are in January 2026. Every sequence has 31 frames and a 15-min physical display span; the complete files retain their original frame data.

Movie	Quantity	Date	Printed time interval	File
S1	δN_e	19 Jan.	18:22:42–18:37:42 UTC	`Movie_S1.gif`
S2	δN_e	19 Jan.	20:22:42–20:37:42 UTC	`Movie_S2.gif`
S3	δN_e	20 Jan.	11:42:42–11:57:42 UTC	`Movie_S3.gif`
S4	dVTEC	20 Jan.	15:22:42–15:37:42 UT	`Movie_S4.gif`
S5	δN_e	20 Jan.	20:22:42–20:37:42 UTC	`Movie_S5.gif`
S6	dVTEC	21 Jan.	11:52:42–12:07:42 UT	`Movie_S6.gif`
S7	dVTEC	21 Jan.	19:52:42–20:07:42 UT	`Movie_S7.gif`



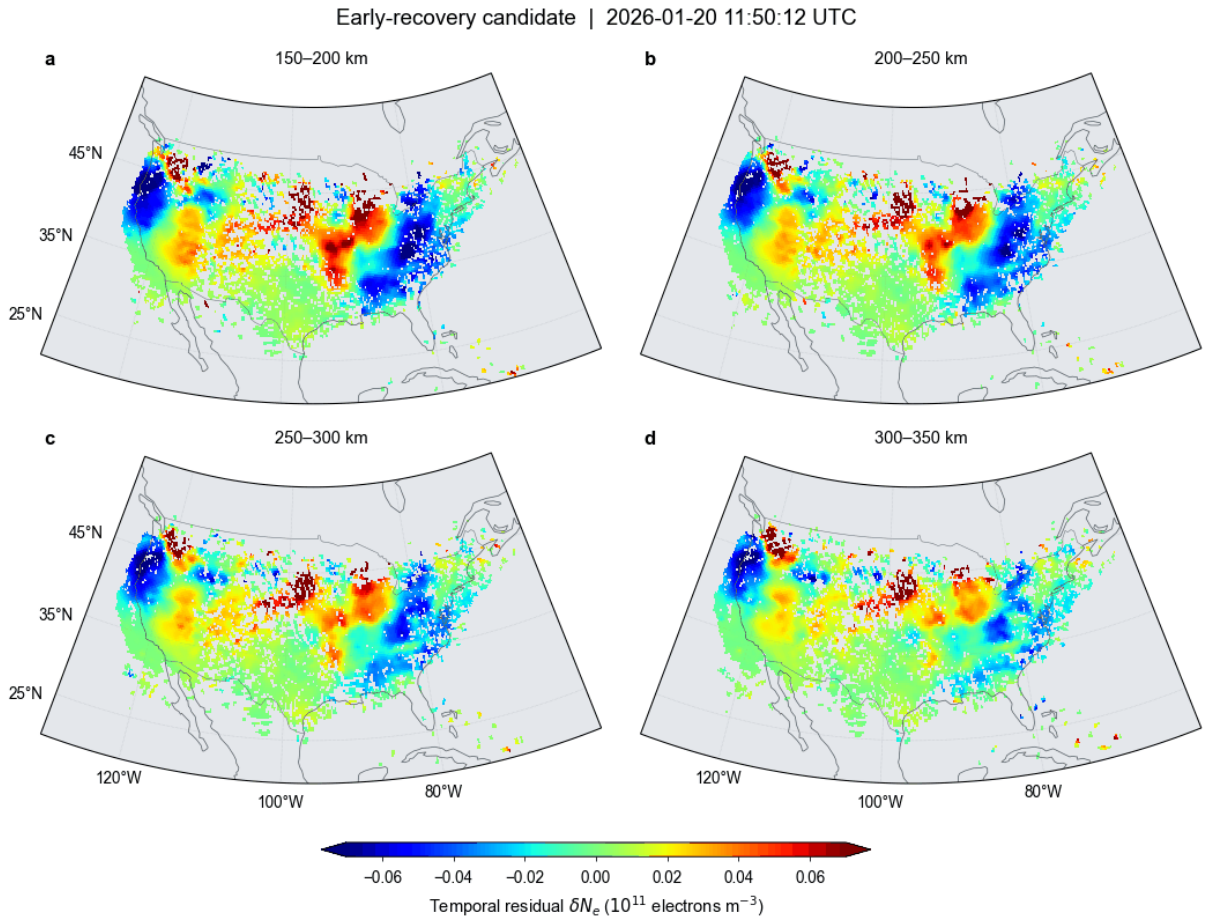
Movie S1: Electron-density residuals during 18:22:42–18:37:42 UTC on 19 January 2026, centered at 18:30:12. Panels (a)–(d) show signed temporal electron-density residuals, δN_e, in the 150–200, 200–250, 250–300, and 300–350 km layers, respectively, on geographic coordinates. The common color bar has labeled ticks from -0.06 to $+0.06$ in increments of 0.02, in units of 10^{11} electrons m^{-3}; one labeled interval therefore represents 2×10^9 electrons m^{-3}. Color-bar extensions indicate values outside the plotted range, and gray areas contain no displayed residual values. The 31 displayed epochs span 15 min with nominal 30-s spacing; the GIF uses 120 ms per frame and a 3.72-s repeating playback cycle. Playback speed is a display setting, not a physical propagation speed. The entire interval precedes the reported 19:24 storm commencement [1]; the original "Initial-phase candidate" heading does not establish an initial-phase classification. The first, middle, and last frames are separated by 7 min 30 s. They show substantial regional residual structure before the reported storm commencement, including negative patches in the southwestern sector and positive features farther east in the south-central region. At the middle epoch, selected southwestern residuals reach the $-(4\text{–}6) \times 10^9$ electrons m^{-3} labeled levels, while the final frame contains stronger positive structure in the western sector. The broad patterns recur across all four displayed heights, with location-dependent differences in contrast. The animation therefore provides a pre-commencement comparison; the presence of a signed residual alone does not identify a storm-driven disturbance or a tracked wave.

Open accompanying GIF: `Movie_S1.gif`.



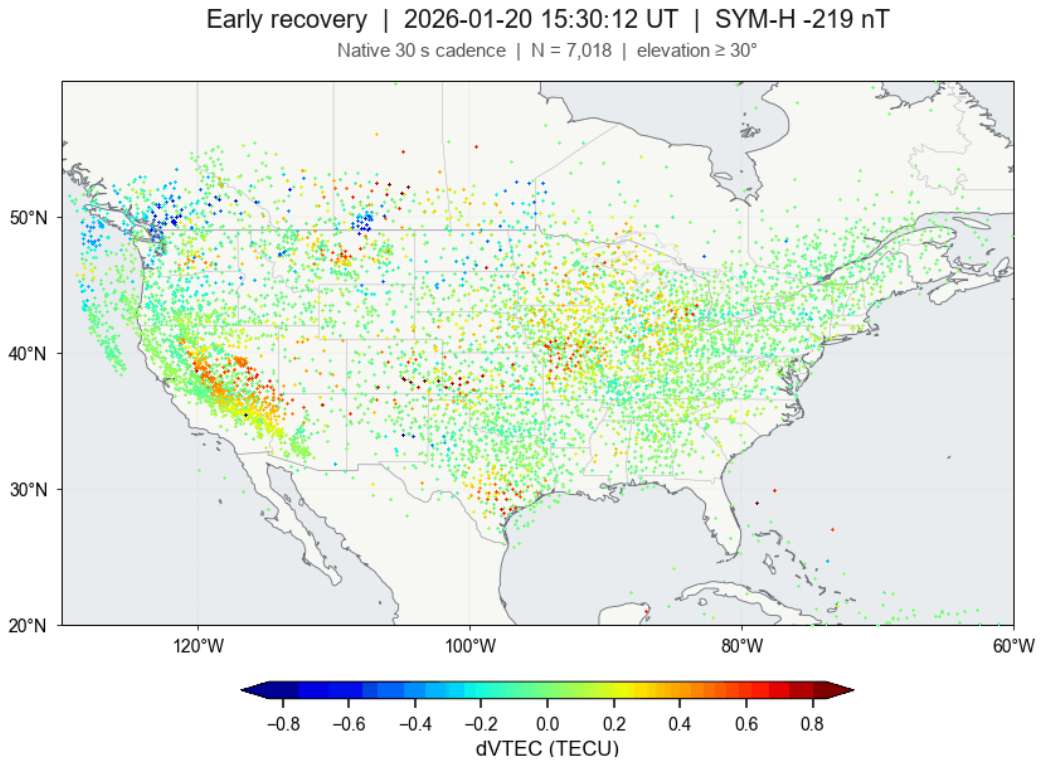
Movie S2: Electron-density residuals during 20:22:42–20:37:42 UTC on 19 January 2026, centered at 20:30:12. Panels (a)–(d) show signed temporal electron-density residuals, δN_e, in the 150–200, 200–250, 250–300, and 300–350 km layers, respectively, on geographic coordinates. The common color bar has labeled ticks from -0.06 to $+0.06$ in increments of 0.02, in units of 10^{11} electrons m^{-3}; one labeled interval therefore represents 2×10^9 electrons m^{-3}. Color-bar extensions indicate values outside the plotted range, and gray areas contain no displayed residual values. The 31 displayed epochs span 15 min with nominal 30-s spacing; the GIF uses 120 ms per frame and a 3.72-s repeating playback cycle. Playback speed is a display setting, not a physical propagation speed. The selected interval lies within the first storm intensification; its timestamps determine the interpretation of the source "Main-phase candidate" heading. The sequence samples a 15-min interval during the first intensification. Positive structures in the western and south-central sectors coexist with negative residuals over parts of the western interior. Local displayed values reach the $\pm 6 \times 10^9$ electrons m^{-3} labeled levels, with some darker positive portions extending beyond the highest labeled tick. The south-central positive feature is conspicuous near the beginning and middle of the sequence, while the final inspected frame contains more extensive negative structure across parts of the interior. Corresponding broad patterns occur in the four 50-km layers, but their contrast varies with altitude. These changing displays describe regional reorganization; they do not provide a fitted velocity or a uniquely identified propagation path.

Open accompanying GIF: `Movie_S2.gif`.



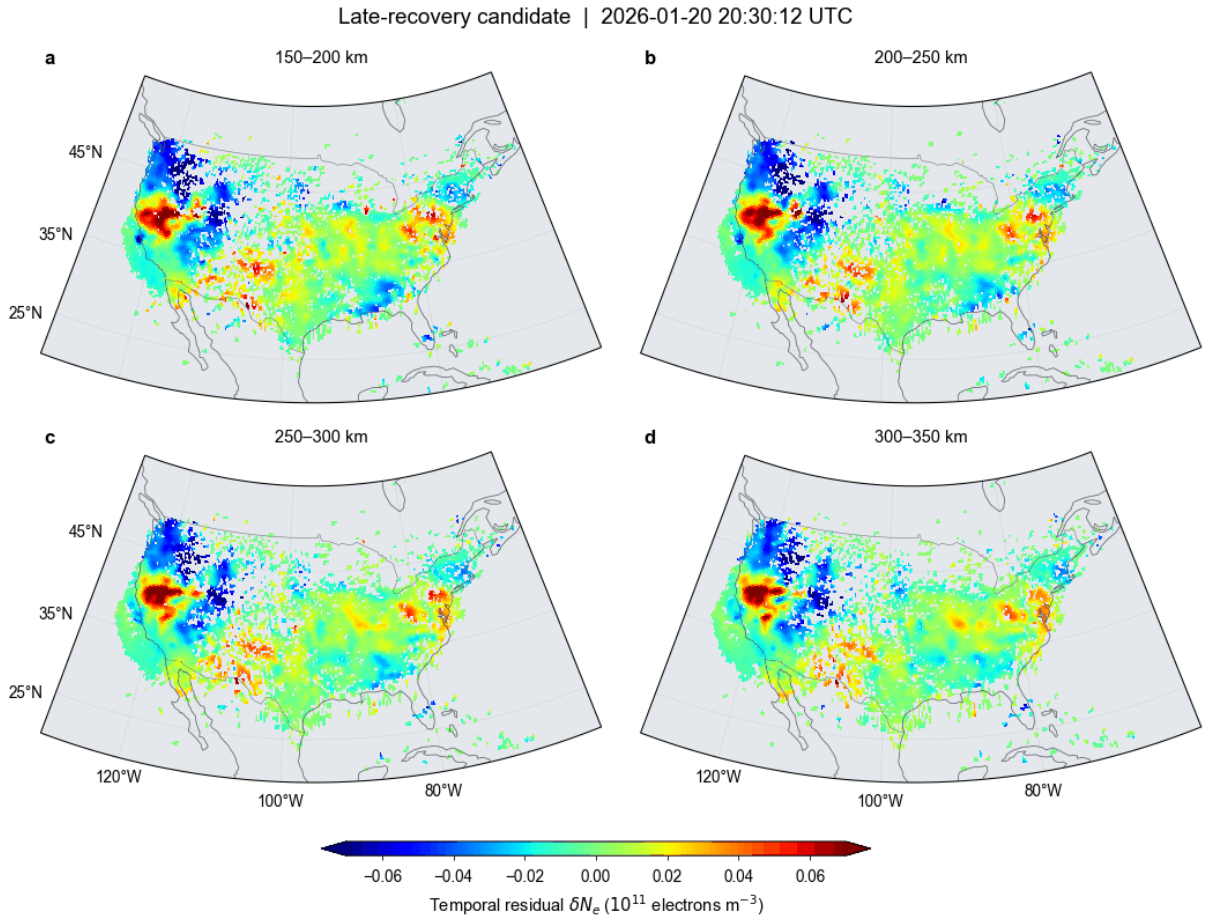
Movie S3: Electron-density residuals during 11:42:42–11:57:42 UTC on 20 January 2026, centered at 11:50:12. Panels (a)–(d) show signed temporal electron-density residuals, δN_e, in the 150–200, 200–250, 250–300, and 300–350 km layers, respectively, on geographic coordinates. The common color bar has labeled ticks from -0.06 to $+0.06$ in increments of 0.02, in units of 10^{11} electrons m^{-3}; one labeled interval therefore represents 2×10^9 electrons m^{-3}. Color-bar extensions indicate values outside the plotted range, and gray areas contain no displayed residual values. The 31 displayed epochs span 15 min with nominal 30-s spacing; the GIF uses 120 ms per frame and a 3.72-s repeating playback cycle. Playback speed is a display setting, not a physical propagation speed. The interval is described by its timestamps near the second strong geomagnetic depression; the source "Early-recovery candidate" heading is not used to define a verified phase boundary. The middle epoch contains pronounced negative structure in the western coastal sector and in parts of the east, separated by positive central features. The four layers provide a direct visual comparison over 150–350 km, with color intervals of 2×10^9 electrons m^{-3}. Positive and negative local features reach the $\pm 6 \times 10^9$ electrons m^{-3} labeled levels. Between the first and final inspected frames, the relative prominence and location of the western and east-central features change substantially. The animation makes the short-term reorganization apparent, but the imagery alone does not establish that individual patches belong to one continuously tracked crest. The 11:50:12 center time should not be conflated with the 15:30:12 center of the separately supplied dVTEC animation.

Open accompanying GIF: `Movie_S3.gif`.



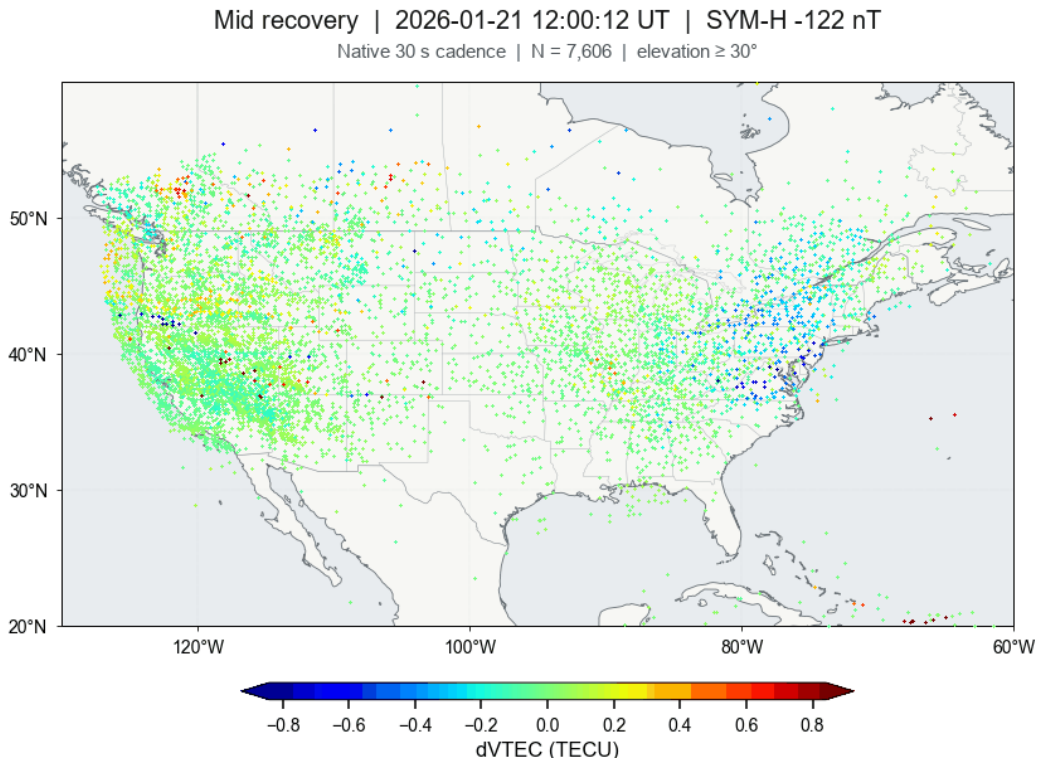
Movie S4: Regional dVTEC during 15:22:42–15:37:42 UT on 20 January 2026, centered at 15:30:12. Colors show detrended vertical total electron content, dVTEC, in TECU on geographic coordinates, with labeled ticks from −0.8 to +0.8 at intervals of 0.2 TECU and extensions at both ends. The displayed region spans approximately 130°–60°W and 20°–60°N. Each frame reports its time in UT, a SYM-H value, the source count N, a native 30-s cadence, and an elevation threshold of 30°. Unplotted locations are not assigned zero perturbation. The 31 displayed epochs span 15 min with nominal 30-s spacing; the GIF uses 120 ms per frame and a 3.72-s repeating playback cycle. Playback speed is a display setting, not a physical propagation speed. The original "Early recovery" heading is retained as a source label; the numerical timestamps and SYM-H annotations govern temporal interpretation. The first, middle, and last frames carry counts $N = 6741$, 7018, and 7244 and SYM-H values of −219, −219, and −226 nT, respectively. Positive dVTEC patches are evident over portions of the southwest and the central sector, while negative features occur locally farther north and in the west. Selected local contrasts approach the labeled 0.6–0.8 TECU intervals. The increasing count and changing spatial sampling are visible alongside the evolving field, so a change in the number of colored locations should not be interpreted as a change in disturbed area alone. The 7-nT decrease between the endpoint SYM-H labels also shows that the source phase heading does not imply monotonic geomagnetic recovery within this 15-min interval.

Open accompanying GIF: `Movie_S4.gif`.



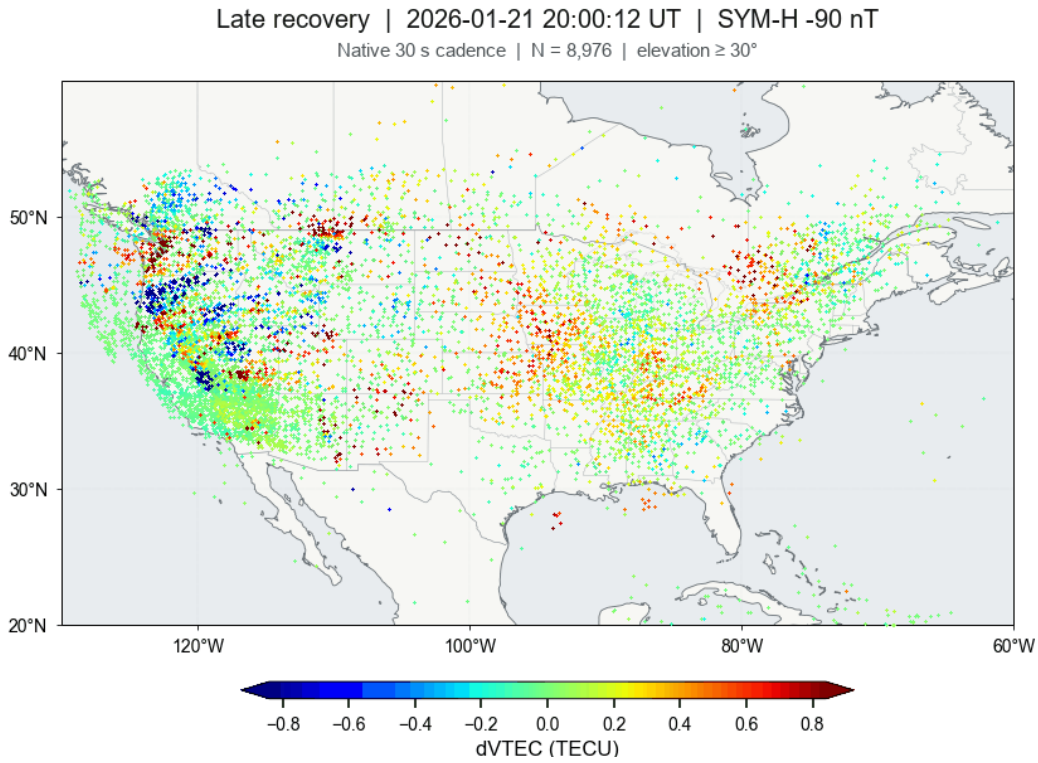
Movie S5: Electron-density residuals during 20:22:42–20:37:42 UTC on 20 January 2026, centered at 20:30:12. Panels (a)–(d) show signed temporal electron-density residuals, δN_e, in the 150–200, 200–250, 250–300, and 300–350 km layers, respectively, on geographic coordinates. The common color bar has labeled ticks from -0.06 to $+0.06$ in increments of 0.02, in units of 10^{11} electrons m^{-3}; one labeled interval therefore represents 2×10^9 electrons m^{-3}. Color-bar extensions indicate values outside the plotted range, and gray areas contain no displayed residual values. The 31 displayed epochs span 15 min with nominal 30-s spacing; the GIF uses 120 ms per frame and a 3.72-s repeating playback cycle. Playback speed is a display setting, not a physical propagation speed. The source "Late-recovery candidate" heading is retained, while the absolute UTC timestamps define this selected comparison interval. The sequence shows a strong positive western feature near 125°–115°W and 35°–40°N adjacent to negative residuals farther north and east. Local positive and negative contrasts reach the $\pm 6 \times 10^9$ electrons m^{-3} labeled levels. Their broad organization is recognizable across the four displayed layers, while smaller features and their amplitudes differ with height. Comparing the first and final inspected frames reveals changes within 15 min without requiring a uniform regional sign or decay rate. The sequence is centered on 20 January and is distinct from the separately supplied dVTEC movie centered at 20:00:12 UT on 21 January; the two "late recovery" filenames do not identify simultaneous products.

Open accompanying GIF: `Movie_S5.gif`.



Movie S6: Regional dVTEC during 11:52:42–12:07:42 UT on 21 January 2026, centered at 12:00:12. Colors show detrended vertical total electron content, dVTEC, in TECU on geographic coordinates, with labeled ticks from −0.8 to +0.8 at intervals of 0.2 TECU and extensions at both ends. The displayed region spans approximately 130°–60°W and 20°–60°N. Each frame reports its time in UT, a SYM-H value, the source count N, a native 30-s cadence, and an elevation threshold of 30°. Unplotted locations are not assigned zero perturbation. The 31 displayed epochs span 15 min with nominal 30-s spacing; the GIF uses 120 ms per frame and a 3.72-s repeating playback cycle. Playback speed is a display setting, not a physical propagation speed. The original "Mid recovery" heading is retained as a source label. The first, middle, and final displayed frames report $N = 7592$, 7606, and 7265 and SYM-H values of −121, −122, and −124 nT. Much of the sampled region lies near the central color levels, while localized negative structure is evident in parts of the eastern sector, including around 85°–70°W and 35°–45°N. The central epoch therefore contains substantial spatial variation even though a broad regional inspection might appear comparatively weak. The source count decreases by 327 between the endpoints, and the small negative change in the SYM-H labels does not determine the regional residual pattern. These values describe the three inspected frames; they are not claimed as the extrema of all 31 frames.

Open accompanying GIF: `Movie_S6.gif`.



Movie S7: Regional dVTEC during 19:52:42–20:07:42 UT on 21 January 2026, centered at 20:00:12. Colors show detrended vertical total electron content, dVTEC, in TECU on geographic coordinates, with labeled ticks from −0.8 to +0.8 at intervals of 0.2 TECU and extensions at both ends. The displayed region spans approximately 130°–60°W and 20°–60°N. Each frame reports its time in UT, a SYM-H value, the source count N, a native 30-s cadence, and an elevation threshold of 30°. Unplotted locations are not assigned zero perturbation. The 31 displayed epochs span 15 min with nominal 30-s spacing; the GIF uses 120 ms per frame and a 3.72-s repeating playback cycle. Playback speed is a display setting, not a physical propagation speed. The original "Late recovery" heading is retained as a source label. At the first, middle, and last inspected frames, the annotated source count decreases from 9180 to 8976 and 8722, whereas the displayed SYM-H values are −89, −90, and −89 nT. Alternating positive and negative structure remains conspicuous in the western sector, with local values reaching the ±0.6–0.8 TECU labeled intervals. Thus, the three nearly equal index values coexist with spatially structured residual maps. The animation illustrates short-term changes within the selected recovery interval, but a 15-min display does not by itself establish persistence of an individual wave throughout the broader storm. Count differences, spatial sampling, and physical field evolution all enter the displayed sequence and must be separated in any subsequent quantitative tracking analysis.

Open accompanying GIF: `Movie_S7.gif`.

Scope of the technical supplement

This supplement gives the complete mathematical specification and a prospective evaluation protocol for multi-pattern propagation–evolution tomography (MP-PET). The declared grids, penalties, solver budgets, synthetic scenarios, and validation partitions are reproducible design choices. They are not evidence that the prescribed experiments, independent tests, or diagnostic calculations have already been performed. The case-study figures supplied with the manuscript contain no associated raw numerical arrays or archived end-to-end MP-PET run; their event morphology and the method's mathematical properties must therefore be assessed separately. Statements below about quantities to retain or report are requirements for an executable run, rather than claims of archived products available with this manuscript. All physical observations and event amplitudes remain empirical inputs or outputs.

The principal state is the summed signed electron-density residual. A unique regularized inner solution, a converged numerical calculation, conditional control sensitivity, and physical component identifiability are distinct concepts. The detailed notation follows the original mathematical derivation, with equation numbers local to this supplement. The one substantive synthetic-design correction is to supply raw data buffers on both sides of the valid inversion interval before applying a centered background filter.

Supplementary Note S7. Slant observations and the disturbance state

S7.1. Phase-continuous slant observations and the background convention

The measurement inputs are 30 s, phase-continuous dual-frequency slant-TEC arcs, observation times, satellite positions, receiver coordinates, elevations, and discontinuity flags. For carrier-phase ranges L_1, L_2 in metres at frequencies $f_1 > f_2$ in hertz, the first-order geometry-free observable can be written as

$$s_j^{\phi} = \frac{L_{1,j} - L_{2,j}}{40.3 \times 10^{16}(f_2^{-2} - f_1^{-2})} = s_j + c_a + \eta_j, \tag{S1}$$

where s_j is physical slant TEC in TECU, c_a is the ambiguity and stable instrumental offset on continuous arc a, and η_j collects measurement errors and unmodeled time-variable contributions. The factor 40.3 is the first-order cold-plasma coefficient in SI units. Frequencies are taken from the recorded signal identifiers and, for frequency-division signals, the recorded channel; they are measurement metadata rather than fitted constants. For example, GPS L1 and L2 use 1575.42 and 1227.60 MHz. A constant arc offset cancels under the background subtraction below, whereas a time-varying bias does not. This distinction is consistent with established TEC-processing practice and documented receiver-bias variability [2, 3]. No ambiguity is carried across a flagged cycle slip, loss of lock, signal-pair change, or receiver reset. A jump is not removed by fitting a density disturbance.

Let $\delta\tau = 30$ s denote the input interval, and let j index the regular samples of an accepted arc. The reference background and detrended observation are

$$b_j = \frac{1}{121} \sum_{p=-60}^{60} s_{j+p}^{\phi}, \qquad y_j = s_j^{\phi} - b_j. \tag{S2}$$

The first and last samples in this centered window are separated by 60 min. The window contains 121 samples, not 120, because both endpoints are included. All samples in a window must be present, valid, and belong to the same continuous arc. Missing observations are not interpolated, windows are not shortened at boundaries, and data from distinct arcs are not joined. An arc must span at least 120 min before detrending; its first and last 30 min cannot supply reference residuals. A gap exceeding 90 s starts a new arc, and even a shorter gap invalidates any filtering window containing the missing samples. A loss-of-lock or discontinuity flag starts a new arc irrespective of gap length. The 35° elevation cutoff is applied to each retained central observation; lower-elevation samples may contribute to its background only when they have passed the same continuity and measurement-validity checks. No residual-amplitude cutoff is used. The 30 and 90 min sensitivity backgrounds use the identical rule with 61 and 181 samples, respectively. Moving-average detrending has published precedent in TID analysis, but its response must be considered when interpreting disturbance amplitudes [4].

For an uninterrupted sinusoid of angular frequency ω, the interior discrete filter has response

$$H_D(\omega) = 1 - \frac{\sin(121\omega\delta\tau/2)}{121\sin(\omega\delta\tau/2)}, \qquad H_D(0) = 0. \tag{S3}$$

Equation (S3) follows by summing the symmetric geometric series in (S2). The reference gains are approximately 0.873 for a 25 min oscillation and 1.161 for a 50 min oscillation. Thus, the filter can attenuate or amplify a retained component; it is not an amplitude-preserving separation of background and disturbance. These values are filter properties, not measured event responses. Centering avoids a phase delay for interior sinusoidal samples, but endpoint rejection, gaps, arc-dependent backgrounds, and subsequent inversion still influence estimated phase. Dividing a reconstruction by H_D is not used to recover lost low-frequency information.

The distinction between a filtered line integral and the line integral of a physical perturbation can be stated explicitly. Let $\mathcal{A}_{30}$ map a full 30 s density sequence to its raw slant observations, let $\mathcal{D}$ apply (S2) only on complete valid windows, and let $\mathcal{R}$ select the corresponding central unfiltered samples. For an absolute background $\mathbf{N}_{\mathrm{bg}}$ and a physical injected perturbation $\boldsymbol{\delta}$,

$$\begin{aligned} \mathbf{y} &= \mathcal{R}\mathcal{A}_{30}\boldsymbol{\delta} + \mathbf{m}_D + \mathcal{D}\boldsymbol{\eta}, \\ \mathbf{m}_D &= \mathcal{D}\mathcal{A}_{30}\mathbf{N}_{\mathrm{bg}} + (\mathcal{D} - \mathcal{R})\mathcal{A}_{30}\boldsymbol{\delta}, \\ \mathrm{Cov}(\mathcal{D}\boldsymbol{\eta}) &= \mathcal{D}\boldsymbol{\Sigma}_\eta\mathcal{D}^{\mathsf{T}}. \end{aligned} \tag{S4}$$

Here, $\boldsymbol{\Sigma}_\eta$ is the raw-observation covariance. The linear maps act on observations from the same arcs, so their row dimensions agree. The discrepancy $\mathbf{m}_D$ contains residual background variability and the distortion of the injected perturbation. This equation does not interchange filtering and ray integration. For independent equal-variance raw errors, the variance of an individual detrended sample is $\sigma^2(1 - 1/121)$, but overlapping windows induce off-diagonal covariance. The practical inversion therefore estimates an observation-compatible residual density relative to the specified background convention, while end-to-end synthetic tests quantify its difference from an injected physical perturbation. It does not estimate the complete low-frequency storm-time departure in absolute electron density.

S7.2. Discrete slant geometry and robust observation fidelity

Let $t = 0, \ldots, T-1$ index reconstruction epochs and $i = 1, \ldots, m_t$ index the accepted receiver–satellite observations at epoch t. The input residual is

$$y_i(t) = \mathrm{sTEC}_i(t) - b_i(t) = 10^{-16} \int_{\mathcal{L}_i(t)\cap\Omega} \delta N_e(\mathbf{r}, t)\, \mathrm{d}\ell + \epsilon_i(t), \tag{S5}$$

where sTEC_i denotes the phase-continuous slant observable and b_i is the centered background in (S2), $\mathcal{L}_i(t)$ is the receiver–satellite ray, Ω is the reconstruction domain, $\mathbf{r}$ is spatial position, and ℓ is distance along the ray in metres. The electron-density disturbance δN_e has units of $\mathrm{el\,m^{-3}}$; y_i and ϵ_i are expressed in TEC units, with 1 TECU = 10^{16} $\mathrm{el\,m^{-2}}$.

The second equality in (S5) is a background-compatibility model, not an identity implied by temporal detrending. It assumes that the retained residuals can be represented approximately by a common three-dimensional disturbance field under the declared background convention. Mismatch between that representation and the supplied backgrounds, disturbances outside Ω, temporal matching error, and voxel discretization error are included in ϵ_i. A temporal filter acting on an arc cannot in general be interchanged with a time-dependent ray-integral operator. Consequently, the estimated state is an observation-compatible disturbance relative to the input background convention, not an estimate of every component of the absolute storm-time density change. Equation (S4) identifies preprocessing contributions to this discrepancy, consistently with the sensitivity of TID diagnostics to detrending [4].

The fixed computational domain extends from 140°W to 50°W, from 10°N to 65°N, and from 100 to 1000 km in WGS84 ellipsoidal height. Horizontal cells are $1^\circ \times 1^\circ$, and the 18 height layers have a thickness of 50 km. Electron density is represented by one constant value per voxel. Thus, $N = 90 \times 55 \times 18 = 89\,100$ unknowns describe each total-field epoch. The grid spacing is a discretization choice, not a claim about achieved spatial resolution. WGS84 uses semimajor axis $a_e = 6\,378\,137$ m and inverse flattening 298.257223563; the squared eccentricity is $e_e^2 = f(2-f)$, where f is flattening. A receiver-to-satellite straight line is intersected with the geodetic voxel surfaces, and the resulting segment lengths are evaluated in Earth-centered Cartesian metres.

With $\ell_{ij,t}$ denoting the ray length inside voxel j, the discrete observation model is

$$\mathbf{y}_t = \mathbf{H}_t\mathbf{u}_t + \boldsymbol{\epsilon}_t, \qquad [\mathbf{A}_t]_{ij} = 10^{-16}\ell_{ij,t}, \qquad \mathbf{H}_t = n_0\mathbf{A}_t, \qquad \mathbf{x}_t = n_0\mathbf{u}_t, \tag{S6}$$

where $\mathbf{y}_t \in \mathbb{R}^{m_t}$, $\mathbf{x}_t \in \mathbb{R}^N$ contains physical electron-density disturbances, $\mathbf{u}_t \in \mathbb{R}^N$ is dimensionless, and $n_0 = 10^{11}$ $\mathrm{el\,m^{-3}}$ is fixed for all epochs. Signed values are admissible; neither nonnegativity nor epoch-wise amplitude normalization is imposed.

Ray intersections are calculated against the actual voxel boundaries. Rays that leave the lateral domain between 100 and 1000 km are excluded. Contributions above and below the modeled shell are not assumed to have been measured as zero. The reference observation selection uses 30 s input sampling, a 35° elevation cutoff, a minimum continuous-arc duration of 120 min, and a new arc after a gap exceeding 90 s. Reconstruction epochs are separated by $\Delta t = 300$ s; at most one sample per arc is retained within 15 s of each epoch. The background in (S2) is applied once. No slant-to-vertical mapping, five-minute averaging, additional amplitude normalization, or selection by perturbation amplitude is introduced in the inversion. Five-minute sampling has a 10 min Nyquist period, but the primary structural interpretation is restricted to periods of 20 min and longer. High-pass detrending is not an anti-aliasing filter; the 8 min synthetic control explicitly tests contamination from faster fluctuations. A 150 s reconstruction-interval sensitivity run uses the same 120 min windows and physical control spacing, with all count normalizations recomputed.

For an arc a, let $z_j = y_j \sin e_j$, where j now indexes its consecutive 30 s samples and e_j is satellite elevation. The working noise scale is

$$\sigma_a = \max\left\{0.03\ \mathrm{TECU}, \frac{1.4826}{\sqrt{6}}\,\mathrm{MAD}(z_{j+1} - 2z_j + z_{j-1})\right\}, \qquad \sigma_i = \frac{\sigma_{a(i)}}{\sin e_i}. \tag{S7}$$

Here, $a(i)$ identifies the arc containing observation i, and $\mathrm{MAD}(z) = \mathrm{median}\,|z - \mathrm{median}(z)|$. At least 40 uninterrupted difference triplets are required; otherwise the arc is excluded rather than assigned an estimated noise value. The factor $\sqrt{6}$ follows from the variance of a second difference of independent equal-variance errors; departures from those assumptions make (S7) a working scale rather than an exact noise estimate. Elevation weighting does not alter the slant quantity in (S5).

Define $\widetilde{\mathbf{H}}_t = \mathrm{diag}(\sigma_i^{-1})\mathbf{H}_t$ and $\widetilde{\mathbf{y}}_t = \mathrm{diag}(\sigma_i^{-1})\mathbf{y}_t$. Observation fidelity uses the Huber loss [5],

$$\rho_c(r) = \begin{cases} \frac{1}{2}r^2, & |r| \le c, \\ c|r| - \frac{1}{2}c^2, & |r| > c, \end{cases} \qquad c = 1.345. \tag{S8}$$

The diagonal scaling is a working precision model, not the inverse of the full covariance in (S4). The floor of 0.03 TECU is a declared numerical lower bound rather than an asserted sensor accuracy. Correlation, time-variable biases, and systematic preprocessing error are to be assessed in held-out residuals and end-to-end controls. The robust loss limits the influence of individual large residuals without deleting observations because their inferred disturbance is large.

Supplementary Note S8. Fixed spatial representation

A climatological coordinate system is constructed from an offline IRI-2020 profile library [6]. Each profile is normalized by its Euclidean norm. With N_c profiles evaluated at the 18 voxel-centre heights, let

$$\mathbf{C} \in \mathbb{R}^{18\times N_c}, \qquad \mathbf{C} = \mathbf{D}_h\mathbf{\Sigma}\mathbf{V}^{\mathsf{T}}, \qquad \mathbf{D}_h^{\mathsf{T}}\mathbf{D}_h = \mathbf{I}_{18}. \tag{S9}$$

The columns of $\mathbf{D}_h$ are the left singular vectors, $\mathbf{\Sigma}$ contains singular values, and $\mathbf{V}$ contains right singular vectors. All 18 vertical vectors are retained. Their ordering follows decreasing singular value, and the largest-magnitude entry of each vector is assigned a positive sign. The actual basis matrix must be archived because repeated singular values can leave rotations undetermined by these conventions alone.

The reference library samples the fifteenth day of every month in 2020, 2022, and 2023, at two-hour UTC intervals, on a $5^\circ \times 5^\circ$ horizontal grid. The adopted options are URSI for the F2 critical frequency, AMTB for peak height, and NeQuick for the topside, with storm correction disabled. The horizontal library nodes include the domain endpoints, giving 19 longitude and 12 latitude values. Together with 36 dates and 12 UTC epochs per date, this gives $N_c = 98\,496$ profiles. The historical solar-index records required by IRI are input data, not fit parameters; their exact values and checksums must be retained with the model version and generated basis. Every requested model profile must be finite at all 18 heights. Invalid profiles stop library generation rather than silently changing its sampling. Neither event GNSS data nor validation profiles enter its construction. Heights supplied to IRI are converted consistently from the geodetic voxel locations to the height convention of the model interface. The complete basis does not impose rank truncation; nevertheless, an ℓ_1 penalty in this basis remains an informative, basis-dependent prior.

The state is ordered by horizontal column with height varying fastest. Let $\mathbf{P}$ permute column-major coefficients into separate horizontal mode maps, $\mathbf{E}$ mirror-extend each 90×55 map to 96×56 without repeating endpoints, and $\mathbf{Q}_H$ apply a three-level orthonormal two-dimensional Haar transform to each extended map. The spatial and increment analysis operators are

$$\mathbf{S} = \mathbf{Q}_H\mathbf{E}\mathbf{P}(\mathbf{I}_{N_{\mathrm{hor}}} \otimes \mathbf{D}_h^{\mathsf{T}}), \qquad \mathbf{S}_q = \mathbf{Q}_H\mathbf{E}\mathbf{P}, \tag{S10}$$

where $N_{\mathrm{hor}} = 90 \times 55$ and $\otimes$ denotes the Kronecker product. In $\mathbf{S}_q$, the maps are physical height layers, so the increment transformation does not mix heights. Both transforms have $P_s = P_q = 96 \times 56 \times 18 = 96\,768$ coefficients per epoch.

For each horizontal map indexed from zero, extension is performed on the high-index side: longitude indices 90–95 copy original indices 88–83, and latitude index 55 copies original index 53. The three Haar levels are then applied to the resulting 96×56 map. This rule fixes the extension origin and avoids repeated endpoints. Mirror extension makes the complete operators non-square. Their transposes therefore require accumulation at repeated boundary samples; cropping is not the adjoint. Low-frequency approximation coefficients and all detail coefficients are penalized. No vertical mode or horizontal scale is forced to vanish.

Supplementary Note S9. Multiple propagation fields and structured evolution

S9.1. Additive disturbance representation

For $K \in \{1, 2\}$, decompose the dimensionless field as

$$\mathbf{u}_t = \sum_{k=1}^{K} \mathbf{u}_t^{(k)}, \qquad \mathbf{y}_t = \mathbf{H}_t \sum_{k=1}^{K} \mathbf{u}_t^{(k)} + \boldsymbol{\epsilon}_t. \tag{S11}$$

The component index k denotes a signed pattern, not a prescribed physical driver. All components are estimated simultaneously from the same slant observations. Their supports may overlap, and their superposition may be constructive or destructive. Orthogonality, disjoint support, and a fixed propagation direction are not imposed.

The temporal model follows a kinematic relation at fixed height,

$$\partial_\tau u^{(k)} + \mathbf{v}_h^{(k)} \cdot \nabla_h u^{(k)} = s^{(k)}, \qquad q_t^{(k)}(\mathbf{r}) = \int_t^{t+\Delta t} s^{(k)}(\boldsymbol{\gamma}^{(k)}(\tau; \mathbf{r}), \tau)\, \mathrm{d}\tau. \tag{S12}$$

Here, τ is continuous time, ∇_h is the physical horizontal gradient, $\mathbf{v}_h^{(k)}$ is pattern velocity, and $\boldsymbol{\gamma}^{(k)}$ is a characteristic ending at $\mathbf{r}$ at $t + \Delta t$. The effective rate $s^{(k)}$ has units of s^{-1} for the dimensionless state. Its integrated value $q_t^{(k)}$ represents changes not explained by horizontal propagation. Integrating the material derivative along a characteristic yields the discrete model below.

Equation (S12) is not a closed electron-continuity equation. Its velocity is not identified with plasma drift, and its effective rate is not decomposed into ionization, recombination, or particle transport. It accommodates the geometric and intensity changes considered in motion-reconstruction and metamorphosis approaches [7, 8]. Earlier motion-aware ionospheric imaging provides a further precedent for incorporating propagation into reconstruction [9].

The kinematic interpretation follows by distinguishing a propagating pattern from the plasma carrying charge. For an absolute electron density N_e, a physical plasma velocity $\mathbf{V}_p$, and production and loss rates $\mathcal{P}_e, \mathcal{L}_e$ in $\mathrm{el\,m^{-3}\,s^{-1}}$, continuity has the form

$$\partial_\tau N_e + \nabla \cdot (N_e \mathbf{V}_p) = \mathcal{P}_e - \mathcal{L}_e. \tag{S13}$$

Writing a true physical perturbation as $N_e = N_{\mathrm{bg}} + n_0 u_{\mathrm{phys}}$ and introducing an arbitrary horizontal pattern velocity $\mathbf{v}_h$ gives, in local physical coordinates,

$$\partial_\tau u_{\mathrm{phys}} + \mathbf{v}_h \cdot \nabla_h u_{\mathrm{phys}} = \frac{\mathcal{P}_e - \mathcal{L}_e - \partial_\tau N_{\mathrm{bg}} - \nabla \cdot (N_{\mathrm{bg}} \mathbf{V}_p)}{n_0} - u_{\mathrm{phys}} \nabla \cdot \mathbf{V}_p + (\mathbf{v}_h - \mathbf{V}_{p,h}) \cdot \nabla_h u_{\mathrm{phys}} - V_{p,z} \partial_h u_{\mathrm{phys}}. \tag{S14}$$

Here, $\mathbf{V}_{p,h}$ and $V_{p,z}$ are the horizontal and vertical components of the plasma velocity. Every term on the right has units of s^{-1}. The expression shows why effective pattern evolution cannot be assigned to a unique physical source: production, loss, compressibility, vertical transport, background evolution, and pattern–plasma velocity differences contribute together. Continuity-based ionospheric models impose additional physical closure [10, 11]. MP-PET does not estimate those separate terms. Its filtered, observation-compatible state also contains the background-convention effects in (S4), and its additive pattern components are not separately conserved plasma populations. The kinematic source in (S12) is therefore a reconstruction variable with a specified physical scale, not a measured production or loss rate.

S9.2. Propagation parameterization and discrete transport

Using longitude λ, latitude φ, and height h, define

$$\mathbf{v}_h^{(k)}(\lambda, \varphi, h, \tau) = v_0 \sum_{a,b,c,d} B_a(\lambda) B_b(\varphi) L_c(h) L_d(\tau) \boldsymbol{\theta}_{abcd}^{(k)}, \qquad \|\boldsymbol{\theta}_{abcd}^{(k)}\|_2 \le 1.5. \tag{S15}$$

The two entries of each control vector are dimensionless eastward and northward velocities. The scale is $v_0 = 1000\ \mathrm{m\,s^{-1}}$. The functions B_a, B_b are open-clamped cubic B-splines with 5° knot spacing; L_c is piecewise linear at 150, 275, 450, and 775 km, with constant extension to the vertical boundaries; and L_d is piecewise linear on 30 min temporal control intervals. All basis functions are nonnegative and form a partition of unity. Thus the velocity magnitude is bounded by $v_{\max} = 1500\ \mathrm{m\,s^{-1}}$ throughout the domain. For the reference 120 min window, the open-clamped cubic mesh contains 21 longitude and 14 latitude controls, giving $N_\theta = 21 \times 14 \times 4 \times 5 \times 2 = 11\,760$ scalar controls per component. The time knots include 0, 30, 60, 90, and 120 min relative to the window start. A zero velocity is admissible. The upper bound is a conservative search constraint, not an assumed event speed or a universal upper limit for ionospheric waves. Published event velocities inform plausibility assessment but are not imposed as reconstruction targets [12].

Characteristics are backtracked in angular coordinates according to

$$\frac{\mathrm{d}\lambda}{\mathrm{d}\tau} = \frac{v_E^{(k)}}{[\mathcal{N}(\varphi) + h] \cos\varphi}, \qquad \frac{\mathrm{d}\varphi}{\mathrm{d}\tau} = \frac{v_N^{(k)}}{\mathcal{M}(\varphi) + h}, \qquad \frac{\mathrm{d}h}{\mathrm{d}\tau} = 0, \tag{S16}$$

where $v_E^{(k)}, v_N^{(k)}$ are physical velocity components and $\mathcal{N}, \mathcal{M}$ are the WGS84 prime-vertical and meridional curvature radii. Angles are in radians during integration. Five explicit-midpoint steps of -60 s cover one reconstruction interval. The height-dependent velocity varies between layers, but the propagation operator does not shift density vertically or align interlayer phases.

Let $(\lambda_j^b, \varphi_j^b)$ be the backtracked position for target voxel j. At the same height, define

$$[\mathbf{W}_t^{(k)} \mathbf{u}]_j = \sum_{p,q} \kappa\left(\frac{\lambda_j^b - \lambda_p}{\Delta\lambda}\right) \kappa\left(\frac{\varphi_j^b - \varphi_q}{\Delta\varphi}\right) u_{pq,h_j}, \tag{S17}$$

where $\Delta\lambda, \Delta\varphi$ are angular grid intervals, (λ_p, φ_q) are grid-centre coordinates, and κ is the Keys cubic-convolution kernel with parameter $-1/2$ [13]. Only the surrounding 4×4 centres contribute. Section S16 gives the kernel and the differentiated characteristic integration. For fixed controls,

$\mathbf{W}_t^{(k)}$ is linear in the state. The warp is a centre-value approximation for the voxel state, not a conservative finite-volume remapping. Its transpose scatters the same interpolation weights and is not interpolation using a reversed velocity.

Let $\mathbf{M} \in \mathbb{R}^{N_M \times N}$ select a fixed interior region whose distance to each lateral boundary exceeds $v_{\max}\Delta t + 2d_{\max}$, where $d_{\max}$ is the largest adjacent-centre horizontal distance at the relevant height. The mask is fixed before optimization and checked against all computed characteristic stencils. Propagation is imposed only through

$$\mathbf{r}_t^{(k)} = \mathbf{M}\left[\mathbf{u}_{t+1}^{(k)} - \mathbf{W}_t^{(k)}(\boldsymbol{\theta}^{(k)})\mathbf{u}_t^{(k)} - \mathbf{q}_t^{(k)}\right]. \tag{S18}$$

The full-grid increments $\mathbf{q}_t^{(k)}$ and residuals $\mathbf{r}_t^{(k)}$ are dimensionless. Their physical increment scale is $n_0\mathbf{q}_t^{(k)}$, in $\mathrm{el\,m^{-3}}$. No periodic wraparound is used. States outside the propagation mask remain in the observation model, but their increments are not interpreted physically.

S9.3. Increment groups

For each component, interval, height, wavelet scale, and horizontal location, the horizontal, vertical, and diagonal detail coefficients of $\mathbf{S}_q\mathbf{q}_t^{(k)}$ form a group g. Each coarsest approximation coefficient forms its own singleton group. Write $\mathcal{G}$ for this partition and $|g|$ for group size. Group sparsity acts within a height layer; it does not require simultaneous activation at different heights or times. Low-frequency coefficients represent broad changes, while detail groups represent localized changes. An increment is estimated jointly with the states rather than set equal to the entire propagation residual.

Supplementary Note S10. Duplication-controlled joint objective

Let $\mathbf{U}^{(k)} = [\mathbf{u}_0^{(k)}, \ldots, \mathbf{u}_{T-1}^{(k)}]$, $\mathbf{Q}^{(k)} = [\mathbf{q}_0^{(k)}, \ldots, \mathbf{q}_{T-2}^{(k)}]$, and let $\mathbf{z} \in \mathbb{R}^{KN(2T-1)}$ stack all component states and increments. The vector $\boldsymbol{\Theta}$ stacks their propagation controls. We minimize

$$\begin{aligned}\mathcal{J}_K(\mathbf{z}, \boldsymbol{\Theta}) = \frac{1}{N_y}\sum_{t,i}\rho_c\left(\left[\widetilde{\mathbf{H}}_t\sum_k \mathbf{u}_t^{(k)} - \widetilde{\mathbf{y}}_t\right]_i\right) \\ + \frac{\lambda_s}{TP_s}\sum_{k,t}\|\mathbf{S}\mathbf{u}_t^{(k)}\|_1 + \frac{K\beta}{2D_d}\sum_{k,t}\|\mathbf{r}_t^{(k)}\|_2^2 \\ + \frac{\lambda_q}{N_q}\sum_{k,t,g}\sqrt{|g|}\,\|[\mathbf{S}_q\mathbf{q}_t^{(k)}]_g\|_2 \\ + \sum_k \mathcal{R}_v(\boldsymbol{\theta}^{(k)}) + \frac{K\varepsilon}{2D_z}\|\mathbf{z}\|_2^2, \qquad \boldsymbol{\Theta} \in \mathcal{C}_K,\end{aligned} \tag{S19}$$

with

$$N_y = \sum_t m_t, \qquad D_d = (T-1)N_M, \qquad N_q = (T-1)P_q, \qquad D_z = N(2T-1). \tag{S20}$$

The set $\mathcal{C}_K$ is the product of the control-vector balls in (S15). All regularization coefficients are positive and dimensionless. The Frobenius norm of each state or increment matrix is equivalent to the Euclidean norm of its vectorization.

Propagation regularization is

$$\mathcal{R}_v(\boldsymbol{\theta}) = \sum_{\nu\in\{h,t,z\}} \frac{\lambda_{v\nu}}{2|\mathcal{E}_\nu|} \sum_{(a,b)\in\mathcal{E}_\nu} \frac{\|\boldsymbol{\theta}_a - \boldsymbol{\theta}_b\|_2^2}{d_{ab,\nu}^2} + \frac{\lambda_{v0}}{2N_\theta}\|\boldsymbol{\theta}\|_2^2. \tag{S21}$$

Here, $\mathcal{E}_h, \mathcal{E}_t, \mathcal{E}_z$ contain adjacent control-point pairs in the horizontal, temporal, and height directions. Their normalized separations are $d_{ab,h} = d_{ab}/(500\text{ km})$, $d_{ab,t} = |\tau_a - \tau_b|/(1800\text{ s})$, and $d_{ab,z} = |h_a - h_b|/(100\text{ km})$, respectively. The physical horizontal distance d_{ab} is geodesic distance between the B-spline Greville control locations. The scalar-control count N_θ is for one component. These penalties constrain velocity variation without averaging the electron-density states between heights.

S10.1. Why the quadratic terms contain a factor of K

A conventional sum of component-wise squared residuals has an undesirable property: replacing one component by K identical copies with amplitudes divided by K divides its quadratic penalty by K, although the observable total field is unchanged. Equation (S19) removes this particular advantage by multiplying the component-wise quadratic terms by K. The one-homogeneous spatial and increment penalties are summed without division by K.

Proposition 1: identical-propagation duplication cannot improve the objective. Suppose all components have the same propagation controls $\boldsymbol{\theta}$. Define $\mathbf{U} = \sum_k \mathbf{U}^{(k)}$ and $\mathbf{Q} = \sum_k \mathbf{Q}^{(k)}$. Then, for identical hyperparameters,

$$\mathcal{J}_K(\mathbf{z}, \boldsymbol{\theta}, \ldots, \boldsymbol{\theta}) \geq \mathcal{J}_1(\mathbf{U}, \mathbf{Q}, \boldsymbol{\theta}) + (K-1)\mathcal{R}_v(\boldsymbol{\theta}). \tag{S22}$$

Proof. The data terms are equal. Linearity gives $\sum_k \mathbf{r}_t^{(k)} = \mathbf{M}[\mathbf{u}_{t+1} - \mathbf{W}_t\mathbf{u}_t - \mathbf{q}_t]$. For arbitrary vectors $\mathbf{a}_k$, Cauchy–Schwarz yields $\|\sum_k \mathbf{a}_k\|_2^2 \leq K\sum_k \|\mathbf{a}_k\|_2^2$; apply it to the propagation residuals and to stacked states and increments. The triangle inequality applies to each spatial ℓ_1 norm and increment-group norm. Finally, the propagation prior is repeated K times. These inequalities establish (S22). For equal copies, the state-dependent terms attain equality. □

This result removes one artificial mechanism for favoring additional components; it does not prove component identifiability. Because of the correction, appending a zero component is not a penalty-preserving embedding. Model counts are therefore compared through held-out predictions, not through a likelihood-ratio interpretation of their regularized objective values. The spatial component penalty also discourages unsupported cancellation by satisfying $\sum_k \|\mathbf{S}\mathbf{u}_t^{(k)}\|_1 \geq \|\mathbf{S}\sum_k \mathbf{u}_t^{(k)}\|_1$, without prohibiting physical interference.

S10.2. Conditional well-posedness

Proposition 2: conditional uniqueness and existence. For fixed K and $\boldsymbol{\Theta}$, (S19) has a unique minimizer in $\mathbf{z}$. A global minimizer of the joint finite-dimensional problem exists on $\mathcal{C}_K$, although the numerical method need not find it.

Proof. The observation, spatial, group, and propagation terms are convex in $\mathbf{z}$ at fixed controls. The final term is μ_K-strongly convex, with $\mu_K = K\varepsilon/D_z > 0$, and the objective is coercive. Thus the conditional minimizer exists and is unique. The control set is compact, and the discretized propagation operator is continuous. Uniform coercivity and continuity imply continuity of the minimized value with respect to the controls, so its minimum is attained on the compact set. Nonlinear dependence on the controls leaves the joint problem nonconvex. □

Supplementary Note S11. Profiled propagation estimation

Define the conditional solution and reduced objective by

$$\mathbf{z}_*(\boldsymbol{\Theta}) = \arg\min_{\mathbf{z}} \mathcal{J}_K(\mathbf{z}, \boldsymbol{\Theta}), \qquad \Phi_K(\boldsymbol{\Theta}) = \mathcal{J}_K(\mathbf{z}_*(\boldsymbol{\Theta}), \boldsymbol{\Theta}). \tag{S23}$$

Every trial propagation field is evaluated after the density fields and increments have been reoptimized, to the reported inner accuracy. This is a variable-projection formulation with nonsmooth inner regularization; its general mathematical foundation is established by van Leeuwen and Aravkin [14].

S11.1. Reduced gradient and curvature

Write $a_d = K\beta/D_d$, and define the linear map $\mathbf{G}_t^{(k)}\mathbf{d} = \mathbf{M}\{[\partial_{\theta^{(k)}}\mathbf{W}_t^{(k)}\,\mathbf{d}]\mathbf{u}_t^{(k)}\}$, where $\mathbf{d}$ is a perturbation of the component controls. At the unique inner optimum,

$$\nabla_{\theta^{(k)}}\Phi_K = -a_d \sum_t [\mathbf{G}_t^{(k)}]^{\mathsf{T}}\mathbf{r}_t^{(k)} + \nabla\mathcal{R}_v(\boldsymbol{\theta}^{(k)}). \tag{S24}$$

Proposition 3: validity of the reduced gradient. On a region where the discretized warp is continuously differentiable in its controls, Φ_K is continuously differentiable and (S24) holds.

Proof. Uniform strong convexity gives a continuous unique inner minimizer. The nonsmooth terms are independent of the propagation controls, and the remaining objective has a continuous partial control derivative. Applying the minimum-value envelope theorem gives $\nabla\Phi_K = \partial_\Theta\mathcal{J}_K(\mathbf{z}_*, \Theta)$. Differentiation of each squared propagation residual contributes $-a_d\mathbf{G}_t^{\mathsf{T}}\mathbf{r}_t$, giving (S24). No differentiation of an arbitrarily selected sparse solution is required. □

Thus the reduced gradient equals the frozen-state partial gradient *at an accurately solved inner optimum*. The distinction from block-coordinate updates is the inner reoptimization at every trial and the resulting reduced curvature, not an additional chain-rule term in the first derivative.

For second derivatives, assume that zero analysis coefficients and zero increment groups have a locally stable active pattern, nonzero groups remain nonzero, Huber residuals avoid $\pm c$, and the controls lie in a smooth interior region. Let $\mathbf{R}$ have orthonormal columns spanning the tangent space of this fixed active face, and write $\mathbf{z} = \mathbf{R}\boldsymbol{\xi}$ locally. Differentiating the inner optimality equation with respect to $\boldsymbol{\Theta}$ gives

$$\frac{\partial\boldsymbol{\xi}_*}{\partial\boldsymbol{\Theta}} = -\mathbf{H}_{\xi\xi}^{-1}\mathbf{H}_{\xi\Theta}, \qquad \nabla^2\Phi_K = \mathbf{H}_{\Theta\Theta} - \mathbf{H}_{\Theta\xi}\mathbf{H}_{\xi\xi}^{-1}\mathbf{H}_{\xi\Theta}. \tag{S25}$$

The blocks are derivatives of $\mathcal{J}_K(\mathbf{R}\xi, \Theta)$ on the active face. The subtraction represents the relaxation obtained when competing state and increment changes are allowed. It is a regularized Schur complement, not a data-only Fisher information matrix. The exact reduced Hessian may be indefinite away from a local minimum. At active-set changes, interpolation second-derivative breakpoints, or active velocity bounds, the unrestricted classical Hessian is not used. Section S17 gives the restricted Hessian and a positive-semidefinite Gauss–Newton approximation.

S11.2. Inner primal–dual solution

Let $\mathcal{L}_y\mathbf{z} = \{\widetilde{\mathbf{H}}_t\sum_k \mathbf{u}_t^{(k)}\}_t$, $\mathcal{L}_s\mathbf{z} = \{\mathbf{S}\mathbf{u}_t^{(k)}\}_{k,t}$, $\mathcal{L}_d(\Theta)\mathbf{z} = \{\mathbf{r}_t^{(k)}\}_{k,t}$, and $\mathcal{L}_q\mathbf{z} = \{\mathbf{S}_q\mathbf{q}_t^{(k)}\}_{k,t}$. Define

$$a_y = N_y^{-1}, \quad a_s = \frac{\lambda_s}{TP_s}, \quad a_q = \frac{\lambda_q}{N_q}, \quad \mu_K = \frac{K\varepsilon}{D_z}, \qquad \mathcal{K}_\Theta = \begin{bmatrix} \sqrt{a_y}\mathcal{L}_y \\ \sqrt{a_s}\mathcal{L}_s \\ \sqrt{a_d}\mathcal{L}_d(\Theta) \\ \sqrt{a_q}\mathcal{L}_q \end{bmatrix}. \tag{S26}$$

For $\mathbf{s} = (\mathbf{s}_y, \mathbf{s}_s, \mathbf{s}_d, \mathbf{s}_q)$, the conditional objective is $F(\mathcal{K}_\Theta \mathbf{z}) + G(\mathbf{z}) + \sum_k \mathcal{R}_v$, where

$$\begin{aligned} F(\mathbf{s}) &= \sum_i \rho_{c\sqrt{a_y}}(s_{y,i} - d_i) + \sqrt{a_s}\|\mathbf{s}_s\|_1 + \tfrac{1}{2}\|\mathbf{s}_d\|_2^2 + \sqrt{a_q}\sum_g \sqrt{|g|}\|\mathbf{s}_{q,g}\|_2, \\ \mathbf{d} &= \sqrt{a_y}\,\widetilde{\mathbf{y}}, \qquad G(\mathbf{z}) = \tfrac{1}{2}\mu_K\|\mathbf{z}\|_2^2. \end{aligned} \tag{S27}$$

The identity $\rho_{c\sqrt{a}}(\sqrt{a}\,r) = a\rho_c(r)$ ensures that (S27) preserves every normalization in (S19).

Applying the primal–dual algorithm of Chambolle and Pock [15] to this particular splitting gives

$$\begin{aligned} \mathbf{p}^{n+1} &= \operatorname{prox}_{\sigma_p F^*}(\mathbf{p}^n + \sigma_p \mathcal{K}_\Theta \bar{\mathbf{z}}^{\,n}), \\ \mathbf{z}^{n+1} &= \frac{\mathbf{z}^n - \tau_p \mathcal{K}_\Theta^{\mathsf{T}} \mathbf{p}^{n+1}}{1 + \tau_p \mu_K}, \\ \bar{\mathbf{z}}^{\,n+1} &= 2\mathbf{z}^{n+1} - \mathbf{z}^n. \end{aligned} \tag{S28}$$

Here, n is the inner iteration index, $\mathbf{p}$ contains the four dual blocks, F^* is the Fenchel conjugate, and $\operatorname{prox}_{\gamma f}(\mathbf{v})$ minimizes $f(\mathbf{p}) + \|\mathbf{p} - \mathbf{v}\|_2^2/(2\gamma)$. The extrapolated primal iterate is $\bar{\mathbf{z}}^{\,n} = 2\mathbf{z}^n - \mathbf{z}^{n-1}$ for $n \geq 1$, with $\bar{\mathbf{z}}^{\,0} = \mathbf{z}^0$. This extrapolation symbol is distinct from the rescaled numerical variable $\mathbf{z}_c$ used below. The positive step sizes satisfy $\tau_p \sigma_p \|\mathcal{K}_\Theta\|_2^2 < 1$.

For the data block, conjugating the shifted Huber term yields $F_y^*(\mathbf{p}) = \frac{1}{2}\|\mathbf{p}\|_2^2 + \mathbf{d}^{\mathsf{T}}\mathbf{p}$ on $|p_i| \leq c\sqrt{a_y}$, and infinity otherwise. Hence, for its incoming argument $\mathbf{v}_y$,

$$[\operatorname{prox}_{\sigma_p F_y^*}(\mathbf{v}_y)]_i = \operatorname{clip}\left(\frac{v_{y,i} - \sigma_p d_i}{1 + \sigma_p}, -c\sqrt{a_y}, c\sqrt{a_y}\right). \tag{S29}$$

For the remaining blocks,

$$\begin{aligned} \operatorname{prox}_{\sigma_p F_s^*}(\mathbf{v}_s) &= \operatorname{clip}(\mathbf{v}_s, -\sqrt{a_s}, \sqrt{a_s}), \\ \operatorname{prox}_{\sigma_p F_d^*}(\mathbf{v}_d) &= \frac{\mathbf{v}_d}{1 + \sigma_p}, \\ [\operatorname{prox}_{\sigma_p F_q^*}(\mathbf{v}_q)]_g &= \frac{\mathbf{v}_{q,g}}{\max\{1, \|\mathbf{v}_{q,g}\|_2 / \sqrt{a_q|g|}\}}. \end{aligned} \tag{S30}$$

These updates follow from the conjugates of the ℓ_1, quadratic, and group norms. No thresholding formula for an orthogonal transform is incorrectly applied to the non-square mirror-extended analysis operators.

At dual feasibility, the primal and dual objective values are

$$\begin{aligned} P &= F(\mathcal{K}_\Theta \mathbf{z}) + \tfrac{1}{2}\mu_K\|\mathbf{z}\|_2^2 + \sum_k \mathcal{R}_v, \\ D &= -\tfrac{1}{2}\|\mathbf{p}_y\|_2^2 - \mathbf{d}^{\mathsf{T}}\mathbf{p}_y - \tfrac{1}{2}\|\mathbf{p}_d\|_2^2 - \frac{\|\mathcal{K}_\Theta^{\mathsf{T}}\mathbf{p}\|_2^2}{2\mu_K} + \sum_k \mathcal{R}_v. \end{aligned} \tag{S31}$$

They satisfy $D \leq \Phi_K(\Theta) \leq P$. The absolute gap $\delta = P - D$ provides

$$\|\mathbf{z}-\mathbf{z}_*\|_2 \le \sqrt{\frac{2\delta}{\mu_K}}, \qquad \|\widetilde{\mathbf{g}}-\nabla\Phi_K\|_2 \le L_{\Theta z}\sqrt{\frac{2\delta}{\mu_K}}, \tag{S32}$$

where $\widetilde{\mathbf{g}} = \partial_\Theta \mathcal{J}_K(\mathbf{z},\Theta)$, and $L_{\Theta z}$ bounds the variation of this partial derivative with the inner state on the relevant level set. A computable conservative bound is given in Section S17. The small stabilization coefficient can make (S32) conservative; a small relative objective gap alone is not presented as a tight component-error or gradient certificate.

S11.3. Safeguarded outer iteration and finite-accuracy decisions

At outer iteration r, the conditional state problem is solved before a control trial is formed. With $\widetilde{\mathbf{g}}^r$ denoting the partial gradient evaluated at that numerical solution, the trial and the available certified acceptance test are

$$\mathbf{\Theta}^+ = \Pi_{\mathcal{C}_K}(\mathbf{\Theta}^r - \alpha\widetilde{\mathbf{g}}^r), \qquad P^+ \le D^r + \nu\,(\widetilde{\mathbf{g}}^r)^{\mathsf{T}}(\mathbf{\Theta}^+ - \mathbf{\Theta}^r), \qquad \nu = 10^{-4}. \tag{S33}$$

The projection rescales each two-dimensional control vector to radius 1.5 only when its norm exceeds that radius. The initial step is $\alpha_0 = 0.1/(\|\widetilde{\mathbf{g}}^r\|_\infty + 10^{-12})$; backtracking multiplies it by 0.5, with at most 20 trials. The state and increment variables are reoptimized at every trial rather than held fixed. If (S33) holds with a negative directional product, the upper and lower bounds imply an actual decrease in the exact profiled objective. This conclusion does not require an exact inner solution.

The global bound in (S32) can be too conservative for a practical stopping rule. For the reference window and $K = 2$, $D_z = 4\,365\,900$ and $\mu_K \simeq 4.58 \times 10^{-13}$. A small mean objective gap must therefore not be converted into an unsupported claim of small componentwise error. Computation uses the exactly equivalent variable $\mathbf{z}_c = \mathbf{z}/\sqrt{D_z}$, the operator $\mathcal{K}_c = \sqrt{D_z}\mathcal{K}_\Theta$, and the ridge coefficient $\mu_{K,c} = K\varepsilon$. This rescaling expresses state errors on a root-mean-square scale; it neither strengthens the observations nor removes ill-conditioning. Primal and dual values are always evaluated for the same original objective.

A finite-budget implementation also retains a practical, explicitly noncertifying acceptance rule. After the current and trial inner solves satisfy their fixed-point residual criteria, a trial must satisfy $P^+ \le P^r + \nu\widetilde{\mathbf{g}}^{\mathsf{T}}(\Theta^+ - \Theta^r)$. Both solves are then refined with tolerances ten times smaller. The refined trial must again decrease the computed objective, and the combined change of the two objective values under refinement must be no more than one quarter of the refined decrease. Otherwise the trial is refined further, backtracked, or rejected. A step meeting this numerical test but not (S33) is classified as a numerically resolved objective decrease, not a certified decrease of the exact profile. The classification is retained in the output.

For exact conditional solves, define the projected-gradient mapping $\mathcal{G}_\alpha(\Theta) = \alpha^{-1}[\Theta - \Pi_{\mathcal{C}_K}(\Theta - \alpha\nabla\Phi_K)]$. On a region with an L_Φ-Lipschitz reduced gradient, the projection inequality and descent lemma give $\Phi_K(\Theta^+) \le \Phi_K(\Theta) - \alpha(1-\alpha L_\Phi/2)\|\mathcal{G}_\alpha(\Theta)\|_2^2$ for $\alpha < 2/L_\Phi$. Bounded admissible step sizes and sufficiently accurate inner solutions support stationary accumulation points, not global optimality. Finite numerical stability, a certified first-order condition, and the existence of a global minimizer are reported as distinct statements.

Supplementary Note S12. Identifiability after accounting for competing explanations

S12.1. Exact propagation–increment ambiguity

The total observation operator contains no propagation parameter when the states are free. Even with a known state sequence, changing the propagation field from θ to θ' can be compensated by

$$\mathbf{q}_t'^{(k)} = \mathbf{q}_t^{(k)} + [\mathbf{W}_t^{(k)}(\theta) - \mathbf{W}_t^{(k)}(\theta')]\mathbf{u}_t^{(k)}. \tag{S34}$$

This leaves both the component sequence and its propagation residual unchanged. Therefore, unrestricted increments preclude structural identification of propagation from these equations. The increment prior selects among competing explanations but does not convert that selection into a direct measurement of a physical velocity.

A second ambiguity is component exchange: $\mathbf{u}_t^{(1)} + \mathbf{a}_t$ and $\mathbf{u}_t^{(2)} - \mathbf{a}_t$ have the same total for any $\mathbf{a}_t$. Labels can also be permuted. Component-wise interpretation consequently requires more than a unique inner minimizer or a low observation residual.

S12.2. Conditional observation sensitivity

For a prescribed propagation field and zero increments in an ideal fully propagated domain, let $\mathbf{P}_0^{(k)} = \mathbf{I}$ and $\mathbf{P}_t^{(k)} = \mathbf{W}_{t-1}^{(k)} \cdots \mathbf{W}_0^{(k)}$. Initial component states are then observed through

$$\mathcal{O}_k = \begin{bmatrix} \widetilde{\mathbf{H}}_0\mathbf{P}_0^{(k)} \\ \widetilde{\mathbf{H}}_1\mathbf{P}_1^{(k)} \\ \vdots \\ \widetilde{\mathbf{H}}_{T-1}\mathbf{P}_{T-1}^{(k)} \end{bmatrix}, \qquad \widetilde{\mathbf{y}} = [\mathcal{O}_1 \ \cdots \ \mathcal{O}_K] \begin{bmatrix} \mathbf{u}_0^{(1)} \\ \vdots \\ \mathbf{u}_0^{(K)} \end{bmatrix} + \widetilde{\boldsymbol{\epsilon}}. \tag{S35}$$

Similar columns of the component observability operators imply poorly separated initial states. In the actual masked model, boundary states, allowed increments, and any admitted changes in evolution mismatch must also be included as nuisance variables; they are not fixed to zero to create artificial observability.

For a specified local admissible disturbance-and-increment family, write the predicted, scaled observations as $\mathbf{f}(\Theta, \boldsymbol{\xi})$, where $\boldsymbol{\xi}$ includes initial fields, permitted increment coefficients, boundary values, and admitted evolution-mismatch variables. The fitted-support family and its prescribed expansions are defined below. At a locally feasible differentiable parameterization, define $\mathbf{B}_\Theta = \partial_\Theta \mathbf{f}$ and $\mathbf{B}_\xi = \partial_\xi \mathbf{f}$. Eliminating a first-order nuisance perturbation gives

$$\begin{aligned} \min_{\delta\xi} \|\mathbf{B}_\Theta \delta\Theta + \mathbf{B}_\xi \delta\xi\|_2^2 &= \delta\Theta^\mathsf{T} \mathbf{I}_{\text{cond}} \delta\Theta, \\ \mathbf{I}_{\text{cond}} &= \mathbf{B}_\Theta^\mathsf{T}(\mathbf{I} - \mathbf{B}_\xi \mathbf{B}_\xi^\dagger)\mathbf{B}_\Theta, \end{aligned} \tag{S36}$$

where † denotes the Moore–Penrose pseudoinverse. The projector removes observation changes that the nuisance variables can reproduce. This is observation sensitivity *conditional on the declared dynamical and support restrictions*. With working rather than full covariance scaling, it is not assigned a likelihood-based confidence interpretation.

Proposition 4: nuisance expansion cannot increase conditional information. If the range of the nuisance Jacobian is enlarged while the propagation perturbation is unchanged,

every directional value of (S36) can only decrease. If all increments or evolution mismatches are unrestricted, (S34) provides an exact compensating perturbation and $\mathbf{I}_{\mathrm{cond}} = \mathbf{0}$.

Proof. Minimizing a squared norm over a larger nuisance space cannot increase its minimum. Under unrestricted increments, differentiation of (S34) cancels any change of the propagation operator while keeping the observed states fixed, so that minimum is zero for every control direction. □

The reference conditional family holds the fitted propagation residual fixed to first order, $\delta\mathbf{r}_t^{(k)} = 0$, while allowing the initial component states and all boundary-state perturbations to vary. Increment coefficients are permitted to vary on the fitted nonzero or numerically ambiguous groups. A group is treated as stably zero only when its coefficient norm is at most $10^{-6}\max\{1, \|\mathbf{S}_q\mathbf{q}\|_2/\sqrt{P_q}\}$ and its dual variable is at least 10% inside the associated dual ball. Spatial analysis rows are treated as stably zero under the analogous coefficient threshold and 10% dual-margin test. Zero-row restrictions are imposed only when the entire row support lies inside the propagation mask; rows intersecting the boundary remain free. All other coefficients are admitted as nuisance variables. These thresholds define a reproducible conditional diagnostic, not physically exact support recovery.

The first support expansion admits increment groups within one neighbouring horizontal coefficient location, one adjacent scale, one adjacent physical height layer, and one adjacent time interval of any admitted group, taking their union. The same expansion releases spatial zero rows whose supports intersect the enlarged increment supports. The second expansion releases all increment coefficients, with the initial and boundary freedoms unchanged. A final check also releases all propagation-residual changes. The unrestricted-increment and unrestricted-residual limits must return zero propagation sensitivity up to the declared numerical tolerance. Thus, a nonzero result is always accompanied by the restrictions that make it possible.

Directional sensitivities are evaluated for spatially uniform eastward and northward velocity perturbations on each of the four height controls and five time controls of a 120 min window. Each named direction changes its control-plane velocity by 100 $\mathrm{m\,s^{-1}}$; its interpolation in height and time remains that of (S15). This produces 40 named directions per component and tests a reproducible low-dimensional subset, not the complete velocity-control space. At an active speed bound, only feasible one-sided perturbations are examined. Constrained least-squares projection uses relative residual tolerance 10^{-8} and at most 2000 iterations. Singular values below 10^{-8} times the largest are classified as numerically unresolved; this threshold does not establish noise-limited recoverability. Failure to resolve the projection is retained as a diagnostic failure rather than replaced by prior curvature.

Accordingly, diagnostics are evaluated for the fitted stable increment support, its prescribed expansions, and the unrestricted-increment limit. A conditional constraint disappearing under modest nuisance expansion is reported as model-dependent. The regularized profile curvature in (S25) and conditional observation sensitivity in (S36) are reported separately; subtracting a velocity-prior Hessian from a regularized Schur complement is not sufficient to obtain data-only information.

On an explicitly chosen control subspace $\mathcal{V}$, suppose $\delta\Theta^{\mathsf{T}}\mathbf{I}_{\mathrm{cond}}\delta\Theta \geq s_*^2\|\delta\Theta\|_2^2$ with $s_* > 0$. First-order observational errors of norm at most δ_y then permit a least-squares control perturbation bounded by δ_y/s_* on that subspace. This statement is conditional and local; it says nothing about directions in the nullspace. In practice, unit perturbations are defined in physical velocity units on named control patterns, and numerical singular-value rank is distinguished from noise-limited practical recoverability.

S12.3. A two-pattern separation limit

An analytical limit can be derived without appealing to the regularized Hessian. Consider a fully sampled horizontal Fourier mode with a known nonzero wavevector $\mathbf{k}$, constant component velocities $\mathbf{v}_1, \mathbf{v}_2$, and no evolution increments. Its temporal samples satisfy

$$b_n = a_1 e^{-\mathrm{i}\Omega_1 n} + a_2 e^{-\mathrm{i}\Omega_2 n} + e_n, \qquad \Omega_k = (\mathbf{k}\cdot\mathbf{v}_k)\Delta t, \qquad n = 0, \dots, T-1. \tag{S37}$$

Here, a_k is a complex Fourier coefficient, b_n is the observed coefficient, e_n is error, and Ω_k is the angular phase advance per sample. The two-column matrix $\mathbf{V}_T$ has entries $[\mathbf{V}_T]_{nk} = e^{-\mathrm{i}\Omega_k n}$. For $\Delta\Omega = \Omega_2 - \Omega_1$, its Gram eigenvalues are

$$\lambda_\pm(\mathbf{V}_T^*\mathbf{V}_T) = T \pm \left|\frac{\sin(T\Delta\Omega/2)}{\sin(\Delta\Omega/2)}\right|, \tag{S38}$$

with the ratio interpreted by continuity at aliased coincidences. The star denotes conjugate transpose. Equation (S38) follows by summing the geometric series of the off-diagonal Gram entry. At $\Delta\Omega = 0 \pmod{2\pi}$, the two columns coincide. For $T \geq 2$ and nonaliased, prescribed component frequencies, the two complex amplitudes are identifiable in this ideal model, with error amplification bounded by $1/\sqrt{\lambda_-}$. This conditional amplitude result does not imply that two epochs identify two unknown component frequencies or full velocity vectors. For small frequency separation,

$$\lambda_- = \frac{T(T^2-1)}{24}(\Delta\Omega)^2 + O((\Delta\Omega)^4). \tag{S39}$$

Thus separation deteriorates quadratically as the projected velocities converge. A small nonzero velocity difference is not sufficient for reliable decomposition within a short window. A single plane wave constrains only $\mathbf{k}\cdot\mathbf{v}$; the velocity perpendicular to $\mathbf{k}$ remains an aperture ambiguity. Actual GNSS geometry, unknown wavevectors, and evolving envelopes can further weaken the ideal separation limit.

Supplementary Note S13. Amplitude, phase, and interference response

Consider the auxiliary full-observation problem $\frac{1}{2}\|u-b\|_2^2 + \frac{1}{2}\lambda_{\text{eff}}\|D_v u\|_2^2$, with uniform sampling, periodic boundaries, a fixed exact translation, and no spatial or increment regularization. Here, b is a complete observed field, $D_v u = u_{t+1} - W_v u_t$, and λ_{eff} is its relative temporal weight. For $u(\mathbf{r}, t) = a\exp[\mathrm{i}(\mathbf{k}\cdot\mathbf{r} - \omega t)]$,

$$\begin{aligned} d_v(\mathbf{k},\omega) &= e^{-\mathrm{i}\omega\Delta t} - e^{-\mathrm{i}\mathbf{k}\cdot\mathbf{v}\Delta t}, \\ H_v(\mathbf{k},\omega) &= \left[1 + \lambda_{\text{eff}}|d_v|^2\right]^{-1} = \left[1 + 4\lambda_{\text{eff}}\sin^2\!\left(\frac{(\omega - \mathbf{k}\cdot\mathbf{v})\Delta t}{2}\right)\right]^{-1}. \end{aligned} \tag{S40}$$

The result follows from the scalar normal equation at each Fourier mode. With $\mathbf{v} = \mathbf{0}$, ordinary temporal smoothing attenuates modes according to their temporal frequency. Propagation compensation instead responds to frequency mismatch in the moving frame. A matched mode has unit gain only in this idealized problem.

For the numerical interpolation operator, replace the exact translation factor by its local Fourier response $\chi(\mathbf{k}, \mathbf{d})$, where $\mathbf{d} = \mathbf{v}\Delta t$. The mismatch becomes $e^{-\mathrm{i}\omega\Delta t} - \chi$. This exposes interpolation-induced attenuation and phase error even when the prescribed velocity is correct.

Wavelet regularization and nonuniform ray geometry prevent (S40) from serving as a global transfer function for the full inversion.

For K components, a complementary quadratic-mode calculation retains the duplication correction. Let

$$\min_{a_1,\dots,a_K} \frac{1}{2}\left|\sum_k a_k - b\right|^2 + \frac{1}{2}\sum_k w_k|a_k|^2, \qquad w_k = K\left(\lambda_{\text{eff}}|d_{v_k}|^2 + \nu_0\right) > 0, \tag{S41}$$

where ν_0 is a positive ridge weight in this auxiliary analysis. Reference response calculations use $\lambda_{\text{eff}} = 1$ and $\nu_0 = 10^{-6}$, with $\lambda_{\text{eff}} = 0.1$ and 10 as sensitivity values; these are not substituted for the normalized weights of (S19). The normal equations give $w_k a_k = b - \sum_\ell a_\ell$. Therefore,

$$a_k = \frac{w_k^{-1}}{1 + \sum_\ell w_\ell^{-1}}\, b, \qquad H_K = \frac{\sum_k w_k^{-1}}{1 + \sum_k w_k^{-1}}. \tag{S42}$$

For identical propagation fields, $w_k = Kw_0$ and $H_K = (1 + w_0)^{-1}$, exactly matching the one-component gain. An improvement cannot arise merely from duplicating the same propagation branch. For distinguishable branches, a mode can preferentially occupy the component with smaller propagation mismatch. If both branches match, allocation becomes ambiguous even when the total gain approaches unity.

A standing pattern illustrates the distinction: $a\cos(\mathbf{k}\cdot\mathbf{r} - \omega t) + a\cos(\mathbf{k}\cdot\mathbf{r} + \omega t)$ consists of oppositely propagating modes. For nonaliased $\omega \neq 0$, no single constant velocity satisfies both $\mathbf{k}\cdot\mathbf{v} = \omega$ and $\mathbf{k}\cdot\mathbf{v} = -\omega$. The one-pattern model must then use evolving amplitude, nonuniform apparent motion, or residual error. Two branches can represent the two translations, but their separation still depends on (S38) and the actual observation operator.

Finally, for a scalar evolution residual r, the auxiliary problem $\min_q\{\frac{1}{2}a_d(r-q)^2 + \lambda|q|\}$ has solution

$$q_* = \operatorname{sign}(r)\max\{|r| - \lambda/a_d, 0\}, \qquad \min_q\left\{\tfrac{1}{2}a_d(r-q)^2 + \lambda|q|\right\} = \begin{cases} \frac{1}{2}a_d r^2, & |r| \le \lambda/a_d, \\ \lambda|r| - \lambda^2/(2a_d), & |r| > \lambda/a_d. \end{cases} \tag{S43}$$

Here, $\lambda > 0$ is the scalar increment-penalty coefficient in this auxiliary problem, not an unnormalized substitution for the grouped penalty in (S19). Eliminating the increment therefore produces a robust evolution penalty. For the full grouped, non-square transform and positive stabilization, increments are obtained by (S28), not by applying (S43) voxel by voxel.

Supplementary Note S14. Model selection, validation, and reported quantities

S14.1. Spatially separated selection

Stations are grouped into fixed $2^\circ \times 2^\circ$ spatial blocks. After sorting block centres by longitude and latitude, seed 20261011 defines their permutation. Block boundaries are anchored at 140°W and 10°N, with half-open 2° intervals. For B occupied blocks, the first $\lceil 0.20B \rceil$ blocks form the independent test set; all their stations, arcs, and epochs are excluded from fitting, initialization, propagation estimation, working-weight fitting, and model selection. The remaining permuted blocks are allocated cyclically to five development folds. At least 25 occupied blocks are required

for the full five-fold design; smaller networks are flagged as insufficient for the prescribed selection rather than replaced by an undocumented split. Test-arc detrending uses only that arc and the frozen preprocessing settings.

For candidate model count K, hyperparameter configuration $\mathbf{h}$, and held-out development block b, define

$$\ell_b(K,\mathbf{h}) = \frac{1}{n_b}\sum_{i\in b}\left(\mathbf{A}_i\, n_0 \sum_k \widehat{\mathbf{u}}_i^{(k)} - y_i\right)^2, \qquad \overline{L}(K,\mathbf{h}) = \frac{1}{5}\sum_{f=1}^{5}\frac{1}{|\mathcal{B}_f|}\sum_{b\in\mathcal{B}_f}\ell_b(K,\mathbf{h}). \tag{S44}$$

Here, n_b is the number of retained validation observations in block b, $\mathbf{A}_i$ is the relevant ray row at its epoch, and $\mathcal{B}_f$ contains the validation blocks of fold f. Scoring uses squared TECU residuals without estimating validation-arc weights from their amplitudes. Averaging block losses prevents a densely instrumented block from dominating the selection.

Let $(K_0,\mathbf{h}_0)$ minimize $\overline{L}$, and let s_0 be the sample standard deviation of its five fold means divided by $\sqrt{5}$. The admissible set is

$$\mathcal{A} = \left\{(K,\mathbf{h}) : \overline{L}(K,\mathbf{h}) \le \overline{L}(K_0,\mathbf{h}_0) + s_0\right\}. \tag{S45}$$

The smallest admissible K is selected. Within that count, ties favor the configuration closest to the reference weights in squared log-ratio distance, followed by lower validation loss and, for an exact numerical tie, the declared configuration order: reference, lower and upper spatial weight, lower and upper propagation weight, lower and upper increment weight, and lower and upper velocity-smoothness weight. This one-standard-error rule is a complexity-control heuristic, not a calibrated confidence test. The selected count and weights are frozen for the event reconstruction, after which all non-test stations are refitted. Failed candidates are reported and cannot acquire a favorable score by excluding difficult folds or windows.

The reference weights are $\lambda_s = 0.01$, $\beta = 5$, $\lambda_q = 0.20$, $(\lambda_{vh},\lambda_{vt},\lambda_{vz}) = (0.05, 0.05, 0.01)$, $\lambda_{v0} = 10^{-4}$, and $\varepsilon = 10^{-6}$. The predefined development set contains nine configurations: the reference configuration; separate factors of 0.5 and 2 for each of $\lambda_s, \beta, \lambda_q$; and factors of 0.5 and 2 applied jointly to the three velocity-smoothness weights. Both model counts receive the same configurations and number of starts. These values are numerical reference settings, not experimentally established optima of the new model.

S14.2. Independent observations and structural errors

The total prediction at an independent station is $\widehat{\mathbf{y}}_{\mathrm{test,t}} = \mathbf{A}_{\mathrm{test,t}}\widehat{\mathbf{x}}_t$. The protocol requires RMSE, mean absolute error, signed residual mean, and residual autocorrelation at 5–30 min lags. Agreement with held-out TEC provides an observational test but cannot alone establish the recovered altitude distribution.

For synthetic truth $x^*_{j,t}$, averaged over each reconstruction voxel, define

$$E_{\mathrm{tot}} = \left[\frac{\sum_{j,t} V_j(\widehat{x}_{j,t} - x^*_{j,t})^2}{\sum_{j,t} V_j (x^*_{j,t})^2}\right]^{1/2}, \qquad E_{\mathrm{cmp}} = \min_{\pi\in\mathfrak{S}_K}\left[\frac{\sum_{k,j,t} V_j(\widehat{x}^{(\pi(k))}_{j,t} - x^{*(k)}_{j,t})^2}{\sum_{k,j,t} V_j (x^{*(k)}_{j,t})^2}\right]^{1/2}. \tag{S46}$$

The voxel volume is V_j, and $\mathfrak{S}_K$ is the set of label permutations. A single permutation is applied to the entire comparison window; voxel-wise relabeling, phase shifts, and amplitude rescaling are prohibited. Component error is evaluated only for compatible component counts and explicitly

separable truth. Count errors and unresolved cases are reported separately. Zero-disturbance controls use absolute error and reconstructed energy instead of a relative denominator.

Amplitude and phase are estimated by least-squares fitting $c\cos\omega t + s\sin\omega t$ at the prescribed synthetic frequency. Under the convention $A\cos(\omega t + \phi)$, $A = (c^2 + s^2)^{1/2}$ and $\phi = \mathrm{atan2}(-s, c)$. Interlayer phase errors use wrapped phase differences. Phase is reported only where the fitted amplitude exceeds the 95th percentile of the maximum fitted amplitude within the structural evaluation domain across 200 noise-only reconstructions processed with the same geometry and settings. The domain-maximum calibration is used to avoid interpreting isolated noise peaks as phase-bearing structure; it is an empirical detection rule, not a voxelwise posterior probability. Velocity error for a plane wave concerns its normal component, not an unidentifiable full vector.

Independent electron-density profiles are compared only when their background definition, vertical support, time, and spatial sampling are compatible. Reference matching tolerances are 5 min and 100 km. Absolute density is not subtracted directly from a disturbance estimate. The same centered 60 min background is applied only when a complete, uniformly sampled reference-profile sequence is available; no temporal interpolation across missing profiles is used to manufacture that background. Profile heights are averaged over the same 50 km reconstruction layers. If more than one profile satisfies the matching tolerances, all qualifying profiles are retained with their actual separations. Without compatible profiles, altitude validation is marked unavailable and is not inferred from TEC agreement.

S14.3. Synthetic and preprocessing controls

Controlled fields are generated independently of the reconstruction dictionary on a $0.25^\circ \times 0.25^\circ \times$ 10 km grid at 30 s spacing. A component wave packet is

$$\delta N_e^{*(k)}(\mathbf{r}, h, t) = A_k g_k(t) \exp\left[-\frac{\|\mathbf{r} - \mathbf{r}_{c,k}(t)\|_2^2}{2s_{h,k}^2} - \frac{(h - h_{c,k})^2}{2s_{z,k}^2}\right] \cos[\mathbf{k}_{h,k} \cdot \mathbf{r} - \omega_k t + k_{z,k}(h - h_{c,k}) + \phi_k]. \tag{S47}$$

Here, A_k is physical density amplitude; g_k is a dimensionless temporal envelope; $\mathbf{r}_{c,k}$ is packet centre; $s_{h,k}, s_{z,k}$ are horizontal and vertical envelope widths; $h_{c,k}$ is central height; $\mathbf{k}_{h,k}$ and $k_{z,k}$ are horizontal and vertical wavenumbers; ω_k is frequency; and ϕ_k is initial phase. Horizontal distances use an azimuthal-equidistant plane centred at $100^\circ\mathrm{W}, 37.5^\circ\mathrm{N}$.

The retained reference cases use $A_k = 5 \times 10^{10}$ el m^{-3}, $h_{c,k} = 325$ km, and $s_{z,k} = 100$ km. A 600 km wave traveling at 400 m/s has a 25 min period; an 1800 km wave at 600 m/s has a 50 min period. Their horizontal widths are 600 and 1200 km. Height-phase cases use $k_z = \pm\pi/(200\text{ km})$. Growth/decay envelopes have 30 min characteristic times. Synthetic azimuths span 0°, 45°, 90°, and 135° clockwise from north, and phases are drawn uniformly on $[0, 2\pi)$.

The experiment set includes sustained propagation, stationary growth and decay, traveling amplitude modulation, height-dependent phase offsets, counter-propagating packets, standing waves, almost coincident propagation branches, and pure noise. Each primary scenario uses 30 independent noise realizations with 0.05 TECU standard deviation. Sensitivities include 0.03 and 0.10 TECU noise, 30% and 60% station removal, 15 and 30 min gaps, serial correlation, and sparse outliers. Almost coincident branches and unrestricted-evolution controls explicitly test whether the algorithm withholds unsupported component interpretations.

All primary packets are centered at the origin of the reference plane at the midpoint $\tau_c = 3600$ s of the 120 min valid evaluation interval. The raw synthetic record extends beyond both ends of this interval by half the centered-filter span. Thus the 60 min reference background requires

180 min of raw data, and the 30 and 90 min sensitivity backgrounds require 150 and 210 min, respectively. These buffers are required even when only the central 120 min are inverted; a 120 min raw record would leave only 60 min with a complete 60 min filter window. With azimuth α measured clockwise from north, define $\mathbf{e}_\alpha = (\sin\alpha, \cos\alpha)$, $\mathbf{k}_h = (2\pi/\Lambda_h)\mathbf{e}_\alpha$, $\omega = 2\pi v/\Lambda_h$, and $\mathbf{r}_c(\tau) = v\mathbf{e}_\alpha(\tau - \tau_c)$. Here, Λ_h is horizontal wavelength in metres and v is the prescribed normal phase speed. The sustained case has $g(\tau) = 1$. The stationary growth-and-decay case uses $\mathbf{r}_c = \mathbf{0}$, $\omega = 0$, and $g(\tau) = \exp[-(\tau - \tau_c)^2/(2 \times 1800^2)]$; the spatial cosine and envelope widths are retained. The traveling amplitude-modulated case uses $g(\tau) = 1 + 0.3\sin[2\pi(\tau - \tau_c)/3600]$. In the reference case $k_z = 0$; height-phase tests replace it by either stated nonzero vertical wavenumber. A fixed phase is drawn once per component and retained throughout a realization.

Counter-propagating and standing-wave controls use the 1800 km wavelength and normal speeds $+600$ and -600 m s^{-1} along the same axis. For the standing wave, both spatial envelopes remain centered at the origin so that the unmodulated carriers form a standing interference pattern. Near-coincident branches use 600 and $600 + \Delta v$ m s^{-1}, with $\Delta v \in \{20, 50, 100\}$. An additional 8 min oscillation is injected as a sampling-stress test and is not treated as a recoverable reference wave at 5 min reconstruction spacing. Its purpose is to measure aliasing leakage into the retained field, not to extend the interpretation band.

The independent errors are Gaussian with the stated TECU standard deviations. Correlated errors follow $\eta_j = 0.8\eta_{j-1} + \zeta_j$, where ζ_j is independent Gaussian noise with variance $0.05^2(1 - 0.8^2)$ TECU^2, giving stationary standard deviation 0.05 TECU. Outlier tests add an independent signed 1 TECU jump to 1% of randomly chosen samples. Station-removal tests remove 30% or 60% of training stations without replacement. Gap tests remove a 15 or 30 min interval centered at τ_c from 25% of training stations; the preprocessing rules remain unchanged, including invalidation of windows crossing gaps. Reference truth fields and validation stations are never reselected to compensate for degraded coverage. Thirty independent realizations are used for each primary scenario, and 200 additional noise-only realizations determine the empirical amplitude-detection threshold. The random generator is PCG64 with seed 20261011 and deterministic child streams assigned in the declared scenario order.

Two types of experiment are distinguished. Algorithm-isolation tests integrate the signed truth directly through the ray model. End-to-end tests add a bounded signed perturbation to an absolute IRI background, generate slant TEC, and apply the identical frozen real-data detrending chain. The positivity-preserving component injection is

$$\delta N_{e,\mathrm{inj}}^{(k)} = \frac{0.3}{K} N_{e,\mathrm{bg}} \tanh\left(\frac{\delta N_e^{*(k)}}{n_0}\right), \qquad N_{e,\mathrm{inj}} = N_{e,\mathrm{bg}} + \sum_k \delta N_{e,\mathrm{inj}}^{(k)}. \tag{S48}$$

Here, $N_{e,\mathrm{bg}} > 0$ is the simulated absolute IRI background. The total perturbation has magnitude at most $0.3N_{e,\mathrm{bg}}$, so the injected density stays positive. The actually injected components, not their pre-transformation packet amplitudes, define the truth for this experiment. Applying the transformation separately preserves an explicit component truth while changing its waveform; the unmodified packet tests isolate the solver's sinusoidal response.

If $\mathcal{D}$ denotes the complete frozen arc-preprocessing map, its incremental observation response is $\mathcal{D}(s_{\mathrm{bg}} + \mathcal{A}x^*) - \mathcal{D}(s_{\mathrm{bg}})$, where s_{bg} is simulated background slant TEC and $\mathcal{A}$ is the fine-grid ray operator. Its difference from $\mathcal{A}x^*$ quantifies preprocessing-induced mismatch. Error relative to the injected field is then an end-to-end error, not attributed solely to the inverse solver. For a fixed accepted-sample mask, $\mathcal{D}$ is the linear centered filter in (S2); rejection and discontinuity decisions are archived separately. The same operator and valid-window mask are applied to the paired background-only and injected records. End-to-end error is reported relative both to the

actual injected density and to its retained slant response, rather than attributing filtering loss to the inverse solver alone.

The complete single-pattern propagation-plus-increment model is the principal baseline. Further ablations remove propagation, increments, or temporal coupling. A post-reconstruction multi-transport decomposition and an equal-budget, more flexible single-pattern fit help distinguish the benefits of joint inversion from additional model capacity. Equal-velocity duplicated components provide a numerical check of Proposition 1. Oracle-velocity experiments are labeled separately. For the fixed-location temporal-smoothing ablation, all velocities and increments are set to zero. The propagation-only ablation sets increments to zero while estimating velocity; the evolution-only ablation sets velocity to zero while estimating increments. Independent-epoch reconstruction omits both temporal terms and uses the initializer objective (S49), followed by its own development selection of the spatial weight. The post-reconstruction control first freezes that independent-epoch total field and then minimizes the same component and velocity regularizers subject to the exact constraint that the components sum to the frozen field. It is a controlled post hoc transport decomposition, not a claim to reproduce every variant of shifted proper orthogonal decomposition [16, 17]. Its total-field prediction cannot improve by construction, so its component errors isolate what can be learned after reconstruction. The flexible one-pattern control uses 2.5° horizontal and 15 min temporal velocity knots, with the same four control heights, physical speed bound, ten starts, and outer iteration budget. Normalization counts are recomputed for its mesh. The same development partitions and numerical acceptance standards apply to every baseline; actual runtime is reported because equal iteration budgets do not imply equal computational cost. Test data, rendering choices, and reconstructed visual appearance do not enter parameter selection.

S14.4. Uncertainty and interpretation flags

Regularized curvature, conditional observability, numerical convergence, and resampling variability are distinct outputs. The uncertainty protocol specifies 2000 paired station/time-block bootstrap resamples for summary-metric intervals with 60 min blocks, following the block-resampling principle for dependent observations [18]; synthetic uncertainty is evaluated across independent noise realizations. Spatially coherent network-wide errors require coherent network-time resampling in a sensitivity analysis. Spatial resampling uses the same 2° station blocks, and time blocks are resampled within fixed 6 h UTC strata to limit mixing of different storm stages. The same sampled blocks are used for every method in a paired comparison; percentile limits are the 2.5th and 97.5th percentiles. A stratum with fewer than four valid one-hour blocks is reported without a stratum-specific interval. These are empirical summary-metric intervals under a local dependence approximation, not voxelwise Bayesian credible intervals or a guarantee of coverage during a nonstationary storm.

For each window, the total field is retained together with numerical and coverage flags. Interpreting the velocities of individual components additionally requires stability across accepted starts and specified nuisance-support expansions, acceptable synthetic separation performance under comparable sampling, and independent validation support for the added model complexity. Numerical rank alone is insufficient. A two-component fit can therefore produce a retained total-field output with `DECOMPOSITION_UNRESOLVED`; this flag does not erase the total field or assert that only one physical wave existed.

Supplementary Note S15. Numerical specification and reconstruction procedure

Each inversion window contains 25 epochs over 120 min and advances by 60 min. Only its central half-open interval, from 30 to less than 90 min, is retained. Adjacent windows are not amplitude-matched or temporally smoothed after reconstruction. Missing endpoint windows are not synthesized by temporal reflection. The estimator uses observations on both sides of the output epochs and is an offline smoother, not a causal forecast.

For each epoch, the independent-epoch initializer is the unique minimizer

$$\mathbf{u}_t^{(0)} = \arg\min_{\mathbf{u}} \left\{ \frac{1}{m_t} \sum_i \rho_c\Big([\widetilde{\mathbf{H}}_t\mathbf{u} - \widetilde{\mathbf{y}}_t]_i\Big) + \frac{\lambda_s}{P_s}\|\mathbf{S}\mathbf{u}\|_1 + \frac{\varepsilon}{2N}\|\mathbf{u}\|_2^2 \right\}. \tag{S49}$$

Only training observations enter this solve. The reference spatial weight and ridge are used, and its primal–dual solve is subject to the same numerical residual criteria as the main reconstruction. An epoch with no accepted training observations is not assigned an observation-based initializer.

Its linearized velocity seed is obtained by fitting the same bounded control basis to $\partial_\tau u^{(0)} + \mathbf{v}_h(\theta)\cdot\nabla_h u^{(0)}$. Time derivatives are centered over 600 s in the interior and one-sided over 300 s at the two endpoints; horizontal derivatives are centered over adjacent grid centers in physical metres. On the propagation mask, the fitted objective is the mean squared residual after multiplying it by 300 s, plus $0.01\|\theta\|_2^2/N_\theta$. The constrained quadratic problem uses projected gradient iterations, at most 2000, and stops when relative control change is below 10^{-6} for five successive iterations. Its step is $0.95/L_{\mathrm{init}}$, where L_{init} is the computable bound $2\|B\|_1\|B\|_\infty/n_r + 0.02/N_\theta$, B is the linear map from controls to the dimensionless residual, and n_r is its row count. If the fit is numerically unresolved, the linearized seed is replaced by zero and that replacement is recorded. The initializer supplies a starting direction, not a velocity observation.

Initialization is deterministic. The independent-epoch solution supplies the total-field starting sequence, and the constrained linearized fit supplies the second velocity seed. Let $\mathbf{e}_j = (\sin\alpha_j, \cos\alpha_j)$ for $\alpha_j \in \{0^\circ, 45^\circ, 90^\circ, 135^\circ\}$. The one-component starts are zero, the linearized field, and $\pm 400\mathbf{e}_j$ m s^{-1}. The two-component starts are the two duplicated seeds, four pairs $(400\mathbf{e}_j, -400\mathbf{e}_j)$, and four pairs $(\mathbf{v}_{\mathrm{lin}} + 300\mathbf{e}_j, \mathbf{v}_{\mathrm{lin}} - 300\mathbf{e}_j)$, all projected to the speed bounds. Both counts thus use ten starts. These values initialize a nonconvex search and do not constrain the recovered directions or speeds. At fixed controls the inner solution is unique; the initial component states merely accelerate its numerical computation.

Each fixed-control solve first permits 5000 primal–dual iterations. Residuals are checked every 50 iterations in the rescaled variable $\mathbf{z}_c$. Let τ_p, σ_p denote the actual primal and dual steps, let $G_c(\mathbf{z}_c) = K\varepsilon\|\mathbf{z}_c\|_2^2/2$, and define the fixed-point residuals

$$\begin{aligned} R_z &= \frac{\|\mathbf{z}_c - \mathrm{prox}_{\tau_p G_c}(\mathbf{z}_c - \tau_p\mathcal{K}_c^{\mathsf{T}}p)\|_2}{\tau_p \max\{1, \|\mathbf{z}_c\|_2, \|\mathcal{K}_c^{\mathsf{T}}p\|_2\}}, \\ R_p &= \frac{\|p - \mathrm{prox}_{\sigma_p F^*}(p + \sigma_p\mathcal{K}_c\mathbf{z}_c)\|_2}{\sigma_p \max\{1, \|p\|_2, \|\mathcal{K}_c\mathbf{z}_c\|_2\}}. \end{aligned} \tag{S50}$$

The practical inner criterion requires $\max\{R_z, R_p\} \le 10^{-5}$ and relative change of $\mathbf{z}_c$ below 10^{-6} over the preceding 50 iterations on two consecutive checks. These are numerical fixed-point tests, not bounds on electron-density error. A line-search refinement tightens both thresholds by a factor of ten and can increase the total budget to 20,000 iterations for that solve; no further unbounded refinement is assumed. The primal–dual gap is evaluated and reported whenever numerically resolved. Failure

to meet the inner criterion invalidates a candidate for comparison until it is resolved, rather than giving that candidate a favorable score by silently dropping difficult windows.

The outer budget is 30 accepted iterations per start. Practical stability requires relative computed-objective change below 10^{-4}, relative total-field change below 10^{-3}, and RMS change of the two physical velocity components below 10 $\mathrm{m\,s^{-1}}$ for three successive accepted iterations. The velocity RMS is evaluated on the fixed grid within the propagation mask; labels are matched by one whole-window permutation before comparison. When a rigorous gradient-error bound e_g is available, a separate first-order certificate requires $\|\widetilde{\mathcal{G}}_1\|_2 + e_g \le 10^{-6}$ and $e_g \le 2.5 \times 10^{-7}$. Practical stability without this certificate is recorded as numerical stability only. Reaching an iteration limit, obtaining a small TEC residual, or seeing little visual change is not recorded as convergence.

Within each fixed model count and hyperparameter configuration, the accepted start with the lowest sufficiently resolved training profile objective is retained. Component labels may be matched between overlapping windows by one whole-window permutation minimizing component field differences. No shift or rescaling is introduced during matching; an ambiguous label match is retained as such.

All operators use double precision. For random compatible vectors $\mathbf{a}, \mathbf{b}$, each adjoint is checked using

$$\frac{|\langle \mathcal{K}\mathbf{a}, \mathbf{b}\rangle - \langle \mathbf{a}, \mathcal{K}^{\mathsf{T}}\mathbf{b}\rangle|}{\max\{1, |\langle \mathcal{K}\mathbf{a}, \mathbf{b}\rangle|, |\langle \mathbf{a}, \mathcal{K}^{\mathsf{T}}\mathbf{b}\rangle|\}} < 10^{-10}. \tag{S51}$$

Control derivatives are compared with centred finite differences over several step sizes, including a dimensionless step of 10^{-6}, away from second-derivative breakpoints. The relative derivative discrepancy target is 10^{-4}. Ray-segment lengths must reproduce the accepted ray's modeled-shell length within 1 m. Identity, constant-field interpolation, duplication, and Fourier-response checks complement these operator tests. They verify implementation consistency, not scientific reconstruction accuracy.

The ray matrix is stored sparsely and applied once to the summed field per data evaluation. Wavelets, local warps, and transpose scatters are matrix-free. For fixed height count and cached warp weights, an inner iteration requires work proportional to $\mathrm{nnz}(\mathcal{H}) + KTN$, with constants from the 18-dimensional vertical transforms and 16-point interpolation stencils; $\mathrm{nnz}(\mathcal{H})$ counts all nonzero ray-matrix entries in the window. Propagation-derivative evaluation adds characteristic integration and local B-spline operations. The two-component state-and-increment vector contains 8,731,800 doubles, approximately 69.9 MB before auxiliary arrays. No dense full space–time matrix or dense inverse Hessian is constructed. Profiled line searches require additional inner solves, so runtime and memory must be measured rather than inferred to improve automatically.

S15.1. Reconstruction procedure

The reconstruction begins by freezing the slant preprocessing, station-block partition, spatial library, propagation mask, and nine candidate configurations. Ray intersections and adjoint identities are checked before any velocity search. For each development fold, each candidate count, and each candidate setting, the ten prescribed starts undergo the same fixed-control reconstruction and profiled velocity updates. State and increment variables are solved at every current and trial velocity, and numerical accuracy is resolved before a trial is compared. The best numerically admissible training objective within that candidate determines its held-out predictions; validation data are not used to initialize or fit its fields.

The block-balanced development score and complexity rule then fix the component count and hyperparameters. All non-test stations are refitted with that frozen configuration. Independent

test predictions are calculated once for the reported comparison without returning those residuals to the optimizer or parameter selection. The same accepted geometry supports the synthetic injections, preprocessing-response tests, and ablations. Conditional nuisance projections and their support expansions distinguish constraints on the total density from assumptions needed to interpret individual propagation components.

The central hour of each solved window supplies the final density sequence in physical units. The output includes the summed disturbance, conditional components, pattern velocities, finite increments, held-out slant predictions, coverage metadata, and the separate numerical and interpretational flags. Overlap comparisons document window seams without shifting, scaling, or smoothing the solutions to improve their appearance. Failed windows and unresolved decompositions remain visible in the evaluation record.

The methodological claim concerns the joint recovery and qualification of superposed disturbance patterns. It does not presume a unique dynamical attribution, lossless recovery, or improved event performance before the comparisons specified above are completed.

Supplementary Note S16. Characteristic integration and operator bounds

S16.1. Geographic characteristic derivatives

For an ellipsoid with semimajor axis a_e and squared eccentricity e_e^2,

$$\mathcal{N}(\varphi) = \frac{a_e}{\sqrt{1 - e_e^2 \sin^2 \varphi}}, \qquad \mathcal{M}(\varphi) = \frac{a_e(1 - e_e^2)}{(1 - e_e^2 \sin^2 \varphi)^{3/2}}. \tag{S52}$$

Let $\mathbf{x} = (\lambda, \varphi)$ and write the right-hand side of (S16) as $\mathbf{f}(\mathbf{x}, \tau; \theta)$. Starting at a target grid centre, one midpoint backstep of signed duration $s = -60$ s is

$$\mathbf{k}_1 = \mathbf{f}(\mathbf{x}_j, \tau_j; \theta), \quad \mathbf{x}_m = \mathbf{x}_j + \tfrac{s}{2}\mathbf{k}_1, \quad \mathbf{x}_{j+1} = \mathbf{x}_j + s\mathbf{f}(\mathbf{x}_m, \tau_j + s/2; \theta). \tag{S53}$$

For a control perturbation $\mathbf{d}$, propagate the directional state derivative $\dot{\mathbf{x}} = \partial_\theta \mathbf{x}\, \mathbf{d}$, initialized to zero at the fixed target. Then

$$\begin{aligned} \dot{\mathbf{k}}_1 &= \mathbf{f}_x \dot{\mathbf{x}}_j + \mathbf{f}_\theta \mathbf{d}, \\ \dot{\mathbf{x}}_m &= \dot{\mathbf{x}}_j + \tfrac{s}{2}\dot{\mathbf{k}}_1, \\ \dot{\mathbf{x}}_{j+1} &= \dot{\mathbf{x}}_j + s\left[\mathbf{f}_x(\mathbf{x}_m, \tau_j + s/2)\dot{\mathbf{x}}_m + \mathbf{f}_\theta(\mathbf{x}_m, \tau_j + s/2)\mathbf{d}\right]. \end{aligned} \tag{S54}$$

The matrices $\mathbf{f}_x$ and $\mathbf{f}_\theta$ differentiate the *discrete* midpoint step, including the curvature and $\cos\varphi$ factors. Differentiating only the velocity numerator would omit a geographic contribution. Reverse-mode accumulation through these steps gives the transpose action required in (S24), without storing a dense control Jacobian.

S16.2. Keys kernel and its derivative

With $a_K = -1/2$ and $r \in \mathbb{R}$,

$$\kappa(r) = \begin{cases} (a_K + 2)|r|^3 - (a_K + 3)|r|^2 + 1, & |r| \le 1, \\ a_K|r|^3 - 5a_K|r|^2 + 8a_K|r| - 4a_K, & 1 < |r| < 2, \\ 0, & |r| \ge 2. \end{cases} \tag{S55}$$

For $r \neq 0$, its derivative is

$$\kappa'(r) = \operatorname{sign}(r) \begin{cases} 3(a_K+2)|r|^2 - 2(a_K+3)|r|, & |r| \le 1, \\ 3a_K|r|^2 - 10a_K|r| + 8a_K, & 1 < |r| < 2, \\ 0, & |r| \ge 2, \end{cases} \qquad \kappa'(0) = 0. \tag{S56}$$

The first derivative is continuous at the knots, whereas the second derivative is piecewise defined. With $\xi_{jp} = (\lambda_j^b - \lambda_p)/\Delta\lambda$ and $\eta_{jq} = (\varphi_j^b - \varphi_q)/\Delta\varphi$, differentiation of an interpolation weight $w_{jpq} = \kappa(\xi_{jp})\kappa(\eta_{jq})$ gives

$$\dot{w}_{jpq} = \kappa'(\xi_{jp})\kappa(\eta_{jq})\frac{\dot{\lambda}_j^b}{\Delta\lambda} + \kappa(\xi_{jp})\kappa'(\eta_{jq})\frac{\dot{\varphi}_j^b}{\Delta\varphi}. \tag{S57}$$

Equations (S54) and (S57) yield $\mathbf{G}_t^{(k)}\mathbf{d}$ by weighted accumulation against the component state. Since the kernel can have negative weights, interpolation does not provide a positivity or monotonicity guarantee; neither is assumed for signed disturbances.

S16.3. A certified step-size bound

Let $c_E = \|\mathbf{E}\|_2^2$, equal to the largest mirror-replication multiplicity, and let $h_*^2 = \max_t \|\widetilde{\mathbf{H}}_t\|_1 \|\widetilde{\mathbf{H}}_t\|_\infty$. Let w_* be an upper bound on $\|\mathbf{M}\mathbf{W}_t^{(k)}\|_2$ for the current controls, computed from the sparse interpolation rows and their induced 1- and infinity-norms. Then

$$\|\mathcal{K}_\Theta\|_2^2 \le L_*^2 = a_y K h_*^2 + a_s c_E + 3a_d(1 + w_*^2) + a_q c_E. \tag{S58}$$

Indeed, the norm squared of summing K fields is at most K, the orthogonal vertical and Haar transforms do not amplify Euclidean norm, and $\|\mathcal{L}_d\mathbf{z}\|_2^2 \le 3[(1+w_*^2)\|\mathbf{U}\|_F^2 + \|\mathbf{Q}\|_F^2]$. Taking $\tau_p = \sigma_p = 0.95/L_*$ satisfies the strict step-size condition. This bound may be conservative but is certified; an unchecked finite power iteration is not assumed to provide an upper bound. Diagonal preconditioning may be used only with its corresponding certified primal–dual step-size condition and consistently rescaled proximal maps.

For the rescaled variable $\mathbf{z}_c$, the corresponding bound is $L_{*,c} = \sqrt{D_z}L_*$, and the reference steps are $\tau_p = \sigma_p = 0.95/L_{*,c}$. All proximal coefficients and the primal ridge are transformed consistently. This is the same optimization problem under a change of variables.

Supplementary Note S17. Profile sensitivity and inner-accuracy control

S17.1. The active-face Hessian

Let $\mathbf{C}_0\mathbf{z} = 0$ collect the locally fixed zero rows of $\mathcal{L}_s\mathbf{z}$ and zero groups of $\mathcal{L}_q\mathbf{z}$. Redundant rows are removed before computing tangent quantities, and $\mathbf{R}$ spans $\ker \mathbf{C}_0$. This face is used only where primal nonzero coefficients and dual inactive margins support local stability; numerically ambiguous coefficients do not justify treating a face as fixed.

For unscaled standardized data residuals $\mathbf{r}_y = \mathcal{L}_y\mathbf{z} - \widetilde{\mathbf{y}}$, define $\mathbf{D}_\rho = \operatorname{diag}(1_{|r_{y,i}|<c})$, away from the Huber transition. At a nonzero increment coefficient group $\mathbf{w}_g$, the Hessian of its norm is

$$\nabla^2\|\mathbf{w}_g\|_2 = \frac{1}{\|\mathbf{w}_g\|_2}\left(\mathbf{I} - \frac{\mathbf{w}_g\mathbf{w}_g^\mathsf{T}}{\|\mathbf{w}_g\|_2^2}\right). \tag{S59}$$

Let $\mathcal{L}_{q,g}$ select that unscaled group and $\mathbf{e}_g = \mathbf{w}_g/\|\mathbf{w}_g\|_2$. The inner active-face Hessian is

$$\mathbf{H}_{\xi\xi} = \mathbf{R}^\mathsf{T}\Bigg[a_y\mathcal{L}_y^\mathsf{T}\mathbf{D}_\rho\mathcal{L}_y + a_d\mathcal{L}_d^\mathsf{T}\mathcal{L}_d + \mu_K\mathbf{I} + a_q\sum_{g:\mathbf{w}_g\neq 0}\frac{\sqrt{|g|}}{\|\mathbf{w}_g\|_2}\mathcal{L}_{q,g}^\mathsf{T}(\mathbf{I}-\mathbf{e}_g\mathbf{e}_g^\mathsf{T})\mathcal{L}_{q,g}\Bigg]\mathbf{R}. \tag{S60}$$

Nonzero ℓ_1 terms are locally linear. Inactive rows are already handled by $\mathbf{R}$, not assigned an artificial second derivative. Strong convexity gives $\mathbf{H}_{\xi\xi} \succeq \mu_K\mathbf{I}$.

Let $\mathbf{G}$ stack the control derivative maps in (S24). If $\mathbf{L}_q^{\mathrm{curv}}$ is any square-root factor for the group-curvature matrix in (S60), define

$$\mathbf{B} = \begin{bmatrix}\sqrt{a_y}\mathbf{D}_\rho^{1/2}\mathcal{L}_y\mathbf{R}\\ \sqrt{a_d}\mathcal{L}_d\mathbf{R}\\ \mathbf{L}_q^{\mathrm{curv}}\mathbf{R}\\ \sqrt{\mu_K}\mathbf{R}\end{bmatrix}, \qquad \mathbf{V} = \begin{bmatrix}\mathbf{0}\\ -\sqrt{a_d}\mathbf{G}\\ \mathbf{0}\\ \mathbf{0}\end{bmatrix}. \tag{S61}$$

Writing $\mathbf{R}_\Theta$ for the Hessian of the propagation prior, the reduced Gauss–Newton matrix is

$$\mathbf{G}_{\mathrm{prof}}^{\mathrm{GN}} = \mathbf{V}^\mathsf{T}[\mathbf{I}-\mathbf{B}(\mathbf{B}^\mathsf{T}\mathbf{B})^{-1}\mathbf{B}^\mathsf{T}]\mathbf{V} + \mathbf{R}_\Theta. \tag{S62}$$

It is positive semidefinite. It is not the exact Hessian in (S25), because residual-weighted second derivatives of the warp are omitted. It is also not data-only: rows for propagation consistency, group curvature, and stabilization occur in $\mathbf{B}$. Inverse actions are computed by constrained linear solves on the tangent space, without forming $\mathbf{R}$, the projector, or inverse matrices densely. At active velocity bounds, sensitivity is restricted to feasible directions; the interior Hessian formula is not interpreted as an unconstrained uncertainty estimate.

S17.2. Bounding the reduced-gradient error

For fixed controls, write the propagation residual as $\mathcal{D}_\Theta\mathbf{z}$, with $\mathcal{D}_\Theta = \mathcal{L}_d(\Theta)$. For scalar control coordinate j, define $\mathbf{B}_j\mathbf{z} = \partial_{\Theta_j}(\mathcal{D}_\Theta\mathbf{z})$. Its gradient contribution is $a_d(\mathbf{B}_j\mathbf{z})^\mathsf{T}\mathcal{D}_\Theta\mathbf{z}$. If $\|\mathbf{z}\|, \|\mathbf{z}_*\| \le R_z$, subtraction and Cauchy–Schwarz give

$$|g_j(\mathbf{z}) - g_j(\mathbf{z}_*)| \le 2a_dR_z\|\mathbf{B}_j\|_2\|\mathcal{D}_\Theta\|_2\|\mathbf{z}-\mathbf{z}_*\|_2. \tag{S63}$$

Consequently, one admissible bound in (S32) is

$$L_{\Theta z} = 2a_dR_z\|\mathcal{D}_\Theta\|_2\left(\sum_j\|\mathbf{B}_j\|_2^2\right)^{1/2}. \tag{S64}$$

A sublevel bound follows from $\mu_K\|\mathbf{z}\|_2^2/2 \le J_{\max}$, giving $R_z \le \sqrt{2J_{\max}/\mu_K}$. Local directional bounds can be much tighter. The small reference ε can make the global bound too conservative for a practical budget; this is reported rather than bypassed by calling an unverified gradient exact.

For a trial displacement $\mathbf{d} = \Theta^+ - \Theta$, requiring the gradient uncertainty bound e_g to satisfy $e_g\|\mathbf{d}\| \le \eta[-\widetilde{\mathbf{g}}^\mathsf{T}\mathbf{d}]$, with a fixed $0 < \eta < 1$, certifies that the approximate direction is a true descent direction. A reference choice is $\eta = 0.25$. Near stopping, the error must instead be smaller than the declared first-order tolerance. For an asymptotic inexact-gradient claim, the error schedule must satisfy the assumptions of the relevant convergence result; a finite fixed gap tolerance does not establish that claim.

S17.3. Constructing the conditional nuisance projection

The practical implementation of (S36) uses directional constrained least squares rather than dense Jacobians. Linearize (S18) at the fitted sequence:

$$\delta \mathbf{r}_t^{(k)} = \mathbf{M}[\delta \mathbf{u}_{t+1}^{(k)} - \mathbf{W}_t^{(k)} \delta \mathbf{u}_t^{(k)} - \delta \mathbf{q}_t^{(k)}] - \mathbf{G}_t^{(k)} \delta \boldsymbol{\theta}^{(k)}. \tag{S65}$$

For each conditional diagnostic, the fitted-support family and successive expansions specified in Section S12.2 determine the admissible increment and residual changes, including all stated boundary freedoms. Equation (S65) is solved together with those linear restrictions. The resulting feasible state perturbation is mapped through $\mathcal{L}_y$. Minimizing its scaled observation norm over nuisance perturbations computes the directional curvature in (S36). Directions for which the chosen restrictions do not admit a local state parameterization are labeled infeasible under that conditional family, not declared observed. Enlarging the permitted increment or residual space supplies the nuisance-expansion sensitivity test of Proposition 4.

The conditional projection may have a very large or complete nuisance range. An output of zero sensitivity is then a valid result. It must not be replaced by a positive prior curvature to obtain apparently informative velocities.

Supplementary Note S18. Reproducibility, coverage, and interpretation

S18.1. Parameter classes and frozen reference choices

The physical state variables are the total and component electron-density disturbances in $\mathrm{el\,m^{-3}}$, pattern velocities in $\mathrm{m\,s^{-1}}$, and finite evolution increments in $\mathrm{el\,m^{-3}}$. They are estimated, not assigned empirical truth values. The fixed normalization scales are $n_0 = 10^{11}\ \mathrm{el\,m^{-3}}$ and $v_0 = 1000\ \mathrm{m\,s^{-1}}$. The fixed analysis configuration is the 89,100-voxel grid, 300 s reconstruction interval, 60 min centered background, complete 18-mode vertical basis, three-level Haar representation, and 120 min overlapping windows specified in the main text. The structural evaluation region is 130°W–60°W, 20°N–55°N; it is not a claim that every voxel within that region is resolved.

The regularization reference is $(\lambda_s, \beta, \lambda_q) = (0.01, 5, 0.20)$, $(\lambda_{vh}, \lambda_{vt}, \lambda_{vz}) = (0.05, 0.05, 0.01)$, $\lambda_{v0} = 10^{-4}$, and $\varepsilon = 10^{-6}$. These are dimensionless numerical choices for the normalized objective, not physical constants or previously demonstrated optima. The nine-configuration development comparison determines the adopted setting using only development data. This rule fixes how a setting is obtained without inventing a data-dependent value before the experiment is performed. Density amplitudes, selected component count, event velocities, actual noise estimates, and available station coverage remain empirical outputs. A missing data product is recorded as unavailable, not replaced by a supposed observational value.

Geometry, input sampling, the noise floor, filter width, model bounds, initialization, optimizer tolerances, support thresholds, and validation partitions are fixed before the independent test set is evaluated. Measurement metadata, ephemerides, observed solar-index inputs to IRI, and instrument discontinuity flags are archived as input data rather than described as tuning parameters. The complete vertical basis and its singular values must be stored alongside the model-version identifier, library dates, input-index records, software version, and floating-point precision. A reproducible run must also retain the station-block assignment, candidate scores, failed cases, accepted starts, iteration traces, and the raw-to-retained observation index map.

S18.2. Coverage and numerical qualification

For each voxel and epoch, coverage metadata include the number of intersecting rays, number of distinct stations, summed path length, and the minimum and maximum accepted elevations. A voxel not intersected at that epoch is marked as having no direct ray support, even when its regularized estimate is finite. A window with fewer than ten distinct training stations, or fewer than ten occupied 2° station blocks, is classified as network-limited. These counts are screening thresholds, not guarantees of vertical resolution. Spatial resolution and component recoverability are assessed by the fine-grid injections under the actual ray geometry; unsupported altitudes are not declared resolved from grid size or temporal smoothness alone.

Each window retains separate flags for measurement coverage, inner fixed-point accuracy, practical outer stability, certified profile decrease, certified first-order stationarity, and the success of conditional-sensitivity solves. A numerically stable solution lacking a first-order certificate is not labeled certified. Numerical instability or an invalid test prediction remains a failed case in summary denominators. Adjacent-window overlap is used to assess seams by the volume-weighted RMS difference divided by the RMS of the two total fields, with a floor of $10^8\ \mathrm{el\,m^{-3}}$ in that denominator. Values above 0.20 are marked as window-inconsistent; no post hoc amplitude matching is applied.

S18.3. Conditional component interpretation

The total disturbance is the primary reported field for either candidate count. A two-component interpretation additionally requires selection of $K = 2$ by the development rule, resolved numerical fits, and consistency under the first nuisance-support expansion. Component stability is evaluated among starts whose final training objectives differ from the best by at most $10^{-3} \max\{1, |P_{\min}|\}$. A single whole-window permutation aligns their labels. If fewer than two distinct velocity starts reach that objective neighbourhood, multi-start component stability is classified as unassessed. Otherwise, a component is labeled unstable when its volume-weighted RMS difference from the selected component exceeds 20% of that component's RMS, or its pattern-velocity RMS difference exceeds $100\ \mathrm{m\,s^{-1}}$. Small components below the noise-only amplitude threshold are not used to assert a second physical wave.

The named directional sensitivity magnitudes are compared before and after the first nuisance-support expansion. A direction whose residualized sensitivity norm decreases by more than 50% is marked as strongly support-dependent; a direction below the numerical rank threshold is unresolved. Such thresholds are disclosure rules for this implementation, not confidence levels or physical identifiability theorems. When two components fail these checks, the summed field can remain reportable while its decomposition is marked unresolved. A zero sensitivity in the unrestricted-increment limit is the expected consequence of (S34), not a failure to be regularized away. Independent test predictions characterize out-of-sample agreement after the configuration is frozen; they do not trigger retrospective model tuning.

References

[1] G. Gordiyenko, Y. Litvinov, M. Zhiganbayev, V. Grichshenko, S. Nurakynov, A. Andreyev, and V. Zhigalov, "Ionospheric response to the severe geomagnetic storm of 19–21 January 2026 over Almaty (Kazakhstan) according to the Digisonde and GPS data," *Sensors*, vol. 26, no. 18, 2026, art. no. 5935, doi: 10.3390/s26185935.

[2] W. Rideout and A. Coster, "Automated GPS processing for global total electron content data," *GPS Solut.*, vol. 10, no. 3, pp. 219–228, 2006, doi: 10.1007/s10291-006-0029-5.

[3] A. Coster, J. Williams, A. Weatherwax, W. Rideout, and D. Herne, "Accuracy of GPS total electron content: GPS receiver bias temperature dependence," *Radio Sci.*, vol. 48, no. 2, pp. 190–196, 2013, doi: 10.1002/rds.20011.

[4] M. Guerra, C. Cesaroni, M. Ravanelli, and L. Spogli, "Travelling ionospheric disturbances detection: A statistical study of detrending techniques, induced period error and near real-time observables," *J. Space Weather Space Clim.*, vol. 14, 2024, art. no. 17, doi: 10.1051/swsc/2024017.

[5] P. J. Huber, "Robust estimation of a location parameter," *Ann. Math. Stat.*, vol. 35, no. 1, pp. 73–101, 1964, doi: 10.1214/aoms/1177703732.

[6] D. Bilitza, M. Pezzopane, V. Truhlik, D. Altadill, B. W. Reinisch, and A. Pignalberi, "The International Reference Ionosphere model: A review and description of an ionospheric benchmark," *Rev. Geophys.*, vol. 60, no. 4, 2022, art. no. e2022RG000792, doi: 10.1029/2022RG000792.

[7] M. Burger, H. Dirks, and C.-B. Schönlieb, "A variational model for joint motion estimation and image reconstruction," *SIAM J. Imaging Sci.*, vol. 11, no. 1, pp. 94–128, 2018, doi: 10.1137/16M1084183.

[8] B. Gris, C. Chen, and O. Öktem, "Image reconstruction through metamorphosis," *Inverse Problems*, vol. 36, no. 2, 2020, art. no. 025001, doi: 10.1088/1361-6420/ab5832.

[9] P. S. J. Spencer and C. N. Mitchell, "Imaging of fast moving electron-density structures in the polar cap," *Ann. Geophys.*, vol. 50, no. 3, pp. 427–434, 2007, doi: 10.4401/ag-3074.

[10] J. Ma, H. Fu, J. D. Huba, and Y. Jin, "A novel ionospheric inversion model: PINN-SAMI3 (physics informed neural network based on SAMI3)," *Space Weather*, vol. 22, no. 4, 2024, art. no. e2023SW003823, doi: 10.1029/2023SW003823.

[11] R. W. Schunk, L. Scherliess, J. J. Sojka, D. C. Thompson, D. N. Anderson, M. Codrescu, C. Minter, T. J. Fuller-Rowell, R. A. Heelis, M. Hairston, and B. M. Howe, "Global Assimilation of Ionospheric Measurements (GAIM)," *Radio Sci.*, vol. 39, no. 1, 2004, art. no. RS1S02, doi: 10.1029/2002RS002794.

[12] X. Guan, W. Xu, W. Li, X. Meng, X. Pan, D. Zhao, and K. Zhang, "Spatiotemporal features of concurrent storm-time ionospheric disturbances over the American sector during the January 2026 geomagnetic storm," *J. Geophys. Res. Space Phys.*, vol. 131, no. 7, 2026, art. no. e2026JA035553, doi: 10.1029/2026JA035553.

[13] R. G. Keys, "Cubic convolution interpolation for digital image processing," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 29, no. 6, pp. 1153–1160, 1981, doi: 10.1109/TASSP.1981.1163711.

[14] T. van Leeuwen and A. Y. Aravkin, "Variable projection for nonsmooth problems," *SIAM J. Sci. Comput.*, vol. 43, no. 5, pp. S249–S268, 2021, doi: 10.1137/20M1348650.

[15] A. Chambolle and T. Pock, "A first-order primal-dual algorithm for convex problems with applications to imaging," *J. Math. Imaging Vis.*, vol. 40, no. 1, pp. 120–145, 2011, doi: 10.1007/s10851-010-0251-1.

[16] J. Reiss, P. Schulze, J. Sesterhenn, and V. Mehrmann, "The shifted proper orthogonal decomposition: A mode decomposition for multiple transport phenomena," *SIAM J. Sci. Comput.*, vol. 40, no. 3, pp. A1322–A1344, 2018, doi: 10.1137/17M1140571.

[17] P. Krah, A. Marmin, B. Zorawski, J. Reiss, and K. Schneider, "A robust shifted proper orthogonal decomposition: Proximal methods for decomposing flows with multiple transports," *SIAM J. Sci. Comput.*, vol. 47, no. 2, pp. A633–A656, 2025, doi: 10.1137/24M164392X.

[18] H. R. Künsch, "The jackknife and the bootstrap for general stationary observations," *Ann. Stat.*, vol. 17, no. 3, pp. 1217–1241, 1989, doi: 10.1214/aos/1176347265.