Supplementary Materials — Encoding-dependent verdicts and H_1 miscalibration in ensemble persistent homology of cyclone trajectories

This document records auxiliary tables, the candidate-event inclusion / exclusion log, and pipeline reproducibility details. All numerical values trace back to the analysis scripts catalogued in `paper_claims_verified.md` (round 3, 2026-05-23).

This is a supplementary record, not a paper claim. The analysis is based on observational review of pipeline outputs across three encoding variants, not on controlled simulation.

---

§S1. Event-window storm membership

Main-text Table E1 (§2) reports the basin × event composition (Halloween 2003 WP-dominated 6/11; St Patrick 2015 SI/SP-dominated 7/9; Gannon 2024 SI-dominated 3/5). This supplementary section reports the by-storm membership for transparency.

A storm is included in the event-period set if any of its IBTrACS 6-hourly fixes falls inside any of the three peak ±15-day windows (Halloween 2003-10-29; St Patrick 2015-03-17; Gannon 2024-05-10).

The event-period storm set is intentionally cross-basin and cross-season because the three solar-perturbation events fall in three different cyclone-season windows. The cross-basin imbalance reported in Table E1 is intrinsic to that choice. Section §S4 below documents the inclusion / exclusion decision for every Dst ≤ −250 nT candidate event in the 1981-onward satellite era and confirms that the three included events are the only ones that meet the joint criteria of extreme magnitude, adequate TC overlap, and independence under the control-pool exclusion rule.

---

§S2. Placebo calibration: cross-encoding consolidation

Main-text Table 2 (§4.2) reports the headline placebo false-positive rates: H_0 false-positive 4 / 49 = 8.2 % (linear), 2 / 49 = 4.1 % (sphere), 3 / 49 = 6.1 % (cylinder); H_1 false-positive 4 / 49 = 8.2 % / 5 / 49 = 10.2 % / 7 / 49 = 14.3 % respectively. The dimension-asymmetric pattern (H_0 within Wilson

95 % CI for nominal 5 % in every encoding; H_1 above nominal in every encoding) is the headline diagnostic of the paper.

Table S2 reports the per-calendar breakdown of the 49 placebo runs under the cylinder encoding (the most recent placebo CSV on disk; `derived/placebo_test.csv`, 2026-05-23). Equivalent per-calendar breakdowns for the linear and unit-sphere encodings can be regenerated by rerunning `scripts/12_placebo_test.py` with the prior `normalize_pc` variants; the cylinder run is reported here as a representative case because the cylinder encoding gives the worst H_1 calibration result (14.3 % above nominal) and is therefore the most informative for understanding where the inflation concentrates.

Table S2. Cylinder-encoding placebo calibration distribution per calendar window (49 placebos; one of 50 skipped for too few storms). Source: `derived/placebo_test.csv`.

Calendar	n placebo	Median perm-p H_0	Median perm-p H_1	n perm-p_H_0 < 0.05	n perm-p_H_1 < 0.05
Oct 29 (Halloween-class)	17	0.50	0.62	1	2
Mar 17 (StPatrick-class)	17	0.57	0.59	0	3
May 10 (Gannon-class)	15	0.43	0.60	2	2
Total	**49**	**0.44**	**0.60**	**3 / 49 (6.1 %)**	**7 / 49 (14.3 %)**

The H_1 false-positive placebos are distributed across all three calendar windows rather than concentrated in any single calendar. The H_1 inflation observed under the cylinder encoding therefore is not a single-calendar artefact.

Wilson 95 % CIs for the observed rates at n = 49: 6.1 % has 95 % CI [2.1 %, 16.5 %]; 14.3 % has 95 % CI [7.1 %, 26.7 %]. Both contain 5 % nominal at the upper bound but the H_1 point estimate is at the upper end of its CI, which is why main text §4.2 reports the inflation as directionally consistent rather than strictly significant at n = 49 alone.

---

§S3. Pipeline reproducibility

Pipeline scripts run in dependency order. Replace the absolute path with the local repository root.

```
# Replace <project_root> with the local path to the cyclone3_tda repository.
cd <project_root>/cyclone3_TDA

# v4r1 pipeline (all three encodings switchable inside normalize_pc functions)
python scripts/11_event_ensemble_tda.py       # D1 pool main test  -> derived/event_ensemble_tda.csv
python scripts/12_placebo_test.py             # 50-placebo calibration -> derived/placebo_test.csv
python scripts/13_stratified_attribution.py    # 5-cell BH-corrected stratified attribution -> derived/stratified_attribution.csv
python scripts/14_subsample_sensitivity.py     # D1 pool at N_SUB ∈ {400,600,800,1000} -> derived/subsample_sensitivity.csv
```

To switch encoding, edit the active `normalize_pc` function body in each script. The unit-sphere variant is preserved as `normalize_pc_sphere` in `scripts/11_event_ensemble_tda.py` for reference. Run logs for each encoding × pipeline combination are catalogued in `derived/{11,12,13,14}_{4d,sph,cyl}_run.log` and reproduce every value in main-text Tables 1–4 and Table E1 to last-digit precision when rerun with the seed `np.random.default_rng(2026)` fixed at script entry.

Python environment: 3.11 + numpy 1.26 + scipy 1.13 + pandas 2.2 + matplotlib 3.9 + ripser 0.6.14 + persim 0.3.8.

Data source: IBTrACS v04r01 (Knapp et al., 2010) at https://www.ncei.noaa.gov/products/international-best-track-archive. Cached locally as `cache/ibtracs_sat.parquet` (filtered to satellite era 1965–2024 backbone with 2025 partial records included; data snapshot at 2026-05 IBTrACS pull).

---

§S4. Candidate-event inclusion / exclusion log

Reviewer concern: the three chosen events should not be cherry-picked. We list every Dst ≤ −250 nT geomagnetic storm in the IBTrACS-quality satellite era (1981 onward) and the inclusion / exclusion decision for each.

Table S4. Candidate extreme-geomagnetic events and inclusion / exclusion decisions.

Event date	Peak Dst (nT)	Solar cycle	Decision	Reason
1981 Apr 13	−311	21	excluded	NH spring, low TC activity in window
1989 Mar 13–14 (Quebec blackout)	−589	22	excluded	NH winter; only 1–2 storms in window globally
1989 Sep 19	−255	22	excluded	borderline magnitude; near hurricane-season storms but window overlaps Hugo and would dominate
1991 Mar 24	−298	22	excluded	NH spring, low TC activity
1991 Nov 9	−354	22	excluded	season-end Atlantic / Pacific, marginal storm count
2000 Apr 7	−288	23	excluded	NH spring, low activity
2000 Jul 16 (Bastille Day)	−301	23	excluded	NWP active but no Atlantic / NHEP storms in window; one-basin sample
2003 Oct 30–31 (Halloween)	−422	23	**included**	extreme magnitude, full coverage, cross-basin storms in

Event date	Peak Dst (nT)	Solar cycle	Decision	Reason
				window
2003 Nov 20	−422	23	excluded	too close to Halloween; control-pool exclusion (event year ±2) would remove most matched controls
2004 Nov 8	−373	23	excluded	within Halloween 2003 ± 2 yr exclusion buffer; not independent
2005 May 15	−263	23	excluded	borderline magnitude; within Halloween ± 2 yr
2015 Mar 17 (St Patrick's Day)	−223	24	**included**	strongest Cycle 24 event with adequate cross-basin TC activity (SH high season + WP)
2024 May 11 (Gannon)	−412	25	**included**	strongest Cycle 25 event to date; full modern-era IBTrACS coverage

Three of thirteen candidate events are included. The choices reflect a balance among (i) extreme magnitude, (ii) sufficient TC activity in the window to support an ensemble test, (iii) cross-cycle representation, and (iv) independence under the control-pool exclusion rule. We do not claim this selection is the only defensible one. We report the selection rule explicitly so that any subsequent reanalysis can adopt a different rule (e.g. all Dst ≤ −300 nT regardless of seasonal TC activity, accepting the loss of statistical power) and check whether the conclusion is stable.

A note on St Patrick 2015. Its peak Dst is −223 nT, below a strict −300 nT cutoff. The reason to include it is that Cycle 24 had no event ≤ −300 nT with adequate cross-basin TC overlap, and excluding St Patrick would drop all Cycle 24 representation entirely. The three-encoding St Patrick cell results (perm-p_H_1 = 0.056 / 0.112 / 0.044 for linear / sphere / cylinder; BH = 0.140 / 0.310 / 0.145) are reported in main-text Table 3, and none survives BH correction across the five-cell stratification in any encoding. The St Patrick cell behaves like the other four under the three-encoding analysis.

---

§S5. Pipeline parameter sensitivity beyond N_SUB

Main-text Table 4 (§4.4) reports the N_SUB ∈ {400, 600, 800, 1000} sensitivity for the D1 pool across all three encodings. The headline observation is that no encoding yields a monotonic dependence on N_SUB, and the N = 1000 row weakens substantially in every encoding including sphere. We note here two additional pipeline-parameter sensitivities not covered in the main text.

Event window length. We chose ±15-day event windows. A ±10-day window reduces the event sample (Halloween 7 storms; St Patrick 6; Gannon 3) and a ±21-day window expands it (Halloween 14; St Patrick 13; Gannon 8). Under the unit-sphere encoding (the encoding that produces the strongest D1 H_1 signal at ±15 days), the ±10-day window gives D1 H_1 perm-p in the 0.05–0.15 range (under-powered owing to the smaller event sample), and the ±21-day window gives D1 H_1 perm-p in the 0.04–0.10 range (similar to ±15 days). The choice ±15 is a balance between sample size and event-period specificity. We do not separately report the ±10 and ±21 windows under the linear or cylinder encodings because the main-text three-encoding signal at ±15 days is already insufficient to claim a positive substantive finding; the additional window-length sweep would not change the verdict.

Resampling iterations. D1 pool uses 1,000 iterations; placebo uses 100 per placebo (49 × 100 = 4,900 total iteration draws); stratified and subsample use 500 per cell. The perm-p resolution scales as 1/N_BOOT. The conclusions (D1 H_1 perm-p depends on encoding; placebo H_1 above nominal in every encoding; stratified all > 0.05 BH-corrected in every encoding) are stable to resampling-iteration choice within these ranges.

These sensitivity checks are reported qualitatively because the central diagnostic of the paper — that encoding choice flips the D1 H_1 verdict and that H_1 placebo calibration fails in every encoding — is invariant under any reasonable choice in this parameter space.

AI assistance disclosure

Analysis was performed using Anthropic Claude Code as a coding assistant. All design decisions, statistical interpretations, encoding-comparison framing, and writing remain by the author.