

Abstract. This supplement documents label provenance, purge and normalization sensitivities, physical preprocessing, the complete threshold–persistence surface, Brienzi information flow, grouped-partition robustness, the temporal null, and reproducibility assertions. It contains no additional Hollin performance claim beyond the single observed event.

S1 Label provenance

- 5 The public Hollin analysis notebook is available at <https://github.com/smouellet/hhdas>. The inspected commit was 2b6179395531a3f07964cbb2c535e5e2d53bff50. It plots selected channels over the event record but provides no explicit sequence-box coordinates. Table S1 therefore records deterministic scenario intervals around approximate anchors. These ranges are not confidence intervals.

Table S1. Target provenance and independence.

Target	Interval (UTC)	Provenance class	Independence
Head-scarp onset	13 Jan 2021 19:00–21:00 UTC	Approximate published point ± 60 min	Circular with DAS input
Rupture-zone trigger	14 Jan 2021 02:00–04:00 UTC	Approximate published point ± 60 min	Circular with DAS input
Peak strain rate	14 Jan 2021 11:00–13:00 UTC	Approximate published point ± 60 min	Circular with DAS input
ShapeArray onset	14 Jan 11:00–14 Jan 13:00 UTC	Two consecutive samples above pre-rain median + 3 unscaled MAD	Separate from evaluated DAS score

- 10 The ShapeArray calculation is repeated independently for each lobe and for 2-, 3-, and 4-MAD thresholds (Table S2). The onset interval starts at the observation preceding the first exceedance and ends at the second, confirming exceedance. This preserves the hourly sampling uncertainty rather than assigning a false point time.

Table S2. ShapeArray onset sensitivity. Thresholds use pre-rain data only.

Lobe	MAD	Threshold (10^{-5} m s $^{-1}$)	Onset interval	Detected	Primary lobe
Eastern	2	5.09	14 Jan 11:00–14 Jan 13:00	yes	yes
Eastern	3	5.90	14 Jan 11:00–14 Jan 13:00	yes	yes
Eastern	4	6.72	–	no	no

S2 Replay and physical preprocessing assertions

S2.1 Stress-test model specification

- 5 The section-split neural inputs are six trailing 5 min frames at 5 min stride. Each frame is nearest-neighbour resampled to 96×96 , robust-normalized separately by cable section using all finite samples from 12 January 2021 11:38:10 through 13 January 05:38:10 UTC (inclusive), and clipped to ± 8 . The frame encoder has convolutional channels 8, 16, and 32 with kernels 5, 3, and 3, batch normalization, GELU activation, two max-pooling stages, global average pooling, and a 48-unit projection. A one-layer 56-unit GRU is pooled by a 24-unit attention network and passed through a 32-unit head with dropout 0.25. The rain variant encodes endpoint rainfall totals over 30 min, 1 h, 6 h, and 24 h into 16 units before concatenation. Rain
- 20 covariates are standardized from sections 1–4 only. Both models use class-weighted binary cross-entropy, AdamW (learning rate 8×10^{-4} , weight decay 10^{-4}), batch size 32, seed 42, and at most six epochs; section 5 validation average precision selects the epoch and section 6 is the reported test. The rainfall logistic uses those four totals, training-set median imputation and standardization, class balancing, seed 42, and 5,000 maximum iterations on the same 1–4/6 split. These models estimate same-event cross-section reconstruction only.

25 The blocked sensitivity uses 6 h test blocks and removes fitting endpoints within 30 min, 6 h, or 24 h of each block. The 24 h purge also covers the longest rainfall-history support. Table S3 reports class support without pooling empty-class folds.

Table S3. Hollin purge sensitivity. Test and train support count folds containing both target classes.

Purge (h)	Blocks	Mixed test	Mixed train	Removes 24 h rain support
0.5	13	2	13	no
6.0	13	2	13	no
24.0	13	2	7	yes

The paired normalization sensitivity refits both neural variants on the same GPU for seeds 42, 61, and 101. Section-specific parameters use each section’s pre-event segment, including the evaluation section. The transferred alternative pools robust centers and scales only across sections 1–4 and applies them unchanged to sections 5 and 6. Table S4 reports transferred minus
30 section-specific differences; the sign changes across models or seeds, so this is sensitivity rather than model selection.

Table S4. Paired normalization-transfer sensitivity. Ranges span the three paired seeds.

Model	Seeds	Δ BACC range	Δ AP range
DAS only	3	-0.034–0.097	-0.055–0.094
DAS plus rain	3	0.001–0.209	-0.003–0.036

S2.2 Replay chronology controls

Every transparent score uses a 60 s trailing window and is emitted at its final sample. Normalization fitting uses [12 January 2021 11:38:10, 13 January 06:00:00) UTC, threshold calibration uses [13 January 06:00, 10:00) UTC, and the background holdout uses [13 January 10:00, 12:00) UTC. The latest threshold-calibration input is 2021-01-13 09:59:39 UTC, before
35 rainfall at 13 January 2021 12:00 UTC. Seven acquisition gaps split the persistence stream. No score, normalization parameter, or threshold uses a later timestamp than the state it supports. Anchor-relative timing requires a new alert-episode start after the scenario anchor; an episode already active at that anchor is retained as pre-anchor burden.

Table S5. Design provenance for the fixed retrospective replay. “Selected using event timing” refers to rupture-time outcome selection, not to whether the design was chosen during a retrospective study.

Choice	Value	Basis	Selected using event timing?	Sensitivity or caveat
Window	60 s trailing	Engineering resolution choice	No	Not varied
Stride	30 s	Replay cadence	No	Persistence varied
Normalize	12 Jan 11:38:10–13 Jan 06:00	Sequential pre-rain split	No	End exclusive
Thresholds	13 Jan 06:00–10:00	Sequential pre-rain split	No	Quantiles 95–99.9%
Background	13 Jan 10:00–12:00	Untouched pre-rain holdout	No	Not used to select thresholds
Aggregation	Top-two mean	Distributed response assumption	Not tuned	Availability sensitivity
Target interval	+/-60 min	Analyst timing scenario	Yes, interval width	Source-author bounds requested
ShapeArray	2 samples above median + 3 MAD	Movement proxy	Rule introduced here	2/3/4 MAD

The physical artifact contains 5400 records with shape 60 by 38 in s^{-1} , plus an equal-sized Boolean mask. Time blocks are arithmetic means over five 1 s samples. Distance blocks use the public along-cable coordinate and 4 m bins after orienting
40 from high to low elevation. Sections shorter than 152 m are padded; the longest section extends 1.35 m beyond the common support. No future-value interpolation fills a missing cell.

S3 DAS-only preprocessing ablation

A small masked autoencoder is fitted only before 13 January 06:00. A held-out pre-rain interval from 06:00 to 12:00 sets the 99th percentile reconstruction threshold. The ablation compares two representations and two normalization modes with
45 identical ten paired seeds. It is an algorithmic-sensitivity experiment, not a multi-event uncertainty analysis.

Table S6. Section availability and monitoring support. Complete support is the fraction of score timestamps with all six sections valid.

Window	Timestamps	Valid sections	Complete support %	Distinct section sets
Normalization fit	2202	6–6	100.0	1
Threshold calibration	480	6–6	100.0	1
Background holdout	240	6–6	100.0	1
Post-rain replay	5796	6–6	100.0	1
Rupture point, pre-6 h	542	6–6	100.0	1

Table S7. Top-two aggregation sensitivity to section support.

Variant	Usable times	Threshold rows	Lower-state burden h	6 h occupancy	Status
reported top two available sections	8718	480	16.06	0.000	computed
fixed complete section subset	8718	480	16.06	0.000	computed
require at least five sections	8718	480	16.06	0.000	computed
conditional null by valid section count	8718	480	–	–	constant support; same as reported

The representation changes both timing and burden. The pooled physical grid gives mean burden 27.7 h and mean anchor-relative declaration delay 3.2 h, compared with 24.0 h and 4.0 h for the pooled 96 by 96 square. The physically scaled grid is therefore not declared superior: it preserves a 5 s by 4 m receptive field and mask, but the single event cannot identify a deployable preprocessing choice.

Table S8. Ten-seed preprocessing summary. Mean and SD describe algorithmic variability only.

Representation	Norm.	Seeds	Burden h, mean (SD)	Delay h, mean (SD)	6 h occupancy
Square 96x96	Pooled	10	24.04 (2.50)	3.96 (1.54)	0.203
Square 96x96	Section	10	22.50 (2.61)	4.74 (1.03)	0.163
Physical 60x38	Pooled	10	27.73 (0.79)	3.24 (0.96)	0.285
Physical 60x38	Section	10	27.39 (0.53)	3.74 (0.50)	0.264

50 **S4 Threshold–persistence surface**

The primary displayed range contains the 20 combinations of pre-rain quantiles 0.95–0.995 and five persistence lengths. The complete archived surface also retains five 0.999 rows as exploratory empirical-resolution sensitivities. The generated CSV records calibration burden, post-rain burden, delay to a new post-anchor episode, pre-anchor burden, and fixed 6 h rupture-horizon occupancy. No row is ranked against rupture timing. At the rupture point anchor every 6 h occupancy is zero. The 0.995 trace is illustrative because the calibration segment has approximate effective sample size 14.7; it is not treated as a stable tail estimate. “Pre-rain-fixed” means that fitted parameters remained frozen during replay, not that the retrospective design was preregistered.

S5 Brienzi split and event details

All grouped and date protocols keep the 15 augmented views together. The published-date test contains 381 label groups, all dated on or after 15 June 2023. One label timestamp at 31 May 2023 23:59:57.120 UTC is outside the literal 1 June training boundary and is excluded, leaving 1,109 training groups. Table S11 identifies where splitting enters the released processing chain and where upstream exposure cannot be recovered.

The event replay contains 97 radar events and 101 merged DAS episodes over 364.6 monitored h. A qualifying output has argmax C2 and probability at least 0.75. Outputs represent 25.6 s windows on a 20.48 s start grid. An episode runs from the first qualifying start through the end of the final qualifying window, and starts merge only when their start-to-start separation is strictly less than 60 s. An episode is eligible when its support intersects a radar interval expanded by 60 s at each boundary. Maximum-cardinality one-to-one assignment then minimizes boundary distance and absolute support-start offset

Section 4, window ending 14 Jan 2021 12:03:09 UTC

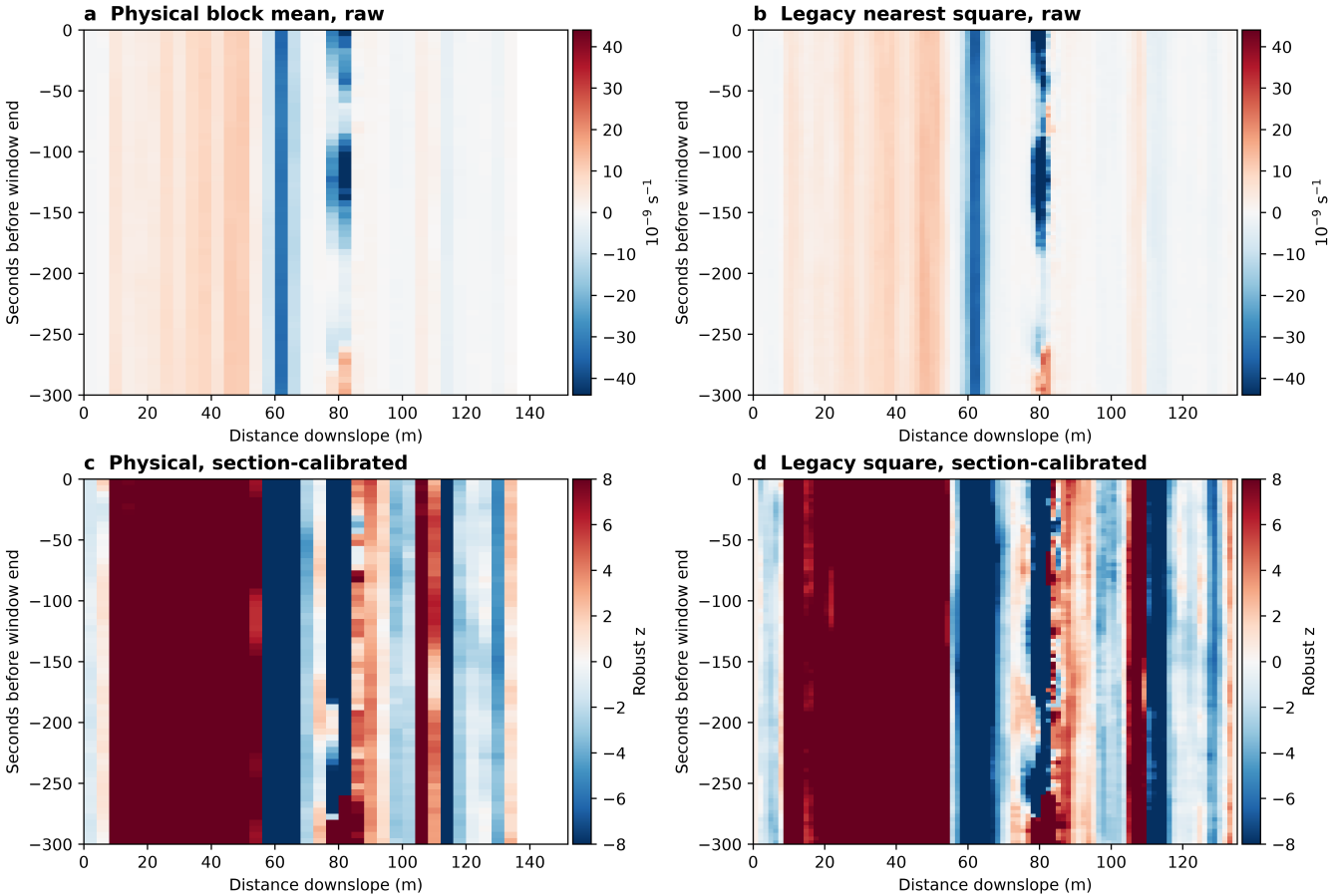


Figure S1. The same section-4 trailing window under the physical 5 s by 4 m representation and the legacy nearest-neighbour square, before and after section-specific pre-rain robust normalization. White cells are masked. The square changes the sampling geometry; normalization does not restore physical scale.

Table S9. Complete fixed surface summarized by threshold quantile. Ranges span persistence settings.

Quantile	Persistence steps	Burden range h	Head-point delay range h
0.950	1, 2, 5, 10, 20	17.54–20.26	0.80–7.33
0.975	1, 2, 5, 10, 20	16.01–18.12	0.80–7.73
0.990	1, 2, 5, 10, 20	15.22–16.82	0.80–7.79
0.995	1, 2, 5, 10, 20	14.89–16.30	0.81–7.82
0.999	1, 2, 5, 10, 20	13.95–15.66	0.81–8.14

lexicographically and is invariant to input row order. Support-start offset is DAS support start minus radar start; the first-window-completion offset adds 25.6 s but excludes acquisition, processing, and telemetry delay. Five-minute aggregation is used only in Figure 5. The primary setting uses no date-level exception.

The original paired test uses one class-stratified partition (seed 20260712), the same 373 held-out groups, and released view 0 for evaluation in every condition (C0/C1/C2/C3 test counts 198/31/96/48). The grouped baseline and each 1/5/14-sibling size-matched exposure control contain 16,770 fitting rows with class counts 8,910/1,380/4,320/2,160; the full 14-sibling exposure contains 21,992 rows with counts 11,682/1,814/5,664/2,832. Removed baseline rows are sampled without replacement within

DAS-only physical preprocessing ablation: every optimization seed

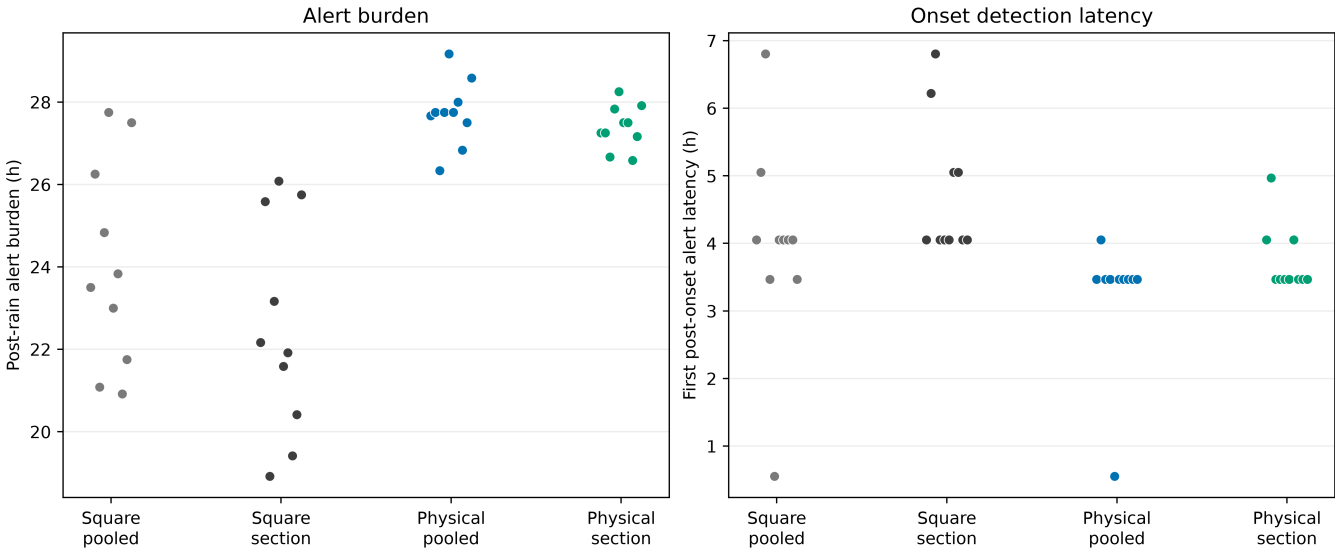


Figure S2. Every random-seed repetition for the DAS-only preprocessing ablation. Each condition contains the same ten seed identifiers. No confidence bars are shown.

Table S10. Fixed rupture-horizon occupancy across 241 possible onsets from 14 January 2021 02:00–04:00 UTC. The point anchor is 03:00 UTC. Guaranteed and maximum are the minimum and maximum scenario occupancies; values are not true-positive rates.

Level	Horizon (h)	Point anchor	Guaranteed	Non-zero scenarios (%)	Maximum
Lower (0.990)	1	0.000	0.000	16.6%	0.333
Lower (0.990)	3	0.000	0.000	16.6%	0.147
Lower (0.990)	6	0.000	0.000	16.6%	0.074
Lower (0.990)	12	0.000	0.000	16.6%	0.032
Upper (0.995)	1	0.000	0.000	15.4%	0.308
Upper (0.995)	3	0.000	0.000	15.4%	0.136
Upper (0.995)	6	0.000	0.000	15.4%	0.068
Upper (0.995)	12	0.000	0.000	15.4%	0.029

Table S11. Brienz information-flow ledger. Exposure status is bounded by the public release.

Stage	Unit	Dimensions	Exposure status
raw DAS record	25.6 s signal window	6000 time samples by 512 channels in source files	not reconstructed
released augmented feature table	label group view	1491 groups by 15 views by 14336 features	upstream feature construction and global exposure unresolved
compact downstream representation	label group view	14 frequency rows by 16 channel blocks by mean and RMS equals 448	deterministic pooling without labels
group partition	manual label group with all 15 views	class stratified 75 25 partition	held out group siblings controlled by protocol
evaluation row	first released view per held out group	one 448 feature row per test group	view 0 is first released view not verified unaugmented
ExtraTrees classifier	released downstream representation row	160 trees four classes	claim limited to sibling exposure within released table

75 class and fixed before the ten model-seed repetitions. Table S14 retains those algorithmic-sensitivity means. The stronger robustness check repeats the group partition ten times, uses three removal draws per partition, and fixes model randomness; Table S15 therefore uses the label-group partition as its variation unit. The structured CSV files retain every confusion matrix and class recall.

Table S12. Brienzy split support and grouped metrics.

Protocol	Train	Test	Overlap	BAcc	Estimand
Naive augmented views	1491	1478	1478	0.943	augmented view interpolation
Grouped labels	1118	373	0	0.900	held out label group interpolation
Published dates	1109	381	0	0.817	future date label group generalization

Table S13. Complete Brienzy partial-catalog results by radar-event size and pre-event RMS-noise stratum. Agreement is matched/catalog events (fraction); timing values are median [IQR] in seconds. Size bins are [0,2.5), [2.5,3.0), and [3.0,10.0].

Stratum	Catalog agreement	Support-start offset	First-window completion offset
All future-date events	43/97 (0.443)	-10.3 [-23.8, 8.4]	15.3 [1.8, 34.0]
Size: small	11/42 (0.262)	-7.6 [-13.3, 10.4]	18.0 [12.3, 36.0]
Size: medium	12/30 (0.400)	-11.2 [-24.7, -8.3]	14.4 [0.9, 17.3]
Size: large	20/25 (0.800)	-12.1 [-36.6, 12.4]	13.5 [-11.0, 38.0]
Noise: low	5/18 (0.278)	-20.4 [-35.5, 11.1]	5.2 [-9.9, 36.7]
Noise: medium	17/26 (0.654)	-12.6 [-39.8, -7.6]	13.0 [-14.2, 18.0]
Noise: high	21/53 (0.396)	-9.4 [-19.7, 21.1]	16.2 [5.9, 46.7]

Table S14. Paired sibling-exposure contrasts. SD and range describe algorithmic variability, not scientific replication.

Contrast	Exposed rows	Baseline rows	Mean Δ BAcc	SD	Range
Full exposure - grouped	21992	16770	0.046	0.013	0.027–0.067
1 sibling, size matched - grouped	16770	16770	0.009	0.012	-0.011–0.032
5 siblings, size matched - grouped	16770	16770	0.047	0.011	0.028–0.066
14 siblings, size matched - grouped	16770	16770	0.047	0.016	0.024–0.071

Table S15. Grouped-partition robustness. Ranges span partition-level mean differences after averaging three class-matched removal draws.

Exposure	Partitions	Median Δ BAcc	Range	Positive partitions
Full exposure	10	0.042	0.018–0.063	10/10
1 sibling, matched	10	0.012	-0.003–0.022	9/10
14 siblings, matched	10	0.045	0.025–0.064	10/10
5 siblings, matched	10	0.032	0.015–0.045	10/10

85 The temporal null applies one common circular shift to each day’s complete episode sequence, preserving episode count, duration, relative spacing, monitored support, and radar intervals. Table S16 reports all event-size strata. The empirical probability is $(1 + \#\{N_{\text{null}} \geq N_{\text{obs}}\})/(5000 + 1)$; it is a conditional timing test, not an event-generalization probability.

Table S16. Timing-preserving Brienzy temporal null. Null values are median [2.5th, 97.5th percentile] matched radar events.

Policy	Size stratum	Radar events	Observed	Null matches	Empirical p
All test days	All	97	43	14 [8, 20]	0.0002
All test days	Small	42	11	4 [1, 7]	0.0002
All test days	Medium	30	12	4 [2, 8]	0.0002
All test days	Large	25	20	6 [2, 11]	0.0018
Excluding 15 June	All	33	9	0 [0, 1]	0.0002
Excluding 15 June	Small	23	5	0 [0, 1]	0.0002
Excluding 15 June	Medium	10	4	0 [0, 1]	0.0002
Excluding 15 June	Large	0	0	0 [0, 0]	1.0000

85 The full Brienzy sensitivity CSV contains 108 combinations, including one-to-one versus many-to-one matching and a labelled day-level 15 June sensitivity. The latter is reported only as a sensitivity to sustained activity on that date, not as the headline strict estimand. Removing 15 June changes the catalog match fraction to 0.273; the day contributes 66.0% of the radar-event count.

Table S17. Reconciliation of the article, released notebook automatic stage, and audit interventions. The article’s 54 manual confirmations are not individually released.

Stage	C2 threshold	Radar	DAS units	Radar matched	Automatic hits	Manual added	DAS confirmation
Article, automatic + manual	0.70	727	993	295	896	54	0.957
Notebook automatic rerun	0.70	727	993	295	896	–	0.902
Future-date windows	0.70	97	345	44	–	–	0.493
Future-date episodes, q=.70	0.70	97	107	43	–	–	0.402
Future-date episodes, q=.75	0.75	97	101	43	–	–	0.426

Table S18. Complete future-date daily counts and leave-one-day-out (LODO) catalog-match fraction. Match entries are numerator/denominator.

Date (UTC)	Radar match	DAS confirmed	Unconfirmed DAS	Monitored h	LODO radar match
15 Jun 2023	34/64	34/83	49	24.00	0.273
16 Jun 2023	4/8	4/7	3	24.00	0.438
17 Jun 2023	4/8	4/4	0	24.00	0.438
18 Jun 2023	0/3	0/0	0	24.00	0.457
19 Jun 2023	0/1	0/0	0	24.00	0.448
20 Jun 2023	0/1	0/0	0	24.00	0.448
21 Jun 2023	0/4	0/0	0	24.00	0.462
22 Jun 2023	0/2	0/0	0	24.00	0.453
23 Jun 2023	1/2	1/1	0	24.00	0.442
24 Jun 2023	0/2	0/1	1	24.00	0.453
25 Jun 2023	0/0	0/0	0	24.00	0.443
26 Jun 2023	0/0	0/0	0	24.00	0.443
27 Jun 2023	0/1	0/2	2	24.00	0.448
28 Jun 2023	0/1	0/1	1	24.00	0.448
29 Jun 2023	0/0	0/2	2	24.00	0.443
30 Jun 2023	0/0	0/0	0	4.62	0.443

Table S19. Brienzt strict one-to-one sensitivity summarized by C2 threshold. Ranges span merge gaps of 30, 60, and 120 s and match tolerances of 30, 60, and 120 s.

C2 threshold	Catalog match fraction	Radar-confirmation fraction	Unconfirmed episodes	Detection-state h
0.50	0.412–0.474	0.234–0.291	100–147	2.69–3.19
0.75	0.402–0.464	0.352–0.506	39–81	2.07–2.51
0.90	0.402–0.464	0.396–0.580	29–67	1.89–2.36

S6 Reproducibility assertions

90 The test suite and artifact generator verify: one global rainfall/label value per timestamp; zero prohibited timestamp overlap at every purge; ordered target intervals; exact reporting of events, storms, unique times, monitored hours, and overlapping frames; no future input in an alert; no credit for a pre-anchor episode; pre-rain-only normalization and thresholds; transferred normalization fitted without held-out sections; grouped Brienz views; future-date Brienz tests; ten grouped partitions; 5,000 temporal-null draws per policy; 2,000 empirical calendar-day resampling rows; ten paired seeds per physical-preprocessing condition; and five main figures plus six main tables. The command manifest and SHA-256 inventories are included in the release candidate.