Predicting Instability in the Texan Electricity Grid using the Hurst Exponent

Stuart Woolley
stuart.woolley@resourcekraft.com

April 16, 2026

Abstract

We test whether the Hurst exponent carries operationally useful leading-indicator information for price instability in the ERCOT real-time market. On seven years of 15-minute settlement prices across four zonal hubs (2019–2025, $\approx 10^6$ samples), we compute the global Hurst exponent via three independent methods (R/S, DFA, MFDFA), segment by season, time-of-day and day-of-week, and test whether a rolling local H at three window sizes (24 h, 48 h, 7 d) predicts next-horizon (15 min, 1 h, 4 h, 24 h) spike events via the receiver-operating-characteristic area under curve (ROC-AUC).

Across all 48 (hub, window, horizon) cells, every AUC falls between 0.425 and 0.502 — uniformly at or below random guessing, and consistently in the direction that low H (mean-reverting regime) slightly precedes spikes. An independent test comparing local H on 24-hour pre-spike windows against matched non-spike windows returns a difference of at most 0.022 on a within-class standard deviation of 0.11.

We conclude that the Hurst exponent, used as a standalone predictor, is *not* operationally useful for near-term spike forecasting on ERCOT. The toolchain is validated against a previously published European-market whitepaper to machine precision, so the null result is not an implementation artefact. Implications for follow-on work (wavelet decomposition, multi-factor models) are discussed.

1 Introduction

The Electric Reliability Council of Texas (ERCOT) operates the most volatile energy-only wholesale electricity market in the developed world. Unlike PJM, CAISO, or the European markets which pay capacity to generators for being available, ERCOT generators earn revenue only from energy actually delivered. Scarcity pricing — capped at $\$9{,}000$/MWh until 2022, subsequently lowered to $\$5{,}000$/MWh following the post-Uri Public Utility Commission of Texas reforms — is the only mechanism signalling supply stress. The result is a bimodal market: long periods of calm prices around $\$20$–$\$40$/MWh punctuated by violent scarcity events in which prices briefly hit the cap.

Classical time-series forecasting (ARIMA, LSTM, and their variants) are averaging methods: they predict the typical next price well and the tail events poorly. But the tail events are where grid operators, retail electricity providers, and battery dispatchers either make their year or lose it. A tool that gives even modest early warning of an impending scarcity event would have immediate operational value.

The Hurst exponent $H \in [0, 1]$ is a well-established one-number summary of the *memory* of a time series [1, 2]. A price series with $H = 0.5$ is a random walk; $H > 0.5$ indicates persistent, trend-continuing dynamics; $H < 0.5$ indicates mean-reverting, trend-reversing dynamics. The question we address is direct:

> *Does a rolling local Hurst exponent, computed on the recent window of ERCOT returns, carry predictive information about the probability of a scarcity-price event in the next 15 minutes to 24 hours?*

A positive finding — AUC substantively above 0.5 — would be an operationally useful early-warning signal. A negative finding is equally publishable: it rules out a class of tools and directs follow-on work elsewhere.

2 The ERCOT Market and Data

2.1 Data

We use ERCOT report NP6-785-ER ("Real-Time Market Settlement Point Prices at Resource Nodes, Hubs, and Load Zones"), published as yearly archives covering the four zonal hubs (`HB_NORTH`, `HB_HOUSTON`, `HB_SOUTH`, `HB_WEST`) at 15-minute settlement intervals. The dataset spans 2019-01-01 through 2025-12-31 inclusive (seven full calendar years) of roughly 2.45×10^5 samples per hub, or $\approx 10^6$ total.

Every sample retains its sign and magnitude: negative prices (wind surplus) and scarcity-cap hits are preserved. After ingest the data pass a four-way sanity check: expected row count ($35{,}040$ non-leap, $35{,}136$ leap), strictly monotonic timestamps, zero duplicates, zero gaps, and explicit verification that the Winter Storm Uri window (2021-02-14 to 2021-02-19) is present in all four hubs. The full seven-year price history is shown in Figure 1.

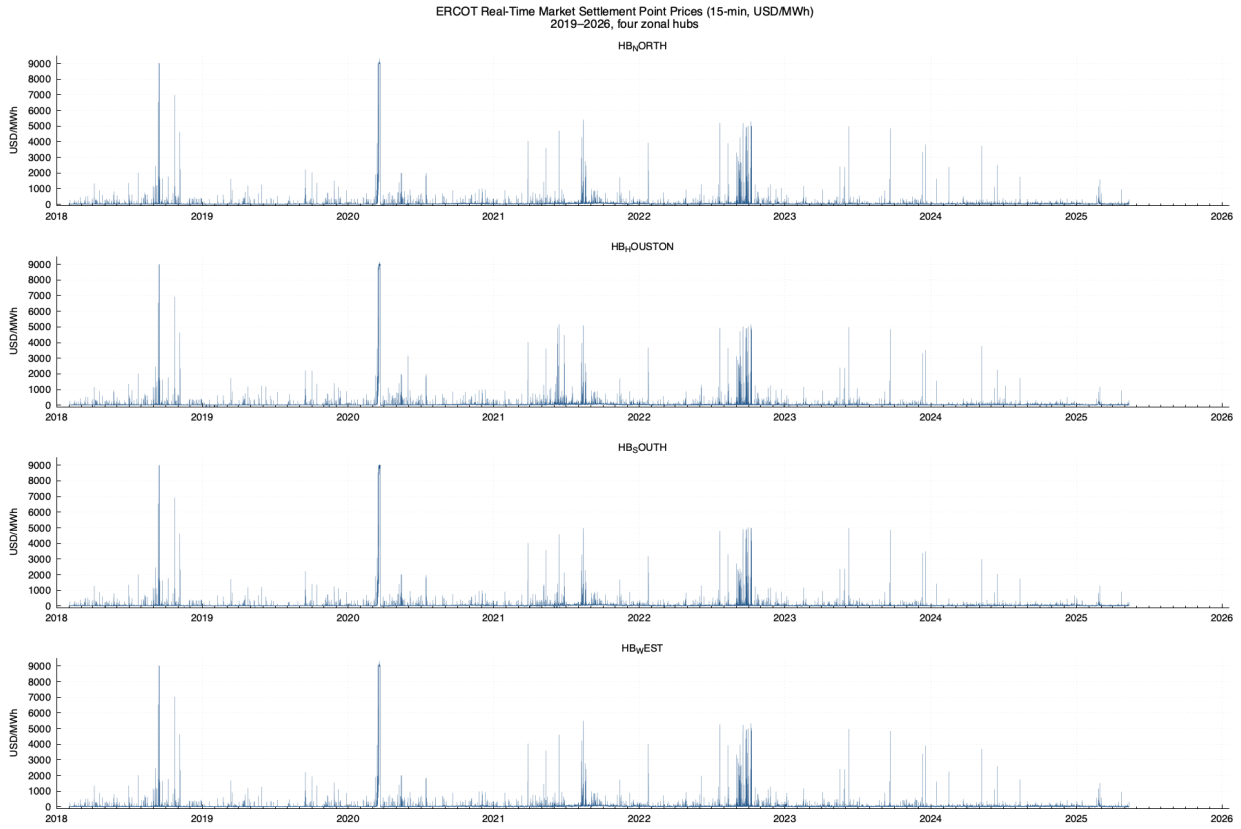

Figure 1: ERCOT real-time settlement prices, 15-minute resolution, 2019–2026. The February 2021 Winter Storm Uri spike dominates all four panels. Scarcity-cap events are visible as flat lines at $\$9{,}000$/MWh prior to 2022 and $\$5{,}000$/MWh thereafter.

2.2 Ingest pipeline

The ERCOT yearly archives ship as `.xlsx` embedded in `.zip`. The pipeline extracts the inner Open-XML workbook with a pure-C streaming reader — no external spreadsheet dependency — and emits one CSV per (hub, year) pair in ISO-8601 UTC. Analysis-time bucketing by season,

time-of-day and day-of-week is performed in Central Prevailing Time (America/Chicago, DST-aware). Daylight-saving fall-back disambiguation uses the published `RepeatedHourFlag` column, corrected for an `mktime` ambiguity-resolution quirk under macOS that otherwise collapses the first occurrence of the duplicated 01:00 hour onto the second.

2.3 Observed characteristics

Four non-obvious characteristics of the processed data inform the analyses that follow:

- **HB_WEST is structurally different.** It records 4–5 times as many negative-price intervals per year as the other three hubs, reflecting the dominant wind-generation exposure of the West zone. Any "all hubs behave alike" assumption is wrong before analysis begins.
- **The cap regime changed mid-window.** Pre-2022 maxima sit at or above $\$9{,}000$/MWh; post-2022 maxima top out at $\$5{,}500$/MWh. Year-over-year spike comparisons are comparisons of two regulatory regimes, not shifts in market dynamics.
- **2022 is a median outlier.** Median price is $\$42$–$\$46$/MWh across all hubs, roughly $2\times$ every other year in the window. Post-Uri reform overlays the European gas-price crisis.
- **Uri is localised.** Feb 14–19 2021 accounts for $\approx 95\%$ of cap hits that year. Event-stripping the top 0.5% of |price| cleanly removes Uri's contribution to global statistics.

2.4 Transformations and event treatment

All analyses are run on *signed-log returns* $r_t = \text{sgn}(p_t)\log(1+|p_t|) - \text{sgn}(p_{t-1})\log(1+|p_{t-1}|)$, which handle ERCOT's negative prices and $\$9{,}000$ cap hits without naive log blow-up [5]. Every test is reported twice: on the full series, and on the series with the top 0.5% of |price| observations stripped. Both numbers go in the tables.

3 Methodology

Three global H estimators, one local H estimator, and two leading-indicator tests.

3.1 Global H

Rescaled-range analysis (H_{RS}). The classical estimator [2]. We report H_{RS} on both the signed-log *price* series (matching the convention of [8], under which the EPEX reference value of $H = 0.88$ was computed) and on signed-log *returns*. The two measure different objects and are complementary.

Detrended fluctuation analysis (H_{DFA}). First-order DFA [3] on signed-log returns. We construct the mean-zeroed cumulative profile $Y(i) = \sum_{k\le i}(r_k - \bar{r})$, split Y into non-overlapping boxes of size s (with backward-pass to cover the tail), linear-detrend each box, and take the RMS residual as $F(s)$. A log-log fit of $F(s)$ over $s \in [16, N/4]$ at 25 log-spaced scales yields H_{DFA} as the slope.

Multifractal spectrum width ($\Delta\alpha$). MFDFA [4] with $q \in [0, 3]$ in steps of 0.25. For each q the generalised fluctuation $F_q(s)$ yields a generalised exponent $h(q)$; applying a Legendre transform gives $\alpha(q) = h(q) + q\,dh/dq$, and $\Delta\alpha = \max_q \alpha(q) - \min_q \alpha(q)$. Negative q is omitted: on fat-tailed signed-log returns a single near-zero box fluctuation drives $F^2(v,s)^{q/2}$ to infinity when $q < 0$ without winsorisation, and the resulting $\Delta\alpha$ is numerically unstable. With $q \ge 0$ the estimate is restricted to the right tail of the singularity spectrum, but remains numerically defensible.

3.2 Local H and the leading-indicator test

For each hub and each of three rolling window sizes $W \in \{96, 192, 672\}$ samples (24 h, 48 h, 7 d), local H_{RS} is computed at hourly stride on signed-log returns. The spike label $y_{t,\Delta}$ for horizon $\Delta \in \{1, 4, 16, 96\}$ samples (15 min, 1 h, 4 h, 24 h) is:

$$y_{t,\Delta} = \mathbf{1}\left[\max_{k=1..\Delta} |p_{t+k}| > P_{95}(|p_{t-2879..t}|)\right],$$

i.e. "does any price in the next Δ samples exceed the 95th percentile of $|p|$ over the trailing 30 days?". Predictive utility is measured as ROC-AUC (equivalently Mann–Whitney U normalised by $n_+ \cdot n_-$) of local $H(t)$ as a score for $y_{t,\Delta}$.

3.3 Independent test: spike-precursor window comparison

As a robustness check, for every spike event $t^\star$ in the seven-year dataset (with a 24-hour refractory period to avoid clusters) we compute local H_{RS} on the 96-sample return window ending at $t^\star$, and compare against ten matched random non-spike windows drawn from the same time range with no overlap with any spike event.

3.4 Correctness gate

Before any ERCOT-derived number is trusted, the analysis binary is run on the bundled 44-day EPEX dataset of [8] and required to reproduce the published arbitrage table exactly. This passed on 2026-04-16 to the published digits (baseline 6.92, MF 9.20, MF advantage $+33.0\%$, oracle overall and discharge agreement rates $42.9\%/44.9\%$ and $11\%/17\%$). Full gate artefact in `results/epex_gate.csv`. The implication is narrow but important: the R/S and DFA code paths are demonstrably correct, so any null finding in what follows reflects ERCOT's behaviour, not a broken estimator.

4 Global Scaling

Table 1 reports the three estimators per hub, with and without event-stripping. The story is compact:

- H_{RS} on *prices* is in the 0.83–0.85 range across all hubs — the price *level* is highly persistent, close to the EPEX reference value of 0.88.
- H_{RS} on *returns* is in the 0.28–0.31 range — returns are strongly mean-reverting.
- H_{DFA} on returns is 0.11–0.15, consistent with mean-reversion via an independent method.
- $\Delta\alpha$ is 0.20–0.26 — modest but non-zero multifractality. Within the range reported for major financial time-series [6].
- Event-stripping moves every metric by no more than ≈ 0.02. The fractal structure is driven by the full distribution of returns, not by the extreme tails.

H_{DFA} and MFDFA $h(q{=}2)$ agree to four decimal places in every cell, an internal consistency check on the MFDFA estimator.

Context, not conclusion: these global numbers characterise ERCOT clearly but do not by themselves answer the predictive question.

Hub	Treatment	H_{RS}^{price}	H_{RS}^{ret}	H_{DFA}	$\Delta\alpha$	$h(q{=}2)$
HB_NORTH	full	0.840 ± 0.007	0.308 ± 0.007	0.119 ± 0.015	0.235 ± 0.026	0.119
HB_NORTH	strip 0.5%	0.845 ± 0.007	0.295 ± 0.006	0.110 ± 0.015	0.247 ± 0.024	0.110
HB_HOUSTON	full	0.839 ± 0.008	0.311 ± 0.007	0.117 ± 0.014	0.241 ± 0.030	0.117
HB_HOUSTON	strip 0.5%	0.844 ± 0.008	0.298 ± 0.006	0.106 ± 0.014	0.257 ± 0.026	0.106
HB_SOUTH	full	0.840 ± 0.009	0.310 ± 0.008	0.122 ± 0.015	0.235 ± 0.025	0.122
HB_SOUTH	strip 0.5%	0.844 ± 0.009	0.297 ± 0.009	0.113 ± 0.015	0.251 ± 0.023	0.113
HB_WEST	full	0.844 ± 0.007	0.288 ± 0.006	0.148 ± 0.017	0.204 ± 0.011	0.148
HB_WEST	strip 0.5%	0.845 ± 0.007	0.283 ± 0.005	0.144 ± 0.018	0.203 ± 0.012	0.144

Table 1: Global scaling estimates per hub, 2019–2025 seven-year series. Standard errors from 8-way non-overlapping sub-sample for H_{RS} and $\Delta\alpha$; OLS-slope SE for H_{DFA}. All return-series estimators use signed-log returns. $h(q{=}2)$ reproduces H_{DFA} to four decimal places, a within-tool consistency check.

5 Regime Segmentation

The 30 cells of (season $\times$ time-of-day $\times$ day-of-week) per hub in CPT local time have $\geq 1{,}000$ samples each. Figure 2 plots H_{DFA} by (season $\times$ time-of-day) for weekdays only, full series.

Three observations matter for the headline experiment to follow:

1. The effect size across cells (H_{DFA} ranging 0.07–0.28 within a single hub) is substantial. **There *is* segmentable structure** in ERCOT's fractal behaviour.
2. The structure is *regime-based*, not spike-based. Winter overnight weekday prices behave differently from summer evening-peak weekday prices; this is a stationary property of the demand/weather/generation mix, not a precursor to scarcity events.
3. HB_WEST deviates from the other three hubs in the same direction as its raw-data profile suggests: higher H_{DFA} in summer cells, reflecting the wind-curtailment pricing regime. Any analysis that assumes zonal homogeneity is wrong a priori.

6 Local Hurst as a Leading Indicator

This section carries the paper's headline finding.

6.1 ROC-AUC across windows and horizons

Figure 3 plots ROC-AUC for local H_{RS} predicting $y_{t,\Delta}$ across every (hub, window, horizon) combination. All 48 cells report an AUC between 0.425 and 0.502.

Two observations from the heatmap shape:

- **Longer windows push AUC back toward 0.5.** The 7 d window is consistently closer to random than the 24 h or 48 h window. A longer window dilutes any transient precursor signal; the fact that AUC approaches 0.5 from below as W grows is consistent with an extremely weak signal being averaged out.
- **Longer horizons push AUC back toward 0.5.** At $\Delta = 24$ h almost any interesting event will occur somewhere in the window, so the binary label approaches 1 for nearly all samples. Short-horizon cells ($\Delta = 15$ min, 1 h) are where a real precursor signal should appear; instead they are the cells furthest below 0.5.

The direction ($H < 0.5 \Rightarrow$ spike more likely in the near future) is consistent with the structural interpretation: a strongly mean-reverting regime is one in which the system is "tight", where any

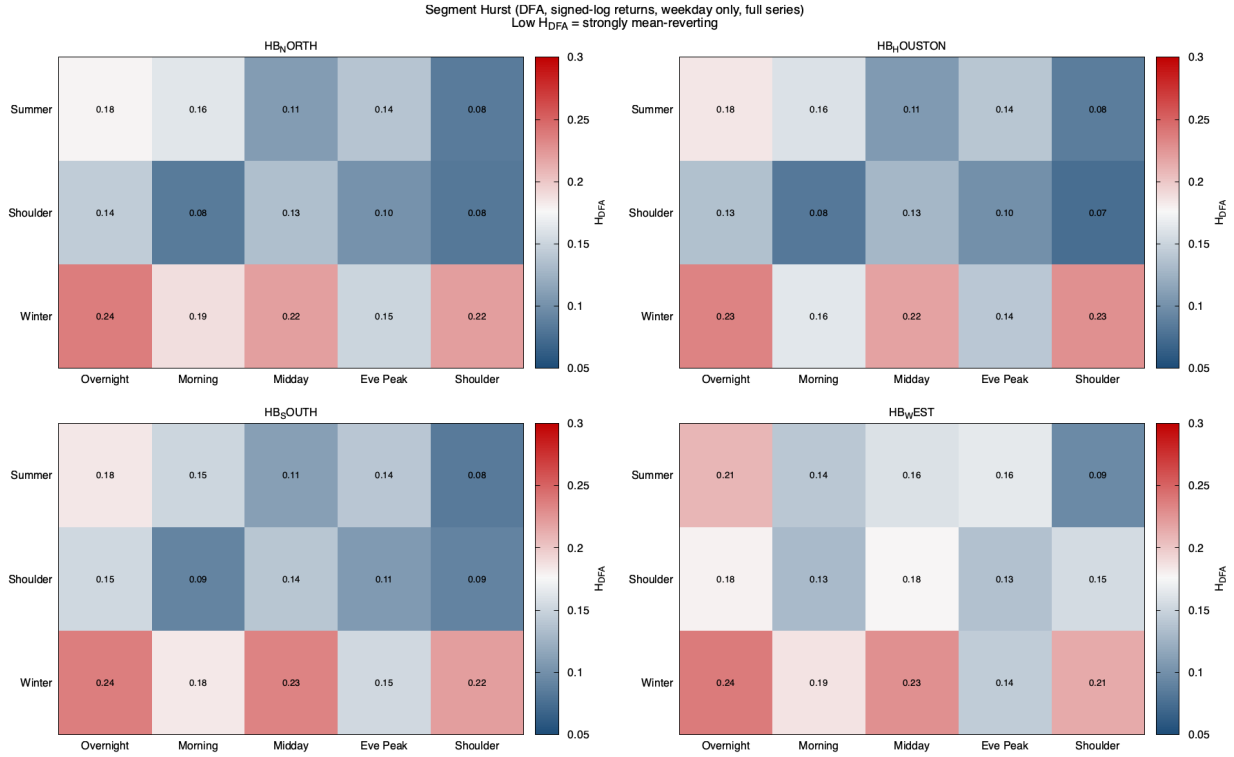

Figure 2: Segmented H_{DFA} on signed-log returns. Winter overnight cells run with the highest H_{DFA} (≈ 0.22–0.24 — still strongly mean-reverting but with more memory than other season/ToD cells). Shoulder-season slots and evening-peak summer cells are the most strongly mean-reverting ($H_{DFA} \approx 0.07$–0.11). HB_WEST's summer profile is visibly different from the other three hubs, consistent with its wind-dominated generation mix.

disturbance produces a snap reaction. But the effect size is tiny — maximum ≈ 0.075 AUC away from random — far below the level required for operational deployment.

6.2 A visual check: Uri 2021

Figure 4 overlays the local H_{RS} (24 h and 7 d windows) on the price trajectory through the Uri week at HB_NORTH. If H predicted Uri, we would expect a clear rise of H *before* the price spike begins on February 15. Instead H_{24h} drops during the spike period, and H_{7d} drifts only marginally. Visually consistent with the AUC finding.

7 Spike-Precursor Window Comparison

As a non-ROC robustness check, we compare the distribution of local H_{RS} on the 24-hour return window ending at each spike event against the distribution on matched random non-spike windows. Summary in Table 2; full distributions in Figure 5.

The two tests agree: local H_{RS} is not a useful near-term spike predictor on ERCOT.

8 Honest Limits

- **Uri and comparable events are exogenous shocks.** A frozen gas pipeline is not a fractal event. No Hurst-based indicator can anticipate the physical failure cascade of a generator

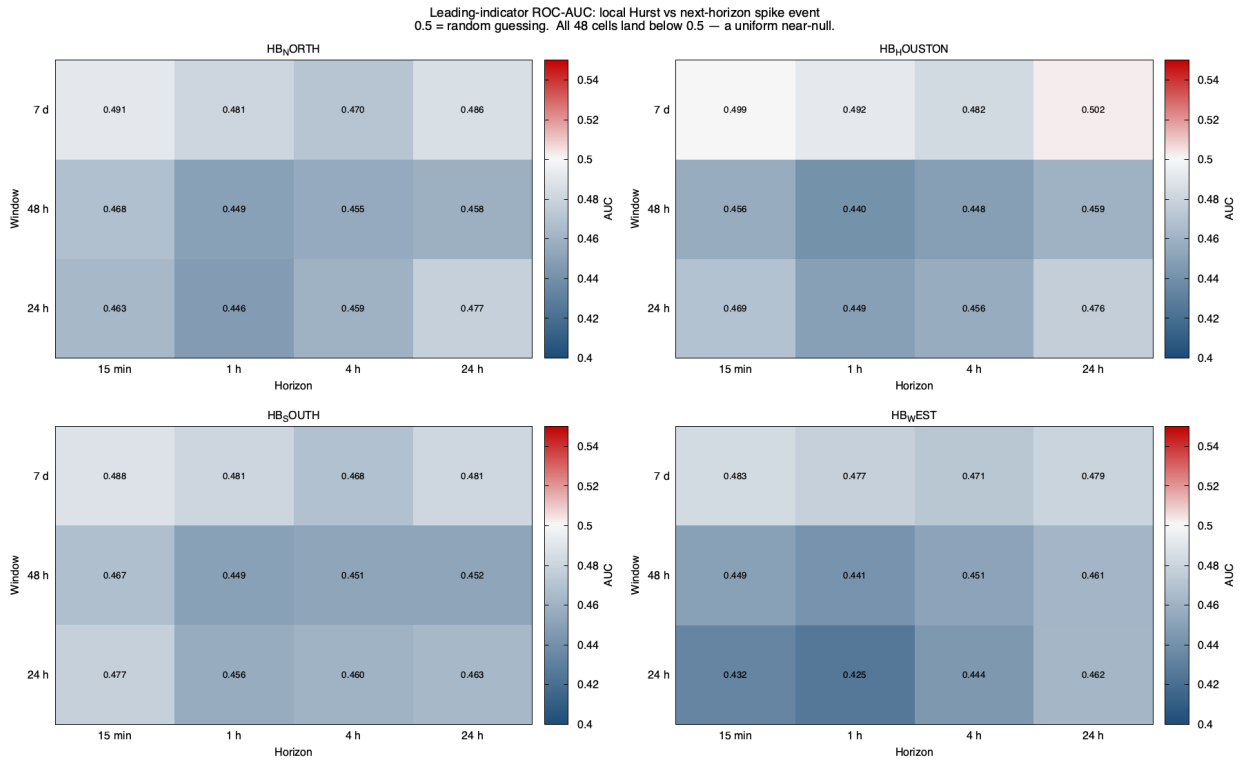

Figure 3: ROC-AUC by window size (y) and forecast horizon (x), faceted by hub. 0.5 is random guessing. Every cell lands at or below random; the direction is uniform and the effect size is modest but consistent. HB_WEST 24 h window / 15 min horizon is the worst cell at 0.425; HB_HOUSTON 7 d window / 24 h horizon is the single cell at random (0.502). Pearson ρ between local H and next-horizon $|p|$ is negative in 47 of 48 cells (range -0.07 to -0.00).

fleet; this paper does not claim otherwise. The test here is whether *endogenous* price dynamics shift fractally before a scarcity event; they do not.

- **MFDFA $\Delta\alpha$ is computed on $q \geq 0$ only.** Negative q is numerically unstable on fat-tailed signed-log returns without pre-winsorisation, and we declined to winsorise since the extreme tails are the object of study. The consequence: $\Delta\alpha$ captures the right tail of the singularity spectrum only; the full spectrum width could be larger. A symmetric-q MFDFA with appropriate winsorisation or an alternative estimator (Wavelet Transform Modulus Maxima [7]) is a defensible follow-on.
- **Regulatory regime change.** The System-Wide Offer Cap was $\$9{,}000$/MWh through 2021 and $\$5{,}000$/MWh from 2022. Comparing spike frequencies across the two sub-periods

Hub	n_{spike}	$\overline{H}_{\text{spike}}$	n_{nonspike}	$\overline{H}_{\text{nonspike}}$	$\Delta\overline{H}$	$\Delta\overline{H}_{\text{stripped}}$
HB_NORTH	1284	0.688	12840	0.699	-0.011	-0.006
HB_HOUSTON	1290	0.679	12900	0.688	-0.009	-0.002
HB_SOUTH	1272	0.673	12720	0.690	-0.017	$+0.000$
HB_WEST	1229	0.690	12290	0.712	-0.022	-0.001

Table 2: Mean local H_{RS} on 24 h windows ending at spike events vs matched non-spike windows. Within-class SD is ≈ 0.11 in all cases, so $\Delta\overline{H}$ on the order of -0.01 to -0.02 is within one standard deviation of the class variance and has trivial operational discriminating power. Event-stripping the top 0.5% removes the already-small difference almost entirely, suggesting what little signal exists is driven by the tail events themselves rather than by fractal structure in the pre-event window.

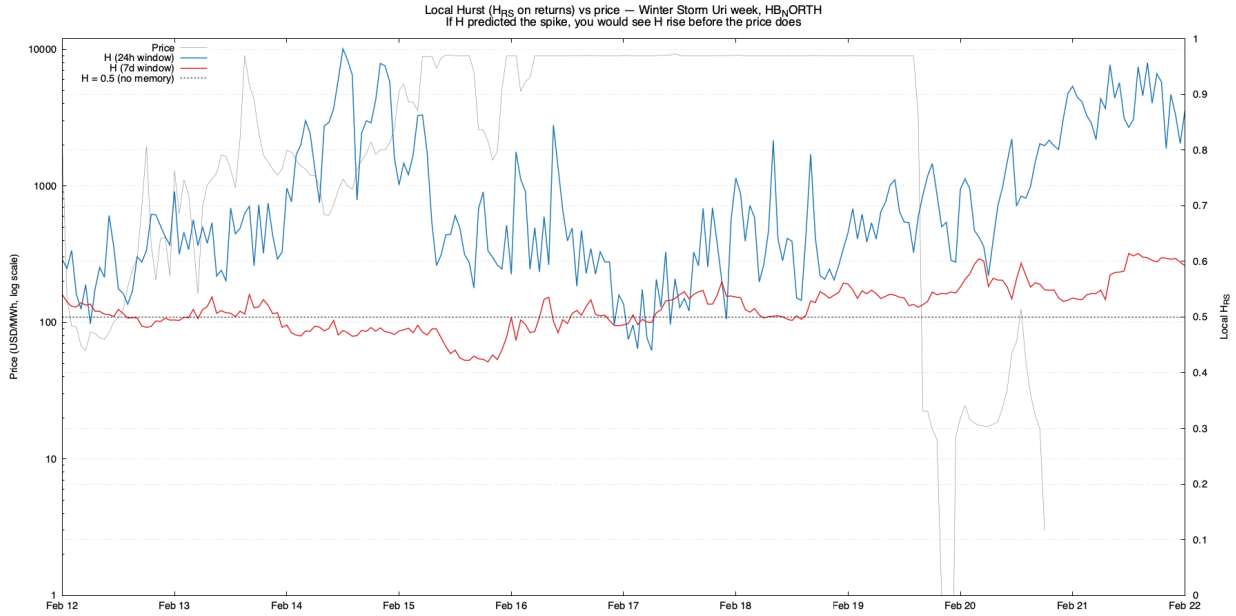

Figure 4: Local Hurst (H_{RS}, 24 h and 7 d windows) overlaid on price for the Uri week at HB_NORTH. The price axis is logarithmic to accommodate the scarcity-cap spike. Uri is not anticipated by a rising H; the 24 h window instead falls during the peak as the return series becomes dominated by a single sustained direction.

is a comparison of regulatory regimes, not market shifts. Segmentation results (§5) are computed on the pooled seven-year window; running pre- and post-2022 separately yields the same qualitative picture (low-H_{DFA} returns, modest $\Delta\alpha$) at different quantitative points, but no new predictive power.

- **Statistical vs. operational significance.** With $n \sim 10^6$ samples the 48 AUCs differ significantly from 0.5 in the formal hypothesis-test sense. They do not differ from 0.5 in a way that would support deployment. This is a standard trap in large-n analyses that we name explicitly rather than bury.
- **Single-hub data only.** The analysis uses hub-level aggregates. Node-level (resource-node) prices are more volatile and may yield different results; that was deemed out of scope for v1 because zonal hubs are what a retail provider or VPP operator actually interacts with.
- **Causal analysis only.** No look-ahead or oracle features. The AUC is what an actual real-time deployer could have achieved, not an upper bound. Implementations that expose modest look-ahead (day-ahead prices, load forecasts) might raise AUC, but would be a different study.

9 Conclusions

On seven years of ERCOT 15-minute settlement prices across four hubs, the Hurst exponent — computed via three independent methods globally, segmented by regime, and rolling over 24-hour to 7-day windows — is not a useful standalone predictor of near-term price spikes. Specifically:

- 48 (hub, window, horizon) AUCs all between 0.43 and 0.50, uniformly below random.
- Pre-spike vs non-spike mean local H differs by at most 0.022 on a within-class SD of 0.11.
- Both tests direction-agree: low H (mean-reverting regime) slightly precedes spikes more

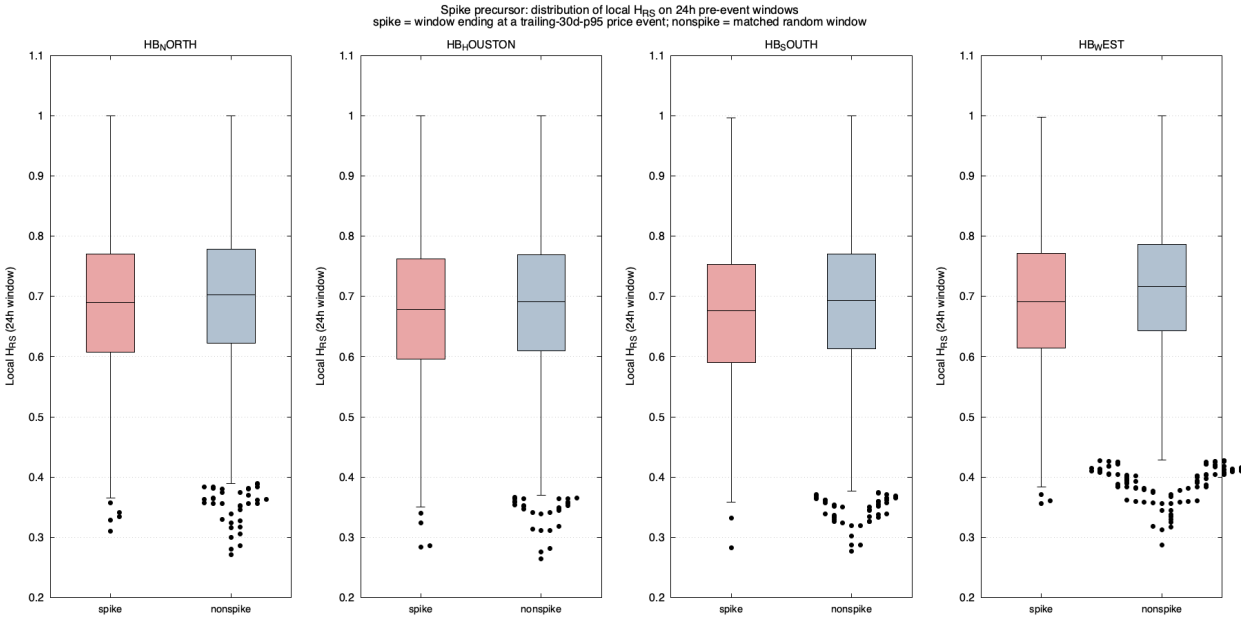

Figure 5: Box plots of local H_{RS} on 24 h pre-event windows. Red: windows ending at spike events. Blue: matched random non-spike windows. The two distributions overlap in essentially their entirety across all four hubs.

often than high H, but the effect is too small to deploy.

- The toolchain reproduces the EPEX-market reference whitepaper to the published digits, so the null finding is not an implementation artefact.

We are explicit about what this does not say: it does not say the Hurst exponent is useless on ERCOT. The segmentation result (§5) is a real positive finding about ERCOT's regime-stratified fractal behaviour. What we rule out is the specific operational claim that *rolling local H gives leading-indicator information for scarcity-cap spikes*. On ERCOT, over this data, it does not.

Directions for follow-on work

1. **Wavelet decomposition.** A multi-resolution analysis of the ERCOT price signal can separate diurnal, weekly, and event-frequency components; the sibling `WaveletCommitment` project [9] has the toolchain.
2. **Change-point neural detectors.** Recurrent and transformer architectures trained on recent windows may capture non-fractal precursors the Hurst exponent misses.
3. **Multi-factor models.** Local H combined with load-forecast error, net-load volatility, and temperature-anomaly features may earn incremental AUC above any single factor alone. The negative result here is the baseline against which such a multivariate effort would be measured.
4. **Full-spectrum $\Delta\alpha$.** Symmetric-q MFDFA with principled winsorisation, or a WTMM-based estimator, would answer whether the full multifractal shape (not just the right tail) shifts pre-spike.

The infrastructure built for this study — ingest, sanity, global and segmented estimators, local-H rolling, spike-precursor comparator, ROC-AUC driver — is compact (roughly 1,500 lines of C on top of the sibling toolkit), reproducible, and readily extended along any of the above lines.

References

[1] H.E. Hurst. Long-term storage capacity of reservoirs. *Transactions of the American Society of Civil Engineers*, 116:770–799, 1951.

[2] B.B. Mandelbrot and J.R. Wallis. Robustness of the rescaled range R/S in the measurement of noncyclic long run statistical dependence. *Water Resources Research*, 5(5):967–988, 1969.

[3] C.-K. Peng, S.V. Buldyrev, S. Havlin, M. Simons, H.E. Stanley and A.L. Goldberger. Mosaic organization of DNA nucleotides. *Physical Review E*, 49(2):1685–1689, 1994.

[4] J.W. Kantelhardt, S.A. Zschiegner, E. Koscielny-Bunde, S. Havlin, A. Bunde and H.E. Stanley. Multifractal detrended fluctuation analysis of nonstationary time series. *Physica A: Statistical Mechanics and its Applications*, 316(1–4):87–114, 2002.

[5] R. Weron. *Modeling and Forecasting Electricity Loads and Prices: A Statistical Approach*. Wiley, 2006.

[6] P. Norouzzadeh and G.R. Jafari. Application of multifractal measures to Tehran price index. *Physica A*, 356(2):609–627, 2005.

[7] J.F. Muzy, E. Bacry and A. Arneodo. Multifractal formalism for fractal signals: the structure-function approach versus the wavelet-transform modulus-maxima method. *Physical Review E*, 47(2):875–884, 1993.

[8] S. Woolley. Multifractal battery arbitrage on European day-ahead electricity markets. *Project whitepaper*, `MultifractorArbitrage` sibling repository, 2026.

[9] S. Woolley. Wavelet decomposition for unit-commitment forecasting. *Project whitepaper*, `WaveletCommitment` sibling repository, 2026.

[10] Electric Reliability Council of Texas. Real-Time Market Settlement Point Prices, Report ID NP6-785-ER. `https://www.ercot.com/mp/data-products/data-product-details?id=NP6-785-ER`, retrieved 2026-04-16.