

Day-Ahead vs Real-Time Price Relationship Integrity

A Reproducible Measurement Method

Delta Motion Analytics — Methods & Reproducibility Whitepaper
Version 1.0 · September 2026

1. What this document is (and is not)

This is a methods-transparency document. Its purpose is to let a technically-fluent reader understand how we measure the integrity of the day-ahead vs real-time (DA-RT) electricity price relationship — the exact sources, windows, statistics, and null construction — well enough to judge the work, without exposing the engine's source code.

It is **not** a sales sheet, a price or spread forecast, a trading signal, a securities or commodity research report, or a recommendation. It makes no directional call and no claim about future prices, spreads, or returns. It is **not** an Independent Market Monitor product and makes no claim to a regulated monitoring role. It is **not** an accusation: a systematic day-ahead premium is a market-structure feature, not evidence of manipulation or misconduct by any participant, venue, or ISO. It measures how a market-structure relationship behaves; **tradability, cost, risk, and market impact were not measured and are not claimed, in either direction.**

2. The problem we measure

A trading desk, utility, or gentailer builds models and hedges on the DA-RT relationship; a data or benchmark provider publishes analytics built on it. Each must be able to demonstrate — to an auditor, a board, a model-risk function, or an institutional client — that the relationship, and the models built on it, are behaving and calibrated. Independent, reproducible evidence of this is hard to obtain: internal QA is not independent; the ISO's Independent Market Monitor reports on the market as a whole, not on your book or model, and its standing is regulatory rather than a second opinion you can commission; data vendors sell the feed and analytics, not an integrity attestation of the relationship; and bespoke consultancy opinions are not built for push-button external reproduction.

The position this document occupies: an **independent, reproducible, method-transparent measurement** of DA-RT relationship integrity, on free public ISO data, with honest reporting against a proper null — a second opinion / validation layer that *complements* the market monitor and never claims its role.

3. What we measure

1. **DA-RT divergence** — the behaviour of the gap between day-ahead and real-time prices over defined windows.
2. **The systematic day-ahead premium** — the average amount by which day-ahead prices sit above real-time, reported as a measured market-structure property, **consistent with a forward risk premium** (hedging pressure, forecast-error and weather uncertainty), and **not attributed** to inefficiency, gaming, or any participant's conduct.
3. **Serial dependence (autocorrelation) of the premium and its persistence** — reported strictly as a **calibration/health property** of the DA-RT relationship: how the relationship behaves over time.

4. **Null-validated throughout** — every claim is tested against what a proper null produces. Tradability, profitability after transaction costs, risk, and market impact are **not measured and are claimed in neither direction**.

4. Method transparency (the reproducibility core)

Method is disclosed faithfully; no engine source code is exposed.

4.1 Data sources. Free public ISO data: **NYISO** (New York City zone, with the Capital zone as a within-grid check) and **CAISO** (the **SP15** and **NP15** hubs). Single-provider-derived analysis, not redistributed raw data.

4.2 Sampling and windows. NYISO: six complete monthly windows, February-July 2026 (winter through summer). **CAISO: July 2026, two hubs (SP15 and NP15).** The series is **deseasonalized before analysis** — hour-of-day and weekday/weekend structure stripped — with **negative prices and price spikes kept** (not winsorized away), so the measured structure is the genuine residual, not an artifact of cleaning.

4.3 The statistics.

- **Day-ahead premium:** the median real-time-minus-day-ahead settlement, in \$/MWh and as a percentage of the day-ahead price, with a bootstrap 95% confidence interval per window. (A negative RT–DA is a positive day-ahead premium: real-time settles below day-ahead.)
- **Serial dependence:** the lag-1 autocorrelation of the deseasonalized residual, with a **high-pass that removes all weekly-and-slower structure applied before the autocorrelation**, so smooth low-frequency structure cannot masquerade as hour-to-hour dependence.

4.4 Null construction and decision rule. A **value-permutation null**: the residual's values are reshuffled i.i.d., destroying serial dependence while preserving the value distribution. (This — not an increment-shuffle random-walk surrogate — is the correct null for a stationary, deseasonalized power-market residual.) The decision bar is pre-registered and fixed before the windows are analysed.

4.5 Artifact guard. A stale / outlier / low-information guard is applied; low-information periods are masked and the test re-run. Negative prices and spikes are retained as genuine market events, not masked.

4.6 Independent reproduction — two distinct senses, kept separate.

- **(a) Cross-grid replication (external validity).** The instrument, run unchanged on a second, independent grid, produced the same qualitative result — see §5.4. This is replication across grids, not adversarial hardening.
- **(b) Different-generative-family reproduction (adversarial hardening).** The result also survived reproduction against a genuinely different family of synthetic generators — stochastic-volatility clustering, Student-t tails, and power-price spikes and negatives — importing the fixed core and the published harness unchanged. Both the premium and the serial dependence survived. This is distinct from, and additional to, the cross-grid replication in (a).

5. Worked demonstration — the DA-RT relationship on NYISO and CAISO (a completed, reproduced worked study)

5.1 Setup. NYISO New York City zone, six monthly windows Feb-Jul 2026 (Capital zone as a within-grid check); CAISO SP15 and NP15 for July 2026; deseasonalized residual;

value-permutation null; high-pass applied before the autocorrelation.

5.2 The day-ahead premium (NYISO). Real-time settled **below** day-ahead — a systematic **positive day-ahead premium** — in **all six windows**; the bootstrap 95% confidence interval excludes zero in every one. The magnitude ranges from about **−2.7% (April) to −6.9% (March)** of the day-ahead price.

Month	RT–DA median (\$/MWh)	95% CI	% of DA	lag-1 ACF	z vs null
Feb 2026	−4.28	[−5.57, −2.60]	−5.2%	+0.65	+17.8
Mar 2026	−3.14	[−4.13, −2.21]	−6.9%	+0.35	+10.3
Apr 2026	−1.03	[−1.39, −0.47]	−2.7%	+0.35	+9.6
May 2026	−2.00	[−3.06, −1.29]	−5.3%	+0.36	+9.9
Jun 2026	−1.80	[−2.66, −1.23]	−4.0%	+0.41	+11.9
Jul 2026	−2.00	[−3.03, −0.91]	−3.7%	+0.50	+15.2

The premium is reported as a measured market-structure property, **consistent with a forward risk premium** — not attributed to inefficiency, gaming, or any participant's conduct.

5.3 Serial dependence (the calibration/health property). The deseasonalized premium exhibits **measurable serial dependence (lag-1 autocorrelation $\approx$ 0.35–0.65 every window, z +9.6 to +17.8 against the value-permutation null)** that persists beyond an i.i.d. reshuffle of its values and **survives a high-pass removing all weekly-and-slower structure in all six windows, with the autocorrelation essentially unchanged** — indicating genuine hour-to-hour forecast-error persistence rather than seasonality or slow drift. This is reported strictly as a **calibration/health property** of the DA–RT relationship — how the relationship behaves over time. **Tradability, profitability after transaction costs, risk, and market impact were not measured and are not claimed, in either direction. No forecast, no directional read, no period-profitability ranking, no bid-shaped output.**

5.4 Cross-grid replication (CAISO). The same qualitative result — a positive day-ahead premium with hour-to-hour serial dependence — appears on a second, independent grid: **CAISO, July 2026, on two hubs. SP15 carried a premium of roughly −9% of the day-ahead price with lag-1 autocorrelation of about 0.30; NP15, the replication hub, showed a smaller premium (about −4%) with stronger persistence (about +0.70). The persistence survives the weekly-and-slower high-pass on both hubs.** This is a **directional cross-grid agreement, not a pooled or meta-analyzed statistic.** The two grids differ: NYISO is milder in percentage terms ($\approx$ −3% to −7%) with somewhat higher persistence ($\approx$ 0.35–0.65) than CAISO ($\approx$ −9%, $\approx$ 0.30). An import-constrained-NYC explanation for the difference is an offered hypothesis, not something this study adjusts for. Note also the asymmetric coverage: NYISO carries the six-month temporal axis; CAISO is a single month across two hubs.

5.5 Different-family reproduction. As §4.6(b): both the premium and the serial dependence survived reproduction against a different family of synthetic generators (stochastic-volatility clustering, Student-t tails, power-price spikes and negatives), core and published harness unchanged.

6. Honest reporting against the null

Every claim here is tested against a proper null (§4.4). The value-permutation null could have returned "no serial dependence beyond chance," in which case that is what we would report; here it returned strong, consistent dependence, and that is reported honestly with its statistics. A method that only ever "finds something" is unfalsifiable; testing against a null that *can* return nothing — and reporting it when it does, across the wider body of work — is the evidence that a positive finding is real and not manufactured.

7. What this is not

- **Not an Independent Market Monitor** — we hold no regulated monitoring mandate; we are an independent second opinion / validation layer that complements the monitor.
- **Not a data vendor** — we do not sell the feed; we measure whether the relationship and the models built on it are genuine and calibrated.
- **Not black-box surveillance** — we pre-register, reproduce on free public data, and report against a proper null; the method is auditable.
- **Not a forecast, signal, or recommendation** — no directional read, no spread or premium forecast, no period-profitability ranking, no bid-shaped output, no "alpha."
- **Not an accusation** — a systematic day-ahead premium is a market-structure feature consistent with a forward risk premium; nothing here implies misconduct by any named participant, venue, or ISO.

8. How an engagement works

1. **Free first read** — a calibration/integrity snapshot on a subset of one market's public DA/RT data, at no cost and no commitment, so your team sees the method live.
2. **Integrity / validation assessment** — a fixed-scope assessment on defined market(s) and windows, delivered as a report formatted for a model-validation file, with a per-engagement reproducibility appendix (sources, windows, statistics, null — no engine code).
3. **Monitoring (optional, later)** — a recurring integrity/calibration re-run once an assessment has landed.

Deeper-scope engagements (additional grids, additional integrity diagnostics) use client-supplied or engagement-specific data, procured only for that engagement.

9. Scope and limitations

1. **Tradability, cost, risk, and market impact were not measured** — no exploitability claim is made in either direction. (This study ran no cost-band clearing; the "sub-cost / not tradable" claim available in other domains is deliberately *not* made here.)
2. The day-ahead premium is **consistent with a forward risk premium** and is **not attributed** to inefficiency, gaming, or any participant's conduct.
3. **Attribution is weather / forecast-error confounded** — what drives the premium is not isolated; the finding is the presence and structure of the premium, not its cause.
4. **Six recent months (Feb-Jul 2026), winter through summer — not a year-round claim.**
5. **One grid plus a correlated companion zone, not a battery of independent grids.** NYISO's Capital zone shares the system energy component with NYC and is a within-grid corroboration, not an independent test. The two genuine axes are temporal (six NYISO months) and cross-grid (NYISO vs CAISO).

6. **CAISO is a single month (July), two hubs (SP15, NP15)** — the cross-grid result is a directional match, not a multi-month CAISO claim and not a pooled cross-grid statistic.
7. This is **exploratory, single-provider-derived analysis of public ISO data** — measurement, not investment or trading advice.
8. Serial dependence is a **statistical property of the historical residual**, reported as calibration/health — it is not a forecast.
9. **Not an Independent Market Monitor function; not an accusation** of manipulation against any participant.

10. Roadmap — not demonstrated in this document

Additional grids, additional months, and additional integrity diagnostics are natural extensions, each measurement-framed and claimed only once built and independently reproduced. Nothing in this document claims them.

Delta Motion Analytics is a cross-border analytics studio providing measurement and reproducibility infrastructure. This document describes method, not source code, and does not constitute investment or trading advice or a recommendation regarding any security, commodity, or asset.