



COMPOUNDING ENERGY LTD

Working Paper Series • CE-NOTE-2026-02

The Graded Record

Pre-Registered Backtests of the CECadence GB BESS Revenue Model, Including the Failures

Dr Christopher T. M. Clack

Founder, Compounding Energy Ltd

christopher@compoundingenergy.com

September 2026

Abstract. Every working paper in this series closes on the same claim: trust the record, not the mechanism. This note publishes that record in one place. The CECadence GB battery revenue model is graded against settled outturn under a protocol whose every rule is frozen in dated commits before the data it will be judged against is fetched, and whose results are published whatever they say. Across the five graded windows from 2022 to the first half of 2026 the annual-mean gate is met in three; the pre-registered held-out year and the held-out half-year both fail on the scarcity tail and are published exactly as they landed. Five pre-registered attempts to re-shape that tail were fitted and rejected under their own frozen adoption rules. We set out the protocol, the table, the diagnosis, what a financing should and should not take from it, and how any reader can check it.

Keywords: pre-registration, backtest, validation record, battery storage, Great Britain, CEC-

adence

Contents

1	Why publish the failures	2
2	The protocol, as publicly committed	2
3	The record	2
4	What the record says	2
5	What a financing should take from it	3
6	How to check it	3

1 Why publish the failures

A revenue model sold to lenders is judged, in the end, by one question: what happened when it was tested on data its authors had not seen? Most vendors answer with the windows that passed. This note answers with all of them. The CECadence model that underpins [Clack \(2026a\)](#), [Clack \(2026b\)](#) and [Clack \(2026c\)](#) is graded in public against the open-data CECadence Quarterly GB BESS Index ([Compounding Energy Ltd, 2026c](#)) and against settled Elexon outturn, under a protocol committed before each test ([Compounding Energy Ltd, 2026b](#)). The record it has produced so far is mixed — and that is the point. A curated record cannot be distinguished from a lucky one; a record that publishes its own failures, under rules that were frozen before the outcomes were known, can.

2 The protocol, as publicly committed

Four commitments define the grade ([Compounding Energy Ltd, 2026b,a](#)). Market references are as of September 2026.

1. **Rules before data.** Every adoption rule is frozen in committed code before any held-out input is fetched; the dated git history proves the gap for each window. The commitment covers the grading convention (a 26-week representative horizon for the full-year windows, halved with the window to 13 weeks for the 2026 half-year), the input-coverage gate, and the pass thresholds.
2. **Three legs, three gates.** Each window is graded on the annual mean, the median (P50) and the scarcity tail (P95) of the revenue distribution, against pre-registered thresholds of 5%, 10% and 20% respectively. A window passes only if every leg passes.
3. **Two bases.** Windows are graded on a *synthetic* basis (the model's own stylised price shape) and, where settled data permit, on a *realised* basis (settled outturn).
4. **Publish whatever it says.** Results ship unmodified. Rejected model changes ship too. A cell that cannot be completed at first execution — as the 2026 realised leg could not, because input coverage was 90.3% against a frozen 95% gate — is published as a gap, and completed only when the gate is crossed by the frozen pipeline (it was, in July 2026).

3 The record

Table 1 reproduces the public record in full. The 2022–2025 cells are in-sample for the level calibration and graded at the frozen 26-week convention; 2025 was the pre-registered held-out year and 2026H1 the pre-registered held-out half-year.

Read across the rows, three facts stand out. The annual mean is within its gate in three of the five windows (2022, 2025 on the realised basis, 2026H1 on the synthetic basis); the 90% confidence interval on the 2026H1 synthetic mean is published alongside it so that the pass is not oversold. The median is within its gate in the in-sample years and misses in both held-out windows. The scarcity tail fails in every window after 2022.

4 What the record says

The public diagnosis is *level right, tail wrong*. The realised-2025 mean is the best of any backtest year; the 2026H1 synthetic mean is the record's first held-out mean pass. The error sits in the shape of the tail: 2025 was a moderate-gas, low-scarcity year whose realised price distribution was flat where the frozen stylised scarcity representation, calibrated on the one genuine scarcity

Table 1: The CECadence graded backtest record, 2022–2026H1, as published (Compounding Energy Ltd, 2026a). Errors are absolute percentage errors of the modelled revenue distribution against outturn on each leg; the pre-registered adoption thresholds are mean $\leq 5\%$, P50 $\leq 10\%$, P95 $\leq 20\%$.

Window	Basis	Mean	P50	P95	Status
2022	synthetic	2.8%	3.7%	7.0%	in-sample; all legs pass (crisis year)
2023	synthetic	16.8%	1.2%	44.5%	in-sample; P95 hot (scarcity knee)
2024	synthetic	10.4%	10.0%	55.8%	in-sample; mean and P95 hot
2025	synthetic	12.0%	8.2%	59.4%	held-out — FAIL on P95
2025	realised	1.6%	24.0%	55.4%	held-out — FAIL on P50/P95; best mean of any year
2026H1	synthetic	1.4%	23.0%	55.6%	held-out — FAIL on P50/P95; first held-out mean pass
2026H1	realised	7.8%	26.3%	54.5%	held-out — FAIL on all three legs

year in the public record, expects a steep one. The 2026 half-year graded exactly that diagnosis out of sample.

The protocol’s response to a chronic tail failure is not to tune it away. Five re-shape attempts were pre-registered, fitted, and rejected under their own frozen adoption rules — including one that closed the 2022 fit to well under one percent but did so by pinning a parameter at a bound in every year, which the rules treat as overfitting rather than as a fix. With only one genuine scarcity year in the public record there is not enough tail data to fit a second shape honestly; the registered unlock is a second genuine high-scarcity window. Until it arrives the tail failure stays published.

5 What a financing should take from it

The record supports a specific reading and no more. The central case — mean and median — is the planning number, and on the realised basis it has been right to within a few percent in the windows where it could be graded. The P95 upside is not bankable on this record and should not be leaned on; the public guidance is to size downside on the P90 exceedance band and to test debt service against it. Nothing in this note changes a revenue level: the graded quantity is the model’s error, not the market’s outcome, and the Index (Compounding Energy Ltd, 2026c) remains the measured baseline against which levels are read.

6 How to check it

Every number in Table 1 is on the public record page (Compounding Energy Ltd, 2026a), with the protocol on its companion page (Compounding Energy Ltd, 2026b); the commits that froze each rule carry their own dates and content hashes, which cannot be edited after the fact. The 2022–2025 cells can be re-run live from the in-product backtest form on a paid plan at the same 26-week convention the committed cells use; free verification runs sample a shorter window and are labelled as such. The 2026H1 cells grade at a 13-week horizon and are not reachable from that form, which takes full settlement years only. This note will be re-issued when the next pre-registered window closes, whatever it says.

References

Christopher T. M. Clack. Cannibalisation as a fixed point: revenue forecasting for battery energy storage in liberalised electricity markets. Working Paper CE-WP-2026-01, Compounding Energy Ltd, 2026a. URL https://www.compoundingenergy.com/library/CE-WP-2026-01_Cannibalisation_Fixed_Point.pdf.

- Christopher T. M. Clack. Co-optimising the GB BESS revenue stack: wholesale energy and the NESO dynamic-response suite at the cannibalisation equilibrium. Working Paper CE-WP-2026-02, Compounding Energy Ltd, 2026b. URL https://www.compoundingenergy.com/library/CE-WP-2026-02_GB_BESS_Revenue_Stack.pdf.
- Christopher T. M. Clack. The cannibalisation of flexibility: Data-centre demand response and the equilibrium value of grid flexibility in Great Britain. Technical Report CE-WP-2026-05, Compounding Energy Ltd, 2026c. URL https://www.compoundingenergy.com/library/CE-WP-2026-05_Cannibalisation_of_Flexibility.pdf.
- Compounding Energy Ltd. CECadence — held-out validation: the 2025 + 2026H1 record, in full, 2026a. URL <https://cadence.compoundingenergy.com/heldout.html>. Public record page; retrieved September 2026.
- Compounding Energy Ltd. CECadence — pre-registration protocol summary, 2026b. URL <https://cadence.compoundingenergy.com/heldout-protocol.html>. Public protocol summary; retrieved September 2026.
- Compounding Energy Ltd. CECadence quarterly GB BESS index, 2025Q3–2026Q2. Technical report, Compounding Energy Ltd, 2026c. URL <https://www.compoundingenergy.com/gb-bess-index-2026q2.html>. Released 22 July 2026; computed from Elexon settlement cashflows and NESO auction records for the full 168-BMU battery registry.