How much power did an AI cluster draw at the grid peak? A two-date reconstruction of the Phoenix field trial

Lluis Eriksson

19 November 2026

Abstract

The Phoenix field trial showed how a 256-GPU cluster can change its power draw. Assessing the coincidence of that response with grid demand requires an absolute date. The authors' figure workbook and standalone CSV contain identical sequences of 400 power values, but all timestamps differ by one day. We carry both dates through an hourly join to US Energy Information Administration (EIA) data. During the daily peak hour of the Salt River Project balancing authority, mean cluster power is 75.8% of a common early reference under the workbook date and 99.0% under the CSV date. Changing the minute endpoint convention gives 75.3% and 99.0%. Calendar choice changes an inference about grid coincidence while preserving the experimental trace. The physical date remains unresolved by the sources examined.

1 Two questions about one trial

Colangelo et al. report that they reduced power by 25% for three hours in an AI-cluster field trial while meeting workload performance requirements [1, 2]. We ask how much power the cluster drew during the hour of highest reported demand on the regional grid. This question concerns the daily maximum of an hourly series. A building's total metered demand, an annual capacity obligation and the counterfactual load without dispatch each require additional observations.

The authors place the Salt River Project (SRP) event on 3 May 2025 in the preprint. A May 2025 presentation by coauthor Ayse Coskun labels the Phoenix SRP demonstration 4 May [3]. Their common authorship matters when interpreting these statements. Neither document supplies an independently timestamped dispatch acknowledgement.

2 Pairing the records

We fixed the authors' repository at commit `34403aef` [4]. The Figure 2 workbook supplies timestamps and total power on its `SRP_granular` sheet. Its companion CSV supplies the same two fields. Every paired power value is numerically identical in the two exports. Both traces have 400 consecutive minute labels from 14:40 to 21:19 with the explicit UTC offset −07:00. Every CSV timestamp falls 86,400 seconds after its workbook counterpart. A conversion between the stated offsets cannot explain the shift.

The workbook and the authors' hourly grid CSV agree on all 24 date, hour and demand triples for 3 May. We separately retrieved the original *Demand (MW)* column for SRP from the EIA-930 January–June 2025 balance file [5]. Each candidate day has 24 distinct local ending hours and one daily maximum. The UTC endpoint is seven hours later than the local endpoint on all 48 rows. Balancing-authority demand and the authors' utility series can have different measurement boundaries or revisions.

3 Observed power during peak demand

Let P_m be a minute observation and h_d the hour ending at the daily grid maximum on date d. We average the 60 observations in $[h_d - 1, h_d)$ and divide by the mean of the first 20 common clock labels, 14:40 through 14:59. Those readings precede both EIA peak intervals and are identical under both calendars. This denominator describes the trace. It cannot establish a settlement baseline.

Calendar and grid source	Peak (MW)	Local hour	Power/reference
3 May, EIA-930	4,963	16:00–17:00	0.7583
4 May, EIA-930	3,888	15:00–16:00	0.9898
3 May, authors' grid*	4,965	17:00–18:00	0.6874

*Treating the authors' hour labels 1–24 as ending hours and converting GW to MW. The third row changes only the grid input relative to the first.

Using $(h_d - 1, h_d]$ gives 0.7534 and 0.9897 for the first two rows and 0.6874 for the third. Dividing the two EIA-selected cluster means directly, in calendar order, yields 0.7661 under start labeling and 0.7613 under end labeling. The contrast therefore survives removal of the early reference. Both values compare hours selected from the same power sequence.

The authors' grid trace is closer to the EIA series for 3 May than to that for 4 May. Under hour-ending alignment their 24-hour mean absolute differences are 48.29 and 478.08 MW, respectively. This comparison supports dating the *grid input*. It does not independently date the cluster. Even for 3 May the two grid sources choose adjacent daily peak hours, changing which minute observations enter the peak calculation.

4 What the comparison establishes

Under the two exported calendars and the EIA peak definition, the normalized means form a two-point set. The data provide no sampling probabilities for these alternatives. A meter log with verifiable absolute time, or a dispatch acknowledgement unambiguously linked to this trace, could select one. The demonstration of power modulation within the trace survives either assignment. Measuring avoided load at a particular time would require an appropriate counterfactual at the same metering boundary.

SRP reports a summer system record on 7 August [6]. The May observation window cannot establish what the cluster delivered in that August hour. Our case illustrates why a procurement assessment of coincident capacity should retain a dispatch identifier, absolute time, meter boundary and the grid series that defines the peak.

Reproduction. The accompanying script fetches pinned trial inputs and the EIA file, checks the input hashes, pairs trace values, verifies UTC offsets and computes both interval conventions. Third-party source files remain with their publishers.

Model assessments and review context

OpenAI Codex assisted in the design, data reconciliation, code and writing. The author administers AIRR. Any assessment accompanying this manuscript is specific to the exact PDF. A reviewer's involvement in producing it bears on the interpretation of its score.

References

[1] P. Colangelo et al., AI data centres as grid-interactive assets, *Nature Energy* 11, 254–261 (2026). doi:10.1038/s41560-025-01927-1.

[2] P. Colangelo et al., Phoenix field demonstration, preprint v1, 2025, section 2.2.

[3] A. Coskun, The Potential for Data Center Load Flexibility, MassCEC Future Grid event materials, May 2025, slide 26.

[4] Emerald AI, Phoenix experiment data and Figure 2 workbook.

[5] US EIA, EIA-930 balancing authority data, January–June 2025 file retrieved 19 November 2026.

[6] Salt River Project, Fiscal 2026 second-quarter report, August–October 2025.