Electric Grid Fragility Analysis - Full Report

An Applied Analysis of Grid Stress Using Public Operational Data

By Miles Wolf
January 10, 2026

1. Executive Summary

This report presents a concise, data-driven framework for identifying and comparing periods of elevated operational stress across large electric power grids using publicly available operational data. Rather than modeling outages or reliability events, the analysis focuses on relative grid fragility—how unusual or demanding system behavior appears compared to each grid's own historical norms.

Using EIA-930 consolidated hourly operations data for calendar year 2025, the framework is applied to two fundamentally different balancing authorities: CAISO and PJM Interconnection. These regions were selected due to their contrasting scale, variability, and operating characteristics, enabling comparison of how stress manifests under different grid regimes. A normalized fragility score is constructed from short-timescale load ramp behavior and deviations from typical operating conditions. Because the metric is normalized within each balancing authority, it allows meaningful comparison across systems with vastly different absolute demand levels. Unsupervised clustering is used to organize operating days into recurring operational regimes, providing context for when and how stress occurs.

Several clear patterns emerge. Grid fragility is episodic rather than persistent, with a small number of days accounting for a disproportionate share of total stress. While both grids experience rare high-fragility events, the structure of these events differs systematically. In CAISO, elevated fragility is primarily associated with abrupt, high-magnitude net load ramps and rapid regime transitions, producing pronounced tail events rather than sustained stress. This behavior reflects the underlying structure of the CAISO system—particularly its high penetration of renewable generation and strong diurnal demand patterns—rather than weaker operational performance. In contrast, PJM fragility more often arises from sustained high-load conditions combined with ongoing ramping, resulting in a higher frequency of elevated-stress days.

The most extreme CAISO stress day identified (July 31, 2025) stands well apart from other high-stress days, exhibiting ramp magnitudes and variability several times larger than typical seasonal extremes despite peak demand remaining within a normal summer range. All ramp-related metrics are internally consistent, indicating genuine operational strain rather than a data artifact.

Regime-level analysis further shows that the highest-fragility regimes occur infrequently but are systematically associated with elevated stress, reinforcing a key operational insight: average conditions obscure risk, and the days that matter most are inherently rare.

This framework does not attempt to predict failures, model physical constraints, or establish causality. Instead, it provides a lightweight, extensible lens for identifying recurring fragility patterns - such as ramp-driven stress events, sustained high-load stress, and regime transitions - offering a foundation for deeper operational, weather-conditioned, or resource-mix analyses.

2. Data and Scope

2.1 Data Sources

This analysis uses publicly available operational data from the **U.S. Energy Information Administration (EIA) Form 930**, which provides near-real-time electricity system metrics reported directly by U.S. balancing authorities (BAs). The EIA-930 dataset includes hourly measurements of system demand and related operational quantities that reflect system-level behavior under varying load and operating conditions.
The analysis relies on the **consolidated EIA-930 balance dataset**, which harmonizes reporting formats across balancing authorities and enables consistent comparison across regions and time.

2.2 Temporal Coverage and Resolution

The dataset spans **January 1, 2025 through December 31, 2025**, covering a full calendar year of operations. This period captures:

- seasonal demand variability,
- typical day-to-day grid behavior,
- and rare but operationally significant stress events.

All metrics are processed at **hourly resolution** and subsequently aggregated to the **daily level** for analysis. Daily aggregation is used to characterize system-level stress patterns while reducing sensitivity to isolated hourly fluctuations. After filtering to the selected regions, the final dataset contains **17,520 hourly records** (8,760 hours per balancing authority).

2.3 Geographic Scope

The analysis focuses on two large and structurally distinct U.S. balancing authorities:

- **California Independent System Operator (CAISO)** — represented in the EIA-930 dataset under the BA code *CISO*. CAISO operates a grid with high penetration of variable renewable generation and pronounced diurnal demand patterns.
  PJM Interconnection (PJM) — serving a large portion of the Mid-Atlantic and Midwest, characterized by high aggregate demand, diverse generation resources, and sustained regional load patterns.

These regions were selected to enable comparison across **fundamentally different operating regimes**, rather than to rank system performance or reliability.

2.4 Data Filtering and Quality Control

Prior to analysis, several preprocessing and quality-control steps were applied:

- filtered the national EIA-930 dataset to include only CAISO and PJM records,
- verified complete hourly timestamp coverage across the full year,
- identified and quantified missing demand observations,
- incorporated EIA-provided demand adjustments where available.

Post-QC results indicate **near-complete temporal coverage** and negligible missing data after adjustment, providing a stable foundation for downstream feature engineering and analysis.

2.5 Analytical Scope and Interpretation

This study focuses on **load-side operational stress**, inferred from short-timescale demand ramp behavior and deviations from typical operating conditions. The resulting fragility metrics are **normalized within each balancing authority**, enabling meaningful comparison of *relative stress* across systems with vastly different absolute demand levels.

Importantly, the analysis distinguishes between:

- **baseline operating behavior**, reflecting typical day-to-day variability, and
- **tail-risk events**, representing rare days with unusually high operational strain.

Both regimes are retained and analyzed explicitly. Extreme observations are not removed; instead, interpretation emphasizes how frequently stress occurs, how extreme it becomes, and how it is distributed across operating regimes.

2.6 Scope Limitations

The analysis does **not** attempt to:

- model generator-level outages or unit commitment constraints,
- represent fuel availability or transmission congestion,
- predict reliability events or customer-facing impacts,
- establish causal attribution between stress and specific resources.

As a result, fragility metrics should be interpreted as **system-level indicators of operational strain**, not as measures of grid reliability or performance.

Future extensions could incorporate persistence metrics, weather conditioning, generation mix data, or resource-specific constraints to further characterize stress dynamics and duration.

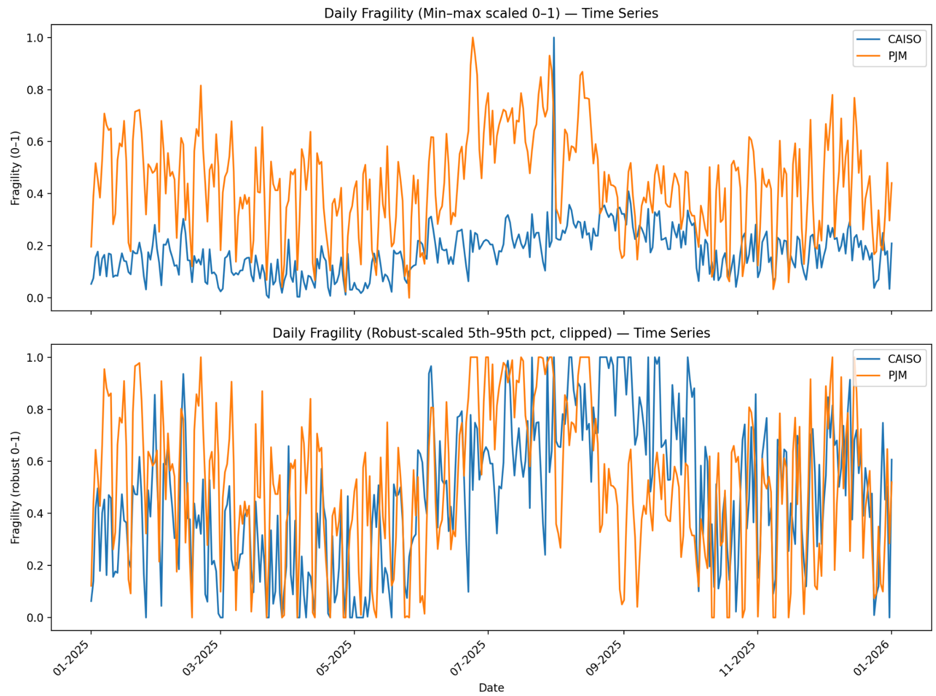

Figure 1. Daily Grid Fragility Time Series (Two Normalizations)
Daily fragility for CAISO and PJM over calendar year 2025 shown using two complementary normalizations: (a) min–max scaling, which emphasizes rare extreme stress days, and (b) robust scaling based on the 5th–95th percentile range, which emphasizes typical day-to-day variability. No observations are removed; scaling affects visualization only.

3. Measuring Grid Fragility

This section defines **grid fragility** as a composite operational stress metric derived from observed electricity demand. Rather than modeling outages or reliability outcomes directly, fragility captures how difficult a power system is to operate under normal conditions by combining multiple indicators of demand magnitude, volatility, and persistence. The emphasis is on operational stress, not system failure.

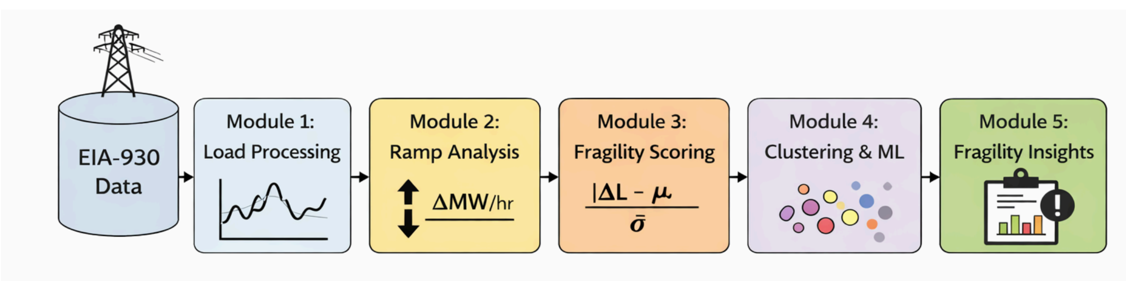

Figure 2. Grid Fragility Analysis Process Flowchart
This flowchart illustrates the full analysis workflow, showing how raw EIA-930 demand data is transformed via successive modules into fragility scores and downstream modeling inputs.

3.1 What the fragility score represents

Grid fragility measures how strongly daily demand behavior deviates from a grid's typical operating conditions. A grid is considered more fragile on days when demand is unusually high, changes rapidly, or remains elevated for extended periods, all of which increase operational burden. Conversely, a grid is less fragile when demand remains moderate, stable, and short-lived in its excursions.

From a controls perspective, fragility reflects the **combined magnitude, variability, and duration of demand stress**, rather than the absolute size of the grid. This framing enables meaningful comparison across systems with very different load scales.

3.2 How the score is constructed

Fragility is computed using daily features derived from hourly load data in the EIA-930 dataset. For each balancing authority, three demand-based stress indicators are calculated:

- **Daily peak load**, capturing demand magnitude
- **Daily maximum load ramp**, capturing short-term volatility
- **Fraction of hours operating near peak demand**, capturing stress persistence

Each feature is independently standardized using z-scores to account for grid-specific scale and variability. The daily fragility score is then defined as the unweighted sum of these standardized components, producing a single, dimensionless indicator of overall operational stress.

This composite formulation intentionally balances intensity, volatility, and duration without privileging any single dimension.

3.3 Why normalization matters

Raw demand levels and ramps vary substantially across grids and seasons. Without normalization, larger systems would appear more stressed simply due to higher absolute loads. Standardizing each feature ensures that fragility reflects **relative stress within each grid**, rather than absolute system size.

This approach allows direct comparison between regions such as PJM Interconnection and California ISO despite large differences in total demand and operating characteristics.

3.4 How to interpret high vs. low values

Fragility values near zero indicate typical operating conditions for a given grid. Positive values indicate above-normal stress, driven by high demand, rapid ramps, sustained peaks, or a combination thereof. Large positive values correspond to rare, high-stress days that may challenge system flexibility. Negative values indicate relatively mild operating conditions.

Fragility does not measure outage probability or reliability outcomes. Instead, it provides a continuous indicator of **how hard the grid is being asked to work** on a given day.

4. Operational Regimes & Stress Mechanisms

This section characterizes grid fragility through **recurring operational regimes** rather than isolated extreme events. Figure 2 provides a conceptual schematic of regime-based grid behavior. Figures 3a and 3b ground this framework by showing how daily operating conditions cluster in feature space for each balancing authority. Figures 4a and 4b then quantify how often these regimes occur and how much fragility they contribute. Together, these results show that fragility arises from **systematic, BA-specific operational mechanisms**.

4.1 Clustering of Operating Days

Daily grid operations cluster into a small number of **repeatable operating regimes** when analyzed across peak load, ramp magnitude, variability, and normalized stress metrics. As illustrated schematically in Figure 2 and observed empirically in Figures 3a and 3b, these regimes represent stable modes of system operation rather than statistical anomalies.

In feature space, most days occupy low- to moderate-stress regimes centered near average load and ramp conditions. Higher-stress regimes appear as coherent clusters at elevated load or ramp intensity, rather than as isolated outliers. This structure indicates that operational stress is **regime-dependent**, reflecting persistent operating conditions rather than rare events.

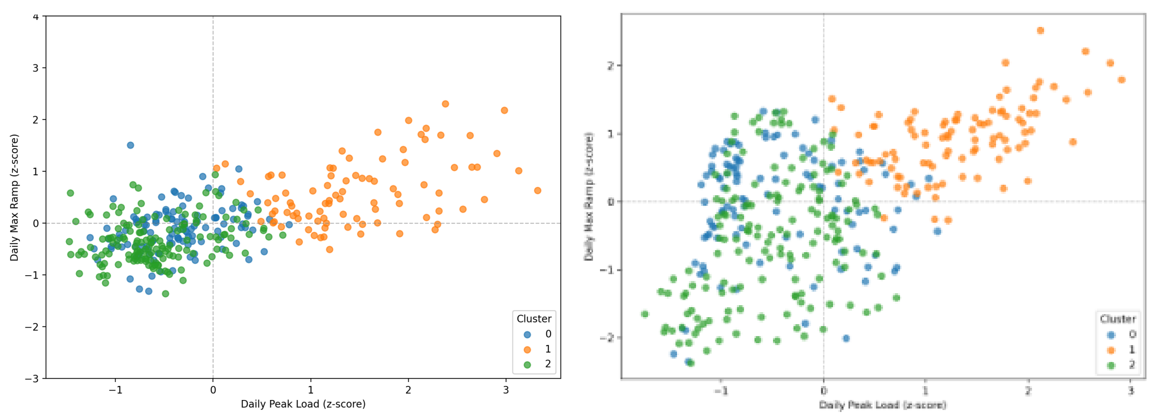

Figure 3. Operational regimes in normalized feature space
(a - left) CAISO *and* ***(b - right) PJM*** *daily operating regimes projected into normalized feature space defined by daily peak load (z-score) and daily maximum ramp magnitude (z-score). Points are colored by cluster assignment, representing operational regimes. Regime geometry differs by balancing authority: CAISO exhibits separation driven primarily by ramp magnitude and short-term variability, while PJM exhibits separation driven primarily by absolute load scale.*
Note: *Rare extreme ramping regime omitted from view for readability.*

4.2 Balancing Authority–Specific Stress Mechanisms

Figures 3a and 3b show the operational regimes for CAISO and PJM plotted side by side in normalized feature space. While both systems exhibit multiple regimes, their **geometry and dominant stress directions differ**.

In CAISO (Figure 3a), regime separation is driven primarily by **ramp magnitude and short-term variability**. High-stress regimes extend vertically in feature space, indicating elevated ramp rates even at moderate load levels. This reflects a system constrained by flexibility, rapid dispatch requirements, and transient dynamics rather than absolute demand.

In PJM (Figure 3b), regime separation is driven primarily by **absolute load scale**. Higher-stress regimes extend horizontally toward elevated peak load, with comparatively less vertical dispersion. This indicates stress arising from sustained demand magnitude, capacity utilization, and reserve margin compression, with ramping playing a secondary role relative to system size.

4.3 Regime Frequency vs. Fragility Contribution

Figures 4a and 4b relate **regime frequency** to **mean fragility**, revealing how different operating modes contribute to overall system stress.

In CAISO (Figure 4a), the most fragile regime is extremely rare, while moderately stressed regimes occur frequently and dominate cumulative fragility exposure. This indicates that CAISO's fragility is driven less by extreme events and more by **repeated operation in variability-stressed regimes**.

In PJM (Figure 4b), higher-fragility regimes are less extreme but occur more regularly at elevated load levels. Fragility is therefore accumulated through **persistent scale-driven stress**, rather than rare excursions into extreme operating states.

These frequency–severity relationships reinforce the interpretation that fragility is a product of **how often the system operates near its dominant constraints**, not just how severe the worst days are.

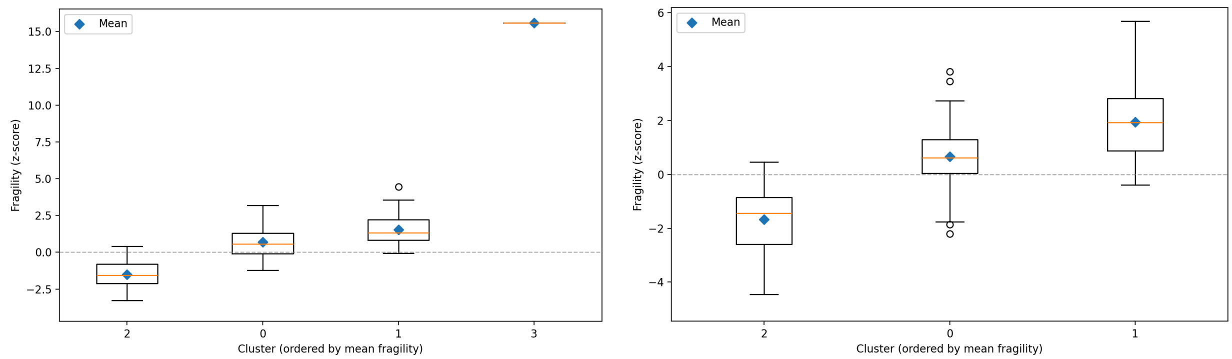

Figure 4. Regime frequency versus mean fragility
(a - left) CAISO (CISO) *and* ***(b - right) PJM*** *relationship between regime frequency (percentage of operating days) and mean fragility (z-score). Each point represents a distinct operational regime, illustrating the tradeoff between how often a regime occurs and the level of stress it imposes. In CAISO, the most fragile regime is rare, while moderately stressed regimes dominate cumulative exposure. In PJM, higher fragility is associated with more frequently occurring, scale-driven regimes.*

4.4 Scale vs. Variability as a Controls Framing

Taken together, Figures 3a–3b and 4a–4b motivate a controls-oriented framing of grid fragility in terms of **scale-driven versus variability-driven stress**.

Scale-driven stress reflects steady-state limitations associated with high absolute demand and infrastructure capacity. Variability-driven stress reflects transient limitations, where rapid ramps and short-timescale fluctuations challenge system responsiveness and control authority.

This distinction explains why uniform reliability metrics do not generalize across regions and highlights the need for **regime-aware analysis and modeling**. Effective prediction and intervention require identifying not only the magnitude of fragility, but the **operational mechanism producing it**.

5. Extreme Fragility Events

This section examines representative **extreme fragility events** to complement the regime-based analysis in Section 4. Rather than recurring operating modes, these events isolate days in the upper tail of fragility, where dominant system constraints are most sharply expressed. Figures 5 and 6 show combined temporal trajectories for load, ramping, fragility, and regime transitions for each balancing authority.

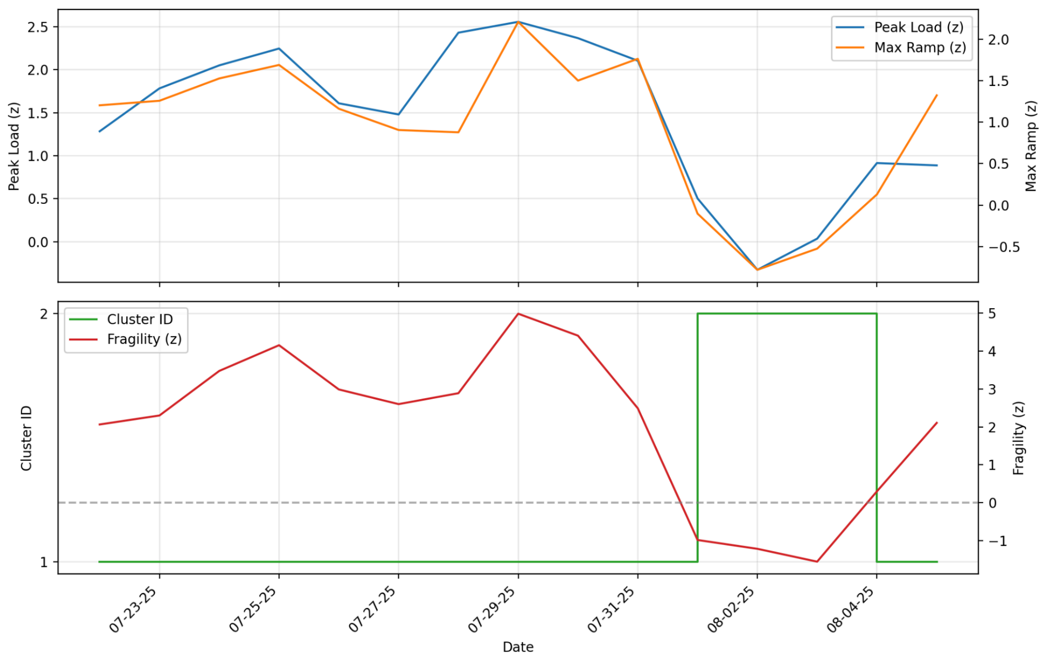

Figure 5. CAISO extreme fragility event — combined temporal trajectory

Daily peak load and maximum ramp magnitude (top), and cluster assignment and fragility (bottom) for a representative high-fragility event in CAISO. Fragility peaks during a brief, isolated ramping spike while peak load remains smooth, indicating stress driven by short-lived net-load variability rather than sustained demand. The cluster transition shows a transient excursion into a rare high-stress regime.

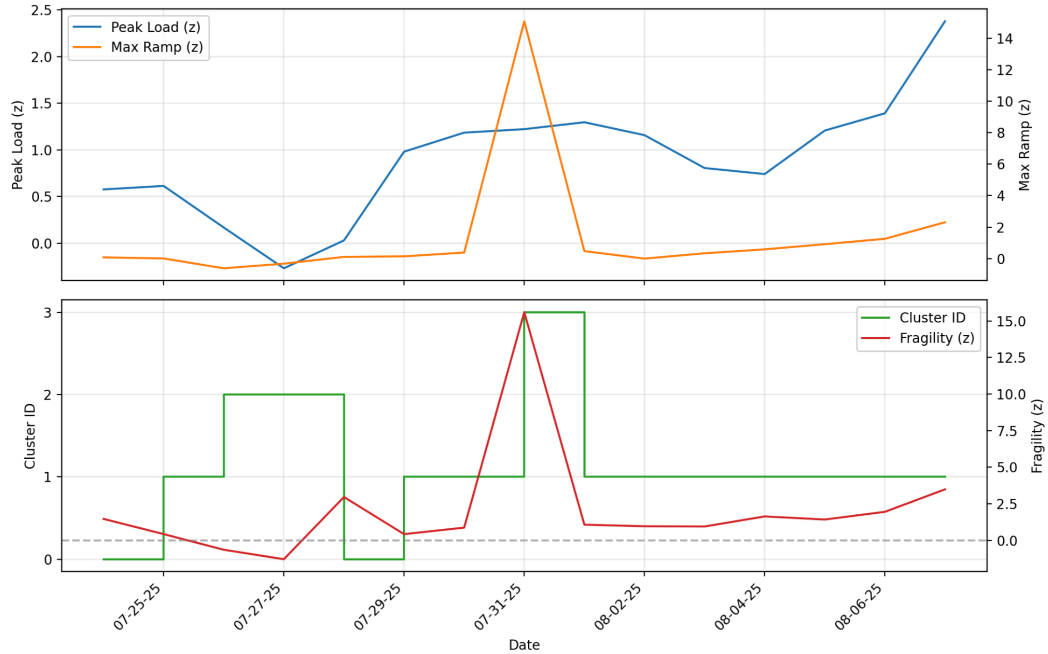

Figure 6. PJM extreme fragility event — combined temporal trajectory

Daily peak load and maximum ramp magnitude (top), and cluster assignment and fragility (bottom) for a representative high-fragility event in PJM. Fragility evolves gradually with elevated load and without sharp ramping spikes, reflecting sustained capacity and reserve pressure. Limited cluster transitions indicate prolonged operation within a high-load regime.

5.1 CAISO extreme fragility event

Figure 5 shows a high-fragility event in **CAISO** characterized by a sharp, isolated spike in ramp magnitude. While peak load is elevated, it evolves smoothly and does not exhibit anomalous behavior. In contrast, ramp magnitude exhibits a single extreme excursion, which coincides directly with the peak in fragility.

The cluster trajectory shows a brief transition into a rare, high-stress regime during this ramping spike, followed by rapid reversion to a more typical operating regime. This indicates that CAISO's extreme fragility is driven by **short-lived variability shocks** rather than sustained demand. Stress is dominated by control speed and flexibility limits, not energy adequacy.

5.2 PJM extreme fragility event

Figure 6 shows a representative high-fragility event in **PJM** with a markedly different structure. Peak load and ramp magnitude evolve smoothly and at comparable scales, with no single dominant spike. Fragility instead rises and falls gradually, tracking periods of elevated load.

Cluster transitions are minimal, indicating prolonged operation within a high-load regime rather than excursions into rare states. Fragility accumulates over multiple days, reflecting **sustained capacity and reserve pressure** rather than transient ramping stress.

5.3 Interpretation

Extreme events reinforce the regime-level findings from Section 4. In CAISO, extreme fragility arises from brief departures into high-variability states that challenge system responsiveness. In PJM, extreme fragility reflects extended operation near capacity limits, where reserve margins compress and stress accumulates over time.

These results demonstrate that extreme fragility is **mechanism-specific**, not universal. Even when fragility magnitudes are comparable, the underlying operational causes differ fundamentally across grids.

6. How Often Stress Occurs

This section examines the frequency distribution of grid stress to clarify the relationship between event rarity and severity. Rather than focusing on representative events or operating regimes, this analysis characterizes how often different levels of fragility occur across the full sample. Figure 7 shows the empirical distribution of daily fragility values for each balancing authority.

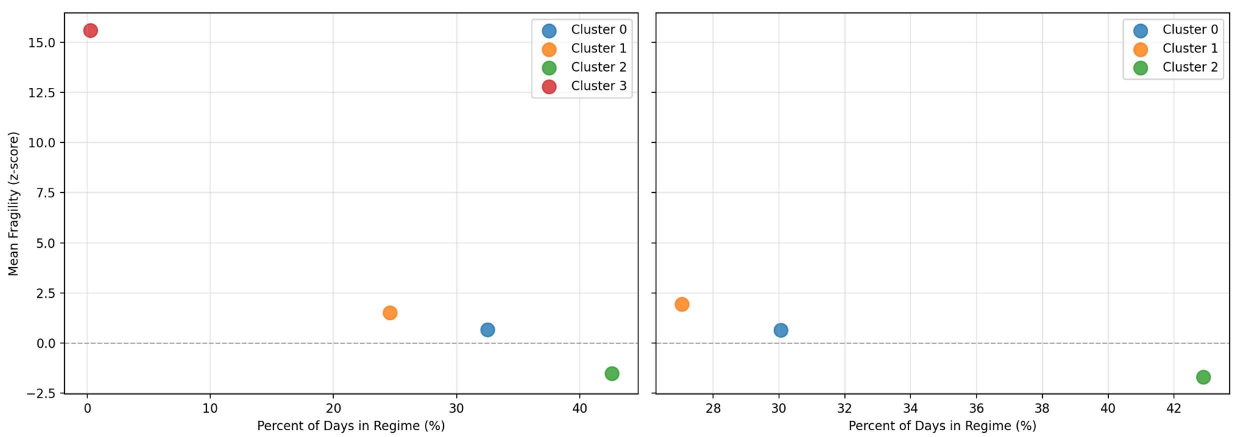

Figure 7. Distribution of daily fragility values

Plot of daily fragility for CAISO and PJM, illustrating the relative frequency of low, moderate, and high-stress conditions. The distributions are strongly right-skewed, with most days clustered at low fragility and a small number of extreme events forming a long upper tail.

6.1 Rarity–severity relationship

Figure 7 shows that high-fragility events are rare relative to typical operating conditions. The majority of days exhibit low to moderate fragility, while only a small fraction of observations populate the upper tail of the distribution. However, these rare events reach magnitudes far exceeding typical conditions, indicating a strong separation between common and extreme system behavior.

This rarity–severity relationship implies that the most consequential stress events occur infrequently but dominate the system's exposure to risk. Grid fragility is therefore not evenly distributed over time, but concentrated in a small number of extreme days.

6.2 Why averages hide risk

Because low-stress days occur far more frequently than high-stress days, traditional summary statistics such as means or medians are dominated by typical operating conditions. As a result, average fragility substantially understates the severity of the most critical events.

Figure 7 demonstrates that system reliability is governed by tail behavior rather than central tendencies. The days that define operational risk and challenge system limits are precisely those that are least represented in average metrics. This explains why aggregate statistics can suggest apparent stability even in systems that experience severe episodic stress.

6.3 Interpretation

The frequency analysis reinforces the need for distribution-aware and extreme-focused metrics when assessing grid performance. Fragility is an episodic phenomenon, driven by rare but severe events rather than persistent moderate stress. Evaluating grid robustness therefore requires attention to the upper tail of the stress distribution, where failures and operational challenges are most likely to occur.

These results motivate the use of fragility metrics and regime-based analysis that explicitly capture tail behavior, rather than relying on average system conditions.

7. Implications & Limitations

This section outlines what the proposed fragility framework enables analytically, clarifies the claims it does not make, and explains why the approach is well suited for scaling across regions and time. Together, these points situate the framework as a diagnostic and exploratory tool rather than a comprehensive reliability model.

7.1 What this framework enables

The fragility framework enables a systematic, quantitative assessment of operational grid stress using demand-side load behavior alone. By expressing stress relative to each balancing authority's own historical operating range, the approach allows meaningful comparisons across systems of vastly different sizes and load profiles.

Because the fragility signal is computed at daily resolution and normalized through z-scores and percentiles, it reveals temporal structure in grid stress, including seasonal patterns and rare extreme events. This makes it possible to study not only typical operating conditions, but also the episodic periods that dominate system risk.

The resulting fragility metrics function as controls-style stress signals that are suitable for downstream analysis. They can be directly used for clustering, anomaly detection, regime classification, or as features in machine learning models without requiring detailed system-specific tuning. As such, the framework provides a flexible analytical layer that can support both descriptive and exploratory modeling objectives.

7.2 What it does not claim

This framework does not attempt to model the physical structure of the power grid, including transmission topology, congestion, or contingency-based reliability criteria such as N-1 conditions. It also does not identify causal drivers of stress, such as individual generators, fuel constraints, market behavior, or policy decisions.

The analysis is intentionally demand-focused and does not explicitly incorporate supply-side dynamics, including renewable intermittency, outages, or reserve margins. As a result, the fragility metrics should be interpreted as indicators of system stress conditions rather than direct measures of reliability or adequacy.

Importantly, the framework does not predict failures or outages. Instead, it highlights operating conditions that are historically associated with elevated stress. Operator-grade reliability assessments and market analyses remain necessary for real-time decision-making and asset-level attribution.

7.3 Why it scales

The framework is built entirely on publicly available, standardized EIA-930 data, which provides consistent coverage across balancing authorities and time. This allows the analysis to be extended beyond individual regions without changes to the underlying methodology.

Computationally, the approach scales linearly with the number of time steps and balancing authorities, making it feasible to apply nationally or across multiple years. The modular pipeline design supports incremental feature additions, such as weather variables or supply-side indicators, without requiring refactoring of the core logic.

Because all metrics are normalized relative to each system's own historical behavior, the framework maintains comparability across grids with different absolute load levels and operating characteristics. This portability makes it well suited for longitudinal studies, regional comparisons, and future data updates.

8. Conclusion

This project presents a modular, reproducible framework for quantifying electric grid fragility using publicly available EIA-930 operational data. By translating hourly demand dynamics into interpretable fragility metrics, the analysis reveals clear and systematic differences in operational stress between balancing authorities, with PJM and CAISO exhibiting distinct load scales, variability, and tail-risk behavior. A key takeaway is that rare, high-stress events—while infrequent—can dominate fragility signals and warrant careful treatment rather than outright removal. Overall, the results show that relatively simple, controls-inspired features can surface meaningful structure in grid behavior, while the project's modular design enables straightforward extension to additional regions, supply-side drivers, and more advanced modeling approaches in future work.