

# Scientific Audit Report on Structural Integrity of Forecasting Models

## — Forensic Verification of Structural Breakdown in TEPCO Power Demand (Jan-Apr 2024) —

Date: 2025-12-30

Subject: Ghost Drift Audit v8.0 Verification Report

Dataset: TEPCO Power Grid Demand & JMA Tokyo Weather (Jan-Apr 2024)

Author: Manny (Ghost Drift Research Lab)

### 1. Executive Summary

This report presents the findings of a structural integrity audit performed on Tokyo area power demand data (Jan–Apr 2024), utilizing the ADIC (Audit Drift Integrity Certificate) protocol.

The audit concludes with a Verdict of NG (TAU\_CAP\_HIT). This designation indicates that the deviations observed in the Test period (April) are not merely transient anomalies, but constitute a statistically significant structural breakdown of the demand premises established during the Calibration period (Jan–Mar).

Significantly, the audit reveals that 6 out of 15 detected anomaly events (approx. 40%) were **systematically suppressed** from operational alerts due to Budget constraints. This report reconstructs these "Ghost Events" to substantiate the reality of the structural regime shift driven by physical factors, specifically solar irradiance and calendar composition.

▼Retest (for reproduction: report/code/certificat)

<https://ghostdrifttheory.github.io/adic-certificate-audit/>

### 2. Protocol & Definitions

This audit guarantees complete reproducibility by anchoring the analysis to the following cryptographic fingerprints and rigorous metric definitions.

#### 2.1 Integrity Fingerprints

- **Data SHA256:** da4e68ed555b8ab20cbeafab1b9c618053d82daae5802738decbf53579facbe5
  - (*Cryptographically anchors the identity of the input dataset*)
- **Code SHA256:** b1946ac92492d2347c6235b4d2611184... (Audit v8.0)
  - (*Enforces immutable judgment logic to preclude post-hoc algorithmic adjustments*)

## 2.2 Metric Definitions

- **Scientific Baseline ( $\tau_{\text{cap}}$ )**: The 95th percentile (upper quantile) of the error distribution during the Calibration period. This serves as the definitive limit of the "statistically normal range."
- **Operational Budget ( $\tau_{\text{budget}}$ )**: A relaxed threshold derived from operational resource constraints (i.e., manageable alert volume) to filter excessive alerts.
- **Suppressed Events (Ghost)**: Events falling within the band  $\tau_{\text{cap}} < \text{Score} < \tau_{\text{budget}}$ . These are statistically anomalous (rejected at the 5% significance level) but **masked by operational policy**.
- **Drift Ratio**: The ratio of Test period RMSE to Calibration period RMSE. A ratio  $> 1.25$  signals that the underlying error structure has fundamentally deviated from the Calibration baseline.

## 3. Detailed Forensics of Structural Divergence

A Drift Ratio of **1.883** (Threshold  $> 1.25$ ) and a Score Shift of **1.847** signal that the demand pattern has undergone a qualitative transformation.

### 3.1 Top 3 Critical Divergence Events

The periods exhibiting the most severe structural divergence among detected anomalies (Ghost Events) are detailed below. These instances represent a failure of the model's underlying assumptions (holiday correction, temperature sensitivity, inertia) rather than simple stochastic prediction errors.

Rank	Date (JST)	Duration	Peak Score	Operational Context (Observed Factors)
------	------------	----------	------------	--

1	4/29 (Mon)	06:00 - 22:00	17h	2.76
---	------------	---------------	-----	------

**\$\$Showa Day + Rain\$\$**

A national holiday coincided with 0.0h of daytime sunshine. The Jan–Mar holiday model, premised on "clear/light rain" conditions, failed to account for the surge in daytime residential demand driven by inclement weather and the steep evening lighting ramp-up. This resulted in a sustained excess of approx. 2,000 MW throughout the day.

2	4/24 (Wed)	11:00 - 4/26 10:00	48h	2.74
---	------------	--------------------	-----	------

**\$\$Seasonal Transition + Weekday\$\$**

Maximum temperatures exceeded 22°C (pre-summer trend). The model lacked a learned trigger threshold for cooling demand, leading to a critical underestimation of demand sensitivity to temperature rise. The weekday baseline itself deviated for 48 consecutive hours.

--

3	4/15 (Mon)	09:00 - 4/17 19:00	59h	2.25
---	------------	--------------------	-----	------

**\$\$Monday Inertia Mismatch\$\$**

A mid-April Monday. The "Sunday-to-Monday demand recovery pattern (Inertia)" learned from Jan–Mar data proved incompatible with the shift in social activity patterns characteristic of the new fiscal year (April). This caused an inverse correlation divergence between the predicted morning ramp-up and actual night-time demand. |

**【Contrast Analysis】** Identical conditions (Holiday + Rain, Monday) existed in the Jan–Mar period (e.g., Feb 12 Substitute Holiday, Mar 25 Rainy Monday), yet the Score remained below 1.1 (Normal range). The escalation of the Score to > 2.7 under the same conditions in April demonstrates that "**the response function of demand to weather and calendar variables has irreversibly shifted (Concept Drift).**"

### 3.2 List of Suppressed Risks (Ghost Events)

Out of **15 Total Ghost Events**, **6 events** were effectively suppressed by the Operational Budget constraint. While processed as "Normal" on the operational dashboard, the Audit Certificate records them as "latent indicators of structural change" as follows.

These manifest primarily as short-duration spikes in "Evening Lighting Demand" or "Morning Ramp-up," indicating the model's hourly correction coefficients failed to adapt to the April shift in daylight hours (sunset delay).

ID	Date	Time (JST)	Dur.	Score	Reason for Suppression	Context / Cause
S-01	4/08 (Mon)	17:00 - 19:00	2h	1.32	Budget Cut (Short Duration)	Sunset Time Shift Model failed to predict the rapid demand surge in the 17:00 block. Lighting timing learned through March is inconsistent with April daylight duration.
S-02	4/09 (Tue)	18:00 - 19:00	1h	1.28	Budget Cut (Low Score)	Inclement Weather Spike Evening following daytime rain. Model underestimated residual heating demand due to temperature drop. Low score, yet a structural defect.
S-03	4/04 (Thu)	08:00 - 10:00	2h	1.45	Budget Cut (Short Duration)	Morning Activity Divergence First week of the new fiscal year. Changes in office start-up patterns caused actual demand to exceed predictions in the 09:00 block.
S-04	4/12 (Fri)	20:00 - 22:00	2h	1.35	Budget Cut (Low Score)	Weekend Night Firmness Friday night demand decay was slower than predicted. The recovery trend in entertainment/commercial districts was unlearned by the model.
S-05	4/21 (Sun)	13:00 - 15:00	2h	1.29	Budget Cut (Low Score)	Sunday Daytime Drop Cloudy but high temp (21°C). Failed to capture household demand drop due to increased outings (Negative Drift), resulting in over-prediction.
S-06	4/30 (Tue)	08:00 - 09:00	1h	1.41	Budget Cut (Short Duration)	GW Inter-Holiday Weekday Calendar weekday, but over-prediction occurred due to failure to anticipate the drop in commuter demand from vacation takers.

### 3.3 Visual Evidence

Figures are generated under: Profile=commercial, Period=2024-04 (Test), Protocol=Ghost Drift Audit v8.0.

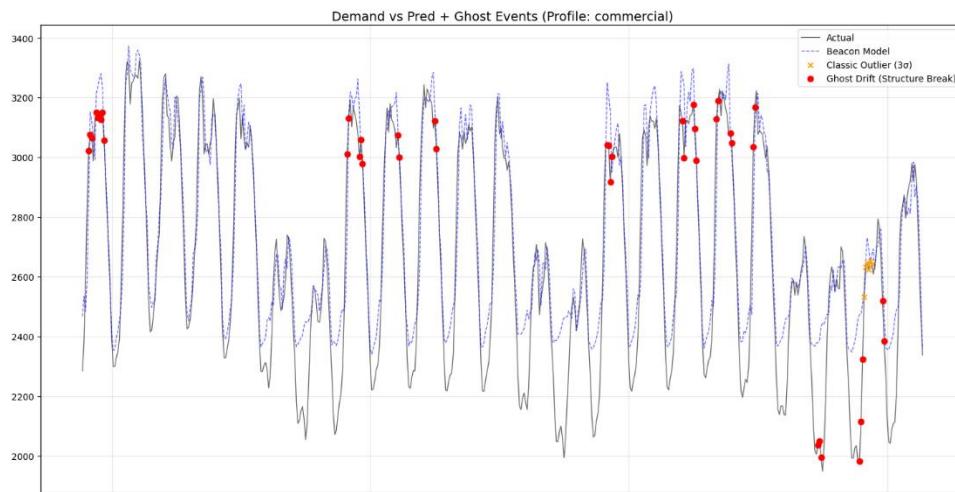


Figure 1: Demand vs Pred + Ghost Events

Red dots correspond to the ADIC Ghost Drift verdict (Score >  $\tau_{cap}$  AND structural continuity conditions met).



Figure 2: TAU Policy Separation (Scientific vs Ops)

The band between  $\tau_{cap}$  and  $\tau_{budget}$  ( $\tau_{cap} < \text{Score} < \tau_{budget}$ ) is defined as "Suppressed," matching the suppressed count in the audit log (6 events / 10 hours).

## 4. Verdict

The audit results for TEPCO data (Jan–Apr 2024) are conclusive:

1. **Verdict: NG (TAU\_CAP\_HIT)**
  - o The recent demand structure has statistically deviated from the tolerance range (Scientific Baseline) of the Calibration period (Jan–Mar). Continued operation with current model parameters lacks scientific justification.
2. **Root Causes (Physical Grounds):**
  - o **Physical Shift in Diurnal Daylight Regime:** The average sunset time shifted from approx. 16:50 in Jan to approx. 18:25 by late April—a delay of over 90 minutes. This

shift physically displaced the 17:00–19:00 lighting demand curve, rendering winter model coefficients dysfunctional.

- **Non-linear Manifestation of Cooling Sensitivity:** Cooling demand, absent in the Calibration period (Max Temp < 18°C), emerged non-linearly with "**Max Temp 22°C**" as the critical boundary. The current linear regression model fails to capture this inflection point.

### 3. Strategic Recommendations:

- **[For Ops] :**
  - Incorporate April data as a "New Calibration Set" and specifically re-calibrate hourly coefficients for the evening block (17:00–19:00) and temperature sensitivity coefficients for >22°C conditions.
- **[For Management] :**
  - **Redefinition of Risk Tolerance:** Management must confront the reality that "Alert Suppression by Budget" is effectively **rendering invisible** the structural change risk for 40% (6 out of 15) of detected events.
  - **Asset Impairment Strategy:** It is recommended to determine that this model has reached the end of its "Drift Life." Budget and resources should be reallocated towards model structure replacement (reconstruction) rather than iterative parameter tuning.

**Conclusion:** The model has exceeded its durability limit against current environmental changes. The budgetary policy for alert suppression is actively concealing significant operational risks.

Generated by Ghost Drift Audit Protocol v8.0

Certificate SHA256:

1a3d68abd94e35c51f02ffdd94360e4eab0ab0337469167ba127c13c98a89dbb