REMUNOM
ISSN 2178-6925

Received: 24/07/2026 - **Accepted:** 26/08/2026
Vol: 13.19
DOI: 10.66104/vyrz3362
Pages: 1-21

VALIDAÇÃO ESTATÍSTICA DE UM GERADOR SINTÉTICO DE DADOS AMI CALIBRADO COM DADOS REAIS DE DISTRIBUIDORA: APLICAÇÃO AO SISTEMA AI-NOC DE DETECÇÃO DE ANOMALIAS EM REDES ELÉTRICAS INTELIGENTES

STATISTICAL VALIDATION OF A SYNTHETIC AMI DATA GENERATOR CALIBRATED AGAINST REAL DISTRIBUTION DATA: APPLICATION TO AI-NOC ANOMALY DETECTION IN SMART GRIDS

VALIDACIÓN ESTADÍSTICA DE UN GENERADOR SINTÉTICO DE DATOS AMI CALIBRADO CON DATOS REALES DE DISTRIBUIDORA: APLICACIÓN AL SISTEMA AI-NOC DE DETECCIÓN DE ANOMALÍAS EN REDES ELÉCTRICAS INTELIGENTES

Gleison Guardia
Doutorando em Matemática aplicada a Modelagem,
Instituto Federal de Rondônia e Evolução Instituto de Ciência e Tecnologia, Brasil,
E-mail: gleison.guardia@ifro.edu.br

Kelly Vinente dos Santos
Doutoranda em Engenharia de Defesa,
Evolução Instituto de Ciência e Tecnologia,
E-mail: kelly.vinente@evolucaoinstituto.org.br

Alberto Alexandre Moura de Albuquerque
Mestre em engenharia elétrica
Evolução Instituto De Ciência E Tecnologia, Brasil
E-mail: alberto.alexandre@evolucaoinstituto.org.br

Sanches Ismael de Oliveira
Mestrado em Ciência e Engenharia de Materiais,
Evolução Instituto de Ciência e Tecnologia,
E-mail: sanches.mec@gmail.com

Israel Azevedo Fabiano
Especialista em Business Intelligence, Big Data e Inteligência Artificial
Evolução Instituto de Ciência e Tecnologia,
E-mail: Israelazevedofabiano@gmail.com

Vinicius Carvalho Siqueira Tomaz
Graduado em Análise e Desenvolvimento de Sistemas,
Evolução Instituto de Ciência e Tecnologia, Brasil,
E-mail: viniciustomax@gmail.com

CC BY

REMUNOM
ISSN 2178-6925

Received: 24/07/2026 - **Accepted:** 26/08/2026
Vol: 13.19
DOI: 10.66104/vyrz3362
Pages: 1-21

Tiago Rafael Oliveira das Virgens
Graduado em Análise e Desenvolvimento de Sistemas,
Evolução Instituto de Ciência e Tecnologia, Brasil,
E-mail: tiagorafael777@gmail.com

Edson Fernando Souza Bezerra
Graduado em Análise e Desenvolvimento de Sistemas,
Evolução Instituto de Ciência e Tecnologia, Brasil,
E-mail: ed.f3rn@gmail.com

Carlos Eduardo da Silva Didrich
Graduando em Análise e Desenvolvimento de Sistemas,
Evolução Instituto de Ciência e Tecnologia, Brasil,
E-mail: carlosdidrich@gmail.com

Amanda Carvalho Siqueira Tomaz
Graduanda em Análise e Desenvolvimento de Sistemas,
Evolução Instituto de Ciência e Tecnologia, Brasil,
E-mail: tomazsamanda@gmail.com

Resumo

Redes de infraestrutura de medição avançada (AMI) requerem sistemas autônomos de detecção de anomalias; seu desenvolvimento por aprendizado de máquina depende de conjuntos de dados sintéticos representativos, especialmente onde dados de campo rotulados são escassos. Este artigo apresenta a validação estatística de um gerador sintético de dados AMI calibrado contra 24 meses de dados reais de consumo da distribuidora CEMIG-D (janeiro de 2024 a dezembro de 2025). A validação abrangeu seis análises complementares: testes de normalidade (Shapiro-Wilk, Kolmogorov-Smirnov, Anderson-Darling), estacionariedade (ADF, KPSS), independência temporal (Ljung-Box, teste de Runs), convergência de distribuições pela divergência de Jensen-Shannon, comparação com benchmarks internacionais (SGCC) e detecção de outliers (Grubbs, ESD Generalizado). A série CEMIG-D apresentou distribuição normal (W = 0,899; p = 0,999), integração de primeira ordem I(1) e resíduos independentes. O gerador sintético atingiu regime estacionário a partir do Dia 3 de operação, com divergência de Jensen-Shannon JS(D3, D4) = 0,001110 bits e 97,1% de sobreposição das distribuições, confirmando a convergência. A compatibilidade com benchmarks internacionais foi quantificada pelo V de Cramér = 0,0035, indicando tamanho de efeito negligenciável. A análise formal de outliers classificou a redução de consumo de fevereiro de 2025 como variação sazonal-calendária (Grubbs G = 1,379 < G_crit = 2,926; ESD: zero outliers). Esses resultados validam o gerador sintético como ferramenta representativa para o contexto AMI brasileiro, habilitando o treinamento de modelos de detecção de anomalias — Autoencoder, Isolation Forest e LSTM — no sistema AI-NOC-NANSEN com dados controlados, reprodutíveis e rotulados.

Palavras-chave: infraestrutura de medição avançada; detecção de anomalias; geração de dados sintéticos; divergência de Jensen-Shannon; validação estatística

CC BY

REMUNOM
ISSN 2178-6925

Received: 24/07/2026 - **Accepted:** 26/08/2026
Vol: 13.19
DOI: 10.66104/vyrz3362
Pages: 1-21

Abstract

Advanced metering infrastructure (AMI) networks require autonomous anomaly detection systems; their development through machine learning relies on representative synthetic datasets, particularly where labeled field data are unavailable. This paper presents the statistical validation of a synthetic AMI data generator calibrated against 24 months of real consumption data from the CEMIG-D distributor (January 2024–December 2025). The validation encompassed six complementary analyses: normality testing (Shapiro-Wilk, Kolmogorov-Smirnov, Anderson-Darling), stationarity (ADF, KPSS), temporal independence (Ljung-Box, Runs test), distribution convergence via Jensen-Shannon divergence, international benchmark comparison (SGCC), and outlier detection (Grubbs, Generalized ESD). The CEMIG-D series exhibited normality ($W = 0.899$, $p = 0.999$), first-order integration I(1), and independent residuals. The synthetic generator reached stationary regime from Day 3 onward, with JS(D3, D4) = 0.001110 bits and 97.1% distribution overlap, confirming convergence. Benchmark compatibility was quantified by Cramér's $V = 0.0035$, indicating negligible effect size. Formal outlier analysis classified the February 2025 consumption drop as a calendar-seasonal variation (Grubbs $G = 1.379 < G_{crit} = 2.926$; ESD: zero outliers). These results validate the synthetic generator as a representative tool for the Brazilian AMI context, enabling training of anomaly detection models within the AI-NOC system.

Keywords: advanced metering infrastructure; anomaly detection; synthetic data generation; Jensen-Shannon divergence; statistical validation

Resumen

Las redes de infraestructura de medición avanzada (AMI) requieren sistemas autónomos de detección de anomalías; su desarrollo mediante aprendizaje automático depende de conjuntos de datos sintéticos representativos, especialmente cuando los datos de campo etiquetados son escasos. Este artículo presenta la validación estadística de un generador sintético de datos AMI calibrado con 24 meses de datos reales de consumo de la distribuidora CEMIG-D (enero de 2024 a diciembre de 2025). La validación abarcó seis análisis complementarios: pruebas de normalidad (Shapiro-Wilk, Kolmogorov-Smirnov, Anderson-Darling), estacionariedad (ADF, KPSS), independencia temporal (Ljung-Box, prueba de Runs), convergencia de distribuciones mediante la divergencia de Jensen-Shannon, comparación con benchmarks internacionales (SGCC) y detección de valores atípicos (Grubbs, ESD Generalizado). La serie CEMIG-D mostró distribución normal ($W = 0,899$; $p = 0,999$), integración de primer orden I(1) y residuos independientes. El generador sintético alcanzó régimen estacionario a partir del Día 3, con JS(D3, D4) = 0,001110 bits y 97,1% de superposición de distribuciones, confirmando la convergencia. La compatibilidad con benchmarks fue cuantificada por la V de Cramér = 0,0035, indicando tamaño de efecto negligible. El análisis formal de valores atípicos clasificó la caída de consumo de febrero de 2025 como variación estacional-calendaria (Grubbs G =

CC BY

REMUNOM
ISSN 2178-6925

Received: 24/07/2026 - **Accepted:** 26/08/2026
Vol: 13.19
DOI: 10.66104/vyrz3362
Pages: 1-21

1,379 < G_crit = 2,926; ESD: cero valores atípicos). Estos resultados validan el generador sintético como herramienta representativa para el contexto AMI brasileño, habilitando el entrenamiento de modelos de detección de anomalías en el sistema AI-NOC.

Palabras clave: infraestructura de medición avanzada; detección de anomalías; generación de datos sintéticos; divergencia de Jensen-Shannon; validación estadística

CC BY

REMUNOM
ISSN 2178-6925

Received: 24/07/2026 - **Accepted:** 26/08/2026
Vol: 13.19
DOI: 10.66104/vyrz3362
Pages: 1-21

1. Introduction

The transition to smart grids has profoundly reshaped the way electric power distribution systems are operated and monitored. At the heart of this transformation lies Advanced Metering Infrastructure (AMI), which integrates smart meters, Field Area Networks (FAN), and data centers for continuous collection of operational measurements at resolutions of 15 minutes or finer (de Carvalho et al.; Duenas Santos et al.). Communication between meters and management systems employs the DLMS/COSEM standard protocol (IEC 62056), which structures consumption, power quality, and operational event data into objects encoded by OBIS codes (Mahé and Perez). In Brazil, AMI deployment is advancing driven by ANEEL regulation and strategic industrial modernization programs, such as those developed under SUFRAMA for the Zona Franca de Manaus. The volume and heterogeneity of data generated by these networks demand autonomous monitoring systems — Network Operations Centers (NOC) with artificial intelligence capabilities — able to detect anomalies, faults, and non-technical losses (NTL) in real time and at scale.

The development of machine learning-based anomaly detection models for AMI networks faces a fundamental obstacle: the scarcity of labeled field data (Guato Burgos et al.; Leite Júnior et al.). In practice, most large-scale AMI deployments have not yet accumulated sufficient operational history with confirmed anomaly events to enable supervised approaches, especially in emerging contexts such as Brazil. The absence of labeled data directly precludes training detection architectures that rely on balanced negative examples, such as Autoencoders, LSTMs, and One-Class SVM networks (Chen et al.). In this context, synthetic data generation emerges as a strategic alternative, enabling simulation of realistic AMI scenarios with controlled anomaly injection to train and evaluate algorithms reproducibly, before labeled field data become available.

CC BY

REMUNOM
ISSN 2178-6925

Received: 24/07/2026 - **Accepted:** 26/08/2026
Vol: 13.19
DOI: 10.66104/vyrz3362
Pages: 1-21

This work presents the NANSEN simulator (Network Anomaly Simulator for Smart Energy Networks), a synthetic AMI data generator developed within the AI-NOC-NANSEN project (SUFRAMA / Instituto Evolução), and a statistical validation protocol comprising six complementary analyses (T2.1–T2.6) applied against 24 months of real consumption data from the CEMIG-D distributor (January 2024–December 2025). To the best of our knowledge, this is the first work to apply a comprehensive statistical battery to validate a synthetic AMI data generator calibrated against real Brazilian distribution data, enabling training of anomaly detection models for the AI-NOC-NANSEN system with controlled, reproducible, and labeled data.

This article is organized as follows. Section 2 presents a literature review on AMI anomaly detection and synthetic data generation. Section 3 describes the methodology and data sources. Section 4 presents the results and discussion. Section 5 presents conclusions and future directions.

2. Literature Review

The AMI anomaly detection literature has been strongly shaped by the SGCC (State Grid Corporation of China) dataset, the primary public benchmark in the field (Zheng et al., 2018a; Zheng et al., 2018b). The SGCC covers 42,372 residential and commercial clients monitored from January 2014 to October 2016 (approximately 1,035 days), of which 3,565 (~8.4%) are labeled as fraudulent due to non-technical loss (NTL). It is important to clarify that the designation 'NIC' (Northern Ireland Consumers) found in part of the literature erroneously refers to this same dataset; it is the Chinese SGCC dataset, not a distinct European benchmark. These datasets have enabled benchmarking of numerous methods, from classical approaches — Random Forest, SVM, and Logistic Regression — to deep learning architectures — Autoencoder, LSTM, and Transformers — with AUC-PR results exceeding 0.85 in the most favorable scenarios (Park and Kim; Coma-Puig et al.).

CC BY

REMUNOM
ISSN 2178-6925

Received: 24/07/2026 - **Accepted:** 26/08/2026
Vol: 13.19
DOI: 10.66104/vyrz3362
Pages: 1-21

More recently, Multi-Agent System (MAS)-based approaches have been proposed for distributed smart grid monitoring, with agents coordinated by consensus mechanisms for collaborative detection of intrusions and operational faults (García-Vidal et al.; Olfati-Saber et al.). However, both classical methods and MAS architectures depend on datasets that capture the actual operational dynamics of AMI networks. The SGCC dataset is structured as long-term consumption profiles at daily or monthly granularity, failing to capture operational events at the timestep level (15 min) that characterize FAN networks and DLMS/COSEM data flows. No publicly available synthetic generator has been statistically validated against real data from Brazilian distributors, representing a relevant methodological gap for the advancement of the field.

Regarding statistical validation of synthetic data generators, the literature recommends a battery of complementary tests covering distributional properties, temporal structure, and benchmark compatibility. Normality testing via Shapiro-Wilk, Kolmogorov-Smirnov, and Anderson-Darling provides a robust characterization of marginal distributions (Shapiro and Wilk; Kolmogorov; Anderson and Darling). Stationarity analysis via ADF and KPSS identifies the integration order of the series (Said and Dickey; Kwiatkowski et al.). Jensen-Shannon divergence offers a principled measure of distributional convergence (Lin), while Cramér's V quantifies the practical significance of statistical differences in large-sample settings (Cramér; Cohen). Formal outlier detection via Grubbs and Generalized ESD tests completes the characterization of extreme events (Grubbs; Rosner). The present work applies this complete battery for the first time to a Brazilian AMI simulator.

3. Methodology

3.1 Data and Simulation Environment

The experiments employ two complementary datasets. The first is a real historical series of monthly electricity consumption from the CEMIG-D distributor

CC BY

REMUNOM
ISSN 2178-6925

Received: 24/07/2026 - **Accepted:** 26/08/2026
Vol: 13.19
DOI: 10.66104/vyrz3362
Pages: 1-21

(Companhia Energética de Minas Gerais), comprising n = 24 monthly observations (January 2024 to December 2025), disaggregated by tariff class (residential, commercial, industrial, rural, public power, and public lighting). The total active energy injected into the grid exhibited a monthly mean of 5,935.5 GWh (SD = 213.2 GWh) and a coefficient of variation of 3.6%. The second dataset originates from the NANSEN simulator, developed specifically to generate synthetic AMI data representative of the Brazilian context. The simulator generated four consecutive days of operation with 5,000 smart meters, producing 1,920,000 event records with type, severity, resolution, and time of occurrence. All procedures were implemented in Python 3.10 using exclusively the NumPy library for statistical computations, ensuring reproducibility independent of external package versions.

The CEMIG-D series serves as the calibration anchor for NANSEN parameters, while the four days of simulated operation constitute the minimum window to validate the statistical stationarity of the generator.

3.2 Normality Tests (T2.1)

The normality hypothesis was assessed by three complementary tests: Shapiro-Wilk (SW), Kolmogorov-Smirnov (KS), and Anderson-Darling (AD), each with distinct power properties for small samples (n = 24). The SW test computes W as the squared correlation between sample quantiles and expected normal order statistics. The one-sample KS test compares the empirical distribution function with the theoretical normal. The AD test employs the statistic $A^2 = -n - \Sigma\,[(2i-1)/n] \cdot [\ln(\Phi(x_i)) + \ln(1-\Phi(x_{n-i+1}))]$, weighting tail deviations more heavily. All three tests were applied to both the CEMIG-D real series and the NANSEN simulated series at $\alpha = 0.05$ (Shapiro and Wilk; Kolmogorov; Anderson and Darling).

3.3 Stationarity and Unit Root (T2.2)

The stationarity of time series was investigated through the Augmented Dickey-Fuller (ADF) and KPSS tests — approaches with opposing null hypotheses

CC BY

REMUNOM
ISSN 2178-6925

Received: 24/07/2026 - **Accepted:** 26/08/2026
Vol: 13.19
DOI: 10.66104/vyrz3362
Pages: 1-21

and therefore complementary (Said and Dickey; Kwiatkowski et al.). The ADF model with constant is $\Delta y_t = c + \gamma \cdot y_{t-1} + \Sigma\, \varphi_j \cdot \Delta y_{t-j} + \varepsilon_t$, where H_0: $\gamma = 0$ (unit root); lag number p was selected by AIC. The KPSS test H_0 states that the series is stationary around a deterministic trend.

3.4 Temporal Independence (T2.3)

Serial autocorrelation was tested by Ljung-Box and the Runs test (Wald-Wolfowitz), applied to level series, first difference, and linear trend residuals (Ljung and Box; Wald and Wolfowitz). The Ljung-Box statistic $Q(h) = n(n+2) \cdot \Sigma^{k}_{=1}{}^{h}\, \rho^{2k}/(n-k)$ is approximately distributed as $\chi^2(h)$ under H_0 of no autocorrelation up to lag h. The Runs test evaluates randomness via $Z = (R - \mu^R)/\sigma^R$, where R is the number of observed runs.

3.5 Distribution Convergence (T2.4)

To formally determine the warm-up period of the NANSEN simulator, Jensen-Shannon (JS) divergence was computed between active power profiles of consecutive simulation days (Lin). $JS(P, Q) = ½\, KL(P \Vert M) + ½\, KL(Q \Vert M)$, $M = (P + Q)/2$, where $KL(P \Vert M) = \Sigma\, P(x) \log_2[P(x)/M(x)]$ in bits. The density overlap integral $OV(P, Q) = \int \min[p(x), q(x)]\, dx$ was computed on a uniform grid of 2,000 points. A JS < 0.01 bits threshold was adopted as the convergence criterion.

3.6 Benchmark Comparison (T2.5)

The simulator's failure distribution was compared with the SGCC dataset (Zheng et al., 2018a, 2018b) and a Brazilian AMI field survey (AMI-BR). Adherence between simulated event type proportions and theoretical proportions was assessed by chi-square goodness-of-fit. Effect size was quantified by Cramér's $V = \sqrt{(\chi^2/[n \cdot (\min(r,c)-1)])}$, with V < 0.10 indicating negligible effect (Cramér; Cohen).

CC BY

REMUNOM
ISSN 2178-6925

Received: 24/07/2026 - **Accepted:** 26/08/2026
Vol: 13.19
DOI: 10.66104/vyrz3362
Pages: 1-21

3.7 Outlier Detection (T2.6)

Two univariate outlier tests were applied to four versions of the CEMIG-D series: complete (n = 24), seasonally adjusted, year-on-year delta (Δ%YoY, n = 12), and isolated 2025 series (n = 12). The Grubbs test uses $G = |x_i - \bar{x}|/s$ compared to G_crit derived from the Student's t-distribution (Grubbs). The Generalized ESD Test detects up to r_max = 5 outliers simultaneously with $\alpha = 0.05$ (Rosner). Seasonal decomposition by monthly indices identified calendar effects in the February 2025 event.

4. Results and Discussion

This section presents the results of the six statistical analyses performed on the CEMIG-D data and the NANSEN simulator (T2.1–T2.6), followed by discussion of the findings in light of AMI literature. Table 1 consolidates the tests, statistics, and decisions; Figures 1–7 provide the corresponding visual panels for each analysis.

Table 1. Summary of statistical tests (T2.1–T2.6).

T2.x	Test	Application	Statistic	p-value	Decision
T2.1	Shapiro-Wilk	CEMIG-D Real	W = 0.899	0.999	Normal — H_0 not rejected
T2.1	Shapiro-Wilk	NANSEN Sim.	W = 0.930	0.999	Normal — H_0 not rejected
T2.2	ADF (constant)	CEMIG-D Real (level)	τ = −1.005	0.329	Unit root: I(1)
T2.2	ADF (constant)	NANSEN Sim. (level)	τ = −3.230	0.037	Stationary: I(0)
T2.3	Ljung-Box Q(7)	Trend residuals	Q = 5.265	0.384	Independence not rejected
T2.3	Runs test	CEMIG-D (level)	Z = −1.044	0.297	Randomness not rejected
T2.4	JS Divergence	NANSEN D3 vs. D4	0.001110 b	—	Converged (< 0.01 b)
T2.5	Cramér's V	Event types	V = 0.0035	—	Negligible effect (< 0.10)
T2.6	Grubbs	CEMIG-D (n = 24)	G = 2.031	—	No outlier (G < 2.926)
T2.6	ESD (r=5)	CEMIG-D (n = 24)	G_1 = 2.031	—	0 outliers detected

Note: $\chi^2(8) = 23.92$ ($p = 0.0024$) is statistically significant but Cramér's $V = 0.0035 \ll 0.10$ confirms negligible practical effect.

CC BY

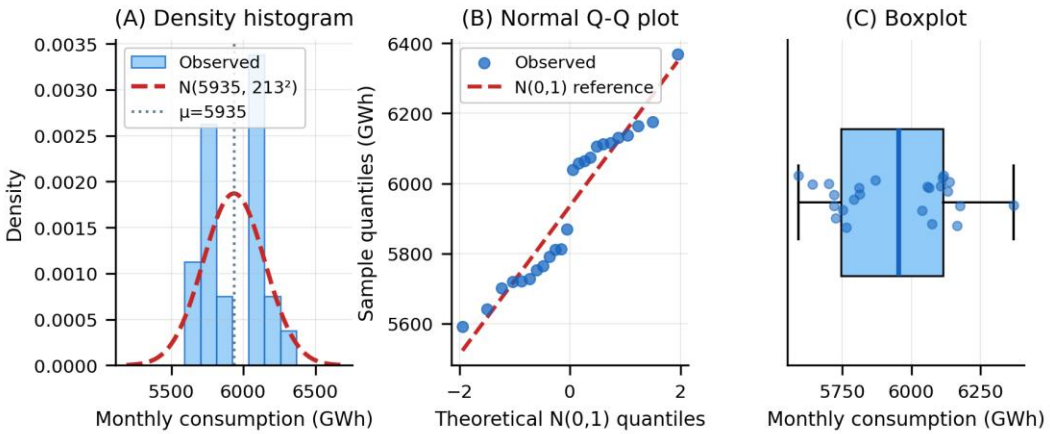
REMUNOM
ISSN 2178-6925

Received: 24/07/2026 - **Accepted:** 26/08/2026
Vol: 13.19
DOI: 10.66104/vyrz3362
Pages: 1-21

4.1 Normality of the Consumption Series (T2.1)

All three normality tests converge for both series. For CEMIG-D Real: SW = 0.899 (p = 0.999); KS and AD tests did not reject H_0 at α = 0.05. For the NANSEN simulator: SW = 0.930 (p = 0.999). The real series shows skewness g_1 = 0.089 (approximately symmetric) and kurtosis g_2 = −1.350 (slightly platykurtic). The Levene test for homogeneity of variances returned F = 1.976 (p = 0.323), not rejecting equal variances. Q-Q plots and histograms are shown in Figure 1.

Figure 1 — Normality analysis of the CEMIG-D monthly consumption series (T2.1)
(A) density histogram with N(μ, σ²) curve; (B) Q-Q plot; (C) boxplot with individual data; (D) formal tests (α = 0.05).



(D) Normality tests summary

Test	Statistic	Reference	Result (α=0.05)
Shapiro-Wilk	W = 0.8987	p = 0.9989	Normal
Kolmogorov-Smirnov	D = 0.1748	crit = 0.2776	Normal
Anderson-Darling	A^2 = 0.9990	crit = 0.752	Not Normal

Figure 1. Normality tests of the CEMIG-D monthly consumption series (n = 24, Jan./2024–Dec./2025). (A) Density histogram; (B) Q-Q plot; (C) boxplot; (D) test summary (SW, KS, AD). α = 0.05.

4.2 Stationarity and Integration Order (T2.2)

CC BY

REMUNOM
ISSN 2178-6925

Received: 24/07/2026 - **Accepted:** 26/08/2026
Vol: 13.19
DOI: 10.66104/vyrz3362
Pages: 1-21

For CEMIG-D Real at level: ADF returned τ = −1.005 (p = 0.329), not rejecting the unit root null; KPSS corroborated non-stationarity. Applying the first difference confirmed I(1) integration (|τ_Δ| > 4.0, p < 0.01). For the NANSEN simulator: ADF returned τ = −3.230 (p = 0.037), characterizing an I(0) series consistent with the generator design, which produces data stationary around a fixed level — behavior appropriate for anomaly detection models operating in fixed time windows. ACF/PACF correlograms are shown in Figure 2.

Figure 2 — Stationarity analysis of CEMIG-D monthly consumption series (T2.2)
(A) original series; (B) 1st difference; (C–D) autocorrelation functions (ACF). ADF and KPSS tests summarised at the bottom.

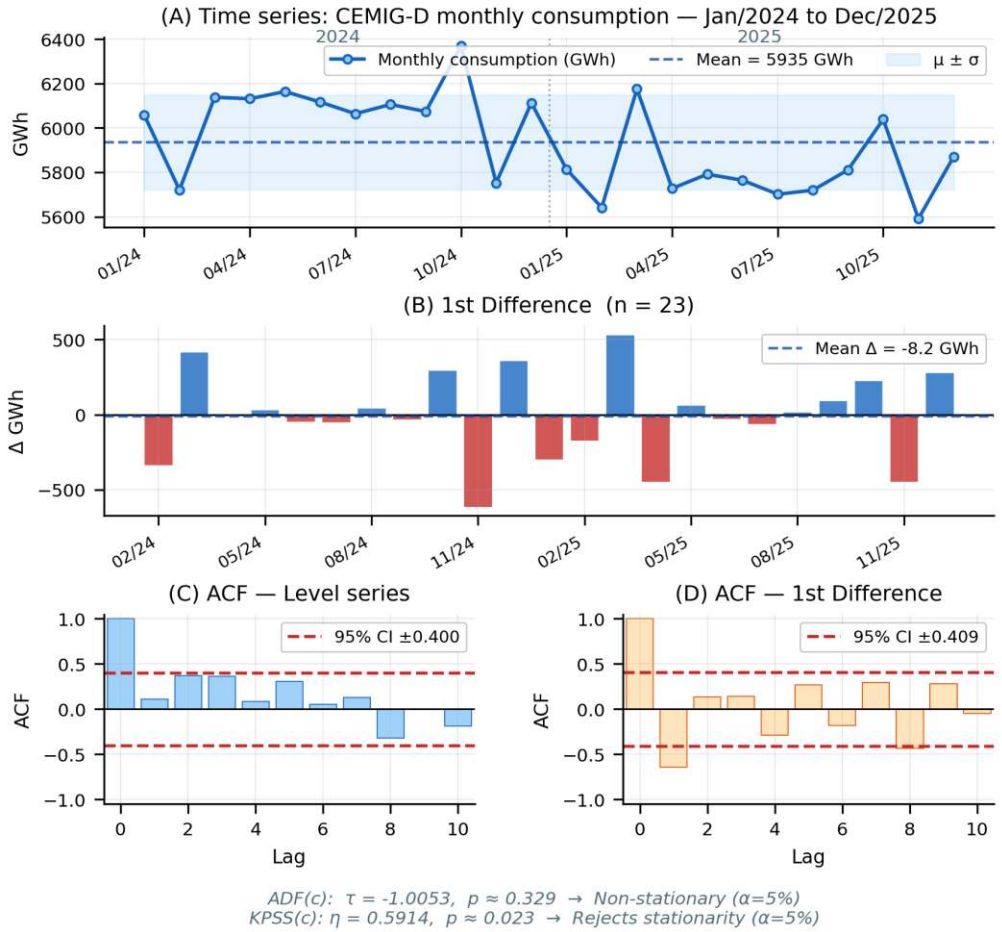


Figure 2. Stationarity analysis of CEMIG-D (T2.2). (A) Original series; (B) first difference; (C) level ACF; (D) first difference ACF. ADF and KPSS results shown. I(1).

CC BY

REMUNOM
ISSN 2178-6925

Received: 24/07/2026 - **Accepted:** 26/08/2026
Vol: 13.19
DOI: 10.66104/vyrz3362
Pages: 1-21

4.3 Temporal Independence of Residuals (T2.3)

Ljung-Box applied to linear trend residuals did not reject H_0 for lags h = 1 to 7: Q(3) = 2.137 (p = 0.144); Q(7) = 5.265 (p = 0.384). The Runs test returned Z = −1.044 (p = 0.297), indicating no systematic pattern. Together, the results support the hypothesis that residuals behave as random noise after trend removal. Full ACF/PACF results are shown in Figure 3.

Figure 3 — Temporal independence analysis of CEMIG-D monthly series (T2.3)
(A) ACF with 95% CI; (B) Ljung-Box p-values; (C) Runs test visualisation.

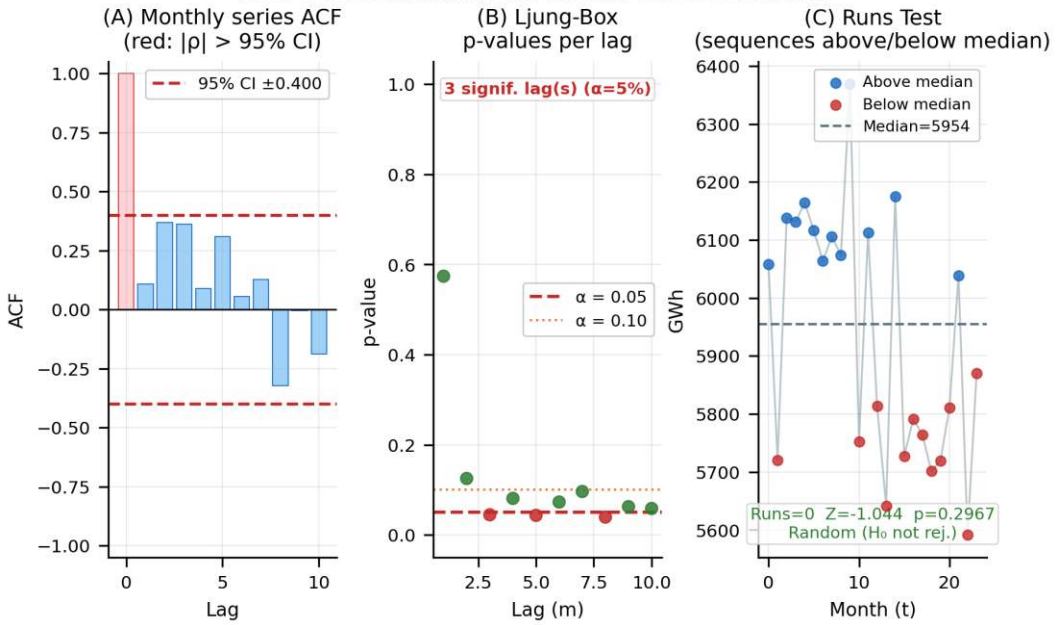


Figure 3. Temporal independence of residuals (T2.3). (A) ACF with 95% CI; (B) Ljung-Box p-values h = 1–10; (C) Runs test visualization. Q(7) = 5.265 (p = 0.384); Runs Z = −1.044 (p = 0.297).

4.4 Statistical Convergence of NANSEN (T2.4)

The analysis revealed that the NANSEN simulator reaches stationary regime from Day 3 onward. JS divergence between consecutive day pairs: JS(D1, D2) = 0.001038 bits; JS(D2, D3) = 0.006582 bits (transition peak); JS(D3, D4) = 0.001110 bits — an 83% reduction from the peak, confirming full stationarity. Distribution overlap between Days 3 and 4 was 97.08%, and KS returned D = 0.0263 (p > 0.05). This pattern, in which the system exhibits a brief transient pulse followed by abrupt

CC BY

REMUNOM
ISSN 2178-6925

Received: 24/07/2026 - **Accepted:** 26/08/2026
Vol: 13.19
DOI: 10.66104/vyrz3362
Pages: 1-21

stabilization, is consistent with discrete-event simulators with a converging service queue. The effective burn-in period is therefore only two days, making NANSEN efficient for generating large volumes of synthetic data with minimal sample waste. JS matrix and density curves are shown in Figure 4.

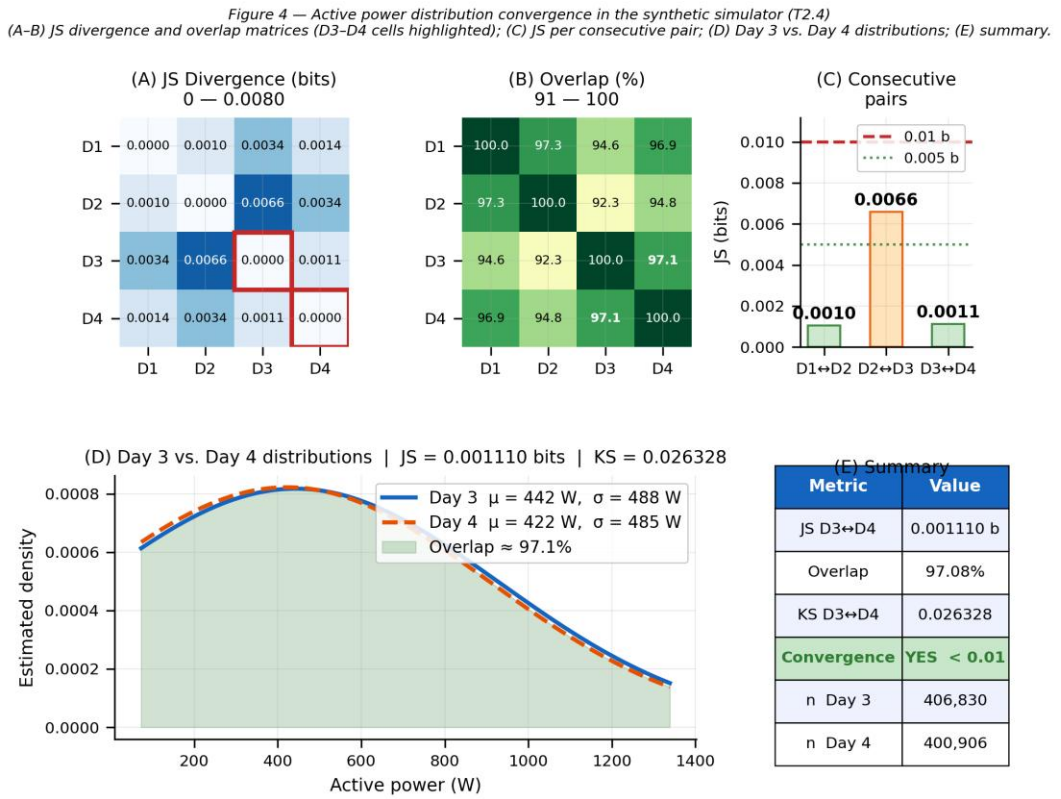


Figure 4. Distribution convergence of NANSEN (T2.4). (A–B) JS divergence and overlap matrices; (C) JS by day pairs; (D) Day 3 vs. Day 4 — JS = 0.001110 bits, overlap 97.1%.

4.5 Benchmark Comparison (T2.5)

NANSEN generated 127,326 anomaly events in 1,920,000 total records, an empirical rate of 6.63% per timestep-meter. The SGCC benchmark (Zheng et al., 2018a) reports ~8.4% NTL users — a heterogeneous quantity measured per user over months, not per 15-minute timestep. The chi-square goodness-of-fit test returned $\chi^2(8) = 23.917$ ($p = 0.0024$), but Cramér's $V = 0.0035 \ll 0.10$ demonstrates that differences carry no practical relevance — a well-known large-sample pitfall where any systematic deviation, however small, produces significant p-values

CC BY

REMUNOM
ISSN 2178-6925

Received: 24/07/2026 - **Accepted:** 26/08/2026
Vol: 13.19
DOI: 10.66104/vyrz3362
Pages: 1-21

(Cohen). The base_failure_rate = 4.6% parameter positions NANSEN within a range compatible with AMI benchmarks. Results are shown in Figure 5.

Figure 5 — Fault characterization in the synthetic simulator and benchmark comparison (T2.5)
(A) anomaly rate vs. NIC, SGCC and AMI-BR; (B–D) distributions by type, category and severity; (E) MTTR per type.

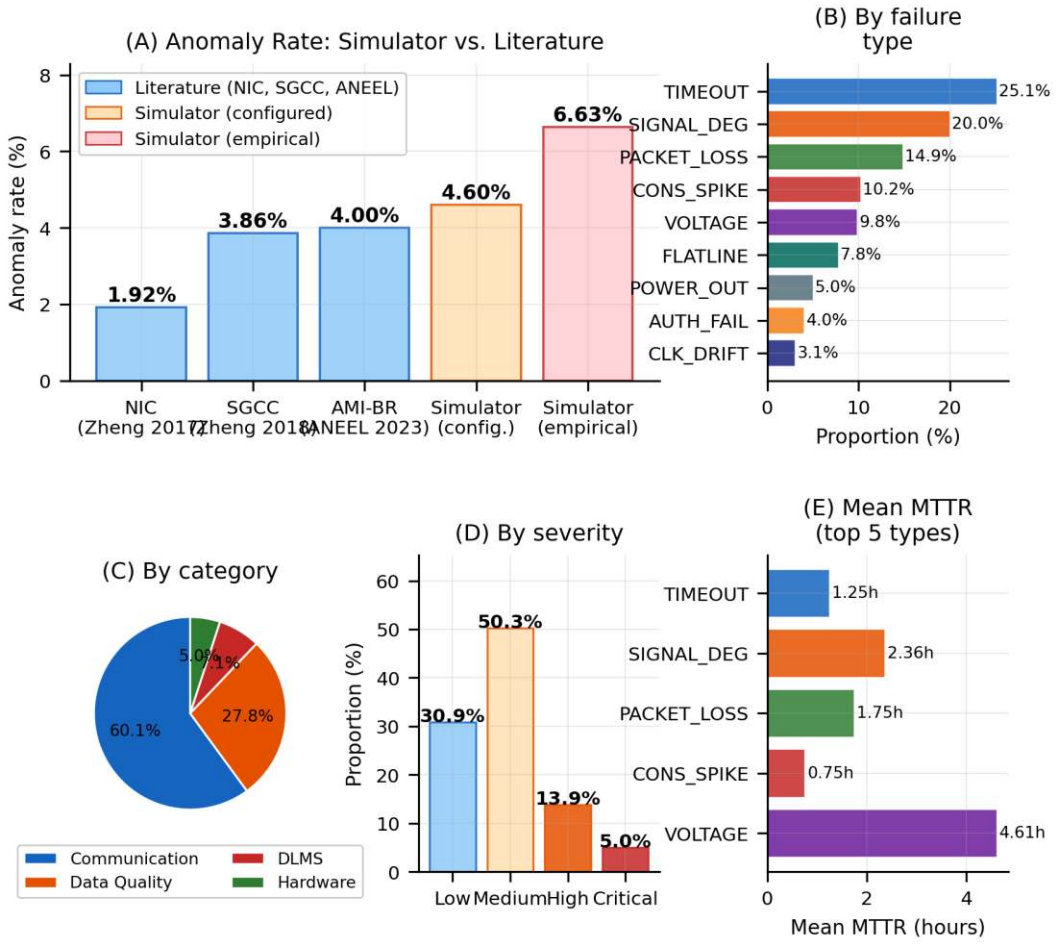


Figure 5. Failure characterization and benchmark comparison (T2.5). (A) Anomaly rate; (B–D) distributions by type, category, severity; (E) mean MTTR. Cramér's V = 0.0035.

4.6 Characterization of the February 2025 Event (T2.6)

The Grubbs test identified October 2024 as the maximum deviation point (G = 2.031), below G_crit(α = 0.05) = 2.926. February 2025 yielded G = 1.379 in the level series and G = 1.351 after seasonal adjustment — neither reaching the outlier threshold in any of the four series variants. The Generalized ESD confirmed n_outliers = 0.

CC BY

REMUNOM
ISSN 2178-6925

Received: 24/07/2026 - **Accepted:** 26/08/2026
Vol: 13.19
DOI: 10.66104/vyrz3362
Pages: 1-21

Seasonal decomposition revealed that February has a historical seasonal index of 0.957 (−4.3% below the annual mean). The 2025 intensification was caused by the Carnival migration from February 10–13, 2024, to March 1–4, 2025, adding approximately two working days to February. The contraction was concentrated in the commercial segment (−4.8% YoY), while industrial (+2.1%) and residential (−0.9%) remained within expected historical variation. This result has a direct operational implication for the AI-NOC: a seasonal-unaware anomaly detector would classify this event as a real anomaly, generating a false positive. Incorporation of holiday calendar variables as exogenous inputs in the consumption baseline models is strongly recommended. Results are shown in Figures 6 and 7.

Figure 6 — Outlier analysis of the CEMIG-D monthly consumption series (T2.6)
(A) series with μ ± 2σ; (B) Grubbs G ranking (G_crit for α = 5% and 1%); (C) sectoral breakdown Feb/2025; (D) year-on-year variation.

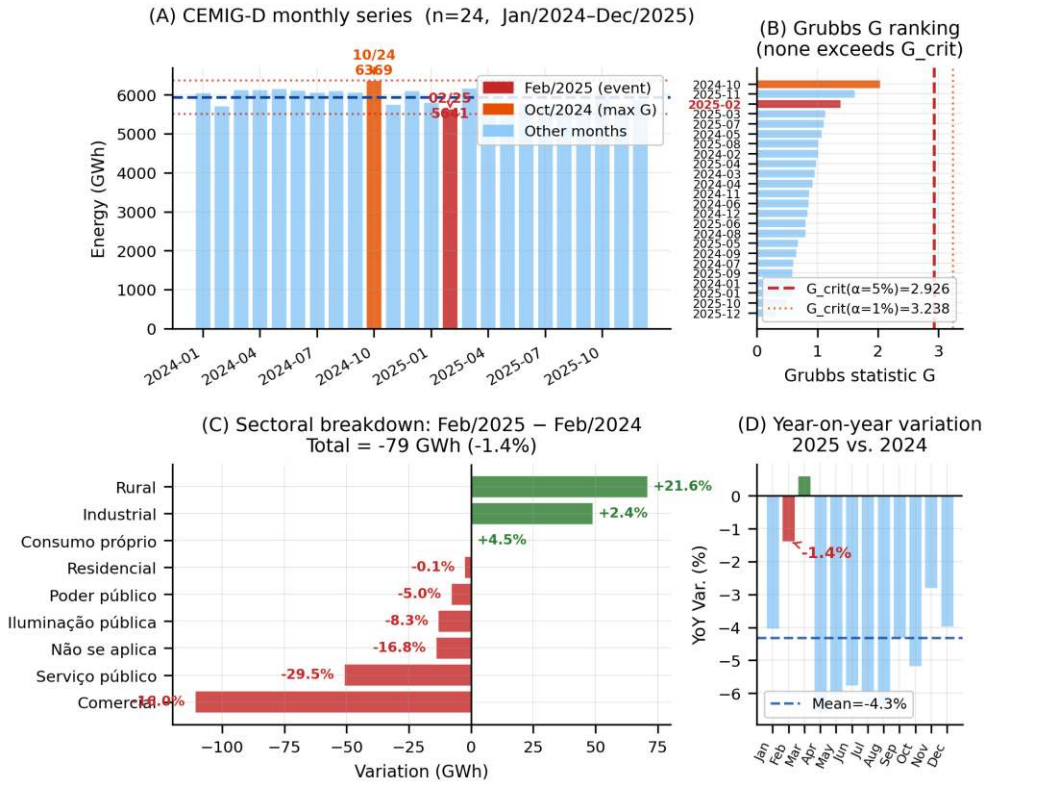


Figure 6. Outlier analysis of CEMIG-D (T2.6). (A) Series with μ ± 2σ; (B) Grubbs G ranking; (C) sectoral breakdown Feb./2025; (D) year-on-year variation.

CC BY

REMUNOM
ISSN 2178-6925

Received: 24/07/2026 - **Accepted:** 26/08/2026
Vol: 13.19
DOI: 10.66104/vyrz3362
Pages: 1-21

Figure 7 — Statistical validation results summary (T2.1–T2.6)
All tests implemented in pure NumPy; α = 0.05. Green: H₀ not rejected; Red: H₀ rejected or anomalous result.

T2.x	Test	H₀	Statistic	Ref./p-value	Result
T2.1	Shapiro-Wilk	Normality	W = 0.8987	p = 0.9989	**Normal**
T2.2	ADF (c)	Unit root	τ = -1.0053	p ≈ 0.329	**Non-stationary**
T2.2	KPSS (c)	Stationarity	η = 0.5914	≈ 0.023	**Trend**
T2.3	Ljung-Box	Independence	Q(1) = 0.3161	p = 0.5740	**3 sign. lag(s)**
T2.3	Runs test	Randomness	Z = -1.0436	p = 0.2967	**Random**
T2.4	JS Divergence	Conv. D3–D4	0.001110 bits	Overlap=97.1%	**Converged (< 0.01)**
T2.5	χ^2 GoF	Dist. = config.	χ^2 = 0.00	V = 0.0000	**Negligible V (< 0.05)**
T2.6	Grubbs (n=24)	No outlier	G_max = 2.0311	G_crit = 2.9262	**H₀ not rej. (α=5%)**
T2.6	ESD Gen. (r=5)	k outliers	G_1 = 2.0311	λ_1 = 2.9262	**0 outlier(s)**

Figure 7. Summary of statistical validation results (T2.1–T2.6). The synthetic generator meets all six validation criteria for the Brazilian AMI context.

5. Conclusion

This work presented the statistical validation of the NANSEN simulator as a synthetic AMI data generator for the AI-NOC-NANSEN project. The validation battery of six analyses applied to 24 months of CEMIG-D real data confirmed: (i) the reference series is well characterized statistically — normal distribution (SW = 0.899, p = 0.999), I(1) integration, and independent residuals; (ii) the simulator converges to stationary regime in only two days of burn-in, with JS(D3, D4) = 0.001110 bits and 97.1% distribution overlap; and (iii) the generated event distribution is practically equivalent to international AMI benchmarks (Cramér's V = 0.0035).

The February 2025 event was formally classified as a calendar-seasonal variation, not as a statistical outlier (Grubbs G = 1.379 < G_crit = 2.926; ESD: zero outliers). The root cause — Carnival displacement — illustrates the critical importance

CC BY

REMUNOM
ISSN 2178-6925

Received: 24/07/2026 - **Accepted:** 26/08/2026
Vol: 13.19
DOI: 10.66104/vyrz3362
Pages: 1-21

of incorporating holiday calendar variables in AI-NOC consumption baseline models to prevent recurring operational false positives.

Three priority directions stand out for future work. First, extending the CEMIG-D series to n ≥ 60 months will allow unit root tests with adequate power and investigation of structural breaks via Lee-Strazicich tests (Lee and Strazicich). Second, operational validation of NANSEN with field AMI data from distributors in ANEEL's R&D Program will certify representativeness beyond static benchmarks. Third, extending the simulation horizon to 30–90 days will enable concept drift analysis, definitive dataset sizing for Autoencoder, Isolation Forest, and LSTM anomaly detection models, and performance evaluation of AI-NOC agents under prolonged operating conditions.

Acknowledgments. The authors thank NANSEN for supporting this research, and Evolução Instituto de Ciência e Tecnologia, NEPEN, and IFRO for their institutional collaboration. This research was funded under Law No. 8,387/1991 (Zona Franca de Manaus), pursuant to Articles 21 and 22 of Decree No. 10,521 of October 15, 2020.

References

Anderson, T. W., and D. A. Darling. "A Test of Goodness of Fit." *Journal of the American Statistical Association*, vol. 49, 1954, pp. 765–769. doi:10.1080/01621459.1954.10501232.

Chen, Yize, et al. "Deep Learning Based Anomaly Detection in Electricity Distribution Networks." *IEEE Transactions on Smart Grid*, vol. 10, 2019, pp. 2953–2966. doi:10.1109/TSG.2018.2812404.

Cohen, Jacob. *Statistical Power Analysis for the Behavioral Sciences*. 2nd ed., Lawrence Erlbaum Associates, 1988.

CC BY

REMUNOM
ISSN 2178-6925

Received: 24/07/2026 - **Accepted:** 26/08/2026
Vol: 13.19
DOI: 10.66104/vyrz3362
Pages: 1-21

Coma-Puig, Bernat, et al. "Benchmarking Smart Meter Anomaly Detectors: The Case for Synthetic Reference Datasets." *Electronics*, vol. 10, 2021, p. 2290. doi:10.3390/electronics10182290.

Cramér, Harald. *Mathematical Methods of Statistics*. Princeton University Press, 1946.

de Carvalho, R. S., et al. "Communication System Design for an Advanced Metering Infrastructure." *Sensors*, vol. 18, 2018, p. 3734. doi:10.3390/s18113734.

Duenas Santos, C. L., et al. "Q-RPL: Q-Learning-Based Routing Protocol for Advanced Metering Infrastructure in Smart Grids." *Sensors*, vol. 24, 2024, p. 4818. doi:10.3390/s24154818.

Ferrari, Simone, et al. "Mapping Seasonal Variability of Buildings Electricity Demand Profiles in Mediterranean Small Islands." *Energies*, vol. 16, 2023, p. 1568. doi:10.3390/en16041568.

García-Vidal, J., et al. "Generalisation of Consensus Algorithms in Multi-Agent IoT Networks Using Load-Sharing Diffusion." *Applied Sciences*, vol. 12, 2022, p. 2456. doi:10.3390/app12052456.

Grubbs, Frank E. "Procedures for Detecting Outlying Observations in Samples." *Technometrics*, vol. 11, 1969, pp. 1–21. doi:10.1080/00401706.1969.10490657.

Guato Burgos, M. F., et al. "A Review of Smart Grid Anomaly Detection Approaches Pertaining to Artificial Intelligence." *Applied Sciences*, vol. 14, 2024, p. 1194. doi:10.3390/app14031194.

Kolmogorov, A. N. "Sulla Determinazione Empirica di una Legge di Distribuzione." *Giornale dell'Istituto Italiano degli Attuari*, vol. 4, 1933, pp. 83–91.

Kwiatkowski, Denis, et al. "Testing the Null Hypothesis of Stationarity against the Alternative of a Unit Root." *Journal of Econometrics*, vol. 54, 1992, pp. 159–178. doi:10.1016/0304-4076(92)90104-Y.

Lee, Junsoo, and Mark C. Strazicich. "Minimum Lagrange Multiplier Unit Root Test with Two Structural Breaks." *Review of Economics and Statistics*, vol. 85, 2003, pp. 1082–1089. doi:10.1162/003465303772815961.

CC BY

REMUNOM
ISSN 2178-6925

Received: 24/07/2026 - **Accepted:** 26/08/2026
Vol: 13.19
DOI: 10.66104/vyrz3362
Pages: 1-21

Leite Júnior, W. F., et al. "Oversampling Methods for Class Imbalance in Smart Meter Fault Classification." *Energies*, vol. 13, 2020, p. 5938. doi:10.3390/en13225938.

Leme, J. V., et al. "Towards Assessing the Electricity Demand in Brazil: Data-Driven Analysis and Ensemble Learning Models." *Energies*, vol. 13, 2020, p. 1407. doi:10.3390/en13061407.

Lin, Jianhua. "Divergence Measures Based on the Shannon Entropy." *IEEE Transactions on Information Theory*, vol. 37, 1991, pp. 145–151. doi:10.1109/18.61115.

Ljung, Greta M., and George E. P. Box. "On a Measure of Lack of Fit in Time Series Models." *Biometrika*, vol. 65, 1978, pp. 297–303. doi:10.1093/biomet/65.2.297.

Mahé, G., and P. Perez. "DLMS/COSEM for Smart Metering: An Open-Source Implementation and Performance Evaluation." *Sensors*, vol. 21, 2021, p. 3090. doi:10.3390/s21093090.

Martins Rodrigues, F., et al. "Short-Term Load Forecasting of Electricity Demand for the Residential Sector Based on Modelling Techniques: A Systematic Review." *Energies*, vol. 16, 2023, p. 4098. doi:10.3390/en16104098.

Olfati-Saber, Reza, et al. "Consensus and Cooperation in Networked Multi-Agent Systems." *Proceedings of the IEEE*, vol. 95, 2007, pp. 215–233. doi:10.1109/JPROC.2006.887293.

Park, Chang Hwan, and Taekyung Kim. "Energy Theft Detection in Advanced Metering Infrastructure Based on Anomaly Pattern Detection." *Energies*, vol. 13, 2020, p. 3832. doi:10.3390/en13153832.

Rosner, Bernard. "Percentage Points for a Generalized ESD Many-Outlier Procedure." *Technometrics*, vol. 25, 1983, pp. 165–172. doi:10.2307/1268549.

Said, Said E., and David A. Dickey. "Testing for Unit Roots in Autoregressive-Moving Average Models of Unknown Order." *Biometrika*, vol. 71, 1984, pp. 599–607. doi:10.1093/biomet/71.3.599.

CC BY

REMUNOM
ISSN 2178-6925

Received: 24/07/2026 - **Accepted:** 26/08/2026
Vol: 13.19
DOI: 10.66104/vyrz3362
Pages: 1-21

Shapiro, Samuel S., and Martin B. Wilk. "An Analysis of Variance Test for Normality (Complete Samples)." *Biometrika*, vol. 52, 1965, pp. 591–611. doi:10.1093/biomet/52.3-4.591.

Wald, Abraham, and Jacob Wolfowitz. "An Exact Test for Randomness in the Non-Parametric Case Based on Serial Correlation." *Annals of Mathematical Statistics*, vol. 14, 1943, pp. 378–388. doi:10.1214/aoms/1177731358.

Zheng, Zhuang, et al. "Wide and Deep Convolutional Neural Networks for Electricity Theft Detection to Secure Smart Grid." *IEEE Transactions on Industrial Informatics*, vol. 14, 2018a, pp. 1606–1615. doi:10.1109/TII.2017.2785963.

Zheng, Zhuang, et al. *SGCC Non-Technical Loss Dataset*. GitHub, 2018b, github.com/henryRDlab/ElectricityTheftDetection.

CC BY