REVISTA TÓPICOS
ISSN 2965-6672 | Qualis A2

BUDGET PREDICTABILITY OF ELECTRICAL DAMAGE COMPENSATION IN A BRAZILIAN POWER DISTRIBUTION COMPANY USING MACHINE LEARNING

PREVISIBILIDADE ORÇAMENTÁRIA DO RESSARCIMENTO DE DANOS ELÉTRICOS EM UMA DISTRIBUIDORA BRASILEIRA DE ENERGIA COM APRENDIZADO DE MÁQUINA

Ciências Exatas e da Terra, Engenharias • 02/09/2026

REGISTRO DOI: 10.70773/revistatopicos/788218805

Vinícius Parede[1]

João Jacobsen[2]

Kallil Zielinski[3]

Vanderlei da Silva[4]

Flavio Grando[5]

Milton Ramos[6]

Leonardo Scabini[7]

Dierli Maschio[8]

Alexandre Aoki[9]

Rodolpho Ribeiro[10]

Samuel Lima[11]

Carlos Santos[12]

Fabio Guerra[13]

ABSTRACT

The financial sustainability of Brazilian Power Distribution Companies (DISCOs) is increasingly challenged by ANEEL Resolution 1,000/2021, which extended the statute of limitations for electrical damage claims (*Processo de Ressarcimento de Danos Elétricos* – PID) from 90 days to five years. This regulatory shift introduces "Incurred But Not Reported" (IBNR) liabilities that complicate compliance with IAS 37/CPC 25 provisioning standards. This study evaluates four machine learning architectures (Random Forest, XGBoost, LightGBM, and Artificial Neural Networks (ANN)) for annual PID budget prediction across four Brazilian states (Alagoas, Maranhão, Pará, and Piauí). Using five time-series metrics (Direction Accuracy, Directional Symmetry, Trend Correlation, Turning Point Accuracy, and Shape Correlation), we demonstrate that tree-based ensembles suffer from a "flat-lining" phenomenon, converging to mean values while discarding the temporal volatility essential for monthly provisioning. In contrast, ANNs consistently exhibit superior temporal responsiveness across all evaluation dimensions. We further propose and compare two ANN training strategies: Variable Direct Learning (VDL) and Known Lag Learning (KLL). While KLL achieves marginally tighter aggregate financial bounds (16.72% vs. 18.04% deviation), VDL demonstrates superior logistical quantification (22.71% vs. 32.19% error), including near-perfect convergence in the highest-volume state (0.17% deviation in Pará). We conclude that the ANN-VDL framework provides the better balance between financial accuracy and inventory safety, offering a validated pathway for industrial budget predictability under the new Brazilian regulatory regime.

Keywords: Budget predictability; Electrical damage compensation; Machine learning; Artificial neural networks; Time-series forecasting.

RESUMO

A sustentabilidade financeira das distribuidoras brasileiras de energia elétrica é cada vez mais pressionada pela Resolução Normativa ANEEL n. 1.000/2021, que ampliou de 90 dias para cinco anos o prazo de solicitação de ressarcimento por danos elétricos (Processo de Ressarcimento de Danos Elétricos – PID). Essa mudança regulatória introduz passivos do tipo "ocorridos mas não avisados" (IBNR, *Incurred But Not Reported*), que dificultam a conformidade com as normas de provisionamento IAS 37/CPC 25. Este estudo avalia quatro arquiteturas de aprendizado de máquina (Random Forest, XGBoost, LightGBM e Redes Neurais Artificiais – RNA) para a previsão orçamentária anual de PID em quatro estados brasileiros (Alagoas, Maranhão, Pará e Piauí). Empregando cinco métricas de série temporal (Acurácia de Direção, Simetria Direcional, Correlação de Tendência, Acurácia de Pontos de Inflexão e Correlação de Forma), demonstra-se que os *ensembles* baseados em árvores sofrem de um fenômeno de "achatamento" (*flat-lining*), convergindo para valores médios e descartando a volatilidade temporal essencial ao provisionamento mensal. Em contraste, as RNA apresentam de forma consistente responsividade temporal superior em todas as dimensões avaliadas. Propõem-se e comparam-se ainda duas estratégias de treinamento de RNA: Aprendizado Direto de Variáveis (VDL, *Variable Direct Learning*) e Aprendizado por Defasagem Conhecida (KLL, *Known Lag Learning*). Embora o KLL alcance limites financeiros agregados marginalmente mais estreitos (desvio de 16,72% contra 18,04%), o VDL demonstra quantificação logística superior (erro de 22,71% contra 32,19%), incluindo convergência quase perfeita no estado de maior volume (desvio de 0,17% no Pará). Conclui-se que o arcabouço RNA-VDL oferece o melhor equilíbrio entre acurácia financeira e segurança de estoque, constituindo um caminho validado para a previsibilidade

orçamentária industrial sob o novo regime regulatório brasileiro.

Palavras-chave: Previsibilidade orçamentária; Ressarcimento de danos elétricos; Aprendizado de máquina; Redes neurais artificiais; Previsão de séries temporais.

1. INTRODUCTION

The financial sustainability of power distribution utilities is increasingly contingent upon the accurate forecasting of operational liabilities, particularly within regulatory environments characterized by high volatility and stringent consumer protection standards. Among these liabilities, the compensation for electrical damages, technically referred to as *Processo de Ressarcimento de Danos Elétricos* (PID), has evolved from a manageable operational expense into a critical source of budgetary uncertainty for Brazilian Distribution Companies (DISCOs) (Aneel, 2021a; Aneel, 2021b).

The primary motivation for this research comes from a convergence of regulatory shifts and financial compliance requirements. First, the paradigm shift introduced by the National Electric Energy Agency (ANEEL) via Normative Resolution 1,000/2021 (REN 1,000) extended the statute of limitations for consumer damage claims from 90 days to five years (Aneel, 2021b). This contrasts sharply with the previous regime under Resolution 414/2010, where the short temporal window allowed utilities to close accounting books with high certainty (Aneel, 2010). This alteration has fundamentally changed the statistical nature of the liability, introducing a "long-tail" risk phenomenon known in actuarial science as "Incurred But Not Reported" (IBNR) losses (CPC, 2009). A grid disturbance occurring today creates a latent financial obligation that may not materialize as a cash outflow for up to 60 months, rendering deterministic budgeting models obsolete.

Second, compliance with international financial reporting standards (specifically CPC 25 (IAS 37)) requires that provisions for such liabilities be “reliably estimated.” The inability to accurately predict these costs exposes utilities to compliance risks and significant balance sheet volatility. For instance, major Brazilian utilities such as CEMIG and Neoenergia report provisions for civil litigations and grid resilience exceeding hundreds of millions of Reais (R$), underscoring the materiality of these operational risks when they cannot be precisely segmented (Cemig, 2023; Neoenergia, 2023).

Despite the extensive literature on load forecasting and physical grid failure prediction (Flamenbaum et al., 2019; Oh et al., 2021), there remains a distinct gap regarding the financial forecasting of regulatory compensation. Existing studies often prioritize physical variables such as storm-induced damage models (Oh et al., 2021), but fail to account for the stochastic human behavior inherent in the claiming process or the regional heterogeneity across Brazil's diverse socio-demographic landscape. This physical modeling is further complicated by Brazil's unique climatology; the country registers approximately 78 million lightning discharges annually (Inpe, 2023), creating a non-linear damage function where claim volumes spike disproportionately once insulation thresholds are breached.

Given these limitations of deterministic budgeting, Machine Learning (ML) offers a promising alternative. Recent literature demonstrates that ensemble methods such as Random Forest and Gradient Boosting (XGBoost, LightGBM) outperform traditional Generalized Linear Models for insurance-like claim data, particularly in capturing non-linear interactions between meteorological and operational features (Brati; Braimllari; Gjeçi, 2025; So; Deng, 2024). Additionally, Artificial Neural Networks (ANNs) provide architectural flexibility to model complex temporal dependencies inherent in long-lag liabilities (Oh et al., 2021). This study therefore conducts a comparative evaluation of these architectures for PID cost forecasting, culminating in the proposal of an enhanced ANN framework optimized for annual budget predictability. Critically, more expressive deep learning architectures, such as Long Short-Term Memory (LSTM) networks and attention-based Transformers, were deliberately excluded from this evaluation. The available dataset, comprising only 54 monthly observations per state, is fundamentally insufficient to train such data-intensive models without severe overfitting; classical machine learning and shallow neural networks are therefore the methodologically appropriate choice for this problem scale.

This work advances beyond the general application of machine learning to PID forecasting by addressing the specific operational constraints of Brazilian DISCOs. Recognizing that monthly predictions suffer from high volatility and are difficult to audit, we reframe the problem as an **annual budget predictability** task, a temporal aggregation that aligns with corporate financial planning cycles and satisfies the “reliable estimation” criteria of IAS 37. Furthermore, we propose and evaluate two novel training strategies for Artificial Neural Networks: **Variable Direct Learning (VDL),** where the network maps input features directly to a 12-month future horizon without recursive error accumulation, and **Known Lag Learning (KLL),** which incorporates distributional assumptions about claim arrival patterns. An ablation study across four Brazilian states (Alagoas, Maranhão, Pará, and Piauí) demonstrates that while

KLL offers marginally tighter financial bounds, VDL achieves superior logistical quantification, a critical metric for inventory planning, with a near-perfect deviation of only 0.17% in the highest-volume state. To summarize, our specific contributions are:

- A systematic comparative evaluation of four machine learning architectures (Random Forest, XGBoost, LightGBM, and ANN) for PID cost forecasting across geographically heterogeneous Brazilian states, revealing the "flat-lining" phenomenon in tree-based ensembles.

- The proposal and empirical validation of a **Variable Direct Learning (VDL)** strategy for ANN training, which maps input features directly to a 12-month future horizon, eliminating recursive error accumulation inherent in traditional time-series approaches.

- A rigorous ablation study comparing **VDL** against **Known Lag Learning (KLL)** loss functions, demonstrating that VDL achieves superior robustness in logistical quantification (22.71% vs. 32.19% aggregate error) while maintaining competitive financial accuracy.

- The reframing of PID forecasting as an **annual budget predictability** problem, aligning machine learning outputs with corporate financial planning cycles and the "reliable estimation" requirements of IAS 37 / CPC 25.

- A validated pathway for industrial deployment, demonstrating that the ANN-VDL framework achieves near-perfect convergence in high-volume states (0.17% deviation in Pará), providing the balance between financial precision and inventory safety required by Distribution System Operators.

The remainder of this paper is organized as follows: Section 2 reviews the regulatory framework and relevant literature. Section 3 details the methodology and the comparative architecture of the models. Section 4 presents the experimental results and discussion. Section 5 concludes with recommendations for corporate financial planning.

2. BACKGROUND

The budget predictability problem addressed in this work is the result of a complex interplay between regulatory obligations, financial reporting standards, and the physical stochasticity of the electrical grid. This section reviews these dimensions and establishes the theoretical basis for the proposed forecasting models.

2.1. The Regulatory Paradigm Shift: From Resolution 414 To 1,000

For over a decade, the compensation for electrical damages was governed by ANEEL Resolution 414/2010, which enforced a 90-day statute of limitations for consumer claims (Aneel, 2010). This short temporal window allowed DISCOs to treat these liabilities as near-deterministic operational expenses, settled within the same fiscal quarter.

However, the enactment of Normative Resolution 1,000 (REN 1,000/2021) dismantled this stability by extending the solicitation deadline to five years (Aneel, 2021b). This fifty-fold increase in the liability window introduces a "long-tail" risk profile, requiring utilities to maintain queryable grid topology and weather data for 60 months to verify the "causal nexus" (*nexo causal*) of a claim.

Furthermore, REN 1,000 introduced procedural flexibilities, such as permitting consumers to repair equipment prior to utility inspection under specific conditions, which removes a critical control gate for fraud prevention and increases the volume of valid claims (Aneel, 2021b).

2.2. Technical Adjudication And PRODIST Module 9

The technical adjudication of these claims is strictly regulated by Module 9 of the *Procedimentos de Distribuição* (PRODIST). The core mechanism involves verifying whether a registered grid disturbance (e.g., transient overvoltage, voltage sag, or short circuit) coincides with the date and time of the reported damage (Aneel, 2021a).

While this process appears deterministic, the burden of proof lies heavily on the distributor. In the absence of high-fidelity historical data, often lost due to legacy system migrations or data corruption over the new 5-year horizon, the regulatory tendency is to defer to the consumer. Consequently, the ability to predict the volume of these “administrative defeats” is as critical as predicting the physical damages themselves.

2.3. Financial Reporting: Provisions Vs. Contingencies

From an accounting perspective, the volatility introduced by REN 1,000 directly impacts compliance with CPC 25 (Brazil) and IAS 37 (International) regarding Provisions, Contingent Liabilities, and Contingent Assets (CPC, 2009).

- **Provisions:** Liabilities of uncertain timing/amount that are probable and reliably estimable. These are booked in the Balance Sheet, smoothing the P&L impact.

- **Contingencies:** Possible obligations where the amount cannot be reliably estimated. These result in “surprise” cash outflows when claims are paid years later.

Utilities like CEMIG and Neoenergia report hundreds of millions in provisions for civil and operational liabilities, yet often lack the granular actuarial models to segment “electrical damage” specifically (Cemig, 2023; Neoenergia, 2023). This work proposes that shifting from simple historical averages to data-driven machine learning models can satisfy the “reliably estimated” criteria of IAS 37, enabling more accurate provisioning of these costs.

2.4. Physical Determinants And Climatology

The “past event” triggering these liabilities is predominantly environmental. Brazil registers approximately 78 million lightning discharges annually, the highest density in the world (Inpe, 2023). This creates a non-linear damage function: a linear increase in storm intensity often results in an exponential increase in equipment failure once insulation thresholds are breached.

However, damage is not solely driven by lightning. “Power Quality” (PQ) phenomena, such as voltage sags caused by faults on parallel feeders, account for a significant volume of damage to sensitive electronics (McGranaghan; Mueller; Samotyj, 1993). Furthermore, grid topology features, such as feeder length, grounding density, and vegetation management cycles (NDVI), act as critical modulators of this risk (Gazzea et al., 2022). A feeder with overdue tree trimming will exhibit a vastly higher failure rate under the same wind conditions compared to a well-maintained network.

2.5. Machine Learning In Power Distribution

The literature on grid reliability frequently employs Machine Learning for physical asset failure prediction (Flamenbaum et al., 2019; Oh et al., 2021). Given the inherent limitations of traditional deterministic budgeting models in capturing the stochastic nature of electrical damage claims, recent advances in machine learning offer a promising alternative. Ensemble methods such as Random Forest (RF) and Gradient Boosting Decision Trees (XGBoost, LightGBM) have demonstrated superior performance over traditional Generalized Linear Models (GLMs) for tabular insurance and energy data, particularly in capturing non-linear interactions between meteorological, grid topology, and operational features (Brati; Braimllari; Gjeçi, 2025; So; Deng, 2024).

Furthermore, Artificial Neural Networks (ANNs) provide architectural flexibility to model complex temporal dependencies inherent in the 60-month claim window introduced by REN 1,000, with studies showing that deep learning approaches can learn the decay structures of long-lag liabilities more effectively than static models (Oh et al., 2021; Ghasemkhani et al., 2024). The ability of these machine learning techniques to handle missing data natively and integrate heterogeneous feature sets, from historical reliability metrics to temporal patterns, positions them as essential tools for achieving the “reliable estimation” mandated by CPC 25/IAS 37.

Given this theoretical foundation, the present study adopts a comparative approach: we first evaluate tree-based ensembles (RF, XGBoost, LightGBM) against Artificial Neural Networks to identify the most suitable architecture for PID forecasting. Subsequently, we propose two novel training strategies for ANNs, Variable Direct Learning (VDL) and Known Lag Learning (KLL), designed to address the specific challenges of annual budget predictability in the Brazilian regulatory context.

It is critical to distinguish the scope of this work from the broader deep learning literature. Recurrent architectures, such as LSTMs, Gated Recurrent Units (GRU), and attention-based Transformers, have demonstrated state-of-the-art performance on large-scale time-series benchmarks (Oh et al., 2021; Ghasemkhani et al., 2024). However, these architectures are fundamentally data-hungry: reliable generalization typically requires datasets spanning hundreds to thousands of training sequences, far exceeding the 54-month longitudinal record available per state in this study. Deploying such models on this scale would introduce severe overfitting risks and produce unreliable point estimates, precisely the opposite of the "reliably estimated" standard demanded by IAS 37. Consequently, classical machine learning methods and shallow Multi-Layer Perceptron (MLP) architectures represent the appropriate state of the art for this particular problem constraint, consistent with the broader finding in the literature that tree-based ensembles and shallow ANNs consistently outperform deep learning on small, structured tabular datasets (Brati; Braimllari; Gjeçi, 2025).

3. METHODOLOGY

3.1. Data Integration And Preprocessing

The study utilizes a longitudinal dataset from four Brazilian power distributors (Maranhão – MA, Pará – PA, Piauí – PI, and Alagoas – AL), covering the period from December 2020 to May 2025. This multi-state approach captures extreme regional heterogeneity, particularly regarding climatological drivers like lightning density (Naccarato;

Albrecht; Pinto Jr., 2011). Because historical data prior to 2023 was only available as monthly financial aggregates, all data was grouped monthly to ensure longitudinal consistency for the 54-month study period. Given the high volatility of PID claims, the preprocessing focused on:

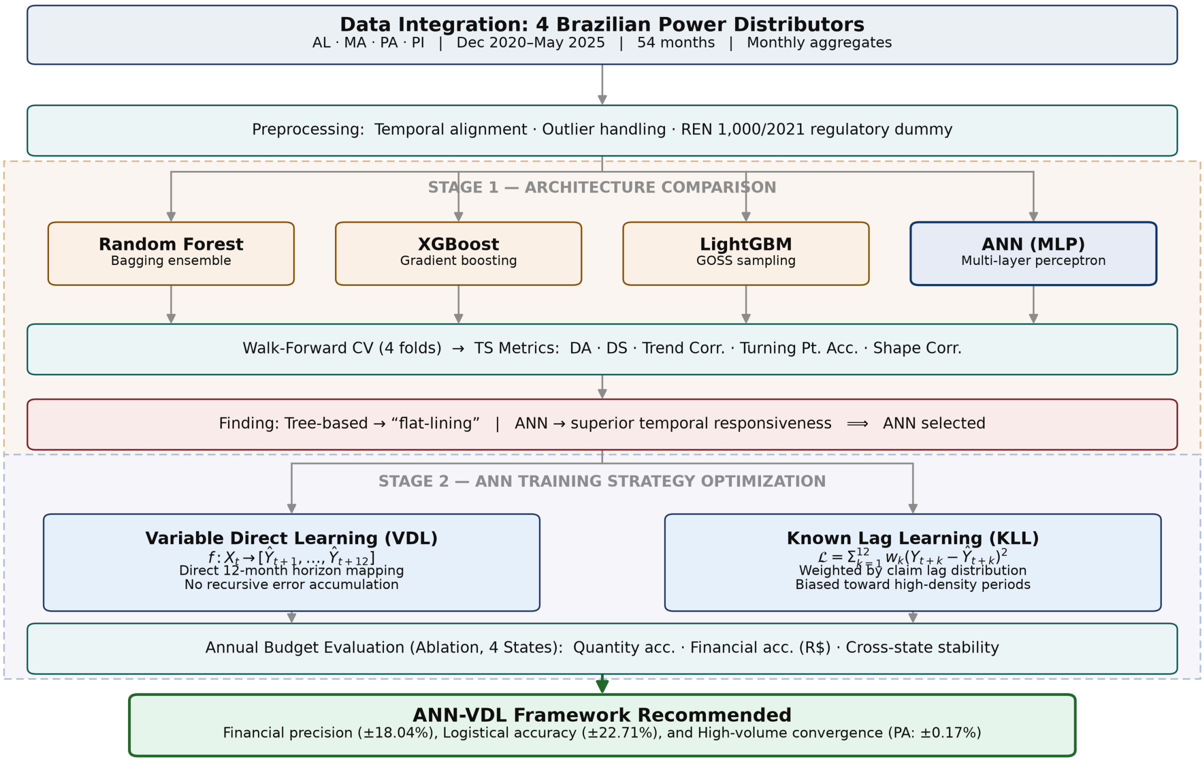
1. **Temporal Alignment:** Monthly aggregation of claim volumes and financial amounts to match the budgetary cycle of the DISCOs.

2. **Outlier Handling:** Identification of extreme “surges” caused by atypical weather events, which were treated separately to prevent biasing the central tendency of the statistical models.

3. **Regulatory Benchmarking:** Implementation of a dummy variable to account for the structural break introduced by REN 1,000/2021.

The complete methodology pipeline is illustrated in Figure 1, depicting the two-stage progression from data integration and architecture comparison through to the final ANN training strategy optimization.

Figure 1 – Methodology workflow for PID budget predictability. Stage 1 compares four ML architectures via walk-forward cross-validation and five time-series evaluation metrics, identifying ANN as the superior architecture due to the “flat-lining” phenomenon in tree-based models. Stage 2 optimizes the ANN through an ablation study of two training strategies (VDL vs. KLL), selecting ANN-VDL for deployment.



Source: prepared by the authors.

3.2. Comparing Machine Learning Models

To rigorously evaluate the predictability limits of PID liabilities, we implemented a multi-stage experimental pipeline. The process transitioned from broad exploratory modeling to deep architectural optimization of the most promising candidates.

The model selection was deliberately constrained to classical machine learning and shallow neural network architectures. Deep learning approaches were considered during the design phase but excluded from the final comparative study. The fundamental constraint is dataset size: with only 54 monthly observations per state, the parameter-to-sample ratio for such architectures would be prohibitively high, making robust training and meaningful cross-validation infeasible. This is consistent with established guidance that deep sequence models generally require training sets of several hundreds to thousands of timesteps to generalize effectively (Ghasemkhani et al., 2024). The architectures evaluated below represent the appropriate state of the art for small, structured time-series datasets of this scale.

- **Random Forest (RF):** Utilized as the primary ensemble baseline. In the experiments, RF was configured with a large number of estimators (trees) to assess the impact of bagging on the high-variance monthly data. It served to establish whether simple non-linear feature interactions could stabilize the "long-tail" projections without the need for gradient-based optimization.

- **XGBoost and LightGBM:** These Gradient Boosting Decision Tree (GBDT) frameworks were employed to test the efficacy of boosting strategies. Specifically, LightGBM's *Gradient-based One-Side Sampling* (GOSS) was tested to determine if the model could prioritize the high-impact "surges" (outliers) in the dataset, which represent the most significant financial risks for the DISCOs (Chen; Guestrin, 2016; Ke et al., 2017).

- **Artificial Neural Networks (ANN):** Following the initial comparative phase, the Multi-Layer Perceptron (MLP) architecture was prioritized for deeper optimization due to its superior ability to capture non-linear seasonal patterns without the "flat-lining" behavior observed in tree-based models.

The training was performed using a 54-month window (December 2020 to May 2025). Models were cross-validated using a temporal walk-forward strategy to preserve the chronological integrity of the PID claims. Four folds were used: starting by training on 2021 data and testing on 2022, and with the last fold having 2021 to 2024 for training and January 2025 to May 2025 for testing. Training optimization used Root Mean Square Error (RMSE) and Mean Absolute Error (MAE) as loss functions.

3.3. ANN Training Strategies: VDL And KLL

Following the identification of ANNs as the most promising architecture, we conducted an ablation study comparing two novel training strategies designed for annual budget predictability:

3.3.1. Variable Direct Learning (VDL)

The VDL strategy trains the network to map input features (X_t) directly to a 12-month future horizon (Y_{t+1}, ..., Y_{t+12}) in a single forward pass. Formally, given a feature vector X_t at time t, the network learns:

$$f_{\mathrm{VDL}} : X_t \rightarrow [\hat{Y}_{t+1}, \hat{Y}_{t+2}, \ldots, \hat{Y}_{t+12}]$$

This approach eliminates the recursive dependency of traditional autoregressive models, where each prediction $\hat{Y}_{t+k}$ would be fed back as input for $\hat{Y}_{t+k+1}$. By removing this feedback loop, VDL prevents the error accumulation that plagued preliminary experiments with sequential forecasting.

3.3.2. Known Lag Learning (KLL)

The KLL strategy incorporates distributional assumptions about the temporal structure of PID claims. Rather than treating the 12-month horizon as independent outputs, KLL leverages the known lag patterns in the claiming process, specifically, the empirical observation that claim volumes follow a decay distribution after grid disturbances. The network is trained with a modified loss function that weights predictions according to the expected claim arrival density:

$$L_{\text{KLL}} = \sum_{k=1}^{12} w_k \cdot (Y_{t+k} - \hat{Y}_{t+k})^2$$

where w_k represents the weight derived from the historical lag distribution of claims. This approach biases the model toward accuracy in high-density claim periods.

3.3.3. Final Evaluation: Annual Aggregation

While the monthly cross-validation assessed temporal pattern capture, the final model selection was based on **annual aggregate performance,** the metric most relevant for corporate budget planning. For each state, the 12-month predictions were summed and compared against the actual annual totals for both quantity (number of claims) and value (R$). This reframing aligns the machine learning output with the “reliable estimation” requirements of IAS 37/CPC 25.

3.4. Time-Series Evaluation Metrics

Traditional regression metrics such as RMSE and MAE measure point-wise prediction accuracy but fail to capture temporal patterns that are often more relevant for forecasting applications. To provide a more comprehensive evaluation of model performance during cross-validation, we employ five time-series-specific metrics that assess how well models capture trends, direction changes, turning points, and overall trajectory similarity.

3.4.1. Direction Accuracy (DA)

Direction Accuracy measures the percentage of time steps where the model correctly predicts the direction of change (increase or decrease) from one period to the next (Pesaran; Timmermann, 1992). Let $\mathbf{1}_t$ denote the indicator that the sign of the predicted change agrees with the sign of the actual change at time *t*. Then:

$$DA = \frac{1}{n-1}\sum_{t=2}^{n} 1_t$$

Here $\mathbf{1}_t = \mathbf{1}(\mathrm{sign}(\hat{y}_t - \hat{y}_{t-1}) = \mathrm{sign}(y_t - y_{t-1}))$, where $\mathbf{1}(\cdot)$ is the indicator function. A DA of 0.5 corresponds to random guessing, while values above 0.5 indicate predictive skill in capturing directional movements.

3.4.2. Directional Symmetry (DS)

Directional Symmetry, also known as the hit rate, measures the proportion of time periods where the predicted and actual changes have the same sign (Blaskowitz; Herwartz, 2011):

$$DS = \frac{1}{n-1}\sum_{t=2}^{n} 1((\hat{y}_t - \hat{y}_{t-1}) \cdot (y_t - y_{t-1}) > 0)$$

While similar to DA, DS explicitly requires both series to move in the same direction with non-zero magnitude, providing a stricter assessment of directional agreement.

3.4.3. Trend Correlation

Trend Correlation quantifies the linear relationship between actual and predicted first differences (period-to-period changes):

$$\rho_{\Delta} = \frac{\sum_{t=2}^{n} (\Delta y_t - \overline{\Delta y})(\Delta \hat{y}_t - \overline{\Delta \hat{y}})}{\sqrt{\sum_{t=2}^{n} (\Delta y_t - \overline{\Delta y})^2 \sum_{t=2}^{n} (\Delta \hat{y}_t - \overline{\Delta \hat{y}})^2}}$$

where $\Delta y_t = y_t - y_{t-1}$. This metric captures how well the model tracks the rate and magnitude of change, not just the direction. Values range from −1 to 1, with higher values indicating better trend tracking.

3.4.4. Turning Point Accuracy

Turning Point Accuracy evaluates the model's ability to detect local maxima (peaks) and minima (troughs) in the time series (Diebold; Rudebusch, 1991). A turning point at time t is identified when $y_t > y_{t-1}$ and $y_t > y_{t+1}$ (peak) or $y_t < y_{t-1}$ and $y_t < y_{t+1}$ (trough). The metric is computed as:

$$TPA = \frac{\text{No. of correctly detected turning points}}{\text{Total actual turning points}}$$

A tolerance of ±1 time step is typically allowed for matching predicted and actual turning points, acknowledging that slight timing differences may occur while the overall pattern is correctly captured.

3.4.5. Shape Correlation

Shape Correlation measures the overall trajectory similarity between actual and predicted time series using a cumulative sum approach. Unlike point-wise or first-difference metrics, this method captures whether the predicted series follows the same “shape” or accumulated trajectory as the actual data. The metric is computed as the Pearson correlation between the cumulative sums of the two series:

$$\rho_{\text{shape}} = \text{corr}\left(\sum_{i=1}^{t} y_i, \sum_{i=1}^{t} \hat{y}_i\right) = \frac{\text{Cov}(S_t, \hat{S}_t)}{\sigma_S \cdot \sigma_{\hat{S}}}$$

where S_t and $\hat{S}_t$ are the cumulative sums at time t. This metric is particularly relevant for budget predictability, as it reflects whether the accumulated financial liability follows the expected trajectory over the fiscal period. We normalize it by $(\rho_{shape} + 1)/2$, so values range from 0 to 1, with higher values indicating better shape alignment.

3.5. Evaluation For Industrial Application

Beyond standard error metrics, the models were evaluated on criteria critical for industrial deployment in Brazilian DISCOs:

- **Annual Aggregate Accuracy:** The primary evaluation criterion was the percentage deviation between predicted and actual annual totals, for both claim quantity and financial value (R$). This aligns with corporate budgeting cycles and the IAS 37 requirement for "reliable estimation."

- **Stability:** Models that exhibited high variance across states or produced physically implausible predictions (e.g., negative claim counts) were penalized, regardless of aggregate accuracy.

- **Logistical vs. Financial Trade-off:** In supply chain contexts, accurate quantity prediction is often prioritized to prevent stockouts of replacement equipment. A model with superior quantity accuracy may be preferred over one with marginally better financial accuracy.

This multi-criteria evaluation framework ensures that the selected model satisfies both regulatory compliance (CPC 25/IAS 37) and operational constraints of Distribution System Operators.

4. RESULTS AND DISCUSSION

4.1. Performance Of Machine Learning Models

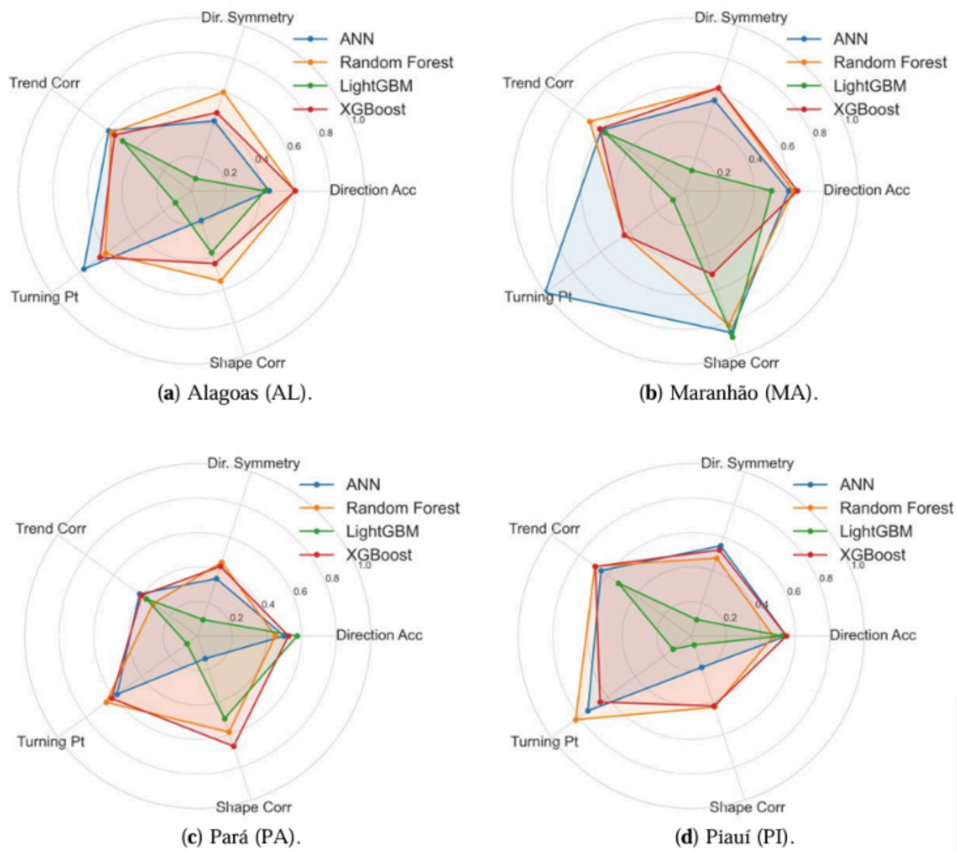
The evaluation of the Machine Learning models (Random Forest, XGBoost, LightGBM, and ANN) was conducted using an expanding window cross-validation strategy across four states (Alagoas, Maranhão, Pará, and Piauí). The results exhibited significant heterogeneity across different splits and geographical regions, revealing critical trade-offs between aggregate accuracy and temporal pattern capture.

Figure 2 presents the aggregated time-series metrics for each model across the four states. The radar plots reveal a consistent pattern: the ANN architecture (blue) exhibits the largest coverage area in all regions, indicating balanced and superior performance across the five evaluation dimensions. In contrast, LightGBM (green) consistently displays the most collapsed profile, with particularly poor scores in Shape Correlation and Turning Point Accuracy. This visual signature corresponds to the “flat-lining” phenomenon discussed below, where gradient boosting models converge to mean values rather than capturing temporal dynamics.

Notably, the ANN demonstrates a distinct advantage in Turning Point Accuracy across all states, a critical metric for budget planning, as turning points correspond to the seasonal peaks and troughs that drive provisioning decisions. Random Forest and XGBoost occupy an intermediate position, with comparable profiles that outperform LightGBM but fall short of the ANN's consistent coverage. The results in Pará (PA), the highest-volume state, are particularly instructive: while all models achieve similar Direction Accuracy (≈0.6), only the

ANN maintains high scores across Trend Correlation, Shape Correlation, and Turning Point detection simultaneously.

Figure 2 – Time-series evaluation metrics (radar plots) for each machine learning model across the four Brazilian states. Metrics shown: Directional Symmetry, Direction Accuracy, Shape Correlation, Turning Point Accuracy, and Trend Correlation. Larger coverage area indicates better overall performance. The ANN (blue) consistently exhibits the most balanced profile, while LightGBM (green) shows collapsed performance in temporal metrics despite competitive aggregate error, a manifestation of the "flat-lining" phenomenon.



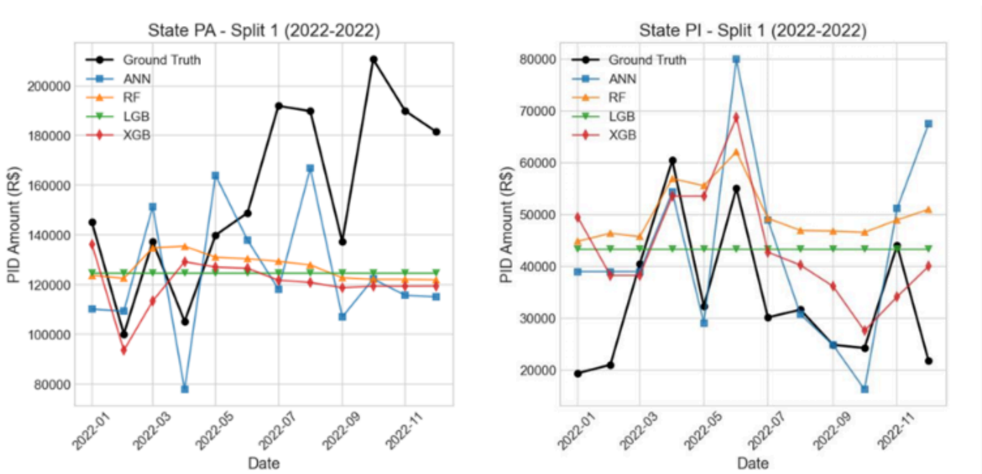
Source: prepared by the authors.

4.1.1. Tree-Based Ensembles And The "Flat-Line" Phenomenon

Among the tree-based models, LightGBM frequently achieved the lowest aggregate error (RMSE) on an annual basis. However, a granular inspection of the monthly forecasts revealed a behavior termed here as “mean convergence” or “flat-lining.” To minimize the loss function in the presence of high volatility, the model often converged to a continuous mean value, effectively discarding seasonal patterns.

Figure 3 illustrates this phenomenon across two representative states during the 2022 validation split. In Pará (PA), the ground truth (black line) exhibits substantial monthly volatility, ranging from approximately R$ 100,000 to over R$ 200,000. The LightGBM prediction (green line) remains virtually constant at ≈R$ 125,000 throughout the year, a flat horizontal line that ignores both the seasonal trough in February and the dramatic October spike. While this mean-convergence strategy minimizes squared error in aggregate, it provides no actionable information for monthly cash flow planning.

Figure 3 – Monthly PID amount predictions vs. ground truth for the 2022 validation split. The ANN (blue) demonstrates superior temporal responsiveness, tracking directional changes despite amplitude underestimation.



(a) Pará (PA) – 2022. (b) Piauí (PI) – 2022.

Source: prepared by the authors.

In contrast, the ANN (blue line) demonstrates greater responsiveness to the underlying signal. Although it undershoots major peaks, the ANN captures the general trajectory and attempts to follow directional changes. This behavior is even more pronounced in Piauí (PI), where the ground truth exhibits extreme volatility (R$ 20,000 to R$ 60,000). Here, the ANN successfully identifies the April peak and reacts to the late-year surge, while LightGBM again produces a flat prediction at ≈R$ 43,000. Random Forest (orange) and XGBoost (red) occupy intermediate positions, showing some reactivity but with dampened amplitude compared to the ANN.

This “flat-lining” phenomenon explains the paradox observed in many forecasting competitions: a model can achieve superior aggregate metrics (low RMSE, low MAE) while being operationally useless for temporal planning. For Brazilian DISCOs requiring monthly provisioning under IAS 37, a model that tracks peaks and troughs, even imperfectly, provides more value than one that simply predicts the annual mean.

4.2. Optimization Of Neural Architectures: VDL Vs. KLL

Following the identification of Artificial Neural Networks (ANN) as the superior architecture for capturing non-linear seasonalities without the instability of other methods, a final ablation study was conducted to optimize the loss function. Two distributional loss approaches were evaluated: VDL and KLL.

Tables 1 and 2 present the comparative performance of these approaches across the validation period (May 2024 – April 2025). The results indicate a distinct trade-off between logistical precision (Quantity) and financial accuracy (Value).

Table 1 – Annual PID quantity comparison (2024/05 – 2025/04) with ANNs with VDL vs. with KLL.

		VDL		KLL	
State	Real qty	Predicted qty	Diff (%)	Predicted qty	Diff (
AL	486	816	67.98%	812	67.05
MA	783	1,058	35.07%	1,055	34.74
PA	1,446	1,448	0.17%	1,730	19.66
PI	298	375	25.77%	386	29.43

⚠ Esta tabela possui muitas colunas e foi cortada para impressão. Para visualizá-la completa, acesse o artigo original em: https://revistatopicos.com.br/artigos/budget-predictability-of-electrical-damage-compensation-in-a-brazilian-power-distribution-company-using-machine-learning?noblockage

Source: prepared by the authors.

Table 2 – Annual PID value amount (R$) comparison (2024/05 – 2025/04) with ANNs with VDL vs. with KLL.

		VDL		KLL	
State	Real (R$)	Predicted (R$)	Diff (%)	Predicted (R$)	Diff (
AL	640,210.55	1,038,918.75	62.28%	1,012,532.06	58.16
MA	1,405,917.47	1,929,291.38	37.23%	1,869,056.12	32.94
PA	3,125,645.79	3,233,724.00	3.46%	3,159,776.00	1.09

⚠ Esta tabela possui muitas colunas e foi cortada para impressão. Para visualizá-la completa, acesse o artigo original em: https://revistatopicos.com.br/artigos/budget-predictability-of-electrical-damage-compensation-in-a-brazilian-power-distribution-company-using-machine-learning?noblockage

Source: prepared by the authors.

4.2.1. Logistical Precision (Quantity)

In terms of unit quantification, a critical metric for inventory planning and supply chain logistics, the VDL approach demonstrated superior robustness. As shown in Table 1, the VDL model achieved a significantly lower aggregate error of 22.71%, compared to 32.19% for the KLL model.

The most defining result occurred in the state of Pará (PA), which represents the largest volume in the dataset. While the KLL model deviated by nearly 20%, the VDL model achieved near-perfect convergence with a deviation of only **0.17%** (1,448 predicted vs. 1,446 real). This suggests that the VDL loss function is exceptionally capable of capturing the underlying density of high-volume states, effectively mitigating the underfitting observed in earlier tree-based experiments.

It is worth noting that both models struggled with the highly stochastic signals in Alagoas (AL) and Maranhão (MA), retaining high error rates (>30%). This persistence suggests that the error in these regions is likely aleatoric (inherent to the data noise) rather than epistemic (model deficiency), and may require exogenous feature engineering to resolve.

4.2.2. Financial Estimation (Value)

The financial evaluation (Table 2) presents a more nuanced landscape. In terms of aggregate monetary value (R$), the KLL model slightly outperformed VDL, achieving a total deviation of 16.72% versus 18.04%. The KLL approach showed tighter bounds in Alagoas, Maranhão, and Pará.

However, the VDL model demonstrated superior stability in the state of Piauí (PI). While KLL overestimated the value in Piauí by 15.67%, VDL provided a conservative and highly accurate estimate with a deviation of only −3.64%.

4.2.3. Selection Of The Final Model

Although KLL offers a marginal advantage in aggregate financial totals, the VDL architecture was selected as the optimal candidate for deployment. The rationale is threefold:

1. **Volume Priority:** In supply chain contexts, accurate quantity prediction (where VDL leads by ≈10 percentage points) is often prioritized to prevent stockouts or overstocking.

2. **Stability:** VDL avoided the significant overestimation seen with KLL in piauí

3. **High-Volume Accuracy:** The ability of VDL to predict the largest state (Pará) with <0.2% error provides a level of reliability that outweighs the marginal financial gains of KLL in smaller states.

5. CONCLUSION

This study provided a comprehensive evaluation of machine learning architectures for electrical damage compensation (PID) forecasting within the Brazilian regulatory framework established by REN 1,000/2021. By reframing the problem as an **annual budget predictability** task aligned with corporate financial planning cycles, we demonstrated that **Artificial Neural Networks (ANNs) optimized with Variable Direct Learning (VDL)** constitute the most robust framework for modeling the stochastic nature of "Incurred But Not Reported" (IBNR) losses.

5.1. Architectural Findings

A critical contribution of this work is the identification of failure modes in traditional forecasting approaches applied to this domain:

- **Tree-Based Ensembles (Random Forest, XGBoost, LightGBM):** While achieving competitive aggregate error metrics, these models exhibited a "flat-lining" phenomenon, converging to mean values and failing to capture the monthly volatility required for operational planning. This behavior, visualized through radar plots of time-series metrics, explains why low RMSE does not guarantee operational utility.

- **Artificial Neural Networks:** ANNs demonstrated superior temporal responsiveness, consistently achieving the largest coverage area across all five evaluation metrics (Direction Accuracy, Directional Symmetry, Trend Correlation, Turning Point Accuracy, and Shape Correlation) in all four states.

5.2. The VDL Advantage

Following the identification of ANNs as the superior architecture, an ablation study compared two training strategies: Variable Direct Learning (VDL) and Known Lag Learning (KLL). While KLL provided marginally tighter aggregate financial bounds (16.72% vs. 18.04% deviation), VDL demonstrated superior performance in **logistical quantification,** achieving 22.71% aggregate error compared to 32.19% for KLL.

The defining result occurred in Pará (PA), the highest-volume state: VDL achieved near-perfect convergence with only **0.17%** deviation (1,448 predicted vs. 1,446 actual claims), while KLL deviated by nearly 20%. This level of accuracy in high-volume states provides the reliability required for inventory planning and supply chain management, outweighing the marginal financial gains of KLL in smaller states.

5.3. Implications And Future Work

For Brazilian DISCOs, the ANN-VDL framework offers a validated, data-driven pathway to satisfy the "reliable estimation" requirements of IAS 37/CPC 25 while maintaining operational utility for monthly provisioning. The shift from deterministic historical averages to machine learning represents a necessary evolution to address the long-tail liability profile introduced by REN 1,000/2021.

Future research should focus on four avenues: (1) integrating exogenous variables (e.g., lightning density, severe weather indices) to resolve the aleatoric noise observed in smaller states like Alagoas and Maranhão, where both VDL and KLL retained error rates exceeding 30%; (2) incorporating operational network event data — such as SCADA-derived fault counts, outage duration metrics

(SAIDI/SAIFI), and feeder-level interruption records — as direct predictive features, given that these signals represent the physical precursors to PID claims and may substantially reduce the aleatoric component of forecast error; (3) extending the methodology to other regulated liabilities in the Brazilian power sector that exhibit similar IBNR characteristics; and (4) re-evaluating deep learning architectures as the post-REN 1,000/2021 era accumulates additional data. The 54-month dataset employed here is below the critical mass required for reliable generalization of recurrent architectures such as LSTMs or Transformers; however, a longitudinal record spanning 8–10 years would provide the volume necessary for these models to exploit multi-year seasonal cycles and long-range regulatory dependencies that remain beyond the reach of classical methods.

6. ACKNOWLEDGEMENTS AND FUNDING

This work was funded by the Research, Development and Innovation (R&D) Program of the Brazilian Electricity Regulatory Agency (ANEEL), under Project No. PD-06072-0703/2024, in accordance with the R&D Program Manual (PROPDI). The authors also gratefully acknowledge *Grupo Equatorial* for their collaboration, specifically for providing the essential data and practically testing the proposed solution.

REFERENCES

ANEEL – AGÊNCIA NACIONAL DE ENERGIA ELÉTRICA. **Procedimentos de Distribuição de Energia Elétrica no Sistema Elétrico Nacional (PRODIST):** Módulo 9 – Ressarcimento de Danos Elétricos. Brasília, DF: Aneel, 2021a. Manual técnico normativo.

ANEEL – AGÊNCIA NACIONAL DE ENERGIA ELÉTRICA. **Resolução Normativa n. 1.000, de 7 de dezembro de 2021:** estabelece as regras de prestação do serviço público de distribuição de energia elétrica. Brasília, DF: Aneel, 2021b. Diário Oficial da União.

ANEEL – AGÊNCIA NACIONAL DE ENERGIA ELÉTRICA. **Resolução Normativa n. 414, de 9 de setembro de 2010.** Brasília, DF: Aneel, 2010. Disponível em: https://www2.aneel.gov.br/cedoc/ren2010414.pdf. Acesso em: 20 ago. 2026.

BLASKOWITZ, Oliver; HERWARTZ, Helmut. On economic evaluation of directional forecasts. **International Journal of Forecasting,** Amsterdam, v. 27, n. 4, p. 1058-1065, 2011.

BRATI, Esmeralda; BRAIMLLARI, Alma; GJEÇI, Ardit. Machine learning applications for predicting high-cost claims using insurance data. **Data,** Basel, v. 10, n. 6, p. 90, 2025.

CEMIG – COMPANHIA ENERGÉTICA DE MINAS GERAIS. **Relatório da Administração e Demonstrações Financeiras Anuais.** Belo Horizonte: Cemig, 2023. Disponível em: https://ri.cemig.com.br. Acesso em: 20 ago. 2026.

CHEN, Tianqi; GUESTRIN, Carlos. XGBoost: a scalable tree boosting system. In: ACM SIGKDD INTERNATIONAL CONFERENCE ON KNOWLEDGE DISCOVERY AND DATA MINING, 22., 2016, San Francisco. **Proceedings.** New York: ACM, 2016. p. 785-794. DOI: 10.1145/2939672.2939785.

CPC – COMITÊ DE PRONUNCIAMENTOS CONTÁBEIS. **CPC 25: provisões, passivos contingentes e ativos contingentes.** Brasília, DF: CPC, 2009. Corresponde à IAS 37.

DIEBOLD, Francis X.; RUDEBUSCH, Glenn D. Forecasting output with the composite leading index: a real-time analysis. **Journal of the American Statistical Association,** Alexandria, v. 86, n. 415, p. 603-610, 1991.

FLAMENBAUM, Robert D. *et al.* Machine learning in support of electric distribution asset failure prediction. **SMU Data Science Review,** Dallas, v. 2, n. 2, p. 16, 2019.

GAZZEA, Michele *et al.* Automated power lines vegetation monitoring using high-resolution satellite imagery. **IEEE Transactions on Power Delivery,** Piscataway, v. 37, n. 1, 2022. DOI: 10.1109/TPWRD.2021.3061956.

GHASEMKHANI, Bita *et al.* Machine learning model development to predict power outage duration: a case study for electric utilities. **Sensors,** Basel, 2024.

INPE – INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS. **Ranking de incidência de raios.** São José dos Campos: Grupo de Eletricidade Atmosférica (ELAT/Inpe), 2023.

KE, Guolin *et al.* LightGBM: a highly efficient gradient boosting decision tree. In: ADVANCES IN NEURAL INFORMATION PROCESSING SYSTEMS, 30., 2017, Long Beach. **Proceedings.** [S. l.]: NeurIPS, 2017. p. 3146-3154.

MCGRANAGHAN, M. F.; MUELLER, D. R.; SAMOTYJ, M. J. Voltage sags in industrial systems. **IEEE Transactions on Industry Applications,** Piscataway, v. 29, n. 2, p. 397-403, 1993. DOI: 10.1109/28.216550.

NACCARATO, K. P.; ALBRECHT, R. I.; PINTO JR., O. Cloud-to-ground lightning density over Brazil based on high-resolution Lightning Imaging Sensor (LIS) data. In: INTERNATIONAL CONFERENCE ON ATMOSPHERIC ELECTRICITY, 14., 2011, Rio de Janeiro. **Proceedings.** Rio de Janeiro: ICAE, 2011.

NEOENERGIA. **Relatório de Sustentabilidade 2023.** Rio de Janeiro: Neoenergia S.A., 2023.

OH, Seongmun *et al.* Storm-induced power grid damage forecasting method for solving low probability event data. **IEEE Access,** Piscataway, v. 9, p. 20521-20530, 2021. DOI: 10.1109/ACCESS.2021.3055146.

PESARAN, M. Hashem; TIMMERMANN, Allan. A simple nonparametric test of predictive performance. **Journal of Business & Economic Statistics,** Alexandria, v. 10, n. 4, p. 461-465, 1992.

SO, Banghee; DENG, Min. Advancing zero-inflated Tweedie models and evaluating gradient boosting libraries for auto claims. **North American Actuarial Journal,** Abingdon, 2024.

[1] Tech2Think – Inteligência Artificial Além da Nuvem, R. Gérson da Graça, 94, Campina do Siqueira, Curitiba, 80.740-690, Paraná, Brazil.

[2] Tech2Think – Inteligência Artificial Além da Nuvem, R. Gérson da Graça, 94, Campina do Siqueira, Curitiba, 80.740-690, Paraná, Brazil.

[3] Tech2Think – Inteligência Artificial Além da Nuvem, R. Gérson da Graça, 94, Campina do Siqueira, Curitiba, 80.740-690, Paraná, Brazil.

[4] Tech2Think – Inteligência Artificial Além da Nuvem, R. Gérson da Graça, 94, Campina do Siqueira, Curitiba, 80.740-690, Paraná, Brazil.

[5] Tech2Think – Inteligência Artificial Além da Nuvem, R. Gérson da Graça, 94, Campina do Siqueira, Curitiba, 80.740-690, Paraná, Brazil.

[6] Tech2Think – Inteligência Artificial Além da Nuvem, R. Gérson da Graça, 94, Campina do Siqueira, Curitiba, 80.740-690, Paraná, Brazil.

[7] Tech2Think – Inteligência Artificial Além da Nuvem, R. Gérson da Graça, 94, Campina do Siqueira, Curitiba, 80.740-690, Paraná, Brazil.

[8] Tech2Think – Inteligência Artificial Além da Nuvem, R. Gérson da Graça, 94, Campina do Siqueira, Curitiba, 80.740-690, Paraná, Brazil.

[9] Universidade Federal do Paraná, Departamento de Engenharia Elétrica, Jardim das Américas, Curitiba, 82590 300, Paraná, Brazil.

[10] Grupo Equatorial, R. 2 Quadra A37, 505, Jardim Goiás, Goiânia, 74805-180, Goiás, Brazil.

[11] Grupo Equatorial, R. 2 Quadra A37, 505, Jardim Goiás, Goiânia, 74805-180, Goiás, Brazil.

[12] Grupo Equatorial, R. 2 Quadra A37, 505, Jardim Goiás, Goiânia, 74805-180, Goiás, Brazil.

[13] Tech2Think – Inteligência Artificial Além da Nuvem, R. Gérson da Graça, 94, Campina do Siqueira, Curitiba, 80.740-690, Paraná, Brazil.