Análise da Influência de Variáveis Socioeconômicas sobre os resultados de Perdas Não Técnicas estimados em Sistemas de Distribuição de Baixa Tensão

Leonardo Nogueira F. da Silva*, Natalia S. Bastos*, Vinicius J. Garcia, Kamila Stromm*, Antônio Kaminski, Jr.*, Daniel P. Bernardon*, Alzenira R. Abaide*, Otacílio O. Carneiro Filho*****

**Programa de Pós-Graduação em Engenharia Elétrica, Universidade Federal de Santa Maria, RS Brasil (e-mails: leo.nogueirafs@gmail.com, bastos.nataliasousa@gmail.com, kamilastromm@hotmail.com, juniorkaminsk@hotmail.com, dpbernardon@ufsm.br, alzenira@ufsm.br).*

***Programa de Pós-Graduação em Engenharia de Produção, Universidade Federal de Santa Maria, RS Brasil (e-mail: viniciusjg@ufsm.br).*

****Centrais Elétricas de Santa Catarina, CELESC, SC, Brasil (e-mail: otacilioocf@celesc.com.br).*

Abstract: Non-technical losses in distribution systems be represented by the consumed energy amount without billing, mainly due to energy theft and connection frauds. These losses affect the entire utility economic chain, also impacting the number of technical losses. Thus, actions such as field inspections are important to mitigate the problem. Analyzing non-technical losses in low voltage consumer contexts, mostly residential units, relationships between socioeconomic aspects and losses are observed, due to the heterogeneity of the population's living conditions, mainly in developing countries like Brazil. Due to advances in data processing techniques and the current volume of information, the influence of these exogenous aspects can be included in loss expectation models, improving the assertiveness of inspection planning, for example. Thus, this paper seeks to present the relationships between non-technical loss results and socioeconomic information, through statistical analysis and the use of Mutual Information applied to non-linear relationship feature selection. The study will be developed to the demographic units that compose a capital city from the south region of Brazil, however, being a model easily replied to other regions. The results showed greater relationships between non-technical losses and aspects of population income and literacy.

Resumo: As perdas não técnicas, em sistemas elétricos de distribuição, são caracterizadas pela parcela de energia consumida sem faturamento, principalmente devido a furtos e fraudes. Essas perdas influenciam toda a cadeia econômica da concessionária, impactando até mesmo o montante de perda técnica observado. Assim, ações como inspeções em campo são importantes para mitigar o problema. Países em desenvolvimento, como o Brasil, apresentam variações representativas nas condições de vida da população. Esse fato indica que relações entre aspectos socioeconômicos e resultados de perdas não técnicas, principalmente no ambiente de baixa tensão, podem ser observadas. Assim, os avanços em modelos de processamento de dados e a disponibilidade atual de informações cria um ambiente propicio para que esses aspectos exógenos possam ser incluídos em modelos de estimativa de perdas, contribuindo na assertividade do planejamento de inspeções de campo, por exemplo. Dessa forma, o presente artigo busca apresentar as relações entre os resultados de perdas não técnicas e informações socioeconômicos, através de análises estatísticas e do uso de Mutual Information, para seleção de atributos em variáveis de relação não linear. O estudo será direcionado a unidades demográficas que compõem uma capital do sul do Brasil, sendo um modelo passível de aplicação para outras regiões. Os resultados apresentaram maiores relações entre perdas não técnicas e aspectos de renda e alfabetização.

Keywords: Feature Selection; Low Voltage Consumers; Mutual Information; Non-Technical Losses; Socio-economic Aspects; Spatial Analysis.

Palavras-chaves: Análise Espacial; Aspectos Socioeconômicos; Consumidores de Baixa Tensão; Informação Mútua; Perdas Não Técnicas; Seleção de Atributos.

1. INTRODUÇÃO

As perdas de energia correspondem ao montante de energia elétrica gerada que circula em linhas de transmissão e redes da distribuição, mas que não é comercializada, seja por motivos técnicos ou comerciais. O transporte da energia, resulta, inevitavelmente, em perdas técnicas relacionadas à transformação de energia elétrica em energia térmica nos condutores (efeito joule), perdas no núcleo dos transformadores, perdas dielétricas, entre outras inerentes aos elementos do sistema. As perdas não técnicas ou comerciais, por sua vez, decorrem principalmente de furto (ligação

clandestina, com desvios diretamente da rede) ou fraude de energia (adulterações no medidor), erros de medição e de faturamento (ANEEL, 2022a).

As perdas não técnicas (PNT) são apuradas pela diferença entre as perdas totais e as perdas técnicas, e seus limites regulatórios são calculados conforme as regras definidas no Submódulo 2.6 do PRORET (Procedimentos de Regulação Tarifária). A Agência Nacional de Energia Elétrica (ANEEL) estima os limites através de uma metodologia de comparação de desempenho das distribuidoras, observando critérios de eficiência e as características socioeconômicas das áreas de concessão. Em virtude de compor o processo de revisão tarifária, o consumidor regular acaba arcando com parte do prejuízo (ANEEL, 2022b).

A fiscalização das perdas comerciais é realizada pelas concessionárias de energia que, periodicamente, realizam programas de inspeção dos consumidores. No entanto, muitas das distribuidoras no Brasil enfrentam dificuldades na fiscalização, dentre elas o difícil acesso as regiões afastadas nas zonas urbanas, o alto risco de exposição dos colaboradores, o número reduzido de equipes disponíveis, e o custo operacional (Ventura e Melo, 2020).

O direcionamento de recursos para inspeções deve estar vinculado, então, a um resultado confiável de estimativa de perdas. Dentre os principais métodos, destaca-se o reconhecimento de padrões históricos de consumo. Pamir et al. (2022) realizam a estimativa de furto de energia no ambiente urbano, considerando o histórico de consumo, e estimando percentuais de "honestidade" dos consumidores, processados por Redes Neurais Recorrentes Bidirecionais. Barros (2022) também aplica históricos de consumo horário em Redes Neurais Perceptron Multicamadas com agrupamento hierárquico de entradas.

Além disso, avanços na qualidade e disponibilidade de informação possibilitam justificar as PNTs com variáveis exógenas ao sistema. Cruz e Ramos (2014) destacam que fatores socioeconômicos associados ao problema das perdas não técnicas, resultam em melhores políticas para redução de perdas, bem como melhores formas de aplicação dessas decisões aos públicos-alvo. Messinis e Hatziargyriou (2018) indicam a necessidade do desenvolvimento de novos métodos para detecção das PNTs, levando em conta o lugar onde essas perdas ocorrem. Nesse contexto, a análise espacial permite a incorporação do espaço, considerando as características das subáreas onde existem as PNTs. Por meio das técnicas de análise espacial e de modelagem dinâmica, é possível determinar quais subáreas da cidade são mais vulneráveis às PNTs (Faria et al., 2014).

Quanto as variáveis consideradas, para Penin (2008) os aspectos sociais, comportamentais, educativos e socioeconômicos, são muito importantes na análise das PNT, onde, estudos comprovam o aumento de ligações irregulares, quando da deterioração econômica da região ou país em que se encontram. Smith (2004) busca compreender como o problema é tratado em diversos países, indicando que as diferenças de PNT entre os países estariam associadas a problemas de governança, tais como aspectos de processos políticos, liberdade civil, burocracia, independência dos serviços públicos às pressões políticas, instabilidade política e violência, além da efetividade do judiciário.

Em outro estudo, Araújo (2007), define que as perdas podem ser explicadas por variáveis socioeconômicas e de infraestrutura, tais como educação, renda, desigualdade, aspectos da área de concessão, tarifa, dentre outras. Cruz e Ramos (2014) desenvolveram um modelo de análise de perdas não técnicas considerando variáveis de violência, renda, aspectos culturais e a organização da empresa, encontrando correlação positiva entre o índice de perdas totais e a violência, e uma correlação negativa entre o nível de organização da empresa, a renda per capita e o percentual de clientes rurais. Faria (2016) apresenta um estudo que correlaciona os indicadores socioeconômicos e as PNT em análises espaciais, que propiciou a criação de distribuições de probabilidade para projeção de ocorrência de perdas.

Razavi e Fleury (2019) utiliza dados de 70 distritos da Índia, avaliando o cruzamento de PNT com taxa de criminalidade, alfabetização, renda, urbanização e consumo médio. Esses dados, aplicados à modelos de regressão de *Machine-Learning*, como Árvores de Decisão, resultaram em 87% da ocorrência de perdas explicadas pelas variáveis supracitadas. Por fim, Gaur e Gupta (2016) aplicam o modelo de regressão linear dos Mínimos Quadrados Generalizados Viáveis, resultando, para a base de dados, que as menores PNTs ocorrem em ambientes de menor corrupção, maior urbanização (número de unidades consumidoras), menores taxas de pobreza, melhores taxas de alfabetização, e maior renda média per capita.

Neste contexto, o presente trabalho busca realizar uma análise de relação entre perdas não técnicas e informações socioeconômicas regionais. Devido a esses aspectos serem inerentes ao indivíduo, serão observadas as PNT estimadas para circuitos de baixa tensão. Processos estatísticos e de seleção de atributos são aplicados em bases de dados espaciais relativos ao sistema elétrico e informações geográficas, definindo as variáveis que melhor representam as PNTs nas áreas de estudo.

2. METODOLOGIA

Conforme definido nos objetivos do presente trabalho, duas bases de dados desacopladas são comparadas a fim de se estabelecer as relações entre as variáveis de PNT estimadas para circuitos de baixa tensão (PNT-BT) e àquelas socioeconômicas. Para tanto, considerando a disponibilidade de informações no Brasil, utiliza-se como referência de dados a Base de Dados Geográfica da Distribuidora (BDGD) e informações do censo demográfico, realizado pelo IBGE. No BDGD, regulamentado pelo módulo 10 dos Procedimentos de Distribuição da ANEEL, se apresentam informações geográficas e não geográficas de infraestruturas de rede e consumidores (ANEEL, 2021). Por sua vez, o IBGE realiza, em períodos decenais, o censo, cujo último consolidado é de 2010. Na aquisição dos dados, o IBGE define como menor área territorial de informação os Setores Censitários, o que propicia observar variações de perfis populacionais em áreas pequenas, como em uma mesma rua (IBGE, 2011).

A Fig. 1 apresenta um diagrama de blocos das variáveis consideradas e dos processos metodológicos. Quanto aos dados BDGD, são consideradas as informações de PNT-BT, por alimentador, e alocação geográfica de cada ponto notável, compreendido principalmente por postes. Por sua vez, quanto as informações do IBGE, no universo de variáveis foram relacionadas às catorze que melhor caracterizam unidades consumidoras, sendo elas relativas a habitantes por domicílio, renda, alfabetização e aspectos de infraestrutura do entorno.

O primeiro processo na Fig. 1 corresponde ao sincronismo geográfico das variáveis. Nesse trabalho propõem-se uma análise de PNT por Setor Censitário (SC), sendo que os consumidores de um mesmo setor podem ser atendidos por diferentes alimentadores. Assim, inicialmente se estimam quantitativo e qualitativamente o número de pontos notáveis por setor, sendo então aplicada (1) para estimativa de PNT.

$$SC_x(PNT) = \frac{\sum_{al=1}^{AL} PN_{al}(SC_x) \cdot PNT_{al}}{PN(SC_x)} \quad (1)$$

Onde $PN_{al}(SC_x)$ corresponde aos pontos notáveis de um alimentador al no Setor SC_x, PNT_{al} as perdas não técnicas do alimentador al, e $PN(SC_x)$ o número total de pontos notáveis.

A partir da sincronização de informações de perdas não técnicas com os respectivos Setores Censitários, são possíveis os estudos com as séries de dados. Inicialmente se estabelece um bloco de estudos de Mutual Information. À medida que se amplia o universo de informação, e o uso de inteligência artificial para classificação e predição de dados, surge a necessidade de definição de um conjunto de variáveis de entrada que melhor representem a saída buscada. Para tanto, técnicas de seleção de atributos vêm sendo desenvolvidas (Pham e Raahemi, 2023).

Uma especificidade das variáveis elencadas é que a relação com as PNT-BT pode ser não-linear. Sendo assim, métodos lineares como a correlação de Pearson tornam-se insuficientes. Dessa forma, se propõem o uso de Mutual Information (MI), que estima a redundância e relevância entre duas variáveis X e Y aleatórias, independentes, baseado no conceito de entropia da informação (Silva et al., 2019). Em (2) é apresentada essa relação.

$$I[X;Y] = \sum_{i=1}^{n}\sum_{j=1}^{m} p_{ij} * \log\frac{p_{ij}}{p_i * p_j} \quad (2)$$

Onde p_{ij} é a probabilidade mútua de amostras i (variável X) e j (variável Y), sendo p_i e p_j as probabilidades individuais. A saída de MI varia, então, entre 0 e a entropia individual mínima de X e Y, no qual, quanto maior o valor de MI, mais relevante é a relação entre as duas variáveis.

Por fim, têm-se a realização de análises estatísticas, baseadas na observação dos diagramas de dispersão das variáveis. Conforme Fig. 1, derivam se, também, as análises de concentração de perdas, no qual os extremos das séries serão avaliados, a fim de observar o impacto das perdas nos piores e melhores SCs frente as variáveis socioeconômicas. Essa análise contribui para avaliar se a tendência da relação é direta ou inversamente proporcional.

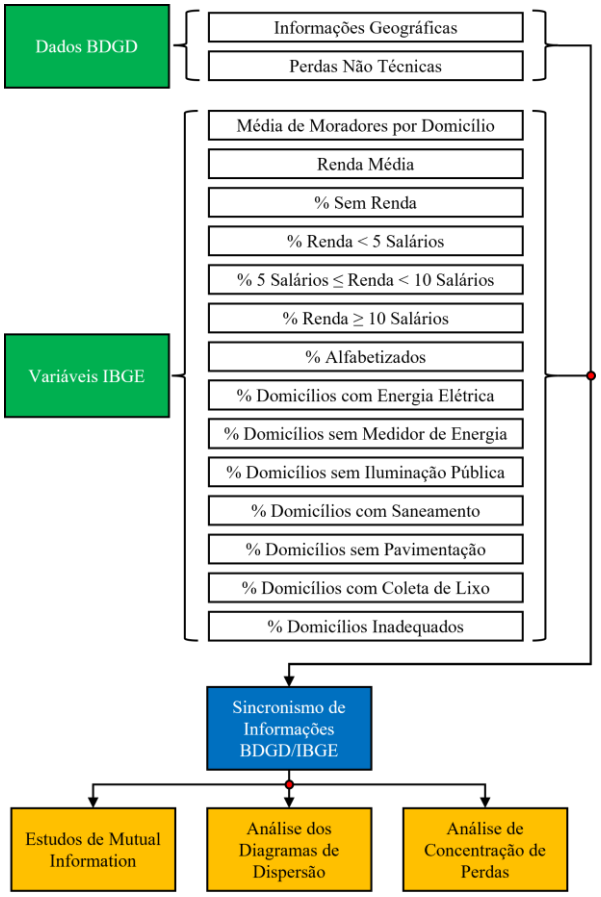

Fig. 1 Diagrama de blocos das relações entre as bases de dados e estudos propostos.

3. ESTUDO DE CASO E RESULTADOS

3.1 Descrição do Estudo de Caso

Para realização dos estudos de relação entre as variáveis socioeconômicas e PNT-BT, propõem-se um estudo de caso composto da área total de uma capital localizada em um estado do sul do Brasil. Ressalta-se que a aplicação dessas etapas metodológicas não está vinculada apenas a esse estudo de caso, sendo possível ampliá-lo para toda área de concessão, ou mesmo todas as concessionárias brasileiras. A Tabela 1 apresenta uma síntese das características do estudo de caso.

A área de análise corresponde a 65 bairros, totalizando 636 SCs, o que indica uma maior granularidade da informação socioeconômica. Foram usados 576 para os estudos, sendo excluídos aqueles sem composição completa de dados. Além disso, os dados do BDGD consideram os resultados para 2021, totalizando um consumo de 1540 TWh para a área de estudo. Esta é composta de 71 alimentadores e 9 subestações, com capacidade de 612 MW. Inicialmente, com auxílio do software QGIS 3.28.6, foram sobrepostos os pontos notáveis e os SCs, a fim da aplicação de (1). A Fig. 2 apresenta o mapa, no qual as áreas em azul correspondem aos setores, cujos limites estão demarcados em vermelho. Os elementos em preto destacam os pontos notáveis da rede.

Tabela 1. Características Globais dos dados do Estudo de Caso

Grupo	Variável	Valor
Informações Geográficas	Área	675,4 km²
	População	508.826
	Nº Bairros	85
Dados BDGD	Conjuntos	6
	Subestações	9
	Alimentadores	71
	Pot. Inst. Subestações	612 MW
	Consumidores de Baixa Tensão	266.726
	Consumo Total 2021	1539,573 TWh
	PNT-BT 2021 (Estimado)	64,95 TWh
	Pontos Notáveis de Baixa Tensão	50.130
Dados IBGE	Setores Censitários TOTAL	636
	Setores Censitários Aval.	576

3.2 Resultados de Mutual Information

Como primeira análise, foi realizado o processamento de MI entre cada variável socioeconômica e as PNT-BT. A Fig. 3 apresenta os resultados considerando as séries de dados completas (576 pontos). Verificam-se 2 grupos distintos. Os dados de melhor relação para a área de estudo foram àqueles relacionados com a característica da população, especificamente a renda média, faixas de salário, alfabetização e média de moradores por residência. As variáveis relacionadas à infraestrutura, como a condição do domicílio, aspectos de pavimentação, e fornecimento de energia mostraram resultados inferiores. Isso se dá em virtude da maior homogeneidade desse segundo conjunto de variáveis, devido ao desenvolvimento da região. Por exemplo, em geral todos os domicílios possuem energia elétrica, não justificando a alta variação de PNT. Contudo, no aspecto renda, comumente é observada uma heterogeneidade em ambientes urbanos, traduzida, nesse caso, para uma relação com as perdas.

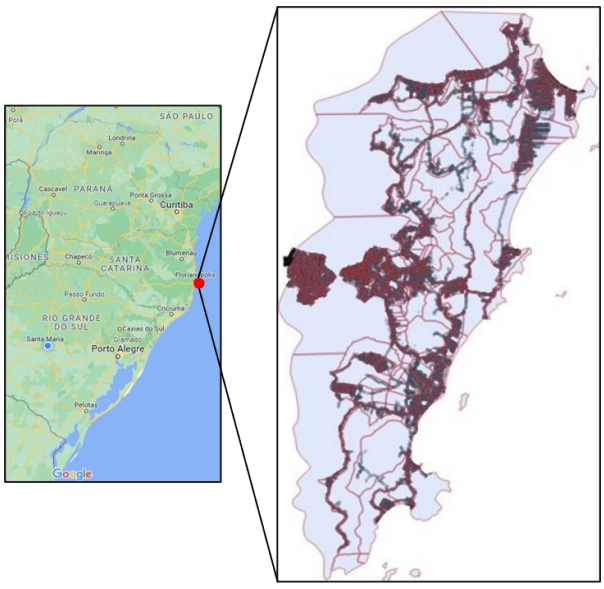

Fig. 2 Representação gráfica da área aplicada no estudo de caso, considerando pontos notáveis e setores censitários.

Visando filtrar essa homogeneidade, foi estabelecida a moda de cada variável socioeconômica. Usando como exemplo a variável percentual de domicílios com energia elétrica, dentre os 576 SCs, apenas 66 apresentam ao menos uma unidade consumidora (UC) sem energia. Assim, uma segunda rodada de MI foi gerada, resultando na Fig. 4. Os resultados, nesse caso, foram iguais em ordem de seleção de atributos, que aqueles da Fig. 3. Contudo, verifica-se uma melhora dos resultados numéricos de conjuntos de variáveis relacionadas a infraestrutura, com os resultados de percentual de domicílios sem pavimentação, sem medidor de energia e sem iluminação pública se aproximando ao primeiro conjunto.

3.3 Análises estatísticas das séries de dados

Os resultados apresentados na subseção anterior indicam a seleção de atributos que representa as PNT na região, sendo aqueles inerentes a população, os de melhor desempenho,

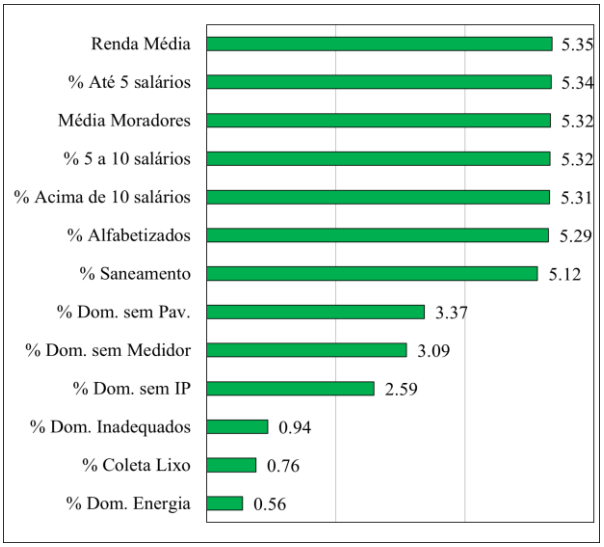

Fig. 3 Resultados de Mutual Information para as variáveis socioeconômicas analisadas, frente as perdas não-técnicas.

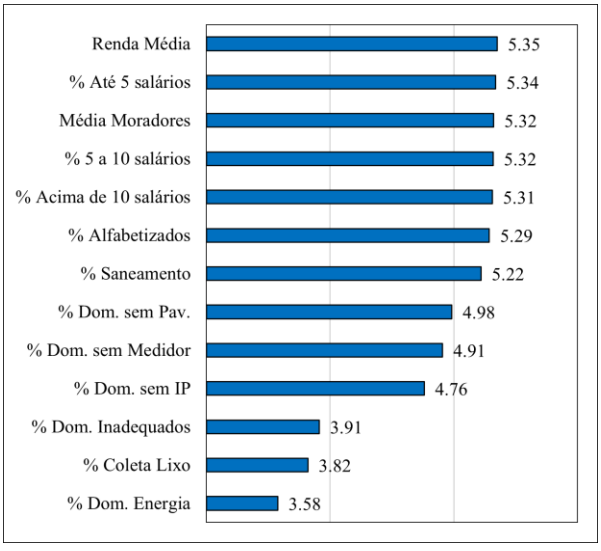

Fig. 4 Resultados de Mutual Information aplicado as variáveis socioeconômicas filtradas, frente as informações de perdas não-técnicas.

principalmente a renda. Contudo, é fundamental analisar o comportamento dessa informação socioeconômica, dado por relações diretas ou inversamente proporcionais.

Para tanto, foram realizados os diagramas de dispersão para todas as 14 variáveis socioeconômicas, em relação a PNT-BT, sendo apresentados três exemplos, em Fig. 5, Fig. 6 e Fig. 7, para média de moradores, renda média, e percentual de alfabetização, respectivamente. Essas variáveis foram escolhidas devido ao melhor desempenho de MI.

Para o caso de Fig. 5, para média de moradores por residência, se observa, através da curva de tendência linear, que há uma relação fraca, mas diretamente proporcional, entre as variáveis, com coeficiente de Pearson estimado em 0,093. Assim, quanto maior a média de pessoas por residência, maior podem ser caracterizadas as PNT-BT na região de estudo. A relação PNT e a média de moradores por residência pode não ser direta, mas sim com os atributos que influenciam essa variável socioeconômica, como a renda dos moradores, a localização desses domicílios e a classe social.

A Fig. 6 apresenta, então, a relação entre PNT-BT e a renda média, variável de maior índice MI na etapa de seleção de atributos. Conforme se observa na curva de tendência, também linear, há uma relação inversamente proporcional nos dados, mais forte que a anterior, com coeficiente de Pearson em 0,167. Dessa forma, quanto menor a renda média da população, maior tendem a ser as perdas não técnicas, por Setor Censitário, para a área de estudo. Esse resultado é similar ao encontrado pelo trabalho de Faria (2016), podendo estar relacionado a localização do domicílio frente a aspectos sociais, e a vulnerabilidade da população de baixa renda.

Por fim, a Fig. 7 apresenta a análise de dispersão da variável percentual de alfabetizados, relacionada a PNT-BT. Novamente, uma análise mais aprofundada dos fatores que afetam a variável de alfabetização é necessária, devido ao fato que a vulnerabilidade social pode indicar um menor acesso à educação, o que gera uma redução da renda média, e a susceptibilidade à ocorrência de PNT. Conforme observado na curva de tendência, apesar da alta dispersão, há uma propensão, nos extremos, de uma maior ocorrência de PNT em regiões de menor percentual de alfabetização.

Uma vez que nos estudos realizados a granularidade de informações socioeconômicas é 8,11 vezes maior que a de perdas (576 SC para 71 alimentadores), é esperado que alguns SCs apresentem o mesmo resultado de PNT, mesmo com variações nos indicadores sociais. Buscando refinar essa análise, propõem-se, por fim, analisar a média de perdas nos extremos das variáveis socioeconômicas. Assim, foi realizado o ranqueamento dos 576 SCs, por indicador, e filtrados os 1%, 5% e 10% maiores e menores valores, do qual se obteve a média aritmética de perdas. A Tabela 2 apresenta esses resultados.

Na Tabela 2 são destacados os resultados nos quais a diferença é maior que 1% de PNT entre os extremos, para 1% dos maiores e menores resultados, uma diferença maior que 0,5%, para os 5% maiores e menores, e uma diferença maior que 0,3% para os extremos em 10%. À medida que se

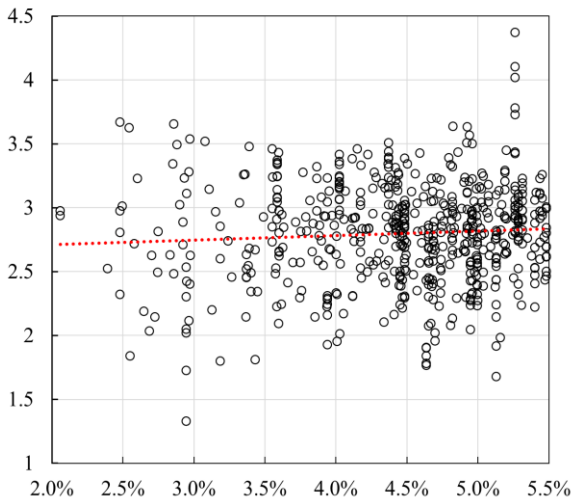

Fig. 5 Diagrama de dispersão da variável Média de Moradores frente a PNT-BT percentual.

Fig. 6 Diagrama de dispersão da variável Renda Média frente a PNT-BT percentual.

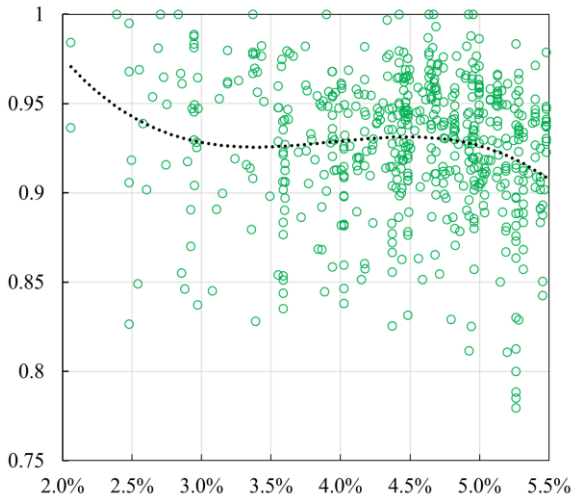

Fig. 7 Diagrama de dispersão da variável Percentual de Alfabetizados frente a PNT-BT percentual.

aumenta essa variação, maior tende a ser a similaridade dos resultados das médias de perdas para os extremos.

Tabela 2. Variação dos resultados de PNT-BT nos extremos 1%, 5% e 10%, das séries de dados

Variável	1%		5%		10%	
	Men	**Mai**	**Men**	**Mai**	**Men**	**Mai**
Média Moradores	4,059%	**5,261%**	4,057%	4,297%	4,184%	4,257%
Renda Média	**5,271%**	3,606%	4,182%	4,100%	**4,405%**	4,066%
% Até 5 salários	3,725%	**4,940%**	3,965%	**4,566%**	4,087%	**4,674%**
% 5 a 10 salários	4,193%	4,839%	4,295%	4,257%	4,386%	4,246%
% Acima de 10 salários	**4,771%**	3,677%	4,470%	4,041%	**4,582%**	4,131%
% Alfabetizados	**5,249%**	3,489%	**4,345%**	3,842%	4,356%	4,075%
% Dom. Energia	4,675%	5,000%	4,417%	4,503%	4,438%	4,244%
% Dom. sem Medidor	4,310%	4,551%	4,410%	4,492%	4,564%	4,455%
% Dom. sem IP	4,269%	5,267%	4,438%	4,604%	4,417%	4,634%
% Saneamento	**5,052%**	4,041%	4,788%	4,387%	**4,843%**	4,428%
% Coleta Lixo	4,468%	4,914%	4,409%	4,470%	4,546%	4,444%
% Dom. sem Pav.	4,658%	4,711%	**4,819%**	4,281%	**4,682%**	4,318%
% Dom. Inadequados	4,900%	5,115%	4,597%	4,471%	**4,905%**	4,516%

Quanto a variável renda média verifica-se que as perdas são maiores para aqueles resultados inferiores de renda, sendo de 5,271% vs. 3,606%, em extremos de dimensão 1%. Essa alta variação também ocorre para a variável percentual de alfabetizados e percentual da população que recebe até 5 salários-mínimos.

Em extremos de dimensão 5% a ocorrência de variações maiores que o limite, nesse caso, 0,5%, diminui, mantendo o destaque das variáveis percentual de alfabetizados e percentual da população com renda até 5 salários-mínimos. Contudo, quando se avaliam os extremos em 10%, voltam as variáveis relacionadas a renda indicando que as perdas são maiores para rendas médias menores, para locais cujo percentual de renda até 5 salários-mínimos é maior, e cujo percentual de renda acima de 10 salários-mínimos é menor.

Assim, destaca-se a variável renda e suas derivações, que nesse estudo de caso, dadas as condições de contorno do problema, indicou uma maior susceptibilidade de PNT em alimentadores localizados em regiões de renda média inferior. Por fim, quando se analisa aspectos de infraestrutura, para extremos em 10%, verifica-se um destaque para uma tendência de maiores perdas em locais com um percentual maior de domicílios inadequados, sem pavimentação e sem saneamento.

Quanto as variáveis censitárias relacionadas ao sistema elétrico, percentual de domicílios sem energia, sem medidor e sem iluminação pública, assim como nos resultados MI, apresentou alto volume de dispersão, tendo em vista fatores já explanados, como a ocorrência de redes de baixa tensão atendendo toda a região de análise.

4. CONCLUSÕES

De forma geral, as perdas na cadeia operacional de sistemas elétricos de potência resultam em redução da eficiência energética, elétrica e econômica. Contudo, diferentemente das perdas técnicas, que são inerentes as características físicas dos elementos que compõem o sistema, as perdas de origem não técnicas, principalmente resultante de fraudes e furtos de energia podem, e devem, ser mitigadas.

Devido a ocorrência da perda não técnica estar relacionada aos consumidores, o grande volume de conexões dificulta a resolução do problema, principalmente pelo número limitado de equipes de inspeção em campo, e recursos financeiros despendidos para tanto. Dessa forma, o aumento da assertividade de pontos de vulnerabilidade do sistema pode proporcionar uma melhor otimização de tempo e recurso. Além disso, ações de conscientização dos grupos-alvo são mais bem direcionadas.

Também, conforme objetivo de estudo do presente artigo, há uma relação entre aspectos exógenos da atividade de distribuição de energia frente à ocorrência de PNTs. As variáveis socioeconômicas podem, então, auxiliar no entendimento dessas ocorrências, compondo os modelos de ranqueamento de pontos de inspeção.

No presente artigo, buscou-se relacionar 14 variáveis socioeconômicas obtidas no censo IBGE 2010, com PNTs estimadas para alimentadores de uma região de estudo. Dentre as variáveis, separam-se dois conjuntos, um relativo as características da população e outro à infraestrutura do ambiente. Observou-se, mesmo com as limitações de informação, uma maior relação entre PNT e variáveis de renda, média de moradores por residência e alfabetização, sendo que índices de PNT foram diretamente proporcionais a média de moradores e inversamente proporcionais a renda e alfabetização. Isso indica um provável ponto de vulnerabilidade social, pelo fato da importância da energia elétrica na sociedade atual. Assim, parcelas populacionais podem estar buscando em alternativas fora do contexto regular o atendimento as suas necessidades energéticas. Ações que já ocorrem, como aplicação de tarifas baixa renda, podem favorecer a redução de PNT. Contudo, a alta dispersão desses dados não exclui a ocorrência de perdas comerciais também em faixas sociais elevadas.

Por fim, para trabalhos futuros, propõem-se a realização do estudo ampliando a área de análise para toda concessão da região em questão, obtendo variações de realidade entre regiões metropolitanas, interior, e extremos sociais. Além disso, se observam espaços para avanços na obtenção de perdas por Setor Censitário, atuando sobre o montante de consumidores geolocalizados, bem como a alternância da ótica de estudo, para análises de alimentadores frente aos indicadores socioeconômicos. Ao final pretende-se obter uma vasta quantidade de informações que melhor caracterizem as perdas não técnicas e auxiliem na tomada de decisão do planejamento de inspeções em campo.

AGRADECIMENTOS

Os autores agradecem ao apoio técnico e financeiro da concessionária Centrais Elétricas de Santa Catarina S.A – CELESC (Programa de P&D ANEEL através do projeto CELESC/UFSM nº 0422/2022), Instituto Nacional de

Ciência e Tecnologia em Sistemas de Geração Distribuída (INCTGD), Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq – nº 465640/2014-1, processo nº150276/2023-0), Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES – nº 23038.000776/2017-54), Fundação de Amparo à Pesquisa do Estado do Rio Grande do Sul (FAPERGS – nº 17/2551-0000517-1) e Universidade Federal de Santa Maria (UFSM), à Fundação de Apoio da Universidade Federal de Minas Gerais (FUNDEP), instituições brasileiras.

REFERÊNCIAS

ANEEL (2021). PRODIST Módulo 10 – Sistema de Informação Geográfica Regulatório. Agência Nacional de Energia Elétrica, Brasília.

ANEEL (2022a). Perdas de Energia. Disponível em: https://www.gov.br/aneel/pt-br/assuntos/distribuicao/perdas-de-energia. Acesso em: 28 abr. 2023.

ANEEL (2022b). Procedimentos de Revisão Tarifária – Submódulo 2.6 Perdas de Energia. REN n 1003/2022. Agência Nacional de Energia Elétrica, Brasília.

Araújo, A. C. M. (2007). *Perdas e Inadimplência na Atividade de Distribuição de Energia Elétrica no Brasil*. Tese de Doutorado. UFRJ -COPPE, Rio de Janeiro.

Barros, V. S. de (2022). *Aplicação de técnica de Machine Learning na detecção de perdas não-técnicas*. Dissertação de Mestrado. Universidade Federal de Pernambuco, Recife.

Cruz, K. E. A., e Ramos, F. de S. (2014). Impacto de variáveis socioeconômicas sobre as perdas comerciais de energia elétrica. *Revista Brasileira de Economia de Empresas*, v. 14, n. 2, p. 31-52.

Faria, L. T., Melo, J. D., e Padilha-Feltrin, A. (2014). Análise espacial de pontos para mapeamento de perdas comerciais. Em *Anais do Simpósio Brasileiro de Sistemas Elétricos 2014*, p. 1-8, Foz do Iguaçu.

Faria, L. T. (2016). *Estimação Espaço-Temporal das Perdas Não Técnicas no Sistema de Distribuição de Energia Elétrica*. Tese de Doutorado. Faculdade de Engenharia, Universidade Estadual Paulista, Ilha Solteira.

Gaur, V., e Gupta, E. (2016). The determinants of electricity theft: An empirical analysis of Indian states. *Energy Policy*, vol. 93, p. 127-136.

IBGE (2011). Base de Informações do Censo Demográfico 2010: Resultados do Universo por setor censitário. Instituto Brasileiro de Geografia e Estatística, Rio de Janeiro.

Messinis, G.M., e Hatziargyriou, N.D. (2018). Review of non-technical loss detection methods. *Electrical Power Systems Research*, v. 158, p. 250-266.

Pamir, N., Almogren, A., e Umar Javed, M. (2022). RFE Based Feature Selection and KNNOR Based Data Balancing for Electricity Theft Detection. *IEEE Access*, v.10, p. 112948-112963.

Penin, C.A.S. (2008). *Combate, prevenção e otimização das perdas comerciais de energia elétrica*. Tese de Doutorado. Escola Politécnica, Universidade de São Paulo, São Paulo.

Pham, T. H., e Raahemi, B. (2023). Bio-Inspired Feature Selection Algorithms with Their Applications: A Systematic Literature Review. *IEEE Access*, vol. 11, 43733-43758.

Razavi, R., e Fleury, M. (2019). Socio-economic predictors of electricity theft in developing countries: An Indian case study. *Energy for Sustainable Development*, v. 49, p. 1-10.

Silva, L. at al. (2019). Impact Evaluation of Feature Selection to Short-Term Load Forecasting Models considering Weather Inputs and Load History. Em *Anais do 2019 54th International Universities Power Engineering Conference (UPEC)*, p. 1-6, Bucharest, Romênia.

Smith, B. (2004). Electricity theft: a comparative analysis. *Energy Policy*, v. 32, p. 15-21.

Ventura, L. O., e Melo, J. D. (2020). Estimação da Vulnerabilidade às Perdas Comerciais por Subárea Utilizando uma Regressão Espacial Hierárquica. Em *Anais do Simpósio Brasileiro de Sistemas Elétricos 2020*, p. 1-6, Santo André.