

Análise de sensibilidade estatística dos atributos físico-elétricos no cálculo dos indicadores de continuidade de sistemas de distribuição

Robson P. Delavechia*. Raul S. Weiland**. Leonardo Silveira*. Bibiana P. Ferraz**.
Maicon J. S. Ramos**, Daniel P. Bernardon*. Laura L. C. Santos*. Rui A. F. Garcia***

*Universidade Federal de Santa Maria, Santa Maria, Brasil

(e-mail: robson.porsch@acad.ufsm.br, leonardo.silveira@acad.ufsm.br, dpbernardon@ufsm.br, laura.santos@ufsm.br)

**Universidade Federal do Rio Grande do Sul, Porto Alegre, Brasil

(e-mail: raul.weiland@ufrgs.br, bibiana.petry@ufrgs.br, maicon.jaderson@ufrgs.br)

***Companhia Estadual de Distribuição de Energia Elétrica, RS, Brasil

(e-mail: ruiifg@ceee.equatorialenergia.com.br)

Abstract: In this paper, results related to the statistical sensitivity analysis of physical-electrical attributes in reliability index calculation of distribution systems will be presented. Considering that electricity companies are inserted in a natural monopoly, the National Agency of Electrical Energy (ANEEL) current methodology proposes a comparative assessment strategy of different sets of consumer units, to classify their performance in relation to the power quality service. From a Stepwise regression model, the statistical sensitivity of different arrays of physical-electrical attributes is assessed. The results obtained for a real case studies demonstrate the impact of regional characteristics on the performance of distribution systems.

Resumo: Neste trabalho serão apresentados resultados relativos à análise de sensibilidade estatística dos atributos físico-elétricos no cálculo dos indicadores de continuidade de sistemas de distribuição. Considerando que as distribuidoras de energia elétrica estão inseridas em um monopólio natural, a metodologia vigente da Agência Nacional de Energia Elétrica (ANEEL) propõe uma estratégia de avaliação comparativa de diferentes conjuntos de unidades consumidoras, de forma a classificar o seu desempenho em relação à qualidade do serviço. A partir de um modelo de regressão por *Stepwise*, é avaliada a sensibilidade estatística de diferentes arranjos de atributos físico-elétricos. Os resultados obtidos para os estudos de casos reais demonstram o impacto de características regionais no desempenho de sistemas de distribuição.

Keywords: power quality; distribution systems; Stepwise regression; benchmarking.

Palavras-chaves: qualidade de energia elétrica; sistemas de distribuição; regressão por Stepwise; avaliação comparativa.

1. INTRODUÇÃO

A qualidade da energia elétrica é um dos fatores mais sensíveis dentro dos sistemas de distribuição, pois está vinculada a constantes debates regulatórios pela busca por melhorias no setor (Pessanha, 2006). O setor de distribuição de energia elétrica tem características de monopólios naturais, e neste ambiente, a regulação econômica é a chave para garantir eficiência de custos, qualidade de abastecimento e uma eficiente precificação da rede de distribuição (Xavier; *et al.*, 2015). Desta forma, o custo operacional eficiente geralmente é baseado em um modelo de *benchmarking*, o qual avalia o desempenho das distribuidoras e classifica-o usando um conjunto de variáveis (Medeiros; *et al.*, 2022). Esse modelo é aplicado em diversos países. Em pesquisa realizada em 40 países em um total de 43 reguladores entre junho e outubro de

2008, descobriu que 51% das empresas estavam aplicando técnicas de *benchmarking* (Costa; Lopes; Matos, 2015). Recentemente, Soroush *et al.* (2021) propuseram uma análise de fronteira estocástica para estimar funções de custo e examinar o desempenho de 107 concessionárias italianas de distribuição de energia elétrica de 2011 a 2015.

No contexto nacional, a Agência Nacional de Energia Elétrica (ANEEL) estabeleceu nos Procedimentos de Distribuição (PRODIST) no Sistema Interligado Nacional (SIN), os indicadores de desempenho como a Duração Equivalente de Interrupção por Unidade Consumidora (DEC) e Frequência Equivalente de Interrupção por Unidade Consumidora (FEC). Tais indicadores mensuram a eficácia de cada conjunto de Unidades Consumidoras (UC), os quais estão agrupados perante a homogeneidade de seus atributos físico-elétricos. Na última década, inúmeras pesquisas buscaram meios de

aprimorar a metodologia de definição dos limites para os indicadores de continuidade DEC e FEC, baseada na clusterização de distribuidoras proposto por Tanure, Thahan e Lima (2006). Costa; Lopes e Matos (2015) propuseram uma análise estatística contemplando teste de hipóteses, teoria de regressão linear e simulações de Monte Carlo para avaliar o modelo de *benchmarking* proposto pela ANEEL. Atualmente, a ANEEL dispõe de uma base de dados com 2.610 conjuntos de 43 distribuidoras, e 146 atributos físico-elétricos, os quais permitem descrever os conjuntos definidos por cada distribuidora brasileira. Um dos principais desafios consiste na seleção de atributos físico-elétricos que descrevem de maneira apropriada os conjuntos, que tenham dados consistentes e atualizados e atendam às premissas estatísticas, resultando em um modelo simplificado e completo.

Nesse contexto, o presente estudo objetiva realizar uma análise de sensibilidade de parâmetros estatísticos para definição de atributos físico-elétricos que caracterizam conjuntos de UC. A partir da metodologia vigente, publicada pela ANEEL na Nota Técnica nº 102/2014, propõe-se avaliar quais atributos melhor representam a variabilidade dos indicadores DEC e FEC, bem como caracterizar os conjuntos de UC, a partir dos registros do último ciclo tarifário.

2. FUNDAMENTAÇÃO TEÓRICA

Nesta seção, serão abordadas as principais definições e conceitos relacionados ao cálculo de indicadores de continuidade, estabelecidos na metodologia vigente da ANEEL, os quais embasarão a metodologia proposta.

2.1 Definições e conceitos

Atributos físico-elétricos representam características de um conjunto, o qual é formado por um determinado número de UC, que permitem analisar semelhanças entre outros conjuntos definidos pela ANEEL. Conforme definido pelo Módulo 8 do PRODIST, o *conjunto* de UC é definido por Subestação de Distribuição (SED). Por intermédio dos atributos, os conjuntos são agrupados pelo método de clusterização, o qual busca reunir os conjuntos mais semelhantes, dando origem a um *cluster*. Além dos atributos físico-elétricos, existem os socioeconômicos que relacionam fatores como: a violência, a renda, domicílios subnormais, entre outros.

A qualidade do fornecimento de energia elétrica em cada conjunto é mensurada por limites pré-estabelecidos, os quais garantem padrões de continuidade mínimos (ANEEL, 2014). Para avaliar a qualidade do serviço por meio dos limites, a ANEEL emprega indicadores coletivos, definidos dentro do Módulo 8 do PRODIST, denominados de DEC e FEC.

A determinação dos limites de cada cluster é feita por intermédio da técnica conhecida como *Yardstick Competition*. Esta técnica baseia-se em determinar a referência do cluster a partir de um limite estatístico. Para os conjuntos conectados ao

SIN, considera-se como medida estatística o segundo decil do cluster.

Deste modo, é fundamental a seleção dos atributos relevantes para o estabelecimento dos limites de continuidade para os conjuntos. Tendo em vista que ocorrem problemas de padronização da forma de envio dos atributos no Sistema de Informação Geográfica Regulatório (SIG-R), assim como alguns atributos apresentam registros vazios ou nulos para determinados conjuntos, ou ainda, bases de dados com atualização pouco frequente (como, por exemplo, atributos retirados do IBGE e do Atlas do Desenvolvimento Humano no Brasil, que possuem atualização decenal), fica evidente a necessidade de redução do número de atributos físico-elétricos utilizados.

2.2 Processo de Seleção dos atributos

O processo de seleção dos atributos publicado pela ANEEL na NT102/14¹ consiste em 7 etapas, ilustradas na Fig. 1. A seguir cada etapa é descrita, demonstrando a seleção de 6 dos 146 atributos, considerando DEC e FEC como variáveis dependentes.

Etapa 1: aquisição da base de dados dos 146 atributos de cada um dos 2.610 conjuntos, conforme definido nas Consultas Públicas n. 043/2009 e n. 008/2010. Com a base de dados consolidada é possível verificar informações associadas as distribuidoras, conjuntos, atributos e valores médios dos indicadores coletivos. Com o objetivo de avaliar o impacto dos atributos socioeconômicos no modelo, foram efetuadas análises com e sem os 11 atributos socioeconômicos, listados na NT102/14. Além disso, perante a dificuldade de obtenção de informações, problema de padronização da forma de envio, intersecção de bases de dados e correlações elevadas, mais de 70 atributos são excluídos na Etapa 1.

Etapa 2: análise de correlação de *Pearson* e *Spearman* entre os atributos e os valores médios de DEC e FEC (considerando os anos de 2011 a 2013), onde a ANEEL levou em consideração uma correlação mínima de 0,20, em módulo, para o atributo permanecer na análise seguinte. Logo, são selecionados 69 atributos para DEC e 67 para o FEC.

Etapa 3: avaliação do impacto de atributos socioeconômicos. Esta etapa intersecciona o problema, em dois fluxos distintos, sendo o fluxo em azul o que representa o modelo com todos os atributos e, o fluxo em verde, o modelo sem os 11 atributos socioeconômicos. Antes da aplicação da análise de regressão é realizada a normalização dos dados, visando uma distribuição mais uniforme. Como a distribuição original dos dados não é normal, aplicou-se o logaritmo natural para obter linearidade. A normalização dos dados foi efetuada por:

$$X_T = \ln(1 + X), \quad (1)$$

onde X_T representa a variável transformada e X representa a variável original. Percebe-se que em (1), a variável é somada

¹ NT102/14: Nota Técnica n. 102/2014, publicada pela ANEEL.

com 1 dentro do argumento, de modo a evitar problemas numéricos nos casos em que a variável é zero. Tendo em vista que alguns conjuntos não apresentam dados para todos os atributos, é necessário retirar os conjuntos que possuem dados vazios.

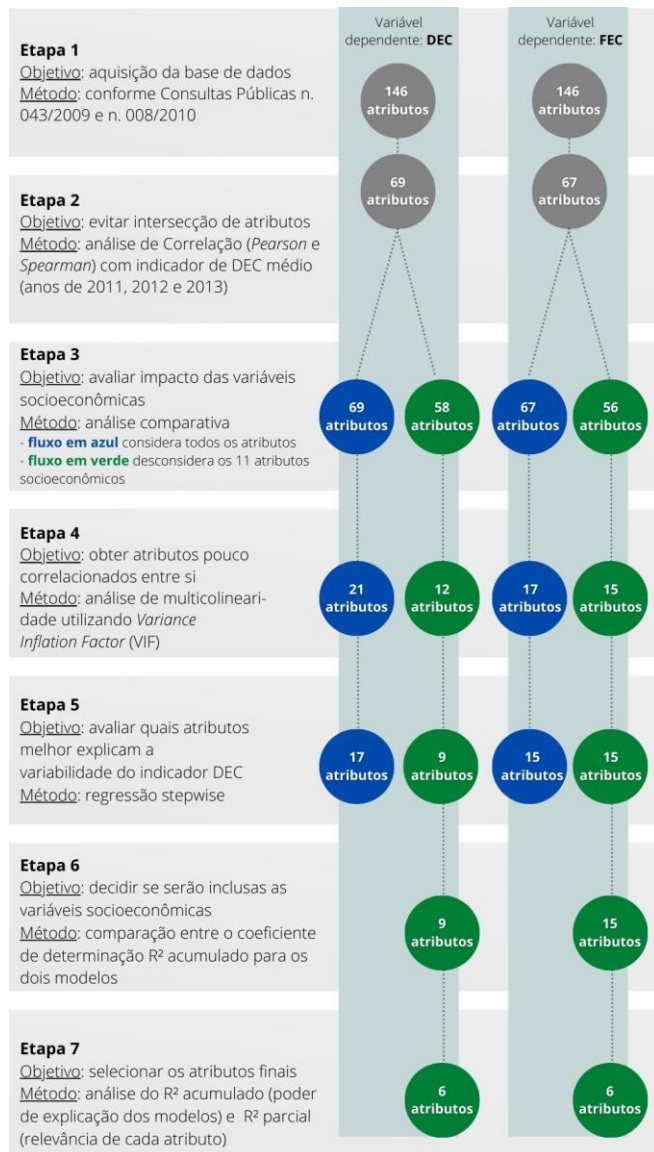


Fig. 1 Principais etapas do processo de seleção dos atributos.

Etapa 4: regressão linear pelo método *Stepwise* para os conjuntos que não possuem dados vazios, onde os atributos são inseridos passo a passo no modelo, de acordo com sua contribuição adicional. Eventualmente, com a inclusão de uma nova variável, uma variável anteriormente inserida pode deixar de ser significativa, sendo retirada do modelo. O critério de significância adotado pela ANEEL é de 1% para a variável entrar ou sair do modelo. Em seguida, visando obter atributos pouco correlacionados entre si, realiza-se uma análise de Multicolinearidade utilizando como métrica a medida estatística denominada de fator de inflação de variância VIF (do inglês, *Variance Inflation Factor*). O atributo de maior VIF é descartado da regressão, a qual é realizada novamente com os atributos que não foram excluídos. Esse processo de

regressão e análise do VIF é realizado até que todos os atributos apresentem VIF inferior a 5. Todo o processo descrito nessa etapa é efetuado para o DEC e FEC, com e sem atributos socioeconômicas.

Etapa 5: repetição do processo da etapa 4, porém com os atributos selecionados na etapa anterior. Além disso, é avaliado o comportamento dos atributos socioeconômicos.

Etapa 6: decisão sobre a inclusão ou não de variáveis socioeconômicas. A partir dos modelos obtidos para DEC e FEC, considerando-se ou não os atributos socioeconômicos, observa-se que os modelos mais completos apresentam um maior valor de coeficiente de determinação acumulado (R² do modelo). Apesar disso, cabe salientar que grande parte dos atributos socioeconômicos disponíveis por município, possuem baixa taxa de atualização dos dados, podendo resultar em distorções nas análises. Logo, observando-se a dificuldade da atualização de dados socioeconômicos, opta-se pelo modelo sem as variáveis socioeconômicas para DEC e FEC.

Etapa 7: avaliação do coeficiente de determinação parcial (R² parcial), a partir do modelo sem atributos socioeconômicos. A ANEEL adotou o valor de 2% para o R² parcial para o DEC e FEC. Esse valor foi escolhido após se verificar que do sexto para o sétimo atributo, o valor do R² parcial cai para menos da metade. Desse modo, optou-se por utilizar apenas os seis atributos mais importantes para os dois modelos. Ressalta-se que, dos seis atributos escolhidos, cinco são comuns aos dois indicadores. A diferença encontrada para os atributos, quando são comparados os modelos para os indicadores coletivos, é que o DEC leva em consideração o número de UC industriais e o FEC o número de UC comerciais.

2.3 Determinação de conjuntos semelhantes e limites de continuidade

A partir dos atributos físico-elétricos selecionados, os conjuntos são agrupados por meio de um método estatístico, o qual leva em consideração a similaridade e a heterogeneidade entre eles. Para avaliar a semelhança entre os conjuntos em relação aos atributos selecionados, adota-se a distância Euclidiana, uma vez que essa métrica é sensível a diferentes escalas ou magnitudes entre as variáveis. Logo, a ANEEL emprega o método dinâmico, o qual visa comparar cada conjunto com os conjuntos mais semelhantes a ele, sendo aplicado separadamente para os indicadores DEC e FEC. Dessa forma, para cada conjunto de referência, será formado um agrupamento de conjuntos que são os mais comparáveis a ele. Além disso, o método conta com a avaliação dos conjuntos heterogêneos e trajetória de redução intensa dos limites. Por fim, o método estabelece os limites de continuidade dos conjuntos de UC.

Por meio das subseções descritas e das etapas dentro do processo de seleção dos atributos, até a definição dos limites de continuidade, verifica-se a importância de se estabelecer uma base de dados sólida para manipulação das informações. Ainda, cabe ressaltar sobre a variação de parâmetros de sensibilidade na etapa de seleção dos atributos, a qual pode ser realizada para se observar o comportamento do modelo. Assim, pode-se verificar se haverá melhorias ou não perante

uma análise estatística dos limites gerados, após os resultados obtidos no processo de agrupamento.

3. METODOLOGIA

O presente estudo visa avaliar o impacto na variação do limite objetivo para o DEC e FEC, por meio da sensibilidade dos parâmetros estatísticos empregados no modelo vigente da ANEEL. Mediante essa avaliação, busca-se comparar os resultados obtidos para todos os conjuntos existentes, bem como analisar os conjuntos de interesse específico. A metodologia proposta é composta de 4 etapas: (i) aquisição de dados, (ii) seleção dos atributos, (iii) clusterização e (iv) análise estatística; as quais são descritas a seguir.

3.1 Aquisição de Dados

A primeira etapa consiste na aquisição do histórico dos atributos físico-elétricos, bem como os dados de DEC e FEC de cada conjunto das distribuidoras de interesse. Conforme descrito nas sete etapas da Fig. 1, antes da aplicação da análise de regressão, deve-se transformar os dados obtidos, visando uma distribuição mais próxima da normal. Cabe ressaltar que durante o processo de normalização dos dados, deve-se observar os conjuntos onde não há informações de seus atributos.

3.2 Seleção dos Atributos

Nas Etapas 1 a 3, ilustradas na Fig. 1, o processo de seleção dos atributos é realizado com base na análise de correlação e na flexibilidade de obtenção dos dados socioeconômicos. Na Etapa 4, o processo de regressão *Stepwise* é executado, o qual conta com os parâmetros de VIF e de significância.

Com o intuito de avaliar o comportamento do processo de regressão perante o aumento ou a diminuição de seus parâmetros estatísticos, nesta metodologia, propõe-se um estudo de sensibilidade do valor do VIF e da significância, nos atributos selecionados e valores associados ao R^2 acumulado e parcial. Por meio dos novos resultados alcançados no processo de seleção, propõe-se compreender a relevância de determinados atributos frente as variações impostas. Dessa forma, procura-se avaliar o número de atributos selecionados, além dos valores de R^2 acumulado e parcial.

Dado

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k + \varepsilon, \quad (2)$$

calcula-se:

$$VIF_{i,j} = \frac{1}{1 - R_{i,j}^2}, \quad (3)$$

onde $VIF_{i,j}$ é o VIF para β_i na j -ésima iteração, e $R_{i,j}$ é o coeficiente de determinação parcial da equação de regressão para X_i na j -ésima iteração. O critério de parada é dado por:

$$VIF_{i,j} < VIF_{ref} \forall i, \quad (4)$$

onde VIF_{ref} é o parâmetro referência, geralmente estabelecido na resolução vigente. O processo iterativo se repete enquanto (4) não estiver satisfeita.

O nível de significância é a probabilidade de rejeitar a hipótese nula quando ela é verdadeira (conhecido como erro do tipo I). No caso, a hipótese nula e a hipótese alternativa, são respectivamente:

$$H_0: \beta_1 = \beta_2 = \dots = \beta_k = 0, \quad (5)$$

$$H_a: \beta_1 = \beta_2 = \dots = \beta_k \neq 0. \quad (6)$$

Rejeita-se a hipótese nula se:

$$F_0 > F_{\alpha}, \quad (7)$$

onde F_{α} é o valor da distribuição F para o nível de significância α , e F_0 calcula-se conforme:

$$F_0 = \frac{MS_i}{MS_E}, \quad (8)$$

onde MS_i é a média quadrática da soma dos quadrados da variabilidade para diferentes níveis do atributo i , e MS_E é a média quadrática da soma dos quadrados da variabilidade dentro de um mesmo tratamento. Na regressão linear por *Stepwise*, para o atributo entrar e permanecer no modelo, é preciso que (7) seja satisfeita para cada iteração.

Dentro da regressão *Stepwise*, (4) e (7) são impactadas pelas alterações dos parâmetros de sensibilidade, posteriormente testados nos estudos de casos descritos na Seção 4.

3.3 Clusterização

A clusterização é o processo executado após as sete etapas destacadas na Fig. 1, sendo o método dinâmico utilizado para o agrupamento dos conjuntos semelhantes em *clusters* (Tanure, 2004). Deve-se salientar que nesse método não há a formação de agrupamentos fixos, como em outros utilizados anteriormente (ANEEL, 2016). Assim, para uma aplicação generalizada do método dinâmico, é preciso estabelecer os seguintes critérios:

- grau de heterogeneidade percentual permitida (refere-se à média das distâncias euclidianas dos conjuntos ao conjunto de referência): 20%;
- número desejável de conjuntos que devem ser selecionados como semelhantes para aplicação da análise comparativa: 100 conjuntos;
- número mínimo de conjuntos que devem ser selecionados como semelhantes para aplicação da análise comparativa: 50 conjuntos.

Assim, destaca-se como dados de entrada e saída para a clusterização, as seguintes informações:

- entrada: atributos selecionados;
- saída: limites de DEC e FEC.

Perante as mudanças sugeridas na etapa de seleção dos atributos, determinadas variáveis são afetadas, impactando (9), a qual é relacionada a padronização dos atributos por meio do Score Z, que transforma cada variável em uma nova variável com média zero e desvio padrão unitário. Dessa maneira, pode-se observar que em:

$$x_{il} = \frac{x_{il}^* - m_l}{s_l}, i = 1, \dots, N; l = 1, \dots, d, \quad (9)$$

o x_{il}^* é o dado original, m_l é a média amostral, s_l é o desvio-padrão amostral, N o número de conjuntos e d o número de atributos. Além disso, (10) também é afetada, a qual é responsável pela medida de similaridade, indicada pela distância Euclidiana:

$$D(x_i, x_j) = \sqrt{\sum_{l=1}^d |x_{il} - x_{jl}|^2}, \quad (10)$$

onde x_i e x_j são os conjuntos de UC de uma matriz de dados com d dimensões, onde d é o total de atributos utilizados.

Além de (9) e (10), há expressões utilizadas para o cálculo da heterogeneidade, técnica *Yardstick competition* para definição dos limites e aplicação da métrica denominada Score ANI, a qual visa avaliar os conjuntos classificados como heterogêneos. Logo, os atributos selecionados e a sua quantidade impactam diretamente no processo de cálculo dentro da clusterização.

3.4 Análise estatística

Com as variações dos parâmetros de sensibilidade, a seleção de outros atributos e em quantidades diferentes é alcançada, repercutindo em novos valores para os limites DEC e FEC. Logo, realiza-se uma análise estatística pela variação do limite objetivo para o DEC e FEC, dado por:

$$\Delta = (DEC \text{ ou } FEC_{\text{Caso}} - DEC \text{ ou } FEC_{\text{Caso base}}), \quad (11)$$

$$\Delta\% = (DEC \text{ ou } FEC_{\text{Melhorou}} - DEC \text{ ou } FEC_{\text{Piorou}}) \quad (12)$$

sendo (11) associada ao cálculo da variação do limite objetivo e (12), à diferença percentual entre os conjuntos que melhoraram e pioraram. Em (11), pode-se observar que é realizado o cálculo pela diferença entre os limites alcançados para o caso em estudo e os limites do caso base, representados pelos valores observados na NT102/14. Assim, por meio de (12) é possível se efetuar a seguinte classificação dos resultados:

- **neutro**: Cenário sem mudanças no limite ($\Delta\%=0$);
- **piorou**: Cenário com redução do limite ($\Delta\%<0$);
- **melhorou**: Cenário com aumento do limite ($\Delta\%>0$).

Por fim, destacam-se como parâmetros para o cálculo do limite objetivo, as seguintes informações:

- **entrada**: limites do caso estudado e do caso base;
- **saída**: Δ (DEC e FEC).

A seguir, a metodologia apresentada é analisada para os estudos de casos reais.

4. DESCRIÇÃO DOS ESTUDOS DE CASOS

Com o objetivo de analisar a sensibilidade estatística dos atributos físico-elétricos no cálculo dos indicadores, o método proposto foi aplicado ao panorama nacional. A aquisição dos dados para o presente estudo, levou em consideração as informações coletadas na Audiência Pública nº 029/2014, onde a ANEEL propôs uma revisão da metodologia de definição dos limites para os indicadores de continuidade DEC e FEC dos conjuntos de UC das distribuidoras, a qual foi oficializada na NT102/14 de continuidade de sistemas de distribuição.

Para efetuar os estudos de casos, foi utilizado o software SAS OnDemand for Academics (SAS, 2021) na fase de seleção dos atributos. Por meio desse, estruturou-se um código para simular a regressão *Stepwise* com laço via SAS de forma automática. No que tange ao processo de clusterização, modelou-se um código no software MATLAB (MATLAB, 2021), o qual permite ao usuário informar os atributos de entrada. Dessa forma, o método dinâmico é executado, fornecendo os limites para os indicadores DEC e FEC.

Perante os tópicos abordados até o momento, os estudos de casos serão analisados para duas situações, sendo a primeira associada a 2.523 conjuntos do Brasil e a segunda para 61 conjuntos de uma distribuidora da região sul do país, ambos levando em consideração os dados da Audiência Pública nº 029/2014. Logo, a partir da pesquisa implementada até o momento, serão implementados quatro estudos de casos, cada qual com seus cenários.

Caso 1. *Proposição de um único atributo e análise do FEC para 5 atributos selecionados*: No Caso 1, foram ilustrados dois cenários, sendo o primeiro associado ao emprego de um único atributo. O embasamento teórico para uso de um único atributo, diretamente, na etapa de clusterização, teve origem no projeto de P&D CELESC / Daimon (ANEEL, 2020). No projeto, conclui-se que o emprego de um único atributo associado a densidade de UC, seria mais relevante para avaliação dos limites de DEC e FEC. Desse modo, utilizou-se o atributo de densidade de UC (NUC_AREA) da NT102/14, por ser o mais semelhante ao P&D em questão. Já o segundo cenário, está associado as simulações efetuadas para reproduzir o modelo de seleção dos atributos da ANEEL no software SAS, onde foi possível encontrar os mesmos 6 atributos utilizados para o DEC. Porém, quando a mesma análise foi aplicada para o FEC, encontrou-se 5 atributos finais, onde nem todos estão entre os 6 atributos utilizados para o FEC pela ANEEL.

Caso 2. Mudança da significância: Para o Caso 2 foi proposto a variação do parâmetro de sensibilidade associado a significância para entrada ou saída do modelo, dentro do processo de regressão *Stepwise*. O valor padrão de significância empregado pela ANEEL é de $\alpha=1\%$, sendo neste caso implementado um estudo para dois cenários, $\alpha=0,5\%$ e $\alpha=3\%$.

Caso 3. Mudança do limite VIF: No Caso 3, foi proposto a variação do parâmetro de sensibilidade associado ao VIF, dentro do processo de regressão *Stepwise*. O valor padrão do VIF empregado pela ANEEL é de 5 ($VIF_{ref} = 5$), sendo, neste caso, implementado um cenário para o valor de $VIF_{ref} = 2,5$ e outro $VIF_{ref} = 8$.

Caso 4. Proposição de um único atributo: Para o Caso 4, buscou-se aplicar, somente, o primeiro atributo a entrar no modelo para o DEC e FEC, diretamente na etapa de clusterização. Logo, realizou-se uma análise para o primeiro atributo a entrar no modelo para o DEC, referente ao atributo percentual de UC em áreas de alta densidade (PUC_NUC_AD), e o primeiro a entrar para o FEC, referente ao atributo percentual de área com vegetação remanescente alta ou média (PC_VRAM).

5. RESULTADOS E DISCUSSÕES

A partir dos estudos de casos estruturados com auxílio dos softwares SAS e MATLAB, analisou-se os resultados nas Tabelas 1 e 2, considerando os 2.523 e 61 conjuntos. As colunas Atrib. DEC / ANEEL e Atrib. FEC / ANEEL contabilizam quantos atributos foram utilizados para o DEC ou FEC em paralelo com quantos atributos são iguais aos utilizados atualmente pela ANEEL. Já nas colunas ΔDEC (%) e ΔFEC (%) é realizada a diferença entre o percentual de conjuntos que melhoraram e pioraram seus limites, conforme (12), sendo o resultado em vermelho para uma maioria pior, em azul melhor e verde neutro.

5.1 Estudo com 2523 conjuntos do Brasil

A base de dados da ANEEL, possui um total de 2610 conjuntos. Porém, foi necessário retirar 87 conjuntos que não possuíam dados dos atributos, resultando em 2.523 conjuntos.

Tabela 1. Resumo para os 2523 conjuntos do Brasil (Neutro / Melhorou / Piorou)

Caso e cenários		Atrib. DEC / ANEEL	Atrib. FEC / ANEEL	Δ DEC (%)	Δ FEC (%)
1	NUC_AREA	1/0	1/0	-4,08	-10,58
	5 Atrib. FEC	6/6	5/3	0	-8,32
2	0,5%	6/4	5/3	1,98	-1,75
	3%	6/6	6/4	0	-9,36
3	2,5	6/3	5/2	-1,15	-13,48
	8	6/6	6/4	0	-9,36
4	PC_NUC_AD	1/6	1/6	-2,54	-2,22
	PC_VRAM	1/6	1/6	-10,58	-6,10

Tabela 2. Resumo para os 61 conjuntos da distribuidora da região sul (Neutro / Melhorou / Piorou)

Caso e cenários		Atrib. DEC / ANEEL	Atrib. FEC / ANEEL	Δ DEC (%)	Δ FEC (%)
1	NUC_AREA	1/0	1/0	32,79	-4,92
	5 Atrib. FEC	6/6	5/3	0	-22,95
2	0,5%	6/4	5/3	-8,20	-32,78
	3%	6/6	6/4	0	-21,32
3	2,5	6/3	5/2	-6,56	-27,87
	8	6/6	6/4	0	-21,32
4	PC_NUC_AD	1/6	1/6	21,31	-3,28
	PC_VRAM	1/6	1/6	8,20	9,83

No que confere ao Caso 1, referente ao cenário com uso, somente, do atributo NUC_AREA, verificou-se uma piora nos limites, tanto para o DEC quanto para o FEC, conforme ilustrado na Tabela 1. Já para o cenário com os 5 atributos encontrados para o FEC, verificou-se uma piora no limite de FEC, conforme a Tabela 1.

No Caso 2, a redução do critério de significância para entrada e saída do modelo para o valor de $\alpha=0,5\%$, resulta, em termos gerais, em um parâmetro mais restritivo para seleção dos atributos. A Tabela 3 apresenta os atributos gerados pelo SAS e os coeficientes de determinação do modelo e parcial para o DEC por iteração.

Tabela 3. R² do modelo e parcial para o DEC no Caso 2 ($\alpha=0,5\%$)

Step	Atributo	R ² modelo	R ² parcial
1	VRAM	0,36	0,362
2	PC_NUC_AD	0,44	0,075
3	PLUV	0,49	0,056
4	CM_NUC_RES	0,54	0,044
5	NUC_RUR	0,57	0,033
6	PC_ERMT_3F	0,59	0,021
7	PC_NUC_COM	0,60	0,012
8	PC_NUC_IND	0,61	0,005
9	NUC_COM	0,61	0,002
10	PC_EST_ARR	0,61	0,003

Com a redução pela metade do critério de significância, 10 atributos são gerados na última simulação em vez de 9 como na NT102/14. Logo, percebe-se que o número de atributos resultantes no final do processo de seleção é maior com a redução da significância.

A Tabela 4 segue as mesmas informações e resultados apresentados para o DEC, porém associados ao FEC. Verifica-se que com a redução pela metade da significância, 14 atributos são gerados na última simulação em vez de 15 como na NT102/14. Ao contrário do DEC, para o FEC o número de atributos foi reduzido com a diminuição da significância.

Tabela 4. R² do modelo e parcial para o FEC no Caso 2
($\alpha = 0,5\%$)

Step	Atributo	R ² modelo	R ² parcial
1	PC_VRAM	0,25	0,253
2	PLUV	0,37	0,120
3	TD_ERMT	0,51	0,133
4	PC_NUC_AD	0,55	0,040
5	PC_NUC_RES	0,58	0,029
6	CM_NUC	0,59	0,011
7	NUC_COM	0,60	0,017
8	PC_ERMT_1F	0,61	0,008
9	TD_1F	0,62	0,005
10	NUC_RUR	0,62	0,003
11	CM_NUC_RES	0,62	0,002
12	DENS_EST_PAV	0,62	0,002
13	ALIM	0,63	0,002
14	PC_ENE_RUR	0,63	0,001

A mesma análise efetuada para a significância de 0,5%, foi realizada considerando 3%, concluindo-se que tornar o modelo mais ou menos restritivo, tratando-se da significância, não justifica, necessariamente, possuir menos ou mais atributos selecionados no final em relação ao caso base. Logo, a mudança percebida nas simulações fica associada a quais atributos foram selecionados e não a quantidade que permaneceu no modelo.

No Caso 3, foi realizada a mesma avaliação que para o caso 2, porém associada a variação do VIF_{ref} . Tanto para o cenário com VIF_{ref} igual a 2,5 como para 8, verificou-se o mesmo que para os cenários do Caso 2, onde se observou que tornar o processo mais ou menos restritivo, tratando-se do valor de VIF, não justifica possuir menos ou mais atributos selecionados no final em relação ao caso base. A mudança fica associada a quais atributos foram selecionados e não a quantidade. Além disso, vale destacar que um valor acima de 8 para o VIF_{ref} , não gera nenhuma mudança para o modelo.

Deve-se salientar, considerando-se o R² do modelo, que próximo do sétimo atributo, a curva tende a adquirir uma característica linear, independente da variação de sensibilidade aplicada, conforme ilustra a Fig. 2. No que confere ao valor de 2%, estabelecido para o R² parcial pela ANEEL, verificou-se que abaixo desse, o coeficiente parcial tende a cair muito, comprovando-se a aplicação do valor atual de corte, conforme indicado pela Fig. 3.

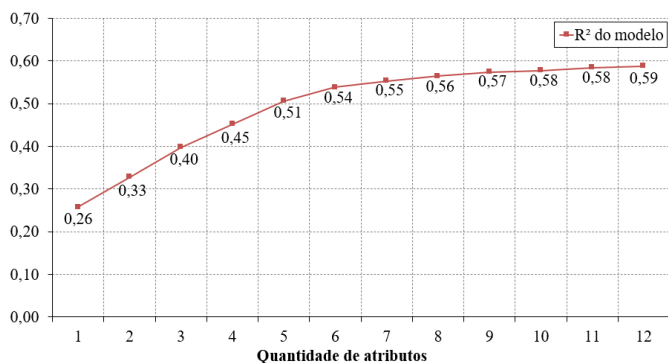


Fig. 2 R² do modelo para o DEC no Caso 3 ($VIF_{ref} = 2,5$).

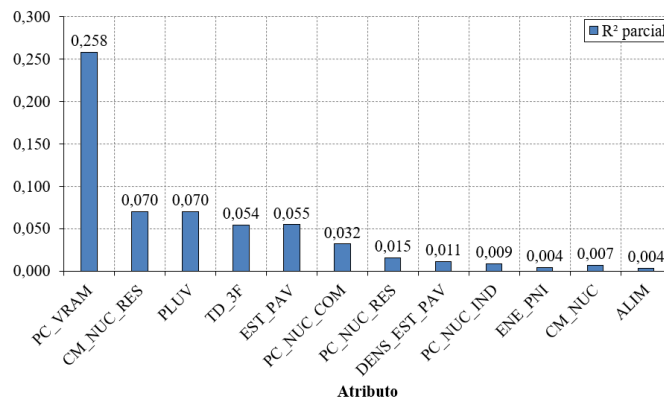


Fig. 3 R² parcial para o DEC no Caso 3 ($VIF_{ref} = 2,5$).

Para o estudo de Caso 4, aplicando, somente, o atributo PC_NUC_AD, observou-se que a análise do percentual dos conjuntos para o DEC, apresentou a maioria dos conjuntos em uma situação pior. Já para o FEC, a maioria dos conjuntos permaneceu neutra, sendo o percentual de conjuntos que pioraram superior aos que melhoraram. Para o cenário, somente, com o atributo PC_VRAM, verificou-se o mesmo resultado para ambos os indicadores. A análise percentual dos conjuntos com as respectivas situações, apresentou uma situação de piora em sua maioria, conforme a Tabela 1.

5.2 Estudo com 61 conjuntos da distribuidora da região sul

Os estudos de casos apresentados até o momento levaram em consideração 2.523 conjuntos. Dessa forma, com o intuito de avaliar os mesmos estudos de casos para os conjuntos de abrangência de uma distribuidora do sul do país, realizam-se as mesmas simulações para os 61 conjuntos.

Para o Caso 1, aplicando, somente, o atributo NUC_AREA, foi possível verificar uma melhoria para a maioria dos conjuntos associado ao DEC, porém para o FEC, a maioria dos conjuntos permaneceu neutra, sendo o percentual de conjuntos que pioraram superior aos que melhoraram. No cenário envolvendo os 5 atributos encontrados para o FEC, os resultados indicaram a mesma situação do FEC para o NUC_AREA.

O Caso 2, associado a redução da significância para 0,5%, resultou em um cenário semelhante para ambos os indicadores. A análise percentual indicou que a maioria dos conjuntos permaneceu com os limites neutros, sendo o percentual de conjuntos que pioraram superior aos que melhoraram. Com o aumento da significância para 3%, observou-se a totalidade dos conjuntos na situação neutra para o DEC. Já para o FEC, o resultado foi semelhante ao cenário de 0,5%.

O Caso 3, envolvendo a alteração de $VIF_{ref} = 2,5$, resultou na maioria do percentual dos conjuntos em uma situação pior para o DEC. Já para o FEC, a maioria dos conjuntos permaneceu neutra, sendo o percentual de conjuntos que pioraram superior aos que melhoraram. Para o cenário com $VIF_{ref} = 8$, a análise percentual para o DEC mostrou a totalidade dos conjuntos na situação neutra. Tratando-se do FEC, o resultado foi semelhante ao FEC com $VIF_{ref} = 2,5$.

No Caso 4, somente, com o atributo PC_NUC_AD, observa-se que a maioria dos conjuntos obteve uma melhoria para o DEC. Para o FEC, a maioria dos conjuntos permaneceu neutra, sendo o percentual de conjuntos que pioraram superior aos que melhoraram. Já para o cenário, somente, utilizando o atributo PC_VRAM, verifica-se que a maioria dos conjuntos obtiveram uma melhoria para o DEC e FEC. Deve-se salientar que esse último cenário foi o único a apresentar melhoria para ambos os indicadores, levando-se em consideração todos os estudos realizados, conforme é ilustrado pela Fig. 4.

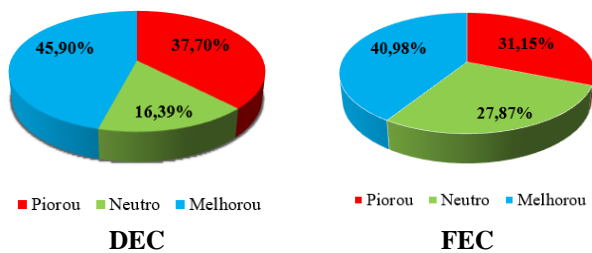


Fig. 4 Resultado percentual para o estudo de caso empregando somente o atributo PC_VRAM.

6. CONCLUSÕES

O presente trabalho, visou avaliar o impacto na variação do limite objetivo para o DEC e FEC, por meio da mudança da sensibilidade dos parâmetros estatísticos no modelo vigente da ANEEL. Em todas as variações de sensibilidade para o DEC, os atributos de precipitação pluviométrica média anual (PLUV) e consumo médio por UC da classe residencial (CM_NUC_RES) permaneceram na faixa de 2% do R^2 parcial. Já para as variações de sensibilidade do FEC, os atributos PLUV e PC_VRAM permaneceram na faixa de 2% do R^2 parcial. O aumento da significância para 3% e do valor de VIF para 8 não afetou o resultado para o DEC, pois os atributos com R^2 parcial acima de 2% permaneceram iguais aos da NT102/14. Quando se avalia os casos propostos, conclui-se que o FEC é mais sensível as mudanças inseridas.

No que confere a análise geral dos casos, a maioria dos conjuntos permaneceu em uma situação neutra ou pioraram seus limites objetivos. Além disso, os resultados alcançados para os 61 conjuntos da distribuidora da região sul e os 2.523 conjuntos do Brasil, apresentam algumas diferenças, indicando que há peculiaridades pertencentes a cada área de concessão. Desse modo, o modelo vigente da agência pode receber contribuições baseadas nos estudos realizados, as quais podem interferir diretamente na qualidade do serviço.

AGRADECIMENTOS

Ao Programa de P&D regulado pela ANEEL e executado pela Companhia Estadual de Distribuição de Energia Elétrica (CEEE-D) via chamada nº 01/2018. Ao suporte técnico e financeiro do Instituto Nacional de Ciência e Tecnologia em Geração Distribuída (INCT-GD) da Universidade Federal de Santa Maria - UFSM (CNPq processo nº 465640/2014-1, CAPES processo nº 23038.000776/2017-54 e FAPERGS nº 17/2551-0000517-1). Ao PIBIC CNPq-UFRGS (Programa Institucional de Bolsas de Iniciação Científica) pelo apoio financeiro.

REFERÊNCIAS

- ANEEL – Agência Nacional de Energia Elétrica., (2014). “Revisão da metodologia de definição de limites para os indicadores de continuidade DEC e FEC das distribuidoras.” Nota Técnica nº102/2014, Agência Nacional de Energia Elétrica – ANEEL.
- ANEEL (2016). “Estabelecimento dos limites para os indicadores de continuidade DEC e FEC dos conjuntos da Companhia Estadual de Distribuição de Energia Elétrica – CEEE-D, para o período 2017 a 2021.” Nota Técnica nº121/2016, Agência Nacional de Energia Elétrica – ANEEL.
- ANEEL (2020). “Definição de conjuntos de unidades consumidoras eficientes sob os enfoques dos consumidores, das empresas e do regulador”. Workshop P&D-05697-0718/2018, Agência Nacional de Energia Elétrica – ANEEL.
- Costa, Marcelo Azevedo; Lopes, Ana Lúcia Miranda; Matos, Giordano Bruno Braz de Pinho (2015). Statistical evaluation of Data Envelopment Analysis versus COLS Cobb–Douglas benchmarking models for the 2011 Brazilian tariff revision. *Socio-Economic Planning Sciences*, v. 49, p. 47-60.
- MATLAB and Statistics Toolbox Release 2021a, The MathWorks, Inc., Natick, Massachusetts, United States.
- Medeiros, Giulia OS *et al* (2022). Efficiency analysis for performance evaluation of electric distribution companies. *International Journal of Electrical Power & Energy Systems*, v. 134, p. 107430.
- Pessanha, J. F. M. (2006). “Um modelo de análise envoltória de dados para estabelecimento das metas de continuidade de fornecimento de energia elétrica”. Doutorado em engenharia elétrica, Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro, Brasil.
- SAS Corporate Compliance. Software de Business Analytics e Business Intelligence. SAS OnDemand for Academics. Cary, NC, USA: SAS Institute Inc, 2021.
- Soroush, Golnoush *et al* (2021). Network utilities performance and institutional quality: Evidence from the Italian electricity sector. *Energy economics*, v. 96, p. 105177.
- Tanure, J. E. P. S. (2004). “Proposta de procedimentos e metodologia para estabelecimento de metas de qualidade (DEC e FEC) para concessionárias de distribuição de energia elétrica através da análise comparativa”. Doutorado em engenharia elétrica, Universidade de São Paulo, São Paulo, Brasil.
- Tanure, J. E. P. S.; Tahan, Carlos Marcio Vieira; Lima, J. W. Marangon (2006). Establishing quality performance of distribution companies based on yardstick regulation. *IEEE Transactions on Power Systems*, v. 21, n. 3, p. 1148-1153.
- Xavier, S. S; Lima, J. W., Lima, L.M., Lopes, A. L. M. (2015). How efficient are the brazilian electricity distribution companies?. *Journal of Control, Automation and Electrical Systems*, v. 26, n. 3, p. 283-296.