

Clusterização Fuzzy Otimizada para Estabelecimento de Limites Coletivos dos Indicadores de Continuidade

Lucas M. Thimoteo* Thais R. Borges** Marley M. Vellasco*
Ricardo Tanscheit*

* *Departamento de Engenharia Elétrica, Pontifícia Universidade Católica do Rio de Janeiro, RJ, (e-mails:lucasthim@yahoo.com, marley@ele.puc-rio.br, ricardo@ele.puc-rio.br).*

** *Departamento de Engenharia Elétrica, Universidade do Estado do Rio de Janeiro, RJ (e-mail: thais.roupe@gmail.com)*

Abstract: This work proposes a new methodology to establish the collective continuity indicators for consumer sets of the electrical energy distribution companies of Brazil. This methodology executes an optimized fuzzy clustering, using the Fuzzy C-Means and the Particle Swarm Optimization algorithms, aiming to make the consumer sets grouping and the target calculation of their continuity index target more flexible. The clustering was done for each geographical region of Brazil and results are comparable to the current dynamic clustering methodology, however with a faster decay throughout the index progression.

Resumo: Este trabalho propõe uma nova metodologia de cálculo dos limites coletivos dos indicadores de continuidade das distribuidoras de energia elétrica no Brasil. Esta metodologia tem como base a execução de uma clusterização fuzzy otimizada, utilizando o algoritmos Fuzzy C-Means e de Otimização por Enxame de Partículas, com o intuito de flexibilizar o agrupamento de conjuntos de unidades consumidoras e o cálculo de metas para seus indicadores coletivos de continuidade. A clusterização foi realizada por regiões do Brasil e os resultados obtidos são comparáveis à metodologia atual, porém com uma trajetória de redução mais acentuada.

Keywords: Fuzzy clustering; Particle swarm; Electricity Supply Quality; Regulation; index target establishment

Palavras-chaves: Clusterização fuzzy; Enxame de partículas; Qualidade de Serviço; Regulação; Estabelecimento de limites;

1. INTRODUÇÃO

O segmento de distribuição de energia elétrica no Brasil é representado por um monopólio natural (concessão de áreas para as empresas). Desta forma, faz-se necessário o estabelecimento de normas quanto à tarifação, licitação, limite de área de concessão e qualidade do serviço prestado.

A Agência Nacional de Energia Elétrica (ANEEL), juntamente com outras entidades governamentais, exerce a função de órgão regulador e fiscalizador do setor elétrico, observando o atendimento, por parte das concessionárias, dos compromissos contratuais estabelecidos Dias (2002). Em relação à regulação da qualidade do serviço prestado, tem-se que uma das principais formas para a sua fiscalização é o estabelecimento de limites para os indicadores coletivos de continuidade, tanto à nível global (para a empresa), quando para seus conjuntos de unidades consumidoras¹.

A metodologia atual para o estabelecimento de limites para os indicadores coletivos de continuidade baseia-se

em uma clusterização dinâmica, estabelecendo trajetória de metas para cada conjunto de unidades consumidoras de uma concessionária, considerando seus semelhantes ANEEL (2014a). Um dos principais desafios, tanto para as distribuidoras quanto para a ANEEL, é definir modelos suficientemente explicativos que possam estabelecer limites condizentes com a realidade da concessionária e, ao mesmo tempo, que possibilitem um incentivo à melhoria contínua e progressiva da prestação de serviço.

A ANEEL busca constantemente aprimoramentos em suas metodologias, contudo poucos trabalhos acadêmicos foram desenvolvidos no âmbito da definição de limites para indicadores de continuidade nos últimos 20 anos. Muitas das fontes de análise a respeito de modificações nos métodos para estabelecimento de limites de continuidade são oriundos das próprias concessionárias, dado que este tema é pouco difundido entre os consumidores finais.

Neste sentido, este trabalho propõe uma clusterização fuzzy otimizada considerando todo o universo de conjuntos do Brasil, além de uma nova métrica de cálculo para os limites de cada conjunto. A Otimização por Enxame de Partículas (*Particle Swarm Optimization* - PSO) é utilizada para mitigar possíveis mínimos locais que podem

¹ Um conjunto de unidades consumidoras é caracterizado por fatores como potência instalada, densidade de carga, área, entre outros, que estão definidos no Módulo 8 do PRODISTANEEL (2018).

ocorrer dependendo da inicialização do algoritmo de clusterização. A clusterização fuzzy permite que uma amostra pertença simultaneamente a dois agrupamentos a partir da definição de um grau de pertinência. Sendo assim, a metodologia proposta fará uso do grau de pertinência dos conjuntos com vistas à ponderar as semelhanças destes em relação aos seus agrupamentos.

As seções a seguir estão divididas em: 2. Qualidade do Serviço, 3. Trabalhos Correlatos, 4. Metodologia Vigente para Estabelecimento de Metas, 5. Metodologia Proposta, 6. Resultados da Clusterização, 7. Simulação para Definição de Limites Coletivos para a Equatorial Pará, 8. Conclusão.

2. QUALIDADE DO SERVIÇO

As concessionárias de distribuição de energia elétrica são avaliadas, entre outros fatores, sob a ótica da qualidade do produto e do serviço durante o período de revisão tarifária. A qualidade do produto busca observar fatores como a forma de onda e as perturbações dos níveis de tensão, o fator de potência, o controle de frequência, entre outros. Já a qualidade do serviço consiste em mensurar a quantidade e a duração das interrupções do fornecimento de energia elétrica. Para o consumidor final, principalmente os residenciais e os comerciais, a qualidade do serviço tende a ser mais valorada, visto que impacta diretamente nas rotinas diárias e nas atividades fins de cada usuário.

2.1 Indicadores Coletivos de Continuidade

A fim de mensurar e de regular a qualidade de serviço prestado pelas concessionárias, foram criados os indicadores de continuidade DNAEE (1978). Eles são divididos em indicadores individuais e coletivos. Os primeiros são obtidos por meio de medição direta feita em cada unidade consumidora. Já os últimos são calculados a partir dos indicadores individuais e para cada conjunto da concessionária, sendo identificados como DEC (Duração Equivalente de Interrupção por Unidade Consumidora) e FEC (Frequência Equivalente de Interrupção por Unidade Consumidora).

$$DEC = \frac{\sum_{i=1}^n Ca(i) * T(i)}{Cs} \quad (1)$$

Em que:

i = quantidade de interrupções de 1 a n ,

$T(i)$ = duração de cada interrupção do conjunto considerado, em horas,

$Ca(i)$ = quantidade de consumidores atingidos do conjunto considerado,

Cs = quantidade total de consumidores do conjunto considerado.

$$FEC = \frac{\sum_{i=1}^n Ca(i)}{Cs} \quad (2)$$

Em que:

i = quantidade de interrupções de 1 a n ,

$Ca(i)$ = quantidade de consumidores atingidos do conjunto considerado,

Cs = quantidade total de consumidores do conjunto considerado.

As demais informações sobre os indicadores podem ser encontradas na seção 8.2 do Módulo 8 do PRODIST ANEEL (2018).

Para caracterizar um serviço de boa qualidade, espera-se apurar patamares baixos para os indicadores DEC e FEC, ou seja, poucas interrupções e de curta duração. Comumente, os indicadores DEC e FEC são calculados para os períodos mensal e anual. A cada revisão tarifária de uma concessionária, são definidos novos limites de indicadores a serem aplicados no ciclo tarifário seguinte. As metas regulatórias somente são revisadas e atualizadas ao final de cada ciclo tarifário, sendo este variante entre 3 a 5 anos, a depender do contrato de concessão de cada distribuidora.

3. TRABALHOS CORRELATOS

Em Sperandio (2004), foi feita uma classificação de conjuntos através de mapas auto-organizáveis. Os resultados aqueles obtidos são comparáveis à clusterização k-means. Os dados analisados foram apenas de uma distribuidora e seus atributos foram relativos aos municípios que faziam parte da área de concessão da companhia. A premissa de separação por municípios diverge da definição considerada pela ANEEL, porém facilita a caracterização e o agrupamentos dos consumidores por permitir o uso de atributos como a área do município e a distância da sede do município à subestação. Os atributos inicialmente considerados (variáveis elétricas e de mercado) foram avaliados e reduzidos por meio de uma análise estatística multivariada. Após o agrupamento, verificou-se que os conjuntos foram divididos em 30 clusters.

Em Pessanha (2007), foi proposto o estabelecimento de limites coletivos utilizando uma análise envoltória de dados (*Data Envelopment Analysis* - DEA), considerando todos o universo de conjuntos do Brasil. Primeiramente, realizou-se uma clusterização das distribuidoras utilizando um mapa auto-organizável e estabeleceu-se metas globais por meio do DEA. Em seguida, foram definidos os indicadores locais segundo regiões da concessão, de maneira que os indicadores de todos os conjuntos de uma mesma distribuidora observassem melhoras significativas. O autor aplicou sua metodologia aos conjuntos da Light SESA, concessionária localizada no estado do Rio de Janeiro. Os resultados foram satisfatórios, pois assemelharam-se à progressão histórica da distribuidora e também à metodologia vigente à época.

4. METODOLOGIA VIGENTE PARA ESTABELECIMENTO DE METAS

A metodologia vigente para a definição de limites coletivos de continuidade do serviço foi proposta, em ANEEL (2014a). À época, foi realizada uma Audiência Pública que permitiu a contribuição, acerca deste tema, de diversas concessionárias de distribuição, bem como de outros agentes do setor.

Atributos suficientemente explicativos foram selecionados - a partir de modelos de regressão - para compor as variáveis da clusterização dinâmica, processo em que o conjunto torna-se o centróide para formar seu cluster. Em seguida, um grau de heterogeneidade do conjunto em relação ao seu cluster é calculado, permitindo mensurar o nível de comparabilidade do conjunto em relação aos seus conjuntos *benchmarkings*. Por fim, define-se um indicador que será classificado como alvo, cuja trajetória de decréscimo do conjunto em análise será definida para alcançar este valor em até 8 anos.

4.1 Clusterização Dinâmica

O objetivo da clusterização dinâmica consiste em agrupar conjuntos que mais se assemelham ao conjunto em análise, sendo este comparado com os demais por meio de um grau heterogeneidade.

É definido que um agrupamento possua no máximo de 100 conjuntos. Entretanto, caso o grau de heterogeneidade calculado para os 100 comparáveis supere 20% (ou 0,2), o conjunto mais distante é retirado da análise, e a heterogeneidade é recalculada. Contudo, considera-se 50 conjuntos como o mínimo de comparáveis necessários. Desta forma, mesmo que o grau de heterogeneidade ainda permaneça superior a 20%, ao chegar a 50 conjuntos não se retira mais nenhum conjunto do agrupamento. Neste último caso, o conjunto em análise é classificado como um conjunto heterogêneo.

Os detalhes da clusterização são encontrados na seção III.3.1 da Norma Técnica 0059/2014 ANEEL (2014a) e na seção III.3.2 da Norma Técnica 0102/2014 ANEEL (2014b).

4.2 Cálculo do DEC/FEC Alvo

A principal etapa para determinar o DEC/FEC alvo do conjunto em análise é definir o percentil de referência do agrupamento. Em suma, define-se que, para conjuntos que sejam interligados ao Sistema Interconectado Nacional (SIN), o percentil inicial a ser considerado é o 20. Já para conjuntos que sejam isolados, ou seja, não possuem conexão com o SIN, atribui-se o percentil 50.

Em seguida, é realizada uma análise de complexidade para os conjuntos que possuam grau de heterogeneidade superior a 20%. Neste caso, a partir do Score de Atributo Normalizado Individual - Score ANI -, avalia-se se o conjunto em análise possui características mais complexas que seus comparáveis. Em caso positivo, o percentil inicialmente estabelecido poderá ser flexibilizado, caso contrário, poderá ser reduzido. O conjunto dentro do *rol* de semelhantes posicionado no percentil definido será o conjunto cujo indicador tornar-se-á o alvo do conjunto em análise.

Os detalhes destas considerações encontram-se na seção III.3.1 da Norma Técnica 0059/2014 ANEEL (2014a) e na seção III.3.2 da Norma Técnica 0102/2014 ANEEL (2014b).

4.3 Determinação da Trajetória de Redução

Após definidos os valores do DEC e do FEC alvos, determina-se uma trajetória linear decrescente com vistas

a alcançar o valor do indicador definido pelo alvo de referência. O ponto de partida considerado é o limite do ano vigente do conjunto em análise. A trajetória é calculada para um período de 8 anos. Vale ressaltar que, caso o valor alvo determinado seja superior ao ponto de partida, este último manter-se-á constante durante todo o período de cálculo.

A Equação 3 apresenta o equacionamento para definição da trajetória de decréscimo dos limites.

$$Limite^t = Limite^0 - t * \left(\frac{Limite^0 - Limite^{alvo}}{T} \right) \quad (3)$$

Em que:

T = período de transição, considerado 8 anos,

t = ano em que se calcula o limite,

$Limite^t$ = limite calculado no ano t ,

$Limite^0$ = limite atual já estabelecido,

$Limite^{alvo}$ = limite alvo calculado pelo percentil.

É importante ressaltar que é efetuado um arredondamento para cima tanto no valor alvo quanto nos valores intermediários da trajetória.

5. METODOLOGIA PROPOSTA

A metodologia proposta para a definição de limites coletivos para os indicadores de continuidade baseia-se na clusterização otimizada dos conjuntos, por meio do algoritmo Fuzzy C-Means e do algoritmo de Enxame de Partículas (*Particle Swarm Optimization* - PSO). A base de dados utilizada é a mesma utilizada pela ANEEL no processo de definição dos limites coletivos.

As fórmulas e algoritmos apresentados foram executados na linguagem de programação Python 3.7, com o auxílio de bibliotecas científicas e matemáticas como o NumPy, SciPy, Pandas, Matplotlib e DEAP.

5.1 Aquisição de Dados e Pré-processamento

Os dados utilizados neste trabalho foram solicitados à ANEEL através do portal e-Sic (Sistema Eletrônico do Serviço de Informações ao Cidadão). Os dados são referentes aos 3.045 conjuntos de unidades consumidoras do Brasil e estão referenciados ao ano de 2018.

Ressalta-se que foram considerados apenas conjuntos com padronização aérea, dado que os conjuntos classificados como subterrâneos performam apenas 1% do universo de conjuntos do Brasil. Além disso, estes conjuntos já configuram um agrupamento específico, pois possuem seus indicadores apurados e limites muito inferiores. Sendo assim, devem ser avaliados separadamente e não serão contemplados neste trabalho.

Para fins de comparação dos resultados, foram utilizados, no processo de clusterização, os mesmos atributos aplicados na metodologia vigente de definição de limites coletivos de continuidade.

Destaca-se que, devido à discrepância entre as escalas das variáveis, fez-se necessário a normalização dos dados com

média igual a 0 e desvio padrão igual a 1. Desta forma, os algoritmos de otimização e de clusterização não serão sensíveis à uma variável em detrimento das outras, por questões de escalas de distâncias.

5.2 Clusterização

O método escolhido para clusterização é o Fuzzy C-Means (FCM), que foi inicialmente proposto por Dunn (1973) e melhorado por Bezdek (1981).

A diferença primordial entre este método de clusterização e os demais métodos *crispy* (como o k-means) é que é possível que uma amostra pertença a mais de um agrupamento simultaneamente. Ou seja, cada amostra possui um grau de pertinência relativo a cada agrupamento.

O algoritmo FCM atualiza sequencialmente os centróides de seus agrupamentos e sua matriz de pertinência até que a convergência seja atingida. Esta acontece quando a atualização dos valores da matriz de pertinência é menor do que um limiar pré-estabelecido.

5.3 Otimização por Enxame de Partículas

A Otimização Por Enxame de Partículas foi inicialmente proposta por Kennedy and Eberhart (1995). O objetivo deste algoritmo é encontrar uma região ótima em um espaço de busca, explorando-o por meio de um enxame de partículas que se movimentam neste espaço.

Cada partícula do enxame é atualizada com base em um melhor local (aspecto cognitivo) e em um melhor global (aspecto social). Desta forma, espera-se que cada partícula explore sua vizinhança, enquanto dirige-se para uma área de ótimo global.

Uma partícula pode ser definida por sua posição e também por sua velocidade. A cada iteração do algoritmo, estes valores são atualizados iterativamente. O espaço vetorial da posição e da velocidade deve ser igual ao número de parâmetros do problema o qual se deseja otimizar.

O critério de parada desta otimização geralmente é definido por um número máximo de iterações (gerações) e/ou por um número máximo de iterações sem que ocorra atualização da melhor partícula encontrada até aquele momento.

5.4 Otimização da Clusterização

Os parâmetros de interesse para otimização são o número de clusters e os centróides resultantes da clusterização. Pela natureza do FCM e de outros algoritmos de clusterização, o sucesso ou a falha do método depende, em partes, da forma como os centróides dos clusters são inicializados. Caso ocorra uma inicialização desfavorável dos centróides, a convergência poderá não ocorrer dentro do número máximo de iterações.

Com vistas à mitigar o risco de mínimos locais durante a execução do FCM, propõe-se uma abordagem híbrida que consiste em executar uma segunda etapa de ajuste dos centróides dos clusters após a etapa do FCM.

Cada partícula da população é codificada pelos centróides da clusterização e seu valor inicial é o resultado final de uma clusterização FCM. A matriz de velocidade

da partícula possui as mesmas dimensões da matriz de posição, sendo inicializada por meio de uma distribuição uniforme entre -1 e 1. O número máximo de gerações para a otimização por enxame é de 150 e a população total é composta por 10 partículas. Caso não haja mudança na melhor partícula obtida em 50 gerações, o algoritmo PSO é finalizado.

Neste processo, ocorre uma inicialização e uma convergência para cada partícula da população do enxame. Cabe ressaltar que se almeja superar a ocorrência de mínimos locais a partir da utilização de diferentes configurações e de uma busca exploratória, ambas oferecidas pelo processo estocástico da PSO.

Em relação à base de dados utilizada, destaca-se que os conjuntos foram subdivididos de acordo com as regiões geográficas do Brasil, a fim de reduzir a heterogeneidade entre os dados. O número mínimo de clusters escolhido foi de 2, enquanto que o máximo foi definido por uma divisão média das amostras, ou seja, de 30 amostras por agrupamento até um máximo de 20.

A Tabela 1 apresenta os intervalos de variação dos números de clusters.

Tabela 1. Clusters por Regiões do Brasil

Região	Conjuntos	Clusters Mínimos	Clusters Máximos
CO	325	2	10
NE	748	2	20
NO	275	2	9
SE	1.227	2	20
SU	441	2	14

Para encontrar o número ótimo de clusters, utilizou-se o coeficiente Silhueta Intra-Inter Generalizada, cuja explicação está disposta na seção a seguir.

Em resumo, tem-se as seguintes etapas para o processo de otimização: (i) escolha do número de clusters; (ii) inicialização das partículas com os resultados dos FCM; (iii) otimização PSO para minimizar a função objetivo do FCM; (iv) escolha da melhor partícula (escolha dos centróides); (v) avaliação da clusterização com a silhueta generalizada; e (vi) escolha do número ótimo de clusters.

5.5 Silhueta Intra-Inter Generalizada

Para encontrar o número de clusters, foi utilizado o coeficiente Silhueta Intra-Inter Generalizada. Este coeficiente foi proposto por Rawashdeh and Ralescu (2012) e é um caso geral do coeficiente de silhueta apresentado em Rousseeuw (1987). Esta métrica busca avaliar tanto o grau de compactidade de um agrupamento (chamado de a_j) quanto o grau de separação entre agrupamentos (chamado de b_j), ponderando a distância entre os pontos (d_{jk}). Assim:

$$s_j = \frac{b_j - a_j}{\max(a_j, b_j)} \quad (4)$$

$$a_j = \min \left\{ \frac{\sum_{k=1}^n \text{IntraDist}_i(j, k) \cdot d_{jk}}{\sum_{k=1}^n \text{IntraDist}_i(j, k)} \right\} \quad (5)$$

$$b_j = \min \left\{ \frac{\sum_{k=1}^n \text{InterDist}_{rs}(j, k) \cdot d_{jk}}{\sum_{k=1}^n \text{InterDist}_{rs}(j, k)} \right\} \quad (6)$$

O valor da silhueta pode variar entre -1 e 1. Valores negativos demonstram que o agrupamento dos dados não foi realizado de maneira correta, pois a medida de distância Intra-Dist é maior que a Inter-Dist, indicando que uma amostra está mais próxima dos demais clusters do que ao seu cluster. Valores próximos de 1 indicam uma boa clusterização, pois a medida Intra-Dist é bem menor que a medida Inter-Dist. Sendo assim, é desejável que sejam obtidos valores superiores a 0 e próximos de 1.

5.6 Cálculo do DEC/FEC Benchmark

Para calcular o *benchmark* de um cluster, primeiramente são selecionados os conjuntos que tenham maior grau de pertinência em um respectivo cluster. Em seguida, estes são colocados em ordem crescente por DEC/FEC (do melhor/menor indicador para o pior/major).

Então, é calculada uma média ponderada entre os n melhores conjuntos do cluster (baseados em um percentil), sendo seus pesos os graus de pertinência ao cluster em questão.

$$benchmark_i = \frac{\sum_{j=1}^n u_{ij} \times indicador_j}{\sum_{j=1}^n u_{ij}} \quad (7)$$

Em que:

$$i = 1, 2, \dots, c,$$

$$j = 1, 2, \dots, n,$$

n = número de amostras selecionadas pelo percentil,

c = número do cluster.

O intuito dessa fórmula é expressar um *benchmark* que seja caracterizado não somente pelos melhores conjuntos, mas também pelos conjuntos que mais representam cada cluster. Logo, caso haja um conjunto com indicadores muito baixos (lembrando que quanto menor, melhor) e grau de pertinência baixo (provavelmente residindo em uma região de fronteira/sobreposição entre clusters), este não terá tanta contribuição no cálculo do *benchmark*.

5.7 Cálculo do DEC/FEC Alvo

O cálculo do DEC/FEC alvo para cada conjunto dá-se por uma média ponderada, em que os pesos são os graus de pertinência relativos a cada cluster. Logo:

$$alvo_{jp} = \frac{\sum_{j=1}^c u_{ij} \times benchmark_i}{\sum_{i=1}^c u_{ij}} \quad (8)$$

Em que:

$$i = 1, 2, \dots, c,$$

$$j = 1, 2, \dots, n,$$

p = percentil dos melhores conjuntos,

n = número de amostras selecionadas pelo percentil,

c = número do cluster,

benchmark = menor valor do DEC/FEC de um cluster.

Destaca-se que, para os conjuntos interligados, são considerados os 20% melhores indicadores de um cluster. Já para os conjuntos isolados, considera-se os 50% melhores.

5.8 Determinação da Trajetória de Limite

Para calcular a progressão dos limites, foi considerada uma trajetória linear decrescente, tal como a metodologia vigente. Além disso, também considerou-se um arredondamento para cima tanto nas progressões intermediárias quanto no valor alvo final.

6. RESULTADOS DA CLUSTERIZAÇÃO

A Figura 1 e a Figura 2 apresentam a variação da silhueta generalizada pelo número de clusters para duas regiões do Brasil.

Figura 1. Silhueta generalizada - DEC - Região CO

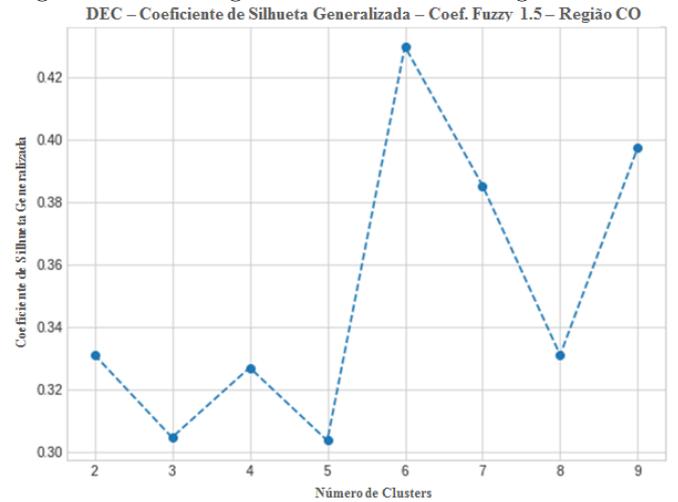
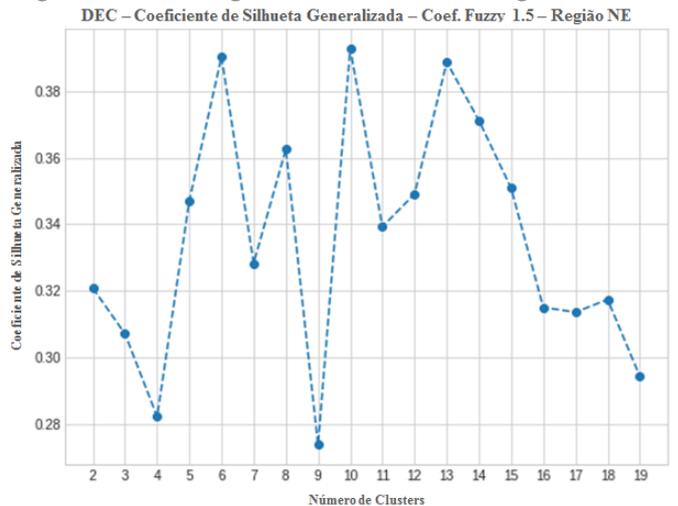


Figura 2. Silhueta generalizada - DEC - Região NE



A Figura 1 apresenta a variação da silhueta para o indicador DEC na região Centro-Oeste. Neste caso, a escolha do número ótimo de clusters fica evidente, dado que o valor máximo da silhueta generalizada (gSil) para 6 clusters é o maior desta distribuição.

Já na Figura 2, é perceptível a presença de três picos. O maior valor da gSil é obtido com 10 clusters. Porém, escolhendo o número de clusters como 6, obtém-se um valor da gSil próximo do valor máximo da Figura 1 e com um número reduzido de clusters.

Para a maioria dos casos, foi escolhido o número de clusters que obteve o maior valor de silhueta generalizada, à exceção de alguns casos em que poderia haver uma redução do número de clusters, sem que houvesse uma queda muito perceptível de performance (até 10% do valor máximo obtido). O coeficiente fuzzy (m) escolhido foi de 1.5, pois não é interessante criar regiões nebulosas (de sobreposição) muito vastas.

Por fim, a Tabela 2 e a Tabela 3 apresentam os resultados da etapa de clusterização. Nota-se que as regiões Sudeste e Sul foram aquelas que verificaram maior quantidade de cluster para o DEC, enquanto que, para o FEC, as regiões Nordeste e a Sudeste foram as que receberam maior quantidade de cluster.

Tabela 2. Resultado da Clusterização - DEC

Região	Clusters	Coef Fuzzy (m)	gSil
CO	6	1.5	0.435
NE	6	1.5	0.391
NO	7	1.5	0.325
SE	10	1.5	0.520
SU	10	1.5	0.410

Tabela 3. Resultado da Clusterização - FEC

Região	Clusters	Coef Fuzzy (m)	gSil
CO	8	1.5	0.400
NE	15	1.5	0.390
NO	8	1.5	0.318
SE	15	1.5	0.445
SU	9	1.5	0.417

7. SIMULAÇÃO PARA DEFINIÇÃO DE LIMITES COLETIVOS PARA A EQUATORIAL PARÁ

Esta seção tem como objetivo avaliar os resultados da metodologia proposta a partir de sua aplicação para a definição dos limites coletivos para a concessionária Equatorial Pará. Esta distribuidora foi selecionada, pois quando da elaboração deste artigo, a companhia estava passando pelo seu quinto ciclo revisional de tarifas e de limites de indicadores de continuidade. Desta forma, seriam definidos novos limites para os indicadores de qualidade do serviço para os próximos 4 anos (2020, 2021, 2022 e 2023) ANEEL (2019). Assim, seria possível a comparação dos resultados oriundos da aplicação da metodologia aqui proposta com os resultados obtidos pela ANEEL.

Vale ressaltar que, assim como a Equatorial Pará, uma distribuidora do grupo CPFL Energia também estava passando por processo de revisão tarifária no período de desenvolvimento deste estudo, podendo então ser selecionada. Contudo, optou-se por selecionar a Equatorial Pará devido aos distintos e discrepantes indicadores apurados por conjunto, os quais verificam patamares de apuração de 30 horas, enquanto outros de 150 horas. Desta forma, o elevado grau de complexidade da concessionária contribuiria para melhores testes de eficácia da metodologia proposta.

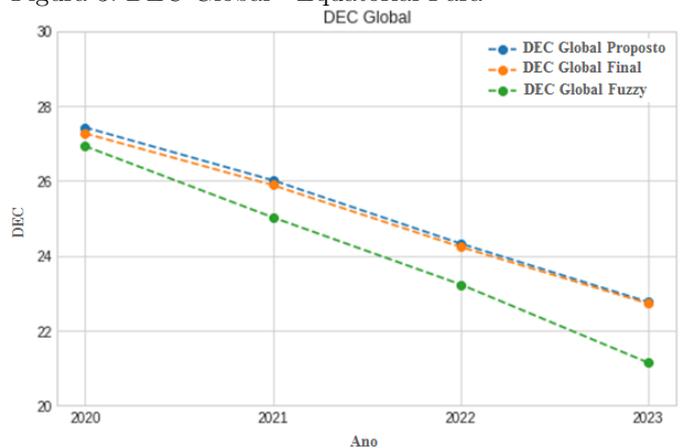
Em linhas gerais, a Equatorial Pará possui um total de 144 conjuntos, sendo 38 classificados como isolados. Segundo ANEEL (2019), a metodologia vigente indica que a maior parte dos conjuntos desta concessionária é classificado como heterogêneo, ou seja, quando comparado aos seus semelhantes, os conjuntos da Equatorial Pará verificam características mais complexas.

Após a aplicação da metodologia proposta, com a realização da clusterização e da definição dos limites coletivos para os conjuntos da distribuidora, alguns resultados são apresentados. A Tabela 4 demonstra as trajetórias de limite para o indicador DEC para alguns conjuntos interligados e que verificam limites moderados (até 30 horas). Já a Tabela 5 apresenta as trajetórias de DEC para alguns dos conjuntos isolados e verificam altos limites (acima de 50 horas).

Por meio dos resultados apresentados nas Tabelas anteriores é notável que a metodologia proposta neste artigo obtém valores próximos dos calculados pelo Regulador, porém com trajetórias de decréscimos um pouco mais severas, principalmente para conjuntos com limites iniciais mais elevados. Este decaimento acelerado deve-se ao fato de que do alvo, por conjunto, é inferior ao obtido pela metodologia vigente.

A Figura 3 apresenta a trajetória global de limites para a Equatorial Pará. Ressalta-se que esta trajetória foi obtida a partir da média ponderada dos limites de cada conjuntos em relação ao número de unidades consumidoras de cada conjunto. É evidente o comportamento de decaimento mais acelerado da trajetória de limites obtida pela metodologia proposta neste artigo.

Figura 3. DEC Global - Equatorial Pará



A linha indicada por "Valor Final" representa os valores corrigidos segundo considerações feitas pela ANEEL e pela própria Equatorial Pará. Em geral, são consideradas flexibilizações de limites para conjuntos com características muito extremas e com difícil acesso, quando a necessidade de deslocamento para o atendimento de uma ocorrência emergencial é dificultado e acarreta na elevação da duração da interrupção e conseqüente majoração do indicador apurado DEC.

Tabela 4. Trajetória DEC - Conjuntos Interligados Moderados

Conjunto	2019		2020		2021		2022		2023	
	Atual	Fuzzy	Atual	Fuzzy	Atual	Fuzzy	Atual	Fuzzy	Atual	
Ourém I	26	25	25	23	24	22	22	20	21	
Outeiro	19	18	18	17	17	16	16	15	16	
Abaetetuba I	30	28	28	26	27	24	25	22	23	
Mãe do Rio	26	25	25	23	24	22	23	20	22	
Coqueiro	17	17	16	16	16	15	15	14	14	

Tabela 5. Trajetória DEC - Conjuntos Isolados com Limites Altos

Conjunto	2019		2020		2021		2022		2023	
	Atual	Fuzzy	Atual	Fuzzy	Atual	Fuzzy	Atual	Fuzzy	Atual	
Santana do Araguaia II	108	99	101	89	94	79	87	69	81	
São Sebastião da Boa Vista	85	78	82	71	79	63	75	56	72	
Muaná	58	53	58	48	54	43	48	38	43	
Óbidos I	59	54	53	48	46	43	40	37	34	
Óbidos II	121	109	114	97	107	85	99	72	92	

8. CONCLUSÃO

Este trabalho apresentou uma proposta alternativa para a definição de limites coletivos para os indicadores coletivos de continuidade do serviço a partir de clusterização fuzzy otimizada. Os parâmetros para a clusterização foram obtidos com sucesso devido à utilização do algoritmo de otimização híbrido e do coeficiente de silhueta generalizada.

A definição dos limites de DEC/FEC, incorporando os graus de pertinência, permitiu auferir mais flexibilidade na avaliação de cada conjunto. No entanto, notou-se que as trajetórias de redução, em relação à metodologia vigente, foram mais acentuadas, ou seja, com reduções mais aceleradas. Tal fato pode ser explicado pela escolha dos percentuais 20% e 50% para o cálculo de *benchmark* dos clusters. Caso fosse selecionada uma faixa maior de valores, ordenados em ordem crescente, seria esperado que o valor do *benchmark* fosse maior, resultando em uma trajetória de limite mais flexível.

Estudos mais aprofundados devem ser feitos para a aplicação dessa nova metodologia. Por exemplo, a escolha de outros percentuais para os *benchmarks*, baseando-se na complexidade absoluta ou relativa do conjunto em relação aos demais comparáveis de seu agrupamento. Outra abordagem interessante, seria a incorporação dos indicadores apurados anteriores de DEC/FEC como variável de entrada para o modelo de clusterização fuzzy.

Por fim, é importante que outras variáveis explicativas de diferentes naturezas (socioeconômicas, geográficas, mobilidade urbana) sejam estudadas. Desta forma, os conjuntos poderão ser caracterizados de maneira mais fidedigna à sua realidade. É importante que se possa mapear problemas como criminalidade e violência pública, dificuldade de acesso devido a más condições de conservação de estradas e rodovias e impacto da confiabilidade da rede básica de suprimento (setor de transmissão).

REFERÊNCIAS

ANEEL (2014a). Nota técnica 0059/2014-SRD/ANEEL.
 ANEEL (2014b). Nota técnica 0102/2014-SRD/ANEEL.
 ANEEL (2018). PRODIST - módulo 8 – qualidade da energia elétrica.

ANEEL (2019). Nota técnica 0060/2019-SRD/ANEEL.
 Bezdek, J.C. (1981). *Pattern Recognition with Fuzzy Objective Function Algorithms*. Advanced Applications in Pattern Recognition. Springer US. doi:10.1007/978-1-4757-0450-1. URL <https://www.springer.com/gp/book/9781475704525>.
 Dias, E.B. (2002). Avaliação de indicadores de continuidade e seu impacto no planejamento de sistemas de distribuição. doi:10.11606/D.3.2002.tde-12062003-173042. URL <http://www.teses.usp.br/teses/disponiveis/3/3143/tde-12062003-173042/>.
 DNAEE (1978). *PORTARIA Nº 46, DE 17 DE ABRIL DE 1978*.
 Dunn, J.C. (1973). A fuzzy relative of the ISODATA process and its use in detecting compact well-separated clusters. doi:10.1080/01969727308546046.
 Kennedy, J. and Eberhart, R. (1995). Particle swarm optimization. In *Proceedings of ICNN'95 - International Conference on Neural Networks*, volume 4, 1942–1948 vol.4. doi:10.1109/ICNN.1995.488968. ISSN: null.
 Pessanha, José Francisco Moreira e Souza, R.C.e.L.L.d.C. (2007). Um modelo de análise envoltória de dados para o estabelecimento de metas de continuidade do fornecimento de energia elétrica. 27(1), 51–83. doi:10.1590/S0101-74382007000100004. URL http://www.scielo.br/scielo.php?script=sci_abstract&pid=S0101-74382007000100004&lng=en&nrm=iso&tlng=pt.
 Rawashdeh, M. and Ralescu, A. (2012). Crisp and fuzzy cluster validity: Generalized intra-inter silhouette index. In *2012 Annual Meeting of the North American Fuzzy Information Processing Society (NAFIPS)*, 1–6. doi:10.1109/NAFIPS.2012.6290969. ISSN: null.
 Rousseeuw, P.J. (1987). Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. 20, 53–65. doi:10.1016/0377-0427(87)90125-7. URL <http://www.sciencedirect.com/science/article/pii/0377042787901257>.
 Sperandio, M. (2004). Classificação de conjuntos consumidores de energia elétrica via mapas auto-organizáveis e estatística multivariada. doi:10.11606/D.3.2002.tde-12062003-173042. URL <https://repositorio.ufsc.br/xmlui/handle/123456789/87238>.