

UNIVERSIDADE FEDERAL DE SANTA MARIA
CAMPUS CACHOEIRA DO SUL
CURSO DE GRADUAÇÃO EM ENGENHARIA ELÉTRICA

Leonardo da Silveira

**ANÁLISE DE MÉTODOS DE AGRUPAMENTO DE CONJUNTOS DE
UNIDADES CONSUMIDORAS APLICADOS NA DEFINIÇÃO DE LIMITES
DOS INDICADORES DE CONTINUIDADE**

Cachoeira do Sul, RS
2021

Leonardo da Silveira

**ANÁLISE DE MÉTODOS DE AGRUPAMENTO DE CONJUNTOS DE UNIDADES
CONSUMIDORAS APLICADOS NA DEFINIÇÃO DE LIMITES DOS INDICADORES DE
CONTINUIDADE**

Trabalho de Conclusão de Curso apresentado ao Curso de Graduação em Engenharia Elétrica da Universidade Federal de Santa Maria (UFSM, RS), como requisito parcial para obtenção do grau de **Bacharel em Engenharia Elétrica**.

ORIENTADORA: Prof.^a Laura Lisiane Callai dos Santos

Cachoeira do Sul, RS
2021

Leonardo da Silveira

**ANÁLISE DE MÉTODOS DE AGRUPAMENTO DE CONJUNTOS DE UNIDADES
CONSUMIDORAS APLICADOS NA DEFINIÇÃO DE LIMITES DOS
INDICADORES DE CONTINUIDADE**

Trabalho de Conclusão de Curso apresentado ao Curso de Graduação em Engenharia Elétrica da Universidade Federal de Santa Maria (UFSM, RS), como requisito parcial para obtenção do grau de **Bacharel em Engenharia Elétrica**.

Aprovado em 25 de agosto de 2021:

Laura Lisiane Callai dos Santos, Dra. (UFSM)
(Presidente/Orientadora)

Maicon Jaderson Silveira Ramos, Dr. (UFRGS)

Nelson Knak Neto, Dr. (UFSM)

Cachoeira do Sul, RS
2021

NUP: 23081.068857/2021-11

Prioridade: Normal

Homologação de Ata

010 - Organização e Funcionamento

COMPONENTE

Ordem	Descrição	Nome do arquivo
1	Ata de defesa de trabalho de conclusão de curso (TCC) (125.322)	Aprovação_Leonardo_TCC2.pdf

Assinaturas

25/08/2021 13:50:28

LAURA LISIANE CALLAI DOS SANTOS (PROFESSOR DO MAGISTÉRIO SUPERIOR)

31.04.00.00.0.0 - COORDENAÇÃO ACADÊMICA - UFSM-CS - C_ACA_CampusCS

26/08/2021 10:47:55

NELSON KNAK NETO (PROFESSOR DO MAGISTÉRIO SUPERIOR)

31.04.00.00.0.0 - COORDENAÇÃO ACADÊMICA - UFSM-CS - C_ACA_CampusCS

27/08/2021 12:07:36

MAICON JADERSON SILVEIRA RAMOS (Pessoa Física)

Usuário Externo (001.***.***.**))



Código Verificador: 808390

Código CRC: be846568

Consulte em: <https://portal.ufsm.br/documentos/publico/autenticacao/assinaturas.html>



AGRADECIMENTOS

A realização desse trabalho foi concretizada com o auxílio, compreensão e dedicação de diversas pessoas. Meu agradecimento a todos que, de alguma forma, contribuíram para a conclusão deste estudo e, de uma maneira especial, agradeço:

Aos meus pais Catarina e Marcos pelo carinho, pela compreensão e pelo apoio que me deram, pois foram fundamentais para o meu sucesso nesta jornada;

À professora Laura Lisiane Callai dos Santos, minha orientadora durante a elaboração deste trabalho, pela colaboração e pela confiança em mim depositada;

Aos demais professores do curso de Engenharia Elétrica da Universidade Federal de Santa Maria - Campus Cachoeira do Sul, que contribuíram para a minha minha formação;

Às grandes e espero que duradouras amizades que construí ao longo dos últimos cinco anos, por terem trazido alegria aos meus dias e me apoiado durante os momentos mais desafiantes;

A todos os meus colegas que me ajudaram no decorrer desses anos, pela compreensão, amizade e parceria que tivemos.

RESUMO

ANÁLISE DE MÉTODOS DE AGRUPAMENTO DE CONJUNTOS DE UNIDADES CONSUMIDORAS APLICADOS NA DEFINIÇÃO DE LIMITES DOS INDICADORES DE CONTINUIDADE

AUTOR: Leonardo da Silveira

ORIENTADORA: Laura Lisiane Callai dos Santos

No contexto da distribuição de energia elétrica no Brasil, cabe a Agência Nacional de Energia Elétrica (ANEEL) regular e fiscalizar a atuação das empresas. A qualidade dos serviços prestados pelas distribuidoras é mensurada por indicadores de continuidade do fornecimento, aos quais são definidos limites de desempenho por meio de uma metodologia de análise comparativa. A premissa dessa metodologia é a de que os conjuntos de unidades consumidoras semelhantes devem possuir desempenhos equivalentes. Tal semelhança entre conjuntos é definida por atributos, que descrevem características do sistema elétrico de distribuição e do mercado consumidor. O descumprimento dos limites regulatórios implica o pagamento de compensações financeiras aos consumidores, que podem representar parte expressiva do faturamento da empresa. Dessarte, a metodologia de cálculo dos limites pode ter influência sobre os resultados financeiros das distribuidoras. Este trabalho analisa a etapa de agrupamento de conjuntos do modelo atual de cálculo de limites dos indicadores DEC e FEC empregado pela ANEEL. Calculam-se os limites para diferentes estudos de caso, em que são variados os atributos selecionados, o método de agrupamento e os parâmetros adotados. Então, é analisada a sensibilidade das variações por meio dos limites calculados em cada estudo de caso. Como principais resultados, observou-se uma maior sensibilidade para variação dos atributos e do método de *clustering* e uma menor sensibilidade para variação dos parâmetros do método dinâmico.

Palavras-chave: Qualidade do serviço. Continuidade. Indicadores. Agrupamento. Limites.

ABSTRACT

ANALYSIS OF GROUPING METHODS FOR SETS OF CONSUMER UNITS APPLIED IN DEFINING LIMITS FOR CONTINUITY INDICATORS

AUTHOR: Leonardo da Silveira

ADVISOR: Laura Lisiane Callai dos Santos

In the context of electricity distribution in Brazil, the National Electric Energy Agency (ANEEL) is responsible for regulating and inspecting the performance of companies. The quality of services provided by distributors is measured by indicators of continuity of supply, to which performance limits are defined using a comparative analysis methodology. The premise of this methodology is that similar sets of consumer units must have equivalent performance. Such similarity between sets is defined by attributes, which describe characteristics of the electric distribution system and the consumer market. Failure to comply with regulatory limits implies the payment of financial compensation to consumers, which can represent a significant part of the company's revenue. Thus, the methodology for calculating the limits can influence the financial results of distributors. This work analyzes the step of grouping sets of the current model used by ANEEL for calculating the limits of the DEC and FEC indicators. The limits for different case studies are calculated, in which the selected attributes, the grouping method and the adopted parameters are varied. Then, the sensitivity of variations is analyzed through the limits calculated in each case study. As main results, it was observed a greater sensitivity for variation of attributes and of the clustering method and a lower sensitivity for variation of the parameters of the dynamic method.

Keywords: Service quality. Continuity. Indicators. Clustering. Limits.

LISTA DE FIGURAS

Figura 2.1 – Mapeamento f entre entradas e saída de um sistema.	18
Figura 2.2 – Divisão dos três <i>clusters</i> obtidos com o método de Ward.	20
Figura 2.3 – Exemplo de <i>clusterização</i>	34
Figura 2.4 – Homogeneidade <i>intracluster</i>	34
Figura 2.5 – Formas de medição da distância entre <i>clusters</i>	36
Figura 2.6 – Representação dos conjuntos no plano cartesiano.	38
Figura 2.7 – Dendograma considerando o método de Ward.	40
Figura 2.8 – Fluxograma do método <i>k-means</i>	42
Figura 2.9 – Diagrama de Voronoi da <i>clusterização</i> pelo método <i>k-means</i>	43
Figura 2.10 – Exemplo da normalização de variáveis.	47
Figura 3.1 – Estrutura da metodologia proposta de agrupamento de conjuntos.	50
Figura 4.1 – Áreas dos conjuntos da distribuidora CEEE-D em 2013.	53
Figura 4.2 – DEC Apurado e DEC Limite Objetivo por conjunto.	57
Figura 4.3 – Histograma de variação dos limites objetivos do Cenário 2.	58
Figura 4.4 – Histograma de variação dos limites objetivos do Cenário 3.	59
Figura 4.5 – Histograma de variação dos limites objetivos do Cenário 4.	61
Figura 4.6 – Histograma de variação dos limites objetivos do Cenário 5.	62
Figura 4.7 – Histograma de variação dos limites objetivos do Cenário 6.	64
Figura 4.8 – Histograma de variação dos limites objetivos do Cenário 7.	65
Figura 4.9 – Comparação dos limites objetivos globais.	66

LISTA DE QUADROS

Quadro 2.1 – Tipos de atributos.	25
Quadro 2.2 – Atributos selecionados para a metodologia comparativa de DEC e FEC.	26
Quadro 2.3 – Parâmetros do método dinâmico.	27
Quadro 2.4 – Equações utilizadas para cálculo do ANI dos atributos selecionados para a metodologia comparativa.	30
Quadro 2.5 – Variação nos percentis a serem utilizados para conjuntos heterogêneos com base no <i>Score ANI</i>	30
Quadro 2.6 – Atributos dos conjuntos a serem analisados.	37
Quadro 2.7 – Matriz de similaridade usando o critério do menor erro.	38
Quadro 2.8 – Matriz de similaridade após o primeiro agrupamento.	39
Quadro 2.9 – Elementos agrupados e níveis.	39
Quadro 2.10 – Distância dos conjuntos aos centroides iniciais.	42
Quadro 2.11 – Distância dos conjuntos aos centroides na segunda iteração.	43
Quadro 2.12 – Matriz de similaridade usando a distância euclidiana.	45
Quadro 2.13 – Tabelas de similaridade para os conjuntos.	46
Quadro 4.1 – Indicadores DEC e FEC médios apurados nos anos de 2011 a 2013. ..	54
Quadro 4.2 – Parâmetros do Cenário 1.	55
Quadro 4.3 – Conjuntos semelhantes ao conjunto em análise.	55
Quadro 4.4 – Conjuntos semelhantes ordenados pelo DEC apurado.	56
Quadro 4.5 – Parâmetros do Cenário 2.	57
Quadro 4.6 – Parâmetros do Cenário 3.	59
Quadro 4.7 – Parâmetros do Cenário 4.	60
Quadro 4.8 – Parâmetros do Cenário 5.	61
Quadro 4.9 – Parâmetros do Cenário 6.	63
Quadro 4.10 – Parâmetros do Cenário 7.	64
Quadro 4.11 – Resumo dos cenários.	67

LISTA DE ABREVIATURAS E SIGLAS

ANEEL	Agência Nacional de Energia Elétrica
BDGD	Base de Dados Geográfica da Distribuidora
CEEE-D	Companhia Estadual de Distribuição de Energia Elétrica
CRTP	Ciclo de Revisões Tarifárias Periódicas
DEC	Duração Equivalente de Interrupção por Unidade Consumidora
DIC	Duração de Interrupção Individual por Unidade Consumidora
DICRI	Duração da Interrupção Individual Ocorrida em Dia Crítico por Unidade Consumidora
DNAEE	Departamento Nacional de Águas e Energia Elétrica
DMIC	Duração Máxima de Interrupção Contínua por Unidade Consumidora
FEC	Frequência Equivalente de Interrupção por Unidade Consumidora
FIC	Frequência de Interrupção Individual por Unidade Consumidora
NUC	Número de Unidades Consumidoras
P&D	Pesquisa e Desenvolvimento
PRODIST	Procedimentos de Distribuição de Energia Elétrica no Sistema Elétrico Nacional
SIG-R	Sistema de Informação Geográfica Regulatório
SIN	Sistema Interligado Nacional
UC	Unidade Consumidora

SUMÁRIO

1	INTRODUÇÃO	11
1.1	CONTEXTUALIZAÇÃO	11
1.2	MOTIVAÇÕES E JUSTIFICATIVA	12
1.3	OBJETIVOS	13
1.3.1	Objetivo Geral	13
1.3.2	Objetivos Específicos	13
1.4	ORGANIZAÇÃO DO TRABALHO	13
2	REVISÃO BIBLIOGRÁFICA	15
2.1	INDICADORES DE CONTINUIDADE DE ENERGIA ELÉTRICA	15
2.2	DEFINIÇÃO DOS LIMITES DE DEC E FEC	17
2.2.1	Modelo Empregado a Partir de 2000	17
2.2.2	Modelo Proposto na Consulta Pública nº 008/2010	19
2.2.3	Modelo Empregado a Partir de 2011	21
2.2.4	Modelo Vigente da Definição de Limites	23
2.2.4.1	<i>Seleção e extração dos atributos</i>	24
2.2.4.2	<i>Aplicação do método dinâmico</i>	26
2.2.4.3	<i>Avaliação dos conjuntos heterogêneos</i>	29
2.2.4.4	<i>Avaliação do suprimento e das trajetórias de redução intensas</i>	31
2.3	MÉTODOS DE CLUSTERIZAÇÃO	32
2.3.1	Métodos Hierárquicos	35
2.3.1.1	<i>Método de Ward</i>	37
2.3.2	Método <i>k-means</i>	40
2.3.3	Método Dinâmico	44
2.3.4	Normalização dos Dados	47
2.4	CONSIDERAÇÕES SOBRE O CAPÍTULO	48
3	METODOLOGIA PROPOSTA	49
3.1	DADOS DE ENTRADA	49
3.2	AGRUPAMENTO DOS CONJUNTOS	50
3.3	DEFINIÇÃO DOS LIMITES	52
3.4	CONSIDERAÇÕES SOBRE O CAPÍTULO	52
4	ESTUDOS DE CASO E RESULTADOS	53
4.1	CENÁRIO 1	55
4.2	CENÁRIO 2	57
4.3	CENÁRIO 3	58
4.4	CENÁRIO 4	60
4.5	CENÁRIO 5	61
4.6	CENÁRIO 6	63
4.7	CENÁRIO 7	64
4.8	COMPARAÇÃO DOS RESULTADOS	66
4.9	CONSIDERAÇÕES SOBRE O CAPÍTULO	68
5	CONCLUSÃO	69
5.1	OPORTUNIDADES DE TRABALHOS FUTUROS	70
5.2	PRODUÇÃO CIENTÍFICA	70
5.2.1	Artigos publicados em anais de eventos	71
	REFERÊNCIAS BIBLIOGRÁFICAS	72

APÊNDICE A – LIMITES OBJETIVOS DE DEC E FEC POR CONJUNTO 75

1 INTRODUÇÃO

1.1 CONTEXTUALIZAÇÃO

A crescente demanda por energia elétrica, a partir da década de 1990, tem exigido do Setor Elétrico Brasileiro uma constante evolução tecnológica e regulatória, objetivando-se melhorar a qualidade da energia elétrica entregue aos consumidores. A criação da Agência Nacional de Energia Elétrica (ANEEL), no ano de 1996, marca um avanço para o mercado de energia, que passou a ser regulado e fiscalizado com maior rigor (BRASIL, 1996; BRASIL, 1997). Além disso, as desestatizações e a desverticalização do setor foram mudanças que contribuíram positivamente para essa evolução.

A modernização tecnológica corrobora com uma operação dos sistemas elétricos melhor planejada e mais eficiente. As distribuidoras utilizam um volume cada vez maior de informações, que, quando bem analisadas, permitem um gerenciamento dos recursos de forma otimizada. Tais informações também podem ser utilizadas pelo Órgão Regulador com a finalidade de criar um ambiente de concorrência em um setor que constitui um monopólio natural.

A qualidade do serviço prestado pelas empresas de distribuição é medida por indicadores de continuidade individuais e coletivos. A ANEEL impõe limites para esses indicadores e, caso não sejam cumpridos, cabe às empresas compensar financeiramente os clientes prejudicados.

As informações geradas pelas distribuidoras e repassadas à ANEEL permitiram, no ano de 2000, a criação de um modelo de definição de limites para os indicadores coletivos: Duração Equivalente de Interrupção por Unidade Consumidora (DEC) e Frequência Equivalente de Interrupção por Unidade Consumidora (FEC). O modelo baseava-se em uma análise comparativa de desempenho: agrupam-se conjuntos semelhantes de unidades consumidoras em função de suas características (atributos) e, então, impõem-se limites de desempenho aos conjuntos.

Os agrupamentos são realizados por técnicas de análise de *cluster*. A análise de agrupamento de dados, também chamada de *clustering*, trata-se de um conjunto de métodos que visam agregar dados em função do grau de semelhança entre eles. No campo do aprendizado de máquina (*machine learning*), a *clusterização* é classificada como uma das abordagens de aprendizado não supervisionado.

Posteriormente a metodologia de definição dos limites passou por atualizações. Após quase dez anos, em 2009, propôs-se a utilização de novos atributos, capazes de representar os conjuntos com maior fidelidade, além da mudança na forma como os conjuntos são agrupados.

A partir de 2011, as distribuidoras de energia passaram a ter de enviar à ANEEL,

uma vez por ano, a Base de Dados Geográfica da Distribuidora (BDGD). A BDGD contém, de forma padronizada, informações relativas aos consumidores e aos ativos das empresas. Essas informações propiciaram a adoção de uma nova metodologia de cálculo de limites por parte da Agência Reguladora, em 2014. O volume maior de dados disponíveis tornou possível uma avaliação de quais atributos eram mais relevantes para a análise comparativa.

Contudo considerando-se o modo como são calculados os limites atualmente, uma parte expressiva dos conjuntos não é capaz de cumpri-los. Assim, é relevante que as concessionárias tenham meios e argumentos para questionar a forma como são definidos os limites e propor mudanças à ANEEL.

1.2 MOTIVAÇÕES E JUSTIFICATIVA

Com o não cumprimento dos limites individuais impostos, as empresas distribuidoras têm grandes despesas com compensações aos clientes. No Rio Grande do Sul, a Companhia Estadual de Distribuição de Energia Elétrica (CEEE-D) pagou, no ano de 2020, mais de R\$ 27 milhões em compensações, o que representou aproximadamente 8% da receita líquida da empresa (ANEEL, 2020; Grupo CEEE, 2020). Assim sendo, a metodologia de definição dos limites de continuidade pode ter relevância expressiva sobre os resultados financeiros das distribuidoras de energia.

Ademais, as constantes alterações da metodologia de cálculo dos limites, ainda que embasadas por diversos estudos, indicam a inexistência de um modelo perfeito e completamente justo às distribuidoras de energia. Nas últimas duas décadas, foram empregados três métodos de *clusterização* diferentes: *k-means*, Ward e dinâmico, contudo não foram encontrados na literatura estudos que investigassem as consequências das alterações nos métodos de agrupamento. Ou seja, desconhece-se qual o impacto causado pela utilização de um novo método em detrimento dos anteriores.

Dessa forma deseja-se avaliar quais as vantagens e desvantagens do método de *clusterização* utilizado na metodologia atual (dinâmico) em face de outros métodos (*k-means* e Ward). Ainda, pretende-se verificar quais são os efeitos do uso de diferentes métodos de *clusterização* na definição dos limites.

1.3 OBJETIVOS

1.3.1 Objetivo Geral

Como objetivo deste trabalho, propõe-se realizar uma análise comparativa de diferentes métodos de agrupamento de conjuntos de unidades consumidoras aplicados ao estabelecimento de limites dos indicadores coletivos de continuidade de energia elétrica.

1.3.2 Objetivos Específicos

A fim de se atingir o objetivo geral, definem-se os seguintes objetivos específicos:

- Realizar uma revisão sobre os indicadores de continuidade coletivos e as metodologias de cálculo de seus limites;
- Revisar os principais métodos de *clusterização* aplicáveis à definição de limites de continuidade;
- Aplicar os métodos de agrupamento de conjuntos a diferentes estudos de caso e calcular os limites e
- Avaliar e comparar o comportamento de diferentes técnicas de análise de *cluster* na definição dos limites dos indicadores.

1.4 ORGANIZAÇÃO DO TRABALHO

Para se alcançar os objetivos propostos, este trabalho é composto por cinco capítulos, dentre os quais está este capítulo introdutório.

No Capítulo 2 apresenta-se uma revisão bibliográfica. Contextualizam-se os indicadores de continuidade de energia elétrica e apresentam-se as metodologias passadas e atual utilizadas pela ANEEL para definição dos limites desses indicadores. São descritos, ainda, três métodos estatísticos de *clusterização* (dinâmico, *k-means* e Ward) empregados no cálculo desses limites.

A metodologia deste trabalho, exposta no Capítulo 3, propõe o emprego de diferentes métodos de agrupamento aplicados à definição dos limites de DEC e FEC e, para isso, realizam-se três etapas. Inicialmente são obtidos os dados de entrada, ou seja, as bases de dados de atributos e indicadores referentes aos conjuntos de unidades consumidoras. Em seguida, realiza-se o agrupamento dos conjuntos com base em suas semelhanças, por meio de técnicas de *clusterização*. Por fim, comparam-se os resultados obtidos e os métodos empregados.

Aplica-se a metodologia proposta para o cálculo dos limites dos conjuntos da CEEE-D em diferentes cenários, os quais são apresentados no Capítulo 4. O primeiro cenário proposto visa replicar o cálculo dos limites da forma como é feito atualmente pela ANEEL. Nos demais estudos de caso, realizam-se alterações no método e calculam-se os limites. Os resultados dos demais casos são comparados aos do cenário base.

O Capítulo 5 traz uma conclusão sobre os resultados obtidos.

2 REVISÃO BIBLIOGRÁFICA

Neste capítulo é apresentada uma revisão bibliográfica. Inicialmente discorre-se sobre o contexto dos indicadores de continuidade de energia elétrica no Brasil. Em seguida apresenta-se a forma como eram e como são definidos os limites de continuidade coletivos. As metodologias de cálculo de limites utilizam-se da análise de *cluster*, assunto abordado nesta revisão. Por fim realiza-se uma breve discussão sobre o capítulo.

2.1 INDICADORES DE CONTINUIDADE DE ENERGIA ELÉTRICA

No Brasil o segmento de distribuição de energia elétrica é regulado e fiscalizado pela ANEEL. Cabe à Agência propiciar as condições favoráveis ao desenvolvimento equilibrado do mercado brasileiro de energia, por meio da regulamentação dos serviços prestados pelas distribuidoras.

O artigo 3º do Anexo I do Decreto nº 2.335, de 06 outubro de 1997, estabelece, nos seguintes incisos, algumas das diretrizes a serem observadas pela ANEEL na execução de suas atividades finalísticas:

- II - regulação e fiscalização realizadas com o caráter de simplicidade e pautadas na livre concorrência entre os agentes, no atendimento às necessidades dos consumidores e no pleno acesso aos serviços de energia elétrica;
- III - adoção de critérios que evitem práticas anticompetitivas e de impedimento ao livre acesso aos sistemas elétricos;
- V - criação de ambiente para o setor de energia elétrica que incentive o investimento, de forma que os concessionários, permissionários e autorizados tenham asseguradas a viabilidade econômica e financeira, nos termos do respectivo contrato (BRASIL, 1997).

Dentre os mecanismos de regulação elaborados pela ANEEL, estão os Procedimentos de Distribuição de Energia Elétrica no Sistema Elétrico Nacional (PRODIST). Constituem o PRODIST onze módulos, que normatizam e padronizam as atividades técnicas relacionadas ao funcionamento e desempenho dos sistemas de distribuição de energia elétrica.

No Módulo 8 do PRODIST, são abordados, em três seções, os procedimentos relacionados à qualidade do produto, à qualidade do serviço prestado e à qualidade do tratamento de reclamações. No que tange à qualidade do serviço, na Seção 8.2 - Qualidade do Serviço, são definidos indicadores de continuidade do fornecimento de energia elétrica individuais e coletivos. A ANEEL utiliza os indicadores coletivos para analisar o desem-

penho dos conjuntos de unidades consumidoras, além do desempenho das distribuidoras, estados, regiões e do país (ANEEL, 2021b).

Os indicadores individuais são apurados por unidade consumidora (UC) ou ponto de conexão e são os seguintes:

- DIC - duração de interrupção individual por unidade consumidora;
- FIC - frequência de interrupção individual por unidade consumidora;
- DMIC - duração máxima de interrupção contínua por unidade consumidora e
- DICRI - duração da interrupção individual ocorrida em dia crítico por unidade consumidora.

As Equações 2.1 e 2.2 expressam a formulação para apuração dos indicadores DIC e FIC respectivamente.

$$DIC = \sum_{i=1}^n t(i) \quad (2.1)$$

$$FIC = n \quad (2.2)$$

Onde:

i = Índice de interrupções da unidade consumidora, no período de apuração, variando de 1 a n ;

n = Número de interrupções da unidade consumidora considerada, no período de apuração;

$t(i)$ = Tempo de duração da interrupção (i) da unidade consumidora considerada, no período de apuração.

Por sua vez, os indicadores de continuidade coletivos são apurados para cada conjunto de unidades consumidoras e estão discriminados a seguir:

- DEC - duração equivalente de interrupção por unidade consumidora e
- FEC - frequência equivalente de interrupção por unidade consumidora.

As formulações dos indicadores DEC e FEC são expressas pelas Equações 2.3 e 2.4 respectivamente.

$$DEC = \frac{\sum_{i=1}^{C_c} DIC(i)}{C_c} \quad (2.3)$$

$$FEC = \frac{\sum_{i=1}^{C_c} FIC(i)}{C_c} \quad (2.4)$$

Onde:

i = Índice da unidade consumidora, no período de apuração, variando de 1 a C_c ;

C_c = Número total de unidades consumidoras atendidas em BT ou MT faturadas do conjunto.

Além dos indicadores citados, são instituídos seus respectivos limites. Os limites de continuidade individuais são definidos pelas tabelas constantes no Anexo I da Seção 8.2 do Módulo 8 do PRODIST. É através dos limites coletivos (DEC e FEC) que se estabelecem os limites dos indicadores individuais, DIC, FIC e DMIC.

No que diz respeito aos limites dos indicadores de continuidade coletivos, esses são definidos por meio de uma metodologia de análise comparativa, a qual é apresentada na Nota Técnica nº 0102/2014-SRD/ANEEL. Uma revisão das metodologias passadas e atual de definição dos limites de DEC e de FEC é descrita na Seção 2.2.

2.2 DEFINIÇÃO DOS LIMITES DE DEC E FEC

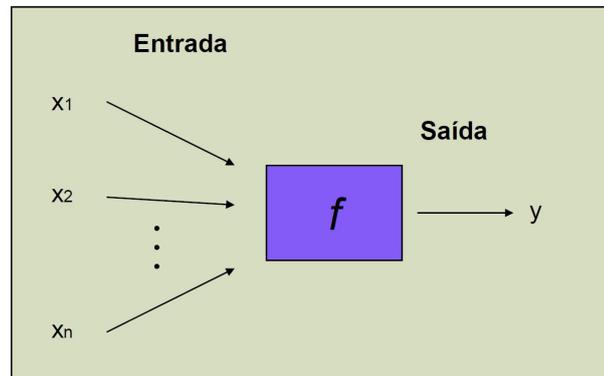
Nesta seção são apresentados um breve histórico dos modelos de definição dos limites de DEC e FEC utilizados pela ANEEL e os principais aspectos da metodologia vigente.

2.2.1 Modelo Empregado a Partir de 2000

A regulamentação da continuidade da energia elétrica no Brasil iniciou-se com a Portaria DNAEE nº 046/78. A partir de 1978, os indicadores coletivos DEC e FEC começaram a ser acompanhados por conjuntos de unidades consumidoras, o que possibilitou a disponibilidade de um histórico de informações. Em 2000 a Resolução ANEEL nº 024/2000 introduziu importantes aperfeiçoamentos sobre o assunto e a apuração dos indicadores no setor de distribuição tornou-se uma prática consolidada.

Definir os limites aos conjuntos, e conseqüentemente à distribuidora, não é uma tarefa simples. Para ilustrar o problema, pode-se analisar o desempenho de um conjunto como um sistema simples, com várias entradas e uma saída. As entradas x_i são as variáveis que, de alguma forma, se relacionam, resultando na saída y , e o mapeamento entre a entrada e a saída é representado por f . Para estabelecimento dos limites dos indicadores coletivos, considera-se as entradas como atributos definidores dos conjuntos, e a saída como o desempenho do indicador. A Figura 2.1 ilustra o exemplo (ANEEL, 2010a).

Figura 2.1 – Mapeamento f entre entradas e saída de um sistema.



Fonte: ANEEL (2010a).

Devido às incertezas associadas aos parâmetros de entrada, a formulação de um modelo funcional para o sistema da Figura 2.1 pode não ser algo simples. É possível citar alguns parâmetros não determinísticos, com probabilidade associada, como eventos climáticos e falhas em equipamentos. Alguns parâmetros são altamente dependentes da gestão da distribuidora, como a eficiência das equipes de atendimento emergencial. Ainda, há complexidade no mapeamento entre a entrada e a saída das variáveis, neste caso, na forma de relacionamento dos parâmetros com a saída do sistema (ANEEL, 2010a).

Considerando a existência de um histórico de valores dos indicadores DEC e FEC, a ANEEL estabeleceu, na Resolução nº 024/2000, uma metodologia comparativa que avalia o desempenho dos conjuntos para propor limites adequados. A premissa básica da metodologia é a seguinte: Conjuntos semelhantes devem apresentar desempenho equivalente (ANEEL, 2010a).

A partir dessa premissa, não se faz necessário conhecer o relacionamento entre as variáveis de entrada e saída. O que se espera é que, uma vez definidos quais conjuntos são semelhantes, os desempenhos desses sejam equivalentes.

O método de comparação de desempenho proposto pela ANEEL é baseado em uma técnica estatística de formação de agrupamentos, denominados *clusters*, que permite identificar conjuntos com características homogêneas. Esta técnica é utilizada quando não é possível estabelecer deterministicamente os valores assumidos pelas variáveis de um determinado problema. Essa metodologia permite superar uma desvantagem natural, decorrente da grande assimetria de informações em relação às distribuidoras, além de permitir a emulação de um ambiente concorrencial que visa a melhoria da qualidade de serviço das empresas (ANEEL, 2010a).

Para definir a semelhança entre os conjuntos, foram definidos cinco atributos físico-elétricos que os caracterizam, os quais são: área (km²), extensão da rede primária (km), potência instalada (kVA), número de unidades consumidoras e consumo médio mensal (MWh). Com base nesses atributos, formam-se agrupamentos de conjuntos. Em cada

cluster os conjuntos possuem características semelhantes, porém os indicadores de desempenhos verificados são diferentes. Em face disso, a metodologia define, para os conjuntos de cada *cluster*, os limites futuros de DEC e FEC, dentro de um determinado intervalo de tempo (ANEEL, 2010a).

A técnica utilizada para se determinar os limites de cada agrupamento, chamada de *Yardstick Competition*, consiste em fixar a referência do *cluster* a partir de um limite estatístico. Para conjuntos atendidos pelo Sistema Interligado Nacional (SIN), os limites correspondem aos valores observados no limite do espaço amostral dos 20% melhores conjuntos daquele *cluster*. No caso de conjuntos atendidos por sistemas isolados, a referência para o estabelecimento dos limites é a mediana do *cluster*, ou seja, o valor observado no espaço amostral de 50% dos conjuntos do referido *cluster* (ANEEL, 2010a).

Por fim, os limites de cada ano são projetados a partir de uma função exponencial decrescente, que partem dos limites previamente estipulados e chegam aos limites estabelecidos para o *cluster* em um período de oito anos. Esse tempo de transição considera o período médio de duas revisões tarifárias. Tanto o período de transição quanto a adoção dos percentis 20 e 50 são flexibilizações para o cumprimento dos limites de forma a garantir a coerência do método (ANEEL, 2010a).

Uma atualização da metodologia descrita foi objeto de discussão a partir do ano de 2009. As novas propostas são apresentadas na próxima seção.

2.2.2 Modelo Proposto na Consulta Pública nº 008/2010

No ano de 2009, a ANEEL abriu a Consulta Pública nº 043/2009 para obter subsídios e informações para o aprimoramento da regulamentação sobre a continuidade do fornecimento de energia elétrica na distribuição. Foram levantados questionamentos acerca da metodologia que estabelece os limites para os indicadores DEC e FEC, como o método de clusterização, a definição de conjuntos de referência nos *clusters* para a fixação dos limites e a escolha dos atributos dos conjuntos. As principais contribuições foram elencadas no anexo à Nota Técnica nº 0064/2009-SRD/ANEEL (ANEEL, 2009).

Tais contribuições embasaram a proposição de uma nova metodologia de definição de limites, apresentada em 2010 no contexto da Consulta Pública nº 008/2010. O documento anexo à Nota Técnica nº 028/2010-SRD/ANEEL reúne as contribuições da comunidade e a formulação do novo modelo (ANEEL, 2010a).

A metodologia proposta baseou-se nos seguintes passos: a) identificação dos atributos que melhor representem as distribuidoras; b) *clusterização* das distribuidoras (*pré-clusterização*); c) análise dos atributos que melhor representem os conjuntos; d) *clusterização* dos conjuntos das distribuidoras em cada *pré-cluster*; e) definição dos limites para os indicadores DEC e FEC (ANEEL, 2010a).

Nota-se que, diferentemente da metodologia anterior, optou-se por realizar duas etapas de seleção de atributos e de *clusterização*. Na primeira etapa, são considerados os dados gerais das distribuidoras com o intuito de considerar atributos que não estão disponíveis por conjunto de unidades consumidoras. Enquanto que na segunda etapa são considerados os dados relativos aos conjuntos (ANEEL, 2010a).

No primeiro passo realizou-se uma análise estatística para se determinar, dentre 28 variáveis previamente escolhidas, quais são os atributos mais relevantes para a pré-*clusterização*. O modelo escolhido foi a análise fatorial. Ao fim da análise, obtiveram-se 12 variáveis (ANEEL, 2010a).

No passo seguinte procedeu-se uma análise de colinearidade entre as variáveis obtidas no primeiro passo. A colinearidade indica a relação entre as variáveis independentes em uma análise de regressão e uma medida que expressa sua existência é o Fator de Inflação de Variância (VIF). Valores altos de VIF indicam alta colinearidade, que é prejudicial em análise de agrupamentos, pois equivale a aplicação de pesos diferenciados entre as características avaliadas (ANEEL, 2010a).

Para o processo de pré-*clusterização*, adotou-se o método de Ward, o qual é apresentado de forma mais detalhada na Seção 2.3.1.1. Nesta etapa as distribuidoras foram classificadas em três *clusters*, apresentados no mapa geográfico do Brasil na Figura 2.2.

Figura 2.2 – Divisão dos três *clusters* obtidos com o método de Ward.



Fonte: ANEEL (2010a).

No terceiro passo, de forma análoga ao primeiro, foi avaliado o poder descritivo de onze variáveis de 3241 conjuntos do Brasil. Após um processo de análise fatorial, foram selecionados seis atributos a serem utilizados na análise comparativa entre os conjuntos de unidades consumidoras (ANEEL, 2010a).

No quarto passo, os conjuntos, previamente classificados em três grupos, são sub-

metidos a uma nova *clusterização*. Para isso, primeiro os dados foram padronizados de forma a terem média zero e desvio padrão unitário. Foram retiradas observações atípicas, denominadas *outliers*, de forma a evitar que subestações atípicas formem *clusters* com quantidades de conjuntos muito reduzidas (ANEEL, 2010a).

Após o tratamento dos dados, aplicou-se o método de clusterização não-hierárquico *k-means*, que é detalhado na Seção 2.3.2. Para a definição do número ideal de *clusters*, adotou-se o método do centroide. Dessa etapa foram obtidos sete *clusters* para o primeiro pré-*cluster*, vinte para o segundo e cinco para o terceiro pré-*cluster* (ANEEL, 2010a).

Por fim, no quinto e último passo, foram definidos os limites de DEC e FEC para cada *cluster*, com base nos percentis 20 e 50 citados. Entretanto optou-se por utilizar uma trajetória linear decrescente até o valor alvo e não mais uma trajetória exponencial. Ainda, propôs-se um decaimento limitado a 15% ao ano, de forma a evitar uma redução brusca dos limites (ANEEL, 2010a).

Os passos aqui apresentados tratam-se de um resumo da proposta apresentada na Nota Técnica nº 028/2010-SRD/ANEEL. Posteriormente a ANEEL realizou a Audiência Pública nº 046/2010 e novas propostas foram feitas. Em especial optou-se pela não adoção das etapas de pré-*clusterização* e de *clusterização*, conforme ANEEL (2010b). A metodologia efetivamente adotada no 3º Ciclo de Revisões Tarifárias Periódicas (3CRTP) é mostrada na Seção 2.2.3.

2.2.3 Modelo Empregado a Partir de 2011

Após as duas Consultas Públicas abertas pela ANEEL para a obtenção de contribuições da sociedade, a Agência Reguladora realizou a Audiência Pública nº 046/2010. Nessa ocasião, foram apresentadas novas propostas de aprimoramento da metodologia de definição dos limites dos indicadores de continuidade coletivos, por meio da Nota Técnica nº 0048/2010-SRD/ANEEL (ANEEL, 2010a).

Nessa revisão da metodologia de análise comparativa, a premissa básica mantém-se a mesma: conjuntos semelhantes devem apresentar desempenhos equivalentes. Assim, objetiva-se formar agrupamentos de conjuntos semelhantes com base em seus atributos.

Durante a Consulta Pública nº 008/2010 foram propostos quinze atributos descritores dos conjuntos. Na verdade, diversas são as variáveis que influenciam na continuidade do fornecimento de energia elétrica. Entretanto algumas premissas devem ser adotadas para a seleção de atributos, dentre as quais estão: possibilidade de definição clara no regulamento, facilidade de obtenção de forma sistemática e viabilidade de fiscalização (ANEEL, 2010b).

Devido à impossibilidade de obtenção de alguns atributos por conjunto, como Ín-

dice Pluviométrico Anual, Densidade de Descarga Atmosférica e Densidade de Estradas Pavimentadas, propôs-se uma etapa de *clusterização* de distribuidoras. Entretanto, devido ao descasamento geográfico e temporal de informações, tais atributos e a etapa de *pré-clusterização* foram descartados da metodologia (ANEEL, 2010b).

Definiram-se, então, sete atributos característicos dos conjuntos (ANEEL, 2010b):

- Área em quilômetros quadrados (km²);
- Extensão da rede de média tensão (MT), segregada em urbana e rural, em quilômetros (km);
- Energia consumida nos últimos 12 meses, segregada pelas classes residencial, industrial, comercial, rural e outras classes, em megawatt-hora (MWh);
- Total de unidades consumidoras atendidas, segregada pelas classes residencial, industrial, comercial, rural e outras classes;
- Potência instalada (kVA);
- Padrão construtivo da rede (aérea ou subterrânea); e
- Localização (sistema isolado ou interligado).

A partir dos atributos físico-elétricos, os conjuntos são agrupados a partir de técnicas estatísticas que buscam dividir os elementos da população em *clusters* a partir da similaridade dos elementos de um dado grupo e da heterogeneidade em relação aos outros grupos. A medida de similaridade utilizada para avaliar a semelhança entre conjuntos em relação aos atributos selecionados é a distância euclidiana. Considerando que essa métrica é sensível a diferentes escalas ou magnitudes entre as variáveis, é necessária uma padronização dos dados. A padronização adotada é o escore Z, obtida pela subtração da média e divisão do desvio-padrão para cada variável (ANEEL, 2010b).

Diferentemente do que era realizado na metodologia anterior, em que se aplicava o método *k-means*, propôs-se a utilização do método de *clusterização* denominado dinâmico. O método dinâmico, abordado em maiores detalhes na Seção 2.3.3, visa comparar cada conjunto com os conjuntos mais semelhantes a ele. A principal diferença do método dinâmico é a de não haver a formação de agrupamentos fixos, ou seja, para cada conjunto é formado um *cluster* e realizada uma comparação independente (ANEEL, 2014a).

Para uma aplicação generalizada do método dinâmico é preciso estabelecer alguns critérios como (ANEEL, 2010b):

- Grau de heterogeneidade percentual permitida (refere-se à média das distâncias euclidianas dos conjuntos ao conjunto de referência);
- Número desejável de conjuntos que devem ser selecionados como semelhantes para aplicação da análise comparativa; e
- Número mínimo de conjuntos que devem ser selecionados como semelhantes para aplicação da análise comparativa.

Adota-se um valor máximo de heterogeneidade permitido para se limitar o número

de conjuntos semelhantes selecionados para determinado conjunto. Dessa forma conjuntos menos semelhantes são desconsiderados da análise. A heterogeneidade percentual é definida pela Equação 2.5.

$$Heterogeneidade_i^P = \frac{\sum_{j=1}^n Dist_i^j}{n \cdot 3\sqrt{k}} \quad (2.5)$$

Em que:

i = Índice do conjunto de referência;

j = Conjuntos próximos ao conjunto i ;

$Dist_i^j$ = Distância euclidiana do conjunto i para o conjunto j ;

n = Número de conjuntos semelhantes ao conjunto i ;

k = Número de atributos que caracterizam o conjunto i ;

Durante a aplicação do método dinâmico, além de se diferenciar os conjuntos localizados em sistemas isolados e interligados, também são diferenciados os conjuntos por padrão construtivo da rede entre aéreo e subterrâneo (ANEEL, 2010b).

Uma vez que os conjuntos estejam agrupados conforme suas características, são definidos os limites de DEC e FEC utilizando-se o critério dos percentis. Como forma de manter a melhoria contínua no modelo, o limite de um ano não pode ser superior ao limite estabelecido para o ano anterior. Assim os conjuntos com desempenho melhor do que a referência terão os limites mantidos no modelo (ANEEL, 2010b).

A última etapa da análise comparativa consiste na análise da consistência dos resultados. Os valores dos limites calculados são disponibilizados por meio de Audiência Pública. Após a análise das contribuições da Audiência Pública, os limites finais de DEC e FEC são estabelecidos (ANEEL, 2010b).

No ano de 2014 houve uma nova revisão da metodologia de análise comparativa, realizada durante a Audiência Pública nº 029/2014. A metodologia, que está em vigor desde 2015, é descrita na Seção 2.2.4.

2.2.4 Modelo Vigente da Definição de Limites

Em 2009 a ANEEL publicou a Resolução Normativa nº 395/2009, criando o Sistema de Informação Geográfica Regulatório (SIG-R). O SIG-R consiste em um agregado de sistemas e bases de dados reunidas pela ANEEL que permitem a obtenção de diversas informações do sistema de distribuição e de seus usuários. Compõem o SIG-R dois principais elementos: a Base de Dados Geográfica da Distribuidora (BDGD) e Dicionário de Dados ANEEL do SIG-R (DDA), além de outras bases de dados que possibilitam ampliar o escopo de análise das informações por parte da Agência Reguladora (ANEEL, 2021a).

A partir de 2011, as distribuidoras passaram a ter a obrigatoriedade de encaminhar anualmente a BDGD, onde estão contidos seus dados de ativos e consumidores de forma geográfica, propiciando à ANEEL ter informações detalhadas de seus sistemas. A estrutura dos dados e o conjunto mínimo de informações que compõem a BDGD estão definidos no Módulo 10 do PRODIST (ANEEL, 2014b).

De posse desse amplo conjunto de informações disponíveis, a Agência vislumbrou a possibilidade de melhorias na metodologia de definição dos limites de DEC e FEC, especialmente em relação aos atributos dos conjuntos. No decorrer do 3º Ciclo de Revisões Tarifárias Periódicas, também foram identificados outros pontos de aprimoramento necessários à metodologia.

Nesse contexto, em 2014, a ANEEL realizou a Audiência Pública nº 029/2014 para receber contribuições de agentes, empresas e associações do setor elétrico, visando alterar a metodologia empregada até então. As Notas Técnicas nº 0059/2014-SRD/ANEEL e nº 0102/2014-SRD/ANEEL relatam os aperfeiçoamentos necessários à metodologia e as contribuições apresentadas.

No momento da confecção da metodologia, vigorava a Revisão 4 do Módulo 8 do PRODIST, o qual define um conjunto de unidades consumidoras, via de regra, por Subestação de Distribuição (SED). Os conjuntos são caracterizados por atributos, os quais serão extraídos das BDGD e de outras bases de dados constituintes do SIG-R disponíveis na ANEEL (ANEEL, 2021b).

A metodologia pode ser dividida em quatro etapas, as quais serão apresentadas nas seções seguintes (ANEEL, 2014b):

- Seleção e extração dos atributos a serem utilizados pela metodologia;
- Aplicação do método dinâmico para definição dos conjuntos semelhantes;
- Avaliação dos conjuntos heterogêneos e
- Avaliação do suprimento aos conjuntos e das trajetórias de redução intensas.

2.2.4.1 Seleção e extração dos atributos

Com o objetivo de avaliar quais atributos melhor explicam a variabilidade dos indicadores DEC e FEC, bem como caracterizar os conjuntos de unidades consumidoras, a ANEEL adotou a média dos anos de 2011, 2012 e 2013 dos indicadores de 2610 conjuntos de 43 distribuidoras, bem como 146 atributos (variáveis independentes). Os atributos podem ser classificados em seis tipos, conforme o Quadro 2.1 os apresenta.

Entende-se que muitas são as variáveis que influenciam na continuidade do fornecimento de energia elétrica, entretanto “para seleção de atributos algumas premissas devem ser adotadas, são elas: possibilidade de definição clara no regulamento, facilidade de obtenção de forma sistemática e viabilidade de fiscalização” (ANEEL, 2010b). Adicional-

mente, conforme detalhado pela ANEEL (2014b), diversos fatores implicaram a exclusão de atributos, destacando-se: inconsistência de registros, falta de padronização, redundância, falta de atualização, multicolinearidade, alto grau de subjetividade e irrelevância.

Quadro 2.1 – Tipos de atributos.

Tipo de Atributo	Total de Atributos	Fonte dos Dados
Elétrico	89	BDGD
Vegetação	20	PROBIO/2008 - Vegetação Nativa
Arruamento	21	Street Base/IMAGEM
Descargas atmosféricas	1	ONS
Precipitação	2	CPTEC/INPE
Socioeconômico	13	IBGE: Censo 2010 e outras

Fonte: ANEEL (2014b).

A partir dos fatores citados, realizou-se a exclusão prévia de 77 dos 146 atributos. Os 69 atributos restantes foram selecionados para a análise seguinte, a qual consiste na aplicação da técnica de regressão *stepwise*. Foram construídos então quatro modelos, dois para o DEC e dois para o FEC, considerando ou não as variáveis socioeconômicas.

Após diversas etapas de exclusão de variáveis por multicolinearidade a cada um dos quatro modelos, as aplicações da regressão *stepwise* resultaram na seleção de diferentes variáveis. Ao final dessa etapa, cada um dos quatro cenários continha entre 9 e 17 atributos.

Posteriormente foram comparados os dois pares de modelos para DEC e FEC, considerando-se o poder de explicação dos atributos em função da quantidade de variáveis no modelo. Concluiu-se que a inclusão de variáveis socioeconômicas nos modelos era pouco relevante, bem como a inserção de mais de seis variáveis. Dessa forma foram escolhidos os seis atributos mais relevantes para DEC e para FEC, sendo que cinco são comuns aos dois indicadores. O Quadro 2.2 apresenta os atributos finais escolhidos para o DEC e o FEC.

A redução de 13 para sete atributos no total em relação à metodologia anterior poderia indicar uma conseqüente redução na precisão da análise. Entretanto verificou-se que a proposta atual foi capaz de reduzir a quantidade de dados necessária para a análise ao passo que eleva substancialmente o poder explicativo dos modelos de DEC e FEC.

Definidos quais atributos serão empregados, realiza-se a extração dos dados dos atributos a partir das BDGD, que são encaminhadas anualmente pelas empresas para o SIG-R da ANEEL. Além dos dados do SIG-R, são utilizadas as bases de vegetação e precipitação pluviométrica disponíveis.

Quadro 2.2 – Atributos selecionados para a metodologia comparativa de DEC e FEC.

DEC	
Sigla	Atributo
PC_NUC_AD	Percentual de NUC em áreas de alta densidade (%)
PC_VRAM	Percentual de área com vegetação remanescente alta ou média (%)
PC_ERMT_3F	Percentual de redes MT trifásicas (%)
PLUV	Precipitação pluviométrica média anual (mm)
CM_NUC_RES	Consumo médio por UC da classe residencial (MWh)
NUC_IND	NUC da classe industrial
FEC	
Sigla	Atributo
PC_VRAM	Percentual de área com vegetação remanescente alta ou média (%)
PLUV	Precipitação pluviométrica média anual (mm)
PC_NUC_AD	Percentual de NUC em áreas de alta densidade (%)
NUC_COM	NUC da classe comercial
CM_NUC_RES	Consumo médio por UC da classe residencial (MWh)
PC_ERMT_3F	Percentual de redes MT trifásicas (%)

Fonte: ANEEL (2016).

2.2.4.2 Aplicação do método dinâmico

A etapa seguinte consiste na aplicação do método dinâmico para a *clusterização* dos conjuntos. Inicialmente cada um dos atributos adotados é padronizado. A padronização utilizada é o *Z-Score*, em que cada variável é transformada em uma nova variável com média zero e desvio padrão unitário, conforme a Equação 2.6. Faz-se necessária essa padronização para que seja removido o efeito da escala da unidade de medida de cada atributo (ANEEL, 2016).

$$x_{il} = \frac{x_{il}^* - m_l}{s_l}, i = 1, \dots, N; l = 1, \dots, d \quad (2.6)$$

Em que:

x_{il}^* = Dado original;

m_l = Média amostral do atributo l ;

s_l = Desvio-padrão amostral do atributo l ;

Os dados normalizados são usados no cálculo do grau de semelhança para cada par de conjuntos, o que é feito com a construção de uma matriz de similaridade. A medida de similaridade adotada é a distância euclidiana, conforme a Equação 2.7.

$$D(x_i, x_j) = \sqrt{\sum_{l=1}^d (x_{il} - x_{jl})^2} \quad (2.7)$$

Em que:

x_i = Denota o conjunto i ;

x_j = Denota o conjunto j ;

j = Índice do atributo, que varia de 1 a d ;

Então deve-se definir com quantos conjuntos deseja-se comparar cada conjunto de unidades consumidoras em análise. Em um estudo apresentado na Nota Técnica nº 0021/2011-SRD/ANEEL, definiu-se como desejável o número de 100 conjuntos (ANEEL, 2011). Dessa maneira, os conjuntos são comparados aos 100 conjuntos mais próximos, desde que respeitado um limite de homogeneidade entre eles.

O grau de heterogeneidade percentual de um agrupamento é calculado de acordo com a Equação 2.8. Definiu-se o valor de 20% como limite de heterogeneidade para uma comparação adequada entre conjuntos. Assim, caso o valor de heterogeneidade calculado para os 100 conjuntos semelhantes supere o limite de 20%, retiram-se os conjuntos mais distantes até que esta condição seja satisfeita. Entretanto considera-se 50 conjuntos como mínimo de conjuntos comparáveis.

$$Heterogeneidade_i^P = \frac{Max(Dist_i^j)}{3 \cdot \sqrt{k}} \quad (2.8)$$

Em que:

i = Índice do conjunto de referência;

j = Índice dos conjuntos próximos ao conjunto i ;

$Dist_i^j$ = Distância euclidiana do conjunto i para o conjunto j ;

k = Número de atributos;

No caso em que o valor de heterogeneidade ainda permaneça superior a 20%, respeitando-se o limite mínimo de 50 conjuntos semelhantes, considera-se que a comparação obtida não foi a ideal, e o conjunto em análise é classificado como conjunto heterogêneo. Tais conjuntos recebem um tratamento particular que será apresentado na Seção 2.2.4.3. O Quadro 2.3 mostra os valores estabelecidos para os parâmetros do método dinâmico (ANEEL, 2016).

Quadro 2.3 – Parâmetros do método dinâmico.

Parâmetro	Valor
Heterogeneidade percentual permitida	20%
Número mínimo de conjuntos semelhantes	50
Número desejável de conjuntos semelhantes	100

Fonte: ANEEL (2016).

Uma vez definidos os conjuntos semelhantes ao conjunto de interesse, inicia-se a definição dos limites. A ANEEL vale-se da técnica denominada *yardstick competition*, em que se estabelece um valor de referência para cada agrupamento, o qual definirá o limite objetivo a ser alcançado pelo conjunto em análise (ANEEL, 2016).

Os conjuntos do agrupamento são ordenados, do melhor para o pior, de acordo com a média dos desempenhos observados (valores apurados de DEC ou FEC dos conjuntos) para os três últimos anos civis disponíveis. A seguir toma-se o enésimo conjunto dessa sequência, cuja posição é definida pela Equação 2.9. Adota-se o percentil 20 para conjuntos interligados aéreos, e o percentil 50 para conjuntos subterrâneos e conjuntos aéreos isolados (ANEEL, 2016).

$$Posição = int((N_{ConjSeme} - 1) * Percentil + 1) \quad (2.9)$$

Em que:

$N_{ConjSeme}$ = Número de conjuntos semelhantes do agrupamento;

O limite objetivo é determinado pelo desempenho do percentil determinado pela Equação 2.9. O período de transição para que o conjunto alcance o limite objetivo é de oito anos, portanto constrói-se uma trajetória de limites para oito anos. Toda via, como os limites são definidos a cada revisão tarifária, tomam-se apenas os anos iniciais da trajetória, que varia de três a cinco anos, em função da periodicidade da revisão tarifária de cada distribuidora (ANEEL, 2016).

A trajetória de limites é linear. Inicia-se no limite atual do conjunto e atinge o limite objetivo em oito anos, conforme a Equação 2.10. A trajetória não deve ser crescente, portanto, caso o limite objetivo seja maior do que o limite inicial, aplica-se a Equação 2.11, ou seja, mantém-se o limite inicial até o final do período tarifário. Os valores resultante da Equação 2.10 são arredondados para números inteiros (ANEEL, 2016).

$$Limite_t = Limite_0 - t \cdot \left(\frac{Limite_0 - Limite_{objetivo}}{T} \right), \text{ se } Limite_0 > Limite_{objetivo} \quad (2.10)$$

$$Limite_t = Limite_0, \text{ se } Limite_0 \leq Limite_{objetivo} \quad (2.11)$$

Em que:

T - período de transição, considerado de 8 anos;

t - ano em que se deseja calcular o limite;

$Limite_t$ - limite a ser calculado para o ano t ;

$Limite_0$ - último limite já estabelecido para o conjunto (limite atual);

$Limite_{objetivo}$ - limite obtido aplicando-se o percentil (limite objetivo).

No caso de determinado conjunto ter sido classificado como heterogêneo, dá-se um tratamento especial. O procedimento para conjuntos heterogêneos é descrito na próxima seção.

2.2.4.3 Avaliação dos conjuntos heterogêneos

A metodologia, ainda, conta com a avaliação dos conjuntos heterogêneos por meio da criação do “Score ANI” (Atributo Normalizado Individual). Essencialmente, a determinação dos limites de DEC e FEC continua sendo feita com base na regra dos percentis, entretanto, nos casos em que os conjuntos são heterogêneos, realiza-se uma variação do percentil, que pode ser de -10% até +30%, com base no Score ANI (ANEEL, 2016).

Inicialmente, para o cálculo do Score ANI, normaliza-se individualmente cada um dos atributos, de modo que possuam valor mínimo igual a 0% e máximo igual a 100%. Essa grandeza normalizada é denominada Atributo Normalizado Individual (ANI). Devido a forte influência causada por valores extremos nos dados, retiram-se os *outliers* do cálculo do ANI. Um *outlier* é definido como um valor que excede a média mais três vezes o desvio padrão de um determinado atributo (ANEEL, 2016).

Leva-se em consideração, no cálculo do ANI, que a complexidade na prestação do serviço de distribuição pode aumentar ou diminuir quando se incrementa o valor de um atributo, em função da natureza da grandeza em questão. Desse modo, para atributos cuja complexidade aumenta conforme se eleva o valor do atributo, o valor do ANI é calculado pela Equação 2.12. Para atributos cuja complexidade diminui com o aumento do valor do atributo, o valor do ANI é dado pela Equação 2.13. O Quadro 2.4 apresenta qual equação é utilizada para o ANI de cada um dos sete atributos selecionados para a metodologia (ANEEL, 2016).

$$ANI_{i,j} = \frac{X_{i,j} - X_{jMin}}{X_{jMax} - X_{jMin}} \cdot 100\% \quad (2.12)$$

$$ANI_{i,j} = 100 - \frac{X_{i,j} - X_{jMin}}{X_{jMax} - X_{jMin}} \cdot 100\% \quad (2.13)$$

Em que:

i = Índice do conjunto;

j = Índice do atributo;

$X_{i,j}$ = Valor do atributo j do conjunto i ;

X_{jMin} = Valor mínimo do atributo j na base de dados;

X_{jMax} = Valor máximo do atributo j na base de dados (excetuando-se os *outliers*);

Quadro 2.4 – Equações utilizadas para cálculo do ANI dos atributos selecionados para a metodologia comparativa.

Sigla do Atributo	Equação para Cálculo do ANI
PC_NUC_AD	Equação 2.13
PC_VRAM	Equação 2.12
PC_ERMT_3F	Equação 2.13
PLUV	Equação 2.12
CM_NUC_RES	Equação 2.13
NUC_IND	Equação 2.13
NUC_COM	Equação 2.13

Fonte: ANEEL (2016).

Após a obtenção dos ANI para todos os atributos de todos os conjuntos da base de dados, é feito o cálculo do indicador *Score ANI* para cada conjunto heterogêneo, conforme a Equação 2.14 (ANEEL, 2016).

$$ScoreANI = \frac{\sum_{j=1}^n (ANI_{i,j}) - \overline{ANI}_j}{n} \quad (2.14)$$

Em que:

$ANI_{i,j}$ = ANI do conjunto heterogêneo para o atributo j ;

\overline{ANI}_j = ANI médio dos conjuntos do agrupamento do conjunto heterogêneo, para o atributo j ;

j = Índice do atributo;

n = Quantidade de atributos;

O *Score ANI* determina se esse conjunto possui, em média, atributos mais ou menos favoráveis quando comparado aos conjuntos de seu agrupamento. Assim, utiliza-se o *Score ANI* para variar o percentil aplicado ao conjunto heterogêneo, de acordo com o Quadro 2.5.

Quadro 2.5 – Variação nos percentis a serem utilizados para conjuntos heterogêneos com base no *Score ANI*.

Score ANI	Variação do percentil
< -3%	-10%
≥-3% e < 3%	0%
≥3% e < 6%	+10%
≥6% e < 9%	+20%
≥9%	+30%

Fonte: ANEEL (2016).

A última etapa da metodologia de definição dos limites de DEC e de FEC consiste na avaliação do suprimento aos conjuntos e na avaliação das trajetórias de redução intensas.

2.2.4.4 Avaliação do suprimento e das trajetórias de redução intensas

Na definição dos limites objetivos, são empregados os valores apurados de DEC e FEC, indicadores os quais levam em consideração as interrupções de origem externa ao sistema de distribuição (interrupções do suprimento). Dessa maneira, nesta etapa é avaliado se a participação do suprimento nos indicadores é expressiva. Nesses casos adota-se uma trajetória de limites atenuados.

Com a entrada em vigor, em 2022, da Resolução Normativa ANEEL nº 925/2021, que altera o Módulo 8 do PRODIST, a avaliação do suprimento deixará de ser realizada. A apuração dos indicadores DEC e FEC deixará de contabilizar as interrupções de origem externa, portanto não se fará necessária tal análise (ANEEL, 2021c).

Avalia-se, ainda, se a trajetória de redução de limites é gradual e factível. Definiu-se um limitador de redução anual de oito horas para o DEC e cinco interrupções para o FEC. Assim, conjuntos que possuam reduções intensas, tem suas trajetórias de redução limitadas a esses valores.

Por fim, os valores calculados por meio da metodologia apresentada são levados à discussão em audiências públicas. Dá-se oportunidade às distribuidoras e aos consumidores de apresentarem contribuições relevantes sobre os limites propostos. As contribuições são avaliadas pela ANEEL, que poderá ajustar os limites resultantes da metodologia a partir das justificativas apresentadas.

Nesse sentido, as concessionárias de energia elétrica investem parte de seus recursos na execução de projetos de pesquisa e desenvolvimento (P&D), a fim de serem avaliadas possíveis melhorias operacionais e, ainda, contribuições às normas impostas pela Agência Reguladora. Tais projetos possibilitam embasar contribuições às consultas públicas realizadas pela ANEEL.

Um dos projetos de P&D das Centrais Elétricas de Santa Catarina (CELESC) é o de nº 5697-0718/2018, intitulado “Definição de conjuntos de unidades consumidoras eficientes sob os enfoques dos consumidores, das empresas e do regulador”. Neste trabalho, chegou-se à conclusão de que um dos principais problemas da metodologia utilizada pela ANEEL é caracterizado pela possibilidade de conjuntos serem compostos por regiões com características bem distintas, tanto de padrão de rede quanto de mercado suprido. Como consequência, conjuntos considerados semelhantes (pelos atributos utilizados) podem ser bem distintos na realidade. Dessa forma, um dos objetivos do projeto consistiu em se identificar novos atributos mais adequados para caracterizar os conjuntos e que possuam maior poder de explicação dos indicadores DEC e FEC.

Outro projeto de P&D, desta vez da Companhia Estadual de Distribuição de Energia Elétrica (CEEE-D), intitulado “Avaliação de novos atributos para metodologia de análise comparativa de desempenho considerando as características técnicas e geográficas da área de concessão da CEEE-D e do Rio Grande do Sul”, tem como objetivo definir novos atributos para a metodologia de análise comparativa de desempenho, considerando as características técnicas e geográficas da área de concessão da CEEE-D.

Diante das metodologias de definição de limites de DEC e FEC expostas, observa-se um ponto comum a elas: o emprego de técnicas de *clusterização*. A Seção 2.3 aborda os métodos utilizados nas metodologias, com exemplos de aplicação. Dentre os métodos hierárquicos, apresenta-se o método de Ward. Em seguida, discorre-se sobre o método não hierárquico *k-means* e, por fim, é apresentado o método dinâmico de *clusterização*.

2.3 MÉTODOS DE CLUSTERIZAÇÃO

A análise de agrupamentos representa um conjunto de técnicas exploratórias que podem ser aplicadas quando se deseja verificar a existência de comportamentos semelhantes entre observações (indivíduos, objetos, empresas, localidades, dentre outros exemplos) em relação a determinadas variáveis. O objetivo principal é o de se criar grupos, ou *clusters*, em que prevaleça a homogeneidade interna (FÁVERO; BELFIORE, 2017).

As técnicas de agrupamento, ou de *clusterização*, são utilizadas em situações nas quais a divisão de um universo de informações em subconjuntos permite uma melhor análise dos dados. A análise de *cluster* é, na verdade, um meio para o tratamento de um grande volume de dados, desempenhando o papel de redução dos dados e de geração de hipóteses (HAIR et al., 2009).

Essas técnicas de análise de dados encontram aplicação em muitos campos de pesquisa. Na biologia são utilizadas na obtenção de taxonomias para agregar todos os organismos vivos. Permitem, no campo da economia, a análise de segmentação de mercados voltada à definição de estratégias de *marketing*, por exemplo. De fato essas técnicas constituem-se como uma poderosa ferramenta de pesquisa, que permitem sugerir modelos estatísticos para descrever populações (HAIR et al., 2009).

No contexto da distribuição de energia elétrica, Tanure (2000) apresenta uma metodologia para o estabelecimento das metas de continuidade a serem cumpridas pelas concessionárias de distribuição. Assim como na metodologia utilizada pela ANEEL, as metas são fixadas para cada conjunto de UC, sendo a meta global da empresa definida como uma média do desempenho dos conjuntos. O autor realiza o agrupamento dos conjuntos por meio dos métodos de *clusterização* de Ward e das k-médias e propõe uma análise comparativa entre regiões similares de diferentes distribuidoras, visando o processo de melhoria contínua dos indicadores. Essa dissertação embasou a metodologia de análise

comparativa adotada inicialmente pela ANEEL.

Sperandio (2004) apresenta uma metodologia para formação de agrupamentos de conjuntos consumidores de energia elétrica, contemplando desde a seleção das variáveis a serem utilizadas no processo até a classificação dos grupos de acordo com a qualidade do fornecimento de seus integrantes. A partir dos dados de 260 conjuntos sob concessão da CELESC, contendo 18 variáveis que descrevem características do mercado consumidor e do sistema elétrico de fornecimento, realiza-se a seleção das variáveis mais significativas por meio da análise estatística multivariada. Então, utilizando uma validação cruzada entre um algoritmo que utiliza paradigmas da inteligência artificial, chamado de Mapa Auto-Organizável, e o tradicional método estatístico das *k*-médias, o autor obtém o agrupamento dos conjuntos consumidores. Esse trabalho é referenciado na Nota Técnica nº 0028/2010-SRD/ANEEL.

A dissertação apresentada por Rosa (2009) propõe o aprimoramento da metodologia usada pela ANEEL no estabelecimento das metas, perante duas alterações: aumento no número de atributos para classificar conjuntos semelhantes e melhoria da metodologia de agrupamento dos conjuntos. O método hierárquico de Ward é proposto no trabalho para o agrupamento, por ser mais robusto e não estar sujeito a instabilidades, visto que o método *k*-médias, então usado pela ANEEL, pode levar a certa instabilidade nos resultados.

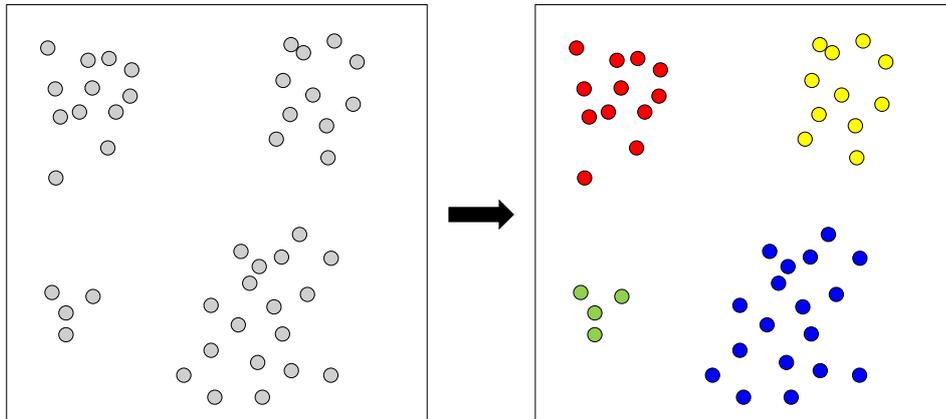
Os autores Granell, Axon e Wallom (2015), Hsu (2015), Al-Jarrah et al. (2017), Cai et al. (2019) utilizam o método *k-means*, além de outros métodos de *clusterização*, em análises de perfis de consumo de energia elétrica.

O agrupamento de elementos semelhantes ocorre com base em variáveis, chamadas também de atributos, as quais descrevem cada um desses elementos. Assim, os atributos a serem considerados no processo de *clusterização* devem ser definidos para cada indivíduo do universo em análise. A escolha dos atributos que comporão o processo de *clusterização* é fulcral para essa técnica de análise, porque as variáveis escolhidas devem ser relevantes para o que se deseja analisar (TANURE, 2000). Técnicas de regressão podem ser empregadas para a seleção dos atributos, como ocorreu na definição da metodologia atual de cálculo de limites de DEC e FEC.

Em geral, as técnicas de agrupamento de dados recorrem aos espaços geométricos para representação do processo. Desse modo é possível analisar mais objetivamente os atributos envolvidos e visualizar os *clusters* formados. A Figura 2.3 mostra um conjunto de pontos que passou por um processo de agrupamento de forma meramente visual. Os grupos, classificados em função da distância entre seus elementos, são representados por cores distintas.

A *clusterização* permite realizar inferências sobre o universo analisado de forma ampla, sem que haja a necessidade de se analisar os elementos individualmente. No exemplo, verifica-se que a divisão do conjunto em quatro grupos facilita a visualização da distribuição dos pontos.

Figura 2.3 – Exemplo de *clusterização*.

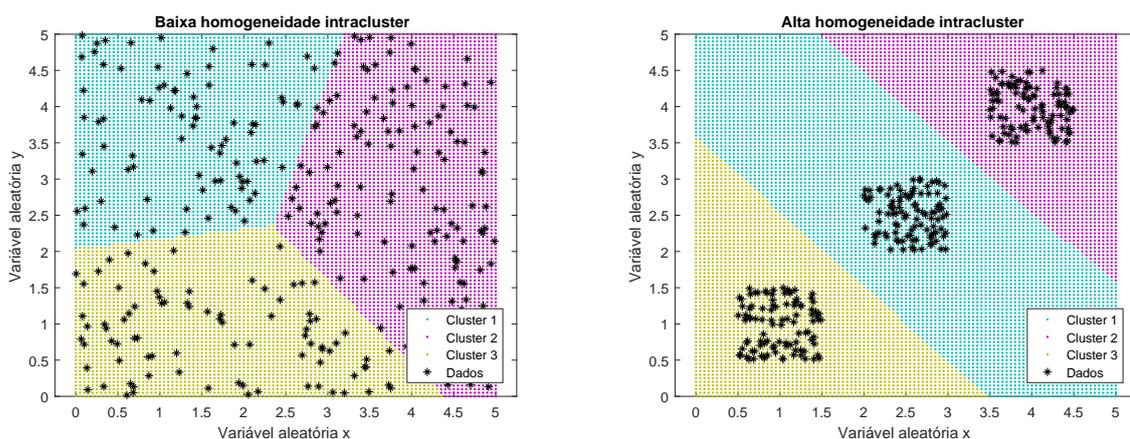


Fonte: Autor.

Em um processo de *clusterização*, almeja-se formar *clusters* com o máximo grau de homogeneidade entre os elementos, ou seja, que os elementos contidos em um mesmo agrupamento sejam os mais semelhantes possível. Essa condição implica que entre os *clusters* deverá haver o maior grau de heterogeneidade possível.

A Figura 2.4 mostra a aplicação de um método de *clusterização* em dois conjuntos de dados distintos, ambos contendo 300 observações. Verifica-se, na Figura 2.4(a), que a classificação dos dados em três *clusters* é ineficiente, pois a homogeneidade entre os elementos de um mesmo *cluster* é baixa. Em contrapartida, na Figura 2.4(b) verificam-se três classes de dados bem definidas, ou seja, que possuem alta homogeneidade *intracluster*.

Figura 2.4 – Homogeneidade *intracluster*.



(a) Baixa homogeneidade *intracluster*

(b) Alta homogeneidade *intracluster*

Fonte: Autor.

Os métodos de agrupamento são classificados em hierárquicos e não hierárquicos. Os métodos não hierárquicos possibilitam formar uma quantidade pré-estabelecida de *clusters*, enquanto os métodos hierárquicos realizam uma sequência de agrupamentos

sucessivos para classificar os dados em um número decrescente de *clusters* (FÁVERO; BELFIORE, 2017).

Os elementos de um universo são agrupados com base na semelhança entre eles. Medidas de similaridade podem se referir à distância entre os membros de um *cluster*, à densidade de elementos em uma área do espaço ou, ainda, à conformidade das observações com distribuições estatísticas. Na maioria dos casos, aplica-se o método da distância euclidiana como medida de similaridade entre dois elementos.

Em uma análise envolvendo m atributos, considera-se um espaço m -dimensional, em que os valores das variáveis são coordenadas desse espaço. Assim, a distância euclidiana entre dois pontos x e y em um espaço m -dimensional é definida pela Equação 2.15.

$$D(x, y) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \dots + (x_m - y_m)^2} \quad (2.15)$$

Em resumo, os métodos de *clusterização* são empregados, via de regra, quando não se é possível escrever um conjunto de equações que permitam definir com exatidão os valores a serem assumidos por determinadas variáveis de um problema. Neste trabalho, são apresentados três métodos de classificação estatística: o método de Ward, o *k-means* e o método dinâmico.

2.3.1 Métodos Hierárquicos

Os métodos hierárquicos consistem de um processo iterativo no qual os elementos de um universo considerado são agregados um a um e de forma crescente, até a completa alocação de todos os elementos em um único *cluster*. Em outras palavras, esses métodos consistem em se produzir uma série de agregações de k *clusters*, promovendo-se, em cada passo, a agregação de um dos elementos anteriores (HAIR et al., 2009).

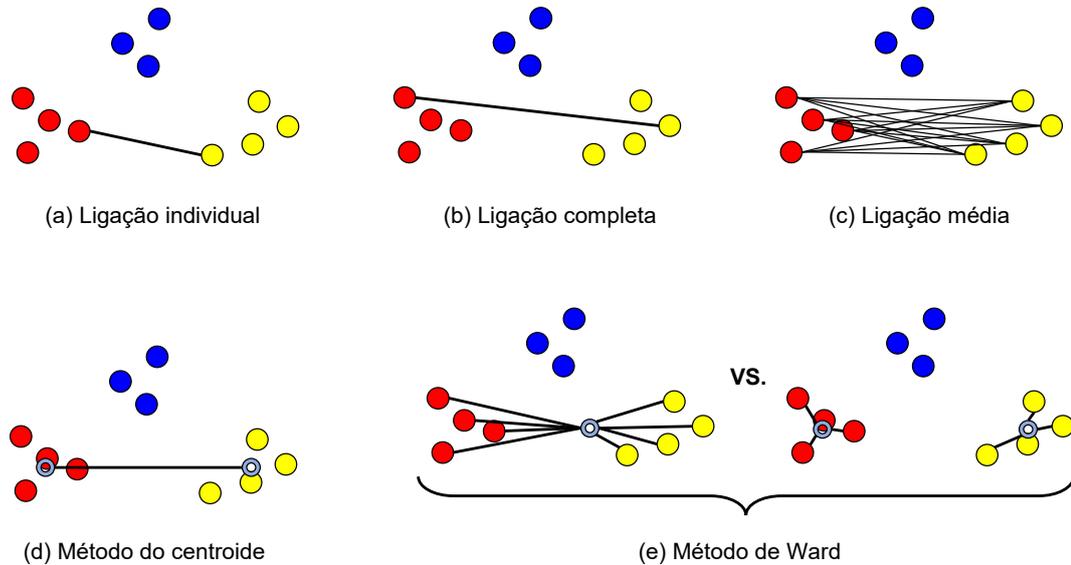
Dada uma população de n objetos, o método inicia-se com a formação de k *clusters*, de forma que cada elemento corresponda a um *cluster*, ou seja, $k = n$. No segundo passo, realiza-se a primeira agregação, resultando-se em $n - 1$ *clusters*. No terceiro, obtém-se $n - 2$ *clusters* e assim sucessivamente até que todos os elementos estejam contidos em um único *cluster*. Assim, são necessárias $n - 1$ iterações para o agrupamento completo (HAIR et al., 2009).

Os *clusters* são agrupados com base na distância entre eles. Os métodos hierárquicos distinguem-se, dentre outros fatores, na forma como se define a distância entre dois *clusters*. Dentre os critérios de agrupamento, citam-se as seguintes possibilidades, ilustradas pela Figura 2.5 (HAIR et al., 2009):

- (a) Menor distância entre elementos de *clusters* diferentes (Ligação individual);
- (b) Maior distância entre elementos de *clusters* diferentes (Ligação completa);
- (c) Distância média entre elementos de *clusters* diferentes (Ligação média);

- (d) Distância entre os centroides de *clusters* diferentes (Método do centroide) e
 (e) Erro gerado pelo agrupamento de *clusters* diferentes (Método de Ward).

Figura 2.5 – Formas de medição da distância entre *clusters*.



Fonte: Autor.

No método de ligação individual, a medida de similaridade entre dois *clusters* é definida pela menor distância entre qualquer ponto do primeiro *cluster* para qualquer ponto do segundo *cluster*. No método de ligação completa, a similaridade define-se pela maior distância entre elementos de *clusters* diferentes.

No método da ligação média, a semelhança trata-se da média de todas as distâncias entre os elementos do primeiro e do segundo *clusters*. No método do centroide a similaridade é definida pela distância entre os pontos médios de ambos os *clusters*.

O método de Ward, também conhecido como método da mínima variância, utiliza como medida de distância a soma das distâncias ao quadrado entre os dois agrupamentos. Avalia-se como a agregação de dois *cluster* influencia no aumento da dispersão dos elementos, ou seja, mede-se o erro associado ao agrupamento.

Dentre os métodos hierárquicos de classificação, o método de Ward é o que possui maior aceitação, pois este minimiza a perda de informação (erro) ao se formar um novo *cluster*. Outra razão é a de que, apesar das diferenças, os métodos produzem sequências de agrupamentos semelhantes e, portanto, não se verifica a necessidade de comparação entre os métodos (TANURE, 2000).

2.3.1.1 Método de Ward

O método de Ward insere-se no contexto dos métodos hierárquicos de *clusterização*. Esse método utiliza o critério de menor erro, ou menor perda de informação, para se determinar a sequência de agrupamento a ser adotada. O erro, conforme a Equação 2.16, é definido como o somatório dos quadrados das distâncias entre cada elemento dos dois *cluster* a serem agrupados e o valor central (centroide) do novo *cluster* formado (HAIR et al., 2009).

$$SQE = \sum_{i=1}^n (X_i - \bar{X})^2 \quad (2.16)$$

Esse critério pode ser expresso pelo seguinte questionamento: “Ao se considerar o agrupamento de dois *clusters*, quanto isso afetará a distância total entre o centroide e os elementos?”.

A fim de se implementar o método de Ward, aplica-se o seguinte algoritmo (HAIR et al., 2009):

- Passo 1 - Considera-se cada elemento como um *cluster*, determina-se qual a combinação de *clusters* dois a dois que resulta em um menor erro, e agrupam-se os dois mais semelhantes;
- Passo 2 - Considera-se o novo agrupamento como um novo *cluster* e guarda-se a informação de sequência de agrupamento e nível;
- Passo 3 - Determina-se qual a combinação de elementos, elementos e *clusters* ou *clusters* que implica um menor erro; e
- Passo 4 - Se ainda existem elementos a serem combinados retorna-se ao passo 2.

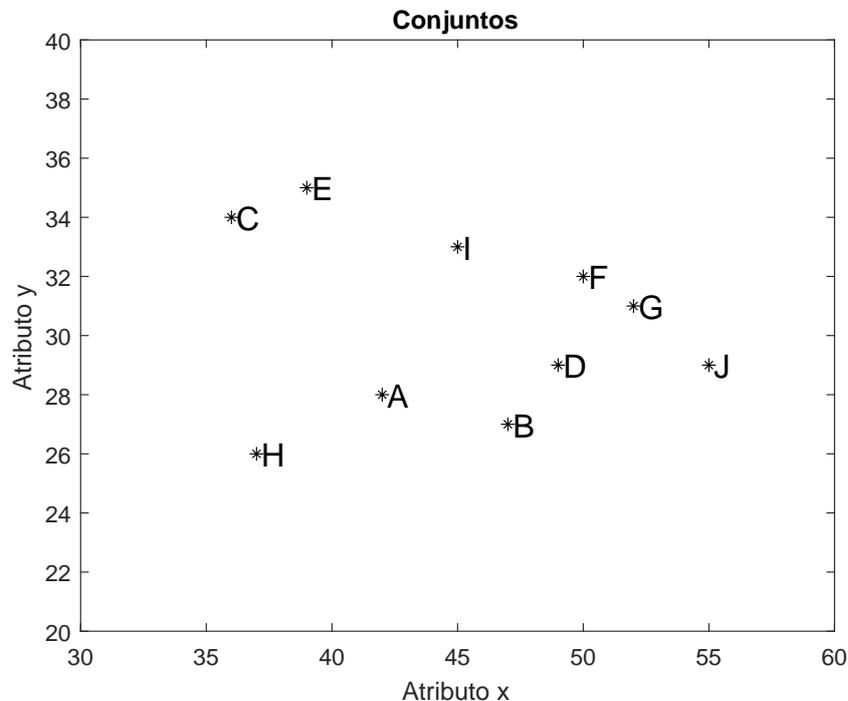
Para exemplificar este método de agregação, são considerados os dados fictícios do Quadro 2.6, os quais descrevem dez conjuntos (A a J) segundo os atributos x e y.

Quadro 2.6 – Atributos dos conjuntos a serem analisados.

Conjunto	Atributo x	Atributo y
A	42	28
B	47	27
C	36	34
D	49	29
E	39	35
F	50	32
G	52	31
H	37	26
I	45	33
J	55	29

Os dados do Quadro 2.6 podem ser representados em um plano cartesiano, no qual os valores das variáveis são coordenadas desse espaço. A Figura 2.6 ilustra essa representação.

Figura 2.6 – Representação dos conjuntos no plano cartesiano.



Fonte: Autor.

Inicialmente, deve-se obter o erro para cada par de *clusters*, em que cada *cluster* corresponde a um dos dez conjuntos. Os erros são apresentados em uma matriz de similaridade, conforme o Quadro 2.7.

Quadro 2.7 – Matriz de similaridade usando o critério do menor erro.

Cluster	1:A	2:B	3:C	4:D	5:E	6:F	7:G	8:H	9:I	10:J
1:A	0	13	36	25	29	40	54,5	14,5	17	85
2:B	13	0	85	4	64	17	20,5	50,5	20	34
3:C	36	85	0	97	5	100	132,5	32,5	41	193
4:D	25	4	97	0	68	5	6,5	76,5	16	18
5:E	29	64	5	68	0	65	92,5	42,5	20	146
6:F	40	17	100	5	65	0	2,5	102,5	13	17
7:G	54,5	20,5	132,5	6,5	92,5	2,5	0	125	26,5	6,5
8:H	14,5	50,5	32,5	76,5	42,5	102,5	125	0	56,5	166,5
9:I	17	20	41	16	20	13	26,5	56,5	0	58
10:J	85	34	193	18	146	17	6,5	166,5	58	0

Fonte: Autor.

Da matriz de similaridade obtém-se que o par de *clusters* que resulta no menor erro

são os conjuntos F e G. Desse modo, F e G formam o primeiro agrupamento de *clusters*, com erro igual a 2,5.

O próximo passo consiste em se construir uma nova matriz de similaridade, desta vez eliminando-se as linhas e colunas referentes ao *clusters* F e G e adicionando-se o *cluster* FG, como apresentado no Quadro 2.8. Recalculam-se os erros entre cada par de *clusters*. Tendo em vista a simetria da matriz de similaridade, estão representados no Quadro 2.8 apenas os elementos da diagonal inferior.

Quadro 2.8 – Matriz de similaridade após o primeiro agrupamento.

Cluster	1:A	2:B	3:C	4:D	5:E	6:FG	7:H	8:I	9:J
1:A	0								
2:B	13	0							
3:C	36	85	0						
4:D	25	4	97	0					
5:E	29	64	5	68	0				
6:FG	93,3	36,3	231,3	10,3	156,3	0			
7:H	14,5	50,5	32,5	76,5	42,5	226,3	0		
8:I	17	20	41	16	20	38,3	56,5	0	
9:J	85	34	193	18	146	22,3	166,5	58	0

Fonte: Autor.

Desta vez, o par de *clusters* que resulta no menor erro são os *clusters* B e D. Assim, B e D formam o segundo agrupamento de *clusters*, com erro igual a 4.

Os passos 2 e 3 do algoritmo devem ser repetidos até que haja um único *cluster* formado, encerrando-se o método de *clusterização* de Ward. Ao final do processo, ter-se-á o registro dos agrupamentos realizados e os respectivos níveis de distância conforme apresenta o Quadro 2.9.

Quadro 2.9 – Elementos agrupados e níveis.

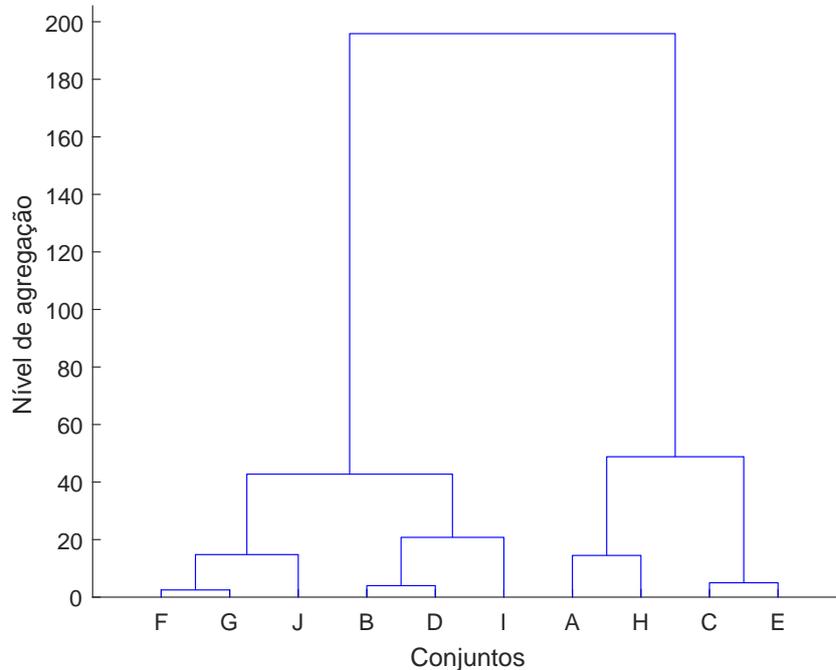
Estágio	Cluster 1	Cluster 2	Nível
1	F	G	2,50
2	B	D	4,00
3	C	E	5,00
4	A	H	14,50
5	J	FG	14,79
6	I	BD	20,79
7	JFG	IBD	42,75
8	AH	CE	48,81
9	AHCE	JFGIBD	195,86

Fonte: Autor.

Usualmente o método hierárquico é representado por um diagrama de árvore denominado dendograma. Com base nas informações do Quadro 2.9, constrói-se o den-

dograma, apresentado na Figura 2.7, resultante do exemplo de aplicação do método de Ward.

Figura 2.7 – Dendograma considerando o método de Ward.



Fonte: Autor.

No exemplo apresentado, aplicou-se o método de Ward a um pequeno conjunto de dados. Para a análise de um grande número de informações, esse método mostra-se demasiadamente trabalhoso, tendo em vista a grande quantidade de operações matemáticas envolvidas. A *clusterização* de 1000 elementos, por exemplo, pelo método hierárquico descrito exigiria calcular o erro mais de 160 milhões de vezes.

Deve-se ressaltar que a escolha do número de *clusters* a serem formados pode ser feita intuitivamente, com base no dendrograma e nos níveis de agregação. Cabe a quem está analisando os dados ponderar a perda de informação em detrimento da quantidade de agrupamentos.

Outro método de *clusterização*, não hierárquico, é exposto na Seção 2.3.2.

2.3.2 Método *k-means*

Nas situações em que o problema a ser estudado engloba um conjunto significativo de dados, os métodos hierárquicos mostram-se inviáveis, visto que a análise e a interpretação dos dendogramas tornam-se trabalhosas. Além disso, a construção das matrizes de similaridade exige um esforço computacional elevado (TANURE, 2000).

As técnicas não hierárquicas surgem no sentido de dar uma solução adequada a esses problemas, procurando encontrar as melhores combinações de elementos que formem *clusters* homogêneos em um número previamente definido (TANURE, 2000). O objetivo é classificar um universo de n elementos em k grupos, de forma que cada unidade pertença a um único grupo (SPERANDIO, 2004).

Preliminarmente, esse pode parecer um problema de combinatória, em que, definido um critério de avaliação, bastaria analisar todas as combinações possíveis de n objetos em k partições. Contudo, em um estudo envolvendo, por exemplo, 50 elementos em 5 *clusters*, seria necessário verificar mais de dois milhões de combinações, o que é uma tarefa praticamente impossível (TANURE, 2000).

O método das k -médias (ou, em inglês, *k-means*) visa contornar esse empecilho. Trata-se de um algoritmo heurístico simples, iterativo e de rápida convergência. A seguir são apresentados o método e sua aplicação aos dados fictícios do Quadro 2.6. O fluxograma da Figura 2.8 ilustra os passos do método das k -médias.

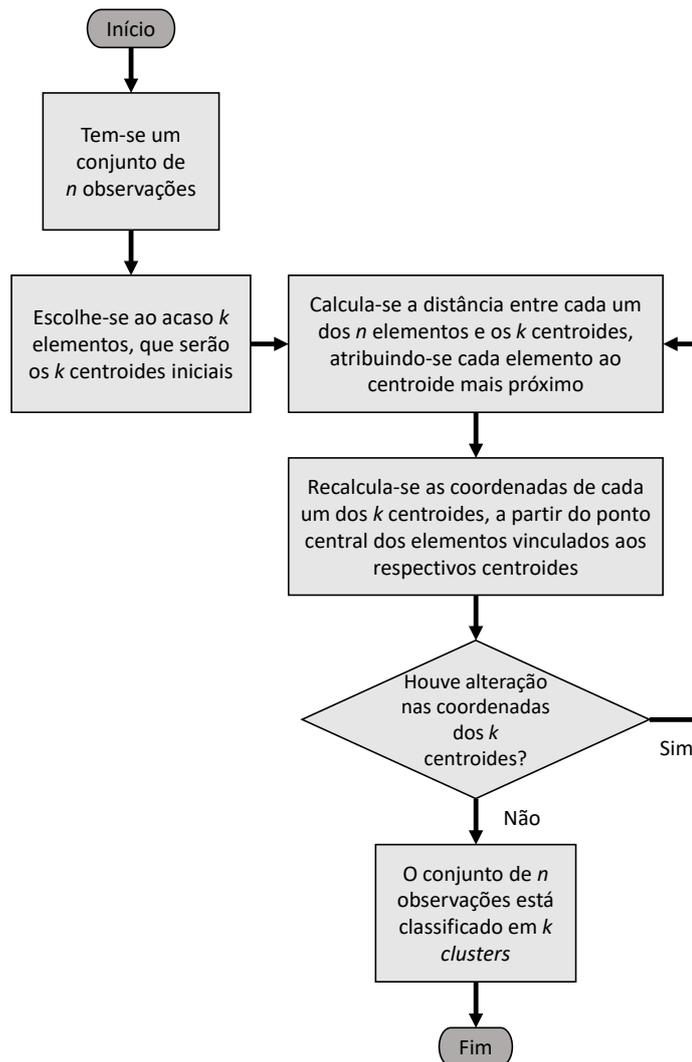
São quatro os passos a serem seguidos para aplicação do método *k-means* (HAIR et al., 2009):

- Passo 1 - Escolhe-se k elementos entre os n elementos do universo considerado. Esses k elementos serão considerados centros dos *clusters* a serem formados (centroides);
- Passo 2 - Calcula-se a distância euclidiana entre cada um dos n elementos e os k centroides, atribuindo-se cada elemento ao centroide mais próximo;
- Passo 3 - Recalcula-se as coordenadas de cada um dos k centroides, a partir do ponto central dos elementos vinculados aos respectivos centroides; e
- Passo 4 - Repete-se os passos 2 e 3 até que não se verifique nenhuma alteração dos elementos agrupados em torno dos k centros. Outra maneira de controlar a convergência do processo é verificando se houve alteração nas coordenadas dos k centros.

Apesar da simplicidade desse processo, a escolha dos k centroides iniciais pode influenciar no arranjo final dos *clusters*. Dentre os possíveis critérios para essa escolha, pode-se escolher os k primeiros dados da base, sorteá-los aleatoriamente, ou, ainda, gerar centroides com base nos autovetores da base de dados (SPERANDIO, 2004).

Para a aplicação dessa técnica de *clusterização* a um exemplo didático, tomam-se os dados referentes a dez conjuntos (Quadro 2.6). Escolheu-se segregar os dez conjuntos em dois agrupamentos ($k = 2$). Como critério de escolha inicial dos centroides, serão tomados os dois primeiros elementos da tabela de dados, ou seja, os conjuntos A e B.

A partir dos pontos centrais escolhidos, calculam-se as distâncias de cada elemento a cada um dos k centros, as quais são apresentadas no Quadro 2.10. Em seguida, associa-se cada um dos conjuntos ao centroide mais próximo.

Figura 2.8 – Fluxograma do método *k-means*.

Fonte: Autor.

Quadro 2.10 – Distância dos conjuntos aos centroides iniciais.

Conjunto	Dist. ao centroide 1	Dist. ao centroide 2	Cluster associado
A	0,0000	5,099	1
B	5,0990	0,0000	2
C	8,4853	13,0384	1
D	7,0711	2,8284	2
E	7,6158	11,3137	1
F	8,9443	5,8310	2
G	10,4403	6,4031	2
H	5,3852	10,0499	1
I	5,8310	6,3246	1
J	13,0384	8,2462	2

Fonte: Autor.

Pode-se observar que a distância entre o conjunto A e o centroide 1 é igual a zero pois tratam-se de pontos coincidentes no espaço. O mesmo vale para o conjunto B e o centroide 2.

Posteriormente, procede-se o passo 3. A partir dos elementos vinculados aos centroides 1 e 2, obtém-se as suas respectivas novas coordenadas: (39,8, 31,2) e (50,6, 29,6). Como houve mudança nas coordenadas de ambos os centroides, procede-se a repetição dos passos 2 e 3. A novas distâncias entre os n conjuntos e o k centroides são mostradas no Quadro 2.11.

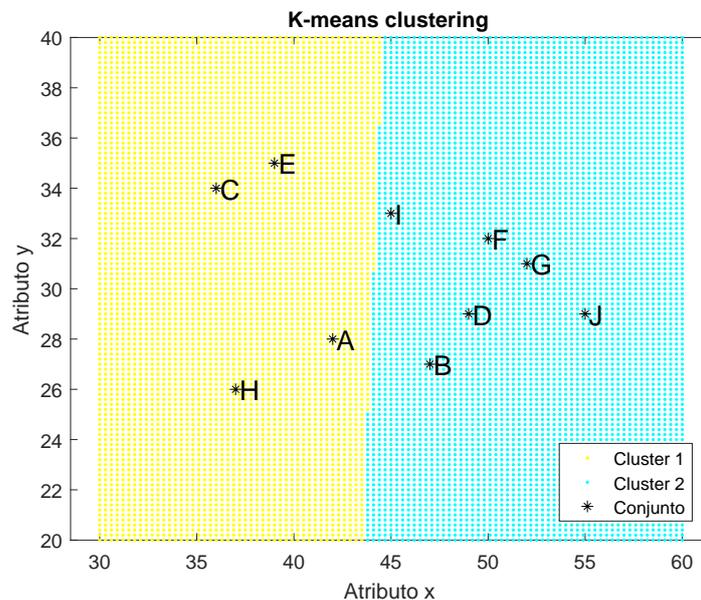
Quadro 2.11 – Distância dos conjuntos aos centroides na segunda iteração.

Conjunto	Dist. ao centroide 1	Dist. ao centroide 2	Cluster associado
A	3,8833	8,7476	1
B	8,3355	4,4407	2
C	4,7202	15,2486	1
D	9,4594	1,7088	2
E	3,8833	12,7953	1
F	10,2313	2,4739	2
G	12,2016	1,9799	2
H	5,9059	14,0684	1
I	5,5027	6,5513	1
J	15,3584	4,4407	2

Fonte: Autor

A Figura 2.9 ilustra a *clusterização* por meio de um diagrama de Voronoi, no qual o plano cartesiano é dividido em dois semiplanos.

Figura 2.9 – Diagrama de Voronoi da *clusterização* pelo método *k-means*.



Fonte: Autor.

Verifica-se que, após a repetição dos passos 2 e 3, não houve alteração dos elementos agrupados em cada um dos dois *clusters* e, portanto, ocorreu a convergência do processo. Desse modo, os dois *clusters* formados são:

- *Cluster* 1: ACEH
- *Cluster* 2: BDFGIJ

O próximo método a ser apresentado neste trabalho é o método dinâmico.

2.3.3 Método Dinâmico

Apesar das vantagens das técnicas de *cluster* hierárquica e não hierárquica, esses métodos podem não ser adequados às análises de elementos localizados nas fronteiras entre os *clusters*. *Outliers*, como são denominados esses elementos, tratam-se de observações atípicas, distantes dos seus respectivos centroides. Uma possível solução para os *outliers* é aumentar o número de agrupamentos formados, de modo a se manter uma relação de compromisso entre o número de *clusters* e o grau de representatividade centroides. O aumento do número de *clusters*, por outro lado, não é viável quando se deseja classificar uma amostra contendo poucos elementos (TANURE, 2004).

Nesse sentido, o método de *cluster* dinâmico visa contornar tais problemas. Esse método tem como premissa formar *clusters* independentes para cada elemento.

Para aplicar esta técnica, realiza-se o seguinte algoritmo (TANURE, 2004):

- Passo 1 - Adota-se cada elemento como centroide de um *cluster* a ser criado;
- Passo 2 - Definido o centroide, determina-se o grau de semelhança (distância) de todos os elementos com este centroide;
- Passo 3 - Estabelece-se um critério de corte para a distância entre o centroide e cada elemento;
- Passo 4 - Agrupa-se a cada centroide os elementos mais representativos com base em suas semelhanças;
- Passo 5 - Extraem-se as características de interesse do *cluster* formado; e
- Passo 6 - Repete-se o processo para cada um dos elementos.

A determinação da medida de semelhança pode ser feita, bem como nos demais métodos de *cluster*, de diversas formas. Em especial, emprega-se a distância euclidiana devido à sua versatilidade.

A partir dos conjuntos e dos atributos do Quadro 2.6, pode-se aplicar o método dinâmico a esses dados. De forma semelhante ao que foi realizado no método de Ward, constrói-se uma matriz de similaridade entre os conjuntos, como apresenta o Quadro 2.12. O grau de semelhança entre os conjuntos é definido pela distância euclidiana entre eles.

Com base no Quadro 2.12, pode-se classificar, para cada conjunto, quais são os

mais semelhantes. O Quadro 2.13 mostra as tabelas de similaridade para os conjuntos de A a J. Em ordem crescente, encontram-se as distâncias absolutas e relativas entre os conjuntos.

Quadro 2.12 – Matriz de similaridade usando a distância euclidiana.

Caso	1:A	2:B	3:C	4:D	5:E	6:F	7:G	8:H	9:I	10:J
1:A		5,10	8,49	7,07	7,62	8,94	10,44	5,39	5,83	13,04
2:B	5,10		13,04	2,83	11,31	5,83	6,40	10,05	6,32	8,25
3:C	8,49	13,04		13,93	3,16	14,14	16,28	8,06	9,06	19,65
4:D	7,07	2,83	13,93		11,66	3,16	3,61	12,37	5,66	6,00
5:E	7,62	11,31	3,16	11,66		11,40	13,60	9,22	6,32	17,09
6:F	8,94	5,83	14,14	3,16	11,40		2,24	14,32	5,10	5,83
7:G	10,44	6,40	16,28	3,61	13,60	2,24		15,81	7,28	3,61
8:H	5,39	10,05	8,06	12,37	9,22	14,32	15,81		10,63	18,25
9:I	5,83	6,32	9,06	5,66	6,32	5,10	7,28	10,63		10,77
10:J	13,04	8,25	19,65	6,00	17,09	5,83	3,61	18,25	10,77	

Fonte: Autor.

No passo 3 do algoritmo, deve-se estabelecer um critério de corte para a distância. Neste exemplo, definiu-se um limite de 40% de afastamento para que dois elementos sejam considerados semelhantes. Dessa forma, são estabelecidos os seguintes *clusters*:

- *Cluster 1*: A, B
- *Cluster 2*: B, D, A
- *Cluster 3*: C, E
- *Cluster 4*: D, B, F, G
- *Cluster 5*: E, C, I
- *Cluster 6*: F, G, D, I
- *Cluster 7*: G, F, D, J, B
- *Cluster 8*: H, A
- *Cluster 9*: I
- *Cluster 10*: J, G, F, D

Em função do critério adotado para a definição do grau de semelhança entre os elementos, nota-se que o conjunto I não possui elementos assemelhados.

Quadro 2.13 – Tabelas de similaridade para os conjuntos.

	Distância	%
A	0,00	0,0%
B	5,10	39,1%
H	5,39	41,3%
I	5,83	44,7%
D	7,07	54,2%
E	7,62	58,4%
C	8,49	65,1%
F	8,94	68,6%
G	10,44	80,1%
J	13,04	100,0%

(a) Conjunto A

	Distância	%
B	0,00	0,0%
D	2,83	21,7%
A	5,10	39,1%
F	5,83	44,7%
I	6,32	48,5%
G	6,40	49,1%
J	8,25	63,2%
H	10,05	77,1%
E	11,31	86,8%
C	13,04	100,0%

(b) Conjunto B

	Distância	%
C	0,00	0,0%
E	3,16	16,1%
H	8,06	41,0%
A	8,49	43,2%
I	9,06	46,1%
B	13,04	66,4%
D	13,93	70,9%
F	14,14	72,0%
G	16,28	82,9%
J	19,65	100,0%

(c) Conjunto C

	Distância	%
D	0,00	0,0%
B	2,83	20,3%
F	3,16	22,7%
G	3,61	25,9%
I	5,66	40,6%
J	6,00	43,1%
A	7,07	50,8%
E	11,66	83,7%
H	12,37	88,8%
C	13,93	100,0%

(d) Conjunto D

	Distância	%
E	0,00	0,0%
C	3,16	18,5%
I	6,32	37,0%
A	7,62	44,6%
H	9,22	54,0%
B	11,31	66,2%
F	11,40	66,7%
D	11,66	68,2%
G	13,60	79,6%
J	17,09	100,0%

(e) Conjunto E

	Distância	%
F	0,00	0,0%
G	2,24	15,6%
D	3,16	22,1%
I	5,10	35,6%
B	5,83	40,7%
J	5,83	40,7%
A	8,94	62,5%
E	11,40	79,6%
C	14,14	98,8%
H	14,32	100,0%

(f) Conjunto F

	Distância	%
G	0,00	0,0%
F	2,24	13,7%
D	3,61	22,1%
J	3,61	22,1%
B	6,40	39,3%
I	7,28	44,7%
A	10,44	64,1%
E	13,60	83,6%
H	15,81	97,1%
C	16,28	100,0%

(g) Conjunto G

	Distância	%
H	0,00	0,0%
A	5,39	29,5%
C	8,06	44,2%
E	9,22	50,5%
B	10,05	55,1%
I	10,63	58,3%
D	12,37	67,8%
F	14,32	78,5%
G	15,81	86,6%
J	18,25	100,0%

(h) Conjunto H

	Distância	%
I	0,00	0,0%
F	5,10	47,3%
D	5,66	52,5%
A	5,83	54,1%
B	6,32	58,7%
E	6,32	58,7%
G	7,28	67,6%
C	9,06	84,1%
H	10,63	98,7%
J	10,77	100,0%

(i) Conjunto I

	Distância	%		Distância	%
J	0,00	0,0%	I	10,77	54,8%
G	3,61	18,4%	A	13,04	66,4%
F	5,83	29,7%	E	17,09	87,0%
D	6,00	30,5%	H	18,25	92,9%
B	8,25	42,0%	C	19,65	100,0%

(j) Conjunto J

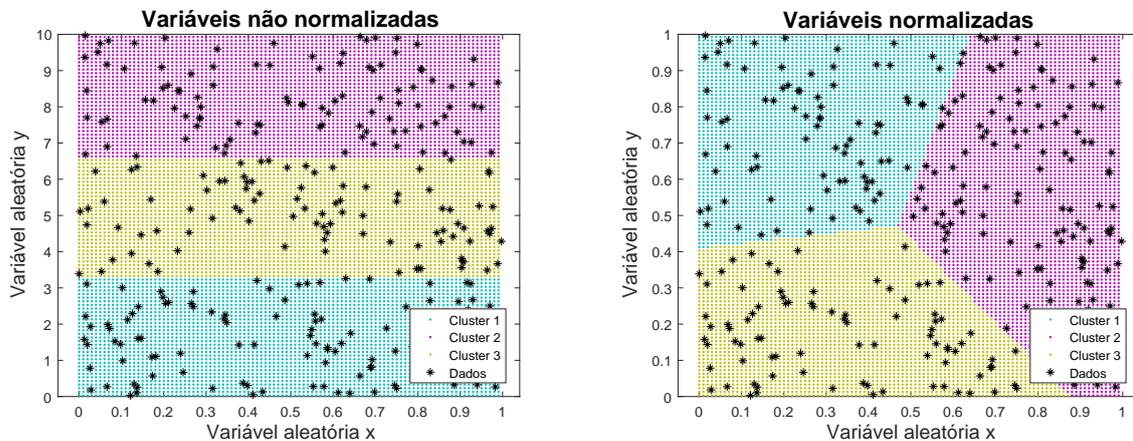
Uma das etapas de um processo de *clusterização* consiste na normalização das variáveis a serem consideradas. O tratamento dos dados é importante para que os métodos não produzam informações incorretas ou tendenciosas.

2.3.4 Normalização dos Dados

Os exemplos de aplicação dos métodos apresentados levaram em consideração as informações contidas no Quadro 2.6. Apesar de se tratarem de dados da mesma ordem de grandeza, os atributos x e y apresentam amplitudes diferentes. Essa diferença de escala influencia diretamente os resultados obtidos na *clusterização*. Uma variável com escala maior do que as demais pode ser interpretada como um atributo que possui um peso maior na classificação dos dados.

A Figura 2.10 apresenta um exemplo de *clusterização* a dois conjuntos de dados distintos, ambos contendo 300 elementos descritos por duas variáveis. Na Figura 2.10(a) a variável x assume valores entre 0 e 1, enquanto que a variável y assume valores entre 0 e 10. Na Figura 2.10(b) os valores de ambas as variáveis estão compreendidos entre 0 e 1.

Figura 2.10 – Exemplo da normalização de variáveis.



(a) Variáveis não normalizadas

(b) Variáveis normalizadas

Fonte: Autor.

Percebe-se que, para o caso da Figura 2.10(a), a estratificação dos dados ocorreu de forma vertical, priorizando a distribuição no eixo y . A normalização dos atributos utilizados na análise visa contornar esse problema. Uma das formas possíveis de se normalizar uma variável consiste em dividir os valores de cada atributo i pelo seu valor máximo, conforme a Equação 2.17.

$$A_i(j) = \frac{A_i(j)}{A_{max}(j)} \quad (2.17)$$

Onde:

$A_i(j)$ = Valor da observação j do atributo i considerado;

$A_{max}(j)$ = Maior valor do atributo i considerado dentre todos os elementos;

Outra maneira de se normalizar um atributo consiste em subtrair cada valor do atributo i da média deste atributo e dividir o resultado pelo desvio padrão do atributo, de acordo com a Equação 2.18.

$$A_i(j) = \frac{A_i(j) - \bar{\chi}(A_i)}{\sigma(A_i)} \quad (2.18)$$

Onde:

$A_i(j)$ = Valor da observação j do atributo i considerado;

$\bar{\chi}(A_i)$ = Média dos valores do atributo i considerado;

$\sigma(A_i)$ = Desvio padrão dos valores do atributo i considerado;

Na literatura, encontram-se outros métodos de normalização de variáveis. Contudo as duas formas citadas são as mais comuns e aplicam-se a maioria dos casos.

2.4 CONSIDERAÇÕES SOBRE O CAPÍTULO

Neste capítulo apresentou-se um breve histórico da evolução da metodologia de definição dos limites de DEC e FEC, que se baseia em uma análise comparativa entre as distribuidoras e utiliza-se de técnicas de agrupamento. A aplicação desse tipo de metodologia em segmentos considerados monopólios naturais permite a emulação de um ambiente concorrencial que objetiva a melhoria da qualidade do serviço das empresas. Dentre as principais mudanças observadas no cálculo dos limites, ao longo do tempo, destacam-se as alterações dos atributos que descrevem os conjuntos e dos métodos de *clusterização* utilizados.

Os três métodos apresentados neste capítulo, que foram utilizados ou propostos pela ANEEL, possuem vantagens e desvantagens. Contudo o método dinâmico é vantajoso frente aos demais devido à sua simplicidade, tendo em vista que sua classificação baseia-se apenas no grau de proximidade dos elementos. Esse método pode ser aplicado a grandes conjuntos de dados e é capaz de solucionar o problema dos *outliers*.

Por tudo isso, verifica-se que os métodos empregados na metodologia da ANEEL não se constituem necessariamente como os melhores ou mais otimizados para os fins a que se destinam. As constantes revisões do método de definição dos limites dos indicadores coletivos de continuidade denotam a inexistência de uma forma perfeitamente justa de estimular a evolução do desempenho das distribuidoras.

3 METODOLOGIA PROPOSTA

O modelo atual de cálculo dos limites de continuidade envolve etapas como a seleção de variáveis, o agrupamento dos conjuntos e a definição das metas. Ao longo de cada etapa, são feitas considerações no sentido de se dar soluções às adversidades que surgem com a aplicação do método. Dessa forma, trata-se de um método complexo.

Devido a essa complexidade, neste trabalho objetiva-se analisar a etapa de agrupamento de conjuntos de UC, por meio de modificações do método de *clusterização* empregado, alterações dos atributos selecionados e variação dos parâmetros de entrada do método dinâmico. Esta metodologia se propõe a reproduzir aquela empregada pela ANEEL para o cálculo dos limites de DEC e FEC.

Almeja-se avaliar como tais mudanças refletem-se nos limites objetivos de cada conjunto. Para se alcançar os objetivos propostos, apresenta-se neste capítulo uma metodologia composta por três etapas, cuja estrutura é exposta na Figura 3.1.

Na primeira etapa, são obtidos os dados de entrada necessários para o cálculo dos limites. Em seguida, na segunda etapa, realizam-se os agrupamentos dos conjuntos empregando-se as técnicas e parâmetros escolhidos. Na última etapa, são definidos os limites por conjunto e são analisados os resultados. A metodologia descrita é aplicada em estudos de caso, apresentados no Capítulo 4, entretanto sua aplicação não se limita aos casos propostos.

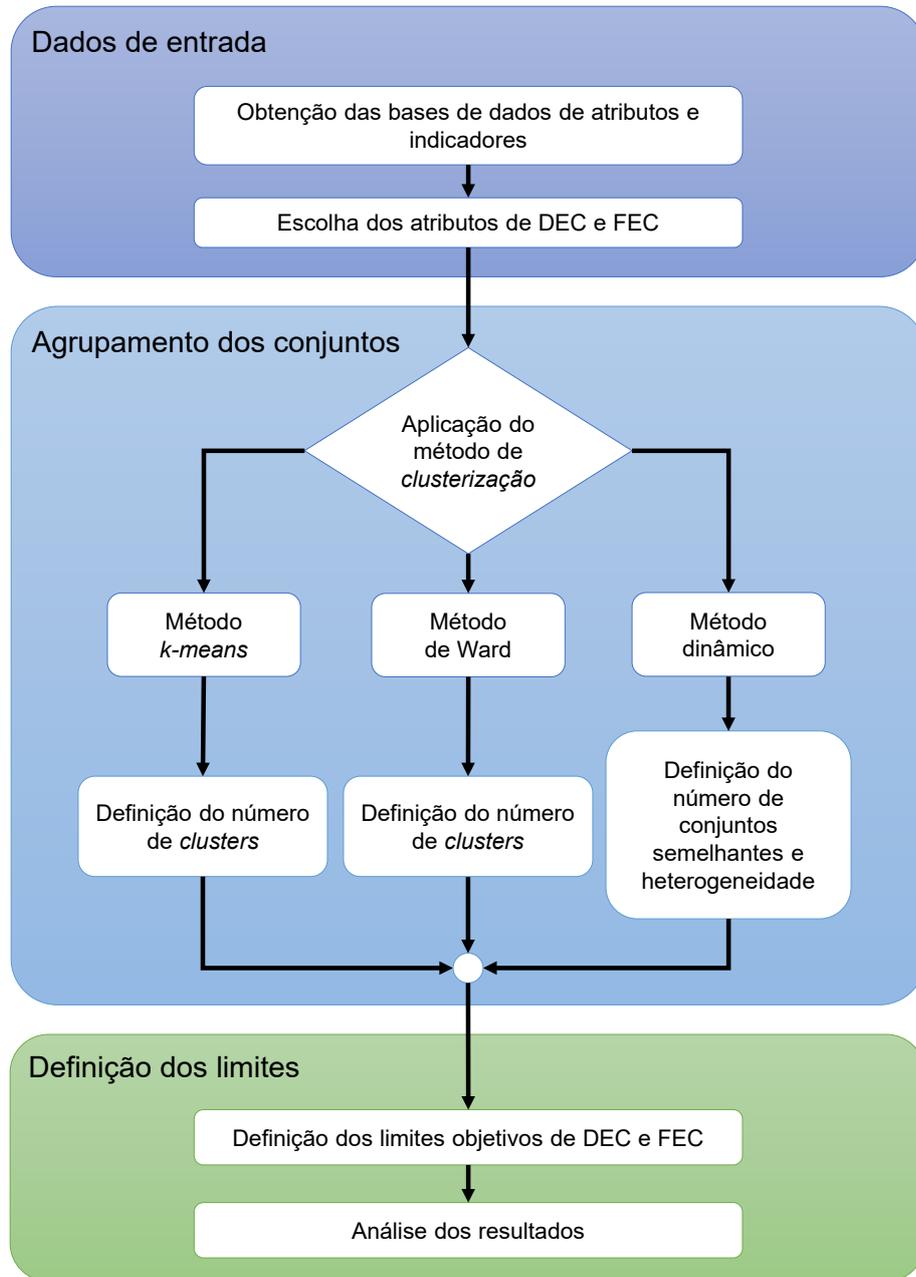
3.1 DADOS DE ENTRADA

A primeira etapa, correspondente aos dados de entrada, consiste na obtenção das bases de dados dos indicadores e dos atributos dos conjuntos. Optou-se por utilizar uma base única, pois o uso de diferentes bases de dados de entrada poderia interferir no processo de comparação dos resultados. As informações são extraídas do Apêndice III da Nota Técnica nº 0102/2014-SRD/ANEEL, que contém os valores de 146 atributos e dois indicadores (DEC e FEC) de 2610 conjuntos de distribuidoras brasileiras (ANEEL, 2014b).

Em seguida dá-se a escolha dos atributos descritores dos conjuntos, que poderão ser os mesmos empregados atualmente na metodologia de análise comparativa, conforme o Quadro 2.2, ou outros, dentre os 146 atributos da base de dados. Por fim, são removidas as eventuais informações faltantes ou incompletas, de modo que sejam considerados apenas os conjuntos cujos dados estejam corretos.

A etapa seguinte consiste no agrupamento dos conjuntos.

Figura 3.1 – Estrutura da metodologia proposta de agrupamento de conjuntos.



Fonte: Autor.

3.2 AGRUPAMENTO DOS CONJUNTOS

A etapa seguinte consiste na aplicação de um dos métodos de *clusterização* apresentados no Capítulo 2: *k-means*, Ward ou dinâmico. Devido à grande quantidade de operações matemáticas envolvidas nesta etapa, desenvolveu-se uma rotina (*script*), utilizando-se o *software* MATLAB, que contempla desde a importação dos indicadores e atributos até a exportação dos limites objetivos por conjunto para uma planilha.

O *script* foi validado utilizando-se bases de atributos e de indicadores disponibilizadas pela ANEEL por meio da Plataforma Integrada de Ouvidoria e Acesso à Informação da Controladoria-Geral da União (CGU). Foi cedida, ainda, a memória de cálculo dos limites da CEEE-D para o 5º Ciclo de Revisões Tarifárias Periódicas (5CRTP), o que possibilitou a comparação dos resultados.

Inicialmente o código realiza a importação de uma tabela contendo os dados dos atributos, os quais são padronizados pelo escore padrão (*Z-Score*), de acordo com a Equação 2.6. Em seguida calcula-se, para cada par de conjuntos, a respectiva distância euclidiana empregando-se a Equação 2.7. Então é feito o agrupamento dos conjuntos. Ressalta-se que são separados os conjuntos aéreos dos subterrâneos na etapa de comparação, ou seja, conjuntos aéreos só são comparados com conjuntos aéreos, e conjuntos subterrâneos só se comparam com subterrâneos.

O agrupamento é feito utilizando um dos três métodos: *k-means*, Ward ou dinâmico. Para cada método escolhido, adotam-se procedimentos distintos. A fim de se aplicar o método *k-means*, define-se inicialmente o número de *clusters* a serem formados *k*. O agrupamento é feito utilizando-se a função *k-means* contida no pacote *Statistics and Machine Learning Toolbox* do software MATLAB.

De modo semelhante ao método *k-means*, a aplicação do método de Ward requer o estabelecimento de critério de corte para o nível de agregação desejado. Neste caso, como a análise envolve uma grande quantidade de elementos, definir-se-á o número de *clusters* *k*. Empregam-se as funções *linkage* e *cluster* contidas no pacote *Statistics and Machine Learning Toolbox* do software MATLAB.

No caso de escolha do método dinâmico, devem ser definidos três parâmetros, conforme a metodologia de análise comparativa vigente: a heterogeneidade percentual permitida (h_{max}), o número desejável de conjuntos semelhantes (n_{max}) e o número mínimo de conjuntos semelhantes (n_{min}). Para cada um dos conjuntos que se deseja calcular os limites, selecionam-se os conjuntos mais semelhantes ao de interesse respeitando-se os parâmetros citados.

A heterogeneidade percentual de cada agrupamento é calculada pela Equação 2.8. Caso o valor de heterogeneidade de um agrupamento supere o valor máximo permitido (h_{max}), classifica-se o conjunto em análise como heterogêneo. Conjuntos heterogêneos recebem um tratamento particular na etapa de definição dos limites, em que ocorre uma variação nos percentis. Adota-se o procedimento descrito na Seção 2.2.4.3.

A última etapa da metodologia proposta consiste na definição dos limites e é descrita na próxima seção.

3.3 DEFINIÇÃO DOS LIMITES

Na última etapa, definem-se os limites objetivos dos indicadores DEC e FEC para os conjuntos de UC a partir dos agrupamentos. Ordenam-se os conjuntos contidos em cada *cluster* de acordo com indicador apurado. Utiliza-se uma média de três anos para cada indicador.

Posteriormente define-se o conjunto de referência de acordo com a Equação 2.9. Aplica-se o percentil 20 para conjuntos interligados aéreos, e o percentil 50 para conjuntos subterrâneos e conjuntos aéreos isolados. O número de conjuntos semelhantes ($N_{ConjSeme}$) trata-se do número de conjuntos contidos no agrupamento analisado.

Os limites objetivos de DEC e FEC do conjunto de interesse são definidos pelo próprio indicador do conjunto de referência. Não será avaliado neste trabalho a trajetória dos limites, mas apenas os limites objetivos por conjunto e o limite objetivo global. Um limite global é calculado pela média dos limites de cada um dos conjuntos ponderada pelos respectivos números de unidades consumidoras.

Por fim, a análise dos resultados envolverá a comparação dos limites calculados em estudos de caso. Pretende-se verificar as variações observadas nos limites globais, que indicam se determinado cenário é mais ou menos favorável para a distribuidora do que outro. Observa-se também o comportamento dos limites por conjunto, por meio da contagem da quantidade de conjuntos que tiveram seus limites aumentados, reduzidos ou mantidos em relação a um cenário base.

3.4 CONSIDERAÇÕES SOBRE O CAPÍTULO

Neste capítulo foi descrita a metodologia proposta para o cálculo dos limites de DEC e FEC. Trata-se de uma reprodução do modelo empregado pela ANEEL, mas com a possibilidade de que sejam testados outros atributos, métodos de agrupamento e parâmetros. Desse modo, a metodologia permite que um grande número de combinações seja testado e avaliado. Optou-se, entretanto, por utilizar uma única base de dados de atributos e de indicadores para que os limites resultantes de diferentes estudos de caso possam ser comparados entre si.

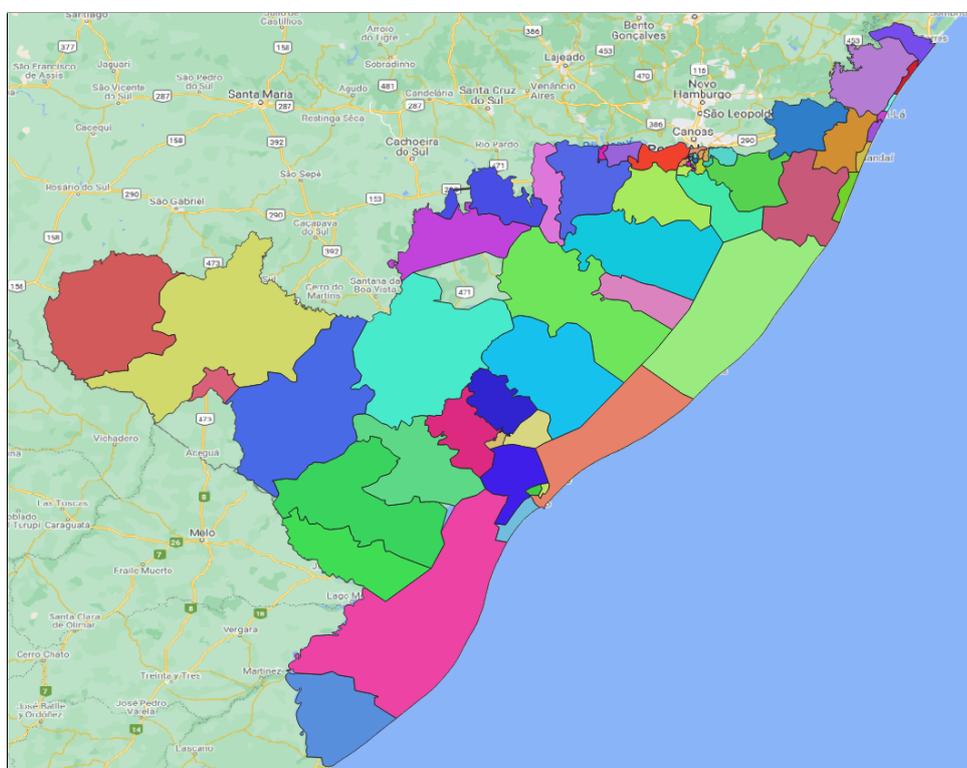
Em cada cenário proposto, calculam-se os limites objetivos por conjunto, além dos limites objetivos de DEC e de FEC globais. Limite objetivo trata-se da meta de desempenho a ser alcançada pelo conjunto num horizonte de oito anos.

A etapa seguinte deste trabalho consiste na proposição de estudos de caso, em que são feitas variações nos atributos, métodos e parâmetros, para se observar o comportamentos dos limites de DEC e FEC calculados.

4 ESTUDOS DE CASO E RESULTADOS

O presente capítulo tem por objetivo apresentar a aplicação da metodologia proposta, descrita no Capítulo 3, em estudos de caso. Para avaliar os resultados obtidos, escolheu-se analisar os conjuntos da Companhia Estadual de Distribuição de Energia Elétrica (CEEE-D), cuja área de atuação e cujos conjuntos elétricos são mostrados na Figura 4.1. Cabe ressaltar que a metodologia pode ser igualmente aplicada a qualquer outra distribuidora do Brasil para o cálculo dos limites.

Figura 4.1 – Áreas dos conjuntos da distribuidora CEEE-D em 2013.



Fonte: Autor.

No ano de 2013, compunham a área de atuação da CEEE-D 61 conjuntos elétricos. Os nomes dos conjuntos e os respectivos DEC e FEC médios apurados nos anos de 2011 a 2013 são apresentados no Quadro 4.1.

Para a aplicação da metodologia, as bases de dados dos indicadores e dos atributos dos conjuntos, que se referem aos dados de entrada, foram obtidas do Apêndice III da Nota Técnica nº 0102/2014-SRD/ANEEL (ANEEL, 2014b). Optou-se por utilizar essa base, ainda que contenha dados antigos, por possuir dados de 146 atributos distintos. Assim podem ser realizados estudos de caso com diferentes combinações de atributos. Foram realizados ao todo sete cenários, nos quais foram variados os atributos selecionados para DEC e FEC, o método de *clusterização* empregado e seus parâmetros.

Quadro 4.1 – Indicadores DEC e FEC médios apurados nos anos de 2011 a 2013.

Conjunto	DEC	FEC	Conjunto	DEC	FEC
Alvorada	24,60	16,15	Porto Alegre 1	7,30	8,56
Areal	35,92	17,27	Porto Alegre 10	8,73	6,22
Arroio do Sal	11,17	9,00	Porto Alegre 12	9,68	11,53
Arroio Grande	31,95	17,30	Porto Alegre 13	10,90	10,78
Atlântida	10,35	8,97	Porto Alegre 14	24,17	18,70
Atlântida Sul	15,15	11,21	Porto Alegre 16	13,09	9,16
Bagé 1	8,57	7,87	Porto Alegre 18	5,93	6,15
Bagé 2	16,20	7,59	Porto Alegre 2	7,41	8,14
Camaquã	34,67	16,66	Porto Alegre 3	12,43	11,00
Camaquã 2	22,20	11,98	Porto Alegre 4 - Centro	2,77	2,48
Canguçu	62,67	27,49	Porto Alegre 4 - Centro 2	5,53	5,57
Capão Novo	13,15	11,34	Porto Alegre 5	5,13	5,03
Capivarita	22,88	15,96	Porto Alegre 6	13,86	11,26
Cassino	25,75	14,83	Porto Alegre 8	10,95	8,38
Charqueadas 2	15,39	12,23	Porto Alegre 9	10,58	8,40
CRM 1 - Minas do Leão	24,84	13,44	Quinta	33,99	21,85
Dom Pedrito	16,59	8,35	Rio Grande 1	17,81	12,47
Eldorado do Sul	30,06	18,45	Rio Grande 2	10,68	9,23
Encruzilhada	17,34	8,10	Rio Grande 3	32,50	18,37
Guaíba	29,44	23,05	Santa Vitória do Palmar	47,30	29,35
Jaguarão	27,99	20,00	Santo Antônio	22,96	16,08
Marmeleiro	86,03	28,74	São Lourenço do Sul	53,19	23,67
Mostardas	59,76	37,27	Terra de Areia	28,07	18,35
Osório 1	9,24	6,74	Torres 1	18,37	10,27
Palmares do Sul	29,16	19,33	Tramandaí	16,33	13,56
Pedro Osório	45,30	22,07	UTE Presidente Médici	39,13	15,13
Pelotas 1	15,33	16,34	UTE São Jerônimo	15,21	15,27
Pelotas 2	13,37	14,31	Vasconcelos	40,59	18,79
Pelotas 3	39,16	23,23	Viamão 1	39,60	30,71
Pelotas 4	53,01	29,41	Viamão 2	24,64	17,25
Pinhal	23,73	17,09			

Fonte: Autor.

O primeiro cenário, apresentado na Seção 4.1, consiste na elaboração de um caso base, que visa replicar a metodologia de definição de limites utilizada pela ANEEL. Esse Cenário 1 é utilizado para comparação com os demais estudos de caso. Nos cenários 2 e 3, modifica-se o método de agrupamento para *k-means* e Ward respectivamente, mantendo-se os mesmos atributos utilizados pela ANEEL. Nos cenários 4, 5 e 6, é empregado o método dinâmico, porém se alteram os atributos selecionados. Por fim, no cenário 7, são feitas alterações nos parâmetros do método dinâmico.

4.1 CENÁRIO 1

O primeiro cenário apresentado neste trabalho objetiva replicar a metodologia de análise comparativa de desempenho utilizada pela ANEEL para o cálculo dos limites. Os resultados obtidos deste cenário servirão de base para comparação com os demais cenários propostos. O Quadro 4.2 resume os parâmetros utilizados no Cenário 1.

Quadro 4.2 – Parâmetros do Cenário 1.

Cenário 1	
Atributos DEC	PC_NUC_AD, PC_VRAM, PC_ERMT_3F, PLUV, CM_NUC_RES, NUC_IND
Atributos FEC	PC_VRAM, PLUV, PC_NUC_AD, NUC_COM, CM_NUC_RES, PC_ERMT_3F
Método de <i>Clusterização</i>	Dinâmico
Parâmetros	$n_{max} = 100, n_{min} = 50, h_{max} = 20\%$

Fonte: Autor.

Empregaram-se os mesmos atributos listados no Quadro 2.2, por terem sido selecionados como os que melhor descrevem DEC e FEC de acordo com a ANEEL (2014b). O método de *clusterização* utilizado é o dinâmico, cujos parâmetros estabelecidos são: número desejável de conjuntos semelhantes igual a 100; número mínimo de conjuntos semelhantes igual a 50 e heterogeneidade percentual máxima igual a 20%. Tais parâmetros são empregados na definição dos limites pela ANEEL (2016).

Definidos os atributos e os parâmetros do método dinâmico, inicia-se a etapa de agrupamento dos conjuntos. Para cada um dos conjuntos que se deseja calcular os limites, deve-se encontrar os respectivos conjuntos mais semelhantes. O Quadro 4.3 apresenta uma amostra dos 100 conjuntos mais semelhantes ao conjunto Alvorada para o DEC. Os conjuntos estão listados da menor para a maior distância ao conjunto de interesse.

Quadro 4.3 – Conjuntos semelhantes ao conjunto em análise.

Nº	Conjunto Semelhante	Heterogeneidade (%)	DEC Apurado (horas)
1	Alvorada	0,00	24,60
2	Alto Lage	5,36	5,02
3	Cachoeirinha 2	6,09	10,06
4	Pelotas 1	6,33	15,33
...
97	Itupeva	10,84	8,80
98	Marília 2-Itambé	10,87	4,13
99	Monumento	10,95	5,35
100	Eldorado do Sul	10,96	30,06

Fonte: Autor.

Verifica-se que, para 100 conjuntos semelhantes, a heterogeneidade percentual não excedeu o critério de $h_{max} = 20\%$, portanto 100 é a quantidade de conjuntos com os quais Alvorada será comparado. Ainda, pode-se afirmar que Alvorada não é um conjunto heterogêneo. Em seguida, os 100 conjuntos semelhantes são ordenados em função dos respectivos indicadores apurados, conforme o Quadro 4.4.

Quadro 4.4 – Conjuntos semelhantes ordenados pelo DEC apurado.

Nº	Conjunto Semelhante	Heterogeneidade (%)	DEC Apurado (horas)
1	Limeira Um	9,31	2,70
2	Roland Davis	7,72	2,83
3	Araras Um	8,23	3,22
...
19	Franca 4-Guanabara	8,69	5,23
20	Votuporanga	9,45	5,24
21	Santa Bárbara 3-Esmeralda	10,47	5,30
...
98	Ceilândia Sul	10,76	26,72
99	Eldorado do Sul	10,96	30,06
100	Viamão 1	7,00	39,60

Fonte: Autor.

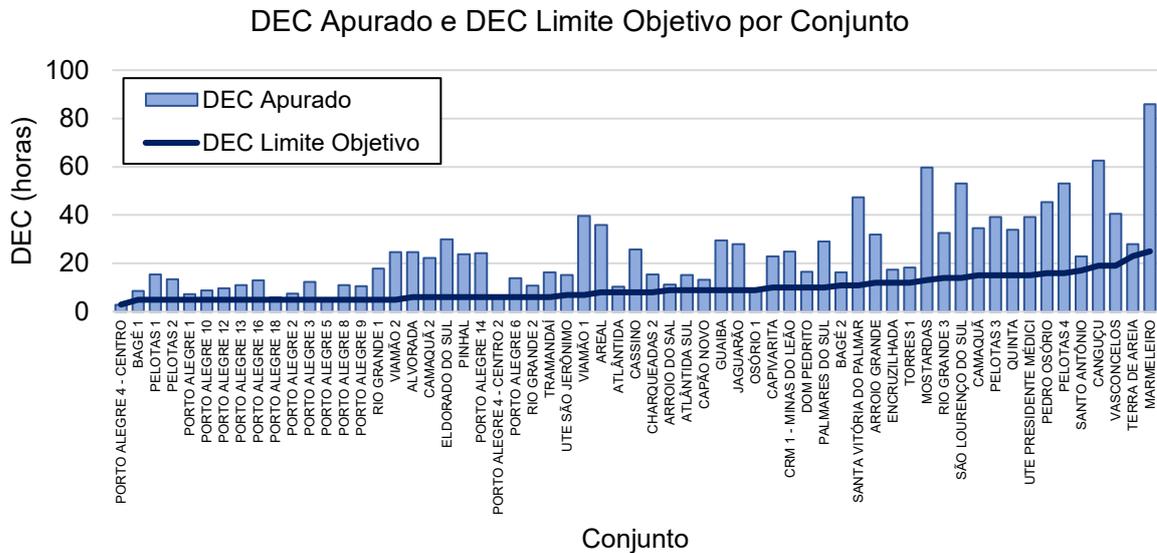
Sabendo-se que o número de conjuntos semelhantes ($N_{ConjSeme}$) é igual a 100 e que Alvorada é um conjunto com rede aérea e interligada ao SIN, encontra-se que a posição do conjunto de referência é igual a 20, de acordo com a Equação 2.9. Desse modo, a referência para o DEC de Alvorada é o conjunto *Votuporanga*. O limite objetivo é obtido do arredondamento para cima do DEC apurado pelo conjunto de referência, ou seja, neste caso o limite é de 6 horas para o DEC.

O procedimento apresentado para o conjunto Alvorada é repetido para o cálculo dos limites de DEC e de FEC dos demais conjuntos da CEEE-D. Como o objetivo deste estudo é comparar os limites objetivos, não se avaliam as interrupções do suprimento nem as trajetórias de redução intensas.

O Apêndice A apresenta os limites calculados para os 61 conjuntos da CEEE-D. Os valores de DEC limite podem ser comparados aos indicadores apurados no gráfico da Figura 4.2. Observa-se que, para 59 dos 61 conjuntos, os limites objetivos são menores do que os DEC apurados.

O DEC limite objetivo global deste cenário é de 7,80 horas, enquanto que o DEC apurado global é de 20,06 horas. Já o FEC limite objetivo global calculado é de 5,48 interrupções, enquanto que o FEC apurado global é de 13,99 interrupções. A diferença entre indicadores apurados e os limites calculados não expressa uma violação dos limites, mas sim que, após um período de oito anos, caso os indicadores se repetissem, haveria transgressões.

Figura 4.2 – DEC Apurado e DEC Limite Objetivo por conjunto.



Fonte: Autor.

Da análise da Figura 4.2, percebe-se que os desempenhos da maior parte dos conjuntos superam os limites calculados. Em alguns casos, o valor apurado corresponde a duas vezes ou mais o limite objetivo. Dessa forma, quaisquer elevações dos limites podem ser favoráveis à distribuidora, pois, possivelmente, haveria redução das compensações financeiras a serem pagas.

O próximo cenário apresentado propõe a mudança do método de *clusterização*.

4.2 CENÁRIO 2

No segundo cenário deste trabalho, utilizam-se os mesmos atributos do Cenário 1, porém os conjuntos são agrupados pelo método *k-means*. Os principais parâmetros do Cenário 2 estão elencados no Quadro 4.5.

Quadro 4.5 – Parâmetros do Cenário 2.

Cenário 2	
Atributos DEC	PC_NUC_AD, PC_VRAM, PC_ERMT_3F, PLUV, CM_NUC_RES, NUC_IND
Atributos FEC	PC_VRAM, PLUV, PC_NUC_AD, NUC_COM, CM_NUC_RES, PC_ERMT_3F
Método de <i>Clusterização</i>	<i>k-means</i>
Parâmetros	$k = 30$

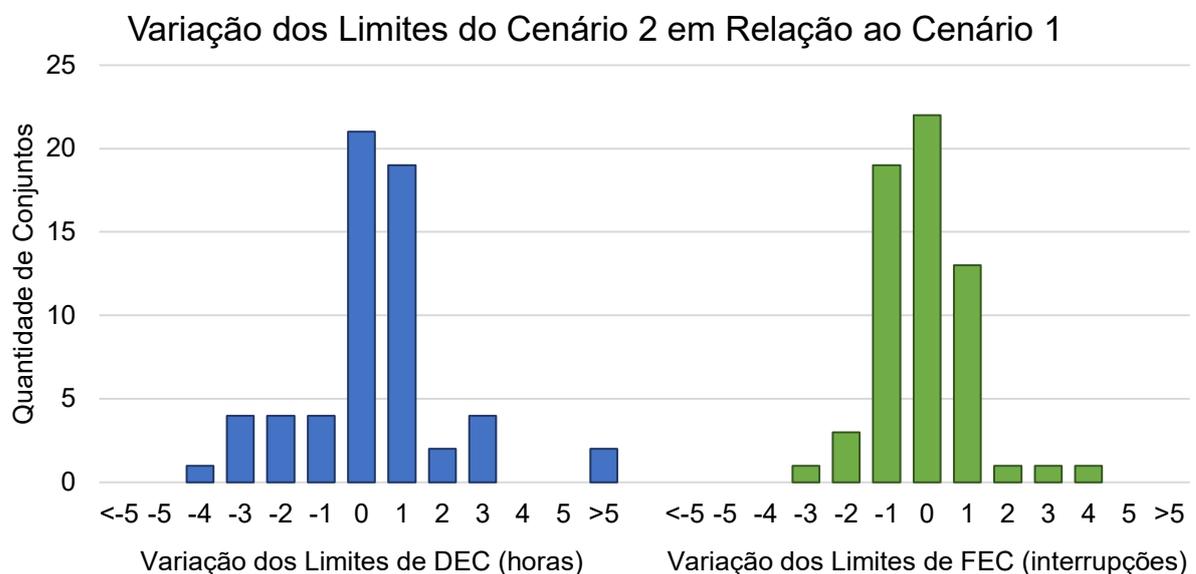
Fonte: Autor.

Aplicou-se o método *k-means* com o número de *clusters* (k) igual a 30, dos quais

29 são para conjuntos aéreos e um para os conjuntos subterrâneos. Esse número foi definido por ter sido utilizado pela ANEEL (2009) em sua metodologia de definição de limites. Foram avaliados outros valores para k , porém não houve alterações significativas nos limites calculados. Os valores de DEC e FEC limites estão listados no Apêndice A.

A Figura 4.3 mostra os histogramas de frequência para a variação limites de DEC e de FEC, calculados no Cenário 2, em relação aos limites do Cenário 1. Este gráfico apresenta os aumentos, as reduções e as manutenções dos limites deste cenário frente aos do cenário base. As variações positivas indicam aumento nos limites, favoráveis à distribuidora, enquanto as variações negativas indicam redução dos limites por conjunto, favoráveis aos consumidores.

Figura 4.3 – Histograma de variação dos limites objetivos do Cenário 2.



Fonte: Autor.

O DEC limite objetivo global para este cenário é de 8,77 horas, enquanto que, no Cenário 1, o limite é de 7,80 horas. Em se tratando do FEC, o limite objetivo global é de 5,17 interrupções, enquanto que, no Cenário 1, é de 5,48 interrupções. Ou seja, observa-se que, para o DEC, em geral, os limites sofreram um aumento, ao passo que, para o FEC, houve uma redução dos limites em relação ao Cenário 1.

No próximo cenário, propõe-se a utilização do método de Ward.

4.3 CENÁRIO 3

No terceiro cenário apresentado, utilizam-se os mesmos atributos do Cenário 1, porém os conjuntos são agrupados pelo método de Ward. Os principais parâmetros do Cenário 3 estão listados no Quadro 4.6.

O emprego do método de Ward exige a definição de um critério para a formação dos agrupamentos, que pode ser pelo nível de agregação ou pelo próprio número desejado de *clusters*. Devido à dificuldade de se estabelecer um nível de agregação máximo, optou-se pela simples escolha do número total de *clusters*. De forma análoga ao cenário anterior, foram formados 30 agrupamentos, sendo 29 deles para conjuntos aéreos e um para os conjuntos subterrâneos.

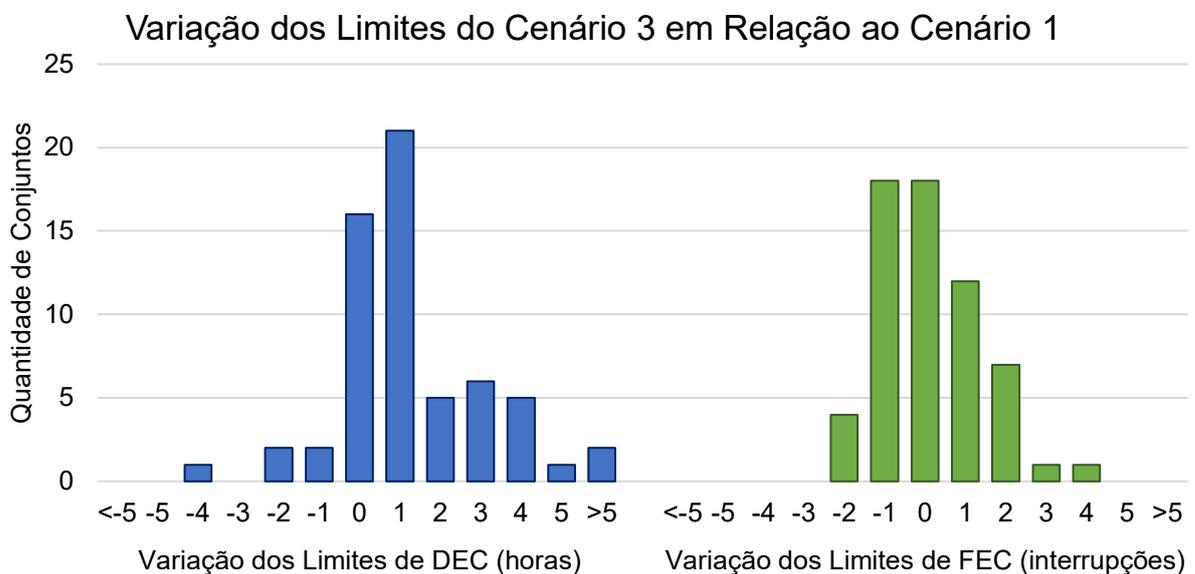
Quadro 4.6 – Parâmetros do Cenário 3.

Cenário 3	
Atributos DEC	PC_NUC_AD, PC_VRAM, PC_ERMT_3F, PLUV, CM_NUC_RES, NUC_IND
Atributos FEC	PC_VRAM, PLUV, PC_NUC_AD, NUC_COM, CM_NUC_RES, PC_ERMT_3F
Método de Clusterização	Ward
Parâmetros	$k = 30$

Fonte: Autor.

A Figura 4.4 ilustra os histogramas de frequência para a variação dos limites de DEC e de FEC, calculados para os conjuntos da CEEE-D.

Figura 4.4 – Histograma de variação dos limites objetivos do Cenário 3.



Fonte: Autor.

Neste cenário, o DEC limite objetivo global é de 8,79 horas e o FEC limite objetivo global é de 5,58 interrupções. No Cenário 1, os limites foram de 7,80 horas para o DEC e 5,48 interrupções para o FEC. Comparando-se ao cenário base, verifica-se um aumento em ambos os limites globais no Cenário 3. Os valores de DEC e FEC limites por conjunto estão listados no Apêndice A.

No cenário a seguir, são empregados atributos diferentes daqueles utilizados na metodologia de análise comparativa da ANEEL, definidos na Nota Técnica nº 0102/2014-SRD/ANEEL.

4.4 CENÁRIO 4

Neste quarto cenário, propõe-se a substituição dos atributos em relação àqueles utilizados atualmente pela ANEEL. O Quadro 4.7 apresenta os parâmetros aplicados no Cenário 4.

Quadro 4.7 – Parâmetros do Cenário 4.

Cenário 4	
Atributos DEC	Correlação mínima: 40%, VIF: 5,0 VRAM, ENE_COM, ERMT_1F, PAREDE_IN
Atributos FEC	Correlação mínima: 40%, VIF: 5,0 PC_ERMT_1F, ERMT_AREA, PAREDE_IN
Método de <i>Clusterização</i>	Dinâmico
Parâmetros	$n_{max} = 100, n_{min} = 50, h_{max} = 20\%$

Fonte: Autor.

Os atributos escolhidos, neste cenário, para comporem a caracterização dos atributos DEC e FEC foram obtidos por Santos (2021), que avaliou a etapa de seleção dos atributos da metodologia comparativa de desempenho atual. Tais atributos foram encontrados a partir da mudança do coeficiente de correlação mínimo (Pearson ou Spearman) de 20% para 40%, em módulo. Ambos os modelos, de DEC e de FEC, consideram também variáveis socioeconômicas.

Para o DEC, são considerados os seguintes atributos:

- VRAM - área com vegetação remanescente alta ou média (km²);
- ENE_COM - energia anual da classe comercial (MWh);
- ERMT_1F - extensão da rede MT monofásica (km) e
- PAREDE_IN - percentual de pessoas em domicílios com paredes inadequadas (%).

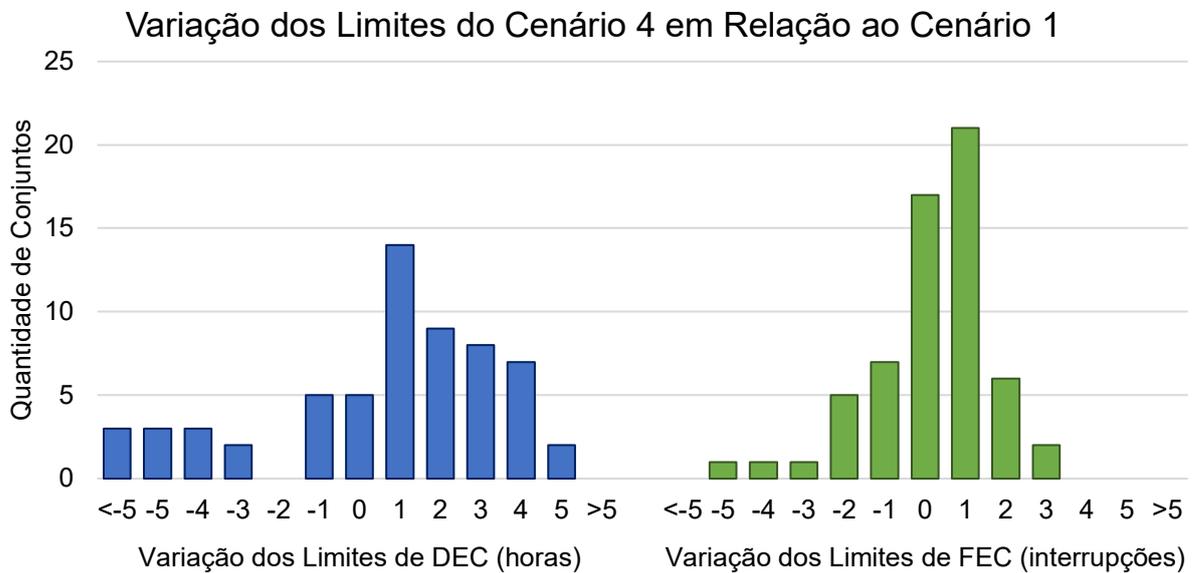
Para o FEC, são considerados os seguintes atributos:

- PC_ERMT_1F - extensão da rede MT monofásica (km);
- ERMT_AREA - densidade de redes MT (km/km²);
- PAREDE_IN - percentual de pessoas em domicílios com paredes inadequadas (%).

Foi aplicado o método de *clusterização* dinâmico, com os mesmos parâmetros do Cenário 1: $n_{max} = 100, n_{min} = 50, h_{max} = 20\%$. O DEC limite objetivo global encontrado foi de 8,52 horas e o FEC limite objetivo global foi de 5,87 interrupções. Os valores de DEC e FEC limites por conjunto estão no Apêndice A. O histograma de variação dos limites,

apresentado na Figura 4.5, indica um aumento geral nos limites quando comparados aos do Cenário 1.

Figura 4.5 – Histograma de variação dos limites objetivos do Cenário 4.



Fonte: Autor.

Percebe-se, dos histogramas de variação dos limites da Figura 4.5, que a maior parte dos conjuntos tiveram seus limites alterados. Para o DEC, apenas 5 conjuntos tiveram seus limites mantidos em relação ao Cenário 1. Para 40 conjuntos houve aumento dos limites de DEC e, para 29, houve aumento dos limites de FEC.

4.5 CENÁRIO 5

Bem como no cenário anterior, neste foi avaliado o uso de novos atributos, mantendo-se o método de agrupamento dinâmico. Os principais parâmetros do Cenário 5 estão listados no Quadro 4.8.

Quadro 4.8 – Parâmetros do Cenário 5.

Cenário 5	
Atributos DEC	Correlação mínima: 20%, VIF: 3,0 PLUV, PC_NUC_URB, PC_NUC_COM, IDHM, LIXO
Atributos FEC	Correlação mínima: 20%, VIF: 3,0 PLUV, PC_VRAM, NUC_COM, PC_ERMT_3F, LIXO
Método de Clusterização	Dinâmico
Parâmetros	$n_{max} = 100$, $n_{min} = 50$, $h_{max} = 20\%$

Fonte: Autor.

Os atributos que compõem este cenário foram encontrados adotando-se um coeficiente de correlação mínimo de 20% e mudando-se o Fator de Inflação de Variância (VIF) aceitável de 5 para 3, na etapa de seleção dos atributos. Também foram consideradas variáveis socioeconômicas para os modelos de DEC e FEC (SANTOS, 2021).

Para o DEC, são considerados os seguintes atributos:

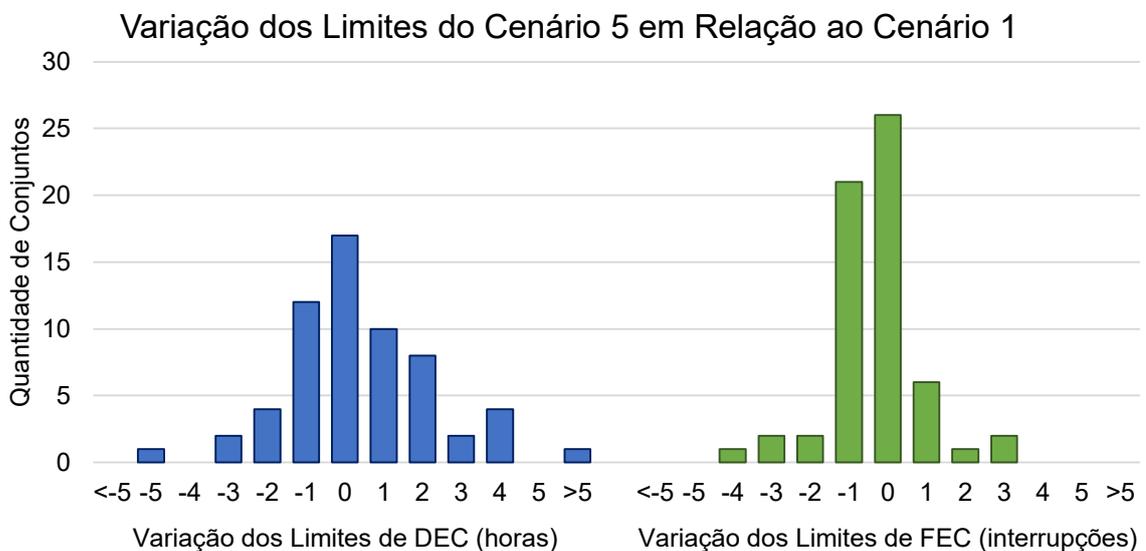
- PLUV - precipitação pluviométrica média anual (mm);
- PC_NUC_URB - percentual de NUC urbano (%);
- PC_NUC_COM - percentual de NUC comercial (%);
- IDHM - índice de desenvolvimento humano municipal e
- LIXO - percentual de domicílios com coleta de lixo (%).

Para o FEC, são considerados os seguintes atributos:

- PLUV - precipitação pluviométrica média anual (mm);
- PC_VRAM - percentual de área com vegetação remanescente alta ou média (%);
- NUC_COM - NUC comercial;
- PC_ERMT_3F - percentual de redes MT trifásicas (%) e
- LIXO - percentual de domicílios com coleta de lixo (%).

Da aplicação do método dinâmico com os mesmos atributos citados, encontrou-se um DEC limite objetivo global de 8,14 horas e um FEC limite objetivo global de 5,11 interrupções. Os limites globais do Cenário 5 indicam um aumento geral nos limites de DEC e FEC quando comparados aos do Cenário 1: 7,80 horas para o DEC e 5,48 interrupções para o FEC. As variações dos limites calculados para cada conjunto, contidos no Apêndice A, estão representados nos histogramas de frequência da Figura 4.6.

Figura 4.6 – Histograma de variação dos limites objetivos do Cenário 5.



Bem como no Cenário 4, verifica-se no histograma de variação dos limites de DEC da Figura 4.6 que parte significativa dos conjuntos teve seus limites alterados em relação ao Cenário 1. Apenas 17 dos 61 conjuntos mantiveram seus limites de DEC no Cenário 5, ao passo que 26 dos 61 conjuntos mantiveram seus limites de FEC.

4.6 CENÁRIO 6

Neste sexto cenário, propõe-se o uso de um único atributo no cálculo dos limites, empregando-se o método dinâmico. Os principais parâmetros do Cenário 6 estão elencados no Quadro 4.9.

Quadro 4.9 – Parâmetros do Cenário 6.

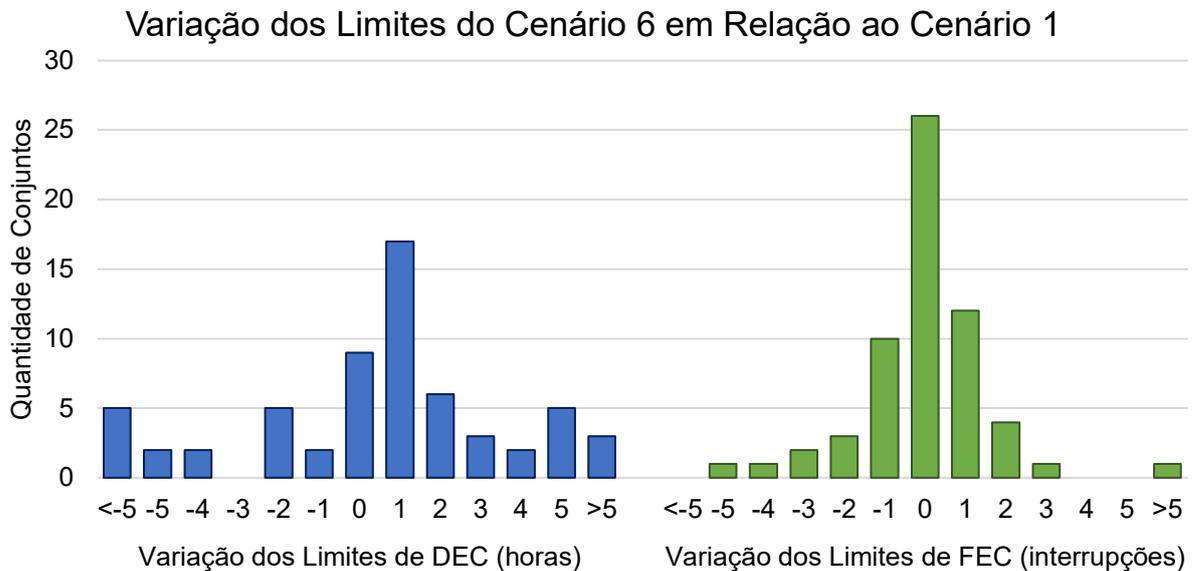
Cenário 6	
Atributos DEC	NUC_AREA
Atributos FEC	NUC_AREA
Método de Clusterização	Dinâmico
Parâmetros	$n_{max} = 100, n_{min} = 50, h_{max} = 20\%$

Fonte: Autor.

A escolha de um único atributo que caracterize ambos os indicadores de continuidade é feita no sentido de se simplificar o modelo avaliado. O atributo densidade de unidades consumidoras (NUC_AREA) é capaz de expressar se determinado conjunto possui alta ou baixa concentração de UC. Em geral, unidades consumidoras em regiões urbanas, onde a densidade de UC por área é maior, são atendidas por redes mais confiáveis e possuem tempo médio de atendimento menor. Entretanto a característica de um conjunto ser mais ou menos denso em termos de UC não é uniforme, pois uma subestação pode atender áreas urbanas e rurais simultaneamente.

Observa-se, nos histogramas da Figura 4.7, um aumento nos limites objetivos de DEC para os conjuntos.

Figura 4.7 – Histograma de variação dos limites objetivos do Cenário 6.



Fonte: Autor.

De fato, no Cenário 1 verificaram-se os limites de 7,80 horas para o DEC e 5,48 interrupções para o FEC. O DEC limite global calculado no Cenário 6 é de 8,13 horas. Já o FEC limite global é de 5,38 interrupções e é levemente inferior ao do Cenário 1. Os limites por conjunto estão no Apêndice A.

No próximo cenário, são modificados parâmetros do método dinâmico em relação aos que são empregados pela ANEEL, conforme a Nota Técnica nº 0121/2016-SRD/ANEEL.

4.7 CENÁRIO 7

No sétimo e último cenário, realiza-se a variação dos parâmetros do método dinâmico e mantém-se os atributos do cenário base. As principais características do Cenário 7 são mostradas no Quadro 4.10.

Quadro 4.10 – Parâmetros do Cenário 7.

Cenário 7	
Atributos DEC	PC_NUC_AD, PC_VRAM, PC_ERMT_3F, PLUV, CM_NUC_RES, NUC_IND
Atributos FEC	PC_VRAM, PLUV, PC_NUC_AD, NUC_COM, CM_NUC_RES, PC_ERMT_3F
Método de Clusterização	Dinâmico
Parâmetros	$n_{max} = 2610$, $n_{min} = 10$, $h_{max} = 10\%$

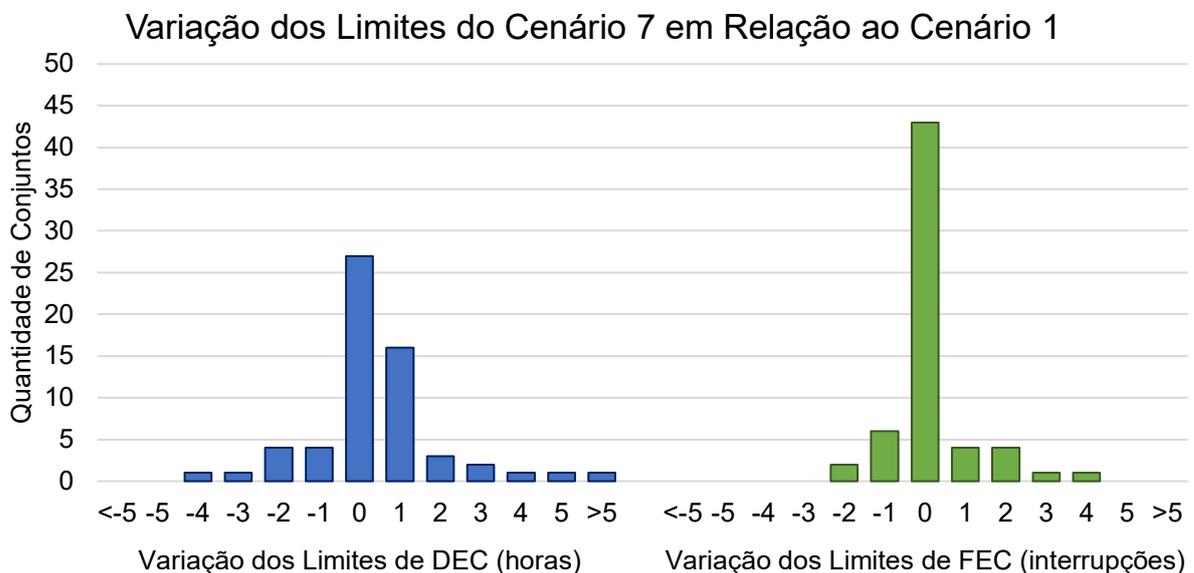
Fonte: Autor.

Para a aplicação do método dinâmico, foram escolhidos os seguintes parâmetros: número desejável de conjuntos semelhantes igual a 2610; número mínimo de conjuntos semelhantes igual a 10 e heterogeneidade percentual máxima igual a 10%. Em outras palavras, a adoção do número desejável de conjuntos semelhantes igual a 2610 é equivalente a não se definir uma quantidade máxima de conjuntos comparáveis. Ou seja, permite-se que sejam incluídos no *cluster* tantos conjuntos quanto forem possíveis antes que se atinja o limite de heterogeneidade.

Reduziu-se ainda, a quantidade mínima de conjuntos semelhantes, de maneira que fosse minimizada a quantidade de conjuntos heterogêneos. A redução na heterogeneidade de 20% para 10% e a diminuição do número mínimo de conjuntos comparáveis de 50 para 10 significam um maior rigor aplicado na metodologia de análise comparativa. Dessa forma, os conjuntos são comparados apenas com outros que possuam grande semelhança.

Em relação ao primeiro cenário, observou-se um aumento nos limites globais de DEC e de FEC. O limite objetivo global calculado é de 8,10 horas para o DEC e 5,64 interrupções para o FEC. A Figura 4.8 mostra os histogramas de variação dos limites objetivos em relação ao cenário base.

Figura 4.8 – Histograma de variação dos limites objetivos do Cenário 7.



Fonte: Autor.

Diante dos cenários apresentados, realiza-se uma comparação entre os resultados obtidos.

4.8 COMPARAÇÃO DOS RESULTADOS

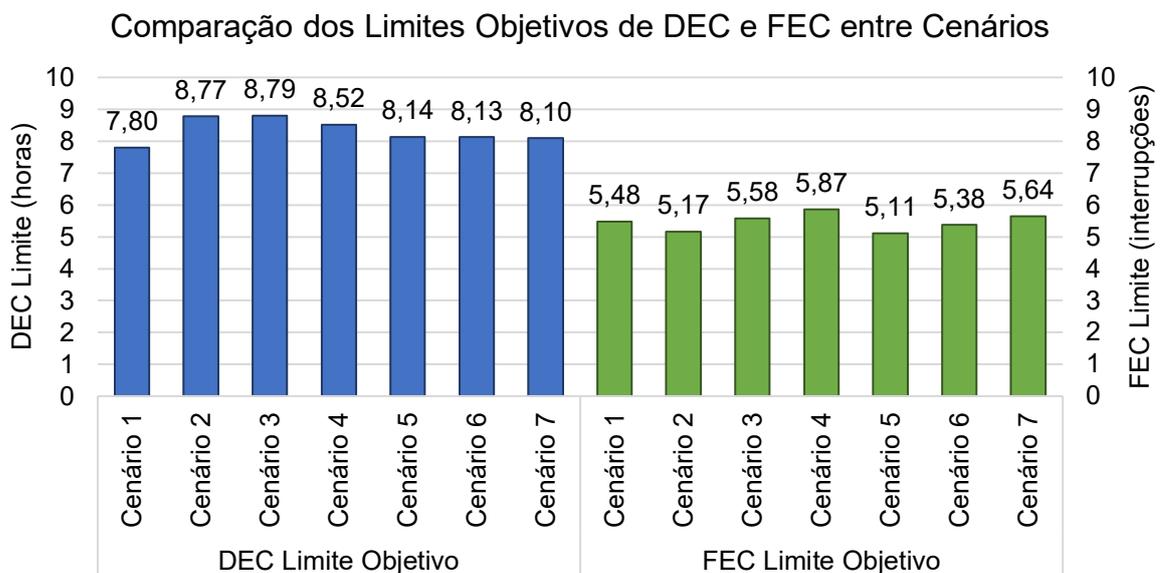
Os cenários apresentados têm como objetivo analisar a sensibilidade da metodologia frente às modificações propostas. Após a simulação e a análise dos resultados, é feita uma comparação entre os cenários apresentados. Os cenários, seus parâmetros e resultados estão resumidos no Quadro 4.11. A Figura 4.9 mostra, em um gráfico, os limites objetivos globais de DEC e FEC calculados nos sete cenários.

Observa-se que, para o DEC, o Cenário 1 é o que apresentou o menor limite global, de 7,80 horas. Nos demais cenários, o limite variou entre 8,10 e 8,79 horas. A maior variação do limite global de DEC ocorreu no Cenário 3, em que se verificou aumento de 12,7% em relação ao limite do Cenário 1.

No caso do indicador FEC, os limites globais oscilaram entre 5,11 interrupções, no Cenário 5, e 5,87 interrupções, no Cenário 4. Em relação ao cenário base, o Cenário 4 apresentou um aumento de 7,12% no limite global de FEC.

Do ponto de vista da distribuidora CEEE-D, os possíveis cenários que mais a favoreceriam são os Cenários 2 e 3, nos quais foram empregados os métodos *k-means* e Ward respectivamente. Nesses cenários foram observados os dois maiores valores de DEC limite objetivo global. A violação dos indicadores de duração da interrupção (DIC e DMIC) costuma representar a maior causa das compensações totais pagas aos clientes, e, assim, faz sentido analisar em quais cenários houve o maior aumento de DEC limite global. No caso da CEEE-D, 77% do montante de compensações corresponderam à violação de DIC ou DMIC no ano de 2020.

Figura 4.9 – Comparação dos limites objetivos globais.



Fonte: Autor.

Quadro 4.11 – Resumo dos cenários.

Cenário	Atributos	Agrupamento	Parâmetros	DEC Limite (horas)	FEC Limite (interrupções)
1	7 atributos ANEEL	Dinâmico	$n_{max} = 100$, $n_{min} = 50$, $h_{max} = 20\%$	7,80	5,48
2	7 atributos ANEEL	<i>k-means</i>	$k = 30$	8,77	5,17
3	7 atributos ANEEL	Ward	$k = 30$	8,79	5,58
4	Correlação = 40% VIF = 5	Dinâmico	$n_{max} = 100$, $n_{min} = 50$, $h_{max} = 20\%$	8,52	5,87
5	Correlação = 20% VIF = 3	Dinâmico	$n_{max} = 100$, $n_{min} = 50$, $h_{max} = 20\%$	8,14	5,11
6	Densidade de UC	Dinâmico	$n_{max} = 100$, $n_{min} = 50$, $h_{max} = 20\%$	8,13	5,38
7	7 atributos ANEEL	Dinâmico	$n_{max} = 2610$, $n_{min} = 10$, $h_{max} = 10\%$	8,10	5,64

Fonte: Autor.

Por outro viés, entende-se que os limites de DEC mais elevados nos casos em que foi alterado o método de agrupamento empregado, Cenários 2 e 3, poderiam não ser verificados se fossem analisadas outras distribuidoras. Além disso, apesar de constituírem-se como cenários favoráveis para a CEEE-D, um aumento de até 12,7% nos limites pode não corresponder a uma redução equivalente das compensações financeiras pela transgressão dos limites de continuidade.

Destaca-se que os métodos de *clusterização k-means* e Ward apresentam problemas quanto ao tratamento de conjunto heterogêneos (*outliers*), um dos motivos pelos quais a ANEEL substituiu essas técnicas pelo método dinâmico. Além disso, observou-se, durante a execução dos *scripts* em MATLAB, um maior esforço computacional requerido quando eram empregados os métodos *k-means* e Ward. A execução do método dinâmico exigiu menos cálculos, pois os agrupamentos foram feitos apenas para os conjuntos de interesse, ou seja, os 61 conjuntos da CEEE-D.

Da análise dos histogramas, os parâmetros dos Cenários 4 e 6 foram os que provocam as maiores alterações dos limites por conjunto de UC dentre os cenários apresentados. Nesses casos, os histogramas apresentam caudas maiores, ou seja, existem mais conjuntos com maiores variações de limites. Tal fato indica que os Cenários 4 e 6 são os que mais destoam do modelo utilizado pela ANEEL.

O Cenário 7, em contrapartida, foi aquele cujos resultados mais se assemelharam aos do cenário base. Ao se observar os histogramas, percebe-se que há uma grande quantidade de conjuntos concentrados em pequenas variações dos limites. Conclui-se,

desse modo, que há uma menor sensibilidade da metodologia frente à variação dos parâmetros do método dinâmico, quando comparada às mudanças dos atributos selecionados ou do método de agrupamento utilizado.

4.9 CONSIDERAÇÕES SOBRE O CAPÍTULO

Diante do exposto, afirma-se que, de modo geral, a metodologia de análise comparativa de desempenho utilizada pela ANEEL tem maior sensibilidade a variações dos atributos descritores dos indicadores. Uma sensibilidade intermediária foi observada na variação do método de *clusterização* e uma pequena sensibilidade, na mudança dos parâmetros do método dinâmico.

Cabe ressaltar que os estudos de caso propostos não findam a discussão acerca da definição dos limites. A metodologia deste trabalho possibilita a criação de inúmeros outros cenários, utilizando-se dados de quaisquer distribuidoras brasileiras. A avaliação de outros cenários contribuiria para a obtenção de resultados mais assertivos.

5 CONCLUSÃO

Neste trabalho explorou-se a metodologia de definição de limites de continuidade de energia elétrica, um importante mecanismo de regulação dos serviços prestados pelas distribuidoras de energia, por parte da Agência Nacional de Energia Elétrica. Tais limites visam estimular investimentos na manutenção e modernização das redes elétricas que corroborem para um padrão ótimo da qualidade do serviço. Objetiva-se, assim, criar um ambiente competitivo entre as empresas.

A metodologia de análise comparativa de desempenho tem como premissa a de que conjuntos semelhantes devem apresentar desempenho equivalente. A semelhança entre conjuntos é definida por suas características físico-elétricas, expressas por atributos. Conjuntos cujos atributos sejam semelhantes entre si, são comparados de tal modo que seus indicadores de continuidade sejam equivalentes. Para isso, estabelece-se uma referência, a qual determinará o objetivo a ser alcançado por cada conjunto em um intervalo de tempo definido.

A etapa de comparação dos conjuntos com os seus mais semelhantes é realizada empregando-se técnicas estatísticas de agrupamento. Os métodos apresentados no decorrer deste texto foram propostos ou utilizados pela ANEEL ao longo dos anos no contexto da definição dos limites dos indicadores. Desde o 3º Ciclo de Revisões Tarifárias Periódicas aplica-se o método dinâmico na metodologia.

Verifica-se que o método como são definidos os limites é decorrente de um longo processo de melhorias, proporcionadas pelo crescente número de informações disponíveis para serem analisadas e pelas diversas contribuições fornecidas por empresas do setor elétrico e pela academia. Contudo não se pode afirmar que o método atual está acabado. Contribuições são necessárias para que os limites propostos sejam justos e possíveis de serem cumpridos.

Nesse sentido, o presente trabalho objetivou analisar três métodos de agrupamento de conjuntos de unidades consumidoras aplicados na definição de limites dos indicadores de continuidade. Avaliou-se o uso dos métodos dinâmico, *k-means* e Ward em estudos de caso considerando os 61 conjuntos da Companhia Estadual de Distribuição de Energia Elétrica (CEEE-D) e dados dos anos de 2011 a 2013. Foram calculados os limites objetivos de DEC e de FEC dos conjuntos para sete diferentes cenários, em que foram variados, além do método de *clusterização*, os atributos selecionados e os parâmetros do método dinâmico.

Dentre os resultados observados, foi percebida uma maior dependência do modelo da ANEEL aos atributos selecionados, o que reforça o argumento de que a escolha de atributos é fundamental para a metodologia de definição dos limites. A escolha de um método de *clusterização* também tem impacto sobre os limites calculados. Em relação aos

parâmetros do método dinâmico, foi observada uma menor sensibilidade do modelo.

Para o caso da distribuidora CEEE-D, os cenários mais favoráveis tratam-se daqueles em que foram utilizados os métodos de agrupamento *k-means* e Ward. Contudo, nos cenários 2 a 7, houve aumento do DEC limite objetivo global, o que indica a possibilidade de redução das compensações pagas pela CEEE-D casos fossem empregados os parâmetros propostos.

Quanto ao emprego dos métodos de *clusterização*, percebeu-se que o *k-means* e o de Ward requerem um esforço computacional mais elevado do que o método dinâmico. Esse fato deve-se a possibilidade de, com o método dinâmico, formar *clusters* tão somente para os conjuntos de interesse, ou seja, para os 61 conjuntos da CEEE-D. Por outro lado, os métodos *k-means* e Ward exigem a formação de agrupamentos com todos os 2610 conjuntos contidos na base.

Uma outra desvantagem da aplicação do método das k-médias está na sua dificuldade de reprodução. Diferentes execuções do método podem resultar em diferentes arranjos de *cluster* a depender da escolha dos centroides iniciais. Esse pode ser um problema para a definição da metodologia de análise comparativa pela ANEEL.

Por fim, destaca-se que, com a metodologia proposta neste trabalho, diversos outros estudos de caso podem ser feitos. Os cenários apresentados foram escolhidos em função da sua relevância e os resultados obtidos não esgotam a discussão acerca da definição de limites de DEC e FEC.

5.1 OPORTUNIDADES DE TRABALHOS FUTUROS

A discussão sobre a metodologia de análise comparativa de desempenho pode ser aprofundada com as seguintes sugestões de trabalhos futuros:

- Expandir a análise dos limites para as demais distribuidoras brasileiras, de forma a identificar em quais condições determinadas empresas possam ser privilegiadas ou prejudicadas pela metodologia;
- Estimar, para cada cenário proposto, as compensações a serem pagas aos consumidores e avaliar a relevância das variações dos parâmetros na metodologia.

5.2 PRODUÇÃO CIENTÍFICA

Durante a elaboração deste trabalho de conclusão de curso, foram desenvolvidos os seguintes trabalhos:

5.2.1 Artigos publicados em anais de eventos

1. SANTOS, J. R; **SILVEIRA, L.**; SANTOS, L. L. C.; TISCHER, C. B.; GARCIA, R. A. F.
Método de Extração de Atributos Técnicos e Geográficos de Distribuidoras de Energia Elétrica In: 13th Seminar on Power Electronics and Control (SEPOC 2021), 2021, Santa Maria - Brasil.
2. **SILVEIRA, L.**; SANTOS, J. R; SANTOS, L. L. C.; GARCIA, R. A. F.
Análise de Sensibilidade do Método Dinâmico na Definição de Limites dos Indicadores de Continuidade In: Conferência Brasileira sobre Qualidade da Energia Elétrica (XIV CBQEE), 2021, Foz do Iguaçu - Brasil.

REFERÊNCIAS BIBLIOGRÁFICAS

AL-JARRAH, O. Y. et al. Multi-layered clustering for power consumption profiling in smart grids. **IEEE Access**, v. 5, p. 18459–18468, 2017. Acesso em 20 jan. 2021. Disponível em: <<https://ieeexplore.ieee.org/document/7947198>>.

ANEEL. **Anexo da Nota Técnica nº 0064, de 23 de junho de 2009**: Revisão da Regulação sobre a Continuidade do Fornecimento de Energia Elétrica na Distribuição. 2009. Acesso em 15 mai. 2021. Disponível em: <https://www2.aneel.gov.br/aplicacoes/consulta_publica/detalhes_consulta.cfm?IdConsultaPublica=162>.

_____. **Anexo da Nota Técnica nº 0028, de 30 de junho de 2010**: Revisão da Metodologia de Estabelecimento dos Limites dos Indicadores Coletivos de Continuidade. 2010. Acesso em 06 ago. 2020. Disponível em: <http://www2.aneel.gov.br/aplicacoes/consulta_publica/documentos/Nota%20T%C3%A9cnica%2028_2010%20-%20SRD%20-%20Anexo.pdf>.

_____. **Nota Técnica nº 0048**: Aprimoramento da metodologia de definição dos limites para os indicadores coletivos de continuidade. 2010. Acesso em 26 jan. 2020. Disponível em: <https://www2.aneel.gov.br/aplicacoes/audiencia/arquivo/2010/046/documento/nota_tecnica_0048_srd.pdf>.

_____. **Nota Técnica nº 0021, de 7 de junho de 2011**: Resultados da segunda parte da Audiência Pública nº 046/2010 relativa ao aprimoramento da metodologia utilizada pela ANEEL para o estabelecimento dos limites dos indicadores de continuidade coletivos e a criação do indicador de desempenho global de continuidade. 2011. Acesso em 14 jul. 2021. Disponível em: <https://www2.aneel.gov.br/aplicacoes/audiencia/dspListaResultado.cfm?attAnoAud=2010&attIdeAud=468&attAnoFasAud=2011&id_area=13>.

_____. **Nota Técnica nº 0059, de 12 de junho de 2014**: Revisão da metodologia de definição de limites para os indicadores de continuidade DEC e FEC das distribuidoras. 2014. Acesso em 13 dez. 2020. Disponível em: <https://www2.aneel.gov.br/aplicacoes/audiencia/arquivo/2014/029/documento/nota_tecnica_0059_2014_srd.pdf>.

_____. **Nota Técnica nº 0102, de 3 de dezembro de 2014**: Revisão da metodologia de definição de limites para os indicadores de continuidade DEC e FEC das distribuidoras. 2014. Acesso em 06 ago. 2020. Disponível em: <http://www2.aneel.gov.br/aplicacoes/audiencia/dspListaResultado.cfm?attAnoAud=2014&attIdeAud=898&attAnoFasAud=2014&id_area=13>.

_____. **Nota Técnica nº 0121, de 24 de agosto de 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. 2016. Acesso em 15 mai. 2021. Disponível em: <https://www2.aneel.gov.br/aplicacoes/audiencia/arquivo/2016/056/documento/nota_tecnica_0121_ceed-d.pdf>.

_____. **Indicadores de Compensação de Continuidade**: Compensação pela transgressão dos limites de continuidade. 2020. Acesso em 02 fev. 2021. Disponível em: <<https://www.aneel.gov.br/indicadores-de-compensacao-de-continuidade>>.

_____. **Procedimentos de Distribuição - PRODIST - Módulo 10**: Sistema de Informação Geográfica Regulatório. 2021. Acesso em 26 jan. 2021. Disponível em: <<https://www.aneel.gov.br/modulo-10>>.

_____. **Procedimentos de Distribuição - PRODIST - Módulo 8**: Qualidade de Energia. 2021. Acesso em 26 jan. 2021. Disponível em: <<https://www.aneel.gov.br/modulo-8>>.

_____. **Resolução Normativa n° 925, de 16 de março de 2021**: Aprova a revisão dos Módulos 1, 6 e 8 dos Procedimentos de Distribuição de Energia Elétrica no Sistema Elétrico Nacional - PRODIST e dos Submódulos 2.5 e 2.5A dos Procedimentos de Regulação Tarifária - PRORET. 2021. Acesso em 14 jul. 2021. Disponível em: <<http://www2.aneel.gov.br/cedoc/ren2021925.pdf>>.

BRASIL. Lei n° 9.427, de 26 de dezembro de 1996. Institui a Agência Nacional de Energia Elétrica - ANEEL, disciplina o regime das concessões de serviços públicos de energia elétrica e dá outras providências. **Diário Oficial da União**, Poder Executivo, Brasília, DF, 1996. Acesso em 18 jan. 2020. Disponível em: <http://www.planalto.gov.br/ccivil_03/leis/19427compilada.htm>.

_____. Decreto n° 2.335, de 6 de outubro de 1997. Constitui a Agência Nacional de Energia Elétrica - ANEEL, autarquia sob regime especial, aprova sua Estrutura Regimental e o Quadro Demonstrativo dos Cargos em Comissão e Funções de Confiança e dá outras providências. **Diário Oficial da União**, Poder Executivo, Brasília, DF, 1997. Acesso em 18 jan. 2020. Disponível em: <http://www.planalto.gov.br/ccivil_03/decreto/d2335.htm>.

CAI, H. et al. Predicting the energy consumption of residential buildings for regional electricity supply-side and demand-side management. **IEEE Access**, v. 7, p. 30386–30397, 2019. Acesso em 20 jan. 2021. Disponível em: <<https://ieeexplore.ieee.org/document/8651280>>.

FÁVERO, L. P.; BELFIORE, P. **Manual de análise de dados: estatística e modelagem multivariada com Excel®, SPSS® e Stata®**. Rio de Janeiro: Elsevier, 2017.

GRANELL, R.; AXON, C. J.; WALLOM, D. C. H. Impacts of raw data temporal resolution using selected clustering methods on residential electricity load profiles. **IEEE Transactions on Power Systems**, v. 30, n. 6, p. 3217–3224, 2015. Acesso em 20 jan. 2021. Disponível em: <<https://ieeexplore.ieee.org/document/6994292>>.

Grupo CEEE. **Dados Econômicos-Financeiros**. 2020. Acesso em 02 fev. 2021. Disponível em: <<http://ri.cee.com.br/Default.aspx?codPagina=341>>.

HAIR, J. F. et al. **Análise multivariada de dados**. Porto Alegre: Bookman, 2009.

HSU, D. Comparison of integrated clustering methods for accurate and stable prediction of building energy consumption data. **Applied Energy**, v. 160, p. 153 – 163, 2015. ISSN 0306-2619. Acesso em 20 jan. 2021. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0306261915010624>>.

ROSA, F. A. F. **Avaliação de Novos Atributos para Estabelecimento de Metas de DEC e FEC em Redes de Distribuição**. 2009. 85 p. Dissertação (Mestrado em Engenharia Elétrica) — Universidade Federal de Itajubá, Itajubá, MG, 2009. Acesso em 19 jan. 2021. Disponível em: <<https://repositorio.unifel.edu.br/xmlui/handle/123456789/1580>>.

SANTOS, J. R. **Análise estatística para seleção dos atributos que caracterizam os indicadores de continuidade do fornecimento de energia elétrica**. 2021. 77 p. Monografia (Trabalho de Conclusão de Curso) — Curso de Graduação em Engenharia Elétrica, Universidade Federal de Santa Maria, Cachoeira do Sul, 2021.

SANTOS, J. R. et al. Método de extração de atributos técnicos e geográficos de distribuidoras de energia elétrica. **11th Seminar on Power Electronics and Control**, v. 1, p. 5, 2021.

SPERANDIO, M. **Classificação de conjuntos consumidores de energia elétrica via mapas Auto-Organizáveis e estatística multivariada**. 2004. 105 p. Dissertação (Mestrado em Engenharia Elétrica) — Universidade Federal de Santa Catarina, Florianópolis, SC,

2004. Acesso em 19 jan. 2021. Disponível em: <<https://repositorio.ufsc.br/xmlui/handle/123456789/87238>>.

TANURE, J. E. P. S. **Análise Comparativa de Empresas de Distribuição para o Estabelecimento de Metas de Desempenho para Indicadores de Continuidade do Serviço de Distribuição**. 2000. 160 p. Dissertação (Mestrado em Engenharia Elétrica) — Universidade Federal de Itajubá, Itajubá, MG, 2000.

TANURE, J. E. P. S. **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**. 2004. 228 f. Tese (Doutorado em Engenharia) — Universidade de São Paulo, São Paulo, 2004. Acesso em 19 jan. 2021. Disponível em: <<https://repositorio.usp.br/item/001415412>>.

APÊNDICE A – LIMITES OBJETIVOS DE DEC E FEC POR CONJUNTO

Tabela A.1 – Limites de DEC calculados para os conjuntos da CEEE-D (horas).

Conjunto	Cen 1	Cen 2	Cen 3	Cen 4	Cen 5	Cen 6	Cen 7
Alvorada	6	6	6	9	7	7	6
Areal	8	10	10	12	12	11	9
Arroio do Sal	9	9	13	10	10	7	10
Arroio Grande	12	13	13	15	13	15	11
Atlântida	8	9	10	9	9	6	9
Atlântida Sul	9	9	10	10	9	8	7
Bagé 1	5	6	6	8	6	8	6
Bagé 2	11	11	14	15	9	15	15
Camaquã	15	16	17	12	17	11	15
Camaquã 2	6	6	10	10	10	12	6
Canguçu	19	22	19	10	25	12	19
Capão Novo	9	9	13	11	9	7	10
Capivarita	10	10	13	14	13	15	11
Cassino	8	9	9	9	8	8	7
Charqueadas 2	8	9	13	8	7	8	8
CRM 1 - Minas do Leão	10	10	10	15	12	14	9
Dom Pedrito	10	10	13	11	9	15	11
Eldorado do Sul	6	6	6	10	7	7	6
Encruzilhada	12	10	13	16	16	13	13
Guaíba	9	9	10	11	8	9	9
Jaguarão	9	9	7	11	8	14	9
Marmeleiro	25	28	33	12	22	45	31
Mostardas	13	10	17	16	16	22	13
Osório 1	9	10	10	8	10	9	10
Palmares do Sul	10	10	10	10	12	15	10
Pedro Osório	16	13	17	12	16	15	19
Pelotas 1	5	6	6	7	7	7	5
Pelotas 2	5	6	9	6	5	7	5
Pelotas 3	15	13	17	12	13	10	15
Pelotas 4	16	13	17	11	15	10	18
Pinhal	6	6	6	9	8	7	11
Porto Alegre 1	5	6	6	5	5	6	5

Continua na próxima página

Tabela A.1 – *Continuação*

Conjunto	Cen 1	Cen 2	Cen 3	Cen 4	Cen 5	Cen 6	Cen 7
Porto Alegre 10	5	6	6	6	5	6	6
Porto Alegre 12	5	5	5	5	5	6	5
Porto Alegre 13	5	6	6	6	5	6	5
Porto Alegre 14	6	6	6	8	6	8	5
Porto Alegre 16	5	6	6	8	6	6	5
Porto Alegre 18	5	5	5	9	4	5	5
Porto Alegre 2	5	6	6	6	4	6	5
Porto Alegre 3	5	6	6	6	5	6	5
Porto Alegre 4 - Centro	3	2	2	2	2	4	1
Porto Alegre 4 - Centro 2	6	6	6	5	5	6	6
Porto Alegre 5	5	5	5	6	4	7	5
Porto Alegre 6	6	6	6	8	6	7	8
Porto Alegre 8	5	6	6	6	5	6	5
Porto Alegre 9	5	6	6	6	5	7	6
Quinta	15	23	14	15	10	10	13
Rio Grande 1	5	6	6	8	7	7	6
Rio Grande 2	6	6	6	8	5	7	6
Rio Grande 3	14	13	17	9	13	12	14
Santa Vitória do Palmar	11	9	7	12	8	11	11
Santo Antônio	17	16	17	13	17	11	15
São Lourenço do Sul	14	10	17	13	18	12	17
Terra de Areia	23	79	26	12	21	11	19
Torres 1	12	9	13	7	10	8	9
Tramandaí	6	9	6	9	7	7	6
UTE Presidente Médici	15	13	17	14	16	20	17
UTE São Jerônimo	7	9	13	9	9	7	8
Vasconcelos	19	22	17	15	19	13	20
Viamão 1	7	6	7	9	7	7	8
Viamão 2	5	6	6	10	7	6	6

Fonte: Autor.

Tabela A.2 – Limites de FEC calculados para os conjuntos da CEEE-D (interrupções).

Conjunto	Cen 1	Cen 2	Cen 3	Cen 4	Cen 5	Cen 6	Cen 7
Alvorada	5	4	4	7	4	5	5
Areal	6	7	5	7	5	7	6
Arroio do Sal	6	5	7	6	7	5	6
Arroio Grande	8	7	7	6	7	8	7
Atlântida	5	7	7	6	5	5	5
Atlântida Sul	6	7	7	6	7	5	5
Bagé 1	5	4	4	7	4	5	5
Bagé 2	7	8	7	8	7	8	7
Camaquã	9	7	11	7	8	6	9
Camaquã 2	5	5	5	6	5	7	5
Canguçu	10	10	11	7	11	8	14
Capão Novo	6	5	7	6	7	5	6
Capivarita	6	7	8	8	6	8	6
Cassino	5	5	6	7	8	5	5
Charqueadas 2	5	6	4	6	5	5	5
CRM 1 - Minas do Leão	6	7	5	8	6	8	6
Dom Pedrito	7	6	5	7	7	8	9
Eldorado do Sul	5	5	5	6	5	5	5
Encruzilhada	7	7	8	8	9	8	7
Guaíba	6	6	7	9	5	6	6
Jaguarão	6	6	8	7	5	7	7
Marmeleiro	11	14	12	7	7	26	9
Mostardas	7	7	8	7	8	10	7
Osório 1	7	7	10	7	6	6	9
Palmares do Sul	6	7	5	7	5	8	6
Pedro Osório	9	10	7	7	8	8	11
Pelotas 1	5	4	5	5	5	5	5
Pelotas 2	4	3	4	5	4	5	4
Pelotas 3	9	7	8	7	7	7	9
Pelotas 4	9	10	11	8	7	7	11
Pinhal	6	6	5	6	5	5	7
Porto Alegre 1	4	4	4	4	4	4	4
Porto Alegre 10	4	4	4	4	4	4	4
Porto Alegre 12	4	4	4	4	4	4	4
Porto Alegre 13	4	4	4	5	4	5	4

Continua na próxima página

Tabela A.2 – *Continuação*

Conjunto	Cen 1	Cen 2	Cen 3	Cen 4	Cen 5	Cen 6	Cen 7
Porto Alegre 14	5	4	4	6	4	5	5
Porto Alegre 16	5	4	4	5	4	5	5
Porto Alegre 18	4	5	4	3	5	3	4
Porto Alegre 2	4	3	4	3	4	4	3
Porto Alegre 3	4	4	4	5	4	5	4
Porto Alegre 4 - Centro	2	2	2	3	2	3	2
Porto Alegre 4 - Centro 2	3	3	4	4	3	4	3
Porto Alegre 5	4	3	4	3	4	4	3
Porto Alegre 6	5	6	5	6	5	5	6
Porto Alegre 8	4	4	4	4	4	4	4
Porto Alegre 9	5	4	4	6	4	5	5
Quinta	8	8	7	9	7	7	6
Rio Grande 1	5	5	4	5	5	5	5
Rio Grande 2	5	4	4	6	4	5	5
Rio Grande 3	8	7	8	8	8	8	8
Santa Vitória do Palmar	7	6	8	7	6	6	7
Santo Antônio	10	7	11	8	7	6	10
São Lourenço do Sul	7	7	11	7	7	8	10
Terra de Areia	12	16	10	7	9	7	12
Torres 1	8	7	7	7	7	5	7
Tramandaí	5	4	4	6	4	5	5
UTE Presidente Médici	9	10	7	8	9	10	9
UTE São Jerônimo	5	6	4	7	8	5	5
Vasconcelos	9	7	11	8	8	8	8
Viamão 1	5	4	7	8	4	5	6
Viamão 2	5	5	5	6	5	5	5

Fonte: Autor.