

Formação de Conjuntos de Sistemas de Distribuição quanto às Variações de Tensão de Curta Duração

Paulo Vitor Grillo de Souza

Marcos Vinícius da Costa

UNIFEI – Universidade Federal de Itajubá
pvgrillo@hotmail.com, marcosvcost@gmail.com

Homero Krauss Ribeiro Filho

José Maria de Carvalho Filho

UNIFEI – Universidade Federal de Itajubá
homerokrauss@yahoo.com.br, jmaria@unifei.edu.br

Resumo — Este artigo apresenta uma proposta de metodologia para o agrupamento de sistemas de distribuição quanto às variações de tensão de curta duração (VTCDs). São estabelecidos critérios para determinar qual o método de formação de conjuntos mais adequado para aplicação em uma base de dados específica. Para testar o procedimento proposto, foram feitas simulações considerando dados reais de uma distribuidora de energia elétrica brasileira. Apresenta-se uma metodologia geral, que pode ser estendida e aplicada a todos os sistemas de distribuição do Brasil, podendo servir como guia orientativo para o órgão regulador ANEEL na definição de novos padrões de referência para o fenômeno, levando em conta as características e as particularidades de cada sistema.

Palavras-chaves — Variações de Tensão de Curta Duração, Qualidade do Produto, Formação de Conjuntos.

I. INTRODUÇÃO

O desenvolvimento de novas tecnologias permitiu a modernização dos processos industriais, aumentando o dinamismo e a produtividade do setor. No entanto, os equipamentos modernos tendem a ser mais vulneráveis aos distúrbios de qualidade da energia elétrica (QEE), principalmente aqueles relacionados às VTCDs.

Em determinados processos industriais, um evento de VTCD com duração de milissegundos pode gerar uma interrupção prolongada por horas, dependendo da sensibilidade dos equipamentos instalados e das particularidades de cada processo. Assim, tem-se atrelados elevados prejuízos financeiros, devido à perda de qualidade dos produtos, perda de insumos, reparos de equipamentos e custos relativos ao tempo de parada da produção [1].

Ciente da relevância desse assunto, recentemente, a ANEEL promoveu uma revisão do módulo 8 do PRODIST. Apesar desse documento abordar diversos fenômenos de QEE, até a sua revisão 7, apenas os indicadores associados às variações de tensão em regime permanente possuíam limites, prazos de regularização e penalidades. Para os indicadores dos demais fenômenos, existiam apenas valores de referência, sem nenhum tipo de penalização quando da sua violação, e no caso das VTCDs, não existiam sequer indicadores, muito menos valores de referência ou limites.

Desse modo, em 2013, a ANEEL contratou serviços técnicos especializados de consultoria, visando a obtenção de

subsídios para o aperfeiçoamento da regulamentação que trata da qualidade do produto constante nesse módulo.

Os resultados desse trabalho foram submetidos à discussão com a sociedade por meio da Consulta Pública (CP) nº 18/2014. Um dos pontos de maior repercussão foi o referente à definição de limites para os indicadores propostos, sendo que a maioria das contribuições se mostraram contrárias a isso.

Por fim, foi instaurada a Audiência Pública (AP) nº 82/2015, onde mais uma vez discutiu-se o assunto. A ANEEL ressaltou que desde a primeira versão do PRODIST, existem valores de referência para os indicadores de qualidade do produto, exceto para as VTCDs. Logo, seria prudente a adoção de valores de referência somente para estes fenômenos, enquanto que para os demais, seria natural a evolução para limites, uma vez que o caráter orientativo do regulamento existe desde 2008 [2].

Assim sendo, no início de 2017 entrou em vigor a revisão 8 do módulo 8 do PRODIST, a qual foi embasada nos trabalhos realizados pela consultoria e nas contribuições recebidas no âmbito da CP nº 18/2014 e da AP nº 82/2015. Com relação às VTCDs, foi definida uma metodologia própria para a avaliação desses fenômenos, considerando a criação de um indicador único, denominado Fator de Impacto (*FI*).

Nesse aspecto, um dos maiores desafios quanto ao indicador *FI* consiste no estabelecimento de valores de referência (futuramente limites) consistentes ao cenário do setor elétrico brasileiro. Dada a vasta extensão do território nacional, da consequente diversidade de relevo, vegetação e clima e dos diferentes níveis de industrialização das regiões, algumas contribuições acenaram no sentido de definir limites diferenciados de acordo com as características de cada sistema.

De fato, o estabelecimento de limites para agrupamentos de sistemas de distribuição, a exemplo do que se faz na área de continuidade, parece ser mais plausível do que o estabelecimento de um único limite para todas as distribuidoras do Brasil. É neste contexto que este trabalho se insere, propondo uma metodologia para a formação de conjuntos considerando os sistemas de distribuição com características semelhantes quanto às VTCDs.

II. FUNDAMENTOS DA FORMAÇÃO DE CONJUNTOS

A análise de agrupamentos tem como objetivo dividir os elementos da amostra, ou população, em grupos de forma que

os elementos pertencentes a um mesmo grupo sejam similares entre si com respeito às variáveis (características) que neles foram medidas, e os elementos em grupos diferentes sejam heterogêneos em relação a estas mesmas características [3].

O primeiro passo para análise de agrupamentos é desenvolver uma escala quantitativa para medir a associação (similaridade) entre os objetos em estudos. A medida de dissimilaridade mais utilizada e, portanto, adotada neste trabalho, é a distância euclidiana.

Supondo um conjunto de dados com n elementos amostrais (objetos), tendo-se medido p variáveis aleatórias em cada um deles, define-se para cada elemento j o vetor de medidas X_j , conforme (1):

$$X_j = [x_{j1}, x_{j2}, \dots, x_{jp}]', \quad j = 1, 2, \dots, n \quad (1)$$

Considerando dois elementos X_l e X_m , com $l \neq m$, a distância euclidiana entre eles é definida por (2):

$$d(X_l, X_m) = \left[\sum_{i=1}^p (x_{li} - x_{mi})^2 \right]^{1/2} \quad (2)$$

Os métodos de agrupamentos são divididos em dois tipos: os não hierárquicos e os hierárquicos e estes são novamente classificados em aglomerativos e divisivos. O algoritmo dos métodos hierárquicos aglomerativos inicia considerando cada objeto como um conjunto, e em cada etapa agrupa os objetos ou conjunto de objetos que mais se assemelham. No final do processo, a um nível de similaridade pequeno, todos os objetos compõem um único conjunto. Em contrapartida, as técnicas hierárquicas divisivas funcionam na direção oposta, ou seja, parte-se de um conjunto único que contém todos os objetos do estudo e a cada passo os separam de acordo com a dissimilaridade, encerrando com conjuntos formados por apenas um objeto. Um exemplo ilustrativo de ambos os processos está apresentado na Fig. 1.

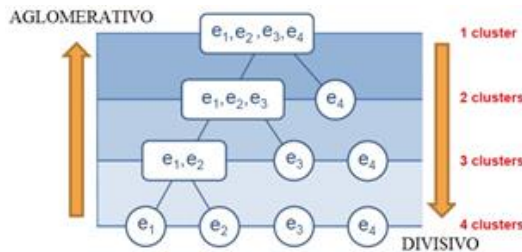


Fig. 1. Ilustração das técnicas hierárquicas aglomerativas e divisivas.

O resultado da aplicação dos métodos hierárquicos pode ser exibido na forma de dendogramas (Fig.2), os quais mostram as uniões ou divisões feitas em níveis sucessivos.

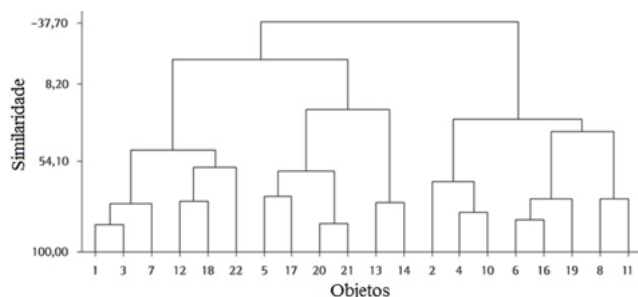


Fig. 2. Modelo de um dendograma.

O algoritmo básico para implementação dos métodos hierárquicos aglomerativos é descrito a seguir:

1. Começa-se com n agrupamentos, cada um deles formado por um único objeto. Então, cria-se uma matriz D , de distâncias (ou de similaridades), $n \times n$;
2. Procura-se na matriz D o par de agrupamentos mais próximos entre si (maior similaridade);
3. Supondo que os agrupamentos U e V sejam os mais próximos, tais agrupamentos são unidos. O novo agrupamento é nomeado como UV , e a matriz D é atualizada, apagando as linhas e colunas correspondentes a U e V , e adicionando a linha e a coluna referente às distâncias entre UV e os demais agrupamentos;
4. Os passos 2 e 3 são repetidos $n-1$ vezes. Os agrupamentos formados e o nível de distância (ou similaridade) em que as uniões ocorrem são registrados. No término, todos os objetos formarão um único agrupamento.

A distância euclidiana determina a distância entre dois objetos. À medida que objetos são agrupados, é necessário determinar como serão definidas as distâncias entre os agrupamentos, já que estes serão compostos por mais de um elemento. Para isso, existem diversos métodos de ligação.

Na sequência são apresentados os métodos utilizados neste trabalho. Para melhor compreensão, suponha dois agrupamentos, um contendo os elementos amostrais 1, 3 e 7 e o outro contendo os elementos 2 e 6, conforme (3):

$$C_1 = \{X_1, X_3, X_7\} \text{ e } C_2 = \{X_2, X_6\} \quad (3)$$

A. Método de Ligação Simples

A distância entre dois agrupamentos é definida como a distância mínima entre dois objetos, um pertencente a cada agrupamento. Logo, a distância entre C_1 e C_2 será dada por (4):

$$d(C_1, C_2) = \min\{d(X_l, X_m), l = 1, 3, 7 \text{ e } m = 2, 6\} \quad (4)$$

Como pode ser visto na Fig. 3, a distância entre os elementos 1 e 6 é a menor em relação às demais, assim, esta é considerada a distância entre os agrupamentos C_1 e C_2 .

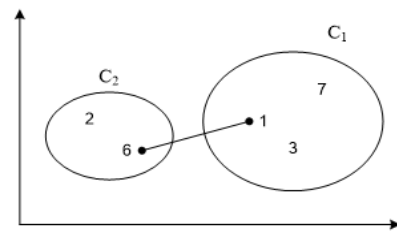


Fig. 3. Ilustração gráfica do método de ligação simples.

B. Método de Ligação Completa

A distância entre dois agrupamentos é definida como a distância máxima entre dois objetos, um pertencente a cada agrupamento. Logo, a distância entre C_1 e C_2 será dada por (5):

$$d(C_1, C_2) = \max\{d(X_l, X_m), l = 1, 3, 7 \text{ e } m = 2, 6\} \quad (5)$$

Nesse caso, a distância entre C_1 e C_2 será igual à distância entre os elementos 2 e 7, os quais são os elementos mais distantes entre si, conforme mostra a Fig. 4.

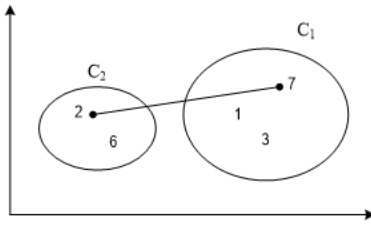


Fig. 4. Ilustração gráfica do método de ligação completa.

C. Método da Média das Distâncias

A distância entre dois agrupamentos é definida como a média das distâncias entre todos os pares de objetos, onde cada membro do par pertence a um dos agrupamentos. Desse modo, a distância entre C_1 e C_2 será dada por (6):

$$d(C_1, C_2) = \frac{1}{6} [d(X_1, X_2) + d(X_1, X_6) + d(X_3, X_2) + d(X_3, X_6) + d(X_7, X_2) + d(X_7, X_6)] \quad (6)$$

D. Método de Ward

A distância entre dois agrupamentos é definida como a soma dos desvios quadrados dos objetos em relação ao centroide.

Supondo que C_1 e C_2 sejam unidos, e que assim seja formado o agrupamento C_i , a soma dos desvios quadrados correspondente a este novo agrupamento será dada por (7):

$$SS_i = \sum (X_j - \bar{X}_i)'(X_j - \bar{X}_i) \quad j = 1, 2, 3, 6, 7 \quad (7)$$

Onde $\bar{X}_i$ representa o centroide do agrupamento C_i . No passo s , a soma dos desvios quadrados total será dada por (8):

$$SSR = \sum_{i=1}^{g_s} SS_i \quad (8)$$

Onde g_s é o número de agrupamentos no passo s . A cada passo do algoritmo, considera-se a união de todos os pares de agrupamentos, e os dois agrupamentos em que a combinação resulta em um menor crescimento de SSR são combinados.

E. Método k -means

O método de ligação k -means é a técnica de agrupamento não hierárquica mais popular. O funcionamento básico desse método consiste na alocação de cada elemento no agrupamento cujo centroide é o mais próximo desse. Segundo [3], o método é composto por quatro passos:

1. Primeiramente, escolhe-se k centroides, chamados de “sementes”, para se inicializar o processo de partição;
2. Cada elemento do conjunto de dados é, então, comparado com cada centroide inicial através de uma medida de distância. O elemento é alocado ao agrupamento cuja distância ao centroide é menor;
3. Depois de aplicar o passo 2 para cada um dos n elementos amostrais, recalcula-se os valores dos centroides para cada agrupamento formado, e repete-se o passo 2, considerando os novos centroides;
4. Os passos 2 e 3 são repetidos até que nenhuma realocação de elementos seja necessária.

O resultado final desse método está fortemente relacionado à escolha das partições iniciais ou à escolha das “sementes”.

III. METODOLOGIA PROPOSTA

A metodologia proposta se resume aos seguintes passos:

1. Escolhe-se um indicador que possa quantificar o desempenho dos sistemas elétricos frente às VTCDs;
2. Define-se através da experiência do pesquisador um universo de variáveis, que a priori, suspeita-se ter relação com esse indicador;
3. Através da análise de regressão linear múltipla, seleciona-se as variáveis para serem entrada dos algoritmos de formação de conjuntos;
4. Aplica-se os algoritmos de formação de conjuntos, considerando os vários métodos que se deseja investigar;
5. Compara-se os métodos através da análise de variância (ANOVA) entre os conjuntos formados, em relação ao indicador escolhido. O método que apresentar o maior valor para o teste de Fisher (F) será considerado o mais adequado para ser aplicado à base de dados em questão;
6. Caso dois ou mais métodos apresentem valores próximos para F , analisa-se qual deles é o mais robusto. Para isso, considera-se a aplicação desses métodos em diversos cenários de perturbação (pequeno ruído) das variáveis de entrada. O método mais robusto é aquele que coincidir o *membership* em cada cenário com o *membership* padrão (sem ruído) o maior número de vezes.

IV. ESTUDO DE CASO

Neste estudo, considerou-se 15 sistemas de distribuição pertencentes a uma distribuidora de energia elétrica do Brasil.

Para quantificar o desempenho destes sistemas frente às VTCDs, foi escolhido como indicador, simplesmente, o número anual de eventos registrados no barramento de 13,8 kV da subestação de distribuição de cada sistema, porém, sem contabilizar as elevações de tensão. Ou seja, foram contabilizados todos os eventos com amplitude inferior a 0,9 p.u e duração inferior a 3 minutos.

A princípio, poderia ser utilizado qualquer outro indicador, inclusive o Fator de Impacto. No entanto, não há um consenso a respeito do “melhor” indicador a ser utilizado, e não é objetivo deste artigo discutir tal questão.

O número de eventos foi obtido através de simulações de curtos-circuitos baseadas no procedimento de Monte Carlo. Para isso, modelou-se os respectivos sistemas no software ASPEN DistriView, sendo representados cenários de faltas referentes a 100 anos de operação.

A. Escolha das variáveis

Foram escolhidas as seguintes variáveis que pressupõe-se ter uma relação com o indicador previamente definido:

1. Potência de curto-circuito (Pcc) do barramento de 13,8 kV que alimenta cada sistema de distribuição, em MVA;
2. Área de vulnerabilidade equivalente (AV eq), que define a região do sistema elétrico, em km, onde a ocorrência de faltas acarreta eventos de amplitude inferior a 0,9 p.u no barramento em questão, sendo calculada pelo somatório dos produtos das áreas de vulnerabilidade de cada tipo de falta pela sua respectiva probabilidade de ocorrência;

3. Taxa de falta (λ), em faltas/100km/ano. Como as faltas são a principal causa de ocorrência das VTCDs, quanto maior a taxa de falta, maior o número de eventos;
4. Extensão da rede, em km. Quanto maior a extensão da rede, maior a quantidade de faltas, para um λ constante.

A Tabela I apresenta, para cada sistema de distribuição, os valores calculados para essas quatro variáveis, bem como a média do número anual de eventos (100 anos de operação).

TABELA I. BASE DE DADOS DOS SISTEMAS DE DISTRIBUIÇÃO.

Sistema de Distribuição	Pcc	AV eq	Taxa de Falta	Extensão da Rede	Média de Eventos
SD_1	75	74	235	365	61
SD_2	155	74	126	886	44
SD_3	152	41	404	165	77
SD_4	39	174	90	1087	117
SD_5	130	140	192	1685	114
SD_6	184	94	206	863	89
SD_7	210	176	212	903	114
SD_8	75	240	184	625	172
SD_9	178	170	279	1800	169
SD_10	180	69	474	323	97
SD_11	166	195	174	1893	124
SD_12	66	114	97	244	94
SD_13	99	175	148	1173	135
SD_14	87	160	203	1137	109
SD_15	81	291	159	1442	165

B. Seleção das variáveis

Após escolher as variáveis, é necessário selecionar aquelas que serão entrada dos algoritmos de formação de conjuntos.

Embora, neste caso, o objetivo seja agrupar os sistemas de distribuição com características semelhantes em relação ao número de VTCDs, este indicador não é utilizado como entrada desses algoritmos. Assim sendo, é fundamental que as variáveis selecionadas tenham uma relação significativa com o indicador escolhido, pois quanto melhor a seleção, mais coerente serão os agrupamentos quanto ao número de eventos. Além disso, essa etapa também tem o intuito de reduzir a dimensionalidade do conjunto de variáveis.

Para isso, realiza-se uma análise de regressão linear múltipla, supondo que o modelo que relaciona essas variáveis à resposta (indicador) seja dado por (9) [4]:

$$y_j = \beta_0 + \beta_1 x_{j1} + \dots + \beta_p x_{jp} + \epsilon_j \quad j = 1, 2, \dots, n \quad (9)$$

Onde β_p é o coeficiente da p -ésima variável, y_j é o valor do indicador para o sistema SDj e ϵ_j é um termo de erro aleatório.

A seleção das variáveis é feita através do P -value (P) correspondente ao seu coeficiente. Assim sendo, tem-se o teste de hipóteses dado por (10):

$$H_0: \beta_i = 0 \text{ e } H_1: \beta_i \neq 0 \quad i = 0, 1, \dots, p \quad (10)$$

Onde H_0 é a hipótese nula e H_1 a hipótese alternativa.

Se o P -value correspondente ao coeficiente de uma variável for menor ou igual a um nível de significância (α) pré-determinado, H_0 é rejeitada e conclui-se que esse coeficiente é diferente de zero, ou seja, a variável em questão é uma adição significativa ao modelo. Caso contrário, H_0 não é rejeitada e conclui-se que tal variável tem um efeito não significativo.

A Fig. 5 mostra o resultado obtido para essa análise, utilizando o software Minitab:

Regression Analysis: Média de Eventos versus Pcc; AV eq; Taxa de Falta; Extensão da Rede						
Backward Elimination of Terms						
	---Step 1---		---Step 2---		---Step 3---	
	Coef	P	Coef	P	Coef	P
Pcc	-0,048	0,670				
AV eq	0,5321	0,000	0,5383	0,000	0,5718	0,000
Taxa de Falta	0,1505	0,013	0,1318	0,000	0,1343	0,000
Extensão da Rede	0,0085	0,479	0,00571	0,553		
Coefficients						
Term	Coef	SE Coef	T-Value	P-Value	VIF	
AV eq	0,5718	0,0379	15,08	0,000	2,18	
Taxa de Falta	0,1343	0,0258	5,20	0,000	2,18	

Fig. 5. Resultado da análise de regressão.

O próprio software disponibiliza uma opção chamada “Stepwise”, onde foi escolhido o procedimento “Backward Elimination of Terms”. Desse modo, a cada passo do processo, realiza-se uma análise de regressão, onde a variável menos significativa é removida do modelo, até que reste somente as variáveis cujo P -value é menor do que α .

Para a base de dados em questão e um α igual a 0,05, foram selecionadas as variáveis “AV eq” e “Taxa de Falta”.

C. Aplicação dos métodos de formação de conjuntos

O pesquisador deve definir o número de conjuntos que serão formados, de acordo com o que lhe parecer mais lógico, e levando em conta o número de sistemas que ele deseja agrupar.

Neste artigo, considerou-se o agrupamento dos sistemas de distribuição em três conjuntos (grupos 1, 2 e 3). Os dendogramas obtidos são apresentados nas Fig. 6 a 9.

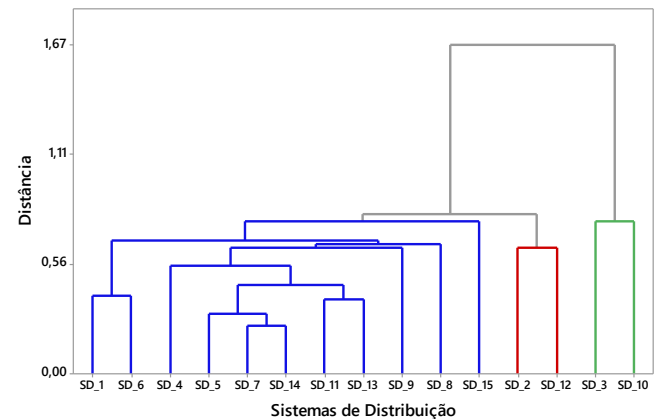


Fig. 6. Dendrograma (método de ligação simples).

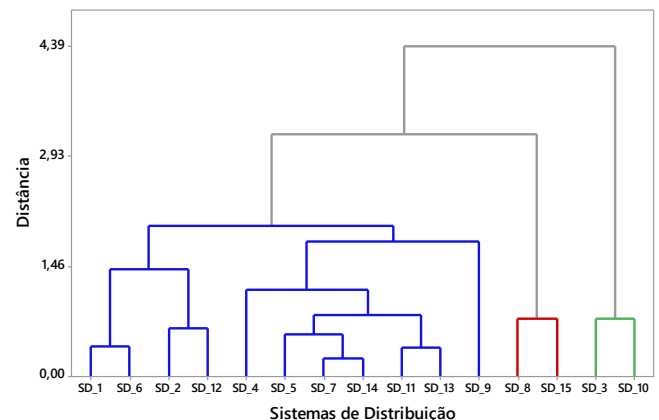


Fig. 7. Dendrograma (método de ligação completa).

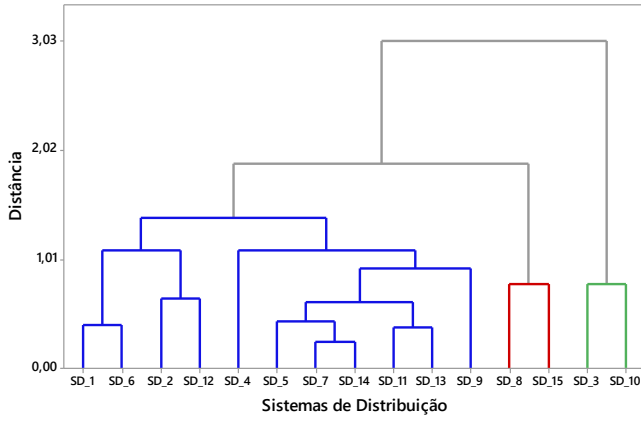


Fig. 8. Dendrograma (método da média das distâncias).

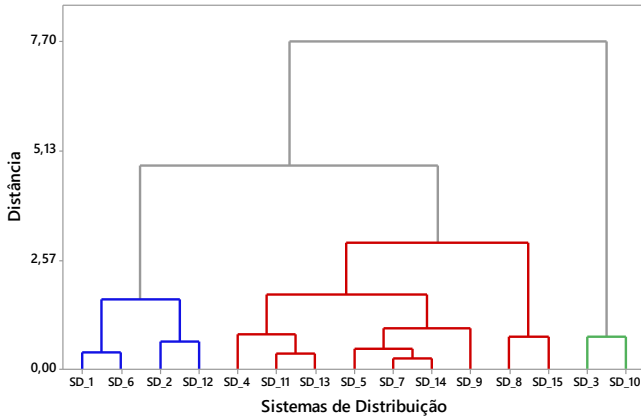


Fig. 9. Dendrograma (método de Ward).

A Tabela II apresenta o *membership* gerado com a aplicação de cada método.

TABELA II. COMPARAÇÃO ENTRE O MEMBERSHIP GERADO EM CADA MÉTODO.

Sistema de Distribuição	Ligação Simples	Ligação Completa	Média das Distâncias	Ward	k-means
SD_1	1	1	1	1	1
SD_2	2	1	1	1	1
SD_3	3	2	2	2	3
SD_4	1	1	1	3	2
SD_5	1	1	1	3	1
SD_6	1	1	1	1	1
SD_7	1	1	1	3	1
SD_8	1	3	3	3	2
SD_9	1	1	1	3	1
SD_10	3	2	2	2	3
SD_11	1	1	1	3	2
SD_12	2	1	1	1	1
SD_13	1	1	1	3	2
SD_14	1	1	1	3	1
SD_15	1	3	3	3	2

D. Comparação entre os métodos

Este passo da metodologia visa determinar qual o método que melhor agrupou os sistemas de distribuição quanto ao indicador escolhido. Assim, para cada algoritmo utilizado, é realizada uma análise de variância entre os grupos 1, 2 e 3.

A comparação entre os métodos é feita com base no valor obtido para o teste de Fisher, o qual é calculado a partir da variação entre e dentro dos grupos formados.

A variação entre os grupos, $SS_{Between}$, é calculada pelo somatório dos desvios quadrados das médias de cada agrupamento em relação à média geral [4], conforme (11):

$$SS_{Between} = \sum_{i=1}^r a_i (\bar{y}_i - \bar{y})^2 \quad (11)$$

Onde a_i e $\bar{y}_i$ correspondem ao número de observações e à média, respectivamente, do i -ésimo agrupamento, r é o número de grupos e $\bar{y}$ é a média geral de todas as observações.

Já a variação dentro dos grupos, SS_{Within} , é calculada pelo somatório dos desvios quadrados de cada observação em relação à média do seu respectivo grupo [4], conforme (12):

$$SS_{Within} = \sum_{i=1}^r \sum_{j=1}^{a_i} (y_{ij} - \bar{y}_i)^2 \quad (12)$$

Onde y_{ij} é o valor da j -ésima observação (valor do indicador para o j -ésimo sistema) pertencente ao i -ésimo grupo.

Assim sendo, o valor de F é obtido de acordo com (13):

$$F = \frac{SS_{Between} / (r - 1)}{SS_{Within} / (n - r)} \quad (13)$$

Quanto maior a variação entre os grupos, e menor a variação dentro dos mesmos, maior será o valor de F .

As Fig. 10 a 14 apresentam os resultados da ANOVA para cada método, onde mais uma vez utilizou-se o Minitab.

One-way ANOVA: grupo1; grupo2; grupo3					
Analysis of Variance					
Source	DF	Adj SS	Adj MS	F-Value	P-Value
Factor	2	6654	3327	3,01	0,087
Error	12	13247	1104		
Total	14	19901			
Means					
Factor	N	Mean	StDev	95% CI	
grupo1	11	124,5	34,3	(102,6; 146,3)	
grupo2	2	69,0	35,4	(17,8; 120,2)	
grupo3	2	87,0	14,1	(35,8; 138,2)	
Pooled StDev = 33,2249					

Fig. 10. Resultados da análise de variância (método de ligação simples).

One-way ANOVA: grupo1; grupo2; grupo3					
Analysis of Variance					
Source	DF	Adj SS	Adj MS	F-Value	P-Value
Factor	2	7984	3991,9	4,02	0,046
Error	12	11917	993,1		
Total	14	19901			
Means					
Factor	N	Mean	StDev	95% CI	
grupo1	11	106,4	34,2	(85,7; 127,1)	
grupo2	2	87,0	14,1	(38,4; 135,6)	
grupo3	2	168,50	4,95	(119,95; 217,05)	
Pooled StDev = 31,5133					

Fig. 11. Resultados da análise de variância (método de ligação completa).

One-way ANOVA: grupo1; grupo2; grupo3					
Analysis of Variance					
Source	DF	Adj SS	Adj MS	F-Value	P-Value
Factor	2	7984	3991,9	4,02	0,046
Error	12	11917	993,1		
Total	14	19901			
Means					
Factor	N	Mean	StDev	95% CI	
grupo1	11	106,4	34,2	(85,7; 127,1)	
grupo2	2	87,0	14,1	(38,4; 135,6)	
grupo3	2	168,50	4,95	(119,95; 217,05)	
Pooled StDev = 31,5133					

Fig. 12. Resultados da análise de variância (método da média das distâncias).

One-way ANOVA: grupo1; grupo2; grupo3					
Analysis of Variance					
Source	DF	Adj SS	Adj MS	F-Value	P-Value
Factor	2	12597	6298,4	10,35	0,002
Error	12	7304	608,7		
Total	14	19901			
Means					
Factor	N	Mean	StDev	95% CI	
grupo1	4	72,0	23,7	(45,1; 98,9)	
grupo2	2	87,0	14,1	(49,0; 125,0)	
grupo3	9	135,44	26,04	(117,53; 153,36)	
Pooled StDev = 24,6715					

Fig. 13. Resultados da análise de variância (método de Ward).

One-way ANOVA: grupo1; grupo2; grupo3					
Analysis of Variance					
Source	DF	Adj SS	Adj MS	F-Value	P-Value
Factor	2	7232	3616	3,43	0,067
Error	12	12669	1056		
Total	14	19901			
Means					
Factor	N	Mean	StDev	95% CI	
grupo1	8	99,3	37,9	(74,2; 124,3)	
grupo2	5	142,6	24,6	(110,9; 174,3)	
grupo3	2	87,0	14,1	(36,9; 137,1)	
Pooled StDev = 32,4919					

Fig. 14. Resultados da análise de variância (método k-means).

A partir desses resultados, conclui-se que o método de Ward é o que melhor agrupou os sistemas de distribuição considerados, pois foi o método para o qual foi obtido o maior valor de F (10,35).

Vale salientar que isso não quer dizer que esse método seja o melhor em todas as ocasiões. Sendo assim, se a base de dados for outra, mesmo que o fenômeno sob análise seja o mesmo, deve-se efetuar a análise seguindo os passos da metodologia proposta, a fim de verificar qual o método mais adequado a ser aplicado.

E. Análise de robustez

Quando dois ou mais métodos apresentam valores de F próximos, recomenda-se analisar qual deles é o mais robusto. Como neste caso, o método de Ward obteve um valor de F bem superior aos demais, decidiu-se analisar somente a robustez desse método, criando diversos cenários de perturbação das variáveis de entrada.

Para isso, foi utilizado o software Matlab. Em cada cenário, variou-se os valores das variáveis “AV eq” e “Taxa de Falta” por meio da adição de pequenos ruídos aleatórios (normalmente distribuídos, com média zero e desvio padrão σ), e em seguida, aplicou-se o método de Ward, com o intuito de verificar se os conjuntos formados anteriormente ainda se mantinham, mesmo com a variação dessas variáveis.

Assim, realizou-se 10 testes para cada valor de desvio padrão, e em cada teste foram criados 10000 cenários. Desse modo, foram obtidos os resultados apresentados na Tabela III:

TABELA III. VALOR DE N PARA ALGUNS VALORES DE DESVIO PADRÃO.

σ	0,2	0,3	0,4	1,0	1,5	2,5	5,0
N	9951	9582	9026	7004	6278	5232	3404

Onde N é a média do número de cenários, nos 10 testes, em que o *membership* gerado coincidiu com *membership* “padrão” (apresentado na 5ª coluna da Tabela II).

Quando dois ou mais métodos são comparados, considera-se o mais robusto aquele para o qual foram obtidos os maiores valores de N.

V. CONCLUSÕES

As VTCDs são os distúrbios de qualidade da energia elétrica que mais impactam os consumidores industriais que possuem cargas sensíveis, gerando prejuízos expressivos. Apesar dos avanços alcançados com a revisão 8 do módulo 8 do PRODIST, ainda há muito o que ser discutido no sentido de evoluir na regulamentação desses fenômenos.

Neste contexto, este artigo propõe uma metodologia para o agrupamento dos sistemas de distribuição com características semelhantes quanto às VTCDs, o que pode ser utilizado na definição de novos padrões de referência para cada conjunto.

A comparação entre os valores obtidos para o teste de Fisher permite que seja escolhido o método que gerou os melhores conjuntos, ou seja, aquele que apresentou maior variação entre os grupos formados e menor variação dentro dos mesmos, o que é justamente o objetivo da análise de agrupamentos.

VI. REFERÊNCIAS

- [1] CARVALHO FILHO, J.M. “Uma contribuição à avaliação do atendimento a consumidores com cargas sensíveis – Proposta de novos indicadores”, 2000. Tese (Doutorado em Engenharia Elétrica) – Escola Federal de Engenharia de Itajubá, Itajubá, 2000.
- [2] AGÊNCIA NACIONAL DE ENERGIA ELÉTRICA - ANEEL “Nota Técnica 85/2016-SRD/ANEEL”. Disponível em: <http://www2.aneel.gov.br/aplicacoes/audiencia/arquivo/2015/082/resultado/nota_tecnica_0085_2016_srd_-_pos_ap.pdf> Acesso em: 25 nov. 2016.
- [3] MINGOTI, S. A. “Análise de Dados Através de Métodos de Estatística Multivariada: Uma Abordagem Aplicada”. 2a. ed. Belo Horizonte: Editora UFMG, 2005.
- [4] MONTGOMERY, D. C.; RUNGER, G. C. “Applied Statistics and Probability for Engineers”. 5a ed. John Wiley & Sons Inc, 2011.