GRUPO DE ESTUDO DE SISTEMAS DE DISTRIBUIÇÃO - GDI

PLANEJAMENTO A PARTIR DE REDES REPRESENTATIVAS DO SISTEMA DE DISTRIBUIÇÃO CONSIDERANDO A PENETRAÇÃO DE RECURSOS ENERGÉTICOS DISTRIBUÍDOS

ANA GABRIELA BEZERRA BENITEZ(1);DANIEL PEREZ DUARTE(1);SÉRGIO MISHIMA DOS SANTOS BARBOSA(1);REGINALDO PRESTES(2);CLEVERSON LUIZ DA SILVA PINTO(2)

SINAPSIS(1);COPEL DISTRIBUIÇÃO(2)

RESUMO

Este trabalho apresenta alguns dos principais desafios do planejamento das redes de distribuição e descreve um modelo para caracterização de alimentadores e avaliação de novas obras, considerando recursos energéticos distribuídos (RED) e priorização multricriterial. A caracterização de alimentadores do sistema é realizada por meio de técnicas de *machine learning* (ML) e análises estatísticas. Para a avaliação de REDs, incorporam-se novos parâmetros de simulação como a capacidade de acomodação das redes e uma metodologia de projeção da difusão de geração distribuída fotovoltaica (GDFV). A partir de indicadores técnicos selecionados, apresenta-se um modelo de priorização de obras com estudo de caso utilizando dados da COPEL-DIS.

PALAVRAS-CHAVE

Modelo de Bass, efeito de *spillover*, geração distribuída, clusterização, avaliação multicritério

1.0 INTRODUÇÃO

O setor elétrico vem passando por diversas evoluções ao longo das últimas décadas. Os anos 1990 foram caracterizados pela abertura do mercado e pela consolidação da figura de um órgão regulador responsável por definir as regras de eficiência e qualidade do serviço provido pelas concessionárias, encorajando investimentos prudentes e a melhoria de indicadores de qualidade. Mais recentemente, observam-se movimentos voltados ao crescimento da geração com baixas emissões de carbono, à implementação de programas de resposta da demanda e à exploração do armazenamento de energia, resultando em um sistema mais inteligente, eficiente e seguro. Para que este cenário seja atingido a contento, são necessárias mudanças de paradigma nos ambientes regulatório, político e de planejamento, com o objetivo de garantir que a infraestrutura seja capaz de acomodar a velocidade de penetração destas novas tecnologias.

Neste contexto de evolução setorial, este artigo tem como objetivo apresentar uma metodologia para caracterização das redes de distribuição e avaliação de novas obras, a partir da aplicação de técnicas de ML e de análises estatísticas, considerando o impacto de REDs por meio da incorporação de um modelo de projeção do crescimento de GDFV e um estudo para a quantificação da capacidade de acomodação (*hosting capacity*) de novos recursos às redes.

A fim de introduzir e qualificar a abordagem metodológica proposta, as próximas subseções se dedicam à apresentação de alguns dos conceitos fundamentais que norteiam o trabalho.

1.1 Modelos de difusão de GDFV

Em se tratando de uma tecnologia nova, que requer a alteração de padrões de comportamento do consumidor para sua introdução, é possível estudar a geração distribuída fotovoltaica sob a ótica de teorias de inovações. Em cenário nacional, Konzen (2014) foi o primeiro a projetar a incorporação da GDFV no mercado residencial de acordo características socioeconômicas da população, baseando-se na Teoria da Difusão de Inovações de Rogers, representada matematicamente pelo modelo de Bass. Essa metodologia vem sido adotada pela Empresa de Pesquisa Energética (EPE) em seus estudos, ampliando as premissas utilizadas a fim de se considerar novos grupos além do residencial, permitindo a obtenção de uma projeção temporal a nível das áreas de concessão das distribuidoras brasileiras.

ana.benitez@sinapsisenergia.com

Apesar da projeção temporal ampla já apresentar um panorama relevante para as distribuidoras, estudos de difusão espacial em menor granularidade se fazem necessários a fim de se analisar o impacto da GDFV na operação e planejamento da expansão em sistemas locais. Na literatura internacional, observa-se uma convergência dos modelos em salientar o impacto da transferência de conhecimento entre usuários e não usuários no processo de difusão, conceito definido como transbordamento tecnológico ou efeito de *spillover* (BALTA-OZKAN et al., 2015; DHARSHING, 2017). Outro aspecto observado é o objetivo em se apresentar resultados por bairros, distritos, estados e áreas geográficas, e não a nível de redes ou subestações, como pretendido e desenvolvido no presente trabalho.

1.2 Clusterização

Análise de clusters, ou agrupamentos, é parte de um conjunto de técnicas multivariadas, cujo objetivo é agrupar indivíduos considerando as características de seus atributos. A clusterização é um método de classificação não supervisionada, em que um cluster é entendido como um grupo de indivíduos com atributos similares entre si e dissimilares em relação aos indivíduos de outro grupo. Existem na literatura diversas estratégias para clusterizar indivíduos, sendo uma das mais exploradas o método das k-médias, ou k-*means*. O objetivo dessa técnica é particionar os grupos de maneira que seus elementos estejam o mais próximo possível da média do cluster (minimizar a distância intra-cluster) e longe da média dos outros clusters (maximizar a distância inter-cluster).

Com o aumento da dimensionalidade do problema, isto é, do número de variáveis avaliadas, algoritmos de clusterização que utilizam fórmulas de distância para a classificação dos elementos (como o k-médias) apresentam dificuldade para obter agrupamentos significativos, uma vez que os indivíduos tendem a ter distâncias muito similares entre si. Nesse caso, podem ser aplicadas técnicas de pré-processamento como a desconsideração de variáveis altamente correlacionadas, que acrescentam redundâncias aos dados, como também algoritmos de redução de dimensionalidade tais como o *Principal Component Analysis* (PCA) ou o *T-Distributed Stochastic Neighbor Embedding* (t-SNE).

Adicionalmente, para se garantir a homogeneidade dos agrupamentos, é possível utilizar processos estatísticos, a fim de se determinar o número ótimo de clusters para um dado conjunto de dados. Um dos métodos mais utilizados é o *Sum of Squared Errors* (SSE), que busca mensurar a média da soma dos erros quadráticos intra-cluster para diferentes quantidades de agrupamentos. De uma análise gráfica, identifica-se a existência de um ponto ótimo (conhecido como *elbow*) a partir do qual a SSE diminui com menor velocidade com o incremento da quantidade de clusters.

1.3 *Hosting Capacity* (HC)

O conceito de *hosting capacity* pode ser definido como a taxa de penetração de recursos energéticos distribuídos que pode ser introduzida a uma rede sem violar as condições normais de operação do sistema ou trazer prejuízo à qualidade de energia (ISMAEL et al., 2019). Destaca-se que o cálculo da HC não se trata de um procedimento fixo, com um único resultado. Em efeito, a HC é calculada para diferentes indicadores de performance como variações de tensão e frequência, limites de carregamento de trechos de rede, perdas técnicas, qualidade de energia (como presença de harmônicas introduzidas à rede) e problemas dos sistemas de proteção.

Com relação ao cálculo da HC propriamente dito, a consideração de incertezas relacionadas à localização, à capacidade instalada e à geração variável de futuros recursos distribuídos é requerida. Nesse processo, podem ser aplicados métodos analíticos, estocásticos ou simplificados, a depender da complexidade do sistema estudado e do nível de detalhamento exigido (ISMAEL et al., 2019). De maneira geral, os modelos de cálculo se distinguem pela precisão e tempo de processamento computacional, variando do estudo barra a barra do sistema (analítico), à estimação de múltiplos cenários de geração (estocástico) e ao estabelecimento de correlações de características elétricas e de consumidores entre a rede alvo e alimentadores previamente avaliados (simplificado). O último modelo, por sua característica, permite estimar a capacidade de acomodação de um sistema com tempo de processamento bastante reduzido, sendo explorado por ferramentas de planejamento comercial como o EPRI DRIVE (RYLANDER et al., 2015).

2.0 METODOLOGIA

Nesta seção serão apresentadas as premissas metodológicas do artigo. A metodologia geral desenvolvida é resumida no fluxograma apresentado na FIGURA 1. Cada uma das etapas é descrita em mais detalhes nos itens dessa seção.

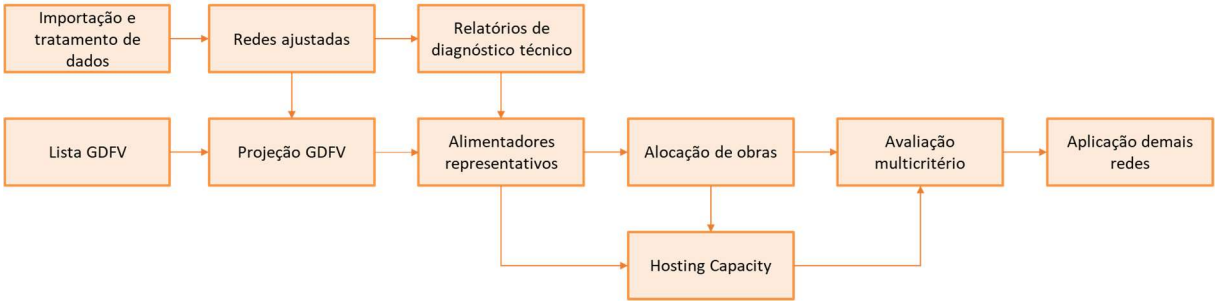

FIGURA 1: Fluxograma geral do projeto

2.1 Importação e tratamento de informações

Nesta subseção são descritas as etapas para a importação, tratamento e geração dos diagnósticos técnicos das redes de média tensão da COPEL-DIS, utilizando o software SINAPgrid. A plataforma SINAPgrid possibilita a análise integrada de redes de energia elétrica de alta, média e baixa tensão.

Para tal, realiza-se primeiramente a importação automatizada das redes, a partir de informações extraídas do GIS. Procede-se então ao ajuste dos valores de demanda de saída dos alimentadores, de acordo com curvas tratadas para caracterização da situação de máxima demanda (patamares horários de curva percentil 90%, considerando dados de 2020). Na sequência, ajustam-se os indicadores de continuidade, com base na importação de arquivo de ocorrências da empresa combinadas aos valores de DEC e FEC apurados por conjunto pela ANEEL em 2020. Por fim, importam-se taxas anuais de crescimento no horizonte de 2021-2030 para cada um dos alimentadores. Ao finalizar todas as etapas de importação e tratamento, executa-se o diagnóstico técnico para as redes selecionadas, exportando-se relatórios a nível de alimentadores (SDMT) e subestações (SED). A descrição detalhada da etapa via fluxograma é apresentada na FIGURA 2.

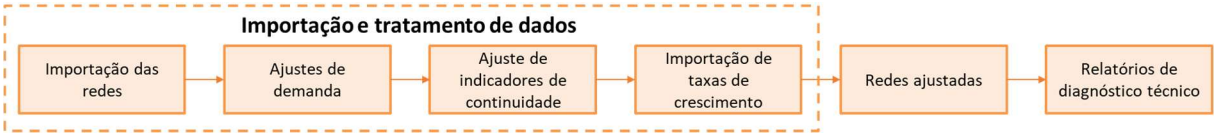

FIGURA 2: Fluxograma das etapas de importação e tratamento de dados

2.2 Projeção da GDFV

Para a projeção da difusão temporal da GDFV na área de concessão da COPEL-DIS, empregou-se uma adaptação da metodologia 4MD da EPE (EPE, 2021). De maneira simplificada, o modelo de Bass desenvolvido se baseia na análise do histórico de adoção da tecnologia para a construção de uma curva base de difusão (curva S ou de Bass) para anos futuros, considerando limitações técnico-econômicas da população, bem como a atratividade do investimento, balizada pelo valor da tarifa, do sistema de compensação de créditos e da evolução do preço dos painéis fotovoltaicos. Como resultado, foram obtidos montantes de novos adotantes e de capacidade instalada ano a ano para a próxima década (até 2030), segmentados em classes de consumo (residencial, rural, comercial e industrial) e subgrupos tarifários (B1, B2, B3, A3a e A4) selecionados. O comparativo entre os grupos pode ser realizado a partir da análise das curvas de Bass.

Ainda, de maneira a satisfazer as necessidades de se analisar o impacto da GDFV nas redes de distribuição, foi proposto um modelo de difusão espacial a nível dos alimentadores da COPEL-DIS, baseado no conceito de transbordamento tecnológico, utilizando como dados de entrada os resultados do modelo de difusão temporal. Para o modelo proposto, primeiramente agregam-se todos adotantes independentemente da classe de consumo e do grupo tarifário a nível das subestações, sendo os novos adotantes alocados de maneira proporcional à taxa de penetração (em termos de adotantes por unidades consumidoras) calculada em 2020. Para distribuição dos adotantes a nível dos alimentadores, atribuem-se pesos pela proporção de unidades consumidoras (UC) do alimentador em relação ao total de UCs supridas pela subestação. Como resultado, obtém-se uma única taxa de penetração futura por alimentador, a qual é utilizada como variável descritiva do modelo de clusterização. Neste artigo, utiliza-se a projeção de penetração de geração distribuída (GD) para o ano de 2025.

Em resumo, a metodologia de difusão de adotantes de GDFV é apresentada no fluxograma da FIGURA 3.

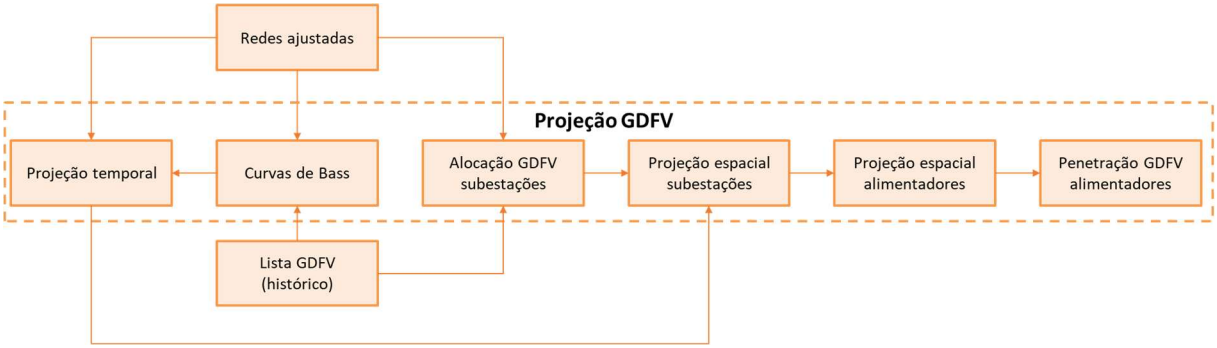

FIGURA 3: Fluxograma da metodologia de projeção da GDFV

2.3 Caracterização e clusterização das redes

As variáveis utilizadas para a clusterização das redes pelo método das k-médias são apresentadas na TABELA 1. São obtidos os valores para cada um dos alimentadores de 13,8 kV da regional Leste da COPEL-DIS, que não possuem grandes recursos energéticos distribuídos alocados, a partir de informações cadastrais das redes, dos resultados dos diagnósticos gerados pela ferramenta SINAPgrid para o ano de 2021 e da taxa de penetração de GDFV obtida do modelo de projeção para o ano de 2025.

TABELA 1: Base de dados para análise das redes de média tensão

Atributo	Descrição
SED-SDMT	Identificação da subestação e do alimentador
Regional	Nome da Regional
Tensão nominal	Tensão nominal do alimentador, em kV
Trechos totais	Quantidade de trechos do alimentador, em km
UC	Quantidade de consumidores do alimentador
Energia	Energia consumida pelas cargas do alimentador, em MWh/mês
Demanda máxima	Demanda máxima verificada no alimentador, em MVA
Religadores	Número de religadores do alimentador
Reguladores	Número de reguladores do alimentador
Penetração GDFV	Penetração de GDFV projetada em 2025
Queda de tensão	Maior queda de tensão percentual observada em uma barra do alimentador
Carregamento	Carregamento máximo percentual observado em uma barra do alimentador
Perdas	Perdas técnicas, em MWh, calculadas para o alimentador considerando 30 dias
DEC	Duração Equivalente de Interrupção por UC, em horas, calculado para o alimentador
FEC	Frequência Equivalente de Interrupção por UC, calculado para o alimentador

Em função do número de atributos trabalhados (12 variáveis classificatórias), são realizadas duas etapas de pré-processamento a fim de selecionar variáveis não redundantes via análise de correlação e de reduzir a dimensionalidade do problema a partir da aplicação do algoritmo t-SNE. Na sequência, os atributos são normalizados e seleciona-se o número ótimo de clusters por inspeção visual da curva de SSE. Finalmente, efetua-se a clusterização pelo método das k-médias, identificando-se o agrupamento a que pertence cada alimentador e a rede representativa de cada grupo (aquela que se encontra mais próxima do valor médio do cluster). A FIGURA 4 apresenta o fluxograma da metodologia desenvolvida.

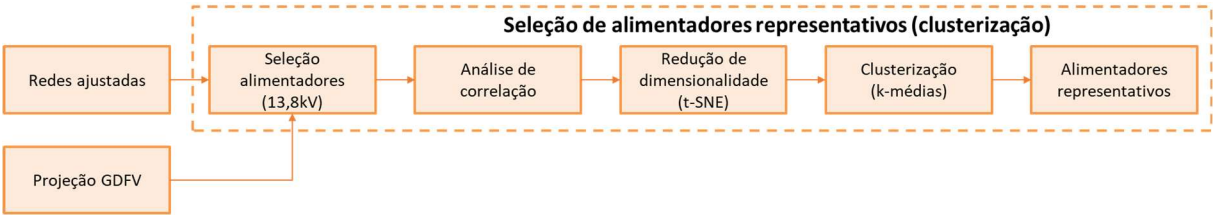

FIGURA 4: Fluxograma da metodologia de caracterização das redes

2.4 Modelo estruturado e aplicação de obras em redes representativas

Selecionados os alimentadores representativos, propõe-se a simulação via SINAPgrid da realização de dois tipos de obras em cada uma das redes, a saber: recondutoramento e divisão de tronco. Para o recondutoramento, altera-se o arranjo do tronco do alimentador para a entrada de cabos protegidos de 185mm^2. Na divisão de tronco, particiona-se a rede para a criação de um novo alimentador com aproximadamente metade do carregamento original, utilizando os mesmos arranjos e caminhamento do tronco antigo. Ambas as obras são conduzidas a fim de proporcionar aumento da oferta, eliminar sobrecarga de trechos, diminuir perdas e quedas de tensão, além de minimizar indicadores de continuidade. A avaliação do benefício da instalação de cada uma das obras é mensurada pelo comparativo de relatórios de diagnóstico técnico da ferramenta SINAPgrid.

2.5 *Hosting Capacity*

No estudo desenvolvido, aplica-se o modelo simplificado do EPRI – com o EPRI DRIVE utilizado como plugin de cálculo integrado ao software SINAPgrid – para determinação da HC das redes representativas elegidas pela clusterização, antes e após a aplicação das obras sugeridas em 2.4. Em aderência aos atuais estudos de planejamento e aos critérios comumente avaliados em estudos de HC (AL-SAFFAR et al., 2019; DUBEY et al., 2015), apura-se a capacidade de acomodação em termos dos limites de sub/sobretensão (0,95 e 1,05 pu), carregamento (100%) e reversão de fluxo. O valor obtido, que representa a máxima potência que pode ser injetada à rede sem impor a superação dos limites operacionais selecionados, é posteriormente utilizado como um dos critérios da priorização multicriterial desenvolvida.

2.6 Análise multicritério para priorização das obras

O objetivo da análise multicriterial proposta é realizar uma comparação entre os benefícios técnicos proporcionados por cada uma das obras selecionadas. Para isso, utilizam-se como atributos os níveis de carregamento, queda de tensão, perdas técnicas e *hosting capacity* antes e após a realização das obras. Como parâmetros adicionais

recuperam-se os quantitativos, em km, de cabos protegidos 185mm² (recondutoramento) e as demandas máximas e número de barras dos alimentadores (divisão de tronco). Assim, o benefício técnico de cada alternativa é calculado pelas equações (1) a (6).

$$BT_{rec} = \left(\frac{C_{orig} - C_{rec}}{C_{orig}} + \frac{dV_{orig} - dV_{rec}}{dV_{orig}} + \frac{PT_{orig} - PT_{rec}}{PT_{orig}} + \frac{HC_{rec} - HC_{orig}}{HC_{orig}}\right) \times \frac{T_{rec}}{\left(T_{rec} - T_{orig}\right)} \quad (1)$$

$$BT_{div1} = \left(\frac{C_{orig} - C_{div1}}{C_{orig}} + \frac{dV_{orig} - dV_{div1}}{dV_{orig}} + \frac{PT_{orig} - PT_{div1}}{PT_{orig}}\right) \times \frac{D_{div1}}{\left(D_{orig} - D_{div1}\right)} \quad (2)$$

$$BT_{div1_hc} = \frac{HC_{div1} - HC_{orig}}{HC_{orig}} \times \frac{D_{div1}}{\left(D_{orig} - D_{div1}\right)} \quad (3)$$

$$BT_{div2} = \left(\frac{C_{orig} - C_{div2}}{C_{orig}} + \frac{dV_{orig} - dV_{div2}}{dV_{orig}} + \frac{PT_{orig} - PT_{div2}}{PT_{orig}}\right) \times \frac{D_{div2}}{\left(D_{orig} - D_{div2}\right)} \quad (4)$$

$$BT_{div2_hc} = \frac{HC_{div2} - HC_{orig}}{HC_{orig}} \times \frac{D_{div2}}{\left(D_{orig} - D_{div2}\right)} \quad (5)$$

$$BT_{div} = \frac{BT_{div1} \times D_{div1} + BT_{div2} \times D_{div2}}{D_{div1} + D_{div2}} + \frac{BT_{div1_hc} \times nb_{div1} + BT_{div2_hv} \times nb_{div2}}{nb_{div1} + nb_{div2}} \quad (6)$$

Onde

- BT: benefício técnico da obra, em termos de percentual de melhoria global dos atributos da rede original;
- C: carregamento máximo percentual observado em uma barra do alimentador;
- dV: maior queda de tensão percentual observada em uma barra do alimentador;
- PT: perdas técnicas calculadas para o alimentador considerando 30 dias, em MWh;
- HC: *hosting capacity* mensurada pela média das potências mínimas acomodadas em cada barra do alimentador;
- T: quantidade, em km, de trechos de cabos protegidos 185mm²;
- D: demanda máxima verificada no alimentador, em MVA;
- nb: número de barras de cada alimentador;
- $orig$: rede original;
- rec: obra de recondutoramento;
- div: divisão de tronco (índices 1 e 2 indicam cada um dos alimentadores resultantes da divisão).

Pelo equacionamento desenvolvido, os benefícios técnicos das obras são adimensionais e ponderados em função do acréscimo de trechos de cabos protegidos ou da redução de demanda promovida pela divisão do tronco. A fim de agregar os resultados dos dois alimentadores resultantes da divisão de tronco, utiliza-se a demanda máxima e o número de barras como critério para a atribuição de pesos. É de se destacar que foi assumido que todos os atributos avaliados possuem mesma relevância, o que pode ser alterado pela eventual alocação de pesos. A priorização das obras é realizada pela comparação numérica simples dos resultados de benefício técnico obtidos.

A partir da aplicação da métrica desenvolvida aos alimentadores representativos, objetiva-se identificar a ordem de prioridade de cada cluster, a qual pode ser extrapolada às demais redes de cada agrupamento. Como forma de validar essa proposição, os benefícios técnicos de ambas as obras são calculados para redes representadas selecionadas. No estudo de caso deste artigo, apresentam-se os resultados obtidos para o cluster 2.

3.0 RESULTADOS E DISCUSSÃO

Após o detalhamento das premissas metodológicas do artigo, são apresentados os resultados e principais discussões acerca do processo de caracterização de redes e priorização de obras adotado.

3.1 Projeção da GDFV

3.1.1 Difusão temporal da GDFV

Dos dados históricos de adotantes de GDFV na área de concessão da COPEL-DIS, foram determinadas as curvas de Bass para cada um dos submercados estudados (FIGURA 5). Da análise da penetração proposta, destaca-se a projeção de saturação do mercado para todos os grupos já na próxima década. É de se salientar que este resultado pode estar atrelado ao *boom* de novos adotantes de GDFV no ano de 2020, como consequência direta da iminente alteração do Sistema de Compensação da MMGD, que diminuirá a atratividade do investimento. Assim, o cenário de corrida do ouro desenhado no último ano pode ter contribuído para a distorção das curvas S no sentido de maximizar o caráter inovador da população, elemento a ser acompanhado pela replicação da metodologia nos próximos anos. Ressalvas metodológicas a parte, das comparações entre grupos é possível identificar características como a menor propensão do consumidor residencial à inovação em relação às classes rural, comercial e industrial, o que é visualizado graficamente a partir do deslocamento da curva S desse grupo à direita.

Ainda assim, em função da maior proporção de unidades consumidoras deste segmento, o mercado residencial representa e deve continuar representando na próxima década, o principal nicho de expansão da GDFV em número de adotantes (483% de crescimento no horizonte quinquenal). A TABELA 2 resume as projeções do modelo de Bass desenvolvido, pontuando os resultados por grupo para o ano de 2020 (último ano de histórico) e de 2025 (ponto futuro para o modelo de difusão espacial).

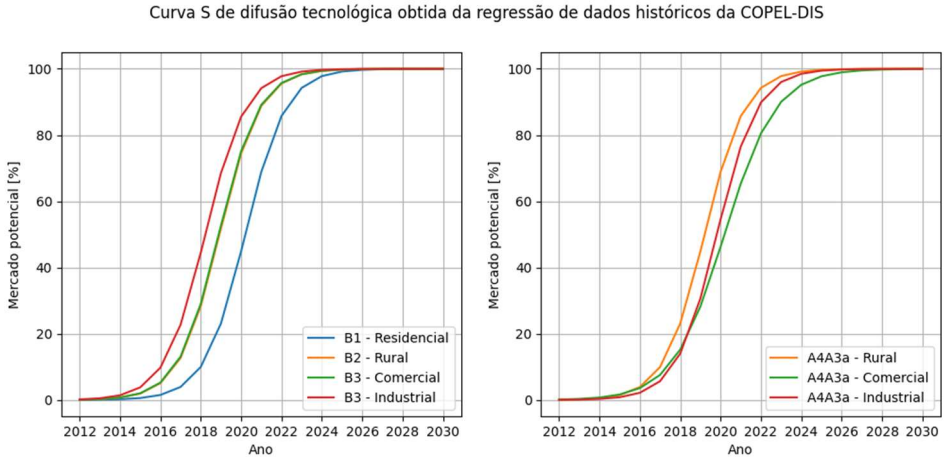

FIGURA 5: Curvas S de difusão obtidas da metodologia de difusão

TABELA 2: Resultados do modelo de difusão temporal por segmento de mercado analisado

		2020			2025			
Subgrupo Tarifário	**Classe de Consumo**	**UCs**	**Adotantes**	**Potência Acumulada [MW]**	**UCs**	**Adotantes**	**Potência Acumulada [MW]**	**Crescimento Adotantes 2025-2020**
B1	Residencial	3.595.837	15.192	144,33	3.970.660	88.523	554,24	483%
B2	Rural	341.933	2.387	61,76	323.966	5.956	160,36	150%
B3	Comercial	401.229	3.900	82,95	435.401	10.256	216,33	163%
B3	Industrial	66.224	815	22,03	53.880	1.243	30,19	53%
A4A3a	Rural	612	30	3,33	663	108	12,38	260%
A4A3a	Comercial	5.993	186	18,64	5.768	910	112,36	389%
A4A3a	Industrial	4.307	149	18,93	3.586	597	86,15	301%

3.1.2 Difusão espacial da GDFV

Como tratado na seção de metodologia, o modelo de difusão espacial proposto utiliza como ponto de partida o agrupamento dos adotantes de todos os segmentos de mercado elencados na TABELA 2 nas subestações da COPEL-DIS, a partir do cruzamento de dados do EGRID e de cadastro de GDs. Na sequência, os novos adotantes previstos no modelo de difusão temporal (no ano de 2025) são alocados de maneira proporcional à taxa de penetração calculada em 2020, de maneira a reproduzir as bases do conceito de transbordamento tecnológico. Na FIGURA 6, apresenta-se um comparativo da distribuição dos adotantes de GDFV no estado do Paraná a nível das subestações em cenário atual (2020) e futuro (2025), com identificação das regionais de planejamento. No mapa, o tamanho das bolhas corresponde ao número de adotantes alocados a cada subestação.

Da análise do cenário atual, destaca-se a concentração de adotantes nos principais centros urbanos do estado – como Curitiba (Leste) e Londrina (Norte) –, ilustrando o vínculo do acesso à informação e de critérios socioeconômicos aos processos de difusão tecnológica. Além disso, o elevado potencial solarimétrico nas regionais Norte, Noroeste e principalmente Oeste justificam a relevante quantidade de empreendimentos nessas porções do estado. Para o ano de 2025, as alterações na FIGURA 6 são sensíveis, mas suficientes para indicar o crescimento no número de adotantes em todas as subestações, com a manutenção das regiões que centralizam os prossumidores.

A forte expansão da GDFV fica mais evidente com a apresentação das taxas de penetração (adotantes por unidades consumidoras) nos dois horizontes avaliados. Na FIGURA 7, esse indicador é levantado após a alocação das GDs atuais e projetadas a nível dos alimentadores. Do histograma, clarifica-se o crescimento do número de alimentadores com taxas de penetração mais elevadas, limitadas a um fator de saturação de 8,17% (KONZEN, 2014). De maneira interessante e condizente à métrica aplicada, reforça-se que as subestações de penetração saturada já apresentavam taxas de penetração superiores à média no ano de 2020. Por fim, salienta-se que somente a taxa de penetração projetada para 2025 é utilizada como variável representativa dos estudos de projeção da GDFV para o modelo de clusterização desenvolvido.

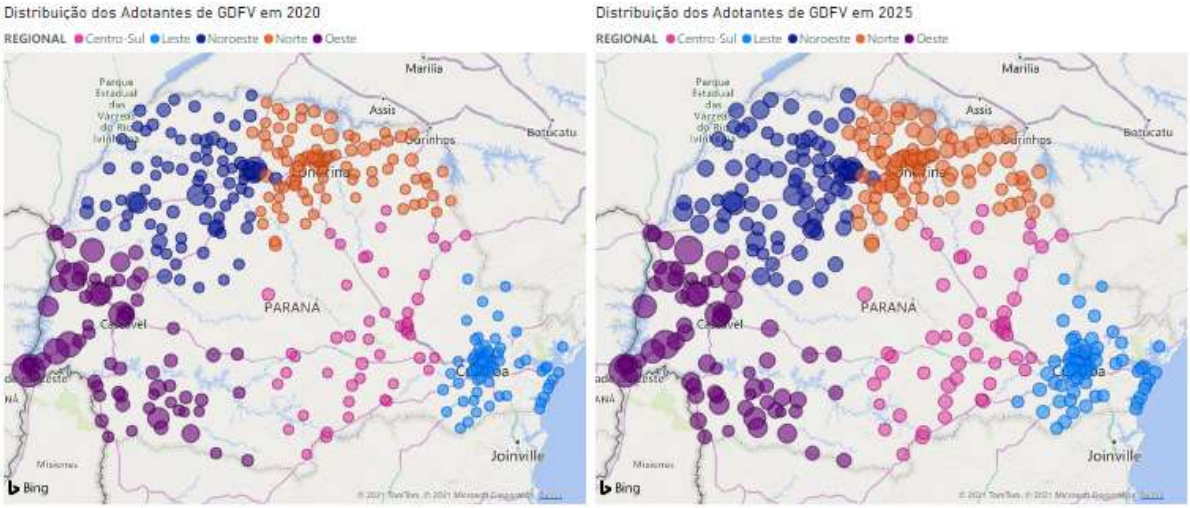

FIGURA 6: Distribuição dos adotantes de GDFV no estado do Paraná em 2020 e 2025

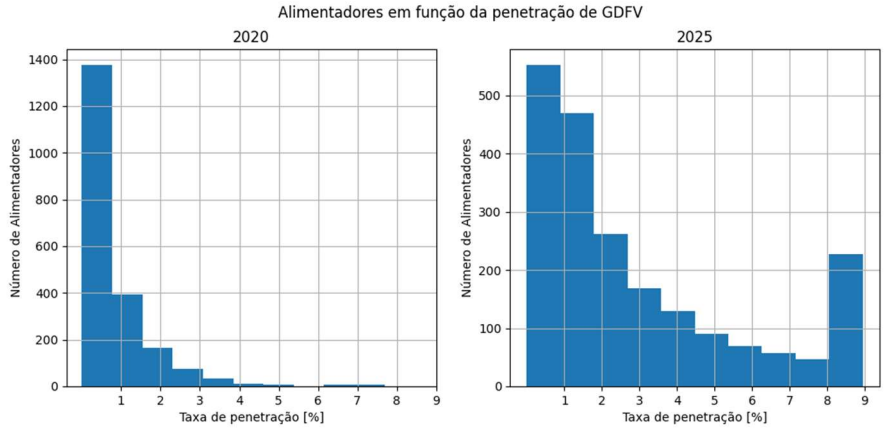

FIGURA 7: Histograma dos alimentadores em função da taxa de penetração em 2020 e 2025

3.2 Estudo de clusterização

Na etapa de pré-processamento, as variáveis classificatórias são selecionadas em função de uma análise de correlação, utilizando o índice de Pearson, que mensura a correlação linear entre as variáveis duas a duas, com coeficientes que variam entre -1 e 1. Assim, permite-se a eliminação de redundâncias dos dados, que dificultariam o processo de clusterização, a partir da classificação baseada apenas nos atributos de trechos totais, número de unidades consumidoras, demanda máxima, penetração de GDFV em 2025, queda de tensão e DEC.

Para finalizar a etapa de pré-processamento, aplica-se o algoritmo de t-SNE para redução do problema para duas dimensões e seleciona-se o número ótimo de 5 clusters por inspeção visual da curva de SSE. Executando-se a clusterização pelo método das k-médias, os agrupamentos podem ser visualizados em duas dimensões na FIGURA 8. Ressalta-se, de maneira importante, que este resultado pode ser levantado graças à aplicação prévia do algoritmo t-SNE. Por fim, com o objetivo de gerar insumos às análises de aplicação de obras e estudo de *hosting capacity*, a TABELA 3 resume os resultados das variáveis descritivas dos alimentadores representativos de cada cluster.

TABELA 3: Resultados dos alimentadores representativos de cada cluster

Cluster	SED-SDMT	Trechos [km]	UC	Demanda Máxima [MVA]	Penetração GDFV [%]	Queda Tensão [%]	DEC [h]
0	SED-055-SDMT-05	11,1	2.112	2,9	0,56	1,01	0,15
1	SED-099-SDMT-06	11,7	4.395	5,2	0,72	1,88	0,46
2	SED-071-SDMT-08	113,1	1.018	4,3	2,49	7,64	2,30
3	SED-067-SDMT-04	10,9	1.146	7,4	1,50	2,46	0,04
4	SED-287-SDMT-08	38,2	6.141	6,6	0,58	4,25	0,20

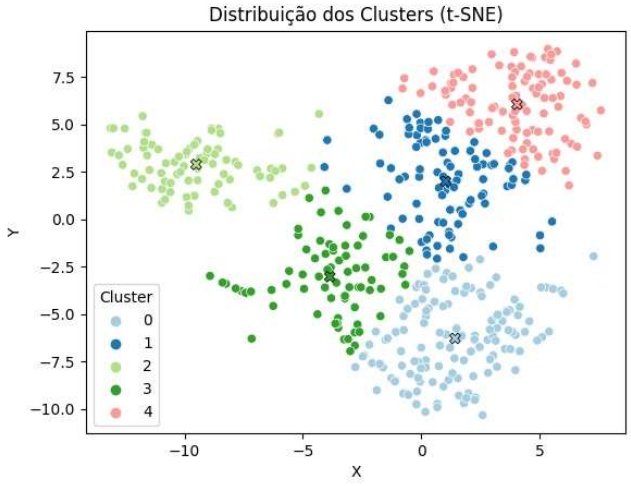

FIGURA 8: Distribuição dos alimentadores nos clusters, com identificação do centroide dos agrupamentos

3.3 Análise multicritério

Os resultados do diagnóstico técnico e do estudo de *hosting capacity* antes e após a realização das obras propostas no alimentador representativo do cluster 2 são resumidos na TABELA 4. No caso da obra de divisão de tronco, identificam-se os resultados dos dois alimentadores resultantes. Assim como indicado na seção 2.6, a *hosting capacity* é mensurada pela média das potências mínimas acomodadas em cada barra do alimentador.

TABELA 4: Resultados do diagnóstico técnico e de *hosting capacity* antes e após as obras

Cluster	SED-SDMT	Obra	Carreg.Max [%]	Queda Tensão [%]	Perdas [MWh]	HC [MW]
2	SED-071-SDMT-08	Original	44,68	7,64	39,7	1,32
		Recondutoramento	40,62	5,67	32,1	1,15
		Divisão do Tronco	45,17	6,08	19,9	1,71
			22,16	1,92	6,5	5,03

Dessa forma, o benefício técnico de cada uma das obras é calculado, indicando-se a obra prioritária na TABELA 5.

TABELA 5: Benefício técnico das obras para rede representante

Cluster	SED-SDMT	Recondutoramento	Divisão do Tronco	Prioridade
2	SED-071-SDMT-08	0,580	2,234	Divisão do Tronco

Em função da similaridade das redes dentro de um mesmo cluster, sugere-se a extrapolação da priorização para os demais alimentadores. Selecionando algumas redes deste agrupamento e efetuando o cálculo de benefício técnico é possível verificar a validade do modelo proposto (TABELA 6).

TABELA 6: Benefício técnico das obras para redes representadas

Cluster	SED-SDMT	Recondutoramento	Divisão do Tronco	Prioridade
2	SED-249-SDMT-01	2,047	2,216	Divisão do Tronco
2	SED-019-SDMT-01	3,455	5,378	Divisão do Tronco
2	SED-069-SDMT-01	1,618	1,805	Divisão do Tronco

Ressalta-se que a alocação exaustiva de obras em todas as redes com o posterior cálculo de benefício técnico pode (e deve) indicar casos em que a priorização obtida via alimentador representativo não é a adequada, o que é justificado pela complexidade das redes trabalhadas, que podem não ser completamente separáveis via processo de clusterização (incorrendo na existência de *outliers*, por exemplo).

4.0 AGRADECIMENTOS

Este trabalho recebeu financiamento e suporte técnico da Copel - Companhia Paranaense de Energia, no âmbito do projeto “PD-02866-0527/2020 Metodologias e ferramental para planejamento automatizado inteligente de redes de distribuição considerando recursos energéticos distribuídos”, que é um programa P&D regulado pela ANEEL.

5.0 CONCLUSÃO

Neste trabalho foi apresentada uma metodologia para caracterização das redes de distribuição e avaliação de novas obras, considerando o impacto de recursos energéticos distribuídos com estudo de caso utilizando dados da COPEL-DIS. No modelo desenvolvido, realizou-se o agrupamento das redes por atributos, a partir da aplicação de técnicas de *machine learning* (análise de correlação para seleção de atributos não redundantes, redução de dimensionalidade via t-SNE e clusterização por k-médias), utilizando a projeção de penetração de GDFV como uma das variáveis classificatórias.

A partir de indicadores técnicos selecionados – entre eles a HC –, foi desenvolvido um modelo de avaliação multicriterial para priorização de obras aplicadas aos alimentadores representativos, cujos resultados podem ser extrapolados às demais redes de cada cluster. Dessa forma, ao se propor a restrição da realização de estudos mais detalhados – como diagnósticos técnicos e estudos de HC – a um número limitado de redes representativas, permite-se a redução de tempo computacional dispendido no processo de planejamento da distribuidora.

6.0 REFERÊNCIAS

[1] KONZEN, G. **Difusão de sistemas fotovoltaicos residenciais conectados à rede no Brasil: uma simulação via modelo de Bass**. 2014. 109 f. Dissertação de Mestrado (Pós-graduação em Energia) – Instituto de Energia e Meio Ambiente da Universidade de São Paulo, São Paulo, 2019. [Orientador: Roberto Zilles].

[2] BALTA-OZKAN, N.; YILDIRIM, J.; CONNOR, P. **Regional distribution of photovoltaic deployment in the UK and its determinants: A spatial econometric approach**. Energy Economics, v. 51, p. 417-429, 2015. 10.1016/j.eneco.2015.08.003.

[3] DHARSHING, S. **Household dynamics of technology adoption: A spatial econometric analysis of residential solar photovoltaic (PV) systems in Germany**. Energy research & social science, v. 23, p. 113-124, 2017. 10.1016/j.erss.2016.10.012.

[4] ISMAEL, S. M. ABDEL ALEEM, S. H. E.; ABDELAZIZ, A. Y.; ZOBAA, A. F. **State-of-the-art of hosting capacity in modern power systems with distributed generation**. Renewable Energy, vol. 130, p. 1002–1020, jan. 2019, doi: 10.1016/j.renene.2018.07.008.

[5] RYLANDER, M.; SMITH, J.; SUNDERMAN, W. **Streamlined method for determining distribution system hosting capacity**. In: 2015 IEEE Rural Electric Power Conference. IEEE, 2015a. p. 3-9.

[6] EPE. Empresa de Pesquisa Energética. **Nota Técnica EPE DEA-SEE 009/2021**. Modelo de Mercado da Micro e Minigeração Distribuída (4MD): Metodologia – Versão PDE 2031. Brasília, p. 27, 2021.

[7] AL-SAFFAR, M.; ZHANG, S.; NASSIF, A.; MUSILEK, P. **Assessment of Photovoltaic Hosting Capacity of Existing Distribution Circuits**. In: 2019 IEEE Canadian Conference of Electrical and Computer Engineering (CCECE). IEEE, 2019. p. 1-4.

[8] DUBEY, A.; SANTOSO, S.; MAITRA, A. **Understanding photovoltaic hosting capacity of distribution circuits**. In: 2015 IEEE power & energy society general meeting. IEEE, 2015. p. 1-5.

DADOS BIOGRÁFICOS

Possui graduação em Engenharia Elétrica - Ênfase em Sistemas de Energia e Automação pela Universidade de São Paulo (2018). Atualmente trabalha como consultora para a Sinapsis Inovação Energia, CEEAC e Softex Recife. Possui experiência nas áreas de Planejamento da Distribuição e Cálculo de Perdas, incluindo tratamento de dados de redes elétricas para simulações, cálculo de perdas e análise de dados, com aplicação de técnicas avançadas, machine learning e bussiness intelligence.

(2) DANIEL PEREZ DUARTE
Graduado em Engenharia Elétrica com ênfase em Sistemas de Potência pela Universidade Mackenzie (2003), Mestrado em Engenharia Elétrica pela Universidade de São Paulo (2008) e Doutorado em Engenharia Elétrica também pela Universidade de São Paulo (2013). Experiência consolidada em consultoria em sistemas de distribuição, particularmente nas áreas de Smart Grids, Planejamento, Automação Avançada da Distribuição (ADA), Sistemas de Proteção, bem como em projetos de sistemas de proteção, controle e serviços auxiliares de usinas e subestações. Atualmente é Diretor Executivo da Sinapsis Inovação em Energia, empresa especializada em fornecer e adaptar soluções inovadoras ao mercado de energia.

(3) SÉRGIO MISHIMA DOS SANTOS BARBOSA
Possui graduação em Engenharia Elétrica - Ênfase em Energia e Automação pela Escola Politécnica da Universidade de São Paulo (2020), com duplo diploma pela CentraleSupélec (2017-2019) e Master Recherche pela Université Paris-Saclay (2018-2019). Atua como consultor para a Sinapsis Inovação em Energia. Possui experiência nas áreas de despacho de energia em microgrids isoladas, sistemas de proteção industriais e planejamento da distribuição, incluindo a concepção de modelos de difusão tecnológica de recursos energéticos distribuídos.

(4) REGINALDO PRESTES
Possui graduação em Engenharia Industrial Elétrica pela Universidade Tecnológica Federal do Paraná (2001). Engenheiro de Planejamento Senior na Copel Distribuição.

(5) CLEVERSON LUIZ DA SILVA PINTO
Possui graduação em pela Universidade Federal de Santa Catarina (1990), mestrado em Engenharia Elétrica pela Universidade de São Paulo (1999) e doutorado em Engenharia Elétrica pela Universidade de São Paulo (2008). Atualmente é engenheiro eletricista consultor da Companhia Paranaense de Energia - COPEL DIS e professor

adjunto no curso de engenharia elétrica da UFPR, em Curitiba. Tem experiência na área de Engenharia Elétrica, com ênfase no planejamento de sistemas de distribuição de média tensão e geração distribuída.