

UNIVERSIDADE FEDERAL DE ITAJUBÁ

PROGRAMA DE PÓS-GRADUAÇÃO EM
ENGENHARIA ELÉTRICA

Paulo Vitor Grillo de Souza

**Metodologia para Estabelecimento de Novos
Fatores de Impacto Base Considerando
Agrupamentos de Sistemas de Distribuição**

Itajubá, março de 2021

UNIVERSIDADE FEDERAL DE ITAJUBÁ

PROGRAMA DE PÓS-GRADUAÇÃO EM
ENGENHARIA ELÉTRICA

Paulo Vitor Grillo de Souza

**Metodologia para Estabelecimento de Novos
Fatores de Impacto Base Considerando
Agrupamentos de Sistemas de Distribuição**

Tese submetida ao Programa de Pós-Graduação em Engenharia Elétrica como parte dos requisitos para obtenção do título de **Doutor em Ciências em Engenharia Elétrica.**

Área de Concentração: Sistemas Elétricos de Potência

Orientador: Prof. Dr. José Maria de Carvalho Filho - UNIFEI

Coorientador: Prof. Dr. Daniel Furtado Ferreira - UFLA

Itajubá, março de 2021

DEDICATÓRIA

Aos meus pais, Paulo e Roseli.

“Cada sonho que você deixa para trás, é um pedaço do seu futuro que deixa de existir.”

Steve Jobs

AGRADECIMENTOS

A Deus, por todas as oportunidades que colocou em meu caminho, pela capacidade que me confiou, por me dar a saúde e força necessária para atingir os objetivos.

Aos meus pais, Paulo e Roseli, que me educaram com muito amor, me deram todo o apoio e depositaram confiança em mim para que eu atingisse este estágio da minha vida.

As minhas irmãs, Karina e Jessica, que apesar da distância, sempre me deram apoio.

A minha namorada, Fabiana, pelo companheirismo e compreensão.

Ao professor, José Maria, pela orientação e pela prontidão em ajudar.

Ao professor, Daniel Furtado Ferreira, pela coorientação, e auxílio em assuntos relacionados à estatística.

A todos os mestres que me ensinaram e contribuíram com a minha evolução ao longo da minha vida.

Aos meus amigos que proporcionaram momentos de descontração essenciais.

À EDP pelo fornecimento dos dados utilizados na pesquisa e pelo suporte financeiro.

À UFLA e à UNIFEI pelo apoio estrutural na pesquisa.

SUMÁRIO

| | |
|--|-----------|
| DEDICATÓRIA..... | I |
| AGRADECIMENTOS..... | III |
| SUMÁRIO..... | IV |
| LISTA DE FIGURAS..... | VI |
| LISTA DE TABELAS..... | VI |
| LISTA DE ABREVIATURAS E SIGLAS..... | VII |
| RESUMO..... | VIII |
| ABSTRACT..... | IX |
| 1. INTRODUÇÃO | 1 |
| 1.1 Considerações Iniciais | 1 |
| 1.2 Relevância do tema | 2 |
| 1.3 Estado da Arte..... | 4 |
| 1.4 Objetivos e Contribuições da Tese | 10 |
| 1.5 Estrutura da Tese | 11 |
| 2. REVISÃO DE CONCEITOS | 13 |
| 2.1 Considerações Iniciais | 13 |
| 2.2 Análise de Regressão Múltipla | 13 |
| 2.3 Análise de Agrupamentos (Método Dinâmico) | 18 |
| 2.4 VTCDs e Indicadores..... | 21 |
| 2.5 Considerações Finais | 25 |
| 3. METODOLOGIA PARA ESTABELECIMENTO DO FATOR DE IMPACTO | |
| BASE..... | 27 |
| 3.1 Considerações Iniciais | 27 |
| 3.2 Escolha da Base de Dados | 27 |
| 3.3 Seleção das Variáveis | 28 |
| 3.4 Formação de Conjuntos | 29 |
| 3.5 Estabelecimento do Fator de Impacto base..... | 29 |
| 3.6 Fluxograma | 30 |
| 3.7 Considerações Finais | 32 |
| 4. ESTUDO DE CASO E RESULTADOS | 33 |
| 4.1 Considerações Iniciais | 33 |
| 4.2 Escolha da Base de Dados | 33 |
| 4.3 Seleção das Variáveis | 35 |

| | | |
|-----------|--|-----------|
| 4.4 | Formação de Conjuntos | 40 |
| 4.5 | Estabelecimento do Fator de Impacto base..... | 42 |
| 4.6 | Considerações Finais | 47 |
| 5. | CONCLUSÕES..... | 48 |
| 5.1 | Conclusões Gerais..... | 48 |
| 5.2 | Proposta para Trabalhos Futuros | 49 |
| 5.3 | Artigos Publicados | 50 |
| | REFERÊNCIAS | 51 |
| | APÊNDICE A – Tabelas de Distância e Heterogeneidade | 54 |
| | APÊNDICE B – Tabela da Frequência de AMTs | 59 |

LISTA DE FIGURAS

| | |
|--|----|
| Figura 2.1 – Fluxograma da estimação <i>stepwise</i> | 17 |
| Figura 2.2 – Semelhança entre elementos na fronteira de grupos diferentes. | 20 |
| Figura 3.1 – Fluxograma da metodologia proposta. | 31 |
| Figura 4.1 – Papel de probabilidade. | 38 |
| Figura 4.2 – Gráfico dos resíduo x valores ajustados. | 39 |
| Figura 4.3 – Gráfico resíduos x ordem de coleta. | 39 |
| Figura 4.4 – Comparação do Fator de Impacto novo e atual com o valor de referência. | 47 |

LISTA DE TABELAS

| | |
|---|----|
| Tabela 2.1 – Tabela de ANOVA | 15 |
| Tabela 2.2 – Classificação das Variações de Tensão de Curta Duração (PRODIST). | 22 |
| Tabela 2.3 – Estratificação dos parâmetros amplitude e duração para contabilização de eventos de VTCD. | 23 |
| Tabela 2.4 – Estratificação das VTCDs com base nos níveis de sensibilidade das diversas cargas. | 24 |
| Tabela 2.5 – Limites mensais de ocorrência por região de sensibilidade. | 25 |
| Tabela 2.6 – Fatores de ponderação e Fator de Impacto Base. | 25 |
| Tabela 4.1 – Base de dados. | 34 |
| Tabela 4.2 – Resultados da regressão <i>stepwise</i> (base de dados com 7 variáveis). | 36 |
| Tabela 4.3 – Resultados da regressão <i>stepwise</i> (base de dados com 9 variáveis). | 37 |
| Tabela 4.4 – Distância e heterogeneidade entre SDs (SD 8 como referência). | 41 |
| Tabela 4.5 – Frequência de AMTs nas regiões de sensibilidade (SD 8 como referência). | 42 |
| Tabela 4.6 – Fatores de ponderação e limites para a frequência de AMTs em cada região de sensibilidade (percentil 20). | 43 |
| Tabela 4.7 – FI_{base} e FI dos sistemas de distribuição (percentil 20). | 43 |
| Tabela 4.8 – Fatores de ponderação e limites para a frequência de AMTs em cada região de sensibilidade (média). | 44 |
| Tabela 4.9 – FI_{base} e FI dos sistemas de distribuição (média). | 45 |
| Tabela 4.10 – FI_{base} e FI dos sistemas de distribuição (regulamentação atual). | 46 |

LISTA DE ABREVIATURAS E SIGLAS

| | |
|---------|---|
| ABRACE | Associação Brasileira de Grandes Consumidores Industriais de Energia e de Consumidores Livres |
| APE | Associação Brasileira de Distribuidores de Energia Elétrica |
| AMT | Afundamento Momentâneo de Tensão |
| ANEEL | Agência Nacional de Energia Elétrica |
| AP | Audiência Pública |
| CP | Consulta Pública |
| DEC | Duração Equivalente de Interrupção por Unidade Consumidora |
| EDP | Energias de Portugal |
| FEC | Frequência Equivalente de Interrupção por Unidade Consumidora |
| IEC | <i>International Electrotechnical Commission</i> |
| IEEE | <i>Institute of Electrical and Electronics Engineers</i> |
| ONS | Operador Nacional do Sistema Elétrico |
| PCA | <i>Principal Component Analysis</i> |
| P&D | Pesquisa e Desenvolvimento |
| PRODIST | Procedimentos de Distribuição |
| p.u. | por unidade |
| QEE | Qualidade da Energia Elétrica |
| SD | Sistema de Distribuição |
| SED | Subestação de Distribuição |
| SRD | Superintendência de Regulação dos Serviços de Distribuição |
| UNIFEI | Universidade Federal de Itajubá |
| VTCD | Variação de Tensão de Curta Duração |

RESUMO

As variações de tensão de curta duração (VTCDs) são eventos que afetam a qualidade da energia elétrica e que podem causar grandes prejuízos financeiros aos consumidores industriais que possuem cargas sensíveis a este distúrbio, além de prejudicar a segurança e a confiabilidade dos processos. A Agência Nacional de Energia Elétrica (ANEEL), reconhecendo a importância do assunto, estabeleceu por meio dos Procedimentos de Distribuição de Energia Elétrica no Sistema Elétrico Nacional (PRODIST) o indicador Fator de Impacto para quantificar o desempenho dos barramentos de um sistema de distribuição sob o ponto de vista das VTCDs. Todavia, o fator de impacto base proposto na regulamentação é único, não considerando as diferenças que os sistemas de distribuição possuem com relação aos potenciais causadores do evento.

Neste contexto, esta tese apresenta uma metodologia para o estabelecimento de fatores de impacto base distintos para cada sistema de distribuição. O trabalho é focado em determinar quais variáveis melhor descrevem o distúrbio e, a partir destas, estabelecer agrupamentos que permitam uma definição de fatores de impacto base mais adequados, considerando as especificidades dos sistemas de distribuição.

Para testar a metodologia proposta, utilizaram-se dados reais de uma concessionária do sistema elétrico brasileiro. Foram monitoradas, por meio de medidores de qualidade da energia elétrica, dezenove barras de média tensão em subestações de distribuição.

A metodologia apresentada é genérica e pode ser aplicada em qualquer sistema de distribuição do Brasil, ou até mesmo de outros países que venham a adotar indicador semelhante ao Fator de Impacto.

Os resultados obtidos no estudo de caso mostraram que, com as variáveis selecionadas e com os agrupamentos formados, é possível estabelecer fatores de impacto base distintos para cada sistema de distribuição, que levam em consideração o desempenho dos sistemas semelhantes.

A aplicação da metodologia traz benefícios para o setor elétrico, principalmente para os consumidores que possuem cargas sensíveis as variações de tensão de curta duração, uma vez que estabelece fatores de impacto base que incentivam a melhora da qualidade da energia elétrica.

Palavras-Chave: Qualidade da energia, Afundamento de tensão, Formação de Conjuntos, Indicadores, Fator de Impacto Base.

ABSTRACT

Short-term voltage variations (STVVs) are events that affect the power quality and can cause great financial losses to industrial consumers who have loads sensitive to this disturbance, in addition to impairing the safety and reliability of processes. ANEEL, recognizing the importance of the matter, established through PRODIST the Impact Factor index to quantify the performance of the buses of a distribution system from the point of view of the STVVs. However, the base impact factor proposed in the regulation is unique, not considering the differences that the distribution systems have in relation to the potential cause of the event.

In this context, this thesis presents a methodology for establishing different base impact factors for each distribution system. The work is focused on determining which variables best describe the disturbance and from these, establish clusters that allow a definition of more suitable base impact factors, considering the specificities of the distribution systems.

To test the proposed methodology, real data from a concessionaire of the Brazilian electrical system was used. Nineteen substation bars were monitored through power quality meters.

The methodology presented is generic and can be applied in any distribution system in Brazil, or even in other countries that may adopt an index similar to the impact factor.

The results obtained in the case study showed that with the selected variables and the formed groups, it was possible to establish different base impact factors for each distribution system, which take into account the performance of similar systems.

The application of the methodology brings benefits to the electricity sector, especially for consumers who have loads that are sensitive to short-term voltage variations, since it establishes basic impact factors that promote the improvement of the power quality.

Keywords: Power Quality, Voltage Sag, Clustering, Indexes, Base Impact Factor.

1. INTRODUÇÃO

1.1 Considerações Iniciais

Devido aos avanços tecnológicos, pautados sempre na melhoria da eficiência produtiva dos processos industriais e do bem estar das pessoas, os dispositivos eletroeletrônicos tiveram grande inserção no ambiente residencial, comercial e principalmente no setor fabril. Todavia, em geral estes equipamentos baseados em eletrônica possuem maior sensibilidade aos distúrbios que afetam a qualidade da energia elétrica (QEE). Assim, cria-se por parte dos consumidores a consciência de que tais eventos, que no passado poderiam não acarretar impacto algum às cargas, atualmente são críticos e de extrema relevância.

Os eventos relacionados à QEE podem ser entendidos como qualquer variação significativa na tensão, corrente ou frequência, que ocasione uma operação imprópria do equipamento do consumidor, podendo ser classificados em (DUGAN et al., 2012):

- Transitórios impulsivos e oscilatórios;
- Variações de tensão de longa duração (sobretensão, subtensão e interrupções);
- Variações de tensão de curta duração (afundamentos de tensão, elevações de tensão e interrupções de curta duração);
- Distorções harmônicas de tensão e de corrente;
- Desequilíbrios de tensão;
- Flutuações de tensão;
- Variações de frequência.

Esta tese tem foco nas variações de tensão de curta duração, mais especificamente os afundamentos de tensão.

Outro fator que pode ser visto como um fomentador no interesse a QEE é a mudança ocorrida no setor elétrico brasileiro, com a privatização das empresas de energia. Pois, devido a este novo modelo, o mercado de energia tornou-se mais competitivo, favorecendo o consumidor em relação a preços e, sobretudo com relação à qualidade do produto ofertado, uma vez que as concessionárias devem atender aos padrões e procedimentos estabelecidos pela ANEEL.

Para os sistemas de distribuição, o papel de órgão regulador é desempenhado pela

ANEEL, que por meio do PRODIST, define normas e padrões para as atividades técnicas relativas ao funcionamento e desempenho das redes de distribuição.

Tendo ciência da importância da QEE, a ANEEL vem desde a elaboração da primeira versão do Módulo 8 do PRODIST, que ocorreu em 2008, aprimorando este documento com ampla participação da comunidade, que por meio de consultas e audiências públicas podem expressar de forma fundamentada as opiniões de sua classe e os anseios quanto a evolução das normas. Neste processo, o módulo 8 já está em sua 12ª versão, buscando regulamentar os principais distúrbios que afetam a QEE.

As VTCDs são a classe de eventos que possui a regulamentação mais recente e ainda incipiente, uma vez que foram estabelecidos na revisão 8, em 2017, o indicador Fator de Impacto (FI) juntamente com seu o valor de referência. Resta ainda a ANEEL estabelecer limites e eventualmente critérios para penalização caso ocorra transgressão por parte das concessionárias.

Sendo assim, neste trabalho é proposta uma metodologia para o estabelecimento de FI_{base} distintos de acordo com as especificidades de cada sistema de distribuição, contribuindo para a discussão e o aprimoramento do indicador FI, cujos benefícios resultantes vem ao encontro às demandas do setor elétrico, sobretudo dos consumidores industriais que possuem cargas sensíveis.

1.2 Relevância do tema

Quando há ocorrência de uma VTCD no sistema elétrico, alguns equipamentos industriais podem apresentar mal funcionamento que comprometa o processo produtivo, ou em casos extremos, até mesmo cessar por completo seu funcionamento, provocando, conseqüentemente, uma parada parcial ou total do processo. Seja qual for o tipo de interrupção ocorrida no processo industrial, existem sempre atrelados a ela os prejuízos devido à perda de produtividade, perda de insumos, reparo e reposição de equipamentos danificados (CARVALHO FILHO, 2000).

Embora as VTCDs provoquem prejuízos econômicos para as indústrias, pela normalização vigente, as concessionárias de energia elétrica não são penalizadas se o consumidor industrial sofrer paradas de processo (ANEEL, 2021).

Devido às perdas financeiras enfrentadas pelo setor industrial resultantes das ocorrências de VTCDs, a ANEEL tem sido pressionada pelos consumidores e suas

associações, para o estabelecimento de limites mais apropriados para o indicador de VTCDs e que reflitam de fato a qualidade da energia elétrica almejada pelos consumidores industriais.

Em 2017 iniciou-se a vigência da revisão 8 do módulo 8 do PRODIST, trazendo evolução significativa em relação às versões anteriores, por apresentar uma metodologia própria para avaliar as VTCDs, pela criação de um novo indicador de QEE chamado Fator de Impacto e pela proposição de um valor de referência único para este indicador. Os estudos que ampararam as modificações da norma foram realizados por um grupo de pesquisa da Universidade Federal de Uberlândia (UFU), que foi contratado em 2013 para prestar consultoria a Superintendência de Regulação dos Serviços de Distribuição (SRD), além de serem recebidas contribuições por meio da Consulta Pública (CP N°18/2014).

Um dos assuntos de maior controvérsia no âmbito da CP n°18/2014, foi justamente o valor de referência de 1 p.u. sugerido para o indicador FI, uma vez que diversas contribuições se posicionaram opostas ao estabelecimento deste valor. Argumentou-se que os dados de medição usados na pesquisa não eram suficientes, pois não correspondiam a um ano completo de medição, ou seja, a característica sazonal do evento não estaria sendo representada, somado a pouca variedade de concessionárias consideradas na análise.

A ABRADDEE (Associação Brasileira de Distribuidores de Energia Elétrica) em sua contribuição enviada salientou que deveria haver um progresso no assunto por meio de medições permanentes nas SEDs (Subestações de Distribuição), visto que esta sugestão proporcionaria um amadurecimento gradual e equilibrado do tema dentro do setor de distribuição. Enquanto a ABRACE (Associação Brasileira de Grandes Consumidores Industriais de Energia e de Consumidores Livres) acenou que o valor de referência proposto é muito brando, não satisfazendo a real necessidade do setor industrial.

Apesar das argumentações contrárias recebidas nas contribuições, a SRD, por meio da Audiência Pública n° 0082/2015, tornou clara a necessidade de definir valores de referência para o indicador para se aperfeiçoar o PRODIST e adverte que refinamentos continuamente serão necessários, mas a busca pela perfeição não deve impedir o avanço do tema.

A questão é como instituir adequadamente limites para o indicador das VTCDs, visto que o Brasil possui grande extensão territorial, é um dos países com maior sistema elétrico interligado, tem muita diversidade em relação à vegetação, clima e relevo das regiões. Embora reconheça o mérito dos avanços realizados na revisão 8 do módulo 8 do PRODIST, a proposição de um único valor de referência para o indicador FI, indistintamente para todos os

sistemas de distribuição no Brasil, certamente não é adequado e coerente tanto do ponto de vista das distribuidoras como dos consumidores industriais.

É neste contexto de aprimoramento do módulo 8 do PRODIST que esta tese se insere, propondo uma metodologia para o estabelecimento de Fatores de Impacto base distintos, considerando agrupamentos de sistemas de distribuição que possuam características semelhantes com relação às variáveis que influenciam a ocorrência de VTCDs.

1.3 Estado da Arte

Feito um levantamento nos principais acervos de pesquisas no campo da engenharia elétrica, foram encontrados dois artigos e uma dissertação, diretamente relacionados com o tema desta tese e alguns artigos que utilizam técnicas de formação de conjuntos para caracterizar os fenômenos relacionados à qualidade da energia elétrica. Após uma seleção rigorosa dos artigos que empregam técnicas próximas às utilizadas na metodologia proposta nesta tese, fez-se uma síntese de cada um destes trabalhos, de forma a apresentar como as técnicas estatísticas de formação de conjuntos estão sendo empregadas na área de QEE.

Em (JASIŃSKI et al., 2016) propõe-se um método para avaliação dos eventos de QEE considerando diferentes condições operativas da rede. Em geral, os métodos convencionais de avaliação se baseiam em um período representativo (normalmente uma semana) no qual se pressupõe condições normais de operação do sistema. Este tipo de análise não é suficiente e não representa o real estado do trecho da rede elétrica considerada, pois, os dados medidos podem depender, para o caso específico da QEE, das mudanças de carregamento, de geração, bem como das diferentes configurações da rede. Por esta razão, o autor do trabalho utiliza as técnicas de formação de conjuntos, para dividir os dados em grupos que ponderem às condições operativas. Especificamente, foram estabelecidos 4 grupos, considerando o perfil de carga (distinguindo entre dias úteis e dias não úteis) e o uso ou não de uma cogeração de 100MW. Para confirmação da metodologia foram utilizados dados reais referentes a três semanas de medições feitas em subestações de distribuição. Os resultados possibilitaram definir a influência das condições operativas do sistema na QEE. Ressalta-se a necessidade de uma análise contínua, por intermédio da formação de conjuntos, juntamente com uma ferramenta que a relacione ao processo de tomada de decisão. É importante reconhecer que, o método proposto utilizando técnicas de formação de conjuntos, pode servir como um complemento para a maneira clássica de avaliação da QEE.

No trabalho (FLORENCIAS-OLIVEROS et al., 2017), foi desenvolvida uma técnica

fundamentada na análise gráfica de agrupamentos, para ser implementada em um analisador de QEE inteligente, com o objetivo de monitorar redes elétricas, onde novos eventos surgem como uma consequência da penetração de geração distribuída. Na presença de uma falta, caracterizada pela transgressão de um limiar previamente selecionado, o equipamento inicia o procedimento de medição e a estatística de ordem superior é calculada no domínio do tempo para permitir a classificação. Os resultados alcançados por meio de teste em um montante de 160 sinais mostraram a divisão em dois grupos de eventos (AMTs e transitórios), com uma precisão de 80%. A análise das intensidades das perturbações dentro do grupo auxilia o operador a caracterizar melhor a base de dados e introduz um critério qualitativo para monitoramento das redes elétricas inteligentes.

O artigo (DUAN et al., 2015) apresenta um algoritmo que utiliza o método de formação de conjuntos *k-means* para reconhecer e classificar os afundamentos de tensão, de dados de medição de uma grande rede elétrica em Shenzhen na China. As distâncias entre diferentes ocorrências de afundamentos de tensão distribuídos em um diagrama (duração x amplitude) são calculadas primeiro. Quando se observa proximidade entre ocorrências, um centróide para o grupo é selecionado para representar estes eventos, em seguida, os valores e locais dos centróides são determinados com base em um processo iterativo. A amplitude e duração do AMT representativo do centróide podem ser consideradas como as características do afundamento de tensão esperado em subestações semelhantes, que representará a condição operativa e permitirá encontrar o elo fraco do sistema elétrico de potência. Os resultados mostraram que praticamente todos os distúrbios de AMTs podem ser classificados em 11 conjuntos que provavelmente representam as características e causas da maioria dos eventos ocorridos em sistemas de distribuição típicos. Resumindo, o método pode efetivamente agrupar toda a massa de dados de medição de AMTs em alguns padrões típicos, simplificando a carga de trabalho na análise da QEE.

Em muitos algoritmos, como *k-means* e *fuzzy c-mean*, o número esperado de grupos deve ser definido com antecedência, como um dado de entrada do algoritmo. Outras metodologias de formação de conjuntos estimam este valor no decorrer do processo utilizando técnicas heurísticas, no entanto, algumas destas podem levar a um mínimo local do espaço de soluções, sem encontrar o verdadeiro número ótimo de grupos. Neste contexto, em (ASHEIBI; STIRLING; SUTANTO, 2008), é apresentado um método desenvolvido para determinar o número ótimo de grupos a serem formados, em dados de medição de QEE, usando um algoritmo de mineração de dados baseado na técnica MML (*minimum message*

length). Para testar o método proposto, utilizaram-se 3 bases de dados distintas, uma criada com dados aleatórios onde se tinha previamente a informação do número ideal de grupos, outra com dados de simulação de um sistema de potência e a última com dados de monitoração de harmônicos em um sistema de distribuição da Austrália. Os resultados dos testes confirmaram a efetividade do método proposto, encontrando o número ótimo de grupos a serem formados.

Foi introduzido em (HARIYANTO; ANGGORO; NOEGROHO, 2014), uma nova abordagem para identificar o perfil de severidade dos afundamentos de tensão nos barramentos, devido a uma falta ocorrida em determinado ponto do sistema de distribuição. Dados contendo 2 variáveis (amplitude e duração) dos afundamentos de tensão, ocasionados por faltas em todos os nós do sistema são gerados. Em seguida, estes dados referentes a todos os barramentos são separados em conjuntos usando a técnica de *k-means*. Os dados agrupados produziram os centróides e os membros do grupo. Para poder estimar o impacto dos eventos na carga, os parâmetros de amplitude e duração do centróide do grupo são convertidos em índice de severidade, que neste caso traduz os efeitos do AMT levando em consideração a curva de sensibilidade CBEMA-ITI. Os dados de certo barramento que sofre AMTs serão classificados com base em sua associação com os grupos existentes. Ao implementar o método, como resultado, é adquirida a informação dos barramentos que possuem menor ocorrência de eventos severos, permitindo então, a escolha de instalação de cargas sensíveis em tais pontos do sistema. Além disso, conhecendo os barramentos mais afetados, ou seja, com maior frequência de eventos severos, a alocação de dispositivos de atenuação como DVR (*Dynamic Voltage Restorer*) pode ser melhor avaliada.

É apresentado em (SEERA et al., 2016) um modelo híbrido para análise de QEE composto por uma modificação do método *fuzzy min-max neural network* (FMM) somado a uma modificação da técnica de *clustering tree* (CT). Inicialmente, usaram-se vários casos teste de referência para ilustrar os padrões dos grupos formados pela metodologia proposta. Em seguida, aplicou-se a um estudo de caso com dados reais relacionados ao monitoramento da qualidade da energia, para avaliar o desempenho do modelo híbrido FMM – CT. Os resultados foram comparados com aqueles obtidos quando aplicadas outras técnicas de formação de conjuntos, indicando uma melhor precisão do novo método proposto. Vale ressaltar, que foram estabelecidas regras explicativas do FMM – CT para justificar suas previsões. Os resultados empíricos indicam a aplicabilidade do modelo exibido em lidar com formação de conjuntos de dados de monitoramento da QEE.

Foi desenvolvido em (MAHELA; SHAIK, 2016) uma metodologia para detecção e classificação de distúrbios de QEE usando transformada Stock-weel's. Os distúrbios foram gerados por meio do Matlab conforme os padrões estabelecidos na IEEE – 1159. Diversas características do sinal foram extraídas da análise multiresolução baseada na transformada-S. Estas características são usadas para classificar os distúrbios por meio do método de formação de conjuntos *fuzzy c-means*. Nota-se que a classificação feita por tal método produz uma precisão satisfatória mesmo sob condições de ruído. Os distúrbios da QEE investigados incluem: afundamentos e elevações de tensão, interrupções, harmônicos, *notch*, *flicker*, transitórios oscilatórios e impulsivos e *spike*. A eficácia do algoritmo proposto foi verificada por resultados satisfatórios de vários estudos de caso, mostrando uma assertividade de 99%. Portanto, esta técnica proposta pode ser implementada em dispositivos de monitoramento da QEE.

Ao identificar uma amostra desconhecida, o algoritmo de *K-nearest neighbors* (KNN) requer o cálculo das distâncias entre essa amostra e todas as amostras no conjunto de treinamento. Assim, este algoritmo não atende aplicações em tempo real, tal como o reconhecimento de distúrbio de qualidade da energia. Para contornar este impasse, (PAN et al., 2017) propõe um novo método para reduzir o tamanho do conjunto de treinamento para o algoritmo KNN. O método proposto é baseado em um processo iterativo. Agrupando o conjunto de treinamento original, se o grupo contiver apenas uma classe, as amostras nesse grupo serão representadas pelo seu centróide, caso contrário, agrupa-se novamente. Ao final deste processo, o conjunto de treinamento original é representado pelo subconjunto que consiste em todos os centróides dos grupos. Resultados experimentais, obtidos com a aplicação de um sinal artificial e de um sinal real de medidor, mostraram que a precisão após a redução da amostra via processo recursivo não teve diferença comparada ao conjunto de treinamento original. No entanto, a classificação de um novo sinal se tornou mais rápida. Para um sinal de um dispositivo real de medição, o tempo de classificação reduziu de 1,35 s para 0,09 s.

O trabalho (DUAN; CHEN, 2014) propõe um método para avaliar de forma abrangente a QEE com base no algoritmo *MaximumTree* para formação de conjuntos pela técnica *fuzzy*. Alguns limites entre objetos são certos, enquanto alguns possuem incertezas. Quando o campo ao qual se está aplicando a formação de conjuntos envolve incertezas nos limites dos objetos sob estudo, a técnica *fuzzy* deve ser empregada. A formação de conjuntos usando uma metodologia *fuzzy* é um dos métodos de análise multivariada presentes na teoria

estatística, capaz de dividir amostras de forma objetiva em vários grupos, determinados pela afinidade entre objetos, que é estabelecida quantitativamente por métricas matemáticas. O primeiro passo da metodologia criada é selecionar índices únicos para avaliação de cada distúrbio de QEE de interesse. Para ilustração o trabalho selecionou 4 indicadores: desvios da tensão nominal, variações de frequência, desequilíbrios de tensão e harmônicos. Em sequência é formada a matriz de dados original, composta por n linhas que caracterizam os objetos (pontos de medição do sistema elétrico) e por m colunas referentes aos indicadores escolhidos no passo anterior. Devido à ordem de grandeza dos indicadores na matriz de dados original serem diferentes, se não padronizados, os de menor ordem poderão ter sua influência subestimada, enquanto os de ordem elevada terão sua influência superestimada. Para evitar isto, é necessário padronizar os índices, sendo que existem diversos métodos de padronização de dados. Neste artigo por melhor adequação, foi utilizada a distância *Hamming* para se obter coeficientes de similaridade entre as células da matriz original. Os resultados atingidos com aplicação em um caso prático provaram a viabilidade do método, que fornece algumas orientações com base científica para o consumidor selecionar a distribuidora de energia e ajustar o preço pago pela energia de acordo com a qualidade ofertada.

O artigo (FILHO et al., 2020) propõe uma metodologia para localização da fonte de afundamentos de tensão. Inicialmente a análise de agrupamentos é utilizada para dividir em grupos os dados de sinais de tensão medidos em diferentes nós do sistema. Então, o conjunto de regras de decisão é definido por meio do algoritmo da árvore de decisão parcial, que irá confrontar as características de cada grupo para definir em qual deles a localização da fonte do distúrbio se adequa. O sistema teste IEEE-34 barras foi utilizado para avaliar a metodologia e os resultados mostraram uma taxa de acerto superior a 98%.

O trabalho (GARCÍA-SÁNCHEZ et al., 2016) propõe e avalia uma metodologia alternativa para caracterizar e classificar afundamentos de tensão. PCA e a técnica de formação de conjuntos *K-means* são aplicadas para identificar os padrões na tensão *RMS* e reduzir o número de perfis de tensão representativos dos eventos considerados. Dados reais de 300 eventos coletados em uma usina eólica na Espanha foram usados para validar a metodologia, que mostrou-se eficiente para avaliar um grande número de eventos.

O artigo (COSTA et al., 2019) propõe uma metodologia para determinar um indicador baseado no Fator de Impacto que considere a sensibilidade dos equipamentos com relação aos AMTs, as variações de seus níveis de sensibilidade e a influencia de afundamentos desequilibrados em cargas trifásicas. Por meio de um procedimento estatístico que considera a

correlação entre o indicador e o número de paradas de funcionamento dos dispositivos, são determinadas diferentes regiões de sensibilidade e fatores de ponderação daqueles estabelecidos no PRODIST.

A pesquisa realizada em (SOUZA, 2016) está relacionada ao tema desta tese, onde foi introduzida uma metodologia para formação de conjuntos de sistemas de distribuição considerando as variáveis relacionadas aos AMTs. A metodologia se resume em quatro processos: seleção das variáveis pela correlação destas com a frequência de afundamentos de tensão, implementação da formação de conjuntos considerando diversos métodos para posterior investigação do mais adequado, avaliação dos métodos que geraram melhores agrupamentos por meio da análise de variância entre a resposta e o *membership* gerado e finalmente análise de robustez feita pela inserção de pequenos ruídos na variável de entrada, observando qual dentre os métodos é mais assertivo nesta condição. O procedimento criado tem como objetivo apontar qual o melhor método de formação de conjuntos a ser aplicado a cada base de dados. Para teste do método, foram considerados dados de um sistema hipotético e outro real com 15 subestações pertencentes à EDP Escelsa. Os resultados obtidos no caso real mostraram que o método de *Ward* foi o mais adequado à base de dados e que existem dois conjuntos de sistemas de distribuição com níveis de AMTs diferentes.

Outro artigo que possui relação com assunto desta tese foi publicado em (FILHO et al., 2016), onde os autores, diante de uma base de dados contendo medições de AMTs de 17 subestações da EDP Escelsa, propõem a aplicação da análise de componentes principais (PCA) para reduzir os dados de entrada de 32 variáveis (com certo nível de redundância) em 7 componentes principais (PCs) que acumulam 97,9% da informação das variáveis originais. A partir destas PCs formam agrupamentos de SEs, utilizando o método de formação de conjuntos denominado *Ward*, considerando a distância Euclidiana entre os elementos. Inicialmente, foram obtidos 5 grupos de SEs motivados pela regra de *Sturges*, porém pela análise de variância feita considerando a frequência de AMTs e o número classificador do conjunto, ficou evidente a existência de 2 grupos bem distintos de SEs, um com elevada e outro com baixa frequência do distúrbio. Adicionalmente, foi feita uma análise das principais variáveis relacionadas com o fenômeno, explicitando para cada uma delas, quais as possíveis ações das concessionárias de energia a fim de melhorar a QEE.

Estudos conduzidos em (ALMEIDA et al., 2020) apresentaram uma nova metodologia para aumentar o poder discriminatório na estimação de padrões de afundamentos de tensão usando funções elipsoidais. O método de *Ward* foi usado para formar grupos de subestações

com similaridade com relação aos afundamentos de tensão, 3 distintos grupos foram encontrados, com pequena, média e grande frequência de AMTs. O trabalho é uma evolução do apresentado em (FILHO et al., 2016). O método mostrou resultados que são mais precisos, estáveis e confiáveis.

De forma geral, pelos resumos das bibliografias correlatas encontradas, nota-se que em 11 trabalhos as técnicas de formação de conjuntos foram utilizadas com um propósito diferente do objetivo desta tese, como monitoração, identificação e classificação de eventos de QEE, localização da fonte e reconhecimento de padrões de afundamentos de tensão. Estas referências foram apresentadas a fim de identificar a aplicação da análise de agrupamentos na área de QEE.

O trabalho (COSTA et al., 2019) é focado em propor diferentes regiões de sensibilidade e fatores de ponderação daqueles estabelecidos no PRODIST. Enquanto esta tese, assumindo que as regiões de sensibilidade e os fatores de ponderação do PRODIST são adequados, usa análise de agrupamentos, para propor novos valores para a frequência máxima de ocorrência de afundamentos de tensão e conseqüentemente um fator de impacto base distinto para cada sistema de distribuição. Portanto, estes trabalhos são distintos, embora complementares.

Os trabalhos (ALMEIDA et al., 2020; FILHO et al., 2016; SOUZA, 2016) testam alguns métodos de formação de conjuntos com o objetivo de avaliar qual deles melhor se adequa para formar grupos de sistemas de distribuição com relação a frequência de afundamentos de tensão. Estes trabalhos são os que possuem maior relação com o tema desta tese, porém eles estão focados apenas na formação dos grupos, enquanto esta tese além de estabelecer os grupos, utiliza-se desta informação para estabelecer fatores de impacto base distintos para cada sistema de distribuição e que consideram o desempenho dos sistemas similares. Portanto esta tese complementa os estudos conduzidos em (ALMEIDA et al., 2020; FILHO et al., 2016; SOUZA, 2016) com o objetivo de propor aprimoramentos para a regulamentação brasileira.

1.4 Objetivos e Contribuições da Tese

O objetivo deste trabalho é propor uma metodologia capaz de selecionar, dentre uma base de dados, as variáveis relativas aos SDs que influenciam a ocorrência de AMTs. Utilizando as variáveis selecionadas formar conjuntos de sistemas de distribuição com características semelhantes e considerando o desempenho dos sistemas presentes no conjunto,

estabelecer FI_{base} adequados e distintos para cada sistema de distribuição.

A contribuição principal da tese é propor e aplicar uma metodologia factível para estabelecimento de FI_{base} distintos de acordo com as características físico elétricas dos sistemas de distribuição. Devido ao caráter generalista com que a proposta de metodologia é feita, tal procedimento poderá ter aplicação em qualquer base de dados que compreenda outros SDs do Brasil, ou até mesmo em outros países que venham adotar indicador similar ao Fator de Impacto.

A aplicação da metodologia gera benefícios para todos os agentes do setor elétrico, os sistemas de distribuição ficarão sujeitos a fatores de impacto base mais coerentes, que consideram os níveis das variáveis que influenciam na ocorrência de AMTs. Já os consumidores industriais que possuem cargas sensíveis perceberão um incentivo a melhora da qualidade da energia elétrica, com possível redução das paradas de processo, pois em geral os fatores de impacto base calculados pela metodologia ora proposta são inferiores ao valor atualmente definido pelo PRODIST.

1.5 Estrutura da Tese

Este trabalho está dividido em cinco capítulos, e a descrição de cada um deles é apresentada nos próximos parágrafos.

No presente capítulo 1 encontra-se a relevância do tema, o estado da arte mostrando como as técnicas de formação de conjuntos foram utilizadas na área de QEE, o objetivo, as contribuições esperadas pela proposta da tese e a estrutura do trabalho.

No capítulo 2, é feita uma breve revisão teórica sobre os principais conceitos utilizados na metodologia proposta, apresentando as técnicas de análise de regressão múltipla, o método dinâmico para formação de conjuntos, a definição de VTCDs e os indicadores.

No capítulo 3, apresenta-se a metodologia proposta para estabelecimento do fator de impacto base distinto para cada sistema de distribuição de acordo com as suas características físico elétricas.

No capítulo 4, aplica-se e valida-se a metodologia proposta utilizando como caso teste dados de 19 sistemas de distribuição reais.

O capítulo 5 é destinado às conclusões da pesquisa e propostas para trabalhos futuros.

Posteriormente são apresentadas as referências utilizadas como base para realização

desta tese.

Finalmente são apresentados o Apêndice A contendo as tabelas de distância e heterogeneidade geradas pela aplicação do método dinâmico e o Apêndice B contendo a tabela de frequência média mensal de AMTs nas regiões de sensibilidade de A à G.

2. REVISÃO DE CONCEITOS

2.1 Considerações Iniciais

Para facilitar o entendimento da tese, este capítulo apresenta uma breve revisão teórica dos principais conceitos que serão utilizados na metodologia proposta no capítulo subsequente.

2.2 Análise de Regressão Múltipla

A análise de regressão tem vasta aplicação nas mais diversas áreas da ciência, consiste em determinar uma função matemática que busca descrever o comportamento de uma determinada variável, denominada dependente ou critério, com base nos valores de uma ou mais variáveis, denominadas independentes, regressoras ou preditoras (CORRAR; PAULO; FILHO, 2009).

Um modelo de regressão que contenha mais de um preditor é chamado de um modelo de regressão múltipla (MONTGOMERY; RUNGER, 2009). Seu objetivo é usar as variáveis independentes cujos valores são conhecidos para prever os valores da variável dependente selecionada pelo pesquisador.

Cada variável independente é ponderada pelo procedimento de análise de regressão para garantir máxima previsão a partir do conjunto de variáveis independentes. Os pesos denotam a contribuição relativa das variáveis independentes para a previsão geral e facilitam a interpretação sobre a influência de cada variável em fazer a previsão apesar da correlação entre as variáveis independentes complicar o processo interpretativo (HAIR et al., 2009).

A análise de regressão múltipla é uma técnica de dependência. Assim, como premissa para empregá-la, o pesquisador deve ser capaz de classificar as variáveis de que ele dispõe em dependentes e independentes (CORRAR; PAULO; FILHO, 2009).

Em geral a variável dependente ou de resposta, y , pode estar relacionada a k variáveis independentes ou preditoras. O modelo genérico apresentado na equação (2.1) é chamado de modelo de regressão linear múltipla com k variáveis preditoras:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k + \epsilon \quad (2.1)$$

A equação (2.1) descreve um hiperplano no espaço k -dimensional das variáveis preditoras. Os parâmetros β_j são chamados de coeficientes parciais de regressão

(MONTGOMERY; RUNGER, 2009). Pode-se interpretar β_j como a mudança esperada em y devido ao aumento de uma unidade em x_j , estando as outras variáveis x_k , $k \neq j$, fixas (CHARNET et al., 2008).

O método dos mínimos quadrados pode ser usado para estimar os coeficientes β_j no modelo de regressão múltipla. No ajuste do modelo é conveniente expressar as operações matemáticas usando notação matricial. Supõe-se que haja k variáveis preditoras e n observações, $(x_{i1}, x_{i2}, \dots, x_{ik}, y_i)$, $i = 1, 2, \dots, n$, e que o modelo relacionando os preditores a resposta seja como o da equação (2.2):

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_k x_{ik} + \epsilon_i \quad (2.2)$$

Esse modelo é um sistema de n equações, que pode ser expresso na notação matricial pela equação (2.3):

$$y = X\beta + \epsilon \quad (2.3)$$

Sendo:

$$y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \quad X = \begin{bmatrix} 1 & x_{11} & x_{12} & \dots & x_{1k} \\ 1 & x_{21} & x_{22} & \dots & x_{2k} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_{n1} & x_{n2} & \dots & x_{nk} \end{bmatrix} \quad \beta = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{bmatrix} \quad \epsilon = \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_n \end{bmatrix}$$

Em geral, y é um vetor ($n \times 1$) das observações, X é uma matriz ($n \times p$) dos níveis das variáveis independentes, β é um vetor ($p \times 1$) dos coeficientes de regressão e ϵ é um vetor ($n \times 1$) dos erros aleatórios (MONTGOMERY; RUNGER, 2009).

A equação (2.4) fornece a estimativa de mínimos quadrados para β (CHARNET et al., 2008):

$$\hat{\beta} = (X'X)^{-1}X'y \quad (2.4)$$

Observe que há $p = k+1$ equações normais para $p = k+1$ incógnitas (os valores de $\hat{\beta}_0, \hat{\beta}_1, \dots, \hat{\beta}_k$). Além disso, a matriz $X'X$ é sempre não singular, de modo que, para inverter essas matrizes, os métodos descritos na literatura específica sobre determinantes e matrizes podem ser usados. Na prática, os cálculos de regressão múltipla são quase sempre realizados em um computador (MONTGOMERY; RUNGER, 2009).

Com o intuito de analisar se há relação linear entre a variável de resposta y e um subconjunto de preditores x_1, x_2, \dots, x_k , testam-se as seguintes hipóteses:

$$H_0: \beta_1 = \beta_2 = \dots = \beta_k = 0$$

$H_1: \beta_j \neq 0$ para no mínimo um j .

Se H_0 é rejeitado, conclui-se que há contribuição significativa de uma ou mais variáveis preditoras no modelo estimado para y .

O teste de significância da regressão é uma generalização do procedimento usado na regressão linear simples.

A soma de quadrados dos erros (SQ_E) representa uma medida do erro de previsão cometido, oposto a tal medida, a soma de quadrados da regressão (SQ_{Reg}) expressa o sucesso de previsão do modelo. Juntas, essas duas medidas, devem se igualar à soma total de quadrados (SQ_T) conforme mostram as equações (2.5) e (2.6) (HAIR et al., 2009):

$$SQ_T = SQ_{Reg} + SQ_E \quad (2.5)$$

$$\sum_{i=1}^n (y_i - \bar{y})^2 = \sum_{i=1}^n (\hat{y}_i - \bar{y})^2 + \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (2.6)$$

A estatística de teste tem, sob H_0 , distribuição F, com k e $(n - p)$ graus de liberdade. Os k graus de liberdade do numerador se devem ao fato de se ter k parâmetros (preditores) sendo testados. No denominador tem-se n subtraído da dimensão do vetor β . As quantidades necessárias para calcular o valor observado dessa estatística são dispostas na Tabela 2.1 denominada tabela de ANOVA (CHARNET et al., 2008).

Tabela 2.1 – Tabela de ANOVA

| Fonte de Variação | Graus de Liberdade | Soma de Quadrados | Quadrado Médio | F ₀ |
|-------------------|--------------------|-------------------|----------------------|---------------------------------|
| Regressão | k | SQ_{Reg} | $\frac{SQ_{Reg}}{k}$ | $\frac{SQ_{Reg}/k}{SQ_E/(n-p)}$ |
| Erro | $n - p$ | SQ_E | $\frac{SQ_E}{(n-p)}$ | |
| Total | $n - 1$ | SQ_T | | |

Fonte: adaptado de (MONTGOMERY; RUNGER, 2009).

A hipótese $H_0: \beta_1 = \beta_2 = \dots = \beta_k = 0$ é rejeitada quando F_0 for maior do que $F_{k,(n-p)}(\alpha)$, o quantil $(1 - \alpha)$ de distribuição F com k e $(n - p)$ graus de liberdade.

Outro modo de expressar o nível de precisão da previsão é com o coeficiente de determinação (R^2), a razão entre a soma de quadrados da regressão e a soma total de quadrados, como mostrado na equação (2.7):

$$R^2 = \frac{SQ_{Reg}}{SQ_T} = 1 - \frac{SQ_E}{SQ_T} \quad (2.7)$$

Assim R^2 é uma estatística global para avaliar o quanto da variabilidade da resposta y é explicada pelas variáveis independentes (CORRAR; PAULO; FILHO, 2009).

A estatística R^2 pode não representar a melhor medida para avaliação do modelo de regressão múltipla, uma vez que ela sempre aumenta quando uma variável é adicionada ao modelo. Portanto, muitos usuários de regressão preferem usar a estatística R^2 ajustada, dada pela equação (2.8) (MONTGOMERY; RUNGER, 2009):

$$R^2_{ajustado} = 1 - \frac{SQ_E/(n-p)}{SQ_T/(n-1)} \quad (2.8)$$

Se a pretensão é a comparação entre várias equações de regressão envolvendo números de variáveis independentes diferentes, ou mesmo tamanhos de amostra diferentes, o valor de R^2 ajustado é extremamente útil, pois ele considera essas especificidades do modelo.

Na maioria das pesquisas, existe um grande número de variáveis independentes disponíveis que podem ser escolhidas para inclusão na equação de regressão. A etapa da seleção de quais variáveis farão parte do modelo se constitui num ponto importante do processo de estimação do modelo (CORRAR; PAULO; FILHO, 2009).

Algumas vezes, experiências prévias ou considerações teóricas em foco podem ajudar o analista a especificar o conjunto de variáveis preditoras a se usar em uma situação particular (MONTGOMERY; RUNGER, 2009).

Um dos métodos de busca sequencial mais popular para seleção de variáveis e que foi utilizado na metodologia proposta nesta tese é o método chamado de estimação *stepwise* ou por etapas (HAIR et al., 2009). Este método ajusta um modelo considerando as variáveis regressoras primeiramente com um conjunto de variáveis independentes e, a partir dele, acrescentam ou eliminam variáveis até alcançarem a melhor medida dentro do critério utilizado (CORRAR; PAULO; FILHO, 2009). Essa técnica fornece um método objetivo para selecionar variáveis que maximiza a previsão ao mesmo tempo em que emprega o menor número de variáveis.

Cuidados adicionais referentes à colinearidade¹ devem ser observados. O critério adotado por esse tipo de método para inclusão ou eliminação de variáveis é a maximização do

¹ Existência de forte correlação entre variáveis independentes.

poder preditivo incremental de cada variável independente adicional, portanto, se duas variáveis regressoras forem altamente correlacionadas entre si e ainda forem correlacionadas com a variável dependente num mesmo grau, é provável que apenas uma delas entre no modelo, pois, quando a primeira delas for incluída, sobrar pouco poder preditivo adicional para a outra (CORRAR; PAULO; FILHO, 2009).

O procedimento constrói iterativamente uma sequência de modelos de regressão pela adição ou remoção de variáveis em cada etapa. O critério para adicionar ou remover uma variável em qualquer etapa é geralmente expresso em termos de um teste parcial F . Sendo f_{entra} o valor da variável aleatória F para adicionar uma variável ao modelo e f_{sai} o valor da variável aleatória F para remover uma variável do modelo. (MONTGOMERY; RUNGER, 2009). As etapas da estimação *stepwise* estão ilustradas no Fluxograma da Figura 2.1.

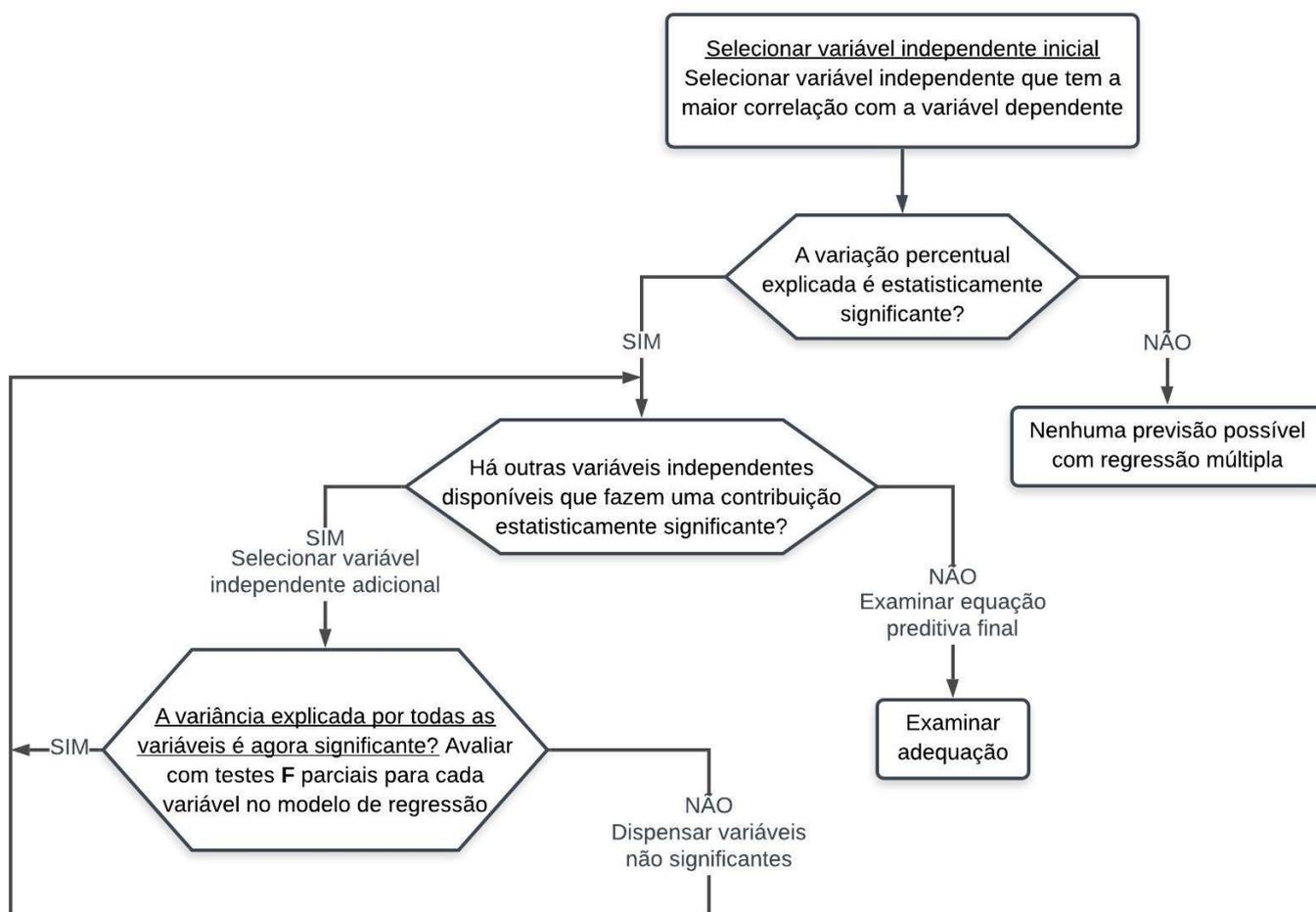


Figura 2.1 – Fluxograma da estimação *stepwise*.

Fonte: adaptado de (HAIR et al., 2009).

O processo se inicia com um modelo de regressão simples onde a variável regressora com o maior coeficiente de correlação com a variável dependente é escolhida. As próximas variáveis regressoras a serem incluídas são selecionadas com base na sua contribuição incremental (correlação parcial) à equação de regressão.

A cada nova variável independente introduzida no modelo é examinado pelo teste F se a contribuição das variáveis que já se encontram no modelo continua significativa, dada a presença da nova variável. Caso não seja, a estimação *stepwise* permite que as variáveis que já estão no modelo sejam eliminadas.

O procedimento continua até que todas as variáveis regressoras ainda não presentes no modelo tenham sua inclusão avaliada e a reação das variáveis já presentes no modelo seja observada quando dessas inclusões (CORRAR; PAULO; FILHO, 2009; HAIR et al., 2009).

Independentemente do método de seleção de variáveis utilizado, o critério mais importante é o bom conhecimento do pesquisador sobre o tema da pesquisa e alguma fundamentação teórica que permita uma perspectiva objetiva e bem informada quanto às variáveis a serem incluídas e aos sinais e magnitude esperados de seus coeficientes. Sem esses conhecimentos, os resultados da regressão podem ter elevada precisão preditiva sem qualquer relevância teórica.

2.3 Análise de Agrupamentos (Método Dinâmico)

Análise de agrupamentos é o conjunto de técnicas multivariadas cuja finalidade principal é agregar objetos, itens ou indivíduos com base nas características que eles possuem. (HAIR et al., 2009). O critério utilizado para agrupar objetos é baseado nas suas semelhanças (FERREIRA, 2011). Assim, objetos pertencentes a um mesmo grupo são similares entre si com relação às variáveis que neles foram medidas, e os elementos de grupos distintos são dissimilares com respeito a estas mesmas variáveis (MINGOTI, 2005).

Uma questão importante refere-se ao critério a ser utilizado para decidir se dois elementos da base de dados podem ser considerados como semelhante ou não, tal questão é respondida com o uso de métricas matemáticas (MINGOTI, 2005). A semelhança entre objetos é quantificada por meio de uma medida de proximidade, que engloba tanto as medidas de similaridade como as de dissimilaridade (distância) (FERREIRA, 2011).

Considere uma base de dados constituída por n elementos amostrais onde foram medidas p variáveis aleatórias para cada um deles. Tem-se, para cada elemento amostral j o

vetor de medidas X_j , conforme equação (2.9):

$$X_j = [X_{1j} X_{2j} \dots X_{pj}]', \quad j = 1, 2, \dots, n \quad (2.9)$$

Em que X_{ij} representa o valor observado da variável i medida no elemento j . Para dar início à análise de agrupamentos, a medida de similaridade ou dissimilaridade a ser utilizada deve ser determinada. Existem vários tipos de medidas e o resultado final dos grupos gerados normalmente se altera de acordo com a escolha do tipo de medida. Neste trabalho foi utilizada a distância euclidiana.

Considerando dois elementos X_l e X_k , $l \neq k$, a distância euclidiana entre eles é definida por (2.10):

$$d(X_l, X_k) = [(X_l - X_k)'(X_l - X_k)]^{\frac{1}{2}} = \left[\sum_{i=1}^p (X_{il} - X_{ik})^2 \right]^{\frac{1}{2}} \quad (2.10)$$

Os dois elementos amostrais são comparados em cada variável pertencente ao vetor de observações.

As técnicas de agrupamentos ou *clusters* são classificadas em dois tipos: as não hierárquicas e as hierárquicas e estas são novamente classificadas em aglomerativas e divisivas (MINGOTI, 2005).

Apesar dos métodos hierárquicos e não hierárquicos possuírem certas vantagens, a aplicação destes pode não gerar bons resultados quando analisados os elementos situados nas fronteiras entre os diversos grupos.

Na Figura 2.2, nota-se que os elementos I e L pertencem ao grupo C1 e F pertence ao grupo C2. Portanto, tais elementos estarão representados pelas características dos seus respectivos centróides. No entanto, é evidente que os elementos I, L e F possuem muito mais semelhança entre si de que com seus centróides.

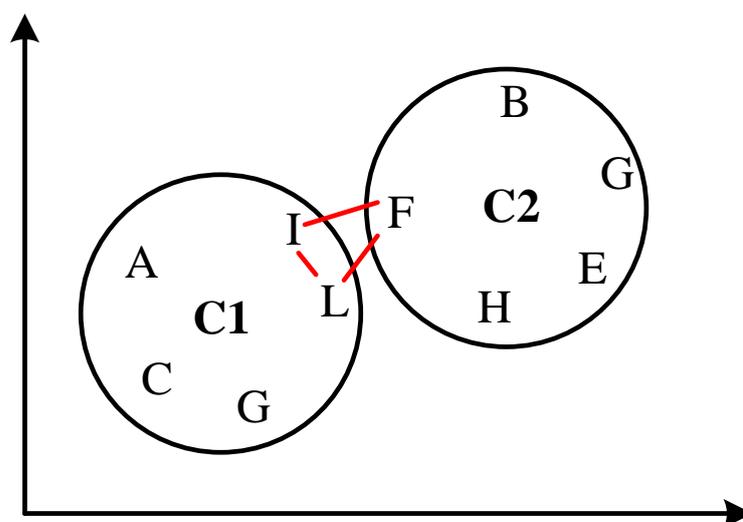


Figura 2.2 – Semelhança entre elementos na fronteira de grupos diferentes.

Fonte: adaptado de (TANURE, 2004).

Este problema tende a reduzir à medida que o número de grupos a ser formado aumenta. Assim, institui-se uma relação de compromisso entre o número de grupos e o grau de representatividade dos elementos com relação aos seus vizinhos e para com seus centróides (TANURE, 2004).

Adicionalmente, outro problema que as técnicas de análise de agrupamentos possuem é referente ao número de elementos da amostra. Quando se pretende classificar amostras compostas por pequeno número de elementos, há pouca representatividade sobre os agrupamentos resultantes.

Visando contornar os dois problemas citados, (TANURE, 2004) propôs um novo método que estabelece grupos dinâmicos (que se alteram) a partir de cada elemento. No método dinâmico, para cada elemento de referência, será formado um agrupamento de elementos que são os mais comparáveis ao denominado elemento de referência. Neste método não há a formação de agrupamentos fixos, é como se existissem agrupamentos distintos para cada elemento (NOTA TÉCNICA N° 0021, 2011). Este método é muito indicado quando for de extrema relevância o senso de pertencimento a cada agrupamento. O algoritmo para esta técnica consiste em:

1. Adota-se cada elemento como centróide de um grupo a ser criado;
2. Definido o centróide, se determina o grau de semelhança (distância ou heterogeneidade percentual) de todos os elementos com relação a este “centróide”;

3. Estabelece-se um critério de corte para o grau de semelhança entre o centróide e cada elemento;
4. Agrupam-se a cada centróide os elementos mais representativos com base em suas semelhanças;
5. Retiram-se as características de interesse do agrupamento formado
6. Repete-se o processo para cada um dos elementos.

A heterogeneidade percentual é obtida dividindo os valores de heterogeneidade pela heterogeneidade máxima. Para padronizar a expressão, a heterogeneidade máxima será a distância entre o elemento de referência e um elemento hipotético cujos atributos padronizados sejam três vezes o valor dos atributos do elemento de referência. Dessa forma, a fórmula da heterogeneidade percentual é apresentada na equação (2.11):

$$Heterogeneidade_i^p = \frac{Max(Dist_i^j)}{\sqrt{k \cdot 3^2}} \quad (2.11)$$

Em que: i – índice do elemento de referência;

j – elementos próximos ao elemento i ;

$Dist_i^j$ – distância euclidiana do elemento i para o elemento j ;

k – número de atributos.

O inconveniente do método dinâmico é que cada elemento da amostra irá gerar um grupo. Então, para aplicações que possuem muitos elementos, é imprescindível que o algoritmo seja implementado computacionalmente.

Ressalta-se que para aspectos regulatórios do setor elétrico, como no processo de estabelecimento de valores base para indicadores, é fundamental que os agentes sintam-se representados pelos agrupamentos propostos, caso contrário, haverá questionamentos por parte das distribuidoras de energia. Por essa razão o método dinâmico configura-se como uma técnica recomendada para este tipo de tratamento.

2.4 VTCDs e Indicadores

As VTCDs são definidas como eventos aleatórios caracterizados por desvios significativos na amplitude do valor eficaz da tensão durante um curto intervalo de tempo e são divididas em: afundamentos, elevações e interrupções de tensão.

A Tabela 2.2 apresenta a classificação das variações de tensão de curta duração de acordo com o Módulo 8 do PRODIST (ANEEL, 2021).

Tabela 2.2 – Classificação das Variações de Tensão de Curta Duração (PRODIST).

| Classificação | Denominação | Duração da Variação | Amplitude da tensão (valor eficaz) em relação à tensão de referência |
|-------------------------------|----------------------------------|--|---|
| Variação Momentânea de Tensão | Interrupção Momentânea de Tensão | Inferior ou igual a três segundos | Inferior a 0,1 p.u. |
| | Afundamento Momentâneo de Tensão | Superior ou igual a um ciclo e inferior ou igual a três segundos | Superior a 0,1 p.u. e inferior a 0,9 p.u. |
| | Elevação Momentânea de Tensão | Superior ou igual a um ciclo e inferior ou igual a três segundos | Superior a 1,1 p.u. |
| Variação Temporária de Tensão | Interrupção Temporária de Tensão | Superior a três segundos e inferior a três minutos | Inferior a 0,1 p.u. |
| | Afundamento Temporária de Tensão | Superior a três segundos e inferior a três minutos | Superior a 0,1 p.u. e inferior a 0,9 p.u. |
| | Elevação Temporária de Tensão | Superior a três segundos e inferior a três minutos | Superior a 1,1 p.u. |

Fonte: (ANEEL, 2021).

Como se observa, os limiares do valor eficaz da tensão de 0,1 p.u., 0,9 p.u. e 1,1 p.u. são utilizados para caracterizar a VTCD nas classes de interrupção, afundamento e elevação, respectivamente. Quanto à duração do evento há uma subdivisão em momentâneos (1 ciclo à 3 segundos) e temporários (3 segundos e inferior a 3 minutos).

Embora haja muitos estudos e recomendações focados em afundamentos de tensão, não se tem um consenso internacional sobre qual indicador melhor caracteriza o distúrbio. O módulo 8 do PRODIST apresenta como parâmetros de uma VTCD, a amplitude (equação 2.12) e a duração do evento (equação 2.13). A frequência de ocorrência de eventos de um barramento ou sistema numa base de tempo apropriada (mensal e anual) é calculada por intermédio da equação (2.14).

$$V_e = \frac{V_{res}}{V_{ref}} \times 100 \quad (2.12)$$

Em que:

V_e – amplitude do evento de VTCD (em %);

V_{res} – tensão residual do evento de VTCD (em Volt);

V_{ref} – tensão de referência (em Volt).

$$\Delta t_e = t_f - t_i \quad (2.13)$$

Em que:

Δt_e – duração do evento de VTCD (em milissegundos);

t_f – instante final do evento de VTCD;

t_i – instante inicial do evento de VTCD.

$$f_e = n \quad (2.14)$$

Em que:

f_e – frequência de ocorrência de eventos de VTCD;

n – quantidade de eventos de VTCD registrados no período de avaliação.

Alguns documentos internacionais como a (IEEE P1564, 2014), (IEC 61000-2-8, 2005), (IEC 61000-4-11, 2020), (NRS 048-2, 2003) recomenda que a estratificação dos eventos seja feita em tabelas com determinadas faixas de amplitude e duração. Para o módulo 8 do PRODIST, o registro dos eventos de VTCDs deve ser estratificado conforme as classes mostradas na Tabela 2.3.

Tabela 2.3 – Estratificação dos parâmetros amplitude e duração para contabilização de eventos de VTCD.

| Amplitude (p.u.) | Duração | | | | | | |
|---------------------|-------------------|-----------------|-----------------|--------------|-----------|-------------|---------------|
| | [16,67ms – 100ms] | (100ms – 300ms] | (300ms – 600ms] | (600ms – 1s] | (1s – 3s] | (3s – 1min] | (1min - 3min] |
| >1,15 | | | | | | | |
| (1,10 – 1,15] | | | | | | | |
| (0,85 – 0,90] | | | | | | | |
| (0,80 – 0,85] | | | | | | | |
| (0,70 – 0,80] | | | | | | | |
| (0,60 – 0,70] | | | | | | | |
| (0,50 – 0,60] | | | | | | | |
| (0,40 – 0,50] | | | | | | | |
| (0,30 – 0,40] | | | | | | | |
| (0,20 – 0,30] | | | | | | | |
| (0,10 – 0,20] | | | | | | | |
| <0,10 | | | | | | | |

Fonte: (ANEEL, 2021).

Com base na estratificação da Tabela 2.3 e levando em consideração as particularidades do sistema elétrico brasileiro, o PRODIST estabelece conforme mostra a Tabela 2.4, nove regiões de sensibilidade, com o intuito de correlacionar a importância de cada evento com os níveis de sensibilidade de diferentes cargas (SOUZA, 2016).

Tabela 2.4 – Estratificação das VTCDs com base nos níveis de sensibilidade das diversas cargas.

| Amplitude (p.u) | Duração | | | | | | |
|--------------------|----------------------|--------------------|--------------------|-----------------|----------------|----------------|------------------|
| | [16,67ms – 100ms] | (100ms – 300ms] | (300ms – 600ms] | (600ms – 1s] | (1seg – 3s] | (3s – 1min] | (1min – 3min) |
| >1,15 | REGIÃO H | | | REGIÃO I | | | |
| (1,10 – 1,15] | REGIÃO H | | | REGIÃO I | | | |
| (0,85 – 0,90] | REGIÃO A | | | REGIÃO G | | | |
| (0,80 – 0,85] | REGIÃO A | | | REGIÃO G | | | |
| (0,70 – 0,80] | REGIÃO B | REGIÃO D | | REGIÃO F | | | |
| (0,60 – 0,70] | REGIÃO B | REGIÃO D | | REGIÃO F | | | |
| (0,50 – 0,60] | REGIÃO C | REGIÃO D | | REGIÃO F | | | |
| (0,40 – 0,50] | REGIÃO C | REGIÃO D | | REGIÃO F | | | |
| (0,30 – 0,40] | REGIÃO E | | | REGIÃO F | | | |
| (0,20 – 0,30] | REGIÃO E | | | REGIÃO F | | | |
| (0,10 – 0,20] | REGIÃO E | | | REGIÃO F | | | |
| < 0,10 | REGIÃO E | | | REGIÃO F | | | |

Fonte: (ANEEL, 2021).

Para caracterização da severidade da incidência de eventos em um único indicador, foi proposto o indicador Fator de Impacto (FI), que possui período de apuração de 30 dias consecutivos, sendo calculado pela equação (2.15) (ANEEL, 2021):

$$FI = \frac{\sum_{i=A}^I (f_{ei} \times fp_i)}{FI_{BASE}} \quad (2.15)$$

Em que:

f_{ei} é a frequência de ocorrência de eventos de VTCD, apuradas por meio de medição apropriada, em um período de 30 dias consecutivos, para cada região de sensibilidade i , com $i = A$ até I ;

fp_i é o fator de ponderação para cada região de sensibilidade;

FI_{BASE} é o fator de impacto base, calculado considerando os fatores de ponderação e a frequência limite de ocorrência de VTCD para cada região de sensibilidade.

A frequência limite de ocorrência de VTCD para cada região de sensibilidade está apresentada na Tabela 2.5.

Tabela 2.5 – Limites mensais de ocorrência por região de sensibilidade.

| Regiões de sensibilidade | Frequência máxima de ocorrências | |
|--------------------------|--|---|
| | $1 \text{ kV} < V_{nom} < 69 \text{ kV}$ | $69 \text{ kV} \leq V_{nom} < 230 \text{ kV}$ |
| A | - | - |
| B | 5 | 4 |
| C | 4 | 3 |
| D | 3 | 2 |
| E | 2 | 1 |
| F | 1 | 1 |
| G | 4 | 1 |
| H | 1 | 1 |
| I | 1 | 1 |

Fonte: (ANEEL, 2021).

O fator de ponderação (fp) para cada região de sensibilidade e também o fator de impacto base são indicados na Tabela 2.6.

Tabela 2.6 – Fatores de ponderação e Fator de Impacto Base.

| Região de Sensibilidade | Fator de Ponderação (fp) | Fator de Impacto Base (FI_{BASE}) | |
|-------------------------|------------------------------|--|---|
| | | $1 \text{ kV} < V_{nom} < 69 \text{ kV}$ | $69 \text{ kV} \leq V_{nom} < 230 \text{ kV}$ |
| A | 0,00 | | |
| B | 0,04 | | |
| C | 0,07 | | |
| D | 0,15 | | |
| E | 0,25 | 2,13 | 1,42 |
| F | 0,36 | | |
| G | 0,07 | | |
| H | 0,02 | | |
| I | 0,04 | | |

Fonte: (ANEEL, 2021).

O valor de referência estabelecido no PRODIST do indicador Fator de Impacto para os sistemas de distribuição, apurado por medição apropriada no período de 30 dias consecutivos, é de 1,0 p.u..

2.5 Considerações Finais

O presente capítulo apresentou inicialmente os conceitos básicos e os objetivos de se realizar uma análise de regressão múltipla.

Também foram exibidos os fundamentos da análise de agrupamentos, com destaque na descrição do método dinâmico que é atualmente utilizado pela ANEEL na definição dos limites para os indicadores de continuidade (FEC e DEC).

Para finalizar foi mostrada a definição de VTCDs e os indicadores atualmente estabelecidos pelo PRODIST para quantificar o fenômeno nos sistemas de distribuição. Nota-

se que o fator de impacto base e o valor de referência estabelecido para os sistemas de distribuição de média tensão são únicos, não levando em consideração as diferenças que os SDs têm com relação às variáveis que contribuem para a ocorrência de VTCDs.

3. METODOLOGIA PARA ESTABELECIMENTO DO FATOR DE IMPACTO BASE

3.1 Considerações Iniciais

Segundo (COSTA, 2018), dados de medição e simulação apontam que os afundamentos de tensão são os eventos mais frequentes dentre as VTCDs, tendo uma taxa de ocorrência muito superior à das elevações de tensão.

Ressalta-se também que (IEEE P1564, 2014) recomenda que o tratamento dos eventos de afundamentos e elevações sejam feitos de forma separada, devido aos efeitos diferentes que estes provocam em equipamentos. Quando expostos às elevações de tensão, os equipamentos são mais propensos a danos permanentes, resultando em queima de componentes ou até mesmo o dano total do equipamento. Quando submetidos a afundamentos de tensão, a probabilidade maior é de ocorrer falha ou desligamento de equipamentos de maneira intempestiva, resultando em parada parcial ou total de processos industriais.

Portanto, com a intenção de questionar a forma atual de cálculo do Fator de Impacto, que congrega em um único indicador afundamentos e elevações de tensão, nesta tese será priorizado o estudo dos afundamentos de tensão (regiões de sensibilidade de A à G). Contudo, matematicamente, a metodologia proposta pode ser empregada para também considerar as elevações de tensão.

Este capítulo é dedicado à apresentação das principais etapas da metodologia para o estabelecimento de novos fatores de impacto base, desde a escolha das variáveis que irão compor a base de dados até o estabelecimento do fator de impacto base específico para cada sistema de distribuição.

3.2 Escolha da Base de Dados

Com o intuito de tornar a metodologia genérica, permitindo sua aplicação em quaisquer sistemas de distribuição, a escolha da base de dados inicial, ou seja, as variáveis que se supõe ter relação com a ocorrência de afundamento de tensão, é feita por um pesquisador da área de qualidade da energia elétrica.

Dentre todas as variáveis possíveis de serem medidas, estimadas ou calculadas dos sistemas de distribuição, o pesquisador escolhe, baseado nos conhecimentos técnicos adquiridos, aquelas que possivelmente influenciam na ocorrência dos afundamentos de tensão. Este passo da metodologia tem certa subjetividade, pois é possível que duas pessoas

qualificadas escolham algumas variáveis diferentes, todavia havendo disponibilidade de dados, as variáveis escolhidas por ambos podem compor a base de dados inicial, desde que a quantidade de variáveis seja inferior ao número de sistemas de distribuição considerados na análise (condição para aplicação do método de seleção de variáveis). Assim, a etapa da escolha das variáveis da base de dados, pode ser feita não somente por um único pesquisador, como também por um grupo destes.

3.3 Seleção das Variáveis

Devido à sensibilidade que os métodos de formação de conjuntos possuem quando é considerado um número grande de variáveis de entrada, propõe-se uma etapa de seleção de variáveis a fim de determinar o menor número de variáveis possível que tenha um bom poder de explicar a variabilidade da resposta. Para esta etapa utiliza-se o procedimento de regressão linear múltipla, considerando como variáveis candidatas ao modelo as escolhidas no passo anterior e como variável resposta a frequência média mensal de afundamentos de tensão, que deve ser obtida através de medição com duração de, no mínimo um ano, para capturar a característica sazonal que o fenômeno possui.

O modelo de regressão é estimado usando o método de busca sequencial *stepwise*, citado no capítulo 2. O nível de significância (α) para entrada e permanência das variáveis no modelo deve ser estabelecido pelo pesquisador, quanto maior este parâmetro, mais variáveis serão incluídas ao modelo. Normalmente, são escolhidos nas aplicações de engenharia os níveis de significância de 0,1; 0,05 ou 0,01 (SHELDON, 2004).

Gerado o modelo, é importante que o pesquisador faça uma análise crítica para verificar: se as variáveis selecionadas são significativas (Valor – P inferior a α), se há baixa colinearidade entre as variáveis selecionadas (VIF^2 inferior a 5), se o modelo tem bom poder de explicação da variabilidade da resposta (R^2 satisfatório), se o sinal e a magnitude dos coeficientes encontrados estão em acordo com as conhecimentos técnicos, e finalmente, fazer uma análise dos resíduos para verificar a adequabilidade do modelo. Tendo o aval do pesquisador, as variáveis selecionadas no modelo seguem para o passo seguinte, de formação de conjuntos de sistemas de distribuição.

Vale ressaltar que o modelo gerado nesta etapa não é usado posteriormente para fazer previsão da variável resposta (frequência mensal de AMTs), serve exclusivamente para selecionar as variáveis a serem utilizadas na formação de conjuntos.

² Fator de inflação da variância.

3.4 Formação de Conjuntos

As variáveis selecionadas no passo anterior são utilizadas como dados de entrada para o método dinâmico de formação de conjuntos. Conforme mencionado no capítulo 2, pela aplicação deste método encontra-se para cada sistema de distribuição tomado como referência, uma lista ordenada dos sistemas de distribuição que possuem semelhança levando em conta as variáveis selecionadas na etapa anterior.

Para o estabelecimento do fator de impacto base, entende-se que o método dinâmico é mais adequado, uma vez que para cada SD tomado como referência encontram-se os SDs que mais se assemelham a este. Esta característica de pertencimento ao conjunto formado é fundamental para os agentes do setor elétrico. Ressalta-se também, que tal método é de conhecimento das distribuidoras e possui boa aceitação, por ser utilizado na metodologia atual para estabelecimento dos limites dos indicadores de continuidade.

Como todos os sistemas de distribuição possuem certa similaridade (mesmo que baixa) com o sistema tomado como referência, a lista ordenada obtida pela aplicação do método contém sempre todos os SDs considerados na pesquisa. Então, para formação dos conjuntos de sistemas de distribuição é necessário estabelecer a heterogeneidade percentual máxima permitida para considerar os SDs pertencentes àquele conjunto. Assim, todos os SDs que possuem heterogeneidade percentual com relação ao SD de referência inferior a máxima permitida, pertencem ao conjunto deste SD.

A heterogeneidade percentual é calculada aplicando-se a equação (2.11), sendo que a escolha do valor da heterogeneidade percentual máxima permitida deve ser feita pelo pesquisador, pois é dependente da quantidade de objetos da amostra. Em amostras que o número de objetos é próximo à população, provavelmente terão diversos objetos semelhantes ao objeto de referência, permitindo que um conjunto com número razoável de objetos seja formado com a escolha de um valor pequeno de heterogeneidade máxima permitida. Já em amostras pequenas, normalmente é necessário elevar a heterogeneidade máxima permitida para obter conjuntos com quantidade adequada de objetos.

3.5 Estabelecimento do Fator de Impacto base

Para o estabelecimento do novo fator de impacto base, para cada SD tomado como referência, é necessária uma tabela contendo a frequência média mensal de afundamentos de tensão em cada região de sensibilidade (A à G) dos SDs semelhantes, que deve ser obtida por meio dos dados de medição anual dos afundamentos de tensão.

Os fatores de impacto base distintos são obtidos atribuindo novos valores para a frequência máxima de ocorrência das regiões de sensibilidade (A à G). Estes valores são calculados levando em consideração o desempenho do conjunto de sistemas de distribuição, pois é de se esperar que SDs com características semelhantes com relação às variáveis que influenciam a ocorrência do fenômeno tenham desempenho similar.

O critério utilizado para estabelecer a frequência máxima de ocorrência deve ser escolhido pelo pesquisador ou em comum acordo entre os agentes do setor elétrico. Dentre as opções que se apresentam pode-se citar as medidas de posição (percentis) e a média.

O novo fator de impacto base do SD tomado como referência é calculado pela soma do produto dos fatores de ponderação indicados na Tabela 2.6 para a regiões de A à G, pela nova frequência máxima de ocorrência calculada por meio do critério escolhido, ou seja, algum percentil ou a média.

A diferenciação dos FI_{base} de cada sistema permite que o valor de referência estipulado pelo módulo 8 do PRODIST de 1 p.u. seja mantido, porém cada SD terá uma meta diferente de acordo com as características que mais contribuem para ocorrência do fenômeno e de acordo com o desempenho dos sistemas que possuem similaridade com relação a estas características.

3.6 Fluxograma

A metodologia proposta pode ser resumida nos seguintes passos:

- O pesquisador deve escolher as variáveis que irão compor a base de dados;
- Realizar a seleção das variáveis, por meio de regressão linear múltipla (*stepwise*);
- Realizar a formação de conjuntos de SDs pelo método dinâmico, utilizando as variáveis selecionadas no passo anterior;
- Estabelecer o novo FI_{base} para cada SD, aplicando o critério escolhido na frequência de AMTs em cada região de sensibilidade (A à G) apuradas nos SDs semelhantes.

A Figura 3.1 mostra o fluxograma detalhando a metodologia proposta para o cálculo dos FI_{base} dos sistemas de distribuição.

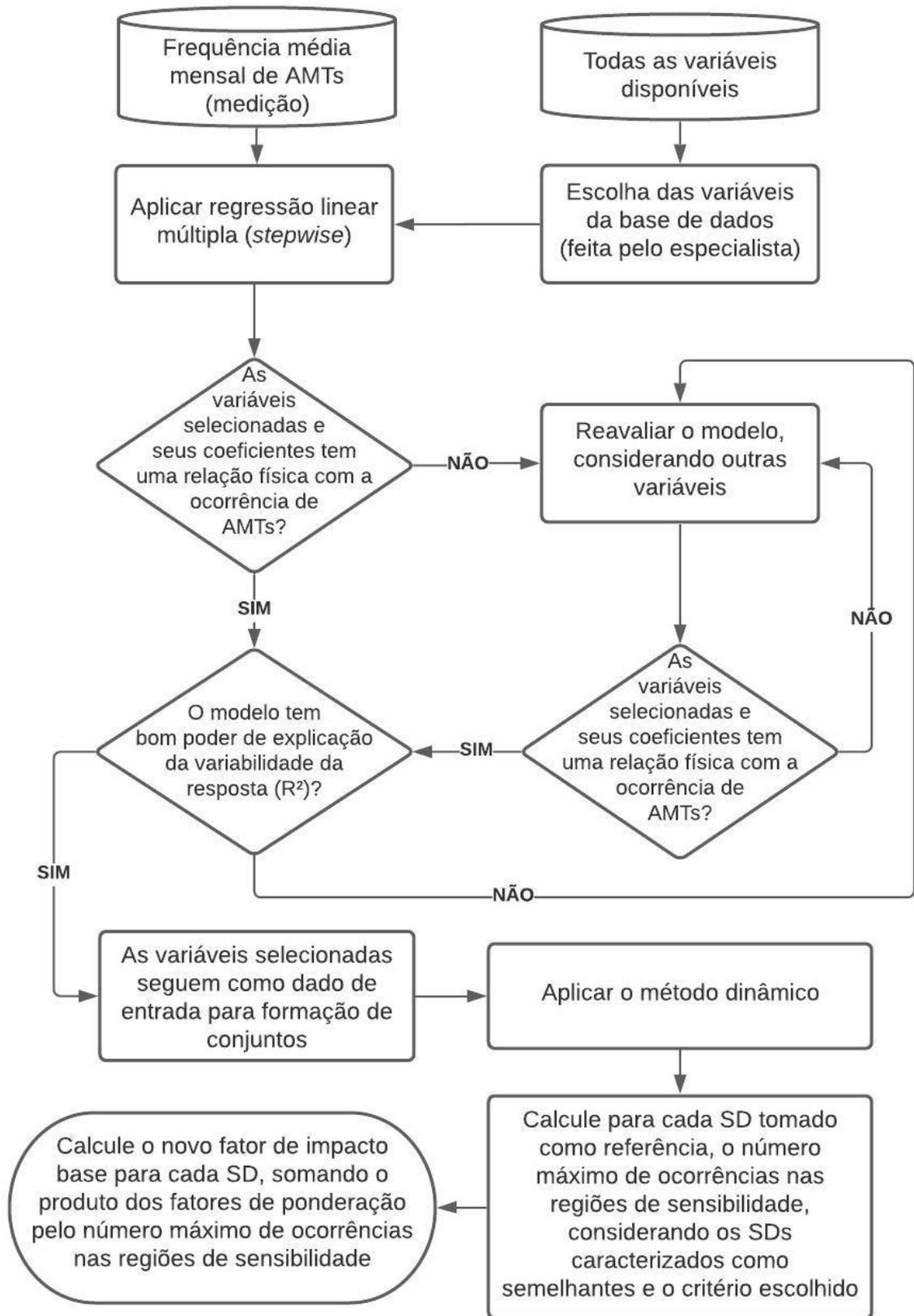


Figura 3.1 – Fluxograma da metodologia proposta.

Fonte: elaborado pelo autor.

3.7 Considerações Finais

O presente capítulo apresentou a metodologia proposta nesta tese para o estabelecimento de fatores de impacto base distintos para cada sistema de distribuição e que consideram o desempenho dos SDs semelhantes.

Os passos do procedimento foram detalhados nas subseções deste capítulo visando facilitar a sua aplicação por outros pesquisadores ou agentes do setor elétrico.

Foi atribuída ao pesquisador a escolha das variáveis para compor a base de dados, o nível de significância na regressão e o critério para estabelecimento do número máximo de ocorrências de afundamentos de tensão. Isto permite que a metodologia seja personalizada de acordo com a disponibilidade de variáveis e com o nível de rigor que se deseja no estabelecimento dos novos fatores de impacto base.

No próximo capítulo a metodologia proposta é aplicada usando dados reais com o intuito de mostrar sua factibilidade.

4. ESTUDO DE CASO E RESULTADOS

4.1 Considerações Iniciais

Neste capítulo é apresentada a aplicação da metodologia proposta para o estabelecimento do fator de impacto base, utilizando dados reais de 19 sistemas de distribuição de média tensão pertencentes a uma concessionária de energia elétrica brasileira.

As etapas da metodologia descritas no capítulo 3 são empregadas para obtenção e análise dos resultados. Fazem-se testes, tanto em relação à abrangência da base de dados considerada quanto ao critério para estabelecimento da frequência máxima de ocorrência do evento, com o objetivo de evidenciar qual o conjunto de variáveis e critérios adotados geram melhores resultados.

É feita uma análise dos resultados obtidos e uma comparação com a forma atual de cálculo do indicador Fator de Impacto, com a finalidade de exibir os ganhos obtidos pela aplicação da metodologia, principalmente com relação ao incentivo a melhora da qualidade da energia e a possível redução das paradas de processos dos consumidores industriais.

4.2 Escolha da Base de Dados

Com o intuito de tornar a metodologia proposta na tese fácil de ser utilizada e aplicável a todos os sistemas de distribuição do Brasil, vislumbrou-se utilizar como base de dados algumas das variáveis que por regulamentação, desde 2011, são enviadas obrigatoriamente para ANEEL, pelas distribuidoras por meio da Base de Dados Geográficos da Distribuidora – BDGD.

Os dados que são enviados pelas concessionárias de energia para o órgão regulador consistem em 146 variáveis, que contemplam informações técnicas da rede de distribuição, ambientais e socioeconômicas. Dentre estas foram escolhidas sete variáveis que se pressupõe ter relação com a ocorrência de afundamentos de tensão. Avaliou-se também a inclusão de duas novas variáveis que atualmente não estão presentes nos dados da BDGD, quais sejam: a área de vulnerabilidade e a taxa de falta do sistema de distribuição.

Portanto, serão testadas duas bases de dados, uma contendo sete variáveis extraídas da BDGD, e outra contendo nove variáveis, sendo as mesmas sete variáveis da BDGD com adição das duas variáveis já citadas, obtidas por meio de cálculo e de dados históricos.

Os valores de frequência média mensal de AMTs foram obtidos a partir de medições realizadas durante um ano em 19 sistemas de distribuição de média tensão ($1 \text{ kV} < V_{\text{nominal}} <$

69 kV) pertencentes a uma concessionária de energia do Brasil, por intermédio de Projeto de Pesquisa e Desenvolvimento – P&D realizado com a parceria da universidade e da distribuidora. A base de dados contendo as variáveis utilizadas nos testes e a frequência média de AMTs é apresentada na Tabela 4.1.

Tabela 4.1 – Base de dados.

| SD | ALIM | NUC_NURB | D_DESC | PC_VRA | PC_TD_1F | PC_TD_NURB | COMP_M_ALIM | TF | AV | AMTs |
|----|------|----------|--------|--------|----------|------------|-------------|-------|--------|------|
| 1 | 4 | 262 | 3,00 | 0,03 | 0,86 | 0,96 | 489,45 | 6,48 | 535,17 | 19 |
| 2 | 4 | 405 | 3,00 | 0,01 | 0,46 | 0,53 | 103,91 | 12,54 | 74,33 | 4 |
| 3 | 4 | 310 | 3,00 | 0,01 | 0,74 | 0,86 | 221,65 | 5,76 | 74,41 | 5 |
| 4 | 2 | 389 | 2,78 | 0,01 | 0,31 | 0,32 | 79,46 | 21,09 | 41,37 | 5 |
| 5 | 4 | 82 | 2,83 | 0,04 | 0,60 | 0,61 | 68,74 | 20,64 | 129,87 | 15 |
| 6 | 2 | 297 | 3,00 | 0,01 | 0,87 | 0,93 | 553,56 | 4,24 | 174,43 | 14 |
| 7 | 3 | 538 | 3,00 | 0,05 | 0,78 | 0,92 | 204,33 | 10,25 | 388,72 | 11 |
| 8 | 5 | 422 | 2,97 | 0,05 | 0,70 | 0,94 | 237,46 | 7,62 | 93,71 | 5 |
| 9 | 5 | 644 | 3,00 | 0,01 | 0,61 | 0,81 | 177,86 | 11,40 | 175,78 | 11 |
| 10 | 4 | 261 | 3,00 | 0,44 | 0,56 | 0,87 | 164,38 | 7,38 | 240,34 | 16 |
| 11 | 11 | 461 | 3,31 | 0,01 | 0,76 | 0,93 | 209,25 | 6,18 | 194,76 | 6 |
| 12 | 3 | 189 | 3,00 | 0,01 | 0,94 | 0,97 | 83,50 | 3,40 | 113,89 | 12 |
| 13 | 17 | 998 | 3,00 | 0,01 | 0,75 | 0,89 | 211,08 | 6,57 | 174,96 | 13 |
| 14 | 4 | 435 | 3,00 | 0,03 | 0,76 | 0,97 | 327,05 | 6,40 | 251,92 | 13 |
| 15 | 4 | 474 | 3,00 | 0,11 | 0,84 | 0,95 | 418,09 | 8,48 | 142,81 | 15 |
| 16 | 6 | 392 | 3,01 | 0,22 | 0,86 | 0,95 | 232,65 | 9,10 | 160,37 | 17 |
| 17 | 6 | 539 | 2,92 | 0,01 | 0,61 | 0,68 | 159,67 | 16,84 | 278,59 | 14 |
| 18 | 1 | 171 | 3,00 | 0,01 | 0,90 | 0,93 | 777,61 | 4,43 | 135,82 | 13 |
| 19 | 3 | 395 | 3,00 | 0,01 | 0,88 | 0,97 | 292,89 | 5,15 | 235,71 | 13 |

Fonte: elaborado pelo autor.

A seguir apresenta-se o significado das siglas das variáveis presentes na Tabela 4.1:

- ALIM – número de alimentadores do sistema de distribuição, esta variável foi extraída da BDGD;
- NUC_NURB – número de unidades consumidoras não urbanas, variável extraída da BDGD;
- D_DESC – densidade de descargas atmosféricas (raios/km²/ano), variável obtida na BDGD;
- PC_VRA – percentual de vegetação remanescente alta, variável presente na BDGD;
- PC_TD_1F – percentual de transformadores monofásicos, variável extraída da BDGD;
- PC_TD_NURB – percentual de transformadores não urbanos, variável obtida na BDGD;
- COMP_M_ALIM – comprimento médio do alimentador (km), variável presente na BDGD;
- TF – taxa de falta (faltas/100 km/ano). Esta variável foi calculada por meio do histórico de ocorrências de faltas de um ano disponibilizado pela distribuidora de energia. Calculou-se o número de faltas que ocorreram durante o ano e dividiu-se este valor pelo comprimento total da rede elétrica do sistema de distribuição, multiplicou-se o resultado por 100 para obter a variável na unidade de medida usual;

- AV – área de vulnerabilidade (km). Esta variável foi calculada por meio do software de simulação de curto-circuito DistriView e de um programa feito em Matlab. Com as redes dos SDs modeladas no software, aplicam-se faltas com impedância nula (pior cenário) igualmente espaçadas em todas as linhas de distribuição do sistema, monitorando para cada falta aplicada o comportamento da tensão no barramento da subestação de interesse. A partir do relatório de saída do programa DistriView identificou-se quais faltas geraram afundamento de tensão na barra monitorada, e com o programa desenvolvido em Matlab contabilizou-se o comprimento da rede no qual a incidência de falta ocasiona AMT no barramento da subestação que alimenta o SD, obtendo-se a área de vulnerabilidade. Todos os tipos de falta foram simulados, portanto a área de vulnerabilidade total foi calculada levando em consideração a probabilidade típica de ocorrência de cada tipo de falta no sistema de distribuição (monofásica = 78%, bifásica = 9%, bifásica-terra = 10% e trifásica = 3%) (GARCIA; ALVES, 2011).
- AMTs – frequência média mensal de AMTs. Esta variável foi obtida por meio de medição no barramento da subestação dos SDs. Foi medida a quantidade de eventos ocorridos durante um ano de operação da rede e dividiu-se o valor obtido por doze meses para encontrar a média mensal. A medição ao longo de um ano foi importante para capturar a característica sazonal da ocorrência dos afundamentos de tensão.

4.3 Seleção das Variáveis

Para esta etapa utilizou-se o procedimento de regressão linear múltipla aplicado por meio do software MINITAB.

Aplicou-se a regressão linear múltipla considerando como variáveis regressoras as escolhidas no passo anterior e como resposta a frequência média mensal de AMTs. Utilizou-se o método de seleção de modelos *stepwise*, com nível de significância de permanência e de entrada das variáveis no modelo iguais a 0,1.

A primeira análise foi feita com a base de dados que contém as sete variáveis escolhidas dentre aquelas disponíveis na BDGD. Os resultados obtidos estão apresentados na Tabela 4.2.

Tabela 4.2 – Resultados da regressão *stepwise* (base de dados com 7 variáveis).

| Análise de Variância | | | | | |
|-----------------------------|--------------------|-------------------|----------------|-----------|-----------|
| Fontes de Variação | Graus de Liberdade | Soma de Quadrados | Quadrado Médio | Valor – F | Valor – P |
| Regressão | 3 | 209,80 | 69,93 | 6,70 | 0,004 |
| D_DESC | 1 | 56,41 | 56,41 | 5,40 | 0,035 |
| PC_VRA | 1 | 78,96 | 78,96 | 7,56 | 0,015 |
| PC_TD_1F | 1 | 151,09 | 151,09 | 14,47 | 0,002 |
| Erro | 15 | 156,62 | 10,44 | | |
| Total | 18 | 366,42 | | | |

| Coefficiente de Determinação | |
|-------------------------------------|------------------------------------|
| R ² | R ² _{ajustado} |
| 57,26% | 48,71% |

| Coefficientes | | | | | |
|----------------------|--------------|-------------------------|-----------|-----------|------|
| Termo | Coefficiente | Erro Padrão Coeficiente | Valor – T | Valor – P | VIF |
| Constantes | 56,3 | 24,4 | 2,31 | 0,036 | |
| D_DESC | -20,11 | 8,65 | -2,32 | 0,035 | 1,27 |
| PC_VRA | 20,02 | 7,28 | 2,75 | 0,015 | 1,02 |
| PC_TD_1F | 19,81 | 5,21 | 3,80 | 0,002 | 1,28 |

| Equação de Regressão |
|--|
| AMTs = 56,3 – 20,11 D_DESC + 20,02 PC_VRA + 19,81 PC_TD_1F |

Fonte: elaborado pelo autor.

Nota-se na Tabela 4.2 que, o Valor – P das variáveis selecionadas são inferiores a 0,05 mostrando serem significativas no modelo. O VIF das variáveis são todos inferiores a 5 demonstrando baixa colinearidade entre elas. O valor encontrado para o coeficiente de determinação ajustado 48,71% é razoável. No entanto o coeficiente atribuído a variável densidade de descarga atmosférica tem sinal negativo, que está em desacordo com o esperado. Segundo (BOLLEN, 1999), uma grande parte dos afundamentos de tensão são devidos a descargas atmosféricas em linhas aéreas. Portanto, espera-se que quanto maior a densidade de descarga atmosféricas maior será o número de ocorrências de afundamentos de tensão.

Com o objetivo de buscar um modelo mais representativo do fenômeno, decidiu-se por incluir na base de dados, outras duas variáveis já citadas e consideradas importantes pelos pesquisadores: a taxa de falta e a área de vulnerabilidade. Portanto, aplicando-se novamente a regressão *stepwise*, considerando agora nove variáveis regressoras, foram obtidos os resultados mostrados na Tabela 4.3.

Tabela 4.3 – Resultados da regressão *stepwise* (base de dados com 9 variáveis).

| Análise de Variância | | | | | |
|-----------------------------|--------------------|-------------------|----------------|-----------|-----------|
| Fontes de Variação | Graus de Liberdade | Soma de Quadrados | Quadrado Médio | Valor – F | Valor – P |
| Regressão | 4 | 271,26 | 67,82 | 9,98 | 0,000 |
| PC_VRA | 1 | 79,71 | 79,71 | 11,73 | 0,004 |
| PC_TD_1F | 1 | 106,12 | 106,12 | 15,61 | 0,001 |
| TF | 1 | 60,99 | 60,99 | 8,97 | 0,010 |
| AV | 1 | 36,81 | 36,81 | 5,42 | 0,035 |
| Erro | 14 | 95,16 | 6,80 | | |
| Total | 18 | 366,42 | | | |

| Coefficiente de Determinação | |
|-------------------------------------|------------------------------------|
| R ² | R ² _{ajustado} |
| 74,03% | 66,61% |

| Coefficientes | | | | | |
|----------------------|--------------|-------------------------|-----------|-----------|------|
| Termo | Coefficiente | Erro Padrão Coeficiente | Valor – T | Valor – P | VIF |
| Constante | -16,25 | 6,16 | -2,64 | 0,020 | |
| PC_VRA | 20,88 | 6,10 | 3,42 | 0,004 | 1,10 |
| PC_TD_1F | 25,89 | 6,55 | 3,95 | 0,001 | 3,11 |
| TF | 0,589 | 0,197 | 3,00 | 0,010 | 2,81 |
| AV | 0,01321 | 0,00568 | 2,33 | 0,035 | 1,18 |

| Equação de Regressão |
|--|
| AMTs = - 16,25 + 20,88 PC_VRA + 25,89 PC_TD_1F + 0,589 TF + 0,01321 AV |

Fonte: elaborado pelo autor.

Observa-se na Tabela 4.3 que todas as variáveis selecionadas apresentaram Valor – P inferior ao limite de 0,05, indicando serem significativas no modelo. Além disto, os valores de VIF são todos inferiores a 5, mostrando baixa colinearidade entre as variáveis selecionadas.

O novo modelo obtido apresentou um valor satisfatório para o parâmetro R²_{ajustado} = 66,61%, cerca de 36,75% superior ao modelo previamente avaliado, mostrando que a adição das variáveis taxa de falta e área de vulnerabilidade na base de dados foi relevante, pois estas foram selecionadas para fazer parte do modelo e aumentou significativamente o poder de explicação da variabilidade da resposta. Possivelmente, a variável densidade de descargas atmosféricas não foi selecionada pela regressão por estar intrinsecamente presente na variável taxa de falta que foi selecionada para compor o modelo (variáveis com elevado grau de colinearidade).

As suposições do modelo ajustado precisam ser validadas para que os resultados sejam confiáveis. Com este objetivo aplicou-se a análise de resíduos que é um conjunto de técnicas utilizadas para investigar a adequabilidade de um modelo de regressão com base nos resíduos.

Resumidamente, se o modelo for apropriado, os resíduos devem refletir as propriedades impostas pelo termo de erro do modelo, que são: normalidade, homocedasticidade e independência.

Fazendo o diagnóstico dos resíduos com relação às três propriedades listadas, têm-se respectivamente os resultados apresentados nas Figuras 4.1 à 4.3.

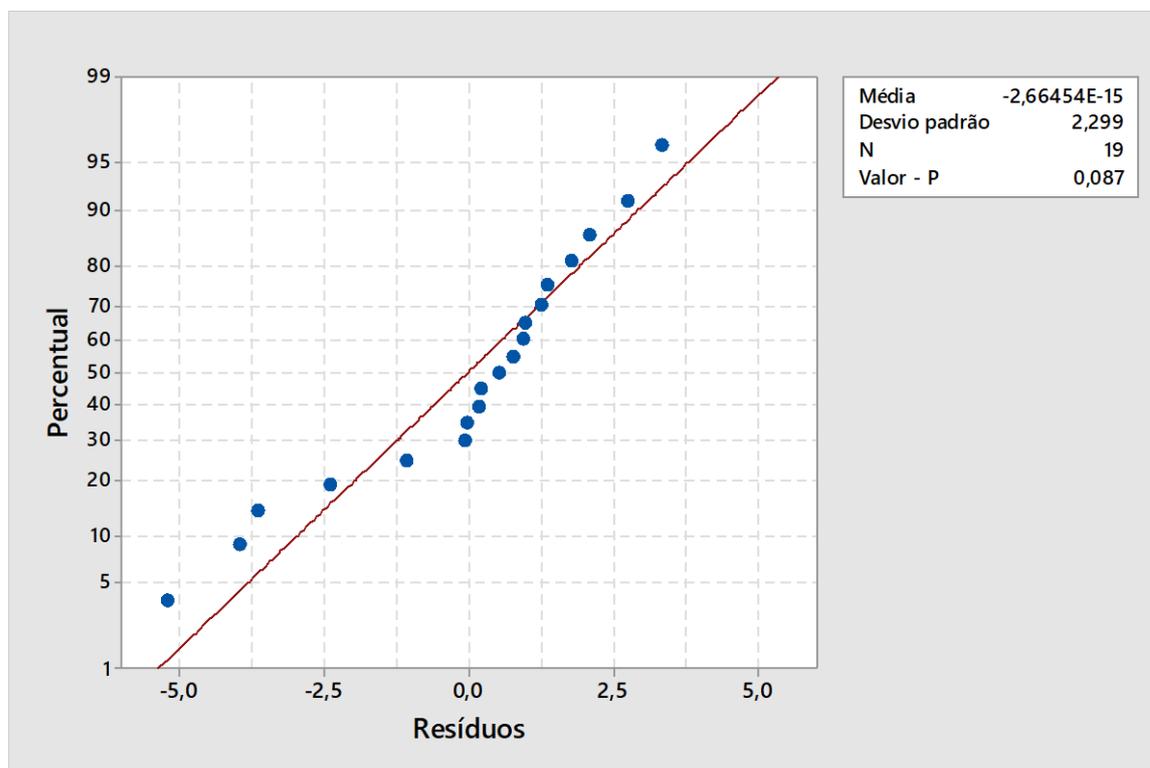


Figura 4.1 – Papel de probabilidade.

Fonte: elaborado pelo autor.

Observa-se na Figura 4.1 que os pontos seguem o comportamento da reta, portanto tem-se indícios que os resíduos são normalmente distribuídos. Adicionalmente, o resultado do teste de Shapiro-Wilk que apresentou Valor – P igual a 0,087 valor superior a 0,05 confirma a normalidade dos resíduos.

Homocedasticidade é o termo para designar variância constante dos erros experimentais para observações distintas. O gráfico dos resíduos versus valores ajustados é uma das principais técnicas utilizadas para verificar a suposição de homocedasticidade, caso não haja nenhum comportamento ou tendência neste gráfico, têm-se indícios que a variância dos resíduos é homocedástica.

Nota-se na Figura 4.2 que os pontos estão aleatoriamente distribuídos em torno do 0, sem apresentar uma tendência, portanto é um indicativo da homocedasticidade dos resíduos.

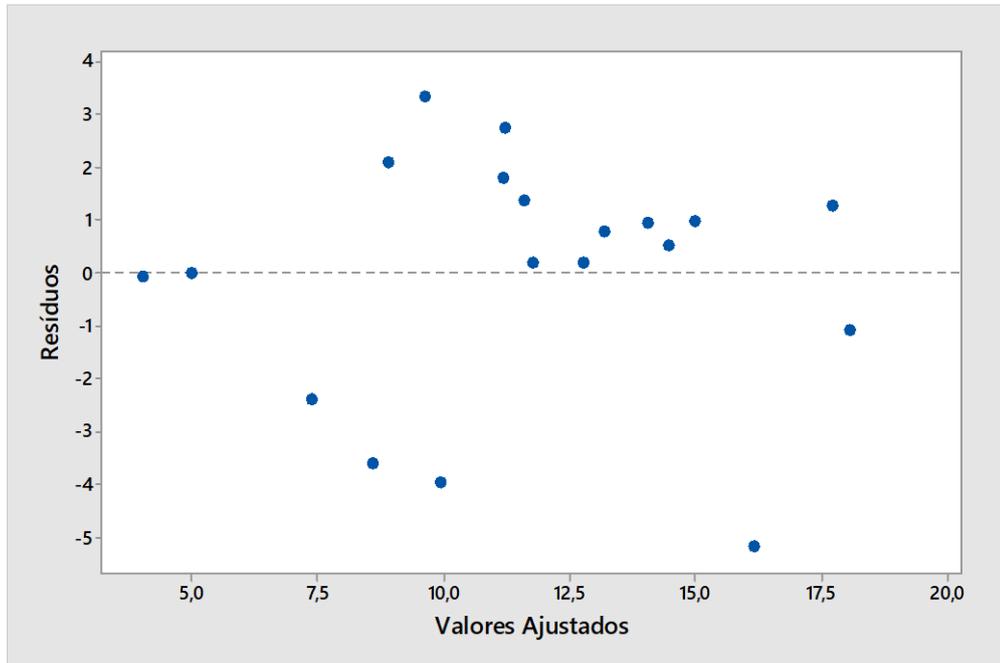


Figura 4.2 – Gráfico dos resíduo x valores ajustados.

Fonte: elaborado pelo autor.

Para verificar a hipótese de independência dos resíduos, pode-se analisar o gráfico dos resíduos versus a ordem de coleta dos dados. Se houver uma tendência dos pontos, ou seja, se estes tiverem um comportamento que se repete em determinada região do gráfico, têm-se indícios de dependência dos resíduos. Analisando a Figura 4.3, observa-se que os pontos do gráfico não parecem ter uma tendência, portanto qualitativamente tem-se indícios de independência dos resíduos.

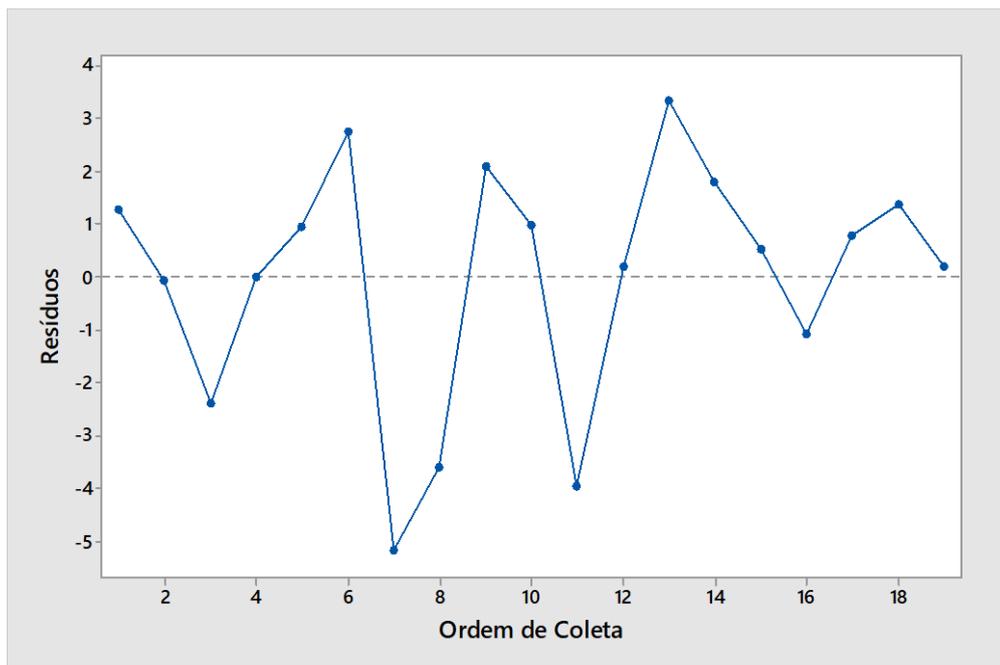


Figura 4.3 – Gráfico resíduos x ordem de coleta.

Fonte: elaborado pelo autor.

Vale ressaltar que qualquer modelo encontrado por método estatístico deve ser apreciado por um especialista da área que está sendo aplicado, para verificação das variáveis selecionadas e seus coeficientes, quanto ao sentido físico que estas de fato possuem com o fenômeno sob análise. Fazendo esta análise crítica do modelo obtido, pode-se citar que, é plausível a seleção da variável percentual de vegetação remanescente alta, visto que uma das fontes de curto-circuito são as árvores que podem entrar em contato com a rede de distribuição. O percentual de transformadores monofásicos traz indiretamente a informação do percentual de redes rurais, pois transformadores monofásicos são normalmente usados nestas. Sendo assim, esta variável também possui explicação do ponto de vista da engenharia elétrica, uma vez que redes rurais estão mais expostas à ação de animais e tendem a ter manutenção menos frequente comparada às redes urbanas. A área de vulnerabilidade está diretamente relacionada com o fenômeno, visto que ela representa a região do sistema elétrico sob a qual a ocorrência de falta irá gerar afundamento de tensão na barra monitorada. De forma semelhante, sabe-se que a taxa de falta também está fortemente relacionada à ocorrência de afundamentos de tensão, uma vez que os curtos-circuitos são as principais causas deste fenômeno.

Vale salientar que o modelo gerado não será utilizado posteriormente para fazer previsão da frequência média de AMTs. Apenas as variáveis selecionadas (percentual de vegetação remanescente alta, percentual de transformadores monofásicos, taxa de falta e área de vulnerabilidade) serão utilizadas na etapa seguinte de formação de conjuntos de sistemas de distribuição.

4.4 Formação de Conjuntos

As variáveis selecionadas no passo anterior são utilizadas como dados de entrada para o método dinâmico de formação de conjuntos. Os cálculos das distâncias euclidianas e heterogeneidades entre SDs assim como o ordenamento destes dados foram feitos em planilhas do Excel.

Para implementação do método dinâmico, faz-se necessária a criação de tabelas ordenando de forma crescente as distâncias entre elementos, para cada elemento (SD) tomado como referência, além do estabelecimento do critério de máxima heterogeneidade permitida. As tabelas contendo as distâncias entre SDs e as heterogeneidades percentuais estão no Apêndice A. Como exemplo, apresenta-se na Tabela 4.4 as distâncias e heterogeneidades em ordem crescente para o sistema de distribuição 8, que foi o SD com maior número de SDs semelhantes.

Tabela 4.4 – Distância e heterogeneidade entre SDs (SD 8 como referência).

| SD | Distância | Heterogeneidade |
|-----------|------------------|------------------------|
| 8 | 0,00 | 0,00% |
| 3 | 0,60 | 9,96% |
| 13 | 0,87 | 14,49% |
| 11 | 1,04 | 17,32% |
| 15 | 1,08 | 17,97% |
| 9 | 1,19 | 19,80% |
| 14 | 1,42 | 23,69% |
| 18 | 1,46 | 24,32% |
| 6 | 1,47 | 24,51% |
| 12 | 1,73 | 28,78% |
| 19 | 1,75 | 29,21% |
| 2 | 1,80 | 29,96% |
| 16 | 1,98 | 33,00% |
| 17 | 2,46 | 40,99% |
| 5 | 2,57 | 42,91% |
| 7 | 2,61 | 43,50% |
| 4 | 3,53 | 58,87% |
| 1 | 3,89 | 64,88% |
| 10 | 3,96 | 65,94% |

Fonte: elaborado pelo autor.

Observa-se na Tabela 4.4 que, pelo limite considerado de 30% de heterogeneidade o sistema de distribuição 8 possui 11 SDs semelhantes.

Pela análise das tabelas do Apêndice A, nota-se que considerando como limite uma heterogeneidade percentual máxima de 30%, onze sistemas de distribuição apresentaram conjuntos com quantidade razoável de SDs semelhantes, seis sistemas apresentaram conjuntos com poucos SDs semelhantes, e os SDs 4 e 10 não possuem nenhum SD semelhante para este nível de heterogeneidade.

Os sistemas que não apresentaram SD semelhante se devem ao número reduzido de SDs na amostra, onde provavelmente nenhum sistema de distribuição considerado possui características semelhantes. Mas como a metodologia proposta neste trabalho visa aplicação mais abrangente, tal impasse certamente não acontecerá quando a base de dados for ampliada contemplando outros SDs do Brasil.

Ressalta-se que, o valor para a heterogeneidade máxima de 30% foi escolhido para mostrar a aplicação da metodologia sobre a amostra de dados de 19 sistemas de distribuição. Portanto, este valor não é uma proposta definitiva, deve ser reavaliado em função do tamanho da amostra. Conforme mencionado no item 3.4, à medida que a amostra aumenta é possível reduzir o valor da heterogeneidade máxima permitida.

4.5 Estabelecimento do Fator de Impacto base

Tomando como exemplo o SD 8, tem-se a Tabela 4.5, que apresenta a frequência média mensal de AMTs medida nos sistemas de distribuição semelhantes, estratificada nas regiões de sensibilidade A à G. Os valores médios mensais da frequência de AMTs nas regiões de sensibilidade de A à G para todos os SDs sob análise estão apresentados na tabela do Apêndice B.

Tabela 4.5 – Frequência de AMTs nas regiões de sensibilidade (SD 8 como referência).

| SD | Frequência de AMTs nas regiões | | | | | | |
|----|--------------------------------|------|------|------|------|------|------|
| | A | B | C | D | E | F | G |
| 8 | 3,08 | 0,00 | 0,00 | 0,50 | 0,25 | 0,50 | 0,58 |
| 3 | 3,25 | 0,08 | 0,00 | 0,50 | 0,00 | 0,17 | 1,25 |
| 13 | 6,67 | 0,67 | 0,33 | 1,92 | 1,75 | 0,42 | 1,42 |
| 11 | 2,08 | 0,75 | 0,17 | 0,83 | 0,67 | 0,92 | 0,83 |
| 15 | 8,55 | 1,36 | 0,09 | 2,73 | 0,18 | 0,18 | 1,45 |
| 9 | 3,58 | 0,17 | 0,08 | 1,33 | 0,33 | 1,50 | 1,25 |
| 14 | 7,60 | 0,40 | 0,00 | 1,10 | 0,30 | 0,70 | 2,70 |
| 18 | 6,13 | 0,75 | 0,13 | 4,25 | 0,38 | 0,63 | 0,88 |
| 6 | 7,08 | 0,67 | 0,33 | 3,58 | 0,33 | 1,00 | 0,92 |
| 12 | 6,00 | 0,75 | 0,42 | 2,25 | 0,67 | 1,58 | 0,67 |
| 19 | 8,89 | 0,33 | 0,22 | 1,33 | 1,11 | 0,56 | 1,00 |
| 2 | 1,83 | 0,50 | 0,25 | 0,92 | 0,08 | 0,25 | 0,25 |

Fonte: elaborado pelo autor.

Considerando inicialmente o percentil 20 dos dados presentes em cada coluna da Tabela 4.5, tem-se para cada região de sensibilidade o número máximo de ocorrências proposto para o SD sob análise. Partindo destes valores e utilizando os fatores de ponderação indicados na Tabela 2.6, calcula-se um novo FI_{base} para o sistema de distribuição 8.

A Tabela 4.6 mostra as regiões de sensibilidade consideradas no cálculo do Fator de Impacto, o fator de ponderação e o número máximo de ocorrências relativo às regiões de sensibilidade de A à G usado pela ANEEL e o calculado pelo procedimento proposto. Destaca-se que foram suprimidas desta tabela as regiões H e I referentes a elevações de tensão, pois como já justificado o foco da pesquisa são os afundamentos de tensão.

Com os valores de número máximo de ocorrências do SD 8, calcula-se o fator de impacto base, por meio do somatório dos produtos do fator de ponderação pelo número máximo de ocorrências, obtendo-se neste caso 0,25.

Tabela 4.6 – Fatores de ponderação e limites para a frequência de AMTs em cada região de sensibilidade (percentil 20).

| Região de Sensibilidade | Fator de Ponderação | Nº Máximo de Ocorrências (ANEEL) | Nº Máximo de Ocorrências (SD 8) |
|--|---------------------|----------------------------------|---------------------------------|
| A | 0 | - | 2,48 |
| B | 0,04 | 5 | 0,12 |
| C | 0,07 | 4 | 0,00 |
| D | 0,15 | 3 | 0,63 |
| E | 0,25 | 2 | 0,12 |
| F | 0,36 | 1 | 0,21 |
| G | 0,07 | 4 | 0,62 |
| Fator de Impacto Base (FI_{base}) | | 2,07 | 0,25 |

Fonte: elaborado pelo autor.

Observa-se que o fator de impacto base encontrado para o sistema de distribuição analisado é muito inferior ao estabelecido pelo PRODIST. Isto se deve ao menor número máximo de ocorrências de AMTs nas regiões A à G calculado pelo critério do percentil 20.

Empregando esta metodologia de cálculo, os FI_{base} e os FI (em p.u.) dos demais SDs deste estudo de caso são mostrados na Tabela 4.7.

Tabela 4.7 – FI_{base} e FI dos sistemas de distribuição (percentil 20).

| SD | FI _{base} | FI |
|----|--------------------|------|
| 1 | - | - |
| 2 | - | - |
| 3 | 0,27 | 0,84 |
| 4 | - | - |
| 5 | - | - |
| 6 | 0,24 | 4,56 |
| 7 | - | - |
| 8 | 0,25 | 1,43 |
| 9 | 0,19 | 4,86 |
| 10 | - | - |
| 11 | 0,27 | 2,67 |
| 12 | 0,24 | 4,92 |
| 13 | 0,27 | 3,79 |
| 14 | 0,30 | 2,32 |
| 15 | 0,30 | 2,28 |
| 16 | - | - |
| 17 | - | - |
| 18 | 0,24 | 4,40 |
| 19 | 0,27 | 2,88 |

Fonte: elaborado pelo autor.

Observa-se que para os SDs que possuem poucos sistemas de distribuição que se assemelham (SD 1, 2, 4, 5, 7, 10, 16 e 17) não é possível calcular o FI_{base} utilizando o critério

do percentil 20.

Nota-se também que, apesar dos fatores de impacto base serem calculados individualmente para cada SD, os sistemas cujo método de formação de conjuntos caracterizou como semelhantes possuem valores de FI_{base} próximos.

Como também se observa na Tabela 4.7, com o cálculo do novo FI_{base} , considerando o percentil 20 da frequência média mensal de AMTs do grupo, apenas 1 SD (valor destacado na cor vermelha) apresentou Fator de Impacto inferior ao valor de referência de 1 p.u.. Portanto, possivelmente o uso do percentil 20 como critério para estabelecimento do FI_{base} seja demasiadamente severo. Ao utilizar o percentil 20 como critério estão sendo selecionados os bons desempenhos de cada região para compor o FI_{base} , gerando assim um valor base subestimado.

Com o objetivo de tornar os valores mais adequados, uma possível solução para contornar o problema identificado é adotar a média da frequência de AMTs dos SDs do conjunto, assim como recomenda a IEEE 1564 para o cálculo de índices de sistema a partir dos valores de índices dos barramentos.

Empregando então a média da frequência de AMTs no grupo dos SDs semelhantes, obtêm-se para o SD 8 os resultados apresentados na Tabela 4.8, onde observa-se que o FI_{base} encontrado foi de 0,75.

Tabela 4.8 – Fatores de ponderação e limites para a frequência de AMTs em cada região de sensibilidade (média).

| Região de Sensibilidade | Fator de Ponderação | Nº Máximo de Ocorrências (ANEEL) | Nº Máximo de Ocorrências (SD 8) |
|---|----------------------------|---|--|
| A | 0 | - | 5,40 |
| B | 0,04 | 5 | 0,54 |
| C | 0,07 | 4 | 0,17 |
| D | 0,15 | 3 | 1,77 |
| E | 0,25 | 2 | 0,50 |
| F | 0,36 | 1 | 0,70 |
| G | 0,07 | 4 | 1,10 |
| Fator de Impacto Base (FI_{base}) | | 2,07 | 0,75 |

Fonte: elaborado pelo autor.

Adotando procedimento semelhante ao descrito anteriormente obtêm-se os valores de FI_{base} e FI apresentados na Tabela 4.9 para todos os sistemas de distribuição, com exceção do SD 4 e 10 que não possuem sistemas semelhantes. Os valores destacados em vermelho correspondem aos SDs que apresentaram FI inferior ao valor de referência de 1 p.u..

Tabela 4.9 – FI_{base} e FI dos sistemas de distribuição (média).

| SD | FI_{base} | FI |
|-----------|-------------------------------|-----------|
| 1 | 0,89 | 1,07 |
| 2 | 0,53 | 0,57 |
| 3 | 0,79 | 0,28 |
| 4 | - | - |
| 5 | 1,56 | 1,26 |
| 6 | 0,84 | 1,30 |
| 7 | 0,81 | 1,02 |
| 8 | 0,75 | 0,48 |
| 9 | 0,68 | 1,37 |
| 10 | - | - |
| 11 | 0,79 | 0,91 |
| 12 | 0,78 | 1,51 |
| 13 | 0,79 | 1,29 |
| 14 | 0,80 | 0,87 |
| 15 | 0,84 | 0,81 |
| 16 | 0,95 | 1,28 |
| 17 | 1,34 | 0,85 |
| 18 | 0,78 | 1,35 |
| 19 | 0,79 | 0,99 |

Fonte: elaborado pelo autor.

Observa-se pela Tabela 4.9 que a proposta de usar como critério a média da frequência de AMTs do grupo parece ser mais adequada para o estabelecimento do fator de impacto base comparado ao uso do percentil 20. Com este procedimento 53% dos SDs apresentaram Fator de Impacto acima do limite de 1 p.u.. Portanto, caso o valor de referência se torne futuramente um limite, a concessionária responsável pelo sistema de distribuição será motivada a implementar ações de melhorias nos SDs como: poda da vegetação próxima a rede, substituição de rede convencional nua por rede compacta ou isolada, melhorias na manutenção da rede, etc.

Para comparação da metodologia proposta, com o estado atual da regulamentação, a Tabela 4.10, mostra o FI_{base} adotado pelo PRODIST e o Fator de Impacto obtido para os SDs sob análise.

Tabela 4.10 – FI_{base} e FI dos sistemas de distribuição (regulamentação atual).

| SD | FI_{base} | FI |
|-----------|-------------------------------|-----------|
| 1 | 2,07 | 0,46 |
| 2 | 2,07 | 0,15 |
| 3 | 2,07 | 0,11 |
| 4 | 2,07 | 0,33 |
| 5 | 2,07 | 0,95 |
| 6 | 2,07 | 0,53 |
| 7 | 2,07 | 0,40 |
| 8 | 2,07 | 0,17 |
| 9 | 2,07 | 0,45 |
| 10 | 2,07 | 0,64 |
| 11 | 2,07 | 0,35 |
| 12 | 2,07 | 0,57 |
| 13 | 2,07 | 0,49 |
| 14 | 2,07 | 0,34 |
| 15 | 2,07 | 0,33 |
| 16 | 2,07 | 0,58 |
| 17 | 2,07 | 0,56 |
| 18 | 2,07 | 0,51 |
| 19 | 2,07 | 0,38 |

Fonte: elaborado pelo autor.

Considerando o valor único de 2,07 do FI_{base} usado no PRODIST, o indicador fator de impacto de todos os SDs usados neste estudo de caso estão abaixo do valor de referência de 1 p.u. (valores destacados na cor vermelha), mostrando que este FI_{base} atual é brando, pois todos SDs estariam em acordo com a regulamentação, não motivando melhorias por parte das distribuidoras de energia. Portanto, um consumidor industrial hipotético que esteja conectado a qualquer sistema de distribuição e que tenha um processo sensível a afundamentos de tensão caracterizados pelas regiões B, C, D, E, F e G, poderá sofrer mais de 19 paradas de processo por mês, sem que o Fator de Impacto exceda 1 p.u.. Em muitos sistemas industriais este número mensal de paradas de produção pode resultar em elevadas perdas financeiras.

Com a metodologia proposta, o número de paradas de processo permitido será reduzido, favorecendo os consumidores que possuem cargas sensíveis aos afundamentos de tensão. A Figura 4.4, mostra uma comparação gráfica dos fatores de impacto calculados considerando a metodologia proposta e o procedimento estabelecido no PRODIST.

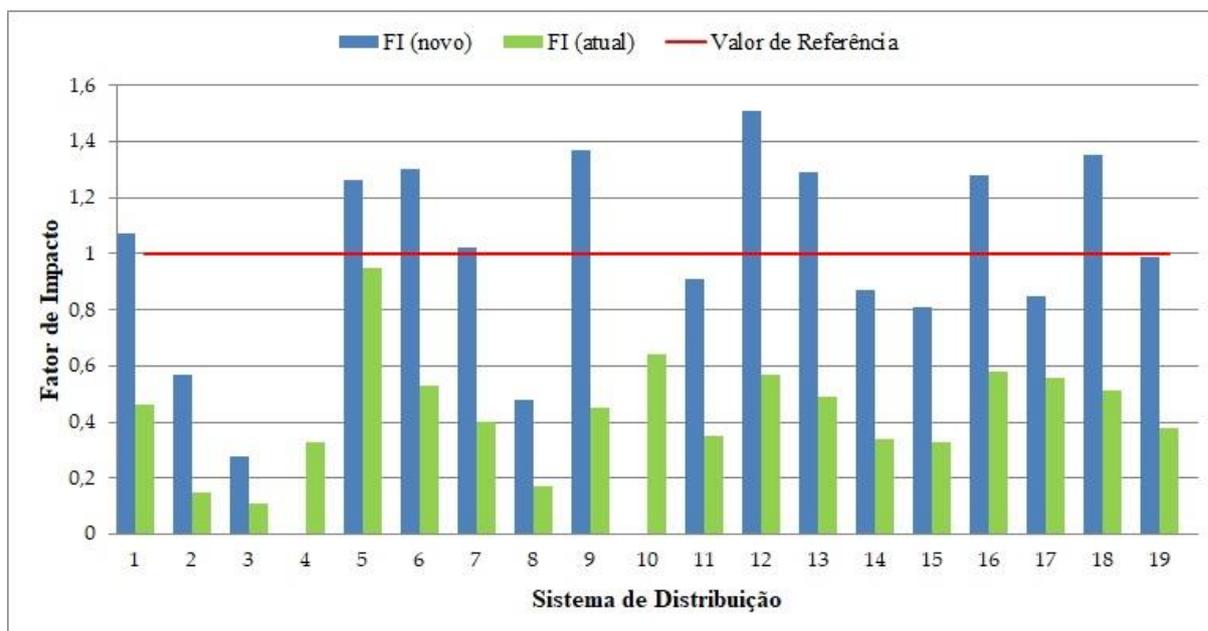


Figura 4.4 – Comparação do Fator de Impacto novo e atual com o valor de referência.

Fonte: elaborado pelo autor.

Observa-se na Figura 4.4 que para todos os SDs sob análise neste estudo de caso, o Fator de Impacto calculado por meio da metodologia proposta é superior ao indicador calculado com os valores da regulamentação atual. Todavia, nota-se que em parte dos SDs os novos fatores de impacto ultrapassam o valor de referência indicando a necessidade de melhoramento nestes por parte da concessionária de energia.

4.6 Considerações Finais

Nota-se que a metodologia proposta é factível e de fácil implementação, pois emprega técnicas conhecidas na área de estatística. Ressalta-se ainda que a metodologia proposta é similar à utilizada para estabelecimento dos limites dos indicadores de continuidade do serviço, facilitando assim, a sua aceitação pelos agentes do setor elétrico para o estabelecimento do FI_{base} .

No processo de seleção de variáveis nota-se que há um ganho significativo no poder de explicação do modelo gerado, quando considerada a base de dados contendo as variáveis taxa de falta de área de vulnerabilidade. Portanto, recomenda-se que estas variáveis sejam incluídas na BDGD.

Pela análise dos resultados obtidos, provavelmente a etapa de estabelecimento do fator de impacto base não poderá utilizar o critério do percentil 20. Uma proposta aparentemente viável seria adotar a média dos valores observados no agrupamento de SDs.

5. CONCLUSÕES

5.1 Conclusões Gerais

Os afundamentos de tensão têm causado grandes prejuízos monetários aos consumidores industriais que possuem cargas sensíveis. Em virtude disto, a ABRACE tem solicitado ao órgão regulamentador, a ANEEL, que tal distúrbio tenha formas coerentes de se quantificar (indicadores) e principalmente valores limites que atendam o nível da qualidade de energia demandada pelas indústrias. A revisão 8 do Módulo 8 do PRODIST, vigente desde o início de 2017, cobre parte desta necessidade ao estabelecer um indicador para caracterização do desempenho dos barramentos frente as VTCDs, o Fator de Impacto, mas estipula um valor de referência único para tal indicador igual a 1 p.u. e um fator de impacto base igual para todos os sistemas de distribuição.

Vislumbra-se que futuramente haja mudanças no PRODIST quanto à regulamentação das VTCDs e acredita-se que o procedimento mais adequado a ser adotado deve ser semelhante ao usado atualmente para determinar os limites dos indicadores de continuidade (DEC e FEC). Assim, o PRODIST poderá evoluir para o estabelecimento de fatores de impacto base distintos para cada sistema de distribuição de acordo com o desempenho dos sistemas que mais se assemelham.

Neste contexto, este trabalho de doutorado se alinha aos anseios do setor elétrico, apresentando uma metodologia para o estabelecimento de fator de impacto base distinto para cada sistema de distribuição, que leva em consideração o desempenho dos sistemas semelhantes, com relação às variáveis que mais influenciam a ocorrência de afundamentos de tensão.

A metodologia exposta apresenta facilidade de aplicação, sendo baseada em métodos estatísticos para seleção das variáveis de maior influência, para o estabelecimento dos SDs semelhantes com relação a tais variáveis e posteriormente para a determinação do fator de impacto base para cada sistema de distribuição.

A capacidade da metodologia proposta de ser estendida a uma base de dados maior que contemple outros sistemas de distribuição é um fator importante, pois viabiliza a sua utilização de forma mais ampla para gerar valores mais adequados para o FI_{base} . Pelos resultados obtidos no estudo de caso, observou-se que as variáveis taxa de falta e área de vulnerabilidade contribuem substancialmente para o aumento do poder de explicação do

modelo, então recomenda-se que órgão regulador solicite a inclusão destas na BDGD, permitindo desta forma a aplicação da metodologia em todos os sistemas de distribuição do Brasil.

A contribuição deste trabalho em estabelecer valores base mais adequados para o indicador de VTCDs visa auxiliar tanto as concessionárias de energia como os consumidores. Os consumidores industriais poderão, mediante o fator de impacto base estabelecido para o sistema de distribuição que estão conectados, realizar estudos dos impactos das VTCDs nos processos produtivos e análises mais realistas da viabilidade econômica de instalação de condicionadores de energia.

As distribuidoras de energia elétrica que realizam boas práticas de manutenção preventiva, poda da vegetação próxima a rede, e que também investem em atualizações da rede, terão sistemas de distribuição com Fator de Impacto provavelmente abaixo do valor de referência, evitando assim eventuais gastos para defesa na esfera judicial, e podendo ainda utilizar estes valores como um fator de *marketing* para atrair novos consumidores industriais.

Caso outros países adotem um indicador similar ao Fator de Impacto, mesmo que as regiões de sensibilidade consideradas, os fatores de ponderação e as variáveis disponíveis não sejam iguais, a metodologia proposta é genérica e poderá ser aplicada seguindo os mesmos passos.

5.2 Proposta para Trabalhos Futuros

Como possíveis desenvolvimentos necessários para dar sequência a linha de pesquisa desta tese, pode-se citar:

- Aplicar e validar a metodologia proposta em uma base de dados que contemple maior número de sistemas de distribuição, preferencialmente pertencentes a diversas empresas distribuidoras de energia elétrica;
- Reavaliar a etapa de seleção de variáveis adicionando novas variáveis que se supõe ter relação com a ocorrência de afundamentos de tensão, com o intuito de aumentar o poder de explicação do modelo gerado;
- Aplicar a metodologia proposta considerando as regiões de sensibilidade e os fatores de ponderação sugeridos em (COSTA et al., 2019);
- Aplicar a metodologia usando simulação pelo método de Monte Carlo para a obtenção da

frequência média mensal de afundamentos de tensão, para os sistemas de distribuição que não realizam medições permanentes do distúrbio.

- Testar outros métodos de agrupamentos (Ward, K-means, etc) e comparar com os resultados obtidos nesta tese.

5.3 Artigos Publicados

Diretamente relacionado com o tema desta tese, foram publicados os artigos:

SOUZA, P. V.; COSTA, M. V.; RIBEIRO, H. K.; CARVALHO, J. M. Formação de Conjuntos de Sistemas de Distribuição quanto às Variações de Curta Duração. **XII Conferência Brasileira sobre Qualidade da Energia Elétrica – CBQEE**, 2017.

SOUZA, P. V.; CARVALHO, J. M.; FERREIRA, D. F.; RIBEIRO, H. K.; FILHO, J. M.; PEREIRA, N. B. Uma Contribuição para o Estabelecimento de Limites para o Indicador de VTCDs. **XIII Conferência Brasileira sobre Qualidade da Energia Elétrica – CBQEE**, 2019.

SOUZA, P. V.; CARVALHO, J. M.; FERREIRA, D. F.; RIBEIRO, H. K.; FILHO, J. M.; PEREIRA, N. B. Cluster-based Method to Determine Base Values for Short-Term Voltage Variation Indices. **Energies - Volume 14, issue 1**, janeiro 2021.

De forma mais abrangente no assunto QEE e geração distribuída, durante o curso de doutorado foram publicados os artigos:

RIBEIRO, H. K.; CARVALHO, J. M.; SOUZA, P. V.; OLIVEIRA, T. C. Método Alternativo para Cálculo de Indicadores Médios de Continuidade e VTCDs em Sistemas de Distribuição. **XII Conferência Brasileira sobre Qualidade da Energia Elétrica – CBQEE**, 2017.

BERTOLIN, A. L.; FERREIRA, S. C.; LEITE, D. F.; SOUZA, P. V.; PACHECO, V. M.; BARBOSA, L. M. Cálculo de Potências de Fornos a Arco: Considerações sobre a Recomendação IEEE 1459-2010 em Sistemas com Flutuação de Tensão e Corrente. **XII Conferência Brasileira sobre Qualidade da Energia Elétrica – CBQEE**, 2017.

MINAMI, J. P.; BRAGA, M. D.; SOUZA, P. V.; RIBEIRO, P. F.; BONATTO, B. D. Aspectos Econômicos da Inserção de Geração Fotovoltaica Considerando Diferentes Cenários. **XII Conferência Brasileira sobre Qualidade da Energia Elétrica – CBQEE**, 2017.

REFERÊNCIAS

- ALMEIDA, F. A. et al. Enhancement of discriminatory power by ellipsoidal functions for substation clustering in voltage sag studies. **Electric Power Systems Research**, v. 185, 2020.
- ANEEL. **Módulo 8 do PRODIST - Procedimentos de Distribuição de Energia Elétrica do Sistema Elétrico Nacional (revisão 12)**, 2021. Disponível em:
<https://www.aneel.gov.br/documents/656827/14866914/M%C3%B3dulo_8-Revis%C3%A3o_12/342ff02a-8eab-2480-a135-e31ed2d7db47>
- ASHEIBI, A.; STIRLING, D.; SUTANTO, D. **Determination of the optimal number of clusters in harmonic data classification**. 2008 13th International Conference on Harmonics and Quality of Power. **Anais...**2008
- BOLLEN, M. H. **Understanding Power Quality Problems**. [s.l.] IEEE, 1999.
- CARVALHO FILHO, J. M. **Uma Contribuição à Avaliação do Atendimento a Consumidores com Cargas Sensíveis – Proposta de Novos Indicadores**. [s.l.] Escola Federal de Engenharia de Itajubá, 2000.
- CHARNET, R. et al. **Análise de Modelos de Regressão Linear: com aplicação**. 2. ed. Campinas: Editora da Unicamp, 2008.
- CORRAR, L. J.; PAULO, E.; FILHO, J. M. D. **Análise Multivariada para os Cursos de Administração, Ciências Contábeis e Economia**. 1. ed. São Paulo: Editora Atlas S.A., 2009.
- COSTA, M. V. **Uma Contribuição à Regulamentação das Variações de Tensão de Curta Duração**. Dissertação (Mestrado em Engenharia Elétrica), Universidade Federal de Itajubá - UNIFEI, 2018.
- COSTA, M. V. et al. A Novel Methodology for Determining the Voltage Sag Impact Factor. **Electric Power Systems Research**, v. 174, 2019.
- DUAN, R. C. et al. **Data mining & pattern recognition of voltage sag based on K-means clustering algorithm**. 2015 IEEE Power & Energy Society General Meeting. **Anais...**2015
- DUAN, X.; CHEN, K. **Research on the application of maximal tree method based on fuzzy clustering for Power Quality Evaluation**. 2014 China International Conference on

- Electricity Distribution (CICED). **Anais...**2014
- DUGAN, R. C. et al. **Electrical Power Systems Quality**. 3rd. ed. NY: McGraw-hill, 2012.
- FERREIRA, D. F. **Estatística Multivariada**. Lavras, MG: Editora UFLA, 2011.
- FILHO, J. L. et al. **Methods for voltage sag source location by Cluster Algorithm and Decision Rule Labeling with a Comparative Approach of K-means and DBSCAN Clustering Algorithms**. 5th International Conference on Smart and Sustainable Technologies (SpliTech). **Anais...Split**: 2020
- FILHO, J. M. et al. A PCA-based approach for substation clustering for voltage sag studies in the Brazilian new energy context. **Electric Power Systems Research**, v. 136, 2016.
- FLORENCIAS-OLIVEROS, O. et al. **Cluster analysis for Power Quality monitoring**. 2017 11th IEEE International Conference on Compatibility, Power Electronics and Power Engineering (CPE-POWERENG). **Anais...**2017
- GARCÍA-SÁNCHEZ, T. et al. Statistical and Clustering Analysis for Disturbances: a Case Study of Voltage Dips in Wind Farms. **IEEE Transactions on Power Delivery**, v. 31, n. 6, p. 2530–2537, 2016.
- GARCIA, C. P.; ALVES, M. F. **Dimensionamento de Compensadores Série Considerando a Frequência de Ocorrência dos Afundamentos**. IX CBQEE - Conferência Brasileira sobre Qualidade da Energia Elétrica. **Anais...**Cuiabá: 2011
- HAIR, J. F. et al. **Análise Multivariada de Dados**. 6. ed. Porto Alegre: Bookman, 2009.
- HARIYANTO, N.; ANGGORO, B.; NOEGROHO, R. **New Probabilistic Approach for Identification Event Severity Index Due To Short Circuit Fault**. IEEE International Conference on Electrical Engineering and Computer Science. **Anais...**Bali, Indonesia: 2014
- IEC 61000-2-8. **Electromagnetic Compatibility (EMC) - Part 2-8: Environment - Voltage dips and short interruptions on public electric power supply systems with statistical measurement results**, 2005.
- IEC 61000-4-11. **Electromagnetic Compatibility (EMC) - Part 4-11: Testing and measurement techniques - Voltage dips, short interruptions and voltage variations immunity tests**, 2020.
- IEEE P1564. **IEEE Guide for Voltage Sag Indices**, 2014.
- JASIŃSKI, M. et al. **Cluster analysis of long-term power quality data**. 2016 Electric Power Networks (EPNet). **Anais...**2016

- MAHELA, O. P.; SHAIK, A. G. **Recognition of power quality disturbances using S-transform and Fuzzy C-means clustering**. 2016 International Conference on Cogeneration, Small Power Plants and District Energy (ICUE). **Anais...**2016
- MINGOTI, S. A. **Análise de Dados Através de Métodos de Estatística Multivariada: Uma Abordagem Aplicada**. 2^a ed. Belo Horizonte: Editora UFMG, 2005.
- MONTGOMERY, D. C.; RUNGER, G. C. **Estatística Aplicada e Probabilidade: para engenheiros**. 4. ed. São Paulo: Editora LTC, 2009.
- NOTA TÉCNICA N° 0021. **Resultados da Audiência Pública n° 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**. Brasília: [s.n.].
- NRS 048-2. **Electricity Supply - Quality of Supply Part 2: Voltage characteristics, compatibility levels, limits and assessment methods**, 2003.
- PAN, D. et al. **Recursive clustering K-nearest neighbors algorithm and the application in the classification of power quality disturbances**. 2017 IEEE Conference on Energy Internet and Energy System Integration (EI2). **Anais...**2017
- SEERA, M. et al. Power Quality Analysis Using a Hybrid Model of the Fuzzy Min–Max Neural Network and Clustering Tree. **IEEE Transactions on Neural Networks and Learning Systems**, v. 27, n. 12, p. 2760–2767, 2016.
- SHELDON, M. R. **Introduction to Probability and Statistics for Engineers and Scientists**. Third ed. [s.l.] Elsevier Academic Press, 2004.
- SOUZA, P. V. G. **Formação de Conjuntos de Sistemas de Distribuição quanto aos Fundamentos de Tensão**. Dissertação (Mestrado em Engenharia Elétrica), Universidade Federal de Itajubá - UNIFEI, 2016.
- TANURE, J. E. **Proposta de Procedimentos e Metodologia para Estabelecimento de Metas de Qualidade (DEC e FEC) para Concessionárias de Distribuição de Energia Elétrica Através de Análise Comparativa**. Tese (Doutorado em Engenharia Elétrica), Universidade de São Paulo, 2004.

APÊNDICE A – Tabelas de Distância e Heterogeneidade

A seguir são apresentadas as tabelas ordenadas de distâncias e heterogeneidades tomando como referência cada um dos sistemas de distribuição considerados no estudo de caso.

| SD | Distância | Heterogeneidade |
|-----------|-------------|-----------------|
| 1 | 0,00 | 0,00% |
| 7 | 1,52 | 25,33% |
| 14 | 2,49 | 41,51% |
| 19 | 2,57 | 42,91% |
| 11 | 2,97 | 49,44% |
| 6 | 3,11 | 51,79% |
| 13 | 3,14 | 52,38% |
| 17 | 3,34 | 55,60% |
| 15 | 3,43 | 57,12% |
| 18 | 3,43 | 57,19% |
| 9 | 3,54 | 59,02% |
| 12 | 3,67 | 61,10% |
| 16 | 3,67 | 61,21% |
| 8 | 3,89 | 64,88% |
| 3 | 4,00 | 66,68% |
| 5 | 4,66 | 77,65% |
| 2 | 4,78 | 79,64% |
| 10 | 4,92 | 81,98% |
| 4 | 6,06 | 100,96% |

| SD | Distância | Heterogeneidade |
|-----------|-------------|-----------------|
| 2 | 0,00 | 0,00% |
| 9 | 1,31 | 21,79% |
| 8 | 1,80 | 29,96% |
| 5 | 1,87 | 31,13% |
| 4 | 1,88 | 31,34% |
| 17 | 2,13 | 35,47% |
| 3 | 2,13 | 35,52% |
| 13 | 2,29 | 38,13% |
| 11 | 2,43 | 40,56% |
| 14 | 2,66 | 44,30% |
| 15 | 2,66 | 44,39% |
| 6 | 3,09 | 51,48% |
| 18 | 3,14 | 52,29% |
| 19 | 3,23 | 53,86% |
| 16 | 3,32 | 55,28% |
| 7 | 3,38 | 56,30% |
| 12 | 3,43 | 57,08% |
| 10 | 4,48 | 74,67% |
| 1 | 4,78 | 79,64% |

| SD | Distância | Heterogeneidade |
|-----------|-------------|-----------------|
| 3 | 0,00 | 0,00% |
| 8 | 0,60 | 9,96% |
| 13 | 0,87 | 14,57% |
| 11 | 1,04 | 17,29% |
| 18 | 1,14 | 19,06% |
| 6 | 1,22 | 20,26% |
| 15 | 1,34 | 22,33% |
| 12 | 1,36 | 22,62% |
| 14 | 1,54 | 25,62% |
| 9 | 1,56 | 26,03% |
| 19 | 1,63 | 27,15% |
| 2 | 2,13 | 35,52% |
| 16 | 2,34 | 38,97% |
| 7 | 2,84 | 47,36% |
| 17 | 2,84 | 47,37% |
| 5 | 3,01 | 50,14% |
| 4 | 3,91 | 65,22% |
| 1 | 4,00 | 66,68% |
| 10 | 4,43 | 73,83% |

| SD | Distância | Heterogeneidade |
|-----------|-------------|-----------------|
| 4 | 0,00 | 0,00% |
| 2 | 1,88 | 31,34% |
| 5 | 1,94 | 32,33% |
| 17 | 2,83 | 47,17% |
| 9 | 2,85 | 47,56% |
| 8 | 3,53 | 58,87% |
| 3 | 3,91 | 65,22% |
| 13 | 4,02 | 66,97% |
| 11 | 4,16 | 69,25% |
| 15 | 4,18 | 69,61% |
| 14 | 4,30 | 71,67% |
| 16 | 4,60 | 76,63% |
| 7 | 4,62 | 77,02% |
| 6 | 4,82 | 80,27% |
| 18 | 4,85 | 80,78% |
| 19 | 4,89 | 81,52% |
| 12 | 5,13 | 85,57% |
| 10 | 5,31 | 88,46% |
| 1 | 6,06 | 100,96% |

| SD | Distância | Heterogeneidade |
|----|-----------|-----------------|
| 5 | 0,00 | 0,00% |
| 17 | 1,50 | 24,99% |
| 9 | 1,83 | 30,56% |
| 2 | 1,87 | 31,13% |
| 4 | 1,94 | 32,33% |
| 8 | 2,57 | 42,91% |
| 15 | 2,79 | 46,49% |
| 13 | 2,88 | 47,93% |
| 11 | 2,99 | 49,80% |
| 3 | 3,01 | 50,14% |
| 14 | 3,06 | 51,03% |
| 7 | 3,16 | 52,72% |
| 16 | 3,19 | 53,23% |
| 19 | 3,54 | 58,97% |
| 6 | 3,57 | 59,47% |
| 18 | 3,59 | 59,91% |
| 12 | 3,89 | 64,89% |
| 10 | 4,62 | 77,06% |
| 1 | 4,66 | 77,65% |

| SD | Distância | Heterogeneidade |
|----|-----------|-----------------|
| 6 | 0,00 | 0,00% |
| 18 | 0,37 | 6,18% |
| 19 | 0,55 | 9,19% |
| 12 | 0,68 | 11,36% |
| 11 | 0,79 | 13,12% |
| 13 | 0,85 | 14,18% |
| 14 | 1,07 | 17,80% |
| 3 | 1,22 | 20,26% |
| 15 | 1,29 | 21,46% |
| 8 | 1,47 | 24,51% |
| 9 | 2,07 | 34,48% |
| 16 | 2,23 | 37,23% |
| 7 | 2,25 | 37,46% |
| 17 | 3,02 | 50,26% |
| 2 | 3,09 | 51,48% |
| 1 | 3,11 | 51,79% |
| 5 | 3,57 | 59,47% |
| 10 | 4,57 | 76,14% |
| 4 | 4,82 | 80,27% |

| SD | Distância | Heterogeneidade |
|----|-----------|-----------------|
| 7 | 0,00 | 0,00% |
| 14 | 1,39 | 23,14% |
| 1 | 1,52 | 25,33% |
| 19 | 1,77 | 29,44% |
| 11 | 1,85 | 30,88% |
| 17 | 1,93 | 32,22% |
| 13 | 1,98 | 33,05% |
| 9 | 2,12 | 35,32% |
| 15 | 2,21 | 36,89% |
| 6 | 2,25 | 37,46% |
| 18 | 2,54 | 42,34% |
| 16 | 2,60 | 43,28% |
| 8 | 2,61 | 43,50% |
| 3 | 2,84 | 47,36% |
| 12 | 2,86 | 47,67% |
| 5 | 3,16 | 52,72% |
| 2 | 3,38 | 56,30% |
| 10 | 4,18 | 69,63% |
| 4 | 4,62 | 77,02% |

| SD | Distância | Heterogeneidade |
|----|-----------|-----------------|
| 8 | 0,00 | 0,00% |
| 3 | 0,60 | 9,96% |
| 13 | 0,87 | 14,49% |
| 11 | 1,04 | 17,32% |
| 15 | 1,08 | 17,97% |
| 9 | 1,19 | 19,80% |
| 14 | 1,42 | 23,69% |
| 18 | 1,46 | 24,32% |
| 6 | 1,47 | 24,51% |
| 12 | 1,73 | 28,78% |
| 19 | 1,75 | 29,21% |
| 2 | 1,80 | 29,96% |
| 16 | 1,98 | 33,00% |
| 17 | 2,46 | 40,99% |
| 5 | 2,57 | 42,91% |
| 7 | 2,61 | 43,50% |
| 4 | 3,53 | 58,87% |
| 1 | 3,89 | 64,88% |
| 10 | 3,96 | 65,94% |

| SD | Distância | Heterogeneidade |
|----|-----------|-----------------|
| 9 | 0,00 | 0,00% |
| 8 | 1,19 | 19,80% |
| 13 | 1,24 | 20,68% |
| 2 | 1,31 | 21,79% |
| 11 | 1,34 | 22,37% |
| 17 | 1,36 | 22,64% |
| 14 | 1,46 | 24,37% |
| 3 | 1,56 | 26,03% |
| 15 | 1,73 | 28,84% |
| 5 | 1,83 | 30,56% |
| 19 | 2,06 | 34,41% |
| 6 | 2,07 | 34,48% |
| 7 | 2,12 | 35,32% |
| 18 | 2,20 | 36,68% |
| 16 | 2,52 | 41,99% |
| 12 | 2,55 | 42,45% |
| 4 | 2,85 | 47,56% |
| 1 | 3,54 | 59,02% |
| 10 | 4,17 | 69,48% |

| SD | Distância | Heterogeneidade |
|----|-----------|-----------------|
| 10 | 0,00 | 0,00% |
| 16 | 2,83 | 47,19% |
| 15 | 3,66 | 60,92% |
| 8 | 3,96 | 65,94% |
| 14 | 4,01 | 66,78% |
| 9 | 4,17 | 69,48% |
| 7 | 4,18 | 69,63% |
| 11 | 4,21 | 70,25% |
| 13 | 4,23 | 70,47% |
| 3 | 4,43 | 73,83% |
| 2 | 4,48 | 74,67% |
| 17 | 4,50 | 74,93% |
| 19 | 4,54 | 75,67% |
| 6 | 4,57 | 76,14% |
| 5 | 4,62 | 77,06% |
| 18 | 4,66 | 77,66% |
| 12 | 4,84 | 80,74% |
| 1 | 4,92 | 81,98% |
| 4 | 5,31 | 88,46% |

| SD | Distância | Heterogeneidade |
|----|-----------|-----------------|
| 11 | 0,00 | 0,00% |
| 13 | 0,19 | 3,20% |
| 14 | 0,52 | 8,75% |
| 6 | 0,79 | 13,12% |
| 19 | 0,83 | 13,80% |
| 18 | 1,03 | 17,18% |
| 3 | 1,04 | 17,29% |
| 8 | 1,04 | 17,32% |
| 15 | 1,16 | 19,35% |
| 9 | 1,34 | 22,37% |
| 12 | 1,39 | 23,18% |
| 7 | 1,85 | 30,88% |
| 16 | 2,14 | 35,64% |
| 17 | 2,35 | 39,09% |
| 2 | 2,43 | 40,56% |
| 1 | 2,97 | 49,44% |
| 5 | 2,99 | 49,80% |
| 4 | 4,16 | 69,25% |
| 10 | 4,21 | 70,25% |

| SD | Distância | Heterogeneidade |
|----|-----------|-----------------|
| 12 | 0,00 | 0,00% |
| 18 | 0,37 | 6,14% |
| 6 | 0,68 | 11,36% |
| 19 | 1,15 | 19,12% |
| 3 | 1,36 | 22,62% |
| 11 | 1,39 | 23,18% |
| 13 | 1,39 | 23,18% |
| 15 | 1,50 | 24,95% |
| 14 | 1,72 | 28,73% |
| 8 | 1,73 | 28,78% |
| 16 | 2,37 | 39,44% |
| 9 | 2,55 | 42,45% |
| 7 | 2,86 | 47,67% |
| 2 | 3,43 | 57,08% |
| 17 | 3,55 | 59,08% |
| 1 | 3,67 | 61,10% |
| 5 | 3,89 | 64,89% |
| 10 | 4,84 | 80,74% |
| 4 | 5,13 | 85,57% |

| SD | Distância | Heterogeneidade |
|-----------|------------------|------------------------|
| 13 | 0,00 | 0,00% |
| 11 | 0,19 | 3,20% |
| 14 | 0,69 | 11,47% |
| 6 | 0,85 | 14,18% |
| 8 | 0,87 | 14,49% |
| 3 | 0,87 | 14,57% |
| 19 | 0,97 | 16,20% |
| 18 | 1,03 | 17,24% |
| 15 | 1,11 | 18,58% |
| 9 | 1,24 | 20,68% |
| 12 | 1,39 | 23,18% |
| 7 | 1,98 | 33,05% |
| 16 | 2,13 | 35,52% |
| 2 | 2,29 | 38,13% |
| 17 | 2,32 | 38,63% |
| 5 | 2,88 | 47,93% |
| 1 | 3,14 | 52,38% |
| 4 | 4,02 | 66,97% |
| 10 | 4,23 | 70,47% |

| SD | Distância | Heterogeneidade |
|-----------|------------------|------------------------|
| 14 | 0,00 | 0,00% |
| 11 | 0,52 | 8,75% |
| 13 | 0,69 | 11,47% |
| 19 | 0,83 | 13,84% |
| 6 | 1,07 | 17,80% |
| 15 | 1,30 | 21,68% |
| 18 | 1,38 | 22,93% |
| 7 | 1,39 | 23,14% |
| 8 | 1,42 | 23,69% |
| 9 | 1,46 | 24,37% |
| 3 | 1,54 | 25,62% |
| 12 | 1,72 | 28,73% |
| 16 | 2,08 | 34,73% |
| 17 | 2,22 | 36,96% |
| 1 | 2,49 | 41,51% |
| 2 | 2,66 | 44,30% |
| 5 | 3,06 | 51,03% |
| 10 | 4,01 | 66,78% |
| 4 | 4,30 | 71,67% |

| SD | Distância | Heterogeneidade |
|-----------|------------------|------------------------|
| 15 | 0,00 | 0,00% |
| 8 | 1,08 | 17,97% |
| 16 | 1,11 | 18,53% |
| 13 | 1,11 | 18,58% |
| 11 | 1,16 | 19,35% |
| 18 | 1,26 | 20,93% |
| 6 | 1,29 | 21,46% |
| 14 | 1,30 | 21,68% |
| 3 | 1,34 | 22,33% |
| 19 | 1,41 | 23,53% |
| 12 | 1,50 | 24,95% |
| 9 | 1,73 | 28,84% |
| 7 | 2,21 | 36,89% |
| 17 | 2,58 | 43,05% |
| 2 | 2,66 | 44,39% |
| 5 | 2,79 | 46,49% |
| 1 | 3,43 | 57,12% |
| 10 | 3,66 | 60,92% |
| 4 | 4,18 | 69,61% |

| SD | Distância | Heterogeneidade |
|-----------|------------------|------------------------|
| 16 | 0,00 | 0,00% |
| 15 | 1,11 | 18,53% |
| 8 | 1,98 | 33,00% |
| 14 | 2,08 | 34,73% |
| 13 | 2,13 | 35,52% |
| 11 | 2,14 | 35,64% |
| 18 | 2,21 | 36,85% |
| 6 | 2,23 | 37,23% |
| 19 | 2,26 | 37,73% |
| 3 | 2,34 | 38,97% |
| 12 | 2,37 | 39,44% |
| 9 | 2,52 | 41,99% |
| 7 | 2,60 | 43,28% |
| 10 | 2,83 | 47,19% |
| 17 | 3,10 | 51,73% |
| 5 | 3,19 | 53,23% |
| 2 | 3,32 | 55,28% |
| 1 | 3,67 | 61,21% |
| 4 | 4,60 | 76,63% |

| SD | Distância | Heterogeneidade |
|-----------|------------------|------------------------|
| 17 | 0,00 | 0,00% |
| 9 | 1,36 | 22,64% |
| 5 | 1,50 | 24,99% |
| 7 | 1,93 | 32,22% |
| 2 | 2,13 | 35,47% |
| 14 | 2,22 | 36,96% |
| 13 | 2,32 | 38,63% |
| 11 | 2,35 | 39,09% |
| 8 | 2,46 | 40,99% |
| 15 | 2,58 | 43,05% |
| 19 | 2,79 | 46,57% |
| 4 | 2,83 | 47,17% |
| 3 | 2,84 | 47,37% |
| 6 | 3,02 | 50,26% |
| 16 | 3,10 | 51,73% |
| 18 | 3,19 | 53,15% |
| 1 | 3,34 | 55,60% |
| 12 | 3,55 | 59,08% |
| 10 | 4,50 | 74,93% |

| SD | Distância | Heterogeneidade |
|-----------|------------------|------------------------|
| 18 | 0,00 | 0,00% |
| 12 | 0,37 | 6,14% |
| 6 | 0,37 | 6,18% |
| 19 | 0,87 | 14,47% |
| 11 | 1,03 | 17,18% |
| 13 | 1,03 | 17,24% |
| 3 | 1,14 | 19,06% |
| 15 | 1,26 | 20,93% |
| 14 | 1,38 | 22,93% |
| 8 | 1,46 | 24,32% |
| 9 | 2,20 | 36,68% |
| 16 | 2,21 | 36,85% |
| 7 | 2,54 | 42,34% |
| 2 | 3,14 | 52,29% |
| 17 | 3,19 | 53,15% |
| 1 | 3,43 | 57,19% |
| 5 | 3,59 | 59,91% |
| 10 | 4,66 | 77,66% |
| 4 | 4,85 | 80,78% |

| SD | Distância | Heterogeneidade |
|-----------|------------------|------------------------|
| 19 | 0,00 | 0,00% |
| 6 | 0,55 | 9,19% |
| 11 | 0,83 | 13,80% |
| 14 | 0,83 | 13,84% |
| 18 | 0,87 | 14,47% |
| 13 | 0,97 | 16,20% |
| 12 | 1,15 | 19,12% |
| 15 | 1,41 | 23,53% |
| 3 | 1,63 | 27,15% |
| 8 | 1,75 | 29,21% |
| 7 | 1,77 | 29,44% |
| 9 | 2,06 | 34,41% |
| 16 | 2,26 | 37,73% |
| 1 | 2,57 | 42,91% |
| 17 | 2,79 | 46,57% |
| 2 | 3,23 | 53,86% |
| 5 | 3,54 | 58,97% |
| 10 | 4,54 | 75,67% |
| 4 | 4,89 | 81,52% |

APÊNDICE B – Tabela da Frequência de AMTs

A seguir está apresentada a tabela com a frequência média mensal de AMTs dos sistemas de distribuição considerados no estudo de caso, distribuída nas regiões de sensibilidade de A à G.

| SD | Frequência de AMTs nas regiões | | | | | | |
|----|--------------------------------|------|------|------|------|------|------|
| | A | B | C | D | E | F | G |
| 1 | 11,80 | 1,60 | 0,30 | 2,20 | 0,50 | 0,90 | 1,30 |
| 2 | 1,83 | 0,50 | 0,25 | 0,92 | 0,08 | 0,25 | 0,25 |
| 3 | 3,25 | 0,08 | 0,00 | 0,50 | 0,00 | 0,17 | 1,25 |
| 4 | 2,33 | 0,17 | 0,08 | 0,50 | 0,67 | 1,08 | 0,50 |
| 5 | 4,73 | 0,91 | 0,36 | 4,18 | 0,73 | 2,82 | 1,09 |
| 6 | 7,08 | 0,67 | 0,33 | 3,58 | 0,33 | 1,00 | 0,92 |
| 7 | 5,09 | 1,55 | 0,00 | 1,73 | 0,45 | 0,82 | 1,45 |
| 8 | 3,08 | 0,00 | 0,00 | 0,50 | 0,25 | 0,50 | 0,58 |
| 9 | 3,58 | 0,17 | 0,08 | 1,33 | 0,33 | 1,50 | 1,25 |
| 10 | 7,67 | 0,50 | 0,00 | 0,75 | 0,58 | 2,25 | 3,42 |
| 11 | 2,08 | 0,75 | 0,17 | 0,83 | 0,67 | 0,92 | 0,83 |
| 12 | 6,00 | 0,75 | 0,42 | 2,25 | 0,67 | 1,58 | 0,67 |
| 13 | 6,67 | 0,67 | 0,33 | 1,92 | 1,75 | 0,42 | 1,42 |
| 14 | 7,60 | 0,40 | 0,00 | 1,10 | 0,30 | 0,70 | 2,70 |
| 15 | 8,55 | 1,36 | 0,09 | 2,73 | 0,18 | 0,18 | 1,45 |
| 16 | 9,25 | 2,25 | 0,17 | 1,67 | 1,42 | 1,25 | 0,75 |
| 17 | 7,09 | 1,18 | 0,09 | 1,55 | 0,64 | 1,73 | 1,18 |
| 18 | 6,13 | 0,75 | 0,13 | 4,25 | 0,38 | 0,63 | 0,88 |
| 19 | 8,89 | 0,33 | 0,22 | 1,33 | 1,11 | 0,56 | 1,00 |