

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
ESCOLA DE ENGENHARIA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

ARY HENRIQUE LAUTENSCHLEGER

**ANÁLISE DA OPERAÇÃO DE
SISTEMAS DE DISTRIBUIÇÃO
CONSIDERANDO AS INCERTEZAS DA
CARGA E DA GERAÇÃO DISTRIBUÍDA**

Porto Alegre
2018

ARY HENRIQUE LAUTENSCHLEGER

**ANÁLISE DA OPERAÇÃO DE
SISTEMAS DE DISTRIBUIÇÃO
CONSIDERANDO AS INCERTEZAS DA
CARGA E DA GERAÇÃO DISTRIBUÍDA**

Dissertação de mestrado apresentada ao Programa de Pós-Graduação em Engenharia Elétrica da Universidade Federal do Rio Grande do Sul como parte dos requisitos para a obtenção do título de Mestre em Engenharia Elétrica.
Área de concentração: Energia

ORIENTADOR: Prof. Dr. Sérgio Luís Haffner

Porto Alegre
2018

ARY HENRIQUE LAUTENSCHLEGER

**ANÁLISE DA OPERAÇÃO DE
SISTEMAS DE DISTRIBUIÇÃO
CONSIDERANDO AS INCERTEZAS DA
CARGA E DA GERAÇÃO DISTRIBUÍDA**

Esta dissertação foi julgada adequada para a obtenção do título de Mestre em Engenharia Elétrica e aprovada em sua forma final pelo Orientador e pela Banca Examinadora.

Orientador: _____
Prof. Dr. Sérgio Luís Haffner, UFRGS
Doutor pela Universidade Estadual de Campinas – Campinas,
Brasil

Banca Examinadora:

Prof. Dr. Antonio Padilha Feltrim, UNESP–Ilha Solteira
Doutor pela Universidade Estadual de Campinas – Campinas, Brasil

Prof. Dr. Alexandre Sanfelici Bazanella, UFRGS
Doutor pela Universidade Federal de Santa Catarina – Florianópolis, Brasil

Prof. Dr. Roberto Chouhy Leborgne, UFRGS
Doutor pela Chalmers University of Technology – Göteborg, Suécia

Coordenador do PPGEE: _____
Prof. Dr. Valner João Brusamarello

Porto Alegre, maio de 2018.

DEDICATÓRIA

Aos meus pais: pelo exemplo, incentivo e, principalmente, por todo o seu sacrifício e superação que me proporcionou condições para lutar pelos meus sonhos.

AGRADECIMENTOS

À minha família, que é a minha fortaleza e fonte de inspiração, especialmente à minha irmã Karin e à Carol, minha amada esposa, por todo incentivo, paciência, apoio e ajuda durante essa jornada, sem a qual o desenvolvimento de várias ideias desta dissertação não teriam sido possíveis.

Ao Prof. Sérgio Luís Haffner, pela confiança, paciência, exigência e exemplar orientação.

Ao Programa de Pós-Graduação em Engenharia Elétrica, PPGEE, pela oportunidade de realização de trabalhos em minha área de interesse.

Aos colegas do GMASP e LASEP, especialmente Bibiana Petry Ferraz, José David Anzanello Junior e Charles Lorenzini pelo companheirismo nos trabalhos e estudos, mas também pelas conversas e momentos de descontração.

À CAPES pela provisão da bolsa de mestrado quando precisei.

À Força Aérea Brasileira e ao Destacamento de Infraestrutura da Aeronáutica de Canoas, onde encontrei colegas de profissão com os quais tenho aprendido imensamente.

À todos os amigos pelos momentos de alegria.

RESUMO

Neste trabalho é apresentado um método probabilístico para avaliação do desempenho de redes de distribuição considerando incertezas na demanda das cargas e na potência gerada por sistemas distribuídos intermitentes. Os consumidores são divididos em agrupamentos por classe e faixa de consumo e a modelagem da demanda horária dos consumidores de cada agrupamento é realizada por uma lei de distribuição acumulada de probabilidade (CDF) adequada. A geração distribuída é contemplada pela consideração de fonte solar fotovoltaica. O procedimento de simulação do Método de Monte Carlo é empregado e a técnica da *Joint Normal Transform* é utilizada na geração de números aleatórios correlacionados, empregados na amostragem da demanda dos consumidores e da energia produzida pelos sistemas de geração distribuídos. O método proposto foi aplicado ao conhecido sistema de 13 barras do IEEE e os resultados dos indicadores de perdas na operação bem como indicadores de violação de tensão crítica e precária obtidos com o modelo probabilístico são comparados aos obtidos com o modelo determinístico convencional. É demonstrado que nem sempre a média é uma descrição suficiente para o comportamento dos componentes de redes de distribuição e que é mais adequado utilizar uma representação com intervalos de confiança para as grandezas de interesse.

Palavras-chave: Redes de Distribuição de Energia, Fluxo de Potência, Simulação Estocástica, Indicadores de Qualidade, Análise da Operação, Geração Distribuída.

ABSTRACT

This work presents a probabilistic method for performance evaluation of distribution networks considering uncertainties in load demand and power generated by intermittent distributed systems. Consumers are divided into clusters by class and consumption range, so the modeling for the hourly demand of the consumers on each cluster is performed by a suitable cumulative probability distribution (CDF). Distributed generation is considered by means of solar photovoltaic sources. The Monte Carlo Simulation (MCS) Method is employed and the Joint Normal Transform technique is applied for correlated random numbers generation, used to sample consumer demand and the energy generated by distributed generation systems. The proposed method was applied in the well-known IEEE 13 node test feeder and the results of the operation losses as well as voltage violation indices obtained by the probabilistic model are compared to those obtained with the conventional deterministic model. It is shown that the mean is not always a sufficient description for the behavior of distribution network components and that it is more appropriate to use confidence intervals for the quantities of interest.

Keywords: Power Distribution Networks, Power Flow, Stochastic Simulation, Power Quality Metrics, Operation Assessment, Distributed Generation.

SUMÁRIO

LISTA DE ILUSTRAÇÕES	9
LISTA DE TABELAS	11
LISTA DE ABREVIATURAS	12
LISTA DE SÍMBOLOS	13
1 INTRODUÇÃO	14
1.1 Contextualização e motivação	14
1.2 Objetivo	18
1.3 Estrutura	18
2 INCERTEZA EM SISTEMAS DE POTÊNCIA	19
2.1 Incerteza dos Eventos	19
2.2 Variáveis Aleatórias (VAs)	20
2.2.1 Valor Esperado, Quantis e Momentos Matemáticos de VAs	20
2.2.2 Relações de Dependência entre Variáveis Aleatórias	21
2.2.3 Distribuições de Probabilidade Notáveis	24
2.2.4 Distribuições Multidimensionais	24
2.2.5 Cópulas	25
2.3 Inferência Estatística	27
2.3.1 Estatísticas Amostrais	28
2.3.2 Lei dos Grandes Números e Teorema do Limite Central	28
2.3.3 Intervalos de Confiança	29
2.3.4 Testes de Hipótese e de Adequação de Ajuste (<i>Goodness of Fit</i>)	29
2.4 Aspectos de Modelagem e Simulação Estocástica	30
2.4.1 Princípio de Amostragem	30
2.4.2 Critério de Parada	31
2.4.3 Construção de Estrutura de Dependência Multidimensional para Variáveis Aleatórias	32
2.4.4 <i>Stochastic Bounds Methodology</i> : Estruturas de Dependência Extrema entre VAs	33
2.4.5 Técnicas de Agrupamento de Variáveis Aleatórias	35
3 ANÁLISE DA OPERAÇÃO DE SISTEMAS DE DISTRIBUIÇÃO DE ENERGIA ELÉTRICA	37
3.1 Análise determinística	37
3.2 Análise probabilística	39

3.2.1	Principais trabalhos investigados	42
3.3	Modelo de resposta das cargas à variação de tensão	48
4	MÉTODO PROPOSTO PARA ANÁLISE PROBABILÍSTICA DAS IN- CERTEZAS	50
4.1	Modelagem da distribuição marginal de probabilidade para os compo- nentes do sistema	52
4.2	Modelagem da estrutura estocástica multidimensional de dependência .	56
4.3	Criação de cenários de demanda e geração distribuída de energia	59
4.3.1	Representação do acoplamento temporal na curva horária da demanda . .	59
4.4	Cálculo do fluxo de potência e dos parâmetros de operação da rede . . .	61
4.5	Aspectos complementares do método	63
5	APLICAÇÃO DO MÉTODO E RESULTADOS OBTIDOS	64
5.1	Rede de Testes e Configurações de Simulação	64
5.1.1	Adaptações para representação da carga	65
5.1.2	Adaptações para inclusão da GD	71
5.2	Simulações e Análise Probabilística	73
5.2.1	Caracterização Estocástica dos Componentes do Sistema e Estrutura de Dependência	73
5.2.2	Amostragem e Composição de Cenários de Demanda e GDFV	81
5.2.3	Convergência das simulações	86
5.2.4	Indicadores Probabilísticos de Operação	89
6	CONCLUSÃO	100
6.1	Trabalhos Futuros	102
	REFERÊNCIAS	104

LISTA DE ILUSTRAÇÕES

Figura 1:	Pesquisas sobre fluxo de potência probabilístico classificadas por método de avaliação.	17
Figura 2:	Diagrama de caixa e linha para uma distribuição normal.	22
Figura 3:	Diagrama de caixa e linha para distribuição lognormal.	22
Figura 4:	Dispersão entre as demandas de energia elétrica de dois consumidores com diferentes relações de dependência.	24
Figura 5:	Exemplo de distribuição conjunta para duas variáveis aleatórias.	25
Figura 6:	Diagrama de dispersão para cópulas independente e de Fréchet.	26
Figura 7:	Exemplo de famílias de cópulas.	26
Figura 8:	Diagrama de dispersão para diferentes correlações das cópulas de banda diagonal.	27
Figura 9:	Procedimento de amostragem de uma VA.	31
Figura 10:	Convergência de uma simulação estocástica.	32
Figura 11:	Ilustração das etapas da JNT para $\rho = 0,75$	33
Figura 12:	Ilustração das etapas da JNT para $\rho = 0,45$	33
Figura 13:	Obtenção de patamares de demanda agregada.	40
Figura 14:	Curva de carga determinística versus probabilística.	41
Figura 15:	Curvas representativas em p.u. para dias úteis de usuários residenciais com consumo entre 220 e 500 kWh.	43
Figura 16:	Curva de frequência acumulada da demanda relativa das 20h às 21h para um cliente residencial em BT com consumo mensal entre 200 e 400 kWh.	44
Figura 17:	Procedimento de amostragem utilizado por Guimarães (2008).	45
Figura 18:	Histograma de frequência relativa das tensões de atendimento.	45
Figura 19:	Diagrama esquemático da análise do sistema.	47
Figura 20:	Relação de conversão de potência de um gerador eólico e PDF da potência gerada.	48
Figura 21:	Fluxograma do Método Proposto.	51
Figura 22:	Divisão de agrupamentos proposta.	54
Figura 23:	Exemplo de curva diária de média e desvio padrão para um agrupamento de consumidores.	55
Figura 24:	Obtenção da CDF da potência horária a partir da lei de distribuição e das curvas típicas diárias.	56
Figura 25:	Resultado da JNT para os agrupamentos.	58
Figura 26:	Procedimento de amostragem das potências dos componentes da rede.	60
Figura 27:	Procedimento de agregação das potências dos componentes nos nós da rede.	60

Figura 28:	Sistema teste IEEE.	64
Figura 29:	Curva de média (μ) e desvio padrão (σ) da demanda dos agrupamentos de consumidores.	70
Figura 30:	Localização da GDFV no sistema IEEE de 13 barras.	72
Figura 31:	Curva de média e desvio padrão horários normalizados da geração dos agrupamentos de GDFV [p.u.].	73
Figura 32:	PDF e CDF da demanda às 9h dos agrupamentos de carga na Barra 692.	74
Figura 33:	PDF e CDF da demanda às 14h dos agrupamentos de carga na Barra 692.	75
Figura 34:	PDF e CDF da demanda às 21h dos agrupamentos de carga na Barra 692.	75
Figura 35:	PDF e CDF em p.u. da geração fotovoltaica por horário.	77
Figura 36:	Dispersão entre U_i gerados para amostragem da potência dos componentes dos GP_i às 9h e 14h.	80
Figura 37:	Demanda agregada na Barra 692 em três horários para diferentes configurações de GDFV.	82
Figura 38:	GDFV na barra 692 para diferentes níveis de penetração de GD e $N=5000$ amostras.	83
Figura 39:	Comportamento probabilístico da demanda agregada no alimentador com ausência de GDFV resultante de $N = 5000$ amostras.	84
Figura 40:	Comportamento probabilístico da demanda agregada no alimentador com 50% de penetração de GDFV resultante de $N = 5000$ amostras.	85
Figura 41:	Comportamento probabilístico da demanda agregada no alimentador com 100% de penetração de GDFV resultante de $N = 5000$ amostras.	85
Figura 42:	Comportamento probabilístico da GDFV no sistema com 50% de penetração para $N = 5000$ amostras.	87
Figura 43:	Comportamento probabilístico da GDFV no sistema com 100% de penetração para $N = 5000$ amostras.	88
Figura 44:	Processo de convergência da demanda média na Fase A da Barra 675 às 21h.	89
Figura 45:	Processo de convergência da energia de perdas diárias no sistema.	90
Figura 46:	Processo de convergência do DRC e DRP na barra 675.	91
Figura 47:	Comportamento probabilístico da potência de perdas horária.	92
Figura 48:	PDF e CDF da perda diária para diferentes configurações de penetração de GDFV.	94
Figura 49:	Perfil de tensão na rede em p.u. para cenários determinísticos de carga na ausência de GDFV.	95
Figura 50:	PDF da tensão em todas as barras às 21h.	96
Figura 51:	PDF da tensão em todas as barras às 14h e às 18h para configurações de penetração de GDFV diversas.	97

LISTA DE TABELAS

Tabela 1:	Sumário dos principais parâmetros de algumas distribuições de probabilidade.	25
Tabela 2:	Valores representativos de carga em p.u. para clientes residenciais dos agrupamentos 1 e 2.	66
Tabela 3:	Valores representativos de carga em p.u. para clientes residenciais dos agrupamentos 3 e 4.	67
Tabela 4:	Valores representativos de carga em p.u. para clientes comerciais dos agrupamentos 5, 6 e 7.	68
Tabela 5:	Valores representativos de carga em p.u. para clientes industriais dos agrupamentos 8 e 9.	69
Tabela 6:	Quantidade definida de clientes de cada agrupamento (GP) em cada barra.	71
Tabela 7:	Fatores de carga e demanda típicos de clientes de redes de distribuição.	76
Tabela 8:	Valores de perda diária resultantes de uma análise determinística.	90
Tabela 9:	Valor médio, desvio padrão e IC da perda diária resultantes de uma análise probabilística.	93
Tabela 10:	Índices de violação de tensão para cenários de carga determinísticos e probabilísticos na ausência de GDFV.	96
Tabela 11:	Índices de violação de tensão para cenários de carga determinísticos e probabilísticos com 50% de penetração de GDFV.	98
Tabela 12:	Índices de violação de tensão para cenários de carga determinísticos e probabilísticos com 100% de penetração de GDFV.	98

LISTA DE ABREVIATURAS

CDF	<i>Cumulative Probability Density Function</i>
DRC	Duração Relativa de Transgressão de Tensão Crítica
DRP	Duração Relativa de Transgressão de Tensão Precária
GD	Geração Distribuída
GDFV	Geração Distribuída Solar Fotovoltaica
IEEE	<i>Institute of Electrical and Electronics Engineers</i>
JNT	<i>Joint-Normal Transform</i>
MEP	<i>Método de Estimação por Pontos</i>
PDF	<i>Probability Density Function</i>
PLF	<i>Power Load Flow</i>
PRODIST	Procedimentos de Distribuição de Energia Elétrica no Sistema Elétrico Nacional
p.u.	Por Unidade
SBM	<i>Stochastic Bounds Methodology</i>
SMC	Simulação de Monte Carlo
VA	Variável aleatória

LISTA DE SÍMBOLOS

$F'(x)$	Derivada da função $F(x)$
\mathbb{R}	Conjunto dos números Reais
\mathcal{C}	Conjunto de condições necessárias para a ocorrência de um determinado evento aleatório
\mathcal{A}	Conjunto de eventos aleatórios elementares que atendem às condições \mathcal{C}
Ω	Conjunto de todos os eventos aleatórios elementares
$P(\mathcal{A})$	Probabilidade de ocorrência de \mathcal{A}

1 INTRODUÇÃO

1.1 Contextualização e motivação

A infraestrutura das redes de distribuição deve atender instantaneamente à solicitação de demanda dos consumidores de energia elétrica, dentro dos padrões de qualidade e confiabilidade, ao longo de todo o tempo, mas especialmente nos horários de pico de consumo, quando o sistema é solicitado de modo mais intenso. Para atingir esse objetivo, é necessário investir em melhorias da rede, tanto na forma de manutenção como através de expansão. Projetos de melhorias da rede costumam ser motivados por violações excessivas dos níveis de tensão, carregamento elevado dos transformadores, fim da vida útil de equipamentos, indicadores de continuidade superiores aos limites normativos bem como troca preventiva e/ou corretiva de condutores de ramais de serviço (BUSATTO, 2015).

Uma vez que os recursos são limitados, frente às necessidades, é preciso estabelecer critérios para a priorização dos projetos executados. Os investimentos são determinados com base no plano de expansão da rede, que contempla obras mandatórias ao atendimento do mercado consumidor por um determinado período (BUSATTO, 2015). A definição da expansão do sistema também considera o desempenho do mesmo frente ao crescimento da demanda, requisitos regulatórios além de critérios estratégicos do planejamento.

No Brasil, o Módulo 2 dos Procedimentos de Distribuição de Energia Elétrica no Sistema Elétrico Nacional (PRODIST) regulamenta os aspectos da expansão dos sistemas elétricos (ANEEL, 2016). Além disso, estudos que apoiam o planejamento envolvem o diagnóstico do sistema existente, com a caracterização da carga e da rede, bem como a avaliação da qualidade do produto e das perdas técnicas. Desse modo, o disposto no Módulo 7 do PRODIST sobre a avaliação das perdas técnicas nos condutores e equipamentos bem como no Módulo 8 do PRODIST sobre os indicadores de qualidade de energia elétrica permite apurar a condição de atendimento dos clientes e identificar os elementos em estado crítico de operação (BUSATTO, 2015; ANEEL, 2015, 2018).

O estabelecimento de metas de expansão contemplando projetos que proporcionem melhoria nos indicadores é interessante tanto para os clientes como para as distribuidoras. Uma vez que há melhora na qualidade do produto e do serviço, existe impacto na redução dos custos de operação juntamente com a redução das compensações financeiras oriundas de índices inadequados. Por ter importante impacto financeiro, é desejável que os métodos envolvidos na obtenção dos indicadores de operação sejam capazes de representar da melhor maneira possível a realidade do sistema. Com isso, permite-se a tomada de decisões mais exatas acerca das melhorias prioritárias e necessárias na rede aumentando o retorno sobre o investimento bem como a satisfação dos clientes.

O Brasil deverá experimentar nos próximos anos um aumento no número de empreendimentos de mini e microgeração distribuída em função da possibilidade de implan-

tação de sistemas *on-grid* de geração distribuída (GD) após regulamentação dada pela Resolução Normativa nº 482 bem como pelo aumento da viabilidade financeira decorrente da redução dos custos associados a tais empreendimentos (ANEEL, 2012, 2017a,b). Assim, ao invés de apenas consumir energia, os clientes passam a desempenhar também o papel de produtores de energia e a ser tratados como prosumidores¹ (LAUTENSCHLEGER, 2013).

Em muitos casos, os empreendimentos de GD costumam ser baseados em fontes de energia primária intermitente, como solar e eólica, introduzindo uma incerteza operacional por ser uma geração não-despachável. A incorporação gradual de quantidade elevada de empreendimentos de GD tem contribuído para a transição de uma estrutura dita vertical das redes elétricas para uma estrutura dita horizontal dos sistemas de distribuição (RODRIGUES, 2008).

Visando avaliar o desempenho de sistemas elétricos, a análise do fluxo de potência permite simular o perfil de tensão nos nós, o fluxo de potência nas linhas e as perdas ao longo da rede de acordo com o perfil de uso das cargas (GUPTA; KUMAR, 2017). Usualmente, as cargas são modeladas por meio de curvas que representam patamares de demanda, os quais variam ao longo do tempo, porém são fixos em cada instante, o que caracteriza uma modelagem determinística. No entanto, a demanda varia continuamente devido a fatores temporais (como estação do ano, dia da semana ou hora do dia), econômicos, dentre outros (GUPTA; KUMAR, 2017). Além disso, diferentemente da transmissão onde a diversidade das cargas apresenta demanda horária agregada com baixa variância, na distribuição o desvio padrão associado à demanda dos componentes do sistema pode acarretar uma representação inadequada de alguns indicadores por um único valor ao invés de um histograma de frequência de ocorrência (PAREJA, 2009; GUIMARÃES, 2008).

Estudos demonstram que o desvio padrão da demanda de clientes da rede secundária de distribuição tem magnitude maior que a média em alguns casos, haja vista que a incerteza na demanda depende do horário na qual a mesma ocorre (PAREJA, 2009). Além disso, a crescente complexidade do sistema elétrico, associada ao acréscimo de fontes intermitentes de energia, aliada a mudanças nos padrões de consumo e ao fácil acesso a diversas tecnologias eletroeletrônicas, têm contribuído para a inclusão de importantes fontes de incerteza na modelagem de sistemas elétricos e introduz um novo aspecto a ser considerado no cálculo do fluxo de potência (MOHAMMADI; SHAYEGANI; ADAMI-NEJAD, 2013).

Análises de operação da rede de distribuição que apoiam a decisão de expansão e investimentos costumam ser realizadas com base em um número reduzido de possibilidades operacionais (JONES; CHOWDHURY, 2008; ALGARNI, 2009; ANEEL, 2016). No entanto, a natureza das cargas e da geração distribuída (GD) não-despachável, principalmente de origem eólica e solar, é incerta. Ao simular poucos cenários, muitas possibilidades operacionais são ignoradas, fazendo com que o resultado obtido possa se distanciar da realidade que o sistema irá enfrentar (PAPAEFTHYMIU, 2007; RODRIGUES, 2008).

Deste modo, os modelos determinísticos de análise tornam-se insuficientes para lidar com o novo paradigma de operação da rede, fazendo com que sejam necessárias abordagens estocásticas (XU; YAN, 2017; CHEN et al., 2018). Neste sentido, o emprego da simulação de Monte Carlo tem sido preferido frente aos métodos analíticos, pois a incerteza

¹ Consumidores com o papel adicional de auto fornecimento (proprietário) de geração de eletricidade e/ou armazenamento para uso privado, necessidades diárias, conforto e necessidades de negócios de Pequenas e Médias Empresas.

na saída de um gerador estocástico ou de uma carga é quantificada em termos numéricos pela análise estatística dos respectivos dados (PAPAEFTHYMIU, 2007; RODRIGUES, 2008).

Um aspecto importante na análise das redes de distribuição diz respeito aos padrões de comportamento entre os consumidores. É possível identificar grupos com alta correlação positiva, i.e., a demanda dos clientes pertencentes ao mesmo grupo apresenta característica estocástica que indica relação de dependência próxima da correlação perfeita. Também é possível observar uma relação de dependência entre os agrupamentos de consumidores em função de suas classes e/ou dos seus consumos médios mensais de energia. Dessa forma, a consideração de independência não é válida visto que existe uma estrutura multidimensional de interdependência estocástica entre as demandas dos consumidores. Essa relação é consequência de fatores externos comuns aos diferentes grupos de consumidores (tais como época do ano e temperatura) que irão induzir hábitos de consumo parecidos entre eles. Para a GD ocorre algo bastante similar, pois a relação de dependência é oriunda da influência que a fonte primária de energia insere nos empreendimentos de GD situados em regiões geográficas próximas e que utilizam a mesma tecnologia de geração (PAPAEFTHYMIU, 2007; RODRIGUES, 2008).

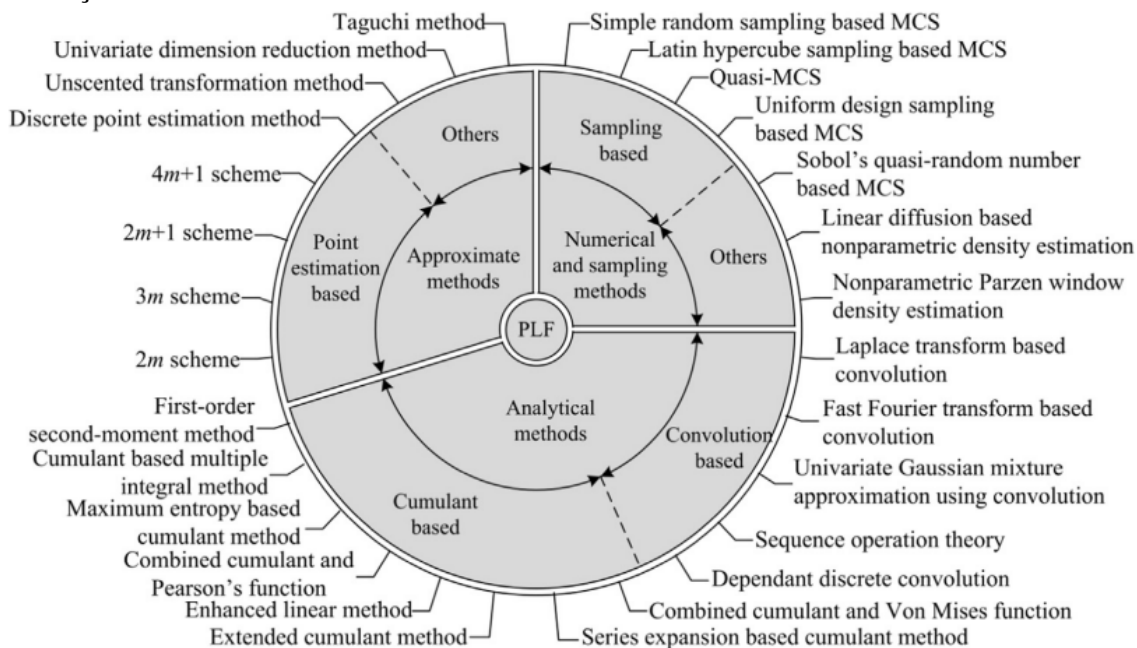
As situações de independência e dependência positiva entre variáveis resultam em demandas agregadas bastante diferentes. No caso de dependência, valores extremos tem maior chance de ocorrência simultânea, surgindo assim cenários com maior carregamento do que aqueles que consideram as variáveis independentes (PAPAEFTHYMIU, 2007; RODRIGUES, 2008).

Uma alternativa ao método de análise determinística de redes elétricas tradicional é o fluxo de potência probabilístico (PLF), cujos principais aspectos contemplam o tipo de incerteza considerada, efeito da relação de dependência entre as variáveis e métodos de avaliação. O início dos estudos sobre PLF se deu com a publicação do trabalho de (BORKOWSKA, 1974) e os estudos realizados sobre o assunto até o ano de 1985 focavam, principalmente, em questões como linearização, interdependência entre potências nodais e efeito das faltas na rede. O foco das pesquisas mudou a partir de então para investigações que buscavam aprimorar os métodos de eficiência computacional e precisão. Com o advento das fontes de energia renovável distribuídas, a década de 2005 a 2015 apresentou grande interesse em verificar o impacto da inclusão destas fontes nas redes elétricas, tanto na transmissão como na distribuição (PRUSTY; JENA, 2016).

Um revisão bibliográfica minuciosa sobre estudos acerca de fluxo de potência probabilístico foi recentemente realizada por Prusty; Jena (2016) em continuação e complementação à revisão de Chen; Chen; Bak-Jensen (2008). É mostrado que o uso de métodos analíticos foi a preferência dos pesquisadores nas primeiras pesquisas realizadas sobre o tema, tendo ocorrido mais recentemente um aumento no interesse por métodos de aproximação bem como por métodos numéricos e de amostragem. A Figura 1 ilustra a divisão das pesquisas de acordo com os métodos de avaliação.

O uso de métodos numéricos e de amostragem é uma tendência crescente em função da inserção de fontes renováveis, sendo o método de Simulação de Monte Carlo (SMC) a referência para comparação e validação de outros métodos de PLF em função de sua exatidão. Contudo, o elevado número de amostras necessário torna o seu uso pouco atrativo em sistemas reais. Nesse aspecto, Cui; Franchetti (2013) demonstra que uma alternativa é o método Quasi-Monte Carlo para a solução do fluxo de potência probabilístico em sistemas de distribuição radiais, tendo esse apresentado maior velocidade de convergência, principalmente no tratamento de variáveis aleatórias com alta variância. Ainda nesse

Figura 1: Pesquisas sobre fluxo de potência probabilístico classificadas por método de avaliação.



Fonte: (PRUSTY; JENA, 2016).

sentido, o fluxo de potência probabilístico pelo método de estimação por pontos (MEP), até então usado apenas para sistemas de transmissão, foi aplicado por Gallego; Echeverri; Feltrin (2012) a sistemas de distribuição trifásicos desbalanceados, sendo verificado que os tempos de execução são inferiores ao da SMC.

As principais incertezas investigadas podem ser divididas entre incertezas dos dados de entradas (geração e demanda) e incertezas da rede (topologia e parâmetros) (PRUSTY; JENA, 2016). O tratamento de incertezas na geração convencional costuma ser expressa em termos de uma taxa forçada de faltas e modela a distribuição de forma discreta através de uma distribuição de Bernoulli quando são considerados dois estados de operação e através de uma distribuição binomial em caso de mais estados. A aleatoriedade associada a fontes não-despacháveis de geração distribuída costuma ser modelada por meio de uma distribuição contínua de probabilidade associada à fonte primária responsável pela produção de energia combinado à característica de conversão do gerador. A incerteza mais marcante em sistemas de energia elétrica é a proveniente das cargas que é derivada, principalmente, das condições ambientais, variações nos equipamentos utilizados e comportamento dos consumidores. Na ausência de dados de medição, o uso de distribuição normal é geralmente adotada. As incertezas da rede são provenientes de faltas ou da variação dos parâmetros como consequência de fatores climáticos. Assim, a mudança na topologia costuma ser caracterizada por uma distribuição discreta de probabilidade (PRUSTY; JENA, 2016).

Um fluxo de potência probabilístico é considerado incompleto na ausência da relação de dependência entre as variáveis do sistema. A maneira geralmente utilizada para modelar a correlação é através do uso de um coeficiente de correlação em conjunto com uma função de cópula. O coeficiente de correlação linear é considerado na maioria das vezes devido à facilidade do uso. No entanto, o coeficiente de correlação de *rank* de Spearman ou de Kendall são mais adequados para relações de dependência com maior

dispersão (PRUSTY; JENA, 2016).

1.2 Objetivo

Nesse contexto, o presente trabalho propõe uma abordagem probabilística para analisar o desempenho de redes de distribuição considerando a natureza incerta da carga e da GD. As contribuições do trabalho incluem o agrupamento dos consumidores e GD que apresentam comportamentos similares (visando reduzir o número de variáveis aleatórias do problema), a consideração da correlação entre os diversos agrupamentos, a determinação de indicadores de operação probabilísticos e sua comparação com os resultados obtidos pelos métodos determinísticos.

1.3 Estrutura

O Capítulo 2 apresenta conceitos importantes acerca de variáveis aleatórias e de operações entre as mesmas, utilizados de forma ampla em métodos de fluxo de potência probabilístico.

Os principais trabalhos envolvendo fluxo de potência probabilístico utilizados no desenvolvimento deste trabalho são explorados em maior profundidade no Capítulo 3.

O Capítulo 4 propõe um método de análise da operação de sistemas de distribuição com o objetivo de obter indicadores probabilísticos das perdas no sistema bem como para a qualidade da tensão de fornecimento. São sorteadas as potências dos componentes do sistema de modo a compor um cenário para a curva de carga. Com isso realiza-se o fluxo de potência e os resultados são utilizados para calcular o que caracteriza uma amostra do indicador de interesse. Após atingir o critério de convergência, uma análise estatística aliada a uma adequação de ajuste permite descrever as nuances dos indicadores para um universo de possibilidades de operação da rede.

No Capítulo 5, o método é aplicado no sistema de 13 barras do IEEE, considerando um cenário de cargas e GDFV típico. O cálculo do fluxo de potência trifásico desequilibrado foi viabilizado pelo uso do programa OpenDSS[®], sendo as demais rotinas implementadas com o uso do programa MATLAB[®].

2 INCERTEZA EM SISTEMAS DE POTÊNCIA

Alguns conceitos de probabilidade e estatística são indispensáveis para o desenvolvimento deste trabalho. Sem a pretensão de esgotar este assunto, os principais conceitos utilizados nesta dissertação são descritos a seguir.

2.1 Incerteza dos Eventos

A ocorrência de algum fenômeno cuja observação se deseja realizar, pode ser chamada de evento. Os eventos podem ser classificados em certos, impossíveis e aleatórios. Se sempre que um conjunto de condições \mathcal{C} se realizar observa-se, necessariamente, também a ocorrência de um determinado evento, diz-se que este evento é certo. Um evento que jamais ocorre concomitante ao conjunto de condições \mathcal{C} é dito impossível. Caso a ocorrência do evento seja observada apenas em algumas ocasiões da realização do conjunto de condições \mathcal{C} , então diz-se que este evento é aleatório (MONTGOMERY; RUNGER, 2007; BECKER, 2015; CAMPOS; RÊGO; MENDONÇA, 2017).

Para vários fenômenos aleatórios pode-se atribuir uma quantidade estimativa da possibilidade de sua ocorrência. Neste caso, não há como prever se o evento ocorrerá, mas sim, com base em experimentos e observações, determinar a chance de sua ocorrência, ou seja, calcular a probabilidade de ocorrência do evento (MONTGOMERY; RUNGER, 2007; BECKER, 2015; CAMPOS; RÊGO; MENDONÇA, 2017).

Todos casos estudados em teoria de probabilidade envolvem um conjunto de condições \mathcal{C} e uma família de eventos que não necessariamente ocorrem a cada realização do conjunto de condições \mathcal{C} . Assim, a família de eventos \mathcal{F} que for fechada às operações de produto, soma e diferença entre eventos é chamada de campo de eventos. Já o campo de eventos \mathcal{F} contendo o evento certo e fechado às operações de produto e soma de infinitos eventos é chamado de σ -álgebra de eventos (BECKER, 2015; LAW; KELTON, 2000).

Desse modo, existe um número não negativo $P(\mathcal{A})$, i.e., a probabilidade de ocorrer o evento \mathcal{A} , associado a cada evento \mathcal{A} na σ -álgebra de eventos \mathcal{F} (axioma da existência). A chance de ocorrência de todos os eventos (Ω) é dada por $P(\Omega) = 1$ (axioma da normalização). Por fim, a probabilidade associada à ocorrência de cada evento pode ser somada e o resultado é igual à probabilidade associada à soma dos eventos (axioma da σ -aditividade). Os três axiomas supracitados são suficientes para construir toda a teoria de probabilidades e todos os demais resultados decorrem da combinação dos mesmos e das operações entre os eventos (MONTGOMERY; RUNGER, 2007; BECKER, 2015; CAMPOS; RÊGO; MENDONÇA, 2017).

2.2 Variáveis Aleatórias (VAs)

Uma variável aleatória (VA) é uma função (ou regra) que atribui um número real (\mathbb{R}) a cada resultado no espaço amostral de um experimento aleatório (MONTGOMERY; RUNGER, 2007; LAW; KELTON, 2000).

A definição formal considera um conjunto de condições \mathcal{C} , um conjunto Ω de eventos elementares e , uma σ -álgebra de eventos \mathcal{F} (subconjunto de Ω) e P a medida de probabilidade definida sobre \mathcal{C} . Um evento é dito elementar quando não pode ser decomposto. Assim, supõe-se que exista um número $X = f(e)$ associado a cada evento elementar $e \in \Omega$, ou alternativamente, que haja uma função f mapeando o conjunto Ω no conjunto dos números reais ($f : \Omega \rightarrow \mathbb{R}$).

Desse modo, diz-se que X é uma variável aleatória (VA) se f é mensurável com relação a P , isto é, o conjunto $\mathcal{A}_x = \{e \in \Omega | f(e) \leq x\} \in \mathcal{F}, \forall x \in \mathbb{R}$. Nesse caso, a probabilidade de \mathcal{A}_x é dada por $P(\mathcal{A}_x) = P(X \leq x) = F(x)$, onde $F(x)$ é chamada de função de distribuição acumulada (CDF – *Cumulative Distribution Function*) da variável aleatória X (BECKER, 2015; MONTGOMERY; RUNGER, 2007; LAW; KELTON, 2000; CAMPOS; RÊGO; MENDONÇA, 2017).

A lei de distribuição é a forma de descrição do comportamento da VA que permite obter sua CDF. Uma VA pode ser discreta ou contínua, mas o maior interesse neste trabalho está nas variáveis aleatórias contínuas, que satisfazem a seguinte equação (BECKER, 2015; MONTGOMERY; RUNGER, 2007):

$$F(x) = \int_{-\infty}^x f(z) dz. \quad (1)$$

A $f(x)$ é uma função real não negativa e é chamada de função densidade de probabilidade (PDF – *Probability Density Function*) da variável aleatória X . Logo, se a CDF da VA X é diferenciável, a PDF $f(x) = F'(x)$, onde $F'(x)$ é a derivada de $F(x)$. A função de sobrevivência (*survival*) $S(x)$ é formalmente expressa conforme (BECKER, 2015; MONTGOMERY; RUNGER, 2007):

$$S(x) = 1 - F(x). \quad (2)$$

2.2.1 Valor Esperado, Quantis e Momentos Matemáticos de VAs

Seja uma variável aleatória contínua X com PDF $f(x)$. Obtém-se a esperança matemática ou simplesmente o valor esperado de X por meio da seguinte integral, quando esta existir (MONTGOMERY; RUNGER, 2007; CAMPOS; RÊGO; MENDONÇA, 2017):

$$E(X) = \int_{-\infty}^{+\infty} x f(x) dx. \quad (3)$$

De forma geral, dado um número real qualquer α e um inteiro positivo γ , diz-se que o valor esperado da variável aleatória $(X - \alpha)^\gamma$ é o γ -ésimo momento da variável aleatória X , denotado por $m_\gamma(\alpha)$, conforme (BECKER, 2015):

$$m_\gamma(\alpha) = E((X - \alpha)^\gamma). \quad (4)$$

Dessa maneira, o primeiro momento em torno da origem ($\alpha = 0$) da VA X é o próprio valor esperado. Tomando $\alpha = E(X)$, obtêm-se os momentos centrais da VA, onde um dos mais utilizados é o segundo momento central, conhecido como variância da VA, dado por:

$$V(X) = E((X - E(X))^2), \quad (5)$$

contudo a sua aplicação é mais conhecida pelo desvio padrão, que é simplesmente a raiz quadrada da variância (BECKER, 2015).

O comportamento das VAs pode ser analisado com o auxílio de alguns operadores, destacando-se o operador Esperança, $E(X)$, o qual possibilita obter o valor esperado de uma soma finita de n VAs dado por (BECKER, 2015):

$$E\left(\sum_{i=1}^n X_i\right) = \sum_{i=1}^n E(X_i). \quad (6)$$

O operador Esperança associa a cada VA o seu valor esperado, i.e., a sua média. Assim como a média, existem outros parâmetros das VAs que auxiliam na descrição da mesma, como os quantis.

Quantis são os $n - 1$ valores de uma VA que dividem os dados ordenados em n subconjuntos de igual proporção. Assim, o k -ésimo n -quantil é o valor x tal que $P(X \leq x) = \frac{k}{n}$ para a VA X , sendo k e n números naturais não-nulos. O quantil de ordem q de uma VA contínua cuja CDF é dada por $F(x)$ é qualquer raiz da seguinte equação (BECKER, 2015; MONTGOMERY; RUNGER, 2007):

$$F(x) = q = \frac{k}{n}, \quad (7)$$

com $q \in (0, 1)$.

Chamam-se quartis os quantis de ordem $\frac{1}{4}$, $\frac{2}{4}$ e $\frac{3}{4}$ (denotados por x_{Q_1} , x_{Q_2} e x_{Q_3}). Assim, a mediana é simplesmente o quantil de ordem $\frac{1}{2}$ (BECKER, 2015). Uma maneira de representar graficamente estas estatísticas é através do diagrama de caixa e linha conforme ilustrado na Figura 2, para uma amostra de população com distribuição normal. O diagrama de caixa e linha é uma representação simplificada que descreve simultaneamente características importantes como mediana, dispersão, simetria e observações não-usuais (*outliers*) para um conjunto de dados (MONTGOMERY; RUNGER, 2007). Na Figura 2, x_{Q_1} e x_{Q_3} são tais que $F(x_{Q_1}) = \frac{1}{4}$ e $F(x_{Q_3}) = \frac{3}{4}$. Além disso, o intervalo interquartil (IQR) é dado pela diferença entre x_{Q_3} e x_{Q_1} . Observa-se que no exemplo da distribuição normal o percentual de amostras no IQR é de 50%, diferente do percentual de 68,27% delimitado em termos de desvio padrão.

Já para uma amostra de uma população com distribuição lognormal, a Figura 3 apresenta o respectivo diagrama de caixa e linha. Neste caso, o diagrama é assimétrico visto que o limite inferior de um diagrama de caixa e linha é o menor valor da amostras dentro do intervalo que se estende até a faixa de 1,5 interquartil abaixo de x_{Q_1} enquanto que o limite superior é o maior valor dentro da faixa de 1,5 interquartil acima de x_{Q_3} (MONTGOMERY; RUNGER, 2007).

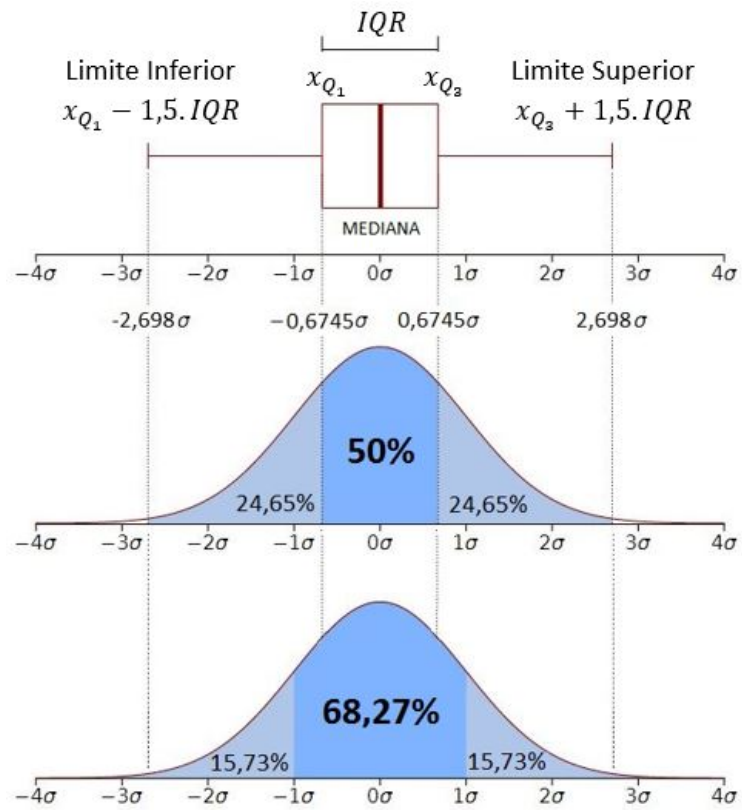
2.2.2 Relações de Dependência entre Variáveis Aleatórias

Sejam X e Y duas variáveis aleatórias contínuas. Diz-se que X e Y são variáveis aleatórias independentes se, para qualquer x e qualquer y , a seguinte expressão é satisfeita (MONTGOMERY; RUNGER, 2007):

$$f_{X,Y}(x, y) = f_X(x)f_Y(y), \quad (8)$$

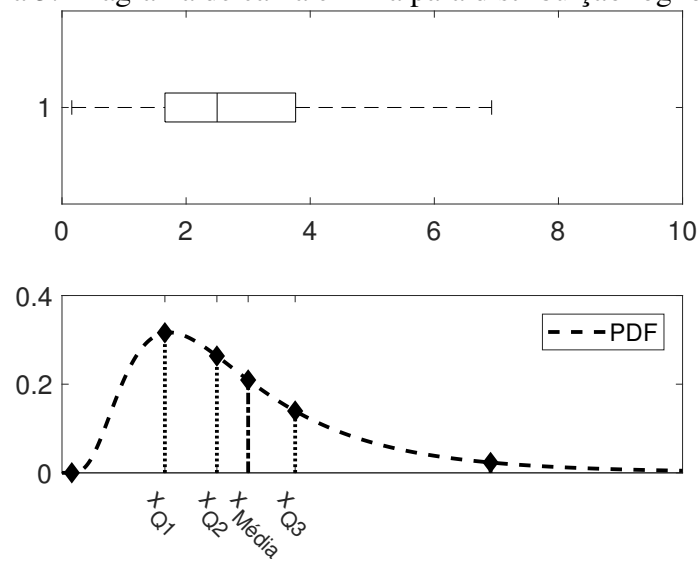
onde $f_{X,Y}(x, y)$ é a função densidade de probabilidade conjunta para as VAs X e Y , $f_X(x)$ é a função densidade de probabilidade marginal da VA X e $f_Y(y)$ é a função densidade de probabilidade marginal da VA Y .

Figura 2: Diagrama de caixa e linha para uma distribuição normal.



Fonte: O autor.

Figura 3: Diagrama de caixa e linha para distribuição lognormal.



Fonte: O autor.

Supondo que tanto as esperanças $E(X)$ e $E(Y)$ quanto as variâncias $V(X)$ e $V(Y)$ existam e sejam finitas,, a covariância entre X e Y é dada por:

$$\text{Cov}(X, Y) = E[(X - E(X))(Y - E(Y))] = E(XY) - E(X)E(Y), \quad (9)$$

sendo válida tanto para o caso discreto como para o caso contínuo. A covariância é uma medida da interdependência de duas variáveis aleatórias. Dessa forma, se X e Y são independentes, então $\text{Cov}(X, Y) = 0$. No entanto, se $\text{Cov}(X, Y) = 0$ não necessariamente X e Y serão independentes (BECKER, 2015; MONTGOMERY; RUNGER, 2007).

A covariância é uma medida da relação linear que depende da escala das variáveis aleatórias. O coeficiente de correlação linear, ou correlação de Pearson, oferece uma medida independente de escala, sendo uma quantidade adimensional para a medida de dependência das variáveis aleatórias X e Y e é expresso por (BECKER, 2015; MONTGOMERY; RUNGER, 2007):

$$\rho_{XY} = \frac{\text{Cov}(X, Y)}{\sqrt{V(X)V(Y)}}, \quad (10)$$

onde ρ_{XY} é a correlação linear entre X e Y , $V(X)$ é a variância de X e $V(Y)$ é a variância de Y . O coeficiente de correlação linear pode assumir valores entre $-1 \leq \rho \leq 1$.

Comonotonicidade é o conceito de dependência que se refere ao caso de dependência perfeita positiva entre variáveis aleatórias, provendo um conceito de dependência estocástica mais amplo que a correlação linear. Variáveis aleatórias comonotônicas¹ são funções crescentes com o mesmo fator subjacente aleatório, isto é, elas sempre variam da mesma maneira, de forma que o crescimento de uma implica o crescimento da outra. No entanto, a proporção de crescimento entre as mesmas não é necessariamente constante, e portanto, a relação de dependência não é linear (PAPAEFTHYMIU, 2007; RODRIGUES, 2008; FARIA, 2014).

A Figura 4 ilustra um exemplo de dispersão entre as demandas de energia elétrica de dois consumidores em situação de independência (pontos vermelhos) e com correlação positiva (cruzes pretas), para 1000 amostras. Conforme ilustra o gráfico (a) da Figura 4, a correlação linear entre as demandas é de $\rho = 0,85$, enquanto que no gráfico (b) é de $\rho = 0,99$. É possível observar no gráfico (b) da Figura 4 que a relação de dependência entre as demandas não é perfeitamente linear, mesmo com um coeficiente de Pearson de $\rho = 0,99$.

O coeficiente de correlação linear fornece uma representação completa da dependência entre VAs com distribuição normal. Para variáveis com distribuições distintas, como é o caso de variáveis comonotônicas, a correlação de *rank*, também conhecida como correlação de Spearman ou de posto, é mais adequada (RODRIGUES, 2008; PAPAEFTHYMIU, 2007).

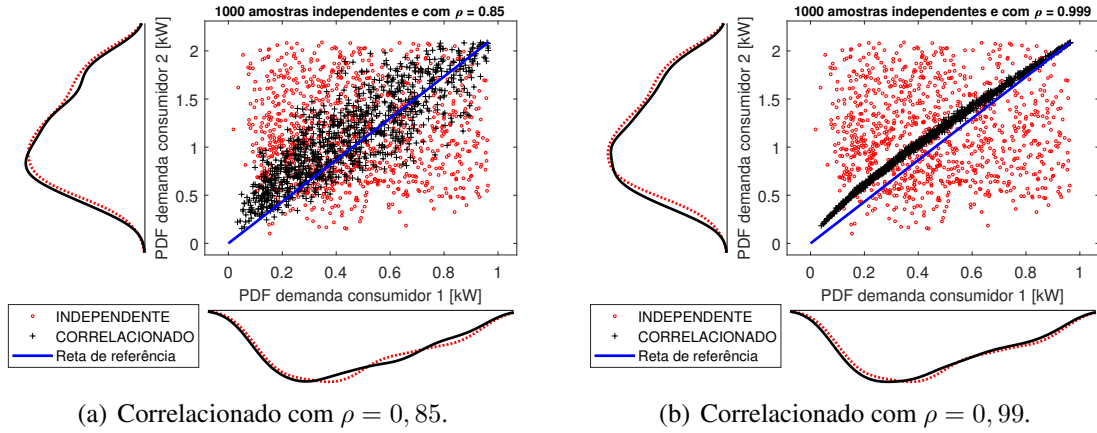
Por definição, para uma variável aleatória X com uma função acumulada de probabilidade (CDF) $F_X(x) = P(X \leq x)$, a variável aleatória $F_X(X)$ segue uma distribuição uniforme no intervalo $[0, 1]$, i.e., $\mathcal{U}(0, 1)$. Sendo $r \in [0, 1]$, a demonstração é dada por (RODRIGUES, 2008; PAPAEFTHYMIU, 2007):

$$P(F_X(X) \leq r) = P(X \leq F_X^{-1}(r)) = F_X[F_X^{-1}(r)] = r. \quad (11)$$

Desse modo, se U_X segue uma distribuição uniforme, $F_X(X) = U_X \leftrightarrow X = F_X^{-1}(U_X)$, ou seja, $F_X^{-1}(U_X)$ segue a distribuição de X (RODRIGUES, 2008; PAPAEFTHYMIU, 2007).

¹A relação entre duas variáveis comonotônicas é descrita por uma função monótona.

Figura 4: Dispersão entre as demandas de energia elétrica de dois consumidores com diferentes relações de dependência.



Fonte: O autor.

A correlação de *rank* ρ_r para as variáveis aleatórias X e Y , com variáveis de distribuições uniformes U_X e U_Y associadas, por conseguinte, é dada por (RODRIGUES, 2008; PAPAETHYMIU, 2007):

$$\rho_r(X, Y) = \rho(U_X, U_Y). \quad (12)$$

Portanto, a correlação de *rank* para X e Y é igual à correlação linear entre as variáveis U_X e U_Y com distribuições uniformes associadas à X e à Y respectivamente. A relação entre correlação linear e correlação de *rank* é dada por (RODRIGUES, 2008; PAPAETHYMIU, 2007):

$$\rho(X, Y) = 2\text{sen}\left(\frac{\pi}{6}\rho_r(X, Y)\right). \quad (13)$$

2.2.3 Distribuições de Probabilidade Notáveis

Dentre todas as distribuições de probabilidade existentes, algumas se destacam por sua ampla aplicação. No caso de sistemas de energia elétrica não é diferente. De forma sucinta as principais distribuições de probabilidade utilizadas na análise probabilística da operação de redes de distribuição, conforme listadas na Tabela 1, são: distribuição uniforme ($\mathcal{U}(a, b)$), normal ($\mathcal{N}(\mu, \sigma)$), lognormal ($\text{Logn}(\mu, \sigma^2)$), Weibull ($\mathcal{W}(\lambda, k)$) e beta ($\text{Beta}(\alpha, \beta)$) (BECKER, 2015; MONTGOMERY; RUNGER, 2007).

2.2.4 Distribuições Multidimensionais

Em alguns casos, um experimento é definido para um conjunto de VAs. Nesse caso, a função densidade de probabilidade do vetor aleatório $[X_1, X_2, \dots, X_n]$ existe se a função $f(x_1, x_2, \dots, x_n)$ é não negativa e, para quaisquer x_1, x_2, \dots, x_n , a seguinte equação é satisfeita (BECKER, 2015; MONTGOMERY; RUNGER, 2007):

$$\begin{aligned} F(x_1, x_2, \dots, x_n) &= \int_{-\infty}^{x_1} \int_{-\infty}^{x_2} \dots \int_{-\infty}^{x_n} f(z_1, z_2, \dots, z_n) dz_n \dots dz_2 dz_1 \\ &= P(\{X_1 \leq x_1, X_2 \leq x_2, \dots, X_n \leq x_n\}). \end{aligned} \quad (14)$$

A função $F(x_1, x_2, \dots, x_n)$ é chamada de função de distribuição acumulada n -dimensional de probabilidade do vetor aleatório $[X_1, X_2, \dots, X_n]$ (BECKER, 2015; MONTGOMERY; RUNGER, 2007).

Tabela 1: Sumário dos principais parâmetros de algumas distribuições de probabilidade.

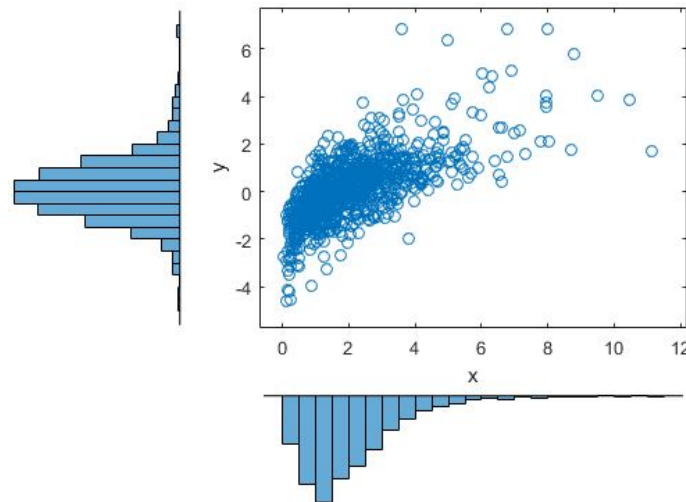
Notação	Parâmetros	PDF	Domínio	Média	Variância
$\mathcal{U}(a, b)$	$a, b \in \mathbb{R} a \leq b$	$\frac{1}{b-a}$	$x \in [a, b]$	$\frac{1}{2}(b+a)$	$\frac{1}{12}(b+a)^2$
$\mathcal{N}(\mu, \sigma)$	$\mu \in \mathbb{R}, \sigma^2 > 0$	$\frac{1}{\sigma\sqrt{2\pi}} e^{-1/2(\frac{x-\mu}{\sigma})^2}$	$x \in \mathbb{R}$	μ	σ^2
$\text{Logn}(\mu, \sigma^2)$	$\mu \in \mathbb{R}, \sigma > 0$	$\frac{1}{x\sigma\sqrt{2\pi}} e^{-1/2(\frac{\ln(x)-\mu}{\sigma})^2}$	$x \in \mathbb{R}_+$	$e^{\mu+\frac{\sigma^2}{2}}$	$e^{\sigma^2-1} e^{2\mu+\sigma^2}$
$\mathcal{W}(\lambda, k)$	$\lambda, k \in \mathbb{R}_+^*$	$\frac{k}{\lambda} (\frac{x}{\lambda})^{k-1} e^{-(x/\lambda)^k}$	$x \in \mathbb{R}_+$	$\lambda\Gamma(1+\frac{1}{k})$	$\lambda^2 [\Gamma(1+\frac{2}{k}) - (\Gamma(1+\frac{1}{k}))^2]$
$\text{Beta}(\alpha, \beta)$	$\alpha, \beta > 0$	$\frac{x^{\alpha-1}(1-x)^{\beta-1}}{B(\alpha, \beta)}$	$x \in [0, 1]$	$\frac{\alpha}{\alpha+\beta}$	$\frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$

Notas: $\Gamma(x) = \int_0^{\infty} s^{x-1} e^{-s} ds$ é a chamada função gama e;

$$B(\alpha, \beta) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}.$$

Tomando um valor de x_k arbitrariamente elevado, o evento $\{X_k \leq x_k\}$ será um evento certo. Este procedimento é chamado de redução de ordem e, estendendo esse processo, é possível determinar a chamada distribuição acumulada marginal de probabilidade $F_{X_k}(x)$ da variável X_k , isto é, $F_{X_k}(x) = F(c_1, c_2, \dots, c_n)$, onde $c_k = x$ e $c_i = \infty$ para todos $i \neq k$ e $1 < i < n$ (BECKER, 2015; MONTGOMERY; RUNGER, 2007). A Figura 5 ilustra a dispersão entre duas VAs e as distribuições marginais nos respectivos eixos cartesianos. Cada circunferência do gráfico representa um par ordenado de uma observação de X associada à Y .

Figura 5: Exemplo de distribuição conjunta para duas variáveis aleatórias.



Fonte: (MATHWORKS, 2017)

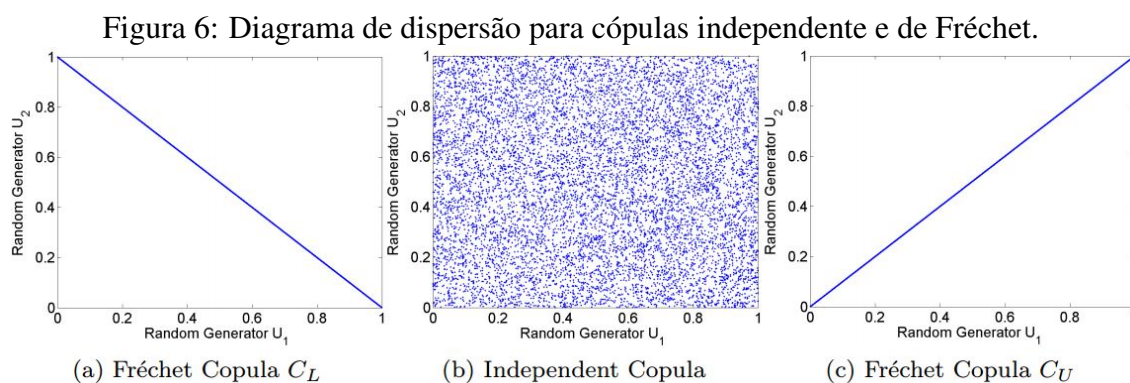
2.2.5 Cópulas

A palavra cópula se refere a uma função que acopla distribuições unidimensionais para formar distribuições multidimensionais. Cópulas são, por definição, funções de distribuição multidimensionais cujas distribuições marginais são uniformes no intervalo $[0, 1]$. Assim, a cópula define uma distribuição multivariável (multidimensional) no cubo unitário $[0, 1]^n$. A informação acerca da estrutura de dependência entre as VAs está contida na distribuição de classificação (*rank*) marginal das mesmas (PAPAEFTHYMIU, 2007; RODRIGUES, 2008; MATHWORKS, 2017).

A definição formal é dada pelo Teorema de Sklar onde: sejam as VAs X e Y com

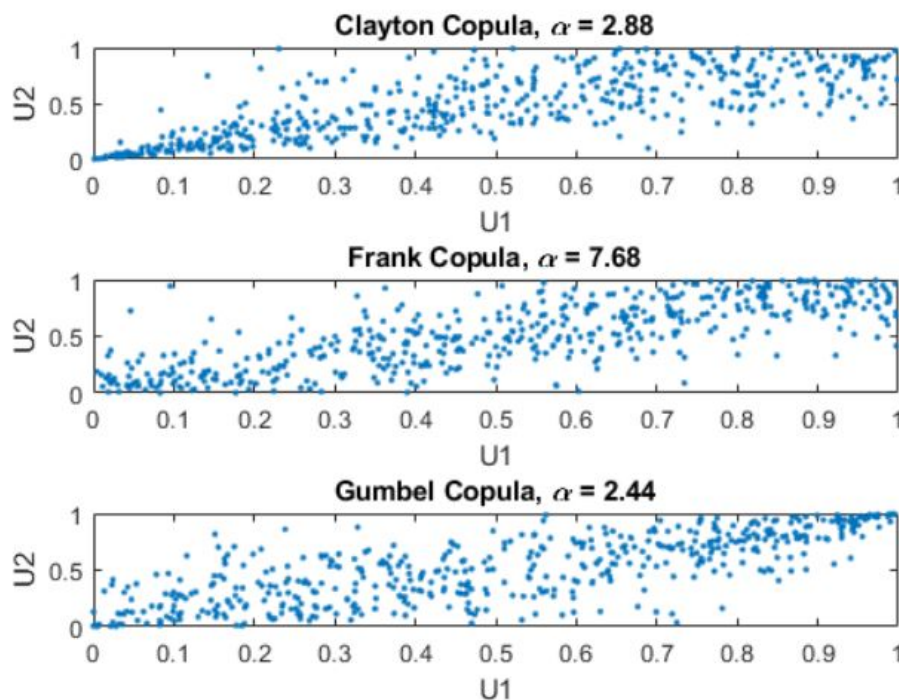
CDF F_x e F_y , respectivamente, existe uma cópula C se sua distribuição conjunta pode ser escrita por $F_{XY}(x, y) = C(F_X(x), F_Y(y))$ (PAPAEFTHYMIU, 2007).

Pode ser demonstrado que para qualquer cópula C , $C_L \leq C \leq C_U$, onde C_L e C_U são distribuições cuja massa é distribuída uniformemente na anti-diagonal e na diagonal principal, respectivamente. Em outras palavras, se as VAs X e Y são acopladas pela cópula C_U (ou C_L), então a correlação de classificação (*rank*) é dada por $\rho_r(X, Y) = 1$ (ou -1). Isso significa que todas as estruturas de dependência entre duas VAs estão delimitadas entre o caso de dependência perfeita positiva (limite superior) e dependência perfeita negativa (limite inferior). Estes limites são chamados fronteiras de Fréchet-Hoeffding, conforme ilustrado na Figura 6 (PAPAEFTHYMIU, 2007; PAPAEFTHYMIU; KUROWICKA, 2009).



Fonte: (PAPAEFTHYMIU, 2007)

Figura 7: Exemplo de famílias de cópulas.

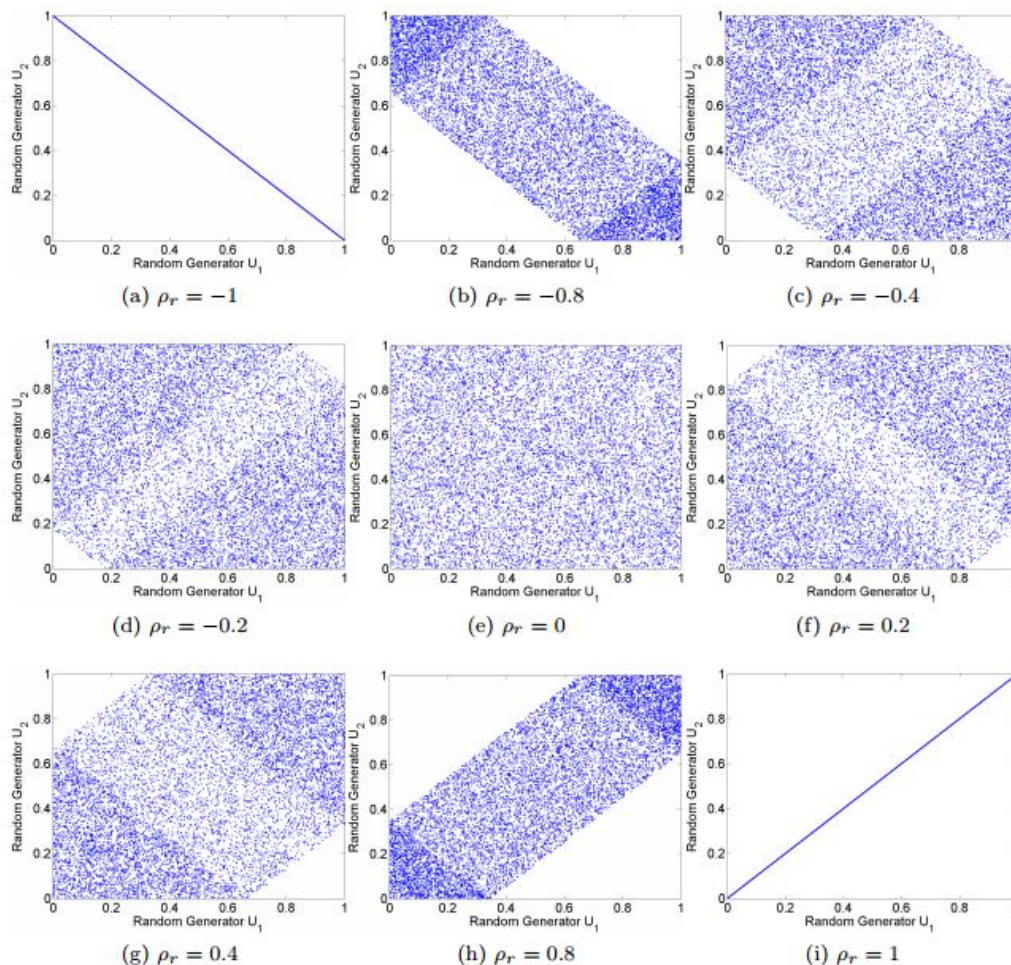


Fonte: (MATHWORKS, 2017)

As cópulas também são referenciadas como representações uniformes ou funções de dependência. Vários pesquisadores se dedicaram a investigar as cópulas, originando o que

se conhece por famílias de cópulas, como as cópulas normais, as elípticas, banda diagonal, entre outras. Este trabalho não tem por objetivo explorar esse assunto em profundidade, mas é importante salientar que cada família resulta em uma distribuição multidimensional diferente, como pode ser observado na Figura 7 e que a correlação entre as variáveis proporciona diferentes efeitos, de acordo com a Figura 8 (PAPAEFTHYMIU, 2007; PAPAEFTHYMIU; KUROWICKA, 2009).

Figura 8: Diagrama de dispersão para diferentes correlações das cópulas de banda diagonal.



Fonte: (PAPAEFTHYMIU, 2007)

A cópula normal é obtida pela transformação das marginais normais padrão das distribuições uniformes multidimensionais em distribuições uniformes classificadas. Esta cópula é empregada na *Joint-Normal Transform* (descrita posteriormente na Seção 2.4.3) na modelagem da dependência multidimensional (RODRIGUES, 2008).

2.3 Inferência Estatística

Nesta seção são abordadas ferramentas para análise dos dados amostrais de uma população. A partir de uma suposição fundamental e suporte de métodos indutivos, considerando um erro probabilístico fixo (confiança estatística), é possível realizar asserções a respeito dessa população.

2.3.1 Estatísticas Amostrais

Dada a amostra para uma população, muitas vezes deseja-se obter alguma estatística desta amostra (CAMPOS; RÊGO; MENDONÇA, 2017). Se representarmos por X a variável da população, uma amostragem aleatória de tamanho n será interpretada como uma sequência de n VAs X_1, X_2, \dots, X_n , sendo cada uma distribuída segundo a variável X , independentes entre si. Nesse caso diz-se que as variáveis (observações) são independentes e identicamente distribuídas (IID). Assim, seja uma amostra de tamanho n da variável aleatória X (denotada por X_1, X_2, \dots, X_n), a média da amostra também será uma variável aleatória dada por (BECKER, 2015; MONTGOMERY; RUNGER, 2007):

$$\bar{X} = \frac{X_1 + X_2 + \dots + X_n}{n} = \frac{1}{n} \sum_{i=1}^n X_i, \quad (15)$$

sendo a variância da amostra dada por:

$$s^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2 = \frac{1}{n-1} \left(\sum_{i=1}^n X_i^2 - n\bar{X}^2 \right). \quad (16)$$

Para um conjunto de observações X_1, X_2, \dots, X_n de uma variável aleatória X com média μ e variância σ^2 , têm-se que a esperança e a variância amostral são dadas por $E[\bar{X}] = \mu$ e $V[\bar{X}] = \frac{\sigma^2}{n}$ (CAMPOS; RÊGO; MENDONÇA, 2017; MONTGOMERY; RUNGER, 2007).

São muitos os casos em que a média e a variância de uma variável aleatória são desconhecidas. Nesses casos, deve-se utilizar algum método de estimação pontual para a obtenção de estimadores desses momentos da variável aleatória (CAMPOS; RÊGO; MENDONÇA, 2017; MONTGOMERY; RUNGER, 2007). Existem dois métodos amplamente utilizados para essa finalidade: o método dos momentos e o método da máxima verossimilhança. Estimadores de máxima verossimilhança são geralmente preferidos aos de momento. No entanto, em função da facilidade de cálculo, o método dos momentos é mais utilizado. O importante é que ambas as técnicas produzem estimadores pontuais não tendenciosos² (MONTGOMERY; RUNGER, 2007).

O método dos momentos estabelece que os estimadores de momento $\hat{\Theta}_1, \hat{\Theta}_2, \dots, \hat{\Theta}_k$, para os parâmetros desconhecidos $\theta_1, \theta_2, \dots, \theta_k$, são determinados igualando os n primeiros momentos da população aos k primeiros momentos da amostra e resolvendo as equações resultantes para os parâmetros desconhecidos (MONTGOMERY; RUNGER, 2007). Desse modo, a determinação dos momentos é realizada com o uso da Equação (4), para $\alpha = 0$ e $1 \leq \gamma \leq k$.

2.3.2 Lei dos Grandes Números e Teorema do Limite Central

Diz-se que a sequência X_1, X_2, \dots, X_n de variáveis aleatórias independentes e identicamente distribuídas com $E(X_i) < +\infty$ para todo i obedece à lei forte dos grandes números se, à medida em que $n \rightarrow +\infty$, a seguinte expressão é satisfeita:

$$\lim_{n \rightarrow +\infty} \left(\frac{1}{n} \sum_{k=1}^n X_k \right) = \frac{1}{n} \sum_{k=1}^n E(X_k). \quad (17)$$

A existência do valor esperado é condição necessária e suficiente para que a sequência de VAs independentes e de mesma distribuição obedeçam à lei forte dos grandes números.

²Diz-se que $\hat{\Theta}$ é um estimador não-tendencioso de θ se $E(\hat{\Theta}) = \theta$.

Se a mesma sequência X_1, X_2, \dots, X_n de variáveis aleatórias independentes e identicamente distribuídas possuem variância σ^2 finita e diferente de zero, a distribuição da média amostral de n amostras tende a uma distribuição normal de probabilidade à medida que $n \rightarrow +\infty$. Este é um dos teoremas mais importantes da estatística e é chamado de Teorema do Limite Central (BECKER, 2015; CAMPOS; RÊGO; MENDONÇA, 2017).

2.3.3 Intervalos de Confiança

Seja X_1, X_2, \dots, X_n uma amostra aleatória com n observações e deseja-se obter um intervalo de confiança para um parâmetro desconhecido θ . Supõe-se que é possível encontrar uma estatística $g(X_1, X_2, \dots, X_n; \theta)$ que depende da amostra e de θ , mas cuja distribuição de probabilidades não depende de θ ou de qualquer outro parâmetro desconhecido. Um parâmetro comumente utilizado é a média μ de uma população com $g(X_1, X_2, \dots, X_n; \theta) = (\bar{X} - \mu)/(\sigma/\sqrt{n})$, onde n é o número de observações e σ/\sqrt{n} é o desvio padrão amostral de \bar{X} cuja variância σ^2 é conhecida (MONTGOMERY; RUNGER, 2007).

A seguir, deve-se determinar as constantes C_L e C_U tal que:

$$P(C_L \leq g(X_1, X_2, \dots, X_n; \theta) \leq C_U) = 1 - \alpha. \quad (18)$$

A manipulação das desigualdades resulta em:

$$P(L(X_1, X_2, \dots, X_n) \leq \theta \leq U(X_1, X_2, \dots, X_n)) = 1 - \alpha. \quad (19)$$

Definindo o intervalo de confiança de $100(1-\alpha)\%$ para θ , os limites inferior e superior de confiança são fornecidos por $L(X_1, X_2, \dots, X_n)$ e $U(X_1, X_2, \dots, X_n)$ (MONTGOMERY; RUNGER, 2007).

Se o tamanho da amostra n de X_1, X_2, \dots, X_n for grande e a média μ e variância σ^2 são grandezas desconhecidas, o Teorema do Limite Central implica que \bar{X} tenha aproximadamente uma distribuição normal com média μ e variância σ^2/n . Apesar de o desvio padrão σ ser desconhecido, a troca de σ pelo desvio padrão amostral S tem pouco efeito na distribuição $Z = (\bar{X} - \mu)/(\sigma/\sqrt{n})$ (MONTGOMERY; RUNGER, 2007). Desse modo, o intervalo de confiança para amostras grandes ($n \geq 40$) da média μ com nível de confiança de $100(1 - \alpha)\%$ é dado por:

$$\bar{x} - z_{\alpha/2} \frac{S}{\sqrt{n}} \leq \mu \leq \bar{x} + z_{\alpha/2} \frac{S}{\sqrt{n}}, \quad (20)$$

em que \bar{x} é a média amostral, $\frac{S}{\sqrt{n}}$ é o erro-padrão amostral e $z_{\alpha/2}$ é o ponto superior com $100(\alpha/2)\%$ da distribuição normal padrão. A aplicação da Equação (20) se mantém independente da forma da distribuição de \bar{X} , mas deve-se considerar $n \geq 40$ para que o resultado do intervalo de confiança seja fidedigno (MONTGOMERY; RUNGER, 2007).

2.3.4 Testes de Hipótese e de Adequação de Ajuste (*Goodness of Fit*)

Um teste de hipóteses estatísticas é um procedimento de tomada de decisão sobre a hipótese acerca de um determinado parâmetro, isto é, se aceita ou rejeita uma afirmação a respeito do mesmo. Dada uma amostra aleatória, testar a hipótese significa calcular uma estatística da variável a partir dos dados amostrais para somente então tomar uma decisão acerca da hipótese nula (MONTGOMERY; RUNGER, 2007).

O teste de hipótese de um modelo estatístico descreve o quão bem este modelo se adequa a uma série de observações. Dessa forma, o teste mostra a discrepância entre

os valores observados e os valores esperados conforme o modelo em questão. Um tipo frequente de teste de hipótese é realizado quando não há conhecimento prévio da distribuição de uma determinada população e deseja-se testar a adequação de uma distribuição particular. O método baseado na distribuição Qui-quadrado foi proposto por K. Pearson (CAMPOS; RÊGO; MENDONÇA, 2017). Para executá-lo, realizam-se n amostras da população, arranjando as mesmas em um histograma de frequências com k intervalos de classe (MONTGOMERY; RUNGER, 2007). Assim, utiliza-se uma medida global para verificar o afastamento global entre a frequência observada O_i para o i -ésimo intervalo de classe e a frequência esperada E_i obtida a partir da distribuição de probabilidade considerada na hipótese. A estatística de teste é dada pela expressão:

$$X_o^2 = \sum_{i=1}^k \frac{(O_i - E_i)^2}{E_i}. \quad (21)$$

Se a distribuição utilizada no teste de hipótese possuir p parâmetros e for adequada para a população, X_o^2 terá uma lei de distribuição aproximada de uma distribuição Qui-quadrado com $k - p - 1$ graus de liberdade, tendo um melhor ajuste à medida que n aumenta (MONTGOMERY; RUNGER, 2007).

2.4 Aspectos de Modelagem e Simulação Estocástica

Os fenômenos do mundo físico tem, por princípio, modelagem analítica complexa. Qualquer teoria desenvolvida terá menor teor de detalhamento que o fenômeno propriamente dito. Se a modelagem é estocástica, inúmeras dependências devem ser consideradas. Todavia, à medida que maior detalhamento das relações de dependência são incluídas nos modelos, mais complicados tornam-se os cálculos, até o ponto de ficarem muito onerosos ou mesmo impossíveis. Desse modo, quanto mais independência é introduzida no modelo probabilístico, maior a chance de realização dos cálculos. Assim, perde-se em qualidade de representação do fenômeno, porém permite-se que o mesmo seja fiel o suficiente a ponto de ser matematicamente tratável (CAMPOS; RÊGO; MENDONÇA, 2017).

A seguir, constam os principais aspectos relacionados à simulação e modelagem dos fenômenos considerados no presente trabalho.

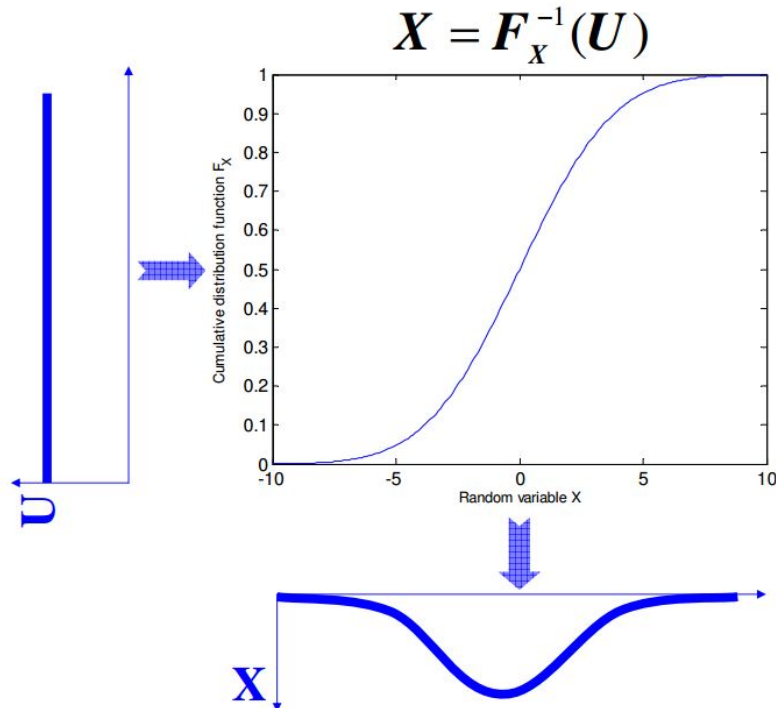
2.4.1 Princípio de Amostragem

O princípio da amostragem em simulações estocásticas é relativamente simples. O conceito é exemplificado na Figura 9, onde números aleatórios de uma distribuição Uniforme $\mathcal{U}(0, 1)$ são utilizados para realizar a amostragem da variável aleatória X com o uso da curva inversa de densidade acumulada de probabilidade $F_X^{-1}(\cdot)$, conforme a Equação:

$$X = F_X^{-1}(\mathcal{U}). \quad (22)$$

A técnica ilustrada na Figura 9 tem aplicação independente da lei de distribuição de $F_X(\cdot)$, desde que a mesma seja inversível. Por conseguinte, dada a distribuição \mathcal{U} e supondo que, porventura, deseja-se alterar a lei de distribuição de F_X , os mesmos números aleatórios \mathcal{U} gerados anteriormente podem ser utilizados para amostrar a variável aleatória X na nova condição (PAPAEFTHYMIU, 2007; RODRIGUES, 2008). Essa técnica também tem a vantagem de proporcionar uma forma universal de sortear amostras para

Figura 9: Procedimento de amostragem de uma VA.



Fonte: (PAPAEFTHYMIU, 2007)

as variáveis aleatórias, independentemente dos valores que estas podem assumir além de que $F_X^{-1}(U)$ será sempre definida uma vez que $0 \leq U \leq 1$ e o contradomínio de F_X é $[0, 1]$ (LAW; KELTON, 2000).

2.4.2 Critério de Parada

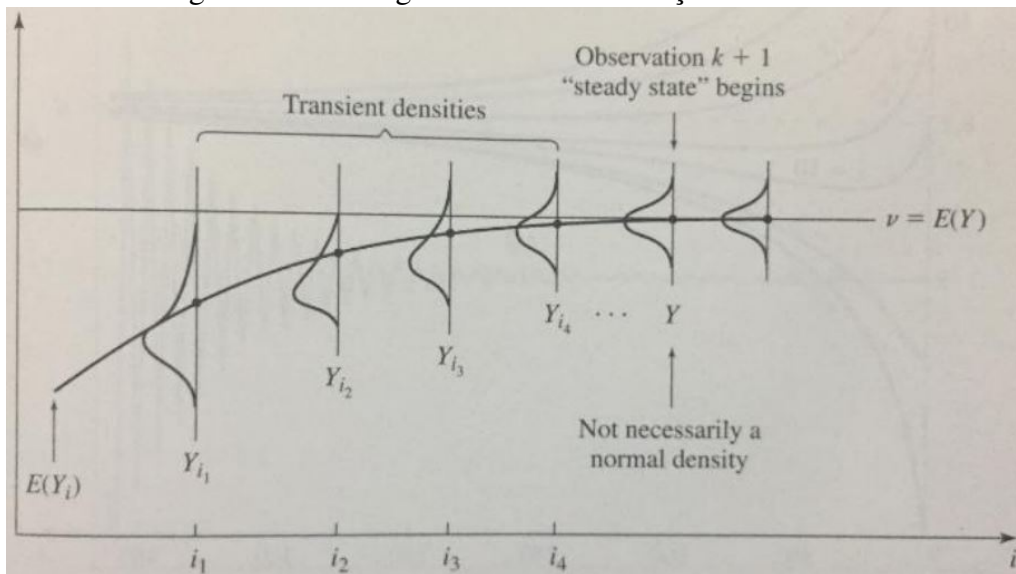
À medida que aumenta a quantidade de amostras das variáveis aleatórias nas simulações estocásticas, é possível observar a convergência de determinadas estatísticas. Estas métricas apresentam densidades transitórias até atingir o regime permanente, instante a partir do qual o acréscimo de amostras não interfere mais de maneira significativa no processo (LAW; KELTON, 2000). A Figura 10 ilustra a convergência de uma simulação estocástica, mostrando também as densidades transitórias da esperança para a variável aleatória Y .

Os critérios de parada da amostragem são baseados na Lei dos Grande Números. Um dos critérios comumente empregados é o coeficiente de variação estatístico (C_v) das variáveis aleatórias em análise, que é dado por:

$$C_v = \frac{s}{\bar{x}\sqrt{n}}, \quad (23)$$

em que s é o desvio-padrão amostral, \bar{x} é a média amostral e n é o número de amostras da variável aleatória X . O valor típico de C_v para estabelecer a convergência é que este seja inferior a 5% (PAREJA, 2009; LAW; KELTON, 2000).

Figura 10: Convergência de uma simulação estocástica.



Fonte: (LAW; KELTON, 2000)

2.4.3 Construção de Estrutura de Dependência Multidimensional para Variáveis Aleatórias

A *Joint Normal Transform* (JNT) é uma técnica para a geração de distribuições Uniformes de probabilidade correlacionadas. Essa técnica é excepcionalmente interessante em simulações estocásticas, pois permite definir a estrutura de dependência estocástica por meio da geração de números aleatórios correlacionados que podem ser utilizados para amostrar as variáveis aleatórias através do princípio de amostragem da Figura 9 (vide Seção 2.4.1).

O procedimento da JNT envolve $(N - 1)N/2$ estruturas de dependência mútua para um sistema com N variáveis aleatórias. O método é dado pelos seguintes passos (PAPAEFTHYMIU, 2007; RODRIGUES, 2008):

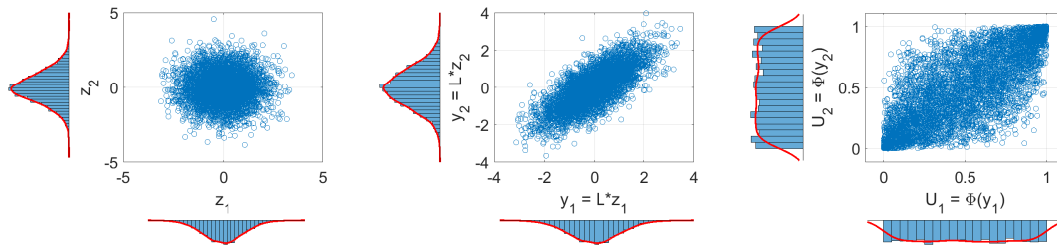
1. Gerar um vetor z de amostras independentes de uma curva normal padrão de probabilidade N -dimensional;
2. Aplicação da transformação linear ao vetor z tal que $y = \mathbf{L} \times z$, sendo \mathbf{L} a matriz triangular inferior da decomposição de Cholesky da matriz \mathbf{R} . A matriz de correlação linear \mathbf{R} é obtida da matriz de correlação de *rank* \mathbf{R}_r pela relação $\mathbf{R} = 2\text{sen}(\mathbf{R}_r\pi/6)$. Os componentes ρ_r da matriz \mathbf{R}_r são a correlação de *rank* entre cada variável aleatória. Assim, os números aleatórios independentes em z passam a estar correlacionados em y pela correlação proveniente da existente entre as variáveis aleatórias;
3. A aplicação da função de distribuição acumulada da normal padrão $\Phi(\cdot)$ aos números aleatórios correlacionados y , resulta em N vetores aleatórios com distribuição Uniforme na forma $u_i = \Phi(y_i)$.

Os números aleatórios correlacionados u_i são utilizados para realizar a amostragem das variáveis aleatórias por meio do modo anteriormente detalhado na Seção 2.4.1. Existe

um requisito matemático de que a matriz de correlação \mathbf{R} seja positiva definida para permitir a decomposição de Cholesky, isto é, para todo $x \in \mathbb{R} \setminus \{0\}$, $x^T \mathbf{R} x > 0$ (PAPAEFTHYMIU, 2007; RODRIGUES, 2008).

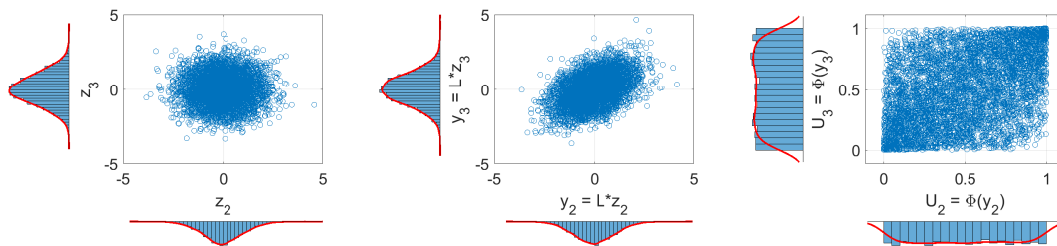
A Figura 11 ilustra os gráficos de dispersão das etapas da JNT para um relação de dependência com $\rho = 0,75$. Observa-se, no gráfico da esquerda que os números aleatórios normalmente distribuídos z_1 e z_2 são independentes. No gráfico central, os números aleatórios obtidos após a transformação linear, y_1 e y_2 , mantém a distribuição normal dos seus geradores (z_1 e z_2), mas evidenciam a correlação positiva, havendo uma concentração de valores no entorno do eixo diagonal. Finalmente, no gráfico da direita, observa-se que os números aleatórios u_1 e u_2 são uniformemente distribuídos e apresentam uma forte correlação positiva. A Figura 12 ilustra as etapas da JNT para uma relação de dependência com $\rho = 0,45$, ou seja, uma relação de dependência menos intensa, sendo esta evidenciada quando se compara os gráficos do centro e da direita, pois os valores do gráfico da esquerda são idênticos nos dois casos.

Figura 11: Ilustração das etapas da JNT para $\rho = 0,75$.



Fonte: O autor.

Figura 12: Ilustração das etapas da JNT para $\rho = 0,45$.



Fonte: O autor.

Fica evidente pela observação da distribuição marginal de probabilidade das distribuições Uniformes obtidas ao final das etapas da JNT, que estas não se alteram apesar da maior ou menor relação de dependência imposta. Assim sendo, o resultado da JNT fornece distribuições uniformes que podem ser utilizadas para amostrar VAs conforme o princípio detalhado na Seção 2.4.1, com a vantagem de considerar a relação de dependência existente entre as VAs do modelo simulado.

2.4.4 *Stochastic Bounds Methodology*: Estruturas de Dependência Extrema entre VAs

Como já mencionado no início da Seção 2.4, há um compromisso entre tratabilidade de um modelo e a qualidade de representação dos fenômenos. No caso de sistemas com

um grande número de variáveis estocásticas de entrada, é necessária a aplicação de alguma técnica que proporcione a redução da ordem da estrutura multidimensional (PAPAETHYMIU, 2007; RODRIGUES, 2008).

Em grande parte dos fenômenos simulados, é possível identificar grupos de variáveis altamente correlacionadas, ou seja, agrupamentos com comportamento e estrutura de dependência observada muito próxima da correlação perfeita. De acordo com a *Stochastic Bounds Methodology* (SBM), todas as estruturas de dependência possíveis para um número de variáveis aleatórias correlacionadas está compreendido entre dois extremos: independência (limite inferior) e comonotonicidade (limite superior). No caso de apenas duas variáveis, o limite inferior para o valor da soma destas variáveis aleatórias corresponde ao caso de correlação perfeita negativa, isto é, a contracomonotonicidade (PAPAETHYMIU, 2007; RODRIGUES, 2008).

2.4.4.1 Limite Superior: Comonotonicidade

Variáveis aleatórias comonotônicas são funções com crescimento baseado no mesmo fator aleatório subjacente. Assim, o mesmo gerador randômico é usado na modelagem do sistema e para a amostragem dessas variáveis (PAPAETHYMIU, 2007; RODRIGUES, 2008). Denotando $(Y_1^C, Y_2^C, \dots, Y_n^C)$ a versão comonotônica do vetor aleatório (Y_1, Y_2, \dots, Y_n) , a amostragem pode ser realizada da seguinte maneira:

$$Y_i^C = F_{Y_i}^{-1}(U), \quad i = \{1, 2, \dots, n\}, \quad (24)$$

em que U é a distribuição Uniforme no intervalo $[0, 1]$ dos números aleatórios sorteados para realizar a amostragem das variáveis aleatórias Y_i a partir de suas respectivas distribuições marginais acumuladas de probabilidade F_{Y_i} .

O valor esperado para $\sum_{i=1}^n Y_i$ e de $\sum_{i=1}^n Y_i^C$ é o mesmo, porém a dispersão nos dois casos é diferente. A amostragem comonotônica oferece o cenário de maior risco uma vez que corresponde ao caso de distribuição extrema da soma das variáveis aleatórias. Esse caso corresponde ao limite C_U da cópula de Fréchet-Hoeffding (vide Seção 2.2.5) que oferece o contorno superior de todas as possíveis cópulas:

$$F_{Y_1, Y_2, \dots, Y_n} = C_U(F_{Y_1}(y_1), F_{Y_2}(y_2), \dots, F_{Y_n}(y_n)) \rightarrow \rho_r(Y_1, Y_2, \dots, Y_n) = 1, \quad (25)$$

em que F_{Y_1, Y_2, \dots, Y_n} é a função de distribuição de probabilidade conjunta para as distribuições acumuladas $F_{Y_1}(y_1), F_{Y_2}(y_2), \dots, F_{Y_n}(y_n)$ das respectivas variáveis aleatórias Y_1, Y_2, \dots, Y_n .

2.4.4.2 Limites Inferiores: Contracomonotonicidade e Independência

No caso de duas variáveis aleatórias, o limite inferior corresponde ao caso de correlação perfeita negativa, isto é, a contracomonotonicidade (PAPAETHYMIU, 2007; RODRIGUES, 2008). Denotando (Y_1^{CM}, Y_2^{CM}) a versão contracomonotônica do vetor aleatório bidimensional (Y_1, Y_2) , a amostragem pode ser realizada da seguinte maneira:

$$Y_1^{CM} = F_{Y_1}^{-1}(U), \quad Y_2^{CM} = F_{Y_2}^{-1}(1 - U), \quad (26)$$

em que U é a distribuição Uniforme no intervalo $[0, 1]$ dos números aleatórios sorteados para realizar a amostragem das variáveis aleatórias Y_1 e Y_2 a partir de suas respectivas distribuições marginais acumuladas de probabilidade F_{Y_1} e F_{Y_2} . Esse caso corresponde

ao limite C_L da cópula de Fréchet-Hoeffding (vide Seção 2.2.5) que oferece o contorno inferior de todas as possíveis cópulas:

$$F_{Y_1, Y_2} = C_L(F_{Y_1}(y_1), F_{Y_2}(y_2)) \rightarrow \rho_r(Y_1, Y_2) = -1, \quad (27)$$

em que F_{Y_1, Y_2} é a função de distribuição de probabilidade conjunta para as distribuições acumuladas $F_{Y_1}(y_1)$ e $F_{Y_2}(y_2)$ das respectivas variáveis aleatórias Y_1 e Y_2 .

A contracomonicidade oferece sempre o contorno inferior para a agregação de duas VAs. Por definição, esse conceito não se aplica ao caso com mais de duas variáveis, já que não é possível haver dependência negativa perfeita para mais de duas variáveis ao mesmo tempo. Nessas situações o contorno inferior é substituído pelo caso de independência (PAPAETHYMIU, 2007; RODRIGUES, 2008). Assim, são usados números aleatórios diferentes na amostragem de cada variável aleatória. Denotando $(Y_1^I, Y_2^I, \dots, Y_n^I)$ como sendo a versão independente do vetor aleatório (Y_1, Y_2, \dots, Y_n) , a amostragem é realizada da seguinte forma:

$$Y_i^I = F_{Y_i}^{-1}(U_i), \quad i = \{1, 2, \dots, n\}, \quad (28)$$

em que U_i é a distribuição Uniforme no intervalo $[0, 1]$ dos números aleatórios sorteados para realizar a amostragem das variáveis aleatórias Y_i a partir de suas respectivas distribuições marginais acumuladas de probabilidade F_{Y_i} .

O caso de independência oferece o cenário de menor risco³ de ocorrência de valores extremos, correspondendo à situação de menor dispersão para a soma das VAs (PAPAETHYMIU, 2007; RODRIGUES, 2008). Todos os outros casos de estruturas de dependência correspondem à somas de VAs, cuja variância estará entre os valores extremos da variância mínima do limite inferior (independência ou contracomonicidade) e da variância máxima do limite superior (comonicidade). O uso de independência para análise de um modelo, cujas variáveis aleatórias são correlacionadas, é uma falácia, uma vez que praticamente todos as ocorrências reais de estrutura de dependência correspondem a situações de impacto mais severo (PAPAETHYMIU, 2007; RODRIGUES, 2008).

A amostragem de acordo com a SBM favorece a redução da ordem do problema a ser tratado, uma vez permite a geração de menor quantidade de distribuições randômicas Uniformes no intervalo $[0, 1]$ para amostragem das variáveis aleatórias. Caso a SBM seja combinada com a JNT por exemplo, a matriz z daquela técnica fica reduzida, diminuindo as operações de todas as etapas subsequentes.

A SBM é um método que favorece a amostragem considerando estruturas de dependência extrema entre variáveis aleatórias baseado no pressuposto de alta correlação entre as mesmas. No caso destas VAs, também é possível tratá-las pelo seu comportamento conjunto, isto é, obtém-se um modelo equivalente de representação para as VAs que são altamente correlacionadas após ter agrupado as mesmas com base em técnicas adequadas.

2.4.5 Técnicas de Agrupamento de Variáveis Aleatórias

A análise dos dados referentes a consumo, demanda, geração de energia dos consumidores/prosumidores, entre outros, permite a classificação dos mesmos em categorias a partir de técnicas específicas para esse fim. O Método das Nuvens Dinâmicas, por exemplo, é um artifício baseado no método K-médias usado para identificar grupos de componentes com comportamento similar, conforme detalhado por Guimarães (2008), sendo possível a partir de um banco de dados com informações dos consumidores/prosumidores

³Para duas VAs o cenário de menor risco é oferecido pela versão contracomotônica.

do sistema em análise. O método das K-médias corresponde a uma técnica de agrupamento onde cada grupo é representado pelo centróide das amostras. Dessa forma, o agrupamento pode ser simplificado conforme o seguinte algoritmo, considerando entradas k centróides, D distâncias, saídas K centróides e objetos de cada fonte:

1. Selecionar arbitrariamente k objetos como os agrupamentos iniciais.
2. Calcular os centróides dos k agrupamentos da posição atual.
3. Associar cada objeto ao grupo (centróide) mais perto (maior similaridade).
4. Retornar ao Passo 2 e parar quando não houver mais mudanças significativas entre os objetos.

Também pode-se dividir os objetos de maneira supervisionada, com base em critérios específicos da preferência do analista. O uso de agrupamentos para clientes de sistemas elétricos também foi empregada por Jardini *et al.* (2000) que utilizou esse artifício para definir curvas de carga típicas para categorias de clientes.

O comparativo entre diferentes técnicas de algoritmos não-supervisionados de divisão em agrupamentos foi investigada por Chicco; Napoli; Piglione (2006). Neste trabalho, os autores investigaram os métodos siga-o-líder modificado, agrupamento hierárquico, K-médias, K-médias fuzzy e mapas auto-organizados concluindo que, em geral, 15 a 20 categorias são suficientes para classificar consumidores de redes de energia elétrica. Além disso, concluem que os algoritmos mais efetivos para tratamento de consumidores destas redes são o siga-o-líder modificado e agrupamento hierárquico, visto que propiciam uma separação altamente detalhada dos agrupamentos, isolando padrões de comportamento incomum e criando grandes grupos contendo os padrões restantes das cargas. Por fim, o estudo de Obrenović; Vidaković; Luković (2017), que inclusive aborda o trabalho de Chicco; Napoli; Piglione (2006), preocupou-se em definir a melhor métrica para o algoritmo das K-Médias++, concluindo que para consumidores de energia elétrica, os melhores índices são obtidos com o uso de distância euclidiana.

3 ANÁLISE DA OPERAÇÃO DE SISTEMAS DE DISTRIBUIÇÃO DE ENERGIA ELÉTRICA

Em geral, a análise da operação de alimentadores de distribuição de energia elétrica consiste no estudo deste sistema sob condições normais de operação em regime permanente (cálculo do fluxo de potência) ou sob condições transitórias (tais como cálculo do curto-circuito ou de indicadores de confiabilidade) (KERSTING, 2007). Neste trabalho, o enfoque é dado à operação sob condições de regime permanente, que pode ser realizada por intermédio de uma abordagem determinística (considerando um valor definido para as cargas e gerações) ou uma abordagem probabilística (considerando as distribuições de probabilidade que caracterizam as cargas e gerações).

3.1 Análise determinística

A análise do fluxo de potência para um alimentador de uma rede de distribuição é similar àquela feita para um sistema de transmissão interconectado, sendo que, tipicamente, as informações conhecidas pré-análise são as tensões trifásicas na subestação, o modelo das cargas e sua potência complexa. Com o cálculo do fluxo de potência é possível determinar (KERSTING, 2007):

- a) módulo e ângulo das tensões em todos os nós do alimentador;
- b) fluxo de potência (kW e kVAr), fator de potência e corrente complexa em cada linha da rede;
- c) perdas de potência em cada ramal;
- d) potência injetada no alimentador em kW e kVAr;
- e) perdas de potência total no alimentador; e
- f) demanda em kW e kVAr com base no modelo considerado das cargas.

Tradicionalmente, a análise de um sistema de distribuição (muitas vezes passivo) depende essencialmente da natureza das cargas. Dependendo do nível de tensão do sistema em análise, a própria definição do que é uma carga pode se alterar. À medida que nos aproximamos dos consumidores da rede secundária, mais pronunciada é a variância da demanda e a ausência de um regime permanente (KERSTING, 2007).

Uma forma de lidar com tal aspecto é considerar a demanda agregada diversificada dos clientes, dada por:

$$D_{\text{div}}(t) = \sum_{k=1}^{NC} D_k(t), \quad (29)$$

em que $D_{\text{div}}(t)$ é a demanda diversificada no tempo t para os NC clientes com demanda individual $D_k(t)$.

Considera-se que cada cliente possui uma curva diária típica e que a soma das demandas individuais de todos os clientes atendidos por um mesmo transformador da rede compõe a chamada curva de demanda diversificada D_{div} daqueles clientes. A máxima demanda diversificada é o maior valor de demanda agregada dos consumidores na Equação (29) e não necessariamente é a soma da máxima demanda individual dos clientes. A soma da máxima demanda individual dos clientes é chamada de máxima demanda não-coincidente, conforme a seguinte expressão:

$$D_{nc} = \sum_{k=1}^{NC} \max(D_k(t)), \quad (30)$$

em que D_{nc} é a máxima demanda não-coincidente para os NC clientes com máxima demanda individual diária $\max(D_k)$.

O fator de diversidade das cargas é dado por:

$$F_{\text{div}} = \frac{D_{nc}}{\max(D_{\text{div}}(t))}. \quad (31)$$

A ideia intrínseca ao fator de diversidade é que, uma vez que a demanda máxima individual seja conhecida, pode-se estimar a máxima demanda diversificada. O fator de diversidade costuma aumentar com o número de consumidores com uma característica de saturação, i.e., a partir de um determinado número de consumidores, o fator de diversidade tende a estabilizar em um determinado valor (KERSTING, 2007). Isso demonstra que quanto menos consumidores são considerados, maior é a influência individual na demanda agregada e que quanto menor for o fator de diversidade, maior é a similaridade no padrão de consumo (relação de dependência alta).

Os conceitos supracitados são típicos de uma análise determinística, pois consideram cenários de carga típica fixa. Esse tipo de abordagem é utilizado nas normas dos Procedimentos de Distribuição de Energia Elétrica no Sistema Elétrico Nacional (PRODIST). O planejamento da expansão do sistema de distribuição é normatizado no Módulo 2 do PRODIST. Nele, as unidades consumidoras são separadas por nível de tensão como sendo pertencentes ao Sistema de Distribuição de Alta Tensão (SDAT, acima de 44 kV), ao Sistema de Distribuição de Média Tensão (SDMT, acima de 1 kV a 44 kV) e ao Sistema de Distribuição de Baixa Tensão SDBT, inferior ou igual a 1 kV). Com relação aos procedimentos para o SDBT, a normativa se limita a estabelecer que o planejamento do mesmo deve definir um plano de obras para um horizonte de cinco anos, com vistas à adequação do sistema existente ao atendimento de novas cargas e à melhora das condições operativas. No entanto, não é definido de que maneira tal objetivo deve ser alcançado (ANEEL, 2016).

O planejamento da rede deve partir do diagnóstico da rede existente, com a caracterização da carga e da rede bem como a avaliação da qualidade e das perdas técnicas. Para o

SDAT e para o SDMT é estabelecido que deve-se considerar patamares de carga (ANEEL, 2016). O Submódulo 5.6 dos Procedimentos de Rede define intervalos horários para os patamares de carga: para a situação sem horário de verão, a carga leve ocorre das 00h às 07h, a carga pesada das 18h às 21h e a carga média nos demais intervalos horários (ONS, 2017). Os patamares obtidos a partir dessa definição podem ser visualizados na Figura 13(b) como resultado da aplicação da definição à curva de carga da Figura 13(a).

Uma forma alternativa de obter os patamares de carga leve, média e pesada se dá a partir da curva de duração: considera-se que a carga pesada ocorra durante três horas e a carga leve durante oito horas, conforme quantidade de horas estabelecida para o horário de ponta e de incentivo às atividades de irrigação e aquicultura justamente em função do baixo consumo (ANEEL, 2010). Sendo a curva de carga a demanda diversificada de todos os clientes atendidos em uma determinada rede, a definição de patamares a partir da curva de duração representa melhor os valores de demanda real do sistema, já que a curva de carga pode apresentar valores de demanda incompatíveis com o intervalo pré-definido de patamar para o qual a respectiva demanda ocorre. Ou seja, uma demanda com magnitude elevada pode ser tratada como carga média simplesmente por ocorrer em um horário pré-definido para tal patamar. Além disso, a curva de duração permite observar a necessidade de ampliação da rede ou mesmo de troca de equipamentos como transformadores devido a sobrecarga, por exemplo (KERSTING, 2007). Sendo assim, essa forma de obtenção dos patamares é empregada neste trabalho. A Figura 13(c) ilustra os patamares obtidos a partir dessa maneira, e é possível constatar que o patamar de carga pesada é consideravelmente superior ao obtido a partir da definição do ONS (2017).

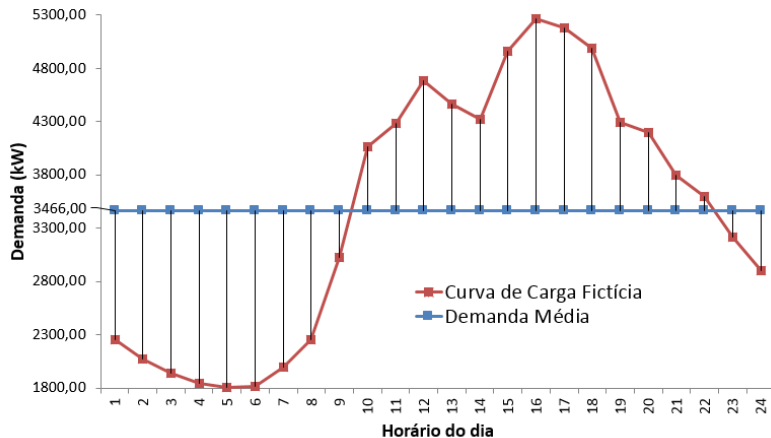
Muitas vezes a demanda complexa agregada (kW e kVAr) é conhecida devido à medição na subestação. Uma maneira de simular o fluxo de potência no sistema é igualar a demanda medida com a calculada a partir dos modelos considerados para os elementos da rede. A razão entre a demanda medida e a demanda calculada pode ser utilizada para multiplicar as demandas das cargas para, de forma iterativa, igualar as duas até atingir uma determinada tolerância. Ou seja, a carga é alocada ao longo da rede de forma proporcional e compatível com o patamar de demanda agregada observado no medidor da subestação (KERSTING, 2007).

Desse modo, a demanda dos clientes é estimada pela divisão proporcional da demanda agregada para cada patamar de consumo considerando os modelos das cargas obtido a partir das campanhas de medição. O cálculo do fluxo de potência é realizado para cada cenário de cargas e os respectivos indicadores de operação como perdas diárias e índices de violação de tensão para o sistema são obtidos pela consideração proporcional ao tempo de permanência em cada cenário de cargas.

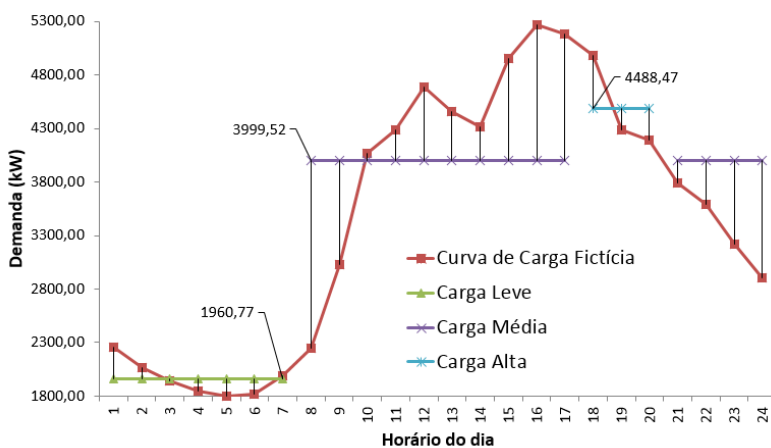
3.2 Análise probabilística

As cargas de um sistema de distribuição de energia elétrica podem ser representadas por valores de demanda fixos ou provenientes de uma curva de carga que varia ao longo de períodos pré-estabelecidos de tempo. Com a mudança nos padrões de consumo em decorrência de tarifas dinâmicas, o acesso à geração distribuída e a tecnologias de automação, entre outros aspectos, os patamares de demanda de um perfil de consumo podem incluir graus de incerteza, o que requer a transição de modelagens tradicionais determinísticas para probabilísticas (CHEN et al., 2018). Para exemplificar a diferença entre as duas abordagens, a Figura 14 ilustra uma curva de carga determinística e uma probabilística. Basicamente, o gráfico superior da Figura 14 ilustra uma curva de carga tradicional, com

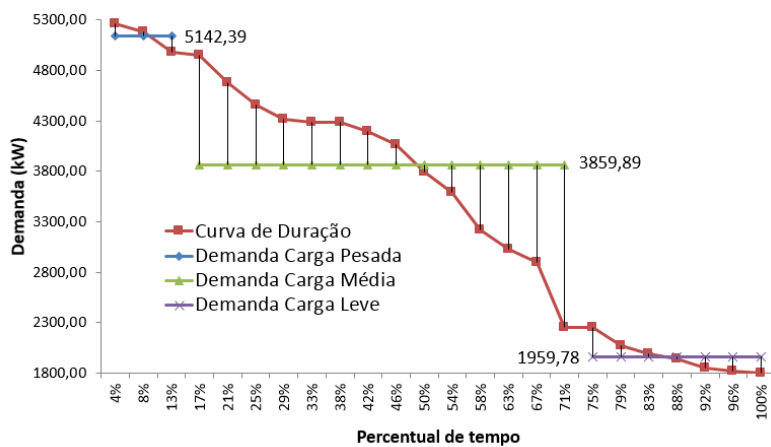
Figura 13: Obtenção de patamares de demanda agregada.



(a) Exemplo de curva de carga diária.



(b) Definição de patamares conforme ONS (2017).

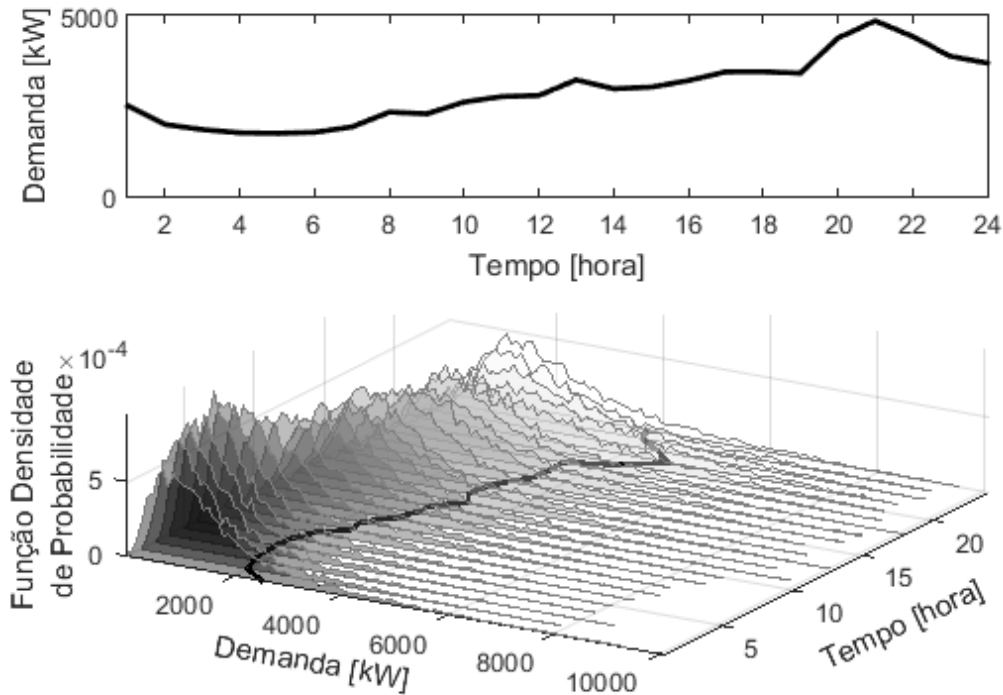


(c) Definição de patamares a partir da curva de duração.

Fonte: O autor.

24 valores de demanda, um para cada hora de um dia típico. Já o gráfico inferior, ilustra a densidade de probabilidade da demanda em cada horário ao longo do dia, incluindo o caso particular determinístico.

Figura 14: Curva de carga determinística versus probabilística.



Fonte: O autor.

Métodos determinísticos consideram que os parâmetros de entrada do sistema são constantes para cada intervalo de tempo (MOHAMMADI; SHAYEGANI; ADAMINE-JAD, 2013). Assim sendo, para a potência ativa e reativa das barras PQ, bem como para as magnitudes de tensão e a potência ativa das barras PV pode-se adotar o valor médio ao longo de período de tempo ou o valor instantâneo do fluxo de potência (SAMET; KHORSHIDSAVAR, 2018). Conforme ilustra a Figura 14, a curva de carga determinística é elaborada a partir de valores de demanda média para cada hora do dia. No entanto, a incerteza das cargas é condicionada no tempo, sendo possível demonstrar em alguns casos que o desvio padrão da demanda tem magnitude maior que seu valor médio (PAREJA, 2009). Logo a média torna-se insuficiente para descrever o comportamento dos componentes do sistema, sendo necessário considerar intervalos de confiança para as grandezas de interesse.

Tendo reconhecido a necessidade de incorporar medidas de incertezas nos modelos matemáticos visando caracterizar o comportamento dinâmico de cargas e geração, pesquisadores têm investigado a análise probabilística do fluxo de carga há algumas décadas (BORKOWSKA, 1974; HONG; FAN, 2016). Dentre os métodos nos quais se considera o fenômeno da incerteza dos sistemas elétricos de potência, a Simulação de Monte Carlo (SMC) é o mais conhecido (PAREJA, 2009).

Monte Carlo é a abordagem tradicional utilizada para simular o fluxo de carga probabilístico e executa o fluxo de carga determinístico repetidamente por um número significativo de vezes para representar a distribuição total das entradas do sistema (KABIR;

MISHRA; BANSAL, 2016). Na SMC, todas as variáveis de interesse são representadas por funções de densidade de probabilidade (CONTI; RAITI, 2007). Essa representação das cargas permite atribuir um intervalo de confiança a cada demanda horária de uma curva de carga (MOHAMMADI; SHAYEGANI; ADAMINEJAD, 2013). Assim, por meio de números aleatórios, os possíveis estados de operação do sistema são obtidos e a cada vez que se gera um estado é necessário executar um fluxo de potência determinístico para estabelecer as condições de operação em regime permanente referentes a este estado (PAREJA, 2009; HONG; FAN, 2016). Este método é bastante preciso e os seus resultados são frequentemente utilizados como referência na comparação com outras técnicas estocásticas e também no planejamento da operação e da expansão de sistemas elétricos de potência (HONG; FAN, 2016; CONTI; RAITI, 2007; CHEN et al., 2018).

3.2.1 Principais trabalhos investigados

Dentre as incertezas que existem nos sistemas de distribuição já mencionadas no Capítulo 1, pode-se considerar a demanda das unidades consumidoras como uma das principais. Assim, o desenvolvimento de um modelo que permita avaliar esse aspecto é essencial na análise da operação de redes de distribuição.

Um fluxo de potência alternativo para redes de distribuição de energia elétrica considerando incerteza na demanda dos consumidores da rede de baixa tensão bem como a incerteza na fase de conexão dos mesmos é proposto por Pareja (2009). Para tal, primeiramente é proposta a divisão dos clientes em categorias pré-definidas com caracterização através de curvas de carga típicas obtidas por campanhas de medição.

A determinação da curva de carga de consumidores individuais e nos transformadores de distribuição utilizada por Pareja (2009) considerou curvas representativas para grupos de clientes conforme preconizado por Jardini *et al.* (2000). As categorias de clientes foram pré-definidas com base nas estratificações determinadas no Módulo 2 do PRODIST para a época da elaboração do trabalho. As curvas de média e desvio padrão da demanda em kW dos consumidores obtidas na campanha de medição são normalizadas de forma que seja possível agrupá-las e obter curvas representativas para os agrupamentos. A potência de base utilizada é dada por:

$$P_{\text{base}} = \frac{1}{24} \sum_{t=1}^{24} M(t) \Delta t = \frac{\text{consumo mensal em kWh}}{24 \times 30}, \quad (32)$$

em que $M(t)$ é a curva de demanda média diária em kW.

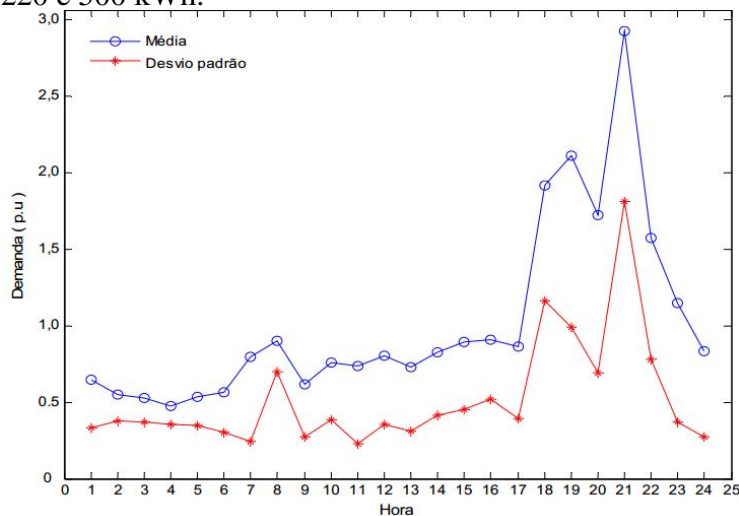
Cabe ressaltar que P_{base} da Equação (32) é diferente da potência aparente de base (S_{base}) utilizada na normalização das grandezas do circuito para cálculo do fluxo de potência. A P_{base} é específica de cada cliente e é utilizada apenas para normalizar as curvas da média e do desvio padrão de sua demanda.

Desse modo, são obtidas curvas representativas normalizadas em p.u. para cada subclasse de consumo. A Figura 15 ilustra um exemplo da representação normalizada da demanda diária típica para os clientes com consumo mensal entre 220 e 500 kWh.

De forma semelhante, com a medição de energia mensal de cada consumidor, pode-se estimar a curva de carga diária em kW do mesmo com base na curva representativa em p.u. associada a uma potência de base P_{base} calculada de acordo com a Equação (32) (JARDINI et al., 2000).

Segundo Pareja (2009), a lognormal é a distribuição de probabilidade que apresentou a melhor adequação de ajuste à demanda horária para maior número de clientes de uma

Figura 15: Curvas representativas em p.u. para dias úteis de usuários residenciais com consumo entre 220 e 500 kWh.



Fonte: (PAREJA, 2009).

rede de distribuição, conforme estudo estatístico dos dados de curvas diárias de demanda obtidas em uma campanha de medição. Além disso, Pareja (2009) mostrou que essa distribuição, com raras exceções, foi a que melhor se adequava aos dados dos clientes, independente de classe ou faixa de consumo.

Os testes foram realizados para dias úteis, sábados, domingos e feriados. Assim, diferentemente de Jardini *et al.* (2000) que assume uma distribuição normal para a demanda dos consumidores, na concepção de Pareja (2009) deve-se realizar uma análise de adequação de ajuste aos dados medidos de demanda (vide Seção 2.3.4) e utilizar a função de distribuição de probabilidade que melhor represente os dados da campanha de medição para modelar a demanda horária dos consumidores.

Desse modo, conhecendo a lei de distribuição que teve melhor ajuste aos dados medidos, tendo definida a curva típica de média e desvio padrão da demanda de cada categoria de consumidores bem como o consumo típico mensal para obter a potência de base P_{base} com a Equação (32), fica definida a modelagem da demanda de todos os consumidores da rede que será empregada neste trabalho.

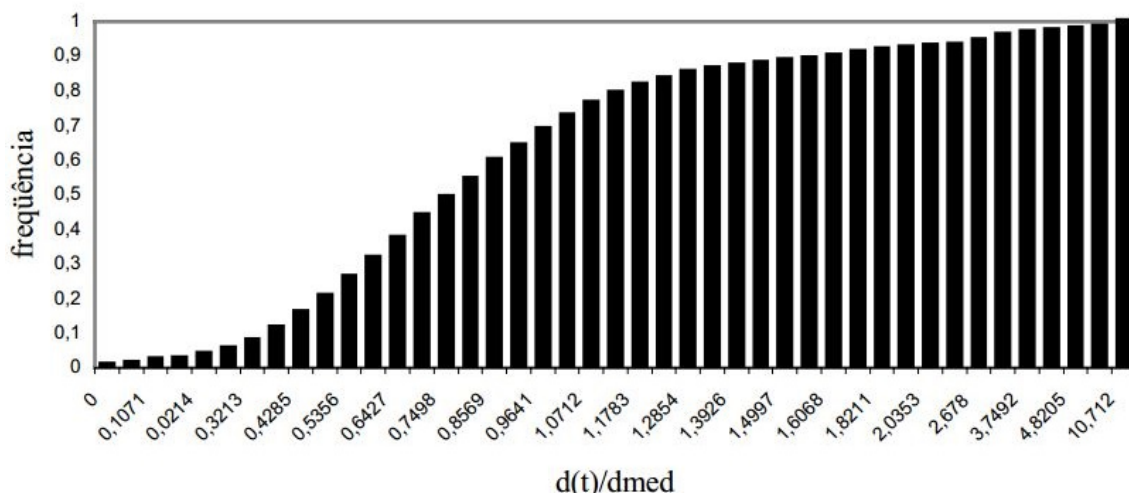
Em seu trabalho, Pareja (2009) propõe o uso da SMC para avaliar as incertezas da rede, de forma que o fluxo de potência é realizado hora a hora para a curva de carga, i.e., existe um problema para cada hora do dia e os mesmos são resolvidos de forma independente. Com o modelo da demanda horária de cada consumidor, a composição dos cenários de cargas para cada horário é realizada conforme o princípio de amostragem da Seção 2.4.1. A demanda de cada cliente é sorteada de forma independente e o algoritmo realiza fluxos de potência determinísticos até que o coeficiente de variação estatístico da Equação (23) seja inferior a uma tolerância especificada. As variáveis de interesse consideradas por Pareja (2009) para determinar a parada das simulações foram as demandas de potência em cada fase do sistema de distribuição. Além disso, o método não prevê e também no estudo de caso do trabalho em questão não é considerada a influência da GD.

Agrupar os clientes é interessante para as simulações pois reduz o tamanho do problema a ser resolvido. Embora a definição dos agrupamentos proposta por Pareja (2009) seja conveniente uma vez que as informações provenientes para a divisão já são naturalmente disponíveis pela coleta de dados das campanhas de medição promovidas conforme

a regularidade normativa estabelecida por ANEEL (2016), essa divisão pode não ser a mais adequada. Nesse aspecto, Guimarães (2008) procede de maneira distinta, valendo-se do Método das Nuvens Dinâmicas para definir os agrupamentos (vide Seção 2.4.5).

Para modelar a demanda das cargas, Guimarães (2008) propõe a utilização da curva de frequência acumulada obtida nas medições para cada agrupamento, normalizando a mesma pela demanda média para cada horário do dia. A Figura 16 ilustra o exemplo de uma curva de frequência acumulada para um consumidor residencial.

Figura 16: Curva de frequência acumulada da demanda relativa das 20h às 21h para um cliente residencial em BT com consumo mensal entre 200 e 400 kWh.



Fonte: (GUIMARÃES, 2008).

Cada curva típica está associada a um fator de carga e a uma participação de mercado das respectivas categorias de clientes. Após obtidas as curvas características, a obtenção da amostra de demanda que compõe o cenário das cargas a ser simulado por Guimarães (2008) é ilustrado na Figura 17. O procedimento de amostragem é similar ao utilizado por Pareja (2009) e descrito com maior detalhe na Seção 2.4.1, porém ao invés de sortear um valor de demanda absoluta em kW, a amostra é da demanda relativa e a conversão para um valor de demanda em kW para compor o cenário de cargas a ser simulado é dado por (GUIMARÃES, 2008):

$$x' = F^{-1}(y'), \quad (33)$$

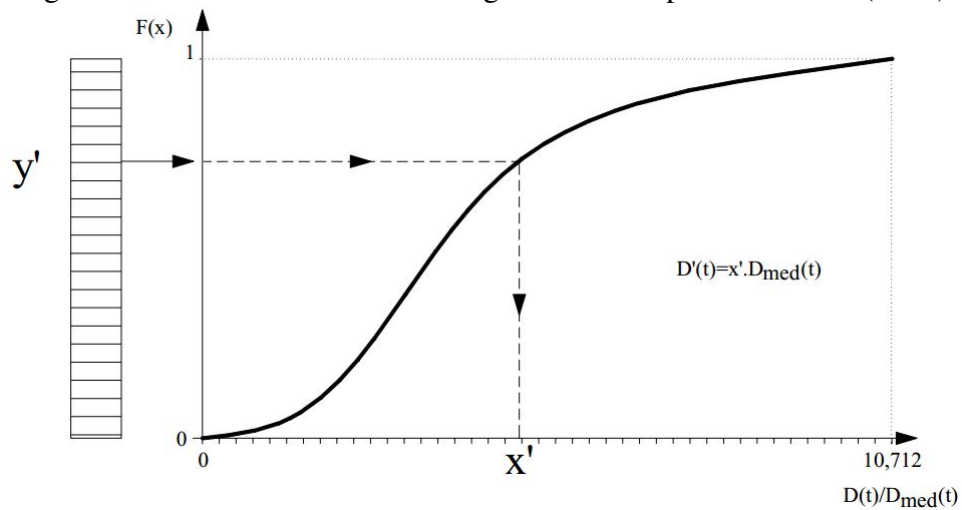
$$D(t) = x' \times D_{\text{med}}(t), \quad (34)$$

em que x' é a amostra da demanda relativa, $F^{-1}(\cdot)$ é a função inversa de distribuição de probabilidade acumulada, y' é o número aleatório sorteado para realizar a amostra, $D(t)$ é a demanda em kW amostrada no horário t e $D_{\text{med}}(t)$ é a demanda média no horário t para o cliente cuja demanda está sendo amostrada.

Guimarães (2008) realiza a simulação de um número pré-fixado de cenários de cargas para cada horário da curva de carga diária e as tensões resultantes são classificadas em adequada, crítica e precária (ANEEL, 2018). Com isso, obtém-se uma distribuição de frequência relativa para as tensões em cada barra do sistema conforme a Figura 18.

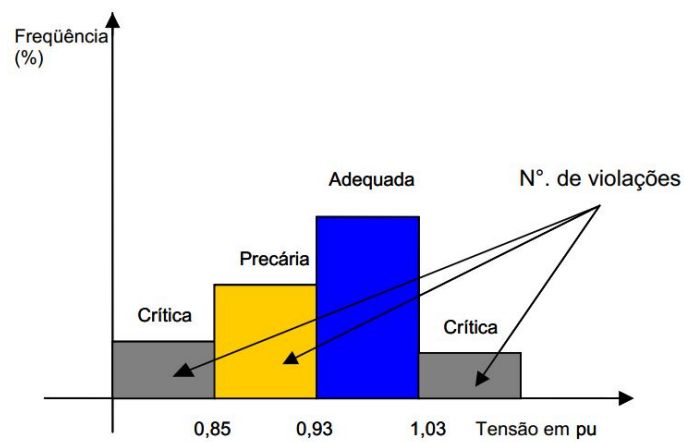
Desse modo, os índices de violação de tensão precária e crítica são estimados por (GUI-

Figura 17: Procedimento de amostragem utilizado por Guimarães (2008).



Fonte: (GUIMARÃES, 2008).

Figura 18: Histograma de frequência relativa das tensões de atendimento.



Fonte: (GUIMARÃES, 2008).

MARÃES, 2008):

$$DRP_{\text{probabilístico}} = \frac{ntp}{nts} \times 100[\%], \quad (35)$$

$$DRC_{\text{probabilístico}} = \frac{ntc}{nts} \times 100[\%], \quad (36)$$

em que ntp é o número de tensões resultantes de simulação com valores situados na faixa precária, ntc é o número de tensões resultantes de simulação com valores situados na faixa crítica, nts é o número total de tensões resultantes da simulação, $DRP_{\text{probabilístico}}$ e $DRC_{\text{probabilístico}}$ são os resultados probabilísticos para a duração relativa de transgressão de tensão precária e crítica respectivamente (ANEEL, 2018).

Os métodos de avaliação probabilística propostos por Pareja (2009) e Guimarães (2008) apresentam modelagem eficiente para as demandas horárias dos clientes quando observa-se o comportamento individual dos mesmos. Considera-se independência entre os consumidores, i.e., a amostragem da demanda de cada cliente não tem relação com os demais. Além disso, não é considerada a inserção de fontes distribuídas intermitentes de geração de energia, realidade que começa a ter maior projeção à medida que as normas de implantação são aperfeiçoadas e os custos reduzidos (MILLER; YE, 2003; ANEEL, 2012).

Com relação a esse aspecto, o trabalho desenvolvido por Papaefthymiou (2007) e explorado por Rodrigues (2008) são de grande relevância. Os autores propõem um método para a consideração de relação de dependência horária entre a demanda dos consumidores bem como a sua relação com a geração distribuída intermitente.

O estudo realizado por Papaefthymiou (2007) é pioneiro, visto que, de acordo com o próprio autor, foi a primeira vez que a solução do problema de modelagem dos recursos estocásticos correlacionados no sistema de potência havia sido investigado. Para tratar as incerteza dos sistemas, Papaefthymiou (2007) é categórico em afirmar que o método mais apropriado é a SMC em função da existência de interdependências complexas entre as variáveis de entrada.

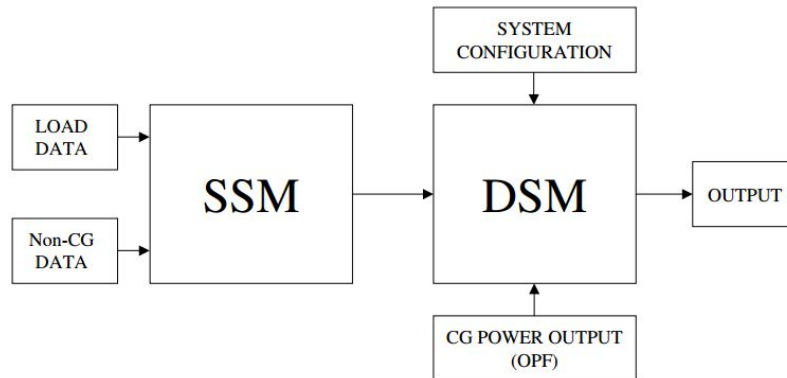
De acordo com Papaefthymiou (2007), o sistema de potência está predominantemente em uma condição operativa que pode ser considerada como regime permanente e que a operação pode ser vista como a transição consecutiva entre estados que representam um cenário das entradas (cargas e gerações). Portanto, a investigação da operação do sistema é a análise do regime permanente para todas as combinações possíveis das entradas. Contudo, ao afirmar isso, deve-se considerar que as simulações de Papaefthymiou (2007) são focadas na média tensão. Caso seja de interesse avaliar a operação na baixa tensão, esse cenário não necessariamente é mantido.

A modelagem é dividida em duas etapas: o modelo determinístico do sistema (DSM¹) e o modelo estocástico do sistema (SSM²). O DSM representa a maneira como as variáveis do sistema são combinadas para produzir as saídas do mesmo, referindo-se a todas as limitações físicas e ações de controle que são executadas durante a operação do sistema. Trata-se do modelo de regime permanente. O SSM descreve o comportamento e a interação entre as entradas incertas do sistema e está relacionado aos mecanismos externos que afetam o comportamento do sistema como padrão de consumo dos clientes, condições climáticas, entre outros. O diagrama esquemático da análise proposta por Papaefthymiou (2007) está mostrado na Figura 19.

¹Do inglês *Deterministic System Model*

²Do inglês *Stochastic System Model*

Figura 19: Diagrama esquemático da análise do sistema.



Fonte: (PAPAEFTHYMIU, 2007).

Os dados de consumidores bem como da geração não-convencional (distribuída), i.e., as entradas com incerteza dos sistema, são fornecidas ao SSM contemplando todo o conjunto de entradas estocásticas. Após tratadas as informações, o DSM recebe as combinações das entradas estocásticas bem como a configuração do sistema e as de geração convencional não-estocástica (obtidas por um algoritmo de fluxo de potência ótimo³) de forma que os dados de saída sejam calculados.

O DSM é a modelagem tradicional utilizada na análise do fluxo de potência. O SSM pode ainda ser dividido em duas etapas: a modelagem do comportamento individual e a da estrutura de dependência conjunta (PAPAEFTHYMIU, 2007). A descrição do comportamento individual é semelhante ao que propuseram Pareja (2009) e Guimarães (2008). No entanto, Papaefthymiou (2007) e Rodrigues (2008) optam por definir uma curva de distribuição acumulada de probabilidade para os barramentos ao invés de fazê-lo para categorias de consumidores, ou seja, a CDF é referente à demanda agregada na barra e, portanto, consideram o sistema AT/MT.

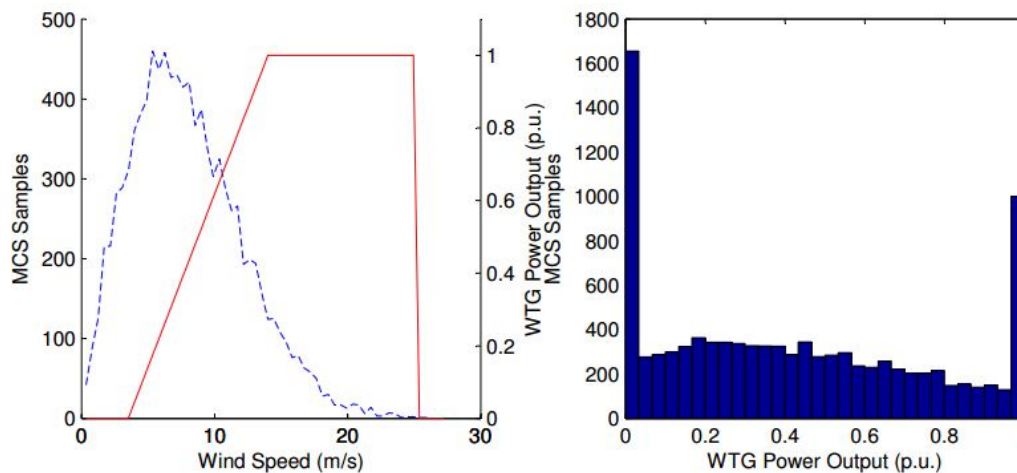
Com relação aos empreendimentos de geração distribuída, Papaefthymiou (2007) os classifica como geradores estocásticos que convertem a energia de uma fonte primária incontroleável em energia elétrica. A potência de saída dos geradores estocásticos depende de dois fatores: a aleatoriedade da fonte motriz primária e do sistema de conversão, i.e., da tecnologia do conversor e das limitações que o mesmo impõe. Isso pode resultar em curvas de distribuição acumulada de probabilidade para a potência gerada não padronizadas. Para uma GD de fonte eólica, por exemplo, a fonte motriz primária (vento) segue uma distribuição de Weibull. O sistema de conversão é modelado por uma relação entre velocidade de vento e potência gerada considerando que há velocidades de vento de *cut-in*, nominal e de *cut-out*, conforme ilustra a Figura 20.

Para o exemplo da Figura 20, velocidades de vento inferiores à *cut-in* e velocidades superiores à *cut-out* resultarão em potência de saída nula, conferindo uma densidade de probabilidade elevada para potência mínima. Da mesma forma, ao atingir a velocidade nominal, o gerador ajusta a posição das pás para que a potência não exceda a máxima, e isso confere uma densidade de probabilidade diferenciada próximo à 1 p.u. (PAPAEFTHYMIU, 2007).

Para modelagem da estrutura de interdependência, Papaefthymiou (2007) utiliza a *Joint Normal Transform* (vide Seção 2.4.3) e a *Stochastic Bounds Methodology* (vide Seção 2.4.4). Desse modo, os comportamentos individuais permanecem inalterados e a

³OPF - *Optimal Power Flow*

Figura 20: Relação de conversão de potência de um gerador eólico e PDF da potência gerada.



Fonte: (PAPAEFTHYMIYOU, 2007).

relação de dependência é inserida através do uso de um mecanismo de amostragem com números aleatórios correlacionados.

3.3 Modelo de resposta das cargas à variação de tensão

Um aspecto importante na análise de operação de sistemas de distribuição é a consideração do modelo adequado de resposta da carga à variação de tensão, com destaque para os modelos de impedância, corrente e potência constante ou mesmo uma combinação das anteriores, o chamado modelo ZIP, dentre outros (SHORT, 2003; GUIMARÃES, 2008; RANGEL, 2015).

Existem divergências na literatura com relação a quais modelos devem ser utilizados e os respectivos coeficientes adotados para cargas de um sistema de distribuição. Contudo, conforme demonstra Rangel (2015), o modelo de resposta da carga à variação de tensão influencia os resultados do fluxo de potência. Assim, a definição adequada destes parâmetros é primordial para a correta análise da operação. Como mencionado no trabalho de Rangel (2015), frequentemente as concessionárias de distribuição adotam para a rede de distribuição de energia elétrica a carga representada por 50% de impedância constante e 50% de potência constante, sem mencionar o horário para o qual esses valores são utilizados.

Milanović *et al.* (2014) apresenta os resultados de um trabalho investigativo com distribuidoras de todos os continentes. Um dos objetivos da pesquisa era identificar práticas comuns na consideração dos modelos de resposta à variação da tensão e também os coeficientes utilizados. Na grande maioria dos casos, o mesmo modelo e também os coeficientes eram utilizados ao longo de toda a rede simulada, em grande parte devido à dificuldade existente em obter informações acuradas acerca das classes de cargas em diferentes nós do sistema. Os resultados mostram que não há uma prática amplamente adotada e que cada distribuidora emprega os modelos que julga mais apropriados para sua área de atuação.

Short (2003) afirma que as cargas da maioria dos circuitos de distribuição podem ser modeladas com algo entre 40 a 60 % de potência constante e 40 a 60% de impedância constante. Já ANEEL (2015) estabelece que o modelo adotado para o cálculo de perdas

na distribuição deva ser o modelo ZIP composto por 100% de impedância constante para a parcela reativa e de 50% potência constante e 50% impedância constante para a parcela ativa da carga.

Percebe-se que, embora existam vários estudos que procuram estabelecer uma prática definitiva, o mais adequado é investigar os hábitos dos consumidores do sistema que será analisado. Caso uma campanha de medição não seja viável, a partir de modelos de resposta da carga e os respectivos parâmetros para cada tipo de utensílio, como realizado por Haffner *et al.* (2009), combinados com a informação da posse de equipamentos e hábitos de uso dos consumidores, como os divulgados por ELETROBRAS (2006), é possível estimar os parâmetros equivalentes para a rede simulada.

4 MÉTODO PROPOSTO PARA ANÁLISE PROBABILÍSTICA DAS INCERTEZAS

O método proposto nesta dissertação para análise probabilística de redes de distribuição é um método numérico de amostragem aleatória simples baseado na SMC e inspirado nas propostas de Papaefthymiou (2007) e Pareja (2009). A carga é modelada por Pareja (2009) conforme tipo de consumidor, permitindo melhor representar as peculiaridades inerentes a cada cliente bem como tornar mais flexível as simulações para diferentes níveis de tensão. Já Papaefthymiou (2007) propõe uma representação da carga concentrada nas barras do sistema além de incluir GD eólica e considerar as relações de dependência entre consumidores e geração.

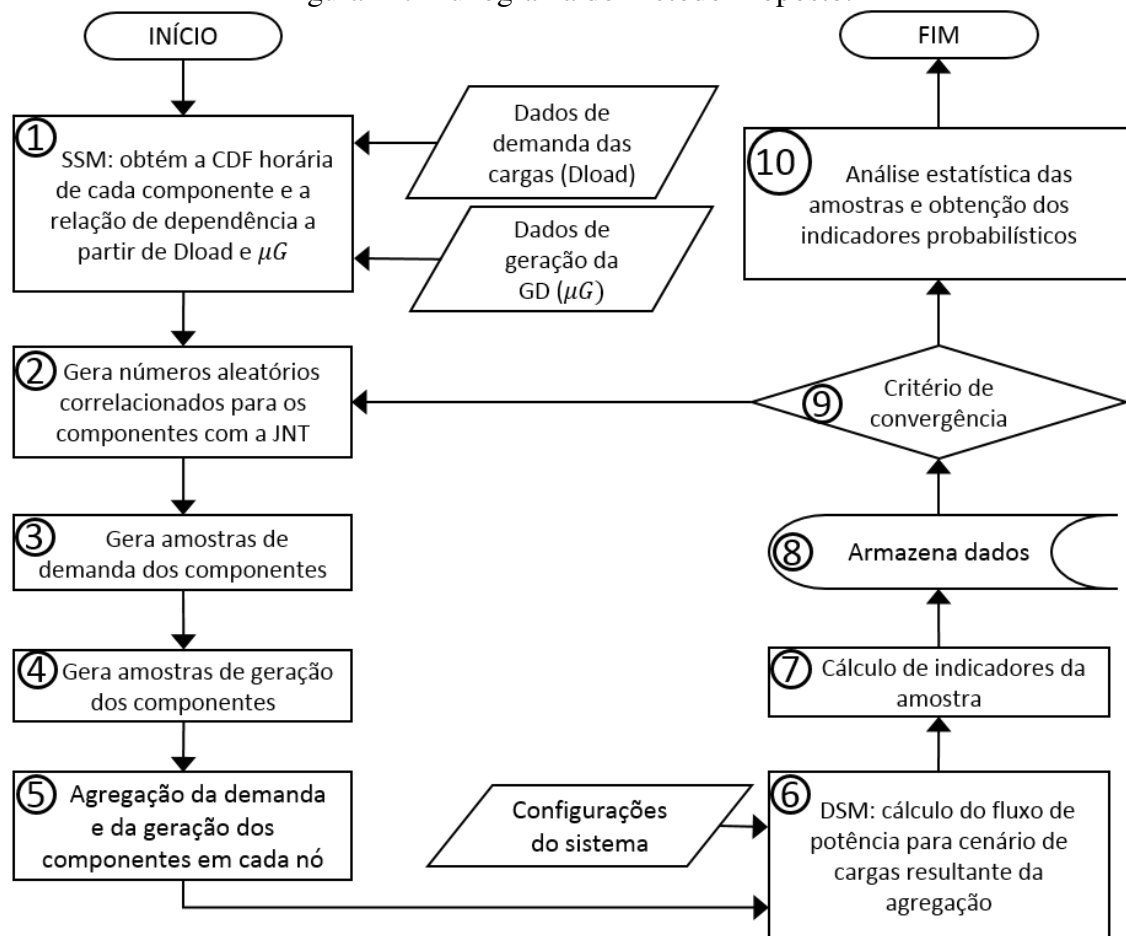
Neste trabalho, considera-se uma representação da carga conforme tipo de consumidor, semelhante ao que propõe Pareja (2009). Desse modo, há possibilidade de simular sistemas em diferentes níveis de tensão pela simples agregação das cargas nos nós da rede. Também foi incluída a relação de dependência entre consumidores e geração, de acordo com o trabalho de Papaefthymiou (2007), sendo considerada, no entanto, a geração fotovoltaica devido à sua maior penetração (ANEEL, 2017a). O método proposto neste trabalho é sintetizado pelo fluxograma da Figura 21.

No passo 1 é obtido o modelo estocástico individual dos componentes do sistema e a relação de dependência entre os mesmos. A partir dos dados das cargas e da GD, são definidos os agrupamentos de componentes com padrão de comportamento semelhante. Na definição dos agrupamentos, a *Stochastic Bounds Methodology* – SBM é utilizada e considera-se que os componentes de um mesmo agrupamento seguem a aleatoriedade da mesma fonte estocástica subjacente. Assim, pode-se prosseguir com a obtenção de curvas diárias normalizadas de média e desvio padrão da potência dos agrupamentos bem como a lei de distribuição de probabilidade que rege o comportamento horário dos mesmos. Cumprida essa etapa, é possível obter a curva CDF da potência dos agrupamentos para cada horário t do período avaliado. O último procedimento do Passo 1 é a obtenção da correlação de rank ρ_r entre todos os agrupamentos para a composição da matriz de correlação \mathbf{R}_t , ou seja, a matriz que contém a essência das relações de interdependência entre os agrupamentos em cada horário t .

No Passo 2, é utilizada a informação da matriz de correlação \mathbf{R}_t para sortear o n -ésimo número aleatório uniforme correlacionado $u_{GP,n}(t)$ para cada agrupamento GP em cada horário t com o uso da técnica da *Joint Normal Transform* – JNT.

A amostra de demanda dos consumidores é realizada no Passo 3 enquanto a amostra da geração dos agrupamentos de GD é realizada no Passo 4 através da curva CDF inversa de cada componente e dos respectivos números aleatórios $u_{GP,n}(t)$ pelo uso da Equação (22).

Figura 21: Fluxograma do Método Proposto.



Fonte: O autor

Com as amostras de potência de cada componente do sistema, procede-se à agregação da potência nos nós para compor o cenário a ser simulado em cada horário t , caracterizando uma amostra N nas simulações. Esse procedimento é realizado no Passo 5.

O passo 6 contempla o modelo determinístico do sistema, ou seja, as configurações do sistema para as quais as incertezas não são consideradas como a topologia e dados de impedância das linhas e equipamentos, conexão de bancos de capacitores, atuação dos reguladores de tensão entre outros. O cálculo do fluxo de potência para cada um dos t cenários da amostra N gerados nos Passos anteriores para a potência dos componentes é realizado e os resultados são utilizados para calcular os indicadores de operação no Passo 7, caracterizando uma amostra N de cada indicador.

No Passo 8 os resultados do cálculo do fluxo de potência para os estados da rede e as amostras dos indicadores são armazenados. No Passo 9 o critério de convergência é verificado. Caso a tolerância não seja atingida, retorna-se ao Passo 2 para a realização de uma nova amostra. Tendo atingido a convergência das simulações, no Passo 10 é realizada a análise estatística das amostras que permite obter os indicadores probabilísticos de operação do sistema.

Devido à especificidade, alguns procedimentos contidos nos Passos do fluxograma merecem especial atenção:

- a) modelagem da distribuição marginal de probabilidade para os componentes do sistema;
- b) modelagem da estrutura estocástica multidimensional de dependência;
- c) criação dos cenários utilizando as informações dos Passos 1 e 2 além das técnicas de amostragem em simulações estocásticas; e
- d) cálculo do fluxo de potência e dos parâmetros de operação da rede.

Cada procedimento compreende passos intermediários, que serão detalhados nas Seções subsequentes.

4.1 Modelagem da distribuição marginal de probabilidade para os componentes do sistema

A distribuição marginal de probabilidade é a descrição do comportamento individual de cada componente (consumidor/prosumidor) do sistema de distribuição. Para que se tenha uma descrição completa do comportamento estocástico, é necessário conhecer a média e o desvio padrão da respectiva demanda/geração além da distribuição de probabilidade que rege o comportamento do componente na condição em análise.

O conjunto formado pela distribuição de probabilidade, média e desvio padrão representa o comportamento do componente em uma determinada situação. Caso o interesse seja observar o comportamento em um dia típico, dia útil ou feriado em diferentes épocas do ano, tais parâmetros devem ser obtidos considerando estas peculiaridades, uma vez que são fatores que afetam o comportamento típico. É necessário atentar ainda para hora do dia, sazonalidade, influência geográfica ou outros aspectos que afetam o padrão de comportamento do sistema em análise. Caso algum destes aspectos causem impactos importantes, pode ser necessário segmentar a análise de acordo com estas características, o que por sua vez aumenta a quantidade de informação necessária para a fiel representação do problema.

Neste trabalho, as unidades consumidoras e de geração distribuída (GD) são agrupadas em categorias¹, conforme semelhança no padrão de demanda/geração para a obtenção do comportamento típico. A escolha da técnica de agrupamento dentre as já mencionadas na Seção 2.4.5 e dos critérios específicos para agrupar as cargas e a geração distribuída são opção de cada analista, porém alguns aspectos precisam ser considerados. As cargas podem ser classificadas com base em critérios como classe (residencial, comercial, industrial, etc), demanda máxima, faixa de consumo mensal, entre outros. Já as unidades de GD podem ser agrupadas em função da fonte energética primária, tecnologia de conversão, proximidade geográfica, entre outros. Também podem ser observados aspectos regulatórios na divisão em categorias.

A divisão dos agrupamentos de cargas apresentada por Papaefthymiou (2007) é diferente da utilizada no método proposto neste trabalho. A representação da demanda agregada das cargas diretamente nos nós, regida por distribuições normais de probabilidade, pode ter limitações. Da Lei dos Grandes Números e do Teorema do Limite Central, pode-se inferir que a soma de uma sequência de variáveis aleatórias independentes e identicamente distribuídas tendem a formar uma distribuição normal de probabilidade (MONTGOMERY; RUNGER, 2007; BECKER, 2015; CAMPOS; RÊGO; MENDONÇA, 2017). Essa representação pode ser aproximada caso haja a contribuição de um número grande de consumidores independentes², o que não ocorre nas extremidades da rede de distribuição onde há contribuição de menor quantidade de consumidores, principalmente em redes radiais. Além disso, a característica do tipo de distribuição de probabilidade que rege o comportamento individual dos componentes fica proeminente, i.e., não necessariamente as variáveis seguem a mesma distribuição de probabilidade, portanto não podem ser consideradas identicamente distribuídas.

Sendo assim, com relação a este aspecto, optou-se por seguir o critério de Guimarães (2008) e Pareja (2009) para as cargas, cuja classificação observou a característica dos clientes por classe e faixa de consumo, ao invés de considerar o perfil da carga diretamente nos nós da rede. Esta opção é justificada pela maior flexibilidade proporcionada com esta abordagem, ou seja, independente da proximidade com a carga e com a extremidade do ramal de alimentação, tanto na rede primária como na secundária, é possível simular a operação da rede com uma boa representação do comportamento dos componentes do sistema. Neste aspecto, a representação do método proposto é mais abrangente por ter aplicação a diversas configurações da rede de distribuição visto que, como demonstraram Pareja (2009) e Abdulkarim; Abdelkader; Morrow (2015), não necessariamente os componentes tem distribuição de probabilidade normal para a demanda horária.

Com relação à divisão das unidades de GD ditas não-despacháveis (de comportamento estocástico) em agrupamentos, este trabalho focou na contribuição da GDFV devido à sua maior penetração observada na rede de distribuição (ANEEL, 2017a; ABGD, 2017). No entanto, a mesma lógica de divisão em agrupamentos se aplica à outras fontes energéticas. Sempre que a unidade de GD for não-despachável e não contar com acumuladores de energia, a geração é altamente dependente da aleatoriedade da fonte primária, seja ela irradiação, vento, etc. Assim, é razoável considerar que unidades de GD com localização geográfica próxima estejam sujeitas à mesma fonte aleatória primária apresentando alta

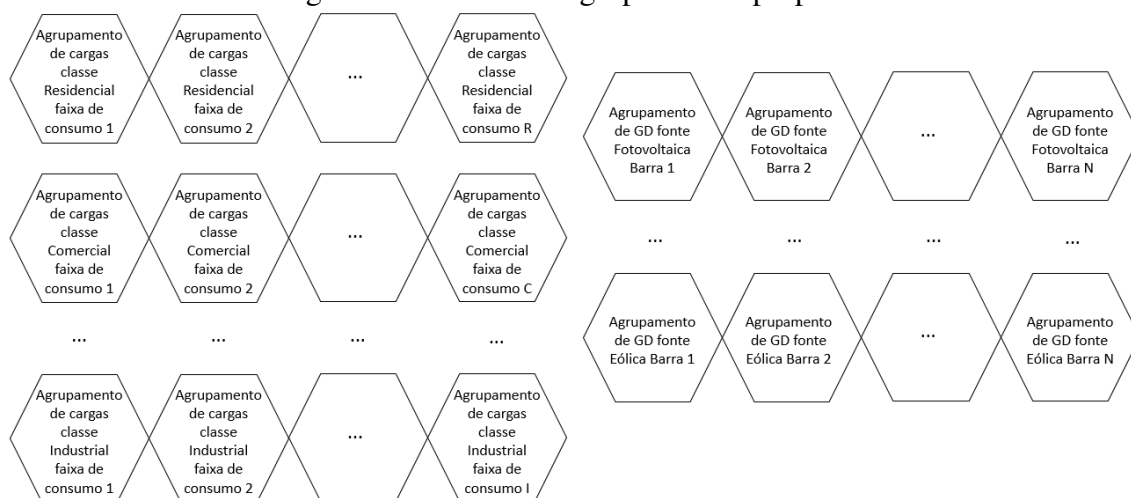
¹Cabe ressaltar que pode-se considerar cada consumidor/prosumidor como sendo um agrupamento e que não é estritamente necessário agrupar as cargas e a GD. A obtenção da descrição de comportamento pode ser realizada para cada componente do sistema individualmente, dependendo o nível de detalhamento que o analista deseja obter. Agrupar os consumidores/prosumidores é apenas uma técnica de redução da dimensão do problema a ser solucionado.

²Sendo o comportamento de cada consumidor representado por uma variável aleatória independente.

dependência. Caso a tecnologia de conversão (função de transferência) entre unidades de GD seja diferente, caberia ainda dividir os agrupamentos considerando esse aspecto. O método proposto assume que a tecnologia de conversão é a mesma para todas unidades de GD da mesma fonte energética e assim, divide-se os agrupamentos por fonte energética primária e localização geográfica, i.e., o barramento ao qual estão conectados.

A Figura 22 ilustra a divisão proposta neste trabalho para os agrupamentos. Conforme mencionado anteriormente, os agrupamentos de cargas são divididos por classe e faixa de consumo enquanto os de GD são divididos por localização geográfica e fonte energética.

Figura 22: Divisão de agrupamentos proposta.



Fonte: O autor

Definidos os critérios que serão utilizados para definir os agrupamentos, procede-se a análise dos padrões de comportamento e posterior agrupamento dos componentes comonotônicos, para que se obtenha a média e desvio padrão horários da demanda/geração, além da distribuição de probabilidade horária que rege o comportamento das categorias. Também, pode-se investigar informações adicionais como potência instantânea máxima de consumo/geração e correlação nas demandas horárias entre os agrupamentos.

A média (μ) e desvio padrão (σ) horários para a demanda/geração são obtidas pela Equação (15) e pela Equação (16), respectivamente, e servem para estabelecer as curvas diárias de cada agrupamento, a exemplo da Figura 23.

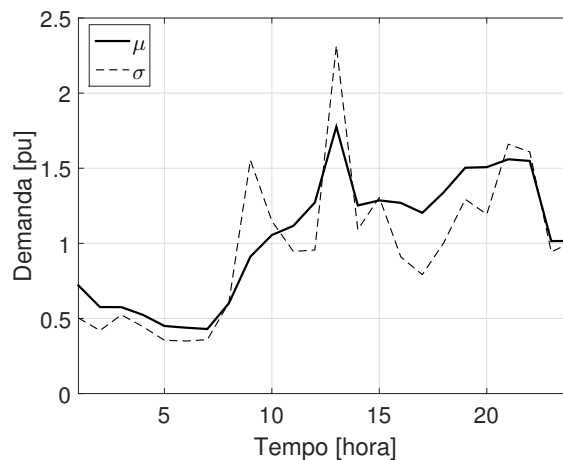
Para isso, a aplicação do método proposto por Jardini *et al.* (2000) é adotada. Dessa maneira, realizam-se os seguintes procedimentos:

1. obtém-se as curvas diárias normalizadas da média e desvio padrão da demanda/geração de cada agrupamento;
2. para se obter o valor da demanda/geração em quilowatts, basta multiplicar o fator da curva diária para o horário analisado pela potência de base de cada componente ³.

A curva de densidade de probabilidade acumulada (CDF – *Cumulative Distribution Function*) horária para a demanda/geração é obtida a partir da média e do desvio padrão horários da demanda/geração de cada grupo de componentes do sistema considerando a

³A potência de base é essencialmente proporcional ao consumo mensal ou à energia produzida pelo consumidor/prosumidor, conforme Equação (32).

Figura 23: Exemplo de curva diária de média e desvio padrão para um agrupamento de consumidores.



Fonte: O autor

distribuição de probabilidade que melhor descreve o comportamento do respectivo agrupamento.

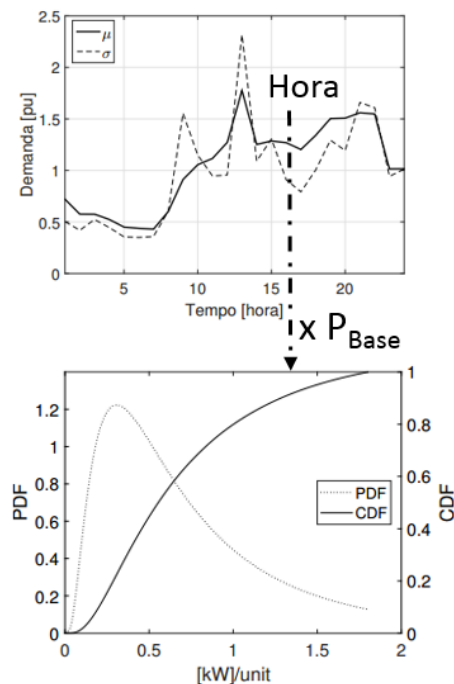
Para obter a distribuição de probabilidade que melhor representa o padrão de comportamento horário da demanda/geração de um componente ou grupo de componentes, é necessário realizar um teste de hipóteses com verificação da adequação de ajuste⁴, conforme detalhado na Seção 2.3.4. (JOHNSON; WICHERN, 1992; MONTGOMERY; RUNGER, 2007; BECKER, 2015).

A média e o desvio padrão da potência de cada horário de um componente são multiplicados pela potência de base (P_{Base}) para, em associação com a lei de distribuição que descreve o comportamento do componente, obter a CDF da potência do componente para cada hora do dia, conforme ilustra a Figura 24. Portanto, cada componente do sistema possuirá uma CDF da potência ativa para cada horário do período simulado. A potência reativa é obtida pelo uso de um fator de potência conhecido conforme o tipo de consumidor/gerador.

Por vezes, não existe um banco de dados ou não há tempo hábil para obter informações detalhadas o suficiente para descrever com exatidão os consumidores/prosumidores do sistema em análise. Nesses casos, pode-se adotar resultados dos estudos investigativos com informações que remetam aos clientes que se pretende analisar e, havendo semelhança, pode-se aplicar os resultados destes estudos ao sistema analisado. O trabalho de Pareja (2009) apresenta informações de média e desvio padrão da demanda horária para várias categorias de consumidores. Da mesma maneira, pode-se combinar resultados de estudos distintos e obter as informações necessárias para a análise pretendida. O Sistema de Informação de Posses e Hábitos de Uso de Aparelhos Elétricos é outro exemplo de fonte de informação acerca de padrão de consumo (ELETROBRAS, 2006). Também são fontes relevantes os resultados dos trabalhos de Jardini *et al.* (2000) e Francisquini (2006) no que se refere a dados de demanda de diversas categorias de clientes. Da mesma forma, existem inúmeras fontes de consulta com informações de GD, das quais podem ser citadas Pfenninger; Staffell (2016), Ekström *et al.* (2016) e LABSOL (2017). É possível ainda confrontar os cenários simulados de demanda agregada com o histórico da operação disponibilizado por ONS (2017).

⁴Do inglês *Goodness of Fit*

Figura 24: Obtenção da CDF da potência horária a partir da lei de distribuição e das curvas típicas diárias.



Fonte: O autor

Cabe lembrar que existem estudos que investigaram a distribuição de probabilidade que rege o comportamento horário de consumidores e da GD. O trabalho de Pareja (2009) demonstrou que independente da classe e da faixa de consumo, a distribuição de probabilidade que melhor descreve o comportamento horário das cargas de um sistema de distribuição é a lognormal. Já Abdulkarim; Abdelkader; Morrow (2015) comprovaram que no caso de GD eólica a distribuição de probabilidade do tipo Weibull é mais adequada enquanto que a distribuição de probabilidade do tipo beta é apropriada para GD fotovoltaica.

4.2 Modelagem da estrutura estocástica multidimensional de dependência

A modelagem estocástica de um sistema de distribuição envolve a obtenção das distribuições marginais de probabilidade bem como o modelo da estrutura de dependência entre as VAs. A primeira é facilmente obtida pelo processamento dos dados reais de radiação solar, velocidade do vento, demanda, etc, enquanto que a obtenção da estrutura de dependência é um tanto mais desafiadora. Definir a estrutura de dependência estocástica é determinar como as fontes subjacentes que determinam os cenários das VAs estão correlacionadas.

Conforme abordado no Capítulo 2, todos os casos reais de dependência entre um conjunto de VAs estarão compreendidos entre os casos extremos de dependência perfeita e independência. Assim, apesar de uma descrição detalhada da relação de dependência poder ser obtida com o uso adequado de cópulas (vide Seção 2.2.5), é possível utilizar artifícios alternativos que permitem obter as condições de contorno da operação do sistema de distribuição, e com isso, operar com resultados mais conservadores.

Conforme indicado na Seção 2.4.5, em geral, os sistemas contém grupos de componentes com uma alta correlação positiva. Assim, a aplicação da SBM (vide Seção 2.4.4) é imediata: o comportamento marginal dos componentes permanece o mesmo, porém a relação de dependência positiva entre os componentes é substituída por uma estrutura de dependência extrema de comonotonicidade. A aplicação desta técnica impõe o cenário de maior risco para a demanda agregada e favorece a redução da ordem de grandeza do problema a ser solucionado (PAPAEFTHYMIU, 2007; RODRIGUES, 2008).

Com a aplicação da SBM, todos os componentes de um agrupamento seguem o mesmo fator aleatório. No entanto, considerar o mesmo para componentes em diferentes agrupamentos é extremamente conservador e a correlação mútua entre os grupos de componentes deve ser mensurada.

A modelagem da estrutura de dependência se completa com a aplicação da técnica da JNT, detalhada na Seção 2.4.3. A análise de dados dos consumidores/prosumidores e a obtenção de informações acerca da correlação entre o comportamento dos agrupamentos é importante para a avaliação sistêmica, visto que não basta avaliar o comportamento individual dos componentes do sistema, mas também da interação entre os mesmos, como já demonstraram Papaefthymiou (2007) e Rodrigues (2008).

A análise da correlação entre os agrupamentos deve ser realizada para cada horário simulado da curva de carga diária, visto que a relação de dependência entre os agrupamentos pode sofrer alterações conforme o horário do dia.

Para uma determinada hora t , utilizando a matriz de correlação \mathbf{R}_t e um vetor de números aleatórios independentes, normalmente distribuídos, \mathbf{z}_t , o respectivo vetor de números aleatórios correlacionados uniformemente distribuídos \mathbf{u}_t é dado por (PAPAEFTHYMIU, 2007; RODRIGUES, 2008):

$$\mathbf{u}_t = \Phi(\mathbf{y}_t) \quad (37)$$

$$\mathbf{y}_t = \mathbf{L}_t \mathbf{z}_t \quad (38)$$

$$\mathbf{L}_t \mathbf{L}_t^T = \mathbf{R}_t \quad (39)$$

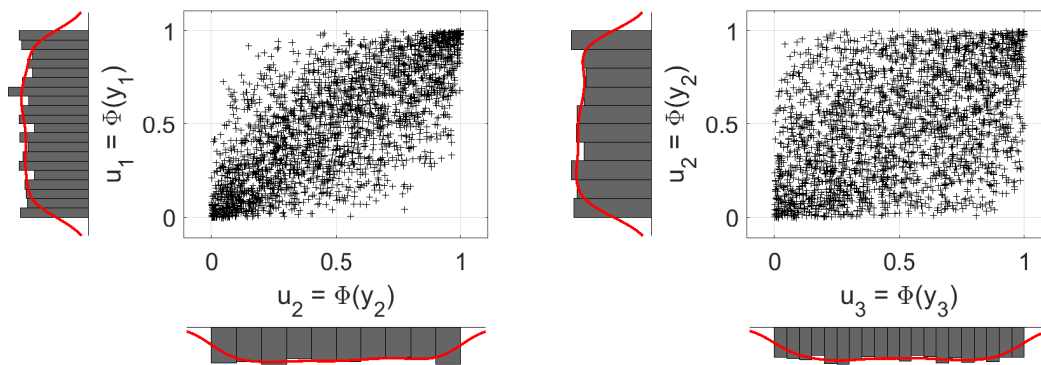
onde $\Phi(\mathbf{y}_t)$ é a CDF de uma distribuição normal para os valores aleatórios correlacionados do vetor \mathbf{y}_t e, \mathbf{L}_t , é a matriz triangular inferior da decomposição de Cholesky da matriz \mathbf{R}_t .

De forma simplificada, pode-se utilizar uma única matriz de correlação diária, o que não contribui para a redução da dimensão do problema. A única vantagem, neste caso, é que uma vez calculada a matriz triangular inferior da decomposição de Cholesky ao aplicar a JNT, esta será única, independente do horário do dia, enquanto que no caso anterior é necessário obter tantas matrizes triangulares inferiores quantos forem os horários da curva de carga simulados.

Considerando a correlação entre os componentes (ou agrupamentos) e fazendo uso da SBM, a JNT permite obter números aleatórios que contém intrinsecamente a informação da estrutura de dependência entre as VAs. A Figura 25 ilustra, a título de exemplo, os números uniformes \mathbf{u}_1 , \mathbf{u}_2 e \mathbf{u}_3 correlacionados resultantes da JNT (vide Seção 2.4.3) para os agrupamentos GP1, GP2 e GP3 no mesmo horário t . Assim como são apresentadas as relações de dispersão para três uniformes referentes aos três agrupamentos, para um sistema com K agrupamentos, haveria uma relação de dispersão entre cada uniforme \mathbf{u}_K com os respectivos uniformes dos demais agrupamentos.

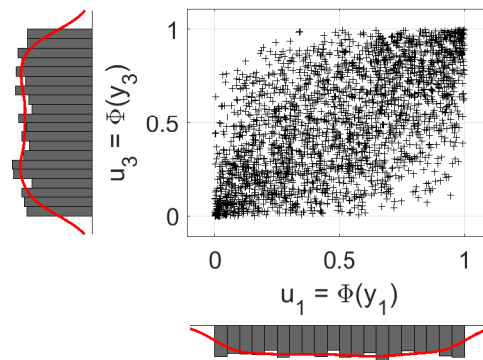
Observa-se que há concentração dos pontos na região da diagonal principal na Figura 25(a), indicando uma correlação $\rho_{1,2}$ entre \mathbf{u}_1 e \mathbf{u}_2 maior do que a correlação $\rho_{2,3}$ existente

Figura 25: Resultado da JNT para os agrupamentos.



(a) Dispersão entre \mathbf{u}_1 e \mathbf{u}_2 .

(b) Dispersão entre \mathbf{u}_2 e \mathbf{u}_3 .



(c) Dispersão entre \mathbf{u}_1 e \mathbf{u}_3 .

Fonte: O autor

entre \mathbf{u}_2 e \mathbf{u}_3 , visto que na Figura 25(b) os pontos possuem maior dispersão. Da mesma forma, é possível inferir que a correlação $\rho_{1,3}$ entre \mathbf{u}_1 e \mathbf{u}_3 é tal que $\rho_{2,3} \leq \rho_{1,3} \leq \rho_{1,2}$.

Assim, quanto mais próximos à diagonal principal estiverem os pontos da dispersão entre dois números aleatórios uniformes, maior será a correlação entre os mesmos. No caso de haver uma dispersão sem tendência de concentração, a relação é de independência.

Aplicada a JNT, está cumprida a modelagem estocástica do sistema e as amostras provenientes deste passo, servirão para a criação dos cenários simulados, conforme detalhado na Seção 4.3.

4.3 Criação de cenários de demanda e geração distribuída de energia

Na Seção 4.1 obteve-se a descrição do comportamento individual de cada componente do sistema e na Seção 4.2 foi modelada a estrutura de dependência entre os mesmos. A partir dos resultados obtidos nestas duas etapas, realiza-se a criação dos cenários da simulação estocástica que serão avaliados.

Como pode ser visualizado na Figura 24, cada componente possui uma curva diária para a média e desvio padrão da demanda/geração. Com a lei característica que rege o comportamento daquele componente e sabendo a potência de base do mesmo, obtém-se uma curva de densidade de probabilidade acumulada (CDF) específica de cada componente para o horário em estudo.

A Figura 26 representa o procedimento de amostragem dos cenários de potência individual a partir da CDF de três componentes da rede, cada um pertencente a um dos agrupamentos GP1, GP2 e GP3 do exemplo anterior, utilizando os números aleatórios correlacionados u_1 , u_2 e u_3 ilustrados na Figura 25 e originados pela JNT, conforme detalhado na Seção 4.2. Estes números seguem uma distribuição uniforme de probabilidade e apresentam relação de dependência entre si iguais à correlação existente entre os consumidores/prosumidores. Com eles é possível amostrar a potência de cada componente de forma similar ao ilustrado na Figura 9 e detalhado na Seção 2.4.1. Desta maneira, têm-se amostras de potência dos componentes correlacionadas entre si para compor o cenário simulado agregando as potências nos nós, conforme mostrado na Figura 27.

O método proposto segue parcialmente a técnica de Jardini *et al.* (2000) e faz uma representação normalizada da curva diária para a média e desvio padrão da demanda/geração dos componentes associados a uma potência de base. Contudo, a demanda é agregada de forma ligeiramente diferente. Jardini *et al.* (2000) consideram que os componentes sigam distribuição idêntica e por isso a soma das VA ocorre pela simples operação entre seus parâmetros estatísticos (média e desvio padrão). Por outro lado, isto não é possível neste trabalho, uma vez que este método tem por premissa considerar que as distribuições de cada componentes podem ser diversas. Logo, a potência agregada em cada nó N deve ser obtida pela soma das potências individuais amostradas dos K componentes adjacentes. Assim, realiza-se a agregação da potência em cada nó de forma que o cenário de demanda e geração distribuída a ser simulado fica definido.

4.3.1 Representação do acoplamento temporal na curva horária da demanda

A modelagem proposta por Papaefthymiou (2007) permite correlacionar as fontes aleatórias primárias apenas dos agrupamentos pertencentes ao mesmo horário quando aplicada a JNT, conforme Seção 4.2. Entretanto, para um mesmo dia de operação, as fontes aleatórias primárias não possuem nenhuma relação temporal, ou seja, o valor amostrado num instante de tempo para um determinado agrupamento não possui acoplamento com

Figura 26: Procedimento de amostragem das potências dos componentes da rede.

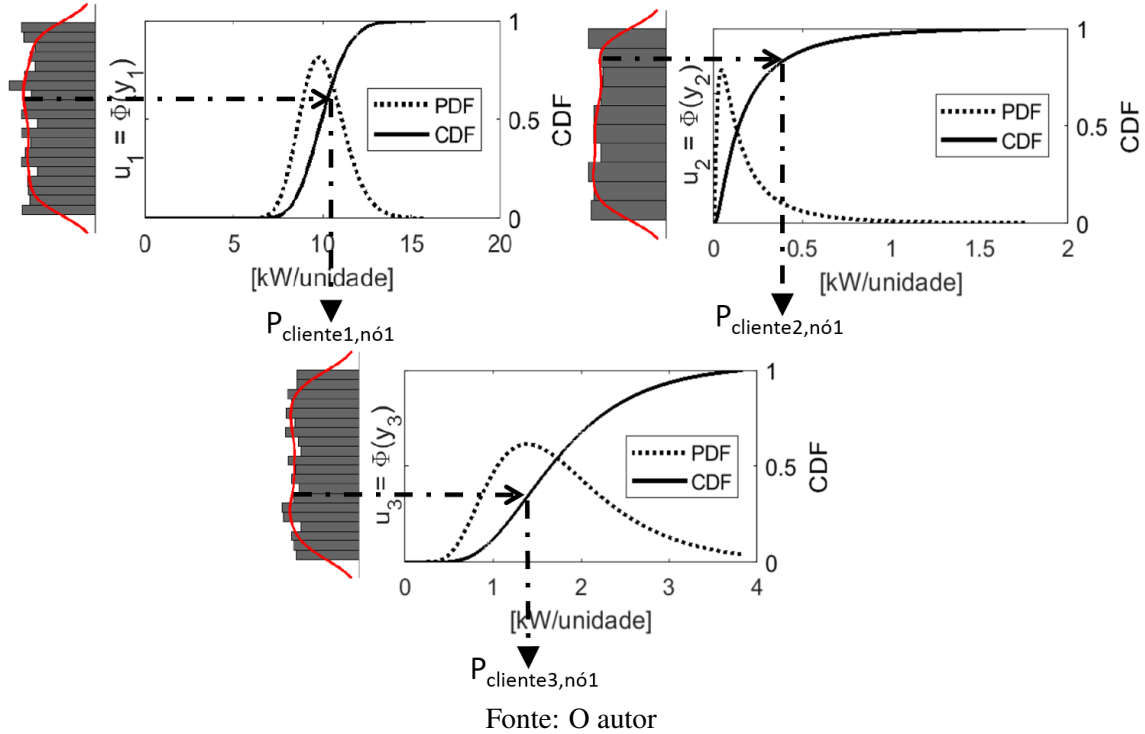
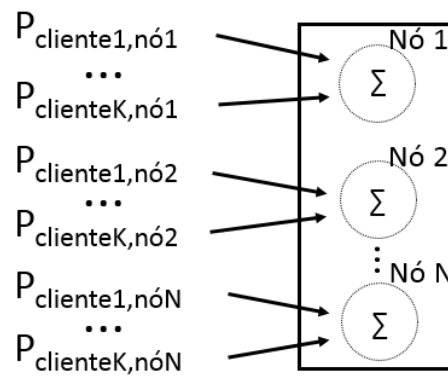


Figura 27: Procedimento de agregação das potências dos componentes nos nós da rede.



os valores amostrados nos instantes anteriores para este mesmo agrupamento. Para permitir representar uma relação entre os valores amostrados nos sucessivos instantes de tempo para os mesmos componentes do sistema, buscou-se considerar um efeito de memória, no qual a demanda de um determinado horário depende em parte da demanda do horário anterior⁵.

Neste trabalho, optou-se por modelar a memória considerando que a demanda de um horário é obtida a partir da combinação da potência amostrada na CDF, utilizando o número aleatório do respectivo horário, juntamente com a potência amostrada na CDF utilizando o número aleatório do horário anterior. Desta forma, a demanda de um consumidor qualquer, para um determinado horário, é dada por:

$$D(t) = CDF_{i,t}^{-1}[u_{i,n}(t)](1 - M) + \dots + CDF_{i,t}^{-1}[u_{i,n}(t - 1)]M \quad (40)$$

em que:

$D(t)$ – é a demanda no horário t ;

$CDF_{i,t}^{-1}(u)$ – é a amostra de potência obtida da curva de densidade de probabilidade acumulada inversa para qualquer consumidor do agrupamento i , no horário t , com o número aleatório u ;

$u_{i,n}(t)$ – é o n -ésimo número aleatório correlacionado no horário t para o agrupamento i ;

M – é a memória considerada, em valor por unidade.

4.4 Cálculo do fluxo de potência e dos parâmetros de operação da rede

Uma vez formulados os cenários, procede-se ao cálculo do fluxo de potência propriamente dito. O fluxo de potência é calculado de forma similar ao caso de análises determinísticas, sendo que a única diferença é a avaliação do sistema quando sujeito a várias possibilidades de cenários de carga e GD.

Após o cálculo do fluxo de potência, são avaliados os indicadores que descrevem a operação do sistema de distribuição, sendo calculados durante o período de tempo utilizado na análise. Sem perda de generalidade, neste trabalho, foi adotado um período de um dia de operação, subdividido em 24 horas. Assim, o fluxo de potência é calculado para cada hora da amostra sorteada na criação dos cenários, sendo os indicadores calculados a partir dos resultados obtidos.

Neste trabalho, foram considerados dois tipos de indicadores:

⁵ Isso porque pode existir uma espécie de memória ou um viés, mesmo que pequeno, da demanda de uma carga ao longo dos horários. Uma forma intuitiva de imaginar esse efeito é observar o que ocorre em dias tipicamente muito quentes, nos quais há uma constante térmica associada e, conseqüentemente, aumento da parcela da demanda utilizada para equipamentos de conforto térmico dos clientes. Assim, para um dia muito quente, há um viés de alta demanda das cargas em todos os horários. Se não considerarmos esse viés e analisarmos o cenário diário com os cenários horários sendo gerados independentemente, há possibilidade de que as amostras das demandas para os consumidores/prosumidores resultem em uma curva de carga pouco coerente. Da mesma forma que é necessário incluir a dependência entre as variáveis aleatórias na análise, é importante considerar a evolução das demandas ao longo dos horários, e não simplesmente avaliar cada horário isoladamente.

- energia de perda diária e
- índices de violação de tensão crítica e precária.

Para cada amostra que descreve um dia de operação, as perdas são integralizadas obtendo-se a energia de perdas diárias nos transformadores e circuitos de distribuição, dados em kWh (ANEEL, 2015). De modo similar, para determinar os índices de violação de tensão crítica (DRC) e de tensão precária (DRP) em cada barra da rede simulada, referente à amostra em análise, são utilizadas as seguintes expressões (ANEEL, 2018):

$$DRC(k) = \frac{nlc(k)}{nl} 100\%, \quad (41)$$

$$DRP(k) = \frac{nlp(k)}{nl} 100\%, \quad (42)$$

onde $nlc(k)$ é o maior número de leituras de tensão na região crítica entre as fases da Barra k , $nlp(k)$ é o maior número de leituras de tensão na região precária entre as fases da Barra k , e nl é o número total de leituras de tensão efetuadas para a barra. Cada hora da curva de carga diária de operação corresponde a uma leitura, logo $nl = 24$, para um dia de operação.

O número mínimo de cenários simulados é tal que permita atingir o regime permanente para as densidades de probabilidade das variáveis aleatórias de interesse nas simulações, conforme ilustrado na Figura 10. À medida que aumenta a quantidade de amostras das variáveis aleatórias na SMC, é possível observar a convergência de determinadas estatísticas, pois durante este processo um comportamento transitório é observado. O critério de convergência para as simulações estocásticas é escolha de cada analista, desde que as condições supracitadas sejam atingidas.

Este método propõe o uso do coeficiente de variação estatístico (C_v) detalhado na Seção 2.4.2. Com a Equação (23), calcula-se a convergência para cada indicador e a parada das simulações para o sistema ocorrerá quando todos indicadores tiverem convergido individualmente, atingindo C_v inferior à tolerância estipulada pelo analista (tipicamente 5%). Portanto, no método proposto, a parada ocorre quando o coeficiente de variação global do sistema, C_v^{global} , dado pela expressão a seguir for inferior à tolerância:

$$C_v^{\text{global}} = \max \left(C_v^{\text{perdas}}, C_v^{DRC(k)}, C_v^{DRP(k)}, C_v^{P_{k,t}} \right), \quad (43)$$

em que C_v^{perdas} é o coeficiente de variação das perdas diárias no sistema, $C_v^{DRC(k)}$ é o coeficiente de variação das violações de tensão crítica de cada Barra k do sistema, $C_v^{DRP(k)}$ é o coeficiente de variação das violações de tensão precária de cada Barra k do sistema e $C_v^{P_{k,t}}$ é o coeficiente de variação da demanda agregada na Barra k no horário t do período avaliado. No caso de considerar outros indicadores de operação, simplesmente é incluído o respectivo coeficiente de variação na Equação (43).

Após a convergência das simulações, é possível obter intervalos de confiança para os indicadores de interesse a partir da distribuição de probabilidade resultante de testes de adequação de ajuste, bem como estimar o percentual de tempo com os indicadores em determinado patamar com o uso das equações do Capítulo 2 (MONTGOMERY; RUNGER, 2007).

4.5 Aspectos complementares do método

O modelo determinístico do sistema envolve a descrição dos parâmetros elétricos da rede e da configuração dos componentes. Este modelo é idêntico ao empregado em análises determinísticas tradicionais, com a definição da matriz de admitância e solução do fluxo de potência por métodos adequados e consolidados para redes de distribuição, como o *backward/forward sweep* usado por Pareja (2009) entre outros (KAGAN; OLIVEIRA; BORBA, 2005; KERSTING, 2007).

O programa de simulação OpenDSS[®], utilizado neste trabalho, difere dos programas tradicionais de solução de fluxo de potência para circuitos radiais à medida que o mesmo calcula com a mesma facilidade o fluxo de potência de redes radiais e de redes malhadas. O mesmo também pode ser usado para resolver redes pequenas a médias com um fluxo de potência nos moldes usados para a transmissão. Existem dois tipos de solução para o fluxo de potência disponíveis: solução direta e fluxo de potência iterativo. A solução iterativa é tipicamente utilizada para o cálculo do fluxo de potência, sendo as cargas e geradores distribuídos tratados como fontes de injeção de corrente. O estado inicial é obtido a partir do resultado do fluxo de potência à vazio, i.e., considerando apenas as admitâncias série. Esse resultado é utilizado para obter a injeção de corrente dos elementos de conversão de potência do sistema e atualizar o vetor de correntes do sistema. A tensão nas barras são então recalculadas e esse processo se repete até atingir a convergência. Os cálculos são realizados mantendo a matriz de admitância inalterada (EPRI, 2013).

A diferença do método de análise probabilística proposto para os métodos tradicionais de análise de sistemas de distribuição é que do modelo estocástico do sistema obtém-se vários cenários, sendo cada um avaliado por intermédio de um modelo determinístico do sistema. Nos métodos de análise determinística, poucos ou um único cenário são avaliados, tanto de cargas como de GD. Além disto, o resultado obtido pela abordagem probabilística consiste de uma distribuição de probabilidade de valores para cada uma das grandezas determinadas.

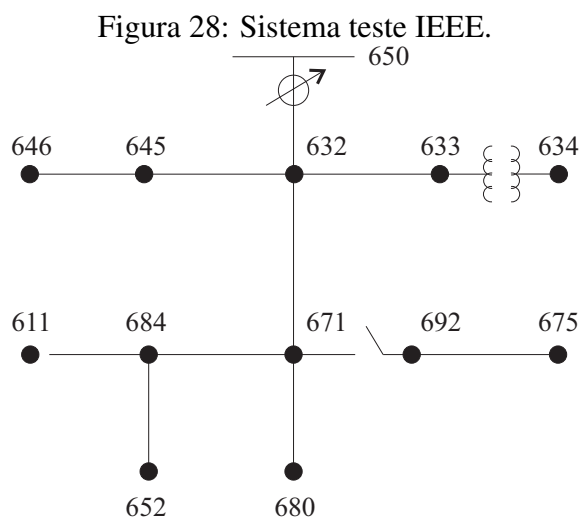
Em princípio, não há necessidade de alterar o modelo determinístico do sistema. Considerando que não sejam contempladas mudanças de configurações da rede devido a contingências, faltas, entre outros, o circuito permanece inalterado e assim a única fonte de incerteza é proveniente do cenário de cargas e GD. Avaliar vários cenários para as cargas e a GD é portanto avaliar resposta do sistema de distribuição frente às incertezas e aleatoriedades proveniente do comportamento dos componentes.

Com relação à maior quantidade de cenários avaliados, o tamanho do problema é uma função da quantidade de pontos da curva de carga simulados (número de horas, por exemplo), do número de cenários (amostras) avaliados, do número de variáveis aleatórias (ou agrupamentos) e da quantidade de cargas e geração distribuída do sistema (tamanho da rede de distribuição). O número de variáveis aleatórias será no máximo igual à quantidade de componentes do sistema, por isso a técnica de agrupamento deve ser empregada para reduzir o universo de possibilidades investigadas.

5 APLICAÇÃO DO MÉTODO E RESULTADOS OBTIDOS

5.1 Rede de Testes e Configurações de Simulação

A fim de avaliar o desempenho do método proposto, foram feitos estudos numéricos em uma rede com a topologia do sistema teste IEEE de 13 nós (IEEE, 2000). O sistema IEEE de 13 barras opera com tensão nominal de 4,16 kV e apresenta circuitos desbalanceados e cargas elevadas (IEEE, 2000), com a topologia mostrada na Figura 28. Desta forma, embora tenha uma dimensão pequena, a rede inclui características comuns aos circuitos reais de distribuição.



Fonte: (IEEE, 2000).

As simulações foram realizadas com apoio dos programas MATLAB[®] e OpenDSS[®]. O sistema de testes de 13 barras do IEEE também foi simulado com o programa OpenDSS[®] por Freitas (2015), Paludo (2014) e por Anzanello (2016). No entanto, a avaliação realizada nestes trabalhos envolvia um cenário típico de carga e de GD.

Algumas adaptações com relação à configuração original da rede foram necessárias para viabilizar uma análise probabilística. As principais adaptações da rede de teste são as seguintes:

- definição da quantidade de consumidores de cada grupo que está conectado em cada nó;
- definição de curvas de média e desvio padrão da demanda diária dos consumidores de cada grupo;

- c) alteração do modelo de resposta à variação de tensão das cargas;
- d) inclusão e posicionamento da GD;
- e) definição de curvas de média e desvio padrão da GD agregada em cada barra; e
- f) definição dos níveis de penetração de GD.

Uma vez que a referência do IEEE apresenta uma única condição de cargas, sem especificar horário, dia ou mesmo período do ano no qual ocorre, os ajustes enumerados foram necessários para viabilizar a análise ao longo de um dia típico de operação da rede (IEEE, 2000). As adaptações referentes às cargas são detalhadas na Seção 5.1.1 enquanto que a Seção 5.1.2 detalha as adaptações acerca da inclusão da GD.

O modelo determinístico da rede foi descrito no OpenDSS utilizando como base o código desenvolvido por Freitas (2015), sendo adaptados o modelo de resposta à tensão das cargas bem como a inclusão de GDFV. No modelo original do IEEE, a linha entre as barras 632 e 671 possui carga distribuída (IEEE, 2000). Para viabilizar as simulações no OpenDSS, a carga foi concentrada na barra 670, criada artificialmente entre as duas barras originais (FREITAS, 2015).

5.1.1 Adaptações para representação da carga

De maneira a viabilizar uma análise estatística, foi necessário definir a quantidade de consumidores em cada grupo que está conectada em cada nó da rede. Tal definição se baseou em informações de um banco de dados acerca de clientes de um sistema de distribuição real (PAREJA, 2009). As informações consideradas foram o padrão de conexão à rede, valor máximo de consumo mensal e curvas diárias de média e desvio-padrão da demanda.

Os clientes do sistema real foram divididos em agrupamentos (GP) de acordo com padrão de consumo típico e similaridade no comportamento (PAREJA, 2009):

- GP1: residencial até 80 kWh;
- GP2: residencial 81-220 kWh;
- GP3: residencial 221-500 kWh;
- GP4: residencial 501-1000 kWh;
- GP5: comercial até 500 kWh;
- GP6: comercial 501-1000 kWh;
- GP7: comercial 1001-5000 kWh;
- GP8: industrial até 500 kWh; e
- GP9: industrial entre 501-1000 kWh.

Cada agrupamento definido por Pareja (2009) apresenta uma curva diária característica de média e desvio padrão para a demanda. Os valores representativos de carga em p.u. utilizados nas simulações deste trabalho foram obtidas do trabalho de Pareja (2009)

Tabela 2: Valores representativos de carga em p.u. para clientes residenciais dos agrupamentos 1 e 2.

Hora	GP 1		GP 2	
	μ	σ	μ	σ
1	1,10234960	0,48706732	0,76391182	0,51997551
2	0,85103670	0,28518272	0,59803076	0,27150329
3	0,73680357	0,30193548	0,57011547	0,32803649
4	0,84532518	0,22423251	0,50784294	0,23183312
5	0,78820851	0,26834148	0,57709422	0,31116840
6	0,61685882	0,25051652	0,63399843	0,40910899
7	0,74251530	0,25066299	0,76230123	0,55220400
8	0,64541716	0,17827780	0,90831950	0,59789356
9	0,67397544	0,27735369	0,58783089	0,28702898
10	0,87388335	0,43542715	0,61843039	0,37937736
11	0,67397539	0,24308114	0,59427296	0,32381393
12	0,75965028	0,29280245	0,69143958	0,52839192
13	0,77107358	0,30193548	0,94536097	0,79410342
14	0,82819015	0,37638472	0,83477352	0,47714861
15	1,04523304	0,35335031	0,81276352	0,52470264
16	0,79392019	0,21055800	0,84443644	0,47117375
17	0,80534360	0,23416049	1,05004324	0,84080714
18	0,91957656	0,42888795	1,02803308	0,68037011
19	0,87959509	0,33059835	1,11446276	0,69119268
20	1,59355205	0,75418198	1,92615276	1,31573417
21	1,87913497	0,71880126	2,25308362	1,83278447
22	1,78774845	1,00067326	1,64002139	0,90521690
23	1,38222084	0,37462543	1,47521402	0,82306363
24	1,34795077	0,43297597	1,39898404	0,81826548

Fonte: (PAREJA, 2009).

Tabela 3: Valores representativos de carga em p.u. para clientes residenciais dos agrupamentos 3 e 4.

Hora	GP 3		GP 4	
	μ	σ	μ	σ
1	1,01046847	0,42182996	1,09421798	0,43867395
2	0,78929032	0,31080252	0,83957436	0,40119556
3	0,74007278	0,30074349	0,82359790	0,39027330
4	0,68784163	0,27953678	0,66611661	0,29026002
5	0,65710582	0,26631757	0,70328610	0,26895298
6	0,66353442	0,28567468	0,68241903	0,30138735
7	0,67418130	0,32953023	0,72513127	0,30809047
8	0,86341814	0,48887696	0,90054523	0,43527384
9	0,93613990	0,56793735	1,12910522	0,75821959
10	0,75453670	0,41286696	0,84935575	0,36554298
11	0,81721397	0,55291537	0,99933790	0,61850783
12	0,75031797	0,40777544	0,73360859	0,34119092
13	1,01287928	0,89240025	0,87739592	0,54446302
14	0,92850599	0,53444592	0,83468359	0,38215066
15	0,79712498	0,39111065	0,79294929	0,30855261
16	0,86562802	0,59533780	0,94064928	0,40160837
17	0,89756938	0,59599687	0,83957428	0,34705275
18	0,87506985	0,53238741	0,94456178	0,40294708
19	1,06511001	0,56578485	0,95075678	0,43757553
20	1,50043538	0,88428106	1,28952090	0,78222247
21	1,73105501	1,00752697	1,64556563	0,98218159
22	1,68625717	0,83616570	1,63904460	0,62083817
23	1,49380594	0,81870454	1,37657573	0,36346277
24	1,41887448	0,69720416	1,32342995	0,52927035

Fonte: (PAREJA, 2009).

Tabela 4: Valores representativos de carga em p.u. para clientes comerciais dos agrupamentos 5, 6 e 7.

Hora	GP 5		GP 6		GP 7	
	μ	σ	μ	σ	μ	σ
1	0,71988951	0,50344102	0,61882486	0,34072696	0,42485002	0,30745042
2	0,57611841	0,41951222	0,56393954	0,33447001	0,36856369	0,27699786
3	0,57529155	0,52408799	0,51425673	0,33200116	0,36697371	0,28582110
4	0,52419866	0,44551151	0,53480618	0,33405079	0,40004592	0,3341447
5	0,45008088	0,35502641	0,49370721	0,31460222	0,37810376	0,27228434
6	0,43854767	0,34975119	0,53896811	0,32912463	0,33231156	0,25136253
7	0,42986688	0,35910284	0,54495079	0,33754770	0,39400381	0,34350717
8	0,60121011	0,60472912	0,70596492	0,55456211	0,43534403	0,28069426
9	0,91140553	1,55531111	1,11071124	0,74434497	0,76892758	0,37395850
10	1,05534205	1,14859232	1,27796811	0,88994208	1,75218825	1,23524803
11	1,11652125	0,94632502	1,62184650	0,94668411	1,91373297	0,88048664
12	1,27186691	0,95576294	1,49360735	0,73946320	2,10039989	1,15681245
13	1,77597592	2,31680705	1,47409853	0,70749176	1,84313660	0,68233516
14	1,25235572	1,09161649	1,33077246	0,52089166	1,44468076	0,22494942
15	1,28604574	1,30560340	1,67022887	0,87548387	1,92390906	0,95591543
16	1,26963480	0,91000826	1,68323453	0,88888630	2,25749290	1,33020508
17	1,20287508	0,79181957	1,78156008	0,96014503	2,30169499	1,44687721
18	1,34077642	1,00429660	1,92254484	0,90842917	2,31155296	1,37055336
19	1,50352130	1,29359789	1,70300373	0,97810289	1,30412416	0,52323993
20	1,50740688	1,19335289	1,33441420	0,77848941	0,91711640	0,41048978
21	1,55982288	1,66062920	1,01342639	0,77398606	0,77942174	0,43957069
22	1,54857917	1,60851050	0,99469790	0,72375314	0,70818937	0,43542385
23	1,01512075	0,94313770	0,87790424	0,68301576	0,56636064	0,42266924
24	1,01615419	1,01101860	0,72625431	0,44828283	0,51166434	0,40165883

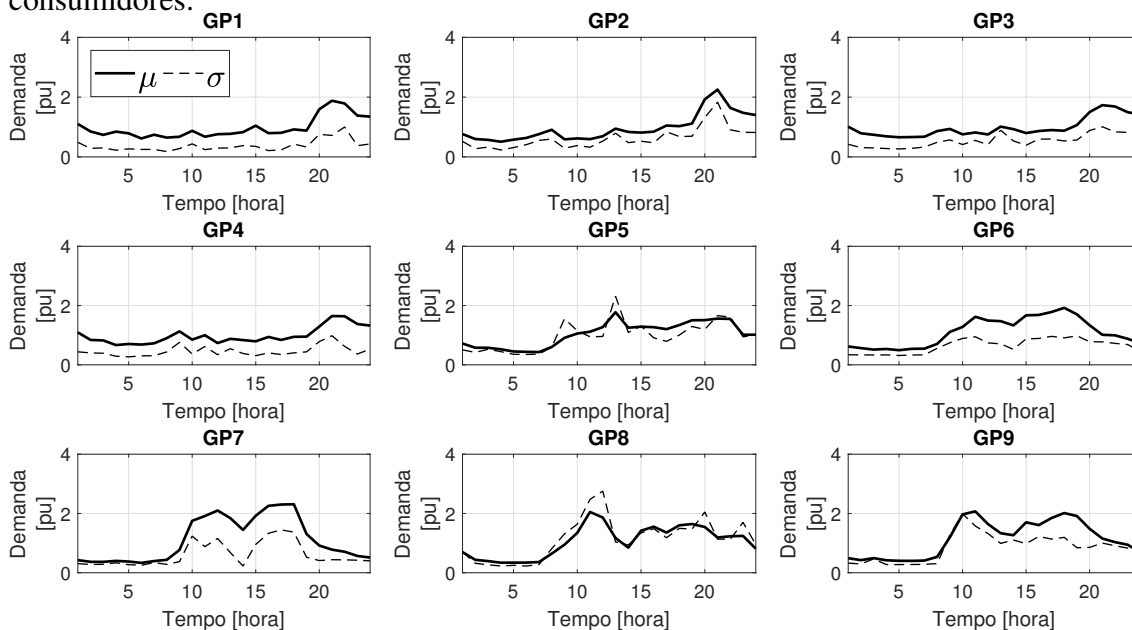
Fonte: (PAREJA, 2009).

Tabela 5: Valores representativos de carga em p.u. para clientes industriais dos agrupamentos 8 e 9.

Hora	GP 8		GP 9	
	μ	σ	μ	σ
1	0,69670188	0,70573747	0,49139762	0,32370526
2	0,43588870	0,31870995	0,42479658	0,29696891
3	0,39199982	0,26846894	0,49366299	0,46732683
4	0,33954913	0,22681054	0,42008474	0,27654312
5	0,34054937	0,25031059	0,40277749	0,27586736
6	0,34194963	0,22726363	0,40005913	0,27788360
7	0,36071343	0,27729476	0,41102339	0,28380900
8	0,64593152	0,81347953	0,53570773	0,30744544
9	0,94759295	1,31239325	1,19555601	1,27834543
10	1,34599394	1,64062860	1,96839966	1,99384111
11	2,05021745	2,47725790	2,07188020	1,57670229
12	1,86349955	2,75416610	1,64209989	1,32174407
13	1,17195849	1,03886224	1,33392315	0,98935791
14	0,84733246	0,93394678	1,26496610	1,10303256
15	1,41732848	1,34403767	1,70199547	0,98488781
16	1,55519651	1,48572787	1,60775725	1,22207161
17	1,35403555	1,18146762	1,84579951	1,13626927
18	1,60024585	1,49786618	2,01470303	1,19363691
19	1,64453488	1,47244932	1,91366880	0,84329049
20	1,54363413	2,04964084	1,48352612	0,85541780
21	1,18988221	1,13384015	1,15450810	1,00115745
22	1,22953029	1,12962107	1,03290462	0,91009483
23	1,24337315	1,69974118	0,94337833	0,83026141
24	0,81136520	0,95150920	0,68367973	0,57176646

Fonte: (PAREJA, 2009).

Figura 29: Curva de média (μ) e desvio padrão (σ) da demanda dos agrupamentos de consumidores.



Fonte: O autor.

e são apresentados nas Tabelas 2 à Tabela 5, sendo as respectivas curvas ilustradas na Figura 29. As curvas de carga da Figura 29 são normalizadas e a potência de base é definida a partir do consumo do respectivo cliente através da Equação (32).

Desse modo, cada carga do sistema de 13 barras original do IEEE foi substituída por um conjunto de clientes, classificados nos agrupamentos GP1 a GP9, conforme detalhado na Tabela 6¹.

Foi considerado que os clientes apresentam o consumo máximo para a categoria à qual pertencem (por exemplo, clientes do GP1 apresentam consumo de 80 kWh). Para definir os agrupamentos existentes em cada Barra, foram identificados no banco de dados do sistema de distribuição real os clientes com o mesmo padrão de conexão da carga original (PAREJA, 2009). Desse modo, a quantidade de clientes² foi ajustada de forma proporcional à participação de mercado dos clientes de cada agrupamento com o padrão de conexão da carga original para que a demanda média no horário de ponta fosse próxima à demanda definida nos dados originais.

Infelizmente não há informação acerca do dia ao qual se referem as curvas de carga típica apresentadas por Pareja (2009) e não é possível afirmar se é acerca de dia útil, sábado ou domingos e feriados (ou se inclui todos os dias de um determinado período de análise). Contudo, isso não prejudica a análise visto que bastaria replicar o método para os demais casos de interesse.

Além do uso de um conjunto equivalente de clientes para a determinação da demanda ao longo do dia nas barras, o modelo de resposta à variação de tensão utilizado para as cargas foi alterado para o modelo ZIP, que é diferente do original, porém é justificado pe-

¹A E_{base} [kWh] é o consumo mensal dos clientes e é utilizada para calcular a Potência de base das curvas de média e desvio padrão da demanda diária.

²A substituição das cargas originais do IEEE por um conjunto equivalente de clientes pertencentes a categorias com padrão característico de consumo é um artifício utilizado para a determinação da demanda na Barra. Uma vez que há uma curva diária de demanda definida para os clientes dos agrupamentos, torna-se possível simular cenários de demanda para um dia típico.

Tabela 6: Quantidade definida de clientes de cada agrupamento (GP) em cada barra.

Barra	Agrupamento (GP)								
	1	2	3	4	5	6	7	8	9
670 A	14	75	22	-	3	-	-	-	-
670 B	14	75	22	-	3	-	-	-	-
670 C	14	75	22	-	3	-	-	-	-
671 ABC	-	180	180	-	180	-	90	-	90
645 B	30	168	48	-	6	-	-	-	-
646 BC	15	141	111	9	3	3	6	6	3
692 AC	10	94	74	6	2	2	4	4	2
675 A	49	271	78	-	10	-	-	-	-
675 B	49	271	78	-	10	-	-	-	-
675 C	49	271	78	-	10	-	-	-	-
611 C	30	168	48	-	6	-	-	-	-
652 A	25	140	40	-	5	-	-	-	-
634 A	24	131	38	-	5	-	-	-	-
634 B	24	131	38	-	5	-	-	-	-
634 C	24	131	38	-	5	-	-	-	-
E_{base} [kWh]	80	220	500	1000	500	1000	5000	500	1000
P_{base} [W]	111,1	305,5	694,4	1388,8	694,4	1388,8	6944,4	694,4	1388,8

los aspectos já discutidos na Seção 3.3. As cargas foram modeladas com 50% de potência constante e 50% de impedância constante e manteve-se o fator de potência constante e igual ao original.

Os limites de tensão que o programa OpenDSS[®] utiliza para alterar o modelo das cargas para garantir a convergência também foram alterados e configurados como 0,6 p.u. e 1,4 p.u., de forma a garantir que a característica ZIP da carga fosse mantida. Os capacitores e reguladores de tensão foram mantidos ativos nas simulações.

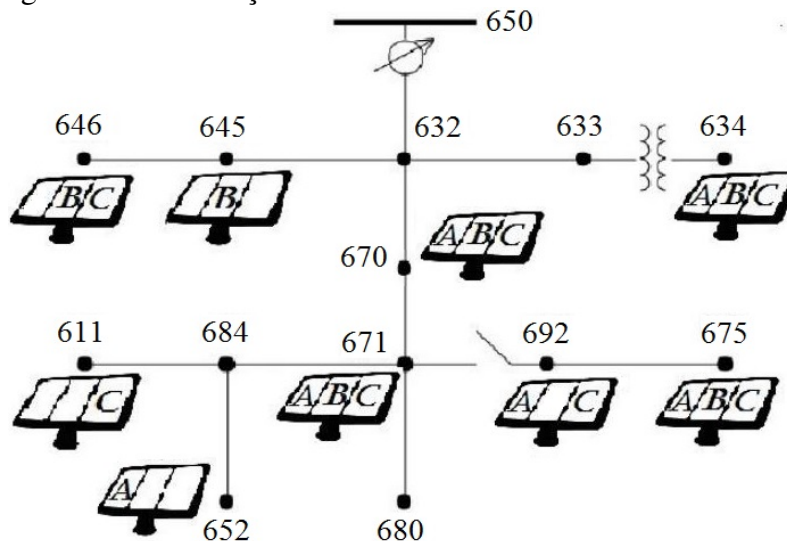
5.1.2 Adaptações para inclusão da GD

Assim como houve a definição dos consumidores, foi estabelecido um cenário de GD para o sistema. As principais incertezas com relação à GD envolvem localização, potência nominal de geração e fonte energética. Neste estudo de caso, somente foi considerada a contribuição da GD de fonte fotovoltaica (GDFV). Com relação à localização, foi considerado que todos os consumidores possuem sistema de GD, passando a ser tratados como prosumidores. A Figura 30 indica os locais considerados da GDFV no sistema.

Buscando avaliar o impacto da GD nos sistemas de distribuição, foram considerados diferentes percentuais de penetração. Conforme Shayani (2010), ainda não há consenso acerca de uma definição para o que se entende por penetração de GD. Em alguns casos é considerada uma relação com capacidade nominal, outros com demanda e até mesmo com energia. Como neste estudo procura-se avaliar a operação para um sistema de distribuição, a definição considerada mais adequada para a penetração é como sendo um percentual da potência da subestação (fonte) que atende a rede.

Diferentemente do que ocorre para os consumidores, para a GDFV é mais apropriado dividir os agrupamentos pela proximidade, tendo em vista a alta dependência imposta pela fonte primária de geração de energia, a irradiação solar. Assim, não é relevante co-

Figura 30: Localização da GDFV no sistema IEEE de 13 barras.



Fonte: Adaptado de Anzanello (2016).

nhecer a quantidade de sistemas fotovoltaicos instalados em cada barra. Como a resposta ao estímulo da fonte primária é o mesmo para todas as GDFV na mesma barra, é como se houvesse uma única unidade geradora no local e basta definir a potência de pico do conjunto de sistemas de geração em cada barra. A potência de GDFV assim definida é o pico de geração que o conjunto de sistemas instalados em cada barra alcança, independente de quantos forem e de suas potências individuais.

Desse modo, os agrupamentos de GDFV são divididos da seguinte maneira:

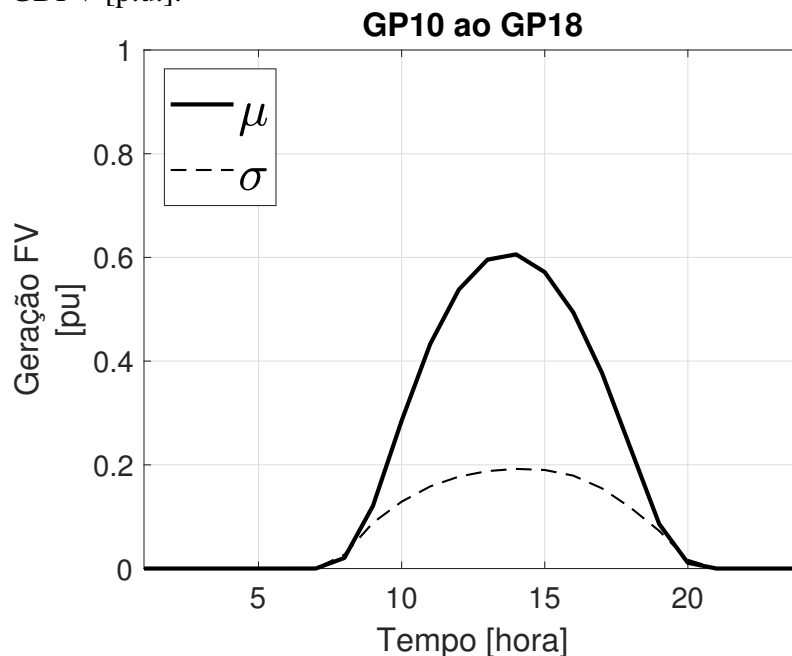
- GP10: GDFV agregada na Barra 670;
- GP11: GDFV agregada na Barra 671;
- GP12: GDFV agregada na Barra 645;
- GP13: GDFV agregada na Barra 646;
- GP14: GDFV agregada na Barra 692;
- GP15: GDFV agregada na Barra 675;
- GP16: GDFV agregada na Barra 611;
- GP17: GDFV agregada na Barra 652; e
- GP18: GDFV agregada na Barra 634.

Dado o percentual de penetração de GDFV do cenário, considera-se que a potência dos sistemas de GDFV esteja distribuída nas barras do sistema de forma proporcional à demanda das cargas no sistema original de 13 barras do IEEE.

Para a obtenção das curvas diárias típicas de produção de energia de sistemas de GDFV, foram compilados os dados obtidos de Pfenninger; Staffell (2016) desde o ano 2000 até 2016 para a geração em Porto Alegre considerando azimute de 180° e uma inclinação de 24°, sendo este o ângulo ótimo para a localidade (CEPEL, 2017). Assim,

a curva média (μ) e do desvio padrão (σ) de geração por horário são obtidas e ilustradas na Figura 31 para cada barra (agrupamento de GDFV). Cabe ressaltar que as curvas normalizadas da geração fotovoltaica em cada barra são idênticas³.

Figura 31: Curva de média e desvio padrão horários normalizados da geração dos agrupamentos de GDFV [p.u.].



Fonte: O autor.

Apesar de existir um modelo no OpenDSS[®] que representa o comportamento de sistemas de GDFV de forma adequada, como já comprovou Anzanello (2016), utilizou-se um modelo de gerador de potência constante. Uma vez que a distribuição horária de geração já havia sido obtida, não é preciso utilizar o modelo específico, evitando assim a necessidade de informações adicionais como temperatura ambiente (PFENNINGER; STAFFELL, 2016; EPRI, 2013). Além disso, os sistemas de GDFV são concebidos para injetar na rede toda a potência produzida pelos painéis, que é proporcional à irradiância, independentemente do nível de tensão ao qual estão conectados (LIU; BEBIC, 2008). Como não há previsão de compensação por injeção de potência reativa na norma, o fator de potência utilizado para os sistemas de GDFV foi unitário (ANEEL, 2012).

5.2 Simulações e Análise Probabilística

Nesta seção, é detalhado o procedimento de análise probabilística da operação aplicando o método proposto para a rede de testes selecionada.

5.2.1 Caracterização Estocástica dos Componentes do Sistema e Estrutura de Dependência

Uma vez que o sistema simulado é fictício, a Seção 5.1.1 e a Seção 5.1.2 apresentaram uma maneira alternativa de obter os agrupamentos dos componentes do sistema bem como a curva média e desvio padrão para a potência dos mesmos, informações essas que

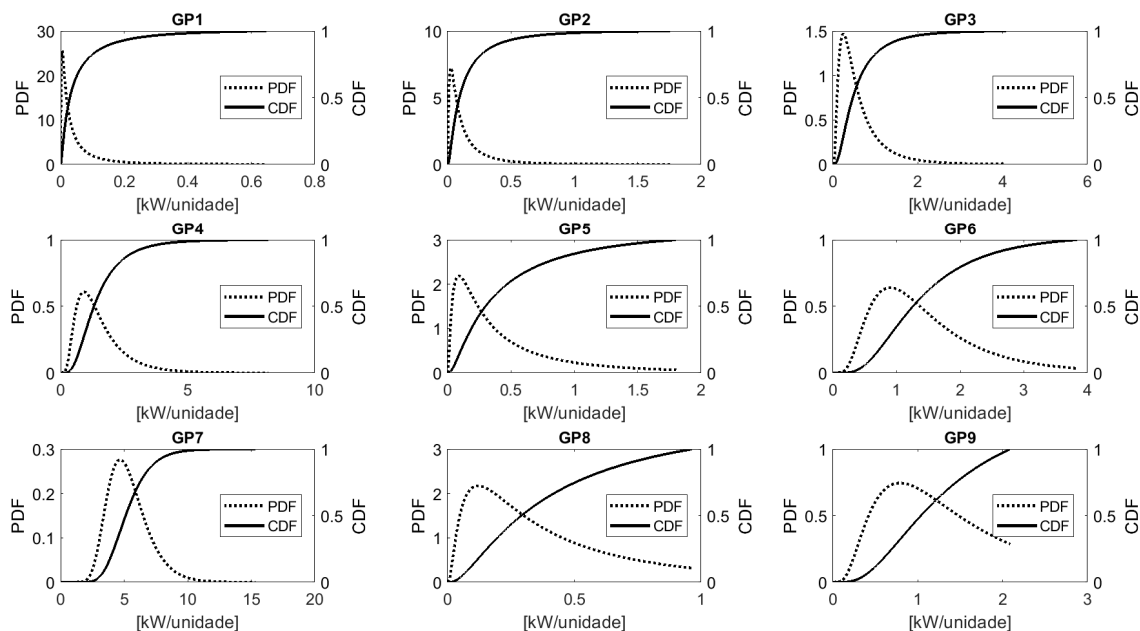
³ Isso porque a fonte primária para a geração de energia é a mesma. O que diferencia a geração em cada barra é a potência de pico dos sistemas da respectiva barra e a influência da fonte primária.

seriam provenientes da análise dos dados de entrada das cargas e da GD no fluxograma da Figura 21, caso dados reais estivessem disponíveis. Nessa situação, também seria necessário analisar qual a lei de distribuição que melhor se ajusta aos dados medidos de potência dos componentes e que permite obter a CDF mais adequada para a modelagem estocástica da demanda do respectivo componente para a hora simulada. Além do mais, seria necessário verificar a relação de dependência entre as potências dos agrupamentos (vide Seção 2.2.2) em cada hora do período simulado para que as simulações possam contemplar o uso de cenários com cargas e geração correlacionadas.

Nesse estudo de caso, a lei de distribuição lognormal foi utilizada para modelar estocasticamente a demanda horária dos consumidores enquanto a lei de distribuição beta foi utilizada para modelar a potência de geração horária dos sistemas de GDFV (PAREJA, 2009; ABDULKARIM; ABDELKADER; MORROW, 2015).

A Figura 29 ilustra as estatísticas de média e desvio padrão normalizados da demanda para diferentes agrupamentos de consumidores a partir dos dados das Tabelas 2 à 5. A multiplicação das estatísticas normalizadas pela potência de base, apresentada na Tabela 6 para cada agrupamento, permite a obtenção da média e do desvio padrão da demanda de cada agrupamento para os diferentes horários do dia. Estas podem ser utilizadas em combinação com a lei de distribuição para obter a curva CDF da demanda horária de cada componente do agrupamento, conforme indicado na Figura 24. Desse modo, a Figura 32 ilustra a PDF e a CDF da demanda em quilowatts por unidade consumidora no horário das 9h para os agrupamentos de carga conectados na Barra 692 da rede. A Figura 33 e a Figura 34 ilustram a PDF e a CDF para o horário das 14h e 21h respectivamente para os mesmos agrupamentos.

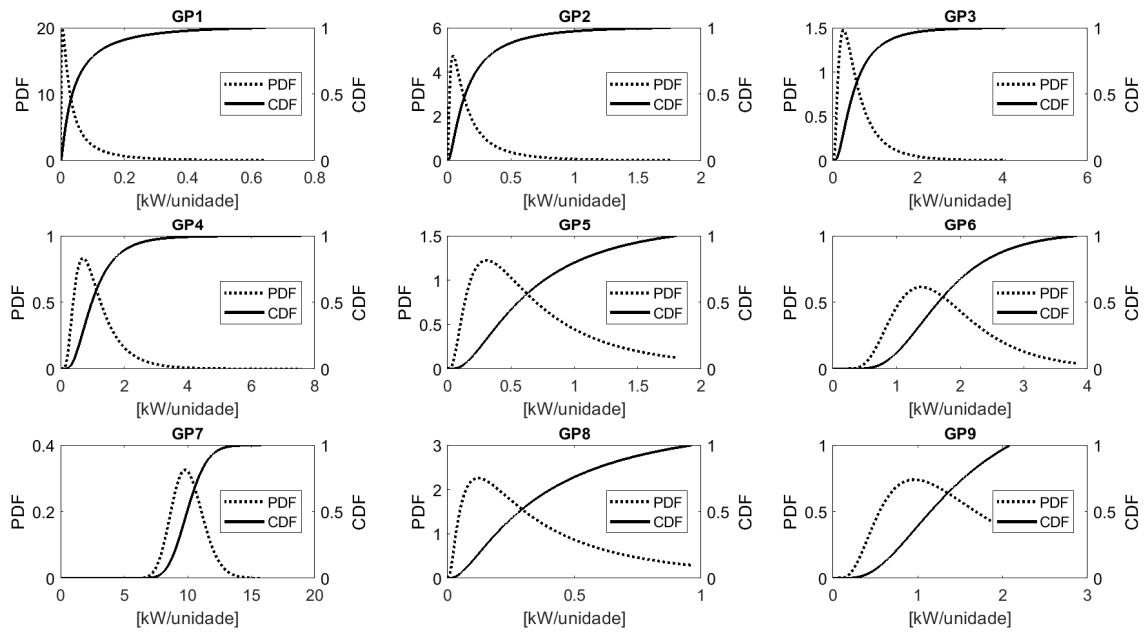
Figura 32: PDF e CDF da demanda às 9h dos agrupamentos de carga na Barra 692.



Fonte: O autor.

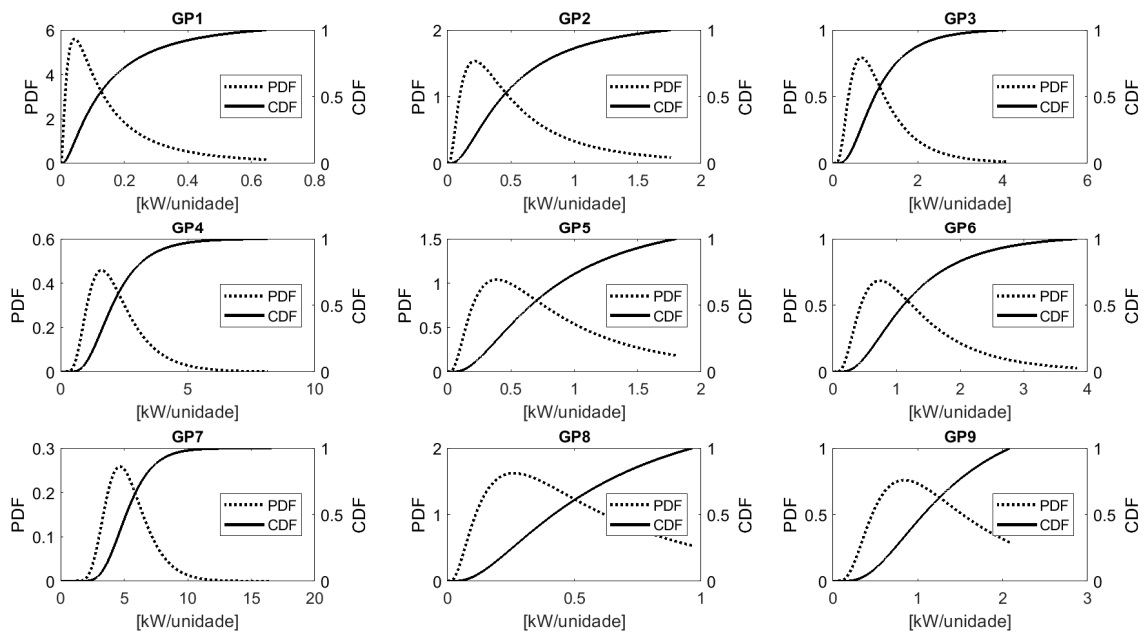
Na modelagem da CDF horária para a demanda dos agrupamentos é ainda aplicado um fator limitador, de modo a evitar valores amostrados extremamente elevados e que não condizem com a realidade. Essa necessidade é decorrente da modelagem da demanda ho-

Figura 33: PDF e CDF da demanda às 14h dos agrupamentos de carga na Barra 692.



Fonte: O autor.

Figura 34: PDF e CDF da demanda às 21h dos agrupamentos de carga na Barra 692.



Fonte: O autor.

rária com uma distribuição lognormal, cujo domínio é $[0; +\infty]$, o que por sua vez poderia resultar em valores de demanda amostrada tendendo a infinito. Sabe-se que existe um limite de demanda para cada consumidor que é imposto, em última instância, pela capacidade do dispositivo de proteção de entrada da instalação. Como não há tal informação para a rede simulada, a limitação foi imposta a partir de valores de fator de carga (FC) típicos para clientes residenciais, comerciais e industriais de uma distribuidora de energia (VILLELA; CARVALHO, 2007). A Tabela 7 apresenta o fator de carga e o fator de demanda (FD) típicos para clientes de diferentes classes de consumo.

Tabela 7: Fatores de carga e demanda típicos de clientes de redes de distribuição.

Classe consumidora	FC	FD
Residencial	0,1670	0,1957
Comercial	0,4000	0,3500
Industrial	0,7600	0,8000

Fonte: (VILLELA; CARVALHO, 2007).

Com a utilização do FC típico, a CDF de todas as cargas em cada horário simulado foi ajustada de forma que o valor unitário correspondesse à máxima demanda possível para aquela carga. Cabe ressaltar que a máxima demanda de uma carga é a mesma para todos os horários.

A proposta de limitar a CDF das cargas é reforçada quando se observa a curva real de um consumidor residencial em baixa tensão (BT), como é ilustrado na Figura 16, que mostra a distribuição acumulada da demanda entre as 20h às 21h para um cliente com consumo mensal de 200 a 400 kWh.

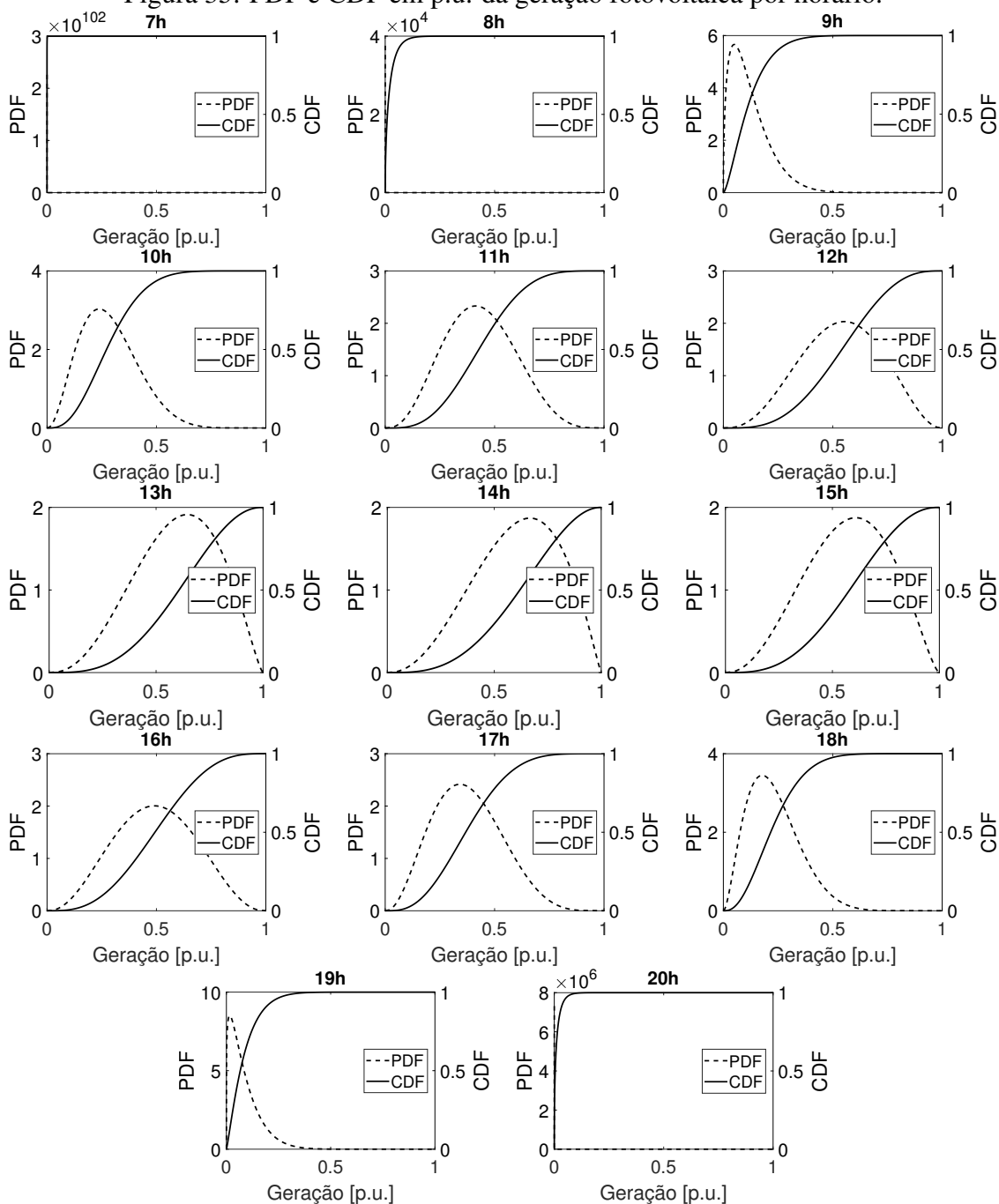
A demanda horária em cada barra do sistema é a agregação da demanda de cada agrupamento de clientes adjacentes àquela barra. A demanda de cada agrupamento de clientes, por sua vez, é resultante da amostragem na CDF do respectivo horário para aquele grupo, considerando a quantidade de clientes e a potência de base conforme Tabela 6.

O total de geração nas barras para cada horário é obtido de maneira semelhante à demanda. A geração horária de energia para sistemas fotovoltaicos segue lei de distribuição beta de probabilidade, com a média e desvio padrão horário de geração ilustrados na Figura 31. Os agrupamentos de GDFV possuem CDF de geração horária em p.u. idêntica e as mesmas são ilustradas na Figura 35 para os horários nos quais existe a geração de energia fotovoltaica (7 às 20h). Assim, para obter a potência gerada em uma barra, basta realizar a amostragem da geração em p.u. e multiplicar pela potência de pico que os sistemas fotovoltaicos daquela barra somam, i.e., a potência de base. A potência de pico dos sistemas fotovoltaicos em cada barra é proporcional à penetração de GD e à demanda da barra no cenário original do IEEE.

A modelagem do comportamento estocástico individual dos componentes se completa com a definição das curvas CDF da potência dos mesmos em todos os horários do período avaliado. A conclusão do Passo 1 do fluxograma para o Método Proposto na Figura 21 se dá com a obtenção da relação de dependência entre os componentes para cada horário simulado e construção da matriz \mathbf{R} de correlação, que contém a essência da relação multidimensional de interdependências. Cabe salientar que pode existir uma matriz \mathbf{R} para cada horário do período analisado.

A matriz de correlação linear \mathbf{R} é obtida da matriz de correlação de *rank* \mathbf{R}_r pela relação $\mathbf{R} = 2\text{sen}(\mathbf{R}_r\pi/6)$. Os componentes $\rho_r(i, j)$ da matriz \mathbf{R}_r são a correlação

Figura 35: PDF e CDF em p.u. da geração fotovoltaica por horário.



Fonte: O autor.

$$\mathbf{R}_r^{KK} = \begin{bmatrix} 1 & 0,8 & 0,8 & 0,8 & 0,8 & 0,8 & 0,8 & 0,8 & 0,8 \\ 0,8 & 1 & 0,8 & 0,8 & 0,8 & 0,8 & 0,8 & 0,8 & 0,8 \\ 0,8 & 0,8 & 1 & 0,8 & 0,8 & 0,8 & 0,8 & 0,8 & 0,8 \\ 0,8 & 0,8 & 0,8 & 1 & 0,8 & 0,8 & 0,8 & 0,8 & 0,8 \\ 0,8 & 0,8 & 0,8 & 0,8 & 1 & 0,8 & 0,8 & 0,8 & 0,8 \\ 0,8 & 0,8 & 0,8 & 0,8 & 0,8 & 1 & 0,8 & 0,8 & 0,8 \\ 0,8 & 0,8 & 0,8 & 0,8 & 0,8 & 0,8 & 1 & 0,8 & 0,8 \\ 0,8 & 0,8 & 0,8 & 0,8 & 0,8 & 0,8 & 0,8 & 1 & 0,8 \\ 0,8 & 0,8 & 0,8 & 0,8 & 0,8 & 0,8 & 0,8 & 0,8 & 1 \end{bmatrix}. \quad (47)$$

Para modelar a estrutura de interdependência estocástica, utilizam-se as relações de dependência compiladas na matriz \mathbf{R} juntamente com o método da JNT (vide Seção 2.4.3), gerando um vetor \mathbf{U} de números aleatórios com distribuição uniforme para cada VA do sistema. Os vetores \mathbf{U} contém intrinsecamente a relação de dependência entre os componentes do sistema. Essa etapa constitui o Passo 2 do fluxograma para o Método Proposto. Posteriormente, os vetores \mathbf{U} são utilizados na criação dos cenários de demanda e geração agregada nas barras através da amostragem da potência individual dos componentes.

O procedimento da JNT para o sistema simulado é dado pelos seguintes passos:

1. Gera um vetor \mathbf{z} de dimensão 18×1 com elementos que são amostras independentes de uma curva normal padrão de probabilidade;
2. Aplica a transformação linear ao vetor \mathbf{z} tal que $\mathbf{y} = \mathbf{L} \times \mathbf{z}$, sendo \mathbf{L} a matriz triangular inferior da decomposição de Cholesky da matriz \mathbf{R} . Assim, os números aleatórios independentes em \mathbf{z} passam a estar correlacionados em \mathbf{y} pela correlação proveniente da matriz \mathbf{R} ;
3. A aplicação da função de distribuição acumulada da normal padrão $\Phi(\cdot)$ aos números aleatórios correlacionados \mathbf{y} , resulta em 18 números aleatórios na forma $\mathbf{u} = \Phi(\mathbf{y})$ pertencentes, cada um, a uma distribuição uniforme \mathbf{U} .

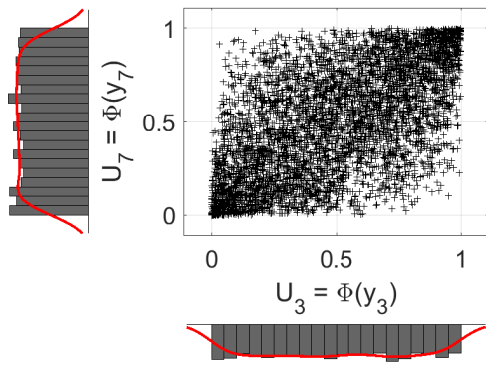
A sequência é repetida para todos horários do período avaliado a cada amostra. Logo, cada elemento $\mathbf{u}(i)$ é correspondente ao respectivo agrupamento i de componentes do sistema. Ao atingir a convergência, terão sido gerados, para cada horário do período simulado, 18 vetores \mathbf{U}_i de dimensão $1 \times (\text{número de amostras})$ cujos elementos $\mathbf{u}(i)$ formam distribuições uniformes, estando os vetores correlacionados entre si e associados ao respectivo agrupamento i de componentes do sistema.

A Figura 36 ilustra a dispersão entre os vetores uniformes \mathbf{U}_i (associados aos GP i) gerados para os horários das 9h e 14h para amostragem dos componentes pertencentes aos agrupamentos GP3, GP7, GP14 e GP18. Cada vetor \mathbf{U}_i possui 5000 números aleatórios.

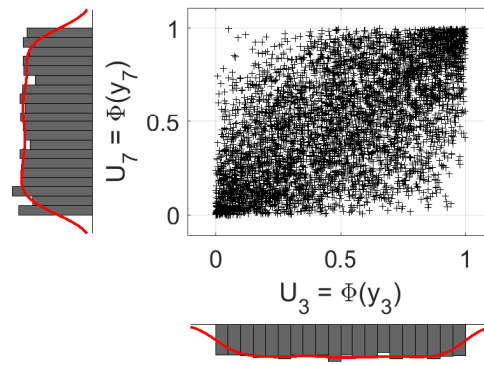
É possível observar a concentração dos pontos dos gráficos à medida que aumenta a relação entre os diferentes agrupamentos. Como uma única matriz \mathbf{R} foi utilizada, não há diferença perceptível entre diferentes horários, sendo as nuances que eventualmente ocorrem decorrentes de um número limitado de amostras. Também fica evidente que, apesar de o número aleatório inicialmente gerado seguir uma distribuição normal de probabilidade, o resultado da JNT é uma distribuição Uniforme e a mesma pode ser empregada na amostragem da potência dos componentes do sistema, independente da lei de distribuição que estes seguem, conforme técnica da Figura 9.

Tendo realizado os procedimentos detalhados anteriormente, o Passo 2 do fluxograma para o Método Proposto é cumprido. Os números aleatórios correlacionados \mathbf{U}_i sorteados

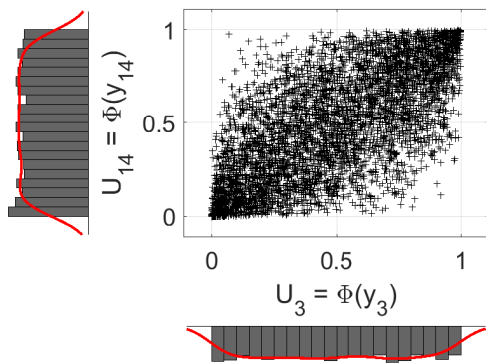
Figura 36: Dispersão entre U_i gerados para amostragem da potência dos componentes dos GPi às 9h e 14h.



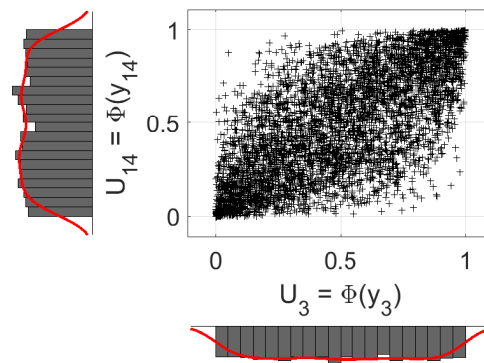
(a) Dispersão entre U_3 e U_7 para as 9h.



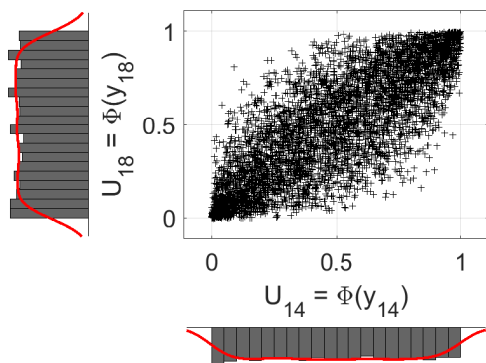
(b) Dispersão entre U_3 e U_7 para as 14h.



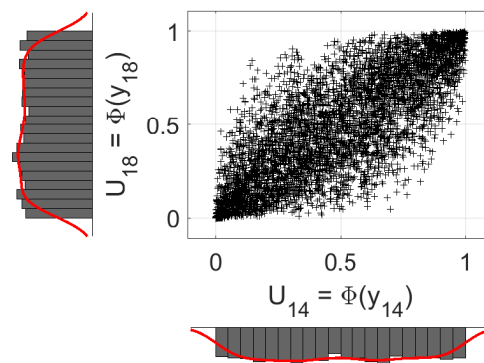
(c) Dispersão entre U_3 e U_{14} para as 9h.



(d) Dispersão entre U_3 e U_{14} para as 14h.



(e) Dispersão entre U_{14} e U_{18} para as 9h.



(f) Dispersão entre U_{14} e U_{18} para as 14h.

Fonte: O autor.

são então utilizados para realizar a amostragem na CDF inversa da potência dos componentes para todas as amostras e horários simulados, sendo esse aspecto abordado a seguir.

5.2.2 Amostragem e Composição de Cenários de Demanda e GDFV

Após definidas as curvas CDF para a potência de todos os agrupamentos de componentes do sistema em cada horário do período avaliado, é possível determinar a demanda agregada em cada nó da rede, obtendo assim, amostras que representam o consumo e a geração total no sistema com enfoque probabilístico.

A determinação dos cenários da demanda agregada e geração distribuída em cada nó é realizada pela sequência dos Passos 3, 4 e 5 do fluxograma para o Método Proposto e pode ser resumida pela Figura 26 e pela Figura 27 do Capítulo 4. A etapa ilustrada na Figura 26 corresponde aos Passos 3 e 4, ou seja, é a amostragem da potência dos componentes da rede em todos os nós com a utilização da CDF inversa e dos números aleatórios U_i sorteados anteriormente para cada horário simulado, conforme Equação (22).

Uma vez amostrada a potência dos componentes, procede-se à agregação das potências nas barras, conforme Figura 27. Para tal, considera-se a nomenclatura passiva, ou seja, injeção de potência tem valor negativo (GDFV). Os cenários amostrados para a demanda dos consumidores são os mesmos nas configurações com ausência e penetração de GDFV. Isso foi assegurado através do controle da semente do pseudo gerador aleatório do programa MATLAB[®], permitindo assim um comparativo entre os indicadores de operação da rede na presença de GD.

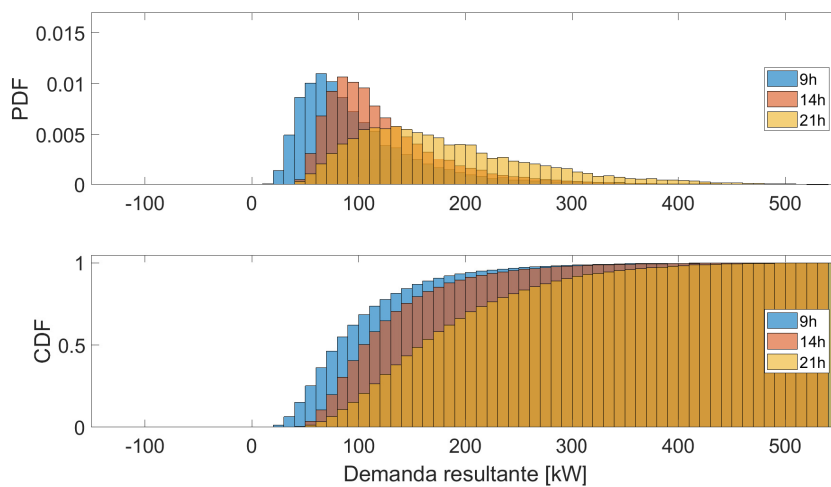
Sendo as demandas dos componentes pertencentes ao mesmo agrupamento obtidas com o mesmo número aleatório (uso da SBM), a demanda agregada em cada barra é resultante da soma das demandas individuais dos clientes para cada uma das N amostras. Cada amostra é composta por 24 valores (número de horários do período avaliado), caracterizando uma curva diária possível para o respectivo componente. Dessa forma, a PDF da demanda agregada na Barra 692, por exemplo, após realizada a amostragem da potência dos agrupamentos e a agregação da demanda e geração, é dada pela Figura 37 para configurações de penetração de GDFV distintas no horário das 9, 14 e 21h com base em $N = 5000$ amostras.

Nos casos em que há penetração de GDFV, a demanda líquida resultante na barra será a diferença entre a demanda dos clientes com o montante de geração dos sistemas distribuídos. Desconsiderando o comportamento econômico dos prosumidores que poderia fazê-los alterar o padrão de consumo após a instalação de um sistema de geração distribuída, em princípio a demanda na barra se mantém independente do nível de penetração de GD. No gráfico (b) da Figura 37 para 50% de penetração é pouco perceptível, mas no gráfico (c) para 100% de penetração de GDFV fica evidente a ocorrência de fluxo reverso de potência na Barra 692.

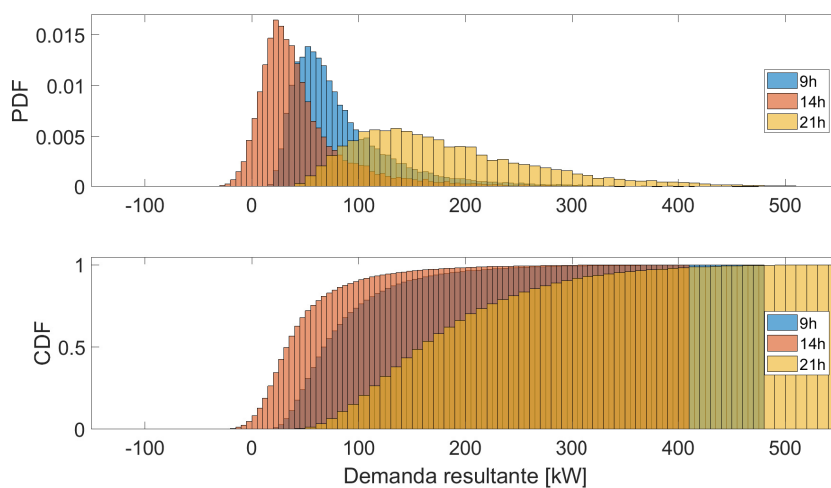
As curvas PDF e CDF de potência gerada pelos sistemas fotovoltaicos na Barra 692 para as configurações com 50% e 100% de penetração de GD são ilustradas na Figura 38. O horário das 14h é o de geração média mais elevada e por isso somente são ilustradas as curvas para os horários entre 9 e 14h. Cabe destacar que as curvas para 50% e 100% tem o mesmo formato, alterando-se apenas a escala.

A combinação adequada das amostras de demanda nas curvas CDF relativas às curvas PDF da Figura 37(a) com as amostras de demanda nas curvas CDF de geração da Figura 38(a) resultam nos cenários que originam a demanda agregada da Figura 37(b). Da mesma forma, a combinação adequada das amostras de demanda nas curvas CDF relativas às curvas PDF da Figura 37(a) com as amostras de demanda nas curvas CDF de geração

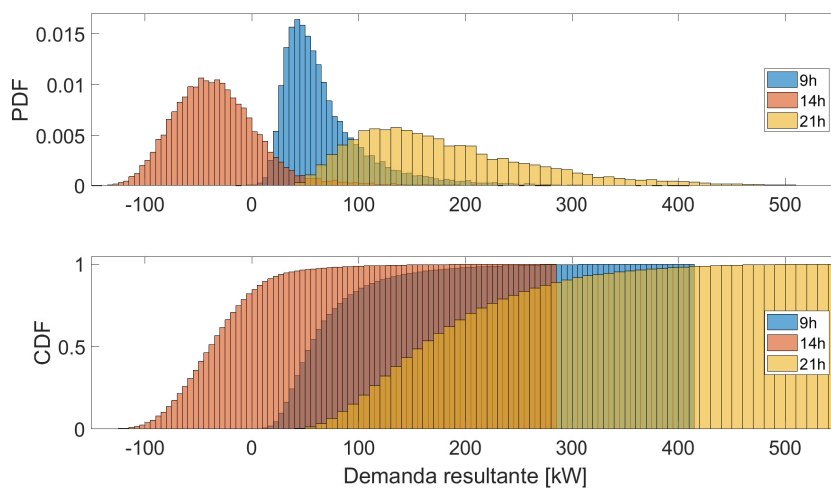
Figura 37: Demanda agregada na Barra 692 em três horários para diferentes configurações de GDFV.



(a) Ausência de GDFV.



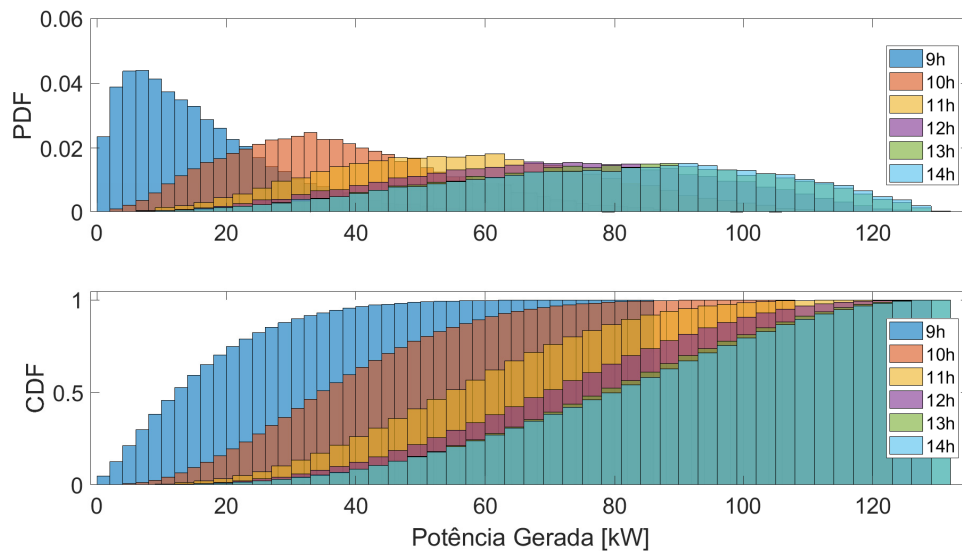
(b) Penetração de 50% de GDFV



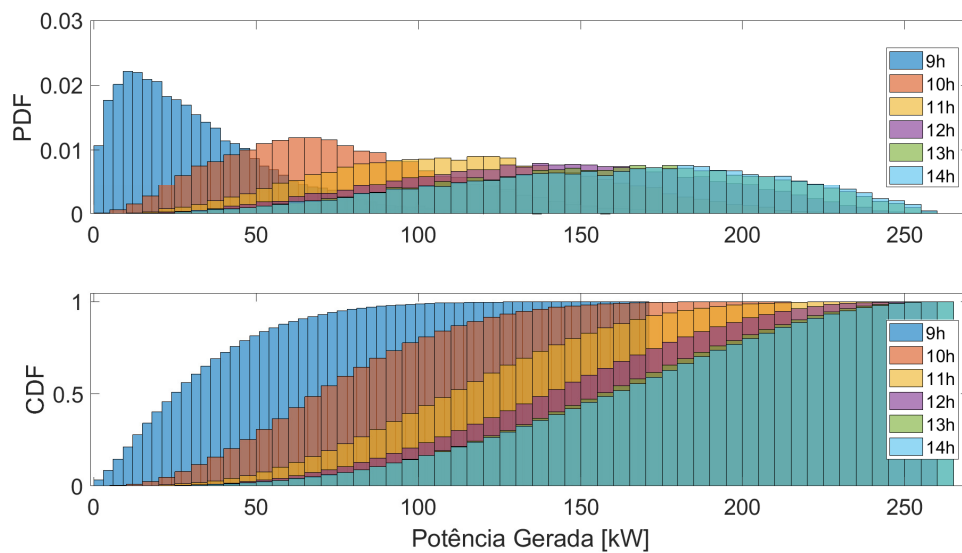
(c) Penetração de 100% de GDFV

Fonte: O autor.

Figura 38: GDFV na barra 692 para diferentes níveis de penetração de GD e N=5000 amostras.



(a) Penetração de 50% de GDFV.



(b) Penetração de 100% de GDFV

Fonte: O autor.

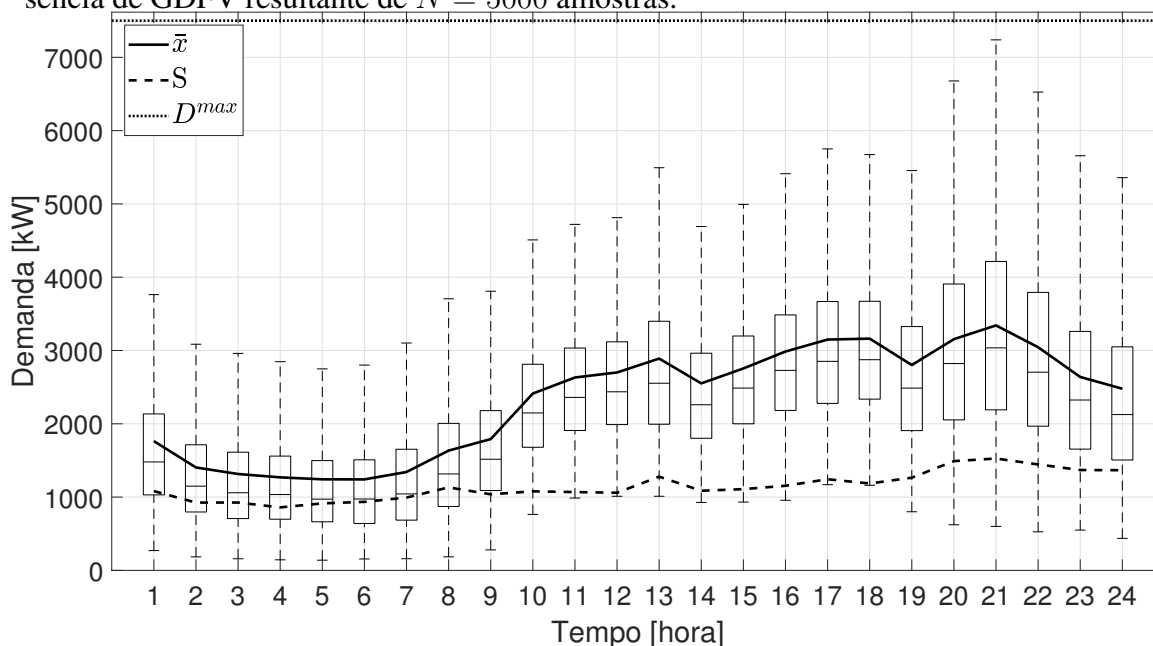
da Figura 38(b) resultam nos cenários que originam a demanda agregada da Figura 37(c).

Na criação dos cenários, foi considerada uma memória arbitrária $M = 10\%$ para o acoplamento temporal entre o comportamento da demanda horária de cada cliente individualmente (vide Seção 4.3.1). A falta de dados reais inviabilizou a investigação minuciosa deste aspecto. Assim, um valor baixo permite não ignorar esse efeito ao mesmo tempo que não compromete os resultados.

A Figura 37 e a Figura 38 ilustram as curvas relativas à demanda e geração agregadas em um dos nós da rede a título de exemplo. Esse mesmo procedimento é replicado em todas as outras barras do sistema para todos os horários do período avaliado.

Como resultado, a Figura 39 ilustra a curva diária probabilística da demanda agregada no alimentador na forma de diagramas de caixa e linha para cada horário do período avaliado, considerando ausência de GDFV e $N = 5000$ amostras. A curva de demanda média (\bar{x}) é a condição que costuma ser avaliada em uma análise determinística. Percebe-se, no entanto, que existe uma expressiva dispersão no entorno do valor médio, sugerindo a necessidade do emprego de um método alternativo para avaliar com maior fidelidade os limites dos indicadores de operação desta rede. É ainda ilustrada a curva do desvio padrão (s) da demanda agregada no alimentador.

Figura 39: Comportamento probabilístico da demanda agregada no alimentador com ausência de GDFV resultante de $N = 5000$ amostras.

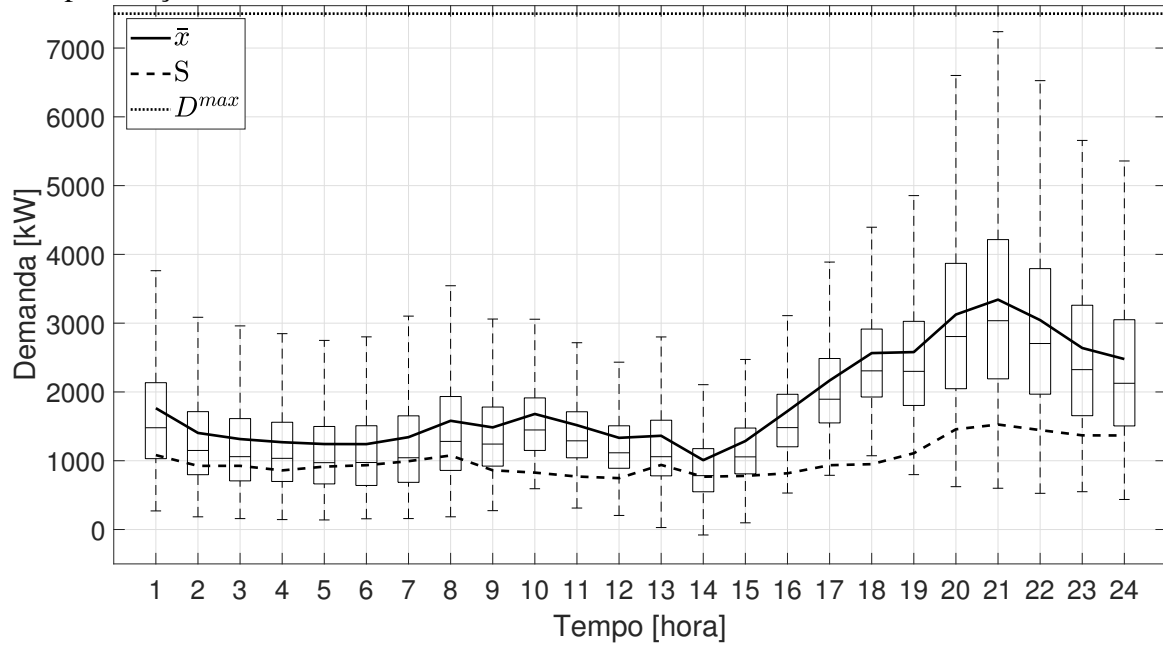


Fonte: O autor.

Também é possível observar que, apesar de a demanda média agregada no horário de pico estar em torno de 3,3 MW, existem amostras nas quais a mesma alcança valores próximos a 7,5 MW. A ABNT (1997) estabelece que o limite de carregamento para transformadores não deve ultrapassar o limiar de 150% da potência nominal em qualquer situação. Isso evidencia que equipamentos operando aparentemente com folga podem estar, na verdade, próximos ao limite normativo de operação, já que a potência nominal do transformador da rede simulada é de 5MVA.

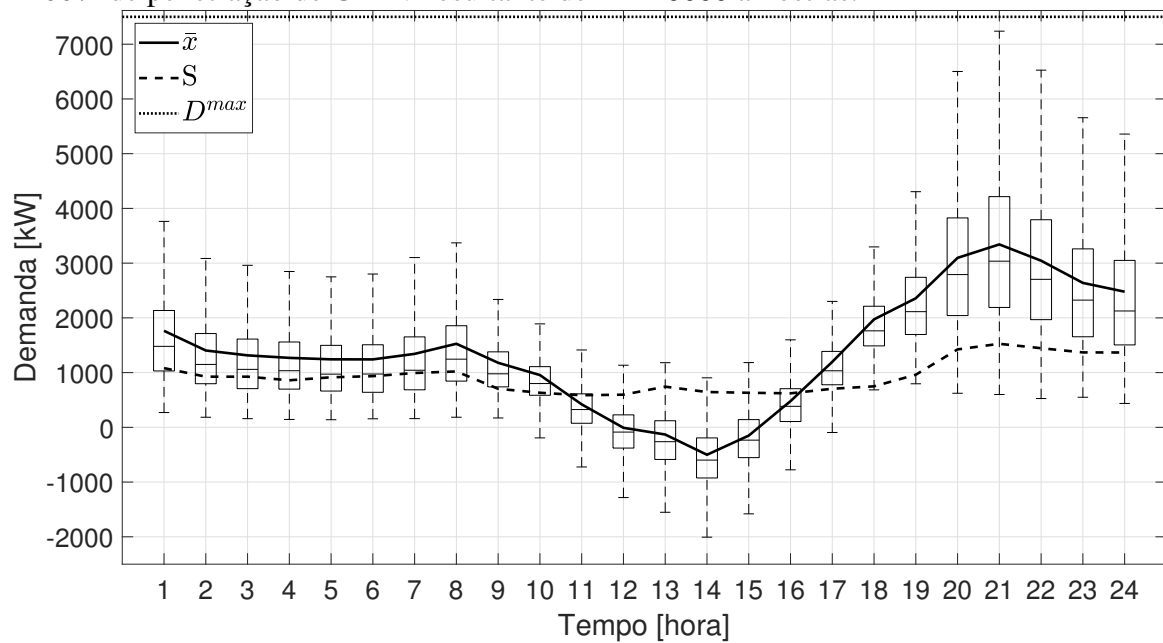
A Figura 40 ilustra a curva diária probabilística da demanda agregada no alimentador para a configuração com 50% de penetração de GDFV enquanto a Figura 41 ilustra o caso com 100% de penetração de GDFV.

Figura 40: Comportamento probabilístico da demanda agregada no alimentador com 50% de penetração de GDFV resultante de $N = 5000$ amostras.



Fonte: O autor.

Figura 41: Comportamento probabilístico da demanda agregada no alimentador com 100% de penetração de GDFV resultante de $N = 5000$ amostras.



Fonte: O autor.

Percebe-se a clara redução na demanda agregada provocada pela penetração da GDFV. No entanto, esta não contribui para a redução da demanda de pico sistêmica devido a ausência de dispositivos de armazenamento de energia. Tais dispositivos não foram contemplados nas simulações em função da legislação atual não prever explicitamente incentivos para isso (ANEEL, 2012). Assim, toda a produção se concentra nos horários de disponibilidade da fonte solar, conforme ilustra a Figura 42, para a configuração com penetração de 50% de GDFV, e a Figura 43, para 100% de penetração. Caso existisse outra fonte de GD no sistema, eventualmente poderia ocorrer compensação no horário de pico de forma natural.

5.2.3 Convergência das simulações

Um aspecto importante acerca de simulações estocásticas diz respeito ao momento de parada, ou seja, a quantidade de amostras N a partir da qual uma nova amostragem não influencia significativamente os indicadores de saída. Uma vez atingido o número de amostras obtidas por método específico, nada impede que mais amostras sejam sorteadas para conferir aumento no nível de confiança estatística aos indicadores. A avaliação da convergência dos indicadores, expressa pela Equação (43), é realizada neste trabalho para as seguintes grandezas:

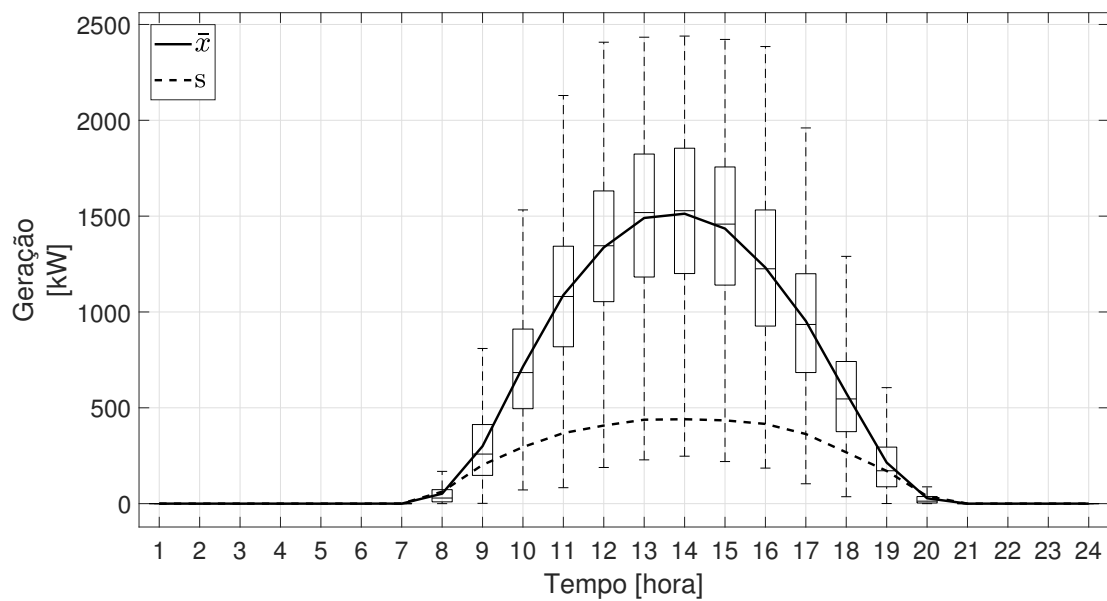
- demanda horária média em cada barra do sistema;
- perdas diárias totais do sistema;
- índices de violação de tensão crítica (DRC) diária em cada barra do sistema e
- os índices de violação de tensão precária (DRP) diária em cada barra do sistema.

Dependendo da estrutura de interdependência estocástica entre as variáveis, o valor típico de 5% para o coeficiente de variação estatístico C_v pode não ser suficiente e a decisão pelo percentual adequado é apoiada graficamente, conforme exemplo ilustrado na Figura 44. Se o critério de 5% fosse utilizado, a convergência da demanda horária média teria ocorrido para 45 amostras, quantidade inferior à necessária e que levaria à consideração de um valor limite equivocado para a demanda no nó. Já com um $C_v = 1,5\%$, foram necessárias 3028 amostras para a convergência de todos os indicadores, como perdas e violações de tensão, detalhados a seguir.

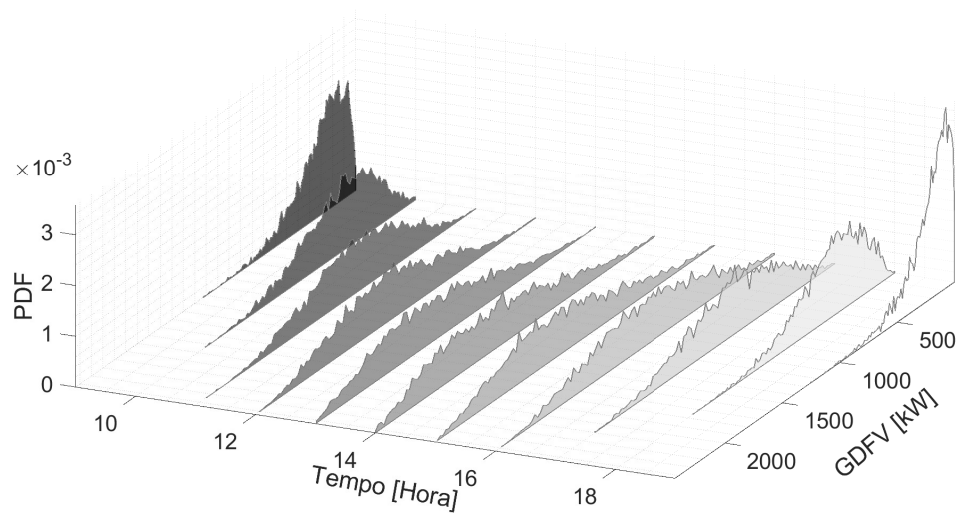
À medida que amostras são realizadas, as variáveis tendem a apresentar uma dinâmica de convergência semelhante. Contudo, naturalmente, alguns indicadores têm evolução mais rápida que outros. Tal aspecto pode ser observado quando se compara a convergência de diferentes indicadores. No caso da energia de perdas diárias no sistema, de acordo com a Figura 45, a convergência ocorre com 345 amostras para $C_v = 1,5\%$ e ausência de GDFV. Também são ilustradas na Figura 45 as convergências nas configurações com 50% e 100% de penetração de GDFV, que ocorrem com 429 e 508 amostras, respectivamente.

Em contrapartida, os índices diários de violação de tensão crítica DRC e precária DRP na Barra 675, por exemplo, exigem 230 e 619 amostras, respectivamente, para atender ao $C_v = 1,5\%$. As Figuras 46(a) e 46(b) ilustram esse processo de convergência para a configuração com ausência de GDFV, sendo possível constatar que a média para as violações de tensão crítica na rede tende a 0,68% enquanto a média para as violações de tensão precária tende a 0,37%, valores inferiores aos limites normativos tolerados, representados por $DRC^{max} = 3\%$ e $DRP^{max} = 0,5\%$ (ANEEL, 2018). São ilustradas, ainda, nas Figuras 46(c) a 46(f), as convergências do DRC e DRP para as configurações com

Figura 42: Comportamento probabilístico da GDFV no sistema com 50% de penetração para $N = 5000$ amostras.



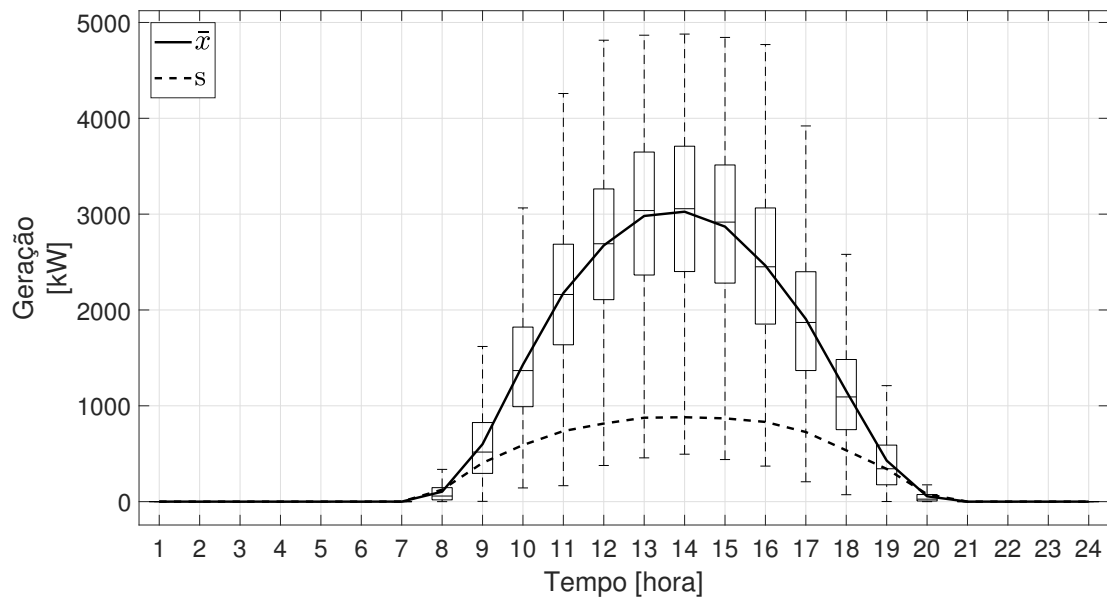
(a) Diagrama de caixa e linha da GDFV no sistema.



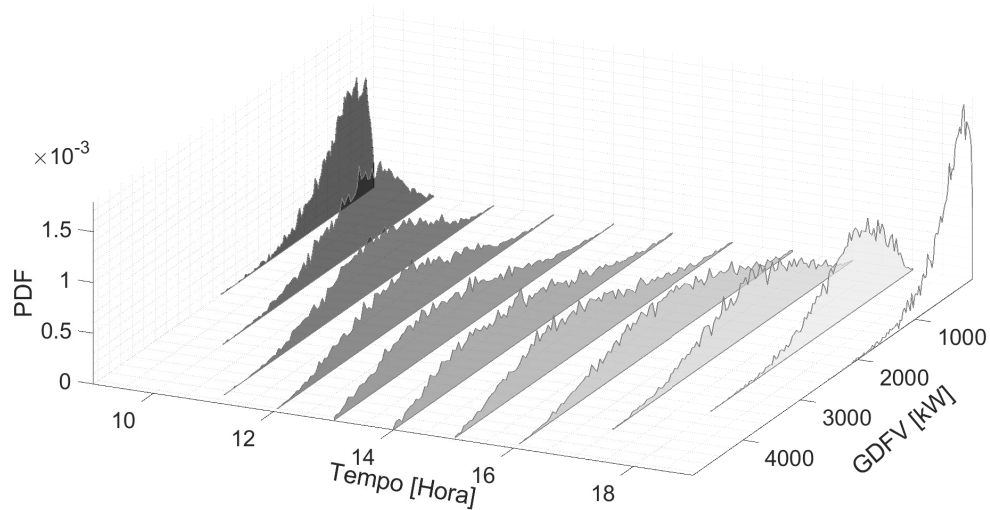
(b) Representação espacial da PDF horária da GDFV no sistema.

Fonte: O autor.

Figura 43: Comportamento probabilístico da GDFV no sistema com 100% de penetração para $N = 5000$ amostras.



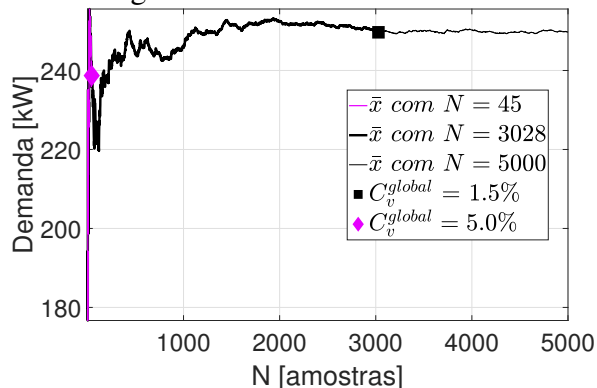
(a) Diagrama de caixa e linha da GDFV no sistema.



(b) Representação espacial da PDF horária da GDFV no sistema.

Fonte: O autor.

Figura 44: Processo de convergência da demanda média na Fase A da Barra 675 às 21h.



Fonte: O autor.

50% e 100% de penetração de GDFV. Por oportuno, observa-se na Figura 46(d) que o coeficiente de variação para o DRP é também o global para esta simulação.

A quantidade de amostras mínima na configuração com 50% de GDFV foi de $N = 765$ enquanto que para 100% de penetração de GDFV foi de $N = 790$. É importante observar que não necessariamente maiores níveis de penetração de GDFV implicam em mais amostras para a convergência. Neste estudo, a configuração com 50% de GDFV apresentou convergência mais rápida do que o caso sem GDFV. Com penetração de 100%, a quantidade de amostras volta a aumentar, porém ainda é inferior ao caso sem a presença de GDFV (quando $N = 3028$ amostras). Conforme investigou Shayani (2010), há um limite de penetração de GD, na condição existente da rede, para manter a operação dentro dos limites normativos. Conforme ilustra a Figura 40, a partir de 50% de penetração começa a existir fluxo reverso em alguns momentos do dia. Esse fato, pode contribuir para aumentar a dispersão dos indicadores e, conseqüentemente, implica a necessidade de utilizar mais amostras para determinar os parâmetros de interesse.

De toda forma, para padronizar os resultados, prosseguiu-se com a realização de 5000 amostras, sendo possível visualizar a evolução dos indicadores após atingida a parada da amostragem pelo critério adotado. As curvas de convergência ilustradas nesta seção para os indicadores representam a média das densidades transitórias para os mesmos, da mesma maneira como ocorre na Figura 10 para uma VA genérica.

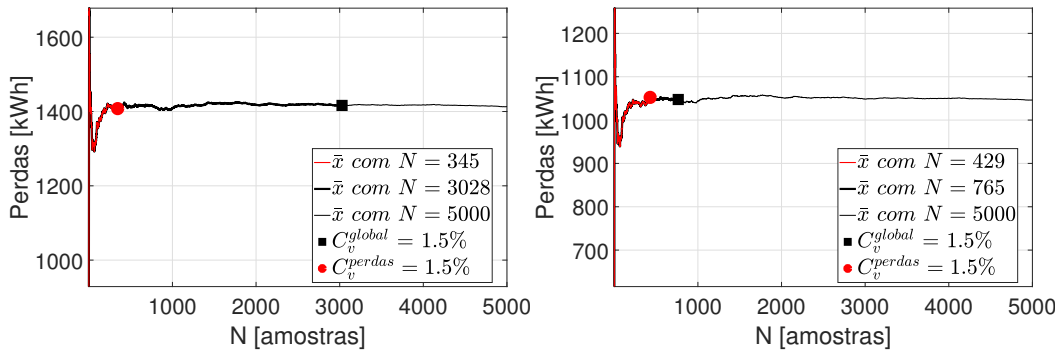
5.2.4 Indicadores Probabilísticos de Operação

A análise probabilística do desempenho de sistemas de distribuição de energia elétrica inclui possibilidades operacionais ignoradas em uma abordagem determinística clássica, enriquecendo o estudo. Desta maneira, a seguir são apresentados os indicadores probabilísticos das perdas ativas e dos níveis de tensão do sistema teste. Os mesmos são ainda comparados aos indicadores obtidos pela abordagem determinística.

5.2.4.1 Perdas Ativas

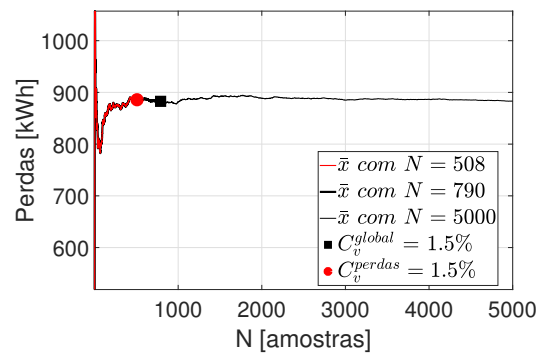
Na abordagem determinística, as perdas são consideradas fixas em cada instante de tempo e apresentam uma curva diária com formato similar ao da curva de perdas média (\bar{x}) ilustrada na Figura 47(a). As perdas nessa situação são provenientes do fluxo de potência para o cenário típico de demanda e geração distribuída. Assim, as curvas de demanda agregada média (\bar{x}) da Figura 39, 40 e 41 bem como as curvas de geração agregada média (\bar{x}) da Figura 42 e 43 foram utilizadas para determinar os cenários de demanda e geração

Figura 45: Processo de convergência da energia de perdas diárias no sistema.



(a) Evolução das perdas na ausência de GDFV.

(b) Evolução das perdas com penetração de 50% de GDFV.



(c) Evolução das perdas com penetração de 100% de GDFV.

Fonte: O autor.

distribuída horária em cada barra para cada configuração de penetração de GDFV da rede. Considerou-se ainda para isso a demanda agregada e a proporção das cargas do cenário original do IEEE. Os resultados da integração das curvas de perda diária correspondentes às configurações com ausência de GDFV bem como penetração de 50% e 100% de GDFV são apresentados na Tabela 8.

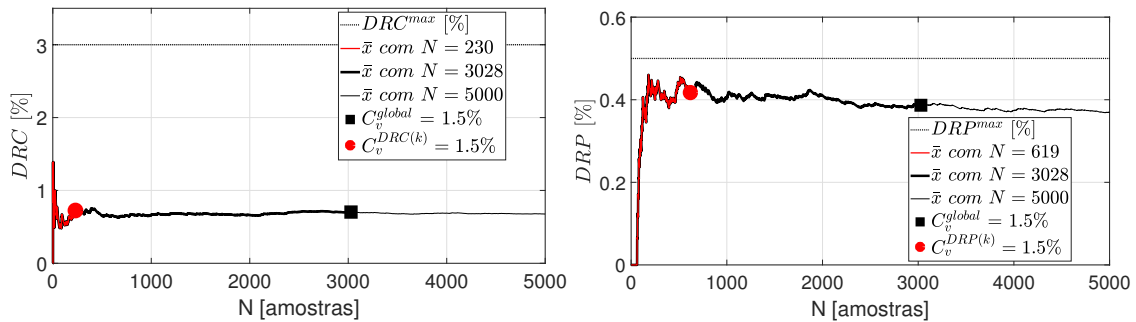
Tabela 8: Valores de perda diária resultantes de uma análise determinística.

Penetração de GDFV	Perdas [kWh]
0%	1193,4
50%	779,2
100%	644,5

Por outro lado, na abordagem probabilística, as perdas ativas são representadas por uma estimativa com intervalo de confiança (IC), em decorrência da atribuição de incertezas nos componentes da rede. Neste caso, as estatísticas de dispersão das perdas ativas complementam a curva de carga média na Figura 47 para cada horário do dia.

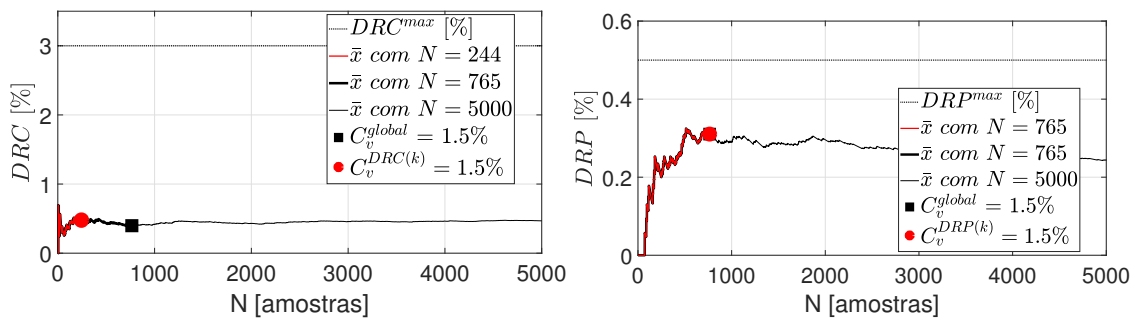
A análise dos dados obtidos nas simulações possibilitam realizar inferências estatísticas para a perda no sistema, obtendo a média \bar{x}_{Perdas} das amostras de perdas com a Equação (15) e o desvio padrão s_{Perdas} com a Equação (16). Também é possível obter um intervalo de confiança para os valores de perda diária IC_{Perdas} em diferentes configurações de penetração de GDFV, conforme apresenta a Tabela 9 para um nível de confiança

Figura 46: Processo de convergência do DRC e DRP na barra 675.



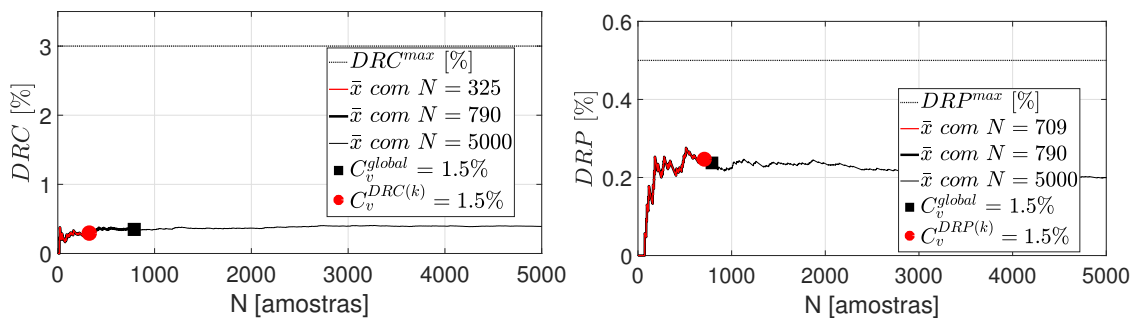
(a) Evolução do DRC na ausência de GDFV.

(b) Evolução do DRP na ausência de GDFV.



(c) Evolução do DRC com penetração de 50% de GDFV.

(d) Evolução do DRP com penetração de 50% de GDFV.

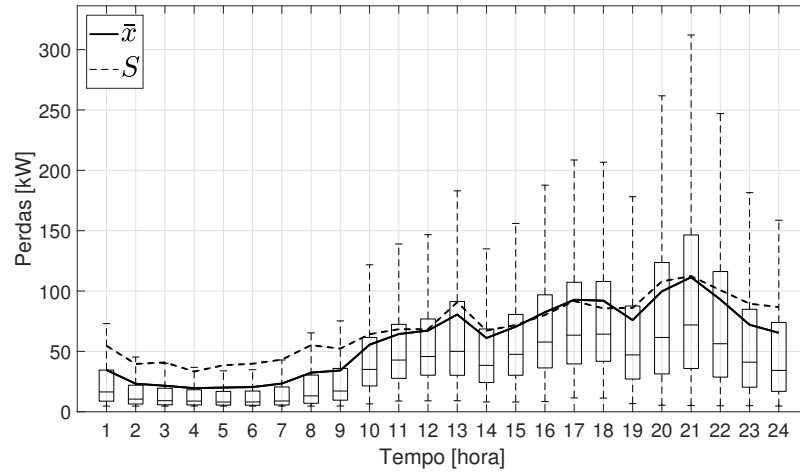


(e) Evolução do DRC com penetração de 100% de GDFV.

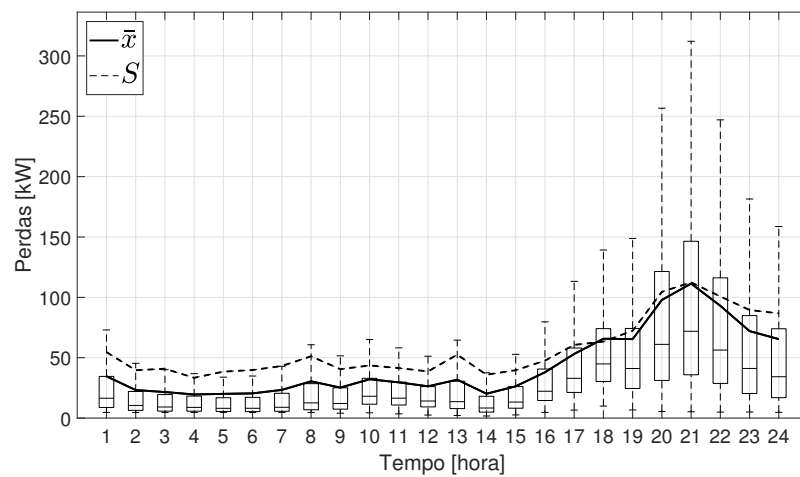
(f) Evolução do DRP com penetração de 100% de GDFV.

Fonte: O autor.

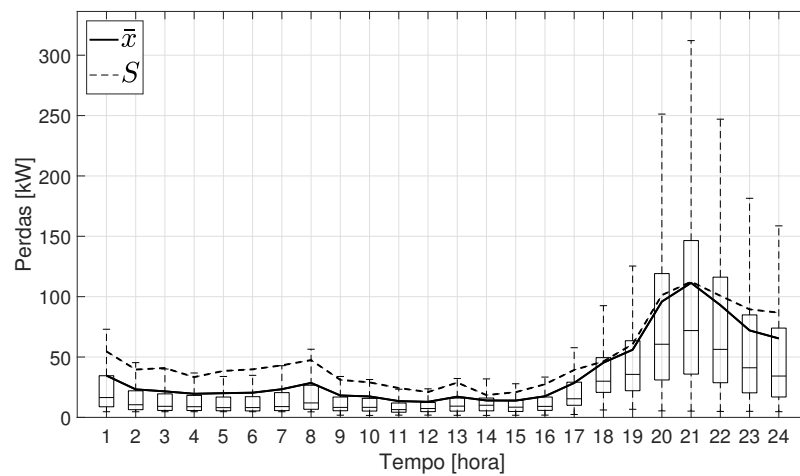
Figura 47: Comportamento probabilístico da potência de perdas horária.



(a) Configuração com ausência de GDFV.



(b) Configuração com 50% de GDFV.



(c) Configuração com 100% de GDFV.

Fonte: O autor.

de 95% com base em $N = 5000$ amostras.

Tabela 9: Valor médio, desvio padrão e IC da perda diária resultantes de uma análise probabilística.

Penetração de GDFV	\bar{x}_{Perdas} [kWh]	s_{Perdas} [kWh]	IC_{Perdas} [kWh]
0%	1413,4	385,5	[809,9 ; 2305,1]
50%	1046,6	323,5	[549,2 ; 1806,6]
100%	883,5	292,7	[446,3 ; 1571,8]

Pela comparação entre os valores de média das perdas da Tabela 8 com os da Tabela 9 é possível perceber que a análise determinística oferece valores de perda para o sistema inferiores aos que são obtidos pela análise de várias possibilidades operativas da rede. Para a configuração sem contribuição de GDFV, o valor obtido para a perda diária pela análise probabilística é 18,4% superior ao obtido pela análise determinística enquanto que com 50% e 100% de penetração os valores calculados probabilisticamente são 34,3% e 37,1% superiores aos determinísticos, respectivamente.

Com os resultados supracitados fica evidente a maior riqueza de informações proporcionada por uma análise probabilística. Na prospecção de opções de investimentos de expansão de uma rede, o patamar de perdas pode ser fator decisivo para rejeitar a execução de uma obra. A decisão apoiada pelos valores do indicador calculados probabilisticamente asseguram maior assertividade na definição da alternativa que minimiza as perdas técnicas e, conseqüentemente, econômicas.

Além disso, a integralização da potência de perdas horárias para os cenários simulados em cada configuração do sistema permite obter, com o apoio de métodos de adequação de ajuste de curvas (vide Seção 2.3.4), a PDF e a CDF da energia de perda diária na rede, conforme ilustra a Figura 48. Com as curvas CDF, é possível estimar, pelo uso da função sobrevivência (2), as chances de haver perda no sistema superior a um determinado valor. Assim, na ausência de GDFV, a probabilidade das perdas diárias serem superiores a 1500 kWh, por exemplo, é de 28,12%. Com a penetração da GDFV, a probabilidade reduz para 7,28% e 2,76%, para 50% e 100% de penetração, respectivamente.

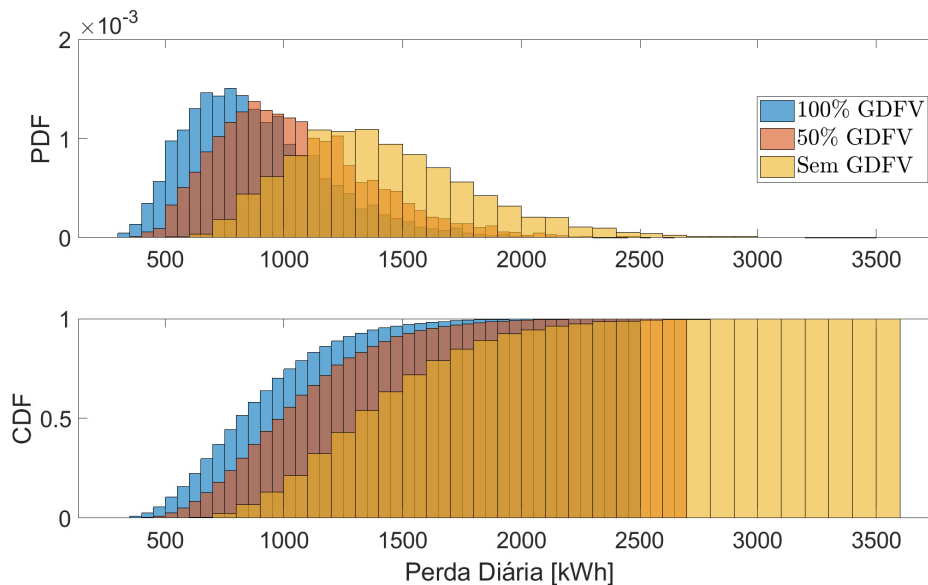
5.2.4.2 Violação de Tensão

O cálculo do fluxo de potência probabilístico evidencia a limitação do método determinístico na avaliação dos índices de violação de tensão. Ao considerar apenas os valores médios de carga, esta abordagem explora em menor profundidade os limites alcançados pelos níveis de tensão, ocultando inúmeras possibilidades operacionais.

A avaliação dos níveis de tensão com uma abordagem determinística é realizada para os cenários de carga leve, média e pesada obtidos da curva de duração da demanda do sistema, conforme ilustrado na Figura 13(c) da Seção 3.1. O perfil de tensão à jusante do regulador de tensão é ilustrado na Figura 49 para uma configuração com ausência de GDFV e diferentes carregamentos. As curvas representam a tensão em cada fase ao longo do alimentador bem como os limites normativos de tensão adequada, i.e., 1,05 p.u. e 0,95 p.u. (ANEEL, 2018).

Por ilustrar o perfil de tensão a partir do regulador, cuja atuação ocorre em cada fase individualmente, a magnitude da mesma no início do alimentador é distinta para cada fase devido, principalmente, ao desequilíbrio das cargas. Esses resultados estão de acordo com os obtidos por Freitas (2015), cujo código fonte de descrição da rede de 13 barras do IEEE

Figura 48: PDF e CDF da perda diária para diferentes configurações de penetração de GDFV.



Fonte: O autor.

no OpenDSS foi utilizado no desenvolvimento desta dissertação. Freitas (2015) validou os resultados obtidos no OpenDSS com valores de referência do IEEE.

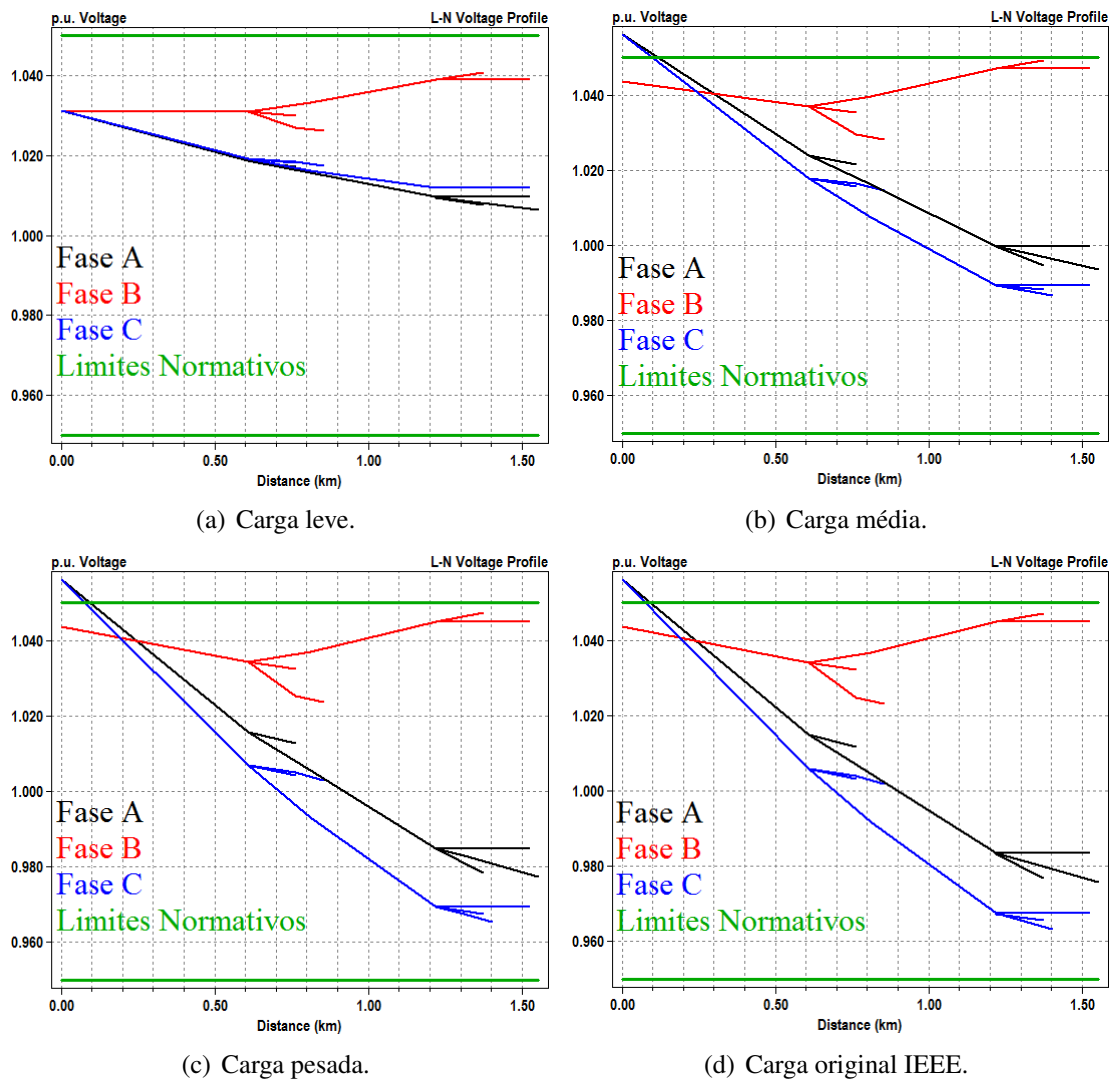
Percebe-se que em nenhum cenário houve violação dos níveis de tensão nas barras com clientes. Também é ilustrado o perfil da tensão no sistema para o cenário de cargas original na rede de 13 barras do IEEE. Observa-se pela Figura 49(c) e pela Figura 49(d) que o perfil de tensão do cenário de carga pesada é praticamente igual ao original do IEEE. Isso comprova que o sistema foi ajustado de forma que a demanda típica dos horários de maior carregamento não superasse a demanda do cenário original da rede, garantindo que a rede não fosse submetida a condições extremas de carregamento para as quais não foi projetada.

A inclusão de incertezas na modelagem da potência dos componentes do sistema permite avaliar o comportamento probabilístico da tensão nas barras com carga e GDFV, exemplificado para as 21h na ilustração da Figura 50, comprovando que tal grandeza pode variar significativamente ao longo de uma hora. Como não há contribuição de GDFV às 21h nas configurações simuladas para o sistema, as respectivas curvas PDF representam as possibilidades para os níveis de penetração de GDFV já citados.

Na Figura 50 são indicados os limites normativos de 0,93pu e 1,05pu para a tensão (ANEEL, 2018). É possível observar que os mesmos possuem maior chance de serem violados nas barras 611, 671, 692 e 675, justamente os pontos de consumo mais distantes da fonte principal de fornecimento de energia na topologia da rede. O gráfico da PDF das tensões no horário das 21h é conveniente para observar as violações de tensão, já que trata-se do horário de pico para a demanda no sistema e, portanto, o de maior exigência da rede. Em algumas curvas é possível observar notoriamente a influência da mudança no *tap* para ajuste da tensão de fornecimento na subestação. Também é possível observar que, a partir de um determinado limite de tensão, a curva PDF sofre uma deformação abrupta, indicando o limite da capacidade de compensação do *tap* em regular a tensão.

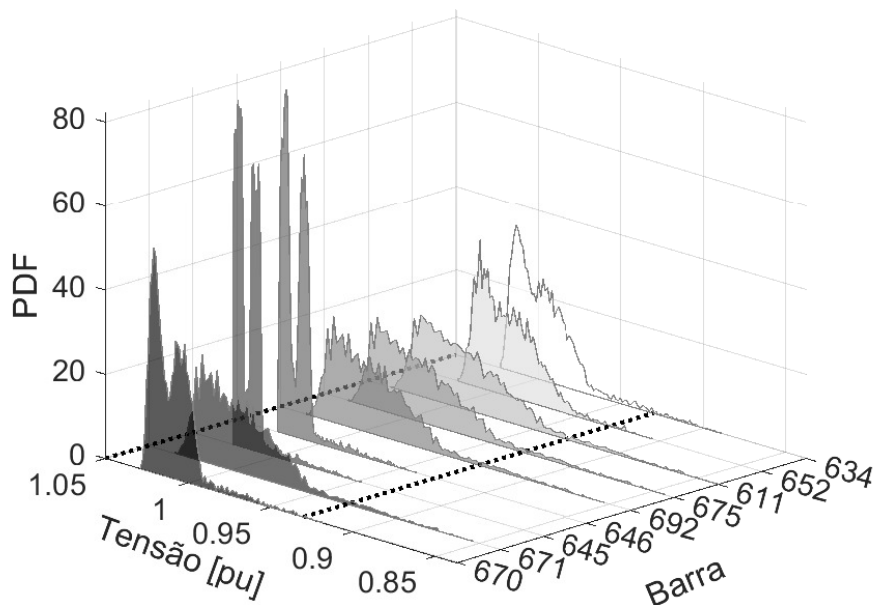
Para verificar a contribuição da GDFV é necessário observar horários alternativos. A Figura 51 ilustra as tensões nas barras para configurações com penetração de GDFV

Figura 49: Perfil de tensão na rede em p.u. para cenários determinísticos de carga na ausência de GDFV.



Fonte: O autor.

Figura 50: PDF da tensão em todas as barras às 21h.



Fonte: O autor.

distintas no horário das 14h e 18h. O pico de geração fotovoltaica ocorre às 14h, sendo esse o horário de maior contribuição da GD. Contudo no horário das 18h, há grande dispersão na demanda dos clientes e ainda existe um nível elevado de geração, tornando mais perceptível nesse horário o impacto da GDFV nas curvas PDF da tensão nas barras.

Dentre os resultados listados na Tabela 10, constam as violações de tensão crítica e precária em cada barra com carga do sistema, para os casos determinístico e probabilístico. Neste caso, os valores médios \bar{x}_{DRC} e \bar{x}_{DRP} são calculados a partir de $N = 5000$ amostras para os índices probabilísticos de violação de tensão crítica (DRC^{prob}) e precária (DRP^{prob}). Também são mostrados os valores da função sobrevivência⁴ $S(DRC^{prob} = 3\%)$ e $S(DRP^{prob} = 0.5\%)$ dos indicadores DRC^{prob} e DRP^{prob} em cada barra, ou seja, a probabilidade do indicador extrapolar o limite normativo.

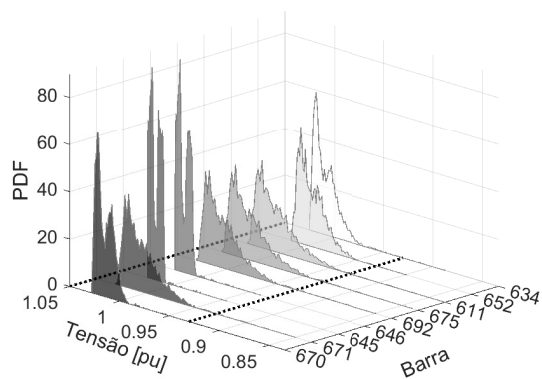
Tabela 10: Índices de violação de tensão para cenários de carga determinísticos e probabilísticos na ausência de GDFV.

Barra	Determinístico		Probabilístico			
	$DRC[\%]$	$DRP[\%]$	$\bar{x}_{DRC} [\%]$	$\bar{x}_{DRP} [\%]$	$S(DRC^{prob} = 3\%)$	$S(DRP^{prob} = 0.5\%)$
670	0	0	0,0525	0,0017	0,0124	0,0004
671	0	0	0,5642	0,2667	0,1282	0,0618
645	0	0	0,0008	0	0,0002	0
646	0	0	0,0025	0	0,0006	0
692	0	0	0,5658	0,2692	0,1284	0,0624
675	0	0	0,6775	0,3708	0,1524	0,0854
611	0	0	0,6683	0,4358	0,1504	0,0992
652	0	0	0,1742	0,0183	0,0412	0,0044
634	0	0	0,2692	0,0442	0,0630	0,0104

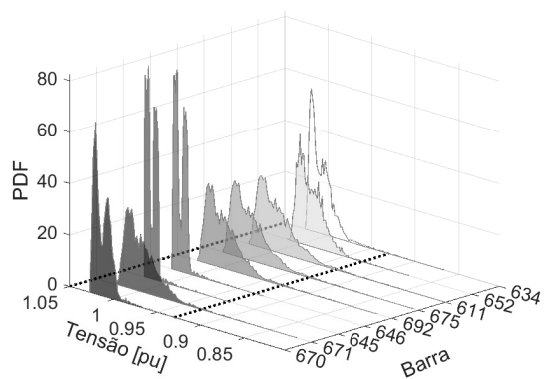
Na abordagem determinística, conclui-se que para um cenário de cargas típicas ao

⁴ $S(DRC^{prob} = 3\%) = P(DRC^{prob} > 3\%)$, conforme Equação (2).

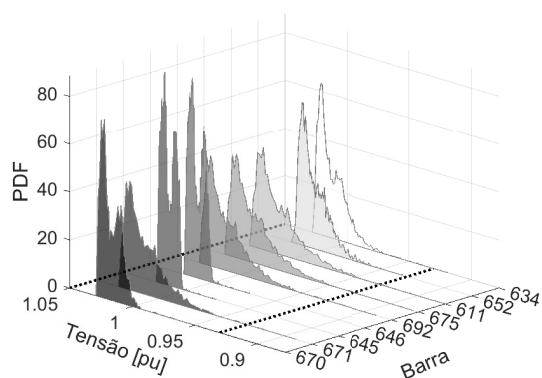
Figura 51: PDF da tensão em todas as barras às 14h e às 18h para configurações de penetração de GDFV diversas.



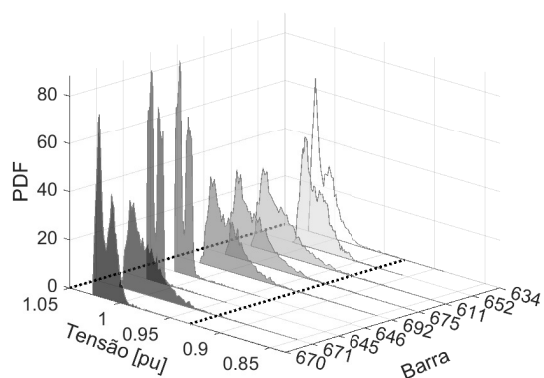
(a) Configuração com ausência de GDFV às 14h.



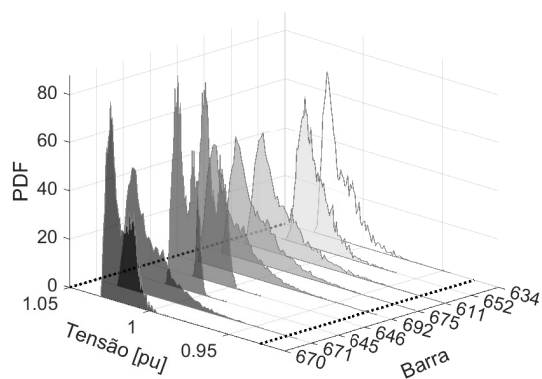
(b) Configuração com ausência de GDFV às 18h.



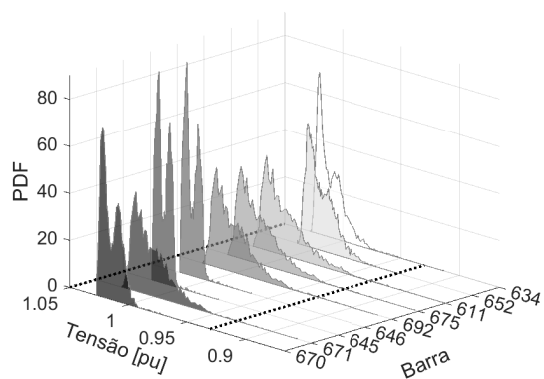
(c) Configuração com 50% de GDFV às 14h.



(d) Configuração com 50% de GDFV às 18h.



(e) Configuração com 100% de GDFV às 14h.



(f) Configuração com 100% de GDFV às 18h.

Fonte: O autor.

longo do dia, não há violação dos índices de tensão no sistema. Contudo, não é possível saber o quão próximo se está do limite normativo. Com o método probabilístico de análise, é possível verificar que, em média, todas as barras estão com os índices adequados.

Além disso, observou-se que o \bar{x}_{DRC} é maior na Barra 675 e que o \bar{x}_{DRP} é maior na Barra 611. Tal fenômeno é evidenciado pelo uso da CDF, obtida por adequação de ajuste para os índices de violação probabilísticos DRC^{prob} e DRP^{prob} nestas duas barras. Verificou-se que a probabilidade do DRC estar acima de 3% e do DRP estar acima de 0,5% na Barra 675 é de 0,1524 e 0,0854, respectivamente. Para a Barra 611 as mesmas probabilidades são de 0,1504 e 0,0992. Desta forma, torna-se possível avaliar a chance de os índices ultrapassarem os limites normativos, complementando o método analítico tradicional (ANEEL, 2018).

A Tabela 11 apresenta os resultados de violações de tensão obtidos pela análise determinística e probabilística para o sistema com uma configuração de penetração de 50% de GDFV. A Tabela 12 apresenta os mesmos resultados para o caso com 100% de penetração de GDFV. Percebe-se que há melhora nos índices à medida que aumenta a penetração de GDFV. Também nesses casos observa-se a incapacidade da análise determinística em obter bons estimadores dos indicadores de violação de tensão, sendo que somente na configuração com 100% de penetração de GDFV o resultado do indicador para as Barras 675 e 611 (barras com os piores indicadores) é não nulo. Somente com a abordagem probabilística é possível compreender o quanto a inserção da GDFV contribui para a melhora nos índices de operação da rede.

Tabela 11: Índices de violação de tensão para cenários de carga determinísticos e probabilísticos com 50% de penetração de GDFV.

Barra	Determinístico		Probabilístico			
	DRC [%]	DRP [%]	\bar{x}_{DRC} [%]	\bar{x}_{DRP} [%]	$S(DRC^{prob} = 3\%)$	$S(DRP^{prob} = 0.5\%)$
670	0	0	0,0225	0	0,0054	0
671	0	0	0,4050	0,1692	0,0942	0,0396
645	0	0	0	0	0	0
646	0	0	0	0	0	0
692	0	0	0,4075	0,1708	0,0948	0,0400
675	0	0	0,4683	0,2450	0,1062	0,0570
611	0	0	0,4967	0,2933	0,1124	0,0680
652	0	0	0,0983	0,0017	0,0234	0,0004
634	0	0	0,2008	0,0192	0,0466	0,0046

Tabela 12: Índices de violação de tensão para cenários de carga determinísticos e probabilísticos com 100% de penetração de GDFV.

Barra	Determinístico		Probabilístico			
	DRC [%]	DRP [%]	\bar{x}_{DRC} [%]	\bar{x}_{DRP} [%]	$S(DRC^{prob} = 3\%)$	$S(DRP^{prob} = 0.5\%)$
670	0	0	0,0192	0	0,0046	0
671	0	0	0,3417	0,1300	0,0800	0,0306
645	0	0	0	0	0	0
646	0	0	0	0	0	0
692	0	0	0,3433	0,1317	0,0804	0,0310
675	0	4,17	0,3925	0,2000	0,0902	0,0468
611	0	4,17	0,4208	0,2358	0,0958	0,0548
652	0	0	0,0775	0,0017	0,0184	0,0004
634	0	0	0,1633	0,0167	0,0382	0,0040

É necessário ressaltar que as curvas PDF obtidas para o DRC^{prob} e o DRP^{prob} descrevem o comportamento dos índices de violação de tensão, sendo o DRC e o DRP medidos conforme estabelece o PRODIST uma simples amostra dentre o universo de possibilidades para os mesmos (ANEEL, 2018).

Com a aplicação do método proposto, pode-se ter o entendimento de quão adequada esta a tensão nos pontos de conexão dos clientes. A partir do cálculo da sobrevivência para os indicadores e violação de tensão nas barras da rede, pode-se ter maior assertividade na definição da estratégia para minimização das transgressões dos limites de tensão de atendimento bem como dos custos decorrentes de compensação financeira, ou seja, disponibilizar apenas os recursos necessários para adequar os indicadores nos pontos com maior chance de infringir o limite. Isso também é benéfico para os consumidores de um modo geral, que passam a pagar uma Tarifa de Uso do Sistema de Distribuição (TUSD) menor nas suas faturas de energia, visto que os custos de operação e expansão da rede tendem a ser minimizados ou postergados para momentos mais oportunos.

6 CONCLUSÃO

Os resultados mostrados neste trabalho comprovam que a consideração da incerteza é muito importante para caracterizar a operação das redes de distribuição de energia em média e baixa tensão, pois as incertezas associadas ao comportamento da carga e da geração distribuída fazem parte da realidade da rede elétrica. A aleatoriedade e as possibilidades de cenários de demanda aumentam com a inserção de fontes de geração distribuída, principalmente nas redes secundárias onde há contribuição de menor quantidade de componentes e a lei dos grandes números tem aplicação limitada.

Diante disso, neste trabalho foi proposto um método para obtenção de indicadores probabilísticos de operação de redes elétricas. Demonstrou-se que a avaliação probabilística da operação torna-se vantajosa ao fornecer mais informações do que a análise determinística tradicional. A inclusão de incertezas na modelagem das cargas e fontes de GD, a geração de amostras correlacionadas e a avaliação estatística de indicadores de perdas ativas e tensões, por meio de critérios de convergência do método estocástico, aprofundaram os resultados médios em estimativas intervalares de confiança.

A modelagem estocástica para o comportamento horário individual dos componentes do sistema (carga e GD), através do uso de leis de distribuição de probabilidade padronizadas, conforme proposto por Pareja (2009), foi facilmente implementada, sendo ainda bastante flexível, podendo-se alterar a mesma, caso seja de interesse do analista.

A abordagem empregada neste trabalho baseia-se em trabalhos anteriores, que comprovaram a validade das leis padronizadas que representam de forma adequada as fontes aleatórias subjacentes que regem o comportamento dos clientes de redes de distribuição bem como das fontes típicas da GD (PAREJA, 2009; ABDULKARIM; ABDELKADER; MORROW, 2015). Entretanto, caso não seja possível representar o comportamento de um determinado componente através de uma lei de distribuição padronizada, pode-se obter a modelagem do comportamento através da adequação de ajuste ao histograma de frequência acumulada para a potência do componente (GUIMARÃES, 2008). Essa abordagem é menos flexível e necessita de maior quantidade de dados de medição dos componentes, porém oferece uma descrição mais personalizada do comportamento.

Com o uso de leis padronizadas de distribuição de probabilidade, para descrever as variáveis estocásticas, eventualmente é necessário inserir uma limitação artificial, pois o limite superior pode ser infinito. Neste trabalho, isso foi realizado a partir da consideração do fator de carga típico dos consumidores. Por outro lado, para a modelagem da GDFV não houve essa necessidade visto que a distribuição beta utilizada já é naturalmente limitada.

A demanda agregada no alimentador é considerada uma variável de saída, resultante da demanda dos componentes conectados na rede simulada. Contudo, o alimentador poderia ser interpretado como um componente e a PDF da demanda agregada utilizada para

simular a aleatoriedade de uma rede maior à qual outros subsistemas também estivessem conectados. Assim, o método proposto é versátil e pode ser aplicado em redes de diferentes níveis de tensão, especialmente nos níveis de tensão mais baixos, onde a variância do comportamento dos componentes tende a se elevar e sua aplicação oferece uma abordagem mais realista.

O fator de diversidade das cargas, amplamente utilizado em estudos de sistemas de distribuição, é um indicativo de que há relação de dependência entre os componentes. Dessa forma, por menor que possa ser, é necessário incluir a correlação na amostragem da potência dos componentes quando são criados os cenários. Não considerar esse aspecto seria realizar uma análise otimista da operação do sistema, visto que a demanda agregada alcança valores maiores à medida que a correlação aumenta, podendo levar o sistema ao limite de sua capacidade operativa mais facilmente. Com o uso da *Joint Normal Transform* – JNT, foi possível criar cenários de demanda considerando a correlação entre os componentes do sistema, obtendo indicadores conservadores da operação da rede. A possibilidade de correlacionar as diversas fontes aleatórias torna-se um grande atrativo da abordagem proposta neste trabalho.

Foi considerada a correlação entre os componentes no mesmo horário. No entanto, sabe-se que podem existir condições externas que impõem um viés de variação conjunta da potência dos componentes que se perpetua ao longo dos horários. Dias ensolarados, por exemplo, tendem a causar um aumento na demanda em função do maior uso de equipamentos de conforto térmico ao mesmo tempo que contribui para o aumento da GDFV. Nesse trabalho, uma tentativa para considerar este efeito foi realizada por intermédio do uso de um fator de memória entre os horários, entretanto este aspecto deveria ser mais bem explorado em trabalhos futuros. Uma das vantagens vislumbradas na consideração do efeito memória é na avaliação de indicadores que possuem memória como, por exemplo, a avaliação da perda de vida útil em transformadores que ocorre em função da elevação de temperatura de trabalho como consequência do histórico de carregamentos (BUSATTO, 2015).

O uso de técnicas de agrupamento dos componentes e o uso da *Stochastic Bounds Methodology* – SBM são essências para limitar o tamanho do problema. Com isso, a ordem da matriz de correlação utilizada na JNT fica restrita ao número de agrupamentos do sistema, ou seja, ao número de fontes aleatórias consideradas.

No estudo de caso, somente foi contemplada a inserção de GDFV, contudo, a avaliação de incertezas pode incluir outras fontes energéticas intermitentes, como as provenientes de geração eólica, e, assim, avaliar com maior fidelidade o impacto da geração distribuída. A inclusão de outras fontes de GD é facilmente implementada adotando os mesmos procedimentos de modelagem utilizados para a fonte solar.

Neste trabalho foi demonstrada a aplicação do método para a análise das perdas do sistema e da tensão de regime permanente na operação de um sistema de distribuição. Foi possível observar que o aumento na penetração de GDFV contribuiu para a melhora nos indicadores de operação da rede. Naturalmente, isso ocorre pelo efeito indireto de desoneração do sistema elétrico através da injeção de potência nos pontos de consumo. Contudo, é preciso atentar ao fato que a distribuição da GDFV considerada nas simulações foi proporcional à carga original no nó da rede e que os indicadores poderiam alcançar valores diferentes caso a concentração estivesse nas extremidades da rede ou concentrada em poucos locais.

Um dos objetivos do planejamento da expansão de sistemas de distribuição é minimizar os custos associados com a operação, ou seja, perdas de energia, violações de tensão

e manutenção (RESENER, 2016). A análise probabilística com o método proposto é adequada para analisar esses aspectos e a discussão apresentada nos capítulos anteriores demonstra como é possível precisar os níveis de perdas no sistema e também identificar a ocorrência de violações de tensão. Os resultados obtidos para a perda técnica diária média no sistema com o método probabilístico foram, pelo menos, 18,4% superiores aos obtidos de forma determinística para o sistema simulado. Desse modo, a estimativa das perdas não-técnicas é distorcida, uma vez que esta corresponde à diferença entre as perdas totais e as perdas técnicas. Visto que as perdas não-técnicas estão diretamente associadas à gestão comercial da distribuidora, esse desvio na estimativa pode impactar o planejamento da mesma. Além disso, é possível estimar valores máximos e mínimos do indicador de operação para um nível de confiança estatística e o uso do método probabilístico permite saber as chances de ultrapassar os índices de violação de tensão normatizados, por exemplo.

Os valores obtidos para os indicadores de operação pelo método probabilístico proposto são resultantes da consideração de correlações oriundas de um estudo distinto e não necessariamente são os mais adequados para a rede analisada nas configurações impostas. Contudo, o aspecto principal a ser considerado não são os valores absolutos, mas sim a perspectiva de análise que defende a incapacidade de representar o comportamento dos clientes de redes elétricas de distribuição de forma plena apenas com a média.

Apesar dos resultados de violação de tensão e perdas obtidos pelo método probabilístico indicarem que o aumento da GD pode ser benéfico para os indicadores de operação do sistema, é necessário verificar outras condições, como a capacidade das linhas e os carregamentos nos transformadores. Assim, o método proposto pode ser aplicado na avaliação do carregamento de condutores bem como em transformadores ou outros equipamentos e com isso, avaliar a perda de vida útil decorrente de sobreaquecimento. Desse modo é possível atuar na manutenção dos equipamentos no momento estatisticamente mais adequado, reduzindo custos e sendo possível até mesmo prever o momento propício para a troca dos mesmos por perda de vida útil ou por subdimensionamento. A combinação do método proposto com a ferramenta de priorização de investimentos desenvolvida por Busatto (2015) tornaria esta mais assertiva por incluir maior quantidade de cenários avaliados para apoiar a decisão de investimento por exemplo.

Por fim, com a aplicação do método proposto, é possível obter uma lei de distribuição para as perdas diárias bem como dos índices de violação de tensão. Os custos associados à operação da rede, seja em função das perdas ou pelos custos de ressarcimento aos clientes por índices inadequados de tensão de fornecimento, podem ser avaliados desse modo com intervalos de confiança. Assim, é possível definir estratégias mais próximas à realidade da rede para a quantificação dos custos operacionais diários da rede de distribuição de energia elétrica, sendo assim possível determinar os investimentos com a melhor relação entre custo e benefício.

6.1 Trabalhos Futuros

Embora o método proposto se apresente como uma ferramenta útil na avaliação da operação de redes elétricas com uma abordagem probabilística, existem aspectos que merecem aperfeiçoamento, sendo sugestões de estudos os seguintes tópicos:

- inclusão de incerteza proveniente de outras fontes de GD e análise do impacto que as mesmas exercem nos indicadores de operação;

- expandir o método para outros indicadores de operação, como perda de vida útil de equipamentos e fator de utilização do sistema;
- aprofundar a investigação do efeito da memória, ou seja, da correlação entre as demandas e GD horárias;
- incluir outras fontes de incerteza, como localização da GD;
- avaliar a incerteza da injeção de potência da GD em caso de uso de sistemas de armazenamento;
- avaliar o nível de penetração da GD que poderia, em combinação com sistemas de armazenamento, funcionar como uma usina virtual e, por meio de despacho local, suprimir a necessidade de componentes de regulação de tensão como banco de capacitores;
- combinar o método proposto com modelos heurísticos para planejamento da expansão de sistemas de distribuição com o objetivo de melhorar os indicadores de qualidade e minimizar os custos de operação e expansão.

REFERÊNCIAS

ABDULKARIM, A.; ABDELKADER, S. M.; MORROW, D. J. Statistical analyses of wind and solar energy resources for the development of hybrid microgrid. In: INTERNATIONAL CONGRESS ON ENERGY EFFICIENCY AND ENERGY RELATED MATERIALS (ENEFM2014), 2., 2015, Oludeniz. **Proceedings...** Cham: Springer International Publishing, 2015. p.9–14.

AGÊNCIA NACIONAL DE ENERGIA ELÉTRICA. **Resolução normativa nº 414, de 09 de setembro de 2010**. Brasília, 2010.

AGÊNCIA NACIONAL DE ENERGIA ELÉTRICA. **Resolução normativa nº 482, de 17 de abril de 2012**. Brasília, 2012.

AGÊNCIA NACIONAL DE ENERGIA ELÉTRICA. **Procedimentos de distribuição de energia elétrica no sistema elétrico nacional - PRODIST: módulo 7 - cálculo de perdas na distribuição**. Brasília, 2015.

AGÊNCIA NACIONAL DE ENERGIA ELÉTRICA. **Procedimentos de distribuição de energia elétrica no sistema elétrico nacional - PRODIST: módulo 2 - planejamento da expansão do sistema de distribuição**. Brasília, 2016.

AGÊNCIA NACIONAL DE ENERGIA ELÉTRICA. **Procedimentos de distribuição de energia elétrica no sistema elétrico nacional - PRODIST: módulo 8 - qualidade da energia elétrica**. Brasília, 2018.

AGÊNCIA NACIONAL DE ENERGIA ELÉTRICA. **Outorgas e registros de geração: micro e minigeração distribuída (REN 482/2012)**, jan. 2017. Disponível em: <<http://www.aneel.gov.br/outorgas/geracao>>. Acesso em: 3 fev. 2018.

AGÊNCIA NACIONAL DE ENERGIA ELÉTRICA. **Dados de geração: empreendimentos em operação**, [dez. 2017]. Disponível em: <<http://www.aneel.gov.br/dados/geracao>>. Acesso em: 3 fev. 2018.

ALGARNI, A. A. S. **Operational and planning aspects of distributions systems in deregulated electricity markets**. 2009. 162 f. Tese (Doutorado) — University of Waterloo, Waterloo, Canada, 2009.

ANZANELLO JÚNIOR, J. D. **Análise de microgeração fotovoltaica em um sistema de distribuição utilizando o OpenDSS**. 2016. 32 f. Monografia (Bacharelado em Engenharia Elétrica) — Universidade Federal do Rio Grande do Sul, Porto Alegre, RS, 2016.

ASSOCIAÇÃO BRASILEIRA DE GERAÇÃO DISTRIBUÍDA. **Mapa do mercado de geração distribuída no Brasil**, abr. 2017. Disponível em: <<https://www.geracaodistribuida.org/mapa-do-mercado>>. Acesso em: 3 fev. 2018.

ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. **NBR 5416**: Aplicação de cargas em transformadores de potência – procedimento. Rio de Janeiro, 1997.

BECKER, J. L. **Estatística básica**: transformando dados em informação. Porto Alegre: Bookman, 2015.

BORKOWSKA, B. Probabilistic load flow. **IEEE Transactions on Power Apparatus and Systems**, New York, v.PAS-93, n.3, p.752–759, maio 1974.

BUSATTO, T. **Metodologia multicritério para priorização de investimentos no planejamento da expansão de sistemas de distribuição de energia elétrica**. 2015. 96 f. Dissertação (Mestrado em Engenharia Elétrica) — Programa de Pós-graduação em Engenharia Elétrica, Universidade Federal do Rio Grande do Sul. Porto Alegre, RS, 2015.

CAMPOS, M. A.; RÊGO, L. C.; MENDONÇA, A. F. de. **Métodos probabilísticos e estatísticos com aplicações em engenharias e ciências exatas**. 1.ed. Rio de Janeiro: LTC, 2017.

CENTRAIS ELÉTRICAS BRASILEIRAS – ELETROBRAS. **Pesquisas de posse e hábitos de consumo de energia (PPHs)**, 2006. Disponível em: <<http://www.procelinfo.com.br/main.asp?View={4A5E324F-A3B0-482A-B1CD-F75A2A150480}>>. Acesso em: 3 set. 2017.

CENTRO DE PESQUISAS DE ENERGIA ELÉTRICA. **Centro de referência para as energias solar e eólica Sérgio de S. Brito**, [2017?]. Disponível em: <www.cresesb.cepel.br/index.php?lang=pt>. Acesso em: 02 dez. 2017.

CHEN, P.; CHEN, Z.; BAK-JENSEN, B. Probabilistic Load Flow: a review. In: INTERNATIONAL CONFERENCE ON ELECTRIC UTILITY DEREGULATION AND RESTRUCTURING AND POWER TECHNOLOGY, 3., 2008, Nanjing. China, **Proceedings ...** New York: IEEE, 2008.

CHEN, Y. et al. Bringing probabilistic analysis capability from planning to operation. **Control Engineering Practice**, Oxford, v.71, p.18–25, fev. 2018. Disponível em: <<https://doi.org/10.1016/j.conengprac.2017.06.006>>. Acesso em: 7 out. 2017.

CHICCO, G.; NAPOLI, R.; PIGLIONE, F. Comparisons among clustering techniques for electricity customer classification. **IEEE Transactions on Power Systems**, New York, v.21, n.2, p.933–940, maio 2006.

CONTI, S.; RAITI, S. Probabilistic load flow using Monte Carlo techniques for distribution networks with photovoltaic generators. **Solar Energy**, Oxford, v.81, n.12, p.1473–1481, dez. 2007. Disponível em: <<https://doi.org/10.1016/j.solener.2007.02.007>>. Acesso em: 7 out. 2017.

CUI, T.; FRANCHETTI, F. A Quasi-Monte Carlo approach for radial distribution system probabilistic load flow. In: IEEE PES INNOVATIVE SMART GRID TECHNOLOGIES CONFERENCE (ISGT), 2013, Washington. **Proceedings...** New York: IEEE, 2013.

EKSTRÖM, J. *et al.* A statistical approach for hourly photovoltaic power generation modeling with generation locations without measured data. **Solar Energy**, Oxford, v.132, p.173-187, jul. 2016. Disponível em: <<https://doi.org/10.1016/j.solener.2016.02.055>>. Acesso em: 7 out. 2017.

ELECTRIC POWER RESEARCH INSTITUTE. **Open distribution system simulator (OpenDSS)**, jun. 2013. Disponível em: <<http://smartgrid.epri.com/SimulationTool.aspx>>. Acesso em: 3 set. 2016.

FARIA, J. M. B. A. de. **VaR para riscos agregados não necessariamente independentes com cópulas**. 2014. 174 f. Dissertação (Mestrado em Estatística) — Universidade de Brasília, Brasília, 2014.

FRANCISQUINI, A. A. **Estimação de curvas de carga em pontos de consumo e em transformadores de distribuição**. 2006. 108 f. Dissertação (Mestrado em Engenharia Elétrica) — Faculdade de Engenharia, Universidade Estadual Paulista, Ilha Solteira, SP, 2006.

FREITAS, P. R. R. de. **Modelos avançados de análise de redes elétricas inteligentes utilizando o software OpenDSS**. 2015. 191 f. Monografia (Bacharelado em Engenharia) — Universidade de São Paulo, São Paulo, 2015.

GALLEGO, L. A.; ECHEVERRI, M. G.; FELTRIN, A. P. Fluxo de potência trifásico probabilístico para redes de distribuição usando o método de estimação por pontos. **Revista Controle e Automação**, Campinas, SP, v.23, n.2, p.179-189, mar./abr. 2012.

GUIMARÃES, R. O. **Avaliação dos riscos de violações de conformidade de tensão em sistemas de distribuição, utilizando método probabilístico e conjuntos fuzzy**. 2008. 129 f. Tese (Doutorado em Engenharia) — Escola Politécnica da Universidade de São Paulo, Universidade de São Paulo, São Paulo, SP, 2008.

GUPTA, A. R.; KUMAR, A. Comparison of deterministic and probabilistic radial distribution systems load flow. **Journal of The Institution of Engineers (India): Series B**, [S.l.], v.98, n.6, p.547-556, dez. 2017. Disponível em: <<https://doi.org/10.1007/s40031-017-0288-2>>. Acesso em: 7 out. 2017.

HAFFNER, S. *et al.* Análise e modelagem de cargas típicas de redes de distribuição de energia elétrica. In: LATIN-AMERICAN CONGRESS ON ELECTRICITY GENERATION AND TRANSMISSION, 8., 2009, Ubatuba, SP. **Anais...** Guaratinguetá: FEG/UNESP, 2009.

HONG, T.; FAN, S. Probabilistic electric load forecasting: a tutorial review. **International Journal of Forecasting**, Amsterdam, v.32, n.3, p.914-938, set. 2016. Disponível em: <<https://doi.org/10.1016/j.ijforecast.2015.11.011>>. Acesso em: 7 out. 2017.

INSTITUTE OF ELECTRICAL AND ELECTRONICS ENGINEERS. **Power Engineering Society (PES) distribution test feeders**. Disponível em: <ewh.ieee.org/soc/pes/dsacom/testfeeders/index.html>. Acesso em: 02 set. 2016.

JARDINI, J. A. *et al.* Daily load profiles for residential, commercial and industrial low voltage consumers. **IEEE Transactions on Power Delivery**, New York, v.15, n.1, p.375-380, jan. 2000.

JOHNSON, R. A.; WICHERN, D. W. **Applied multivariate statistical analysis**. 3rd.ed. Upper Saddle River: Prentice-Hall International, 1992.

JONES, G. W.; CHOWDHURY, B. H. Distribution system operation and planning in the presence of distributed generation technology. In: IEEE/PES TRANSMISSION AND DISTRIBUTION CONFERENCE AND EXPOSITION, 2008, Chigado. **Proceedings...** New York: IEEE, 2008.

KABIR, M.; MISHRA, Y.; BANSAL, R. Probabilistic load flow for distribution systems with uncertain PV generation. **Applied Energy**, Oxford, v.163, p.343–351, fev. 2016. Disponível em: <<https://doi.org/10.1016/j.apenergy.2015.11.003>>. Acesso em: 7 out. 2017.

KAGAN, N.; OLIVEIRA, C. C. B. de; BORBA, E. J. **Introdução aos sistemas de distribuição de energia elétrica**. 2.ed. São Paulo: Edgard Blucher, 2005.

KERSTING, W. H. **Distribution system modeling and analysis**. 2.ed. Boca Raton, Florida: CRC Press LLC, 2007.

LABORATÓRIO DE ENERGIA SOLAR (LABSOL) - UFRGS. **Programa Radiasol2**, [2017?]. Disponível em: <<http://www.solar.ufrgs.br/#radiasol>>. Acesso em: 13 jan. 2018.

LAUTENSCHLEGER, A. H. **Projeção de demanda de energia elétrica da classe residencial considerando a inserção de mini e microgeração fotovoltaica**. 2013. 95 f. Monografia (Bacharelado em Engenharia) — Universidade Federal do Rio Grande do Sul, Porto Alegre, RS, 2013.

LAW, A. M.; KELTON, W. D. **Simulation modeling and analysis**. 3.ed. New York: McGraw-Hill, 2000.

LIU, E.; BEBIC, J. **Distribution system voltage performance analysis for high-penetration photovoltaics**. New York: NREL, 2008.

MATHWORKS. **Copulas: generate correlated samples**, [2017?]. Disponível em: <www.mathworks.com/help/stats/copulas-generate-correlated-samples.html#buqq6p7-1>. Acesso em: 10 nov. 2017.

MILANOVIĆ, J. V. *et al.* **Modelling and aggregation of loads in flexible power networks**. Paris: CIGRÉ, 2014.

MILLER, N.; YE, Z. **Report on distributed generation penetration study**. Cole Boulevard Golden, CO: National Renewable Energy Laboratory (NREL), 2003. Disponível em: <<https://www.nrel.gov/docs/fy03osti/34715.pdf>>. Acesso em: 7 mar. 2017.

MOHAMMADI, M.; SHAYEGANI, A.; ADAMINEJAD, H. A new approach of point estimate method for probabilistic load flow. **International Journal of Electrical Power & Energy Systems**, Oxford, v.51, p.54–60, out. 2013. Disponível em: <<https://doi.org/10.1016/j.ijepes.2013.02.019>>. Acesso em: 7 out. 2017.

MONTGOMERY, D. C.; RUNGER, G. C. **Applied statistics and probability for engineers**. 4.ed. Hoboken: John Wiley & Sons, 2007.

OBRENOVIĆ, N.; VIDAKOVIĆ, G.; LUKOVIĆ, I. The choice of metric for clustering of electrical power distribution consumers. In: DATA SCIENCE - ANALYTICS AND APPLICATIONS, 2017, Wiesbaden. **Proceedings...** Wiesbaden: Springer Vieweg, 2017. p.71-76.

OPERADOR NACIONAL DO SISTEMA. **Procedimentos de Rede**: submódulo 5.6 - consolidação da previsão de carga para a elaboração do programa mensal da operação energética. Brasília, 2017.

OPERADOR NACIONAL DO SISTEMA. **Histórico da operação**, [2017]. Disponível em: <<http://ons.org.br/pt/paginas/resultados-da-operacao/historico-da-operacao>>. Acesso em: 3 set. 2017.

PALUDO, J. A. **Avaliação dos impactos de elevados níveis de penetração de geração fotovoltaica no desempenho de sistemas de distribuição de energia elétrica em regime permanente**. 2014. 186 f. Dissertação (Mestrado em Ciências) — Escola de Engenharia de São Carlos, Universidade de São Paulo, São Carlos, SP, 2014.

PAPAEFTHYMIU, G. **Integration of stochastic generation in power systems**. 2007. 178 f. Tese (Doutorado) — Delft University of Technology, Delft, The Netherlands, 2007.

PAPAEFTHYMIU, G.; KUROWICKA, D. Using copulas for modeling stochastic dependence in power system uncertainty analysis. **IEEE Transactions on Power Systems**, New York, v.24, n.1, p.40-49, fev. 2009.

PAREJA, L. A. G. **Fluxo de potência em redes de distribuição de energia elétrica considerando incertezas**. 2009. 156 f. Tese (Doutorado em Engenharia) — Programa de Pós-Graduação em Engenharia Elétrica, Faculdade de Engenharia de Ilha Solteira, Universidade Estadual Paulista, Ilha Solteira, SP, 2009.

PFENNINGER, S.; STAFFELL, I. **Renewables Ninja**: dados de geração solar fotovoltaica mundial, 2016. Disponível em: <<https://www.renewables.ninja/>>. Acesso em: 02 dez. 2017.

PRUSTY, B. R.; JENA, D. A critical review on probabilistic lod flow studies in uncertainty constrained power systems with photovoltaic generation and a new approach. **Renewable and Sustainable Energy Reviews**, Oxford, v.69, p.1286-1302, mar. 2017. Disponível em: <<https://doi.org/10.1016/j.rser.2016.12.044>>. Acesso em: 7 out. 2017.

RANGEL, C. A. S. **Modelo aproximado de sistemas de distribuição considerando geração distribuída e o efeito da tensão na carga**. 2015. 98 f. Dissertação (Mestrado em Engenharia) — Programa de Pós-graduação em Engenharia Elétrica, Universidade Federal do Rio Grande do Sul, Porto Alegre, RS, 2015.

RESENER, M. **Modelo linearizado para problemas de planejamento da expansão de sistemas de distribuição**. 2016. 199 f. Tese (Doutorado em Engenharia) — Programa de Pós-graduação em Engenharia Elétrica, Universidade Federal do Rio Grande do Sul, Porto Alegre, RS, 2016.

RODRIGUES, P. M. L. M. R. **Probabilistic load flow considering distributed renewable generation**. 2008. 85 f. Dissertação (Mestrado em Engenharia) — Universidade Técnica de Lisboa, Lisboa, 2008.

SAMET, H.; KHORSHIDSAVAR, M. Analytic time series load flow. **Renewable and Sustainable Energy Reviews**, Oxford, v.82, p.3886–3899, fev. 2018. Disponível em: <<https://doi.org/10.1016/j.rser.2017.10.084>>. Acesso em: 7 out. 2017.

SHAYANI, R. A. **Método para determinação do limite de penetração da geração distribuída fotovoltaica em redes radiais de distribuição**. 2010. 161 f. Tese (Doutorado em Engenharia) — Departamento de Engenharia Elétrica, Faculdade de Tecnologia, Universidade de Brasília, Brasília, DF, 2010.

SHORT, T. A. **Electric power distribution handbook**. 1.ed. Boca Raton, Florida: CRC Press LLC, 2003.

VILLELA, A. G.; CARVALHO, V. A. de. **Cálculo de demanda rede para acréscimo de novas cargas na rede de distribuição de baixa tensão**. São Paulo: AES Eletropaulo, 2007, 24 f. Disponível em: <<https://www.aeseletropaulo.com.br/padrees-e-normas-tecnicas/manuais-normas-tecnicas-e-de-seguranca/Documents/Padr%C3%B5es%20e%20Normas%20T%C3%A9cnicas/NT-2018.pdf>>. Acesso em: 24 nov. 2017.

XU, X.; YAN, Z. Probabilistic load flow calculation with quasi-Monte Carlo and multiple linear regression. **International Journal of Electrical Power & Energy Systems**, Oxford, v.88, p.1–12, jun. 2017. Disponível em: <<https://doi.org/10.1016/j.ijepes.2016.11.013>>. Acesso em: 7 out. 2017.