



Gabriel Calvo Martinez

**Geração de cenários sintéticos com
granularidade horária para produção energética
de usinas de biomassa**

Dissertação de Mestrado

Dissertação apresentada como requisito parcial para a obtenção do grau de Mestre pelo Programa de Pós-graduação em Engenharia de Produção da PUC-Rio .

Orientador : Prof. Reinaldo Castro Souza
Co-orientador: Prof. Paula Medina Maçaira Louro

Rio de Janeiro
setembro de 2021



Gabriel Calvo Martinez

**Geração de cenários sintéticos com
granularidade horária para produção energética
de usinas de biomassa**

Dissertação apresentada como requisito parcial para a obtenção do grau de Mestre pelo Programa de Pós-graduação em Engenharia de Produção da PUC-Rio . Aprovada pela Comissão Examinadora abaixo.

Prof. Reinaldo Castro Souza

Orientador

Departamento de Engenharia Industrial – PUC-Rio

Profa. Paula Medina Maçaira Louro

Co-orientadora

Departamento de Engenharia Industrial – PUC-Rio

Prof André Luís Marques Marcato

UFJF

Prof José Francisco Moreira Pessanha

UERJ

Rio de Janeiro, 14 de setembro de 2021

Todos os direitos reservados. É proibida a reprodução total ou parcial do trabalho sem autorização da universidade, do autor e do orientador.

Gabriel Calvo Martinez

O autor é graduado em engenharia elétrica na Pontifícia Universidade Católica, PUC-Rio, em Dezembro de 2018. Durante a graduação, o autor trabalhou como estagiário no laboratório Lambda do Departamento de Engenharia Elétrica da PUC-Rio, realizando simulação de operação de subestações de energia elétrica.

Ficha Catalográfica

Martinez, Gabriel Calvo
Geração de cenários sintéticos com granularidade horária para produção energética de usinas de biomassa / Gabriel Calvo Martinez; orientador: Reinaldo Castro Souza; co-orientadora: Paula Medina Maçaira Louro. – Rio de Janeiro: PUC-Rio, Departamento de Engenharia Industrial, 2021.

v., 60 f: il. color. ; 30 cm

Dissertação (mestrado) - Pontifícia Universidade Católica do Rio de Janeiro, Departamento de Engenharia Industrial.

Inclui bibliografia

1. Engenharia de Produção – Teses. 2. Granularidade Horária;. 3. Simulação;. 4. Séries Temporais.. I. Souza, Reinaldo Castro . II. Louro, Paula Medina Maçaira. III. Pontifícia Universidade Católica do Rio de Janeiro. Departamento de Engenharia Industrial. IV. Título.

CDD: 658.5

Dedico esta dissertação a minha mãe, Maria Josefa Martinez Fernandez, e a meu pai, José Carlos Barreiro Calvo, por todo o apoio fornecido durante o período em que esta dissertação foi elaborada.

Agradecimentos

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001.

Agradeço especialmente ao meu orientador, Prof. Reinaldo Castro Souza e a co-orientadora, Profa. Paula Medina Maçaira, pela oportunidade de realizar esta dissertação e o privilégio de receber a orientação de ambos.

Direciono meus agradecimentos especiais também a Mariana Alves Londes.

Resumo

Martinez, Gabriel Calvo ; Souza, Reinaldo Castro ; Louro, Paula Medina Maçaira. **Geração de cenários sintéticos com granularidade horária para produção energética de usinas de biomassa.** Rio de Janeiro, 2021. 60p. Dissertação de Mestrado – Departamento de Engenharia Industrial, Pontifícia Universidade Católica do Rio de Janeiro.

As fontes renováveis de energia tem adquirido grande relevância no contexto mundial devido as diretrizes tomadas pelas políticas globais. Embora sejam termoeletricas, as usinas a biomassa são consideradas fontes de energia renováveis por manter os níveis de gás carbônico na atmosfera. Além disso, as usinas termoeletricas têm mais autonomia na geração do que as demais fontes renováveis como hidrelétricas, eólicas e solares que dependem de incertezas como vazões de água, velocidade do vento e irradiação solar, respectivamente. Este conjunto de características torna a biomassa uma fonte de energia indispensável para futuros planejamentos de operação do sistema. No entanto, esta fonte apresenta períodos intermitentes por estar usualmente atrelada a cultura de cana-de-açúcar. Poucos são os estudos de previsão e simulação desta fonte de energia. Dessa forma, a dissertação visa a implementação do modelo de cadeias de Markov com simulações de Monte Carlo.

Palavras-chave

Granularidade Horária; Simulação; Séries Temporais.

Abstract

Martinez, Gabriel Calvo ; Souza, Reinaldo Castro (Advisor); Louro, Paula Medina Maçaira (Co-Advisor). **Synthetic scenarios generation with hourly granularity for energy production from biomass power plants**. Rio de Janeiro, 2021. 60p. Dissertação de Mestrado – Departamento de Engenharia Industrial, Pontifícia Universidade Católica do Rio de Janeiro.

Renewable energy sources have acquired great relevance in the world context due to the guidelines taken by global policies. Although they are thermoelectric, biomass power plants are considered renewable energy sources because they maintain the levels of carbon dioxide in the atmosphere. In addition, thermoelectric plants have more autonomy in generation than other renewable sources such as hydro, wind and solar, which depend on uncertainties such as water flow, wind speed and solar irradiation, respectively. This set of characteristics makes biomass an indispensable energy source for future system operation planning. However, this source has intermittent periods as it is usually linked to the sugarcane crop. There are few prediction and simulation studies of this energy source. Thus, the dissertation aims to implement the Markov chain model with Monte Carlo simulations.

Keywords

Hourly Granularity; Simulation; Time Series.

Sumário

1	Introdução	12
2	Revisão de Literatura	16
2.1	Resumo do Capítulo	19
3	Modelagem Matemática	21
3.1	Subconjunto do Histórico	21
3.2	Clusterização k-means	22
3.3	Matriz de Transição	23
3.4	Simulação de Monte Carlo	25
3.5	Resumo do Capítulo	26
4	Análise de Dados	28
4.1	Análise de Tendência e Sazonalidade	29
4.2	Distribuição Probabilística	32
4.3	Divisão em Etapas	35
4.4	Resumo do Capítulo	36
5	Resultados	37
5.1	Clusterização via k-means	37
5.2	Matrizes de Transição	39
5.3	Avaliação das simulações	41
5.4	Resumo do Capítulo	47
6	Conclusão	48
7	Referências bibliográficas	50

Lista de figuras

Figura 1.1	Matriz energética europeia, ano 2016. Adaptado da fonte:[1]	13
Figura 1.2	Matriz elétrica dos Estados Unidos da América, ano 2016. Adaptado da fonte:[2]	14
Figura 1.3	Matriz elétrica brasileira, ano 2018. Adaptado da fonte:[3]	15
Figura 3.1	Fluxograma da metodologia	21
Figura 4.1	Histórico de Geração de Energia Elétrica a Biomassa.	28
Figura 4.2	Boxplot mensal.	29
Figura 4.3	Boxplot Horário sem o Período Intermitentes.	30
Figura 4.4	Decomposição dos Fatores Sazonais da Geração.	31
Figura 4.5	Densidade de Probabilidade - Início de Geração	33
Figura 4.6	Densidade de Probabilidade - Geração Plena	33
Figura 4.7	Densidade de Probabilidade - Término de Geração	34
Figura 4.8	Subconjuntos do Histórico de Geração de Energia Elétrica.	35
Figura 5.1	WSS dos clusters estimados para a Etapa Inicial.	37
Figura 5.2	WSS dos clusters estimados para a Etapa de Geração Plena.	38
Figura 5.3	WSS dos clusters estimados para a Etapa Final.	38
Figura 5.4	Densidade de Clusters - Início de Geração.	44
Figura 5.5	Densidade de Clusters - Geração Plena.	44
Figura 5.6	Densidade de Clusters - Término de Geração.	45
Figura 5.7	Curvas Comparativas de geração de energia a Biomassa.	46

Lista de tabelas

Tabela 5.1	Matriz de Probabilidade de Transição - Início de Geração.	39
Tabela 5.2	Matriz de Probabilidade de Transição - Geração Plena.	39
Tabela 5.3	Matriz de Probabilidade de Transição - Término de Geração.	40
Tabela 5.4	Matriz de Probabilidade Acumulada - Início de Geração.	40
Tabela 5.5	Matriz de Probabilidade Acumulada - Geração Plena.	41
Tabela 5.6	Matriz de Probabilidade Acumulada - Término de Geração.	41
Tabela 5.7	Comparação das Estatísticas Básicas entre histórico e simulação.	42
Tabela 5.8	Teste Hipóteses sobre as diferenças entre Médias, teste-t, e Variâncias, teste-F, dos Histórico e Simulação.	43

*Só na grandeza da vossa misericórdia coloco
toda a minha esperança. Dai-me o que me
ordenais e ordenai-me o que quiserdes.*

Santo Agostinho, Confissões.

1

Introdução

As fontes renováveis de energia têm gerado grande repercussão no contexto global por serem base do conceito de desenvolvimento sustentável. Dentre essas fontes, estão inclusas nesta categoria a energia hidrelétrica, eólica e biomassa. Segundo [4], a biomassa é renovada através do ciclo do carbono, em que o carbono liberado pela queima do bagaço produzido é reabsorvido pela cultura da própria cana-de-açúcar e regiões florestais, não aumentando as concentrações de carbono na atmosfera como o caso dos combustíveis fósseis.

A biomassa se apresenta como uma alternativa ao uso de combustíveis fósseis pelas características autônomas de operação de ambas usinas. A partir desta perspectiva, trabalhos como [5] e [6] buscaram estabelecer fatores econômicos afetados pelas flutuações de mercado que, por sua vez, impactam no preço e produção de biomassa destinada a geração de energia. [7] aponta que a substituição de combustíveis fósseis por biomassa se mostrou viável economicamente.

Ressalta-se que, pelo fato da biomassa ser frequentemente oriunda de atividades agrícolas, o produtor pode optar por produzir a matriz energética para atender o mercado de alimentos. Nesse contexto, [8] realizou uma análise de sensibilidade às variações de mercado no Kansas para que os produtores locais que originalmente produzem alimentos, passassem a produzir, como alternativa, os biocombustíveis.

Analisando o contexto de matriz energética pelo mundo, não se encontra um cenário favorável para as energias renováveis. Na Europa, o perfil da matriz energética da União Europeia em 2016, ano de referência representada pela Figura 1.1 segundo a Comissão Europeia [1], aponta uma grande ênfase em fontes não-renováveis de energia, que representavam naquele ano 83% da matriz energética europeia. Embora a biomassa continue sendo a principal fonte de energia renovável, apenas 1,34% de toda a matriz energética a biomassa é destinada a energia elétrica propriamente dita.

Recentemente, a EUROSTAT [9] informou em seu mais recente balanço, que em 2018 houve um aumento na participação de renováveis, que chegaram a representar 18,7% da composição da matriz energética europeia, com expectativa de 20% para o ano de 2020, dado ainda a ser levantado, e 32% até 2030. A Agenda 2030 [10] estabeleceu metas específicas neste âmbito, em que se propõe estabelecer tarifas de energia mais acessíveis e a utilização de fontes renováveis de energia. Dentre essas metas, está o aumento da participação da

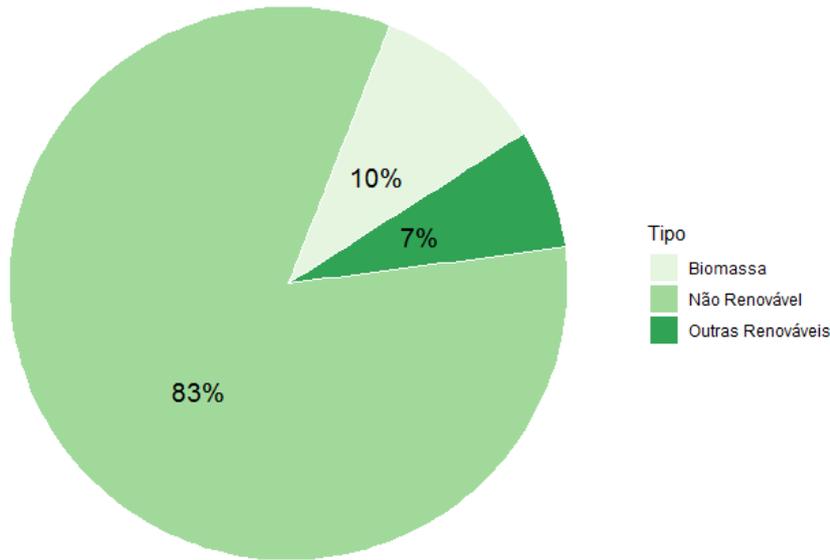


Figura 1.1: Matriz energética europeia, ano 2016. Adaptado da fonte:[1]

matriz energética renovável para 27%. No entanto, as energias renováveis representaram 38% da matriz elétrica do ano de 2020 sendo apenas 2.2% oriunda da biomassa.

Nos Estados Unidos foi estabelecido em 2007 o plano de ação 25X'25 [11] que visa diversificar a matriz energética de modo que 25% de toda energia consumida nos EUA seja oriunda de fazendas, florestas e sítios enquanto produzem comida de forma abundante até 2025. Tal proposta se mostra um desafio uma vez que o produtor da agricultura deve decidir entre produzir alimentos e gerar energia elétrica a partir da biomassa oriunda dos tipos de plantação. Ainda em 2007, o Comitê de Clima do Senado dos EUA [12] anunciou um plano de emissão zero de gás carbônico na geração de energia até 2050. No entanto, devido as dificuldades do setor de biomassa, contam a curto prazo apenas com fontes renováveis como energia solar e eólica.

Em 2016, segundo a EIA [2], vide a Figura 1.2, apenas 2% de toda energia elétrica gerada era proveniente de biomassa. Dentre essa quantidade, a madeira 33% se encontra como a principal fonte, seguida pelo licor negro, 27%, gás de aterros sanitários, 16%, resíduos sólidos biogenéticos e não-biogenéticos, 20%, e finalmente, as outras fontes com os demais 5%.

Em vista das dificuldades de adesão a geração de energia elétrica a biomassa, políticas globais visaram incentivar o uso de renováveis. Eventualmente, as políticas de passaram a ser objeto de estudos de diversos trabalhos que buscaram avaliar sua eficiência e seus impactos no mercado de energia. Dentre alguns trabalhos, [13] avaliou os impactos das políticas do mercado do carbono estabelecidos na união europeia na contribuição de energias renováveis.

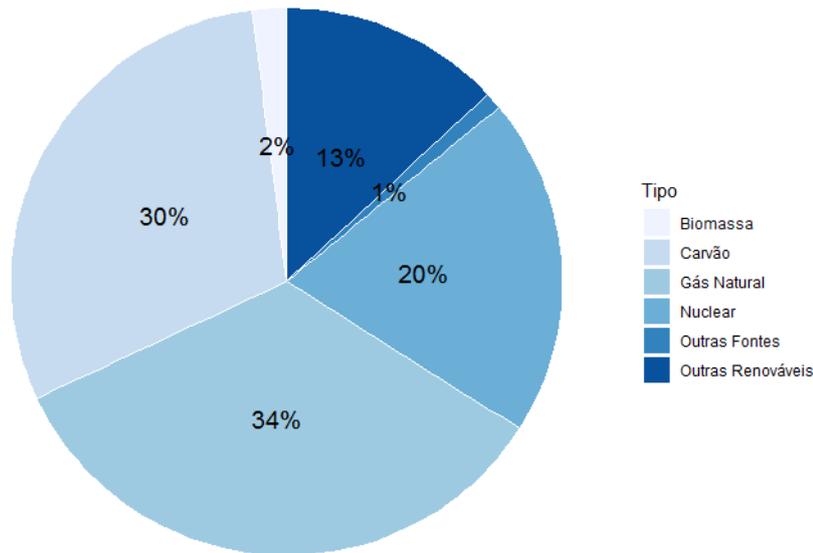


Figura 1.2: Matriz elétrica dos Estados Unidos da América, ano 2016. Adaptado da fonte:[2]

Em paralelo as análises de mercado que incentive o desenvolvimento sustentável, a política de financiamento foi abordado por [14] a fim de estabelecer uma metodologia para avaliação de investimento em portfólio baseado em energias renováveis, o estudo denota a importância de acompanhar a evolução das tecnologias para cada fonte renovável. Apesar da iniciativa política, [15] abordou as estratégias para desenvolvimento sustentável sob cenários de probabilidade alertando através de seus resultados que todos os setores de energia aumentariam, menos o consumo de biomassa.

O Brasil, ao contrário dos exemplos anteriores, apresenta grande diversificação em sua matriz elétrica. Ainda assim, estão sendo estimuladas pela lei RenovaBio, elaborada pela Presidência da República [2017] [16]. o Balanço Energético Nacional, BEN [3], indica que a contribuição da biomassa é bem menor com seus 8,2%, vide a Figura 1.3. Especificamente, o uso do bagaço da cana-de-açúcar tem sido a principal fonte de biomassa destinada a matriz elétrica. Como consequência desta aderência, [17] analisou a produção brasileira e avaliou a contribuição do país para o mercado mundial com exportação de biomassa para a produção de energia em outros países.

Além disso, as diretrizes do mercado brasileiro de energia elétrica têm se voltado para a questão da precificação horária que, segundo a Câmara de Comercialização de Energia Elétrica, CCEE [18], foi proposto pela Comissão Permanente para Análise de Metodologias e programas Computacionais do Setor Elétrico – CPAMP para que o cálculo fosse realizado a partir de janeiro de 2021 como definido pela Portaria MME 301/2019 [19]. De forma a atender

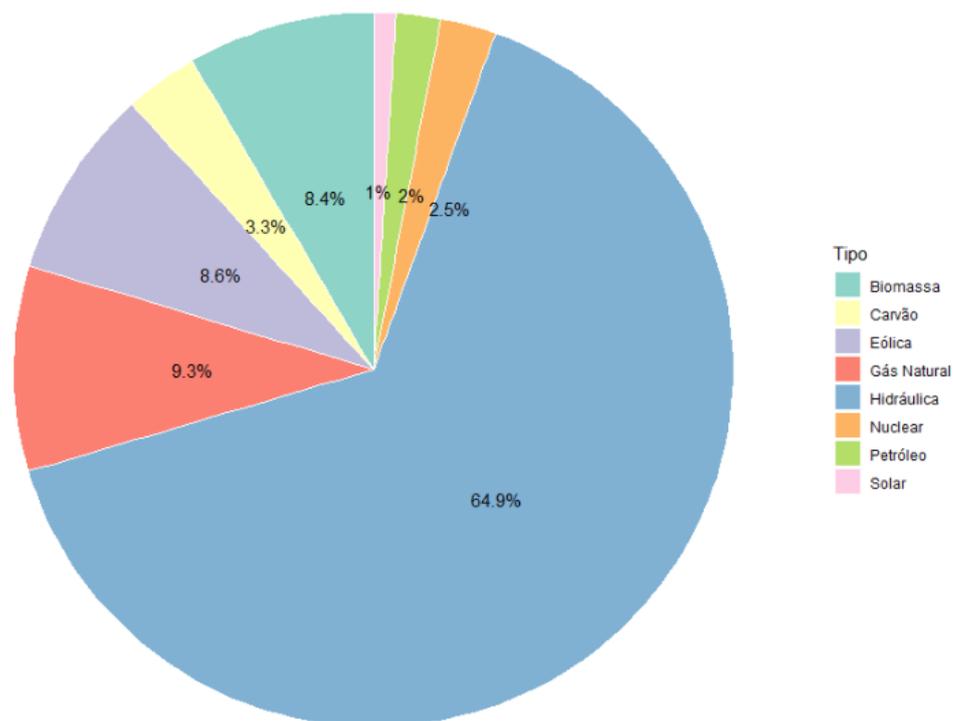


Figura 1.3: Matriz elétrica brasileira, ano 2018. Adaptado da fonte:[3]

as novas tendências, esta dissertação visa aplicar as técnicas de simulação para a geração termoelétrica de usinas a biomassa para gerar cenários de granularidade horária.

Este documento segue a seguinte estrutura: O capítulo 1 faz a introdução contextual no âmbito das usinas a biomassa; O capítulo 2 traz uma revisão da literatura, em que são destacados artigos relevantes e é definido o estado da arte desta dissertação. O capítulo 3 contém os modelos e técnicas selecionadas para a realização das simulações. O capítulo 4 apresenta toda a análise de dados de modo a destacar as peculiaridades do estudo de caso. O capítulo 5 contém os resultados e as comparações quanto ao desempenho dos modelos. O capítulo 6 faz as conclusões e ressalvas quanto aos resultados de simulação e desempenho. Cada um dos capítulos posteriores, com exceção do capítulo de conclusão, possuirá uma seção com o resumo.

2

Revisão de Literatura

A revisão da literatura desta dissertação tem o objetivo de destacar as metodologias utilizadas para previsão e simulação de séries de geração de energia elétrica a biomassa. Para esta energia renovável foram identificadas diversas abordagens para a previsão ao longo da revisão, partindo de contextos macroeconômicos até as abordagens peculiares a unidades geradoras através de metodologias autoregressivas.

A primeira abordagem se baseia em aspectos espaciais e macroeconômicos de uma região ou país. Dessa forma, é natural que esta abordagem não considere nuances de operação de uma determinada usina, mas um aglomerado de unidades geradoras avaliados em balanços energéticos mensais ou até anuais. Esta abordagem se baseia no aspecto de tomada de decisão para gerar ou não a energia, sob a condição da disponibilidade da matriz energética. Portanto, em muitos casos, a previsão da quantidade de biomassa muitas vezes também é endereçada a geração de energia oriunda a essa matriz.

A revisão da literatura traz uma breve comparação com as gerações de energia, eólica e solar. O motivo desta comparação é avaliar como a energia a biomassa tem sido abordada frente as demais fontes renováveis que aumentaram sua participação no mercado de energia recentemente.

Um extenso trabalho apresentado em [20] apresenta as modelagens fisicamente os diversos processos de geração de energia a biomassa, onde foram avaliados as diversas matrizes energéticas e sua respectiva eficiência. No entanto, pelo fato das usinas operarem sobre a tomada de decisão, a modelagem física não pode ser atrelada as estocasticidades de velocidade de vento, nem em irradiação solar como os apresentados em [21] e [22].

Como mencionado anteriormente, as séries de geração a biomassa eram abordadas como problemas de tomada de decisão. Uma evidência disso foram trabalhos que atrelaram previsão a geração com a projeções futuras sobre a produtividade nacional da matriz energética. Como exemplos destes trabalho, [23], [24], [25] e [26] levantaram informações geográficas sobre as produções da matriz energética de forma a estimar o potencial energético de países a longo prazo. já em [27] abrangiu a metodologia estimando a produção de energia elétrica anual para as fontes renováveis solar, eólica e diversas fontes de biomassa como cultivos e florestais.

A segunda metodologia de previsão de energia a biomassa foi a modelagem matemática. Esta técnica aborda o problema através de restrições físicas

representadas por equações e inequações que representam o tratamento dos recursos da usina. O objetivo da modelagem é maximizar o potencial de geração de biomassa a partir de decisões tomadas considerando novamente as projeções estatísticas dos recursos, esta abordagem está representada em trabalhos como [28], [29] e [30].

[Melhorar] O embasamento da produção energética sobre projeções de recursos a médio a longo prazo se deve pela suposição de que o operador da usina assume a arbitragem de gerar energia consumindo toda a matriz energética disponível. Tal suposição está coerente, dado que a matriz energética é perecível, no entanto, não é adequada para se tratar do curto prazo a médio prazo, considerando as sazonalidades e o despacho de energia.

A modelagem matemática também foi aprimorada através da implementação de distintos cenários plausíveis, atribuindo-se, pois, a definição de otimização com estocasticidade. Tais cenários podem ser obtidos através de estatísticas ou simulações de Monte Carlo como em [31], [32], [33] e [34].

Alguns trabalhos visaram tratar apenas de certas usinas deixando o contexto macroeconômico e passaram a abordar projeções de variáveis de mercado para a viabilização das usinas. Em [35] relata-se um estudo de caso em que se buscou obter a viabilidade financeira considerando a comercialização de energia brasileira e períodos de safra através de simulações. Já em [36], as usinas dispunham também de uma fonte solar, concluiu-se que o uso exclusivo de biomassa só era afetado em caso de alto custo de aquisição da matriz energética. Além disso, [37] chegou a realizar um estudo sobre a flexibilidade de uma cultura de cana de açúcar para a produção de bagaço destinada a geração de energia elétrica ou etanol, para o transporte.

Ao adotar a estocasticidade, metodologias mais sofisticadas tomaram espaço na literatura, a primeira delas foi a regressão dinâmica. Em [38] utilizou-se a metodologia de regressão dinâmica em dados em painel de países desenvolvidos e em desenvolvimento para a previsão anual de demanda e geração de energia a biomassa. Enquanto [39] realizou a comparação da regressão dinâmica com os modelos da família ARMA e gray sistem GM, longo prazo prever quantidade de biomassa [40] visou o mesmo objetivo, mas utilizando previsões baseadas no modelo de holt-winters. [41], diferentemente dos demais, realizou previsões a curto prazo através de modelos regressivos, algoritmo genético.

O aprendizado de máquina, especificamente tratando de redes neurais, também tomou protagonismo na previsão de geração de energia elétrica. No contexto da energia a biomassa, [42] realizou previsões da energia renovável através com uma abordagem generalizada, utilizando nenhuma variável expli-

cativa ou suposição. Enquanto [43] e [44] aderiram informações geográficas ao modelo, como localização e variáveis climáticas.

Trabalhos como [45], [46], [47] e [48] utilizaram redes neurais com lógica Fuzzy embuída na metodologia, FNN, para realizar previsões em unidades de geração [49] e [50] utilizaram redes neurais para prever a produção de biogás oriundo de diferentes processos químicos a fim de estimar um potencial de geração. [51] realizou testes com diversas metodologias de redes neurais a fim de estabelecer fatores essenciais para realizar as previsões de geração de energia a biomassa, não sendo muito conclusivo. Embora as Redes Neurais tenham se apresentado expressivamente no contexto de previsões, ressalta-se que há as dificuldades de implementação de simulação de cenários de geração para metodologias de inteligência artificial, [52].

[Ressalva para as demais fontes solar e eólica] Importante ressaltar que as metodologias também foram aplicadas as demais fontes renováveis. A regressão dinâmica está presente também na energia solar, [53], [54], [55] e [56]. Os modelos ARIMA se encontram tanto na eólica, [57], [58], [59] e [60], quanto na solar, [61], [62], [63] e [64]. E finalmente, as redes neurais que foram intensamente exploradas em ambas fontes de energia eólica, [65], [66], [67], [68], e solar, [69], [70] e [71].

Em paralelo as novas tendências da aplicação do aprendizado de máquinas, as Cadeias de Markov, cuja metodologia será apresentada no próximo capítulo. Elas foram utilizadas em diversos contextos de previsão, dentre eles, originalmente, a área de finanças para as previsões de índices de valores mobiliários, [72], [73], [74] e [75]. Apesar disso, ressaltamos os estudos em que as cadeias de Markov foram utilizados para prever cenários de geração de energia renovável.

O primeiro trabalho identificado foi [76], que combinou o uso das Cadeias de Markov com modelo auto-regressivo discreto utilizando atributos de velocidade e direção do vento. Trabalhos como [77], [78] e [79] deram sequência a aplicação nesta área de previsões a curtíssimo prazo relatando que a modelagem poderia ser diretamente implementada a série de geração de energia em vez de sua matriz energética, velocidade de vento.

A aplicação das cadeias de Markov tomou espaço na geração fotovoltaica, onde [80] e [81] utilizaram uma classificação de estados através de variáveis climáticas para a previsão da geração a longo prazo, enquanto [82] e [83] utilizaram apenas a irradiação e temperatura para estabelecer as condições de operação, como estados transitórios, para previsões de 15 em 15 minutos. Também foram identificados trabalhos feitos em cima de cultivos agrícolas para estimar o potencial energético de energia a biomassa em [84] e [85]

Introduzindo o conceito de simulações, cada metodologia previsão apresentada está atrelada a uma ou mais metodologia de simulação adequada a abordagem do problema. Para o caso das cadeias de Markov, foram implementadas a metodologia de simulações de Monte Carlo. Dentre os trabalhos de simulação voltados para geração de energia, [86], [87], [88] e [89] demonstraram que o uso exclusivo das simulações de Monte Carlo é uma forma eficiente para a construção de cenários de geração em sistemas isolados com fontes renováveis como, por exemplo, a micro-geração solar.

Posteriormente, [90] incorporou as simulações de Monte Carlo as Cadeias de Markov, recebendo a nomenclatura de MCMC (Markov Chain Monte Carlo) aplicando diretamente na série temporal de geração de energia eólica. [91] seguiu a inovação metodologia do MCMC aplicando uma distribuição probabilística acumulada da velocidade do vento para simulações de energia elétrica. [92] abordou o problema através da aplicação do algoritmo de Kernel para estimar a densidade probabilística da geração eólica. [93] inovou o processo de criação dos estados aplicando a Clusterização em k-means como um pré-processamento para o MCMC.

Dentre as aplicações do MCMC, não foi identificado nenhum estudo de caso relativo a geração de energia a biomassa. Dessa forma, a dissertação visa preencher este espaço da literatura aplicando uma versão adaptada do MCMC em uma fonte de energia renovável pouco estudada recentemente devido a disposição de dados de alta frequência.

2.1 Resumo do Capítulo

A revisão bibliográfica visou não só cobrir as abordagens e metodologias aplicadas as séries temporais do contexto de energia elétrica, mas também apontar a lacuna na produção científica quanto a geração a biomassa que tem acompanhado ao passo do avanço de novas tecnologias para as demais fontes de energia renovável. Parte desta lacuna se deve a premissa de que a usina a biomassa executa uma arbitragem de gerar sempre que há biomassa disponível por se tratar de uma operação controlada desde seu fornecimento, até seu gerenciamento. De fato, esta premissa não é totalmente falsa, porém desconsidera as nuances sobre a operação da usina, permitindo apenas realizar previsões apenas de médio a longo prazo.

Conforme os conceitos de estocasticidade foram considerados nas estimativas, as modelagens deixaram de tratar a geração a biomassa em escalas nacionais com baixa granularidade e passaram a realizar estudos de casos para viabilidade financeira de certas unidades geradoras. A partir desta nova tendên-

cia, foram identificados artigos que abordaram desde as metodologias clássicas, como modelos de regressão e de análise de séries temporais, até o estado da arte, como inteligência artificial e redes neurais. Esta inovação revelou a complexidade que estava por trás do gerenciamento de uma usina a biomassa que era desconsiderado ao relevar apenas as estatísticas territoriais de uma região abrangente.

As gerações eólica e solar foram brevemente abordadas na revisão da literatura, de modo a estabelecer um paralelo entre as abordagens, pois, apesar de apresentarem usualmente capacidades de geração próximas, possuem suas peculiaridades quanto a operação. As gerações eólica e solar são impactadas diretamente pelos atributos da natureza, velocidade de vento e irradiação, respectivamente, enquanto a produção de energia a biomassa depende indiretamente das safras e tomada de decisão por parte do gerenciamento manual. Embora sofram diferentes impactos na geração, a dissertação foi inspirada em trazer uma nova metodologia para a aplicação no contexto de geração de energia elétrica a biomassa através das simulações de Monte Carlo aplicadas as Cadeias de Markov.

3 Modelagem Matemática

A metodologia se baseia na simulação de Monte Carlo aplicada as Cadeias de Markov. Esta metodologia, apesar de ser matematicamente elaborada de forma simples, traz resultados suficientemente satisfatórios em questão de aderência dos cenários as características intrínsecas de uma série temporal. Outra vantagem a ser considerada é que o uso desta técnica em simulações se mostrou tão eficiente quanto as simulações em inteligência artificial. , [94].

A metodologia proposta está representada pela Figura 3.1 consiste em quatro etapas. A primeira é a divisão do histórico em subconjuntos para cada estágio de geração. Esta etapa está atrelada a análises de tendência e sazonalidade da série temporal. A segunda é a Clusterização, este processo visa agrupar de maneira a reduzir o número de estados do processo. A terceira é a construção da matriz de transição de estados, elaborada pelo conceito das cadeias de Markov. Finalmente, é realizada a simulação de Monte Carlo com base nas estatísticas estimadas da construção da matriz de transição.

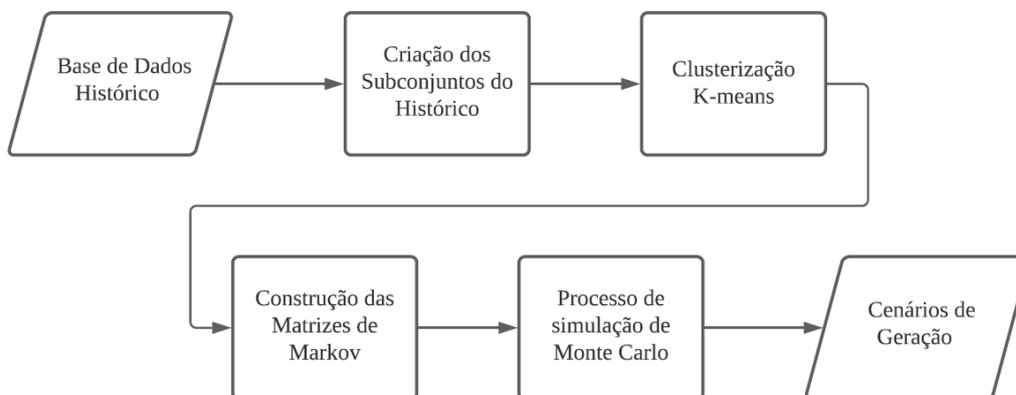


Figura 3.1: Fluxograma da metodologia

3.1 Subconjunto do Histórico

Nesta etapa, o objetivo é a identificação de diferentes tendências ao longo do histórico que será realizado ao passo das análises de dados. Ao serem identificadas as tendências da geração, será estabelecido uma divisão em etapas de modo que possam descrever a sazonalidade da série durante o período de safras.

Uma questão bastante recorrente a esta aplicação é a possibilidade de se determinar de uma janela de tempo homogênea para estimar as matrizes de probabilidade de transição, [95] trabalhou em cima desta hipótese para índices da bolsa de valores. No entanto, esta dissertação propôs determinar arbitrariamente as janelas de tempo atribuídas para as etapas de geração de acordo com as análises de dados.

Com a definição de etapas de geração, visa-se a separação de diferentes distribuições probabilísticas identificadas ao longo do processo. As distribuições são resultados da evolução do processo ao longo do tempo e, portanto, cada distribuição identificada, acarreta a necessidade de uma matriz de transição diferente a ser estimada. Após uma separação bem definida, é esperado que as matrizes tenham uma maior eficiência para que possam reproduzir as tendências da série pertencente a cada etapa.

No próximo capítulo, estão detalhadas as análises de dados que levaram a divisão em etapas. Foram identificadas até três etapas, são elas a de início de geração, geração plena e término de geração. Para cada estágio de geração serão aplicados os processos de Clusterização, construção de matriz de transição e simulação de Monte Carlo.

3.2 Clusterização k-means

O objetivo da Clusterização é definir uma quantidade finita de valores, tornando os dados de geração em valores discretos. Isto viabiliza a construção de uma matriz de transição de dimensões $k \times k$, onde k é o número de clusters. A metodologia foi elaborada por Lloyd em 1957, porém foi formalmente publicada apenas em 1982 por [96]. Posteriormente, [97] e [98] fizeram versões alternativas do algoritmo de Clusterização, mas somente Hartigan e Wong desenvolveram uma versão mais eficiente em [99], a qual foi utilizada nesta dissertação pela linguagem de programação R.

Importante ressaltar que, originalmente, a Clusterização é utilizada na classificação de grupos dentro de uma amostra com diversos atributos. No entanto, para o estudo de caso em questão, há apenas um único atributo a ser classificado, a geração de energia a biomassa. Como consequência disso, o algoritmo k-means irá classificar a série através de médias, emulando o processo de discretização.

A Clusterização k-means propõe dividir os dados obtidos de geração em uma quantidade finita de grupos, K . Diferente de um processo de discretização simples, a classificação se dará de forma que a soma dos quadrados das distâncias entre os valores ao centróide seja minimizada. O centróide, por sua

vez corresponde a média do conjunto, M_k , e ao permanecerem fixos após certo número de iterações, o processo é finalizado. Ao final, os dados do histórico são substituídos pelos centróides correspondentes ao seu cluster, c_k construindo assim, um histórico com valores discretos, GD_i . A Equação 3-1 ilustra este processo:

$$GD_i = \begin{cases} M_1, & \text{se } G_i \in c_1 \\ M_2, & \text{se } G_i \in c_2 \\ \vdots & \\ M_k, & \text{se } G_i \in c_k \end{cases} \quad (3-1)$$

Uma condição para determinar que a Clusterização tenha um impacto positivo é a de que os valores contínuos do histórico sejam bem representados pelo centróide. Esta condição implica no fato de que o centróide, por ser uma média de valores atribuídos a um cluster, deve apresentar uma pequena dispersão de valores dentro do próprio cluster. Este requisito pode ser traduzido em um menor desvio padrão calculado para cada cluster.

Ao longo da revisão da literatura [100], um critério de avaliação condizente com a necessidade foi identificado, sendo denominado de "elbow method". O método se baseia na minimização da soma total dos desvios-padrões dos clusters. São feitas diversas clusterizações, aumentando a quantidade imposta de clusters a serem estimados. A quantidade é definida quando o desvio padrão total converge, ou seja, quando não há mais variações significativas ao incluir mais um cluster na partição.

Através do "elbow method", visa-se a minimização dos desvios-padrões de cada clusters tenham sido conciliados com a minimização da soma e, ao mesmo tempo, a definição de um número reduzido de estados. Como resultado, tem-se que os centróides passam a ter mais acurácia na representação dos valores históricos agrupados, ocasionando no aumento na eficiência da Clusterização.

3.3

Matriz de Transição

Seguindo para o próximo passo da metodologia, assume-se que cada cluster é interpretado como um estado de um processo estocástico. Segundo [101], o processo estocástico é por definição a evolução de uma variável aleatória ao longo do tempo, particularmente a variável se trata da geração discretizada. A partir dessa definição, segue que a evolução de um processo estocástico resulta em uma determinada distribuição de probabilidade ao analisar toda a amostragem. Um processo estocástico cujo número de estados é finito é denominado uma Cadeia de Markov.

Seja a probabilidade de transição, $P_{a,b}$, a probabilidade do processo estocástico assumir o valor do estado b , partindo de um estado a . Cada probabilidade é um elemento da matriz e é calculado pela Equação 3-2:

$$P_{a,b} = \frac{n_{a,b}}{\sum_k n_{a,k}} \quad (3-2)$$

Onde $k \in K$, conjunto de todos os estados, $n_{a,b}$ é o número de vezes em que o processo assumiu o valor b , partindo do estado a e $n_{a,k}$ o número de transições em que o processo assumiu qualquer estado k a partir de a . A Equação 3-2 resulta na probabilidade empírica da transição de estado, por definição, todas as probabilidades estão compreendidas em 0 e 1 e a soma de todas elas dado o estado inicial a é igual a 1.

Ao serem estimadas, as probabilidades de transição são arranjadas construindo assim a matriz de probabilidade de transição, P , descrita pela Equação 3-3:

$$P = \begin{pmatrix} P_{1,1} & P_{1,2} & \cdots & P_{1,K} \\ P_{2,1} & P_{2,2} & & P_{2,K} \\ \vdots & & \ddots & \vdots \\ P_{K,1} & P_{K,2} & \cdots & P_{K,K} \end{pmatrix} \quad (3-3)$$

A aplicação das cadeias de Markov exige certos cuidados para se tratar de um processo estocástico. Dentre eles, uma forma de inviabilizar a metodologia é a existência de um estado absorvente, isto é, quando a probabilidade de um estado transitar para ele mesmo é igual a 1. Em termos práticos, um estado pode ser considerado absorvente caso a probabilidade seja muito próxima a 1.

A presença de um estado absorvente prejudica a simulação de Monte Carlo de forma que ao atingir este estado, a simulação reproduz esse valor até o final do horizonte de tempo estipulado. Segue, portanto, que a distribuição probabilística das simulações e a evolução ao longo do tempo são comprometida, fato a ser analisado durante as análises de resultados. De forma a contornar esta complicação, os períodos atribuídos para cada etapa de geração devem estar bem definidos.

Caso as estimativas levem ao caso de sucesso, em que não há estados permanentes, é dito que a matriz trata de um processo ergódico. Uma característica interessante sobre este tipo de classe de cadeia de Markov é que o processo tende a uma distribuição probabilística. Tal fato requer que as etapas sejam bem definidas a ponto de consolidar os dados em uma distribuição probabilística única representativa de sua respectiva etapa.

Seguindo a metodologia de [91] e [93], é demonstrada a construção das matriz de probabilidade acumulada. Esta matriz é construída através da Equação 3-4 apresentada abaixo:

$$P_{ac} = \begin{pmatrix} P_{1,1} & P_{1,1} + P_{1,2} & \cdots & P_{1,1} + P_{1,2} + \cdots + P_{1,K} \\ P_{2,1} & P_{2,1} + P_{2,2} & & P_{2,1} + P_{2,2} + \cdots + P_{2,K} \\ \vdots & & \ddots & \vdots \\ P_{K,1} & P_{K,1} + P_{K,2} & \cdots & P_{K,1} + P_{K,2} + \cdots + P_{K,K} \end{pmatrix} \quad (3-4)$$

Eventualmente, a última coluna apresenta o valor unitário em todas as linhas, por justamente a soma total de probabilidades ser igual a 1.

3.4 Simulação de Monte Carlo

As simulações realizadas nesta dissertação seguiram a metodologia de Monte Carlo sob a distribuição probabilística uniforme. [102], [103] e [104] apresentam extensas revisões dos trabalhos feitos em cima de simulações de cadeias de Markov.

Foram encontrados alguns trabalhos que analisaram as simulações e os fatores que tornam as simulações aderentes a metodologia. Dentre eles, [105] ressalta que as simulações de Monte Carlo devem respeitar as probabilidades de transição em cadeias de markov para realizar resultados sem viés, enquanto [106] e [107] estabeleceram os aspectos que aumentam a eficácia das simulações de cadeias de markov quanto a aderência.

O processo de simulação de Monte Carlo, segundo [108], é um processo iterativo, em que a cada instante t é realizado uma amostragem sob uma densidade de probabilidade uniforme P_{Unif} entre 0 e 1, equação 3-5:

$$p_{t+1} \sim P_{Unif[0,1]} \quad (3-5)$$

Este valor é classificado de acordo com os patamares da distribuição de probabilidade acumulada de acordo com o estado atual. Se a probabilidade amostrada for maior que seu estado antecessor e menor ou igual a um estado sucessor, é definido o estado sucessor para atribuir o valor do respectivo centróide no instante em $t + 1$, portanto, o valor da simulação em Sim_{t+1} será um dos valores discretos da Clusterização. A equação 3-6 demonstra o processo:

$$Sim_{t+1} = M_k \iff P_{ac[s,k-1]} < p_{t+1} \leq P_{ac[s,k]} \quad (3-6)$$

Onde $k \in K$, e s é o estado que o processo assume no instante t . $P_{ac[s,k]}$ é o valor da probabilidade acumulada dado o estado s no instante t para cada cluster k . Após a classificação, é atribuído o centróide M_k correspondente ao cluster selecionado.

Ao final, as simulações são analisadas quanto a sua aderência as estatísticas do histórico e evolução do processo ao longo do tempo. As avaliações estatísticas serão feitas através de testes hipóteses em que serão avaliadas as diferenças de médias e variâncias de amostras, histórico e cenários de simulação, que naturalmente possuem tamanhos diferentes.

3.5

Resumo do Capítulo

O capítulo traz de maneira didática a metodologia utilizada em [93] passo a passo. Como esta metodologia teve seu desempenho reconhecido no contexto da geração de energia, a dissertação tem a intenção de estender para as subcategorias de energia renovável. Como citado inicialmente, diferente de uma abordagem mensal, a metodologia aborda a geração em etapas, devido a um esperado comportamento de geração constante devido as pressuposições citadas no capítulo anterior.

Trata-se de uma modelagem simples, porém ainda robusta, considerando o processo de criação de estados de transição. O processo de Clusterização k-means é um dos clássicos algoritmos na área de ciência de dados e, particularmente, por este caso se tratar de um problema univariado, sua interpretação é interpretável como patamares de geração. Outro fator que determina a eficiência do algoritmo é a aplicação do "elbow method" para estabelecer um número adequado de clusters conciliando a dispersão dos dados em cada clusters, desvio-padrão.

Com o número de clusters bem definidos, os processos de construção das matrizes não perderam sua eficiência na interpretação de suas probabilidades de transição. A princípio, a redução para um número de estados finitos menor pode soar contraintuitivo nos quesitos de precisão do modelo. No entanto, a forma de como os clusters foram estabelecidos, citada anteriormente, resulta em uma acurácia grande o suficiente para que, mesmo com um número menor de estados, seja possível reproduzir as tendências da série temporal.

Finalmente, a simulação com base em Monte Carlo busca reproduzir as tendências da série temporal de acordo com as matrizes de transição. O processo de simulação de Monte Carlo tem sido frequentemente utilizadas em diversas áreas devido a sua confiabilidade. Ao final, as simulações são avaliadas

por sua aderência a série temporal de acordo com a medidas estatísticas e distribuição probabilística.

4

Análise de Dados

Este capítulo traz toda a análise da base de dados fornecida por uma concessionária de energia localizada na região nordeste do Brasil. A base de dados consiste em uma série temporal de geração de energia elétrica oriunda de uma usina a biomassa com intervalos de 15 em 15 minutos. O bagaço da cana-de-açúcar, monocultura expressiva da própria região, é a matriz energética utilizada. Estas usinas apresentam períodos diferentes no histórico, elas não entraram em operação mutuamente.

A usina apresenta uma geração contratada a 3000 kW e um histórico que compreende um período de três anos. A figura 4.1 apresenta o históricos da usina a biomassa. O histórico apresenta uma sazonalidade anual bastante evidente, com meses em que a geração cessa completamente. Estes meses recebem a denominação de um período intermitente, no entanto, uma análise mais apurada é necessária para a identificação do horizonte de tempo. Enquanto nos períodos de geração, esta se encontra em patamares estáveis próximas a capacidade contratada.

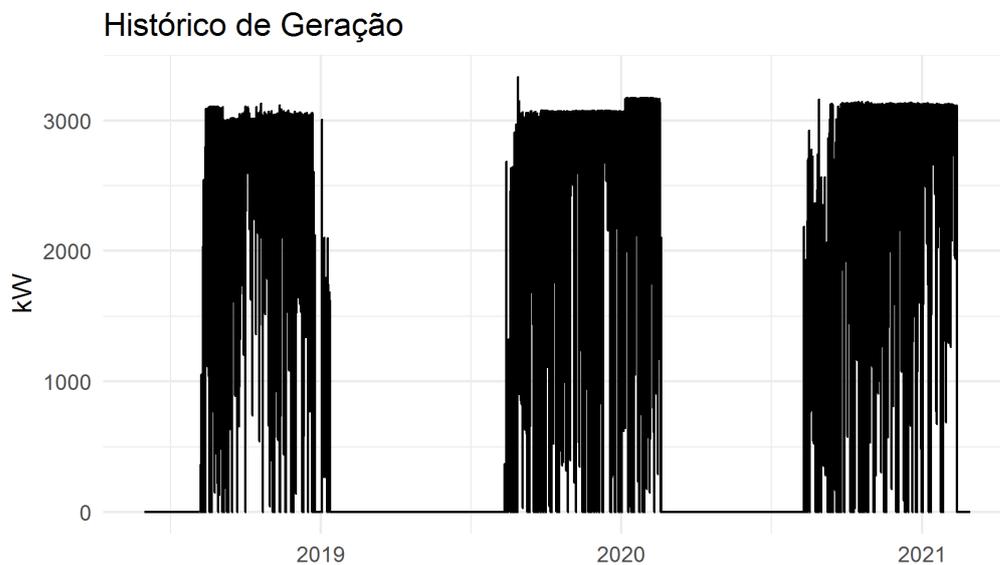


Figura 4.1: Histórico de Geração de Energia Elétrica a Biomassa.

4.1 Análise de Tendência e Sazonalidade

Como a geração está atrelada às safras de produção agrícola, assume-se uma periodização anual devido às estações do ano definirem o ciclo de cultivo das plantações. Dado esta afirmativa, foi realizado o boxplot relativo a cada mês do histórico de modo a identificar os meses que compõem o período intermitente. É considerado o período intermitente os meses consecutivos que apresentam a sua respectiva mediana e quartis nulos, apontando o período em que a geração é definitivamente nula. A figura 4.2 apresenta o boxplot mensal da usina.

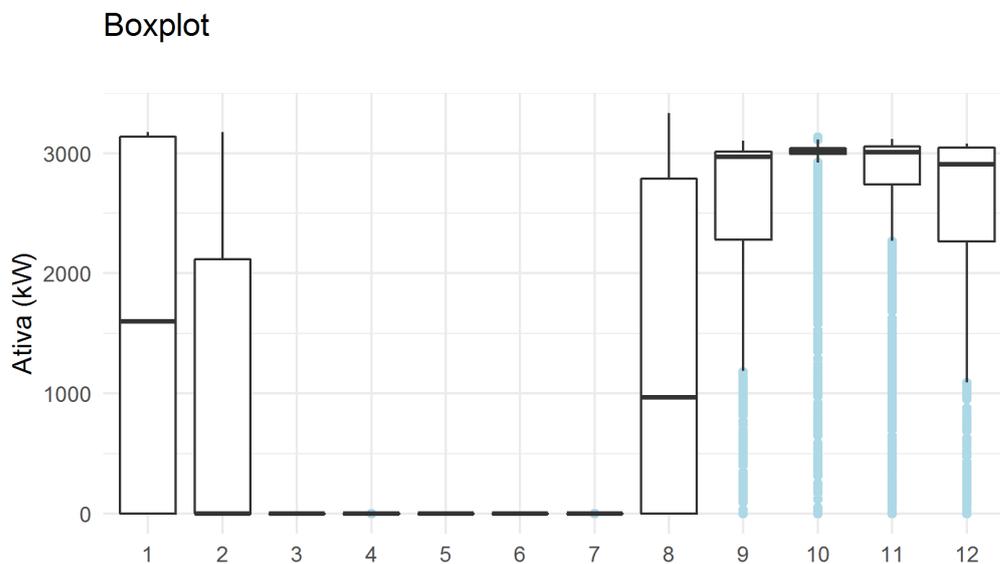


Figura 4.2: Boxplot mensal.

De acordo com o critério estabelecido para a Figura 4.2, o período intermitente corresponde aos meses de fevereiro a julho. Em contrapartida, observa-se que a geração se mantém estável e próxima à capacidade instalada de setembro a dezembro. Além disso, o gráfico traz a informação de que a geração não começa nem termina abruptamente entre um mês e outro, havendo assim uma transição entre a intermitência e o período de geração. Deve-se também considerar que a geração começa e termina em meados dos meses de agosto e janeiro, respectivamente.

Um olhar mais minucioso é trazido para a análise da série ao desconsiderar o período intermitente. Desta vez a Figura 4.3 representa o boxplot horário, após omitir o período de meados de fevereiro até meados de agosto. Dentre os objetivos deste gráfico, estaria o questionamento sobre a influência das horas na operação da usina.

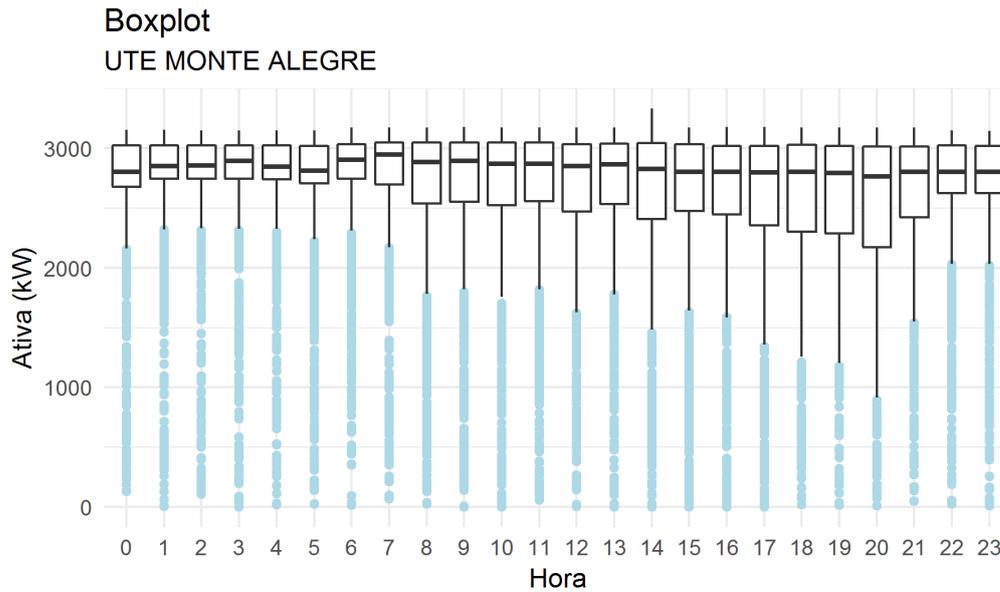


Figura 4.3: Boxplot Horário sem o Período Intermitentes.

A Figura 4.3 apontou curvas estáveis tanto para as medianas quanto para os quartis horários. Este comportamento significa que a usina apresenta um processo estável de abastecimento. As medianas e quartis assumiram valores muito próximos, indicando uma geração estável, enquanto são apontados "outliers", em azul, cujos valores são muito abaixo da capacidade de geração. Estas discrepâncias são oriundas de ligeiras, porém significativas, variações ocorridas durante todo o período de geração. Em suma, é possível afirmar que a série não sofre influências horárias.

A fim de trazer uma abordagem mais mensurável, é feita a decomposição STL sob o histórico de geração. O algoritmo de decomposição estima todos os fatores de nível, tendência e sazonalidades presentes na série. A Figura 4.4 apresenta a decomposição da série de acordo com os fatores enumerados.

Na Figura 4.4 são apresentadas as seguintes séries temporais, de cima para baixo tem-se: O histórico, a tendência, a sazonalidade anual, a sazonalidade semanal, sazonalidade diária, sazonalidade horária e o resíduo. Nela podemos observar a magnitude e a forma em que estão arranjadas as estimativas, ao lado, as barras de cada gráfico representa o quanto da variação da série está contida na decomposição. Quanto menor o tamanho desta barra, maior será a contribuição do componente para a variação da série temporal. Ressalta-se que o estas barras são relativas às discrepâncias de cada série e portanto estas devem ser relevadas para concluir qual componente mais influência a série.

Dentre as séries apresentadas pela Figura 4.4, duas se mostram relevantes, são elas as curvas de tendência e sazonalidade anual. Os fatores que mais

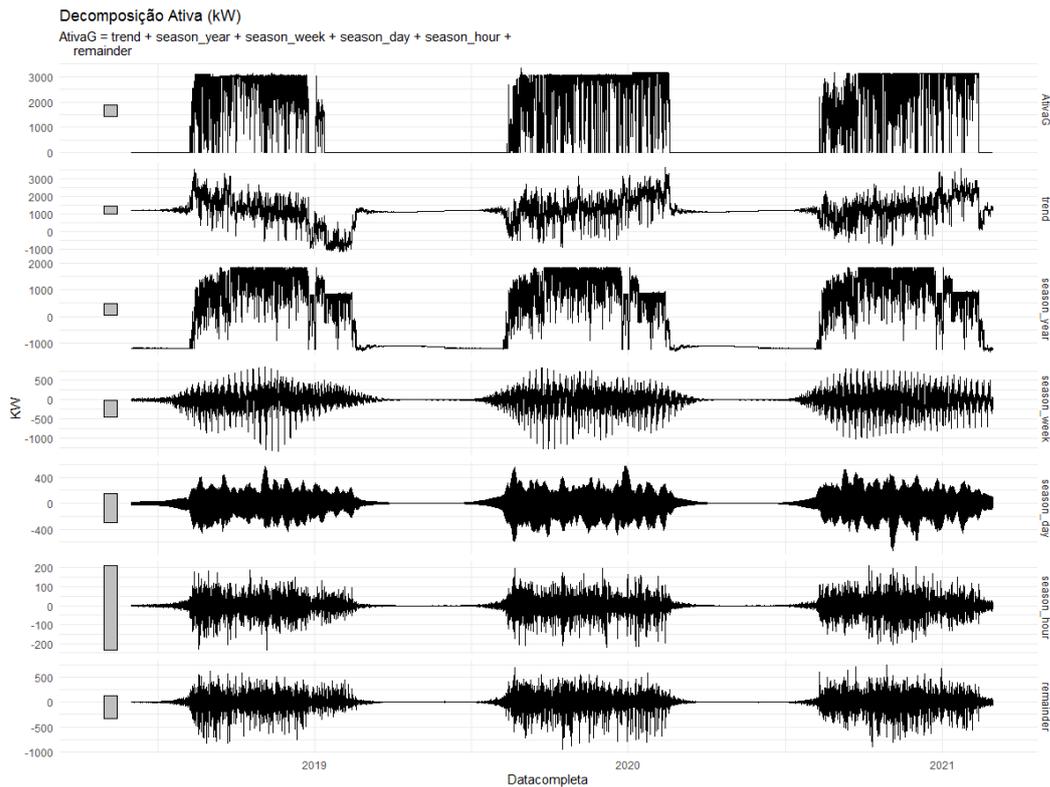


Figura 4.4: Decomposição dos Fatores Sazonais da Geração.

destacam essas curvas são a ordem de grandeza dos valores, variando de -1MW a 3MW, e a forma que estas séries se desenvolvem com o tempo, seguindo um arranjo linear. Apesar de serem os fatores relacionados a uma granularidade menor, anual e mensal, são os fatores que mais influentes para a reprodução da série histórica.

Por outro lado, nota-se que os fatores que só poderiam ser identificados em uma alta granularidade, como sazonalidade diária e intradiária, são os menos relevantes. Segundo a Figura 4.4, esses fatores mais se assemelham ao resíduo, que supostamente é tido por hipótese como um ruído, do que com qualquer outro dos fatores sazonais de baixa granularidade, tanto em ordem de grandeza, que desta vez se encontra dez vezes menor do que as componentes na ordem de -400kW a 400kW, quanto em dispersão cujo desenvolvimento ao longo do tempo se dá de maneira aleatória.

Embora seja contraintuitivo, este fenômeno já era esperado segundo as conclusões retiradas das Figuras 4.2 e 4.3. Ambos os gráficos, trouxeram a informação de que a sazonalidade anual exerce mais influência na produção de energia do que as variações intradiárias mesmo após a retirada do período intermitente. Tal conclusão é consolidada pela figura 4.4, que estimou os dos fatores sazonais de alta e baixa granularidade.

Em suma, as Figuras , demonstram que, embora a série possua dados de

15 em 15 minutos, ela pode ser dividida por extensos períodos, abrangendo de semanas até meses.

De maneira a preservar o período de geração plena, o qual se estende de meados de setembro até meados de janeiro, a divisão seguiu os padrões identificados pela figura 4.2. Assim sendo, a série foi dividida em três períodos que representam um período de geração plena, e dois períodos de transição, início e término de geração, que supostamente conciliam a mudança do período de geração com a intermitência.

A divisão segue a seguinte configuração: foi estipulado um período de um mês para a etapa inicial de geração, que se inicia em meados de agosto até meados de setembro; também foi feita a mesma atribuição para a etapa final de geração, a qual se inicia em meados de janeiro e se encerra em fevereiro; por fim, a etapa de geração plena, que compreende um horizonte de quatro meses, indo de meados de setembro até meados de janeiro.

4.2

Distribuição Probabilística

Seguimos as análises considerando a divisão descrita acima para avaliar as distribuições de cada etapa. Esta análise serve para o embasamento da aplicação das cadeias de Markov, como dito na seção anterior, os processos estocásticos resultam em uma distribuição probabilística ao final de sua amostragem. Embora a literatura não apresente uma distribuição probabilística clássica para a geração a biomassa, diferentes curvas de distribuição em cada etapa de geração é suficiente para justificar a estimativa de diferentes matrizes de transição.

Dessa forma, as curvas de densidade probabilística foram levantadas para cada etapa de geração. As etapas de início de geração, geração plena e término de geração são apresentadas, respectivamente pelas Figuras 4.5, 4.7 e 4.6.

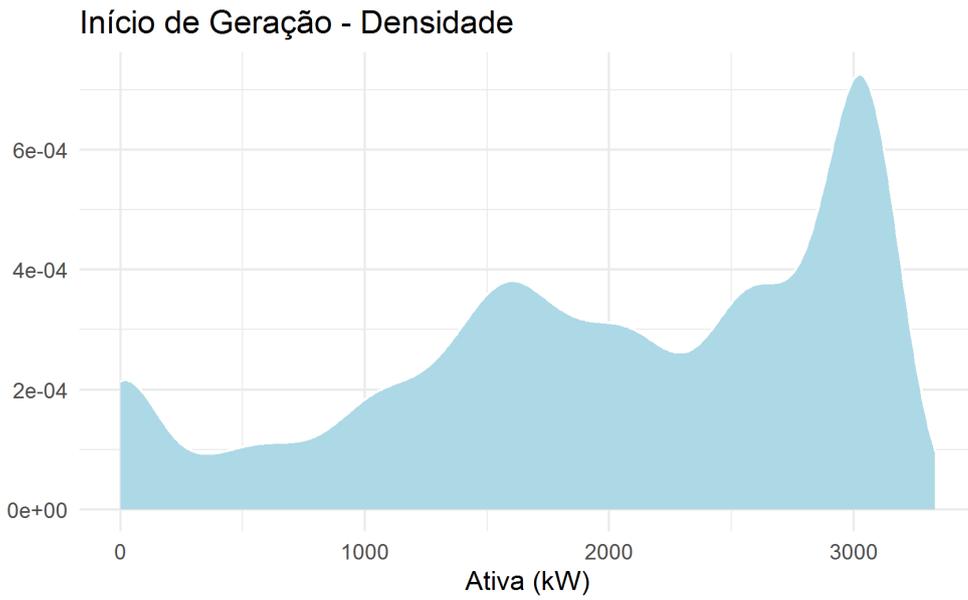


Figura 4.5: Densidade de Probabilidade - Início de Geração

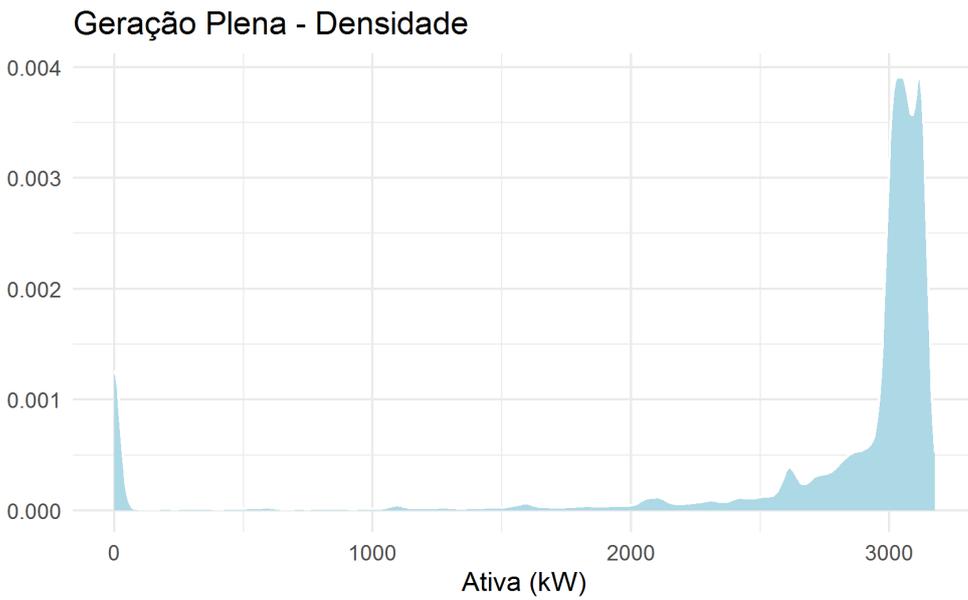


Figura 4.6: Densidade de Probabilidade - Geração Plena

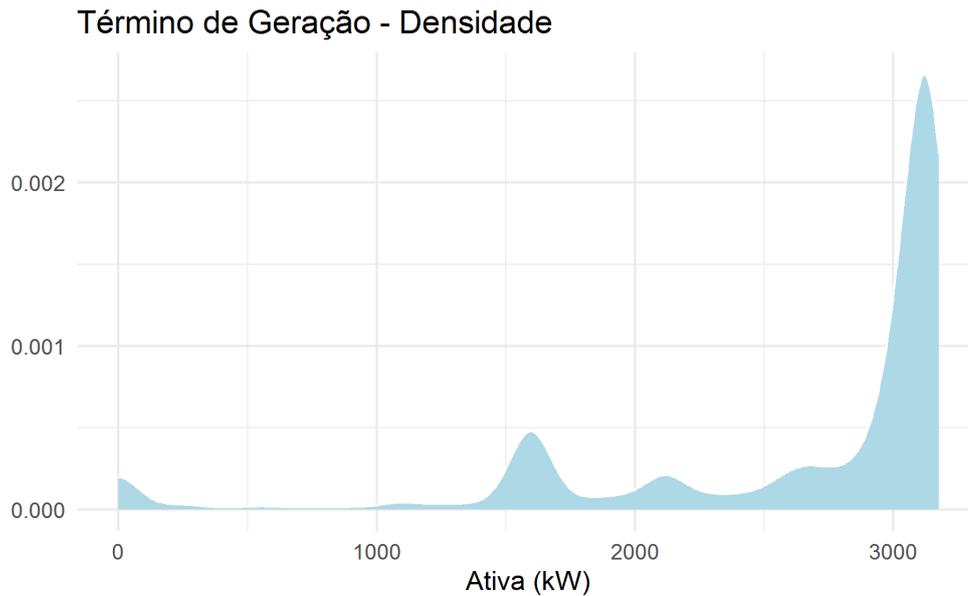


Figura 4.7: Densidade de Probabilidade - Término de Geração

Os resultados são evidentemente favoráveis a divisão em etapas de geração por terem apresentado diferentes densidades para cada etapa do processo. Inicialmente, a Figura 4.5 é a que mais diverge das demais, apresentando inúmeros picos de probabilidades, modas locais. É argumentável de que a divisão esteja envolvendo mais de uma tendência, porém, se tratando de um curto espaço de tempo de um mês cujos extremos divergem na ordem de 3 MW, essa separação se torna conveniente por abranger uma complexidade inerente a retomada da operação da usina.

Em segundo lugar, a Figura 4.6 apresentou uma análise promissora para a aplicação da metodologia. Em contraparte a etapa inicial, que apresentou uma densidade probabilística complexa em um curto período, a etapa de geração plena apresentou uma densidade probabilística concentrada próxima a capacidade instalada durante todos os meses de setembro a janeiro. Algumas ressalvas devem ser feitas, uma vez que ainda são identificados valores nulos, resultado da volatilidade apresentada o histórico.

Finalmente, tem-se a Figura 4.7 com uma densidade similar a etapa de geração plena. Tal fenômeno se deve pela usina estar partindo de um longo período de geração estável, conseqüente de um possível gerenciamento da matriz energética. No entanto, a divisão desta etapa de término de geração cumpre o objetivo de reter as nuances de operação da usina durante o período de transição, sendo esta peculiaridade apresentada por modas locais mais discretas do que as observadas na etapa inicial.

4.3 Divisão em Etapas

Em suma, conclui-se que a divisão da série temporal está dividida adequadamente em três etapas, início, meio e fim. A Figura 4.8 ilustra o histórico com a divisão das etapas ao longo dos anos de geração.

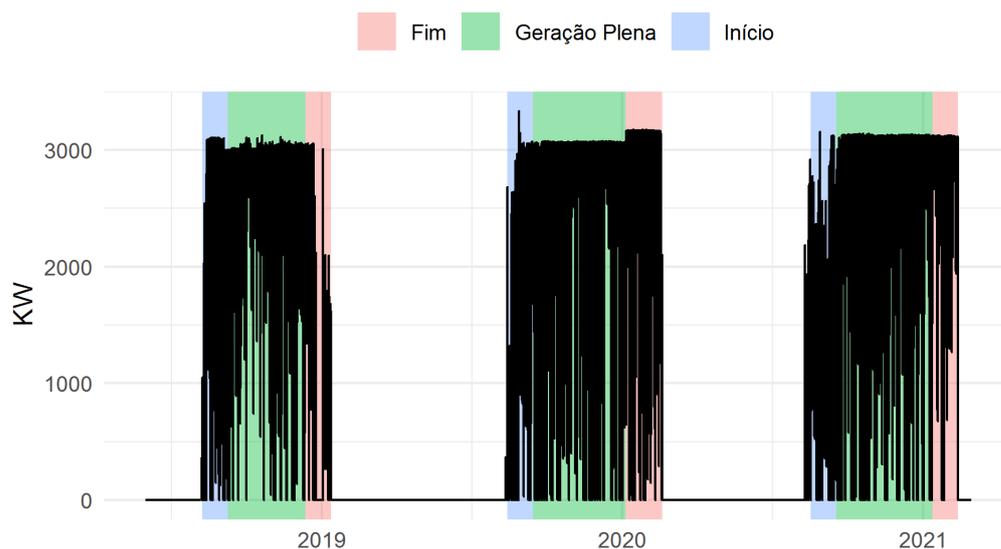


Figura 4.8: Subconjuntos do Histórico de Geração de Energia Elétrica.

4.4

Resumo do Capítulo

As séries temporais que possuem uma granularidade intra-horária tem sido a tendência para a implementação das novas tecnologias de previsão. No entanto, poucos estudos foram feitos sobre esses tipos de dados, mais raro ainda se tratando de geração a biomassa.

O capítulo da análise de dados trouxe informações sobre a operação de uma usina a biomassa. Foi notado que ao início de cada período de geração, a usina apresenta uma relutância para atingir a capacidade de geração. Tal relutância persiste até um horizonte de um mês, causando impactos na distribuição probabilística da série. O mesmo fenômeno foi percebido ao final da série, quando a usina está à beira de consumir o resto da biomassa.

As análises de distribuição probabilística dessas etapas levaram a conclusão de que se faz necessário a divisão da série em etapas. Esta divisão, visou preservar o período de geração em que a usina se encontra fornecendo próxima a sua capacidade máxima por meses. Tal período recebeu a definição de geração plena por justamente apresentar uma distribuição probabilística concentrada próxima aos 3000kW nominais da usina.

Ressalta-se que, após a divisão, em cada etapa será aplicado demais passos da metodologia e os resultados estão apresentados na próxima seção.

5 Resultados

Esta seção visa trazer os resultados da aplicação da metodologia em cada etapa da geração - início, meio e fim - que foram estabelecidos com base nas análises dos dados detalhadas na seção anterior.

Para cada etapa, será aplicado o algoritmo de Clusterização k-means na geração de energia. Pelo fato de estar tratando uma única variável, o resultado da classificação corresponde a um agrupamento por média. Conseqüentemente, os valores passam a ser representados pela média do cluster correspondente, levando a discretização da série.

5.1 Clusterização via k-means

A Clusterização se deu através do critério de avaliação do "elbow method", como especificado no capítulo da metodologia, Trata-se de uma avaliação utilizando recursos gráficos em que é avaliado a evolução da soma total de desvios-padrões, WSS, ao longo do incremento do número de clusters no algoritmo do k-means. As figuras 5.1, 5.2 e 5.3 apresentam o comportamento do WSS da etapas de início, meio e fim, respectivamente.

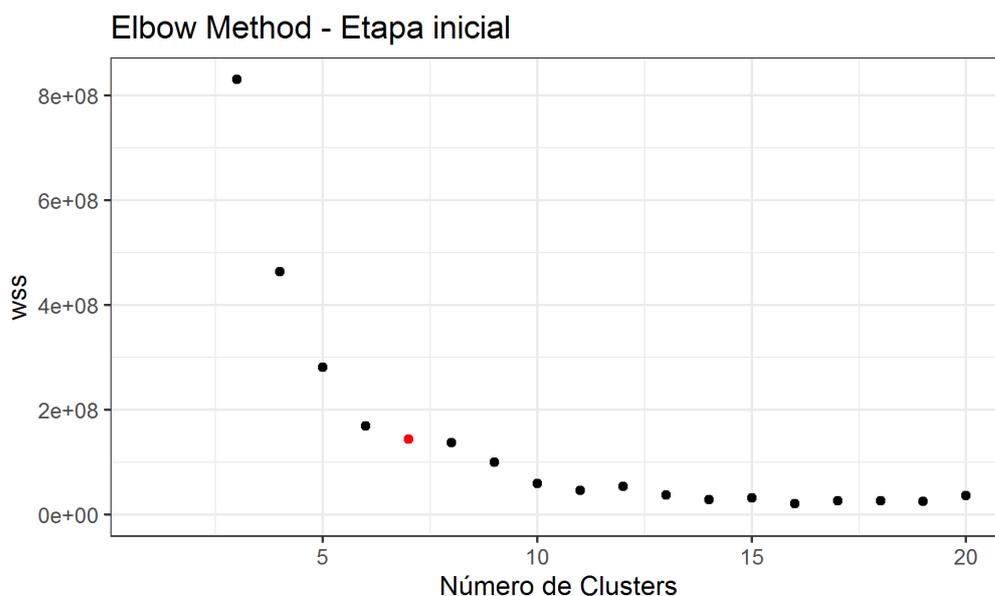


Figura 5.1: WSS dos clusters estimados para a Etapa Inicial.

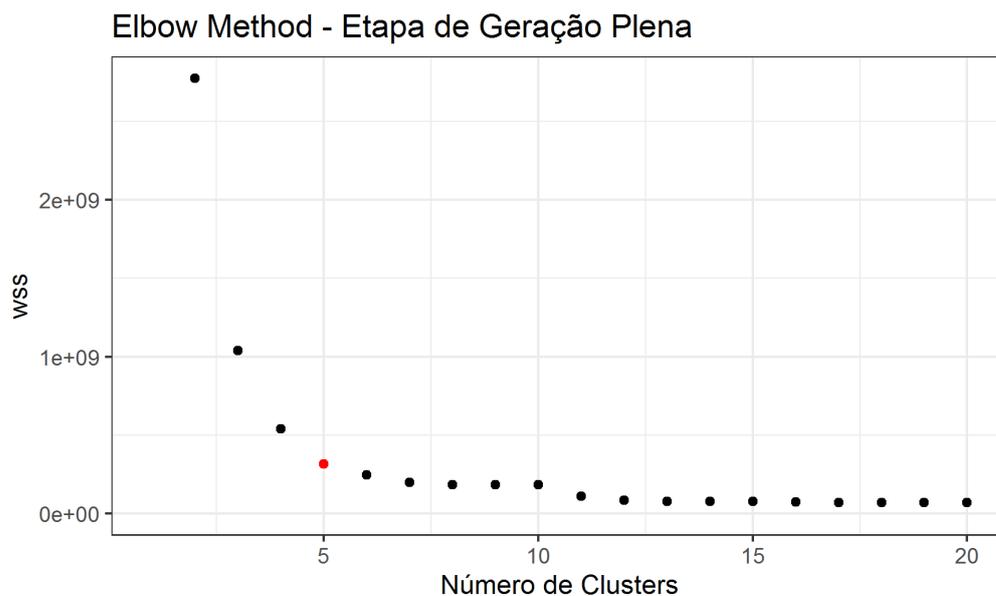


Figura 5.2: WSS dos clusters estimados para a Etapa de Geração Plena.

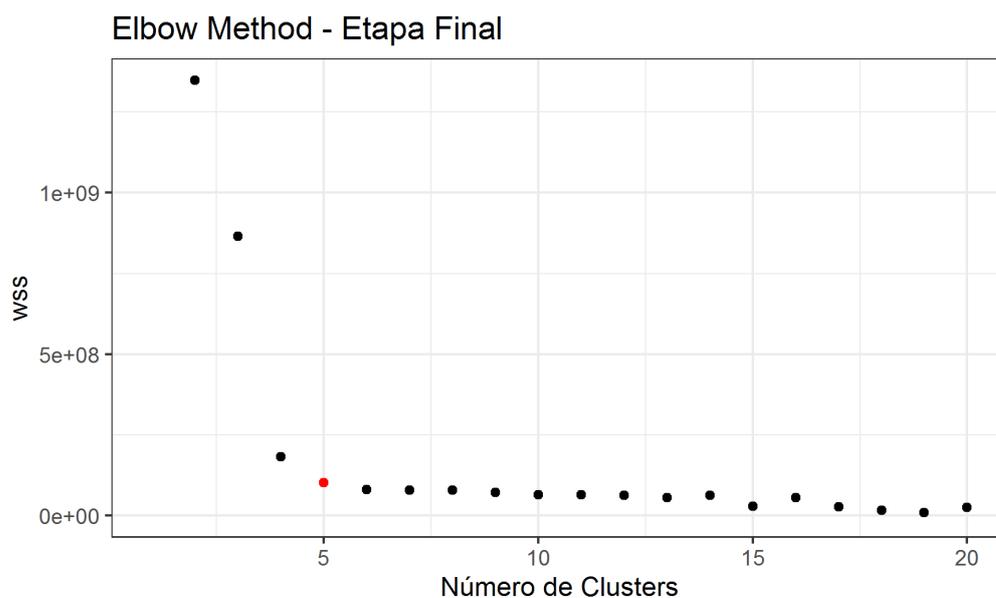


Figura 5.3: WSS dos clusters estimados para a Etapa Final.

É observado que o WSS assume comportamento assintótico conforme o aumento do número de clusters, fazendo a alusão de um cotovelo, "elbow", esta região é onde a assíntota se encontra mais notável. A metodologia consiste em escolher um número de clusters que pertença a esta região e ao mesmo tempo represente o momento de inflexão da assíntota, isto é, quando há decréscimos menos discrepantes no WSS a partir deste ponto. Para que esta condição fosse atendida, foi escolhido o primeiro número de clusters que apresentou a menor variação no WSS com seu sucessor

Como consequência, foram definidos diferentes números de clusters para cada etapa, este resultado ressalta a importância de tratar a série temporal considerando uma sazonalidade através de uma divisão em etapas. Este fenômeno era esperado uma vez que as análises individuais de cada etapa apontaram para diferentes densidades probabilísticas. Em suma, foram estabelecidos 7 clusters para representar o processo de geração durante a etapa inicial, 5 clusters para a etapa de geração plena e mais 5 para a etapa final de geração, a qual precede a intermitência.

5.2 Matrizes de Transição

Após a discretização de valores, são construídas as matrizes de transição seguindo as Equações 3-2 e 3-3. As matrizes de probabilidade de transição das etapas de início, meio e fim estão representadas pelas Tabelas 5.1, 5.2 e 5.3, respectivamente.

M_k	95	899	1504	1974	2430	2707	3037
95	0.846	0.102	0.019	0.012	0.011	0.006	0.001
899	0.082	0.738	0.130	0.026	0.007	0.007	0.007
1504	0.012	0.070	0.775	0.110	0.010	0.007	0.012
1974	0.009	0.018	0.123	0.709	0.092	0.028	0.017
2430	0.010	0.007	0.020	0.132	0.682	0.109	0.036
2707	0.006	0.004	0.009	0.037	0.109	0.720	0.111
3037	0.003	0.006	0.006	0.010	0.015	0.039	0.919

Tabela 5.1: Matriz de Probabilidade de Transição - Início de Geração.

M_k	42	1301	2146	2717	3058
42	0.923	0.039	0.016	0.012	0.007
1301	0.106	0.482	0.204	0.093	0.112
2146	0.039	0.098	0.496	0.222	0.143
2717	0.007	0.017	0.070	0.657	0.246
3058	0.001	0.004	0.010	0.043	0.940

Tabela 5.2: Matriz de Probabilidade de Transição - Geração Plena.

M_k	104	1540	2120	2678	3102
104	0.842	0.046	0.019	0.040	0.051
1540	0.025	0.908	0.027	0.015	0.024
2120	0.012	0.056	0.659	0.160	0.111
2678	0.022	0.016	0.109	0.492	0.359
3102	0.003	0.004	0.013	0.059	0.918

Tabela 5.3: Matriz de Probabilidade de Transição - Término de Geração.

Uma vantagem trazida pelo processo de Clusterização é a interpretação das tabelas de transição devido a número reduzido de estados do processo. Observando os valores da diagonal das matrizes, em todas as etapas há uma grande probabilidade de um estado atual em t permanecer nele mesmo um passo a frente, $t + 1$. Este fenômeno condiz com a realidade operacional de uma geração automatizada mais estável, que atrelada ao gerenciamento de estoque da biomassa.

Outro fator a ser ressaltado é que, mesmo após os valores estarem discretizados, foram estimadas probabilidades de transição entre clusters cujos centróides apresentavam valores muito discrepantes. Estas estimativas estão congruentes as características da série temporal uma vez que a série apresenta breves volatilidade em todas as etapas de geração. Portanto, isto também indica que a Clusterização está adequada a ponto de permitir que as matrizes sejam influenciadas pelo aspecto ruidoso do histórico.

Através da Equação 3-4, foram estimadas as matrizes de probabilidade acumulada para as simulações. As Tabelas 5.4, 5.5 e 5.6 apresentam respectivamente as matrizes de probabilidade acumulada para as etapas de início, meio e fim da geração.

M_k	95	899	1504	1974	2430	2707	3037
95	0,846	0,949	0,968	0,980	0,992	0,998	1
899	0,082	0,820	0,951	0,978	0,985	0,992	1
1504	0,012	0,082	0,858	0,969	0,980	0,987	1
1974	0,009	0,028	0,151	0,860	0,953	0,982	1
2430	0,010	0,018	0,038	0,170	0,853	0,963	1
2707	0,006	0,010	0,020	0,058	0,167	0,888	1
3037	0,003	0,009	0,015	0,026	0,041	0,080	1

Tabela 5.4: Matriz de Probabilidade Acumulada - Início de Geração.

M_k	42	1301	2146	2717	3058
42	0,923	0,963	0,979	0,992	1
1301	0,106	0,589	0,793	0,887	1
2146	0,039	0,137	0,633	0,856	1
2717	0,007	0,024	0,095	0,753	1
3058	0,001	0,005	0,016	0,059	1

Tabela 5.5: Matriz de Probabilidade Acumulada - Geração Plena.

M_k	104	1540	2120	2678	3102
104	0,842	0,889	0,908	0,948	1
1540	0,025	0,933	0,960	0,975	1
2120	0,012	0,068	0,728	0,888	1
2678	0,022	0,039	0,148	0,640	1
3102	0,003	0,008	0,021	0,081	1

Tabela 5.6: Matriz de Probabilidade Acumulada - Término de Geração.

As matrizes de probabilidade acumuladas serão utilizadas para a construção dos cenários da simulação de Monte Carlo de acordo com a Equação 3-6, gerando 100 cenários para cada etapa. Para o início e fim da geração, as simulações foram geradas com horizontes de um mês, respeitando o critério estabelecido na seção anterior. Enquanto, para a geração plena, as simulações se estendem em um horizonte de até três meses. Os horizontes correspondem a separação descrita pela Figura 4.8.

Com as matrizes de transição calculadas, são realizadas as simulações de Monte Carlo de acordo com a Equação 3-6. Foram construídos 100 cenários de geração de energia respeitando o horizonte de tempo estabelecido para cada etapa na seção anterior, sendo elas um mês para o início e término e quatro meses para a geração plena.

5.3

Avaliação das simulações

No entanto, é necessário avaliar se as simulações conseguem replicar as características do histórico de geração. De forma a demonstrar a aderência das simulações, elas devem estar coerentes quanto as estatísticas e a evolução do processo ao longo do tempo. Ao longo da revisão da literatura, não foi

encontrado nenhuma medida estatística coerente ao contexto de geração a biomassa nem que tratasse de uma distribuição probabilística a mesma.

Diante da dificuldade de não haver uma métrica estatística, a aderência das simulações foi avaliada através da comparação de medidas estatísticas básicas como média e desvio padrão. O motivo desta abordagem simples é avaliar se as simulações respeitam o nível de geração usual da usina bem como a dispersão desses valores. Ressalta-se que a reprodução destas estimativas estatísticas é muito impactante ao considerar a metodologia de Clusterização, a probabilidade de transição e o processo de simulação.

A Tabela 5.7 apresenta a comparação das estatísticas para cada etapa de geração entre as simulações e o histórico relativo ao último período de geração, meados de 2020 até início de 2021.

	Média		Desvio Padrão	
	Histórico	Simulação	Histórico	Simulação
Início da Geração	1731	1729	855	845
Geração plena	2651	2643	938	938
Término da Geração	2848	2846	723	708

Tabela 5.7: Comparação das Estatísticas Básicas entre histórico e simulação.

A Tabela 5.7 aponta que tanto as médias quando os desvios padrões de cada etapa estão muito próximos. O resultado se tornar ainda mais relevante por considerar ser uma comparação com as médias do histórico, que são de natureza contínua, com as simulações cujos valores são discretizados. Embora as estatísticas tenham tido um excelente desempenho, é necessário tornar evidente a aderência através de testes hipóteses sobre as estatísticas mensuradas.

Os testes hipóteses avaliam as estatísticas apresentadas na Tabela 5.7 de acordo com a distribuição probabilística. De acordo com o Teorema Central do Limite, TCL, um grande número de amostragens tende a uma distribuição normal. Dessa forma, as estatísticas de teste t e F podem ser aplicados para o teste de diferenças de média e variâncias, respectivamente, entre o histórico e as simulações. A tabela 5.8 apresenta os resultados dos testes hipóteses para as diferenças de média e variâncias para cada etapa, com o valor das estatísticas e o respectivo valor da probabilidade, "p-value".

	Teste t		Teste F	
	p-value	Estatística t	p-value	Estatística F
Início da Geração	0,8466	0,1934	0,4805	1,013
Geração plena	0,2506	1,1489	0,5075	0,9937
Término da Geração	0,8278	0,2175	0,3319	1,018

Tabela 5.8: Teste Hipóteses sobre as diferenças entre Médias, teste-t, e Variâncias, teste-F, dos Histórico e Simulação.

Os testes estatísticas trabalham em cima de duas hipóteses, a hipótese nula e a hipótese alternativa. A hipótese nula para o teste t é a de que não há diferenças entre as médias das populações do histórico e simulações. Em contrapartida, a hipótese alternativa é justamente a de que há uma diferença de médias.

A hipótese nula é avaliada através do "p-value" que, caso seja menor que 0,05, isto é 5%, significa que a hipótese nula foi rejeitada a um intervalo de confiança de 95%. Este evento se torna uma evidência de que as simulações diferem do histórico e, portanto, não há aderência, favorecendo a hipótese alternativa. O mesmo pode ser inferido ao teste F que, por sua vez, avalia a hipótese nula de que não há divergência entre as amostras por conta da variância.

O resultado da Tabela 5.8 indica que em todos os testes hipóteses para cada etapa, a hipótese nula foi favorável, indicando que não há diferenças em média nem em variâncias. Portanto, é evidente de que as simulações foram totalmente aderentes a simulação pelo fato de todos os "p-values" serem maiores que 0,05.

Há também o teste de distribuição de Wilcoxon, que avalia alterações na distribuição probabilística. Infelizmente, este teste só poderia ser aplicado caso as amostras tivessem o mesmo tamanho para serem pareadas ou caso fossem dependentes onde não se pode assumir uma distribuição normal das amostras. Nota-se que o estudo de caso não se encontra em nenhuma das situações citadas, portanto, os histogramas visam suprir de maneira gráfica essa ausência para avaliar se a simulação respeitou a distribuição probabilística de clusters obtidos pelo k-means.

Assim sendo, as Figuras 5.4, 5.5 e 5.6 representam os histogramas comparativos das etapas de início, meio e fim, respectivamente. Vale ressaltar que para esta análise, o histórico também se encontra discretizado a fim de fazer uma comparação coerente.

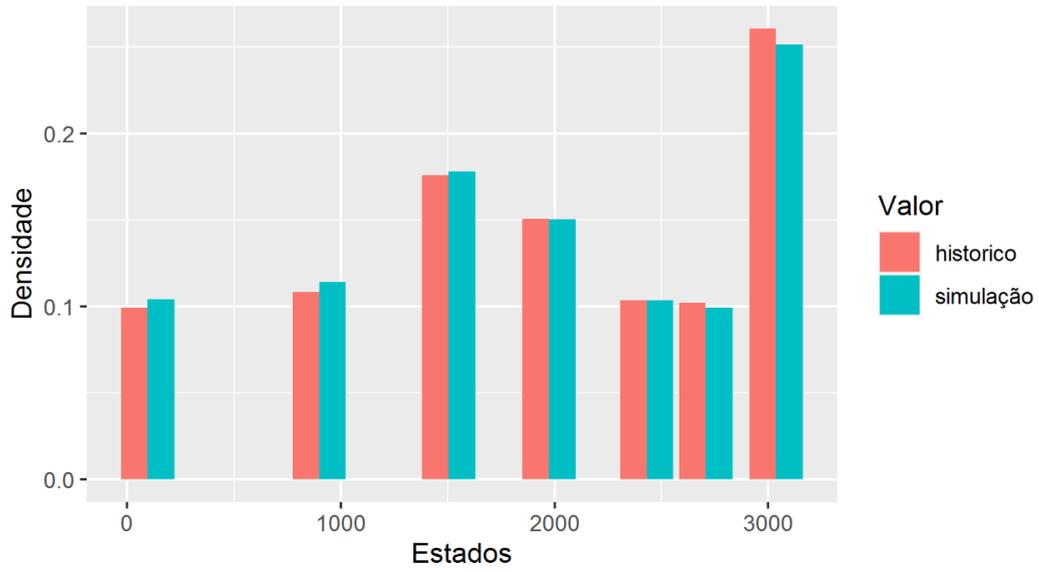


Figura 5.4: Densidade de Clusters - Início de Geração.

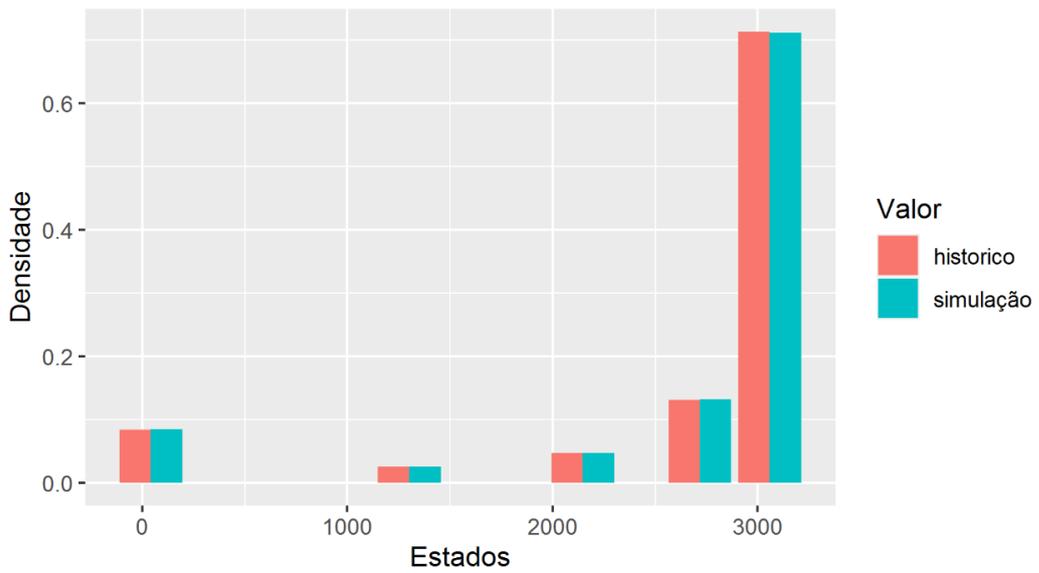


Figura 5.5: Densidade de Clusters - Geração Plena.

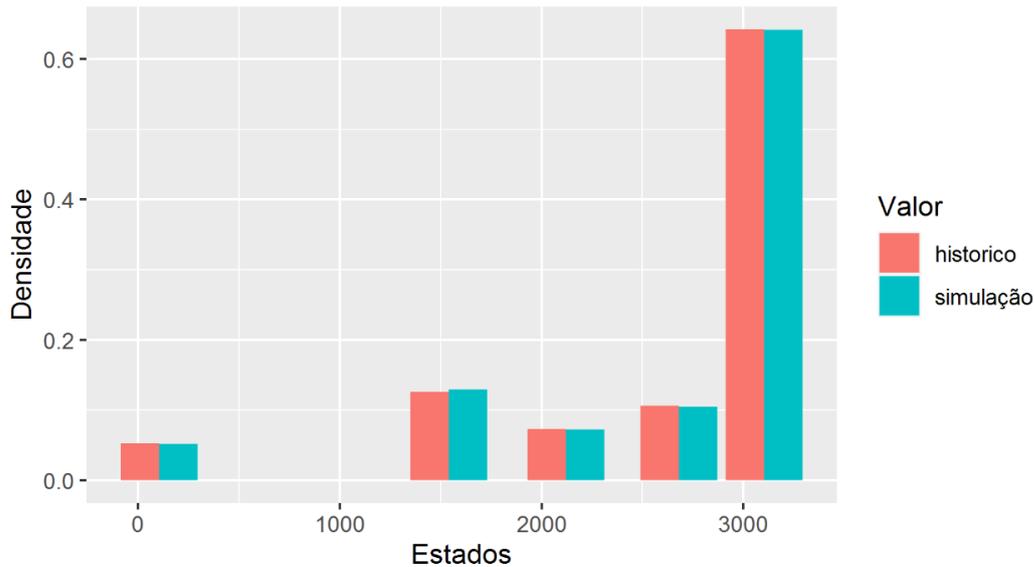


Figura 5.6: Densidade de Clusters - Término de Geração.

Uma análise inicial dos histogramas traz semelhanças com as curvas de densidade obtidas na análise de dados, Figuras 4.5, 4.7 e 4.6. Nota-se que os centróides junto as respectivas probabilidades visam acompanhar as curvas de densidade probabilística. Por conseguinte, alguns dos centróides estão próximos das modas locais das curvas de probabilidade. Este fenômeno pode ser interpretado de forma que a Clusterização acarretou não somente a discretização da série, mas em uma discretização bastante verossímil da distribuição probabilística contínua.

Como consequência disso, nota-se que as diferenças entre as clusterizações e suas respectivas distribuições podem ser explicadas devido as peculiaridades da operação da usina em cada etapa. Na Figura 5.4, o maior número de clusters visa acompanhar a tendência de aumento sutil da geração até que atinja níveis próximos a sua capacidade. Já as Figuras 5.5 e 5.6 apresentam sua geração próxima aos níveis da capacidade instalada na maioria das vezes. Este fato está de acordo com o contexto de uma usina termoeletrica, que tem sua geração controlada submetida a parâmetros de contrato de energia.

De maneira geral, foram notadas ligeiras variações na distribuição de Clusters principalmente na etapa inicial de geração, Figura 5.4. Este fenômeno se deve ao fato da comparação estar sendo feita em cima apenas do último ano de geração, podendo ser afetado por nuances da operação. No entanto, é possível afirmar que os resultados estão coerentes por não haver discrepâncias significativas tanto por parte das medidas estatísticas quanto por parte das distribuições probabilísticas.

Finalmente, de forma a concluir as análises de resultados da simulação, foram levantadas as curvas estatísticas para avaliar a evolução do processo ao

longo do tempo. Ressalta-se que, como visto no capítulo de análise de dados, as estimativas mensais são suficientes para avaliação, pois foi observado que, dentro de um período de geração, a geração não sofre influência intradiária.

A Figura 5.7 apresenta as curvas de médias mensais do histórico e das simulações, bem como as curvas de limite inferior e superior, que correspondem a quantis de 15 e 85% respectivamente. O motivo pelo uso dos quantis de 15 e 85% é buscar uma demonstração mais impactante do que o uso dos quantis de 5 e 95%, Como observados nos gráficos da distribuição dos clusters, 5.4, 5.5 e 5.6, em todas as etapas o cluster com a menor média apresenta uma densidade maior do que 5%, a qual se aproxima de 0, como consequência, o quantil de 5% resulta em uma linha constante próxima de zero, não tendo uma interpretação muito conclusiva.

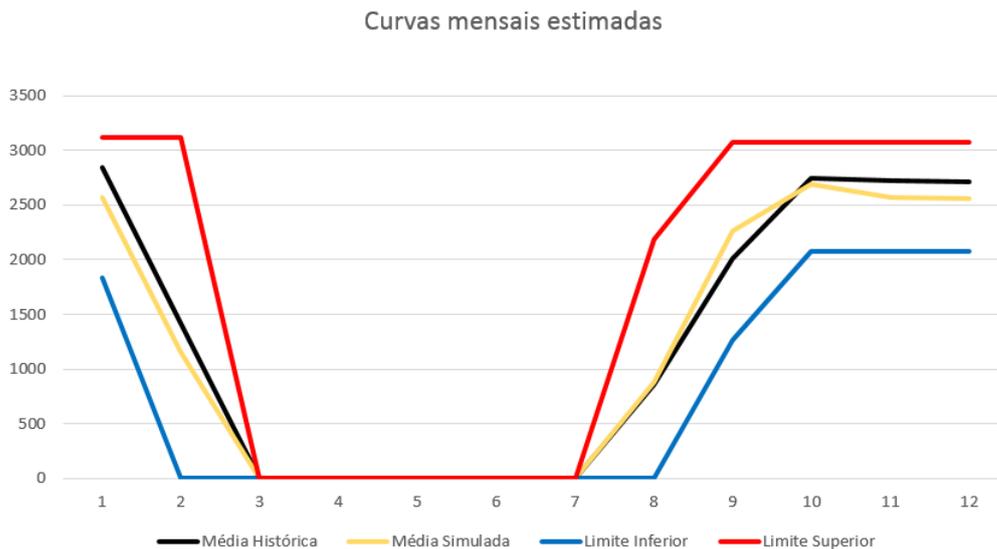


Figura 5.7: Curvas Comparativas de geração de energia a Biomassa.

Em suma, as simulações se mostraram bastante aderentes ao histórico de geração tanto estatisticamente quanto a reprodução das tendências ao longo do ano. A Figura 5.7 complementa a análise, mostrando que os limites compreendem a média histórica enquanto a mesma é acompanhada pela média dos cenários durante todo o período de geração.

Vários fatores estão por trás do excelente desempenho da modelagem, dentre eles, pode-se citar principalmente o processo de Clusterização. Estabelecer um centróide com bastante acurácia foi traduzida posteriormente como um fator estatístico após estimar as matrizes de transição e realizar as simulações. Em segundo lugar, considerando que as simulações obedeceram a probabilidade de transição acumulada, é possível afirmar que a metodologia é consistente em sua aplicação ao replicar fidedignamente a distribuição de clusters do histórico.

Finalmente, deve-se citar a versatilidade da aplicação de Markov Chain. Apesar de ser uma série temporal bastante peculiar, uma sazonalidade bem definida foi suficiente para que o modelo executasse as simulações. No entanto, dentre as limitações do modelo, a existência de um limite de geração por parte do dimensionamento da matriz resulta que o modelo não contempla a expansão, nem redução da capacidade de geração, seja por reajuste de demanda contratada ou por alterações da capacidade física. Ressalta-se que estas limitações não são exclusivas desta modelagem e que esta conseguiu reproduzir todos os atributos da geração a biomassa.

5.4

Resumo do Capítulo

O capítulo apresentou os resultados da metodologia adaptada do MCMC, que, por sua vez, se mostraram completamente coerentes com os atributos da geração a biomassa. Partindo de uma modelagem univariada, sem qualquer informação sobre o potencial de geração a biomassa, os resultados se mostram surpreendentes. Os resultados obtidos de cada etapa da metodologia serão ressaltados neste resumo.

A Clusterização se mostrava teoricamente coerente pelo fato do estado ser representado por um valor de mesma grandeza da geração, kW. Tal processo teve seu desempenho otimizado ao utilizar o "elbow method" para estabelecer o número de clusters, pois esta técnica visa exatamente melhorar a acurácia do centróide em representar um conjunto de valores contínuos. Como resultado, diferentes Clusterização, tanto em números como centróides estimados foram notadas, reforçando a importância de dividir a série em diferentes etapas de geração.

As matrizes de transição estimadas foram definidas como ergódicas, isto é, tendem a uma distribuição probabilística a longo prazo. O fato de não haver a presença de nenhum estado absorvente foi definitivo para seguir a metodologia. Este fato é justificado pela retirada de todo o período intermitente, que se estendeu até meados de agosto e fevereiro, não respeitando uma sazonalidade mensal.

Finalmente, o mais importante resultado foram as simulações, as quais aderiram fortemente as suas respectivas seções do histórico. As simulações apresentaram estatísticas bastante precisas ao considerarem a média e o desvio-padrão. Além disso, as simulações também apresentaram uma forte aderência da série ao longo do período de geração estipulado, acompanhando as tendências de cada etapa. Portanto, pode-se dizer que a geração a biomassa pode ser modelada por uma cadeia de markov.

6

Conclusão

A dissertação teve como objetivo a construção de cenários de geração de energia a biomassa via simulação da metodologia adaptada de MCMC. Os resultados das simulações foram promissores ao avaliar a aderência à base de dados fornecida pela unidade geradora. No entanto, algumas ressalvas sobre a metodologia devem ser notadas para a aplicação prática e estudos futuros.

A princípio, a Clusterização k-means tornou viável a aplicação o conceito de transição de estados de um processo markoviano para uma série temporal cujos valores de abrangem grandes magnitudes, MW. A redução para um número finito de estados através da Clusterização é feita de forma consistente, por se tratar de uma série univariada, o k-means agrupou os valores por nível de geração, o que acarretou em uma discretização da série. A Clusterização foi estabelecida visando reduzir a variância total de cada cluster. Dessa forma, o centróide, que representa o aglomerado de valores em um cluster através de uma média, possui uma alta acurácia.

As matrizes de transição, por sua vez, dependem tanto da Clusterização quanto da janela de tempo atribuída para a estimação das probabilidades. Como notado na seção de análise de dados, a densidade probabilística evolui a cada etapa de geração. Tal fenômeno afetou diretamente a Clusterização obtida pelo k-means, sendo estabelecidos diferentes quantidades de clusters em diferentes níveis de geração. As diferentes Clusterizações resultaram, portanto, em diferentes matrizes de transição que estão condizentes com a natureza operacional da usina.

A geração elétrica cessa por breves momentos durante o período de safras, ocasionando em uma grande volatilidade na geração. Esta volatilidade se encontra ainda mais recorrentes durante a etapa inicial, assim sendo, a divisão da geração em três etapas - início, meio e fim - influenciou positivamente nos resultados desta dissertação ao julgar pelas distribuições de clusters e a aderência das medidas estatísticas das simulações avaliadas pelo teste hipótese. Apesar desta complexidade, as matrizes de transição capturaram o aspecto ruidoso de geração observada ao longo do histórico, mesmo após a Clusterização.

Tratando-se das limitações desta modelagem, é notável que esta se encontra em um estudo de caso muito específico, porém não muito raro, de geração de energia. A primeira limitação intrínseca é o fato de não haver qualquer conciliação quanto as alterações de parâmetros de geração, seja ela a redução

ou o aumento da capacidade instalada. Além disso, diferentes tipos de nível, tendência e sazonalidades em um mesmo período em diferentes anos afetam a estimativa das probabilidades de transição, comprometendo posteriormente, as simulações. Como dito inicialmente, são limitações aparentemente significativas, porém a alteração das capacidade instalada e sazonalidades não são usuais no contexto de geração de energia.

Do ponto de vista técnico, a metodologia aborda a geração de maneira univariada utilizando apenas o próprio histórico para as simulações, sem considerar qualquer variável explicativa, como supostamente um potencial de geração a partir do estoque de cana-de-açúcar. No entanto, o modelo adaptado MCMC se mostrou robusto o suficiente, em paralelo com a divisão da geração em etapas, através dos resultados obtidos das simulações. Estas foram avaliadas de maneira simples, considerando apenas estatísticas básicas, como média e desvio padrão, pelo fato da Clusterização dificultar a aplicação de teste hipóteses sobre a distribuição probabilística. De forma a contornar esta problemática, foi utilizado recurso gráfico para analisar as distribuições de clusters bem como a evolução do processo pelo tempo. Tais resultados, denotam que as simulações replicaram fidedignamente o comportamento da geração a biomassa.

Devido às limitações e dificuldades encontradas no desenvolvimento da metodologia, são sugeridas propostas para o aperfeiçoamento e estudos futuros. A primeira seria o desenvolvimento de teste hipótese para a distribuição probabilística sobre a medidas estatísticas para avaliar a aderência das simulações ao histórico, vale ressaltar que a aderência dos resultados quanto a distribuição probabilística se deu com base em avaliações gráficas. A segunda proposta é sobre a aplicação de variáveis explicativas que estimem o potencial de energia a ser gerado, como o gerenciamento de estoque da biomassa ou simplesmente a produção da mesma. Como forma de promover a metodologia, finalmente sugere-se a comparação com outras metodologias de séries temporais, como inteligência artificial ou modelos mais clássicos como o ARIMA.

Referências bibliográficas

- [1] COMMITION, E.. **Brief on biomass for energy in the european union**. Technical report, 2019. (document), 1, 1.1
- [2] EIA. **Biomass and waste fuels made up 2% of total u.s. electricity generation in 2016**. <https://www.eia.gov/todayinenergy/detail.php?id=33872>, 2017. Acessado: 2020-08-18. (document), 1, 1.2
- [3] MME. **Balanco energético nacional**. Technical report, 2018. (document), 1, 1.3
- [4] LEEMANS, R.; VAN AMSTEL, A.; BATTJES, C.; KREILEMAN, E. ; TOET, S.. **The land cover and carbon cycle consequences of large-scale utilizations of biomass as an energy source**. *Global Environmental Change*, 6(4):335–357, 1996. *Integrated Scenarios of Global Chage: Results from the IMAGE 2 Model*. 1
- [5] LI, C.; HAYES, D. J. ; JACOBS, K. L.. **Biomass for bioenergy: Optimal collection mechanisms and pricing when feedstock supply does not equal availability**. *Energy Economics*, 76:403–410, 2018. 1
- [6] DORNBURG, V.; VAN VUUREN, D.; VAN DE VEN, G.; LANGEVELD, H.; MEEUSEN, M.; BANSE, M.; VAN OORSCHOT, M.; ROS, J.; VAN DEN BORN, G. J.; AIKING, H. ; OTHERS. **Bioenergy revisited: key factors in global potentials of bioenergy**. *Energy & Environmental Science*, 3(3):258–267, 2010. 1
- [7] STYLES, D.; THORNE, F. ; JONES, M. B.. **Energy crops in ireland: an economic comparison of willow and miscanthus production with conventional farming systems**. *Biomass and bioenergy*, 32(5):407–421, 2008. 1
- [8] CALDAS, M. M.; BERGTOLD, J. S.; PETERSON, J. M.; GRAVES, R. W.; EARNHART, D.; GONG, S.; LAUER, B. ; BROWN, J. C.. **Factors affecting farmers' willingness to grow alternative biofuel feedstocks across kansas**. *Biomass and Bioenergy*, 66:223–231, 2014. 1
- [9] EUROSTAT. **Renewable energy in the eu in 2018**. Technical report, 2020. Acessado: 2020-09-2. 1

- [10] NATIONS, U.. **Agenda 2030**. https://ec.europa.eu/sustainable-development/goal17_en, 2015. Acessado: 2020-05-20. 1
- [11] 25X'25 NATIONAL STEERING COMMITTEE. **25x'25 action plan: Charting america's energy future**. Technical report, 2004. 1
- [12] ON THE CLIMATE CRISIS, S. D. S. C.. **The case for climate action building a clean economy for the american people**. <https://www.democrats.senate.gov/climate>, 2015. Acessado: 2020-08-18. 1
- [13] RABE, M.; STREIMIKIENE, D. ; BILAN, Y.. **Eu carbon emissions market development and its impact on penetration of renewables in the power sector**. *Energies*, 12(15):2961, 2019. 1
- [14] MAZZUCATO, M.; SEMIENIUK, G.. **Financing renewable energy: Who is financing what and why it matters**. *Technological Forecasting and Social Change*, 127:8–22, 2018. 1
- [15] DESFIANDI, A.. **Building an energy consumption model and sustainable economic growth in emerging countries**. 2019. 1
- [16] PRESIDÊNCIA DA REPÚBLICA. **Lei ordinária nº 13.576/2017.**, 2017. <http://www.planalto.gov.br>. 1
- [17] WELFLE, A.. **Balancing growing global bioenergy resource demands-brazil's biomass potential and the availability of resource for trade**. *Biomass and Bioenergy*, 105:83–95, 2017. 1
- [18] CCEE. **Preço sombra**. <https://www.ccee.org.br>, 2019. 1
- [19] MME. **Portaria nº 301, de 31 de julho de 2019**. Technical report, 2019. Acessado: 2020-05-20. 1
- [20] SIKARWAR, V. S.; ZHAO, M.; CLOUGH, P.; YAO, J.; ZHONG, X.; MEMON, M. Z.; SHAH, N.; ANTHONY, E. J. ; FENNELL, P. S.. **An overview of advances in biomass gasification**. *Energy & Environmental Science*, 9(10):2939–2977, 2016. 2
- [21] LEI, M.; SHIYAN, L.; CHUANWEN, J.; HONGLING, L. ; YAN, Z.. **A review on the forecasting of wind speed and generated power**. *Renewable and Sustainable Energy Reviews*, 13(4):915–920, 2009. 2

- [22] SOBRI, S.; KOOHI-KAMALI, S. ; RAHIM, N. A.. **Solar photovoltaic generation forecasting methods: A review**. *Energy Conversion and Management*, 156:459–497, 2018. 2
- [23] AVCIOĞLU, A. O.; TÜRKER, U.. **Status and potential of biogas energy from animal wastes in turkey**. *Renewable and Sustainable Energy Reviews*, 16(3):1557–1561, 2012. 2
- [24] JIANG, D.; ZHUANG, D.; FU, J.; HUANG, Y. ; WEN, K.. **Bioenergy potential from crop residues in china: Availability and distribution**. *Renewable and Sustainable Energy Reviews*, 16(3):1377–1382, 2012. 2
- [25] BERINGER, T.; LUCHT, W. ; SCHAPHOFF, S.. **Bioenergy production potential of global biomass plantations under environmental and agricultural constraints**. *Gcb Bioenergy*, 3(4):299–312, 2011. 2
- [26] THRÄN, D.; SCHALDACH, R.; MILLINGER, M.; WOLF, V.; ARENDT, O.; PONITKA, J.; GÄRTNER, S.; RETTENMAIER, N.; HENNENBERG, K. ; SCHÜNGEL, J.. **The milestones modeling framework: An integrated analysis of national bioenergy strategies and their global environmental impacts**. *Environmental Modelling Software*, 86:14–29, 2016. 2
- [27] DE VRIES, B. J.; VAN VUUREN, D. P. ; HOOGWIJK, M. M.. **Renewable energy sources: Their global potential for the first-half of the 21st century at a global level: An integrated approach**. *Energy policy*, 35(4):2590–2610, 2007. 2
- [28] WELFLE, A.; GILBERT, P. ; THORNLEY, P.. **Securing a bioenergy future without imports**. *Energy Policy*, 68:1–14, 2014. 2
- [29] DENG, Y. Y.; KOPER, M.; HAIGH, M. ; DORNBURG, V.. **Country-level assessment of long-term global bioenergy potential**. *Biomass and Bioenergy*, 74:253–267, 2015. 2
- [30] CHEN, X.. **Economic potential of biomass supply from crop residues in china**. *Applied Energy*, 166:141–149, 2016. 2
- [31] KAMJOO, A.; MAHERI, A.; DIZQAH, A. M. ; PUTRUS, G. A.. **Multi-objective design under uncertainties of hybrid renewable energy system using nsga-ii and chance constrained programming**. *International Journal of Electrical Power Energy Systems*, 74:187–194, 2016. 2

- [32] SANAJAOBA SINGH, S.; FERNANDEZ, E.. Modeling, size optimization and sensitivity analysis of a remote hybrid renewable energy system. *Energy*, 143:719–731, 2018. 2
- [33] AVCIOĞLU, A.; DAYIOĞLU, M. ; TÜRKER, U.. Assessment of the energy potential of agricultural biomass residues in turkey. *Renewable Energy*, 138, 08 2019. 2
- [34] ELKADEEM, M.; WANG, S.; SHARSHIR, S. W. ; ATIA, E. G.. Feasibility analysis and techno-economic design of grid-isolated hybrid renewable energy system for electrification of agriculture and irrigation area: A case study in dongola, sudan. *Energy Conversion and Management*, 196:1453–1478, 2019. 2
- [35] WATANABE, M. D.; MORAIS, E. R.; CARDOSO, T. F.; CHAGAS, M. F.; JUNQUEIRA, T. L.; CARVALHO, D. J. ; BONOMI, A.. Process simulation of renewable electricity from sugarcane straw: Techno-economic assessment of retrofit scenarios in brazil. *Journal of Cleaner Production*, 254:120081, 2020. 2
- [36] NIXON, J.; DEY, P. ; DAVIES, P.. The feasibility of hybrid solar-biomass power plants in india. *Energy*, 46(1):541–554, 2012. 2
- [37] FURLAN, F. F.; TONON FILHO, R.; PINTO, F. H.; COSTA, C. B.; CRUZ, A. J.; GIORDANO, R. L. ; GIORDANO, R. C.. Bioelectricity versus bioethanol from sugarcane bagasse: is it worth being flexible? *Biotechnology for Biofuels*, 6(1):1–12, 2013. 2
- [38] ALSALEH, M.; ABDUL-RAHIM, A. ; MOHD-SHAHWAHID, H.. An empirical and forecasting analysis of the bioenergy market in the eu28 region: Evidence from a panel data simultaneous equation model. *Renewable and Sustainable Energy Reviews*, 80:1123–1137, 2017. 2
- [39] ZHANG, F.; LI, C.; YU, Y. ; JOHNSON, D. M.. Resources and future availability of agricultural biomass for energy use in beijing. *Energies*, 12(10):1828, 2019. 2
- [40] FLORES HERNANDEZ U., JAEGER, D.; SAMPERIO, J.. Modeling forest woody biomass availability for energy use based on short-term forecasting scenarios. *Waste Biomass Valor*, 11(5):2137–2151, 2020. 2

- [41] HUANG, Y.; YANG, L.; GAO, C.; JIANG, Y. ; DONG, Y.. A novel prediction approach for short-term renewable energy consumption in china based on improved gaussian process regression. *Energies*, 12(21):4181, 2019. 2
- [42] MA, J.; OPPONG, A.; ACHEAMPONG, K. ; ABRUQUAH, L.. Forecasting renewable energy consumption under zero assumptions. *Sustainability*, 10:576, 02 2018. 2
- [43] SENOCAK, A. A.; GUNER GOREN, H.. Forecasting the biomass-based energy potential using artificial intelligence and geographic information systems: A case study. *Engineering Science and Technology, an International Journal*, 2021. 2
- [44] HUNTINGTON, T.; CUI, X.; MISHRA, U. ; SCOWN, C. D.. Machine learning to predict biomass sorghum yields under future climate scenarios. *Biofuels, Bioproducts and Biorefining*, 14(3):566–577, 2020. 2
- [45] DE, S.; KAIADI, M.; FAST, M. ; ASSADI, M.. Development of an artificial neural network model for the steam process of a coal biomass cofired combined heat and power (chp) plant in sweden. *Energy*, 32(11):2099–2109, 2007. 2
- [46] FAST, M.; PALME, T.. Application of artificial neural networks to the condition monitoring and diagnosis of a combined heat and power plant. *Energy*, 35(2):1114–1120, 2010. 2
- [47] NIKPEY, H.; ASSADI, M.; BREUHAUS, P. ; MØRKVED, P.. Experimental evaluation and ann modeling of a recuperative micro gas turbine burning mixtures of natural gas and biogas. *Applied Energy*, 117:30–41, 2014. 2
- [48] BALDINELLI, A.; BARELLI, L.; BIDINI, G.; BONUCCI, F. ; ISKENDE-ROĞLU, F. C.. Regarding solid oxide fuel cells simulation through artificial intelligence: A neural networks application. *Applied Sciences*, 9(1):51, 2019. 2
- [49] DJATKOV, D.; EFFENBERGER, M. ; MARTINOV, M.. Method for assessing and improving the efficiency of agricultural biogas plants based on fuzzy logic and expert systems. *Applied energy*, 134:163–175, 2014. 2

- [50] ANGELINE, A. A.. Power generation from combusted “syngas” using hybrid thermoelectric generator and forecasting the performance with ann technique. *Journal of Thermal Engineering*, 4(4):2149–2168, 2018. 2
- [51] LAWAL, A. I.; ALADEJARE, A. E.; ONIFADE, M.; BADA, S. ; IDRIS, M. A.. Predictions of elemental composition of coal and biomass from their proximate analyses using anfis, ann and mlr. *International Journal of Coal Science & Technology*, 8(1):124–140, 2021. 2
- [52] HIPPERT, H.; PEDREIRA, C. ; SOUZA, R.. Neural networks for short-term load forecasting: a review and evaluation. *IEEE Transactions on Power Systems*, 16(1):44–55, 2001. 2
- [53] TRAPERO, J. R.; KOURENTZES, N. ; MARTIN, A.. Short-term solar irradiation forecasting based on dynamic harmonic regression. *Energy*, 84:289–295, 2015. 2
- [54] WANG, G.; SU, Y. ; SHU, L.. One-day-ahead daily power forecasting of photovoltaic systems based on partial functional linear regression models. *Renewable Energy*, 96:469–478, 2016. 2
- [55] LI, Y.; HE, Y.; SU, Y. ; SHU, L.. Forecasting the daily power output of a grid-connected photovoltaic system based on multivariate adaptive regression splines. *Applied Energy*, 180:392–401, 2016. 2
- [56] MASSIDDA, L.; MARROCU, M.. Use of multilinear adaptive regression splines and numerical weather prediction to forecast the power output of a pv plant in borkum, germany. *Solar Energy*, 146:141–149, 2017. 2
- [57] GUOYANG, W.; YANG, X. ; SHASHA, W.. Discussion about short-term forecast of wind speed on wind farm. *Jilin Electric Power*, 181(5):21–24, 2005. 2
- [58] DING, M.; ZHANG, L.-J. ; WU, Y.-C.. Wind speed forecast model for wind farms based on time series analysis. *Electric Power Automation Equipment*, 25(8):32–34, 2005. 2
- [59] TORRES, J. L.; GARCIA, A.; DE BLAS, M. ; DE FRANCISCO, A.. Forecast of hourly average wind speed with arma models in navarre (spain). *Solar energy*, 79(1):65–77, 2005. 2

- [60] COSTA, A.; CRESPO, A.; NAVARRO, J.; LIZCANO, G.; MADSEN, H. ; FEITOSA, E.. **A review on the young history of the wind power short-term prediction.** *Renewable and Sustainable Energy Reviews*, 12(6):1725–1744, 2008. 2
- [61] PHINIKARIDES, A.; MAKRIDES, G.; KINDYNI, N.; KYPRIANOU, A. ; GEORGHIOU, G. E.. **Arima modeling of the performance of different photovoltaic technologies.** In: 2013 IEEE 39TH PHOTOVOLTAIC SPECIALISTS CONFERENCE (PVSC), p. 0797–0801. IEEE, 2013. 2
- [62] YANG, D.; JIRUTITIJAROEN, P. ; WALSH, W. M.. **Hourly solar irradiance time series forecasting using cloud cover index.** *Solar Energy*, 86(12):3531–3543, 2012. 2
- [63] YANG, C.; THATTE, A. A. ; XIE, L.. **Multitime-scale data-driven spatio-temporal forecast of photovoltaic generation.** *IEEE Transactions on Sustainable Energy*, 6(1):104–112, 2014. 2
- [64] AGOUA, X. G.; GIRARD, R. ; KARINIOTAKIS, G.. **Short-term spatio-temporal forecasting of photovoltaic power production.** *IEEE Transactions on Sustainable Energy*, 9(2):538–546, 2017. 2
- [65] SHUANG, H.; YONGQIAN, L. ; YONGPING, Y.. **Taboo search algorithm based ann model for wind speed prediction.** In: 2007 2ND IEEE CONFERENCE ON INDUSTRIAL ELECTRONICS AND APPLICATIONS, p. 2599–2602. IEEE, 2007. 2
- [66] FLORES, P.; TAPIA, A. ; TAPIA, G.. **Application of a control algorithm for wind speed prediction and active power generation.** *Renewable Energy*, 30(4):523–536, 2005. 2
- [67] MABEL, M. C.; FERNANDEZ, E.. **Analysis of wind power generation and prediction using ann: A case study.** *Renewable energy*, 33(5):986–992, 2008. 2
- [68] XIAO, Y.-S.; WANG, W.-Q. ; HUO, X.-P.. **Study on the time-series wind speed forecasting of the wind farm based on neural networks.** *Energy Conservation Technology*, 25(2):106–108, 2007. 2
- [69] IZGI, E.; ÖZTOPAL, A.; YERLI, B.; KAYMAK, M. K. ; ŞAHIN, A. D.. **Short–mid-term solar power prediction by using artificial neural networks.** *Solar Energy*, 86(2):725–733, 2012. 2

- [70] NOTTON, G.; PAOLI, C.; IVANOVA, L.; VASILEVA, S. ; NIVET, M. L.. **Neural network approach to estimate 10-min solar global irradiation values on tilted planes.** *Renewable Energy*, 50:576–584, 2013. 2
- [71] AMROUCHE, B.; LE PIVERT, X.. **Artificial neural network based daily local forecasting for global solar radiation.** *Applied energy*, 130:333–341, 2014. 2
- [72] HUANG, J.-C.; HUANG, W.-T.; CHU, P.-T.; LEE, W.-Y.; PAI, H.-P.; CHUANG, C.-C. ; WU, Y.-W.. **Applying a markov chain for the stock pricing of a novel forecasting model.** *Communications in Statistics-theory and Methods*, 46(9):4388–4402, 2017. 2
- [73] DAI, Y.; HAN, D. ; DAI, W.. **Modeling and computing of stock index forecasting based on neural network and markov chain.** *The Scientific World Journal*, 2014, 2014. 2
- [74] GANESAN, K.; ANNAMALAI, U. ; DEIVANAYAGAMPILLAI, N.. **An integrated new threshold fcms markov chain based forecasting model for analyzing the power of stock trading trend.** *Financial Innovation*, 5(1):1–19, 2019. 2
- [75] TSAUR, R.-C.; OTHERS. **A fuzzy time series-markov chain model with an application to forecast the exchange rate between the taiwan and us dollar.** *International journal of innovative computing, information and control*, 8(7):4931–4942, 2012. 2
- [76] CASTINO, F.; FESTA, R. ; RATTO, C.. **Stochastic modelling of wind velocities time series.** *Journal of Wind Engineering and Industrial Aerodynamics*, 74-76:141–151, 1998. 2
- [77] CARPINONE, A.; LANGELLA, R.; TESTA, A. ; GIORGIO, M.. **Very short-term probabilistic wind power forecasting based on markov chain models.** In: 2010 IEEE 11TH INTERNATIONAL CONFERENCE ON PROBABILISTIC METHODS APPLIED TO POWER SYSTEMS, p. 107–112. IEEE, 2010. 2
- [78] YODER, M.; HERING, A. S.; NAVIDI, W. C. ; LARSON, K.. **Short-term forecasting of categorical changes in wind power with markov chain models.** *Wind energy*, 17(9):1425–1439, 2014. 2

- [79] CARPINONE, A.; GIORGIO, M.; LANGELLA, R. ; TESTA, A.. **Markov chain modeling for very-short-term wind power forecasting.** *Electric power systems research*, 122:152–158, 2015. 2
- [80] BHARDWAJ, S.; SHARMA, V.; SRIVASTAVA, S.; SASTRY, O.; BANDYOPADHYAY, B.; CHANDEL, S. ; GUPTA, J.. **Estimation of solar radiation using a combination of hidden markov model and generalized fuzzy model.** *Solar Energy*, 93:43–54, 2013. 2
- [81] LI, Y.-Z.; NIU, J.-C.. **Forecast of power generation for grid-connected photovoltaic system based on markov chain.** In: 2009 ASIA-PACIFIC POWER AND ENERGY ENGINEERING CONFERENCE, p. 1–4, 2009. 2
- [82] SANJARI, M. J.; GOOI, H.. **Probabilistic forecast of pv power generation based on higher order markov chain.** *IEEE Transactions on Power Systems*, 32(4):2942–2952, 2016. 2
- [83] MING, D.; NINGZHOU, X.. **A method to forecast short-term output power of photovoltaic generation system based on markov chain.** *Power System Technology*, 35(1):152–157, 2011. 2
- [84] MATIS, J.; BIRKETT, T. ; BOUDREAUX, D.. **An application of the markov chain approach to forecasting cotton yields from surveys.** *Agricultural Systems*, 29(4):357–370, 1989. 2
- [85] MATIS, J.; SAITO, T.; GRANT, W.; IWIG, W. ; RITCHIE, J.. **A markov chain approach to crop yield forecasting.** *Agricultural systems*, 18(3):171–187, 1985. 2
- [86] MALEKI, A.; KHAJEH, M. G. ; AMERI, M.. **Optimal sizing of a grid independent hybrid renewable energy system incorporating resource uncertainty, and load uncertainty.** *International Journal of Electrical Power Energy Systems*, 83:514–524, 2016. 2
- [87] ALHARBI, H.; BHATTACHARYA, K.. **Stochastic optimal planning of battery energy storage systems for isolated microgrids.** *IEEE Transactions on Sustainable Energy*, 9(1):211–227, 2018. 2
- [88] SAMADI GAZIJAHANI, F.; SALEHI, J.. **Optimal bilevel model for stochastic risk-based planning of microgrids under uncertainty.** *IEEE Transactions on Industrial Informatics*, 14(7):3054–3064, 2018. 2

- [89] EHSAN, A.; YANG, Q.. **Scenario-based investment planning of isolated multi-energy microgrids considering electricity, heating and cooling demand.** *Applied Energy*, 235:1277–1288, 2019. 2
- [90] PAPAETHYMIU, G.; KLOCKL, B.. **Mcmc for wind power simulation.** *IEEE Transactions on Energy Conversion*, 23(1):234–240, 2008. 2
- [91] ALMUTAIRI, A.; HASSAN AHMED, M. ; SALAMA, M.. **Use of mcmc to incorporate a wind power model for the evaluation of generating capacity adequacy.** *Electric Power Systems Research*, 133:63–70, 2016. 2, 3.3
- [92] YANG, X.; MA, X.; KANG, N. ; MAIHEMUTI, M.. **Probability interval prediction of wind power based on kde method with rough sets and weighted markov chain.** *IEEE Access*, 6:51556–51565, 2018. 2
- [93] MEDINA MAÇAIRA, P.; MONTEIRO CYRILLO, Y.; CYRINO OLIVEIRA, F. L. ; CASTRO SOUZA, R.. **Including wind power generation in brazil’s long-term optimization model for energy planning.** *Energies*, 12(5), 2019. 2, 3.3, 3.5
- [94] JONES, D.; LORENZ, M.. **An application of a markov chain noise model to wind generator simulation.** *Mathematics and Computers in Simulation*, 28(5):391–402, 1986. 3
- [95] WILINSKI, A.. **Time series modeling and forecasting based on a markov chain with changing transition matrices.** *Expert Systems with Applications*, 133:163–172, 2019. 3.1
- [96] LLOYD, S.. **Least squares quantization in pcm.** *IEEE Transactions on Information Theory*, 28(2):129–137, 1982. 3.2
- [97] FORGY, E.. **Cluster analysis of multivariate data: efficiency versus interpretability of classifications.** *Biometrics*, 21:768–780, 1965. 3.2
- [98] MACQUEEN, J. B.. **Some methods for classification and analysis of multivariate observations.** In: Cam, L. M. L.; Neyman, J., editors, *PROC. OF THE FIFTH BERKELEY SYMPOSIUM ON MATHEMATICAL STATISTICS AND PROBABILITY*, volumen 1, p. 281–297. University of California Press, 1967. 3.2

- [99] HARTIGAN, J. A.; WONG, M. A.. **Algorithm as 136: A k-means clustering algorithm.** Journal of the Royal Statistical Society. Series C (Applied Statistics), 28(1):100–108, 1979. 3.2
- [100] KETCHEN, D. J.; SHOOK, C. L.. **The application of cluster analysis in strategic management research: An analysis and critique.** Strategic Management Journal, 17(6):441–458, 1996. 3.2
- [101] GAGNIUC, P. A.. **Markov Chains: From Theory to Implementation and Experimentation.** USA, NJ: John Wiley Sons, 2017. 3.3
- [102] ATHREYA, K. B.; DOSS, H. ; SETHURAMAN, J.. **On the convergence of the Markov chain simulation method.** The Annals of Statistics, 24(1):69 – 100, 1996. 3.4
- [103] WALLER, E. N. K.; ADABLAH, P. D. ; KESTER, Q.-A.. **Markov chain: forecasting economic variables.** In: 2019 INTERNATIONAL CONFERENCE ON COMPUTING, COMPUTATIONAL MODELLING AND APPLICATIONS (ICCA), p. 115–1154. IEEE, 2019. 3.4
- [104] CHIB, S.; GREENBERG, E.. **Markov chain monte carlo simulation methods in econometrics.** Econometric Theory, 12(3):409–431, 1996. 3.4
- [105] BUCKLEW, J. A.; NEY, P. ; SADOWSKY, J. S.. **Monte carlo simulation and large deviations theory for uniformly recurrent markov chains.** Journal of Applied Probability, 27(1):44–59, 1990. 3.4
- [106] ALDOUS, D.. **On the markov chain simulation method for uniform combinatorial distributions and simulated annealing.** Probability in the Engineering and Informational Sciences, 1(1):33–46, 1987. 3.4
- [107] AU, S. K.. **Probabilistic failure analysis by importance sampling markov chain simulation.** Journal of Engineering Mechanics, 130(3):303–311, 2004. 3.4
- [108] VAN RAVENZWAAIJ, DON; CASSEY, P. B. S. D.. **A simple introduction to markov chain monte-carlo sampling.** Psychonomic Bulletin Review, 25:143–154, 2016. 3.4