



Tuany Esthefany Barcellos de Carvalho Silva

**Modelagem da relação de dependência entre
as variáveis de velocidade do vento e a geração
de energia eólica: Uma aplicação da teoria de
cópulas**

Dissertação de Mestrado

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

Orientador : Prof. Reinaldo Castro Souza
Co-orientador: Prof. Fernando Luiz Cyrino Oliveira

Rio de Janeiro
setembro de 2022



Tuany Esthefany Barcellos de Carvalho Silva

**Modelagem da relação de dependência entre
as variáveis de velocidade do vento e a geração
de energia eólica: Uma aplicação da teoria de
cópulas**

Dissertação apresentada como requisito parcial para 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

Prof. Fernando Luiz Cyrino Oliveira

Co-orientador

Departamento de Engenharia Industrial - PUC-Rio

Prof. Guilherme Armando de Almeida Pereira

UFES

Prof. Marco Aurélio dos Santos Sanfins

UFF

Rio de Janeiro, 01 de setembro de 2022

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

Tuany Esthefany Barcellos de Carvalho Silva

Tuany Esthefany Barcellos de Carvalho Silva graduou-se em Estatística no ano de 2019 pela Universidade Federal Fluminense (UFF). Aluna de mestrado da PUC-Rio desde de agosto de 2020.

Ficha Catalográfica

Silva, Tuany Esthefany Barcellos de Carvalho

Modelagem da relação de dependência entre as variáveis de velocidade do vento e a geração de energia eólica: Uma aplicação da teoria de cópulas / Tuany Esthefany Barcellos de Carvalho Silva; orientador: Reinaldo Castro Souza; co-orientador: Fernando Luiz Cyrino Oliveira. – 2022.

107 f: il. color. ; 30 cm

Dissertação (mestrado) - Pontifícia Universidade Católica do Rio de Janeiro, Departamento de Engenharia Industrial, 2022.

Inclui bibliografia

1. Engenharia de Produção – Teses. 2. Teoria de Cópulas. 3. Simulação. 4. Energia eólica. 5. Estrutura de dependência. I. Castro Souza, Reinaldo. II. Cyrino Oliveira, Luiz Fernando. III. Pontifícia Universidade Católica do Rio de Janeiro. Departamento de Engenharia Industrial. IV. Título.

CDD: 004

Agradecimentos

Consagre ao senhor tudo que você faz e seus planos serão bem-sucedidos (Provérbio 16:3)

Início o meu agradecimento, agradecendo primeiramente a Deus, por me amparar, fortalecer e capacitar. A ele toda glória e minha gratidão.

Aos meus pais, Ediléa e Ismael por todo amor e dedicação incondicional. Obrigada por me cuidar, e se dedicarem tanto para que eu pudesse realizar meus sonhos, essa conquista jamais seria possível sem vocês ao meu lado.

Aos meus avós, Cléia e Valdelém por todo carinho, auxílio e principalmente por toda oração para que eu alcançasse meus objetivos e conseguisse realizar minhas tarefas.

As minhas primas Paôlla e Simone, por me acompanharem nesta caminhada e por vibrarem comigo a cada conquista.

Agradeço ao meu orientador e amigo professor Reinaldo Castro, por todo carinho, ensinamento, oportunidade e dedicação, para que eu realizasse este, trabalho. Ao meu coorientador Fernando Cyrino, por todo auxílio acadêmico, confiança e oportunidade.

Ao meu querido professor Marco Aurélio Sanfins por dedicar seu tempo a me ensinar teoria de cópulas e acreditar que eu conseguiria realizar este trabalho. Você foi fundamental nesta caminhada, obrigada pela dedicação, confiança e oportunidade.

Ao professor Guilherme Almeida, membro da banca, por aceitar o convite e dedicar seu tempo a fazer contribuições de alto valor.

A professora Paula Maçaíra por todo apoio acadêmico e por me auxiliar no desafio de entender os dados de energia eólica.

A minha querida e brilhante amiga e professora Daiane Rodrigues, por me acompanhar nessa trajetória, dividindo comigo seus conhecimentos. Obrigada por toda parceria e por todos os momentos, você é um exemplo pra mim!

A minhas amigas da PUC Rio Daiane Oliveira, Fernanda Nucci e Ana Beatriz, pelos conhecimentos divididos, trabalhos realizados e companheirismo durante esse período.

A minha amiga Karla Karolina, por me amparar e acalmar todas as vezes que eu precisei, me ajudando em todas as apresentações.

Ao meu querido amigo Tiago, por toda paciência em me ajudar quando eu não encontrava o erro, e por sempre me incentivar a pensar “fora da caixa”.

Aos professores e funcionários do DEI da PUC Rio.

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

Hoje sinto-me privilegiada, a vocês os meus mais sinceros agradecimentos!

Resumo

Silva, Tuany Esthefany Barcellos de Carvalho; Castro Souza, Reinaldo; Cyrino Oliveira, Luiz Fernando. **Modelagem da relação de dependência entre as variáveis de velocidade do vento e a geração de energia eólica: Uma aplicação da teoria de cópulas**. Rio de Janeiro, 2022. 107p. Dissertação de Mestrado – Departamento de Engenharia Industrial, Pontifícia Universidade Católica do Rio de Janeiro.

A preocupação com o aquecimento global e a poluição tem aumentado significativamente o interesse no desenvolvimento de fontes renováveis de energia. Este estudo tem como eixo principal a energia eólica, o uso dessa energia elimina resíduos indesejados e prejudiciais à saúde e ao meio ambiente causados por outras fontes de energia, como carvão e usinas nucleares. Este trabalho objetiva analisar a relação de dependência entre a velocidade do vento e a geração de energia eólica, esta é uma relação bastante complexa, por isso busca-se entender a natureza estocástica de ambas as variáveis. Como ferramenta metodológica foi utilizada a teoria da cópula. O estudo baseia-se na análise e modelagem da dependência entre dados de velocidade do vento e geração de energia eólica, para um banco de dados horário de um parque eólico do estado da Bahia, no período de janeiro a dezembro de 2017, após encontrar a cópula correspondente a estrutura de dependência para o ano completo e para cada mês individualmente, foram geradas simulações e apresentadas as probabilidades de ocorrência dos cenários em intervalos pré-definidos, os resultados obtidos foram significativos, testes estatísticos adequados foram realizados, evidenciando a qualidade do ajuste.

Palavras-chave

Teoria de Cópulas; Simulação; Energia eólica; Estrutura de dependência.

Abstract

Silva, Tuany Esthefany Barcellos de Carvalho; Castro Souza, Reinaldo (Advisor); Cyrino Oliveira, Luiz Fernando (Co-Advisor). **Modeling the dependency relationship between the wind speed variables and the generation of wind energy: An application of the theory of copulations**. Rio de Janeiro, 2022. 107p. Dissertação de Mestrado – Departamento de Engenharia Industrial, Pontifícia Universidade Católica do Rio de Janeiro.

Concern about global warming and pollution has significantly increased interest in developing renewable energy sources. This study has wind energy as its main axis, the use of this energy eliminates unwanted and harmful waste to health and the environment caused by other energy sources, such as coal and nuclear power plants. This work aims to analyze the dependence relationship between wind speed and wind energy generation, this is a very complex relationship, so we seek to understand the stochastic nature of both variables. As a methodological tool, the copula theory was used. The study is based on the analysis and modeling of the dependence between wind speed data and wind energy generation, for an hourly database of a wind farm in the state of Bahia, from January to December 2017, after finding the copula corresponding to the dependency structure for the entire year and for each month individually, simulations were generated and the probabilities of occurrence of the scenarios were presented at pre-defined intervals, the results obtained were significant, adequate statistical tests were performed, evidencing the quality of the fit .

Keywords

Copula theory; Simulation; Windenergy; Dependencystructure.

Sumário

1	Introdução	11
1.1	Motivação	13
1.2	Objetivo do trabalho	15
1.3	Organização da dissertação	15
2	Revisão da Literatura	17
3	Energia Eólica	21
3.1	Vento e a conversão da energia eólica	25
3.2	Capacidade eólica mundial	27
3.3	Potência Extraída do Vento	30
3.4	Distribuição Weibull	32
4	Metodologia	34
4.1	Cópulas	34
4.2	Medidas de dependência e concordância	41
4.3	Principais Famílias de Cópulas	46
4.4	Métodos de estimação	54
4.5	Métricas de desempenho	55
4.6	Testes	57
4.7	Estimador kernel da função de densidade	59
5	Análise e Resultados	61
5.1	Análises e resultados para o período anual	61
5.2	Análise e resultados para o período mensal	68
5.3	Probabilidade de Ocorrência dos Intervalos Simulados	89
6	Conclusão e trabalhos futuros	96
7	Referências bibliográficas	101

Lista de figuras

Figura 3.1	Matriz elétrica brasileira	22
Figura 3.2	Evolução da Capacidade Instalada em MW	25
Figura 3.3	Evolução da nova Capacidade Instalada mundial em GW	28
Figura 3.4	Capacidade instalada acumulada total mundial em GW	29
Figura 5.1	Geração de energia eólica em KW	62
Figura 5.2	Velocidade do vento em m/s	62
Figura 5.3	Densidade das variáveis	64
Figura 5.4	Densidade da Cópula	66
Figura 5.5	Simulação via cópulas (Potência vs Velocidade)	67
Figura 5.6	Comparação das Curvas de Nível Cópula Frank e Cópula Empírica (Jan à Jun)	68
Figura 5.7	Gráfico das séries de Potência e Velocidade do vento (Jan à Jun)	70
Figura 5.8	Histograma e correlação entre as variáveis (Jan à Jun)	72
Figura 5.9	Densidade da Cópula Frank (Jan à Jun)	75
Figura 5.10	Simulação via cópulas (Jan à Jun)	76
Figura 5.11	Comparação das Curvas de Nível Cópula Frank e Cópula Empírica (Jan à Jun)	77
Figura 5.12	Gráfico das séries de Potência e Velocidade do vento (Jul à Set)	78
Figura 5.13	Histograma e correlação entre as variáveis (Jul à Set)	79
Figura 5.14	Densidade da Cópula Frank (Jul à Set)	81
Figura 5.15	Simulação via cópulas (Jul à Set)	82
Figura 5.16	Comparação das Curvas de Nível Cópula Frank e Cópula Empírica (Jul à Set)	83
Figura 5.17	Gráfico das séries de Potência e Velocidade do vento (Out à Dez)	84
Figura 5.18	Histograma e correlação entre as variáveis (Out à Dez)	85
Figura 5.19	Densidade da Cópula Frank (Out à Dez)	87
Figura 5.20	Simulação via cópulas (Out à Dez)	88
Figura 5.21	Comparação das Curvas de Nível Cópula Frank e Cópula Empírica (Jan à Jun)	89
Figura 6.1	Tabela do cálculo das probabilidades - Ano completo - Outros intervalos	100

Lista de tabelas

Tabela 3.1	Capacidade instalada e Número de Parques por estado	24
Tabela 3.2	Capacidade instalada e Número de Parques por estado	26
Tabela 3.3	Capacidade instalada dos países em GW	30
Tabela 5.1	Estatística descritiva	63
Tabela 5.2	Estatística descritiva - Velocidade do vento (Jan à Jun)	71
Tabela 5.3	Estatística descritiva - Potência (Jan à Jun)	71
Tabela 5.4	Resultado do p-valor referente a distribuição Weibull (Jan à Jun)	73
Tabela 5.5	Parâmetro e Tau de Kendal da Cópula Frank (Jan à Jun)	75
Tabela 5.6	Estatística descritiva - Velocidade do vento (Jul à Set)	78
Tabela 5.7	Estatística descritiva - Potência (Jul à Set)	79
Tabela 5.8	Resultado do p-valor referente a distribuição Weibull (Jul à Set)	80
Tabela 5.9	Parâmetro e Tau de Kendal da Cópula Joe (Jul à Set)	81
Tabela 5.10	Estatística descritiva - Velocidade do vento (Out à Dez)	84
Tabela 5.11	Estatística descritiva - Potência (Out à Dez)	84
Tabela 5.12	Resultado do p-valor referente a distribuição Weibull (Out à Dez)	86
Tabela 5.13	Parâmetro e Tau de Kendal da Cópula Frank (Out à Dez)	87
Tabela 5.14	Probabilidade de ocorrência dos intervalos (Ano completo)	90
Tabela 5.15	Probabilidade de ocorrência dos intervalos (Janeiro)	91
Tabela 5.16	Probabilidade de ocorrência dos intervalos (Fevereiro)	91
Tabela 5.17	Probabilidade de ocorrência dos intervalos (Março)	92
Tabela 5.18	Probabilidade de ocorrência dos intervalos (Abril)	92
Tabela 5.19	Probabilidade de ocorrência dos intervalos (Maio)	93
Tabela 5.20	Probabilidade de ocorrência dos intervalos (Junho)	93
Tabela 5.21	Probabilidade de ocorrência dos intervalos (Julho)	93
Tabela 5.22	Probabilidade de ocorrência dos intervalos (Agosto)	94
Tabela 5.23	Probabilidade de ocorrência dos intervalos (Setembro)	94
Tabela 5.24	Probabilidade de ocorrência dos intervalos (Outubro)	94
Tabela 5.25	Probabilidade de ocorrência dos intervalos (Novembro)	95
Tabela 5.26	Probabilidade de ocorrência dos intervalos (Dezembro)	95

1

Introdução

Incentivados por uma demanda crescente de energia, visando a preservação do meio ambiente, a busca por alternativas limpas de obtenção de energia vem se intensificando ao decorrer dos anos. Algumas fontes renováveis, surgiram no passado como uma alternativa útil aos recursos escassos de combustíveis fósseis, uma dessas fontes renováveis é a energia eólica, ou seja, aquela derivada da velocidade do vento. Segundo Duca (2020), houve um aumento significativo de estudos voltados para tipos de energias renováveis nos anos 70, muitos deles focados na geração de energia eólica. Com os avanços tecnológicos e investimentos governamentais, seu potencial tem sido explorado em várias regiões, originando inúmeras pesquisas abordando o tema. Por meio deste método de geração, a força do vento é transformada em energia elétrica através de aerogeradores, que são turbinas com hélices dispostas em zonas onde há grandes deslocamentos de ar. Para que a produção de energia eólica seja eficiente e economicamente viável, é necessário concentrar os aerogeradores em um mesmo ponto estratégico, sendo este um agrupamento conhecido como parque eólico.

Segundo a Agência Internacional de energia (EIA, 2018), o consumo de energia cresce de forma intensiva e acelerada, chegando a aproximadamente 524 quadrilhões de unidades térmicas britânicas (Btu) nos últimos anos, sendo previsto um aumento de 56% até 2040. Com isso, surgem inúmeras questões como o esgotamento de fontes tradicionais de energia, por exemplo a hidráulica e os combustíveis fósseis. De acordo com Aguilar (2015), existem preocupações com o meio ambiente, e com a segurança dos sistemas de suprimento de energia, o que causa um crescimento significativo na busca pela geração de energia eólica. Estudo realizado pela Associação Brasileira de Energia Eólica (ABEEólica, 2020), apresenta que o Brasil possui uma matriz energética com

um vasto potencial hidrelétrico, e inúmeras fontes de energia renováveis a serem exploradas. As hidrelétricas representam cerca de 58% da produção de energia do país, seguida pela geração de energia eólica com aproximadamente 10,3%. Segundo Duca (2020), essa fonte de energia é altamente capaz de abastecer o Sistema Interligado Nacional - (SIN), com cerca 14% de sua produção registrada durante o período conhecido como “colheita do vento”, que ocorre entre junho e novembro, essa época do ano é caracterizada por ventos mais fortes e constantes.

De acordo com Conselho Global de Energia Eólica (GWEC, 2021), o Brasil é um país altamente promissor para geração de energia eólica, pesquisas abordando o tema foram apresentadas nos últimos anos, a fim de contribuir e acompanhar o crescimento da geração desta fonte renovável de energia. Como esta, trata-se de uma geração dependente de fatores naturais, sendo o principal a velocidade do vento, são inúmeros os elementos que afetam diretamente a geração dos recursos eólicos. Componentes como a região, terreno e temperatura, impactam significativamente a geração desta energia. Distintos estudos foram realizados, a fim de explorar o potencial eólico para diversas regiões do Brasil e do mundo. Porém, existe uma limitação em relação a informações necessárias sobre as variáveis de interesse, muitas vezes essas informações são confidenciais e com acesso estritamente limitado, dificultando estudos e análises das mesmas.

A relação de dependência entre a velocidade do vento, e a geração de energia eólica vem sendo foco de pesquisas realizadas por diversos autores como, (KARAKAŞ, 2019) e (BAHRAOUI, 2018), esses estudos apontam a metodologia de Cópulas, como algo inovador e eficiente para a análise de dependências das variáveis aqui mencionadas. Neste estudo, objetiva-se utilizar a teoria de cópula para lidar com a estrutura de dependência, além da relação entre a variável natural velocidade do vento, e a geração de energia eólica. A função cópula foi introduzida no conceito de análise multivariada por Sklar

(1959), porém só no ano de 1986 os autores Genest e MacKay implementaram esta metodologia, realizando uma aplicação na área financeira, explorando as propriedades da cópula arquimediana (GENEST MACKAY, 1986).

Por conseguinte, as aplicações foram estendidas a diversos campos de pesquisa. A metodologia passou a ser utilizada para análises de variáveis naturais. Uma das propriedades mais relevantes da teoria de cópulas, é que a mesma possibilita separar a função de dependência das funções de distribuição marginais, sendo usado este princípio para análise das variáveis observadas neste trabalho, buscando estimar uma cópula adequada para a velocidade do vento e a geração de energia eólica. Dentro desse contexto, este trabalho traz contribuições para a área de conhecimento de energia renovável e sustentabilidade, os resultados permitem a observação de simulações em intervalos distintos, por meio da metodologia implementada, onde é apresentada a estrutura de dependência entre as variáveis aleatórias abordadas.

1.1 Motivação

A preocupação com o aquecimento global e a poluição, têm aumentado o interesse no desenvolvimento de fontes renováveis e menos agressivas ao meio ambiente, como eólica, solar, hidrelétrica, geotérmica, hidrogênio, e a biomassa como substitutos dos combustíveis fósseis. A energia eólica pode fornecer soluções adequadas para os problemas supramencionados. A utilização desta energia elimina resíduos não desejados e prejudiciais à saúde e ao meio ambiente, provenientes das outras fontes energéticas como o carvão e usinas nucleares (resíduos radioativos). Adicionalmente, ao benefício aqui citado, a diversificação da matriz energética proveniente da inserção ou até mesmo o aumento da participação dessa fonte de energia na matriz existente, reduz a dependência de combustíveis fósseis, que estão sujeitos à instabilidade de preços e fornecimento.

Conforme descrito no Relatório (*Global Wind* de 2021 da GWEC (*Global*

Wind Energy Council), a pandemia da COVID-19 trouxe um maior reconhecimento das consequências do desenvolvimento humano, e dos efeitos em cascata que um evento pode exercer na economia mundial, nos meios de subsistência e segurança. Conforme os formuladores de políticas traçam a saída da pandemia, há um acordo sem precedentes que a mudança climática é uma verdadeira emergência global. O conceito de uma ameaça mundial agora não é apenas credível, mas factível. De acordo com o relatório, isso fez com que o ONU reforçasse o apelo para uma ação urgente, a fim de chegar a zero as emissões de gases de efeito estufa (GEE) em 2050 - uma chamada que desde então foi ecoada por mais de 120 países, que representam aproximadamente mais da metade de PIB global. Reduzir as emissões de CO_2 relacionadas com a produção de energia, de acordo com o relatório Irena (2019), *Future of wind: Deployment, investment, technology, grid integration and socio-economic aspects*, está no centro da transformação energética mundial. O mundo está mudando rapidamente, diminuindo o consumo de combustíveis fósseis que causam drásticas mudanças climáticas, adotando cada vez mais energias limpas e renováveis. De acordo com o relatório, novas formas de geração de energia são a chave para que o mundo alcance os objetivos climáticos acordados até o presente momento.

Isto posto, este trabalho adotará a simulação em intervalos distintos, de geração de energia eólica e velocidade do vento, por meio da teoria de cópulas, sendo este o eixo principal de desenvolvimento desta pesquisa. Visto que, torna-se cada vez mais necessário análises sobre a estrutura de dependência entre variáveis naturais, como velocidade do vento e geração de energia eólica, dado que a sustentabilidade, e o foco em fontes de energias renováveis trata-se de um tema de relevância mundial.

A análise da estrutura de dependência da geração de energia eólica e da sua fonte natural, traz significativos benefícios operacionais, sociais e econômicos, permitindo um gerenciamento eficaz da energia eólica. O emprego da teoria de copulas na modelagem das variáveis aqui abordadas, tem a capaci-

dade de capturar a complexa relação multivariada não linear, com base em suas distribuições marginais univariadas; através do uso da cópula encontrada, será possível realizar uma simulação eficiente em inúmeros intervalos, contribuindo significativamente para o cálculo das probabilidades de ocorrência dos mesmos, auxiliando na redução da dependência de geração de energia a partir de combustíveis fósseis, podendo assim minimizar as emissões de CO_2 .

1.2

Objetivo do trabalho

O presente trabalho objetiva implementar a teoria de cópulas, para que através desta modelagem seja possível identificar a estrutura de dependência entre as variáveis geração de energia eólica e velocidade do vento, para o ano completo de 2017 com frequência horária e em seguida para cada mês de forma individual, também com frequência horária. Buscando assim, verificar se a família da cópula que representa a estrutura de dependência é a mesma para o ano completo e todos os meses abordados na pesquisa. Após evidenciar a família de cópulas correspondente busca-se simular cenários e apresentar a probabilidade de ocorrência dos mesmos, em intervalos pré-definidos.

1.3

Organização da dissertação

Este trabalho está organizado em 6 capítulos. O capítulo 1 apresenta a introdução, onde é contextualizado de modo geral o tema que pretende-se abordar ao longo do texto, assim como a motivação e o objetivo da pesquisa. O capítulo 2, traz uma revisão da literatura feita na base de dados Scopus com palavras específicas como: “*copula*” e “*wind energy*”, neste capítulo busca-se quantificar e analisar o volume de produções científicas que tratam o tema cópula e geração de energia eólica de maneira conjunta. No capítulo 3 aborda-se a Energia eólica no Brasil e no mundo, são apresentadas estatísticas sobre esta fonte renovável de energia. No capítulo 4 contém toda metodologia utilizada para a análise e obtenção dos resultados deste estudo, nele encontra-se seções

abordando a teoria de cópulas de maneira ampla, métodos de estimação e desempenho, testes estatísticos e o método de estimação de densidade por Kernel. O capítulo 5, traz as análises e os resultados após a implementação da teoria de cópulas. No capítulo 6 encontra-se a conclusão da pesquisa.

2

Revisão da Literatura

Segundo Lima (2018), o conceito de sustentabilidade, tornou-se amplamente difundido no ano de 1987 após o Relatório *Brundtland*, o mesmo define o desenvolvimento sustentável como o “desenvolvimento que satisfaz as necessidades presentes, sem comprometer a capacidade das gerações futuras de suprir suas próprias necessidades”. Desde então a preocupação com o aquecimento global, poluição e geração de energia, tornaram-se pauta para estudos e pesquisas, onde busca-se ampliar o desenvolvimento de fontes renováveis e menos agressivas ao meio ambiente. Ao contrário das fontes de combustível fóssil, as fontes de energia renováveis são incontestavelmente mais limpas, sendo assim uma ótima alternativa para manter a geração de energia. Entre estas fontes renováveis está a potência do vento, que após capturada possibilita a conversão de ar e energia cinética em eletricidade por meio de turbinas.

Buscando quantificar e analisar o volume de produções científicas que abordam o tema cópula e geração de energia eólica, iniciou-se em dezembro de 2021 uma pesquisa na base de dados Scopus, os termos utilizados na busca foram “copula” e “wind energy”. De acordo com a pesquisa na plataforma, os termos supracitados, foram descritos de forma conjunta em um trabalho acadêmico, pela primeira vez em 2010, e até 02 de dezembro de 2021 um total de 59 trabalhos relevantes para esta pesquisa foram publicados abordando tema.

A partir do século XX diversas metodologias foram implementadas, para melhor compreender a geração de energias sustentáveis, assim como a relação de dependência entre a velocidade do vento e a geração de energia eólica. Após pesquisa bibliográfica, nota-se um aumento considerável nos estudos de energia renovável na década de 1970, devido a crescente preocupação das partes governamentais com a limitação das fontes de combustível fóssil e o impacto da

queima desses combustíveis no ambiente global. Diante desse cenário altamente pessimista, o foco na energia eólica cresceu devido à resultados alcançados em países como Dinamarca e Alemanha.

Em seus estudos Schindler et. al. (2021), examinaram o potencial do uso complementar dos recursos eólicos disponíveis em 33 países europeus de 1971 a 2010, identificando as escalas de tempo que explicam a maior parte da variação nas séries de rendimento diário de energia eólica nos países supramencionados. Os autores ajustaram cópulas as séries dos rendimentos anuais de energia eólica, (*AEY - annual wind energy yield (TWh/ano)*), a componente sazonal da *EY* (*SEY - seasonal component of EY*) e a componente diária (*DEY - daily component of EY (TWh/ano)*) do rendimento diário de energia eólica. Os autores utilizaram a implementação da teoria de cópulas para encontrar os limiares do parâmetro de correlação linear (ρ), para distribuições bivariadas do rendimento energético nacional dos países europeus em escalas anual (AEY), sazonal (SEY) e diária (DEY). Este resultado em conjunto com outras métricas aplicadas, auxiliaram os autores na obtenção de resultados relevantes no que tange o potencial de uso complementar dos recursos eólicos nacionais disponíveis nestes países.

Mi et. al (2021), apresentaram em seu estudo um método de avaliação de acomodação de energia renovável, considerando a correlação de energia eólica e solar. Para avaliar a relação entre as energias renováveis, os autores estabeleceram modelos estocásticos com base na função cópula. No artigo em questão, a teoria de cópula é aplicada para estabelecer uma função de densidade de probabilidade conjunta que considera a correlação espacial da energia eólica e solar. Com base nesta função de densidade, o método de Monte Carlo é aplicado para capturar a sequência do vento, e produção de energia solar considerando a correlação com a energia eólica. Os resultados da simulação demonstram que o modelo proposto pelos autores no artigo tem um bom desempenho, sendo está uma aplicação relevante da teoria de cópulas no

setor de energia.

O estudo recente dos autores Ma et. al. (2021), apresenta uma programação estocástica ajustável de risco para o dia seguinte, (RSDS) (*risk adjustable stochastic day-ahead scheduling RSDS*) para equilibrar os requisitos de risco dos sistemas de energia (*power system participant*) PSPs, e propõem uma abordagem de Amostragem de Hipercubo - Sobol-*augmented Latin Hypercube Sampling* (SaLHS), para melhorar a eficiência do processo de geração de cenários no RSDS. Os autores apresentaram a abordagem SaLHS combinada com cópulas D-vine, para gerar cenários de erro de previsão da energia eólica *wind power-WF* para RSDS, considerando correlações de parques eólicos.

No caso de vários parques eólicos, de acordo com Ma et. al. (2021), os erros de previsão da velocidade do vento de parques distintos são dependentes. Para os pesquisadores, uma maneira correta de abordar as correlações dos erros de previsões de inúmeros parques eólicos é obter a distribuição conjunta, e a teoria da copula fornece uma maneira conveniente de modelar distribuição conjunta multivariada. A utilização da abordagem SaLHS combinada com cópulas D-vine, foi uma das contribuições mais relevantes de pesquisas abordando o tema para a literatura atual.

O autor Grothe (2011), apresenta uma aplicação interessante, no qual investigou-se as dependências estocásticas da velocidade do vento para um grande conjunto de dados de estações meteorológicas *on e offshore* alemãs. O estudo apontou que essas dependências não tornam-se extremamente não lineares, mas constantes ao longo do tempo, o uso da teoria da cópula possibilitou reproduzir planos de alocação ideal para parques eólicos, descobrindo assim que a alocação otimizada de parques eólicos pode equilibrar substancialmente o fornecimento total de energia eólica na frequência diária, e também horária de uma localidade.

Haghi (2010), afirma que a crescente influência da geração renovável em sistemas de energia requer uma modelagem por meio da teoria de cópulas, onde

exista a integração de turbinas fotovoltaicas e eólicas, difundidas em uma rede de distribuição. O mesmo propõe investigar as dependências relacionadas a variáveis de energias renováveis, com base na modelagem estocástica usando uma cópula, para este estudo o autor optou pela implementação da cópula Arquimediana, obtendo resultados satisfatórios para sua pesquisa.

De acordo com Cherubini e Luciano (2001), umas das principais vantagens da implementação da metodologia de cópulas, é a maneira com que a mesma representa uma distribuição conjunta de probabilidade. Autores como Ozawa (2017) e Santos (2019), indicam em seus estudos a distribuição Weibull como uma distribuição eficiente para modelar a natureza estocástica da velocidade do vento. Dado que para encontrar a cópula que explique a relação de dependência entre a velocidade do vento e a geração de energia eólica, é necessário evidenciar a distribuição de probabilidade marginal das variáveis supramencionadas, é de extrema importância estudos como estes, que apontam uma distribuição eficaz, auxiliando significativamente na modelagem.

Duca (2020), afirma que a geração de energia eólica mesmo com suas inúmeras vantagens como fonte renovável, também apresenta deficiências. Uma limitação importante a ser ressaltada está relacionada à natureza estocástica da energia eólica, e a grande incerteza na previsão do fornecimento desta energia. Outro fator é a incapacidade de armazenar o vento para uma geração futura, o que torna o planejamento muito dependente de previsões confiáveis de curto prazo. Deste modo, torna-se evidente a necessidade da implementação de metodologias que permitam, analisar de forma eficiente a estrutura de dependência das variáveis relacionadas a geração de energia, possibilitando fazer simulações mais condizentes com a realidade.

3

Energia Eólica

Com o passar do tempo o consumo de energia tornou-se extremamente necessário, fundamental e indispensável na vida da população, este trata-se de um dos insumos essenciais para o desenvolvimento socioeconômico de um país (SATHYAJITH, 2014). Estudos apontam que a demanda de energia no mundo cresce de forma significativamente acelerada. A Agência Internacional de energia (EIA, 2018), apresenta previsão de um aumento de aproximadamente 56% no consumo até 2040.

A energia eólica é aquela obtida a partir da força do vento, a mesma é produzida pela junção da energia cinética com o aquecimento eletromagnético do sol, fazendo com que as pás de captadores se movimentem e gerem energia útil. A utilização da energia eólica apresenta inúmeras vantagens quando comparada com o modo de produção de energia tradicional. O benefício mais notável é que está trata-se de uma energia com fonte inesgotável, pois a mesma depende apenas de recurso natural, o vento. Outra vantagem a ser mencionada é a não emissão de gases poluentes, evitando contribuições negativas para o meio ambiente, amenizando os impactos que geram o aquecimento global (SOUZA, 2013). A energia eólica categoriza-se dentro do grupo de fontes de energias renováveis, representando uma das alternativas mais promissoras e favoráveis, sua natureza limpa e barata faz com que ela se torne extremamente atraente ao atendimento da demanda energética global, que atualmente encontra-se em um movimento crescente significativo (AGUILAR, 2015).

Segundo Gomes (2017), mesmo a utilização da energia renovável apresentando inúmeros benefícios, socioeconômicos e ambientais, infelizmente o uso da energia decorrente de combustíveis fósseis ainda é altamente predominante na matriz energética mundial. Porém instabilidades climáticas, e a alta volatilidade no preço do petróleo atraem o interesse governamental para as fontes de

energias renováveis (FER), fazendo com que os mesmos invistam na implementação dessas novas fontes em suas matrizes energéticas, possibilitando assim a redução das emissões de carbono, minimizando a dependência econômica relacionada ao petróleo (FAGIANI, 2013).

Uma matriz elétrica é formada pelo conjunto de fontes disponíveis apenas para a geração de energia elétrica em um país, estado ou no mundo. A geração de energia elétrica mundial é baseada, principalmente, em combustíveis fósseis como carvão, óleo e gás natural, em termelétricas. No entanto no Brasil, as fontes de energia mais usadas são as renováveis (AGUILAR, 2016). Segundo relatório EPE (2021) o país usa 83% de fontes renováveis para a produção de energia, enquanto a média mundial é de 27%. Além dos inúmeros benefícios sustentáveis, isso também gera menos custos de operação para as usinas que escolhem a energia renovável. De acordo com o relatório anual disponibilizado pela Associação Brasileira de Energia Eólica (ABEEOLICA, 2021), pode-se observar que a produção de energia eólica representou cerca de 11,8% da produção de energia do país, ocupando o segundo lugar de maior geração entre as fontes renováveis.

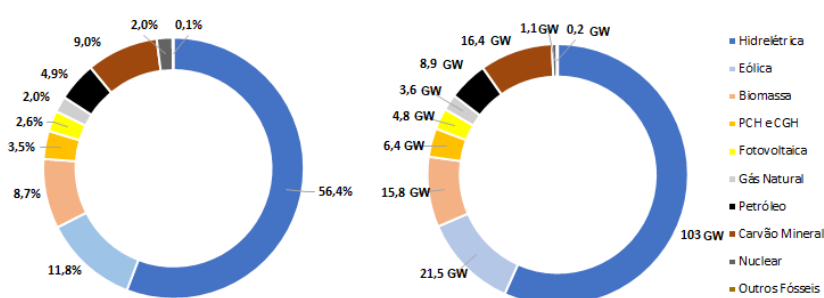


Figura 3.1: Matriz elétrica brasileira

Fonte: Adaptado da Associação Brasileira de Energia Eólica, 2021.

São notórios e inúmeros os benefícios proporcionados pela produção de energias limpas e sustentáveis. Segundo o boletim anual da ABEEólica (2021), o aumento de investimentos para geração de energia eólica impacta positivamente a economia e o meio ambiente, gerando renda e melhoria de

vida para proprietários de terra com arrendamento para colocação das torres, outro fator muito importante, é que tratando-se de uma energia renovável, a mesma não polui, colaborando para que o Brasil cumpra seus objetivos no Acordo do Clima. Segundo ABEEólica (2021) a instalação de novos parques eólicos contribui significativamente para o aumento do Produto Interno Bruto (PIB) e do Índice de Desenvolvimento Humano do Município (IDHM). O estudo apresenta um comparativo entre um grupo de municípios de distintos estados brasileiros que receberam parques eólicos com outros que ainda não foram contemplados, o mesmo identificou que nos municípios que receberam a instalação o PIB real aumentou cerca de 21,15% no período de 1999 a 2017 e o IDHM cresceu em torno de 20% entre 2000 e 2010. Na tabela 3.1 é possível verificar a capacidade instalada e o número de Parques por estado, nota-se que o Rio Grande do Norte e a Bahia possuem juntos 53,38% dos Parques eólicos do país.

Tabela 3.1: Capacidade instalada e Número de Parques por estado

UF	Potência (MW)	Parques
RN	5.154,2	182
BA	4.879,6	189
PI	2.275,9	79
CE	2.179,3	84
RS	1.835,9	80
PE	798,4	34
MA	426,0	15
SC	238,5	14
PB	157,2	15
SE	34,5	1
RJ	28,1	1
PR	2,5	1
Total	18.010,1	695

Fonte: Adaptado do boletim ABEEólica (2021)

O Brasil é considerado um país altamente promissor na geração de energia eólica, segundo o Conselho Global de Energia Eólica (GWEC, 2021), o país está em 6º lugar no Ranking mundial de capacidade instalada de energia eólica. A pesquisa aponta que 80% dos parques eólicos brasileiros estão situados no Nordeste, esta região possui um dos melhores ventos do mundo para produção de energia eólica, pois são ventos mais constantes, com uma velocidade estável e não mudam de direção com frequência. Os ventos no território brasileiro apresentaram um fator de capacidade média de 42,7% em 2019, já o fator de capacidade média mundial é de aproximadamente 34%. Na figura 2 é possível observar a evolução da Capacidade Instalada em MW no período de 2005 a 2024, nota-se um crescimento constante ao longo do tempo. É necessário ressaltar que os dados futuros apresentados na figura referem-se a contratos

viabilizados em leilões já realizados e no mercado livre (ABEEólica 2021).



Figura 3.2: Evolução da Capacidade Instalada em MW

Fonte: : Adaptada da ANEEL/Associação Brasileira de Energia Eólica, 2021.

Sendo assim, evidencia-se a necessidade de estudos que possibilitem compreender através da implementação de metodologias adequadas, a relação de dependência entre os componentes da geração de energia eólica.

3.1

Vento e a conversão da energia eólica

A definição clássica do vento apresenta o mesmo como o ar em movimento. Este trata-se do deslocamento frequente e contínuo do ar na superfície. Embora não seja visível o ar em movimento, é possível notar sua existência e intensidade por diversos meios. A troposfera é a camada interna da atmosfera com aproximadamente 12 km de altura, a mesma é composta por ar que torna-se mais denso de acordo com a altura, movimentando-se devido às diferenças de pressão existentes, gerando então o vento (AGUILAR, 2016).

A velocidade do vento é o bem natural mais importante para geração de energia eólica, sendo a velocidade dos ventos locais uma variável fundamental para avaliação do potencial eólico de uma região. Outro fator essencial é a rugosidade superficial do terreno, os ventos de até 100 metros de altura são significativamente influenciados por esta rugosidade, ou seja, quanto maior a rugosidade do terreno, menor a velocidade do vento, o que afeta diretamente

a geração de energia. Na tabela 3.2, observa-se a classificação da rugosidade do terreno e a quantidade de energia que pode ser aproveitada da velocidade do vento, segundo o tipo de superfície.

Tabela 3.2: Capacidade instalada e Número de Parques por estado

Classe	Energia (%)	Tipo de solo
0	100	Plano de água
1	52	Planície
2	39	Planície com casas
3	24	Floresta com vila
4	13	Grande cidade

Fonte: Adaptado do PER, (2013)

Isto posto, segundo relatório da Confederação Nacional da Indústria (CNI, 2018), tem-se que o potencial de geração disponível é função cúbica da velocidade, ou seja, se a velocidade do vento dobrar, o potencial aumenta aproximadamente oito vezes. Por este motivo, o regime dos ventos é o fator determinante para avaliar se o aproveitamento da fonte eólica é promissor. Via de regra, considera-se a implementação de aerogeradores em locais cuja velocidade média anual dos ventos supera 6 m/s.

Segundo Filho (2019), o aproveitamento total da energia eólica requer a implantação eficiente e estruturada de parques eólicos, geralmente com cerca de dez a cem aerogeradores, com potência unitária de 300 a 750 kW. Sendo estes aerogeradores instalados a 200 m uns dos outros, para evitar interferências entre os mesmos. Uma densidade de instalação típica é 10 MW/km.

Os barcos a vela foram uma das primeiras formas de aproveitamento da energia eólica. As velas são capazes de capturar a energia do vento para impulsionar o barco pela água. Outra maneira eficaz de utilização da energia eólica se deu por meio dos primeiros moinhos de vento (AGUILAR, 2016).

Deste modo, tem-se que as estações do ano, assim como as horas

do dia impactam significativamente na quantidade de energia disponível no vento. Como supramencionado a topografia e a rugosidade do solo também apresentam influência na distribuição dos ventos e de sua velocidade em um local.

3.2

Capacidade eólica mundial

Devido a preocupação mundial sobre as mudanças climáticas e a escassez de recursos, medidas sustentáveis tornaram-se extremamente importantes, visto que o futuro depende do trabalho sustentável que é exercido atualmente. De acordo com Gsanger (2021), a eficiência energética está em pauta mundialmente, a geração de energia eólica apresentou um crescimento significativo em 2020, dado que alguns órgãos governamentais mostram-se fortemente comprometidos com fontes de energia renováveis, acredita-se então que este será um crescimento constante. Nesse sentido, há grande possibilidade do mundo entrar no caminho para cumprir os objetivos do acordo de Paris, limitando assim o aquecimento global a $1,5^{\circ}\text{C}$ (WWEA, 2021).

De acordo com estatísticas relacionadas a energia eólica publicadas pela *World Wind Energy Association* (2021), o mundo alcançou um novo recorde em número de novas instalações de energia eólica, sendo adicionado aproximadamente 93 gigawatts de novas turbinas eólicas em 2020. Países como: China, Estado Unidos da América e Rússia estabeleceram novos recordes de instalação, já os mercados europeus registraram um crescimento modesto.

Isto posto, ainda segundo o relatório WWEA (2021) o mercado global de novas turbinas atingiu um volume total de 93 gigawatts, sendo este número cerca de 50% a mais que no ano de 2019, atingindo o recorde de instalação em um único ano. Dessa forma, tem-se que a capacidade total de todos os parques eólicos do mundo atingiu 744 gigawatts, sendo este número o suficiente para gerar 7% da demanda mundial de eletricidade.

Na figura 3.3, pode-se observar o crescente aumento da nova capacidade

instalada de energia eólica mundial no período de 2016 a 2020. Porém, segundo relatório GWEC mesmo a indústria eólica global alcançando o recorde de nova capacidade em 2020, o mundo ainda precisa instalar no mínimo mais 180 GW a cada ano, para então evitar os impactos extremos das mudanças climáticas.

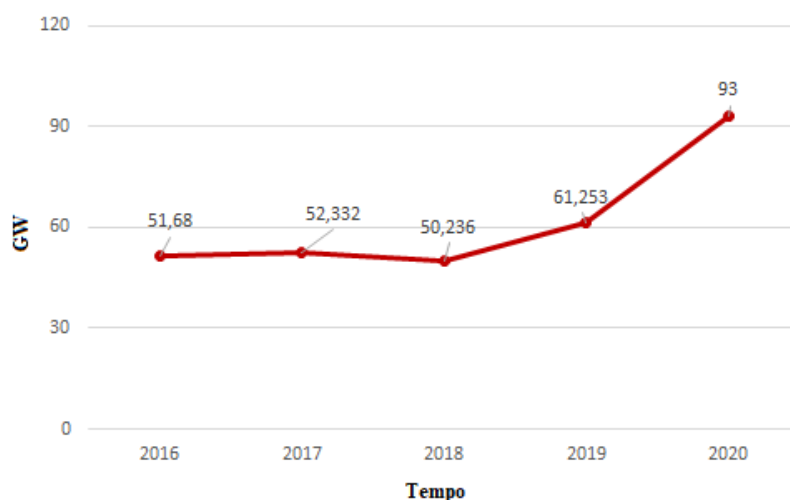


Figura 3.3: Evolução da nova Capacidade Instalada mundial em GW

Fonte: Adaptado de World Wind Energy Association, 2021.

A figura 3.4 apresenta a capacidade instalada acumulada total mundial. A capacidade total de energia eólica alcançou uma marca de 744 GW, ajudando o mundo a evitar mais de 1,1 bilhão de toneladas de CO_2 por ano, este número é equivalente às emissões anuais de carbono de toda a América do Sul, segundo as estimativas apresentadas no relatório GWEC.

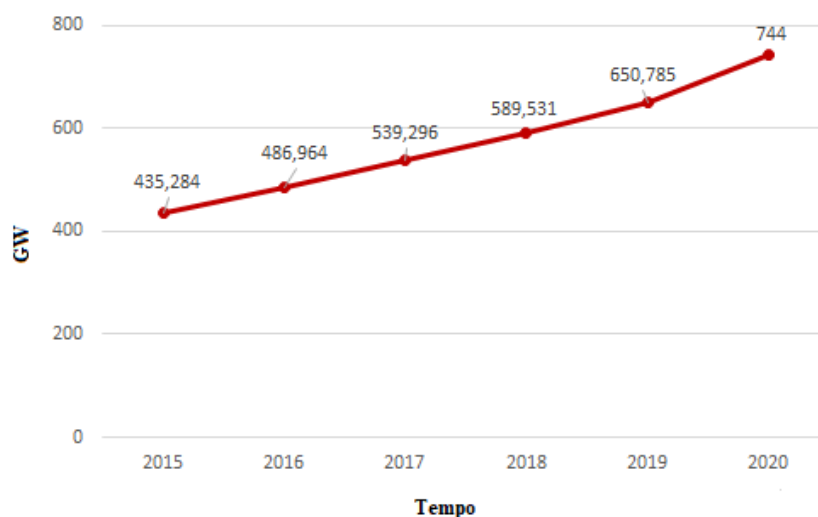


Figura 3.4: Capacidade instalada acumulada total mundial em GW

Fonte: Adaptado de World Wind Energy Association, 2021.

Mesmo com a pandemia da covid 19 o mundo alcançou um forte desenvolvimento eólico. Na China foram construídos cerca de 52 gigawatts em um ano, o que corresponde a uma participação de mercado de 56%. Com esta adição o país passou a ter capacidade eólica instalada de 289 gigawatts, representando 39% da capacidade global. O mercado dos Estados Unidos também se desenvolveu, crescendo cerca de 17 gigawatts em 2020. Com esse novo recorde e os parques eólicos já existentes o país totalizou 122 gigawatts, consolidando sua posição de número 2 em todo o mundo. Na tabela 3.3, pode-se observar a capacidade instalada dos países nos anos de 2018 a 2020.

Tabela 3.3: Capacidade instalada dos países em *GW*

Países	2020	2019	2018
China	290,000	237,029	209,529
Estados Unidos	122,328	105,433	91,363
Alemanha	62,784	61,357	59,313
Índia	38,625	37,529	35,129
Espanha	27,446	25,808	23,494
Reino Unido	24,167	23,515	20,743
França	17,949	16,646	15,313
Brasil	18,010	15,452	14,707
Canadá	13,588	13,413	12,816
Itália	10,850	10,512	9,958
Turquia	9,305	8,056	7,369
Outros países	110,000	96,0355	84,814
Total	744,000	650,785	589,547

Fonte: Adaptado de WWEA, 2021.

3.3

Potência Extraída do Vento

De acordo com Dutra (2008), a turbina eólica possibilita a captura da energia cinética gerada através do movimento das pás do aerogerador. A energia cinética de uma massa de ar em movimento é denotada por m e velocidade V é dada pela equação (3-1).

$$E = \frac{1}{2}mV^2 \quad (3-1)$$

Segundo Lima (2011), a velocidade instantânea do vento V pode ser descrita como um valor médio acrescentado de um desvio a partir da média (equação 3-2).

$$V = \bar{V} + v' \quad (3-2)$$

No qual, \bar{V} trata-se da velocidade média do vento e v' é a flutuação.

A potência do vento pode ser calculada de acordo com a equação (3-3)

$$P_d = \frac{1}{2} \times \rho \times A \times V^3 \quad (3-3)$$

Sendo,

P_d = potência média do vento disponível em Watts (W);

ρ = densidade do ar seco equivalente a $1,225 \text{ kg/m}^3$;

A = área de varredura do rotor (m^2);

V = velocidade média do vento (m/s).

Tem-se então que velocidade do ar é descontínua de V para V_0 e a taxa de fluxo de massa de ar através das lâminas rotativas (tm) é, portanto, resultado da multiplicação da densidade com a velocidade média, de acordo com a equação (3-4).

$$T_m = \rho \times \frac{V + V_0}{2} \quad (3-4)$$

Assim, a potência P_a extraída pelo rotor pode ser calculada de acordo com a equação (3-5)

$$P_a = \frac{1}{2} \times \left[\rho \times A \frac{V+V_0}{2} \right] \times (v^2 - v_0^2) \quad (3-5)$$

Tem-se então que a quantidade de energia contida no vento apresenta uma variação de acordo com o cubo da velocidade média do mesmo, porém não é possível capturar toda esta potência através de um aerogerador. Desta maneira, existe um máximo de energia cinética do vento que pode ser convertido para energia mecânica por uma turbina eólica, esta captura máxima é determinada pela "Lei de Betz". A fração da potência obtida é totalmente dependente do Coeficiente de cada turbina eólica, aqui denominado por C_p ;

então o coeficiente de potência C_p indica a fração da potência eólica disponível, ou seja, P_d convertida em potência mecânica. Portanto, a equação (3-5) pode ser reescrita da seguinte forma:

$$P_a = \frac{1}{2} \times \rho \times A \times V^3 \times C_p(V) \quad (3-6)$$

Sendo, P_a a potência eólica aproveitável e o coeficiente $C_p(V)$ o desempenho da turbina eólica. Segundo Aguilar (2015), o máximo valor teórico é de aproximadamente 16/27 ou 59%, este valor é conhecido como "Limite de Betz". Deve-se ressaltar que as turbinas eólicas possuem limites de operação como velocidade de partida e de corte.

O limite de Betz é baseado em um atuador de disco aberto. Se um difusor for usado para coletar o fluxo de vento adicional e direcioná-lo através da turbina, mais energia pode ser extraída, porém o limite ainda se aplica à seção transversal de toda a estrutura.

Desta forma, pode-se calcular o valor de $C_p(V)$ de acordo com a equação (3-7).

$$C_p(V) = \frac{\left(1 + \frac{V_0}{V}\right) \times \left(1 - \left(\frac{V_0}{V}\right)^2\right)}{2} \quad (3-7)$$

3.4 Distribuição Weibull

De acordo com Tuller (1984), a literatura especializada em fontes de energia renováveis aponta a distribuição Weibull como a mais utilizada para modelar a velocidade do vento. Esta distribuição tem sido implementada de maneira significativa e eficiente para análise da velocidade do vento, sendo frequentemente empregada na avaliação do potencial eólico (DE SOUZA, 2019).

O conhecimento sobre o modelo de distribuição de probabilidade que melhor caracteriza o comportamento do regime dos ventos é extramente

importante na avaliação do potencial eólico. Deste modo, segundo a literatura a distribuição Weibull (w) é uma das distribuições mais utilizadas, devido a simplicidade para a estimação dos parâmetros de seu modelo e a boa aderência do modelo às inúmeras e distintas bases de dados de velocidade do vento (NASCIMENTO, 2020).

Neste estudo trabalha-se com duas variáveis, sendo uma delas a velocidade do vento, logo, é de suma importância identificar a distribuição adequada, para então prosseguir com a análise. A função densidade de probabilidade Weibull (f.d.p.) com três parâmetros Weibull é apresentada a seguir:

$$f_w(\nu; \alpha, \beta) = \frac{\alpha}{\beta} \left(\frac{\nu}{\beta} \right)^{\alpha-1} e^{-\left(\frac{\nu}{\beta} \right)^{\alpha}} \quad (3-8)$$

No qual, função densidade acumulada (f.d.a.) é dada por:

$$F_{\nu} = 1 - e^{-\left(\frac{\nu}{\beta} \right)^{\alpha}} \quad (3-9)$$

Sendo, ν corresponde a observação de velocidade do vento, α o parâmetro de forma, e β o parâmetro de escala.

4

Metodologia

Neste capítulo será apresentada a metodologia utilizada para análise dos dados de interesse.

4.1

Cópuas

Esta seção apresenta a definição de cópuas, bem como suas propriedades. A função cópula pode ser definida como sendo uma função de ligação entre as distribuições marginais univariadas que tem como resultado distribuições multivariadas. Em Nelsen (2006), esta teoria é definida por dois pontos de vista: o primeiro aborda cópuas como sendo funções que juntam ou acoplam distribuições marginais às suas distribuições multivariadas. A outra forma de interpretação afirma que cópuas são distribuições multivariadas cujas distribuições marginais são uniformes no intervalo $[0,1]$.

A teoria de cópula, como uma função de dependência entre variáveis aleatórias, foi introduzida por Sklar (1959), entretanto, suas aplicações são recentes em diversas áreas. Em sua pesquisa Sklar estudava sobre a distribuição tridimensional conjunta, e com isso introduziu funções auxiliares definidas no suporte unitário, ligando a função de distribuição às suas marginais (SKLAR, 1959). Segundo Nelsen (2006), uma cópula equivale a uma função de distribuição multivariada com marginais uniformes em $[0, 1]$, as mesmas contém toda a estrutura de dependência entre as variáveis aleatórias observadas. A teoria em questão tem sido amplamente utilizada para casos bivariados, sendo este o foco deste estudo, dado que o objetivo principal é observar a estrutura de dependência entre duas variáveis aleatórias, sendo elas a velocidade do vento e a geração de energia eólica. Deste modo, tem-se que a cópula $C(u, v)$ trata-se de uma função de distribuição bivariada de duas variáveis aleatórias u e v . Por conseguinte, uma cópula bidimensional para quaisquer $0 \leq u_1 \leq u_2 \leq 1$ e

$0 \leq v_1 \leq v_2 \leq 1$, possui as seguintes características:

1. $C(u, 0) = \int_0^u \int_0^0 c(u, v) \, dvdu = 0 = \int_0^v \int_0^0 c(u, v) \, dudv = C(0, v)$;
2. $C(u, 1) = u = \int_0^u \int_0^1 c(u, v) \, dvdu$ e $C(1, v) = v = \int_0^v \int_0^1 c(u, v) \, dudv$;
3. $C(u_2, v_2) - C(u_1, v_2) - C(u_2, v_1) + C(u_1, v_1) \geq 0$.

Então, tem-se que a característica (1), resultará em $C(0, v)$, sempre que ao menos uma das componentes de $\{u\}$ for nula, a característica (2), aponta que $C(u, 1) = u$ e $C(1, v) = v$ ocorrerá sempre que todas as componentes de $\{u\}$ forem iguais a 1, exceto a i -ésima, ou todas as componentes de $\{v\}$ forem iguais a 1, exceto a i -ésima, e juntas as características (1) e (2) mostram que a cópula apresenta distribuições marginais uniformes em $[0, 1]$. Já a característica (3) evidencia que a cópula é uma função crescente. Sendo assim, todo o conceito supracitado sobre cópulas é oriundo do teorema de Sklar(1959).

Desta maneira, a cópula foi definida como uma função real apresentando algumas características específicas. Como mencionado anteriormente, sua interpretação probabilística será dada através do Teorema de Sklar. Porém, torna-se necessário ressaltar antes da apresentação do teorema de Sklar uma breve introdução conceitual sobre a interpretação probabilística da função de cópula (NELSEN, 2006).

Sendo x e y números reais, pode-se então definir X e Y como duas variáveis aleatórias com funções distribuições acumuladas $F_X(x) = P[X \leq x]$ e $F_Y(y) = P[Y \leq y]$, cuja função de distribuição conjunta é definida por $F(x, y) = P[X \leq x, Y \leq y]$. Logo, cada par de números (x_i, y_i) gera um ponto $(F_X(x_i), F_Y(y_i))$ contido no intervalo $[0, 1] \times [0, 1]$. Então, tem-se que cópulas são funções que associam o ponto $(F_X(x_i), F_Y(y_i))$ a um valor específico da sua função conjunta $F(x, y)$, definida no intervalo $[0, 1]$

4.1.1

Teorema de Sklar

A teoria de cópulas é fundamentada pelo teorema de Sklar (1959), de acordo com este teorema, pode-se tratar uma distribuição multivariada pela sua estrutura de dependência, a cópula, e suas marginais.

Teorema 4.1 *Então, para uma distribuição conjunta n -dimensional F de um vetor de variáveis aleatórias $\mathbf{X} = (X_1, \dots, X_n) \in \mathbb{R}^n$ com marginais $F_1(x_1), \dots, F_n(x_n)$, existe uma função cópula C tal que:*

$$F(x_1, \dots, x_n) = C(F_1(x_1), \dots, F_n(x_n)) \quad (4-1)$$

Caso, F_1, \dots, F_n sejam contínuas, a C é única.

A partir do resultado supracitado, obtêm-se a função cópula, aqui representada por C .

$$C(u_1, \dots, u_n) = F(F_1^{-1}(u_1), \dots, F_n^{-1}(u_n)), \quad \forall u = (u_1, \dots, u_n) \in [0, 1]^n \quad (4-2)$$

Sendo, F_i^{-1} , uma função inversa generalizada, ou seja, $F_1^{-1}(u_1), \dots, F_n^{-1}(u_n)$, são as funções inversas de $F_1(x_1), \dots, F_n(x_n)$.

Com isto, a distribuição multivariada das variáveis transformadas, é denominada de cópula. A densidade associada a uma cópula é definida pelas derivadas parciais em relação a \mathbf{u} da função $C(u_1, \dots, u_n)$ (Equação 4-2):

$$c(u_1, \dots, u_n) = \frac{\partial C^n(u_1, \dots, u_n)}{\partial u_1 \dots \partial u_n} \quad (4-3)$$

Isto posto, segundo o Teorema de Sklar (1959), a densidade do vetor aleatório \mathbf{X} podendo ser escrita da seguinte forma:

$$f(x_1, \dots, x_n) = c(u_1, \dots, u_n) \prod_{i=1}^n f_i(x_i) \quad (4-4)$$

No qual, $u_1 = F_1(x_1; \theta_1), \dots, u_n = F_n(x_n; \theta_n)$, sendo θ_i os vetores p_i dimensionais dos parâmetros das marginais, $i = 1, 2, \dots, n$

Tendo em foco o caso bivariado, tem-se que o teorema de Sklar para o mesmo pode ser apresentado da seguinte forma:

Teorema 4.2 *Sejam X e Y v.a's que possuem função distribuição conjunta F e distribuições marginais F_X e F_Y , respectivamente. Então existe uma cópula C tal que:*

$$F(x, y) = C(F_X(x), F_Y(y)) \quad (4-5)$$

Para todo x, y em \mathbb{R} . Então se X e Y são contínuas, C é única. Logo, se C é uma cópula e F_X e F_Y são funções de distribuição, então F trata-se da distribuição conjunta com marginais F_X e F_Y .

Do Teorema proposto por Sklar (nesta pesquisa apresentado no teorema 4.1) segue o corolário que, dada uma cópula que satisfaça a equação 4-5, a mesma pode ser obtida por inversão, conforme a equação a seguir.

$$C(u, v) = F(F_X^{-1}(u), F_Y^{-1}(v)) \quad (4-6)$$

Sendo, $F_X^{-1}(u), \dots, F_Y^{-1}(v)$, funções pseudo inversas de $F_X(x), F_Y(y)$.

Por conseguinte, apresenta-se a relação entre cópulas, distribuições conjuntas e marginais, possibilitando observar brevemente a origem da equação 4-5, base do Teorema de Sklar (Demonstração formal em Cherubini, Luciano e Vecchiato (2004) ou Nelsen (2006)).

Primeiramente, dado que X e Y são duas variáveis aleatórias que possuem uma distribuição conjunta $F(x, y)$, pode-se então escrever $F(x, y)$ da seguinte maneira.

$$F(x, y) = P(X \leq x, Y \leq y) \quad (4-7)$$

Se $F_X(x)$ e $F_Y(y)$ forem definidos como as distribuições marginais de X e Y , ou seja, $F_X(x) = P[X \leq x]$ e $F_Y(y) = P[Y \leq y]$, pode-se dizer então que ambas seguem uma distribuição uniforme padrão, $F(X) \sim U$ e $F(Y) \sim V$, e que suas funções inversas podem ser escritas como $F_X^{-1}(U) \sim X$ e $F_Y^{-1}(V) \sim Y$, o que permite reescrever a equação 4-7.

$$F(x, y) = P(F_X^{-1}(U) \leq x, F_Y^{-1}(V) \leq y) \quad (4-8)$$

Ou então:

$$F(x, y) = P(U \leq F_X(x), V \leq F_Y(y)) \quad (4-9)$$

Sendo então, uma cópula definida como uma função de distribuição conjunta cujas distribuições marginais são uniformes no intervalo $[0, 1]$, pode-se assim substituir o segundo termo da equação 4-9 pela cópula C , logo tem-se que uma cópula é uma função que associa uma função de distribuição conjunta às suas funções de distribuição marginais, conforme a equação abaixo.

$$F(x, y) = C(F_X(x), F_Y(y)) \quad (4-10)$$

Desta forma, nota-se que uma das principais vantagens da implementação da teoria de cópulas é a maneira como as mesmas representam uma distribuição conjunta de probabilidades. Segundo Cherubini e Luciano (2001), as cópulas possibilitam uma análise ampla da relação de dependência entre variáveis, pois a escolha das distribuições marginais pode ser feita de maneira independente da modelagem da estrutura de associação das variáveis observadas.

Exemplo 1 (*GRUGEL, 2015*) Seguindo a teoria apresenta anteriormente, considera-se um vetor aleatório bi-dimensional (X, Y) , com função de distribuição conjunta dada por:

$$F_{X,Y}(x,y) = \begin{cases} 1 - e^{-\lambda y} - e^{-\lambda x} + e^{-\lambda(x+y)} & , \text{ se } x > 0, y > 0 \\ 0 & , \text{ caso contrário.} \end{cases}$$

Com isto, pode-se encontrar as marginais da seguinte maneira.

$$F_X(x) = \lim_{y \rightarrow \infty} F_{X,Y}(x,y) = \lim_{y \rightarrow \infty} (1 - e^{-\lambda y} - e^{-\lambda x} + e^{-\lambda(x+y)}) = 1 - e^{-\lambda x}.$$

e

$$F_Y(y) = \lim_{x \rightarrow \infty} F_{X,Y}(x,y) = \lim_{x \rightarrow \infty} (1 - e^{-\lambda y} - e^{-\lambda x} + e^{-\lambda(x+y)}) = 1 - e^{-\lambda y}.$$

Então, nota-se que as distribuições marginais de X e Y são:

$$X \sim \text{Exp}(\lambda) \text{ e } Y \sim \text{Exp}(\lambda)$$

Encontrando $x = h(u)$ (x em função de u e $y = g(v)$ (y em função de v , tem-se que:

$$u = 1 - e^{-\lambda x}$$

$$u - 1 = -e^{-\lambda x}$$

$$1 - u = e^{-\lambda x}$$

$$\log(1 - u) = -\lambda x$$

$$x = \frac{-\log(1 - u)}{\lambda}$$

$$\text{Analogamente, } y = \frac{-\log(1 - v)}{\lambda}$$

Desta maneira, pode-se encontrar a cópula associada ao vetor aleatório (X, Y) :

$$\begin{aligned}
C(u, v) &= F\left(\frac{-\log(1-u)}{\lambda}, \frac{-\log(1-v)}{\lambda}\right) \\
&= 1 - e^{\frac{\lambda \log(1-v)}{\lambda}} - e^{\frac{\lambda \log(1-u)}{\lambda}} + e^{\lambda(\frac{\log(1-u)}{\lambda} + \frac{\log(1-v)}{\lambda})} \\
&= 1 - (1-v) - (1-u) + (1-u)(1-v) \\
&= 1 - 1 + v - 1 + u + 1 - v - u + uv \\
&= uv
\end{aligned}$$

Com base no Teorema de Sklar, como $C(u, v) = uv$, pode-se dizer que X e Y são independentes.

4.1.2

Limites de Fréchet-Hoeffding

Após apresentar os conceitos fundamentais da teoria de cópulas, assim como suas propriedades, torna-se necessário destacar um resultado importante, sendo este os limites que uma função cópula pode assumir. Dado que C é uma cópula, então para todo ponto (u, v) pertencente a $[0, 1] \times [0, 1]$, tem-se que as funções $W(u, v) = \max(0, u + v - 1)$ e $M(u, v) = \min(u, v)$ são cópulas bivariadas e possuem a propriedade de limitar em que conjunto o intervalo $C(u, v)$. Assim, se $C(u, v)$ é uma cópula de duas dimensões, logo:

$$W(u, v) \leq C(u, v) \leq M(u, v) \quad (4-11)$$

Isto, para quaisquer u e v no intervalo unitário.

Tratando-se do caso multivariado também existem cópulas limítrofes dadas por:

$$W(u_1, u_2, \dots, u_n) = \max\left(1 - n + \sum_{i=1}^n u_i, 0\right) \quad (4-12)$$

$$M(u_1, u_2, \dots, u_n) = \min(u_1, u_2, \dots, u_n) \quad (4-13)$$

De maneira que $W(\mathbf{u}) \leq C(\mathbf{u}) \leq M(\mathbf{u})$.

A apresentação de estudos mais aprofundados sobre a questão da estrutura de dependência que envolve os limites de Fréchet-Hoeffding encontra-se amplamente detalhada em Nelsen (2006).

4.2

Medidas de dependência e concordância

Segundo Joe (2015), as medidas de dependência e concordância são de extrema importância, pois possibilitam quantificar a relação existente entre variáveis observadas. De acordo com Nelsen (2006), o estudo sobre a relação de dependência entre variáveis aleatórias é um dos temas mais abordados em teoria das probabilidades e estatística. Dado que a função cópula objetiva representar essa relação, pode-se notar que a mesma está diretamente relacionada ao estudo de medidas de dependência entre variáveis aleatórias.

4.2.1

Medidas de dependência

Os conceitos de dependência tornam-se necessários e fundamentais para o estudo das distribuições multivariadas, os mesmos serão apresentados brevemente nesta seção. Estes conceitos, assim como a demonstração dos mesmos podem ser encontrados de maneira detalhada em Nelsen(2006), Cherubini (2004) e Joe(2015).

4.2.1.1

Coeficiente de Correlação Linear

A associação entre variáveis aleatórias pode ser realizada por meio do coeficiente de correlação linear ou de Pearson. Supondo então duas variáveis aleatórias X e Y , o mesmo é calculado com base na covariância e desvio-padrão

das variáveis observadas. Sendo este dado por:

$$\rho(X, Y) = \frac{Cov(X, Y)}{\sqrt{\sigma_X^2 \sigma_Y^2}} \quad (4-14)$$

No qual, $Cov(X, Y) = E(XY) - E(X)E(Y)$ é a covariância entre X e Y , σ_X^2, σ_Y^2 correspondem as variâncias de X e Y , respectivamente. Com isto, seu estimador pode ser representado da seguinte maneira:

$$\hat{\rho}(X, Y) = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sum_{i=1}^n (X_i - \bar{X})^2 (Y_i - \bar{Y})^2} \quad (4-15)$$

Sendo, \bar{X} e \bar{Y} os estimadores do valor esperado de X e Y , respectivamente, e n trata-se do número de observações da amostra.

Isto posto, tem-se que o coeficiente de correlação linear apresenta as seguintes propriedades.

1. $-1 \leq \rho \leq 1$;
2. X e Y são positivamente correlacionadas se $\rho(X, Y) > 0$, negativamente correlacionadas se $\rho(X, Y) < 0$ e não correlacionadas se $\rho(X, Y) = 0$;
3. se X e Y forem independentes logo, $\rho(X, Y) = 0$;
4. se X e Y são perfeitamente correlacionados então $|\rho(X, Y)| = 1$;
5. ρ é invariante para transformações lineares das variáveis aleatórias, *i.e.*,

$$\rho(aX + b, Y) = \text{sin}(a)\rho(X, Y), a \neq 0 \quad b \in \mathbb{R}$$

A função $\text{sin}(x)$ retorna o sinal da variável x , visto que se $x < 0$, o valor é -1 , se $x = 0$, o valor é 0 e se $x > 0$, o valor é 1 .

Tem-se então, que está trata-se de uma medida de dependência, muito utilizada devido à seu método simplificado de estimação, porém a mesma apresenta algumas limitações como:

- O coeficiente de correlação linear não é invariante para transformações não lineares das variáveis aleatórias;
- Não é uma medida de associação robusta, pois uma única observação pode ter uma influência significativa no valor de ρ , logo, não se pode utilizá-lo quando a amostra apresenta *outliers*.

4.2.1.2

Coeficiente de Dependência na Cauda

Este conceito concentra-se na relação entre os valores extremos das variáveis. De maneira geral, supõe-se X e Y duas variáveis aleatórias contínuas, a dependência na cauda é dada pela probabilidade condicional de que a função distribuição de X exceda um dado valor v , assumindo que esse valor tenha sido excedido pela função distribuição de Y .

Com isto, o conceito de coeficiente de dependência na cauda está relacionado com a dependência existente na cauda inferior ou superior de uma distribuição conjunta bivariada. Tem-se então, que o coeficiente de dependência na cauda superior de duas variáveis aleatórias, sendo elas X e Y , com funções de distribuição F_X e F_Y , respectivamente, é dado pela seguinte equação.

$$\lambda_u = \lim_{\alpha \rightarrow 1^-} P[Y > F_Y^{-1}(\alpha) | X > F_X^{-1}(\alpha)] \quad (4-16)$$

Logo, analogamente o coeficiente de dependência na cauda inferior é dado por:

$$\lambda_l = \lim_{\alpha \rightarrow 0^+} P[Y \leq F_Y^{-1}(\alpha) | X \leq F_X^{-1}(\alpha)] \quad (4-17)$$

Sendo, $F_X^{-1}(\alpha)$ e $F_Y^{-1}(\alpha)$ as funções quantílicas de F_X e F_Y , respectivamente. Dito isto, tem-se que os coeficientes de dependência na cauda apresentam as seguintes propriedades:

1. $\lambda_u, \lambda_l \in [0, 1]$;

2. se $\lambda_u \in [0, 1]$, então, X e Y são assintoticamente relacionados na cauda superior; se $\lambda_u = 0$, então, X e Y são assintoticamente independentes na cauda superior;
3. se $\lambda_l \in [0, 1]$, então tem-se que X e Y são assintoticamente correlacionados na cauda inferior; se $\lambda_l = 0$, então, X e Y são assintoticamente independentes na cauda inferior;
4. os coeficientes λ_u e λ_l são invariantes para transformações estritamente crescentes das variáveis aleatórias.

4.2.2

Medidas de concordância

De acordo com Cherubini (2004), dado um par de variáveis aleatórias X e Y , o mesmo é considerado concordante se valores grandes de X estão associados a valores grandes Y , e se valores pequenos de X estão associados a valores pequenos de Y . De modo mais formal, sendo (x_1, y_1) e (x_2, y_2) observações de (X, Y) , pode-se dizer que existe concordância se $(x_1 - y_1)(x_2 - y_2) > 0$, e se $(x_1 - y_1)(x_2 - y_2) < 0$ afirma-se a existência de uma discordância. As medidas mais conhecidas baseadas neste conceito são o Tau de Kendall e o Rho de Sperman.

4.2.2.1

Tau de Kendall

A medida de concordância Tau de Kendall é definida na literatura como a diferença entre a probabilidade de concordância e a probabilidade de discordância de duas variáveis aleatórias. Uma definição formal para a versão populacional desta medida pode ser observada na equação a seguir.

$$\tau = P((X_1 - X_2)(Y_1 - Y_2) > 0) - P((X_1 - X_2)(Y_1 - Y_2) < 0) \quad (4-18)$$

Sendo (X_1, Y_1) e (X_2, Y_2) vetores aleatórios iid, ou seja, independentes e identicamente distribuídos, formados a partir das observações (x_i, y_i) e (x_i, y_j) do vetor de variáveis aleatórias (X, Y) .

A versão amostral desta medida de concordância pode ser facilmente calculada, de modo que dado o vetor de variáveis aleatórias contínuas (X, Y) , e supondo uma amostra contendo n observações $\{(x_1, y_1), \dots, (x_n, y_n)\}$, pode-se então afirmar que existem $\binom{n}{2}$ pares distintos (x_i, y_i) e x_j, y_j . Logo, c é o número de pares concordantes e d o número de pares discordantes. Com isto, tem-se que a versão amostral do Tau de Kendall se dá pela equação a seguir.

$$\tau = \frac{(c - d)}{\binom{n}{2}} = \frac{2}{n(n-1)}(c - d) \quad (4-19)$$

Assim, a medida tau de Kendall é a probabilidade de concordância menos a probabilidade de discordância entre os pares (X_1, Y_1) e (X_2, Y_2) podendo ser estimada conforme a equação abaixo.

$$\hat{\tau} = \binom{t}{2}^{-1} \sum (\text{sinal}[(X_i - X_j)(Y_i - Y_j)]) \quad (4-20)$$

No qual, $\binom{t}{2}$ corresponde aos pares distintos (x_i, y_i) e (x_j, y_j) e $\sum(\text{sinal}[(X_i - X_j)(Y_i - Y_j)])$ trata-se da diferença entre o número de pares concordantes e discordantes.

Isto posto, sejam X e Y variáveis aleatórias contínuas com distribuição conjunta dada pela cópula C , tem-se que a versão do tau de Kendall, estabelecida como uma função cópula (τC) , é dada pela seguinte equação.

$$\tau = \tau C = 4E(C(U, V)) - 1 \Leftrightarrow \iint_{I^2} C(u, v) dC(u, v) - 1 \quad (4-21)$$

Logo, a estimação do Tau de Kendall pode ser utilizada para a estimação da cópula que relaciona as variáveis aleatórias X e Y , já que esta relação é válida entre o Tau de Kendall e a cópula (demonstração completa apresenta

em Nelsen (2006)).

4.2.2.2

Rho de Sperman

O coeficiente de correlação de Spearman, também conhecido correlação ordinal, é definido como o coeficiente de correlação linear entre as ordens de X e Y .

Definição 4.1 *Seja (X, Y) vetores aleatórios reais. É chamado de coeficiente de correlação de Spearman entre X e Y o número real.*

$$\rho_S(X, Y) = \rho_L(F_x(X), F_y(Y)) \quad (4-22)$$

Proposição 4.1 *Seja, (X, Y) um vetor aleatório real com cópula C , é possível escrever ρ_S em função da cópula pertinente:*

$$\rho_S(X, Y) = 12 \int_0^1 \int_0^1 \{C(u, v) - uv\} dudv \quad (4-23)$$

4.3

Principais Famílias de Cópuas

De acordo com o Teorema de Sklar, uma distribuição multivariada pode ser tratada através da sua estrutura de dependência, cópula, e por suas marginais. Com este tratamento pode-se explicar a alta flexibilidade da modelagem via funções cópulas. Segundo Cherubini (2004), isto ocorre tanto nas aplicações empíricas quanto do ponto de vista teórico. Com isso, existem inúmeras funções cópulas propostas pela literatura. Nesta seção serão apresentados os principais tipos de cópulas bivariadas e a quais famílias pertencem.

4.3.1

Cópuas Elípticas

A classe de distribuições elípticas disponibiliza uma variedade de distribuições multivariadas, as mesmas compartilham de inúmeras propriedades

com a distribuição normal multivariada, o que proporciona uma facilidade de simular dados provenientes destas distribuições (EMBRECHTS, 2001b).

Seja H uma função de distribuição elíptica bivariada com marginais F_X e F_Y . Então, C é uma cópula elíptica, com parâmetro de dependência θ , dada pela seguinte equação:

$$C(u, v) = H(F_X^{-1}(u), F_Y^{-1}(v); \theta) \quad \forall (u, v) \in [0, 1]^2 \quad (4-24)$$

As principais distribuições elípticas são as distribuições multivariadas Normal ou Gaussiana e t-Student, as mesmas possuem o coeficiente de correlação linear como medida de dependência.

Definição 4.2 *Tem-se por definição sendo $\mu \in \mathbb{R}$ e Σ uma matriz real 2×2 , simétrica e definida positiva. Afirma-se que o vetor aleatório real $X = (X_1, X_2)$ tem uma distribuição elítica não degenerada com vetor de localização μ e matriz de dispersão de Σ se*

$$X = \mu + YAU \quad (4-25)$$

No qual A trata-se de uma matriz real 2×2 tal que $\Sigma = AA^T$, sendo então Y uma variável aleatória real não negativa e $U = (U_1, U_2)$ um vetor aleatório uniformemente distribuído no conjunto $S_d = \{x \in \mathbb{R}^2 : \|x\|_e^2 = 1\}$, sendo este independente de Y .

4.3.1.1 Cópula Gaussiana

A cópula Gaussiana bivariada é determinada tomando inicialmente a função de distribuição Normal bivariada como base, feito isto, substitui-se as variáveis pela sua inversa generalizada, como apresentado anteriormente. Deste modo, esta cópula é classificada como uma cópula elíptica. Logo, a cópula Gaussiana tem a seguinte expressão:

$$C_G(u, v; \rho) = \int_{-\infty}^{\phi^{-1}(u)} \int_{-\infty}^{\phi^{-1}(v)} \frac{1}{2\pi(1 - \rho^2)} \exp \left\{ -\frac{x^2 - 2\rho xy + y^2}{2(1 - \rho^2)} \right\} dv du \quad (4-26)$$

No qual, (u, v) trata-se de um vetor aleatório bidimensional $\in [0, 1]^2$, ρ é o coeficiente de correlação linear, e ϕ^{-1} representa a inversa da distribuição de densidade Normal padrão univariada.

A cópula Gaussiana possui parâmetro de correlação $-1 \leq \rho \leq 1$ e é simétrica; a dependência aumenta com o valor de ρ e não possui dependência na cauda, indicando que as variáveis aleatórias não possuem dependência em eventos considerados extremos.

4.3.1.2

Cópula t-Student

A cópula T-Student assim como a normal é uma cópula simétrica, porém sua vantagem é que a mesma possibilita a modelagem eficiente de um maior grau de dependência em suas extremidades. Sua equação é expressa por:

$$C_T(u, v; \rho, \nu) = \int_{-\infty}^{T_\nu^{-1}(u)} \int_{-\infty}^{T_\nu^{-1}(v)} \frac{\Gamma^{\frac{\nu+2}{2}}}{\Gamma(\frac{\nu}{2})\sqrt{(\pi\nu^2)(1-\rho^2)}} \exp\left\{1 + \frac{x^2 - 2\rho xy + y^2}{\nu(1-\rho^2)}\right\} dv du \quad (4-27)$$

No qual, ν representa os graus de liberdade, T^ν é a distribuição multivariada $t_\nu - Student$, já T_ν^{-1} trata-se da inversa da distribuição $t_\nu - Student$.

A cópula t-Student, possui parâmetros de correlação $-1 \leq \rho \leq 1$. Desta maneira, a implementação desta cópula possibilita a modelagem de um maior grau de dependência presente nas caudas. Os coeficientes de dependência são dados por:

$$\lambda_l = \lambda_u = 2\bar{t}_{\nu+1} \left(\sqrt{\nu+1} \sqrt{\frac{1-\rho}{1+\rho}} \right) \quad (4-28)$$

Sendo que $\bar{t}_{\nu+1} = 1 - t_{\nu+1}$ indica a simetria e o coeficiente de dependência na cauda é crescente em ρ e decrescente em ν .

4.3.2 Cópulas Arquimedianas

A família das cópulas Arquimedianas apresentam propriedades matemáticas interessantes, essa classe tem significativas aplicações em modelagens de problemas econômicos, financeiros e atuariais. Essa vasta gama de aplicações se dá pela sua simples lei de formação, pelo grande número de famílias de cópulas com esta classificação e das propriedades desejáveis que possuem.

As Cópulas Arquimedianas podem ser construídas de um modo simplificado, esta facilidade se deve ao fato da representação da Cópula Arquimediana permitir reduzir o estudo de cópula multivariada ao estudo de uma função univariada. A equação a seguir apresenta a função cópula arquimediana.

Definição 4.3 *Seja $\varphi : \mathbb{I} \rightarrow [0, \infty]$ uma função contínua e estritamente decrescente tal que $\varphi(1) = 0$. A pseudo-inversa de φ é a função $\varphi^{[-1]} : [0, \infty] \rightarrow \mathbb{I}$ que se dá por:*

$$\varphi^{[-1]}(t) = \begin{cases} \varphi^{-1}(t) & , \quad \text{se} \quad 0 \leq t \leq \varphi(0) \\ 0 & , \quad \text{se} \quad \varphi(0) \leq t \leq \infty \end{cases} \quad (4-29)$$

Dito isto, nota-se que $\varphi^{[-1]}$ é contínua e não crescente em $[0, \infty]$ e estritamente decrescente em $[0, \varphi(0)]$. Além disso,

$$\varphi^{[-1]}(\varphi(u)) = u, \quad \forall u \in \mathbb{I} \quad (4-30)$$

$$e \quad \varphi(\varphi^{[-1]}(t)) = \begin{cases} t & , \quad \text{se} \quad 0 \leq t \leq \varphi(0) \\ \varphi(0) & , \quad \text{se} \quad \varphi(0) \leq t \leq \infty \end{cases} \quad (4-31)$$

Consequentemente, se $\varphi(0) = \infty$ então, $\varphi^{[-1]} = \varphi^{-1}$.

Lema 4.1 *Seja φ uma função contínua e estritamente decrescente de $\mathbb{I} \rightarrow [0, \infty]$ no qual $\varphi(1) = 0$ e $\varphi^{[-1]}$ é a pseudo inversa de φ . Seja C a função de \mathbb{I}^2 para \mathbb{I} dada por*

$$C(u, v) = \varphi^{[-1]}(\varphi(u) + \varphi(v)) \quad (4-32)$$

Lema 4.2 *Dado que C é uma Cópula Arquimediana continua, então a sua densidade é dada por:*

$$c(u, v) = \frac{\partial C^2(u, v)}{\partial u \partial v} = \frac{\varphi''(C(u, v))\varphi'(u)\varphi'(v)}{[\varphi'(C(u, v))]^3} \quad (4-33)$$

4.3.2.1 Cópula Gumbel

Tem-se então que para a Cópula Gumbel, a função geradora é dada por (GENEST; FAVRE, 2007; NELSEN, 2013)

$$\varphi(v_i) = (-\ln v_i)^\theta, \forall \theta \geq 1 \quad (4-34)$$

No qual, $v_i = F_{X_i}(x_i)$

Sendo sua inversa dada pela equação:

$$\varphi^{-1}(t) = \exp\left(-t^{\frac{1}{\theta}}\right) \quad (4-35)$$

Fazendo a substituição das Equações 4-37 e 4-38 na Equação 4-35 obtém-se a chamada cópula de Gumbel-Hougaard e sua distribuição acumulada conjunta é dada pela seguinte equação:

$$F_{X_i, Y_i}(x_i, y_i) = C(u, v) = \exp\left\{-\left[\sum_{i=1}^k (-\ln(u, v))^\theta\right]^{\frac{1}{\theta}}\right\} \quad (4-36)$$

Na cópula de Gumbel-Hougaard, o grau de dependência entre as variáveis é medido a partir do parâmetro θ . Então, tem-se que para $\theta = 1$ obtém-se a cópula de marginais independentes e no caso de $\theta \rightarrow \infty$ obtém-se a cópula dedicada à dependência perfeita positiva. Desta maneira, a cópula de Gumbel-Hougaard permite modelar desde situações de independência até de dependência perfeita positiva. Ela é a única cópula Arquimediana de valores extremos e apresenta dependência na cauda superior (NELSEN, 2013).

4.3.2.2

Cópula Frank

É importante ressaltar que a cópula Frank é uma das mais importantes deste trabalho, a mesma representa a estrutura de dependência de alguns meses analisados.

A cópula de Frank é uma cópula de Arquimedes simétrica dada por:

$$C_{\alpha}(u, v) = -\frac{1}{\alpha} \ln \left(1 + \frac{(e^{-\alpha u} - 1)(e^{-\alpha v} - 1)}{e^{-\alpha} - 1} \right) \quad (4-37)$$

Logo, sua função geradora é dada por:

$$\varphi_{\alpha}(t) = -\ln \left(\frac{\exp(-\alpha t) - 1}{\exp(-\alpha) - 1} \right) \quad (4-38)$$

No qual, $\alpha(-\infty, \infty)$

Na cópula de Frank, o grau de dependência entre as variáveis (ou preditores) também é medido por um único parâmetro denotado aqui por α .

4.3.2.3

Cópula Joe

A Cópula de Joe, é caracterizada por permitir dependência somente na cauda superior, pode ser definida da seguinte forma

$$C_{\theta}^{Joe}(u, v) = 1 - [(1 - u)^{\theta} + (1 - v)^{\theta} - (1 - u)^{\theta}(1 - v)^{\theta}]^{1/\theta} \quad (4-39)$$

No qual, $\theta \geq 1$, com $\theta = 1$ correspondendo a independência e o coeficiente de dependência da cauda superior dado por $\lambda_U = 2 - 2^{1/\theta}$.

4.3.3

Cópulas de Valores Extremos

A designação desta classe de cópulas indicia que são definidas com base na Teoria de Valores Extremos. Então considerando (X, Y) um vetor aleatório real absolutamente contínuo com função de distribuição F e funções de distribuição marginais $F_i, i = 1, 2$. Seja ainda C_F a respectiva cópula.

Definição 4.4 A Cópula de Valores Extremos pode ser definida da seguinte maneira.

$$C(u, v) = \lim_{n \rightarrow \infty} C_F^n(u^{1/n}, v^{1/n}) \quad \forall (u, v) \in [0, 1]^2 \quad (4-40)$$

Com isto, afirma-se que C_F pertence ao domínio de atração de C . Além disso, a noção de cópula de valores extremos está diretamente relacionada com a de cópula máx-estável, seguidamente apresentada.

Definição 4.5 Uma cópula C é de valores extremos se e somente se é máx-estável, caracterizando as cópulas de valores extremos à custa das cópulas máx-estáveis, o que facilita o seu manuseamento.

$$C(u, v) = C^n(u^{1/n}, v^{1/n}) \quad \forall (u, v) \in [0, 1]^2, n \in \mathbb{N} \quad (4-41)$$

4.3.3.1 Cópula Tawn

A cópula Tawn foi definida por Tawn em 1988, e é uma extensão da cópula Gumbel. Em particular, a cópula Tawn (ou cópula logística assimétrica) é baseada nas chamadas funções de dependência de Pickands (BERNARD, 2015). A equação 4-45 apresenta a forma como se pode calcular a densidade no espaço de probabilidade usando uma função de Pickands A :

$$C(u, v) = (u, v)^{A(\omega)}, \text{ com } \omega = \frac{\ln(u)}{\ln(uv)} \quad (4-42)$$

A função pickand da cópula Tawn é escrita como:

$$A(t) = (1 - \psi_2)(1 - t) + (1 - \psi_1)t + [(\psi_1(1 - t))^\theta + (\psi_2 t)^\theta]^{\frac{1}{\theta}} \quad (4-43)$$

Então, $t \in [0, 1]$, $0 \leq \psi_1, \psi_2 \leq 1$ e $\theta \in [1, \infty)$, a cópula Tawn é na verdade uma cópula Gumbel com dois parâmetros adicionais de assimetria: ψ_1 e ψ_2 . Se estes dois parâmetros forem iguais à unidade, obtém-se a cópula Gumbel (TAILLON, 2019).

É importante ressaltar que os parâmetros ψ_1 e ψ_2 podem ser considerados como parâmetros de assimetria. Assim, podemos dividir a cópula de Tawn em

dois tipos: “Tipo 1”, com $\psi_1 = 1$ (assimetria negativa) e “Tipo 2”, $\psi_2 = 1$ (assimetria positiva).

4.3.4

Cópula Condicional

Toda abordagem da teoria de cópulas aqui apresentada, pode ser estendida às distribuições condicionais (CARVALHO, 2018). Partindo da teoria de cópulas condicionais apresentada com detalhes em Patton (2015), a partir do caso condicional do Teorema de Sklar.

Para o caso bivariado, seja o vetor aleatório $X = X_1, X_2$ e as variáveis aleatórias $X_1 | W$ e $X_2 | W$, com suas respectivas distribuições condicionais $F_1(x_1 | w)$ e $F_2(x_2 | w)$, e a distribuição conjunta condicional F de $X | W$, logo existe uma função cópula C tal que:

$$F(x_1, x_2 | w) = C(F_1(x_1 | w), F_2(x_2 | w) | w), \forall (x_1, x_2) \in \mathbb{R}^{\neq}, w \in \Omega \quad (4-44)$$

Sendo, F contínua, a função $C(\cdot)$ é única. A função cópula C , obtida por meio da transformação integral de probabilidade em cada distribuição marginal para distribuição Uniforme $[0, 1]^2$, sendo assim, representada por:

$$C(u, v | w) = F(F_1^{-1}(u | w), F_2^{-1}(v | w) | w), \forall (x_1, x_2) \in \mathbb{R}^{\neq}, w \in \Omega \quad (4-45)$$

Neste cenário, a função cópula é dada por $C(u, v) : [0, 1] \times [0, 1] \times \Omega \rightarrow [0, 1]$.

- para todo ponto (u, v) e $w \in \Omega$ tem-se que $C(u, 0 | w) = C(0, v | w) = 0$; $C(u, 1 | w) = u$ e $C(1, v | w) = v$;
- para todo u_1, u_2, v_1, v_2 , com $u_1 \leq u_2$ e $v_1 \leq v_2$ e $w \in \Omega$ tem-se que $V_c([u_1, u_2] \times [v_1, v_2] | w) = C(u_2, v_2 | w)C(u_2, v_1 | w)C(u_1, v_2 | w) + C(u_1, v_1 | w) \geq 0$.

Tem-se que a densidade da conjunta condicional F é dada pela equação a seguir:

$$f(x_1, x_2 | w) = c(u, v | w) f_1(x_1 | w) f_2(x_2 | w) \quad (4-46)$$

sendo, $c(u, v | w)$ a densidade associada a uma cópula, definida pelas derivadas parciais em relação a u e v da função $C(u, v)$, e $u = F_1(x_1 | w)$ e $v = F_2(x_2 | w)$.

Segundo Patton (2012), esta abordagem é utilizada, na análise da estrutura e comportamento de séries temporais. As cópulas condicionais apresentam inúmeras utilidades, as mesmas podem ser usadas, para modelar a dependência entre valores sucessivos de uma série temporal univariada, ou para modelar a dependência condicional de um vetor aleatório com algumas informações sobre o seu passado, dentre outros modos.

4.4

Métodos de estimação

Para o caso bivariado da implementação da Teoria de Cópulas, o conjunto $\Xi = \{(x_{1,n}, x_{2,n}), n = 1, \dots, N\}$ de N observações independentes e identicamente distribuídas de uma distribuição bivariada serão usadas para estimar as cópulas deste estudo. O método de estimação pode ser paramétrico, por meio da cópula de máxima verossimilhança e também não paramétrico, através da cópula empírica.

4.4.1

Estimação Empírica

Suponha $(X_{11}, \dots, X_{1p}), \dots, (X_{n1}, \dots, X_{np})$ amostras aleatórias. Então a cópula estimada de maneira empírica pode ser definida como:

$$C_n(u_1, \dots, u_p) = \frac{1}{n} \sum_{i=1}^n \prod_{j=1}^p 1 \{F_{j,n}(X_{ij}) \leq u_j\} \quad (4-47)$$

No qual, $F_{j,n}(x_j) := \frac{1}{n+1} \sum_{i=1}^n 1 \{X_{ij} \leq x_j\}$ para $1 \leq j \leq p$

Tratando-se do caso bivariado, tem-se a seguinte equação,

$$C_n(u, v) = \frac{1}{n} \sum_{i=1}^n 1 \{F(X) \leq u, G(Y) \leq v\} \quad (4-48)$$

4.4.2

Estimação via Método de Máxima Verossimilhança (EMV)

Tendo a função de densidade da cópula, seja $(\alpha_1, \alpha_2, \delta)$ um vetor de $k \times 1$ parâmetros e Θ o espaço paramétrico a ser utilizado, para qualquer distribuição bivariada, o logaritmo da função de verossimilhança $L_t(\alpha_1, \alpha_2, \delta)$, de uma observação única pode ser escrita como:

$$L(\alpha_1, \alpha_2, \delta; (x_{1,t}, x_{2,t})) = \ln(c(F_1(x_{1,t}, \alpha_1), (F_2(x_{2,t}, \alpha_2), \delta))) + \sum_{j=1}^2 \ln f_j(x_{j,t}; \alpha_j) \quad (4-49)$$

Sendo, c é a densidade da cópula, $f_j, j = 1, 2$ são as densidades marginais, α_1 e α_2 os vetores de parâmetros das distribuições marginais e δ o vetor de parâmetros que tem relação com a estrutura de dependência.

Dadas T observações de X_1 e X_2 , isto é, o conjunto Ξ de observações, a função de log-verossimilhança pode ser definida da seguinte maneira:

$$l(\alpha_1, \alpha_2, \delta; \Xi) = \sum_{t=1}^T L(\alpha_1, \alpha_2, \delta; (x_{1,t}, x_{2,t})) \quad (4-50)$$

Logo, tem-se que a função de log-verossimilhança é representada pela seguinte equação:

$$l(\alpha_1, \alpha_2, \delta; \Xi) = \sum_{t=1}^T \ln(c(F_1(x_{1,t}, \alpha_1), (F_2(x_{2,t}, \alpha_2), \delta))) + \sum_{t=1}^T \sum_{j=1}^2 \ln(f_j(x_{j,t}; \alpha_j)) \quad (4-51)$$

Torna-se necessário ressaltar que os estimadores de máxima verossimilhança $(\hat{\alpha}_1, \hat{\alpha}_2, \hat{\delta})$ satisfazem a equação.

$$l(\hat{\alpha}_1, \hat{\alpha}_2, \hat{\delta}; \Xi) \geq l(\alpha_1, \alpha_2, \delta; \Xi) \quad \forall (\alpha_1, \alpha_2, \delta) \in \Theta \quad (4-52)$$

4.5

Métricas de desempenho

Dentre os critérios para seleção de modelos, os critérios baseados no máximo da função de verossimilhança (MFV) são os mais utilizados, com maior ênfase o Critério de Informação de Akaike (AIC) e o Critério Bayesiano

de Schwarz (BIC).

4.5.1

Critério de Informação de Akaike (AIC)

A métrica AIC estima a quantidade de informações perdidas por um determinado modelo, ou seja, quanto menor a perda de informações de um determinado modelo, maior será sua qualidade, o que resultará em um valor de AIC menor. O Critério de Informação de Akaike aumenta conforme a soma dos quadrados dos erros (SQE) aumenta, sendo assim, valores menores de AIC são preferíveis. A métrica AIC é definida por

$$AIC = -2 \log(L_p) + 2p \quad (4-53)$$

Sendo, L_p a função de máxima verossimilhança do modelo com p representando o número de parâmetros do modelo ajustado.

4.5.2

Critério de Informação Bayesiano (BIC)

A métrica BIC assim como o AIC tem seu valor aumentado conforme soma do quadrado dos erros (SQE) aumenta, o mesmo penaliza modelos com muitas variáveis. Como modelos com muitas variáveis tendem a produzir menor SQE, porém usam mais parâmetros, a melhor escolha é balancear o ajuste com a quantidade de variáveis, onde a sua equação se dá por

$$BIC = -2 \log(L_p) + p \log n \quad (4-54)$$

No qual, L_p trata-se da função de máxima verossimilhança do modelo, p representa o número de parâmetros a serem estimados e n é o número de observações da amostra.

4.6

Testes

Após a implementação da Teoria de Cópulas, torna-se necessário a verificação da qualidade do ajuste, a mesma pode ser realizada de diversas maneiras, nessa seção será mostrado duas diferentes formas de avaliação do ajuste, utilizando-se ferramentas gráficas e testes estatísticos.

4.6.1

Testes Gráficos

Um teste gráfico pode ser definido como uma análise sobre determinados dados, sendo realizado a partir de um ou mais gráficos, se os gráficos apresentarem algum padrão desejado específico, o teste terá um resultado positivo.

4.6.2

Teste de Bondade de Ajuste

De acordo com Hofert (2018), seja $X = (X_1, \dots, X_n)$ um vetor n -dimensional com função de distribuição F , marginais F_1, \dots, F_n e com uma cópula subjacente única C . Seja também $U_i := F_i(X_i), i = 1, \dots, n$, então a distribuição conjunta de $U = (U_1, \dots, U_n)$ é C . Então, define-se um Teste de Bondade de Ajuste (Goodness-of-Fit Test, GOF) como um teste de hipóteses que verifica se a cópula C que é sugerida como a melhor para os dados é, de fato, uma cópula C , ou seja, testando a hipótese $H_0 : C \in \mathbb{C}$, como pode ser visto a seguir:

$$\begin{cases} H_0 : C \in \mathbb{C} = \{C_\theta : \theta \in O\} \\ H_1 : C \notin \mathbb{C} \end{cases} \quad (4-55)$$

Sobre H_0 , o vetor U é distribuído como C_θ para algum $\theta \in O$.

4.6.3

Teste Anderson-Darling

O teste de Anderson-Darling ou teste de adequação de Anderson-Darling trata-se de um teste estatístico utilizado para determinar se a distribuição de frequência dos dados em uma amostra se desvia de uma determinada distribui-

ção de probabilidade hipotética. O mesmo basea-se em uma transformação dos valores classificados por magnitude na amostra em uma distribuição uniforme usando a função de distribuição da distribuição de probabilidade hipotética dada. A distância entre os dados da amostra transformada e a função de distribuição da distribuição uniforme atua como uma variável de teste (ARSHAD, 2002).

Este teste valoriza às caudas da distribuição de frequência, sendo mais recomendado para distribuições assintóticas. O teste de Anderson-Darling faz uso de uma distribuição específica para o cálculo de valores críticos.

Tem-se então que as hipóteses a serem testadas são:

$$\begin{cases} H_0 : A \text{ v.a. estudada segue a distribuição específica.} \\ H_1 : A \text{ amostra não é proveniente da distribuição especificada.} \end{cases} \quad (4-56)$$

4.6.4

Teste de Kolmogorov-Smirnov

O Teste de Kolmogorov-Smirnov consiste em observar a máxima diferença absoluta entre a função de distribuição acumulada assumida para os dados, e a função de distribuição empírica dos dados. Como critério, compara-se esta diferença com um valor crítico, para um dado nível de significância. O nível de mensuração da variável deve seguir ao menos uma escala ordinal (mais comum em intervalar ou de razão). A distribuição a ser testada deve ser plenamente especificada, este teste é aplicável sem restrições, a pequenas amostras (DURBIN, 1975).

Suponha $F(x)$ a distribuição acumulada esperada (teórica) e $F_0(x)$ a distribuição acumulada observada. O objetivo é testar se a amostra é proveniente de uma população com distribuição plenamente especificada $F(x)$, ou seja, as hipóteses a serem testadas são:

$$\begin{cases} H_0 :: F_0(x) = F(x), \text{ para todo } -\infty < x < \infty \\ H_1 : F_0(x) \neq F(x), \text{ para pelo menos um valor de } x. \end{cases} \quad (4-57)$$

4.7

Estimador kernel da função de densidade

Após análise exploratória dos dados abordados neste trabalho, notou-se uma bimodalidade na variável de geração de energia eólica quando separada em períodos mensais com frequência horária, ou seja, uma base completa de 1 ano foi dividida em 12 meses, tornando-se 12 bases distintas com frequência horária. Porém, não foi possível identificar as distribuições marginais da variável de geração quando separada mensalmente, tornando assim necessária a utilização de métodos estatísticos que possibilitassem a estimação da densidade dessa variável, para que então fosse possível dar continuidade na implementação da teoria de cópulas.

O método selecionado para obter a estimação da função de densidade foi Kernel. Em Estatística, a estimativa de densidade por Kernel (EDK) trata-se de uma forma não-paramétrica para estimar a função densidade de probabilidade (FDP) de uma variável aleatória. A estimativa da densidade por Kernel é um problema fundamental de suavização de dados onde inferências sobre a população são feitas com base em uma amostra de dados finita (ROSENBLATT, 1956).

De acordo com Lucambio (2008), o estimador kernel ou núcleo tem sido bastante utilizado nos últimos tempos. Fatores que contribuíram significativamente para esta ampla utilização são a simplicidade e as boas propriedades do estimador \hat{f} .

Definição 4.6 *Seja (x_1, x_2, \dots, x_n) uma amostra de uma variável aleatória independente e identicamente distribuída com uma densidade desconhecida.*

Busca-se estimar a forma desta função f . A estimativa de densidade por Kernel pode ser feita através da seguinte equação,

$$\hat{f}_h(x) = \frac{1}{n} K_h(x - x_i) = \frac{1}{nh} \sum_{i=1}^n K\left(\frac{x-x_i}{h}\right) \quad (4-58)$$

No qual, K é um kernel, e $h > 0$ é um parâmetro de distância chamado de largura de banda, que controla a influência dos vizinhos em um kernel. Um kernel com índice h é um kernel que já foi escalonado pela largura de banda, e definido como $K_h(x) = \frac{1}{h} K(\frac{x}{h})$. Intuitivamente procura-se escolher um valor de h tão pequeno quanto a distribuição dos dados em uma amostra permitem. No entanto, há sempre um compromisso entre o viés da estimativa e sua variância.

Para os dados deste estudo, seguiu-se os seguintes passos para estimação da densidade, e obtenção dos parâmetros que serão utilizados futuramente para implementação da teoria de cópulas.

- Utilizando a função *Kde* do pacote *Ks* do *software R-Studio*, estimou-se a densidade dos dados, gerando 2.500 observações com intervalo de 0,01 com base na variável de interesse (geração de energia eólica);
- com a função *Kcde* do pacote *Ks* do *software R-Studio*, obteve-se a função de distribuição acumulada, também com base na variável de interesse;
- feito isto, utilizou-se tanto a densidade quanto a distribuição acumulada para a implementação da teoria de cópulas. Este processo será explicado mais detalhadamente na seção de resultados.

5

Análise e Resultados

Neste capítulo serão dispostas as análises e os resultados. Os dados utilizados referem-se a velocidade do vento e a geração de energia eólica no estado da Bahia no ano de 2017. Esses dados apresentam observações a cada 10 minutos, totalizando mais de 50.000 medições. Para análise considerou-se a média horária dos mesmos, ou seja, trabalhou-se com uma frequência horária. Valores atípicos resultantes de erros de medição ou armazenamento, foram classificados como velocidades do vento ou geração de energia eólica irrealistas. Seguindo a abordagem realizada em DUCA (2020), buscando não perder informações, optou-se por adotar uma estratégia para variável de velocidade do vento acima de 30 m/s , limitando-as a esse valor, porém para valores irreais de potência, a decisão foi a exclusão dos mesmos, por serem resultados incompatíveis com a variável, já para os valores faltantes utilizou-se a média.

5.1

Análises e resultados para o período anual

Para análise abordando o ano completo de 2017, considerou-se a média horária, reduzindo a base a um total de 8.723 observações. Para compreender o comportamento estocástico de cada variável, realizou-se uma breve análise descritiva dos dados. Na figura 5.1, pode-se observar o comportamento da série horária de potência em KW do parque eólico do estado da Bahia.

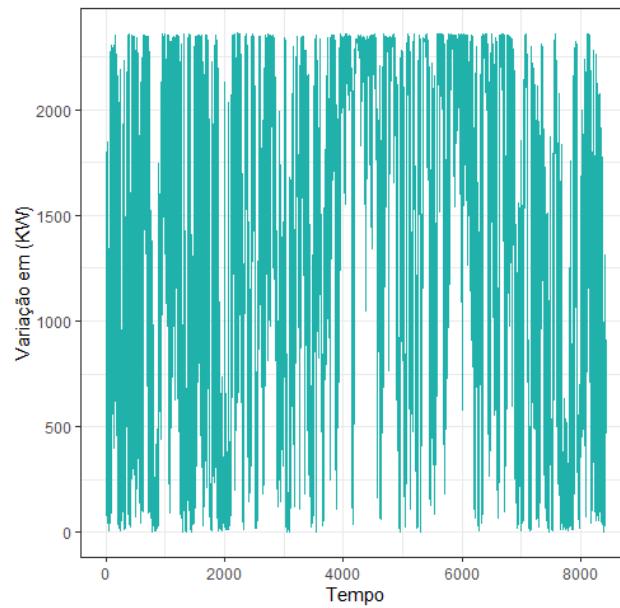


Figura 5.1: Geração de energia eólica em KW

Fonte: Elaboração própria com base nos dados analisados.

A figura 5.2 apresenta a série de velocidade do vento em m/s , nota-se que o ponto mais alto de velocidade ocorre no meio do ano entre os meses de junho e julho.

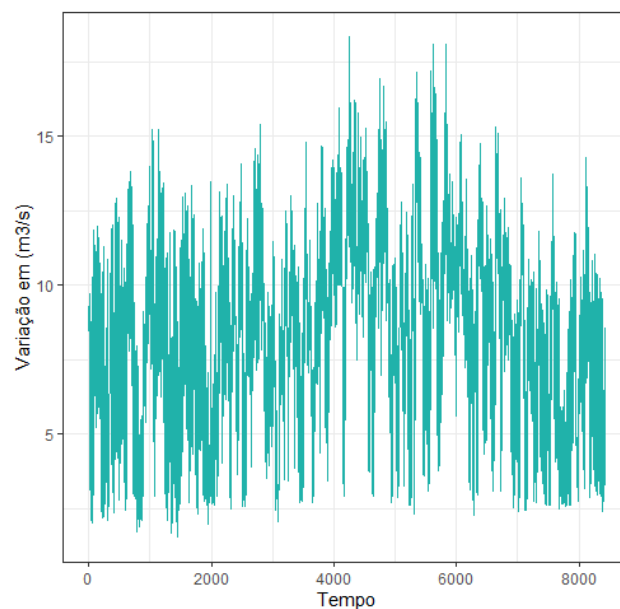


Figura 5.2: Velocidade do vento em m/s

Fonte: Elaboração própria com base nos dados analisados.

Na tabela 5.1 observa-se a estatística descritiva dos dados. Torna-se im-

portante ressaltar que nesta análise está sendo abordada a média horária, logo nota-se que em média a velocidade máxima do vento é de aproximadamente $18,32 \text{ m/s}$ nesta região e a potência máxima é de $2.361,3 \text{ KW}$.

Tabela 5.1: Estatística descritiva

Medidas	Potência (KW)	Velocidade m/s
Mínimo	0	1,53
Mediana	1.417,3	8,487
Média	1.346,6	8,356
Máximo	2.361,3	18,326

Fonte: Elaboração própria com base nos dados analisados

De acordo com a literatura a distribuição Weibull é significativamente eficiente na modelagem dos dados de velocidade do vento e geração de energia (DOS SANTOS, 2019). Foi feito um painel (Figura 5.3) com três gráficos e uma medida de correlação. No qual, o lado superior esquerdo apresenta um histograma da velocidade do vento referente, no lado inferior direito tem-se o análogo para variável de geração de energia eólica, já no lado inferior esquerdo um gráfico de dispersão entre as variáveis de velocidade e a geração, e por fim no lado superior direito, o coeficiente de correlação de Pearson.

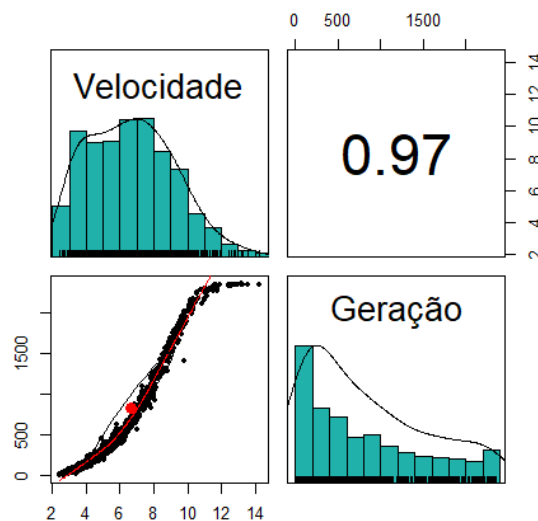


Figura 5.3: Densidade das variáveis

Fonte: Elaboração própria com base nos dados analisados.

Buscando verificar se esta distribuição Weibull realmente se ajusta de maneira eficiente aos dados, estimou-se os parâmetros pelo método de máxima verossimilhança. Isto posto, aplicou-se o teste de kolmogorov Smirnov, onde as hipóteses testadas foram:

$$\begin{cases} H_0 : \text{possivelmente os dados são extraídos da distribuição testada} \\ H_1 : \text{possivelmente os dados não são extraídos da distribuição testada} \end{cases}$$

Considerando um $\alpha = 0,01$, o teste para distribuição Weibull apresentou p-valor igual a 0,03 para a variável de geração e p-valor de 0,09 para variável velocidade, logo ao nível de significância de 1%, não rejeita-se a hipótese nula, ou seja, possivelmente os dados são extraídos da distribuição testada, neste caso a Weibull. Nota-se que o resultado vai de encontro ao apontado pela literatura.

Após a seleção das distribuições marginais correspondentes a cada variável aleatória, o próximo passo é a escolha da função cópula. Realizou-se o teste de Anderson-Darling (4.6.3), nas variáveis transformadas u e v , o resultado do

p-valor do teste apontou que, ao nível de significância de 1% não existem evidências para rejeitar a hipótese nula, logo, as variáveis transformadas seguem uma distribuição uniforme. O resultado é significativo e atende uma das premissas da Teoria de Cópulas. Feito isto, foi realizada a estimação dos parâmetros da cópula através do método de máxima verossimilhança, sendo selecionada aquela que apresentou o melhor ajuste aos dados, a ser avaliado pela análise gráfica e por outros métodos de avaliação do ajustamento de cópulas aos dados de uma amostra, como BIC e AIC.

Os resultados do BIC e do AIC, apresentaram valores significativamente baixos para a cópula bivariada *tawn type 2* quando comparada com as demais, sendo esta a cópula que melhor se ajustou a base de dados horária do ano completo de 2017. A cópula Tawn é uma extensão da cópula Gumbel com três parâmetros. Para simplificar, implementou-se duas versões da cópula Tawn com dois parâmetros. Cada tipo tem um dos parâmetros de assimetria fixados em 1, de modo que a densidade da cópula correspondente seja inclinada para a esquerda ou para a direita (em relação à diagonal principal). No caso dos dados analisados a cópula é *tawn type 2*, ou seja, com densidade inclinada para direita. Na figura 5.4, pode-se observar a densidade da cópula encontrada, a mesma apresenta o primeiro parâmetro de 12,41, e segundo de 0,99, com um tau de 0,91.

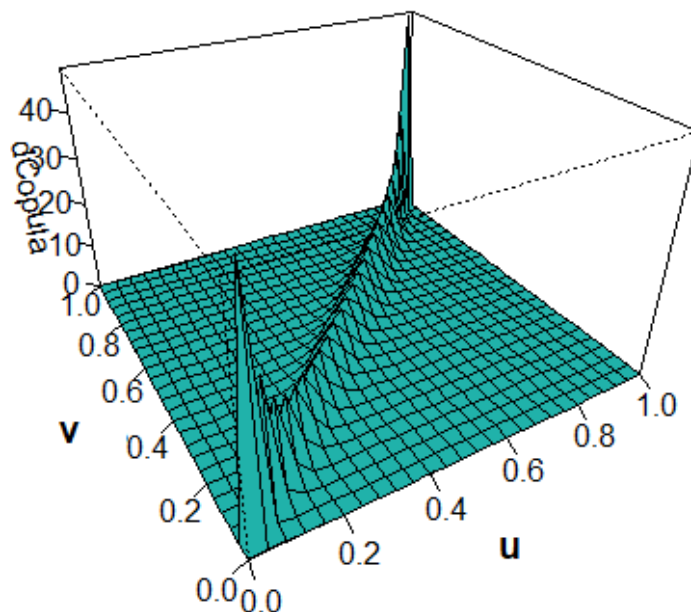


Figura 5.4: Densidade da Cópula

Fonte: Elaboração própria com base nos dados analisados.

Por conseguinte, utilizando a cópula que melhor se ajustou os dados, foram simulados 8.723 observações. A figura 5.5 apresenta o gráfico de dispersão final dos dados, sob a suposição de distribuições marginais weibull e a cópula *tawn type 2* para a estrutura de dependência, como pode-se observar, esta cópula apresenta os resultados simulados próximos às observações reais, porém nota-se que no extremo superior direito os dados simulados não acompanham totalmente os dados reais. Desta forma, buscando melhorar o ajuste optou-se por reproduzir toda implementação do método abordado para uma base de dados horária mês a mês.

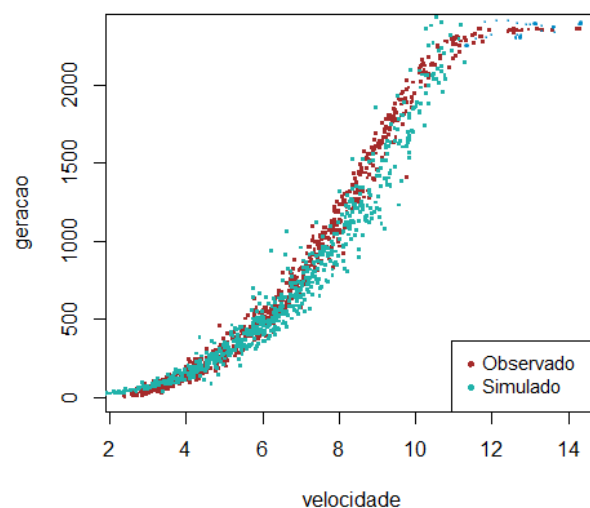


Figura 5.5: Simulação via cópulas (Potência vs Velocidade)

Fonte: Elaboração própria com base nos dados analisados.

A fim de verificar a qualidade do ajuste, realizou-se um teste gráfico para verificar o ajuste, no qual as curvas de nível da cópula *tawn type 2* são comparadas com as curvas de nível da cópula empírica que foi calculada sobre os dados desse estudo. Essa comparação pode ser observada na figura 5.6. Nota-se que as curvas de nível da cópula que foi *tawn type 2* se ajustam bem as curvas da cópula empírica, indicando um bom ajuste.

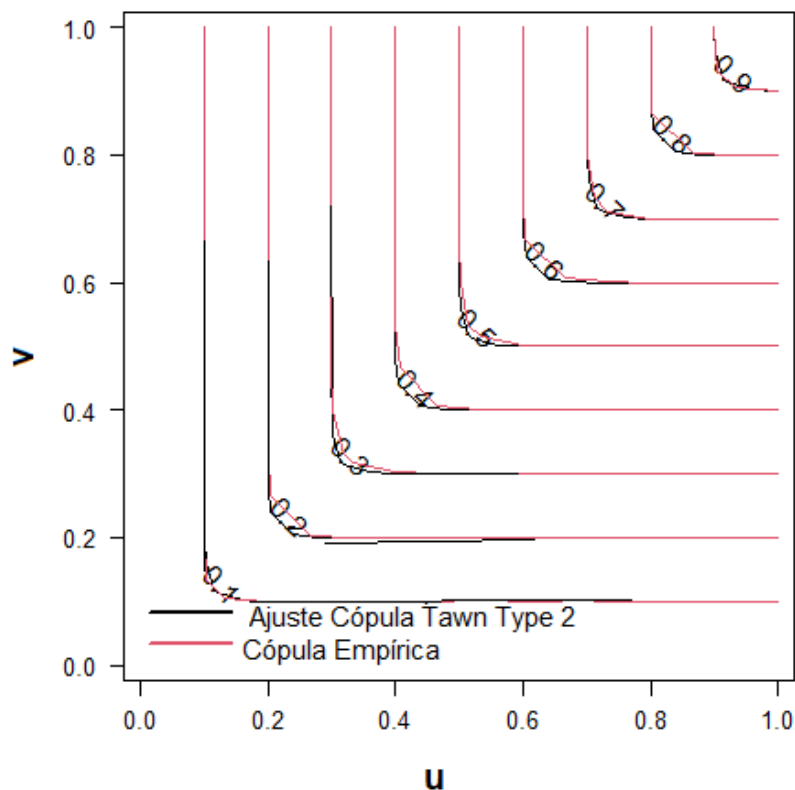


Figura 5.6: Comparação das Curvas de Nível Cópula Frank e Cópula Empírica (Jan à Jun)

Fonte: Elaboração própria com base nos dados analisados.

Para complementar a verificação do ajuste foi realizado o teste de bondade do ajuste (4.6.2), onde o resultado do p-valor foi de 0,68, ou seja, ao nível de significância de 1% não existem evidências para rejeitar a hipótese nula do teste (H_0), logo a cópula encontrada representa a estrutura de dependência entre as variáveis aqui estudadas.

5.2

Análise e resultados para o período mensal

Nesta seção serão apresentadas as análises e resultados de maneira individual para cada mês do ano de 2017. É importante ressaltar que a frequência permanece horária, porém o período consiste do dia 1 ao dia 30 ou 31 de cada mês, com exceção de fevereiro, cuja análise foi realizada do dia 1 ao dia 28.

Para melhor organização dos resultados aqui expostos, os dados separados de forma mensal foram divididos em três grupos e seus resultados serão dispostos de maneira conjunta na seguinte ordem: Tópico 1 consiste nos meses de Janeiro à Junho, o tópico 2 apresentará os resultados dos meses de julho, agosto e setembro, já o tópico 3 trará as análises dos meses de outubro, novembro e dezembro.

- Tópico 1) Análise para os meses de Janeiro à Junho

Para compreender o comportamento estocástico das variáveis aqui abordadas, realizou-se uma breve análise descritiva dos dados. Na figura 5.7, observa-se o comportamento das séries horárias de velocidade do vento em m/s e de potência em KW do parque eólico do estado da Bahia.

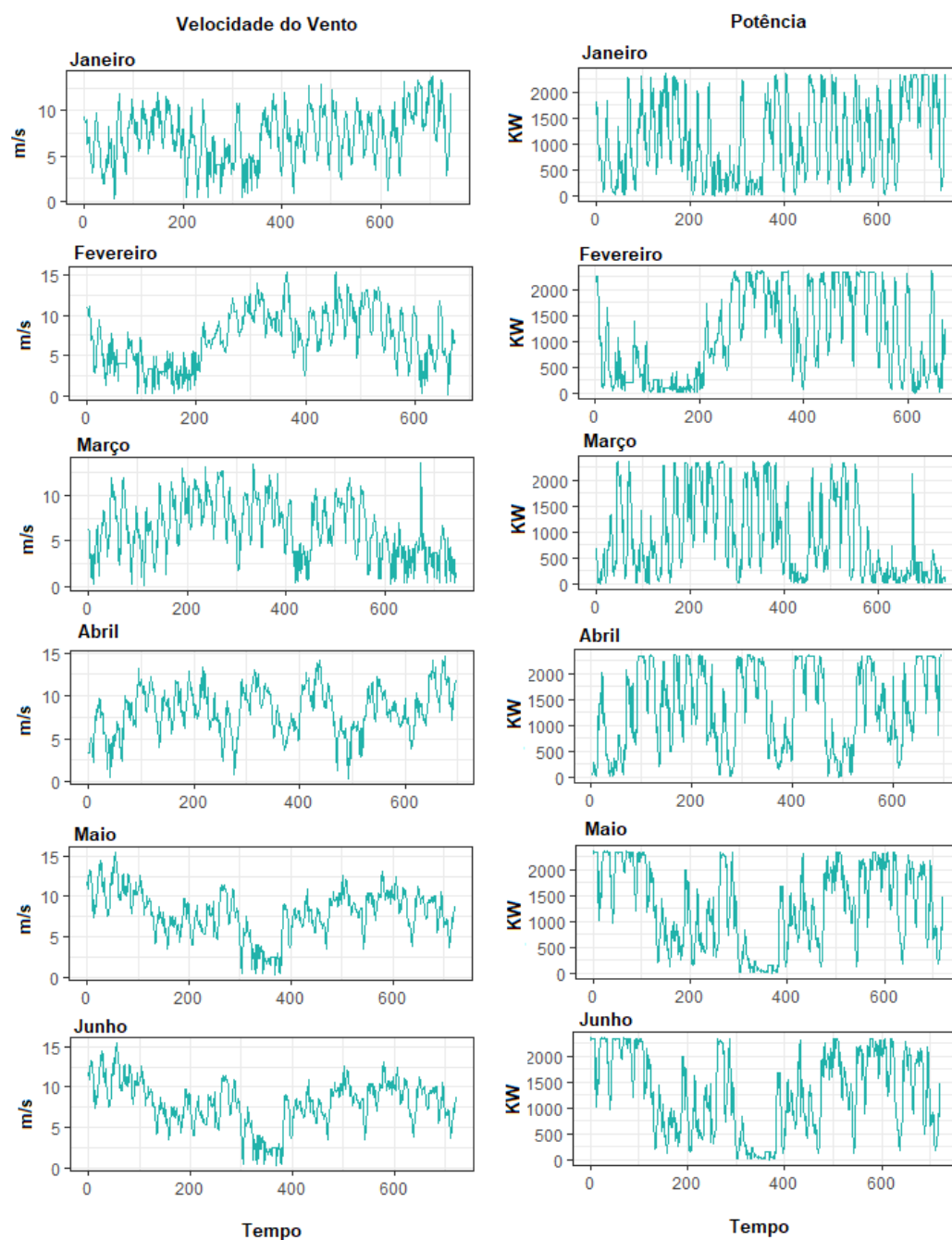


Figura 5.7: Gráfico das séries de Potência e Velocidade do vento (Jan à Jun)

Fonte: Elaboração própria com base nos dados analisados.

As tabelas 5.2 e 5.3 apresentam a estatísticas descritivas dos dados de velocidade do vento e de geração de energia eólica. Torna-se importante ressaltar que nesta análise está sendo abordada a média horária.

Tabela 5.2: Estatística descritiva - Velocidade do vento (Jan à Jun)

Meses	Mínimo	Mediana	Média	Máximo
Janeiro	0,3428	7,1621	7,2085	13,8018
Fevereiro	0,2794	7,0969	7,1881	15,2435
Março	1,451	8,371	8,388	14,791
Abril	0,4312	8,3781	8,2834	14,5824
Maio	0,3433	8,1592	7,9266	15,4109
Junho	1,6335	8,525	8,744	15,333

Fonte: Elaboração própria com base nos dados analisados

Tabela 5.3: Estatística descritiva - Potência (Jan à Jun)

Meses	Mínimo	Mediana	Média	Máximo
Janeiro	1,522	956,535	1.805,747	2.359,849
Fevereiro	0,3569	938,8580	02,3298	2.358,2575
Março	22,54	1.358,55	1.334,26	2.354,38
Abril	2,889	1.354,700	1.346,556	2.361,342
Maio	1,535	1.284,666	1.278,086	2.354,319
Junho	22,91	1.378,37	1.339,23	2.354,277

Fonte: Elaboração própria com base nos dados analisados

Buscando realizar uma análise inicial da relação entre as variáveis, foi feito um painel (Figura 5.8) com três gráficos e uma medida de correlação. No qual, o lado superior esquerdo apresenta um histograma da velocidade do vento referente a cada mês, no lado inferior direito tem-se o análogo para variável de geração de energia eólica, já no lado inferior esquerdo um gráfico de dispersão entre as variáveis de velocidade e a geração, e por fim no lado superior direito, o coeficiente de correlação de Pearson. Por meio deles, pode-se observar que as variáveis se comportam de maneira semelhante e que elas possuem uma correlação linear positiva extremamente alta.

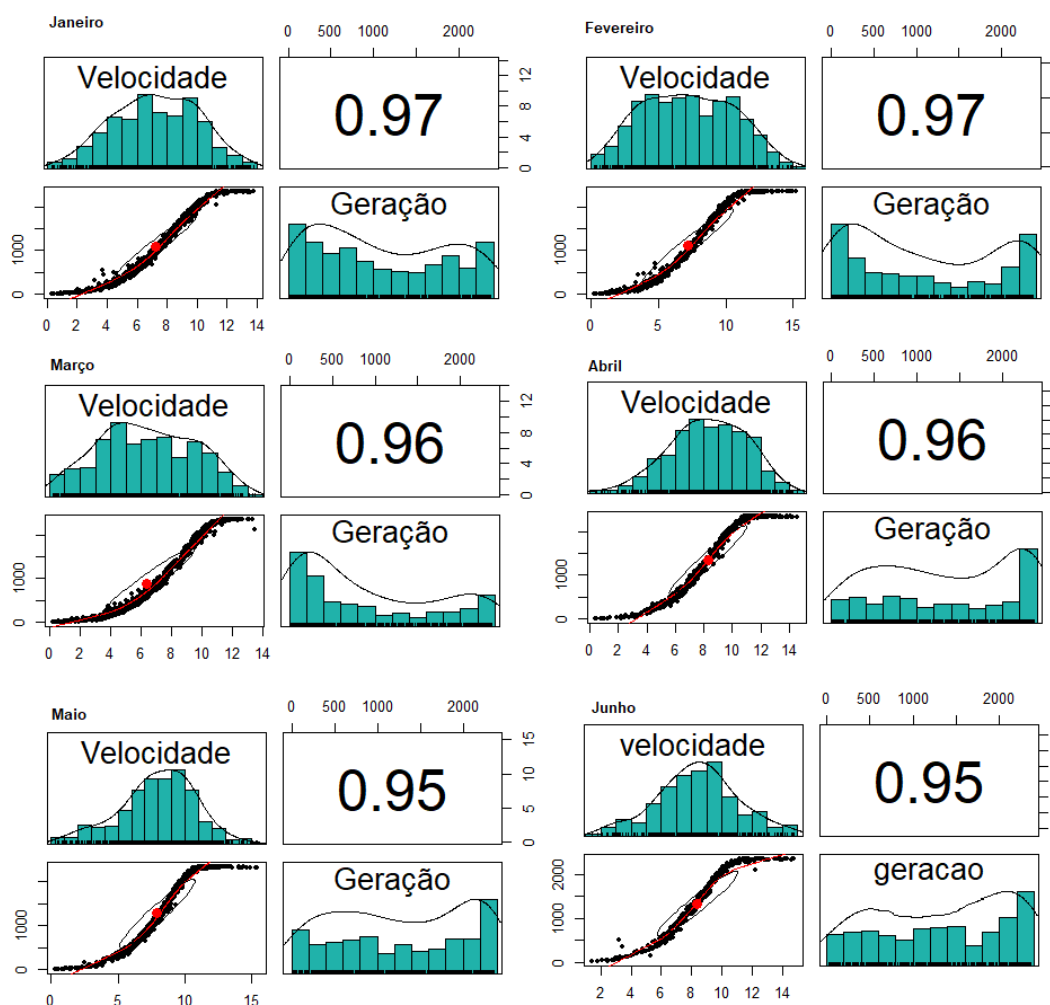


Figura 5.8: Histograma e correlação entre as variáveis (Jan à Jun)

Fonte: Elaboração própria com base nos dados analisados.

Isto posto, faz-se necessário verificar qual distribuição de probabilidade se ajusta de maneira eficiente aos dados. Iniciando a análise pelos dados de velocidade do vento, estimou-se os parâmetros pelo método de máxima verossimilhança, logo após foi aplicado o teste de kolmogorov Smirnov.

Considerando um $\alpha = 0,01$, o teste para distribuição Weibull apresentou p-valor superior ao α para a variável de velocidade, logo ao nível de significância de 1%, não rejeita-se a hipótese nula, ou seja, existem evidências de que os dados são extraídos da distribuição testada, neste caso a Weibull. Ressaltando que outras distribuições foram testadas, porém a weibull foi a que melhor se ajustou aos dados. Na tabela 5.4, pode-se observar o resultado do p-valor para

cada mês de maneira individual.

Tabela 5.4: Resultado do p-valor referente a distribuição Weibull (Jan à Jun)

Meses	p-valor
Janeiro	0,0933
Fevereiro	0,0874
Março	0,0629
Abril	0,2627
Maio	0,0756
Junho	0,5825

Após análise gráfica e testes estatísticos, devido a sua bimodalidade, não encontrou-se uma distribuição de densidade conhecida para a variável de geração de energia eólica. Buscando identificar a densidade da variável e também sua distribuição de probabilidade acumulada, aplicou-se o alisamento de Kernel (seção 4.7). Primeiramente estimou-se a densidade (definição 4.6), foram gerados 2.500 pontos com intervalo de 0,01, feito isto, utilizou-se a função de densidade encontrada para obter a distribuição acumulada das 2.500 observações. Tendo a densidade e distribuição acumulada, selecionou-se os pontos correspondentes a variável aleatória de geração. Faz-se necessário ressaltar que este processo foi realizado individualmente para cada mês. Com isto, obteve-se as informações necessárias sobre o a variável de geração para que fosse possível dar continuidade as análises e implementar a teoria de cópulas.

Partindo da seleção das distribuições correspondentes de cada variável aleatória, utilizando a distribuição weibull e a distribuição encontrada através do método de alisamento de Kernel para obter u e v conforme demonstrado na seção 4.1.1, o próximo passo é a escolha da função cópula. Torna-se importante ressaltar que o teste de Anderson-Darling (4.6.3) foi realizado nas variáveis transformadas u e v , o resultado do p-valor do teste apontou que, ao nível de significância de 1% não existem evidências para rejeitar a hipótese nula,

logo para os meses de Janeiro à Junho as variáveis transformadas seguem uma distribuição uniforme. O resultado é favorável e atende uma das premissas da Teoria de Cópulas.

Isto posto, realizou-se a estimação dos parâmetros da cópula através do método de máxima verossimilhança, sendo selecionada aquela que apresentou o melhor ajuste aos dados, a ser avaliado pela análise gráfica, por testes estatísticos e por outros métodos de avaliação do ajustamento de cópulas aos dados de uma amostra, como BIC e AIC.

Os resultados do BIC e do AIC, apresentaram valores significativamente baixos para a cópula bivariada Frank quando comparada com as demais, sendo esta a cópula que melhor se ajustou a base de dados horária para os meses de modo individual de Janeiro à Junho. A cópula Frank é simétrica na estrutura de dependência, então é uma cópula adequada quando as variáveis têm a mesma probabilidade de apresentarem valores conjuntamente baixos ou altos. Na figura 5.9 é possível observar a densidade da cópula Frank, e em seguida a tabela 5.5 apresenta o parâmetro e tau de kendal da cópula Frank para cada mês.

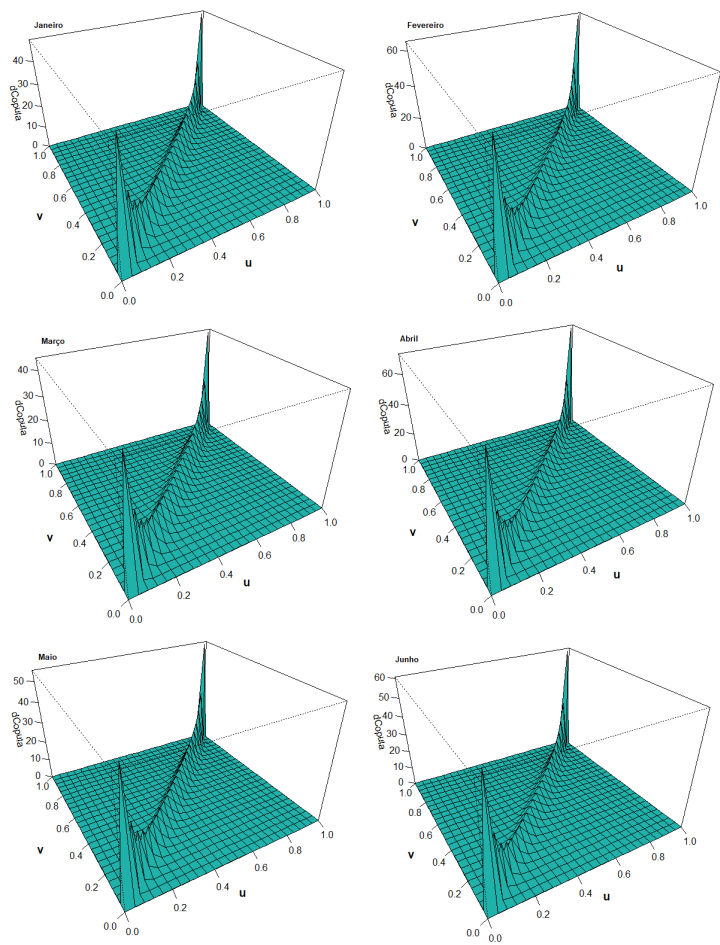


Figura 5.9: Densidade da Cópula Frank (Jan à Jun)

Fonte: Elaboração própria com base nos dados analisados.

Tabela 5.5: Parâmetro e Tau de Kendal da Cópula Frank (Jan à Jun)

Meses	Parâmetro cópula Frank (α)	Tau de Kendal (τ)
Janeiro	34,99994	0,891975
Fevereiro	34,99994	0,891975
Março	34,99993	0,891975
Abril	34,99994	0,891975
Maio	34,99993	0,891975
Junho	34,99996	0,891975

Fonte: Elaboração própria com base nos dados analisados

Utilizando a cópula Frank, foram simulados aproximadamente 730 ob-

servações para cada mês individualmente. A figura 5.10 apresenta o gráfico de dispersão final dos dados, sob a suposição de distribuição marginal weibull para a variável de velocidade do vento, a distribuição encontrada através do alisamento de Kernel para a variável de geração e a cópula Frank para a estrutura de dependência, como pode-se observar, os resultados simulados são próximos às observações reais.

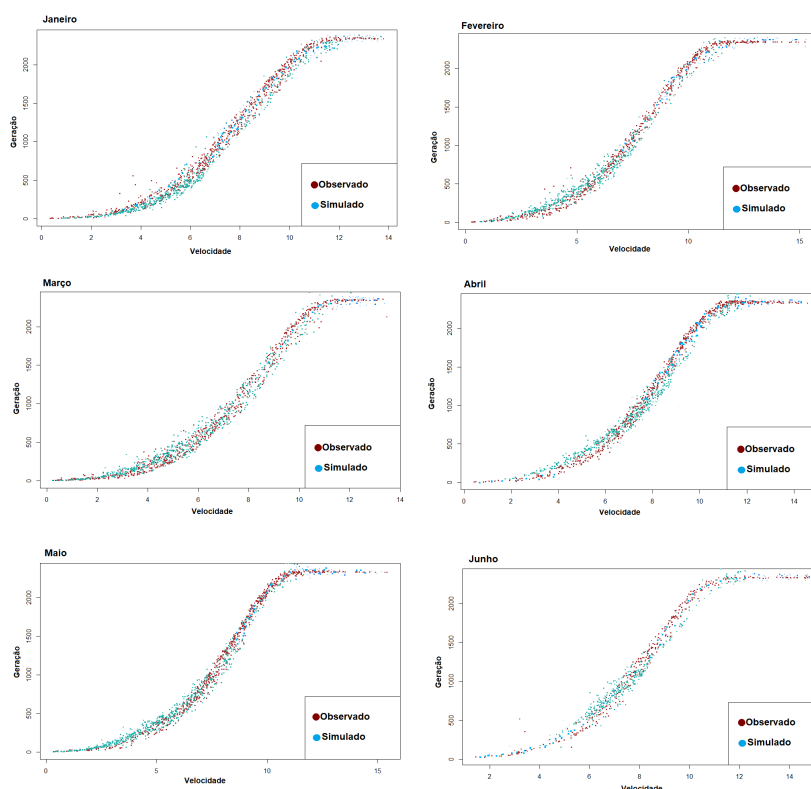


Figura 5.10: Simulação via cópulas (Jan à Jun)

Fonte: Elaboração própria com base nos dados analisados.

Buscando verificar a eficiência do ajuste, foi feito um teste gráfico, no qual as curvas de nível da cópula Frank são comparadas com as curvas de nível da cópula empírica que foi calculada sobre os dados desse estudo. Essa comparação pode ser observada na figura 5.11, onde as curvas de nível de ambas as cópulas são plotadas em um mesmo gráfico, nota-se que as curvas de nível da cópula que foi encontrada se ajustam bem as curvas da cópula empírica.

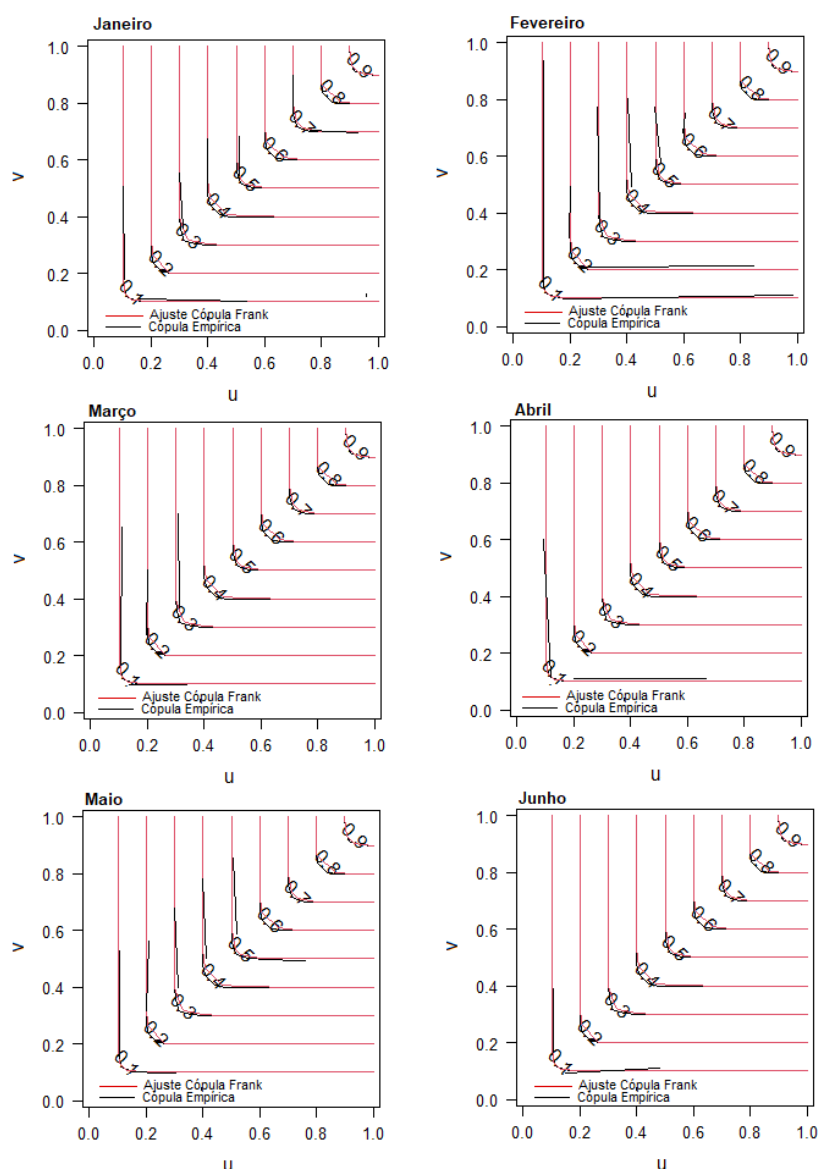


Figura 5.11: Comparação das Curvas de Nível Cópula Frank e Cópula Empírica (Jan à Jun)

Fonte: Elaboração própria com base nos dados analisados.

Para complementar a verificação do ajuste foi realizado o teste de bondade do ajuste, onde o resultado do p-valor foi maior que o alfa de 0,01 para os meses de janeiro a junho, ou seja, ao nível de significância de 1% não existem evidências para rejeitar a hipótese nula do teste, logo a cópula encontrada representa a estrutura de dependência entre as variáveis aqui estudadas.

- Tópico 2) Análise para os meses de Julho a setembro

Neste tópico serão apresentadas as análises e resultados referentes aos

meses de julho, agosto e setembro. Para compreender o comportamento das variáveis de vento e energia, pode-se observar a 5.12, nota-se que nos meses de julho e setembro a velocidade do vento permanece a maior parte do tempo entre 10 e 15 m/s e consequentemente obteve-se uma maior geração de energia eólica neste período.

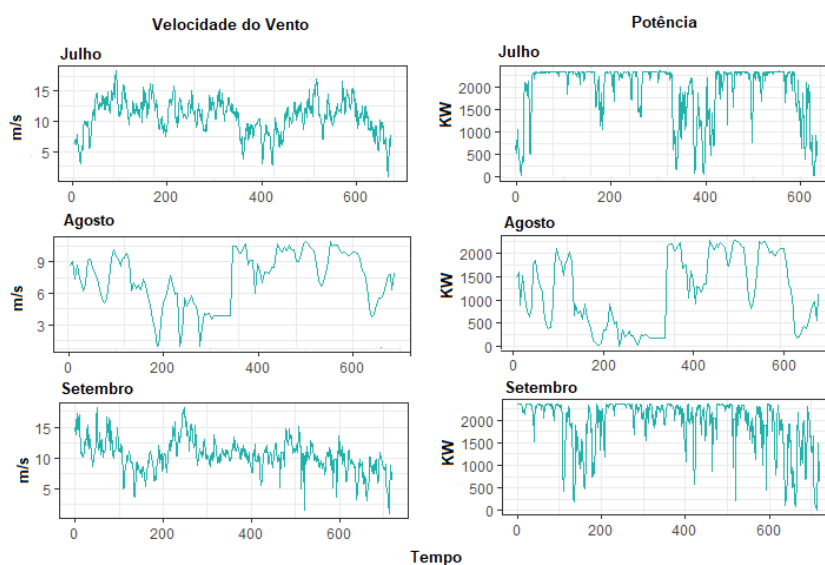


Figura 5.12: Gráfico das séries de Potência e Velocidade do vento (Jul à Set)

Fonte: Elaboração própria com base nos dados analisados.

As tabelas 5.6 e 5.7 apresentam a estatísticas descritivas dos dados de velocidade do vento e de geração de energia eólica. Nota-se que nos meses de julho e setembro apresentaram a velocidade máxima do vento acima de 18 m/s , sendo a maior entre os meses abordados nesta pesquisa.

Tabela 5.6: Estatística descritiva - Velocidade do vento (Jul à Set)

Meses	Mínimo	Mediana	Média	Máximo
Julho	1,005	11,0902	10,7970	18,3264
Agosto	0,9064	7,8174	7,3544	10,8989
Setembro	1,155	10,395	10,370	18,080

Fonte: Elaboração própria com base nos dados analisados

Tabela 5.7: Estatística descritiva - Potência (Jul à Set)

Meses	Mínimo	Mediana	Média	Máximo
Julho	14,61	2.294,62	1.930,47	2.370,17
Agosto	6,713	1.109,034	1.161,912	2.281,132
Setembro	16,07	2.173,43	1.928,26	2.368,15

Fonte: Elaboração própria com base nos dados analisados

Na figura 5.13 pode-se observar o painel contendo o histograma de velocidade do vento e o de geração de energia eólica, o gráfico de dispersão entre as variáveis e o coeficiente de correlação de Pearson. Nota-se, que as variáveis se comportam de maneira semelhante e que elas possuem uma correlação linear positiva extremamente alta.

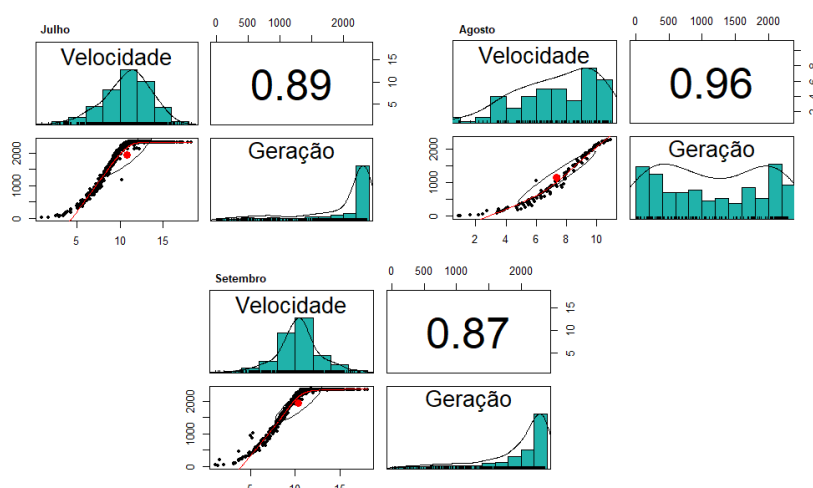


Figura 5.13: Histograma e correlação entre as variáveis (Jul à Set)

Fonte: Elaboração própria com base nos dados analisados.

Verificou-se qual distribuição de probabilidade se ajusta de maneira eficiente aos dados. Iniciando pela variável de velocidade do vento, estimou-se os parâmetros pelo método de máxima verossimilhança, depois foi realizado o teste de kolmogorov Smirnov. Foi considerado um $\alpha = 0,01$, o resultado do teste para distribuição Weibull apresentou p-valor superior ao α , logo ao nível de significância de 1%, não rejeita-se a hipótese nula, ou seja, existem

evidências de que os dados são extraídos da distribuição testada, neste caso a Weibull. Na tabela 5.8, pode-se observar o resultado do p-valor para os meses de julho a setembro.

Tabela 5.8: Resultado do p-valor referente a distribuição Weibull (Jul à Set)

Meses	p-valor
Julho	0,1348
Agosto	0,0622
Setembro	0,1121

Assim como os dados de Janeiro à Junho, não encontrou-se uma distribuição para a variável de geração de energia eólica para os dados referentes aos meses de julho, agosto e setembro. Para identificar a densidade da variável e sua distribuição de probabilidade acumulada, aplicou-se o alisamento de Kernel (seção 4.6). O método possibilitou a estimação da densidade, gerando 2.500 pontos com intervalo de 0,01, feito isto, também utilizando Kernel obteve-se a distribuição acumulada das 2.500 observações. Tendo a densidade e distribuição acumulada, selecionou-se os pontos correspondentes a variável aleatória de geração. Sendo assim, é possível dar continuidade as análises e implementar a teoria de cópulas.

Após obter as distribuições correspondentes de cada variável aleatória, utilizando a distribuição weibull e a distribuição encontrada através do método de alisamento de Kernel para obter u e v conforme demonstrado na seção 4.1.1, o próximo passo é a seleção da função cópula. Feito isto, aplicou-se o teste de Anderson-Darling (4.6.3), nas variáveis transformadas u e v , o resultado do p-valor do teste apontou que, ao nível de significância de 1% não existem evidências para rejeitar a hipótese nula, ou seja, para os meses de julho a setembro as variáveis transformadas seguem uma distribuição uniforme.

Isto posto, realizou-se a estimação dos parâmetros da cópula através do método de máxima verossimilhança, sendo selecionada aquela que apresentou

o melhor ajuste aos dados. Os resultados do BIC e do AIC, apresentaram melhores valores para a cópula bivariada Joe em comparação com as demais, sendo está a cópula que melhor se ajustou a base de dados horária para os meses de julho, agosto e setembro. A cópula Joe pertence a família de cópulas arquimedianas, a mesma apresenta dependência de cauda assimétrica, ou seja, a dependência de cauda superior e inferior não são iguais. A figura 5.14 apresenta a densidade da cópula Joe, e na tabela 5.9 pode-se o parâmetro e tau de kendal da cópula para os meses de julho à setembro.

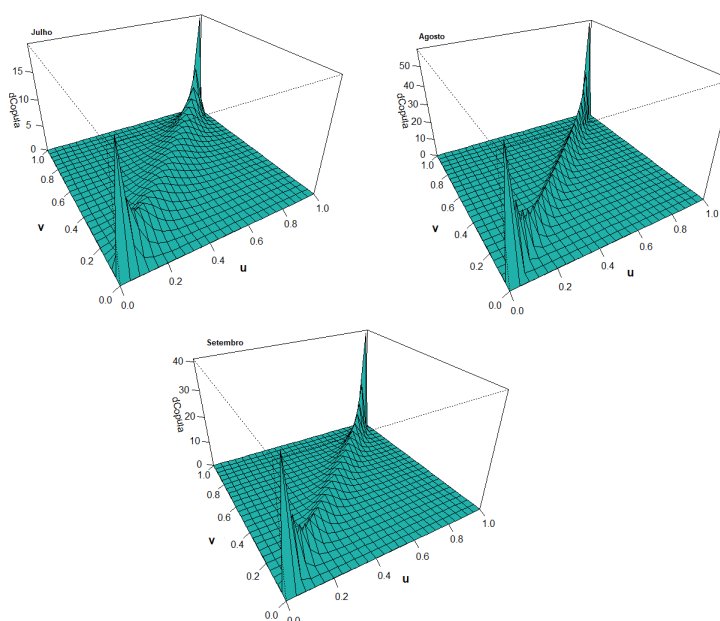


Figura 5.14: Densidade da Cópula Frank (Jul à Set)

Fonte: Elaboração própria com base nos dados analisados.

Tabela 5.9: Parâmetro e Tau de Kendal da Cópula Joe (Jul à Set)

Meses	Parâmetro cópula Joe (θ)	Tau de Kendal (τ)
Julho	9,020	0,804
Agosto	17,87	0,890
Setembro	13,75	0,866

Fonte: Elaboração própria com base nos dados analisados

Por meio da cópula Joe, que melhor se ajustou os dados no período, foram simulados aproximadamente 730 observações para cada mês individualmente.

Na figura 5.15 pode-se observar o gráfico de dispersão final dos dados, sob a suposição de distribuição marginal weibull para a variável de velocidade do vento, a distribuição encontrada através do alisamento de Kernel para a variável de geração e a cópula Joe para a estrutura de dependência, nota-se que a simulação apresenta os resultados próximos às observações reais.

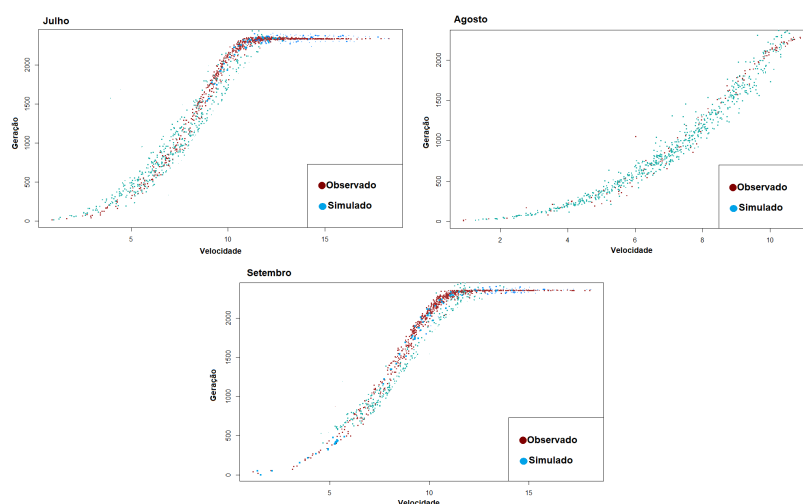


Figura 5.15: Simulação via cópulas (Jul à Set)

Fonte: Elaboração própria com base nos dados analisados.

Para verificar a qualidade do ajuste, foi feito um teste gráfico, no qual as curvas de nível da cópula Joe são comparadas com as curvas de nível da cópula empírica que foi calculada sobre os dados desse estudo (FIGURA 5.16), nota-se que as curvas de nível da cópula Joe se ajustam bem as curvas da cópula empírica.

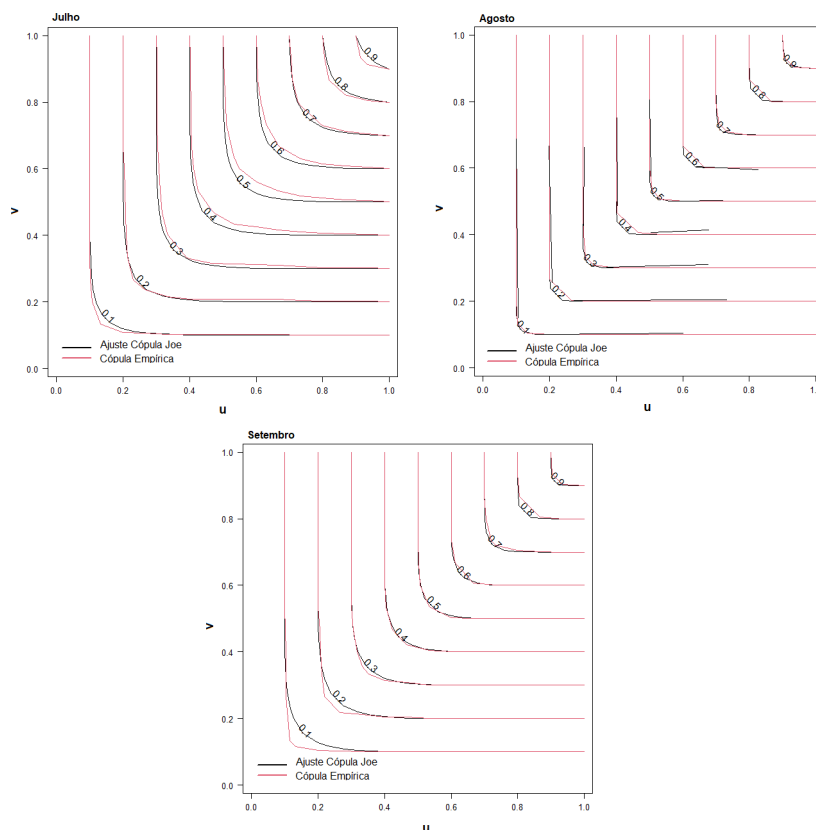


Figura 5.16: Comparação das Curvas de Nível Cópula Frank e Cópula Empírica (Jul à Set)

Fonte: Elaboração própria com base nos dados analisados.

Buscando complementar a verificação do ajuste, foi realizado o teste de bondade do ajuste, onde o resultado do p-valor foi maior que o alfa de 0,01 para os meses de julho a setembro, ou seja, ao nível de significância de 1% não existem evidências para rejeitar a hipótese nula do teste, logo a cópula encontrada representa a estrutura de dependência entre as variáveis aqui estudadas.

- Tópico 3) Análise para os meses de outubro a dezembro

Este tópico apresenta as análises e resultados referentes aos últimos meses do ano, outubro, novembro e dezembro. Para compreender o comportamento das variáveis de vento e energia, pode-se observar a figura 5.17.

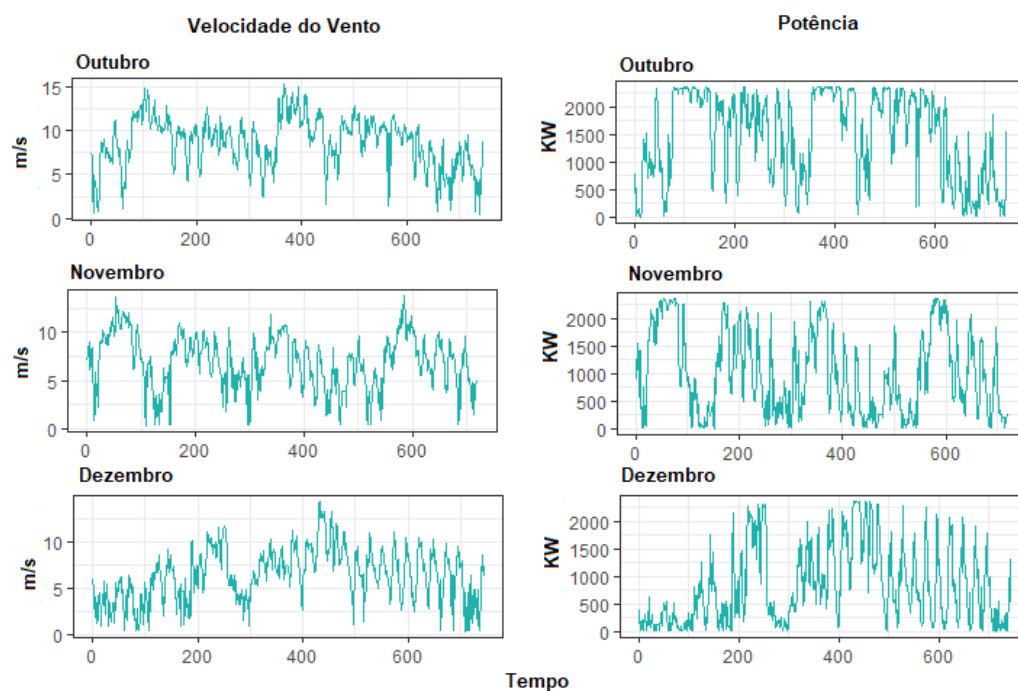


Figura 5.17: Gráfico das séries de Potência e Velocidade do vento (Out à Dez)

Fonte: Elaboração própria com base nos dados analisados.

Nas tabelas 5.10 e 5.11 pode-se observar as estatísticas descritivas dos dados de velocidade do vento e de geração de energia eólica.

Tabela 5.10: Estatística descritiva - Velocidade do vento (Out à Dez)

Meses	Mínimo	Mediana	Média	Máximo
Outubro	0,4057	7,3425	6,9998	13, 8528
Novembro	0,4086	7,1389	6,9479	13,7353
Dezembro	0,3968	6,4429	8,4050	14,2798

Fonte: Elaboração própria com base nos dados analisados

Tabela 5.11: Estatística descritiva - Potência (Out à Dez)

Meses	Mínimo	Mediana	Média	Máximo
Outubro	0,9523	885,678	1.015,555	2.358,218
Novembro	0,9457	884,587	1.013,484	2.356,361
Dezembro	0,8226	594,4874	809,0342	2.358,2619

Fonte: Elaboração própria com base nos dados analisados

A figura 5.18 apresenta um painel com três gráficos e uma medida de correlação, observa-se que as variáveis se comportam de maneira semelhante e que elas possuem uma correlação linear positiva extremamente alta.

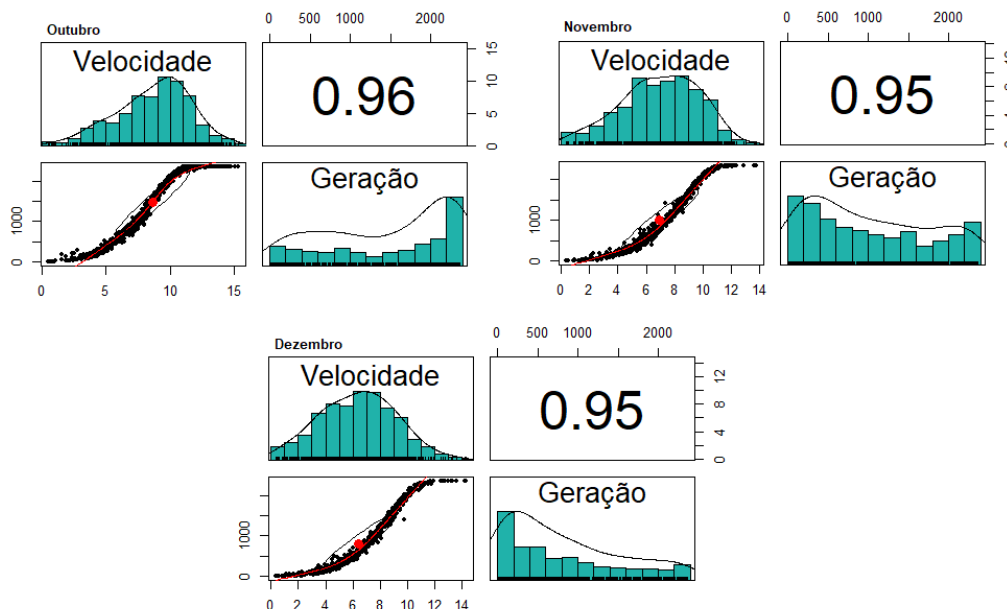


Figura 5.18: Histograma e correlação entre as variáveis (Out à Dez)

Fonte: Elaboração própria com base nos dados analisados.

Verificou-se qual distribuição de probabilidade se ajusta de maneira eficiente aos dados. Para os dados de velocidade do vento, estimou-se os parâmetros pelo método de máxima verossimilhança, feito isto, o teste de kolmogorov Smirnov foi aplicado.

Para a análise considerou-se um $\alpha = 0,01$, o teste para distribuição Weibull apresentou p-valor superior ao α para a variável de velocidade, logo ao nível de significância de 1%, não rejeita-se a hipótese nula, ou seja, existem evidências de que os dados são extraídos da distribuição testada. O resultado p-valor do teste pode ser observado na tabela 5.12.

Tabela 5.12: Resultado do p-valor referente a distribuição Weibull (Out à Dez)

Meses	p-valor
Outubro	0,0775
Novembro	0,0601
Dezembro	0,07161

Assim como nos meses supramencionados, devido a sua bimodalidade, não encontrou-se uma distribuição para a variável de geração de energia eólica, então optou-se por utilizar a metodologia de alisamento de Kernel para estimar a densidade, foram gerados 2.500 pontos com intervalo de 0,01, feito isto, também por Kernel obteve-se a distribuição acumulada das 2.500 observações. Tendo a densidade e distribuição acumulada, selecionou-se os pontos correspondentes a variável aleatória de geração.

Com isto, utilizando a distribuição weibull e a distribuição encontrada através do método de alisamento de Kernel para obter u e v conforme demonstrado na seção 4.1.1, o próximo passo foi selecionar a cópula. Foi realizado o teste de Anderson-Darling (4.6.3), nas variáveis transformadas u e v , o resultado do p-valor do teste apontou que, ao nível de significância de 1% não existem evidências para rejeitar a hipótese nula, logo para os dados referentes aos meses de outubro a dezembro as variáveis transformadas seguem uma distribuição uniforme. O resultado atende uma das premissas da Teoria de Cópuas.

Então, realizou-se a estimação dos parâmetros da cópula através do método de máxima verossimilhança, sendo selecionada aquela que apresentou o melhor ajuste aos dados. Os resultados do BIC e do AIC, apontaram a cópula bivariada Frank como a que melhor se ajustou aos dados para os dados de outubro a dezembro. A figura 5.19 apresenta a densidade da cópula Frank, e em seguida pode-se observar os parâmetro e tau de kendal da cópula Frank para cada mês na tabela 5.13.

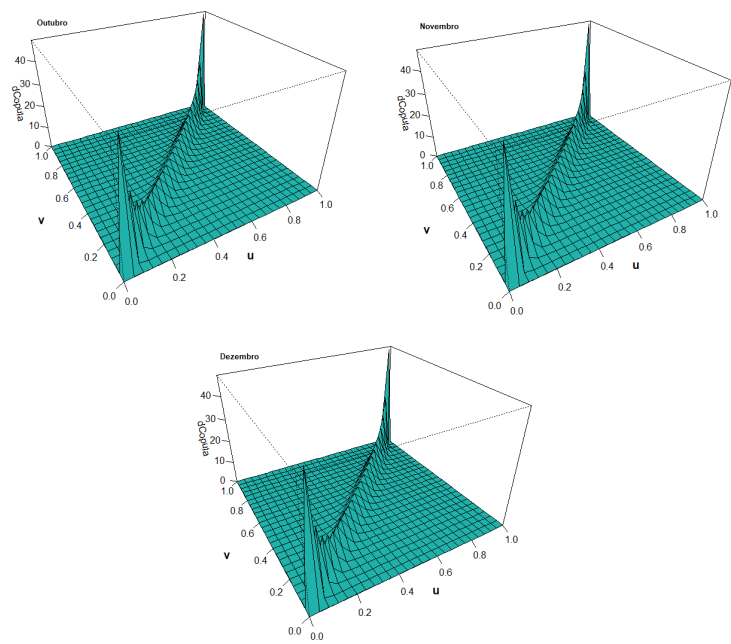


Figura 5.19: Densidade da Cópula Frank (Out à Dez)

Fonte: Elaboração própria com base nos dados analisados.

Tabela 5.13: Parâmetro e Tau de Kendal da Cópula Frank (Out à Dez)

Meses	Parâmetro cópula Frank (α)	Tau de Kendal (τ)
Outubro	34,99994	0,891975
Novembro	34,99993	0,891975
Dezembro	34,99993	0,891975

Fonte: Elaboração própria com base nos dados analisados

Utilizando a cópula Frank, foram simulados aproximadamente 730 observações. Na figura 5.20 pode-se observar o gráfico de dispersão final dos dados, sob a suposição de distribuição marginal weibull para a variável de velocidade do vento, a distribuição encontrada através do alisamento de Kernel para a variável de geração e a cópula Frank para a estrutura de dependência. Ressalta-se que os resultados simulados são próximos às observações reais.

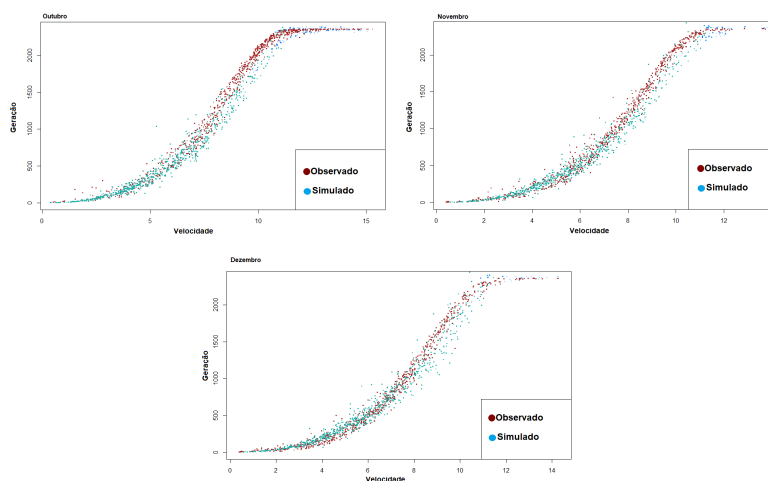


Figura 5.20: Simulação via cópulas (Out à Dez)

Fonte: Elaboração própria com base nos dados analisados.

Realizou-se um teste gráfico para verificar o ajuste, no qual as curvas de nível da cópula Frank são comparadas com as curvas de nível da cópula empírica que foi calculada sobre os dados desse estudo. Essa comparação pode ser observada na figura 5.21. As curvas de nível da cópula que foi Frank se ajustam bem as curvas da cópula empírica.

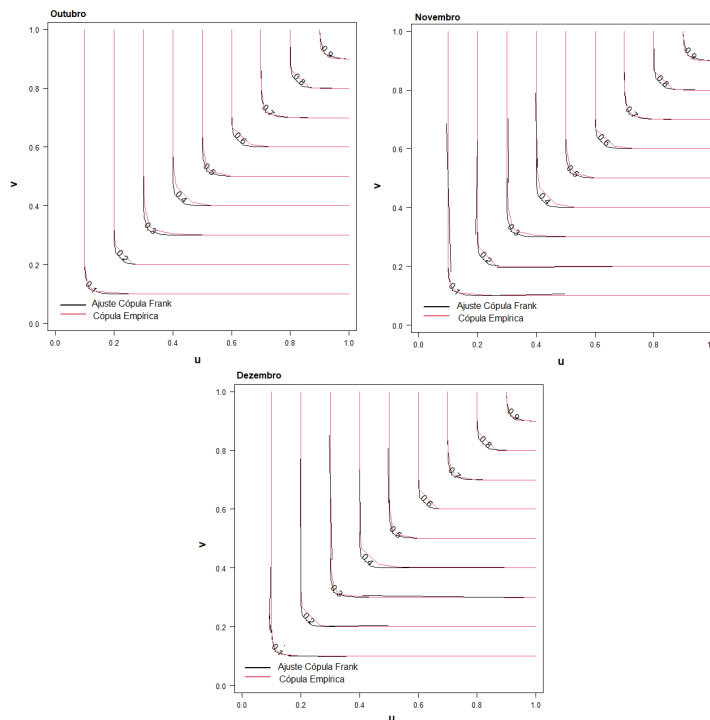


Figura 5.21: Comparação das Curvas de Nível Cópula Frank e Cópula Empírica (Jan à Jun)

Fonte: Elaboração própria com base nos dados analisados.

Complementando a verificação do ajuste, foi realizado o teste de bondade do ajuste, onde o resultado do p-valor foi maior que o alfa de 0,01 para os meses de outubro a dezembro, ou seja, ao nível de significância de 1% não existem evidências para rejeitar a hipótese nula do teste, logo a cópula encontrada representa a estrutura de dependência entre as variáveis aqui estudadas.

5.3

Probabilidade de Ocorrência dos Intervalos Simulados

Nesta seção serão apresentadas as probabilidades de ocorrências de alguns intervalos gerados, é importante ressaltar que é possível calcular a probabilidade de inúmeros e diversos intervalos, ou seja, após a seleção da função cópula e a simulação dos intervalos é possível obter toda e qualquer probabilidade associada a ele.

Para melhor organização e visualização, as probabilidades estão expostas em tabelas, referentes ao ano completo de 2017 e a cada mês de maneira

individual. O intervalo selecionado para potência foi de 100 k/w à 2.500 k/w , sendo exposto 100, 500, ..., 2500. Já para a variável de velocidade, na tabela representada por vel , o intervalo selecionado foi de: $2 \leq vel < 5$, $5 \leq vel < 8$, $8 \leq vel < 11$ e $11 \leq vel \leq 14$, ressaltando que qualquer outro intervalo poderia ser escolhido, tanto para a variável de geração quanto para velocidade do vento.

Para iniciar o cálculo das probabilidades, utilizou-se os vetores u e v gerados através da cópula que melhor representa a estrutura de dependência das variáveis no período abordado, ou seja, para o ano completo utilizou-se a cópula *tawn type 2*, já para os meses de janeiro a junho e outubro a dezembro, foi utilizada cópula Frank, e para os meses de julho a setembro a cópula de referência foi a Joe.

- Probabilidade dos intervalos simulados (Ano Completo)

A tabela 5.14 é referente ao ano completo, e apresenta as probabilidades de ocorrência dos intervalos selecionados. Nota-se que quando a velocidade do vento for maior ou igual a 2 e menor que 5 m/s , a probabilidade de se obter uma geração acima de 100 k/w é de 67,5%, e acima de 500 k/w é 0. Porém quando a velocidade do vento for maior ou igual a 11 e menor ou igual a 14 m/s a probabilidade de obter geração acima de 100 é 100 %.

Tabela 5.14: Probabilidade de ocorrência dos intervalos (Ano completo)

Potência	$2 \leq vel < 5(\%)$	$5 \leq vel < 8(\%)$	$8 \leq vel < 11(\%)$	$11 \leq vel \leq 14(\%)$
≥ 100	67,5	99,9	100	100
≥ 500	0	59,79	100	100
≥ 1000	0	4,37	93,78	100
≥ 1500	0	0,02	43,92	99,7
≥ 2000	0	0	13,89	98,60

Fonte: Elaboração própria com base nos dados analisados

- Probabilidade dos intervalos simulados (Janeiro)

Nota-se que no mês de janeiro se a velocidade do vento estiver acima de 5 m/s a probabilidade de geração acima de 100 k/w é 100%. Pode-se observar que se a velocidade do vento estiver entre $11 \leq vel \leq 14(\text{m/s})$ a probabilidade de se obter uma geração maior ou igual 2000 k/w é de 99,35% (Tabela 5.15).

Tabela 5.15: Probabilidade de ocorrência dos intervalos (Janeiro)

Potência	$2 \leq vel < 5(\%)$	$5 \leq vel < 8(\%)$	$8 \leq vel < 11(\%)$	$11 \leq vel \leq 14(\%)$
≥ 100	68,28	100	100	100
≥ 500	0,08	61,04	99,96	100
≥ 1000	0	4,37	93,2	100
≥ 1500	0	0,02	43,49	100
≥ 2000	0	0	13,94	99,35

Fonte: Elaboração própria com base nos dados analisados

- Probabilidade dos intervalos simulados (Fevereiro)

No mês de fevereiro, pode-se observar que mesmo com baixas probabilidades, quando a velocidade do vento for maior ou igual a 2 e menor que 5 m/s sempre existe uma probabilidade não nula de geração (Tabela 5.16).

Tabela 5.16: Probabilidade de ocorrência dos intervalos (Fevereiro)

Potência	$2 \leq vel < 5(\%)$	$5 \leq vel < 8(\%)$	$8 \leq vel < 11(\%)$	$11 \leq vel \leq 14(\%)$
≥ 100	86,5	99,97	100	100
≥ 500	1,96	80,13	99,51	100
≥ 1000	0,33	13,12	95,94	99,29
≥ 1500	0,08	0,7	51,83	98,89
≥ 2000	0,06	0,15	13,13	96,26

Fonte: Elaboração própria com base nos dados analisados

- Probabilidade dos intervalos simulados (Março)

A tabela 5.17, apresenta as probabilidades associadas aos intervalos referentes ao mês de março. Nota-se que quando a velocidade do vento

encontra-se entre 11 e 14 m/s a probabilidade de se obter uma geração acima de 1500 K/w é de 99,85%.

Tabela 5.17: Probabilidade de ocorrência dos intervalos (Março)

Potência	$2 \leq vel < 5(\%)$	$5 \leq vel < 8(\%)$	$8 \leq vel < 11(\%)$	$11 \leq vel \leq 14(\%)$
≥ 100	76,06	99,97	100	100
≥ 500	0,92	73,92	99,99	100
≥ 1000	0	12,88	95,66	100
≥ 1500	0	0,37	55,22	99,85
≥ 2000	0	0,14	17,98	97,75

Fonte: Elaboração própria com base nos dados analisados

- Probabilidade dos intervalos simulados (Abril)

Na tabela 5.18, pode-se observar as probabilidades referentes ao mês de abril. Assim como o mês de fevereiro mesmo com baixa probabilidade pode-se obter geração quando a velocidade do vento for maior ou igual a 2 e menor que 5 m/s .

Tabela 5.18: Probabilidade de ocorrência dos intervalos (Abril)

Potência	$2 \leq vel < 5(\%)$	$5 \leq vel < 8(\%)$	$8 \leq vel < 11(\%)$	$11 \leq vel \leq 14(\%)$
≥ 100	91,23	100	100	100
≥ 500	1,22	83,89	99,82	99,93
≥ 1000	0,26	13,73	97,66	99,57
≥ 1500	0,24	0,44	51,80	99,15
≥ 2000	0,17	0,18	11,06	99,05

Fonte: Elaboração própria com base nos dados analisados

As tabelas a seguir, representam as probabilidades de ocorrência dos intervalos simulados em cada mês de maneira individual.

- Probabilidade dos intervalos simulados (Maio)

Tabela 5.19: Probabilidade de ocorrência dos intervalos (Maio)

Potência	$2 \leq vel < 5(\%)$	$5 \leq vel < 8(\%)$	$8 \leq vel < 11(\%)$	$11 \leq vel \leq 14(\%)$
≥ 100	75,82	99,97	1	100
≥ 500	0,19	73,34	1	100
≥ 1000	0	11,54	97,9	100
≥ 1500	0	0,12	54, 25	99,85
≥ 2000	0	0	16,97	99,40

Fonte: Elaboração própria com base nos dados analisados

- Probabilidade dos intervalos simulados (Junho)

Tabela 5.20: Probabilidade de ocorrência dos intervalos (Junho)

Potência	$2 \leq vel < 5(\%)$	$5 \leq vel < 8(\%)$	$8 \leq vel < 11(\%)$	$11 \leq vel \leq 14(\%)$
≥ 100	94,55	100	100	100
≥ 500	2,68	89,89	99,97	100
≥ 1000	0,29	17,38	97,55	99,80
≥ 1500	0,1	0,41	48,66	99,46
≥ 2000	0,1	0,15	6,62	95,08

Fonte: Elaboração própria com base nos dados analisados

- Probabilidade dos intervalos simulados (Julho)

Tabela 5.21: Probabilidade de ocorrência dos intervalos (Julho)

Potência	$2 \leq vel < 5(\%)$	$5 \leq vel < 8(\%)$	$8 \leq vel < 11(\%)$	$11 \leq vel \leq 14(\%)$
≥ 100	94,5	99,94	100	100
≥ 500	13,76	86,58	99,36	99,93
≥ 1000	2,75	25,87	89,5	97,34
≥ 1500	1,65	4,68	47,25	93,61
≥ 2000	0,92	1,28	10,45	78,63

Fonte: Elaboração própria com base nos dados analisados

- Probabilidade dos intervalos simulados (Agosto)

Tabela 5.22: Probabilidade de ocorrência dos intervalos (Agosto)

Potência	$2 \leq vel < 5(\%)$	$5 \leq vel < 8(\%)$	$8 \leq vel < 11(\%)$	$11 \leq vel \leq 14(\%)$
≥ 100	84,55	100	100	100
≥ 500	0,18	80,37	99,97	100
≥ 1000	0	18,15	99,63	100
≥ 1500	0	0,07	65,29	100
≥ 2000	0	0	26,5	99,84

Fonte: Elaboração própria com base nos dados analisados

- Probabilidade dos intervalos simulados (Setembro)

Tabela 5.23: Probabilidade de ocorrência dos intervalos (Setembro)

Potência	$2 \leq vel < 5(\%)$	$5 \leq vel < 8(\%)$	$8 \leq vel < 11(\%)$	$11 \leq vel \leq 14(\%)$
≥ 100	100	100	100	100
≥ 500	44,12	99,88	100	100
≥ 1000	1,18	25,25	97,78	99,88
≥ 1500	1,16	1,62	51,49	99,08
≥ 2000	0,59	0,61	8,88	90,37

Fonte: Elaboração própria com base nos dados analisados

- Probabilidade dos intervalos simulados (Outubro)

Tabela 5.24: Probabilidade de ocorrência dos intervalos (Outubro)

Potência	$2 \leq vel < 5(\%)$	$5 \leq vel < 8(\%)$	$8 \leq vel < 11(\%)$	$11 \leq vel \leq 14(\%)$
≥ 100	74,36	99,97	100	100
≥ 500	0,64	66,73	99,9	100
≥ 1000	0	8,1	93,66	100
≥ 1500	0	0,13	48,47	100
≥ 2000	0	0	14,49	98,46

Fonte: Elaboração própria com base nos dados analisados

- Probabilidade dos intervalos simulados (Novembro)

Tabela 5.25: Probabilidade de ocorrência dos intervalos (Novembro)

Potência	$2 \leq vel < 5(\%)$	$5 \leq vel < 8(\%)$	$8 \leq vel < 11(\%)$	$11 \leq vel \leq 14(\%)$
≥ 100	72,31	100	100	100
≥ 500	0,2	68,02	100	100
≥ 1000	0	7,63	95,57	100
≥ 1500	0	0,1	45,07	100
≥ 2000	0	0	13,32	98,67

Fonte: Elaboração própria com base nos dados analisados

- Probabilidade dos intervalos simulados (Dezembro)

Tabela 5.26: Probabilidade de ocorrência dos intervalos (Dezembro)

Potência	$2 \leq vel < 5(\%)$	$5 \leq vel < 8(\%)$	$8 \leq vel < 11(\%)$	$11 \leq vel \leq 14(\%)$
≥ 100	72,27	100	100	100
≥ 500	0,23	67,99	100	100
≥ 1000	0	7,71	95,39	100
≥ 1500	0	0,01	41,12	100
≥ 2000	0	0	13,31	98,67

Fonte: Elaboração própria com base nos dados analisados

6

Conclusão e trabalhos futuros

O presente trabalho apresentou uma proposta de implementação da teoria de cópula, a fim de estudar e evidenciar a estrutura de dependência entre as variáveis de velocidade do vento e a geração de energia eólica. A pesquisa objetivou gerar simulações e calcular as probabilidades de ocorrência dos intervalos por meio da função cópula. Os dados eram provenientes de um parque eólico no estado da Bahia e o período foi de 1 ano completo, sendo este o ano de 2017, a frequência utilizada para as análises foi a média horaria, totalizando aproximadamente 8.723 observações.

Primeiramente iniciou-se a análise para o ano completo, e através de testes gráficos e do Teste Kolmogorov-Smirnov, identificou-se a distribuição marginal Weibull como a que melhor se ajustou aos dados, tanto para a variável de velocidade do vento quanto para a variável de geração. Feito isto estimou-se através do método de máxima verossimilhança os parâmetros da cópula que melhor se ajusta aos dados com a informação de que as distribuições marginais são provenientes de uma distribuição weibull. Seguindo o critério de avaliação BIC e AIC a cópula Tawn Type 2 é a que melhor representa a estrutura de dependência dos dados analisados, foi realizado o teste de bondade do ajuste e o mesmo confirmou que a cópula Tawn se ajusta bem aos dados, logo após aplicou-se o teste de Anderson Darling nas variáveis transformadas, a fim de verificar se as mesmas atendiam a premissa de uniformidade, o resultado do teste evidenciou que as variáveis transformadas pertenciam a uma distribuição uniforme. Isto posto, gerou-se observações utilizando a cópula Tawn Type 2, depois foram selecionados alguns intervalos e calculou-se a probabilidade de ocorrência, nota-se que quando a velocidade do vento for maior ou igual a 2 e menor que 5 m/s , a probabilidade de se obter uma geração acima de 100 k/w é de 67,5%, e acima de 500 k/w é 0, já quando o cenário de velocidade do

vento for maior ou igual a 11 e menor ou igual a 14 m/s a probabilidade de se obter geração acima de 100 é 100 %, de se obter geração acima de 2000 é 98,6%. Com isto, pode-se observar a probabilidade de ocorrência de inúmeros e diversos intervalos.

Após, a implementação da teoria de cópulas para os dados referentes ao ano completo, buscou-se analisar cada mês de maneira individual. Feita toda a análise exposta nesta pesquisa para cada mês com uma frequência horária, notou-se que a variável de velocidade do vento também segue uma distribuição weibull para os dados no período mensal, porém os dados de geração quando analisados mensalmente apresentam uma bimodalidade, então utilizou-se o método de alisamento de Kernel para estimar a função de densidade e a distribuição acumulada desta variável, isto foi feito para todos os meses. O resultado da estimação da cópula utilizando a distribuição marginal Weibull e a distribuição selecionada através do método de kernel, apontou a cópula Frank e a cópula Joe (cópulas da família arquimediana) como as que melhor se ajustaram aos dados.

Os meses de janeiro a junho, e de outubro a dezembro apontam a cópula Frank como a que melhor representa sua estrutura de dependência, todos os testes supramencionados foram aplicados e esta foi a cópula que apresentou o melhor resultado. Foram simuladas aproximadamente 730 observações para cada mês, e as probabilidades de ocorrência dos intervalos selecionados estão expostas na seção de resultados deste trabalho. Um resultado muito relevante que torna-se importante ressaltar é a similaridade da estrutura de dependência entre as variáveis de velocidade do vento e geração de energia eólica no período de janeiro a junho, visto que o parâmetro ajustado permaneceu constante e a família de cópula é a mesma para todos esses 6 meses. E além disso o coeficiente tau de kendall que mede a concordância e discordância dos vetores u e v estimado foi idêntico, pretende-se analisar esses resultado em trabalhos futuros, pois faz-se necessário um base de dados com um período

mais extenso. Para os meses de julho a setembro a cópula que melhor se ajustou aos dados foi a cópula Joe, a mesma também faz parte da família das cópulas arquimedianas, porém a cópula Joe contempla a assimetria, ou seja, a dependência de cauda superior e inferior não são iguais. Através dessa cópula foram geradas inúmeras simulações e calculada a probabilidade de ocorrências dos intervalos selecionados. É importante ressaltar que os meses de julho a setembro apresentam o melhor potencial eólico entre os meses analisados, nestes meses os ventos são mais constantes, aumentando a geração de energia, isto pode ser notado nas probabilidades de ocorrência dos intervalos, para o mês de julho existe probabilidade de ocorrência de geração eólica para qualquer cenário simulado, o mesmo ocorre para o mês de setembro. Para as análises mensais também foram feitos os testes gráficos, os testes de bondade do ajuste e o teste de Anderson Darling, os resultados indicam um bom ajuste.

A abordagem da aplicação da teoria de cópulas aqui apresentada, auxilia na gestão da geração de energia eólica, proporcionando menos custos de operação para as usinas que utilizam métodos no qual é possível identificar o comportamento das variáveis de geração e velocidade do vento de maneira conjunta, a análise permite observar padrões relevantes na série, e a simulação de intervalos apresentando suas probabilidades de ocorrência auxilia no processo de decisão, mitigando incertezas. O uso da metodologia adequada possibilita melhorias nas gestões, proporcionando inúmeros benefícios sustentáveis como a redução de CO^2 . O gerenciamento eficaz de um parque eólico além de propiciar uma geração mais eficiente, contribui significativamente para um mundo mais sustentável.

Appendices

Potência	3 ≤ Vel < 4	4 ≤ Vel < 5	5 ≤ Vel < 6	6 ≤ Vel < 7	7 ≤ Vel < 8	8 ≤ Vel < 9	9 ≤ Vel < 10	10 ≤ Vel < 11	11 ≤ Vel < 12	12 ≤ Vel < 13	13 ≤ Vel < 14
≥ 100	60,0%	99,7%	100,0%	100,0%	100,0%	100,0%	100,0%	100,0%	100,0%	100,0%	100,0%
≥ 200	1,5%	60,4%	99,1%	99,8%	100,0%	100,0%	100,0%	100,0%	100,0%	100,0%	100,0%
≥ 300	1,2%	8,6%	80,0%	99,4%	99,9%	100,0%	100,0%	100,0%	100,0%	100,0%	100,0%
≥ 400	1,2%	0,7%	33,9%	96,9%	99,8%	100,0%	100,0%	100,0%	100,0%	100,0%	100,0%
≥ 500	0,0%	0,0%	6,1%	74,7%	99,4%	100,0%	100,0%	100,0%	100,0%	100,0%	100,0%
≥ 600	0,0%	0,0%	0,8%	39,4%	97,5%	100,0%	100,0%	100,0%	100,0%	100,0%	100,0%
≥ 700	0,0%	0,0%	0,2%	12,1%	85,5%	99,5%	100,0%	100,0%	100,0%	100,0%	100,0%
≥ 800	0,0%	0,0%	0,2%	2,9%	60,8%	98,3%	100,0%	100,0%	100,0%	100,0%	100,0%
≥ 900	0,0%	0,0%	0,1%	0,7%	34,0%	95,4%	99,8%	100,0%	100,0%	100,0%	100,0%
≥ 1000	0,0%	0,0%	0,0%	0,2%	13,5%	85,8%	99,8%	100,0%	100,0%	100,0%	100,0%
≥ 1100	0,0%	0,0%	0,0%	0,1%	4,0%	67,5%	97,1%	100,0%	100,0%	100,0%	100,0%
≥ 1200	0,0%	0,0%	0,0%	0,0%	1,2%	45,4%	93,0%	100,0%	100,0%	100,0%	100,0%
≥ 1300	0,0%	0,0%	0,0%	0,0%	0,2%	23,3%	82,1%	99,8%	99,9%	100,0%	100,0%
≥ 1400	0,0%	0,0%	0,0%	0,0%	0,1%	9,9%	65,4%	99,6%	99,6%	100,0%	100,0%
≥ 1500	0,0%	0,0%	0,0%	0,0%	0,1%	3,7%	46,0%	99,0%	99,6%	100,0%	100,0%
≥ 1600	0,0%	0,0%	0,0%	0,0%	0,0%	1,3%	27,4%	98,4%	99,6%	100,0%	100,0%
≥ 1700	0,0%	0,0%	0,0%	0,0%	0,0%	0,4%	14,3%	96,3%	99,6%	100,0%	100,0%
≥ 1800	0,0%	0,0%	0,0%	0,0%	0,0%	0,3%	5,8%	91,4%	99,2%	100,0%	100,0%
≥ 1900	0,0%	0,0%	0,0%	0,0%	0,0%	0,1%	2,5%	81,4%	97,8%	100,0%	100,0%
≥ 2000	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,7%	63,3%	97,5%	100,0%	100,0%
≥ 2100	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,4%	45,8%	96,1%	100,0%	100,0%
≥ 2200	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,2%	32,7%	93,6%	100,0%	100,0%
≥ 2300	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,2%	18,7%	85,5%	100,0%	100,0%
≥ 2400	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%
≥ 2500	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%

Figura 6.1: Tabela do calculo das probabilidade - Ano completo - Outros intervalos

Referências bibliográficas

AGUILAR, S. Previsão da distribuição da densidade de probabilidade da geração de energia eólica usando técnicas não paramétricas. **Pontifícia Universidade Católica do Rio de Janeiro.**, 2015.

BAHRAOUI, Z.; JADIDA, E. et al. Modeling wind energy using copula. **Open Access Library Journal**, Scientific Research Publishing, v. 5, n. 11, p. 1, 2018.

BECKER, R. Generation of time-coupled wind power infeed scenarios using pair-copula construction. **IEEE Transactions on Sustainable Energy**, v. 9, n. 3, p. 1298–1306, 2018. Disponível em: <<https://www.scopus.com/inward/record.uri?eid=2-s2.0-85038854721&doi=10.1109\%2fSTE.2017.2782089&partnerID=40&md5=2ec7c5daedc1cd6a9e9a24d6c1684731>>.

BERNARD, C.; CZADO, C. Conditional quantiles and tail dependence. **Journal of Multivariate Analysis**, Elsevier, v. 138, p. 104–126, 2015.

BOLETIM Trimestral da Energia Eólica. <https://www.epe.gov.br/sites-pt/publicacoes-dados-abertos/publicacoes/PublicacoesArquivos/publicacao-233/topico-520/Boletim_2020_Q1.pdf>. Accessed: 2022-05-20.

BRAZILIAN Wind Energy Association (ABEEólica) - Annual Wind Energy Report of 2020. <<http://abeeolica.org.br/dados-abeeolica/>>. Accessed: 2022-05-20.

BRAZILIAN Wind Energy Association (ABEEólica) - Annual Wind Energy Report of 2021. <<http://abeeolica.org.br/dados-abeeolica/>>. Accessed: 2022-05-20.

CAPÉRAÀ, P.; FOUGÈRES, A.-L.; GENEST, C. A nonparametric estimation procedure for bivariate extreme value copulas. **Biometrika**, Oxford University Press, v. 84, n. 3, p. 567–577, 1997.

CARLIN, P.; LAXSON, A.; MULJADI, E. **The history and state of the art of variable-speed wind turbine technology**. National Renewable Energy Laboratory. [S.l.], 2001.

CARNEIRO, J. A. Energia eólica. 2013.

CONRADSEN, K.; NIELSEN, L.; PRAHM, L. Review of weibull statistics for estimation of wind speed distributions. **Journal of Applied Meteorology and Climatology**, v. 23, n. 8, p. 1173–1183, 1984.

COUNCIL, G. W. E. Gwec| global wind report 2021. 2021.

D'AMICO, G.; PETRONI, F.; SOBOLEWSKI, R. Optimal control of a dispatchable energy source for wind energy management. **Stochastics and Quality Control**, v. 34, n. 1, p. 19–34, 2019. Disponível em: <<https://www.scopus.com/inward/record.uri?eid=2-s2.0-85061738922&doi=10.1515\%2feqc-2019-0001&partnerID=40&md5=255e6728e2175d227eda2b961b1448a3>>.

DEMARTA, S.; MCNEIL, A. J. The t copula and related copulas. **International statistical review**, Wiley Online Library, v. 73, n. 1, p. 111–129, 2005.

DÍAZ, G.; GÓMEZ-ALEIXANDRE, J.; COTO, J. Statistical characterization of aggregated wind power from small clusters of generators. **International Journal of Electrical Power & Energy Systems**, Elsevier, v. 62, p. 273–283, 2014.

DJOKIC, S. et al. On the importance of correlating wind speed and wind direction for evaluating uncertainty in wind turbine power output. In: . [s.n.], 2019. p. 36–45. Disponível em: <<https://www.scopus.com/inward/record.uri?eid=2-s2.0-85075371146&doi=10.1109/%2fICCEP.2019.8890086&partnerID=40&md5=1f8dea81c80ab9493b50aa03e08f6aa1>>.

DONG, W. et al. Regional wind power probabilistic forecasting based on an improved kernel density estimation, regular vine copulas, and ensemble learning. **Energy**, v. 238, 2022.

DUCA, V. E. L. de A. Non-gaussian dynamical modeling of wind power generation.

D'ÂMICO, G. et al. Managing wind power generation via indexed semi-markov model and copula. **Energies**, v. 13, n. 6, 2020. Disponível em: <<https://www.scopus.com/inward/record.uri?eid=2-s2.0-85090923083&doi=10.3390/%2fen13164246&partnerID=40&md5=85c88c3e4a608dc4da4011c3de567467>>.

FILHO, A. C. de L. Um breve estudo sobre o avanço da energia eólica. **Revista Engenharia e Tecnologia Aplicada-UNG-Ser**, v. 3, n. 1, p. 31–38, 2019.

GENEST, C.; MACKAY, R. J. Copules archimédiennes et familles de lois bidimensionnelles dont les marges sont données. **Canadian journal of statistics**, Wiley Online Library, v. 14, n. 2, p. 145–159, 1986.

GENEST, C.; RIVEST, L.-P. Statistical inference procedures for bivariate archimedean copulas. **Journal of the American statistical Association**, Taylor & Francis, v. 88, n. 423, p. 1034–1043, 1993.

GOYAL, N. et al. Reliability measures and profit exploration of windmill water-pumping systems incorporating warranty and two types of repair. **Mathematics**, v. 9, n. 8, 2021. Disponível em: <<https://www.scopus.com/inward/record.uri?eid=2-s2.0-85104837179&doi=10.3390/%2fmath9080822&partnerID=40&md5=b68a76046a7e07724ca3c01727c30b08>>.

GROTHER, O.; SCHNIEDERS, J. Spatial dependence in wind and optimal wind power allocation: A copula-based analysis. **Energy Policy**, v. 39, n. 9, p. 4742–4754, 2011.

GRUGEL, D. M. et al. Estudo de dependência de extremos via teoria de cópulas. 2015.

HAGHI, H. V. et al. Using copulas for analysis of large datasets in renewable distributed generation: Pv and wind power integration in iran. **Renewable Energy**, Elsevier, v. 35, n. 9, p. 1991–2000, 2010.

HAGSPIEL, S. et al. Copula-based modeling of stochastic wind power in europe and implications for the swiss power grid. **Applied Energy**, v. 96, p. 33–44, 2012. Disponível em: <<https://www.scopus.com/inward/record.uri?eid=2-s2.0-84861657566&doi=10.1016/j.apenergy.2011.10.039&partnerID=40&md5=0cd22c2f21d7e8744a7c9d13390b80c1>>.

HAN, Q.; CHU, F. Directional wind energy assessment of china based on nonparametric copula models. **Renewable Energy**, v. 164, p. 1334–1349, 2021. Disponível em: <<https://www.scopus.com/inward/record.uri?eid=2-s2.0-85095777721&doi=10.1016/j.renene.2020.10.149&partnerID=40&md5=bfeac0b018b83ddc7cb93699ac521404>>.

HANIF, W. et al. Nonlinear dependence and connectedness between clean/renewable energy sector equity and european emission allowance prices. **Energy Economics**, v. 101, 2021. Disponível em: <<https://www.scopus.com/inward/record.uri?eid=2-s2.0-85111952117&doi=10.1016/j.eneco.2021.105409&partnerID=40&md5=2f23859a863bbe3cb624b282db8386f6>>.

HENZE, J. et al. Probabilistic upscaling and aggregation of wind power forecasts. **Energy, Sustainability and Society**, v. 10, n. 1, 2020. Disponível em: <<https://www.scopus.com/inward/record.uri?eid=2-s2.0-85082123340&doi=10.1186/s13705-020-00247-4&partnerID=40&md5=5ee8e63a748b5f128088e7af3370e581>>.

HOFERT, M. et al. **Elements of copula modeling with R**. [S.l.]: Springer, 2018.

IRENA, I. Future of wind: Deployment, investment, technology, grid integration and socio-economic aspects. International Renewable Energy Agency Abu Dhabi, 2019.

JR, J. P. H. Some aspects of wind power statistics. **Journal of Applied Meteorology and Climatology**, v. 16, n. 2, p. 119–128, 1977.

JUNG, C.; SCHINDLER, D. Development of a statistical bivariate wind speed-wind shear model (wsws) to quantify the height-dependent wind resource. **Energy Conversion and Management**, v. 149, p. 303–317, 2017. Disponível em: <<https://www.scopus.com/inward/record.uri?eid=2-s2.0-85025138915&doi=10.1016/j.enconman.2017.07.033&partnerID=40&md5=d4b7acf0a53e07d238b34ef5bd6c5331>>.

KARAKAŞ, A. M. Using copulas for modeling dependence in wind power. **Asian Journal of Engineering and Technology**, v. 7, n. 1, 2019.

KRISMANTO, A. et al. Copula-monte carlo based probabilistic oscillatory stability analysis of microgrid. **International Journal of Intelligent Engineering and Systems**, v. 14, n. 5, p. 479–491, 2021. Disponível em: <<https://www.scopus.com/inward/record.uri?eid=2-s2.0-85114735888&doi=10.22266/ijies2021.1031.42&partnerID=40&md5=a984475a7ddfbad6274ea6eecd4536dd>>.

LI, X.; ZHANG, W. Long-term fatigue damage assessment for a floating offshore wind turbine under realistic environmental conditions. **Renewable Energy**, v. 159, p. 570–584, 2020. Disponível em: <<https://www.scopus.com/inward/>

record.uri?eid=2-s2.0-85086894630&doi=10.1016\%2fj.renene.2020.06.043&partnerID=40&md5=9931e525ac868b9656370c06267c524a>.

LI, Y. et al. A study on probability distribution models for multiple wind speeds. In: . [s.n.], 2020. p. 2054–2059. Disponível em: <<https://www.scopus.com/inward/record.uri?eid=2-s2.0-85088555102&doi=10.1109\%2fACPEE48638.2020.9136432&partnerID=40&md5=89d08aa9ac2cf63c5e00c7124b8d2940>>.

LI, Y.; ZHAO, B.; BAI, H. A wind speed simulation approach based on markov sequence model and archimedean copula. In: . [s.n.], 2018. p. 411–416. Disponível em: <<https://www.scopus.com/inward/record.uri?eid=2-s2.0-85070314305&doi=10.1109\%2fCIEEC.2018.8745949&partnerID=40&md5=c4a2d3f7969e6f82beb085f5c3222c5c>>.

LIMA, L. F.; GUIMARÃES, S. Potência extraída de turbinas eólicas baseada na-comparação de diferentes tipos de velocidade dos ventos. **IX CEEL. Uberlândia:[sn]**, 2011.

LU, Q. et al. Wind power uncertainty modeling considering spatial dependence based on pair-copula theory. In: IEEE. **2014 IEEE PES General Meeting| Conference & Exposition**. [S.l.], 2014. p. 1–5.

LUCIANO, E.; CHERUBINI, U. Multivariate option pricing with copulas. **Available at SSRN 269868**, 2000.

MALUF, Y. S. Distribuição de funções de variáveis aleatórias dependentes e r-vines cópulas. 2015.

Marius Hofert; Martin Mächler. Nested archimedean copulas meet R: The nacopula package. **Journal of Statistical Software**, v. 39, n. 9, p. 1–20, 2011. Disponível em: <<https://www.jstatsoft.org/v39/i09/>>.

MI, Y. et al. Evaluation method of renewable energy accommodation considering wind-solar correlation. In: . [S.l.: s.n.], 2021. p. 1205–1209.

MIAO, S. et al. Modeling wind direction randomness and correlation using a probabilistic approach. In: . [s.n.], 2020. p. 292–299. Disponível em: <<https://www.scopus.com/inward/record.uri?eid=2-s2.0-85101695992&doi=10.1109\%2fICPES51309.2020.9349700&partnerID=40&md5=f82d4b45c3f88d72098083897b5b4d02>>.

MOLNAR, P.; CAMARGO, L.; RAMOS, D. Applying copulas functions for wind and hydro complementarity evaluation: A brazilian case. In: . [s.n.], 2015. v. 2015-August. Disponível em: <<https://www.scopus.com/inward/record.uri?eid=2-s2.0-84952029416&doi=10.1109\%2fEEM.2015.7216743&partnerID=40&md5=cae42f49e8cc9816a980779dc412d071>>.

MUDANÇAS Climáticas e Desdobramentos sobre os Estudos de Planejamento Energético: Considerações Iniciais. :<http://www.epe.gov.br/sites-pt/publicacoes-dados-abertos/publicacoes/PublicacoesArquivos/publicacao-227/topico-457/Mudancas%20Climaticas%20e%20Planejamento%20Energetico.pdf>.

NASCIMENTO, K. K. F. do et al. Comparação de métodos e modelos de distribuição para a modelagem de dados de velocidade do vento no município de petrolina, nordeste brasileiro. **Research, Society and Development**, v. 9, n. 7, p. e308974221–e308974221, 2020.

NELSEN, R. B. Dependence and order in families of archimedean copulas. **Journal of Multivariate Analysis**, Elsevier, v. 60, n. 1, p. 111–122, 1997.

OZAWA, M. T. **Aplicação de parâmetros da distribuição de Weibull na análise do potencial energético de um microgerador eólico em cidades do estado do Paraná**. Dissertação (B.S. thesis) — Universidade Tecnológica Federal do Paraná, 2017.

PARK, H.; BALDICK, R. Optimal capacity planning of generation system integrating uncertain solar and wind energy with seasonal variability. **Electric Power Systems Research**, v. 180, 2020. Disponível em: <<https://www.scopus.com/inward/record.uri?eid=2-s2.0-85076767797&doi=10.1016/j.epsr.2019.106072&partnerID=40&md5=354922e2a190b0c074a8294421d09b2f>>.

PASQUALETTI, M.; RIGHTER, R.; GIPE, P. History of wind energy. In: _____. [S.l.: s.n.], 2004. p. 419–433.

PATTON, A. J. Copula-based models for financial time series. In: **Handbook of financial time series**. [S.l.]: Springer, 2009. p. 767–785.

R Core Team. **R: A Language and Environment for Statistical Computing**. Vienna, Austria, 2021. Disponível em: <<https://www.R-project.org/>>.

R., L. A. et al. Complementarity modeling of monthly streamflow and wind speed regimes based on a copula-entropy approach: A brazilian case study. **Applied Energy**, v. 259, 2020. Disponível em: <<https://www.scopus.com/inward/record.uri?eid=2-s2.0-85075956003&doi=10.1016/j.apenergy.2019.114127&partnerID=40&md5=da342c44187f03df74f374484ec0b183>>.

ROSA, A. C. M. et al. **Uma introdução à teoria das copulas e sua aplicação à estimação do Valor em Risco e do Déficit Esperado**. Tese (Doutorado) — Universidade de Coimbra, 2019.

ROSENBLATT, M. Remarks on some nonparametric estimates of a density function. **The annals of mathematical statistics**, JSTOR, p. 832–837, 1956.

SAMPAIO, K. R. A.; BATISTA, V. O atual cenário da produção de energia eólica no brasil: Uma revisão de literatura. **Research, Society and Development**, v. 10, n. 1, 2021.

SANFINS, M. A. **Cópias para Distribuições Generalizadas: de Valores Extremos Multidimensionais**. [S.l.]: Novas Edições Acadêmicas, 2020.

SANTOS, F. S. dos et al. Análise estatística da velocidade do vento em petrolina-pe utilizando as distribuições weibull e a burr. **Journal of Environmental Analysis and Progress**, p. 057–064, 2019.

SCHINDLER, D.; BEHR, H.; JUNG, C. On the spatiotemporal variability and potential of complementarity of wind and solar resources. **Energy Conversion and Management**, v. 218, 2020. Disponível em: <<https://www.scopus.com/inward/record.uri?eid=2-s2.0-85085574105&doi=10.1016/j.enconman.2020.113016&partnerID=40&md5=60e505f12d9e657e2825691433da2534>>.

SCHINDLER, D.; JUNG, C. Copula-based estimation of directional wind energy yield: A case study from germany. **Energy Conversion and Management**, v. 169, p. 359–370, 2018. Disponível em: <<https://www.scopus.com/inward/record.uri?eid=2-s2.0-85047622117&doi=10.1016/j.enconman.2018.05.071&partnerID=40&md5=dbf1c9fb7974a09d475552d5c87ed2e4>>.

SCHINDLER, D.; SCHMIDT-ROHR, S.; JUNG, C. On the spatiotemporal complementarity of the european onshore wind resource. **Energy Conversion and Management**, v. 237, 2021.

SKLAR, M. Fonctions de repartition an dimensions et leurs marges. **Publ. inst. statist. univ. Paris**, v. 8, p. 229–231, 1959.

SOUZA, A. de; OLIVEIRA, S. S. de; OZONUR, D. Análise estatística de parâmetros de weibull para avaliação de potencial de energia eólica em campo grande. **Journal of Environmental Analysis and Progress**, p. 168–179, 2019.

SUN, Y. et al. A wind speed simulation model based on condition dependence and its application. **Taiyangneng Xuebao/Acta Energiae Solaris Sinica**, v. 38, n. 11, p. 3131–3137, 2017. Disponível em: <<https://www.scopus.com/inward/record.uri?eid=2-s2.0-85045093010&partnerID=40&md5=9708c45aea11ea51340450b1f70a3392>>.

TAILLON, G. et al. Statistical analysis of cavitation erosion impacts in a vibratory apparatus with copulas. In: IOP PUBLISHING. **IOP Conference Series: Earth and Environmental Science**. [S.l.], 2019. v. 240, n. 6, p. 062035.

TULLER, S. E.; BRETT, A. C. The characteristics of wind velocity that favor the fitting of a weibull distribution in wind speed analysis. **Journal of Applied Meteorology and Climatology**, v. 23, n. 1, p. 124–134, 1984.

VELASCO, L. L. H. et al. Ajuste bayesiano para cópulas bivariadas. [sn], 2016.

WANG, Z. et al. Forecasted scenarios of regional wind farms based on regular vine copulas. **Journal of Modern Power Systems and Clean Energy**, v. 8, n. 1, p. 77–85, 2020. Disponível em: <<https://www.scopus.com/inward/record.uri?eid=2-s2.0-85089277760&doi=10.35833/jmpce.2017.000570&partnerID=40&md5=d53b3737f2151331d6e08aa5049b5383>>.

XIE, K.; LI, Y.; LI, W. Modelling wind speed dependence in system reliability assessment using copulas. **IET Renewable Power Generation**, v. 6, n. 6, p. 392–399, 2012. Disponível em: <<https://www.scopus.com/inward/record.uri?eid=2-s2.0-84878352708&doi=10.1049/iet-rpg.2011.0156&partnerID=40&md5=238717df50609b3bf87f852ada13b590>>.

XIE, K. et al. A two-stage wind speed model for multiple wind farms considering autocorrelations and cross-correlations. In: . [s.n.], 2016. Disponível em: <<https://www.scopus.com/inward/record.uri?eid=2-s2.0-85015165842&doi=10.1109\%2fPMAPS.2016.7764225&partnerID=40&md5=c2d3520756d507b7b3e3296a40ad2b29>>.

XIE, M. et al. Modeling of multi wind farm output correlation based on copula and its application in power system economic dispatch. **Dianwang Jishu/Power System Technology**, v. 40, n. 4, p. 1100–1106, 2016. Disponível em: <<https://www.scopus.com/inward/record.uri?eid=2-s2.0-84968548101&doi=10.13335\%2fj.1000-3673.pst.2016.04.018&partnerID=40&md5=dad4c2295cb7e75ea8ee85b78992ae5d>>.

XIE, Z. et al. Quasi-monte carlo based probabilistic optimal power flow considering the correlation of wind speeds using copula function. **IEEE Transactions on Power Systems**, v. 33, n. 2, p. 2239–2247, 2018. Disponível em: <<https://www.scopus.com/inward/record.uri?eid=2-s2.0-85028975516&doi=10.1109\%2fTPWRS.2017.2737580&partnerID=40&md5=9e282989013973f09f10b4ad5e56c939>>.

YUNDAI, X.; YUE, Y. Analysis of aggregated wind power dependence based on optimal vine copula. In: . [s.n.], 2019. p. 1788–1792. Disponível em: <<https://www.scopus.com/inward/record.uri?eid=2-s2.0-85074917427&doi=10.1109\%2fISGT-Asia.2019.8881069&partnerID=40&md5=cea3a66d1e4bcf92c1dcec0ceb635852>>.

ZHANG, N. et al. Modeling conditional forecast error for wind power in generation scheduling. **IEEE Transactions on Power Systems**, IEEE, v. 29, n. 3, p. 1316–1324, 2013.

ZHANG, Y. et al. Transmission planning considering spatial correlation of wind power. In: . [s.n.], 2017. Disponível em: <<https://www.scopus.com/inward/record.uri?eid=2-s2.0-85028763223&doi=10.1109\%2fTDCLLM.2016.8013227&partnerID=40&md5=0b6898cd9fefac4fc069b62b4ee5cf6e>>.

ZHU, X. et al. Learning-based probabilistic power flow calculation considering the correlation among multiple wind farms. **IEEE Access**, v. 8, p. 136782–136793, 2020. Disponível em: <<https://www.scopus.com/inward/record.uri?eid=2-s2.0-85090394789&doi=10.1109\%2fACCESS.2020.3011511&partnerID=40&md5=9a9ff0e164f0e0d0b1e84223b8d005c4>>.