

# RELAÇÃO DE EXTREMOS DE VENTOS E A OPERAÇÃO DO T-27 NO 1º ESQUADRÃO DE INSTRUÇÃO AÉREA<sup>1</sup>

## *RELATIONSHIP BETWEEN WIND EXTREMES AND THE OPERATION OF THE T-27 IN THE 1st AIR INSTRUCTION SQUADRON*

**Rafael Luiz Xavier<sup>2</sup>**  
João Paulo M. dos Santos<sup>3</sup>  
Marcus Vinícius de A. Lima<sup>4</sup>

### RESUMO

Eventos extremos de velocidades de ventos são prejudiciais ao desenvolvimento da atividade humana, seja ela uma edificação, onde máximas de ventos podem causar danos estruturais, ou em captação de energia eólica, em que mínimas de ventos interferem na geração de energia. Na atividade de instrução aérea na Academia da Força Aérea (AFA), que foi o objeto de estudo desta pesquisa, as máximas de ventos inviabilizam as decolagens dos Cadetes com os instrutores e geram um atraso no progresso operacional dos pilotos. Baseado nesse contexto, o estudo e análise das intensidades e durações dos ventos, por meio de dados reais, permite o cálculo das curvas de probabilidades de eventos de máximas de ventos, separados por estações do ano, a fim de obter um conhecimento prévio de possíveis interrupções nas instruções aéreas. O tratamento dos dados foi pautado pelo uso do método do limiar e separado por estações do ano. A proposta de trabalho é delimitada pela restrição dos cálculos às funções de densidade de probabilidades de Gumbel, Pareto, Fréchet e Distribuição Generalizada de Valores Extremos (GEV). A construção dos itens anteriormente discutidos foi direcionada pela linguagem Python, fazendo uso do Colab, juntamente com as bibliotecas específicas para a análise de dados.

**Palavras-chave:** Instrução aérea; Distribuição de valores extremos; Estatística; Máximas de vento.

---

<sup>1</sup> Artigo de Conclusão de Curso apresentado ao Curso de Formação de Oficiais Aviadores (CFOAv) da Academia da Força Aérea (AFA).

<sup>2</sup> Cadete Aviador do 4º Esquadrão (Turma Orthrus, 2023).

<sup>3</sup> Prof Dr João Paulo M. dos Santos. Doutor em Ciências. Academia da Força Aérea. E-mail: [jpmdosantos@yahoo.com.br](mailto:jpmdosantos@yahoo.com.br).

<sup>4</sup> Prof Dr Marcus Vinícius de A. Lima. Doutor em Ciências. Academia da Força Aérea. E-mail: [marcuslima.afa@gmail.com](mailto:marcuslima.afa@gmail.com).

## ABSTRACT

Extreme wind speeds can be detrimental to human activities, whether it is a building where maximum wind speeds can cause structural damage or in wind energy capture, where minimum wind speeds interfere with power generation. In the aerial instruction activity at the Air Force Academy (AFA), which was the object of study in this research, maximum wind speeds make it impossible for cadets to take off with instructors, leading to operational delays for pilots. Based on this context, studying and analyzing wind intensities and durations using real data allows for the calculation of probability curves for maximum wind events, separated by seasons, in order to gain prior knowledge of possible interruptions in aerial instructions. The data treatment was based on the threshold method and separated by seasons. The proposed work is delimited by the constraint of calculating the probability density functions of Gumbel, Pareto, Fréchet, and Generalized Extreme Value (GEV) distributions. The construction of the aforementioned items was guided by the Python language, using Colab along with specific libraries for data analysis.

**Keywords:** Air instruction; Extreme value distribution; Statistics; Maximum wind speed.

## INTRODUÇÃO

Segundo Gipe (1995), desde os primórdios, os seres humanos aprenderam a utilizar os meios naturais para benefícios próprios. Acredita-se que o vento, por exemplo, foi utilizado pela primeira vez, de forma prática, no ano de 2800 AC pelos egípcios, usando velas para ajudar a força dos remos dos escravos. Eventualmente, as velas apoiavam o trabalho da força animal em tarefas como moagem de grãos e bombeamento de água. Entretanto, quando se registra velocidades elevadas de vento, os seus efeitos, geralmente, podem ser danosos e gerar sinistros, como naufrágios de embarcações e tsunamis nas regiões costeiras (FUENTES; BITENCOURT; FUENTES, 2013).

De forma análoga, o vento ocupa posição de destaque no contexto da aviação. Eventos de máximo, quando acima dos limites suportados pelas aeronaves, são prejudiciais ao bom desenvolvimento da atividade aérea, principalmente na área de instruções básicas de voo. A velocidade do vento é um termo relativo à posição do observador e utiliza um valor de referência, adotado por organização internacional (SANT'ANNA, 2018). A caracterização quantitativa detalhada dos ventos leva em conta a direção, magnitude, unidades e metodologia de medição, intensidade, frequência, probabilidade de ocorrência e distribuição de velocidades segundo a direção (FERREIRA, 2005).

De forma geral, a estimativa dos parâmetros das curvas de densidade que descrevem os eventos de máximas de ventos é de extrema importância para o planejamento das atividades sujeitas a seus efeitos desfavoráveis, e uma forma de modelar esses eventos, foi utilizar a teoria dos valores extremos proposta por Fisher e Tippett (1928). De acordo com esta teoria, existem três tipos de distribuições assintóticas de valores extremos, a tipo I de Gumbel, a tipo II de Fréchet e a tipo III de Weibull (BAUTISTA; ZOCCHI; ANGELOCCI, 2004). Além dessas distribuições, foram utilizadas as funções de densidade de Pareto e a Distribuição Generalizada de Valores Extremos (GEV), desenvolvida por Jenkinson (1955) (LISKA; BORTOLINI; SÁ-FADI; BEIJO, 2013).

O desenvolvimento e cálculo dos parâmetros das funções de densidade pôde ser realizado pela incorporação de uma linguagem de programação e a adoção de bibliotecas específicas para a computação científica. Nesse sentido, as bibliotecas componentes foram utilizadas por meio da linguagem Python, fazendo uso da plataforma Colab e tomando por base as referências (BRANCALHÃO; SANTOS; FIRMIANO; 2021) e (SANTOS *et al.*, 2021), a fim de utilizar esse artigo como ponto de início para uma abordagem mais profunda na questão do impacto do vento nas atividades aéreas.

O presente estudo foi desenvolvido com o propósito principal de analisar os dados de ventos obtidos da estação de medição de São Carlos, obter a probabilidade de ocorrência sazonal de valores extremos de velocidades de vento acima dos valores previstos para instrução no 1º Esquadrão de Instrução Aérea (1º EIA), de acordo com o Programa de Instrução e Manutenção Operacional da AFA (PIMO) e o Manual de procedimentos do 1º Esquadrão de Instrução Aérea (MAPRO), de forma a identificar, como objetivo geral, as principais curvas de densidade de probabilidade que modelam os fenômenos de extremos de vento nas diferentes estações do ano.

Para atender ao objetivo geral da pesquisa foi fundamental cumprir com alguns objetivos específicos, dentre eles, a compreensão dos conceitos estatísticos fundamentais como média, desvio padrão e análise do gráfico de percentis e ajuste de curvas e teste de hipóteses, além de compreender a linguagem de programação Python, e a plataforma *Colaboratory* (Colab), a fim de desenvolver o código computacional.

A adoção da Estação Meteorológica de São Carlos se deve à disponibilidade pública dos dados no site do INMET (<https://portal.inmet.gov.br/>) e pelo fato de que a cidade de São Carlos se encontra dentro da área de instrução da AFA. Deve ser ressaltado que, os desenvolvimentos alcançados podem ser prontamente empregados para os dados da Academia da Força Aérea,

bastando que os dados sejam disponibilizados para livre acesso. Até o momento do início do presente artigo, o endereço eletrônico do INMET não continha tais informações e, por isso, optou-se pela utilização dos dados da estação meteorológica automática de São Carlos. Esses dados servem de ferramenta para a elucidação do seguinte problema de pesquisa: de que maneira a velocidade de vento, de acordo com a estação do ano e sua curva de densidade de probabilidade, interferem na instrução aérea?

## **1 REVISÃO DA LITERATURA**

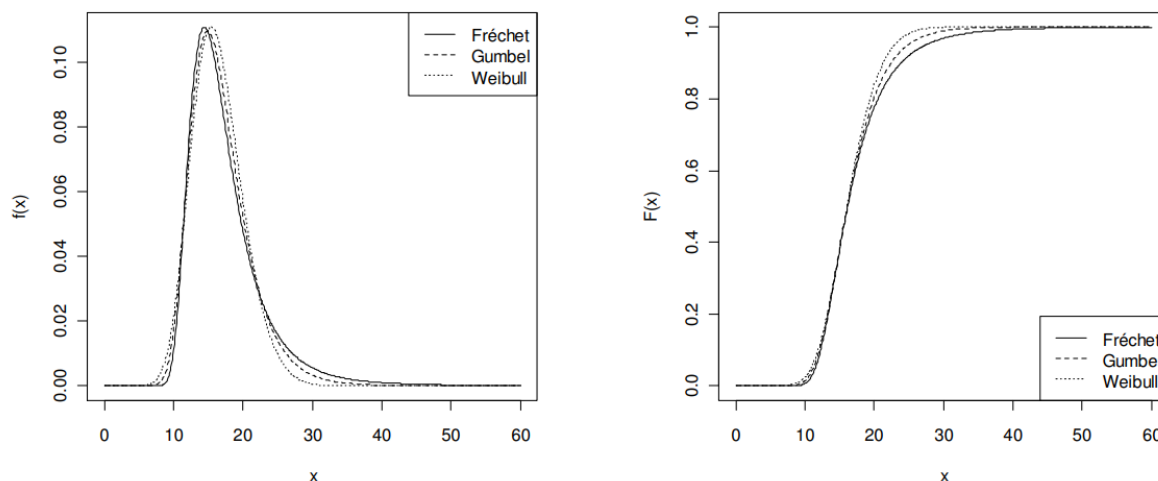
“A Teoria Probabilística dos Valores Extremos tem uma enorme aplicação envolvendo fenômenos naturais como chuvas, inundações, rajadas de vento, poluição do ar e corrosão, pois fornece um resultado em processos pontuais e funções variadas” (KOTZ; NADARAJAH, 2000, p.01). Conforme Kendall e Stuart (1946), essa teoria é uma abordagem estatística que se concentra na modelagem e análise de eventos raros ou extremos em um conjunto de dados.

De acordo com Bautista, Zocchi e Angelocci (2004), a teoria dos valores extremos desempenha um papel fundamental na modelagem de eventos associados a probabilidades de eventos raros. De forma geral, a previsão probabilística da ocorrência de ventos extremos é de vital importância para o planejamento das atividades sujeitas a seus efeitos adversos.

Em conjunto com a teoria dos valores extremos, para fins de estimativas probabilísticas, existem dois vieses de análise dos ventos extremos: máximos acima do limiar e máximos por blocos de tempo. O primeiro estabelece um limite inferior para que um valor seja classificado como um evento de extremo de máximo, enquanto o segundo consiste em estabelecer um período de tempo e obter o máximo valor no período (HECKERT; SIMIU; YEO, 2020).

Segundo Liska et al. (2013), a distribuição generalizada de valores extremos (GEV) atrelada aos três tipos de distribuições de valores extremos, as distribuições Weibull, Gumbel e Fréchet, mostradas à esquerda no gráfico 1, geram uma das funções de densidade de probabilidade acumulada (mostradas à direita no gráfico). A questão que surge naturalmente é qual dos três tipos melhor descreve os máximos de velocidades de vento (PAPALEXIOU; KOUTSOYIANNIS, 2013). A alteração dos parâmetros da GEV dá origem às famílias de funções citadas, ou seja, Fréchet, Weibull e Gumbel são casos particulares da GEV, obtidos por meio da alteração de algum parâmetro específico.

**Gráfico 1** Gráfico da função densidade de probabilidade (esquerda) e probabilidade acumulada (direita) da distribuição GVE nos casos Fréchet, Weibull e Gumbel



Fonte: Liska *et al.* (2013).

Rowinski *et al.* (2002), após análise dos resultados, obteve que a distribuição de probabilidade do tipo Gumbel é a função de densidade mais adequada para a descrição dos fenômenos meteorológicos que envolvam máximos, entre os quais se enquadram as velocidades máximas de vento, entretanto há aplicações específicas que apontam que outras distribuições também são adequadas. Estas observações e referências são apontadas em (BRANCALHÃO; SANTOS; FIRMIANO; 2021) e (SANTOS *et al.*, 2021) para a utilização de outras funções de densidade. A justificativa é baseada no fato de que os parâmetros, obtidos por meio das estimativas fazendo uso de procedimentos estatísticos, geram curvas que não oferecem condições suficientes para descartar a hipótese testada e várias curvas podem satisfazer tal critério. Como consequência é necessário um procedimento estatístico para testar a qualidade dos parâmetros calculados.

Os parâmetros para um conjunto de funções de densidade, utilizando método de máxima verossimilhança, foram obtidos em Brancalhão e Santos (2021) considerando dados de velocidade de vento, sem separação por estações, por meio do método dos máximos acima de um limiar previamente estabelecido. Em seguida, as funções de densidade são classificadas de acordo com a ordenação dos valores  $p$  provenientes do teste de Kolmogorov-Smirnov. Estes resultados mostram quais curvas, dentre aquelas testadas, que não podem ser descartadas segundo o parâmetro  $\alpha=0.05$  e possuem, por consequência, características que devem ser consideradas para a representação do conjunto de dados. Estes resultados estão diretamente

relacionados a escolha das curvas de densidade, pois a adoção do tipo errado pode levar a uma subestimação ou superestimação dos resultados obtidos (PAPALEXIOU; KOUTSOYIANNIS, 2013).

## 1.1 DESCRIÇÃO GERAL

Os ventos são descritos por meio de sua intensidade (geralmente expressa em km/h ou nós (kt)), direção (seu sentido, de onde vem) e caráter (se é um vento contínuo ou uma rajada). Essas características podem ser obtidas através de um aparelho chamado anemômetro, mas podem também ser indicadas através da biruta do aeroporto.

Como o avião deve pousar e decolar de uma pista, tem apenas uma escolha: pode escolher a cabeceira da pista a usar, isto é, o sentido da pista a usar. A escolha da cabeceira é feita em função do vento que sopra na ocasião da operação. Uma vez escolhida a cabeceira, de modo que se opere contra a componente de vento da direção da pista, temos que conviver com a componente do vento perpendicular ao eixo da pista (SÓRIA, 2003, p. 03).

De acordo com a Agência Nacional de Aviação Civil (ANAC), o vento de proa é o mais indicado para as operações de pousos e decolagens, pois ele permite um ganho de sustentação antecipado, se comparado com o vento de cauda. A operação contra o vento gera parte da elevação necessária, resultando em uma menor velocidade no solo e menor distância de pista para decolagem. O vento de proa pode proporcionar também uma maior razão de subida. Pousar contra o vento pode também proporcionar uma operação em que é necessário utilizar menor distância de pista e menor velocidade ao atingir a zona de toque.

Portanto, como o intuito do trabalho foi analisar do ponto de vista probabilístico, os casos em que os ventos atrapalham a atividade aérea. A base da análise são os ventos de través, que não favorecem o bom desempenho do avião nos momentos de pousos e decolagens.

## 1.2 DESCRIÇÃO DAS VARIÁVEIS

### 1.2.1 Descrição das variáveis gerais

A caracterização das variáveis de interesse é feita nos parágrafos seguintes baseando-se no Manual de Procedimentos do 1º EIA (Mapro), vide Quadro 2. O Mapro define apenas a intensidade e direção do vento toleráveis para a instrução aérea na AFA no que tange pousos e decolagens. O presente estudo não se prendeu apenas a essas duas fases do voo, mas sim, de uma forma geral, incluindo os exercícios na área de instrução.

**Tabela 1** Limites de vento no pouso e decolagem

Cálculo da Componente de TRAVÉS							
Vel. Kts	20°	30°	40°	50°	60°	70°	80°
10			6	8	9	9	10
11			6	8	9	9	10
12		6	7	9	10	11	12
13		6	8	10	11	12	13
14		7	9	11	12	13	14
15		8	10	11	13	14	15
16		8	10	12	14	15	16
17	6	8	11	13	15	16	17
18	6	9	12	14	16	17	18
19	7	9	12	14	16	18	19
20	7	10	13	15	17	19	20
21	7	11	14	16	18	20	21
22	7	11	14	17	19	21	22
23	8	12	15	17	20	22	23
24	8	12	15	18	21	22	24
25	9	13	16	19	22	23	25
26	9	13	17	20	22	24	
27	9	13	17	21	23	25	
28	10	14	18	21	24		
29	10	14	19	22	25		
30	10	15	19	23			
31	11	15	20	24			
32	11	16	21	25			
	70°	60°	50°	40°	30°	20°	10°
Cálculo da Componente de PROA/CAUDA							
<div style="display: flex; justify-content: space-between; padding: 0;"> <div style="width: 20%;"><span style="background-color: #c6e0b4; border: 1px solid black; display: inline-block; width: 10px; height: 10px; margin-right: 5px;"></span> Fora para pouso na ala</div> <div style="width: 20%;"><span style="background-color: #4472c4; border: 1px solid black; display: inline-block; width: 10px; height: 10px; margin-right: 5px;"></span> Fora para Instrução</div> <div style="width: 20%;"><span style="background-color: #e74c3c; border: 1px solid black; display: inline-block; width: 10px; height: 10px; margin-right: 5px;"></span> Fora para operação</div> </div>							

Fonte: Manual de Procedimentos do 1° EIA, 2022.

I) **Direção horária do vento:** Indica a direção de onde o vento sopra, de acordo com os 360 graus definidos em uma circunferência, para cobrir todas as direções e no sentido horário. Por exemplo, uma medição que indica 290 graus, demonstra que o vento está soprando praticamente de Oeste, enquanto uma medição que indica 130 graus, está entre Leste e Sul ou exatamente soprando do Sudeste.

II) **Velocidade horária do vento:** As velocidades demonstradas nas tabelas são obtidas por meio de instrumentos específicos como anemógrafo e anemômetro e cuja operacionalidade em termos técnicos não cabe descrever na presente monografia, dada a complexidade e especificidade do assunto. Por meio de medições em intervalos de tempo específicos e conversões em tabelas já preconizadas, são fornecidas as velocidades médias horárias do vento de hora em hora em metros por segundo. Ou seja, se houver a indicação de 3.5, por exemplo, significa que a velocidade média do vento durante a hora de medição foi de 3.5 metros por segundo.

III) **Rajada máxima horária:** Indica a velocidade máxima que o vento atingiu, em metros por segundo, que o vento atingiu naquele intervalo de hora de medição, ou seja, uma indicação de 10.4 na hora 2:00, por exemplo, indicará que, durante o intervalo que compreende 2:00 até às 3:00 horas, o vento, dentre todas rajadas e velocidades, atingiu em algum momento a velocidade máxima de 10.4 metros por segundo.

### 1.2.2 Descrição das estações do ano

O início das estações do ano foi definido da seguinte forma: Outono: 20 de março de 2020, Inverno: 20 de junho de 2020, Primavera: 22 de setembro de 2020 e Verão: 21 de dezembro de 2020, de acordo com definição da Fiocruz (<https://portal.fiocruz.br/>). Vale ressaltar que os minutos e segundos foram descartados, utilizando apenas os dias completos.

## 1.3 HISTOGRAMAS E BOX-PLOTS

De acordo com Morettin e Bussab (2017), histogramas e box-plots são dois tipos comuns de gráficos usados na análise exploratória de dados em estatística. Histogramas são gráficos que mostram a frequência de ocorrência de valores em um conjunto de dados contínuos, enquanto box-plots são gráficos que mostram a distribuição dos dados resumidos em cinco estatísticas e se destacam por sua facilidade de visualização dos *outliers*, que são os dados que mais se distanciam da média. Esses gráficos são úteis para visualizar a distribuição dos dados e identificar padrões, como clusters, picos ou assimetrias.

## 1.4 ESTIMATIVA DOS PARÂMETROS DAS FUNÇÕES DE DENSIDADE DE PROBABILIDADE

O tratamento dos dados será pautado pelo uso do método dos máximos acima de um limiar conforme Heckert, Simiu e Yeo (2020) e separado por estações do ano. O método proposto foi o valor máximo maior ou igual a um limiar de valor fixo de 10,3 m/s. Esse valor foi selecionado pelo autor devido às dificuldades enfrentadas pelos alunos durante as missões de instrução na AFA. Ventos acima de 10,3 m/s (20 kt) prejudicam o raciocínio dos alunos em relação à intensidade e direção do vento, impactando negativamente a execução dos exercícios.

A proposta de trabalho é delimitada pela restrição dos cálculos às funções de densidade de probabilidades de Gumbel, Pareto, Fréchet, Distribuição Generalizada de Valores Extremos (GEV), Weibull e Weibull invertida.

Geralmente, as curvas mencionadas são utilizadas para descrever eventos extremos. Embora certas distribuições tenham sido indicadas e favorecidas em contextos específicos, como Fréchet, Weibull e Gumbel para análises de extremos meteorológicos, os resultados de Alam, Sabbir e Khan (2014) sugerem que a avaliação de outras distribuições pode fornecer resultados igualmente interessantes, que se adequam melhor aos valores  $p$  e à análise do gráfico quantil-quantil.

### 1.5 ANÁLISE DA QUALIDADE DO AJUSTE: KS-TESTE E O GRÁFICO QQ-PLOT

Os parâmetros para um conjunto de funções de máximos de ventos, utilizando método de máxima verossimilhança, foram obtidos em Brancalhão e Santos (2021). Em seguida, os valores foram ordenados em ordem crescente para o valor  $p$  retornado pelo teste de Kolmogorov-Smirnov.

De acordo com Montgomery e Runger (2010), o teste de Kolmogorov-Smirnov é um teste estatístico utilizado para comparar a distribuição empírica de uma amostra com uma distribuição teórica. O teste compara a função de distribuição acumulada (FDC) empírica da amostra com a FDC da distribuição teórica e calcula a estatística de teste  $D$ , que representa a maior diferença absoluta entre as duas FDCs. O valor de  $D$  é comparado com um valor crítico obtido de tabelas de distribuição para o nível de significância desejado. Se o valor de  $D$  for maior que o valor crítico, rejeita-se a hipótese nula de que as duas distribuições são iguais. Caso contrário, não se rejeita a hipótese nula.

## 2 MATERIAIS E MÉTODOS

De acordo com Gil (2019), a pesquisa caracteriza-se como explicativa dado que visa analisar as velocidades máximas de vento e o impacto desses eventos na operação de instrução aérea, e de estrutura quantitativa, visto que visa expor os dados em forma de números,

atribuindo valores ao objeto de estudo. No que se refere aos procedimentos técnicos, a pesquisa é de caráter descritiva e estatística inferencial.

A estrutura metodológica quanto aos procedimentos e análises que foram executadas na pesquisa, segundo Brancalhão e Santos (2021), está descrita a seguir levando em conta o desenvolvimento, os aspectos matemático, estatístico e computacionais: quanto aos aspectos Matemático e Estatístico, as seguintes etapas de desenvolvimento foram elaboradas:

I) obtenção dos dados: o levantamento de dados primários com a utilização dos dados em domínio público de velocidade do vento. Estes foram utilizados por (BRANCALHÃO; SANTOS; FIRMIANO; 2021) e (SANTOS *et al.*, 2021);

II) revisar os conceitos estatísticos elementares de medidas de tendência central e variabilidade bem como os conceitos de medidas separatrizes para cada uma das estações do ano;

III) interpretação e análise dos gráficos de box-plot, histograma e quantil-quantil para cada uma das estações do ano;

IV) estimar os parâmetros por meio do método dos estimadores de máxima verossimilhança para cada uma das estações do ano;

V) verificar a bondade do ajuste por meio do teste de Kolmogorov-Smirnov para cada uma das estações do ano.

Quanto aos aspectos computacionais, os seguintes elementos essenciais foram elencados:

I) compreender os requisitos básicos da utilização do Python dentro da linguagem de programação do Colab;

II) adaptar o código para ajuste de curvas de densidade de probabilidade ao caso proposto;

III) utilizar os recursos de visualização gráfica para avaliação dos gráficos quantil-quantil.

Quanto à análise e discussão dos resultados, as etapas são resumidas em:

I) analisar e interpretar os resultados numéricos obtidos com os desenvolvimentos matemáticos e computacionais;

II) analisar os resultados do teste KS e dos resultados gráficos do QQ-plot;

III) interpretar os resultados dos tempos de retorno à luz das probabilidades.

## 2.1 PRÉ-PROCESSAMENTO

O objetivo desta seção é explicar como os dados foram preparados e organizados para serem analisados estatisticamente. Para isso, foram realizadas diversas etapas de pré-processamento dos dados utilizando ferramentas computacionais como o Python e o Colab, bibliotecas Pandas e o artigo (BRANCALHÃO; SANTOS; FIRMIANO; 2021) e (SANTOS *et al.*, 2021). No início, os dados obtidos do INMET foram filtrados retirando os dados incompletos e interpretados em tabelas para ajudar nas discussões e facilitar a análise estatística posterior. As seguintes etapas iniciais de preparação dos dados envolveram a transferência dos arquivos recebidos para o ambiente do Colab, de modo a permitir sua análise computacional.

Posteriormente, os dados foram organizados por meio da renomeação das colunas, interpretação de sua representação computacional e exclusão dos dados não numéricos. Uma das etapas importantes foi a separação dos dados em meses do ano e a eliminação de dados não numéricos, que são representados na tabela por letras ou "nan" (*not a number* em inglês). Esses dados não numéricos podem resultar em valores incorretos nos cálculos estatísticos e na criação de elementos estatísticos, como histogramas e box-plots, por isso foram removidos e reorganizados.

No final dos processos de adaptação dos valores, os dados analisados foram divididos em três conjuntos: treino, validação e teste, sendo o conjunto de treino utilizado para estimar os parâmetros e os conjuntos de validação e teste recomendados para análises posteriores.

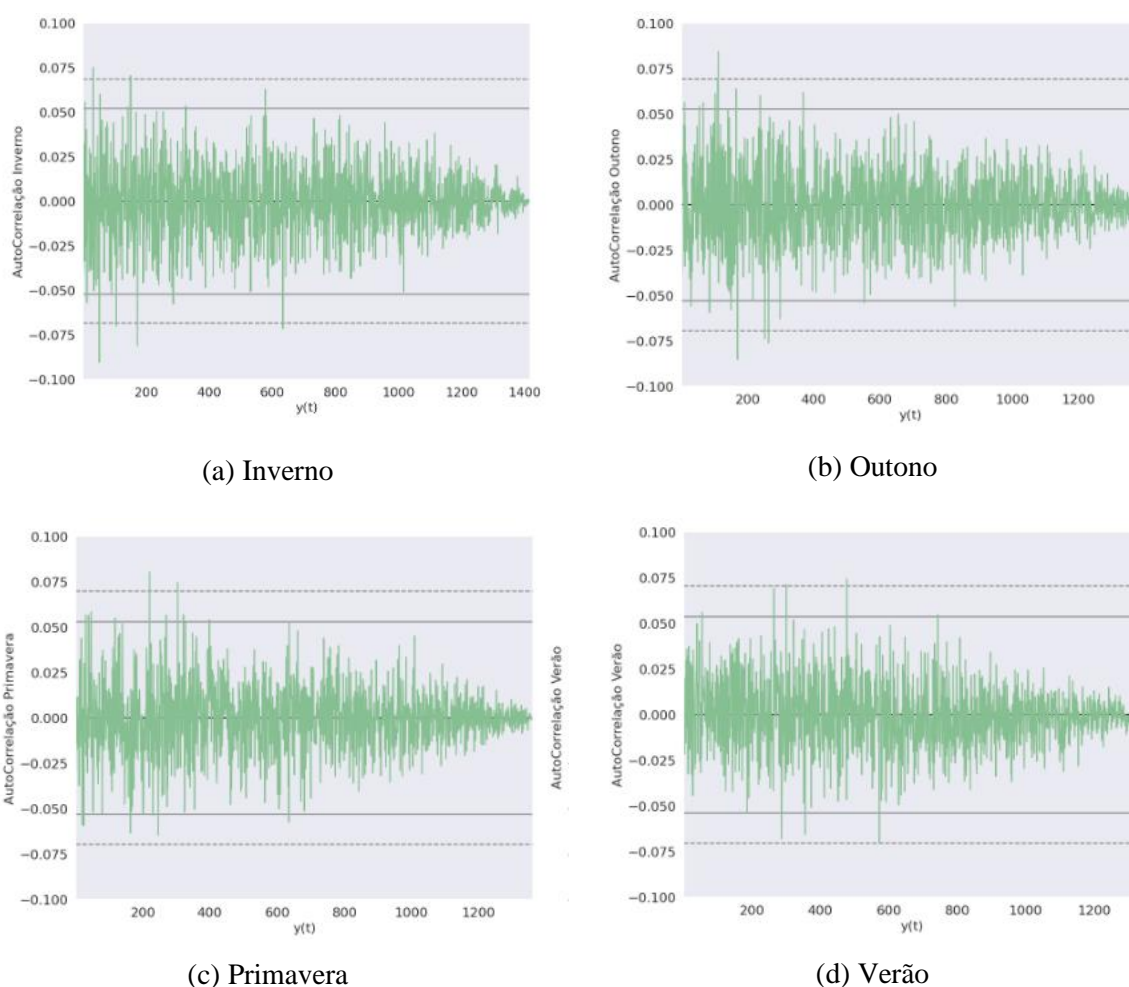
### **3 RESULTADOS E DISCUSSÃO**

Nesta seção, apresentam-se os resultados da análise estatística, iniciando com a análise descritiva dos dados pré-processados e acima do limiar de 10,3 m/s. Em seguida, é realizada uma análise gráfica descritiva. Por fim, é apresentada a análise das curvas de densidade de probabilidades com base nos valores p das cinco primeiras funções. Foram construídos os primeiros instrumentos gráficos de análise do conjunto de dados: os histogramas. Visualmente, há uma demanda de códigos maior e uma maior compreensão da linguagem é necessária nesse ponto, entretanto, as escolhas simplificadoras como a de usar a biblioteca Pandas mostraram-se bem sucedidas a esse ponto de interpretação.

Com base na análise dos resultados, o gráfico 2 apresenta o diagrama de autocorrelação. Os resultados não fornecem evidências estatisticamente significativas para rejeitar a hipótese de que a amostra de dados seja aleatória.

O conceito de autocorrelação, definido pelo National Institute of Standards and Technology (n.d.), indica que quando os valores analisados se aproximam de 1, pode-se inferir que os dados estão inter-relacionados. Nesse caso, é possível utilizar a informação do primeiro dia para prever os eventos dos próximos dias. Por outro lado, se os valores se aproximam de 0, os eventos são essencialmente aleatórios e não possuem uma relação entre si. Por exemplo, a ocorrência de ventos fortes em um dia não implica necessariamente que ventos fortes ocorrerão no dia seguinte.

**Gráfico 2** Diagrama de autocorrelação das estações do ano



Fonte: Elaboração própria.

O resumo estatístico descritivo é mostrado nas tabelas a seguir. Nelas estão o número de linhas de dados disponíveis, após a eliminação dos nan's, as médias, os desvios padrão, os valores mínimos, o primeiro quartil, o segundo quartil ou mediana, o terceiro quartil e o valor

máximo. Na Tabela 2, são considerados os dados gerais e na Tabela 3, os dados foram divididos em estações do ano.

**TABELA 2** Tabela de estatísticas descritivas geral: número de linhas sem nan's, média, desvio padrão, mínimo, quartis e valores máximos

GERAL			
	DirVento	Vmax	Vhora
Count	110215	110215	110215
Mean	162,06	4,91	1,81
Std	105,42	2,71	1,28
Min	1	0	0
25%	80	2,9	0,8
50%	126	4,8	1,7
75%	273	6,7	2,7
Max	360	24,3	9,3

Fonte: Adaptado de Brancalhão e Santos (2021).

**TABELA 3** Tabela de estatísticas descritivas das estações do ano: número de linhas sem nan's, média, desvio padrão, mínimo, quartis e valores máximos

	INVERNO			VERÃO			PRIMAVERA			OUTONO		
	DirVento	Vmax	Vhora	DirVento	Vmax	Vhora	DirVento	Vmax	Vhora	DirVento	Vmax	Vhora
Count	28338	28338	28338	26673	26673	26673	27237	27237	27237	27655	27655	27655
Mean	150,83	4,78	1,81	167,34	4,94	1,77	166,33	5,37	1,98	164,13	4,55	1,69
Std	103,48	2,79	1,33	108,25	2,63	1,23	105,05	2,78	1,34	103,98	2,56	1,21
Min	1	0	0	1	0	0	1	0	0	1	0	0
25%	72	2,7	0,8	79	3	0,8	85	3,4	1	86	2,6	0,7
50%	116	4,7	1,7	131	4,9	1,7	133	5,4	1,9	127	4,5	1,6
75%	244	6,6	2,7	288	6,6	2,6	278	7,2	2,9	273	6,3	2,5
Max	360	22,5	8,3	360	24,3	8,5	360	24,3	9,3	360	18,2	8

Fonte: Elaboração própria.

Mesmo tendo a possibilidade de se analisar as três variáveis presentes nas tabelas, o artigo foca apenas nas velocidades máximas. A Tabela 2 foi utilizada no artigo de Brancalhão e Santos (2021) e chegou-se à conclusão de que:

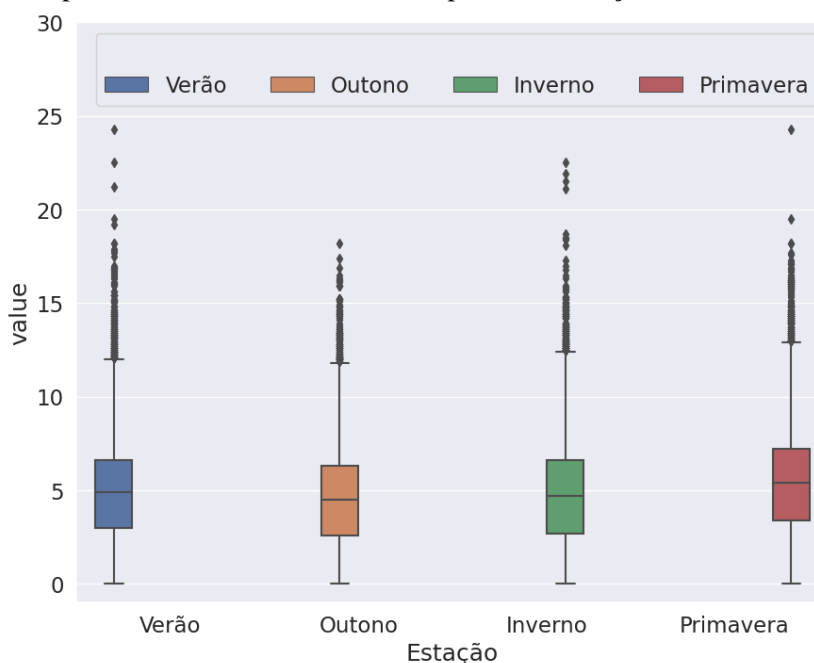
A análise de Vmax da Tabela 2 mostra um intervalo de variação amplo, entre 0 m/s e 24.30 m/s, e uma média que é distante do valor máximo por cerca de 7 desvios-padrão. Por sua vez, a velocidade horária é menos acentuada, mas a média permanece cerca de 5 desvios do valor máximo. Ambos os casos

indicam a presença de valores discrepantes baseados no critério do box-plot de Tukey (SANTOS *et al.*, 2021).

Comparando os dados gerais com as estações do ano em separado, verifica-se que o valor máximo de vento foi obtido no verão e na primavera, com intensidade de 24,3 m/s. Os intervalos de variações continuaram amplos entre os valores mínimos e máximos, e se mantiveram cerca de 7 desvios-padrões. Todas as estações do ano indicaram valores discrepantes baseados no mesmo critério do box-plot de Tukey analisado. Observa-se melhor a presença de pontos discrepantes da média de acordo com o limite superior CS do diagrama do box-plot nos gráficos 3 e 4.

Baseado nos dados de  $V_{max}$ , o limiar  $\varphi = 10,3$  m/s corresponde aos valores de rajada de vento com velocidades superiores ou iguais a 10,3 m/s.

**Gráfico 3** Box-plot dos máximos de velocidade para cada estação do ano



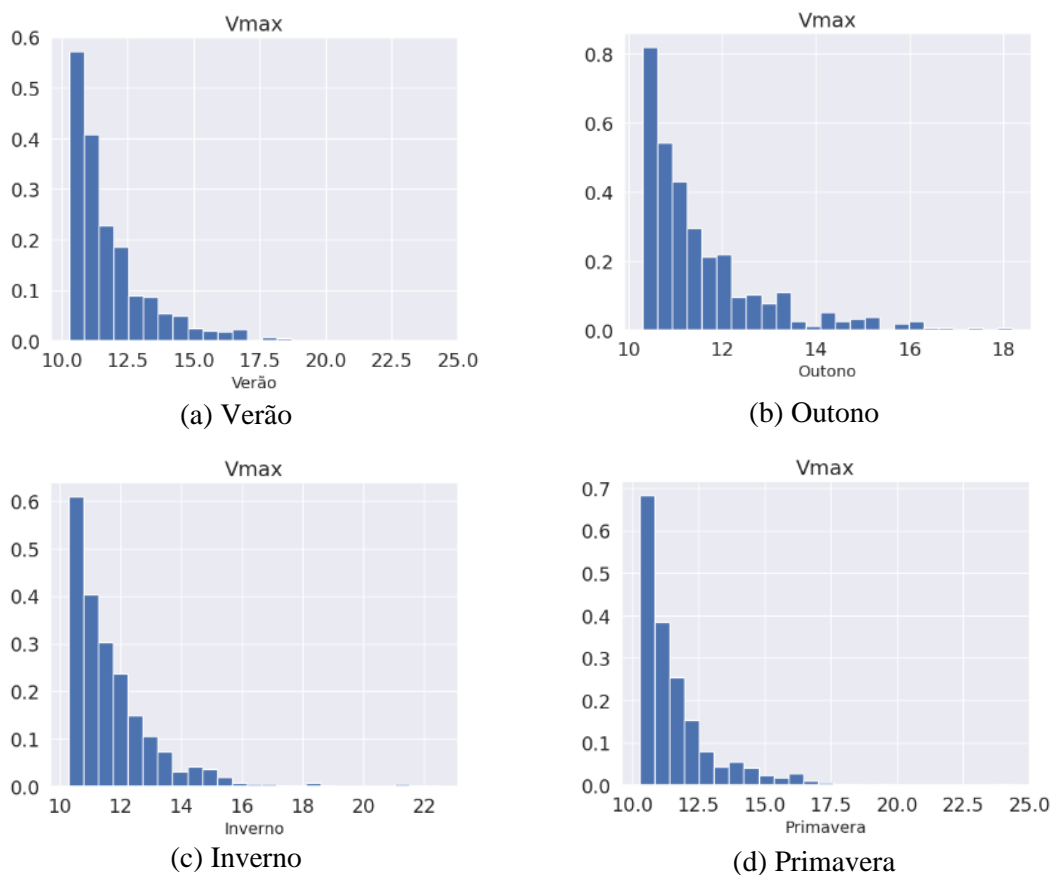
Fonte: Elaboração própria.

A partir da análise do box-plot percebe-se que o valor limiar  $\varphi$ , que está localizado na região superior do gráfico 3, restringe uma porção dos dados de  $V_{max}$  menor do que aquela delimitada pela constante CS. Esses valores acima do limiar, apesar de serem consideráveis, representam uma frequência relativa baixa, que pode ser melhor observada no gráfico 4.

Os resultados obtidos a partir das estimativas dos parâmetros das funções de densidade especificadas anteriormente são apresentados na Tabela 4, sendo o eixo “x” a velocidade

máxima e o eixo “y” a relação de ocorrência, de acordo com a classificação do valor p do teste de Kolmogorov-Smirnov.

**Gráfico 4** Histograma dos valores acima do limiar de velocidade máxima para as estações do ano



Fonte: Elaboração própria.

**TABELA 4** Tabela ordenada pelos valores p de algumas das curvas em estudo

VERÃO				PRIMAVERA			
#	Distribution	p-value	D-value	#	Distribution	p-value	D-value
1	pareto	0,33482	0,07043	1	pareto	0,61358	0,055736
2	gompertz	0,3015	0,07252	2	genextreme	0,26928	0,076
3	invweibull	0,22164	0,0783	3	invweibull	0,26917	0,076
OUTONO				INVERNO			
#	Distribution	p-value	D-value	#	Distribution	p-value	D-value
1	gompertz	0,6461	0,06557	1	pareto	0,56653	0,05988
2	pareto	0,6461	0,06557	2	gompertz	0,56653	0,05988
3	invweibull	0,31848	0,08533	3	invweibull	0,19947	0,08206

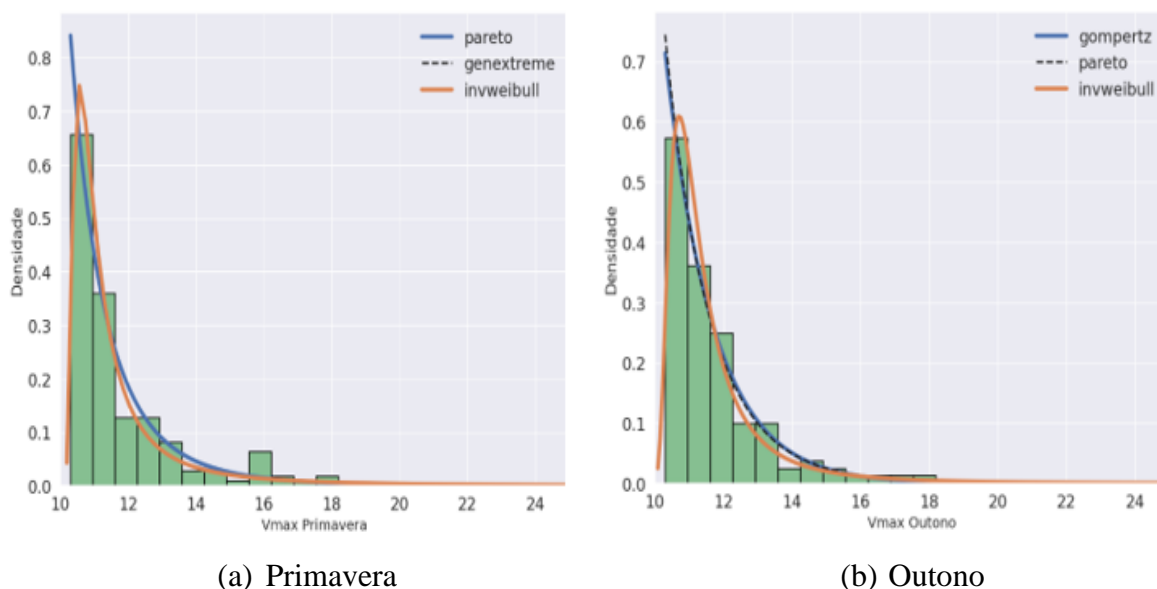
Fonte: Elaboração própria.

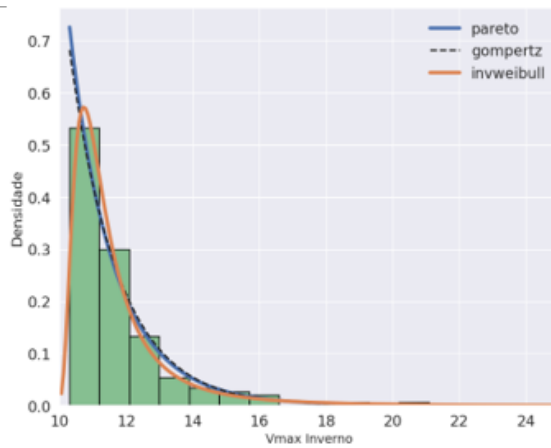
A análise dos resultados da Tabela 4 mostra que todos os valores  $p$  são tais que  $p \geq 0,05$ . Portanto, não há evidências para rejeitar a hipótese de que as distribuições sejam dos tipos especificados. Com base na análise dos valores  $p$  do teste de Kolmogorov-Smirnov, os maiores valores foram obtidos, respectivamente, para as funções de densidade que modelam a distribuição de velocidades de vento para outono, inverno, primavera e verão, respectivamente. As funções de densidade identificadas como mais adequadas, seguindo o critério de valor  $p$  ordenado em ordem crescente, correspondem à distribuição de Pareto para as estações verão, primavera e inverno e a distribuição de Gompertz para a estação outono.

Apesar da distribuição de Pareto ser indicada em três estações, os parâmetros respectivos são distintos, indicando que modelam diferentes magnitudes de vento. Estas diferentes magnitudes podem ser observadas na Tabela 3 que mostra as estatísticas descritivas de velocidades de vento de acordo com as estações.

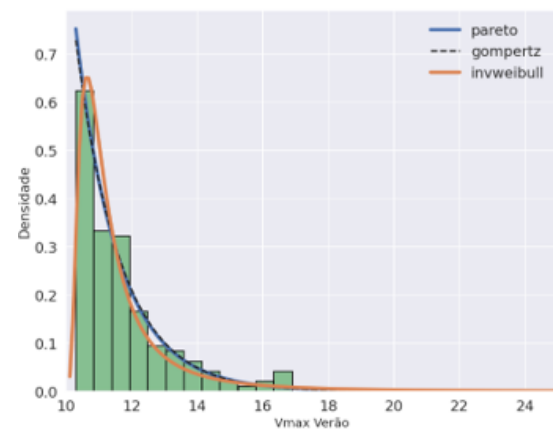
O gráfico 5 apresenta as funções de densidade ajustadas que melhor se adaptam aos dados de cada estação do ano conforme a Tabela 4.

**Gráfico 5** Funções de densidade para os valores máximos acima do limiar com parâmetros estimados





(c) Inverno

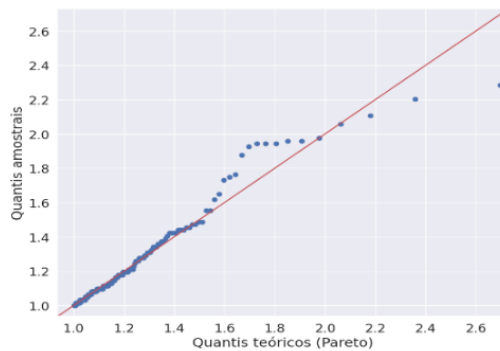


(d) Verão

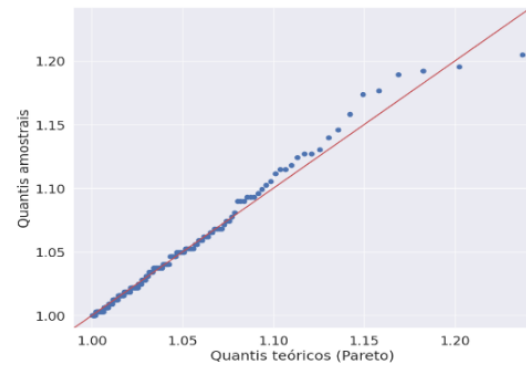
Fonte: Elaboração própria.

A seguir, estão os diagramas de probabilidade-probabilidade para as funções indicadas para cada estação do ano. Com base na análise, as funções de densidade não serão rejeitadas. O ajuste, em cada caso, possui desvios apenas nas caudas superiores das funções de densidade.

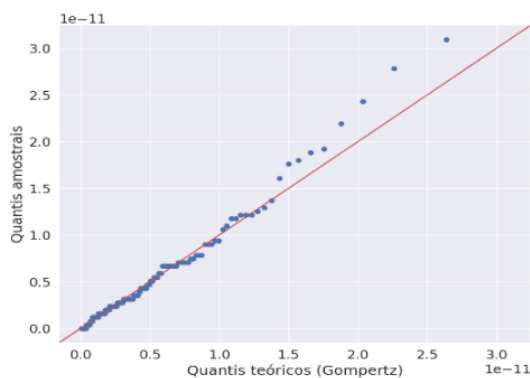
**Gráfico 6** Gráficos qq-plot das curvas com parâmetros estimados



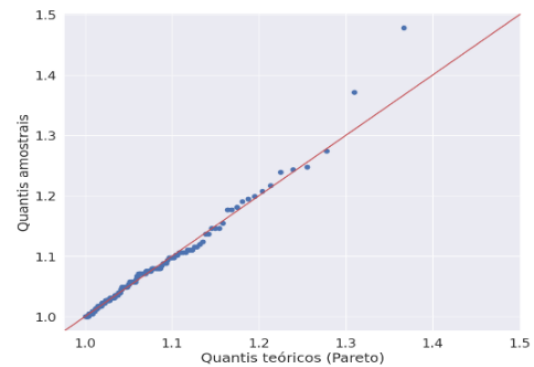
(a) Primavera



(b) Verão



(c) Outono



(d) Inverno

Fonte: Elaboração própria.

Os gráficos QQ-plot mostram razoável aderência à linha bissetriz do primeiro quadrante apresentando um desvio mais acentuado em (c) à esquerda.

Com base nas análises anteriores, os parâmetros estimados, com 2 casas decimais, para cada uma das distribuições são fornecidos na tabela a seguir. Essas curvas para as funções de densidade são descritas por meio das ordenadas “(b, loc, scale)”.

**Tabela 5** Parâmetros estimados das distribuições de probabilidade adequadas, segundo valor p, em cada uma das estações

	Função	b	loc	scale
Verão	Pareto	24,29	-21,95	32,25
Primavera	Pareto	5,17	4,15	6,14
Outono	gompertz	18,35	10,29	25,52
Inverno	Pareto	16,41	-12,29	22,59

Fonte: Elaboração própria.

É possível calcular as probabilidades de ocorrer um máximo de vento inferior ou superior a um determinado valor C, escolhido arbitrariamente, utilizando métodos estatísticos apropriados. Para ilustrar este cálculo, podemos considerar os intervalos de velocidade do vento em diferentes estações tal que a fórmula para calcular a probabilidade de o máximo de vento ( $V_{max}$ ) ser inferior a um determinado valor ( $valor_1$ ) utilizando a função de densidade de probabilidade com parâmetros estimados é dada por:

$$P(V_{max} \leq valor_1) = \int_{-\infty}^{valor_1} f(V_{max})dV_{max}$$

A variável  $V_{max}$  está no intervalo de menos infinito até  $valor_1$ , resultando em uma probabilidade no intervalo entre 0 e 1. Essa probabilidade é representada pela área abaixo da curva até o  $valor_1$  da função de densidade de probabilidade de Pareto.

Da mesma forma, para calcular a probabilidade de o máximo de vento ser superior a outro valor ( $valor_2$ ) é:

$$P(V_{max} \geq valor_2) = 1 - \int_{-\infty}^{valor_2} f(V_{max})dV_{max}$$

Nesse caso, subtraímos a probabilidade calculada (integral da função de densidade de probabilidade) de 1 para obter a probabilidade de o máximo de vento ser maior que  $valor_2$ . Para fins de cálculo, o pacote `scipy.stats` foi utilizado, por meio da função computacional `.ppf`, ou seja, `scipy.stats.pareto.ppf([0.05,0.95],b,loc,scale)` retorna a probabilidade da função de Pareto para os parâmetros  $b$ ,  $loc$  e  $scale$ . Analogamente se aplica para a função de gompertz.

A tabela a seguir, mostra os valores tais que ambas as probabilidades acima resultam em valor 0,05 para cada uma das estações do ano:

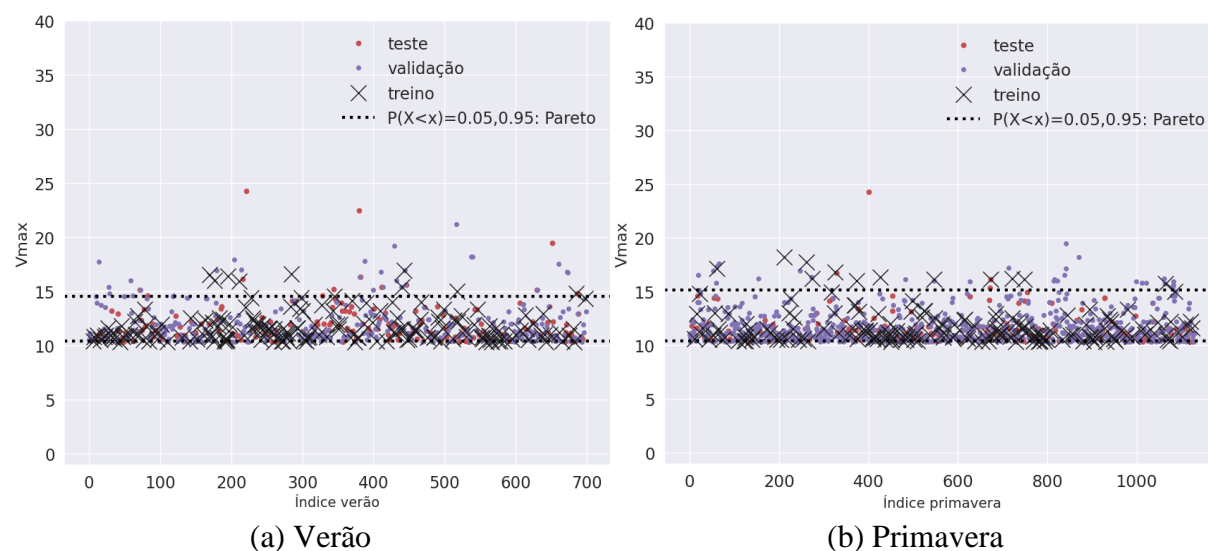
**Tabela 6** Probabilidade para os valores de velocidade

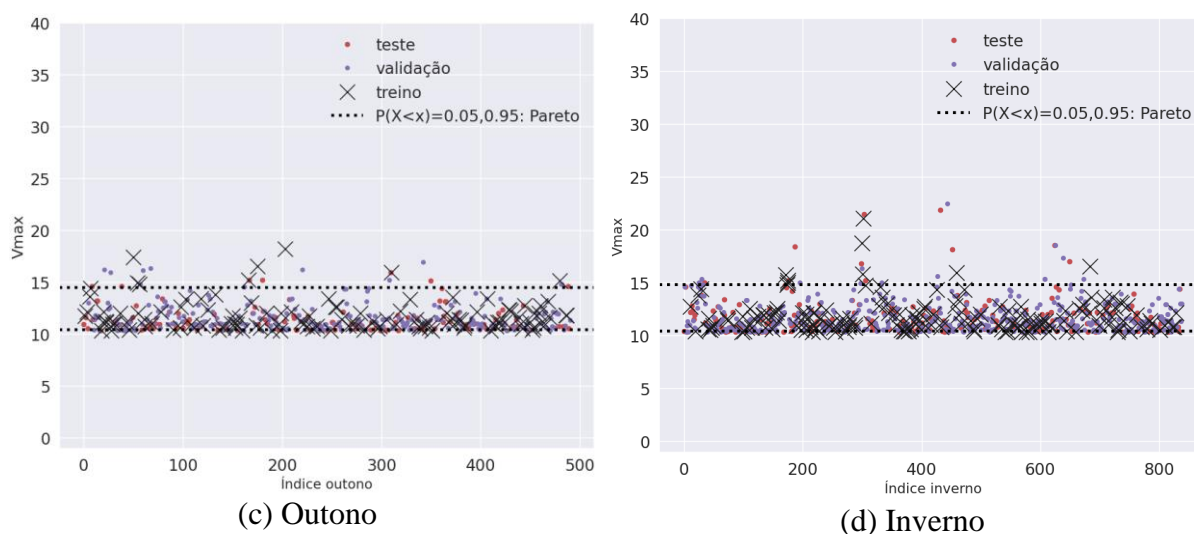
$P(valor1 < V_{max} < valor2) = 0,9$	Valor1	Valor2
Verão	10,44	14,53
Primavera	10,42	15,11
Outono	10,44	14,49
Inverno	10,44	14,82

Fonte: Elaboração própria.

O gráfico de dispersão apresentado a seguir demonstra as retas horizontais traçadas com base nos quantis, juntamente com os dados dos conjuntos de treinamento, validação e teste

**Gráfico 7** Dados de treinamento e teste sobrepostos às retas definidas pelos percentis 5 e 95, respectivamente





Fonte: Elaboração própria.

Como citado na metodologia, os dados analisados foram divididos em três conjuntos: treino, validação e teste, sendo o conjunto de treino utilizado para estimar os parâmetros e os conjuntos de validação e teste recomendados para análises posteriores.

A observação visual do gráfico sugere que existe uma concordância considerável entre os conjuntos de dados de teste e validação. Isso significa que os intervalos percentis fornecem uma ideia razoável do comportamento dos dados.

A comparação dos dados gerais com as estações do ano, as análises gráficas e os valores  $p$  e  $D$  indica uma concordância satisfatória entre os dados de teste e validação, com os intervalos percentis fornecendo uma boa base do comportamento dos dados.

Finalmente, após a obtenção das curvas de densidade, é possível abordar um outro ponto importante: a análise dos tempos de retorno, risco de ocorrência de ventos extremos. O tempo de retorno é um conceito utilizado em análises de risco e probabilidade para descrever a frequência esperada de ocorrência de um evento específico em um determinado intervalo de tempo.

O período de retorno é definido como o inverso da probabilidade ser igualada ou superada, ou seja,  $T=1/p$ , em que  $p$  é definido como a última integral, ou seja, associada ao  $valor_2$ . Assim, o tempo de retorno para a velocidade de 10,44 m/s e estação verão é 1/0,95, ou seja, aproximadamente 1 ano; por outro lado o período de retorno para 14,53m/s é de 20 anos. Períodos de retorno mais longos podem ser determinados da mesma forma, porém não são objeto de estudos do artigo. Analogamente, os períodos de retorno para as demais estações

podem ser determinados. Assim, de posse das curvas de densidade, outros estudos podem ser efetuados com finalidade diversa.

#### **4 CONSIDERAÇÕES FINAIS**

Os dados de rajadas de vento foram submetidos a um processo de ajuste utilizando um conjunto específico de funções de densidade de probabilidade. A escolha dessas funções levou em consideração as sugestões encontradas nas referências bibliográficas que indicaram quais funções melhores se encaixam em estudos meteorológicos. A análise dos dados foi conduzida através da aplicação do método dos máximos acima de um limiar, utilizando bibliotecas Python que oferecem recursos para análise, ajuste e visualização dos resultados de forma gráfica. Os resultados foram realizados no ambiente Colab.

Foram utilizadas as seguintes funções de densidade de probabilidade: Valor Extremo Generalizado (GEV), Gompertz, Gumbel e Fréchet, Pareto, Weibull Máximo e Weibull invertida (também conhecida como Fréchet ou Tipo II GEV). Essas funções foram selecionadas para realizar o ajuste dos dados de rajadas de vento.

O principal propósito do artigo consistiu em avaliar um conjunto de dados referentes à velocidade do vento em relação aos valores máximos utilizando a abordagem acima de um limiar. Além disso, buscou-se realizar uma comparação sazonal a fim de identificar a estação do ano que apresentou a maior ocorrência de ventos intensos. Tais objetivos foram plenamente alcançados e delineados por meio de uma análise preliminar seguida de uma análise descritiva. Por meio da extração das informações amostrais, foi possível constatar que, pelos comprimentos dos vetores, a estação do verão se destacou como aquela que registrou a maior incidência de ventos superiores a 10,3 m/s durante os últimos 10 anos.

No contexto da instrução aérea no T-27, é importante destacar que a estação do verão pode ser especialmente desafiadora, devido às condições de máximas de ventos observadas. Essa situação pode levar a possíveis dificuldades de desempenho por parte dos Cadetes.

Com os resultados alcançados, torna-se possível uma preparação mais adequada por parte dos Cadetes em relação à incidência de vento durante a execução dos exercícios nas áreas de instrução. Além disso, é recomendável realizar uma análise meteorológica mais minuciosa antes dos voos, especialmente durante esse período do ano, a fim de mitigar os riscos associados à intensidade do vento durante a instrução aérea. Essas medidas visam aumentar a segurança e

a eficiência das operações aéreas, garantindo uma abordagem mais precisa e informada em relação às condições de vento.

De maneira ampla, o presente artigo pode ser relacionado às aulas de cálculo e estatísticas ministradas na AFA como um exemplo a ser seguido na abordagem dessas disciplinas. Além disso, o trabalho oferece uma base sólida para a realização de novas pesquisas no campo da estatística, explorando diversos fenômenos que impactam a atividade aérea. Essas contribuições proporcionam perspectivas promissoras para o aprofundamento dos estudos estatísticos nesse contexto, fornecendo dados valiosos para o entendimento e aprimoramento das operações aéreas.

## REFERÊNCIAS

- ALAM, S.; KHAN, M. S. M. Statistical characterization of extreme hydrologic parameters for the peripheral river system of Dhaka City. **Journal of Water Resources and Ocean Science**, v. 3, n. 3, p. 30-37, 2014.
- BAUTISTA, E. A. L.; ZOCCHI, S. S.; ANGELOCCI, L. R. A distribuição generalizada de valores extremos aplicada ao ajuste dos dados de velocidade máxima do vento em Piracicaba, São Paulo, Brasil. **Revista Matemática e Estatística**, v. 22, n. 1, p. 95-111, 2004.
- BRASIL. M. D. Comando da Aeronáutica. Academia da Força Aérea. **Manual de Procedimentos do 1º EIA**. Pirassununga, 2022.
- BRASIL. M. D. Comando da Aeronáutica. Academia da Força Aérea. **Programa de Instrução e Manutenção Operacional da AFA**. Pirassununga, 2022.
- FERREIRA, E. T. **Estudo comparativo entre a velocidade básica do vento estabelecida pela NBR. 6123 e a obtida de estações meteorológicas na Paraíba - Impactos nos âmbitos do projeto estrutural, do meio ambiente e dos custos**. Dissertação (Mestrado) — Universidade Federal da Paraíba, João Pessoa-PB, 2005.
- FISHER, R. A.; TIPPETT, L. H. C. Limiting forms of the frequency distribution of the largest or smallest member of a sample. In: **Mathematical proceedings of the Cambridge philosophical society**, v. 24, n. 2, p. 180-190. Cambridge University Press, 1928.
- FUENTES, E. V.; BITENCOURT, D. P.; FUENTES, M. V. Análise da velocidade do vento e altura de onda em incidentes de naufrágio na costa brasileira entre os estados do Sergipe e do Rio Grande do Sul. **Revista Brasileira de Meteorologia**, v. 28, p. 257-266, 2013.
- Fundação Oswaldo Cruz. **Estações do ano**. Disponível em: <http://www.fiocruz.br/biosseguranca/Bis/infantil/estacoes-ano.htm>. Acesso em: 13, maio, 2023.
- GIL, A. C. **Métodos e Técnicas de Pesquisa Social**. 7.ed. São Paulo: Atlas, 2019.
- GIPE, P. **Wind energy comes of age**. John Wiley & Sons, 1995.
- GUMBEL, E. J. **Statistics of extremes**. New York: Columbia University Press, p. 375, 1958.
- HECKERT, A. S.; SIMIU, E.; YEO, D. **Extreme Winds and Wind Effects on Structures**. 2020. Disponível em: <https://www.itl.nist.gov/div898/winds/overview.htm>. Acesso em: 11, setembro, 2022.
- INSTITUTO NACIONAL DE PADRONIZAÇÃO E TECNOLOGIA. **Autocorrelation**. Disponível em: <https://www.itl.nist.gov/div898/handbook/eda/section3/autocopl.htm>. Acesso em: 10, junho, 2023.

JENKINSON, A. F. The frequency distribution of the annual maximum (or minimum) values of meteorological elements. **Quarterly Journal of the Royal Meteorological Society**, v. 81, n. 348, p. 158-171, 1955.

KENDALL, M. G. et al. The advanced theory of statistics. **The advanced theory of statistics**, n. 2nd Ed, 1946.

KOTZ, S.; NADARAJAH, S. **Extreme value distributions: theory and applications**. World Scientific Publishing Company, 2000.

LEOTTI, V. B.; COSTER, R.; RIBOLDI, J. Normalidade de variáveis: métodos de verificação e comparação de alguns testes não-paramétricos por simulação. **Revista HCPA**. Porto Alegre. Vol. 32, n. 2 (2012), p. 227-234, 2012.

LISKA, G. R. et al. Estimativas de velocidade máxima de vento em Piracicaba-SP via Séries Temporais e Teoria de Valores Extremos. **Revista Brasileira de Biometria**, v. 2, p. 295-309, 2013.

MONTGOMERY, D. C.; RUNGER, G. C. **Applied statistics and probability for engineers**. John Wiley & Sons, 2010.

MORETTIN, P. A.; BUSSAB, W. O. **Estatística básica**. São Paulo: Saraiva, 2017.

PAPALEXIOU, S. M.; KOUTSOYIANNIS, D. Battle of extreme value distributions: A global survey on extreme daily rainfall. **Water Resources Research**, v. 49, n. 1, p. 187-201, 2013. Disponível em: <https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2012WR012557>. Acesso em: 06, setembro, 2022.

ROWINSKI, P. M.; STRUPCZEWSKI, W. G.; SINGH, V. P. A note on the applicability of log-Gumbel and log-logistic probability distributions in hydrological analyses: I. Know pdf. **Hydrological Science Journal**, Walingford, Inglaterra, v.47, n.1, p.107-122, 2002.

SANT'ANNA, C. L. M. **Estimativa de quantis extremos de vento como suporte ao dimensionamento de edificações**. 2018. [Trabalho de Conclusão de Curso]. Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2018. Disponível em: <http://repositorio.poli.ufrj.br/monografias/monopoli10023818.pdf>. Acesso em: 12, janeiro, 2023.

SANTOS, J. P. M.; BRANCALHÃO, S. É.; DE JESUS, A. F. Ajuste de Distribuições de Valores Extremos em Velocidade de Vento. **Proceeding Series of the Brazilian Society of Computational and Applied Mathematics**, v. 8, n. 1, 2021.

SÓRIA, M. H. A. **Orientação em pistas**. 2003. Disponível em: <http://repositorio.eesc.usp.br/items/fd9614b1-46ab-4db6-a51b-26ec4c3ce048>. Acesso em: 15, fevereiro, 2023.