



UNIVERSIDADE  
FEDERAL RURAL  
DE PERNAMBUCO



Kimbelly Emanuelle Avelino Ferraz

**Processo de Renovação Generalizado baseado  
na distribuição Gumbel como modelo de  
estimativas de ocorrências de ondas de calor  
para auxiliar no processo de tomada de decisão  
do cultivo de manga no Sertão de Pernambuco**

Recife

2023

Kimbelly Emanuelle Avelino Ferraz

**Processo de Renovação Generalizado baseado na  
distribuição Gumbel como modelo de estimativas de  
ocorrências de ondas de calor para auxiliar no processo  
de tomada de decisão do cultivo de manga no Sertão de  
Pernambuco**

Monografia apresentada ao Curso de Bacharelado em Sistemas de Informação da Universidade Federal Rural de Pernambuco, como requisito parcial para obtenção do título de Bacharel em Sistemas de Informação.

Universidade Federal Rural de Pernambuco – UFRPE

Departamento de Estatística e Informática

Curso de Bacharelado em Sistemas de Informação

Orientador: Cláudio Tadeu Cristino

Coorientador: -

Recife

2023

Dados Internacionais de Catalogação na Publicação  
Universidade Federal Rural de Pernambuco  
Sistema Integrado de Bibliotecas  
Gerada automaticamente, mediante os dados fornecidos pelo(a) autor(a)

---

F381p Ferraz, Kimbelly Emanuelle Avelino  
Processo de Renovação Generalizado baseado na distribuição Gumbel como modelo de estimativas de ocorrências de ondas de calor para auxiliar no processo de tomada de decisão do cultivo de manga no Sertão de Pernambuco / Kimbelly Emanuelle Avelino Ferraz. - 2023.  
66 f. : il.

Orientador: Claudio Tadeu Cristino.  
Inclui referências.

Trabalho de Conclusão de Curso (Graduação) - Universidade Federal Rural de Pernambuco, Bacharelado em Sistemas da Informação, Recife, 2023.

1. ondas de calor. 2. Processo de Renovação Generalizado. 3. Gumbel. 4. GuGRP. 5. manga. I. Cristino, Claudio Tadeu, orient. II. Título

---

CDD 004

Kimbelly Emanuelle Avelino Ferraz

**Processo de Renovação Generalizado baseado na distribuição Gumbel  
como modelo de estimativas de ocorrências de ondas de calor para  
auxiliar no processo de tomada de decisão do cultivo de manga no  
Sertão de Pernambuco**

Monografia apresentada ao Curso de Bacharelado em Sistemas de Informação da Universidade Federal Rural de Pernambuco, como requisito parcial para obtenção do título de Bacharel em Sistemas de Informação.

Aprovada em: 08 de Maio de 2023.

**BANCA EXAMINADORA**

Orientador: Cláudio Tadeu Cristino  
Departamento de Estatística e Informática  
Universidade Federal Rural de Pernambuco

Marcelo Gama da Silva  
Departamento de Estatística e Informática  
Universidade Federal Rural de Pernambuco

# Agradecimentos

Agradeço a Deus por estar sempre guiando meus passos e ser minha força nos desafios encontrados em toda a trajetória e por me conceder persistência.

À minha família e amigos.

Agradeço ao meu orientador por ser paciente e me orientar na dissertação, mesmo com todas as dificuldades, sempre me incentivando.

A todas as pessoas e professores que encontrei e que, de alguma forma, contribuíram para o meu aprendizado.

*“A persistência é o caminho do êxito.”*  
*(Charles Chaplin)*

# Resumo

Diversos tipos de eventos podem prejudicar o plantio, a colheita ou manejo das plantas e frutos nas áreas agrícolas, sendo um deles inclusive o evento chamado ondas de calor, que se caracteriza como um fenômeno meteorológico prolongado e relativamente incomum com temperaturas extremamente altas para região e persistentes durante vários dias ou até semanas. Dada a importância da agricultura, este trabalho busca por meio da análise dos dados de temperatura máxima da região de Petrolina, do estudo da plantação da manga, do evento da Onda de Calor por meio do percentil 90, de algoritmos de otimização e dos processos de renovação generalizados e Gumbel, estimar esse evento contribuindo para a tomada de decisão do agricultor e na otimização da produção da Manga. O modelo proposto usa o processo de renovação generalizado baseado na distribuição Gumbel (GuGRP) para modelar os intervalos de tempo entre as ondas de calor, considerando que os eventos consecutivos são condicionalmente independentes. Esse modelo se mostrou aderente para modelar os eventos com nível de significância de 0,05 e o  $P - Value$  de 0,28 por meio do teste de aderência de Kolmogorov-Smirnov nos dados de adequação adaptados ao GuGRP. Os parâmetros do modelo foram estimados pelo Log-Verossimilhança utilizando algoritmos de otimização, também testando especificamente o algoritmo de Enxame de partículas.

**Palavras-chave:** ondas de calor, processos de renovação generalizados, Gumbel, GuGRP, manga, agricultura, algoritmo, otimização

# Abstract

Several types of events can harm the planting, harvesting or handling of plants and fruits in agricultural areas, one of them including the event called heat waves, which is characterized as a prolonged and relatively uncommon meteorological phenomenon with extremely high temperatures for the region and persistent for several days or even weeks. Given the importance of agriculture, this work seeks, through the analysis of the maximum temperature data in the Petrolina region, the study of the mango plantation, the Heat Wave event through the 90th percentile, optimization algorithms and the processes of generalized renewal and Gumbel, estimating this event contributing to the farmer's decision making and optimization of Mango production. The proposed model uses the generalized renewal process based on the Gumbel distribution (GuGRP) to model the time intervals between heat waves, considering that consecutive events are conditionally independent. This model proved to be adherent to model events with a significance level of 0.05 and a  $P - Value$  of 0.28 through the Kolmogorov-Smirnov adherence test on the adequacy data adapted to the GuGRP. The model parameters were estimated by Log-Likelihood using optimization algorithms, also specifically testing the Particle Swarm algorithm.

**Keywords:** heat waves, generalized renewal processes, Gumbel, GuGRP, mango, agriculture, profit, production, algorithm, optimization.

# Listas de ilustrações

Figura 1 – Região do Vale do São Francisco . . . . .	24
Figura 2 – Linha do tempo do PR e PRG. . . . .	28
Figura 3 – Representação dos eventos de interesse (ondas de calor) no modelo GuGRP. . . . .	33
Figura 4 – Possível representação da ocorrência de ondas de calor. Fonte: a autora.	33
Figura 5 – Analogia de um bando de pássaros em um espaço de busca de 3D. Considerando um problema de maximização, o pássaro azul é a partícula de solução ótima do problema. . . . .	38
Figura 6 – Fluxograma do algoritmo de Enxame de partículas. . . . .	39
Figura 7 – Representação de um gráfico <i>box plot</i> , $Q_1$ é o primeiro quartil, a mediana é o segundo quartil, $Q_2$ , e $Q_3$ é o terceiro quartil. $L_i$ e $L_s$ limites inferior e superior, respectivamente. $RIC$ é a região do intervalo de confiança. . . . .	41
Figura 8 – Temperaturas médias máximas de cada mês. . . . .	45
Figura 9 – Box plot de temperaturas máximas de todos os meses e da série histórica entre 2008-2022. . . . .	46
Figura 10 – Histograma das temperaturas máximas médias. . . . .	47
Figura 11 – Gráfico de pizza de temperaturas máximas que ultrapassaram as normais climatológicas separadas por estação . . . . .	48
Figura 12 – Gráfico de barras da intensidade em graus da Onda de calor no decorrer dos anos . . . . .	49
Figura 13 – Gráfico de barras da duração da Onda de calor no decorrer dos anos . . . . .	50
Figura 14 – Gráfico comparando os dados transformados com uma distribuição exponencial com parâmetro 1 . . . . .	52
Figura 15 – Box plot das estimativas da distribuição x o que de fato ocorreu . . . . .	53
Figura 16 – Box plot do preço de Venda - Manga Tommy. . . . .	54
Figura 17 – Box plot do preço de Venda - Manga Palmer. . . . .	54

# **Lista de tabelas**

Tabela 1 – Descrição básica dos dados de temperaturas máximas da série histórica analisada. . . . .	46
Tabela 2 – Parâmetros estimados . . . . .	51
Tabela 3 – Tabela do teste de aderência dos dados transformados na distribuição exponencial com parâmetro 1. . . . .	51

# **Lista de abreviaturas e siglas**

OMM	Organização Meteorológica Mundial
IPCC	Painel Intergovernamental sobre Mudanças Climáticas
LabMet	Laboratório de Meteorologia
Univasf	Universidade Federal do Vale do Rio São Francisco
Embrapa	Empresa Brasileira de Pesquisa Agropecuária
Abrafrutas	Associação Brasileira dos Produtores e Exportadores de Frutas e Derivados
Adagro	Agência de Defesa e Fiscalização Agropecuária do Estado de Pernambuco
Inmet	Instituto Nacional de Meteorologia
MET OFFICE	Serviço Nacional de Meteorologia do Reino Unido
IPMA	Instituto Português do Mar e da Atmosfera
Cepea	Centro de Estudos Avançados em Economia Aplicada
GEV	Distribuição de valor extremo generalizado
FDP	Função de Densidade de Probabilidade
CDF	Função de distribuição cumulativa
PDF	Função densidade de probabilidade
EDF	Função de distribuição empírica
MTBF	Tempo Médio entre Falhas
PR	Processo de Renovação
PRG	Processo de Renovação Generalizado
VA	Variável Aleatória
EMV's	Estimadores de Máxima Verossimilhança
E.M.V	Estimador de Máxima Verossimilhança

Max	Máximo
Min	Minimo
MM	Método dos Momentos
MV	Método de máxima verossimilhança
Q	Período de Quinze dias
GuGRP	Modelo que usa um processo de renovação generalizada baseado na distribuição Gumbel
PSO	Particle swarm optimization ou enxame de partículas em português
teste KS	Teste Kolmogorov–Smirnov
K2	Teste do Qui-Quadrado
Outliers	Em português significa valores atípicos.

# Sumário

<b>Lista de ilustrações . . . . .</b>	<b>6</b>
<b>1 INTRODUÇÃO . . . . .</b>	<b>13</b>
<b>1.1 Motivação e Justificativa . . . . .</b>	<b>16</b>
<b>1.2 Objetivos . . . . .</b>	<b>16</b>
<b>2 TRABALHOS RELACIONADOS . . . . .</b>	<b>18</b>
<b>3 FUNDAMENTAÇÃO TEÓRICA . . . . .</b>	<b>20</b>
<b>3.1 Manga . . . . .</b>	<b>20</b>
3.1.1 Mercado brasileiro da manga . . . . .	21
3.1.2 Etapas da produção da manga . . . . .	21
3.1.3 Condições ideais de temperaturas e como altas temperaturas podem afetar a manga . . . . .	22
3.1.4 Importância da manga para o Vale do São Francisco e Petrolina . . . . .	23
<b>3.2 Época de floração colheita da manga - Brasil . . . . .</b>	<b>25</b>
<b>3.3 Ondas de calor . . . . .</b>	<b>25</b>
<b>3.4 Normais climatológicas . . . . .</b>	<b>27</b>
<b>3.5 Percentil . . . . .</b>	<b>27</b>
<b>3.6 Separação em treinamento e teste . . . . .</b>	<b>27</b>
<b>3.7 Processo de renovação generalizado . . . . .</b>	<b>28</b>
<b>3.8 Distribuição de probabilidade . . . . .</b>	<b>30</b>
3.8.1 Termos PDF, CDF e EDF . . . . .	30
3.8.2 Teoria dos valores extremos . . . . .	30
3.8.3 A distribuição Gumbel . . . . .	31
<b>3.9 Processo de Renovação Generalizada baseada na distribuição Gumbel (GuGRP) . . . . .</b>	<b>32</b>
3.9.1 Teste de aderência para o GuGRP . . . . .	34
<b>3.10 Técnicas de estimação de parâmetros . . . . .</b>	<b>34</b>
3.10.1 Método de Máxima Verossimilhança . . . . .	35
3.10.2 Log-Verossimilhança . . . . .	36
<b>3.11 Algoritmos de otimização . . . . .</b>	<b>37</b>
3.11.1 Algoritmo de enxame de partículas . . . . .	37
<b>3.12 Testes de aderência . . . . .</b>	<b>39</b>
3.12.1 Teste de Kolmogorov-Smirnov . . . . .	40
<b>3.13 Box plot . . . . .</b>	<b>40</b>

<b>3.14</b>	<b>Definição de Risco . . . . .</b>	<b>41</b>
<b>4</b>	<b>METÓDOS E FERRAMENTAS . . . . .</b>	<b>42</b>
<b>4.1</b>	<b>Seleção da base de dados . . . . .</b>	<b>42</b>
<b>4.2</b>	<b>Pré-processamento de Dados . . . . .</b>	<b>42</b>
<b>4.3</b>	<b>Processamento dos dados . . . . .</b>	<b>43</b>
4.3.1	Análises iniciais de temperaturas . . . . .	43
4.3.2	Análises iniciais do Preço de Venda da Manga - Brasil . . . . .	43
4.3.3	Cálculo para as probabilidades de ocorrência de Ondas de Calor . . . . .	43
4.3.3.1	Contagem de Ondas de Calor que ocorreram em Petrolina . . . . .	43
4.3.3.2	Separação de dados em treinamento e teste . . . . .	43
4.3.3.3	Utilização - Gumbel, PRG, MV, Teste KS . . . . .	43
4.3.4	Cruzamento das estimativas do Evento com os investimentos e preços da Manga . . . . .	44
<b>5</b>	<b>RESULTADOS E DISCUSSÕES . . . . .</b>	<b>45</b>
<b>5.1</b>	<b>Análises iniciais . . . . .</b>	<b>45</b>
5.1.1	Observações importantes . . . . .	45
5.1.2	Gráfico de colunas de temperaturas médias máximas . . . . .	45
5.1.3	Desvio padrão, média e mediana dos dados . . . . .	45
5.1.4	Box plot das temperaturas . . . . .	46
5.1.5	Histogramas das temperaturas . . . . .	47
5.1.6	Dias com temperatura máxima acima da normal climatológica . . . . .	47
5.1.7	Gráficos de intensidade e duração de uma Onda de Calor . . . . .	48
5.1.7.1	Intensidade . . . . .	49
5.1.7.2	Duração . . . . .	50
<b>5.2</b>	<b>Estimativas de temperaturas . . . . .</b>	<b>50</b>
5.2.1	Aplicação do Método de Máxima Verossimilhança e Log-Verossimilhança . .	50
5.2.2	Utilização do algoritmo de otimização enxame de partículas . . . . .	50
5.2.3	Parâmetros estimados . . . . .	51
5.2.4	Teste de aderência - Teste de Kolmogorov-Smirnov . . . . .	51
5.2.4.1	Gráfico da aderência . . . . .	52
5.2.5	Box plot estimativas . . . . .	52
5.2.5.1	Box plot real x treinamento . . . . .	52
<b>5.3</b>	<b>Análises básicas do mercado da Manga . . . . .</b>	<b>53</b>
5.3.1	Análise de preço - Manga Tommy . . . . .	53
5.3.2	Análise de preço - Manga Palmer . . . . .	54
<b>6</b>	<b>POSSÍVEIS SITUAÇÕES - TEMPERATURAS ELEVADAS X PREÇOS DE MERCADO POR MÊS . . . . .</b>	<b>55</b>

7	<b>CONCLUSÃO</b>	56
---	------------------	----

<b>REFERÊNCIAS</b>	58
--------------------	----

# 1 Introdução

Um dos frutos mais apreciados pelo mercado interno e internacional, a manga, foi a fruta mais exportada pelo Brasil em 2021, as informações são da Associação Brasileira dos Produtores e Exportadores de Frutas e Derivados ([ABRAFRUTAS, 2022](#)). Grande responsável por este dado, o Vale do São Francisco, entre os estados de Pernambuco e Bahia, produziu 87% do volume embarcado ([GLOBO RURAL, 2021](#)).

A alta produção de Manga na região do Vale do São Francisco teve como contribuinte o clima, que é favorável a sua produção. Em entrevista ao CORREIO, o pesquisador da Embrapa Semiárido, João Ricardo Ferreira de Lima foi questionado sobre até que ponto as condições climáticas do Vale do São Francisco e da região sudoeste da Bahia favorecem o cultivo de manga. Em resposta, o pesquisador João Ricardo respondeu :

“Estas duas regiões possuem condições climáticas extremamente favoráveis, como a alta radiação solar, a luminosidade, baixa pluviosidade, que leva a menores problemas com praga e doenças. Estes fatores, em união com um conjunto de tecnologias disponibilizadas ao produtor, fazem com que se possa produzir mangas de elevado padrão de qualidade ao longo de todo o ano, de forma escalonada. Assim, é possível abastecer o mercado em qualquer período do ano.” ([ABRAFRUTAS, 2019](#))

Embora a Manga seja um fruto bastante adaptável ao clima de Petrolina com alta radiação solar, as condições ideais para seu cultivo são temperaturas médias entre 24°C a 30°C ([EMBRAPA, 2014](#)). Em condições não rotineiras e ideais, a Mangueira tolera temperaturas entre -6°C e 40°C, existindo variedades que suportam até 48°C mas em etapas específicas da produção da manga. Fora deste intervalo, ocorrem danos em plantas adultas e morte de plantas jovens ([RODRIGUES; SOUZA; LIMA, 2013](#)) ([CAMPBELL R.J.; CAMPBELL, 2006](#)) ([RAMÍREZ, 2010](#)). Por exemplo, temperaturas acima de 33 °C prejudicam a etapa da formação do grão de pólen, reduzindo em 50% sua viabilidade ([RODRIGUES; SOUZA; LIMA, 2016](#)). Temperaturas elevadas também pode provocar altas taxas de partenocarpia (frutos que se desenvolvem sem o embrião), originando frutos pequenos e sem valor comercial ([FILHO; TEIXEIRA, 2021](#)).

Na etapa de frutificação da mangueira por exemplo, sendo a etapa de formação e crescimento dos frutos, temperaturas acima ou muito abaixo da faixa recomendada são prejudiciais ao crescimento, ao desenvolvimento e à frutificação ([COSTA et al., 2014](#)). Temperaturas elevadas também facilitam o aparecimento de tripés e outras doenças.

O cultivo da Manga portanto, tem alta relação com as temperaturas da região, sendo sensível a eventos raros como por exemplo inundações e Ondas frio ou de calor. Esse

último evento se caracteriza como um fenômeno climático prolongado com temperaturas extremamente altas para a época e região observada, persistindo durante vários dias ou até semanas. Para a identificação desse evento cada agência meteorológica possui seus critérios, não possuindo uma fórmula literalmente global para tal identificação. Em ([OFFICE, 2022](#)), por exemplo, uma Onda de Calor é definida como temperaturas máximas diárias elevadas para à época e região, considerando limiares definidos, que tenham duração de no mínimo 3 dias consecutivos. Já pela Organização Meteorológica Mundial, esse evento se caracteriza quando num intervalo de pelo menos seis dias, a temperatura máxima diária é 5°C superior ao valor médio diário do período de referência ([PETERSON, 2001](#)). Em ([PERKINS; ALEXANDER, 2013](#)) utilizando percentil 90 nas temperaturas máximas da região para a caracterização do evento.

Ainda sobre temperaturas elevadas, recentemente aconteceram temperaturas não usuais na região de Petrolina em setembro de 2021, chegando a temperaturas máximas de 41°C ([FOLHAPE, 2021](#)). Infelizmente eventos como este tendem a ser cada vez mais comuns, confirmado inclusive pelo secretário-geral da OMM, Petteri Taalas, ao dizer que as ocorrências de Ondas de Calor “serão normais e, inclusive, mais fortes” ([CMJORNAL, 2022](#)). Segundo o relatório publicado pelo IPCC (Painel Intergovernamental sobre Mudanças Climáticas), a frequência de certos tipos de eventos climáticos extremos está crescendo devido às alterações climáticas globais antrópicas do ser humano ([IPCC, 2022](#)). Por estar cada vez mais acontecendo, as altas temperaturas estão ocasionando incêndios e mortes na região da Europa ([CMJORNAL, 2022](#)) e frutos danificados ou que vegetaram antes de florir, como ocorreu na safra em Monte Alto/Taquaritinga (SP) com a manga Palmer ([ABRAFRUTAS, 2020](#)).

Tendo em vista cada vez mais a possibilidade do aparecimento de eventos extremos, ocorre a necessidade de aprimorar e estabelecer sistemas de alertas para minimizar impactos negativos, fato citado inclusive pela ([OMS, 2015](#)). Buscando portanto entender e predizer eventos como esses, este trabalho testou a aplicação de uma distribuição da teoria dos extremos, mais propriamente a de Gumbel em conjunto com o Processo de Renovação Generalizada, em estimações e previsões de temperaturas elevadas e Ondas de calor na região do Vale do São Francisco, em Petrolina. Buscando-se com esse possível modelo, apoiar o agricultor da manga na decisão da melhor época de plantio, de venda ou cuidados especiais com o fruto.

Para a realização das análises para as previsões foi escolhido o banco de dados de temperaturas máximas disponibilizadas pelo INMET, Instituto Nacional de Meteorologia em detrimento do Laboratório de Meteorologia da Universidade Federal do Vale do São Francisco, por possuir mais dados. Atrelado a isso, foi utilizado o método de máxima verossimilhança e Log-verossimilhança para estimar os parâmetros do processo de renovação generalizada baseada em Gumbel, combinação GuGRP (abreviação do inglês *Gumble (ba-*

*(sed) generalized renewal process).* O algoritmo escolhido para a otimização dos parâmetros sendo o enxame de partículas mas utilizando também o método `findMaximum` do Wolfram Mathematica. Para testar à aderência dos dados à distribuição, o teste de Kolmogorov-Smirnov foi utilizado. Por último, algumas análises de preços de venda da manga Tommy e Palmer, em cima dos dados disponibilizados do ([HORTIFRUTI/CEPEA, 2023](#)), para possíveis situações do agricultor na escolha da melhor época para venda, considerando possíveis riscos de temperaturas elevadas que possam danificar o fruto.

O modelo GuGRP se mostrou aderente para modelar a chegada dos eventos de Ondas de Calor a um nível de significância de 0,05 do Kolmogorov-Smirnov, o  $P - value$  obteve 0,28 ou 28% de aderência aos dados transformados e adaptados para o teste. A caracterização e identificação do evento baseou-se em encontrar as temperaturas máximas acima do percentil 90 para a região de Petrolina. Também foram realizados alguns testes utilizando temperaturas máximas acima das normais climatológicas à título de conhecimento. Para o processo de tomada de decisão para o agricultor, com as estimativas das probabilidades e possíveis épocas de ocorrências do evento em mãos, é possível auxiliá-lo na tomada de decisão com base nos riscos.

## 1.1 Motivação e Justificativa

A motivação na escolha deste trabalho está baseada no entendimento da importância da agricultura para o ser humano e como estimativas de eventos raros ou temperaturas tidas como elevadas poderiam ajudar os agricultores na tomada de decisões referentes a produtividade de um fruto, evitando grandes prejuízos e perdas. Inclusive porque eventos como ondas de calor tendem a se tornar cada vez mais comuns (SOARES; SANTOS; NEVES, 2022).

Sabe-se também que o Processo de Renovação Generalizado é bastante utilizado para modelar o comportamento de falha/reparo de sistemas reparáveis e Gumbel para eventos extremos. Já os algoritmos de otimização são ótimas ferramentas matemáticas que podem procurar uma solução ótima, ou pelo menos satisfatória para um problema, sendo um fator chave em um método de máxima verossimilhança.

Alinhando portanto estes estudos e simulações de possíveis eventos como ondas de calor e temperaturas não ideais em cima de dados de temperaturas máximas de Petrolina, acredita-se que os resultados percebidos em um estudo como este podem contribuir para uma tomada de decisão mais clara ao agricultor e otimizar a produtividade do fruto.

## 1.2 Objetivos

Este estudo utiliza o Processo de renovação generalizado baseado em Gumbel, modelo GuGRP, para investigar e diagnosticar probabilidades de eventos como ondas de calor na região de Petrolina, buscando à melhoria da tomada de decisão do agricultor acerca da produtividade, cultivo e venda da manga. Para isso, o trabalho tem como objetivos gerais:

- Avaliar parâmetros de modelos preditivos de ondas de calor a partir de algoritmos de otimização utilizando a base de dados de temperatura da região de Petrolina disponibilizadas no site do Instituto Nacional de Meteorologia(INMET).

Quanto aos objetivos específicos deste estudo, busca-se:

- Formatar o banco de dados de temperaturas da região de Petrolina.
- Compreender o conceito de onda de calor e temperaturas elevadas para o fruto.
- Entender o processo de cultivo, plantio, manejo e venda da Manga.
- Fazer análises iniciais sobre os dados de temperatura.

- Aplicar Processos de renovação generalizados baseado em Gumbel para predição de eventos raros.
- Utilizar e aplicar algoritmos de otimização no método de máxima verossimilhança para estimar os parâmetros do modelo.
- Fazer simulações de Eventos.
- Estabelecer métricas de tomada de decisão para o objeto de estudo.

## 2 Trabalhos relacionados

Nesta seção são abordados alguns trabalhos relacionados a utilização da Distribuição de Gumbel em eventos raros, onde pode-se encontrar pontos relevantes relacionados ao tema deste trabalho. Foram encontrados diversos trabalhos que de alguma forma contribuíram no entendimento do estudo, entre eles estão:

No artigo ([CAMARGO et al., 1993](#)), foi estudado a probabilidade de ocorrência de temperaturas mínimas absolutas mensais e anuais no estado de São Paulo utilizando Gumbel, método de máxima verossimilhança e o teste de Kolmogorov-Smirnov onde o modelo se mostrou bom no ajuste entre as freqüências estimadas e as observadas. Neste estudo, as maiores probabilidades de ocorrência de baixas temperaturas foram observadas em junho-agosto, com valores entre 60 e 95% para as regiões serranas e sudoeste do Estado..

Já no trabalho ([ROCHA JÚNIOR et al., 2022](#)), o assunto era a análise de extremos de temperatura do Sul do Brasil, estudando a variabilidade climática nessa região, buscando tendências e utilizando dados em grade de alta resolução.

No estudo ([SANSIGOLO, 2008](#)), foi utilizado diversas distribuições teóricas de probabilidade incluindo a Normal, Gumbel, Fréchet, Weibull, LogNormal e Pearson 3, foram ajustadas aos extremos anuais de precipitação máxima diária, temperatura máxima e mínima absolutas e velocidade instantânea do vento em Piracicaba, SP (1917-2006).Foram utilizados também o MV e os testes X<sub>2</sub> e KS. A distribuição de Gumbel foi a que melhor se ajustou aos dados de temperaturas máximas absolutas.

Outro trabalho semelhante ao anterior mas com enfoque a Iguatu - CE e somente às temperaturas máximas foi ([ARAÚJO et al., 2010](#)) que também aplicou diversas distribuições de probabilidade, sendo 6 testadas. A série estudada abrangeu um período de 45 anos (1961 a 2005). A distribuição que melhor se adequou a série de dados foi a Normal e a de pior desempenho foi a distribuição Gumbel tipo I. Foram utilizados testes de aderência Qui-quadrado e Lilliefors.

Em ([ALVES et al., 2013](#)), foram analisados dados de precipitações máximas na cidade de Cuibá - MT, este trabalho utilizou a distribuição Gumbel e GEV para comparação. Além disso utilizou também Método de Máxima Verossimilhança e de Momentos para analisar qual o que se saia melhor na estimação de parâmetros e para comprovar a aderência, Anderson-Darling e Kolmogorov-Smirnov. Como resultado a distribuição GEV estimada pelo método da máxima verossimilhança foi a que melhor se adequou a série histórica presente.

No artigo ([CRISTINO; ŹEBROWSKI; WILDEMEERSCH, 2020](#)), os intervalos de tempo entre recessões econômicas foram estudados e estimados utilizando o Processo de Renovação Generalizada, Gumbel e a estimação de parâmetros utilizando o Método de Máxima Verossimilhança. O GuGRP mostrou-se bem na captura de sucessão de recessões.

Todos esses trabalhos e artigos aqui mencionados foram citados por usar distribuições de probabilidade como Gumbel ou por fazer análises de temperaturas máximas ou mínimas. Não foram encontrados muitos artigos que utilizem ao mesmo tempo Gumbel e Processo de Renovação Generalizada para dados de ondas de calor, principalmente no Brasil. Esses trabalhos relacionados somados com a fundamentação teórica do capítulo à seguir, facilitaram em entender o contexto do problema e possíveis modelos e análises que poderiam ser feitas nas estimativas de ondas de calor do presente estudo.

### 3 Fundamentação teórica

Esta seção apresenta os conhecimentos necessários para um melhor entendimento desse trabalho. A seção inicia com a apresentação das condições ideais de cultivo, manejo e colheita da manga, mercado da manga no Brasil, do entendimento do que é ondas de calor, de algoritmos de otimização e Processos de Renovação Generalizados. Também é apresentado os conceitos da distribuição de Gumbel, Método da Máxima Verossimilhança e testes de aderência. Esses conceitos serão a base para buscar estimar a possibilidade de ondas de calor durante o ano, possibilitando uma tomada de decisão do agricultor.

#### 3.1 Manga

A manga é reconhecida como a fruta fresca mais consumida em todo o mundo ([ABRAFRUTAS, 2023](#)). Ela é o fruto da mangueira (*Mangifera indica L.*, pertencente à família *Anacardiaceae*), popularmente conhecida como "pé de manga", de origem asiática, mais precisamente indiana, estima-se que chegou ao Brasil no século XVII por meio dos colonizadores portugueses. Por sua origem, o fruto se adaptou facilmente ao clima tropical do Brasil.

Segundo diversos estudos como ([OLIVEIRA et al., 2011](#)), ([GONÇALVES et al., 1998](#)), ([REIS et al., 2019](#)), a fruta traz diversos benefícios à saúde, ela possui fibras, antioxidantes, vitaminas como A e C e são fontes de minerais como cálcio, zinco, potássio, ferro e magnésio. Entre os benefícios estão:

- Imunidade, por possuir vitaminas A e C.
- Alivia a constipação por conter fibras.
- Faz bem para o coração pelos antioxidantes e fibras.
- Protege a visão, pois possui carotenoides e vitamina A.
- Retarda o envelhecimento da pele - vitaminas A e C.
- Melhora a concentração e a memória - pela composição do ácido glutâmico (glutamina).
- Aumenta a absorção do ferro - Vitamina C.

### 3.1.1 Mercado brasileiro da manga

Segundo dados da Abrafrutas ([ABRAFRUTAS, 2022](#)), a manga foi a fruta mais exportada pelo Brasil em 2021. Existindo diversas variedades do fruto, os principais tipos no Brasil são a tommy, palmer, rosa, espada e Keitt.

A manga é cultivada em todas as regiões do País e, segundo dados da Empresa Brasileira de Pesquisa Agropecuária (Embrapa), atingiu o ápice da sua expansão entre as décadas de 1980 e 1990. ([SEMIÁRIDO, 2010](#)). Hoje, é uma das queridinhas entre as frutas tropicais, uma vez que seu cultivo tem grande potencial de rentabilidade.

Em termos de exportação, a variedade de manga mais comercializada é a tommy, preferida de 85% dos consumidores europeus e a segunda é a palmer, mais procurada entre os asiáticos ([ABRAFRUTAS, 2019](#)). Graças à irrigação e a indução floral, a produção de manga no Nordeste dura o ano todo, o que fez ter vantagens à frente de outros países como a África do Sul, que só a produzia durante quatro meses ([ABRAFRUTAS, 2018](#))([CORREIO, 2018](#)).

### 3.1.2 Etapas da produção da manga

De maneira bastante resumida as principais etapas da produção de manga são as seguintes:

**I. Preparo do solo:** Quando é decidido plantar e cultivar a manga é necessário ter uma etapa de pré produção, consiste em escolher e preparar o terreno. É importante levar em consideração as propriedades físicas, químicas e biológicas e os aspectos eda-foambientais (insolação, altitude, precipitação e distribuição das chuvas). Também pode incluir o uso de adubos orgânicos e químicos, aração, nivelamento e irrigação para preparação do terreno.

No Submédio do Vale do São Francisco, em solos de textura arenosa, como os Neossolos Quartzarênicos, faz-se a limpeza da área por meio do destocamento e roçagem da vegetação, 3 a 4 meses antes do plantio, sem o uso da aração e da gradagem ([EMBRAPA SEMIÁRIDO, 2015](#)).

**II. Produção de mudas:** Pode ser feita por sementes ou pelo método da enxertia.

**III. Plantio:** As mudas são plantadas em covas com distância específica entre as plantas.

**IV. Cuidados com a planta:** Os cuidados envolvem irrigação, adubação e o controle de pragas e doenças. Para facilitar a colheita e a remoção de galhos doentes ou mortos ou até mesmo controlar o tamanho e a forma da planta, podem ser feitos algumas podas na mangueira.

**V. Florada:** É o período em que a planta produz as flores e acontece a polinização.

VI. **Frutificação:** Época em que ocorre a formação e o desenvolvimento dos frutos à partir das flores polinizadas.

VII. **Colheita:** A idade do fruto é um método bastante seguro para avaliar a maturação de mangas. Geralmente, dos 110 aos 120 dias após a floração, os frutos encontram-se em ponto de colheita ([EMBRAPA SEMIÁRIDO, 2004](#)).

Mas, o tempo ainda pode variar de acordo com a variação plantada da manga e com as condições climáticas. Os frutos devem ser cuidadosamente manipulados para evitar danos e contaminações.

VIII. **Pós-colheita:** A pós-colheita da manga envolve a seleção e classificação dos frutos; Eliminação de pedúnculo; Limpeza; Tratamento fitossanitário, controle de fungos e moscas ; Aplicação de cera; Pré-resfriamento; Embalagem; Armazenamento.

Em condições ideais, as árvores com menos de 10 anos de idade podem florescer e frutificar regularmente a cada ano. O florescimento é fortemente afetado pelo clima.

### 3.1.3 Condições ideais de temperaturas e como altas temperaturas podem afetar a manga

Segundo informações sobre a produção de manga no site da Embrapa ,[Teixeira \(2021\)](#), a faixa de temperatura considerada ideal para o cultivo da mangueira situa-se entre 24°C a 30°C.

Mas é importante lembrar que o intervalo de temperatura ideal varia conforme a etapa da produção em que se encontra o fruto. Algumas doenças por exemplo, podem aparecer na manga, como o fungo *L. theobromae*, e é atribuído às condições da planta durante a indução floral, pois altas incidências da doença são observadas em condições de estresse hídrico e nutricional ([BATISTA; BARBOSA., 2008](#)) pode causar morte de mudas quando a infecção se dá na região da enxertia.

As temperaturas muito elevadas (> 32°C) quando associadas à baixa umidade relativa e ventos intensos poderão prejudicar o florescimento e frutificação da manga ([MANCIN; MELO; SOUZA, 2003](#)), afetando negativamente o desenvolvimento dos frutos e a qualidade dos mesmos, algumas das possíveis consequências incluem:

- Redução da produção de flores: Afetar a produção de flores na mangueira, resultando em uma redução na produção de frutos.
- Queda precoce dos frutos: Afetar o desenvolvimento dos frutos, fazendo com que caiam precocemente da árvore antes de atingirem a maturidade.

- Frutos pequenos e de má qualidade: Durante a frutificação podem afetar o crescimento dos frutos, fazendo com que sejam menores e de má qualidade, com menor teor de açúcar e maior suscetibilidade a doenças e pragas.
- Retardo no amadurecimento dos frutos: Temperaturas muito altas durante a etapa de frutificação podem retardar o amadurecimento dos frutos, fazendo com que demorem mais tempo para atingir a maturidade e ficarem prontos para a colheita.

Para evitar esses problemas, é importante monitorar as condições climáticas durante a etapa de frutificação e tomar medidas para reduzir o estresse térmico na planta. Isso pode incluir a utilização de técnicas de irrigação eficientes, o uso de cobertura para reduzir a incidência direta do sol nas plantas, e a aplicação de fertilizantes e nutrientes adequados para manter a saúde da planta ([INSTITUTE, 1995](#)).

### 3.1.4 Importância da manga para o Vale do São Francisco e Petrolina

O Vale do São Francisco compreende a região drenada pelo rio São Francisco e seus afluentes localizado nos estados de Minas Gerais, Bahia, Pernambuco, Sergipe e Alagoas. Já o Submédio do Vale do São Francisco é mais delimitado na região sertaneja, no semiárido do Nordeste do Brasil e norte do estado da Bahia, compreendido portanto entre o estado de Pernambuco, onde a cidade de Petrolina faz parte e também o estado da Bahia.

Como no presente trabalho foi escolhido analisar os dados de temperatura da região de Petrolina e essa está contida no Vale, foi preferido contextualizar os resultados do Vale do São Francisco.

Figura 1 – Região do Vale do São Francisco



Fonte: Imagem retirada do site da Embrapa

No Vale do São Francisco, Santos conta que cerca de 90% das grandes empresas do setor e que cultivam manga, atuam para exportar, ficando apenas uma fatia de 10% para atender ao mercado interno ([FRUTA, 2022](#)).

O Submédio do Vale do São Francisco é o maior exportador de uvas finas de mesa e mangas do Brasil. Sendo inclusive a manga a fruta mais exportada do Brasil em 2021, com mais de 270 mil toneladas exportadas ([ABRAFRUTAS, 2022](#)). Em virtude do desenvolvimento tecnológico e da disponibilidade hídrica, a região se destaca na produção de frutas, gerando emprego e renda na parte central do Semiárido brasileiro, contribuindo, assim, para a redução da pobreza.

Estima-se que a fruticultura do Vale do São Francisco é responsável por gerar cerca de 60 mil empregos diretos e 200 mil indiretos ([ABRAFRUTAS, 2023](#)). Paulo Roberto Lima, diretor presidente da Adagro em sua fala também reforçou a grande importância do Vale na economia([ABRAFRUTAS, 2023](#)):

“A produção é exportada para mais de 30 países e gera uma receita média anual de 180 milhões de dólares, impactando positivamente no PIB agrícola de Pernambuco e do Brasil”.

Especificando mais algumas condições favoráveis que fazem a região do Submédio do Vale do São Francisco ter destaque na produção de frutas segundo a Revista A Lavoura:

Insolação anual de três mil horas – equivalentes a 300 dias de sol por ano, temperatura média de 26° C, umidade relativa de 50% e precipitação

média anual de 450 mm fazem do Vale do Submédio São Francisco uma região única para a produção de frutas. ([SOCIEDADE NACIONAL DE AGRICULTURA, 2014](#))

Além disso, devido ao desenvolvimento da agricultura irrigada e aplicação no Vale do Submédio São Francisco, foi possibilitado a produção da manga o ano todo, o que os da uma grande vantagem competitiva pois podem vender para diferentes localidades.

### 3.2 Época de floração colheita da manga - Brasil

A floração natural da mangueira na região do Vale do São Francisco ocorre com maior intensidade entre junho e agosto, mas por indução floral pode acontecer em diferentes períodos como no período de maio a agosto ([MOUCO; ALBUQUERQUE, 2002](#)). A época portanto pode variar de acordo com o plantio e a época em que o agricultor quer colher as mangas.

A época de colheita da manga pode variar se utiliza-se tecnologia de indução da floração, das condições naturais ou da variedade da manga, além do objetivo do agricultor para produção para exportação ou importação. No geral, no Nordeste a colheita ocorre de outubro a fevereiro (em condições naturais) e de agosto a outubro (com indução artificial da floração), e no Centro-Sul, de novembro a dezembro ([MANCIN; MELO; SOUZA, 2003](#)).

### 3.3 Ondas de calor

A identificação de uma onda de calor varia de uma agência meteorológica para outra, cada uma possuindo seus critérios para tal identificação.

Em geral, uma definição genérica, comum entre todas as agências, é a caracterização da onda de calor como um fenômeno meteorológico prolongado e relativamente incomum com temperaturas extremamente altas para região e persistentes durante vários dias ou até semanas.

Ondas de calor intenso podem levar a diversos problemas como secas, desequilíbrio de biomas, incêndios florestais, desidratação e mortes ([UNDRR, 2020](#)). A Europa por exemplo está sentindo cada vez mais a intensidade de ondas de calor, gerando focos de incêndio e alertas máximos. No ano de 2022 o Reino Unido excedeu 40°C pela primeira vez na história([PORTAL G1, 2022](#)). A problemática é tão grande que estima-se que a Onda de calor provocou mais de 15.000 mortes na Europa ([OMM, 2023](#)) em 2022 e 72 mil mortes em 2003 ([UNDRR, 2020](#)).

O fenômeno da onda de calor não é algo exclusivo da Europa e é sentido em todo o mundo. No Brasil não é comum ocorrer óbitos, mas ainda assim é algo que afeta o corpo

e também as plantas e os seres vivos em geral.

Constantemente o Inmet emite alertas e resumos climáticos do Brasil, em 2020 o instituto mostrou que mais de 50 cidades brasileiras registraram as maiores temperaturas do ano em setembro, muitas ultrapassando temperaturas acima de 40°C ([INMET, 2022](#))([INMET, 2022a](#)). Normalmente o Inmet também pode associar a onda de calor com umidade e temperaturas elevadas. Altas temperaturas também ocorreram nas cidades do Nordeste inclusive em algumas que fazem parte do Vale do São Francisco.

A qualidade ou previsibilidade dos produtos agrícolas são afetados diretamente pelas variações bruscas de temperaturas, na manga por exemplo, as plantas podem vegetar ao invés de florir se os eventos extremos ocorrem durante a fase de vegetação , efeito esse que foi sentido em 2020 pela safra em Monte Alto/Taquaritinga (SP) para o tipo de manga palmer ([ABRAFRUTAS, 2020](#)). Para piorar a situação, temperaturas elevadas em determinadas fases ou solos podem acarretar maiores chances de incidência de doenças nos frutos.

Voltando sobre a definição de ondas de calor, várias agências meteorológicas possuem suas fórmulas diferentesumas das outras e discordâncias quanto ao tempo de duração de temperaturas elevadas para ser identificado o evento. Pela Organização Mundial de Meteorologia (OMM)([PETERSON, 2001](#)) por exemplo, suas fórmulas se baseiam em temperaturas máximas diárias com pelo menos 5°C acima da normal climatológica da temperatura máxima da região e do mês de referência, com duração de no mínimo 5 dias. Já em ([OFFICE, 2022](#)) e ([METSUL, 2022](#)), o evento é definido por temperaturas máximas diárias, que tem duração de no mínimo 3 dias consecutivos e ultrapassam os limiares estabelecidos de acordo com as últimas normais climatológicas antes da data analisada para a região. Em ([DINIZ; GONÇALVES, 2018](#)) também são usadas a referência de duração de no mínimo 3 dias e a variável de temperatura máxima, mas usando o limiar de acima do 90º percentil da temperatura máxima diária do período climatológico de referência. Além dessas citadas, existem várias outras com duração e cálculos diferentes.

Neste trabalho será usado a técnica do percentil 90 semelhante ao trabalho de ([DINIZ; GONÇALVES, 2018](#)) e de outros que também utilizam essa técnica. Será considerado portanto uma Onda de Calor, temperaturas máximas diárias maiores que o percentil 90 para à época e região de Petrolina. A utilização dos 5 graus centígrados do trabalho ([PETERSON, 2001](#)) acima da temperatura máxima média do mês e da região não foi escolhida por não ter tantos dados que entrassem nos critérios, pelo menos para a área de Petrolina, logo não sendo suficientes para a modelagem.

### 3.4 Normais climatológicas

Para cada vez mais entender a variabilidade climática e fazer comparações no decorrer dos anos, foram definidos normais climatológicas, em resumo, dados climáticos médios dentro de um período de tempo. Sendo a atualização e acompanhamento das normais extremamente importantes para o agronegócio, risco climático de seguros agrícolas, mercado financeiro, setor de geração de energia, entre outros (INMET, 2022b).

Um guia de instruções para o estabelecimento e cálculos para normais climatológicas podem ser encontrados em (OMM, 2017). Nesse guia, é recomendado usar períodos de referência de 30 anos nas definições para as normais. E pelo site do INMET é possível encontrar os períodos de 1961 a 1990, 1981 a 2010 e 1991 a 2020 para as normais climatológicas brasileiras (INMET, 2022b).

Para temperaturas, o cálculo estabelecido nos guias (INMET, 2022b) e (OMM, 2017) são:

$$X_{ij} = \frac{\sum_k X_{ijk}}{N} \quad (3.1)$$

em que ( $X_{ijk}$ ) é o valor observado da variável ( $X$ ) (temperatura, por exemplo), no dia ( $k$ ), do mês ( $i$ ), do ano ( $j$ ), e ( $N$ ) é o número de dias no mês ( $i$ ), do ano ( $j$ ), para os quais se dispõe de observações.

A OMM recomenda que, nestes casos, se adote a regra 3:5, descartando-se os meses com ausência de dados em três ou mais dias consecutivos, ou cinco ou mais dias alternados.

Entendendo essas referências quanto à normais, é possível verificar anomalias e a variabilidade dos dados no decorrer dos anos.

### 3.5 Percentil

O percentil é uma medida estatística que divide um conjunto de dados em 100 partes iguais. Essa medida é utilizada em diferentes aplicações e para identificações de eventos como Ondas de calor. Para tal identificação, utilizando temperaturas máximas acima do 90 percentil, ou seja, acima de 90% dos dados, como no trabalho de (PERKINS; ALEXANDER, 2013) e que também será utilizado no presente estudo.

### 3.6 Separação em treinamento e teste

A separação dos dados em treinamento e teste serve para avaliar a capacidade de generalização de um modelo, pois a separação em dois subconjuntos distintos ajuda a evitar o problema de sobreajuste (*overfitting*), ou seja, quando um modelo se ajusta muito bem aos dados de entrada mas não consegue generalizar bem para novos dados.

O sobreajuste pode ser identificado quando o modelo ao ser avaliado usando o conjunto de teste, tem o seu desempenho significativamente pior do que no conjunto de treinamento, indicando a presença do sobreajuste, precisando ser reajustado.

### 3.7 Processo de renovação generalizado

Para explicar sobre o Processo de renovação generalizado é importante explicar a sua base, o processo de renovação, esse último, se baseia em modelar o tempo entre os eventos de falha ou reparo de um sistema. Ela é muito útil para otimizar a programação de manutenção preventiva ou para prever o desempenho futuro do sistema, na figura 2 tem uma representação simplista do ciclo de vida, linha do tempo de um sistema.

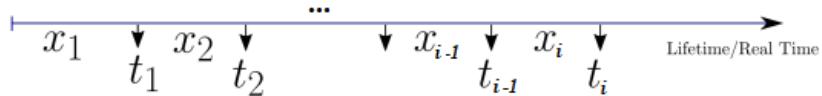


Figura 2 – Linha do tempo do PR e PRG.

Fonte: Montagem da própria autora com base no ciclo de vida do sistema em PR e PRG.

Um sistema, após um reparo, pode retornar ao funcionamento em uma das seguintes formas ([YANEZ; JOGLAR; MODARRES, 2002](#)):

- tão bom quanto novo;
- tão ruim quanto velho;
- melhor do que velho mas pior do que novo;
- melhor do que novo;
- pior do que velho.

O processo de renovação puro e simples não reflete tão bem a influência das intervenções, do tempo decorrido desde a última renovação ou o número de falhas anteriores. Baseado nisso, surgiu o Processo de Renovação Generalizada sendo uma extensão do PR mas introduzindo, para melhor interpretar as suposições de reparo, o conceito de idade virtual nos processos pontuais ([VEBER; NAGODE; FAJDIGA, 2008](#)).

A idade virtual é uma medida de tempo que indica a média ponderada da idade atual de um sistema com base no tempo que já se passou desde a última renovação ou reparo. Uma possível representação para a fórmula da idade virtual é ([YANEZ; JOGLAR; MODARRES, 2002](#)):

$$V = t + (1 - F(t))/f(t) \quad (3.2)$$

em que:

- $V$  é a idade virtual do sistema;
- $t$  é o tempo decorrido desde a última renovação ou reparo;
- $F(t)$  é a função de distribuição acumulada do tempo entre as renovações ou reparos;
- $f(t)$  é a função densidade de probabilidade do tempo entre as renovações ou reparos.

O conceito de idade virtual junto com um parâmetro de envelhecimento/rejuvenescimento  $q$ , mostra a relação e o impacto da falha observada dentro do sistema com as outras falhas e reparos antecedentes para explicar o processo presente no tempo operacional.

Os valores assumidos pelo parâmetro  $q$ , grau de renovação da idade atual do sistema, possibilitam a representação dos tipos de reparos que pode ser feito da seguinte maneira ([MOURA, 2007](#)):

- $q = 0$ : corresponde a um reparo perfeito, já que a idade virtual  $x_n$  é sempre anulada após a n-ésima ação de reparo;
- $0 < q < 1$ : corresponde a um reparo imperfeito, já que  $x_n$  é uma fração da idade real  $tn$  ;
- $q = 1$ : corresponde a um reparo mínimo, já que  $x_n$  é exatamente igual à idade real  $t_n$  .
- $q < 0$ : reparo melhor do que antes. Esse caso considera uma mudança de tecnologia do sistema, por exemplo;
- $q > 1$ : reparo pior que antes da intervenção.

O processo de renovação generalizado por ser um processo de contagem estocástica de eventos pontuais, descreve a evolução de probabilidade ao longo do tempo, aos quais entre tempos de eventos ou falhas consecutivas são variáveis aleatórias (r.v) e seguem alguma distribuição probabilística. As aplicações do PRG são assumidas quando o tempo do evento de parada é instantâneo em relação ao tempo operacional e assumem que a distribuição é renovada em intervalos regulares.

É possível avaliar o i-ésimo tempo  $t_i$  de falha através da função de distribuição acumulada condicionada na idade virtual  $x_{i-1}$ , como segue ([MOURA, 2007](#); [KIJIMA; SUMITA, 1986](#)):

$$F(t_i | x_{i-1}) = P(T \leq t_i | T > t_{i-1}) = \frac{F(t_i + x_{i-1}) - F(x_{i-1})}{1 - F(x_{i-1})} \quad (3.3)$$

## 3.8 Distribuição de probabilidade

Existem muitas distribuições de probabilidade, cada uma adequada para descrever diferentes tipos de fenômenos aleatórios.

### 3.8.1 Termos PDF, CDF e EDF

PDF, CDF e EDF são termos relacionados a probabilidade e estatística e descrevem diferentes aspectos de distribuições de probabilidade.

PDF - em inglês Probability Density Function - ou função densidade de probabilidade, é utilizada principalmente para variáveis aleatórias contínuas. Ela é uma função matemática que descreve a probabilidade de uma variável aleatória cair em um intervalo específico de valores.

CDF - em inglês Cumulative Distribution Function - ou função de distribuição cumulativa , é utilizada tanto para variáveis aleatórias contínuas quanto discretas. A CDF é uma função matemática que descreve a probabilidade acumulada de uma variável aleatória ser menor ou igual a um determinado valor.

EDF - em inglês Empirical Distribution Function - ou função de distribuição empírica, é uma função que descreve a distribuição de probabilidade empírica de um conjunto de dados. Em outras palavras, a EDF é uma função que descreve a probabilidade acumulada de um conjunto de dados. A EDF é uma estimativa não paramétrica da CDF e é frequentemente usada em análises exploratórias de dados para descrever a distribuição de probabilidade de um conjunto de dados.

Em resumo, PDF descreve a probabilidade de uma variável aleatória assumir valores em um intervalo específico, CDF descreve a probabilidade acumulada de uma variável aleatória ser menor ou igual a um determinado valor e EDF descreve a distribuição de probabilidade empírica de um conjunto de dados.

### 3.8.2 Teoria dos valores extremos

A teoria dos valores extremos também conhecida como análise de valores extremos ou análise de eventos raros é uma área da estatística que se concentra no estudo de valores extremos, ou seja, valores incomuns. Na análise gráfica é possível observar como os chamados *outliers*, que estão muito acima ou abaixo da média dos outros dados. A teoria dos valores extremos usa as distribuições de Gumbel e Fréchet para modelar o comportamento de eventos extremos e tem várias aplicações práticas em várias áreas. Para ilustrar essa teoria, considere o seguinte Teorema:

**Teorema 3.8.1.** *Seja  $(Z_n)$  uma sequência de variáveis aleatórias, v.a.'s, independentes e identicamente distribuídas com distribuição comum  $F$ . Se existe constantes normalizadoras*

$c_n > 0$ ,  $d_n \in \mathbb{R}$  e alguma distribuição não degenerada  $H$  tais que:

$$c_n^{-1}(M_n - d_n) \xrightarrow{d} H \quad (3.4)$$

então  $H$  pertence a um dos tipos de distribuições a seguir:

- Fréchet:  $\Phi_\alpha(z) = \begin{cases} 0, & z \leq 0 \\ \exp(-z^{-\alpha}), & z > 0, \alpha > 0 \end{cases}$
- Weibull:  $\Psi_\alpha(z) = \begin{cases} \exp(-(-z)^\alpha), & z \leq 0 \\ 1, & z > 0 \end{cases}$
- Gumbel:  $\Lambda(z) = \exp(-e^{-z})$ ,  $z \in \mathbb{R}$ .

$E$ ,

$$Z \sim \Phi_\alpha \Leftrightarrow \log(Z) \sim \Lambda \Leftrightarrow -Z^{-1} \sim \Psi_\alpha.$$

As funções de distribuição  $\Phi_\alpha, \Psi_\alpha, \Lambda$  como apresentadas no Teorema 3.8.1 são chamadas de Distribuição Padrão de Valores Extremos e as correspondentes v.a.'s de V.A.'s de Padrão Extremo. A prova do Teorema 3.8.1 pode ser encontrada em (EMBRECHTS; KLÜPPELBERG; MIKOSCH, 2013).

### 3.8.3 A distribuição Gumbel

De acordo com (COLES et al., 2001), a distribuição de Gumbel é uma distribuição de valores extremos e é frequentemente usada para modelar eventos extremos. A função de distribuição acumulada para o modelo Gumbel é dada por:

$$F_X(x | \alpha, \beta) = 1 - \exp \left[ -\exp \left( \frac{x - \alpha}{\beta} \right) \right]. \quad (3.5)$$

A correspondente função densidade de probabilidade é:

$$f_X(x | \alpha, \beta) = \frac{1}{\beta} \exp \left[ -\frac{x - \alpha}{\beta} - \exp \left( \frac{x - \alpha}{\beta} \right) \right] \quad (3.6)$$

O parâmetro de escala ( $\beta$ ) representa a dispersão dos dados e afeta a altura e a forma da curva de distribuição. Por exemplo, um valor maior de escala indica que os dados são mais dispersos e a curva de distribuição é mais achatada e ampla. Já um valor menor de escala indica que os dados são menos dispersos e a curva de distribuição é mais estreita e alta (GÓMEZ; BOLFARINE; GÓMEZ, 2019).

### 3.9 Processo de Renovação Generalizada baseada na distribuição Gumbel (GuGRP)

Relembrando a Seção 3.7, o *processos de renovação generalizada* (GRP - sigla em inglês) é um processo de contagem de eventos pontuais para os quais os tempos entre dois eventos consecutivos são variáveis aleatórias não negativas, seguindo algum modelo probabilístico. A distribuição do  $i$ -ésimo incremento de tempo depende somente do valor da soma parcial  $S_{i-1} = \sum_{j=1}^{i-1} X_j$  e é baseada no conceito de *idade virtual* (KIJIMA; MORIMURA; SUZUKI, 1988). Tais processos explicitam o impacto de intervenções sobre o sistema de acordo com o grau da atual “idade” do sistema (FERREIRA; FIRMINO; CRISTINO, 2015; OLIVEIRA; FIRMINO; CRISTINO, 2019).

Nesse trabalho será considerada a seguinte estratégia de modelagem. Seja  $T_{i-1}$  a idade atual de um sistema no qual ocorre o  $(i-1)$ -ésimo evento, e quando a respectiva idade efetiva é  $v_{i-1}$ , chamada *idade virtual*. Seja  $x$  ( $\geq 0$ ) o incremento de tempo até o  $i$ -ésimo evento. Então (KIJIMA; MORIMURA; SUZUKI, 1988) propõe que o sistema tem tempo acumulado até o  $i$ -ésimo evento,  $T_i$ , tem distribuição probabilística de acordo com a seguinte definição:

$$\begin{aligned} F_{T_i}(x + v_{i-1} | v_{i-1}) &= F_{T_1}(x + v_{i-1} | v_{i-1}) \\ &= \frac{F_{T_1}(x + v_{i-1}) - F_{T_1}(v_{i-1})}{1 - F_{T_1}(v_{i-1})}. \end{aligned} \quad (3.7)$$

A partir da primeira Equação (3.7), pode-se ver que o GRP é baseado na consideração eu a soma de tempos entre a primeira intervenção e a idade virtual está totalmente determinada pela família de distribuições que modelam tempos entre intervenções.

O modelo GuGRP modela o tempo de chegada da ocorrências de eventos por meio do Processo de Renovação Generalizado baseado na distribuição de Gumbel. Este modelo foi testado para predizer os eventos de recessões econômicas em (CRISTINO; ŻEBROWSKI; WILDEMEERSCH, 2020).

Nesse trabalho será utilizado esse modelo para predizer também os eventos de ondas de calor. A Figura 4, é a representação gráfica dos eventos de falha do sistema (ondas de calor) no decorrer do tempo, em que  $X$  é o tempo que o sistema está abaixo do limiar para ser considerado uma onda de calor,  $T$  é o tempo de duração da onda de calor e  $L$  é a intensidade da onda de calor (temperatura máxima alcançada durante a falha - limiar). O limiar aqui considerado serão as temperaturas máximas acima da temperatura média máxima do mês de referência das normais climatológicas, mas também poderiam ser facilmente trocadas para qualquer limiar de temperatura máxima não desejados.

Figura 3 – Representação dos eventos de interesse (ondas de calor) no modelo GuGRP.

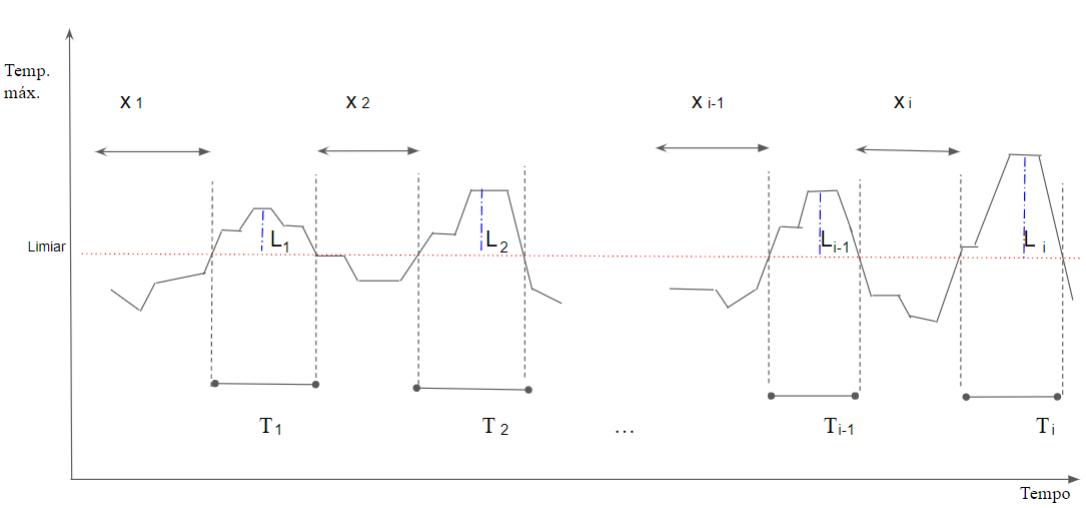


Figura 4 – Possível representação da ocorrência de ondas de calor. Fonte: a autora.

O GuGRP definido por (CRISTINO; ŹEBROWSKI; WILDEMEERSCH, 2020) considera o modelo de distribuição Gumbel como base do processo de renovação generalizada. Nesse caso, tem-se como função de distribuição:

$$\begin{aligned} F_{T_i}(x + v_{i-1} | v_{i-1}; \alpha, \beta) &= \frac{1 - \exp \left[ -\exp \left( \frac{x + v_{i-1} - \alpha}{\beta} \right) \right] - \left( 1 - \exp \left[ -\exp \left( \frac{v_{i-1} - \alpha}{\beta} \right) \right] \right)}{1 - \left( 1 - \exp \left[ -\exp \left( \frac{v_{i-1} - \alpha}{\beta} \right) \right] \right)} \\ &= 1 - \exp \left\{ \exp \left[ \frac{v_{i-1} - \alpha}{\beta} \right] - \exp \left[ \frac{x - \alpha + v_{i-1}}{\beta} \right] \right\} \end{aligned} \quad (3.8)$$

A respectiva função densidade de probabilidade é:

$$\begin{aligned} f_{T_i}(x + v_{i-1} | v_{i-1}; \alpha, \beta) &= \frac{d}{dx} F_{T_i}(x + v_{i-1} | v_{i-1}; \alpha, \beta) \\ &= \frac{1}{\beta} \exp \left\{ \frac{x - \alpha + v_{i-1}}{\beta} - \exp \left[ \frac{x - \alpha + v_{i-1}}{\beta} \right] + \exp \left[ \frac{-\alpha + v_{i-1}}{\beta} \right] \right\} \end{aligned} \quad (3.9)$$

Os mesmos autores definem as seguintes idades virtuais:

$$v'_i = v'_{i-1} + q \left( 1 - \exp \left[ -(\lambda(\tau_i, L_i)) \right] \right) x_i, \quad (3.10)$$

$$v''_i = q \left[ v''_{i-1} + \left( 1 - \exp \left[ -(\lambda(\tau_i, L_i)) \right] \right) x_i \right], \quad (3.11)$$

para as quais  $\lambda(\tau_i, L_i)$  é uma função do tempo de duração do evento de interesse,  $\tau_i$  e a intensidade do evento,  $L_i$ . Verificou-se que  $\lambda(\tau_i, L_i) = \delta\tau_i + \delta L_i$  é uma função adequada para esse trabalho, com  $\delta_1, \delta_2 \geq 0$ . Os valores de  $q$ , parâmetro de envelhecimento/rejuvenescimento, é como discutido na Seção 3.7

Como em (FERREIRA; FIRMINO; CRISTINO, 2015) e (OLIVEIRA; CRISTINO; FIRMINO, 2016) a idade virtual é uma combinação convexa das  $v'_i$  e  $v''_i$ :

$$v_i = \gamma v'_i + (1 - \gamma) v''_i, \quad (3.12)$$

com  $\gamma \in [0, 1]$ .

### 3.9.1 Teste de aderência para o GuGRP

A fim de se concluir que determinada sequência de tempos entre ocorrências de eventos de interesse são aderentes ao modelo GuGRP utiliza-se o seguinte Teorema:

**Teorema 3.9.1.** *Seja  $(W_1, W_2, \dots, W_n)$ ,  $n \geq 1$ , um vetor aleatório tal que*

$$W_i = \exp \left[ \frac{X_i + V_{i-1} - \alpha}{\beta} \right] - \exp \left[ \frac{V_{i-1} - \alpha}{\beta} \right] \quad (3.13)$$

*em que  $X_i + V_{i-1}$ ,  $i = 0, 1, \dots, n$  condicionadas à  $(i-1)$ -ésima idade virtual  $(V_i)_{i=0}^{n-1}$  segue um GuGRP com parâmetros  $\alpha$  e  $\beta$ . Então:*

1.  $W_i$  seguem uma distribuição exponencial com parâmetro 1;
2.  $W_1, \dots, W_n$  são identicamente distribuídos e mutuamente independentes.

Fórmula de transformação : A demonstração deste do Teorema 3.9.1 pode ser encontrada em (CRISTINO; ŹEBROWSKI; WILDEMEERSCH, 2020)

No Algoritmo 1 a seguir, é apresentado o modelo geral de aplicação do Teorema 3.9.1 como teste de aderência proposto. No Algoritmo,  $(x_i)$  são os tempos entre eventos (de interesse) consecutivos,  $(\tau_i)$  é a duração de cada evento  $i$  e  $(L_i)$  é a “intensidade” do evento, definida no contexto da aplicação.

## 3.10 Técnicas de estimação de parâmetros

As técnicas de estimação de parâmetros são amplamente utilizadas em diversas áreas do conhecimento, como estatística, econometria, engenharia, entre outras. Essas técnicas visam estimar os valores dos parâmetros de um modelo estatístico com base em dados observados. Existem várias técnicas de estimação de parâmetros, incluindo :

- Mínimos Quadrados Ordinários (MQO).
- Máxima Verossimilhança (MV).
- Métodos Bayesiano.
- Método dos Momentos.

---

**Algoritmo 1:** Teste de aderência para o processo de renovação generalizado baseado na distribuição Gumbel.

---

```

input      :  $(x_i, \tau_i, L_i)_{i=1..n}$ , para cada evento; o nível de significância para o teste de
              aderência para o GuGRP,  $\zeta$ .
output    : A conclusão do teste de aderência para o GuGRP.

begin
  GugrpFitTest( $(x_i, \tau_i, L_i)_{i=1..n}$ ):
  /* O melhor modelo GuGRP ajustado a  $(x_1, \dots, x_n)$  baseado no estimador
     de máxima verossimilhança é tomado */  

   $(\hat{\alpha}, \hat{\beta}, \hat{\gamma}, \hat{\delta}_1, \hat{\delta}_2, \hat{q}) \leftarrow \text{MLE}((x_i), (\tau_i), L_i)$ 
  /* O vetor proposto pela transformação do Teorema 3.9.1
      $w = (w_1, \dots, w_n)$  é tomado */  

   $v_0 = 0$ 
  for  $i \leftarrow 1$  to  $n$  do
     $v_i = \hat{\gamma} \cdot \hat{q} \sum_{r=1}^i (1 - e^{-\hat{\delta}_1 \tau_r - \hat{\delta}_2 L_r}) x_r + (1 - \hat{\gamma}) \sum_{r=1}^i \hat{q}^{i-r+1} (1 - e^{-\hat{\delta}_1 \tau_r - \hat{\delta}_2 L_r}) x_r.$ 
    // veja Equação (3.12)
     $w_i = \exp\left[\frac{x_i + v_{i-1} - \hat{\alpha}}{\hat{\beta}}\right] - \exp\left[\frac{v_{i-1} - \hat{\alpha}}{\hat{\beta}}\right]$  // Veja Teorema 3.9.1
  end
  /* O p-valor,  $p^*$ , do GugrpFitTest de  $w$  para a distribuição
     Exponencial é tomado */  

   $p^* \leftarrow \text{ExponentialPValue}(w, 1)$  // O nível de significância de  $w$  para a
     distribuição Exponencial com média 1
  if  $p^* \geq \zeta$  then
    | Não rejeite  $H_0$ : existe evidência que  $x_1, \dots, x_n$  vem de uma GuGRP com
      | parâmetros  $(\hat{\alpha}, \hat{\beta}, \hat{\gamma}, \hat{\delta}_1, \hat{\delta}_2, \hat{q})$ 
  else
    | Rejeite  $H_0$ : existe evidência que  $(x_1, \dots, x_n)$  não vem de um GuGRP ...
  end
end

```

---

### 3.10.1 Método de Máxima Verossimilhança

Máxima Verossimilhança (MV) é uma técnica de estimação de parâmetros amplamente utilizada em modelos probabilísticos. A MV estima os valores dos parâmetros do modelo que maximizam a verossimilhança ou probabilidade de observar os dados que foram observados. (CASELLA; BERGER, 2001)

Sendo a distribuição de Gumbel escolhida nesse trabalho, esse necessita que seus parâmetros sejam estimados a partir das amostras, sendo o Método de Máxima Verossimilhança bastante útil para a obtenção desses estimadores dos parâmetros da distribuição. Vários métodos podem ser utilizados para a obtenção dos estimadores dos parâmetros da distribuição de Gumbel, sendo o método dos momentos e o método da máxima verossimilhança os mais utilizados.

Embora o método dos momentos tenha sua implementação mais simples, em ge-

ral, o método da máxima verossimilhança é considerado mais eficiente do que o método dos momentos porque leva em consideração mais informações sobre os dados observados, como a variância e a covariância dos dados. A abordagem do MV baseia-se na ideia de que a distribuição que melhor explica os dados observados é aquela que tem a maior probabilidade de gerar esses dados, por sua vez a dos momentos baseia-se na ideia de que os momentos da amostra fornecem uma estimativa confiável dos momentos teóricos da distribuição.

Seja  $X = (X_1, \dots, X_n)$  um vetor aleatório e seja  $f(X | \theta)$  sua função densidade de probabilidade conjunta. Aqui  $\theta$  é um parâmetro ou um conjunto de parâmetros. No caso da estimação de  $\theta$ , a partir de uma amostra  $x = (x_1, \dots, x_n)$ , considera-se a seguinte função:

$$L(\theta | x) = f(x | \theta) \quad (3.14)$$

denominada função de verossimilhança de  $\theta$  relativa a  $X$ . Deve-se notar que se as variáveis  $X_i$ 's são mutuamente independentes, então:

$$L(\theta | x) = \prod_{i=1}^n f_{X_i}(x_i | \theta), \quad (3.15)$$

e se tais variáveis são identicamente distribuídas:

$$L(\theta | x) = \left[ f_{X_1}(x_1 | \theta) \right]^n. \quad (3.16)$$

A fórmula de máxima verossimilhança para estimar os parâmetros de um modelo pode ser escrita da seguinte forma:

$$\hat{\theta} = \operatorname{argmax} L(\theta | x), \quad (3.17)$$

em que  $L$  é a função de verossimilhança que descreve a probabilidade dos dados observados  $x$  dado um conjunto de parâmetros  $\theta$ , e  $\operatorname{argmax}$  indica o valor de  $\theta$  que maximiza a função de verossimilhança. Aqui  $\hat{\theta}$  indica que se trata de uma estimativa e não do valor verdadeiro dos parâmetros do modelo. Em outras palavras, o estimador de máxima verossimilhança considera o conjunto de parâmetros que melhor explica a amostra.

A função de verossimilhança é o produto das funções de densidade ou de probabilidade de cada observação no conjunto de dados. A forma matemática da função de densidade de probabilidade ou de massa de probabilidade depende do tipo de modelo estatístico utilizado e dos parâmetros do modelo. Como escolhido a distribuição GuGRP para o estudo, a função utilizada será (3.8).

### 3.10.2 Log-Verossimilhança

O logaritmo da função de verossimilhança simplesmente é a transformação logarítmica da função de verossimilhança.

A utilização do logaritmo na verossimilhança do modelo é bastante utilizada para facilitar os cálculos matemáticos de números grandes da fórmula original. A transformação logarítmica transforma o produto em uma soma, em consequência pode ajudar a evitar problemas de *underflow* ou *overflow* numérico.

Em termos de equivalência, maximizar a verossimilhança é o mesmo que maximizar o logaritmo da verossimilhança, logo produzem em teoria as mesmas estimativas.

## 3.11 Algoritmos de otimização

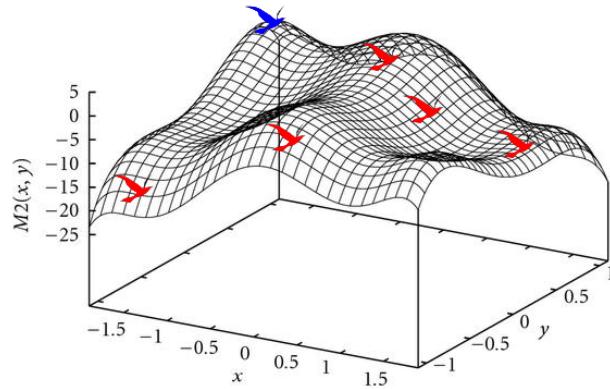
A área de Otimização trata do estudo para buscar a solução ideal para um determinado problema, encontrando a melhor combinação de variáveis que minimize ou maximize a função objetivo, respeitando as restrições impostas (PEDREGAL, 2004). Diversos algoritmos tratam dessa área cada um contendo seus prós e contras, cabendo ao pesquisador identificar qual o ideal para o problema proposto.

### 3.11.1 Algoritmo de enxame de partículas

O algoritmo de enxame de partículas é baseado em comportamentos sociais de enxames de animais e funcionam criando uma população de partículas que se movem pelo espaço de busca para encontrar a solução ótima. Eles podem ser eficazes para lidar com eventos extremos, pois são capazes de explorar uma ampla gama de soluções e encontrar soluções incomuns (CLERC, 2010).

O PSO começa com um conjunto de partículas aleatórias em um espaço de busca multidimensional. Cada partícula representa uma solução candidata para o problema de otimização. As partículas se movem pelo espaço de busca em busca da melhor solução possível. A posição de cada partícula é atualizada em cada iteração com base na sua posição atual, na melhor posição já encontrada pela partícula e na melhor posição já encontrada por todo o enxame (KENNEDY; EBERHART, 1995).

Figura 5 – Analogia de um bando de pássaros em um espaço de busca de 3D. Considerando um problema de maximização, o pássaro azul é a partícula de solução ótima do problema.



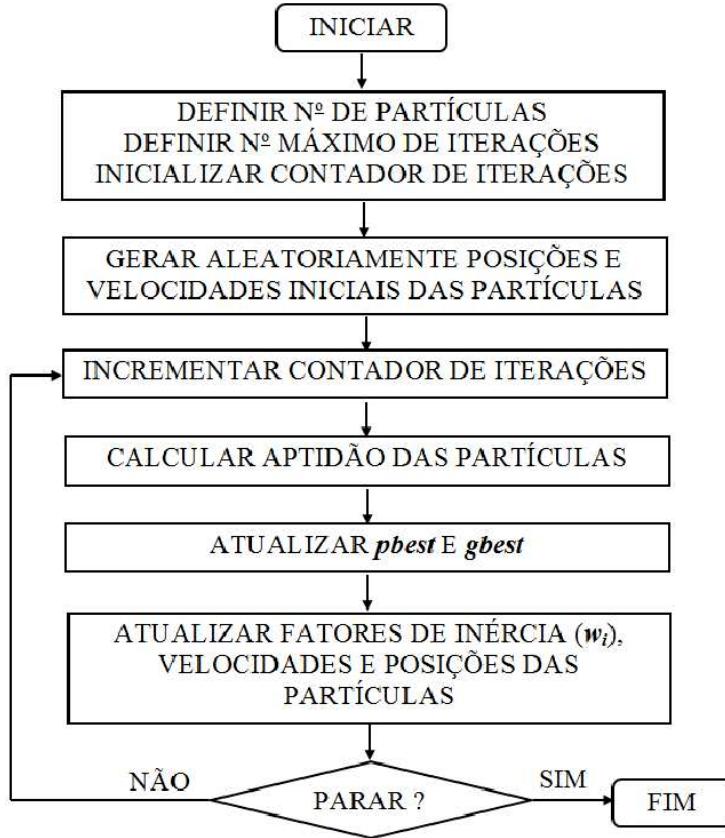
Fonte: Imagem retirada do site - Computação Inteligente

O enxame é amplamente utilizado em problemas de otimização não-lineares e não-convexos, em problemas de treinamento de redes neurais artificiais e em outras aplicações de aprendizado de máquina. É uma técnica de otimização global, ou seja, pode encontrar a solução ótima global em vez de ficar preso em um mínimo local ([CLERC, 2010](#)).

Entre os parâmetros do algoritmo de Enxame de Partículas são o número de partículas da população, número máximo de iterações e dimensão da função teste ([SILVA, 2017](#)).

Na figura 6 está um fluxograma do algoritmo apresentado no trabalho ([NETO et al., 2012](#)) explicando como funciona.

Figura 6 – Fluxograma do algoritmo de Enxame de partículas.



Fonte: Imagem retirada do trabalho - Melhoria de segurança dinâmica baseada em análise estocástica e metaheurística

### 3.12 Testes de aderência

Os testes de aderência são utilizados para verificar se uma amostra de dados segue uma determinada distribuição teórica. Entre os testes de aderência é possível citar:

- Teste de Kolmogorov-Smirnov.
- Teste de Lilliefors.
- Teste de Shapiro-Wilk.
- Teste de Anderson-Darling.

Para testar a aderência de uma amostra de dados ao Processo de Renovação Generalizada baseada na distribuição Gumbel, o teste escolhido nesse estudo foi o teste de Kolmogorov-Smirnov por ser mais simples e mais fácil de interpretar do que os outros testes de aderência e ao mesmo tempo poderoso o suficiente para detectar desvios da distribuição teórica.

### 3.12.1 Teste de Kolmogorov-Smirnov

O Teste não paramétrico de kolmogorov-Smirnov é utilizado para comparar o quão próximo os dados amostrais estão de uma distribuição de probabilidade de referência ([FRANCISCO; SIMÕES, 2021](#)). Ele pode ser aplicado até mesmo para pequenas amostras.

A estatística de teste Kolmogorov-Smirnov (KS):

$$D_n = \sup_x [|F(x) - S(x)|] \quad (3.18)$$

em que  $F(x)$  é a distribuição acumulada teórica e  $S(x)$  é a distribuição acumulada empírica. O teste KS utiliza o teste de hipótese intrinsecamente, em que:

$H_0 : F(x) = S(x)$  - A amostra provém de uma distribuição teórica específica.

$H_1 : F(x) \neq S(x)$  - A amostra não provém de uma distribuição teórica específica.

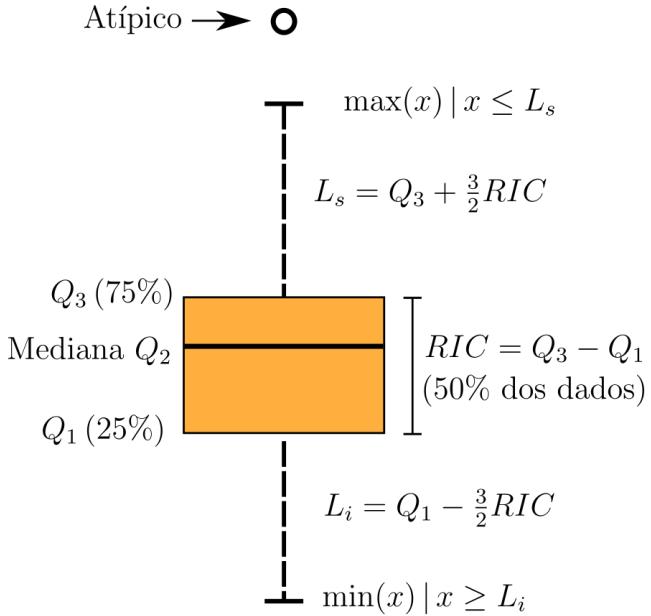
Neste trabalho a distribuição teórica específica é Gumbel.

A estatística do teste espera que quando  $H_0$  é verdadeira, as diferenças entre  $S_N(X_i)$  e  $F_0(X_i)$  sejam pequenas e estejam dentro do limite dos erros aleatórios. ([SCUDINO, 2008](#)) O teste focaliza o maior dos desvios chamado de desvio máximo.

## 3.13 Box plot

Um gráfico box plot possibilita o entendimento da média dos dados, mediana e dos valores discrepantes dentro de uma distribuição. Na [7](#) é possível ver como se deve fazer uma leitura desse gráfico. Onde os dados são separados em torno da média e mediana, 25% até a mediana, sendo o primeiro quartil, da mediana 50% até 75% sendo o terceiro quartil. Os valores maiores do terceiro quartil sendo os limites superiores, atípicos, assim como os valores abaixo do primeiro quartil.

Figura 7 – Representação de um gráfico *box plot*,  $Q_1$  é o primeiro quartil, a mediana é o segundo quartil,  $Q_2$ , e  $Q_3$  é o terceiro quartil.  $L_i$  e  $L_s$  limites inferior e superior, respectivamente.  $RIC$  é a região do intervalo de confiança.



Fonte: a autora.

### 3.14 Definição de Risco

Define-se *risco* de um evento como ([HENLEY; KUMAMOTO, 1996](#)):

$$R(E) = P(E) \times C(E), \quad (3.19)$$

em que  $R$  é o valor da possível perda atrelado ao item (nesse caso, a manga),  $P$  é a probabilidade do evento ocorrer um evento (indesejado)  $E$  e  $C$  o custo, consequência ou o valor investido desse evento.

Esta definição é utilizada como fator de decisão em muitos casos. Neste trabalho, o risco de um exemplo pode ser usado na tomada de decisão de agricultores em fazer ou não o manejo de seus cultivos em determinada época do ano, levando-se em consideração os preços de venda na época da colheita, o custo de uma queda de produção devido a eventos condicionantes, tais como ondas de calor, na época da floração das plantas.

# 4 Metódos e ferramentas

O método empregado para a execução deste estudo é constituído pela utilização de estatística, do Processo de Renovação Generalizado baseado em Gumbel, do Método de Máxima Verossimilhança e Log-Verossimilhança utilizando o algoritmo de otimização de Enxame de Partículas e o `findMaximum` do Wolfram e o teste de Kolmogorov-Smirnov. A linguagem de programação é Python, versão 3.8.3, além das bibliotecas de Python Numpy, Pandas, Pandas Profiling, Feature Engine, Scikit-Learn, Imbalanced-Learn, Yellowbrick, Seaborn e Matplotlib, nas versões mais recentes e também com a utilização do programa Wolfram Mathematica na versão Notebook Edition.

## 4.1 Seleção da base de dados

A base de dados escolhida é das temperaturas da região de Petrolina disponíveis no site do instituto Nacional de Meteorologia - INMET. Essa base de dados no site, abrange os anos de abril de 2003 até os dias atuais, a medição é realizada diariamente todos os dias do ano pelo laboratório ([INMET, 2023](#)). Já os dados referentes a preço de Venda da Manga no mercado interno foram retirados da fonte Hortifruti/Cepea.

## 4.2 Pré-processamento de Dados

Conforme foram sendo analisadas as bases de dados disponíveis no site do INMET foi necessário fazer a limpeza e padronização dos dados, retirando cabeçalhos desnecessários, renomeando colunas, contando valores vazios da tabela de meses que não foram preenchidos na base como por exemplo o ano de 2003 que não contém os meses de janeiro e fevereiro até o dia 21 de fevereiro. Assim como 2003, o ano de 2023 não contém o preenchimento dos dados de temperatura de todo o ano de 2023 já que o ano ainda não terminou no momento em que este trabalho estava sendo feito.

Além disso, a base de dados de temperaturas continha os dados de temperaturas máximas medidas de hora em hora, para não atrapalhar no momento da leitura diária das temperaturas foi realizada a filtragem da maior temperatura máxima do dia e colocada em outro dataframe para continuar as análises.

Nos dados dos valores de venda da Manga foram separadas os dados pelas variações Palmer e Tommy para uma melhor análise de preços.

Diversas verificações de inconsistências foram executadas sobre os dados observados por meio da linguagem de programação Python e das bibliotecas de Python Pandas

e Pandas Profiling durante a etapa de pré processamento de dados.

## 4.3 Processamento dos dados

### 4.3.1 Análises iniciais de temperaturas

Foram realizados análises exploratórias dos dados com objetivo de entender um pouco melhor as características dos dados, insights e possíveis problemas que possam afetar a validade dos resultados.

### 4.3.2 Análises iniciais do Preço de Venda da Manga - Brasil

A base de dados dos preços de venda da manga do Vale do São Francisco foram plotados em gráficos de linha, coluna e box plot para leitura visual rápida de quais meses tiveram valores com maiores médias a título de comparação também.

### 4.3.3 Cálculo para as probabilidades de ocorrência de Ondas de Calor

#### 4.3.3.1 Contagem de Ondas de Calor que ocorreram em Petrolina

Para começar a contagem dos eventos extremos, mais precisamente a Onda de Calor, foi necessário definir qual seria o conceito para aplicar na metodologia:

O conceito escolhido foi a definição de Onda de Calor como um Evento em que as temperaturas máximas diárias estariam acima do 90 percentil do mês de referência para a região de Petrolina. Além disso, sendo necessário que as temperaturas extremamente elevadas ocorressem por pelo menos três dias consecutivos. Também à título de informação, alguns gráficos foram feitos mostrando os dados que ficaram acima das normais climatológicas de referência.

#### 4.3.3.2 Separação de dados em treinamento e teste

Será feita a separação dos dados na proporção 80% treinamento e 20% teste. Isso ajuda a verificar se o modelo é capaz de generalizar para novos dados. A separação será feita pela cronologia dos dados.

#### 4.3.3.3 Utilização - Gumbel, PRG, MV, Teste KS

- Foi realizado a contagem de Ondas de Calor ocorridas, guardando 2 informações de cada evento para serem passadas para a função de máxima verossimilhança: o tempo em anos entre o estado de falha do dado atual e anterior(Onda de Calor); a duração do evento em dias. Também foi guardadas outros dados como a intensidade, essa última definida como a diferença de temperatura máxima do evento

e da temperatura máxima média para à época de referência, ano do evento e mês mas não foram passadas para a função objetiva. Os dados depois da contagem dos eventos foram separados em treinamento e teste.

- Posteriormente, com a função de Máxima Verossimilhança do Processo de Renovação Generalizada baseada em Gumbel, utiliza-se o algoritmo Enxame de partículas ou `findMaximum` do Wolfram para estimar os parâmetros da distribuição pelos dados de treinamento.
- Para se calcular a idade virtual de cada evento, é feito uma transformação dos dados utilizando a função [3.13](#).
- Para ser testada a aderência da distribuição, o teste de Kolmogorov Smirnov é utilizado .
- É feito a extrapolação das previsões para testar as previsões para os dados de teste e cálculo das probabilidades do evento ocorrer em determinado tempo.

Para a estimação de parâmetros com o Método de Máxima Verossimilhança e Log-Verossimilhança a função foi escrita e definida baseada na função de densidade da probabilidade do modelo GuGRP em perspectiva de estimação de parâmetros pelo Log-Verossimilhança, sendo passada como função objetivo para o algoritmo de otimização escolhido, o Enxame de Partículas, para isso utilizando a biblioteca PySwarms e a utilização do método `findMaximum` do Wolfram Mathematica. Para a validação do modelo foi realizado o teste de Kolmogorov-Smirnov.

#### 4.3.4 Cruzamento das estimações do Evento com os investimentos e preços da Manga

Com o modelo de estimações do Evento em mãos é possível calcular o risco e deixar com que o agricultor decida se é muito alto ou não para se enfrentar. Caso não queira se arriscar em uma "grande" perda, ele poderá escolher outra época ou mês em que o risco seja menor.

# 5 Resultados e discussões

## 5.1 Análises iniciais

### 5.1.1 Observações importantes

É importante destacar que diversos dias estão mal formatados ou com o não preenchimento da coluna de temperaturas máximas. Na contagem da série histórica de 2004 a 2023 foram encontrados 390 dias distribuídos em vários anos, nessa condição.

### 5.1.2 Gráfico de colunas de temperaturas médias máximas

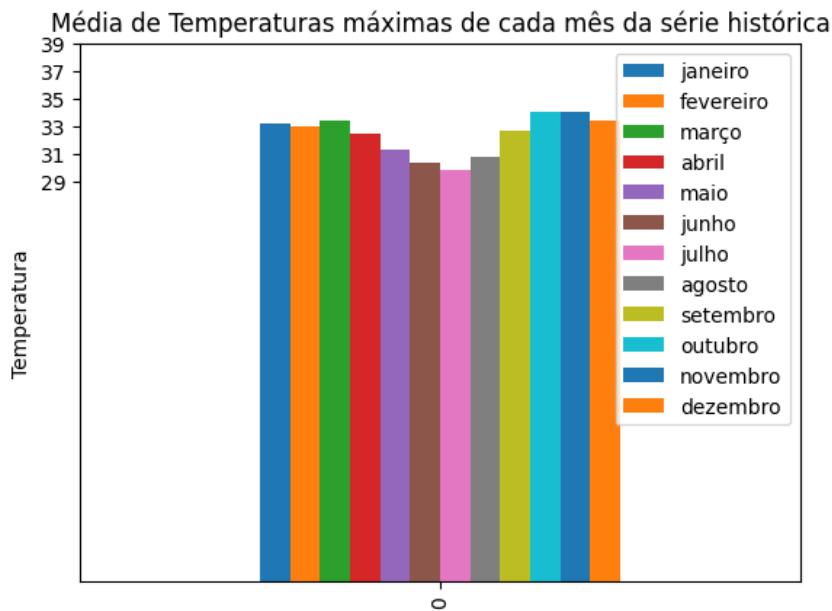


Figura 8 – Temperaturas médias máximas de cada mês.

Fonte: Análise da própria autora.

Gráfico de barras plotados [8](#) em cima da análise das temperaturas máximas médias de cada mês de toda a amostra coletada, é possível notar que maiores médias de temperaturas máximas foram nos meses de outubro e novembro e a menor em julho. Do mês de abril a julho com tendência a decrescer as temperaturas e de agosto a outubro/novembro crescer.

### 5.1.3 Desvio padrão, média e mediana dos dados

Na tabela [1](#) tem um breve resumo dos dados já filtrados, ou seja, dados que já passaram pelo tratamento de linhas vazias, entre outros.

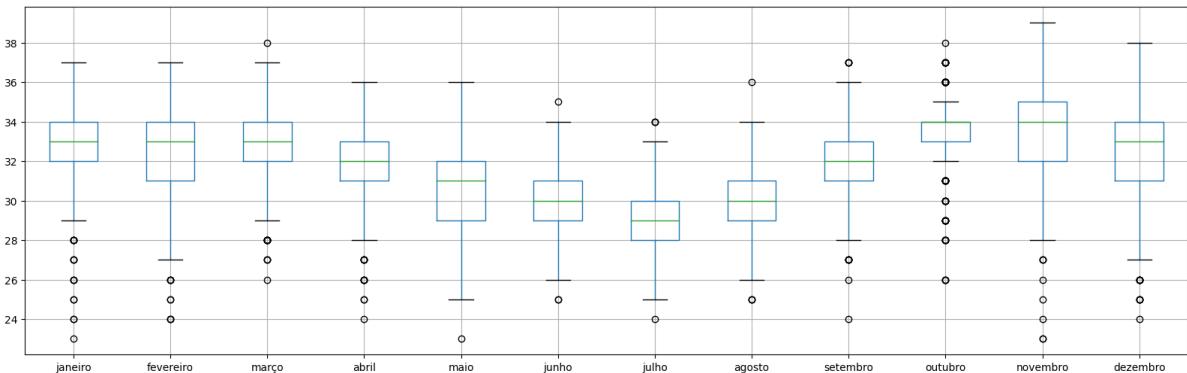
Tabela 1 – Descrição básica dos dados de temperaturas máximas da série histórica analisada.

	Temperatura máxima
Quantidade	6379
Mediana	31.512102
Desvio padrão	2.493103
Min	22.30
25%	29.70
50%	31.70
75%	33.50
Máximo	38.00

Fonte: Elaborada pela autora.

#### 5.1.4 Box plot das temperaturas

Figura 9 – Box plot de temperaturas máximas de todos os meses e da série histórica entre 2008-2022.

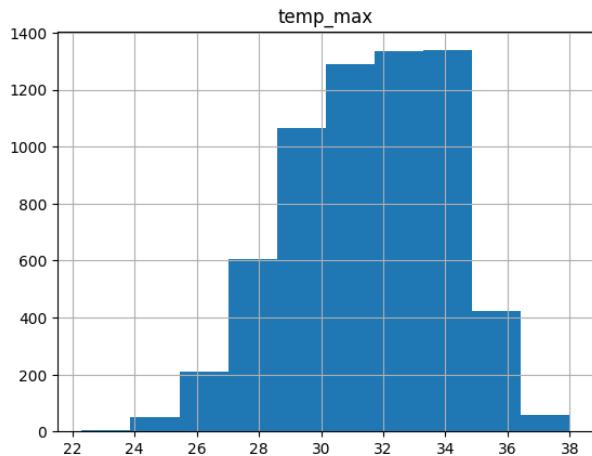


Fonte: Análise da própria autora.

Na plotagem do box plot de temperaturas máximas de todos os meses entre os anos de 2008 e 2022 na figura 9 é possível notar a presença de alguns outliers, ou seja, alguns valores atípicos de temperaturas máximas tanto "bem" abaixo quanto alguns muito "acima". Lembrando que os outliers plotados no box plot da 9 representam as temperaturas máximas individuais que foram atípicas e não exatamente indicam a presença explícita do Evento de Onda de Calor, já que para ser caracterizado assim precisaria obedecer os critérios determinados nesse trabalho.

### 5.1.5 Histogramas das temperaturas

Figura 10 – Histograma das temperaturas máximas médias.



Fonte: Análise da própria autora.

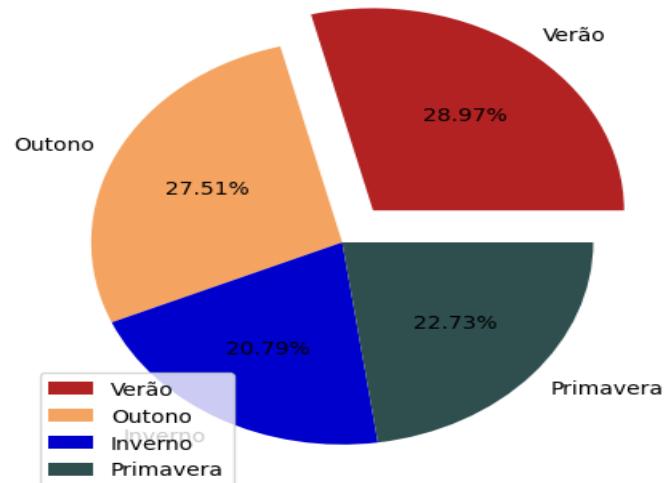
Pelo histograma 10 de temperaturas máximas médias de todos os meses da série histórica coletada é possível observar a distribuição dos dados. A maior frequência de temperaturas médias máximas está centrada no centro do gráfico, ou seja perto da média (medida central), o que caracteriza o histograma como simétrico. Esse resultado pode dar indício que a distribuição dos dados é normal, mas para confirmar se segue a normal ou outra é necessário a validação no teste de Kolmogorov-Smirnov.

Em resumo, os dados então principalmente entre as temperaturas médias máximas de 30°C e 36°C, contendo poucas frequências com valores menores que 28°C ou maiores que 36°C.

### 5.1.6 Dias com temperatura máxima acima da normal climatológica

No gráfico de pizza 11, do total dos 2323 dados acima, 28,97% das anomalias dos dados de temperaturas máximas acima da normal climatológica de referência ficaram com a estação do verão, em seguida o outono com 27,51%, primavera com 22,73% e o inverno com 20,79%.

Figura 11 – Gráfico de pizza de temperaturas máximas que ultrapassaram as normais climatológicas separadas por estação



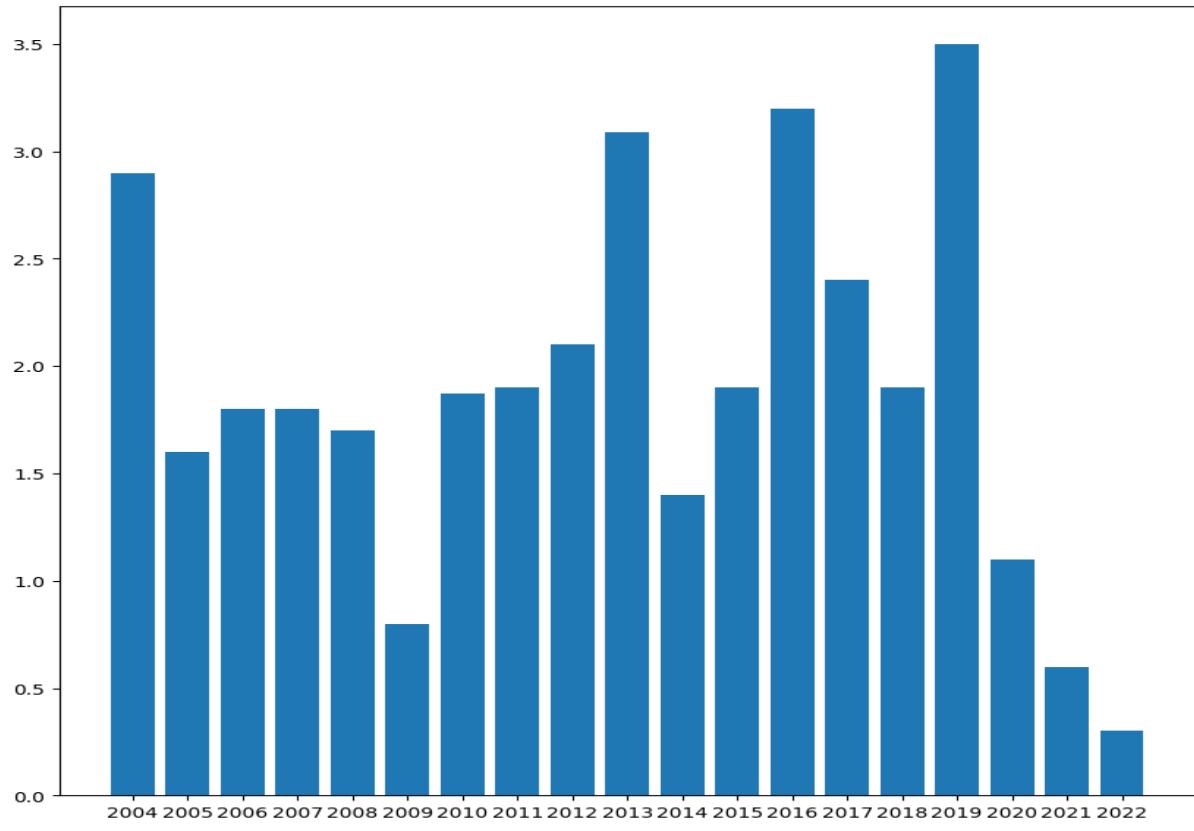
Fonte: Análise da própria autora.

### 5.1.7 Gráficos de intensidade e duração de uma Onda de Calor

Utilizando da técnica de percentil 90, as temperaturas máximas acima e consecutivas podem ser observadas do ponto de vista de intensidade e duração segundo os gráficos abaixo.

### 5.1.7.1 Intensidade

Figura 12 – Gráfico de barras da intensidade em graus da Onda de calor no decorrer dos anos

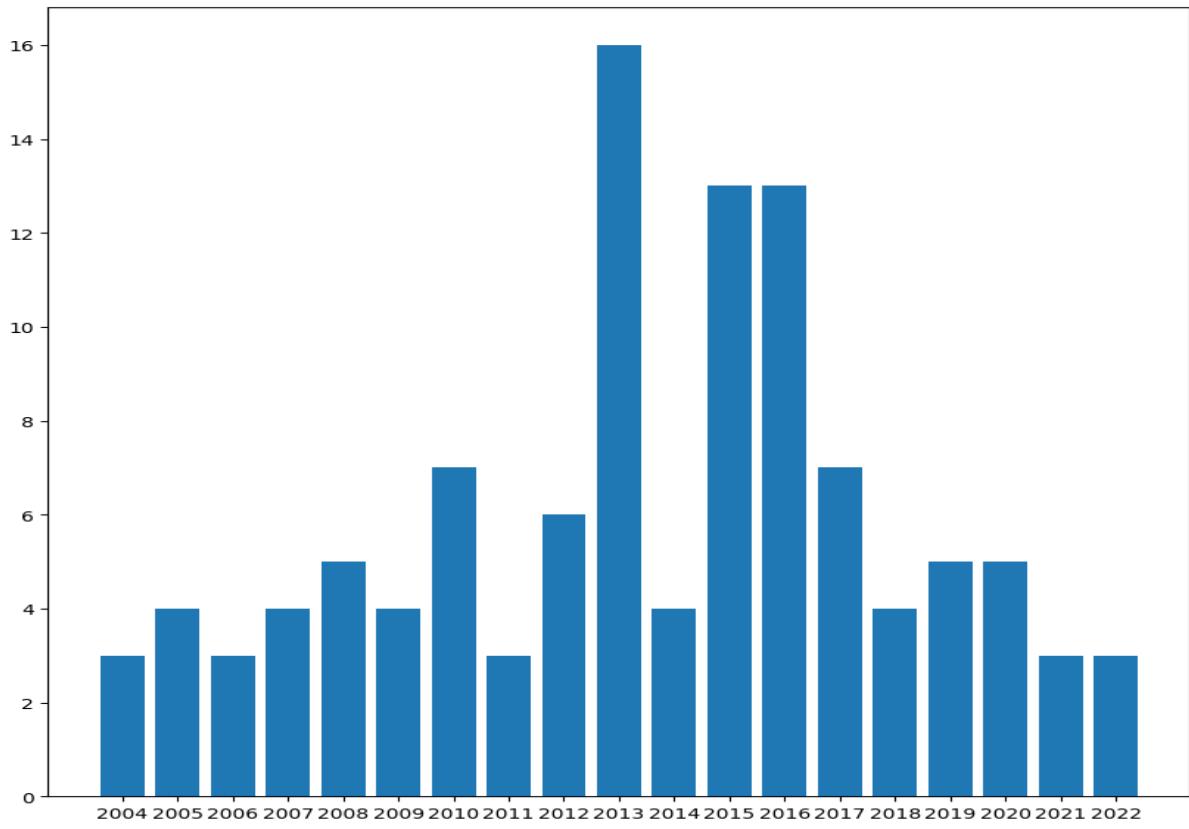


Fonte: Análise da própria autora.

Na figura 13, em termos de intensidade, ou seja, a variação em graus das temperaturas máximas acima do percentil 90, os maiores valores de variação foram os anos 2019, 2016, 2013 e 2004 em Petrolina.

### 5.1.7.2 Duração

Figura 13 – Gráfico de barras da duração da Onda de calor no decorrer dos anos



Fonte: Análise da própria autora.

Pela técnica do percentil, utilizando os dados do Inmet, é possível observar uma alta duração do evento de Onda de Calor no ano de 2013, seguido de 2015 e 2016 em detrimento dos outros anos.

## 5.2 Estimativas de temperaturas

### 5.2.1 Aplicação do Método de Máxima Verossimilhança e Log-Verossimilhança

Com a função de log-verossimilhança estabelecida, os dados necessários como tempo de uma onda de calor para outra e a duração do evento, foram aplicados e lidos na função, formando uma função objetivo para algoritmos de otimização estimar os parâmetros.

### 5.2.2 Utilização do algoritmo de otimização enxame de partículas

Na utilização do algoritmo enxame de partículas utilizando a biblioteca PySwarms, utilizou-se a função de log-verossimilhança como função objetiva mas não obteve as melhores estimativas dos quatro parâmetros, que tornassem as distribuições transformadas

mais aderentes a uma distribuição exponencial de parâmetro 1. Entretanto, a função do `findMaximum` do Wolfram Mathematica estimou os parâmetros e obteve boa aderência para o modelo e a distribuição esperada. A função `findMaximum` do Wolfram utiliza diversos algoritmos de otimização para estimar os parâmetros da função do método de máxima verossimilhança ou log-verossimilhança.

### 5.2.3 Parâmetros estimados

Os valores estimados para cada um dos 4 parâmetros utilizados na função de densidade de probabilidade do modelo GuGRP na perspectiva de log de máxima verossimilhança estão apresentados em valores aproximados na tabela 2.

Tabela 2 – Parâmetros estimados

$\rho$	$\alpha$	$\beta$	$\delta$
$\approx 1.53483$	$\approx -345.71036$	$\approx 64.84899$	$\approx 17.29964$

Fonte: Dado obtido no Wolfram Notebook.

### 5.2.4 Teste de aderência - Teste de Kolmogorov-Smirnov

Ao nível de significância de 0.05, o teste de Kolmogorov-Smirnov representado na tabela 3 apontou que não tinha evidências para rejeitar a hipótese nula que os dados transformados não seguiam uma distribuição exponencial com parâmetro 1, em outras palavras, significando que os dados se mostraram aderentes ao modelo GuGRP com a medida de  $P-value$  de 0.280101.

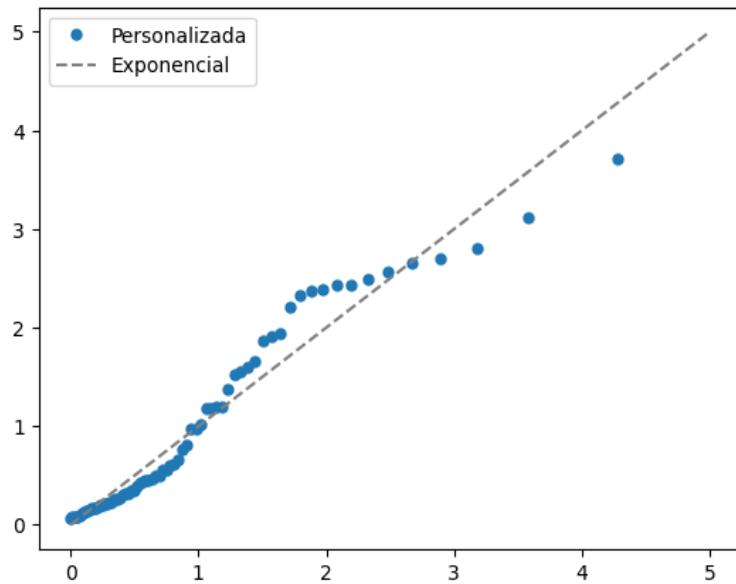
Tabela 3 – Tabela do teste de aderência dos dados transformados na distribuição exponencial com parâmetro 1.

Teste	Statistic	P-value
Anderson Darling	1.38515	0.20604
Cramér-von Mises	0.236713	0.206334
Kolmogorov-Smirnov	0.113701	0.280101
Kuiper	0.178486	0.0717222
Pearson $x^2$	12.3151	0.340433
Watson $U^2$	0.217264	0.0268756

Fonte: Dado obtido no Wolfram Notebook.

#### 5.2.4.1 Gráfico da aderência

Figura 14 – Gráfico comparando os dados transformados com uma distribuição exponencial com parâmetro 1



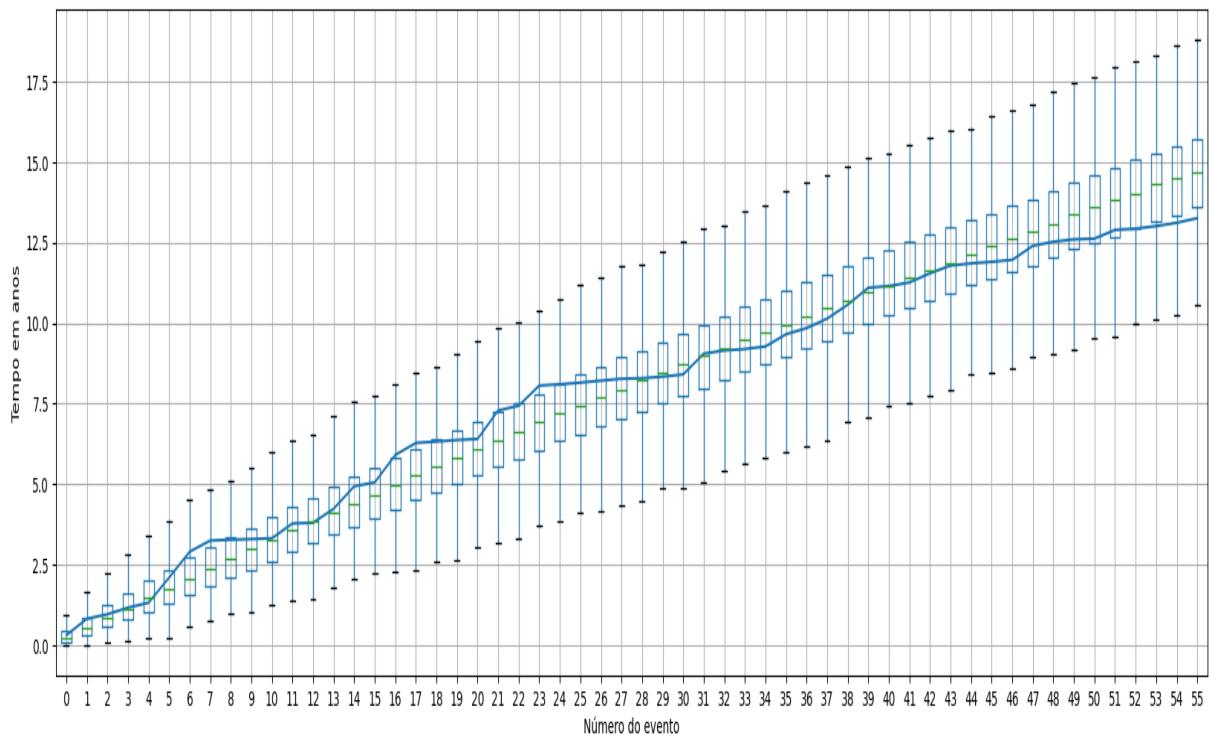
Fonte: Análise da própria autora em código python.

#### 5.2.5 Box plot estimativas

##### 5.2.5.1 Box plot real x treinamento

No box plot 15 do treinamento, O eixo  $x$  representa o número dos eventos e o  $y$  os anos. O box representa a distribuição das estimativas para o evento utilizando a distribuição acumulada do modelo, enquanto a linha contínua, a distribuição com o que de fato ocorreu de onda de calor. Para fazer os limites do box plot se assemelhando a figura, foram realizados 999 interações em cada número de evento de onda de calor. É possível observar com isso, que as estimativas se adequaram bem ao que de fato ocorreu.

Figura 15 – Box plot das estimativas da distribuição x o que de fato ocorreu



Fonte: Análise da própria autora em código python.

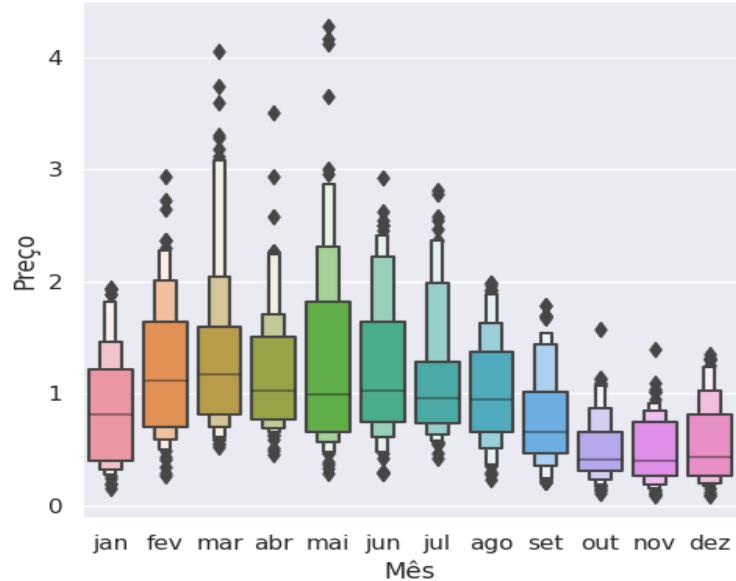
## 5.3 Análises básicas do mercado da Manga

### 5.3.1 Análise de preço - Manga Tommy

A partir dos dados de preço de venda da manga da região do Vale São Francisco, sendo esses dados disponibilizados pelo Hortifrutti Cepea em seu site, na Figura 16 são apresentados os *Box plot* do preço dessa espécie de manga ao longo do ano.

Pelo gráfico nota-se que a mediana de preço de venda da manga Tommy se encontra nos meses de fevereiro e março e as menores medianas entre os meses de outubro, novembro e dezembro. Nota-se também a presença de vários outliers, que normalmente são causados por fatores externos como eventos extremos sejam eles encheres, ondas de calor ou outros eventos.

Figura 16 – Box plot do preço de Venda - Manga Tommy.



Fonte: Análise da própria autora.

### 5.3.2 Análise de preço - Manga Palmer

No boxplot 17 do preço de venda da manga Palmer da região do Vale São Francisco também disponibilizados pelo Hortifrutti Cepea, a maior mediana encontra-se no mês de abril e os menores assim como a Tommy, nos meses de novembro e dezembro. A maior discrepância dos outliers no mês de maio.

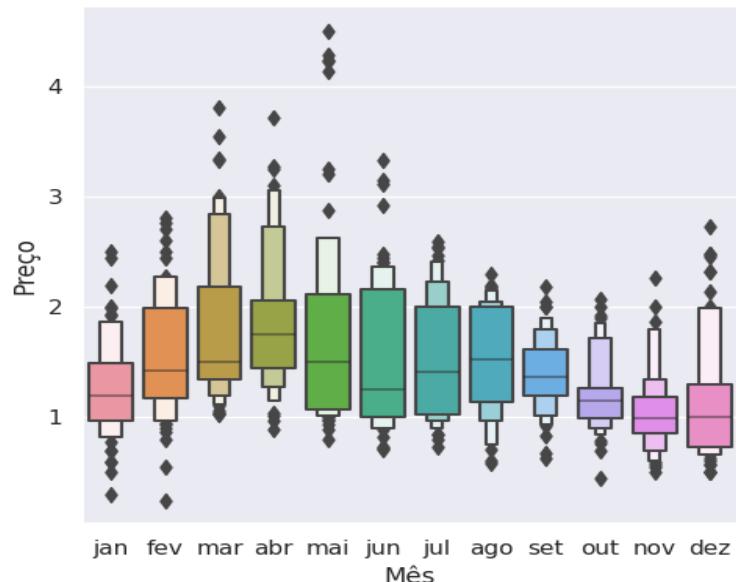


Figura 17 – Box plot do preço de Venda - Manga Palmer.

Fonte: Análise da própria autora.

## 6 Possíveis situações - Temperaturas elevadas x Preços de Mercado por mês

Utilizando da distribuição do modelo pode-se realizar estimativas de quando um evento de onda de calor possivelmente pode ocorrer, pela extração dos dados.

Um exemplo prático que poderia utilizar é: Se pela extração conflitar com a etapa de polinização e frutificação da manga à uma probabilidade de 28% pode-se calcular o risco pela fórmula já mencionada [3.19](#).

Onde, em uma situação hipotética, o custo, investimento realizado na produção da manga for cem mil reais, substituindo os valores com a probabilidade do evento que possa prejudicar a manga (uma onda de calor) de 28%, obtém-se uma possível perda de 28 mil reais como representado na equação [6.1](#) e [6.2](#). Ao mesmo tempo, se o evento não ocorre-se, podendo estimar o lucro a se obter pela média de valores de venda do produto menos o custo

$$R(E) = 0,28 * 100000 \quad (6.1)$$

$$R(E) = 28000 \quad (6.2)$$

$$Lu = (qtd * Vunit) - (Cunit * qtd) \quad (6.3)$$

Onde em [6.3](#),  $Lu$  representa o lucro,  $qtd$  a quantidade de mangas,  $Vunit$  o valor unitário para venda e o  $Cunit$  o custo unitário para produzir.

Sabendo portanto dessa probabilidade e da possível perda ou lucro, cabe ao agricultor escolher se cabe o risco de perder ou adiantar em alguns meses ou até atrasar a produção, para não sofrer possível prejuízo ou também ter um cuidado mais especial nesse possível evento.

É importante destacar que essas probabilidades como o próprio nome já diz, são estimativas do evento, não significando que de fato irá acontecer com 100% de certeza, por isso, utiliza-se os níveis de significância e probabilidade.

## 7 Conclusão

A presente dissertação teve o objetivo de encontrar tendências ou estimativas de temperaturas extremamente elevadas, Ondas de Calor, no plantio, manejo ou cultivo da Manga na região do Vale do São Francisco e portanto tornando um processo de apoio à decisão na produção da manga. Para tal, utilizando análises básicas e o modelo GuGRP, Processo de Renovação Generalizado baseado na distribuição de Gumbel, tendo seus parâmetros estimados pelo método de máxima verossimilhança por algoritmos de otimização e testes de aderências para garantir uma melhor estimativa desses eventos e apoio à decisão ao agricultor.

Com relação às dificuldades para o desenvolvimento do trabalho, os dados coletados pelo INMET não possui tantos anos de coleta das temperaturas disponíveis em seu site, o que pode ocasionar alguns erros de precisão nas estimativas. Em termos de processamento, ocorreram alguns overflows e limites de processamento para calcular os dados no Google Collab, tendo uma melhor performance no Wolfram Mathematica. Além disso, ocorreu uma dificuldade em validar os Eventos de Ondas de Calor encontrados nos dados com os noticiários da região. A qualidade dos dados de temperatura é de extrema importância para garantir uma análise mais realista da estimação.

Os dados de temperatura máxima que se mostraram mais aderentes a distribuição GuGRP, com os seus parâmetros sendo maximizados e estimados pelo método de Máxima Verossimilhança, foi utilizando a função `findMaximum` do Wolfram Mathematica, que utiliza diversos algoritmos de otimização para encontrar as estimativas para os parâmetros em detrimento do método de otimização enxame de partículas utilizando a biblioteca PySwarms. A aderência dos dados foram comprovadas pelo teste de Kolmogorov-Smirnov.

Pela extração dos dados da distribuição do modelo utilizando temperaturas máximas, é possível estimar possíveis eventos de Ondas de Calor. Além disso, à título de apoio a decisão para o agricultor, os meses de março, abril e maio para a Manga Palmer e os meses fevereiro, março e maio para a Manga Tommy são os que possuem maiores médias de preço na venda da Manga ou discrepâncias de valores para o mercado interno(brasileiro). Sabendo da probabilidade desses eventos é possível ter uma maior atenção no manejo e na produção da Manga ou até mesmo adiantando o plantio e consequentemente venda do produto para evitar possibilidades de maiores prejuízos.

Como possíveis trabalhos futuros, com o objetivo de aprofundar e facilitar ainda mais as tomadas de decisões do agricultor da manga, poderiam ser realizados os seguintes trabalhos:

- Comparativos de distribuições de probabilidade na estimação de Ondas de Calor.
- Previsão e estimativa de temperaturas mínimas.
- Relações entre diferentes variáveis climáticas na qualidade do fruto.
- Predição de intensidade das Ondas de calor.
- Utilização de teoria dos jogos na negociação entre envolvidos do mercado da Manga.

## Referências

- ABRAFRUTAS. *Novo mercado para mangas nordestinas*. 2018. Url <https://abrafrutas.org/2018/08/novo-mercado-para-mangas-nordestinas/>. Accessed: 2023 feb 03. Citado na página 21.
- ABRAFRUTAS. *Ascenza lança no Brasil produto que aumenta rendimento da manga*. 2019. Url <https://abrafrutas.org/2019/07/ascenza-lanca-no-brasil-produto-que-aumenta-rendimento-da-manga/>. Accessed: 2023 feb 03. Citado na página 21.
- ABRAFRUTAS. *Excesso de mangas faz preços caírem até 40% no mercado*. 2019. Url <https://abrafrutas.org/2019/10/excesso-de-mangas-faz-precos-cairem-ate-40-no-mercado/>. Accessed: 2022 Sept 10. Citado na página 13.
- ABRAFRUTAS. *Hortifrutti: Onda de calor pode afetar a produção de frutas e hortaliças?* 2020. Accessed: 2023 april 10. Citado 2 vezes nas páginas 14 e 26.
- ABRAFRUTAS. *Manga foi a fruta mais exportada pelo Brasil em 2021*. 2022. Url <https://abrafrutas.org/2022/01/manga-foi-a-fruta-mais-exportada-pelo-brasil-em-2021/>. Accessed: 2022 Sept 10. Citado 3 vezes nas páginas 13, 21 e 24.
- ABRAFRUTAS. *Acordo fortalece fruticultura no Vale do São Francisco*. 2023. Url <https://abrafrutas.org/2023/01/acordo-fortalece-fruticultura-no-vale-do-sao-francisco/>. Accessed: 2023 mar 03. Citado na página 24.
- ABRAFRUTAS. *Conheça os tipos de manga mais comuns no Brasil*. 2023. Url <https://abrafrutas.org/2023/01/conheca-os-tipos-de-manga-mais-comuns-no-brasil/>. Accessed: 2023 Mar 03. Citado na página 20.
- ALVES, A. V. P. et al. Análise dos métodos de estimação para os parâmetros das distribuições de gumbel e gev em eventos de precipitações máximas na cidade de cuiabá-mt. *Revista Ciência Agronômica*, v. 6, n. 1, 2013. Citado na página 18.
- ARAÚJO, E. M. et al. Aplicação de seis distribuições de probabilidade a séries de temperatura máxima em iguatu-ce. *Revista Ciência Agronômica*, SciELO Brasil, v. 41, p. 36–45, 2010. Citado na página 18.
- BATISTA, D. D. C.; BARBOSA., M. A. G. Doenças da mangueira. In: *SIMPÓSIO INTERNACIONAL DE VITIVINICULTURA, 1.; FEIRA NACIONAL DA AGRICULTURA IRRIGADA - FENAGRI, 2008, Petrolina. Minicursos. Petrolina: Prefeitura Municipal: ValeXport: Embrapa Semi-Árido, 2008. [S.l.: s.n.], 2008.* Citado na página 22.
- CAMARGO, M. B. P. de et al. Probabilidade de ocorrência de temperaturas mínimas absolutas mensais e anual no estado de são paulo. *Scielo*, n. 1993, 1993. Citado na página 18.
- CAMPBELL R.J.; CAMPBELL, C. The tommy atkins mango. *ournal of the American Pomological Society*, v. 60, p. 55–57, 2006. Citado na página 13.

- CASELLA, G.; BERGER, R. L. *Statistical Inference*. 2. ed. Boston, MA: Cengage Learning, 2001. Citado na página 35.
- CLERC, M. *Particle swarm optimization*. New York, NY: John Wiley & Sons, 2010. v. 93. Citado 2 vezes nas páginas 37 e 38.
- CMJORNAL. *Ondas de calor serão mais frequentes e intensas pelo menos até 2060*. 2022. Url <https://www.cmjornal.pt/sociedade/clima/detalhe/ondas-de-calor-serao-mais-frequentes-e-intensas-pelo-menos-ate-2060>. Accessed: 2022-08-21. Citado na página 14.
- COLES, S. et al. *An introduction to statistical modeling of extreme values*. New York, NY: Springer, 2001. v. 208. Citado na página 31.
- CORREIO. *Novo mercado para mangas nordestinas*. 2018. Url <https://www.correio24horas.com.br/noticia/mid/novo-mercado-para-mangas-nordestinas/>. Accessed: 2023 feb 03. Citado na página 21.
- COSTA, A. N. d. et al. Recomendações técnicas para a produção de manga. n. 155, 2014. ISSN 1519-2059. Citado na página 13.
- CRISTINO, C. T.; ŹEBROWSKI, P.; WILDEMEERSCH, M. Assessing the time intervals between economic recessions. *PloS one*, v. 15, n. 5, p. e0232615, 2020. Accessed: 2022 Jul 19. Citado 4 vezes nas páginas 19, 32, 33 e 34.
- DINIZ, F. R.; GONÇALVES, F. L. T. Ondas de calor na cidade de são paulo: análises no presente e projeções perante o cenário rcp8,5 das mudanças climáticas. In: . [S.l.: s.n.], 2018. Citado na página 26.
- EMBRAPA. Cultivo da mangueira. *Embrapa Semi-Árido*, n. 1807-0027, 2014. Citado na página 13.
- EMBRAPA SEMIÁRIDO. *Colheita e pós-colheita*. 2004. Url [http://www.cpatsa.embrapa.br:8080/sistema\\_producao/spmanga/colheita.htm](http://www.cpatsa.embrapa.br:8080/sistema_producao/spmanga/colheita.htm). Accessed: 2022 april 11. Citado na página 22.
- EMBRAPA SEMIÁRIDO. *Manejo do Solo*. 2015. Url [https://www.spo.cnptia.embrapa.br/conteudo?p\\_p\\_id=conteudoportlet\\_WAR\\_sistemasdeproducaolf6\\_1ga1ceportlet&p\\_p\\_lifecycle=0&p\\_p\\_state=normal&p\\_p\\_mode=view&p\\_p\\_col\\_id=column-1&p\\_p\\_col\\_count=1&p\\_r\\_p\\_-76293187\\_sistemaProducaoId=7743&p\\_r\\_p\\_-996514994\\_topicoId=8290](https://www.spo.cnptia.embrapa.br/conteudo?p_p_id=conteudoportlet_WAR_sistemasdeproducaolf6_1ga1ceportlet&p_p_lifecycle=0&p_p_state=normal&p_p_mode=view&p_p_col_id=column-1&p_p_col_count=1&p_r_p_-76293187_sistemaProducaoId=7743&p_r_p_-996514994_topicoId=8290). Accessed: 2023 april 02. Citado na página 21.
- EMBRECHTS, P.; KLÜPPELBERG, C.; MIKOSCH, T. *Modelling extremal events: for insurance and finance*. New York, NY: Springer Science & Business Media, 2013. v. 33. Citado na página 31.
- FERREIRA, R. J.; FIRMINO, P. R. A.; CRISTINO, C. T. A mixed kijima model using the weibull-based generalized renewal processes. *PLoS ONE*, Public Library of Science, v. 10, n. 7, p. e0133772, 2015. Citado 2 vezes nas páginas 32 e 33.

- FILHO, J. M. P. L.; TEIXEIRA, A. H. de C. *Temperatura*. 2021. Url<https://www.embrapa.br/agencia-de-informacao-tecnologica/cultivos/manga/pre-producao/caracteristicas/clima/temperatura>. Accessed: 2023 mar 03. Citado na página 13.
- FOLHAPE. *Com 41°C, Petrolina bate recorde de temperatura registrada em Pernambuco no ano*. 2021. Url <https://www.folhape.com.br/noticias/com-41oc-petrolina-bate-recorde-de-temperatura-registrada-em/198898/>. Accessed: 2022 ago 05. Citado na página 14.
- FRANCISCO, S.; SIMÕES, S. Testes de aderência aplicados à distribuição da profundidade de trinca em tubos do gerador de vapor de uma planta nuclear pwr). *Vetor*, v. 31, n. 1, 2021. Citado na página 40.
- FRUTA, T. *Manga exportação gera mais lucro ao produtor mas exige investimento e atenção*. 2022. Url <https://www.todafruta.com.br/manga-exportacao-gera-mais-lucro-ao-produtor-mas-exige-investimento-e-atencao>. Accessed: 2023 feb 03. Citado na página 24.
- GLOBO RURAL. *Vale do São Francisco é responsável por 87% do volume de manga exportado no país*. 2021. Url <https://g1.globo.com/economia/agronegocios/globo-rural/noticia/2021/04/25/vale-do-sao-francisco-e-responsavel-por-87percent-do-volume-de-manga-exportado-no-pais.ghtml>. Accessed: 2022 sept 10. Citado na página 13.
- GÓMEZ, Y. M.; BOLFARINE, H.; GÓMEZ, H. W. Gumbel distribution with heavy tails and applications to environmental data. *Mathematics and Computers in Simulation*, Elsevier, v. 157, p. 115–129, 2019. Citado na página 31.
- GONÇALVES, N. et al. Caracterização física e química dos frutos de cultivares de mangueira (mangifera indica). *Ciência e Agrotecnologia*, v. 22, n. 1, 1998. Citado na página 20.
- HENLEY, E. J.; KUMAMOTO, H. *Probabilistic risk assessment and management for engineers and scientists*. [S.l.]: IEEE Press (2nd Edition), 1996. Citado na página 41.
- HORTIFRUTI/CEPEA. *Hortifruti Brasil e Centro de Estudos Avançados em Economia Aplicada (Cepea), da Escola Superior de Agricultura Luiz de Queiroz – Universidade de São Paulo (Esalq-USP)*. 2023. Url <https://www.hfbrasil.org.br/br/banco-de-dados-precos-medios-dos-hortifruticolas.aspx>. Accessed: 2023 Mar 30. Citado na página 15.
- INMET. *Alerta de Onda de calor*. 2022. Url <https://alertas2.inmet.gov.br/40139>. Accessed: 2023 april 01. Citado na página 26.
- INMET. *Normais Climatológicas*. 2022. Url <https://portal.inmet.gov.br/normais>. Accessed: 2023 April 11. Citado na página 27.
- INMET. *Onda de calor: mais de 50 cidades brasileiras registram as maiores temperaturas do ano até o momento*. 2022. Url <https://portal.inmet.gov.br/noticias/onda-de-calor-mais-de-50-cidades-registraram-nesta-segunda-feira-12-as-maiores-temperaturas-do-ano-até-o-momento>. Accessed: 2023 april 01. Citado na página 26.

- INMET. *DADOS HISTÓRICOS ANUAIS*. 2023. Url <https://portal.inmet.gov.br/dadoshistoricos>. Accessed: 2023-31-03. Citado na página 42.
- INSTITUTE, P. . P. *International Soil Fertility Manual*. Piracicaba: Potafos, 1995. Citado na página 23.
- IPCC. *Mudanças Climáticas 2022: Impactos, Adaptação e Vulnerabilidade*. 2022. Url <https://www.ipcc.ch/report/ar6/wg2/>. Accessed: 2023 Apr 10. Citado na página 14.
- KENNEDY, J.; EBERHART, R. Particle swarm optimization. In: IEEE. *Proceedings of ICNN'95-international conference on neural networks*. [S.l.], 1995. v. 4, p. 1942–1948. Citado na página 37.
- KIJIMA, M.; MORIMURA, H.; SUZUKI, Y. Periodical replacement problem without assuming minimal repair. *European Journal of Operational Research*, Elsevier, v. 37, n. 2, p. 194–203, 1988. Citado na página 32.
- KIJIMA, M.; SUMITA, N. A useful generalization of renewal theory: counting process governed by non-negative markovian increments. *Journal of Applied Probability*, v. 23, p. 71–88, 1986. Citado na página 29.
- MANCIN, C.; MELO, B.; SOUZA, O. P. *CULTURA DA MANGUEIRA*. 2003. Url <http://www.fruticultura.iciag.ufu.br/manga.html>. Accessed: 2023 feb 03. Citado 2 vezes nas páginas 22 e 25.
- METSUL. *Ondas de Calor mais frequentes, intensas e duradouras no Brasil*. 2022. Url <https://metsul.com/ondas-de-calor-mais-frequentes-intensas-e-duradouras-no-brasil/>. Accessed: 2023 March 02. Citado na página 26.
- MOUCO, M. A. do C.; ALBUQUERQUE, J. A. S. de. Manejo da floração. *Embrapa Semi-Árido*, n. 2, 2002. ISSN 1807-0027. Citado na página 25.
- MOURA, M. D. C. e. a. Avaliação bayesiana da eficácia da manutenção via processo de renovação generalizado). *Pesquisa Operacional*, v. 27, n. 03, p. 569–589, 2007. Citado na página 29.
- NETO, C. A. da S. et al. *Melhoria da segurança dinâmica baseada em análise estocástica e metaheurística*. 2012. Url <https://doi.org/10.1590/S0103-17592012000200008>. Accessed: 2023 mar 03. Citado na página 38.
- OFFICE, M. *What is a heatwave?* 2022. Url <https://www.metoffice.gov.uk/weather/learn-about/weather/types-of-weather/temperature/heatwave>. Accessed: 2023 April 13. Citado 2 vezes nas páginas 14 e 26.
- OLIVEIRA, C. C. F.; CRISTINO, C. T.; FIRMINO, P. R. A. In the kernel of modelling repairable systems: a goodness of fit test for weibull-based generalized renewal processes. *Journal of Cleaner Production*, Elsevier, v. 133, p. 358–367, 2016. Citado na página 33.
- OLIVEIRA, C. C. F.; FIRMINO, P. R. A.; CRISTINO, C. T. A tool for evaluating repairable systems based on generalized renewal processes. *Reliability Engineering & System Safety*, Elsevier, v. 183, p. 281–297, 2019. Citado na página 32.

- OLIVEIRA, D. d. S. et al. Vitamina c, carotenoides, fenólicos totais e atividade antioxidante de goiaba, manga e mamão procedentes da ceasa do estado de minas gerais. *Health Sciences*, v. 33, n. 1, 2011. Citado na página 20.
- OMM. *WMO Guidelines on the Calculation of Climate Normals*. 2017. Url [https://library.wmo.int/doc\\_num.php?explnum\\_id=4166](https://library.wmo.int/doc_num.php?explnum_id=4166). Accessed: 2023 April 12. Citado na página 27.
- OMM. *O excesso de mortalidade relacionada ao calor na Europa ultrapassou 15.000 mortes*. 2023. Url <https://public.wmo.int/es/media/comunicados-de-prensa/el-informe-anual-de-la-omm-pone-de-relieve-el-avance-continuo-del-cambio>. Accessed: 2023 april 23. Citado na página 25.
- OMS. *OMS lança perfis de países para monitorar ações em saúde e mudanças climáticas*. 2015. Url <https://www.who.int/news/item/16-11-2015-who-launches-country-profiles-to-monitor-action-on-health-and-climate-change>. Accessed: 2022 Dec 21. Citado na página 14.
- PEDREGAL, P. *Introduction to Optimization*. 1. ed. New York, NY: Springer, 2004. Citado na página 37.
- PERKINS, S. E.; ALEXANDER, L. V. On the measurement of heat waves. *Journal of Climate*, v. 26, n. 13, 2013. Citado 2 vezes nas páginas 14 e 27.
- PETERSON, T. Report on the activities of the working group on climate change detection and related rapporteurs. *World Meteorological Organization (WMO)*, n. 1071, 2001. Citado 2 vezes nas páginas 14 e 26.
- PORTAL G1. *Com 40,2°C, Reino Unido bate recorde na história do país em meio a onda de calor pela Europa*. 2022. Url <https://g1.globo.com/mundo/noticia/2022/07/19/calor-europa-incendio-temperatura-recorde.ghtml>. Accessed: 2023 april 10. Citado na página 25.
- RAMÍREZ, F. D. Mango (mangifera indica l.) flowering physiology. *Scientia Horticulturae*, v. 126, p. 65–72, 2010. Citado na página 13.
- REIS, R. et al. *Atributos físico-químicos de frutos de variedades de mangueira cultivadas em sistema orgânico*. [S.l.]: Cruz das Almas, BA: Embrapa Mandioca e Fruticultura, 2019., 2019. Citado na página 20.
- ROCHA JÚNIOR, R. L. da et al. Análise de extremos de temperatura no sul do brasil. *Revista brasileira de climatologia*, n. 2022, 2022. Citado na página 18.
- RODRIGUES, J. C.; SOUZA, P. J. D. O. P. D.; LIMA, R. T. D. Estimativa de temperaturas basais e exigência térmica em mangueiras no nordeste do estado do pará. 2013. Citado na página 13.
- RODRIGUES, J. C.; SOUZA, P. J. de Oliveira Ponte de; LIMA, R. T. de. Estimativa de temperaturas basais e exigência térmica em mangueiras no nordeste do estado do pará. *Revista Brasileira de Meteorologia*, v. 31, n. 4, 2016. Citado na página 13.
- SANSIGOLO, C. A. Distribuições de extremos de precipitação diária, temperatura máxima e mínima e velocidade do vento em piracicaba, sp (1917-2006). *Revista Brasileira de Meteorologia*, v. 23, n. 3, 2008. Citado na página 18.

- SCUDINO, P. A. *A utilização de alguns testes estatísticos para análise da variabilidade do preço do mel nos municípios de Angra dos Reis e Mangaratiba, Estado do Rio de Janeiro.* 51 p. Monografia (Trabalho de Conclusão de Curso de Graduação) — Universidade Federal Rural do Rio de Janeiro, Rio de Janeiro, 2008. Citado na página 40.
- SEMIÁRIDO, E. *Cultivo da Mangueira.* 2010. Url <https://www.infoteca.cnptia.embrapa.br/bitstream/doc/884451/1/CultivodaMangueira.pdf>. Accessed: 2023 feb 03. Citado na página 21.
- SILVA, A. A. M. d. *UM ESTUDO DO MÉTODO ENXAME DE PARTÍCULAS PARA SE DETERMINAR MÍNIMOS GLOBAIS DE FUNÇÕES NÃO LINEARES.* 72 p. Monografia (Trabalho de Conclusão de Curso de Graduação) — UNIVERSIDADE FEDERAL FLUMINENSE, Santo Antônio de Pádua, 2017. Citado na página 38.
- SOARES, M. R.; SANTOS, J. R.; NEVES, C. S. *Incêndios em Portugal. A situação ao minuto.* 2022. Url [https://www.rtp.pt/noticias/pais/incendios-em-portugal-a-situacao-ao-minuto\\_e1420749](https://www.rtp.pt/noticias/pais/incendios-em-portugal-a-situacao-ao-minuto_e1420749). Accessed: 2022 Ago 23. Citado na página 16.
- SOCIEDADE NACIONAL DE AGRICULTURA. Ovinos/caprinos: Como aumentar a produção de cordeiros. *A Lavoura*, n. 703, 2014. Citado na página 25.
- TEIXEIRA, J. M. P. L. A. H. de C. *Temperatura - Manga.* 2021. Url <https://www.embrapa.br/agencia-de-informacao-tecnologica/cultivos/manga/pre-producao/caracteristicas/clima/temperatura>. Accessed: 2022 Dez 23. Citado na página 22.
- UNDRR. Relatório Técnico, *O custo humano dos desastres: uma visão geral dos últimos 20 anos (2000-2019).* 2020. Url <https://www.undrr.org/publication/human-cost-disasters-overview-last-20-years-2000-2019>. Citado na página 25.
- VEBER, B.; NAGODE, M.; FAJDIGA, M. Processo generalizado de renovação de sistemas reparáveis baseado em mistura finita de weibull). *Engenharia de Confiabilidade e Segurança do Sistema*, v. 93, n. 10, p. 1461–1472, 2008. Citado na página 28.
- YANEZ, M.; JOGLAR, F.; MODARRES, M. Generalized renewal process for analysis of repairable systems with limited failure experience. *Reliability Engineering & System Safety*, Elsevier, v. 77, n. 2, p. 167–180, 2002. ISSN 0951-8320. Citado na página 28.