

UNIVERSIDADE FEDERAL DE MINAS GERAIS
Instituto de Ciências Exatas
Programa de Pós-Graduação em Estatística

Leonardo Brandão Freitas do Nascimento

**Classe de Distribuições Multivariadas para Dados Extremos Dependentes,
Limitados e não Estacionários**

Belo Horizonte
2023

Leonardo Brandão Freitas do Nascimento

**Classe de Distribuições Multivariadas para Dados Extremos Dependentes,
Limitados e não Estacionários**

Tese apresentada ao Programa de Pós-Graduação em Estatística da Universidade Federal de Minas Gerais, como requisito parcial à obtenção do título de Doutor em Estatística.

Orientador: Luiz Henrique Duczmal
Coorientador: Max Sousa de Lima

Belo Horizonte
2023

Nascimento, Leonardo Brandão Freitas do.

N244c

Classe de distribuições multivariadas para dados extremos dependentes, limitados e não estacionários [recurso eletrônico / Leonardo Brandão Freitas do Nascimento. – 2023.

1 recurso online (129 f. il, color.) : pdf.

Orientador: Luiz Henrique Duczmal.

Coorientador: Max Sousa de Lima

Tese (doutorado) - Universidade Federal de Minas Gerais, Instituto de Ciências Exatas, Departamento de Estatística.

Referências: f. 121-127

1. Estatística – Teses. 2. Distribuições estáveis – Teses. 3. Distribuição (Probabilidades) - Valores extremos – Teses. 4. Mín - Estável. I. Duczmal, Luiz Henrique. II. Lima, Max Sousa de. III. Universidade Federal de Minas Gerais, Instituto de Ciências Exatas, Departamento de Estatística. III. Título.

CDU 519.2(043)

Ficha catalográfica elaborada pela bibliotecária Belkiz Inez Rezende Costa
CRB 6/1510 Universidade Federal de Minas Gerais - ICEx



UNIVERSIDADE FEDERAL DE MINAS GERAIS

PROGRAMA DE PÓS-GRADUAÇÃO EM ESTATÍSTICA



FOLHA DE APROVAÇÃO

Classe de distribuições multivariadas para dados extremos dependentes, limitados e não estacionários.

LEONARDO BRANDÃO FREITAS DO NASCIMENTO

Tese submetida à Banca Examinadora, designada pelo Colegiado do Programa de Pós-Graduação em ESTATÍSTICA, como requisito para obtenção do grau de Doutor em ESTATÍSTICA, área de concentração ESTATÍSTICA E PROBABILIDADE.

Aprovada, em 03 de agosto de 2023, pela banca constituída pelos membros:

Prof. Luiz Henrique Duczmal - Orientador
DEST/UFMG

Prof. Max Sousa de Lima - Co-orientador
ICE/UFAM

Prof. Elias T. Krainski
Dept. Statistics - King Abdullah University of Sciences

Prof. Alexandre Louzeiros Rodrigues
DEST/UFES

Prof. Vinícius Diniz Mayrink
DEST/UFMG

Prof. Fabio Nogueira Demarqui
DEST/UFMG

Belo Horizonte, 3 de agosto de 2023.

Agradecimentos

Agradeço à Deus pela saúde, sabedoria e perseverança durante esse tempo de estudo e dedicação. À minha esposa e filha pelo carinho, amor e paciência. Aos meus pais pelas orações. Aos meus irmãos Lennon e Victor pelas conversas e companhia. Aos meus orientadores Max Lima (UFAM) e Luiz Duczmal (UFMG) pela orientação e amizade. Aos meus amigos da pós-graduação pelo conhecimento e momentos compartilhados. Por fim, agradeço à Fundação de Amparo à Pesquisa do Estado do Amazonas (FAPEAM) pelo apoio financeiro e ao Laboratório de Sistemas Embarcados da Universidade do Estado do Amazonas (UEA) pelo apoio de equipamentos.

Resumo

O modelo proposto neste trabalho pertence à classe das distribuições mín-Estável e foi obtido através da marginalização de uma distribuição baseline V condicionada a um campo aleatório latente com distribuição α -estável positiva. Ele é adequado para modelar valores extremos dependentes, não estacionário e pertencentes a intervalos limitados. O modelo pode ser aplicado em séries temporais e dados com interação espaço-tempo. Considerando o intervalo $(0, 1)$, foram obtidas propriedades que medem a dependência, como o extremograma temporal e espacial, além de probabilidades condicionais úteis para realizar inferências e obter resultados práticos relevantes. No contexto de séries temporais, um algoritmo EM foi desenvolvido para estimar os parâmetros do modelo e uma expressão analítica para a matriz de informação observada foi obtida, permitindo que intervalos de confiança e testes de hipótese possam ser realizados através da distribuição assintótica dos estimadores. A aplicação do modelo aos dados de umidade relativa do ar em Manaus identificou meses com alta probabilidade de ocorrência de umidade mínima acima de 0.70, indicando potenciais riscos à saúde humana. Na análise espaço-tempo, o modelo mostrou-se adequado para modelar a dependência espacial nos mínimos extremos. A análise revelou padrões espaciais climáticos extremos de umidade no estado do Amazonas que podem representar riscos à saúde humana. Estimativas dos parâmetros foram obtidas usando o algoritmo MCEM condicional, onde a variância e testes de hipóteses dos estimadores foram realizados com o método *bootstrap* condicional.

Palavras-chave: Mín-Estável; Distribuição Estável; Umidade, Extremos

Abstract

The model proposed in this work belongs to the class of Min-stable distributions and was obtained through the marginalization of a baseline distribution V conditioned on a latent random field with a positive α -stable distribution. It is suitable for modeling dependent extreme values that are non-stationary and within limited intervals. The model can be applied to time series and data with space-time interactions. Considering the interval $(0, 1)$, properties measuring dependence, such as the temporal and spatial extremogram, were obtained, along with conditional probabilities useful for making inferences and obtaining relevant practical results. In the context of time series, an EM algorithm was developed to estimate the model parameters, and an analytical expression for the observed information matrix was obtained, enabling confidence intervals and hypothesis tests to be conducted based on the asymptotic distribution of the estimators. Applying the model to air relative humidity data in Manaus identified months with a high probability of minimum humidity exceeding 0.70, indicating potential risks to human health. In the space-time analysis, the model proved suitable for modeling spatial dependence in extreme minima. The analysis revealed extreme climatic spatial patterns of humidity in the state of Amazonas that could pose risks to human health. Parameter estimates were obtained using the conditional MCEM algorithm, where the variance and hypothesis tests of the estimators were performed using the conditional bootstrap method.

Keywords: Min-Stable; Stable Distribution, Humidity, Extremes

Lista de Figuras

3.1	Função de excedência (lado esquerdo) e função densidade (lado direito) do modelo Kumaraswamy Exponencializada Multivariado para diferentes valores de α	37
3.2	Geração de amostras bivariadas do modelo Kumaraswamy Exponencializada Multivariado para diferentes valores de α	43

Lista de Algoritmos

2.1	Algoritmo para gerar amostras da distribuição α -estável positiva.	20
2.2	Algoritmo EM.	29
2.3	Algoritmo ECM.	30
2.4	Algoritmo MCEM.	31

Sumário

1	Introdução	11
1.1	Contextualização	11
1.2	Objetivo e Contribuição	12
1.3	Divisão do Trabalho	14
2	Fundamentação Teórica	15
2.1	Distribuição Kumaraswamy	15
2.1.1	Parametrização Padrão	15
2.1.2	Reparametrização: mediana - dispersão	17
2.2	Distribuição α –Estável	18
2.3	Distribuições Máx e Mín Estáveis	20
2.4	Distribuições Condicionais G -Exponencializada	24
2.5	Marginais Univariadas de Distribuições Condicionais G -Exponencializadas	26
2.6	Métodos de Estimações	27
2.6.1	Algoritmo EM	27
2.6.2	Algoritmo EM Condicional	28
2.6.3	Algoritmo Monte Carlo EM	29
2.6.4	Matriz de Informação Observada	30
3	Classe de Distribuições Multivariadas para Dados Extremos Dependentes, Limitados e não Estacionários	33
3.1	Construção da Classe	33
3.1.1	Classe Kumaraswamy Exponencializada	35
3.2	Propriedades	36
3.2.1	Função Densidade Conjunta	36
3.2.2	Distribuições do Mínimo e Marginal	39
3.2.3	Função Densidade Por Grupos	40
3.2.4	Geração de Vetores Aleatórios Por Grupos	40
3.2.5	Inferência Por Grupos	42
4	Artigo 1	44
5	Artigo 2	93
6	Conclusão	119

Referências	121
Apêndice A Demonstração da Função Densidade por Grupos	128

CAPÍTULO 1

Introdução

1.1 Contextualização

A Teoria do Valor Extremo (TVE) é uma área da estatística que realiza modelagem e análise do comportamento de eventos ou valores extremos. Segundo Salvadori et al. (2007), a teoria clássica univariada de valores extremos foi desenvolvida por Fréchet (1927) e Fisher and Tippett (1928) e estudos posteriores sobre distribuições assintóticas foram apresentados em Gnedenko (1943), James (1959) e De Haan (1976).

O principal objetivo dessa teoria é oferecer modelos, propriedades e métodos para melhor compreender e prever a distribuição da cauda de uma variável aleatória. Diante disso, a suposição principal da TVE é que eventos extremos são regidos por propriedades estatísticas distintas em relação aos demais valores observados. Alguns livros (Embrechts et al., 1997; Coles et al., 2001; Beirlant et al., 2004; Haan and Ferreira, 2006) e artigos (Davison et al., 2012; Cooley et al., 2012; Davison and Huser, 2015; Wadsworth and Tawn, 2022) já foram escritos sobre a teoria e aplicações de modelos univariados, multivariados e extremos espaciais. Diante disso, percebe-se a TVE é particularmente útil quando lidamos com eventos que ocorrem com pouca frequência, mas possuem um impacto significativo, como grandes perdas financeiras, eventos climáticos extremos ou falhas raras em sistemas de engenharia. Nesse sentido, nota-se que TVE tem aplicações em diversas áreas.

No contexto da análise de séries temporais são feitas suposições sobre a distribuição dos dados a partir do pressuposto de erros Gaussianos ou normalmente distribuídos. No entanto, muitos fenômenos como retornos financeiros e eventos climáticos extremos apresentam um comportamento de caudas pesadas, nesses casos, eventos extremos ocorrem com mais frequência do que o previsto por uma distribuição Gaussiana. Nesse sentido, a Teoria do Valor Extremo pode ser usada para estudar e prever valores ou eventos extremos que ocorrem ao longo do tempo. Uma classe comumente usada de modelos é baseada na Distribuição Pareto Generalizada (GPD) ou na Distribuição Geral de Valores Extremos (GEV). Essas distribuições fornecem um quadro flexível para modelar o comportamento

da cauda dos dados através do índice ou parâmetro de cauda. Tal parâmetro quantifica a intensidade e o decaimento das probabilidades de caudas. Além disso, controla a estrutura de dependência dos dados, pois a ocorrência de um evento extremo em determinado momento pode afetar a ocorrência de eventos extremos em momentos subsequentes no tempo.

Porém, algumas quantidades extremas como temperatura, precipitação ou poluição do ar são medidas em estações de monitoramento que possuem lugares determinados. Mais especificamente, assuma que $Y_1(s), Y_2(s), \dots$ seja uma sequência de processos aleatórios independentes definido em torno de uma região $\mathcal{S} \subseteq \mathbb{R}^d$ e $s_1, \dots, s_D \in \mathcal{S}$ uma coleção de localizações. Na análise de extremos dentro do contexto de séries temporais, $Y_t(s_j)$, $t = 1, \dots$, pode ser analisada separadamente em cada localização s_j . Porém, pode ser mais razoável modelar as séries conjuntamente em um modelo que relaciona os dados em diferentes localizações. Nesses casos, faz-se necessária a modelagem espacial.

A análise de valores extremos espacial é a combinação de duas áreas da estatística: Geoestatística e a Teoria do Valor Extremo. Em relação a geoestatística (Banerjee et al., 2004; Cressie, 1993; Gelfand et al., 2010) fornece ferramentas para descrever e prever padrões espaciais, ou seja, é o ramo da estatística que analisa a relação de dependência entre valores em diferentes localidades, tendo como premissa que locais vizinhos tendem a ter valores mais similares em comparação com locais mais distantes. Basicamente, a dependência espacial é modelada através de uma função de covariância que regula a associação entre os valores em diferentes locais. Porém, ao considerar eventos extremos são necessários modelos que considerem as propriedades de dependência adequada das distribuições extremas. Nesse sentido, segundo Thibaud et al. (2013), o processo máx-Estável (Davison and Gholamrezaee, 2012; De Haan, 1984; Haan and Ferreira, 2006) é uma generalização da teoria de valores extremos (Fisher and Tippett, 1928) para o contexto espacial, e portanto, fornece distribuições multivariadas para máximos em qualquer dimensão. Estudos iniciais sobre esse processo são apresentados em Smith (1990); Coles and Tawn (1996); Coles (1993). Schlather and Tawn (2003) apresenta uma forma de analisar a estrutura de dependência de processos máx-Estáveis através da função de coeficientes extremos. Outras estruturas são apresentadas em Ledford and Tawn (1996); Schlather and Tawn (2002). Uma revisão teórica é apresentado em Ribatet (2013) e algumas aplicações recentes podem ser vistas em Skahill et al. (2023); Portero Serrano et al. (2020) e Jurado et al. (2020).

1.2 Objetivo e Contribuição

Diante do cenário apresentado em 1.1, percebe-se que dados extremos necessitam de propriedades específicas para melhor compreensão da natureza que os formulam. Além disso, modelos que apresentem propriedades para dados limitados também se faz

necessário para capturar comportamentos das observações e, apesar de que a literatura apresente modelos para dados limitados, (Rocha and Cribari-Neto, 2009; Bayer et al., 2017; Alanzi et al., 2022), estes não estão no contexto de valores extremos ou são para variáveis aleatórias independentes e identicamente distribuídas. Diante disso, a classe proposta terá como base o trabalho de Fougères et al. (2009) que propôs uma distribuição de valores extremos multivariados para dados dependentes a partir de misturas de distribuições α -estável positiva (Nolan, 2020) tendo como base a distribuição Gumbel. Portanto, o presente trabalho tem como objetivo apresentar uma classe de distribuições de valores extremos multivariados mín-Estável para dados dependentes, limitados e não estacionários. Tal classe, será obtida a partir da marginalização de uma distribuição baseline V condicionada a um campo aleatório latente utilizando as propriedades da distribuição α -estável positiva (Nolan, 2020; Fougères et al., 2009). Um caso particular é apresentado quando V é uma função de distribuição de uma variável uniforme no intervalo $(0,1)$, resultando no modelo Kumaraswamy Exponencializada. Assim sendo, as principais contribuições dessa classe é fornecer propriedades que possuem uma aplicação relevante no contexto de valores extremos, como probabilidades condicionais de ocorrerem eventos por duas vezes consecutivos na área de séries temporais e um índice que controla a dependência entre observações em localizações (espacial) ou instantes de tempos (séries temporais) diferentes por exemplo. Além disso, essa classe possibilita utilizar covariáveis para melhor explicar os parâmetros de tendência temporal/espacial e dispersão temporal presentes em dados não estacionários. Por fim, um método de estimação (para séries temporais de valores extremos) é proposto, a partir do algoritmo EM (Dempster et al., 1977), e a variância assintótica dos estimadores é calculada. Destaca-se que tal método não depende da ordem da série. Sob o ponto de vista prático, esse modelo pode ser aplicado em qualquer observações pertencentes a intervalos limitados. Em particular, será aplicado nos dados de umidade relativa, $(0,1)$, do Estado do Amazonas. Essa análise se torna importante devido ao fato de que a umidade relativa é influenciada por diversos fatores, como temperatura, a pressão atmosférica, altitude e a proximidade de fontes de água, como oceanos, rios ou lagos. Ela desempenha um papel importante nas condições climáticas e pode afetar nosso conforto e bem-estar. Baixas umidades relativas, por exemplo, podem resultar em uma sensação de pele seca, irritação nos olhos, garganta seca e ressecamento das vias respiratórias. Além disso, a baixa umidade pode levar à desidratação mais rápida, especialmente em climas quentes ou durante atividades físicas intensas. Por outro lado, altas umidades relativas podem causar desconforto devido à dificuldade de evaporação do suor, o que dificulta o resfriamento do corpo. Isso pode resultar em sensação de abafamento, transpiração excessiva, dificuldade em respirar e um ambiente propício ao crescimento de mofo e ácaros. Nesse cenário, o estudo de eventos extremos, umidades acima de 0.70 ou abaixo de 0.30, são fundamentais para compreender a natureza dos dados.

1.3 Divisão do Trabalho

No Capítulo 2 é apresentada a revisão teórica das principais definições utilizadas para o processo de inferência e desenvolvimento do modelo . Mais especificamente, apresenta a definição de distribuições α , mín e máx estáveis, além de revisar o algoritmo EM e suas variações. O Capítulo 3 apresenta a classe de distribuições para dados extremos, limitados e não estacionário. Propriedades dessa classe são formuladas e uma análise gráfica é realizada. O Capítulo 4 mostra o primeiro artigo de aplicação intitulado: **Modelo de Regressão *Power* Mín-Estável para Séries Temporais com Valores Extremos em Intervalo Limitado**. Neste, é apresentada uma aplicabilidade do modelo proposto no contexto de séries temporais. O Capítulo 5 mostra o segundo artigo intitulado **Modelo Multivariado Espaço-Temporal Mín-Estável para Indicadores de Climats Extremos**. Nesse caso o modelo proposto foi aplicado no contexto espaço-tempo. Por fim, o Capítulo 6 apresenta a conclusão e considerações finais do trabalho.

CAPÍTULO 2

Fundamentação Teórica

2.1 Distribuição Kumaraswamy

A distribuição Kumaraswamy é uma classe de distribuições de probabilidade que foi proposta em Kumaraswamy (1980). A princípio, foi aplicada em diversos problemas hidrológicos como vazões de rios e precipitação. De forma geral, espera-se que essa classe seja aplicável em processos que são limitados inferiormente e superiormente. Diante disso, essa distribuição tem sido amplamente utilizada em diferentes áreas devido à sua flexibilidade e capacidade de modelar uma variedade de fenômenos. Em De Araujo Lima-Filho and Mariano Bayer (2021) foi proposta uma carta de controle para monitorar o comportamento de dados ambientais, mais especificamente a umidade relativa. Outro estudo sobre a umidade relativa também é apresentado em Bayer et al. (2017). Aplicações na área de confiabilidade são apresentados em Dey et al. (2017); Wang et al. (2017). Outros exemplos são apresenatdos em Fletcher and Ponnambalam (1996); Sánchez et al. (2007); Seifi et al. (2000); Ganji et al. (2006); Sundar and Subbiah (1989); Courard-Hauri (2007).

2.1.1 Parametrização Padrão

Segundo Mitnik and Baek (2013), as funções de densidade de probabilidade, distribuição acumulada e quantil da forma geral da distribuição Kumaraswamy, isto é, a forma da distribuição com suporte em qualquer intervalo aberto da reta real com limite superior d e limite inferior c são as seguintes:

$$g(y) = (d - c)^{-1} p q \left(\frac{y - c}{d - c} \right)^{p-1} \left[1 - \left(\frac{y - c}{d - c} \right)^p \right]^{q-1}$$

$$G(y) = 1 - \left[1 - \left(\frac{y - c}{d - c} \right)^p \right]^q$$

$$y = G^{-1}(u) = c + (d - c) \left[1 - (1 - u)^{\frac{1}{q}} \right]^{\frac{1}{p}}, \quad (2.1)$$

onde $c < y < d$, $0 < u < 1$, e $p > 0$ e $q > 0$ são parâmetros de forma. Denotaremos esta forma geral da distribuição por $Kw(p, q, c, d)$.

Neste caso, segue que a esperança e variância da forma geral da distribuição são, respectivamente,

$$E(Y) = c + (d - c) a B \left(1 + \frac{1}{p}, q \right)$$

e

$$Var(Y) = (d - c)^2 \left\{ q B \left(1 + \frac{2}{p}, \beta \right) - \left[q B \left(1 + \frac{1}{p}, b \right) \right]^2 \right\},$$

onde $B(j, v) = \int_0^1 s^{j-1} (1 - s)^{v-1} ds = \frac{\Gamma(j)\Gamma(v)}{\Gamma(j+v)}$ é a função Beta e $\Gamma(v) = \int_0^\infty t^{v-1} e^{-t} dt$ é a função Gama. Além disso, da Equação 2.1 segue que a mediana é dada por

$$md(Y) = \kappa = c + (d - c) \left(1 - 0.5^{\frac{1}{q}} \right)^{\frac{1}{p}}.$$

Para mais informações sobre os demais momentos da distribuição, vê Mitnik and Baek (2013). Neste trabalho será adotada um caso particular em que $c = 0$ e $d = 1$. Neste caso, funções de densidade de probabilidade, distribuição acumulada, quantil, excedência e taxa de falha são, respectivamente, apresentadas nas Definições 2.1.1, 2.1.2, 2.1.3, 2.1.4 e 2.1.5

Definição 2.1.1 (Função densidade de probabilidade). *Uma variável aleatória tem distribuição Kumaraswamy, $Y \sim Kw(p, q)$ com parâmetros $p > 0$ e $q > 0$, se sua função densidade de probabilidade é dada por*

$$g(y) = \begin{cases} pqy^{p-1}(1 - y^p)^{q-1} & , \quad 0 < y < 1, \\ 0 & , \quad \text{caso contrário.} \end{cases}$$

Definição 2.1.2 (Função de distribuição acumulada). *Se Y é uma variável aleatória tal que $Y \sim Kw(p, q)$, então sua função de distribuição acumulada é expressa por*

$$G(y) = 1 - (1 - y^p)^q, \quad y \in (0, 1).$$

Definição 2.1.3 (Função quantil). *Se a função de distribuição acumulada é dada pela Definição 2.1.2, então a função quantil é:*

$$y = G^{-1}(u) = \left[1 - (1 - u)^{\frac{1}{q}} \right]^{\frac{1}{p}},$$

onde $0 < u < 1$. Observe que se $u = 0.5$ então $md(Y) = \kappa = \left[1 - (1 - 0.5)^{\frac{1}{q}} \right]^{\frac{1}{p}}$ é a mediana de Y .

Definição 2.1.4 (Função de excedência). *Seja Y uma variável aleatória com função de distribuição G . Definimos função de excedência da variável Y , indicada por $\bar{G}$, como sendo o complementar da sua respectiva função distribuição, isto é,*

$$\bar{G}(y; \boldsymbol{\theta}) = 1 - G(y; \boldsymbol{\theta}) = 1 - [1 - (1 - y^p)^q] = (1 - y^p)^q, \quad y \in (0, 1).$$

Definição 2.1.5 (Função de risco). *A função de risco de uma variável aleatória Y com densidade $g(y)$ e função de excedência $\bar{G}(y)$ é dada por*

$$r(y; \boldsymbol{\theta}) = \frac{g(y; \boldsymbol{\theta})}{\bar{G}(y; \boldsymbol{\theta})} = \frac{pqy^{p-1}}{1 - y^p}, \quad y \in (0, 1),$$

onde $\boldsymbol{\theta}$ é o vetor de parâmetros.

2.1.2 Reparametrização: mediana - dispersão

Sem perda de generalidade, será assumido o caso particular para $c = 0$ e $d = 1$. Importante observar que a expressão para $E(Y)$ torna inviável uma reparametrização baseada na média. Neste caso, Mitnik and Baek (2013) apresenta duas reparametrizações através da mediana e do parâmetro de dispersão $1/p$ ou $1/q$. De fato, considerando a Definição 2.1.3 segue que

$$\begin{aligned} \kappa &= \left(1 - 0.5^{\frac{1}{q}}\right)^{\frac{1}{p}}, \\ \kappa^p &= \left(1 - 0.5^{\frac{1}{q}}\right), \\ 0.5^{\frac{1}{q}} &= 1 - \kappa^p, \\ \frac{1}{q} &= \frac{\ln(1 - \kappa^p)}{\ln 0.5} \\ q &= \frac{\ln 0.5}{\ln(1 - \kappa^p)}. \end{aligned} \tag{2.2}$$

Analogamente

$$p = \frac{\ln(1 - 0.5^{1/q})}{\ln \kappa}. \tag{2.3}$$

Considerando $d_p = p^{-1}$ e $d_q = q^{-1}$. Substituindo sucessivamente as Equações (2.3) e (2.2) nas Definições 2.1.1, 2.1.2 e 2.1.3 são obtidas duas reparametrizações. A primeira reparametrização denotada por $Kw(\kappa, p^{-1}, 0, 1)$ é:

$$\begin{aligned} g_p(y) &= \frac{p \ln(0, 5)}{\ln(1 - \kappa^p)} y^{p-1} (1 - y^p)^{\frac{\ln(0, 5)}{\ln(1 - \kappa^p)} - 1}, \\ G_p(y) &= 1 - (1 - y^p)^{\frac{\ln(0, 5)}{\ln(1 - \kappa^p)}}, \\ G_p^{-1}(u) &= \left[1 - (1 - u)^{\frac{\ln(1 - \kappa^p)}{\ln(0, 5)}}\right]^{\frac{1}{p}}. \end{aligned}$$

Enquanto a segunda reparametrização denotada por $Kw(\kappa, q^{-1}, 0, 1)$ é:

$$g_q(y) = \frac{\ln\left(1 - 0.5^{\frac{1}{q}}\right)}{\ln \kappa} q y^{\frac{\ln\left(1 - 0.5^{\frac{1}{q}}\right)}{\ln \kappa} - 1} \left[1 - y^{\frac{\ln\left(1 - 0.5^{\frac{1}{q}}\right)}{\ln \kappa}}\right]^{q-1},$$

$$G_q(y) = 1 - \left[1 - y^{\frac{\ln\left(1 - 0.5^{\frac{1}{q}}\right)}{\ln \kappa}}\right]^q,$$

$$G_q^{-1}(u) = y = \left[1 - (1 - u)^{\frac{1}{q}}\right]^{\frac{\ln \kappa}{\ln\left(1 - 0.5^{\frac{1}{q}}\right)}}.$$

2.2 Distribuição α -Estável

Segundo Nolan (2020), a classe de distribuições estáveis foi introduzida por Levy (1925) em seus estudos de somas de variáveis aleatórias. Além disso, Zolotarev (1986); Uchaikin and Zolotarev (1999); Meerschaert and Scheffler (2001) fornecem grande parte da teoria sobre distribuições estáveis nos anos seguintes.

Devido a sua característica de modelar dados de caudas pesadas, simétricos ou com outliers, essa classe de distribuição tem sido aplicada na área de economia, engenharia, física e em questões climáticas para analisar eventos extremos ou desastres naturais.

No entanto, as funções de densidade e distribuição que fazem parte dessa classe, com exceção dos tipos Normal, Cauchy e Lévy, não possuem formas analíticas tratáveis. Porém, sua função característica é bem estabelecida e conhecida. Em Nolan (2020) são apresentadas mais detalhes sobre as propriedades dessa classe e as diferentes parametrizações para as funções características. Para esse trabalho, em particular, foi adotada a forma apresentada na Definição 2.2.1.

Definição 2.2.1. *Uma variável aleatória Z possui distribuição estável, denotada por $S(\alpha, \varrho, \sigma, \mu)$ se sua função característica for dada por:*

$$\phi_Z(u) = \begin{cases} \exp \left\{ -\sigma^\alpha |u|^\alpha \left(1 + i\varrho \tan\left(\frac{\pi\alpha}{2}\right) \text{sign}(u) (|\sigma u|^{1-\alpha} - 1) + i\mu u \right) \right\}, & \text{se } \alpha \neq 1, \\ \exp \left\{ -\sigma |u| \left(1 + i\varrho \left(\frac{2}{\pi}\right) \text{sign}(u) \log(\sigma |u|) + i\mu u \right) \right\}, & \text{se } \alpha = 1, \end{cases}$$

com $\text{sign}(u) = \frac{u}{|u|} \forall u \in \mathbb{R}$, $\alpha \in (0, 2]$, $\varrho \in [-1, 1]$, $\sigma > 0$ e $\mu \in \mathbb{R}$.

Observe que a Definição 2.2.1 apresenta quatro parâmetros para descrever a função característica da distribuição estável: α é o parâmetro chamado de índice de estabilidade ou expoente característico, ele define o nível da intensidade local, ϱ é o parâmetro de

assimetria (se $\varrho = 0$ a distribuição é totalmente simétrica em torno de zero, se $\varrho = -1$ é assimétrica negativa, se $\varrho = 1$ é assimétrica positiva), μ é o parâmetro de locação e σ é o parâmetro de escala.

Outra característica das distribuições α -estáveis é sua utilidade para construir classes de modelos de valores extremos com propriedades bem definidas. Para isso, considere a Definição 2.2.2 e observe que esta definição é um caso particular da Definição 2.2.1.

Definição 2.2.2. *Seja $Z \sim S(\alpha, \varrho, \sigma, \mu)$. Quando $\alpha \in (0, 1]$, $\varrho = 0$, $\sigma = 1$ e $\mu = 0$ dizemos que a v.a $Z > 0$ possui distribuição α -estável positiva e denotamos por $Z \sim S(\alpha, 0, 1, 0)$ ou Z_α . Neste caso, sua distribuição fica especificada pela seguinte transformada de Laplace:*

$$\mathbb{E}(e^{-\dot{a}Z}) = e^{-\dot{a}^\alpha}, \quad \dot{a} \geq 0. \quad (2.4)$$

Segundo Kanter et al. (1975) a função densidade que apresenta a transformada de Laplace da Equação (2.4) poder ser dada por:

$$p(z|\alpha) = \frac{1}{\pi} \frac{\alpha}{1-\alpha} \left(\frac{1}{z}\right)^{(1-\alpha)^{-1}} \int_0^\pi a(\varphi) \exp\left\{-\left(\frac{1}{z}\right)^{\alpha/(1-\alpha)} a(\varphi)\right\} d\varphi, \quad (2.5)$$

onde

$$a(\varphi) = \left\{ \frac{\sin(\alpha\varphi)}{\sin(\varphi)} \right\}^{\frac{1}{1-\alpha}} \frac{\sin\{(1-\alpha)\varphi\}}{\sin(\alpha\varphi)}. \quad (2.6)$$

Além disso, Kanter et al. (1975) apresenta uma forma de gerar da distribuição α -estável positiva, vê Colorolário 2.2.1 e o Algoritmo 2.1. Outro resultado importante, apresentado em Williams (1977), está relacionado ao k -ésimo momento da variável $1/Z_\alpha$, ou seja,

$$\mathbb{E}(Z_\alpha^{-k}) = \frac{\Gamma(k\alpha^{-1} + 1)}{\Gamma(k + 1)}.$$

Corolário 2.2.1. *Sejam V e U variáveis aleatórias independentes em que $V \sim \exp(1)$ e $U \sim \text{Uniforme}[0, \pi]$. Então para $\alpha \in (0, 1)$ tem-se que $Z = \left(\frac{a(U)}{V}\right)^{(1-\alpha)/\alpha} \sim \alpha$ -estável positiva. Em que a função $a(\cdot)$ é dada por (2.6).*

```

n ← tamanho da amostra
V ← Gere n amostras da distribuição exponencial (1)
U ← Gere n amostras da distribuição Uniforme (0,π)
Z ←  $\left(\frac{a(U)}{V}\right)^{(1-\alpha)/\alpha}$ 

```

Algoritmo 2.1: Algoritmo para gerar amostras da distribuição α -estável positiva.

2.3 Distribuições Máx e Mín Estáveis

De forma geral, a Teoria de valores extremos está relacionada com a modelagem e análise estatística de eventos extremos, como os valores máximos ou mínimos observados em um conjunto de dados. Dessa forma, a teoria fornece ferramentas e técnicas para estimar o comportamento das caudas das distribuições e fazer previsões sobre eventos raros.

Um conceito importante na TVE é a distribuição máx-Estável (*max-stable*). Nesse caso, de maneira geral, uma distribuição é considerada máx-Estável se o seu valor máximo, devidamente normalizado, converge para uma distribuição específica Marshall and Olkin (1983). Assim sendo, um exemplo de distribuição máx-Estável é a Distribuição Generalizada de Valores Extremos (GEV). A distribuição GEV é frequentemente utilizada para modelar o valor máximo ou mínimo de um conjunto de variáveis aleatórias independentes e identicamente distribuídas, oferecendo ferramentas e propriedades flexíveis para modelar tais dados. Mais especificamente, ela possui três parâmetros que controlam a locação, escala e forma da distribuição, permitindo capturar uma ampla gama de comportamentos de cauda, e consequentemente englobando três tipos de distribuições: Gumbel, Fréchet e Weibull. Outros exemplos de distribuições máx-Estável incluem a distribuição Pareto generalizada e os processos maximamente estáveis, como o processo de Brown-Resnick e o processo de Smith Ribatet (2013). Em relação a distribuições mín-Estáveis (*min-stable*), são menos abordadas na literatura em comparação com as distribuições máx-Estáveis. No entanto, o conceito é semelhante ao das distribuições máx-Estáveis, mas com ênfase nos valores mínimos. Portanto, uma distribuição mín-Estável é uma distribuição de probabilidade em que o mínimo, devidamente normalizado, converge para uma distribuição de excedência específica Marshall and Olkin (1983). Segundo De Haan and Pickands III (1986), um exemplo de distribuição mín-Estável é o inverso de uma distribuição máx-Estável, ou seja, se uma distribuição máx-Estável converge para uma distribuição específica, então tomar a inversa dessa distribuição resultará em uma distribuição mín-Estável que converge para a inversa da distribuição original. Logo, a distribuição de valores extremos generalizada (GEV) e a distribuição Pareto generalizada, que são máx-Estável, podem ser transformadas em uma distribuição mín-Estável considerando seu inverso.

A teoria dos valores extremos teve início com os resultados de Fisher and Tippet

(1928) sobre as distribuições limite de valores extremos sob normalização linear. Para isso, considere $Y_1, Y_2, Y_3, \dots$ uma sequência de variáveis aleatórias independente e identicamente distribuídas com função de distribuição F e função de excedência $\bar{F}$. Seja $M_n = \max(Y_1, \dots, Y_n)$ máximo parcial de $Y_1, Y_2, \dots$. Se existem constantes $a_n > 0$ e $b_n \in \mathbb{R}$ tal que

$$\lim_{n \rightarrow \infty} P(M_n \leq a_n y + b_n) = \lim_{n \rightarrow \infty} F_{M_n}(a_n y + b_n) = \lim_{n \rightarrow \infty} [F(a_n y + b_n)]^n \rightarrow G_{\max}(y),$$

com $G_{\max}(y)$ uma função de distribuição não degenerada, então $G_{\max}(y)$ é uma distribuição máx-Estável sob normalização linear ou simplesmente uma distribuição l-max estável. Nesse caso $G_{\max}(y)$ deve ser uma distribuição de valores extremos generalizada, denotada por $\text{GEV}(\mu, \sigma, \xi)$, com a função de distribuição dada por:

$$G_{\max}(y) = \begin{cases} \exp \left[-\{1 + \xi(y - \mu)/\sigma\}_+^{-1/\xi} \right], & \xi \neq 0 \\ \exp[-\exp\{-(y - \mu)/\sigma\}], & \xi = 0 \end{cases} \quad (2.7)$$

onde $\bar{a}_+ = \max(0, \bar{a})$, $S_G = \{y \in \mathbb{R} : 1 + \xi(y - \mu)/\sigma > 0\}$ é o suporte, $\mu \in \mathbb{R}$ e $\sigma > 0$ são parâmetros de localização e escala respectivamente e $\xi \in \mathbb{R}$ é o parâmetro de forma que determina a taxa de decaimento da cauda: $\xi = 0$, cauda leve (Gumbel); $\xi > 0$, cauda pesada (Fréchet); $\xi < 0$, cauda curta (Weibull). Mais especificamente, sem perda de generalidade, tem-se que a forma padrão :

- Tipo I Gumbel:

$$G_{\max,1}(y) = \exp\{-\exp(-y)\}, \quad y \in \mathbb{R};$$

- Tipo II Fréchet:

$$G_{\max,2}(y) = \begin{cases} 0, & y < 0, \\ \exp\{-y^{-\dot{\alpha}}\}, & y \geq 0 \end{cases} \quad \text{para } \dot{\alpha} > 0;$$

- Tipo III Weibull:

$$G_{\max,3}(y) = \begin{cases} \exp\{-(-y)^{\dot{\alpha}}\}, & y < 0, \\ 1, & y \geq 0 \end{cases} \quad \text{para } \dot{\alpha} > 0.$$

Uma propriedade chave nesta classe de distribuições é o de “máxima-estabilidade”, ou seja, para qualquer número real $t > 0$, existem constantes $A_t > 0$ e $B_t \in \mathbb{R}$ tais que para todo valor de y

$$G_{\max}^t(A_t y + B_t) = G_{\max}(y).$$

Isto implica que as distribuições $G_{\max}$ e $G_{\max}^t$ pertence a mesma família de locação e escala, ou seja, $G_{\max}$ e $G_{\max}^t$ são distribuições GEV com mesmo parâmetro de forma ξ , mas com diferentes parâmetros de locação e escala. Mais especificamente, considere

$Y_1, \dots, Y_N$ como a máxima anual (N anos) de alguma variável de interesse para algum local fixo. Suponha que essas máximas são modeladas utilizando as distribuições GEV $G_{max}(y)$. Como as distribuições GEV são máx-Estáveis, então o máximo de n anos ($n < N$), $\max(Y_1, \dots, Y_n)$, tem distribuição $G_{max}(\frac{y-B_n}{A_n})$ para constantes A_n e B_n que podem ser estimadas a partir dos dados.

Segundo De Haan and Pickands III (1986) e Marshall and Olkin (1983), os resultados apresentados acima também se aplicam ao mínimo. Seja $m_n = \min(Y_1, \dots, Y_n)$ o mínimo parcial de $Y_1, Y_2, Y_3, \dots$. Se existem constantes $a_n > 0$ e $b_n \in \mathbb{R}$ tal que

$$\lim_{n \rightarrow \infty} \bar{F}^n(a_n y + b_n) \rightarrow \bar{G}_{max}(y),$$

com $\bar{G}_{max}(y)$ uma função de excedência não degenerada, então $\bar{G}_{max}(y)$ é uma distribuição de excedência mínimo estável sob normalização linear. Nesse caso, $G_{min}(y)$ é a função de distribuição do mínimo e deve pertencer a classe distribuição de valores extremos generalizada. De fato, se $m_n = \min(Y_1, \dots, Y_n)$ e $M_n = \max(-Y_1, \dots, -Y_n)$ então $m_n = -M_n$, logo

$$\begin{aligned} F_{m_n}(a_n y + b_n) &= P(m_n < a_n y + b_n) \\ &= P(-M_n < a_n y + b_n) \\ &= P(M_n > -(a_n y + b_n)) \\ &= 1 - P(M_n \leq -(a_n y + b_n)) \\ &= 1 - F_{M_n}(-(a_n y + b_n)). \end{aligned} \tag{2.8}$$

Como M_n representa o máximo de variáveis aleatórias $-Y_1, \dots, -Y_n$ então $F_{M_n}(a_n y + b_n)$ converge para uma distribuição GEV $G_{max}(y)$ (Equação 2.7). Portanto, $1 - F_{M_n}(-(a_n y + b_n))$ converge para $1 - G_{max}(-y) = G_{min}(y)$. Nesse caso, a forma padrão é dada por:

- Tipo I Gumbel:

$$G_{min,1}(y) = 1 - \exp\{-\exp(y)\}, \quad y \in \mathbb{R};$$

- Tipo II Fréchet:

$$G_{min,2}(y) = \begin{cases} 1, & y \geq 0, \\ 1 - \exp\{-(y)^{-\alpha}\}, & y < 0 \end{cases} \quad \text{para } \alpha > 0;$$

- Tipo III Weibull:

$$G_{min,3}(y) = \begin{cases} 1 - \exp\{-y^\alpha\}, & y \geq 0, \\ 0, & y < 0 \end{cases} \quad \text{para } \alpha > 0.$$

Pantcheva (2006), Mohan and Ravi (1993) e Pancheva (1994) estenderam o resultado de normalização linear de M_n para uma normalização do tipo potência (Peng et al., 2015) (power normalization). Nesse caso, se existem constantes $a_n > 0$ e $b_n > 0$ tais que

$$\begin{aligned}
\lim_{n \rightarrow \infty} P \left\{ \left| \frac{M_n}{a_n} \right|^{\frac{1}{b_n}} \text{sign}(M_n) \leq y \right\} &= \lim_{n \rightarrow \infty} F_{M_n}(a_n |y|^{b_n} \text{sign}(y)) \\
&= \lim_{n \rightarrow \infty} [F(a_n |y|^{b_n} \text{sign}(y))]^n \\
&\rightarrow G^{pw}(y),
\end{aligned}$$

onde $\text{sign}(y) = -1$ se $y < 0$, $\text{sign}(y) = 0$ se $y = 0$ e $\text{sign}(y) = 1$ se $y > 0$. Neste caso $G^{pw}(y)$ é max estável sob normalização do tipo potência ou p -máx estável. Pantcheva (2006) mostrou que $G^{pw}(y)$ pertence a uma dessas classes de valores extremos para $\beta > 0$:

- Tipo I):

$$G_{1,\beta}^{pw}(y) = \begin{cases} 0, & y \leq 1, \\ \exp \{-(\log y)^{-\beta}\}, & y > 1. \end{cases}$$

- Tipo II) :

$$G_{2,\beta}^{pw}(y) = \begin{cases} 0, & y \leq 0, \\ \exp \{-(\log y)^{\beta}\}, & 0 < y \leq 1, \\ 1, & y > 1. \end{cases}$$

- Tipo III) :

$$G_{3,\beta}^{pw}(y) = \begin{cases} 0, & y \leq -1, \\ \exp \{-(\log(-y))^{-\beta}\}, & -1 < y \leq 0, \\ 1, & y > 0. \end{cases}$$

- Tipo IV) :

$$G_{4,\beta}^{pw}(y) = \begin{cases} \exp \{-(\log(-y))^{\beta}\}, & y \leq -1, \\ 1, & y > -1. \end{cases}$$

- Tipo V) :

$$G_5^{pw}(y) = \begin{cases} 0, & y \leq 0 \\ \exp(-y^{-1}), & y > 0. \end{cases}$$

- Tipo VI) :

$$G_6^{pw}(y) = \begin{cases} \exp(y), & y \leq 0 \\ 1, & y > 0. \end{cases}$$

Portanto duas distribuições K_1 e K_2 são p -máx estáveis se existe $A > 0$ e $B > 0$ tais que $K_1(y) = K_2(A|y|^B \text{sign}(y))$ para todo $y \in \mathbb{R}$.

2.4 Distribuições Condicionais G -Exponencializada

Na literatura, encontramos várias técnicas que permitem expandir famílias de distribuições e construir modelos mais flexíveis, capazes de incorporar propriedades específicas de interesse. Uma estratégia comumente empregada é a exponencialização de uma função de distribuição acumulada G , introduzida por Lehmann (1953).

Encontramos diversos estudos que abordam diferentes famílias de distribuições pertencentes à classe de distribuições exponencializadas. Alguns exemplos incluem: Gupta et al. (1998); Gupta and Kundu (2001) propuseram as distribuições pareto e exponencial exponencializada, respectivamente. Resultados e discussões sobre as distribuições Exponencial, Weibull, Gumbel, Gama e Fréchet exponencializadas são apresentados em Nadarajah and Kotz (2006). A Gama-exponencializada Exponencial foi apresentada em Ristić and Balakrishnan (2012). A distribuição Binomial-Exponencial exponencializada foi proposta em Bakouch et al. (2012). Pinho et al. (2012) introduziram a distribuição Gama-exponencializada Weibull e uma generalização de várias destas distribuições é a família de distribuições $T - X$ exponencializada desenvolvida em Alzaghal et al. (2013). Mais recentemente, Tovar-Falon and Martínez-Flórez (2022) apresentou a distribuição Beta-Skew-Laplace exponencializada que pode ser utilizada para dados trimodais. Propriedades matemáticas e aplicações das distribuições Exponencial Weibull exponencializada e Mukherjee-Islam exponencializada são apresensentadas em Al-Sulami (2020); Rather and Subramanian (2018) respectivamente. Modelos exponencializados utilizando a distribuição kumarawamy são apresentados em Bursa and Gamze (2017); Kumar et al. (2021). Importante observar que essas distribuições foram desenvolvidas e aplicadas em contextos univariados e em amostras de variáveis independente e identicamente distribuídas.

Neste trabalho será adotado a ideia apresentada em Lehmann (1953), mas ao invés de elevar a distribuição acumulada à uma constante positiva, o nosso objetivo é elevar a uma variável aleatória Z , conforme a Definição 2.4.1.

Definição 2.4.1. *Sejam Y e $Z > 0$ duas variáveis aleatórias. Dizemos que condicional a Z , Y possui distribuição G -Exponencializada, se a Função de distribuição de $Y|Z$ for dada por:*

$$F(y; \boldsymbol{\theta}|Z) = P(Y \leq y|Z) = [G(y; \boldsymbol{\theta})]^Z, \quad (2.9)$$

ou, se sua função de excedência é da forma

$$\begin{aligned} \bar{F}(y; \boldsymbol{\theta}|Z) = P(Y > y|Z) &= [1 - G(y; \boldsymbol{\theta})]^Z \\ &= [\bar{G}(y; \boldsymbol{\theta})]^Z, \end{aligned} \quad (2.10)$$

em que G e $\overline{G}$ são chamadas, respectivamente, de função de distribuição e função de excedência de base dependendo do vetor de parâmetros θ .

Note que apartir da Definição 2.4.1, segue que a função de densidade e risco condicional são dadas, respectivamente por

$$f(y; \theta|Z) = Z [\overline{G}(y; \theta)]^{Z-1} g(y; \theta)$$

e

$$h(y; \theta|Z) = Zh(y; \theta).$$

A seguir são apresentados alguns exemplos usando a Definição 2.4.1.

Exemplo 2.4.1. *Seja $G \sim Kw(p, q)$, isto é, a distribuição de base é Kumaraswamy. Então, $Y|Z \sim Kw(p, qZ)$, ou seja,*

$$P(Y \geq y|Z) = (1 - y^p)^{qZ} \quad (2.11)$$

Demonstração.

$$\begin{aligned} P(Y \geq y|Z) &= [\overline{G}(y; \theta)]^Z \\ &= [(1 - y^p)^q]^Z \\ &= (1 - y^p)^{qZ}. \end{aligned}$$

$\therefore Y|Z \sim Kw(p, qZ)$.

Exemplo 2.4.2. *Seja $G \sim Gumbel(\mu, \sigma)$, isto é, a distribuição de base é Gumbel. Então, $Y|Z \sim Gumbel(\mu + \sigma \log Z, \sigma)$, ou seja,*

$$P(Y \leq y|Z) = \exp \left\{ - \exp \left\{ - \frac{y - (\mu + \sigma \log Z)}{\sigma} \right\} \right\}. \quad (2.12)$$

Demonstração.

$$\begin{aligned} P(Y \leq y|Z) &= [G(y; \theta)]^Z \\ &= \left[\exp \left\{ - \exp \left\{ - \left(\frac{y - \mu}{\sigma} \right) \right\} \right\} \right]^Z \\ &= \exp \left\{ - Z \exp \left\{ - \left(\frac{y - \mu}{\sigma} \right) \right\} \right\} \\ &= \exp \left\{ - \exp \{ \log Z \} \times \exp \left\{ - \left(\frac{y - \mu}{\sigma} \right) \right\} \right\} \\ &= \exp \left\{ - \exp \left\{ \frac{\sigma \log Z}{\sigma} - \left(\frac{y - \mu}{\sigma} \right) \right\} \right\} \\ &= \exp \left\{ - \exp \left\{ - \frac{y - (\mu + \sigma \log Z)}{\sigma} \right\} \right\}. \end{aligned}$$

$\therefore Y|Z \sim Gumbel(\mu + \sigma \log Z, \sigma)$.

2.5 Marginais Univariadas de Distribuições Condicionais G -Exponencializadas

Observe que, nos Exemplos 2.4.1 e 2.4.2, Z é uma variável aleatória positiva (vê Definição 2.4.1). Em particular, nesse trabalho, será considerado que $Z \sim \alpha$ -estável positiva (vê Definição 2.2.2), conforme Fougères et al. (2009).

Para isso, note que podemos expressar a Equação (2.9) da seguinte maneira:

$$\begin{aligned}
 F(y; \boldsymbol{\theta}|Z) &= [G(y; \boldsymbol{\theta})]^Z \\
 &= \exp \{ \log [G(y; \boldsymbol{\theta})]^Z \} \\
 &= \exp \{ Z \log [G(y; \boldsymbol{\theta})] \} \\
 F(y; \boldsymbol{\theta}|Z) &= \exp \{ -Z \log [1 / (G(y; \boldsymbol{\theta}))] \}.
 \end{aligned} \tag{2.13}$$

Aplicando a esperança em ambos os lados da Equação (2.13) e utilizando a transformada de Laplace, apresentada na Definição 2.2.2, segue que a distribuição marginal é dada por:

$$\begin{aligned}
 \mathbb{E}(F(y; \boldsymbol{\theta}|Z)) &= \mathbb{E}[G(y; \boldsymbol{\theta})^Z] \\
 &= \mathbb{E}\{\exp[-Z \log(1 / (G(y; \boldsymbol{\theta})))]\} \\
 &= \mathbb{E}\{\exp[-Z(-\log(G(y; \boldsymbol{\theta})))]\} \\
 &= \exp\{-[-\log(G(y; \boldsymbol{\theta}))]^\alpha\} \\
 &= F(y; \boldsymbol{\theta}).
 \end{aligned}$$

Analogamente, usando a Equação (2.10) segue que

$$\bar{F}(y; \boldsymbol{\theta}) = P(Y > y) = \mathbb{E}[\bar{F}(y; |Z)] = \exp\{-[-\log(\bar{G}(y; \boldsymbol{\theta}))]^\alpha\}. \tag{2.14}$$

A seguir vamos apresentar alguns exemplos de marginais univariadas. As notações $Y \sim G(y; \boldsymbol{\theta})$ ou $Y \sim \bar{G}(y; \boldsymbol{\theta})$ será utilizada para especificar que G ou $\bar{G}$ é a distribuição de base para a variável Y .

Exemplo 2.5.1. Se $Z \sim \alpha$ -estável positiva e $Y \sim KW(p, q)$ então $Y|Z \sim Kw(p, qZ)$. Portanto, por (2.14) a distribuição marginal é dada por:

$$\bar{F}(y; \boldsymbol{\theta}, \alpha) = \exp\{-p^\alpha[-\log(1 - y^q)]^\alpha\}, \quad y \in (0, 1).$$

Exemplo 2.5.2. Se a distribuição de base é $Y \sim \text{Gumbel}(\mu, \sigma)$ com $\mu \in \mathbb{R}$ e $\sigma > 0$, isto é

$$G(y; \boldsymbol{\theta}) = \exp\left\{-\exp\left[-\left(\frac{y - \mu}{\sigma}\right)\right]\right\}, \quad y \in \mathbb{R}, \tag{2.15}$$

então,

$$F(y; \boldsymbol{\theta}|Z) = \exp \left\{ -\exp \left[-\left(\frac{y - (\mu + \sigma \log Z)}{\sigma} \right) \right] \right\} = \exp \left\{ -Z \exp \left[-\left(\frac{y - \mu}{\sigma} \right) \right] \right\}$$

isto é, $Y|Z \sim \text{Gumbel}(\mu + \sigma \log Z, \sigma)$. Logo, aplicando (2.14), segue que

$$F(y; \boldsymbol{\theta}) = \exp \left\{ -\exp \left[-\left(\frac{y - \mu}{\sigma/\alpha} \right) \right] \right\}, \quad y \in \mathbb{R}$$

$$\therefore Y \sim \text{Gumbel}(\mu, \frac{\sigma}{\alpha}).$$

Observe que no Exemplo 2.5.1 a distribuição marginal não pertence a mesma família da distribuição condicional, além disso α é um parâmetro de forma. No entanto, no Exemplo 2.5.2 a distribuição marginal pertence a mesma família da distribuição condicional. Neste caso α não representa um parâmetro de forma. Deste modo, dizemos que a distribuição estável positiva simplesmente aumentou ($0 < \alpha \leq 1$) o parâmetro de escala de σ para σ/α .

2.6 Métodos de Estimações

Nesta seção são apresentados alguns métodos que foram utilizados nos artigos apresentados nos Capítulos 4 e 5.

2.6.1 Algoritmo EM

O algoritmo EM (*Expectation-Maximization*) é uma ferramenta computacional usada para calcular de forma iterativa a Estimativa de Máxima Verossimilhança (EMV), sendo especialmente útil em problemas que envolvem dados incompletos ou variáveis não observadas (variáveis latentes). Ele permite que as informações das observações incompletas sejam utilizadas para estimar os parâmetros do modelo de forma mais precisa. O algoritmo foi apresentado por Dempster et al. (1977) e desde então tem sido aplicado em várias áreas. Alguns exemplos podem ser vistos em Liu et al. (2020); Ng and McLachlan (2004); Ding and Song (2016); Sengupta and Cressie (2013).

Basicamente, o algoritmo EM é um processo iterativo que envolve duas etapas principais: Esperança (E) e Maximização (M). Na etapa E, são calculadas as esperanças das variáveis latentes condicionais aos dados observados e aos parâmetros atuais do modelo. Essas esperanças são utilizadas para estimar os valores esperados das variáveis latentes. Na etapa M, os parâmetros do modelo são estimados através da maximização da função de log-verossimilhança completa, considerando as estimativas das variáveis latentes obtidas na etapa E. Essas etapas são repetidas até que a convergência seja alcançada. Para melhor compreensão do algoritmo, considere os seguintes pontos:

- Suponha que $\mathbf{y}$ seja um vetor de dados observados da variável $\mathbf{Y}$ e $\mathbf{z}$ um vetor de dados não observados da variável $\mathbf{Z}$;
- Seja $\mathbf{x} = (\mathbf{y}, \mathbf{z})$ os dados completos;
- Considere $\boldsymbol{\theta}$ um vetor de parâmetros e as variáveis $\mathbf{Y}|\boldsymbol{\theta}$, $\mathbf{Y}|\mathbf{Z}$, $\mathbf{X}|\boldsymbol{\theta}$, $\mathbf{Z}|\mathbf{y}, \boldsymbol{\theta}$ e $\mathbf{Z}|\boldsymbol{\theta}$ com densidades $f_{\mathbf{Y}}(\mathbf{y}|\boldsymbol{\theta})$, $f_{\mathbf{Y}|\mathbf{Z}}(\mathbf{y}|\mathbf{z})$, $f_{\mathbf{X}}(\mathbf{x}|\boldsymbol{\theta})$, $f_{\mathbf{Z}|\mathbf{Y}}(\mathbf{z}|\mathbf{y}, \boldsymbol{\theta})$ e $f_{\mathbf{Z}}(\mathbf{z}|\boldsymbol{\theta})$ respectivamente. Além disso, denote por $f_{\mathbf{x}}(\mathbf{X}|\boldsymbol{\theta}) = L(\boldsymbol{\theta}|\mathbf{X})$ a verossimilhança dos dados completos;
- Observe que a densidade condicional $\mathbf{Z}|\mathbf{Y}$ é dada por

$$f_{\mathbf{Z}|\mathbf{Y}}(\mathbf{z}|\mathbf{y}, \boldsymbol{\theta}) = \frac{f_{\mathbf{X}}(\mathbf{x}|\boldsymbol{\theta})}{f_{\mathbf{Y}}(\mathbf{y}|\boldsymbol{\theta})} \quad (2.16)$$

- O objetivo é maximizar a verossimilhança

$$\begin{aligned} L(\boldsymbol{\theta}|\mathbf{y}) &= \int f_{\mathbf{X}}(\mathbf{x}|\boldsymbol{\theta}) d\mathbf{z} \\ &= \int f_{\mathbf{X}}(\mathbf{y}, \mathbf{z}|\boldsymbol{\theta}) d\mathbf{z} \\ &= \int f_{\mathbf{Y}|\mathbf{Z}}(\mathbf{y}|\mathbf{z}, \boldsymbol{\theta}) f_{\mathbf{Z}}(\mathbf{z}|\boldsymbol{\theta}) d\mathbf{z} \end{aligned} \quad (2.17)$$

- Quando integral da Equação (2.17) não é analiticamente tratável, utilizamos o algoritmo EM para maximizar $L(\boldsymbol{\theta}|\mathbf{y})$ iterativamente conforme o Algoritmo 2.2

2.6.2 Algoritmo EM Condicional

Segundo McLachlan and Krishnan (2008), o algoritmo *Expectation-Conditional Maximization* (ECM), proposto por Meng and Rubin (1993), é uma extensão do algoritmo EM em situações onde a maximização na etapa M é relativamente simples quando condicionada a alguma função dos parâmetros. Nesse caso, o algoritmo ECM substitui a etapa M do algoritmo EM por uma série de etapas de maximização condicional (CM) computacionalmente mais simples. Embora convirja mais lentamente em termos de número de iterações, pode ser mais rápido em termos do tempo total de computação. Importante ressaltar que o algoritmo ECM preserva as propriedades de convergência do algoritmo EM. Algumas aplicações podem ser vistas em Zeephongsekul et al. (2016); Matsumura et al. (2021). O Algoritmo 2.3 apresenta os passos do ECM para um caso particular.

1. Obtenha:

$$Q(\boldsymbol{\theta} | \boldsymbol{\theta}) \leftarrow \mathbb{E}_{\mathbf{Z}|\mathbf{Y}} [\log L(\boldsymbol{\theta} | \mathbf{X}) | \mathbf{y}, \boldsymbol{\theta}]$$

2. Seja $\boldsymbol{\theta}^{(k)}$ a estimativa na iteração k , para $k = 0, 1, \dots$

3. Inicie o algoritmo na iteração $\boldsymbol{\theta}^{(0)}$

4. Passo E: Atualize

$$Q(\boldsymbol{\theta} | \boldsymbol{\theta}^{(k)}) \leftarrow \mathbb{E}_{\mathbf{Z}|\mathbf{Y}} [\log L(\boldsymbol{\theta} | \mathbf{X}) | \mathbf{y}, \boldsymbol{\theta}^{(k)}]$$

5. Passo M: Obtenha $\boldsymbol{\theta}^{(k+1)}$ maximizando $Q(\boldsymbol{\theta} | \boldsymbol{\theta}^{(k)})$ com respeito $\boldsymbol{\theta}$ e faça $k = k + 1$;

6. Repita os passos E e M até obter convergência para sequência de estimativas $\boldsymbol{\theta}^k$, segundo algum critério de parada, em particular $\|\hat{\boldsymbol{\theta}}^{(k+1)} - \hat{\boldsymbol{\theta}}^{(k)}\| < \epsilon$ em que ϵ é um valor, maior do que zero, especificado para o erro de aproximação e $\|a\|$ é a norma do vetor a .

Algoritmo 2.2: Algoritmo EM.

2.6.3 Algoritmo Monte Carlo EM

Observe que as operações de integração e otimização constituem aspectos cruciais do algoritmo EM, ou seja, o passo E é o cálculo de esperança da log verossimilhança (integração) e o passo M é um cálculo de otimização. Diante disso, a etapa E (cálculo da integral) pode ser complexa e não permitir uma solução analítica direta, ou seja, a função $Q(\boldsymbol{\theta} | \boldsymbol{\theta}^{(k)})$ não pode ser calculada explicitamente em certas aplicações do algoritmo EM. Uma abordagem para contornar essa dificuldade é utilizar o método de Monte Carlo (MC), proposto em Metropolis and Ulam (1949), para estimar a função $Q(\boldsymbol{\theta} | \boldsymbol{\theta}^{(k)})$. Essa variação do algoritmo EM é denominada Monte Carlo EM (MCEM) e foi apresentada por Wei and Tanner (1990). Alguns exemplos utilizando essa abordagem podem ser vistos em Mota et al. (2018); Da Silva and Montenegro (2021). O algoritmo apresentado em 2.4 mostra os passos MCEM.

1. Considere $\boldsymbol{\theta} = (\boldsymbol{\theta}_1, \boldsymbol{\theta}_2)$ em que $\boldsymbol{\theta}_1$ e $\boldsymbol{\theta}_2$ é um subvetor dos parâmetros.

2. Obtenha:

$$Q(\boldsymbol{\theta} | \boldsymbol{\theta}) \leftarrow \mathbb{E}_{\mathbf{Z}|\mathbf{Y}} [\log L(\boldsymbol{\theta} | \mathbf{X}) | \mathbf{y}, \boldsymbol{\theta}]$$

3. Seja $\boldsymbol{\theta}^{(k)}$ a estimativa na iteração k , para $k = 0, 1, \dots$

4. Inicie o algoritmo na iteração $\boldsymbol{\theta}^{(0)}$

5. Passo E: Atualize

$$Q(\boldsymbol{\theta} | \boldsymbol{\theta}^{(k)}) \leftarrow \mathbb{E}_{\mathbf{Z}|\mathbf{Y}} [\log L(\boldsymbol{\theta} | \mathbf{X}) | \mathbf{y}, \boldsymbol{\theta}^{(k)}]$$

6. Passo CM: Obtenha $\boldsymbol{\theta}^{(k+1)}$ maximizando $Q(\boldsymbol{\theta} | \boldsymbol{\theta}^{(k)})$ condicionalmente, ou seja,

$$\boldsymbol{\theta}_1^{(k+1)} = \operatorname{argmax}_{\boldsymbol{\theta}_1} Q(\boldsymbol{\theta} | \boldsymbol{\theta}_2^{(k)})$$

e

$$\boldsymbol{\theta}_2^{(k+1)} = \operatorname{argmax}_{\boldsymbol{\theta}_2} Q(\boldsymbol{\theta} | \boldsymbol{\theta}_1^{(k+1)})$$

faça $k = k + 1$;

7. Repita os passos E e M até obter convergência para sequência de estimativas $\boldsymbol{\theta}^k$, segundo algum critério de parada, em particular $\|\hat{\boldsymbol{\theta}}^{(k+1)} - \hat{\boldsymbol{\theta}}^{(k)}\| < \epsilon$ em que ϵ é um valor, maior do que zero, especificado para o erro de aproximação e $\|a\|$ é a norma do vetor a .

Algoritmo 2.3: Algoritmo ECM.

2.6.4 Matriz de Informação Observada

Segundo McLachlan and Krishnan (2008) o algoritmo EM não fornece automaticamente uma matriz de covariância dos estimadores de máxima verossimilhança. Uma alternativa para estimar a matriz de covariância é calcular a informação observada $-l''(\hat{\boldsymbol{\theta}}|\mathbf{y})$, em que l'' é a matriz de segunda derivada de $\log(L(\boldsymbol{\theta}|\mathbf{y}))$.

Para isso, note que a partir da Equação (2.16)

$$\log f_{\mathbf{Y}}(\mathbf{y} | \boldsymbol{\theta}) = \log f_{\mathbf{X}}(\mathbf{x} | \boldsymbol{\theta}) - \log f_{\mathbf{Z}|\mathbf{Y}}(\mathbf{z} | \mathbf{y}, \boldsymbol{\theta}). \quad (2.18)$$

Denote por $l(\boldsymbol{\theta}|\mathbf{Y}) = \log(f_{\mathbf{Y}}(\mathbf{Y} | \boldsymbol{\theta}))$ e $l(\boldsymbol{\theta}|\mathbf{X}) = \log(f_{\mathbf{X}}(\mathbf{x} | \boldsymbol{\theta}))$. Diferenciando em relação a $\boldsymbol{\theta}$ duas vezes o negativo da Equação (2.16), temos que

$$\mathbf{I}(\boldsymbol{\theta}; \mathbf{y}) = \mathbf{I}_c(\boldsymbol{\theta}; \mathbf{x}) + \partial^2 \log f(\mathbf{z} | \mathbf{y}; \boldsymbol{\theta}) / \partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^T, \quad (2.19)$$

1. Seja $\boldsymbol{\theta}^{(k)}$ a estimativa na iteração k , para $k = 0, 1, \dots$
2. Inicie o algoritmo na iteração $\boldsymbol{\theta}^{(0)}$
3. Passo E:
 - Na (k) -ésima iteração, gere uma amostra $\mathbf{z}_1^k, \dots, \mathbf{z}_M^k$ da condicional $f_{\mathbf{Z}|\mathbf{Y}}(\mathbf{z}|\mathbf{y}, \boldsymbol{\theta}^{k-1})$. Cada $\mathbf{z}_m^k$ é um vetor de dados faltantes. Então $\mathbf{X}_m^k = (\mathbf{y}, \mathbf{z}_m^k)$.
 - Aproxime $Q(\boldsymbol{\theta} | \boldsymbol{\theta}^{(k)})$ via Monte Carlo, ou seja,

$$Q(\boldsymbol{\theta} | \boldsymbol{\theta}^{(k)}) = \frac{1}{M} \sum_{m=1}^M L(\boldsymbol{\theta} | \mathbf{X}_m^k)$$

4. Passo M: Obtenha $\boldsymbol{\theta}^{(k+1)}$ maximizando $Q(\boldsymbol{\theta} | \boldsymbol{\theta}^{(k)})$ com respeito $\boldsymbol{\theta}$ e faça $k = k + 1$;
5. Repita os passos E e M até obter convergência para sequência de estimativas $\boldsymbol{\theta}^k$, segundo algum critério de parada, em particular $\|\hat{\boldsymbol{\theta}}^{(k+1)} - \hat{\boldsymbol{\theta}}^{(k)}\| < \epsilon$ em que ϵ é um valor, maior do que zero, especificado para o erro de aproximação e $\|a\|$ é a norma do vetor a .

Algoritmo 2.4: Algoritmo MCEM.

onde

$$\mathbf{I}(\boldsymbol{\theta}; \mathbf{y}) = -\partial^2 l(\boldsymbol{\theta} | \mathbf{Y}) / \partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^T$$

e

$$\mathbf{I}_c(\boldsymbol{\theta}; \mathbf{x}) = -\partial^2 l(\boldsymbol{\theta} | \mathbf{X}) / \partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^T.$$

Aplicando a esperança, em relação a condicional $\mathbf{Z}|\mathbf{Y}$, em ambos os lados da Equação 2.19 segue que

$$\mathbf{I}(\boldsymbol{\theta}; \mathbf{y}) = \mathcal{I}_c(\boldsymbol{\theta}; \mathbf{y}) - \mathcal{I}_m(\boldsymbol{\theta}; \mathbf{y}), \quad (2.20)$$

onde

$$\mathcal{I}_c(\boldsymbol{\theta}; \mathbf{y}) = \mathbb{E} \{ \mathbf{I}_c(\boldsymbol{\theta}; \mathbf{X}) | \mathbf{y} \}$$

é a esperança condicional da matrix de informação dos dados completos, e

$$\mathcal{I}_m(\boldsymbol{\theta}; \mathbf{x}) = -\mathbb{E} \{ \partial^2 \log f(\mathbf{z} | \mathbf{y}; \boldsymbol{\theta}) / \partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^T | \mathbf{y} \}$$

é a matriz de informação esperada para $\boldsymbol{\theta}$ baseada nos dados não observados $\mathbf{z}$. Portanto, a informação observada é igual à informação completa menos a informação faltante. Esse resultado é denominado princípio da falta de informação (Orchard and Woodbury, 1972).

Seja $S_{\mathbf{X}}(\boldsymbol{\theta}) = -\partial l(\boldsymbol{\theta}|\mathbf{X})/\partial \boldsymbol{\theta}$ e $\hat{\boldsymbol{\theta}}$ a estimativa de $\boldsymbol{\theta}$. Segundo Louis (1982), $\mathcal{I}_m(\boldsymbol{\theta}; \mathbf{y})$ pode ser expressa como

$$\begin{aligned} \mathcal{I}_m(\boldsymbol{\theta}; \mathbf{y}) &= \text{cov}[S_{\mathbf{X}}(\boldsymbol{\theta})] \\ &= \mathbb{E} \left[S_{\mathbf{X}}(\boldsymbol{\theta}) S_{\mathbf{X}}(\boldsymbol{\theta})^T \mid \mathbf{y}, \hat{\boldsymbol{\theta}} \right] \\ &\quad - \mathbb{E} \left[S_{\mathbf{X}}(\boldsymbol{\theta}) \mid \mathbf{y}, \hat{\boldsymbol{\theta}} \right] \mathbb{E}^T \left[S_{\mathbf{X}}(\boldsymbol{\theta}) \mid \mathbf{y}, \hat{\boldsymbol{\theta}} \right]. \end{aligned} \quad (2.21)$$

Desde que $\mathbb{E} \left[S_{\mathbf{X}}(\boldsymbol{\theta}) \mid \mathbf{y}, \hat{\boldsymbol{\theta}} \right] = 0$ e substituindo a Equação 2.21 na Equação 2.20

$$\mathbf{I}(\hat{\boldsymbol{\theta}}; \mathbf{y}) = \mathcal{I}_c(\hat{\boldsymbol{\theta}}; \mathbf{y}) - \mathbb{E} \left[S_{\mathbf{X}}(\hat{\boldsymbol{\theta}}) S_{\mathbf{X}}(\hat{\boldsymbol{\theta}})^T \mid \mathbf{y}, \hat{\boldsymbol{\theta}} \right]. \quad (2.22)$$

Uma forma para desenvolver a matriz de Informação observada para o algoritmo MCEM é apresentado em McLachlan and Krishnan (2008).

CAPÍTULO 3

Classe de Distribuições Multivariadas para Dados Extremos Dependentes, Limitados e não Estacionários

As distribuições apresentadas na Seção 2.4 foram desenvolvidas e aplicadas em um contexto univariado, considerando amostras de variáveis independentes e identicamente distribuídas. No entanto, existem conjuntos de dados multivariados que apresentam algum tipo de dependência entre eventos, indivíduos ou grupos de indivíduos. Portanto, é necessário utilizar distribuições multivariadas que acomodem essa dependência conjunta a fim de obter inferências mais robustas. Um exemplo disso é a modelagem de valores extremos ambientais ou meteorológicos, pois dados como precipitação, temperatura ou poluição do ar são medidos em locais específicos, como estações de monitoramento, e esses valores são multivariados e podem apresentar correlação temporal, espacial ou espaço-temporal. Exemplos de distribuições multivariadas no contexto espacial podem ser vistas em Fougères et al. (2009) e Mota et al. (2018). Neste caso, ambos autores utilizaram a Definição 2.2.2 para a construção das classes de distribuições Gumbel e log-Weibull, respectivamente.

Neste capítulo será apresentado, utilizando a Definição 2.2.2, uma classe de distribuições multivariadas para valores extremos, correlacionados, limitados inferiormente e superiormente e não estacionários. Em particular, serão apresentadas propriedades para observações no intervalo $(0, 1)$. Aplicações dessa classe são apresentadas nos Capítulos 4 e 5.

3.1 Construção da Classe

Sejam $\mathcal{A}$ e $\mathcal{B}$ conjuntos de índices discretos finitos, $\{\omega_{a,b} : a \in \mathcal{A} \text{ e } b \in \mathcal{B}\}$ constantes não-negativas tais que $\sum_{b \in \mathcal{B}} \omega_{a,b} = 1$ e $\{Z_b : b \in \mathcal{B}\}$ um conjunto de variáveis

independentes α_b -estáveis positivas. Assuma que $\{Y_a|H_a : a \in \mathcal{A}\}$ é um conjunto de variáveis condicionalmente independentes com suporte no intervalo $(0, 1)$, parâmetro de forma $p_a > 0$, H_a um processo não observado, V uma função de distribuição e $u_a \sim U(0, 1)$. Assim sendo, uma possível representação do processo $Y_a|H_a$ é

$$Y_a = [1 - V(u_a^{1/H_a})]^{1/p_a}. \quad (3.1)$$

Portanto, a função de distribuição de $Y_a|H_a$ é dada por:

$$\begin{aligned} P(Y_a \leq y_a|H_a) &= P(1 - V(u_a^{1/H_a}) \leq y_a^p) \\ &= P(V(u_a^{1/H_a}) \geq (1 - y_a^p)) \\ &= P(u_a \geq [V^{-1}(1 - y_a^p)]^{H_a}) \\ &= 1 - [V^{-1}(1 - y_a^p)]^{H_a}. \end{aligned} \quad (3.2)$$

Logo a função de excedência é dada por

$$\begin{aligned} \bar{F}(y_a; \boldsymbol{\theta}|H_a) &= P(Y_a \geq y_a|H_a) \\ &= \exp \left\{ -H_a \log [V^{-1}(1 - y_a^p)] \right\}. \end{aligned} \quad (3.3)$$

As escolhas de V e H_a determinam a classe de modelos multivariados induzida pela lei de potência u_a^{1/H_a} . Considere $H_a = q_a \sum_{b \in \mathcal{B}} \omega_{a,b} Z_b$, para o parâmetro de forma $q_a > 0$, e $\mathcal{H}$ como um vetor cujas coordenadas são marginais compostas por misturas de distribuições α_b -estáveis positivas, logo:

$$\begin{aligned} \bar{F}(\mathbf{y}; \boldsymbol{\theta}|\mathcal{H}) &= \prod_{a \in \mathcal{A}} \bar{F}(y_a; \boldsymbol{\theta}|H_a) \\ &= \prod_{a \in \mathcal{A}} \exp \left\{ -H_a \left\{ -\log [V^{-1}(1 - y_a^p)] \right\} \right\} \\ &= \prod_{b \in \mathcal{B}} \exp \left\{ -Z_b \sum_{a \in \mathcal{A}} \omega_{a,b} \left\{ -\log [V^{-1}(1 - y_a^p)]^{q_a} \right\} \right\}. \end{aligned}$$

Aplicando a esperança em ambos os lados da igualdade acima, temos que marginalmente ao vetor $\mathcal{H}$

$$\begin{aligned} \bar{F}(\mathbf{y}; \boldsymbol{\theta}, \boldsymbol{\alpha}) &= \mathbb{E}(\bar{F}(\mathbf{y}; \boldsymbol{\theta}|\mathcal{H})) \\ &= \mathbb{E} \left(\prod_{b \in \mathcal{B}} \exp \left\{ -Z_b \sum_{a \in \mathcal{A}} \omega_{a,b} \left\{ -\log [V^{-1}(1 - y_a^p)]^{q_a} \right\} \right\} \right). \end{aligned}$$

Como $Z_b, b \in \mathcal{B}$ são variáveis aleatórias independentes α_q -estáveis positivas, pela Definição 2.2.2 obtemos o seguinte resultado:

$$\bar{F}(\mathbf{y}; \boldsymbol{\theta}, \boldsymbol{\alpha}) = \prod_{b \in \mathcal{B}} \exp \left\{ - \left[- \sum_{a \in \mathcal{A}} \omega_{a,b} \log [V^{-1} (1 - y_a^{p_a})]^{q_a} \right]^{\alpha_b} \right\} \quad (3.4)$$

Se V é um função de distribuição de uma variável $U(0, 1)$, então o modelo acima é equivalente ao modelo descrito a seguir.

3.1.1 Classe Kumaraswamy Exponencializada

Defina $H_a = \sum_{b \in \mathcal{B}} \omega_{a,b} Z_b$, isto é, para cada $a \in \mathcal{A}$, H_a é uma mistura de distribuições α_b -estáveis positivas. Assuma que $\{Y_a | H_a : a \in \mathcal{A}\}$ é um conjunto de variáveis condicionalmente independentes com distribuições $Kw(p_a, q_a H_a, 0, 1)$, ou seja, possui função de excedência

$$\bar{F}(y_a; \boldsymbol{\theta} | H_a) = \exp \{ -H_a \{ -\log (1 - y_a^p)^q \} \}. \quad (3.5)$$

Considere $\mathcal{H}$ como um vetor cujas coordenadas são marginais compostas por misturas de distribuições α_b -estáveis positivas, então:

$$\begin{aligned} \bar{F}(\mathbf{y}; \boldsymbol{\theta} | \mathcal{H}) &= \prod_{a \in \mathcal{A}} \bar{F}(y_a; \boldsymbol{\theta} | H_a) \\ &= \prod_{a \in \mathcal{A}} \exp \{ -H_a \{ -\log (1 - y_a^p)^q \} \}. \\ &= \prod_{b \in \mathcal{B}} \exp \left\{ -Z_b \sum_{a \in \mathcal{A}} \omega_{a,b} \{ -\log (1 - y_a^p)^q \} \right\}. \end{aligned}$$

Aplicando a esperança em ambos os lados da igualdade acima, temos que marginalmente ao vetor $\mathcal{H}$

$$\begin{aligned} \bar{F}(\mathbf{y}; \boldsymbol{\theta}, \boldsymbol{\alpha}) &= \mathbb{E}(\bar{F}(\mathbf{y}; \boldsymbol{\theta} | \mathcal{H})) \\ &= \mathbb{E} \left(\prod_{b \in \mathcal{B}} \exp \left\{ -Z_b \sum_{a \in \mathcal{A}} \omega_{a,b} \{ -\log (1 - y_a^p)^q \} \right\} \right). \end{aligned}$$

Como $Z_b, b \in \mathcal{B}$ são variáveis aleatórias independentes α_q -estáveis positivas, pela Definição 2.2.2 obtemos o seguinte resultado:

$$\bar{F}(\mathbf{y}; \boldsymbol{\theta}, \boldsymbol{\alpha}) = \prod_{b \in \mathcal{B}} \exp \left\{ - \left[- \sum_{a \in \mathcal{A}} \omega_{a,b} \log (1 - y_a^p)^q \right]^{\alpha_b} \right\}. \quad (3.6)$$

O modelo obtido é chamado de Kumaraswamy Exponencializada Multivariada e será denotada por $\mathbf{Y} \sim KEM(\boldsymbol{\alpha}, \boldsymbol{\theta})$. A Figura 3.1 mostra a função de excedência e a função densidade do modelo bivariado com $b = 1$, $p_1 = p_2 = 15$ e $q_1 = q_2 = 100$ e $\alpha \in \{0.25, 0.50, 0.75\}$. Observe que o formato da curva da densidade conjunta se torna menos elíptico à medida que $\alpha \rightarrow 1$. Esse comportamento indica que há uma diminuição na relação de dependência entre as variáveis.

3.2 Propriedades

3.2.1 Função Densidade Conjunta

Considere as seguintes variáveis para o cálculo da função densidade conjunta:

$$\nu_{a,b} = -\omega_{a,b} \log(1 - y_a^p)^q \quad (3.7)$$

e

$$\zeta_b = \left(\sum_{a \in \mathcal{A}} \nu_{a,b} \right)^{\alpha_b}. \quad (3.8)$$

Assim, a função de excedência conjunta, definida por (3.6), pode ser escrita da seguinte maneira:

$$\bar{F}(\mathbf{y}; \boldsymbol{\theta}, \boldsymbol{\alpha}) = \exp \left(- \sum_{b \in \mathcal{B}} \zeta_b \right), \quad \zeta_b > 0. \quad (3.9)$$

A partir do conhecimento da função de excedência, é possível determinar a função de densidade de probabilidade multivariada de um vetor aleatório $\mathbf{Y} = (Y_1, Y_2, \dots, Y_n)$ por meio da seguinte fórmula:

$$f(y_1, \dots, y_n; \boldsymbol{\theta}, \boldsymbol{\alpha}) = \frac{(-1)^n \partial^{(n)} \bar{F}(\mathbf{y}, \boldsymbol{\theta}, \boldsymbol{\alpha})}{\partial y_1 \partial y_2 \dots \partial y_n}. \quad (3.10)$$

Proposição 3.2.1. A função densidade conjunta de (3.9) é expressa por:

$$f(\mathbf{y}; \boldsymbol{\theta}, \boldsymbol{\alpha}) = \bar{F}(\mathbf{y}, \boldsymbol{\theta}, \boldsymbol{\alpha}) \phi_{1,2,\dots,A}(\mathbf{y}; \boldsymbol{\theta}, \boldsymbol{\alpha}), \quad (3.11)$$

em que A é a cardinalidade do conjunto $\mathcal{A}$,

$$\phi_{1,2,\dots,A}(\mathbf{y}, \boldsymbol{\theta}, \boldsymbol{\alpha}) = \phi_A(\mathbf{y}, \boldsymbol{\theta}, \boldsymbol{\alpha}) \cdot \phi_{1,2,\dots,(A-1)}(\mathbf{y}, \boldsymbol{\theta}, \boldsymbol{\alpha}) - \frac{\partial}{\partial y_A} \phi_{1,2,\dots,(A-1)}(\mathbf{y}, \boldsymbol{\theta}, \boldsymbol{\alpha})$$

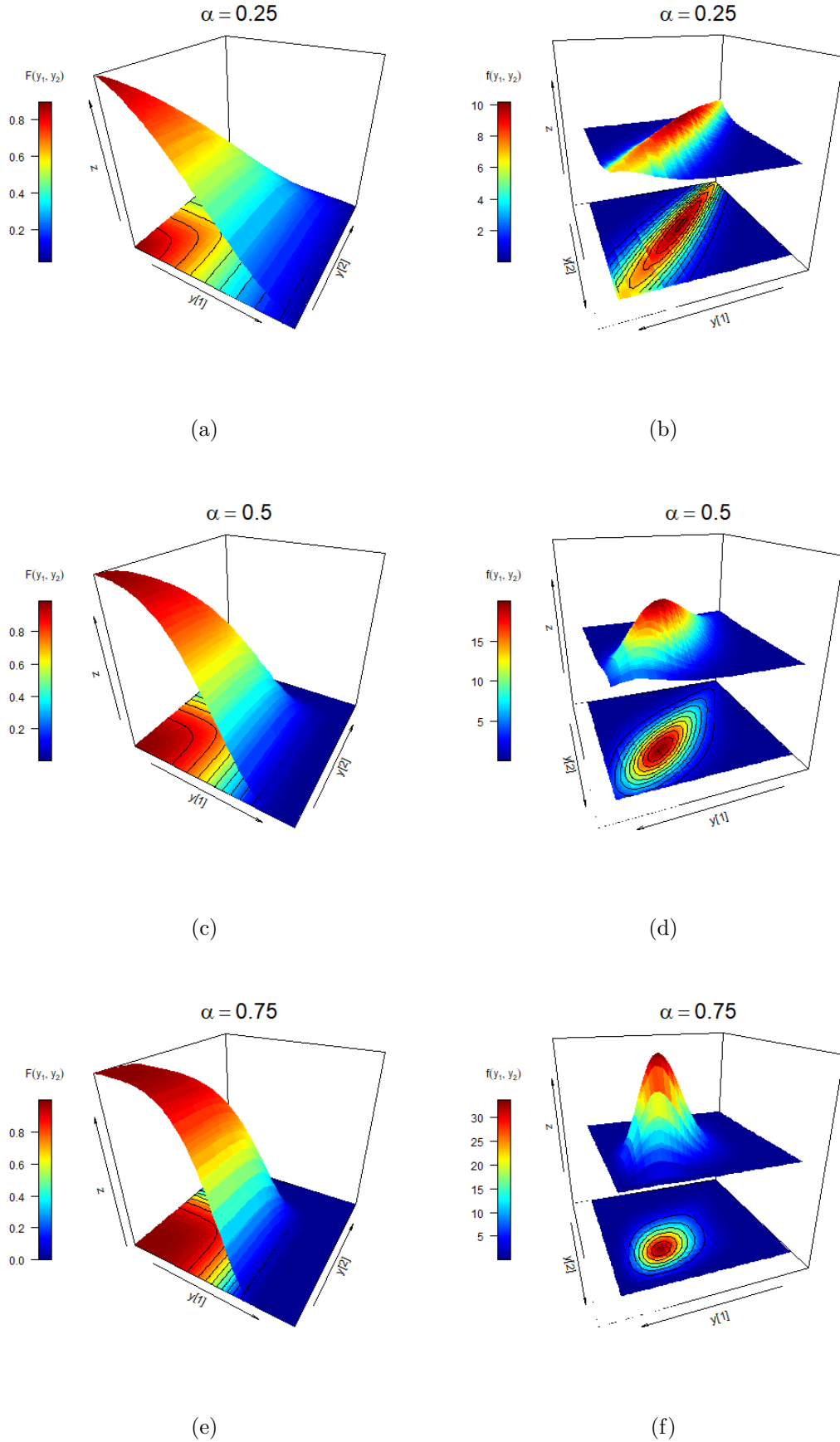


Figura 3.1: Função de excedência (lado esquerdo) e função densidade (lado direito) do modelo Kumaraswamy Exponencializada Multivariado para diferentes valores de α .

e

$$\phi_a(\mathbf{y}; \boldsymbol{\theta}; \boldsymbol{\alpha}) = h(y_a; \boldsymbol{\theta}) \cdot \Psi(\mathbf{y}; \boldsymbol{\theta}, \boldsymbol{\alpha}) \quad \text{para } a = 1, 2, 3, \dots, A$$

com

$$\Psi(\mathbf{y}; \boldsymbol{\theta}, \boldsymbol{\alpha}) = \sum_{b \in \mathcal{B}} \omega_{a,b} \alpha_b z_b^{1 - \frac{1}{\alpha_b}}$$

e

$$h(y_a; \boldsymbol{\theta}) = \frac{g(y_a; \boldsymbol{\theta})}{\bar{G}(y_a; \boldsymbol{\theta})} = \frac{pqy_a^{p-1}}{1 - y_a^p} \quad (3.12)$$

é a função de risco associada a distribuição $Kw(p_a, q_a)$.

A prova dessa propriedade é similar a Proposição 1 de Fougères et al. (2013).

Demonstração. Vamos usar o princípio da indução finita sobre $A \geq 2$. Com efeito, para $A = 2$, temos que por (3.10)

$$f(y_1, y_2; \boldsymbol{\theta}, \boldsymbol{\alpha}) = (-1)^2 \frac{\partial^{(2)} \bar{F}(y_1, y_2; \boldsymbol{\theta}, \boldsymbol{\alpha})}{\partial y_1 \partial y_2} = \frac{\partial}{\partial y_2} \left(\frac{\partial \bar{F}(y_1, y_2; \boldsymbol{\theta}, \boldsymbol{\alpha})}{\partial y_1} \right).$$

Note que

$$\frac{\partial \bar{F}(\mathbf{y}; \boldsymbol{\theta}, \boldsymbol{\alpha})}{\partial y_a} = -\bar{F}(\mathbf{y}; \boldsymbol{\theta}, \boldsymbol{\alpha}) \phi_a(\mathbf{y}; \boldsymbol{\theta}; \boldsymbol{\alpha}) \quad \text{para } a = 1, 2, 3, \dots, A. \quad (3.13)$$

Assim,

$$\begin{aligned} f(y_1, y_2; \boldsymbol{\theta}, \boldsymbol{\alpha}) &= \frac{\partial}{\partial y_2} \left(-\bar{F}(y_1, y_2; \boldsymbol{\theta}, \boldsymbol{\alpha}) \phi_1(y_1, y_2; \boldsymbol{\theta}, \boldsymbol{\alpha}) \right) \\ &= \left[-\frac{\partial}{\partial y_2} \bar{F}(y_1, y_2; \boldsymbol{\theta}, \boldsymbol{\alpha}) \right] \phi_1(y_1, y_2; \boldsymbol{\theta}, \boldsymbol{\alpha}) \\ &\quad - \bar{F}(y_1, y_2; \boldsymbol{\theta}, \boldsymbol{\alpha}) \left[\frac{\partial}{\partial y_2} \phi_1(y_1, y_2; \boldsymbol{\theta}, \boldsymbol{\alpha}) \right] \\ &= \bar{F}(y_1, y_2; \boldsymbol{\theta}, \boldsymbol{\alpha}) \\ &\quad \times \left[\phi_2(y_1, y_2; \boldsymbol{\theta}, \boldsymbol{\alpha}) \phi_1(y_1, y_2; \boldsymbol{\theta}, \boldsymbol{\alpha}) - \frac{\partial}{\partial y_2} \phi_1(y_1, y_2; \boldsymbol{\theta}, \boldsymbol{\alpha}) \right]. \end{aligned}$$

Logo,

$$f(y_1, y_2; \boldsymbol{\theta}, \boldsymbol{\alpha}) = \bar{F}(y_1, y_2; \boldsymbol{\theta}, \boldsymbol{\alpha}) \phi_{1,2}(y_1, y_2; \boldsymbol{\theta}, \boldsymbol{\alpha})$$

e, portanto, o resultado é válido.

Suponha que o resultado é válido para $A = k$ com $k \in \mathbb{N}$, isto é,

$$\begin{aligned} f(y_1, y_2, \dots, y_k; \boldsymbol{\theta}, \boldsymbol{\alpha}) &= \frac{(-1)^k \partial^{(k)} \bar{F}(y_1, y_2, \dots, y_k; \boldsymbol{\theta}, \boldsymbol{\alpha})}{\partial y_1 \partial y_2 \dots \partial y_k} \\ &= \bar{F}(y_1, y_2, \dots, y_k; \boldsymbol{\theta}, \boldsymbol{\alpha}) \cdot \phi_{1,2,\dots,k}(y_1, y_2, \dots, y_k; \boldsymbol{\theta}, \boldsymbol{\alpha}). \end{aligned}$$

Devemos provar que vale para $A = k + 1$, $k \in \mathbb{N}$. De fato, por diferenciação recursiva temos que

$$\begin{aligned} f(y_1, y_2, \dots, y_{k+1}; \boldsymbol{\theta}, \boldsymbol{\alpha}) &= \frac{-\partial}{\partial y_{k+1}} \left(\frac{(-1)^k \partial^{(k)} \bar{F}(y_1, y_2, \dots, y_{k+1}; \boldsymbol{\theta}, \boldsymbol{\alpha})}{\partial y_1 \partial y_2 \dots \partial y_k} \right) \\ &= \frac{-\partial}{\partial y_{k+1}} \left(\bar{F}(y_1, y_2, \dots, y_{k+1}; \boldsymbol{\theta}, \boldsymbol{\alpha}) \cdot \phi_{1,2,\dots,k}(y_1, y_2, \dots, y_{k+1}; \boldsymbol{\theta}, \boldsymbol{\alpha}) \right) \end{aligned}$$

de modo que

$$f(y_1, y_2, \dots, y_{k+1}; \boldsymbol{\theta}, \boldsymbol{\alpha}) = \bar{F}(y_1, y_2, \dots, y_{k+1}; \boldsymbol{\theta}, \boldsymbol{\alpha}) \phi_{1,2,\dots,k+1}(y_1, y_2, \dots, y_{k+1}; \boldsymbol{\theta}, \boldsymbol{\alpha})$$

Portanto, segue o resultado para todo $A \in \mathbb{N}$. $\square$

3.2.2 Distribuições do Mínimo e Marginal

Considere $\mathbf{Y} \sim \text{KEM}(\boldsymbol{\alpha}, \boldsymbol{\theta})$. As distribuições do mínimo e marginais sobre qualquer subconjunto $\mathcal{A}_0 \subset \mathcal{A}$, tal que $\mathcal{A}_0 \neq \emptyset$, pertencem à mesma classe Kumaraswamy Exponencializada Multivariada.

Demonstração. Com efeito, observe que

$$\bar{F}(y; \boldsymbol{\theta}, \boldsymbol{\alpha}) = \mathbb{E} \left(P \left(\min_{a \in \mathcal{A}_0} Y_a > y \mid \mathcal{H} \right) \right).$$

Como $\{Y_a \mid H_a : a \in \mathcal{A}_0\}$ são variáveis condicionalmente independentes com distribuições $Kw \sim (p, qH_a)$, então por (3.5)

$$\bar{F}(y; \boldsymbol{\theta}, \boldsymbol{\alpha}) = \mathbb{E} \left(\prod_{a \in \mathcal{A}_0} \exp \{ -H_a (-\log(1 - y_a^p))^q \} \right). \quad (3.14)$$

Substituindo-se $H_a = \sum_{b \in \mathcal{B}} \omega_{a,b} \zeta_b$ e em seguida aplicando a Definição 2.2.2, temos

$$\bar{F}(y; \boldsymbol{\theta}, \boldsymbol{\alpha}) = \prod_{b \in \mathcal{B}} \exp \left\{ - \left[- \sum_{a \in \mathcal{A}_0} \omega_{a,b} \log(1 - y_a^p)^q \right]^{\alpha_b} \right\},$$

e pertence à mesma família $KEM(\boldsymbol{\alpha}, \boldsymbol{\theta})$.

Em particular, se $\mathcal{A}_0$ possui apenas um ponto então a distribuição marginal de Y_a é

$$\bar{F}_{Y_a}(y_a; \boldsymbol{\theta}, \boldsymbol{\alpha}) = \prod_{b \in \mathcal{B}} \exp \{ - [-\omega_{a,b} \log(1 - y_a^p)^q]^{\alpha_b} \},$$

também pertence à mesma família. $\square$

3.2.3 Função Densidade Por Grupos

Suponha que a família de subconjuntos $\mathcal{A}_b$ $b \in \mathcal{B}$ seja uma partição de $\mathcal{A}$ e para cada $b \in \mathcal{B}$ tenhamos observações na forma $\mathbf{y}_b = (y_1, y_2, \dots, y_{A_b})$ em $\mathcal{A}_b$. Então, de (3.9) tem-se que a função de distribuição marginal do vetor aleatório $\mathbf{Y}_b \sim KEM(\boldsymbol{\theta}, \alpha_b)$, é dada por:

$$\bar{F}(\mathbf{y}_b; \boldsymbol{\theta}, \alpha_b) = \exp \{ -\zeta_b \}, \quad \zeta_b > 0, \quad (3.15)$$

em que $\zeta_b = \left(\sum_{a \in \mathcal{A}_b} \nu_{a,b} \right)^{\alpha_b}$ e $\nu_{a,b}$ é dado em (3.7).

Proposição 3.2.2. *A função densidade do vetor aleatório $\mathbf{Y}_b \sim KEM(\boldsymbol{\theta}, \alpha_b)$, $b \in \mathcal{B}$ é expressa por*

$$f(\mathbf{y}_b; \boldsymbol{\theta}_b, \alpha_b) = \alpha_b^{A_b} \left(\prod_{a \in \mathcal{A}_b} \omega_{a,b} h(y_{a,b}; \boldsymbol{\theta}) \right) \exp(-\zeta_b) \zeta_b^{1 - \frac{A_b}{\alpha_b}} Q_{A_b}(\zeta_b, \alpha_b) \quad (3.16)$$

em que $\boldsymbol{\theta}_b$ é o subvetor de parâmetros, A_b é a cardinalidade do conjunto $\mathcal{A}_b$, $h(\cdot)$ é dada como em (3.12) e $Q_1(\zeta_b, \alpha_b) = 1$ com

$$Q_{A_b}(\zeta_b, \alpha_b) = \left(\frac{A_b - 1 - \alpha_b}{\alpha_b} + \zeta_b \right) Q_{A_b-1}(\zeta_b, \alpha_b) - \zeta_b \frac{\partial}{\partial \zeta_b} Q_{A_b-1}(\zeta_b, \alpha_b), \quad A_b \geq 2. \quad (3.17)$$

Esta função densidade conjunta por grupos (3.16) é encontrada de forma similar a Shi (1995), em que $Q_{A_b}(\zeta_b, \alpha_b)$ é um polinômio de ordem $(A_b - 1)$ em ζ_b .

A demonstração desse resultado é consequência da Proposição (3.2.1) e se encontra no Apêndice A.

3.2.4 Geração de Vetores Aleatórios Por Grupos

Para cada conjunto de observações $\mathbf{y}_b = (y_1, y_2, \dots, y_{A_b})$ que pertencem a $\mathcal{A}_b$, apresentamos a seguinte transformação:

$$(w_1, w_2, \dots, w_{A_b-1}, \zeta_b) = \left(\frac{\nu_{1,b}}{\sum_{a \in \mathcal{A}_b} \nu_{a,b}}, \frac{\nu_{2,b}}{\sum_{a \in \mathcal{A}_b} \nu_{a,b}}, \dots, \frac{\nu_{A_b-1,b}}{\sum_{a \in \mathcal{A}_b} \nu_{a,b}}, \left(\sum_{a \in \mathcal{A}_b} \nu_{a,b} \right)^{\alpha_b} \right). \quad (3.18)$$

Note que $w_1 + w_2 + \dots + w_{A_b} = 1$ e $\nu_{a,b} = w_a \zeta_b^{\frac{1}{\alpha_b}}$. Assim,

$$\frac{\partial y_{a,b}}{\partial w_j} = \frac{\partial y_{a,b}}{\partial \nu_{a,b}} \times \frac{\partial \nu_{a,b}}{\partial w_a} = \begin{cases} \frac{\partial y_{a,b}}{\partial \nu_{a,b}} \times \zeta_b^{\frac{1}{\alpha_b}}, & \text{se } a = j, \\ 0, & \text{se } a \neq j. \end{cases}$$

e, como $\nu_{a,b} = \zeta_b^{\frac{1}{\alpha_b}} - \sum_{j \neq a} \nu_{j,b}$, segue que:

$$\frac{\partial y_{a,b}}{\partial \zeta_b} = \frac{\partial y_{a,b}}{\partial \nu_{a,b}} \times \frac{\partial \nu_{a,b}}{\partial \zeta_b} = \frac{\partial y_{a,b}}{\partial \nu_{a,b}} \times \frac{1}{\alpha_b} \zeta_b^{\frac{1}{\alpha_b} - 1}.$$

Portanto, o módulo do jacobiano da transformação $(y_{1,b}, y_{2,b}, \dots, y_{A_b,b}) \mapsto (w_1, \dots, w_{A_b-1}, \zeta_b)$ é dado por

$$|\mathbf{J}| = \frac{\zeta_b^{\frac{A_b}{\alpha_b} - 1}}{\alpha_b \left(\prod_{a=1}^{A_b} \left| \frac{\partial y_{a,b}}{\partial \nu_{a,b}} \right| \right)^{-1}}. \quad (3.19)$$

Observe que

$$\left| \frac{\partial y_{a,b}}{\partial \nu_{a,b}} \right|^{-1} = \left| \frac{\partial \nu_{a,b}}{\partial y_{a,b}} \right| = \omega_{a,b} h(y_a; \boldsymbol{\theta}).$$

Logo,

$$f(w_1, \dots, w_{A_b-1}, \zeta_b) = \Gamma(A_b) \times Q_{A_b}(\zeta_b, \alpha_b) \frac{\alpha_b^{A_b-1}}{\Gamma(A_b)} e^{-\zeta_b}. \quad (3.20)$$

Portanto, $\mathbf{W}_b$ é independente de ζ_b . Além disso, temos que $\mathbf{W}_b \sim \text{Dirichlet}(1, \dots, 1)$ simétrica e ζ_b é uma mistura de distribuições Gama, (Shi, 1995), isto é,

$$f(\zeta_b) = \sum_{a \in A_b} q_{A_b,a} \frac{1}{\Gamma(a)} \zeta_b^{a-1} e^{-\zeta_b}, \quad \zeta_b > 0$$

em que os pesos $q_{A_b,a}$ são determinados por $Q_{A_b}(\zeta_b, \alpha_b)$.

Por exemplo, se $A_b = 2$, teremos

$$\zeta_b \sim (1 - \alpha_b)Gama(1, 1) + \alpha_b Gama(2, 1) \quad (3.21)$$

$$w_{1,b} \sim Beta(1, 1) \quad \text{e}$$

$$w_{2,b} = 1 - w_{1,b}.$$

A partir de (3.7) e da transformação (3.18) obtemos que

$$\begin{aligned} -\omega_{a,b} \log(1 - y_a^p)^q &= w_a \zeta_b^{\frac{1}{\alpha_b}} \\ y_a &= \left\{ 1 - \left[\exp\left(-\frac{w_a \zeta_b^{1/\alpha_b}}{\omega_{a,b}}\right) \right]^{1/q} \right\}^{1/p}. \end{aligned} \quad (3.22)$$

Então, geramos valores para as distribuições de $\mathbf{W}_b$ e ζ_b . Em seguida utilizamos o método da transformação inversa descrito na Equação (3.22) para obter o conjunto de observações $\mathbf{y}_b$. A Figura 3.2 mostra a geração bivariada com $b = 1$, $p_1 = p_2 = 15$ e $q_1 = q_2 = 100$ e $\alpha \in \{0.15, 0.25, 0.50, 0.75, 0.85, 0.95\}$. Observe que a associação entre as variáveis deixa de ser linear à medida que o valor de $\alpha \rightarrow 1$. Esse comportamento indica que há uma diminuição na relação de dependência entre as variáveis.

3.2.5 Inferência Por Grupos

Uma abordagem adequada para realizar o ajuste do modelo é utilizar o método da máxima verossimilhança por grupos. Para isso, considere a densidade por grupo apresentada em 3.16. Nesse caso, a função de log-verossimilhança do modelo Log-kumaraswamy G-exponencializada Multivariada por grupo pode ser escrita da seguinte forma:

$$\ell(\boldsymbol{\theta}, \alpha_b; \mathbf{y}_b) = A_b \log \alpha_b + \sum_{a \in \mathcal{A}_b} \log[\omega_{a,b} h(y_{a,b}; \boldsymbol{\theta})] - \zeta_b + \left(1 - \frac{A_b}{\alpha_b}\right) \log \zeta_b + \log Q_{A_b}(\zeta_b, \alpha_b).$$

A fórmula recursiva da densidade mencionada é útil para programação computacional quando não há ajuste do modelo por meio de covariáveis. No entanto, ao incluirmos covariáveis, os algoritmos convencionais de otimização geralmente enfrentam dificuldades no processo de maximização.

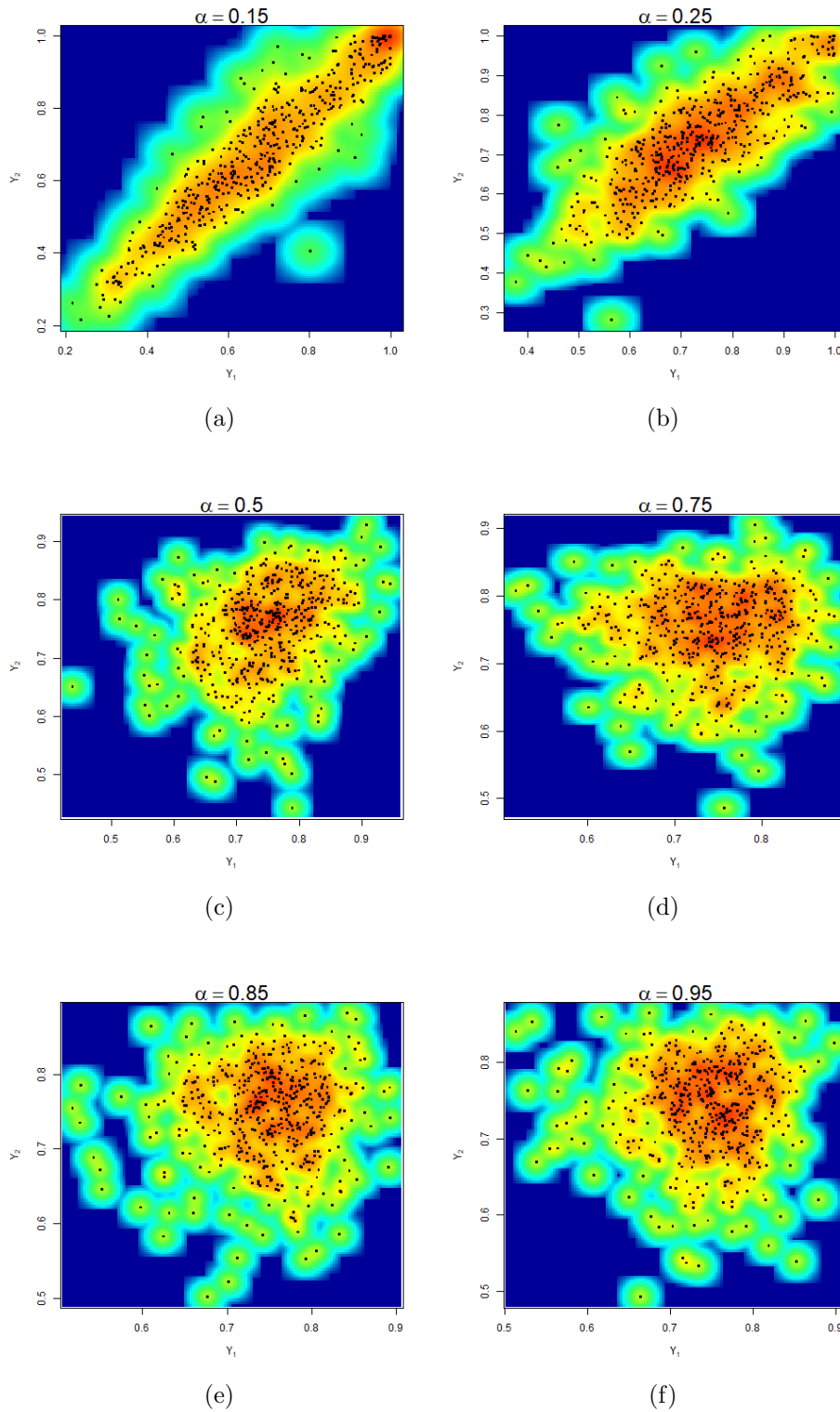


Figura 3.2: Geração de amostras bivariadas do modelo Kumaraswamy Exponencializada Multivariado para diferentes valores de α .

CAPÍTULO 4

Artigo 1

Modelo de Regressão *Power* Mín-Estável para Séries Temporais com Valores Extremos em Intervalo Limitado

NASCIMENTO, L. B. F^a, LIMA, M.S. ^a, and DUCZMAL, L H^b

^a*Departamento de Estatística, Universidade Federal do Amazonas, 69077-000, Manaus, AM, Brazil, Tel.: +55-92-33052850, Fax: +55-92-33052840;*

^b*Departamento de Estatística, Universidade Federal de Minas Gerais, Belo Horizonte, MG, Brazil*

Resumo

Neste paper, é proposto um modelo de regressão *power Min-stable* para uma Série Temporal de valores extremos observados em intervalo limitado. O modelo é útil, por exemplo, para ser utilizado em situações onde a variável de interesse é o valor mínimo de uma série restrita ao intervalo (0,1) e está relacionada com outras variáveis através de uma estrutura de regressão. A dependência nos extremos da série é induzida através da marginalização da distribuição Kumaraswamy condicionada a um processo latente com distribuição α -estável. Algumas propriedades do modelo são apresentadas, assim como o Extremograma da série. Procedimentos para estimação e inferência são discutidos e um algoritmo EM, disponível na linguagem R Core Team (2023), foi desenvolvido para estimar os parâmetros. Como um caso particular, o modelo foi utilizado para análise da mínima umidade relativa do ar observada em Manaus no Estado do Amazonas-Brasil.

Palavras chaves

Kumaraswamy; Distribuição Estável; *Power Min-Stable*; Extremos.

1. Introdução

Modelos estatísticos para Séries Temporais de valores extremos têm recebido muita atenção nas áreas de Seguros, Finanças, Hidrologia, Química Atmosférica, dados ambientais, entre outras. Embora os métodos para dados extremos univariados estacionários estejam bem desenvolvidos (Tavares, 1980; Sim, 1986; Balkema and Haan, 1988; Davis and Resnick, 1989; Smith, 1989; Küchenhoff and Thamerus, 1996; Hall and Tajvidi, 2000; Hall et al., 2002; Eastoe, 2007; Eastoe and Tawn, 2009), a modelagem de dados extremos dependentes no tempo, ainda, é uma área ativa de pesquisa em série temporal. Recentemente, um modelo para extremos condicionais que permite uma ampla gama de possíveis estruturas de dependência em uma série estacionária é apresentado em Auld and Papastathopoulos (2023) que realiza inferência no modelo via método de Monte Carlo. Mhatre and Cooley (2020) desenvolvem modelos de séries temporais para extremos por meio de combinações lineares transformadas e Regularmente Variantes que são adequados para modelar dados extremos que exibem dependência assintótica ou nas caudas. Modelos para detectar extremos heterocedásticos são descritos em Bücher and Jennessen (2022); Einmahl et al. (2016). Bortot and Gaetan (2014) apresenta uma abordagem hierárquica para modelar extremos de uma série temporal

estacionária. As distribuições marginais desse modelo hierárquico pertencem à família Pareto Generalizada que está na classe de valores extremos. A inferência no modelo é realizada via *Pairwise likelihood*. Um outro modelo, aparentemente útil para modelagem de extremos é o autoregressivo Mittag-Leffler descrito em Sebastian and Gorenflo (2016). No entanto os autores não apresentam discussão sobre inferência no modelo. Modelos marginais autoregressivos e de médias móveis para extremos baseados na distribuição Gumbel são apresentados em Fougères et al. (2009). Chavez-Demoulin and Davison (2012) apresentam uma boa revisão sobre a modelagem de séries temporais de valores extremos e o extemograma de uma série é proposto em Davis and Mikosch (2009). Uma abordagem comum, na análise de séries de extremos, consiste em modelar tais eventos por máximos ou mínimos de um conjunto de dados obtidos por blocos contíguos (dias, semanas, meses) usando a teoria de valores extremos. O principal resultado dessa teoria é que a escolha dos modelos é limitada a três famílias de distribuições representadas por sua generalização chamada de Valor Extremo Generalizado (GEV). Esta família de distribuições pertence a classe de normalização linear max-estável (De Haan, 1984; Davis and Resnick, 1993), linear mín-Estável (De Haan and Pickands III, 1986; Penrose, 1992) que foram estendidas para normalização não-linear (Pantcheva, 1985; Pancheva, 1994) e têm uma longa história na modelagem de tais eventos assumindo valores extremos estacionários em um intervalo ilimitado. No entanto, existem valores extremos que necessitam de uma flexibilidade para modelagem de dados não estacionários com *spread* temporal não constante e limitados ao intervalo (a,b) , $a < b$. Uma variável de interesse nesse contexto é a umidade relativa do ar, a qual pertence ao intervalo $(0,1)$, sendo extremamente importante para engenharia, agricultura, meio ambiente e saúde humana. Particularmente, no Amazonas-Brasil, a umidade tem características não estacionárias e sazonais apresentando padrões muito diferentes quando comparados a outras regiões. Tem-se observado nos últimos anos uma variabilidade ou dispersão cada vez maior na umidade do ar devido a variações climáticas ocasionadas pela destruição das áreas verdes e efeitos *la Niña*. A saúde humana é a mais afetadas por estas variações. Uma umidade alta (acima de 0.7), pode gerar vários problemas de saúde, como respiratórios, alergias, sinusites, asma e favorecem a proliferação rápida de fungos e bactérias. Quando a umidade é muito baixa (menor que 0.3), pode ocorrer estresse por calor, sensação desagradável, elevação da temperatura corporal, doenças térmicas brandas e relacionadas com o quadro de desidratação e hipertermia. Assim, é importante modelar os riscos consecutivos da mínima umidade exceder (ou ser menor que) limites específicos, acima (ou abaixo) dos quais pode gerar efeitos adversos a saúde. A teoria para valores extremos, com suporte limitado tem sido proposta Kunin (2004) e Peng et al. (2015) apresenta uma versão unificada chamada p-max estável distribuições. Modelos de regressão na média e mediana para valores limitados no contexto de séries temporais foram desenvolvidos, respectivamente, em Rocha and Cribari-Neto (2009) e Bayer et al. (2017) mas não acomodam a dependência nos extremos. Nossa proposta, neste paper, é um modelo de regressão para uma série temporal de valores mínimos limitados que são não estacionários, sazonais e com risco de excedência auto-regressivos. A dependência na série é induzida através da marginalização da distribuição Kumaraswamy (Kumaraswamy, 1980) condicionada a um processo latente com distribuição α -estável positiva (Uchaikin and Zolotarev, 1999). Algumas propriedades do modelo são apresentadas, assim como o Extremograma da série. Procedimentos para Inferência Estatística são discutidos e um algoritmo disponível em R Core Team (2023), foi desenvolvido para estimação dos parâmetros. No mais, o restante deste paper está organizado da seguinte forma: Na Seção 2, descrevemos o modelo e suas propriedades; Na Seção 3 propomos métodos

para estimação dos parâmetros; Uma avaliação numérica do modelo é apresentado na Seção 4, na Seção 5 a metodologia proposta é utilizada para modelar os dados de mínima unidade relativa do ar observada em Manaus-AM e, por fim descrevemos nossas conclusões e considerações sobre o modelo na Seção 6.

2. Modelo Proposto

Para definir nossa proposta, introduziremos um modelo condicionalmente independente para extremos e depois acomodamos a estrutura de dependência temporal na série através da função de excedência multivariada do processo marginal.

2.1. Modelo Condicional para Extremos

Seja $\mathbf{Y}_t = \{Y_t : t \in \mathbb{N}\}$ uma série resposta de um processo não estacionário de valores extremos $\in (0, 1)$ com processo não estocástico $\{(\mathbf{X}_t, \mathbf{W}_t) : t \in \mathbb{N}\}$ de covariáveis associadas. Onde, $\mathbf{X}_t = (X_{t,1}, X_{t,2}, \dots, X_{t,p})^\top$ é um vetor p -dimensional, $\mathbf{W}_t = (W_{t,1}, W_{t,2}, \dots, W_{t,k})^\top$ é um vetor k -dimensional e Y_t é um bloco mínimo, ou seja, o mínimo de muitas observações realizadas no tempo t , como o mínimo mensal da umidade relativa do ar. Considere que $\mathbf{y} = (y_1, y_2, \dots, y_n)$ representa uma amostra das primeiras n observações do processo $\mathbf{Y}_t$. Expressamos hierarquicamente Y_t por um modelo lei de potência (Power Law):

$$\begin{aligned} Y_t &= (1 - U_t^{\frac{1}{Z_t}})^{1/\nu_t}, \nu_t > 0, t = 1, 2, \dots, n, \\ U_t &\stackrel{i.i.d}{\sim} \text{Uniforme}(0, 1), \\ Z_t &= \frac{Z}{\delta_t} \text{ com } Z \sim \alpha - \text{Stable}, \alpha \in (0, 1), \end{aligned} \quad (1)$$

tal que U_t é independente do processo latente (ou um fator ambiental) Z_t com *spread* temporal $\delta_t > 0$. Para um melhor entendimento do modelo condicional em (1), observe que $Y_t|Z_t \sim \text{Kumaraswamy}(\nu_t, Z/\delta_t) \doteq K(\nu_t, Z/\delta_t)$ e defina $\Psi(a) = -\log(-\log(a))$. Então,

$$\Psi(y_t) = \log(\nu_t) + \varepsilon_t,$$

no qual $\varepsilon_t = \Psi\left(1 - U_t^{1/Z_t}\right) \sim \mathbb{EG}(1, 0, Z/\delta_t)$ é o erro não estacionário Gumbel exponencializado com locação 0, escala 1 e potência fracional Z/δ_t . Na prática a potência fracional torna os erros mais flexíveis para modelos que podem ter caudas mais leve (platykurtic) ou mais pesada (leptokurtic) do que a Gumbel padrão (Nadarajah and Kotz (2006)) e esperamos com isso capturar o efeito e a variação dos extremos (inferior e superior) na série temporal. No contexto de um modelo de regressão, podemos escrever, $\Psi(y_t) = \mathbf{x}_t^\top \beta + \varepsilon_t$ no qual $\log(\nu_t) = \mathbf{x}_t^\top \beta$ é o preditor para tendência temporal ou sazonalidade e, ainda, $\log(\delta_t) = \mathbf{w}_t^\top \gamma$ é o preditor para dispersão temporal com $\beta = (\beta_1, \dots, \beta_p)^\top$ e $\gamma = (\gamma_1, \dots, \gamma_k)^\top$. Se os coeficientes da regressão forem não significativos, o processo será estacionário. Seja $p_t = \mathbb{P}(Y_t > y_t|Z = z) = (1 - y_t^{\nu_t})^{\frac{z}{\delta_t}}$ a probabilidade condicional de excedência que avalia o risco de ocorrência de valores de Y_t superiores a um valor crítico y_t . Então (usando os momentos da distribuição

$\log z)$, a correlação $\rho(\Psi(p_t), \Psi(p_{t-1})) = 1 - \alpha^2$ de modo que $\alpha \in (0, 1)$ é uma medida de dependência entre os valores de excedência da série. Apesar do modelo (1) ser uma importante caracterização do nosso processo de extremos. Estamos interessados no comportamento conjunto $\mathbf{y} = (y_1, y_2, \dots, y_n)$ que pode ser avaliado pela função de excedência conjunta, $\bar{F}(\mathbf{y}) = \mathbb{P}(Y_1 > y_1, \dots, Y_n > y_n)$ através da integral,

$$\bar{F}(\mathbf{y}) = \int_z \bar{F}(\mathbf{y}|z) g_\alpha(z) dz = \int_z \mathbb{P}(Y_1 > y_1, \dots, Y_n > y_n | Z = z) g_\alpha(z) dz. \quad (2)$$

Onde $g_\alpha(z)$ é a densidade de Z com $\mathbb{E}(e^{-aZ_\alpha}) = e^{-a^\alpha}$, $a > 0$ (Stephenson, 2009). A análise da série através do processo conjunto por $\bar{F}(\mathbf{y})$ possui a vantagem que a característica autoregressiva do risco de excedência e o extremograma da série podem ser apresentados de uma forma analiticamente tratável. Na próxima seção, apresentamos as características desse novo modelo.

2.2. Modelo de Pontência mín-Estável

Nosso objetivo é construir um modelo conjunto para mínimos extremos no intervalo $(0, 1)$ observados ao longo do tempo. No entanto, podemos redefinir $(Y_t - a)/(b - a)$ quando $Y_t \in (a, b)$. Assuma que o processo $\mathbf{Y}_t$ satisfaz 1 tal que $Y_t | Z, t = 1, 2, \dots, n$ são condicionalmente independentes e $Z \sim \alpha - \text{stable}$. Defina $\Lambda_t(y_t, \nu_t, \delta_t) = -\delta_t^{-1} \log(1 - y_t^{\nu_t}) > 0$ com (ν_t, δ_t) dados na Seção 2.1. Então, a integral em (2) é:

$$\begin{aligned} \bar{F}(\mathbf{y}) &= \int_z \left(\prod_{t=1}^n (1 - y_t^{\nu_t})^{z/\delta_t} \right) g_\alpha(z) dz \\ &= \int_z e^{-z \sum_{t=1}^n \Lambda_t(y_t, \nu_t, \delta_t)} g_\alpha(z) dz \\ &= e^{-(\sum_{t=1}^n \Lambda_t(y_t, \nu_t, \delta_t))^\alpha}, \end{aligned} \quad (3)$$

que é uma medida de probabilidade somente para $0 < \alpha \leq 1$ (Hougaard, 1986). Note que $\Lambda_t(y_t, \nu_t, \delta_t)$ é simplesmente a função de excedência acumulada para a distribuição $K(\nu_t, 1/\delta_t)$ com taxa de excedência $\lambda_t = \frac{\partial}{\partial y_t} \Lambda_t = \nu_t y_t^{\nu_t-1} (1 - y_t^{\nu_t})^{-1} / \delta_t$. Então quando $\alpha = 1$ os componentes do processo $\mathbf{Y}_t$ são independentes com $Y_t \sim K(\nu_t, 1/\delta_t)$ e outros modelos do tipo $\text{Beta}(\nu_t, 1)$ ou $\text{Beta}(1, 1/\delta_t)$ também podem ser obtidos. Uma propriedade importante no contexto de valores extremos ocorre quando as marginais e o mínimo pertencem a mesma classe da conjunta. Por isso, mostramos duas proposições a seguir que garantem essas propriedades e, ainda, o mínimo dos blocos converge para uma distribuição mín-Estável sob uma transformação não linear.

Proposição 2.1. *Para o modelo obtido em (3), a distribuição do mínimo sobre qualquer bloco ou subconjunto $n_0 \subset \mathbb{N}$ pertence a mesma classe.*

Para prova, defina $Y_0 = \min(Y_1, Y_2, \dots, Y_{n_0})$. Então,

$$\mathbb{P}(Y_0 \leq y_0) = 1 - \mathbb{P}(Y_1 > y_0, Y_2 > y_0, \dots, Y_{n_0} > y_0) = 1 - e^{-(\sum_{t=1}^{n_0} \Lambda_t(y_0, \nu_t, \delta_t))^\alpha}$$

de modo que $\bar{F}_{Y_0}(y_0) = e^{-(\sum_{t=1}^{n_0} \Lambda_t(y_0, \nu_t, \delta_t))^\alpha}$. Em particular, se n_0 é um conjunto de um ponto, temos que as marginais são do tipo $\bar{F}_{Y_t}(y_t) = e^{-(\Lambda_t(y_t, \nu_t, \delta_t))^\alpha}$ que pertencem a mesma classe, assim como, as distribuições bivariadas. Nas distribuições marginais do

modelo, usando a transformação integral $1 - U_t = \bar{F}_{Y_t}(y_t)$, com $U_t \sim \text{Uniforme}(0,1)$. Obtemos que,

$$y_t = \left[\underbrace{1 - \exp \left\{ - \left(\frac{-\log(1 - U_t)}{(1/\delta_t)^\alpha} \right)^{1/\alpha} \right\}}_{b_t} \right]^{1/\nu_t}, \quad (4)$$

com mediana κ_t igual a

$$\kappa_t = \left[\underbrace{1 - \exp \left\{ - \left(\frac{-\log(0.5)}{(1/\delta_t)^\alpha} \right)^{1/\alpha} \right\}}_{c_t} \right]^{1/\nu_t}. \quad (5)$$

Podemos escrever

$$\Psi(y_t) = \Psi(\kappa_t) + \zeta_t,$$

onde $\zeta_t = \Psi(b_t) - \Psi(c_t)$ com $\mathbb{E}[\zeta_t] \approx 0$. Fazamos $\Psi(\kappa_t) = \mathbf{x}_t^\top \vartheta$. Então, $\Psi(y_t) = \mathbf{x}_t^\top \vartheta + \zeta_t$, que tem forma similar a Seção 2.1. Observe que no modelo Condicional e Marginal, Y_t é expresso em lei de potência (*Power Law*). A estrutura de regressão permanece a mesma para os modelos considerando a função de ligação $\Psi(\kappa_t)$ no modelo marginal e $\log \nu_t$ no condicional com $\beta = \vartheta$ e, ainda, valores para y_t podem ser gerados usando a expressão (4)

Proposição 2.2. *O processo $\mathbf{Y}_t = \{Y_t : t \in \mathbb{N}\}$ com função de excedência dado em (3) é Power Min-Stable no sentido Marshall and Olkin (1983) e Pantcheva (1985).*

Para a prova, suponha que $V \sim \text{Weibull}(1, \alpha)$, $\alpha \in (0, 1)$ e defina $Y = (1 - e^{-vb})^{1/a}$, com $a, b > 0$. Então, $y \in (0, 1)$ e $\bar{F}_Y(y) = \bar{F}_V(-b^{-1} \log(1 - y^a))$ que tem a mesma distribuição de (3) em cada coordenada do vetor $(y_1, y_2, \dots, y_n)$. Logo, deve existir uma transformação não linear $d_n(v) \in (0, 1)$ com $v > 0$ tal que,

$$\lim_{n \rightarrow \infty} \bar{F}_{\mathbf{y}}^n(d_n(\mathbf{v})) = \bar{H}(\mathbf{v}) = e^{-(\sum_{i=1}^n v_i)^\alpha}, v_i > 0, \quad (6)$$

onde $\bar{H}$ é uma função de excedência multivariada Weibull. Sem perda de generalidade, defina a função potência $d_n(v) = [1 - (1 - v/c_n)^{c_n}]^{1/a}$, com $(1 - v/c_n) < 1$ e $c_n \rightarrow \infty$. Então, em (3) aplique $d_n(v)$ em cada y , substitua δ_t por $\delta_{t,n} = n^{1/\alpha} \delta_t$ e obtenha o limite em (6). Chamamos esse modelo de α -Power Min-Stable, **PMS**(α), porque é mínimo-Estável (Marshall and Olkin (1983)) para uma transformação do tipo potência (vê, Pantcheva (1985)) com dependência extrema controlada por α (see, section 2.1). Algumas propriedades condicionais úteis do modelo **PMS**(α) para realização da inferência são apresentadas a seguir considerando que $\Lambda_t = \Lambda_t(y_t, \nu_t, \delta_t)$, $\lambda_t = \frac{\partial}{\partial y_t} \Lambda_t = \nu_t y_t^{\nu_t-1} (1 - y_t^{\nu_t})^{-1} / \delta_t$, o par (ν_t, δ_t) são funções das covariáveis dadas na Seção (2.1), $\boldsymbol{\theta} = (\beta, \gamma) \subset \mathbb{R}^{p+k}$ é o vetor de coeficientes da regressão e $\Lambda_n(\mathbf{y}; \boldsymbol{\theta}) = \sum_{t=1}^n \Lambda_t(y_t, \nu_t, \delta_t)$.

- (1) **P1: Densidade:** A função de densidade conjunta de $\mathbf{y} = (y_1, y_2, \dots, y_n)$ é dada por

$$f(\mathbf{y}_n; \boldsymbol{\theta}, \alpha) = (-1)^n \frac{\partial^{(n)} \bar{F}}{\partial y_1, \dots, \partial y_n} = e^{-\Lambda_n^\alpha(\mathbf{y}; \boldsymbol{\theta})} D_{1,2,\dots,n}(\mathbf{y}; \boldsymbol{\theta}, \alpha),$$

onde,

$$D_{1,2,\dots,n}(\mathbf{y}; \boldsymbol{\theta}, \alpha) = D_{1,2,\dots,(n-1)}(\mathbf{y}; \boldsymbol{\theta}, \alpha) D_n(\mathbf{y}; \boldsymbol{\theta}, \alpha) + \frac{\partial}{\partial y_n} D_{1,2,\dots,n-1}(\mathbf{y}; \boldsymbol{\theta}, \alpha)$$

e $t = 1, 2, \dots, n$.

$$D_t(\mathbf{y}; \boldsymbol{\theta}, \alpha) = \alpha \lambda_t \Lambda_n^{\alpha-1}(\mathbf{y}; \boldsymbol{\theta}), \quad D_0(\mathbf{y}; \boldsymbol{\theta}, \alpha) = 1.$$

Para prova dessa propriedade nota-se por (3) que $\frac{\partial \bar{F}}{\partial y_t} = e^{-\Lambda_n^\alpha(\mathbf{y}; \boldsymbol{\theta})} D_t(\mathbf{y}; \boldsymbol{\theta}, \alpha)$. Procedendo por diferenciação recursiva, teremos que se

$$\frac{\partial^{(n-1)} \bar{F}}{\partial y_1, \dots, \partial y_{n-1}} = e^{-\Lambda_n^\alpha(\mathbf{y}; \boldsymbol{\theta})} D_{1,2,\dots,n-1}(\mathbf{y}; \boldsymbol{\theta}, \alpha),$$

então,

$$\begin{aligned} \frac{\partial^{(n)} \bar{F}}{\partial y_1, \dots, \partial y_n} &= \frac{\partial}{\partial y_n} \left[e^{-\Lambda_n^\alpha(\mathbf{y}; \boldsymbol{\theta})} D_{1,2,\dots,n-1}(\mathbf{y}; \boldsymbol{\theta}, \alpha) \right] \\ &= e^{-\Lambda_n^\alpha(\mathbf{y}; \boldsymbol{\theta})} D_{1,2,\dots,n-1}(\mathbf{y}; \boldsymbol{\theta}, \alpha) D_n(\mathbf{y}; \boldsymbol{\theta}, \alpha) \\ &\quad + e^{-\Lambda_n^\alpha(\mathbf{y}; \boldsymbol{\theta})} \frac{\partial}{\partial y_n} D_{1,2,\dots,n-1}(\mathbf{y}; \boldsymbol{\theta}, \alpha) \\ &= e^{-\Lambda_n^\alpha(\mathbf{y}; \boldsymbol{\theta})} D_{1,2,\dots,n}(\mathbf{y}; \boldsymbol{\theta}, \alpha). \end{aligned}$$

Assim, a log-verossimilhança do modelo é dada por

$$\begin{aligned} l(\boldsymbol{\theta}, \alpha; \mathbf{y}) &= -\Lambda_n^\alpha(\mathbf{y}; \boldsymbol{\theta}) + \log(D_{1,2,\dots,n}(\mathbf{y}; \boldsymbol{\theta}, \alpha)) \\ &= -\sum_{t=1}^n \left(-\frac{1}{\delta_t} \log(1 - y_t^{\nu_t}) \right)^\alpha + \log(D_{1,2,\dots,n}(\mathbf{y}; \boldsymbol{\theta}, \alpha)). \end{aligned} \quad (7)$$

Essa fórmula recursiva da densidade é conveniente para programação computacional quando não há ajuste do modelo por covariáveis. No entanto, ao incluirmos covariáveis os algoritmos convencionais de otimização tendem a falhar no processo de maximização.

- (2) **P2: Densidades Bivariadas, Condicionais e verossimilhanças:** As densidades bivariadas e condicionais do modelo $\mathbf{PMS}(\alpha)$ podem ser utilizadas para estimação dos parâmetros do processo. Pela proposição (2.1), derivando a função de excedência com respeito a y_t obtemos que as densidades são da forma $f(y_t) = \alpha \lambda_t \Lambda_t^{\alpha-1} e^{-\Lambda_t^\alpha}$, $t = 1, 2, \dots, n$. Usando **P1**, para qualquer par $(h, t) \in \mathbb{N}$, a densidade bivariada é dada por,

$$f(y_t, y_h | \boldsymbol{\theta}, \alpha) = \alpha \lambda_t \lambda_h e^{-(\Lambda_h + \Lambda_t)^\alpha} (\Lambda_h + \Lambda_t)^{\alpha-2} [\alpha (\Lambda_h + \Lambda_t)^\alpha + (1 - \alpha)],$$

com $h = t-1$ temos que a distribuição condicional de primeira ordem do processo é

$$f(y_t|y_{t-1}, \boldsymbol{\theta}, \alpha) = \lambda_t \Lambda_{t-1}^{1-\alpha} e^{\Lambda_{t-1}^\alpha - (\Lambda_{t-1} + \Lambda_t)^\alpha} (\Lambda_{t-1} + \Lambda_t)^{\alpha-2} [\alpha(\Lambda_{t-1} + \Lambda_t)^\alpha + (1-\alpha)]$$

de modo que a *pairwise composite log-likelihood* e a *conditional log-likelihood* de primeira ordem, são dadas, respectivamente, por:

$$l_{pc}(\boldsymbol{\theta}, \alpha; \mathbf{y}) = \sum_{t=1}^{n-1} \sum_{h=t+1}^n \log f(y_t, y_h | \boldsymbol{\theta}, \alpha), \quad (8)$$

$$\begin{aligned} l_{cd}(\boldsymbol{\theta}, \alpha; \mathbf{y}) &= \log f(y_1 | \boldsymbol{\theta}, \alpha) + \sum_{t=2}^n \log f(y_t | y_{t-1}, \boldsymbol{\theta}, \alpha) \\ &= \log(\alpha \lambda_t \Lambda_t^{\alpha-1} e^{-\Lambda_t^\alpha}) + \sum_{t=2}^n \log f(y_t | y_{t-1}, \boldsymbol{\theta}, \alpha). \end{aligned} \quad (9)$$

Estas funções, também, podem ser maximizadas para inferência no modelo. Generalizando para $h = 1, 2, \dots, t-1$, teremos diretamente de **P1**:

$$\begin{aligned} f(y_t | y_{t-1}, \dots, y_{t-h}) &= \frac{D_{t-h, \dots, t}(\mathbf{y}_t; \boldsymbol{\theta}, \alpha)}{D_{t-h, \dots, t-1}(\mathbf{y}_{t-1}; \boldsymbol{\theta}, \alpha)} \\ &\times \exp \left\{ \left(\sum_{j=t-h}^{t-1} \Lambda_j \right)^\alpha - \left(\Lambda_t + \sum_{j=t-h}^{t-1} \Lambda_j \right)^\alpha \right\}. \end{aligned} \quad (10)$$

- (3) **P3: excedência de primeira e segunda ordem** : No modelo, a função de excedência condicional de primeira ordem é $\mathbb{P}(Y_t > y_t | Y_{t-1} > y_{t-1}) = \frac{\mathbb{P}(Y_t > y_t, Y_{t-1} > y_{t-1})}{\mathbb{P}(Y_{t-1} > y_{t-1})} = e^{\Lambda_{t-1}^\alpha - (\Lambda_{t-1} + \Lambda_t)^\alpha}$, a função de excedência condicional instatânea (pontual) é:

$$\begin{aligned} \mathbb{P}(Y_t > y_t | Y_{t-1} = y_{t-1}) &= \frac{\mathbb{P}(Y_t > y_t, Y_{t-1} = y_{t-1})}{f(y_{t-1})} \\ &= \Lambda_{t-1}^{1-\alpha} (\Lambda_{t-1} + \Lambda_t)^{\alpha-1} e^{\Lambda_{t-1}^\alpha - (\Lambda_{t-1} + \Lambda_t)^\alpha}, \end{aligned} \quad (11)$$

de modo que, fazendo a transformação de variável $u = (\Lambda_{t-1} + \Lambda_t)^\alpha - \Lambda_{t-1}^\alpha$ obteremos $\mathbb{E}(\Lambda_t | Y_{t-1}) = \int_0^1 \lambda_t \mathbb{P}(Y_t > y_t | Y_{t-1} = y_{t-1}) dy_t = \alpha^{-1} \Lambda_{t-1}^{1-\alpha}$. Fazendo $u = (\Lambda_{t-1} + \Lambda_t)^\alpha$ teremos $\mathbb{E}(\Lambda_t | Y_{t-1} > y_{t-1}) = \int_0^1 \lambda_t \mathbb{P}(Y_t > y_t | Y_{t-1} > y_{t-1}) dy_t = \alpha^{-1} \Gamma(1/\alpha, \Lambda_{t-1}^\alpha) e^{\Lambda_{t-1}^\alpha}$, onde $\Gamma(s, a)$ é a função gama incompleta superior. Estas esperanças condicionais são úteis para previsão, um passo a frente, do risco de excedência do processo (veja, Seção 3.4). A excedência bivariada condicional, ou de segunda ordem, é dada por:

$$\begin{aligned} \mathbb{P}(Y_t > y_t, Y_{t-1} > y_{t-1} | Y_{t-2} = y_{t-2}) &= \Lambda_{t-2}^{1-\alpha} (\Lambda_{t-2} + \Lambda_{t-1} + \Lambda_t)^{\alpha-1} \\ &\times e^{\Lambda_{t-2}^\alpha - (\Lambda_{t-2} + \Lambda_{t-1} + \Lambda_t)^\alpha}. \end{aligned}$$

Essa probabilidade é útil para previsão de riscos adversos com dois períodos de antecedência permitindo o ajuste de medidas de prevenção e controle. Obviamente, essa probabilidade pode ser generalizada para k períodos de antecedência usando a expressão:

$$\frac{\mathbb{P}(Y_t > y_t, \dots, Y_{t-k+1} > y_{t-k+1} | Y_{t-k} = y_{t-k})}{f(y_{t-k})}.$$

Onde o numerador é $-\frac{\partial}{\partial y_{t-k}} \mathbb{P}(Y_t > y_t, \dots, Y_{t-k+1} > y_{t-k+1}, Y_{t-k} > y_{t-k})$.

- (4) **P4 Extremograma do Modelo:** A quantificação da dependência extrema tem sido um tema central nos métodos probabilísticos para Séries de Valores Extremos. O correlograma para eventos extremos ou simplesmente extremograma da série (Davis and Mikosch, 2009) é definido por $\mathbb{P}(Y_t > y_u | Y_{t-h} > y_u)$, $h = 1, 2, 3, \dots, t-1$, onde y_u é um ponto de alta excedência (ou limite superior de excedência), ou também, no caso de baixa excedência $\mathbb{P}(Y_t < y_l | Y_{t-h} < y_l)$, onde y_l é um ponto de baixa excedência (ou limite inferior de excedência). No primeiro caso, temos que

$$\begin{aligned} \mathbb{P}(Y_t > y_u | Y_{t-h} > y_u) &= e^{\Lambda_{t-h}^\alpha(y_u) - (\Lambda_{t-h}(y_u) + \Lambda_t(y_u))^\alpha} \\ &= e^{\Lambda_{t-h}^\alpha(y_u) \left[1 - \left(1 + \frac{\Lambda_t(y_u)}{\Lambda_{t-h}(y_u)} \right)^\alpha \right]}. \end{aligned} \quad (12)$$

Note em (12) que se o processo é estacionário, tal que $\nu_t = \nu$ e $\delta_t = \delta$ para todo $t \in \mathbb{N}$, então, $\mathbb{P}(Y_t > y_u | Y_{t-h} > y_u) = e^{(-\delta^{-1} \log(1-y_u^\nu))^\alpha [1-2^\alpha]}$. Logo, se $\alpha \rightarrow 0$, $\mathbb{P}(Y_t > y_u | Y_{t-h} > y_u) \rightarrow 1$, e os extremos são totalmente dependentes. Quando $\alpha \rightarrow 1$, $\mathbb{P}(Y_t > y_u | Y_{t-h} > y_u) \rightarrow (1 - y_u^\nu)^{1/\delta}$ que é a excedência do processo $K(\nu, \delta^{-1})$ indicando a independência. Note, ainda, que o log-Extremograma é da forma,

$$\chi_h(y_u) = \log \mathbb{P}(Y_t > y_u | Y_{t-h} > y_u) = \Lambda_{t-h}^\alpha(y_u) \left[1 - \left(1 + \frac{\Lambda_t(y_u)}{\Lambda_{t-h}(y_u)} \right)^\alpha \right],$$

e $\Lambda_t(y_u)/\Lambda_{t-h}(y_u) = [e^{-(\mathbf{w}_t - \mathbf{w}_{t-h})^\top \gamma}] (\log(1 - y_u^{\nu_t}) / \log(1 - y_u^{\nu_{t-h}}))$ de modo que γ controla diretamente a dispersão nos extremos. Se $\mathbf{w}_t^\top = (1, t)$ com $\gamma = (\gamma_1, \gamma_2)^\top$. Então, $\Lambda_t(y_u)/\Lambda_{t-h}(y_u) = [e^{-h\gamma_2}] (\log(1 - y_u^{\nu_t}) / \log(1 - y_u^{\nu_{t-h}}))$ que decresce com h . Façamos ainda, $\nu_t = \nu$, então, $\chi_h(y_u) = \Lambda_{t-h}^\alpha(y_u) [1 - (1 + e^{-h\gamma_2})^\alpha]$. No segundo caso, o Extremograma para limites inferiores de excedência é dado por

$$\begin{aligned} \mathbb{P}(Y_t < y_l | Y_{t-h} < y_l) &= \frac{\mathbb{P}(Y_t < y_l, Y_{t-h} < y_l)}{\mathbb{P}(Y_{t-h} < y_l)} \\ &= 1 + \frac{1 - \mathbb{P}(Y_t > y_l) - [1 - \mathbb{P}(Y_t > y_l, Y_{t-h} > y_l)]}{1 - \mathbb{P}(Y_{t-h} > y_l)}. \end{aligned}$$

Usando a aproximação $\log(p) = 1 - p$, obtemos

$$\mathbb{P}(Y_t < y_l | Y_{t-h} < y_l) \approx 1 + \left(\frac{\Lambda_t(y_l)}{\Lambda_{t-h}(y_l)} \right)^\alpha - \left(1 + \frac{\Lambda_t(y_l)}{\Lambda_{t-h}(y_l)} \right)^\alpha. \quad (13)$$

Se o processo é estacionário, $\mathbb{P}(Y_t < y_l | Y_{t-h} < y_l) \approx 2 - 2^\alpha$ o qual é similar ao Extremograma do modelo extremo logístico (Coles et al., 1999).

3. Inferência Estatística do Modelo

3.1. Estimação dos Parâmetros

A estimação dos parâmetros do modelo $\mathbf{PMS}(\alpha)$ pode ser realizada através da maximização da função de verossimilhança e pode depender da ordem e forma de dependência especificada para o modelo. Por exemplo, para verossimilhança composta aos pares dada em (8) ou condicional de h -ésima ordem dada em (10) podemos utilizar os *Composite Likelihood Methods* (Varin et al., 2011). Se o interesse é analisar a dependência condicional de primeira ordem expressando a verossimilhança como o produto de densidades condicionais como em (9). Então, é relativamente fácil calcular as estimativas de máxima verossimilhança e podemos explorar o fato de que sua derivada com relação aos parâmetros do modelo é um martingale integrável com média zero, veja capítulo 6 de Hall and Heyde (1980). O problema na utilização destas verossimilhanças é que sua ordem de dependência pode ser mal especificada e afetar o cômputo da matriz de covariâncias dos coeficientes da regressão. Assim utilizaremos um algoritmo **EM** (Dempster et al., 1977) que nos fornece uma forma analiticamente visível para avaliar a variância assintótica dos estimadores sem depender da estrutura de dependência.

3.1.1. Estimação dos Parâmetros da Regressão via algoritmo EM

Considere o modelo $\mathbf{PMS}(\alpha)$ em (3) com $\alpha \in (0, 1)$, $g_\alpha(z)$ a densidade do fator latente (ou dado faltante) $Z \sim \alpha$ -estável. , $Y_t | Z = z \sim K(\nu_t, z/\delta_t)$ com $\log(\nu_t) = \mathbf{x}_t^\top \beta$, $\log(\delta_t) = \mathbf{w}_t^\top \gamma$ descritos na Seção 2.1 e $\boldsymbol{\theta} = (\beta, \gamma) \in \mathbb{R}^{p+k}$. Marginalmente a Z o modelo tem distribuição $\mathbf{PMS}(\alpha)$. Então a log-verossimilhança completa é, $l(\boldsymbol{\theta}, \alpha; \mathbf{y}, z) = \sum_{t=1}^n \log f(y_t | z, \boldsymbol{\theta}, \alpha) + \log g_\alpha(z) = l(\boldsymbol{\theta}; \mathbf{y}, z) + l(\alpha; z)$. Onde $l(\alpha; z) = n \log z + \log g_\alpha(z)$ e $l(\boldsymbol{\theta}; \mathbf{y}, z) = \sum_{t=1}^n l(\boldsymbol{\theta}; y_t, z)$ com

$$l(\boldsymbol{\theta}; y_t, z) = -z\Lambda_t(y_t, \nu_t, \delta_t) - \log \delta_t + \log \nu_t - \nu_t \log y_t - \log(1 - y_t^{\nu_t}) - \log y_t. \quad (14)$$

que é linear em z e, ainda, os coeficientes da regressão podem ser estimados de forma independente do parâmetro α . A distribuição de $(Z | \mathbf{y}, \boldsymbol{\theta}, \alpha)$ é Gamma-Tilted- Stable (Qu et al., 2021) com densidade condicional (vê Proposição A.4 no Apêndice A) dada por:

$$f_Z(z | \mathbf{y}, \boldsymbol{\theta}, \alpha) = \frac{z^n e^{-z\Lambda_n(\mathbf{y}, \boldsymbol{\theta})} g_\alpha(z)}{G(n, \Lambda_n(\mathbf{y}, \boldsymbol{\theta}))}, \quad z > 0. \quad (15)$$

de forma que

$$\tilde{z} = \mathbb{E}(Z | \mathbf{y}, \boldsymbol{\theta}, \alpha) = -\frac{\sum_{j=1}^{n+1} q_{n+1,j} \alpha^j v^{j\alpha}}{\Lambda_n(\mathbf{y}, \boldsymbol{\theta}) \sum_{j=1}^n q_{n,j} \alpha^j v^{j\alpha}}, \quad (16)$$

onde $\Lambda_n(\mathbf{y}, \boldsymbol{\theta}) = \sum_{t=1}^n \Lambda_t(y_t, \nu_t, \delta_t) > 0$ e $G(n, \Lambda_\cdot) = \int_0^\infty z^n e^{-z\Lambda_\cdot} g_\alpha(z) dz$, e

$$q_{n,j} = \begin{cases} 0 & \text{se } n = 1, j = 0 \\ q_{n-1,j-1} + (n-1-j\alpha)q_{n-1,j} & \text{se } j = 1, 2, \dots, n-1 \\ 1 & \text{se } n = j. \end{cases}$$

Para um valor fixado $\tilde{\alpha}$ a $\mathbb{E}_{Z|\mathbf{y}, \boldsymbol{\theta}, \tilde{\alpha}}[l(\boldsymbol{\theta}; \mathbf{y}, z)] = l(\boldsymbol{\theta}; \mathbf{y}, \tilde{z})$. Então, o algoritmo iterativo **EM** para $\boldsymbol{\theta}$ na $(k+1)$ -ésima iteração é:

- **E-step**: Computamos $Q(\boldsymbol{\theta}; \boldsymbol{\theta}^{(k)}) = l(\boldsymbol{\theta}; \mathbf{y}, \tilde{z}^{(k)})$;
- **M-step**: Selecione $\boldsymbol{\theta}^{(k+1)}$ que maximiza $Q(\boldsymbol{\theta}; \boldsymbol{\theta}^{(k)})$.

Os passos **E** e **M** são alternados repetidamente até a convergência. O **M**-step requer o cômputo do vetor escore e da matriz hessiana para os dados completos. O vetor Escore, $S_c(\boldsymbol{\theta}) = (S_c(\beta), S_c(\gamma))^\top$, para os dados completos é obtido pela soma das derivadas parciais da expressão (14). Seja $\mathbf{D}$ o operador diagonal de um vetor tal que $\mathbf{D}_a = \text{diag}(\mathbf{a})$ onde $\mathbf{a} = (a_1, a_2, \dots, a_n)^\top$. Então,

$$\begin{aligned} S(\beta_p; y_t, z) &= \frac{\partial}{\partial \beta_p} l(\boldsymbol{\theta}; \mathbf{y}, z) = \sum_{t=1}^n \frac{\partial}{\partial \nu_t} l(\boldsymbol{\theta}; y_t, z) \times \frac{\partial \nu_t}{\partial \beta_p} \\ &= \sum_{t=1}^n \underbrace{\left(-\frac{z}{\delta_t} \frac{\nu_t y_t^{\nu_t} \log y_t^{\nu_t}}{(1-y_t^{\nu_t})} + \frac{\nu_t \log y_t^{\nu_t}}{(1-y_t^{\nu_t})} + \nu_t \right)}_{b_t} \frac{1}{\nu_t} X_{t,p}, \end{aligned}$$

e logo $S_c(\beta) = \mathbf{X}^\top \mathbf{D}_\nu^{-1} \mathbf{b}$, no qual $\mathbf{b} = (b_1, b_2, \dots, b_n)^\top$. Para o vetor γ ,

$$\begin{aligned} S(\gamma_k; y_t, z) &= \frac{\partial}{\partial \gamma_k} l(\boldsymbol{\theta}; \mathbf{y}, z) = \sum_{t=1}^n \frac{\partial}{\partial \delta_t} l(\boldsymbol{\theta}; y_t, z) \times \frac{\partial \delta_t}{\partial \gamma_k} = \sum_{t=1}^n \left(-\frac{z}{\delta_t} \log(1-y_t^{\nu_t}) - 1 \right) W_{t,k} \\ &= \sum_{t=1}^n \delta_t^{-1} (\Lambda_t(y_t, \nu_t, z^{-1}) - \delta_t) W_{t,k}, \end{aligned}$$

de forma que $S_c(\gamma) = \mathbf{W}^\top \mathbf{D}_\delta^{-1} (\Lambda_z - \boldsymbol{\delta})$. Em que Λ_z é a função de excedência acumulada do processo aplicada no ponto $\delta_t = z^{-1}$. As derivadas parciais de segunda ordem são:

$$\begin{aligned} \mathcal{I}_{(\beta_p, \gamma_k)} &= \frac{\partial^2}{\partial \beta_p \partial \gamma_k} l(\boldsymbol{\theta}; \mathbf{y}, z) = \sum_{t=1}^n \frac{\partial}{\partial \delta_t} S(\beta_p; y_t, z) \times \frac{\partial \delta_t}{\partial \gamma_k} \\ &= z \sum_{t=1}^n X_{t,p} \underbrace{\left(\frac{\nu_t y_t^{\nu_t} \log y_t}{\delta_t (1-y_t^{\nu_t})} \right)}_{c_t} W_{t,k} = z \sum_{t=1}^n X_{t,p} c_t W_{t,k}. \end{aligned}$$

Logo, $\mathcal{I}_{(\beta, \gamma)} = z \mathbf{X}^\top \mathbf{D}_c \mathbf{W}$. Para o vetor β , a Hessiana é da forma $\mathcal{I}_{(\beta, \beta)} = \mathbf{X}^\top \mathbf{D}_d \mathbf{X}$ onde, $d_t = \left(-\frac{z \log y_t (y_t^{\nu_t} - y_t^{2\nu_t} + \nu_t y_t^{\nu_t} \log y_t)}{\delta_t (1-y_t^{\nu_t})^2} + \frac{1-y_t^{\nu_t} + \nu_t y_t^{\nu_t} \log y_t}{(1-y_t^{\nu_t})^2} \right) \nu_t$. Por fim,

$$\mathcal{I}_{(\gamma_j, \gamma_k)} = \frac{\partial^2}{\partial \gamma_j \partial \gamma_k} l(\boldsymbol{\theta}; \mathbf{y}, z) = \sum_{t=1}^n \frac{\partial}{\partial \delta_t} S(\gamma_k; y_t, z) \times \frac{\partial \delta_t}{\partial \gamma_j} = z \sum_{t=1}^n W_{t,j} \Lambda_t(y_t, \nu_t, \delta_t) W_{t,k}$$

de modo que $\mathcal{I}_{(\gamma, \gamma)} = z \mathbf{W}^\top \mathbf{D}_\Lambda \mathbf{W}$. Assim, a matriz Hessiana ou matriz de Informação observada do modelo completo é dada por,

$$\mathcal{I}_c(\boldsymbol{\theta}) = \begin{pmatrix} \mathcal{I}_{(\beta, \beta)} & \mathcal{I}_{(\beta, \gamma)} \\ \mathcal{I}_{(\gamma, \beta)} & \mathcal{I}_{(\gamma, \gamma)} \end{pmatrix}. \quad (17)$$

Agora, seja $\tilde{S}_c(\boldsymbol{\theta}) = S_c(\boldsymbol{\theta})|_{z=\tilde{z}}$ e $\tilde{\mathcal{I}}_c(\boldsymbol{\theta}) = \mathcal{I}_c(\boldsymbol{\theta})|_{z=\tilde{z}}$. Então, iterativamente a otimização no passo **M** pode ser dada por: $\boldsymbol{\theta}^{(k+1)} = \boldsymbol{\theta}^{(k)} - \tilde{\mathcal{I}}_c^{-1}(\boldsymbol{\theta}^{(k)}) \tilde{S}_c(\boldsymbol{\theta}^{(k)})$.

O algoritmo EM não fornece automaticamente uma estimativa da matriz de covariância do Estimador de Máxima verossimilhança de $\boldsymbol{\theta}$. No entanto, no modelo proposto essa estimativa pode ser obtida através do princípio da informação faltante (Orchard and Woodbury, 1972). Para tanto, seja $\mathcal{I}(\boldsymbol{\theta}; \mathbf{y})$ a matriz de informação para os dados observados com log-verossimilhança dada em (7), $\mathcal{I}_m(\boldsymbol{\theta}; \mathbf{y})$ a matriz de informação para os dados faltantes (missing) ou fator ambiental Z e $\mathcal{I}_c(\boldsymbol{\theta}; \mathbf{y})$ a matriz de informação para os dados completos $(\mathbf{y}, z)$. Então, pelo princípio da informação faltante, podemos escrever $\mathcal{I}(\boldsymbol{\theta}; \mathbf{y}) = \mathcal{I}_c(\boldsymbol{\theta}; \mathbf{y}) - \mathcal{I}_m(\boldsymbol{\theta}; \mathbf{y})$, de modo que a matriz de informação estimada para os dados observados simplifica em (veja capítulo IV de McLachlan and Krishnan (2008)),

$$\mathcal{I}(\hat{\boldsymbol{\theta}}; \mathbf{y}) = \tilde{\mathcal{I}}_c(\hat{\boldsymbol{\theta}}) - \left[\mathbb{E}_{Z|\mathbf{y}, \boldsymbol{\theta}, \tilde{\alpha}} \left(S_c(\boldsymbol{\theta}) S_c^\top(\boldsymbol{\theta}) \right) \right]_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}}. \quad (18)$$

Sob as usuais condições de regularidade para a estimação por Máxima Verossimilhança, quando o tamanho da amostra da série é grande,

$$\begin{pmatrix} \hat{\beta} \\ \hat{\gamma} \end{pmatrix} \sim \text{asymptotically } \mathcal{N}_{p+k} \left[\begin{pmatrix} \beta \\ \gamma \end{pmatrix}, \mathcal{I}^{-1}(\boldsymbol{\theta}; \mathbf{y}) \right], \quad (19)$$

onde $\mathcal{N}_d$ denota a distribuição normal d -dimensional. Se $\mathcal{I}(\hat{\boldsymbol{\theta}}; \mathbf{y})$ é positiva definida, então, a utilizamos como estimador para $\mathcal{I}(\boldsymbol{\theta}; \mathbf{y})$. As quantidades $\tilde{I}_c(\boldsymbol{\theta})$, $S_c(\boldsymbol{\theta})$, são dadas na Seção 3.1.1. O valor esperado na equação (18) é com respeito a distribuição condicional de $(Z|\mathbf{y}, \boldsymbol{\theta}, \tilde{\alpha})$ e dependem do primeiro e segundo momento dessa distribuição que tem densidade dada por (15). O primeiro momento é dada na equação (16) O segundo momento é (Proposição A.3 no Apêndice A)

$$\tilde{z}^2 = \mathbb{E}(Z^2|\mathbf{y}, \boldsymbol{\theta}, \alpha) = \frac{\sum_{j=1}^{n+2} q_{n+2,j} \alpha^j v^{j\alpha}}{\Lambda_*(\mathbf{y}, \boldsymbol{\theta})^2 \sum_{j=1}^n q_{n,j} \alpha^j v^{j\alpha}}. \quad (20)$$

Se n cresce, os cálculos destes momentos podem ser computacionalmente complexos e, neste caso, o problema pode ser contornado gerando uma amostra $(z_1, z_2, \dots, z_J)$ da distribuição de $(Z|\mathbf{y}, \hat{\boldsymbol{\theta}}, \tilde{\alpha})$ e aproximar $\tilde{z} = J^{-1} \sum_{j=1}^J z_j$ e $\tilde{z}^2 = J^{-1} \sum_{j=1}^J z_j^2$. Um algoritmo eficiente para gerar dados da distribuição $(Z|\mathbf{y}, \boldsymbol{\theta}, \alpha) \sim \text{Gamma-Tilted-Stable}$ é apresentado em Qu et al. (2021).

3.1.2. Estimação do Parâmetro α

Na estimação do parâmetro α , notamos na Seção 3.1.1 que a log-verossimilhança $l(\alpha; z) = n \log z + \log g_\alpha(z)$ é função de um único valor não observado z . Esta função pode ser maximizada substituindo z por $\tilde{z}$. No entanto, esse estimador não será consis-

tente pois é obtido através de uma única observação. Uma estratégia para contornar este problema são descritas em dois passos. **Passo 1)** Notamos pelas expressões (2) e (3) que a função de excedência condicional conjunta é da forma $\bar{F}(\mathbf{y}|z) = [\bar{F}_0(\mathbf{y})]^z$, onde $\bar{F}_0(\mathbf{y}) = \prod_{t=1}^n (1 - y_t^{\nu_t})^{1/\delta_t}$ é o produto de n variáveis aleatórias independentes com distribuição $K(\nu_t, 1/\delta_t)$. Isto é, $\bar{F}_0(\mathbf{y})$ é uma espécie de distribuição baseline para o processo observado $\mathbf{y} = (y_1, \dots, y_n)$. Então, a maximização da verossimilhança dessa distribuição baseline deve nos fornecer bons estimadores iniciais $(\beta_0, \gamma_0)^\top$ para o vetor $(\beta, \gamma)^\top$. **Passo 2)** O parâmetro α não depende da estrutura de dependência do modelo, apenas controla a dependência na série de extremos $\mathbf{y} = (y_1, \dots, y_n)$. Então, estimamos α por $\hat{\alpha}$ maximizando a log-verossimilhança condicional, $l_{cd}(\boldsymbol{\theta}, \alpha; \mathbf{y})$, dada em (9), com $\boldsymbol{\theta} = (\beta, \gamma)^\top$ inicial em $\boldsymbol{\theta}_0 = (\beta_0, \gamma_0)^\top$ obtido no **Passo 1**. Como a derivada dessa log-verossimilhança é um processo *martingale* integrável com média zero; veja capítulo 6 de (Hall and Heyde, 1980). Segue que $\hat{\alpha}$ é um estimador consistente e assintoticamente $\mathcal{N}(\alpha, I^{-1}(\alpha))$ onde $I(\alpha)$ é a Informação observada que pode ser facilmente obtida pela função de otimização **optim** disponível em R Core Team (2023). Essa estratégia de estimação é avaliada através de estudos numéricos descritos na Seção 4.

3.1.3. Algoritmo para Estimação dos Parâmetros

Simplificando o processo de estimação apresentado nas seções 3.1.1 e 3.1.2 o Algoritmo iterativo para obter $(\hat{\alpha}, \hat{\beta}, \hat{\gamma})$ é da seguinte forma:

- **Passo 1:** Maximizamos a distribuição baseline na Seção 3.1.2 e obtemos $(\beta_0, \gamma_0)^\top$;
- **Passo 2:** Obtemos $\hat{\alpha}$, maximizando $l_{cd}(\boldsymbol{\theta}, \alpha; \mathbf{y})$ com $\boldsymbol{\theta} = (\beta_0, \gamma_0)^\top$ inicial;
- **Passo 3:** Com a distribuição assintótica de $\hat{\alpha}$ testamos a hipótese $H_0 : \alpha = 1$. Se não rejeitamos H_0 , paramos o processo e o modelo **PMS**(α) não é adequado. Caso contrário, seguimos para o **Passo 4**;
- **Passo 4:** Com $\hat{\alpha}$ computamos $(\tilde{z}, \tilde{z}^2)$ e obtemos os estimadores $\hat{\boldsymbol{\theta}} = (\hat{\beta}, \hat{\gamma})$ via o E-M.
- **Passo 5:** Com $\tilde{z}^2$ e $\hat{\boldsymbol{\theta}}$ computamos $\mathcal{I}(\hat{\boldsymbol{\theta}}; \mathbf{y})$ na equação (18) e obtemos a distribuição assintótica de $(\hat{\beta}, \hat{\gamma})^\top$ na expressão (19).

Este algoritmo está implementado na linguagem R Core Team (2023) e pode ser solicitado diretamente dos autores.

3.2. Critério de Seleção do Modelo

A comparação e seleção entre os vários modelos concorrentes testados usando o **PMS**(α), pode ser realizado através da verossimilhança $l(\boldsymbol{\theta}, \alpha; \mathbf{y})$ em (7) e aplicando sobre ela os critérios de seleção disponíveis na literatura. A ideia básica é selecionar um modelo que seja bem ajustado, com fácil interpretabilidade e tenha um número pequeno de parâmetros. Dois critérios de seleção de modelos bastante utilizados são o *Akaike Information Criterion* (AIC) (Akaike, 1974) e o *Bayesian Information Criterion* (BIC) (Schwarz, 1978), que podem ser computados para o **PMS**(α) por:

$$AIC = -2l(\hat{\boldsymbol{\theta}}, \hat{\alpha}; \mathbf{y}) + 2(p + k + 1) \quad \text{and} \quad BIC = -2l(\hat{\boldsymbol{\theta}}, \hat{\alpha}; \mathbf{y}) + (p + k + 1)\log(n),$$

onde a log-verossimilhança dada em (7) é:

$$l(\boldsymbol{\theta}, \alpha; \mathbf{y}) = -\Lambda_n^\alpha(\mathbf{y}; \boldsymbol{\theta}) + \log(D_{1,2,\dots,n}(\mathbf{y}; \boldsymbol{\theta}, \alpha)).$$

Selecionamos o modelo ajustado identificado pelo menor valor do AIC ou BIC.

3.3. Análise dos Resíduos

A análise de resíduos é uma importante ferramenta para verificar se as suposições do modelo são válidas e também usada para identificar observações mal ajustadas. Para o proposto modelo $\mathbf{PMS}(\alpha)$, consideramos o resíduo quantílico (Dunn and Smyth, 1996) especificado por: $r_t^{(q)} = \Phi^{-1}(1 - \mathbb{P}(Y_t > y_t | Y_{t-1} = y_{t-1}))$, onde $\mathbb{P}(Y_t > y_t | Y_{t-1} = y_{t-1})$ é a excedência condicional de primeira ordem dada pela expressão (11) que também pode ser usada para produzir resíduos do modelo (Crowder, 1989). Nesse caso, $\Phi^{-1}(\cdot)$ denota a função quantílica da distribuição normal padrão. Se o modelo ajustado está bem especificado esses resíduos devem ter aproximadamente uma distribuição normal com média zero e variância unitária e seu gráfico não deve exibir padrão perceptível.

3.4. Previsão

Considere o problema de prever um valor para a resposta observada h passos à frente. Previsões de observações futuras para o modelo $\mathbf{PMS}(\alpha)$ podem ser obtidas aplicando as propriedades descritas em **P3**.

Para a previsão, dois métodos foram propostos. Considere h_0 o horizonte de previsão. Além disso, assuma que as covariáveis $\mathbf{x}_t$ e $\mathbf{w}_t$ estão disponíveis para $t = n + 1, \dots, n + h_0$. O Primeiro método consiste em calcular a inversa do risco (*Reverse Risk*):

- Passo 1: Considere $t = 1, \dots, n$. Obtenha os preditores de tendência temporal e da dispersão temporal a partir de $\hat{\nu}_t = \exp(\mathbf{x}_t^T \hat{\beta})$ e $\hat{\delta}_t = \exp(\mathbf{w}_t^T \hat{\gamma})$, onde os valores para $\hat{\beta}$ e $\hat{\gamma}$ são as estimativas obtidas;
- Passo 2: Considere $h = 1, \dots, h_0$. Obtenha os preditores de tendência temporal e da dispersão temporal a partir de $\hat{\nu}_{n+h} = \exp(\mathbf{x}_{n+h}^T \hat{\beta})$ e $\hat{\delta}_{n+h} = \exp(\mathbf{w}_{n+h}^T \hat{\gamma})$;
- Passo 3: Usando **P3**, calcule $\mathbb{E}(\Lambda_{n+h} | Y_{n+h-1} = y_{n+h-1}) = \hat{\alpha}^{-1} \Lambda_{n+h-1}^{1-\hat{\alpha}}$ e estime

$$\Lambda_{t+h}(y_{t+h}, \hat{\nu}_{t+h}, \hat{\delta}_{t+h}) = \mathbb{E}(\Lambda_{n+h} | Y_{n+h-1} = y_{n+h-1})$$

$$\text{para obter } \hat{Y}_{n+h} = (1 - \exp\{\mathbb{E}(\Lambda_{n+h} | Y_{n+h-1} = y_{n+h-1})\}^{\hat{\delta}_{t+h}})^{1/\hat{\nu}_{t+h}}.$$

Para o segundo método, considere os passos 1 e 2 apresentados acima. O seguinte método consiste em calcular a mediana condicional (*Condition Median*) das observações. Portanto, segue que:

- Passo 3: Considere a função de probabilidade apresentada em (11). Obtenha a mediana $\hat{y}_{n+h}$ a partir da solução numérica da equação

$$1 - 0.5 - \mathbb{P}(Y_{n+h} > y_{n+h} | Y_{n+h-1} = y_{n+h-1}) = 0.$$

Note que uma estimativa para qualquer quantil condicional de primeira ordem $q_t|y_{t-1}$ pode ser obtida substituindo 0.5 por $q_t|y_{t-1}$, onde $q_t|y_{t-1} = \mathbb{P}(Y_t < y_t | Y_{t-1} = y_{t-1})$.

4. Avaliação Numérica

Nesta seção apresentamos um estudo com dados simulados para avaliar a performance de dois métodos de estimação no modelo $\mathbf{PMS}(\alpha)$. O método 1 consiste em utilizar a verossimilhança condicional de primeira ordem apresentado na expressão (9) e o método 2 consiste na utilização do algoritmo **EM** apresentado na seção 3.1.1. Para isso, foram simuladas 3000 réplicas de 3 modelos com covariáveis de tendência e sazonalidade apresetados na Tabela 1. Cada modelo foi combinado com cinco valores de $\alpha = (0.35, 0.5, 0.65, 0.85, 0.95)$. Foram considerados três tamanhos de série: $n = 84, 144$ e $n = 168$. Essas quantidades representam, respectivamente, 6, 12 e 14 anos de análise, totalizando 135 mil séries geradas. Considerando $t = 1, \dots, n$, as covaráveis de tendência temporal, associadas ao modelo I são $\mathbf{x}_t = (\sin(2\pi t)/12, \cos(2\pi t)/12)^T$ com $\beta = (\beta_1, \beta_2)^T$ e as covariáveis associadas a dispersão temporal é uma variável dummy indicando 0 para o primeiro semestre do ano e 1 para o segundo semestre, ou seja, $\mathbf{w}_t = (1, s_{2,t})^T$ com $\gamma = (\gamma_0, \gamma_1)^T$. Para o modelo II, $\mathbf{x}_t = (1, \sin(2\pi t)/12, \cos(2\pi t)/12)^T$ com $\beta = (\beta_0, \beta_1, \beta_2)^T$ e $\mathbf{w}_t = (s_{1,t}, s_{2,t})^T$, onde $s_{i,t} = 1$ se t pertence ao semestre i e $s_{i,t} = 0$, caso contrário, com $\gamma = (\gamma_1, \gamma_2)^T$ e $i \in \{1, 2\}$. No modelo III, $\mathbf{x}_t = (1, \log(t))^T$ com $\beta = (\beta_0, \beta_1)^T$ e $\mathbf{w}_t$ é o mesmo do Modelo I. A idéia de segundo parâmetro do vetor γ em todos os modelos é capturar o aumento (ou diminuição) na dispersão temporal ao mudarmos do estado 1 (primeiro semestre) para o estado 2 (segundo semestre). Para simular uma série de tamanho n com dependência de primeira ordem no modelo $\mathbf{PMS}(\alpha)$, consideramos a função de excedência condicional apresentada em 11 e o seguinte algoritmo:

- Passo 0: Obtenha os preditores de tendência temporal e da dispersão temporal a partir de $\nu_t = \exp(\mathbf{x}_t^T \beta)$ e $\delta_t = \exp(\mathbf{w}_t^T \gamma)$, onde os valores para β e γ são apresentados na Tabela 1.
- Passo 1: Gere $u_1 \sim U(0, 1)$ e obtenha y_1 pela transformação na expressão (4),

$$y_1 = \bar{F}_{Y_1}^{-1}(1 - u_1)$$

- Passo 2: Gere $u_2 \sim U(0, 1)$ e obtenha y_2 a partir da solução da equação

$$1 - u_2 - \mathbb{P}(Y_2 > y_2 | Y_1 = y_1) = 0$$

- Passo 3: Gere $u_3 \sim U(0, 1)$ e obtenha y_3 a partir da solução da equação

$$1 - u_3 - \mathbb{P}(Y_3 > y_3 | Y_2 = y_2) = 0$$

onde y_2 é obtido no Passo 2.

⋮

- Passo n: Gere $u_n \sim U(0, 1)$ e obtenha y_n a partir da solução da equação

$$1 - u_n - \mathbb{P}(Y_n > y_n | Y_{n-1} = y_{n-1}) = 0$$

onde y_{n-1} é obtido no Passo n-1.

Tabela 1. Valores reais para os parâmetros dos modelos I, II e III

Model	Predictor	Variable	Parameter	Real values
I	Trend	sent	β_1	0.10
		cost	β_2	-0.10
	Dispersion	(Intercept)	γ_0	-3.50
		semester2	γ_1	-0.25
II	Trend	(Intercept)	β_0	2.50
		sent	β_1	0.10
		cost	β_2	-0.10
	Dispersion	semester1	γ_1	-3.50
		semester2	γ_2	-3.75
III	Trend	(Intercept)	β_0	2.00
		$\log(t)$	β_1	0.20
	Dispersion	(Intercept)	γ_0	-3.50
		semester2	γ_1	-0.25

O vetor gerado $\mathbf{y} = (y_1, y_2, \dots, y_n)$ é uma amostra aproximada de primeira de ordem para o modelo $\mathbf{PMS}(\alpha)$. Amostras com dependência de k -ésima ordem podem ser geradas utilizando a *probability integral transformation multivariate* descrita em Rosenblatt (1952). Os códigos foram escritos na linguagem de programação R Core Team (2023) e estão disponíveis mediante solicitação ao primeiro autor. Na Figura B1 são apresentadas exemplos das séries simuladas dos modelos para diferentes combinações do α . Note que para o Modelo I (sem β_0) a série apresenta resultados próximos de zero com picos variando de acordo com o valor do α , ou seja, quando $\alpha \rightarrow 0.95$ os tamanhos dos picos diminuíam e consequentemente a variabilidade dos dados. Para o modelo II (sem γ_0) os dados apresentam flutuações em torno de 0.75 com a variabilidade diminuindo a medida que $\alpha \rightarrow 0.95$. No modelo III (tendência) também foi observado a diminuição da variabilidade dos dados quando $\alpha \rightarrow 0.95$. De forma geral foi observado que quando $\alpha \rightarrow 0.35$ os dados tendem a ficar mais próximo dos extremos e que o modelo sem β_0 apresenta observações próximas de zero com picos variando de acordo com o valor do α . Na prática, em dados meteorológicos que é o foco de nossa pesquisa, os valores extremos se comportam de forma similar ao modelo II.

Nas Tabelas B1 a B15 são apresentadas as estimativas dos parâmetros. As métricas de avaliações utilizadas foram: *mean*, *median*, *relative bias* e o *mean square error* (MSE). Para as estimativas do α , obteve-se melhores resultados para todos os modelos com $\alpha > 0.35$ e considerando todas combinações de n . Para $\alpha = 0.35$ o valor do menor e maior viés relativo são -0.187 e -0.506 , respectivamente. De forma geral, em **relação ao método 1**, as estimativas para os parâmetros dos preditores de tendência temporal e dispesão temporal apresentaram bons resultados para as diferentes combinações de n e α . No entanto, destacamos algumas observações para os parâmetros associados a dispesão temporal. Quando comparado o γ_0 dos modelos I (sem β_0) e III (tendência), foi observado que quando $\alpha \rightarrow 0.35$ o viés diminui para o Modelo III e quando $\alpha \rightarrow 0.95$ foi observado que quando $\alpha \rightarrow 0.35$ o viés diminui para o Modelo I. Para o γ_1 (semester1-delta) do modelo II (sem γ_0) foi observado que o viés diminui quando $\alpha \rightarrow 0.35$. Para o γ_1 (semester2-delta) do modelo I foi observado que para $n = 84$ e $n = 144$, o viés diminui quando $\alpha \rightarrow 0.95$ e para $n = 168$ o menor viés foi em $\alpha = 0.65$. Para o γ_1 (semester2-delta) do modelo III todas as combinações apresentaram boas estimativas e o maior viés (-0.102) está em $n = 168$ e $\alpha = 0.35$. Para o γ_2 (semester2-delta) do modelo II foi observado que o viés diminui quando $\alpha \rightarrow 0.35$. De forma geral, em **relação ao método 2**, ocorre a diminuição do viés para as estimativas dos parâmetros dos preditores de tendência a medida que $\alpha \rightarrow 0.95$. Além disso, foi observado que os modelos II (sem

γ_0) e III (tendência) apresentaram melhores estimativas aos parâmetros associados às covariáveis. Em relação aos parâmetros da dispersão temporal. Quando comparado o γ_0 dos modelos I (sem β_0) e III (tendência), foi observado que quando $\alpha \rightarrow 0.95$ o método apresenta melhores estimativas para o modelo I. Esse comportamento acontece no modelo III apenas para $n = 84$, para os demais valores de n os menores vícios ocorrem em $\alpha = 0.35$ e $\alpha = 0.95$. Para o γ_1 (semester1-delta) do modelo II (sem γ_0) foi observado que o viés diminui quando $n = 84$ e $\alpha \rightarrow 0.95$. Para $n = 144$ e $n = 168$ as estimativas melhoram para $\alpha = 0.95$ e $\alpha = 0.35$. Para o γ_1 (semester2-delta) dos modelos I e III foram observados que o viés diminui quando $\alpha \rightarrow 0.95$. Para o γ_2 (semester2-delta) do modelo II foi observado que o viés diminui quando $n = 84$ e $\alpha \rightarrow 0.95$. Para $n = 144$ e $n = 168$ as estimativas melhoram para $\alpha = 0.95$ e $\alpha = 0.35$. Após a análise geral dos resultados, observou-se que, de forma geral, **método 1** apresenta melhores estimativas. Mas, isso era esperado devido as amostras serem geradas com dependência de primeira ordem. O viés das estimativas não necessariamente diminui com o aumento do tamanho da série e que o valor do α pode influenciar nos resultados estimados. Além disso, observou-se melhores estimativas quando $\alpha \geq 0.65$ para o **método 2**. Neste caso, o **método 2** foi competitivo ao **método 1** e o importante é que ele pode ser usado independente da verdadeira ordem do modelo nos dados. As Figuras B2, B3 e B4 mostram as distribuições dos estimadores para os três modelos com $\alpha = 0.80$.

5. Aplicação

A série temporal analisada consiste na mínima umidade relativa do ar mensal observada na cidade de Manaus no período de janeiro de 2009 a dezembro de 2020. Nesse caso, temos inicialmente o extremo diário que é a soma do maior valor diário com o menor valor diário e o resultado é dividido por 2. Esses dados correspondem ao tamanho de amostra $n = 144$ blocos mensais e $y_t, t = 1, 2, \dots, n$ é o menor valor observado em cada bloco. O modelo **PMS**(α) proposto foi ajustado com as 132 primeiras observações e as 12 últimas foram separadas para comparações de previsões. Os dados estão disponíveis no site do Instituto Nacional de Meteorologia (INMET, <https://portal.inmet.gov.br/>) para estação meteorológica de Manaus nº 82331 operante a uma altitude de 61,25m. Na Figura 1 é apresentada a série, o comportamento sazonal, acf e pacf dos dados. A partir da análise da série observada, percebeu-se um suave decrescimento na mínima umidade no decorrer dos últimos meses, Figura 1a. Além disso, foi constatado que a umidade tende a apresentar maiores valores nos primeiros seis meses do ano, diminuindo nos meses de agosto, setembro e outubro; voltando a exibir um crescimento nos meses de novembro e dezembro, Figura 1b. Fato importante é que a umidade é um indicador de conforto térmico que expressa o quão agradável ou desagradável são as condições térmicas do lugar e, em Manaus a presença de vegetação é de grande importância na variação desse indicador que devido a urbanização da cidade tem variado bastante ao longo dos últimos anos (Mandú and Gomes, 2019). Assim, é razoável supor $Z_t = Z/\delta_t$, onde Z é um efeito ambiental de vegetação compartilhado por toda série com dispersão temporal δ_t devido ao efeito do processo de urbanização da cidade. Diante disso, foi proposto o modelo com $\mathbf{x}_t = (1, \log(temp), \sin(2\pi t/12), \cos(2\pi t/12))^T$, onde $temp$ é a temperatura extrema do bulbo seco observada no dia em que houve a menor umidade extrema diária. O $\log(temp)$ é adotado devido a grande variação térmica diária que ocorre em Manaus. Os vetores de parâmetros são da forma $\beta = (\beta_0, \beta_1, \beta_2, \beta_3)^T$, $\mathbf{w}_t = (1, \log(t))^T$ com

$\gamma = (\gamma_0, \gamma_1)^\top$. Nesse modelo, β_1 captura a variação na mínima umidade em função da log-variação térmica diária, o par (β_2, β_3) mede o efeito sazonal mensal através do harmônico simples e γ_1 representa a variação na dispersão temporal. Na Tabela

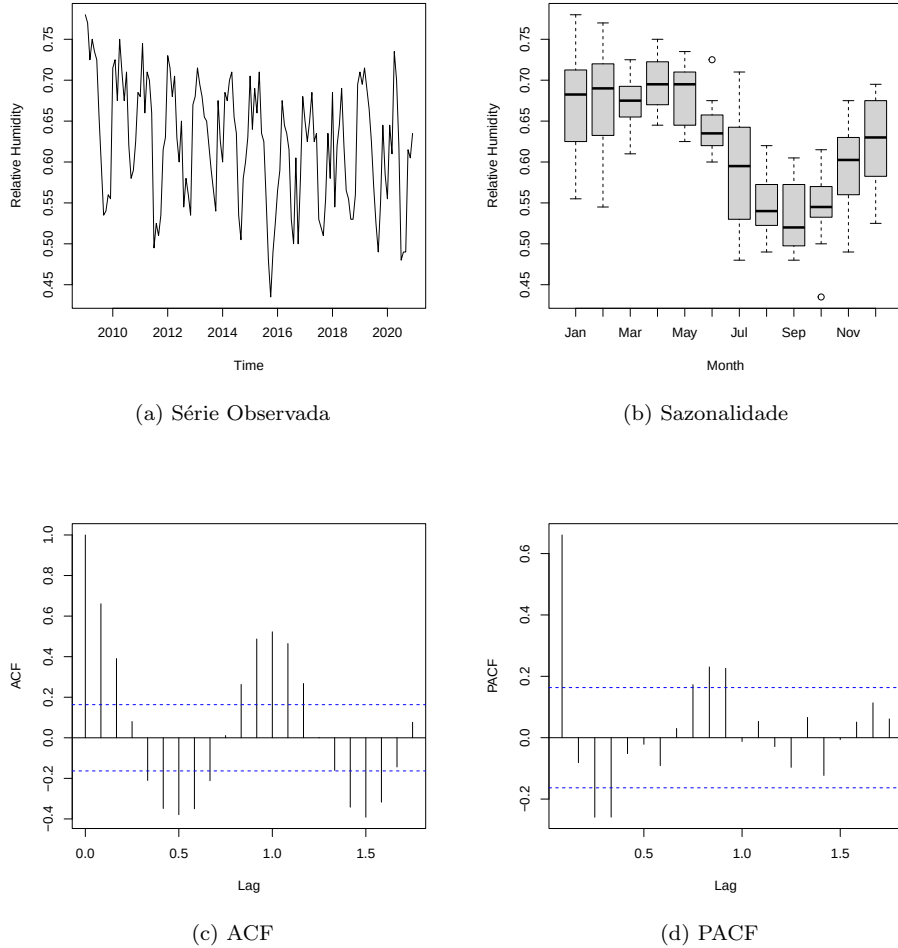


Figura 1. Série observada da umidade relativa de Manaus-AM

2 apresentamos as estimativas para os parâmetros e os valores do AIC e BIC estão descritos na Tabela 3. Observe que, comparando os dois métodos (1 e 2) de estimação, os valores das estimativas estão próximos, indicando possivelmente uma concordância na dependência de primeira ordem do modelo suavemente detectada por $\hat{\alpha} = 0.8352$. O método 1 estima um modelo mais parcimonioso. Mas o método 2 apresentou menores valores para o AIC, BIC e, ainda, uma melhor qualidade na previsão. Uma suave diminuição na dispersão temporal ou amplitude da mínima umidade foi detectada por $\hat{\gamma}_1$ e o efeito da variação térmica diária foi significativamente estimado por $\hat{\beta}_1$. Nas Figuras C1 e C2 são apresentados os diagnósticos dos resíduos para os métodos 1 e 2, respectivamente. Os resultados são similares, então os comentários a seguir servem para ambos. Considerando o método 2. Na Figura C2a é exibido o gráfico dos resíduos quantílicos em relação ao tempo. Foi observado que não há nenhum padrão evidente no decorrer do tempo e os resíduos apresentam o comportamento típico de ruído branco. A Figura C2b mostra o gráfico da distribuição normal em comparação

Tabela 2. Modelo ajustado para os dados de umidade relativa

Method	Variables	Parameters	Estimates	Standard Deviation	Interval 95%
Method 1	(Intercept)	β_0	14.8349	1.2259	(12.4443, 17.2255)
	log(temp)	β_1	-3.3549	0.3555	(-4.0481, -2.6617)
	sent	β_2	0.1174	0.0221	(0.0743, 0.1604)
	cost	β_3	0.0268	0.0151	(-0.0026, 0.0562)
	(Intercept)	γ_0	-8.5413	0.9311	(-10.3569, -6.7257)
	log(t)	γ_1	-0.8213	0.1575	(-1.1285, -0.5141)
	-	α	0.8352	0.0636	(0.7112, 0.9591)
Method 2	(Intercept)	β_0	14.8230	1.1602	(12.5606, 17.0854)
	log(temp)	β_1	-3.3972	0.3366	(-4.0536, -2.7407)
	sent	β_2	0.1154	0.0194	(0.0776, 0.1532)
	cost	β_3	0.0283	0.0129	(0.0032, 0.0534)
	(Intercept)	γ_0	-7.2807	0.6122	(-8.4746, -6.0868)
	log(t)	γ_1	-0.7345	0.1136	(-0.9561, -0.5129)
	-	α	0.8352	0.0636	(0.7112, 0.9591)

Tabela 3. AIC e BIC do modelo ajustado

Method	AIC	BIC
Method 1	-534.1752	-513.9956
Method 2	-560.0014	-539.8218

com os quantis empíricos. É possível observar uma linha reta, indicando bom ajuste do modelo aos dados. Nas Figuras C2c e C2d são apresentadas a ACF e PACF, respectivamente. Essas medidas auxiliam na verificação visual da hipótese de ruído branco residual. Além disso, o teste de Ljung-Box (Ljung and Box, 1978) foi utilizado para testar a hipótese de independência dos resíduos. Para esse caso, o p-valor do teste foi de 0.2919 (0.3763 método 1) com o lag considerado de 20. Esses resultados indicam que o modelo proposto pode ser utilizado para realizar previsões e outras medidas.

As Figuras 2 e 3 exibem os valores preditos e a previsão do modelo para observações que não estão na amostra. Nesse caso, a previsão foi realizada para o ano de 2020 ($h = 12$). É possível observar que os valores ajustados e a previsão conseguiram capturar o padrão sazonal presente nos dados reais. O erro quadrático médio (MSE) e o erro percentual médio (MAPE) entre os valores observados e os valores previstos são apresentados na Tabela 4. Observou-se que mediana condicional (vê, Seção 3.4) estimada pelo método 2 apresentou melhor performance no processo de previsão. Além desses resultados tradicionais, o modelo proposto possibilita calcular probabilidades que são importantes sob o ponto de vista prático. Em particular, estamos interessado na probabilidade (de primeira ordem) de que a menor umidade extrema mensal no tempo t seja maior que 0.70 condicionado no valor observado no tempo $t - 1$. Outra medida é a probabilidade de excedência de segunda ordem. Nesse caso, estamos interessados na probabilidade de excedência conjunta de que a menor umidade extrema mensal nos tempos t e $t - 1$ seja maior que 0.70, condicionado no valor observado no tempo $t - 2$. Essas medidas são importante pois sabe-se que umidades relativas acima de 0.70 são prejudiciais para a saúde humana e sua ocorrência em tempos consecutivos requerem medidas de atenção para o surgimento de surtos de doenças causadas por fungos e bactérias. Na Figura 4 são apresentadas as séries extremas para as probabilidades de excedência condicionais de primeira e segunda ordem, respectivamente. Em ambas as probabilidades, foi observado que, de forma geral, há maior possibilidade de haver o evento “umidade extrema mensal maior que 0.70” no primeiro semestre. Para probabilidades de primeira ordem, as chances desse evento ocorrer ultrapassam

0.5 quando $t = \{\text{Feb, Mar}\}$ e decai no decorrer do ano. Para probabilidade de segunda ordem, há probabilidades maiores que 0.25 de ocorrer o evento “umidade extrema mensal maior que 0.70 por dois períodos consecutivos” quando considerados $t = \{\text{Feb, Mar}\}$. Importante observar que ambos os eventos têm poucas chances de ocorrerem quando considerado os meses do segundo semestre (período de poucas chuvas). Neste caso, o mais interessante é avaliar essas probabilidades para um valor menor que 0.3. Por fim, um estudo bootstrap foi realizado para analisar distribuição assintótica das estimativas. A partir dos parâmetros estimados pelo método 2, foram geradas 1000 simulações conforme os procedimentos apresentados na Seção 4. A partir da análise da Figura C3, percebe-se que as distribuições dos estimadores se aproximam da distribuição normal.

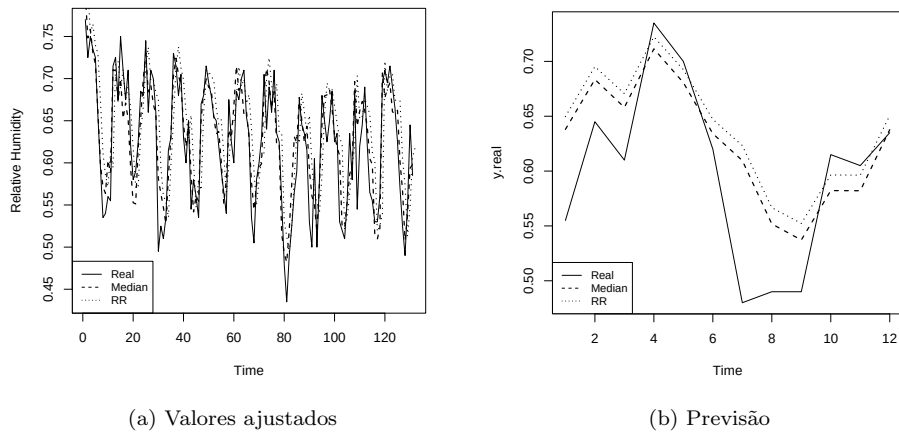


Figura 2. (a) Valores ajustados e (b) previsão para Método 1. RR é o risco reverso

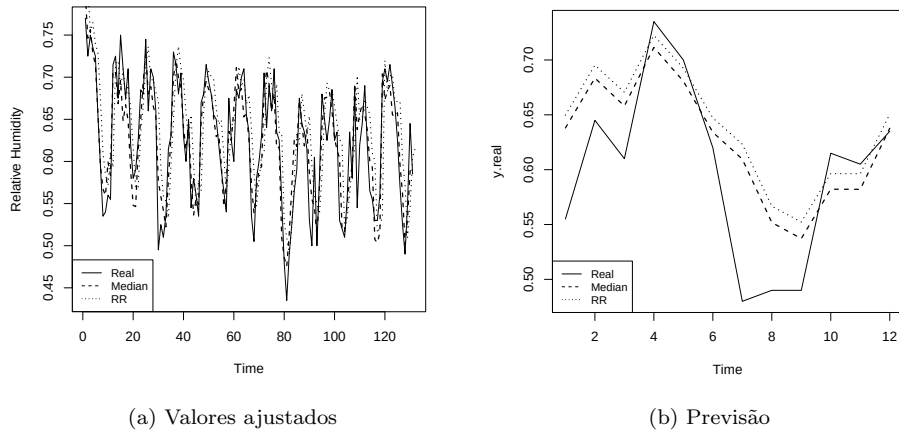
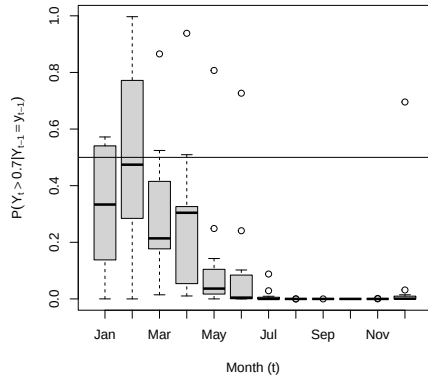


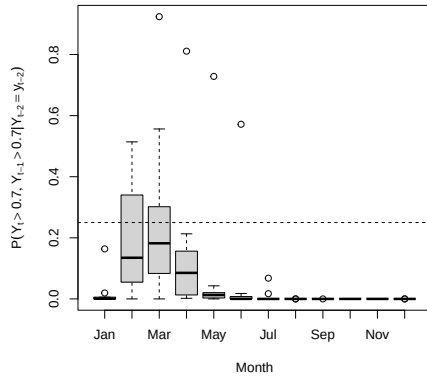
Figura 3. (a) Valores ajustados e (b) previsão para Método 2. RR é o risco reverso

Tabela 4. Comparação da performance de previsão entre os diferentes métodos

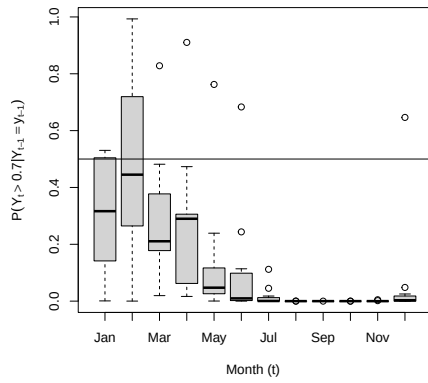
Method	Forecast type	MAPE	MSE
Method 1	Condition Median	0.0547314	0.0030228
	Reverse Risk	0.0771508	0.0039213
Method 2	Condition Median	0.0462608	0.0027683
	Reverse Risk	0.0721913	0.0036834



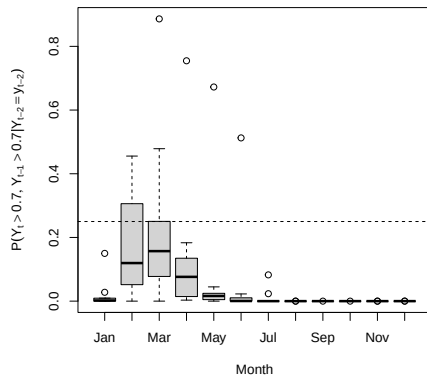
(a) Probabilidade de primeira ordem



(b) Probabilidade de segunda ordem



(c) Probabilidade de primeira ordem



(d) Probabilidade de segunda ordem

Figura 4. Probabilidade de primeira e segunda ordem para o método 1 (cima) e método 2 (embaixo)

6. Conclusão

Neste trabalho foi proposto um modelo de regressão *Power min-stable* - **PMS**(α) caracterizado por uma função de excedência multivariada para séries temporais de valores extremos com dependência $\alpha \in (0, 1)$. Essa função foi obtida a partir da marginalização da distribuição de uma série temporal com componentes condicionalmente independentes para extremos utilizando as propriedades da distribuição α -Stable positiva. O modelo é útil para a análise de mínimos extremos não estacionários observados ao longo do tempo com suporte no intervalo $(0,1)$, podendo ser generalizado para o intervalo (a,b) . Ele também acomoda estruturas de covariáveis nos parâmetros de tendência temporal ou sazonal e dispersão temporal possibilitando um melhor entendimento do comportamento da série. Uma expressão analítica para o extremograma da série é apresentado. Medidas de dependência entre os valores de excedência da série foram obtidas e, ainda, descrevemos algumas probabilidades condicionais que podem ser utilizadas para realizar inferências e obter resultados práticos relevantes. Em relação ao processo de estimação, dois métodos foram propostos. O primeiro é baseado na verossimilhança condicional de primeira ordem e o segundo utiliza um algoritmo EM que não considera a ordem de dependência da série e apresenta uma expressão analítica para a matriz de informação observada de modo que intervalos de confiança e testes de hipótese podem ser realizados através da distribuição assintótica dos estimadores. Um estudo de simulação Monte Carlo foi conduzido para avaliar o desempenho dos métodos de estimação e estes mostraram-se competitivos quando $\alpha \geq 0.65$. Seleção de modelos, análise de resíduos e formas de previsão também são discutidas. Por fim, sua utilidade, foi demonstrada através de uma aplicação aos dados de mínima umidade relativa do ar mensal, observada na cidade de Manaus no período de janeiro de 2009 a dezembro de 2020, onde identificamos meses com grandes probabilidades da mínima umidade ser maior que 0.70, podendo gerar vários problemas de saúde, (respiratórios, alergias, sinusites, asma, proliferação de fungos) e por isso requerem medidas de atenção para a saúde humana.

Referências

- Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19(6):716–723. doi: 10.1109/TAC.1974.1100705.
- Auld, G. and Papastathopoulos, I. (2023). Time series conditional extremes. Mar. 8, 2023. doi: 10.48550/arXiv.2303.04447.
- Balkema, A. A. and Haan, L. D. (1988). Almost sure continuity of stable moving average processes with index less than one. *The Annals of Probability*, 16(1):333–343.
- Bayer, F. M., Bayer, D. M., and Pumi, G. (2017). Kumaraswamy autoregressive moving average models for double bounded environmental data. *Journal of Hydrology*, 555:385–396.
- Bortot, P. and Gaetan, C. (2014). A latent process model for temporal extremes. *Scandinavian Journal of Statistics*, 41(3):606–621.
- Bücher, A. and Jennessen, T. (2022). Statistics for Heteroscedastic Time Series Extremes. Apr. 20, 2022. doi: <http://arxiv.org/abs/2204.09534>.
- Chavez-Demoulin, V. and Davison, A. (2012). Modelling the time series extremes. *Revstat Statistical Journal*, 10:109–133. doi: 10.57805/revstat.v10i1.113.
- Coles, S., Heffernan, J., and Tawn, J. (1999). Dependence measures for extreme value analyses. *Extremes*, 2:339–365.
- Crowder, M. (1989). A multivariate distribution with weibull connections. *Journal of the Royal Statistical Society: Series B*, 51(1):93–107.
- Davis, R. A. and Mikosch, T. (2009). The extremogram: A correlogram for extreme events. *Bernoulli*, 15(4):977–1009.
- Davis, R. A. and Resnick, S. I. (1989). Basic properties and prediction of max-arma processes. *Advances in Applied Probability*, 21(4):781–803.
- Davis, R. A. and Resnick, S. I. (1993). Prediction of stationary max-stable processes. *The Annals of Applied Probability*, 3(2):497–525.
- De Haan, L. (1984). A spectral representation for max-stable processes. *The annals of probability*, 12:1194–1204.
- De Haan, L. and Pickands III, J. (1986). Stationary min-stable stochastic processes. *Probability Theory and Related Fields*, 72(4):477–492.
- Dempster, A. P., Laird, N. M., and Rubin, D. B. (1977). Maximum likelihood from incomplete data via the em algorithm. *Journal of the royal statistical society: series B*, 39(1):1–22.
- Dunn, P. K. and Smyth, G. K. (1996). Randomized quantile residuals. *Journal of Computational and Graphical Statistics*, 5(3):236–244.
- Eastoe, E. (2007). *Statistical models for dependent and non-stationary extreme events*. PhD thesis, Lancaster University. Available at https://eprints.lancs.ac.uk/id/eprint/26855/1/thesis_final.pdf.
- Eastoe, E. F. and Tawn, J. A. (2009). Modelling non-stationary extremes with application to surface level ozone. *Journal of the Royal Statistical Society: Series C*, 58(1):25–45.
- Einmahl, J. H. J., de Haan, L., and Zhou, C. (2016). Statistics of heteroscedastic extremes. *Journal of the Royal Statistical Society. Series B*, 78(1):31–51.
- Fougères, A.-L., Nolan, J. P., and Rootzén, H. (2009). Models for dependent extremes using stable mixtures. *Scandinavian Journal of Statistics*, 36(1):42–59.
- Hall, P. and Heyde, C. (1980). *Martingale Limit Theory and its Application*. Probability and Mathematical Statistics: A Series of Monographs and Textbooks. Academic Press, New York, NY, 1 edition. doi: <https://doi.org/10.1016/>

- B978-0-12-319350-6.50004-9.
- Hall, P., Peng, L., and Yao, Q. (2002). Moving-maximum models for extrema of time series. *Journal of Statistical Planning and Inference*, 103(1):51–63.
- Hall, P. and Tajvidi, N. (2000). Nonparametric analysis of temporal trend when fitting parametric models to extreme-value data. *Statistical Science*, 15(2):153–167.
- Hougaard, P. (1986). A class of multivariate failure time distributions. *Biometrika*, 73(3):671–678.
- Küchenhoff, H. and Thamerus, M. (1996). Extreme value analysis of Munich air pollution data. *Environmental and Ecological Statistics*, 3(2):127–141.
- Kumaraswamy, P. (1980). A generalized probability density function for double-bounded random processes. *Journal of Hydrology*, 46(1):79–88.
- Kunin, B. I. (2004). Extreme value type distributions with bounded support. *International Journal of Pure and Applied Mathematics*, 12:499–514.
- Ljung, G. M. and Box, G. E. P. (1978). On a measure of lack of fit in time series models. *Biometrika*, 65(2):297–303.
- Mandú, T. B. and Gomes, A. C. D. S. (2019). Identificação de tendências no conforto térmico na região norte do brasil: Estudo de caso em manaus-am/identification of thermal comfort trends in the region north do brazil: case study in manaus-am. *Revista geonorte*, 10(34):63–81.
- Marshall, A. W. and Olkin, I. (1983). Domains of attraction of multivariate extreme value distributions. *The Annals of Probability*, 11(1):168–177.
- McLachlan, G. and Krishnan, T. (2008). *The EM algorithm and extensions*. Wiley series in probability and statistics. Wiley, Hoboken, NJ, 2 edition.
- Mhatre, N. and Cooley, D. (2020). Transformed-Linear Models for Time Series Extremes. Oct. 25, 2021. doi <http://arxiv.org/abs/2012.06705>.
- Nadarajah, S. and Kotz, S. (2006). The exponentiated type distributions. *Acta Applicandae Mathematica*, 92:97–111.
- Orchard, T. and Woodbury, M. A. (1972). *A Missing Information Principle: Theory and Applications*, volume 1, pages 697–716. University of California Press, Berkeley. doi:10.1525/9780520325883-036.
- Pancheva, E. (1994). *Extreme Value Limit Theory with Nonlinear Normalization*, pages 305–318. Springer US, Boston, MA.
- Pancheva, E. (1985). Limit theorems for extreme order statistics under nonlinear normalization. In Kalashnikov, V. V. and Zolotarev, V. M., editors, *Stability Problems for Stochastic Models*, pages 284–309, Berlin, Heidelberg. Springer Berlin Heidelberg.
- Peng, Z., Jiang, Q., and Nadarajah, S. (2015). A unified representation of p-max stable distributions. *MATH. REPORTS*, 17(3):277–293. Available at: http://imar.ro/journals/Mathematical_Reports/Pdfs/2015/3/3.pdf.
- Penrose, M. D. (1992). Semi-min-stable processes. *The Annals of Probability*, 20(3):1450–1463.
- Qu, Y., Dassios, A., and Zhao, H. (2021). Random variate generation for exponential and gamma tilted stable distributions. *ACM Transactions on Modeling and Computer Simulation (TOMACS)*, 31(4):1–21.
- R Core Team (2023). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. Available at: <https://www.R-project.org/>.
- Rocha, A. V. and Cribari-Neto, F. (2009). Beta autoregressive moving average models. *Test*, 18(3):529–545.
- Rosenblatt, M. (1952). Remarks on a multivariate transformation. *The annals of*

-
- mathematical statistics*, 23(3):470–472.
- Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics*, 6(2):461–464.
- Sebastian, N. and Gorenflo, R. (2016). Time series models associated with Mittag-Leffler type distributions and its properties. *Communications in Statistics - Theory and Methods*, 45(24):7210–7225.
- Sim, C.-H. (1986). Simulation of weibull and gamma autoregressive stationary process. *Communications in Statistics - Simulation and Computation*, 15(4):1141–1146. doi: 10.1080/03610918608812565.
- Smith, R. L. (1989). Extreme value analysis of environmental time series: an application to trend detection in ground-level ozone. *Statistical Science*, 4(4):367–377.
- Stephenson, A. G. (2009). High-dimensional parametric modelling of multivariate extreme events. *Australian and New Zealand Journal of Statistics*, 51(1):77–88.
- Tavares, L. V. (1980). An exponential Markovian stationary process. *Journal of Applied Probability*, 17(4).
- Uchaikin, V. V. and Zolotarev, V. M. (1999). *Chance and Stability: Stable Distributions and their Applications*. De Gruyter, Berlin, Boston.
- Varin, C., Reid, N., and Firth, D. (2011). An overview of composite likelihood methods. *Statistica Sinica*, 21(1):5–42. <http://www.jstor.org/stable/24309261>.

Apêndice A.

Proposição A.1. *Considere a função $f(v) = e^{-v^\alpha}$, $v > 0$, $0 < \alpha < 1$ e f^n a n -ésima derivada da função $f(v)$. Portanto, a n -ésima derivada de $f(v)$ é dada por:*

$$\frac{d^n}{dv^n} [e^{-v^\alpha}] = (-1)^n e^{-v^\alpha} v^{-n} \sum_{j=1}^n q_{n,j} \alpha^j v^{j\alpha}$$

onde

$$q_{n,j} = \begin{cases} 0 & \text{se } n=1, j=0 \\ q_{n-1,j-1} + (n-1-j\alpha)q_{n-1,j} & \text{se } j=1,2,\dots,n-1 \\ 1 & \text{se } n=j \end{cases}$$

Demonstração. Tem-se que a derivada de primeira a quinta ordem de $f(v)$ são dados por:

$$f^1 = -e^{-v^\alpha} \alpha v^{\alpha-1},$$

$$f^2 = e^{-v^\alpha} \{ \alpha^2 v^{2\alpha-2} + (1-\alpha) \alpha v^{\alpha-2} \},$$

$$\begin{aligned} f^3 = e^{-v^\alpha} \{ & \alpha^3 v^{3\alpha-3} \\ & + \alpha^2 v^{2\alpha-3} [3(1-\alpha)] \\ & + \alpha v^{\alpha-3} [(2-\alpha)(1-\alpha)] \}, \end{aligned}$$

$$\begin{aligned} f^4 = e^{-v^\alpha} \{ & \alpha^4 v^{4\alpha-4} \\ & + \alpha^3 v^{3\alpha-4} [6(1-\alpha)] \\ & + \alpha^2 v^{2\alpha-4} [2(1-\alpha)(1-\alpha) + 3(3-2\alpha)] \\ & + \alpha v^{\alpha-4} [(3-\alpha)(2-\alpha)(1-\alpha)] \}, \end{aligned}$$

$$\begin{aligned} f^5 = e^{-v^\alpha} \{ & \alpha^5 v^{5\alpha-5} \\ & + \alpha^4 v^{4\alpha-5} [6(1-\alpha) + 4(1-\alpha)] \\ & + \alpha^3 v^{3\alpha-5} [(2-\alpha)(1-\alpha) + 3(3-2\alpha)(1-\alpha) + 6(4-3\alpha)(1-\alpha)] \\ & + \alpha^2 v^{2\alpha-5} [(3-\alpha)(2-\alpha)(1-\alpha) + (4-2\alpha) [(2-\alpha)(1-\alpha) + 3(3-2\alpha)]] \\ & + \alpha v^{\alpha-5} [(4-\alpha)(3-\alpha)(2-\alpha)(1-\alpha)] \}. \end{aligned}$$

Portanto, de forma geral, tem-se que n -ésima derivada é dada por:

$$\frac{d^n}{dv^n} [e^{-v^\alpha}] = (-1)^n e^{-v^\alpha} \sum_{j=1}^n q_{n,j} \alpha^j v^{j\alpha-n}$$

□

Proposição A.2. *Seja Z uma v.a α -estável com densidade $p_\alpha(Z)$ e $v > 0$. Então, para $n \geq 2$*

$$\begin{aligned} \int_0^\infty z^n e^{-zv} p_\alpha(z) dz &= \int_0^\infty \frac{d^n}{dv^n} [e^{-zv}] p_\alpha(z) dz \\ &= \frac{d^n}{dv^n} \int_0^\infty e^{-zv} p_\alpha(z) dz \\ &= \frac{d^n}{dv^n} \mathbb{E}(e^{-zv}) \\ &= \frac{d^n}{dv^n} [e^{-v^\alpha}]. \end{aligned}$$

Portanto, o resultado segue pela Proposition A.1

Proposição A.3. *Considere que a distribuição de $(Z|\mathbf{y}, \boldsymbol{\theta}, \alpha)$ é Gamma-Tilted- Stable (Qu et al., 2021) com $z > 0$ e densidade condicional dada por*

$$f_Z(z|\mathbf{y}, \boldsymbol{\theta}, \alpha) = \frac{z^n e^{-z\Lambda_n(\mathbf{y}, \boldsymbol{\theta})} g_\alpha(z)}{G(n, \Lambda_n(\mathbf{y}, \boldsymbol{\theta}))},$$

onde $\Lambda_n(\mathbf{y}, \boldsymbol{\theta}) = \sum_{t=1}^n \Lambda_t(y_t, \nu_t, \delta_t) > 0$ e $G(n, \Lambda_\cdot) = \int_0^\infty z^n e^{-z\Lambda_\cdot} g_\alpha(z) dz$. Então:

$$\begin{aligned} a) \quad \mathbb{E}(Z|\mathbf{y}, \boldsymbol{\theta}, \alpha) &= -\frac{\sum_{j=1}^{n+1} q_{n+1,j} \alpha^j v^{j\alpha}}{\Lambda_n(\mathbf{y}, \boldsymbol{\theta}) \sum_{j=1}^n q_{n,j} \alpha^j v^{j\alpha}}. \\ b) \quad \mathbb{E}(Z^2|\mathbf{y}, \boldsymbol{\theta}, \alpha) &= \frac{\sum_{j=1}^{n+2} q_{n+2,j} \alpha^j v^{j\alpha}}{\Lambda_\cdot(\mathbf{y}, \boldsymbol{\theta})^2 \sum_{j=1}^n q_{n,j} \alpha^j v^{j\alpha}}. \end{aligned}$$

Demonstração.

$$\begin{aligned} \mathbb{E}(Z|\mathbf{y}, \boldsymbol{\theta}, \alpha) &= \int_0^\infty z f_Z(z|\mathbf{y}, \boldsymbol{\theta}, \alpha) dz \\ &= \int_0^\infty z \frac{z^n e^{-z\Lambda_n(\mathbf{y}, \boldsymbol{\theta})} g_\alpha(z)}{G(n, \Lambda_n(\mathbf{y}, \boldsymbol{\theta}))} dz \\ &= \frac{1}{G(n, \Lambda_n(\mathbf{y}, \boldsymbol{\theta}))} \int_0^\infty z^{n+1} e^{-z\Lambda_n(\mathbf{y}, \boldsymbol{\theta})} g_\alpha(z) dz. \end{aligned}$$

Portanto, utilizando as Proposições A.2 e A.1, segue que

$$\tilde{z} = \mathbb{E}(Z|\mathbf{y}, \boldsymbol{\theta}, \alpha) = -\frac{\sum_{j=1}^{n+1} q_{n+1,j} \alpha^j v^{j\alpha}}{\Lambda_n(\mathbf{y}, \boldsymbol{\theta}) \sum_{j=1}^n q_{n,j} \alpha^j v^{j\alpha}}.$$

De forma análoga pode ser demonstrado para $\mathbb{E}(Z^2|\mathbf{y}, \boldsymbol{\theta}, \alpha)$.

□

Proposição A.4. Considere $f(y_t|z, \boldsymbol{\theta}, \alpha)$ a função densidade da variável $Y_t|Z_t \sim \text{Kumaraswamy}(\nu_t, Z/\delta_t)$. Se $Z \sim \alpha - \text{Stable}$, $\alpha \in (0, 1)$, com função densidade $g_\alpha(z)$, então $(Z|\mathbf{y}, \boldsymbol{\theta}, \alpha)$ Gamma-Tilted- Stable (Qu et al., 2021) com $z > 0$ e sua densidade condicional é dada por

$$f_Z(z|\mathbf{y}, \boldsymbol{\theta}, \alpha) = \frac{z^n e^{-z\Lambda_n(\mathbf{y}, \boldsymbol{\theta})} g_\alpha(z)}{G(n, \Lambda_n(\mathbf{y}, \boldsymbol{\theta}))},$$

onde $\Lambda_n(\mathbf{y}, \boldsymbol{\theta}) = \sum_{t=1}^n \Lambda_t(y_t, \nu_t, \delta_t) > 0$ e $G(n, \Lambda_\cdot) = \int_0^\infty z^n e^{-z\Lambda_\cdot} g_\alpha(z) dz$.

Demonstração.

$$\begin{aligned} f_Z(z|\mathbf{y}, \boldsymbol{\theta}, \alpha) &\propto \prod_{t=1}^n f(y_t|z, \boldsymbol{\theta}, \alpha) g_\alpha(z) \\ &\propto \prod_{t=1}^n z e^{-z\Lambda_t(y_t, \nu_t, \delta_t)} g_\alpha(z) \\ &\propto z^n e^{-z\Lambda_n(\mathbf{y}, \boldsymbol{\theta})} g_\alpha(z). \end{aligned} \tag{A1}$$

Neste caso, a constante de proporcionalidade é dada pela $\int_0^\infty z^n e^{-z\Lambda_\cdot} g_\alpha(z) dz = G(n, \Lambda_\cdot)$. Portanto, segue o resultado. $\square$

Apêndice B.

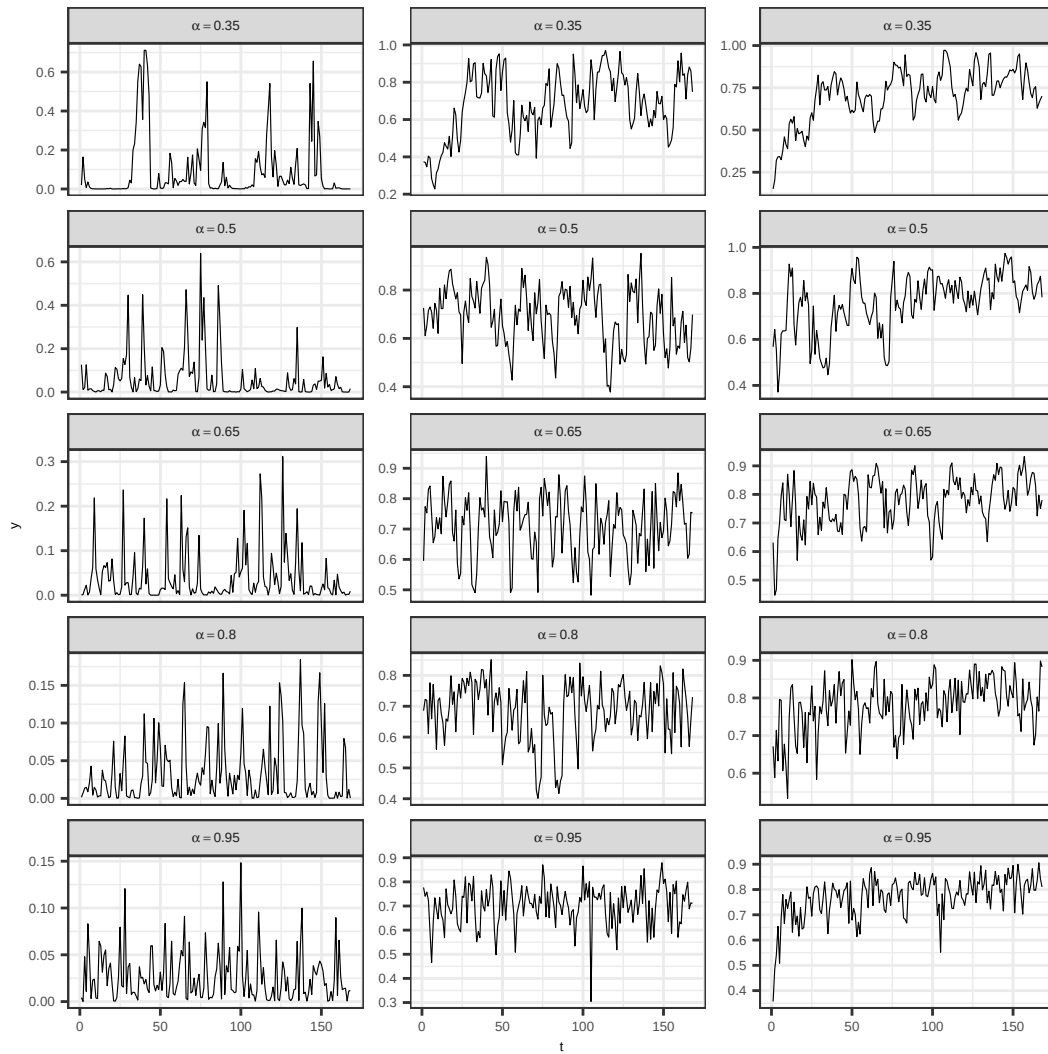


Figura B1. Exemplos de séries simuladas para o Modelo I (esquerda), Modelo II (meio) e Modelo III (direita).

Tabela B1. Estimacões dos parâmetros para o Modelo I com $\alpha = 0.35$

Method	n	Parameter	Par Real	Mean	Median	Relative Bias	Mean Squared Error
Method 1	84	β_1	0.10	0.079	0.079	0.210	0.009
		β_2	-0.10	-0.069	-0.070	0.310	0.008
		γ_0	-3.50	-3.744	-3.766	-0.070	0.881
		γ_1	-0.25	-0.298	-0.310	-0.192	0.289
		α	0.35	0.490	0.486	-0.400	0.022
	144	β_1	0.10	0.084	0.084	0.160	0.005
		β_2	-0.10	-0.070	-0.071	0.300	0.005
		γ_0	-3.50	-3.841	-3.898	-0.097	0.676
		γ_1	-0.25	-0.338	-0.331	-0.352	0.218
		α	0.35	0.468	0.467	-0.337	0.015
	168	β_1	0.10	0.081	0.080	0.190	0.004
		β_2	-0.10	-0.070	-0.070	0.300	0.004
		γ_0	-3.50	-3.913	-3.951	-0.118	0.642
		γ_1	-0.25	-0.369	-0.352	-0.476	0.210
		α	0.35	0.462	0.460	-0.320	0.014
Method 2	84	β_1	0.10	0.074	0.075	0.260	0.015
		β_2	-0.10	-0.068	-0.067	0.320	0.014
		γ_0	-3.50	-3.825	-3.789	-0.093	1.132
		γ_1	-0.25	-0.341	-0.326	-0.364	0.506
		α	0.35	0.491	0.488	-0.403	0.023
	144	β_1	0.10	0.060	0.056	0.400	0.024
		β_2	-0.10	-0.057	-0.060	0.430	0.018
		γ_0	-3.50	-5.517	-5.495	-0.576	4.788
		γ_1	-0.25	-0.594	-0.688	-1.376	0.880
		α	0.35	0.469	0.468	-0.340	0.015
	168	β_1	0.10	0.046	0.043	0.540	0.025
		β_2	-0.10	-0.061	-0.062	0.390	0.015
		γ_0	-3.50	-5.334	-5.327	-0.524	4.156
		γ_1	-0.25	-0.703	-0.724	-1.812	0.870
		α	0.35	0.505	0.502	-0.443	0.025

Tabela B2. Estimacões dos parâmetros para o Modelo I com $\alpha = 0.5$

Method	n	Parameter	Par Real	Mean	Median	Relative Bias	Mean Squared Error
Method 1	84	β_1	0.10	0.084	0.085	0.160	0.009
		β_2	-0.10	-0.076	-0.076	0.240	0.008
		γ_0	-3.50	-3.563	-3.532	-0.018	0.339
		γ_1	-0.25	-0.322	-0.327	-0.288	0.251
		α	0.50	0.562	0.559	-0.124	0.008
	144	β_1	0.10	0.083	0.083	0.170	0.005
		β_2	-0.10	-0.074	-0.074	0.260	0.005
		γ_0	-3.50	-3.672	-3.652	-0.049	0.280
		γ_1	-0.25	-0.319	-0.314	-0.276	0.158
		α	0.50	0.539	0.535	-0.078	0.004
	168	β_1	0.10	0.081	0.080	0.190	0.005
		β_2	-0.10	-0.078	-0.078	0.220	0.004
		γ_0	-3.50	-3.663	-3.646	-0.047	0.279
		γ_1	-0.25	-0.294	-0.300	-0.176	0.133
		α	0.50	0.538	0.535	-0.076	0.004
Method 2	84	β_1	0.10	0.078	0.078	0.220	0.013
		β_2	-0.10	-0.075	-0.076	0.250	0.010
		γ_0	-3.50	-3.550	-3.507	-0.014	0.387
		γ_1	-0.25	-0.405	-0.401	-0.620	0.426
		α	0.50	0.562	0.559	-0.124	0.008
	144	β_1	0.10	0.064	0.060	0.360	0.018
		β_2	-0.10	-0.068	-0.067	0.320	0.012
		γ_0	-3.50	-5.214	-5.371	-0.490	3.963
		γ_1	-0.25	-0.572	-0.590	-1.288	0.689
		α	0.50	0.540	0.536	-0.080	0.004
	168	β_1	0.10	0.039	0.040	0.610	0.023
		β_2	-0.10	-0.066	-0.069	0.340	0.013
		γ_0	-3.50	-5.488	-5.452	-0.568	4.287
		γ_1	-0.25	-0.735	-0.796	-1.940	0.788
		α	0.50	0.552	0.547	-0.104	0.004

Tabela B3. Estimacões dos parâmetros para o Modelo I com $\alpha = 0.65$

Method	n	Parameter	Par Real	Mean	Median	Relative Bias	Mean Squared Error
Method 1	84	β_1	0.10	0.086	0.086	0.140	0.009
		β_2	-0.10	-0.089	-0.090	0.110	0.007
		γ_0	-3.50	-3.509	-3.497	-0.003	0.151
		γ_1	-0.25	-0.328	-0.319	-0.312	0.239
		α	0.65	0.681	0.684	-0.048	0.007
	144	β_1	0.10	0.089	0.088	0.110	0.006
		β_2	-0.10	-0.086	-0.085	0.140	0.004
		γ_0	-3.50	-3.499	-3.473	0.000	0.094
		γ_1	-0.25	-0.290	-0.304	-0.160	0.124
		α	0.65	0.672	0.673	-0.034	0.004
	168	β_1	0.10	0.089	0.090	0.110	0.004
		β_2	-0.10	-0.086	-0.087	0.140	0.003
		γ_0	-3.50	-3.505	-3.478	-0.001	0.079
		γ_1	-0.25	-0.255	-0.247	-0.020	0.103
		α	0.65	0.675	0.676	-0.038	0.004
Method 2	84	β_1	0.10	0.077	0.080	0.230	0.013
		β_2	-0.10	-0.087	-0.086	0.130	0.008
		γ_0	-3.50	-3.503	-3.483	-0.001	0.165
		γ_1	-0.25	-0.387	-0.388	-0.548	0.338
		α	0.65	0.681	0.684	-0.048	0.007
	144	β_1	0.10	0.083	0.082	0.170	0.008
		β_2	-0.10	-0.083	-0.083	0.170	0.005
		γ_0	-3.50	-3.467	-3.451	0.009	0.094
		γ_1	-0.25	-0.335	-0.338	-0.340	0.193
		α	0.65	0.673	0.675	-0.035	0.004
	168	β_1	0.10	0.070	0.075	0.300	0.010
		β_2	-0.10	-0.082	-0.082	0.180	0.005
		γ_0	-3.50	-4.805	-4.976	-0.373	2.533
		γ_1	-0.25	-0.424	-0.413	-0.696	0.286
		α	0.65	0.677	0.678	-0.042	0.004

Tabela B4. Estimacões dos parâmetros para o Modelo I com $\alpha = 0.8$

Method	n	Parameter	Par Real	Mean	Median	Relative Bias	Mean Squared Error
Method 1	84	β_1	0.10	0.085	0.087	0.150	0.009
		β_2	-0.10	-0.093	-0.093	0.070	0.004
		γ_0	-3.50	-3.488	-3.475	0.003	0.084
		γ_1	-0.25	-0.306	-0.322	-0.224	0.225
		α	0.80	0.836	0.837	-0.045	0.006
	144	β_1	0.10	0.086	0.087	0.140	0.005
		β_2	-0.10	-0.092	-0.091	0.080	0.003
		γ_0	-3.50	-3.469	-3.465	0.009	0.047
		γ_1	-0.25	-0.312	-0.314	-0.248	0.124
		α	0.80	0.823	0.823	-0.029	0.003
	168	β_1	0.10	0.090	0.091	0.100	0.004
		β_2	-0.10	-0.095	-0.094	0.050	0.002
		γ_0	-3.50	-3.480	-3.474	0.006	0.042
		γ_1	-0.25	-0.280	-0.280	-0.120	0.107
		α	0.80	0.822	0.820	-0.027	0.003
Method 2	84	β_1	0.10	0.082	0.083	0.180	0.010
		β_2	-0.10	-0.092	-0.091	0.080	0.004
		γ_0	-3.50	-3.525	-3.505	-0.007	0.094
		γ_1	-0.25	-0.323	-0.330	-0.292	0.246
		α	0.80	0.837	0.837	-0.046	0.006
	144	β_1	0.10	0.082	0.083	0.180	0.006
		β_2	-0.10	-0.091	-0.090	0.090	0.003
		γ_0	-3.50	-3.482	-3.476	0.005	0.052
		γ_1	-0.25	-0.341	-0.357	-0.364	0.147
		α	0.80	0.823	0.823	-0.029	0.003
	168	β_1	0.10	0.089	0.090	0.110	0.005
		β_2	-0.10	-0.094	-0.094	0.060	0.002
		γ_0	-3.50	-3.500	-3.493	0.000	0.044
		γ_1	-0.25	-0.285	-0.290	-0.140	0.119
		α	0.80	0.822	0.820	-0.027	0.003

Tabela B5. Estimacões dos parâmetros para o Modelo I com $\alpha = 0.95$

Method	n	Parameter	Par Real	Mean	Median	Relative Bias	Mean Squared Error
Method 1	84	β_1	0.10	0.098	0.096	0.020	0.008
		β_2	-0.10	-0.096	-0.097	0.040	0.002
		γ_0	-3.50	-3.513	-3.504	-0.004	0.059
		γ_1	-0.25	-0.248	-0.246	0.008	0.185
		α	0.95	0.967	0.988	-0.018	0.001
	144	β_1	0.10	0.097	0.095	0.030	0.004
		β_2	-0.10	-0.098	-0.097	0.020	0.001
		γ_0	-3.50	-3.497	-3.503	0.001	0.033
		γ_1	-0.25	-0.262	-0.258	-0.048	0.109
		α	0.95	0.963	0.972	-0.014	0.001
	168	β_1	0.10	0.091	0.089	0.090	0.004
		β_2	-0.10	-0.095	-0.096	0.050	0.001
		γ_0	-3.50	-3.483	-3.486	0.005	0.029
		γ_1	-0.25	-0.288	-0.279	-0.152	0.093
		α	0.95	0.961	0.968	-0.012	0.001
Method 2	84	β_1	0.10	0.097	0.097	0.030	0.008
		β_2	-0.10	-0.096	-0.097	0.040	0.003
		γ_0	-3.50	-3.552	-3.548	-0.015	0.062
		γ_1	-0.25	-0.246	-0.250	0.016	0.190
		α	0.95	0.967	0.988	-0.018	0.001
	144	β_1	0.10	0.096	0.097	0.040	0.004
		β_2	-0.10	-0.097	-0.097	0.030	0.001
		γ_0	-3.50	-3.537	-3.538	-0.011	0.034
		γ_1	-0.25	-0.268	-0.260	-0.072	0.114
		α	0.95	0.963	0.972	-0.014	0.001
	168	β_1	0.10	0.092	0.091	0.080	0.004
		β_2	-0.10	-0.095	-0.096	0.050	0.001
		γ_0	-3.50	-3.524	-3.529	-0.007	0.030
		γ_1	-0.25	-0.284	-0.274	-0.136	0.094
		α	0.95	0.961	0.968	-0.012	0.001

Tabela B6. Estimacões dos parâmetros para o modelo II com $\alpha = 0.35$

Method	n	Parameter	Par Real	Mean	Median	Relative Bias	Mean Squared Error
Method 1	84	β_0	2.50	2.538	2.538	-0.015	0.012
		β_1	0.10	0.102	0.102	-0.020	0.007
		β_2	-0.10	-0.104	-0.103	-0.040	0.006
		γ_1	-3.50	-3.635	-3.532	-0.039	0.604
		γ_2	-3.75	-3.895	-3.789	-0.039	0.630
		α	0.35	0.429	0.426	-0.226	0.011
	144	β_0	2.50	2.531	2.529	-0.012	0.008
		β_1	0.10	0.100	0.100	< 0.001	0.004
		β_2	-0.10	-0.102	-0.099	-0.020	0.003
		γ_1	-3.50	-3.495	-3.453	0.001	0.255
		γ_2	-3.75	-3.751	-3.712	< 0.001	0.270
		α	0.35	0.419	0.417	-0.197	0.007
	168	β_0	2.50	2.528	2.525	-0.011	0.007
		β_1	0.10	0.101	0.102	-0.010	0.003
		β_2	-0.10	-0.103	-0.104	-0.030	0.003
		γ_1	-3.50	-3.451	-3.431	0.014	0.183
		γ_2	-3.75	-3.698	-3.678	0.014	0.192
		α	0.35	0.417	0.417	-0.191	0.007
Method 2	84	β_0	2.50	1.660	1.645	0.336	0.757
		β_1	0.10	0.096	0.097	0.040	0.022
		β_2	-0.10	-0.101	-0.102	-0.010	0.013
		γ_1	-3.50	-2.109	-2.016	0.397	2.220
		γ_2	-3.75	-2.253	-2.171	0.399	2.533
		α	0.35	0.453	0.447	-0.294	0.014
	144	β_0	2.50	1.338	1.362	0.465	1.422
		β_1	0.10	0.084	0.087	0.160	0.030
		β_2	-0.10	-0.103	-0.102	-0.030	0.024
		γ_1	-3.50	-3.576	-3.544	-0.022	0.286
		γ_2	-3.75	-3.761	-3.697	-0.003	0.309
		α	0.35	0.469	0.463	-0.340	0.015
	168	β_0	2.50	1.892	1.666	0.243	0.658
		β_1	0.10	0.087	0.096	0.130	0.008
		β_2	-0.10	-0.096	-0.094	0.040	0.007
		γ_1	-3.50	-3.725	-3.658	-0.064	0.398
		γ_2	-3.75	-3.880	-3.863	-0.035	0.361
		α	0.35	0.510	0.531	-0.457	0.029

Tabela B7. Estimacões dos parâmetros para o modelo II com $\alpha = 0.5$

Method	n	Parameter	Par Real	Mean	Median	Relative Bias	Mean Squared Error
Method 1	84	β_0	2.50	2.526	2.523	-0.010	0.012
		β_1	0.10	0.099	0.100	0.010	0.009
		β_2	-0.10	-0.103	-0.105	-0.030	0.008
		γ_1	-3.50	-3.682	-3.654	-0.052	0.393
		γ_2	-3.75	-3.952	-3.918	-0.054	0.427
		α	0.50	0.530	0.526	-0.060	0.006
	144	β_0	2.50	2.512	2.512	-0.005	0.007
		β_1	0.10	0.100	0.098	< 0.001	0.005
		β_2	-0.10	-0.103	-0.103	-0.030	0.004
		γ_1	-3.50	-3.600	-3.574	-0.029	0.195
		γ_2	-3.75	-3.865	-3.832	-0.031	0.211
		α	0.50	0.521	0.518	-0.042	0.003
	168	β_0	2.50	2.512	2.513	-0.005	0.006
		β_1	0.10	0.100	0.099	< 0.001	0.005
		β_2	-0.10	-0.103	-0.103	-0.030	0.004
		γ_1	-3.50	-3.593	-3.567	-0.027	0.165
		γ_2	-3.75	-3.837	-3.815	-0.023	0.164
		α	0.50	0.519	0.515	-0.038	0.003
Method 2	84	β_0	2.50	1.857	1.854	0.257	0.454
		β_1	0.10	0.094	0.093	0.060	0.020
		β_2	-0.10	-0.102	-0.102	-0.020	0.011
		γ_1	-3.50	-2.298	-2.261	0.343	1.614
		γ_2	-3.75	-2.475	-2.441	0.340	1.814
		α	0.50	0.533	0.529	-0.066	0.006
	144	β_0	2.50	1.506	1.488	0.398	1.067
		β_1	0.10	0.094	0.094	0.060	0.023
		β_2	-0.10	-0.100	-0.099	< 0.001	0.015
		γ_1	-3.50	-3.795	-3.772	-0.084	0.270
		γ_2	-3.75	-3.942	-3.963	-0.051	0.255
		α	0.50	0.525	0.520	-0.050	0.003
	168	β_0	2.50	1.504	1.500	0.398	1.024
		β_1	0.10	0.080	0.080	0.200	0.018
		β_2	-0.10	-0.092	-0.098	0.080	0.013
		γ_1	-3.50	-3.785	-3.761	-0.081	0.169
		γ_2	-3.75	-3.969	-3.965	-0.058	0.138
		α	0.50	0.571	0.565	-0.142	0.006

Tabela B8. Estimacões dos parâmetros para o modelo II com $\alpha = 0.65$

Method	n	Parameter	Par Real	Mean	Median	Relative Bias	Mean Squared Error
Method 1	84	β_0	2.50	2.529	2.529	-0.012	0.012
		β_1	0.10	0.101	0.100	-0.010	0.010
		β_2	-0.10	-0.100	-0.101	< 0.001	0.007
		γ_1	-3.50	-3.643	-3.621	-0.041	0.260
		γ_2	-3.75	-3.908	-3.880	-0.042	0.271
		α	0.65	0.673	0.675	-0.035	0.007
	144	β_0	2.50	2.518	2.518	-0.007	0.007
		β_1	0.10	0.099	0.099	0.010	0.006
		β_2	-0.10	-0.102	-0.101	-0.020	0.004
		γ_1	-3.50	-3.599	-3.585	-0.028	0.140
		γ_2	-3.75	-3.864	-3.841	-0.030	0.155
		α	0.65	0.659	0.658	-0.014	0.004
	168	β_0	2.50	2.520	2.522	-0.008	0.006
		β_1	0.10	0.100	0.099	< 0.001	0.005
		β_2	-0.10	-0.100	-0.101	< 0.001	0.004
		γ_1	-3.50	-3.589	-3.574	-0.025	0.121
		γ_2	-3.75	-3.856	-3.851	-0.028	0.127
		α	0.65	0.658	0.658	-0.012	0.003
Method 2	84	β_0	2.50	2.110	2.114	0.156	0.175
		β_1	0.10	0.097	0.098	0.030	0.015
		β_2	-0.10	-0.099	-0.099	0.010	0.008
		γ_1	-3.50	-2.607	-2.594	0.255	0.929
		γ_2	-3.75	-2.799	-2.774	0.254	1.043
		α	0.65	0.673	0.675	-0.035	0.007
	144	β_0	2.50	2.061	2.071	0.176	0.211
		β_1	0.10	0.096	0.096	0.040	0.010
		β_2	-0.10	-0.101	-0.101	-0.010	0.005
		γ_1	-3.50	-2.520	-2.506	0.280	1.039
		γ_2	-3.75	-2.714	-2.700	0.276	1.164
		α	0.65	0.659	0.658	-0.014	0.004
	168	β_0	2.50	1.912	1.958	0.235	0.408
		β_1	0.10	0.093	0.093	0.070	0.011
		β_2	-0.10	-0.099	-0.099	0.010	0.006
		γ_1	-3.50	-4.016	-4.061	-0.147	0.450
		γ_2	-3.75	-4.186	-4.244	-0.116	0.391
		α	0.65	0.659	0.659	-0.014	0.003

Tabela B9. Estimacões dos parâmetros para o modelo II com $\alpha = 0.8$

Method	n	Parameter	Par Real	Mean	Median	Relative Bias	Mean Squared Error
Method 1	84	β_0	2.50	2.540	2.539	-0.016	0.013
		β_1	0.10	0.099	0.099	0.010	0.009
		β_2	-0.10	-0.100	-0.101	< 0.001	0.004
		γ_1	-3.50	-3.656	-3.629	-0.045	0.230
		γ_2	-3.75	-3.925	-3.904	-0.047	0.250
		α	0.80	0.819	0.820	-0.024	0.006
	144	β_0	2.50	2.534	2.535	-0.014	0.008
		β_1	0.10	0.099	0.098	0.010	0.005
		β_2	-0.10	-0.100	-0.100	< 0.001	0.003
		γ_1	-3.50	-3.623	-3.614	-0.035	0.134
		γ_2	-3.75	-3.878	-3.859	-0.034	0.139
		α	0.80	0.808	0.808	-0.010	0.003
	168	β_0	2.50	2.531	2.529	-0.012	0.007
		β_1	0.10	0.099	0.098	0.010	0.005
		β_2	-0.10	-0.099	-0.100	0.010	0.002
		γ_1	-3.50	-3.616	-3.612	-0.033	0.119
		γ_2	-3.75	-3.871	-3.867	-0.032	0.116
		α	0.80	0.805	0.805	-0.006	0.002
Method 2	84	β_0	2.50	2.308	2.310	0.077	0.051
		β_1	0.10	0.099	0.098	0.010	0.012
		β_2	-0.10	-0.100	-0.099	< 0.001	0.005
		γ_1	-3.50	-3.035	-3.021	0.133	0.341
		γ_2	-3.75	-3.256	-3.245	0.132	0.391
		α	0.80	0.819	0.820	-0.024	0.006
	144	β_0	2.50	2.288	2.289	0.085	0.053
		β_1	0.10	0.098	0.098	0.020	0.006
		β_2	-0.10	-0.101	-0.100	-0.010	0.003
		γ_1	-3.50	-2.949	-2.951	0.157	0.376
		γ_2	-3.75	-3.163	-3.155	0.157	0.425
		α	0.80	0.808	0.808	-0.010	0.003
	168	β_0	2.50	2.281	2.281	0.088	0.055
		β_1	0.10	0.098	0.099	0.020	0.006
		β_2	-0.10	-0.099	-0.098	0.010	0.003
		γ_1	-3.50	-2.963	-2.950	0.153	0.361
		γ_2	-3.75	-3.182	-3.169	0.151	0.401
		α	0.80	0.805	0.805	-0.006	0.002

Tabela B10. Estimções dos parâmetros para o modelo II com $\alpha = 0.95$

Method	n	Parameter	Par Real	Mean	Median	Relative Bias	Mean Squared Error
Method 1	84	β_0	2.50	2.578	2.580	-0.031	0.018
		β_1	0.10	0.099	0.097	0.010	0.008
		β_2	-0.10	-0.100	-0.101	< 0.001	0.003
		γ_1	-3.50	-3.796	-3.788	-0.085	0.300
		γ_2	-3.75	-4.066	-4.041	-0.084	0.312
		α	0.95	0.944	0.954	0.006	0.002
	144	β_0	2.50	2.562	2.560	-0.025	0.011
		β_1	0.10	0.099	0.099	0.010	0.004
		β_2	-0.10	-0.099	-0.099	0.010	0.002
		γ_1	-3.50	-3.718	-3.705	-0.062	0.176
		γ_2	-3.75	-3.995	-3.972	-0.065	0.187
		α	0.95	0.943	0.948	0.007	0.002
	168	β_0	2.50	2.561	2.561	-0.024	0.009
		β_1	0.10	0.097	0.097	0.030	0.004
		β_2	-0.10	-0.099	-0.100	0.010	0.001
		γ_1	-3.50	-3.717	-3.701	-0.062	0.152
		γ_2	-3.75	-3.999	-3.978	-0.066	0.168
		α	0.95	0.941	0.945	0.009	0.002
Method 2	84	β_0	2.50	2.485	2.485	0.006	0.009
		β_1	0.10	0.100	0.098	< 0.001	0.008
		β_2	-0.10	-0.100	-0.101	< 0.001	0.003
		γ_1	-3.50	-3.542	-3.524	-0.012	0.148
		γ_2	-3.75	-3.790	-3.775	-0.011	0.155
		α	0.95	0.944	0.954	0.006	0.002
	144	β_0	2.50	2.465	2.463	0.014	0.007
		β_1	0.10	0.099	0.101	0.010	0.004
		β_2	-0.10	-0.099	-0.099	0.010	0.002
		γ_1	-3.50	-3.455	-3.443	0.013	0.081
		γ_2	-3.75	-3.703	-3.681	0.013	0.090
		α	0.95	0.943	0.948	0.007	0.002
	168	β_0	2.50	2.463	2.462	0.015	0.006
		β_1	0.10	0.098	0.098	0.020	0.004
		β_2	-0.10	-0.100	-0.101	< 0.001	0.001
		γ_1	-3.50	-3.442	-3.436	0.017	0.067
		γ_2	-3.75	-3.700	-3.691	0.013	0.078
		α	0.95	0.941	0.945	0.009	0.002

Tabela B11. Estimações dos parâmetros para o modelo III com $\alpha = 0.35$

Method	n	Parameter	Par Real	Mean	Median	Relative Bias	Mean Squared Error
Method 1	84	β_0	2.00	2.036	2.014	-0.018	0.157
		β_1	0.20	0.198	0.200	0.011	0.012
		γ_0	-3.50	-3.524	-3.466	-0.007	0.463
		γ_1	-0.25	-0.254	-0.252	-0.016	0.225
		α	0.35	0.433	0.428	-0.236	0.012
	144	β_0	2.00	2.028	2.015	-0.014	0.118
		β_1	0.20	0.199	0.201	0.004	0.007
		γ_0	-3.50	-3.432	-3.423	0.019	0.198
		γ_1	-0.25	-0.254	-0.259	-0.016	0.130
		α	0.35	0.421	0.420	-0.202	0.008
	168	β_0	2.00	2.029	2.019	-0.014	0.102
		β_1	0.20	0.199	0.201	0.006	0.006
		γ_0	-3.50	-3.404	-3.412	0.027	0.136
		γ_1	-0.25	-0.276	-0.283	-0.102	0.116
		α	0.35	0.415	0.415	-0.187	0.006
Method 2	84	β_0	2.00	1.158	1.164	0.421	1.202
		β_1	0.20	0.197	0.196	0.013	0.038
		γ_0	-3.50	-2.084	-2.046	0.405	2.200
		γ_1	-0.25	-0.120	-0.122	0.520	0.127
		α	0.35	0.456	0.448	-0.302	0.015
	144	β_0	2.00	1.634	1.683	0.183	0.577
		β_1	0.20	0.029	0.031	0.855	0.062
		γ_0	-3.50	-3.560	-3.502	-0.017	0.406
		γ_1	-0.25	-0.205	-0.196	0.181	0.173
		α	0.35	0.466	0.462	-0.331	0.014
	168	β_0	2.00	1.830	1.844	0.085	0.268
		β_1	0.20	0.084	0.101	0.579	0.036
		γ_0	-3.50	-3.580	-3.535	-0.023	0.301
		γ_1	-0.25	-0.408	-0.410	-0.632	0.183
		α	0.35	0.527	0.536	-0.506	0.033

Tabela B12. Estimações dos parâmetros para o modelo III com $\alpha = 0.5$

Method	n	Parameter	Par Real	Mean	Median	Relative Bias	Mean Squared Error
Method 1	84	β_0	2.00	2.025	2.020	-0.012	0.129
		β_1	0.20	0.198	0.201	0.010	0.010
		γ_0	-3.50	-3.608	-3.578	-0.031	0.322
		γ_1	-0.25	-0.262	-0.256	-0.050	0.214
		α	0.50	0.540	0.536	-0.080	0.007
	144	β_0	2.00	2.008	1.999	-0.004	0.099
		β_1	0.20	0.201	0.200	-0.003	0.006
		γ_0	-3.50	-3.594	-3.575	-0.027	0.193
		γ_1	-0.25	-0.256	-0.246	-0.024	0.116
		α	0.50	0.522	0.518	-0.044	0.003
	168	β_0	2.00	2.000	2.000	< 0.001	0.089
		β_1	0.20	0.202	0.202	-0.012	0.005
		γ_0	-3.50	-3.589	-3.575	-0.025	0.166
		γ_1	-0.25	-0.263	-0.269	-0.051	0.106
		α	0.50	0.520	0.518	-0.039	0.003
Method 2	84	β_0	2.00	1.372	1.383	0.314	0.636
		β_1	0.20	0.194	0.195	0.028	0.018
		γ_0	-3.50	-2.291	-2.273	0.345	1.595
		γ_1	-0.25	-0.146	-0.136	0.417	0.116
		α	0.50	0.542	0.537	-0.084	0.007
	144	β_0	2.00	1.472	1.478	0.264	0.564
		β_1	0.20	0.091	0.092	0.545	0.031
		γ_0	-3.50	-3.664	-3.679	-0.047	0.347
		γ_1	-0.25	-0.162	-0.153	0.351	0.112
		α	0.50	0.526	0.522	-0.052	0.003
	168	β_0	2.00	1.693	1.716	0.153	0.301
		β_1	0.20	0.084	0.084	0.582	0.027
		γ_0	-3.50	-3.917	-3.913	-0.119	0.300
		γ_1	-0.25	-0.252	-0.250	-0.010	0.116
		α	0.50	0.563	0.557	-0.126	0.005

Tabela B13. Estimções dos parâmetros para o modelo III com $\alpha = 0.65$

Method	n	Parameter	Par Real	Mean	Median	Relative Bias	Mean Squared Error
Method 1	84	β_0	2.00	2.015	2.015	-0.007	0.079
		β_1	0.20	0.203	0.202	-0.013	0.006
		γ_0	-3.50	-3.602	-3.580	-0.029	0.203
		γ_1	-0.25	-0.258	-0.251	-0.031	0.143
		α	0.65	0.675	0.675	-0.039	0.006
	144	β_0	2.00	2.009	2.011	-0.004	0.060
		β_1	0.20	0.201	0.201	-0.007	0.003
		γ_0	-3.50	-3.584	-3.568	-0.024	0.118
		γ_1	-0.25	-0.248	-0.258	0.010	0.083
		α	0.65	0.661	0.660	-0.017	0.003
	168	β_0	2.00	2.008	2.006	-0.004	0.054
		β_1	0.20	0.201	0.202	-0.005	0.003
		γ_0	-3.50	-3.574	-3.555	-0.021	0.105
		γ_1	-0.25	-0.253	-0.249	-0.014	0.074
		α	0.65	0.658	0.659	-0.012	0.003
Method 2	84	β_0	2.00	1.601	1.608	0.200	0.263
		β_1	0.20	0.200	0.198	-0.001	0.007
		γ_0	-3.50	-2.571	-2.560	0.265	0.970
		γ_1	-0.25	-0.179	-0.181	0.283	0.087
		α	0.65	0.676	0.675	-0.039	0.006
	144	β_0	2.00	1.576	1.591	0.212	0.256
		β_1	0.20	0.194	0.194	0.030	0.004
		γ_0	-3.50	-2.510	-2.505	0.283	1.039
		γ_1	-0.25	-0.169	-0.166	0.326	0.053
		α	0.65	0.661	0.660	-0.017	0.003
	168	β_0	2.00	1.608	1.613	0.196	0.241
		β_1	0.20	0.164	0.171	0.181	0.008
		γ_0	-3.50	-3.986	-4.046	-0.139	0.421
		γ_1	-0.25	-0.188	-0.181	0.249	0.062
		α	0.65	0.660	0.660	-0.015	0.003

Tabela B14. Estimções dos parâmetros para o modelo III com $\alpha = 0.8$

Method	n	Parameter	Par Real	Mean	Median	Relative Bias	Mean Squared Error
Method 1	84	β_0	2.00	2.027	2.035	-0.014	0.049
		β_1	0.20	0.203	0.204	-0.014	0.003
		γ_0	-3.50	-3.656	-3.631	-0.045	0.202
		γ_1	-0.25	-0.250	-0.244	< 0.001	0.099
		α	0.80	0.817	0.818	-0.021	0.005
	144	β_0	2.00	2.027	2.029	-0.014	0.032
		β_1	0.20	0.201	0.201	-0.005	0.002
		γ_0	-3.50	-3.608	-3.598	-0.031	0.118
		γ_1	-0.25	-0.272	-0.270	-0.088	0.053
		α	0.80	0.805	0.808	-0.007	0.003
	168	β_0	2.00	2.034	2.036	-0.017	0.029
		β_1	0.20	0.201	0.200	-0.004	0.001
		γ_0	-3.50	-3.628	-3.610	-0.036	0.101
		γ_1	-0.25	-0.262	-0.262	-0.050	0.048
		α	0.80	0.805	0.806	-0.006	0.002
Method 2	84	β_0	2.00	1.794	1.802	0.103	0.093
		β_1	0.20	0.201	0.201	-0.007	0.003
		γ_0	-3.50	-3.002	-2.976	0.142	0.364
		γ_1	-0.25	-0.201	-0.200	0.195	0.071
		α	0.80	0.817	0.818	-0.021	0.005
	144	β_0	2.00	1.783	1.786	0.108	0.083
		β_1	0.20	0.199	0.198	0.005	0.002
		γ_0	-3.50	-2.926	-2.917	0.164	0.406
		γ_1	-0.25	-0.216	-0.211	0.137	0.037
		α	0.80	0.805	0.808	-0.007	0.003
	168	β_0	2.00	1.791	1.795	0.104	0.076
		β_1	0.20	0.199	0.198	0.007	0.001
		γ_0	-3.50	-2.972	-2.971	0.151	0.350
		γ_1	-0.25	-0.210	-0.213	0.159	0.034
		α	0.80	0.805	0.806	-0.006	0.002

Tabela B15. Estimções dos parâmetros para o modelo III com $\alpha = 0.95$

Method	n	Parameter	Par Real	Mean	Median	Relative Bias	Mean Squared Error
Method 1	84	β_0	2.00	2.088	2.091	-0.044	0.035
		β_1	0.20	0.200	0.200	-0.001	0.001
		γ_0	-3.50	-3.839	-3.835	-0.097	0.282
		γ_1	-0.25	-0.259	-0.255	-0.037	0.069
		α	0.95	0.943	0.953	0.007	0.002
	144	β_0	2.00	2.074	2.074	-0.037	0.024
		β_1	0.20	0.201	0.200	-0.004	0.001
		γ_0	-3.50	-3.771	-3.756	-0.077	0.162
		γ_1	-0.25	-0.260	-0.254	-0.041	0.035
		α	0.95	0.940	0.944	0.011	0.002
	168	β_0	2.00	2.069	2.069	-0.034	0.022
		β_1	0.20	0.200	0.200	-0.002	0.001
		γ_0	-3.50	-3.749	-3.733	-0.071	0.141
		γ_1	-0.25	-0.266	-0.264	-0.064	0.030
		α	0.95	0.939	0.942	0.011	0.002
Method 2	84	β_0	2.00	1.986	1.992	0.007	0.025
		β_1	0.20	0.200	0.199	< 0.001	0.001
		γ_0	-3.50	-3.543	-3.548	-0.012	0.117
		γ_1	-0.25	-0.233	-0.231	0.067	0.056
		α	0.95	0.943	0.953	0.007	0.002
	144	β_0	2.00	1.968	1.970	0.016	0.017
		β_1	0.20	0.201	0.200	-0.003	0.001
		γ_0	-3.50	-3.468	-3.466	0.009	0.053
		γ_1	-0.25	-0.237	-0.234	0.052	0.029
		α	0.95	0.940	0.944	0.011	0.002
	168	β_0	2.00	1.963	1.965	0.018	0.017
		β_1	0.20	0.200	0.200	0.001	0.001
		γ_0	-3.50	-3.445	-3.435	0.016	0.050
		γ_1	-0.25	-0.241	-0.241	0.037	0.025
		α	0.95	0.939	0.942	0.011	0.002

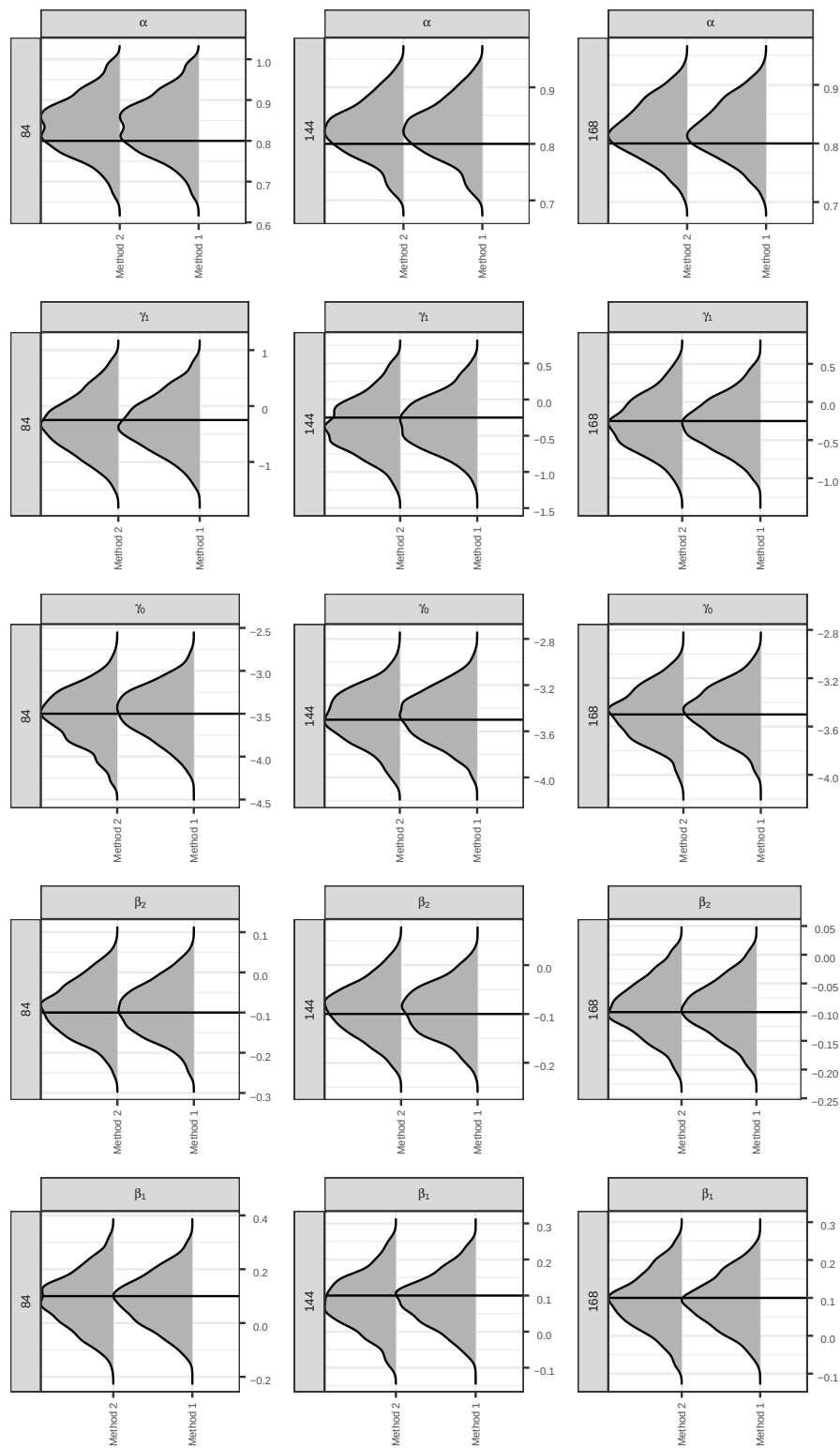


Figura B2. Distribuições dos estimadores dos parâmetros para o modelo I e $\alpha = 0.80$, linha horizontal representa o valor real dos parâmetros

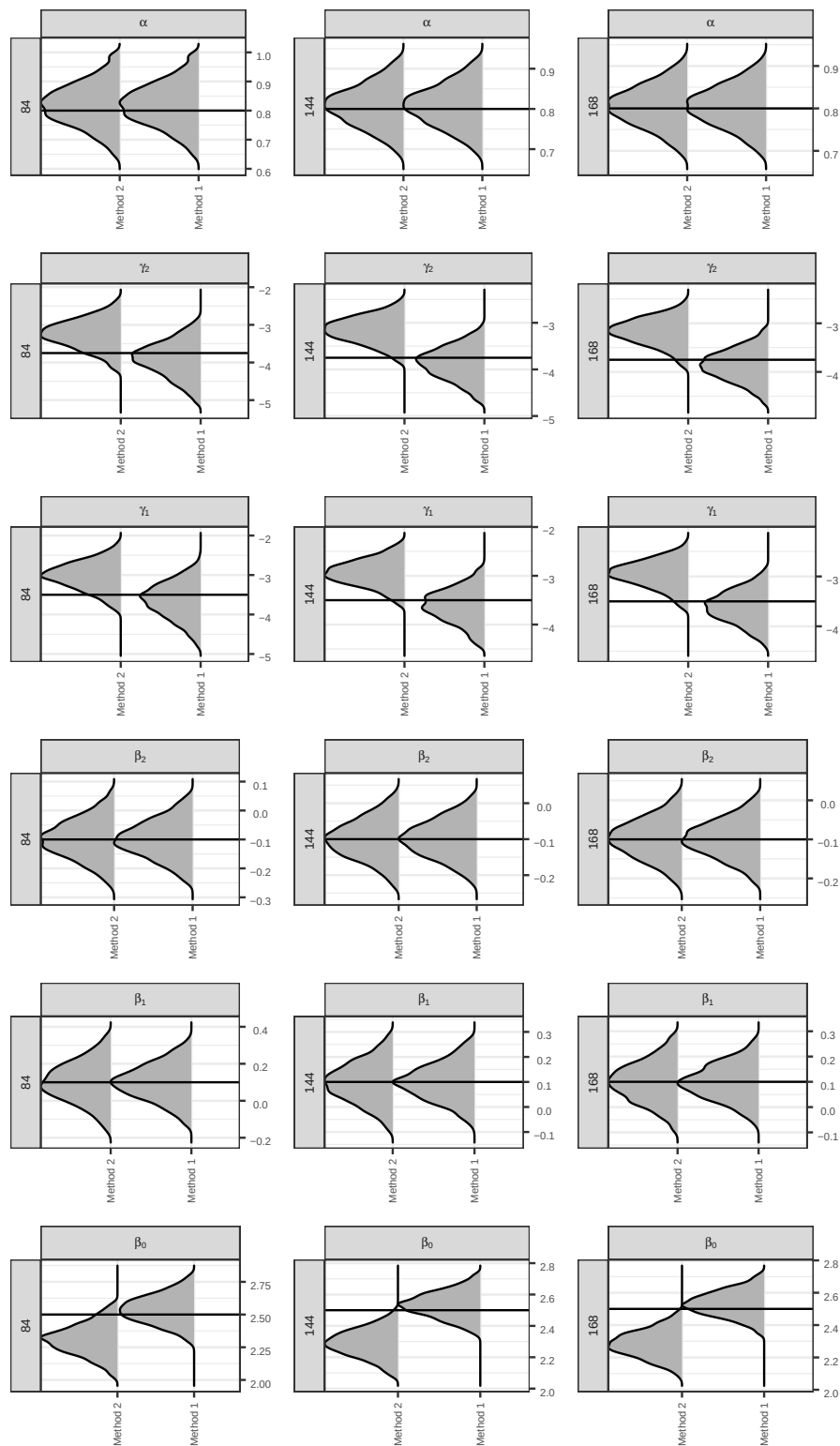


Figura B3. Distribuições dos estimadores dos parâmetros para o modelo II e $\alpha = 0.80$, linha horizontal representa o valor real dos parâmetros

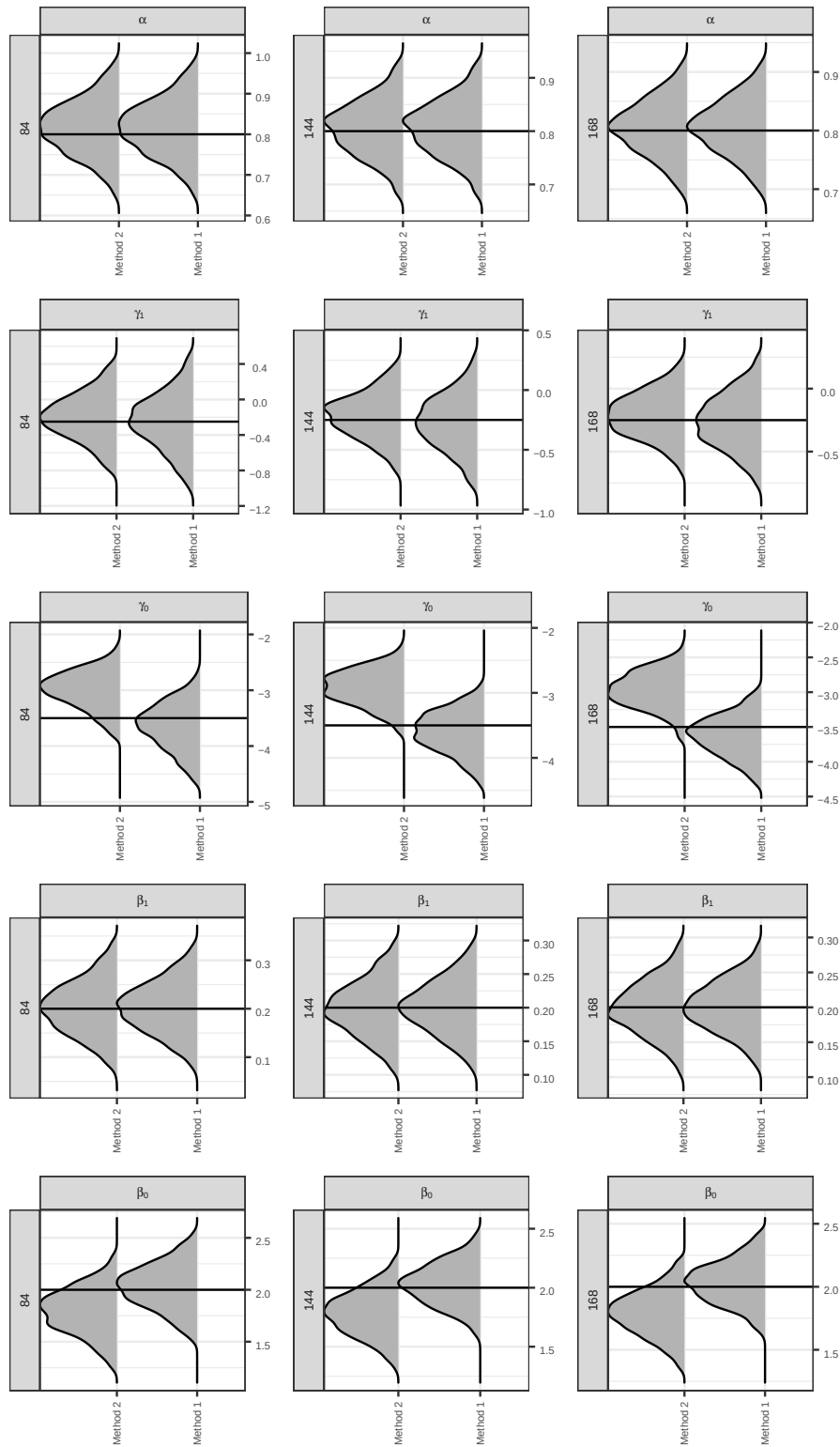


Figura B4. Distribuições dos estimadores dos parâmetros para o modelo III e $\alpha = 0.80$, linha horizontal representa o valor real dos parâmetros

Apêndice C.

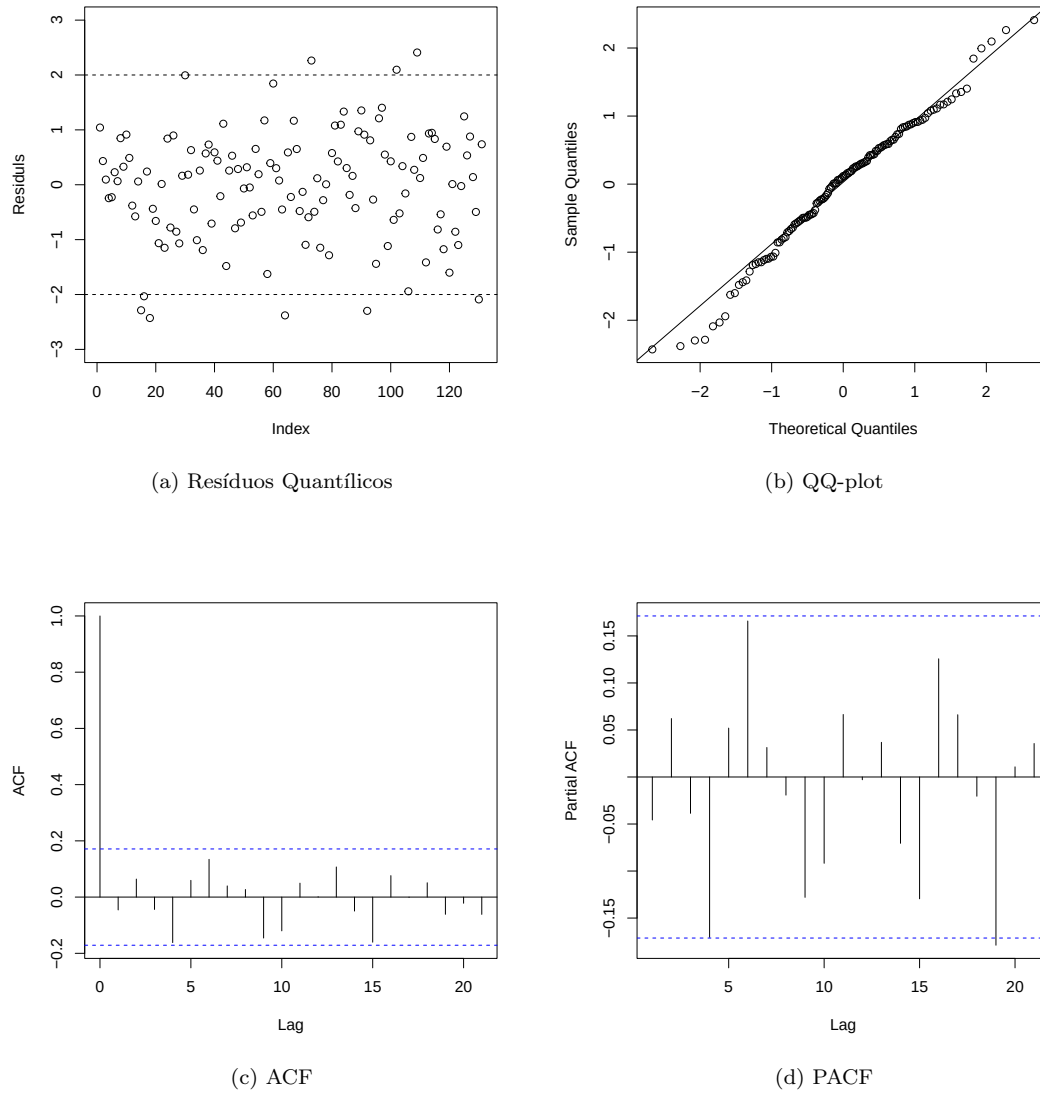
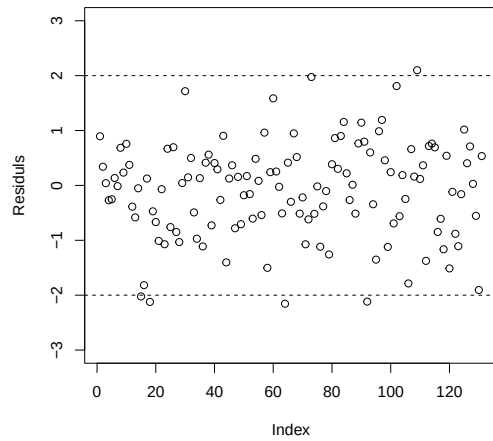
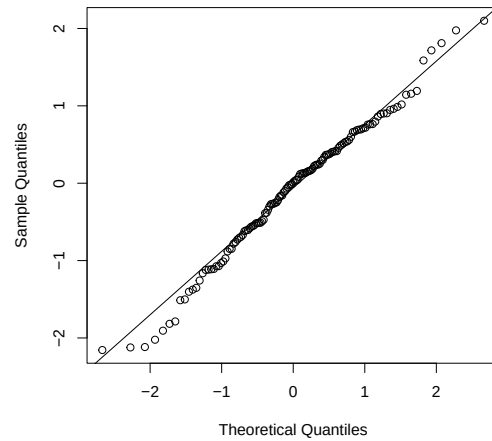


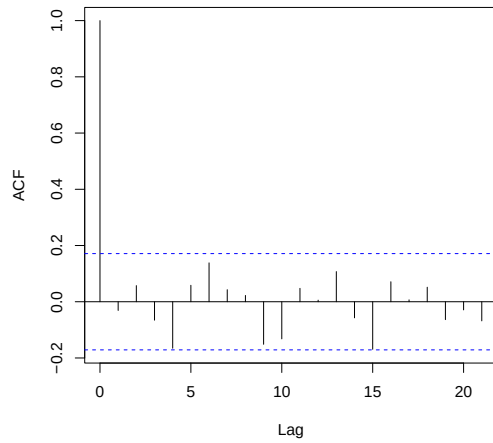
Figura C1. Diagnóstico, do modelo ajustado, baseado nos resíduos quantílicos - Método 1.



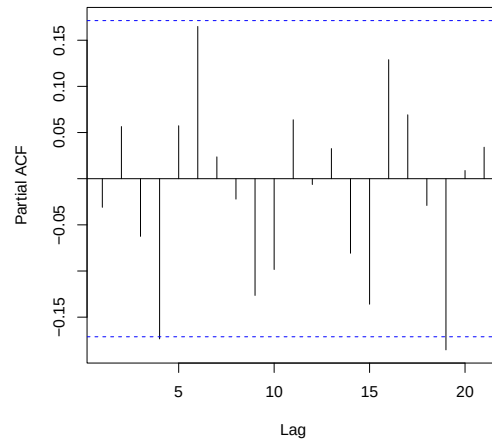
(a) Resíduos Quantílicos



(b) QQ-plot



(c) ACF



(d) PACF

Figura C2. Diagnóstico, do modelo ajustado, baseado nos resíduos quantílicos - Método 2.

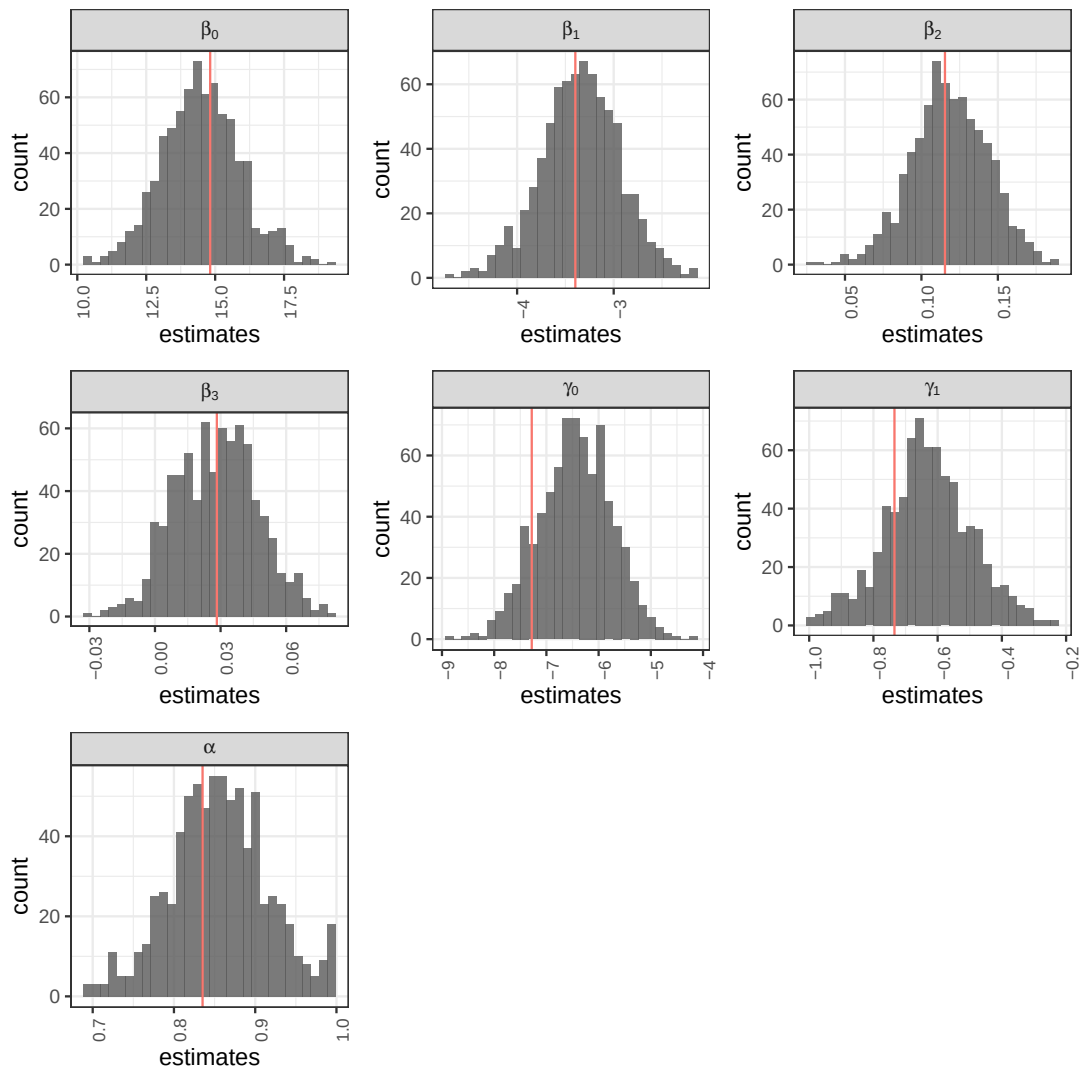


Figura C3. Distribuições dos estimadores dos parâmetros do modelo ajustado considerando o método 2 de estimação. Linha vertical vermelha representa o valor estimado

CAPÍTULO 5

Artigo 2

Modelo Multivariado Espaço-Temporal Mín-Estável para Indicadores de Climas Extremos

NASCIMENTO, L. B. F^a, LIMA, M.S. ^{a*}, and DUCZMAL, L H^b

^a*Departamento de Estatística, Universidade Federal do Amazonas, 69077-000, Manaus, AM, Brazil, Tel.: +55-92-33052850, Fax: +55-92-33052840;*

^a*Departamento de Estatística, Universidade Federal de Minas Gerais, Belo Horizonte, MG, Brazil*

Resumo

Predizer a variação espacial e temporal de indicadores para valores extremos climáticos (Temperatura, precipitação, umidade e velocidade do vento) é crucial para engenharia, agricultura, meio ambiente, saúde e quase toda atividade humana. Em particular, a saúde humana é negativamente afetada pela grande variação em indicadores climáticos que, geralmente, são limitados a um intervalo (a,b). Processos de natureza climática e ambiental, são medidos em estações meteorológicas e tem estrutura multivariada caracterizada por dependência espacial e/ou temporal. Por isso, é proposto, neste artigo, um modelo estatístico multivariado para predição espaço-temporal de indicadores (taxas, proporções) de valores extremos dependentes temporalmente e/ou espacialmente usados para modelar valores climáticos limitados em algum intervalo. A estrutura espacial do modelo é construída marginalizando uma distribuição baseline G condicionada a um campo aleatório latente com distribuição α -estável espacialmente compartilhada. A dependência temporal é introduzida por meio de uma função de ligação autorregressiva e média móvel (ARMA) na mediana do processo. Algumas propriedades do modelo são apresentadas, procedimentos para estimação e inferência são discutidos, e um algoritmo MCEM condicional é desenvolvido para estimar os parâmetros. Como aplicação, o modelo proposto é utilizado para predição espaço-temporal da umidade relativa do ar observada nas estações meteorológicas no estado do Amazonas, Brasil. Todas as rotinas para implementação do método foram escritas pelos autores na linguagem R Core Team (2023) e estão disponíveis sob consulta para reprodução dos resultados.

Palavras chaves

Processos Mín-Estáveis; Distribuição Estável; Predição Espacial; Estações meteorológicas; umidade.

1. Introdução

Processos climáticos e ambientais, como precipitação, temperatura e umidade relativa, medidos em estações meteorológicas, são multivariados e caracterizados por variabilidade espacial e/ou temporal (Papalexiou and Serinaldi (2020); Sun et al. (2012)). A complexidade da modelagem espacial e espaço-temporal de processos dessa natureza tem sido abordada por meio de campos aleatórios para valores extremos (Bopp et al., 2021; Cooley et al., 2010; Opitz, 2016; Papalexiou et al., 2021; Reich et al., 2014) campos Gaussianos e extensões (Alegría et al., 2017; Rue and Held, 2005; Zhang and El-Shaarawi, 2010), campos aleatórios de Laplace (Opitz, 2016; Saber et al.,

2021), cópulas (Carreau and Toulemonde, 2020; Krupskii and Genton, 2019) e campos aleatórios para dados positivos e limitados inferiormente (Mota et al. (2018)). Esses campos espaciais e/ou espaço-temporais têm uma longa história na modelagem de tais processos e, no entanto, há poucos campos aleatórios que oferecem flexibilidade suficiente para modelar valores extremos não-estacionários no espaço-tempo e limitados em algum intervalo do tipo (a,b). A modelagem espacial e/ou espaço-temporal de indicadores de valores extremos (como proporções e taxas extremas) é o foco deste paper onde, por exemplo, uma variável de interesse é a umidade relativa do ar pertencente ao (0,1), medida continuamente no tempo em várias localizações espaciais (estações meteorológicas), a qual é crucial para engenharia, agricultura, meio ambiente e saúde humana. Em particular, a saúde humana é negativamente afetada por essa variável. Uma umidade alta pode desencadear problemas respiratórios, alergias, sinusite e asma, além de favorecer a rápida proliferação de fungos e bactérias. Quando a umidade está muito baixa, ocorre estresse térmico, sensação desagradável, elevação da temperatura corporal, doenças térmicas leves e aquelas relacionadas à desidratação e hipertermia De Camargo (2007). Portanto, é importante modelar no espaço-tempo a ocorrência de umidade extrema. A teoria para valores extremos com suporte limitado foi proposta em Kunin (2004) e apresenta uma versão unificada chamada distribuição p-max estável proposta em Peng et al. (2015). Distribuições min-estáveis e max-estáveis multivariadas têm sua representação dada em Marshall and Olkin (1983). Processos max-stable foram usados pela primeira vez em análise espacial de valores extremos por Smith (1990) e desde então, várias modelagens de extremos espaciais via processos max-estáveis foram propostas e uma revisão destes modelos pode ser vista em Mathieu (2013). Uma classe de distribuições Multivariadas para valores extremos é apresentada em Fougères et al. (2010). Modelos marginais max-estáveis para extremos dependentes baseados na distribuição Gumbel e usando mistura de distribuições estáveis positivas como efeitos aleatórios foram propostos em Fougères et al. (2009). Essa abordagem max-estável é estendida em Reich and Shaby (2012) através de uma estrutura hierárquica Bayesiana. No contexto de análise de sobrevivência, Crowder (1989) propôs um modelo multivariado min-estável com marginais Weibul para certas escolhas dos parâmetros de interesse e um *Power Min-Stable Models for Time Series of Bounded Extremes* tem sido proposto em Nascimento et al. (2023). Nós estendemos esta abordagem para acomodar a dependência entre os mínimos extremos no espaço-tempo, além de ajustes por covariáveis e efeitos sazonais. Nossa modelagem é dinâmica, permitindo a predição espaço-temporal em locais e tempos onde não foram medidas a variável de interesse. Para ilustrar o tipo de modelagem proposta, considere uma região $\mathcal{S} \in \mathcal{R}^2$ na qual somente para um conjunto de posições fixas s_l , para $l \in \{1, \dots, L\}$, e tempos $t \in \{1, \dots, T\}$ são conhecidas as medidas de interesse de um processo estocástico no espaço-tempo $\mathbf{Y}_t = \{Y_t(s_1), \dots, Y_t(s_L)\}$. O objetivo é fornecer a estimativa $Y_t(s_0)$ a qualquer localização $s_0 \in \mathcal{S}$ e $t \geq 1$, onde não se conhece o valor da medida de interesse, a partir do processo observado $\mathbf{Y}_t$. Em nosso modelo min-estável assumimos que $Y_t(\mathbf{s})$ é o mínimo de um bloco, ou seja, o mínimo de muitas observações feitas no local $\mathbf{s}$ no tempo t , como por exemplo o mínimo mensal de indicadores diários como taxas, proporções ou índice extremos de umidade relativa do ar. Sem perda de generalidade, assumamos que $Y_t(s_l) \in (0, 1)$ e considere uma função de distribuição qualquer denotada por G . Admitimos a seguinte representação condicional para o processo espaço-temporal: $Y_t(s) = [1 - G(u_t^{z_t(s)})]^{1/\nu_t(s)} \in (0, 1)$. Onde $u_t, t \in \{1, \dots, T\}$, são i.i.d $\sim U(0, 1)$, $\nu_t(s)$ é uma função da mediana do processo assumindo uma estrutura ARMA e $z_t(s)$ é o processo latente introduzido para acomodar a variabilidade espacial local e temporal de larga escala. A função de excedência condicional do processo é

da forma $\Pr(Y_t(s) > y | z_t(s)) = [G^{-1}(1 - y^{\nu_t(s)})]^{z_t(s)}$. As escolhas de G e $z_t(s)$ nos levam a uma classe de modelos multivariados com interação espaço-temporal induzida pela lei de potência $u_t^{z_t(s)}$. Dependendo dessas escolhas, a correlação temporal na mediana é introduzida por meio de uma função de ligação autorregressiva. Quando $Y_t(s) \in (a, b)$, podemos usar sua versão reescalada $(Y_t(s) - a)/(b - a)$. Concentramos nossos esforços para desenvolver o modelo quando o processo $z_t(s)$ é espacialmente compartilhado Mota et al. (2018) com distribuição estável positiva que nos leva a um modelo marginal Multivariado Min-Stable para valores extremos limitados. O modelo que desenvolvemos no restante deste artigo permite uma condicional representação em termos de efeitos aleatórios espacialmente compartilhados que torna possível o desenvolvimento de um algoritmo MCEM (Monte Carlo Expectation-Maximization) para estimação dos parâmetros e uma simulação condicional via *Bootstrap* para inferência, que é útil para previsão em locais onde as medidas de interesse não são observadas. No que segue, este artigo está organizado da seguinte forma: na Seção 2 apresentamos o modelo proposto e suas propriedades. A Seção 3 aborda a Estimativa dos Parâmetros e Predição no espaço-tempo. Na Seção 4, ilustramos a metodologia na modelagem de dados reais e um estudo de validação cruzada é elaborado para comparar a modelagem proposta com outro modelo mostrando seu potencial. Por fim, as considerações finais e conclusões são discutidas na Seção 5.

2. Construção do modelo proposto

2.1. Modelo com efeito aleatório espacialmente compartilhado e Excedência Marginal

Sejam $\mathbf{s}_1, \dots, \mathbf{s}_L$ um conjunto de L localizações espaciais em $\mathcal{S} \in \mathcal{R}^2$ (considerado aqui como localizações de estações meteorológicas), onde para cada tempo t temos o processo $\mathbf{Y}_t(\mathbf{s}) = (Y_t(\mathbf{s}_1), Y_t(\mathbf{s}_2), \dots, Y_t(\mathbf{s}_L))$ e $\mathcal{F}_{t-1} = \{\mathbf{Y}_{t-1}(\mathbf{s}), \mathbf{Y}_{t-2}(\mathbf{s}), \dots, \mathbf{x}_t(\mathbf{s}), \mathbf{x}_{t-1}(\mathbf{s}), \dots, \mathbf{w}_t(\mathbf{s}), \mathbf{w}_{t-1}(\mathbf{s}), \dots\}$ é a sigma-álgebra gerada por valores passados da série resposta e, também, valores passados e presentes de um conjunto conhecido de covariáveis (Kedem and Fokianos, 2005). A variável resposta, ou sua versão reescalada, $Y_t(\mathbf{s}_l) \in (0, 1)$ é um indicador de valor mínimo medido em uma semana, mês, etc... Considere que $\mathcal{H} = \{\mathbf{h}_1, \dots, \mathbf{h}_M\}$ é um conjunto com M localizações espaciais no domínio de $\mathcal{S}$ e $\mathbf{Z}_t(\mathbf{s}) = \{\omega_m(\mathbf{s})Z(\mathbf{h}_m)/\delta_t\}$, para $m \in \{1, \dots, M\}$. Na linguagem da estatística espacial, dizemos que $\mathbf{Y}_t(\mathbf{s})$ é o campo aleatório observado e $\mathbf{Z}_t(\mathbf{s})$ é um campo aleatório latente ou não observado com constantes, $0 < \omega_m(\mathbf{s}) < 1$, representando uma medida de distância entre o par de localizações $(\mathbf{s}, \mathbf{h}_m)$. No conjunto $\{\mathbf{h}_1, \dots, \mathbf{h}_M\}$ não há medição de dados. Mas, ele pode ser uma fonte de uma informação importante para o processo. Na modelagem de dados meteorológicos, podemos pensar em $Z(\mathbf{h}_m)$ como um efeito espacialmente compartilhado por todas as medidas obtidas nas estações meteorológicas próximas de $\mathbf{h}_m$. Nesse caso em $\mathbf{Z}_t(\mathbf{s})$, $Z(\mathbf{h}_m)$ pode representar um efeito de uma cobertura vegetal, efeito de um curso de rio ou, ainda, efeito do solo com $\omega_m(\mathbf{s}_l)$ controlando o alcance ou a contribuição desse efeito na medida de interesse na l -ésima estação meteorológica e δ_t monitora a dispersão temporal do efeito. Se $Z(\mathbf{h}_m)$ é espacialmente uniforme, podemos escolher $\mathbf{h}_m$ como o baricentro da subregião em $\mathcal{S}$ que contém o grupo de estações mais próximas $\mathbf{h}_m$ (veja, figura 3 para detalhes). Para acomodar a dependência espacial, seja $\mathbf{s}_l \sim \mathbf{h}_m$ o conjunto de estações meteorológicas mais próximas de $\mathbf{h}_m$ do que de qualquer outro ponto $\mathbf{h} \in \mathcal{H}$. Definimos $K(d_m(\mathbf{s}_l); \tau_m)$ como sendo o valor de uma função núcleo Gaussiano centrada em

$d_m(s_l)$, onde $d_m(s_l) = \log(\|\mathbf{s}_l - \mathbf{h}_m\|)$ é o logaritmo da distância Euclidiana entre as localização s_l da l -ésima estação meteorológica e o ponto $\mathbf{h}_m$. O parâmetro τ_m , assumido conhecido, é chamado de parâmetro de suavização da função núcleo e obtido de maneira similar a descrita em (Bowman and Azzalini, 1997). Os valores $\omega_m(s_l)$ são então definidos por $\omega_m(s_l) = K(d_m(s_l); \tau_m) / \sum_{s_j \sim \mathbf{h}_m} K(d_m(s_j); \tau_m)$, $0 < \omega_m(s_l) < 1$. Completamos o modelo admitindo a seguinte representação condicional para o processo com (t, s) fixos: $Y_t(s) = [1 - G(u_t^{z_t(s)})]^{1/\nu_t(s)} \in (0, 1)$. Onde u_t são i.i.d $\sim U(0, 1)$, $\nu_t(s)$ é uma função da mediana do processo e $\{Z(\mathbf{h}_1), \dots, Z(\mathbf{h}_M)\}$ um campo aleatório latente construído por variáveis independentes α_m -Estável com $E(e^{-aZ_m}) = e^{-a^{\alpha_m}}$, $\alpha_m \in (0, 1)$ e $a > 0$ Uchaikin and Zolotarev (1999). Então o processo $\mathbf{Y}_t(\mathbf{s})$ possui distribuição de excedência conjunta marginal

$$\bar{F}(\mathbf{y}_t(\mathbf{s}); \boldsymbol{\theta}_t(\mathbf{s}), \boldsymbol{\alpha}) = \exp \left[- \sum_{m=1}^M \left\{ \sum_{\mathbf{s}_l \sim \mathbf{h}_m} \left(\frac{-\omega_m(s_l) \log G^{-1}(1 - y_t(s_l)^{\nu_t(s_l)})}{\delta_t} \right) \right\}^{\alpha_m} \right], \quad (1)$$

com $\boldsymbol{\theta}_t(\mathbf{s}) = (\nu_t(s_l), \delta_t)$, $l \in \{1, \dots, L\}$, e $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \dots, \alpha_M)$. O campo espacial de excedência acumulado para o modelo é $-\log[\bar{F}(\mathbf{y}_t(\mathbf{s}); \boldsymbol{\theta}_t(\mathbf{s}), \boldsymbol{\alpha})] = \sum_{m=1}^M \Omega_{m,t}$, onde

$$\Omega_{m,t} = \left\{ \sum_{\mathbf{s}_l \sim \mathbf{h}_m} \left(\frac{-\omega_m(s_l) \log G^{-1}(1 - y_t(s_l)^{\nu_t(s_l)})}{\delta_t} \right) \right\}^{\alpha_m}.$$

A distribuição do mínimo $Y_{t,0}$ sobre qualquer subconjunto $\mathcal{S}_0 \subset \mathcal{S}$ pertence à mesma classe marginal. Para verificar isso, observe que se $Y_{t,0} = \min_{s_l \in \mathcal{S}_0} \mathbf{Y}_t(\mathbf{s})$. Então a função de excedência $\bar{F}(y_{t,0}; \boldsymbol{\theta}_t(s_l), \alpha_m)$ é simplesmente $\mathbb{E}_{\mathbf{Z}}[P(\min_{s_l \in \mathcal{S}_0} \mathbf{Y}_t(\mathbf{s}) > y_{t,0} | \mathbf{Z})]$ que é dado por

$$\bar{F}(y_{t,0}; \boldsymbol{\theta}_t(s_l), \alpha_m) = \exp \left[- \sum_{m=1}^M \left\{ \sum_{\{\mathbf{s}_l \in \mathcal{S}_0\} \sim \mathbf{h}_m} \left(\frac{-\omega_m(s_l) \log G^{-1}(1 - y_{t,0}^{\nu_t(s_l)})}{\delta_t} \right) \right\}^{\alpha_m} \right],$$

que pertence à mesma família de distribuição em (1). Se $\mathcal{S}_0 = \{s_l\}$ é um ponto com m fixo. Então, a distribuição marginal univariada para cada $Y_t(s_l)$ é expressa por

$$\bar{F}(y_t(s_l); \boldsymbol{\theta}_t(s_l), \alpha_m) = \exp \left\{ - \left(\frac{-\omega_m(s_l) \log G^{-1}(1 - y_t(s_l)^{\nu_t(s_l)})}{\delta_t} \right)^{\alpha_m} \right\}, \quad (2)$$

$\forall \mathbf{s}_l \sim \mathbf{h}_m$.

2.2. Modelo Espaço-Temporal Mín-Estável para predição de extremos limitados

Um caso particular de interesse do modelo em (1) abordado aqui, ocorre quando $G(a) = a$, é a função da distribuição uniforme no intervalo $(0, 1)$. Assim, a excedência

marginal multivariada é a da forma

$$\bar{F}(\mathbf{y}_t(\mathbf{s}); \boldsymbol{\theta}_t(\mathbf{s}), \boldsymbol{\alpha}) = \exp \left[- \sum_{m=1}^M \left\{ \sum_{\mathbf{s}_l \sim \mathbf{h}_m} \left(\frac{-\omega_m(s_l) \log(1 - y_t(s_l)^{\nu_t(s_l)})}{\delta_t} \right) \right\}^{\alpha_m} \right], \quad (3)$$

e para cada m e t fixos o modelo espacial em (3) é Power Min-Stable. A prova dessa propriedade é similar a descrita em Nascimento et al. (2023). Por (2) a excedência marginal univariada é

$$\bar{F}(y_t(s_l); \boldsymbol{\theta}_t(s_l), \alpha_m) = \exp \left\{ - \left(\frac{-\omega_m(s_l) \log(1 - y_t(s_l)^{\nu_t(s_l)})}{\delta_t} \right)^{\alpha_m} \right\}, \quad \forall \quad \mathbf{s}_l \sim \mathbf{h}_m. \quad (4)$$

de modo que para $u \sim U(0, 1)$, podemos representar,

$$y_t(s) = \left[\underbrace{1 - \exp \left\{ - \left(\frac{-\log(1 - u)}{(\omega_m(s_l)/\delta_t)^{\alpha_m}} \right)^{1/\alpha_m} \right\}}_{b_{t,s}} \right]^{1/\nu_t(\mathbf{s})}, \quad (5)$$

com mediana $\kappa_t(\mathbf{s})$ igual a

$$\kappa_t(s) = \left[\underbrace{1 - \exp \left\{ - \left(\frac{-\log(0.5)}{(\omega_m(s_l)/\delta_t)^{\alpha_m}} \right)^{1/\alpha_m} \right\}}_{c_{t,s}} \right]^{1/\nu_t(\mathbf{s})}.$$

Pela aproximação de primeira ordem, $E[y_t(s) - \kappa_t(s)] \approx 0$. Considere $\Phi(a) = -\log(-\log(a))$ e $\nu_t(\mathbf{s}) = \nu_{\mathbf{s}} \times \nu_t$ separáveis no espaço-tempo. Então,

$$\Phi(y_t(s)) = \Phi(\kappa_t(s)) + \underbrace{\zeta_{t,s}}_{\Phi(\kappa_t(s))} = \log(\nu_{\mathbf{s}}) + \log(\nu_t) + \zeta_{t,s},$$

onde $\zeta_{t,s} = \Phi(b_{t,s}) - \Phi(c_{t,s})$ e, novamente, pela aproximação de primeira ordem a $E[\zeta_{t,s}] \approx 0$. No contexto de um modelo linear generalizado com componente temporal auto-regressivo e média móvel **ARMA**(1, 1) introduzido na mediana, podemos escrever

$$\Phi(\kappa_t(\mathbf{s})) = \beta_0 + \mathbf{x}_t^\top(\mathbf{s})\beta + \psi[\Phi(y_{t-1}(\mathbf{s}_l)) - \mathbf{x}_{t-1}^\top(\mathbf{s})\beta] + \rho\epsilon_{t-1,s},$$

no qual $\log(\nu_{\mathbf{s}}) = \eta_{\mathbf{s}} = \mathbf{x}_{\mathbf{s}}^\top \vartheta$ é o preditor para tendência espacial, $\log(\nu_t) = \eta_t = \mathbf{x}_t^\top \varphi$ é o preditor para tendência temporal, β_0 é uma tendência global que independe do espaço-tempo, ψ é o coeficiente autorregressivo e ρ representa o parâmetro de média móvel, $\epsilon_{t-1,s} = \Phi(y_{t-1}(s)) - \Phi(\kappa_{t-1}(s))$. Nessa representação, $\mathbf{x}_t(\mathbf{s}) = (\mathbf{x}_{\mathbf{s}}, \mathbf{x}_t)$ e $\beta = (\vartheta, \varphi)^\top$. Uma regressão $\log \delta_t = \mathbf{w}_t^\top \gamma$ é assumida para alguma covariável $\mathbf{w}_t$ associada ao vetor de coeficientes γ . Assim, a estrutura do modelo proposto é dinâmica condicionada em $\mathcal{F}_{t-1}$ com dependência temporal na função de ligação da mediana.

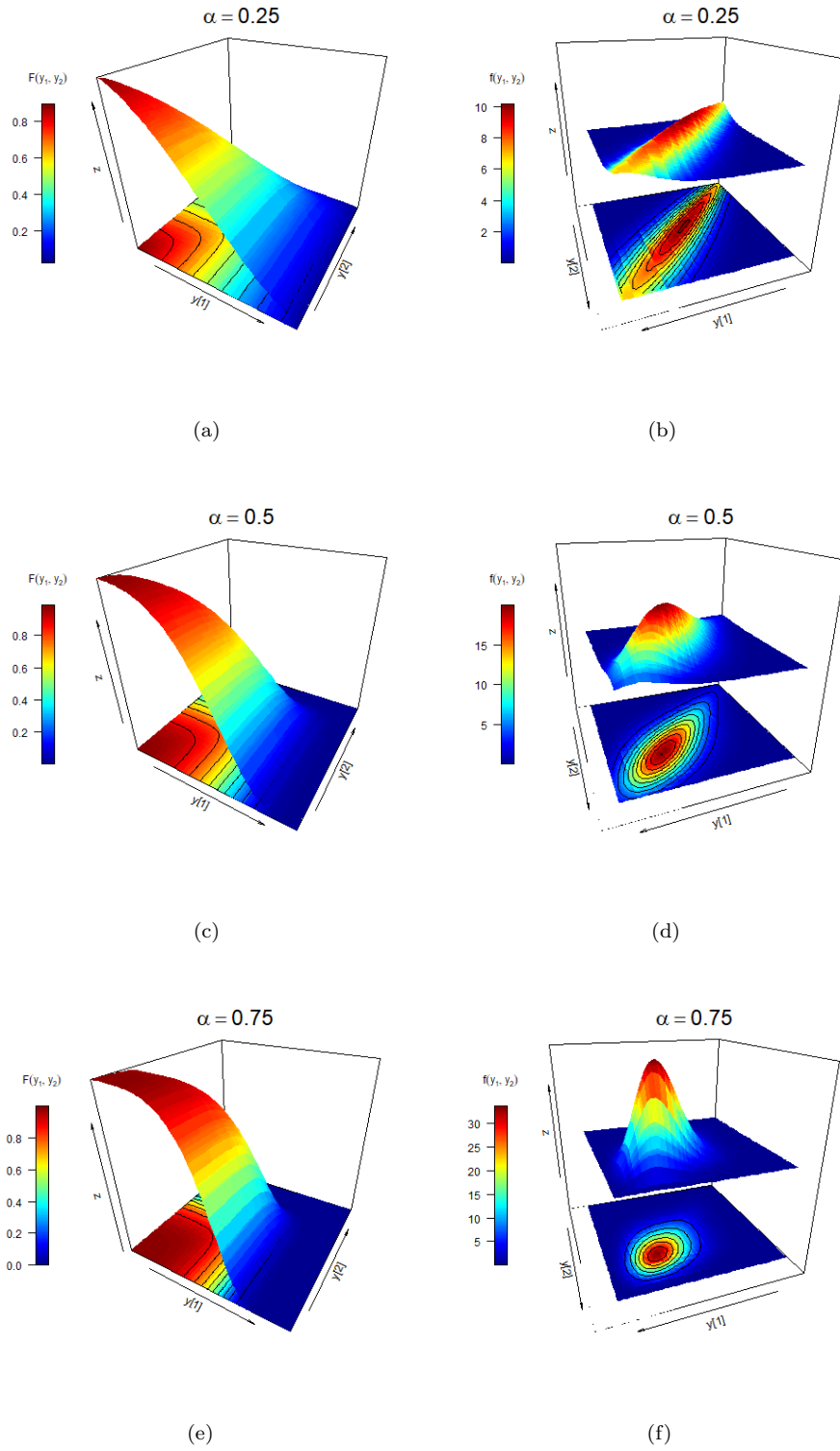


Figura 1. Função de excedência bivariada (lado esquerdo) e função densidade (lado direito) para uma realização do modelo espacial Min-Stable proposto com diferentes valores de α .

Apresentamos na Fig. 1, um gráfico da distribuição de excedência bivariada desse modelo para uma realização do processo através do par $(y_t(s_1), y_t(s_2)) \equiv (y_1, y_2)$ com mediana 0.7, $\delta = 0.01$, $\omega(s_1) = \omega(s_2) = 0.5$, valores de $\alpha \in \{0.25, 0.50, 0.75\}$ e sem ajuste por covariáveis. Observe que o formato da curva da densidade conjunta se torna menos elíptico à medida que $\alpha \rightarrow 1$. Esse comportamento indica que há uma diminuição na relação de dependência entre os extremos das variáveis. Estas características de dependência no modelo são discutidas na próxima seção.

2.3. Propriedades de Dependência

Nesta seção apresentaremos algumas propriedades de dependência espacial do modelo marginal proposto em (3) e mostramos que o parâmetro α desempenha um papel crucial na determinação do tipo de dependência.

(1) Dependência Espaço-Temporal Induzida:

A dependência espaço-temporal induzida no modelo pelo efeito aleatório espacialmente compartilhado pode ser obtida através da covariância marginal entre $\Psi(y_t(s_l), \nu_t(s_l))$ e $\Psi(y_t(s_j), \nu_t(s_j))$ tal que $(s_l, s_j) \sim \mathbf{h}_m$. Onde, $\Psi(y_t(\mathbf{s}), \nu_t(s)) = -\log(1 - y_t(s)^{\nu_t(s)})$. Note, pela distribuição marginal em (4), que

$$\Psi(y_t(s_l), \nu_t(s_l)) \sim \text{Weibull}(\omega_m(s_l)/\delta_t, \alpha_m).$$

Então, $\Psi(y_t(s_l), \nu_t(s_l))|Z(\mathbf{h}_m) \sim \exp(\omega_m(s)Z(\mathbf{h}_m)/\delta_t)$. Onde $Z(\mathbf{h}_m)$ é uma α_m -estável positiva com variância de $1/Z(\mathbf{h}_m)$ igual a $0.5\Gamma(1 + 2\alpha_m^{-1}) - \Gamma^2(1 + \alpha_m^{-1})$, veja Williams (1977). Pela independência condicional em $Z(\mathbf{h}_m)$, temos que $E\{\text{Cov}[\Psi(y_t(s_l), \nu_t(s_l)), \Psi(y_t(s_j), \nu_t(s_j))]|Z(\mathbf{h}_m)\} = 0$. Então, podemos escrever

$$\begin{aligned} \text{Cov}[\Psi(y_t(s_l), \nu_t(s_l)), \Psi(y_t(s_j), \nu_t(s_j))] &= \text{Cov}[E\{\Psi(y_t(s_l), \nu_t(s_l))|Z(\mathbf{h}_m)\}, \\ &\quad E\{\Psi(y_t(s_j), \nu_t(s_j))|Z(\mathbf{h}_m)\}]. \end{aligned}$$

Portanto,

$$\begin{aligned} \text{Cov}[\Psi(y_t(s_l), \nu_t(s_l)), \Psi(y_t(s_j), \nu_t(s_j))] &= \frac{\delta_t^2}{\omega_m(s_l)\omega_m(s_j)} \text{Var}(1/Z(\mathbf{h}_m)) \\ &= \frac{\delta_t^2}{\omega_m(s_l)\omega_m(s_j)} \\ &\quad \times [0.5\Gamma(1 + 2\alpha_m^{-1}) - \Gamma^2(1 + \alpha_m^{-1})]. \end{aligned}$$

Note que $\alpha_m \rightarrow 1$. Então $\text{Cov}[\Psi(y_t(s_l), \nu_t(s_l)), \Psi(y_t(s_j), \nu_t(s_j))] \rightarrow 0$ e com $\alpha_m \in (0, 1)$ temos

$$\text{Cov}[\Psi(y_t(s_l), \nu_t(s_l)), \Psi(y_t(s_l), \nu_t(s_j))] > 0$$

e não estacionária. Como $\Psi(y_t(s_l), \nu_t(s_l))$ é uma função não decrescente em $y_t(\mathbf{s}_l)$ segue que $y_t(\mathbf{s}_l)$ e $y_t(\mathbf{s}_j)$ são espacialmente dependentes (Shaked, 1982). Desde que $0 < \omega_m(s_l) \times \omega_m(s_j) < 1$. Quanto mais próximo $\mathbf{s}_l$ e $\mathbf{s}_j$ maior é a covariância.

(2) Função de excedência cruzada bivariada:

Sejam t e m fixos, e $(\mathbf{s}_l, \mathbf{s}_j) \sim \mathbf{h}_m$. Para o par $(y_t(s_l), y_t(s_j))$ que compartilham o

mesmo efeito aleatório $Z(\mathbf{h}_m)$ considere a função de excedência cruzada definida por,

$$\vartheta_{l,j} = \frac{r(y_t(s_l)|Y_t(s_j) = y_j)}{r(y_t(s_l)|Y_t(s_j) \geq y_j)} = \frac{\bar{F} \times \bar{F}_{l,j}}{\bar{F}_l \times \bar{F}_j}.$$

Alguns autores em Clayton (1978); Gupta (2018); Oakes (1989) usam essa medida no contexto da análise de sobrevivência. Aqui, esta medida pode ser interpretada como uma comparação entre a taxa de risco de excedência da distribuição condicional de $Y_t(s_l)$, dado que $Y_t(s_j)$ é igual a y_j , versus a taxa de risco dado que $Y_t(s_j)$ excedeu y_j . Onde, $\bar{F} = \bar{F}(y_t(s_l), y_t(s_j); \boldsymbol{\theta}_t, \alpha_m)$ é interpretada como a função de excedência conjunta bivariada, $\bar{F}_{l,j} = \partial^2 \bar{F} / \partial y_t(s_l) \partial y_t(s_j)$, $\bar{F}_l = \partial \bar{F} / \partial y_t(s_l)$, $\bar{F}_j = \partial \bar{F} / \partial y_t(s_j)$ e r é a taxa de risco de excedência. Sem perda de generalidade, defina $\Lambda_{t,l} = \Psi(y_t(s_l), \nu_t(s_l)) \omega_m(s_l) / \delta_t$ e $\Lambda_{t,j} = \Psi(y_t(s_j), \nu_t(s_j)) \omega_m(s_j) / \delta_t$ com as respectivas derivadas $\lambda_{t,l} = \partial \Lambda_{t,l} / \partial y_t(s_l)$ e $\lambda_{t,j} = \partial \Lambda_{t,j} / \partial y_t(s_j)$. Portanto, segue da distribuição dada na expressão (3) que:

$$\bar{F} = \exp \{ - (\Lambda_{t,l} + \Lambda_{t,j})^{\alpha_m} \},$$

$$\bar{F}_l = -\alpha_m (\Lambda_{t,l} + \Lambda_{t,j})^{\alpha_m - 1} \exp \{ - (\Lambda_{t,l} + \Lambda_{t,j})^{\alpha_m} \} \lambda_{t,l},$$

$$\bar{F}_j = -\alpha_m (\Lambda_{t,l} + \Lambda_{t,j})^{\alpha_m - 1} \exp \{ - (\Lambda_{t,l} + \Lambda_{t,j})^{\alpha_m} \} \lambda_{t,j}$$

e

$$\begin{aligned} \bar{F}_{l,j} &= \lambda_j \lambda_l \alpha_m (\Lambda_{t,l} + \Lambda_{t,j})^{\alpha_m - 2} \exp \{ - (\Lambda_{t,l} + \Lambda_{t,j})^{\alpha_m} \} \\ &\quad \times (\alpha_m (\Lambda_{t,l} + \Lambda_{t,j})^{\alpha_m} + (1 - \alpha_m)). \end{aligned}$$

Logo, o modelo marginal bivariado em (3) possui a seguinte excedência cruzada:

$$\begin{aligned} \vartheta_{l,j} &= \frac{\bar{F} \times \bar{F}_{l,j}}{\bar{F}_l \times \bar{F}_j} \\ &= \frac{\alpha_m^2 (\Lambda_{t,l} + \Lambda_{t,j})^{2\alpha_m - 2} + \alpha_m (\Lambda_{t,l} + \Lambda_{t,j})^{\alpha_m - 2} (1 - \alpha_m)}{\alpha_m^2 (\Lambda_{t,l} + \Lambda_{t,j})^{2\alpha_m - 2}} \\ &= 1 + \frac{1 - \alpha_m}{\alpha_m (\Lambda_{t,l} + \Lambda_{t,j})^{\alpha_m}} \\ &= 1 + \frac{(1 - \alpha_m)}{\alpha_m \left(\frac{\Psi(y_t(s_l), \nu_t(s_l))}{\delta_t / \omega_m(s_l)} + \frac{\Psi(y_t(s_j), \nu_t(s_j))}{\delta_t / \omega_m(s_j)} \right)^{\alpha_m}}, \quad \forall \quad (\mathbf{s}_l, \mathbf{s}_j) \sim \mathbf{h}_m. \end{aligned}$$

Se $\alpha_m \rightarrow 1$, então $\vartheta_{l,j} \rightarrow 1$ e $y_t(s_l)$ com $y_t(s_j)$ são independentes (Gupta, 2018). A Fig. 2 apresenta o comportamento da excedência cruzada para uma realização do processo com mediana 0.7, $\delta = 0.01$, $\omega(s_1) = \omega(s_2) = 0.5$ e valores de $\alpha \in \{0.15, 0.25, 0.50, 0.75, 0.85, 0.95\}$. Observamos que a dependência nos valores

mínimos cresce quando α diminui. Isso mostra que o modelo é adequado para modelar dependência nos mínimos extremos.

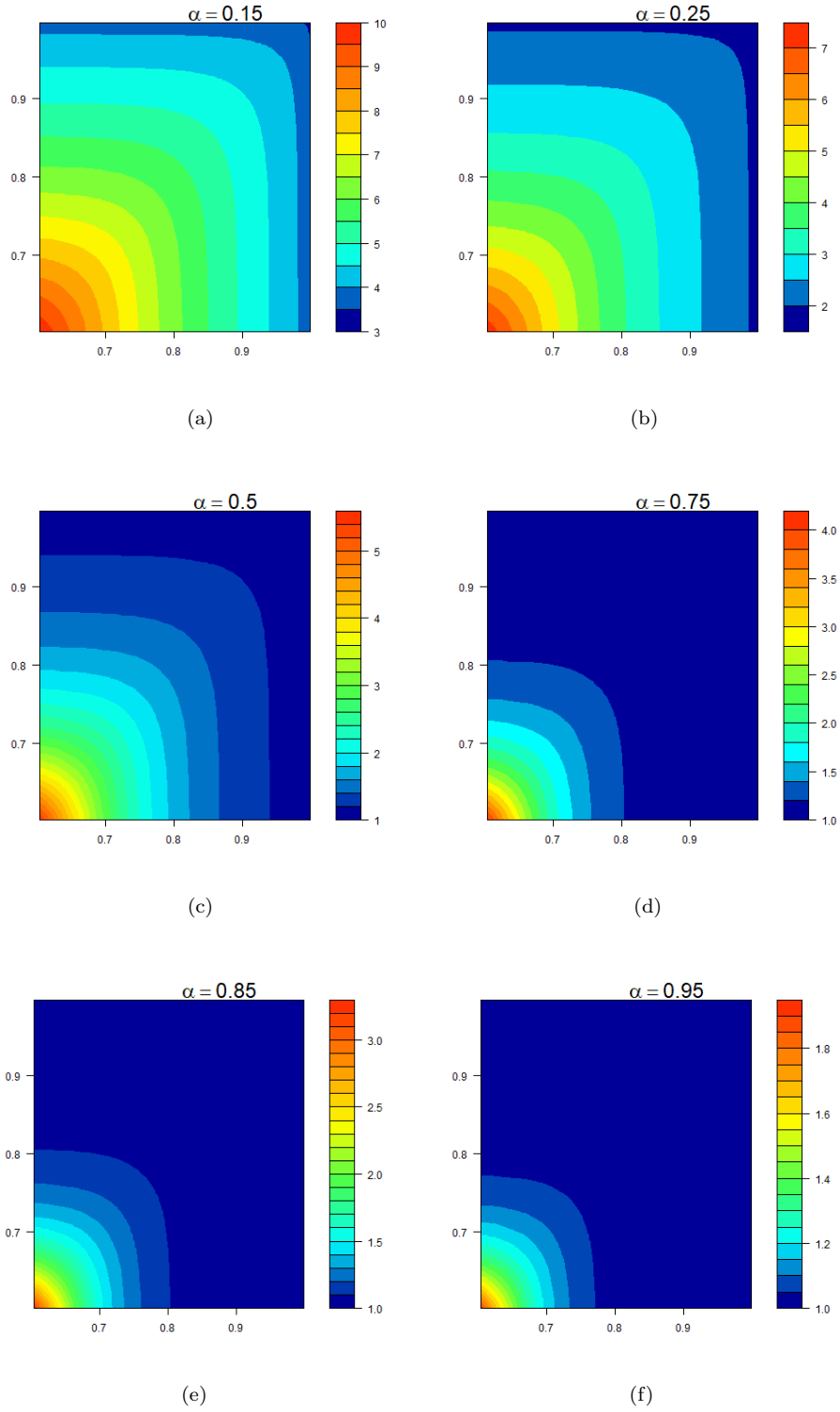


Figura 2. Função de excedência cruzada bivariada para diferentes valores de α

(3) Extremograma Espacial:

Outra características de dependência espacial do modelo proposto é a quantificação da dependência espacial extrema representada pela probabilidade condicional de excedência conjunta de um quantil fixo chamado de Spatial Extremogram (Coles et al., 1999) denotada por $\chi(y_u) = \Pr(Y_t(s_l) > y_u | Y_t(s_j) > y_u)$ para $(s_l, s_j) \sim \mathbf{h}_m$, onde y_u é um ponto de alta excedência (ou limite superior de excedência). Nesse caso, temos que

$$\begin{aligned} \chi(y_u) &= \exp \left\{ \left(\frac{\omega_m(s_j)}{\delta_t} \Psi(y_u, \nu_t(s_j)) \right)^{\alpha_m} \right. \\ &\quad \left. - \left(\frac{\omega_m(s_l)}{\delta_t} \Psi(y_u, \nu_t(s_l)) + \frac{\omega_m(s_j)}{\delta_t} \Psi(y_u, \nu_t(s_j)) \right)^{\alpha_m} \right\} \\ &= \exp \left[\left\{ \frac{\omega_m(s_j)}{\delta_t} \Psi(y_u, \nu_t(s_j)) \right\}^{\alpha_m} \right. \\ &\quad \left. \times \left\{ 1 - \left(1 + \frac{\omega_m(s_l) \Psi(y_u, \nu_t(s_l))}{\omega_m(s_j) \Psi(y_u, \nu_t(s_j))} \right)^{\alpha_m} \right\} \right]. \end{aligned}$$

Note que se $\alpha \rightarrow 0$ implica que $\chi(y_u) \rightarrow 1$ e os extremos são totalmente dependentes. Quando $\alpha \rightarrow 1$, $\chi(y_u) \rightarrow \Pr(Y_t(s_l) > y_u)$ indicando a independência. Suponha que o par (s_l, s_j) está uma mesma distância $\mathbf{d}$ de $\mathbf{h}_m$, isto é, $\omega_m(s_l) = \omega_m(s_j) = \omega_m(\mathbf{d})$ e considere que o processo observado é espacialmente estacionário tal que $\nu_t(s_l) = \nu_t(s_j) = \nu_t$. Então, a probabilidade condicional de excedência simplifica em

$$\chi(y_u) = \exp \left\{ \left(-\frac{\omega_m(\mathbf{d})}{\delta_t} \log(1 - y_u^{\nu_t}) \right)^{\alpha_m} (1 - 2^{\alpha_m}) \right\},$$

que decresce com o aumento da distância $\mathbf{d}$, mostrando que $\chi(y_u)$ é uma medida de dependência espacial válida.

3. Inferência Estatística

3.1. Estimação dos parâmetros via Algoritmo MCEM

Para cada realização $t \in \{1, \dots, T\}$, do processo espaço-temporal $\mathbf{Y}_t(\mathbf{s})$ com estrutura temporal $\mathbf{ARMA}(1, 1)$ na mediana $\kappa_t(s)$ apresentado na Seção 2.2. A função de densidade conjunta de $\mathbf{Y}_t(\mathbf{s}) | \mathcal{F}_{t-1}$ é dada por,

$$f(\mathbf{y}_t(\mathbf{s}); \boldsymbol{\theta}_t(\mathbf{s}), \boldsymbol{\alpha} | \mathcal{F}_{t-1}) = (-1)^L \frac{\partial^L \bar{F}(\mathbf{y}_t(\mathbf{s}); \boldsymbol{\theta}_t(\mathbf{s}), \boldsymbol{\alpha} | \mathcal{F}_{t-1})}{\partial y_t(s_1), \dots, \partial y_t(s_L)}.$$

Essa densidade conjunta não possui uma forma analiticamente tratável, pois depende de algumas quantidades obtidas de forma recursiva e quando incorporamos covariáveis, os algoritmos de otimização convencionais tendem a falhar no processo de maximização para encontra o Estimador de Máxima Verossimilhança. Assim, o processo de maximização será realizado através de um algoritmo **MCEM** condicional, onde o passo

E é realizado via aproximação Monte Carlo e o passo **M** é a maximização de forma condicional da log-verossimilhança completa. Inicialmente, Considere o campo latente $\mathbf{Z}_t(\mathbf{s})$ com $Z(\mathbf{h}_m) \doteq Z_m \sim$ positive α_m -stable e $\mathbf{y}_c = (\mathbf{y}, \mathbf{z})$ o vetor dos dados completos no espaço-tempo. Seja $B(\mathbf{h}_m) \doteq B_m \in (0, 1)$ uma variável auxiliar tal que a distribuição marginal de Z_m é

$$p(Z_m; \alpha_m) = \frac{\alpha_m}{(1 - \alpha_m)} Z_m^{-\frac{1}{1-\alpha_m}} \int_0^1 c(B_m) e^{-c(B_m) Z_m^{-\frac{\alpha_m}{(1-\alpha_m)}}} dB_m \quad \text{para } Z_m > 0,$$

onde

$$c(B_m) = \left\{ \frac{\sin(\pi \alpha_m B_m)}{\sin(\pi B_m)} \right\}^{\frac{1}{1-\alpha_m}} \frac{\sin\{(1 - \alpha_m)\pi B_m\}}{\sin(\pi \alpha_m B_m)}.$$

Então, marginalmente a B_m , $Z_m \sim \alpha_m$ -estável (Stephenson, 2009). Neste caso, o modelo completo em (3) tem densidade

$$\begin{aligned} f(\mathbf{y}, \mathbf{Z}; \boldsymbol{\theta}, \boldsymbol{\alpha}) &= f(\mathbf{y}; \boldsymbol{\theta} | \mathbf{Z}) p(\mathbf{Z}; \boldsymbol{\alpha}) \\ &= \prod_{m=1}^M \left[\left\{ \prod_{t=1}^T \prod_{s_l \sim \mathbf{h}_m} \frac{\omega_m(s_l) Z_m}{\delta_t} (G^{-1}(1 - y_t(s_l)^{\nu_t(s_l)}))^{\frac{\omega_m(s_l) Z_m}{\delta_t}} \right. \right. \\ &\quad \left. \left. \times \frac{g(y_t(s_l); \nu_t(s_l))}{G^{-1}(1 - y_t(s_l)^{\nu_t(s_l)})} \right\} p(Z_m; \alpha_m) \right], \end{aligned}$$

onde $(G^{-1}(1 - y_t(s_l)^{\nu_t(s_l)})) = (1 - y_t(s_l)^{\nu_t(s_l)})$ e $g(y_t(s_l), \nu_t(s_l)) = \nu_t(s_l) y_t(s_l)^{\nu_t(s_l)-1}$ é a derivada da função interna em G^{-1} . Observe que expressamos a densidade completa em termos da generalização do modelo apresentado na Seção 2.1 de modo que o algoritmo proposto pode ser estendido para qualquer forma da função G . Em nosso caso específico, suprimindo a dependência em $\mathcal{F}_{t-1}$, temos que função de log-verossimilhança completa é dada por

$$\begin{aligned} \ell_c(\boldsymbol{\theta}, \boldsymbol{\alpha}; \mathbf{y}, \mathbf{Z}) &= \sum_{m=1}^M \left[T \sum_{s_l \sim \mathbf{h}_m} \log \omega_m(s_l) - Z_m v^{(m)}(\boldsymbol{\theta}) + \sum_{t=1}^T \sum_{s_l \sim \mathbf{h}_m} \nu_t(s_l) \log y_t(s_l) \right. \\ &\quad - \sum_{t=1}^T \sum_{s_l \sim \mathbf{h}_m} \log \left(1 - y_t(s_l)^{\nu_t(s_l)} \right) \\ &\quad + \sum_{t=1}^T \sum_{s_l \sim \mathbf{h}_m} \log \nu_t(s_l) + \sum_{t=1}^T \sum_{s_l \sim \mathbf{h}_m} \log y_t(s_l) - L_m \sum_{t=1}^T \log \delta_t \left. \right] \\ &\quad + \underbrace{\sum_{m=1}^M (\log p(Z_m; \alpha_m) + T L_m \log Z_m)}_{\ell_c(\boldsymbol{\alpha}; \mathbf{Z})} \\ &= \ell_c(\boldsymbol{\theta}; \mathbf{y}, \mathbf{Z}) + \ell_c(\boldsymbol{\alpha}; \mathbf{Z}), \end{aligned}$$

em que $v^{(m)}(\boldsymbol{\theta}) = \sum_{t=1}^T \sum_{s_l \sim \mathbf{h}_m} \omega_m(s_l) \frac{1}{\delta_t} [-\log(1 - y_t(s_l)^{\nu_t(s_l)})]$ e L_m é a cardinalidade do conjunto $\{s_l \sim \mathbf{h}_m\}$. Considerando $\tilde{\ell}_c(\boldsymbol{\theta}, \boldsymbol{\alpha}; \mathbf{y}, \mathbf{Z}) = \mathbb{E}_{\mathbf{Z}|\mathbf{y}, \boldsymbol{\theta}, \boldsymbol{\alpha}}(\ell_c(\boldsymbol{\theta}, \boldsymbol{\alpha}; \mathbf{y}, \mathbf{Z}))$, temos

$$\tilde{\ell}_c(\boldsymbol{\theta}, \boldsymbol{\alpha}; \mathbf{y}, \mathbf{Z}) = \tilde{\ell}_c(\boldsymbol{\theta}; \mathbf{y}, \mathbf{Z}) + \tilde{\ell}_c(\boldsymbol{\alpha}; \mathbf{Z}),$$

onde $\tilde{\ell}_c(\boldsymbol{\theta}; \mathbf{y}, \mathbf{Z}) = \ell_c(\boldsymbol{\theta}; \mathbf{y}, \mathbf{Z})|_{Z_m = \mathbb{E}_{m,1}}$ com $\mathbb{E}_{m,1} = \mathbb{E}_{\mathbf{Z}|\mathbf{y}, \boldsymbol{\theta}, \boldsymbol{\alpha}}(Z_m)$ e

$$\tilde{\ell}_c(\boldsymbol{\alpha}; \mathbf{Z}) = \sum_{m=1}^M \left[\log \left(\frac{\alpha_m}{1 - \alpha_m} \right) - \left(\frac{1}{1 - \alpha_m} \right) \mathbb{E}_{m,2} + \mathbb{E}_{m,3} - \mathbb{E}_{m,4} + TL_m \mathbb{E}_{m,2} \right].$$

Neste caso, $\mathbb{E}_{m,2} = \mathbb{E}_{\mathbf{Z}|\mathbf{y}, \boldsymbol{\theta}, \boldsymbol{\alpha}}(\log(Z_m))$, $\mathbb{E}_{m,3} = \mathbb{E}_{\mathbf{Z}|\mathbf{y}, \boldsymbol{\theta}, \boldsymbol{\alpha}}(\log(c(B_m)))$ e $\mathbb{E}_{m,4} = \mathbb{E}_{\mathbf{Z}|\mathbf{y}, \boldsymbol{\theta}, \boldsymbol{\alpha}}(Z_m^{-\frac{\alpha_m}{1-\alpha_m}} c(B_m))$. As esperanças $\mathbb{E}_{m,1}$, $\mathbb{E}_{m,2}$, $\mathbb{E}_{m,3}$ e $\mathbb{E}_{m,4}$ são integrais obtidas por Monte Carlo e apresentadas no Apêndice B. Para o algoritmo **MCEM** condicional, considere $\dot{\boldsymbol{\theta}} = (\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \boldsymbol{\theta}_3, \boldsymbol{\theta}_4)$, onde $\boldsymbol{\theta}_1 = (\alpha_1, \dots, \alpha_m)$, $\boldsymbol{\theta}_2 = (\beta_0, \dots, \beta_i)$, $\boldsymbol{\theta}_3 = (\psi, \rho)$ e $\boldsymbol{\theta}_4 = (\gamma_0, \dots, \gamma_p)$. Assim, os passos E e M são dados por:

- **Passo E:** Inicialize o processo iterativo com $\dot{\boldsymbol{\theta}}^{(0)}$ e na k -ésima iteração, obtenha $v^{(m)}(\dot{\boldsymbol{\theta}}^{(k)})$. Gere $\{(B_m^{j_k}, Z_m^{j_k})\}$, onde $Z_m^{j_k} \sim \text{Gamma}(TL_m + 1, v^{(m)}(\dot{\boldsymbol{\theta}}^{(k)}))$ e $B_m^{j_k} \sim U(0, 1)$, para $j_k \in \{1, \dots, N\}$. Compute $\mathbb{E}_{m,1}^{j_k}, \mathbb{E}_{m,2}^{j_k}, \mathbb{E}_{m,3}^{j_k}$ e $\mathbb{E}_{m,4}^{j_k}$ e aproxime $\mathbb{E}_{m,1}$, $\mathbb{E}_{m,2}$, $\mathbb{E}_{m,3}$ e $\mathbb{E}_{m,4}$ por Monte Carlo,

$$\begin{aligned} \tilde{\mathbb{E}}_{m,1}^{(k)} &= \frac{1}{N} \sum_{j_k=1}^N \mathbb{E}_{m,1}^{j_k}, & \tilde{\mathbb{E}}_{m,2}^{(k)} &= \frac{1}{N} \sum_{j_k=1}^N \mathbb{E}_{m,2}^{j_k} \\ \tilde{\mathbb{E}}_{m,3}^{(k)} &= \frac{1}{N} \sum_{j_k=1}^N \mathbb{E}_{m,3}^{j_k}, & \tilde{\mathbb{E}}_{m,4}^{(k)} &= \frac{1}{N} \sum_{j_k=1}^N \mathbb{E}_{m,4}^{j_k} \end{aligned}$$

- **Passo M:** Obtenha $\dot{\boldsymbol{\theta}}^{(k+1)}$ maximizando $Q(\dot{\boldsymbol{\theta}} | \boldsymbol{\theta}^{(k)}) = \tilde{\ell}_c(\boldsymbol{\theta}, \boldsymbol{\alpha}; \mathbf{y}, \mathbf{Z})$ condicionalmente, ou seja,

$$\boldsymbol{\theta}_1^{(k+1)} = \operatorname{argmax}_{\boldsymbol{\theta}_1} Q(\dot{\boldsymbol{\theta}} | \boldsymbol{\theta}_2^{(k)}, \boldsymbol{\theta}_3^{(k)}, \boldsymbol{\theta}_4^{(k)}),$$

$$\boldsymbol{\theta}_2^{(k+1)} = \operatorname{argmax}_{\boldsymbol{\theta}_2} Q(\dot{\boldsymbol{\theta}} | \boldsymbol{\theta}_1^{(k+1)}, \boldsymbol{\theta}_3^{(k)}, \boldsymbol{\theta}_4^{(k)})$$

$$\boldsymbol{\theta}_3^{(k+1)} = \operatorname{argmax}_{\boldsymbol{\theta}_3} Q(\dot{\boldsymbol{\theta}} | \boldsymbol{\theta}_1^{(k+1)}, \boldsymbol{\theta}_2^{(k+1)}, \boldsymbol{\theta}_4^{(k)})$$

e

$$\boldsymbol{\theta}_4^{(k+1)} = \operatorname{argmax}_{\boldsymbol{\theta}_4} Q(\dot{\boldsymbol{\theta}} | \boldsymbol{\theta}_1^{(k+1)}, \boldsymbol{\theta}_2^{(k+1)}, \boldsymbol{\theta}_3^{(k+1)}).$$

Repita os passos E e M até obter convergência para sequência de estimativas $\hat{\theta}^k$, segundo algum critério de parada, em particular $\|\hat{\theta}^{(k+1)} - \hat{\theta}^{(k)}\| < \epsilon$ em que ϵ é um valor, maior do que zero, especificado para o erro de aproximação e $\|c\|$ é a norma do vetor c .

3.2. Intervalo de Confiança Bootstrap

Os estimadores de Máxima Verossimilhança obtidos pelo algoritmo **MCEM**, não apresentam uma forma analítica para a matriz de covariância dos estimadores. Neste caso, os Intervalos de Confiança para os parâmetros são obtidos via método Bootstrap condicional. Para isso, reamostramos valores do processo utilizando o fato que se $[G^{-1}(1 - y_t(s_l)^{\nu_t(s_l)})] = (1 - y_t(s_l)^{\nu_t(s_l)})$. Então, para cada m e t , $\mathbf{Y}_t(\mathbf{s})|\mathbf{Z}, \mathcal{F}_{t-1} \sim \mathbf{K}(\nu_t(s), \omega_m(s)Z_m/\delta_t)$. Onde $\mathbf{K}$ denota a distribuição Kumaraswami. Portanto, usando a relação $\Phi(\kappa_t(s)) = \log(\nu_t(s))$ na Seção 2.2, o seguinte algoritmo pode ser usado para gerar amostras Bootstrap do processo:

- 1 Considere as covariáveis $\{\mathbf{x}_t(\mathbf{s}), \mathbf{x}_{t-1}(\mathbf{s}), \dots, \mathbf{w}_t, \dots, \mathbf{w}_{t-1}\}$ observadas e $\hat{\theta}$ o vetor de parâmetros estimados com os dados reais.
- 2 Obtenha $\widehat{\nu_t(s)} = \exp\{\Phi(\widehat{\kappa_t(s)})\}$.
- 3 Para $m \in \{1, \dots, M\}$, gere $Z_{m,1}, \dots, Z_{m,N} \sim \hat{\alpha}_m$ —estável positiva e tome $\tilde{Z}_m$ como a média trimed de $Z_{m,1}, \dots, Z_{m,N}$.
- 4 Para todo $s_l \sim \mathbf{h}_m$ gere $y_t^*(s_l) \sim \mathbf{K}(\widehat{\nu_t(s_l)}, \omega_m(s)\tilde{Z}_m/\delta_t)$, $t \in \{1, \dots, T\}$.
- 5 Realize os passos 2, 3 e 4 B vezes e estime $\hat{\theta}_1^*, \hat{\theta}_2^*, \dots, \hat{\theta}_B^*$ para obter a distribuição Bootstrap do processo.

A média trimed é usada no passo 3, devido a distribuição α -estável ser de caudas pesadas possibilitando a geração de outliers na amostra Bootstrap.

3.3. Predição Espaço-Temporal

Duas alternativas são propostas para Predição Espaço-Temporal no modelo. A primeira é baseada na mediana das amostras do bootstrap condicional. Assim, um preditor $\widehat{Y_t(s_0)}$ na localização s_0 próxima de $\mathbf{h}_m$ e tempo t pode ser obtido pelo seguinte algoritmo:

- 1 Considere as covariáveis $\{\mathbf{x}_t(\mathbf{s}_0), \mathbf{x}_{t-1}(\mathbf{s}_0), \dots, \mathbf{w}_t, \dots, \mathbf{w}_{t-1}\}$ observadas e $\hat{\theta}$ o vetor de parâmetros estimados com os dados reais.
- 2 Obtenha $\widehat{\nu_t(s_0)} = \exp\{\Phi(\widehat{\kappa_t(s_0)})\}$.
- 3 Para $m \in \{1, \dots, M\}$, gere $Z_{m,1}, \dots, Z_{m,N} \sim \hat{\alpha}_m$ —estável positiva e tome $\tilde{Z}_m$ como a média trimed de $Z_{m,1}, \dots, Z_{m,N}$.
- 4 Para todo $s_0 \sim \mathbf{h}_m$ gere $y_t^*(s_0) \sim \mathbf{K}(\widehat{\nu_t(s_0)}, \omega_m(s_0)\tilde{Z}_m/\delta_t)$, $t \in \{1, \dots, T\}$.
- 5 Realize os passos 2, 3 e 4 B vezes e obtenha $y_{t,1}^*(s_0), y_{t,2}^*(s_0), \dots, y_{t,B}^*(s_0)$. Nesse caso, tomamos $\widehat{Y_t(s_0)}$ como sendo a mediana das amostra bootstrap geradas no passo 4.

Na segunda abordagem para predição usamos a representação estocástica de $y_t(s)$ dada na expressão (5) da Seção 2.2. Usando o algoritmo:

- 1 Considere as covariáveis $\{\mathbf{x}_t(\mathbf{s}_0), \mathbf{x}_{t-1}(\mathbf{s}_0), \dots, \mathbf{w}_t, \dots, \mathbf{w}_{t-1}\}$ observadas e $\hat{\boldsymbol{\theta}}$ o vetor de parâmetros estimados com os dados reais.
- 2 Obtenha $\widehat{\nu}_t(\mathbf{s}_0) = \exp\{\Phi(\widehat{\kappa}_t(\mathbf{s}_0))\}$.
- 3 Gere $u \sim U(0, 1)$.
- 4 Obtenha para todo $\mathbf{s}_0 \sim \mathbf{h}_m$,

$$\widetilde{y_t(\mathbf{s}_0)} = \left[1 - \exp \left\{ - \left(\frac{-\log(1-u)}{(\omega_m(\mathbf{s}_0)/\widehat{\delta}_t)^{\widehat{\alpha}_m}} \right)^{1/\widehat{\alpha}_m} \right\} \right]^{1/\widehat{\nu}_t(\mathbf{s}_0)}$$

- 5 Realize os passos 2, 3 e 4 B vezes e obtenha $\widetilde{y_{t,1}(\mathbf{s}_0)}, \widetilde{y_{t,2}(\mathbf{s}_0)}, \dots, \widetilde{y_{t,B}(\mathbf{s}_0)}$. Nesse caso, tomamos $\widetilde{Y_t(\mathbf{s}_0)}$ como sendo a mediana das amostra geradas no passo 4.

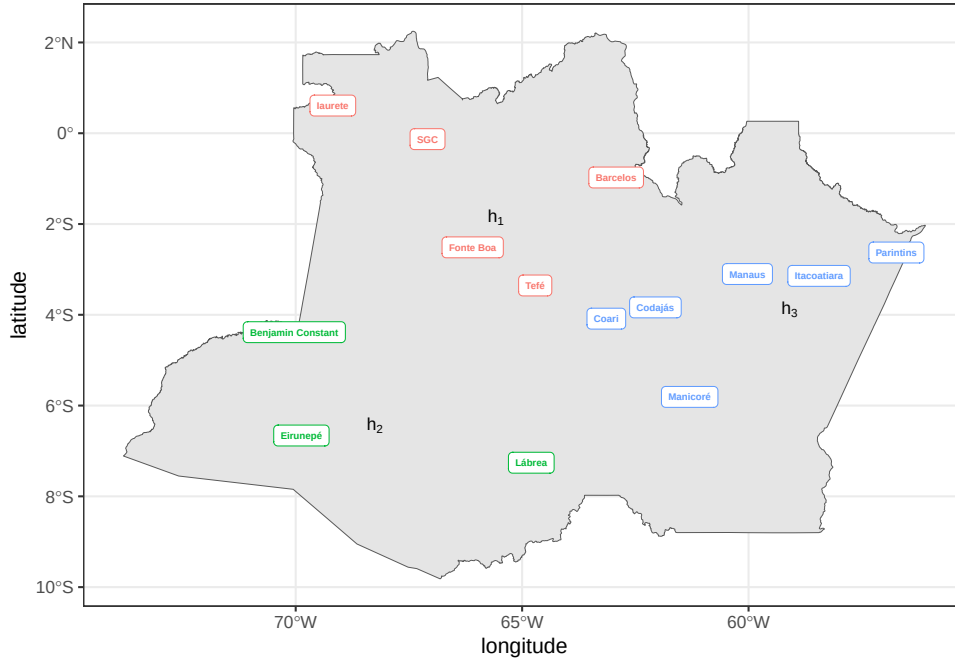


Figura 3. Representação gráfica do efeito compartilhado espacialmente pelas estações. As estações foram divididas em três sub-regiões, cada uma com seu centróide em h_1 , h_2 e h_3 , respectivamente. Cada sub-região possui variáveis latentes correspondentes α_m -estáveis, para $m \in \{1, 2, 3\}$. Essa divisão foi realizada levando em consideração o padrão pluviométrico do estado.

4. Aplicação em dados de Mínima umidade relativa

4.1. Dados e Motivação

Os dados são de umidade relativa do ar medidas nas estações meteorológica do estado do Amazonas. Especificamente a medida consiste no indicador de valor mínimo dos extremos diários (máximo diário mais o mínimo diário e o resultado dividido por dois) de um determinado mês. Eles foram obtidos nas $L = 14$ estações meteorológicas presentes em 13 municípios do estado. Os dados estão disponíveis na Base de dados do Instituto Nacional de Meteorologia do Brasil (<https://portal.inmet.gov.br/>). Consideramos o período de 01/2000 até 12/2008 ($T = 108$), pois algumas estações foram desativadas (Coari, Parintins e Iauarete) a partir de 2016 e, além disso, haviam muito dados faltante. Note que neste caso temos $108 \times 14 = 1512$ observações no total. Portanto, para $t \in \{1, \dots, 108\}$ e $l \in \{1, \dots, 14\}$, o processo no espaço-tempo é $\{Y_t(s_1), \dots, Y_t(s_{14})\}$. As estações meteorológicas foram agrupadas de acordo com Mota et al., ou seja, em $M = 3$ subregiões com centros em $\mathbf{h}_1, \mathbf{h}_2, \mathbf{h}_3$. A partição está de acordo com o padrão de chuvas do estado. Observe que $\mathbf{h}_m \sim \alpha_m$ —estável para $m \in \{1, 2, 3\}$. Além disso, $L_1 = 5$, $L_2 = 3$ e $L_3 = 6$ representam a quantidade de estações em cada grupo, vê Fig 3. Nossa motivação é modelar estes dados para identificar regiões com padrões climáticos extremos de umidade com potenciais riscos a saúde humana.

4.2. Ajuste e predição do Modelo

Nas Fig. 4 e 5 são apresentadas as séries e o comportamento sazonal do indicador nas diferentes estações meteorológicas, respectivamente. A partir da análise da série observada, percebeu-se uma tendência decrescimento e crescimento na mínima umidade no decorrer dos anos. Além disso, foi constatado que a umidade tende a apresentar maiores valores nos primeiros seis meses do ano, diminuindo nos meses de agosto, setembro e outubro e voltando a exibir um crescimento nos meses novembro e dezembro. Diante disso, foi proposto o modelo com $\mathbf{x}_t(s_l) = (1, \text{Altitude}_l, \sin(2\pi t/12), \cos(2\pi t/12))^\top$, onde $\mathbf{x}_s = \text{Altitude}$, é a altura em relação ao nível do mar da estação localizada em s , $\mathbf{x}_t = (\sin(2\pi t/12), \cos(2\pi t/12))$ e $\mathbf{w}_t = (1, \text{sem}_{2,t})^\top$, onde $\text{sem}_{r,t} = 1$ se t pertence ao semestre r e $\text{sem}_{r,t} = 0$ caso contrário, para $r \in \{1, 2\}$. Os vetores de parâmetros dos ajustes por covariáveis são da forma $\beta = (\beta_0, \beta_1, \beta_2, \beta_3)^\top$ e $\gamma = (\gamma_0, \gamma_1)^\top$. Nesse modelo $\log(\nu_s) = \eta_s = \mathbf{x}_s^\top \beta_1$ é o preditor para tendência espacial do indicador de umidade em diferentes localidades (estações meteorológicas), o par $\varphi = (\beta_2, \beta_3)$ mede o efeito sazonal mensal através do harmônico simples de modo que $\log(\nu_t) = \eta_t = \mathbf{x}_t^\top \varphi$ é o preditor para tendência temporal da umidade e γ_1 representa a variação na dispersão temporal da umidade quando mudamos do primeiro para o segundo semestre considerando todo o estado do Amazonas.

Na Tabela 1 apresentamos as estimativas dos parâmetros dos modelos com seus respectivos Intervalos de Confiança Bootstrap considerando $N = 500$, $B = 400$ e $\epsilon = 0.0001$. De forma geral, à medida que a altitude aumenta, ocorre uma diminuição na umidade relativa do ar. Esse comportamento foi suavemente detectado de forma significativa por β_1 . Além disso, uma pequena variação da dispersão temporal ou amplitude da mínima umidade do estado foi detectada por $\hat{\gamma}_1$, ou seja, quando comparado a dispersão temporal do primeiro com o segundo semestre, há uma diminuição na variação de aproximadamente 34% no segundo semestre. O comportamento sazonal também foi significativo. Em relação a dependência espacial, as observações do Grupo 3

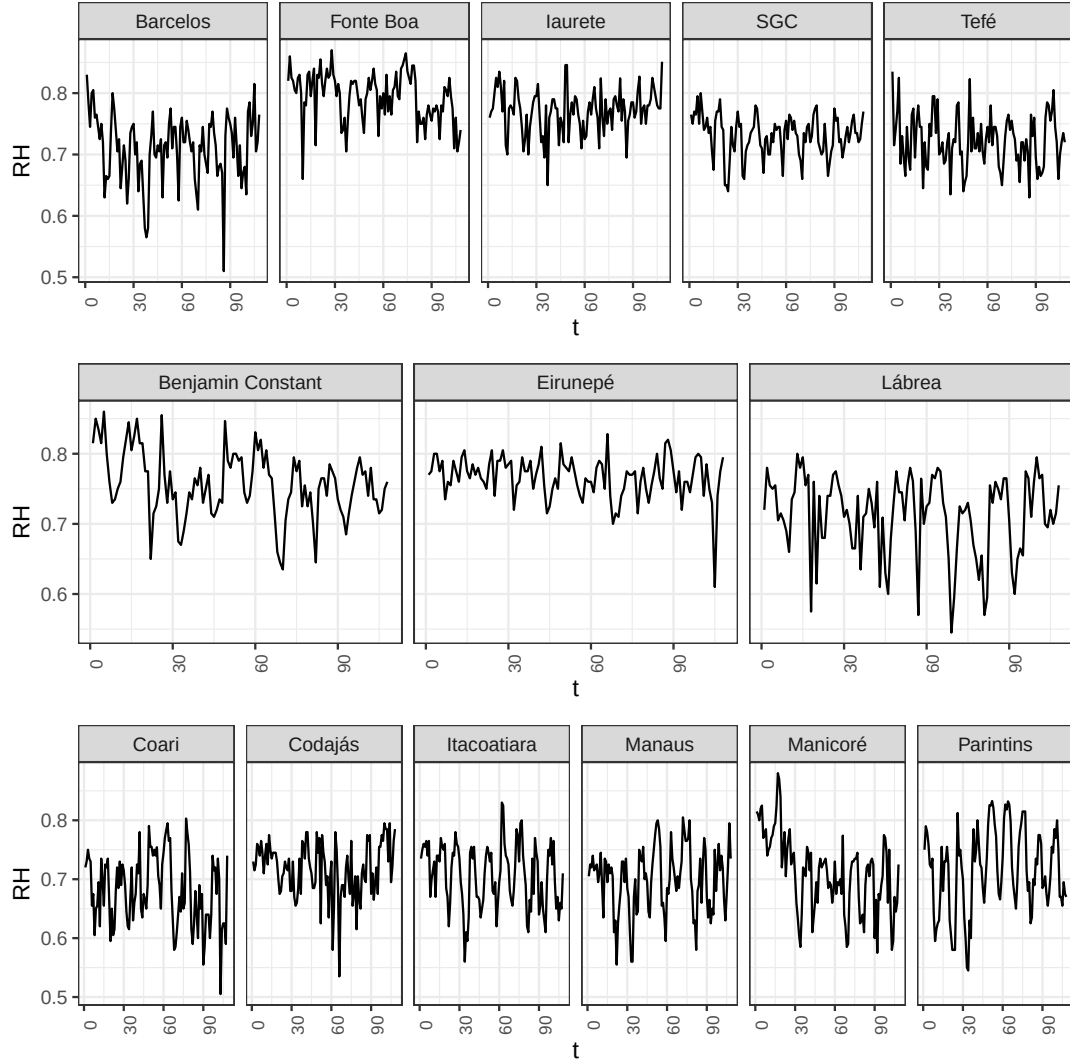


Figura 4. Umidade relativa das estações meteorologia do Estado do Amazonas entre jan/2000 a dez/2008

Tabela 1. Estimativas dos parâmetros

	Variables	Parameter	Estimate	Sd	Interval 95%
1	-	ψ	0.5793	0.0096	(0.5487;0.5863)
2	-	ρ	-0.0981	0.0101	(-0.1172;-0.0785)
3	(Intercept)	β_0	1.2801	0.0099	(1.2781;1.3172)
4	Altitude	β_1	-0.0010	0.0001	(-0.0011;-5e-04)
5	sent	β_2	0.1408	0.0042	(0.1327;0.1494)
6	cost	β_3	-0.0496	0.0033	(-0.0512;-0.0388)
7	(Intercept)	γ_0	-5.7316	0.0922	(-5.7646;-5.4175)
8	sem-2	γ_1	-0.2997	0.1631	(-0.6683;-0.0428)
9	-	α_1	0.3866	0.0034	(0.3836;0.3968)
10	-	α_2	0.5254	0.0078	(0.5062;0.5363)
11	-	α_3	0.3018	0.0027	(0.3005;0.3114)

($m = 3$) são suavemente mais dependentes que as demais, $\hat{\alpha} = (0.3866, 0.5254, 0.3018)$. A dependência temporal de primeira ordem e o efeito de média móvel também foram significativos. Fig. C1 apresenta as distribuições das estimativas. A partir da análise gráfica, percebe-se que as estimativas se aproximam de uma distribuição normal. A

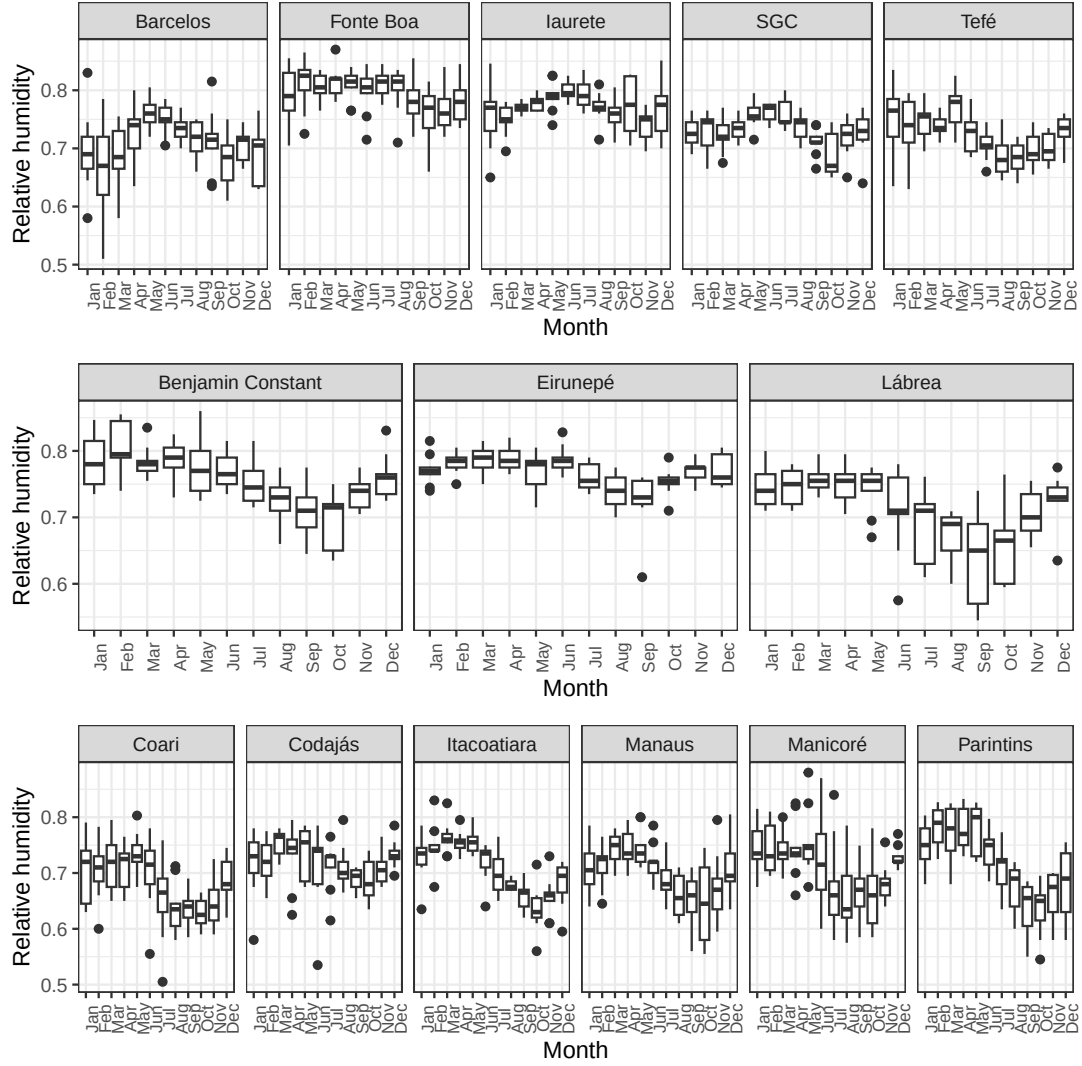


Figura 5. Umidade reltiva das estações meteorológica do Estado do Amazonas de acordo com o Mês.

partir desses resultados, foram utilizados os métodos apresentados na Seção 3.3 para realizar as estimativas para $Y_t(s_l)$. Para o Preditor $\hat{Y}_t(s_l)$ foram utilizados a média *trimed* (Leonowicz et al., 2005) considerando o corte nas caudas das estimativas de 0.15, 0.10 e 0.18 para os grupos 1, 2 e 3 respectivamente. Tal escolha foi feita a partir de uma análise de sensibilidade, no qual os valores de cortes selecionados apresentaram o menor erro percentual médio para as estimaivas. O uso da média *trimed* foi necessário devido ao fato de que o processo de estimação gera valores da distribuição de cauda pesada α -estável positiva, portanto o processo é afetado por outliers e consequentemente influencia na média aritmética. A Fig. 6 apresenta o comportamento dessa estimativa ao longo do tempo. Importante destacar a variabilidade dos valores estimados, a partir do intervalo de cobertura, para as estações em diferentes grupos. Nesse caso, por exemplo, o grupo 2 (três de estações) apresenta maior variabilidade nas estimativas quando comparado ao grupo 3 (seis estações), evidenciando que a quantidade de estações podem influenciar na variabilidade das estimações.

Além disso, a Tabela 2 mostra uma comparação entre $\hat{Y}_t(s_l)$ e $\tilde{Y}_t(s_l)$ considerando o

Tabela 2. Erro médio percentual para os dois métodos de estimativas de $Y_t(s_l)$.

	Grupos	Estação - $Y_t(s_l)$	$\hat{Y}_t(s_l)$	$\tilde{Y}_t(s_l)$
1	Group 1	Barcelos	0.0571	0.0676
2		Fonte Boa	0.0364	0.0329
3		Iaurete	0.0390	0.0535
4		SGC	0.0364	0.0574
5		Tefé	0.0439	0.0544
6	Group 2	Benjamin Constant	0.0322	0.0320
7		Eirunepé	0.0303	0.0287
8		Lábrea	0.0549	0.0569
9	Group 3	Coari	0.0526	0.0717
10		Codajás	0.0480	0.0538
11		Itacoatiara	0.0413	0.0768
12		Manaus	0.0436	0.0702
13		Manicoré	0.0454	0.0593
14		Parintins	0.0499	0.0870

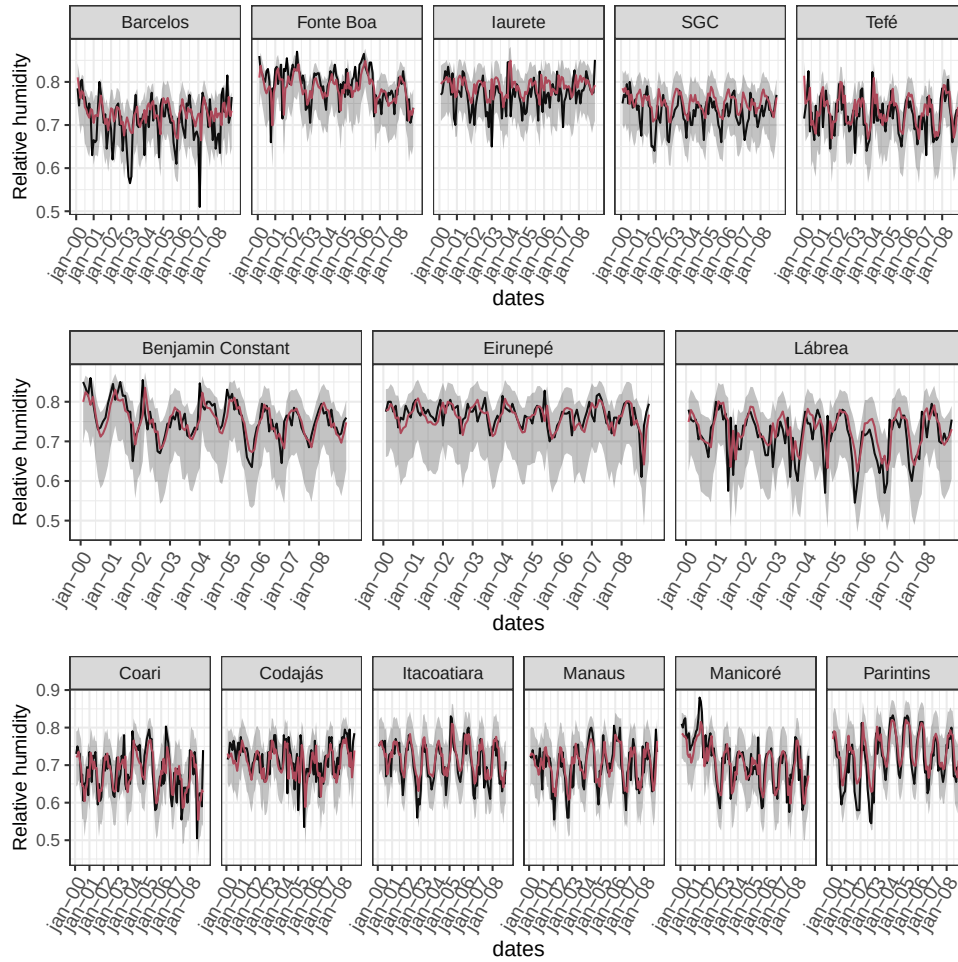


Figura 6. Estimativas, $\hat{Y}_t(s_l)$, (linha vermelha), observações (linha preta) e intervalo de cobertura (parte sombreada.)

erro percentual médio. De forma geral, percebe-se que $\hat{Y}_t(s_l)$ possui uma leve vantagem no processo de predição. Diante disso, a partir das predições feita com $\hat{Y}_t(s_l)$ no centro dos polígonos que delineiam cada município do Amazonas, foi utilizado uma krigagem simples para a interpolação para todo o Estado. A Fig. 7 mostra a predição espacial da

mínima umidade no Estado para o ano de 2008. Nesse caso, nota-se claramente o efeito sazonal, meses do primeiro semestre são mais chuvosos e por isso apresentam umidade mais alta, em particular os meses de Fevereiro, Março e Abril. No entanto, os meses do segundo semestre correspondem ao período de verão da região e consequentemente poucas chuvas, menor umidade e um Estado muito seco gerando um grande desconforto térmico e afetando de forma negativa a saúde humana. Extremos de mínima umidade são observados principalmente no sul do estado do Amazonas no período de agosto a outubro.

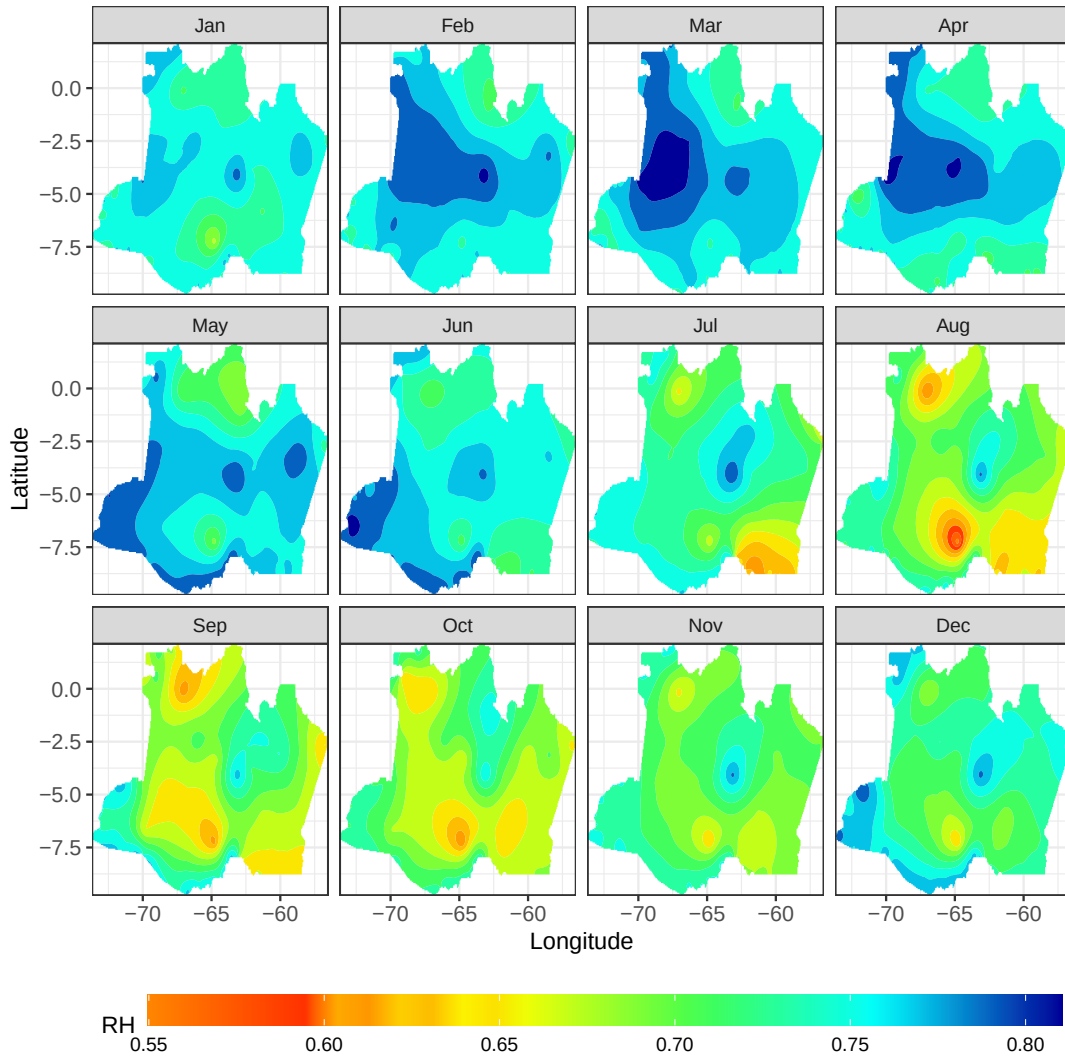


Figura 7. Mapa da umidade relativa para o ano de 2008

4.3. Validação do Modelo

Em problemas de predição Espaço-Temporal é mais importante validar o poder de predição do Modelo. Assim, guiados pelos dados reais analisamos o processo de predição do modelo proposto para um ponto $Y_t(s_0)$, através de um estudo de validação

cruzada, ou seja, foi retirado da base de dados uma estação localizada em s_l , estimamos os parâmetros do modelo e utilizamos $\hat{Y}_t(s_l)$ e $\tilde{Y}_t(s_l)$ para fazer a predição para a estação removida. Esse processo foi realizado para as 14 estações. Esses métodos foram comparados com a bem conhecida *Inverse Distance Weighting (IDW) function for spatio-temporal prediction*. Esta função está disponível na linguagem R Core Team (2023) no pacote *geosptdb* e fornece uma interpolação considerando o espaço e o tempo das observações.

A Tabela 3 apresenta o erro percentual médio para os três métodos. Percebe-se que para o Grupo 1 as estimativas ficaram bem próximas para Barcelos e Fonte Boa, porém diferentes para Tefé. Em relação ao Grupo 2 a IDW apresentou um melhor desempenho para estações Benjamin Constant e Eirunepé mas não para Lábrea. No Grupo 3 IDW teve um desempenho inferior aos métodos propostos para todas as estações, tendo uma maior diferença em Coari. De forma geral, o método proposto foi superior para prever a mínima umidade nos grupos 1 e 3 e essa superioridade, aparentemente cresce com o número de estações avaliadas na região de estudo. Essa comparação sugere que o modelo proposto fornece resultados consistentes quando comparado ao método padrão IDW para espaço tempo.

Tabela 3. Erro médio percentual para os três métodos de estimativas de $Y_t(s_0)$

	Grupos	Estação - $Y_t(s_0)$	$\hat{Y}_t(s_0)$	$\tilde{Y}_t(s_0)$	IDW
1	Group 1	Barcelos	0.0652	0.0642	0.0643
2		Fonte Boa	0.1093	0.1055	0.0900
3		Iaurete	0.0435	0.0399	0.0561
4		SGC	0.0392	0.0375	0.0575
5		Tefé	0.0446	0.0403	0.1038
6	Group 2	Benjamin Constant	0.0795	0.0696	0.0441
7		Eirunepé	0.0989	0.0861	0.0429
8		Lábrea	0.0559	0.0558	0.0855
9	Group 3	Coari	0.0600	0.0684	0.0924
10		Codajás	0.0774	0.0590	0.0801
11		Itacoatiara	0.0474	0.0465	0.0511
12		Manaus	0.0520	0.0537	0.0547
13		Manicoré	0.0658	0.0639	0.0850
14		Parintins	0.0691	0.0682	0.0624

5. Conclusão

Neste trabalho foi proposto um modelo multivariado, caracterizado por uma função de excedência, para predição espaço-temporal de taxas extremas dependentes temporalmente e/ou espacialmente usado para modelar valores climáticos em intervalos limitados. Para isso, a estrutura espacial do modelo foi formulada a partir da marginalização de uma distribuição baseline G condicionada a um campo aleatório latente com distribuição α -estável positiva espacialmente compartilhada. O modelo é útil para a análise de mínimos extremos no espaço-tempo não estacionários com suporte no intervalo $(0,1)$, podendo ser generalizado para o intervalo (a,b) . Para acomodar a não estacionariedade no espaço-tempo, a estrutura do modelo permite a incorporação de covariáveis nos parâmetros de tendência espaço/tempo e dispersão temporal. Além disso, algumas propriedades que permitem calcular a relação de dependência entre observações em locais diferentes foram apresentadas, em particular uma expressão analítica para o extremograma. Em relação as estimativas dos parâmetros, foi utilizado o algoritmo MCEM condicional, onde a densidade completa foi expressada em termos da distribuição baseline G e aplicado ao caso particular $G(y) = y$. Para calcular a variabilidade e realizar teste de hipótese das estimativas, utilizou-se o método bootstrap condicional, onde os valores foram gerados a partir de uma distribuição Kumaraswami obtida através do caso particular $G(y) = y$. De forma geral, as distribuições das estimativas se aproximaram da distribuição normal.

Por fim, uma aplicação nos dados de umidade relativa do estado do Amazonas obtidos nas $L = 14$ estações meteorológicas presentes em 13 municípios do estado entre o período de 01/2000 até 12/2008 foi realizada para analisar a qualidade das estimativas do modelo. Para isso, os métodos propostos foram comparados com a *Inverse Distance Weighting (IDW) function for spatio-temporal prediction* disponível na linguagem de programação R Core Team (2023). De forma geral, o erro médio percentual dos métodos foram parecidos, sendo o método proposto suavemente superior em 9 das 14 estações. Sob o ponto de vista prático, com base nos dados analisados, foi possível identificar padrões climáticos extremos de umidade no estado do Amazonas que podem representar potenciais riscos à saúde humana. Durante o primeiro semestre, correspondente ao período chuvoso, observou-se que algumas regiões apresentaram um índice de umidade superior a 70%. Esse alto nível de umidade pode contribuir para a proliferação de doenças respiratórias, como a pneumonia e a bronquite, além de aumentar a incidência de mofo e fungos nos ambientes, o que pode agravar alergias e problemas respiratórios pré-existentes. No segundo semestre, que compreende o período quente do estado, verificou-se uma redução da umidade para valores abaixo de 60%. Essa diminuição da umidade também pode acarretar problemas de saúde relacionados à desidratação e ressecamento das mucosas, como o ressecamento dos olhos, nariz e garganta, além de aumentar a susceptibilidade a infecções respiratórias. Portanto, é essencial monitorar de perto as regiões com esses padrões climáticos extremos de umidade no estado do Amazonas, a fim de implementar medidas preventivas e mitigar os potenciais riscos à saúde humana.

Referências

- Alegría, A., Caro, S., Bevilacqua, M., Porcu, E., and Clarke, J. (2017). Estimating covariance functions of multivariate skew-gaussian random fields on the sphere. *Spatial Statistics*, 22:388–402.
- Bopp, G. P., Shaby, B. A., and Huser, R. (2021). A Hierarchical Max-Infinitely Divisible Spatial Model for Extreme Precipitation. *Journal of the American Statistical Association*, 533(116):93–106.
- Bowman, A. W. and Azzalini, A. (1997). Applied Smoothing Techniques for Data Analysis: The Kernel Approach with S-Plus Illustrations. *Oxford University Press Inc, New York*.
- Carreau, J. and Toulemonde, G. (2020). Extra-parametrized extreme value copula : Extension to a spatial framework. *Spatial Statistics*, 40:100410.
- Clayton, D. G. (1978). A model for association in bivariate life tables and its application in epidemiological studies of familial tendency in chronic disease incidence. *Biometrika*, 65(1):141–151.
- Coles, S., Heffernan, J., and Tawn, J. (1999). Dependence measures for extreme value analyses. *Extremes*, 2:339–365.
- Cooley, D., Davis, R., and Naveau, P. (2010). The pairwise beta distribution: A flexible parametric multivariate model for extremes. *Journal of Multivariate Analysis*, 101:2103–2117.
- Crowder, M. (1989). A multivariate distribution with Weibull connections. *Journal of the Royal Statistical Society*, 51(1):93–107.
- De Camargo, M. (2007). *Análise da percepção térmica dos carteiros do CDD-Londrina em relação ao uniforme utilizado em ambiente quente*. Dissertação (mestrado), Programa de Pós-Graduação em Desenho Industrial, Universidade Estadual Paulista. Online available at: <https://www.faac.unesp.br/Home/Pos-Graduacao/Design/Dissertacoes/maristela.pdf>. accessed at 28/08/2023.
- Fougères, A.-L., Mercadier, C., and Nolan, J. P. (2010). Dense classes of multivariate extreme value distributions. *Journal of Multivariate Analysis*, 116:109–129.
- Fougères, A.-L., Nolan, J. P., and Rootzén, H. (2009). Models for dependent extremes using stable mixtures. *Scandinavian Journal of Statistics*, 36(1):42–59.
- Gupta, R. C. (2018). Association measures in the bivariate correlated frailty model. *REVSTAT-Statistical Journal*, 16(2):257–278.
- Kedem, B. and Fokianos, K. (2005). Regression models for time series sanlysis. *John Wiley Sons, New Jersey*.
- Krupskii, P. and Genton, M. G. (2019). A copula model for non-Gaussian multivariate spatial data. *Journal of Multivariate Analysis*, 169:264–277.
- Kunin, B. I. (2004). Extreme Value Type Distributions with Bounded Support. *International Journal of Pure and Applied Mathematics*, 12(4):495–507.
- Leonowicz, Z., Karvanen, J., and Shishkin, S. L. (2005). Trimmed estimators for robust averaging of event-related potentials. *Journal of neuroscience methods*, 142(1):17–26.
- Marshall, A. W. and Olkin, I. (1983). Domains of Attraction of Multivariate Extreme Value Distributions. *The Annals of Probability*, 1(11):168–177.
- Mathieu, R. (2013). Spatial extremes: Max-stable processes at work. *Journal de la Société Française de Statistique*, 154(2):156–177.
- Mota, A. L., De Lima, M. S., Demarqui, F. N., and Duczmal, L. H. (2018). The poly-log weibull model applied to space-time interpolation of temperature. *Environmetrics*, 29(7):e2520.

- Nascimento, L. B. F., Lima, M., and Duczmal, L. H. (2023). Power Min-Stable Regression Models for Time Series of Bounded Extremes. Submitted to *Environmetrics*.
- Oakes, D. (1989). Bivariate survival models induced by frailties. *Journal of the American Statistical Association*, 84(406):487–493.
- Opitz, T. (2016). Modeling asymptotically independent spatial extremes based on laplace random fields. *Spatial Statistics*, 16:1–18.
- Papalexiou, S. M. and Serinaldi, F. (2020). Random fields simplified: Preserving marginal distributions, correlations, and intermittency, with applications from rainfall to humidity. *Water Resources Research*, 56(2):e2019WR026331.
- Papalexiou, S. M., Serinaldi, F., and Porcu, E. (2021). Advancing space-time simulation of random fields: From storms to cyclones and beyond. *Water Resources Research*, 57(8):e2020WR029466.
- Peng, Z., Jiang, Q., and Nadarajah, S. (2015). A Unified Representation of P-max Stable Distributions. *MATH. REPORTS*, 17(3):277–293. Available at: http://imar.ro/journals/Mathematical_Reports/Pdfs/2015/3/3.pdf.
- R Core Team (2023). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. Available at: <https://www.R-project.org/>.
- Reich, B. J. and Shaby, B. A. (2012). A Hierarchical Max-Stable Spatial Model for Extreme Precipitation. *The Annals of Applied Statistics*, 6:1430–1451.
- Reich, B. J., Shaby, B. A., and Cooley, D. (2014). A Hierarchical Model for Serially-Dependent Extremes: A Study of Heat Waves in the Western US. *Journal of Agricultural, Biological, and Environmental Statistics*, 19(1).
- Rue, H. and Held, L. (2005). *Gaussian Markov random fields: theory and applications*. Chapman and Hall/CRC.
- Saber, M. M., Nematollahi, A., and Mohammadzadeh, M. (2021). Extended generalized skew laplace random field: Spatial autoregressive and moving average model for prediction of missing data in skew and heavy tailed data. *Journal of Sciences, Islamic Republic of Iran*, 32(2):169–178.
- Shaked, M. (1982). A General Theory of Some Positive Dependence. *Journal of Multivariate Analysis*, 12:199–218.
- Smith, R. L. (1990). Max-stable processes and spatial extremes. Online available at <https://www.rls.sites.oasis.unc.edu/postscript/rs/spatex.pdf>; accessed at 28/08/2023.
- Stephenson, A. G. (2009). High-dimensional parametric modelling of multivariate extreme events. *Australian & New Zealand Journal of Statistics*, 51(1):77–88.
- Sun, Y., Joachimski, M. M., Wignall, P. B., Yan, C., Chen, Y., Jiang, H., Wang, L., and Lai, X. (2012). Lethally hot temperatures during the early triassic greenhouse. *Science*, 338(6105).
- Uchaikin, V. V. and Zolotarev, V. M. (1999). *Chance and Stability: Stable Distributions and their Applications*. De Gruyter, Berlin, Boston.
- Williams, E. (1977). Some representations of stable random variables as products. *Biometrika*, 64(1):167–169.
- Zhang, H. and El-Shaarawi, A. (2010). On spatial skew-gaussian processes and applications. *Environmetrics*, 21(1):33–47.

Apêndice A. Demonstração da distribuição Condicional $Z|\mathbf{y}, \boldsymbol{\theta}, \boldsymbol{\alpha}$

A distribuição condicional de $Z|\mathbf{y}, \boldsymbol{\theta}, \boldsymbol{\alpha}$ é dada por

$$p(\mathbf{Z}; \boldsymbol{\alpha}|\mathbf{y}, \boldsymbol{\theta}) = \prod_{m=1}^M \frac{(v^{(m)}(\boldsymbol{\theta}))^{TL_m+1} Z_m^{TL_m} e^{-Z_m v^{(m)}(\boldsymbol{\theta})} \varphi_0(Z_m, \alpha_m)}{\Gamma(TL_m + 1) \mathbb{E}_{\mathcal{G}}(\varphi_0(Z_m, \alpha_m))},$$

onde $\mathcal{G} \sim \text{Gamma}(TL_m + 1, v^{(m)}(\boldsymbol{\theta}))$.

Note que

$$\begin{aligned} p(\mathbf{Z}|\mathbf{y}, \boldsymbol{\theta}, \boldsymbol{\alpha}) &\propto f(\mathbf{y}; \boldsymbol{\theta}|\mathbf{Z}, \boldsymbol{\alpha}) p(\mathbf{Z}; \boldsymbol{\alpha}) \\ &= c \prod_{m=1}^M Z_m^{TL_m} e^{-Z_m v^{(m)}(\boldsymbol{\theta})} \varphi_0(Z_m, \alpha_m), \end{aligned}$$

onde c é a constante de normalização. Afirmamos que $c = \prod_{m=1}^M \frac{(v^{(m)}(\boldsymbol{\theta}))^{TL_m+1}}{\Gamma(TL_m + 1) \mathbb{E}_{\mathcal{G}}(\varphi_0(Z_m, \alpha_m))}$ sendo que $\mathcal{G} \sim \text{Gama}(TL_m + 1, v^{(m)}(\boldsymbol{\theta}))$.
Com efeito,

$$\begin{aligned} c^{-1} &= \prod_{m=1}^M \int_0^\infty \varphi_0(Z_m, \alpha_m) Z_m^{TL_m} e^{-Z_m v^{(m)}(\boldsymbol{\theta})} dZ_m \\ &= \prod_{m=1}^M \frac{\Gamma(TL_m + 1) \mathbb{E}_{\mathcal{G}}(\varphi_0(Z_m, \alpha_m))}{(v^{(m)}(\boldsymbol{\theta}))^{TL_m+1}}, \end{aligned}$$

e, portanto

$$c = \prod_{m=1}^M \frac{(v^{(m)}(\boldsymbol{\theta}))^{TL_m+1}}{\Gamma(TL_m + 1) \mathbb{E}_{\mathcal{G}}(\varphi_0(Z_m, \alpha_m))}.$$

Apêndice B. Cálculo das Esperanças $\mathbb{E}_{m,1}, \mathbb{E}_{m,2}, \mathbb{E}_{m,3}$ e $\mathbb{E}_{m,4}$

Considere as seguintes funções:

$$\begin{aligned} \varphi_0(Z_m, \alpha_m) &= Z_m^{-\frac{1}{1-\alpha_m}} \int_0^1 c(B_m) e^{-c(B_m) Z_m^{-\frac{\alpha_m}{(1-\alpha_m)}}} dB_m. \\ \varphi_1(Z_m, \alpha_m) &= Z_m^{-\frac{1}{1-\alpha_m}} \int_0^1 \log(c(B_m)) \times c(B_m) e^{-c(B_m) Z_m^{-\frac{\alpha_m}{(1-\alpha_m)}}} dB_m. \\ \varphi_2(Z_m, \alpha_m) &= Z_m^{-\frac{1}{1-\alpha_m}} \int_0^1 c^2(B_m) e^{-c(B_m) Z_m^{-\frac{\alpha_m}{(1-\alpha_m)}}} dB_m. \end{aligned} \tag{B1}$$

Então, temos que:

$$\begin{aligned}
\mathbb{E}_{m,1} &= \mathbb{E}_{\mathbf{Z}|\mathbf{y},\boldsymbol{\theta},\boldsymbol{\alpha}}(Z_m) = \int_0^\infty \dots \int_0^\infty Z_m p(\mathbf{Z}|\mathbf{y},\boldsymbol{\theta},\boldsymbol{\alpha}) d\mathbf{Z} \\
&= \int_0^\infty Z_m p(Z_m|\mathbf{y},\boldsymbol{\theta},\boldsymbol{\alpha}) dZ_m \\
&= \frac{(v^{(m)}(\boldsymbol{\theta}))^{TL_m+1}}{\Gamma(TL_m+1) \mathbb{E}_{\mathcal{G}}(\varphi_0(Z_m, \alpha_m))} \int_0^\infty Z_m \varphi_0(Z_m, \alpha_m) Z_m^{TL_m} e^{-Z_m v^{(m)}(\boldsymbol{\theta})} dZ_m \\
&= \frac{1}{\mathbb{E}_{\mathcal{G}}(\varphi_0(Z_m, \alpha_m))} \left(\frac{(v^{(m)}(\boldsymbol{\theta}))^{TL_m+1}}{\Gamma(TL_m+1)} \int_0^\infty Z_m \varphi_0(Z_m, \alpha_m) Z_m^{TL_m} e^{-Z_m v^{(m)}(\boldsymbol{\theta})} dZ_m \right) \\
&= \frac{\mathbb{E}_{Z_m|\mathbf{y},\boldsymbol{\theta},\boldsymbol{\alpha}}(Z_m \times \varphi_0(Z_m, \alpha_m))}{\mathbb{E}_{\mathcal{G}}(\varphi_0(Z_m, \alpha_m))}.
\end{aligned}$$

$$\begin{aligned}
\mathbb{E}_{m,2} &= \mathbb{E}_{\mathbf{Z}|\mathbf{y},\boldsymbol{\theta},\boldsymbol{\alpha}}(\log(Z_m)) = \int_0^\infty \log(Z_m) p(Z_m|\mathbf{y},\boldsymbol{\theta},\boldsymbol{\alpha}) dZ_m \\
&= \frac{(v^{(m)}(\boldsymbol{\theta}))^{TL_m+1}}{\Gamma(TL_m+1) \mathbb{E}_{\mathcal{G}}(\varphi_0(Z_m, \alpha_m))} \int_0^\infty \log(Z_m) \varphi_0(Z_m, \alpha_m) Z_m^{TL_m} e^{-Z_m v^{(m)}(\boldsymbol{\theta})} dZ_m \\
&= \frac{1}{\mathbb{E}_{\mathcal{G}}(\varphi_0(Z_m, \alpha_m))} \left(\frac{(v^{(m)}(\boldsymbol{\theta}))^{TL_m+1}}{\Gamma(TL_m+1)} \int_0^\infty \log(Z_m) \varphi_0(Z_m, \alpha_m) Z_m^{TL_m} e^{-Z_m v^{(m)}(\boldsymbol{\theta})} dZ_m \right) \\
&= \frac{\mathbb{E}_{Z_m|\mathbf{y},\boldsymbol{\theta},\boldsymbol{\alpha}}(\log(Z_m) \times \varphi_0(Z_m, \alpha_m))}{\mathbb{E}_{\mathcal{G}}(\varphi_0(Z_m, \alpha_m))}.
\end{aligned}$$

$$\begin{aligned}
\mathbb{E}_{m,3} &= \mathbb{E}_{\mathbf{Z}|\mathbf{y},\boldsymbol{\theta},\boldsymbol{\alpha}}(\log(c(B_m))) = \int_0^\infty \log(c(B_m)) p(Z_m|\mathbf{y},\boldsymbol{\theta},\boldsymbol{\alpha}) dZ_m \\
&= \frac{(v^{(m)}(\boldsymbol{\theta}))^{TL_m+1}}{\Gamma(TL_m+1) \mathbb{E}_{\mathcal{G}}(\varphi_0(Z_m, \alpha_m))} \int_0^\infty \varphi_1(Z_m, \alpha_m) Z_m^{TL_m} e^{-Z_m v^{(m)}(\boldsymbol{\theta})} dZ_m \\
&= \frac{1}{\mathbb{E}_{\mathcal{G}}(\varphi_0(Z_m, \alpha_m))} \left(\frac{(v^{(m)}(\boldsymbol{\theta}))^{TL_m+1}}{\Gamma(TL_m+1)} \int_0^\infty \varphi_1(Z_m, \alpha_m) Z_m^{TL_m} e^{-Z_m v^{(m)}(\boldsymbol{\theta})} dZ_m \right) \\
&= \frac{\mathbb{E}_{Z_m|\mathbf{y},\boldsymbol{\theta},\boldsymbol{\alpha}}(\varphi_1(Z_m, \alpha_m))}{\mathbb{E}_{\mathcal{G}}(\varphi_0(Z_m, \alpha_m))}.
\end{aligned}$$

$$\begin{aligned}
\mathbb{E}_{m,4} &= \mathbb{E}_{\mathbf{Z}|\mathbf{y},\boldsymbol{\theta},\boldsymbol{\alpha}}(Z_m^{-\frac{\alpha_m}{1-\alpha_m}} c(B_m)) = \int_0^\infty c(B_m) Z_m^{-\frac{\alpha_m}{1-\alpha_m}} p(Z_m|\mathbf{y},\boldsymbol{\theta},\boldsymbol{\alpha}) dZ_m \\
&= \frac{(v^{(m)}(\boldsymbol{\theta}))^{TL_m+1}}{\Gamma(TL_m+1) \mathbb{E}_{\mathcal{G}}(\varphi_0(Z_m, \alpha_m))} \int_0^\infty Z_m^{-\frac{\alpha_m}{1-\alpha_m}} \varphi_2(Z_m, \alpha_m) Z_m^{TL_m} e^{-Z_m v^{(m)}(\boldsymbol{\theta})} dZ_m \\
&= \frac{1}{\mathbb{E}_{\mathcal{G}}(\varphi_0(Z_m, \alpha_m))} \left(\frac{(v^{(m)}(\boldsymbol{\theta}))^{TL_m+1}}{\Gamma(TL_m+1)} \int_0^\infty Z_m^{-\frac{\alpha_m}{1-\alpha_m}} \varphi_2(Z_m, \alpha_m) Z_m^{TL_m} e^{-Z_m v^{(m)}(\boldsymbol{\theta})} dZ_m \right) \\
&= \frac{\mathbb{E}_{Z_m|\mathbf{y},\boldsymbol{\theta},\boldsymbol{\alpha}}(Z_m^{-\frac{\alpha_m}{1-\alpha_m}} \varphi_2(Z_m, \alpha_m))}{\mathbb{E}_{\mathcal{G}}(\varphi_0(Z_m, \alpha_m))}.
\end{aligned}$$

Apêndice C. Figuras

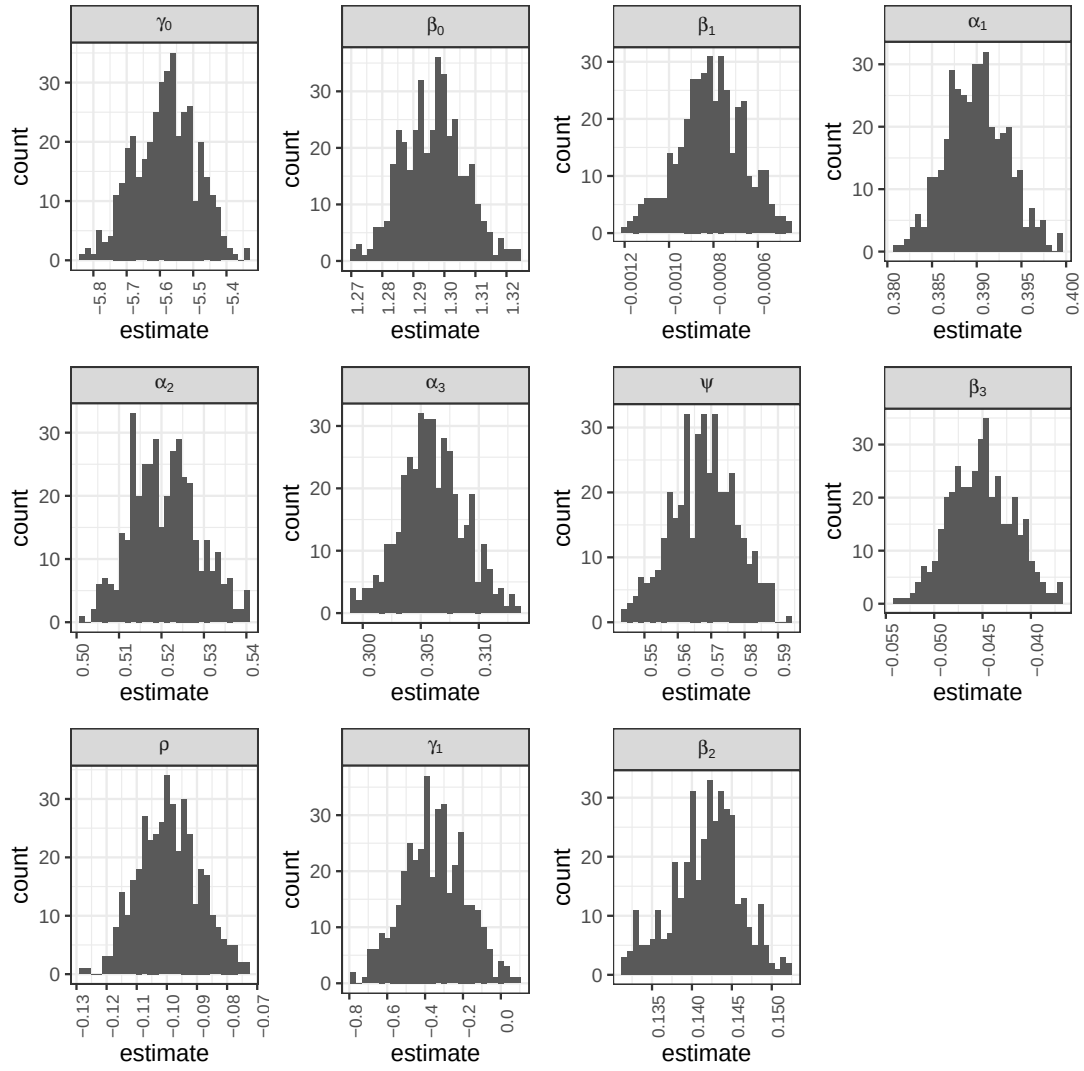


Figura C1. Estimativas Bootstrap do modelo apresentado na Tabela 1.

CAPÍTULO 6

Conclusão

O modelo proposto neste trabalho pertence à classe das distribuições mín-Estável e foi obtido através da marginalização de uma distribuição baseline V condicionada a um campo aleatório latente com distribuição α -estável positiva. A escolha de V e do campo aleatório latente resulta em um modelo multivariado induzido por uma lei de potência. Um caso particular é apresentado quando V é uma função de distribuição de uma variável uniforme no intervalo $(0, 1)$, resultando no modelo Kumaraswamy Exponencializada.

O modelo é adequado para modelar valores extremos em intervalos limitados, levando em consideração a dependência e a não estacionariedade das observações. Ele pode ser aplicado tanto em séries temporais quanto em dados com interação espaço-tempo. Além disso, sua estrutura permite a incorporação de covariáveis nos parâmetros de tendência espaço/tempo e dispersão temporal, permitindo uma compreensão mais completa dos dados. Foram obtidas propriedades analíticas que medem a dependência entre os valores de excedência, como o extremograma temporal e espacial, além da função de excedência cruzada bivariada espacial. Também foram calculadas probabilidades condicionais que podem ser utilizadas para inferências e obtenção de resultados práticos relevantes.

No contexto de séries temporais, foi desenvolvido um algoritmo EM que não considera a ordem de dependência da série. Além disso, uma expressão analítica para a matriz de informação observada foi obtida. Isso permite realizar intervalos de confiança e testes de hipóteses com base na distribuição assintótica dos estimadores. Estudos de simulação Monte Carlo foram conduzidos para avaliar o desempenho do método de estimação, demonstrando resultados satisfatórios para valores de α acima de 0.65. Além disso, uma análise bootstrap das estimativas dos parâmetros foi realizada. Os resultados mostraram boa aproximação da distribuição dos estimadores com a distribuição normal. Por fim, uma aplicação aos dados de umidade relativa do ar em Manaus foi conduzida. A partir das propriedades condicionais fornecidas pelo modelo, foi possível identificar meses com alta probabilidade de ocorrência de umidade mínima acima de 0.70, o que pode representar

riscos à saúde humana.

Na análise espaço-tempo, o modelo pode ser utilizado para analisar a variação espacial da umidade extrema. Nesse caso, medidas de dependência espacial foram analisadas considerando um caso particular da distribuição baseline V . Os resultados mostraram que o modelo é adequado para modelar a dependência nos mínimos extremos. As estimativas dos parâmetros foram obtidas usando o algoritmo MCEM condicional. A variabilidade e testes de hipóteses foram realizados usando o método bootstrap condicional. Com base nos dados analisados, foi possível identificar padrões climáticos extremos de umidade no estado do Amazonas que podem representar potenciais riscos à saúde humana.

Por fim, o modelo proposto neste trabalho se mostrou uma ferramenta adequada para caracterizar, modelar e analisar padrões climáticos extremos que podem representar riscos à saúde humana. Ele oferece uma compreensão mais aprofundada dos dados climáticos e fornece informações úteis para orientar medidas de prevenção e mitigação desses riscos. A aplicação do modelo demonstrou sua utilidade na identificação de períodos críticos de umidade extrema, contribuindo para a adoção de estratégias adequadas de proteção à saúde.

Referências

- Al-Sulami, D. (2020). Exponentiated exponential weibull distribution: mathematical properties and application. *American Journal of Applied Sciences*, 17(1):188–195.
- Alanzi, A., Rafique, M., Tahir, M., Sami, W., and Jamal, F. (2022). A new modified kumaraswamy distribution: Actuarial measures and applications. *Journal of Mathematics*, 2022. doi: 10.1155/2022/4288286.
- Alzaghal, A., Famoye, F., and Lee, C. (2013). Exponentiated t - x family of distributions with some applications. *International Journal of Statistics and Probability*, 2(3):31.
- Bakouch, H. S., Ristić, M. M., Asgharzadeh, A., Esmaily, L., and Al-Zahrani, B. M. (2012). An exponentiated exponential binomial distribution with application. *Statistics & Probability Letters*, 82(6):1067–1081.
- Banerjee, S., Gelfand, A., Knight, J. R., and Sirmans, C. (2004). Spatial modeling of house prices using normalized distance-weighted sums of stationary processes. *Journal of Business & Economic Statistics*, 22(2):206–213.
- Bayer, F. M., Bayer, D. M., and Pumi, G. (2017). Kumaraswamy autoregressive moving average models for double bounded environmental data. *Journal of Hydrology*, 555:385–396. doi: <https://doi.org/10.1016/j.jhydrol.2017.10.006>.
- Beirlant, J., Goegebeur, Y., Segers, J., and Teugels, J. L. (2004). *Statistics of extremes: theory and applications*. Wiley Series in Probability and Statistics. John Wiley & Sons, Hoboken, New Jersey, 1 edition. doi: 10.1002/0470012382.
- Bursa, N. and Gamze, O. (2017). The exponentiated kumaraswamy-power function distribution. *Hacettepe Journal of Mathematics and Statistics*, 46(2):277–292.
- Coles, S., Bawa, J., Trenner, L., and Dorazio, P. (2001). *An introduction to statistical modeling of extreme values*. Springer Series in Statistics. Springer, London, 1 edition. doi: <https://doi.org/10.1007/978-1-4471-3675-0>.
- Coles, S. G. (1993). Regional modelling of extreme storms via max-stable processes. *Journal of the Royal Statistical Society: Series B*, 55(4):797–816.

- Coles, S. G. and Tawn, J. A. (1996). Modelling extremes of the areal rainfall process. *Journal of the Royal Statistical Society: Series B (Methodological)*, 58(2):329–347.
- Cooley, D., Cisewski, J., Erhardt, R. J., Jeon, S., Mannshardt, E., Omolo, B. O., and Sun, Y. (2012). A survey of spatial extremes: Measuring spatial dependence and modeling spatial effects. *REVSTAT-Statistical Journal*, 10(1):135–165.
- Courard-Hauri, D. (2007). Using monte carlo analysis to investigate the relationship between overconsumption and uncertain access to one’s personal utility function. *Ecological Economics*, 64(1):152–162.
- Cressie, N. (1993). *Statistics for spatial data*. Series in Probability and Statistics. John Wiley & Sons, Hoboken, New Jersey.
- Da Silva, D. M. and Montenegro, L. C. C. (2021). Mcem and saem algorithms for geostatistical models under preferential sampling. *arXiv preprint arXiv:2108.01510*.
- Davison, A. C. and Gholamrezaee, M. M. (2012). Geostatistics of extremes. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 468(2138):581–608.
- Davison, A. C. and Huser, R. (2015). Statistics of extremes. *Annual Review of Statistics and its Application*, 2:203–235.
- Davison, A. C., Padoan, S. A., and Ribatet, M. (2012). Statistical Modeling of Spatial Extremes. *Statistical Science*, 27(2):161 – 186. doi: 10.1214/11-STS376.
- De Araujo Lima-Filho, L. M. and Mariano Bayer, F. (2021). Kumaraswamy control chart for monitoring double bounded environmental data. *Communications in Statistics-Simulation and Computation*, 50(9):2513–2528.
- De Haan, L. (1976). Sample extremes: an elementary introduction. *Statistica Neerlandica*, 30(4):161–172. doi: <https://doi.org/10.1111/j.1467-9574.1976.tb00275.x>.
- De Haan, L. (1984). A spectral representation for max-stable processes. *The Annals of Probability*, 12(4):1194–1204.
- De Haan, L. and Pickands III, J. (1986). Stationary min-stable stochastic processes. *Probability Theory and Related Fields*, 72(4):477–492.
- Dempster, A. P., Laird, N. M., and Rubin, D. B. (1977). Maximum likelihood from incomplete data via the em algorithm. *Journal of the royal statistical society: series B (methodological)*, 39(1):1–22.

- Dey, S., Mazucheli, J., and Anis, M. (2017). Estimation of reliability of multicomponent stress–strength for a kumaraswamy distribution. *Communications in Statistics-Theory and Methods*, 46(4):1560–1572.
- Ding, W. and Song, P. X.-K. (2016). Em algorithm in gaussian copula with missing data. *Computational Statistics & Data Analysis*, 101:1–11.
- Embrechts, P., Klüppelberg, C., and Mikosch, T. (1997). *Modelling extremal events: for insurance and finance*, volume 33 of *Stochastic Modelling and Applied Probability*. Springer, Berlin, Heidelberg, 1 edition.
- Fisher, R. A. and Tippett, L. H. C. (1928). Limiting forms of the frequency distribution of the largest or smallest member of a sample. In *Mathematical proceedings of the Cambridge philosophical society*, volume 24, pages 180–190. Cambridge University Press.
- Fletcher, S. and Ponnambalam, K. (1996). Estimation of reservoir yield and storage distribution using moments analysis. *Journal of Hydrology*, 182(1-4):259–275.
- Fougeres, A.-L., Mercadier, C., and Nolan, J. P. (2013). Dense classes of multivariate extreme value distributions. *Journal of Multivariate Analysis*, 116:109–129.
- Fougeres, A.-L., Nolan, J. P., and Rootzén, H. (2009). Models for dependent extremes using stable mixtures. *Scandinavian Journal of Statistics*, 36(1):42–59.
- Fréchet, M. (1927). Sur la loi de probabilité de l’écart maximum. In *Annales de la Societe Polonaise de Mathematique*, volume 6, pages 93–116.
- Ganji, A., Ponnambalam, K., Khalili, D., and Karamouz, M. (2006). Grain yield reliability analysis with crop water demand uncertainty. *Stochastic Environmental Research and Risk Assessment*, 20(4):259–277.
- Gelfand, A. E., Diggle, P., Guttorp, P., and Fuentes, M. (2010). *Handbook of spatial statistics*. CRC press, Boca Raton, Florida, 1 edition. doi: <https://doi.org/10.1201/9781420072884>.
- Gnedenko, B. (1943). Sur la distribution limite du terme maximum d’une serie aleatoire. *Annals of Mathematics*, 44:423–453.
- Gupta, R. C., Gupta, P. L., and Gupta, R. D. (1998). Modeling failure time data by lehman alternatives. *Communications in Statistics-Theory and methods*, 27(4):887–904.
- Gupta, R. D. and Kundu, D. (2001). Exponentiated exponential family: an alternative to gamma and weibull distributions. *Biometrical Journal*, 43(1):117–130.

- Haan, L. and Ferreira, A. (2006). *Extreme value theory: an introduction*, volume 3 of *Series in Operations Research and Financial Engineering*. Springer, New York, NY, 1 edition.
- James, G. S. (1959). *Journal of the Royal Statistical Society. Series A (General)*, 122(2):243–245.
- Jurado, O. E., Ulrich, J., Scheibel, M., and Rust, H. W. (2020). Evaluating the performance of a max-stable process for estimating intensity-duration-frequency curves. *Water*, 12(12):3314.
- Kanter, M. et al. (1975). Stable densities under change of scale and total variation inequalities. *The Annals of Probability*, 3(4):697–707.
- Kumar, D., Jain, N., Nassar, M., and Abo-Kasem, O. E. (2021). Parameter estimation for the exponentiated kumaraswamy-power function distribution based on order statistics with application. *Annals of Data Science*, 8:785–811.
- Kumaraswamy, P. (1980). A generalized probability density function for double-bounded random processes. *Journal of Hydrology*, 46(1-2):79–88.
- Ledford, A. W. and Tawn, J. A. (1996). Statistics for near independence in multivariate extreme values. *Biometrika*, 83(1):169–187.
- Lehmann, E. L. (1953). The power of rank tests. *The Annals of Mathematical Statistics*, pages 23–43.
- Levy, P. (1925). *Calcul des probabilités*, page 113. Gauthier Villars, Paris.
- Liu, D., Vu, M. T., Chatterjee, S., and Rasmussen, L. K. (2020). Neural network based explicit mixture models and expectation-maximization based learning. May. 24, 2020 doi: <https://arxiv.org/abs/1907.13432>IEEE.
- Louis, T. A. (1982). Finding the observed information matrix when using the em algorithm. *Journal of the Royal Statistical Society: Series B (Methodological)*, 44(2):226–233.
- Marshall, A. W. and Olkin, I. (1983). Domains of attraction of multivariate extreme value distributions. *The Annals of Probability*, 11:168–177.
- Matsumura, T., Nagamura, N., Akaho, S., Nagata, K., and Ando, Y. (2021). Spectrum adapted expectation-conditional maximization algorithm for extending high-throughput peak separation method in xps analysis. *Science and Technology of Advanced Materials: Methods*, 1(1):45–55.

- McLachlan, G. and Krishnan, T. (2008). *The EM algorithm and extensions*. Wiley series in probability and statistics. Wiley, Hoboken, NJ, 2 edition.
- Meerschaert, M. M. and Scheffler, H.-P. (2001). *Limit distributions for sums of independent random vectors: Heavy tails in theory and practice*, volume 321 of *Series in Probability and Statistics*. John Wiley & Sons, Hoboken, New Jersey.
- Meng, X.-L. and Rubin, D. B. (1993). Maximum likelihood estimation via the ecm algorithm: A general framework. *Biometrika*, 80(2):267–278.
- Metropolis, N. and Ulam, S. (1949). The monte carlo method. *Journal of the American Statistical Association*, 44(247):335–341.
- Mitnik, P. A. and Baek, S. (2013). The kumaraswamy distribution: median-dispersion reparameterizations for regression modeling and simulation-based estimation. *Statistical Papers*, 54:177–192.
- Mohan, N. and Ravi, S. (1993). Max domains of attraction of univariate and multivariate p-max stable laws. *Theory of Probability & Its Applications*, 37(4):632–643.
- Mota, A., De Lima, M., Demarqui, F., and Duczmal, L. (2018). The poly-log weibull model applied to space-time interpolation of temperature. *Environmetrics*, 29(7):e2520.
- Nadarajah, S. and Kotz, S. (2006). The exponentiated type distributions. *Acta Applicandae Mathematicae*, 92(2):97–111.
- Ng, S.-K. and McLachlan, G. J. (2004). Using the em algorithm to train neural networks: misconceptions and a new algorithm for multiclass classification. *IEEE transactions on neural networks*, 15(3):738–749.
- Nolan, J. P. (2020). *Univariate Stable Distributions: Models for Heavy Tailed Data*. Series in Operations Research and Financial Engineering. Springer Nature, Switzerland, 1 edition.
- Orchard, T. and Woodbury, M. A. (1972). *A missing information principle: theory and applications*, pages 697–716. University of California Press, Berkeley. doi:10.1525/9780520325883-036.
- Pancheva, E. (1994). *Extreme Value Limit Theory with Nonlinear Normalization*, pages 305–318. Springer US, Boston, MA.
- Pancheva, E. (2006). Limit theorems for extreme order statistics under nonlinear normalization. In *Stability Problems for Stochastic Models: Proceedings of the 8th International Seminar held in Uzhgorod, USSR, Sept. 23–29, 1984*, pages 284–309. Springer.

- Peng, Z., Jiang, Q., and Nadarajah, S. (2015). A unified representation of p-max stable distributions. *MATH. REPORTS*, 17(3):277–293. Available at: http://imar.ro/journals/Mathematical_Reports/Pdfs/2015/3/3.pdf.
- Pinho, L., Cordeiro, G., and Nobre, J. (2012). The gamma-exponentiated weibull distribution. *Journal of Statistical Theory and Applications*, 11(4):379–395.
- Portero Serrano, J., Acero Díaz, F. J., and García García, J. A. (2020). Analysis of extreme temperature events over the iberian peninsula during the 21st century using dynamic climate projections chosen using max-stable processes. *Atmosphere*, 11(5):506.
- Rather, A. A. and Subramanian, C. (2018). Exponentiated mukherjee-islam distribution. *Journal of Statistics Applications & Probability*, 7(2):357–361.
- Ribatet, M. (2013). Spatial extremes: Max-stable processes at work. *Journal de la Société Française de Statistique*, 154(2):156–177.
- Ristić, M. M. and Balakrishnan, N. (2012). The gamma-exponentiated exponential distribution. *Journal of Statistical Computation and Simulation*, 82(8):1191–1206.
- Rocha, A. V. and Cribari-Neto, F. (2009). Beta autoregressive moving average models. *Test*, 18(3):529–545.
- Salvadori, G., De Michele, C., Kottegoda, N. T., and Rosso, R. (2007). *Extremes in nature: an approach using copulas*, volume 56 of *Water Science and Technology Library*. Springer Science & Business Media, Dordrecht.
- Sánchez, S., Ancheyta, J., and McCaffrey, W. C. (2007). Comparison of probability distribution functions for fitting distillation curves of petroleum. *Energy & Fuels*, 21(5):2955–2963.
- Schlather, M. and Tawn, J. (2002). Inequalities for the extremal coefficients of multivariate extreme value distributions. *Extremes*, 5:87–102.
- Schlather, M. and Tawn, J. A. (2003). A dependence measure for multivariate and spatial extreme values: Properties and inference. *Biometrika*, 90(1):139–156.
- Seifi, A., Ponnambalam, K., and Vlach, J. (2000). Maximization of manufacturing yield of systems with arbitrary distributions of component values. *Annals of Operations research*, 99(1):373–383.
- Sengupta, A. and Cressie, N. (2013). Empirical hierarchical modelling for count data using the spatial random effects model. *Spatial Economic Analysis*, 8(3):389–418.
- Shi, D. (1995). Multivariate extreme value distribution and its fisher information matrix. *ACTA Mathematicae Applicatae Sinica*, 11(4):421–428.

- Skahill, B., Smith, C. H., Russell, B. T., and England, J. F. (2023). Impacts of max-stable process areal exceedance calculations to study area sampling density, surface network precipitation gage extent and density, and model fitting method. *Hydrology*, 10(6):121.
- Smith, R. L. (1990). Max-stable processes and spatial extremes. Online available at <https://www.rls.sites.oasis.unc.edu/postscript/rs/spatex.pdf>; accessed at 28/08/2023.
- Sundar, V. and Subbiah, K. (1989). Application of double bounded probability density function for analysis of ocean waves. *Ocean Engineering*, 16(2):193–200.
- Thibaud, E., Mutzner, R., and Davison, A. C. (2013). Threshold modeling of extreme spatial rainfall. *Water Resources Research*, 49(8):4633–4644.
- Tovar-Falon, R. and Martínez-Flórez, G. (2022). A new class of exponentiated beta-skew-laplace distribution. *Anais da Academia Brasileira de Ciências*, 94. doi: <https://doi.org/10.1590/0001-3765202220191597>.
- Uchaikin, V. V. and Zolotarev, V. M. (1999). *Chance and Stability: Stable Distributions and their Applications*. De Gruyter, Berlin, Boston.
- Wadsworth, J. L. and Tawn, J. (2022). Higher-dimensional spatial extremes via single-site conditioning. *Spatial Statistics*, 51:100677.
- Wang, B. X., Wang, X. K., and Yu, K. (2017). Inference on the kumaraswamy distribution. *Communications in Statistics-Theory and Methods*, 46(5):2079–2090.
- Wei, G. C. and Tanner, M. A. (1990). A monte carlo implementation of the em algorithm and the poor man’s data augmentation algorithms. *Journal of the American statistical Association*, 85(411):699–704.
- Williams, E. J. (1977). Some representations of stable random variables as products. *Biometrika*, 64(1):167–169.
- Zeepongsekul, P., Jayasinghe, C. L., Fiondella, L., and Nagaraju, V. (2016). Maximum-likelihood estimation of parameters of nhpp software reliability models using expectation conditional maximization algorithm. *IEEE Transactions on Reliability*, 65(3):1571–1583.
- Zolotarev, V. M. (1986). *One-dimensional stable distributions*, *Translations of Mathematical Monographs*, volume 65. American Mathematical Society, Boston, Massachusetts.

APÊNDICE A

Demonstração da Função Densidade por Grupos

Neste apêndice, será realizada a demonstração da função densidade por grupos fornecida em (3.16) utilizando a técnica de derivadas recursivas de forma sucessiva.

Demonstração.

$$\begin{aligned} \frac{\partial \bar{F}(\mathbf{y}_b; \boldsymbol{\theta}, \alpha_b)}{\partial y_{a,b}} &= \frac{\partial \bar{F}(\mathbf{y}_b; \boldsymbol{\theta}, \alpha_b)}{\partial \zeta_b} \times \frac{\partial \zeta_b}{\partial \nu_{a,b}} \times \frac{\partial \nu_{a,b}}{\partial y_{a,b}} \\ &= -\alpha_b \omega_{a,b} h(y_{a,b}; \boldsymbol{\theta}) \zeta_b^{1-\frac{1}{\alpha_b}} e^{-\zeta_b}, \quad \text{para } a = 1, 2, \dots, A_b. \end{aligned} \quad (\text{A.1})$$

Vamos usar o método da indução finita sob $A_b \geq 2$.

Para $A_b = 2$, temos que:

$$\begin{aligned} f(y_{1,b}, y_{2,b}; \boldsymbol{\theta}, \alpha_b) &= \frac{\partial^{(2)} \bar{F}(y_{1,b}, y_{2,b}; \boldsymbol{\theta}, \alpha_b)}{\partial y_1 \partial y_2} = \frac{\partial}{\partial y_2} \left(\frac{\partial \bar{F}(y_{1,b}, y_{2,b}; \boldsymbol{\theta}, \alpha_b)}{\partial y_1} \right) \\ &= \frac{\partial}{\partial y_2} \left(-\alpha_b \omega_{1,b} h(y_{1,b}; \boldsymbol{\theta}) \zeta_b^{1-\frac{1}{\alpha_b}} e^{-\zeta} \right) \quad (\text{por } \text{A.1}) \\ &= -\alpha_b \omega_{1,b} h(y_{1,b}; \boldsymbol{\theta}) \frac{\partial}{\partial y_2} \left(\zeta_b^{1-\frac{1}{\alpha_b}} e^{-\zeta} \right) \\ &= -\alpha_b \omega_{1,b} h(y_{1,b}; \boldsymbol{\theta}) \left[\left(\frac{\partial \zeta_b^{1-\frac{1}{\alpha_b}}}{\partial y_2} \right) e^{-\zeta} + \zeta_b^{1-\frac{1}{\alpha_b}} \left(\frac{\partial e^{-\zeta}}{\partial y_2} \right) \right] \\ &= -\alpha_b \omega_{1,b} h(y_{1,b}; \boldsymbol{\theta}) \\ &\quad \times \left[e^{-\zeta} (\alpha_b - 1) \zeta_b^{1-\frac{2}{\alpha_b}} \omega_{2,b} h(y_{2,b}; \boldsymbol{\theta}) - e^{-\zeta} \alpha_b \omega_{2,b} \zeta_b^{2-\frac{2}{\alpha_b}} h(y_{2,b}; \boldsymbol{\theta}) \right] \\ &= \alpha_b^2 \left(\prod_{a=1}^2 \omega_{a,b} h(y_{a,b}; \boldsymbol{\theta}) \right) e^{-\zeta} \zeta_b^{1-\frac{2}{\alpha_b}} \left(\frac{1-\alpha_b}{\alpha_b} + \zeta \right). \end{aligned}$$

Como $Q_2(\zeta; \alpha_b) = \left(\frac{1-\alpha_b}{\alpha_b} + \zeta \right)$, o resultado é válido para $A_b = 2$.

Suponha que o resultado é válido para $A_b = k$, isto é

$$f(y_{1,b}, y_{2,b}, \dots, y_{k,b}; \boldsymbol{\theta}_b, \alpha_b) = \alpha_b^k \left(\prod_{a=1}^k \omega_{a,b} h(y_{a,b}; \boldsymbol{\theta}) \right) \exp(-\zeta) \zeta^{1-\frac{k}{\alpha_b}} Q_k(\zeta, \alpha_b) \quad (\text{A.2})$$

Mostraremos que vale para $A_b = k + 1$. De fato,

$$\begin{aligned} f(y_1, y_2, \dots, y_k, y_{k+1}; \boldsymbol{\theta}, \alpha_b) &= \frac{(-1)^{(k+1)} \partial^{(k+1)} \overline{F}(y_1, y_2, \dots, y_k, y_{k+1}; \boldsymbol{\theta}, \alpha_b)}{\partial y_1 \partial y_2 \dots \partial y_{k+1}} \\ &= -\frac{\partial}{\partial y_{k+1}} \left(\frac{(-1)^k \partial^{(k)} \overline{F}(y_1, y_2, \dots, y_k, y_{k+1}; \boldsymbol{\theta}, \alpha_b)}{\partial y_1 \dots \partial y_k} \right) \\ &= -\frac{\partial}{\partial y_{k+1}} \left(\alpha_b^k \left(\prod_{a=1}^k \omega_{a,b} h(y_{a,b}; \boldsymbol{\theta}) \right) \exp(-\zeta) \zeta^{1-\frac{k}{\alpha_b}} Q_k(\zeta, \alpha_b) \right) \\ &= -\alpha_b^k \left(\prod_{a=1}^k \omega_{a,b} h(y_{a,b}; \boldsymbol{\theta}) \right) \frac{\partial}{\partial y_{k+1}} \left(\exp(-\zeta) \zeta^{1-\frac{k}{\alpha_b}} Q_k(\zeta, \alpha_b) \right) \quad (\text{A.3}) \end{aligned}$$

Mas,

$$\begin{aligned} \frac{\partial}{\partial y_{k+1}} \left(e^{-\zeta} \zeta^{1-\frac{k}{\alpha_b}} \right) &= \left(\frac{\partial e^{-\zeta}}{\partial y_{k+1}} \right) \zeta^{1-\frac{k}{\alpha_b}} + e^{-\zeta} \left(\frac{\partial \zeta^{1-\frac{k}{\alpha_b}}}{\partial y_{k+1}} \right) \\ &= \alpha_b \omega_{k+1,b} h(y_{k+1,b}; \boldsymbol{\theta}) \zeta e^{-\zeta} \left(\frac{\alpha_b - k}{\alpha_b} - \zeta \right) \quad (\text{A.4}) \end{aligned}$$

e,

$$\begin{aligned} \frac{\partial Q_k(\zeta, \alpha_b)}{\partial y_{k+1}} &= \frac{\partial Q_k(\zeta, \alpha_b)}{\partial \zeta_b} \times \frac{\partial \zeta_b}{\partial y_{k+1,b}} \times \frac{\partial y_{k+1,b}}{\partial y_{k+1}} \\ &= \frac{\partial Q_k(\zeta, \alpha_b)}{\partial \zeta_b} \left(\alpha_b \omega_{k+1,b} h(y_{k+1,b}; \boldsymbol{\theta}) \zeta^{1-\frac{1}{\alpha_b}} \right) \quad (\text{A.5}) \end{aligned}$$

Após realizar algumas manipulações algébricas em (A.3) e utilizando os resultados obtidos em (A.4) e (A.5), chega-se à seguinte expressão:

$$f(y_1, y_2, \dots, y_k, y_{k+1}; \boldsymbol{\theta}, \alpha_b) = \alpha_b^{k+1} \left(\prod_{a=1}^{k+1} \omega_{a,b} h(y_{a,b}; \boldsymbol{\theta}) \right) e^{-\zeta} \zeta_b^{1-\frac{k+1}{\alpha_b}} Q_{k+1}(\zeta, \alpha_b)$$

e, portanto, o resultado é válido para todo $A_b \in \mathbb{N}$. □